



The  
University  
Of  
Sheffield.

# **Investigations into some applications of redundancy in ptychography**

By:

**Yangyang Mu**

Supervised by:

Prof John Rodenburg FRS

Dr Andrew Maiden

A thesis submission for the degree of Doctor of Philosophy

February 2021

## **Abstract**

Ptychography is a computational microscopic imaging technique which has developed very quickly during the past decade. It is a form of coherent diffraction imaging (CDI), but which uses many diffraction patterns measured from many different overlapping areas on the specimen. Ptychography can do lens less imaging, which allows it to avoid the lens-caused aberration and resolution limitation. Furthermore, ptychography can self-calibrate experiential errors in the calculation. The reason for realizing these advantages is that there is a huge amount of information redundancy in the recorded data. Because of this large information diversity, it is hard for any error or noise to disturb the calculated reconstruction of the specimen image.

In this thesis, we investigate some novel applications of the redundancy in ptychography. We investigate one of the most important redundancy applications in ptychography – mixed state ptychography. We engineer some known mode patterns to speed up ptychography that uses partially coherent illumination, which we test experimentally with synchrotron X-ray radiation. We also use a theoretical model of mixed state object modes in ptychography to reduce the influence of specimen damage in the image reconstruction. Furthermore, we remove the non-linear intensities in the data by using the redundancy in the data itself, without any other calibration measurement.

## **Declaration**

I confirm the work present in this thesis is my own based on my research carried out in Professor John Rodenburg' lab, Department of Electronic and Electrical Engineering, the University of Sheffield. No part has previously been submitted for any other degree or qualification. It is all my own work unless explicit reference is given.

Yangyang mu

18/02/2020

## **Acknowledgement**

It was lucky for me to meet such lot of talented and helpful persons during the last four years of my PhD studies. I would like to express my gratitude to them for their support and help.

I would like to express the depth of my gratitude to my supervisor John Rodenburg FRS. I am grateful that he gave me this opportunity. He is the wisest person I have met, and his optimistic spirit and serious attitude have deeply affected me. He not only taught me how to do research, but also how to be a researcher. I am also grateful for his approachability and tolerance. The days of drinking in the Red Deer are my most precious memories.

I would like to thank my second supervisor Dr. Andrew Maiden. He helped me a lot at the coding and modeling, and the activities he organized to strengthen the communication of the team. I thank Dr. Peng Li who gave me a lot of very important guidance, which allowed me to quickly enter my role in my first year.

My gratitude also goes to Dr. Shaohong Cao, Dr. Samuel McDermott, Mr. Zhuoqun Zhang, Mr. Freddrick Allars, Mr. Shengbo You, and Mr. Wenjie Mei. It is lucky to work with you in the research office and I thank you very much for your help over the years.

I would like to thank my parents for their support, encouragement and understanding. I would like to express my special thanks to my girlfriend for her support, encouragement, understanding, and company.

Finally, I would like to thank the University of Sheffield for providing me with convenient learning and living conditions, and also the effective protection and safety of the students and staff during the COVID-19. I am honoured to be part of it.

Yangyang mu

18/02/2020

# Table of Contents

<b>Abstract</b> .....	<i>ii</i>
<b>Declaration</b> .....	<i>iii</i>
<b>Acknowledgement</b> .....	<i>iv</i>
<b>Chapter 1 Introduction</b> .....	<b>3</b>
1.1. The outline of this thesis.....	6
<b>Chapter 2 General background</b> .....	<b>9</b>
2.1. Lens imaging.....	9
2.2. Resolution limitation.....	11
2.3. X-Ray synchrotron and FZP .....	14
2.3.1. X-ray sources .....	14
2.3.2. Properties of synchrotron x-rays .....	18
2.3.3. Properties of the FZP .....	19
2.4. Calculation approximates of wave propagation.....	22
2.5. Conclusion .....	26
<b>Chapter 3 Coherent diffractive imaging and ptychography</b> .....	<b>27</b>
3.1. Introduction .....	27
3.2. Single shot CDI .....	28
3.3. Ptychography.....	32
3.3.1. Iterative ptychography.....	33
3.3.2. Sampling in ptychography.....	37
3.3.3. Super resolution in ptychography.....	39
3.3.4. Upsampling in ptychography.....	40
3.3.5. Scan position correction .....	41
3.3.6. Error metric for ptychography .....	42
3.3.7. Fourier ptychography .....	43
3.4. Conclusion .....	45
<b>Chapter 4 Modal decomposition of partially coherent wave field via ptychography</b> .....	<b>47</b>
4.1. Coherence .....	47
4.1.1. Temporal coherence .....	48
4.1.2. Spatial coherence .....	50
4.1.3. The Van Cittert–Zernike theorem.....	52
4.2. Mixed-state ptychography .....	53
4.2.1 ePIE-based modal decomposition algorithm.....	53
4.3. Probe modes .....	56
4.3.1. Probe mode power separation in the reconstruction.....	62
4.3.2. Factors influencing modal decomposition.....	64
4.4. A sample of object state modes.....	70

4.5. Conclusion .....	71
<b>Chapter 5 Engineering illumination modes in X-ray ptychography.....</b>	<b>73</b>
5.1. Simulation theory .....	74
5.2. Round aperture tests .....	74
5.3. Corners simulation experiment.....	81
5.4. Corner aperture lens simulation .....	84
5.5. The corner-aperture lens-simulation experiment.....	85
5.6. Simulation discussion .....	89
5.7. Diamond experiment.....	90
5.7.1. The Diamond Light Source and beamline I13 .....	90
5.7.2. Experiment setup.....	91
5.7.3. Ideas to prove with experiments.....	93
5.7.4. The experiment processes.....	94
5.5. Conclusion .....	105
<b>Chapter 6 Damaged specimen reconstruction by ptychography.....</b>	<b>106</b>
6.1. The interaction between electrons and specimen atoms .....	106
6.2. Type of radiation damage in TEM .....	108
6.3. Method used to minimise radiation damage.....	110
6.4. Simulation.....	112
6.4.1. The damage probability model .....	112
6.4.2. The damage probability between big and small probes .....	113
6.5. Using mixed state ptychography to improve the reconstruction of the damaged specimen .....	117
6.5.1. Simulation without noise .....	118
6.5.2. Simulation with noise .....	122
6.6. Analysis .....	124
6.7. Conclusion .....	125
<b>Chapter 7 Detector correction algorithm.....</b>	<b>126</b>
7.1. The detector and the detector mechanism.....	126
7.2. Detector Correction.....	130
7.3. The simulation .....	132
7.3.1. Non-linear noiseless detector .....	135
7.3.2. Non-linear and random gain on all detector pixels .....	137
7.3.4. Non-linear, random gain and noise on all detector pixels.....	139
7.4. Conclusion .....	140
<b>Chapter 8 Conclusion.....</b>	<b>141</b>
<b>References .....</b>	<b>147</b>

# Chapter 1 Introduction

One aspect of human nature is to continually explore the unknown world – from the enormity of the universe to a microscopic atom. As the only tool that enables humans to explore atomic structure, microscopy has hugely boosted the development of biology sciences, materials sciences, and medicine. Resolution is the most important property of a microscopy.

According to Abbé's imaging theory [1], the resolution of optical microscopy is limited by the diffraction of light; more specifically, it is limited by the wavelengths of visible light and the numerical aperture of lenses. As lenses for light can be manufactured with a large numerical aperture, resolution is mainly only limited by the wavelengths of light currently. Therefore, resolution can be further improved by using electromagnetic waves with a smaller wavelength, like X-rays, or particle waves, like high energy electrons. Unfortunately, lenses for these wavelengths are very hard to make, so their numerical aperture is very small (see below).

However, a new technique known as the coherent diffraction imaging (CDI) method has gradually attracted more attention from researchers. This method can present a high-resolution reconstruction without the use of a lens. This thesis will study applications of ptychography, which is now a widely recognized CDI method that provides the highest resolution and contrast images possible.

Ptychography has experienced rapid development and wide and increasing application over the last decade. Ptychography is unlike conventional microscopy because lenses are not a necessary component in the formation of an image, so it has a huge numerical aperture in theory and can avoid aberrations caused by a lens of poor quality. The resolution of ptychography is only limited by the size of the detector.

Ptychography works not only in visible light but also for x-ray and electron imaging. Historically, at these wavelengths, the maximum aperture size of conventional lenses has been minuscule (e.g. using 0.1mrad for x-rays and 30mrad for electrons). It is at these wavelengths where ptychography proves to be so effective. Most recently, in 2018, a Cornell University research group, led by Muller, achieved a resolution record that had never been seen before [2] in any type of transmission microscope. See a commentary about this achievement in [3].

The secret behind this resolution is in the redundancy of the ptychography data, where there is much more recorded information than is necessary to produce an image. In this thesis, we will focus on the application of the redundancy of these types of data. First, we developed a method to speed up the ptychography experiment using synchrotron x-ray ptychography; this method has been proved effective by experiments we conducted at the Diamond Light Source synchrotron in Oxford. Second, we proposed a method to minimise the effect of poor partially coherence by engineering some known pattern of illumination modes. Third, we attempted to improve the image quality of electron ptychography when knock-on damage occurred. Finally, we developed an algorithm to correct non-linear detector response by using the redundancy of ptychography data without any other independent calibration measurement. The first three these studies were conducted by employing a multimodal decomposition method, where we explored the limits of modal decomposition. The last application was based on the self-consistent response characteristics of the detector; we recorded a series of diffraction patterns using the same pixels on the detector, but from many parts of the object.

As a diffractive imaging method, coherence is extremely important for ptychography. However, in most experiments, the irradiating beam is more or less partially coherent, which strongly disturbs image reconstruction and reduces image quality. Modal decomposition has been proven and employed to overcome the influence of partially coherence in optics [4,5].

In 1992, Rodenburg and Bates proved that in theory a ptychographic dataset could be used to remove the effects of partially coherence in the illuminating beam [6]. A more effective method was developed in 2013 by Thibault et al. [7] using iterative solution

methods. However, if we have no information regarding the degree of partially coherence in the beam, many modes will be needed for the reconstruction, which will require computers to have greater calculation abilities and may take a longer time. In this thesis, we examine in detail mixed state ptychography in the context of engineering illumination modes to show that we can retain a reconstruction of good quality and can speed up the scan using only a few modes with partially coherent illumination. We demonstrate this experimentally with X-rays.

With the exception of using modes in probes, we also explore the use of modes in the object's reconstruction. Radiation is often the main experimental limitation in electron microscopy, especially when biological structures are considered [8]. The interaction between radiation and specimen is also very important in diffractive imaging methods. However, this interaction also can lead to radiation damage on or in the specimen, which must be taken seriously. This is especially true in the electron microscopy field, for the electron is a kind of particle with a rest mass, and charge damage can be quick and severe. Radiation damage can change the structure and properties of a specimen by heating and or charging it, or even knocking atoms completely out of their positions [9,10]. The damage can be minimised by using low dosage techniques, coating the specimen, and reducing the current or energy of the illumination beam [11]. In this thesis, we propose another method to minimise the effects of radiation damage on a reconstruction: treating the damaged specimen as another object state in a multimodal decomposition.

Apart from partial coherence, a redundancy in ptychography data can contribute to many different applications, and one such application is to correct detector non-linearity. The quality of the detector also limits reconstructions made by ptychography. Because of the manufacturing process, a detector pixel may produce some non-linear effects – creating more apparent flux or producing different levels of flux sensitivity – which could guide the iterative ptychography algorithms to a wrong solution. Therefore, it is necessary to correct the non-linearity and influence of random gains of the data. Thanks to the huge diversity of information on ptychography data, we can use anyone-pixel's recorded intensity collected from many different areas of the specimen to amend abnormal values.

## 1.1. The outline of this thesis

This thesis investigated the mechanism and influence factor of the mixed state decomposition method, which is one of the most important applications of the redundancy. Meanwhile, based on the research, a simulation experiment has been designed to speed up the ptychography experiment under a partial coherent illumination by engineering illumination modes. Later, a synchrotron x-ray experiment was completed to prove the simulation results that using engineered illumination modes can not only improve the reconstruction quality, but also speed up the ptychography experiment. This thesis demonstrates another important implication of the mixed stated decomposition method that it can be used to represent the undamaged object image if the object has been damaged during the experiment. In this thesis, except the mixed stated decomposition method, another application of the redundancy has been demonstrated that the rich redundancy in ptychography data can be employed to correct any non-linear response of detector pixels.

**Chapter 2** provides the general background knowledge of some important concepts and devices. This chapter is separated into three main parts: the lens imaging, the synchrotron, and the propagation approximation. First, we introduce the lens imaging system and derive the most important aspects of Abbé's theory of imaging from it. Then, we introduce the structure of a synchrotron facility, the properties of a synchrotron x-ray and the design of a Fresnel zone plate (FZP). Finally, we conclude the wave propagation and diffraction approbation based on a Fourier transform, which is helpful to understanding the calculation processes in this thesis.

**Chapter 3** gives a review of the CDI method and provides more details of ptychography. We begin the review from the phase problem, which is the purpose of CDI methods. To solve the phase problem, we introduce two important phase retrieval algorithms: the Gerchberg and Saxton (GS) algorithm [12] and the hybrid input–output (HIO) algorithm [13]. We then introduce ptychography, which is a type of CDI method but has a unique overlapping scan process. Compared to other CDI techniques, ptychography has a great deal of information redundancy within its data, which provides it with the ability to correct experimental errors in the reconstruction, produce

a large field of view, and enable high noise tolerance. We demonstrate two widely employed algorithms in ptychography: the ptychographic iterative engine (PIE) [14] and the extended ptychographic iterative engine (ePIE) [15]. We review the sampling requirements of ptychography as well as some important applications and sources.

**Chapter 4** discusses spatially mixed state ptychography. There are two kinds of application in spatially mixed state ptychography: the probe mode and the object mode. To begin, we introduce the ePIE-based mixed state algorithm. Then, by using this algorithm, we attempt to engineer a partially coherent source and examine the source modes. Because we know the exact source mode in these modelling experiments, we can investigate the modal decomposition of the mechanism of the probe modes, the factors that affect the performance of modal decomposition (such as the sampling condition) and the specimen and probe mode structure. We ascertain that each mode will conflict with another during reconstruction. Finally, we give an example of how object modes work.

**Chapter 5** demonstrates a proposed x-ray ptychography experiment in which we use a certain number of engineered illumination modes in a partially coherent ptychographic x-ray experiment to improve the reconstruction quality, reduce the requirement for calculation ability, and increase the speed of the experiment. This chapter comprises two parts. The first part is based on the simulation before the experiment in which we compare the reconstruction quality of a normal, large aperture and four engineered apertures under a partially coherent illumination with and without lenses using the Fourier ring correlation. In this test, the four engineered apertures provide four modes that are self-coherent but totally incoherent with others. The second part is based on the experiment itself, where we introduce some information regarding the Diamond Light Source's beam I13. The experiment's setup is described in detail. The idea behind this experiment is that by engineering suitable modes of illumination, we can reduce both the computational requirements and the time the experiment takes to be completed. We use two tests, using hole masks and wires to block parts of the beam, and we demonstrate our experiment's results using the wired test set, in which we compare the reconstruction quality with a wire crossing the beam, first parallel and then horizontal to the coherence direction, using an FZP-only

experiment as a reference. The results of our experiment support our idea and have the potential to speed up scans.

**Chapter 6** examines an idea that it should be possible to use object modes to reconstruct an image of a damaging object, say in the electron microscope, before it was damaged. At the beginning of this chapter, we introduce the interactions between electrons and specimen atoms, demonstrate how these interactions can cause damage to the specimen and explain how we can increase the accuracy of an image obtained through TEM experiments. Also, we suggest a function that can calculate the probability of damage based on the flux, dose rate or exposure time. We then explain why a large probe has a reduced chance of damaging the specimen compared to a small probe if the total flux is the same. Next, we review the reconstruction results of a modelled damaged dataset by using object modes.

**Chapter 7** suggests a detector non-linearity correction method for ptychography data as ptychography data contains extremely diverse information regarding the features of the specimen, so it has the ability to self-correct any error in the data. In this chapter, we use a polynomial to fit the recorded values by each detector pixel with a polynomial function. We test the performance of the algorithm for a modulus image and phase image ptychography data with sine function non-linearity on the detector pixels; sine non-linearity and random gain on each detector pixel; and sine non-linearity, random gain, and some background offset on the detector pixels. The results show that the algorithm works well for modulus specimen data in all situations, but there are some limitations to the phase-only specimen data.

**Chapter 8** concludes the above-mentioned works of this thesis and summarises the research that has been conducted.

## Chapter 2 General background

In this chapter, we will introduce the lens optical imaging system and its resolution limitations. We will briefly discuss the synchrotron light source regarding the properties of a synchrotron x-ray, the insertion device, and the characteristics of an FZP. Furthermore, we will review the scalar wave theory and propagation approximations.

### 2.1. Lens imaging

In conventional imaging, the lens has played an essential role in focusing and magnifying an image of a specimen since the advent of the microscope. Figure 1 presents a simple model of optical microscopy, where the focusing and magnification optical components exist similarly in electron microscopy and x-ray microscopy as an electromagnetic lens [16] and an FZP [17] respectively.

In Figure 1, a light wave emitted from a point source passes the condenser aperture and can be focused by the condenser lens onto the specimen. The exit wave goes through the objective aperture and is collected by the objective lens to form an inverted, visible image of the specimen. The inverted image is further magnified by the projector's lens to present a larger image, where the magnification is equal to

$$M = \frac{\textit{magnified image size}}{\textit{real specimen size}}. \quad (1)$$

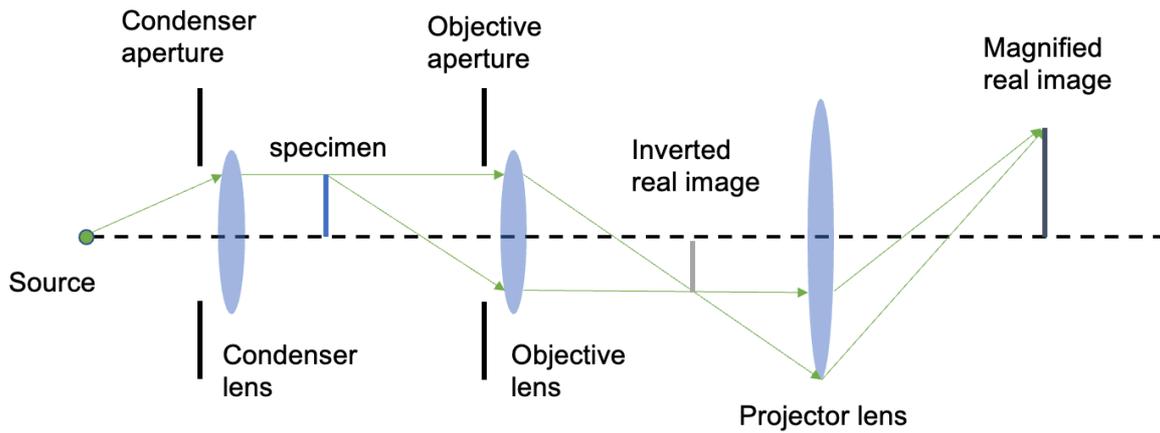


Figure 1. A simple lens-based microscopy schematic diagram. The radiation is emitted from the source and passed through the condenser aperture, which controls which fraction of the radiation is used in the following process. Radiation that passes through the condenser aperture is collected by the condenser lens, diverting the radiation toward the specimen to create more illumination on the specimen. Then, the diffracted light is selected by the objective aperture before entering through the objective lens to form a magnified image. The magnified image is produced through the projector's lens to provide further magnification. (There is usually no projector lens in an X-ray microscope.)

However, the distance between specimen and lens must lie between the focal length and twice the focal length. Figure 2 (see next page) shows a single lens magnification process, in which  $L_s$  is the distance between specimen and lens,  $L_f$  is the focal length of the lens and  $L_i$  is the distance between lens and magnified image. According to Newton's lens equation

$$\frac{1}{L_f} = \frac{1}{L_s} + \frac{1}{L_i}, \quad (2)$$

the magnification can also be calculated by

$$M = \frac{L_i}{L_s} = \frac{L_f}{L_s - L_f}. \quad (3)$$

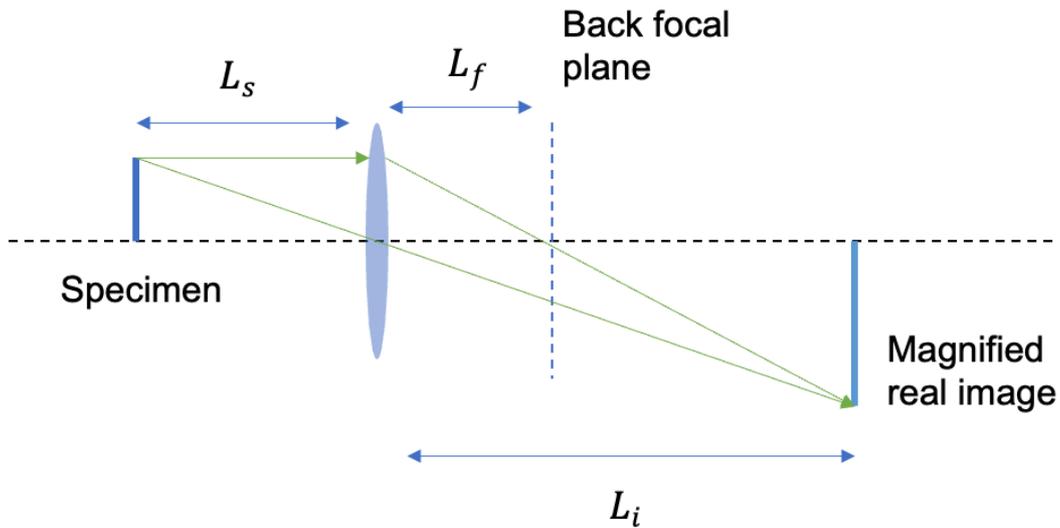


Figure 2. A single lens magnification process schematic diagram. To produce a magnified image of the specimen at the  $L_i$  distance, the  $L_s$  (the distance between specimen and lens) must be between  $L_f$  and  $2L_f$ .

## 2.2. Resolution limitation

In 1873, Abbé pointed out that the image formation in microscopy can be separated into two interfere imaging stages [1]: the diffraction into the back focal plane followed by re-interfering in the image plane. When the specimen is illuminated by plane waves, each frequency component of the specimen can diffract the radiation, and the smaller components (higher frequency) produce a larger diffraction angle. The scattering angle of the diffracted waves also depends on the wavelength of the illumination.

After the diffraction process, the specimen's structure is stored as spatial frequency components in reciprocal space, where a higher spatial frequency highlights the specimen's finer details. An objective lens is fixed downstream of the specimen to collect and redirect the diffracted waves onto the image plane, as shown in Figure 3, where the final image on the image plane results from the interfere between different diffracted waves.

Essentially, the whole process is two Fourier transform processes [18,19]. The first Fourier transform gives the frequency distribution at the Fourier domain, which is

reprojected to the spatial domain by another Fourier transform. The finest resolution, according to Abbé's theory, is given by

$$resolution = \frac{\lambda}{2NA'} \quad (4)$$

where  $\lambda$  is the wavelength the illumination wave and  $NA$  is the numerical aperture of the objective lens, which describes the half angle that the maximum light cone can be collected by a lens, as shown in Figure 4. This can be equivalent to

$$NA = n\sin(\theta), \quad (5)$$

where  $n$  is the refractive index for the wave in the medium and  $\theta$  is the largest semi-angle of light that can be collected by the lens.

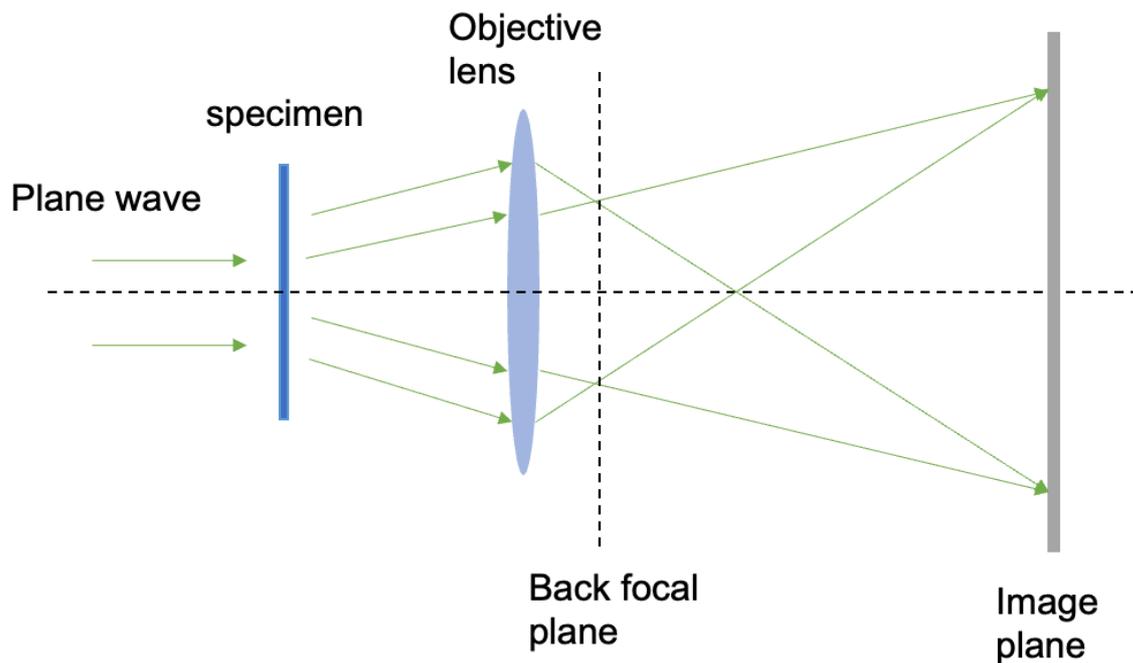


Figure 3. The schematic diagram of Abbé's image theory. The objective lens collects and focuses the interfere of diffracted waves at the rear focal plane when a specimen is illuminated by a parallel light, which produces the first interfere image of the specimen. Then, another interfere occurs between the rear focal plane and image plane, and a second interfere image is captured on the image plane. The second interfere is necessary to image formation; without it, it would be impossible to produce an image of the specimen.

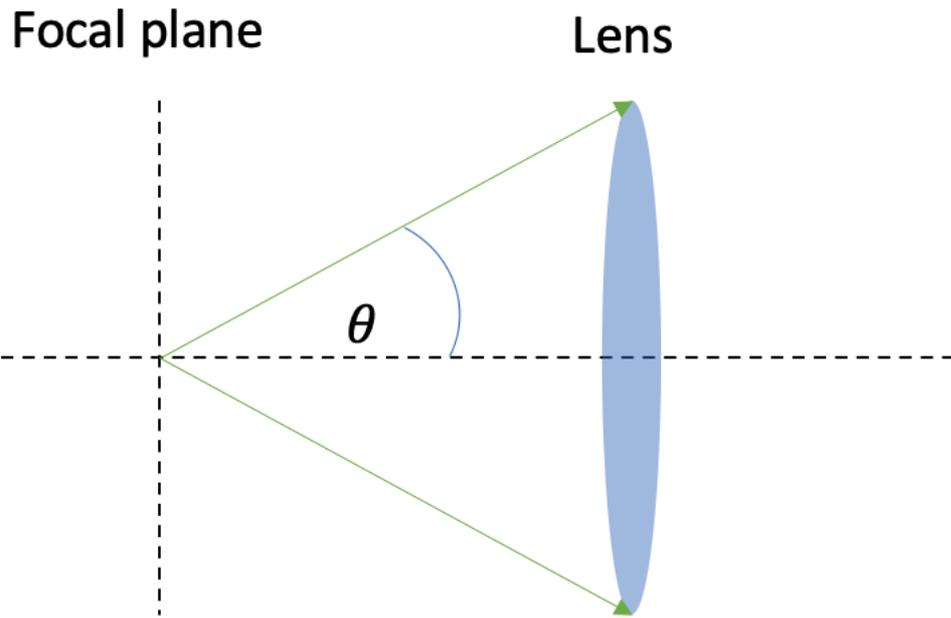


Figure 4. A schematic diagram of a numerical aperture. If the light is emitted from one point, the numerical aperture is the half angle that the maximum light cone can be collected by a lens.

In practice, because the lens has a finite size and each wave passes through a lens or aperture, diffraction is determined by the widest part of the lens or aperture. Hence, the wave is focused as a disk rather than a point, which is called the Airy disk. If there are two points on the specimen that are too close separate and where Airy disks overlap, the resolution can be calculated as

$$resolution = \frac{0.61\lambda}{NA}, \quad (6)$$

which is called the Rayleigh resolution limitation [20].

From Abbé's theory and Rayleigh's resolution limitation, it is easy to ascertain that the resolution only relates to the wavelengths of light and the manufacturing technique of the lens. Today, it is possible to produce a high NA lens. As a result, the resolution of optical microscopy is mostly only limited by the wavelengths of light. To overcome the limitations of resolution in optical microscopy, x-ray microscopy and electron microscopy have been examined and systemically developed.

Nevertheless, the lens itself has some shortcomings that can restrict performance, including aberrations. As described by Goodhew [21], the most common aberrations in electron lenses include the spherical aberration, astigmatism, and chromatic aberration. Some of these are also observed in x-ray microscopy [22] [23].

## 2.3. X-Ray synchrotron and FZP

### 2.3.1. X-ray sources

X-ray imaging was first found by German physicist Wilhelm Rontgen in 1895, and it has widely been used in biological and medical research over the last century. X-ray imaging, as a form of electromagnetic radiation, has a higher energy and smaller wavelength than visible light, so as Abbé's theory suggests, an x-ray can provide a higher resolution when used to replace the visible light in microscopy.

Early x-rays comprised x-ray tubes and a rotating anode source, where free electrons with a given kinetic energy would pass through a vacuum and hit the target metal anode until they came to rest. During this interaction, Bremsstrahlung x-ray [21] is generated. However, the issue with x-ray tubes and rotating anode sources is that they have much less brilliance (a measure of brightness) than a synchrotron light source, which we will discuss later [24,25]. Brilliance is a key parameter in measuring how an x-ray source performs, and the brilliance is described by

$$brilliance = \frac{j}{\theta^2 s^2 (0.1\% B_e)}, \quad (7)$$

where  $j$  is the x-ray flux,  $\theta$  is the spread speed of the beam,  $s$  is the cross section of the beam and  $B_e$  is the energy's bandwidth. A higher brilliance means that there is more focused flux within a given wavelength and within a unit of time.

Before X-Ray Free Electron Laser (XFEL), synchrotron light sources were recognised as the most advanced x-ray source. A synchrotron light source has key advantages, such as high brilliance, high coherence degree, broad energy spectrum and

collimation [26]. Rather than bremsstrahlung x-rays, a synchrotron produces an x-ray by directing charged particles (electrons or positrons) to move in a curved trajectory within a periodic magnetic field [27,28,29] called an insertion device. The structure of a synchrotron light source can be divided into an electron source, booster ring, storage ring, beam line, insertion device and end station, which includes the experimental hutch and workstations, as shown in Figure 5.

Electrons produced from the electron source, which is normally an electron gun, are injected into the booster ring after being passed through a linear accelerator. In the booster ring, electrons are accelerated to the required energy of the storage ring before travelling to the storage ring. In the meantime, some new electrons are injected into the booster ring to maintain the beam's current. The storage ring is the most important component of a synchrotron light source. It is in this ring that electrons travel almost at the speed of light in the vacuum and a very high kinetic energy measured in GeV. The electron moves in a closed-end polygon tube, in which some magnetic components bend and focus the electron beam. Electrons in the storage ring gradually lose their energy and emit a variety of radiations, including a spectrum of x-rays. The produced x-rays pass through a different beamline to the end stations for different applications, such as to explore the structure of a biology specimen or materials, x-ray absorption spectroscopy [30] or x-ray microscopy [31,32].

To produce high brilliance x-rays, some insertion devices (ID) are used, and the x-ray is emitted when the electron beam passes each magnet. X-rays are then guided by beamlines to the end station. This method is called the third-generation synchrotron light source, as shown in Figure 6.

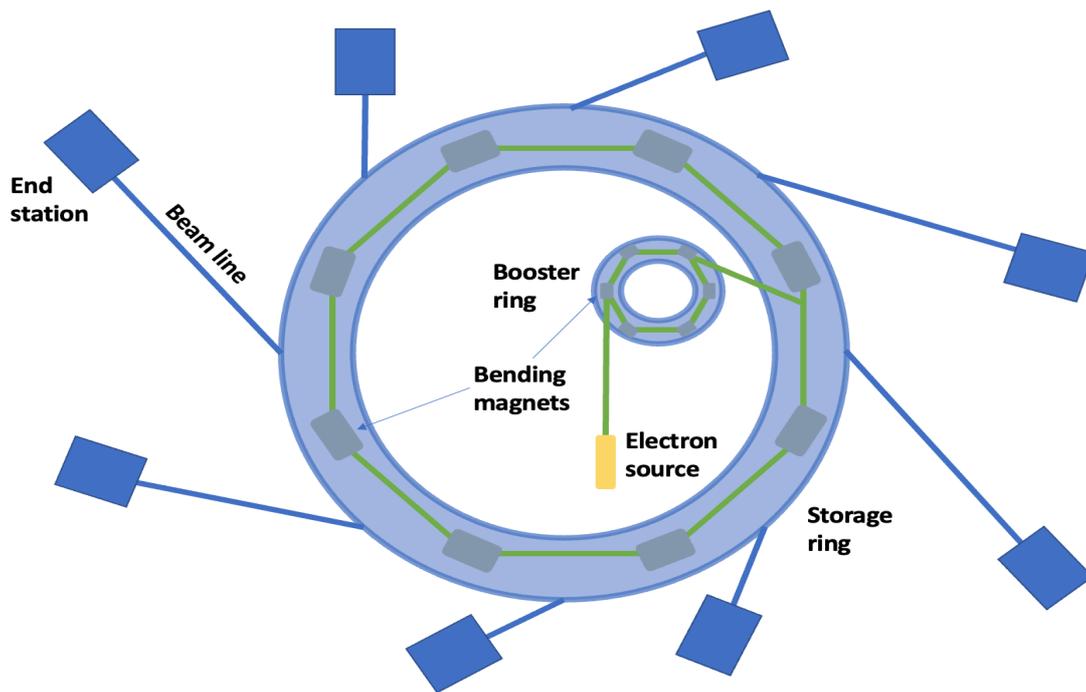


Figure 5. A simple schematic of a synchrotron light source. The main component of a synchrotron light source includes an electron source, booster ring, storage ring, beamlines and end stations. The electron beam is emitted from the electron source and accelerated by the boost ring to a speed that is high enough to allow it to be injected into the storage ring. The radiation from the x-ray is generated in the storage ring as the electrons gradually lose energy and are passed along the beamlines to the end station.

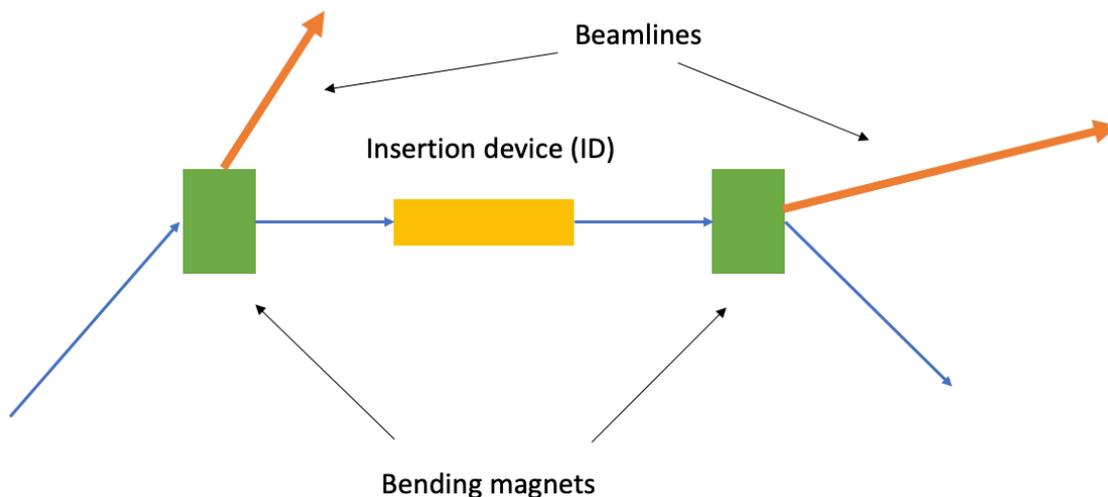


Figure 6. A simple schematic of a third-generation synchrotron light source storage ring to produce radiations. The electron beam in the storage ring is bent by the bending magnets to maintain its trajectory. Between the two bending magnets is an ID device, which is also like a well-designed magnet to enhance the brilliance and coherence of the emitted x-rays.

Widely employed IDs include a wiggler and undulator [28,29], in which are well-designed periodic magnetic devices. This is shown in Figure 7, where IDs are inserted into the storage ring between bending magnets to continually deflect the electron beam. IDs can help to improve the brilliance, coherence, and spectral range of synchrotron radiations without disturbing the circulating beam's trajectory through displacement or deflection. Wiggler magnets have a stronger magnetic field compared to bending magnets, so the energy spectrum of the electron beam will be enlarged. Some spectrums have a higher photon energy, which helps to improve the overall output power of the electron beam. In contrast, an undulator performs more like an energy spectrum filter, where magnets have fewer magnetic fields than bending magnetics but are denser. Because the excursion angle is smaller than the bending angle, some individual undulations to the input electron beam, and the undulations will interfere others downstream constructively or destructively. Therefore, in some frequencies, the power of the electron beam is enhanced to produce a high brightness, dependent on the design of the undulator. To examine what types of IDs are in the storage ring, an undulator strength parameter can be used, which is a representation of the motion characteristics of the electron beam. If  $K$  is the undulator strength parameter

$$K = \frac{eB_U\lambda_U}{2\pi m_r c} \quad (8)$$

then  $e$  is the charge of an electron,  $B_U$  is the strength of the magnetic field of the undulator,  $\lambda_U$  is the spacing distance between two magnetic fields,  $m_r$  is the rest mass of an electron and  $c$  is the speed of light in a vacuum. If  $K > 1$ , the magnets perform like a wiggler, but if  $K < 1$ , the magnets play a role of undulator. If  $K \sim 1$ , the magnets behave like a normal bending magnet.

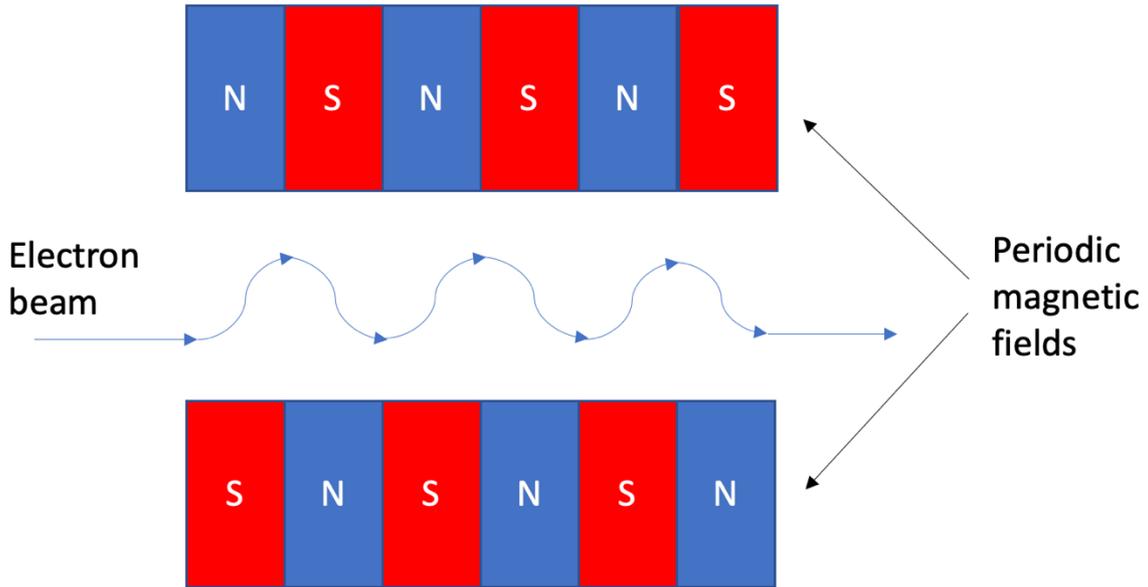


Figure 7. The schematic of the magnetic fields of a wiggler and undulator. When an electron beam passes through the periodic magnetic fields, it is deflected by a certain angular excursion, but the magnetic fields in the ID device do not deflect or displace the beam's trajectory in the storage ring.

More recently, diffraction-limited storage rings have been designed by Hettel [33,34], who is inspired by the multiple bend achromatic concept [35] of improving the coherence degree and brilliance of current synchrotron sources. Next-generation synchrotron light sources, called x-ray-free electron lasers (XFELs) [36,37] that are under construction in several countries can provide about  $10^{22}$  times more brilliance.

### 2.3.2. Properties of synchrotron x-rays

We will now introduce some of the properties of a synchrotron light source. As introduced by Einstein [38], the kinetic energy of a particle is described as

$$E = \frac{mc^2}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \quad (9)$$

where  $m$  is the mass and  $v$  is the speed of the particle. However, with the speed being close to the speed of light ( $c$ ),  $m$  will approach infinity. In the storage ring, when  $v$  is almost equal to  $c$ , the centripetal force is provided by the Lorentz force, and we can understand the relationship through [39]

$$E = ecBR, \quad (10)$$

in which  $B$  is the magnetic induction and  $R$  is the radii of the storage ring. So, the beam energy has a positive relationship with the magnetic density and size of the storage ring. Further, the critical energy of a synchrotron light source can be measured as the half power of a magnet, which is used to bend the beam, for

$$E_c = 0.665E^2M. \quad (11)$$

For a synchrotron light source, Paganin [26] describes the early research about it as the critical wavelength of a synchrotron light source is equal to

$$\lambda_c = \frac{4\pi r_M}{3} \left( \frac{mc^2}{E} \right)^3, \quad (12)$$

Where  $r_M$  is the arc radius of electrons in the bending magnetic field. And that the concentrated half angle of a synchrotron light source is

$$\theta \approx \sqrt{1 - \frac{|v|^2}{c^2}}. \quad (13)$$

### 2.3.3. Properties of the FZP

In an x-ray experiment, resolution is mainly limited by the quality of the optical system, which depends on the requirement of an experiment. However, glass lenses cannot be used because of absorption. Thus, some new optical elements are used to replace the lens, such as the FZP [17] and the Kirkpatrick Baez (KB) mirror [40]. As we have not used a KB mirror in this thesis, we will not discuss it here. The FZP is formed by a

series of concentric circles – called zones – that are opaque and transparent in succession. In an x-ray experiment, the FZP optical system is fixed, as shown in Figure 8. From left to right, we see the x-ray beam, the FZP with a central stop as the condenser lens in this system, the order selecting aperture (OSA), the specimen and a detector at the far end.

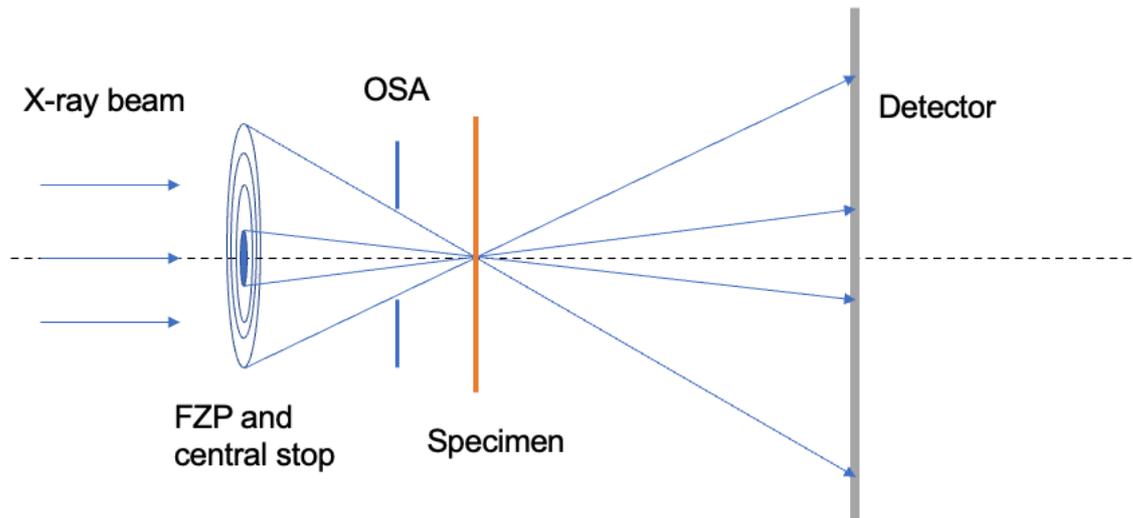


Figure 8. A simple schematic of the FZP optical system. The coherent parallel x-ray radiations pass through the FZP, which acts as a radiation-focusing device and be focused on the specimen. Before reaching the specimen, an OSA helps to block unwanted diffractions following the FZP. A detector is placed on the far end of the diffraction plane to collect the diffraction.

To focus the light, the FZP depends on the light's diffraction rather than refraction; where the light passes through transparent zones, the exit waves should produce constructive interference at the focus. If  $r_n$  is the radius of  $n$ th zone, where  $\lambda$  is the wavelength of the radiation and  $L_f$  is the first order focal length, to keep the light in focus, the light passing through the zone plane should have this relationship:

$$\sqrt{r_n^2 + L_f^2} = L_f + \frac{n\lambda}{2}. \quad (14)$$

Finally, we achieve the equation [24]

$$r_n^2 = n\lambda \left( L_f + \frac{n\lambda}{4} \right) \quad (15)$$

by expanding the above relationship, where  $L_f$  can be calculated by

$$L_f = \frac{4nr^2}{\lambda} \quad (16)$$

and  $r$  is the width of outer-most zone. Due to Abbé's theory, if the wavelength is fixed, a large NA will create the highest resolution, and the NA of the FZP is controlled by the outer most zone's width by

$$NA = \frac{\lambda}{2r}. \quad (17)$$

So, the resolution of the FZP is equal to

$$resolution = 1.22r, \quad (18)$$

which means that the resolution of the FZP is directly related to the outer-most zone's width.

The central stop and OSA are used to block the zero- and negative-order diffractions, where the zero-order component is a straight light and the negative-order components diverge, so they cannot be focused on one focus point. To form a unique, strong focal point, all the orders have strong positive interference with the first order zone will be kept, and the other orders, that will disturb above interference, need to be blocked. Therefore, the central stop is placed before the zone plate to block the zero-order straight light, and the OSA is fixed after this to block unwanted diffractions from different zones.

## 2.4. Calculation approximates of wave propagation

Before we discuss decoding algorithms, we must understand the main approximations regarding wave diffraction. Many books have presented a wave equation derivation process using Maxwell's equations [41,42]. A brief derivation process is presented here to show how the Fourier transform can be used to analyse optics. Maxwell's equations suggest the propagation behaviours of optical waves, and it can be simplified as scalar diffraction function

$$\nabla^2 g(p, t) - \frac{n^2}{c^2} \frac{\partial^2 g(p, t)}{\partial t^2} = 0, \quad (19)$$

in which  $\nabla^2$  is the Laplace operator,  $n$  is the index of refraction of a medium,  $c$  is the speed of light in vacuum and  $g(p, t)$  is the disturbance of point  $p$  at time  $t$  in a linear, isotropic, homogeneous, nondispersive, and nonmagnetic situation. For monochromatic light,  $g(p, t)$  can be written as

$$g(p, t) = \text{Re}\{A(p)e^{i\phi(p)-i2\pi ft}\}, \quad (20)$$

where  $A(p)$ ,  $\phi(p)$  and  $f$  are the amplitude, phase, and frequency of the wave, and  $\text{Re}$  indicates the real part. A time-invariant scalar wave equation can then be factored out:

$$(\nabla^2 + k^2)Ae^{i\phi} = 0. \quad (21)$$

This equation is called the Helmholtz function, where  $k$  is the wave number:

$$k = \frac{2\pi n f}{c} = \frac{2\pi}{\lambda}, \quad (22)$$

in which  $\lambda$  is the wavelength of a wave. A general solution of the Helmholtz function is called the Rayleigh–Sommerfeld diffraction formula, which applies when there are two parallel planes with distance  $d$ . The left and right planes are the producing and receiving planes, respectively, as shown in Figure 9.

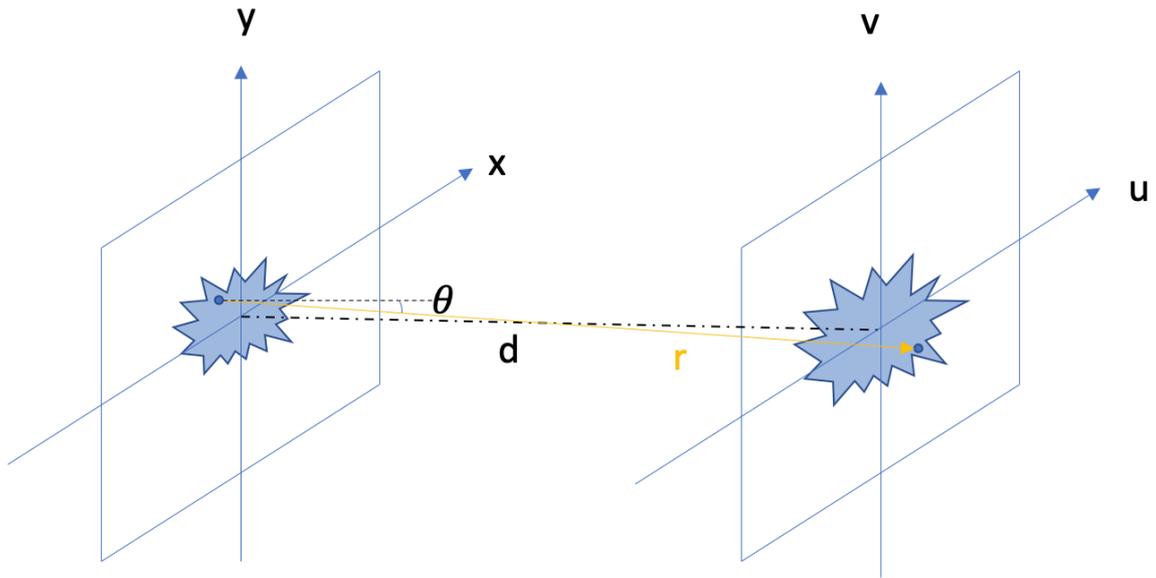


Figure 9. The wave propagation in free space between two parallel planes. The diffraction approximation depends on the distance between the two planes.

If a wave is emitted from a point at the start plane with angle  $\theta$  and is received at the receiving plane with a propagation distance,  $r$ , then the Rayleigh–Sommerfeld diffraction formula is

$$U_{receive}(u, v) = \frac{d}{i\lambda} \iint U_{start}(x, y) \frac{e^{ikr}}{r^2} \cos(\theta) dx dy, \quad (23)$$

which gives an accurate description of the wave propagation between planes and, inside the formula,

$$r = \sqrt{(u - x)^2 + (v - y)^2 + d^2}, \quad (24)$$

$$\theta = \arccos\left(\frac{r}{d}\right) \quad (25)$$

And

$$U = Ae^{i\phi}. \quad (26)$$

If  $d$  is small, the  $\cos(\theta) \approx 1$ . The Rayleigh–Sommerfeld diffraction formula will then be the same as the Huygens–Fresnel principle. To simplify the Huygens–Fresnel principle, the Taylor expansion can be used to give an approximation of  $r$ , resulting in

$$U_{receive}(u, v) = \frac{e^{ikd}}{i\lambda d} \iint U_{start}(x, y) e^{\frac{ik}{2d}[(u-x)^2+(v-y)^2]} dx dy, \quad (27)$$

which is the expression of the Fresnel diffraction approximation. Using the Fresnel approximation to model a scalar diffraction, the distance between two planes must satisfy

$$d^3 \gg \max \left\{ \frac{\pi}{4} [(u-x)^2 + (v-y)^2]^2 \right\}, \quad (28)$$

which arises when the higher order Taylor expansion terms are ignored.

However, if  $d$  is quite large –

$$d \gg \max \left[ \frac{k(x^2 + y^2)}{2} \right] - \quad (29)$$

the Huygens–Fresnel principle can be simplified as

$$U_{receive}(u, v) = \frac{e^{ikd}}{i\lambda d} e^{\frac{ik}{2d}(u^2+v^2)} \iint U_{start}(x, y) e^{\left[ \frac{-i2\pi}{\lambda d}(ux+vy) \right]} dx dy. \quad (30)$$

This result is the Fraunhofer diffraction approximation. Except for the phase factor at the start, the expression can be written as Fourier transform results if

$$f_x \rightarrow \frac{u}{\lambda d} \quad f_y \rightarrow \frac{v}{\lambda d}. \quad (31)$$

A very useful tool that people can use to choose the correct scalar diffraction approximation is called the Fresnel number [43]. If a wave with wavelength  $\lambda$  passes

through an aperture with radius  $r$  to a detector plane at distance  $l$ , the Fresnel number is calculated as

$$F = \frac{r^2}{l\lambda}. \quad (32)$$

If  $F \approx 1$ , the Fresnel propagation should be used, but if  $F \ll 1$ , the propagation is closer to the Fraunhofer diffraction approximation. If  $F \gg 1$ , the angular spectrum method can better describe the propagation of the wave, but this method does not relate to this thesis and will not be discussed further.

The lens was the most indispensable component in the past century for microscopy, and it is still widely used to diverge and direct light beams in experiments. Therefore, it is important to know the transmittance function of a lens. As described by Goodman [42], as the refractive index of a lens is larger than that of air, when a wave penetrates the lens, there will be a delay. The degree of that delay depends on the thickness of the lens at the incident position. In other words, the lens produces a phase change to the incident wave, as

$$p_l(x, y) = e^{-i\frac{k}{2L_f}(x^2+y^2)}, \quad (33)$$

where  $k$  is the number of waves,  $L_f$  is the focal length of the lens and  $x$  and  $y$  are the space coordinates on the lens. This equation has an analogous application to an FZP. So, if the lens is illuminated by a plane wave,  $U(x, y)$ , and because the lens has a finite size, we use a pupil function,  $P(x, y)$ , to identify the light that passes through the lens. The exited wave can be calculated as

$$U'(x, y) = U(x, y)P(x, y)p_l(x, y). \quad (34)$$

To calculate the diffraction in the back focal plane, we placed the above equation (34) into an equation mentioned earlier (27) and replaced the propagation distance,  $d$ , by  $L_f$ :

$$U(u, v) = \frac{1}{i\lambda L_f} e^{\frac{ik}{2L_f}(u^2+v^2)} \iint U(x, y) P(x, y) e^{\left[\frac{-i2\pi}{\lambda L_f}(ux+vy)\right]} dx dy. \quad (35)$$

Except for the pupil function,  $P(x, y)$ , this expression matches exactly the Fraunhofer diffraction at the distance  $L_f$ , which means that we not only achieve the Fraunhofer diffraction of a plane wave at a long distance, but we can also obtain it through a suitable lens.

## 2.5. Conclusion

This chapter has reviewed some general information about imaging systems, beginning with the lens-based imaging system. We introduced the basic structure of a lens imaging system and the resolution limitations based on the theories of Abbé and Rayleigh. We then introduced some of the characteristics of the x-ray synchrotron light source and the FZP; later, we will use these devices experimentally to demonstrate the theoretical parts of this thesis. Further, we have introduced the wave transfer function and diffraction approximation, which are helpful to understanding the propagation calculation in this thesis.

# Chapter 3 Coherent diffractive imaging and ptychography

In this chapter, we will introduce coherent diffractive imaging (CDI). We begin with the phase problem and some widely employed algorithms that aim to solve it. Finally, we review the history of ptychography and some important tools in my ptychography research works.

## 3.1. Introduction

To achieve a higher resolution and to avoid small numerical aperture and aberrations caused by the X-ray and electron lenses, the CDI method has gradually attracted more attention from researchers. The general paradigm is shown in Figure 10, where the detector is employed to record the interference pattern of the object's scattered light under a substantial coherent source. This is the encoding process; a suitable algorithm should then be used to solve the object image.

From Figure 10, we discover an essential issue of the CDI method – the phase problem. When the light is scattered by the object, the exit wave is a complex one, which includes the modulus and phase information. At the exit wave of the specimen the modulus information gives the transmission of the specimen, and the phase gives the refractive index changed phase difference, they are both very important for measuring the structure and characteristics of the object. The detector in the farthest field can only measure the intensity values, and all phase information is lost. Thus, an essential part of the decoding process is to solve the phase problem. Once this is

completed in the farthest field, the solution can be backpropagated to the specimen's exit wave.

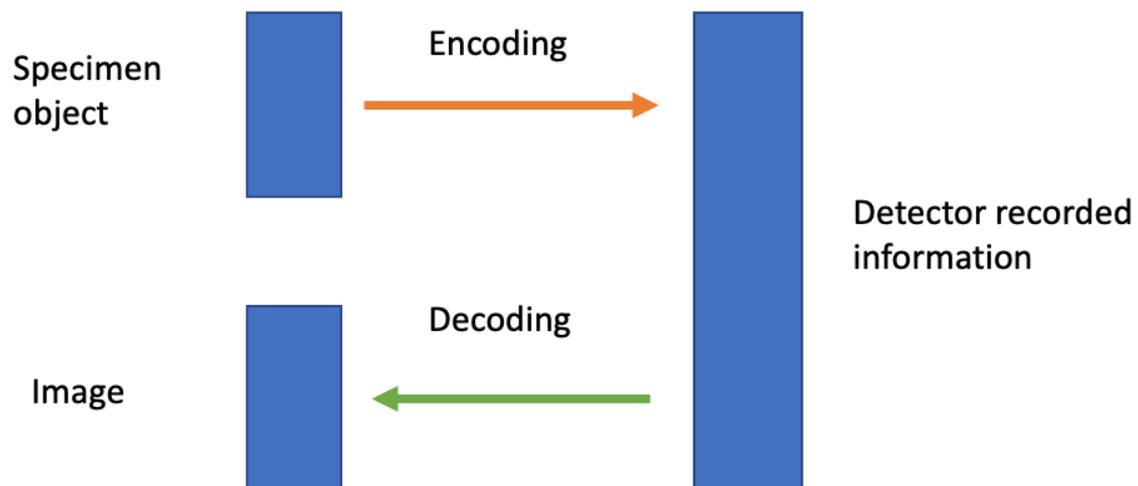


Figure 10. A general paradigm of the CDI method. In CDI methods, only the detector records the interference pattern, which provides much information regarding the specimen and radiation, but this can only be recorded in intensity. Due to the loss of phase information, the final image is obtained by decoding calculations.

### 3.2. Single shot CDI

The first widely known and successful phase-retrieval algorithm for solving the phase problem was published by Gerchberg and Saxton in 1972 [12]. The aim of this algorithm was to solve the phase problem in electron microscopy experiments. Solving this problem depends on recording both the image and diffraction pattern intensity. The GS algorithm is an iterative procedure; the basic idea of this algorithm is that an amplitude change in one domain will lead to a change in the amplitude and phase of another domain. To begin, the GS algorithm has three lots of input data: the generated random phase  $\phi$  that contains random values between  $\pi$  and  $-\pi$ , the recorded intensity in the object plane,  $I_o$ , and the diffraction plane,  $I_d$ . The iterative schematic is shown in Figure 11.

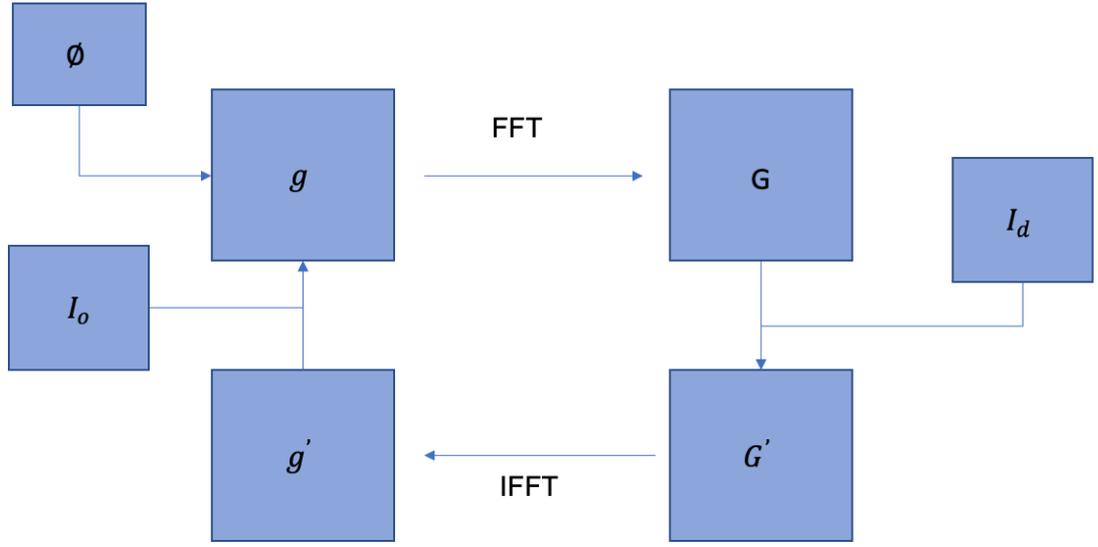


Figure 11. The iterative schematic of the GS algorithm. The real space is on the left, where the constraints on the image are employed. Reciprocal space is on the right-hand side, where the measured intensity as reciprocal space restrictions is applied. In each iteration, the phase of the reconstruction is not changed, only the (measured) modulus is readjusted.

At first, it will generate the guess image,  $g_n$ , using

$$g_n(u) = \sqrt{I_o} * \exp(i * \phi(u)), \quad (36)$$

where,  $u$  is the real space coordinate and  $n$  is the iteration number. Then, we apply the fast Fourier transform [44] to the guess image as the guessed exit wave propagate to the detector plane:

$$G_n(v) = f(g_n(u)). \quad (37)$$

At the detector plane, we apply the reciprocal plane's constraint to the  $G_n(v)$  by replacing the modulus section with the measured modulus value, but the phase remains the same. Then, we apply an inverse Fourier transform to get the updated exit wave  $g'_n(u)$ , for

$$g'_n(u) = f^{-1} \left( \sqrt{I(v)} \frac{G_n(v)}{|G_n(v)|} \right). \quad (38)$$

Finally, we apply the object plane constraint to the updated exit wave and repeat the whole procedure several hundred times, for

$$g_{n+1}(u) = \frac{\sqrt{I_o}}{|g'_n(u)|} * g'_n(u). \quad (39)$$

To prove that the error decreases by using their algorithm, Gerchberg and Saxton tracked the calculation error by using the square error, which is given by measuring the total amplitude's difference between the measured intensity and the calculated amplitude:

$$Er(n) = \sum |I(v) - |G_n(v)||^2. \quad (40)$$

When the error value is closer to zero, the reconstruction is closer to the correct value. However, the algorithm has its shortcomings, including the fact that the convergence of this algorithm may not go in the right direction. This could mean that when it has collected less information, the algorithm could drop into a local minimum, where the error can drop to zero or jump into an endless ambiguity solution loop without reducing the error.

Based on the GS algorithm, Fienup [13] provided a new object constraint to improve it, resulting in what is called the error reduction algorithm. Rather than update the whole object function, the new real space constraint can be presented as

$$g_{n+1}(u) = \begin{cases} g'_n(u), & \text{if } u \in s \\ 0, & \text{if } u \notin s \end{cases}, \quad (41)$$

where  $s$  is the support area. This assumes the object is finite, lying only within the support. So, during the updated process, only the supported area has been corrected.

Other values remain the same. To measure the error value during a calculation, the mean squared error method can be used:

$$Er_n = \frac{\sum \left| |G_n(v)| - \sqrt{I(v)} \right|^2}{\sum I(v)}. \quad (42)$$

However, although the error-reducing algorithm reduces the error quickly at the beginning, it gradually slows down. Therefore, to achieve a good reconstruction, the error-reducing algorithm needs a large number of iterations, and anyway can converge on local minima in the error function space. Fienup [13,45] further updated the error-reducing algorithm to a HIO algorithm. The new object updated function is given by

$$g_{n+1}(u) = \begin{cases} g_n(u) - \alpha g'_n(u), & \text{if } u \in s \\ g'_n(u), & \text{if } u \notin s \end{cases}, \quad (43)$$

in which  $\alpha$  is a feedback constant. Although the HIO algorithm improves the reconstruction quality, it still has some flaws, such as that the field of view is smaller.

It can be argued that the proposal of the CDI concept can be traced back to 1952, when Sayer found that the Shannon sampling theory could be used in crystallography to uniquely solve object structure by recording the intensity of the specimen's interference pattern on a detector [46]: the diffraction pattern. This idea was very close to Hoppe's concept of ptychography. However, it is debatable whether different objects have similar diffraction patterns [47,48], which would cause residual ambiguity, so it may need more measuring or constraints to guaranty a unique solution. Therefore, some different CDI methods have appeared under the basic CDI concept, like Fresnel CDI, Bragg CDI, scanning CDI and ptychography, which we will now explore in much greater detail. Ptychography is much more powerful than single shot CDI and is central to all the work that follows.

### 3.3. Ptychography

Ptychography is a branch of CDI, so it has all the advantages of diffractive imaging, including a high resolution and improved phase retrieval contrast. Also, ptychography has a huge amount of redundancy, which is its main advantage when compared with other CDI methods. Redundancy is based on the large number of diffraction patterns, so many experimental errors can be corrected during the reconstruction process [49]. Ptychography has better noise and partial coherence tolerance. It also makes lenses unnecessary, produces complex images and can self-calibrate. However, in the past, the application of ptychography has been limited by the calculation capacity of the computer being used and detector technologies.

Unlike other conventional CDI methods, in the scanning process, each of the two adjoining scan positions must have a certain level of overlap [6,50], as shown in Figure 12. When the radiation passes through an aperture, it is driven by a condenser lens to the surface of the specimen, either in focus or defocused, and the exit wave contains the information of probe and specimen in probe position will be record by a detector. The beam is then moved to the next position, overlapping on the previous position, and the next interference pattern is recorded by the detector. The process continues until the target area has been fully scanned.

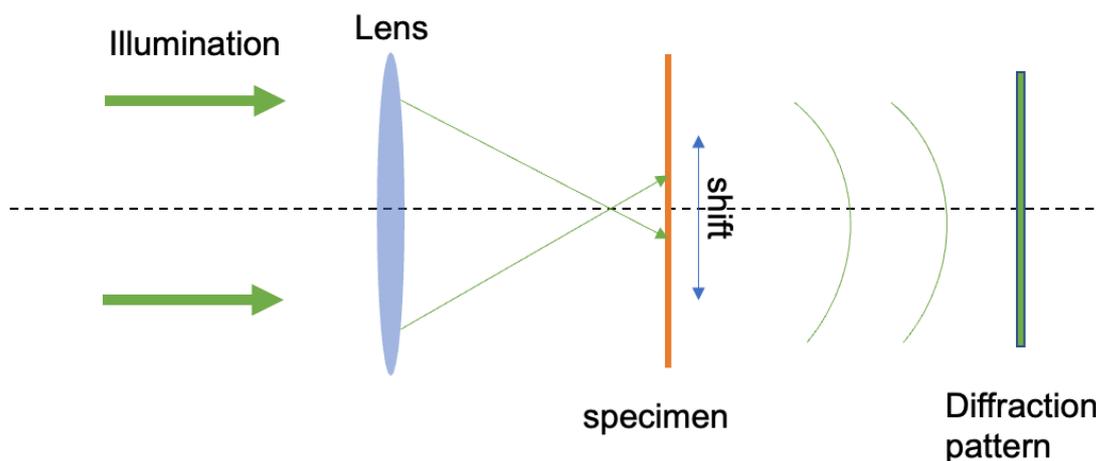


Figure 12. A sketch map of the ptychography scanning process. The light is focused by a lens onto the specimen, and a related diffraction pattern is recorded at the far end. By moving the

specimen, the light can scan each area of it and generate a series of diffraction patterns, where each of the two scan positions must have a certain level of overlap.

The concept of ptychography was proposed by Walter Hopper [51] to solving the phase problem in crystallography. The name 'ptycho' comes from the Greek language and means 'a fold'. This name is used because the diffraction pattern can be recognised as the convoluted results of the FT of the probe and the FT of the specimen when illuminated, which is like a folding process. Hoppe failed to extend the use of ptychography to non-crystalline specimens and ceased his research into ptychography [52]. In 1989, Rodenburg [53] compared many phase retrieval methods and realised that the geometry in Figure 12 can provide a solution linearly. In 1992, Bates and Rodenburg [6] published the Wigner distribution deconvolution method (WDDM), which was the start of modern ptychography, and the WDDM proved itself in the optical [54], x-ray [55] and electron [56] fields. The most remarkable advantage of the WDDM is that it solves the phase problem linearly, but it requires a very dense scan to provide a very large 4D dataset (for example, when used for a 256x256x256x256 byte dataset), which is a challenge for computers. Further, noise may negatively affect the deconvolution process. Thus, the development of the WDDM method has gradually been surpassed by iterative ptychography. Recently, however, due to an increase in the calculation abilities of computers, the WDDM has gradually returned to the spotlight [57,58,59].

### 3.3.1. *Iterative ptychography*

Later, Rodenburg and Faulkner proposed an iterative algorithm for ptychography to solve the phase problem – the PIE [14,60]. The success of iterative ptychography has been demonstrated in optical [61,62], x-ray [63,64] and electron wavelengths [65,66]. The PIE algorithm uses the information from an HIO algorithm [13] as there are two constraints in each iteration: the real space support and the modulus constraint. According to the Fourier shift theorem, if the phase of a pixel is laterally shifted in real space, the phase offset will occur in reciprocal space to affect the phase retrieval. Therefore, the real space support is used to block everything outside of the support. The modulus constraint described as the calculated intensity must equal the measured intensity for each scan position. An example of a PIE algorithm is demonstrated in

Figure 13, where, if the probe function is  $P$ , the object guess function is  $O$ , the exit wave is  $\psi$ , the intensity in the Fourier transform is  $\Psi$ , the real and reciprocal space coordinates are  $u$  and  $v$ , respectively, the related scan position is  $r$ , the record intensity is  $I$  and the iteration number is  $n$ .

- In the  $n$ th iteration, the probe illuminates the specimen guess at position  $r$ , and the exit wave is given as

$$\psi(u) = P(u - r)O_n(u). \quad (44)$$

- Then, the exit wave propagates the detector, and the propagation process is taken by a forward Fourier transform:

$$\Psi(v) = f(\psi(u)). \quad (45)$$

- In the reciprocal space, the modulus constraint is applied to replace the real part of the Fourier transform of the previous exit wave by the square root of the collected real intensity. However, the phase component is left to update the previous exit wave:

$$\Psi'(v) = \sqrt{I(v)} \frac{\Psi(v)}{|\Psi(v)|}. \quad (46)$$

- The updated exit wave is then propagated back to the object plane using an inverse Fourier transform:

$$\psi'(u) = f^{-1}(\Psi'(v)). \quad (47)$$

- The object function is updated by

$$O_{n+1}(u) = O_n(u) + \alpha \frac{|P(u - r)|}{\max|P(u - r)|} \frac{P^*(u - r)}{(|P(u - r)|^2 + c)} [\psi'(u) - \psi(u)], \quad (48)$$

where  $*$  means the complex conjugate,  $c$  is a very small number to avoid zero and  $\alpha$  is a feedback controller used for the update function, the value of  $\alpha$  being between 0 and 1.

- The illumination wave then moves on to the next position, where the new position overlaps with the previous.
- The process is repeated several hundred times, and the calculation error after each iteration is measured by the sum square error (SSE), calculated by

$$SSE = \frac{(I(v) - |\psi(v)|^2)}{N}, \quad (49)$$

in which  $N$  is the number of pixels in the wave function. To cease updating, the SSE value must be lower than a predetermined minimum value, which is usually defined by the level of noise in the experiment.

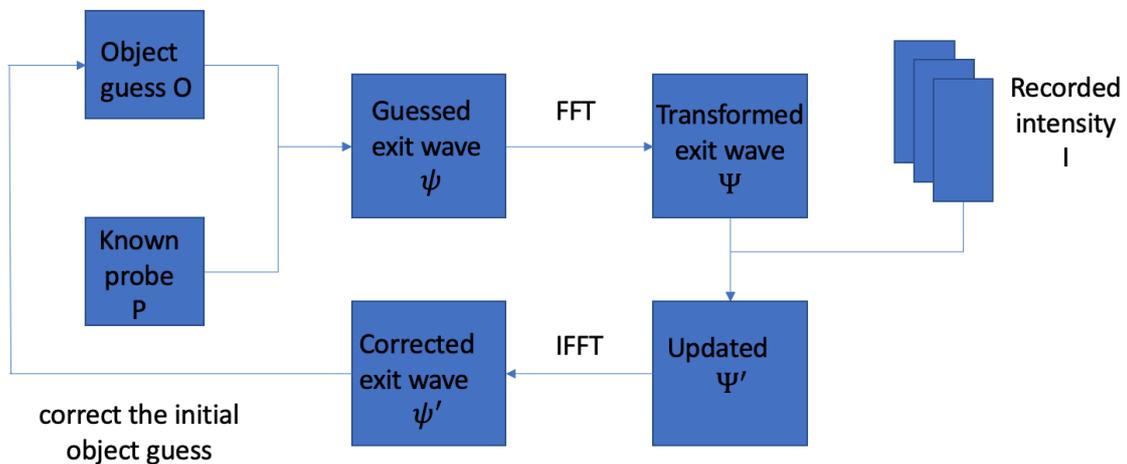


Figure 13. A flow chart of the PIE algorithm. At the real space, a guessed object function will be provided to the algorithm, where the real space constraint is applied. After the guesSED exit wave has been propagated to the reciprocal space, the Fourier space constraint will be applied and will backpropagated to the updated exit wave, to real space, to update the guesSED object function.

The PIE algorithm is much easier to calculate in a computer than the WDDM method (at least when it was invented), so it gave a greater boost to the development of

ptychography. However, as the above shows, to use the PIE algorithm, the exact probe function must be known, which limits the work of the PIE. Guizar-Sicairos and Fienup solved the probe function by using the conjugate gradient method [67]. In 2008, Thibault et al. [68] used a difference-map algorithm [69,70] to solve the probe function. In 2009, Maiden and Rodenburg [15] published the ePIE to solve the probe function, based on the PIE method. The ePIE algorithm is mainly used in this work. It is an important milestone for PIE-related algorithms as it provides a probe-updated function to estimate the accurate probe function from a rough guess. Also, ePIE has high speed convergence and good noise tolerance. The process of the ePIE is quite like the PIE algorithm – the only differences being in the initial guessing element and the updating aspect. The two updating functions are the probe-update function and the object-update function:

$$P_{n+1} = P_n + \alpha \frac{O_n^*(u - r)}{\max|O_n(u - r)|^2} [\psi'(u) - \psi(u)] \quad (50)$$

and

$$O_{n+1} = O_n + \beta \frac{P_{n+1}^*(u - r)}{\max|P_n(u - r)|^2} [\psi'(u) - \psi(u)], \quad (51)$$

where  $\alpha$  and  $\beta$  are the feedback controllers of the updated process. This algorithm is quickly accepted in visible light [71,72], x-ray [73] and electron experiments [74,75].

As is described above, PIE-related algorithms are serial algorithms, which means that the algorithms update each exit wave serially while updating the object and probe guess. Thus, this kind of algorithm creates large updates through small changes. However, there are similar algorithms that correct the object and probe guess through updating all exit waves in parallel simultaneously. There are two important parallel algorithms: the conjugate gradient method [67] and the difference-map algorithm [69,70]. The conjugate gradient method was the first algorithm to provide probe function updating, and the difference-map algorithm is one of the most popular algorithms in the x-ray ptychography community. In addition to the above, other popular algorithms in ptychography include the Proximal Block Implicit–Explicit

(PHEBIE) algorithm [76], hybrid projection reflection (HPR) method and Relaxed averaged alternating reflections method (RAAR) [77].

### 3.3.2. Sampling in ptychography

Sampling is a vital constraint in CDI methods, where the intensity of an object or the illuminated area of an object on the detector must be sampled by a suitable pixel matrix pitch, which records enough information to efficiently recover the object function. In other words, the detector size decides the real space object or illumination size. In the conventional CDI method, the sampling condition is controlled by the pixel pitch on the detector [78,79,80], in which if the radiation wavelength is  $\lambda$ , the distance between object and detector is  $d$ , the real space specimen size is  $X$  and the detector pixel pitch is  $\Delta u$  because the physical size of a detector pixel pitch has an angular relationship with the specimen, where for a small diffraction angle,

$$\Delta\theta = \frac{\lambda}{X} = \frac{\Delta u}{d}. \quad (52)$$

According to this relationship, we can conclude that to meet the Nyquist frequency on the detector plane, the object or illumination size must be smaller than

$$X \leq \frac{\lambda d}{\Delta u}. \quad (53)$$

However, the wave function is a complex value, but the detector can only record the intensity, so this sampling condition does not supply enough information to solve for the (complex) numbers required. Later, Sayre pointed out that the Fourier transform of the intensity is the autocorrelation of the process. The autocorrelation results are twice the size of the specimen, so the equation (53) should be changed to

$$X \leq \frac{\lambda d}{2\Delta u}. \quad (54)$$

To explore the sampling requirement, we must consider the special overlapping scan process of ptychography, which contrasts other CDI methods. Bunk et al. [81] proved that the more overlap there is, the better the reconstruction quality in ptychography experiments. This is with the exception of a 100% overlap, for this becomes more like a repeating single shot CDI process than ptychography. From the perspective of the richness of data, the larger overlap, in the ptychography there is more information about the overlapping area. The huge volume of information in ptychography data not only improves reconstruction quality but also provides strong noise tolerance and error correction abilities. Also, we see the scan process as a real space sampling process. When there is a greater overlap on the object plane, there are fewer intensity changes on the detector, which reduces the sampling requirement on the detector plane. Edo et al. [82] used experiment results to demonstrate that

$$\Delta x = \frac{1}{2\Delta u}, \quad (55)$$

which does not depend on the probe size, and where  $\Delta x$  is the pixel pitch on the specimen. Because only intensity is collected, a factor is required been added to recover the complex (real and imaging) parts of the wave. Similar results were also demonstrated by Silva and Menzel [83].

This is the sampling condition of the diffraction intensity. However, when iterative ptychography collects diffraction patterns, it scans the specimen cross-over with a certain degree of overlap, and the scanning process is also a form of the real space sampling process in ptychography. Therefore, to achieve an accurate measurement of the sampling condition in ptychography, it is better to combine real space and reciprocal space sampling, for

$$S_{pty} = \frac{1}{2\Delta s\Delta u}, \quad (56)$$

where the  $\Delta s$  is the step size on the specimen.

### 3.3.3. Super resolution in ptychography

At the Nyquist sampling condition, ptychography is highly over redundant, which means that in ptychography, there is much information redundancy. This is a big treasure for the ptychography to improve its performance. Edo et al. [84] proposed an experimental change to improve the resolution of ptychography, using a curved light source to replace the parallel light, which is quite similar to the tilt beam method proposed by Kirkland et al. [85]. In a parallel illumination, one detector pixel will record only one of the scattering vectors. Therefore, the detector's size will limit the resolution of ptychography. However, if the light is curved, there is no longer a one-to-one relationship between the detector's pixels and the scattering vectors, and the special frequency information is spread into several detector pixels, which means that a detector of the same size can record more than a parallel wave. Based on the above research, Edo proposed a method to achieve a higher resolution using PIE, where each of the diffraction patterns are attached on a larger all-zero values matrix. Then, the new diffraction patterns are used in the PIE algorithm to recover the missing special frequency information caused by the finite size of the detector, which is due to the missing spatial frequency information of one diffraction pattern has been stored in the data and spread in several overlapped scan position diffraction patterns.

Later, Maiden et al. [86] published the ePIE-based super-resolution algorithm. Essentially, the idea of this was to computationally enlarge the size of diffraction patterns. This was named the super-resolution ptychographical iterative engine (SRPIE). Two procedures in the SRPIE deserve attention. First, when the diffraction is enlarged several times, the probe function also resizes to the same size, so the probe's position may not be an integer value and there may need to be a subpixel shift, which is achieved in the SRPIE by adding a linear phase ramp on the Fourier transform of the probe function. When the probe function is updated, this approach is necessary to shift the probe back. The second procedure is through adding a constraint on the exit wave at the detector plane and applying a Fourier constraint:

$$\Psi'(v) = \begin{cases} \sqrt{I(v)}\Phi(\Psi(v)), & \text{if } u \in s \\ \Psi(v), & \text{if } u \notin s \end{cases}, \quad (57)$$

where  $s$  represents the area containing the intensity information. This equation means the valid area will be updated as usual, but the constraints area does not change, which can help to avoid the Fourier repeating and reducing high frequency noise.

### 3.3.4. Upsampling in ptychography

In ptychography, as described above, the probe size is unrelated with the sampling condition, but it limits the sampling condition in reciprocal space. In some situations, the probe's structure in the reciprocal space is smaller than the size of a detector pixel, which could lead to the information on the detector being confused, for detector-recorded information is the integrated result of different structural components. This may be troublesome for conventional CDI methods, but due to the overlapping scanning method, each detector pixel is employed several times, so there is abundant redundant information in the data, which may be useful to refine the detector pixels' computationally. Upsampling is used to improve the sampling conditions of such a coarse detector, proposed by Batey et al. [87]. The basic idea of upsampling is quite easy to understand: A large pixel is separated into several smaller pixels. This is shown in Figure 14, where a big pixel, A, has been equally separated to four smaller pixels, and each small pixel is isolated from the others.

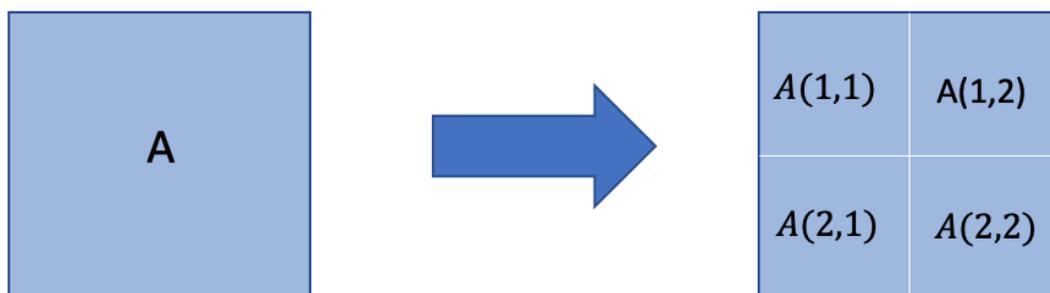


Figure 14. A schematic diagram of pixel separation. A big pixel is separated equally into four smaller pixels. The total value of the small pixels is equal to that of the big pixel.

The number of separated pixels depends on the over sampling condition in the ptychography data, which can be estimated by using equations (52) and (54):

$$\eta = \frac{\lambda}{2\Delta\theta d}. \quad (58)$$

Thus, if  $\eta = 2$ , we can separate a big pixel into  $2 \times 2$  small pixels. The reconstruction algorithm of upsampling is based on ePIE, where we must assume that each detector is divided into  $\eta \times \eta$  small pixels. Therefore, at the detector plane, we must update each of the small pixels of the exit wave by recording the intensity as

$$\Psi'_{(x,y)}(v) = \frac{\Psi_{(x,y)}(v)\sqrt{I(v)}}{\sqrt{\sum_x \sum_y |\Psi_{(x,y)}(v)|^2}} \quad (59)$$

where  $x$  and  $y$  are the coordinates of the order of small pixels. The upsampling method is a solution to some poor detector ptychography experiments that cannot meet the minimum requirements of the Nyquist sampling conditions. However, implementing it for general ptychographic reconstruction should be avoided, for it artificially separates a large pixel into small pixels; the performance of upsampling is strongly affected by real detector data quality.

### 3.3.5. Scan position correction

The first factor that limits the performance of ptychography is an error with the scan position. In this ptychography experiment, the only data collected were the diffraction patterns. Each diffraction pattern uniquely corresponds to the real space's scan position, so the accuracy of the specimen's position limits the quality of reconstruction. This has also been revealed by another research [88,89,90]. However, several algorithms have been used to solve this problem, such as the conjugate gradient method [91] and cross correlation-based methods [92].

We will now introduce an ePIE-based position correction method, which was created by Maiden et al. [93], called 'Jiggle'. To begin the Jiggle method, we generate several random positions surrounding one of the original positions with a certain radius. These are random positions, and after each position correction iteration, the corrected

position will move closer to the real position, so we also need to reduce the annealing position area radius. We then propagate the exit waves of the original position and the surrounding annealing positions to the detector plane to calculate the intensities. Once we have obtained all the guessed intensities, we compare the recorded intensity at this position with the calculated intensities to find the closest result. Next, we apply the Fourier constraint to the most comparable exit wave at the detector plane, and backpropagate to the real space to update the object and probe guess. The same procedure will be performed for all diffraction patterns to address the real space's scan position at each iteration.

### 3.3.6. Error metric for ptychography

To analyse quantitatively the convergence of a ptychographic algorithm, an error measurement is necessary. This is because the only data measured in a ptychography experiment is the intensity of the diffracted exit waves from the different scan positions of the specimen. Hence, the most straightforward way to measure error is to calculate the recorded intensities with the calculated exit wave on the detector plane, such as the R-factor, which is used in crystallography to measure the reconstruction quality,

$$R = \frac{\sum |F_{obs} - F_{cal}|}{\sum |F_{obs}|}, \quad (60)$$

in which  $F_{obs}$  is the amplitude of measure data and  $F_{cal}$  is the amplitude of the calculated data.

The error matrix used in this work was refined by Maiden et al. [15] from the root mean square error (RMES), which calculates the normalised error between measurement and calculation after each iteration:

$$\varepsilon = \sqrt{\frac{\sum_j \sum_v \left| |\Psi_j(v)|^2 - I_j(v) \right|}{\sum_j \sum_v I(v)^2}}, \quad (61)$$

where  $j$  denotes the scan position. However, as a normalised, calculated-based error metric, it can give us a good indication regarding accuracy. However, some influential factors must be considered, such as the noise and local minima. If there are several better solutions, the algorithm may choose one that is not the best and have no to chance to jump out. Therefore, it is better to combine this information with other measurements to analyse the reconstruction.

Another widely employed error metric is the Fourier ring correlation (FRC) [94,95], which is an amended 2D version of the Fourier shell correlation (FSC) [96,97]. The basic idea behind the FRC and FSC is to calculate the normalised cross correlation between two objects that are on the same target structure volume of Fourier space by separating the spatial frequency content of them into several calculation-concentric rings or shell regions. The cross correlation at each calculation ring is given by

$$FRC(r_m) = \frac{\sum_{r \in r_m} F_{ref}(r) F_{cal}^*(r)}{\sqrt{\sum_{r \in r_m} |F_{ref}(r)|^2 \sum_{r \in r_m} |F_{cal}(r)|^2}}, \quad (62)$$

in which  $m$  is the order of the ring,  $r$  is the special frequency at this ring zone,  $F_{ref}$  is the first object (normally the reference object) and  $F_{cal}$  is the second object (the reconstruction in this work). The FRC is quite a good error measurement and has some advantages. First, it can measure the performance of reconstructions quantitatively. Secondly, its measurement is made at the Fourier space, which can help to avoid subjective bias to the objects as it depends on the structure of specimen as well as the illumination. Third, it can measure the performance of a reconstruction without knowing the noise criterion in advance.

### 3.3.7. Fourier ptychography

An important derivative technique is Fourier ptychography, where the real space overlapping scan process is replaced by a spectrum shift in Fourier space [98]. According to the shift theorem, the spectrum shift in Fourier space can be performed by tilting the illumination angle toward the real space. Therefore, Fourier ptychography requires a series of tilted lights to illuminate the specimen. The setup of a Fourier

ptychography experiment is shown in Figure 15, where the diffracted exit waves are collected by an objective lens and focused onto the rear focal plane, where an aperture is placed to filter the spectrum of the exit waves. A detector at the far end of the aperture records the object's image.

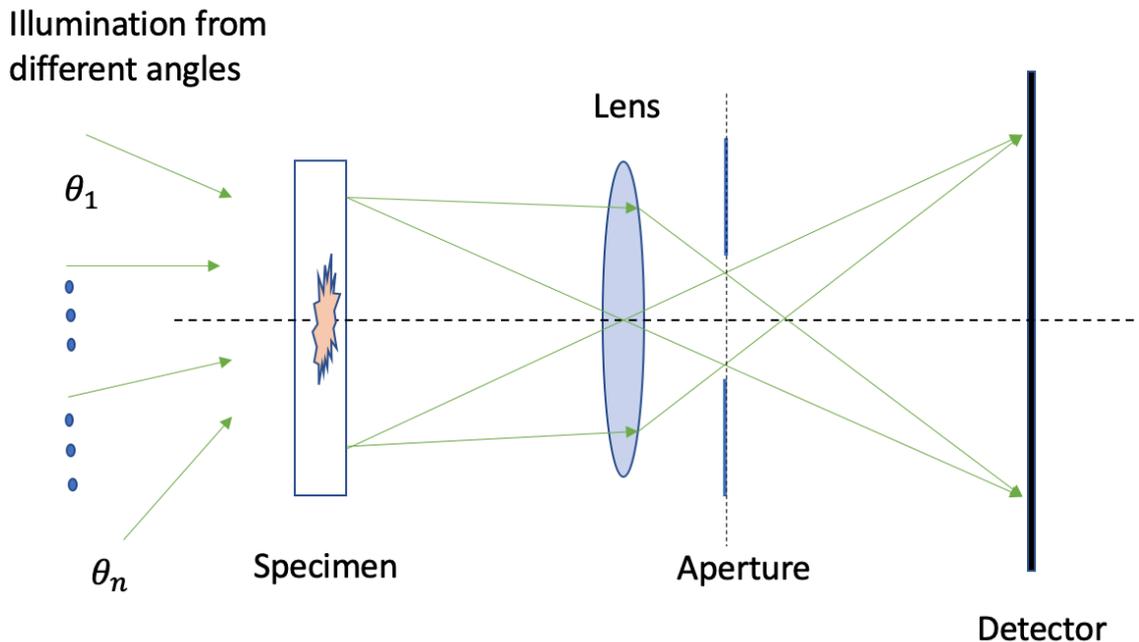


Figure 15. A sample setup of a Fourier ptychography experiment. As the tilted illumination in real space is equal to the spectrum shift in Fourier space, when the light is focused on the specimen from different angles, the diffractions are collected by an imaging lens and sorted by an aperture at the rear focal plane of the lens to filter the spectrum of diffractions. A detector is fixed on the image plane to record low-resolution images.

Following the illumination from each tilted light, a low-resolution image is recorded on the image detector. This low resolution is caused by the aperture on the back focal plane, through which only a small portion of the specimen's spectrum can pass. Therefore, the resolution of Fourier ptychography is limited by the tilting angle of the light. A higher tilting angle collects higher frequency components from the spectrum of the specimen, but where counts might be low. Low angle collects components of a lower frequency, which may reduce the signal-to-noise ratio. To easily measure the resolution of a Fourier ptychography experiment, Konda et al. [99] provided this equation:

$$res = \frac{\lambda}{NA_L + NA_I}, \quad (63)$$

in which  $NA_L$  is the numerical aperture of the lens and  $NA_I$  is the highest tilted angle illumination numerical aperture and which also proves that, with a fixed illumination wavelength and lens, enlarging the tilting angle of the light helps to improve the resolution in Fourier ptychography.

Fourier ptychography has some advantages. First, it can produce a high-resolution, complex image at gigapixel levels. Second, fewer requirements are placed on the detector as it is used to collect the specimen image rather than the diffraction, of which the latter has larger dynamic range. Third, as there are no scanning processes on a specimen in Fourier ptychography, it can produce a specimen reconstruction much quicker; the scanning process is replaced by the spectrum shift created by switching on an off an array of sources. Nowadays, Fourier ptychography progress has been made in many ways, such as through fast Fourier ptychography [100,101,102] and 3D Fourier ptychography [103, 104].

### 3.4. Conclusion

In this chapter, based on the phase problem, we have reviewed the basic models of CDI and some decoding processes aimed at solving the phase problem, including the GS algorithm and the HIO algorithm. Next, we discussed ptychography, which is the core technique in this thesis; all the works are based on it. To begin, we introduced the special overlapping data collection method and a brief history of ptychography.

The iterative ptychography method was then demonstrated – the main technique used to solve the phase problem – based on ptychography-based research. At this point, the PIE and ePIE algorithms were introduced and were used in this work. Following the decoding algorithm, we reviewed the sampling conditions in conventional CDI methods and ptychography, which showed that ptychography data has a much higher sampling condition than single shot CDI methods. We introduced the super resolution, upsampling and scan position corrections techniques, which I have studied in my past research experience, but which has not contributed to this thesis. They are the best

examples of the potential of ptychography due to their high information redundancy features. Finally, we gave a brief introduction of Fourier ptychography, which is an important derivative technique and a key research topic for ptychography, but I have not done any research about it.

# Chapter 4 Modal decomposition of partially coherent wave field via ptychography

As a type of CDI method, partially coherence is critically limiting reconstruction quality. In this chapter, we use two classical light experiments to gain a general understanding of temporal and spatial coherence. As recent research suggests, we can separate multiple object and illumination states at the same time in spatially mixed-state ptychography, where the mixed states come from spatial instability, and they are solved by the information redundancy of the data. The spatially mixed states include the probe and object states, and the overall mixed state is the mixed interaction result of each probe state with each object state. Therefore, if we treat the partial coherence as a mixture of many coherent illumination states, the spatially mixed-state ptychography can properly deal with it. In this chapter, we theoretically and experimentally demonstrate how the algorithm works. Further, we explore the patterns of the source illumination mode and the influencing factors in the modal decomposition and reconstruction using engineered point sources.

## 4.1. Coherence

In optics, coherence describes the characteristic that any two points in the wave front have a defined phase change at any time. This is very important to the CDI method because a good coherent beam can provide high contrast and a stable interference pattern on a detector. Normally, coherence can be measured by interferometers as the intensity form [105]

$$I(\tau) = I(p_1) + I(p_2) + 2\Gamma_{12}(\tau), \quad (64)$$

where  $I(p)$  is the intensity at point  $p$ ,  $\tau$  is the time advance between two points and  $\Gamma_{12}(\tau)$  is the mutual coherence function. The mutual coherence function was proposed by Wolf [106] to give a general description of the phase correlation of two random points on the illumination wave, along and transverse to the propagation, as

$$\Gamma_{12}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T U(p_1, t + \tau) U^*(p_2, t) dt, \quad (65)$$

in which  $U(p, t)$  is a time-independent complex function of a random point,  $p$ , at a fixed time  $t$ , and  $T$  is the time. The degree of coherence can be calculated as

$$C = \frac{\Gamma_{12}(\tau)}{\sqrt{I(p_1, t) I(p_2, t)}}. \quad (66)$$

The value of  $C$  is between 0 and 1 to represent incoherence, partial coherence, and coherence.

#### 4.1.1. Temporal coherence

Coherence can be divided into temporal coherence and spatial coherence. Temporal coherence measures the phase correlation between two points in the direction of propagation. According to Saleh and Teich [107], the temporal coherence function can be written as

$$f(\tau) = \int U(t) U^*(t + \tau) dt, \quad (67)$$

and the temporal coherence degree is calculated by

$$C_t = \frac{f(\tau)}{f(0)}, \quad (68)$$

where  $f(0)$  represents the fully coherent temporal coherence function. This is a key definition in the length of temporal coherence, which describes that the illumination wave can be recognised as coherent during a coherence time. The length of coherence of light can easily be measured using the Michelson–Morley experiment [108,109], which is helpful for understanding temporal coherence. However, it cannot be used with electrons or x-rays. A simplified module is shown in Figure 16, in which a monochromatic illumination is emitted from the light source and propagated to the half silver mirror,  $m_0$ . At  $m_0$ , half of the illumination passes through  $m_0$  and is propagated to reflection mirror,  $m_1$ . The other half of the illumination is reflected to the second reflection mirror,  $m_2$ . The illumination reflected by mirror  $m_1$  back to  $m_0$  is reflected to a detector at the bottom, while the illumination reflected by  $m_2$  will pass through  $m_0$  directly to the detector.

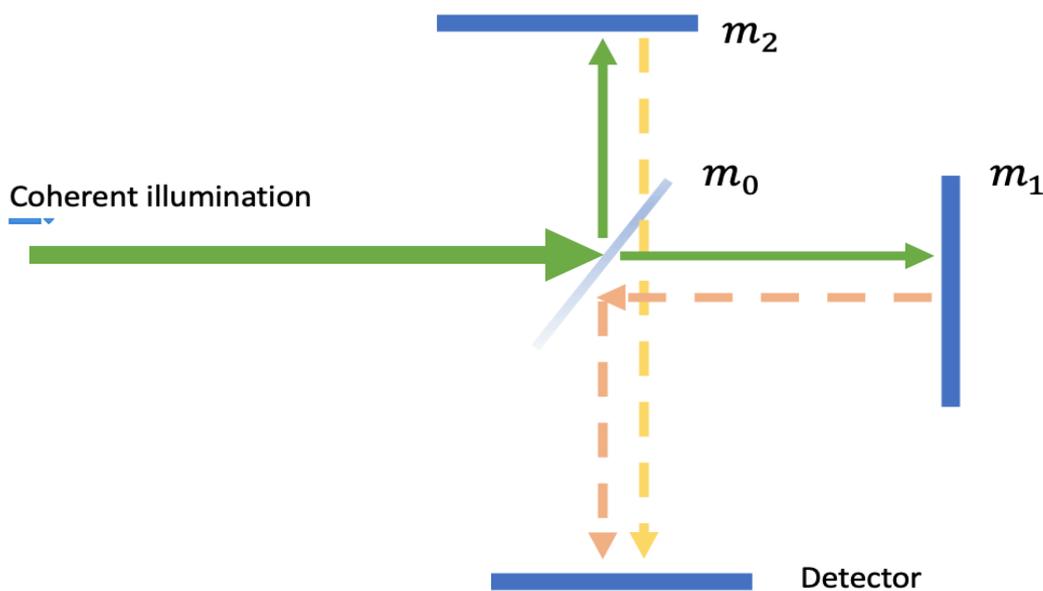


Figure 16. Schematic of the Michelson-Morley experiment. In this diagram,  $m_0$  is a half silver mirror that can split a light beam into two. The reflect mirrors,  $m_1$  and  $m_2$ , have different path lengths to  $m_0$ . The detector at the bottom collects the interfere pattern of the two reflected beams.

The two illuminations will interfere with each other at the detector to form a diffraction pattern. To generate a good coherent diffraction, the path difference between  $m_0$  to  $m_1$  and  $m_0$  to  $m_2$  must have a relationship in the form

$$L_c = 2d = ct_c, \quad (69)$$

where  $d$  is the path difference from  $m_0$  to  $m_1$  and  $m_2$ ,  $c$  is the light speed in a vacuum and  $t_c$  is the coherence time that is given by

$$t_c = \int |C_t|^2 d\tau. \quad (70)$$

The coherence time also has a relationship with the light illumination bandwidth defined by

$$t_c B \geq 1. \quad (71)$$

Therefore, the coherence length is approximately

$$L_c \geq \frac{c}{B}, \quad (72)$$

where  $B$  is the bandwidth of the light illumination.

#### 4.1.2. Spatial coherence

Spatial coherence is the phase difference between any two points on the wave front at a point in time. To describe the spatial coherence, Young's double slits experiment is widely employed example [110,111,112], as shown in Figure 17.

In Figure 17,  $x$  and  $y$  are coordinates; the left side of the mask shows a plane wave;  $P_1$  and  $P_2$  are two pin holes on a block, with distance of  $D$ ;  $F$  is the complex-valued wavefront incident at point  $y$  on the detector and  $L$  is the distance between the block and the detector. The path difference between the radiation passing through  $P_1$  and  $P_2$  to point  $y$  is

$$D_{diff}(y) = \left| \sqrt{L^2 + \left(\frac{D}{2} - y\right)^2} - \sqrt{L^2 + \left(\frac{D}{2} + y\right)^2} \right|, \quad (73)$$

and therefore, the phase difference is

$$\Phi_{diff} = \frac{2\pi}{\lambda} D_{diff}, \quad (74)$$

where  $\lambda$  is the wavelength of the radiation. If the phase difference is  $2\pi$ , the fringe period can be calculated as

$$y_{fringe} = \frac{\lambda L}{D}. \quad (75)$$

The intensity at point  $F(y)$  is

$$I_F(y) = |F(y)|^2 = \frac{I_0}{2} \left( 1 + \cos \left( \Phi_{diff}(y) \right) \right), \quad (76)$$

in which  $I_0$  is the recorded intensity of the radiation passing through only one pinhole. Michelson has proposed a function to measure the quality of the interfere pattern as visibility,  $V$ , which is represent as

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \quad (77)$$

where  $I_{max}$  and  $I_{min}$  are the recorded maximum and minimum intensity values.

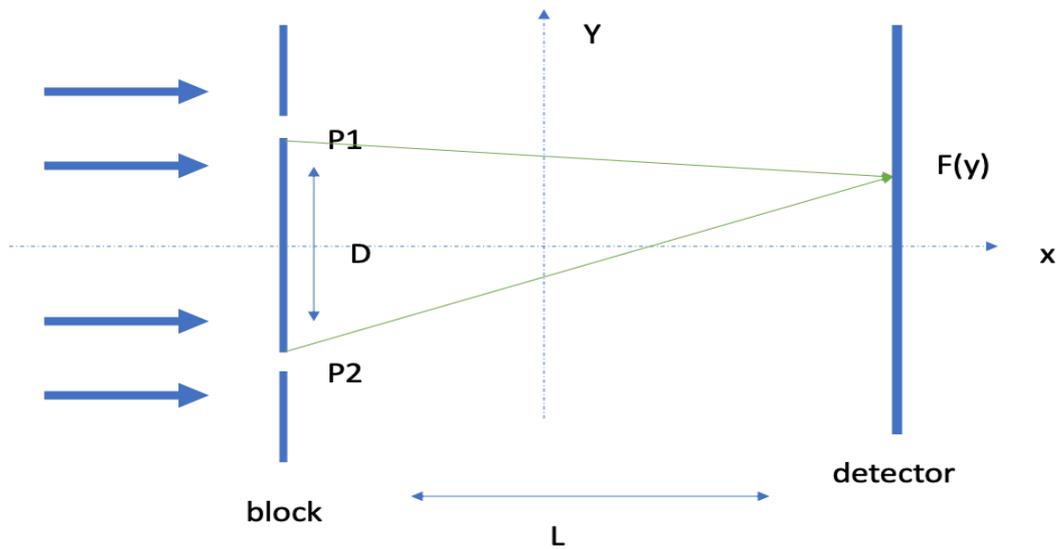


Figure 17. Schematic diagram of Young's double slits experiment. On the left-hand side is a light block with two pin holes,  $P_1$  and  $P_2$ . The distance between the two pin holes is  $D$ . On the right-hand side is a detector to record the interfere pattern of the beams emitted from the two holes. The distance between the light block and the detector is  $L$ .

#### 4.1.3. The Van Cittert–Zernike theorem

To measure the degree of coherence, the Van Cittert–Zernike theorem [113, 114, 115], which assumes that spatial coherence dominates the coherence properties of a wave (as in the synchrotron source in the next chapter), can be used to characterise the degree of coherence of an electromagnetic wave. Matter waves can also be calculated using this theorem. The Van Cittert–Zernike theorem describes that, in some situations, the complex visibility of a long-distance incoherent source is equal to the Fourier transformation of its intensity distribution as

$$\Gamma_{12}(u, v, 0) = \iint I(m, n) e^{-2\pi(mu+nv)} dmdn, \quad (78)$$

where  $u$  and  $v$  are the distance between the source and the observation plane on the  $x$  and  $y$  dimensions,  $m$  and  $n$  are the source direction cosines of a source point and  $I$  is the intensity of the source. The Van Cittert–Zernike theorem gives a Fourier transform relationship of the mutual coherence and the source intensity distribution.

## 4.2. Mixed-state ptychography

In an ideal ptychography experiment, the illumination is purely coherent. Therefore, there is only one probe state and specimen state in the experiment spatially, where we use the probe to replace the illumination, as the illumination will be focused on the specimen by the lenses. However, partially coherent illumination is employed in x-ray and electron experiments to balance the requirements of the degree of illumination coherence and flux. In partially coherent illumination, there is more than one probe state in the experiment, and each probe mode will interact with the specimen function to contribute to the recorded intensity. For the phase problem, the detector can only record an intensity, which is the sum of all the intensities caused by the multiple probe states. All the phase information is lost. To solve the modulus and phase information of the specimen and all probe states, the dataset must contain as much information redundancy as possible. Fortunately, one of the most important characteristics of ptychography, when compared with CDI methods, is that the dataset has rich information diversity. Mandel and Wolf [116] showed that any partially coherent wavefield can be divided into several totally incoherent wave components that are self-coherent but incoherent with each other.

More recently, Thibault and Menzel [7] proved that the modal decomposition method can be applied to ptychography because of the high information diversity in its dataset. Peng et al. [117] have further improved the theory to resolve the ambiguities in the original reconstruction process and experiment setup of multi-state ptychography experiments. In this chapter, the reconstruction algorithm will be demonstrated, first based on the ePIE algorithm, then some artificially built modes will be used to discuss the mechanism of and influence of parameters on modal decomposition experiments.

### 4.2.1 *ePIE-based modal decomposition algorithm*

The reconstruction algorithm of modal decomposition is similar to the ePIE algorithm. The only difference is that the measured intensity is contributed to by several partially coherent probe modes, rather than one probe mode. Therefore, it must be scaled to

different intensities because of multi-probe states. If the probe guess function is  $P$  with  $n$  states, the object guess function is  $O$  with  $m$  states, the exit wave is  $\psi$ , the measured intensity is  $\Psi$ , the real and reciprocal space coordinates are  $u$  and  $v$ , the related scan position is  $r$ , the record intensity is  $I$  and the iteration number is  $i$ , then the ePIE-based modal decomposition process is as shown in Figure 18, and

- In the  $i$ th iteration, the  $n$  probe states on the  $m$  specimen guess states at position  $r$  are such that the multi-state exit wave will be

$$\psi_i^{(n,m)}(u) = P_i^n(u-r)O_i^m(u). \quad (79)$$

- The multi-state exit wave will propagate to the detector as a forward Fourier transform in far field as

$$\Psi_i^{(n,m)}(v) = f\left(\psi_i^{(n,m)}(u)\right). \quad (80)$$

- In reciprocal space, the recorded intensity must be scaled to several small parts according to the scaled weight of each mode, as defined by Mandel and Wolf [116], to perform the Fourier constraint individually. The total calculated intensity must be equal to the measured intensity, as shown in Figure 19. The phase part of the mixed-state exit waves will keep updating the pervious exit waves as

$$\Psi_i'^{(n,m)}(v) = \sqrt{I(v) \frac{|\Psi_i^{(n,m)}(v)|^2}{\sum_n \sum_m |\Psi_i^{(n,m)}(v)|^2} \frac{\Psi_i^{(n,m)}(v)}{|\Psi_i^{(n,m)}(v)|}} \quad (81)$$

- Propagating the updated multi-state exit wave back to the object and probe state guesses as

$$\psi_i'^{(n,m)}(u) = f^{-1}\left(\Psi_i'^{(n,m)}(v)\right). \quad (82)$$

- Updating the probe and object state guesses by

$$P_{i+1}^n(u) = P_i^n(u) + \alpha \frac{\sum_m O_i^{m*}(u-r)}{\max[\sum_m |O_i^m(u-r)|^2]} [\psi_i'^{(n,m)}(u) - \psi_i^{(n,m)}(u)] \quad (83)$$

and

$$O_{i+1}^m(u) = O_i^m(u) + \beta \frac{\sum_n P_i^{n*}(u)}{\max[\sum_n |P_i^n(u)|^2]} [\psi_i'^{(n,m)}(u) - \psi_i^{(n,m)}(u)], \quad (84)$$

where \* means the complex conjugate,  $\alpha$  and  $\beta$  are the feedback controllers for the update functions and the values are between 0 and 1.

- Then move to the next joint position and repeat the process (79–84) several hundred times.

For the probe modes, we can perform an orthogonalizing process to get the unique probe modes, where the probe modes are unique and self-coherent but incoherent with others.

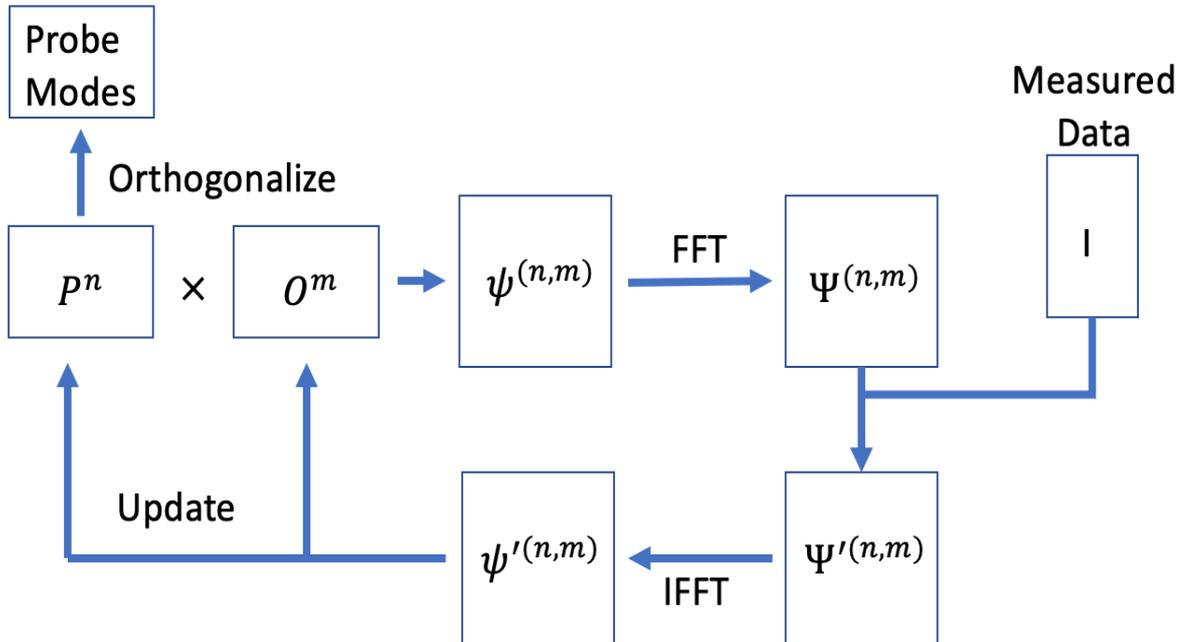


Figure 18. Flow chart of the ePIE-based modal decomposition process. In general, the modal decomposition process is a series of parallel ePIE processes at each scan position. However, there are two differences. First, at the reciprocal space, an intensity separation process must

be applied to each calculated intensity, as shown in Figure 19. Second, to get the isolated probe modes, an orthogonalizing process is employed to the calculated probe functions.

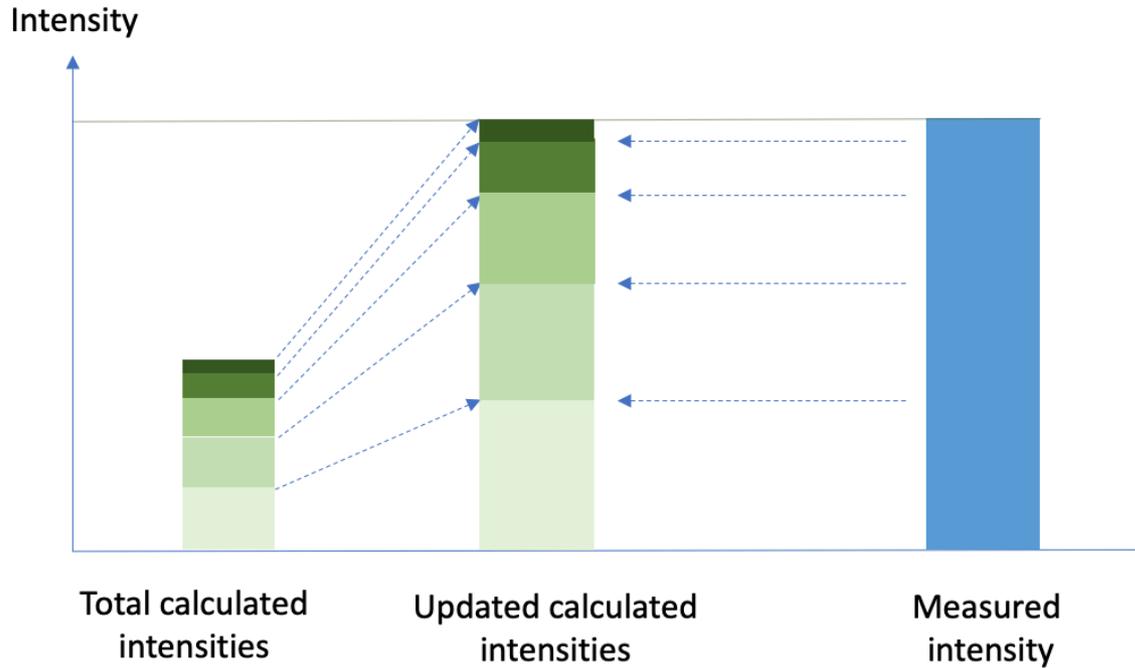


Figure 19. The intensity separation in the Fourier constraint process for each scan position. As the total calculated intensities must equal the measured intensity in each scan position, the measured intensity must be separated in each calculation, according to the weight of each calculated intensity in the total calculated intensity, to perform the Fourier constraint, where the phase in each calculation keeps the same.

### 4.3. Probe modes

Here, we separate a partially coherent illumination source into a finite number of point sources on the source plane and model the radiation. A monochromatic plane wave on the detector can be recognised as a plane wave that emits from a single point on the source plane. Consequently, the partially coherent wave is the result of the addition of the intensity of different monochromatic waves during propagation. Therefore, each mode shows a type of interference pattern on the detector. We select three, four and five joint points with the same value on the source plane as examples and transfer each point individually, then inversely transform them back. On the image plane, we use singular value decomposition (SVD) to orthogonalize the backward results to get the source modes, where SVD is a widely used method to perform Eigen decomposition to extract the eigenvalue and the eigenvector [118].

Figure 20 shows the point source and corresponding source mode patterns, where the points on the source plane display as different distribution patterns, and the source modes show the interference pattern of different waves after they pass through a circle mask. They are shown in descending order of power in each mode. The power of each mode gives the participation of the mode pattern in a partially coherent source, where the mode patterns and the power distribution can be affected by the source points' spatial distribution on the source plane and the intensity of each source point. Therefore, we can change the mode patterns and related power distribution by changing the spatial distribution or intensity of the source points. For example, in Figure 20 there are three types of source patterns and related illumination modes in each row. Comparing the first row and other point sources, the first-row source has no vertical source point, so the mode has not shown the influence of any horizontally interfere. Also, the first mode of each row shows more power than others because the source points are closer to each other. Therefore, the detector will record lots of similar information. In the second and third rows, the second and third modes look symmetrical with equal power, as point sources are symmetrical. Therefore, they will create the same strength interfere from all directions.

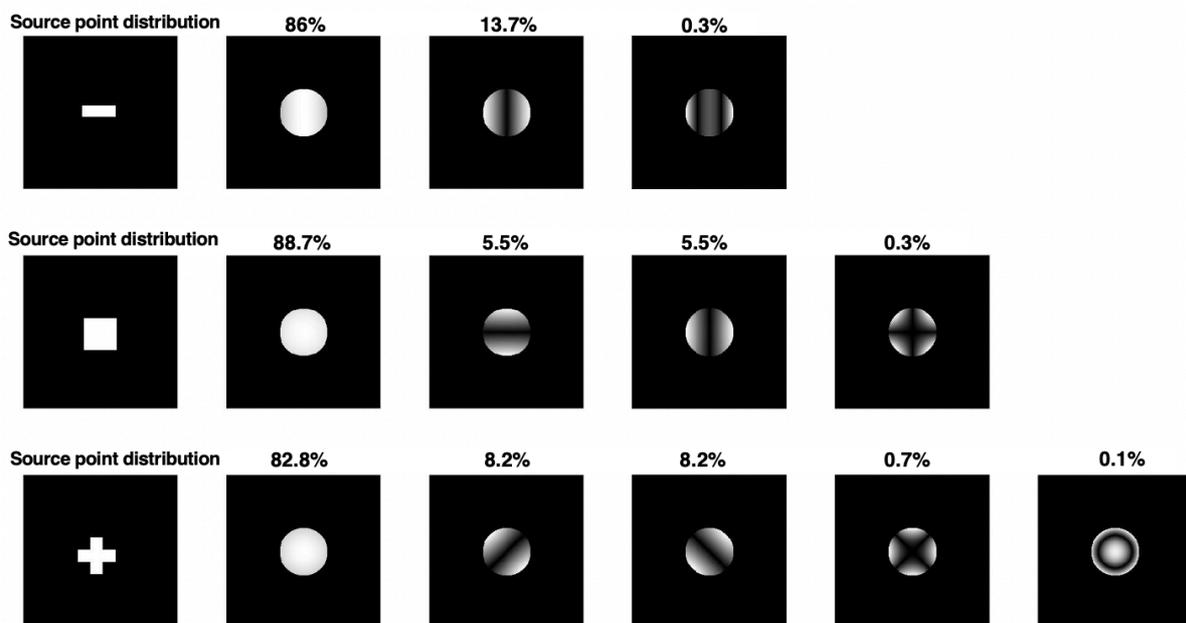


Figure 20. The point source patterns and corresponding source modes. From top to bottom they are Three-point source pattern and three mode patterns with power, four-point source

pattern and four mode patterns with power, and five-point source pattern and five mode patterns with power.

Unlike with source modes, by using ptychography reconstruction, we can put a random number of probe functions into the reconstruction to solve the dataset. However, there are only a few modes that strongly contribute to the reconstruction process, while other modes are empty modes, which may record artificial noise information in the dataset.

Here, the five-point source (the third row in Figure 20) is used as above. We employ the ePIE algorithm with eight modes at random to show the modal decomposition results. after five hundred iterations. The reconstruction results are shown in Figure 21 (see next page), in which the first image shows the reconstruction result of Lena, which has been widely used in the digital image processing community. In these results, the power of the last three probe modes is significantly smaller than the first five, as they do not have a clear interfere pattern because they are pure artefacts and do not exist in the original data. This means that they contribute no information to the object reconstruction, which is why they are called the empty modes.

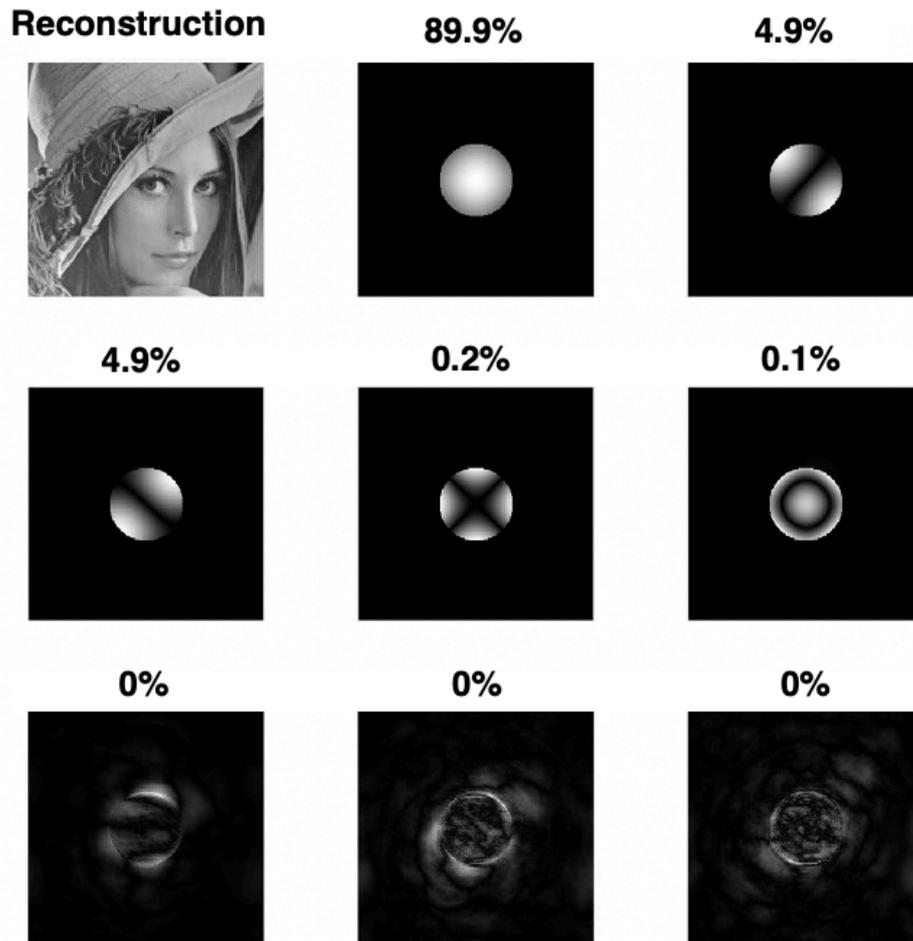


Figure 21. The object reconstruction and eight orthogonalized probe modes. Modes one to five are almost the same as the source mode, and modes six to eight are the empty modes, which contributed nothing to the reconstruction.

As introduced above, power is a measure of the importance of the mode's pattern to the reconstruction process. From Figure 21, we learned that although we know there should be five mode patterns, more than 99% of the power is concentrated in the first three patterns. From this conclusion, we can infer that we may not need to calculate all the modes to achieve a good reconstruction result. To prove this idea, we performed another simulation, in which we engineered a new source mode pattern, as shown in Figure 22, in which there are no symmetrical mode patterns with the same power. In this test, we only tried to extract two to five modes each time and compare the reconstruction image quality using Fourier ring correlation (FRC).

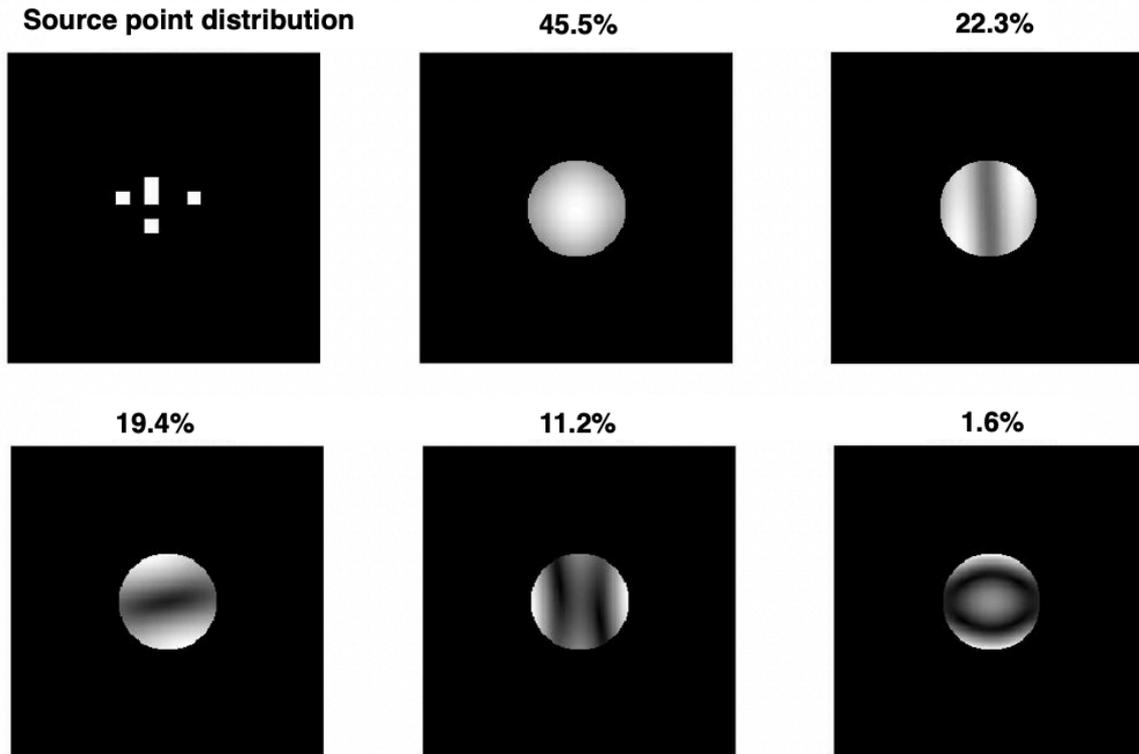


Figure 22. Five-point source and mode patterns. By changing the spatial distribution of the source points, we can avoid any symmetrical mode patterns with the same power.

The results are shown in Figures 23 and 24. Figure 23 shows the reconstructions from two to five mode calculations, but only the centre part of each result, as this area has been scanned more times than the boundary part, which helps to record more information in the data. As a result, any artificial effect will be spread over more diffraction patterns giving the best possible results in the reconstructions. From the results, it can be intuitively seen that the last two results are better than the first two, and that by inputting more modes into the calculation, the reconstruction improves.

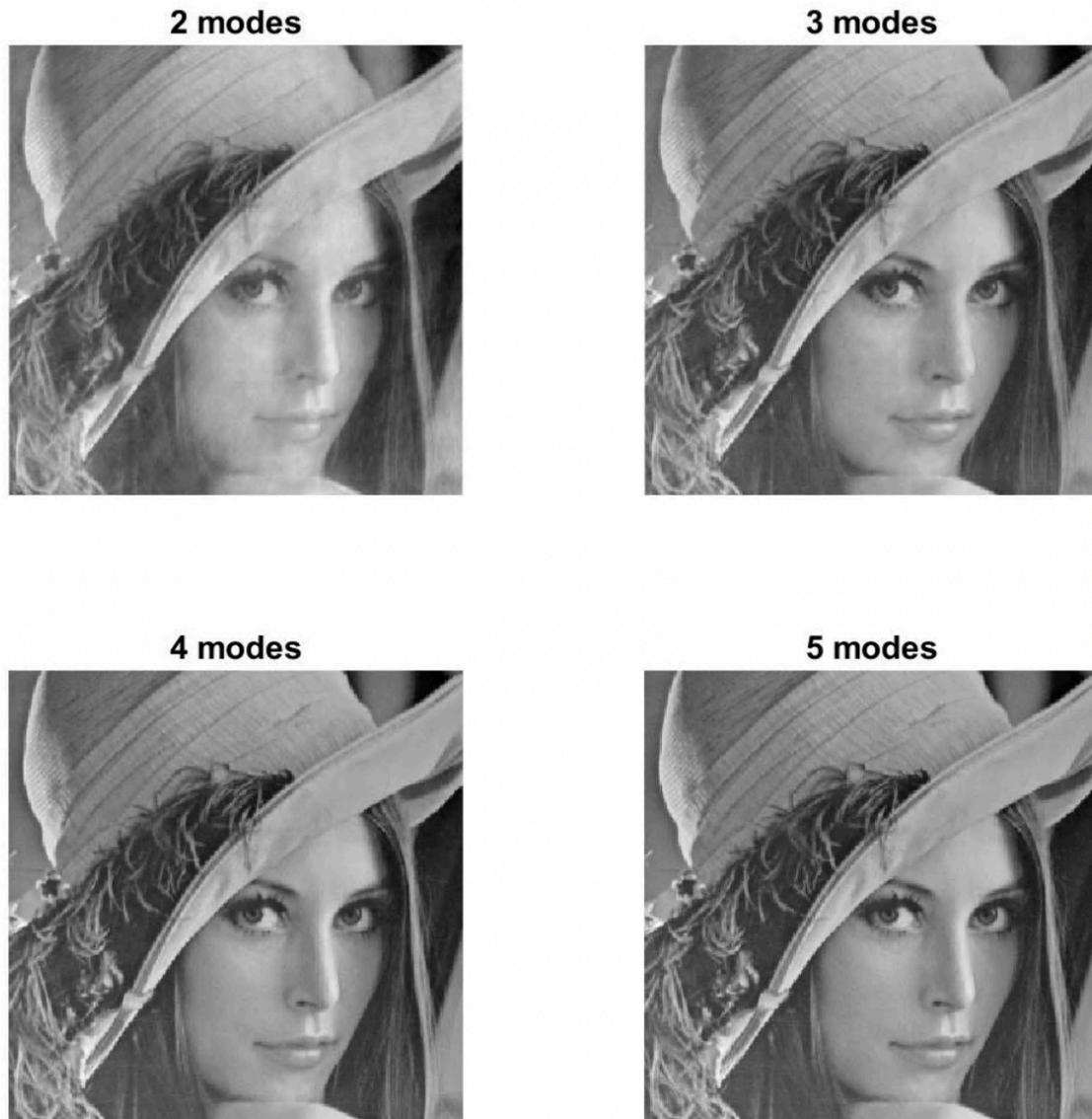


Figure 23. The image reconstruction results with different modes in the calculation. This clearly shows that with more modes in the calculation, the reconstruction is improved.

Figure 24 displays the FRC results of each reconstruction result, with the object image. From Figure 24, the reconstruction quality using two modes is significantly lower than the others. The reason for this is that using only two modes cannot fully overcome partially coherence in all directions. However, when we use three modes, the three-mode patterns have contributed more than 86% to the reconstruction. In other words, if we can calculate three modes, we can reduce the negative effect of partially coherence of the reconstruction by more than 86%. With more modes in the construction, the influence of partially coherence can be further reduced. However, as

the last two modes have less influence, they cannot improve the reconstruction to a great degree.

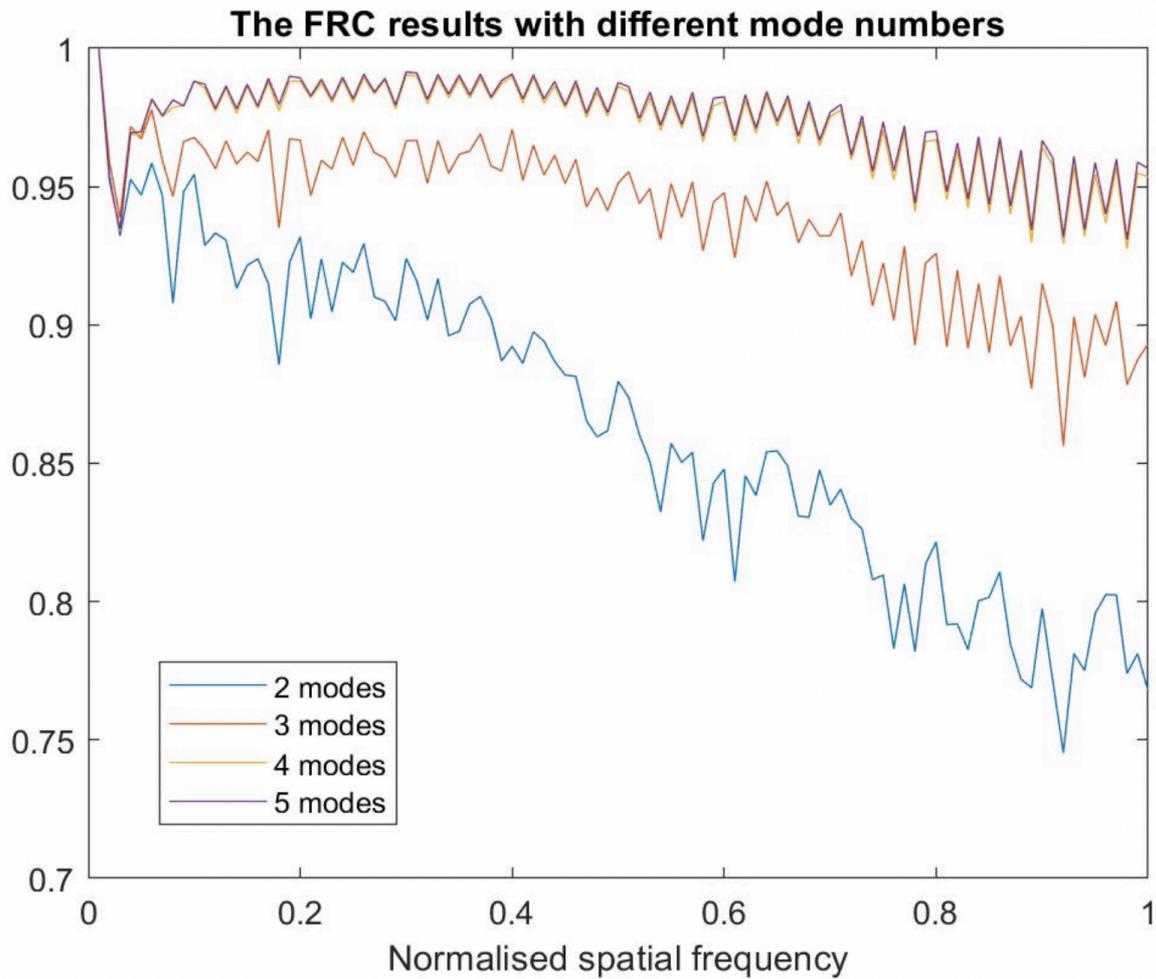


Figure 24. The Fourier ring correlation result. The result compares the centre part of the reconstruction result with the specimen image using different numbers of modes. From the results, using more modes in the calculation results in a better reconstruction and that the improvement depends on the power of mode, where more power results in a greater improvement. The negative peak common to all reconstructions is because we use only part of the whole image to check the autocorrelation results, where we break the smooth distribution of the frequency component.

#### 4.3.1. Probe mode power separation in the reconstruction

From the above results, if we put more probe functions into the reconstruction, we can achieve a better a result. But what is hidden in the modal decomposition? To answer this question, a tracking experiment was done to plot the mode power distribution during the modal decomposition process at each iteration. A new five-point source

was used to give equal mode powers, and the modal decomposition results are shown in Figure 25, where five source modes with equal power were used. We used equal power modes in this test because it can give a more obvious result. After each iteration, the power of each orthogonalized probe mode was recorded. The recorded data were then plotted together, where the order of the mode pattern may vary because they have approximately the same power.

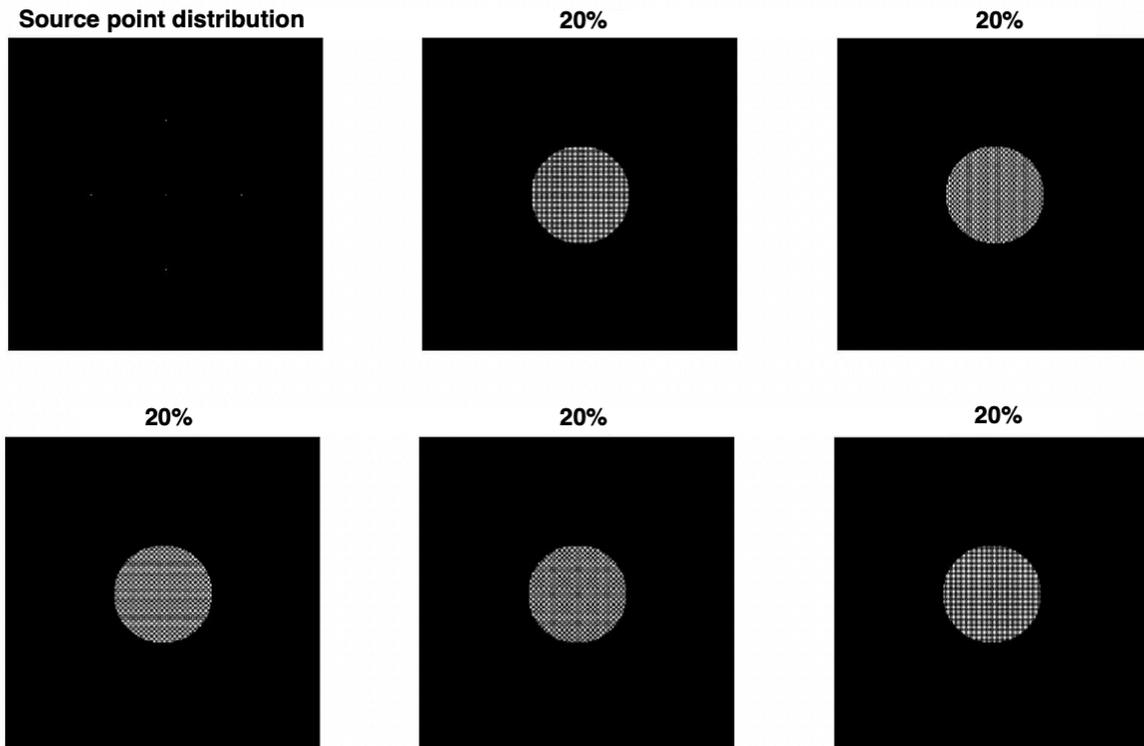


Figure 25. The point source pattern and source modes. The five source modes have approximately equal power. The source points on the source plane have some intensity and a symmetrical spatial distribution.

The plot result is shown in Figure 26, where the five source modes have approximately equal power at the end of 500 iterations. In the beginning, most of the intensity is occupied by one of the mode patterns, and the other seven modes only have only a little. This is as a result of the probe functions that were used having a similar structure when they are used in the calculation. With more iterations, more mode patterns have been separated out and can direct the reconstruction in a more accurate direction. In the beginning, the power of each mode depends on the guessed probe functions. However, instead of working out all the mode patterns at the same time, the power is gradually shared to all other modes. The larger power modes are easier to generate

because they have a greater contribution to the reconstruction. Finally, all the valid modes will be represented by approximately equal power. The remaining modes received almost nothing from this process and did not contribute to the reconstruction. Therefore, they are empty modes in this test.

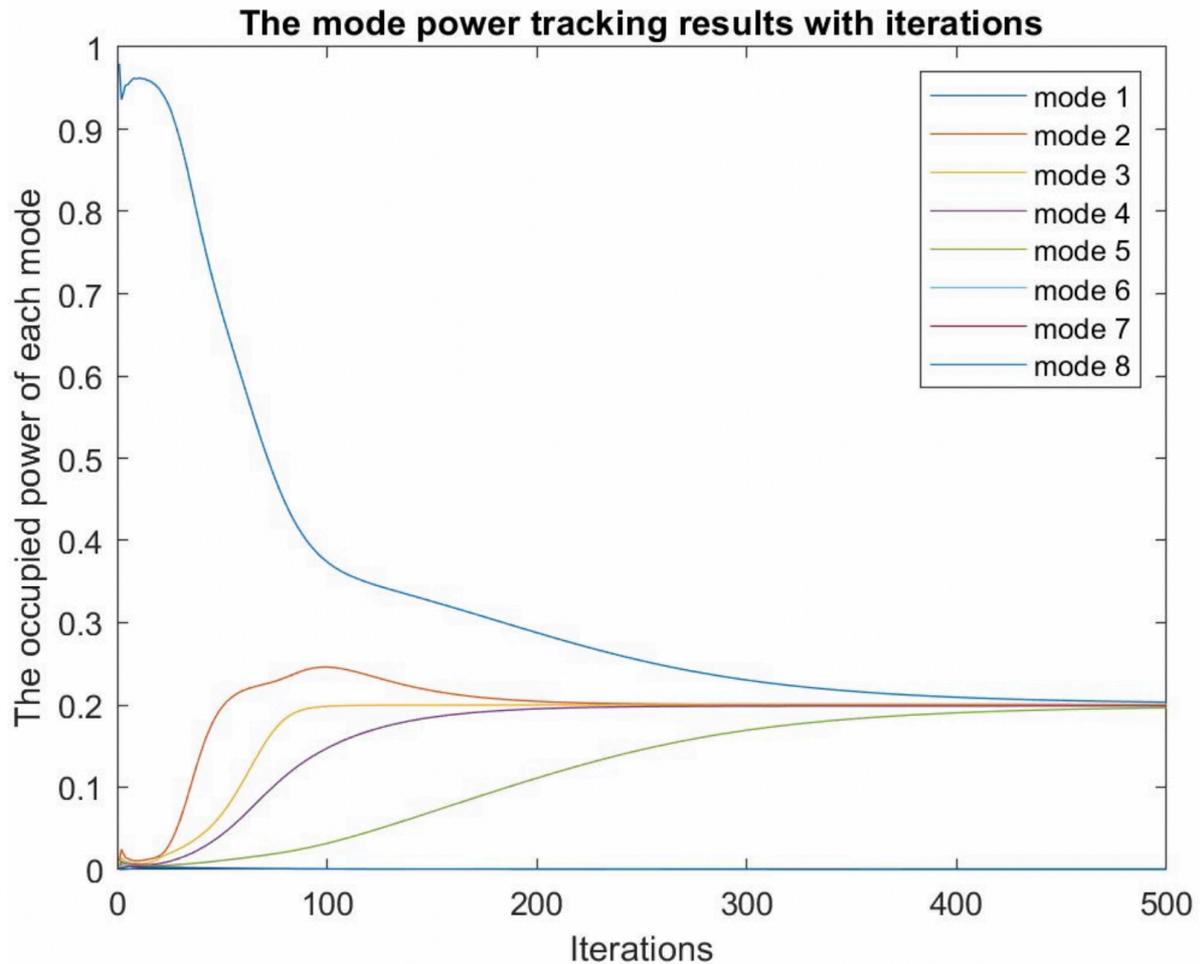


Figure 26. The plot of the power distribution of the nine modes. Modes 1–5, as the source modes, have approximately equal power at the end. Modes 6–8 are empty modes. Except for the beginning probe function guess, through the iterations, the modes with greater power participate in the reconstruction earlier.

#### 4.3.2. Factors influencing modal decomposition

To explore the parameters that could have an effect on modal decomposition, we realised that modal decomposition is a ‘number-of-numbers’ issue. In other words, how many mode functions can we solve given a certain number of measurements? If we want to rebuild an  $L \times L$  pixel specimen using  $K$  modes and each mode is  $M \times M$  pixels, the total number we are going to solve is  $2 \times L^2 + 2 \times K \times M^2$ , where every

value is multiplied by two because they are complex values. For this reason, the minimum requirement is that our data should record more data than needs to be solved. To prove the number-of-numbers argument, we are going to reduce the data by reducing the overlapping levels to check if we can still work out all the known modes. In this test, we use a square probe function. The point source is in Figure 22. However, we reduce the overlap level from 80% to 30%. The table below shows that the overlap levels, the number of diffraction patterns (dp) and the number of modes has been successfully calculated.

Table 1. The overlap levels, the number of diffraction patterns (dp) and the number of modes has been successfully calculated.

Overlap levels	Number of collected dp	Number of true modes
80%	400	5
70%	169	5
60%	100	5
50%	64	4
40%	49	0
30%	36	0

From the Table 1 and Figure 27, it can be seen that when the overlap of any two adjoining scan positions is reduced, it is more difficult to work out all the mode patterns. If too little information is recorded, it is not possible to work out even one correct mode pattern. In addition, the real-space scan process can be treated as the real-space sampling for ptychography, which identifies the first factor affecting the modal decomposition, the sampling. The sampling in the reciprocal space is the density of pixels on the detector, which will assist in recording more frequency-component information. If the probe function that is sampled in the reciprocal space is much smaller than the Nyquist sampling condition, much of the collected data will be empty data, meaning that the data are not enough to solve the number-of-numbers argument. According to the minimum sampling requirement of ptychography [82]

$$\Delta x = \frac{1}{2\Delta u'} \quad (85)$$

where  $\Delta x$  is the sampling unit in the real space and  $\Delta u$  is the sampling unit in the reciprocal space. Therefore, as the sampling increases in one, it will decrease in the other. Therefore, when the overlap on the specimen is reduced, the sampling requirement on the detector is increased. If the detector cannot satisfy the new requirement, there will not be enough information to solve the object and probe functions.

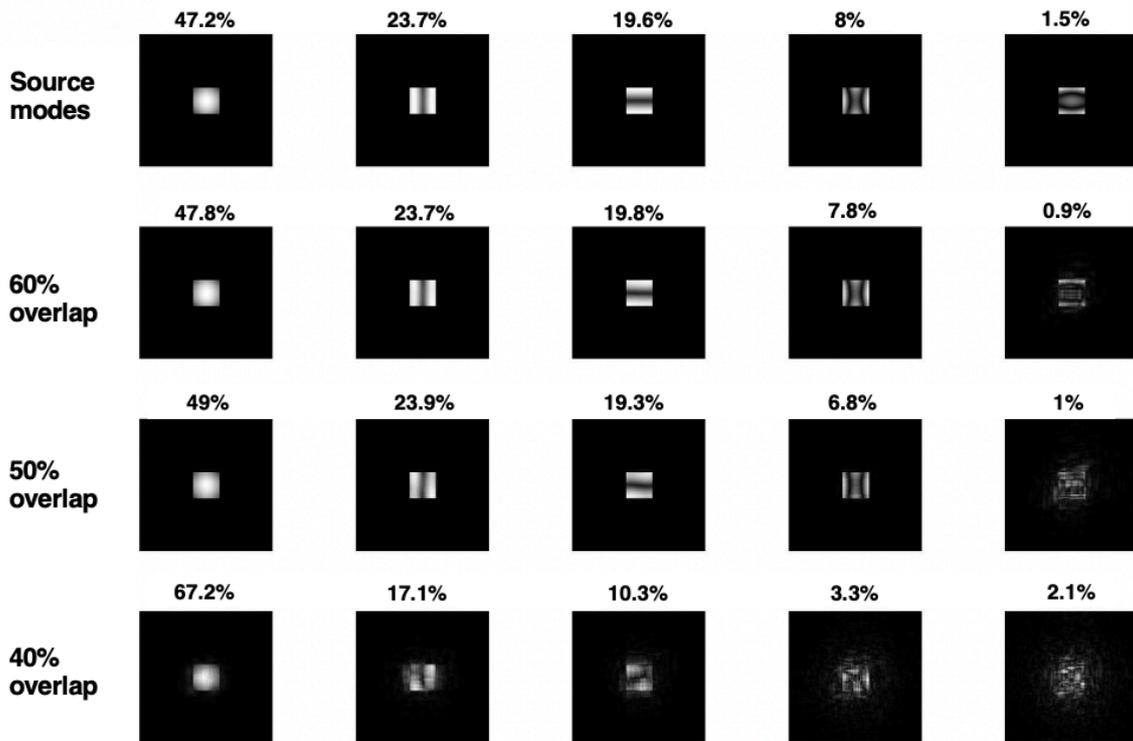


Figure 27. Square probe source modes and the modal decomposition result with 60%, 50% and 40% overlap.

Next, we artificially built two orthogonal moduli-only probe modes and checked the reconstruction result, which is shown in Figure 28. The images on the left show the artificially built probe modes in this experiment, where the two probes have two different values, 1 and  $-1$ . The two images on the right show the modal decomposition results of the two artificial modes. In theory, because the two probe modes are orthogonal probes, the modal decomposition results should be exactly the same as the probe patterns. However, the result shows two orthogonal mixed states of the probe mode patterns, which proves that during each iteration the information in the different modes conflicts with each other, and the final modal decomposition results

show the result of a compromise. This is a possible shortcoming of modal decomposition, which has no effect on the modal decomposition and reconstruction but could result in being unable to uniquely separate all the similarly structured modes from orthogonalization.

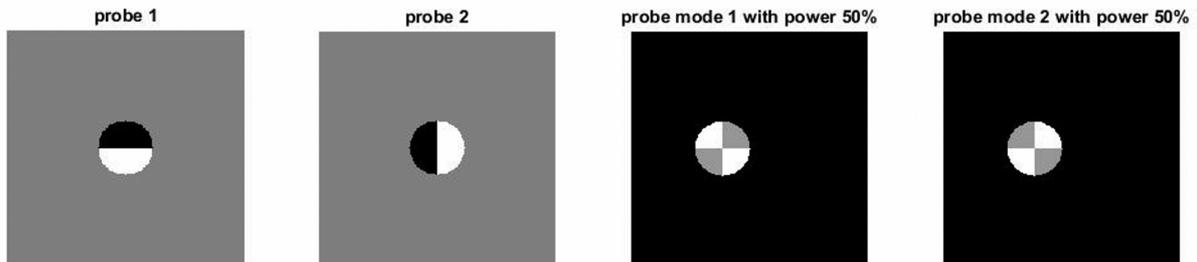


Figure 28. The artificial probes and orthogonal modes. The two images on the left show the artificial probe patterns. The two images on the right show the orthogonal probe modes. The results show that the states in different modes come into conflict with each other during reconstruction.

Even if the minimum data requirements are met, the diversity of the image structure and mode structures could affect the modal decomposition. In other words, several modes can only be separated when the diffraction pattern has enough valid information. To prove this idea, we created two simulation experiments. The first experiment was used to show that the structure of the specimen will affect the modal decomposition. The experiment used a white slash line as the specimen and used the five-point source shown in Figure 20 to generate the partially coherent illumination for the experiment. The modal decomposition results are shown in Figure 29.

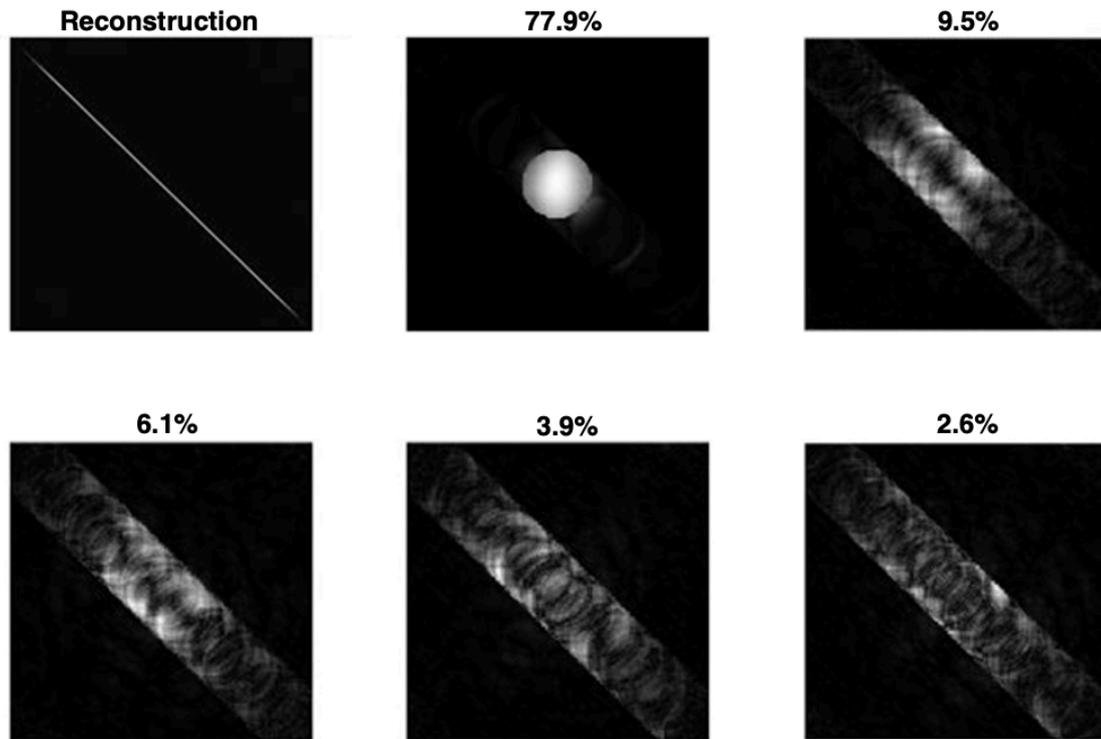


Figure 29. The modal reconstruction and modal decomposition results of a slash line phase image. In this figure, although no correct mode pattern has been calculated, we can also get a good reconstruction because there is enough information in the data to solve a simple structured specimen.

Figure 29 shows that the phase image is represented, but all the modes are destroyed. This is because of the lack of diversity in the diffraction data along the direction of the line. This can also be explained by the number-of-numbers argument that the dataset has not recorded enough information to solve all the mode patterns, and most of the valid diffraction patterns record the same structure information. Therefore, mode pattern ambiguities will occur along the direction of the line. The second experiment showed how the probe mode structure could affect the reconstruction. As shown in Figure 30, the same slash line example was used but with different sources, to give different partially coherent illumination. The reconstruction and modal decomposition results are shown in Figure 31. The source mode has the same direction as the slash line with different patterns. The reconstruction result was worse than the previous one, with some ghost parallel line structure, which is due to the lack of referenced structure information along the other directions in the mode pattern and the large amount of ambiguity along the direction of the line and between different mode patterns.

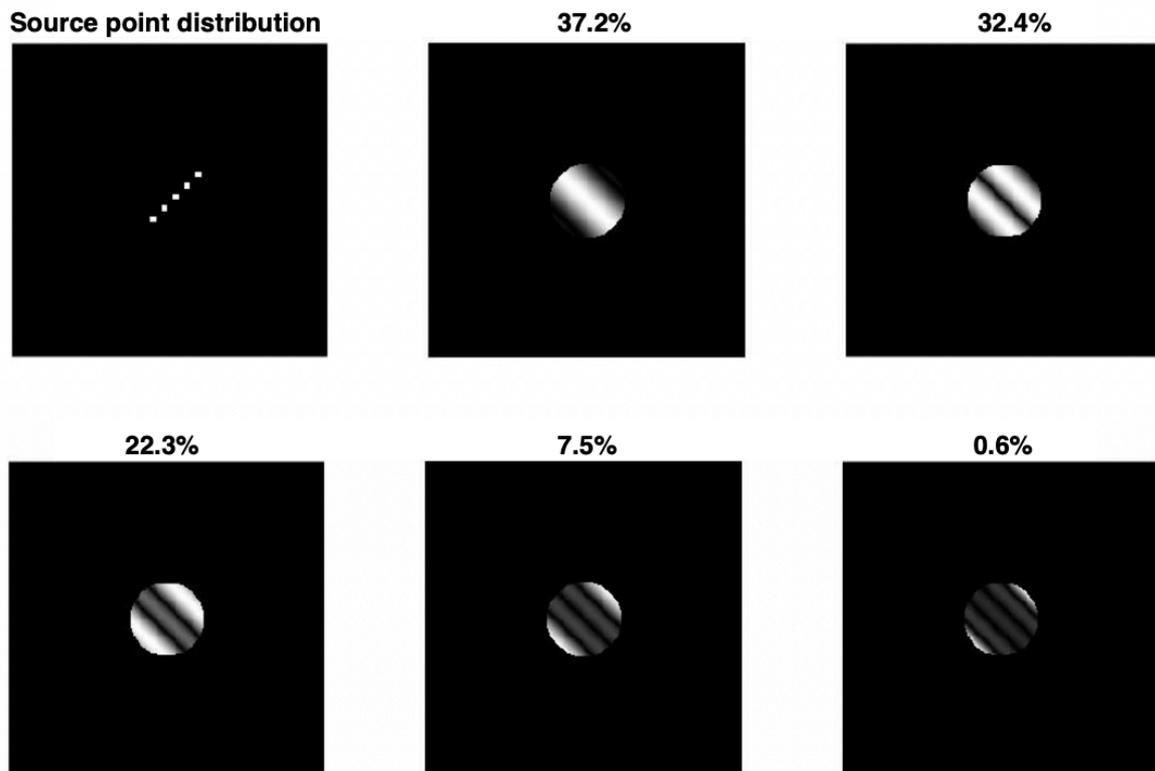


Figure 30. A new, engineered partially coherent source and related source modes with powers. In this source, all the probe modes have the same interfere features along the direction of the slash line.

To plot a graph that describes the relationship between the object structure and the minimum number of probe modes that will give an acceptable reconstruction result, we have searched for a unique solution to give a quantitative description of the structural diversity of the specimen and the probe, including the size of the compressed image, the Shannon entropy, and the frequency histogram. It has been noted that the diffraction pattern contains all the structural information that is needed because the diffraction pattern is the result of the convolution of the object and probe functions. Therefore, we attempted to analyse the diffraction pattern using the peak signal-to-noise ratio of the diffraction pattern, the structural similarity of the diffraction patterns and the probe function on the detector, but the result was not very precise. Therefore, it is not possible to give a precise number for the number of modes needed in a reconstruction because this depends on the source and object structures.

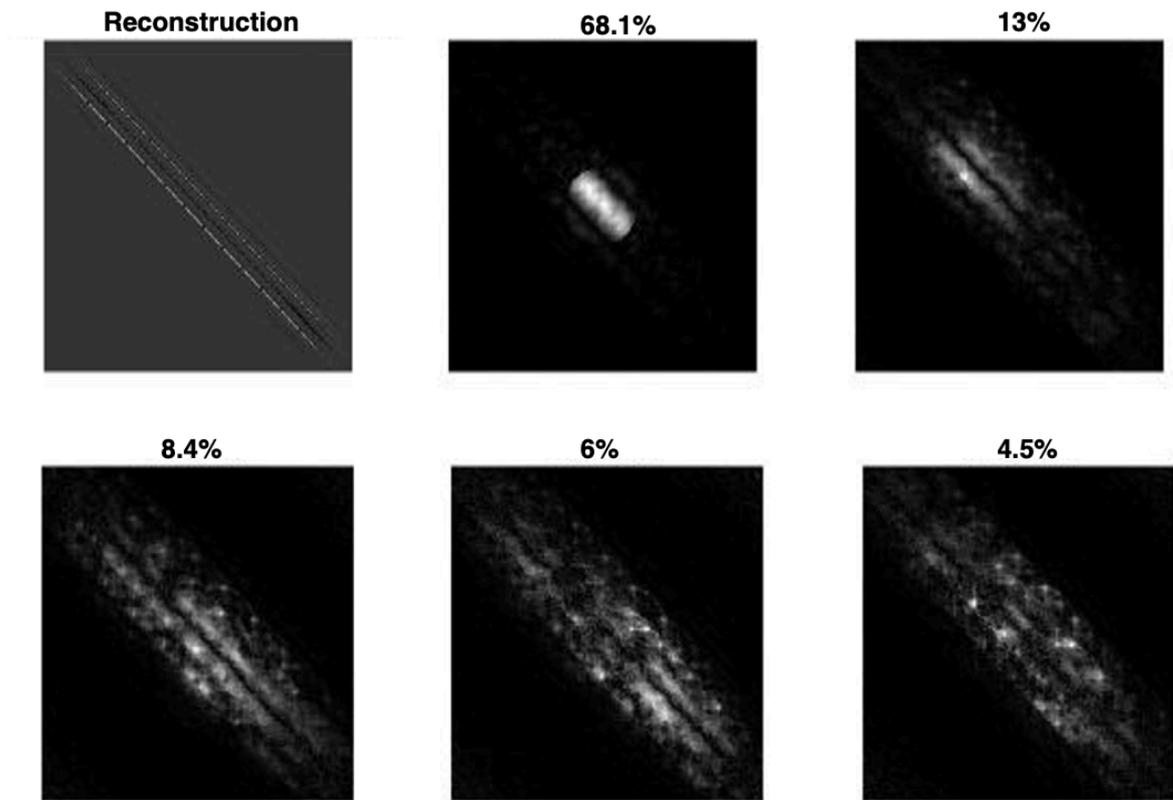


Figure 31. The reconstructed phase image and modal decomposition results. In this figure, the reconstruction has some ghost parallel lines with the real slash line. In this test, this is due to the object function and probe modes not being able to provide reference position information in directions other than in the direction of the slash line.

#### 4.4. A sample of object state modes

The object modes can be understood as having more than one specimen or noise information recorded in the same dataset. To simplify things, the illumination in this part is purely coherence. Figure 32 shows example results of two-object state modal decomposition and reconstruction. The top row shows the two-object phase image, and the bottom row gives the reconstruction results, where it is clear that all the object phase images are fully separated. However, due to the reconstruction, object mode order is random. Therefore, the order of the reconstruction images may not be the same as the object images. Therefore, the modal decomposition is treated as a number-of-numbers problem. All the parameters that affect the reconstruction of the probe modes will also affect the object modes. There are some potential uses for

object state modes, such as reconstruction denoising and dealing with radiation damage.



Figure 32. The object mode example. The top row shows the two object phase images, and the bottom row shows the reconstruction results.

#### 4.5. Conclusion

In this chapter, we examined the mechanism of modal decomposition, where the order of the probe modes in modal decomposition deal with different modes in descending order of power. We showed some parameters that could affect the modal decomposition and proved that, because most of the power is occupied by a few

modes, it is not necessary to generate all the modes to achieve a good reconstruction quality of the specimen. In addition, it was found that during reconstruction, different modes conflict with each other. In some complex situations this may generate some compromised modes rather than real modes. Aside from the collected data, it was also shown that the diversity of the structures of the object specimen probe modes can affect the modal decomposition. It was also shown that modal decomposition might not work for some very simple, structured object images. However, in the beginning we aimed to find a unique value to represent the structure of the specimen and the mode pattern to quantitatively analyse the relationship between the specimen or the probe mode pattern and the reconstruction quality. However, this was not practical given the strong dependence of the reconstruction quality on the structure of the probe and the sample. Sampling conditions and iteration algorithms also affect modal decomposition. These need to be treated with care during the reconstruction.

## **Chapter 5 Engineering illumination modes in X-ray ptychography**

As a kind of coherent diffractive imaging (CDI) method, ptychography has widely been used in the x-ray microscopy community. As the name shows, coherence is important to all CDI methods; so, in theory, the quality of a ptychography experiment is highly limited by the degree of beam coherence. To achieve a highly spatially coherent beam, x-ray-synchrotron-experiment facilities must be a large distance from the source, and a suitable aperture must be used to contain the radiation that passes through the aperture relatively coherently because, compared with the beam diameter, the coherence width is small. As described above, mixed-state ptychography methods are widely used to deal with partial coherence in ptychography experiments, as any partially coherent beam can be separated into a series of probe modes, which are self-coherent but incoherent with one another. However, an adequate number of modes must be used to fully cover the partial coherence and avoid the ill-condition of the reconstruction. A key application of ptychography-data redundancy in this thesis is the artificial engineering of probe modes to optimally solve the partial coherent problem. To prove this idea, we designed some simulation experiments in which some positive results were gained and realised some insufficiencies where improvement is needed.

The real experiment was done at Diamond Light Source's I13 beamline, where two set of tests were conducted using hole masks and wires, and our experiment results were demonstrated with the wire test set in which the reconstruction quality was compared with a wire parallel with and cross the coherence direction, and an FZP-only experiment as reference. The results of our experiment bear out our idea and have the potential to speed up the scan.

## 5.1. Simulation theory

If a partially coherent beam passes through an aperture that has a wider diameter ( $2D$ ) than the beam coherence width  $D$ , the exit beam is mostly incoherent as most of the beam pairs are incoherent with others. However, if the partially coherent beam passes through an aperture with a diameter that is the same as the beam-coherence width, the exit beam will become relatively coherent. Therefore, if we engineer a certain number of totally incoherent but self-coherent modes for our partially coherent radiation, we only need to find the number of modes necessary to avoid having many unnecessary modes.

This method has some advantages. The first is overcoming the negative influence of partial coherence, as we use more coherent flux in the experiment; the second is reducing the calculation, as of a large number of modes are a waste of calculation in the mixed-state ptychography method; the third is a larger field of view; the fourth is that we can speed up the experiment; the last is that we can achieve better reconstruction with only a few modes.

## 5.2. Round aperture tests

We started by modelling various experimental aperture set-ups. In the first experiment, we prepared two aperture masks, as show in Figure 33; one of them is a large-diameter circle aperture and the another has four small-diameter circle apertures. We set the mask-area pixel value to 0, or opaque, and the aperture-area pixel value to 1, or transparent. Also, the total aperture area of both masks is the same to ensure that the same amount of flux can pass through them at the same time. The distance between two apertures, vertically and horizontally, is the same as the aperture diameter of the four-aperture mask to minimise any interference in the test.

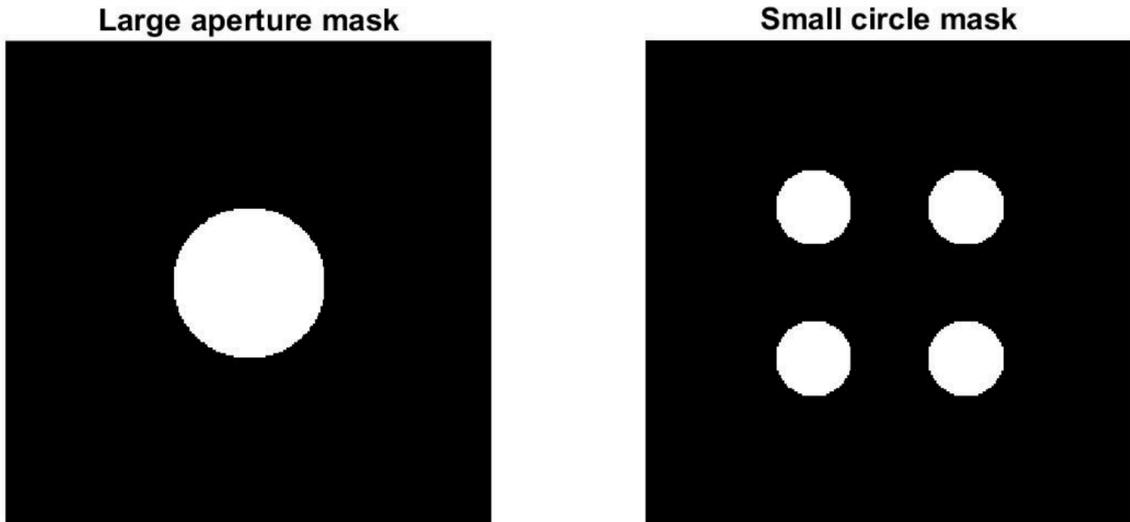


Figure 33. Two designed aperture masks. On the left is a large aperture mask, and on the right is a four-aperture mask. They have the same aperture area.

The partially coherent radiation source function we used here is the truncated cos function

$$S(x, y) = C \times \left( 1 + \cos \left( \beta \times \sqrt{x^2 + y^2} \right) \right), \quad (86)$$

where  $C$  controls the intensity of the source,  $\beta$  is the source width in angle space, and  $x$  and  $y$  are real space coordinates, where they are simulation parameters which model the broad qualities of the for the purposes of our calculations experiment source. In Figure 34, we give an example image of the source function and the related partially coherent function at  $C = 0.5$  and  $\beta = 20$ .

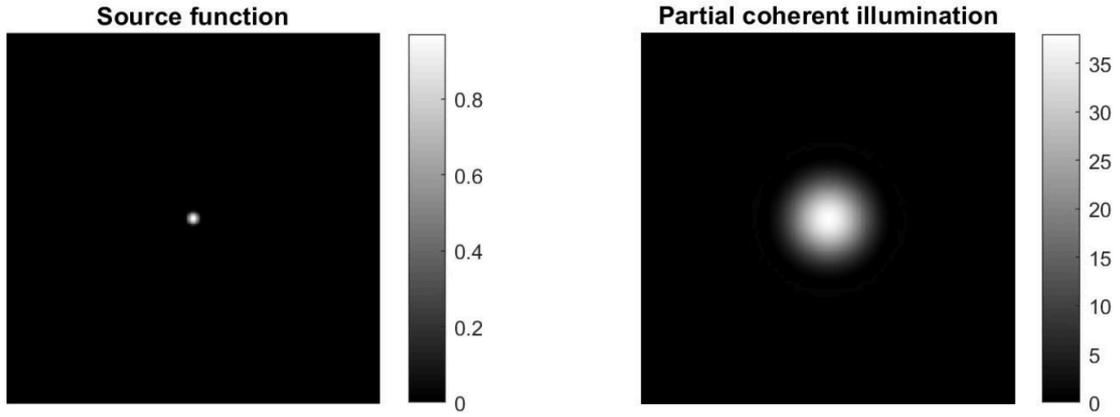


Figure 34. The source function generated by equation (1) at  $C = 0.5$  and  $\beta = 20$  and the related partially coherent function. The image on the left is the source function and the one on the right is the partially coherent function.

The source propagates to the aperture as function

$$b(u, v) = f(S(x, y)), \quad (87)$$

in which  $u$  and  $v$  are the coordinates in the Fourier space, and  $f$  is the Fourier transform. When the illumination passes through the apertures and lens, the one-point source probe function is formed by

$$p(x, y) = f^{-1}[b(u, v) \times M \times e^{i\phi}], \quad (88)$$

where  $M$  is the aperture mask function, and  $e^{i\phi}$  is the phase change caused by the lens. Then, the partially coherent function is given by the sum of the probe functions of all points:

$$p(S(x, y)) = \sum_a p(x, y). \quad (89)$$

On the detector, if there is not a lens in the experiment, the diffraction pattern is given by the convolution of a coherent diffraction pattern with the partially coherence source function as

$$I = |f(p(a))|^2 \otimes S, \quad (90)$$

in which  $\otimes$  denotes the convolution. If there is a lens in the experiment, the diffraction pattern is equal to

$$I = \sum_a |f(p(a)O)|^2, \quad (91)$$

where  $O(x, y)$  is the object function. The experiment is designed as Figure 35. From the left to the right, the figure shows the partially coherent source, one of the masks from Figure 33 with a big aperture or four smaller apertures, a condenser lens, which is not necessary for a ptychography experiment, the specimen at the specimen holder and the detector at the far end.

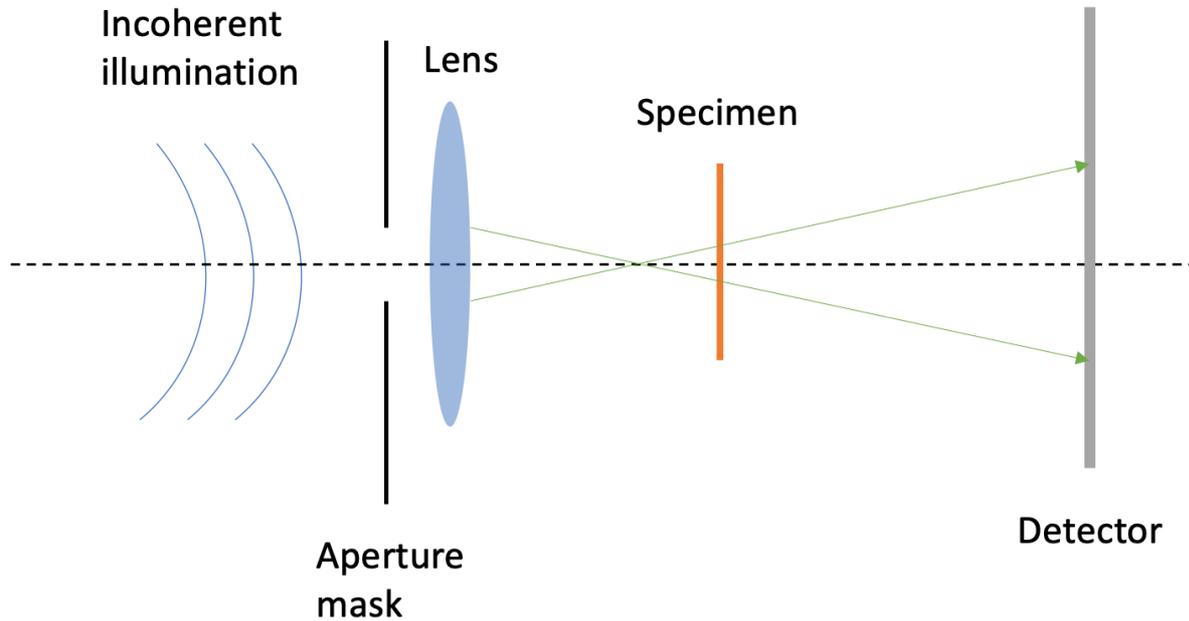


Figure 35. The setup of the simulation experiments. The incoherent illumination passes through the aperture and lens, which is used to focus the relatively coherent illumination on the specimen, and the diffraction pattern is recorded on the detector at the far end.

Here, we use a bird photo as the phase of a pure-phase complex specimen, and the phase changes between 0 and  $0.1\pi$ . At the start of the reconstruction, we guess four

isolated aperture probes and each one on the right position of one hole. Therefore, in theory, because the diameters of each hole are close to the coherence width, each of the probe functions should be self-coherent but totally incoherent to one another. So, when the reconstruction is finished, we can attain the same four apertures without interfere from the other hole probes. However, when we did the four-aperture mask simulation test, we found that it does not work as well as we thought. This is due to the fact that the Fourier transform of a round aperture at any spatial position is the same, and it is hard to separate the recorded data of hypothesised aperture functions. Therefore, we artificially cut parts of the aperture functions to make four dissimilar apertures, and we kept the shape of the large aperture, as show in Figure 36.

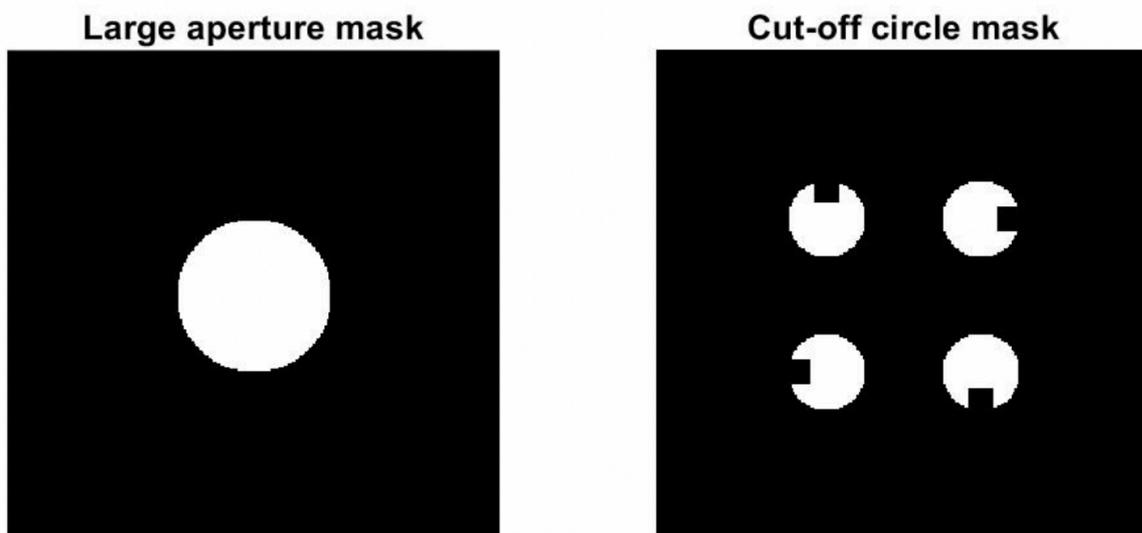


Figure 36. Two redesigned aperture masks. On the left is a large-aperture mask, and on the right is a four-aperture mask with cut-offs.

When the coherence width of the beam is near the diameter of the cut-off apertures, we do normal ptychography for each aperture mask with four mode functions in the calculation and keep the same step size to gain a fair result. Figure 37 provides the object-phase image and phase-retrieval results from the large-aperture mask, cut-off aperture mask and four-aperture mask. According to the result, the reconstruction quality of the cut off aperture is significantly better than that of the others. First, we compare the reconstruction quality of the cut-off aperture and large aperture, where, as the size of the cut-off aperture is quite close to the coherence width of our simulated illumination, the illumination is more relatively coherent passed the cut-off aperture

than the large aperture. Also, we only put four modes into the calculation, which may not be enough to deal with the partial coherence in the large aperture's reconstruction. Therefore, we are going to compare the cut-off aperture and four-round-aperture tests. As mentioned before, the Fourier transform result of a round aperture is not related to the spatial position of the aperture, which means that by doing inverse Fourier transform, there are several ambiguous 'correct' results. Meanwhile, with the reconstruction one position gives more stronger, the algorithm treats the four-aperture modal decomposition process as one small-aperture reconstruction with four modes, so all of modes are gradually calculated in the same way and becomes the other states of one illumination mode, as show in Figure 38.

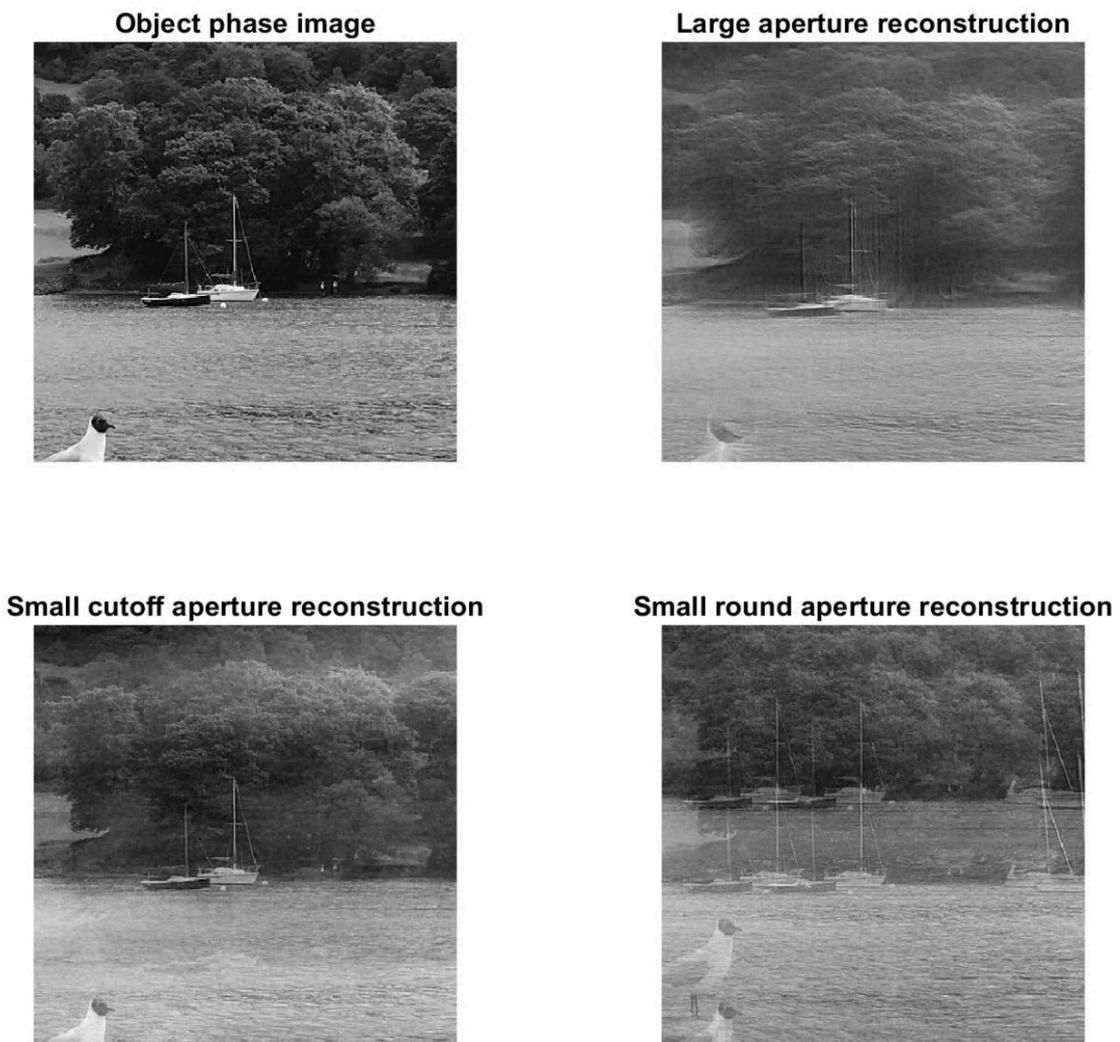


Figure 37. The object-phase image and phase-retrieval results, achieved by using the big aperture, four, small, cut-out apertures and four, small, round apertures.

Comparing the large-aperture and cut-off-mask reconstruction, the cut-off reconstruction is more ideal, as the illumination is more coherently passed through the cut-off aperture. Also, four modes are not enough in the calculation for the large aperture to overcome the partial coherence. Comparing the cut-off and round aperture, the contrast in the round aperture still good because of the coherence width. However, the round aperture is not represented at the right position and causes some overlapped repeat features, because the Fourier transform result of a round aperture is not related to the spatial position of the aperture, which can break the ambiguity of the centre symmetric diffraction pattern.

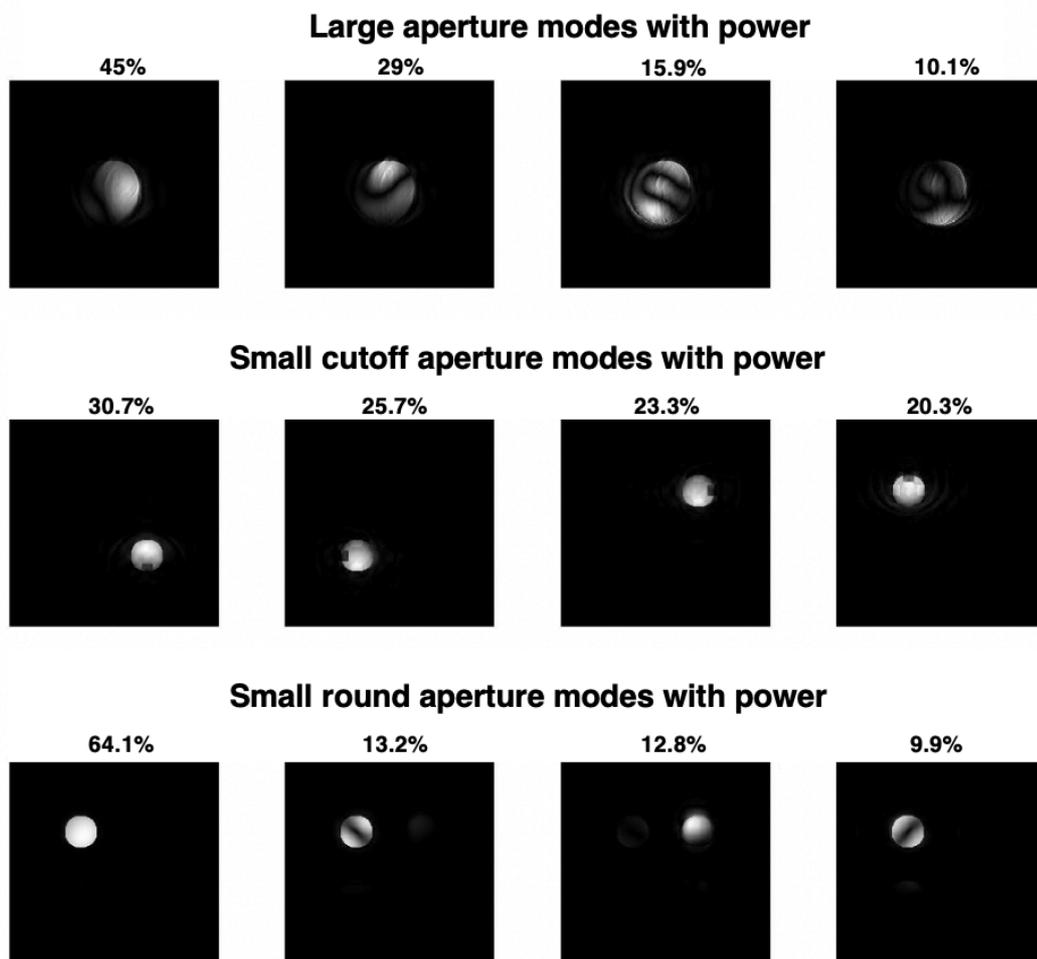


Figure 38. The modal-decomposition results of the big aperture, four, small, cut-off apertures and four, small, round apertures. The most remarkable result is that we separated the four isolated coherent modes of the cut-off aperture. As described, this is due the fact that each of the isolated modes are self-coherent but incoherent with others. Also, it proved our idea that

by using small apertures with the same size as the width of the illumination coherence, or smaller, we can remove the partially coherence from the reconstruction.

### 5.3. Corners simulation experiment

Another experiment we proposed to overcome the Fourier ambiguity in the Fourier domain of four round apertures, is to split the large aperture into four aperture corners, as shown in Figure 39, where the left side of the large aperture is the same as in Figure 36, and on the right side, we separate the same big aperture into four aperture corners; any two corners have a gap distance equal to the side length of the corner. We did this corner shape because in the real experiment we will use some optical components, and the corner can easily be built by adding wires to the lens.

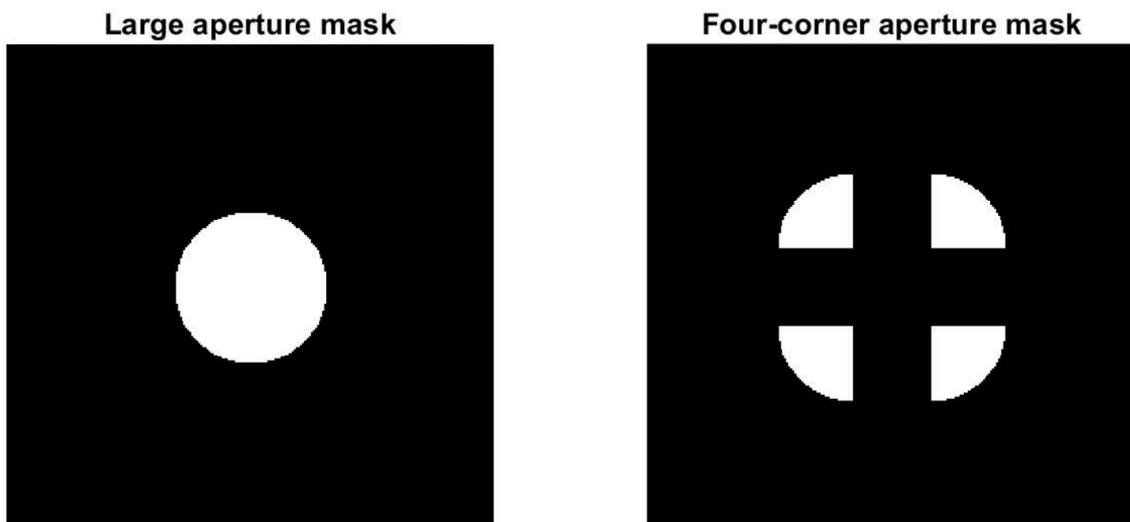


Figure 39. The aperture masks used in the simulation. On the left, the large aperture is the same as in Figure 36, but on the right, four corners are used to replace the artificially engineered round apertures and each aperture has one-fourth the large aperture's area to provide the same flux in both experiments.

We use the same source to do the same experiment by using the corner mask and comparing the reconstruction results with the large-aperture reconstruction in Figure 40. In Figure 40, on the left-hand side is the reconstruction of the large aperture as in Figure 37, and on the right-hand side is the reconstruction of the corner aperture. The reconstruction results are as we thought: the corners use the coherence area on the

partially coherent beam more than the large aperture and break the ambiguity of the centre symmetric diffraction pattern.

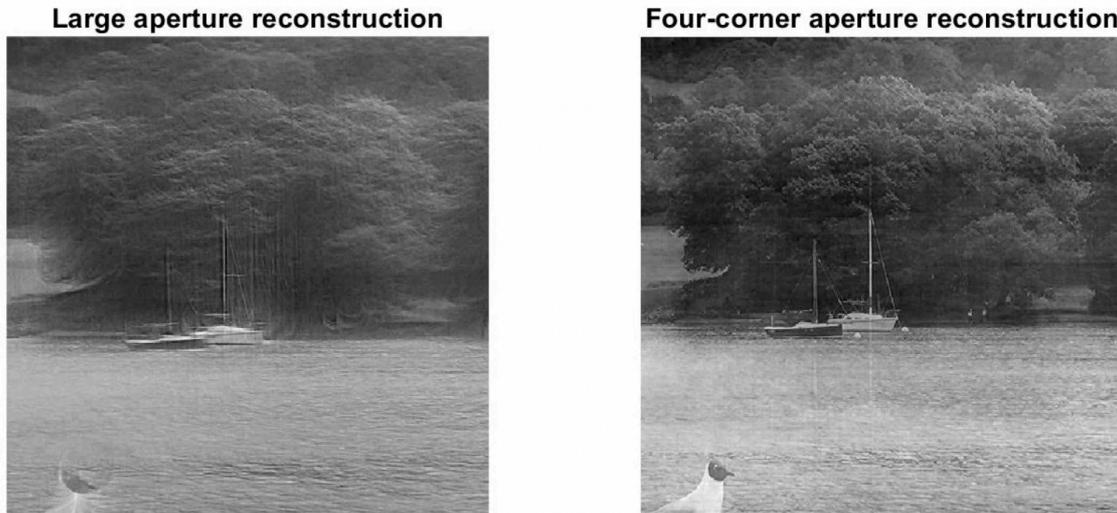


Figure 40. The reconstruction results of the corner aperture compared to the large aperture. The results prove our idea once more and confirm that the corner aperture is available to be applied in our experiment.

We also provide an FRC result to quantitatively analyse the difference in reconstruction in Figure 41 (see the next page). The top line is the FRC-reconstruction result of the corner aperture, and the bottom line is the FRC result of the large aperture. It is clear that, in any normalised spatial frequency, the corner aperture can provide a much better result than the large aperture because the beam is more coherent when it passes the corner aperture. In figure 42, we demonstrate the modal decomposition results of large aperture and corner aperture to verify that the corner aperture is as good as cutoff apertures.

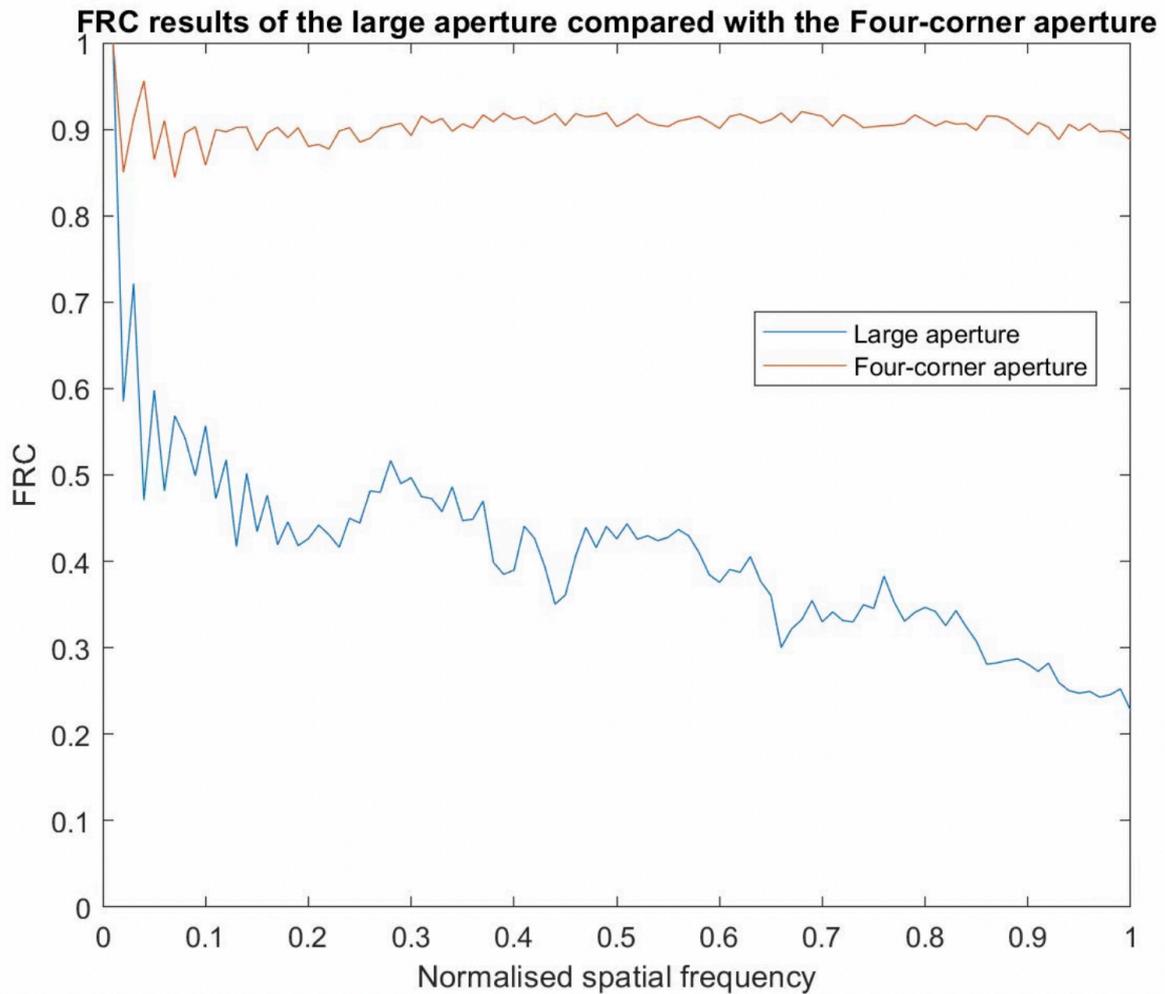


Figure 41. The FRC results of the large-aperture and corner-aperture reconstructions with the object function. The line on the top is the FRC result of the corner aperture, which shows steady, high reconstruction quality. The line at the bottom is the FRC result of the large aperture's reconstruction. There is a significant gap between the corner and large-aperture results, as the illumination is more coherent past the corner aperture than the large aperture.

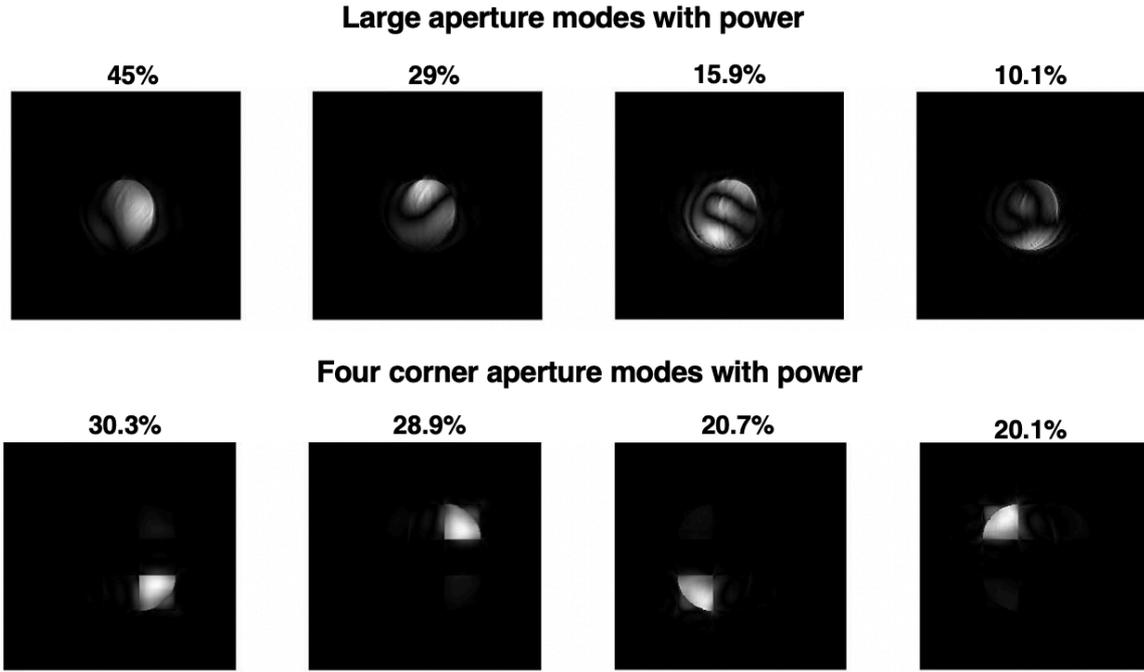


Figure 42. The modal decomposition result of the large aperture and the corner aperture. It is clear that the illumination passes each corner more coherently than the large aperture, and any two corner apertures are as incoherent as four isolated coherent beams.

#### 5.4. Corner aperture lens simulation

We now examine a more realistic experimental set-up using a lens. These calculations emulate the true experiments we present later. We test if the lens works well. It causes a phase-change defocus function in the form according to

$$D\phi = e^{-1i*\lambda*dl*(x^2+y^2)}, \quad (92)$$

where the  $dl$  is the defocus length,  $\lambda$  is the illumination wavelength and  $x$  and  $y$  are the coordinates in the Fourier space.

In this part, we model real data to conduct the simulation experiments. The x-ray energy in the simulation is 9.7keV, and the wavelength, after calculation, is  $1.278 \times 10^{-10}$  metres. The lens (FZP) diameter is  $400 \times 10^{-6}$  metres, the focal length  $470 \times 10^{-3}$ metres and the NA is 0.0043 mrad. The detector pixel size is  $55 \times 10^{-6}$  metres, and there is a total of 256 x 256 pixels. The defocus is  $4 \times 10^{-1}$ metres, and

the camera length is 15 metres. These values are typical for the I13 beamline at the Diamond Light Source where the physical experiments take place.

### **5.5. The corner-aperture lens-simulation experiment**

Here, we use the above lens function for the four-corners mask experiment and compare two pairs of results to test whether a lens can improve the performance of ptychography at low coherence levels.

The first pair of results are the reconstruction and FRC results of the corner and large-aperture masks at a coherent level, where beta ( $\beta$ ) in our source function (86) is equal to 18. In another pair of results, we change  $\beta$  to 12 to simulate a less-coherent level. The reconstructions are compared in Figure 43, in which the upper row contains the phase-retrieval results of the large aperture and mask at a partially coherent level, where  $\beta$  in equation (86) is set to 18. The bottom row contains the results at a lower coherence level, where  $\beta$  in equation (86) is set to 12. At first, we compare the reconstruction of the corner mask with that of the large aperture mask at the same coherence level. We find that in any coherence level, the corner aperture reconstruction is better than that of the large aperture, as it is relatively coherent. Also, we find that the corner mask can provide a larger field of view, as it is more widely distributed. Then, by comparing the large and small apertures' reconstructions at different coherence levels, we can find that the large aperture mask's reconstruction deteriorates more quickly than that of the corner mask. One reason is the difference in coherence, as the illumination is more coherent past the corner aperture, so the collected data will record more useful information as redundancy. Another possible reason is that the strong interference between any two corners, as show in Figure 44, also contributes to the diffraction pattern, so relative overlap is larger than in the case of the large aperture.

**Large aperture: beta = 18**



**Large aperture: beta = 12**



**Four-corner aperture: beta = 18**



**Four-corner aperture: beta = 12**



Figure 43. The reconstruction results of the large and corner apertures with the lens at different coherence levels. We put four modes into the calculation, and the results shows that the corner aperture always performs better than the large aperture at different coherent levels. Also, because of the beam passed through the corner apertures is more relatively coherent than passed through the large aperture, so the reconstruction quality is better at any time.

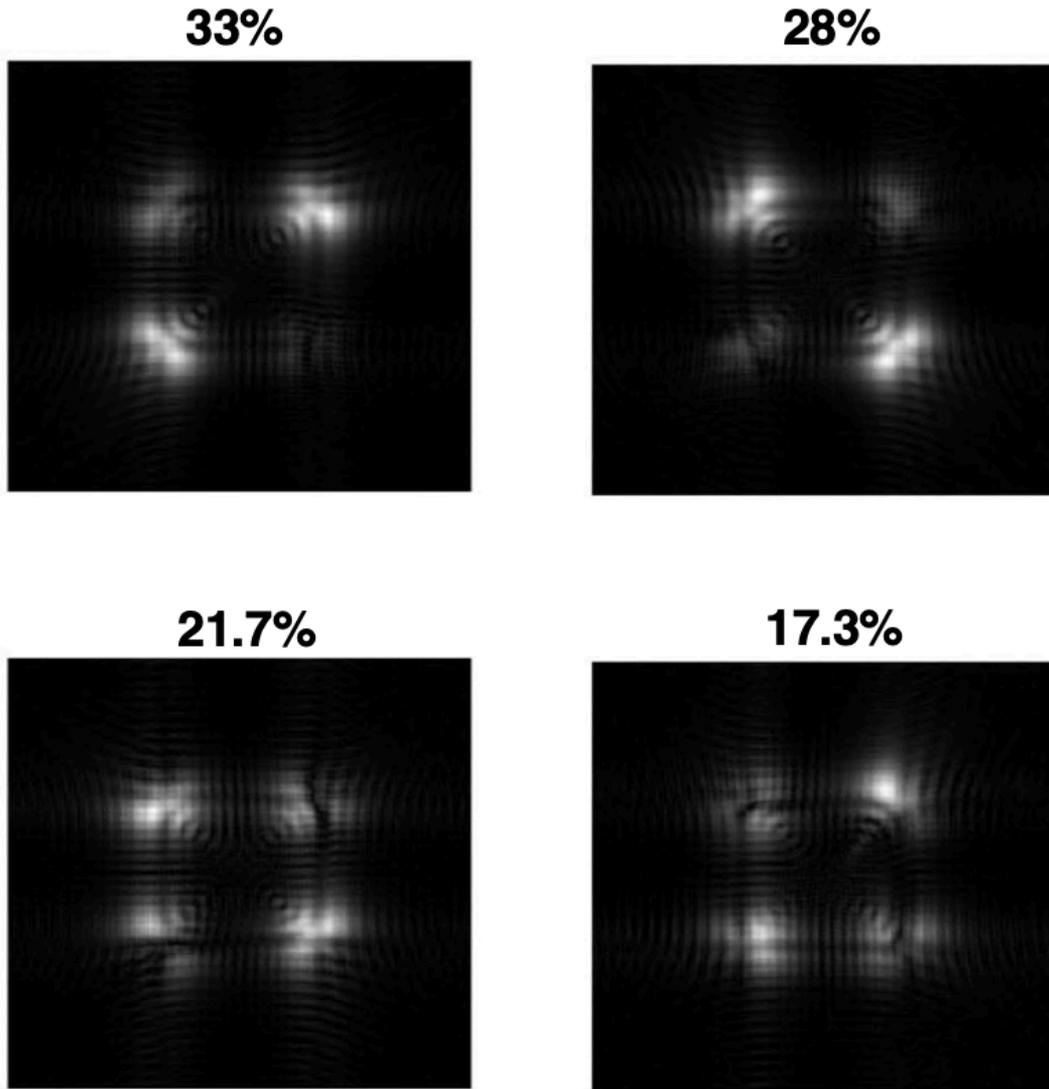


Figure 44. The modal decomposition result of the corner mask with the lens and with beta  $\beta = 18$ . In this result, the strong inference between any two apertures is caused by the diffraction at the border of the apertures when the illumination passes the aperture mask and is magnified by the lens.

In Figure 45, we use the FRC to quantitatively analyse the reconstruction difference, where from the top to bottom, the lines represent the corner-aperture mask at beta  $\beta = 18$ , large aperture mask at beta  $\beta = 18$ , corner aperture mask at beta  $\beta = 12$  and large aperture mask at beta  $\beta = 12$ . From the result, we can find that the reconstruction quality of the corner-aperture mask is better than that of the large aperture at the same coherence level, and reconstruction quality drops slowly when decreasing coherence level.

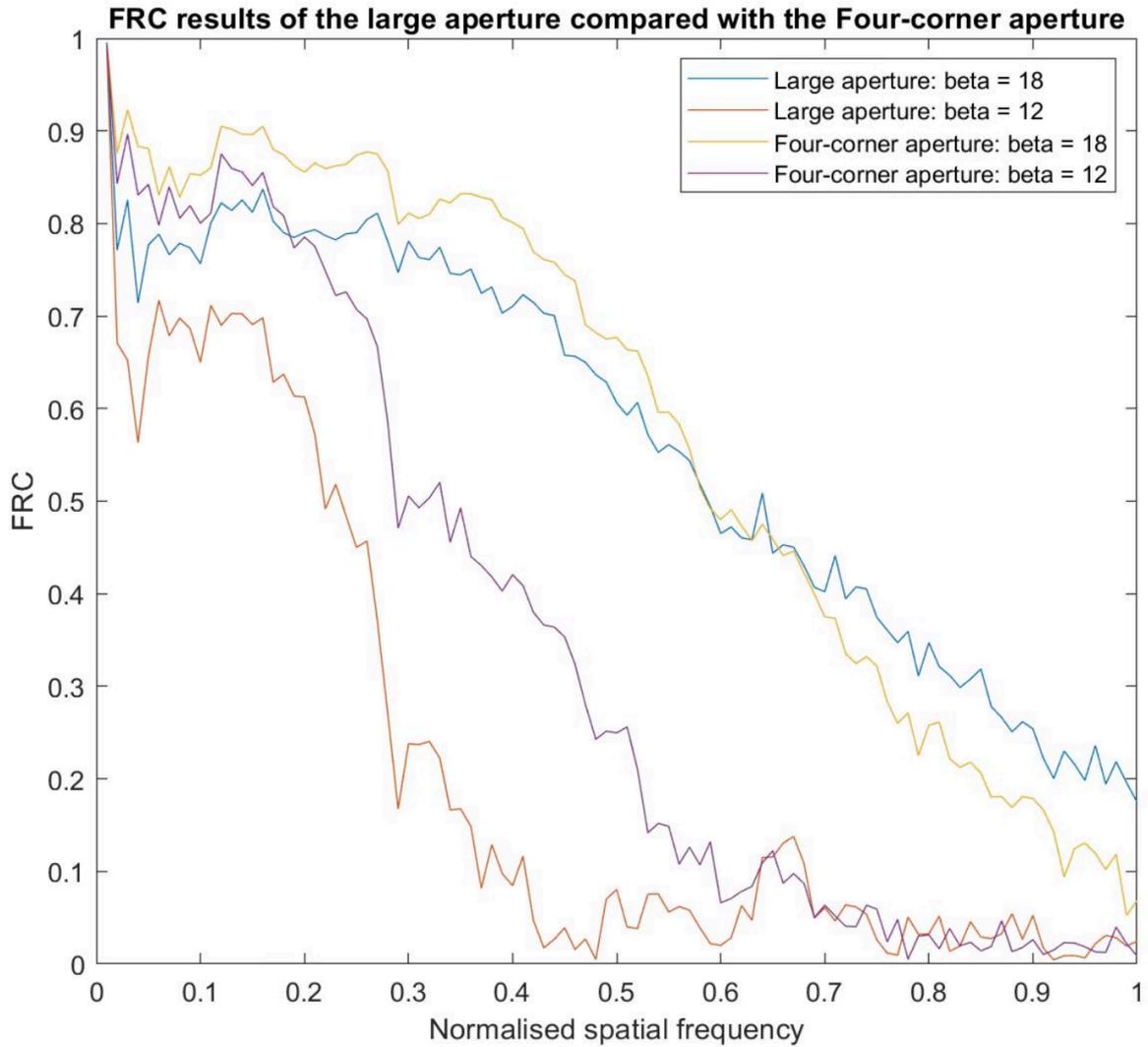


Figure 45. The FRC results of the large and corner apertures at two different coherence levels. From top to bottom, the lines represent the corner-aperture mask at beta  $\beta = 18$ , large aperture mask at beta  $\beta = 18$ , corner aperture mask at beta  $\beta = 12$  and large aperture mask at beta  $\beta = 12$ . The corner-aperture results are better than those of the large aperture at two different partially coherence levels and when the coherence level is decreasing; influence on the corner aperture is smaller as well. There is a cross point between big aperture and four corner aperture experiment when  $\beta = 18$ , which is due to the FRC frequency is more continue in the big aperture than corner experiment, as there are some interferences between corner apertures.

To find what is behind the modal decomposition, we conducted a test at beta  $\beta = 18$  and in a larger step size situation to amplify the effect of each probe mode, and the modal decomposition result is shown in Figure 46. In the figure, we can find that there is an interference pattern between the different corner probes along horizontal, vertical

and diagonal directions, which is the reason why we cannot separate each isolated mode.

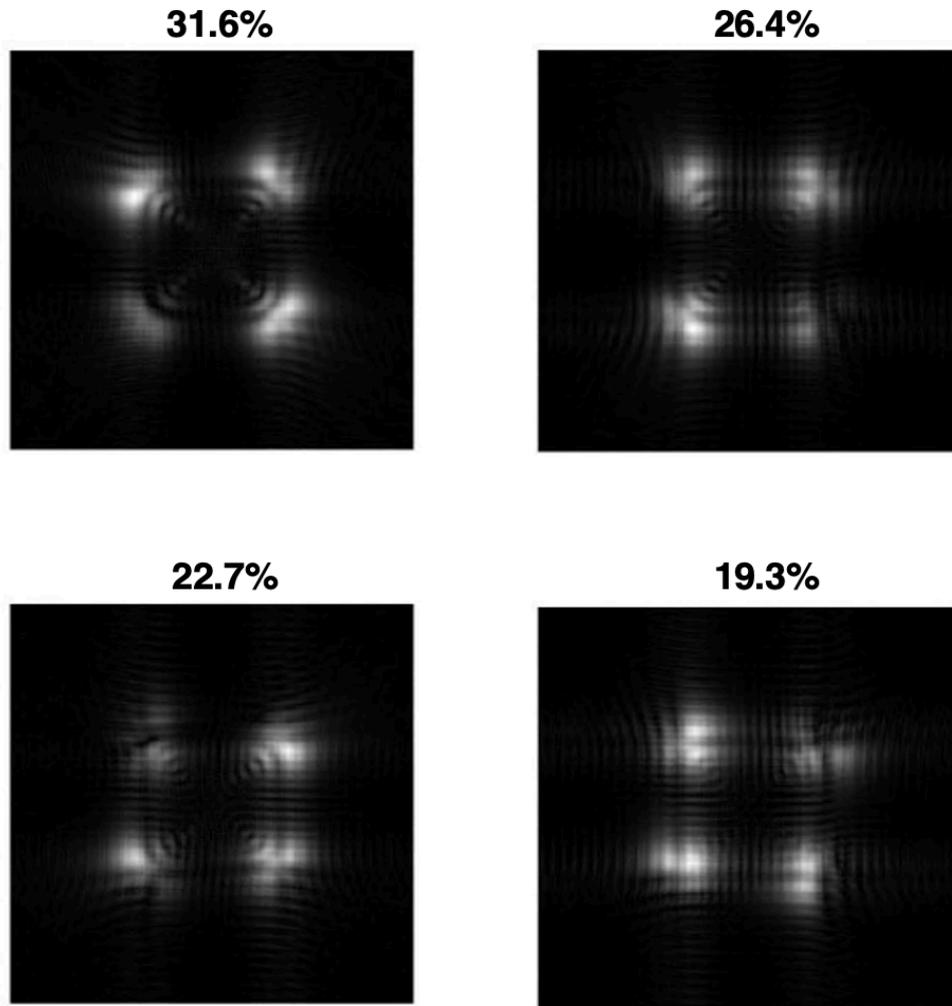


Figure 46. The modal decomposition result of the corner mask in a purely coherent situation. From the results, it is clear that there is some interfere between different probes horizontally, vertically and diagonally.

## 5.6. Simulation discussion

The above results prove that our idea is tenable in the true experiment. By using small apertures or corners, the reconstruction quality has been significantly improved and provides a better tolerance to partial coherence in the beam as more information is recorded by the coherent illuminations. The number of necessary modes was significantly reduced to reduce calculation requirements and ill-conditioned reconstructions. Also, the small aperture or corners give a large field of view with a

lens, as we believed. The simulation results show us how to conduct the experiment in a real facility.

## 5.7. Diamond experiment

### 5.7.1. The Diamond Light Source and beamline I13

The Diamond Light Source is a UK government-funded national third-generation synchrotron facility, which can create light millions of times brighter than the sun. The name Diamond is an abbreviation of dipole and multipole output for the nation at Daresbury (Daresbury was the site of the first UK synchrotron and the original planned location of the Diamond). The first beamline in the Diamond Light Source was opened in 2007, and, currently, there are 32 beamlines in operation. Each beamline directs generated light or x-rays to the lab facilities and have different purposes in research and study. An image of the Diamond Light Source from the official website is presented in Figure 47. The storage ring has a circumference 561.517 metres, in which the electrons have 3GeV energy that gives 300mA maximum beam current, and the revolution frequency is 533.8KHz.



Figure 47. Bird's-eye-view photo of the Diamond Light Source on the official website.

The beamline I13 is the longest beamline at the Diamond Light Source (about 250 metres) and is divided into two beamlines. The first is coherence-branch line I13-1, which can provide highly coherent x-ray radiation to do CDI experiments, such as classical CDI, ptychography and Bragg CDI, at 6-20keV. Another beamline is Diamond Manchester Imaging Branch Line I13-2, which can provide an 8-20KeV energy range for real space imaging and 3D tomography. For our experiment, we used beam line I13-1 as we need to change the coherence degree of the x-ray beam and compare the results at different coherent degrees of the large and multi-aperture probes.

### 5.7.2. Experiment setup

The general experiment setup is quite close to our simulation. The only difference is that zone-plate optics are used to replace the lens; Figure 48 presents a schematic of zone-plate optics and Figure 49 the real setup. In the Figure 48, from along the beam front, the main components are as follows: the Fresnel zone aperture (FZA), which is a mask to isolate the zone plate; the central stop (CS), which is used to stop the direct beam; the FZP, which is used to replace the lens, the masks used in this experiment to engineer illumination modes; the order-sorting aperture (OSA) to block unwanted illumination; the camera to correct to the position of the zone-plate-optics components; the specimen; and a detector at the far end to collect the diffraction patterns.

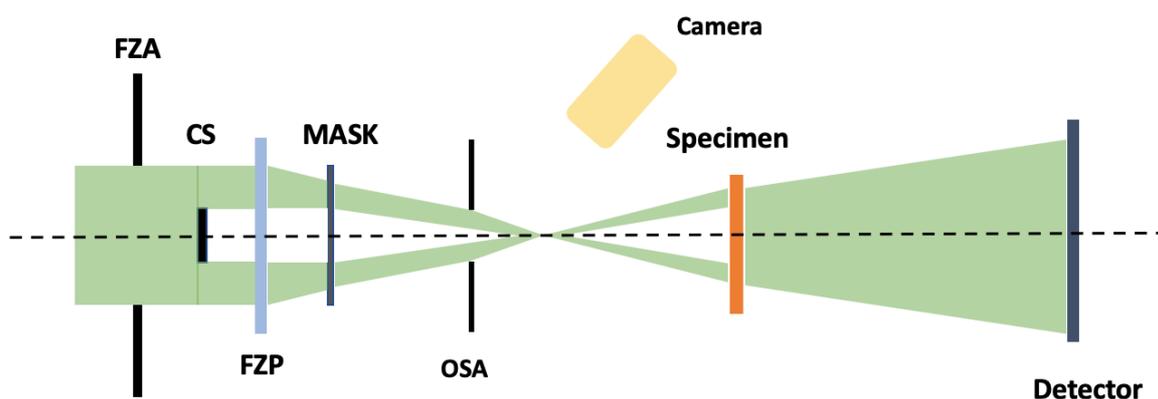


Figure 48. A schematic of zone-plate optics. In the zone-plate-optics system, there is a Fresnel zone aperture, central stop, Fresnel zone plate and an order-sorting aperture. A camera is applied to correct the position of the zone-plate-optics components and a detector at far end.

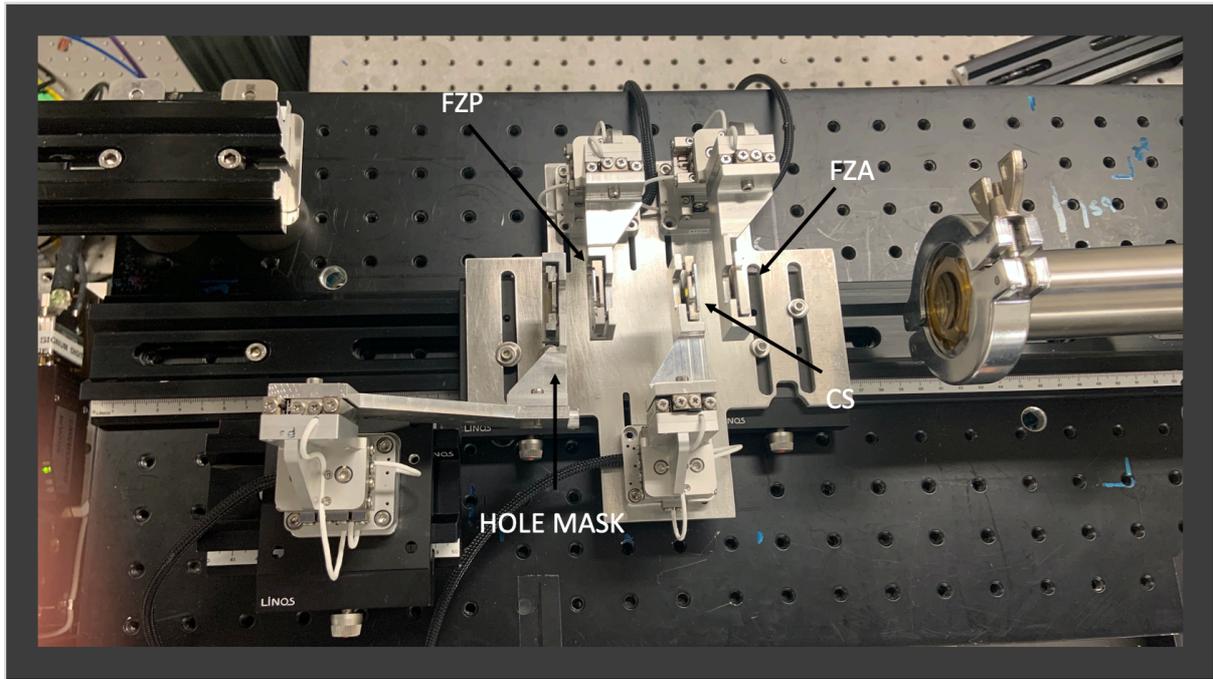


Figure 49. The real experiment's zone-plate-optics system setup.

The source size of beamline I13-1 is 302.7 and 3.6 micrometres horizontally and vertically, as shown in figure 50, respectively, with divergences of 18.5 and 2.2 urad, where the shape of the source size is determined by the optics of the main ring. In this experiment, we used an energy of 9.6keV. The FZP has a diameter of 400 microns and a 150-nanometre outer zone width, which gives a focal length of 47 centimetres at a voltage 9.6KeV. The centre stop size of the FZP system has a 60-micron diameter. Before the experiment, we prepared some Tungsten hole foils of 10 by 10 millimetres and drilled holes of different sizes into each foil, as shown in Table 2. We used a 5-micron-diameter OSA and a Siemens star-test specimen, fixed 16 millimetres downstream of the OSA. The detector we used is the Excalibur photo-counting x-ray detector, which was 14.55 metres downstream of the FZP optics system.

Table 2. The mask characteristics of the experiment

Hole matrix	Thickness ( $\mu\text{m}$ )	Diameter ( $\mu\text{m}$ )	Hole pitch ( $\mu\text{m}$ )
5X5	200	50	100
10X10	100	20	40
20X20	100	10	20
40X40	100	5	10

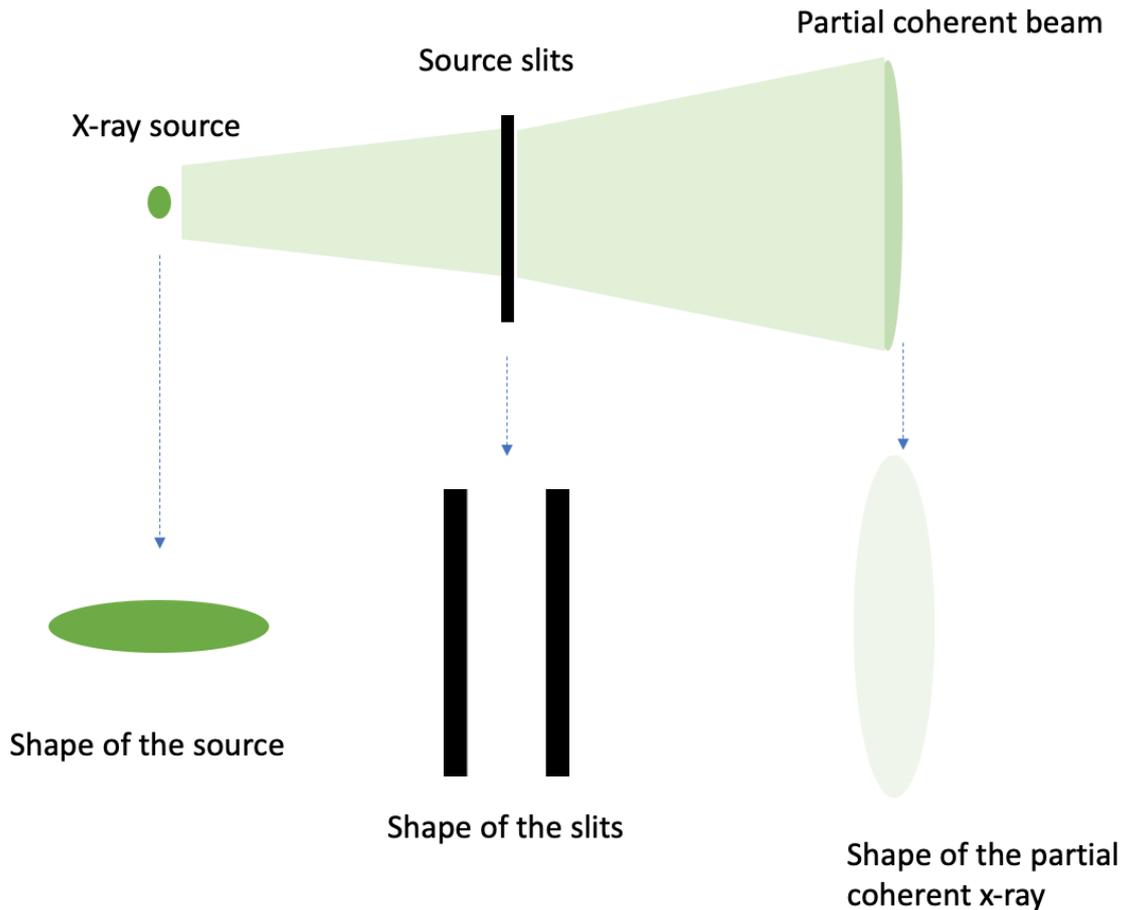


Figure 50. X-ray source, source slits, and the partially coherent beam. Due to the shape of x-ray source is oval, we use the gap between to source slits to control the coherence level of the x-ray beam. Smaller slits gap gives better coherence level of the beam.

### 5.7.3. Ideas to prove with experiments

There are some ideas we would like to prove in this experiment. The main goal is to prove that small apertures can provide better reconstruction quality than large apertures in partially coherent situations and the same flux through the specimen. This is because if the coherence width of the beam is about the size of each small aperture, the illuminations that pass each aperture are roughly coherent, when the beam illuminated to the aperture mask. On the other hand, the coherence width of the beam is much smaller than the size of the large aperture, which means the illumination that has passed the large aperture will not as coherent as in the case of the small aperture.

The second idea is to check if we can speed up ptychography reconstruction. As described above, if the illumination passes through the small aperture, it is relatively

coherent; the recorded data in the small-aperture experiment has a much higher known-to-unknown ratio. This means there is a much lower redundancy requirement to work out the reconstruction, so a good reconstruction can be achieved at a low-level overlap, which helps speed up the data collection of the ptychographic scan. Also, as the illuminations that pass through each small aperture are self-coherent but incoherent with others, we can use only a few modes to achieve good reconstruction, which reduces calculation time.

The last idea we want to prove is that when there are at least two orthogonal modes, each mode can be separated individually. We would like to find a relationship between coherence width and number of modes. In theory, at a small source-slits width, x-ray is coherent, while due to its self-correction ability, ePIE can reduce the weak, partial coherent impact; therefore, no more modes are necessary. However, if the source-slit widths are too wide, the reconstruction cannot be well-represented by using only ePIE. We must add more modes to the calculation to reduce the negative effect of the partially coherence.

#### *5.7.4. The experiment processes*

##### *5.7.4.1. Experiments using hole masks*

The first experiment we did was using the 50-micron aperture set to test our proposition. Before collecting data, we adjusted the zone plate optics (ZPO). To begin with, we used a camera to find the aperture position and marked the position as a reference to correct the position of the ZPO components, as shown in Figure 51, in which the x and y coordinates give the specimen spatial position in microns; the number above the coordinates is the experiment number, and the blue line is the mark we made as a reference. We put the OSA under the camera, and we used the camera to find the OSA's position and move it to the reference point. Once we found the position to be correct, we moved the OSA out of the line of the camera and put the FZP and the CS in the correct positions. When we centred all of the elements of the FZP, we moved the camera away and switched the detector on to collect data and adjust the specimen to the desired experimental defocus position. Then, we conducted

a ptychography scan with 16 by 16 scan positions and the step size was fixed to 1 micron with an under-0.1-second exposure time and a 16-millimetre defocus. The reconstruction results are shown in Figure 52, where the left-hand side is the probe reconstruction, and the right-hand side is the specimen phase result. From this result, we can see the strong interference between any two adjoint holes, as the illumination is diffracted at the boundary of each hole. Also, we can achieve reconstruction by using the aperture mask and approach a 0.2-micro resolution.

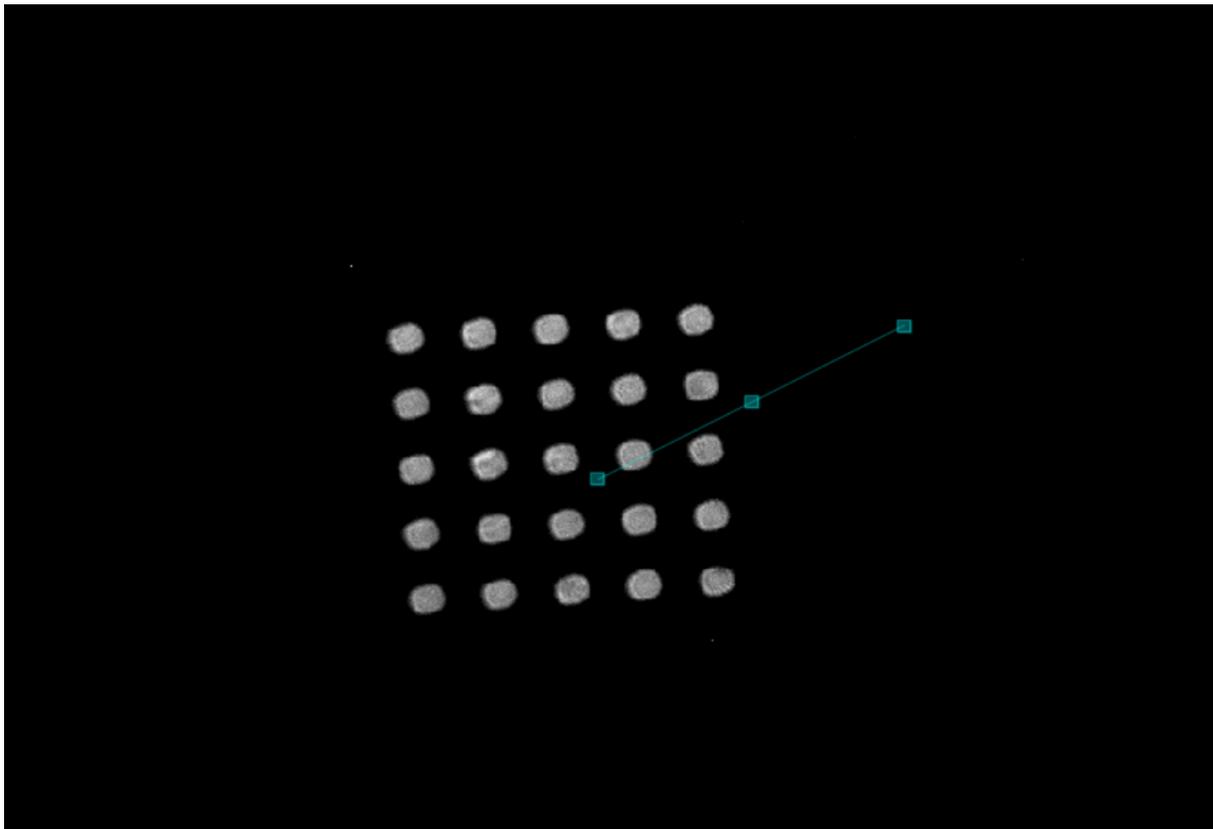


Figure 51. The aperture mask and reference mark in front of the camera. The reference mark is used to find the correct ZPO-component positions.

Then, we change the defocus and scan positions to change the size of the probe function on the specimen and the number of diffraction patterns. In Figure 53 and 54, we show the reconstruction results of the experiment with a defocus value of 6 and 16 millimetres under a 0,01-second exposure time. When we change the probe defocus, we change the overlap levels, as the large defocus gives more overlap. In Figure 53, the overlap level is too small to record enough information to solve the object and probe functions; therefore, the reconstruction quality is quite low and the field of view smaller. However, when we increase the defocus, as shown in Figure 54, any two,

joint scan positions have a large overlap and record more data, which gives better reconstruction quality than in Figure 53.

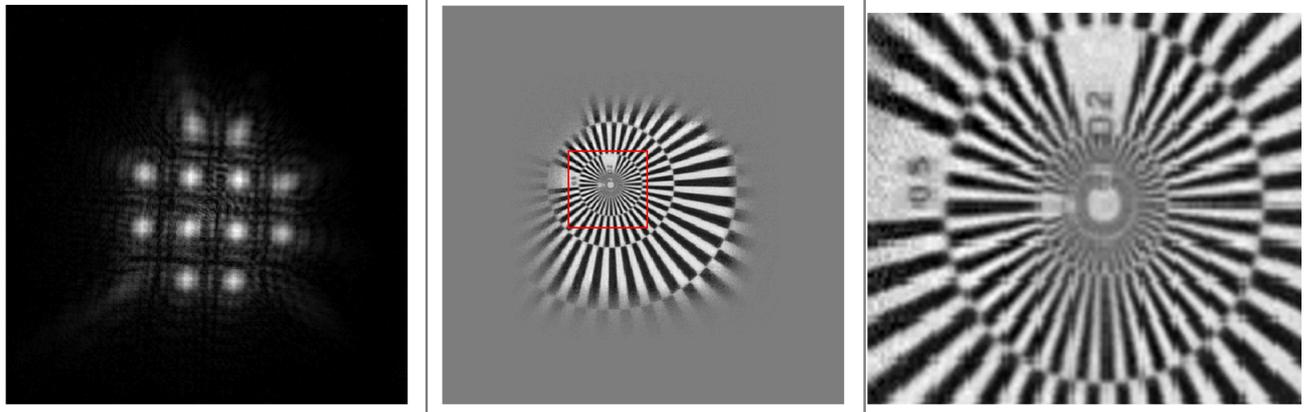


Figure 52. The object and probe reconstruction of our first experiment in beamline I13. In the probe function, we can see strong inference between any two adjoint holes, and we can approach a resolution of 0.2 micro. The crosses on the right image are reconstruction result, because of the interference between the probe functions passed the apertures, as shown on the left image.

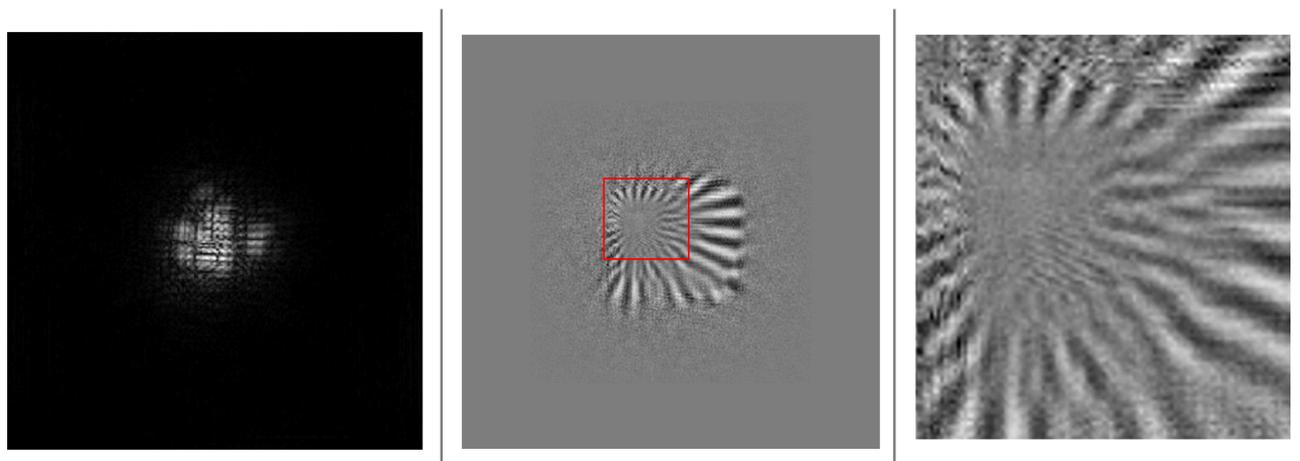


Figure 53. Reconstruction at a 6-millimetre defocus, 1-micon step size and 0.01-second exposure time. These results show that the defocus is too small, so the step size is too large to collect enough information to solve the object and probe functions and provide a large field of view.

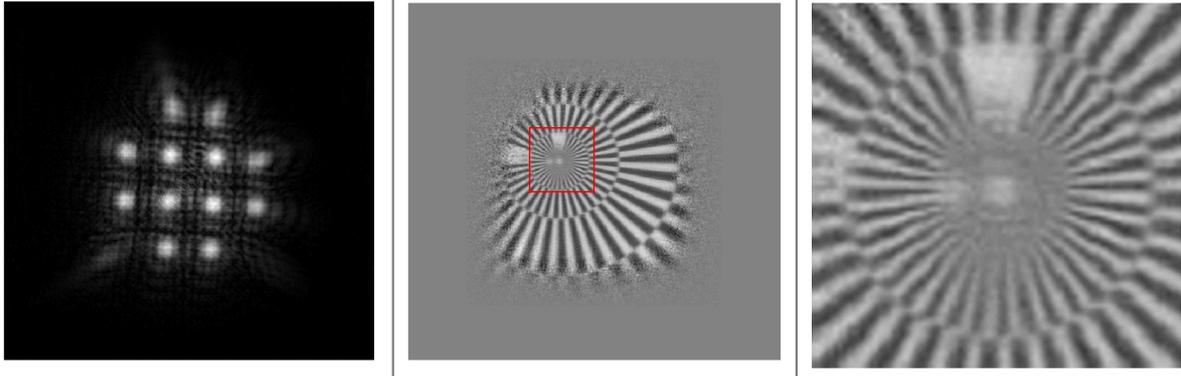


Figure 54. Reconstruction at a 16-millimetre defocus, 1-micron step size and 0.01-second exposure time. These results show that with a longer defocus, the probe function gradually becomes clear, and more features of the reconstruction are visible because with the defocus increase, any adjoining positions have more overlap, and more information are recorded.

In Figures 55 and 56, we use a 16-millimetre defocus and 0.1-second exposure time but change the step size from 5 to 10 microns. The theory is as same as with the defocus: changing the step size affect the overlap levels, and a larger step size results in smaller overlap levels. Therefore, the reconstruction in Figure 55 is better than that in Figure 56, as it has smaller step size and larger overlap.

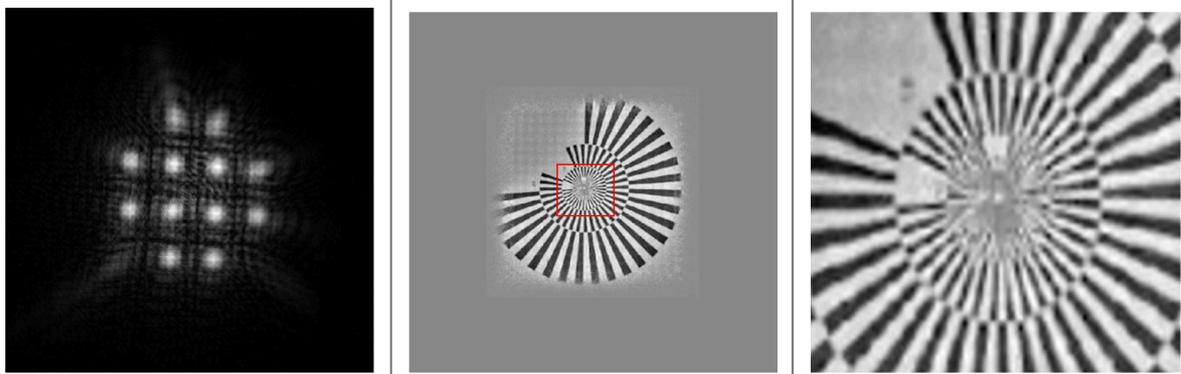


Figure 55. Reconstruction at a 16-millimetre defocus, 0.1-second exposure time and 5-micron step size. When we increase the step size, we reduce the known and unknown ratio of the structure of the specimen and illumination functions by reducing the overlap levels.

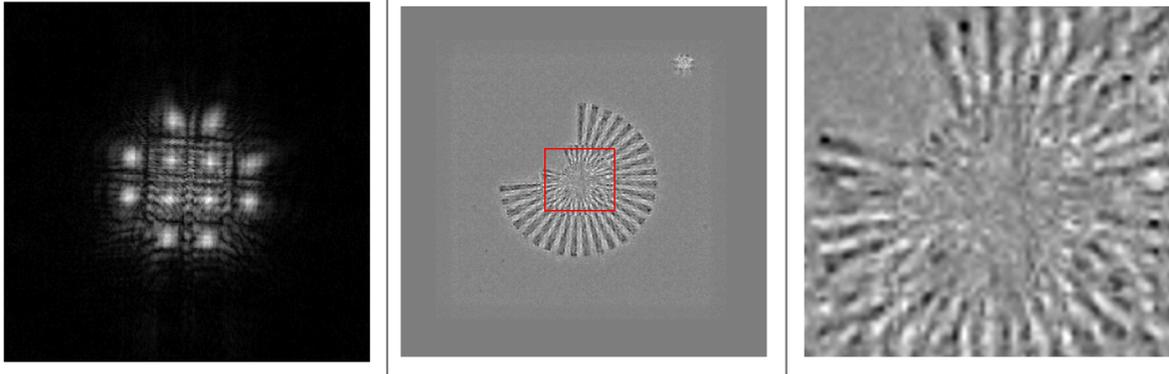


Figure 56. Reconstruction at a 16-millimetre defocus, 0.1-second exposure time and 10-micron step size. When we further increase the step size, the redundancy in the data is not enough to provide good reconstruction.

Then, we moved the aperture 178 millimetres downstream, and there are only four holes under the illumination. We did a test experiment to check the performance, where we used a 16-by-16-strip grid with a step size of 1 micron, an under-0.1-second exposure time and a source-slit size of 40 microns. The results are in Figure 57 – on the left is the probe reconstruction, and on the right, the object phase result.

We then did the first modal decomposition calculation, where the scan grid is 32 by 32 positions with a 1-micron step size and under-0.025-second exposure time, but the source-slits gap changed from 40 to 640 microns as we are going to test the results at different coherence levels. In Figures 58, the object-phase reconstruction is shown as well as modal decomposition results with 4 modes in a 640-micron slit-gap situation.

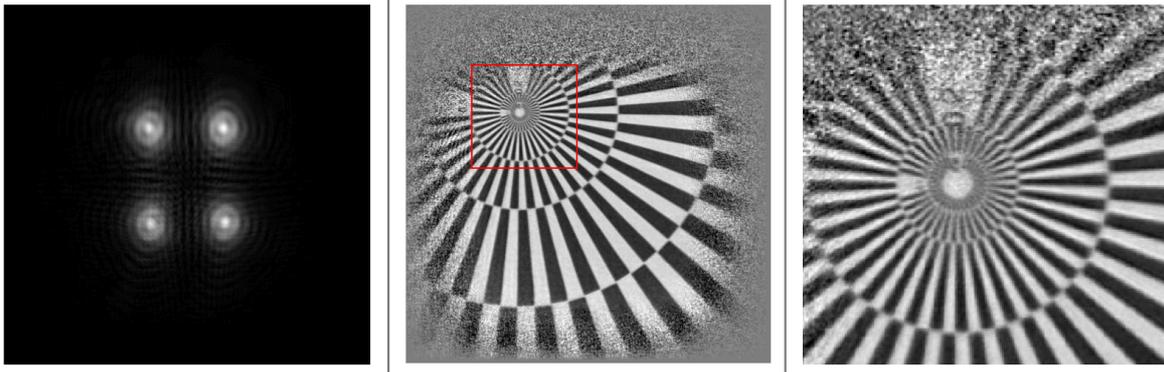


Figure 57. The first four-hole ptychography reconstruction. From the result, we can see that the reconstruction is much clearer, so the setup is acceptable for this experiment, and we will do the other experiments at this defocus.

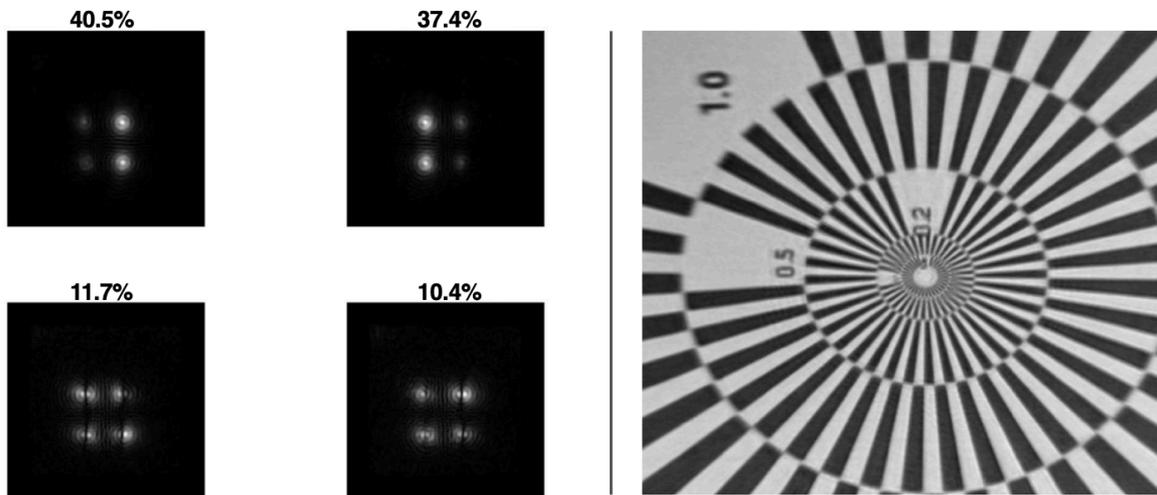


Figure 58. The modal decomposition result and reconstructions at the source-size of 640 microns. We can see a clear structure and 0.2-micron resolution. Also, as the coherent width of the beam is much larger vertically than horizontally, the first two modes show that one of the columns is brighter than the other.

#### 5.7.4.2. Halfway discussion

Until now, we have correctly tested all the key experiment elements and they work well. The conclusions of Figures 53, 54, 55 and 56 are as we predicted, that more exposure time gives more flux, which gives better reconstruction quality. Also, larger step size and small defocus reduces reconstruction quality, which is due to the step size being larger and the overlap level smaller, so the amount of information on the

ratio of known and unknown real space reduces and there is not enough to give a unique solution.

When we used four holes to do ptychography reconstruction and added four modes to the modal decomposition calculation, there was strong interference between any two joint small probes in the reconstructions in Figures 57 and 58, which is the same as our four-corner simulation results, so we cannot clearly separate each small probe mode. However, because the beam coherence shape is oval instead of round, the modal decomposition results show two strong and weak vertical-direction small probe combinations, which give us a hint that we can possibly replace our aperture mask with a semicircle mask to fit the illumination shape. With this idea, we designed a second experiment.

#### *5.7.4.3. Experiments using wires*

Because the source size is a horizontal oval, we changed our multi-probe experiment to use wire experiments, where we are going to put a 100-micron-diameter gold wire behind the FZP instead of the apertures so that the gold wire divides the FZP into two small semicircles. We designed two types of wire experiments. In the first experiment, we put a wire vertically in the direction of the coherence width; in the second, we put a wire horizontally across the direction of the coherence width.

Because the vertical wire is in the same direction as the coherence width's direction, it is possible to isolate two coherent modes. To test this idea, we did two modal decomposition tests in a 50- and 640-micron slit-gap situation, where the scan grid is 32 by 32 with a step size of 1 micron and an under-0.25-second exposure time, and the modal decomposition results are shown in Figure 59. The left side of the figure gives the modal decomposition result at the 50-micron slit gap, where the beam coherence width in the horizontal direction is much larger than the largest width of the two semicircles. Due to the high coherence, we find that we can almost clearly isolate two coherent modes, even if there is still interference, where the third and fourth modes provide interference and noise patterns. When we enlarge the source-slit size to 640 microns, the beam becomes less coherent; the modal decomposition results are

shown on the right side of Figure 59. In this test, because the coherent width of the beam is smaller than the largest width of the two semicircles, the illumination passing through each semicircle becomes less coherent. Therefore, except for the interfere, there are more than one illumination states in each semicircle.

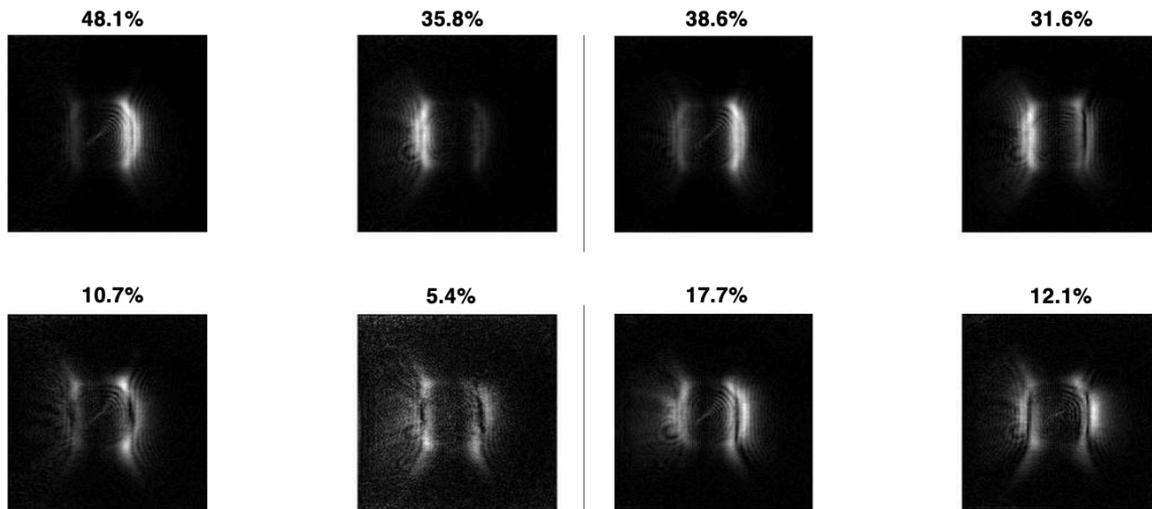


Figure 59. The modal decomposition of the lens in the vertical wire experiment at source slits of 50 and 640 microns. The left-hand side of the Figure the is the result of the slits being at 50 microns, and on the right-hand side at 640. Note the percentage of power in each mode. The results prove we can separate two coherent illumination modes by using a vertical wire. Also, if the coherence level is too low, there are more illumination states in each semicircle.

It seems that the wire experiment works as we proposed, so we are going to use the wire to replace the aperture mask and test our initial targets. Except for the horizontal and vertical wires, we used the FZP-only experiment as a reference experiment for testing the result. In Figure 60, a diffraction pattern model of each experiment we collected from the synchrotron is shown to provide a general understanding of how the wire is placed and separates the illumination. In each diffraction pattern, except for the wire-blocked areas. (The lines crossing the images are gaps on the detector, which is made from several chips.)

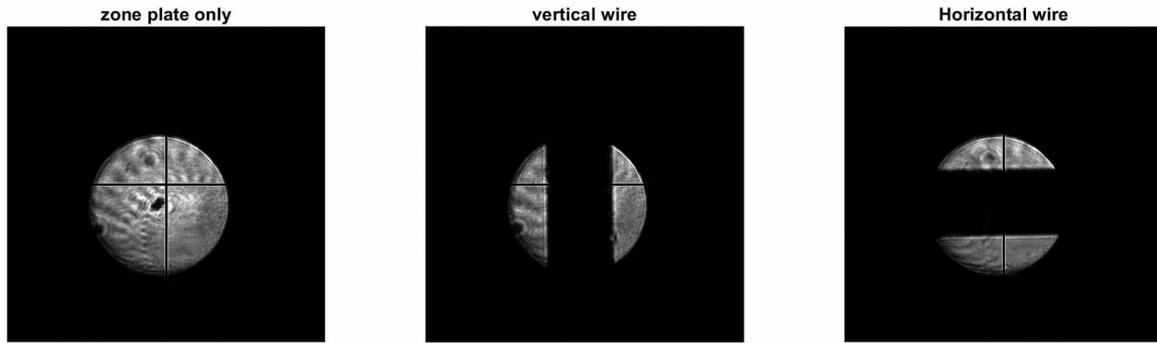


Figure 60. A real diffraction-pattern sample of the zone-plate-only experiment, vertical wire experiment and horizontal wire experiment. The light lines on each diffraction pattern are the chip gaps on the detector. The radial lines are the shadow image of the specimen.

So, we are going to collect the data of the FZP-only, vertical wire and horizontal wire experiment data, where we use a 32 by 32 scan grid with a 0.5-micron step size and an under-0,025-second exposure time, but the slits-gap changes from 20 to 640 microns with a slit-step size of 5 microns. In this experiment, we are going to show that if the vertical wire is in the same direction as the coherence width, the illumination passing through the vertical-wire lens is more coherent than that which passes through the horizontal wire and lens. Therefore, we can use less modes to achieve good reconstruction.

Because the illumination on the specimen is more coherent in the vertical wire experiment, the data has a much higher known-to-unknown ratio, which means it has a much lower redundancy requirement than others. This may help speed up the ptychography experiment.

#### 5.7.4.4. *The experiment results and discussion*

We are going to demonstrate how the modes affect the phase-retrieval process with different number of modes used in the calculation of the first experiment. Here, we are going to use an ePIE algorithm to process the single-wire lens only – vertical and horizontal results with 2 and 10 modes in the 640-micron slit-gap situation – which are the most incoherent results we can collect, by 500 iterations. Figure 61 shows that

while the first row provides the reconstruction of the horizontal wire, vertical wire and lens with 10 modes, the second row provides the reconstruction with 2. When we compare the reconstruction results with 10 modes, the reconstruction qualities are quite close, but the lens result shows some uneven background features, which is because, at this partially coherent degree, 10 modes cannot fully overcome the partial coherence in the reconstruction. However, for the wire experiment, the ratio between the beam coherent width and the aperture size is larger than in the lens-only experiment. Therefore, in general, the degree of coherence in the horizontal and vertical wire experiments is higher than in the lens only experiment, and they collected more information to solve the object function at a high quality. However, comparing the 2 modes results in 10 mode results; the vertical-wire results retain good quality, but the horizontal and lens-only results worsen.

Because the vertical-wire-separated two parts are relatively coherent, it is not necessary to use lots of modes to solve the partially coherence; also, as the beam is more coherent on the specimen in the vertical-wire experiment than in the others, it records lots of useful information as redundancy to contribute to the reconstruction. In the lens experiment, there are not enough modes to solve the partial coherence, so the reconstruction worsens. The horizontal-wire results are a special situation, where because the wire in the experiment crosses the coherence width, lots of coherent illumination is blocked by it and the rest of the illumination is incoherent.

The second experiment is about whether the vertical wire can achieve good reconstruction in a situation with less overlap. As discussed above, because the illumination in the vertical-wire experiment had more coherence than in the other two, it recorded more useful information to contribute to reconstruction. In this experiment, we use the same data with 10 modes but reduce the overlap level between any two-joint scan positions to one-fourth. The results are shown in Figure 62. Comparing the vertical-wire results with the others, it shows a significant difference. The vertical wire retains acceptable reconstruction quality, but the others do not.

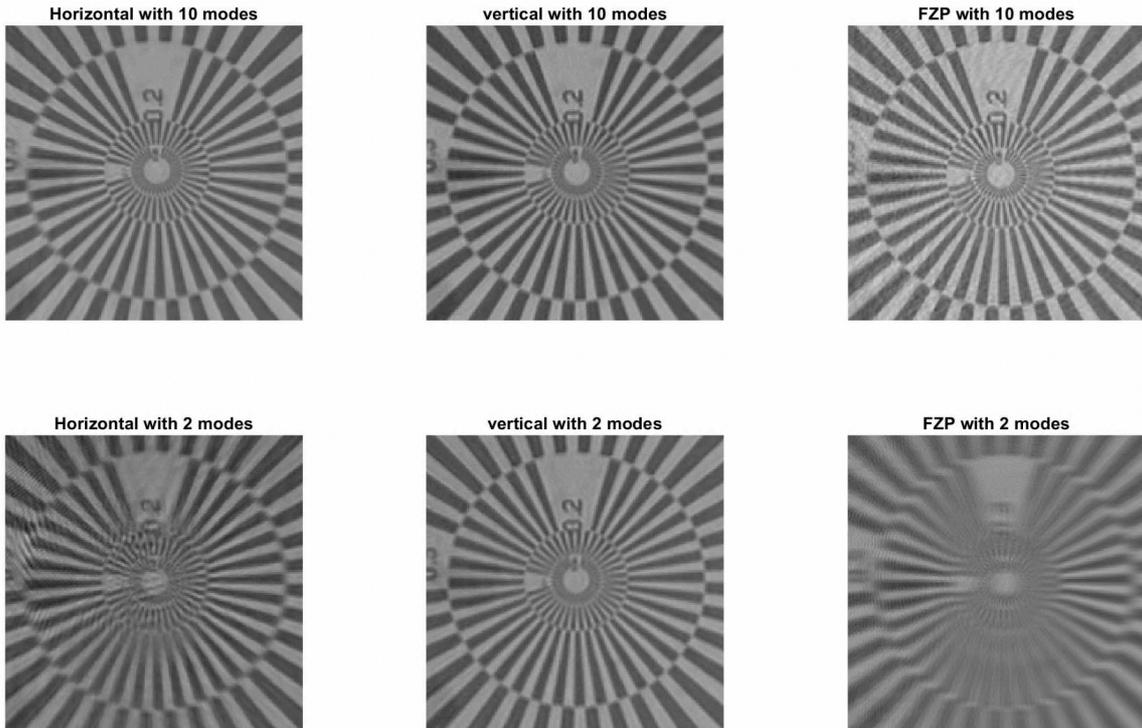


Figure 61. Reconstruction of the horizontal-wire, vertical-wire and FZP-only experiments with 2 and 10 modes. Because the illumination in the vertical-wire experiment has more coherence than other two, it records more useful information to contribute to reconstruction and does not need many modes to solve the partially coherence. For the other two experiments, because the illumination is not highly coherent, many modes are necessary to retain good reconstruction quality. So, if we reduce the number of modes in the calculation, they do not have enough ability to solve the partial coherence, and the reconstruction worsens.

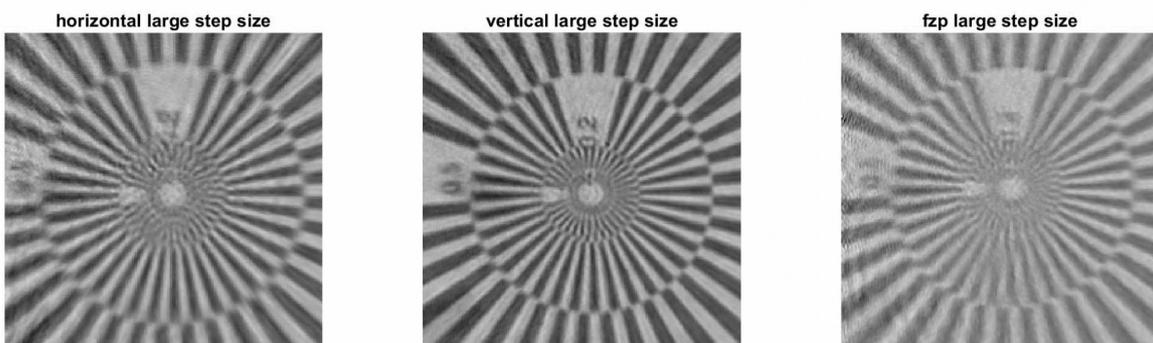


Figure 62. Reconstruction of the horizontal-wire, vertical-wire and FZP experiments with 10 modes in a lower-overlap situation. From the result, because the coherent beam in the vertical-wire experiment recorded more information, although we reduced the overlap, it still has enough data to solve the object function.

## 5.5. Conclusion

In this chapter, we introduced our experiment process, experiment setup, experiments, and related results. We compared the vertical-wire, horizontal-wire and lens-only reconstruction results at a low, partially coherent level. Also, we reduced the overlap level of the three different experiment setups at the same partially coherent level. From the results, we find that by adding a wire in the direction of coherence, we can separate a large, partially coherent probe function into two, separate, relatively coherent probe functions, which can help efficiently use the coherent part of the partially coherent beam. Furthermore, as the probe's coherence level was improved with the wire, a large number of modes are not necessary. This way, only a few modes contributed to the calculation, so reducing the number of modes has a negligible influence on reconstruction. Also, if we reduce the degree of overlap, we reduce the known and unknown real-space-information ratio of the data, and it does not have enough bandwidth to calculate with many modes. In this situation, because the vertical wire improves the probes' overall degree of coherence, only a small number of modes need to be solved, which requires less data. This result shows us that we can speed up the ptychography experiment by engineering mode patterns.

# **Chapter 6 Damaged specimen reconstruction by ptychography**

This chapter provides a general knowledge about radiation damage in TEM experiments. It also demonstrates a proposed method, based on mixed state ptychography, to improve the image reconstruction of a damaging specimen. The focus is on exploring the use of ptychography, for both the scan and reconstruction processes, to determine whether the effect of radiation damage can be minimised in the data recorded. Pixel information relating to the specimen was recorded several times in the ptychography experiment to test if large amounts of diverse information in the multiple imaging collection system could self-calibrate any errors in the dataset and correct any damaged pixels at reconstruction.

## **6.1. The interaction between electrons and specimen atoms**

As a particle, an electron beam can approach minute de Broglie wavelengths (of the order of  $10^{-9}$  m). This has been used in TEM to provide higher resolutions than electromagnetic waves to boost research in a variety of fields, such as chemistry, biological science, materials science, and pharmacy. While TEM can assist researchers to obtain an atomic resolution, a specimen can experience radiation damage when it has been over exposed in TEM, including a surface or inner structure change on the specimen. Since radiation damage negatively affects the reconstruction of the specimen, the chance of damage is reduced by using effective coping methods.

There are four types of interactions between electrons and specimens: non-interaction, elastic scattering, inelastic scattering, and absorption. Non-interaction is the simplest case as the electrons pass through a gap in a specimen directly, without any interactions.

Elastic interaction is important in TEM when forming the diffraction pattern and image. Elastic scattering occurs when the incident beam passes through the electron cloud of a specimen atom. Because of the coulomb force of the nucleus and orbital electrons, the incident electrons will change direction and get closer to the nucleus or orbital electrons, increasing the angle. Hall [119] provided two equations to calculate the scattering field radius of the nucleus,  $R_n$ , and the cloud electrons,  $R_e$ , being

$$R_e = \frac{e}{V\phi}, \quad (93)$$

and

$$R_n = \frac{Ne}{V\phi}, \quad (94)$$

where  $e$  is the electron charge,  $V$  is the acceleration voltage of the incident beam,  $\phi$  is the scattered angle of the incident beam, and  $N$  is the number of cloud electrons. Goodhew et al. [120] points out that the probability of an incident beam being scattered through a certain angle is

$$p(\phi) \propto \frac{1}{E^2 \sin(\phi)^4}, \quad (95)$$

in which  $E$  is the kernel energy of the incident beam. The above equations show that the incident beam is more likely to be scattered at a small angle, and an electron with more energy has less chance of being scattered at a large angle.

Inelastic scattering arises from the interaction between an incident beam, atom nucleus and orbital electrons. Unlike elastic scattering, energy is lost and transferred during the interaction, representing both heat and other useful signals, including x-ray fluorescence, Auger electrons or secondary electrons that could contribute to other research. Since inelastic scattering is the reason for radiation damage, this is further discussed below.

Finally, during the interaction between an incident beam, atom nucleus and orbital electrons in a thick specimen, the incident beam electron can lose all its energy and be 'absorbed' by the specimen. Absorption, however, is rare because electron energy in a TEM will exceed 100keV.

## 6.2. Type of radiation damage in TEM

Radiation damage is a significant research area in the physics, materials science, biological science, and chemistry fields. It is mainly caused by inelastic scattering, which can lead to surface or structure changes of a specimen in a TEM experiment. Generally, radiation damage can be attributed to heating, atomic displacement, and radiolysis.

Heating is the most common cause in TEM experiments because the energy lost during collision between the incident beam electrons and atoms are converted to heat and emitted to the environment. The 'generated' heat is described as [121]

$$T = \bar{E}_l \frac{2R}{d(4\pi k\lambda)}, \quad (96)$$

where  $\bar{E}_l$  is the average energy lost during inelastic scattering of a single incident electron,  $R$  is the radial distance,  $k$  is the specimen thermal conductivity and  $\lambda$  is the interacted electron mean-free path. The heating could be affected by several variables, such as the conductivity of the specimen, the energy, diameter and current in the beam, and/or the specimen surface conditions.

Another common result of radiation damage is atomic displacement caused by inelastic scattering. As the elastic scattering process conserves energy and momentum, deflecting an electron in an atomic nucleus field must transfers some energy to the nucleus. According to Egerton and Malac [9,122] the transferred energy can be calculated as

$$E_t = E_{max} \sin \left( \frac{\theta}{2} \right)^2, \quad (97)$$

and

$$E_{max} = \frac{E(1.02 + \frac{E}{10^6})}{(465.7A)}, \quad (98)$$

in which  $E_t$  is the transferred energy in eV,  $E_{max}$  is the largest amount of transferred energy when the deflected angle is 180 degrees and the electron back is scattered,  $E$  is the kernel energy of the incident beam and  $A$  is the specimen atomic mass number. Banhart [123] gives another equation to calculate the maximum transfer energy as

$$E_{max} = \frac{2E(E + 2mc^2)}{Mc^2}, \quad (99)$$

where  $m$  and  $M$  are the mass of electrons and atoms, respectively. Therefore, if the transferred energy is more than the intermolecular force between specimen atoms, what remains is the specimen structure. The nature stable can be determined by [124]

$$E_d = \frac{\left( \frac{100 + AE_m}{5} \right)^{0.5} - 10}{20}, \quad (100)$$

where  $E_m$  is the displacement energy of the specimen. The specimen atom nucleus will be removed from its position, affecting the structure the specimen. Atomic displacement is called knock-on damage. Knock on damage from the surface is called sputtering, where the detached atom nucleus at the surface of the specimen flies out to surround the space. Williams and Carter [125] and Medlin and Howitt [126] indicated that a crater will be caused if more than half of the surface atoms sputter out towards the surface. Compared with other materials, a crystalline structure specimen is more likely to sputter [127,128].

There are other radiation damage results, such as radiolysis, in which the chemical bonds of specimen atoms are broken by inelastic interactions. Water radiolysis is a

basic radiolysis model where water atoms are separated into hydrogen and oxyhydroxides under the ionizing radiation [129,130,131]. Electrostatic charging is caused by inelastic scattering. The beams interact with the specimens and create multiple secondary electrons that can either be kept in the specimen at high-kernel energy incident beams or escape to the surrounding area when the beam-kernel energy is at a low level and must be rebalanced. If the specimen has been charged, there will be an inner electric field inside the specimen that could destroy it.

### 6.3. Method used to minimise radiation damage

Before discussing the strategies to minimise radiation damage, the electron dose must be considered. The electron dose describes the number of electrons per unit area on the specimen. The dose rate of the specimen is crucial to the probability of radiation damage and the imaging results. According to Egerton [132,133] pointed out the dose-limited resolution to maintain balance and control radiation damage. If the critical dose specimen can tolerant before damage is  $D$ , and

$$D = \frac{It}{d^2}, \quad (101)$$

the maximum recorded electrons should be

$$N = \frac{FD r^2}{e}, \quad (102)$$

and the dose-limited resolution is described as

$$r = \sqrt{N \frac{FD}{e}} = SNR(DQE)^{-0.5} C^{-1} \left( \frac{FD}{e} \right)^{-0.5}, \quad (103)$$

where  $I$  is the beam current,  $t$  is the irradiating time of the  $d$  diameter beam on the specimen,  $F$  is the single-collection device's efficiency of a TEM,  $DQE$  is a

measurement of the efficiency of the detector,  $C$  is the cross contrast of the unit area, and  $SNR$  is the record signal to noise ratio, the latter of which is represented as

$$SNR = \sqrt{DQE} \frac{|\Delta N|}{\sqrt{N}} = \sqrt{DQE} \frac{CN}{\sqrt{N}} = C\sqrt{(DQE)N}, \quad (104)$$

where an  $SNR$  value larger than 5 is sufficient for purposes of this experiment [134].

The most common strategy to minimise radiation damage is cooling the specimen, which is useful for materials that are sensitive to heat. By cooling the specimen, the critical dose can be increased to improve the resolution in TEM experiments [121].

Next, changes need to be made on the incident beam. First, the most direct way of reducing the energy of an incident beam was proven by Bradley and Zaluzec [135], who stated that if the incident-beam energy is less than twice the damage requirement, it reduces the probability of radiation damage. Second, the incident-beam current can be reduced because the probability of damage depends on the dose rate and accumulation. Salisbury et al. [136] and Cazaux [137] pointed out that if the beam current is smaller than the damaged intensity requirement of the specimen, radiation damage can be avoided. In addition, a low-beam current will reduce the chance of radiation radiolysis. Some results indicate that the damage probability of a high-dose-rate experiment can be reduced within a short recording time [138]. However, reducing the incident-beam current has some disadvantages, including longer recording times for experiments that may increase the risk of instability affecting the image. Lastly, the beam diameter can be reduced. While this has been explored in previous work [139,140], the mechanism is not yet clear.

Changes can also be made to the specimen. Coating is widely used to protect the specimen. In the TEM experiment the specimen is thin, and sputtering becomes the main result of the knock-on damage. Some results show that the damage can be significantly reduced by coating the specimen in carbon or carbon nanotubes [141]. Furthermore, coating some non-conductive materials can reduce the effect of electrostatic charging by increasing the critical dose value, which can also be reduced by using SIO coating [142] or reducing the beam current. In addition to the above

methods to reduce the probability of radiation damage, low dose techniques are a rapidly developing technology, especially in the biotechnology and healthcare fields.

## 6.4. Simulation

The above introduction is relevant for all types of TEM experiments. The first step is to build a damage probability model, which compares the damage probability between using big or small probes. With that result, a probe is chosen to generate damaged data and use object or probe modes to minimise the effect of the damage on the reconstruction. To achieve this goal, the specimen is treated at different damage levels, allowing the possibility of using two or more modes to extract the object function in different states of damage.

### 6.4.1. The damage probability model

To simulate radiation damage, we can use a simple to calculate the no damage probability, where it just a model to simulate the damage probability. The assumption is that the number of damage events remains constant for  $T$  every second of exposure time. For each exposure,  $p$  and  $q$  are the probability of damage, non-damage, and the process, is memoryless, meaning that the probability  $p$  and  $q$  will not change for each exposure. If  $U(t)$  is the probability of non-damage after time  $t$ ,  $t = 0$  and  $U(t) = 1$ . If the time gap is  $\Delta t$ , when  $t = n\Delta t$ :

$$U(t) = q^t, \quad (105)$$

$$\int q^t dt = \frac{q^t}{\ln q} + c, \quad (106)$$

$$\Delta U = -q\Delta t U \quad \text{if } c = 0, \quad (107)$$

$$dU = -qUdt, \quad (108)$$

$$\int \frac{1}{U} dU = -q \int dt, \quad (109)$$

$$\ln U = -qt + c, \quad (110)$$

and

$$U = e^{-qt} \text{ if } c = 0. \quad (111)$$

The above will result in the probability of non-damage for exposure time  $t$ . The next step is to generate a random number between 0 and 1. If the random value is larger than the non-damage probability, the atom in a specimen will be damaged during the exposure. In contrast, if the random number is smaller than the non-damage probability, the specimen remains intact. In its subsequent exposure, the probability is calculated from resetting  $t = 0$ . Except for the exposure time, this model can also be used to analyse the flux and probe energy of each pixel.

#### 6.4.2. *The damage probability between big and small probes*

When assessing the effects of different probe sizes, larger sizes were expected to allow each atom the ability to provide more information before being damaged. This influences the best way of performing an electron-ptychography experiment, being with either a focused or defocused probe. Figure 63 illustrates a straightforward method of understanding scanning with two schematic diagrams using a big and small probe. As Figure 63 shows, for the same atom in a specimen, a larger probe means that the atom will contribute five times more diffraction patterns than the small aperture, provided that the total flux per atom is the same in both scan geometries.

The same setup was used in the simulation, with changes that defocused the degrees to change the size of the probe. The simulation setup is shown in Figure 64, indicating the incident beam, the objective lens, and the specimen, from left to right. Defocusing allowed for a larger probe size on the specimen.

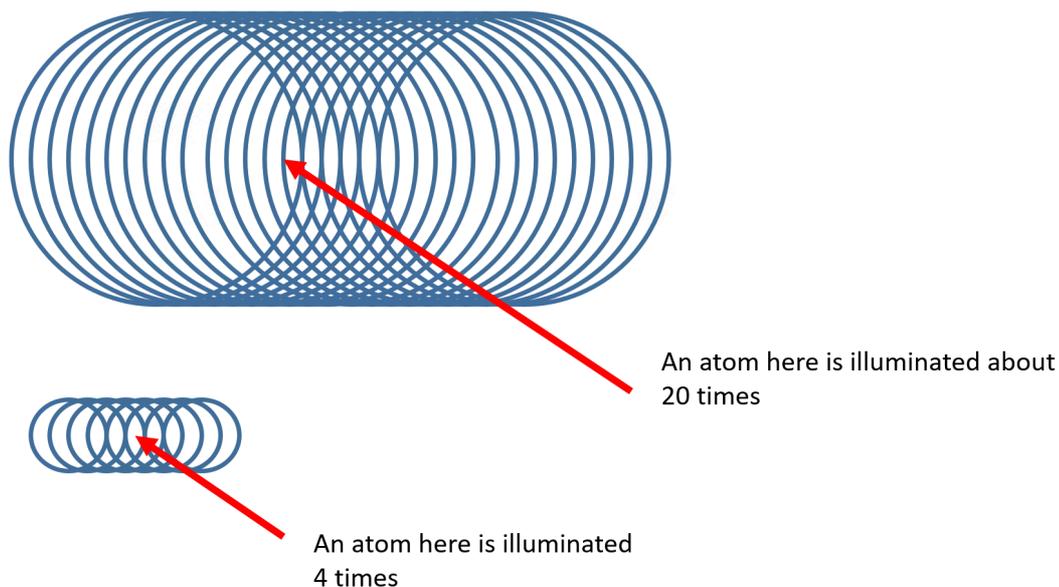


Figure 63: The scanning process illustrating the difference between big and small probes for the same atom. Using a larger probe provides more diffraction patterns that record the information of an atom before it is damaged, allowing for ptychography to solve for the damaged atom at the correct position.

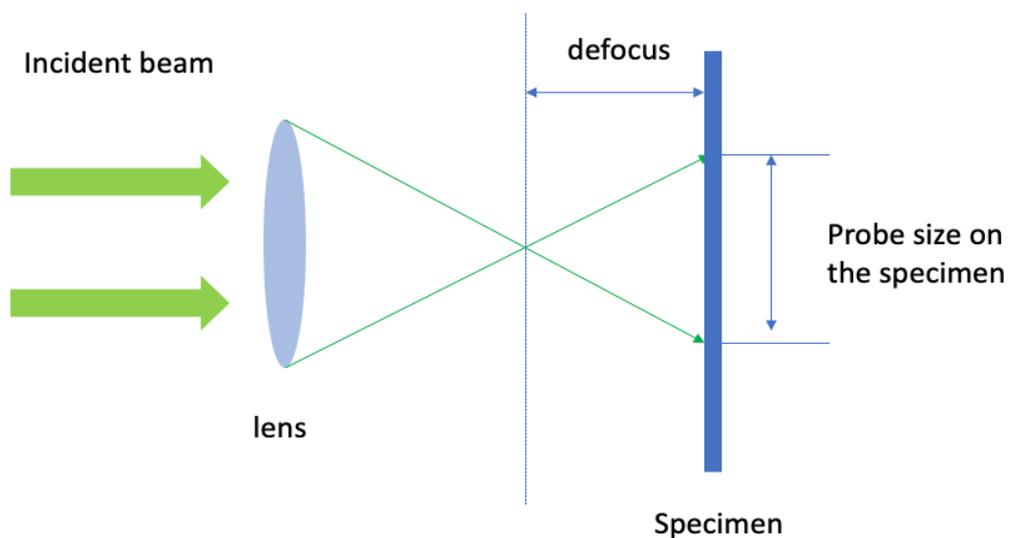


Figure 64: The setup of the damage-probability test. When an incident beam is focused by a lens, the focused illumination size can be enlarged or reduced by moving the specimen close to, or further away from, the back focal plane of the lens.

Because the probability of damage is measured by the number of electrons that hit the specimen, we assumed a total of  $N$  electrons in each scan. Figure 65 illustrates that when using two probes, where the left and right images show the effects of the small and large probes, respectively, the large probe has twice the amount of defocus than the small probe. Because of the rough relationship between defocus and probe radius on the specimen,

$$\tan(\theta) = \frac{\text{probe radius}}{\text{defocus}} \quad (112)$$

if the defocus is doubled, the probe radius would also approximately double, and the illumination area would become four times larger. The average electrons per atom would be about of  $1/4$  the average electrons per atom when using the small probe. Therefore, in each scan position, using a bigger aperture is more likely to reduce the probability of radiation damage, while also reducing the flux on each scan position atom.

After checking the average electron per atom, the next step is to check the number of accumulated electrons per pixel. This is done on the same scan step size to keep the total intensity in this experiment the same. Here, the total electron per scan is normalized to be 1, then the probe is scaled according to the probability of electron distribution. The result of the accumulated electrons per pixel of the big and small probe is shown in Figure 66, where the left and right images indicate the accumulated electron distribution of the small and large probes, respectively, while the total electrons per scan is normalized to be 1. The result clearly shows that when using a big aperture, each atom will receive fewer electrons, before exposure time.

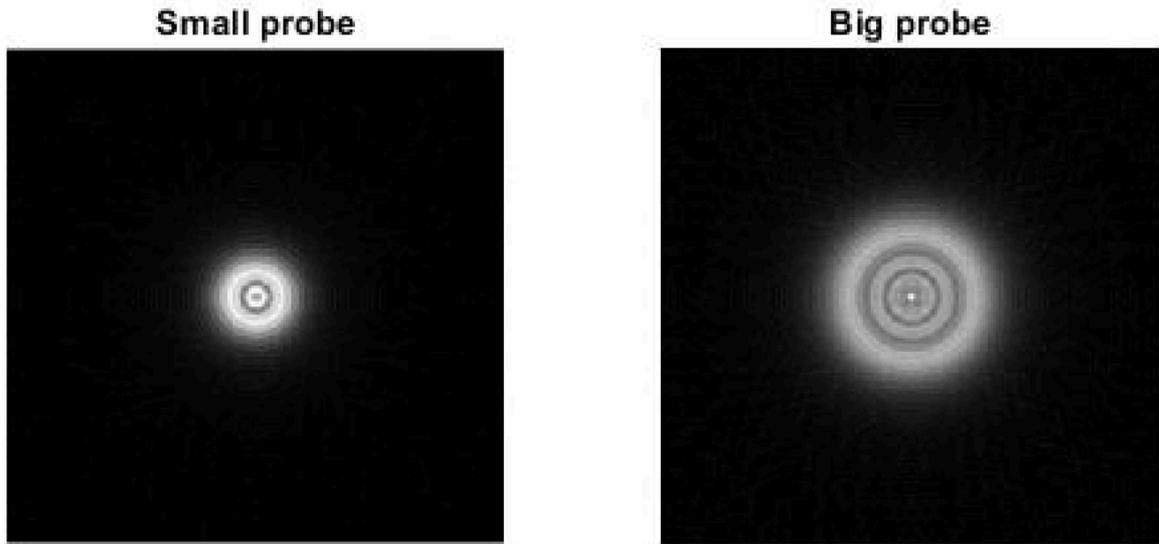


Figure 65: The big and small probe in this simulation. The big probe is built by a large defocus.

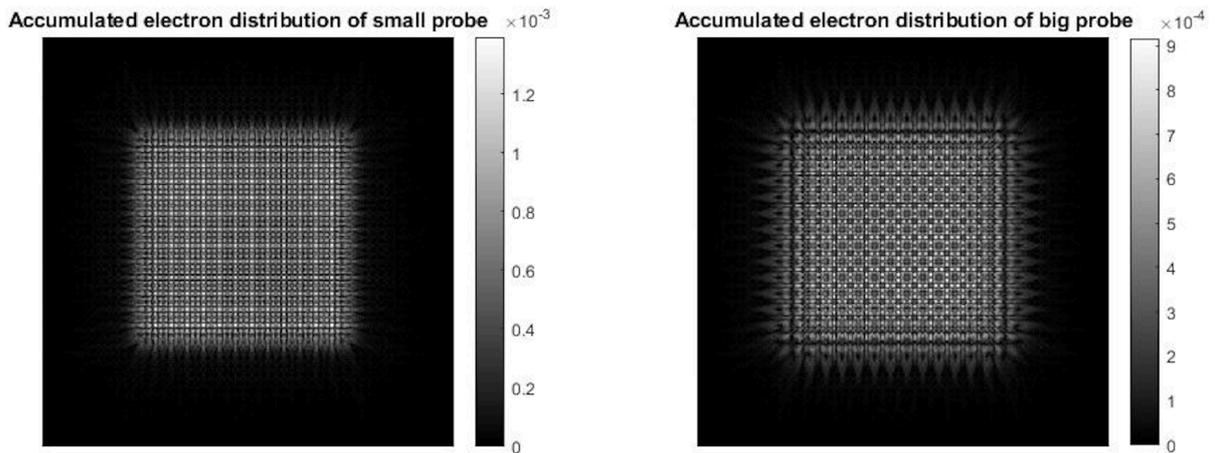


Figure 66: The accumulated electron distribution of the small and big probes in this simulation, where the total electron pre-scan position is 1. The grey bar indicates that, by using a big aperture, each atom will receive fewer electrons before exposure time.

Using the small probe, each pixel receives, on average, four times the number of electrons than when using the big probe. However, except for the pixels on the outer part of the scan, each pixel will be scanned more times when using the big probe. This results in more recorded diffraction patterns in the data for the pixel, which is important when proposing the below method to recover the damaged specimen.

## 6.5. Using mixed state ptychography to improve the reconstruction of the damaged specimen

Pixel information on the specimen has been recorded several times in this model ptychography experiment. Ptychography has many advantages, including the abundant and diverse information obtained using the multiple imaging collection system. This means that the ptychography can self-calibrate any errors in the dataset, making it able to correct the damaged pixel at reconstruction.

The first idea is that the damaged and undamaged specimens are treated as two individual specimen states, meaning that two object modes can be used to represent the damaged and undamaged specimen images as the data in the reconstruction. The specimen object image in the simulation is a phase image, being an object modulus as an all-in-one matrix, which is a simulated atomic image created by convolving a regular grid of dots with a gaussian filter and a standard deviation of 3 pixels. (Note that in this model, we are not putting in real atomic dimensions.) The specimen phase image is shown in Figure 67, where the phase changes within a range of 0 and 1.9.

During the experiment, the big probe function is used, with an approximate 20% overlap for any two joint probe positions. By using the damaged probability model, and assuming the radiation damage will happen, on average, after  $N$  electrons hit a pixel, the exposure time  $t$  is changed to a number of electrons that is greater than  $N$  to calculate the undamaged probability after removing any damaged pixels on the atom. In this simulation,  $q$  is equal to 0.15, which give a high damage rate and the specimen will be damaged quickly in a reasonable level of damage in the simulation setup. By generating a random number between 0 and 1, with a random value higher than the non-damage probability, the specimen in the test will be damaged during scanning. If the random number is smaller than the non-damage probability, the specimen remains, and the next position is scanned.

simulated phase image

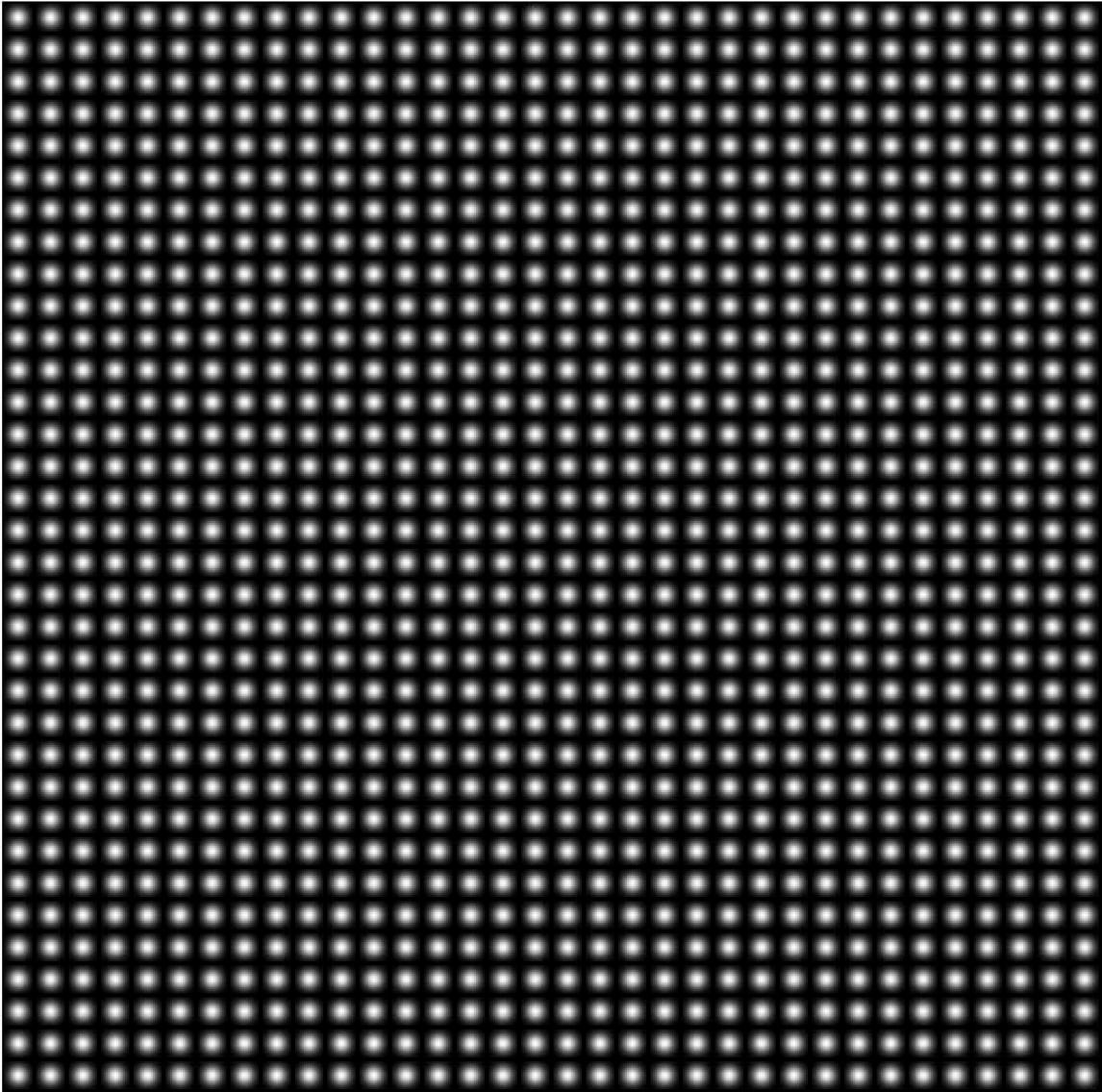


Figure 67: The simulated atom specimen phase image

#### 6.5.1. *Simulation without noise*

The first experiment tests the damage data reconstruction using two object modes, where the damaged specimen is treated as another state of the specimen. If this holds true, two results will be obtained: one with some or less damage, and another with all the damaged atom positions, the latter of which is related to the number of diffraction patterns recorded with valid information about the missing atom. For example, if an atom has been kicked out at the first scan, there are no diffraction patterns with

information about the atom. In Figure 68, the object phase image illustrates the before- and after-radiation damage, where the empty holes are the atoms that have been knocked out from the specimen. All the damage appears in the middle of the specimen because the pixels in the centre have been scanned more times, making it more prone to damage.

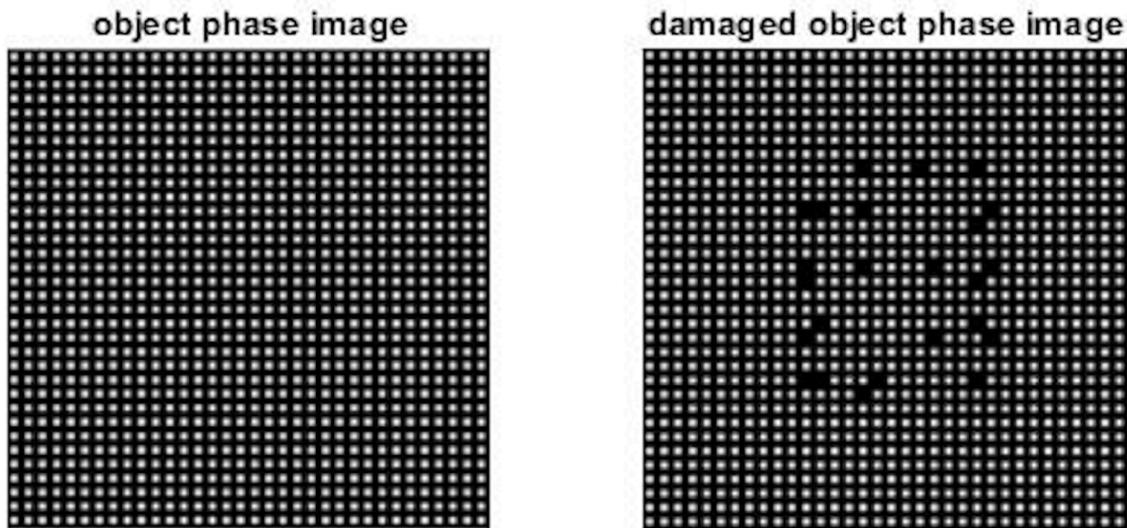


Figure 68: The object phase image before and after radiation damage. The damaged phase image shows empty holes where the atoms have been knocked out.

Figure 69 shows the modal decomposition results of the reconstruction. The left and right indicate the first and second object modes, respectively. The first object mode shows the specimen before damage, where every missing atom has been successfully recovered. The second object mode is a mixed-state mode, where there will be an individual state of the specimen after a single atom has been damaged and has caused blurs. The contrast in the second mode depends on the information ratio in the data regarding the missing atom positions. Where there is a higher information ratio in the diffraction pattern with a lower contrast, such as in the top three missing atoms in the figure on the right, there is a larger contrast among the atoms because there is less information about the atoms in the data. The results show the results we expected.

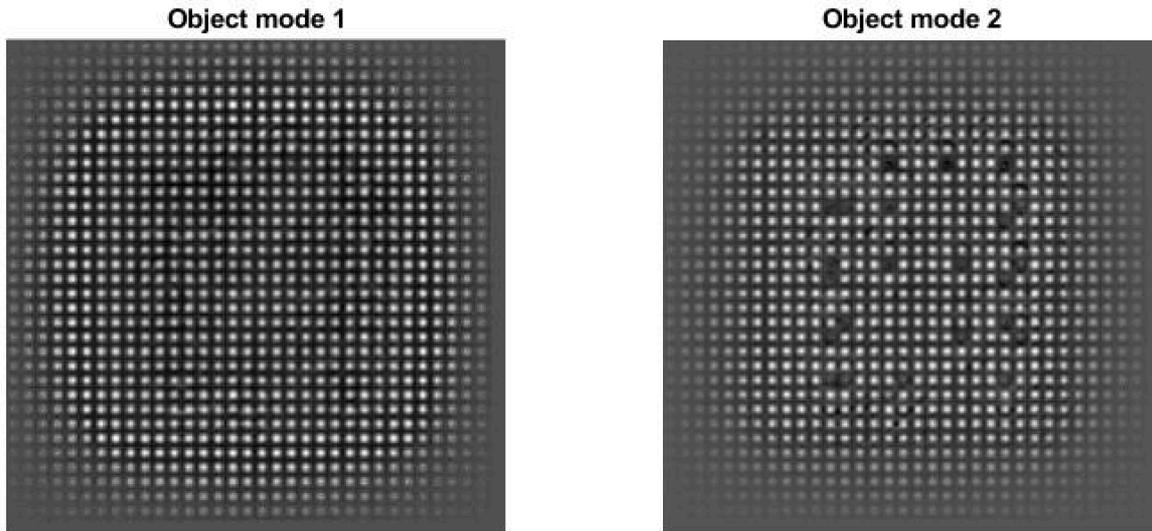


Figure 69: The modal decomposition result of missing atoms. The first mode has been generated as the object state before damage. The second object mode gives the mixed state of missing atoms, where the contrast of missing atoms depends on the valid information ratio in the data, and a larger information ratio results in less contrast.

Figure 70 illustrates the use of object mode for some isolated atoms, with the red circles indicating the positions of the damaged atoms. In the previous experiment, the specimen phase image indicated repeating structures that could be helpful using an iterative algorithm. However, the isolated atoms have almost no repetitive structure, making it fair to test the performance in the object mode.

The modal decomposition results are shown in Figure 71. The left and right images show the first and second object modes, respectively, with red circles highlighting the difference between the two. The results indicate that modal decomposition works with an isolated atom image as much as with regular structure atom images. However, both images result in some background blurs because the missing atoms cause reconstruction ambiguities in the area.

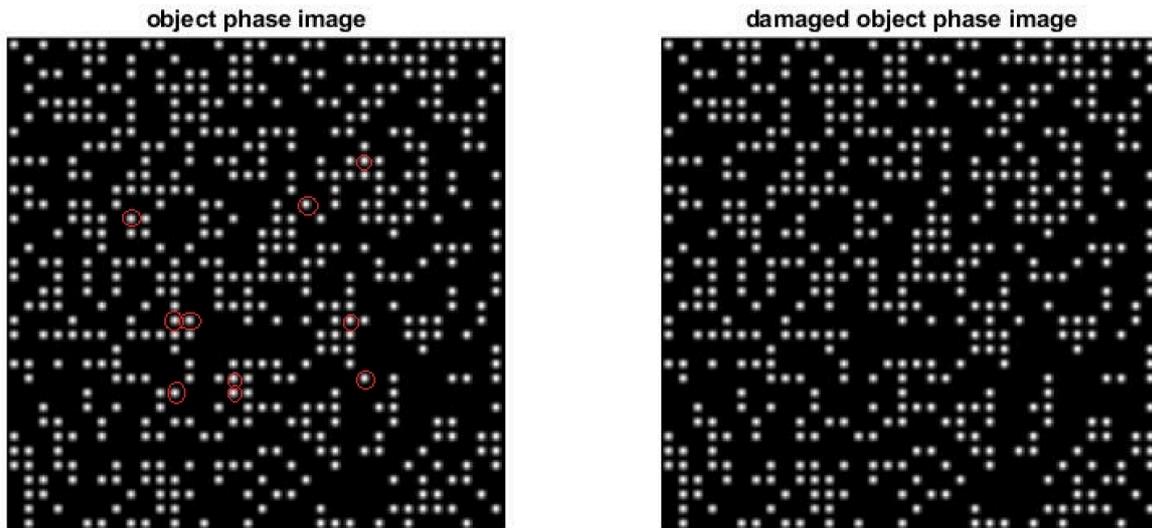


Figure 70: The object phase image before and after radiation damage. The red circles in the left image show the atoms that are missing on the right image.

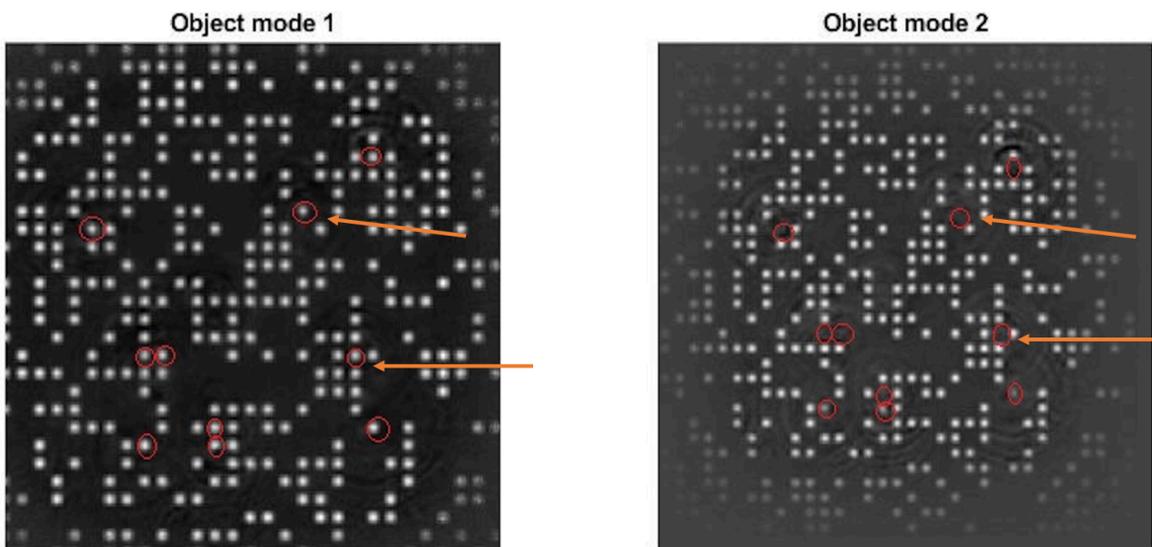


Figure 71: The modal decomposition result of missing atoms. The first mode has been generated as the object state before damage, while the second gives the mixed state of missing atoms. Red circles, two of which have been emphasised with pointer, show the position of damaging atoms. The contrast of missing atoms in the second mode depends on the validity of information in the data, with more information resulting in less contrast. The blurs in both images are because of the missing atoms causing reconstruction ambiguities in the area.

### 6.5.2. Simulation with noise

The second experiment tests the reconstruction of object modes in noisy environments. Noise is a significant issue affecting the reconstruction of objects. As the control-radiation-damage part is introduced, the damage probability is reduced by the flux in the experiment, which reduces the chance of damage along with the signal to noise ratio of each diffraction pattern. The same damage condition is used in this experiment with the addition of some Poisson noise to each diffraction pattern. Figure 72 shows the object-phase image before damage on the left, and the damaged-object-phase image, with holes indicating the damaged atom positions, on the right.

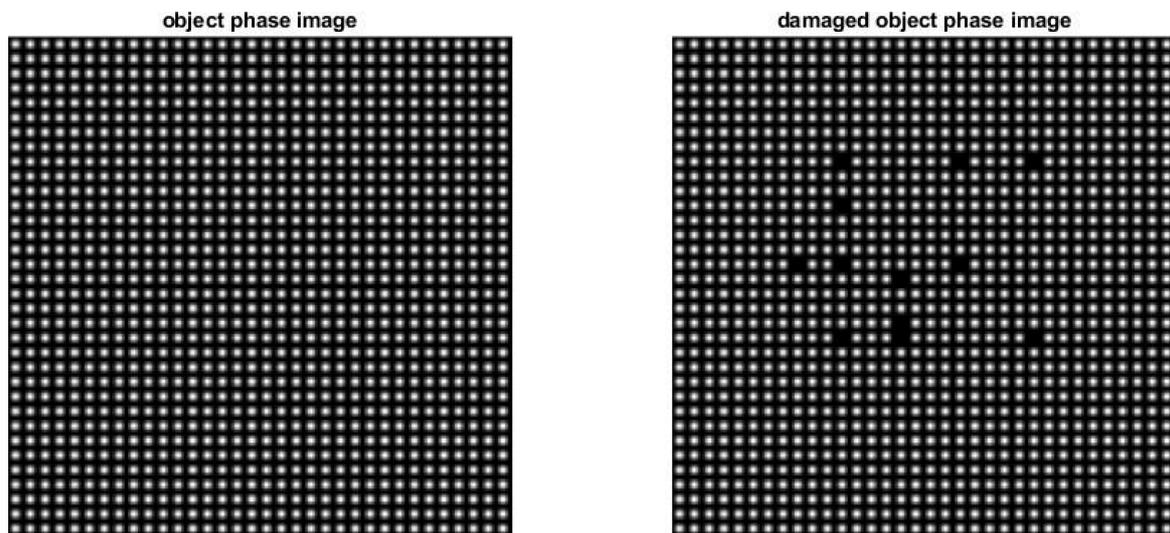


Figure 72: The object phase image before and after radiation damage. The damaged phase image shows holes where the atoms have been knocked out.

The modal decomposition result is shown in Figure 73. The left and right images illustrate the first and second object modes, respectively. The results indicate that, for repeating structured objects, noise has less of an effect in the modal decomposition reconstruction as the repeat structure provides a large amount of universal corrected information to the data.

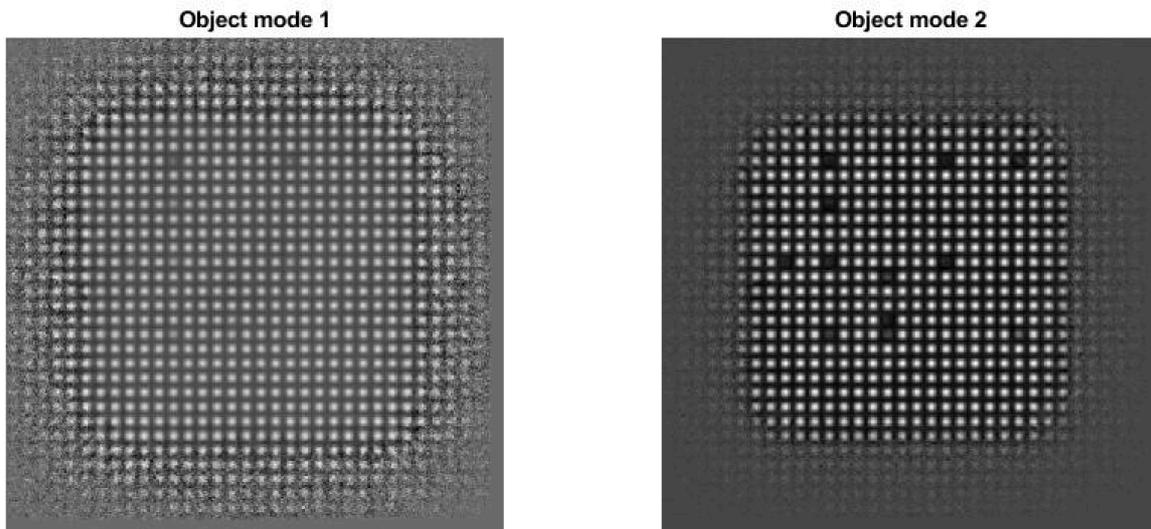


Figure 73: The modal decomposition result of missing atoms. Because of the repeating structure, there are numerous redundancies in the data required to solve the object modes.

The next experiment tests the object modes on some isolated atoms. Figure 74 shows the damaged atom positions using red circles.

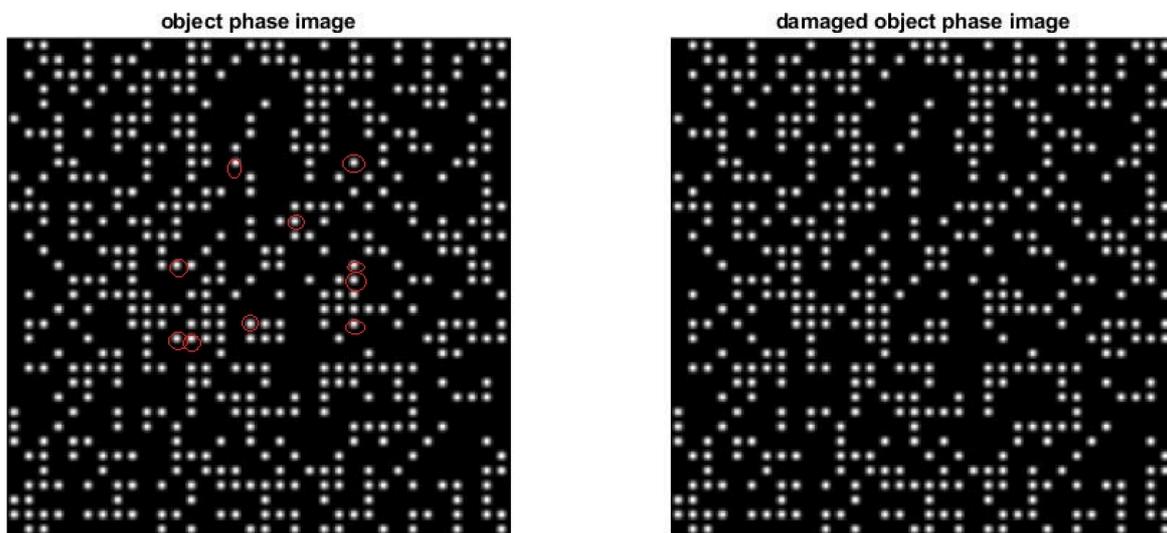


Figure 74. The object phase image before and after radiation damage. Isolated atoms are used to avoid repetitive structures, with the missing atoms highlighted using red circles.

The modal decomposition result is shown in Figure 75. The left and right images illustrate the first and second object modes, respectively. The results indicate that

noise affects the reconstruction of low information ratio atoms, such as the top two missing atoms near the border with less illumination. However, the atom in the middle of the specimen has a large information ratio in the data and thus has a strong noise tolerance and a clearer reconstructed image.

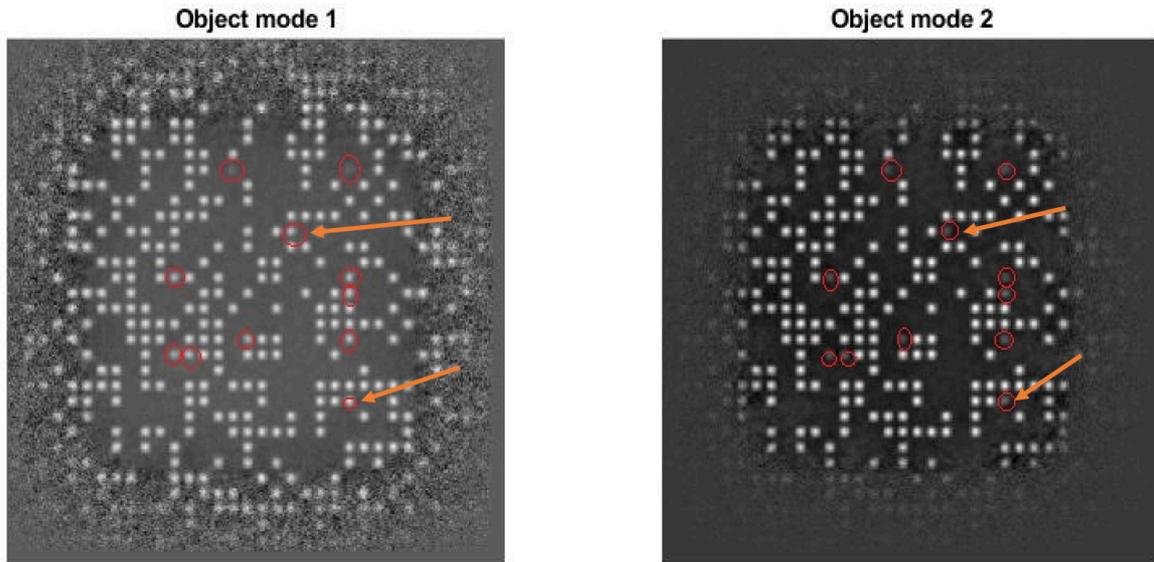


Figure 75: The modal decomposition result of missing atoms. The atom in the centre has a higher information ratio in the data, with better noise tolerance and clearer reconstruction than the atoms in the border. The red circles are described in the caption of Fig 74.

## 6.6. Analysis

Based on the above results, the object mode has the potential to represent the object state before damage, even if the damaged areas have less contrast than other undamaged structures. However, the performance of the additional object mode method depends on several factors. The first issue relates to when radiation happens. If the structure has less information across the whole data, it is harder for it to be clearly represented. If the radiation damage happens at the first scan, there is no diffraction pattern record of any information on the existence of the structure, so it cannot be worked out. The second factor is the damage position. The reason is the same as the first because, during the ptychography experiment, the centre was scanned more times than the edge, so more diffraction patterns could record the information about the centre structure of the specimen. The last factor is the noise,

which is a common influencing factor for all experiments. Generally, noise reduces the valid signal of the object structure.

## **6.7. Conclusion**

In this chapter, the interaction between primary electrons and the specimen atoms was introduced, which helped to form the interference pattern about the inner structure of the specimen on the detector. In addition, the interactions can damage the specimen by changing the structure or the chemical property thereof. The methods to reduce the probability of radiation damage in the experiments was reviewed. Except for changing the experimental equipment, the results indicated that the effect of damage on a specimen can be reduced through reconstruction because of the wide bandwidth of ptychography data, where the specimen can be treated before and after damage as two individual states. To prove this idea, mixed-state ptychography was simulated to reduce the radiation-damage effect on the reconstruction. The results showed that the object mode works for both no-noise or noisy data at varying levels, depending on the illumination number and damaged position.

# Chapter 7 Detector correction algorithm

Previous chapters examined a major use of redundancy in ptychographic data. However, there are many more applications of this redundancy. While the other chapters mostly focus on examining the capabilities of modal decomposition, this chapter explores a slightly different issue in ptychography related to faults that can occur in the detector. As a CDI method, the phase retrieval computation results of ptychography are really depends on the quality of the raw data, which is collected by the detector pixels. Therefore, the detector imperfection is a serious problem to ptychography. In this chapter, a new application of redundancy is examined and shown capable of calibrating a detector's non-linearity responds.

## 7.1. The detector and the detector mechanism

As a type of CDI method, ptychography uses recorded data to represent the specimen after calculation, rather than from the image directly. At each scanning position, the interference pattern is recorded by a detector. Because most radiations are invisible, an intermediate material was required to convert the invisible radiation to quantifiable data.

Historically, detectors have developed from gas filled detectors, including the ionization chamber [143], the proportional counter [144] and the Geiger-Muller tube [145], to liquid-filled detectors, including the liquid ionization chamber, the liquid proportional counter, the Cherenkov detector [146], the Bubble chamber [147] and the liquid-scintillator detector, to the solid-state detector, including the semiconductor detector and the chemiluminescent detector. The detector described in this chapter is based on a semiconductor detector. More information about these various types of

detectors can be found in the following references [143,144,148]. Here, a brief description of a TEM detector is provided.

The core part of the semiconductor detector is the P-N junction. There are two types of semiconductor materials: P-type and N-type. The N-type semiconductor material contains more free electrons and holes, due to the electron band jumps and a positive charge leave. As the name suggests, the P-N junction is created by joining P-type and N-type semiconductors together.

If there is no bias, the free electrons will move from N to P, leaving positive electrons and holes on the N-side and negative charges on the P side. Then, the holes will combine on both sides with free electrons to give a positional balance condition at the P-N interference. The balanced area is called the depletion layer. There is a tiny current flow in the P-N junction due to thermal agitation, called the dark current. The mechanism of the radiation detector enlarges the radiation-caused current, which should be high enough to facilitate measurement. A common method used to build a high electron field in the P-N junction is by adding a reverse bias. The theory suggests that, when the radiation passes through the depletion region, numerous free electrons and holes are created and form a measurable current that has a positive relationship with the energy of radiation. It adds a reverse bias to generate a larger depletion region, providing more current flow in the P-N junction.

When the electrons are incident upon the reverse-bias-based detector, the charges are added to the P side and form new electron-hole pairs and a pulse that can be measured. The pulse can be calculated using Ramo's theorem [149], being

$$i = cvE_w, \quad (113)$$

in which  $c$  is the electron charges or holes,  $v$  is the moving speed of charges or holes and  $E_w$  is the weighting electric field at the  $v$  direction.  $E_w$  is calculated by

$$E_w = \frac{1}{d} \quad (114)$$

at unit potential situation, and  $d$  is the depth of depletion region.  $d$  is the sum of the depth at the P- and N-junction sides as

$$d = d_c + d_h, \quad (115)$$

$$d_c = \sqrt{\frac{2\epsilon V}{e\rho_a \left(1 + \frac{\rho_a}{\rho_d}\right)}}, \quad (116)$$

and

$$d_h = \sqrt{\frac{2\epsilon V}{e\rho_d \left(1 + \frac{\rho_a}{\rho_d}\right)}}, \quad (117)$$

where  $d_c$  and  $d_h$  are the P- and N-side junction depths,  $\rho_a$  and  $\rho_d$  are the impurity concentrations at acceptor and donor, respectively,  $e$  is the charge of an electron,  $V$  is the potential difference, and  $\epsilon$  is the permittivity of the P-N junction.

The moving time for a charge and a hole at a distance of  $x$  away from a collection of electrodes are

$$t_c = \frac{x}{v_e}, \quad (118)$$

And

$$t_h = \frac{d - x}{v_h}, \quad (119)$$

where  $v_e$  and  $v_h$  are the moving speed of charges and holes, respectively. Therefore, the total charge for the charges and holes are

$$q_c = i_c t_c, \quad (120)$$

and

$$q_h = i_h t_h, \quad (121)$$

where  $i_c$  and  $i_h$  are the induced current of charges and holes, respectively, and the cumulative total charge is

$$q = q_c(t) + q_h(t) = (i_c + i_h)t. \quad (122)$$

As previously mentioned, there is always some dark current on the detector. Apart from dark current, environmental background noise and the efficiency distinction of different detector pixels or chips will significantly affect the performance of ptychography phase retrieval processes. This is due to the relationship between different record data that are severely damaged because of the above perturbations.

One of the most significant advantages of ptychography is the extremely abundant and diverse information in its multiple imaging collection system, allowing it to self-calibrate any errors in the dataset, such as partial coherence, scan position errors and missing data. As a result, it may be possible to correct any detector-pixel-record errors.

During the scanning process, the illumination scans different areas on the specimen while the exit wave is recorded by the same detector. In other words, each detector pixel records different information for each scan position. In the phase retrieval process, there will be one or more estimated object and probe functions in the beginning that produce the estimated exit wave recorded by the detector using the recorded intensity as a constraint to update the estimations. In an ideal situation, the final intensity estimate should equate to the real intensity for each detector pixel, meaning that the plot of estimated intensity against the real intensity should be a straight line with a gradient of one, as shown in Figure 76.

However, if the detector is not perfect, all the perturbations could disturb the linear relationship and deteriorate the reconstruction quality. Any record diffraction pattern

can be recognised as the convolution result of the probe function and related illuminated area on the specimen. If a part of the specimen has been scanned several times, parts of this information will be spread over many diffraction patterns, resulting in a real-space specimen constraint that can help to correct any imperfect detector-pixel measured intensity as the Fourier constraint. Based on this, a new algorithm Detector-pie is proposed. It is based on the ePIE algorithm, with an external detector correction process included therein.

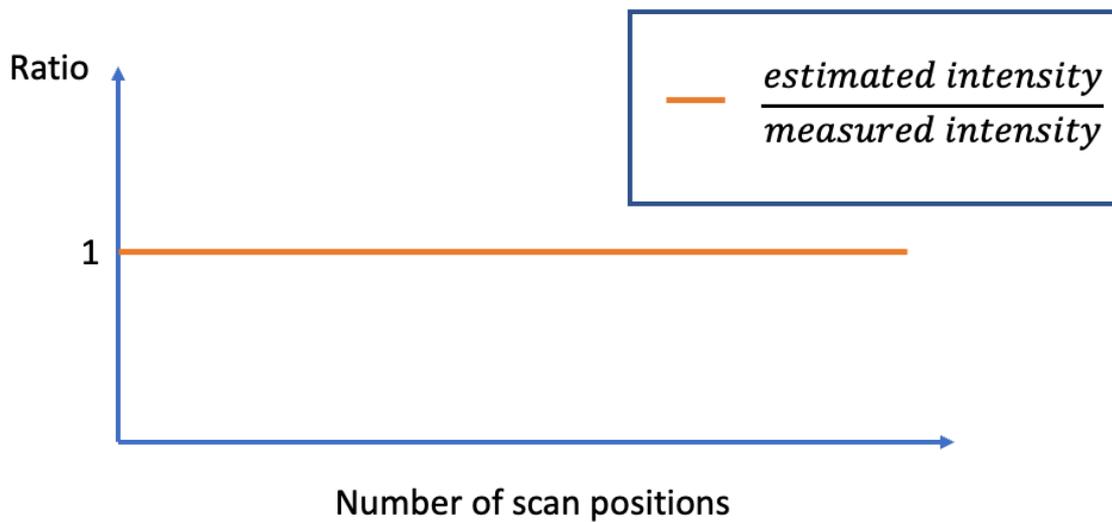


Figure 76: The measured versus the estimated intensities. In an ideal situation, the measured intensity should be equal to the estimated intensity.

## 7.2. Detector Correction

The detector-correction algorithm is divided into two parts: the iterative-phase-retrieval and the detector-correction parts, per Figure 77. The detector-correction begins by calculating the exit wave, being the product of the current probe function  $p_i(r - d)$  and the current estimate of the object function  $o_i(r)$ . The exit wave,  $\psi_i(r)$ , is then forward propagated to the detector and recorded as  $I_e(r)$ , where

$$I_e(r) = f(\psi_i(r))f^*(\psi_i(r)). \quad (123)$$

Normally, the exit wave at the Fourier space will be updated by the measured data  $I_m(r)$ . However, in the detector-pie, an attempt is made to correct the imperfect measured data  $I_d(r)$  by the forward estimated calculation  $I_e(r)$ . The relationship between  $I_m(r)$  and  $I_e(r)$  is modelled as a Taylor expansion, being

$$I_e(r) = c_0 + c_1 I_m(r) + c_2 I_m^2(r) + c_3 I_m^3(r) + c_4 I_m^4(r) + \dots + c_n I_m^n(r), \quad (124)$$

where  $c_n$  is the  $n^{\text{th}}$  order detector pixel response that provides better convergence accuracy. The best fit curve is modelled as

$$I_e(r) = P(c_n, I_m(r)), \quad (125)$$

in which  $P$  is a polynomial function. At the beginning,

$$I_m^1 = I_d, \quad (126)$$

once the reconstruction has converged, then the next estimate of  $I_m^n$  is given by

$$I_m^{n+1} = I_m^n + \gamma(I_m^{n-1} - I_m^n), \quad (127)$$

where  $\gamma$  is the updated step size between 0 and 1. The next iteration will use the  $I_m^{n+1}$  to determine the Fourier constraint as  $I_{mc}(r)$  to provide a more accurate reconstruction. After the detector correction, the updated exit wave at the Fourier space  $\psi'_i(r)$ , is equal to

$$\psi'_i(r) = \sqrt{I_{mc}(r)} \frac{\psi_i(r)}{|\psi_i(r)|}, \quad (128)$$

and will be employed to update  $p_i(r - d)$  and  $o_i(r)$ .

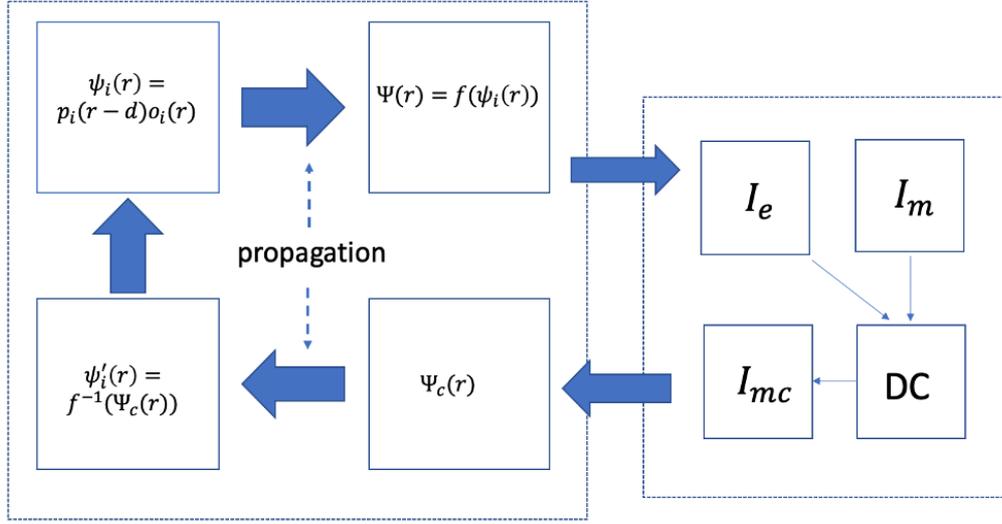


Figure 77: The flow chart of the detector-correction.  $p_i(r-d)$  and  $o_i(r)$  are the estimated probe and object functions,  $\psi(r)$  and  $\Psi(r)$  are the exit waves at the real and Fourier spaces,  $I_e$  is the calculated intensity,  $I_m$  is the intensity using the Fourier constraint, DC represents the detector-correction process and  $I_{mc}$  is the corrected intensity using the next-iteration calculation.

### 7.3. The simulation

The first step of the simulation is making the data non-linear. In this simulation experiment, a sine non-linearity is inserted into the test dataset, where choosing sine non-linearity is due to the sine function using a relatively simple Taylor expansion of

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots, \quad (129)$$

while only using the first four terms in the 7<sup>th</sup> degree. Before making all the pixels non-linear, a few pixels were randomly chosen to perform non-linearity, where the number of pixels began from 1, then 10, 100, 1,000, until the break down number was found, beyond which the algorithm did not work. The non-linear pixels are generated as

$$I_d = A + B\sin(CI_r), \quad (130)$$

where  $A, B$  and  $C$  are constants. In the simulation,  $A$  was set as the noise on the detector without any illumination and  $B$  and  $C$  were set as the amplification effects of the detector pixels. Because  $CI_r$  should not be more than  $\frac{\pi}{2}$ , the maximum of  $CI_r$  should be less than  $\frac{\pi}{2}$ , as

$$C \max(I_r) \leq \frac{\pi}{2}, \quad (131)$$

and

$$C \leq \frac{\frac{\pi}{2}}{\max(I_r)}. \quad (132)$$

Here, the  $\max(I_r)$  of the test dataset is  $I_{max}$  and  $I_L = 1.1 \times I_{max}$ , then

$$C = \frac{\pi}{2I_L}. \quad (133)$$

If  $A = 0$ , for some small value,

$$\sin(x) \approx x, \quad (134)$$

and

$$B \sin(CI_r) = BCI_r = I_r. \quad (135)$$

Therefore,

$$B = \frac{1}{C} = \frac{2I_L}{\pi}. \quad (136)$$

For the perfect data, Lenna was used as the object image. The reconstruction with collected diffraction patterns, as shown in Figure 78, was used to generate a test dataset without any perturbation as  $I_r$ , where all the detector pixels work perfectly. The test dataset provides the best possible reconstruction as a reference. Next, a continuous set of data was generated with the intensity value from 1 to 100 counts, with a step size of 1 count. The intensity of the linear situation was checked before and after the sine non-linearity. The results are shown in Figure 79. In the figure, the blue line reflects the real intensity values, and the red line is the sine non-linear value. Clearly, the sine non-linear works better with a wider range of intensities in the data and has less of an affect for a low value.

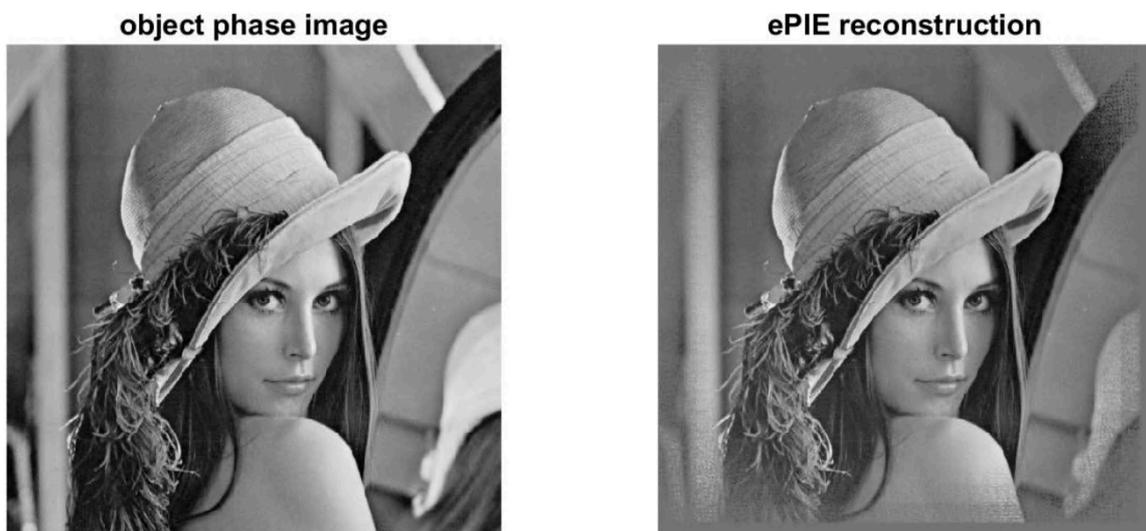


Figure 78: The object function and reconstruction using generated test data. This is the reference image where all the detector pixels work perfectly.

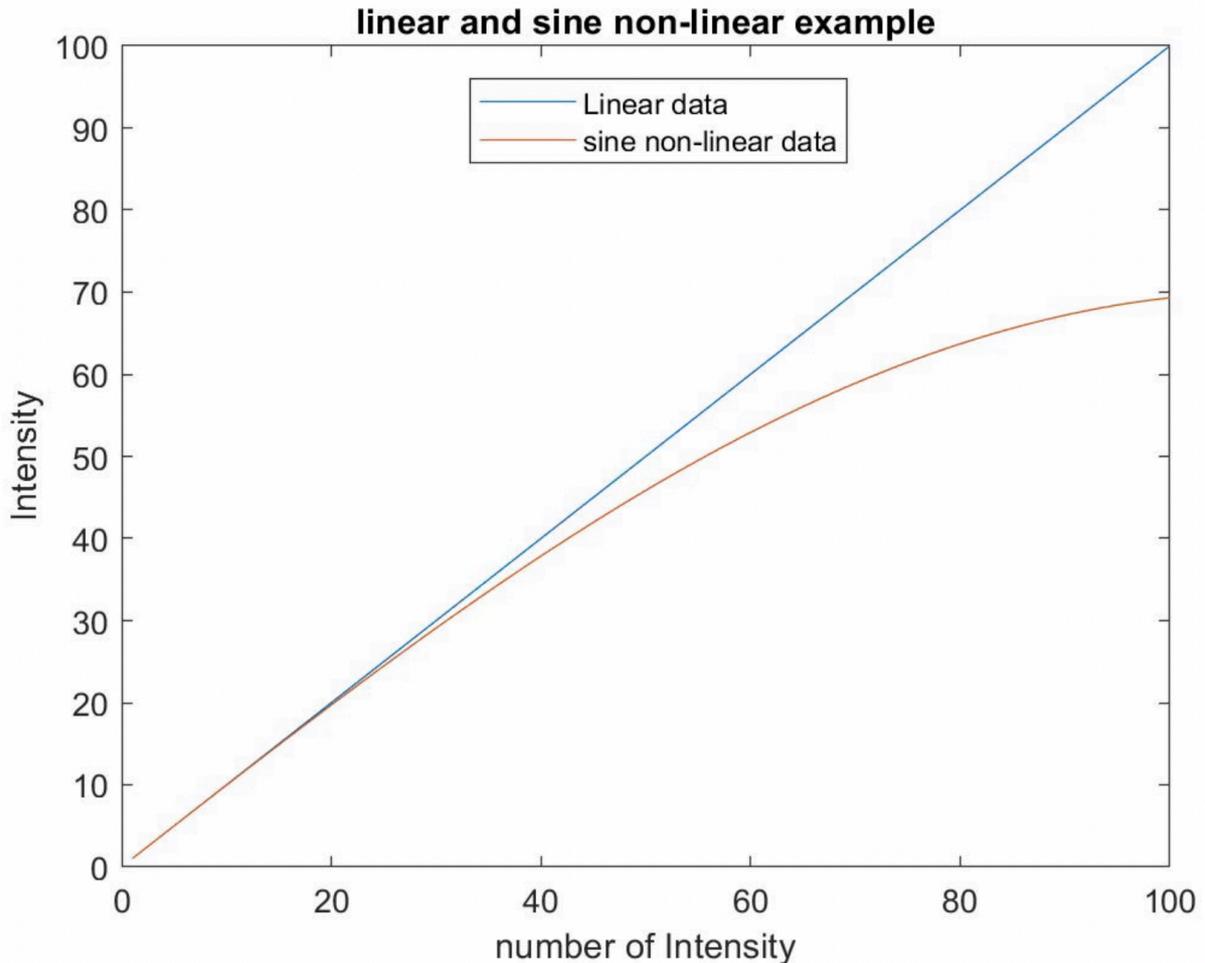


Figure 79: Simulation results of linear data and sine non-linear data. The two lines indicate that the sine non-linear has a stronger effect on larger values.

### 7.3.1. Non-linear noiseless detector

In the experiment, a complex image, with a size of 512 by 512 pixels, was used as the object. The modulus image is white with a value of 1 being totally transparent, and the phase image is Lenna with a phase change from 0 to  $0.05\pi$  as a weak phase. The probe function is a defocus probe, where any two joint probe positions have approximately 70% overlap to provide sufficient information content during the scanning. The detector size is 256 by 256 pixels, and there is the sine non-linear on the detector. There are three individual experiments, being the ideal data reconstruction, the ePIE reconstruction without detector correction and the ePIE reconstruction with detector correction.

There was a total of 500 iterations during the reconstruction. For the detector correction experiment, the detector correction began at the 300<sup>th</sup> iteration and was employed after every 10 iterations. In the detector correction process, only two-degree polynormal fittings were used in the calculation. Figure 80 gives the reconstruction result using non-linear data without and with detector correction, being the left and right images, respectively. Comparing the two results, reconstruction without detector correction shows a messy image with some non-existent features. In contrast, the correction removes almost all the messy and non-existent features and provides a better contrast, meaning that the algorithm works well, and the non-linear intensity values have gradually been corrected.



Figure 80: The phase reconstruction result with non-linear test data with and without detector correction. After the correction, the reconstruction shows better contrast, and the blurs in the left image have been removed when compared to the image on the right.

The same experiment was performed for a modulus image, where the modulus image is Lenna, with a value from 0 to 255. The same probe and overlap level were used to perform the reconstruction with 500 iterations, and the detector correction began from the 300<sup>th</sup> iteration. The results are shown in Figure 81, with the left and right images showing the reconstruction results without and with correction, respectively. This test shows that there was not a significant difference between the images because for the

modulus image, the transferred power in the incident wave had more various values for all probe positions, so there was more information to calibrate the detector.

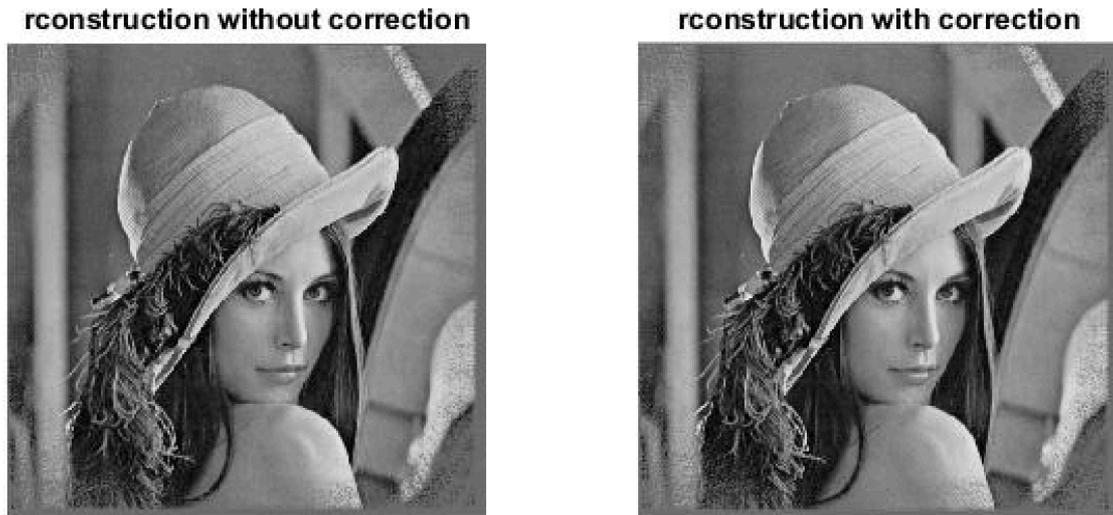


Figure 81: The modulus image reconstruction results with sine non-linear detector. The results show almost no difference, because at each scan position, every detector pixel records various values that provide more information about how to calibrate the detector.

### 7.3.2. *Non-linear and random gain on all detector pixels*

In the next test, an extra random gain, being the intensity sensitivity of the pixel, was added to each pixel to enlarge the gain effect. With gain values between 0 and 2 normal distributions, the detector pixels could be dead or record large and unrealistic intensities. The same experiments were performed on the data, and the reconstructions after 500 iterations are shown in Figure 82. In the figure, the left and right images show the results before and after reconstruction, respectively. The results show that after detector correction, reconstruction gives a better reconstruction and large field of view. This can further promote the reconstruction quality by providing more corrections.

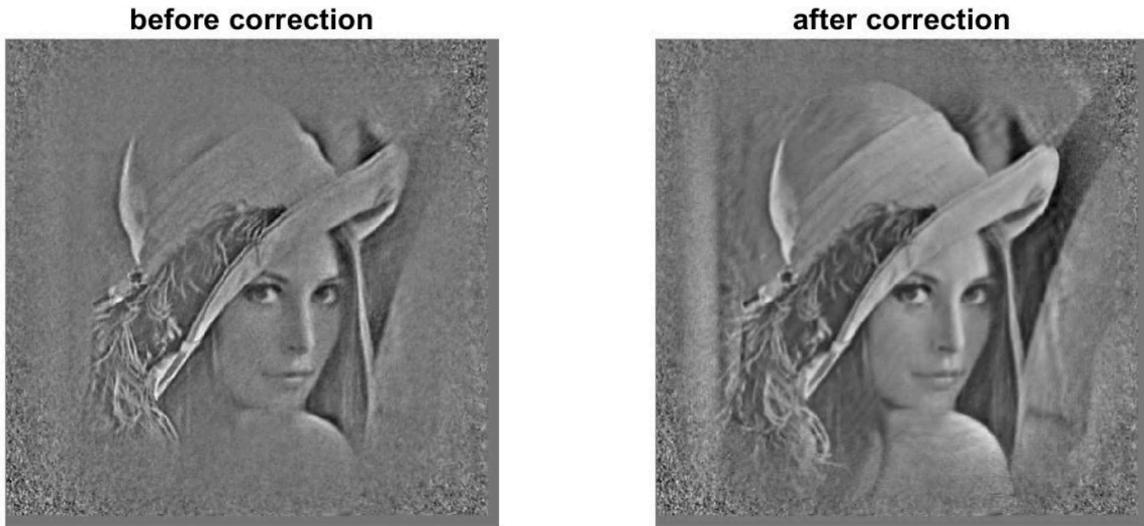


Figure 82: The phase reconstruction result with a non-linear and random gain on each pixel test data with and without detector correction. After correction, the phase image is clearer and shows a larger field of view than the image before correction.

The same experiment was performed for the modulus image object, except for the random gain on the detector that was randomly distributed between 0 and 2. The results were compared after 500 iterations, with the detector correction starting at the 300<sup>th</sup> iteration, and each 10<sup>th</sup> iteration thereafter. The results shown in Figure 83 indicate that, after correction, reconstruction provides clearer details on the image, better contrast, and a larger field of view.



Figure 83: The modulus image reconstruction results with non-linear test data, with and without correction.

#### 7.3.4. Non-linear, random gain and noise on all detector pixels

In addition to the sine non-linear and random gains, this test includes an offset to the detector, being a value between 0 and 30. The modulus reconstruction results are illustrated in Figure 84, where the sine is non-linear, the gain distributed on the detector is between 0 and 2 and the random offset value is between 0 and 30, where the maximum is about 1.5% of the average recorded intensities. The reconstruction results are almost identical to the non-linear and gain experiments. This is because the offset value is not strong.

After the test, the detector correction was found to not work in this situation as some phase information disappeared, as shown in Figure 85. By tracking the correction process for the phase image, the intensity difference in each scan position was almost constant in the ideal situation. However, when gain and noise were added to each diffraction pattern during the best fitting process, the measured intensity yielded incorrect linear values.



Figure 84: The modulus image reconstruction results with a non-linear test data, without and with correction.

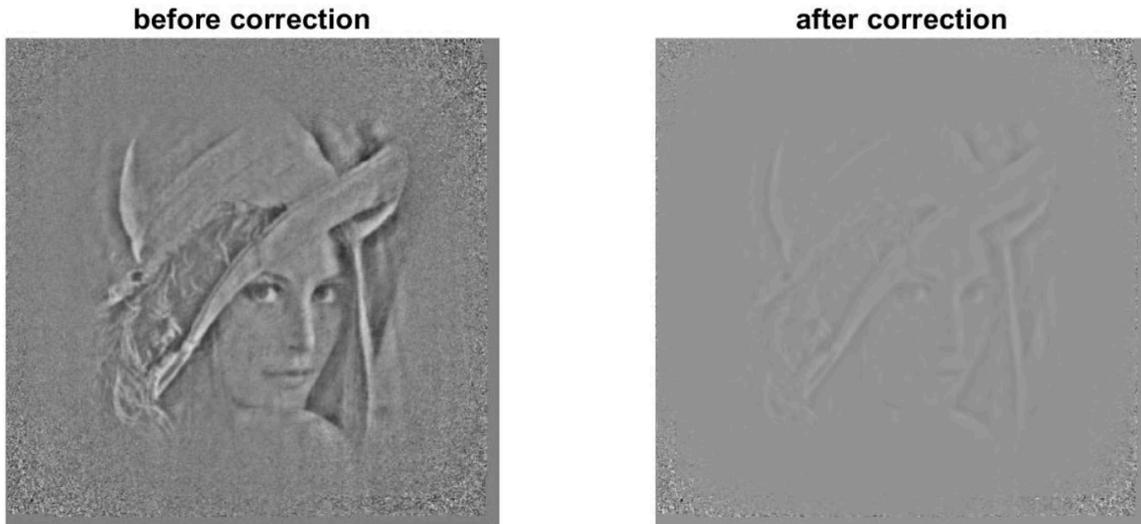


Figure 85: The phase image reconstruction results with a non-linear test data, without and with correction. Because the intensity of the phase image has a less intense difference at various scan positions, when noise and gain are included on the detector, more intensity differences are artificially included into different diffractions. Therefore, when best fitting the data, the intensity values are pushed in the wrong directions.

## 7.4. Conclusion

In this chapter, we tested our detector correction idea for non-linear or bad quality detectors in a model ptychographic experiment. This was done to improve the reconstruction quality, given the abundant and diverse information available in ptychography data. We used a polynormal to best fit the nonlinear data and gradually correct the non-linear data to linear data. We explored the correction method for a non-linear detector, a non-linear detector with random pixel gains and a non-linear detector with random pixel gains and dark current across all the detector pixels. We investigated modulus-only and phase-only images. The results showed that the detector-correction method works well for modulus images in any situation but does not work so well for a phase image with dark current. Therefore, to make the algorithm work well, it is better to have a pre-recorded diffraction pattern about the dark current on the detector.

## Chapter 8 Conclusion

This thesis has focused on the use of the redundancy in ptychography data to solve for partial coherence (and thus increase the speed of an experiment), to improve the reconstruction quality for a damaged specimen and to facilitate the correction of nonlinearities in detectors.

The key application of the ptychographic redundancy is the spatially mixed state ptychography, where the mixed states come from spatial instability or spatial oscillations, or from partially coherence in a wavefield. The overall mixed state is the mixed result of the interaction of each probe state with each object state. To calculate the probe states and object states from the data, we used the ePIE algorithm, a variant of one of the commonly used reconstruction algorithms. Except for the Fourier constraint part, where the total calculated intensity of each mode in Fourier space needs to be kept the same as the recorded real intensity, the mixed stated algorithm is the same as the ePIE.

We began by theoretically exploring spatially mixed states of the probe/illumination modes. In order to simplify the results, our illumination source was simulated as a series of different numbers of source points. Each point has a certain intensity and a different spatial distribution to the other point sources, and each of them gives a different state in the probe. If the exact number and state of probe modes are known, then a calculation will give some 'empty' probe modes when more probe mode states are used in the reconstruction than the real number of underlying modes. The empty modes are pure artefacts without a clear pattern, and they have no contribution to the reconstruction. Meanwhile, each probe mode patten has a related power value to measure the participation in the reconstruction process; the larger the power, the more important the related mode pattern is in the reconstruction process. Therefore, we used the calculation results to show that, except for the empty modes, acceptable reconstructions can be obtained by using just a few large power modes.

To give a direct understanding of how the modal decomposition works, we tracked the power changes of each mode pattern in the calculation with iterations by using a series of equal power modes. The tracking results showed that the mode patterns with strong power will be worked out more quickly, while the empty modes do not participate in the calculation until the end.

After establishing the above findings in relation to probe modes, we studied the influencing factors in the modal decomposition process. The first factors we considered were the overlap and step size. A calculation showed that if the overlap between two adjoint scan positions is reduced, this will reduce the known to unknown ratio in the dataset. So, if the modal decomposition process is treated as a 'number to number argument, with a small overlap, then there will not be enough information to solve the specimen function and probe mode functions.

Next, we investigated how the specimen and probe mode structures affect the modal decomposition. We used the simplest object imaginable – a single line – and we used two different spatially distributed points as the sources to carry out the modal decomposition. The two results demonstrated that if there is more structural diversity in the probe modes, it is easier to get a good reconstruction. We also examined this with an object with less complexity; surprisingly, even though it was very simple and had the same influence on the probe mode, it did not substantially affect the probe reconstruction.

As mentioned above, the probe modes can be employed to treat remove the effects of partial coherence, and it may not be necessary to put a large number of modes into calculations in order to get a good reconstruction. We realised that it may be possible to artificially engineer some known mode patterns in order to overcome the partially coherence by using just a few modes. Because ptychography is a type of CDI method, the coherence critically limits the reconstruction quality; therefore, it would be better if the influence of partially coherence on the reconstructions could be reduced. To achieve this goal, we designed a four-aperture mask, with the diameter of each aperture being about the same as the coherence width of the illumination. Meanwhile, we also used a large aperture mask, with the same area as the total area of the four

apertures, to ensure that the same amount of flux could pass through both masks. (In this stage, we did not use a lens, only the aperture masks.) After the calculation, the four-aperture mask did not show a better reconstruction than the large aperture mask, because the illumination that passed each aperture on the four-aperture mask recorded a very similar interfere pattern on the detector, which misleads the reconstruction direction.

Later, we used cut-off apertures, which were not identical, to replace the round aperture on the four-aperture mask, as they generated different interfere patterns on the detector. The calculation was then repeated. By using FRC to give a quantitative comparison, the cut-off aperture worked better than the large aperture in the situation where the coherence width of illumination was roughly equal to the diameter of the cut-off aperture. Continuing this approach, we used four-corner masks to replace the cut-off aperture; this had the advantages of breaking symmetry as before and was also a very easy way of implementing the concept experimentally, as it was simple to build by adding two wires on to an aperture or lens. We did the same calculation using the corner masks, and the results proved successful. To make the simulation closer to a real-world situation in x-ray ptychography, we placed a lens downstream of the masks and used some real parameters to do the model calculation again. The method proved as effective as our aperture-only calculation, and, indeed, the separation of the modes was even clearer. After conducting another test, we realised that by using a lens there would be some connected interfere fringes between each two probe modes, which helped us considerably in understanding the true experiment.

The true experiment was conducted at Diamond Light Source's I13 beamline. Before we went to the synchrotron, we ordered several masks with holes in them for the experiment. Initially, we conducted some experiments by using the masks with holes to check the influence of defocus and overlap. If the exposure time, the defocus and the partially coherence level were fixed, the masks with holes could give a very good reconstruction, but the interfere between each hole was too strong to separate each small probe mode out.

Because the coherence width of the x-ray beam in the I13 is much larger in the vertical direction than in the horizontal direction, we proposed to use a wire to replace the

mask to separate the beam horizontally and vertically. At the same time, we used a lens-only experiment as a reference to compare the reconstruction quality when there was no wire blocking the beam into separate modes. We collected a large amount of data by changing the source slit's width to vary the source coherence at fixed defocus and overlap situations. Following the calculation, we demonstrated that, as the vertical wire has the same direction as the coherence width, the illumination passing through the vertical wire and lens is more coherent than that passing through the horizontal wire and lens. Therefore, it was possible to get a good reconstruction using only two modes. For the other two experiments, they required more modes to treat the partial coherence, due to there being less coherence in the data. On the other hand, because the data in the vertical wire experiment had a much higher known to unknown ratio of numbers, it had a much lower redundancy requirement than the other, so a good reconstruction could be worked out at a low-level overlap, which could help speed up the data collection in ptychographic scans. This is a key finding from this thesis.

We then went on to show that not only the probe modes but also the object modes can help to improve the reconstruction in ptychography. In the real world, radiation damage is an inevitable issue in microscopy experiments, especially for electron microscopy experiments. A lot of research has focused on how to describe the mechanism of radiation damage in electron microscopy and how to reduce the risk of damaging the specimen. Here, we investigated how to improve the reconstruction quality in ptychography when the damage was occurring. In other words, rather than trying to stop the damage by physical means, we aimed to use the redundancy in the ptychographic data set to reconstruct what the specimen looked like before it was damaged. The main idea is that different states in the object appear as different object modes. It is logical to think that a pre-damaged mode can be found even in data that has lots of damaged specimen information.

To begin, we built a mathematical model to calculate the damage probability of each atom on the specimen during the scanning. By comparing the accumulated electron distribution result using a big probe function and a small probe function, we found that using a big probe function can reduce the chance of knock-on damage but creates a lower reconstruction quality as each atom will receive fewer electrons. However, we decided to use a big probe function to do the following tests. In the first test, we used

two object modes to do the modal decomposition, with evenly distributed damage and randomly distributed atoms. Following the calculation, we found that the whole specimen function can be represented even if it has been damaged. However, the reconstructed missing atoms have a very low contrast compared to the undamaged atoms, as there are only a few diffraction patterns that have recorded the missing atom, so there is not enough information to allow it to be worked out. Also, the reconstruction quality of missing atoms is related to their position; because the centre part of the specimen has been scanned more times than the boundary area, more information has been recorded for the centre atoms. Next, we put Poisson noise into the modelled data, and showed that it still works. However, due to a shortage of time, a more detailed study of this topic will have to await further research.

Apart from modal decomposition, the redundancy in ptychography data can also be applied to other aspects of the technique. The last chapter endeavoured to exploit the redundancy to correct non-linearities in the detector. Detector non-linearity is quite common in electron microscopy instruments, especially those that use a charged coupled device bonded via optical fibres to a scintillator. In short, the detector non-linearity is caused by the deviating signal or the dark current on the detector.

To correct the non-linearity, we have created an algorithm, based on ePIE, with an extra best-fitting process, which we have modelled theoretically. Before applying the Fourier constraint, we supposed that the calculated intensity is closer to the real intensity than the measured non-linear diffraction patterns, because the calculated intensity involves data from very many probe positions, not just one. We used a polynomial function to fit the calculated intensity to the measured intensity. Then we used the fitted intensity to do the Fourier constraint and the update. In this project, we put the sine non-linear to the measured intensity and in three different situations to test the performance of this algorithm for the modulus image and phase image.

After the calculation, we found that the algorithm works very well for the modulus image in the presence of non-linearities only, and in the non-linear with random gain and non-linear with random gain and offset situations. However, it does not work as well for phase images as for modulus images. This is because for the phase image,

the transferred power in the incident wave is about constant for all probe positions, so there is not much information to calibrate the detector.

There are some meaningful future works for my research. The most straightforward work is to prove my damaged object reconstruction results and detector correction algorithm by real experiments.

There is more work that can be extended in my synchrotron experiment. Firstly, because the wire will block most of the flux in the wire experiment, the total flux in the two experiments is different. So, there should be an extra experiment to compare the reconstruction quality of wired experiments and FZP only experiments with different numbers of mode and different speed at same flux condition. Secondly, if the total experiment flux is fixed, to speed up a ptychography experiment, we can not only use large step size with long exposure time, but also small step size with short exposure time, where, in a unit time, the same number of photons will go through each object pixels in both situations. Therefore, we can compare the performance of the two experiments to test if dose fractionation can be applied in ptychography and the modal decomposition method. Finally, in our experiment, we use only one wire. So, we can try more wire tests, which could be more persuasive.

## References

- [1]. Abbé, E. (1873). Contributions to the theory of the microscope and microscopic detection (Traduction from German). *Archiv Für Mikroskopische Anatomie*.
- [2]. Jiang, Y., Chen, Z., Han, Y., Deb, P., Gao, H., Xie, S., ... Muller, D. A. (2018). Electron ptychography of 2D materials to deep sub-ångström resolution. *Nature*.  
<https://doi.org/10.1038/s41586-018-0298-5>
- [3]. Rodenburg, J. (2018). A record-breaking microscope news-and-views. *Nature*, 559(7714), 334–335. <https://doi.org/10.1038/d41586-018-05711-y>
- [4]. Wolf, E. (1986). New theory of partial coherence in the space-frequency domain Part II: Steady-state fields and higher-order correlations. *Journal of the Optical Society of America A*, 3(1), 76. <https://doi.org/10.1364/josaa.3.000076>
- [5]. Williams, G. J., Quiney, H. M., Peele, A. G., & Nugent, K. A. (2007). Coherent diffractive imaging and partial coherence. *Physical Review B - Condensed Matter and Materials Physics*, 75(10), 1–7. <https://doi.org/10.1103/PhysRevB.75.104102>
- [6]. Rodenburg, J. M., & Bates, R. H. T. (1992). The theory of super-resolution electron microscopy via Wigner-distribution deconvolution. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 339(1655), 521–553. <https://doi.org/10.1098/rsta.1992.0050>
- [7]. Thibault, P., & Menzel, A. (2013). Reconstructing state mixtures from diffraction measurements. *Nature*, 494, 68. Retrieved from <http://dx.doi.org/10.1038/nature11806>
- [8]. Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*, 28(2), 171–193. <https://doi.org/10.1017/s003358350000305x>
- [9]. Egerton, R. F., Li, P., & Malac, M. (2004). Radiation damage in the TEM and SEM. In *Micron* (Vol. 35, pp. 399–409). <https://doi.org/10.1016/j.micron.2004.02.003>
- [10]. Egerton, R. F. (1976). Measurement of inelastic/elastic scattering ratio for fast electrons and its use in the study of radiation damage. *Physica Status Solidi (A)*, 37(2), 663–668. <https://doi.org/10.1002/pssa.2210370237>
- [11]. Egerton, R. F. (2013). Control of radiation damage in the TEM. *Ultramicroscopy*, 127, 100–108. <https://doi.org/10.1016/j.ultramic.2012.07.006>
- [12]. Gerchberg, R. W., & Saxton, W. O. (1972). A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35(2), 237–246. <https://doi.org/10.1070/QE2009v039n06ABEH013642>
- [13]. Fienup, J. R. (1982). Phase retrieval algorithms: a comparison. *Appl. Optics*, 21(15), 2758–2769. <https://doi.org/10.1364/AO.21.002758>

- [14]. Rodenburg, J. M., & Faulkner, H. M. L. (2004). A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20), 4795–4797. <https://doi.org/10.1063/1.1823034>
- [15]. Maiden, A. M., & Rodenburg, J. M. (2009). An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10), 1256–1262. <https://doi.org/10.1016/j.ultramic.2009.05.012>
- [16]. Williams, D. B., & Carter, C. B. (2009). *Transmission Electron Microscopy: A Textbook for Materials Science. Materials Science* (Vol. 1–4). [https://doi.org/10.1007/978-1-61779-415-5\\_23](https://doi.org/10.1007/978-1-61779-415-5_23)
- [17]. Michette, A. G., & Buckley, C. J. (Christopher J. . (1993). *X-ray science and technology*. Institute of Physics Pub. Retrieved from [https://inis.iaea.org/search/search.aspx?orig\\_q=RN:25016340](https://inis.iaea.org/search/search.aspx?orig_q=RN:25016340)
- [18]. Born, M., & Wolf, E. (1999). *Principles of Optics 7th edition. Principles of Optics Electromagnetic Theory of Propagation Interference and Diffraction of Light*. [https://doi.org/10.1016/S0030-3992\(00\)00061-X](https://doi.org/10.1016/S0030-3992(00)00061-X)
- [19]. Köhler, H. (1981). On Abbé's theory of image formation in the microscope. *Optica Acta*, 28(12), 1691–1701. <https://doi.org/10.1080/713820514>
- [20]. Rayleigh, Lord. (1879). LVI. Investigations in optics, with special reference to the spectroscope. *Philosophical Magazine Series 5*, 8(51), 477–486. <https://doi.org/10.1080/14786447908639715>
- [21]. Goodhew, P. J., Humphreys, J. F., & Beanland, R. (2001). Electron Microscopy and Analysis. *Annals of Physics*, 54, 254. Retrieved from <http://books.google.com/books?hl=en&lr=&id=zuxPIVmGLGsC&oi=fnd&pg=PR9&dq=Electron+microscopy+and+analysis&ots=-U-Itc0MZJ&sig=ibc15cL-3PUTTnXb-0VEbsmESEw>
- [22]. Matsuyama, S., Yasuda, S., Yamada, J., Okada, H., Kohmura, Y., Yabashi, M., ... Yamauchi, K. (2017). 50-nm-resolution full-field X-ray microscope without chromatic aberration using total-reflection imaging mirrors. *Scientific Reports*, 7(November 2016), 2–4. <https://doi.org/10.1038/srep46358>
- [23]. Scherzer, O. (1949). The theoretical resolution limit of the electron microscope. *Journal of Applied Physics*. <https://doi.org/10.1063/1.1698233>
- [24]. Mobilio, S., Boscherini, F., & Meneghini, C. (2015). *Synchrotron radiation: Basics, methods and applications. Synchrotron Radiation: Basics, Methods and Applications*. <https://doi.org/10.1007/978-3-642-55315-8>
- [25]. Shenoy, G. (2003). Basic Characteristics of Synchrotron Radiation. *Structural Chemistry*. <https://doi.org/10.1023/A:1021656723964>
- [26]. Paganin, D. (2009). *Coherent X-Ray Optics (Oxford Series on Synchrotron Radiation)*, 1–424.

- [27].Luchini, P., Motz, H., & Motz, L. E. P. E. H. (1990). *Undulators and Free-Electron Lasers*. Clarendon Press. Retrieved from <https://books.google.co.uk/books?id=DOZRAAAAMAAJ>
- [28].Onuki, H., & Elleaume, P. (Eds.). (2003). *Undulators, wigglers and their applications*.
- [29].Wille, K. (1996). Introduction to insertion devices. In *CERN Accelerator School: Course on Synchrotron Radiation and Free-electron Lasers* (pp. 61–75).
- [30].Bunker, G. (2016). X-ray absorption spectrometers. In *Encyclopedia of Spectroscopy and Spectrometry*. <https://doi.org/10.1016/B978-0-12-409547-2.05066-6>
- [31].Snigireva, I., & Snigireva, A. (2004). Microscopy Techniques - X-Ray Microscopy. In *Encyclopedia of Analytical Science: Second Edition* (pp. 151–158). Elsevier Inc. <https://doi.org/10.1016/B0-12-369397-7/00729-9>
- [32]Horowitz, P., & Howell, J. A. (1972). A scanning x-ray microscope using synchrotron radiation. *Science*, 178(4061). <https://doi.org/10.1126/science.178.4061.608>
- [33].Hettel, R. (2014). Challenges in the design of diffraction-limited storage rings. In *IPAC 2014: Proceedings of the 5th International Particle Accelerator Conference*.
- [34].Hettel, R. (2014). DLSR design and plans: An international overview. *Journal of Synchrotron Radiation*. <https://doi.org/10.1107/S1600577514011515>
- [35].Einfeld, D., Plesko, M., & Schaper, J. (2014). First multi-bend achromat lattice consideration. *Journal of Synchrotron Radiation*. <https://doi.org/10.1107/S160057751401193X>
- [36].Mcneil, B. W. J., & Thompson, N. R. (2010). X-ray free-electron lasers. *Nature Photonics*. <https://doi.org/10.1038/nphoton.2010.239>
- [37].Pellegrini, C. (2012). The history of X-ray free-electron lasers. *European Physical Journal H*. <https://doi.org/10.1140/epjh/e2012-20064-5>
- [38].Arons, A. B., & Peppard, M. B. (1965). Einstein's Proposal of the Photon Concept—a Translation of the Annalen der Physik Paper of 1905. *American Journal of Physics*, 33(5). <https://doi.org/10.1119/1.1971542>
- [39].Bharti, A., & Goyal, N. (2019). Fundamental of Synchrotron Radiations. In *Synchrotron Radiation - Useful and Interesting Applications*. <https://doi.org/10.5772/intechopen.82202>
- [40].KIRKPATRICK, P., & BAEZ, A. V. (1948). Formation of optical images by X-rays. *Journal of the Optical Society of America*. <https://doi.org/10.1364/JOSA.38.000766>
- [41].Attwood, D. (2000). Soft X-Rays and Extreme Ultraviolet Radiation: Principles and Applications. *System*.

- [42]. Goodman, J. W. (2005). Introduction to Fourier Optics 3ed. *Roberts & Company Publishers*.
- [43]. Hecht, E. (2002). *Optics*. Addison-Wesley. Retrieved from <https://books.google.co.uk/books?id=7aG6QgAACAAJ>
- [44]. Cooley, J. W., & Tukey, J. W. (1965). An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*. <https://doi.org/10.2307/2003354>
- [45]. Fienup, J. R. (1978). Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters*, 3(1), 27. <https://doi.org/10.1364/OL.3.000027>
- [46]. Sayre, D. (1952). Some implications of a theorem due to Shannon. *Acta Crystallographica*. <https://doi.org/10.1107/s0365110x52002276>
- [47]. Bates, R. H. T. (1982). Fourier phase problems are uniquely solvable in more than one dimensional: underlying theory. In *Optik (Jena)* (Vol. 61).
- [48]. Nugent, K. A., Peele, A. G., Quiney, H. M., & Chapman, H. N. (2005). Diffraction with wavefront curvature: A path to unique phase recovery. *Acta Crystallographica Section A: Foundations of Crystallography*, 61(3). <https://doi.org/10.1107/S010876730501055X>
- [49]. Rodenburg, J., & Maiden, A. (2019). *Ptychography*. *Springer Handbooks*. [https://doi.org/10.1007/978-3-030-00069-1\\_17](https://doi.org/10.1007/978-3-030-00069-1_17)
- [50]. Rodenburg, J. M., & Faulkner, H. M. L. (2004). A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20), 4795–4797. <https://doi.org/10.1063/1.1823034>
- [51]. Hoppe, W. (1969). Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten. *Acta Crystallographica Section A*, 25(4), 508–514. <https://doi.org/https://doi.org/10.1107/S0567739469001069>
- [52]. Hegerl, R., & Hoppe, W. (1970). Dynamische Theorie der Kristallstrukturanalyse durch Elektronenbeugung im inhomogenen Primärstrahlwellenfeld. *Berichte Der Bunsengesellschaft Für Physikalische Chemie*, 74(11), 1148–1154. <https://doi.org/10.1002/bbpc.19700741112>
- [53]. Rodenburg, J. M. (1989). The phase problem, microdiffraction and wavelength-limited resolution - a discussion. *Ultramicroscopy*. [https://doi.org/10.1016/0304-3991\(89\)90009-0](https://doi.org/10.1016/0304-3991(89)90009-0)
- [54]. McCallum, B. C., & Rodenburg, J. M. (1992). Two-dimensional demonstration of Wigner phase-retrieval microscopy in the STEM configuration. *Ultramicroscopy*, 45(3–4), 371–380. [https://doi.org/10.1016/0304-3991\(92\)90149-E](https://doi.org/10.1016/0304-3991(92)90149-E)
- [55]. Chapman, H. N. (1996). Phase-retrieval X-ray microscopy by Wigner-distribution deconvolution. *Ultramicroscopy*. [https://doi.org/10.1016/S0304-3991\(96\)00084-8](https://doi.org/10.1016/S0304-3991(96)00084-8)

- [56].Nellist, P. D., McCallum, B. C., & Rodenburg, J. M. (1995). Resolution beyond the “information limit” in transmission electron microscopy. *Nature*, 374(6523), 630–632. <https://doi.org/10.1038/374630a0>
- [57].Pennycook, T. J., Lupini, A. R., Yang, H., Murfitt, M. F., Jones, L., & Nellist, P. D. (2015). Efficient phase contrast imaging in STEM using a pixelated detector. Part I: Experimental demonstration at atomic resolution. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2014.09.013>
- [58].Yang, H., Pennycook, T. J., & Nellist, P. D. (2015). Efficient phase contrast imaging in STEM using a pixelated detector. Part II: Optimisation of imaging conditions. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2014.10.013>
- [59].Yang, H., MacLaren, I., Jones, L., Martinez, G. T., Simson, M., Huth, M., ... Nellist, P. D. (2017). Electron ptychographic phase imaging of light elements in crystalline materials using Wigner distribution deconvolution. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2017.02.006>
- [60].Faulkner, H. M. L., & Rodenburg, J. M. (2004). Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm. *Physical Review Letters*. <https://doi.org/10.1103/PhysRevLett.93.023903>
- [61].Rodenburg, J. M., Hurst, A. C., & Cullis, A. G. (2007). Transmission microscopy without lenses for objects of unlimited size. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2006.07.007>
- [62].Thibault, P., Dierolf, M., Bunk, O., Menzel, A., & Pfeiffer, F. (2009). Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2008.12.011>
- [63].Rodenburg, J. M., Hurst, A. C., Cullis, A. G., Dobson, B. R., Pfeiffer, F., Bunk, O., ... Johnson, I. (2007). Hard-X-ray lensless imaging of extended objects. *Physical Review Letters*. <https://doi.org/10.1103/PhysRevLett.98.034801>
- [64].Thibault, P., Dierolf, M., Menzel, A., Bunk, O., David, C., & Pfeiffer, F. (2008). High-resolution scanning X-ray diffraction microscopy. *Science*. <https://doi.org/10.1126/science.1158573>
- [65].Hüe, F., Rodenburg, J. M., Maiden, A. M., Sweeney, F., & Midgley, P. A. (2010). Wave-front phase retrieval in transmission electron microscopy via ptychography. *Physical Review B - Condensed Matter and Materials Physics*. <https://doi.org/10.1103/PhysRevB.82.121415>
- [66].Humphry, M. J., Kraus, B., Hurst, A. C., Maiden, A. M., & Rodenburg, J. M. (2012). Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging. *Nature Communications*. <https://doi.org/10.1038/ncomms1733>

- [67]. Guizar-Sicairos, M., & Fienup, J. R. (2008). Phase retrieval with transverse translation diversity: a nonlinear optimization approach. *Optics Express*, 16(10). <https://doi.org/10.1364/oe.16.007264>
- [68]. Thibault, P., Dierolf, M., Menzel, A., Bunk, O., David, C., & Pfeiffer, F. (2008). High-resolution scanning x-ray diffraction microscopy. *Science (New York, N.Y.)*, 321(5887), 379–382. <https://doi.org/10.1126/science.1158573>
- [69]. Thibault, P., Dierolf, M., Bunk, O., Menzel, A., & Pfeiffer, F. (2009). Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*. <https://doi.org/10.1016/j.ultramic.2008.12.011>
- [70]. Elser, V., Rankenburg, I., & Thibault, P. (2007). Searching with iterated maps. *Proceedings of the National Academy of Sciences of the United States of America*. <https://doi.org/10.1073/pnas.0606359104>
- [71]. Hesse, R., Russell Luke, D., Sabach, S., & Tam, M. K. (2015). Proximal heterogeneous block implicit-explicit method and application to blind ptychographic diffraction imaging. *SIAM Journal on Imaging Sciences*. <https://doi.org/10.1137/14098168X>
- [72]. Maiden, A. M., Rodenburg, J. M., & Humphry, M. J. (2010). A new method of high resolution, quantitative phase scanning microscopy. In *Scanning Microscopy 2010* (Vol. 7729). <https://doi.org/10.1117/12.853339>
- [73]. Claus, D., Maiden, A. M., Zhang, F., Sweeney, F. G. R., Humphry, M. J., Schluesener, H., & Rodenburg, J. M. (2012). Quantitative phase contrast optimised cancerous cell differentiation via ptychography. *Optics Express*, 20(9). <https://doi.org/10.1364/oe.20.009911>
- [74]. Giewekemeyera, K., Thibault, P., Kalbfleisch, S., Beerlink, A., Kewish, C. M., Dierolf, M., ... Salditta, T. (2010). Quantitative biological imaging by ptychographic x-ray diffraction microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 107(2). <https://doi.org/10.1073/pnas.0905846107>
- [75]. Hübner, F., Rodenburg, J. M., Maiden, A. M., & Midgley, P. A. (2011). Extended ptychography in the transmission electron microscope: Possibilities and limitations. *Ultramicroscopy*, 111(8). <https://doi.org/10.1016/j.ultramic.2011.02.005>
- [76]. Humphry, M. J., Kraus, B., Hurst, A. C., Maiden, A. M., & Rodenburg, J. M. (2012). Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging. *Nature Communications*. <https://doi.org/10.1038/ncomms1733>
- [77]. Luke, D. R. (2005). Relaxed averaged alternating reflections for diffraction imaging. *Inverse Problems*, 21(1). <https://doi.org/10.1088/0266-5611/21/1/004>
- [78]. Sayre, D. (1952). Some implications of a theorem due to Shannon. *Acta Crystallographica*. <https://doi.org/10.1107/s0365110x52002276>

- [79].Miao, J., Charalambous, P., Kirz, J., & Sayre, D. (1999). Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature*. <https://doi.org/10.1038/22498>
- [80].Miao, J., & Sayre, D. (2000). On possible extensions of X-ray crystallography through diffraction-pattern oversampling. *Acta Crystallographica Section A: Foundations of Crystallography*, 56(6). <https://doi.org/10.1107/S010876730001031X>
- [81].Bunk, O., Dierolf, M., Kynde, S., Johnson, I., Marti, O., & Pfeiffer, F. (2008). Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy*, 108(5), 481–487. <https://doi.org/10.1016/j.ultramic.2007.08.003>
- [82].Edo, T. B., Batey, D. J., Maiden, A. M., Rau, C., Wagner, U., Pešić, Z. D., ... Rodenburg, J. M. (2013). Sampling in x-ray ptychography. *Physical Review A - Atomic, Molecular, and Optical Physics*. <https://doi.org/10.1103/PhysRevA.87.053850>
- [83].da Silva, J. C., & Menzel, A. (2015). Elementary signals in ptychography. *Optics Express*, 23(26). <https://doi.org/10.1364/oe.23.033812>
- [84].Edo, T. B., Zhang, F., & Rodenburg, J. M. (2010). Resolution improvement in coherent diffractive imaging (ptychography). *Scanning Microscopy 2010*, 7729(June 2010), 77291H. <https://doi.org/10.1117/12.853346>
- [84].Kirkland, A. I., Saxton, W. O., Chau, K. L., Tsuno, K., & Kawasaki, M. (1995). Super-resolution by aperture synthesis: tilt series reconstruction in CTEM. *Ultramicroscopy*, 57(4), 355–374. [https://doi.org/10.1016/0304-3991\(94\)00191-O](https://doi.org/10.1016/0304-3991(94)00191-O)
- [86].Maiden, A. M., Humphry, M. J., Zhang, F., & Rodenburg, J. M. (2011). Superresolution imaging via ptychography. *Journal of the Optical Society of America A*, 28(4), 604–612. <https://doi.org/10.1364/JOSAA.28.000604>
- [87].Batey, D. J., Edo, T. B., Rau, C., Wagner, U., Pešić, Z. D., Waigh, T. A., & Rodenburg, J. M. (2014). Reciprocal-space up-sampling from real-space oversampling in x-ray ptychography. *Physical Review A - Atomic, Molecular, and Optical Physics*. <https://doi.org/10.1103/PhysRevA.89.043812>
- [88].Putkunz, C. T., D'Alfonso, A. J., Morgan, A. J., Weyland, M., Dwyer, C., Bourgeois, L., ... Allen, L. J. (2012). Atom-scale ptychographic electron diffractive imaging of boron nitride cones. *Physical Review Letters*, 108(7). <https://doi.org/10.1103/PhysRevLett.108.073901>
- [89].Schropp, A., Boye, P., Goldschmidt, A., Hönig, S., Hoppe, R., Patommel, J., ... Schroer, C. G. (2011). Non-destructive and quantitative imaging of a nano-structured microchip by ptychographic hard X-ray scanning microscopy. *Journal of Microscopy*, 241(1). <https://doi.org/10.1111/j.1365-2818.2010.03453.x>
- [90].Takahashi, Y., Suzuki, A., Zettsu, N., Kohmura, Y., Senba, Y., Ohashi, H., ... Ishikawa, T. (2011). Towards high-resolution ptychographic x-ray diffraction microscopy. *Physical Review B - Condensed Matter and Materials Physics*, 83(21). <https://doi.org/10.1103/PhysRevB.83.214109>

- [91].Tripathi, A., McNulty, I., & Shpyrko, O. G. (2014). Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods. *Optics Express*, 22(2). <https://doi.org/10.1364/oe.22.001452>
- [92].Zhang, F., Peterson, I., Vila-Comamala, J., Diaz, A., Berenguer, F., Bean, R., ... Rodenburg, J. M. (2013). Translation position determination in ptychographic coherent diffraction imaging. *Optics Express*, 21(11). <https://doi.org/10.1364/oe.21.013592>
- [93].Maiden, A. M., Humphry, M. J., Sarahan, M. C., Kraus, B., & Rodenburg, J. M. (2012). An annealing algorithm to correct positioning errors in ptychography. *Ultramicroscopy*, 120, 64–72. <https://doi.org/10.1016/j.ultramic.2012.06.001>
- [94].Koho, S., Tortarolo, G., Castello, M., Deguchi, T., Diaspro, A., & Vicidomini, G. (2019). Fourier ring correlation simplifies image restoration in fluorescence microscopy. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-11024-z>
- [95].Banterle, N., Bui, K. H., Lemke, E. A., & Beck, M. (2013). Fourier ring correlation as a resolution criterion for super-resolution microscopy. *Journal of Structural Biology*, 183(3), 363–367. <https://doi.org/10.1016/j.jsb.2013.05.004>
- [96].Saxton, W. O., & Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of Microscopy*, 127(2), 127–138. <https://doi.org/10.1111/j.1365-2818.1982.tb00405.x>
- [97].Van Heel, M., & Schatz, M. (2005). Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3), 250–262. <https://doi.org/10.1016/j.jsb.2005.05.009>
- [98].Zheng, G., Horstmeyer, R., & Yang, C. (2013). Wide-field, high-resolution Fourier ptychographic microscopy. *Nature Photonics*, 7, 739. Retrieved from <http://dx.doi.org/10.1038/nphoton.2013.187>
- [99].Konda, P. C., Loetgering, L., Zhou, K. C., Xu, S., Harvey, A. R., & Horstmeyer, R. (2020). Fourier ptychography: current applications and future promises. *Optics Express*, 28(7). <https://doi.org/10.1364/oe.386168>
- [100].Kellman, M. R., Bostan, E., Repina, N., & Waller, L. (2018). Physics-based learned design: Optimized coded-illumination for quantitative phase imaging. *ArXiv*. <https://doi.org/10.1109/tci.2019.2905434>
- [101].Chowdhury, S., & Izatt, J. (2013). Structured illumination quantitative phase microscopy for enhanced resolution amplitude and phase imaging. *Biomedical Optics Express*, 4(10). <https://doi.org/10.1364/boe.4.001795>
- [102].Guo, K., Dong, S., Nanda, P., & Zheng, G. (2015). Optimization of sampling pattern and the design of Fourier ptychographic illuminator. *Optics Express*, 23(5). <https://doi.org/10.1364/oe.23.006171>
- [103].Li, P., Batey, D. J., Edo, T. B., & Rodenburg, J. M. (2015). Separation of three-dimensional scattering effects in tilt-series Fourier ptychography. *Ultramicroscopy*, 158. <https://doi.org/10.1016/j.ultramic.2015.06.010>

- [104]. Tian, L., & Waller, L. (2015). 3D Fourier Ptychographic imaging from light field measurements in an LED array microscope. <https://doi.org/10.1364/cosi.2015.ct3f.2>
- [105]. Born, M., & Wolf, E. (1999). *Principles of Optics 7th edition. Principles of Optics Electromagnetic Theory of Propagation Interference and Diffraction of Light*. [https://doi.org/10.1016/S0030-3992\(00\)00061-X](https://doi.org/10.1016/S0030-3992(00)00061-X)
- [106] Mandel, L., & Wolf, E. (1965). Coherence properties of optical fields. *Reviews of Modern Physics*. <https://doi.org/10.1103/RevModPhys.37.231>
- [107] Saleh, B. E. A., & Teich, M. C. (2007). *Fundamentals of Photonics, 2nd Edition*. Wiley.
- [108] Michelson, A. A., & Morley, E. W. (1887). On the relative motion of the Earth and the luminiferous ether. *American Journal of Science*. <https://doi.org/10.2475/ajs.s3-34.203.333>
- [109] Robertson, H. P. (1949). Postulate versus observation in the special theory of relativity. *Reviews of Modern Physics*. <https://doi.org/10.1103/RevModPhys.21.378>
- [110] Young, T. (1802). On the Theory of Light and Colours. *Phil. Trans. R. Soc. Lond*. <https://doi.org/10.1098/rstl.1802.0004>
- [111] Lederman, L. M. (2011). *Quantum Physics for Poets*. Prometheus. Retrieved from [https://books.google.co.uk/books?id=qY\\_yOwHg\\_WYC](https://books.google.co.uk/books?id=qY_yOwHg_WYC)
- [112] Feynman, R. P., Leighton, R. B., Sands, M., & Lindsay, R. B. (1966). The Feynman Lectures on Physics, Vol. 3: Quantum Mechanics. *Physics Today*. <https://doi.org/10.1063/1.3047826>
- [113] Zernike, F. (1938). The concept of degree of coherence and its application to optical problems. *Physica*. [https://doi.org/10.1016/S0031-8914\(38\)80203-2](https://doi.org/10.1016/S0031-8914(38)80203-2)
- [114] van Cittert, P. H. (1934). Die Wahrscheinliche Schwingungsverteilung in Einer von Einer Lichtquelle Direkt Oder Mittels Einer Linse Beleuchteten Ebene. *Physica*. [https://doi.org/10.1016/S0031-8914\(34\)90026-4](https://doi.org/10.1016/S0031-8914(34)90026-4)
- [115] James, J. F. (2011). *A student's guide to fourier transforms: With applications in physics and engineering, third edition*. <https://doi.org/10.1017/CBO9780511762307>
- [116]. Mandel, L., & Wolf, E. (1965). Coherence properties of optical fields. *Reviews of Modern Physics*. <https://doi.org/10.1103/RevModPhys.37.231>
- [117] Li, P., Edo, T., Batey, D., Rodenburg, J., & Maiden, A. (2016). Breaking ambiguities in mixed state ptychography. *Optics Express*, 24(8), 9038. <https://doi.org/10.1364/OE.24.009038>
- [118] Banerjee, S., & Roy, A. (2014). *Linear Algebra and Matrix Analysis for Statistics*. Taylor & Francis. Retrieved from <https://books.google.co.uk/books?id=iIOhAwAAQBAJ>

- [119]. Hall, C. E. (1953). *Introduction to Electron Microscopy*. McGraw-Hill. Retrieved from <https://books.google.co.uk/books?id=oRIRAAAAMAAJ>
- [120] Goodhew, P. J., Humphreys, J. F., & Beanland, R. (2001). Electron Microscopy and Analysis. *Annals of Physics*, 54, 254. Retrieved from <http://books.google.com/books?hl=en&lr=&id=zuxPIVmGLGsC&oi=fnd&pg=PR9&dq=Electron+microscopy+and+analysis&ots=-U-Itc0MZJ&sig=ibc15cL-3PUTTnXb-0VEbsmESEw>
- [121] Egerton, R. F. (2013). Control of radiation damage in the TEM. *Ultramicroscopy*, 127, 100–108. <https://doi.org/10.1016/j.ultramic.2012.07.006>
- [122] Egerton, R. F., McLeod, R., Wang, F., & Malac, M. (2010). Basic questions related to electron-induced sputtering in the TEM. *Ultramicroscopy*, 110(8), 991–997. <https://doi.org/10.1016/j.ultramic.2009.11.003>
- [123] Banhart, Florian. “Formation and Transformation of Carbon Nanoparticles under Electron Irradiation.” *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, vol. 362, no. 1823, 2004, pp. 2205–2222. JSTOR, [www.jstor.org/stable/4142439](http://www.jstor.org/stable/4142439). Accessed 14 Feb. 2021.
- [124]. Hobbs, L. W. (1979). Radiation Effects in Analysis of Inorganic Specimens by TEM. In J. J. Hren, J. I. Goldstein, & D. C. Joy (Eds.), *Introduction to Analytical Electron Microscopy* (pp. 437–480). Boston, MA: Springer US. [https://doi.org/10.1007/978-1-4757-5581-7\\_17](https://doi.org/10.1007/978-1-4757-5581-7_17)
- [125] Williams, D. B., & Carter, C. B. (2009). *Transmission Electron Microscopy: A Textbook for Materials Science*. Materials Science (Vol. 1–4). [https://doi.org/10.1007/978-1-61779-415-5\\_23](https://doi.org/10.1007/978-1-61779-415-5_23)
- [126] Medlin, D. L., & Howitt, D. G. (1991). The role of sputtering and displacement damage in the electron SCRIBE process. *Philosophical Magazine Letters*, 64(3), 133–141. <https://doi.org/10.1080/09500839108214679>
- [127]. Cherns, D., Minter, F. J., & Nelson, R. S. (1976). Sputtering in the high voltage electron microscope. *Nuclear Instruments and Methods*, 132, 369–376. [https://doi.org/10.1016/0029-554X\(76\)90760-6](https://doi.org/10.1016/0029-554X(76)90760-6)
- [128] Cherns, D., Finnis, M. W., & Matthews, M. D. (1977). Sputtering of gold foils in a high voltage electron microscope A comparison of theory and experiment. *Philosophical Magazine*, 35(3), 693–714. <https://doi.org/10.1080/14786437708235999>
- [129] Le Caër, S. (2011). Water Radiolysis: Influence of Oxide Surfaces on H<sub>2</sub> Production under Ionizing Radiation. *Water*, 3(1), 235–253. <https://doi.org/10.3390/w3010235>
- [130] Dzaugis, M. E., Spivack, A. J., & D’Hondt, S. (2015). A quantitative model of water radiolysis and chemical production rates near radionuclide-containing solids. *Radiation Physics and Chemistry*, 115, 127–134. <https://doi.org/https://doi.org/10.1016/j.radphyschem.2015.06.011>

- [131] Hamill, W. H. (1969). Model for the radiolysis of water. *The Journal of Physical Chemistry*, 73(5), 1341–1347. <https://doi.org/10.1021/j100725a027>
- [132] Egerton, R. F. (2012). Mechanisms of radiation damage in beam-sensitive specimens, for TEM accelerating voltages between 10 and 300 kV. *Microscopy Research and Technique*, 75(11), 1550–1556. <https://doi.org/10.1002/jemt.22099>
- [133] Egerton, R. F. (2012). The importance of dose-limited resolution, *110*(2010), 1–2.
- [134] Natusch, M. K. H., Humphreys, C. J., Menon, N., & Krivanek, O. L. (1999). Experimental and theoretical study of the detection limits in electron energy-loss spectroscopy. *Micron*, 30(2), 173–183. [https://doi.org/https://doi.org/10.1016/S0968-4328\(99\)00021-9](https://doi.org/https://doi.org/10.1016/S0968-4328(99)00021-9)
- [135] Bradley, C. R., & Zaluzec, N. J. (1989). Atomic sputtering in the Analytical Electron Microscope. *Ultramicroscopy*, 28(1), 335–338. [https://doi.org/https://doi.org/10.1016/0304-3991\(89\)90320-3](https://doi.org/https://doi.org/10.1016/0304-3991(89)90320-3)
- [136] Salisbury, I. G., Timsit, R. S., Berger, S. D., & Humphreys, C. J. (1984). Nanometer scale electron beam lithography in inorganic materials. *Applied Physics Letters*, 45(12), 1289–1291. <https://doi.org/10.1063/1.95115>
- [137] Cazaux, J. (1995). Correlations between ionization radiation damage and charging effects in transmission electron microscopy. *Ultramicroscopy*, 60(3), 411–425. [https://doi.org/https://doi.org/10.1016/0304-3991\(95\)00077-1](https://doi.org/https://doi.org/10.1016/0304-3991(95)00077-1)
- [138] Fryer, J. R. (1987). The effect of dose rate on imaging aromatic organic crystals. *Ultramicroscopy*, 23(3), 321–327. [https://doi.org/https://doi.org/10.1016/0304-3991\(87\)90242-7](https://doi.org/https://doi.org/10.1016/0304-3991(87)90242-7)
- [139] Varlot, K., Martin, J. M., Quet, C., & Kihn, Y. (1997). Towards sub-nanometer scale EELS analysis of polymers in the TEM. *Ultramicroscopy*, 68(2), 123–133. [https://doi.org/https://doi.org/10.1016/S0304-3991\(97\)00019-3](https://doi.org/https://doi.org/10.1016/S0304-3991(97)00019-3)
- [140] Egerton, R. F., Wang, F., & Crozier, P. A. (2006). Beam-induced damage to thin specimens in an intense electron probe. *Microscopy and Microanalysis : The Official Journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada*, 12(1), 65–71. <https://doi.org/10.1017/S1431927606060065>
- [141] Fryer, J. R., McNee, C., & Holland, F. M. (1984). The further reduction of radiation damage in the electron microscope. *Ultramicroscopy*, 14(4), 357–358. [https://doi.org/https://doi.org/10.1016/0304-3991\(84\)90221-3](https://doi.org/https://doi.org/10.1016/0304-3991(84)90221-3)
- [142] Fryer, J.R., & Holland, F. (1984). High resolution electron microscopy of molecular crystals III. Radiation process at room temperature. *Proc.R.Soc.lond.A*, 353-369. <http://doi.org/10.1098/rspa.1984.0062>
- [143] Knoll, G. F. (2000). *Radiation Detection and Measurement*. Wiley. Retrieved from <https://books.google.co.uk/books?id=HKBVAAAAMAAJ>

- [144] Ahmed, S. N. (2015). 2 - Interaction of radiation with matter. In S. N. Ahmed (Ed.), *Physics and Engineering of Radiation Detection (Second Edition)* (Second Edition, pp. 65–155). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-801363-2.00002-4>
- [145] Rutherford, E., & Geiger, H. (1908). An electrical method of counting the number of particles from radio-active substances. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 81(546), 141–161. <https://doi.org/10.1098/rspa.1908.0065>
- [146] Ekelöf, T., Séguinot, J., Tocqueville, J., & Ypsilantis, T. (1981). The Cerenkov Ring-Imaging Detector: Recent Progress and Future Development. *Physica Scripta*, 23(4B), 718–726. <https://doi.org/10.1088/0031-8949/23/4b/023>
- [147] Glaser, D. A. (1952). Some Effects of Ionizing Radiation on the Formation of Bubbles in Liquids. *Phys. Rev.*, 87(4), 665. <https://doi.org/10.1103/PhysRev.87.665>
- [148] Ahmed, S. N. (2015). 4 - Liquid-filled detectors. In S. N. Ahmed (Ed.), *Physics and Engineering of Radiation Detection (Second Edition)* (Second Edition, pp. 233–258). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-12-801363-2.00004-8>
- [149] Ramo, S. (1939). Currents Induced by Electron Motion. *Proceedings of the IRE*, 27(9), 584–585. <https://doi.org/10.1109/JRPROC.1939.228757>

