The University Of Sheffield.

Investigations and improvements in ptychographic imaging

Peng Li

A thesis submitted for the degree of

Doctor of Philosophy

The University of Sheffield

Department of Electronic and Electrical Engineering

August 2016

Contents

Abstract

This thesis has been devoted to investigate and improve ptychography, which is a newly developed coherent diffractive imaging technique that can achieve quantitative imaging (both modulus and phase) at diffraction-limited resolution without imaging lenses.

In particular, this thesis has first looked into two solutions of partial coherence in ptychography: the Wigner distribution deconvolution method (WDDM) and mixed state decomposition. WDDM is a non-iterative solution and with it partial coherence was first mathematically demonstrated solvable. We have improved the performance of WDDM, especially in the presence of noise, by proposing three tools that can be used together: a projection strategy, design of a favourable probe, and an iterative method. Furthermore, the reconstruction of spatial partial coherence via WDDM has been successfully demonstrated using a model calculation for the first time.

Mixed state decomposition is an iterative solution. It provides much more flexibility and is able to solve any experimental instability (not just partial coherence) that can be modelled as a set of mutually orthogonal states. According to the formation of the mixed states, it can be divided into spatially mixed state ptychography and temporally mixed state ptychography. For spatially mixed state reconstruction, we have mathematically and experimentally demonstrated an inherent linear ambiguity in the reconstructions and also that the ambiguity can be broken by using an orthogonality constraint or phase-only constraint. Besides, the effects of a diffused probe on the reconstructions have been investigated using a spatial partial coherent x-ray experiment. For temporally mixed state ptychography, we have mathematically and experimentally demonstrated the breakdown of the linear ambiguity. In addition, an iterative algorithm

has been proposed to remove the static background noise from the measurements by treating the background as the diffraction pattern from an extra temporal state.

Moreover, this thesis has also explored two ways to extend ptychography for three-dimensional (3D) imaging: multislice ptychography and ptychographic tomography. The multislice method has already been introduced into ptychography to provide 3D information before. In this thesis, we have further extended it into a Fourier variant of ptychography – Fourier ptychography – by applying a parallel update for the aperture reconstruction and reforming the iterative algorithm to involve the specimen plane. Also, the reconstruction resolution has been discussed via the Ewald sphere construction and demonstrated via model calculations.

Ptychographic tomography utilises ptychography to acquire 2D projection images at different orientations and makes use of tomography to achieve isotropic 3D reconstructions at high resolution. In this thesis, we have demonstrated this technique step by step via an x-ray experiment and shown how the inherent ptychographic reconstruction ambiguities are removed prior to the tomographic reconstruction. The possibility of electron ptychographic tomography is also discussed based on the scale calculation with the x-ray experiment.

Declaration

I confirm that this thesis is my own work based on the research carried out in the Ultimate Imaging Centre, Department of Electronic and Electrical Engineering, the University of Sheffield, United Kingdom. The materials in this thesis have not been submitted in part, or in whole for any other degree or qualification and references are made to materials that are not the work of the author.

Peng Li

25th of August 2016

Acknowledgements

I would like to express my sincere thanks to those people who assisted and supported me through the development of this thesis work.

I would like firstly to give my deepest gratitude to my supervisor Professor John Rodenburg for his patient guidance, generous assistance and invaluable advice throughout my PhD research. I have learnt so much from him not just in how to be a professional researcher, but also in how to be an open-minded person. I feel lucky and proud to be his research student. Such wonderful and pleasure experience working with him will keep inspiring and encouraging me in the future.

Furthermore, I would like to thank those wonderful colleagues that I have been working with. In particular, my thanks go to Dr. Andrew Maiden for his assistance with the optical experiments, generous offer of parts of his code and numerous fruitful discussions; Dr. Tega Edo for his extremely helpful advice and support, both academic and non-academic, and his willingness to share and pass on his expertise; Dr. Darren Batey for the insightful and stimulating discussions; Dr. Daniel Clause for his help at the early stage of my PhD. My gratitude also goes to Shaohong Cao, Daniel Johnson and Samuel McDermott for providing such lovely working atmosphere in the office.

I am also very grateful for the opportunities to carry out x-ray experiments in Diamond Light Source in UK and for the help I was offered during the experiments by Christoph Rau, Ulrich Wagner and Aaron Parsons.

Last but not least, I would like to thank my wonderful parents, Defeng Tang and Zongjun Li, and my amazing sister, Rong Li, for their constant support and encouragement during all these years. I would also like to express my special thanks to

my beloved wife, Gaoyan He, and my lovely daughter, Grace Li, for their company and the happiness they bring into my life. I cannot be where I am today without them. This thesis is dedicated to them.

Publications

Parts of the work and results presented in this thesis have led to the following journal paper publications and my contributions are also accordingly given:

1. **Peng Li**, Darren J Batey, Tega B Edo, Aaron D Parsons, Christoph Rau, and John M Rodenburg, "Multiple mode x-ray ptychography using a lens and a fixed diffuser optic," *Journal of optics* 18(5):054008, 2016. DOI:10.1088/2040-8978/18/5/054008

My contributions: substantially involved with the experiments and the theoretical ideas, writing all the codes, processing the experimental data and writing the draft of the paper.

2. **Peng Li**, Tega Edo, Darren Batey, John Rodenburg, and Andrew Maiden, "Breaking ambiguities in mixed state ptychography," *Optics Express* 24(8):9038, 2016. DOI:10.1364/OE.24.009038

My contributions: developing the theory and maths, implementing all the experiments, writing all the codes, processing the experimental data and writing the draft of the paper.

3. **Peng Li**, Darren J. Batey, Tega B. Edo, and John M. Rodenburg, "Separation of three-dimensional scattering effects in tilt-series Fourier ptychography," *Ultramicroscopy* 158, 2015. DOI:10.1016/j.ultramic.2015.06.010

My contributions: substantially involved with the theoretical ideas, implementing all the experiments, writing all the codes, processing the experimental data and writing the draft of the paper.

4. **Peng Li**, Tega B. Edo, and John M. Rodenburg, "Ptychographic inversion via Wigner distribution deconvolution: Noise suppression and probe design," *Ultramicroscopy* 147, 2014. DOI:10.1016/j.ultramic.2014.07.004

My contributions: substantially involved with the theoretical ideas, writing all the codes, simulating all the model calculations and writing the draft of the paper.

Conference publications are listed as follows:

1. **Peng Li**, and John M. Rodenburg, "Mitigation of partial coherence effects in ptychography," *International conference on the state and future of ptychography*, 2013.
2. **Peng Li**, and John M. Rodenburg, "Optimal illumination functions for electron ptychography," *The Electron Microscopy and Analysis Group Conference*, 2013.
3. Andrew M. Maiden, and **Peng Li**, "The Remarkably Flexible Ptychographic Data Set," *Laser Science*, 2015.
4. Shaohong Cao, **Peng Li**, Andrew M. Maiden, and John M. Rodenburg, "Defocus and probe-position coupling in electron ptychography," *European Microscopy Congress*, 2016. (accepted)

1 Introduction

In the context of imaging techniques, resolution has been always the factor that people care much about. In this chapter we will start with the brief review of conventional lens-based imaging and how its resolution is limited according to Abbé's theory. Diffractive imaging is then introduced as a solution to the resolution problem of conventional imaging. Apart from high resolution, another advance of diffractive imaging is that it delivers quantitative phase images, which not only provide much better contrast (compared to the brightness images produced by conventional imaging) but also allow quantitative analysis. Different types of diffractive imaging techniques are reviewed to give a general picture of how iterative ptychography stands out from them.

1.1 Conventional imaging

For a conventional transmission microscope, a typical configuration is schematically shown in Fig. 1.1. A thin specimen is illuminated by a plane wave radiation. The scattered waves are collected by the objective lens and form the spectrum of the specimen at the back focal plane. Then the spectrum re-interferes and evolves into an image at the image plane. The image is usually a magnified reflection of the absorption or attenua-

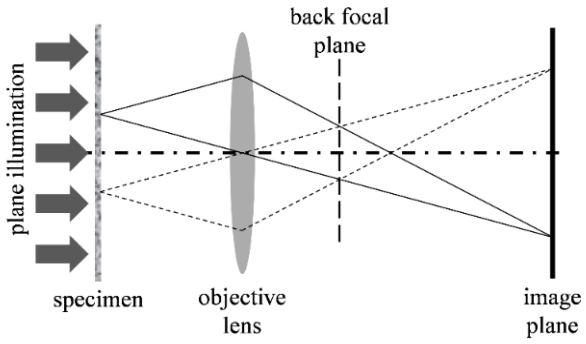

Figure 1.1: A schematic diagram of a conventional transmission microscope.

tion power of the materials in the specimen with the magnification determined by the ratio of the distance from the specimen plane to the lens principal plane (the vertical symmetric axis of the lens) and the distance from the lens principal plane to the image plane.

1.1.1 Abbé's theory

According to Abbé, the formation of an image in a microscope can be broken down into two stages, analysis and synthesis [1]. The analysis stage is the process from the specimen plane to the lens back focal plane. During this stage, the objective lens focuses the parallel diffracted waves into a point at the back focal plane. This process acts exactly like the Fourier transform. The synthesis stage is from the lens back focal plane to the image plane. During this stage, the waves in the back focal, i.e. the diffraction pattern or the spectrum of the specimen, re-interfere and evolve into an image of the specimen at the image plane, which can be considered as another Fourier transform. After two successive Fourier transforms, the image is a central symmetric version of the specimen. The Abbé resolution [2] of the image is given by

$$\gamma = \frac{\lambda}{2n\sin\theta}, \tag{1.1}$$

where λ is the wavelength of the radiation, n is the refractive index of the medium between the objective lens and the specimen, and θ is the half angle subtended by the objective lens. The product $n\sin\theta$ is defined as the numerical aperture (NA) of the objective lens. The resolution is limited by the wavelength of the radiation and the highest scattering angle that can be collected by the objective lens. Abbé's theory provides us guidance to evaluate or improve the performance of a microscope.

1.1.2 Resolution limit

According to Eq. (1.1), the resolution limit is defined by the wavelength of the radiation source and the numerical aperture (NA) of the objective lens. In visible light regime, although good objective lenses with high NAs can be manufactured, the long wavelengths (400nm-700nm) still limit the achievable resolution. To pursue higher resolution, shorter wavelengths, like x-rays and electrons, are desired. However, for these sources good objective lenses are not available. Or if available, they certainly come at the price of an enormous increase in cost and complexity.

For high-resolution x-ray microscopes, a zone plate is the predominant choice for the objective lens [3]. It is made up of a set of concentric rings, known as Fresnel zones, which alternate between opaque and transparent (or between 0 and π phase changes). Light hitting the zone plate will diffract around the opaque zones. The width of zones decreases along the radial direction so that the diffracted light constructively interferes at the desired focus. The resolution of an x-ray microscope is determined by the size of the focus of the Fresnel zone plate used and the size of the focus is of the same order of the outer zone width. For high-resolution x-ray imaging, the manufacture of narrow outer rings of a zone plate becomes very challenging, especially for hard x-rays because of the increased thickness of the zone plate.

In electron microscopes, an electromagnetic field is used as the objective lens to guide the charged electrons [4]. The simplest magnetic lens is a donut-shaped coil through which the electron beam passes, preferably along the axis of the coil. An electric current is passed through the coil to generate the magnetic field, which enables to focus the electron beams. The magnetic field needs to be stable enough to maintain the path difference of the scattered electron beams, so that they can properly re-interfere to form

an image. In 1936, Scherzer showed that round magnetic lenses inevitably have spherical aberration [5], which means that the focusing power increases along the radial direction of the lenses. As a result of spherical aberration, electron beams of different angles are focused to different points along the optic axis, seriously limiting the resolution of electron microscopes. Later, Scherzer also pointed out that a non-round lens is able to compensate for spherical aberration [6, 7], but it was not until 1997 for its successful demonstration [8]. Although great progress has now been made in the manufacture of non-round aberration correctors, the cost of such systems is high and the gain in resolution is limited.

1.2 Diffractive imaging

To overcome the resolution limit imposed by the poor objective lens, an unintuitive way is to get rid of the lens during imaging. From this lens-less setup, rather than obtain the image of the specimen directly, we make some other forms of measurement (like diffraction patterns or holograms), from which we deduce or calculate the image. When the measurement is the diffraction pattern that is related to the exit wave emanating from the specimen by a Fourier transform, we refer to the corresponding imaging method as diffractive imaging. However, a detector is only able to record the intensity that is proportional to the square of the modulus part of the complex diffraction pattern. The information about the phase part is lost and this is referred to as the well-known phase problem. Without the lenses, rephasing the diffraction pattern to form an image is replaced by reconstruction algorithms (or “phase retrieval algorithms”).

In diffractive imaging, the resolution is limited by the NA of the detector (instead of the NA of the objective lens as in conventional imaging) and the wavelength of the radiation. Normally, the NA of the detector is big enough to capture all the diffraction

orders of the specimen, so the resolution is only limited by the wavelength. Apart from high resolution, diffractive imaging also provides quantitative phase information of the specimen. Phase images present much better contrast than the brightness images. It thus allows us to see transparent structures without staining or tagging that would affect the natural state of the specimen under investigation. Conventional methods, like Zernike phase contrast microscopy [9] and differential interference contrast (DIC) or Nomarski microscopy [10], converts the invisible phase shifts experienced by the wave passing through the specimen to intensity changes in the image, hence enhancing contrast. However, the obtained image is often not linearly related to the phase information and show the presence of artefacts, like the halo effects in Zernike phase contrast microscopy and the shadows in DIC microscopy [11]. Diffractive imaging methods, on the other hand, create two images: the modulus image and the phase image. The phase image is linearly related to the optical thickness of the sample and it allows us to quantitatively study the sample [11].

1.2.1 Crystallography

A representative diffractive imaging method would be crystallography that is used to determine the atom arrangement in crystalline objects via analysing the diffraction pattern measured by illuminating the crystals with a coherent beam [12]. The diffraction pattern contains a set of bright spots, called Bragg peaks, resulting from the constructive interference of the waves reflected from a set of atomic planes. Due to the phase problem, a direct reconstruction by applying an inverse Fourier transform is not possible. An iterative process of modelling and refinement is often used to reconstruct the atomic structure [12]. In this process, a hypothesized structure of the crystal is modelled and its diffraction pattern is calculated to compare with the measured one. Refinement of the model is repeated until the calculated diffraction pattern matches the measurement to a

great degree. It is a painstaking process. Fortunately, some *a priori* information about the crystals, like the knowledge of the constituent atoms, can be utilised to generate the initial guess models, so that less iterations are needed to find the solutions and the number of candidate solutions is reduced [12]. One obvious drawback of the technique is the stringent requirement that the sample is crystalline, because not only crystallising a sample can be extremely difficult, but also most samples of interest are not crystalline and cannot be crystallised. The word 'crystal' is derived from the Greek word 'κρύσταλλος' meaning 'ice' or 'rock crystal' and 'graphy' derived from 'γραφή' meaning 'writing' or 'drawing'.

1.2.2 Crystalline ptychography

Walter Hoppe first conceived the concept of ptychography to solve the phase problem in electron microscopy [13, 14, 15, 16, 17]. The basic idea is to let the diffraction orders (i.e. the Bragg peaks) of the crystal interfere and from the interference to work out the phase relation between them. The interference is realised by using a convergent illumination to illuminate the crystal. The convergent illumination is formed by a condenser lens with an aperture at its back focal plane. In the far field, the diffraction pattern is formed by the convolution of the Bragg peaks with the aperture function. If an appropriate size is chosen for the aperture so that in reciprocal space different diffraction orders only overlap (i.e. interfere) with their adjacent ones, then the phase problem is solvable. It should be noted that the condenser lens is used to provide a range of incident angles that only needs to be slightly bigger than the scattering angle between any two diffraction orders. It has much lower requirements as the objective lens does. 'Ptycho' comes from the Greek word 'πτυξ' meaning 'to fold' that refers to the convolution in the diffraction plane.

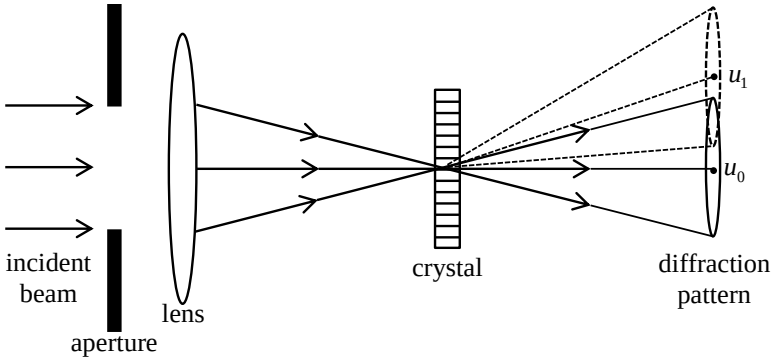

Figure 1.2: The schematic depiction of the crystalline ptychography. Only two diffraction orders are considered here. The lens introduces a range of angles, which is determined by the size of the aperture, into the incident beam and expands each Bragg peak into a circle. When the size of the aperture is properly chosen, the diffraction circles will overlap. From the intensity of the overlap area, the phase relation between the two orders can be recovered.

As shown in Fig. 1.2, if we only consider two diffraction orders u_0 and u_1, whose complex values are denoted as Ψ_{u_0} and Ψ_{u_1}, the intensity at the overlap region can be expressed as

$$\begin{aligned}&\left|A(u-u_0)\Psi_{u_0}+A(u-u_1)\Psi_{u_1}\right|^2 = \\ &\left|A(u-u_0)\Psi_{u_0}\right|^2+\left|A(u-u_1)\Psi_{u_1}\right|^2+2\left|A(u-u_0)\Psi_{u_0}\right|\left|A(u-u_1)\Psi_{u_1}\right|\cos\left(\phi_{u_0}-\phi_{u_1}\right),\end{aligned} \tag{1.2}$$

where u is the coordinate in reciprocal space. It can be calculated by $u=\sin\theta/\lambda=n/a$, where λ is the wavelength of the incident beam, a is the side length of the unit cell of the crystal and θ is the scattering angle of the nth diffraction order. So $u_0=0$ is the position of the zeroth diffraction order and $u_1=1/a$ the position of the first diffraction order. A is the aperture function, and ϕ_{u_0} and ϕ_{u_1} are the phases of Ψ_{u_0} and Ψ_{u_1} respectively. As we can see, the intensity value depends on the phase difference between the two orders. If we further measure the intensities at the non-overlap region, the value of the cosine term can be determined. However due to the even property of the cosine function, the sign of the phase difference remains ambiguous. An easy way to solve this sign ambiguity is to introduce a phase ramp across the diffraction orders by laterally

shifting the illumination relative to the object by a distance of σ, as a result the intensity value at the overlap now becomes

$$\left|A(u-u_0)e^{-i2\pi\sigma(u-u_0)}\Psi_{u_0}+A(u-u_1)e^{-i2\pi\sigma(u-u_1)}\Psi_{u_1}\right|^2=\left|A(u-u_0)\Psi_{u_0}\right|^2+ \\ \left|A(u-u_1)\Psi_{u_1}\right|^2+2\left|A(u-u_0)\Psi_{u_0}\right|\left|A(u-u_1)\Psi_{u_1}\right|\cos\left[\left(\phi_{u_0}-\phi_{u_1}\right)+2\pi\sigma(u_0-u_1)\right]. \quad (1.3)$$

Likewise, we can calculate the value of the cosine term in Eq. (1.3). Meanwhile, the value of $2\pi\sigma(u_0-u_1)=-2\pi\sigma/a$ is known. So the value of $\phi_{u_0}-\phi_{u_1}$ can be completely determined without any ambiguity. However, it should be noted that when the shift σ is equal to an integral multiple of the unit cell size a, the sign ambiguity still remains [18]. In this way, the phase difference between any two overlapped diffraction orders can be obtained. If we assign an arbitrary phase to the zeroth diffraction order, we have the knowledge of the whole complex diffraction pattern. Although the absolute phase is lost, it has no effect on determining the structure of the crystal. Again, this method only applies to crystalline objects and the unit cell size must be known. Moreover, the transfer function of the lens will not be a perfect aperture function in real situations and it will complicate the intensity measurements recorded.

1.2.3 Direct ptychography

Hoppe attempted to extend crystalline ptychography to non-crystalline objects. But since the very simple interference phenomena that occur in crystalline ptychography no longer apply, an explicit solution for all the relevant phases did not at that time appear to be available. Even in his later discussion of these matters, Hoppe seemed to conclude that the aperiodic ptychographic problem was intractable [19]. It was not until 1989 when Bates and Rodenburg [20] proposed a direct (non-iterative) reconstruction algorithm, the Wigner Distribution Deconvolution Method (WDDM), did ptychography become applicable to non-crystalline objects [21, 22, 23, 24]. In this thesis, we refer to

the ptychography solved via WDDM as direct ptychography to differentiate from the ptychography solved by iterative algorithms that will be referred to as iterative ptychography (see section 1.2.6). Direct ptychography requires the specimen to be scanned over a grid of 2D positions separated by the desired resolution in the final reconstruction and a 2D diffraction intensity to be recorded at each scan position, giving rise to a very large 4D dataset. Taking the Fourier transforms of this 4D dataset with respect to the scanning coordinates and the diffraction coordinates (spatial frequency), we have

$$H(\boldsymbol{r},\boldsymbol{\rho}') = \mathcal{F}_{\boldsymbol{r}'}^{-1}\mathcal{F}_{\boldsymbol{\rho}}\left\{\left|M(\boldsymbol{r}',\boldsymbol{\rho})\right|^2\right\} = \chi_a(\boldsymbol{r},-\boldsymbol{\rho}')\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}') \tag{1.4}$$

with the quantity χ_f called the Wigner Distribution Function (WDF) [25] of f and defined by $\chi_f(\boldsymbol{r},\boldsymbol{\rho}') = \mathcal{F}_{\boldsymbol{\rho}}\left\{f(\boldsymbol{r}+\boldsymbol{\rho})f^*(\boldsymbol{\rho})\right\}$, where $\boldsymbol{r}'$ and $\boldsymbol{\rho}$ are respectively the diffraction coordinates and the scanning coordinates, a and ψ are respectively the illumination function and the specimen function, $\left|M(\boldsymbol{r}',\boldsymbol{\rho})\right|^2$ is the recorded 4D intensity dataset, $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transform and the inverse Fourier transform respectively with the subscript indicating the coordinate they operate upon, and the superscript asterisk represents complex conjugate. A deconvolution step allows us to separate out the effects of the illumination (e.g. aberrations) from the specimen by

$$\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}') = \frac{\chi_a^*(\boldsymbol{r},-\boldsymbol{\rho}')H(\boldsymbol{r},\boldsymbol{\rho}')}{\left|\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')\right|^2+\epsilon}, \tag{1.5}$$

where ϵ is a small constant to avoid division by zero. A further Fourier transform of $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ with respect to $\boldsymbol{r}$ gives us the complex product between every point in the diffraction pattern and the complex conjugate of every other point in the diffraction pattern:

$$D(\boldsymbol{r}', \boldsymbol{\rho}') = \mathcal{F}_{\boldsymbol{r}} \left\{ \chi_{\psi}(\boldsymbol{r}, \boldsymbol{\rho}') \right\} = \Psi^*(\boldsymbol{r}' - \boldsymbol{\rho}')\Psi(\boldsymbol{r}'), \tag{1.6}$$

where Ψ is the Fourier transform of ψ, i.e. the diffraction pattern of the specimen. This quantity includes all the information of the interference between any two diffraction orders separated by $-\boldsymbol{\rho}'$ and it allows us to solve for the phase problem.

Due to the partial coherence, experimental instabilities or the finite extent of the lens, the frequency transmitted through the system is limited along $\boldsymbol{\rho}'$ [18, 24, 26, 27]. This is why the bright field image of a conventional microscope has an information limit. For $D(\boldsymbol{r}', \boldsymbol{\rho}')$, even though along $\boldsymbol{\rho}'$ the frequency information beyond certain limit is lost, it can be recovered along $\boldsymbol{r}'$ and eventually an aberration-free image of the specimen can be reconstructed at diffraction-limited resolution [26]. Back in the 1990s, this frequency information recovery was realised by a method called stepping out in which a route containing all the specimen frequency information is chosen and the frequencies along the chosen route can be recovered sequentially one after another after assigning an arbitrary phase to the start point [18, 21, 22, 26]. However, it only makes use of a small fraction of the 4D data $D(\boldsymbol{r}', \boldsymbol{\rho}')$ and the error will accumulate as the stepping out procedure progresses if a wrong data segment is chosen to start with [26]. A full derivation of WDDM and detailed introduction of a new stepping out strategy will be given in chapter 3. Nevertheless, the intensive scan and computation needed for this method make it less popular compared to iterative ptychography.

1.2.4 Fourier holography

Holography is a very elegant lens-less imaging method. It was invented by Gabor as a way to circumvent the resolution limit imposed by the spherical aberration in electron microscope [28]. In his invention, a convergent electron beam with good coherence illuminates a specimen that is highly transparent so that most part of the incident

electron beam goes straight through without interacting with the specimen. This part of electron beam is called reference wave. In the recording plane, the part of beam that is scattered by the specimen interferes with the reference wave. In this way, the phase information of the scattered wave is encoded as intensity variations and the record of them is called hologram. The specimen can be reconstructed by illuminating the hologram with the reference wave. However, the reconstruction is corrupted by its conjugate symmetric twin image because of the inline setup that means the reference wave coincides with the scattered wave during the hologram recording process. Leith and Upatnieks discovered that a small angle between the reference wave and the scattered wave can be used to avoid the overlap of the twin images, hence separating out the object reconstruction [29]. This optical arrangement is called the off-axis configuration. The word 'holo' originates from the Greek word 'ὅλος' meaning 'whole'.

The analogous form of optical holography in diffractive imaging is called Fourier holography (or Fourier-transform holography) [30, 31]. It requires a scatterer (normally a small pinhole) isolated from the specimen to be illuminated together with the specimen by a coherent plane wave. The resulting diffraction pattern is recorded in the far field. Here the isolated scatterer is the counterpart of the off-axis reference wave in optical holography and the measured diffraction pattern is the counterpart of hologram. The image reconstruction is realised by taking a Fourier transform of the measured diffraction pattern, which is analogous to illuminating the hologram with a reference plane wave and observing the reconstruction in the far field. The Fourier transform results in the autocorrelation of whatever in the specimen plane, i.e. the specimen plus the reference scatterer, according to the autocorrelation theorem in section 2.1.2. A simple numerical calculation is shown in Fig. 1.3 to demonstrate the implementation of Fourier holography.

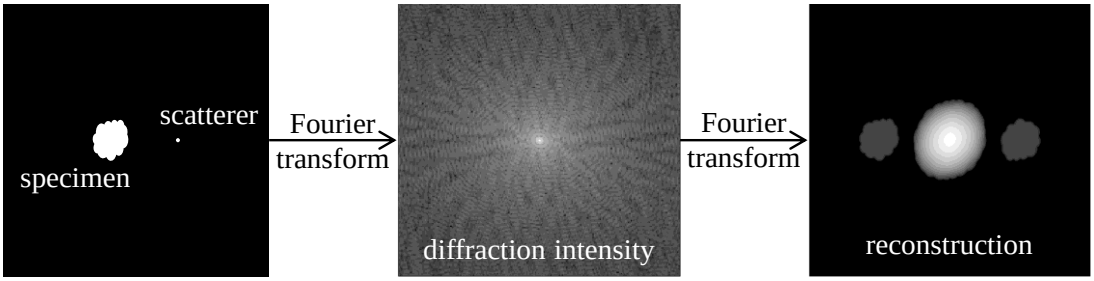

Figure 1.3: The flow chart of Fourier holography. The recorded data is the intensity of the diffraction pattern that is a Fourier transform of the specimen and the reference scatterer. The reconstruction is accomplished by taking a Fourier transform on the recorded diffraction intensity that leads to the cross-correlation of the function composed of the specimen and the scatterer. When the scatterer is small and far enough from the specimen, one of the side functions in the cross-correlation can be considered as the specimen function.

Suppose $R(\boldsymbol{r})$ and $O(\boldsymbol{r})$ respectively denotes the reference scatterer and the specimen function, the corresponding autocorrelation (written in the form of convolution) is given by

$$\begin{aligned}&\left[O(\boldsymbol{r})+R(\boldsymbol{r}-\boldsymbol{r}_0)\right]^*\otimes\left[O(-\boldsymbol{r})+R(-\boldsymbol{r}+\boldsymbol{r}_0)\right]=\\&O^*(\boldsymbol{r})\otimes O(-\boldsymbol{r})+R^*(\boldsymbol{r}-\boldsymbol{r}_0)\otimes R(-\boldsymbol{r}+\boldsymbol{r}_0)+O^*(\boldsymbol{r})\otimes R(-\boldsymbol{r}+\boldsymbol{r}_0)+O(-\boldsymbol{r})\otimes R^*(\boldsymbol{r}-\boldsymbol{r}_0),\end{aligned} \tag{1.7}$$

where the asterisk denotes the complex conjugate, $\otimes$ represents the convolution and $\boldsymbol{r}_0$ is the separation between the specimen and the reference scatterer. When the reference scatterer $R(\boldsymbol{r})$ is small enough to be approximated as the Dirac function, the third and the fourth terms in Eq. (1.7) respectively become $O^*(\boldsymbol{r}-\boldsymbol{r}_0)$ and $O(\boldsymbol{r}_0-\boldsymbol{r})$, i.e. the complex conjugate shifted version of the object function and its flipped shifted version. The first two terms in Eq. (1.7) are the autocorrelations of the object function and reference scatterer respectively. They are superimposed at the origin with a transverse extent twice as big as the size of the object function (in both Cartesian coordinates). To avoid the overlap between the four terms in Eq. (1.7), $\boldsymbol{r}_0$ cannot be smaller than one and half times the object size.

Although Fourier holography is able to image aperiodic specimen at high resolution, several drawbacks are associated with this imaging method. 1. The resolution of this imaging method is limited by the size of the reference scatter. The reconstruction is really the cross-correlation of the object function with the scatterer, not truly the object function. In hard x-ray regime, fabricating a small enough scatterer is challenging. Besides, when the scatterer is small, the transmitted x-ray flux becomes less, reducing the signal-to-noise ratio in the recorded data. The problem can be mitigated by using multiple reference scatterers [32] or by using extended reference with sharp features [33]. 2. The spatial coherence requirement is very high. The coherent beam needs to cover both the specimen and the scatterer. Since the distance between the scatterer and the specimen needs to be at least one and half times the specimen size, the extent of the beam with good coherence (i.e. the spatial coherence length) needs not be smaller than twice the specimen size. 3. The sampling requirement in the detector plane is very stringent, i.e. the pixel pitch of the detector needs to be very small. According to Shannon sampling condition (see section 2.1.4), the pixel pitch needs to be smaller than the inverse of the extent of the illuminated area. Since the beam extent needs to be twice bigger than the specimen size, as we explained in the second point, the pixel pitch needs to be twice smaller than the Shannon sampling pitch set by the specimen size.

1.2.5 Iterative coherent diffractive imaging

Iterative coherent diffractive imaging (CDI) is able to image non-crystalline objects at diffraction-limited resolution without use of any reference. It can be considered as an extension of crystallography to image aperiodic objects. Sayre [34] first suggested that the possibility of diffractive imaging of aperiodic objects given that the diffraction intensity is sufficiently sampled. Bruck et al [35] and Bates [36] have proved that the phase problem can be uniquely solvable for the two-dimensional case (or higher

dimensions) apart from some trivial ambiguities. The first observation of x-ray diffraction from an isolated object was reported in 1987 [37] but it was not until 1999 that an image of an array of fabricated gold dots was reconstructed from a measured diffraction intensity by Miao et al [38] using an iterative phase retrieval algorithm called the hybrid input output (HIO) algorithm proposed by Fienup [39, 40]. From then on, this imaging method has grown enormously in both size and scope (until it is replaced by iterative ptychography), e.g. imaging of cells [41] and 3D samples [42].

A typical setup of iterative CDI with an opaque aperture is shown in Fig. 1.4. A coherent plane wave illumination incidents on an opaque aperture, which sits right up against the specimen, so that an isolated area of the specimen can be illuminated. The exit wave emanating out of the specimen propagates over free space to the detector plane forming a diffraction pattern. The detector is positioned in the far field of the specimen, so that the diffraction pattern is estimated by a Fourier transform of the exit wave. The image reconstruction is performed using iterative phase retrieval algorithms. During the image reconstruction, only one diffraction pattern measurement is needed, so it is sometimes referred to as single-shot CDI.

Apart from the developments of hardware, like bright sources with good coherence,

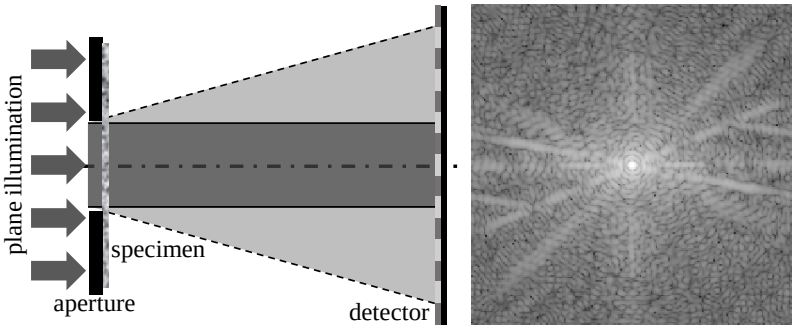

Figure 1.4: A typical pinhole setup for iterative coherent diffractive imaging. An isolated area of the specimen is illuminated by a coherent plane wave illumination and the resulting diffraction pattern is recorded by a detector placed somewhere downstream. Then the recorded diffraction pattern is used to reconstruct the specimen image via iterative phase retrieval algorithms.

high efficient detectors with small pixels and fast computers with large memories, iterative phase retrieval algorithms have played a very important role in the realisation of iterative CDI [43, 44]. Most iterative algorithms to reconstruct an image from its Fourier modulus are mainly built on an idea proposed by Gerchberg and Saxton for phase retrieval in electron imaging [45]. It is referred to as the GS algorithm. It requires two intensity measurements – an image of the specimen in real space and a diffraction intensity measurement in reciprocal space. The algorithm involves iterative Fourier transform back and forth between real space and reciprocal space and in between the applications of the real-space constraint and the reciprocal-space constraint. The real-space constraint is to replace the modulus of the current reconstruction with the square root of the measured image, meanwhile keeping the phase part intact. The reciprocal-space constraint is to replace the modulus of the Fourier transform of the current reconstruction with the square root of the measured diffraction intensity, meanwhile keeping the phase part untouched.

The block diagram of the GS algorithm is shown in Fig. 1.5. With an estimate of the specimen image g, one loop of the iteration consists of four steps: 1. Take the Fourier transform of the specimen estimate g to obtain the complex diffraction pattern G; 2. Replace the modulus of G with the square root of the recorded diffraction intensity to

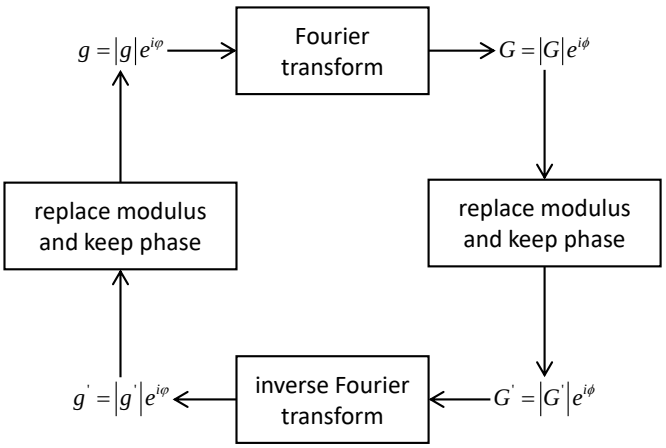

Figure 1.5: The block diagram of the GS algorithm.

give an updated diffraction pattern G'; 3. Take the inverse Fourier transform of the updated diffraction pattern and get a new specimen estimate g'; 4. Replace the modulus of the new specimen estimate with the square root of the measured image to provide an updated specimen estimate, serving as the starting point for the next iteration.

Fienup realised that the GS algorithm can be applied to any problem in which partial constraints (in the form of measured data or *a priori* information) are known in each of two domains, usually real space and reciprocal space. One simply transforms back and forth between the two domains, satisfying the constraints in one before returning to the other. In particular, the error reduction (ER) algorithm [39, 40] applies the support constraint instead of the modulus constraint in real space as used in the fourth step of the GS algorithm. The support constraint means beyond certain range the reconstruction has no values but zero. It can be expressed by

$$g_{k+1}(x) = \begin{cases} g'_k(x), & \text{if } x \in S \\ 0, & \text{if } x \notin S \end{cases}, \tag{1.8}$$

where S represents the support region and k denotes the iteration number. It allows the phase retrieval to start with only one intensity measurement (the diffraction intensity), and no longer needs the real-space measurement. The name error reduction is based on the fact that the calculated diffraction intensity from the reconstructed image is closer to the measured one each time after it is updated. However the algorithm easily gets stuck in local minima, causing stagnations [39, 40].

To avoid the stagnation problem of the ER algorithm, Fienup further developed the hybrid input output (HIO) algorithm [39, 40]. Instead of setting the values outside the support region to zero, one assigns the difference between the previous estimate g_k and its update g'_k to these values. It can be formularised as

$$g_{k+1}(x) = \begin{cases} g'_k(x), & \text{if } x \in S \\ g_k(x) - \beta g'_k(x), & \text{if } x \notin S \end{cases}, \tag{1.9}$$

where β is a feedback constant and is normally set to a value smaller than 1. This difference outside the support region acts like a strong modulator to the support region during the Fourier transform. By minimising this difference, i.e. the values are driven to zero, the HIO algorithm manages to avoid stagnations and converge much faster than the ER algorithm. In this generalisation, Fienup treated the first three steps in the ER algorithm as an nonlinear system, and thus the input and output of this nonlinear system are respectively g_k and g'_k. The strategy to constraint the region outside the support is a mixture of the input and output. This is why it is called the hybrid input output algorithm.

In real space, any other *a priori* information about the specimen can be added as constraints together with the support constraint to facilitate the convergence of the algorithm. The positivity constraint is the one often used for fabricated sample in hard x-ray experiments where the electron density of the specimen is real and positive [46]. It is quite effective in removing the twin image ambiguity that means the true image and its conjugate symmetry simultaneously exists in the reconstruction. The ambiguity is caused by the fact that the twin images have the same Fourier magnitudes and from diffraction intensity alone they cannot be distinguished. Twin image ambiguity can also be removed by using a non-centrosymmetric support [47].

Another very general projection-based algorithm called the difference map (DM) has been developed by Elser [48]. It is designed to solve problems that can be expressed as the search for the intersection point between two constraint sets. In iterative CDI, the two constraint sets are the support constraint and the modulus constraint. A detailed comparison between DM and HIO can be seen in [49].

However, iterative CDI has some drawbacks: 1. to fulfil the sampling requirement the support area needs to be smaller than half of the illuminated area (see section 2.3.4). This requires either an isolated small sample (very challenging to prepare) or an opaque aperture placed on top of an extended sample. Due to the limited spatial coherence in the experiments, the size of the illuminated area is restricted and so in either way only a small field of view is available; 2. For the object with a complex transmission function, the positivity constraint is not applicable, thus the twin image ambiguity could seriously corrupt the reconstruction. In this case, asymmetric apertures are required to get reasonable results [47, 50]; 3. For a 3D object, the reconstruction suffers defocus ambiguity [50]. The focus plane of the reconstruction is determined by which axial plane to impose the support constraint; 4. The presence of noise affects the probability that the unique solution will be found [50].

1.2.6 Iterative ptychography

Thanks to iterative phase retrieval algorithms, ptychography can be applied to non-crystalline objects in a much more practical way, because we no longer need to scan the sample at a step size of the resolution reconstructed. In 2004, Faulkner and Rodenburg developed an iterative algorithm called the ptychographical iterative engine (PIE) [51]. It has been successfully demonstrated with simulations using a sharp aperture [51] and a convergent illumination [52] and then with experiments using visible light [53, 54], x-ray [55] and electron [56]. It marked a revolutionary breakthrough for diffractive imaging. Compared to iterative CDI, ptychography makes measurements of multiple diffraction patterns from a set of overlapped illuminated area by laterally scanning the specimen with respect to the illumination, so it is also known as scanning CDI [57]. In this way, ptychography not only eliminates the ambiguity and convergence problems in iterative CDI, but also increases the robustness to noise and extends the field of view.

Fig. 1.6a shows a typical experimental configuration of iterative ptychography operated in transmission mode. It requires a localised illumination to be scanned over an extended specimen with the illuminated areas from adjacent scan positions partially overlapped. The scan is achieved by mounting the specimen on a translation stage with the scan plane perpendicular to the propagation direction of the illumination. At each scan position, the intensity of the diffraction pattern is recorded and the phase information is lost. Provided the overlap is sufficient, iterative algorithms can re-phase the recorded diffraction data and reconstruct a complex image of the specimen at a resolution only limited by the angular extent of the detector and the wavelength of the illumination. The flow chart of iterative ptychography is shown in Fig. 1.6b.

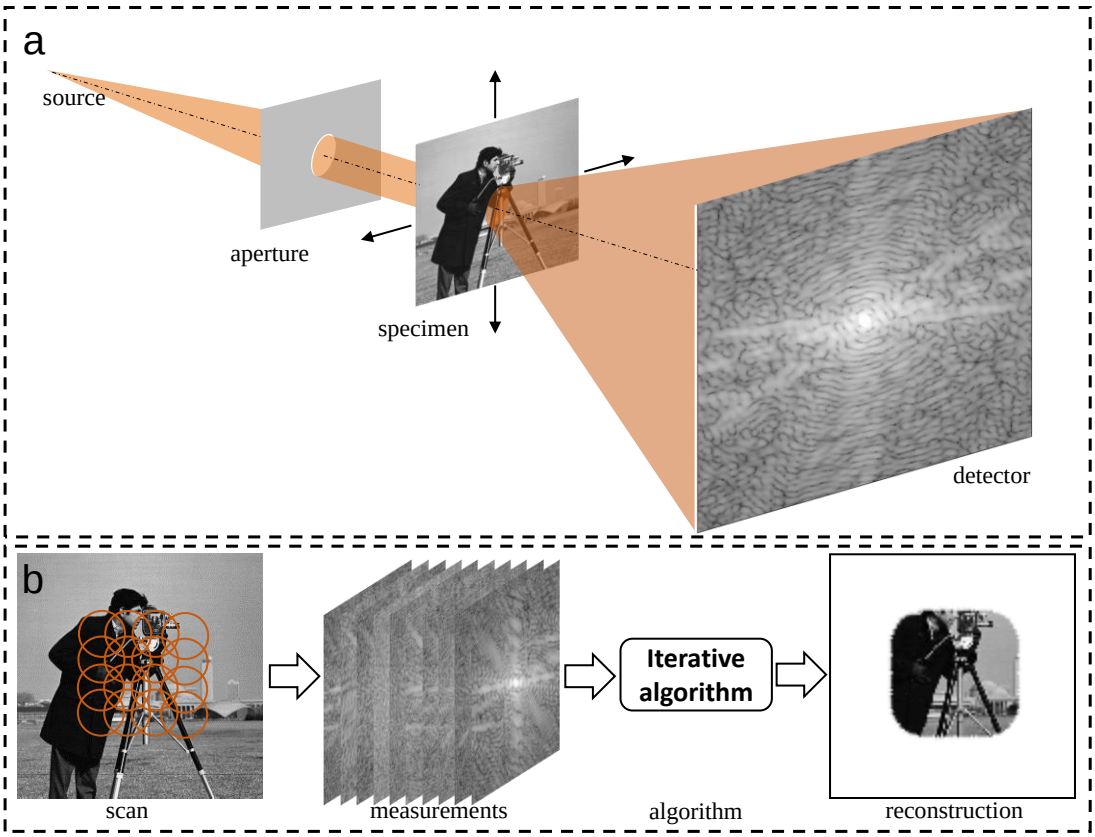

Figure 1.6: Iterative ptychography. (a) A typical experimental setup of iterative ptychography. A localised illumination is provided by an opaque aperture. The specimen is scanned laterally with respect to the illumination. At each scan position, the diffraction pattern is recorded in the far field. (b) The flow chart of ptychographic imaging. Iterative algorithms are adopted to reconstruct the specimen function from the measured diffraction intensities.

Since the overlapped areas are measured more than once, the recorded diffraction data has certain amount of redundancy. During the past decade, algorithmic developments allow us to keep exploring the redundancy in a ptychographic dataset and recovering more and more information of the experiment, making ptychography a very powerful and useful imaging technique. Alongside the object function, the illumination function [58, 59, 60], the scan positions [58, 61, 62, 63] and the partial coherence [64] can also be recovered from the dataset. In addition, ptychography can extrapolate a limited size detector to gain resolution improvements in the object reconstruction [65, 66], image 3D objects via the multislice method [67, 68, 69] without rotation of the specimen, deal with noisy data [70, 71], and achieve information multiplexing [72], dynamic object imaging [73] and continuous scan [74, 75, 76] via mixed state decomposition. It has also been modified to Bragg Ptychography [77] and Fourier Ptychography [78], and combined with tomography to achieve isotropic 3D imaging [79].

1.3 Outline of thesis

This thesis focuses on ptychography, particularly its application in the presence of partial coherence and its extension to 3D imaging.

As a coherent diffractive imaging technique, ptychography needs a radiation source with good coherence (both spatial and temporal). However, in real situations, especially in the regime of x-rays and electrons, most radiation sources are partially coherent. As a result, it compromises the quality of the images reconstructed. Although it has been demonstrated that good reconstructions are obtainable in the presence of partial coherence using iterative CDI via mixed state decomposition, a good model of partial coherence is necessary prior to the reconstruction [80, 81, 82]. Demonstration even shows the partial coherence can be reconstructed via a blind deconvolution using

iterative CDI, but the coherence function must be modelled as a Gaussian function [83]. Partial coherence was first proved solvable by Rodenburg and Bates using ptychography via the Wigner distribution deconvolution method in 1992 [26]. Because of this advantage, this discarded method is revisited and substantial improvements are made to it. The mixed state decomposition method used in iterative CDI has also been modified to adapt the iterative algorithms of ptychography and successful demonstration of this method has been made using an x-ray experiment by Thibault et al [64]. This thesis also investigates this iterative method and relevant contributions are made as well.

With the development during the past decade, iterative ptychography is mature enough to routinely deliver high quality 2D complex images (both modulus and phase) of samples at diffraction-limited resolution [84]. Not only that, but it also shows great potential in 3D imaging. When the specimen under investigation has certain thickness that is not big enough to violate the multiplicative approximation – the exit wave of the specimen when an illumination is shining upon it is estimated by the multiplication of the illumination function and the specimen function – the 2D image produced by ptychography is the projection of the 3D transmission function of the specimen at a specific orientation [26, 57]. This property makes it perfect to be combined with tomography. In addition, the quantitative phase image ptychography produces not only provides great contrast of the sample, but also allows quantitative analysis of the sample [85, 86]. The combination of ptychography and tomography, referred to as ptychographic tomography in this thesis, is particularly successful in the regime of x-rays and a lot of its applications have been explored in this regime [79, 87, 88, 89, 90]. This thesis will demonstrate ptychographic tomography step by step using a synchrotron x-ray experiment of a micro-capillary filled with glass beads that implicates great

possibility of imaging the 3D atom distributions within amorphous materials using electron ptychographic tomography. Furthermore, ptychography on its own is able to provide 3D information as well. When the specimen is thick enough to break the multiplicative approximation, the multislice method (widely used in electron microscopy [91]) can be applied to circumvent the problem by computationally sectioning the thick specimen into a set of thin slices [67]. All the thin slices can be reconstructed, hence providing 3D information of the specimen. A great advantage of this technique is that no rotation of the specimen is involved. It is particularly successful in the regime of visible light because of the strong scattering inside the specimen and high NAs available for the illumination [68], although a proof-of-principle experiment has successfully demonstrated using x-rays [69]. This thesis will modify the iterative algorithms to adapt the multislice method for Fourier ptychography that is a Fourier variant of the conventional real-space ptychography [78].

The outline of this thesis is detailed as follows:

Chapter 2 provides the theoretical and experimental backgrounds for ptychography. The most important mathematical tool, the Fourier transform, is first introduced and the associated properties are also given. Diffraction theory and different wave propagation models are then reviewed. Furthermore, the fundamental concepts and experimental concerns about ptychography are highlighted at the end.

Chapter 3 discusses ptychographic inversion via the Wigner distribution deconvolution method (WDDM). The full derivation of WDDM is reviewed and a condition for the separation of the Wigner distribution functions (WDF) of the specimen and the illumination is presented. To suppress the errors introduced by noise, three ways at different stages of WDDM are proposed: a projection strategy makes use of all the

available data to mitigate the inconstancy of noisy data; design of a favourable probe avoids division of zero by evening the WDF of the illumination; and an iterative method utilises the information with higher accuracy (less affected by the noise) to recover the information with lower accuracy. The reconstruction of spatial partial coherence via WDDM is also demonstrated for the first time using a model calculation.

Chapter 4 explores spatially mixed state ptychography. The reconstruction algorithm and the reconstruction structure are described. The modification of modulus constraint used in the coherent case enables the reconstruction of partial coherence via mixed state decomposition. However, since only the intensity addition of all the mixed states is constrained, linear ambiguity arises in the individual state reconstruction. *A priori* information of the experiment, like the orthogonalisation of the probe states in spatial partial coherent experiments and the phase-only response of the object states, can break the ambiguity and lead to true reconstructions. Furthermore, in this chapter the effects of a diffused probe on the reconstructions in the presence of spatial partial coherence are investigated. The use of diffuser increases the angular range of the illumination and enriches the information content encoded in the recorded intensity measurements. Eventually, it helps the reconstruction, especially when the specimen has very simple structures.

Chapter 5 investigates temporally mixed state ptychography. The reconstruction algorithm and the reconstruction structure are given. The linear reconstruction ambiguity is broken because each wavelength only has one specific response from the specimen. An iterative algorithm is proposed to remove the static background noise from the measurements by treating the background as the diffraction pattern from an extra temporal state. This enables reconstruction without the step of background subtraction.

Chapter 6 looks into Fourier ptychography and its extension to 3D imaging via the multislice method. Fourier ptychography applies the ptychographic scan in the Fourier space of specimen, i.e. on the spectrum of the specimen. A parallel update is adopted in the iterative algorithm for the reconstruction of the aperture to avoid the dramatic slow convergence caused by the fact that the spectrum normally has a very big value for the zeroth diffraction order and presents very high dynamic range. Also, the iterative algorithm is extended to involve three planes (specimen plane, back focal plane and image plane) in the reconstruction, so that the multislice method can be incorporated to realise 3D imaging.

Chapter 7 demonstrates tomography and ptychographic tomography. A background review of tomography is first given. In ptychographic tomography, ptychography acquires the projections of the 3D sample for all the orientations, so that tomographic reconstruction can be carried out on these projections. However, ptychographic reconstruction has inherent ambiguities, such as a phase ramp, phase offset and lateral shifts. These ambiguities need to be removed prior to the tomographic reconstruction. A synchrotron x-ray experiment of a micro-capillary filled with glass beads is used to demonstrate the implementation of ptychographic tomography. The possibility of electron ptychographic tomography is also discussed based on the scale calculation with the x-ray experiment.

Chapter 8 summarises the work presented in this thesis and discusses the implication of the work for future research.

2 Background

In this chapter, we will introduce the computational basis of diffractive imaging, the Fourier transform, and some frequently used properties associated with it. The scalar wave theory and propagation approximations, which formularise diffractive imaging, will also be reviewed. Furthermore, the fundamental concepts and some key experimental concerns about iterative ptychography are highlighted.

2.1 The 2D Fourier transform

The Fourier transform plays an essential role in diffractive imaging. It is a central part of the formularisation of the propagation theory. It also relates a specimen in real space and its spectrum in reciprocal space (the domain of the original function is referred to as real space and the domain after the transform as reciprocal space or Fourier space). Since only 2D wave fields or images are considered, the discussions about the Fourier transform are limited to two dimensions here.

2.1.1 Definition

The Fourier transform, denoted by $\mathcal{F}$, of a 2D function $\psi\left(r_x, r_y\right)$ is defined as

$$\Psi\left(f_x, f_y\right) = \mathcal{F}\left\{\psi\left(r_x, r_y\right)\right\} = \iint \psi\left(r_x, r_y\right) \exp\left[-i2\pi\left(r_x f_x + r_y f_y\right)\right] \mathrm{d}r_x \mathrm{d}r_y, \qquad (2.1)$$

where f_x and f_y are the coordinates in reciprocal space corresponding to real space represented by r_x and r_y and $i = \sqrt{-1}$. The inverse Fourier transform, denoted by $\mathcal{F}^{-1}$, is similarly defined as

$$\psi\left(r_x, r_y\right) = \mathcal{F}^{-1}\left\{\Psi\left(f_x, f_y\right)\right\} = \iint \Psi\left(f_x, f_y\right) \exp\left[i2\pi\left(r_x f_x + r_y f_y\right)\right] \mathrm{d}r_x \mathrm{d}r_y. \qquad (2.2)$$

The only difference between these two transforms is the sign in the exponential term.

2.1.2 Properties

There are well-known theorems associated with the Fourier transform [92]. Here we list some of them that will find wide use later in this thesis. If we define two Fourier transform pairs $\Psi(f_x, f_y) = \mathcal{F}\{\psi(r_x, r_y)\}$ and $\Gamma(f_x, f_y) = \mathcal{F}\{\gamma(r_x, r_y)\}$, some basic properties of the Fourier transform can be described as:

1. Linearity theorem:

$$\mathcal{F}\{a\psi(r_x, r_y) + b\gamma(r_x, r_y)\} = a\Psi(f_x, f_y) + b\Gamma(f_x, f_y). \tag{2.3}$$

The Fourier transform of a weighted sum of two functions is equal to the identical weighted sum of their individual transforms.

2. Similarity theorem:

$$\mathcal{F}\{\psi(ar_x, br_y)\} = \frac{1}{|ab|}\Psi\left(\frac{f_x}{a}, \frac{f_y}{b}\right). \tag{2.4}$$

A stretch of the real-space coordinates leads to a contraction of the reciprocal-space coordinates, plus an overall scaling of the Fourier amplitude.

3. Shift theorem:

$$\mathcal{F}\{\psi(r_x - a, r_y - b)\} = \Psi(f_x, f_y)\exp\left[-i2\pi(af_x + bf_y)\right]. \tag{2.5}$$

A lateral shift in real space introduces a linear phase offset in reciprocal space.

4. Convolution theorem:

$$\mathcal{F}\left\{\iint \psi(\xi, \eta)\gamma(r_x - \xi, r_y - \eta)\mathrm{d}\xi\mathrm{d}\eta\right\} = \Psi(f_x, f_y)\Gamma(f_x, f_y). \tag{2.6}$$

The Fourier transform of the convolution of two functions in real space is equivalent to the multiplication of their individual transforms in reciprocal space.

5. Autocorrelation theorem:

$$\mathcal{F}\left\{\iint \psi(\xi,\eta)\psi^{*}(\xi-r_x,\eta-r_y)\mathrm{d}\xi\mathrm{d}\eta\right\}=\left|\Psi\left(f_x,f_y\right)\right|^2. \tag{2.7}$$

This can be considered as a special case of the convolution theorem in which we convolve $\psi\left(r_x,r_y\right)$ with $\psi^{*}\left(-r_x,-r_y\right)$.

6. Parseval's theorem:

$$\iint\left|\psi\left(r_x,r_y\right)\right|^2\mathrm{d}r_x\mathrm{d}r_y=\iint\left|\Psi\left(f_x,f_y\right)\right|^2\mathrm{d}f_x\mathrm{d}f_y. \tag{2.8}$$

The energy in real space is preserved after transformed into reciprocal space.

7. Fourier integral theorem:

$$\psi\left(r_x,r_y\right)=\mathcal{F}\left\{\mathcal{F}^{-1}\left[\psi\left(r_x,r_y\right)\right]\right\}=\mathcal{F}^{-1}\left\{\mathcal{F}\left[\psi\left(r_x,r_y\right)\right]\right\}. \tag{2.9}$$

The successive transform and inverse transform of a function yield the original function.

8. Central symmetry theorem:

$$\psi\left(-r_x,-r_y\right)=\mathcal{F}\left\{\mathcal{F}\left[\psi\left(r_x,r_y\right)\right]\right\} \tag{2.10}$$

Two successive transforms of a function produce a central symmetric version of itself.

9. Conjugate symmetry theorem:

$$\Psi^{*}\left(-f_x,-f_y\right)=\mathcal{F}\left\{\psi^{*}\left(r_x,r_y\right)\right\} \tag{2.11}$$

The Fourier transform of the conjugate of a function results in a conjugate symmetric version of the Fourier transform of the original function.

10. Friedel's law:

$$\Psi\left(f_x,f_y\right)=\Psi^{*}\left(-f_x,-f_y\right),\quad \text{if } \psi\left(r_x,r_y\right)=\psi^{*}\left(r_x,r_y\right). \tag{2.12}$$

The Fourier transform of a real image is conjugate-symmetric.

2.1.3 Discrete Fourier transform

Since all the calculations in this thesis are computer-aided and a computer only handles discrete functions, here we need to introduce the definition of Discrete Fourier transform (DFT). It is defined as

$$\Psi(u,v) = DFT\{\psi(x,y)\} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\psi(x,y)\exp\left[-i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right], \qquad (2.13)$$

where x, y and u, v are the discrete coordinates in real space and reciprocal space respectively, and M and N are the numbers of discrete points in real space along x and y directions. Similarly, the inverse Discrete Fourier transform is defined as

$$\psi(x,y) = DFT^{-1}\{\Psi(u,v)\} = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}\Psi(u,v)\exp\left[i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right]. \qquad (2.14)$$

For the DFT, both real space and reciprocal space are discrete and periodic. The implementation of the DFT usually applies the fast Fourier transform (FFT) algorithm [93]. For this reason, the two terms, “DFT” and “FFT”, are often used interchangeably. It should be noted that the theorems introduced in section 2.1.2 for the continuous case are still valid for the discrete case DFT here.

2.1.4 The Shannon sampling theorem

Sampling is an operation to convert a continuous signal into a discrete sequence. The Shannon sampling theorem points out that the sampling interval has to be small enough to recover the original signal from its sampled version without loss of information [94]. Here we demonstrate the theorem using a one-dimensional function, but it is straightforward to extend it to two dimensions.

As shown in Fig. 2.1, a continuous signal $f(x)$ has a bandwidth of $2B$, so

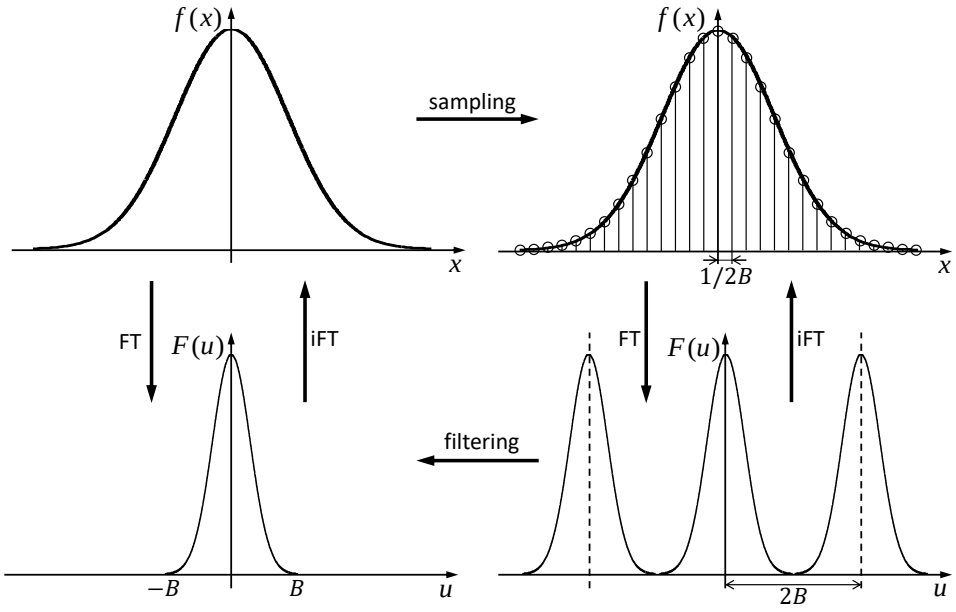

Figure 2.1: Illustration of the Shannon sampling theorem. For a continuous function *f(x)* with a bandwidth of *2B*, it can be fully recovered from its discrete version when the sampling interval is smaller than $1/2B$. FT stands for Fourier transform and iFT for inverse Fourier transform.

$$F(u) = \mathcal{F}\{f(x)\} = 0, \qquad \text{if } |u| > B. \tag{2.15}$$

We convert the continuous signal into a discrete sequence by sampling $f(x)$ on a set of equally spaced points such as

$$f[n] = \sum_{n=-\infty}^{\infty} f(x)\,\delta(x - n\Delta x), \qquad n \in \mathbb{N}, \tag{2.16}$$

where Δx is the sampling interval and δ is the Dirac delta function.

Performing the Fourier transform on the discrete sequence and utilising the convolution theorem of the Fourier transform, we have

$$\mathcal{F}\{f[n]\} = F(u) \otimes \sum_{n=-\infty}^{\infty} \delta\left(u - \frac{n}{\Delta x}\right) = \sum_{n=-\infty}^{\infty} F\left(u - \frac{n}{\Delta x}\right), \tag{2.17}$$

where $\otimes$ denotes the convolution operation. The Fourier transform of the sampled signal $f[n]$ is basically a repeat of the spectrum of the continuous signal, i.e. $F(u)$, on another set of discrete points with a separation of $1/\Delta x$. If all the repeated spectrums do

not overlap with each other (otherwise it is known as aliasing), filtering with a top-hat function $T(u)$ with a support region of $2B$ (unity on the support region, zero elsewhere) can pick out one complete spectrum for the recovery of the original function. To avoid the overlap (or aliasing) means the separation between the spectrums must be bigger than the width of the spectrum, i.e. $1/\Delta x \geq 2B$. This is equivalent to saying the sampling interval in real space must be smaller than the inversion of the bandwidth of the spectrum, i.e. $\Delta x \leq 1/2B$. In this case, the original function can be recovered without any loss of information from its sampled discrete version by a convolution with a sinc function

$$f(x) = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{f[n]\right\}T(u)\right\} = \sum_{n=-\infty}^{\infty} f[n]\,\mathrm{sinc}\left[2B\left(x - n\Delta x\right)\right]. \qquad (2.18)$$

This is the so called Shannon sampling theorem, which was originally pointed out by Whittaker [95] and was later popularised by Shannon [94].

2.2 Wave propagation

The formularisation of the wave propagation is the basis of the computational imaging. In this section we will give a brief introduction of the scalar diffraction theory that leads to the fundamental Rayleigh-Sommerfeld diffraction formula. Then we show that a series of approximations depending on the propagation distance can be applied to the diffraction formula to simplify the calculation. Another propagation method, angular spectrum propagation, is also introduced when the propagation distance is too short to adopt any other approximations on the diffraction formula.

2.2.1 Scalar diffraction theory

In a medium that is linear, dielectric, isotropic, homogeneous, nondispersive and nonmagnetic, all components of the electric and magnetic field behave identically and their behaviour is fully described by a single scalar wave equation [92, 96]

$$\nabla^2 u(\boldsymbol{r},t) - \frac{n^2}{c^2}\frac{\partial^2 u(\boldsymbol{r},t)}{\partial t^2} = 0, \tag{2.19}$$

where $\boldsymbol{r}$ represents the space coordinate, t denotes the time and n is the refractive index of the medium, defined by

$$n = \sqrt{\mu\epsilon}, \tag{2.20}$$

where μ and ϵ are the permeability and permittivity of the medium respectively, and c is the speed of light in vacuum, given by

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}}, \tag{2.21}$$

where μ_0 and ϵ_0 are the permeability and permittivity of the vacuum respectively. A detailed derivation of the scalar wave equation from Maxwell's equations can be found in standard textbooks of optics [92, 96]. If the scalar field is monochromatic, its disturbance is represented as

$$u(\boldsymbol{r},t) = U(\boldsymbol{r})\exp[-i2\pi\nu t], \tag{2.22}$$

where ν is the optical frequency and $U(\boldsymbol{r})$ represents the space dependent part of the disturbance. Since the disturbance must satisfy the scalar wave equation, the substitution of Eq. (2.22) into Eq. (2.19) leads to the time-independent Helmholtz equation, given by

$$\left(\nabla^2 + k^2\right)U(\boldsymbol{r}) = 0, \tag{2.23}$$

with the wave number *k* defined as

$$k = 2\pi n \frac{\nu}{c} = \frac{2\pi}{\lambda}, \tag{2.24}$$

where λ is the wavelength in the medium. As a result of the mathematical treatment involving Green's integral theorem, a general solution to the Helmholtz equation is given by the Rayleigh-Sommerfeld diffraction formula [92, 96]

$$U(u,v,z) = \frac{1}{i\lambda}\iint U(x,y,0)\frac{e^{ikr}}{r}\cos(\vec{n},\vec{r})\mathrm{d}x\mathrm{d}y \tag{2.25}$$

where $U(x,y,0)$ denotes the field in the diffraction plane, $U(u,v,z)$ is the field in the observation plane, $\cos(\vec{n},\vec{r})$ represents the cosine of the angle between the propagation direction $\vec{n}$ and the vector $\vec{r}$ joining the diffraction point to the observation point, and r is the length of the vector $\vec{r}$. Except for the cosine term that is referred to as obliquity factor, the equation has a physical interpretation identical to the Huygens-Fresnel principle [96]. It expresses the observed field $U(u,v,z)$ as a superposition of divergent spherical waves e^{ikr}/r originating from secondary sources located at each and every point of the diffraction field $U(x,y,0)$.

2.2.2 Fresnel propagation

A series of approximations can be applied to Eq. (2.25) to simplify the calculation. As depicted in Fig. 2.2, a slice of the diffraction field lies in the plane P1 (indexed by *x* and *y*) and propagates along the positive *z* direction. We will calculate the field across the observation plane P2 (indexed by *u* and *v*) that is parallel to P1.

The distance from a point (x,y) in P1 to a point (u,v) in P2 is denoted as *r* and it can be calculated by

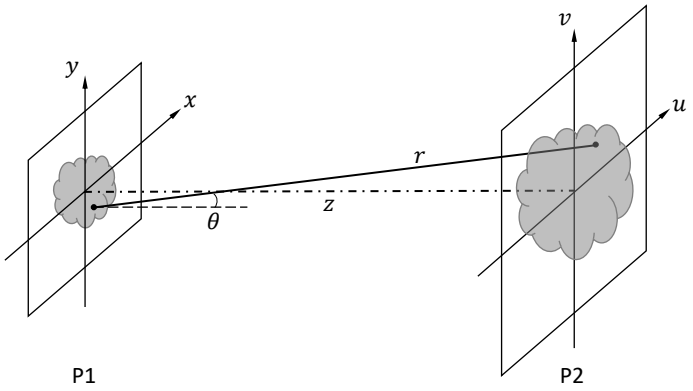

Figure 2.2: Illustration of wave propagation over free space between two planes. Different approximations can be applied depending on the propagation distance z and the wave extent in the transverse plane.

$$r=\sqrt{(u-x)^2+(v-y)^2+z^2}=z\sqrt{1+(\frac{u-x}{z})^2+(\frac{v-y}{z})^2}. \tag{2.26}$$

The angle between the vector joining these two points and propagation direction is denoted as θ and its cosine can be calculated by

$$\cos\theta=\frac{z}{r}. \tag{2.27}$$

In the paraxial approximation [92, 96], i.e. all the light rays are almost parallel to the *z*-axis, $\cos\theta\approx 1$. Therefore if the field in P1 is represented by $U(x,y)$, according to Eq. (2.25), the field $U(u,v)$ across P2 can be expressed as

$$U(u,v)=\frac{1}{i\lambda z}\iint U(x,y)\exp\left[ikz\sqrt{1+(\frac{u-x}{z})^2+(\frac{v-y}{z})^2}\right]\mathrm{d}x\mathrm{d}y. \tag{2.28}$$

Besides, we can make use of the binominal expansion (or Taylor expansion) of the square root to expand the expression of the distance *r* to the second term as

$$r\approx z\left[1+\frac{1}{2}\left((\frac{u-x}{z})^2+(\frac{v-y}{z})^2\right)\right]. \tag{2.29}$$

Substitution of Eq. (2.29) into Eq. (2.28) gives

$$U(u,v)=\frac{e^{ikz}}{i\lambda z}\iint U(x,y)\exp\left\{i\frac{k}{2z}\left[(u-x)^2+(v-y)^2\right]\right\}\mathrm{d}x\mathrm{d}y. \tag{2.30}$$

This is the Fresnel propagation approximation. It is a convolution of the field with a parabolic phase term. So this approximation basically uses parabolic phase surface to approximate the spherical phase surface. Eq. (2.30) can be further reformed as

$$U(u,v)=\frac{e^{ikz}}{i\lambda z}e^{i\frac{k}{2z}(u^2+v^2)}\iint\left[U(x,y)e^{i\frac{k}{2z}(x^2+y^2)}\right]\exp\left[-i\frac{k}{z}(ux+vy)\right]\mathrm{d}x\mathrm{d}y. \tag{2.31}$$

It is recognised as the Fourier transform of the product of the field and a quadratic phase term. During this approximation, we neglect the contribution from the third and the higher terms of the square root expansion, which requires

$$z\gg\sqrt[3]{\frac{\pi}{4\lambda}\left[(u-x)^2+(v-y)^2\right]^2_{\max}}. \tag{2.32}$$

When the Fresnel approximation is valid, P2 is said to be in the near field of P1. For this reason, the Fresnel approximation can also be called near-field approximation.

2.2.3 Fraunhofer propagation

If the parabolic phase term $\exp[i\frac{k}{2z}(x^2+y^2)]$ in Eq. (2.31) can be assumed be flat, which requires

$$z\gg\frac{(x^2+y^2)_{\max}}{2}, \tag{2.33}$$

the calculation can be further simplified as a Fourier transform

$$U(u,v)=\frac{e^{ikz}}{i\lambda z}e^{i\frac{k}{2z}(u^2+v^2)}\iint U(x,y)\exp\left[-i\frac{2\pi}{\lambda z}(ux+vy)\right]\mathrm{d}x\mathrm{d}y. \tag{2.34}$$

This is the Fraunhofer propagation approximation. For this approximation to be valid, P2 has to be in the far field of P1. So the Fraunhofer approximation is also called far-field approximation.

2.2.4 Angular spectrum propagation

The Fresnel and Fraunhofer approximations have simplified the calculation of the wave propagation and it is very easy to implement them in a computer. However, the propagation distance has to be long enough to validate these approximations. For short propagation distances, we use angular spectrum method to implement the calculation for the sake of computational convenience in a computer.

The angular spectrum of a wave field is defined as the Fourier transform of the wave field [92, 96]. So the angular spectrum of the wave in P1 is expressed as

$$A(f_x, f_y; 0) = \iint U(x, y, 0) \exp\left[-i2\pi(f_x x + f_y y)\right] \mathrm{d}x\mathrm{d}y. \tag{2.35}$$

The Fourier transform here can be regarded as a decomposition of the wave field into a series of simple complex-exponential functions. According to the Fourier integral theorem described in Eq. (2.9), the wave field is given as the inverse Fourier transform of the angular spectrum

$$U(x, y, 0) = \iint A(f_x, f_y; 0) \exp\left[i2\pi(f_x x + f_y y)\right] \mathrm{d}f_x\mathrm{d}f_y. \tag{2.36}$$

The complex-exponential function $\exp\left[i2\pi(f_x x + f_y y)\right]$ can be viewed as a plane wave with direction cosines

$$\alpha = \lambda f_x, \quad \beta = \lambda f_y, \quad \gamma = \sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}. \tag{2.37}$$

So the inverse Fourier transform in Eq. (2.36) can be regarded as recomposing the wave field by summing a series of plane waves weighted by its angular spectrum. Here we rewrite the angular spectrum in terms of direction cosines instead of spatial frequencies

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};0\right)=\iint U\left(x,y,0\right)\exp\left[-i2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}x\mathrm{d}y. \tag{2.38}$$

So the angular spectrum of the wave field in P2 can be expressed as

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=\iint U\left(u,v,z\right)\exp\left[-i2\pi\left(\frac{\alpha}{\lambda}u+\frac{\beta}{\lambda}v\right)\right]\mathrm{d}u\mathrm{d}v. \tag{2.39}$$

And likewise, the wave field in P2 can be recomposed by an inverse Fourier transform

$$U\left(x,y,z\right)=\iint A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)\exp\left[i2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}\frac{\alpha}{\lambda}\mathrm{d}\frac{\beta}{\lambda}. \tag{2.40}$$

Since $U\left(x,y,z\right)$ has to satisfy the Helmholtz equation, it leads to

$$\frac{\mathrm{d}^2}{\mathrm{d}z^2}A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)+\left(\frac{2\pi}{\lambda}\right)^2\left[1-\alpha^2-\beta^2\right]A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=0. \tag{2.41}$$

An elementary solution to the differential equation can be written in the form

$$A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};z\right)=A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};0\right)\exp\left[i\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}z\right]. \tag{2.42}$$

There are two different classes of solutions. For $\alpha^2+\beta^2>1$, the exponential term becomes real and decays to zero rapidly. This is the non-propagating solution that is also called evanescent waves [92, 96]. For $\alpha^2+\beta^2<1$, the effect of propagation over distance *z* is simply a change of the relative phases of the various components of the angular spectrum. Substitution of Eq. (2.42) into Eq. (2.40) gives

$$U(x,y,z)=\iint A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda};0\right)\exp\left[i\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}z\right]\exp\left[i2\pi\left(\frac{\alpha}{\lambda}x+\frac{\beta}{\lambda}y\right)\right]\mathrm{d}\frac{\alpha}{\lambda}\mathrm{d}\frac{\beta}{\lambda}. \tag{2.43}$$

This is the angular spectrum propagation. Its implementation can be broken down into three simple steps: 1. Take the Fourier transform of the wave field in P1; 2. For $\alpha^2+\beta^2<1$, multiply with the phase term $\exp\left[i\frac{2\pi}{\lambda}\sqrt{1-\alpha^2-\beta^2}z\right]$; 3. Take the inverse Fourier transform to give the wave field in P2. The angular spectrum method is applicable for any propagation distance. However, due to the successive Fourier transform and inverse Fourier transform, P1 and P2 are both in the same space, having the same sampling. As the propagation goes on, the size of the wave field will rapidly grow, eventually exceeding the calculation window, causing aliasing. This method is therefore mainly used for short propagation distance over which the Fresnel approximation is not valid.

2.3 Iterative ptychography

In this section we will consider some fundamental concepts, basic theories and experimental concerns of iterative ptychography.

2.3.1 The phase problem

A monochromatic electromagnetic wave field is completely characterised by a complex scalar function $\psi(\boldsymbol{r})=|\psi(\boldsymbol{r})|\exp[i\phi(\boldsymbol{r})]$, containing both the modulus part $|\psi(\boldsymbol{r})|$ and the phase part $\phi(\boldsymbol{r})$, where $\boldsymbol{r}$ represents a position vector in the wave field. With the whole information of $\psi(\boldsymbol{r})$, one can predict the wave field at any plane along the propagation direction by adopting an appropriate propagation model (see section 2.2). Unfortunately, a detector device is only able to record the intensity of the wave field, i.e. $I(\boldsymbol{r})=|\psi(\boldsymbol{r})|^2=\psi(\boldsymbol{r})\psi^*(\boldsymbol{r})$, where the asterisk denotes the complex conjugate, i.e.

$\psi^*(\boldsymbol{r}) = |\psi(\boldsymbol{r})| \exp[-i\phi(\boldsymbol{r})]$. The phase part $\phi(\boldsymbol{r})$ is lost during the recording process. This is referred to as the phase problem.

In ptychography, the detector is positioned somewhere downstream of the specimen to record the intensity of the diffraction pattern that is normally related to the exit wave of the specimen by a Fourier transform. Without the phase information of the diffraction pattern, one cannot acquire the knowledge of the exit wave via an inverse Fourier transform, and hence cannot deduce the illumination function and the specimen function from the exit wave that is estimated by the multiplication of the two functions. To reconstruct the image of the specimen, the phase information of the diffraction pattern needs to be recovered and algorithms to do the phase recovery are referred to as phase retrieval algorithms.

2.3.2 Coherence

The retrieval of phase information relies on the interference between the different diffracting components. The interference has to be stable during the integration time of the recording process. This requires a good coherence, both temporal and spatial, for the radiation source. The general description of coherence is based on evaluating the cross-correlation between two points in a wave field *U* separated in both space $\boldsymbol{r}$ and time *t* over certain time period *T*, which is defined as the mutual coherence function [97], given by

$$\Gamma(\boldsymbol{r}_1, \boldsymbol{r}_2, \tau) = \langle U(\boldsymbol{r}_1, t), U^*(\boldsymbol{r}_2, t+\tau) \rangle_T = \frac{1}{2T} \int_{-T}^{T} U(\boldsymbol{r}_1, t) U^*(\boldsymbol{r}_2, t+\tau) \mathrm{d}t, \qquad (2.44)$$

where $*$ denotes the complex conjugate and τ the time offset between the measurement of the wave field at the two points $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. The complex degree of coherence is defined as the normalisation of the mutual coherence function

$$\gamma(\boldsymbol{r}_1,\boldsymbol{r}_2,\tau)=\frac{\Gamma\left(\boldsymbol{r}_1,\boldsymbol{r}_2,\tau\right)}{\sqrt{\Gamma\left(\boldsymbol{r}_1,\boldsymbol{r}_1,0\right)\Gamma\left(\boldsymbol{r}_2,\boldsymbol{r}_2,0\right)}}. \tag{2.45}$$

If a monochromatic or quasi-monochromatic wavefield is considered, the dependence on the time offset τ can be dropped with $\gamma(\boldsymbol{r}_1,\boldsymbol{r}_2)$ denoting the degree of the spatial coherence. We have $\left|\gamma(\boldsymbol{r}_1,\boldsymbol{r}_2)\right|=1$ for full coherence, $\left|\gamma(\boldsymbol{r}_1,\boldsymbol{r}_2)\right|=0$ for complete incoherence and any value between for partial coherence in space. According to Wolf [98], for a spatially partial coherent wavefield, its mutual coherence function can be expressed as a weighted summation of a set of orthonormal modes

$$\Gamma\left(\boldsymbol{r}_1,\boldsymbol{r}_2\right)=\sum\nolimits_n \eta_n\varphi_n(\boldsymbol{r}_1)\varphi_n^{*}(\boldsymbol{r}_2), \tag{2.46}$$

with

$$\int_{\boldsymbol{r}} \varphi_m(\boldsymbol{r})\varphi_n^{*}(\boldsymbol{r})\mathrm{d}\boldsymbol{r}=\delta_{mn}, \tag{2.47}$$

where η_n is the power occupation of the mode $\varphi_n(\boldsymbol{r})$ and δ_{mn} is the Kronecker symbol. According to Eq. (2.46), the intensity of the partially coherent wave field is given as

$$I(\boldsymbol{r})=\Gamma\left(\boldsymbol{r},\boldsymbol{r}\right)=\sum\nolimits_n \eta_n\left|\varphi_n(\boldsymbol{r})\right|^2, \tag{2.48}$$

i.e. the incoherent summation of the intensities of all the modes. Therefore, the diffraction pattern measured from a partially coherent radiation can be treated as the incoherent summation of a series of diffraction patterns measured from each individual mode.

Coherence can also be estimated by a characteristic length scale called coherence length that is defined as the propagation distance over which a coherent wavefield maintains a high degree of coherence [99]. Spatial coherence is estimated by transverse coherence length. A source with lateral extension in the plane transverse to the propagation

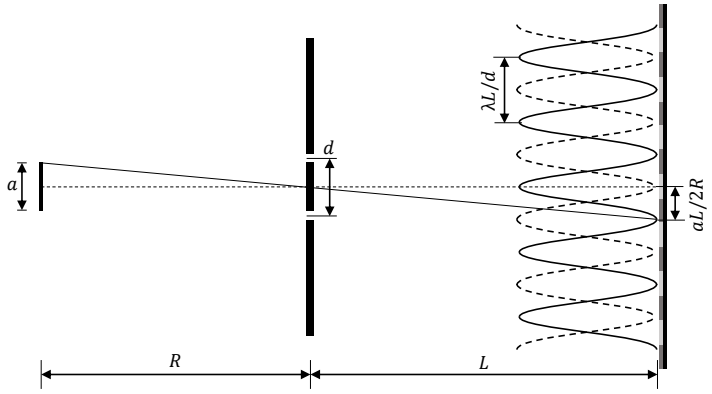

Figure 2.3: Demonstration of transverse coherence length via the Young's double slits experiment with a line source. The diffraction patterns from two narrow slits separated by a distance of *d*, originating from the centre of the source (dashed curve) and from the edge of the source (solid curve). The transverse coherence length is defined to be the slit distance *d* is when the two patterns are in antiphase.

direction will degrade the spatial coherence. Let us consider the Young's double slits experiment with a line source, as shown in Fig. 2.3.

The source extent is denoted as *a*, the separation between the two slits as *d*, the wavelength of the radiation as λ, the distance from the source to the slits as *R* and the distance from the slits to the detector as *L*. The central point in the line source will generate an interference pattern in the detector plane. The constructive interference happens when $d\sin\theta = n\lambda$. *n* is integers, representing different diffraction orders. θ is the scattering angle and under the small-angle approximation $\theta \approx \sin\theta = n\lambda/d$. Different points in the line source will generate the same interference pattern with a lateral shift in the detector plane. If the minima of the interference pattern from the extreme point of the line source coincide with the maxima of the interference pattern from the central point, i.e. $aL/2R = \lambda L/2d$, the separation between the two slits is defined as the transverse coherence length d_t and it is given by

$$d_t = \frac{\lambda R}{a}. \tag{2.49}$$

Eq. (2.49) implies that a long wavelength source with small extent or large distance from the source to the specimen could give a good spatial coherence in the specimen plane. That is why in synchrotron x-ray coherent experiments an upfront slit is required to control the source size and the experimental hatch needs to be built far away from the source, or in electron microscope a small spot size is desired for high spatial coherence.

The temporal coherence is characterised by the longitudinal coherence length [99]. A radiation with a broad bandwidth will cause degradation of temporal coherence. Let us consider two waves with wavelengths λ and $\lambda+\Delta\lambda$. They depart from a point at the same time and after a distance of *d* the two waves are in antiphase, as shown in Fig. 2.4. This distance *d* is defined as the longitudinal coherence length.

To calculate the propagation distance *d*, we assume the wave with wavelength of $\lambda+\Delta\lambda$ propagates *N* wavelengths, and then the wave with wavelength of λ propagates $N+1/2$ wavelengths. The two distances are equal, i.e. $N(\lambda+\Delta\lambda)=(N+1/2)\lambda$, which, as a result, gives $N=\lambda/2\Delta\lambda$. The propagation distance can then be calculated by $d=N(\lambda+\Delta\lambda)=\lambda^2/2\Delta\lambda+\lambda/2$. Normally the wavelength is much bigger than the wavelength difference, i.e. $\lambda\gg\Delta\lambda$, so the longitudinal coherence length d_l can be approximately given by

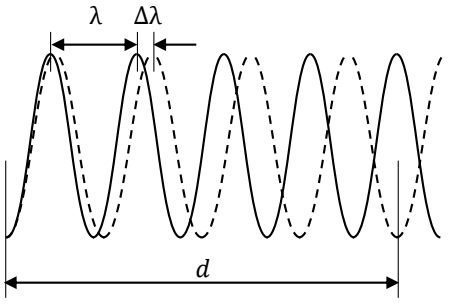

Figure 2.4: Demonstration of longitudinal coherence length via two waves with wavelengths λ and $\lambda+\Delta\lambda$. The two waves depart at the same point and at the same time. The longitudinal coherence length is defined to be the propagation distance when the two waves are in antiphase.

$$d_l \approx \frac{\lambda^2}{2\Delta\lambda}. \tag{2.50}$$

So a radiation with small bandwidth is desirable to ensure good temporal coherence. That is why in synchrotron x-ray coherent experiments monochromator is needed to select a narrow band of wavelengths, or in electron microscope stable high-voltage supply is required to suppress the energy spread.

2.3.3 Geometries of the image and detector

As we mentioned earlier, for computational imaging aided by a computer, both the input data and the output image are discrete. The data fed into the calculation for ptychography is the diffraction pattern recorded by a pixelated detector. To simplify the description, here we assume both the image and the detector are square with square pixels. Because they are related by a calculation of a discrete Fourier transform, the number of pixels is same for both of them. As shown in Fig. 2.5, we denote the sizes of the image and the detector as X and U respectively, and the pixel sizes as Δx and Δu. The number of pixels along one dimension is given as M, so the sizes of the image and the detector are given by $X = M\Delta x$ and $U = M\Delta u$ respectively.

Using the relation described in Eq. (2.37), the spatial frequency can be calculated by $f = \theta/\lambda$ with $\theta \approx \sin\theta = U/L$ under the paraxial approximation, where θ is the angle subtended by the detector from the image plane and L is the distance from the image to the detector. According to the Shannon sampling theorem, the sampling in real space, i.e. the image pixel size Δx is calculated by

$$\Delta x = \frac{1}{f} = \frac{\lambda L}{U} = \frac{\lambda L}{M\Delta u}. \tag{2.51}$$

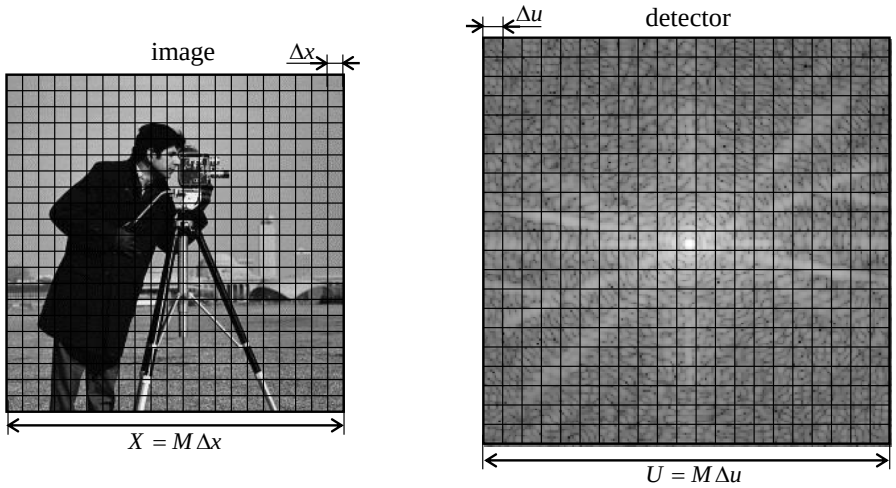

Figure 2.5: Geometries of the reconstructed image and the detector. Both the image reconstruction and the detector are discrete functions. The application of Shannon sampling theorem on both of them results in a relation between the pixel size in the image plane and the pixel size in the detector plane.

The image pixel size is inversely proportional to the detector size. If we define the resolution as the smallest feature of the specimen that can be resolved, the pixel size can be assumed as the resolution of the image, but only when the detector is just big enough to cover the whole diffraction pattern, because a bigger detector cannot record more information. In another words, a detector with a size bigger than the diffraction pattern cannot increase the resolution, although it can cause smaller pixel size of the reconstructed image.

2.3.4 Sampling requirement in ptychography

We consider the sampling of the intensity measurement for only one diffraction pattern first, i.e. the sampling for iterative CDI [34, 38, 100, 101]. It is denoted by κ_u and equals to the angle θ subtended by the detector sampling pitch (or the separation between two detector pixels) at the specimen plane, given by

$$\kappa_u = \frac{\sin\theta}{\lambda} = \frac{W}{\lambda L}, \tag{2.52}$$

where W is the separation between two detector pixels, λ is the wavelength of the radiation, and L is the distance from the specimen plane to the detector plane. If we use a detector with densely packed pixels (so no gap between two pixels, normally is the case in real situations), then W is equal to the pixel pitch Δu and it can be expressed by reforming Eq. (2.51) as $W = \Delta u = \lambda L / X$. As a result, the sampling in the detector plane is

$$\kappa_u = \frac{1}{X}. \tag{2.53}$$

It is basically the Shannon sampling in reciprocal space for a given image size X in real space. However, this sampling requirement is set provided all the information about the diffraction pattern (both modulus and phase) is known. In another words, Eq. (2.53) is the sampling of the complex diffraction pattern. Since only the intensity is measured, the sampling is now determined by the autocorrelation function of the image, not the image itself [34], because according to Eq. (2.7) the Fourier transform of the intensity of the diffraction pattern gives the autocorrelation of the image. The autocorrelation size is twice as big as the image size, so the sampling of the diffraction intensity is now becoming

$$\kappa_u = \frac{1}{2X}. \tag{2.54}$$

In the literature of iterative CDI, this is referred to as oversampling from the perspective of the sampling of a complex diffraction pattern [38, 100, 101]. To satisfy this, an isolated specimen or an opaque aperture with a support size smaller than half of the calculation window is needed in each dimension.

The degree of oversampling can be characterised by a quantity called the oversampling ratio σ in real space and it is defined in each dimension by [100, 101]

$$\sigma = \frac{\text{linear extent of the whole image}}{\text{linear extent of the support region}}. \tag{2.55}$$

According to the sampling requirement described in Eq. (2.54), the oversampling ratio needs to be bigger than 2 in each direction to make the phase retrieval possible.

For iterative ptychography, one needs to scan the illumination patch over the sample. When the illumination patch does not overlap for the adjacent scans, each scan can be considered as an independent iterative CDI and the sampling for each scan is thus the same as iterative CDI. However, when the illumination patch overlaps, the information content increases because the overlapped areas are measured more than once, hence relaxing the sampling requirement in the detector plane. If we consider the scan of the illumination patch in the specimen plane as a sampling in real space, then ptychography has two types of sampling: sampling in real space and sampling in reciprocal space. Edo et al [102] points out that these two types of sampling can trade for one another without loss of information. A ptychographic sampling ratio σ_{Π} can also be defined and in each dimension given by

$$\sigma_{\Pi} = \frac{1}{\kappa_u \kappa_x}, \tag{2.56}$$

where κ_u and κ_x represent the sampling pitches in reciprocal space and real space respectively. κ_u is determined by the inverse of the linear extent of the illumination patch (according to Shannon sampling) and κ_x equals the step size of the scan. So Eq. (2.56) can be reformed as

$$\sigma_{\Pi} = \frac{1}{1-\eta}, \tag{2.57}$$

with the overlap η defined as

$$\eta = \frac{\left(\text{linear extent of the illumination}\right) - \left(\text{scan step size}\right)}{\text{linear extent of the illumination}}. \qquad (2.58)$$

The phase problem suggests that we need at least twice as many measurements to recover phase information. Thus, the ptychographic sampling ratio needs to be bigger than 2 in each direction. According to Eq. (2.56), it is obvious that increasing the sampling in real space helps relax the sampling in reciprocal space, and vice versa.

2.3.5 Experimental geometries for the illumination

According to the Shannon sampling theorem, the limited pixel pitch size of the recording detector imposes an upper limit on the size of the illumination used in ptychography. Although this upper limit, as we mentioned above, can be increased by using smaller scan step size [102, 103], we always try to fulfil the fundamental Shannon sampling in an experiment. An illumination of limited size can be formed by passing an extended beam through a pinhole or by condensing an extended beam via a lens as depicted in Fig. 2.6.

The first way is referred to as pinhole setup. It is very easy to implement. Because of the diffraction of the edges of the aperture, the distance from the aperture to the specimen has to be as small as possible to stop the illumination from evolving too big when it approaches the specimen. Due to the unscattered beam, this setup will produce a diffraction pattern with a very high dynamic range – high counts in the central spot but very low counts at higher scattered angles [38]. Since the detector has a limited dynamic range, this will cause a very low signal-to-noise ratio for the high scattered signals that correspond to the high-resolution information in the specimen. As a consequence, a noisy reconstruction will be produced. Although multiple exposures can mitigate this problem (see section 2.3.10), the experimental time duration will increase, which could cause radiation damage and instabilities.

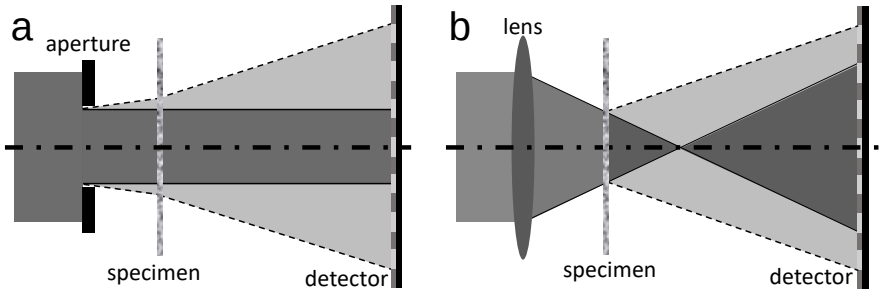

Figure 2.6: Optical setups for creating localised illumination required in ptychography. (a) Pinhole setup where an opaque aperture is placed in the proximal front plane of the specimen to form a confined illumination patch on the specimen plane. (b) Lens setup where a lens is used to condense an extended wavefield to a localised patch on the specimen plane. The illumination size can be controlled via the defocus.

The second way is referred to as lens setup. If the spherical aberration of the lens is not considered, placing the specimen either upstream or downstream of the focus is equivalent in terms of producing an illumination with certain size. This setup can alleviate the high dynamic range problem substantially, because the high unscattered counts are now spread over a disc. Besides, this setup has flexible control over the illumination size via changing the defocus. Furthermore, from the perspective of information encoding, the lens setup also defeats the pinhole setup by having a range of incident angles in the illumination. From the lens setup, the probe will consist of a sum of plane waves with different angles, i.e. a range of *k*-vectors of k_i in reciprocal space [104]. When it illuminates an object, the resulting scattered wave will also have a range of *k*-vectors of k_s that is recorded on a detector (a specific pixel of the detector records a specific k_s). The structure information of the object is encoded in the scattering vectors $K = k_s - k_i$, which means the structure information is expressed at many different pixels in this case. If one of the pixels is 'dead' (i.e. no information is recorded on that pixel), it can still be recovered from other pixels that express the same structure information of the object. However, from the pinhole setup, the probe mainly contains one k_i. As a result, the structure information is only expressed at one k_s. If the

particular pixel that measures the structure information is dead, the corresponding information will be lost and cannot be recovered.

All these advances of the lens setup are basically coming from the large range of incident angles. Therefore these advantages can be further enhanced by inserting a static diffuser upstream of the specimen. Considering the diffraction of it, the diffuser should be as close as possible to the specimen to stop the illumination getting too big when it reaches specimen.

2.3.6 The Ptychographic Iterative Engine (PIE)

The PIE algorithm was first proposed by Faulkner and Rodenburg and demonstrated using a sharp aperture function as the illumination function [51]. However, a sharp aperture illumination is not applicable for a real experiment. For this reason, the PIE algorithm was further improved by Rodenburg and Faulkner to allow the use of a realistic illumination function [52]. Here we denote the illumination function as P, the object transmission function as O, the exit wave as ψ, the diffraction pattern as Ψ, the measured intensity as M, the iteration number as n, the reciprocal-space coordinate as $\boldsymbol{q}$, the real-space coordinate as $\boldsymbol{r}$ and the scan position as $\boldsymbol{r}_j$. The iterative procedures of the PIE algorithm are given as follows:

1. In the nth iteration, the exit wave guess at a probe position of $\boldsymbol{r}_j$ is formed by the product of the current object transmission function guess and the known illumination function at the corresponding scan position:

$$\psi_j(\boldsymbol{r}) = P(\boldsymbol{r} - \boldsymbol{r}_j) O_n(\boldsymbol{r}). \tag{2.59}$$

2. In the far field the diffraction pattern guess is formed by taking a Fourier transform of the exit wave guess:

$$\Psi_j(\boldsymbol{q}) = \mathcal{F}\left[\psi_j(\boldsymbol{r})\right]. \tag{2.60}$$

3. Apply the modulus constraint, i.e. replace the modulus of the diffraction pattern guess with the square root the measured intensity and keep the phase (it is exactly the same as the reciprocal-space constraint used in iterative CDI (see section 1.2.5)), and back propagate to real space to get an updated exit wave ψ':

$$\psi'_j(\boldsymbol{r}) = \mathcal{F}^{-1}\left[\sqrt{M_j(\boldsymbol{q})}\frac{\Psi_j(\boldsymbol{q})}{\left|\Psi_j(\boldsymbol{q})\right|}\right]. \tag{2.61}$$

4. The object update function is given as:

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + \alpha\frac{\left|P(\boldsymbol{r}-\boldsymbol{r}_j)\right|}{\left|P(\boldsymbol{r}-\boldsymbol{r}_j)\right|_{\max}}\frac{P^*(\boldsymbol{r}-\boldsymbol{r}_j)}{\left(\left|P(\boldsymbol{r}-\boldsymbol{r}_j)\right|^2+\epsilon\right)}\left[\psi'_j(\boldsymbol{r})-\psi_j(\boldsymbol{r})\right], \tag{2.62}$$

where ϵ is an arbitrary constant with a small value to avoid division by zero, α is also a constant ranging between 0 and 1 to control the feedback of the update, and $\left|P(\boldsymbol{r}-\boldsymbol{r}_j)\right| \Big/ \left|P(\boldsymbol{r}-\boldsymbol{r}_j)\right|_{\max}$ is the normalised modulus of the illumination function. The normalisation favours the areas of the specimen that have been strongly illuminated. This strategy, according to Rodenburg and Faulkner [52], could suppress the high errors that would otherwise arise when the illumination is weak. Repeat 1 to 4 for the next scan position until all the positions are looped through and it is considered as one loop of the PIE algorithm. When the illumination function is a sharp aperture and only one scan position is considered, the PIE algorithm reduces to the ER algorithm. It is the scanning and the overlapping of the illumination function in real space that enable the PIE algorithm to defeat the ER algorithm in terms of convergence speed, uniqueness of solution and field of view.

The PIE algorithm was successfully demonstrated using visible light [53, 54], x-rays [55] and electrons [56]. However, it has a serious disadvantage that is the requirement of accurate illumination function. At best, this could be very time-consuming, whilst at worst an accurate illumination model may be impossible to get. Algorithms with simultaneous reconstruction of the illumination and the object were developed to address this problem [58, 59, 60]. Among these algorithms, the extended PIE (ePIE) algorithm [60] and the difference map (DM) algorithm [59] give comparably good results. Since the ePIE algorithm is used for all the reconstructions shown in this thesis, we will give a detailed description of it.

2.3.7 The extended PIE and its relation with the gradient descent method

To eliminate the necessity of accurate illumination function in the PIE algorithm, Maiden and Rodenburg [60] extended it by introducing a probe update function, which is referred to as the ePIE algorithm. PIE and ePIE share the first three steps, except PIE uses the known probe function $P(\boldsymbol{r})$ to calculate the exit waves while ePIE uses an updated probe function $P_n(\boldsymbol{r})$. In addition to the introduced probe update function, ePIE also changes the object update function so that no division is present (division by constant is not counted). Using the same notations as above, the updated probe function and object function are given as:

$$\left.\begin{aligned} P_{n+1}(\boldsymbol{r}) &= P_n(\boldsymbol{r}) + \alpha \frac{O_n^*(\boldsymbol{r}+\boldsymbol{r}_j)\left[\psi'_j(\boldsymbol{r}) - \psi_j(\boldsymbol{r})\right]}{\left(\left|O_n(\boldsymbol{r}+\boldsymbol{r}_j)\right|^2\right)_{\max}} \\ O_{n+1}(\boldsymbol{r}) &= O_n(\boldsymbol{r}) + \beta \frac{P_n^*(\boldsymbol{r}-\boldsymbol{r}_j)\left[\psi'_j(\boldsymbol{r}) - \psi_j(\boldsymbol{r})\right]}{\left(\left|P_n(\boldsymbol{r}-\boldsymbol{r}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{2.63}$$

The update functions given in step 4 can be considered as the gradient descent search method. At each scan position, the algorithm is trying to minimize the difference between the $\psi_j(\boldsymbol{r})$ and $\psi'_j(\boldsymbol{r})$, i.e.

$$E = \left| P_n(\boldsymbol{r} - \boldsymbol{r}_j) O_n(\boldsymbol{r}) - \psi'_j(\boldsymbol{r}) \right|^2 . \tag{2.64}$$

Take a partial derivative with respect to the illumination function and the object function respectively, the resulting gradients are

$$\left.\begin{aligned} G_P &= 2O_n^*(\boldsymbol{r} + \boldsymbol{r}_j)\left[\psi_j(\boldsymbol{r}) - \psi'_j(\boldsymbol{r})\right] \\ G_O &= 2P_n^*(\boldsymbol{r} - \boldsymbol{r}_j)\left[\psi_j(\boldsymbol{r}) - \psi'_j(\boldsymbol{r})\right] \end{aligned}\right\}. \tag{2.65}$$

Apply the standard iterative form of the gradient descent search method [105], i.e.

$$x_{n+1} = x_n - \sigma \nabla F(x_n), \tag{2.66}$$

where x is the variable that needs to be searched to minimize the objective function $F(x)$, $\nabla F(x_n)$ is the gradient value of $F(x)$ with respect to x at the point of x_n and σ is the step size for the iterative search. Obviously, the two update functions in step 4 are choices of σ as

$$\left.\begin{aligned} \sigma_P &= \frac{\alpha}{2\left(\left|O_n(\boldsymbol{r} + \boldsymbol{r}_j)\right|^2\right)_{\max}} \\ \sigma_O &= \frac{\beta}{2\left(\left|P_n(\boldsymbol{r} - \boldsymbol{r}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{2.67}$$

The robustness of the ePIE algorithm demonstrates the effectiveness of the choices for the search step sizes. Parameters α and β are normally set to a value between 0 and 1. Higher values will boost the convergence, while lower values stabilise the reconstructions. When the recorded diffraction patterns are very good (for example, the signal-to-

noise ratio is very high), they can be set to 1 to lean on the fast convergence, otherwise lower values are preferred to obtain good reconstructions.

2.3.8 Reconstruction ambiguities

Because of the symmetry of the multiplication of the illumination function and the object function, some reconstruction ambiguities will occur. Suppose an illumination function $P(\boldsymbol{r})$ is scanned over the object $O(\boldsymbol{r})$ and a set of diffraction patterns are recorded. However, any pair of two functions of $\tilde{O}(\boldsymbol{r}) = u(\boldsymbol{r})O(\boldsymbol{r})$ and $\tilde{P}(\boldsymbol{r}) = P(\boldsymbol{r})/u(\boldsymbol{r})$ will produce the same set of diffraction patterns. Because of this, several inherent ambiguities will be present in the reconstructions:

1. *Modulus scaling and phase offset*. If $u(\boldsymbol{r})$ is equal to a complex constant $ae^{i\theta}$, we then have $\tilde{O}(r) = ae^{i\theta}O(\boldsymbol{r})$ and $P(\boldsymbol{r}) = P(\boldsymbol{r})/ae^{i\theta}$. So the two reconstructed functions could have opposite modulus scaling and phase offset that will be cancelled in the multiplication. This ambiguity cannot be eliminated unless part of either function is known. However, these reconstructions will be still considered as right reconstructions because only the relative modulus and the phase are important.

2. *Raster scan pathology*. When $u(\boldsymbol{r})$ is not a constant function but a periodic function, i.e. $u(\boldsymbol{r}) = u(\boldsymbol{r} - \boldsymbol{r}_j)$, reconstruction ambiguities will take place when $\boldsymbol{r}_j$ happens to be the scan positions. These ambiguous reconstructions cannot be considered right. This case is very common in practice, where raster scans are frequently used – this is referred to as raster scan pathology [59, 106]. The way to avoid this ambiguity is to break the periodicity of the scan grids. Two common methods are: 1. add a small amount of random offsets to a set of raster scan grids [60]; 2. use a circular scan geometry [79, 106].

Apart from the multiplication symmetry, some other ambiguities could take place due to the loss of the phase information of the measured diffraction patterns. During the reconstruction, the modulus constraint requires the calculated diffraction patterns from the reconstructed illumination and object to have the same intensity distributions as the measured ones, imposing no constraints on the phases. As a result, the following ambiguities could happen in the reconstructions:

1. *Arbitrary phase offset*. Since the Fourier magnitude is not sensitive to phase offsets, the reconstructed illumination and object are allowed to have arbitrary phase offsets and they cannot be cancelled in the multiplication. This ambiguity cannot be eliminated unless we have some priori knowledges about both the object and the illumination functions. However, only the relative phase is important to determine the structure of the specimen, so the reconstructions with arbitrary phase offsets can still be considered right.

2. *Lateral shift.* According to the shift theorem of the Fourier transform, a shift in real space will cause a phase ramp in reciprocal space. A phase ramp does not affect the intensity distribution, so the lateral shift version of the original functions can be produced in the reconstructions without compromise the modulus constraint. Determined by the relative shift between the illumination and the object during the ptychographic scan, the lateral shifts in the reconstructions of both functions have to be same. Again, this ambiguity does not affect the integrity of the reconstructions.

3. *Phase ramp*. If the illumination and the object have opposite phase ramps, the multiplication of them at different scan positions will cancel out the phase ramps but leave different constant phase offsets [107]. Since these phase offsets do not affect the intensity distributions, they will be possible reconstructions. These ambiguous phase

ramps can be calculated and removed if either the illumination function or the object function has a flat region. In addition, the phase ramp can also happen when the measured diffraction patterns are not at the centre of the detector, as the shift theorem of the Fourier transform indicates. This phase ramp could be either on the illumination or on the object. To stop this from happening, the diffraction patterns need to be centred before reconstruction.

According the conjugate symmetry theorem of the Fourier transform (see Eq. (2.11)), $P^*\left(-\boldsymbol{r}+\boldsymbol{r}_j\right)O^*\left(-\boldsymbol{r}\right)$ will produce a set of diffraction patterns with the same intensity distributions as $P\left(\boldsymbol{r}-\boldsymbol{r}_j\right)O\left(\boldsymbol{r}\right)$ will. So the conjugate symmetric versions of the illumination and the object could be reconstructed if the orientations of the scan positions are flipped. It should be noted that it is different from the case in CDI where the original function and its conjugate symmetry simultaneously exist in the reconstruction. Accurate knowledge of the orientation of the coordinate system can remove this ambiguity.

There is another kind of ambiguity resulting from the propagation model used for the space between the specimen plane and the detector plane [59]. Normally we use Fourier propagation. However, when the detector is in the near field of the specimen plane, the use of Fourier propagation is still able to give a right object reconstruction but a wrong illumination reconstruction with an extra parabolic phase offset according to Eq. (2.31) and (2.34). This is not critical when we only care about the object reconstruction. But in the situations where the illumination reconstruction matters, such as multislice ptychography [67] and optics wave front characterisation [108, 109, 110], Fresnel propagation has to be used instead.

2.3.9 Error metrics for the reconstructions

For a ptychographic reconstruction, normally the only available information is the recorded diffraction intensity. To assess the quality of the reconstructions, the direct way is to calculate the normalised root-mean-square-error between the calculated diffraction intensities (from the reconstructions) and the measurements [59, 60]. Here we call this error metric as *diffraction intensity error metric* and it is given by

$$\varepsilon = \sqrt{\frac{\sum_j \sum_{\boldsymbol{q}} \left| \left|\Psi_j(\boldsymbol{q})\right|^2 - M_j(\boldsymbol{q}) \right|^2}{\sum_j \sum_{\boldsymbol{q}} \left[M_j(\boldsymbol{q}) \right]^2}}, \tag{2.68}$$

where $\Psi_j(\boldsymbol{q})$ and $M_j(\boldsymbol{q})$ are respectively the calculated diffraction pattern and the measured intensity at the probe scan position *j*, as described in section 2.3.6. It is often used to monitor the convergence of the algorithm, so that when the error is smaller than a preset accuracy, the algorithm will stop the iteration. However, noise is inevitable in real situations. As a consequence, the measurements themselves are not accurate and this error metric is no longer able to reflect the quality of the reconstructions.

In this case, *Fourier ring correlation* (FRC) would be a better alternative to compare the reconstructions [111]. FRC is the 2D analogue of the Fourier Shell Correlation for 3D objects [112]. It calculates the normalised cross-correlation coefficient between reconstructions from two independent experiments over corresponding rings ($\boldsymbol{q}_i$) in Fourier space as a function of spatial frequency ($\boldsymbol{q}$) and it is given by

$$FRC(\boldsymbol{q}_i) = \frac{\sum_{\boldsymbol{q}\in\boldsymbol{q}_i} F_1(\boldsymbol{q}) F_2^*(\boldsymbol{q})}{\sqrt{\sum_{\boldsymbol{q}\in\boldsymbol{q}_i} \left|F_1(\boldsymbol{q})\right|^2 \sum_{\boldsymbol{q}\in\boldsymbol{q}_i} \left|F_2(\boldsymbol{q})\right|^2}}, \tag{2.69}$$

where $F(\boldsymbol{q})$ denotes the spectrum of the object reconstruction. It indicates the repeatability of the reconstructions.

In some situations, like model calculations or real situations where a calibration experiment is needed, the reference image of the object reconstruction is available. In these cases, we calculate the normalised root-mean-square-error between the reconstruction and the reference to assess the reconstruction quality and it is referred to as *complex image error metric* here. However, due to the reconstruction ambiguities, a direct calculation would fail to reflect the quality. An error metric invariant to the phase offset and lateral shift was developed by Fienup [113], but it is not applicable when the phase ramp ambiguity exists. Here we adopt similar idea to derive a more general error metric. Note that the modulus scaling ambiguity can be eliminated by normalising the reconstruction to its maximum modulus, so in the derivation we will not include the scaling. Suppose the reference image is $f(x,y)$ and the reconstruction image is $\exp[ia]g(x-r_x,y-r_y)\exp\left[i2\pi\left(ux/M+vy/N\right)\right]$, where a is the constant phase offset, (r_x,r_y) is the lateral shift, (u,v) is the linear coefficient for the phase ramp, and (M,N) is the dimension of the image. The normalised square error between the reference and the reconstruction is given by

$$\varepsilon^2=\frac{\sum_x\sum_y\left|f(x,y)-\exp[ia]g(x-r_x,y-r_y)\exp\left[i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right]\right|^2}{\sum_x\sum_y|f(x,y)|^2}. \qquad (2.70)$$

The error metric should be zero with respect to any of these ambiguities. That is, we find the parameters a, r_x, r_y, u and v that minimise the error ε^2. Eq. (2.70) can be simplified as

$$\varepsilon^2 = \frac{\sum_x \sum_y |f(x,y)|^2 + \sum_x \sum_y |g(x,y)|^2 - \left[\exp(-ia) W_{fg} + \exp(ia) W_{fg}^*\right]}{\sum_x \sum_y |f(x,y)|^2}, \quad (2.71)$$

with

$$W_{fg}(r_x, r_y, u, v) = \sum_x \sum_y f(x,y) g^*(x - r_x, y - r_y) \exp\left[-i2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right]. \quad (2.72)$$

W_{fg} is the Wigner distribution function (see section 3.1.2 for details) between $f(x,y)$ and $g(x,y)$. For a given lateral shift (r_x, r_y) and phase ramp (u,v) we determine the constant phase a that minimises the error by setting the partial derivative of $\varepsilon^2\left(a, r_x, r_y, u, v\right)$ with respect to a to zero:

$$-\left[-i\exp(-ia) W_{fg}(r_x, r_y, u, v) + i\exp(ia) W_{fg}^*(r_x, r_y, u, v)\right] = 0, \quad (2.73)$$

which has the solution

$$a = \arg\left[W_{fg}(r_x, r_y, u, v)\right] + n\pi, \quad (2.74)$$

where n is an integer and $\arg[*]$ represents the phase. The corresponding minimum error is

$$\varepsilon^2 = \frac{\sum_x \sum_y |f(x,y)|^2 + \sum_x \sum_y |g(x,y)|^2 - 2\left|W_{fg}(r_x, r_y, u, v)\right|_{\max}}{\sum_x \sum_y |f(x,y)|^2}. \quad (2.75)$$

The square root of the minimum error, i.e. ε, is the *complex image error metric*. It is invariant to all the reconstruction ambiguities. Compared to the other two, this error metric is most reliable to assess the reconstruction quality, but a reference image is needed.

2.3.10 Some key experimental parameters

To perform a ptychographic experiment, several key parameters need to be carefully taken care of:

1. *Scan pattern*. To date, two types of scan patterns are widely used: raster scan [60] and circular scan [79]. Raster scan is very straightforward and easy to implement. However, periodic positions in a raster scan can cause ambiguous reconstructions (raster scan pathology) in ptychography. The circular scan is proposed and adopted to fix the problem. Another easy way to fix it is to add small amount of random offsets to perturb the regular positions in a raster scan. In this thesis, all the experiments presented use the raster scan plus small random offsets.

2. *Scan step size*. The characteristic parameter of a raster scan is the step size between two scan positions. For an illumination of a specific size, scan step size determines the overlap of the adjacent scans. The overlap is given by $1-s/d$ (as defined in section 2.3.4), where s is the scan step size and d is the diameter of the illumination. It affects the sampling in real space, i.e. the redundancy in a ptychographic dataset, and hence the reconstruction quality. It has been studied for the PIE algorithm by Bunk et al that 60% overlap is required for a good reconstruction [54]. For the ePIE algorithm, it has not been systematically studied, but from personal experience at least 70% overlap is necessary for good reconstructions. To solve for more information, more overlap is needed.

3. *Random offsets*. As we mentioned above, to avoid raster scan pathology, random offsets need to be added. Normally, they are given a magnitude of from 10% to 20% of the scan step size. Too small would not be able to get rid of raster scan pathology. Too big would affect the overlap of the scan.

4. *Camera length*. This means the distance from the specimen plane to the detector plane. Decrease of camera length will enlarge the angle subtended by one detector pixel at the specimen plane and also the angle subtended by the whole detector. The former would cut down the maximum allowable illumination size according to the Shannon sampling theorem. The latter will cause a decrease of the pixel size (and also cause an increase of the resolution if the scattering strength falls outside of the detector) of the reconstruction image (see section 2.3.3). Of course, camera length also determines which propagation model would be used to form the diffraction pattern (see section 2.2).

5. *Exposure time*. This is defined as the time period over which the detector is exposed to the radiation. It determines how many photon/electron counts can be collected by the detector because the counts are proportional to the exposure time. The photon counting obeys Poisson statistics and the fluctuation of the counting numbers is called Poisson noise. The standard deviation of a Poisson distribution is equal to the square root of the expectation that is the average photons a pixel receives during the exposure time. This means the signal-to-noise ratio (SNR) for a specific count is its square root. Therefore, to increase the SNR, one needs to increase the counts and hence the exposure time.

6. *Multiple exposures*. As we mentioned earlier, a diffraction pattern normally has a very high dynamic range (the difference between the largest value and the smallest value), especially for the pinhole setup. A detector also has a limited dynamic range determined by the bit numbers of the analogue-to-digital-converter used. To detect a complete diffraction pattern, its dynamic range must be smaller than that of the detector. This means the exposure time cannot be infinitely increased, even though doing so can increase SNR, because otherwise the highest value in the diffraction pattern will saturate the detector. The signals of the high scattered angles correspond to the high frequencies in the specimen. They normally have relatively low counts and hence small

SNR compared with the signals of the low scattered angles. As a consequence, bad reconstructions are produced. The common way to fix this is to take multiple exposures with different exposure time, so for each exposure the subsequent set of recorded diffraction patterns will have a different non-saturated area with high SNR. Then these areas will be scaled according to the exposure time and combined to give a complete diffraction pattern but with much higher dynamic range.

7. *Background noise*. In the absence of radiation, the detector may still have some constant response that is referred to as the background noise. This is mainly caused by the thermal heat. Most detectors, like CCD and CMOS, utilise the photoelectric effect that converts photons into electrons. The detected signals are proportional to the electrons stored in the detector. However, due to the thermal heat produced by the detector circuitry, electrons can be generated as well without being hit by photons. The signals correspond to these electrons are referred to as thermal heat noise (or dark current). The ambient light in the laboratory could also contribute to the background noise. This noise will superimpose on the true signals to give the diffraction intensities we measure during the experiment. In order to compensate for it, we need to measure a dark frame without turning on the radiation and subtract it from the diffraction intensity measurements.

3 Direct Ptychography

In this chapter we will review the direct (non-iterative) solution for ptychography called the Wigner Distribution Deconvolution Method (WDDM), which has been discarded for almost twenty years, but recently has regained some attention [114, 115]. WDDM, as we described in section 1.2.3, requires to scan the illumination (or the specimen) over a very dense grid. The diffraction intensity recorded on each pixel of the detector will vary as a function of the scan position. An important step in WDDM is to take a Fourier transform of the intensity with respect to the scan position, which gives access to the phase differences between all the pairs of the diffraction components of the specimen. Then a deconvolution step allows us to further separate out the phase differences information and eventually solve the phase problem.

WDDM is remarkable for two reasons: 1. it solves the quadratic phase problem with linear calculations. 2. It can decouple the influence of the illumination from the specimen. The success of this algorithm was experimentally demonstrated with optical light [21], x-rays [22] and electrons [23, 24]. However, there were a few problems: 1. a very big 4D dataset needs to be recorded, which imposes challenges not only on the memory but also on the computation of a computer. 2. The deconvolution step involved is often ill-conditioned, especially in the presence of noise. Although this deconvolution step can be skipped, it only works weak phase objects and only double-resolution (compared to the resolution limit set by the imaging system) can be obtained [23, 26, 114]. 3. The phase retrieval step following the ill-conditioned deconvolution step is not very reliable, because of the fact that only a fraction of the big 4D dataset can be used. In this chapter we will try to fix these problems. The work described in this chapter has led to a journal paper publication [116].

3.1 Direct ptychography of aperiodic objects

In this section we will review the derivation of direct ptychography and point out a prerequisite for this method to work that has been missing from the previous derivations [20, 26]. Also a novel method will be demonstrated to use all the information from the measured data to solve for the phase problem, instead of just the small subset used in the original 'stepping out' procedure [18, 21, 22, 26].

3.1.1 The 4D intensity dataset

To keep consistent with the literature, here we keep all the nomenclature from Rodenburg and Bates [26], some of which will be different from that used in other chapters. As shown in Fig. 3.1, a radiation source $s(\boldsymbol{\rho})$ illuminates an aperture $A(\boldsymbol{r}')$ sitting in the back focal plane of a condenser lens. The illumination is then focused by the condenser lens onto the specimen plane as a probe function $a(\boldsymbol{r})$. The probe function can be scanned to various positions $\boldsymbol{\rho}$ laterally with respect to the specimen $\psi(\boldsymbol{r})$. At each scan position, the intensity of the corresponding diffraction pattern $M(\boldsymbol{r}', \boldsymbol{\rho})$ is recorded by a pixelated detector lying in the far field of the specimen.

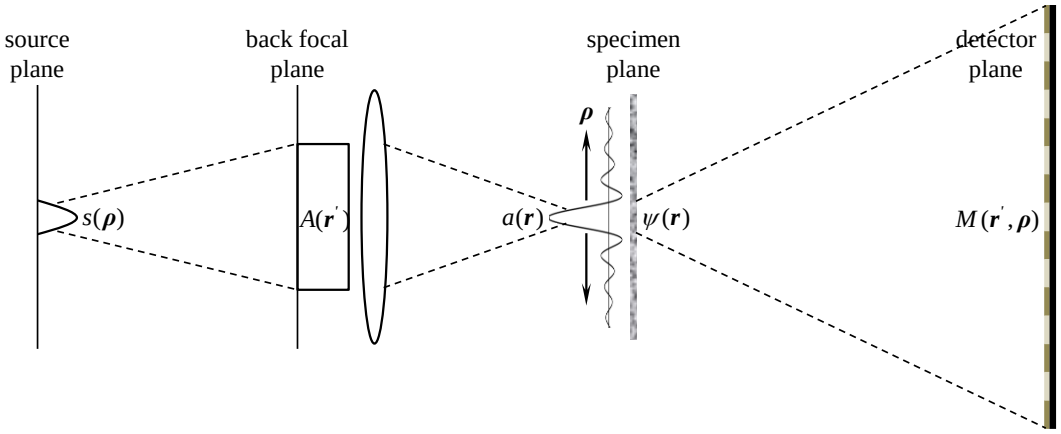

Figure 3.1: The coordinate definition for different planes in a lens setup ptychography. We define the coordinate in the source plane to be the same as the scan position, because scanning a point in the source plane is equivalent to scanning the probe across the specimen plane. The lens back focal plane and the detector plane are defined as reciprocal space of the specimen plane.

According to the Fraunhofer propagation approximation (see section 2.2.3), we have the following relation

$$M(\boldsymbol{r}', \boldsymbol{\rho}) = \int a(\boldsymbol{r} - \boldsymbol{\rho})\psi(\boldsymbol{r})e^{-i2\pi \boldsymbol{r}\boldsymbol{r}'}\mathrm{d}\boldsymbol{r}. \tag{3.1}$$

It should be noted that both $\boldsymbol{r}$ and $\boldsymbol{\rho}$ are 2D coordinates. As a result, we acquire a 4D intensity dataset $\left|M(\boldsymbol{r}', \boldsymbol{\rho})\right|^2$.

For a specific scan position $\boldsymbol{\rho}_0$ in the dataset, $\left|M(\boldsymbol{r}', \boldsymbol{\rho}_0)\right|^2$ is the intensity of the corresponding diffraction pattern. For a particular pixel $\boldsymbol{r}'_0$ in the dataset, according to the reciprocity, $\left|M(\boldsymbol{r}'_0, \boldsymbol{\rho})\right|^2$ is the intensity of the image that is obtained by placing the detector in the source plane and having the source in the position of the pixel $\boldsymbol{r}'_0$. This is a tilt illumination condition in the conventional electron microscope. Obviously, each image has the same resolution limited by the aperture at the back focal plane of the lens, but with different frequency information of the specimen. WDDM is a method to combine all this information to give a complex image of the specimen at a resolution limited by the recording angle of the detector.

As we will explain later, WDDM relies on the equivalence of the coordinates $\boldsymbol{r}'$ and $\boldsymbol{\rho}'$ (the reciprocal coordinate of $\boldsymbol{\rho}$) to recover the high frequency components in the spectrum of the specimen. This requires the equivalence of the coordinates $\boldsymbol{r}$ and $\boldsymbol{\rho}$, which means this method needs to scan the specimen at a step size ($\boldsymbol{\rho}$) of the pixel size of the reconstructed image ($\boldsymbol{r}$). In addition, the dimensions of $\boldsymbol{r}$ and $\boldsymbol{\rho}$ (or $\boldsymbol{r}'$ and $\boldsymbol{\rho}'$) need to be the same as well.

To help understand this, here a model calculation is simulated based on the configuration shown in Fig. 3.1. Since WDDM needs to do a very dense scan, it

imposes challenges on both the computation and the memory of the computer. To alleviate this problem, we set the probe and the object to be 64×64 pixels, as shown in Figs. 3.2a and 3.2b. The pixel size in real space is 2Å. The electron beam has a wavelength of 0.0037nm (equivalent to beam energy of 100keV). The aperture size at the back focal plane of the objective lens is chosen to be 3mrad that is about one third of the entire reciprocal space, which can be seen from the unscattered beam in one of the recorded diffraction patterns as shown in Fig. 3.2c. The focused beam at the defocus of 1000Å is chosen as the probe. Since $\boldsymbol{r}$ and $\boldsymbol{\rho}$ need to be equivalent and have the same dimension, the scan grid is set to 64×64 with a step size of 1 pixel. At each scan position, the intensity of the diffraction is recorded. In this way, a simulated 4D intensity data is acquired.

Fig 3.2c is a 2D slice of the 4D intensity data at $\boldsymbol{\rho} = 0$. It is basically the intensity of the diffraction pattern at the scan position of $\boldsymbol{\rho}_x = 0$ and $\boldsymbol{\rho}_y = 0$. Fig 3.2d is a 2D slice of the 4D intensity data at $\boldsymbol{r}' = 0$. It is basically, according to the reciprocity, the intensity

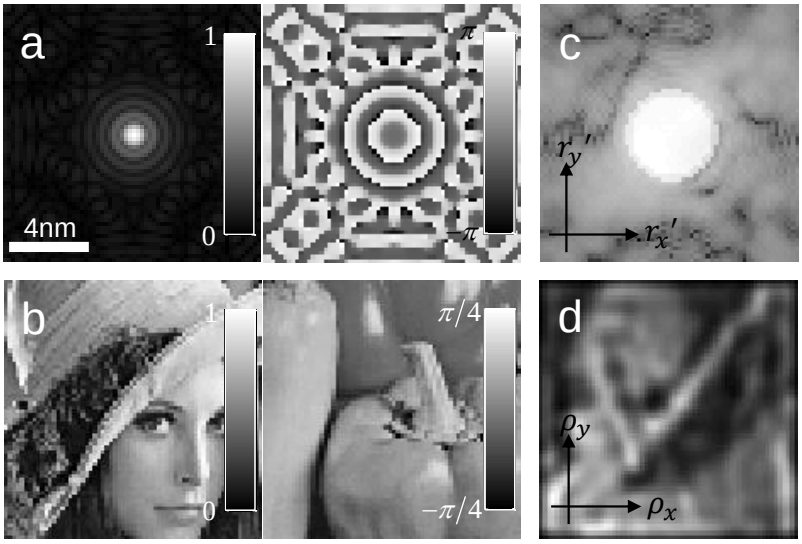

Figure 3.2: The 4D intensity data simulations. (a) The probe model. The left is the modulus part and the right phase part. The modulus part is square rooted for display purpose. (b) The object model. The left is the modulus part and the right phase part. (c) The 2D slice of the 4D intensity data at $\boldsymbol{\rho} = 0$(logarithmic scale). (d) The 2D slice of the 4D intensity data at $\boldsymbol{r}' = 0$.

of the image recorded at the source plane using a point source placed at the position of $\boldsymbol{r}'_x = 0$ and $\boldsymbol{r}'_y = 0$ (in this case, the illumination is on the optical axis, not tilted). It should be noticed that the image is a rotated version of the object and that the resolution is limited by the small aperture size. Other 2D slices at different $\boldsymbol{r}'$ positions are the images of tilted illuminations.

3.1.2 The derivation of WDDM

According to Eq. (3.1), the 4D intensity dataset $\left|M(\boldsymbol{r}', \boldsymbol{\rho})\right|^2$ can be expressed as

$$\left|M(\boldsymbol{r}', \boldsymbol{\rho})\right|^2 = \iint a(\boldsymbol{\xi} - \boldsymbol{\rho}) a^*(\boldsymbol{\eta} - \boldsymbol{\rho}) \psi(\boldsymbol{\xi}) \psi^*(\boldsymbol{\eta}) e^{-i2\pi(\boldsymbol{\xi} - \boldsymbol{\eta})\boldsymbol{r}'} \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{\eta}, \tag{3.2}$$

where the asterisk denotes the complex conjugate, and $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are the dummy variables of the specimen coordinate $\boldsymbol{r}$. The key mathematical step in WDDM is to take a Fourier transform of the intensity dataset with respect to the scan positions $\boldsymbol{\rho}$ and an inverse Fourier transform with respect to $\boldsymbol{r}'$, such that

$$H(\boldsymbol{r}, \boldsymbol{\rho}') = \iint \iint a(\boldsymbol{\xi} - \boldsymbol{\rho}) a^*(\boldsymbol{\eta} - \boldsymbol{\rho}) \psi(\boldsymbol{\xi}) \psi^*(\boldsymbol{\eta}) e^{-i2\pi(\boldsymbol{\xi} - \boldsymbol{\eta})\boldsymbol{r}'} e^{-i2\pi\boldsymbol{\rho}\boldsymbol{\rho}'} e^{i2\pi\boldsymbol{r}\boldsymbol{r}'} \mathrm{d}\boldsymbol{\xi} \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{\rho} \mathrm{d}\boldsymbol{r}' \tag{3.3}$$

We separate out the part of Eq. (3.3) dependent on $\boldsymbol{r}'$ and make use of the following relation

$$\int e^{-i2\pi(\boldsymbol{\xi} - \boldsymbol{\eta} - \boldsymbol{r})\boldsymbol{r}'} \mathrm{d}\boldsymbol{r}' = \delta\left(\boldsymbol{\xi} - \boldsymbol{\eta} - \boldsymbol{r}\right), \tag{3.4}$$

where δ is the Dirac function, Eq. (3.3) can be simplified to

$$H(\boldsymbol{r}, \boldsymbol{\rho}') = \iint a(\boldsymbol{r} + \boldsymbol{\eta} - \boldsymbol{\rho}) a^*(\boldsymbol{\eta} - \boldsymbol{\rho}) \psi(\boldsymbol{r} + \boldsymbol{\eta}) \psi^*(\boldsymbol{\eta}) e^{-i2\pi\boldsymbol{\rho}\boldsymbol{\rho}'} \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{\rho}. \tag{3.5}$$

Using the variable substitution $\boldsymbol{\sigma} = \boldsymbol{\eta} - \boldsymbol{\rho}$ for Eq. (3.5), we have

$$H(\boldsymbol{r}, \boldsymbol{\rho}') = \iint a(\boldsymbol{r} + \boldsymbol{\sigma}) a^*(\boldsymbol{\sigma}) \psi(\boldsymbol{r} + \boldsymbol{\eta}) \psi^*(\boldsymbol{\eta}) e^{-i2\pi(\boldsymbol{\eta} - \boldsymbol{\sigma})\boldsymbol{\rho}'} \mathrm{d}\boldsymbol{\eta} \mathrm{d}\boldsymbol{\sigma}. \tag{3.6}$$

If we separate the two integrals in Eq. (3.6) and define a quantity for any general function $q(\boldsymbol{r})$ as

$$\chi_q(x,v) = \int q(x+y)q^*(y)e^{-i2\pi yv}\mathrm{d}y, \tag{3.7}$$

Eq. (3.6) can be then rewritten as

$$H(\boldsymbol{r},\boldsymbol{\rho}') = \chi_a(\boldsymbol{r},-\boldsymbol{\rho}')\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}'). \tag{3.8}$$

Eq. (3.7) is the definition of Wigner distribution function (WDF) [25]. The advantage of forming $H(\boldsymbol{r},\boldsymbol{\rho}')$ becomes obvious now because it makes the probe and the specimen separable.

Using the simulated model in the above section, we calculate the WDF of the probe $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ as shown in Fig. 3.3a, the WDF of the object $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ as shown in Fig. 3.3b, and the product of the two WDFs as shown in Fig 3.3c. Also, according to its definition in Eq. (3.2), we calculate $H(\boldsymbol{r},\boldsymbol{\rho}')$ by taking the Fourier transform of the 4D

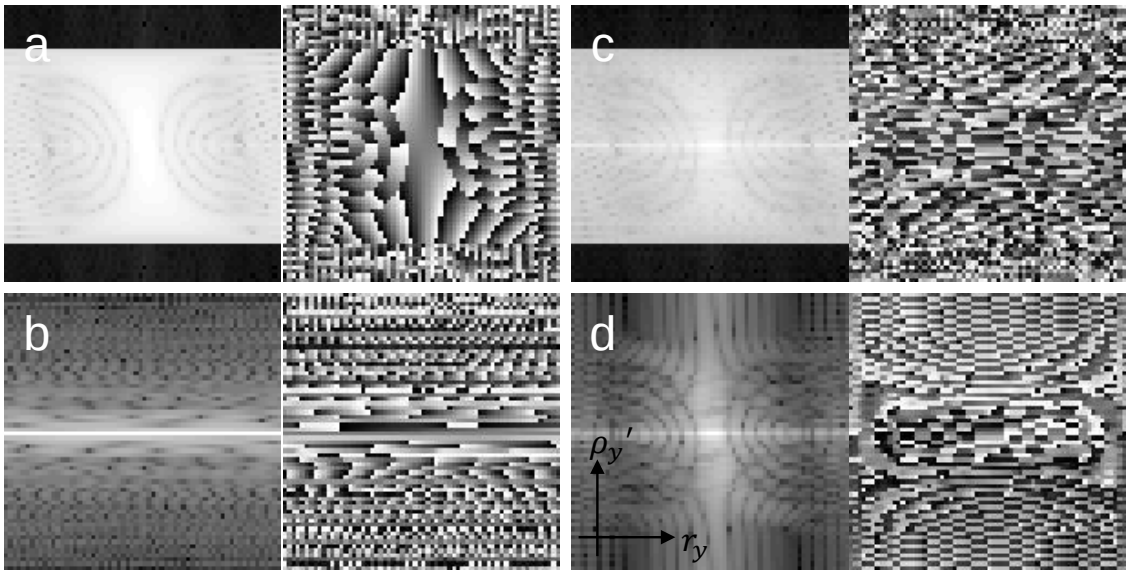

Figure 3.3: The model calculations of the two WDFs and the Fourier transform of the 4D intensity data (logarithmic scale). (a) The WDF of the probe. (b) The WDF of the object. (c) The product of the two WDFs. (d) The Fourier transform of the 4D intensity data, i.e. $H(\boldsymbol{r},\boldsymbol{\rho}')$. For all the figures, the left is the modulus part and the right the phase part. The coordinate indicator applies to all the images.

intensity data as shown in Fig. 3.3d. It should be noted that all these quantities are 4D and that the corresponding images shown are their central 2D slices with $\boldsymbol{r}'_x = 0$ and $\boldsymbol{\rho}_x = 0$ at logarithmic scale. Hereinafter, if it is not explicitly stated, the image of a 4D data will be shown in the same fashion. According to Eq. (3.8), the calculated $H(\boldsymbol{r}, \boldsymbol{\rho}')$ should be equal to the product of the two WDFs, i.e. Fig. 3.3c should be the same as Fig. 3.3d. But clearly, they are not the same.

The reason is that the two variables, $\boldsymbol{\sigma}$ and $\boldsymbol{\eta}$, were treated as two independent variables from Eq. (3.6) to Eq. (3.8), as in the original derivation [20, 26]. However, the variable substitution $\boldsymbol{\sigma} = \boldsymbol{\eta} - \boldsymbol{\rho}$ we used above indicates that the two variables are related. The separation of them has conditions. We separate out the part of Eq. (3.8) dependent on $\boldsymbol{\sigma}$ and convert the Fourier integral into discrete Fourier transform as

$$U(\boldsymbol{r}, \boldsymbol{\rho}') = \sum_{\boldsymbol{\sigma}=\boldsymbol{\eta}-M+1}^{\boldsymbol{\eta}} a(\boldsymbol{r}+\boldsymbol{\sigma}) a^*(\boldsymbol{\sigma}) e^{i2\pi\boldsymbol{\sigma}\boldsymbol{\rho}'/M}. \tag{3.9}$$

Clearly, the value of $U(\boldsymbol{r}, \boldsymbol{\rho}')$ relies on $\boldsymbol{\eta}$ unless $\boldsymbol{\sigma}$ being a cyclic coordinate. This means the scan coordinate $\boldsymbol{\rho}$ $(= \boldsymbol{\eta} - \boldsymbol{\sigma})$ needs to be cyclic as well. However, in a real experiment the scan is linear. The strategy to make this separation possible is to impose constraint on the sizes of the probe function and the specimen function so that a linear scan is equivalent to a circular scan. A circular scan means the parts of probe that exceed the extent of the specimen will wrap around to the opposite side of the specimen. This is equivalent to a linear scan with a periodic specimen (see Fig. 3.4). Obviously, when the sum of the probe support size Ω_a and the specimen support size Ω_ψ is smaller than the calculation window size Ω, a linear scan will generate the same dataset as a circular scan does. Therefore, the condition for the separation of $\boldsymbol{\sigma}$ and $\boldsymbol{\eta}$ in Eq. (3.6) becomes

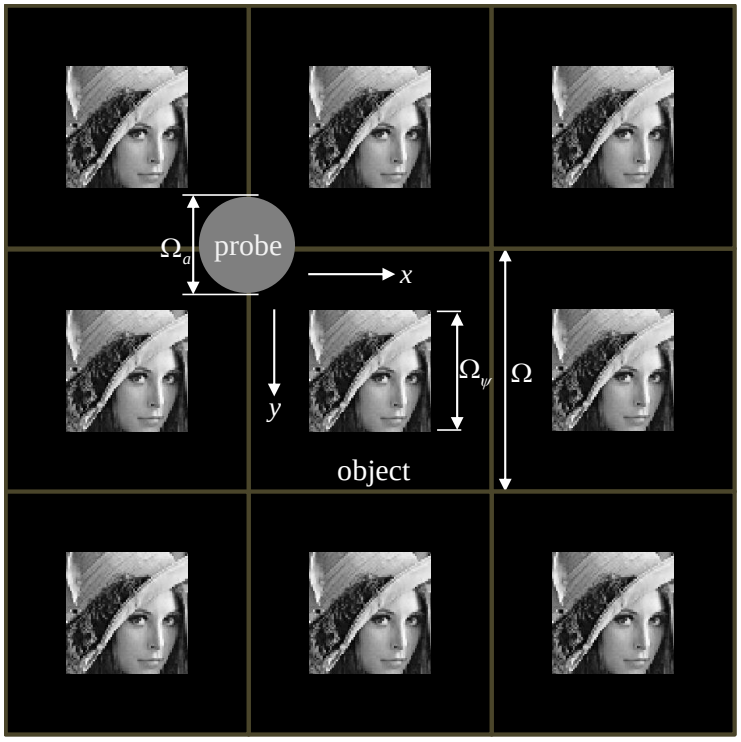

Figure 3.4: Illustration of the connection between circular and linear scan using a periodic specimen. The transmission coefficient of the dark margin is zero. The white arrows indicate the scanning directions. A linear scan over the periodic specimen is equivalent to circular scan over its one period. The separation of probe function and specimen function in WDDM relies on linear scan being equivalent to circular scan.

$$\Omega_a + \Omega_\psi \le \Omega \,. \tag{3.10}$$

This requirement applies to both the x and y coordinates of $\boldsymbol{\rho}$ and it allows us to separate the two integrals in Eq. (3.6). This actually is the sampling condition for the Fourier transform acting on $\boldsymbol{\rho}$, analogous to the sampling condition for the Fourier transform acting on $\boldsymbol{r}$ which is the probe size should be smaller than the calculation window size (i.e., $\Omega_a \le \Omega$). Therefore, violation of Eq. (3.10) will cause aliasing in $\boldsymbol{\rho}$.

The aliasing problem can be mitigated by multiplying the 4D intensity data with a window function $w(\boldsymbol{\rho})$ that is unity at its central region and falls smoothly to zero at its edges [21, 26], because the separation error is bigger when the scan position $\boldsymbol{\rho}$ is at the edges than at the centre. To make sure the separation condition is fully fulfilled, we crop

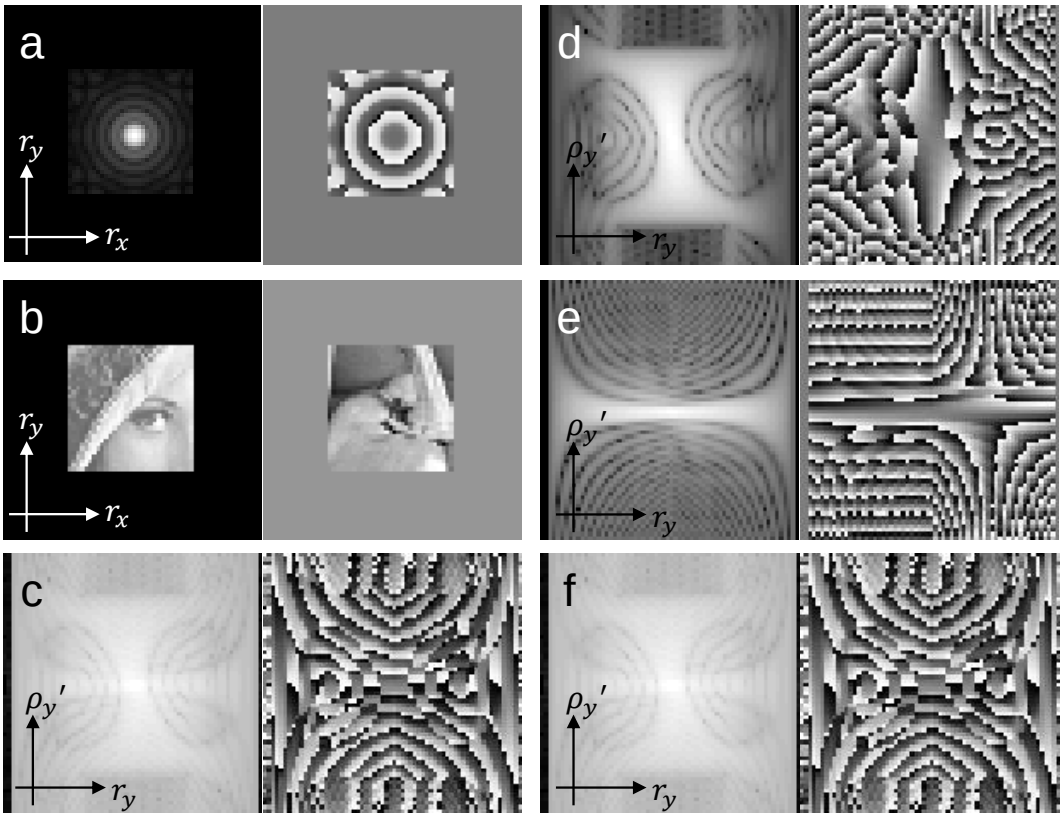

Figure 3.5: The new simulated models. (a) The cropped probe model. (b) The cropped object model. (c) The Fourier transform of the 4D intensity data, i.e. $H(\boldsymbol{r},\boldsymbol{\rho}')$. (d) The WDF of the probe. (e) The WDF of the object. (f) The product of the two WDFs. For all the figures, the left is the modulus part and the right phase part.

both the probe and the object into half of the calculation window as shown in Figs. 3.5a and 3.5b, although they might not be realistic in real experiments. A 4D intensity data is accordingly generated. In the same, we calculate the two WDFs (respectively shown in Figs. 3.5d and 3.5e), their product (shown in Fig. 3.5f), and $H(\boldsymbol{r},\boldsymbol{\rho}')$ (shown in Fig. 3.5c). Now, it is obvious that the calculated $H(\boldsymbol{r},\boldsymbol{\rho}')$ is the same as the product of the two WDFs. The following calculations will use this new model.

3.1.3 The frequency cut-off

If taking the Fourier transform of $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ with respect to $\boldsymbol{r}$, we have

$$D(\boldsymbol{r}',\boldsymbol{\rho}') = \Psi^*(\boldsymbol{r}'-\boldsymbol{\rho}')\Psi(\boldsymbol{r}'), \tag{3.11}$$

where $\Psi(\boldsymbol{r}')$ is the Fourier transform of the specimen function $\psi(\boldsymbol{r})$. The quantity $D(\boldsymbol{r}', \boldsymbol{\rho}')$ includes all the information of the interference between any two diffraction orders separated by $-\boldsymbol{\rho}'$. It is this information that allows us to solve for the phase problem. To obtain this information, we need to perform a deconvolution on $H(\boldsymbol{r}, \boldsymbol{\rho}')$ to separate out the WDF of the probe function. The probe function can be estimated given the experimental parameters are known, so $\chi_a(\boldsymbol{r}, -\boldsymbol{\rho}')$ can be calculated according to Eq. (3.7). The deconvolution is carried out via a division as

$$\chi_\psi(\boldsymbol{r}, \boldsymbol{\rho}') = \frac{\chi_a^*(\boldsymbol{r}, -\boldsymbol{\rho}')H(\boldsymbol{r}, \boldsymbol{\rho}')}{\left|\chi_a(\boldsymbol{r}, -\boldsymbol{\rho}')\right|^2 + \epsilon}, \tag{3.12}$$

where ϵ is some small constant to avoid the division by zero. The deconvolved $\chi_\psi(\boldsymbol{r}, \boldsymbol{\rho}')$ of the new model calculation is shown in Fig. 3.6a and it is the same as the WDF of the object shown in Fig. 3.5e that is calculated according to its definition. The Fourier transform of the deconvolved $\chi_\psi(\boldsymbol{r}, \boldsymbol{\rho}')$ with respect to $\boldsymbol{r}$, i.e. $D(\boldsymbol{r}', \boldsymbol{\rho}')$ according to Eq. (3.11), is shown in Fig. 3.6b.

Prior to the deconvolution, we will give a little more introduction about the probe function and its WDF. In a real situation, due to the partial coherence, experimental instabilities and the finite extent of the lens, in the back focal plane of the lens there is a

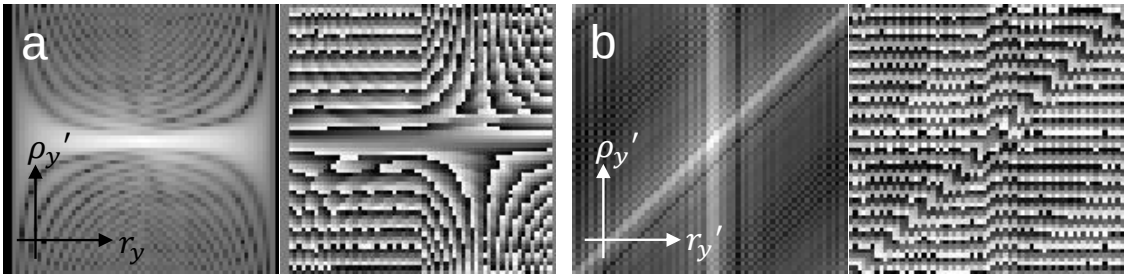

Figure 3.6: The quantities after the deconvolution of the new model calculation. (a) The deconvolved $\chi_\psi(\boldsymbol{r}, \boldsymbol{\rho}')$. (b) The corresponding $D(\boldsymbol{r}', \boldsymbol{\rho}')$ by taking the Fourier transform of (a) with respect to $\boldsymbol{r}$. For all the figures, the left is the modulus part and the right phase part.

cut-off beyond which the spatial frequencies cannot be expressed in the image [18, 24, 26, 27]. This is why the bright field image of a conventional microscope has an information limit (as shown in Fig. 3.2d). This frequency cut-off also causes a finite extent of $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ along $\boldsymbol{\rho}'$. To understand this better, we rewrite $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ as

$$\chi_a(\boldsymbol{r},-\boldsymbol{\rho}') = \int A^*(\boldsymbol{r}'+\boldsymbol{\rho}')A(\boldsymbol{r}')e^{i2\pi\boldsymbol{r}\boldsymbol{r}'}\mathrm{d}\boldsymbol{r}', \tag{3.13}$$

where $A(\boldsymbol{r}')$ is the Fourier transform of the probe function $a(\boldsymbol{r})$. If we denote the frequency cut-off as $\boldsymbol{\rho}'_{\max}$, we can infer that $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ has a support size of $2\boldsymbol{\rho}'_{\max}$ along $\boldsymbol{\rho}'$. This can be clearly seen in the calculation shown in Fig. 3.3a that has a finite support region twice the size of the aperture of the objective lens along $\boldsymbol{\rho}'$. This, in return, results in a finite support size of $2\boldsymbol{\rho}'_{\max}$ on $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ along $\boldsymbol{\rho}'$ after the deconvolution, i.e. a limited number of spatial frequencies can be accessible along $\boldsymbol{\rho}'$. However, the cut-off is not seen in the new calculation shown in Fig. 3.5d and neither in the deconvolved $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ and its Fourier transform shown in Fig. 3.6. This is because the new probe model is cropped in real space. However, to demonstrate the frequency cut-off, three fourths of the data $D(\boldsymbol{r}',\boldsymbol{\rho}')$ at high $\boldsymbol{\rho}'$ are manually thrown away, as shown in Fig. 3.7.

In the next section, we will introduce a projection strategy to overcome this frequency

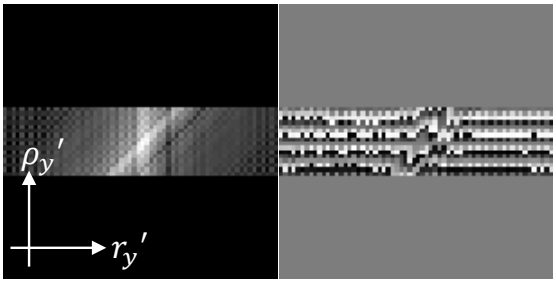

Figure 3.7: The $D(\boldsymbol{r}',\boldsymbol{\rho}')$ with three fourths of its complete data thrown away at high $\boldsymbol{\rho}'$. The left is the modulus part and the right the phase part.

cut-off problem and obtain the information of the frequencies higher than the cut-off by exploring the data $D(\boldsymbol{r}', \boldsymbol{\rho}')$ along $\boldsymbol{r}'$.

3.1.4 The projection strategy

Although the high frequency information of the specimen is lost along $\boldsymbol{\rho}'$, the detector is able to collect the high-angle, high-resolution information in $\boldsymbol{r}'$. All this information is stored in the dataset $D(\boldsymbol{r}', \boldsymbol{\rho}')$. We need an algorithm to utilize the data lying in the $\boldsymbol{r}'$ direction while avoiding the constraint in the $\boldsymbol{\rho}'$ direction. Rodenburg and Bates used a method called 'stepping out' that only uses a subset in $D(\boldsymbol{r}', \boldsymbol{\rho}')$ [26]. For example, consider a specimen that contains five frequency components Ψ_{-2}, Ψ_{-1}, Ψ_0, Ψ_1, Ψ_2 and its $D(\boldsymbol{r}', \boldsymbol{\rho}')$ has a support size that only allows the expression of three components along $\boldsymbol{\rho}'$ (see Fig. 3.8, the unshaded region between $-\boldsymbol{\rho}'_{\max}$ and $\boldsymbol{\rho}'_{\max}$). In the stepping out method, one chooses a route, as indicated by the black arrows, containing all the frequency information of the specimen. The modulus of Ψ_2 can be calculated from $\Psi_2{}^*\Psi_2$. By assigning an arbitrary phase to $|\Psi_2|$, one then can derive Ψ_1 from $\Psi_2{}^*\Psi_1$. In the same way, Ψ_0, Ψ_{-1} and Ψ_{-2} can be recovered in turn from

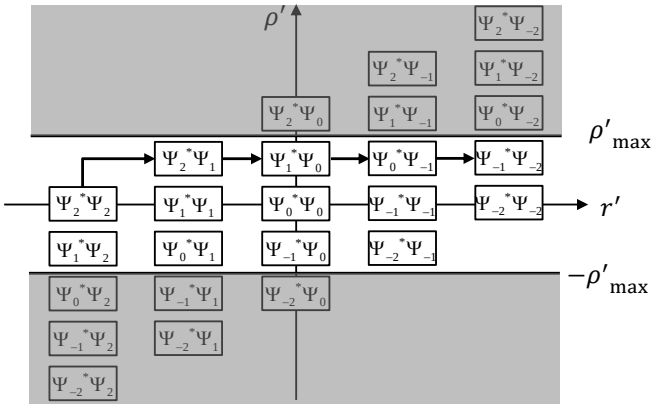

Figure 3.8: Demonstration of the stepping out method using the dataset $D(\boldsymbol{r}', \boldsymbol{\rho}')$ of a simple specimen containing five frequency components. The arrow line shows one of the routes to implement the stepping out method.

$\Psi_1^*\Psi_0$, $\Psi_0^*\Psi_{-1}$ and $\Psi_{-1}^*\Psi_{-2}$ respectively. Although the stepping out method can manage to recover all the frequency components, one has to manually pick a suitable fraction from the dataset $D(\boldsymbol{r}', \boldsymbol{\rho}')$ and cannot make full use of the information available. In the presence of noise, if a wrong data segment is chosen to start with, the error will accumulate as the stepping out procedure progresses.

Here we improve the stepping out method by introducing a projection strategy that recovers all the information as well as improving the robustness of the method to noise. The projection method is expressed as

$$\Psi(\boldsymbol{r}') = \frac{\sum_{\boldsymbol{\rho}'=-\boldsymbol{\rho}'_{\max}}^{\boldsymbol{\rho}'_{\max}} \Psi(\boldsymbol{r}'-\boldsymbol{\rho}')D(\boldsymbol{r}', \boldsymbol{\rho}')}{\sum_{\boldsymbol{\rho}'=-\boldsymbol{\rho}'_{\max}}^{\boldsymbol{\rho}'_{\max}} \left|\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')\right|^2}. \tag{3.14}$$

The two summations in the above equation are carried out along $\boldsymbol{\rho}'$, like a projection in this direction. This is where the name projection strategy comes from. It should be noticed that because of the cut-off $\pm\boldsymbol{\rho}'_{\max}$ the implementation of the projection only recovers the frequency of $\Psi(\boldsymbol{r}')$ up to $2\boldsymbol{\rho}'_{\max}$ each time, so it may need to be repeated if one implementation of the projection is not enough to recover all the frequencies. Using the same example in Fig. 3.8, the implementation of the projection method is demonstrated in Fig. 3.9. By assigning an arbitrary phase to the modulus of Ψ_0 that is given by the square root of $D(0,0)$, i.e. $\Psi_0^*\Psi_0$, we obtain the first component of $\Psi(\boldsymbol{r}')$. Making use of the data lying in the column at $\boldsymbol{r}'=0$, i.e. $D(0,\boldsymbol{\rho}')$, we can extract two more components by doing the calculation $\left[0,\Psi_{-1},\Psi_0,\Psi_1,0\right]=\Psi_0^*D(0,\boldsymbol{\rho}')\big/\left|\Psi_0\right|^2$, still losing the two highest components. Then we apply the projection strategy by calculating

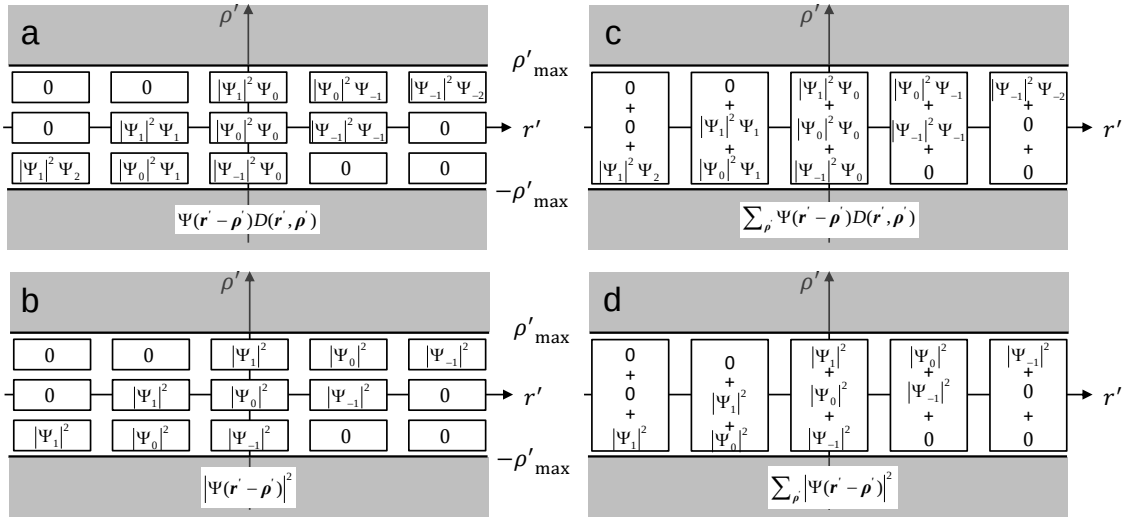

Figure 3.9: Implementation of the projection strategy using the dataset $D(\boldsymbol{r}',\boldsymbol{\rho}')$ of a simple specimen containing five frequency components. First, we calculate (a) $\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')D(\boldsymbol{r}',\boldsymbol{\rho}')$ and (b) $\left|\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')\right|^2$. We then sum along $\boldsymbol{\rho}'$ to get (c) $\sum_{\rho'}\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')D(\boldsymbol{r}',\boldsymbol{\rho}')$ and (d) $\sum_{\rho'}\left|\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')\right|^2$ respectively. Finally, the division of (c) and (d) gives all the frequencies of $\Psi(\boldsymbol{r}')$.

$$\left[\Psi_{-2},\Psi_{-1},\Psi_0,\Psi_1,\Psi_2\right]=\sum_{\boldsymbol{\rho}'=-1}^{1}\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')D(\boldsymbol{r}',\boldsymbol{\rho}')\Big/\sum_{\boldsymbol{\rho}'=-1}^{1}\left|\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')\right|^2.$$

In this case, one projection is enough to solve all the components in the specimen. Once recovering all the frequencies, we only need to implement an inverse Fourier transform to get the complex image of the specimen.

The new model calculation is also used to demonstrate the projection strategy as shown in Fig. 3.10. The low-resolution image shown in Fig. 3.10a corresponds to the frequencies available lying on $D(0,\boldsymbol{\rho}')$ that has only one fourth of the complete frequency information at the low angles. Using this available low frequency information on $D(0,\boldsymbol{\rho}')$, after we apply the projection strategy by three times (each time it recovers one fourth of the frequency information), all the frequency information is recovered. The corresponding image after each time of projection is respectively shown in Figs. 3.10b

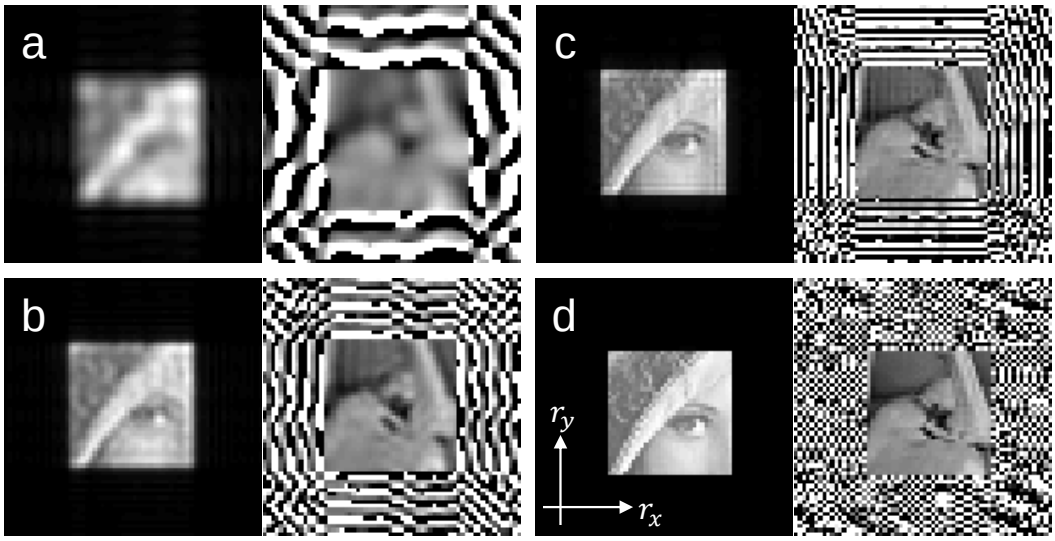

Figure 3.10: The reconstructed object functions after applying the projection strategy by (a) zero time, (b) one time, (c) two times, and (d) three times. For all the figures, the left is the modulus part and the right the phase part.

(one time), 3.10c (two times) and 3.10d (three times). Fig. 3.10d has complete frequency information and it looks the same as the reference given in Fig. 3.5b.

3.2 Noise effects on WDDM

The ideal simulation shown above demonstrates the success of the projection strategy and the reconstruction via WDDM. However, we have not considered a very practical problem in the real situations, that is, the statistical noise produced during the recording process. In this section, we will consider the noise effects and introduce two ways to suppress them: probe design and an iterative method.

3.2.1 Noise suppression via the probe design

Using the above calculation model, here we add Poisson noise to the measured intensity data with a level of 10^6 total counts per diffraction pattern (about 244 counts per pixel). The reconstructions are shown in Fig. 3.11 (the probe function and its WDF are shown here again for comparison). Compared to the ideal case (see Fig. 3.6a), the deconvolved $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$, as sown in Fig. 3.11c, has serious numerical errors, especially at the edges

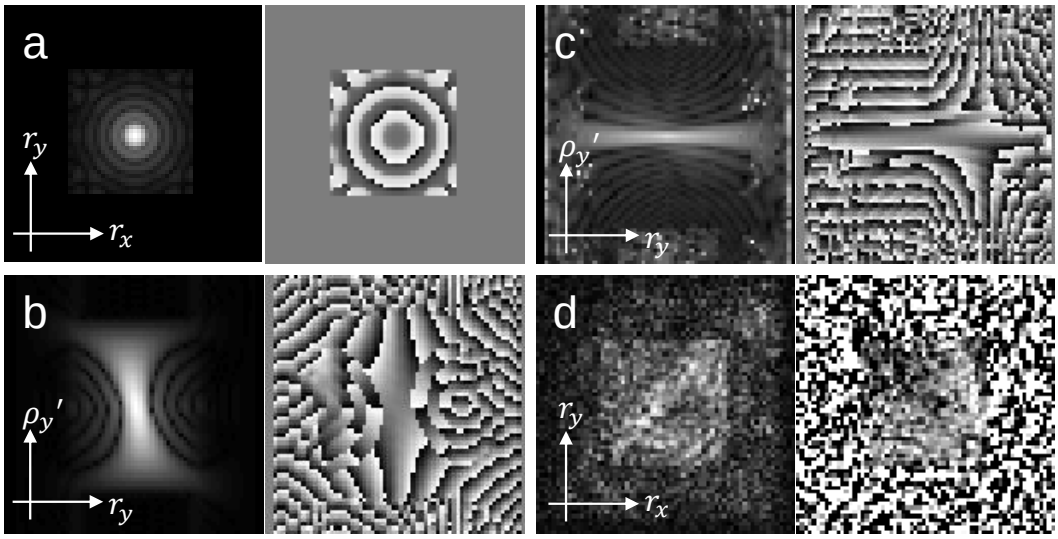

Figure 3.11: The WDDM reconstruction of the data generated from the defocused probe. Poisson noise with total counts of 10^6 for each diffraction pattern is added. (a) The defocus probe. (b) The WDF of the probe function. (c) The deconvolved WDF of the specimen function. (d) The final reconstruction of the specimen function. For each subset, the left image represents the modulus part and the right one represents the phase part. The moduli of (b) and (c) are square rooted for display purpose.

where $\chi_a(\boldsymbol{r}, -\boldsymbol{\rho}')$ has small values. It eventually results in an unrecognisable reconstruction of the specimen (see Fig. 3.11d).

Although the projection strategy makes full use of the redundancy in the big dataset and can mitigate any inconsistency by averaging all the available data along the projection direction, WDDM still has very poor performance in the presence of noise due to the numerical instability caused by the division of small values during the deconvolution process [26]. An instinctive way to improve this is to produce a probe that would have an even distribution across its WDF, i.e. a rather flat $\chi_a(\boldsymbol{r}, -\boldsymbol{\rho}')$, so that there will be no division by small values. The easiest way to make such probe is via a diffuser. Using the same parameters as used above for the defocused probe, we insert a random phase diffuser (ranging from $-\pi/2$ to $\pi/2$) at the back focal plane of the lens and get a diffused probe (see Fig. 3.12a). The corresponding WDF (see Fig. 3.12b) stretches out more compared with the WDF of the defocused probe. As a result, the deconvolved $\chi_\psi(\boldsymbol{r}, \boldsymbol{\rho}')$

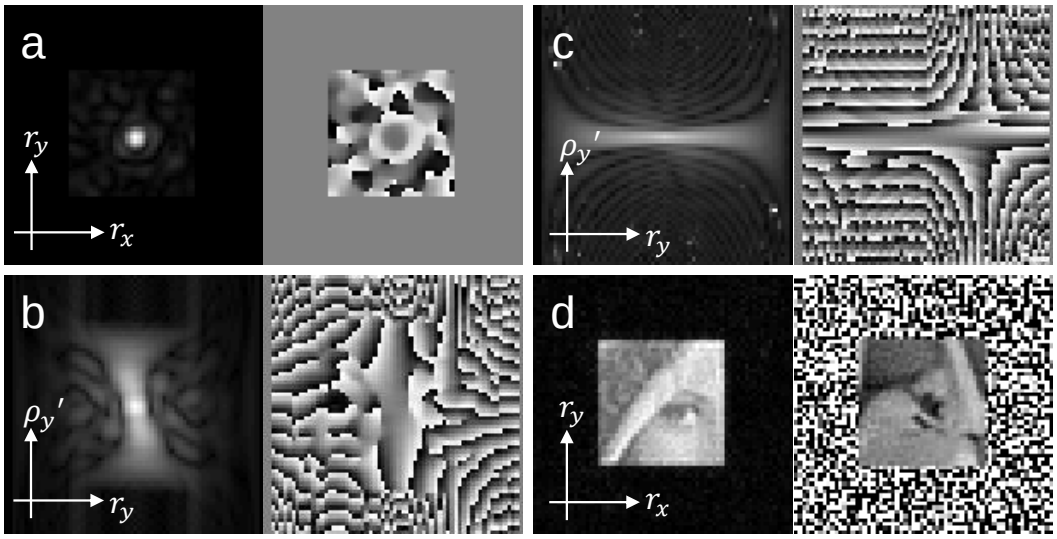

Figure 3.12: The WDDM reconstruction of the data generated from the diffused probe. Poisson noise with total counts of 10^6 for each diffraction pattern is added. (a) The diffused probe. (b) The WDF of the probe function. (c) The deconvolved WDF of the specimen function. (d) The final reconstruction of the specimen function. For each subset, the left image represents the modulus part and the right one represents the phase part. The moduli of (b) and (c) are square rooted for display purpose.

has much higher accuracy at the edges and also a much better specimen image is reconstructed (see Figs. 3.12c and 3.12d).

Although a diffused probe has substantially suppressed the noise and improved the reconstruction, there are still some noticeable errors in the specimen image. In order to optimize and engineer such a probe, we have found it is possible to use an iterative method similar to a conventional iterative ptychographic reconstruction. If reviewing the definition of WDF in Eq. (3.7) and the diffraction pattern in Eq. (3.1), we will find that they have the same formulations. In another words, the WDF of a probe function can be treated as the diffraction patterns obtained by scanning the probe over its conjugate function. This means we can use the iterative ptychographic reconstruction algorithms to find the favourable probe given a desired WDF as the diffraction intensity data. Here we use the square root of the modulus of the WDF from the above diffused probe as the desired $\chi_a(\boldsymbol{r}, -\boldsymbol{\rho}')$ (the use of the square root is trying to flatten the WDF

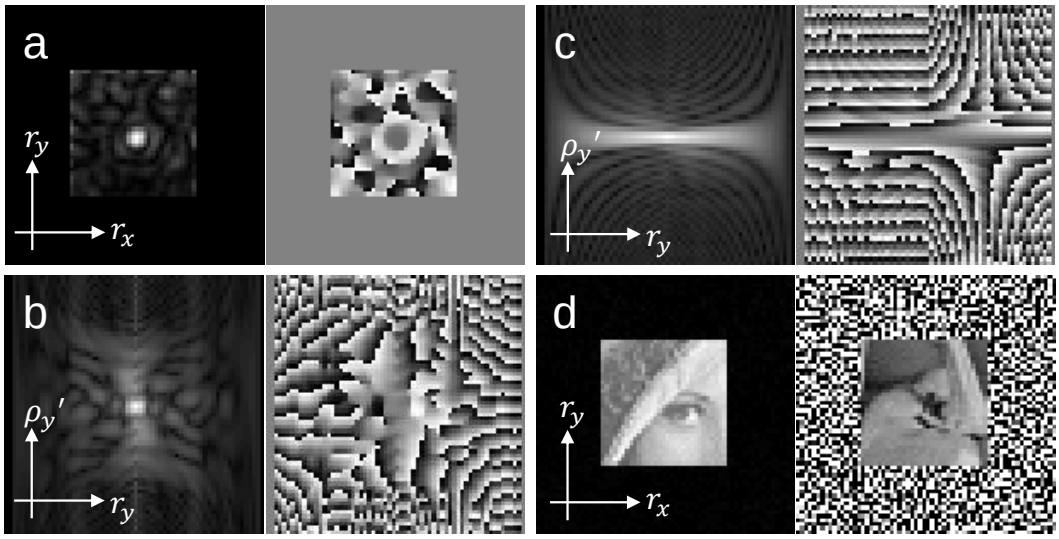

Figure 3.13: The WDDM reconstruction of the data generated from the designed probe. Poisson noise with total counts of 10^6 for each diffraction pattern is added. (a) The designed probe. (b) The WDF of the probe function. (c) The deconvolved WDF of the specimen function. (d) The final reconstruction of the specimen function. For each subset, the left image represents the modulus part and the right one represents the phase part. The moduli of (b) and (c) are square rooted for display purpose.

distribution) and apply the ePIE algorithm to solve for the favourable probe function. Because of the square root operation, the desired $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ may not correspond to a function's WDF, which means the ePIE does not necessarily converge to a solution that gives the exact $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ as we want. For this reason, we only run the ePIE for 4 iterations. In this way, we generate a third probe that we call the designed probe here (see Fig. 3.13a). The corresponding WDF completely stretches out and fills in the whole window. As a consequence, we cannot even observe any noise effect in the deconvolved $\chi_\psi(\boldsymbol{r},\boldsymbol{\rho}')$ and the specimen reconstruction (see Figs. 3.13c and 3.13d).

3.2.2 Noise suppression via an iterative method

Because of the projection operation, a reconstructed specimen function can average out some errors caused by the noise. To put it in another way, the deconvolved data with big errors is refined by the areas with small errors. This means the calculated WDF

$\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ using the reconstructed specimen will have higher accuracy than the deconvolved WDF $\chi_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ in the places where $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ has small values, while for the places where $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ has big values $\chi_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ has higher accuracy than $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$. Making use of this property, we come up with an iterative method to suppress the residual errors introduced by the deconvolution. First we need to go through all the usual steps of WDDM and obtain a non-optimally deconvolved WDF $\chi_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ and a noisy reconstructed specimen image. Then the iterative method loops through the following three procedures:

1. Use the reconstructed specimen image to calculate the WDF $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$;
2. Generate a new WDF by replacing the values of $\chi_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ in the places where $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ is smaller than a pre-set threshold *T* with the values of $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$;
3. Go through the usual WDDM steps using the new WDF to reconstruct an improved specimen image.

A block diagram of the method is shown in Fig. 3.14. As the iteration runs, the reconstructed specimen image will be improved. As a result, the accuracy of all the values in the calculated WDF $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ slowly gets higher. For this reason, we adopt

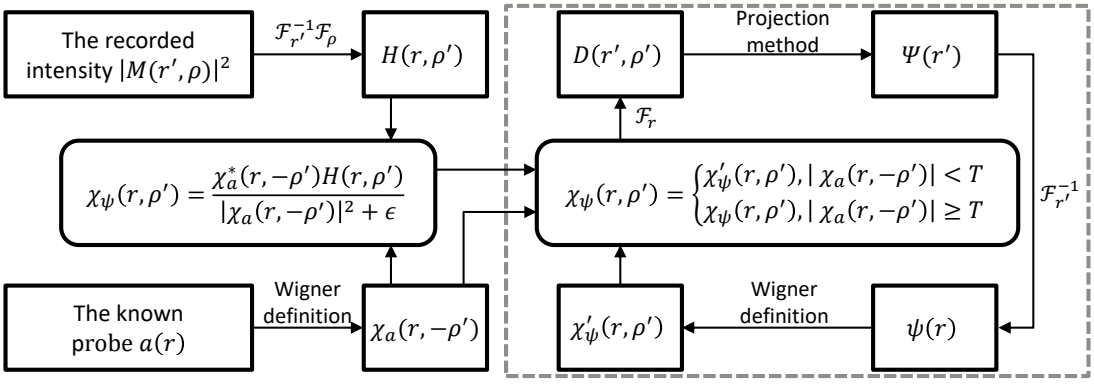

Figure 3.14: Block diagram of the iterative noise suppression method. The part in the dashed square shows the noise suppression loop. The inputs of the loop are the two WDFs and the output of the loop is the final reconstruction of the specimen function. In the diagram, $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier transform and the inverse Fourier transform respectively, and their subscripts indicate the coordinate on which the transform operates on.

an increasing threshold, which means more and more areas of $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$ for the places where $\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')$ has big values are assumed as accurate as the iteration proceeds. Eventually, the specimen image is reconstructed accurately and it gives an accurate $\chi'_{\psi}(\boldsymbol{r},\boldsymbol{\rho}')$. We apply this algorithm on the above three noisy datasets with different probes for 50 iterations and the reconstructions are shown in Fig. 3.15. Compared to the previous reconstructions (see Figs. 3.11d, 3.12d and 3.13d), the improvements are obvious, especially for the cases of the defocused probe (see Fig. 3.15a) and the diffused probe (see Fig. 3.15b).

3.2.3 Model calculations for different noise levels

The combination of the probe design and the iterative method has greatly suppressed detrimental effects of the noise introduced during the deconvolution and considerably enhanced the performance of WDDM. In this section, we will look into the noise tolerance of this combination.

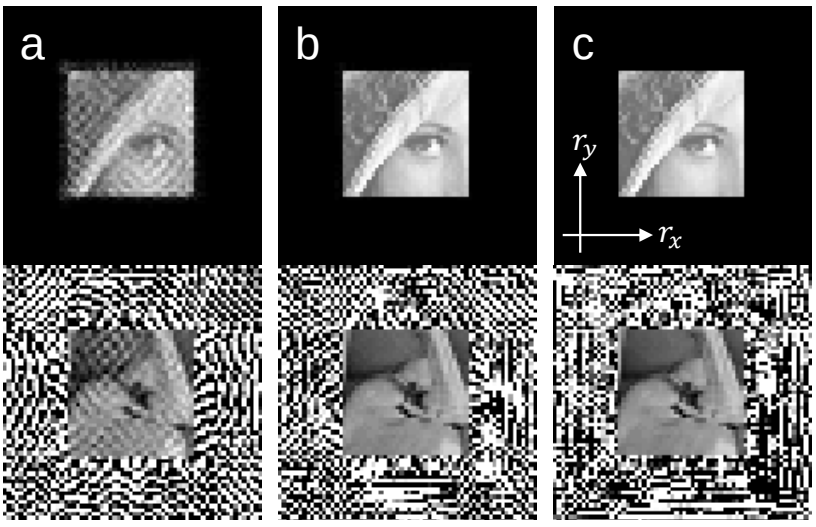

Figure 3.15: The WDDM reconstructions using the iterative noise suppression method from the three different probes: (a) the defocused probe, (b) the diffused probe and (c) the designed probe. 50 iterations are implemented for both the noise suppression WDDM and the ePIE algorithms. For each subset, the top image represents the modulus part and the bottom one represents the phase part. The coordinate indicator applies to all images.

Here we still use the above three types of probe and three more noise levels were simulated with total counts of 10^5 (about 24.4 counts per pixel), 10^4 (about 2.44 counts per pixel) and 10^3 (about 0.24 counts per pixel) in each diffraction pattern. The iterative noise suppression algorithm is run for 50 iterations on each dataset. For comparison, we also run the ePIE algorithm on the designed probe datasets. Since the probe function is assumed known during the implementation of WDDM, to be fair we also feed ePIE with the known probe function and do not update it through the whole reconstruction. The reconstructions are shown in Fig. 3.16. For the defocused probe, the reconstruction is barely recognisable when the noise level only increases to 10^5 (see Fig. 3.16a). The reconstruction quality degrades dramatically. For the diffused probe, the iterative algorithm manages to give a good reconstruction until the noise level reaches to 10^3 (see Fig. 3.16j). The noise tolerance of the designed probe is much better than the other two probes and it still produces a satisfactory reconstruction for the noise level of 10^3 (see Fig. 3.16k). For the low noise levels (10^5 and 10^4), the designed probe is able to recover the specimen image without any observable detrimental noise effect (see Fig. 3.16c and 3.16g). It even gives better reconstructions than the ePIE algorithm does for high noise levels (10^4 and 10^3, see the comparisons between Figs. 3.16g and 3.16h, and Figs. 3.16k and 3.16i). The reason is because the increasing threshold strategy in the iterative method actually improves the measured intensity dataset at the places where have low signal-to-noise ratio using the information from the places where have high signal-to-noise ratio, while the ePIE algorithm makes no changes to the measured dataset and keeps using it to update the specimen image. This implies that the ePIE algorithm also has space for improvements when processing very noisy datasets.

To compare the reconstructions quantitatively, here we calculate the complex image errors for all the reconstructions with respect to the simulated specimen image (see Fig.

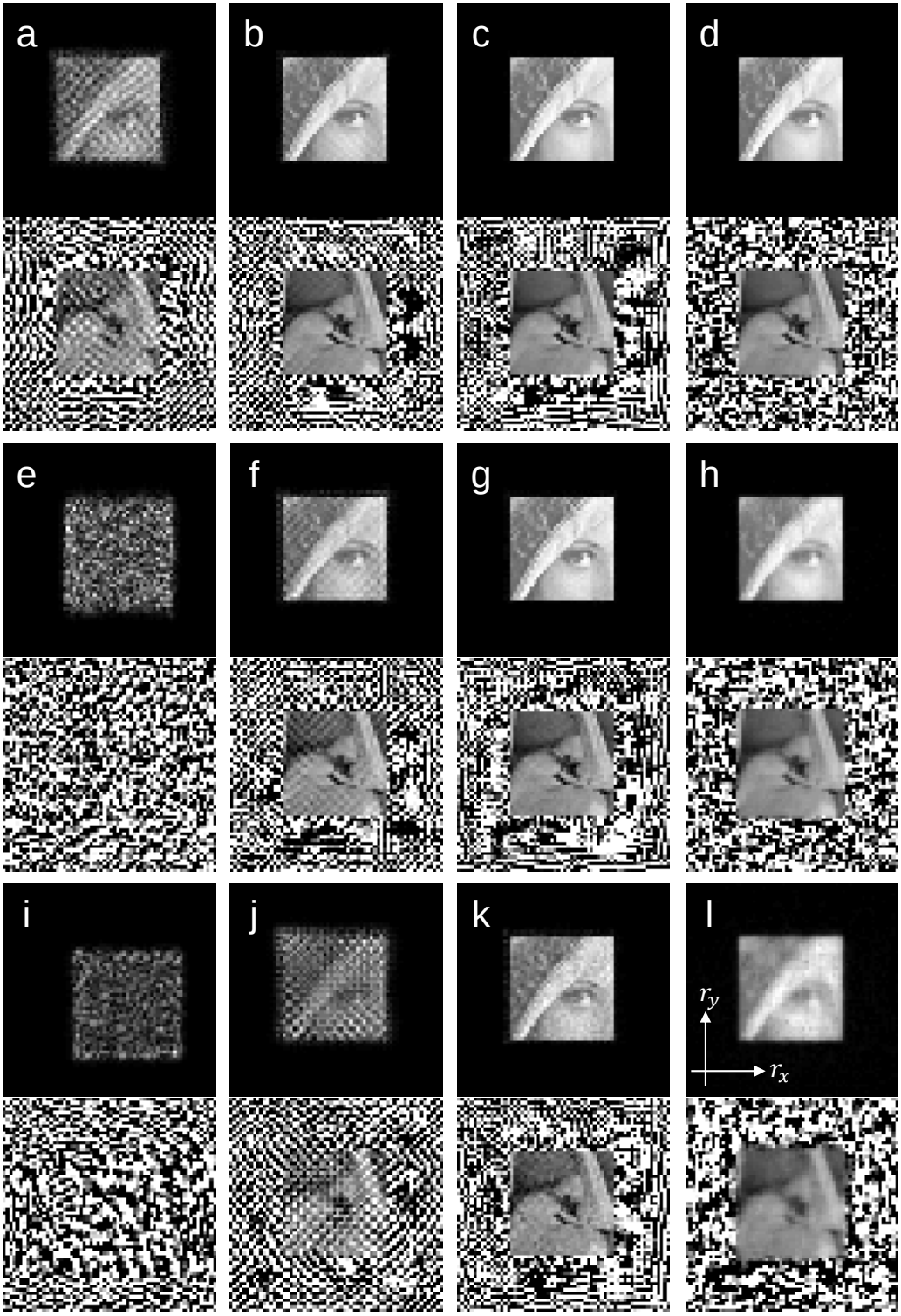

Figure 3.16: The reconstructions from the three different probes with different noisy datasets. The first column to the third column respectively represent the results from the defocused probe, the diffused probe and the designed probe using iterative noise suppression WDDM. The last column represents the results from the designed probe using ePIE. From (a) to (d), (e) to (h), and (i) to (l), the intensity data respectively has a noise level of 10^5, 10^4 and 10^3. For each subset, the top image represents the modulus part and the bottom one represents the phase part. The coordinate indicator applies to all the images.

3.5b). The errors are listed in Table 3.1. The quality degradation trend of the reconstructions is consistent with the visual comparisons indicated in Fig. 3.16. In addition, the stability of the WDDM performance under an increasing noise level for different probes is much more obvious in the tabulated values. The optimal probe is no doubt the most stable case, because from a noise level of 10^6 to 10^3 the reconstruction error only increases by a factor of 4, while it is 20 for the case of the optimal probe

Table 3.1: Normalized mean square errors of different reconstructions from datasets with different noise levels

Noise level	WDDM with normal probe	WDDM with diffused probe	WDDM with optimal probe	ePIE with optimal probe
10^6	4.6×10^{-2}	4.0×10^{-3}	3.7×10^{-3}	2.5×10^{-3}
10^5	1.2×10^{-1}	6.7×10^{-3}	3.8×10^{-3}	2.9×10^{-3}
10^4	6.0×10^{2}	2.5×10^{-2}	5.0×10^{-3}	1.1×10^{-2}
10^3	7.4×10^{4}	3.7×10^{-1}	1.5×10^{-2}	5.8×10^{-2}

using the ePIE algorithm, 100 for the diffused probe and it obviously diverges for the normal probe. It is interesting to note that for the optimal probe case with low noise levels (10^6 and 10^5), WDDM has higher reconstruction errors compared to ePIE. It is because we throw out the high frequency data of $D(\boldsymbol{r}', \boldsymbol{\rho}')$ along $\boldsymbol{\rho}'$ to simulate the cut-off as we mentioned in section 3.1.3, which means WDDM uses much less data than ePIE during the reconstruction.

3.3 Solving for spatial partial coherence via WDDM

It has been considered by Rodenburg and Bates [26] and Nellist et al [24, 27] that the effect of an extended incoherent source can be accounted in WDDM by a convolution over the scan position $\boldsymbol{\rho}$. The measured intensity data, therefore, is modified as

$$\left|M_{\text{incoh}}(\boldsymbol{r}', \boldsymbol{\rho})\right|^2 = \left|M(\boldsymbol{r}', \boldsymbol{\rho})\right|^2 \otimes_{\boldsymbol{\rho}} \left|s(\boldsymbol{\rho})\right|^2, \tag{3.15}$$

where $s(\boldsymbol{\rho})$ is the source function and $\otimes_{\rho}$ denotes the convolution over $\boldsymbol{\rho}$. With this incoherent intensity dataset we apply the Fourier transform operations as described in Eq. (3.3) and it leads to

$$H_{\text{incoh}}(\boldsymbol{r},\boldsymbol{\rho}') = S(\boldsymbol{\rho}')\chi_a(\boldsymbol{r},-\boldsymbol{\rho}')\chi_{\psi}(\boldsymbol{r},\boldsymbol{\rho}'), \tag{3.16}$$

where $S(\boldsymbol{\rho}')$ is the Fourier transform of the source intensity profile $|s(\boldsymbol{\rho})|^2$ and it corresponds to the mutual coherence function of the source according to Van-Cittert-Zernike theorem [117]. Proceeding to do the deconvolution (see Eq. (3.12)) and the Fourier transform with respect to $\boldsymbol{r}$ (see Eq. (3.11)), we are left with

$$D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}') = S(\boldsymbol{\rho}')\Psi^*(\boldsymbol{r}'-\boldsymbol{\rho}')\Psi(\boldsymbol{r}'). \tag{3.17}$$

It is obvious now that the coherence function is separable from the WDF of the specimen. It can be calculated by

$$\begin{aligned} |S_N(\boldsymbol{\rho}')| &= \sum_{\boldsymbol{r}'} \frac{|D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}')|}{\sqrt{D_{\text{incoh}}(\boldsymbol{r}',0)D^*_{\text{incoh}}(\boldsymbol{r}'-\boldsymbol{\rho}',0)}} \\ &= \sum_{\boldsymbol{r}'} \frac{|S(\boldsymbol{\rho}')\Psi^*(\boldsymbol{r}'-\boldsymbol{\rho}')\Psi(\boldsymbol{r}')|}{|S(0)||\Psi(\boldsymbol{r}'-\boldsymbol{\rho}')||\Psi(\boldsymbol{r}')|} = \frac{|S(\boldsymbol{\rho}')|}{|S(0)|}, \end{aligned} \tag{3.18}$$

where $S_N(\boldsymbol{\rho}')$ denotes the normalisation of $S(\boldsymbol{\rho}')$ because $S(0)$ is the maximum amplitude of $S(\boldsymbol{\rho}')$. According to the definition descried in section 2.3.2, $S_N(\boldsymbol{\rho}')$ is the complex coherence degree. It should be noted here the phase information of $S_N(\boldsymbol{\rho}')$ is lost. However, if the source intensity profile (positive-real valued) is symmetric and smoothly varying, $S_N(\boldsymbol{\rho}')$ will be positive and real, which means $S_N(\boldsymbol{\rho}') = |S_N(\boldsymbol{\rho}')|$. A division is needed here to obtain $D(\boldsymbol{r}',\boldsymbol{\rho}')$ from $D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}')$ and it is given by

$$D(\boldsymbol{r}',\boldsymbol{\rho}') = \frac{D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}')}{S_N(\boldsymbol{\rho}')+\epsilon}, \tag{3.19}$$

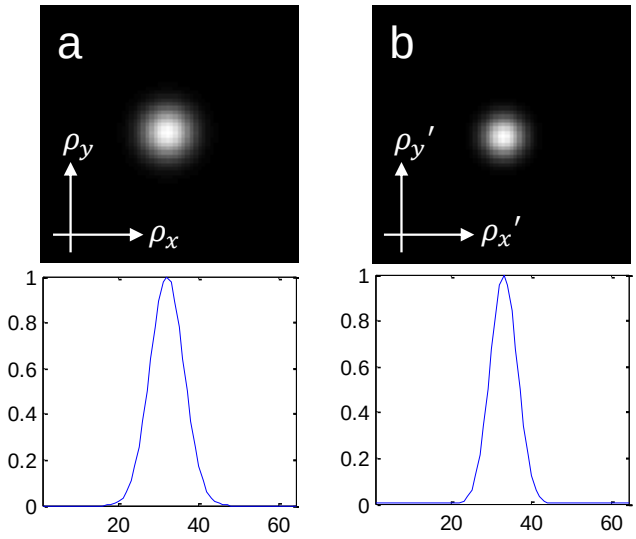

Figure 3.17: The source model. (a) The source profile (top) and its horizontal cross-section plot (bottom). (b) The coherence degree function (top) and its horizontal cross-section plot (bottom).

where ϵ is a small constant to avoid division by zeros. For an extended incoherent source, its mutual coherence function $S(\boldsymbol{\rho}')$ will have a finite support, which means it will impose information cut-off in $\boldsymbol{\rho}'$ direction on $D(\boldsymbol{r}', \boldsymbol{\rho}')$. However, as we have already demonstrated above, this will not affect the final resolution of the reconstructed specimen image.

Here we use a model calculation to demonstrate how the spatial partial coherence can be solved via WDDM. We use a Gaussian distribution with a half width of 4.5 pixels in both two Cartesian coordinates to simulate the extended source profile (see Fig. 3.17a). Its coherence degree function, given by the normalisation of the Fourier transform of the source intensity, is shown in Fig. 3.17b. The rest of the configuration remains the same as the coherent model (see section 3.1.2).

Each point in the extended incoherent source will produce the same probe function at the specimen plane with only a shift in $\boldsymbol{\rho}$ corresponding to the position of the source point. Then each of these probes will generate a spatially coherent diffraction pattern at the detector plane and the recorded intensity is the incoherent superposition of the

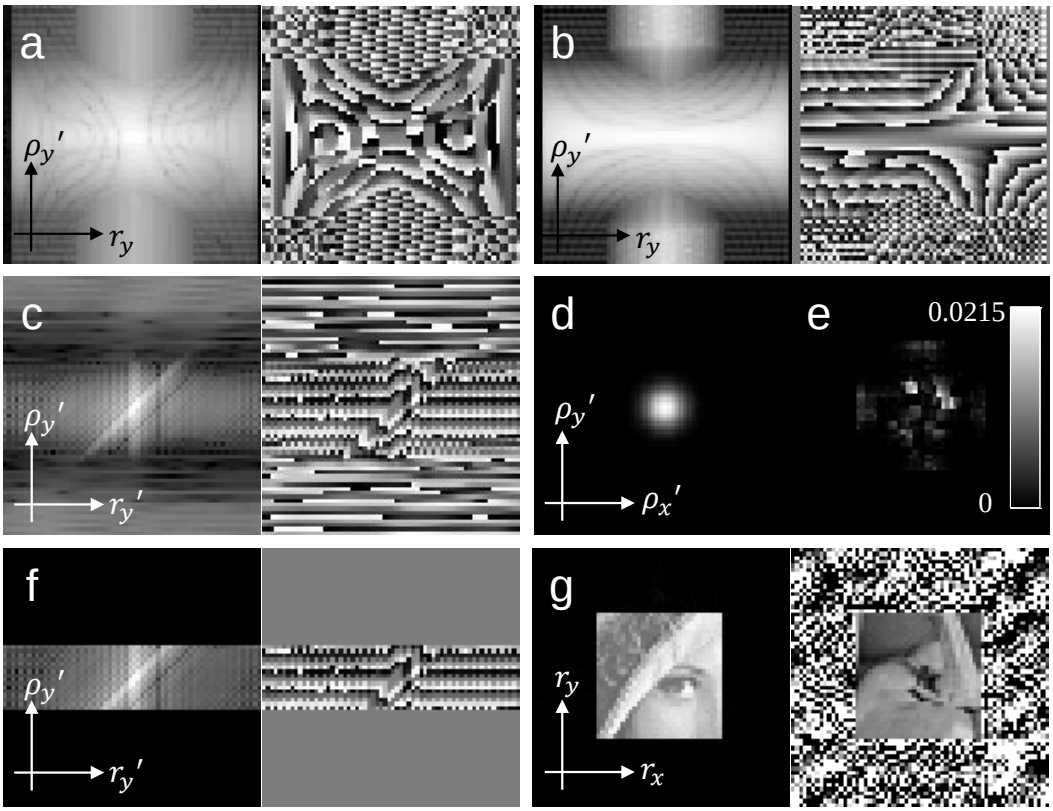

Figure 3.18: The reconstructions of the spatially partial coherent dataset via WDDM. (a) The dataset $H(\boldsymbol{r},\boldsymbol{\rho}')$. (b) The deconvolved WDF of the specimen function. (c) The complete dataset $D(\boldsymbol{r}',\boldsymbol{\rho}')$. (d) The retrieved coherence function $\Gamma(\boldsymbol{\rho}')$. (e) The difference between the reference of the coherence function and the retrieved version. (f) The cropped dataset $D(\boldsymbol{r}',\boldsymbol{\rho}')$. (g) The high-resolution reconstruction of the specimen. For all the subsets, except (d) and (e), the left image represents the modulus part and the right one represents the phase part.

intensities from all these diffraction patterns. In this way we acquire the 4D incoherent intensity dataset (noise is not considered here). However, it should be noted that Eq. (3.10) is violated here due to a bigger effective probe size caused by the extended source size. For the extended incoherent source, we need to modify Eq. (3.10) as

$$\Omega_a + \Omega_\psi + \Omega_s \le \Omega\,, \tag{3.20}$$

where Ω_s is the support size of the source. Because of this violation, aliasing occurs along $\boldsymbol{\rho}'$ direction, as we can see in $H_{\text{incoh}}(\boldsymbol{r},\boldsymbol{\rho}')$ (Fig. 3.18a) and $D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}')$ (Fig. 3.18b). Compared with the coherent case (see Fig. 3.5), it is not difficult to observe a Gaussian decay (i.e. the coherence envelope) in $H_{\text{incoh}}(\boldsymbol{r},\boldsymbol{\rho}')$ and $D_{\text{incoh}}(\boldsymbol{r}',\boldsymbol{\rho}')$ along $\boldsymbol{\rho}'$

direction. The partial coherence effect is very obvious in $D_{\text{incoh}}(\boldsymbol{r}', \boldsymbol{\rho}')$ (Fig. 3.18c), causing information cut-off in $\boldsymbol{\rho}'$. We apply Eq. (3.18) to obtain the coherence degree function as shown in Fig. 3.18d. The difference with respect to the reference (see Fig. 3.17b) is shown in Fig. 3.18e and it indicates the reconstruction errors are very small. Using this calculated coherence function, we carry out the division expressed by Eq. (3.19) over a finite region (about one fourth of the window size) where the coherence function has substantial values and it leaves us with $D(\boldsymbol{r}', \boldsymbol{\rho}')$ (see Fig. 3.18f) to proceed with the projection method. Eventually, applying the projection method three times produces the high-resolution specimen image reconstruction (see Fig. 3.18g).

3.4 Conclusions

In this chapter, we have reviewed the direct (non-iterative) solution of ptychography called WDDM. We have presented the mathematical derivation, which formularises the Fourier transform of the 4D intensity dataset to be the product of the WDFs of the specimen and the probe. However, there is an essential condition for this formularisation, namely, the sum of the specimen size and the probe size should be smaller than the calculation window size. This condition actually is the sampling requirement for WDDM. Violating it when taking the Fourier transform with respect to $\boldsymbol{\rho}$ (the scan coordinate) will cause aliasing in $\boldsymbol{\rho}'$, resulting in the same effect as the spatial partial coherence or a finite aperture at the lens back focal plane does.

WDDM allows us to separate out the influence of the probe from the specimen by a deconvolution. However, in the presence of noise the deconvolution introduces numerical errors in the places where the probe WDF has small values. We have presented a means to design a favourable probe that is able to produce an even distribution WDF, thus minimizing the effects of division by small numbers. The design

is based on that the WDF has a form analogous to the conventional ptychographic dataset. As a result, we could use ePIE to find the favourable probe given a desired probe WDF distribution.

Once the deconvolution step has been implemented, we are left with a 4D dataset representing all pairs of phase difference between all pixels in the diffraction plane. This dataset allows us to solve the phase problem and recover the high frequency information of the specimen. Unlike previous work, in which only a fraction of the 4D dataset can be used, a projection strategy proposed here is able to use all the available data, thus mitigating the inconsistencies and increasing its robustness. An iterative noise suppression method has also been presented to further improve the performance by utilising the redundancy of the 4D dataset. The basic idea is to use the information with higher accuracy to recover the information with lower accuracy.

Furthermore, we have also reconsidered the effect of an extended source size and for the first time demonstrated the reconstruction of the coherence function via WDDM using a model calculation. However, the phase part of the coherence function is not recoverable. As a result, it needs the source to be symmetric and smoothly varying, so the coherence function will not have phase information.

WDDM provides us with an insightful way to understand how information is expressed in a ptychographic dataset and in return it will also be very useful and helpful to seek improvements for the iterative methods.

4 Spatially Mixed State Ptychography

Recent algorithmic development allows ptychography to image multiple states simultaneously in both the object and the illumination [64]. The coherent case discussed before can be considered as a single state ptychography where both the probe and the object are single state. When either the probe or the object has more than one state, we refer to it as mixed state ptychography. Spatially mixed state ptychography means the mixed states are caused by or equivalent to the spatial variation in the experiment. In spatially mixed state ptychography, every probe state will interact with every object state and all the resulting states contribute to the experiment incoherently. This technique not only substantially relaxes the requirements on the experiments, such as the instrumental or environmental stability and the coherence of the radiation [64], but also further extends the applications of ptychography, such as imaging dynamic objects [73] and fly scan ptychography [74, 75, 76]. In this chapter we will theoretically and experimentally reveal the inherent ambiguities in the reconstruction algorithm and demonstrate how the ambiguities can be broken using additional constraints. Furthermore, we will also experimentally explore the effects of a diffuser on the reconstructions for a lens setup x-ray ptychography in the presence of spatial partial coherence, which is a mixed state ptychography with multiple probe states but one object state [64].

4.1 Spatially mixed state ptychography

In this section, we will give a description of the modified reconstruction algorithm for spatially mixed state ptychography based on ePIE and its relation with the gradient descent method, and present the derivation of the linear reconstruction ambiguity and

the means to break the ambiguity. Experimental demonstration will also be presented. The work presented in this section has led to a journal paper publication [118].

4.1.1 The reconstruction algorithm

Both ePIE and DM have been modified to take account of the spatially mixed state effect by Thibault et al [64]. The most critical part is the modification of the modulus constraint. Since the measured intensity is the incoherent addition of multiple diffraction states, we cannot assign this measured intensity to any single diffraction state, as we did before for single state ptychography. In the modified algorithms, the strategy is to scale the measured intensity according to the forward intensity calculations of all the diffraction states. To detail the algorithm, here we denote the *k*th illumination state as $P^{(k)}$, the *l*th object state as $O^{(l)}$, the corresponding exit wave state as $\psi^{(k,l)}$, the corresponding diffraction state as $\Psi^{(k,l)}$, the measured intensity as *M*, the iteration number as *n*, the reciprocal-space coordinate as $\boldsymbol{q}$, the real-space coordinate as $\boldsymbol{r}$ and the scan position as $\boldsymbol{r}_j$. The iterative procedures are as follows:

1. In the *n*th iteration, the exit wave guess at a probe position of $\boldsymbol{r}_j$ is formed by the product of the current object state guess and the illumination state guess at the corresponding scan position:

$$\psi_j^{(k,l)}(\boldsymbol{r}) = P_n^{(k)}(\boldsymbol{r}-\boldsymbol{r}_j)O_n^{(l)}(\boldsymbol{r}). \tag{4.1}$$

2. In the far field the diffraction pattern guess is then formed by Fourier transforming the exit wave guess:

$$\Psi_j^{(k,l)}(\boldsymbol{q}) = \mathcal{F}\left[\psi_j^{(k,l)}(\boldsymbol{r})\right]. \tag{4.2}$$

3. Apply the modulus constraint, i.e. replace the modulus of the diffraction pattern guess with the square root of the scaled intensity measurement and keep the phase part, and back propagate to real space to get an updated exit wave $\psi'^{(k,l)}$:

$$\psi'^{(k,l)}_j(\boldsymbol{r}) = \mathcal{F}^{-1}\left[\sqrt{\frac{\left|\Psi^{(k,l)}_j(\boldsymbol{q})\right|^2}{\sum_k\sum_l\left|\Psi^{(k,l)}_j(\boldsymbol{q})\right|^2}M_j(\boldsymbol{q})}\frac{\Psi^{(k,l)}_j(\boldsymbol{q})}{\left|\Psi^{(k,l)}_j(\boldsymbol{q})\right|}\right]. \tag{4.3}$$

4. The update functions for the illumination states and the object states are given as:

$$\left.\begin{aligned} P^{(k)}_{n+1}(\boldsymbol{r}) &= P^{(k)}_n(\boldsymbol{r}) + \alpha\frac{\sum_l O^{(l)*}_n(\boldsymbol{r}+\boldsymbol{r}_j)\left[\psi'^{(k,l)}_j(\boldsymbol{r}) - \psi^{(k,l)}_j(\boldsymbol{r})\right]}{\left(\sum_l\left|O^{(l)}_n(\boldsymbol{r}+\boldsymbol{r}_j)\right|^2\right)_{\max}} \\ O^{(l)}_{n+1}(\boldsymbol{r}) &= O^{(l)}_n(\boldsymbol{r}) + \beta\frac{\sum_k P^{(k)*}_n(\boldsymbol{r}-\boldsymbol{r}_j)\left[\psi'^{(k,l)}_j(\boldsymbol{r}) - \psi^{(k,l)}_j(\boldsymbol{r})\right]}{\left(\sum_k\left|P^{(k)}_n(\boldsymbol{r}-\boldsymbol{r}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{4.4}$$

In the same way as the update functions of ePIE derived in section 2.3.7, in this case the objective function is to minimise the difference between $\psi^{(k,l)}_j(\boldsymbol{r})$ and $\psi'^{(k,l)}_j(\boldsymbol{r})$, i.e.

$$E = \sum_k\sum_l\left|P^{(k)}_n(\boldsymbol{r}-\boldsymbol{r}_j)O^{(l)}_n(\boldsymbol{r}) - \psi'^{(k,l)}_j(\boldsymbol{r})\right|^2. \tag{4.5}$$

Take partial derivative with respect to each illumination state and each object state respectively, the gradients are given as

$$\left.\begin{aligned} G^{(k)}_P &= 2\sum_l O^{(l)*}_n(\boldsymbol{r}+\boldsymbol{r}_j)\left[\psi^{(k,l)}_j(\boldsymbol{r}) - \psi'^{(k,l)}_j(\boldsymbol{r})\right] \\ G^{(l)}_O &= 2\sum_k P^{(k)*}_n(\boldsymbol{r}-\boldsymbol{r}_j)\left[\psi^{(k,l)}_j(\boldsymbol{r}) - \psi'^{(k,l)}_j(\boldsymbol{r})\right] \end{aligned}\right\}. \tag{4.6}$$

So the update functions given in step 4 correspond to the choices of the search step size of

$$\left.\begin{aligned}\sigma_P^{(k)} &= \frac{\alpha}{2\left(\sum_l \left|O_n^{(l)}\left(\boldsymbol{r}+\boldsymbol{r}_j\right)\right|^2\right)_{\max}} \\ \sigma_O^{(k)} &= \frac{\beta}{2\left(\sum_k \left|P_n^{(k)}\left(\boldsymbol{r}-\boldsymbol{r}_j\right)\right|^2\right)_{\max}}\end{aligned}\right\}. \tag{4.7}$$

The parameters α and β can be adjusted to find a balance between the convergence speed and the reconstruction stability. They are normally set to a value between 0 and 1.

4.1.2 The reconstruction structure

Attention should be drawn to step 3 in the above section, where the algorithm only sets constraint on the intensity addition of all the diffraction states, not on the intensity of the individual diffraction state. As a result, it allows the reconstruction of certain linear combination of the underlying object/probe states that still satisfies the modulus constraint. In this section we will show it is this degree of freedom in the algorithm that leads to ambiguous reconstructions.

Here we consider a general case of *K* probe states and *L* object states. A particular set of solutions that satisfy the modulus constraint is given as $\left\{P^{(k)}\left(\boldsymbol{r}\right)\right\}_{k=1}^{K}$ for probe states and $\left\{O^{(l)}\left(\boldsymbol{r}\right)\right\}_{l=1}^{L}$ for object states. Therefore, each probe state will interact with each object state to give a single diffraction state, i.e. $\Psi^{(k,l)}\left(\boldsymbol{q}\right)=\mathcal{F}\left[P^{(k)}(\boldsymbol{r})O^{(l)}(\boldsymbol{r})\right]$, and the intensity summation of all the diffraction states equals to the measured intensity, i.e. $\sum_{k=1}^{K}\sum_{l=1}^{L}\left|\Psi^{(k,l)}\left(\boldsymbol{q}\right)\right|^2 = M(\boldsymbol{q})$. Since the above relation applies to every scan position, the position index *j* is dropped here. It should be noted that the multiplication between the probe state $P^{(k)}(\boldsymbol{r})$ and the object state $O^{(l)}(\boldsymbol{r})$ is pixel-wise. As shown in Fig. 4.1, if we assemble the probe states as a row vector $\boldsymbol{P}=\left[P^{(1)}(\boldsymbol{r}),P^{(2)}(\boldsymbol{r}),\cdots,P^{(K)}(\boldsymbol{r})\right]$, the

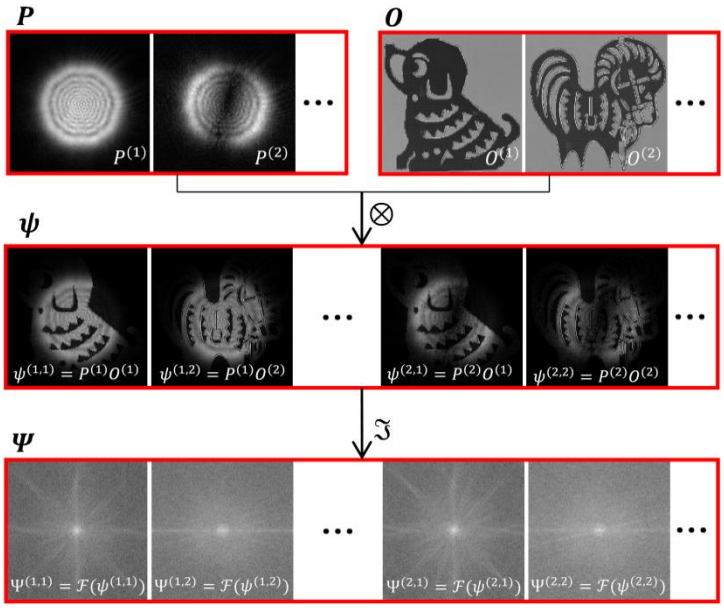

Figure 4.1: Schematic demonstration of Kronecker product (denoted by $\otimes$) and the formation of exit waves (denoted by $\boldsymbol{\psi}$) and diffraction waves (denoted by $\boldsymbol{\Psi}$). $\boldsymbol{P}$ and $\boldsymbol{O}$ denote the illumination states and the object states respectively. $\mathcal{F}$ denotes the Fourier transform. The diffraction patterns are displayed at logarithmic scale.

object states as a row vector $\boldsymbol{O} = \left[O^{(1)}(\boldsymbol{r}), O^{(2)}(\boldsymbol{r}), \cdots, O^{(L)}(\boldsymbol{r}) \right]$, and the diffraction states as a row vector $\boldsymbol{\Psi} = \left[\Psi^{(1,1)}(\boldsymbol{q}), \cdots, \Psi^{(1,L)}(\boldsymbol{q}), \cdots, \Psi^{(K,1)}(\boldsymbol{q}), \cdots, \Psi^{(K,L)}(\boldsymbol{q}) \right]$, then the relation between the three state vectors can be described as a state-wise Kronecker product (denoted by $\otimes$):

$$\boldsymbol{\Psi} = \left[\Im\left(\boldsymbol{P} \otimes \boldsymbol{O} \right) \right], \tag{4.8}$$

where $\Im(*)$ represents the Fourier transform operating on each exit wave state, i.e. $\mathcal{F}\left[P^{(k)}(\boldsymbol{r}) O^{(l)}(\boldsymbol{r}) \right]$. The elements of the three row vectors are defined to be states and the state-wise Kronecker product of the vectors operates on the states, not on the individual pixels within each state. Hereinafter, when the element of a vector is defined to be a state, any calculation on this vector will be considered to operate state-wise whilst multiplications between states will be pixel-wise. Therefore, the intensity summation of all the diffraction states can be expressed as

$$\boldsymbol{\Psi}\boldsymbol{\Psi}^{\dagger} = \sum_{k=1}^{K}\sum_{l=1}^{L}\Psi^{(k,l)}(\boldsymbol{q})\Psi^{(k,l)}(\boldsymbol{q})^{\dagger} = \sum_{k=1}^{K}\sum_{l=1}^{L}\left|\mathcal{F}\left[P^{(k)}(\boldsymbol{r})O^{(l)}(\boldsymbol{r})\right]\right|^2, \tag{4.9}$$

where † denotes the Hermitian conjugate.

Using the same vector representation as above, we denote the set of reconstructed probe states as a row vector $\tilde{\boldsymbol{P}} = \left[\tilde{P}^{(1)}(\boldsymbol{r}), \tilde{P}^{(2)}(\boldsymbol{r}), \cdots, \tilde{P}^{(\tilde{K})}(\boldsymbol{r})\right]$ and the set of reconstructed object states as a row vector $\tilde{\boldsymbol{O}} = \left[\tilde{O}^{(1)}(\boldsymbol{r}), \tilde{O}^{(2)}(\boldsymbol{r}), \cdots, \tilde{O}^{(\tilde{L})}(\boldsymbol{r})\right]$, where $\tilde{K}$ and $\tilde{L}$ are respectively the number of probe states and object states assigned in the algorithm. As previously, the reconstructed diffraction wave states $\tilde{\boldsymbol{\Psi}}$ can then be represented by a state-wise Kronecker product as

$$\tilde{\boldsymbol{\Psi}} = \left[\Im\left(\tilde{\boldsymbol{P}} \otimes \tilde{\boldsymbol{O}}\right)\right]. \tag{4.10}$$

We assume that the reconstructed states ($\tilde{\boldsymbol{P}}$ and $\tilde{\boldsymbol{O}}$) and our particular set of solution states ($\boldsymbol{P}$ and $\boldsymbol{O}$) have the following linear relationship

$$\left.\begin{aligned}\tilde{\boldsymbol{P}} &= \boldsymbol{P}\boldsymbol{U}_P \\ \tilde{\boldsymbol{O}} &= \boldsymbol{O}\boldsymbol{U}_O\end{aligned}\right\}, \tag{4.11}$$

where $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ are matrices with each element being a complex-valued number and the their row-by-column dimensions are $K \times \tilde{K}$ and $L \times \tilde{L}$ respectively. Since the elements of $\tilde{\boldsymbol{P}}$, $\tilde{\boldsymbol{O}}$, $\boldsymbol{P}$ and $\boldsymbol{O}$ are states, Eq. (4.11) can be expressed explicitly as

$$\left.\begin{aligned}\tilde{P}^{(k)}(\boldsymbol{r}) &= \sum_i P^{(i)}(\boldsymbol{r})u_{ik} \\ \tilde{O}^{(l)}(\boldsymbol{r}) &= \sum_j O^{(j)}(\boldsymbol{r})v_{jl}\end{aligned}\right\}, \tag{4.12}$$

where u_{ik} is the element of $\boldsymbol{U}_P$ on row *i* and column *k*, and v_{jl} is the element of $\boldsymbol{U}_O$ on row j and column l. The Fourier transform of each state in $\tilde{\boldsymbol{\Psi}}$ is then calculated as follows:

$$\mathcal{F}\left[\tilde{P}^{(k)}(\boldsymbol{r})\tilde{O}^{(l)}(\boldsymbol{r})\right]=\sum_i\sum_j \mathcal{F}\left[P^{(i)}(\boldsymbol{r})O^{(j)}(\boldsymbol{r})\right]u_{ik}v_{jl}. \tag{4.13}$$

For a set of reconstructed states to be a set of solutions, it has to satisfy the modulus constraint, i.e. $\tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Psi}}^\dagger=\boldsymbol{\Psi}\boldsymbol{\Psi}^\dagger$. The substitution of Eq. (4.10) and (4.11) into $\tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Psi}}^\dagger$ gives

$$\tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Psi}}^\dagger=\left[\Im\left(\tilde{\boldsymbol{P}}\otimes\tilde{\boldsymbol{O}}\right)\right]\left[\Im\left(\tilde{\boldsymbol{P}}\otimes\tilde{\boldsymbol{O}}\right)\right]^\dagger=\left[\Im\left(\boldsymbol{P}\boldsymbol{U}_P\otimes\boldsymbol{O}\boldsymbol{U}_O\right)\right]\left[\Im\left(\boldsymbol{P}\boldsymbol{U}_P\otimes\boldsymbol{O}\boldsymbol{U}_O\right)\right]^\dagger. \tag{4.14}$$

Since the Fourier transform does not operate on the matrices $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$, as we can see in Eq. (4.13), Eq. (4.14) can be rewritten as

$$\tilde{\boldsymbol{\Psi}}\tilde{\boldsymbol{\Psi}}^\dagger=\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]^\dagger\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]^\dagger. \tag{4.15}$$

The modulus constraint requires

$$\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]^\dagger\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]^\dagger=\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]\left[\Im\left(\boldsymbol{P}\otimes\boldsymbol{O}\right)\right]^\dagger, \tag{4.16}$$

which leads to the following relation

$$\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]\left[\boldsymbol{U}_P\otimes\boldsymbol{U}_O\right]^\dagger=\boldsymbol{U}_P\boldsymbol{U}_P{}^\dagger\otimes\boldsymbol{U}_O\boldsymbol{U}_O{}^\dagger=\boldsymbol{I}_{KL}, \tag{4.17}$$

where $\boldsymbol{I}$ denotes the identity matrix with its dimension indicated by the subscript. This means that the matrices $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ have to satisfy

$$\left.\begin{aligned}\boldsymbol{U}_P\boldsymbol{U}_P{}^\dagger&=c\boldsymbol{I}_K\\ \boldsymbol{U}_O\boldsymbol{U}_O{}^\dagger&=\boldsymbol{I}_L/c\end{aligned}\right\}, \tag{4.18}$$

where *c* represents some arbitrary positive real number. $\boldsymbol{I}_K$ has full rank, so the rank of $\boldsymbol{U}_P\boldsymbol{U}_P{}^\dagger$ is *K*. According to the properties of the matrix rank, $K\le\mathrm{rank}(\boldsymbol{U}_P)$ and

$\text{rank}(\boldsymbol{U}_P) \le \min(K, \tilde{K})$, which leads to $\text{rank}(\boldsymbol{U}_P) = K$ and $K \le \tilde{K}$. For matrix $\boldsymbol{U}_O$, we have $\text{rank}(\boldsymbol{U}_O) = L$ and $L \le \tilde{L}$. This means as long as we assign a number no smaller than the number of the underlying object/probe states to the number of the reconstruction states in the algorithm, the modulus constraint can be fulfilled but each reconstruction state will be an arbitrary linear combination of the underlying states, i.e. ambiguous reconstructions. Since the magnitude of the diffraction wave is only sensitive to the relative phase not the absolute phase, the constant phase ambiguity will still exist in each reconstruction state. Furthermore, the substitution of Eq. (4.11) into Eq. (4.18) leads to

$$\left.\begin{aligned} \tilde{\boldsymbol{P}}\tilde{\boldsymbol{P}}^{\dagger} &= c\boldsymbol{P}\boldsymbol{P}^{\dagger} \\ \tilde{\boldsymbol{O}}\tilde{\boldsymbol{O}}^{\dagger} &= \boldsymbol{O}\boldsymbol{O}^{\dagger} / c \end{aligned}\right\}. \tag{4.19}$$

This means that although the reconstruction ambiguity happens for each individual state, the normalised intensity addition of all the reconstruction states is equal to that of the underlying states for both the probe and the object.

4.1.3 Breaking the reconstruction ambiguities

The reconstruction ambiguity basically originates from the modulus constraint where the only condition is on the intensity addition of all the diffraction waves. Consequently, the reconstruction recomposes the original probes and objects linearly in a manner determined by the initial probe and object guesses. The ambiguity is the inherent property of the reconstruction algorithm. To break it, we have to make use of the known information about the experiment. The known information can be formed as extra constraints added into the algorithm. Hopefully, these constraints can help to eliminate the ambiguity, producing unique true reconstructions. Here two such constraints are discussed: phase-only and orthogonalisation.

Phase-only constraint means the amplitude part is known to be flat for a complex signal. It is very common for the objects in high-resolution x-ray [119] and electron imaging [120]. Here we use a simple example of two phase object states as the demonstration. The two states are $O^{(1)}(\boldsymbol{r}) = \mathrm{A}^{(1)}(\boldsymbol{r})e^{i\Phi^{(1)}(\boldsymbol{r})}$ and $O^{(2)}(\boldsymbol{r}) = \mathrm{A}^{(2)}(\boldsymbol{r})e^{i\Phi^{(2)}(\boldsymbol{r})}$, where $\mathrm{A}^{(1)}(\boldsymbol{r})$, $\mathrm{A}^{(2)}(\boldsymbol{r})$ are the flat amplitudes and $\Phi^{(1)}(\boldsymbol{r})$, $\Phi^{(2)}(\boldsymbol{r})$ are the phases. If the matrix $\boldsymbol{U}_O = [u_{11}, u_{12}; u_{21}, u_{22}]$, according to Eq. (4.11) the two reconstruction states will have the following relation

$$\left.\begin{aligned}\tilde{O}^{(1)}(\boldsymbol{r}) &= u_{11}O^{(1)}(\boldsymbol{r}) + u_{21}O^{(2)}(\boldsymbol{r})\\ \tilde{O}^{(2)}(\boldsymbol{r}) &= u_{12}O^{(1)}(\boldsymbol{r}) + u_{22}O^{(2)}(\boldsymbol{r})\end{aligned}\right\}. \tag{4.20}$$

The amplitudes of the two reconstruction states are

$$\left.\begin{aligned}&\left|O^{(1)}(\boldsymbol{r})\right| = \\ &\sqrt{\left|u_{11}\mathrm{A}^{(1)}(\boldsymbol{r})\right|^2 + \left|u_{21}\mathrm{A}^{(2)}(\boldsymbol{r})\right|^2 + 2u_{11}\mathrm{A}^{(1)}(\boldsymbol{r})u_{21}\mathrm{A}^{(2)}(\boldsymbol{r})\cos\left[\Phi^{(1)}(\boldsymbol{r}) - \Phi^{(2)}(\boldsymbol{r})\right]}\\ &\left|O^{(2)}(\boldsymbol{r})\right| = \\ &\sqrt{\left|u_{12}\mathrm{A}^{(1)}(\boldsymbol{r})\right|^2 + \left|u_{22}\mathrm{A}^{(2)}(\boldsymbol{r})\right|^2 + 2u_{12}\mathrm{A}^{(1)}(\boldsymbol{r})u_{22}\mathrm{A}^{(2)}(\boldsymbol{r})\cos\left[\Phi^{(1)}(\boldsymbol{r}) - \Phi^{(2)}(\boldsymbol{r})\right]}\end{aligned}\right\}. \tag{4.21}$$

Since the amplitudes $\mathrm{A}^{(1)}(\boldsymbol{r})$ and $\mathrm{A}^{(2)}(\boldsymbol{r})$ are flat, the cross-terms in Eq. (4.21) must be zero, which gives $u_{11}u_{21} = 0$ and $u_{12}u_{22} = 0$. Besides, the matrix $\boldsymbol{U}_O$ satisfying Eq. (4.18) requires $|u_{11}|^2 + |u_{12}|^2 = 1$ and $|u_{21}|^2 + |u_{22}|^2 = 1$. The scaling constant c is omitted for simplicity and its absence will not affect the results. So the phase-only constraint leads to $\boldsymbol{U}_O = [1, 0; 0, 1]$ or $\boldsymbol{U}_O = [0, 1; 1, 0]$. Both the solutions are unambiguous, only with the ordering of the two states swapped. Although an example of only two states is demonstrated here, it can be readily extended to cases with more than two states. However, this constraint is unlikely applicable to the probe states for a real situation.

We have to resort to other constraints to break the reconstruction ambiguity of the probe states.

Orthogonalisation constraint means the states are known to be mutually orthogonal (i.e. the dot product of two states will be zero). The spatial partial coherence is a well-known example for this constraint [98]. As we stated in section 2.3.2, a spatial partial coherent field can be decomposed into a set of mutually orthogonal modes. Each mode is totally coherent itself but completely incoherent with the others. It is a common constraint to break the ambiguity in the probe reconstructions and gives a unique set of probe modes [64].

Suppose the solution probe states are mutually orthogonal, which means their dot product is zero. In order to formularise the orthogonality, we write each state $P^{(k)}(\boldsymbol{r})$ as a column vector $\boldsymbol{s}^{(k)} = \left[p^{(k)}(\boldsymbol{r}_1), p^{(k)}(\boldsymbol{r}_2), \cdots, p^{(k)}(\boldsymbol{r}_M) \right]^T$, where *T* represents the transpose operation, $\boldsymbol{r}_i$ denotes the *i*th pixel and *M* is the number of pixels in the state. So the mutual orthogonality means

$$\boldsymbol{s}^{(k)\dagger}\boldsymbol{s}^{(l)} = \begin{cases} 0, & k \neq l \\ \eta, & k = l \end{cases}, \tag{4.22}$$

where η is the power of the mode and it is a positive real number. Assembling this set of probe states as a matrix $\boldsymbol{S}$ (to differentiate from the notation $\boldsymbol{P}$ used in section 4.1.2, where it is only a state vector, not a matrix, but the relation described in Eq. (4.11) still holds here) such that

$$\boldsymbol{S} = \left[\boldsymbol{s}^{(1)}, \boldsymbol{s}^{(2)}, \cdots, \boldsymbol{s}^{(M)} \right] = \begin{bmatrix} p^{(1)}(\boldsymbol{r}_1) & p^{(2)}(\boldsymbol{r}_1) & \cdots & p^{(K)}(\boldsymbol{r}_1) \\ p^{(1)}(\boldsymbol{r}_2) & p^{(2)}(\boldsymbol{r}_2) & \cdots & p^{(K)}(\boldsymbol{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ p^{(1)}(\boldsymbol{r}_M) & p^{(2)}(\boldsymbol{r}_M) & \cdots & p^{(K)}(\boldsymbol{r}_M) \end{bmatrix}, \tag{4.23}$$

so now we can form the matrix $\boldsymbol{D}$ containing all the dot products of the probe states as

$$\boldsymbol{D} = \boldsymbol{S}^{\dagger}\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}^{(1)\dagger}\boldsymbol{s}^{(1)} & \boldsymbol{s}^{(1)\dagger}\boldsymbol{s}^{(2)} & \cdots & \boldsymbol{s}^{(1)\dagger}\boldsymbol{s}^{(K)} \\ \boldsymbol{s}^{(2)\dagger}\boldsymbol{s}^{(1)} & \boldsymbol{s}^{(2)\dagger}\boldsymbol{s}^{(2)} & \cdots & \boldsymbol{s}^{(2)\dagger}\boldsymbol{s}^{(K)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{s}^{(K)\dagger}\boldsymbol{s}^{(1)} & \boldsymbol{s}^{(K)\dagger}\boldsymbol{s}^{(2)} & \cdots & \boldsymbol{s}^{(K)\dagger}\boldsymbol{s}^{(K)} \end{bmatrix} = \begin{bmatrix} \eta_1 & 0 & \cdots & 0 \\ 0 & \eta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_K \end{bmatrix}. \quad (4.24)$$

And this matrix $\boldsymbol{D}$ is a diagonal matrix. Making use of the relation described in Eq. (4.11), we can form a matrix $\tilde{\boldsymbol{D}}$ for the set of reconstruction states as

$$\tilde{\boldsymbol{D}} = \tilde{\boldsymbol{S}}^{\dagger}\tilde{\boldsymbol{S}} = \boldsymbol{U}_P{}^{\dagger}\boldsymbol{D}\boldsymbol{U}_P. \quad (4.25)$$

Eq. (4.25) gives us the form of the Eigen decomposition of the matrix $\tilde{\boldsymbol{D}}$. The diagonal elements in matrix $\boldsymbol{D}$ are the eigen-values and the columns in $\boldsymbol{U}_P{}^{\dagger}$ are the corresponding eigen-vectors. So through the Eigen decomposition of the matrix $\tilde{\boldsymbol{D}}$, we can obtain the matrix $\boldsymbol{U}_P{}^{\dagger}$. Utilising the relation in Eq. (4.11), we then can calculate the orthogonal states from the reconstructed ones via

$$\boldsymbol{S} = \tilde{\boldsymbol{S}}\boldsymbol{U}_P{}^{\dagger}. \quad (4.26)$$

It is worth mentioning that the orthogonalisation of the probe states described here is equivalent to the diagonalisation of the density matrix of the coherence function. According to the definition, the density matrix is given by $\tilde{\boldsymbol{S}}\tilde{\boldsymbol{S}}^{\dagger}$. Using the relation in Eq. (4.11) and Eq. (4.18) we have

$$\tilde{\boldsymbol{S}}\tilde{\boldsymbol{S}}^{\dagger} = \boldsymbol{S}\boldsymbol{U}_P\boldsymbol{U}_P{}^{\dagger}\boldsymbol{S}^{\dagger} = \boldsymbol{S}_N\boldsymbol{D}\boldsymbol{S}_N^{\dagger}, \quad (4.27)$$

with

$$\boldsymbol{S}_N = \left[\boldsymbol{s}^{(1)}/\sqrt{\eta_1}, \boldsymbol{s}^{(2)}/\sqrt{\eta_2}, \cdots, \boldsymbol{s}^{(K)}/\sqrt{\eta_K}\right]. \quad (4.28)$$

So the eigenvectors of the density matrix are the original illumination states normalised and the eigenvalues are the power of the corresponding states.

4.1.4 Optical experiments

Optical experiments are carried out to demonstrate the reconstruction ambiguities. We used a diode laser (with wavelength of 675nm) coupled with a step-index multi-mode optical fibre (10μm core diameter, 0.1 NA) as the source. Since the exit wave from a multi-mode fibre is spatially partial coherent [121], this will produce an illumination with multiple states. In order to generate multiple object states, we use a phase-only spatial light modulator (SLM, a Holoeye Pluto reflective liquid crystal device) to display two patterns (a dog and a horse) alternatively at a rate of 20Hz. The optical setup is shown in Fig. 4.2. It is a reflection configuration for ptychography. A linear polarizer is inserted into the beam to make sure the polarization of the illumination matches the polarization of the SLM. Two standard doublet lenses (f1=100mm and f2=45mm) are arranged in a 4f configuration to focus the illumination. An aperture of 4mm diameter is used to limit the size and numerical aperture of the probe. The SLM is positioned at a slight distance after the focus, where the beam diameter is approximately 900μm. A beam-splitter is inserted between the second lens and the SLM to allow the diffraction patterns reflected from the SLM to be recorded on a CCD detector (an AVT Pike F421B, with 2048×2048 pixels, each of 7.4μm^2. However, for reconstruction the diffraction patterns are binned by a factor of 4 to 512×512). The SLM is mounted on an *x-y* translation stage to implement the ptychographic scan.

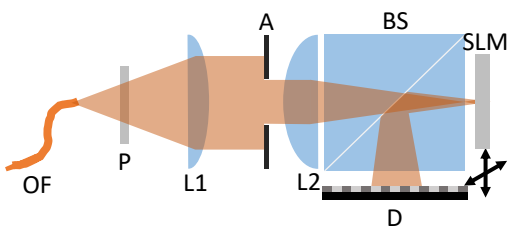

Figure 4.2: The experimental setup of ptychography operating in the reflection mode. OF stands for the optical fibre, P for polarizer, L for the lens, A for aperture, BS for the beam splitter, D for the detector and SLM for the spatial light modulator. The focal length of L1 is 100mm and that of L2 is 45mm.

To obtain the reference images for the probe states and the two object states, we perform two calibration experiments first, one with a static dog pattern displayed on the SLM and the other with a static horse pattern displayed. The illumination is scanned over the patterns on a grid of 13×13. At each scan, two exposures (0.25s and 1.0s) are taken and combined to increase the dynamic range of the detector. The step size is about 100μm with ±20% random offsets.

As explained above, the number of reconstruction states should be no smaller than that of underlying states in the experiment to ensure the convergence of the reconstruction (i.e. fulfil the modulus constraint). However, for this experiment the number of underlying probe states is unknown. Therefore, we try a set of reconstructions with an increasing number of probe states from 1 to 7 in the algorithm. The reconstruction began with 100 iterations of the ePIE algorithm, assuming a single probe state. The results were then used as the starting point for 900 iterations of the modified ePIE algorithm described in section 4.1.1, when we assign a number bigger than 2 to the probe states. The orthogonalised probe modes from different reconstructions are shown in Fig. 4.3. We note the following: the first mode, containing most of the power in the illumination, looks similar for all the reconstructions; the modes with lower power evolve as the number of states assigned in the reconstruction is increased, until five modes is included, after which the first five modes remain more or less constant; modes 6 and 7 varied randomly and contain very little power (totalling less than 3% of the incident beam). We conclude that the latter low power, randomly shaped modes account for imperfections in the data sets, such as the background noise in the detector [122],and their inclusion therefore helps ‘clean’ the image quality in the dominant object and probe states. As a result, in subsequent reconstructions we assign six to the number of

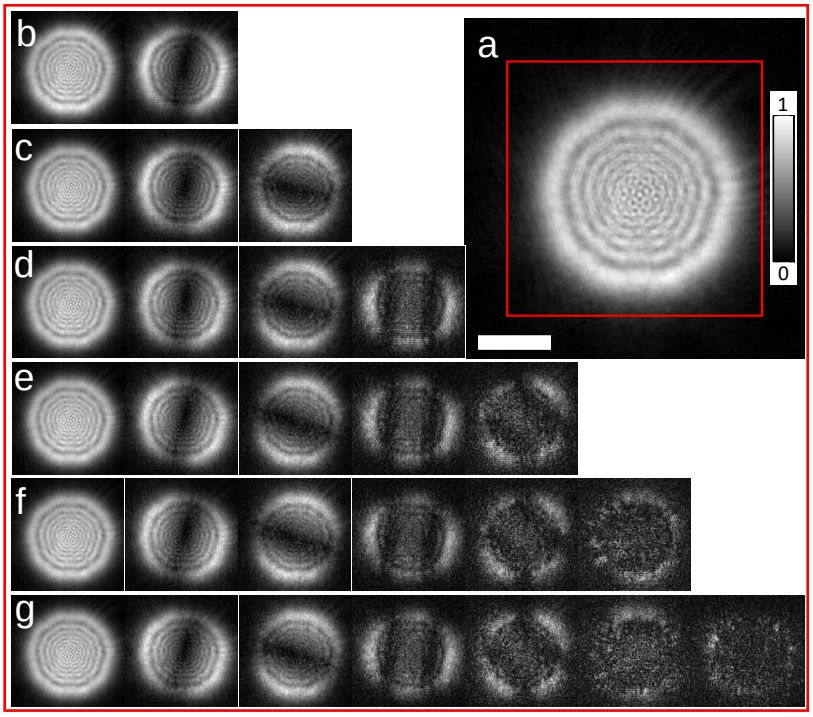

Figure 4.3: The modulus parts of the probe modes reconstructions after orthogonalisation with different numbers of modes assigned to the reconstruction algorithm. From (a) to (g), the numbers are accordingly from 1 to 7. The scale bar indicates a length of 300μm.

the probe states, five to model the beam emitted from the optical fibre and a sixth to manage inconsistency in the recorded diffraction data.

The five illumination modes and the two object states from the calibration experiments are shown in Fig. 4.4, where the colour wheel depiction shows phase variations as changes in hue and modulus variations as changes in brightness. Because of the strong phase curvature, the structures of the probe modes cannot be observed clearly, so the phase of the first mode is used as a reference to remove all the phase curvatures, as shown in Fig. 4.4d. (For display purposes we will use this same reference to remove the phase curvature in the probe reconstructions that follow.) Although the SLM is a phase modulator, our images do show some amplitude modulation around the edges of the dog and horse patterns, caused by scattering beyond the extent of the detector from these regions; close inspection also reveals the regular 8μm spacing of the SLM pixels.

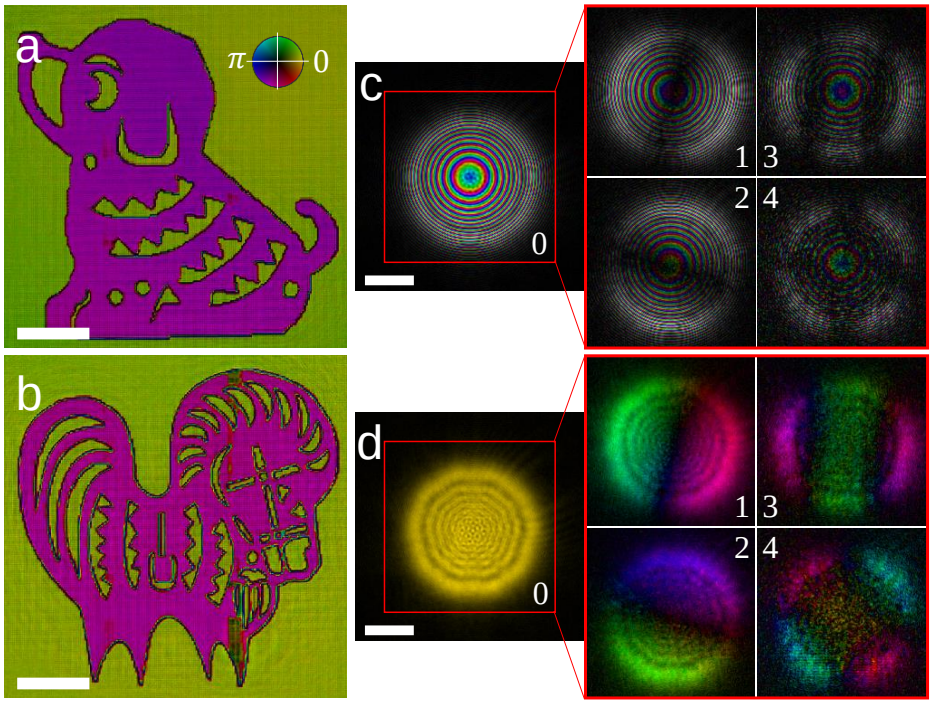

Figure 4.4: The reconstructions from the calibration experiments. (a) The dog pattern reconstruction. (b) The horse pattern reconstruction. (c) The illumination modes reconstructions. (d) The illumination states reconstructions with the phase curvature removed. The scale bars indicate a length of 300μm. The colour wheel applies to all the images.

Using the same experimental parameters as above, we repeat our experiment with the SLM programmed to alternately display the dog and horse images on a 20Hz cycle, ensuring that both patterns are displayed several times during the 0.25s and 1.0s exposures. Again, we run the ePIE algorithm for the first 100 iterations. Then the results are fed into the modified ePIE algorithm, running for 900 iterations. With this dataset, two sets of reconstructions are performed with different initial guesses and the reconstructions are shown in Fig. 4.5. Obviously, the two reconstructions are different for both the probe states and the object states and also different from the calibrated results, demonstrating the ambiguities.

According to Eq. (4.19) the normalised intensity addition of both reconstructed probe states and reconstructed object states should be unique despite the ambiguity of the individual state. To demonstrate this, we calculate the normalised intensity additions for both the two calibrated object states (the main figure in Fig. 4.6a) and the five calibrated

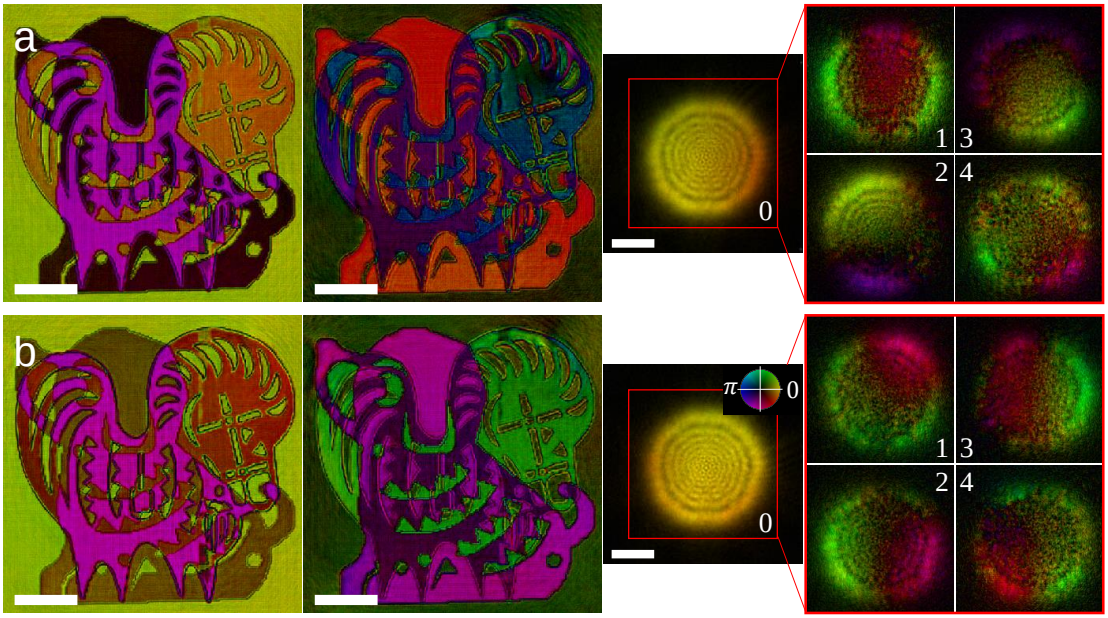

Figure 4.5: Two different sets of reconstructions from the same dataset using different initial guesses. (a) The first set of raw reconstructions. (b) The second set of raw reconstructions. The scale bars indicate a length of 300μm. The colour wheel applies to all the images. To aid in display, the curvature of the probe phases has been removed.

illumination modes (the inset in Fig. 4.6a). The same calculations were done for the two sets of reconstructions shown in Fig. 4.5. It is interesting to notice that except the edges of the patterns the intensity additions of the two object states can be assumed flat, which is consistent with the fact that the SLM only had phase response in the experiment. These edges are due to the high frequencies scattered off the detector. From visualisation, there is no major difference among these intensity additions. They can

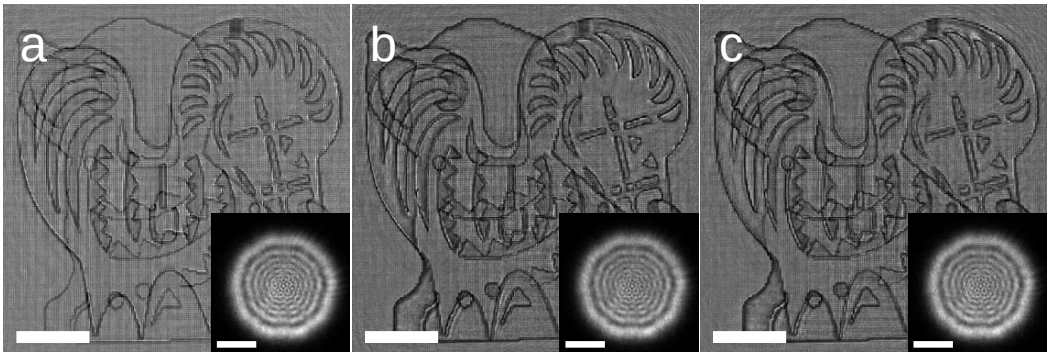

Figure 4.6: Comparison of the normalised intensity additions of the states. (a) The intensity additions of the calibrated probe modes and the calibrated object states. (b) The intensity additions of the probe states and the object states from the first reconstruction. (c) The intensity additions of the probe states and the object states from the second reconstruction. The scale bars indicate a length of 300μm.

also be quantitatively compared using the complex image error metric (see section 2.3.9). Between the calibrations and the first reconstructions, i.e. between Figs. 4.6a and 4.6b, the error for the normalised intensity addition of the objects is 0.214 and the error for that of the probes is 0.0767. Between the calibrations and the second reconstructions, i.e. between Figs 4.6a and 4.6c, the error for the normalised intensity addition of the objects is 0.210 and the error for that of the probes is 0.0721. The errors are higher for the object case, due mainly to the discrepancy from the edges in the patterns. These errors can be assumed small considering the presence of imperfections in the experiment, such as statistic noise, background noise, readout noise of detector and the signals scattering beyond the detector.

Since the SLM we used in the experiment is a phase modulator (although the border of each SLM pixel does affect the amplitude of the reflected beam to some extent), an assumption of flat amplitude for the object states can be used as *a priori* knowledge in the reconstruction. Forcing the amplitudes of the two object states to be flat during the reconstruction produces the results as shown in Fig. 4.7a. Evidently, the phase-only constraint has successfully broken the ambiguity, although there is some cross-talk between the two states because the pixelation of the SLM does somewhat modulate the object state amplitudes. To mitigate this problem, the results of Fig. 4.7a are used as accurate initial guesses to feed into a second run of the algorithm, this time without the phase-only constraint, with the results shown in Fig. 4.7b. Having broken the object state ambiguity and generated images very close to the calibrated reference states, this additional step largely eliminates background artefacts, whilst the slight cross-talk between the two states still exists because of its presence in the initial guesses.

The slight inaccuracy in the assumption of flat amplitude for the SLM experiments leads us to simulate an experiment that demonstrates that the phase-only constraint

Figure 4.7: Reconstructions of the two object states with applying the phase-only constraint. (a) The two object states reconstructions with the phase-only constraint through the whole reconstruction process. (b) The two object states reconstructions using (a) as the initial guesses respectively without the phase-only constraint. The scale bars indicate a length of 300μm. The scale bar and the colour wheel apply to all the images.

works perfectly when the amplitudes are ideally flat. Moreover, to demonstrate this constraint can be applied successfully to more than two object states, in this simulation four object states are used. The phases, shown in Fig. 4.8a, are set to a range between $-\pi/2$ to $\pi/2$ and the amplitudes set to 1. The probe mode with the most power in the reconstruction, as shown in Fig. 4.4c, is used as the single probe state in this simulated experiment, and the scan positions from the optical experimental are also reused, producing 169 diffraction patterns from 13×13 overlapping regions. First the reconstruction is carried out without the phase-only constraint – the phase parts of the results are shown in Fig. 4.8b. Ambiguous reconstructions with different combinations of the original four object states are reconstructed. Then, imposing the phase-only constraint another reconstruction is carried out, with the results shown in Fig. 4.8c; the four object states are successfully reconstructed without any artefacts or cross-talk – the only ambiguity remaining is the order in which the states appear in the reconstruction.

Figure 4.8: The simulation and the reconstructions. (a) The phase parts of the four object states. The amplitudes are all set to 1. (b) The phase parts of the four states reconstructions without applying the phase only constraint. (c) The phase parts of the four states reconstructions with applying the phase only constraint.

Some other constraints have also been tested to try to break the object ambiguity. Among these, the real-valued constraint (the object states only have real part), positivity constraint (both the real part and imaginary part are both positive) and support constraint (the object states have a limited extent) fail to break the ambiguity. Another constraint that allows small fluctuation in the amplitudes (the fluctuation range we tried is from 0.9 to 1.0) can successfully break the ambiguity.

Since the illumination exiting the step-index multi-mode optical fibre is spatial-partially coherent, here we can orthogonalise the probe states reconstructions to obtain the orthogonal probe modes. The orthogonal modes from the two ambiguous reconstructions are shown in Fig. 4.9. The structures of the modes are consistent across the two experiments and agree with the calibrated modes, apart from arbitrary constant phase

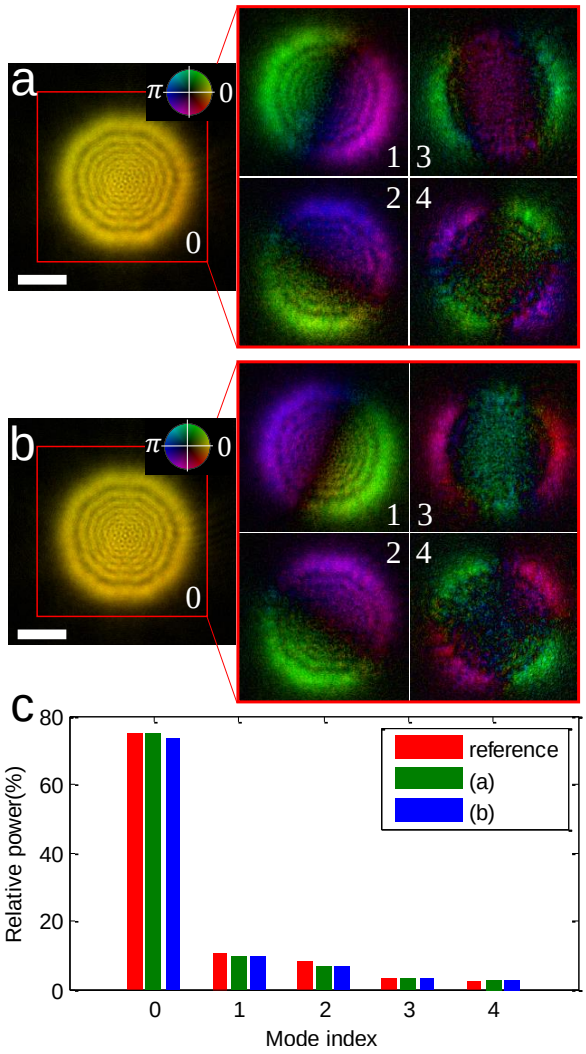

Figure 4.9: Comparison of the illumination modes from the two different sets of reconstructions after orthogonalisation. (a) Orthogonal probe modes from the first set of reconstructions. (b) Orthogonal probe modes from the second set of reconstructions. (c) The power distribution of the original probe modes and the two sets of reconstructed probe modes. The scale bars indicate a length of 300µm.

offsets. A quantitative comparison of each mode, using the complex image error metric (see section 2.3.9), gives errors for the first set of modes, compared to the calibrated modes, as respectively 0.0428, 0.0282, 0.0245, 0.0219 and 0.0224; the errors for the second set of modes are respectively 0.0386, 0.0292, 0.0253, 0.0183 and 0.0197. These errors are normalised to the total power of the calibrated modes and can be assumed very small in the presence of noise in the data set, demonstrating that orthogonalisation breaks the probe reconstruction ambiguity. It also implies that, as expected, the two sets

of reconstructed probe states are different linear combinations of the original probe modes.

It should be noted that the orthogonalisation can also lead to a unique set of object reconstructions. However, unlike the orthogonal illumination modes for partial coherent illumination, the orthogonalisation of the object states does not give interpretable results in this case, because the two images used in the experiment were not themselves orthogonal.

4.2 The effects of a diffuser in the presence of spatial partial coherence

The photon flux available in the experiments is very important for x-ray diffractive imaging. The increase of flux could shorten the exposure time and hence the whole experimental duration, reducing possible instabilities in the experiment like drift and radiation damage. It also increases the signal-to-noise ratio for the measurements. For these reasons, to make full use of the x-ray coherence, a lens setup should be adopted, because pinhole setup cuts out useful flux. For example, a synchrotron x-ray source typically has a horizontal coherence width no smaller than 50μm (the vertical coherence width is normally bigger than the horizontal width by several times). However, the common size for an illumination function is around 30μm or less to make sure the sampling condition in the detector plane is well satisfied. In this case, the pinhole will cut off almost 50% of the useful counts. Moreover, partial coherence can be solved via mixed state ptychography. This means we are able to increase the flux by intentionally introducing small amount of partial coherence into the experiment. In this section we explore the effects of a diffuser on the reconstructions for a lens setup x-ray experiment in the presence of spatial partial coherence (the content of this section may not fit in with this chapter quite well, but because of the spatial partial coherent condition, we

place it here). The work presented in this section has led to a journal paper publication [123].

4.2.1 The benefits of a diffuser

In section 2.3.5, we have discussed the benefits of the lens setup against the pinhole setup. The lens setup defeats the pinhole setup by introducing a much larger range of incident angles in the illumination, so it could reduce the dynamic range in the diffraction pattern and encode the scattering information of the specimen into more pixels. All these benefits can be enhanced by having a diffuser inserted into the beam, because the angular range of the illumination is further broadened. Therefore, a diffuser can help to make full use of the limited dynamic range of a detector, increase signal-to-noise ratio for the signals scattered to high angles, and solve more unmeasured pixels much more reliably.

An example of recovering unmeasured pixels is to extrapolate a detector of limited size to gain resolution improvement [65, 66]. The ultimate size that the detector can be extended to is, in theory, equal to the width of the detector plus the angular range of the illumination. The stronger a diffuser is, the broader the angular range will be produced and hence the bigger the ultimate size is. When the scattering strength of the specimen exceeds the collection angle of the detector, this detector extrapolation strategy can help the gain of the resolution of the object reconstruction [65, 66].

It has also been demonstrated that a diffuser is very helpful in improving the depth resolution in 3D visible light ptychography via the multislice method [67, 68]. As it will be explained in detail in chapter 5, the depth resolution is related to both the scattering strength of the specimen and the angular range of the illumination. Thus, increasing either of them provides a gain in the depth resolution.

For the Wigner distribution deconvolution method (see chapter 3), we have demonstrated that a diffused probe is better than a normal defocused probe in terms of avoiding division by small values during the deconvolution process in the presence of noise. For iterative algorithms, it has been shown that tailoring the beam using a sharp aperture to broaden the angular range of the illumination can improve the reconstruction in the context of x-ray coherent ptychography [124]. Burdet et al [125] have reached the same conclusion using model calculations in the presence of partial coherence, a result we will explore experimentally in this section.

Making use of the properties of the diffraction pattern speckle from a random surface, we can estimate the degree of the partial coherence in real time by using a diffuser. We will also demonstrate this advantage of using a diffuser in this section.

4.2.2 X-ray experiments

We implement the x-ray experiments on the I13 beamline at the Diamond Light Source in UK. The monochromated x-ray energy is selected at 9.7keV. In the horizontal direction, due to a larger size of the source, the beamline has lower coherence compared to the vertical direction. For the experiment, we intentionally open up the upfront horizontal slits to introduce partial coherence. A Fresnel zone plate (FZP) with a focal length of 470mm is inserted downstream of the slits to condense the beam. A central-stop (CS) is used to block the transmitted beam and an order-select-aperture (OSA) is used to choose the first order of the diffracted beam as the probe for the experiment. The sample is a Siemens star with 500 nm thick gold structures and it is mounted about 16mm upstream of the FZP focus. A piece of paint is used as the diffuser and is inserted about 1mm upstream of the Siemens star. The experimental setup is shown in Fig. 4.10. The photon-counting ‘Merlin’ detector is composed of 2×2 arrays of MediPix3 chips

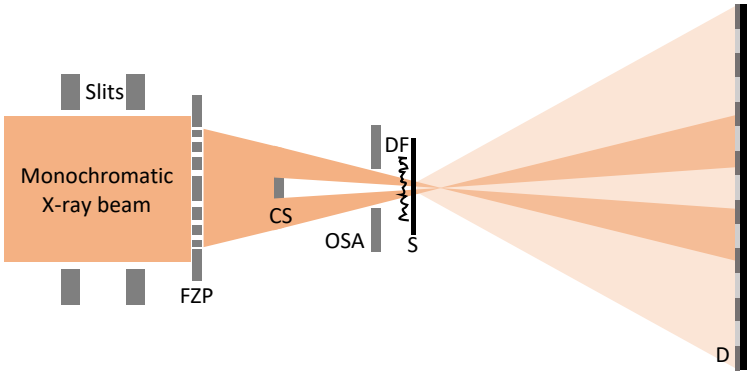

Figure 4.10: The experimental configuration. FZP stands for the Fresnel Zone Plate, CS for the Central Stop, OSA for Order Selected Aperture, S for the specimen and D for the detector.

(each with 256×256 pixels of 55μm^2), which is positioned 14.33m downstream of the sample.

Two experiments are carried out using the above setup with the diffuser ('D_{in}' configuration) and without the diffuser ('D_{out}' configuration). The centre of the Siemens star is scanned over a grid of 30×30. The step size is 1um plus ±10% random offsets. For experimental reasons that do not affect the present work, each diffraction pattern is taken with 256 exposures of 0.032 seconds (i.e. total exposure is 8.192 seconds). For each experiment, the dataset takes about 5 hours to collect. Unfortunately, the probe is not consistent during this long scan (which we discovered when we process the data). It is probably caused by the drift of the optics or the radiation damage of the diffuser. Because of this, we select a subset scan of 12×24 positions that has negligible probe variation to assess the performance of the diffuser. We divide the 256 exposures into two 128 exposures to implement the Fourier ring correlation comparison. An example of the diffraction patterns from each experiment at the same scan position is shown in Fig. 4.11. The two perpendicular dark lines in the centre are the gaps between the 4 detector chips. For the diffuser configuration, the diffraction pattern extends well beyond the unscattered beam and fills up a larger proportion of the detector. Judging from this, the paint acts as a good diffuser.

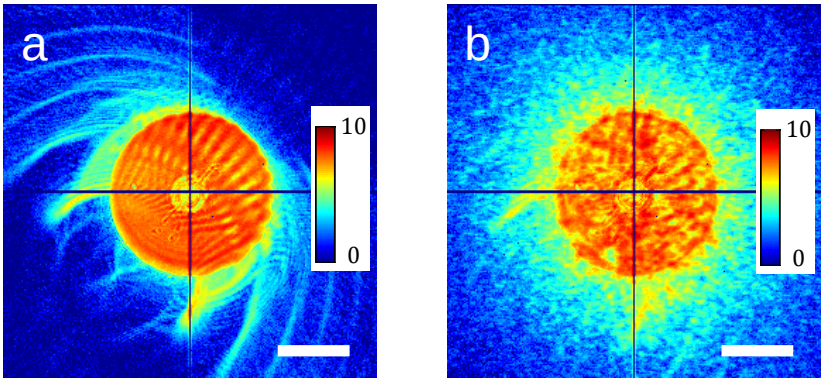

Figure 4.11: An example of the recorded diffraction patterns at logarithmic scale from (a) without diffuser and (b) with diffuser at the same scan position. With the diffuser in, the diffraction pattern extends well beyond the unscattered beam and fills up a larger proportion of the detector. The scale bars represent 3 μm^{-1}.

100 iterations of the normal ePIE algorithm are run first on the two datasets without accounting for the partial coherence. The results are then served as the initial guesses for the modified ePIE algorithm. 300 iterations are run and 12 probe modes are used in the algorithm. Fig. 4.12 shows the reconstruction of eight dominant probe modes after orthogonalisation from the D_{out} (Fig. 4.12a) and D_{in} (Fig. 4.12b) configurations. Except random perturbation exists for the D_{in} configuration, the two sets of probe modes have the same structures, which is that more lobes present in the horizontal direction. The results indicate that the experiments have a shorter coherence length along the horizontal direction, which is consistent with the fact that the synchrotron source has a much bigger horizontal width than its vertical width.

The corresponding specimen reconstructions are shown in Fig. 4.13. Both two reconstructions are fairly good considering the low degree of coherence in the datasets. With inspection of the zoom in figures, we can see that the reconstruction from the D_{out} configuration has ringing effects and blurred structures, whereas the reconstruction from the D_{in} configuration has cleaner background and sharper details, which is particularly obvious in the central disc of the Siemens star in the top right corner of Fig. 4.13(a2) and 4.13(b2).

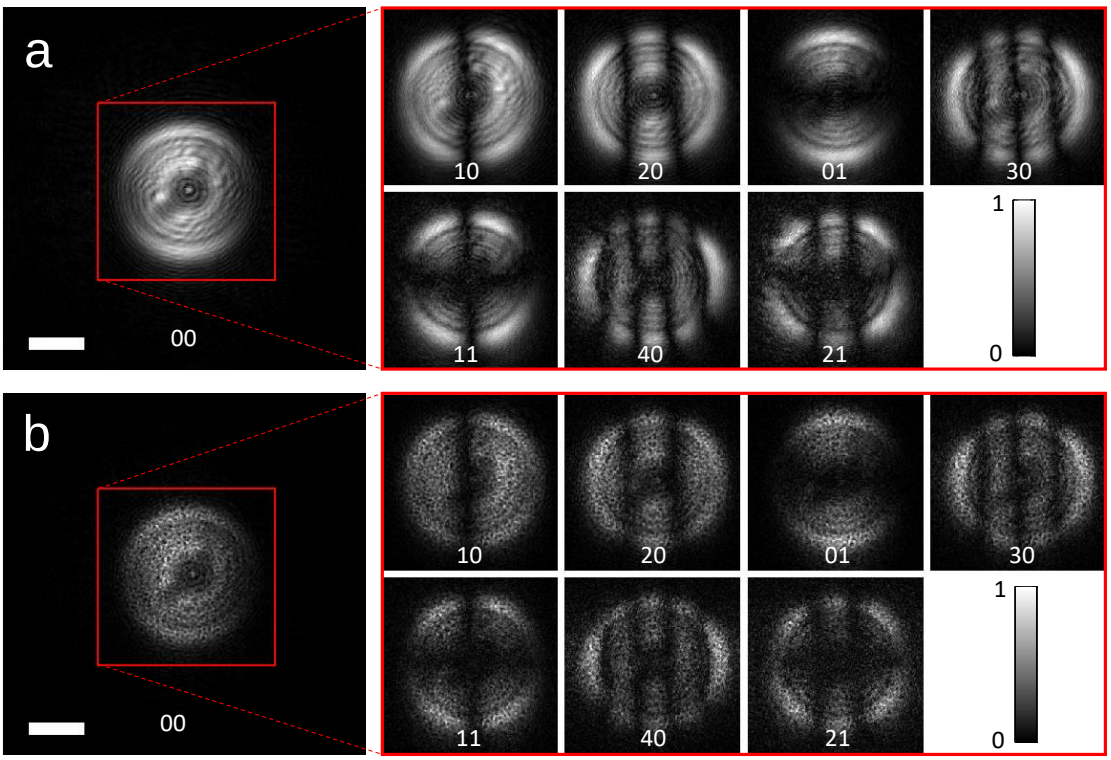

Figure 4.12: The moduli of the probe modes reconstructions using the data from (a) D_{out} configuration and (b) D_{in} configuration. More lobes present in horizontal direction in the probe modes indicate that the horizontal size of the synchrotron source is bigger than its vertical size. The scale bars represent 5μm.

Since we have two sets of datasets with the same exposures for each experiment (256 exposures divided into two 128 exposures as we mentioned above), we can calculate the Fourier ring correlation to compare the reconstructions. The correlation curves are shown in Fig. 4.14 and they also indicate that the diffused probe produces better reconstructions. However, the overall improvement in the reconstruction is rather marginal.

The reason is probably because that the specimen is complicated (structure-wise) enough so that the complexity of the probe does not matter that much. To investigate this, we should use a simple object (for example having large areas of flat contrast and sparse features). Therefore, another pair of experiments with both D_{in} and D_{out} configurations is carried out at the edge of the Siemens star where the features are much sparser. The sample this time is scanned over a grid of 16×16 and only one exposure of

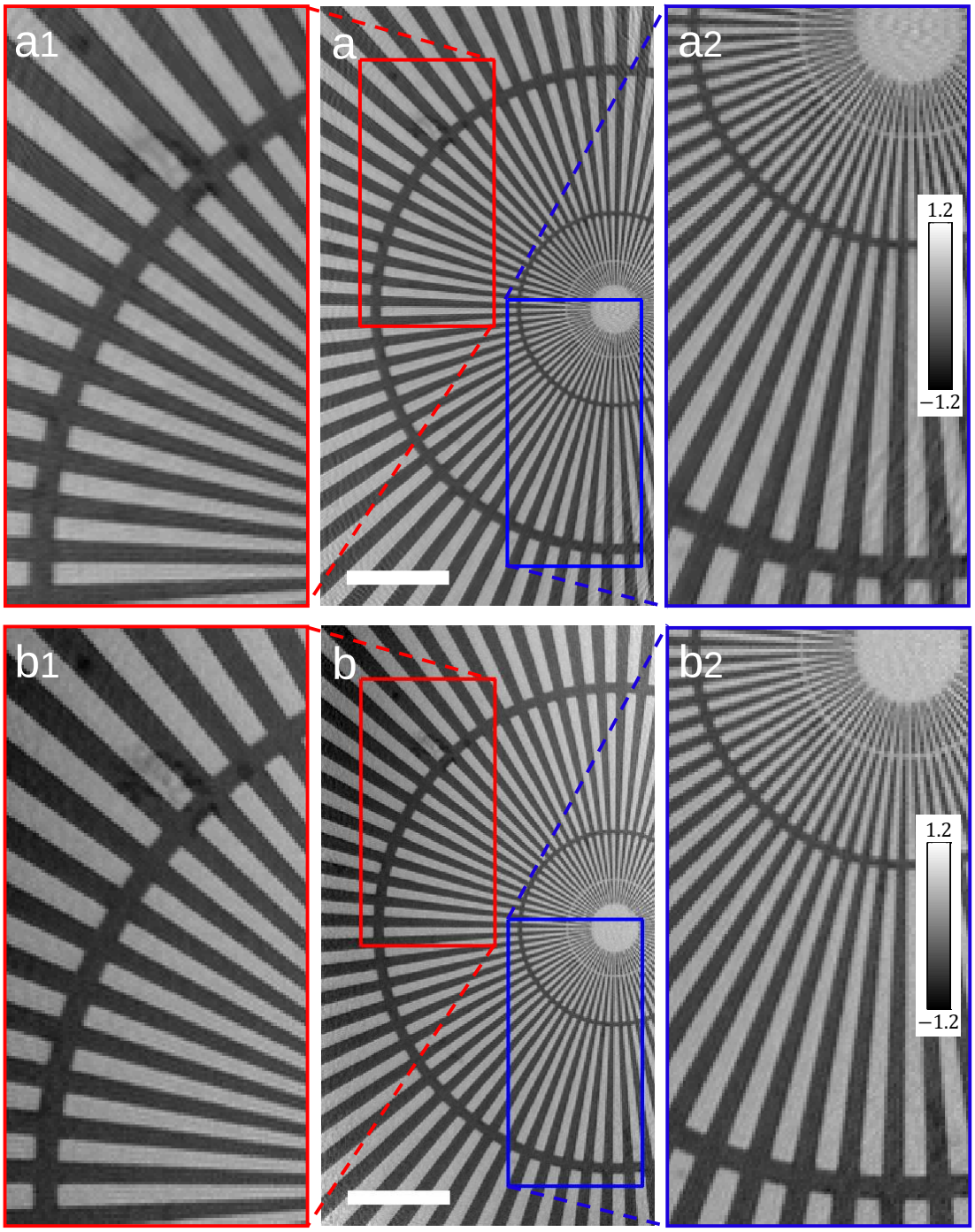

Figure 4.13: The phase reconstructions of the Siemens star from (a) the D_{out} configuration and (b) the D_{in} configuration. The scale bars represent 5μm.

5 seconds is taken for each diffraction pattern. The same reconstruction strategy as above is used here and the results are shown in Fig. 4.15. For the D_{out} configuration, the specimen reconstruction is blurred. When the diffuser is present, the spikes in the specimen are reconstructed much more sharply. Clearly, the two sets of orthogonalised

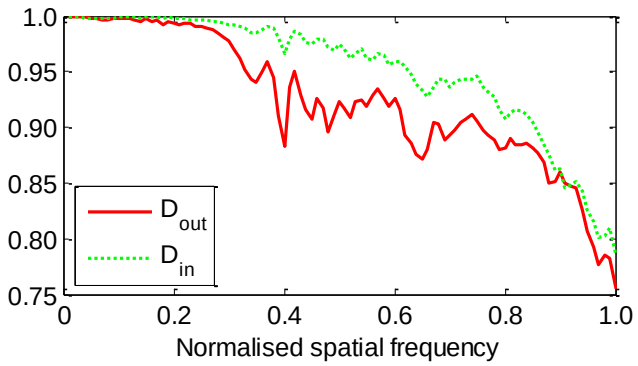

Figure 4.14: A comparison of the Fourier ring correlation (FRC) plots taken from the object reconstructions in the D_{out} configuration and the D_{in} configuration.

probe modes are not well reconstructed, because we would expect them to show the same structures as in Fig. 4.12 provided low coherence in the horizontal direction. However, some vertical lobes present in the orthogonalised probe modes reconstructed from the D_{in} configuration indicates that they are closer to the expected modes than the ones reconstructed from the D_{out} configuration. This means that highly structured illumination helps both the object reconstruction and the probe mode reconstructions.

As stated above, another advance of using a diffuser is that we can estimate the partial coherence of the experiment in real time. The phenomenon of coherent speckle in x-ray scattering is very well established [126], and its theory is very well understood [127, 128]. The intensity histogram of a coherent diffraction pattern from a random object should have an exponential profile. This means the probability of the measured intensities in the diffraction pattern falls as an exponential; the highest probability thus locates at the zero intensity. However, when any form of incoherence is present, zero intensity is never measured. As a result, the highest probability shifts to some finite intensity and the histogram will have a rising hump. To demonstrate this, we plot the histogram of one of the diffraction patterns measured from our first experiment with diffuser in and also the histogram of the same diffraction pattern calculated from the first probe mode reconstructed, as shown in Fig. 4.16. With only one mode, the

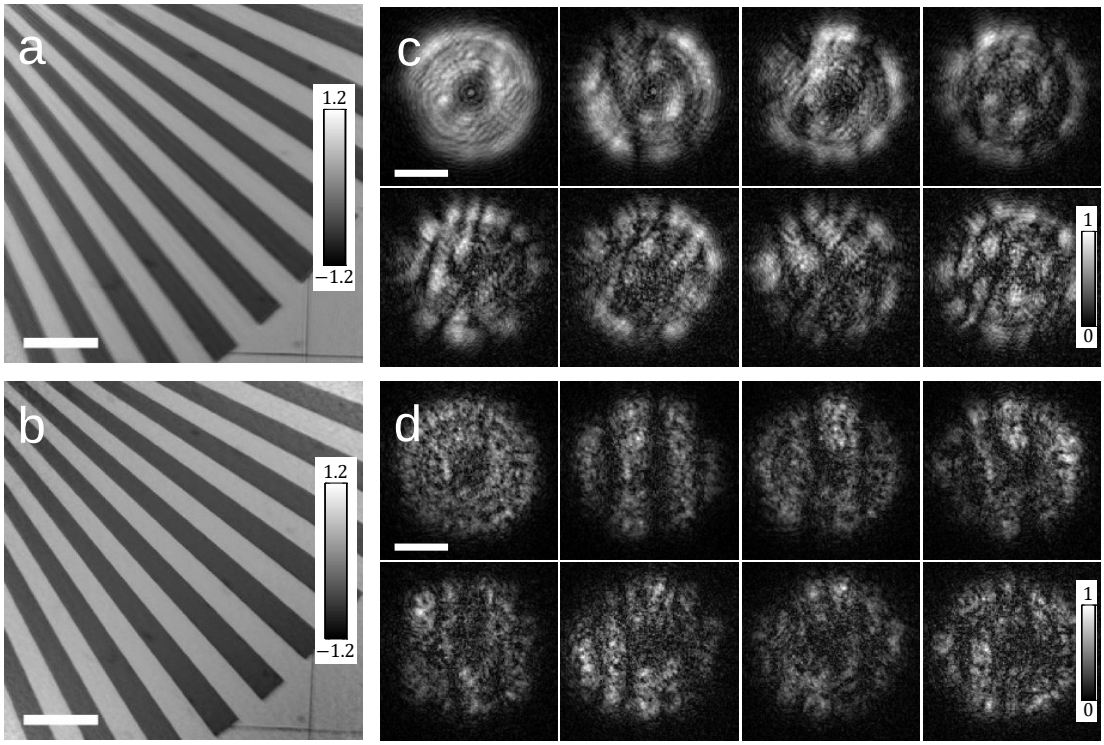

Figure 4.15: Reconstructions when the specimen has sparser features. (a) The phase reconstruction of the object from the D_{out} configuration. (b) The modulus reconstruction of the probe modes from the D_{out} configuration. (c) The phase reconstruction of the object from the D_{in} configuration. (d) The modulus reconstruction of the probe modes from the D_{in} configuration. The scale bars represent 5μm.

diffraction pattern is coherent and consequently the histogram is an exponential decay with the peak located at zero (see the red plot), whereas the histogram of the diffraction pattern measured under low degree of coherence has a rising hump (see the green plot).

When the histogram is plotted as a function of the number of coherent modes for a specific experiment, we can use it as an empirical indicator to estimate the degree of coherence of the experiment, although the exact modelling of the plots is depending on many factors, like the detector pixel size and the ratio between the typical structure size in the diffuser and the probe size. The histogram can be plotted in real time, which means we can adjust the coherence of the experiment qualitatively before collecting data.

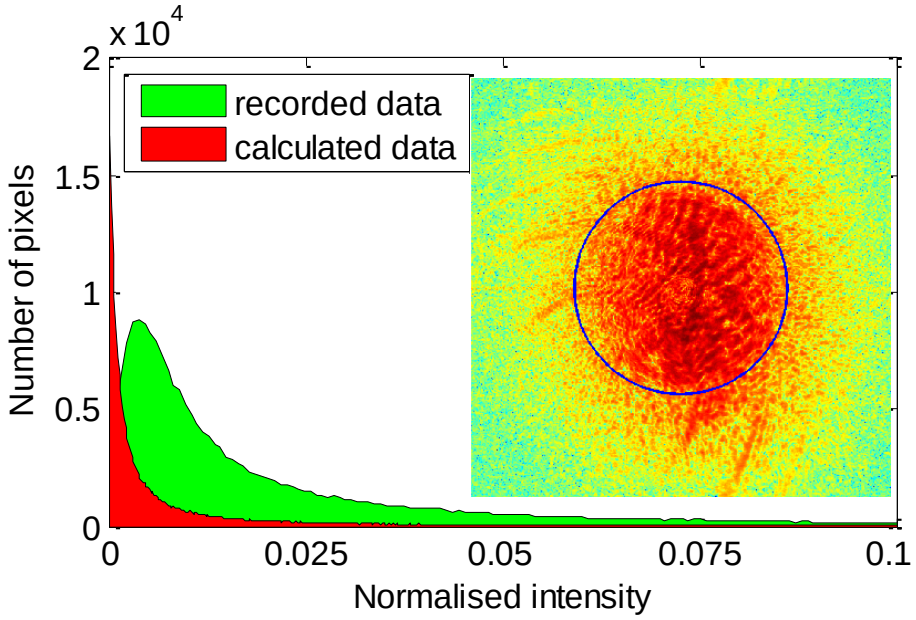

Figure 4.16: A comparison of the histograms from a typical diffraction pattern, as recorded (green) and as calculated using the first probe mode (red). The inset is the calculated diffraction pattern relating to the first probe mode. The corresponding recorded diffraction pattern is shown in Fig. 4.11b. The histograms are plotted using the region outside of the blue circle. Note that in order to show the different location of the peaks, the histogram plots are cut to the bottom 10% of the intensities for display purpose.

4.3 Conclusions

In this chapter, we have mathematically shown when a multi-state probe interacts with a multi-state object and only the intensity addition of all the resulting diffraction states is constrained during the reconstruction, some linear combinations of the underlying states (both object and probe) also satisfy the modulus constraint, leading to reconstruction ambiguities. Despite the linear reconstruction ambiguity, the normalised intensity addition of the all the states reconstructions is unique and true for both the probe case and the object case. Also, we have derived how the phase-only constraint and orthogonality constraint break the reconstruction ambiguity. An optical experiment that involves a dynamic object composed of two states (not mutually orthogonal) and a spatially partial coherent illumination with five modes has successfully demonstrated the validity of these conjectures.

Moreover, to test the performance of a diffuser in ptychographic reconstruction in the presence of partial coherence, we have carried out x-ray experiments to compare the configuration with only a FZP and that with a FZP plus a diffuser. The introduction of spatial partial coherence helps increase the flux, but the information encoded in the diffraction intensity measurements is corrupted by the partial coherence. In this case, the experimental results have shown that a diffused probe produces better reconstructions than a non-diffused probe, especially when the specimen has very simple and sparse structures. This is attributed to the fact that the diffuser increases the angular range in the probe that helps enrich the information content contained in the intensity measurements.

Furthermore, we demonstrate the advantage of estimating the degree of partial coherence by using a diffuser. It makes use of the histogram plot of the diffraction pattern. When the experiment is coherent, the histogram has an exponential profile with the peak located at zero. However, when the experiment is partial coherent, the peak shifts to a higher intensity and the histogram will have a rising hump. If the relationship between the histogram plot and the number of coherent modes for a specific experiment can be found, we are able to adjust the coherence of the experiment in real time.

5 Temporally Mixed State Ptychography

When the radiation source is temporally partial coherent, each wavelength will cause a different response from the specimen to give a single diffraction state in the detector plane and the measurement will be the incoherent summation of the states from all the wavelengths [80, 82]. In this case, the probe/object state is wavelength-dependent and each probe state only interacts with the object state triggered by the specific wavelength. We will mathematically show in this chapter that it is this wavelength dependent state interaction that eliminates the linear reconstruction ambiguity. Experiments using two lasers with wavelengths of 633nm and 675nm are carried out and successfully demonstrate the validity of the mathematics. Also, by treating the static background noise in the diffraction patterns as an extra temporal state, we develop an iterative update function for the background in the reconstruction algorithm. In this way, we manage to reconstruct the background and remove it from the signals iteratively without the background subtraction step.

5.1 Temporally mixed state ptychography

In this section, we will give a description of the modified reconstruction algorithm for temporally mixed state ptychography based on ePIE and its relation with the gradient descent method, and present the derivation of how the linear reconstruction ambiguity is broken in this case. Experimental demonstration will also be presented.

5.1.1 The reconstruction algorithm

The ePIE algorithm has been adapted to multi-wavelength ptychography by Batey et al [72]. The modulus constraint is modified in the same way as the spatially mixed state

ptychography. Using the same notation as used in section 4.1.1, here we present the iterative reconstruction algorithm for temporally mixed state ptychography in detail as:

1. In the *n*th iteration, the exit wave guess at a probe position $\boldsymbol{r}_j$ is formed by the product of the current object state guess and the illumination state guess at the corresponding scan position:

$$\psi_j^{(k)}(\boldsymbol{r}) = P_n^{(k)}(\boldsymbol{r}-\boldsymbol{r}_j) O_n^{(k)}(\boldsymbol{r}). \tag{5.1}$$

2. In the far field the diffraction pattern guess is then formed by Fourier transforming the exit wave guess:

$$\Psi_j^{(k)}(\boldsymbol{q}) = \mathcal{F}\left[\psi_j^{(k)}(\boldsymbol{r})\right]. \tag{5.2}$$

3. Apply the modulus constraint, i.e. replace the modulus of the diffraction pattern guess with the square root of the scaled intensity measurement and keep the phase part, and back propagate to real space to get an updated exit wave $\psi'^{(k)}_j$:

$$\psi'^{(k)}_j(\boldsymbol{r}) = \mathcal{F}^{-1}\left[\sqrt{\frac{\left|\Psi_j^{(k)}(\boldsymbol{q})\right|^2}{\sum_k \left|\Psi_j^{(k)}(\boldsymbol{q})\right|^2} M_j(\boldsymbol{q})}\, \frac{\Psi_j^{(k)}(\boldsymbol{q})}{\left|\Psi_j^{(k)}(\boldsymbol{q})\right|}\right]. \tag{5.3}$$

4. The update functions for the illumination states and the object states are given as:

$$\left.\begin{aligned} P_{n+1}^{(k)}(\boldsymbol{r}) &= P_n^{(k)}(\boldsymbol{r}) + \alpha \frac{O_n^{(k)*}(\boldsymbol{r}+\boldsymbol{r}_j)\left[\psi'^{(k)}_j(\boldsymbol{r}) - \psi_j^{(k)}(\boldsymbol{r})\right]}{\left(\left|O_n^{(k)}(\boldsymbol{r}+\boldsymbol{r}_j)\right|^2\right)_{\max}} \\ O_{n+1}^{(k)}(\boldsymbol{r}) &= O_n^{(k)}(\boldsymbol{r}) + \beta \frac{P_n^{(k)*}(\boldsymbol{r}-\boldsymbol{r}_j)\left[\psi'^{(k)}_j(\boldsymbol{r}) - \psi_j^{(k)}(\boldsymbol{r})\right]}{\left(\left|P_n^{(k)}(\boldsymbol{r}-\boldsymbol{r}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{5.4}$$

In the same way as the update functions of ePIE derived in section 2.3.7, in this case the objective function is to minimise the difference between $\psi_j^{(k)}(\boldsymbol{r})$ and $\psi'^{(k)}_j(\boldsymbol{r})$, i.e.

$$E = \sum_k \left| P_n^{(k)}(\boldsymbol{r}-\boldsymbol{r}_j) O_n^{(k)}(\boldsymbol{r}) - \psi'^{(k)}_j(\boldsymbol{r}) \right|^2 . \tag{5.5}$$

Take partial derivative with respect to each illumination state and each object state respectively, the gradients are given as

$$\left.\begin{aligned} G_P^{(k)} &= 2O_n^{(k)*}(\boldsymbol{r}+\boldsymbol{r}_j)\left[\psi_j^{(k)}(\boldsymbol{r}) - \psi'^{(k)}_j(\boldsymbol{r})\right] \\ G_O^{(k)} &= 2P_n^{(k)*}(\boldsymbol{r}-\boldsymbol{r}_j)\left[\psi_j^{(k)}(\boldsymbol{r}) - \psi'^{(k)}_j(\boldsymbol{r})\right] \end{aligned}\right\}. \tag{5.6}$$

So the update functions given in step 4 correspond to the choices of the search step size of

$$\left.\begin{aligned} \sigma_P^{(k)} &= \frac{\alpha}{2\left(\left|O_n^{(k)}(\boldsymbol{r}+\boldsymbol{r}_j)\right|^2\right)_{\max}} \\ \sigma_O^{(k)} &= \frac{\beta}{2\left(\left|P_n^{(k)}(\boldsymbol{r}-\boldsymbol{r}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{5.7}$$

The parameters α and β can be adjusted to find a balance between the convergence speed and the reconstruction stability. They are normally set to a value between 0 and 1.

5.1.2 The reconstruction structure

During the reconstruction of temporally mixed state ptychography, it is the same as spatially mixed state ptychography that the modulus constraint only constrains the intensity summation of all the states, not the intensity of individual state. However the interaction between the probe states and the object states is different here. A specific object state is triggered by a specific wavelength, so this object state will only respond to the probe state corresponding to that wavelength. There will be no cross-interaction

between different wavelengths. The resulting diffraction states are a subset of those in spatially mixed state ptychography. From a different perspective, we can assume temporally mixed state ptychography as a special case of spatially mixed state ptychography with a much more stringent constraint on the state interactions. In this section, we will mathematically show this constraint is able to break the linear reconstruction ambiguity and leads to unique solution.

Suppose a temporally partial coherent radiation source, which can be modelled by K wavelengths, will produce K probe states $\left\{P^{(k)}(\boldsymbol{r})\right\}_{k=1}^{K}$ and K object states $\left\{O^{(k)}(\boldsymbol{r})\right\}_{k=1}^{K}$. The corresponding diffraction states $\left\{\Psi^{(k)}(\boldsymbol{q})\right\}_{k=1}^{K}$ will be $\left\{\mathcal{F}\left[P^{(k)}(\boldsymbol{r})O^{(k)}(\boldsymbol{r})\right]\right\}_{k=1}^{K}$. The diffraction intensity measurement $M(\boldsymbol{q})$ would be

$$M(\boldsymbol{q})=\sum_{k=1}^{K}\left|\Psi^{(k)}(\boldsymbol{q})\right|^2=\sum_{k=1}^{K}\left|\mathcal{F}\left[P^{(k)}(\boldsymbol{r})O^{(k)}(\boldsymbol{r})\right]\right|^2. \tag{5.8}$$

Again, since the above relation applies to every scan position, the corresponding index will not be considered here. If assembling the underlying probe states as a row vector $\boldsymbol{P}=\left[P^{(1)}(\boldsymbol{r}),P^{(2)}(\boldsymbol{r}),\cdots,P^{(K)}(\boldsymbol{r})\right]$, the underlying object states as a row vector $\boldsymbol{O}=\left[O^{(1)}(\boldsymbol{r}),O^{(2)}(\boldsymbol{r}),\cdots,O^{(K)}(\boldsymbol{r})\right]$, the reconstructed probe states as a row vector $\tilde{\boldsymbol{P}}=\left[\tilde{P}^{(1)}(\boldsymbol{r}),\tilde{P}^{(2)}(\boldsymbol{r}),\cdots,\tilde{P}^{(K)}(\boldsymbol{r})\right]$, the reconstructed object states as a row vector $\tilde{\boldsymbol{O}}=\left[\tilde{O}^{(1)}(\boldsymbol{r}),\tilde{O}^{(2)}(\boldsymbol{r}),\cdots,\tilde{O}^{(K)}(\boldsymbol{r})\right]$, and also assuming the linear relation expressed by Eq. (4.11) between the underlying states and the reconstructed states is still applicable here, we then have the intensity summation $\tilde{M}(\boldsymbol{q})$ of all the reconstructed diffraction states described by

$$\begin{aligned}\tilde{M}(\boldsymbol{q}) &= \sum_{k=1}^{K}\left|\tilde{\Psi}^{(k)}(\boldsymbol{q})\right|^2 \\ &= \sum_{k=1}^{K}\left|\mathcal{F}\left[\tilde{P}^{(k)}(\boldsymbol{r})\tilde{O}^{(k)}(\boldsymbol{r})\right]\right|^2 = \sum_{k=1}^{K}\left|\mathcal{F}\left[P^{(i)}(\boldsymbol{r})U_P^{ik}O^{(j)}(\boldsymbol{r})U_O^{jk}\right]\right|^2.\end{aligned} \tag{5.9}$$

The modulus constraint requires $\tilde{M}(\boldsymbol{q}) = M(\boldsymbol{q})$, which leads to

$$\sum_{k=1}^{K}\left|U_P^{ik}U_O^{jk}\right|^2 = \begin{cases}1, & i = j\\ 0, & i \neq j\end{cases}. \tag{5.10}$$

It should be noted here that both $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ are full rank square matrix. This means there will be no columns or rows full of zeros. Under this condition, for $i \neq j$ in Eq. (5.10), $\sum_{k=1}^{K}\left|U_P^{ik}U_O^{jk}\right|^2 = 0$, which is equivalent to $U_P^{ik}U_O^{jk} = 0$, requires that in each row and each column of both $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ there is exactly one non-zero entry (this is the definition of generalised permutation matrix) and that the positions of the non-zero entries in $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ are the same (i.e. the non-zero patterns for $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ are the same). As a result, the linear reconstruction ambiguity is broken in temporally mixed state ptychography. The positions of the non-zero entries in $\boldsymbol{U}_P$ and $\boldsymbol{U}_O$ determine the order of the reconstructed probe/object states, which will stay ambiguous. Furthermore, for $i = j$ in Eq. (5.10), $\sum_{k=1}^{K}\left|U_P^{ik}U_O^{jk}\right|^2 = 1$ is now equivalent to $\left|U_P^{ik}U_O^{ik}\right| = 1$, which means scale and phase offset are allowed for the interacting pairs of probe and object states in the reconstruction. As we can see, each state reconstruction in temporally mixed state ptychography is independent and it is exactly the same as the reconstruction structure of the coherent single state ptychography.

5.1.3 Optical experiments

Optical experiment with multi-wavelength has been demonstrated before using three lasers with distinctive wavelengths (410 nm, 530 nm and 633 nm) and the reconstructions also clearly showed the breakdown of the ambiguity [72]. Here we demonstrate temporally mixed state ptychography by using a 633nm HeNe laser and a 675nm diode laser as the source. The setup is shown in Fig. 5.1. The two lasers are coupled via a beam splitter to provide the illumination. A circular aperture with a diameter of 150μm is inserted upstream of the specimen to form a localised illumination. The transmitted diffraction patterns are recorded on a CCD detector (an AVT Pike F421B, with 2048×2048 pixels, each of 7.4μm^2) positioned at about 26mm downstream of the specimen. A microscopy slide of a clam gill section is used as the specimen and it is mounted on an *x*-*y* translation stage to implement the ptychographic scan.

In order to obtain reference images of two wavelengths for later comparisons, we perform two calibration experiments with only one wavelength switched on each time during the ptychographic scan. The illumination is scanned over the patterns on a grid of 20×20. At each scan, four exposures (0.8ms, 4ms, 20ms and 50ms) are taken and

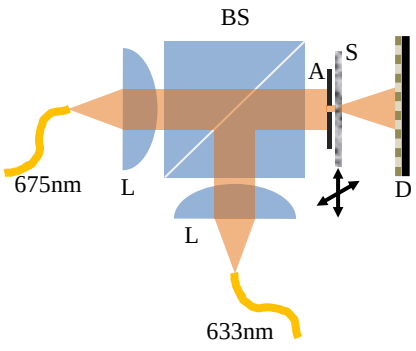

Figure 5.1: The experimental configuration. Two lasers, 675nm and 633nm, are coupled via a beam splitter as the radiation source. A circular aperture with a diameter of 150μm is placed upstream of the specimen to provide the localised illumination. A CCD detector is positioned at about 25mm downstream of the specimen to record the diffraction patterns. L stands for lens, BS for beam splitter, A for aperture, S for specimen and D for detector.

combined to increase the dynamic range of the detector. The step size is about 30μm with ±20% random.

For the reconstructions, the diffraction patterns are binned by a factor of 4 to 512×512 (the sampling requirement is still well fulfilled) and only a central region of 256×256 (outside of which very little photons are detected) is used. The initial guess for the object is free space (i.e. 1 everywhere) and for the probe is a circular aperture with a diameter of 150μm. 300 iterations of the ePIE algorithm are run and the reconstructions are shown in Fig. 5.2. The colour wheel depiction shows phase variations as changes in hue and modulus variations as changes in brightness. As we can see, the specimen has different responses for different wavelengths. Also, the scale of the two reconstructions is different due to the difference in wavelength causing different pixel sizes in the reconstructed images, which is very obvious in the illumination modulus reconstructions. Different phase ramps are reconstructed in the illumination phases, indicating the misalignment in the propagation direction of the two laser beams, given that no obvious phase ramp ambiguity exists in the object reconstructions.

We repeat our experiment with both two wavelengths switched on during the scan. This

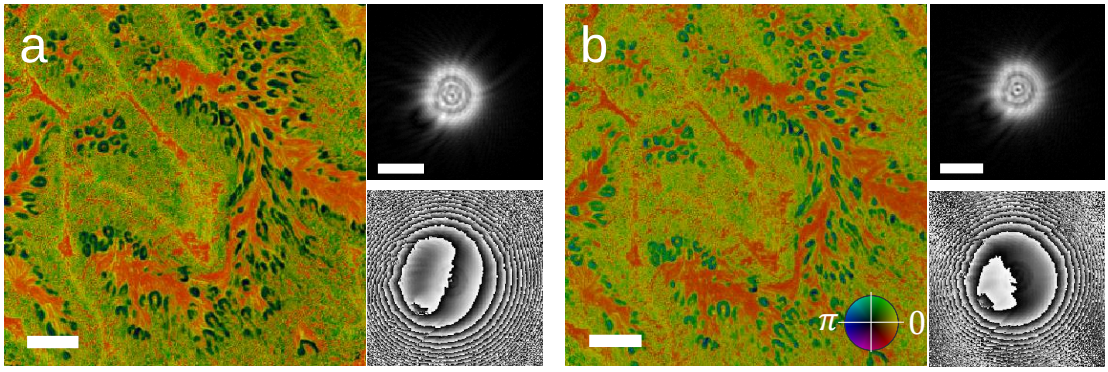

Figure 5.2: The reconstructions from the two single wavelength experiments. (a) 633nm. (b) 675nm. In each figure, the left panel is the object reconstruction and the right panel is the probe reconstruction with the modulus on top and the phase at the bottom. The scale bars indicate a length of 100um. The colorwheel applies to all the figures of the object reconstructions.

experiment, in theory, is equivalent to combining the two calibrated experiments into one. Again, the diffraction patterns are binned by a factor of 4 and a central region of 256×256 is cropped out for the reconstructions. Also, the initial guesses for the two object states are free space (i.e. 1 everywhere) and for the two probe states are given by the product of the 150μm circular aperture with two different arrays of random values. We use the modified ePIE algorithm to run the reconstruction for 300 iterations. It should be noted here that the algorithm does not require the initial guesses for either the two probe states or the two object states to be different, because the different pixel sizes in the reconstructions will help the algorithm to differentiate the states. The results are shown in Fig. 5.3 and they have a good agreement with the calibrated results shown in Fig. 5.2. The two wavelengths are clearly well separated, indicating that the discrepancy between the specimen responses to the two wavelengths is big enough to be assumed as two states for the reconstruction algorithm, because otherwise the linear reconstruction ambiguity will occur for the probe states reconstructions.

This could be very useful, because it shows the possibility of spectroscopic imaging via ptychography using just one scan [129]. It could also relax the requirement for the tem-

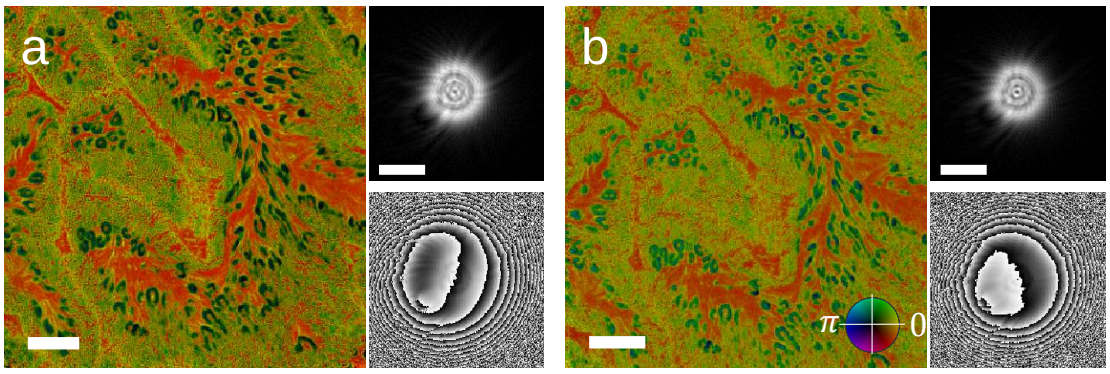

Figure 5.3: The reconstructions from the double wavelengths experiment. (a) 633nm. (b) 675nm. In each figure, the left panel is the object reconstruction and the right panel is the probe reconstruction with the modulus on top and the phase at the bottom. The scale bars indicate a length of 100um. The colorwheel applies to all the figures of the object reconstructions.

poral coherence in the experiment. To further explore it, hopefully we could carry out experiments using a source with continuous spectrum of broad bandwidth in the near future.

5.2 An iterative method to remove background noise

As we explained in section 2.3.10, the background noise is the constant response of the detector in the absence of radiation. We need to measure it by taking a dark frame of the detector without turning on the radiation and then remove it by subtracting it from the diffraction intensity measurements. As an inspiration from the temporally mixed state ptychography, by treating the background noise as the diffraction pattern from an extra temporal state, we are able to reconstruct and remove the background noise iteratively in the algorithm without the need of measuring and subtracting the dark frame.

5.2.1 The background noise update function

For the extra temporal state corresponding to the background noise, we denote its illumination function as P^{BG} and its object response as O^{BG}. According to the modified ePIE algorithm for the temporally mixed state ptychography, the update function for the illumination is

$$P_{n+1}^{BG}\left(\boldsymbol{r}\right)=P_{n}^{BG}\left(\boldsymbol{r}\right)+\alpha\frac{O^{BG*}\left(\boldsymbol{r}+\boldsymbol{r}_{j}\right)\left[\psi'^{BG}_{j}\left(\boldsymbol{r}\right)-\psi_{j}^{BG}\left(\boldsymbol{r}\right)\right]}{\left(\left|O^{BG}\left(\boldsymbol{r}+\boldsymbol{r}_{j}\right)\right|^{2}\right)_{\max}}. \tag{5.11}$$

Since the background noise is the same for all the scan positions, for this extra temporal state its object response O^{BG} should be constant everywhere and we do not need to update it during the iteration. As a result, the update function can be simplified as

$$P_{n+1}^{BG}\left(\boldsymbol{r}\right)=P_{n}^{BG}\left(\boldsymbol{r}\right)+\alpha\left[P'^{BG}_{n}\left(\boldsymbol{r}\right)-P_{n}^{BG}\left(\boldsymbol{r}\right)\right], \tag{5.12}$$

where P'^{BG}_n denotes the illumination function after applying the modulus constraint. Since we only care about the background noise (i.e. the intensity of the Fourier transform of the illumination function), not the illumination function itself, the illumination update function can be transferred to its Fourier space by taking the Fourier transform of both sides of the equal sign in Eq. (5.12)

$$\overline{P^{BG}_{n+1}(\boldsymbol{q})} = \overline{P^{BG}_{n}(\boldsymbol{q})} + \alpha\left[\overline{P'^{BG}_{n}(\boldsymbol{q})} - \overline{P^{BG}_{n}(\boldsymbol{q})}\right], \tag{5.13}$$

where the overline denotes the Fourier transform of the corresponding functions. The modulus square of $\overline{P^{BG}_{n+1}}$ and $\overline{P^{BG}_{n}}$ are basically the current and previous estimates of the background noise respectively. The modulus square of $\overline{P'^{BG}_{n}}$ is actually the assigned intensity of the background noise and it is given by

$$I^{BG}(\boldsymbol{q}) = \frac{\left|\Psi^{BG}_{j}(\boldsymbol{q})\right|^2}{\left|\Psi^{BG}_{j}(\boldsymbol{q})\right|^2 + \left|\Psi^{S}_{j}(\boldsymbol{q})\right|^2} M_j(\boldsymbol{q}), \tag{5.14}$$

where Ψ^S is the forward calculation of the diffraction pattern of the temporal state from the signal, Ψ^{BG} is that from the background noise, and M is the diffraction intensity measurement. Since the phase parts of $\overline{P^{BG}_{n}}$ and $\overline{P'^{BG}_{n}}$ are the same, if we take the modulus square of both sides of the equal sign in Eq. (5.13), we have

$$B_{\mathrm{n+1}}(\boldsymbol{q}) = B_{\mathrm{n}}(\boldsymbol{q})\left[1 + \alpha\left(\sqrt{\frac{I^{BG}(\boldsymbol{q})}{B_{\mathrm{n}}(\boldsymbol{q})}} - 1\right)\right]^2, \tag{5.15}$$

with $B_n(\boldsymbol{q}) = \left|\overline{P^{BG}_{n}(\boldsymbol{q})}\right|^2$ and $B_{\mathrm{n+1}}(\boldsymbol{q}) = \left|\overline{P^{BG}_{n+1}(\boldsymbol{q})}\right|^2$, where B denotes the background noise update. Adding this extra update function into the ePIE algorithm, we are able to reconstruct the background noise and separate it from the signal.

5.2.2 Optical Experiments

To demonstrate the background noise reconstruction, we reuse the ptychographic data from the calibration experiment of 633nm HeNe laser in section 5.1.3, but without the background subtraction. The background noise and an example of the diffraction intensity measurements with and without the background subtraction are shown in Fig. 5.4. As we can see, the background noise seriously corrupts the diffraction data, especially at the high scattering angles, where very low counts can be detected.

The reconstruction of the uncorrupted diffraction data (with the background subtraction) is already done in section 5.1.3 and it is shown in Fig. 5.2a. The corrupted diffraction data (without the background subtraction) is first reconstructed by running the normal ePIE algorithm for 300 iterations and the results are shown in Fig. 5.5. The object is very blurry and it is clearly not well reconstructed. As for the probe, in the modulus a bright spot disguises the real structures that can still be observed although not well developed in the phase part. The intensity of the Fourier transform of the probe (see Fig. 5.5b) shows a resemblance to the background noise. This means the algorithm treats the background noise as part of the spectrum of the probe. This is easy to understand. Because during the ePIE reconstruction the probe function is the same through all the

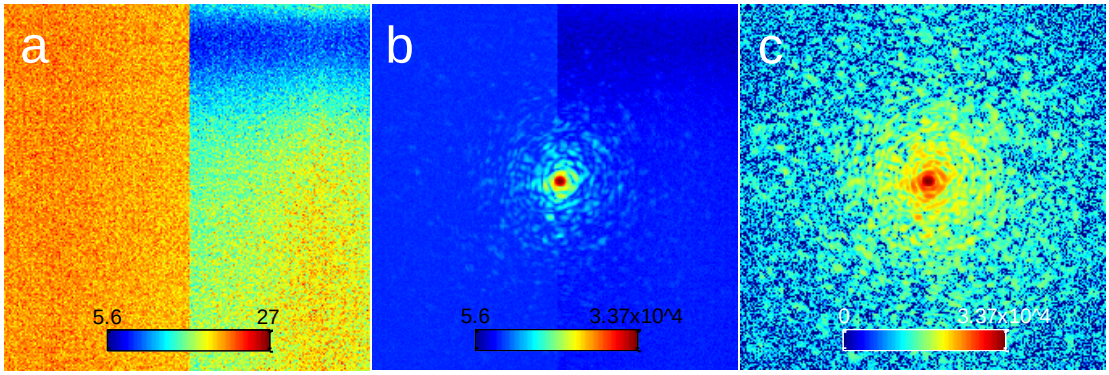

Figure 5.4: The measured background noise (a) and one example of diffraction intensity measurements with (b) and without (c) the background subtraction. The background noise has relatively small values. However, it seriously corrupts the diffraction data of the signals, especially at the high scattering angles.

scan positions, the algorithm will assume any unchanged part in the diffraction pattern is from the probe function. Since the background noise spans the entire frequency space and has relatively uniform response, it causes a bright spot in the probe reconstruction. As this reconstructed probe function scanned over the specimen, different area of the specimen will interact with the bright spot, still causing slightly difference in the diffraction pattern corresponding to the background noise. So the ePIE algorithm in its traditional way is not able to accurately account for the background noise. In addition, the bright spot has another detrimental effect on the reconstruction. The update function for the object in ePIE (see section 2.3.7) is normalised to the square of the maximum value of the modulus of the probe. The bright spot will dramatically slow down the update of the object function.

The corrupted diffraction pattern is then reconstructed by running the modified ePIE algorithm for 300 iterations and the results are shown in Fig. 5.6. Both the object and the probe are very well reconstructed and cannot be visually differentiated from the reference reconstructions (shown in Fig. 5.2a). In addition, the background is also reconstructed. Except at the low frequencies, the reconstructed background noise has a

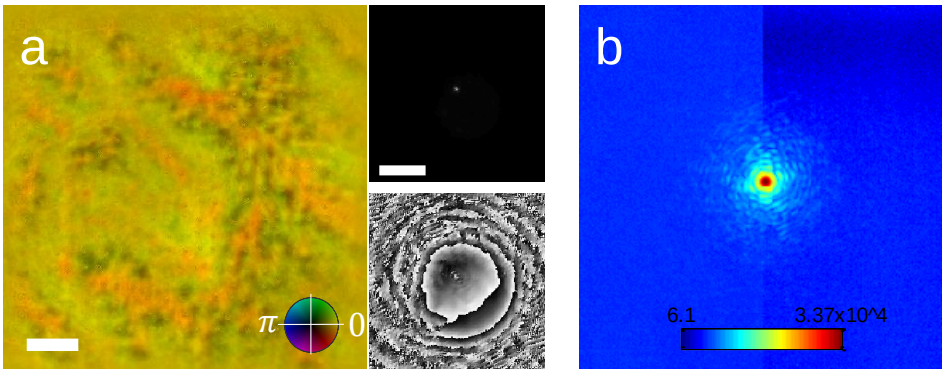

Figure 5.5: The reconstructions from the corrupted diffraction data using the normal ePIE algorithm. (a) The object and probe reconstructions. The left panel is the object reconstruction and the right panel is the probe reconstruction with the modulus on top and the phase at the bottom. The scale bars indicate a length of 100um. (b) The intensity of the Fourier transform of the reconstructed probe. It has a good resemblance to the background noise, indicating the algorithm treats the background noise as part of the spectrum of the probe.

good agreement with the measured one. The static background noise reconstruction highly relies on the variation of the signal from one scan position to another. Otherwise the algorithm cannot differentiate the signals from the background noise. For the low frequencies, the diffraction pattern normally varies little between different scan positions. This is because the low frequencies in the specimen do not have enough sampling due to the sparse scan compared to the dense scan needed by WDDM (see chapter 3). Because of this slow variation of the signals at the low frequencies, the background noise is not very well reconstructed at these areas. However, the low frequencies normally have very high counts, so that the background noise has very little effects on them. For this reason, the reconstruction shown in Fig. 5.6 can be assumed successful.

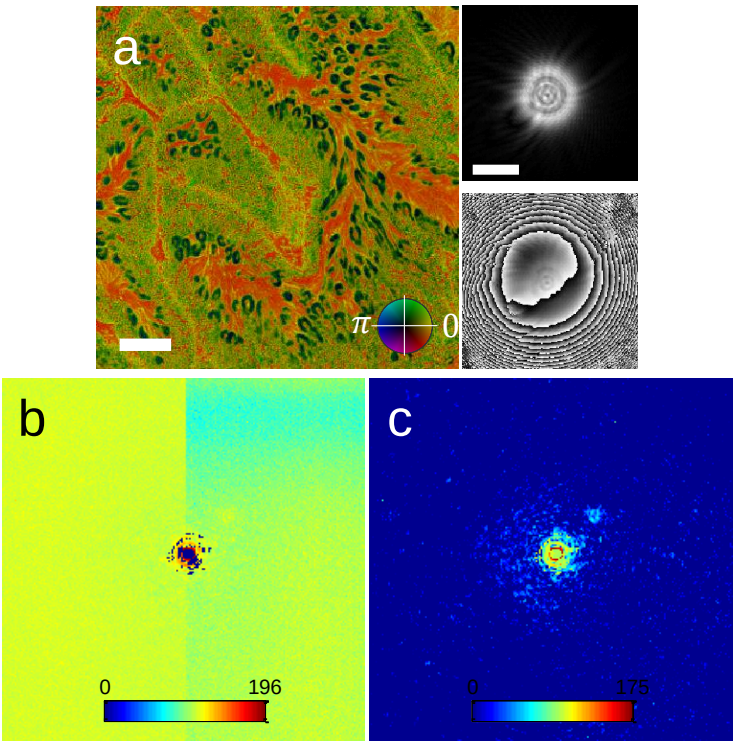

Figure 5.6: The reconstructions from the corrupted diffraction data using the modified ePIE algorithm with background noise update function. (a) The object and probe reconstructions. The left panel is the object reconstruction and the right panel is the probe reconstruction with the modulus on top and the phase at the bottom. The scale bars indicate a length of 100um. (b) The intensity of the reconstructed background noise. (c) The difference between the measured background noise and the reconstructed version.

5.3 Conclusions

In this chapter, we have mathematically proven that the linear reconstruction ambiguity occurred in spatially mixed state ptychography is broken in the case of temporal mixed state ptychography, although the modulus constraint is applied in the same way for the two cases. It is because the temporal mixed state ptychography has a non-cross interaction between the states from the different wavelengths. For a specific wavelength, the triggered object state only responds to the corresponding probe state. This *a priori* information is fed into the algorithm acting like a strong constraint and it breaks the linear reconstruction ambiguity.

Here we have carried out an optical experiment using a clam gill section with two lasers, 633nm and 675nm, to demonstrate the breakdown of the reconstruction ambiguity. The good reconstructions have shown that this technique is able to differentiate the two wavelengths for this specific specimen. To further explore it, experiments using a source with continuous broadband spectrum would be ideal, so that we could show that the requirement for the temporal coherence can be relaxed. It could also indicate the possibility of spectroscopic imaging via ptychography using just one scan.

As an inspiration from the temporally mixed state ptychography, we have developed an iterative method to reconstruction the background noise in the diffraction intensity measurements by treating it as the diffraction pattern from an extra temporal state. In this way, we are able to implement the reconstruction without the need of measuring and subtracting the background noise from diffraction intensity measurements. Experimental data reconstructions have been carried out and the results have shown that the recovery of the background noise relies on the variations of the signals through the

scan. No variation or even slow variations, like the low frequency areas in the diffraction patterns, will cause high errors in the reconstructed background noise.

6 3D Fourier Ptychography

Optical 3D imaging is generally realised via two ways: optical tomography and optical sectioning. The common idea of optical tomography is to record multiple images for various angles of illumination with respect to the specimen, and then to reconstruct 3D structure with this set of angular images. Optical sectioning, on the other hand, involves collecting a set of images at different longitudinal positions by an axial scanning and then assembling the set of axial images to form the 3D reconstruction.

Ptychography has been adapted to produce 3D reconstructions by combining with the multislice method [67, 68, 69]. Multislice ptychography belongs to the category of optical sectioning, but it does not need to perform a physical scan along the longitudinal axis. The sectioning is numerically carried out in the reconstruction algorithm. It thus saves the data acquisition time compared with the conventional optical sectioning methods like confocal scanning microscopy [130]. Moreover, multislice ptychography delivers quantitative phase reconstruction of a 3D structure. The phase reconstruction is able to provide good contrast of the structure at its nature state without any assistance like fluorescence used in confocal scanning microscopy and light sheet microscopy. The quantitative information is further able to allow us to study the structure quantitatively. These properties are very important in biologic imaging applications, like the live cell imaging [131].

Fourier ptychography is a Fourier domain variant of ptychography [78]. To help differentiate these two types of ptychography, here we refer to the conventional ptychography as real-space ptychography. Instead of scanning a localised illumination function over an extended specimen and recording the intensities of the diffraction patterns in real-

space ptychography, in Fourier ptychography one scans a small aperture over the spectrum of the extended specimen and records the intensities of the resulting low-resolution images. Methodologically, the two forms of ptychography are the same: a localised function is scanned over a fixed extended function with certain amount of overlap and at each scan position the measurement is made in the corresponding reciprocal space with the phase information missing. Therefore the ePIE algorithm can be applied to Fourier ptychography.

However, unlike the specimen function in real-space ptychography, the spectrum of the specimen in Fourier ptychography has a very high dynamic range and a dramatically bright spot (the DC component) in the centre. This will cause slow convergence in the reconstruction of the aperture function if we directly use the ePIE algorithm in Fourier ptychography, because the update function will be scaled to the maximum intensity value. Furthermore, it is not obvious how to extend the ePIE algorithm of Fourier ptychography to produce 3D reconstruction of the specimen via the multislice method. In this chapter, we will address these problems. The work presented in this chapter has led to a journal paper publication [132].

6.1 2D Fourier ptychography

A typical setup for Fourier ptychography is shown in Fig. 6.1. Provided the specimen is thin enough, the spectrum lying in the back focal plane of the lens is related to the specimen by a Fourier transform when the specimen is illuminated by a plane wave radiation. According to the shift theorem of the Fourier transform (see section 2.1.2), tilting the plane wave illumination corresponds to a shift of the spectrum. An aperture is inserted at the back focal plane with its centre coinciding with the optic axis to filter the spectrum. Meanwhile a detector is placed at the image plane to record the low-

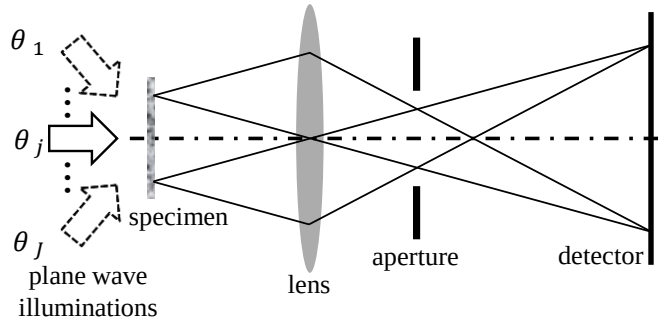

Figure 6.1: The optical setup of Fourier ptychography. When the specimen is thin enough, tilting the plane wave illumination corresponds to a shift of the spectrum. Meanwhile an aperture is inserted in the back focal plane of the imaging lens and a detector is placed at the image plane to record the resulting low-resolution images. In this way, a ptychographical scan can be performed in the back focal plane.

resolution image. Carefully choosing the tilt angle interval to ensure enough overlap of the specimen spectrum filtered by the aperture, in this way, we can perform a ptychographic scan by tilting the plane wave illumination. The aperture here is the counterpart of the localised illumination in real-space ptychography. The spectrum and low-resolution images here have the same roles as the specimen and the diffraction patterns do respectively in real-space ptychography.

Recorded images have low resolution due to the small aperture in the back focal plane and they contain different frequency information of the specimen spectrum. The different frequency information can be tiled to form a spectrum with much wider region than the size of the aperture. The larger the illumination tilt angle is, the higher the frequency can be extended. The resolution of the reconstructed specimen image is determined by the highest angle of the synthetic spectrum and it can be much smaller than the pixel size of the detector. This method is very similar to the aperture synthesis approach used in astronomy radio [133] and radar [134]. It was originally adopted in electron microscopy to overcome the resolution limit imposed by the small NAs of the electron lenses [135, 136]. Although at that time it was recognised as a reciprocal analogue of ptychography [137], but a good reconstruction algorithm was not available.

Fourier ptychography uses the ePIE algorithm to tile the different frequency information and hence achieve high resolution [78]. The outputs of the ePIE algorithm in Fourier ptychography are the aperture function and the synthetic specimen spectrum. To obtain the specimen image, an inverse Fourier transform needs to be carried out on the reconstructed spectrum.

Fourier ptychography, like real-space ptychography, is capable of delivering quantitative phase images at diffraction limited resolution. However, due to the specimen image having a very much lower dynamic range compared to its spectrum, Fourier ptychography has much lower requirement on the detector dynamic range compared to conventional ptychography. But when the illumination is tilted to high angles, very little scattered amplitude goes through the aperture. As a result, longer exposures might be needed for high tilted angles to ensure good signal-to-noise ratio. In real-space ptychography, this is compensated by the multiple exposures strategy (see section 2.3.10). Moreover, unlike real-space ptychography, the different angles of illumination can be supplied by LED arrays placed far away from the specimen, meaning that there is no need of scan: the whole imaging procedure can be controlled electronically without moving parts (of course this is only true for visible light). Furthermore, small defocus errors in the lens, which means the aperture slightly deviate from the back focal plane, can be reconstructed as part of the aperture function because they stay fixed substantially relative to the aperture when the plane wave illumination is tilted to different angles. As a result, the reconstructed specimen image is always in focus.

6.1.1 The reconstruction algorithm

As mentioned earlier, the ePIE algorithm can be adopted here for Fourier ptychography reconstruction. If we only consider the aperture plane (i.e. the back focal plane) and the

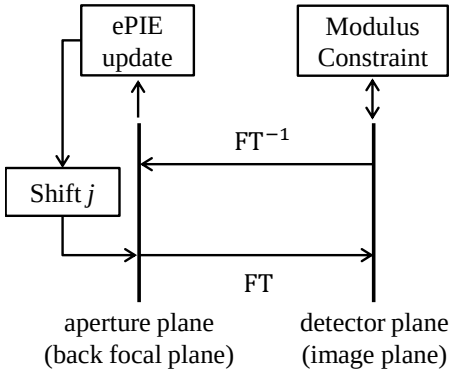

Figure 6.2: The flow chart of the ePIE algorithm for Fourier ptychography. Same as the ePIE algorithm used for conventional ptychography: modulus constraint is applied in the detector plane where the phase information is lost; ePIE update in the aperture plane, where the scan is performed, could reconstruct the aperture function and the spectrum of the specimen.

detector plane (i.e. the image plane), the flow chart of the ePIE algorithm for Fourier ptychography is depicted in Fig. 6.2. A tilt of the illumination becomes a shift of the spectrum in the back focal plane. The exit wave $S_E(\boldsymbol{q})$, i.e. the filtered spectrum, is given as the product of the aperture $A(\boldsymbol{q})$ and the spectrum $S_I(\boldsymbol{q})$, where $\boldsymbol{q}$ denotes the coordinate at the back focal plane. At the image plane, a low-resolution image is formed by the Fourier transform of the filtered spectrum. The square root of the intensity of the measured low-resolution image is used to replace the modulus of the formed image and the phase part remains untouched. By taking an inverse Fourier transform of the modified image, an updated version of the filtered spectrum $S'_E(\boldsymbol{q})$ is obtained at the back focal plane. If we use the update functions described in section 2.3.7, the aperture and the spectrum can be reconstructed.

Here for Fourier ptychography, we use the above notation and rewrite the ePIE update functions in Eq. (2.63) as

$$\left.\begin{aligned} A_{n+1}(\boldsymbol{q}) &= A_n(\boldsymbol{q}) + \alpha \frac{S^*_{n,I}(\boldsymbol{q}+\boldsymbol{q}_j)\left[S'_{j,E}(\boldsymbol{q}) - S_{j,E}(\boldsymbol{q})\right]}{\left(\left|S_{n,I}(\boldsymbol{q}+\boldsymbol{q}_j)\right|^2\right)_{\max}} \\ S_{n+1,I}(\boldsymbol{q}) &= S_{n,I}(\boldsymbol{q}) + \beta \frac{A^*_n(\boldsymbol{q}-\boldsymbol{q}_j)\left[S'_{j,E}(\boldsymbol{q}) - S_{j,E}(\boldsymbol{q})\right]}{\left(\left|A_n(\boldsymbol{q}-\boldsymbol{q}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{6.1}$$

where n indexes the iteration number and $\boldsymbol{q}_j$ denotes the lateral shift of the spectrum in the back focal plane when the jth tilted plane wave illuminated the specimen. So far, it has been exactly the same as real-space ptychography. However, different from the specimen function in real space, the specimen spectrum normally has high dynamic range with a very bright spot in the centre (corresponds to the unscattered beam). The iterative search step sizes for all the pixels in $A(\boldsymbol{q})$ are scaled by the square modulus of this brightest pixel. This will cause dramatically small search step sizes for the pixels that have very small modulus, hence slowing the convergences for these pixels. A parallel update version of the ePIE algorithm, which is first proposed by P. Godard et al to enhance the performance in the presence of noise [70], can mitigate this problem by using spatially variant search step sizes. Instead of trying to minimise the difference between the filtered spectrum $S_E(\boldsymbol{q})$ and its updated version $S'_E(\boldsymbol{q})$ for each tilted illumination separately, the parallel version minimises the difference for all the illumination angles at once. The objective function for the minimisation can be described as

$$E = \sum_{j=1}^{J} \left| A(\boldsymbol{q}) S_I(\boldsymbol{q}+\boldsymbol{q}_j) - S'_E(\boldsymbol{q}) \right|^2. \tag{6.2}$$

If we take the partial derivative of the objective function E with respect to the aperture function $A(\boldsymbol{q})$, we have the gradient expressed as

$$G_A(\boldsymbol{q}) = 2S_I^*(\boldsymbol{q}+\boldsymbol{q}_j)\left[S_E(\boldsymbol{q}) - S'_E(\boldsymbol{q})\right]. \tag{6.3}$$

The spatially variant search step sizes are chosen to be

$$\sigma_A(\boldsymbol{q}) = \frac{\alpha}{2\left(\sum_{j=1}^{J}\left|S_I(\boldsymbol{q}+\boldsymbol{q}_j)\right|^2 + \epsilon\right)}, \tag{6.4}$$

where ϵ is a small constant to prevent division by zero. Therefore the parallel update for the aperture function written in the form of gradient descend search method is given by

$$A_{n+1}(\boldsymbol{q}) = A_n(\boldsymbol{q}) + \alpha \frac{\sum_{j=1}^{J} S_{n,I}^*(\boldsymbol{q}+\boldsymbol{q}_j)\left[S'_{j,E}(\boldsymbol{q}) - S_{j,E}(\boldsymbol{q})\right]}{\left(\sum_{j=1}^{J}\left|S_{n,I}(\boldsymbol{q}+\boldsymbol{q}_j)\right|^2 + \epsilon\right)}. \tag{6.5}$$

Now the search step sizes for different pixels are scaled by different values of the same magnitude. In this way we make sure of relatively even convergences for all the pixels. However, unlike the sequential ePIE where the function is updated for each scan position, the parallel ePIE only updates the function once for all the positions. This property of parallel update on the contrary slows the convergence [70]. But luckily in Fourier ptychography, the aperture used is normally a circular function (with smoothly variant phase if defocus error exists), which is a very simple structure, and it thus does not need to be updated very frequently. For the spectrum reconstruction, since it does not have the high dynamic range problem as the aperture reconstruction does, we stick to the sequential update for the sake of convergence speed.

6.1.2 Optical experiment

An optical experiment is implemented here to compare the sequential ePIE algorithm (conventional version) and the parallel ePIE algorithm (modified version) as described above. The experiment was set up as sketched in Fig. 6.1. The illumination was

provided by a HeNe laser ($\lambda = 633\text{nm}$) coupled with an optical fibre via a 20× Olympus objective lens. The exit of the fibre, which produces a divergent point source, was mounted on a stepper motor driven *x*-*y* translation stage. When the specimen is far enough from the source (the distance was about 70mm for the experiment), by scanning the stage we can generate plane waves with different tilt angles at the specimen plane. The specimen was a microscope slide of a fern leaf and a standard doublet lens with a focal length of 30mm was positioned downstream at a distance of approximate 45mm from the specimen. At the back focal plane of the lens, a diaphragm with its diameter set to 2mm (amounts to an angle of 0.09rad) was placed. A CCD detector (2048×2048 pixels, each of 7.4μm^2) in the image plane (215mm from the diaphragm) was used to record the low-resolution images. The magnification of this imaging system was calibrated and it was about 8. A raster scan of 15×15 was carried out with a nominal step size of 1mm (equivalent to an angle of 0.02rad, resulting in a 78% overlap in the back focal plane) with ±20% random offsets. One example of the recorded low-resolution images is shown in Fig. 6.3a.

For the reconstruction, the recorded images were binned by a factor of 4 to 512×512 pixels and only a central region of 200×200 pixels were used to avoid the vignetting effect. Free space (1 everywhere) was used as the initial guess for the specimen and a Gaussian function as the aperture initial guess. Both the algorithms with sequential update (sequential ePIE) and parallel update (parallel ePIE) for the aperture reconstruction were run for 200 iterations. The reconstructions in Fourier space are shown in Figs. 6.3b and 6.3c respectively. As we can see from the size of the bright regions in the reconstructions, the specimen spectrum is expanded by about 3 times ((15-1)×(1-78%)=3.08) the size of the aperture in both *x* and *y* directions. Apart from the spectrum, the aperture function is also reconstructed. The polygonal shape implies a

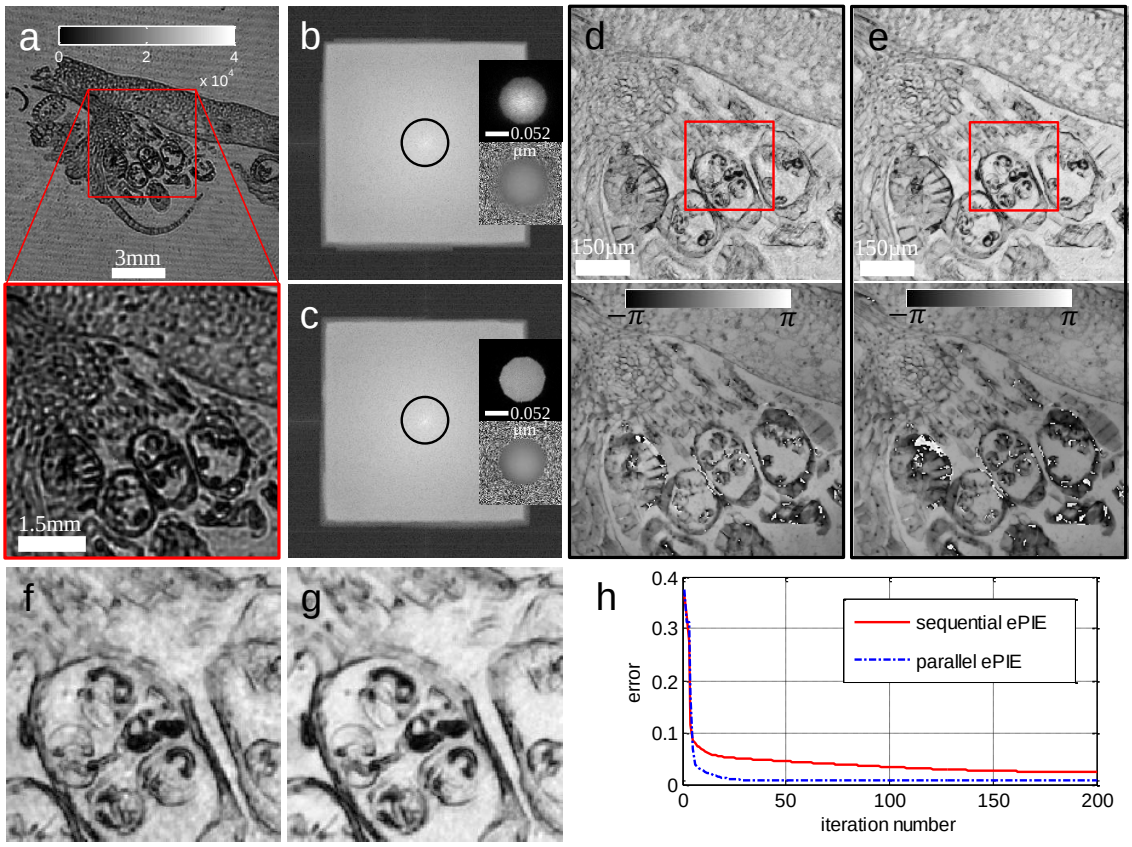

Figure 6.3: Comparison between the sequential ePIE algorithm and the parallel version using a visible light experiment. (a) One example of the recorded low-resolution images. (b) and (c) are the reconstructions of sequential ePIE and parallel ePIE respectively: the main figure is the specimen spectrum reconstruction; the top row of the inset is the modulus part of the aperture reconstruction and the bottom row is the phase part. The black circles in the spectrums indicate the size of the aperture. (d) and (e) are the specimen reconstructions obtained by respectively taking Fourier transform of (b) and (c). (f) and (g) are respectively the enlarged views for the highlighted regions in (d) and (e). (h) The image intensity error plots of the two reconstructions.

good agreement with the fact that we are using a diaphragm as the aperture. With close inspection of the aperture phase reconstructions, we can observe a small phase curvature that is accounting for the slight defocus error of the imaging system. Judging from the flatness and sharpness in the modulus parts of the aperture reconstructions, it is easy to tell that parallel ePIE outperforms the sequential ePIE. Moreover, when the spectra are Fourier transformed into real space, the comparison between them becomes obvious. From the enlarged view shown in Figs. 6.3f and 6.3g, it is not difficult to see that the modulus of the specimen reconstruction from the parallel ePIE has a much cleaner

background. Furthermore, the image intensity errors versus the iterations for both reconstructions are also plotted. Parallel ePIE reaches convergence almost at the 30^{th} iteration while sequential ePIE makes very slow progress and converges at the 160^{th} iteration still with a higher error.

Compare to the recorded raw image, the reconstructed image in real space has a much higher resolution: it is increased by about 19 times, given the pixel size is reduced from 29.6μm to 1.57μm that is calculated by $200\times 29.6/8\times 470$, where $200\times 29.6\mu m$ is the size of the recorded image, 8 is the magnification of the imaging system and 470 is the pixel number of the reconstructed image. However, it should be noted that the actual pixel size is slightly bigger than 1.57μm because the dark frames in the spectrum reconstructions (see Figs. 6.3b and 6.3c) do not transmit any frequency information. If this is taken into account, the actual pixel size is about 1.72μm, still much smaller than that of the raw image. In addition, the phase of the specimen, which was lost during the recording process, is also recovered.

6.2 3D Fourier ptychography

Fourier ptychography only works when the specimen is thin enough, so that the spectrum at the back focal will be a constant function with different lateral shifts during the tilting of the plane wave illumination. However, when the object is thick, the spectrum will be changing rather than just shifting. An easy way to think of this is that when the illumination is tilted the projected image of a thick object will changes. The Ewald sphere constructions can be used here to describe how information is expressed in reciprocal space for Fourier ptychography, as shown in Fig. 6.4. Tilting the illumination in reciprocal space is equivalent to rocking the Ewald sphere. However, the lens in Fourier ptychography always lets through the same range of angles of the

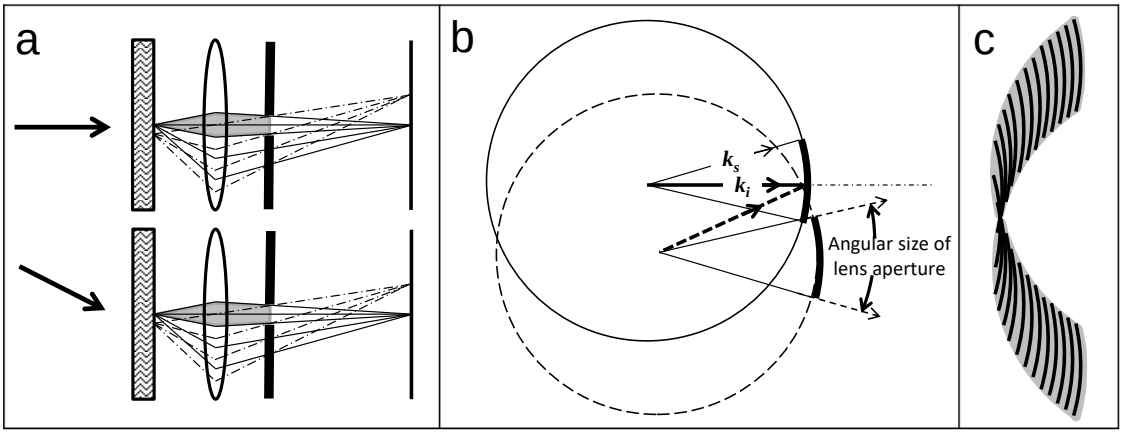

Figure 6.4: The Ewald sphere construction for tilt-series configuration. (a) Illustration of the experimental setup for the tilt-series configuration. (b) The data representation on the Ewald sphere of tilt-series configuration. k_i is the k-vector of an incident plane wave along the optic axis. Allowed scattering vectors, k_s, must lie on the Ewald sphere (seen in cross-section, solid circle) associated with k_i. For a tilted incident k-vector, the dotted Ewald sphere applies. The cone of scattered wave vectors admitted by the object lens remains constant, picking out a series of spherical caps in 3D reciprocal space. (c) Total volume in reciprocal space (seen in cross-section) spanned by the tilt-series configuration.

scattered beams, which is related to each image recorded on the detector by a Fourier transform. Since the Ewald sphere is different under different illumination angle, each recorded image corresponds to a different section of reciprocal space. In this case, conventional ptychographic reconstruction algorithms, which depend on the constancy of the object function whether in real space or in reciprocal space, cannot be used.

If we keep the illumination angle constant, the Ewald sphere will be fixed. Scanning the aperture across the back focal plane then allows us to explore different section of the fixed Ewald sphere, as shown in Fig. 6.5. In this way, the conventional ptychographic reconstruction algorithms become applicable now, but this method solves for the full scattering field of the object [138].This is the same information that is admitted by the objective of a conventional microscope. Although the wave can then be computationally propagated to different layers within the object, like going through focus in a conventional microscope there will always be Fresnel-type artefacts arising from other

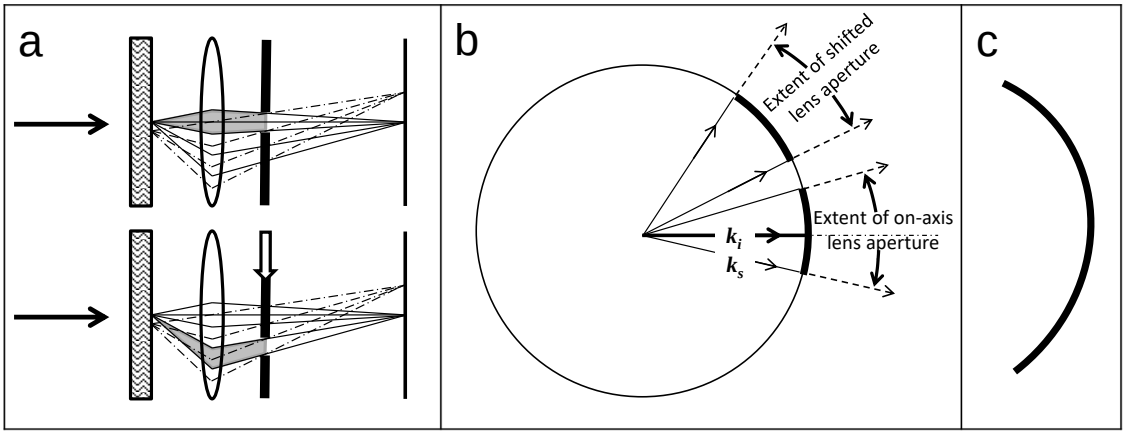

Figure 6.5: The Ewald sphere construction for aperture-scanning configuration. (a) Illustration of the experimental setup for aperture-scanning configuration. (b) The data representation on the Ewald sphere of aperture-scanning configuration. The method solves for the information over the surface of the Ewald sphere: the same information that is expressed in a conventional transmission microscope. (c) Total volume in reciprocal space (seen in cross-section) spanned by the aperture-scanning configuration.

out-of-focus planes of the object, especially given that the illumination needs to be coherent for ptychography to work. A further disadvantage of this approach is that the experimental elegance of the tilted illumination configuration is lost: a mechanical stepping motor is required to move the filter, unlike an electronically controlled illumination array. We refer to this approach as the aperture-scanning method.

6.2.1 The multislice method in real-space ptychography

Real-space ptychography has a similar problem with the specimen thickness. During the formation of the specimen exit wave, a multiplicative approximation [26, 57] – when an illumination propagates through the specimen the exit wave at the instant plane after the specimen can be factorised as the multiplication of the illumination function and the transmission function of the specimen – is used. When the specimen is too thick, the approximation breaks down. As a consequence, the conventional ePIE algorithm cannot give right reconstructions anymore [67]. The multislice method, widely used in electron microscopy to account for the multiple elastic scattering effects [91], can circumvent the

problem by sectioning the thick specimen into a set of thin slices, as shown in Fig. 6.6 (reproduced from the original 3D real-space ptychography work by Maiden et al [67]).

In the forward calculation, as shown in Fig. 6.6a, the exit wave of a thick specimen can no longer be estimated by a simple multiplication of the incident illumination and the object transmission function. Instead, we section the thick specimen into a set of thin slices, so that the multiplication approximation is valid for each thin slice. Therefore, the exit wave of each thin slice can be calculated by the multiplication of its incident wave function and its transmission function. The resulting exit wave is then propagated to the next slice to be the incident wave of that slice. This process is repeated until the last slice is reached and its exit wave (i.e. the exit wave of the thick specimen) is propagated to the detector plane to give the diffraction pattern in the detector plane.

In the update calculation, as shown in Fig. 6.6b, the modulus constraint is first applied to the diffraction pattern in the detector plane and backpropagation of the accordingly corrected diffraction pattern gives the updated exit wave of the last slice. Then the conventional ePIE update can be used to reconstruct the current slice and its incident wave, and backpropagation of the updated incident wave to the slice upstream gives the

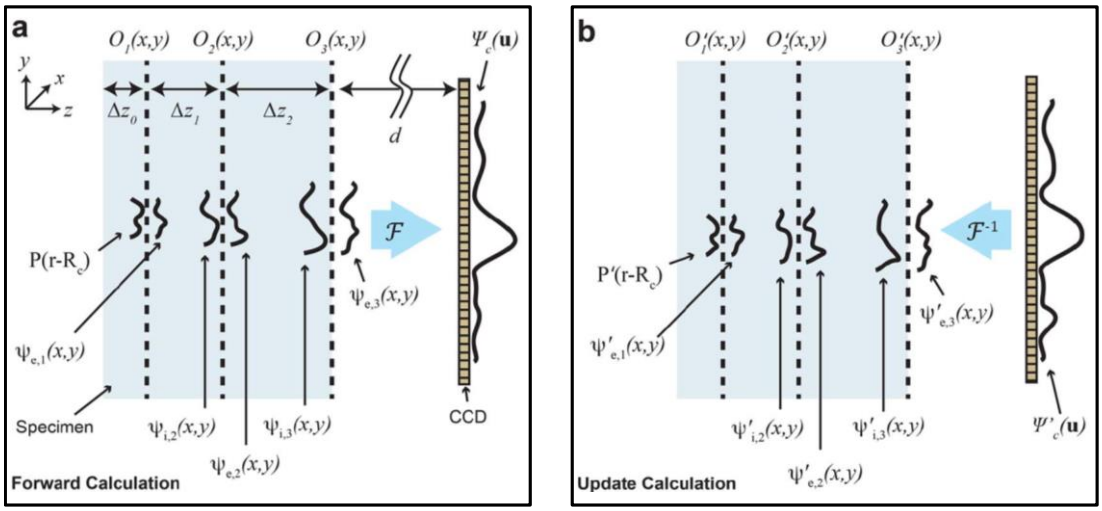

Figure 6.6: The schematic demonstration of real-space multislice ptychography. Reproduced from [67].

updated exit wave of that slice. This reverse process is also repeated until all the slices are updated. In this way, real-space ptychography is able to break the thickness limit and provide 3D reconstruction of the specimen [67].

6.2.2 The reconstruction algorithm of 3D Fourier ptychography

Actually, the thickness problem in Fourier ptychography can also be considered as the breakdown of the multiplication approximation. When the specimen is thick, the exit wave of the specimen is no longer equal to the product of the plane wave illumination and the specimen. In this case, the multislice method can be adopted to estimate the exit wave of the specimen, just like real-space ptychography. However, this is not compatible with the ePIE algorithm described in section 6.1.1, because only the aperture plane and the image plane are included in that algorithm. To use the multislice method, the specimen plane needs to be involved in the reconstruction algorithm as well. Therefore, a new algorithm is developed to include all the three planes and the flow chart is shown in Fig. 6.7.

A tilted plane wave illuminates the specimen and the exit wave estimated by the multiplication of the illumination and the specimen. The imaging lens performs a

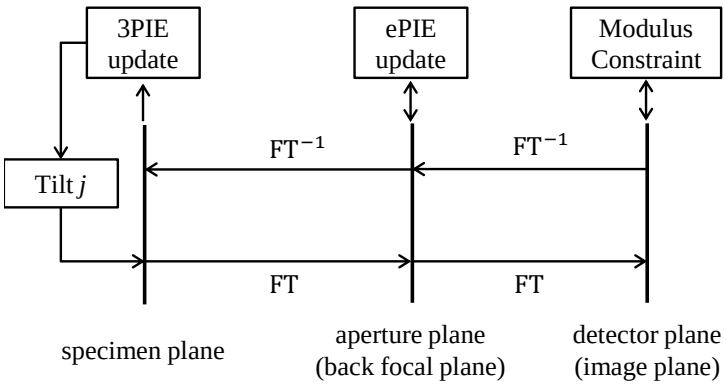

Figure 6.7: The flow chart of the reconstruction algorithm for multislice 3D Fourier ptychography. Compared to the 2D reconstruction algorithm, it extends to three planes, so that it takes into account of the specimen plane where we can apply the multislice method.

Fourier transform on the exit wave and transforms it into a spectrum in the back focal plane. The spectrum is then filtered by a small aperture and it is also estimated by a multiplication. The propagation from the filtered spectrum to the detector plane is modelled by another Fourier transform, so a low-resolution image of the specimen is formed and detected. When the specimen is thin, the Fourier transform from the specimen plane to the aperture plane converts the tilts of the plane wave illumination into lateral shifts of the spectrum. This is exactly the same as the algorithm described in section 6.1.1. However, when the specimen is thick, the multislice method can now be used to estimate the exit wave. In this way, the 3D information can be extracted, just like 3D real-space ptychography.

For simplicity, we use an object composed of two slices to demonstrate the algorithm, but extension to the cases with more than two slices is straightforward. With reference to Fig. 6.8, here we describe the algorithm step by step:

1. For the *n*th iteration, at tilt angle *j* the plane wave $P_{j,1}$ is incident on the first layer of the current object guess $O_{n,1}$ and the exit wave $\psi_{j,1}$ is calculated as

$$\psi_{j,1}(\boldsymbol{r}) = P_{j,1}(\boldsymbol{r}) O_{n,1}(\boldsymbol{r}). \tag{6.6}$$

2. The exit wave is then propagated over a distance of *z* to the second layer forming the

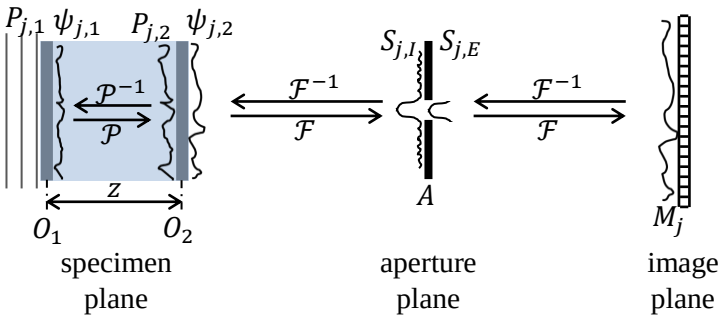

Figure 6.8: Schematic detailing the nomenclature used to describe the new 3D algorithm.

incident wave $P_{j,2}$

$$P_{j,2}(\boldsymbol{r}) = \mathcal{P}_z\left[\psi_{j,1}(\boldsymbol{r})\right]. \tag{6.7}$$

Because the distance z is normally a very small value, the propagation will be modelled by the angular spectrum method (see section 2.2.4).

3. The exit wave for the second layer is calculated as the product of the incident wave $P_{j,2}$ with the second layer transmission function $O_{n,2}$

$$\psi_{j,2}(\boldsymbol{r}) = P_{j,2}(\boldsymbol{r}) O_{n,2}(\boldsymbol{r}). \tag{6.8}$$

4. The spectrum $S_{j,I}$ at the back focal plane is related to the exit wave $\psi_{j,2}$ by a Fourier transform

$$S_{j,I}(\boldsymbol{q}) = \mathcal{F}\left[\psi_{j,2}(\boldsymbol{r})\right]. \tag{6.9}$$

5. The spectrum is then filtered by the current aperture guess A_n at the back focal plane

$$S_{j,E}(\boldsymbol{q}) = S_{j,I}(\boldsymbol{q}) A_n(\boldsymbol{q}). \tag{6.10}$$

6. The low-resolution image guess G_j at the detector plane is formed by taking a Fourier transform of the filtered spectrum

$$G_j(\boldsymbol{r}) = \mathcal{F}\left[S_{j,E}(\boldsymbol{q})\right]. \tag{6.11}$$

7. The modulus constraint is applied at the detector plane, i.e. replace the modulus of the low-resolution image guess by the square root of the recorded image intensity M_j and keep the phase unchanged. Then an inverse Fourier transform is performed to give the new filtered spectrum $S'_{j,E}$

$$S'_{j,E}(\boldsymbol{q}) = \mathcal{F}^{-1}\left[\sqrt{M_j(\boldsymbol{r})}\frac{G_j(\boldsymbol{r})}{\left|G_j(\boldsymbol{r})\right|}\right]. \tag{6.12}$$

8. The aperture function is reconstructed via the parallel ePIE update and the spectrum reconstructed via the sequential ePIE update

$$\left.\begin{aligned} A_{n+1}(\boldsymbol{q}) &= A_n(\boldsymbol{q}) + \alpha\frac{\sum_{j=1}^{J} S_{j,I}^{*}(\boldsymbol{q}+\boldsymbol{q}_j)\left[S'_{j,E}(\boldsymbol{q}) - S_{j,E}(\boldsymbol{q})\right]}{\left(\sum_{j=1}^{J}\left|S_{j,I}(\boldsymbol{q}+\boldsymbol{q}_j)\right|^2 + \epsilon\right)} \\ S'_{j,I}(\boldsymbol{q}) &= S_{j,I}(\boldsymbol{q}) + \alpha\frac{A_n^{*}(\boldsymbol{q}-\boldsymbol{q}_j)\left[S'_{j,E}(\boldsymbol{q}) - S_{j,E}(\boldsymbol{q})\right]}{\left(\left|A_n(\boldsymbol{q}-\boldsymbol{q}_j)\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{6.13}$$

9. The reconstructed spectrum is back propagated to the specimen plane by an inverse Fourier transform

$$\psi'_{j,2}(\boldsymbol{r}) = \mathcal{F}^{-1}\left[S'_{j,I}(\boldsymbol{q})\right]. \tag{6.14}$$

10. Apply the conventional ePIE update to reconstruct the second layer and its incident wave

$$\left.\begin{aligned} O_{n+1,2}(\boldsymbol{r}) &= O_{n,2}(\boldsymbol{r}) + \alpha\frac{P_{j,2}^{*}(\boldsymbol{r})\left[\psi'_{j,2}(\boldsymbol{r}) - \psi_{j,2}(\boldsymbol{r})\right]}{\left(\left|P_{j,2}(\boldsymbol{r})\right|^2\right)_{\max}} \\ P'_{j,2}(\boldsymbol{r}) &= P_{j,2}(\boldsymbol{r}) + \alpha\frac{O_{n,2}^{*}(\boldsymbol{r})\left[\psi'_{j,2}(\boldsymbol{r}) - \psi_{j,2}(\boldsymbol{r})\right]}{\left(\left|O_{n,2}(\boldsymbol{r})\right|^2\right)_{\max}} \end{aligned}\right\}. \tag{6.15}$$

11. The reconstructed incident wave for the second layer is back propagated to the first layer to form the updated exit wave $\psi'_{j,1}$ for the first layer

$$\psi'_{j,1}(\boldsymbol{r}) = \mathcal{P}_z^{-1}\left[P'_{j,2}(\boldsymbol{r})\right]. \tag{6.16}$$

12. The illumination for the first layer is the known plane wave, so we only need to reconstruct the first layer using the ePIE update

$$O_{n+1,1}(\boldsymbol{r}) = O_{n,1}(\boldsymbol{r}) + \alpha \frac{P^*_{j,1}(\boldsymbol{r})\left[\psi'_{j,1}(\boldsymbol{r}) - \psi_{j,1}(\boldsymbol{r})\right]}{\left(\left|P_{j,1}(\boldsymbol{r})\right|^2\right)_{\max}}. \quad (6.17)$$

Here $\boldsymbol{r}$ and $\boldsymbol{q}$ represent the real-space and reciprocal-space coordinates respectively. The constant α in the update functions can be altered to adjust the feedback of the update and it is normally assigned a value between 0 and 1. The forward calculation is from step 1 to step 6. Except the different incident illuminations (in Fourier ptychography it is a set of plane waves with different tilt angles, while in real-space ptychography it is a localised function with different lateral offsets), the forward calculation of the multislice method is basically the same for real-space ptychography and Fourier ptychography. However, for real-space ptychography the lateral scan is already performed in the first three procedures, Fourier ptychography needs further two procedures, step 4 and step 5, to accomplish the scan. The propagation in step 6 is again the same as in real-space ptychography, to calculate the wave field in the detector plane, except that the wave field is an image in Fourier ptychography but a diffraction pattern in real-space ptychography. The update calculation is from step 7 to step 13, which is basically to reverse all the steps in the forward calculation. The outputs of this algorithm are the 3D function of the specimen and the aperture function in the back focal plane. These two reconstructions are in difference spaces. In real-space ptychography, the outputs are the 3D function of the specimen and the localised illumination function and both two functions are in real space.

6.2.3 Optical experiments

A proof-of-principle experiment was implemented here using a specimen made up by two microscope slides (Hydrilla stem tip and dicotyledon leaf) to test the algorithm. The separation between the two slides was about 1mm. The same setup and scan parameters

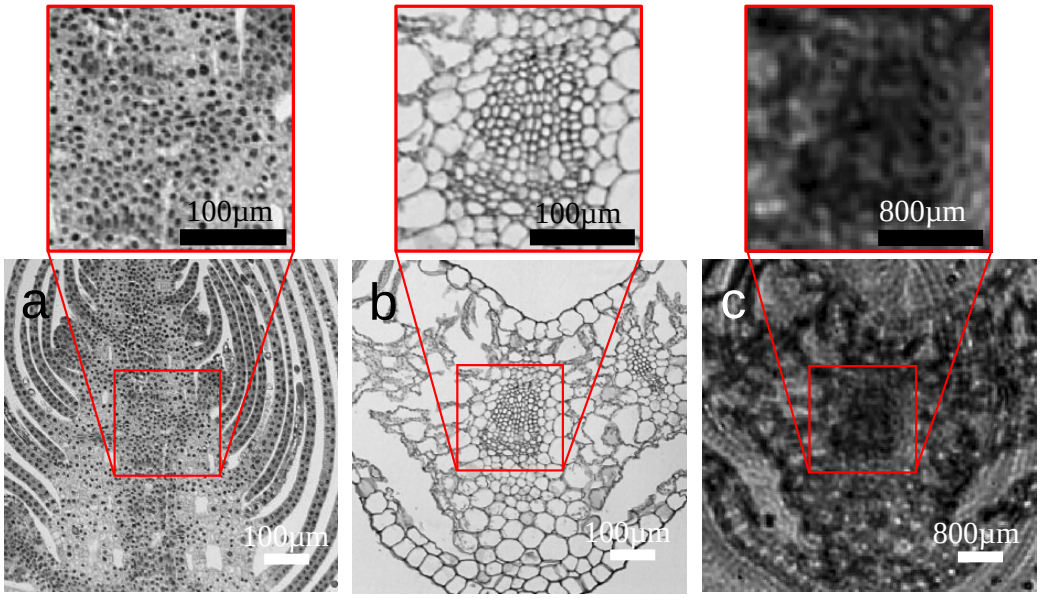

Figure 6.9: The microscopic images (taken with a 10× (0.25 NA) objective lens) of the two slides, (a) and (b), that are made up the thick specimen used for the optical experiment and (c) one example of the recorded low-resolution images.

as used in section 6.1.2 were used here. One of the recorded images is shown in Fig. 6.9 together with the microscopic images of the two slides.

First, we tried a reconstruction assuming single slice in the specimen and the results are shown in Figs. 6.10a (specimen reconstruction) and 6.10c (aperture reconstruction). The specimen reconstruction seems like it is focused on the second layer with out-of-focus features from the first layer. However, it is not a right reconstruction because the algorithm is assuming a constant spectrum at the back focal plane is applicable here. This is clearly shown in the reconstruction of the aperture that is obviously very bad. Then we tried the modified algorithm assuming two slices with a separation of 1mm in the specimen on the dataset. The reconstructions are shown in Figs. 6.10b (the specimen) and 6.10d (the aperture). The algorithm successfully separated the two slides, producing reconstructions with both modulus and phase images and comparable resolution with the microscopic images taken with a 10× objective lens (0.25 NA) for each slide. It is easy to tell from the aperture reconstruction that the modified algorithm improves the

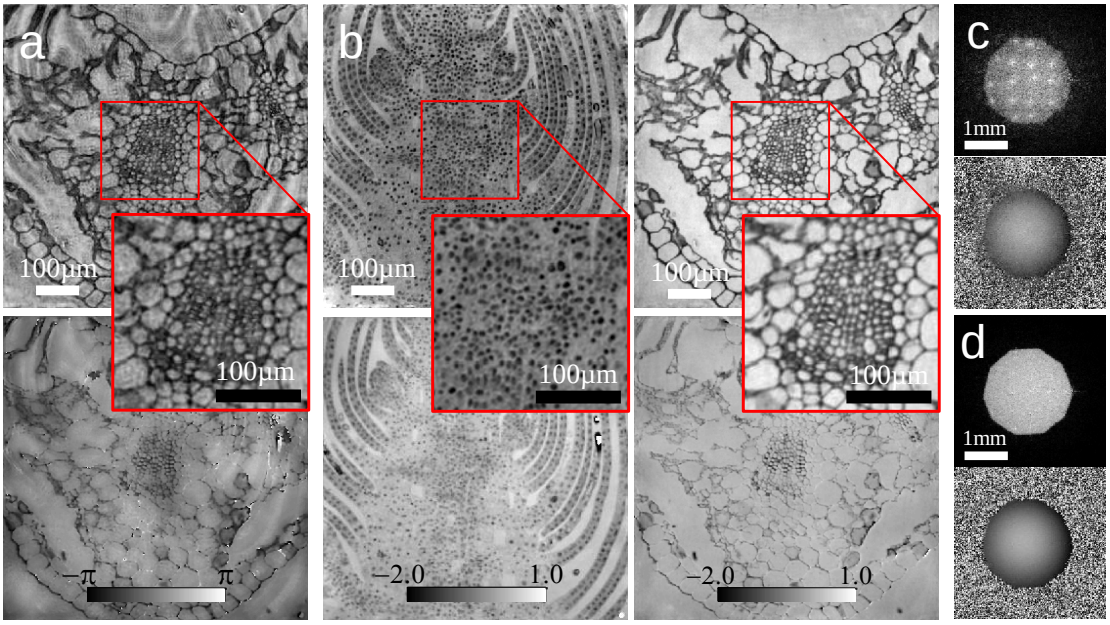

Figure 6.10: The reconstructions of 3D Fourier ptychography. (a) The specimen reconstructions and (c) the aperture reconstructions without taking account of 3D effects using the multislice method. (b) The specimen reconstructions of the first slice (left column) and the second slice (the right column) and (d) the aperture reconstructions with taking account of 3D effects using the multislice method. For all the figures, the top row is the modulus part and the bottom row is phase part.

reconstructions by a substantial amount. Again, the small phase curvature in the aperture reconstruction indicates a slight defocus error in the experimental setup.

To compare it with the aperture-scanning method, here we carried out another experiment with the illumination angle fixed but scanning the aperture at the back focal plane. Due to the limited space in the setup, unfortunately, the diaphragm will cause interference with other components during the scan. We therefore changed it to a 2mm circular aperture. This will not affect the comparison with the tilt illumination experiment, because only the size of the aperture (not the shape) matters. We also adjusted the step size (from 1mm to 435μm, so the overlap for both experiments is about 78%) to match the lateral shift of the spectrum in tilt illumination experiment. The rest of the parameters remained the same.

In this configuration, the function lying in back focal plane is constant during the scan, so the conventional algorithm (described in section 6.1.1) is applicable on this dataset now. The reconstructions are shown in Fig. 6.11a. Among them, the specimen reconstruction is actually the exit wave (not the transmission function) of the 3D specimen. Since the whole complex field is retrieved, we can propagate this exit wave to each slice to obtain the images of the two slides, as shown in Figs. 6.11b and 6.11c. As we can see, the fine structures are disrupted by the out-of-focus features from the other slice. Moreover, the aperture reconstruction cannot account for the defocus error as indicated by its flat phase reconstruction. This is because in this aperture-scanning configuration the defocus error stays fixed with the scattering field rather than with the aperture (as in the tilt-series configuration) during the scan.

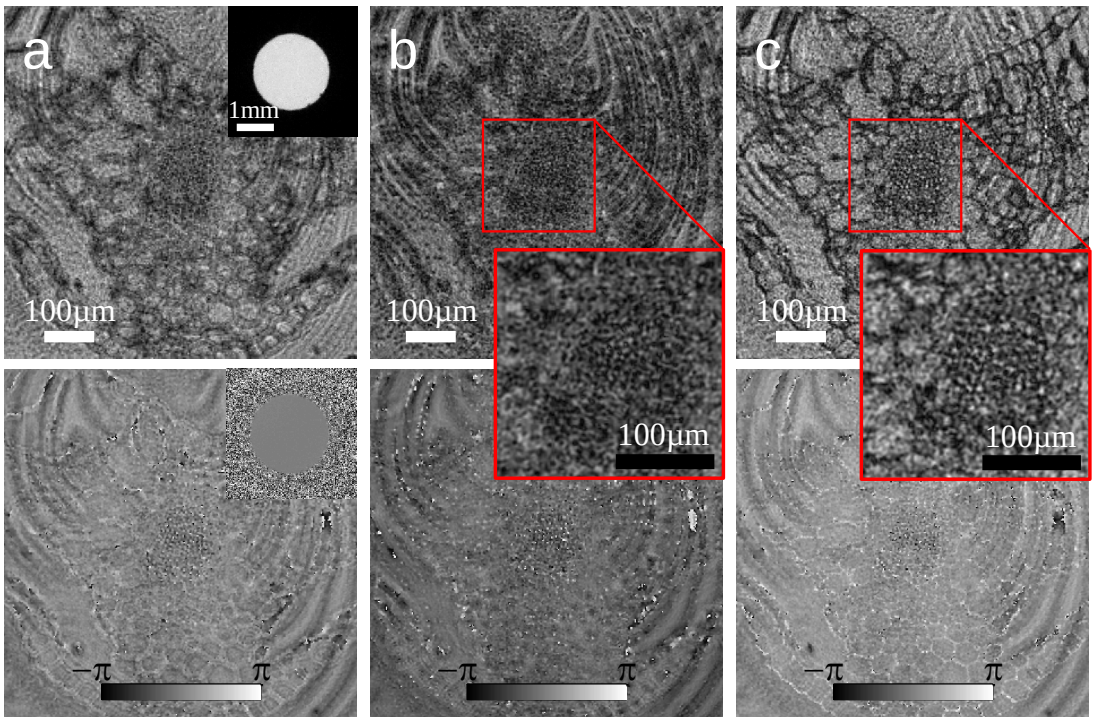

Figure 6.11: The reconstructions of aperture scanning method. (a) The raw outputs of the 2D reconstruction algorithm (described in section 6.1.1). The inset in the top right corner is the aperture reconstruction. (b) The first layer reconstructions with the raw outputs propagated to the first slice. (c) The second layer reconstructions with the raw outputs propagated to the second slice. For all the figures, the top row is the modulus part and the bottom row is phase part.

6.2.4 Reconstruction resolution

As we define the resolution of a 2D reconstruction (see Eq. (2.51)), the transverse resolution for a 3D reconstruction is defined to be the inverse of the frequency bandwidth in the plane perpendicular to the propagation direction. Likewise, the longitudinal (depth) resolution is defined to be the inverse of the frequency bandwidth along the propagation direction, as discussed by Chapman et al [42] in iterative CDI, and by Takahashi et al [69] and Tian et al [139] in multislice ptychography via the Ewald sphere constructions. In other words, a frequency bandwidth of B in reciprocal space is able to resolve a feature size of $1/B$ in real space. Therefore, to acquire the knowledge of reconstruction resolution we need to know the frequency bandwidth in the corresponding direction for 3D Fourier ptychography.

Here we redraw the Ewald spheres for 3D Fourier ptychography as shown in Fig. 6.12. The radii of the Ewald spheres are $1/\lambda$, with λ being the wavelength of the radiation. θ_I denotes the half maximum tilt angle of the plane wave illumination and θ_A the half maximum angle subtended by the aperture at the back focal plane. $\Delta q_{x,y}$ represents the frequency bandwidth in the transverse direction (e.g. x and y) and Δq_z the frequency bandwidth in longitudinal direction (e.g. z). If we define the numerical aperture of the tilt illumination as $\mathrm{NA}_I = \sin\theta_I$ and that of the aperture as $\mathrm{NA}_A = \sin\theta_A$, the two bandwidths $\Delta q_{x,y}$ and Δq_z can be calculated by

$$\left.\begin{aligned} \Delta q_{x,y} &= \frac{2\left(\mathrm{NA}_I + \mathrm{NA}_A\right)}{\lambda} \\ \Delta q_z &= \frac{2 - \sqrt{1-\mathrm{NA}_I{}^2} - \sqrt{1-\mathrm{NA}_A{}^2}}{\lambda} \end{aligned}\right\}. \tag{6.18}$$

So the corresponding resolutions are

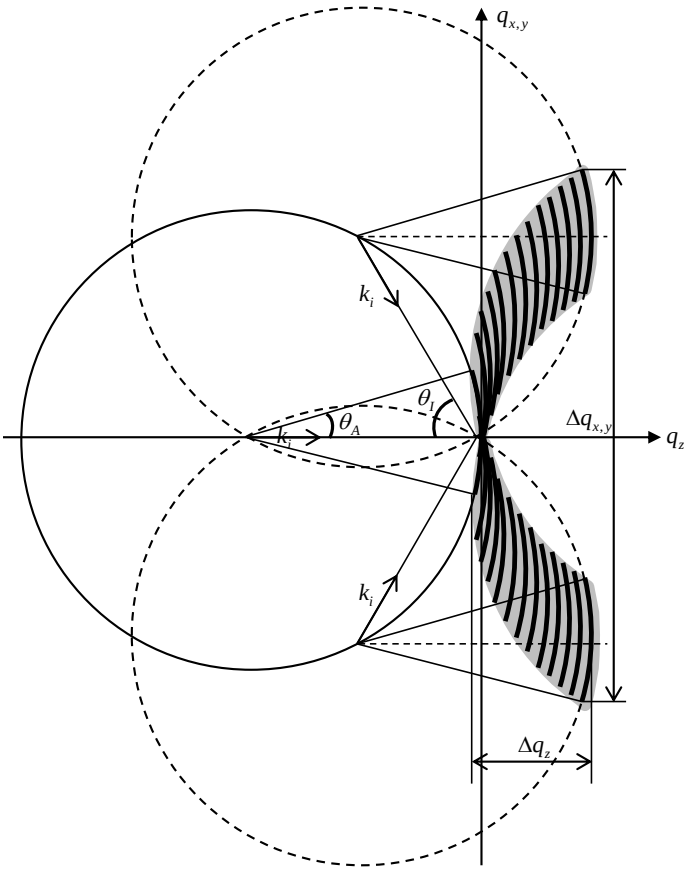

Figure 6.12: The Ewald sphere construction for Fourier ptychography. The shaded area is the total volume of the accessible region in reciprocal space (seen in cross-section). The radii of the Ewald spheres are 1/λ, with λ being the wavelength of the radiation. θ_I denotes the half maximum tilt angle of the plane wave illumination and θ_A the half maximum angle subtended by the aperture at the back focal plane. $\Delta q_{x,y}$ represents the frequency bandwidth in the transverse direction and Δq_z the frequency bandwidth in longitudinal direction.

$$\left.\begin{aligned} \Delta x = \Delta y &= \frac{\lambda}{2\left(\mathrm{NA}_I + \mathrm{NA}_A\right)} \\ \Delta z &= \frac{\lambda}{2 - \sqrt{1-\mathrm{NA}_I{}^2} - \sqrt{1-\mathrm{NA}_A{}^2}} \end{aligned}\right\}. \tag{6.19}$$

The transverse resolution (Δx and Δy) for Fourier ptychography is equivalent to a conventional microscope having an objective with an numerical aperture of $\mathrm{NA}_I + \mathrm{NA}_A$. However, the longitudinal resolution (Δz) for Fourier ptychography is not equivalent to that conventional microscope whose longitudinal resolution would be $\lambda \Big/ 1 - \sqrt{1-\left(\mathrm{NA}_I + \mathrm{NA}_A\right)^2}$, which is slightly smaller. As we mentioned in section 6.1.2,

the spectrum is expanded by 3.08 times the size of the aperture, which means $\mathrm{NA}_I = 3.08\mathrm{NA}_A$, and the transverse resolution is calculated to be 1.72μm. As a result, we have $\mathrm{NA}_A = 633\times10^{-9}/2\times4.08\times1.72\times10^{-6} = 0.045$ and $\mathrm{NA}_I = 3.08\mathrm{NA}_A = 0.139$. The achievable longitudinal resolution for this imaging system is therefore about 59μm. However, this calculation only sets a limit for the highest achievable longitudinal resolution, because for different transverse spatial frequencies the bandwidth in the longitudinal direction is different. This means different features in the specimen will have different longitudinal resolution. The features with lower spatial frequency will have smaller longitudinal bandwidth, hence having worse longitudinal resolution. The zero spatial frequency has zero longitudinal bandwidth, which leads to infinite longitudinal resolution. This means that a Fourier ptychography imaging system can never resolve a 3D object along its longitudinal axis if the object is uniform along its transverse direction.

To demonstrate the achievable longitudinal resolution for the experimental imaging system, a set of simulations with different distances between the two slices is carried out using the same parameters used in the optical experiments (to make sure the same NAs for the illumination and the aperture). The results are shown in Fig. 6.13. A very clear trend can be observed: the distance between the two slices increases the separation of them becomes cleaner in the reconstructions. The two slices cannot be distinguished until the distance reaches the theoretical resolution limit (i.e. 59μm), although each slice still has some leftover features from the other slice. As we increase the distance, the leftover features reduce. Upon close inspection, we could find that these leftover features are very smooth and they correspond to the low spatial frequency of the specimen spectrum. As we pointed out before, the low spatial frequency has a large longitudinal resolution. This means a long distance is necessary to ensure a clean

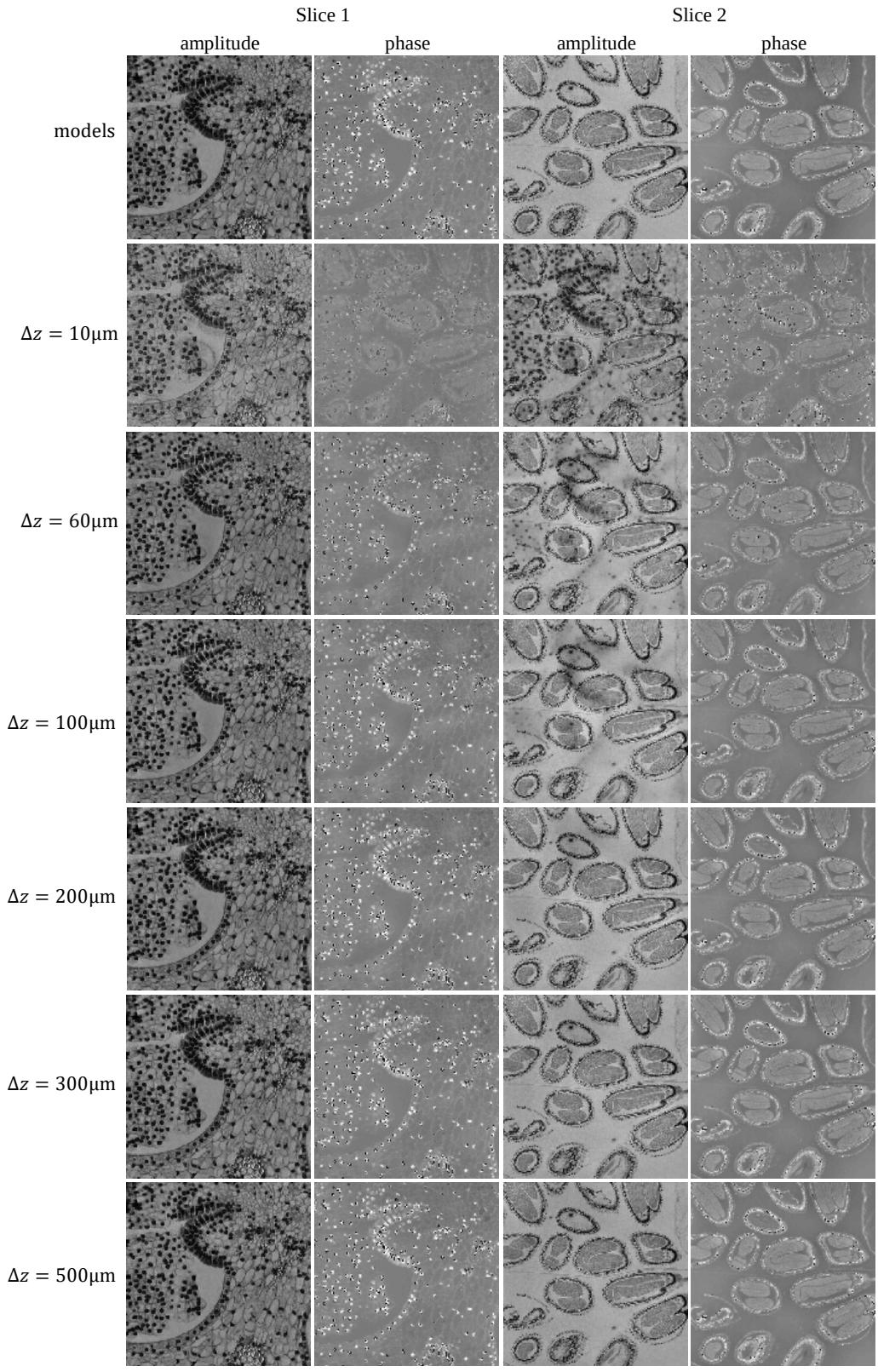

Figure 6.13: The simulation results using the same illumination NA and aperture NA used in the optical experiments. When the distance between the two slices increases the separation becomes cleaner in the reconstructions. A distance of at least 60μm is required to distinguish the two slices, although slightly cross talk between them still exists. Features with different spatial frequencies need different distances to ensure a clean separation.

separation of the two slices. For this particular simulation, 500μm is able to completely separate out the two slices.

To increase the resolution, high NAs are needed for both the illumination and the aperture. In the extreme case where the illumination is able to be tilted to a solid π angle and the aperture is fully opened up, the accessible volume has a shape of a doughnut with the cross-section corresponding to two discs of the diameter of the Ewald sphere touching at the origin, as shown in Fig. 6.14. As a result, a longitudinal resolution of half wavelength is achievable (because $\mathrm{NA}_I = \mathrm{NA}_A = 1$). Good volumetric imaging has been demonstrated for 3D real-space ptychography using the multislice method [68]. By reciprocity, the accessible volume in Fourier ptychography is the same as in real-space ptychography. Therefore, 3D Fourier ptychography could achieve similar performance if we manipulate a system with a very wide range of incident illumination angles.

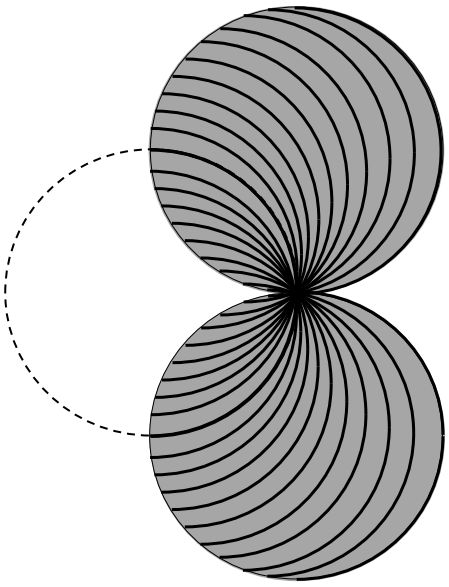

Figure 6.14: Total accessible volume in reciprocal space for Fourier ptychography when the illumination can be tilted to a π solid angle and the aperture in the back focal plane is fully opened up. In this extreme case, the illuminate NA and the aperture NA are both 1. As a result, the achievable transverse and longitudinal resolutions are respectively quarter and half wavelength.

6.3 Conclusion and discussion

In this chapter, we have explored Fourier ptychography in both 2D and 3D forms. Fourier ptychography is a Fourier domain version of real-space ptychography. Methodologically, the two forms of ptychography are the same. This also means they can share the same reconstruction algorithms. However, in Fourier ptychography the spectrum of the specimen and the aperture function are reconstructed, instead of the specimen and the illumination in real-space ptychography. Since the spectrum normally has a very bright spot in the centre and the search step size for the aperture update function in the ePIE algorithm is scaled to the brightest pixel, this dramatically diminishes the search step sizes for the pixels with small moduli, hence slowing the convergences for these pixels during the aperture reconstruction. A parallel update version can mitigate this problem by scaling the search step sizes with different values to make sure of relatively even convergences for all the pixels. Optical experiments have been carried out and verified the effectiveness of the modified ePIE algorithm.

Like real-space ptychography, Fourier ptychography also has an upper limit for the thickness of specimen. When exceeding this limit, the specimen spectrum will be changing rather than just shifting during the tilting of the plane wave illumination. As a consequence, the 2D reconstruction algorithm is no longer applicable. By directly scanning the aperture in the back focal plane, instead of tilting the illumination, one can enable the validity of applying the 2D reconstruction algorithm, because the spectrum function is now constant through the whole scanning process. However, in this way only the exit wave of the thick specimen is reconstructed. Although the reconstructed exit wave can be propagated to different axial positions within the specimen to bring a particular plane in focus, there will be out-of-focus artefacts arising from other planes.

Furthermore, the experimental elegance of the tilting illumination configuration (i.e. no physical scan involved) is lost. Here in this chapter, we have modified the 2D reconstruction algorithm to include the specimen plane, so that the multislice method can be used to account for the thickness problem, which would otherwise not be possible. Optical experiments have been performed and demonstrated the success of the modified 3D reconstruction algorithm compared to the aperture scanning method.

The reconstruction resolution of Fourier ptychography has also been discussed. It is given as the inverse of the frequency bandwidth accessible in reciprocal space along the corresponding direction. The transverse resolution is equivalent to that of a conventional microscope with an objective of the sum of the illumination NA and the aperture NA. The longitudinal resolution depends on the spatial frequency. Features with lower spatial frequency normally will have a worse longitudinal resolution and uniform 3D space (that has zero spatial frequency) cannot be resolved via Fourier ptychography. Simulations based on the configuration of the optical experiments have been modelled to demonstrate the achievable longitudinal resolution of the imaging system and the results have agreed with the theoretical calculation and analysis.

7 Ptychographic Tomography

Ptychography is suitable for any wavelength, but it is particularly successful in the regime of x-rays. There are two main reasons: 1. Short wavelength leads to high resolution. Compared to visible light, x-rays, especially hard x-rays, have a wavelength more than thousand times shorter. 10nm resolution via x-ray ptychography has been achieved using 0.2nm hard x-rays by Holler et al [140]. 2. The simple configuration makes the experiments very easy to implement. A localised illumination, a two dimensional scanning stage and a detector are enough to produce a high-resolution image [57]. No high quality imaging lenses and no complicated optical setup are needed. Although electrons have shorter wavelength, the electron microscope has extremely complicated optical arrangements, especially the lenses in between the specimen and the detector that distort and drift the diffraction patterns, imposing big challenges on the experiments and reconstructions.

The quantitative phase signal recovery is a very important advance for x-ray ptychography, because for light elements in biological samples the absorption based imaging mechanism cannot provide enough contrast for us to observe the structures. Conversely, the phase shifts introduced by the samples are able to produce much better contrast [79, 107]. The linear relationship between the reconstructed phase signal and the physical parameter (the real part of the refractive index) of the samples allows us to quantitatively study the samples (e.g. the electron density [107, 141] or sample thickness [85]). It does not require the samples to be weakly scattering to retain the quantitative relationship like the Zernike phase contrast imaging does [142], and it provides much higher resolution than the grating interferometer methods do [143].

Tomography provides 3D reconstruction of a sample by using a series of 2D projection images taken at different sample orientations [144]. It is an important tool to detect and observe the internal structures of 3D samples non-invasively. Because of its high penetration power, x-rays are most frequently used for this technique. However, quite often the detected signal is attenuation-based (i.e. absorption contrast imaging) [145]. Ptychography provides a means to extract the phase signal at high resolution, so if we use it to acquire the 2D projection phase images for different sample orientations, tomography is able to reconstruct the 3D structure information of the sample at high resolution and with good contrast. This form of combination of these two techniques is referred to as ptychographic tomography. In this chapter, we will demonstrate and explore this technique with a synchrotron x-ray experiment using a micro-capillary filled with glass beads as the sample. The possibility of electron ptychographic tomography is also discussed based on the scale calculation with the x-ray experiment.

7.1 Theoretical basics of tomography

In this section, we will review some important theoretical background of tomography and the sampling requirement for ptychographic tomography.

7.1.1 The calculation coordinates

In practice, a tomographic experiment is normally implemented by rotating the sample relative to the radiation and detector as shown in Fig. 7.1a. However, during the calculation it is easier to fix the image (i.e. sample) while rotating the radiation and the detector in the opposite direction as shown in Fig. 7.1b. The two systems are equivalent. Here is this chapter all the calculations are applying the coordinates shown in Fig. 7.1b. In other words, when we say at the rotation angle of θ, it means the radiation and the detector are rotated anti-clockwise by an angle θ, but the sample is fixed.

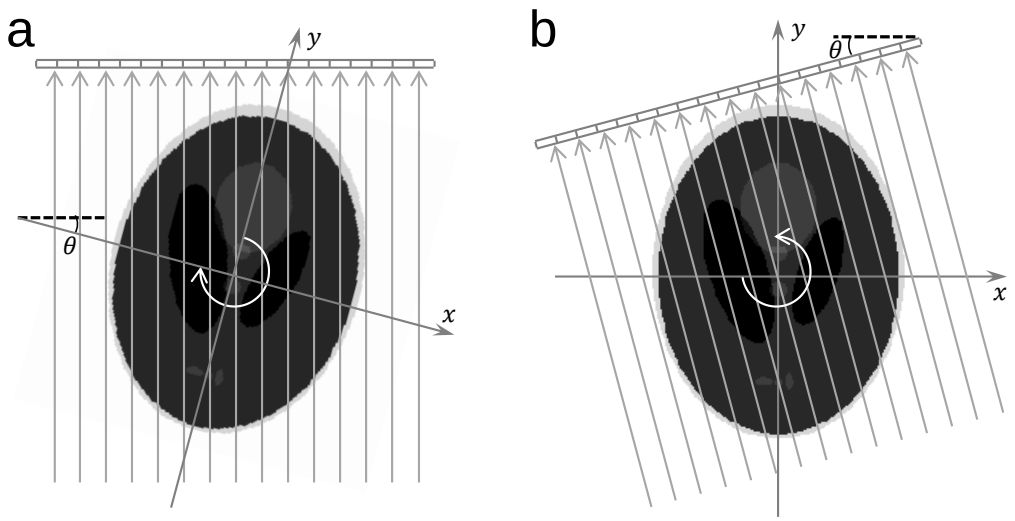

Figure 7.1: Coordinate systems in tomography. (a) Experimental coordinate system. During the data acquisition, the radiation source and the detector are fixed, while the specimen is rotated clockwise to different angles. (b) Calculation coordinate system. During the reconstruction, the image of the sample is treated still while the radiation source and the detector are rotated anti-clockwise. It is equivalent to the experimental coordinate system, but much easier for the calculation.

7.1.2 The Radon transform

In mathematics, the Radon transform is an integral calculation along straight lines, i.e. a projection along the direction parallel to the lines [146, 147]. As shown in Fig. 7.2, if we use coordinate s to denote the detector pixel position, the parallel lines at the rotational angle of θ_i are given by

$$x\cos\theta_i + y\sin\theta_i - s = 0. \tag{7.1}$$

If we use Dirac delta function δ to represent a point, then the lines can also be represented by $\delta\left(x\cos\theta_i + y\sin\theta_i - s\right)$. Given a 2D function $f(x,y)$, its Radon transform along these parallel lines is defined as

$$p(s,\theta_i) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta_i + y\sin\theta_i - s)\mathrm{d}x\mathrm{d}y. \tag{7.2}$$

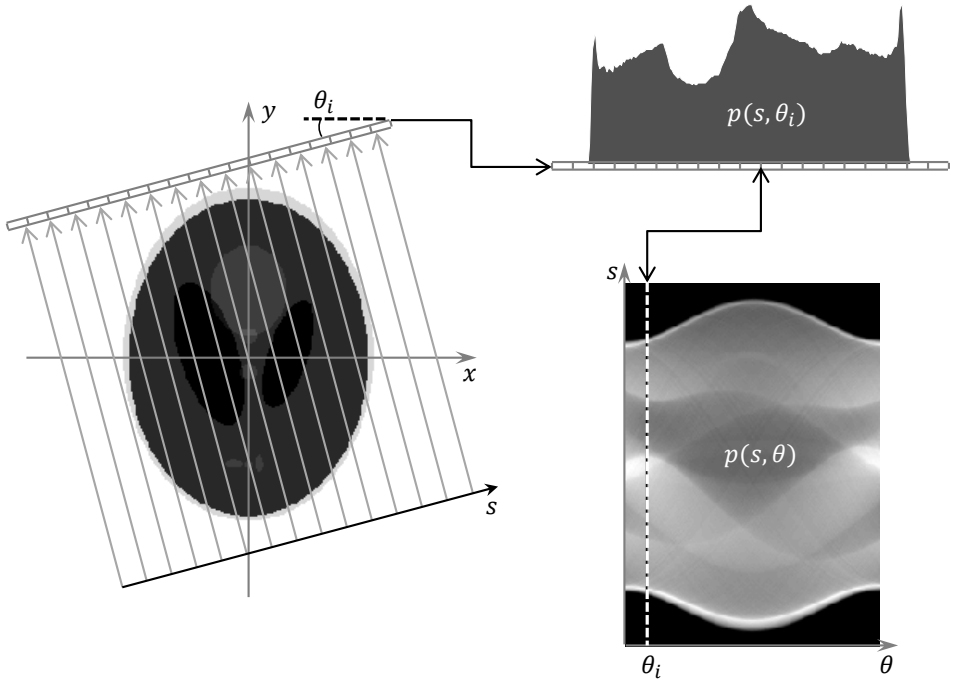

Figure 7.2: Schematic representation of the Radon transform. It is mathematically expressed as the line integrations along the direction of the propagation of the radiation source. These line integrations are called the projection of the sample and it is a 1D function. For a 2D sample, its projection at a specific orientation will be a line. Assembling these lines in the order of the sample orientation angles, the 2D diagram we get is called sinogram.

This is basically the projection of $f(x,y)$ at rotational angle of θ_i. If we rotate the sample from 0° to 180° at an angular step of $\Delta\theta$ and measure the projection at each angle, assembling these projections gives the projection dataset $p(s,\theta)$ as shown in Fig. 7.2.

The projection dataset $p(s,\theta)$ is also called a sinogram [144], because for a specific point (x_i, y_i) in the sample its projection position s_i on the detector is related to the projection angle θ by a sine function

$$s_i = x_i \cos\theta + y_i \sin\theta. \tag{7.3}$$

As we can see from Fig 7.2, the projections of the small features, especially the bright edges of the sample, appear graphically as a number of blurred sine waves with different amplitudes and phases.

7.1.3 Fourier slice theorem

If the 2D Fourier transform of the sample $f(x,y)$ is given as $F(u,v)$ and the 1D Fourier transform of $p(s,\theta_i)$ is $P(w,\theta_i)$, the Fourier slice theorem states that $P(w,\theta_i)$ equals a slice of $F(u,v)$ going through its origin at the same angle of θ_i [144]. According to the definition, the 1D Fourier transform of $p(s,\theta_i)$ is given by

$$P(w,\theta_i)=\int_{-\infty}^{\infty} p(s,\theta_i)e^{-i2\pi sw}\mathrm{d}s. \tag{7.4}$$

Substitution of Eq. (7.2) into the above equation leads to

$$P(w,\theta_i)=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta_i+y\sin\theta_i-s)\mathrm{d}x\mathrm{d}y\right\}e^{-i2\pi sw}\mathrm{d}s. \tag{7.5}$$

Switching the integral order gives

$$P(w,\theta_i)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\left\{\int_{-\infty}^{\infty}\delta(x\cos\theta_i+y\sin\theta_i-s)e^{-i2\pi sw}\mathrm{d}s\right\}\mathrm{d}x\mathrm{d}y \tag{7.6}$$

Since δ is zero everywhere except at zero, we have

$$P(w,\theta_i)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-i2\pi\left(x\cos\theta_i+y\sin\theta_i\right)w}\mathrm{d}x\mathrm{d}y. \tag{7.7}$$

According to the definition of a 2D Fourier transform, the above equation can be re-written as

$$P\left(w,\theta\right)=F\left(u,v\right)\Big|_{u=w\cos\theta,v=w\sin\theta}. \tag{7.8}$$

If we use the polar coordinates w and θ to replace the Cartesian coordinates u and v (w and θ are related to u and v by $u=w\cos\theta$ and $v=w\sin\theta$), $F(u,v)$ can be written as

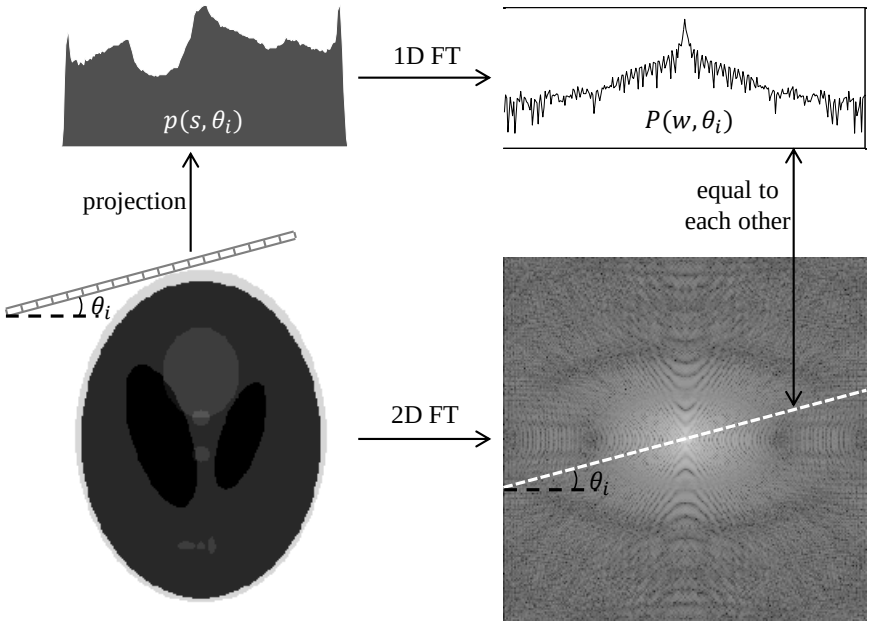

Figure 7.3: Schematic representation of Fourier slice theorem. For a 2D sample, it says the Fourier transform of the projection of the sample (the projection is 1D, so the Fourier transform is also 1D) at a specific orientation angle is equal to the slice of the 2D Fourier transform of the sample that goes through its origin at the same angle.

$F_P(w,\theta)$, where F_P denotes the Fourier transform represented in the polar coordinate system. Therefore the Fourier slice theorem is simply described by

$$P(w,\theta) = F_P(w,\theta). \tag{7.9}$$

A schematic representation of the Fourier slice theorem is shown in Fig. 7.3.

7.1.4 Image reconstruction via filtered back projection

Image reconstruction reverses the projection process, i.e. retrieving the image of the sample from the projection dataset. According to Fourier slice theorem, if we measure the projection data of the sample from 0° to 180° and perform the 1D Fourier transform on the projection dataset, we are able to fill in the 2D Fourier space of the sample slice by slice. To obtain the specimen, only an inverse Fourier transform is needed. However in this way the 2D Fourier transform of the specimen is represented in the polar coordinate system. To do the inverse Fourier transform using FFT, we need to convert

the polar coordinates to the Cartesian coordinates [148]. This conversion requires interpolations in Fourier space that will introduce substantial errors and result in serious artefacts in the reconstruction [149].

A widely-used well-known reconstruction algorithm called filtered back projection (FBP) is able to produce very satisfactory reconstructions [150]. If we adopt the same idea as above, the image of the sample $f(x,y)$ is reconstructed by performing an inverse Fourier transform on $F_P(w,\theta)$ in the polar coordinate system as

$$f(x,y)=\int_0^{2\pi}\int_0^{\infty} F_P(w,\theta)e^{i2\pi w(x\cos\theta+y\sin\theta)}w\mathrm{d}w\mathrm{d}\theta. \tag{7.10}$$

Since $F_P(w,\theta)=F_P(-w,\theta+\pi)$, Eq. (7.10) can be rewritten as

$$f(x,y)=\int_0^{\pi}\int_{-\infty}^{\infty} F_P(w,\theta)|w|e^{i2\pi w(x\cos\theta+y\sin\theta)}\mathrm{d}w\mathrm{d}\theta. \tag{7.11}$$

Substituting Eq. (7.9) into Eq. (7.11) leads to

$$f(x,y)=\int_0^{\pi}\int_{-\infty}^{\infty} P(w,\theta)|w|e^{i2\pi w(x\cos\theta+y\sin\theta)}\mathrm{d}w\mathrm{d}\theta. \tag{7.12}$$

Using the following substitutions

$$Q(w,\theta)=P(w,\theta)|w| \tag{7.13}$$

and

$$q(s,\theta)=\int_{-\infty}^{\infty} Q(w,\theta)e^{i2\pi ws}\mathrm{d}w, \tag{7.14}$$

we have

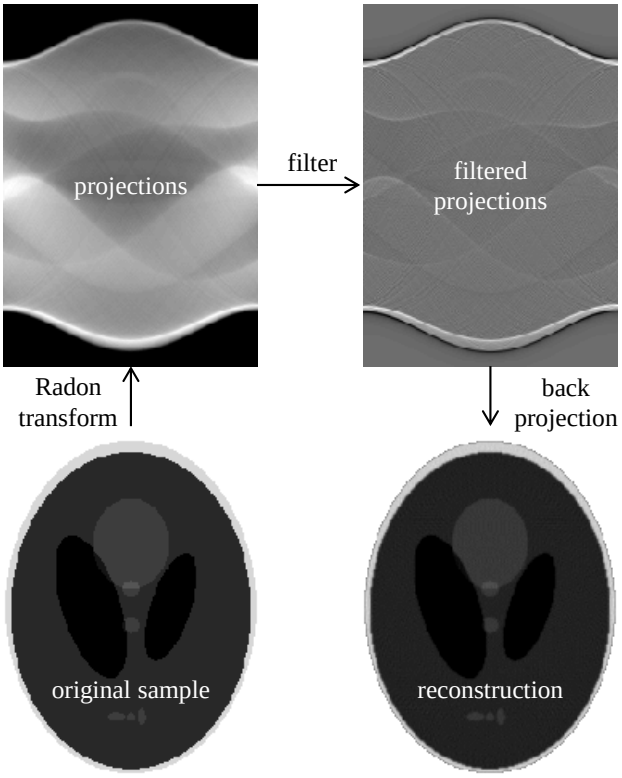

Figure 7.4: The flow chart of tomographic reconstruction via FBP algorithm. The data acquired from the experiments provides us the projections of the sample at different orientation angles. At each angle, the projection is first filtered along the detector index using the ramp filter. The filtered projections are then back projected to form the tomographic images of the sample. For a 3D sample, the above procedures need to be repeated for all the slices in the sample.

$$f(x,y)=\int_0^{\pi} q(s,\theta)\Big|_{s=x\cos\theta+y\sin\theta}\,\mathrm{d}\theta. \tag{7.15}$$

Here $q(s,\theta)$ is a filtered version of the projection dataset $p(s,\theta)$. The transfer function of the filter is $|w|$, which suppresses the oversampled low frequencies. Because of its shape, the filter is often called ramp filter [150]. Eq. (7.15) is basically a back projection process: smearing the filtered projections back over the paths along which the Radon transform is performed. Eq. (7.13) to Eq. (7.15) mathematically expresses the FBP algorithm and Fig. 7.4 shows the tomographic reconstruction process via the FBP algorithm.

7.1.5 Sampling requirement for ptychographic tomography

For tomography, the measurement of the projections is discrete and so is the Fourier space of the sample. So there are two types of sampling: the translational sampling for the projections and the angular sampling for the Fourier space. The translational sampling space Δx is the size of the pixel pitch of the detector where the projections are recorded, as shown in Fig. 7.5. According to Shannon sampling theorem, we have

$$\Delta x = \frac{1}{N\Delta k}, \tag{7.16}$$

where N is the number of sampling spaces and Δk is the sampling space in reciprocal space. In ptychographic tomography, the projections are the reconstructed images from ptychography, thus the translational sampling is the pixel size of the ptychographic images, which is given in Eq. (2.51). It is determined by the highest frequency that can be measured in the diffraction plane.

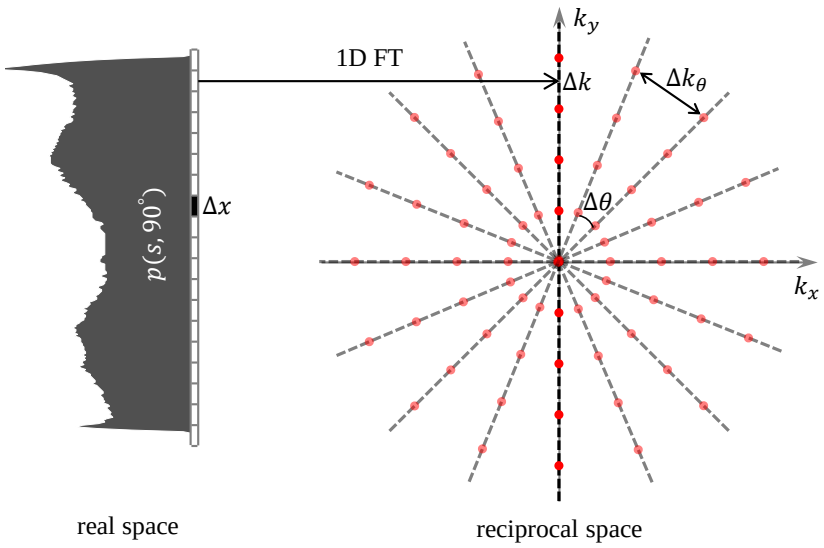

Figure 7.5: Sampling requirements of tomography. There are two samplings: translational sampling Δk for the recorded projections and angular sampling $\Delta\theta$ for the Fourier space of the sample. The resulting sampling space Δk_θ is proportional to the angular sampling $\Delta\theta$ and the highest frequency in the projection image (i.e. $1/\Delta x$). In the experiments, we need to match these two samplings, which requires $\Delta k = \Delta k_\theta$.

The angular sampling space $\Delta\theta$ for tomography is

$$\Delta\theta = \frac{\pi}{M}, \tag{7.17}$$

where M is the number of projection measurements over 180°. Corresponding to $\Delta\theta$, the sampling space Δk_θ in reciprocal space of the specimen is calculated by

$$\Delta k_\theta = \frac{N\Delta k}{2}\Delta\theta = \frac{\Delta\theta}{2\Delta x}. \tag{7.18}$$

It is proportional to the angular sampling space $\Delta\theta$ and the highest frequency of the projection image (i.e. $1/\Delta x$). To match the translational sampling and the angular sampling, we require $\Delta k_\theta = \Delta k$ and this leads to

$$M = \frac{\pi N}{2}. \tag{7.19}$$

For a given sample, if its diameter is D and the reconstruction pixel size is Δx, Eq. (7.19) can be written as

$$M = \frac{\pi D}{2\Delta x}. \tag{7.20}$$

This sets a minimum number of projections needs to be measured for a specific sample size D to be reconstructed at a specific resolution Δx [148].

7.2 X-ray ptychographic tomography experiment of glass beads

As a demonstration of ptychographic tomography, an x-ray experiment was performed on the I13 beamline at the Diamond Light Source in UK. The sample was a 42/52μm (inner/outer diameter) glass capillary filled with 1μm glass beads.

7.2.1 Data acquisition

The monochromated x-ray energy was selected at 9.12KeV ($\lambda = 0.136\text{nm}$). The setup is shown in Fig. 7.6a. A Fresnel zone plate (FZP) with NA of 4.6×10^{-4} was used to focus the beam. A central-stop (CS) was used to block the transmitted beams and an order-select-aperture (OSA) was used to choose the first order of the diffracted beams from the FZP. Upstream of the FZP, two sets of slits were adjusted to select a coherent region of the beam. Located at about 2.3mm upstream of the FZP focus, the sample was mounted on a translational piezo stage that itself was mounted on a rotational stage. Because of this arrangement, at each sample orientation the translational scan needs to be adjusted to compensate for the rotation angle. The single-photon counting 'Merlin' detector (composed of 2×2 arrays of MediPix3 chips each with 512×512 pixels of $55\mu\text{m}^2$) was positioned 7.284m downstream of the sample to measure the diffraction patterns. To resolve the 1μm glass beads in a 42μm glass tube (the thickness of the glass tube is not included), Eq. (7.20) gives a minimal number of projection measurements of 132 (the pixel size Δx has to be smaller than the radius of the glass beads). Here 180 sample orientations spaced by 1 degree ranging from -90 degrees to 90 degrees were

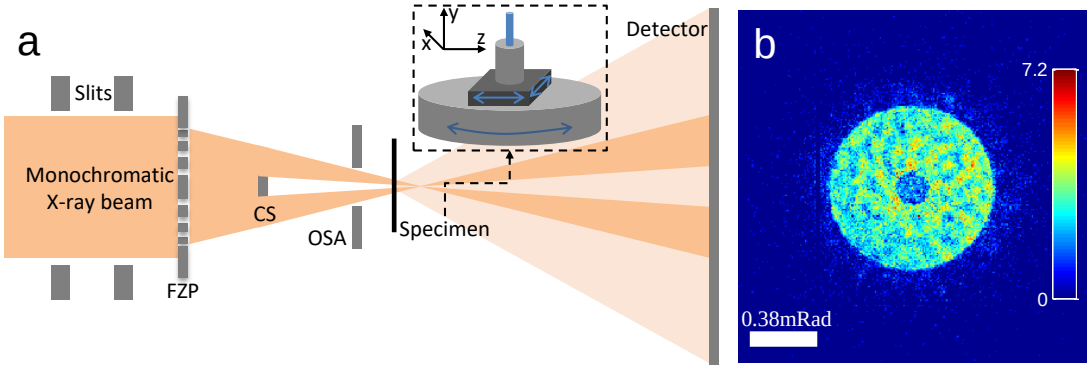

Figure 7.6: The x-ray ptychographic tomographic experiment. (a) The optical configuration of the experiment. The three dimensional scanning stage, including the coordinate system, is shown in the inset enclosed by the dashed rectangular. (b) An example of the diffraction intensity measurements (square rooted for display purpose).

measured. At each orientation, the sample was scanned over a 12×26 raster grid with a nominal step size of 2.5μm plus ±20% random offsets, covering a field of view of 30μm×65μm. At each scan position, the exposure time was 0.25s, giving average total counts of 1.1×10^5 per diffraction pattern. The whole measurements took about 12 hours. One example of the recorded diffraction patterns (cropped to a region of 256×256, where significant photons were detected) is shown in Fig. 7.6b.

7.2.2 Ptychographic reconstruction

Since the sample has a thickness of 42μm, prior to reconstruction, we need to make sure the projection approximation is well satisfied. Given the fact that the NAs of the FZP and the object are very small, according to the calculation in section 6.2.3 the maximal allowable thickness is approximately given by

$$\Delta z \approx \frac{2\lambda}{\mathrm{NA}_O^{\,2} + \mathrm{NA}_P^{\,2}} \tag{7.21}$$

where NA_O and NA_P are the numeric apertures of the object and the FZP respectively. The cropped diffraction pattern has an NA of $128\times55\times10^{-6}/7.284 = 9.7\times10^{-4}$ and its NA is equal to the addition of NA_O and NA_P, so $\mathrm{NA}_O = (9.7-4.6)\times10^{-4} = 5.1\times10^{-4}$. The maximal allowable thickness is 577μm, so the reconstruction is well modelled by the projection approximation.

At each angle θ, the ptychographic reconstruction of the object is the representation of the projected transmission function of the sample along the beam propagation direction, which can be expressed as [151]

$$t(x,y,\theta) = \exp\left[\frac{i2\pi}{\lambda}\int\left(n(\boldsymbol{r})-1\right)\mathrm{d}z\right], \tag{7.22}$$

where λ is the wavelength, *x*, *y* and *z* are the experimental coordinates (the orientation is indicated in Fig. 7.6a), $\boldsymbol{r}$ is the sample coordinates and is related to the experimental coordinates by $\boldsymbol{r} = (r_x, r_y, r_z) = (x\cos\theta - z\sin\theta, y, z\cos\theta + x\sin\theta)$, and $n = 1 - \alpha + i\beta$ is the complex refractive index of the sample. α and β are determined by physical properties of the material under investigation and the distributions of α and β can provide different contrast images of the sample: the imaginary part β causes attenuations of the incident x-rays, giving an absorption image of the sample; the real part α introduces phase shifts on the incident beam, forming a phase image of the sample. For hard x-rays, α is normally much bigger than β, especially for light elements. As a consequence, phase images always have much better contrast than absorption images in x-ray regime.

At each sample orientation, 400 iterations of ePIE were used to reconstruction the projected complex transmission function of the sample as well as the illumination function. Fig. 7.7 shows the ptychographic reconstructions when the sample is rotated to

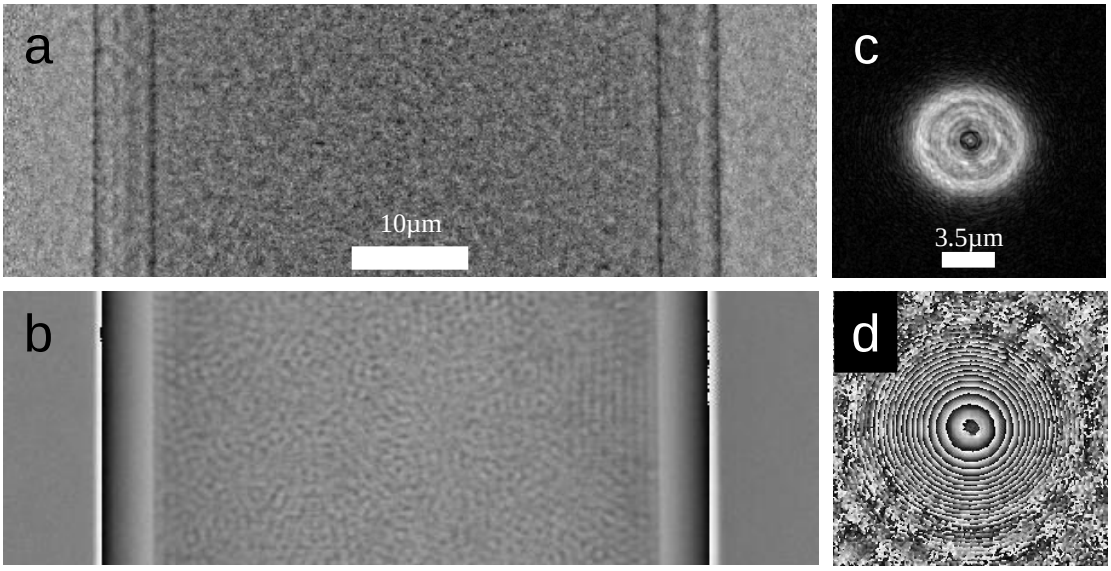

Figure 7.7: Ptychographic reconstructions at the rotational angle of -90 degree. (a) The modulus part of the reconstructed object. (b) The phase part of the reconstructed object. (c) The modulus part of the reconstructed probe. (d) The phase part of the reconstructed probe.

-90 degrees. As we can see, the phase part (determined by α) indeed has much better contrast than the modulus part (determined by β) and better signal-to-noise ratio as well. It is interesting to notice that at the edge of the tube some glass beads are crystallised. The discontinuities of the phase at the edge of the tube are caused by phase wrapping, which happens when the sample introduces a phase change that exceeds the range $[-\pi, \pi]$.

7.2.3 Tomographic reconstruction

Here for the tomographic reconstruction, we use the phase images

$$\phi(x, y, \theta) = \arg[t(x, y, \theta)] = -\frac{2\pi}{\lambda}\int \alpha(x\cos\theta - z\sin\theta, y, z\cos\theta + x\sin\theta)\mathrm{d}z. \quad (7.23)$$

As we discussed in section 2.3.8, the simultaneous reconstruction of the illumination and the specimen in ptychography has some inherent ambiguities, such as a phase ramp, phase offset and lateral shifts. These ambiguities need to be eliminated before the tomographic reconstruction [107].

1. Removal of phase ramp and phase offset

The ptychographic scan has covered the whole sample in *x* direction with some extra free space at each side (see Fig. 7.7b). The free space provides an easy way to remove the phase ramp and offset. First we cut out a *M*×*N* rectangular region from the free space reconstruction. Suppose the phase ramp and the phase offset are expressed as

$$\Phi(x, y) = \exp[i(ax + by + c)], \quad (7.24)$$

where *a* and *b* are the linear coefficients of the phase ramp along *x* and *y* directions respectively, *c* is the phase offset. Then performing a discrete Fourier transform (DFT) on the cut out region, we have

$$\mathrm{DFT}\left\{\exp\left[i\left(ax+by+c\right)\right]\right\}=\exp(ic)\delta(u-\frac{aM}{2\pi},v-\frac{bN}{2\pi}), \tag{7.25}$$

where δ is Dirac delta function, u and v are the reciprocal coordinates. So the linear coefficients of the phase ramp can be calculated from the shift of the brightest spot of the Fourier amplitude from the image origin. The phase offset c is directly given by the phase value of the brightest spot. A higher accuracy of the linear coefficients of the phase ramp can be obtained by using a DFT subpixel refinement method [152]. After obtaining a, b and c, we manually form an opposite phase $conj\left\{\Phi(x,y)\right\}$ and use it to multiply the reconstruction to remove the phase ramp and phase offset. Since these phase ambiguities can be different for each ptychographic reconstruction at different sample orientations, this step needs to be applied to all the ptychographic reconstructions.

2. Phase unwrapping

Phase unwrapping is to reverse the wrapping process and return the phase signal to a continuous form. The phase is well reconstructed here and has very few phase residues - inconsistent points in the unwrapped phase [153] - at the edges of the field of view. Since phase residues cause problems and difficulties for the phase unwrapping, here we cropped out a region with no residues as shown in Fig. 7.7b. With this region, the phase unwrapping becomes very easy. The unwrapping of a two dimensional phase image can be broken down into two one dimensional phase unwrapping (row and column).

Suppose $\phi_{i,j}$ represents the phase value located at row i and column j of a 2D phase image, the unwrapping of the ith row of this phase image can be described by the following steps:

1) Calculate the phase difference between two adjacent pixels via $E = \phi_{i,j+1} - \phi_{i,j}$ from $j = 1$;

2) If E is bigger than π when $j = m$, then $\phi_{i,j} = \phi_{i,j} - 2\pi$ for $j \geq m$;

3) If E is smaller than -π when $j = n$, then $\phi_{i,j} = \phi_{i,j} + 2\pi$ for $j \geq n$;

4) Repeat the above steps until j reaches to the last pixel of the ith row.

After going through all the rows, the phase unwrapping along the row direction is considered finished. Now we do the same implementation for the column direction. Then the phase unwrapping of a 2D image is completed. Fig. 7.8 shows the phase unwrapping results of the phase image shown in Fig. 7.7b. The top image is the result of applying phase unwrapping along column direction. Noticeable changes are the two vertical discontinuity lines, which become straight after phase unwrapping. The bottom image is the phase unwrapping result of the top image along the row direction. This is the phase image that can be used for the tomographic reconstruction.

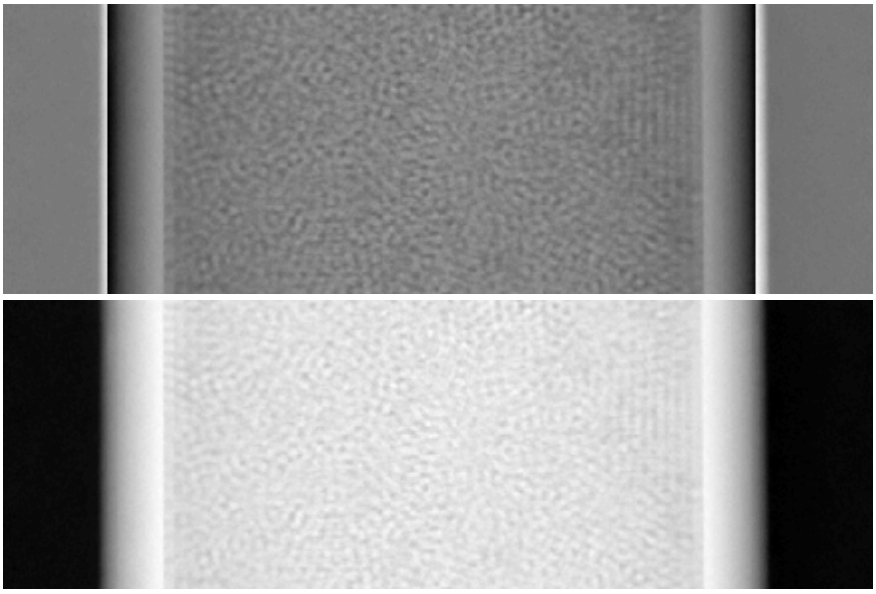

Figure 7.8: Phase unwrapped images of the object reconstruction. The top one is the result after applying phase unwrapping on the phase image of the reconstructed object along vertical direction. The bottom one is the result after applying phase unwrapping on the top one along horizontal direction.

3. Align projections

Due to the mechanical errors of the stage and lateral shifts ambiguity in the reconstructions, the phase images need to be aligned before the tomographic reconstruction. The alignments include both *x* and *y* directions. The randomly positioned glass beads lead to projections with random features (see Fig. 7.8). Alignment via cross-correlation [154] or fiducial markers [155] is not very effective here. Since the field of view covers the whole sample along *x*, the integration of the unwrapped phase images along *x* should be the same for all the measurement angles. This provides us a way to align the phase images along *y*.

We sum the unwrapped phase images along *x* for all the angles, which can be formularised as

$$\bar{\phi}(y,\theta) = \int \phi(x,y,\theta)\mathrm{d}x. \tag{7.26}$$

As shown in Fig. 7.9a, the distribution of $\bar{\phi}$ for different angle θ look very similar, but with different offsets along *y*. To align them, we choose a section (enclosed in the red rectangular) with big variations and arbitrarily pick a vertical line (or the average of all the vertical lines) as reference. Then for a specific vertical line we shift it to different positions and for each position we calculate the difference with respect to the reference within the selected section. The position that corresponds to the smallest difference will be the aligned position. If we denote the line reference as $\bar{\phi}(y,\Theta)$, and the selected section as a top hat function *w*, we can formularise the alignment as

$$\min_{\Delta y} \sum_{y} w\left[\bar{\phi}(y-\Delta y,\theta) - \bar{\phi}(y,\Theta)\right]^2, \tag{7.27}$$

where Δy denotes the offset along *y*. Repeat the implementations for all the angles to complete the alignment for all the phase projections as shown in Fig. 7.9b.

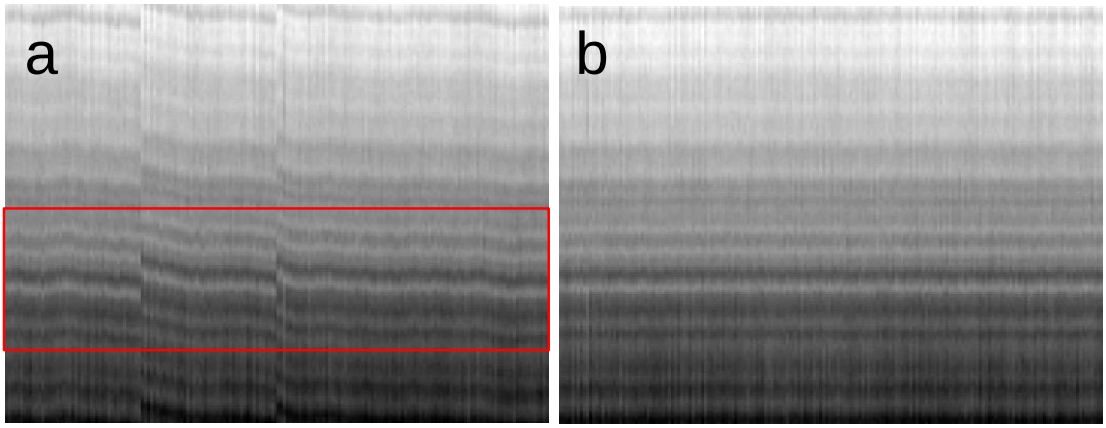

Figure 7.9: Phase projections alignment along *y* (the vertical direction). The distribution of $\bar{\phi}(y,\theta)$ is shown (a) before and (b) after the alignment.

Since the integration of the unwrapped phase images along *y* will not be the same for different angles, we cannot use the above method for the alignment along *x*. However, there is a hidden fiducial marker – the projection centre-of-mass – that allows us to do the alignment along *x* [156]. For angle θ, the projection centre-of-mass is defined as

$$\bar{x}(\theta)=\frac{\int x\bar{\phi}(x,\theta)\mathrm{d}x}{\int\bar{\phi}(x,\theta)\mathrm{d}x}, \tag{7.28}$$

with

$$\bar{\phi}(x,\theta)=\int\phi(x,y,\theta)\mathrm{d}y. \tag{7.29}$$

According to the definition of projection in Eq. (7.2), we have

$$\bar{\phi}(x,\theta)=\iint f(r_x,r_z)\delta\left(r_x\cos\theta+r_z\sin\theta-x\right)\mathrm{d}r_x\mathrm{d}r_z, \tag{7.30}$$

with

$$f(r_x,r_z)=-\frac{2\pi}{\lambda}\int\alpha(r_x,r_y,r_z)\mathrm{d}r_y. \tag{7.31}$$

Substitution of Eq. (7.30) into Eq. (7.28) leads to

$$\bar{x}(\theta)=\frac{\int x\bar{\phi}(x,\theta)\mathrm{d}x}{\int \bar{\phi}(x,\theta)\mathrm{d}x}=\frac{\iint f(r_x,r_z)\left[r_x\cos\theta+r_z\sin\theta\right]\mathrm{d}r_x\mathrm{d}r_z}{\iint f(r_x,r_z)\mathrm{d}r_x\mathrm{d}r_z}=\bar{r}_x\cos\theta+\bar{r}_z\sin\theta, \quad (7.32)$$

with

$$\left.\begin{aligned}\bar{r}_x&=\frac{\iint r_x f(r_x,r_z)\mathrm{d}r_x\mathrm{d}r_z}{\iint f(r_x,r_z)\mathrm{d}r_x\mathrm{d}r_z}\\ \bar{r}_z&=\frac{\iint r_z f(r_x,r_z)\mathrm{d}r_x\mathrm{d}r_z}{\iint f(r_x,r_z)\mathrm{d}r_x\mathrm{d}r_z}\end{aligned}\right\}. \quad (7.33)$$

$\bar{r}_x$ and $\bar{r}_z$ are the centre-of-mass of $f(r_x,r_z)$ along r_x and r_z respectively. According to Eq. (7.3), the relation described by Eq. (7.32) indicates that the centre-of-mass of $f(r_x,r_z)$ will always be projected onto $\bar{x}(\theta)$ at angle θ. However, the relations described by Eq. (7.3) and Eq. (7.32) are based on the assumption that the rotation centre will be projected onto the centre of the detector. If there is an offset Δs between the rotation centre projection and the detector centre, as the real case always does, the relation shown in Eq. (7.32) is then modified as

$$\bar{x}(\theta)=\Delta s+\bar{r}_x\cos\theta+\bar{r}_z\sin\theta, \quad (7.34)$$

In an ideal situation, the distribution of $\bar{x}(\theta)$ over the measurement angle θ should be sinusoidal. Due to the misalignment, the distribution will be wobbly around the sinusoid. The alignment along x is basically trying to correct the random offset of $\bar{\phi}(x,\theta)$ for each angle. The strategy is to use the linear least squares fitting method to find the sinusoid, i.e. to find out the values of Δs, $\bar{r}_x$ and $\bar{r}_z$. This can be written as a set of simultaneous equations in matrix form as

$$\underbrace{\begin{bmatrix} 1 & \cos\theta_0 & \sin\theta_0 \\ 1 & \cos\theta_1 & \sin\theta_1 \\ \vdots & \vdots & \vdots \\ 1 & \cos\theta_{n-1} & \sin\theta_{n-1} \end{bmatrix}}_{\boldsymbol{A}} \underbrace{\begin{bmatrix} \Delta s \\ \bar{r}_x \\ \bar{r}_z \end{bmatrix}}_{\boldsymbol{\sigma}} = \underbrace{\begin{bmatrix} \bar{x}(0) \\ \bar{x}(1) \\ \vdots \\ \bar{x}(n-1) \end{bmatrix}}_{\boldsymbol{\rho}}, \tag{7.35}$$

where n is the number of angular measurements. The standard linear least squares solution [157] to Eq. (7.35) is

$$\boldsymbol{\sigma} = \left[\boldsymbol{A}^T \boldsymbol{A}\right]^{-1} \boldsymbol{A}^T \boldsymbol{\rho} \tag{7.36}$$

For all the angles (from 0° to 179°, 180 measurements), we can calculate the $\bar{x}(\theta)$, $\cos\theta$ and $\sin\theta$ and thus find solutions for Δs, $\bar{r}_x$ and $\bar{r}_z$. According to Eq. (7.34), the estimated projection centre-of-mass $\bar{x}_e(\theta)$ can then be calculated for each angle, so the offset along x for each angle is

$$\Delta x = \bar{x}_e(\theta) - \bar{x}(\theta). \tag{7.37}$$

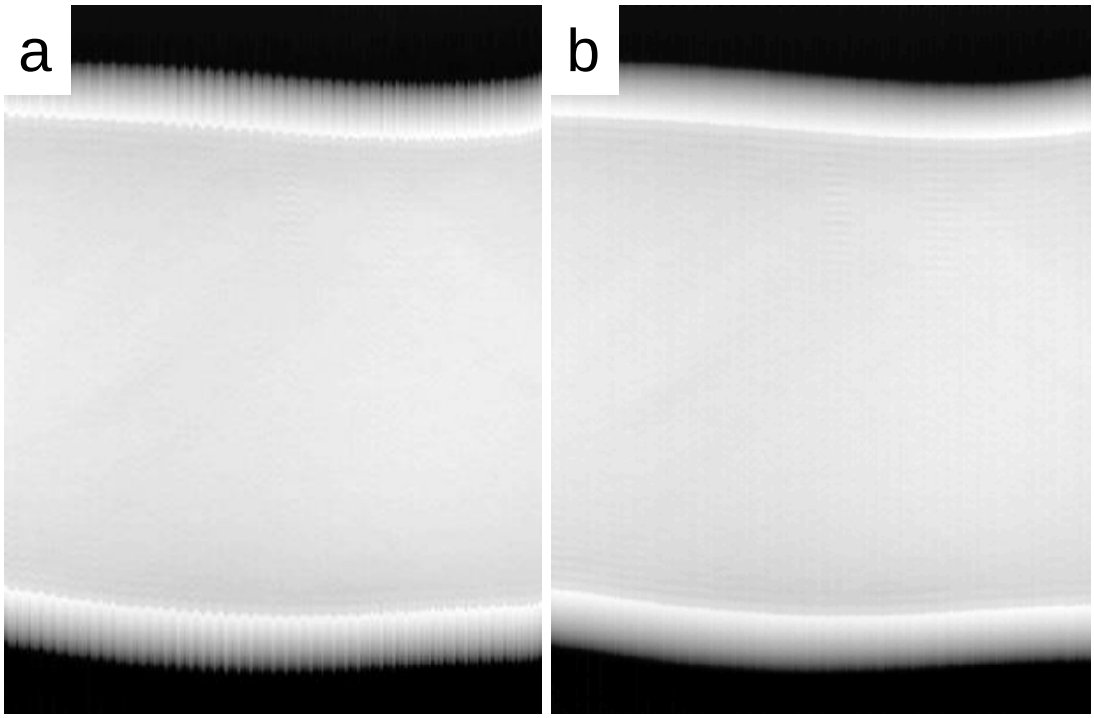

Figure 7.10: Phase projections alignment along x (the horizontal direction). The distribution of $\bar{\phi}(x,\theta)$ is shown (a) before and (b) after the alignment.

Fig. 7.10 shows the distribution of $\bar{\phi}(x,\theta)$ before and after alignment along x. The difference is obvious at the edges. It should be noted that the accuracy of the alignment along x is dependent on the accuracy of the alignment along y. If the y direction is not well aligned, the projection $f(r_x, r_z)$ as defined in Eq. (7.31) will be different for different angles, then the basis for the alignment of x direction described in Eq. (7.32) is violated. Moreover, to ensure Eq. (7.32), here the alignment along y is done before the

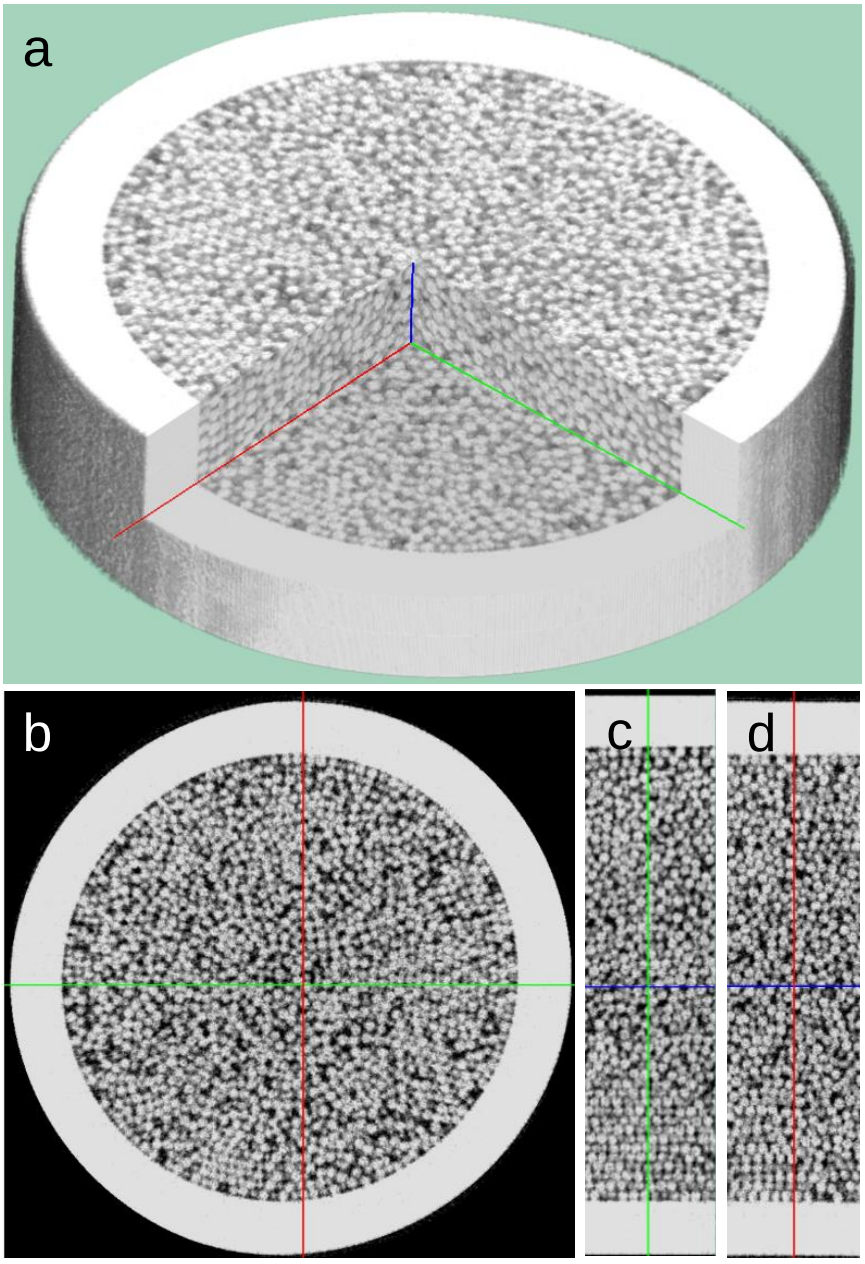

Figure 7.11: The tomographic reconstructions of the glass beads in the 42/52um micro-capillary. (a) 3D rendering of the tomographic reconstruction. (b) The top view. (c) The side view. (d) The front view.

alignment along *x*.

After all these pre-processes described above, the phase images are ready to be fed into FBP algorithm for the tomographic reconstruction. Each row in the unwrapped phase images from all the angles will give a slice reconstruction of the 3D distributions of the component α of the sample. Assembling all the slice reconstructions gives the whole 3D reconstruction as shown in Fig. 7.11.

7.2.4 The possibility of electron ptychographic tomography

In high resolution transmission electron microscopy, the image of an amorphous material becomes uninterpretable and meaningless when the specimen reaches certain thickness, because the superimposition of too many layers of random structures will completely wash out meaningful signals and become noise [158, 159]. This actually can be clearly seen in the x-ray ptychographic reconstruction of the glass beads as shown in Fig. 7.7b. Except the crystalline structures at both inner edges of the capillary, the reconstruction in the central region is completely unidentifiable because of the overlap of so many beads. However, the tomographic reconstruction, as shown in Fig. 7.11, surprisingly reveals the arrangement of every single bead in the bulk. This prompts a possibility of seeing the atom arrangement in amorphous materials via electron ptychographic tomography.

To demonstrate the possibility, we use the x-ray experiment as a scaled experiment for the electron case. The important factor of this demonstration is to show whether an atom can be well resolved enough for it to be seen clearly in the 3D reconstruction for the electron case. From the 3D reconstruction of the x-ray experiment, we know that the beads are well resolved enough. Here we use the ratio of the diameter of the target (i.e. the bead in x-ray case or the atom in electron case) versus the pixel size in the

reconstruction to make the comparison between the x-ray case and the electron case. The ratio basically shows how many pixels across the target in the reconstruction. For the x-ray case, the pixel size, according to Eq. (2.51), is 70nm (calculated by λ/2NA). The ratio of the x-ray case is therefore about 14.3.

For the electron case, suppose the wavelength is 1.97pm (i.e. the beam energy is 300keV) and the scattering has a semi-angle of 30mrad. This gives a pixel size of 0.33Å in the reconstruction. If the atom has a diameter of 2.4Å, the ratio of the electron case is 7.27, which is about half of the ratio of the x-ray case. To match these two ratios, we need to reduce the scattering angle of the x-ray experiment by half. This can be realised by cropping down the diffraction patterns from 256×256 to 128×128. For this reason, we did another reconstruction using the same diffraction data as above but with only the central region of 128×128. The reconstruction is shown in Fig. 7.12. Compared to the result shown in Fig. 7.11, the resolution of each individual bead has not been compromised too much by the cropping of the diffraction patterns. If reviewing the diffraction pattern shown in Fig. 7.6b, we will know the reason is because most of the counts are in the unscattered disc, which means the cropping only cut out very few useful counts. This means the electron ptychographic tomography is able to resolve 2.4Å atoms in a 3D amorphous material given a wavelength of 1.97pm and a semi-scattering-angle of 30mrad.

The capillary in the above x-ray experiment contains about 40 layers of beads. To further test the ability of this technique by using more layers of beads, another set of x-ray ptychographic tomography experiment is carried out using a 117/141μm (inner/outer diameter) capillary filled with 1μm glass beads. The capillary diameter is still much smaller than the maximal allowable thickness set by the projection approximation. The experimental setup remains the same as above. According to Eq.

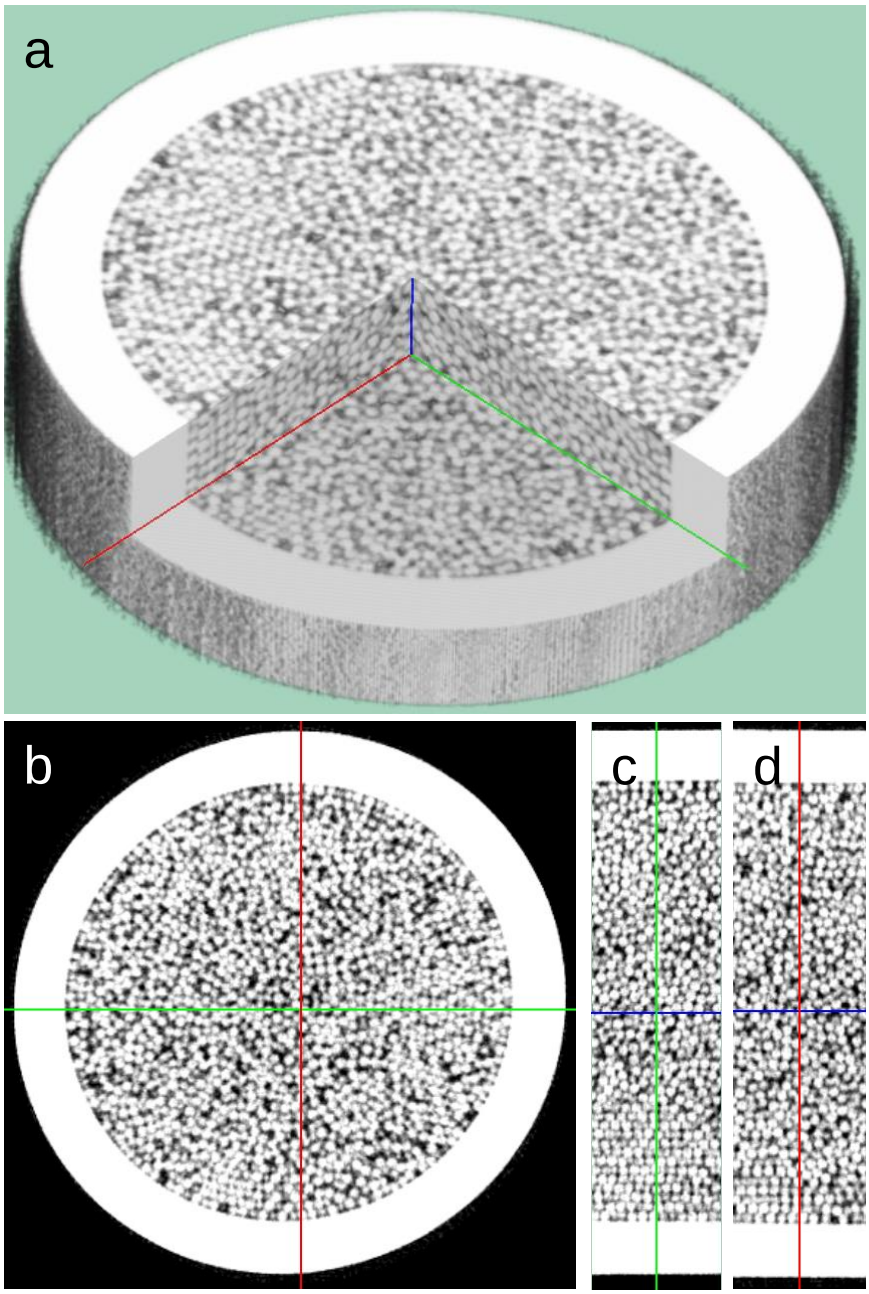

Figure 7.12: The tomographic reconstructions of the glass beads in the 42/52um micro-capillary using the diffraction data with only the central region of 128×128. (a) 3D rendering of the tomographic reconstruction. (b) The top view. (c) The side view. (d) The front view.

(7.20), at least 367 projection measurements are needed. Here 360 sample orientations spaced by 0.5 degree ranging from -90 degrees to 90 degrees were measured (slightly undersampled). At each orientation, the sample was scanned over a 12×70 raster grid with a nominal step size of 2.5μm plus ±20% random offsets, covering a field of view

of 30μm×175μm. At each scan position, the exposure time was 0.15s, giving average total counts of 5.4×10^4 per diffraction pattern.

The diffraction patterns were cropped down to 128×128 for the ptychographic reconstructions. The decrease of the exposure time leads to slightly noisy reconstructions, which imposes challenges in the phase unwrapping and the projection alignment. The increase of the sample dimension and the slight undersample in the angular measurement also cause degradation in the reconstruction. The first attempt of the tomographic reconstruction is shown in Fig. 7.13.

However, there are many challenges and problems present in the electron tomographic experiments [160] or even just in the electron ptychographic experiments, like the sample preparation, the instability of the sample stage and the projection alignment at such small scale (0.33Å), the limited dose to avoid radiation damage, the missing wedge of information due to a restricted tilt range (if the sample has a shape of plate), etc. Although the x-ray experimental results show great potential of electron ptychographic tomography for revealing atom arrangement in amorphous materials in theory, there is still a long way to go in practice.

7.3 Conclusions

In this chapter, we have briefly reviewed some important theoretical basis of tomography and mathematically derived the Fourier slice theorem and the filtered back projection (FBP) reconstruction algorithm. In addition, we have also discussed the two sampling requirements, translational sampling and angular sampling, in tomography. The match of these two samplings imposes a limit on the size of the object under observation for a desired reconstruction resolution with a specific number of projection measurements.

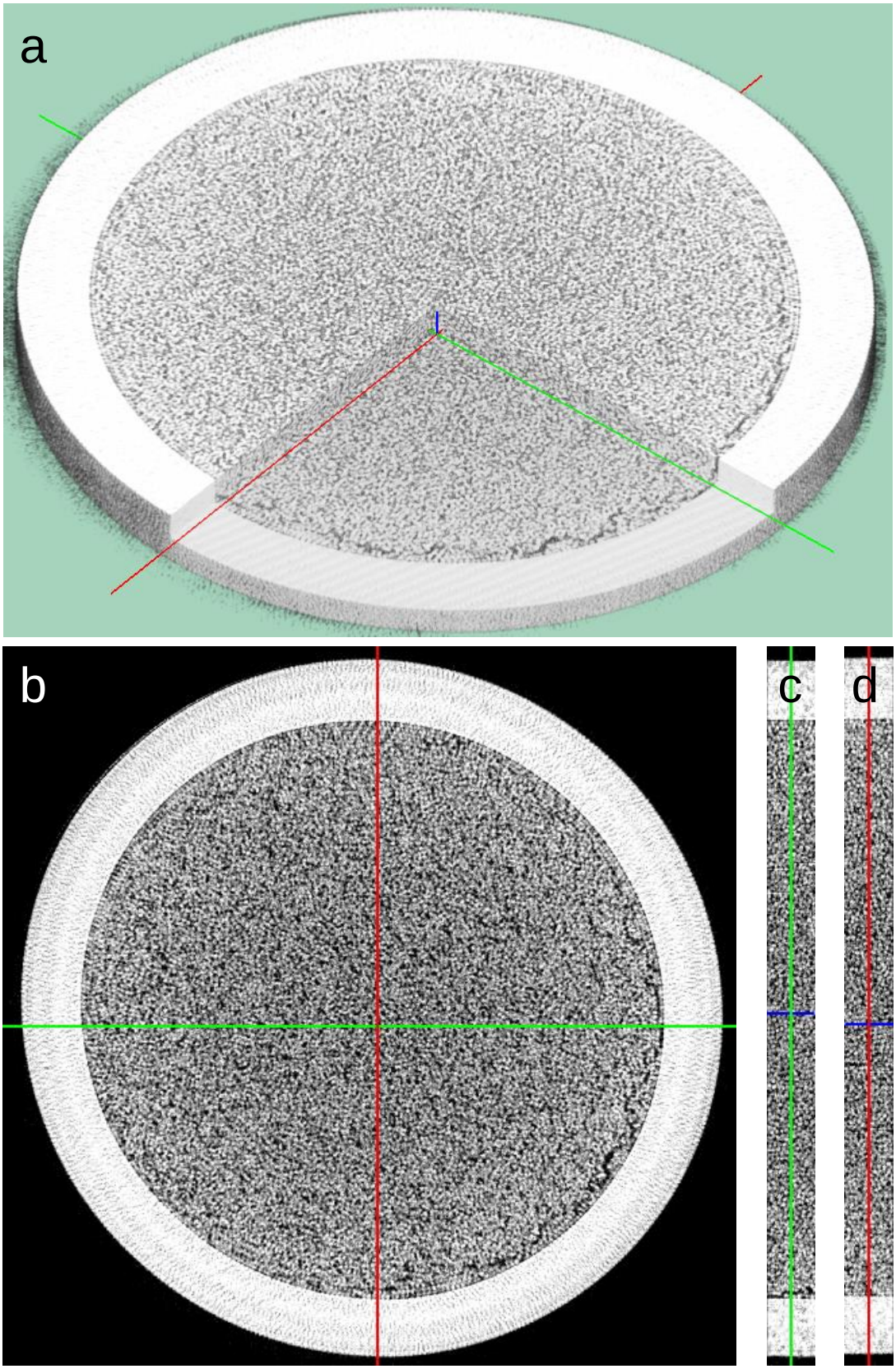

Figure 7.13: The tomographic reconstructions of the glass beads in the 117/141um micro-capillary. (a) 3D rendering of the tomographic reconstruction. (b) The top view. (c) The side view. (d) The front view.

Furthermore, we have demonstrated the combination of ptychography and tomography via an x-ray experiment. Since ptychography has phase ramp and offset ambiguities in the phase reconstruction and also the phase values are wrapped in the range between $-\pi$ and π, we need to remove the ambiguities and unwrap the phase. The ambiguities can be eliminated using the free space in the ptychographic reconstruction as the reference, because the free space should have flat phase of zeros. Phase unwrapping is easy when no residues are present and it can be broken down into two one-dimensional unwrapping processes respectively along vertical and horizontal directions.

Even though the ambiguities are removed and the phases are unwrapped, the phase projections need to be aligned before the tomographic reconstruction due to the mechanical errors of the stage and lateral shifts ambiguity in the ptychographic reconstructions. The alignment is also broken down into two steps: vertical alignment and horizontal alignment. The vertical alignment is realised by making use of the consistency of $\bar{\phi}(y,\theta)$, i.e., the distribution of the phase projection summation along horizontal direction for different sample orientations. The horizontal alignment is achieved by utilising the centre-of-mass of $f(r_x, r_z)$ will always be projected onto $\bar{x}(\theta)$ at angle θ. At this point, the phase projections are ready to be fed into FBP to produce tomographic reconstructions.

Because of the overlap of the beads, projection images produced via ptychography have unidentifiable random structures, but tomographic reconstruction is able to recover the 3D arrangement of the beads in the bulk. This shows great possibility of seeing the 3D arrangement of atoms in amorphous materials via electron ptychographic tomography. The scale calculation between the x-ray experiment and the electron case further strengthens the possibility in theory. Another set of x-ray ptychographic tomography

experiment is carried out on a bigger capillary to further demonstrate the power of this technique. However, in real situations, there are still a lot of problems present in the electron tomographic experiments or even just in the electron ptychographic experiments.

Ptychographic tomography, possessing the strengths of both ptychography and tomography, is able to deliver 3D quantitative phase contrast images at very high resolution. It will be beneficial to many other fields, like life science and material science.

8 Summary and Future Work

8.1 Summary

The thesis was focused on the developments of ptychography to solve partial coherence and achieve 3D imaging.

Partial coherence is important because of its common existence in real experiments, especially in x-rays and electrons. It seriously detriments the quality of the reconstructed images, so to remove it from the reconstructions becomes very important and beneficial. In ptychography partial coherence can be solved non-iteratively using WDDM or iteratively via mixed state decomposition. This thesis investigated both methods.

WDDM is an elegant linear solution to solve the quadratic phase problem, but a dense scan at a step size of the final resolution in real space is needed to collect a big 4D intensity dataset. The Fourier transform of this 4D dataset gives us the product of the WDFs of the specimen and the illumination. During this formulisation, we found that a condition needs to be fulfilled and that is the sum of the specimen size and the probe size needs to be smaller than the calculation window size in each direction. In this formulisation, we can separate out the influence of the probe from the specimen by a deconvolution. However, in the presence of noise the deconvolution is ill-conditioned. It introduces serious errors in the places where the illumination WDF has relatively small values. To mitigate this problem, we have presented a means to design a favourable illumination whose WDF has an even distribution. After the deconvolution, we are presented with a 4D dataset that contains the phase difference between all pixels in the diffraction plane. It is this dataset allows us to solve the phase problem. To

recover the high frequency information lost during the experiment, we proposed a projection strategy that makes use of all the available data to do the job, and hence it outperforms the stepping out method used in 1990s by averaging out the inconsistencies present. Also, we have presented an iterative method to suppress noise effects by exploiting the high redundancy of the 4D dataset. The basic idea is to utilise the well deconvolved data to iteratively refine the badly deconvolved data. Furthermore, we have revisited the effect of spatial partial coherence and for the first time demonstrated the reconstruction of the coherence function using a model calculation via WDDM. However, it needs the source to be symmetric and smoothly varying, so that the coherence function will not have phase information that is not recoverable using WDDM.

Mixed state ptychography provides much more flexibility in solving experimental instabilities (including partial coherence), as long as the instability can be decomposed into a set of states that contribute to the experiment incoherently. There are two types of mixed state ptychography: spatially mixed state ptychography and temporally mixed state ptychography. Spatially mixed state ptychography means the mixed states are caused by the spatially variation (or anything equivalent, such as spatial partial coherence) in the experiment. Every probe state will interact with every object state in spatially mixed state ptychography. We have given the reconstruction algorithm modified based on ePIE and its relation with the gradient descent method. The sum of the Fourier intensities of all the mixed states is kept to equal to the corresponding intensity measurement during the reconstruction, and for each individual state its Fourier intensity is scaled according to their proportion in the sum. It is this modification of the modulus constraint that enables the reconstruction of the mixed states. However, we have mathematically shown that some linear combinations of the

underlying mixed states (both object and probe) are also able to satisfy the modulus constraint, leading to reconstruction ambiguities. We have also mathematically and experimentally demonstrated that the orthogonalisation of the probe states in spatial partial coherence and the phase-only response of the object states can break the ambiguity and lead to true reconstructions. Furthermore, we have looked into the effects of a diffused probe on the reconstructions in the presence of spatial partial coherence. The use of diffuser increases the angular range of the illumination and enriches the information content encoded in the recorded intensity measurements. Eventually, it helps the reconstruction, especially when the specimen has very simple structures. Also, a diffuser can help us estimate the degree of partial coherence in real-time by looking at the histogram distribution of the diffraction pattern.

In temporally mixed state ptychography, the mixed states are wavelength dependent. This means a particular wavelength will only trigger one specific object response. The ePIE-based modified reconstruction algorithm and its relation with the gradient descent method have been given. The reconstruction algorithm is very similar to that used for spatially mixed state ptychography, but with each probe state only interacting with one specific object state (not all the object states). We have mathematically proven that, because of this, the linear reconstruction ambiguity occurred in spatially mixed state ptychography is broken in temporally mixed state ptychography, although the modulus constraint is applied in the same way for the two cases. We have also experimentally demonstrated the breakdown of the linear reconstruction ambiguity using two lasers (633nm and 675nm). The good results have suggested the great potential of this technique to solve for temporal partial coherence. Moreover, as an inspiration of the temporally mixed state ptychography, an iterative method has been developed to reconstruction the static background noise in the intensity measurements by treating it

as the diffraction pattern from an extra temporal state. In this way, we are able to implement the reconstruction without having to measure and subtract the background noise. Experimental results have shown that the reconstruction of the background noise relies on the variation of the superimposed signals through the scan. The low frequencies of the diffraction patterns vary very slowly and this causes high errors in the reconstruction of the background noise.

A specimen in its natural state is usually 3D. Having 2D images, either under the projection approximation or by sectioning thin slices from a thick sample, is not able to completely characterize it. Ptychography has two ways to extend its ability to the depth dimension: the multislice method and ptychographic tomography. This thesis explored both techniques.

Fourier ptychography is a Fourier variant of real-space ptychography. A conventional microscope can be easily modified to implement this technique by inserting a small aperture in the back focal plane of the objective lens. The ptychographic scan is carried out on the spectrum of the specimen by tilting parallel illumination to different angles and the intensity measurement takes places in the image plane. Methodologically, this is the same as real-space ptychography where the small aperture would lie in the specimen plane, the ptychographic scan would be realised by scanning the specimen, and the intensity measurement would occur in the back focal plane. Although this in theory allows the two forms of ptychography to share the same reconstruction algorithms, the different nature of the specimen and its spectrum determine that different considerations need to be taken. The spectrum function normally has a very bright spot in the centre and presents very high dynamic range. In the ePIE algorithm, because the search step size is scaled to the biggest value, this dramatically diminishes the search step sizes (hence slowing down the convergences) of the pixels with small moduli during the

aperture reconstruction. We have adopted a parallel update to avoid this problem by scaling the search step sizes to different values, resulting in relatively even convergences for all the pixels. Besides, we have also extended the reconstruction algorithm to involve three planes: the specimen plane, the back focal plane and the image plane. In this way, the multislice method can be used to provide 3D information when the multiplicative approximation is broken because of a big thickness of the specimen. This is not possible for the original form of the reconstruction algorithm where only the back focal plane and the image plane are considered. Experimental comparisons have been made between this multislice method and a method that directly scans the aperture in the back focal plane without tilting the illumination. The latter method is only able to reconstruct the exit wave of a thick specimen, whilst the multislice method can reconstruct the 3D transmission function of the specimen. Furthermore, the reconstruction resolution has been discussed via the Ewald sphere construction. It is defined as the inverse of the frequency bandwidth accessible in reciprocal space along the corresponding direction. The depth resolution has been further explored via simulations. The results have suggested that the achievable depth resolution is feature-dependent. Features with lower spatial frequency normally will have a worse depth resolution and uniform 3D space (that has zero spatial frequency) cannot be resolved via Fourier ptychography.

Under the multiplicative approximation, a 2D image produced by ptychography is the projection of the 3D transmission function of a specimen. If the specimen is rotated to different orientations and a set of 2D projection images are obtained via ptychography, an isotropic 3D reconstruction of the specimen can be achieved via tomography. We have demonstrated this technique step by step with a synchrotron x-ray experiment of a micro-capillary filled by 1μm glass beads. The capillary has an inner/outer diameter of

42/52μm and the glass beads. 180 sample orientations spaced by 1 degree ranging from -90 degrees to 90 degrees were measured. For each orientation, ptychography was used to reconstruct the corresponding 2D projection image. The phase part of the ptychographic reconstruction of the specimen shows great contrasts between the beads and the capillary, whilst the modulus part is very noisy and barely provides any contrast. However, the simultaneous reconstruction of the illumination and the specimen in ptychography causes inherent ambiguities, such as a phase ramp, phase offset and lateral shift. Phase ramp and phase offset are random, and they cause inconsistency between the reconstructions of different orientations. Fortunately, the field of view (FOV) of ptychography covers the whole sample in horizontal direction, providing us some extra free space at both sides. Free space should be flat, making it very easy to remove the phase ramp, and free space should have zero phase change, making it very easy to remove the phase offset. Apart from these two phase reconstruction ambiguities, phase wrapping also causes inconsistency. We therefore need to unwrap the phase as well. Since we cropped out a well reconstructed region with no phase residues, phase unwrapping is rather easy. A 2D phase unwrapping can be divided into two 1D phase unwrapping implemented one after another and the order does not matter. For 1D phase unwrapping, it is realised by sequentially calculating the phase difference of two adjacent pixels, once the phase difference exceeds π, the phase value of the current pixel will be added 2π or -2π depending on whether the phase difference is smaller than –π or bigger than π. Furthermore, the lateral shift ambiguity causes misalignment of the object reconstructions of different orientations. Even if there is no lateral shift ambiguity, misalignment could happen because of the mechanical instabilities in the stages. The misalignment is in both horizontal and vertical directions and the strategies of the alignment for the two directions are different. The alignment for the vertical direction is

relatively easy because the integration along horizontal direction of the phase reconstructions at different orientations should be the same. The alignment for the horizontal direction is slightly complicated. It utilises a hidden fiducial marker called the projection centre-of-mass. When a 2D image is projected to a set of lines at different angles, for a specific angle the centre-of-mass of the 2D image is always projected onto the centre-of-mass of the corresponding line. The distribution of the centre-of-mass of the lines versus orientation angles should be a sinusoidal. The actual distribution can be calculated according to the experimental geometry and a linear least squares fitting method can then be used to find the right sinusoidal distribution. In this way, the alignment for horizontal direction is accomplished. At this point, the phase reconstructions of different orientations are ready to be fed into FBP algorithm for the tomographic reconstruction. The tomographic reconstruction of the beads, especially the scale calculation between the x-ray and the electron, demonstrates great potential of electron ptychographic tomography for seeing the 3D arrangement of atoms in amorphous materials, even though there are still plenty of practical problems that need to be solved before it can really happen.

8.2 Future work

Some implications for future research of ptychography are discussed below.

Ptychography is a lensless phase imaging technique. The lensless feature of ptychography has gained huge success in high resolution x-ray imaging, because it not only substantially reduces the cost to manufacture high performance x-ray lenses (like Fresnel zone plate), but also simplifies the imaging system without compromising the achievable resolution. The lensless feature also gives plenty of space for complex sample operation, like installing complicated sample mounting system or storing sample

in culture dish. In the aspect of reducing the complexity and the cost of imaging systems, ptychography has great potentials and will play important roles.

Although lensless feature is not viewed as important in electrons and visible lights, the phase imaging feature still enables ptychography to become popular in these wavelengths. Phase imaging allows to see transparent structures without applying contrast agents (like stain and phosphor), which not only simplifies the sample preparation, but also images the samples at their nature states. Moreover, the phase image provided by ptychography is of high quality and quantitative. The quantitative phase information is linearly related to the optical thickness of the sample and it allows us to quantitatively study the sample. As an important quantitative phase imaging tool, ptychography will find its wide applications in material science, life science and etc.

Another very important feature of ptychography is that it is able to reconstruct the illumination function. This can separate out any effects of the illumination on the specimen and improve the image quality. Accurate model of the illumination function needs complete investigation of the optical components placed upstream. Anything like aberration or misalignment would cause a wrong illumination model, and hence affecting the reconstructed images. Being able to reconstruct the illumination saves all the troubles to obtain an accurate illumination model. Moreover, comparing the actual illumination wavefield (the reconstructed illumination function) and the ideal illumination model would also allow us to characterize the upstream optics and correct the defeats in the optics. This wavefront sensing feature of ptychography is becoming a powerful tool for optics characterization [161] and optical metrology [162].

Compared to iterative algorithms, WDDM is much quicker to reconstruct the images. However, a big 4D dataset needs to be obtained from a dense scan, which makes

WDDM not experimental friendly. But as a theoretical tool, WDDM provides us insightful understanding about ptychography. It could be a very useful tool to systematically study the sampling of ptychography which further includes the effects of the illumination size and the detector pixel size, not just about the scan step size in specimen plane and the sampling pitch in detector plane as investigated in [102]. Moreover, WDDM on its own may not seem very practical, but it can be combined with other imaging method that shares the same experimental implementations, like scanning transmission electron microscopy (STEM), to provide phase contrast information about the specimen under inspection. STEM performed in annular dark-field (ADF) imaging mode is able to provide contrast of atomic number (i.e. Z-contrast) which helps to identify heavy atoms. Phase contrast image, on the other hand, gives good contrast of the light elements. These two images can be obtained simultaneously in one single experiment [163].

Mixed state ptychography has substantially relaxed the requirements on the experiments and extended the ability of ptychography. Multi-state models arise in many physical situations where mixed state ptychography may become a useful tool. Examples of the application of multiple probe states include modelling partial spatial coherence from an extended source, lateral specimen vibrations and detector point spread [64]. Examples where multiple object states apply are also wide ranging and include axially vibrating structures – the tympanic membrane being one interesting instance [164] – and flicker in red blood cell membranes [165]. Models relying on both multiple probe and multiple object states are of particular relevance to electron microscopy, where the electron beam is always partially spatially coherent and its interaction with a specimen generates multiple states in the form of plasmon and phonon resonances and structural instabilities, for example in the case of a bistable dislocation core [166].

So far, temporally mixed state ptychography has not been fully explored and only discrete wavelength spectrum has been demonstrated. To further explore it, experiments using continuous wavelength spectrum would be ideal and necessary. As a result, high quality lasers used in optical ptychography maybe can be replaced with low cost LEDs, or gigantic synchrotron x-rays sources used in x-ray ptychography can be replaced by laboratory table-top x-rays. Besides, this also indicates the the possibility of spectroscopic imaging via ptychography in one single experiment, instead of one ptychographic scan for each energy [129]. Moreover, ptychography has only been demonstrated to solve spatially and temporally partial coherence separately, but it is very likely that ptychography is able to handle both of them simultaneously. The reconstruction structure can be envisaged that the linear reconstruction ambiguity will occur between the spatially mixed states within each wavelength, but not between different wavelengths. This could substantially relax both the spatial and temporal coherence requirement on the radiation source, like the field emission gun used in electron ptychography.

Multislice ptychography can achieve 3D imaging without the need of sample rotation. From the perspective of data acquisition, it is a great advantage. However, the longitudinal resolution is rather low compared with the transverse resolution. Ptychographic tomography, on the other hand, can produce isotropic 3D high resolution, but the data acquisition is very time-consuming, because the sample need to be rotated 180 degrees at a small angular step and a ptychographic scan needs to be performed at each sample orientation. A combination of these two techniques could increase the longitudinal resolution and meanwhile reduce the angular measurements. It looks very appealing and definitely worth further exploration. Whilst the attention is dragged to improve the longitudinal resolution, multislice ptychography has another very important

potential and that is to extend the depth of field. High resolution imaging techniques have very small depth of field and as a consequence only a thin layer of the specimen can be imaged in focus. Multislice ptychography, however, can extend the limited depth of field by sectioning the out of focus features into one layer or more (depending on the optical setup and sample thickness) [167].

Lastly, the concept of ptychography is being generalised and adopted to many other imaging techniques. Conventional ptychography requires an overlapping scan of a localised illumination across the specimen at a grid of positions and the measurements of the transmitted or reflected signals somewhere downstream. The overlapping scan provides diversity in the measurements that allows to reconstruct the image of the specimen. The measurements need to take place at a different plane from the scanning plane, so that the lost information can be encoded in the measurements. A successful modification of conventional ptychography is Fourier ptychography which utilises a series of tilted plane waves and a small aperture to accomplish the overlapping scan in Fourier space of the specimen. The use of LED array to generate the tilted plane waves makes it very easy to be compatible with most existing microscope platforms, no major changes to be made. As a result, a normal microscope can deliver high resolution quantitative phase images [78]. Moreover, the concept of ptychography can be applied to incoherent imaging as well. An epifluorescence microscope has been successfully modified to achieve resolution improvement using randomly structured illumination [168]. This implies that the general concept of ptychography can be applied to many other problems to retrieve the lost information during the measurements.

Bibliography

[1] E. Abbé. Contributions to the theory of the microscope and microscopic perception. *Arch Für Mikrosk Anat*, 9(1): 413-418, 1873.

[2] S. G. Lipson, H. Lipson, and D. S. Tannhauser. *Optical Physics*. Cambridge University Press, Cambridge, third edition, 1995.

[3] A. Sakdinawat, and D. Attwood. Nanoscale X-ray imaging. *Nature Photonics*, *4*(12): 840-848, 2010.

[4] R. F. Egerton. *Physical Principles of Electron Microscopy: An introduction to TEM, SEM, and AEM*. Springer US, 2005.

[5] O. Scherzer. Some defects of electron lenses. *Z. Phys.*, 101(9-10): 593-603, 1936.

[6] O. Scherzer. Spherical and chromatic correction of electron lenses. *Optik*, 2: 114-132, 1947.

[7] O. Scherzer. The theoretical resolution limit of the electron microscope. *Journal of Applied Physics*, 20(1): 20-29, 1949.

[8] O. L. Krivanek, N. Dellby, A. J. Spence, R. A. Camps, and L. M. Brown. Aberration correction in the STEM. *Institute of Physics Conference*, 153: 35-40, 1997.

[9] F. Zernike. Phase-contrast, a new method for microscopic observation of transparent objects. *Physica*, 9(7): 686-698, 1942.

[10] G. Nomarski. Microinterféromètre différentiel à ondes polarisées. *Journal de Physique et le Radium*, 16: S9-S13, 1955.

[11] A. Barty, K. A. Nugent, D. Paganin, and A. Roberts. Quantitative optical phase microscopy. *Optics Letters*, 23(11): 817-819, 1998.

[12] C. Giacovazzo. *Fundamentals of crystallography*. Oxford University Press, New York, third edition, 2002.

[13] W. Hoppe. Diffraction in inhomogeneous primary wave fields. 1. Principle of phase determination from electron diffraction interference. *Acta Crystallogr. A*, 25: 495-501, 1969.

[14] W. Hoppe, and G. Strube. Diffraction in inhomogeneous primary wave fields. 2. Optical experiments for phase determination of lattice interferences. *Acta Crystallogr. A*, 25: 502-507, 1969.

[15] W. Hoppe. Diffraction in inhomogeneous primary wave fields. 3. Amplitude and phase determination for nonperiodic objects. *Acta Crystallogr. A*, 25: 508-515, 1969.

[16] R. Hegerl, and W. Hoppe. Dynamische Theorie der Kristallstrukturanalyse durch Elektronenbeugung im inhomogen Primarstrahlwellenfeld. Ber. Bunsenges. *Physik. Chemie*, 74: 1148-1154, 1970.

[17] R. Hegerl, and W. Hoppe. Phase evaluation in generalized diffraction (ptychography). *Proc. Fifth Eur. Cong. Electron Microscopy*, 628-629, 1972.

[18] B. C. McCallum, and J. M. Rodenburg. Error analysis of crystalline ptychography in the STEM mode. *Ultramicroscopy*, 52(1): 85-99, 1993.

[19] W. Hoppe. Trace structure analysis, ptychography, phase tomography. *Ultramicroscopy*, 10: 187-198, 1982.

[20] R. H. T. Bates, and J. M. Rodenburg. Sub-Ångström transmission microscopy: a Fourier transform algorithm for microdiffraction plane intensity information. *Ultramicroscopy*, 31(3): 303-307, 1989.

[21] B. C. McCallum, and J. M. Rodenburg. Two-dimensional demonstration of Wigner phase-retrieval microscopy in the STEM configuration. *Ultramicroscopy*, 45(3): 371-380, 1992.

[22] H. N. Chapman. Phase-retrieval X-ray microscopy by Wigner-distribution deconvolution. *Ultramicroscopy*, 66(3): 153-172, 1996.

[23] J. M. Rodenburg, B. C. McCallum, and P. D. Nellist. Experimental tests on double-resolution coherent imaging via STEM. *Ultramicroscopy*, 48(3): 304-314, 1993.

[24] P. D. Nellist, B. C. McCallum, and J. M. Rodenburg. Resolution beyond the 'information limit' in transmission electron microscopy. *Nature*, 374: 630-632, 1995.

[25] J. A. Johnston. Wigner distribution and FM radar signal design. *IEE Proceedings F-Radar and Signal Processing*, 136(2): 81-88, 1989.

[26] J. M. Rodenburg, and R. H. T. Bates. The theory of super-resolution electron microscopy via Wigner-distribution deconvolution. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 339(1655): 521-553, 1992.

[27] P. D. Nellist, and J. M. Rodenburg. Beyond the conventional information limit: the relevant coherence function. *Ultramicroscopy*, 54(1): 61-74, 1994.

[28] D. Gabor. A new microscopic principle. *Nature*, 161(4098): 777-778, 1948.

[29] E. N. Leith, and J. Upatnieks. Reconstructed wavefronts and communication theory. *JOSA*, 52(10): 1123-1130, 1962.

[30] I. McNulty, and C. Jacobsen. High-resolution imaging by Fourier transform X-ray holography. *Science*, 256(5059): 1009, 1992.

[31] S. Eisebitt, J. Lüning, W. F. Schlotter, M. Lörgen, O. Hellwig, W. Eberhardt, and J. Stöhr. Lensless imaging of magnetic nanostructures by X-ray spectro-holography. *Nature*, 432(7019): 885-888, 2004.

[32] S. Marchesini, S. Boutet, A. E. Sakdinawat, M. J. Bogan, S. Bajt, A. Barty, H. N. Chapman, M. Frank, S. P. Hau-Riege, A. Szöke, and C. Cui. Massively parallel X-ray holography. *Nature Photonics*, 2(9): 560-563, 2008.

[33] M. Guizar-Sicairos, and J. R. Fienup. Holography with extended reference by autocorrelation linear differential operation. *Optics Express*, 15(26): 17592-17612, 2007.

[34] D. Sayre. Some implications of a theorem due to Shannon. *Acta Crystallographica*, 5(6): 843–843, 1952.

[35] Yu M. Bruck, and L. G. Sodin. On the ambiguity of the image reconstruction problem. *Optics Communications*, 30(3): 304-308, 1979.

[36] R. H. T. Bates. Fourier phase problems are uniquely solvable in mute than one dimension. I: Underlying theory. *Optik (Stuttgart)*, 61: 247-262, 1982.

[37] W. B. Yun, J. Kirz, and D. Sayre. Observation of the soft X-ray diffraction pattern of a single diatom. *Acta Crystallographica Section A: Foundations of Crystallography*, 43(1): 131–133, 1987.

[38] J. Miao, P. Charalambous, J. Kirz, and D. Sayre. Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. *Nature*, 400(6742): 342-344, 1999.

[39] J. R. Fienup. Reconstruction of an object from the modulus of its Fourier transform. *Optics Letters*, 3(1): 27-29, 1978.

[40] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15): 2758-2769, (1982).

[41] P. Thibault, V. Elser, C. Jacobsen, D. Shapiro, and D. Sayre. Reconstruction of a yeast cell from x-ray diffraction data. *Acta Crystallographica Section A: Foundations of Crystallography*, 62(4): 248-261, 2006.

[42] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. Spence, and U. Weierstall. High-resolution ab initio three-dimensional x-ray diffraction microscopy. *JOSA A*, 23(5):1179-1200, 2006.

[43] P. Thibault, and V. Elser. X-ray diffraction microscopy. *Condensed Matter Physics*, 1, 2010.

[44] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane. Beyond crystallography: Diffractive imaging using coherent x-ray light sources. *Science*, 348(6234): 530-535, 2015.

[45] R. W. Gerchberg. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35: 237, 1972.

[46] J. Miao, and D. Sayre. On possible extensions of X-ray crystallography through diffraction-pattern oversampling. *Acta Crystallographica Section A: Foundations of Crystallography*, 56(6): 596-605, 2000.

[47] J. R. Fienup. Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint. *JOSA A*, 4(1): 118-123, 1987.

[48] V. Elser. Phase retrieval by iterated projections. *JOSA A*, 20(1): 40-55, 2003.

[49] S. Marchesini. Invited article: A unified evaluation of iterative projection algorithms for phase retrieval. *Review of Scientific Instruments*, 78(1): 011301, 2007.

[50] J. C. H. Spence, U. Weierstall, and M. Howells. Phase recovery and lensless imaging by iterative methods in optical, X-ray and electron diffraction. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 360(1794): 875-895, 2002.

[51] H. M. L. Faulkner, and J. M. Rodenburg. Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm. *Physical review letters*, 93(2): 023903, 2004.

[52] J. M. Rodenburg, and H. M. Faulkner. A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20): 4795-4797, 2004.

[53] J. M. Rodenburg, A. C. Hurst, and A. G. Cullis. Transmission microscopy without lenses for objects of unlimited size. *Ultramicroscopy*, 107(2): 227-231, 2007.

[54] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer. Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy*, 108(5): 481-487, 2008.

[55] J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson. Hard-x-ray lensless imaging of extended objects. *Physical Review Letters*, 98(3): 034801, 2007.

[56] F. Hüe, J. M. Rodenburg, A. M. Maiden, Francis Sweeney, and P. A. Midgley. Wave-front phase retrieval in transmission electron microscopy via ptychography. *Physical Review B*, 82(12): 121415, 2010.

[57] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer. High-resolution scanning x-ray diffraction microscopy. *Science*, 321(5887): 379-382, 2008.

[58] M. Guizar-Sicairos, and J. R. Fienup. Phase retrieval with transverse translation diversity: a nonlinear optimization approach. *Optics Express*, 16(10): 7264-7278, 2008.

[59] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer. Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*, 109(4): 338-343, 2009.

[60] A. M. Maiden, and J. M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10): 1256-1262, 2009.

[61] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus, and J. M. Rodenburg. An annealing algorithm to correct positioning errors in ptychography. *Ultramicroscopy*, 120: 64-72, 2012.

[62] F. Zhang, I. Peterson, J. Vila-Comamala, A. Diaz, F. Berenguer, R. Bean, B. Chen, A. Menzel, I. K. Robinson, and J. M. Rodenburg. Translation position determination in ptychographic coherent diffraction imaging. *Optics Express*, 21(11): 13592-13606, 2013.

[63] A. Tripathi, I. McNulty, and O. G. Shpyrko. Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods. *Optics Express*, 22(2): 1452-1466, 2014.

[64] P. Thibault, and A. Menzel. Reconstructing state mixtures from diffraction measurements. *Nature*, 494(7435): 68-71, 2013.

[65] T. B. Edo, F. Zhang, and J. M. Rodenburg. Resolution improvement in coherent diffractive imaging (ptychography). *International Society for Optics and Photonics*, Scanning Microscopy: 77291H-77291H, 2010.

[66] A. M. Maiden, M. J. Humphry, F. Zhang, and J. M. Rodenburg. Superresolution imaging via ptychography. *JOSA A*, 28(4): 604-612, 2011.

[67] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg. Ptychographic transmission microscopy in three dimensions using a multi-slice approach. JOSA A, 29(8): 1606-1614, 2012.

[68] T. M. Godden, R. Suman, M. J. Humphry, J. M. Rodenburg, and A. M. Maiden. Ptychographic microscope for three-dimensional imaging. *Optic Express*, 22(10): 12513-12523, 2014.

[69] A. Suzuki, S. Furutaku, K. Shimomura, K. Yamauchi, Y. Kohmura, T. Ishikawa, and Y. Takahashi. High-resolution multislice x-ray ptychography of extended thick objects. *Physical Review Letters*, 112(5): 053903, 2014.

[70] P. Godard, M. Allain, V. Chamard, and J. Rodenburg. Noise models for low counting rate coherent diffraction imaging. *Optics Express*, 20(23): 25914-25934, 2012.

[71] P. Thibault, and M. Guizar-Sicairos. Maximum-likelihood refinement for coherent diffractive imaging. *New Journal of Physics*, 14(6): 063004, 2012.

[72] D. J. Batey, D. Claus, and J. M. Rodenburg. Information multiplexing in ptychography. *Ultramicroscopy*, 138: 13-21, 2014.

[73] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson. Dynamic imaging using ptychography. *Physical Review Letters*, 112(11): 113901, 2014.

[74] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson. Continuous scanning mode for ptychography. *Optics Letters*, 39(20): 6066-6069, 2014.

[75] P. M. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel. On-the-fly scans for X-ray ptychography. *Applied Physics Letters*, 105(25): 251101, 2014.

[76] X. Huang, K. Lauer, J. N. Clark, W. Xu, E. Nazaretski, R. Harder, I. K. Robinson, and Y. S. Chu. Fly-scan ptychography. *Scientific Reports*, 5, 2015.

[77] P. Godard, G. Carbone, M. Allain, F. Mastropietro, G. Chen, L. Capello, A. Diaz, T. H. Metzger, J. Stangl, and V. Chamard. Three-dimensional high-resolution quantitative microscopy of extended crystals. *Nature Communications*, 2: 568, 2011.

[78] G. Zheng, R. Horstmeyer, and C. Yang. Wide-field, high-resolution Fourier ptychographic microscopy. *Nature Photonics*, 7(9): 739-745, 2013.

[79] M. Dierolf, A. Menzel, P. Thibault, P. Schneider, C. M. Kewish, R. Wepf, O. Bunk, and F. Pfeiffer. Ptychographic X-ray computed tomography at the nanoscale. *Nature*, 467(7314): 436-439, 2010.

[80] B. Chen, R. A. Dilanian, S. Teichmann, B. Abbey, A. G. Peele, G. J. Williams, P. Hannaford, L. Van Dao, H. M. Quiney, and K. A. Nugent. Multiple wavelength diffractive imaging. *Physical Review A*, 79(2): 023809, 2009.

[81] L. W. Whitehead, G. J. Williams, H. M. Quiney, D. J. Vine, R. A. Dilanian, S. Flewett, K. A. Nugent, A. G. Peele, E. Balaur, and I. McNulty. Diffractive imaging using partially coherent x rays. *Physical Review Letters*, 103(24): 243902, 2009.

[82] B. Abbey, L. W. Whitehead, H. M. Quiney, D. J. Vine, G. A. Cadenazzi, C. A. Henderson, K. A. Nugent, E. Balaur, C. T. Putkunz, A. G. Peele, and G. J. Williams. Lensless imaging using broadband X-ray sources. *Nature Photonics*, 5(7): 420-424, 2011.

[83] J. N. Clark, and A. G. Peele. Simultaneous sample and spatial coherence characterisation using diffractive imaging. *Applied Physics Letters*, 99(15): 154103, 2011.

[84] P. Thibault, M. Guizar-Sicairos, and A. Menzel. Coherent imaging at the diffraction limit. Journal of Synchrotron Radiation, 21(5): 1011-1018, 2014.

[85] J. Marrison, L. Räty, P. Marriott, and P. O'Toole. Ptychography-a label free, high-contrast imaging technique for live cells using quantitative phase information. *Scientific Reports*, 3, 2013.

[86] T. M. Godden, A. Muñiz-Piniella, J. D. Claverley, A. Yacoot, and M. J. Humphry. Phase calibration target for quantitative phase imaging with ptychography. *Optics Express*, 24(7): 7679-7692, 2016.

[87] P. Trtik, A. Diaz, M. Guizar-Sicairos, A. Menzel, and O. Bunk. Density mapping of hardened cement paste using ptychographic X-ray computed tomography. *Cement and Concrete Composites*, 36: 71-77, 2013.

[88] B. Chen, M. Guizar-Sicairos, G. Xiong, L. Shemilt, A. Diaz, J. Nutter, N. Burdet, S. Huo, J. Mancuso, A. Monteith, and F. Vergeer. Three-dimensional structure analysis and percolation properties of a barrier marine coating. *Scientific Reports*, 3, 2013.

[89] M. Esmaeili, J. B. Fløystad, A. Diaz, K. Høydalsvik, M. Guizar-Sicairos, J. W. Andreasen, and D. W. Breiby. Ptychographic X-ray tomography of silk fiber hydration. *Macromolecules*, 46(2): 434-439, 2013.

[90] A. Diaz, M. Guizar-Sicairos, A. Poeppel, A. Menzel, and O. Bunk. Characterization of carbon fibers using X-ray phase nanotomography. *Carbon*, 67: 98-103, 2014.

[91] J. M. Cowley, and A. F. Moodie. The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Crystallographica*, 10(10): 609-619, 1957.

[92] J. W. Goodman. *Introduction to Fourier Optics*. Roberts and Company, Englewood, Colorado, 3rd edition, 2005.

[93] J. W. Cooley, and J. W. Tukey. An algorithm for machine calculation of complex Fourier series. *Math. Comput.*, 19(90): 297-301, 1965.

[94] C. E. Shannon. Communication in the presence of noise. *Proc. Inst. Radio Eng.*, 37(1): 10-21, 1949.

[95] E. T. Whittaker. Xviii.-on the functions which are represented by the expansions of the interpolation-theory. *Proceedings of the Royal Society of Edinburgh*, 35: 181-194, 1915.

[96] M. Born, and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, Cambridge, 7th edition, 2006.

[97] L. Mandel, and E. Wolf. Coherence properties of optical fields. *Reviews of Modern Physics*, 37(2): 231, 1965.

[98] E. Wolf. New theory of partial coherence in the space-frequency domain. Part I: spectra and cross spectra of steady-state sources. *JOSA*, 72(3): 343-351, 1982.

[99] F. Van der Veen, and F. Pfeiffer. Coherent x-ray scattering. *Journal of Physics: Condensed Matter*, 16(28): 5003, 2004.

[100] J. Miao, D. Sayre, and H. N. Chapman. Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects. *J. Opt. Soc. Am. A*, 15(6): 1662-1669, 1998.

[101] J. Miao, and D. Sayre. On possible extensions of X-ray crystallography through diffraction-pattern oversampling. *Acta Crystallographica Section A Foundations of Crystallography*, 56(6): 596-605, 2000.

[102] T. B. Edo, D. J. Batey, A. M. Maiden, C. Rau, U. Wagner, Z. D. Pešić, T. A. Waigh, and J. M. Rodenburg. Sampling in x-ray ptychography. *Physical Review A*, 87(5): 053850, 2013.

[103] D. J. Batey, T. B. Edo, C. Rau, U. Wagner, Z. D. Pešić, T. A. Waigh, and J. M. Rodenburg. Reciprocal-space up-sampling from real-space oversampling in x-ray ptychography. *Physical Review A*, 89(4): 043812, 2014.

[104] P. P. Ewald. Introduction to the dynamical theory of X-ray diffraction. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, 25(1): 103-108, 1969.

[105] J. Snyman. *Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms*. Science and Business Media, Springer, 2005.

[106] M. Dierolf, P. Thibault, A. Menzel, C. M. Kewish, K. Jefimovs, I. Schlichting, K. Von Koenig, O. Bunk, and F. Pfeiffer. Ptychographic coherent diffractive imaging of weakly scattering specimens. *New Journal of Physics*, 12(3): 035017, 2010.

[107] M. Guizar-Sicairos, A. Diaz, M. Holler, M. S. Lucas, A. Menzel, R. A. Wepf, and O. Bunk. Phase tomography from x-ray coherent diffractive imaging projections. *Optics Express*, 19(22): 21345-21357, 2011.

[108] C. M. Kewish, P. Thibault, M. Dierolf, O. Bunk, A. Menzel, J. Vila-Comamala, K. Jefimovs, and F. Pfeiffer. Ptychographic characterization of the wavefield in the focus of reflective hard X-ray optics. *Ultramicroscopy*, 110(4): 325-329, 2010.

[109] J. Vila-Comamala, A. Diaz, M. Guizar-Sicairos, A. Mantion, C. M. Kewish, A. Menzel, O. Bunk, and C. David. Characterization of high-resolution diffractive X-ray optics by ptychographic coherent diffractive imaging. *Optics Express*, 19(22): 21333-21344, 2011.

[110] Andreas. Schropp, R. Hoppe, V. Meier, J. Patommel, F. Seiboth, H. J. Lee, B. Nagler, E. C. Galtier, B. Arnold, U. Zastrau, and J. B. Hastings. Full spatial characterization of a nanofocused x-ray free-electron laser beam by ptychographic imaging. *Scientific Reports*, 3, 2013.

[111] W. O. Saxton, and W. Baumeister. The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of Microscopy*, 127(2): 127-138, 1982.

[112] M. Van Heel, and M. Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3): 250-262, 2005.

[113] J. R. Fienup. Invariant error metrics for image reconstruction. *Applied Optics*, 36(32): 8352-8357, 1997.

[114] T. J. Pennycook, A. R. Lupini, H. Yang, M. F. Murfitt, L. Jones, and P. D. Nellist. Efficient phase contrast imaging in STEM using a pixelated detector. Part 1: Experimental demonstration at atomic resolution. *Ultramicroscopy*, 151: 160-167, 2015.

[115] H. Yang, T. J. Pennycook, and P. D. Nellist. Efficient phase contrast imaging in STEM using a pixelated detector. Part II: Optimisation of imaging conditions. *Ultramicroscopy*, 151: 232-239, 2015.

[116] P. Li, T. B. Edo, and J. M. Rodenburg. Ptychographic inversion via Wigner distribution deconvolution: noise suppression and probe design. *Ultramicroscopy*, 147: 106-113, 2014.

[117] F. Zernike. The concept of degree of coherence and its application to optical problems. *Physica*, 5(8): 785-795, 1938.

[118] P. Li, T. Edo, D. Batey, J. Rodenburg, and A. Maiden. Breaking ambiguities in mixed state ptychography. *Optics Express*, 24(8): 9038-9052, 2016.

[119] A. Pogany, D. Gao, and S. W. Wilkins. Contrast and resolution in imaging with a microfocus x-ray source. *Review of Scientific Instruments*, 68(7): 2774-2782, 1997.

[120] D. L. Misell. On the validity of the weak-phase and other approximations in the analysis of electron microscope images. *Journal of Physics D: Applied Physics*, 9(13): 1849, 1976.

[121] H. Yoshimura, T. Asakura, and N. Takai. Spatial coherence properties of light from optical fibres. *Optical and Quantum Electronics*, 24(6): 631-646, 1992.

[122] B. Enders, M. Dierolf, P. Cloetens, M. Stockmar, F. Pfeiffer, and P. Thibault. Ptychography with broad-bandwidth radiation. *Applied Physics Letters*, 104(17): 171104, 2014.

[123] P. Li, D. J. Batey, T. B. Edo, A. D. Parsons, C. Rau, and J. M. Rodenburg. Multiple mode x-ray ptychography using a lens and a fixed diffuser optic. *Journal of Optics*, 18(5): 054008, 2016.

[124] M. Guizar-Sicairos, M. Holler, A. Diaz, J. Vila-Comamala, O. Bunk, and A. Menzel. Role of the illumination spatial-frequency spectrum for ptychography. *Physical Review B*, 86(10): 100103, 2012.

[125] N. Burdet, , X. Shi, D. Parks, J. N. Clark, X. Huang, S. D. Kevan, and I. K. Robinson. Evaluation of partial coherence correction in X-ray ptychography. *Optics Express*, 23(5): 5452-5467, 2015.

[126] M. Sutton, S. G. J. Mochrie, T. Greytak, S. E. Nagler, L. E. Berman, G. A. Held, and G. B. Stephenson. Observation of speckle by diffraction with coherent X-rays. *Nature*, 352(6336): 608-610, 1991.

[127] J. W. Goodman. *Statistical properties of laser speckle patterns*. Laser Speckle and Related Phenomena, Springer Berlin Heidelberg, 1975.

[128] J. W. Goodman. Some fundamental properties of speckle. *JOSA*, 66(11): 1145-1150, 1976.

[129] A. M. Maiden, G. R. Morrison, B. Kaulich, A. Gianoncelli, and J. M. Rodenburg. Soft X-ray spectromicroscopy using ptychography with randomly phased illumination. *Nature Communications*, 4: 1669, 2013.

[130] G. S. Kino, and T. R. Corle. *Confocal scanning optical microscopy and related imaging systems*. Academic Press, California, USA, 1996.

[131] D. J. Stephens, and V. J. Allan. Light microscopy techniques for live cell imaging. Science, 300(5616): 82-86, 2003.

[132] P. Li, D. J. Batey, T. B. Edo, and J. M. Rodenburg. Separation of three-dimensional scattering effects in tilt-series Fourier ptychography. *Ultramicroscopy*, 158: 1-7, 2015.

[133] M. Ryle. 5-km radio telescope at Cambridge. *Nature*, 239(5373): 435-438, 1972.

[134] E. Shearman, and J. Clarke. Aperture synthesis in ionospheric radar. *Nature*, 219: 143-144, 1968.

[135] A. I. Kirkland, W. O. Saxton, K. L. Chau, K. Tsuno, and M. Kawasaki. Super-resolution by aperture synthesis: tilt series reconstruction in CTEM. *Ultramicroscopy*, 57(4): 355-374, 1995.

[136] A. I. Kirkland, W. O. Saxton, and G. Chand. Multiple beam tilt microscopy for super resolved imaging. Journal of Electron Microscopy, 46(1): 11-22, 1997.

[137] S. J. Haigh, H. Sawada, and A. I. Kirkland. Atomic structure imaging beyond conventional resolution limits in the transmission electron microscope. *Physical Review Letters*, 103(12): 126101, 2009.

[138] S. Dong, R. Horstmeyer, R. Shiradkar, K. Guo, X. Ou, Z. Bian, H. Xin and G. Zheng. Aperture-scanning Fourier ptychography for 3D refocusing and super-resolution macroscopic imaging. *Optics Express*, 22(11): 13586-13599, 2014.

[139] L. Tian, and L. Waller. 3D intensity and phase imaging from light field measurements in an LED array microscope. *Optica*, 2(2): 104-111, 2015.

[140] M. Holler, A. Diaz, M. Guizar-Sicairos, P. Karvinen, E. Färm, E. Härkönen, M. Ritala, A. Menzel, J. Raabe, and O. Bunk. X-ray ptychographic computed tomography at 16 nm isotropic 3D resolution. *Scientific Reports*, 4, 2014.

[141] A. Diaz, P. Trtik, M. Guizar-Sicairos, A. Menzel, P. Thibault, and O. Bunk. Quantitative x-ray phase nanotomography. *Physical Review B*, 85(2): 020104, 2012.

[142] M. Stampanoni, R. Mokso, F. Marone, J. Vila-Comamala, S. Gorelick, P. Trtik, K. Jefimovs, and C. David. Phase-contrast tomography at the nanoscale using hard x rays. *Physical Review B*, 81(14): 140105, 2010.

[143] F. Pfeiffer, O. Bunk, C. David, M. Bech, G. Le Duc, A. Bravin, and P. Cloetens. High-resolution brain tumor visualization using three-dimensional x-ray phase contrast tomography. *Physics in Medicine and Biology*, 52(23): 6923, 2007.

[144] A. C. Kak, and M. Slaney. *Principles of computerized tomographic imaging*. IEEE Press, New York, 1988.

[145] B. P. Flannery, H. W. Deckman, W. G. Roberge, and K. L. D'AMICO. Three-dimensional X-ray microtomography. *Science*, 237(4821): 1439-1444, 1987.

[146] J. Radon. 1.1 über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten. *Classic Papers in Modern Diagnostic Radiology*, 5, 2005.

[147] R. N. Bracewell. Strip integration in radio astronomy. *Australian Journal of Physics*, 9(2): 198-217, 1956.

[148] R. A. Crowther, D. J. DeRosier, and A. Klug. The reconstruction of a three-dimensional structure from projections and its application to electron microscopy.

Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 317(1530): 319-340, 1970.

[149] H. Stark, J. Woods, I. Paul, and R. Hingorani. Direct Fourier reconstruction in computer tomography. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(2): 237-245, 1981.

[150] L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *JOSA A*, 1(6): 612-619, 1984.

[151] D. Attwood. *Soft x-rays and extreme ultraviolet radiation: principles and applications*. Cambridge University Press, Cambridge, 2007.

[152] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup. Efficient subpixel image registration algorithms. *Optics Letters*, 33(2): 156-158, 2008.

[153] R. M. Goldstein, H. A. Zebker, and C. L. Werner. Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science*, 23(4): 713-720, 1988.

[154] J. Frank, and B. F. McEwen. Alignment by cross-correlation. In Electron Tomography: 205-213, Springer, US, 1992.

[155] D. Ress, M. L. Harlow, M. Schwarz, R. M. Marshall, and U. J. McMahan. Automatic acquisition of fiducial markers and alignment of images in tilt series for electron tomography. *Journal of Electron Microscopy*, 48(3): 277-287, 1999.

[156] S. G. Azevedo, D. J. Schneberk, J. P. Fitch, and H. E. Martz. Calculation of the rotational centers in computed tomography sinograms. *IEEE Transactions on Nuclear Science*, 37(4): 1525-1540, 1990.

[157] P. R. Bevington, D. K. Robinson, J. M. Blair, A. J. Mallinckrodt, and S. McKay. Data reduction and error analysis for the physical sciences. *Computers in Physics*, 7(4): 415-416, 1993.

[158] O. L. Krivanek, P. H. Gaskell, and A. Howie. Seeing order in amorphous materials. *Nature*, 262: 454-457, 1976.

[159] A. Howie. High resolution electron microscopy of amorphous thin films. *Journal of Non-Crystalline Solids*, 31(1-2): 41-55, 1978.

[160] P. A. Midgley, and R. E. Dunin-Borkowski. Electron tomography and holography in materials science. *Nature Materials*, 8(4): 271-280, 2009.

[161] J. Vila-Comamala, A. Diaz, M. Guizar-Sicairos, A. Mantion, C. M. Kewish, A. Menzel, O. Bunk and C. David,. Characterization of high-resolution diffractive X-ray optics by ptychographic coherent diffractive imaging. *Optics express*, 19(22), 21333-21344, 2011.

[162] D. Claus, D. J. Robinson, D. G. Chetwynd, Y. Shuo, W. T. Pike, J. De J. José, and J. M. Rodenburg. Dual wavelength optical metrology using ptychography. *Journal of Optics*, 15(3): 035702, 2013.

[163] H. Yang, R. N. Rutte, L. Jones, M. Simson, R. Sagawa, H. Ryll, M. Huth T. J. Pennycook, M.L.H. Green, H. Soltau, Y. Kondo, B. G. Davis and P. D. Nellist. Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures. *Nature Communications*, 7 (2016).

[164] S. M. Solís, F. Mendoza Santoyo, and M. del Socorro Hernández-Montes. 3D displacement measurements of the tympanic membrane with digital holographic interferometry. *Optics Express*, 20(5): 5613-5621, 2012.

[165] D. Boss, , A. Hoffmann, B. Rappaz, C. Depeursinge, P. J. Magistretti, D. Van de Ville, and P. Marquet. Spatially-resolved eigenmode decomposition of red blood cells membrane fluctuations questions the role of ATP in flickering. *PLoS One*, 7(8): e40667, 2012.

[166] D. B. Williams, C. B. Carter and P. Veyssiere. Transmission electron microscopy: a textbook for materials science. *MRS Bulletin-Materials Research Society*, 23(5): 47, 1998.

[167] E. H. R. Tsai, I. Usov, A. Diaz, A. Menzel and M. Guizar-Sicairos. X-ray ptychography with extended depth of field. *Optics Express*, 24(25): 29089-29108, 2016.

[168] S. Dong, P. Nanda, R. Shiradkar, K. Guo and G. Zheng. High-resolution fluorescence imaging via pattern-illuminated Fourier ptychography. *Optics Express*, 22(17): 20856-20870, 2014.