



The
University
Of
Sheffield.

UNIVERSITY OF SHEFFIELD

Large Field of View Electron Ptychography

Shaohong Cao

Supervised by

Prof. John M Rodenburg and Dr. Andrew M Maiden

A thesis submitted for the degree of Doctor of
Philosophy

August 2017

Abstract

Electron ptychography can overcome the limits of the conventional electron microscopy in terms of both resolution and phase quantitative measurements. There are two ways to implement ptychography with electrons. One employs a focused probe and the other uses a large probe. The advantage of focused probe electron ptychography is allowing to analyse spectrum while collecting the data. The biggest advantage of large probe electron ptychography is much larger field of view with the same scanning positions. In this thesis, we investigate the applications of the large probe ptychography in three modes, which are a transmission electron microscope in the selected area diffraction mode (SAD ptychography), a scanning electron microscope in the transmission mode (SEM ptychography), and a scanning transmission electron microscope (STEM ptychography).

The thesis includes the detailed experimental procedures to collect ptychographic data in the three modes, as well as the investigation and evaluation of the experimental parameters. It presents extensive experimental data and results, which includes the decomposition of a partially coherent electron source via the modal decomposition ptychography with the SAD ptychography, the improvement of the delocalization issue with the SEM ptychography, and the atomic resolution reconstruction with the STEM ptychography. The challenges of the implementation and the reconstruction of ptychography in the three modes are also discussed. The main achievement of the thesis is the modal decomposition of matter wave, which has never been done before.

Declaration

I confirm that all the work in this thesis is my own research. I confirm that all the work in this thesis is original except where the reference is explicitly given.

Shaohong Cao

15/08/2017

Acknowledgements

Now, it comes to the end of my PhD. I was full of joy the moment when I finished the last chapter of this thesis and put the whole thesis together. A lot of feeling has come to me and a lot of people I want to express my acknowledgments to.

I would like to express my sincere gratitude to my supervisor Prof. John Rodenburg. I have received tremendous support from John during my PhD. John has transferred me not only a vast amount of knowledge about the scientific research, but also a solid attitude to conduct scientific research. It has been a great pleasure to work with John together; he is always calm, patient and very easy-going; he is a supervisor who would like to spend time with students. We have had numerous discussion in the pub called Red Deer, or at the dining table in some Café or restaurant. John's influence on me from day by day, from conversation by conversation; it will stay with me for the rest of my life.

I would like to express my special thanks to my second supervisor, Dr. Andrew Maiden. I was Andy's first PhD student, so I have got the special attention that the first 'child' always has. In my first year, Andy has spent a lot of time to supervise my algorithms and to make sure that I have an extensive literature review. Andy has given me great support in the

research projects in this thesis. He is a supervisor who involved experiments with students together, debugged algorithms with students together, who supports students a lot and is very patient in supervising. It is my great luck to have supervisors like John and Andy.

I would like to express my thanks to my colleague Dr. Peng Li. Peng has helped me considerably. Peng sat next to me in the office; I cannot count how many times in a day that I just turned my head to him, to begin a question, a discussion or a complaint. Peng has given me a lot of very useful suggestion on coding, data processing and the experimental design. He is like a third supervisor to me.

I would like to express my special thanks to Dr. Ian Ross for his patient and detailed training on the JEOL R005. I would like to express my special thanks to Dr. Pieter Kok, who has given the exhaustive nomenclature of the quantum mechanics derivation of the modal decomposition.

I would like to express my thanks to my previous colleagues Dr. Darren Batey, Dr. Tega Edo and Dr. Francis Sweeney, for the helpful conversations and the encouragements. I would like to give my special thanks to my colleagues Danny Johnson and Samuel McDermott for the discussion we have had in the office, and especially for your grammar review on Chapter 4 and Chapter 5 of this thesis. It is really a pleasant experience to be working in the office with you people together.

I would like to express my thanks to my friends and my family, thank you for giving me the strength when sometimes I nearly ran out, for your being there to share my joy, for encouraging me when I felt I was small, for helping me clear my mind when I felt lost.

Last but not least, I am truly grateful to Phase Focus Ltd. and the Electrical and Electronic Engineering department of the University of Sheffield for funding my PhD. Phase Focus Ltd

also gave me quite significantly technical support during my first year. I sincerely appreciate it.

List of Symbols

P the complex illumination/probe function

O the complex object transmission function

P_e extra mode of the probe function

O_e extra mode of the object function

L camera length

ψ complex exit wave

Ψ the Fourier transform of ψ

\mathcal{F} Fourier transform

λ wavelength

h Planck's constant

m_0 rest mass of electron

d lattice spacing

R the radius of the diffraction ring

x, y real space coordinates

\mathbf{u}, \mathbf{v} Fourier space coordinates

\mathbf{H} transfer function

\mathbf{j} register of scanning positions

Γ mutual coherence function

α object function update scale

β probe function update scale

\mathbf{G} Gaussian profile

I_{coh} coherent diffraction intensities

I_m partially coherent diffraction intensities

\mathbf{c} detector pedestal

Table of Contents

Abstract	2
Declaration	4
Acknowledgements	6
List of Symbols	9
Chapter 1 Introduction	16
1.1 Motivation	16
1.2 Thesis Outline	20
Chapter 2 Background	23
2.1 The Fourier transform and its properties	23
2.2 Electrons and electron waves	26
2.2.1 Definition of the electron wave	26
2.2.2 Description of electron waves and wave propagation	27
2.2.3 Diffraction approximations	29
2.3 Partially Spatially Coherent illumination	34
2.3.1 van Cittert-Zernike theorem	35

2.3.2 Density matrix.....	36
2.4 Wave front aberrations and the transfer function	39
2.4.1 Perfect lens function	39
2.4.2 Aberrations.....	41
2.5 Electron specimen interaction.....	44
2.5.1 Elastic scattering and inelastic scattering	44
2.5.2 Diffraction.....	45
2.5.3 Weak Phase Object Approximation.....	48
2.5.4 WPOA image	48
2.5.5 Contrast transfer function	50
2.6 Electron microscopy	52
2.6.1 Why use an electron microscope	52
2.6.2 The components of an electron microscope.....	55
2.6.3 TEM and STEM.....	61
2.7 Ptychography	70
2.7.1 Introduction.....	70
2.7.2 Review	74
2.7.3 Iterative Phase Retrieval Algorithms	81
2.8 Conclusion	102
Chapter 3 Ptychography in SAD mode.....	103
3.1 Modal decomposition of a propagating matter wave.....	103

3. 2 Collecting SAD ptychographic data	132
3.2.1 Specimen.....	132
3.2.2 Calibration.....	132
3.2.3 Alignment	137
3.3 Theoretical evaluation on modes reconstruction	138
3.3.1 The properties of the eigenstates.....	138
3.3.2 Factors affecting the Reconstruction of the eigenstates.....	145
3.4 Supplementary experimental Results.....	150
3.4.1 Data description	150
3.4.2 Reconstruction results.....	151
3.5 How low counts can ptychography go?.....	162
3.5.1 Introduction.....	162
3.5.2 Experiment and results.....	164
3.5.3 Discussion	168
Chapter 4 Image feature delocalization in defocused probe ptychography	170
4.1 Introduction.....	170
4.2 Factors result in delocalization	173
4.2.1 Source transverse partial coherence.....	173
4.2.2 Detector point spread	182
4.2.3 Detector pedestal.....	184
4.2.4 Data truncation.....	190

4.2.5 Positions scaling error.....	193
4.3 Experimental result	197
4.4 Conclusion	202
 Chapter 5 Ptychography on a STEM	204
5.1 Introduction.....	204
5.2 Experimental details.....	215
5.3 Reconstruction and results	220
5.3.1 Camera length calibration	220
5.3.2 Scanning positions examination	222
5.3.3 Initial probe generation	222
5.3.4 Algorithms and reconstruction	224
5.4 Discussion	234
5.5 Conclusion	244
 Chapter 6 The convolution in ptychography	245
6.1 Introduction.....	245
6.2 The convolution in ptychography	248
6.3 Tolerance of Ptychography on breaking the convolution	254
6.3.1 Partial diffraction patterns.....	255
6.4 conclusion	266
 Chapter 7 Conclusion and future work	268
References.....	272

Appendix.....	288
TEM Image alignment	288
SAD diffraction alignment.....	293
collect data	294

Chapter 1 Introduction

1.1 Motivation

Nowadays, with the better quality of the electromagnetic lens and the aberration corrector, a transmission electron microscope can achieve an imaging resolution of 0.05nm (Williams and Carter, 2009). However, conventional electron microscopy has its limits in some applications. Firstly, the atomic resolution and location of the conventional electron microscopy is not reliable. The transfer function of a transmission electron microscope (TEM) is affected by the defocus, the aberrations and the coherence of the source; some image frequencies are lost in the transmission electron microscopy. The high angle annular dark field (HAADF) image of a scanning transmission electron microscope (STEM) gives relatively true image of the sample, however, only the electrons that are scattered with the selected angles are used to form the image; light atoms are hard to observe. Moreover, the size of the focused beam spot, which determines the resolution of the HAADF image, is affected by the aberrations of the lens, the source coherence and the convergence angle. The scanning stability of the beam will affect the HAADF image directly. An accurate measure of the atom location is very important to analyse the properties of some materials, for example, some catalysts that are used in the automobile exhausts, they are nanoparticles (Pt, Co), which has similar size to the sample we

test in this thesis; the efficiency of catalyst depends on the surface area of the nanoparticles and their orientations. For steel, a slightly different chemical composition or adding a tiny portion of another element may significantly affect its mechanical properties. In these cases, it will be greatly helpful for the material scientist if there is a method to obtain a true image of the atomic composition of the material.

The second limit of the conventional electron microscopy is quantitative phase measurement, for example, quantitative characterization of the magnetic tunnel junctions in the magnetic memory cells (Gallagher et al., 1997), defect detection of p-n junctions in semiconductor manufacturing (Volkl et al., 1999), and the strain measurement in some electronic devices (Hytch et al., 2008). When high-energy electron beam goes through the sample, these structures barely change the energy of the electron beam; but they do change the scattering direction of the electrons to a certain degree (Williams and Carter, 2009; Spence, 2003); so, the field is ‘electron beam transparent’ in the intensity image. Another property we cannot obtain from the direct transmission electron image is the three-dimensional atomic structure of the sample. Even though the samples of the transmission electron microscope are usually very thin, most of them cannot be single-layer atoms. The transmission image projected on the detector is an averaged image over the sample thickness (Frank, 2006) through tens or hundreds of atoms.

Electron ptychography (Rodenburg, 1989; Rodenburg, McCallum and Nellist, 1993; Rodenburg, 2004) is a technique that can thoroughly and quantitatively image a material beyond atomic resolution and location, phase characterization and three-dimensional structure, for an early review see (Rodenburg, 2008). It is a diffraction imaging technique which uses a localised electron beam scanning across the sample and collects diffraction patterns at each scanning position, then reconstructs the complex probe function and the complex illumination function by inverse calculation. Electron ptychography has the

advantages that 1) it can retrieve both the object and the illumination (Thibault et al., 2008; Maiden and Rodenburg, 2009); this is important to make the measurement of the sample not influenced by the illumination condition; 2) its resolution is not limited by the transfer function of the lenses (Humphry et al., 2012); this is important to achieve good resolution when there is no aberration correctors in an electron microscope; 3) it can reconstruct weak or pure phase object (Yang et al., 2016); this is important to precisely measure some light atoms like the carbon nanotube (Yang et al., 2016), field or strain; 4) it has an unlimited field of view; not like holography (Gabor, 1948) or propagating series reconstruction (Ishizuka and Allman, 2005), which have a limited field of view, the field of view of ptychography is unlimited; it is able to achieve a gigapixel reconstruction as well as high resolution (Zheng et al., 2013); 5) it can reconstruct both the two-dimensional and the three-dimensional structure of the sample (Dierolf et al., 2010; Holler et al., 2017).

There are two ways to do ptychography with electron sources; one way is to focus the electron beam on the sample and scan the beam across the sample densely (Rodenburg and Nellist, 1993; Nellist et al., 1995; H. Yang et al., 2015, 2016 and 2017), which is called focused probe ptychography in this thesis; the other way is to use a relatively large probe and scan across the sample with a large step size (Rodenburg and Faulkner, 2004; Humphry et al., 2012; Maiden et al., 2015), which is called large probe ptychography in this thesis. Compared with focused probe ptychography, large probe ptychography has several advantages. Firstly, large probe ptychography processes far fewer data (Putkunz ey al., 2011; Yang et al., 2017); with a same field of view, the size of the data focused probe needs is about 1500 times of what large probe ptychography needs. Secondly, large probe ptychography brings less radiation damage to the sample (Pelz et al., 2017). Much smaller data and less radiation damage is very important for three dimensional ptychographic imaging via a tomographic configuration. Thirdly, large probe ptychography retrieves both the illumination function and

the object function; the reconstruction of the object is free of illumination aberrations or partial coherence. Finally, large probe ptychography employs iterative algorithms for the reconstruction. It is flexible to correct the imperfections in the experiments, for example, the scanning positions error or the data noise. This thesis investigates large probe ptychography on electron microscopes.

To date, there are just a few publications with the large probe electron ptychography (Humphry et al., 2012; Putkunz et al., 2012; D’Alfonso et al., 2014; Maiden et al., 2015). Even though resolution beyond the information limit and large field of view accurate phase imaging has been reported, there are still some issues existing in these published results, for example, defects caused by too few diffraction patterns (Putkunz et al., 2012; D’Alfonso et al., 2014) and the gold fringes delocalization in the reconstruction (Humphry et al., 2012).

This thesis includes the implementation of large probe electron ptychography with three setups – plane wave illumination using the transmission electron microscope selected area diffraction mode (SAD), convergent beam illumination with the scanning transmission electron microscope (STEM) and convergent beam illumination with scanning electron microscopy in the transmission mode (SEM). It investigates the issues that exist in large probe electron ptychography including the influence of the source partial coherence, detector response and noise, information lost in the diffraction patterns, and errors in the camera length and scanning positions, and aims to solve these problems via either the experiment or the algorithm. In this thesis, we also try to determine a minimum dose for large probe electron ptychography. The main achievement of this thesis is that, for the first time, modal decomposition of a propagating matter wave was measured experimentally with electron ptychography.

1.2 Thesis Outline

In Chapter 2, all the background of the application and the reconstruction of ptychography with electron waves has been included. There are three sections in the chapter. The first section introduces the optics involved in ptychography, which is necessary to understand how ptychography works. It includes the electron wave, electron wave propagation, electron wave diffraction, electron wave interaction with the samples and Fourier optics. The second section is about conventional electron microscopy, which is necessary to understand the experimental setup of ptychography. We talk about the functions of the electron source, the lenses, and the detectors in the electron microscopes, the contrast transfer function and the transmission electron microscope and the scanning transmission electron microscope (TEM/STEM) in the imaging mode and the diffraction mode. The third section reviews the electron ptychography literature and the development of ptychographic iterative reconstruction algorithms.

Chapter 3 presents an application of electron ptychography in the selected area diffraction mode of a transmission electron microscope (TEM). We decomposed a partially coherent electron source into a set of independent modes via a modal decomposition algorithm. This was the first time that the coherence function of a propagating matter wave was completely retrieved. The result was published in the journal Physical Review A. The supplementary part of this chapter details the procedures of aligning the TEM in the SAD mode and collecting the ptychographic data in the SAD mode. We represent the complete results when changing the parameters of the experiment. We conclude that the count limit of the SAD ptychography can be as low as $5 e^-/\text{\AA}^2$. The modes and their relationship with the experimental setup have also been explored in simulation.

Chapter 4 diagnoses the delocalization issue in the defocused convergent probe ptychographic reconstruction based on simulation and experimental data which was collected

on a 30 KeV scanning electron microscope (SEM) by Phase Focus Ltd. We explore the influence of the transverse partial coherence of the electron source, the detector point spread, the detector pedestal, the truncated diffraction patterns and the scanning position global scaling error on the fringes reconstruction delocalization. We find that each of these factors has a different effect. We propose two methods that can effectively improve the quality of the reconstruction according to each factor.

Chapter 5 is about the implementation of ptychography on a scanning transmission electron microscope (STEM). In this chapter, we explore the possible setups for defocused convergent beam ptychography on a STEM with an accelerating voltage of 300 KeV. The detailed procedure for collecting ptychographic data in STEM is included and the author's experience of tuning the experimental parameters and the difficulties encountered. Various data has been collected. The imperfect reconstructions from two sets of data are presented. The chapter also explores the possible reasons for the problems existing in the reconstruction.

Chapter 6 is a theoretical chapter to study the convolution in ptychography. We prove that ptychography relies on the overlap both in real space and in reciprocal space using simulations. Then we explore the tolerance of the iterative ptychographic reconstruction when breaking the convolution in the reciprocal space by three means 1) have a gap between the interference area, 2) add a noise pedestal to the data and 3) Poisson noise.

Chapter 7 concludes all the work that has been done in this thesis and the research that needs to be investigated in the future.

The arrangement of each chapter is slightly different. In Chapter 3, the published paper was inserted as the first section; this is a format recently introduced by the University of Sheffield. The supplement of the paper and the low counts experiments form the other sections. The

contents of Chapter 4, Chapter 5, and Chapter 6 are arranged as scientific papers, with an introduction, experiment and results, and discussion sections.

Chapter 2 Background

2.1 The Fourier transform and its properties

Throughout this thesis, we use Fourier transform extensively. The whole simulation of the optics propagation and the diffraction are based on the Fourier transform, and some of the Fourier transform theorems (Goodman, 2004). So, in the first section of this thesis, we give the general introduction of the Fourier transform and its properties.

The analytic Fourier transform of a two-dimensional function $g(x, y)$, where x, y are the coordinates that define real space, is

$$G(f_X, f_Y) = \iint_{-\infty}^{+\infty} g(x, y) e^{-2\pi i(f_X x + f_Y y)} dx dy, \quad (2.1.1)$$

where f_X, f_Y are the coordinates which define frequency space.

The inverse Fourier transform is given by

$$g(x, y) = \iint_{-\infty}^{+\infty} G(f_X, f_Y) e^{2\pi i(f_X x + f_Y y)} df_X df_Y. \quad (2.1.2)$$

Eq. (2.1) and (2.2) can be denoted as

$$G(f_X, f_Y) = \mathcal{F}\{g(x, y)\}, \quad (2.1.3)$$

$$g(x, y) = \mathcal{F}^{-1}\{G(f_X, f_Y)\}. \quad (2.1.4)$$

The Fourier transform is an operator which transforms a function from its description in real space coordinates x, y to its representation **in frequency coordinates** f_X, f_Y . The functions which can be applied with Fourier transform must satisfy these conditions:

- 1) The function must be absolutely integrable over the infinite range of x and y ;
- 2) The function must have only a finite number of discontinuities;
- 3) The function must have no infinite discontinuities.

When we read a function (signal) into a computer, the signal will not be a continuous function anymore, but has to be decomposed into a series of discrete points. The process of decomposing a continuous function into a series of points is called sampling. The smaller the sampling interval means higher sampling frequencies, which will preserve the information of the original signal better. However, in signal processing, higher frequencies need larger storage space of the computer and more expensive computing. The lowest frequency needed when sampling a signal and preserve the information equals double of the frequency of the signal.

$$\Delta x < \frac{1}{2B_x}, \Delta y < \frac{1}{2B_y}, \quad (2.1.5)$$

where $\Delta x, \Delta y$ are the sampling intervals in x, y directions, B_x, B_y are the bandwidths of the continuous signal in x, y directions. This is called the Shannon-Nyquist sampling theorem.

The Fourier transform of discrete function $g(m\Delta x, n\Delta y)$ is

$$G(u\Delta f_X, v\Delta f_Y) = \sum_{m=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} g(m\Delta x, n\Delta y) e^{-2\pi i(u\Delta f_X m\Delta x + v\Delta f_Y n\Delta y)}, \quad (2.1.6)$$

where $\Delta f_X, \Delta f_Y$ are the sampling intervals of the frequency space.

There are several theorems of Fourier transform that are relevant to optics, as follows:

Linearity

$$\mathcal{F}\{af(x, y) + bg(x, y)\} = a\mathcal{F}\{f(x, y)\} + b\mathcal{F}\{g(x, y)\}. \quad (2.1.7)$$

Similarity

$$\mathcal{F}\{g(ax, by)\} = \frac{1}{|ab|} G\left(\frac{f_X}{a}, \frac{f_Y}{b}\right). \quad (2.1.8)$$

Shift

$$\mathcal{F}\{g(x - a, y - b)\} = G(f_X, f_Y)e^{-2\pi i(a f_X + b f_Y)}. \quad (2.1.9)$$

Parseval's theorem

$$\iint |g(x, y)|^2 dx dy = \iint |G(f_X, f_Y)|^2 df_X df_Y. \quad (2.1.10)$$

Convolution

$$\mathcal{F}\{\iint f(\zeta, \eta)g(x - \zeta, y - \eta)d\zeta d\eta\} = F(f_X, f_Y)G(f_X, f_Y). \quad (2.1.11)$$

2.2 Electrons and electron waves

2.2.1 Definition of the electron wave

Most of the work presented in this thesis is dealing with the electron waves (Chapter 3, 4 and 5). In this section, we give some brief information about the physical properties of electrons and electron waves.

An electron is an elementary particle, with a charge of $1.602 \times 10^{-19} C$, rest mass of $9.109 \times 10^{-31} kg$. Electrons show both particle and wave characteristics. From Broglie's wave-particle duality, we know that the wavelength of massive particle is related to the momentum of the particle by

$$\lambda = \frac{h}{p}, \quad (2.2.1)$$

where λ is the wavelength of the particle, p is its momentum and h is Planck's constant.

If the electron is accelerated through a potential of E from rest, the kinetic energy of the electron will equal E , thus we have

$$E = \frac{1}{2}m_0v^2, \quad (2.2.2)$$

where m_0 is the mass of the electron and v is its velocity.

The momentum of the electron p is the product of the mass and the velocity:

$$p = m_0v, \quad (2.2.3)$$

From Eq. 2.2.1, Eq. 2.2.2 and Eq. 2.2.3, we have

$$\lambda = \frac{h}{\sqrt{2eVm_0}}. \quad (2.2.4)$$

This is the equation to calculate the wavelength of accelerated massive particles. We can see that as we increase the kinetic energy of the electron, its wavelength becomes shorter.

The calculation of the electron wavelength in Eq. (2.2.4) does not consider the effect of the relativity. However, when the potential energy is higher than 100 keV, the velocity of the electron v is greater than half of the speed of light $c = 2.998 \times 10^8 \text{ m/sec}$. In this case, we cannot ignore the relativistic effect, so the Eq. (2.2.4) is modified as

$$\lambda = \frac{h}{\sqrt{2eVm_0(1+\frac{eV}{2m_0c^2})}}. \quad (2.2.5)$$

For example, when the electrons are accelerated by the potential energy of 30 keV, 80 keV, 200 keV and 300 keV, the corresponding wavelengths are $6.98 \times 10^{-12} \text{ m}$, $4.18 \times 10^{-12} \text{ m}$, $2.55 \times 10^{-12} \text{ m}$ and $1.97 \times 10^{-12} \text{ m}$ respectively.

2.2.2 Description of electron waves and wave propagation

We use electron waves as the illumination for the experiments we have done in this thesis. In this section, we briefly introduce the equations of the plane electron waves and the spherical electron waves, which are the two kinds of wavefronts that are often used in this thesis. We also introduce the propagation of the two kinds of wavefronts in the free space.

The electromagnetic wave equation is derived from Maxwell's equations (Born and Wolf, 1999), which can be rearranged as

$$\frac{\partial^2}{\partial x^2} \vec{E}(x, t) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E}(x, t) = 0, \quad (2.2.6)$$

where $\vec{E}(x)$ is the electric field along direction x , c is the velocity of light.

Maxwell's equations describe the wave equation of photons. One of its solutions is a plane wave.

$$\vec{E}(x, t) = \vec{E}_0 e^{i(kx - \omega t)}. \quad (2.2.7)$$

Schrodinger proposed the wave equation of the massive particles as

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x, t) \psi(x, t), \quad (2.2.8)$$

where $V(x, t)$ is the potential of the atoms.

Schrodinger equation is derived from these three relationships:

Planck's equation

$$E = \hbar\omega, \quad (2.2.9)$$

De Broglie's hypothesis

$$\vec{p} = \hbar\vec{k}, \quad (2.2.10)$$

and the classical momentum-energy relation

$$E = \frac{p^2}{2m} \Rightarrow \hbar\omega = \frac{\hbar^2 k^2}{2m}, \quad (2.2.11)$$

where $\hbar = \frac{h}{2\pi}$, m is the mass of the particle.

In this thesis, the potential energy of the particle in the physical world is not changing with time, so, we can separate t and x when we solve the Schrodinger equation (2.2.8). The time-independent Schrodinger equation can be written as

$$\hat{H}\psi_E(x) = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi_E(x) = E\psi_E(x), \quad (2.2.12)$$

where E is the energy of the system.

The plane wave solution of a point source, which is a source that has in infinite small spatial dimension, in the Cartesian coordinates is

$$\psi(r, t) = \psi_0 e^{i(kr - \omega t)}. \quad (2.2.13)$$

The spherical wave solution of a point source in the polar coordinates is

$$\psi(r, t) = \frac{\psi_0}{r} e^{i(kr - \omega t)}. \quad (2.2.14)$$

The plane waves and the spherical waves are the two solutions of the Schrodinger equation that we have applied in this thesis.

2.2.3 Diffraction approximations

In this section, we introduce the propagation of the plane wavefront and the spherical wavefront in the free space. We give the analytical propagation equation and its approximations at different conditions. More details of the derivations that follow can be found elsewhere (e.g. Born and Wolf, 1999).

2.2.3.1 Huygens-Fresnel principle

According to Huygens-Fresnel principle, a wave front can be taken as the composition of a series of points; each point is a new source emitting spherical wave, the next wave front is the tangency plane of the superposition of all these small wave fronts (as shown in Figure 2.2.1).

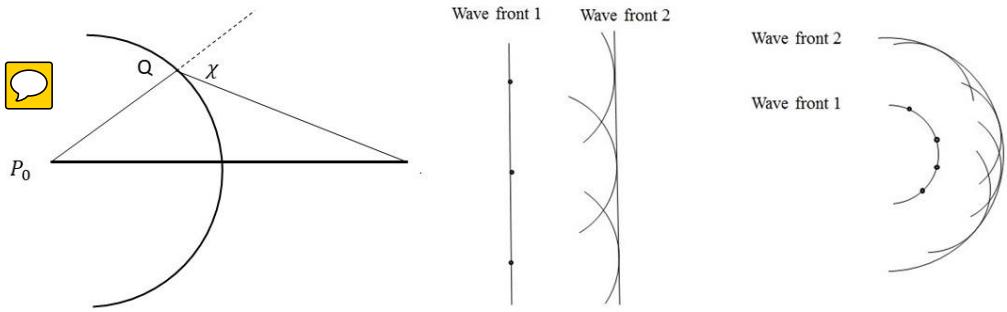


Figure 2.2.1 Huygens-Fresnel principle. The wave propagates from plane 1 to plane 2.

Supposing p_0 is a point source, at distance r there is a point p ; ψ_0 is the complex magnitude at the point p_0 ; the wave function at point p propagating from p_0 is

$$\psi(r) = \frac{\psi_0 e^{ikr}}{r}, \quad (2.2.15)$$

where $k = 2\pi/\lambda$ is the wave number. This equation means at the distance r , the magnitude of the wave decreases inversely to r . According to Huygens theory and superposition principle, the wave function propagating to point p from the previous wave front s is

$$\psi = -\frac{i}{\lambda} \psi(r) \int_s \frac{e^{iks}}{s} K(\chi) ds, \quad (2.2.16)$$

where s is the surface of the sphere, $K(\chi) = \frac{1}{2}(1 + \cos \chi)$.

2.2.3.2 Rayleigh-Sommerfeld diffraction

Thinking of a two-dimensional source plane (x, y) , the source distribution function is given by the function $u_1(x, y)$, and the plane (ξ, η) is an observation plane at the distance z (as shown in Figure 2.2.2).

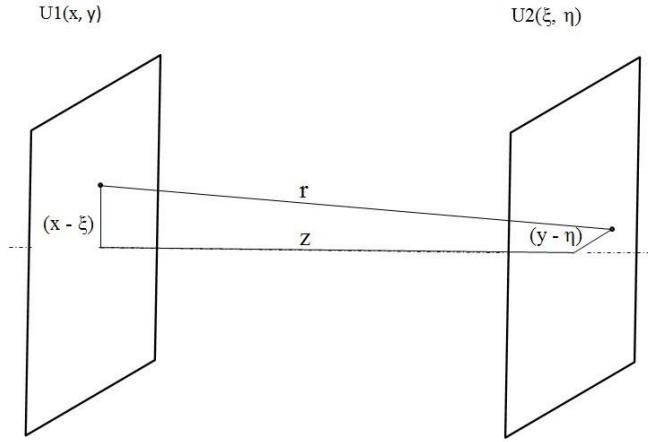


Figure 2.2.2 **U1** is the source plane, **U2** is the observing plane. **Z** is the distance between **U1** and **U2**. **r** is the distance from any point in plane **U1** to plane **U2**.

The Rayleigh-Sommerfeld expression of the field distribution function $u_2(\xi, \eta)$ at the observation plane is

$$U_2(\xi, \eta) = \frac{z}{j\lambda} \iint U_1(x, y) \frac{e^{ikr}}{r} dx dy, \quad (2.2.17)$$

$$r = \sqrt{z^2 + (\xi - x)^2 + (\eta - y)^2}, \quad (2.2.18)$$

which is derived from the Huygens-Fresnel principle. Here, r is the distance from an arbitrary point at the source plane to an arbitrary point at the observation plane.

The Rayleigh-Sommerfeld solution is a precise description on how wavefronts propagate from one plane to the next plane. The only prerequisite for Rayleigh-Sommerfeld description is $r \gg \lambda$, which is satisfied in most conditions for electrons.

2.2.3.3 Fresnel approximation

The Fresnel approximation of the wave propagation is a simplified form of the Rayleigh-Sommerfeld expression, by replacing the square root term with its polynomial expansion approximation according to the Binomial theorem:

$$(1 + x)^m = 1 + \frac{m!}{1!(m-1)!}x + \frac{m!}{2!(m-2)!}x^2 + \dots + \frac{m!}{(m-1)!1!}x^{m-1} + \frac{m!}{m!0!}x^m. \quad (2.2.19)$$

If we keep the first two terms to approximate in the transformation, we get r :

$$r = \sqrt{z^2 + (\xi - x)^2 + (\eta - y)^2} = z\left(1 + \frac{1}{2}\left(\frac{\xi - x}{z}\right)^2 + \frac{1}{2}\left(\frac{\eta - y}{z}\right)^2\right). \quad (2.2.20)$$

Replacing r in Eq. 2.2.17 with Eq. 2.2.20, we obtain the Fresnel approximation description:

$$U_2(\xi, \eta) = \frac{e^{jkz}}{j\lambda z} \iint U_1(x, y) e^{\frac{jk}{2z}[(\xi-x)^2 + (\eta-y)^2]} dx dy. \quad (2.2.21)$$

The prerequisite for Fresnel approximation is $\frac{w^2}{\lambda z} < 1$, where w is the half width of the aperture dimension in the source plane.

2.2.3.4 Fraunhofer approximation

If we expand the square term in the exponent part of Eq. 2.2.21, we get

$$U_2(\xi, \eta) = \frac{e^{jkz}}{j\lambda z} e^{\frac{jk}{2z}(\xi^2 + \eta^2)} \iint U_1(x, y) e^{\frac{jk}{2z}(x^2 + y^2)} e^{\frac{-jk}{z}(x\xi + y\eta)} dx dy. \quad (2.2.22)$$

When the distance between the source plane and the observation plane z is very far:

$$z \gg [\frac{k(x^2+y^2)}{2}]_{\max}, \quad (2.2.23)$$

then $e^{\frac{jk}{2z}(m^2+n^2)} \approx 1$, thus Eq. 2.2.22 becomes:

$$U_2(\xi, \eta) = \frac{\exp(j\lambda z)}{j\lambda z} \exp\left(j\frac{k}{2z}(\xi^2 + \eta^2)\right) \times \iint U_1(x, y) \exp[-j\frac{2\pi}{\lambda z}(x\xi + y\eta)] dx dy. \quad (2.2.24)$$

This is the Fraunhofer propagation approximation.

Comparing the Fraunhofer approximation (Eq. 2.2.24) with the Fourier transform

$$G(f_x, f_y) = \iint_{-\infty}^{\infty} g(x, y) \exp[-j2\pi(f_x x + f_y y)] dx dy, \quad (2.2.25)$$

we can find that the diffraction in the very far field U_2 can be taken as the Fourier Transform of the source plane field U_1 .

In the setup of Chapter 3, the detector is placed in the near field relative the sample, so we use the Fresnel diffraction approximation in the calculations. In Chapter 4, 5, and 6, the detector was placed in the far field relative to the sample, so we use the Fraunhofer diffraction approximation in the calculation.

2.3 Partially Spatially Coherent illumination

In this section, we discuss the spatial coherence and the temporal coherence of a source. The electron sources we have used in this thesis were not ideal sources. The electron sources have certain dimensions, so they are partially spatially coherent; the energy of the electrons emitted from the source is slightly different, so the electron sources are also partially temporally coherent.

The coherence degree of a source can be described by the mutual coherence function Γ (Zernike, 1938),

$$\Gamma_{12}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E_1(t)E_2^*(t - \tau)dt, \quad (2.3.1)$$

where E_1 and E_2 are any two measurements at the wavefronts of the source; τ is the time delay between the two measurements. The mutual coherence function expresses the correlation between any two components emitted from the source. We can use function γ to measure the coherence degree:

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1}\sqrt{I_2}}, \quad (2.3.2a)$$

$$I_1 = E_1 E_1^*, I_2 = E_2 E_2^*, \quad (2.3.2.b)$$

where I_1, I_2 are the intensities of each measurement. When $\gamma_{12} = 0$, it means the two waves are completely incoherent; when $\gamma_{12} = 1$, it means the two waves are fully coherent.

The coherence of the source determines the visibility of the diffraction patterns and images. For example, in Young's double-slit experiment, when the source is narrower, the diffraction fringes or rings will have better contrast.

There are two ways to characterize a partially coherent source; one is an approximation to measure the spatially partial coherence according to the van Cittert-Zernike theorem, and the other is mixed states theorem. In the following section, we will discuss the two methods respectively.

2.3.1 van Cittert-Zernike theorem

In the paper (van Cittert, 1934; Zernike, 1938), P.H. van Cittert and F. Zernike gave the concept of the degree of coherence; they gave the theorem that the coherence degree of any extended light source equals to the amplitude of a certain diffraction pattern, and that it will not be changed by the lens but will depend on the aperture. It was proved later that van Cittert-Zernike theorem could also be applied to matter wave (Taylor et al., 1994).

The expression of van Cittert-Zernike theorem in equation is that

$$\Gamma_{12}(u, v, 0) = \iint I(l, m) e^{-2\pi i(ul+vm)} dl dm, \quad (2.3.3)$$

where u, v are the x-distance and y-distance between the observation plane to the source plane; l, m are the direction cosines of a point on a distant source in the source plane; I is the intensity of the source. Eq. 2.3.3 points out that the coherence function of a source equals the Fourier transform of its intensities.

The van Cittert Zernike theorem was built on the assumptions that

- 1) the source is partial coherent;
- 2) the observing plane is distant from the source

$$R \gg \frac{D^2}{\lambda}, \quad (2.3.4)$$

where R is the distance between the source plane and the observing plane; D is the characteristic size of the observing area, and λ is the wavelength. This condition is the same with the Fraunhofer diffraction approximation condition.

- 3) the angular size of the source is small;
- 4) the source is quasi-monochromatic and two-dimensional;
- 5) the propagation medium is homogeneous.

The van Cittert-Zernike theorem indicates that the coherence of the source can be characterized from the visibility of the diffraction fringes, as what we can see from the Young's slit experiment.

$$\text{visibility} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (2.3.5)$$

2.3.2 Density matrix

2.3.2.1 Pure state

Supposing ψ is a coherent source, a coherent source ψ can be defined as a pure state $|\psi\rangle$. The density matrix of this pure state is (Wolf, 1982)

$$\rho = |\psi\rangle\langle\psi|, \quad (2.3.6)$$

The trace of the density matrix of a pure state is $Tr_\rho = 1$.

2.3.2.2 Mixed state from pure state

Supposing there are two sources ψ_a and ψ_b , the density matrix of the state mixed with these two sources (mixed state) is

$$\rho = \begin{bmatrix} \psi_a \\ \psi_b \end{bmatrix} [\psi_a^* \quad \psi_b^*] = \begin{bmatrix} \psi_a \psi_a^* & \psi_a \psi_b^* \\ \psi_b \psi_a^* & \psi_b \psi_b^* \end{bmatrix}. \quad (2.3.7)$$

If the two sources are completely incoherent $\psi_a \psi_b^* = 0$, the density matrix of the mixed state is

$$\rho = \begin{bmatrix} \psi_a \psi_a^* & 0 \\ 0 & \psi_b \psi_b^* \end{bmatrix}, \quad (2.3.8)$$

which is **the probabilistic addition** of the two pure states $\psi_a \psi_a^*$ and $\psi_b \psi_b^*$. After normalization, the trace of ρ can be calculated as

$$p_a = \frac{\psi_a \psi_a^*}{\psi_a \psi_a^* + \psi_b \psi_b^*}, \quad (2.3.9a)$$

$$p_b = \frac{\psi_b \psi_b^*}{\psi_a \psi_a^* + \psi_b \psi_b^*}, \quad (2.3.9b)$$

$$p_a + p_b = 1. \quad (2.3.9c)$$

2.3.2.3 Mixed state from mixed state

If the two sources are partially coherent: $\psi_a = \psi_1 + \psi_2$, $\psi_b = \psi_1 - \psi_2$, where ψ_1, ψ_2 are two pure states, then, $\psi_a \psi_b^* \neq 0$.

$$\begin{aligned} \rho &= \begin{bmatrix} \psi_1 + \psi_2 \\ \psi_1 - \psi_2 \end{bmatrix} [\psi_1^* + \psi_2^* \quad \psi_1^* - \psi_2^*] \\ &= \begin{bmatrix} (\psi_1 + \psi_2)(\psi_1^* + \psi_2^*) & (\psi_1 + \psi_2)(\psi_1^* - \psi_2^*) \\ (\psi_1 - \psi_2)(\psi_1^* + \psi_2^*) & (\psi_1 - \psi_2)(\psi_1^* - \psi_2^*) \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} (\psi_1 + \psi_2)(\psi_1^* + \psi_2^*) & \psi_1\psi_1^* - \psi_2\psi_2^* \\ \psi_1\psi_1^* - \psi_2\psi_2^* & (\psi_1 - \psi_2)(\psi_1^* - \psi_2^*) \end{bmatrix}, \quad (2.3.10)$$

Since ρ is a symmetric matrix, ρ can be transformed with singular value decomposition by

$$\rho = U\rho'V, \quad (2.3.11)$$

where U and V are unitary matrixes, ρ' is a diagonal matrix

$$\rho' = \begin{bmatrix} p_1\psi_1\psi_1^* & 0 \\ 0 & p_2\psi_2\psi_2^* \end{bmatrix}, \quad (2.3.12)$$

$p_1 \geq 0$; $p_2 \geq 0$; and $p_1 + p_2 = 1$.

In conclusion, any mixed state is the superposition of more than one pure states. Partially coherent sources can be projected into any dimensional coordinates, that is to be decomposed into several mixed states or pure states. To decompose a partially coherent source into several pure states is the minimum representation of the source; the coordinates where the pure states are presented are called as the eigen coordinates (Whitehead et al., 2009).

In this thesis, we have retrieved the coherence function of the electron sources with two methods, which are called the Gaussian blind convolution method and the modal decomposition method. The Gaussian blind convolution method is based on the van Cittert-Zernike theorem; the modal decomposition method is based on the mixed states theory.

2.4 Wave front aberrations and the transfer function

A freely propagating wave front is without any aberrations; its wave front can be approximated as a perfect polynomial profile. In an imaging system, we need to change the path of the wave through the lens. For example, we use the optical lens in the optical microscope or the magnetic field lens in an electron microscope to control the wave front. The lens will introduce some aberrations into the wave front. Aberration is one of the terms that determine the transfer function of an imaging system; it will affect the contrast and resolution of the images and diffraction patterns.

In this section, we discuss several common aberrations introduced by the lens: astigmatism, spherical aberration, coma and some other higher-order aberrations.

2.4.1 Perfect lens function

When the wave propagation meets a perfect lens, the lens does two things to the wave: 1) it gives the wave field a limited transmission area, which is usually a top hat function; we call it a lens pupil function; 2) it adds a phase curvature to the wave front. A perfect lens function can be described mathematically as (Goodman, 2004):

$$L(x, y) = A(x, y)e^{-\frac{jk}{2f}(x^2+y^2)}, \quad (2.4.1)$$

where $A(x, y)$ is the lens pupil function; k is the wave number; f is the focal length of the lens which indicates how strongly the lens can change the wave phase; and the negative sign means the convergent direction (positive means divergent) (Voelz, 2011). Supposing there is an object wave front $U_1(x, y)$, the image plane of U_1 is $U_2(\xi, \eta)$, which is formed by a lens with focal length equals f , as shown in Fig. 2.4.1:

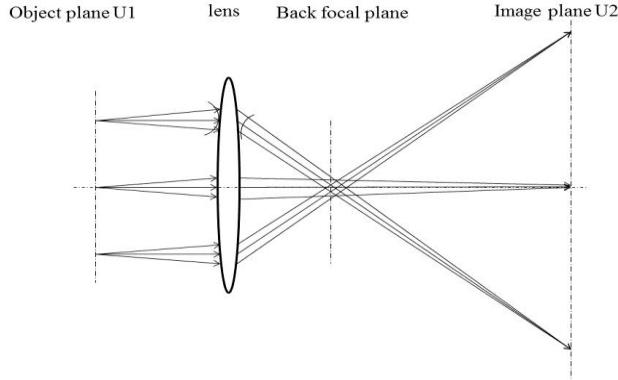


Figure 2.4.1 A perfect lens will make all the rays come out from a point source into an image of the source at the image plane, and all the parallel rays meet up at the back focal plane. The distance between the lens plane (thin lens) and the back focal plane is the focal lens f .

The wave immediately emitted from the lens is

$$U_1(x, y)L(x, y) = U_1(x, y)A(x, y)e^{-\frac{jk}{2f}(x^2+y^2)}. \quad (2.4.2)$$

According to the Fresnel approximation, the diffraction at the observation plane is

$$U_2(\xi, \eta) = \frac{e^{j k z}}{j \lambda z} \iint U_1(x, y)A(x, y) e^{-\frac{jk}{2f}(x^2+y^2)} e^{\frac{jk}{2z}[(\xi-x)^2+(\eta-y)^2]} dx dy, \quad (2.4.3)$$

when $z = f$,

$$\begin{aligned} U_2(\xi, \eta) &= \frac{e^{j k f}}{j \lambda f} \iint U_1(x, y)A(x, y) e^{-\frac{jk}{2f}(x^2+y^2)} e^{\frac{jk}{2f}(\xi^2+\eta^2+x^2+y^2-2(x\xi+y\eta))} dx dy \\ &= \frac{e^{j k f}}{j \lambda f} e^{\frac{jk}{2f}(\xi^2+\eta^2)} \iint U_1(x, y)A(x, y) e^{\frac{-jk}{f}(x\xi+y\eta)} dx dy, \end{aligned} \quad (2.4.4)$$

which is the Fraunhofer approximation propagation, also the Fourier transform of the source plane. We call this plane the back focal plane of a lens. We know that there are two situations where we can use the Fraunhofer approximation: the very far field or the back focal plane of a lens.

2.4.2 Aberrations

The common aberrations introduced by the lens are the spherical aberration, coma, and astigmatism. The other common aberration is called the chromatic aberration. It is not introduced by the lens, but by the energy spread of the source, which results in different wavelengths present in the beam.

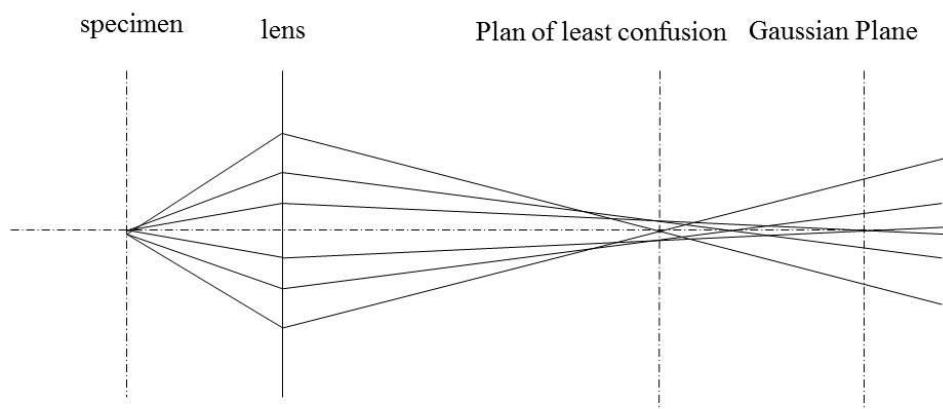


Figure 2.4.2 Spherical aberration. The edge part of the lens bends the rays stronger than the central part, which makes the rays come out from the edge part meet earlier. The Gaussian plane is where the beams from the central part meet. The plane of least confusion plane is where the ring has the smallest diameter.

Spherical aberration comes from the fact that the edge part of the lens changes the path of the beam more strongly than the central part of the lens does, as shown in Figure 2.4.2 (Williams and Carter, 2009). Astigmatism occurs when the lens function on the wave front is not cylindrically symmetrical – the lens bends the beam in one direction more strongly than in other direction (Figure 2.4.3(d)). A lens with coma will have the effect that makes the image of the source or the object look like it has a tail (Figure 2.4.3(c)). Ideally, the phase curvature of the lens function is a polynomial function(Eq.2.4.1). When the aberrations are introduced into the lens, the lens function will become

$$L(x, y) = A(x, y)e^{-jkw}, \quad (2.4.5)$$

Where w is the phase curvature function such that

$$\begin{aligned} w(x, y, u_0) = & w_d(x^2 + y^2) + w_{040}(x^2 + y^2)^2 + w_{131}u_0(x^2 + y^2)x + w_{222}u_0^2x^2 + \\ & w_{220}u_0^2(x^2 + y^2) + w_{311}u_0^3x, \end{aligned} \quad (2.4.6)$$

where x, y are the coordinates in the lens plane; u_0 is the dimension of the normalized image along u direction in the image plane coordinates u, v . The coefficients of each term are defined as (Voelz, 2010)

Defocus	Spherical	Coma	Astigmatism	Field Curvature	Distortion
w_d	w_{040}	w_{131}	w_{222}	w_{220}	w_{311}

Table 2.4.1. We number the name of the different orders of aberration in polar coordinates. $\rho = \sqrt{x^2 + y^2}$; $\rho \cos \theta = x$. So, spherical aberration $(x^2 + y^2)^2 = \rho^4$, which is named as w_{040} ; Astigmatism $u_0^2x^2 = u_0^2\rho^2 \cos \theta^2$, which is named as w_{222} .

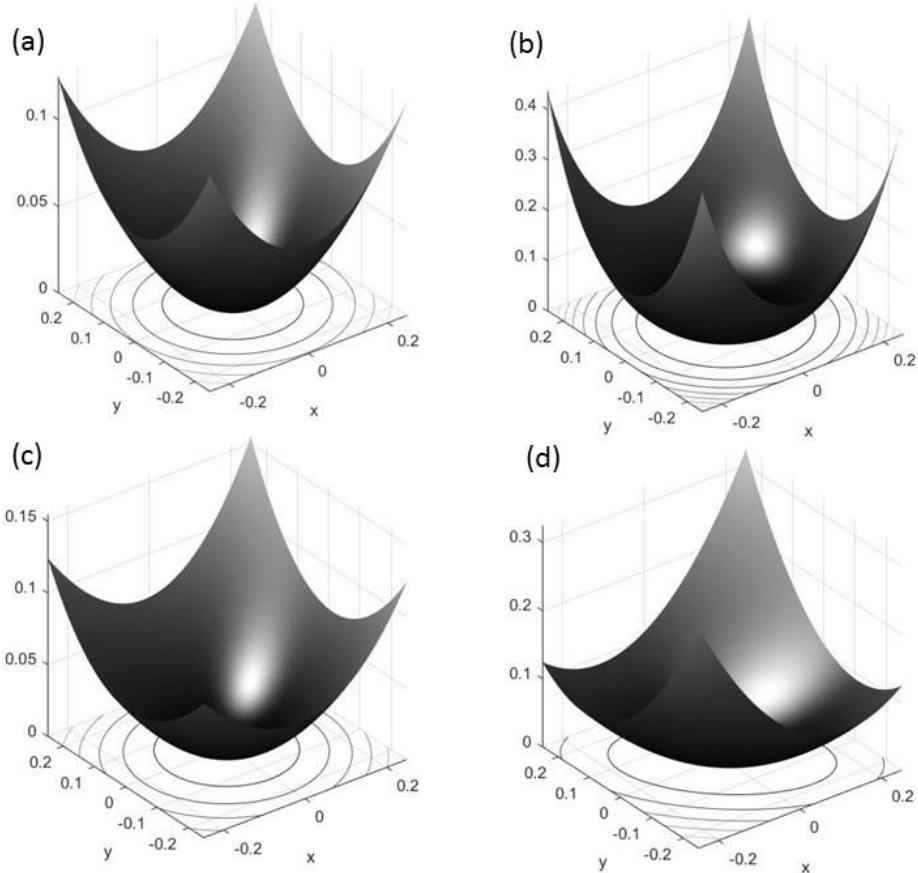


Figure 2.4.3 shows the wave fronts with aberrations. (a) is a defocused wave front; (b) is a defocused wave front with spherical aberration; (c) is defocused wave front with coma; (d) is defocused wave front with astigmatism. **x and **y** axes label frequencies and **z** axis labels phase.**

In the experiments of this thesis, the wavefronts generated by the transmission electron microscope or the scanning electron microscope were not perfect. There exist some aberrations in it. The most common aberrations we have observed in the images and the diffraction patterns are astigmatism and coma. We have observed that how these aberrations changed with the defocus and their influence on the reconstruction.

2.5 Electron specimen interaction

When the electrons beam goes through the sample, the scattering directions and the energy of the electrons may change because of the interaction between the electrons and the sample.

These scattered electrons carry the structure information of the sample. They are what we used in this thesis to retrieve the object function and the illumination function. In this section, we will discuss what may happen to the electrons beam when they interact with the sample.

2.5.1 Elastic scattering and inelastic scattering

The scattering phenomenon happens because of the Coulomb force between the incident electrons and the atomic potential or nuclei of the atoms.

The specimen we talk about here is a thin specimen that is used for transmission electron microscopy (TEM). After the electron beam goes through a thin specimen, there are several components of the electron exit waves: 1) Most of the electrons will go through the specimen without interaction with the atoms. This part of the exiting electron beam is called the direct beam. Both the energy and the direction of the direct beam does not change compared to the incident beam. 2) Some of the electrons have undergone elastic scattering with the atoms in the specimen. These electrons do not lose the energy, but they are scattered into different directions. Those which go through the outer layer of the electron cloud of the atoms inside the specimen (as shown in Figure 2.5.1), will interact with weak Coulomb force, thus they will be scattered by a small angle ($1^\circ - 10^\circ$); usually, these electrons remain coherent. Those which go through the inner layer of the electron cloud of the atoms inside the specimen will interact with stronger Coulomb force of the nucleus, thus they will be scattered by a large angle ($> 10^\circ$). These large angle elastic scattered electrons are incoherent. There are also some of the electrons which go so close to the nucleus of the atoms that they are back

scattered; these back scattered electrons are also incoherent. 3) Some of the electrons, when interacting with the atoms of the specimen, will lose some energy. We call this as inelastic scattering event. The inelastic scattering electrons usually have at a low angle ($< 1^\circ$) and are usually incoherent (Williams and Carter, 2009).

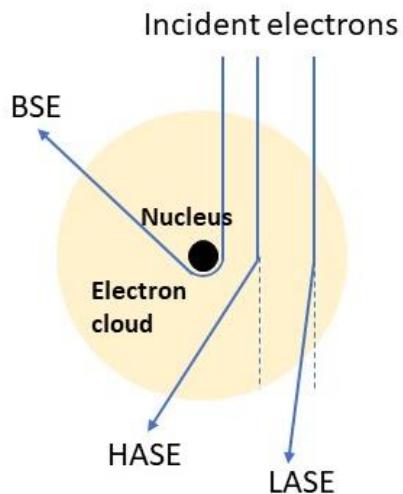


Figure 2.5.1 shows the potential paths when an electron goes through the specimen. Some electrons that go through close to the nucleus will be back scattered; some other electrons that go through the electron cloud will be scattered by different angles in the forward directions (HASSE is the high angle scattered electrons and LASE is the low angle scattered electrons).

Transmission electron microscopy mostly uses the signals from forward scattering, like the image or diffraction pattern viewed on the screen, or the detected the X-Ray spectrum or the energy-loss spectrum. This will be discussed in detail in the later chapters. The back scattered electrons are fundamental in scanning electron microscopy. As the specimen gets thicker, there are more back scattered electrons. So, the specimen used for scanning electron microscopy is thicker than the specimen used in the transmission electron microscopy.

2.5.2 Diffraction

The diffraction happens because of the wave characteristic of electrons. When the electron wave encounters an obstacle or a slit which is of comparable size with the electron wavelength, the diffraction occurs. Here we use two-point obstacles to illustrate the diffraction phenomenon. Supposing there are two points separated by a distance d (Figure 2.5.2), the path difference between the two scattered wavelets along the direction shown in the figure is $d \sin \theta$; the phase difference is $d \sin \theta / \lambda$; we define the wavelet from one point as

$$\psi_1 = A(x, t)e^{i\phi(x, t)}, \text{ where } A \text{ is the amplitude of the wave at a point } x \text{ and at time } t.$$

Then, the wave at the other point can be described as

$$\psi_2 = A(x, t)e^{i(\phi(x, t) + d \sin \theta / \lambda)}. \quad (2.5.1)$$

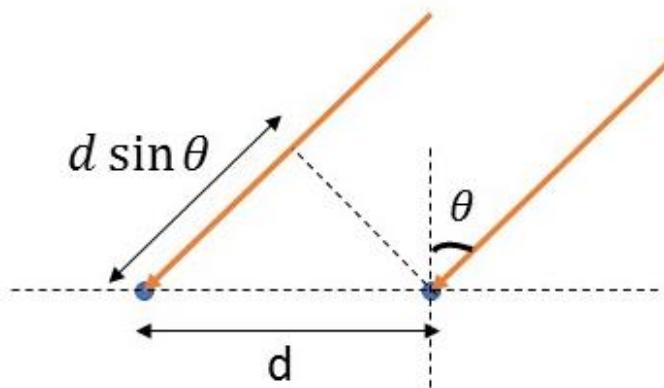


Figure 2.5.2 shows the path difference of incident plane wave with an angle θ when scattering by two atoms separated by distance d .

The diffraction wave behind the two points is the vector sum of the two scattered wavelets.

When

$$\frac{d \sin \theta}{\lambda} = 2n, \quad (2.5.2)$$

Where n is integer the two scattered wavelets will be in phase again, the amplitude of the diffracted wave will be $2A$; when $\frac{d \sin \theta}{\lambda} = n$, the two scattered waves will be totally out of phase, the amplitude of the diffracted wave will be 0.

In the transmission electron microscope, diffraction happens when the electron wave goes through the specimen. Depending on where the detector is placed, we can classify the diffraction pattern, which is the intensity of the diffraction wave, as either near field diffraction patterns (Fresnel diffraction) or far field diffraction patterns (Fraunhofer diffraction). The Fresnel diffraction pattern of an aperture illuminated with the parallel beam is shown in Figure 2.5.3(a), it is like a shadow image of the aperture, but we can see the Fresnel fringes at the edge. The Fraunhofer diffraction pattern of an aperture is shown in Figure 2.5.3(b). The rings in the diffraction patterns are called Airy rings. The first Airy ring is called the Airy disc.

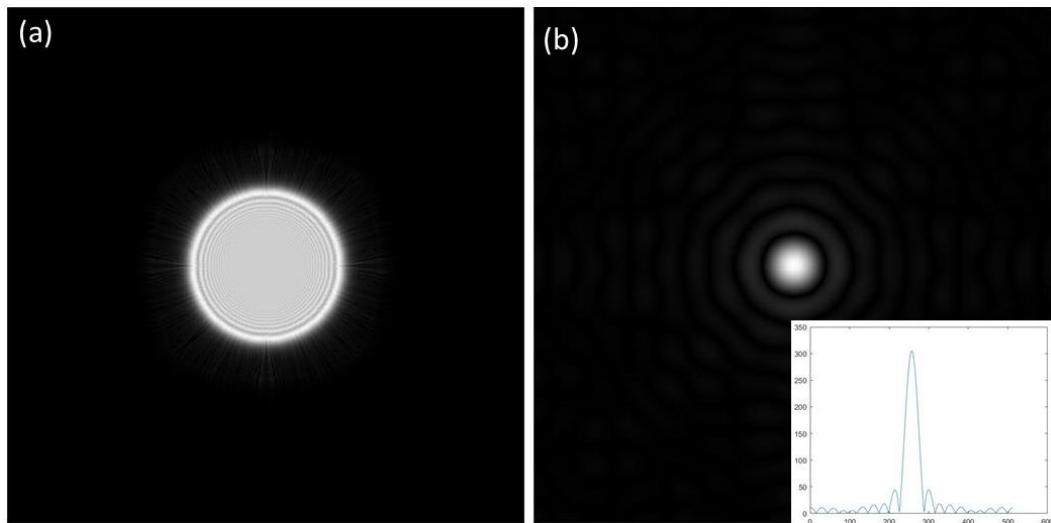


Figure 2.5.3 shows the near field Fresnel diffraction pattern (a) and Fraunhofer pattern (b) of an aperture. Figures are plotted with MATLAB.

2.5.3 Weak Phase Object Approximation

In electron microscopy, since the specimen is very thin (several tens to several hundreds of nanometers), and the electrons are accelerated by an 80-300 KeV voltage, when the high-energy electrons interact with the specimen, the amplitude of the electron wave changes very little and usually can be ignored. In this case, the specimen can be approximated as a phase object approximation (POA). If we use $f(x, y)$ to describe the object function, x, y is the 2-D specimen coordinate, and then we have

$$f(x, y) = e^{i\sigma v(x, y)}, \quad (2.5.3)$$

where $v(x, y)$ is the specimen potential field function integrated along the optical axis.

Mathematically, the object function $f(x, y)$ can be expanded with the Taylor expansion. For a weak phase object, we take the first order and abandon the higher order terms in the expansion, which is the weak phase object approximation (WPOA):

$$f(x, y) = 1 + i\sigma v(x, y), \quad (2.5.4)$$

where σ is a very small constant (Williams and Carter, 2009).

2.5.4 WPOA image

Supposing there is a specimen which is just a point, when this point goes through the microscope we can get an image of this point which is a disc. This is because the lenses in the microscope have limited diameters, which means that some high frequencies are lost when forming the image. If we use a function to describe this imaging system which images a point into a disc, then we get the definition of the point spread function or impulse response function $h(x, y)$. Usually, the specimen is a 2-D plane consists of a lot of different points

which is described as $f(x, y)$. When all these points go through the microscope system $h(x, y)$, the image of the specimen $g(x, y)$ is produced. The image function is a convolution of the object function and the imaging system function:

$$g(x, y) = f(x, y) \otimes h(x, y). \quad (2.5.5)$$

When we use a camera to detect the image function, the phase is lost, what we detect is the intensity, which is $|g(x, y)|^2$. In Fourier space, according to the convolution theorem of Fourier transform (Eq. 2.1.11), we have

$$G(u, v) = F(u, v)H(u, v), \quad (2.5.6)$$

where $G(u, v)$, $F(u, v)$, $H(u, v)$ are the Fourier transform of $g(x, y)$, $f(x, y)$, $h(x, y)$ respectively; u, v are the Fourier space frequency coordinate. $H(u, v)$ is defined as the transfer function of the imaging system.

$$h(x, y) = e^{i\varphi(x, y)} = \cos\varphi(x, y) + i\sin\varphi(x, y). \quad (2.5.7)$$

So, the image function of a weak phase object is

$$g(x, y) = (1 + i\sigma v(x, y)) \otimes (\cos\varphi(x, y) + i\sin\varphi(x, y)) = 1 + i\sigma v(x, y) \otimes \cos\varphi(x, y) - \sigma v(x, y) \otimes \sin\varphi(x, y). \quad (2.5.8)$$

The intensity of the image wave is

$$I(x, y) = g(x, y)g^*(x, y) = (1 - \sigma v(x, y) \otimes \sin\varphi(x, y))^2 + (\sigma v(x, y) \otimes \cos\varphi(x, y))^2. \quad (2.5.9)$$

Since σ is small, we can ignore the second order terms, so that the intensity is given by

$$I(x, y) = 1 - \sigma v(x, y) \otimes 2\sin\varphi(x, y), \quad (2.5.10)$$

which means that actually, only the sine term of the system transfer function contributes to the image intensity. This is important for simplifying the contrast transfer function later.

2.5.5 Contrast transfer function

The contrast transfer function of an imaging system $H(u, v)$ defines the frequencies of the object that can pass the imaging system, and thus defines the resolution and the phase contrast of the imaging system. The contrast transfer function of an electron microscope usually consists of three terms: the aperture transfer function $A(u, v)$, the envelope transfer function $E(u, v)$ and the aberration transfer function $B(u, v)$. The envelope function describes the coherence of the source and the instability properties of the microscope. The aberration function in this section will only discuss the defocus term and the spherical aberration term in Eq. 2.4.6.

$$H(u, v) = A(u, v)E(u, v)B(u, v), \quad (2.5.11)$$

The aberration transfer function $B(u)$ is

$$B(u, v) = e^{i\chi(u, v)} = \cos \chi(u, v) + i \sin \chi(u, v). \quad (2.5.12)$$

In the image function of a weak phase object (Eq. 2.5.10), the intensity of the image is only affected by the sine term of equation 2.5.9, so we can simplify $B(u, v)$ as

$$B(u, v) = 2 \sin \chi(u, v) = 2 \sin \left(\frac{\pi c_s \lambda^3 (u^2 + v^2)^2}{2} + \pi \Delta f \lambda (u^2 + v^2) \right), \quad (2.5.13)$$

Where Δf is the defocus, c_s is the spherical aberration constant of the lens.

Figure 2.3 plots the aberration term of the contrast transfer function along one dimension, to show how the transfer function changes with c_s , λ , and Δf .

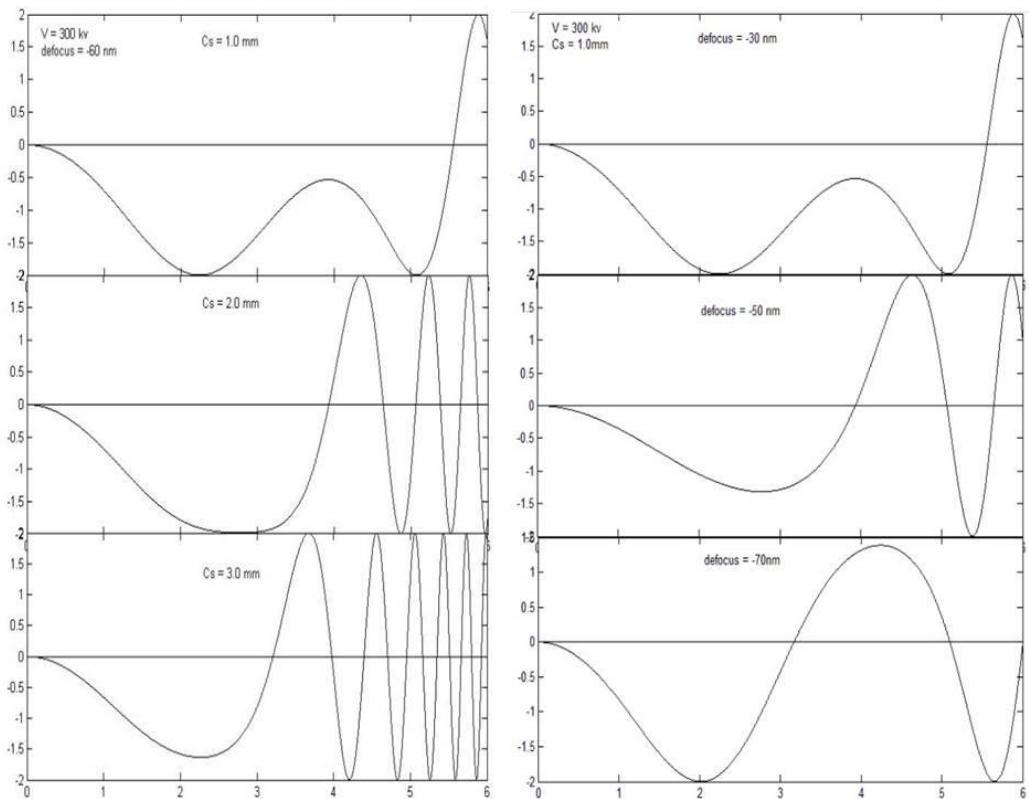


Figure 2.3 the contrast transfer function of the electron microscope. The left figure shows how it changes when $C_s = 10\text{mm}$, 20mm , and 30mm . The right figure shows how it changes when the defocus = -30nm , -50nm , and -70nm . Figures are calculated by the author with MATLAB. x axis plots the frequency and y axis plots the transfer degree.

The contrast and resolution of the imaging from the conventional electron microscopy are defined by the contrast transfer function. When we do the conventional imaging on the electron microscope, we need to correct the aberrations of the lenses and change the defocus, to achieve the best results. However, the contrast transfer function of ptychography is free of the influence of aberrations; in theory, it is only defined by the size of the detector (Humphry et al., 2012).

2.6 Electron microscopy

2.6.1 Why use an electron microscope

According to the Rayleigh criterion in light microscopy, the smallest distance δ that can be resolved by an optical system is

$$\delta = \frac{0.61\lambda}{\mu \sin \beta}, \quad (2.6.1)$$

where λ is the wavelength of the illumination, μ is the refractive index of the viewing medium, and β is the semi collection angle. Usually, $\mu \sin \beta$ is called the numerical aperture.

We can see from Eq. 2.6.1 that, when the numerical aperture equals 1, the resolution is about half the wavelength of the illumination. The wavelength of green light is about 550nm, so the highest resolution that can be achieved in light microscopy is about 300nm. In material science, some properties of the material are closely related with its nanostructure. It is necessary to know the micro features of the material at atomic scale for better understanding the properties of the material. The radius of the atom is about 0.1-0.5nm (carbon 0.22nm, gold 0.166nm, silicon 0.21nm). The resolution of light microscopy is far from the atomic scale. However, as we described in the previous section, the electrons with high kinetic energy has a wavelength on the pm scale. The trajectory of the electrons can be controlled by an appropriate magnetic field, which behaves like the lens in a light microscope. All these make an electron microscope possible.

The wave-like characteristic of the electron was proposed by Louis de Broglie in 1925; in 1927, it was proved experimentally via the classical electron diffraction experiment. In 1932 the idea of electron microscope was proposed and demonstrated by Knoll and Ruska. To date, there are different kinds of electron microscopes used for different applications. For example, transmission electron microscopes (TEM) is used for transmission and diffraction imaging;

the accelerating voltage of TEM varies between 80keV and 400keV. The higher accelerating voltage means shorter wavelength and higher resolution. Scanning electron microscopes (SEM) for topography imaging usually run at a voltage of 1keV-30keV. Scanning transmission electron microscopes (STEM) are a combination of TEM and STEM. Analytical electron microscopes (AEM) use the X-Ray energy dispersive spectrometry (XEDS) and the electron energy-loss spectrometry (EELS) to analyze the chemical properties of the material. Nowadays, an advanced electron microscope can combine TEM, AEM, and STEM together. For example, the electron microscope used for the experiments in this thesis, JEOL R005 HREM, we can obtain a bright field image, a dark field image, a high angle annular dark field image, a selected area diffraction pattern or a convergent beam electron diffraction pattern.

Electron microscopy has lots of advantages compared with light microscopy, X-Ray microscopy, and other microscopy methods. Firstly, as we already mentioned, electron microscopy can achieve a high-resolution image of the specimen; today we can get atomic scale resolution on an HREM. Secondly, the electron beam is an ionizing radiation; when the electrons interact with the specimen, there is a wide range of secondary signals will be derived; these signals can be detected for XEDS or EELS, which is useful for analyzing the chemical properties of the material. Thirdly, electron microscopes usually have a large depth of field; the TEM images usually appear in focus even though the specimen has a certain thickness; however, this characteristic also brings about inconvenience sometimes – it may give confusing images when we cannot tell the depth of the structures. Finally, because of the short wavelength of the electron wave, it is also easy to get the diffraction image of the specimen at the atomic level, which can be used for various applications: the diffraction patterns can be used to analyse the structure of the crystal; if we know the spacing structures of the crystal, we can also use its diffraction patterns to calibrate the parameters of the electron microscope; by retrieving the phase of the diffraction patterns, we can obtain super

high-resolution images without the critical demand for the electron lenses (Gabor, 1948; Spence, 2003).

Despite all these advantages of electron microscopy, there are still some unfavorable factors for electron microscopy. Here we list some of the drawbacks of the conventional transmission electron microscopy: Firstly, it is expensive, not like the light microscopy, of which the illumination is cheap, the illumination in an electron microscope is the electrons accelerating with high energy. Secondly, the electron lenses are of very bad quality. There are significant aberrations of the magnetic electron lens, like astigmatism, spherical aberration, chromatic aberrations and other higher orders aberrations. The aberration of the electron lenses is the primary factor that limits the resolution of an electron microscope, and the worse news is that the aberrations of the electron lens can never be completely removed. Nowadays, some advanced S/TEM is assisted with aberration correctors to improve the function of the objective lens and image lens. Thirdly, the specimen used for transmission electron microscope has to be very thin, usually not thicker than 100 nm, and the thinner, the better. The specimens we used for the experiments in this thesis are standard TEM specimen with thickness around 500nm. Fourthly, the ionizing electron radiation can cause damage to the specimen. This is very easy to be observed during a STEM imaging, for example, gold particles can be melted by the nanobeam in a few seconds. The damage during the TEM imaging, in which the electron beam is spread widely, is much weaker; but we can still see the specimen get contamination from the heated carbon film after several minutes. Finally, the images or diffraction patterns we collect directly from a TEM, which are an averaged image through the thickness of the specimen, do not have depth sensitivity – they will not give any information about the depth of the specimen.

In this section, we explained why we need electron microscopy, the common electron microscopes, and the advantages and drawbacks of transmission electron microscope. In the

next section, we will talk about the different roles in an electron microscope: the source, the lens, the aperture and the detector.

2.6.2 The components of an electron microscope

2.6.2.1 Source

There are two ways to make the electrons overcome the work function (usually several electron volts), which stops the electrons being emitted. One is called thermionic emission, which works by heating the material to a certain temperature. The other is called field emission, which works by applying electronic field with certain strength on a sharp tip. The apparatus generates the electron beam in an electron microscope is called electron gun. Based on these two kinds of electron emission, the electron guns can be classified as a thermionic emission gun or a field emission gun. Tungsten and LaB_6 are the common materials used for the thermionic emission gun. The material of the tip in a field emission gun is usually tungsten. To allow field emission, the surface of the tip has to be pristine, which means free of contamination and oxide. The field emission gun is kept at a low temperature; that is why we call cold field emission gun. In practice, when the emission of the gun is out of condition, for example, the emission is too weak or asymmetric, we can clean the tip by heating it up either with a relatively low temperature for a longer time (low flash), or with a high temperature for a short time (high flash).

To evaluate the properties of an electron source, we need to consider the brightness, the temporal coherence, the energy spread, the spatial coherence and the stability.

The brightness is flux per unit area per solid angle. The brightness is particularly important in AEM and high-resolution TEM. AEM works creates the signals emanating from the

specimen by the irradiation of the electron beam; in high-resolution TEM imaging, the magnification can be above 800k. Thus, the image on the detector is an extremely small part of the specimen, which needs certain intensity per unit area to make sure the signal noise ratio in the image. In practice, there are three ways to improve the brightness of a cold field emission source: 1) we can increase the voltage on the anodes, which controls how many electrons are extracted from the source. Usually, they are named as A_1, A_2 on the microscope. 2) we can change the magnification of the condenser lenses, which is similar to control how far the source is away from the specimen. This is usually named as spot size in the microscope. 3) we can change the size of the condenser aperture, which is inserted behind the condenser lenses to limit the electrons that go through the illumination system. The larger the condenser aperture size is, the higher flux is. All these three ways to increase the intensity of the illumination are at the expense of decreasing the spatial coherence of the illumination.

The temporal coherence of the electron source is about the wavelength spread of the electrons. As we already know that the wavelength of the electron is determined by its kinetic energy E . If the electrons emanated from the gun obtain the same energy, the wavelengths of the electrons will be the same; in this case, the electron illumination is fully temporally coherent. However, with either thermionic emission or field emission, there always exists an energy spread ΔE , thus the electrons will have wavelength difference $\Delta\lambda$ of

$$\Delta\lambda = \frac{vh}{\Delta E} , \quad (2.6.2)$$

where v is the velocity of the electron, h is Planck's constant.

The temporal coherence is about the wavelength difference of the electrons, while the spatial coherence is about whether the electrons emanated from the same spatial point. If all the electrons are emanated from a point source, then the illumination will be fully spatial coherent. However, all the electron sources in the electron microscopes have finite physical

size. The cold field emission gun has a smaller physical source size compared with the thermionic emission gun, thus it has better spatial coherence. The coherence of the illumination affects the resolution of TEM imaging and the contrast of the diffraction patterns by limiting the contrast transfer function; it also affects the resolution of SEM and STEM imaging by limiting the focal size of the nanobeam. For a specific electron microscope, to improve the spatial coherence, we can either use lower anodes voltage, smaller spot size or insert a smaller condenser aperture. We need to take a balance between brightness and spatial coherence during an experiment. There are several ways to measure the spatial coherence of an electron source, which we will discuss in section 2.7.

The stability is about the voltage supply for the source. The thermionic emission gun usually has more stable supply than a field emission gun. Here Table 2.6.1 compares the brightness, energy spread, spatial coherence and stability of the common tungsten thermionic emission gun, LaB_6 thermionic emission gun and field emission gun. We can see that cold field emission guns have smaller energy spread and better spatial coherence. The data in this thesis are collected with a cold field emission gun (S/TEM, Chapter 3 and 5) and thermally assisted field emission gun (SEM, Chapter 4).

	Units	Tungsten	LaB_6	Field emission
Brightness	A/m^2sr	10^9	5×10^{10}	10^{13}
Energy spread	eV	3	1.5	0.3
Crossover size	μm	50	10	<0.01
Stability	%/hr	<1	<1	5

Table 2.6.1 compares the properties of different kinds of electron sources.

2.6.2.2 Lenses

In this section, we will discuss the lenses in S/TEM (the condenser lens, the objective lens, the projector lens and the deflectors) and their function in the electron microscope.

The magnetic lens affects the accelerated electrons beam in a similar way to the optical lens affecting light. It either converges the parallel beams into a focus or forms an image with a certain magnification of an object.

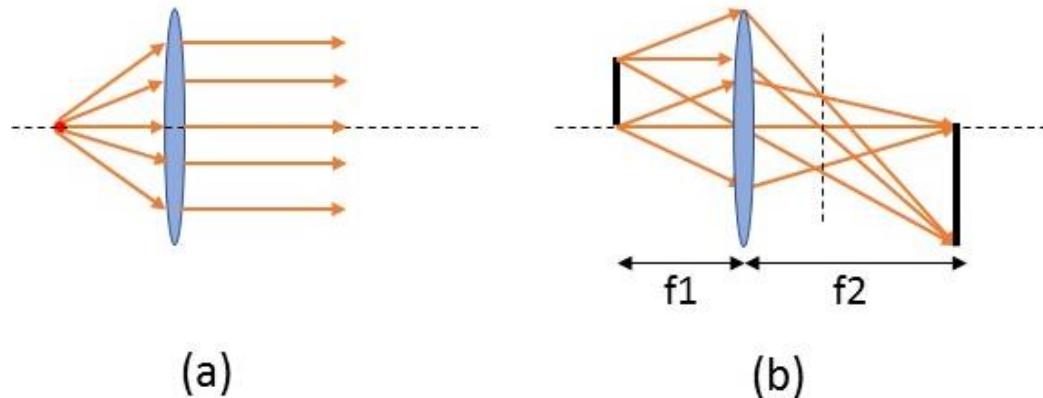


Figure 2.6.1 the functions of the lens: (a) to form parallel illumination from a point source, or inversely, converge parallel illumination into a point source; - transfers between image space and reciprocal space; (b) to form a magnified or demagnified image of an object – transfers in the same space.

The lenses that are used to form the illumination in the electron microscope are called condenser lenses. We call the source and the condenser lenses together as the illumination system. There are usually 2 or 3 condenser lenses working together to give the parallel or focused illumination in S/TEM. Figure 2.6.3(a) shows the illumination system in TEM in imaging mode. It consists of two condenser lenses C_1 and C_2 . C_1 forms a demagnified images of the electron source. We can change the strength of C_1 by changing the spot size button on the microscope operation panel. The stronger C_1 is, the further distance from the crossover to the specimen; the smaller the spot size is, the better the spatial coherence is. C_2

controls the convergence angle of the electrons beam illuminated on the specimen. We can change convergence angle (the brightness) of the beam by changing the strength of C_2 with the C_2 on the operation panel, which is used quite often during the imaging process.

Diffraction imaging in selected area mode needs parallel illumination; in this case, the pre-objective lens will behave as part of the illumination system (Figure 2.6.3(b)). C_2 forms a second image of the electron source at the back focal plane of the pre-objective lens. Figure 2.6.3(c) is a simplified ray diagram to show the illumination system in scanning image mode (SEM or STEM). The source is demagnified and focused into a nanobeam spot on the specimen. There are two series of condenser apertures with size [50-500um] inserted behind C_1 and C_2 respectively. The electron lens has strong aberration near its edges; the condenser aperture with an appropriate size can reduce the aberrations, and improve the coherence of the illumination.

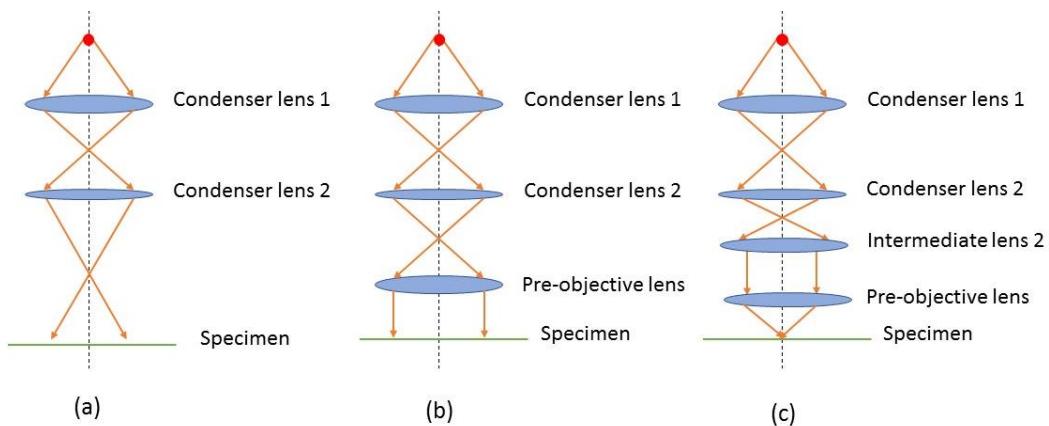


Figure 2.6.3 shows the condenser system of TEM in imaging mode (a) and (SAD) diffraction mode (b), and STEM (c).

The objective lens is the most important lens. It forms the first image of the specimen in TEM and defines the smallest size of the scanning nanobeam in STEM – it is the critical part which determines the resolution of the electron microscope.

The projector lens is the lens behind the objective lens; it determines whether the image plane or the diffraction pattern plane is projected onto the camera. The lenses in the imaging system magnify the image plane, via the projector lens, onto the detector. In a TEM, there is a series of selected area apertures with sizes of [10-50um] inserted at the first image plane of the objective lens; in the selected area diffraction (SAD) mode, the selected area aperture localizes an area in the image of the sample to form the diffraction pattern.

2.6.2.3 Detector

We discuss some properties of the Charge-coupled Devices (CCD) in this section, which is the most popular detector used in the electron microscopes nowadays. All the detectors that are used to collect either images or diffraction patterns in this thesis are CCD detectors.

A CCD detector consists of many units; these units are called as pixels. All the pixels on the detector are isolated from each other by creating potential wells. Each pixel will accumulate the electrons charged on it and produce a readout value of the signal. So, there is a response function between signal-in S_{in} and signal-out S_{out} for every pixel. After one acquisition, the CCD need some time to empty the signal for the next exposure; we call this frame do-well time (or frame time), which can be as short as 0.01 seconds.

There are usually several detectors installed in one electron microscope. For example, the JEOL R005 HRTEM in this thesis is used, has a CCD detector D_1 for the bright field imaging below the specimen in the chamber, another CCD detector D_2 for the diffraction imaging in the chamber at a different height from D_1 , an annular CCD D_3 for dark field imaging, and CCD for detecting the back scattered electron and secondary electrons. These CCD detectors have 2048*2048 pixels; each pixel has a dimension of $7.4 \mu m$. The pixels can be binned into

larger pixels in the acquisition for saving the storage space (binning 2, image size 1024*1024; binning 4 image size 512*512).

There exist several problems of the CCD detector that may affect the quality of the images or diffraction patterns: 1) dark current (detector pedestal), which is that the pixel may have a readout when no electrons arrive on it; 2) response nonlinearity; 3) CCD noise; 4) point spread. These properties will be discussed in more detail in section 2.7.

2.6.3 TEM and STEM

In this section, we will discuss the common application of TEM and STEM in imaging and diffraction. We will include the setup to obtain selected area diffraction pattern (SAD), convergent beam electron pattern (CBED), TEM bright field image (BF), TEM dark field image (DF), STEM BF image, STEM annular dark field image (ADF), STEM high angle annular dark field image (HAADF), and Ronchigram.

2.6.3.1 Selected area diffraction (SAD)

Selected area diffraction is a standard diffraction technique in TEM. The imaging system diagram is shown in Figure 2.6.4. The illumination is spread as a parallel beam on the specimen. The objective lens forms a magnified image of the specimen. There is a set of selected area apertures placed at the first image plane, which is used to select the area of interest to project the diffraction pattern on the detector. There is an intermediate lens (IL) behind the objective lens, the strength of which can be changed to make the projector lens focuses on its back focal plane. The projector lens is the lens to project the back focal plane of the intermediate lens with a certain magnification (camera length) onto the detector. In

practice, there are usually several intermediate lenses between the projector lens and the detector to assist the controlling of the magnification (camera length).

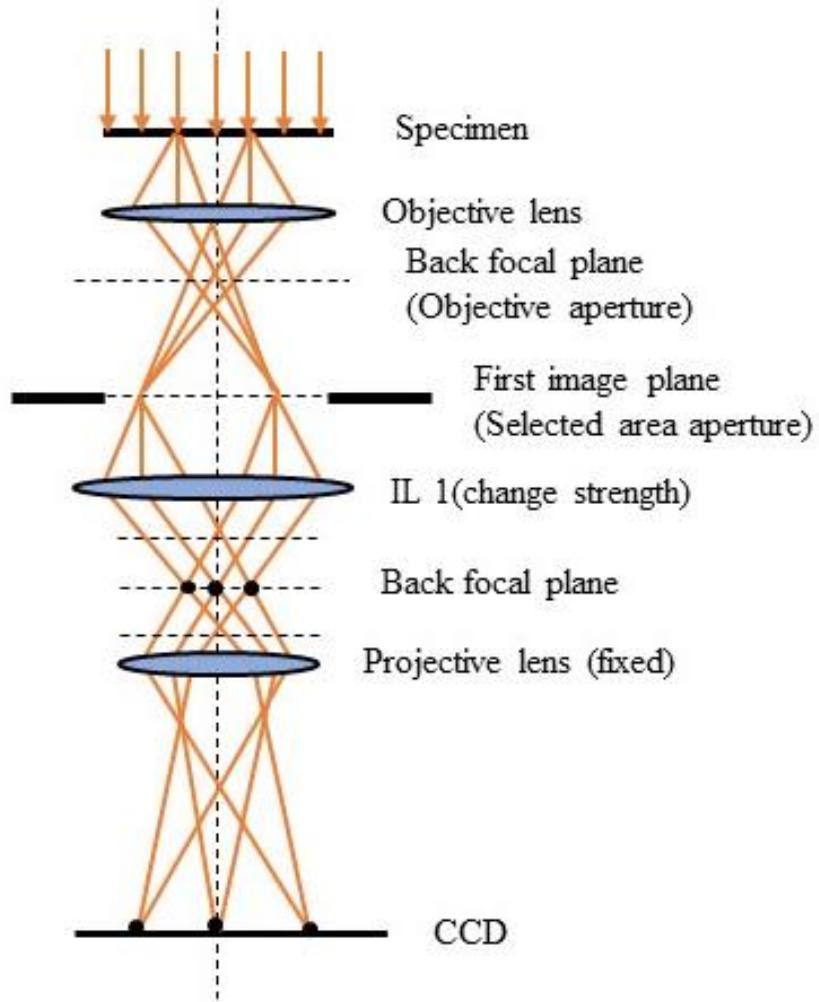


Figure 2.6.4 shows the ray diagram of TEM imaging system in SAD mode. The objective lens forms the first image plane, which is the same as TEM image mode; the projective lens project the back focal plane of the intermediate lens onto the detector.

Figure 2.6.5(a) shows SAD patterns of Thallous chloride crystals and gold particles obtained in the JEOL R005. Because the specimen is illuminated with a (nearly) parallel beam in selected area diffraction mode, in the diffraction pattern of the crystal, the diffraction spots are points; a lot of these diffraction points connect to form a thin ring as shown in Figure 2.6.5(a). To collect a SAD pattern on the CCD, usually we need a beam stop to block the

central spot as shown in Figure 2.6.5(a); that is because the central spot is too bright, which may saturate the detector. The camera length can be changed by altering the projector lens system, from 8cm to 2m. The strength of the intermediate lens can be changed with the diffraction defocus, thus we can choose to project either the defocus side or the over focus side of the back focal plane onto the detector, rather than the exact back focal plane; we call this kind of diffraction pattern as a near field diffraction (Fresnel) pattern. Figure 2.6.5(b) shows a defocused near field diffraction pattern collected at 2m camera length, which is the kind of diffraction patterns used to do the ptychographic reconstruction in Chapter 3.

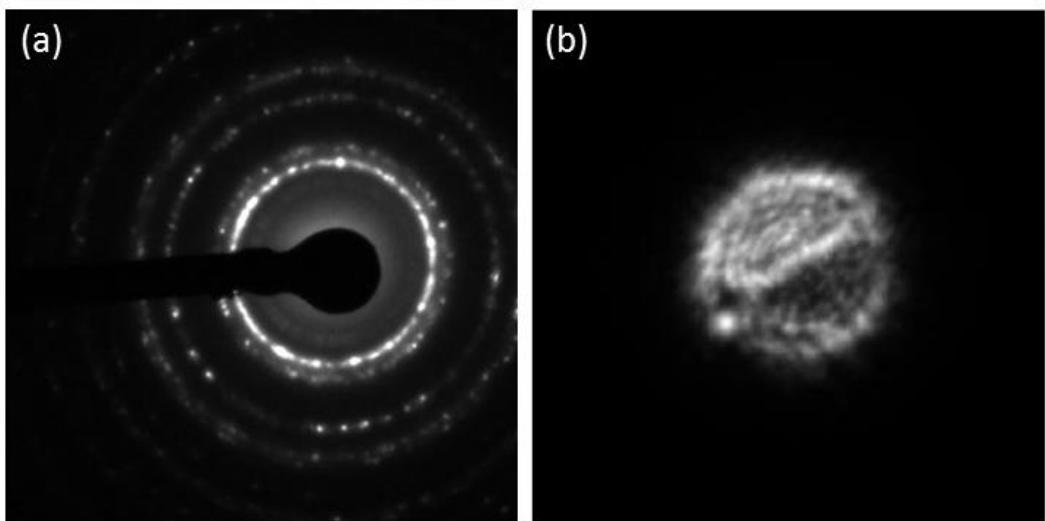


Figure 2.6.5 (a) shows the diffraction patterns at the back focal plane; there is a beam stop blocked the central part to protect the detector from saturation. **(b)** shows the pattern that when the projective lens projects not exactly at the back focal plane, but a defocused plane. Images were obtained on the JEOL R005.

2.6.3.2 Convergent beam electron diffraction (CBED)

Convergent beam electron diffraction is the other diffraction technique that can be implemented on a TEM. The difference of the setup to form a CBED pattern from the setup to form a SAD pattern is that, instead of spreading parallel illumination on the specimen in

SAD mode, the beam is focused on the specimen (Figure 2.6.6 and 2.6.3(a)). Since in the illumination beam there is a range of angles, the diffraction from a crystal will be a disc. Compared with the SAD pattern, the CBED pattern contains more information about the thickness of the specimen, the structure of the crystal (unit cells and lattice parameters). However, the drawback of the convergent beam diffraction is that the illuminated area can contaminate or damage more easily. Figure 2.6.7 show one example of a CBED pattern.

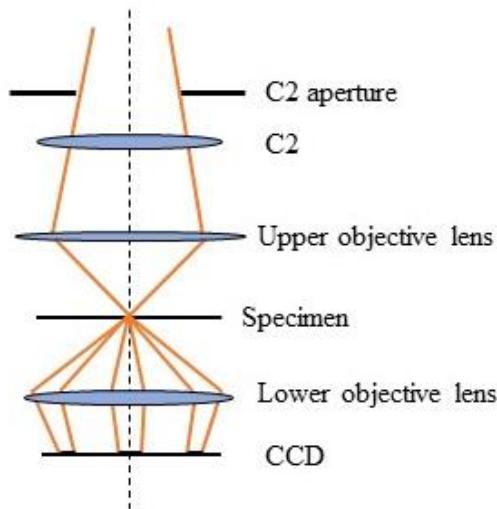


Figure 2.6.6 the ray diagram of CBED mode; the projective lens forms a nanobeam on the specimen. The detector collects the CBED patterns transmitting the specimen.

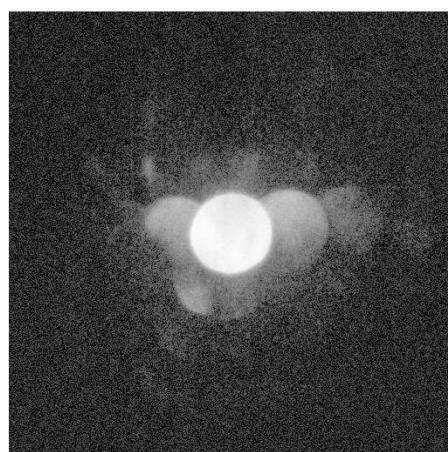


Figure 2.6.7 shows one example of CBED pattern. Plot in logarithm scale; beam energy 300KeV, camera length 12cm, exposure time 1s.

2.6.3.3 TEM bright field image and dark field image

TEM imaging mode is similar to the function of the conventional optical microscope, which forms a magnified image of the sample. With high-resolution S/TEM, we can see the nanostructures of the materials directly. Taking the JEOL R005 (300KeV) as an example, the magnification range at the conventional imaging mode is 6000 – 800k. The functions of each lens in TEM imaging mode is shown in Figure 2.6.9. The objective lens forms the first image of the sample; the projector lenses project the first image onto the detector. So, the image quality is mainly decided by the objective lens and the alignment of the projector lens (including the intermedia lenses). When we change the magnification, we are changing the strength of the projector lens and the intermedia lenses.

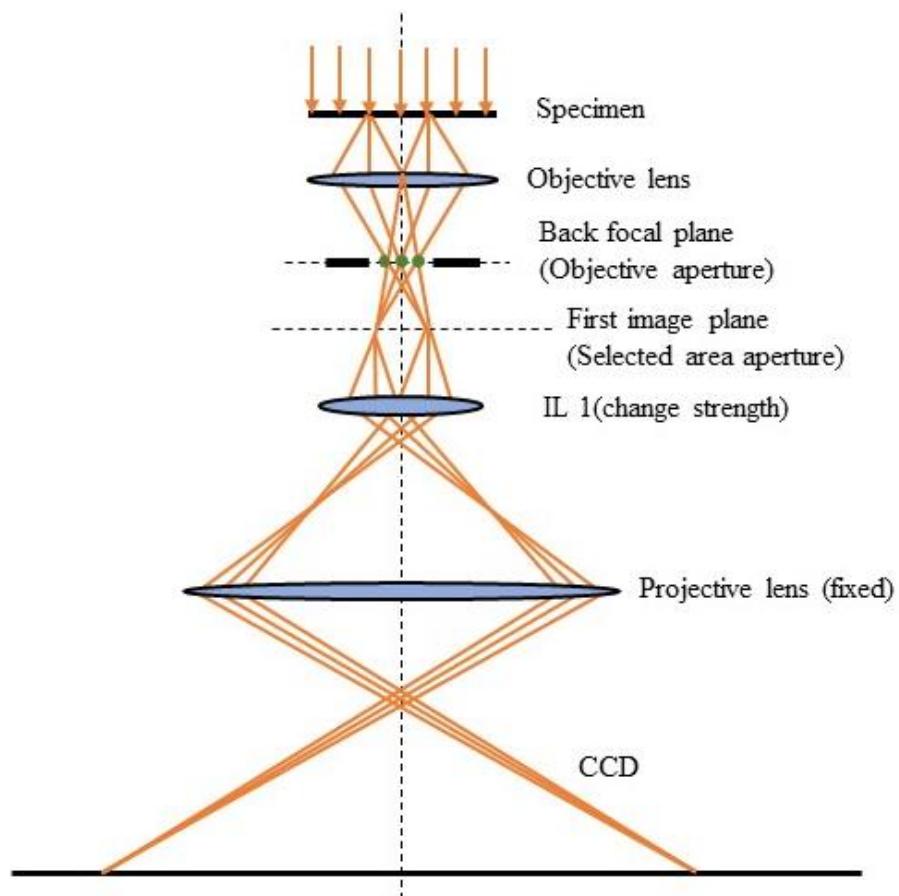


Figure 2.6.8. The ray diagram of TEM imaging mode. Objective lens forms the first image; the projective lenses project the first image plane onto the detector.

There are two kinds of image that can be obtained from a TEM depending on which frequencies at the back focal plane of the objective lens are selected to form the image: the bright field image or the dark field image. As shown in Figure 2.6.9, there is a series of objective apertures at the back focal plane of the objective lens. Here we have several choices: (1) do not insert any objective aperture (Figure 2.6.10(a)); in this case, all the frequencies will contribute to the image; this is one kind of bright field imaging; (2) insert a centred aperture (Figure 2.6.10(b)), thus only the low frequencies will contribute to the image; this is the other kind of bright field image; (3) insert a shifted objective aperture (Figure 2.6.10(c)), thus only the high frequencies contribute to the image; this is dark field imaging; (4) instead of placing the aperture to one side, we still insert a centred aperture, but change the incident angle of the beam (Figure 2.6.10(d)); in this case, still the high frequencies contribute to the image, this is the other kind of dark field imaging.

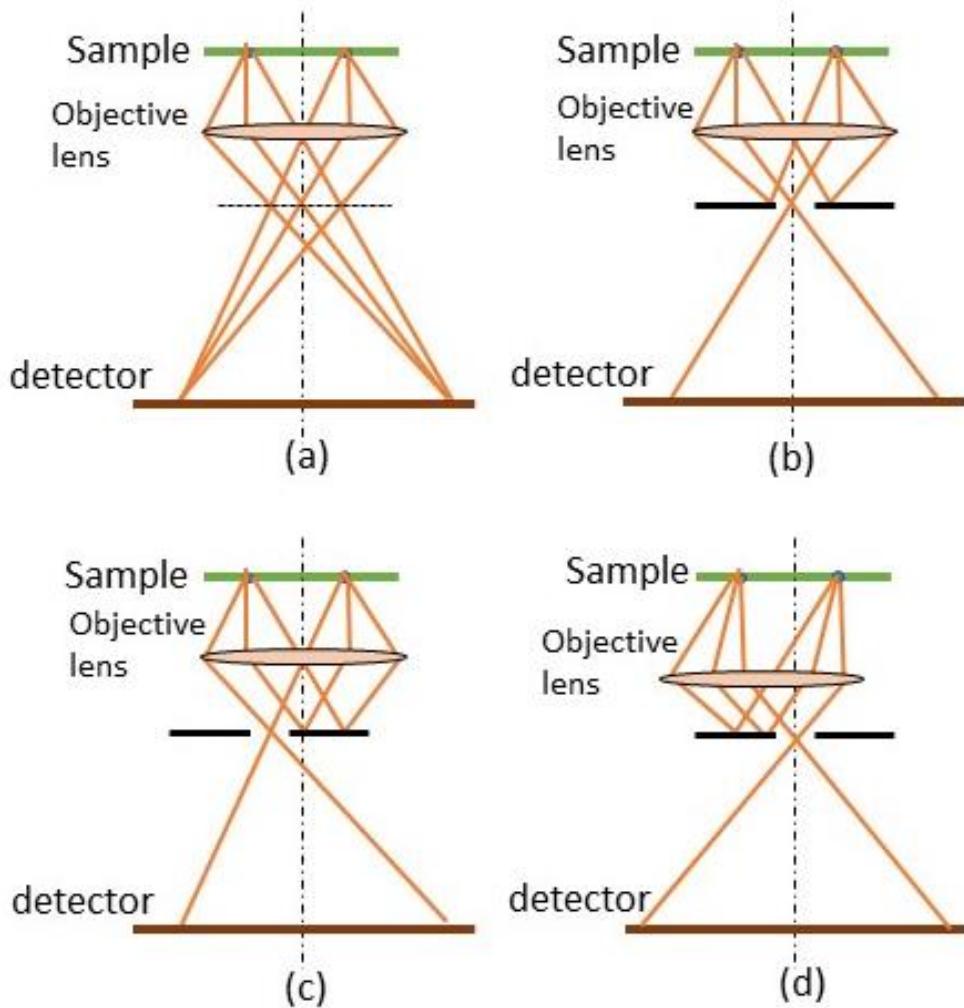


Figure 2.6.9 shows the ray diagram of BF imaging and DF imaging (figures from Springer). (a-b) shows the bright field imaging – Without using objective aperture or use centered objective aperture; (c-d) shows the dark field imaging – either use a side aperture or use a centered aperture but with a tilted beam.

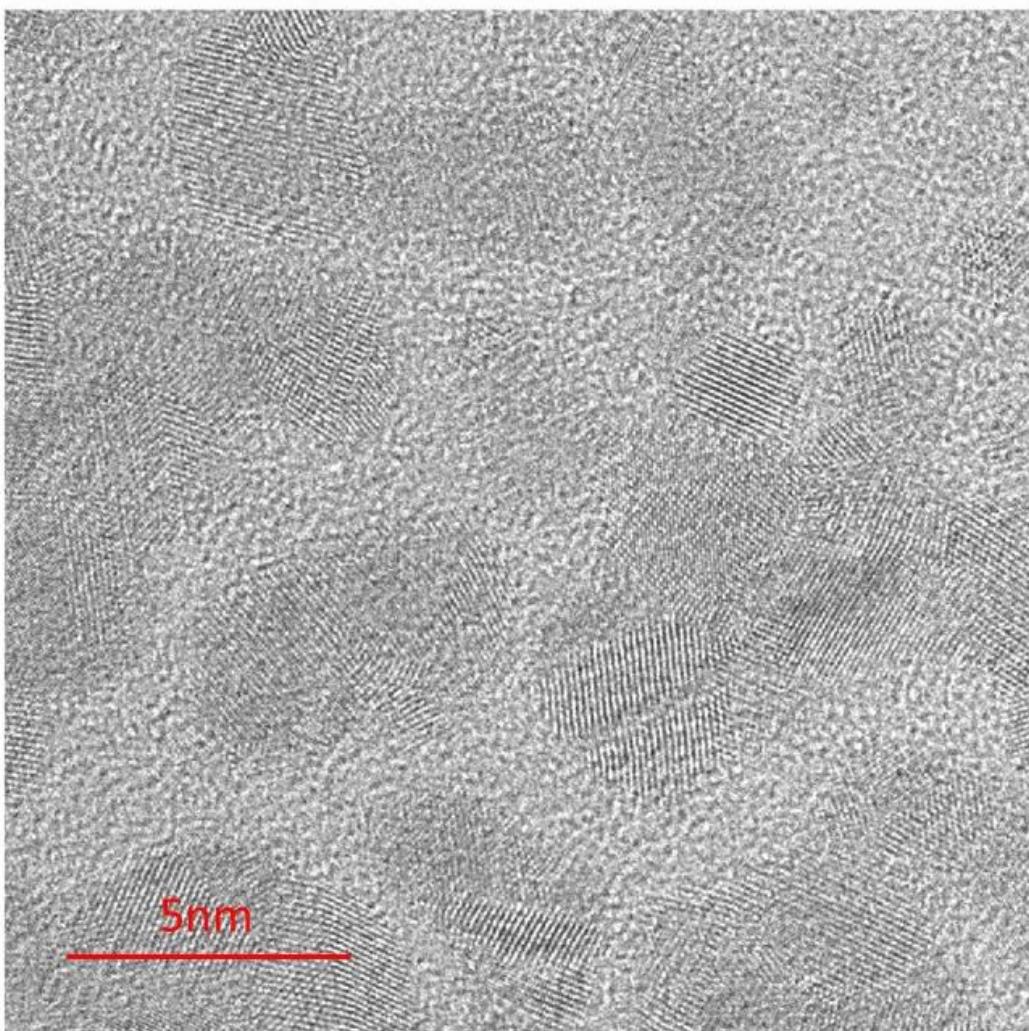


Figure 2.6.10 shows a BF from TEM (JEOL R005) at magnification 600k. The specimen is gold particles on the carbon film. We can see the **gold lattice** and the amorphous carbon film.

2.6.3.4 STEM BF/DF/ADDF imaging

In STEM imaging mode, the detectors play the roles of the objective apertures in TEM imaging mode, to select which frequencies contribute to the image. As shown in Figure 2.6.12, there is bright field detector to collect the directly scattered electrons to form bright field image; annular dark field detector collects high order scattered electrons to form dark field image (ADF); there is also high angle annular detector, which collects the scattered electrons with higher angle to form high angle annular dark field image (HAADF).

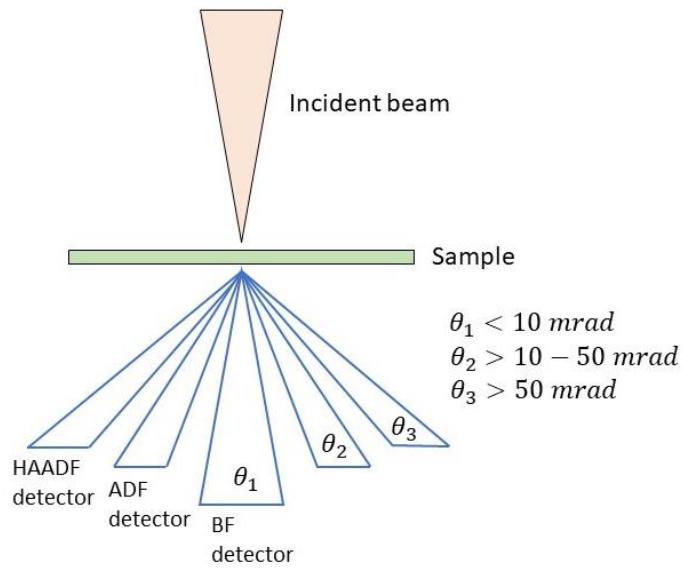


Figure 2.6.11 shows the diagram of BF, ADF, and HAADF detectors in a STEM imaging system.

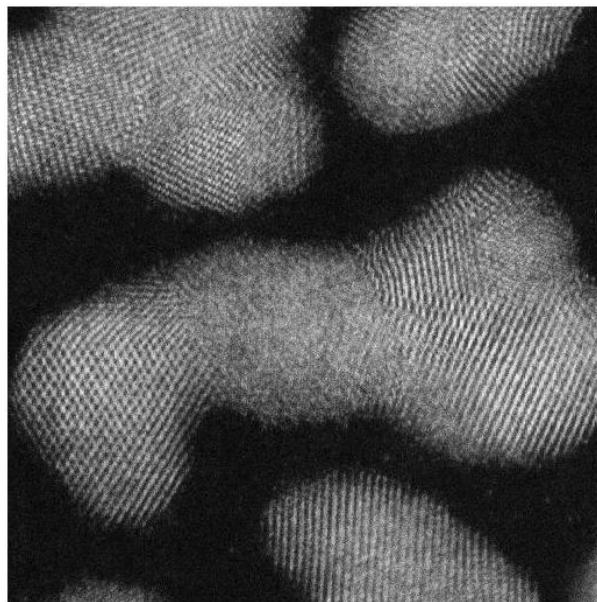


Figure 2.6.12 shows an HAADF image from STEM (JEOL R005). The specimen was gold particles on the amorphous carbon film. We can see the true structure of gold particles, but we cannot see the light carbon atoms.

2.7 Ptychography

2.7.1 Introduction

Direct imaging microscopy uses the lenses to form the image of the sample. For example, the optical lens for light microscopy, Fresnel Zone plates (Kirz and Attwood, 2015), Kirkpatrick-Baez mirror for X-Ray microscopy (Kirkpatrick and Baez, 1948), and the magnetic field lens for electron microscopy. In light microscopy field, nowadays nearly perfect optical lenses can be manufactured, so the resolution is not limited by the quality of the lens. However, it still has the essential limits: we cannot ‘see’ a pure (weak) phase specimen, for examples, a piece of glasses, contact lens, or organic cells, directly under an optical microscope unless by introducing a certain defocus at the sacrifice of losing some low frequencies. In the X-Ray and electron microscopy fields, the cost to manufacture the lens is high, and the transfer function of the lens is bad because of the aberrations, especially the magnetic field lens for electron microscopy: the transfer function of the lens and the coherence of the source limit the resolution (Frank, 1973; Frank, 2006). In X-Ray and electron field there also exists a vast of demands for phase imaging (Spence, 2003), such as of light atoms, magnetic or electronic fields, strain fields and so on.

There are some direct methods to image weak phase object: the simplest way is to insert a Zernike phase plate at the back focal plane of the lens to form the image (objective lens in an electron microscope) (Danev and Nagayama, 2001); holography is a method to split the incident beam into two beams – the reference beam and the transmission beam, by interfering the exit wave with the reference beam to characteristic the weak phase object; the limits of holography is the critical requirement for the illumination coherence and stability (Gabor, 1948; Gabor et al., 1971; Lichte, 1986; Lin and Cowley, 1986); there are also other methods

(Shibata et al., 2012; Lubk and Zweck, 2015) like the differential phase contrast (DPC) method and the differential interference contrast (DIC) method.

A completely different concept of imaging technique differing from the conventional microscopy is called coherent diffraction imaging (CDI) (Fienup, 1978; Miao, 1999; Zuo and Huang, 2012; McBride et al., 2004; Morgan et al., 2013), which is an imaging technique without lens (lensless imaging) or reference beam. The principle of CDI is based on the interference of the illumination with the specimen other than the interference of the reference wave with the exit wave (Fienup, 1982): a camera is placed behind the specimen to detect the Fresnel or Fraunhofer diffraction patterns; numerical methods are used to model the interaction of the wave with the object and the wave propagation; algorithms apply the constraints of the measurements in both real space and Fourier space to phase the diffraction pattern thus work out the transmission function of the specimen (Gerchberg and Saxton, 1972; Fienup, 1978; Marchesini, 2003). CDI retrieves both the modulus and the phase information of the object, and its resolution, in theory, is only limited by the detector size (diffraction limit resolution, without considering super-resolution here).

Conventional CDI employs a single diffraction pattern (Miao, 1999). The algorithm reconstructs the object function with a constraint in the Fourier domain (the diffraction pattern) and a real space support (the location of the object and the illuminated area boundaries), as shown in Figure 2.7.1; more details of the calculation will be discussed in Section 2.7.3. The limits of conventional CDI are that it has a very small field of view and that it is not convenient to obtain accurate support in real space experimentally.

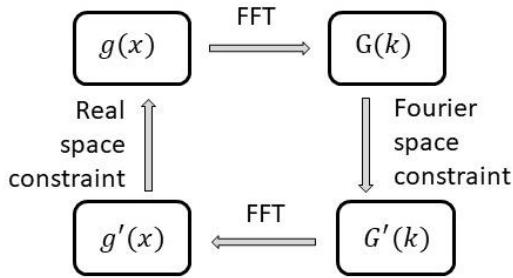


Figure 2.7.1. The flow chart of iterative computing algorithm for single shot coherent diffraction imaging.

Ptychography was developed from the conventional CDI: instead of employing a single diffraction pattern, ptychography collects a series of diffraction patterns scanning cross the object by either shifting the illumination or shifting the specimen while making sure that the adjacent illuminated area has a certain amount of overlap (Figure 2.7.2) (Rodenburg and Bates, 1992; Rodenburg and Faulkner, 2004; Maiden and Rodenburg, 2009). Because of these overlaps, each part of the object is measured more than once (the measurements depend on the overlapping times), so there is a lot of redundancy in the ptychographic data. Ptychography can retrieve both the complex object transmission function and the illumination function with the reciprocal space constraint (the detected intensities) and the real space constraint (the known scanning positions) (Thibault et al., 2008; Maiden and Rodenburg 2009). The reconstructed field of view is unlimited and the resolution, in theory, is only limited by the diffraction information.

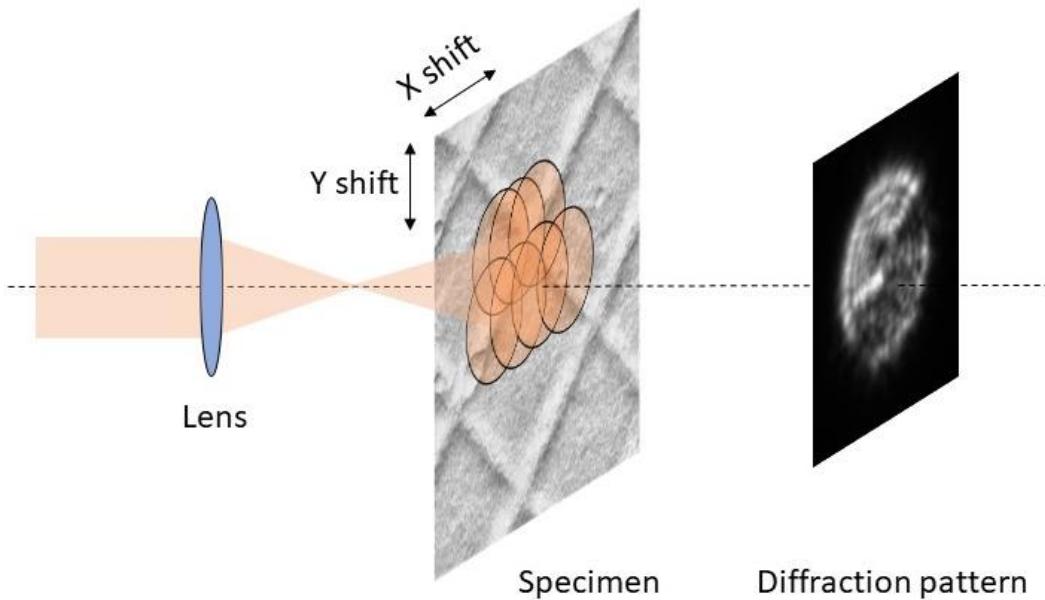


Figure 2.7.2 shows a ptychographic setup: either lens or aperture (here shows lens) is used to form localized illumination on the specimen. A detector is placed behind the specimen either in the near field or far field. We can either shift the specimen or shift the illumination and keep a certain overlapping between each adjacent illuminated area; we collect one or several diffraction patterns at each position.

Other similar techniques to ptychography are the transport intensity equations (TIE) (Teague, 1983; Beleggia et al., 2004; Ishizuka and Allman, 2005; Mcvitie and Ngo, 2007) or the focal series reconstruction (Dyck, 1992; Dyck et al., 2012), which also apply multiple correlative measurements and retrieve the complex transmission function based on the interference of the incident wave with the specimen. Unlike ptychography, which makes the multiple correlative measurements by shifting the specimen or the illumination in x and y plane (Figure 2.7.2), TIE and focal series reconstruction methods collect a series of images of Fresnel diffraction patterns by shifting the specimen in the z direction. The reconstruction algorithm is similar to ptychographic iterative reconstruction algorithms (Spence, 2003). The limits of TIE and focal series compared with ptychography are 1) TIE and focal series have a small field of view; 2) the number of the measurements is limited; 3) loss of low frequencies; 4) resolution limited by lens.

Section 2.7.2 gives a general literature review about the development of ptychography in light, X-Ray and electron field; section 2.7.3 discusses the ptychography reconstruction algorithms; we will in detail discuss the algorithm applied to process the data in this thesis - ePIE and the improvements.

2.7.2 Review

Since the idea of ptychography (to solve the phase problem by shifting the probe) was first proposed as the third method to solve the phase problem in electron microscopy field by Rodenburg(Rodenburg, 1989), - in the paper he summarised three ways to solve the phase problem: 1) interference with reference wave, for example, holography (Gabor, 1948; Gabor et al., 1971; Midgley, 2001; Lin and Cowley, 1986); 2) single diffraction pattern CDI (Zuo and Huang, 2012; McBride et al., 2004; Morgan et al., 2013); 3) ptychography; ptychography now has developed significantly in light (Rodenburg et al., 2007; Thibault et al., 2008; Maiden et al., 2012), X-Ray (Thibault et al., 2008; Dierolf et al., 2010; Thibault and Menzel, 2013), and electron (Nellist, McCallum and Rodenburg, 1995; Yang et al., 2015, 2016, and 2017; Humphry et al., 2012; Maiden et al., 2015) imaging field both on the experimental setups and the reconstruction algorithms. To begin with (1989-2004), ptychography applied a focused illumination as the point scanning probe; The scanning positions had to be very close to each other. In this thesis, we call this setup as focus-scan ptychography (Nellist, McCallum and Rodenburg, 1995; Yang et al., 2015, 2016 and 2017). The experimental setup of focus-scan ptychography is shown in Figure 2.7.3 (Nellist, McCallum and Rodenburg, 1995). The reconstruction method for focus-scan ptychography with this type of data was the Wigner Distribution Deconvolution method (WDD) (Rodenburg and Bates, 1992; Rodenburg, 2005); this was the only known inversion method at that time, which is an analytical method by

Fourier transform the 4-D dataset. During that time, focus-scan ptychography was mainly applied on a scanning transmission electron microscope (STEM) . In the paper by Nellist McCallum and Rodenburg (1995), 0.136 nm resolution was achieved on a STEM which had a conventional information limit of 0.33 nm, but only for a crystal specimen, which requires much less data than a full two-dimensional scan of a general object.

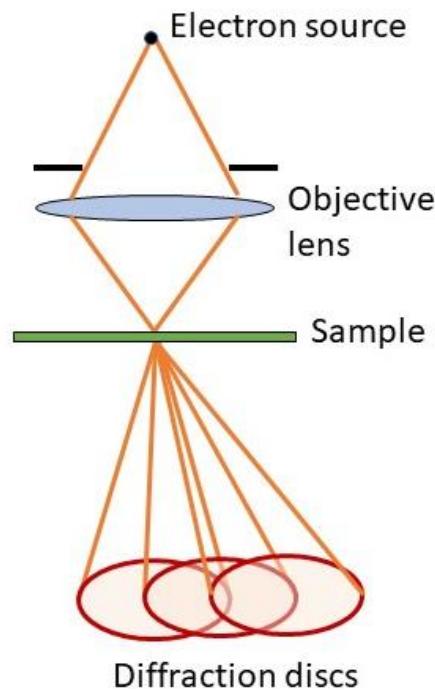


Figure 2.7.3 shows focus-scan ptychographic setup; while the beam crossover scan over the specimen, the detector in the far field collects the 4D data (2D in specimen space and 2D in the reciprocal space).

The problems existing in focus-scan ptychography are that: 1) the dataset is huge: the size of the focused electron beam in STEM is several angstroms; so scanning over a gold particle, focused scanning may need a 256*256 scanning positions. Supposing the size of the detector is 256*256. The size of the dataset is then 256*256*256*256, which is 32 GB. To do WDD with such large data will be difficult on a common computer. 2) a focused high energy

electron beam can damage the specimen quickly; in papers by Yang et al. (2015), (2016) and (2017), the counts are as high as $10^4 \text{ e}/\text{\AA}^2$.

The introduction of iterative reconstruction algorithm brought ptychography to the next stage (2004 - now) – large probe ptychography, which is the popular way to implement ptychography in the optical and X-Ray field nowadays. There are two ways to form a localized large probe on the specimen: one is using an aperture to localize the illumination - the illumination is widely spread on the specimen, which can be approximated as a plane wave; an aperture is placed closely in front of the specimen, as shown in Figure 2.7.4(a). We call this setup as parallel beam ptychography (Batey et al., 2014; Maiden et al., 2015; Li and Maiden, 2017). When the illumination is a plane wave, diffraction patterns in the far field with a bright spot in the center may saturate the camera, so, usually, the camera is placed in the near field. Or, a method to collect far field patterns and get rid of the bright spot is to place a diffuser film onto the aperture as shown in Figure 2.7.4(b); several advantages of diffused illumination have been discussed in the paper by Li et al. (2016). The other way to form a largely localized probe is to use the lens to form defocused convergent beam as shown in Figure 2.7.4(c). We call this setup as defocus-scan ptychography (Rodenburg and Faulkner, 2004; Humphry et al., 2012; Wilke et al., 2013; Li and Maiden, 2017).

Compared with focus-scan ptychography, the advantages of large probe ptychography are that – much fewer scanning positions were needed to achieve a same field of view with the focused scan, so the data is much smaller. For example, in the results that already be published, to scan a gold particle, focused scan needs scan 256*256 positions (Yang et al., 2015 and 2017), while a defocused scan needs only 7*7 positions (D'Alfonso et al., 2014). Moreover, the iterative algorithm computes one position by one position, not like WDD, which need to compute the whole data at once, thus the computer memory needed for iterative reconstruction is much less.

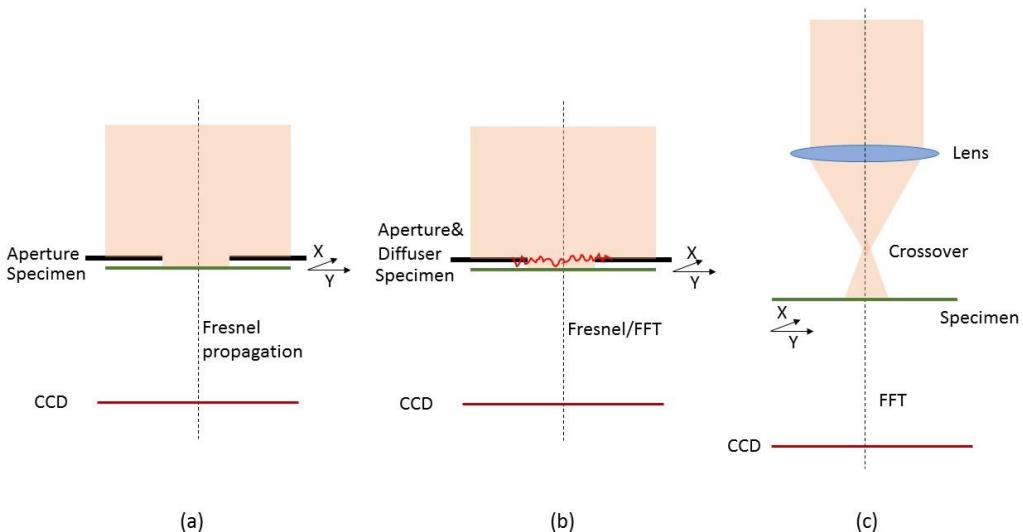


Figure 2.7.4 shows the setups of large probe ptychography; (a) an aperture is used to localize the parallel illumination on the specimen; the detector has to be placed in the near field to avoid saturation without inserting a beam stop; (b) a diffuser is inserted behind/in the aperture and in front of the specimen to increase the angles of the illumination; the detector may be placed either in the near field or the far field; (c) a lens is used to form defocused illumination on the specimen; we can change the defocus to change the size of the probe.

Large probe ptychography nowadays has become a standard diffraction imaging method in light and X-Ray field. Various kinds of experimental setups have been explored apart from the conventional transmission ptychography (detector is placed behind the specimen in the far field), like 1) Near field ptychography (Stockmar et al., 2012; Clare et al., 2015), where the detector is placed behind the specimen at a shorter distance. Near field ptychography has the advantages of shorter propagation distance, better counts in the diffraction patterns and fewer measurements. To increase the diversity of the measurements usually a diffuser is placed before the specimen to increase the structural diversity of the illumination. 2) Fourier ptychography (Zheng et al., 2013; Ou et al., 2015), where instead of shifting the specimen or the illumination to collect the diffraction patterns of the illuminated area, an aperture is placed in the back focal plane of the specimen, we can either tilt the illumination or shift the aperture to measure the intensities of the Fourier transform of the localized area at the back

focal plane of the specimen, which are low resolution images. Then we reconstruct the back focal plane of the specimen with the iterative algorithm, and Fourier transform the back focal plane to get a high-resolution complex specimen reconstruction. 3) Reflection ptychography (Roy et al., 2011), where the detector is placed in front of the specimen at some angle. Reflection patterns are collected while the convergent beam scan cross the specimen. 4) Fly-scan ptychography or continuous scan ptychography (Huang et al., 2015), where the specimen or illumination is shifting continuously while collecting the diffraction patterns. 5) Ptycho-tomography (Donnelly et al., 2017), which is a method to combine ptychography and tomography together to obtain a three dimensional reconstruction of the sample; 6) Lateral and longitudinal scan ptychography (Robisch et al., 2015), which involves not only shifting the specimen in x and y directions, but also in z direction; this is an idea to combine ptychography and the transport intensity equation (TIE).

The iterative reconstruction algorithm has also been improved a lot since the original ptychographic iterative engine (PIE) (Rodenburg and Faulkner, 2004), in which only the object function was retrieved and the illumination function was supposed known. Here we generally list some improvements of the algorithm, more details about these methods will be discussed later in section 2.7.3.2: 1) ePIE (Maiden and Rodenburg 2009) and Difference Map (DM) (Thibault et al. 2008), both retrieve the object function and the illumination function; 2) annealing and correlation methods (Maiden et al. 2012; Zhang, Yamaguchi, and Yaroslavsky 2004), to correct for random scanning positions errors; 3) super resolution (Maiden et al. 2011), improves the reconstructed resolution by extending the size of the diffraction patterns; 4) upsampling (Edo et al., 2013; Batey et al., 2014), shows that ptychography allows sparser sampling than Shannon-Nyquist sampling frequency; 5) detector background matching (Maiden et al. 2015), retrieves the detector background noise; 6) maximum likelihood (P. Thibault and Guizar-Sicairos 2012), reconstructs noisy data; 7) Gaussian convolution (Burdet

et al., 2015; Maiden et al., 2015) and modal decomposition (Li et al., 2016; Thibault and Menzel, 2013), retrieve the source partial coherence and detector point spread.

For electron waves, for the moment, ptychography has to be implemented on a transmission electron microscope (TEM/STEM or SEM in the transmission mode); the illumination can be either plane wave or convergent beam. Plane electron wave ptychography is implemented in the selected area diffraction mode, where the plane electron wave illumination is localized by a selected area aperture (Maiden et al., 2015). The characteristics of plane electron wave ptychography are that: 1) the scanned object is not the specimen, but the magnified image of the specimen; 2) the probe size is limited by the size of the selected area aperture and the magnification of the objective lens. For example, with a 10um selected area aperture and objective lens with the magnification of 80, the size of the probe is limited to 125 nm; 3) the counts in plane electron wave ptychography is low – good reconstruction can be obtained from data with counts as low as $5 e/\text{\AA}^2$ (Chapter 3), while the counts in focus-scan ptychography was $1.3 \times 10^4 e/\text{\AA}^2$ in the paper (Yang et al., 2016).

There are two ways to implement the convergent beam electron ptychography on a STEM or SEM. One way is the classical focus scan reconstructed with Wigner distribution deconvolution (WDD) method. Great successes have been achieved in focus-scan electron ptychography recently thanks to the development of direct detector, which is a direct detector that can count electrons individually, thus it has a highly efficient response property (Modulation transfer function); the conventional CCD detects the electron signal by transferring the electrons arriving at the pixel into photons first, and then counting the photons. A direct detector can achieve an acquisition frequency of 20,000 frames per second (FPS). The direct detector allows focus-scan electron ptychography on a STEM to collect the diffraction patterns efficiently and quickly while the focused beam scans over the specimen,

to simultaneously collect the HAADF image. The extreme short exposure time will bring less damage to the sample. The paper by Yang et al. (2017) reported a single gold nanoparticle reconstructed with focus-scan electron ptychography on a STEM and aberration corrected WDD algorithm; the resolution was comparable with HAADF image. The paper by Yang et al. (2016) reported reconstructions with z contrast through focus-scan electron ptychography on a STEM and WDD optical sectioning algorithm.

The other way of performing ptychography on STEM/SEM is to scan the specimen with a defocused probe. The probe size varies according to the defocus plane - usually it is several nanometres in diameter. Several results have been published: the paper by Humphry et al. (2012) showed the reconstruction on a 30keV SEM operated in the transmission mode, reconstructed with ePIE algorithm; 0.24nm resolution was obtained with ptychography while the SEM has a 1.2nm resolution limit. The paper by D'Alfonso et al. (2014) showed the reconstruction on a 300keV STEM; the data only covered 7*7 positions.

What hampers defocus-scan electron ptychography for wide application is the difficulty to collect the data with accurate parameters. For example, the accuracy of the scan positions: supposing the diameter of the probe size is 5nm, thus the scanning step size is about 1nm; either by shifting the specimen with a piezo stage or by shifting the illumination, because of the hysteresis, drifting or the calibration of the shifting pitch, it is difficult to measure the scanning positions accurately. Even though some methods have been proposed in the light or X-Ray community to correct the positions errors in the algorithm, the efficiency of the methods is affected by the structure of the specimen, noise in the data, and so on. More problems of defocus-scan electron ptychography will be discussed in Chapter 4 and Chapter 5.

2.7.3 Iterative Phase Retrieval Algorithms

After the idea of ptychography was proposed in 1970 by Hoppe (Hegerl and Hoppe, 1970), computation methods to solve the complex object function with the detected intensities (image or diffraction patterns) have also been developed correspondingly. In this section, we give an introduction and conclusion of the iterative algorithm we used to process large probe ptychographic data in this thesis, which was firstly proposed by Gerchberg and Saxton, (1972).

In the late 1980s, the idea of ptychography, which is a multiple measurement coherent diffraction imaging method, was proposed by Rodenburg (Rodenburg, 1989). Similar to CDI, there also exists two kinds of computations in the ptychography community: the analytical computation method. Wigner distribution deconvolution (WDD), was proposed by Bates and Rodenburg in 1992 by Rodenburg and Bates (1992), and was further improved recently (Li et al., 2014; Yang et al., 2015, 2016 and 2017); shortly after WDD, the iterative algorithm PIE, which retrieves the complex object function supposing the illumination function is known, was introduced by Rodenburg and Faulkner (2004). Since then, the iterative algorithms have been developed very significantly (Maiden et al., 2009, 2011 and 2012; Thibault et al., 2009, 2012 and 2013).

In the following sections we will discuss the iterative ptychographic algorithms: in section 2.3.3.1, we generally introduce the idea of iterative phase retrieval algorithms; the popular algorithm ePIE and DM are taken as examples to explain in detail about the steepest descent search and the gradient descent search; in section 2.3.3.2, we discuss the improvements in the iterative ptychographic algorithms based on ePIE, including positions correction, drifting correction, super resolution, upsampling, partial coherence, detector response correction and noise.

2.7.3.1 Ptychographic iterative phase retrieval algorithms

The iterative algorithm is a computing method to search the solution with some constraints from an arbitrary initial guess solution (Gerchberg and Saxton, 1972; Fienup, 1982; Elser, 2003; Teague, 1983; Marchesini, 2007). The calculated solution will be updated iteratively, until the calculation arrives a point where the calculated solution satisfies all the constraints. In practice, because of the noise or measured errors, the calculation solution cannot exactly satisfy all the constraints; in this case, the algorithm will give a solution which has a minimum error within the constraints.

In ptychography, the constraints are the knowns and measurements in real space and reciprocal space (Rodenburg and Faulkner, 2004; Maiden and Rodenburg, 2009; Thibault et al., 2008). The initial guesses are an arbitrary object function O and an arbitrary illumination function P . Iterative phase retrieval algorithms try to find the object function and the illumination function that match the detected intensities I , by propagating the guessed exit wave between real space and reciprocal space back and forth with Fourier transform: when propagated to real space, the calculated solution is constrained by the knowns (scanning positions) in real space; when propagated to reciprocal space, the calculated solution is constrained by the measurements (the detected diffraction patterns intensities) in reciprocal space. In Fourier ptychography, the constraint in real space is the detected image intensities; the constraint in Fourier space is the tilted angles of the illumination. In TIE, the constraint in real space is the focal depths, the constraint in reciprocal space are the near field diffraction patterns (Bostan et al., 2014).

While we apply the constraints to the calculated solutions, it is like that the solution is guided to ‘walk’ toward the solution space, attempting to find the minimum error; according to its

‘walking’ direction, the iterative algorithms can be divided into the steepest descent methods and other gradient descent methods; according to how far and in what direction it walks (steps).

Figure 2.7.5 shows the choices when applying constraint *a* to the current calculation: the red and blue arrows show the steepest direction (perpendicular to the constraint) but with different step size; searching algorithms using the steepest direction are called the steepest descent search. The green arrow shows the searching along a different direction from the steepest direction, which is called other gradient descent search. Among other gradient descent search, there is a special gradient, which is called conjugate gradient descent search.

Figure 2.7.6 concludes and classifies some popular iterative algorithms (Marchesini, 2007).

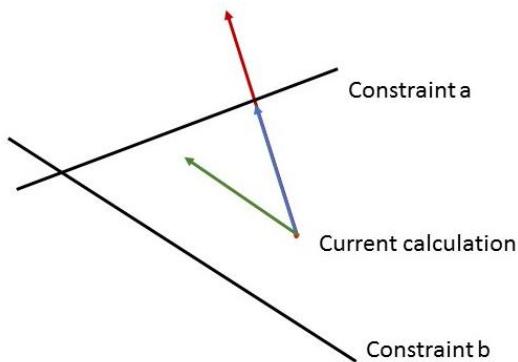


Figure 2.7.5 shows the choices when applying constraint *a* to the current calculation.

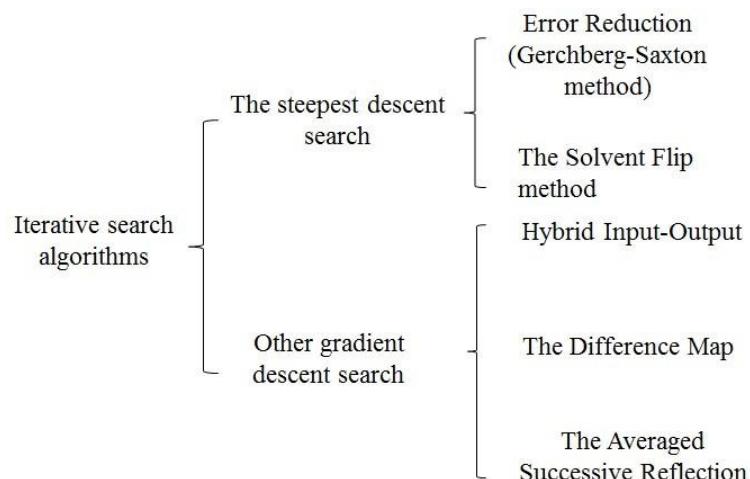


Figure 2.7.6. The classification of the iterative search methods.

The steepest descent search - ePIE

Similar to Figure 2.7.5, here we also use two lines to represent the constraints of ptychography: one is the constraint in real space, and the other is the constraint in the reciprocal space. The intersection of the two lines is the solution that the algorithm is searching for; in the Figure 2.7.7, it is the only point that satisfies both constraints. While in practice, due to noise or the measurements errors, the intersection of the two constraints is not a point but a set of points, which means there exist more than one solutions that satisfy the constraints. In this case, the algorithm is searching for a point that has the minimum error to match the constraints (Fienup and Wackerman, 1986). Here we only discuss the ideal case where no error, no noise, and where a unique solution exists.

The error reduction method (Gerchberg-Saxton method, simplified as GS method) shown in Figure 2.7.6 searches towards the smallest error direction at each step; the step size is exactly the distance between the current position and the constraint (Fig 2.7.7.a). The Solvent-Flip method (Figure 2.7.7.b) is different from GS by the step size; its step size is the GS step size scaled by a certain factor.

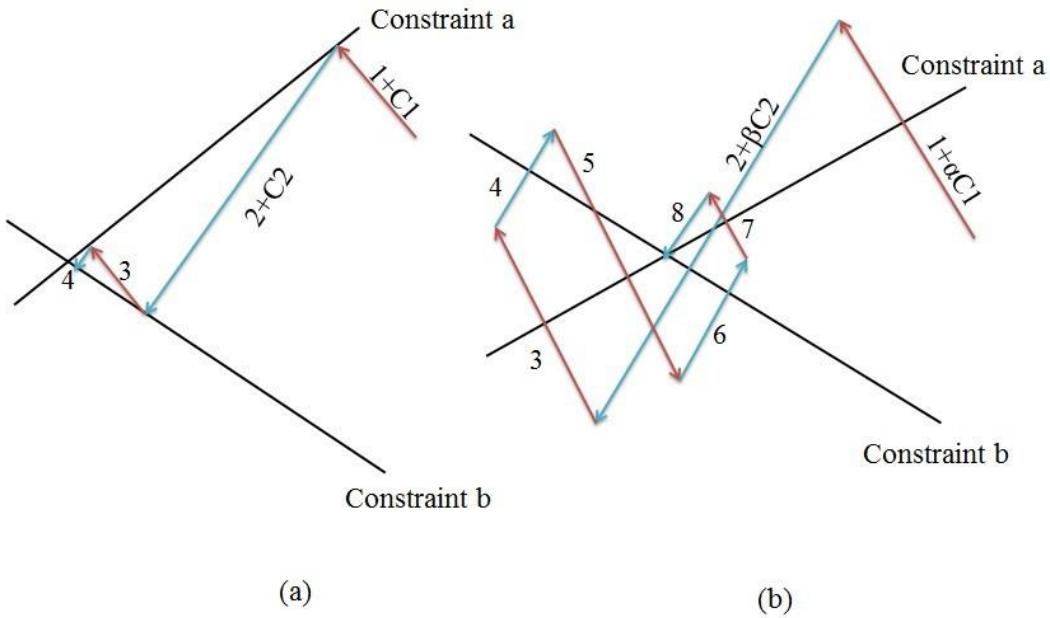


Figure 2.7.7. (a) error reduction search: the searching direction is the steepest descent direction, and the step size is the distance to each constraint; (b) solvent flip search: the searching direction is the same as ER, but the step size is proportional to the distance to the constraints.

One of the most popular steepest descent search ptychographic iterative algorithm is ePIE (Maiden and Rodenburg, 2009). The algorithm applied to process the data in the later chapters of this thesis is based on ePIE. So here below we take ePIE as an example to explain how steepest searching works with ptychography; Figure 2.7.8 shows the algorithm flow chart of ePIE.

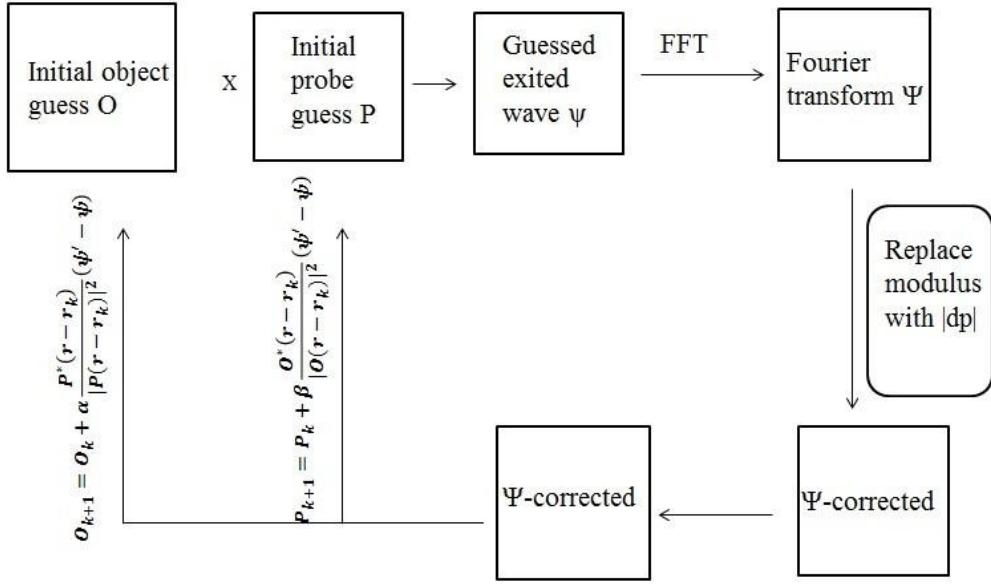


Figure 2.7.8. The flow chart of ePIE. See text for details.

The real space overlap constraint is that

$$\psi_j = O_j P_j, \quad (2.7.1)$$

where O is the object function; P is the illumination function; ψ is the exit wave, and j denotes the scanning positions.

The exit wave is a complex function is that

$$\psi_j = |\psi_j| e^{i\varphi_j}, \quad (2.7.2)$$

where φ_j is the phase.

Supposing we measure the diffraction patterns in the far field, the Fourier transform of the exit wave is

$$\Psi_j = \mathcal{F}\{\psi_j\}. \quad (2.7.3)$$

Then we use the measured intensities to constrain Ψ in the Fourier space by replacing the modulus of Ψ with the square root of the measurements I while keeping the phase as

$$\Psi'_j = \sqrt{I_j} \cdot \frac{\Psi_j}{|\Psi_j|}. \quad (2.7.4)$$

Then we propagate Ψ'_j back to the real space. Thus, we obtain the updated exit wave function as

$$\psi'_j = \mathcal{F}(\Psi'_j). \quad (2.7.5)$$

Then we update the illumination function and the object function as

$$O'_j = O_j + \alpha \frac{P_j^*}{|P_j|^2} (\psi'_j - \psi_j),$$

$$P'_j = P_j + \beta \frac{O_j^*}{|O_j|^2} (\psi'_j - \psi_j), \quad (2.7.6)$$

where P^* , O^* are the conjugations of P and O ; α , β are the update coefficients between 0 and 1 to scale the searching step size.

Then the calculation moves to update the next position until the whole area is reconstructed; this is a complete iteration. A reconstruction usually requires several hundred of iterations.

The error metric in ePIE is that (Maiden and Rodenburg, 2009)

$$\text{error} = \frac{\sum_j \sum_u [\Psi_j(u) - \sqrt{I_j(u)}]^2}{(\sum_j \sum_u I_j(u))^2}, \quad (2.7.7)$$

where j labels the position of the probe, and u are the pixels on the detector.

In conclusion, if we define π_o as the constraint in real space, and π_f as the constraint in Fourier space, then the reconstruction procedure of ePIE can be described as

$$\psi_{n+1} = \psi_n + \alpha \Delta = \psi_n + \alpha \pi_f[\pi_o(\psi_n)], \quad (2.7.8)$$

where Δ is the searching direction, and α scales the step size.

Difference Map search

The other popular iterative algorithm in the ptychography community is Difference Map (Thibault et al., 2008), the update function of which follows the conjugate gradient direction.

The searching procedure is shown in Figure 2.7.9. Supposing that the initial guess is at position 1, if it is projected to constraint a first, and then it is projected to constraint b , we will get the updated solution A_{ab} ; but if it is projected to constraint b first, and then it is projected to constraint a , we will get another updated solution A_{ba} . If both A_{ab} and A_{ba} are the solution, then no matter which constraint we apply first, they should be the same.

However, if A_{ab} and A_{ba} are not the same, it means neither of them is the right solution. In this case, we take the differential vector of A_{ba} and A_{ab} as the updating direction for the next step (Marchesini, 2009).

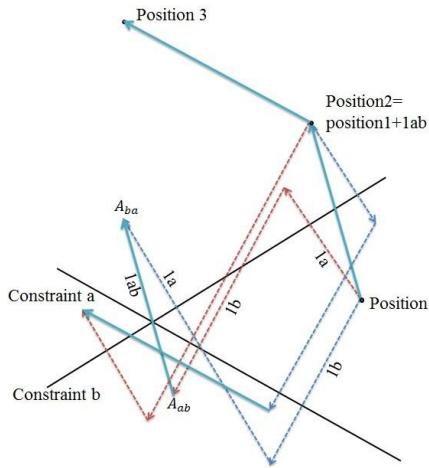


Figure 2.7.9 DM search. Position 1 is the initial guess. A_{ab} is the updated solution when the initial guess is projected to Constraint a first, and then projected to Constraint b ; A_{ba} is the updated solution when the initial guess is projected to Constraint b first, and then projected to Constraint a . The Vector from A_{ab} to A_{ba} is the searching direction of the Difference Map.

We can write the searching strategy of DM as

$$\Delta = \pi_o[\pi_f^*(\psi_n)] - \pi_f[\pi_o^*(\psi_n)], \quad (2.7.9)$$

where

$$\pi_f^*(\psi_n) = \pi_f(\psi_n) + \gamma_f[\pi_f(\psi_n) - \psi_n] = (1 + \gamma_f)\pi_f(\psi_n) - \gamma_f\psi_n,$$

$$\pi_o^*(\psi_n) = \pi_o(\psi_n) + \gamma_o[\pi_o(\psi_n) - \psi_n] = (1 + \gamma_o)\pi_o(\psi_n) - \gamma_o\psi_n, \quad (2.7.10)$$

where γ_f and γ_o are constants.

The DM update function is that

$$\psi_{n+1} = \psi_n + \alpha\Delta = \psi_n + \alpha\{\pi_o[\pi_f^*(\psi_n)] - \pi_f[\pi_o^*(\psi_n)]\}. \quad (2.7.11)$$

According to the paper by Marchesini (2009), when $\gamma_o = \alpha^{-1}$ and $\gamma_f = -\alpha^{-1}$, the optimal convergence is achieved.

2.7.3.2 Improvement in ePIE

Scanning position correction

The scanning positions are the very important constraints in the real space of ptychography (here we are not discussing Fourier ptychography) to locate the updated area at each position. The accuracy of the scanning positions directly decides the reconstruction quality of the sample. When we collect the data either by shifting the specimen or by shifting the illumination, there are some cases in which we cannot measure the positions accurately, because of 1) the hysteresis in the moving of the mechanical/piezo stage (when shifting the specimen) or the **electromagnetic** field (when shifting the illumination), 2) the drifting of the

specimen, which may be caused by the thermal expansion of the sample, the external mechanical vibration of the microscope, or the drifting of the illumination; 3) the calibration error of the stage/illumination moving pitch or the magnification, for example, a shifting step size of 3 nm may be calibrated as 3.2 nm. Of these three kinds of positions errors, 1 and 2 will introduce random errors to each position; 3 will introduce an error to the step size, which is a global scaling error to the whole positions map.

Some research has been done to investigate the influence of the scanning positions errors on the ptychographic reconstruction. Several scanning positions correction methods have been proposed. The paper by Faulkner and Rodenburg (2005) explored the influence of random positions errors on the error metric (sum squared error (SSE) between the calculated intensities and the detected intensities) in the PIE (Rodenburg and Faulkner, 2004) reconstruction; The paper by Hue et al. (2011) investigated the influence of random and the global scaled positions errors in ePIE reconstruction. The conclusion of these two investigations was consistent: the scanning positions are the dominant factor that determines the reconstruction quality in the experimental parameters. These papers (Hurst et al., 2009; F. Zhang et al., 2013; Maiden et al., 2012; Tripathi et al., 2014; Marchesini and Schi^Ozek, 2012) proposed different methods to correct the random and global positions errors, which can be classified as 1) the annealing method (Maiden et al., 2012), the idea of which is to generate several positions randomly around the original position, then do the calculation with all the generated and the original positions, find out the position which has the minimum error, take it as the current correct position. The limit of the annealing method is the computing cost. 2) cross correlation of the exit wave before and after the updating at each position (F. Zhang et al., 2013); the idea is based on that, when there is an error in the measured position, the exit wave after being enforced with the Fourier constraint, will appear shifted compared with the exit wave that has not been applied the Fourier constraint; this shift

is used as a guide for the position correction. The limit of this method is that the specimen has to be very structured. 3) cross correlation of the convergent beam diffraction patterns (Hue et al., 2011); when the illumination is the convergent beam, the diffraction patterns in the far field will show some structures of the sample inside the central discs; by cross correlating the structures inside the central discs, we can achieve the shift between each two adjacent positions. The limit of this method is that the specimen has to be structured, and it is difficult to calculate the dimension of the pixel in the diffraction pattern plane; moreover, this method is only limited to the convergent beam patterns. 4) cross correlation the Gabor holograms (Hue et al., 2011); the idea is that propagating each diffraction pattern to the real space to get the Gabor holograms at each position, and then cross correlate these Gabor holograms; the limits of this method are the same as the method 3. 5) The conjugate gradient method (Tripathi et al., 2014).

The scanning positions correction method applied in this thesis is the annealing method – ‘jiggle’ (Maiden et al., 2012). The realization of jiggle with ePIE (Maiden et al., 2009; Maiden et al., 2012) has the following steps. j denotes the number of the positions; m denotes the iterations; k denotes the number of annealing positions at each measured position.

- 1) Around each original position, we generate k positions randomly at a certain radius of r . As more iterations the algorithm runs, the corrected positions will get closer to the true positions, then the radius r will be reduced smaller. In Figure 2.7.10, at the j^{th} measured position in iteration m^{th} , $x_0(j, m)$, k annealing positions $x_k(j, m)$ are generated at the radius r_m .

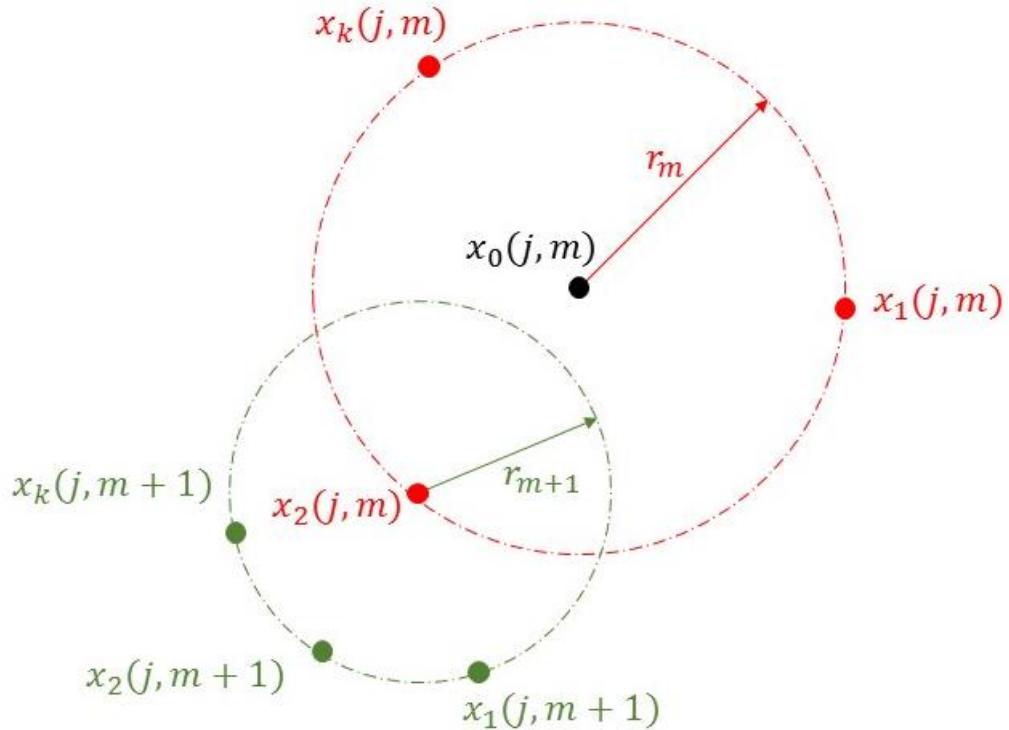


Figure 2.7.10. The red circle shows the radius of a set of annealing positions that are generated at the j^{th} measured position x_0 in the m^{th} iteration. The green circle shows where the annealing positions at the corrected position x_2 when the algorithm runs the $(m + 1)^{th}$ iteration.

- 2) Then for the forward propagation, we in total have $k+1$ exit waves ψ_{k+1} at each position; the $k+1$ exit waves are propagated to the reciprocal space, to generate the $k+1$ calculated intensities $|\Psi_{k+1}|^2$.
- 3) We calculate the errors between the $k+1$ calculated intensities and the detected intensity I by

$$e_{k+1} = \sum_u (|\mathcal{F}(\sqrt{I})| - |\mathcal{F}(|\Psi_{k+1}|)|)^2 \text{ or } e_{k+1} = \sum_u ((I - |\Psi_{k+1}|^2)^2); \quad (2.7.12)$$

- 4) We compare the $k+1$ error metrics; the one with the minimum error will be chosen as the currently corrected positions for the m^{th} iteration.
- 5) We apply steps 1-4 to each position, to find out the corrected positions for each measurement in the m^{th} iteration, then the calculation moves to the $(m + 1)^{th}$

iteration, generates a set of k new annealing positions at each corrected position from the m^{th} iteration with the radius of r_{m+1} , as shown as the green circle in Figure 2.7.12.

Super-resolution

In the diffraction imaging field, super-resolution means to obtain higher resolution by extracting more information from the redundancy of the extant measurements (Gerchberg, 1974; Lindberg, 2012). In this section, we discuss the application of super-resolution in ptychography.

In theory, the resolution of diffraction imaging is only limited by the size of the detector (the spectrum truncation), which is called as the diffraction resolution limit. However, it has been discovered that, in fact, the reconstructed resolution from diffraction imaging can be beyond the diffraction resolution limit. This is because of the redundancy in the diffraction pattern measurement(s). In the conventional CDI with a single exposure, when the sampling frequency of the diffraction pattern is higher than Shannon-Nyquist sampling criteria (we call this as over-sampling), there will result in redundancy in the data. The iterative algorithm of super-resolution reconstruction was firstly proposed by Gerchberg in the conventional coherent diffraction imaging field (Gerchberg, 1974). Basically, the idea was that we extend the detector with a larger size than the actual measurement; when the exit wave is propagated alternately between the real and the reciprocal space, we apply the constraints only to the area we have measured, and keep the area unmeasured as the calculated values (Greenbaum et al., 2013). The division of the maximum detector size that can be extended by the original detector size, is called super-resolution ratio.

$$\sigma_r = \frac{\text{extended dimension of the detector}}{\text{original dimension of the detector}}; \quad (2.7.13)$$

σ_r is decided by the redundancy in the data (Gerchberg, 1974).

The iterative computing of super-resolution of Gerchberg was introduced to ptychography by Maiden (Maiden et al., 2012). Since the redundancy in ptychographic data is much higher than single exposure CDI, due to the multiple measurements and the real space oversampling, the super-resolution ratio is also higher in ptychography. For a specific ptychography reconstruction, σ_r is decided by the overlapping in real space, the sampling in the detector, the illumination structure, the specimen structure, the illumination coherence and the detector noise. We have applied super-resolution method of the reconstructions in Chapter 4 and Chapter 6.

Up-sampling

The sampling criteria of conventional CDI is Shannon-Nyquist sampling - the sampling frequency is no less than twice the highest frequency of the signal (Zuo and Huang, 2012).

$$\Delta\theta \leq \lambda/2D, \quad (2.7.14)$$

where $\Delta\theta$ is collecting angle of a single pixel; λ is the wavelength; D is the object or the probe width.

The sampling criteria in ptychographic reconstruction are relaxed, because of the over sampling in the real space - each unit of the sample is measured more than once in a ptychographic experiment. Papers (Edo et al., 2013; Batey et al., 2014) gave the sampling criteria of ptychography:

$$\Delta\theta\Delta R \leq \frac{\lambda}{2}, \quad (2.7.15)$$

where $\Delta\theta$ is the sampling at the detector; ΔR is the sampling in the real space.

The idea of up-sampling is that when the sampling at the detector is coarse, which means that when the pixel dimension does not satisfy the Nyquist sampling, with the ptychographic reconstruction, we can divide each detector pixel into several smaller pixels; thus, we generate a synthetic detector which has the same dimension with the original detector, but each pixel has a smaller size. To divide the number of pixels in the synthetic detector by the number of pixels in the original detector, is the up-sampling ratio σ_s (Batey et al., 2014).

$$\sigma_s = \frac{\text{number of pixels in the synthetic detector}}{\text{number of pixels in the original detector}}. \quad (2.7.16)$$

Similar to the super-resolution ratio, the maximum up-sampling ration is also decided by the redundancy in the data.

Supposing that a pixel m in the original detector is divided into N sub-pixels ($\sigma_s = N$), during the ptychographic reconstruction with ePIE, in the Fourier domain, we need to replace the modulus of the calculated diffraction spectrum $\Psi(m, n)$ with the square root of the detected intensities I_m , thus the modulus of each sub-pixel is assigned as

$$\Psi'(m, n) = \frac{\Psi(m, n)\sqrt{I_m}}{\sqrt{\sum_{n=1}^N |\Psi|_{m,n}^2}}, \quad (2.7.17)$$

where m is the pixel in the original detector; n labels the sub-pixel in the synthetic detector (Edo et al., 2013; Batey et al., 2014).

In the reconstructions of this thesis, our sampling criteria satisfy the Nyquist sampling, so we have not applied up-sampling methods.

Diffraction patterns drifting correction

Diffraction pattern drifting is a very common phenomenon in the electron microscopy. The drifting may be caused by 1) the illumination instability, 2) the specimen drifting for the

Fresnel patterns and 3) the instability of the imaging lenses, which are the lenses that project the back focal plane of the diffraction lens to the detector with a certain magnification.

During a ptychographic data acquisition, for example, the data collected in Chapter 3, the collecting of 15*15 SAD patterns took 20-30 minutes; the drifting distance from the first diffraction pattern to the last diffraction pattern was 30-50 pixels. The diffraction patterns drifting will relax the constraints in the real and the reciprocal space, thus, it will cause distortion in the reconstruction. The drifting that is caused by the illumination and specimen will result in the inaccuracy of the scanning positions measurements; we have discussed the correction methods in the previous section. Here we discuss the method to solve the diffraction patterns drifting caused by the imaging lens instability.

The method to correct the drifting caused by the imaging lens we applied in this thesis is to cross correlate the calculated spectrum $I_{cal} = |\Psi|^2$, with the detected pattern I at each position, find out the drifting distance between the detected pattern and the calculated pattern in each position, and then shift the detected pattern in line with the calculated pattern before we apply the Fourier constraint. Here are the procedures at each position as the following steps (Maiden et al., 2012).

- 1) Cross correlation I_{cal} and I :

$$X = \sum_r I_{cal}(r)I^*(r - r_0); \quad (2.7.18)$$

- 2) Find out the location of the maximum value of X ;

$$[u_0, v_0] = FindCoordinate\{X_{max}\};$$

- 3) Shift I with u_0, v_0 in each direction.

We have applied the diffraction patterns drifting correction to the reconstructions of the experimental data in Chapter 3, 4 and 5.

Modal decomposition and *Blind deconvolution*

Modal decomposition is a method to solve the mixed state issue in the coherent diffraction imaging reconstruction (Wolf, 1982). The mixed state of an imaging system is composed of 1) the partial coherence of the source, 2) the detector point spread, and 3) the specimen multiple states (Thibault and Menzel, 2013). To describe the mixed state of an imaging system, here we borrow some definitions from the coherence theory (Born and Wolf, 1999). We define the mutual coherence function of an imaging system as

$$\Gamma = \frac{Q_1 Q_2}{\sqrt{Q_1^* Q_1} \sqrt{Q_2^* Q_2}}, \quad (2.7.19)$$

where Q_1 is not only a state that describes the partially coherent source, but a state that describes the transfer function of the imaging system, which includes the source partial coherence, the detector point spread and the specimen multiple states.

A mixed-state imaging system will result in the blurring images or diffraction patterns, thus, some features or frequencies in the images or diffraction patterns will be lost. The effect of partial coherence can be mitigated at the expense of reducing the counts by inserting condenser apertures or taking the source optically further away during the electron imaging. There is no way in the direct imaging field to remove the effect of partial coherence completely.

Supposing that the detected intensity (diffraction pattern or image) of a pure state imaging system is I_c ; a mixed-state imaging system can be decomposed into K states Q_k (Q_k can be a mixed state or pure state). There are two ways to express the detected intensity from a mixed-state imaging system I_{pc} – the modal convolution and the modal decomposition (Zernike, 1938; Whitehead, 2009; Suvorov, 2010; Clark and Peele, 2011; Thibault and Menzel, 2013).

1) Model convolution

$$I_{pc} = I_c \otimes G , \quad (2.7.20)$$

where G is the model of the mixed states; the common model is Gaussian profile, or Cauchy profile, or Gaussian-Cauchy profile.

2) Modal decomposition

$$I_{pc} = \sum_k |\mathcal{F}(Q_k \cdot O)|^2 = \sum_k I_k, \quad (2.7.21)$$

where O is the object function; \mathcal{F} is the propagation operator; when the detector is in the far field, \mathcal{F} is Fourier transform operator.

The model convolution is an approximation measurement of the mixed state; the modal decomposition is a complete measurement of the mixed state. The computing procedures of the two methods based on ePIE in the ptychographic reconstruction are as follows.

The computing steps of the model (Gaussian) blind convolution with ePIE (Maiden et al., 2015):

- (1) Guess an arbitrary initial object function O , pure state illumination function P and an arbitrary 2-D Gaussian model;
- (2) At j^{th} scanning position, form an exit wave $\psi_j = P \cdot O_j$;
- (3) Propagate the exit wave to the detector space $\Psi_j = \mathcal{F}\{\psi_j\}$, and get the calculated intensities $I_j = |\Psi_j|^2$;
- (4) Generate several random standard derivations for the Gaussian profile, keep the one that matches I_j and the detected intensities I_j^{pc} with the minimum error as the current correct Gaussian profile $I_j^{pc} \leftrightarrow I_j \otimes G$.

- (5) Update Ψ_j with the Fourier constraint $\Psi'_j = \sqrt{I_j^{pc}} \frac{\Psi_j}{|\Psi_j|}$, then propagate back to real space, update the object function and illumination function;
- (6) Move to next position, repeat (2-5), till update all the positions, this is a complete iteration;
- (7) Repeat (2-6) for several hundreds of iterations.

The computing steps of the modal decomposition with ePIE (Li et al., 2016):

- (1) Guess an arbitrary object function and k arbitrary illumination function P^k ;
- (2) At the j^{th} position, form k exit wave $\psi_j^k = O_j \cdot P^k$;
- (3) Propagate k exit waves to the detector plane $\Psi_j^k = \mathcal{F}\{\psi_j^k\}$; Apply Fourier constraint to each sub exit wave by $\Psi_j^{k'} = \sqrt{I_j^k} \frac{\Psi_j^k}{|\Psi_j^k|}$, where $I_j^k = I_j \cdot \frac{|\Psi_j^k|^2}{\sum_k |\Psi_j^k|^2}$;
- (4) Propagate back to the real space $\psi_j^{k'} = \mathcal{F}^{-1}\{\Psi_j^{k'}\}$; update the object function and each probe function by

$$P'_k = P_k + \alpha \frac{\psi'_k \cdot O^*}{|O|_{max}^2}, O' = O + \beta \cdot \frac{\sum_k (\psi'_k \cdot P_k^*)}{\sum_k |P_k|^2}; \quad (2.7.22)$$

- (5) Repeat (2-4) over all the positions;
- (6) Repeat (2-5) for several hundreds of iterations;
- (7) Orthogonalize the k probe functions, to obtain the eigenstates of the imaging system, which is the minimum representation required

$$\sum_k Q_k = \sum_m M_m, \text{ where } m \leq k, \text{ and } M_m \text{ are orthogonal with each other.}$$

Detector pedestal

The pixels on the detector transform the electrons or photons that arrive on them as the digital signal.

$$S_{out} = H\{S_{in}\},$$

where H is the transfer operator. Each pixel has a transfer function. Ideally, the transfer function of the detector is unity. The operator H can be represented in the form of Taylor expansion.

$$S_{out} = A + \alpha \cdot S_{in} + \beta \cdot S_{in}^2 + \gamma \cdot S_{in}^3 + \dots, \quad (2.7.23)$$

where A is the constant term (detector pedestal); $\alpha \cdot S_{in}$ is the linear term; the others are the nonlinear terms. The constant term and nonlinear terms will result in the signal distortion. Distorted diffraction patterns will relax the constraint in the reciprocal space, thus, it will give ambiguous solutions in iterative ptychographic reconstruction. Here we discuss the methods that have been applied with ePIE in this thesis to solve the detector pedestal issue.

In the Fourier space, supposing that the detected intensities are $I(u)$, where u represents the pixel locations, the calculated intensities $I_{cal}(u) = \sum_k |\Psi_k(u)|^2$, where k represents the states, and the detector pedestal $C(u)$. At each scanning position, the algorithm will find out $C(u)$ that match $I_{cal}(u)$ with $I(u)$: $I(u) \leftrightarrow I_{cal}(u) + C(u)$. There are two ways to update $C(u)$: using an extra mode or scale matching.

An extra mode is the same with the modal decomposition concept. We suppose that there is an extra probe mode and extra object mode, which satisfy $C(u) = (\mathcal{F}\{P_e \cdot O_e\})^2$. We update $C(u)$ by updating P_e and O_e at each position the same as we update each probe function and object function.

Scaling matching is to set an arbitrary initial background $c(u)$ with an arbitrary scaling number η : $C(u) = \eta \cdot c(u)$; at each position, we update $c(u)$ by

$$c(u) = \gamma(I(u) - I_{cal}(u)), \quad (2.7.24)$$

$$\gamma = \frac{\sum_u [I(u) \cdot I_{cal}(u)]}{\sum_u [I(u) \cdot I(u)]}. \quad (2.7.25)$$

In this thesis, we have applied both the extra mode method and the scaling matching method to retrieve the detector pedestal.

2.8 Conclusion

In this Chapter, we have introduced the theory background of electron ptychography from three aspects. In the first part, we introduced the background of the Fourier optics, which includes the electron waves function, its propagation and diffraction. The whole calculation of this thesis is based on the Fourier optics. Furthermore, we also introduced the imperfection of the wavefronts, the transfer function of an imaging system and the representation of the partial coherence of the illumination. This is important to understand the simulation of an imaging system. Since all the experiments in this thesis are executed on the electron microscopes, in the second part, we introduced the background of the electron microscopy. We introduced the functions of each part of the microscope, which includes the function of the source, the lenses and the detector. We also described the transmission electron microscope (TEM), the scanning electron microscope (SEM) and the scanning transmission electron microscope (STEM) both in the imaging mode and in the diffraction mode. This is important to understand the imaging system of an electron microscope and hence the parameters in the experiments. The third part is to introduce ptychography, which is the technique we are going to investigate in this thesis. We presented the developments of the electron ptychography since 1989 both from the experimental setup and from the reconstruction algorithms. We introduced the setups that can be used to implement the electron ptychography and the algorithms that we employed in this thesis in detail.

Chapter 3 Ptychography in SAD mode

3.1 Modal decomposition of a propagating matter wave

This paper is with the cooperation of other authors Dr. P. Kok, Dr. P. Li, Dr. Andrew Maiden and Prof. John Rodenburg.

Dr. P. Kok contributed the quantum mechanics theory part of the paper.

As the first author of the paper, I carried out the experiments, the reconstruction, and the writing of the first draft, which was reviewed and revised by my supervisor Prof. John Rodenburg.

Dr. A. M. Maiden and Dr. P. Li provided the original version of the Gaussian blind deconvolution code and the modal decomposition code.

Modal decomposition of a propagating matter wave via electron ptychography

S. Cao¹, P. Kok², P. Li¹, A. M. Maiden¹, J. M. Rodenburg¹

¹*Department of Electronic and Electrical Engineering, University of Sheffield, S1 3JD, United Kingdom*

²*Department of Physics and Astronomy, University of Sheffield, S3 7RH, United Kingdom*

Abstract:

We employ ptychography, a phase-retrieval imaging technique, to show experimentally for the first time that a partially coherent high-energy matter (electron) wave emanating from an extended source can be decomposed into a set of mutually independent modes of minimal rank. Partial coherence significantly determines the optical transfer properties of an electron microscope and so there has been much work on this subject. However, previous studies have employed forms of interferometry to determine spatial coherence between discrete points in the wavefield. Here we use the density matrix to derive a formal quantum mechanical description of electron ptychography and use it to measure a full description of the spatial coherence of a propagating matter wavefield, at least to the within the fundamental uncertainties of the measurements we can obtain.

1. INTRODUCTION

We investigate experimentally the decomposition of a freely-propagating partially coherent electron wave into a set of mutually orthogonal wave function components (modes) which are incoherent with each other. By diagonalising the density matrix, we measure the most compact representation of this mixed state. We use the decomposition to determine explicitly, without making any assumptions, the spatial coherence properties of a field emission electron source, at least within the quantum mechanical uncertainties of the wave we physically measure, which propagates through an aperture of finite lateral extent. Although the principle of the modal decomposition is well known in optics (visible light and X-ray) [1,2] including partial coherence [3, 4], and its relation to the density matrix formalism has been mentioned in the context of X-ray ptychography [5], this is the first time it has been demonstrated experimentally for a matter wave.

The partial coherence of a high-energy electron wave has critical importance in the field of electron microscopy. It defines the transfer properties of the imaging lens in a transmission electron microscope (TEM), determining the extent of the ‘information limit’ – i.e. the highest spatial frequency that can be transferred through the electron optics, and hence the highest resolution that can

be obtained in the image [6,7]. This is particularly important in the field of imaging nanoscale biological structures such as large proteins and viruses because contrast in the image of such weakly-scattering structures must be enhanced by large defocus in the object lens [8], which, in the presence of partial coherence, greatly reduces the width of the information limit and hence limits resolution. In the case of scanning transmission electron microscopy (STEM), the intensity of the focused probe is convolved with the demagnified source intensity, which also imposes a limit in resolution. Perhaps most important of all, the entire field of electron holography depends wholly upon an interference effect that is determined by the coherence properties of the electron wave [9].

In the field of light optics, a common approach for quantifying coherence is to consider the temporal correlation and the underlying statistics of fluctuations of the electric field between pairs of points separated in space. If the source is quasi-monochromatic and spatially localised (as in the experiments described here), many approximations can be employed, greatly abridging the formal treatment. In particular, lateral spatial coherence between points lying within a plane perpendicular to the direction of propagation can be characterised by a single two-dimensional, shift invariant function: the mutual coherence function [10]. For electrons one way to achieve a quantitative measurement of this is to place a wire, which acts as a bi-prism, within the plane of the wavefront, as used in electron holography [9]. As a voltage on the wire is increased, parts of the wavefield either side of it are deflected so that increasingly separated parts of the original wave overlap with one another in a measurement plane downstream of the beam splitter. The degree of coherence is expressed in the decreasing depth of the resulting interference fringes as a function of bi-prism voltage [11]. The incoherent source profile can then be obtained via Fourier transformation of this function (i.e. by using the van Cittert-Zernike theorem [10]). Another method of achieving the interference (this time in reciprocal space) is via convergent beam electron diffraction [12], where a crystal is used as an interferometer. Other less direct methods, e.g. by observing the depth and extent of Fresnel fringes in the image plane, can also be used [13].

Partial coherence can alternatively be modelled as a series of independent modes that are entirely mutually incoherent with respect to one another, but which propagate through the optical system

independently as pure coherent states [1,2]. This methodology, which we adopt here, is widely used in the optics community, say for describing modes in a laser source [14]. In quantum mechanics, modes in a matter wave are more naturally – though equivalently – handled by the density matrix formalism [15]. In a particular mode decomposition, the density matrix for the electron wave is not generally diagonal, with off diagonal terms arising from coherence between the modes. However, there is a unique mode decomposition in which the density matrix is diagonal and where the associated eigenmodes are perfectly incoherent and orthogonal with respect to one another (assuming non-degenerate eigenvalues). If the state in each mode is normalised, the eigenvalues of the density matrix are proportional to the probability of finding the system in that mode.

In this paper, we measure the diagonalised density matrix of a freely-propagating mixed state electron wave after it has passed through a confining aperture. We calculate the orthogonal pure states relating to the diagonal terms as they express themselves within the plane of the aperture. This represents the most compact and complete description of the spatial partial coherence in the beam. We demonstrate that as the preparation of the incoherent source is changed (by optically altering its shape) the relative probability of these states change as expected.

The experimental technique we employ is a form of electron ptychography [16-18]. We place a two-dimensional object (a transmission electron microscope test specimen) within the electron beam and measure many diffraction patterns scattered from it as it is moved relative to the aperture. (In fact, the aperture lies in an image plane of the specimen, but the measured data are equivalent.) The distance of the specimen movement between exposures is much less than the size of the aperture, so any one point in the object contributes to many diffraction patterns. By applying the constraints that the illuminating wave and the object function remain constant, it is well known that it is easy to solve for both the object and illumination [19,20]. If the illumination is partially coherent, it is impossible to separate the corresponding intensity at any single detector pixel in any single diffraction pattern into separate modes. However, ptychography makes many measurements, sampling diverse scattering conditions as the object is moved laterally, and since the illumination (and the modes within it) remain constant throughout the process, this gives sufficient data to decompose the illumination modes, as

previously shown in the case of photons [5,21]. As an imaging technique that offers wavelength-limited resolution and a very sensitive phase image, electron ptychography is hampered by the fact that electron sources are substantially incoherent: although field emission sources are often described in the literature as fully spatially coherent [16,22], in practice they are not. In fact, the sort of modal decomposition we apply here can be used to mitigate the effects of partial coherence, which can be used to improve any result obtained by electron ptychography [22].

2.THEORY

Consider an idealised experimental setup illustrated in Figure 1.a. An incoherent source of electrons of finite size creates a freely propagating wave that impinges upon an aperture. We assume that the distance from the source to the aperture is sufficiently large so that the aperture is effectively lying in the Fraunhofer plane relative to the source. A detector lies downstream of the aperture at a large enough distance so that it lies in the Fraunhofer plane of the aperture.

Even though there are many electrons in the beam, they can be considered non-interacting and are therefore well-described by single particle states. First suppose that the electrons are each in a pure state. We may choose to expand this state in any set of basis functions; for the present discussion we use as an example

$$|\psi_S\rangle = \sum_{lm} d_{lm} |l,m\rangle , \quad (1)$$

where $\{|l,m\rangle\}$ is the Laguerre-Gaussian transverse mode expansion such that $\langle x|l,m\rangle \equiv L_{lm}(x)$ are the Laguerre-Gaussian mode functions. These mode functions are convenient in describing approximately cylindrically symmetric sources. When the state propagates to the aperture plane, it becomes

$$|\psi_A\rangle = \mathfrak{I}|\psi_S\rangle = \sum_{lm} \gamma_{lm} |l,m\rangle , \quad (2)$$

where \mathfrak{I} is the Fourier transform, and γ_{lm} are the transformed d_{lm} . The aperture A will block the electron beam at certain positions according to the function

$$q(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \quad (3)$$

The aperture can then be described by the high-rank projector

$$P_A = \sum_x q(x) |x\rangle\langle x|. \quad (4)$$

After the aperture, the state must be renormalized, and becomes

$$|\Psi\rangle = \frac{P_A |\psi_A\rangle}{\sqrt{\|P_A \psi_A\|}} \equiv \eta^{-\frac{1}{2}} P_A |\psi_A\rangle, \quad (5)$$

where η indicates the transmittance of the aperture given the state $|\psi_A\rangle$, or how many of the electrons in the state pass through the aperture. We can write $|\Psi\rangle$ in terms of the position states $|x\rangle$ at the aperture:

$$|\Psi\rangle = \eta^{-\frac{1}{2}} \sum_x \sum_{lm} q(x) \gamma_{lm} L_{lm}(x) |x\rangle. \quad (6)$$

In the detector plane the state is modified by another Fourier transform, and becomes

$$|\Psi_D\rangle = \mathfrak{I}(\Psi) = \eta^{-\frac{1}{2}} \sum_x \sum_{lm} q(x) \gamma_{lm} L_{lm}(x) \mathfrak{I}|x\rangle. \quad (7)$$

Finally, the intensity at the position u in the detector plane is

$$I(u) = |\langle u | \Psi_D \rangle|^2 = \eta^{-1} \left| \sum_{lm} q(x) \gamma_{lm} L_{lm}(x) \langle u | \mathfrak{I} | x \rangle \right|^2. \quad (8)$$

The matrix elements of the Fourier transform are

$$\langle u | \mathfrak{J} | x \rangle = \frac{1}{N} \exp \left[\frac{2\pi i (xu_x + yu_y)}{N} \right], \quad (9)$$

where N is the number of pixels of the detector. This describes the case of a source of completely coherent electron wave functions.

We now extend the above treatment to take account of the extended nature of the source, which renders the electron beam partially coherent, by considering the case of the mixed states. The density matrix for single electron state for an extended source is a probabilistic mixture of pure states, and can be written as

$$\rho_s = \sum_j p_j |\psi_{S,j}\rangle \langle \psi_{S,j}| \quad \text{with} \quad |\psi_{S,j}\rangle = \sum_{lm} d_{lm}^{(j)} |l,m\rangle, \quad (10)$$

and in the aperture plane just before the state becomes

$$\rho_A = \mathfrak{J} \rho_s \mathfrak{J}^{-1}. \quad (11)$$

Applying the aperture projector P_A as before, the state must be renormalized to

$$R_A = \frac{P_A \rho_A P_A}{\|P_A \rho_A P_A\|} = \eta^{-1} \sum_{x,x'} q(x) q(x') \rho_A(x, x') |x\rangle \langle x'|, \quad (12)$$

where $\rho_A(x, x') = \langle x | \rho_A | x' \rangle$. In the detector plane, we again apply a Fourier transform to R_A to obtain

$$R_D = \mathfrak{J} R_A \mathfrak{J}^{-1} = \eta^{-1} \sum_{x,x'} q(x) q(x') \rho_A(x, x') \mathfrak{J} |x\rangle \langle x'| \mathfrak{J}^{-1}. \quad (13)$$

The intensity at position u in the detector plane then becomes

$$\begin{aligned} I(u) &= \text{Tr} [|u\rangle \langle u| R_D] = \eta^{-1} \sum_{x,x'} q(x) q(x') \rho_A(x, x') \langle u | \mathfrak{J} | x \rangle \langle x' | \mathfrak{J}^{-1} | u \rangle \\ &= \frac{1}{\eta N^2} \sum_{x,x'} q(x) q(x') \rho_A(x, x') e^{2\pi i [u_x(x-x') + u_y(y-y')] / N} . \end{aligned} \quad (14)$$

The matrix element $\rho_A(x, x')$ can be expressed as

$$\rho_A(x, x') = \sum_j \sum_{l_j, m_j} \sum_{l'_j, m'_j} p_j \gamma_{l_j m_j}^{(j)} \gamma_{l'_j m'_j}^{(j)*} L_{l_j m_j}(x) L_{l'_j m'_j}^*(x). \quad (15)$$

Note that we can in principle use different transverse mode functions $L(x)$, which will have corresponding different amplitudes γ .

Finally, we include an object with complex transmission function $O(x)$ immediately following the aperture. The transmission operator acting on the state can be written as

$$\hat{E}_O(v) = \sum_x O(x - v) |x\rangle\langle x| \quad \text{with} \quad 0 \leq |O(x)| \leq 1. \quad (16)$$

Moving the object in the aperture plane along a vector v_j modifies the object function according to

$$O(x) \rightarrow O(x - v_j). \quad (17)$$

We assume that the aperture and object are sufficiently close together that there is negligible propagation between them, and $[P_A, \hat{E}_O(v)] = 0$ for all v . The intensity $I_j(u)$ at the position u in the detector plane for the j^{th} object shift is then given by

$$I_j(u) = \frac{1}{\eta_j N^2} \sum_{x, x'} q(x) q(x') O(x - v_j) O^*(x' - v_j) \rho_A(x, x') \times \exp \left\{ \frac{2\pi i [u_x(x - x') + u_y(y - y')]}{N} \right\}, \quad (18)$$

where

$$\eta_j = \text{Tr} \left[|\hat{E}_O(v_j)|^2 P_A \rho_A(x, x') \right] \quad (19)$$

is the modified normalisation function due to the shifted object function. In the case of electron scattering through a thin object, the modulus of the transmission function $O(x)$ is often assumed to be unity. The phase of the transmission is also often assumed to be weak (much less than π), although in practice it is often larger than this. To a good approximation we can therefore write $\hat{E}_O(v_j) = \exp[i\varphi(x)]$, and as a consequence $\eta_j = \eta$ in this case.

The diffraction patterns in equation (18) are fully specified by the aperture function $q(x)$, the object function $O(x)$ and the state ρ_A whose matrix elements are given by equation (15).

Rather than expanding the transverse mode shape of the electron beam in Laguerre-Gaussian modes, we can expand it into eigen-modes of the density operator. Formally, this can be written as

$$\rho_S = \sum_k s_k |m_k\rangle\langle m_k|. \quad (20)$$

Ptychography can determine the probabilities s_k and the mode functions $m_k(x) = \langle x|m_k\rangle$, as described below. We assume for simplicity that the s_k are non-degenerate, so that the $|m_k\rangle$ are unique. In practice, this is typically a valid assumption.

Just before the aperture, the density matrix becomes

$$\rho_A = \sum_k s_k \mathfrak{I} |m_k\rangle\langle m_k| \mathfrak{I}^{-1}. \quad (21)$$

The matrix element $\rho_A(x, x')$ can be written as

$$\rho(x, x') = \sum_k s_k \langle x | \mathfrak{I} | m_k \rangle \langle m_k | \mathfrak{I}^{-1} | x' \rangle. \quad (22)$$

If we define $\mathfrak{I}|m_k\rangle = |M_k\rangle$ and $\langle x|M_k\rangle = M_k(x)$, then M_k is the Fourier transform of m_k , and the matrix element takes on the simple form

$$\rho_A(x, x') = \sum_k s_k M_k(x) M_k^*(x'). \quad (23)$$

This expression can now be substituted into equation (18). Solving for $M_k(x)$ will allow us to retrieve the mode functions $m_k(y)$ via a simple inverse Fourier transform, where y are the transverse coordinates in the source plane.

As an aside, we can define the purity of the source as $\wp(\rho_s) = \text{Tr}[\rho_s^2]$, which is easily calculated as

$$\wp(\rho_s) = \text{Tr}[\rho_s^2] = \sum_k s_k^2 \leq 1. \quad (24)$$

This is a useful figure of merit for single-mode behaviour of the source. The smaller \wp , the more mixed the source. This is a standard quantity in quantum information theory [23].

In what follows, we will calculate the $M_k(x)$ from the experimental intensities $I_j(u)$ that we measure. Ptychography is an exceptionally powerful solution of the Fourier domain phase problem [24], which allows us to solve for both the amplitudes of the modes and their phases. The iterative reconstruction algorithm we employ is the widely-used ePIE [19]. At a given point during the iterative procedure, we have a running estimate of the object function and the illumination function (which in this case will be composed of multiple modes within the aperture). By computationally propagating the corresponding exit waves (one for each mode) to the detector plane, we obtain a set of estimated complex wavefunctions incident upon the detector, $\Phi_{j,k}(u)$, for a particular k^{th} mode and j^{th} specimen position. In general, the sum of the intensity of these estimated modes will not correspond to the measured intensity of our experimental data (equation 18), i.e. $I_j(u) \neq \sum_k |\Phi_{j,k}(u)|^2$. We therefore calculate new estimates of the complex modes incident upon the detector

$$\Phi'_{j,k}(u) = \frac{\Phi_{j,k}(u) \sqrt{I_j(u)}}{\sqrt{\sum_{k=1}^K |\Phi_{j,k}(u)|^2}}, \quad (25)$$

where $I_j(u)$ is the intensity component arising on the detector at the j^{th} position [5]. This constraint has the effect of scaling the intensity of each mode by a scaling factor so that their sum matches the measured intensity, while preserving their estimated (and immeasurable) phases. Next, each mode is back-propagated to object/aperture plane, where the object and each mode estimate are altered to be more closely commensurate with the Fourier constraint according to an update procedure, as shown in the flow diagram in Figure 2. The update function in real space is detailed for a single mode in [25]: the same update is in this case applied to all modes separately, but the object function is kept constant for all the updates. The process is repeated over all specimen positions many (typically a few hundred) times until an error metric, characterising the difference between the estimated detected intensity and the measured intensity reaches a certain threshold (see section 3 below); convergence properties are discussed in [25].

Diversity in the data, and the fact that the modes are added in intensity, ensures that each mode is independent (in the sense of being incoherent) with respect to every other mode. The modified Fourier constraint gives us K modes, where K is the number of illumination modes we have chosen to include in the reconstruction. There is an infinite combination of modes that can satisfy the overall intensity constraint, because they can be made up of any linear combination of each other [26].

However, the most compact description of the partially coherent state of the illumination is the set of orthogonal modes corresponding to the diagonalised density matrix (equation 20) and their corresponding illumination modes $M_k(x)$, of which there may be only κ pure states with significant probability, with $\kappa < K$. Indeed, if the experiment is perfectly coherent we have by definition $\kappa = 1$. If the reconstruction uses $K < \kappa$, then both the object and the mode reconstructions are imperfect, leading to a higher error metric: the partial coherence has not been fully accounted for. However, using a large K incurs a high computational overhead. The $M_k(x)$ give us the most compact representation of the partial coherence in the illumination, which is what we measure here.

It is important to note that the sum of the intensities of the source modes, $m_k(y)$, is not the same as the physical intensity distribution of the source, despite them being the back Fourier transform of the

$M_k(x)$. The modes derive from the entire experiment, i.e. the information that arrives at the detector. A full description of the source requires us to capture the entire wavefunction, including the part that falls outside the aperture. Or, equivalently, the source we reconstruct is the real source as seen backwards from detector, which is an image that is diffraction limited by the aperture.

3. EXPERIMENT

Provided position-dependent aberrations in the image forming optics (coma, etc.) can be ignored and the specimen is suitably thin, the idealised experiment of Figure 1.a can be realised in practice using the TEM geometry shown in Figure 1.b; the projection of the selected area aperture back through the objective lens gives a virtual aperture in the specimen plane equivalent to the pre-specimen aperture in Figure 1.a.

We implement the setup of Figure 1.b using 300keV electrons in a JEOL R005 high-resolution transmission electron microscope (TEM) equipped with a cold field emission gun and dual aberration correctors. A condenser aperture with a diameter of 100 μm was inserted to ensure a reasonable coherence of the electron wavefield illuminating the specimen, which was a thin layer of gold evaporated on amorphous carbon. The energy spread of the electron wave is small ($\Delta E/E < 10^{-5}$), meaning that the path length differences (maximum 10^{-7} m given the geometry of the experiment) between any of the interference phenomena we measure are much smaller than longitudinal coherence length (10^{-6} m).

Diffraction patterns were collected in the selected area diffraction mode at a calibrated camera length of 2.44m, using a 10 μm selected area aperture. Usually in this mode a beam stop is needed to prevent the intense central spot of the diffraction pattern from saturating or damaging the detector, resulting in loss of the low spatial frequency information that this bright peak contains. To combat this here, instead of imaging the diffraction pattern in the far-field Fourier domain, as would be conventional, we adjusted the microscope intermediate lens to image a plane at a slight defocus from the microscope's primary image plane, resulting in near-field diffraction patterns whose much lower dynamic range removes the need for a beamstop. Equations (9) and (15) therefore do not formally

apply, but this is not fundamental: any unitary operator (whether Fresnel or Fourier) can be used to propagate the $M_k(x)$ to the detector plane. In the reconstruction process we use the angular spectrum method [10] to calculate the near-field diffraction pattern before applying the intensity constraint in equation (25). For an example diffraction pattern, see Figure 3.

Diffraction patterns were recorded using a Gatan Orius SC2003 CCD camera with 2048×2048 pixels on a 7.4 μm pitch. The detector was binned by 2 during the acquisition and the central 512 by 512 region of each pattern extracted. The specimen was placed on a holder driven by a piezo motor with a minimum incremental step of 0.3 nm and was translated through a 15×15 grid of positions in a raster fashion, with a diffraction pattern captured by the CCD at each of the 225 specimen positions. When projected back to the specimen plane the selected area aperture covered a disc of 113nm on the specimen surface, so the position grid was programmed with a nominal step size of 30 nm to ensure significant overlap of the selected specimen areas between positions. The stage repeatability was determined to be much better than its absolute accuracy, so we obtained a reasonable initial estimate of the true position grid by removing the selected area aperture and recording a set of defocussed brightfield images with the same moving instructions as for the subsequent ptychography experiments. These images were then cross-correlated to give a set of relative offsets for use as initial grid positions for the ptychographic algorithms, which further improved these initial position estimates using a simulated annealing method [25].

Four sets of diffraction patterns were used for the results presented here: dataset 1 and dataset 2 were collected at spot size 2 and spot size 4 respectively; dataset 3 was collected at a stronger strength of the second condenser lens at spot size 3, while dataset 4 was collected at a weaker strength of the second condenser lens at spot size 3. The spot size, controlled by the condenser lenses, determines the angular size of the source as seen from the specimen and aperture, and hence the spatial coherence of the illumination. The exposure time was varied between each set of data to make sure the highest pixel value in any given diffraction pattern was within the linearity range of the detector. The longest exposure was 3 seconds (dataset 2) and the shortest was 1 second (dataset 1). Due to slowly varying diffraction lens current instability and a relatively long data collection time, the centre of the

diffraction patterns drifted during our experiments by around 50 pixels – this drift was corrected algorithmically during the reconstruction process, as detailed in [27].

The ‘ePIE’ implementation of the multi-mode ptychography algorithm [26] was used to process our data sets. However, for the first 500 iterations we used an approximate method for accounting for the partial coherence using a Gaussian convolution of the diffraction pattern, as described elsewhere [27,28]. With reference to Figure 4, which shows the error metric as a function of iteration number, we see that when the multi-mode method is switched on, the error drops significantly after an initial instability, indicating that the modal decomposition is much more effective at handling the coherence properties of the beam. We use $K=16$ illumination modes and a single object mode. It is computationally impractical to diagonalise the density matrix itself in the aperture plane (composed of 512×512 pixels), so after the reconstruction process is complete, we use an equivalent orthogonalisation method derived from principle component analysis. The modes are put into an X by K matrix, where X is the number of pixels in the aperture plane. We form the $K \times K$ covariance matrix of the modes, and then find its eigenvalues and eigenvectors. The K orthogonalised modes are then obtained by projecting the original modes onto the eigenvectors of the covariance matrix [26].

Figure 5 shows the first six components of the four partially coherent electron wavefields. The pixel sizes in the reconstructions are calculated according to

$$\delta u = \frac{\delta x \times D}{M_o \times d} , \quad (26)$$

where δx is the pixel size of the detector after binning; M_o is the magnification of the objective lens; D is the physical size of the aperture and d is the physical size of the central diffraction patterns disc. This gives pixel dimensions of 0.54nm, 0.48nm, 0.37nm and 0.38nm for datasets 1-4 respectively, and explains the discrepancy in the apparent size of the apertures shown in Figure 5. Figure 6 plots the eigenvalues of the reconstructed modes representing the partially coherent electron wavefield, which illustrates that as we change the strength of the condenser lenses, the degree of partial coherence changes: as the first condenser lens is strengthened or the second condenser lens is weakened, the

effective size of the source is smaller, which results in better transverse coherence. Figure 7 compares ptychographic reconstructions from the spot 2 dataset using the multi-mode method (with Figure 7.a, showing the phase and 7.b showing the amplitude of the reconstruction) versus a conventional single mode reconstruction (with Figure 7.c showing the phase and 7.d showing the amplitude). The phase contrast and resolution of the multi-mode specimen reconstruction are clearly improved over the single mode version. The effective source functions finally resulted from these reconstruction (i.e. the intensity summation of the Fourier transform of the 16 modes) are clearly not symmetric – an assumption usually made in the electron microscopy literature. The variation in their size corresponds to the change of magnification of the source as the condenser lenses are varied, just as we would expect.

4. DISCUSSION

The degree of partial coherence in propagating electron waves has previously been inferred by measuring the transverse mutual intensity function via holography or other interferometric methods, usually only in one dimension. In electron coherent diffractive imaging, another approach to handle the spatial coherence envelope is to model it as a Gaussian function of the illumination convergence angle that dampens the transfer of high spatial frequencies through the TEM [7]. However, this Gaussian model assumes the profile of the source is symmetric; the results shown here clearly show that in general it is not.

Here we use a comprehensive method to model partial coherence, which combines ptychography with modal decomposition to measure the components of a partially coherent electron wavefield explicitly, without any assumptions about the source shape or its spatial coherence properties. Uniquely amongst the various forms of coherent diffractive imaging, the redundancy and diversity of the ptychographic dataset is capable of supplying the greater amount of information required to retrieve multiple modes from diffraction data. The particular representation of the modes we have calculated, corresponding to the pure states associated with the eigenvalues of density matrix, is the most comprehensive and compact description of the partial coherence in a propagating matter wave. We emphasise that the

wavefield reconstructed from our data is not a representation of the actual electron source, but rather a diffraction limited image of source as it is seen through the aperture from the detector; the effective sources shown in Figure 5 are the convolution of the source function with the Fourier transform of the aperture function.

The formal theoretical analysis we derive explicitly includes the spatial coherence of the source, but can be expanded in a straightforward manner to include other effects. For example, our source has chromatic spread, the lenses employed may have high frequency instabilities and the specimen will have generated inelastic scattering events that may partly mask the coherent (elastic) scattering signal. We have also not accounted for the pixel size and fill factor of the detector. These effects will have to be the subject of further work. However, in our experience to date with visible light and X-ray ptychography, other sources of incoherence usually manifest themselves as modes that bear no relation to the size or shape of the illumination.

The partial coherence of the source is a key obstacle to accurate phase retrieval imaging. Ptychography combined with multi-mode reconstruction techniques offers a highly flexible way forward, since it considers the multiple sources of partial coherence within a single framework that does not require *a priori* coherence measurements or a parameterisation of the coherence envelope based on simplifying assumptions. The goal of our future research is therefore to optimise our experimental process, maximising information content in the diffraction data to realise high resolution electron phase imaging whilst retaining the accuracy and sensitivity already demonstrated at lower resolutions.

ACKNOWLEDGEMENTS

This work was funded by Phase Focus Ltd, EPSRC Grant No. EP/I022562/1 and the Department of Electronic and Electrical Engineering, University of Sheffield. JMR declares a shareholding in Phase Focus Ltd, which owns patents relating to some of the methods described in this paper.

REFERENCES

- [1] E. Wolf, J. Opt. Soc. Am. **72**, 343 (1982).
- [2] L. W. Whitehead, G. J. Williams, H. M. Quiney, D. J. Vine, R. A. Dilanian, S. Flewett and K. A. Nugent, Phys. Rev. Lett. **103**, 243902 (2009).
- [3] G. J. Williams, H. M. Quiney, A. G. Peele and K. A. Nugent, Phys. Rev. B. **75**, 104102 (2007).
- [4] J. N. Clark, X. Huang, R. J. Harder and I. K. Robinson, Phys. Rev. Lett. **112**, 113901 (2014).
- [5] P. Thibault and A. Menzel, Nature **494**, 68 (2013).
- [6] A. F. de Jong and D. V. Dyck, Ultramicroscopy **49**, 66 (1993).
- [7] J. Frank, OPTIK **38**, 519 (1973).
- [8] J. Frank, *Three-Dimensional Electron microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their native State* (Oxford, Oxford University Press, 2006).
- [9] E. Volkl, L. F. Allard and D. C. Joy, *Introduction to Electron Holography* (New Orleans, Springer, 1999).
- [10] M. Born and E. Wolf, *Principles of Optics* (Cambridge, Cambridge University Press, 1999).
- [11] G. F. Missirolii, G. Pozzi and U. Valldre, J. Phys. E: Sci. **14**, 649 (1981).
- [12] C. Maunders, C. Dwyer, P. C. Tiemeijer and J. Etheridge, Ultramicroscopy **111**, 1437 (2011).
- [13] F. M. Ross and W. M. Stobbs, Ultramicroscopy **36**, 331 (1991).
- [14] D. J. Batey, D. Claus and J. M. Rodenburg, Ultramicroscopy **138**, 13 (2014).
- [15] P. Kok and B. W. Lovett, *Introduction to Optical Quantum Information Processing* (Cambridge, Cambridge University Press, 2010).
- [16] A. J. D' Alfonso, A. J. Morgan, A. W. C. Yan, P. Wang, H. Sawada, A. I. Kirkland and L. J. Allen, Phys. Rev. B **89**, 064101 (2014).
- [17] F. Hue, J. M. Rodenburg, A. M. Maiden, F. Sweeney and P. A. Midgley, Phys. Rev. B **82**, 121415 (2010).

- [18] H. Yang, P. Ercius, P. Nellist and C. Ophus, Ultramicroscopy **171**, 117 (2016).
- [19] A. M. Maiden and J. M. Rodenburg, Ultramicroscopy, **109**, 1256 (2009).
- [20] P. Thibault, M. Dierolf, O. Bunk, A. Menzel and F. Pfeiffer, Ultramicroscopy **109**, 338 (2009).
- [21] P. Li, D. J. Batey, T. B. Edo, A. D. Parsons, C. R. and J. M. Rodenburg, Journal of Optics **18**, 054008 (2016).
- [22] M. J. Humphry, B. Kraus, A. C. Hurst, A. M. Maiden and J. M. Rodenburg, Nat. Commun. **3**, 1 (2012).
- [23] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States* (Cambridge, Cambridge University Press, 2006).
- [24] J. M. Rodenburg, Advances In Imaging and Electron Physics **150**, 87 (2005).
- [25] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus and J. M. Rodenburg, Ultramicroscopy **120**, 64 (2012).
- [26] P. Li, T. B. Edo, D. J. Batey, J. M. Rodenburg and A. M. Maiden, Opt. Express **24**, 9038 (2016).
- [27] A. M. Maiden, M. C. Sarahan, M. D. Stagg, S. M. Schramm and M. J. Humphry, Sci. Rep. **5**, 14690 (2015).
- [28] N. Burdet, X. Shi, D. Parks, J. N. Clark, X. Huang, S. D. KeVan and I. K. Robinson, Opt. Express **23**, 5452 (2015).

Figure Captions

Figure 1.a the idealised configuration. The upper part includes the source and the aperture which lies in the far field corresponding to the source; the lower part is the detector lies in the Fourier domain of the aperture.

Figure 1.b the setup of the experiment. The specimen was shifted by the piezo stage in x and y directions. The selected area aperture was on the first image plane with a slight defocus. Data 1 and data 2 were collected at spot size 2 and spot size 4, which can be taken as if the source were at different heights; the higher one is the spot 4. Data 3 and data 4 were collected at a different convergence angle by changing the strength of the condenser lens.

Figure 2. The flow chart when multiple probe modes engaged in ePIE. N modes were reconstructed after applying the calculation in the frame at every position for several hundreds of iterations. Then the N reconstructed probes (mixed states) will be orthogonalised to give the eigen modes.

Figure 3. One of the raw experimental diffraction patterns, which were collected in the near field (see text for more information.).

Figure 4. The error metric of data 4 with 500 iterations Gaussian convolution partial coherence correction and 800 iterations multiple modes partial coherence correction. The error at 500th iteration $e_G = 0.003833$; the final error $e_m = 0.002802$. After 500 iterations Gaussian corrector then from the 501st iteration modes corrector began without updating the object until the 901st iteration.

Figure 5. The modes and the effective sources of the 4 sets of data. In any set of modes, each of them is orthogonal to each other. The numbers above each mode show the contribution percentage of each mode. The effective sources are the convolution of the source with the aperture.

Figure 6. The plots of the 16 eigenvalues of the 4 data sets. The more coherent the source is, the more contribution comes from the first mode, and less come from the other modes.

Figure 7. The specimen reconstructions. (a) and (b) show the phase and modulus of the specimen when 16 modes were engaged. (c) and (d) show the phase and modulus reconstructed with single mode. The main images are the magnification of the squared area by the red dashed line in the whole field of view inserted in the top corner.

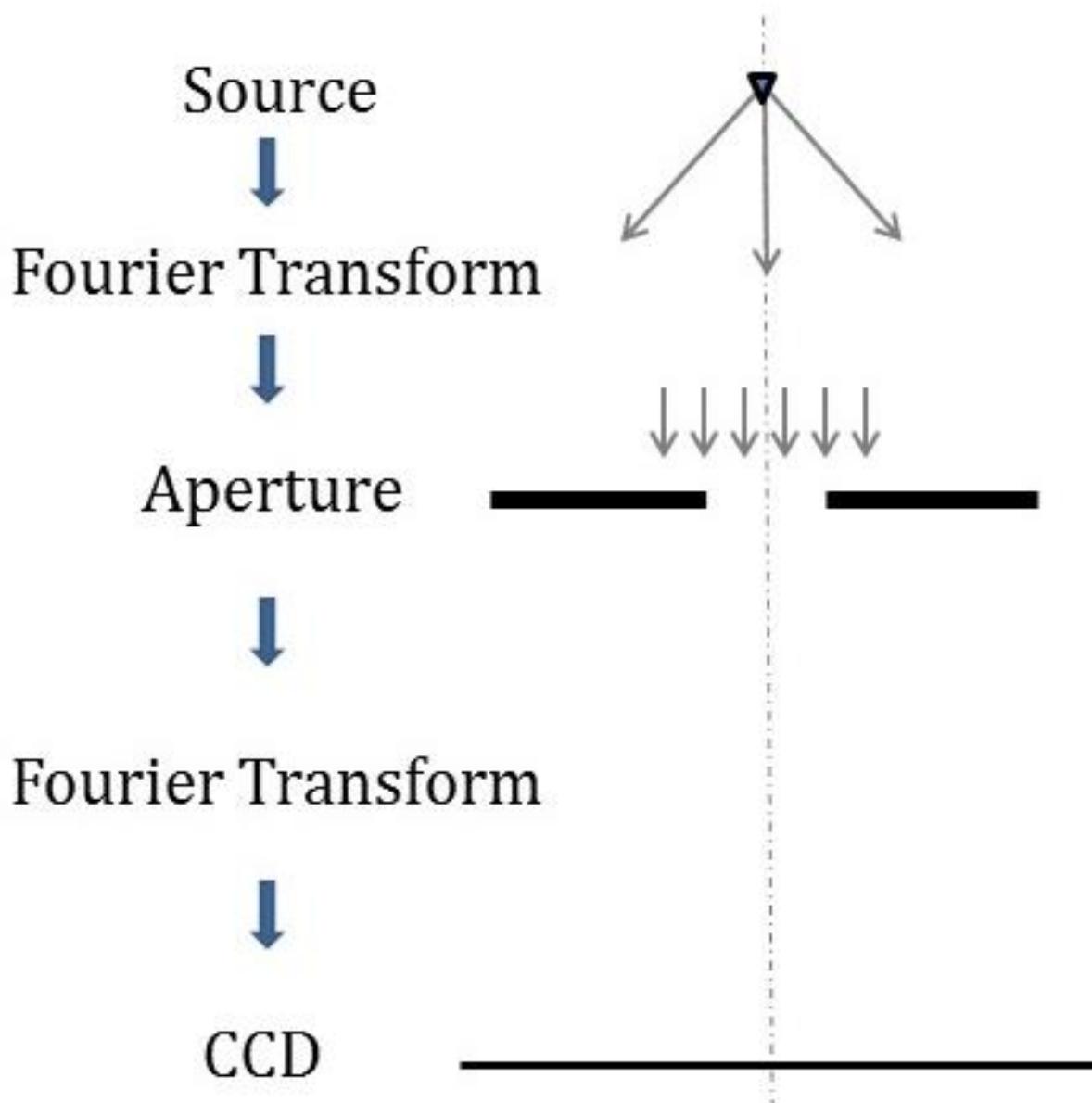


Figure 1.a

Electron source (spot 4)

Electron source (spot 2)

Condenser lens

Specimen

Objective lens

SA aperture at image
plane

Angular spectrum
propagator

CCD

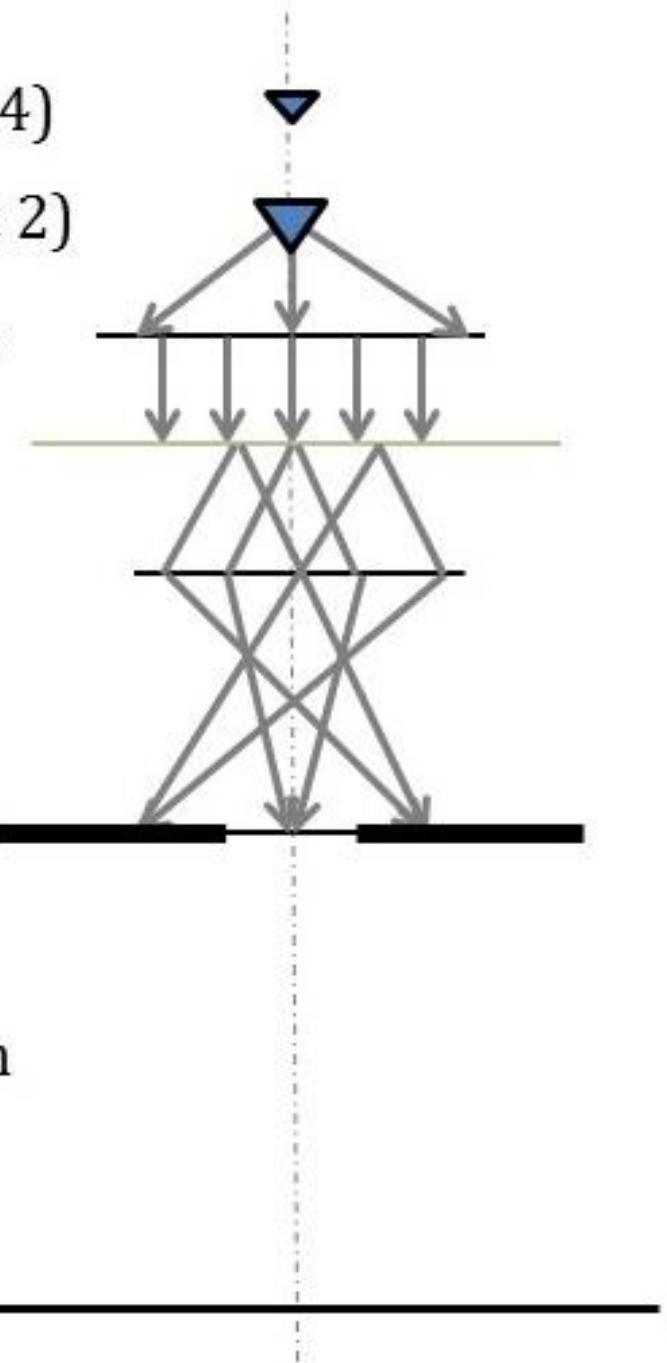


Figure 1.b

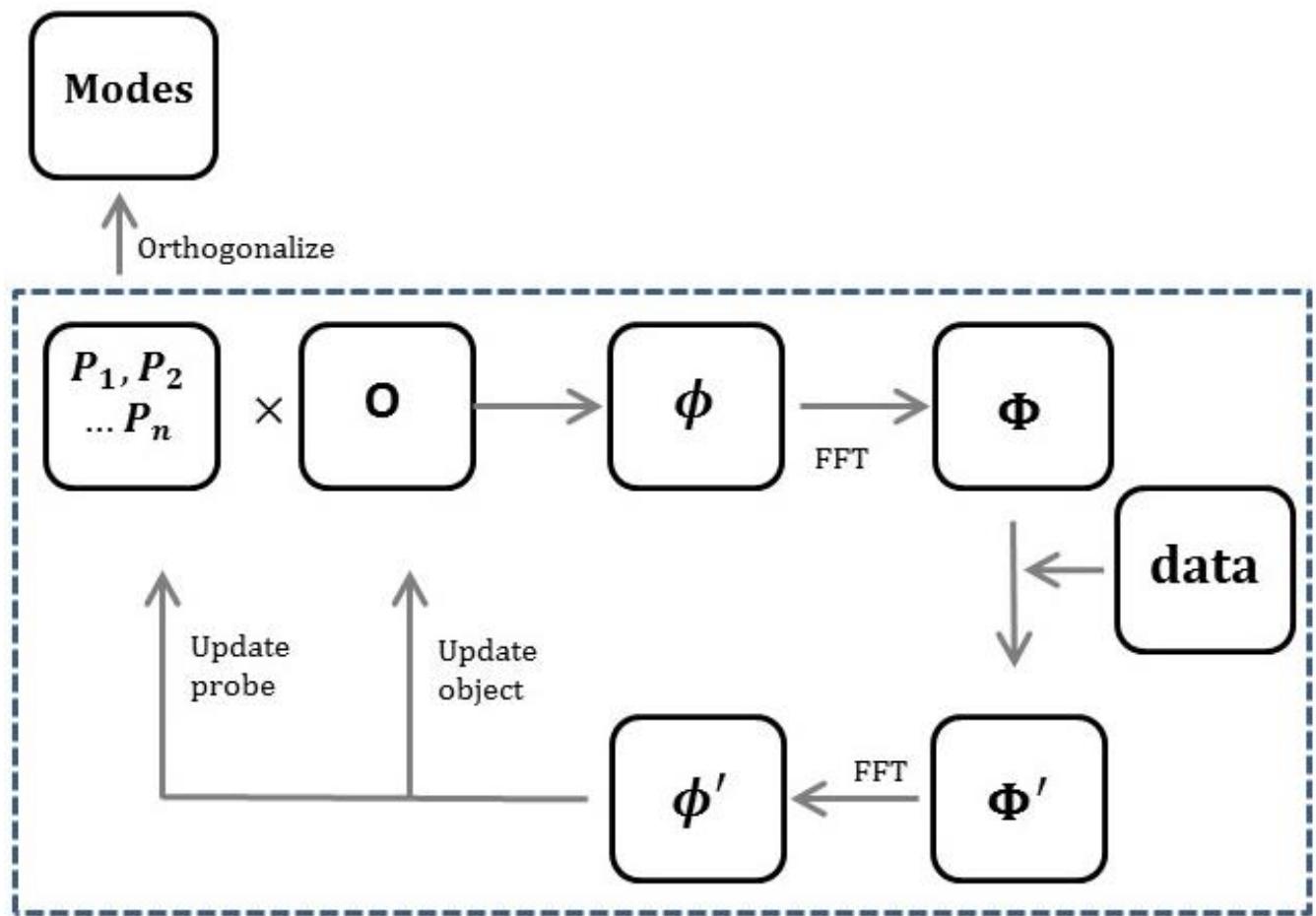


Figure 2

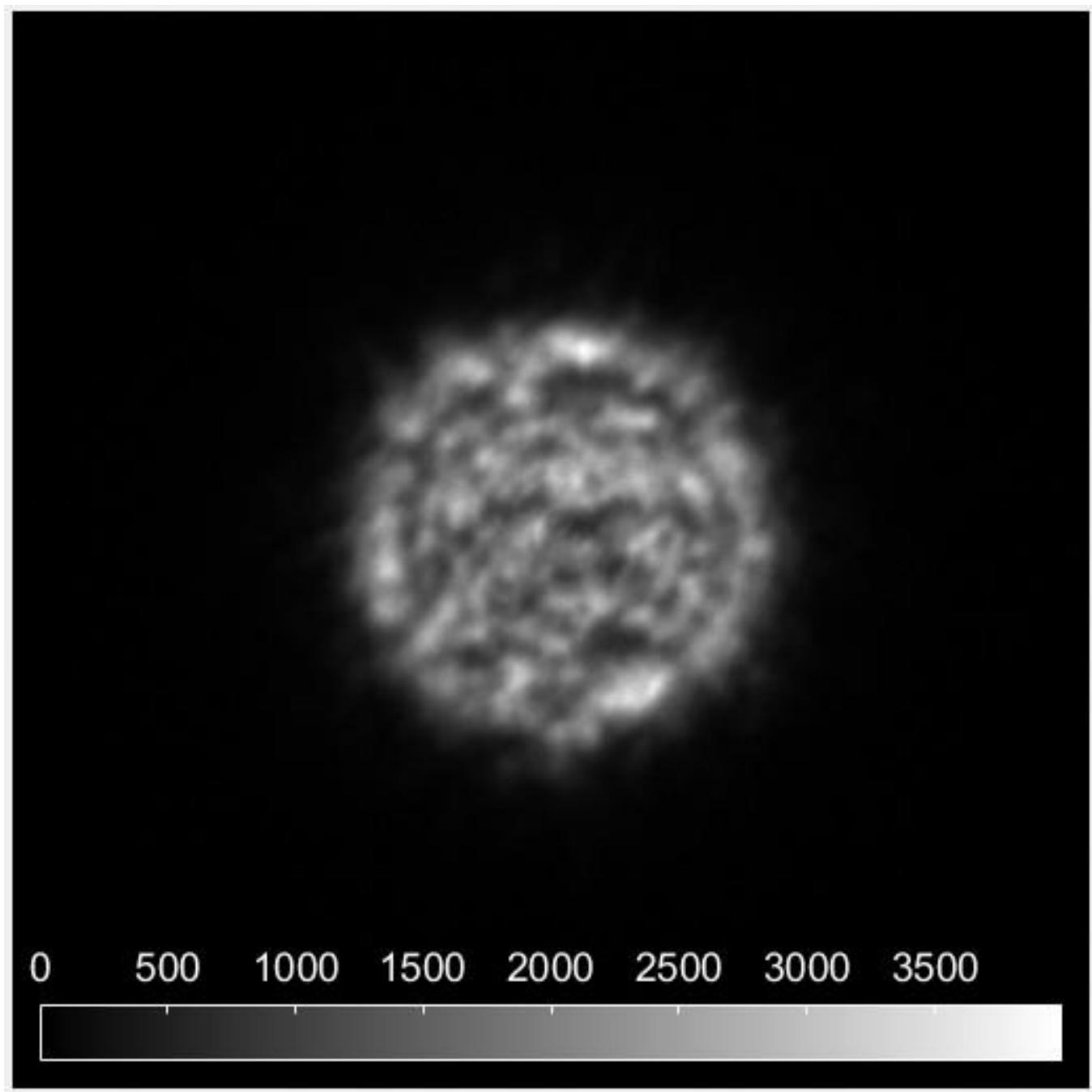


Figure 3

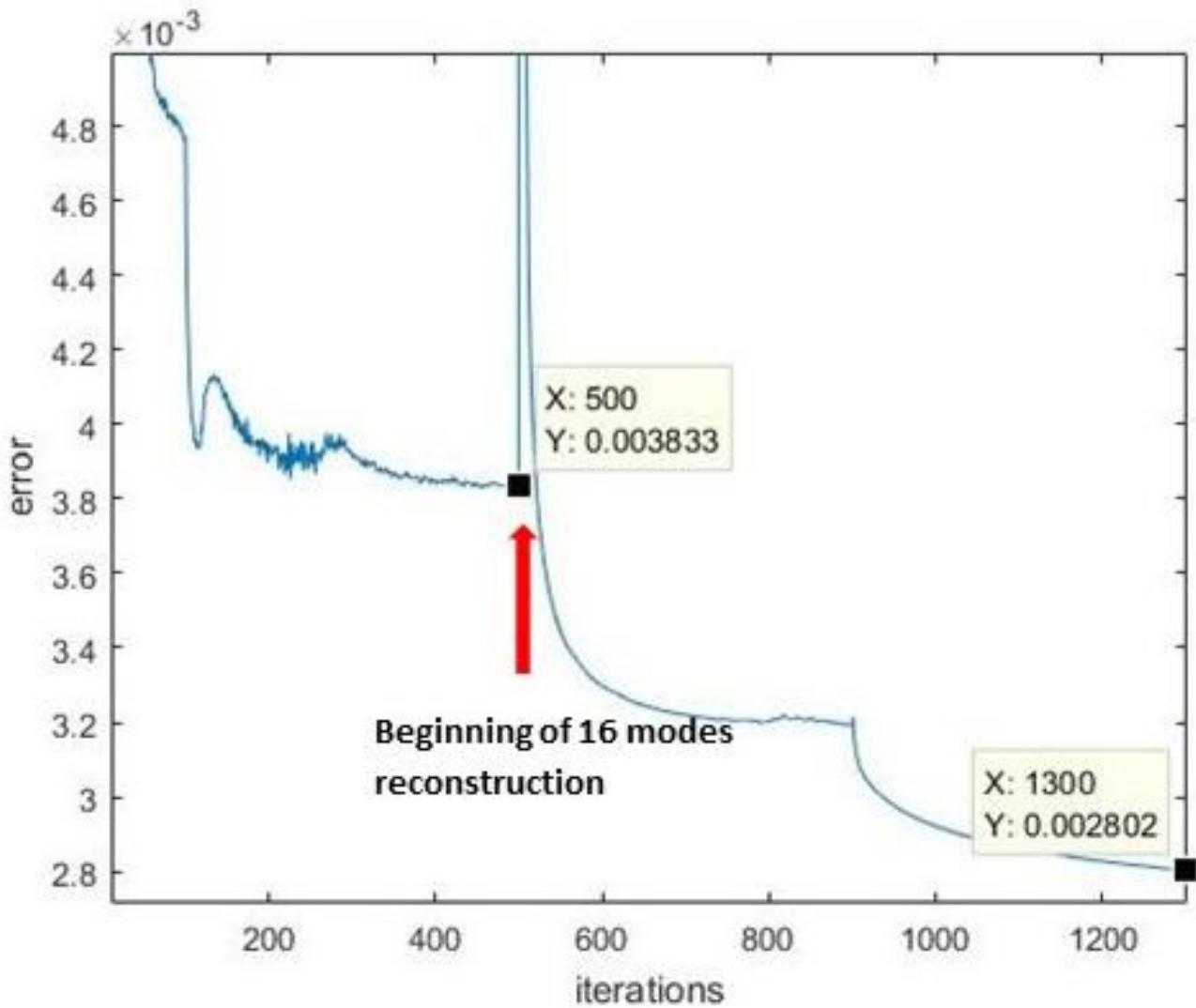


Figure 4

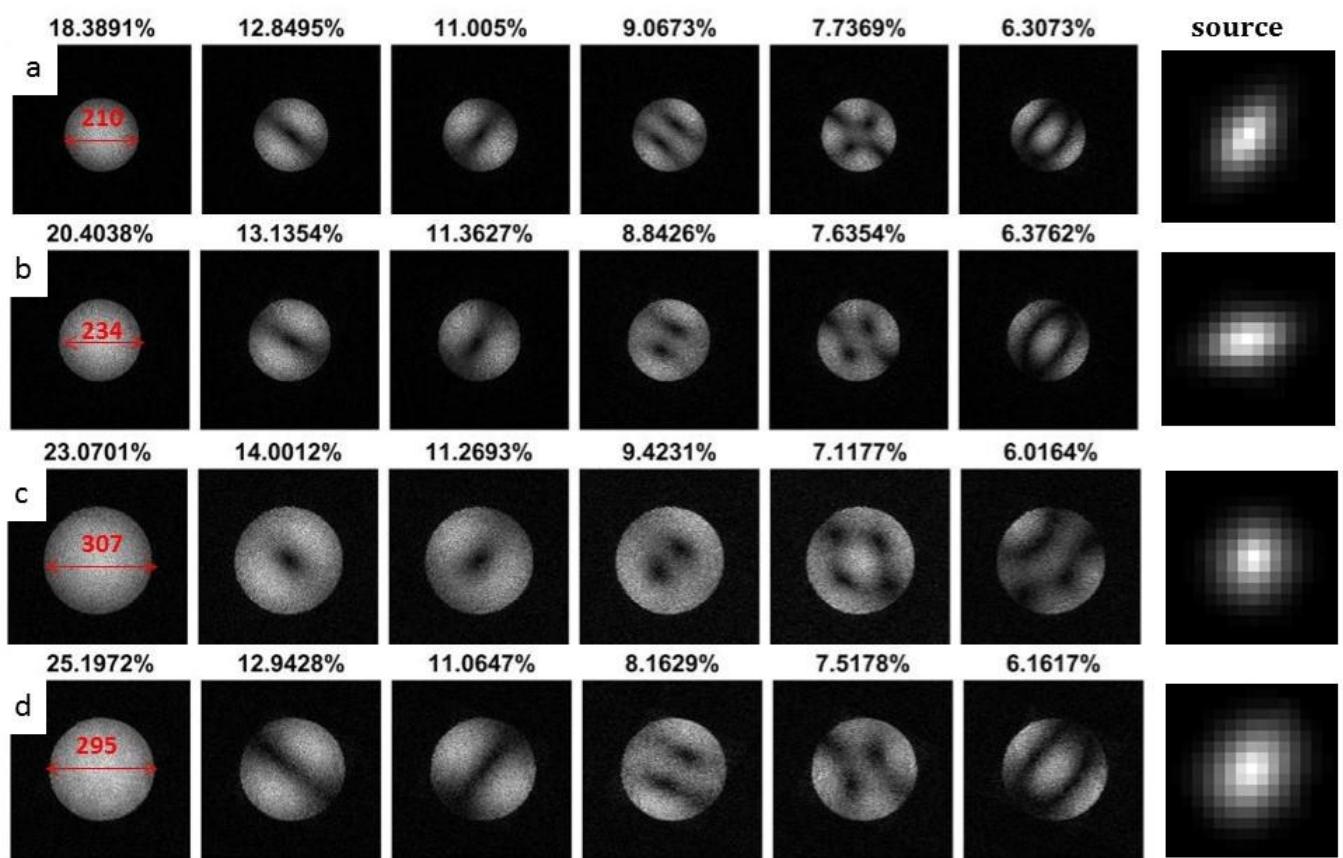


Figure 5

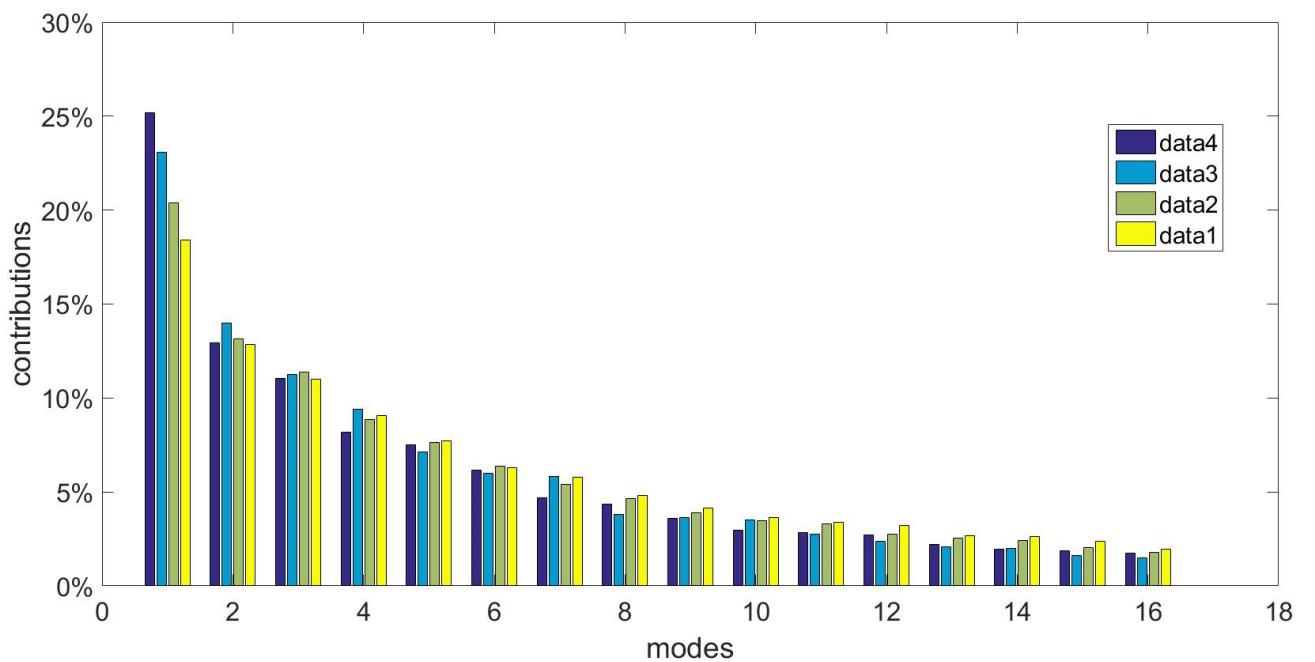


Figure 6

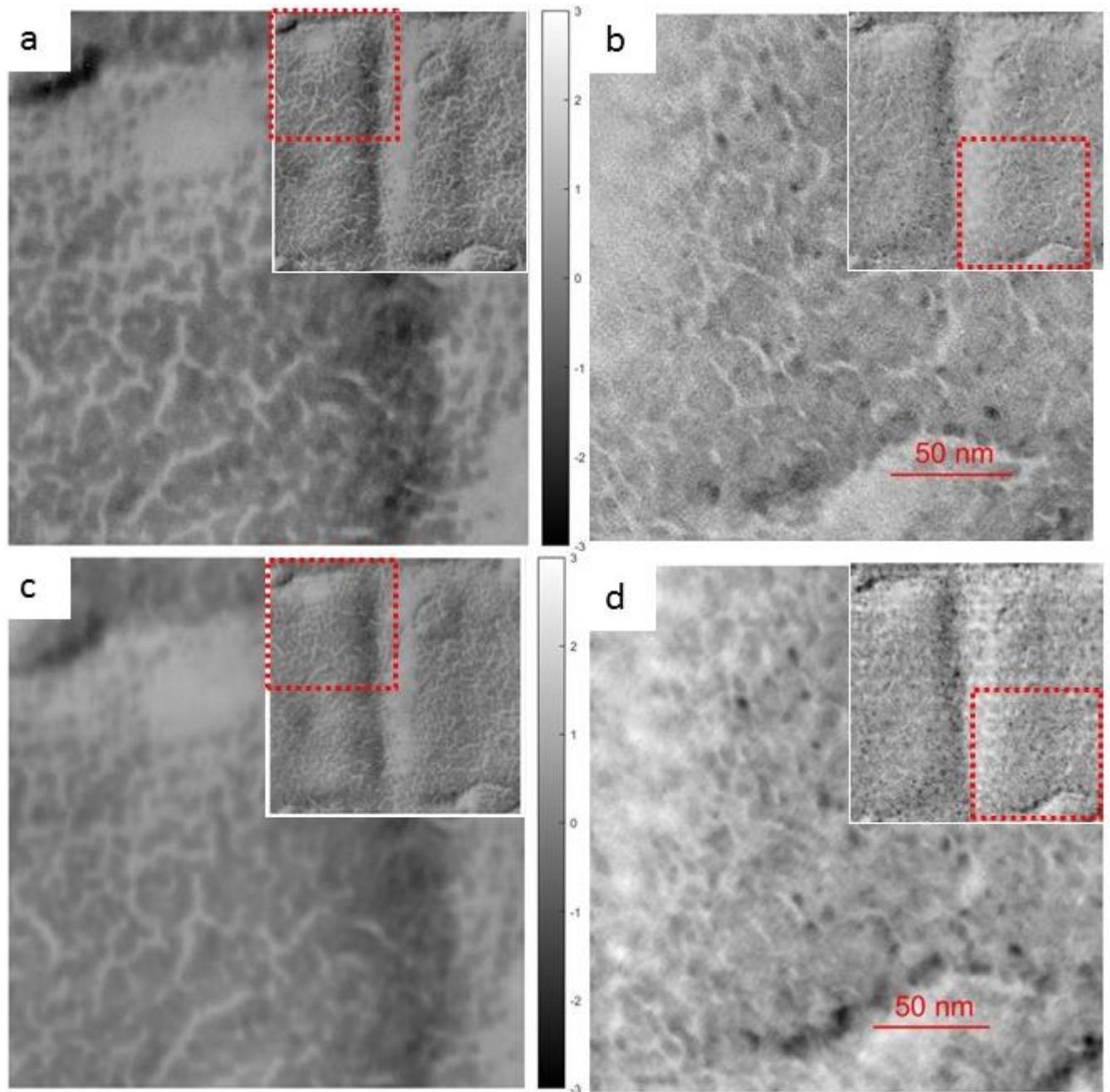


Figure 7

3.2 Collecting SAD ptychographic data

3.2.1 Specimen

The specimen that we have used in the experiment in this chapter is a 500nm Diffraction Grating Replica with Latex Spheres. Latex sphere is about 260nm in diameter. The area we have scanned constituted thick gold nanoparticles that are stacked together. We estimate the thickness is about several tens of nanometres. For the purpose of decomposing the density matrix of the source via ptychography, a specimen that has more structure is easier for ptychographic reconstruction.

3.2.2 Calibration

3.2.2.1 Camera length calibration

An accurate measurement of camera length is essential for the ptychographic reconstruction. The camera length readout from the microscope usually is not accurate; it needs to be calibrated before the reconstruction. The theory to calibrate the camera length involves Bragg's law.

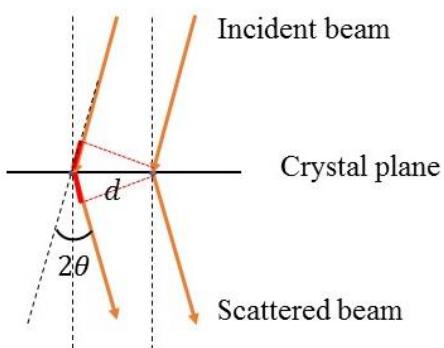


Figure 3.2.1. Bragg's law. A Parallel illumination beam is scattered by the crystal lattices with an angle of 2θ . The lattice spacing is d .

As shown in Figure 3.2.1, when a parallel beam illuminates a crystal specimen with lattice spacing of d , the path difference between the scattered beam is $2dsin\theta$, where 2θ is the scattered angle. When the path difference equals the product of an integer number with the wavelength λ , there will be a peak intensity in the diffraction pattern. This is Bragg's law.

$$2d \sin \theta = n\lambda. \quad (3.2.1)$$

When the scattering angle 2θ is small, we can take the approximation that

$$\sin 2\theta = \tan 2\theta = 2\theta = \frac{R}{L}, \quad (3.2.2)$$

where L is the camera length, and R is the radius of the diffraction spots associated with a particular lattice spacing. Thus, we obtain the equation to calibrate the camera length as

$$\lambda L = dR. \quad (3.2.3)$$

The specimen that we used to calibrate the camera length was a Thallous Chloride crystal, whose lattice spacings are shown in Table 3.2.1.

N	1	2	3	4	5	6	8	9	10
Lattice spacing (nm)	0.384	0.272	0.222	0.192	0.172	0.157	0.136	0.128	0.121

Table 3.2.1. the lattice spacings of the Thallous chloride crystal.

Figure 3.2.2 shows one example of the SAD patterns of the Thallous chloride crystal specimen. The calibration result of the camera length of the JEOL R005 in the SAD mode with a range of 15cm – 200cm is shown in Table 3.2.2.

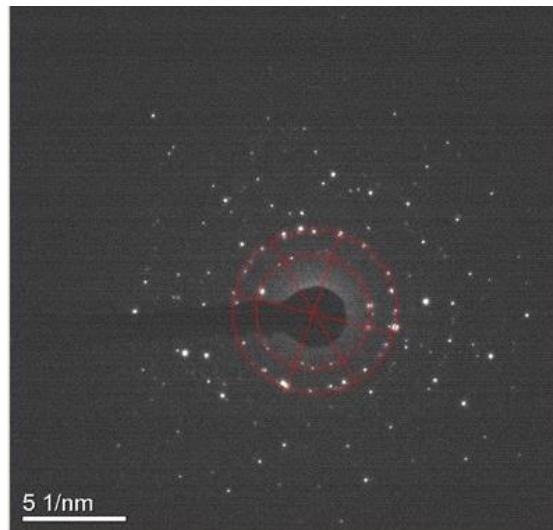


Figure 3.2.2 shows one example of the SAD patterns of the Thallous chloride sample. The red circles are used to work out the diameters of the diffraction rings.

Read/cm	15	20	30	40	50	60	80	100	120	150	200
Correct/cm	23.6	31	44.9	60.4	74.9	95.4	115.9	144.8	179.9	216.4	275.9
Error/cm	8.6	11	14.9	20.4	24.9	35.4	35.9	44.8	59.9	66.4	75.9

Table 3.2.2 the reading out camera length and the calibrated camera length of JEOL R005 in SAD mode.

3.2.2.2 Piezo stage movement calibration

In the SAD ptychographic experiment, we shifted the specimen with a piezo stage to perform a raster scan; the step size was around 15 – 20 nm. In order to make sure that the movement of the stage was accurate and repeatable, we calibrated and corrected the shifting of the piezo stage in the x and y directions before the experiment. The specimen that we used to calibrate the stage shift was a 500nm diffraction grating replica with latex spheres. The method to work out the true relative movement between each position was cross correlation. The procedures of the calibration were:

- 1) Generate a position map with a step size of 10% - 20% of the diameter of the probe; assign this position map to the DigitalMicrograph software, to shift the specimen with the piezo stage and collect a set of TEM bright field images;
- 2) Use a cross correlation method to work out the relative shifting distance of this set of bright field images; name it as positions map 1;
- 3) Calibrate the shift scale factor of the piezo stage, which is the scale factor between the assigned step size and the actual step size; we find out the scale factors by $\frac{\text{assigned step}}{\text{actual step}}$ in both the x and y directions;
- 4) Check the repeatability; repeat step 1), to collect another set of bright field images with the same assigned positions map; cross correlate this set of bright field images and work out the relative shifts; – name it as positions map 2; check whether positions map 2 is the same as positions map1; if they are not the same, we need to calibrate the repeatability;
- 5) Calibrate the repeatability; the reason that the stage movement is not repeatable may because of hysteresis and inertia; the method we applied to improve the shifting repeatability is to move the stage in both directions some distance before the scanning, and to shift the stage backwards some distance when finishing one acquisition before moving to the next position.

After the calibration, the piezo stage is able to achieve an accurate and repeatable scanning step size of 10 – 20nm. Figure 3.2.3 shows the positions from the cross correlation of two sets of bright field images; the step sizes in both directions are the same, except there is a slight rotation between the two sets of positions, which can be corrected by ‘jiggle’ (Maiden et al., 2012) during the reconstruction.

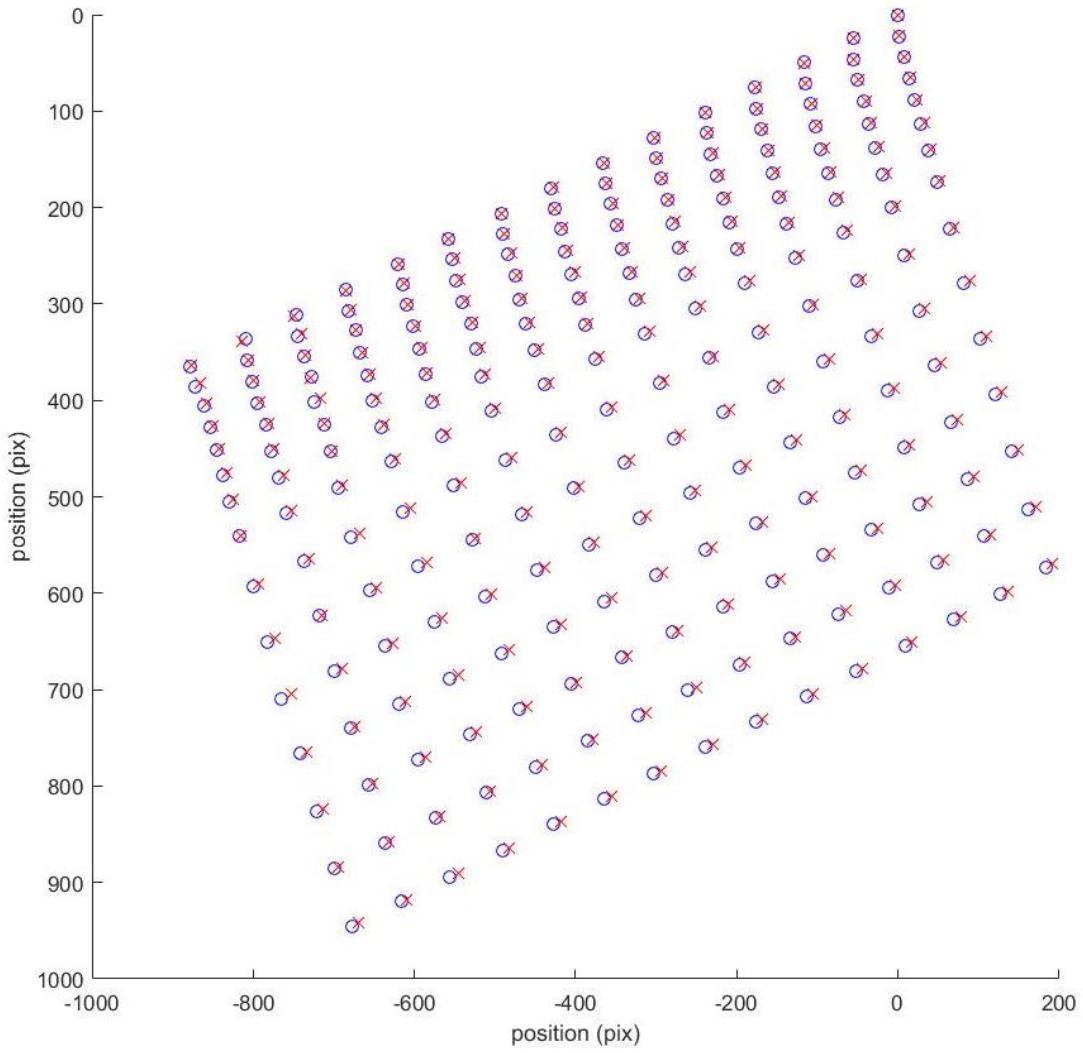


Figure 3.2.3 shows two sets of scanning positions after the calibration. They were worked out from the cross correlation of two sets of bright field images collected with the same assigned positions map. The first scanning position was at the top right corner. The positions are plotted in the unit of pixels; the pixel dimension was 0.35nm. We can see that the positions in the first 8 rows, the shifting has hysteresis in y direction; the step size in y direction was about 9nm, which became about 20nm since the 9th row. The shifting in the x direction was stable; the step size in x direction was about 12nm. The two sets of positions repeated very well, except for a slightly relative rotation.

3.2.3 Alignment

Good optics alignment both in image mode and in diffraction mode is necessary to obtain diffraction patterns for the ptychographic reconstruction. In SAD ptychography, the alignment in the image mode determines the first image, which is the object function that will be reconstructed. Even though ptychography can reconstruct both the object function and the illumination function, the aberrations and distortions in the diffraction lens will relax the Fourier constraints in the reconstruction, which may cause the reconstruction to fail. Here we give the alignment procedures and standards that we have used for ptychographic experiments in the SAD mode.

The alignment procedures of TEM imaging mode and SAD mode see Appendix 1.

3.3 Theoretical evaluation on modes reconstruction

The paper in Chapter 3.1 has derived the formal quantum mechanical description of the modal decomposition theory and has reported the experimental measurements of the modal decomposition of the partially coherent electron source. This supplementary section reports some theoretical demonstration and investigation on the modal decomposition theory. It contains the work that 1) we demonstrate theoretically that a partially coherent source can be described in infinite coordinates; however, there exists a unique coordinate system, where the partially coherent source can be described with the minimum number of states; we call this coordinate system the eigen coordinate, and the states expressed in the eigen coordinate as the eigenstates; 2) we show that the eigenstates of a partially coherent source are characterised by the physical source size, the source shape and the size of the condenser aperture; 3) each eigenstate can be propagated separately and freely as a pure state in the imaging system. When adding up the intensities of all the eigenstates at the source plane, we achieve the effective source, which is the convolution of the physical source with the Fourier transform of the condenser aperture. 4) we show that the ptychographic reconstruction of the eigenstates with the modal decomposition method is affected by the number of the diffraction patterns (the reconstruction area), the structure of the specimen, the probe scanning step size and the propagation distance.

3.3.1 The properties of the eigenstates

We take Figure 3.3.1 as the configuration for all the calculations in this section. In the first calculation, we simulated a source consisting of 9 delta functions (Figure 3.3.2.a); each of the point sources was propagated separately through an aperture; then they were added up together to form the mixed states. Figure 3.3.3 shows three probability expressions of the mixed states that exiting behind the aperture. if we orthogonalize any of the mixed-state sets

in Figure 3.3.3, we will get the unique eigenstates of the partially coherent source, which is shown in Figure 3.3.4.

Key conclusion: Any arbitrary linear superposition of these eigenstates is a possible expression of the mixed-state illumination.

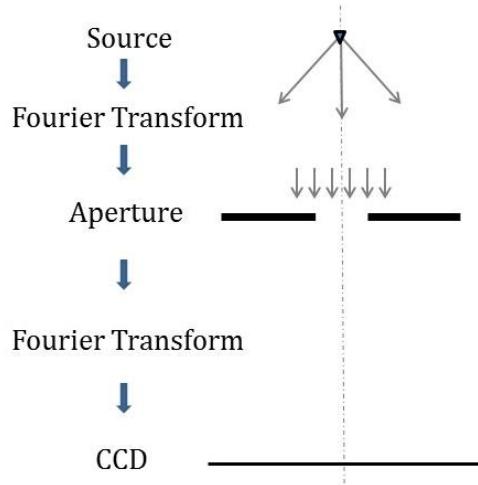


Figure 3.3.1. the setup used for all the simulations in this section. the specimen is placed at the far field (close behind the aperture) relative to the source. An aperture is placed to localize the illumination on the specimen. The detector is placed at the far field relative to the specimen.

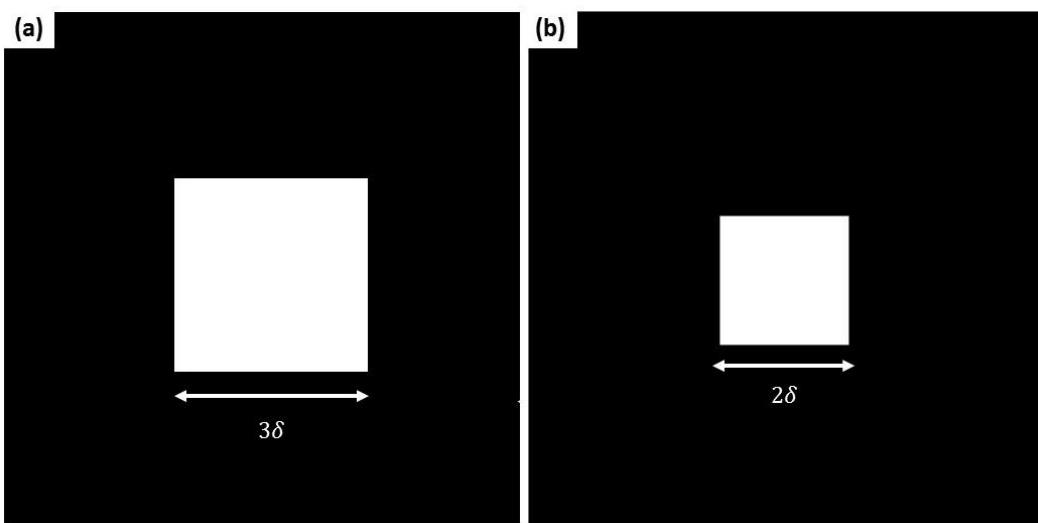


Figure 3.3.2. The simulated transverse partially coherent sources with limited width; (a) the width is 3 pixels in both dimensions; (b) the width is 2 pixels in both dimensions.

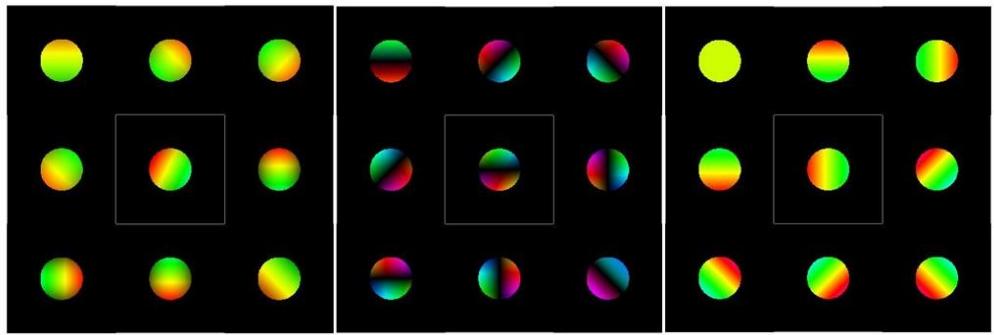


Figure 3.3.3. The 3 probability distributions of the source **Figure 3.3.2.a.** Each of these states is a mixed state. The addition of the 9 mixed states in each set is the partially coherent source that goes through the aperture.

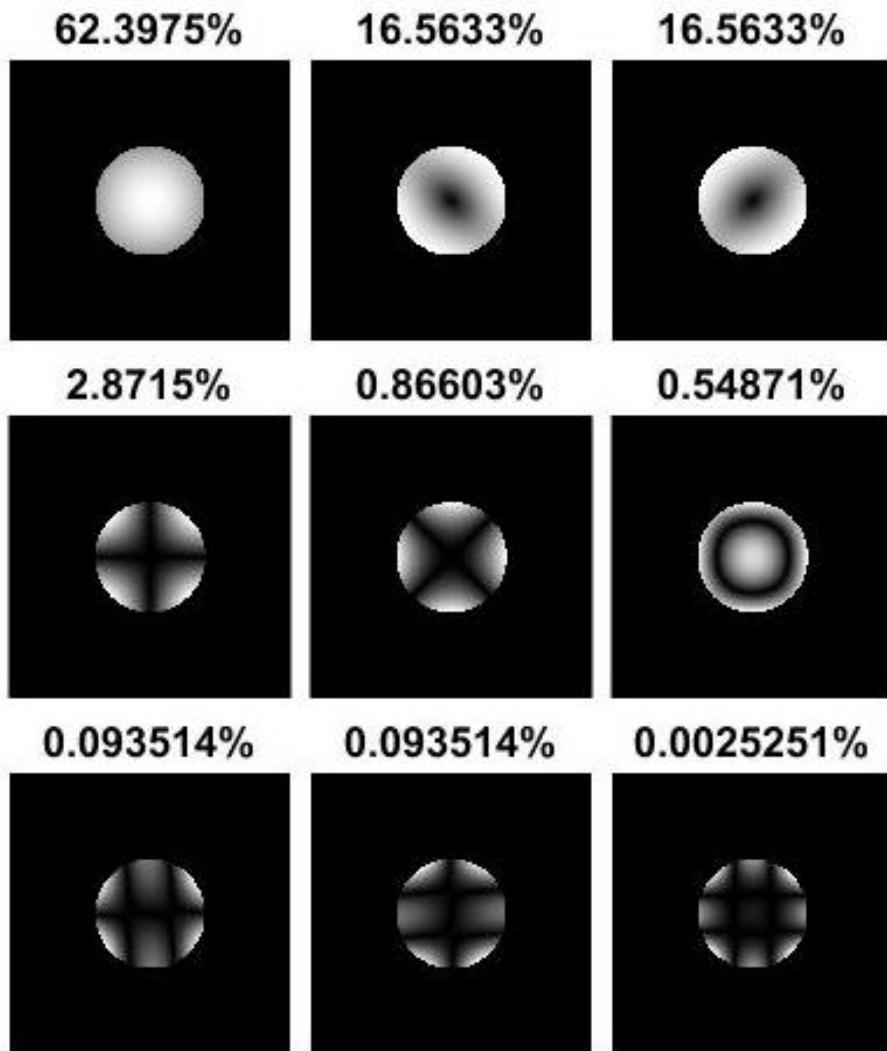


Figure 3.3.4. the eigenstates of the source shown in **Figure 3.3.2.a**, which is obtained by orthogonalizing any set of the mixed state that is shown in **Figure 3.3.3.**

When we make the physical size of the source smaller, so it consists of 4 point sources, as shown in Figure 3.3.2.b, the eigenstates at the aperture plane are as shown in Figure 3.3.5. We can see that the contribution of the first eigenstate, which is 62.4% of source in Figure 3.3.2.a, now has increased to 82.87%; the shapes of the eigenstates of these two sources with different physical size is consistent.

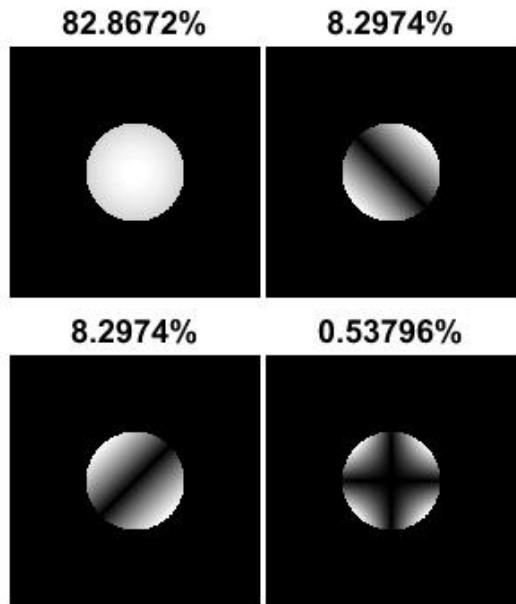


Figure 3.3.5. The eigenstates of the source shown in Figure 3.3.2.b. when the source has a smaller spatial size, the contribution of the first mode is larger: 62.3975% for source consisting of 9 point sources, 82.8672% for source consisting of 4 point sources.

Figure 3.3.6 shows some partially coherent sources of different shapes and their corresponding eigenstates at the aperture plane.

Key note: We can see that, as the shape of the source changes, the shapes of eigenstates will change as well.

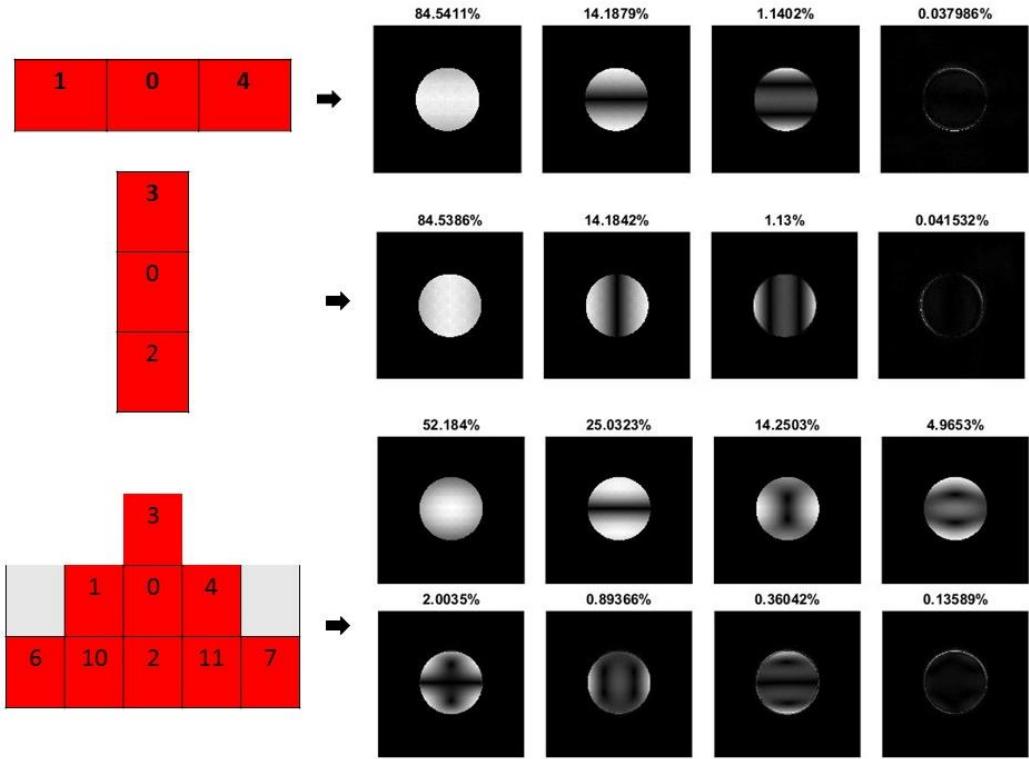


Figure 3.3.6 shows that the shapes of the eigenstates are related with the shapes of the sources. When the source spread is lateral, the shapes of the modes are also lateral; when the source spread is longitudinal, the shapes of the modes are also longitudinal.

Figure 3.3.7 shows the eigenstates of source that is shown in Figure 3.3.2.a, when the diameter of the aperture is twice as large. We can see that the shapes of the modes are consistent with Figure 3.3.4, but the contribution of the first mode reduces to as low as 23%. So, the condenser aperture, which limits the electrons going through the imaging system, may affect the coherence.

Key note: A smaller condenser aperture with a proper size may improve the coherence of the illumination. In the experiment, we need a balance between the coherence of the illumination and the counts when we choose the size of the condenser aperture.

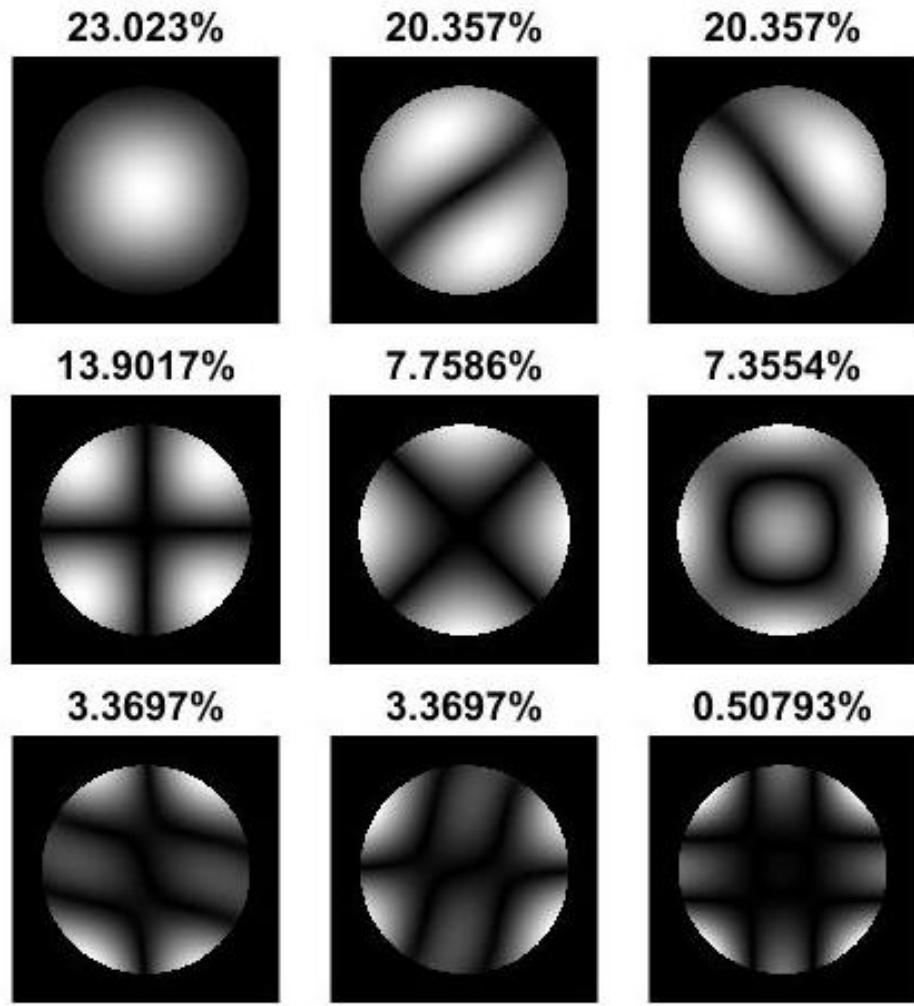


Figure 3.3.7 shows the eigenstates of the source in Figure 3.3.2.a when replacing the aperture with a larger one. As the aperture diameter goes twice larger, the coherence goes worse. The contribution of the first mode reduces to 23.023%.

Figure 3.3.8 shows a partially coherent source consisting of several sparse point sources and its eigenstates at the aperture plane. When propagating each of these eigenstates separately to the source plane and adding up the intensities, we get the effective source, as shown in 3.3.8.c. Figure 3.3.8.d is the convolution of the source with the aperture, which is the same with the effective source.

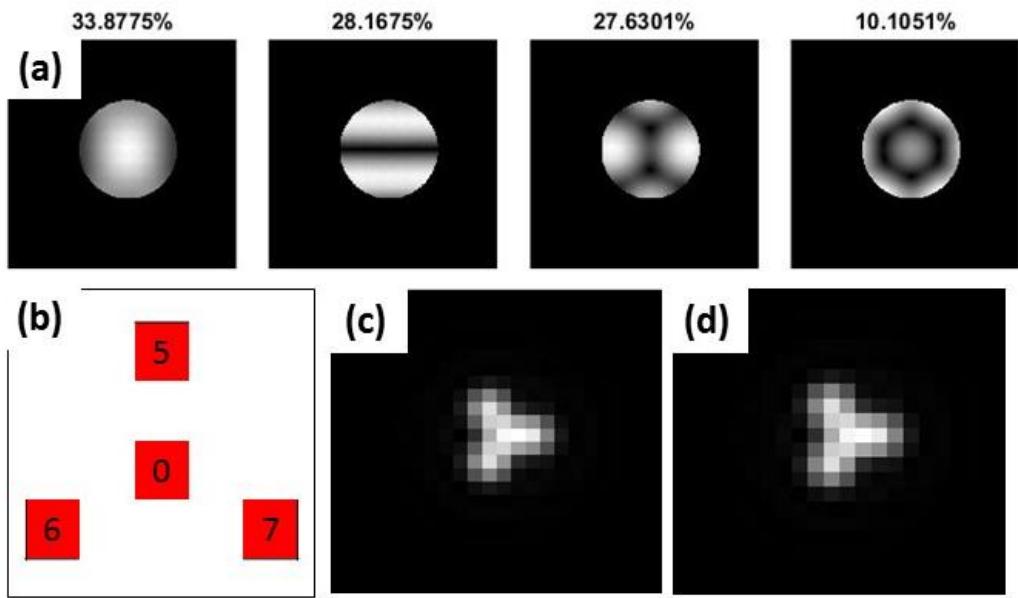


Figure 3.3.8 shows that the superposition of the intensities of each eigenstate when propagated to the source plane, is the convolution of the source with the aperture. (a) the eigenstates of source shown in (b). (c) the superposition of the intensities of all the eigenstates propagated to the source plane. (d) the convolution of the source with the aperture.

In conclusion, the properties of the eigenstates are that

- 1) the eigenstates are the unique and the minimum-number expression of a partially coherent source; any arbitrary linear superposition of the eigenstates can be an expression of the source.
- 2) The contribution of the first eigenstate shows the coherence of the source: the more contribution the first eigenstate takes, the better coherence the illumination has.
- 3) The shapes of the eigenstates are related to the shapes of the sources.
- 4) The size of the aperture may affect the coherence of the illumination.
- 5) The superposition of the intensities of all the eigenstates when propagated to the source plane, equals the convolution of the source with the aperture.

3.3.2 Factors affecting the Reconstruction of the eigenstates

In this section, we investigate the factors in a ptychographic experiment which may affect the reconstruction quality of the eigenstates. These factors are 1) the probe scanning step size (real space overlapping of the illumination position), 2) the structure of the specimen, 3) the number of scanning positions (reconstruction area), and 4) the propagation distance (near field diffraction pattern or far field diffraction pattern).

To investigate computationally the influence of the probe scanning step size on the modal decomposition reconstruction, we modelled two sets of diffraction patterns in the far field, which have 10*10 diffraction patterns with a raster scan for each set of data; dataset 1 was collected with a step size of 50% of the diameter of the illumination, while dataset 2 was collected with a step size of 70% of the diameter of the illumination. Figure 3.3.9.a and 3.3.9.b show the reconstructed eigenstates of the two datasets. Dataset 2 with 70% step size failed the modes reconstruction.

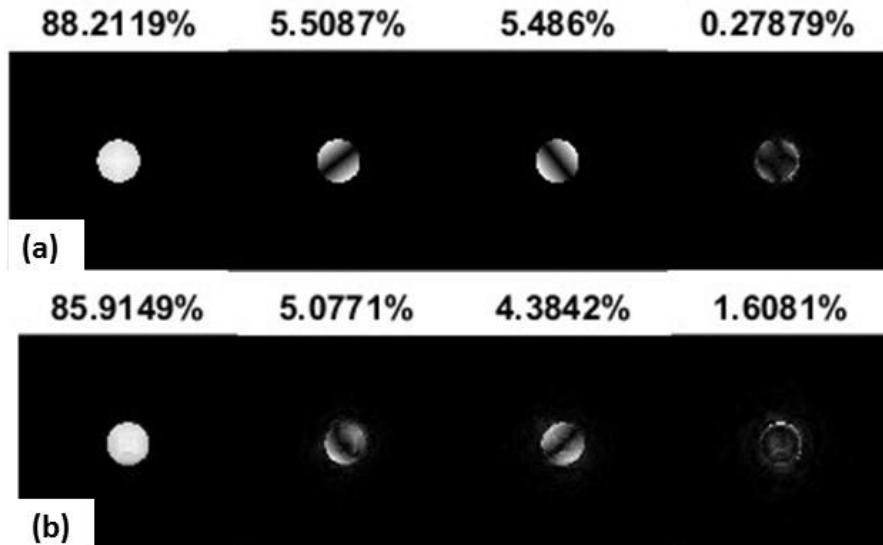


Figure 3.3.9 shows the reconstruction of the modes with data (a) 50% overlap, and (b) 30% overlap.

When the overlap in the real space went as low as 30%, the algorithm failed to retrieve the modes.

To investigate the influence of the specimen structure on the modal decomposition reconstruction, we collected diffraction patterns of two different specimens; specimen 1 was a pure phase weak scattering specimen (Figure 3.3.10.a); specimen 2 was a strong scattering complex image of Lena (Figure 3.3.10.b). 10*10 diffraction patterns with a 70% real space overlapping raster scan in the far field were collected for each of the datasets. The reconstructions of the modes are shown in Figure 3.3.11.a and 3.3.11.b. We can see that specimen with stronger scattering property and more diverse structures is easier for the modal decomposition reconstruction.

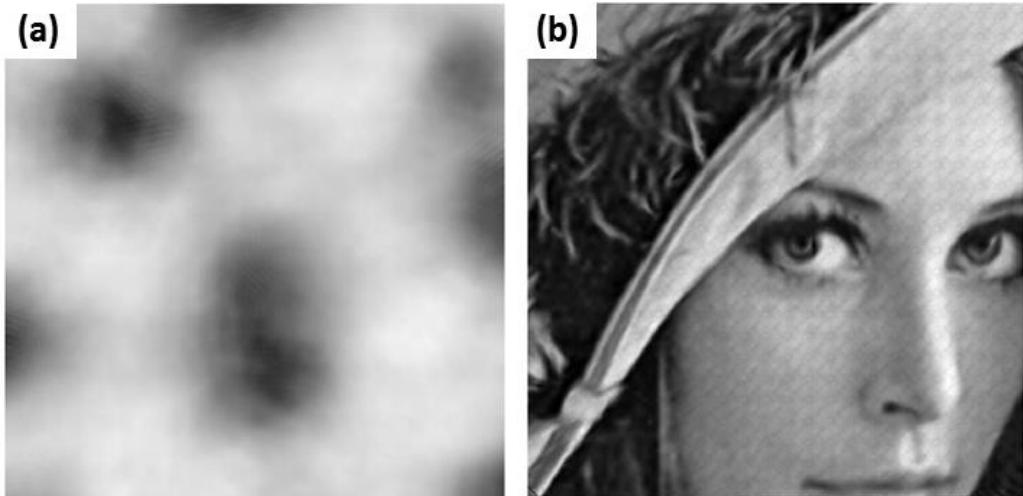


Figure 3.3.10 shows the specimen that have used to test the modes reconstruction: (a) weak pure phase object, and (b) object with both strong modulus and phase.

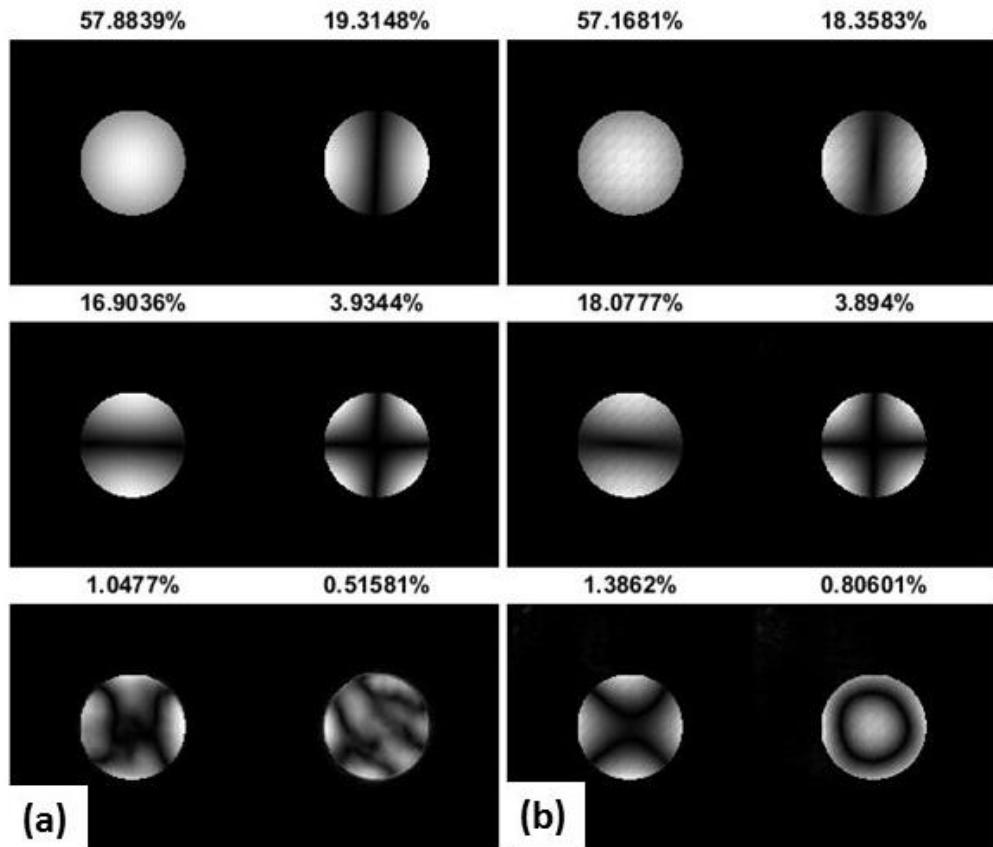


Figure 3.3.11 shows the modes reconstruction with the weak phase object (a), and with the strong scattering object (b). The modes with the strong structure object were reconstructed better.

To investigate the influence of diffraction patterns numbers, or the influence of the reconstructed area, on the modal decomposition reconstruction, we calculated two sets of diffraction patterns with the weak phase specimen shown in Figure 3.3.10.a in the near field; dataset 1 collected 10×10 diffraction patterns; dataset 2 collected 20×20 diffraction patterns. The reconstructions of the eigenstates from the two sets of data are shown in Figure 3.3.12.a and 3.3.12.b. We can see that when the specimen has less structure and weak scattering property, a larger reconstructed area (more scanning positions) is helpful for the modes reconstruction.

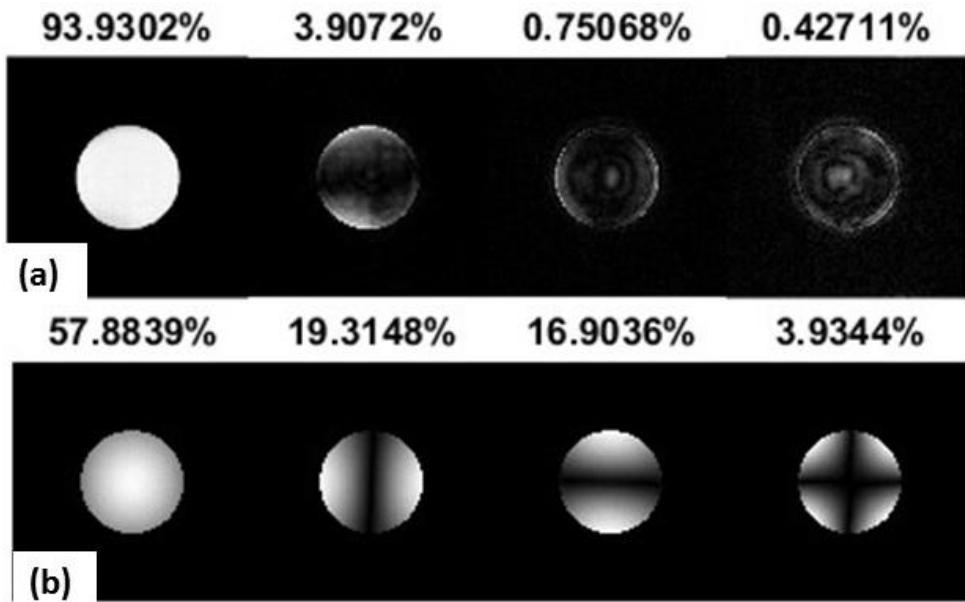


Figure 3.3.12 shows the reconstruction with 10×10 scanning positions (a) and 20×20 scanning positions (b) at the same step size. The algorithm failed to reconstruct the modes with 10×10 scanning positions. Larger scanning area makes it easier for the modes reconstruction.

To investigated the influence of the propagation distance on the modal decomposition reconstruction, we collected 2 sets of 10×10 of diffraction patterns; one was collected in the near field, and the other was collected in the far field. The reconstructions of the modes with the two sets of data are shown in Figure 3.3.13. We can see that the data collected in the far field gives a better mode reconstruction.

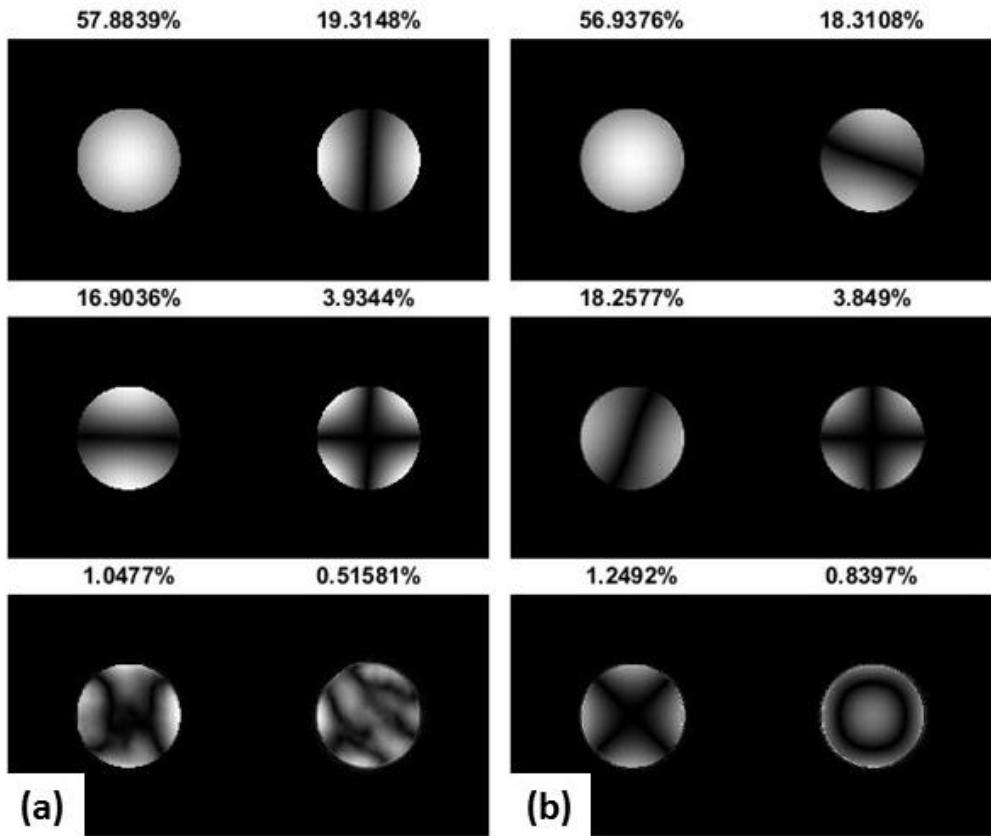


Figure 3.3.14 shows the modes reconstruction with the diffraction patterns collected in the near field (a) and in the far field (b). Far field diffraction patterns make the modes reconstruction easier.

In conclusion, a proper overlap in the real space, a highly structured specimen, larger scanning area and far field diffraction patterns enable the algorithm to reconstruct the complete decomposition of the partially coherent source. This is because ptychography needs a certain measurements (knowns) to reconstruct the object and the illumination (unknowns). A highly structured sample enables more convolution in reciprocal space.

3.4 Supplementary experimental Results

We collected 8 sets of data in total in the SAD ptychographic experiment at different coherence conditions. 4 sets of them were collected by changing the spot size, while keeping other parameters the same; the other 4 sets of data were collected by changing the convergence angle of the condenser lens C2, while keeping the other parameters the same. In the paper (chapter 3.1), we only presented the first six modes of 4 sets data (2 sets by changing the spot size and 2 sets by changing the convergence angle), while the algorithm engaged 16 modes during the reconstruction. Here as a supplementary, we present the complete 16 reconstructed modes of the 8 sets of data. Reconstructions with the modal decomposition method (Thibault et al., 2014; Li et al., 2016) and the Gaussian blind deconvolution method (Maiden et al., 2015; Burdet et al., 2015) are compared.

3.4.1 Data description

All the data were collected with the following experimental conditions

Machine: JEOL R005 HRTEM

Camera: Gatan Orius SC200D

Condenser Aperture: CL2 2 (diameter 100um)

Selected area aperture: SA 4 (diameter 10um)

Magnification in Image mode: 30K (pixel pitch on CCD 0.35nm)

Nominal camera length: 2m (before calibration)

Scanning positions: 15*15 raster scan

Data Series 1: 4 sets of data collected by changing the convergence angle of the condenser lens (C2) at spot size 3.

Data set	Dp1	Dp2	Dp3	Dp4
Exposure (s)	1.5	1.0	2.0	3
Current (pA/cm^2)	32	45.2	18.4	9.6
$M = M_{obj} * M_{proj}$	80.6*495	80.6*465	80.6*480	80.6*470

Table 3.4.1 the parameters of datasets when changing C2.

Data Series 2: 4 sets of data collected by changing the spot size.

Data set	Spot 2	Spot 3	Spot 4	Spot 5
M	80.6*340	80.6*375	80.6*380	80.6*390

Table 3.4.2 the parameters of dataset when changing the spot size.

3.4.2 Reconstruction results

As mentioned in the paper, for the SAD experimental setup, there is another way to remove the source transverse partial coherence – the Gaussian blind deconvolution method (Burdet et al., 2015; Maiden et al., 2015). In this section, we will show the complete 16 reconstructed modes of the 8 sets of data. Moreover, we also process the 8 sets of data with a two-dimensional Gaussian profile convolution method, and compared the results of the two ptychographic reconstruction methods. As described in chapter 2.7, with the partially coherent illumination, the diffraction pattern in the far field can be estimated as the convolution of the diffraction pattern from coherent illumination with a specific Gaussian profile.

$$I_m = I_{coh} \otimes G, \quad (3.4.1)$$

where I_m is the intensity of the m^{th} diffraction pattern from partially coherent illumination;

I_{coh} is the intensities of the diffraction pattern from the coherent illumination; G is a Gaussian profile.

$$G = e^{-\frac{(u^2+v^2)}{\sigma^2}}, \quad (3.4.2)$$

where σ is the standard deviation; u, v are the coordinates in the detector plane. So, in the specimen domain, we have

$$|\psi_m|^2 = |\psi_c|^2 \cdot g, \quad (3.4.3)$$

where ψ_m, ψ_c represent the exit wave from the sample with partially coherent illumination and coherent illumination respectively; g is the Fourier transform of G

$$g = e^{(y^2+x^2)\cdot\sigma^2}, \quad (3.4.4)$$

where x, y are the coordinates in the sample plane. The two-dimensional Gaussian profile can be described as

$$g = e^{\frac{x^2+y^2}{\sigma_x^2+\sigma_y^2}}, \quad (3.4.5)$$

where σ_x, σ_y are the standard deviations in the x and y directions. During the reconstruction, we will search the σ_x, σ_y that matches I_m to I_{coh} with the smallest error.

3.4.2.1 The eigenstates of data when changing the illumination convergence angle

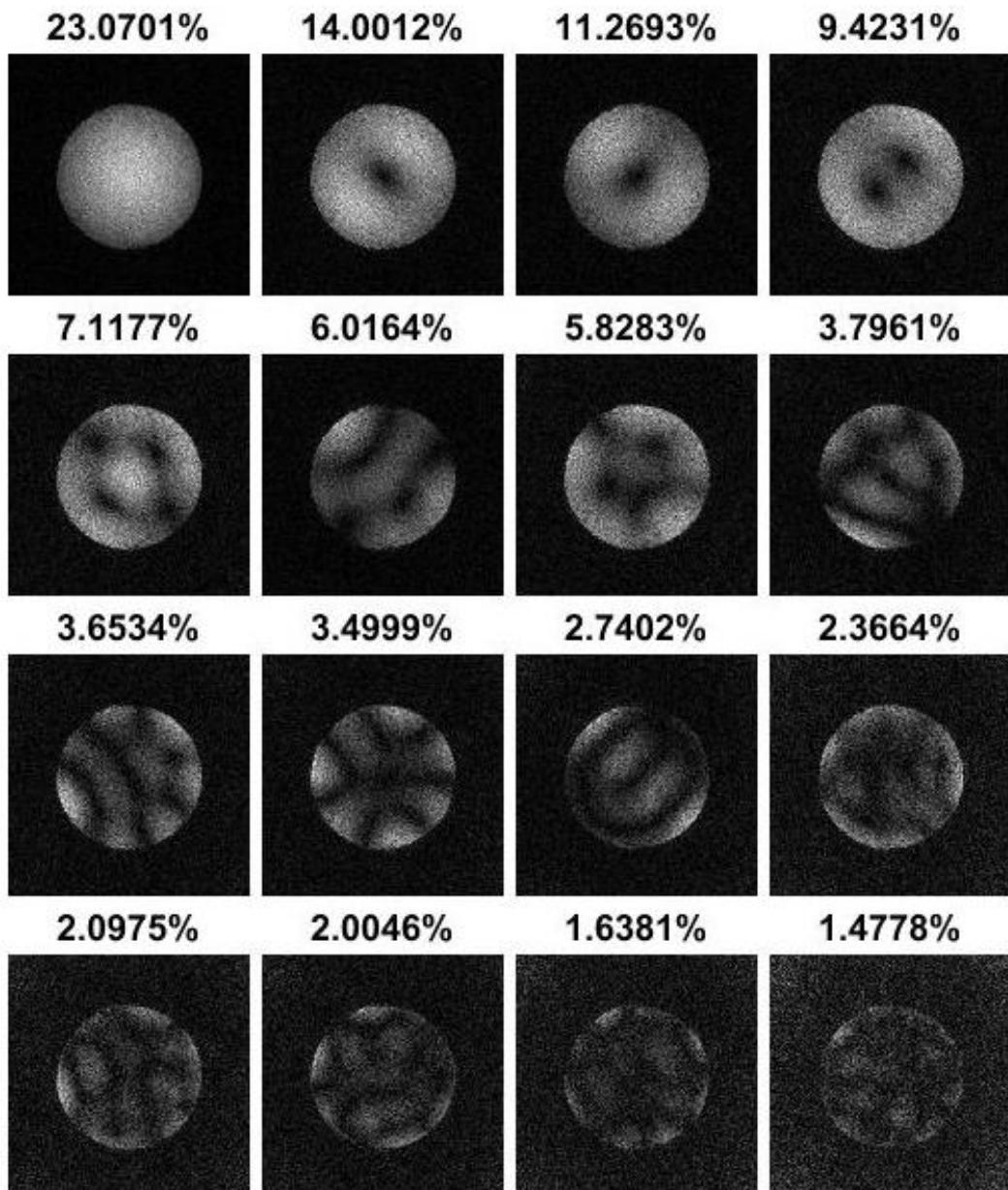


Figure 3.4.1. Eigenstates of dp1.

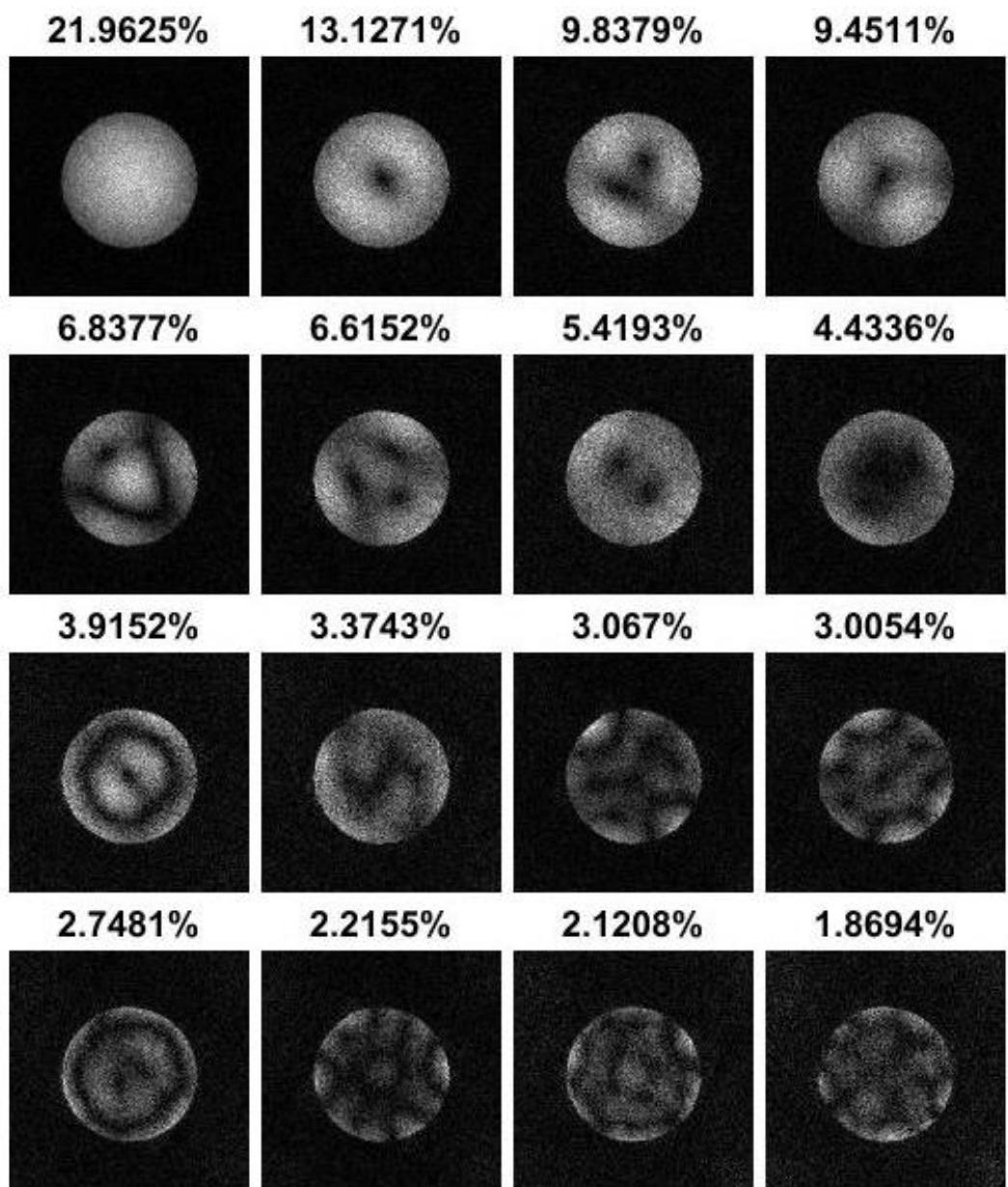


Figure 3.4.2. Eigenstates of dp2.

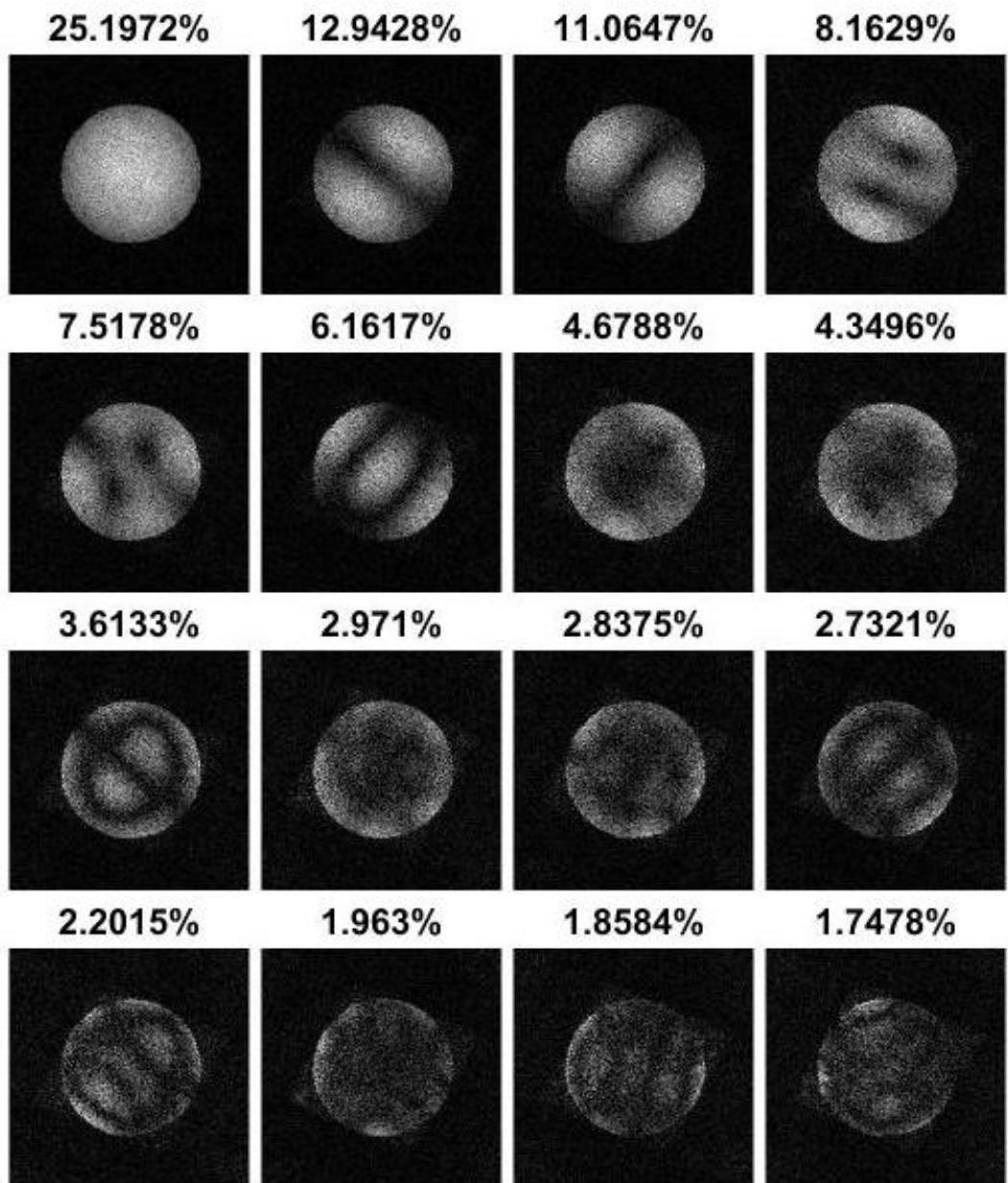


Figure 3.4.3. Eigenstates of dp3.

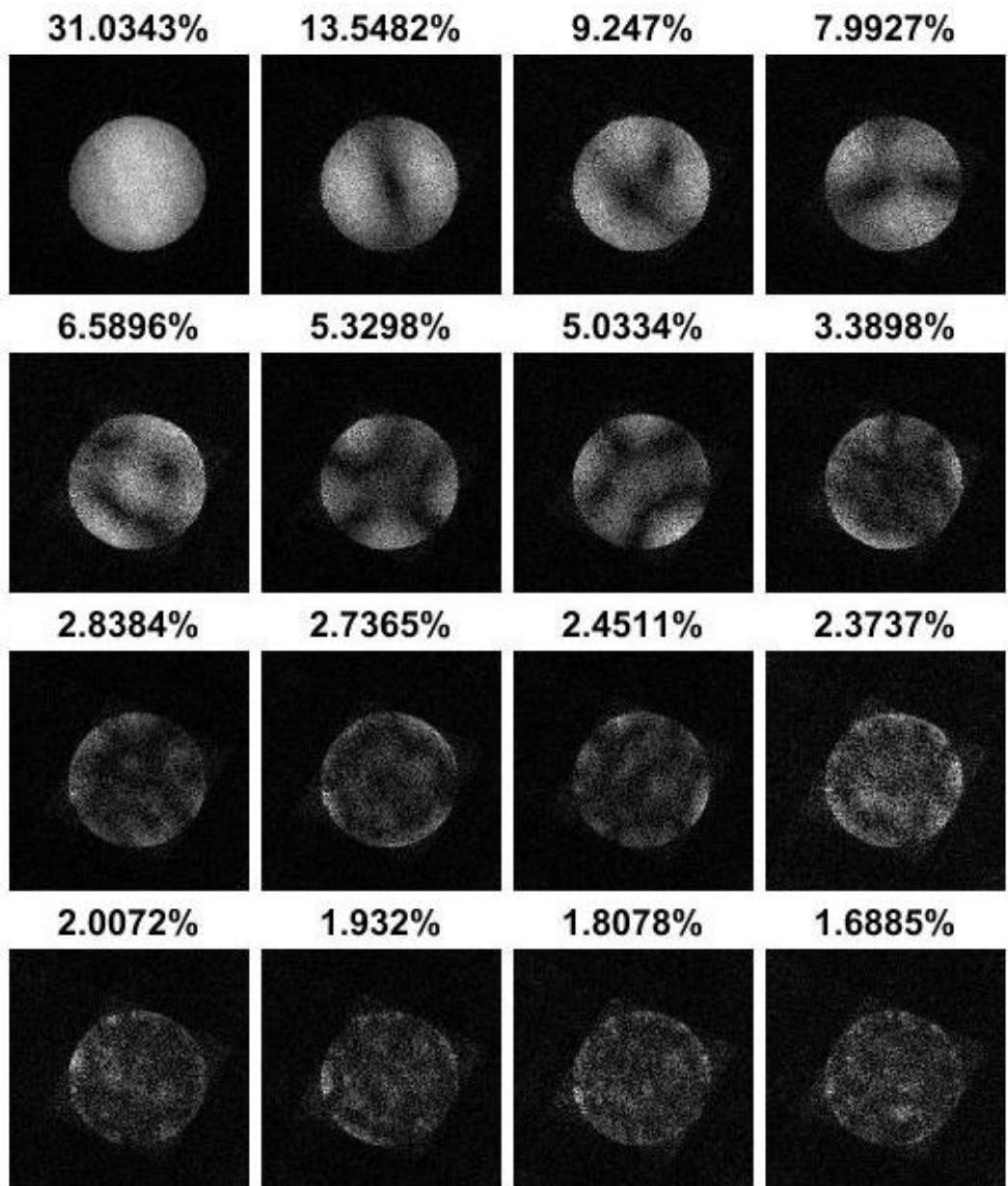


Figure 3.4.4. Eigenstates of dp4.

3.4.2.2 The eigenstates of data when changing the spot size

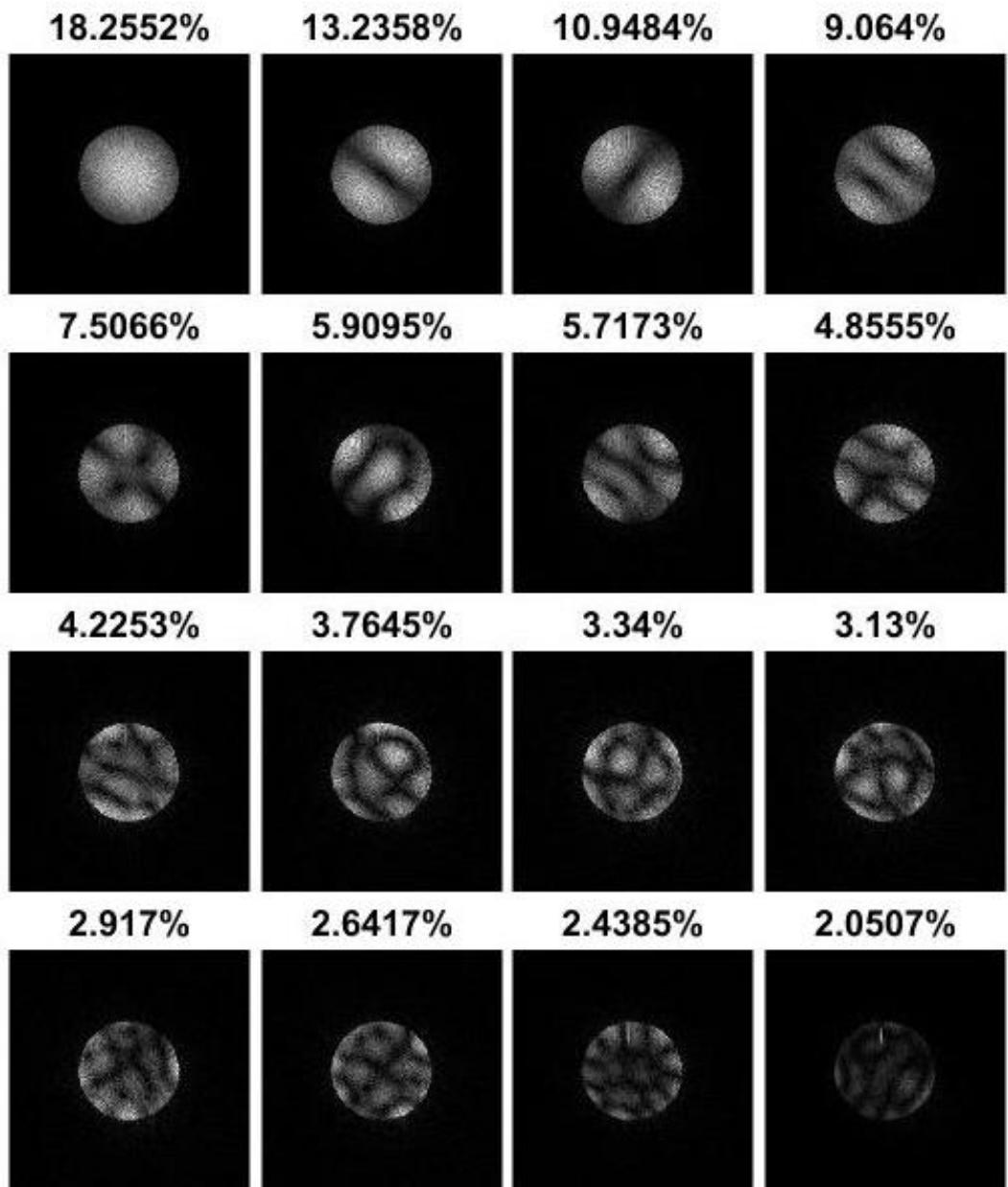


Figure 3.4.5. Eigenstates of spot2.

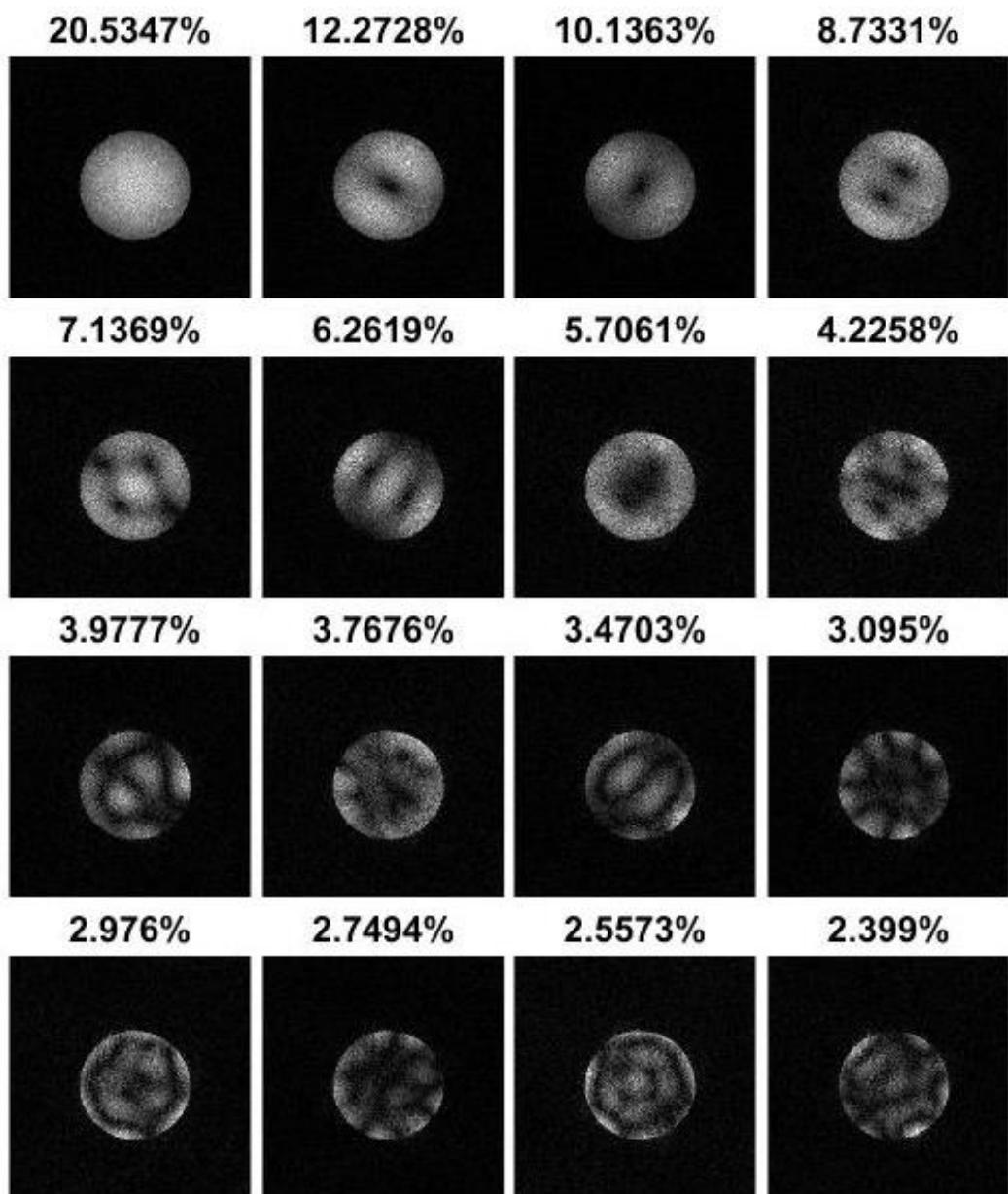


Figure 3.4.6. Eigenstates of spot 3.

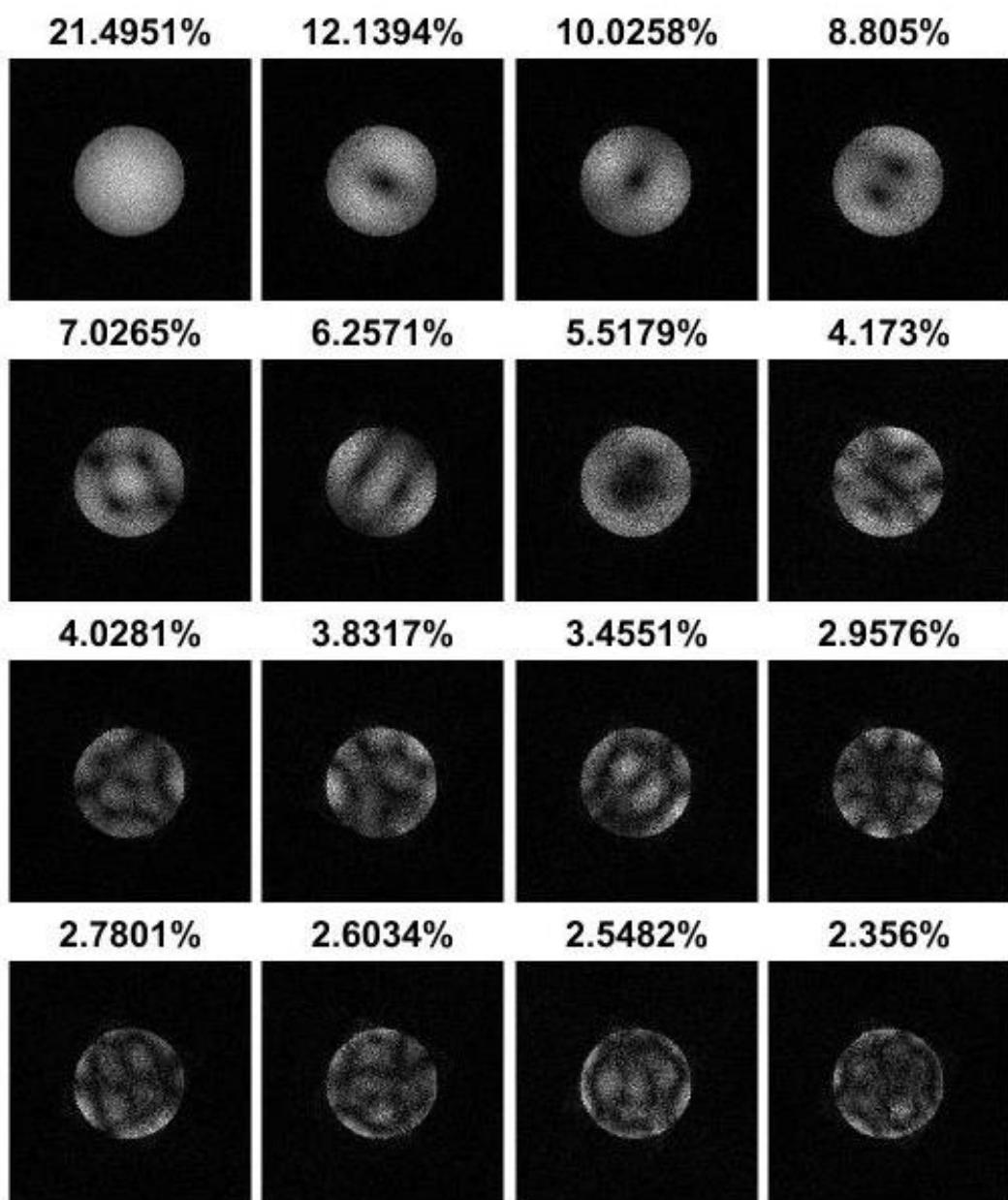


Figure 3.4.7. Eigenstates of spot4.

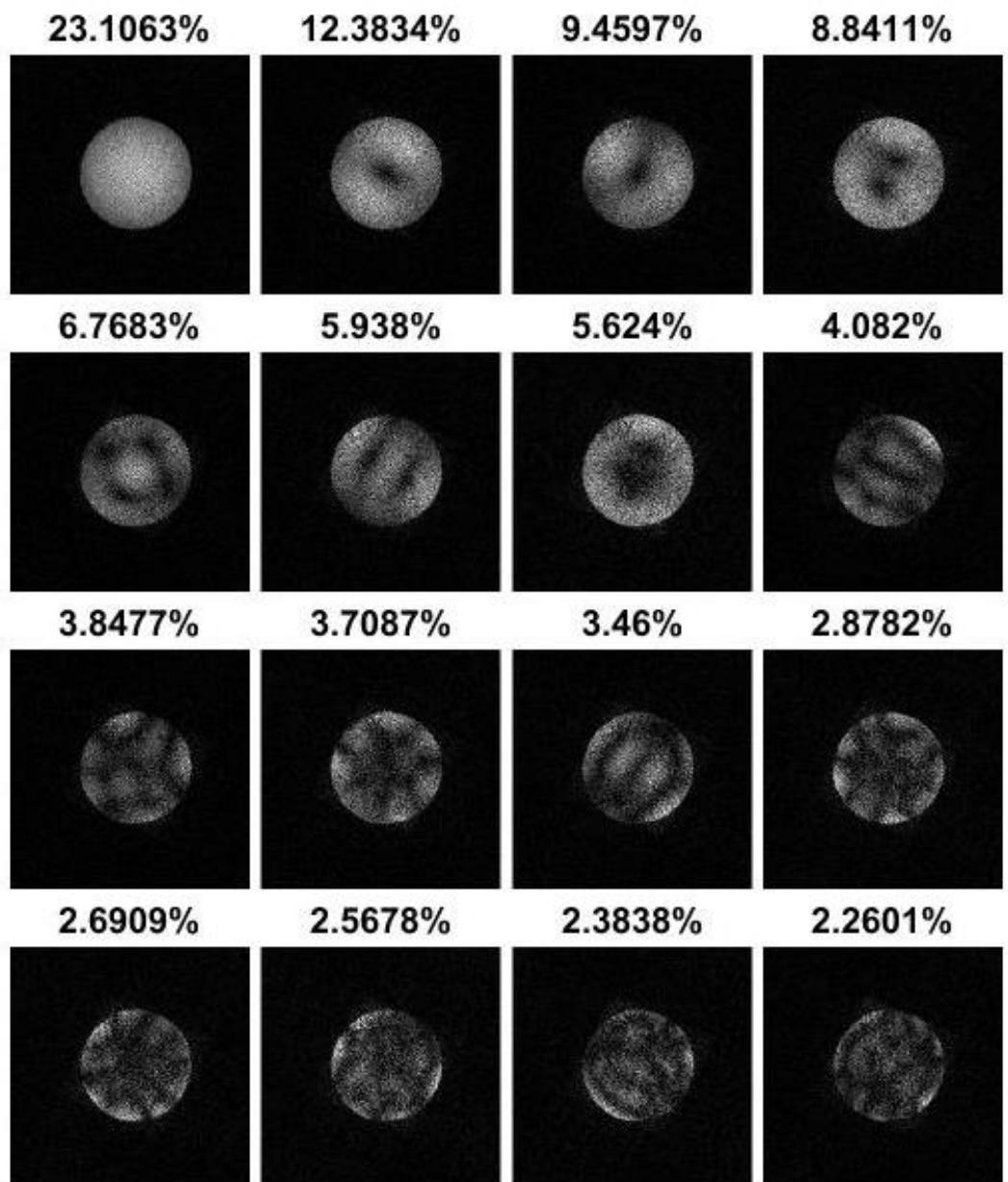


Figure 3.4.8. Eigenstates of spot5.

3.4.2.3 Reconstruction of the sample with ePIE, Gaussian and modes

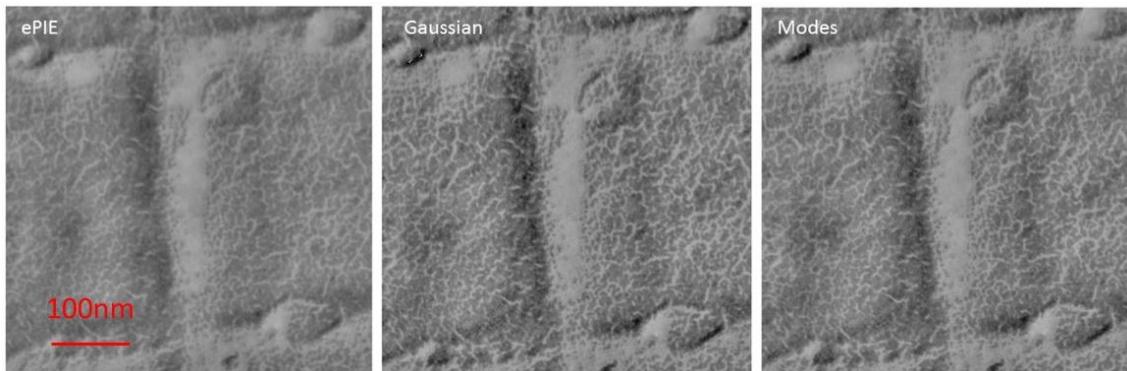


Figure 3.4.9 shows the phase reconstruction with ePIE, Gaussian blind deconvolution and modal decomposition method respectively. The Gaussian blind deconvolution and modes reconstruction results show better contrast than the conventional ePIE reconstruction.

Data	Dp1	Dp2	Dp3	Dp4	Spot2	Spot3	Spot4	Spot5
1 st ratio	23.0701%	21.9625%	25.1972%	31.0343%	18.2552%	20.5347%	21.4951%	23.1063%
σ	0.8180	0.9730	0.7838	0.7599	1.4237	1.3895	1.2896	1.1605

Table 3.4.3 compares the contribution of the 1st mode with the reconstructed standard deviation of the Gaussian profile from the 8 datasets. Dp4 has the largest 1st mode contribution, thus it also has the Gaussian profile with the smallest standard deviation. So, Dp4 has the most coherent illumination; Spot2 has the least coherent illumination.

3.5 How low counts can ptychography go?

3.5.1 Introduction

As we have discussed in Chapter 2.6, one of the disadvantages of electron microscopy is the radiation damage. There are different kinds of damage to the sample from electron radiation, for example, knock-on damage, radiolysis, sample heating and gas emission (Carlson and Evans, 2012; Buban et al., 2010; Downing, 1986; Egerton et al., 2004; Chiu and Jeng, 1982). The knock-on damage comes from when the electrons beam knock on the nucleus; if the beam energy is above the threshold energy, it will result in atom displacement (Carlson and Evans, 2012). The radiolysis is because the electron beam interacts with the valence electrons and breaks the chemical bonds. When the electrons have inelastic interaction either with the nucleus or with the valence electrons, there will be some secondary radiation generated.

Sample heating arises when the secondary radiation propagate inside the sample; gas emission comes from the sample decomposition from the radiation. The electron beam can also result the sample charging, which comes from both elastic scattering and inelastic scattering (Carlson and Evans, 2012; Buban et al., 2010). Increasing the accelerated voltage of the electron beam (above 300KeV) will increase the knock-on damage, but the radiolysis damage will be reduced. An electron beam accelerated with a voltage below 80KeV can avoid most knock-on damage; however, the low-energy electron beam will bring more radiolysis (Carlson and Evans, 2012).

The heavy atoms have higher dose tolerance than the light atoms; for example, carbon, hydrogen, oxygen and nitrogen, which are the main elements making up organic and biological material. For most biological specimens, the dose tolerance is $6e/\text{\AA}^2$. Taking protein as an example, it begins to have radiolysis damage when exposed to a dose of $10e/\text{\AA}^2$; some of its high-frequency diffraction spots will disappear; when the dose arrives $12e/\text{\AA}^2$,

there will some gas bubbles emission. The dose tolerance of polystyrene is $36e/\text{\AA}^2$; the dose tolerance of anthracene is $42e/\text{\AA}^2$; the dose tolerance of poly-xylene is $120e/\text{\AA}^2$. (Carlson and Evans, 2012; Buban et al., 2010; Egerton et al., 2004; Chiu and Jeng, 1982)

To avoid radiation damage to the sample, some low dose imaging techniques have been developed. Cryogenic-Electron Microscopy (Cryo-EM) is a transmission electron microscopy method which observes the sample at cryogenic temperature. The sample usually is kept at liquid nitrogen temperature. The low temperature will increase the radiation tolerance of the sample. For example, at a temperature of about 4K, the dose tolerance of protein arrives to $20e/\text{\AA}^2$ (Chiu and Jeng, 1982). Other techniques try to obtain an image with a low electron dose at normal temperatures, such as spot scanning imaging, electron crystallography and single shot dynamic TEM.

A recent paper by Pelz et al. (2017) summarised the doses used in some published ptychographic results. As shown in Figure 3.5.1. All these listed results were obtained with convergent beam electron ptychography either on a STEM or on a SEM. The results of D'Alfonso 2014 and Putkunz 2011 were obtained on a 300 KeV STEM with defocus-scan ptychography; the result of Humphry 2012 was obtained on a 30KeV SEM with defocus-scan ptychography; the result of Yang 2016 was obtained with focus-scan ptychography. All the tested specimens were heavy atoms. We can see that the dose in convergent beam electron ptychography was at least at the magnitude of $10^3 e/\text{\AA}^2$. The dose is still as high as $10^4 e/\text{\AA}^2$ even though with a direct camera (Yang et al., 2016); this is because the scanning is too intensive in focus-scan ptychography. These doses were far beyond what organic or biological specimen can stand.

Reference	Resolution	$e/\text{\AA}^2$
D'Alfonso et al., 2014	$\sim 1.5 \text{ \AA}$	3.98×10^4
Yang et al., 2016	Atomic	1.3×10^4
Putkunz et al., 2011	$\sim 1 \text{ \AA}$	9.2×10^6
Humphry et al., 2012	$\sim 2.3 \text{ \AA}$	3.33×10^3
SAD ptychography	$\sim 5 \text{ nm}$	1~5

Table 3.5.1 shows the electron dose and reconstructed resolution in some published ptychographic results.

As we compared in chapter 2.7.1 and 3.1, the selected area diffraction ptychography has the advantage of low dose when compared with the focused probe ptychography, since the electron beam is widely spread on the sample in the SAD ptychographic experiment. In this chapter, we systematically investigate how low dose can the SAD ptychography can handle, and the influence that it was on the reconstruction.

3.5.2 Experiment and results

Data was collected on the JEOL R005 HRTEM (300KeV), which has a cold field emission gun, the probe and image double aberration correctors. The specimen we tested contained gold particles on amorphous carbon film. The experimental setup was the same as shown in Figure 1.(b) in chapter 3.1; the TEM ran in SAD mode; a condenser aperture with diameter of 100um was inserted; a medium spot size (spot 3) was selected; the electrons beam was widely spread on the specimen; the objective lens formed an image of the sample with 80.6 magnification; the selected area aperture with diameter of 10um was inserted at the first image plane; the detector was placed behind the specimen with a nominal camera length of 2m; we changed the strength of the objective lens to project a defocus plane near the back focal plane onto the detector. The detector was a Gatan Orius SC200, with a dimension of

2048*2048 pixels and each pixel is of 7.4um; during the acquisition, the detector was binned by 2, so the frame size of the original diffraction pattern was 1024*1024.

Once the experiment is set up, there are four ways to change the counts of the intensities on the detector; 1) change the beam convergence angle by changing the strength of the condenser lens; the more spread the beam, the less dose to the specimen; 2) change the projection plane by changing the strength of the projector lens; the closer the projected plane to the back focal plane, the more counts on the detector; 3) change the exposure time; shorter exposure time, less counts; 4) change the camera length; shorter camera length, more counts on the detector per unit area. To characterise the organic or biological sample, we can reduce the radiation dose by using more widely spread beam, a shorter exposure time, a closer plane to the back focal plane and a shorter camera length.

In this chapter, we mean to test the counts that ptychographic reconstruction algorithm ‘ePIE’ (Maiden et al., 2009) requires; we reduce the counts on the detector simply by reducing the exposure time, while keeping all the other parameters unchanged. In this case, we did not reduce the radiation damage to the sample; our sample was exposed under the radiation all the time. We collected 3 sets of diffraction patterns of the same area on the sample, with a 15*15 grid scanning by shifting the sample. The scaled illumination size was about 120nm; the scanning step size was 10-20nm. Dataset 1 was collected with an exposure time of 0.1s; the maximum readout value of the diffraction patterns was about 350. Dataset 2 was collected with an exposure time of 0.05s; the maximum readout value of the diffraction patterns was 180. Dataset 3 was collected with an exposure time of 0.01s; the maximum readout value was 40. Figure 3.5.1 shows the examples of the diffraction patterns from each dataset.

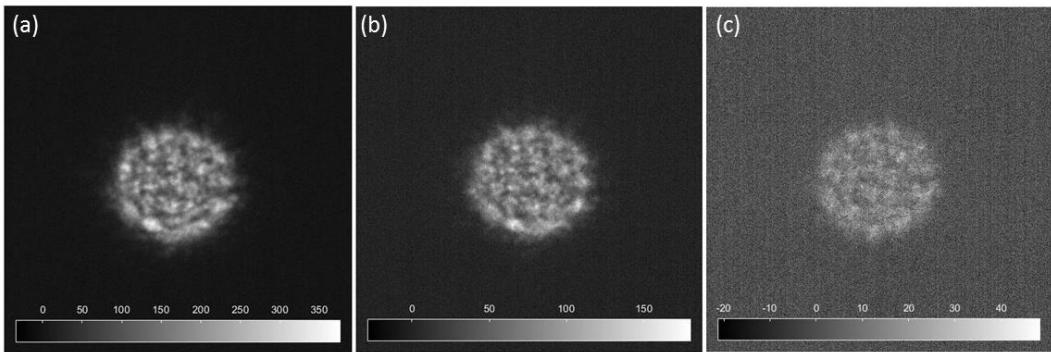


Figure 3.5.1 shows the examples of the diffraction patterns from the 3 sets of data, with exposure time of 0.1s, 0.05s and 0.01s respectively; the corresponding highest counts of the 3 sets of data are 350, 150 and 40.

The reconstruction algorithm we have applied to process the data was ePIE (Maiden and Rodenburg, 2009), with ‘jiggle’ (Maiden et al., 2012), diffraction pattern drift correction, and Gaussian blind convolution decoherence (Maiden et al., 2015; Burdet et al., 2015).

We calculated the counts with the same method that was used in the paper by Pelz et al. (2017) by calculating how many electrons were detected on the CCD, which is not an accurate way to measure the radiation dose to which the specimen was exposed to. The doses of the three sets of data are $10e/\text{\AA}^2$, $5e/\text{\AA}^2$ and $1e/\text{\AA}^2$ respectively. The figures below show the reconstruction results of the 3 sets of data.

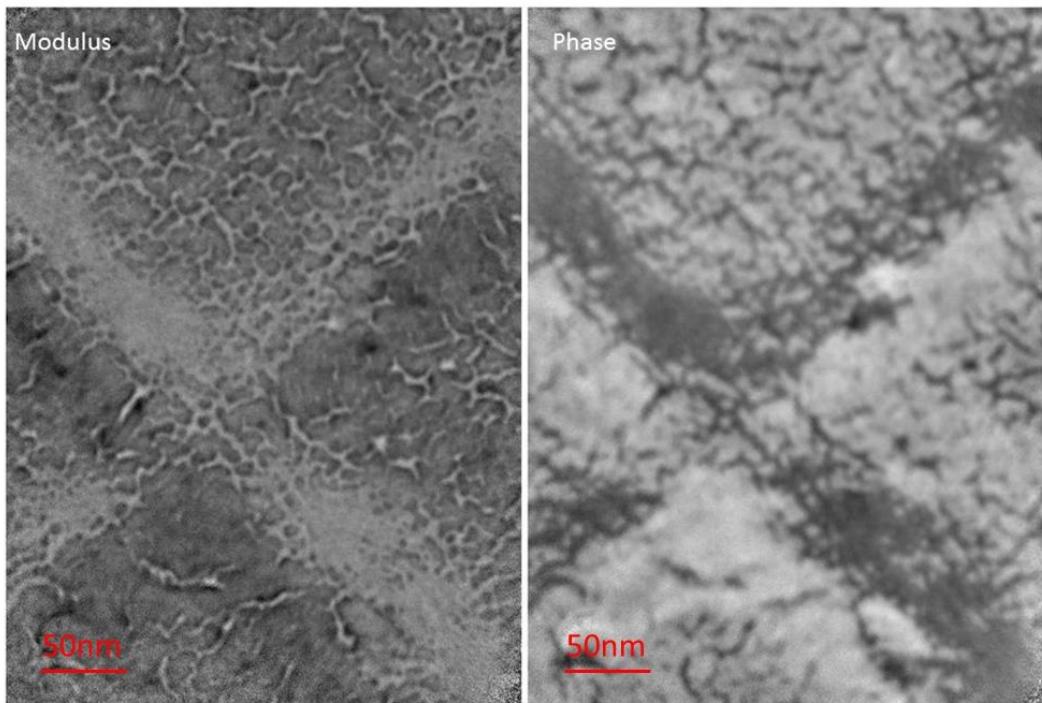


Figure 3.5.2 shows the modulus and phase reconstruction from dataset with a radiation dose of about $10e/\text{\AA}^2$. The smallest gold nanoparticle that can be seen is of dimension about 5nm.

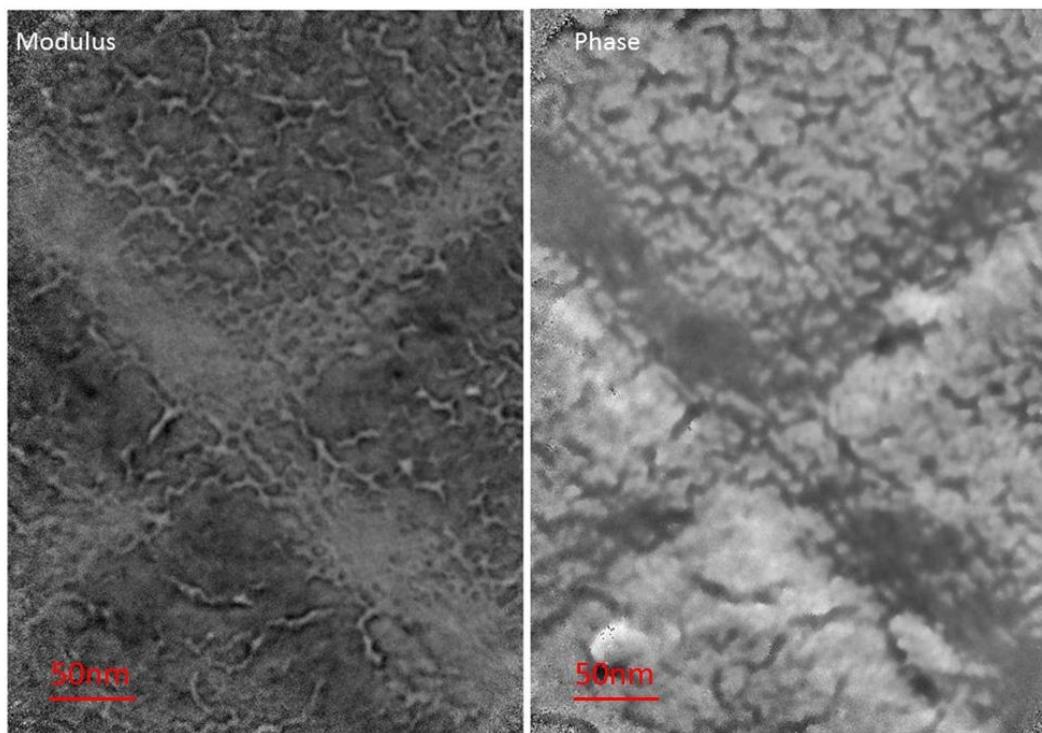


Figure 3.5.3 shows the modulus and phase reconstruction from dataset with a radiation dose of about $5e/\text{\AA}^2$.

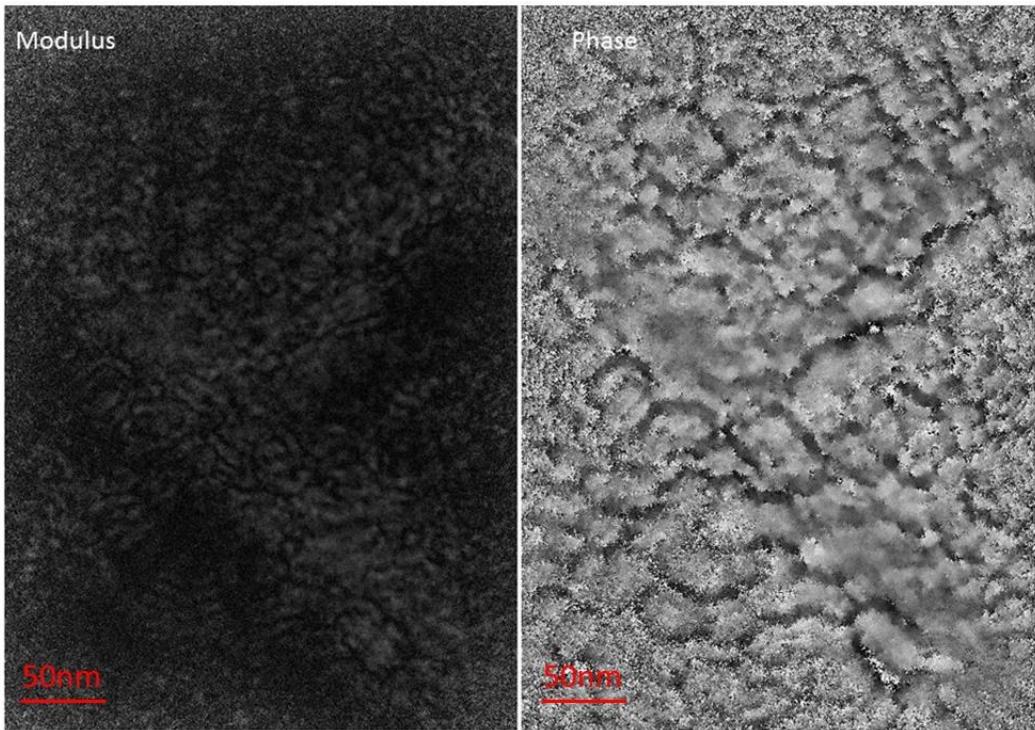


Figure 3.5.4 shows the modulus and phase reconstruction from dataset with a radiation dose of about $1e/\text{\AA}^2$.

3.5.3 Discussion

When we reduce the dose from $10e/\text{\AA}^2$ to $5e/\text{\AA}^2$, the reconstructed contrast of the sample is reduced, but there is not obvious resolution degradation. When we further reduce the dose to $1e/\text{\AA}^2$, ePIE cannot reconstruct the structures of the specimen any more. So, the dose limit of ptychography iterative reconstruction algorithm (ePIE) is about $5 e/\text{\AA}^2$.

The SAD ptychography gives much lower dose radiation to the specimen than the focused beam setup on a SEM (Humphry et al., 2012) or STEM (D’Alfonso et al., 2014; Yang et al., 2016 and 2017). However, there are some issues that we have to consider with the SAD ptychography. Firstly, the resolution of SAD ptychography cannot be comparable with the convergent beam scan ptychography; of the results that have been presented in this chapter, the resolution is about 5nm, which is the size of the smallest gold nanoparticle. This

resolution is far behind the information limit of the objective lens (0.07nm). However, if we can have a smaller selected area aperture to localize a smaller area on the sample, or if we can have an objective lens with a higher magnification, the reconstruction resolution of SAD ptychography can be further improved. Secondly, the dose we calculated in this paper is how many electrons per area arrived on the detector; in the SAD ptychographic data collection, the whole region of interest of the sample is exposed under the electron beam all the time. With a conventional CCD camera, to collect 225 diffraction patterns, it takes about 20 minutes when the exposure time is 1s; during such a long time, even though the dose is low, it still will cause contamination or even damage to light atoms such as carbon film or organic samples.

There are some options in the experiment which may be used to further reduce the radiation exposed on the sample while maintaining the counts on the detector no less than $5e/\text{\AA}^2$; 1) shorten the camera length. The data was collected at the camera length of 2m in this chapter. It is possible to collect the data at a shorter camera length. The electron microscope has a camera length range between 8cm to 2m. 2) change the projective lens defocus closer to the back focal plane. When we scan a light specimen (organic, biological, or materials made of light atoms), to reduce the radiation damage, we can collect the diffraction patterns at a defocus plane as close as possible to the focal plane, while also considering the dynamic range of the detector and the sampling condition (Edo et al., 2013; Batey et al., 2014).

Chapter 4 Image feature delocalization in defocused probe ptychography

4.1 Introduction

There are a number of different optical configurations possible for electron ptychography. A lens can be used to focus a diffraction-limited spot (a probe) into the plane of the specimen, with diffraction data being recorded in the Fraunhofer plane far downstream from the object (Nellist, McCallum and Rodenburg, 1995). In order to satisfy the overlap requirement, the probe must be scanned on a very fine grid, with two-dimensional diffraction patterns being recorded from every pixel required for the final reconstruction. Two closed form inversion methods for this type of data, which is densely sampled in real-space, were developed during the 1990s (Bates and Rodenburg, 1989). The methods were in those days greatly hampered by poor detector technology and the small computer memory available. There is now renewed interest in this type of ptychography: recent results using the latest detector technology have proved very successful (Yang et al., 2016 and 2017). With the advent of iterative ptychographic reconstruction techniques (Rodenburg and Faulkner, 2004; Thibault et al., 2008; Maiden and Rodenburg, 2009), it is also possible to illuminate an extended area of the object and move it in large steps; i.e. the sampling in real space can be made coarse. One

of the benefits of this approach is speed, because a large field of view can be spanned very quickly. There is also a much smaller requirement on the amount of data that needs to be collected and stored. Furthermore, unlike in the focused probe configuration, the lens used to condense the electrons can have a small angular size if the principal source of high-resolution information arises from the dark-field region of the diffraction pattern lying outside the bright central Ronchigram.

A possible application for this type of large step size (LSS) electron ptychography is to convert a machine of moderate specification, such as a conventional scanning electron microscope (SEM), into a good quality transmission electron microscope (TEM) simply by placing a two-dimensional detector behind a specimen holder that can accommodate a transmission specimen. A proof-of-principle of this approach was demonstrated some years ago (Humphry et al., 2012). However, the community expressed some doubt about the results, specifically that the fringes in the gold particles that were used as a test specimen were, in some instances, delocalised, extending beyond the boundaries of the particles.

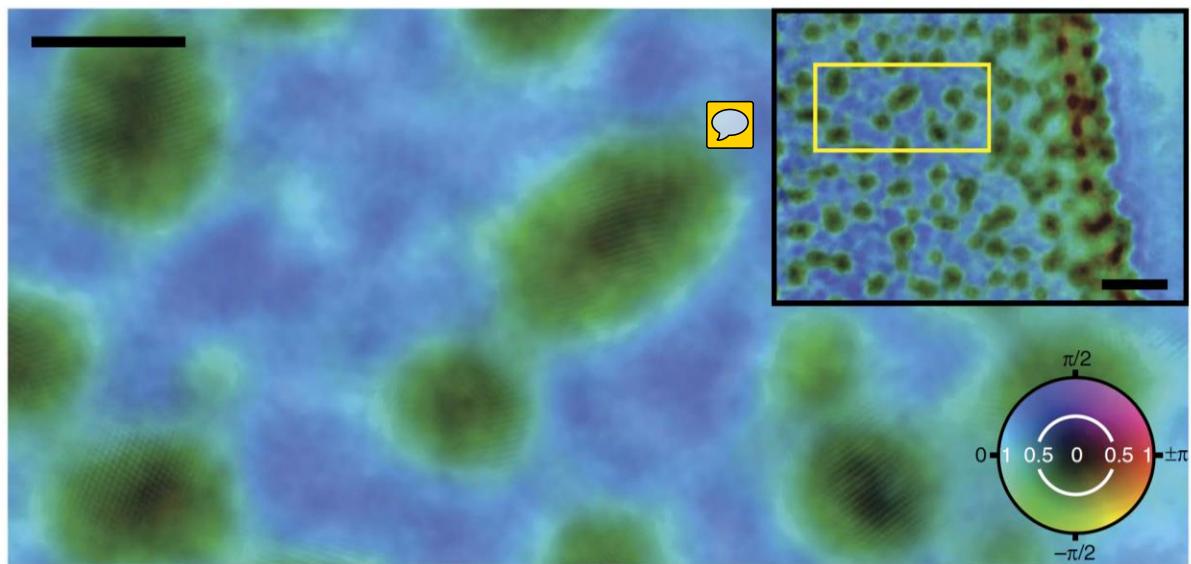


Figure 4.1 shows the reconstruction result from paper (Humphry et al., 2012). Some particles appear lattice fringes delocalization problem.

We investigate various influences that affect the fidelity of electron ptychographic reconstructions in this defocused probe configuration, especially as they relate to delocalisation of atomic fringes. Clearly, if fringes (i.e. Fourier components comprising the image) are laterally shifted, which is equivalent to their diffracted intensities being assigned the wrong phase, then the image is meaningless as far as interpreting the position of atomic columns, rather like an out of focus conventional bright-field image that has negative regions in its contrast transfer function. We test here via model calculations the effects of 1) partial spatial coherence caused by the source size, 2) a noise pedestal in the detector output, 3) the point spread function of the detector, 4) abrupt truncation of the signal at the detector edges, and 5) errors in the probe positions. We then examine the efficacy of various methods for removing these imperfections in the data with some simulations, before applying the same methods to the experimental data.

4.2 Factors result in delocalization

4.2.1 Source transverse partial coherence

In the previous reconstruction of the data collected on the SEM (Humphry et al., 2012), the illumination was assumed to be fully coherent, which means the source was assumed to be a point source. In reality, the source of an electron microscope is never fully coherent (Nguyen, Findlay and Etheridge, 2014). The source generated by a cold field emission gun has a physical size of about 10nm, which can be demagnified into a beam spot of several angstroms by the condenser lenses and condenser aperture; the source generated by the thermionic emission gun is larger (~100nm). Figure 4.2.a shows a simulated symmetric source with a dimension of about 5Å*5Å. The incoherent illumination at a defocused plane formed with this source is shown in Figure 4.2.b, which was simulated by propagating each point source to the Fourier domain, multiplying with an aperture, propagating to a defocused plane, and superpositioning the illumination formed from each point source:

$$P_a = \mathcal{F}^{-1}\{\mathcal{F}\{S_a\} \times A \times e^{i\theta}\}; \quad (4.1)$$

$$P = \sum_{a=1}^m P_a, \quad (4.2)$$

where \mathcal{F} represents the Fourier transform, S_a represents the a^{th} point source in the source planes; A represents the aperture; $e^{i\phi}$ is the object potential function.

When the partially coherent illumination is transmitted through an object, the diffraction pattern intensities on the detector are

$$I_{pc} = \sum_{a=1}^m |\mathcal{F}\{P_a \times O\}|^2, \quad (4.3)$$

where I_{pc} is the partially coherent diffraction pattern; O is the object (Born and Wolf, 1999).

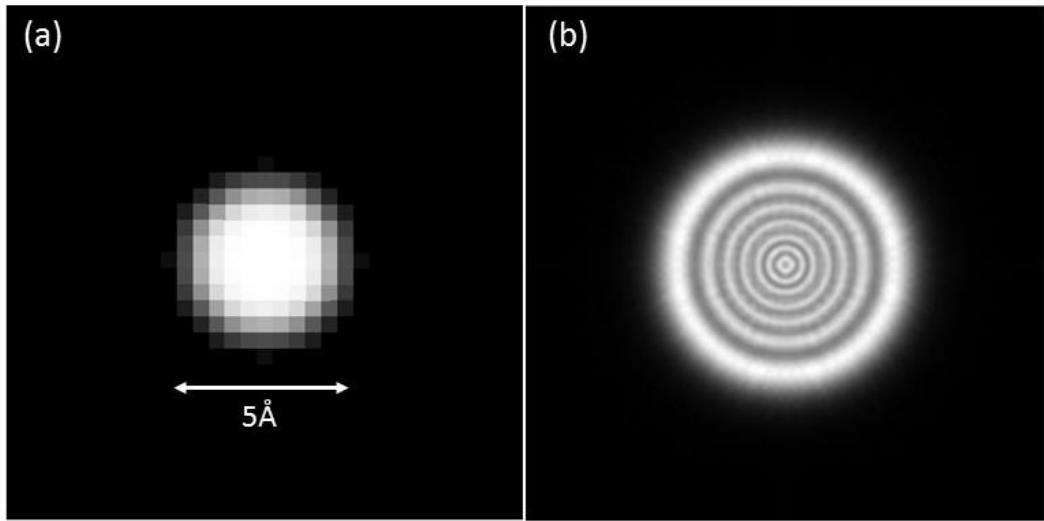


Figure 4.2. (a) is the transverse partially coherent source used for the simulations; the dimension of the source is about $5\text{\AA} \times 5\text{\AA}$. (b) shows the partially coherent illumination generated from (a) according to the setup shown in Figure 4.4; this is the illumination used to simulate the partially coherent diffraction patterns for the ptychographic reconstruction.

Figure 4.3 shows the object used for the simulations in this paper; it is a complex object, with the modulus varying between 0 and -1 and phase varying between -1 and 1. The red crosses on the object show the probe scanning positions when collecting the ptychographic data. We modelled this object to simulate the atomic fringes along a perpendicular projection. Figure 4.4 shows the experimental setup for the simulations.



Figure 4.3 shows the modulus of the complex object used the simulations. The modulus of the object varies between [0, 1]; the phase of the object is the same pattern with the modulus, varying [-1, 1]. The red crosses show the scanning positions when collecting the ptychographic data.

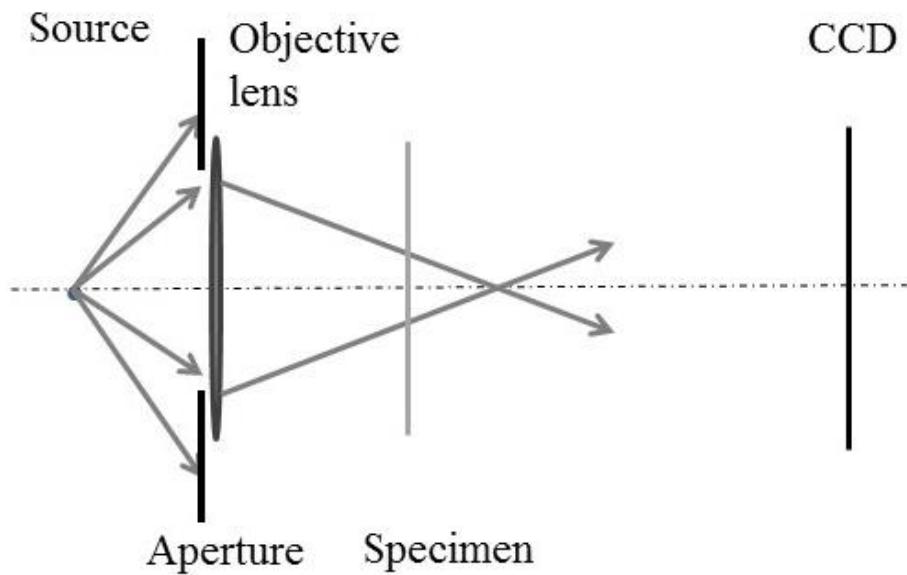


Figure 4.4 shows the setup engaged for the simulations, which is also a simplified configuration of the experimental setup: a condenser aperture was inserted in the far field with respect to the source domain; the objective lens formed a probe at a slightly defocused plane on the specimen; the detector was placed in the far field respect to the sample.

Figure 4.5.a and 4.5.b compare the diffraction patterns generated from coherent illumination and partially coherent illumination shown in Figure 4.2.b. We can see that, when the illumination was fully coherent, the diffracted structures inside the central disc of the Ronchigram have clear fringes; these fringes were related to the structure of the sample. When the illumination is partially coherent, the diffracted fringes inside the central disc of the Ronchigram are blurred. When we use the blurred diffraction patterns from the partially coherent source to do ptychographic reconstruction, while assuming that the source is fully coherent, it is like we are searching for the solution with a relaxed Fourier constraint. In this case, there will be a lot of ambiguous solutions, but none of them is correct. Figure 4.6.a shows the ptychographic reconstruction from the partially coherent data with the conventional ePIE algorithm when assuming the illumination was fully coherent. The reconstruction quality is quite bad; the features are blurred; moreover, the fringes are

delocalized. So, the source transverse partial coherence is one of the reasons that result in the reconstructed fringes delocalization problem.

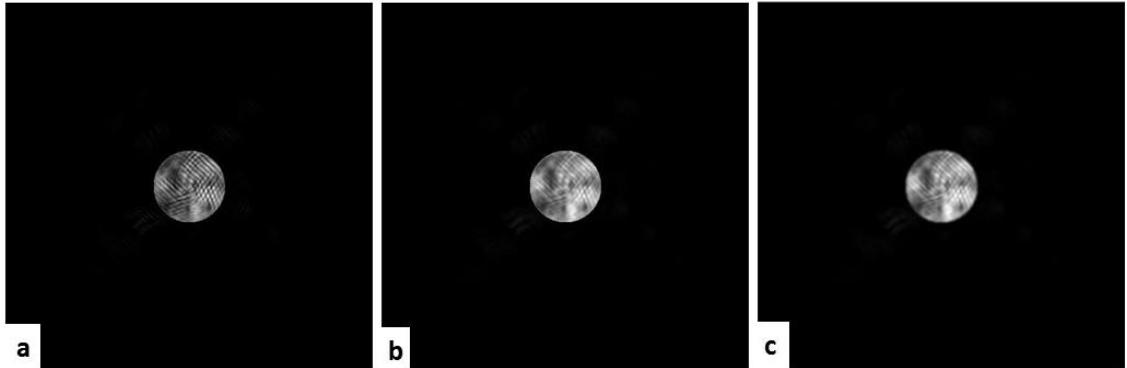


Figure 4.5 shows the examples of the diffraction patterns generated from (a) fully coherent illumination and (b) partially coherent illumination. (c) shows one diffraction pattern generated from coherent illumination, while considering the point spread function of the detector; the detector point spread function was approximated as a 2-dimensional Gaussian profile with the derivation of 1.5 pixels in both dimensions.

In chapter 2.7 and chapter 3, we have discussed two methods for decoherence in ptychographic reconstruction: Gaussian blind convolution (Burdet et al., 2015; Maiden et al., 2015) and modal decomposition (Thibault and Menzel, 2013; Li et al., 2016). Here we apply the modal decomposition method based on ePIE to improve the reconstruction of the partially coherent data. Instead of engaging a single illumination function in the conventional ePIE reconstruction, we used six illumination profiles during the reconstruction. Thus, in the Fourier domain, the update function at each scanning position becomes

$$\Psi'_n = \sqrt{I} \frac{\Psi_n}{\sqrt{\sum_n |\Psi_n|^2}}, \quad (4.4)$$

where n registers the order of each probe functions; I is the detected intensity, and Ψ is the Fourier transform of the exit wave.

Then we propagated Ψ' to real space by Fourier transform, to get the updated exit wave function ψ' . The next step is to update each of the probe functions and the object function by

$$P'_n = P_n + \alpha \frac{\psi'_n \cdot O^*}{|O|_{max}^2}, \quad (4.5)$$

$$O' = O + \beta \cdot \frac{\sum_n (\psi'_n \cdot P_n^*)}{\sum_n |P_n|^2}, \quad (4.6)$$

where P'_n is the n^{th} updated probe function; O^* is the conjugate of the object function; α and β are scaling factors to control the update step. We took $\alpha = 1$ and $\beta = 1$ during the reconstruction of the simulated data.

The specimen reconstruction on the partially coherent data with modal decomposition method is shown in Figure 4.6.b; compared with the conventional ePIE (shown in Figure 4.6.a), the decoherence effect of the modal decomposition method is apparent; the features of the sample are clearer, and the fringes are more localized. However, from the magnified features we can see that the fringes are still slightly delocalized, which means that the effect of the source partial coherence has not been removed completely. Figure 4.7 compares the Fourier ring correlation between the true object with the ePIE reconstruction and the modal decomposition reconstruction respectively. Modal decomposition method improves the reconstruction quality of the partially coherent data in both low frequencies and high frequencies significantly. As investigated in Chapter 3, the efficiency of the modal decomposition method is affected by 1) the specimen structure diversity, 2) the propagation distance, 3) the overlap amount in the real space and 4) the reconstructed area. We analyse that it is the specimen structure diversity affecting the modal decomposition reconstruction efficiency in this simulation. Figure 4.8.a shows the reconstructed decompositions of the partially coherent illumination function in the eigencoordinates at the sample plane. The values above each mode show the percentage of the corresponding mode contributes to the

illumination function; the larger percentage the first mode takes, the better coherency the effective source has. In this simulation, when the source size is about 5 Å, and the aperture size is about one fifth of the frame window, the first mode takes about a 47% contribution of the illumination. For comparison, the first mode of an electron source generated by a cold field emission gun (the results in chapter 3) takes 20%-30%; the first mode of the synchronous X-ray source takes about 30% (Li et al., 2016); the first mode of the diode laser source takes about 95%.

We can conclude that the transverse partial coherence of the source will result in the features delocalization in the ptychographic reconstruction (Figure 4.6.a). Considering the coherence degree of the electron source of the SEM for collecting the data in the paper (Humphry et al., 2012) source partial coherence is most likely one of the factors that result in the features delocalization. As has been proved in the simulation data, modal decomposition is an efficient method for decoherence; it has improved the reconstruction quality significantly.

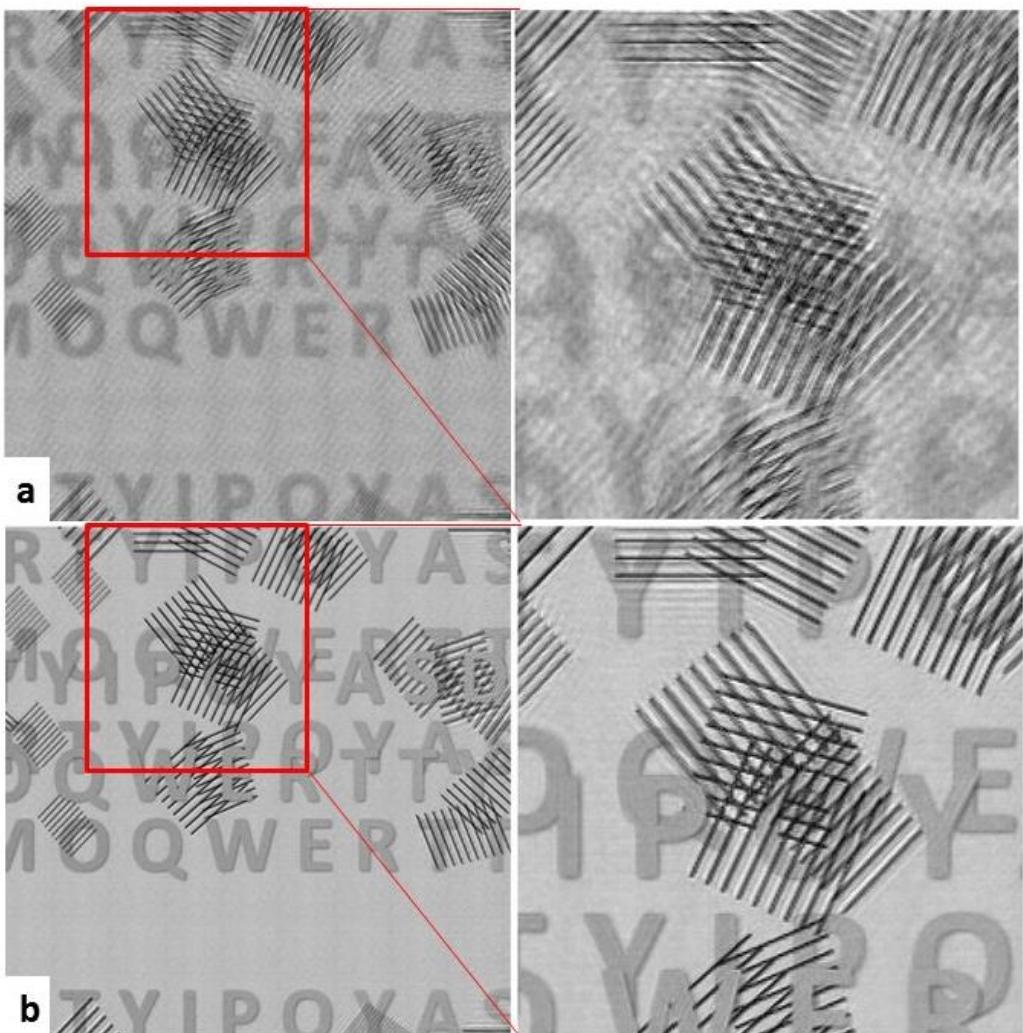


Figure 4.6 compares the reconstructions from the partially coherent dataset with the conventional ePIE (a) and with the modal decomposition method (b). The figures on the right show the magnified features; we can see that when applying the conventional ePIE to reconstruct the partially coherent data, the reconstruction quality is very bad; the features are unclear and delocalized; while the modal decomposition method has improved the reconstruction a lot, even though it has not completely removed the influence of the source partial coherence.

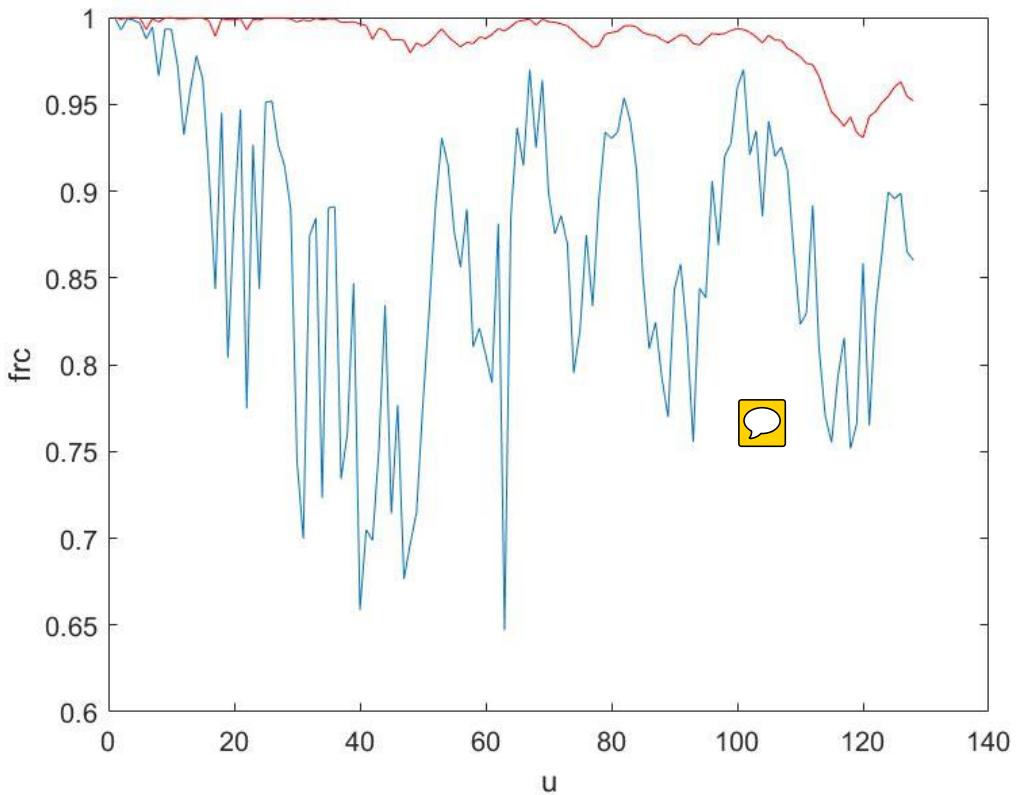


Figure 4.7 shows the Fourier Ring Correlation between the true object with the reconstruction results from ePIE and modal decomposition respectively. The blue line plots the FRC between the true object with the ePIE reconstruction; the red line plots the FRC between the true object with the modal decomposition reconstruction.

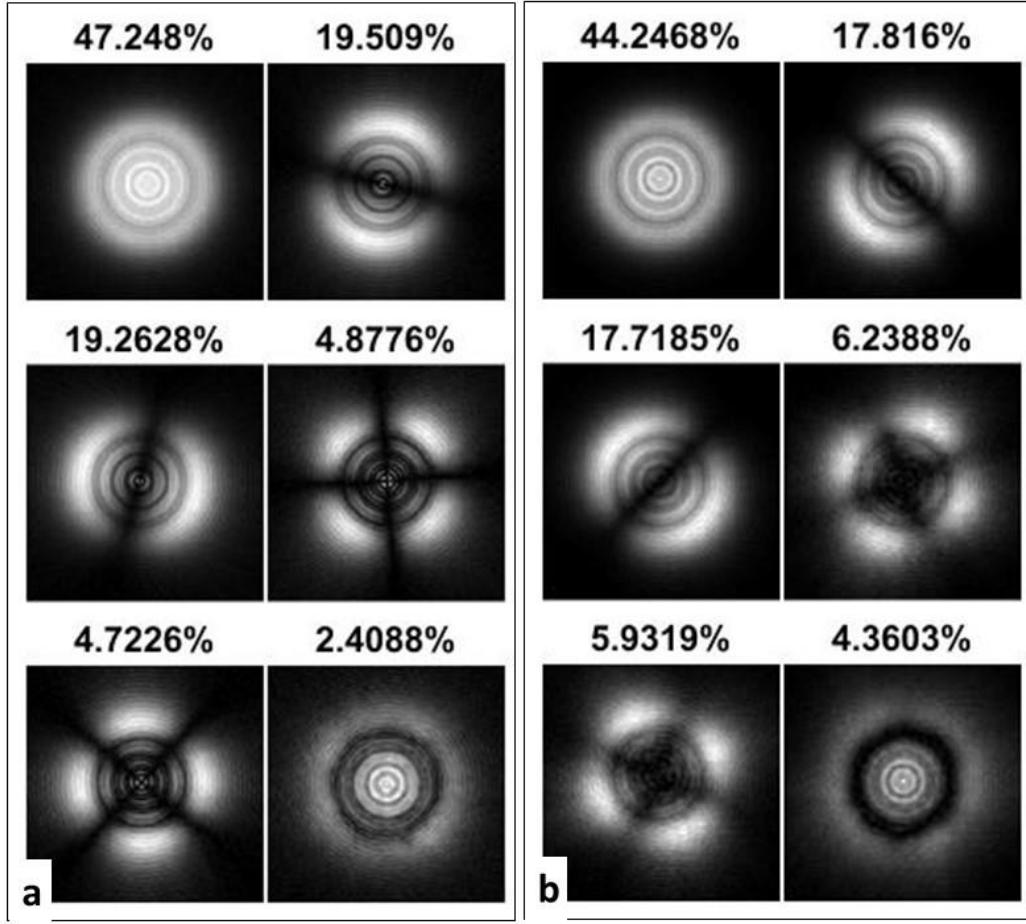


Figure 4.8 shows the modes reconstructed with modal decomposition method at the sample plane. (a) shows the modes of the partially coherent source; the first mode takes about 47%. (b) shows the modes reconstructed from the data with the effect of detector point spread; the detector point spread was simulated as a Gaussian profile convolution with a standard derivation of 1.5 pixels; the first mode takes about 44%; the detector point spread has a similar effect with the source transverse partial coherence.

4.2.2 Detector point spread

The detector response property affects the quality of the diffraction pattern or the image directly. Detector point spread is a common phenomenon existing when the Charge Coupled Device detector is excited by photons or electrons. The influence of the detector point spread on the signal can be estimated as the convolution of the true signal with a Gaussian profile:

$$I_m = I \otimes G_\sigma, \quad (4.7)$$

where I_m represents the diffraction pattern with the influence of detector point spread; G_σ is a Gaussian profile with standard deviation equals σ . The effect of detector point spread is similar to the source partial coherence; both of them will make the diffraction pattern blurred as shown in Figure 4.5.c.

In the simulation, we generated a two-dimensional Gaussian profile with the standard derivation equal to 1.5 pixels, to simulate the detector point spread effect. The reconstruction of the sample with the conventional ePIE algorithm is shown in Figure 4.9.a; the reconstruction quality is similar to the result with the partially coherent data shown in Figure 4.6.a; the features and the boundaries of the fringes are blurred. As demonstrated in the previous work (Thibault and Menzel, 2013; Li et al., 2016), the modal decomposition method also works to remove the detector point spread effect on the specimen reconstruction, by retrieving the detector point spread as decomposed modes of the illumination. Figure 4.9.b shows the reconstruction of the same data with the modal decomposition method; the features are clear and localized; the modal decomposition method has almost completely removed the effect of the detector point spread. Figure 4.8.b shows the reconstructed modes from the data with the detector point spread effect; the modes have the same patterns with the modes reconstructed from the partially coherent data.

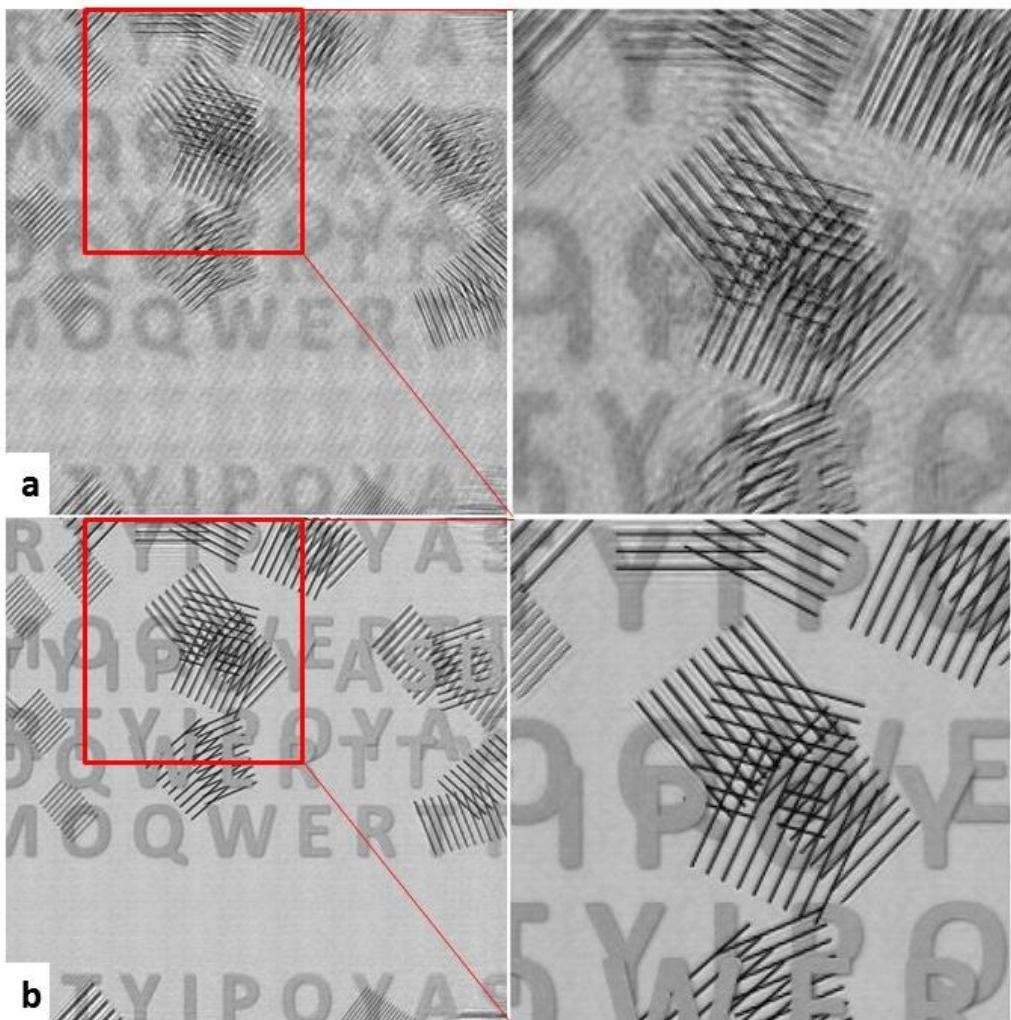


Figure 4.9 shows the reconstructed phase from the data with the detector point spread effect. (a) is the reconstruction with the conventional ePIE algorithm; the features on the sample are blurred, and we cannot see the clear boundaries of fringes. (b) shows the reconstruction with the modal decomposition algorithm; the features and the boundaries are clear.

4.2.3 Detector pedestal

Another problem of the detector is the response pedestal, which is the problem that some pixels of the detector will have read-out values without the excitation of the photons or electrons. The detector pedestal may vary between different pixels, and it may change nonlinearly according to the strength of the excitation; thus, it will result in a distortion of the

signal. As described in Chapter 2.7 and Chapter 3, we can characterize the response property of each pixel on the detector as

$$I_{out}(u) = I(u) + c(u), \quad (4.8)$$

where u denotes the pixels on the detector; I_{out} denotes the readout value; $I(u)$ denotes the true response; $c(u)$ denotes the pedestal at each pixel. Figure 4.10 compares the diffraction pattern when adding pedestal, c , of 0, 0.005, 0.01, and 0.5. We can see that, as the value of the pedestal increases, more high-frequency diffraction speckles are submerged. One ptychographic reconstruction with the conventional ePIE from the dataset when $c = 0.005$ is shown in Figure 4.11.a; the reconstruction is very noisy; the features on the sample are blurred; some features, for example, the letters in the background, have nearly disappeared.

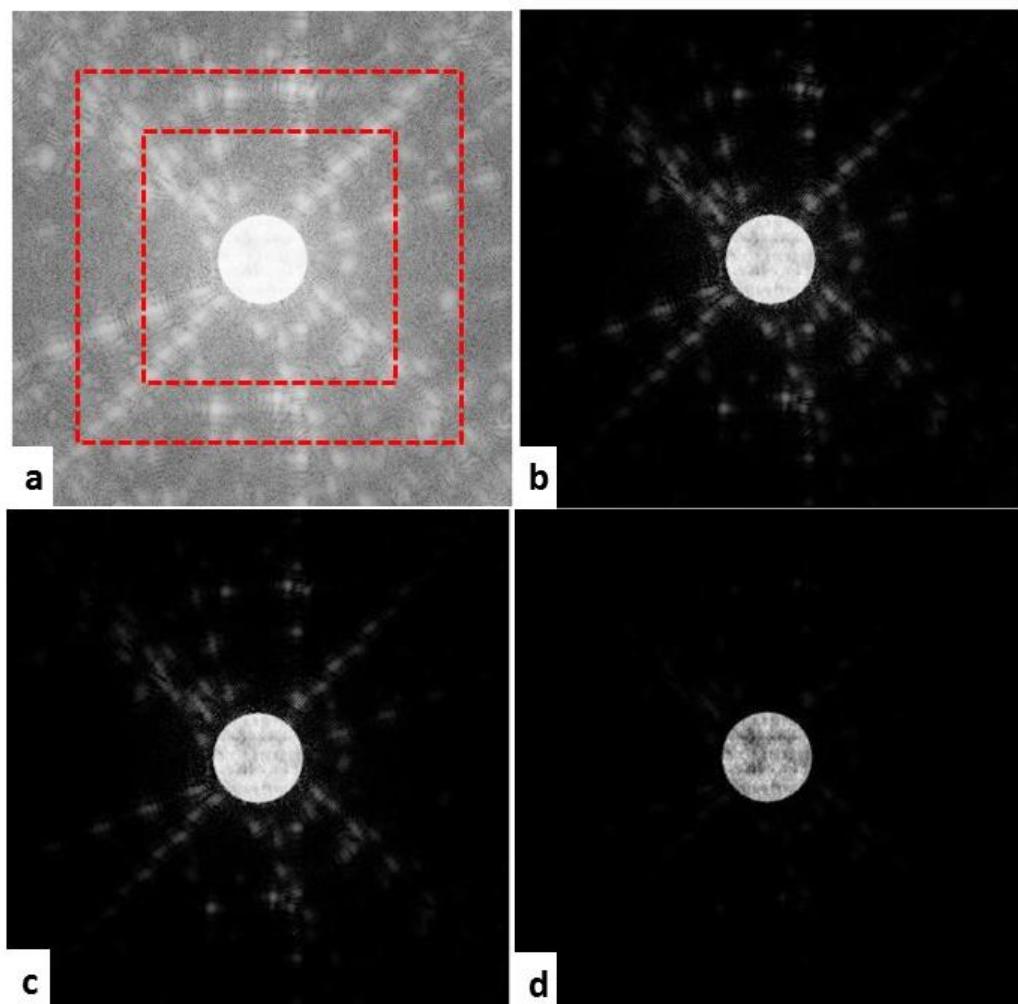


Figure 4.10 shows the examples of the diffraction pattern when adding the response pedestal; the diffraction patterns are plotted in logarithm scale. (a) plots the clean diffraction pattern without pedestal; (b - d) plot the same diffraction pattern shown in (a) when adding constant pedestals of 0.005, 0.01 and 0.5. The pedestal will emerge some high-order diffraction speckles at the edge part of the pattern; as we increase the pedestal, more speckles disappear. The red dashed lines in (a) show the data reconstructed in the later section that were truncated by $\frac{1}{2}$ or $\frac{3}{4}$; the data truncation will also result in the loss of the high-order diffraction speckles.

One experimental way to reduce the influence of the detector pedestal is to remove a reference from the image or diffraction pattern. Usually, after we set up the microscope ready before we collect the diffraction patterns, we find an empty area on the sample, then take an image as the reference. The reference will be removed from each diffraction pattern before

processing. Since the pedestal may vary with the excitation, this method cannot remove the pedestal accurately.

The other way to improve the quality of the data is to remove an artificial pedestal before the processing, if the dataset is very noisy, or even has some negative values. This method improves the reconstruction a lot when the dataset is very noisy. However, it cannot remove the pedestal accurately. Moreover, to find out the pedestal with a proper constant value, we may have to try several values if the constant value is set too high, we will lose some high-frequency information.

As has already been introduced in Chapter 2.7 and has been applied in Chapter 3, we have proposed two efficient methods to improve the ptychographic reconstruction with the pedestal data: the pedestal matching method (Maiden et al., 2015) and the extra mode method (Li et al., 2016). Both of the methods try to work out the background $c(u)$ to match the calculated intensities I_c at each pixel with the detected intensities I ; the difference between the two methods lies at the method of updating $c(u)$. The pedestal matching method updates the $c(u)$ at each scanning position by

$$c(j, u) = \gamma_j (I(j, u) - I_c(j, u)), \quad (4.9)$$

where j registers the probe positions; I_c is the calculated intensities; γ is the scaling factor update at each position by

$$\gamma_j = \frac{\sum_u [I(j, u) * I_c(j, u)]}{\sum_u [I(j, u) * I(j, u)]}. \quad (4.10)$$

The method to update the pedestal at each scanning position of the extra mode method is to set an extra probe function P_e and extra object function O_e , to make the calculated intensities match the detected intensities by

$$I = I_c + I_e, \quad (4.11)$$

where I_e is the intensity from the extra probe and extra object.

$$I_e = |\mathcal{F}\{P_e \times O_e\}|^2. \quad (4.12)$$

At each position, we update the pedestal I_e by updating the extra object function and the extra probe function by

$$O'_e = O_e + \alpha \frac{P_e^*}{|P_e|^2} (\psi'_e - \psi_e), \quad (4.13)$$

$$P'_e = P_e + \beta \frac{O_e^*}{|O_e|^2} (\psi'_e - \psi_e) \quad (4.14)$$

We retrieved the pedestal from the simulated data with both methods. Figure 4.11.b shows the reconstruction when removing the pedestal with the extra mode method; Figure 4.11.c shows the reconstruction when removing the pedestal with the pedestal matching method. Both of the reconstructions have been improved considerably; the features on the sample are clear, and the letters in the background are well reconstructed. Figure 4.12 shows the reconstructed pedestal at one position from the extra mode method. The retrieved value of the pedestal outside the central disc is about 0.005, which is exactly the pedestal we added up to the original diffraction patterns. However, in the central disc area, the retrieved value of the pedestal is about 0, which means that the algorithm failed to retrieve the pedestal inside the disc. The reason it resulted in this may be because the intensities of the diffraction pattern inside the central disc have an average value of 1.5, so when a pedestal of 0.005 is added it is 0.3% of the detected intensities, and so the algorithm is not sensitive enough to retrieve such a small error.

Now we can get the conclusion that the pedestal of the detector will make the ptychographic reconstruction very noisy, and will decrease the resolution. Both the pedestal matching

method and the extra mode method can retrieve the pedestal accurately, and improve the reconstruction efficiently.

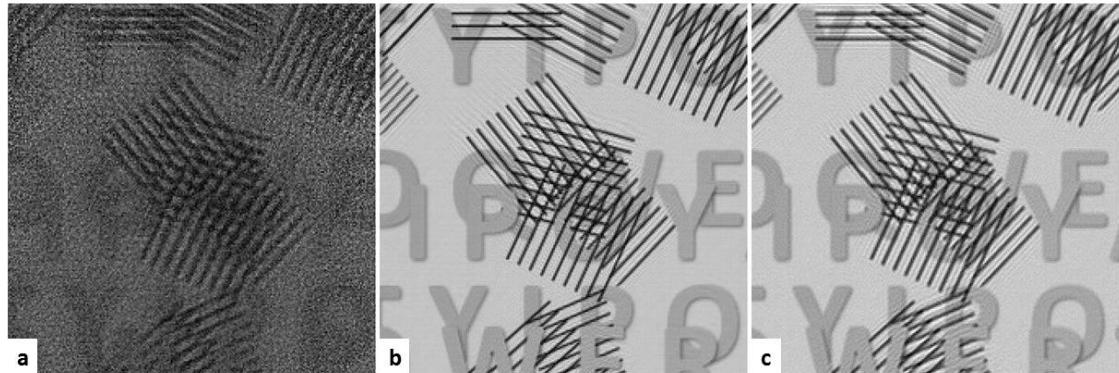


Figure 4.11 shows the reconstructions from the data shown in Figure 4.10.b. (a) shows the reconstruction with the conventional ePIE; the reconstruction is very noisy; the fringes are blurred, and the letters in the background have nearly disappeared. (b) shows the reconstruction with the extra mode method; (c) shows the reconstruction with the pedestal matching method. Both the two methods have improved the reconstruction quality a lot.

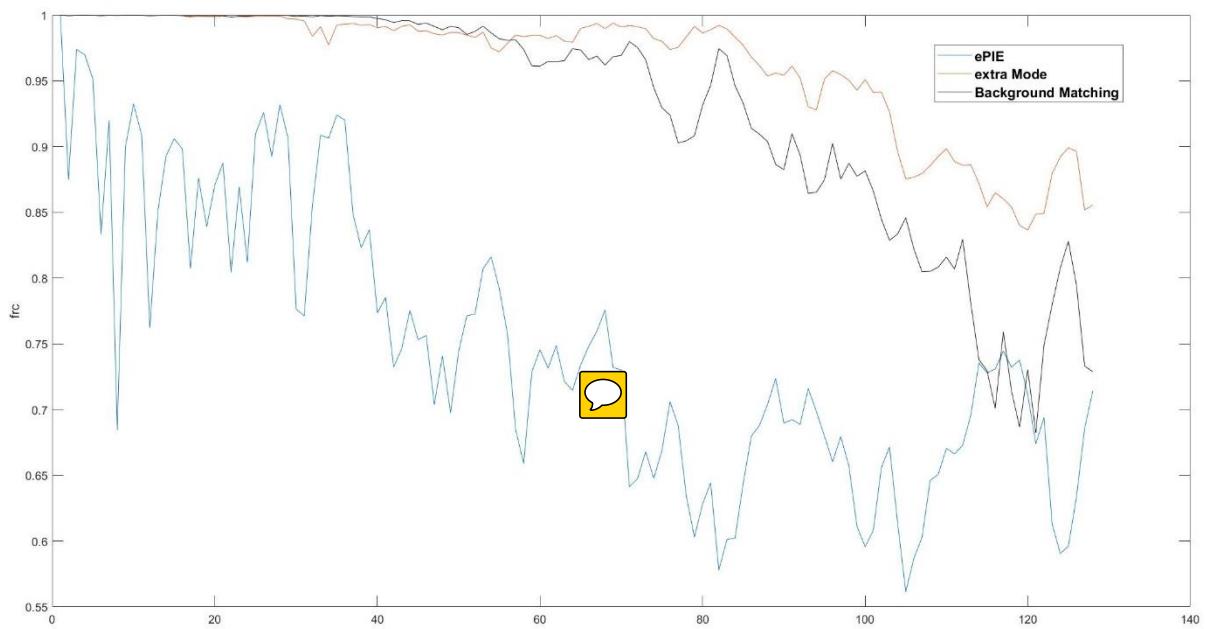


Figure 4.12. The blue line shows the FRC between the true object with the ePIE reconstruction; the red line shows the FRC between the true object with the extra mode reconstruction; the black line shows the FRC between the true object with the background matching reconstruction. We can see that with both

the extra mode method and the background matching method the reconstruction quality has been improved significantly. The extra mode method works better on the higher frequencies reconstruction.

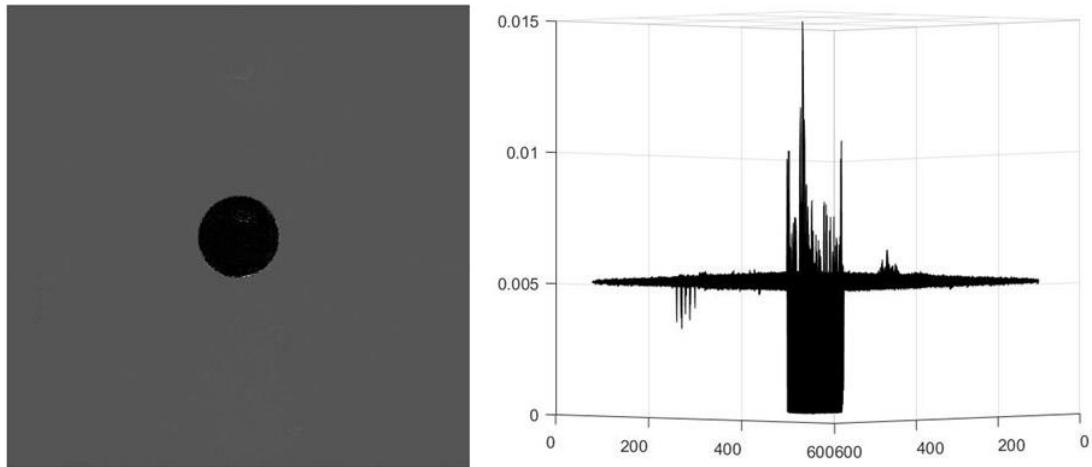


Figure 4.13 shows the retrieved detector pedestal at one position with the extra mode method; the right image plots the retrieved values of the pedestal; the values outside the central disc area are about 0.005, which have been retrieved accurately; the values inside the central disc area are nearly 0; which failed to be retrieved by the algorithm.

4.2.4 Data truncation

The experimental electron data processed for the previous result was not saved as the full detector size because of the limit of the computer memory. When the data is truncated artificially, some of the high-frequency information on the diffraction pattern will be lost. As shown in Figure 10.a, when the data was truncated as $\frac{3}{4}$ or $\frac{1}{2}$ of the original size, some speckles at the edge part will be lost. Since these speckles define the location of the crystals, we suspect that the truncated diffraction patterns may be another reason that resulted in the feature delocalization.

Figure 4.14.a shows the ePIE reconstruction from the half-truncated data; it is obvious that there are some artificial fringes lying outside the boundaries and that the features have lost resolution.

The method we applied to improve the reconstruction of the truncated data is called super resolution, which is a method that reconstructs beyond the diffraction limit resolution by extending the frame dimension of the diffraction patterns (Maiden et al., 2011). We can extend the frame size of the diffraction patterns in ptychographic reconstruction because of the redundancy in the dataset; more about this has been investigated in Chapter 6. In this simulation, we put the truncated data into a box with a larger size as shown in Figure 4.15: the 256*256 diffraction pattern was put in a 512*512 box. We use the extended diffraction patterns to do the reconstruction with ePIE. In the update with the Fourier domain constraint, we replace the modulus of the calculated intensities inside the original frame size area with the detected intensities I , while in the extended area, we leave the same as the calculation. This is the equation for the Fourier domain constraint update with the super resolution method:

$$\Psi'_j = M \cdot \sqrt{I} \frac{\Psi_j}{\sqrt{\sum_j |\Psi_j|^2}} + (1 - M) \cdot |\Psi_j| \frac{\Psi_j}{\sqrt{\sum_j |\Psi_j|^2}}, \quad (4.15)$$

where M is a 1-0 filter, the pass area of which is the original diffraction pattern area, which the block area is the extended area.

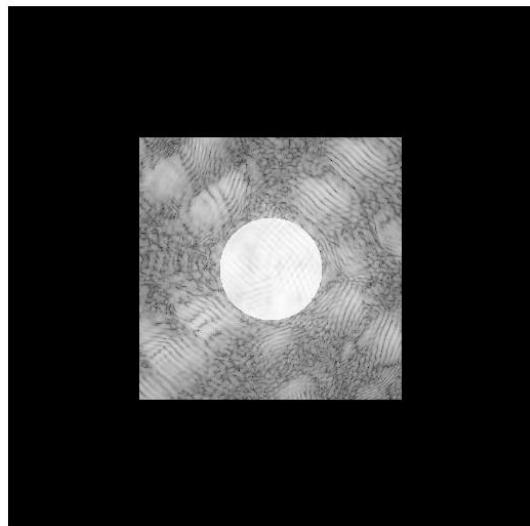


Figure 4.14 shows an example of the $\frac{1}{2}$ truncated diffraction pattern when embedding in a 512*512 box for the super resolution reconstruction.

The reconstruction from the $\frac{1}{2}$ truncated dataset with the super resolution method is shown in Figure 4.15.b. Compared with the result without extending the frame size in Figure 4.14.a, most of the artificial fringes have disappeared, and the features are sharp; the resolution has been improved. However, there are still some slight fringes outside the boundaries. The super resolution ratio σ_r , which is the extended data size divided by the original data size, is decided by the redundancy in the dataset.

In conclusion, data truncation will result in both the features delocalization and the resolution lost in the ptychographic reconstruction. The Super-Resolution method can improve the reconstruction because of the information redundancy in the dataset. However, there is a limit on the truncated area for the super resolution method. The efficiency of the super resolution method depends on how much information is lost and the overlap in real space, the illumination and sample structure, etc..

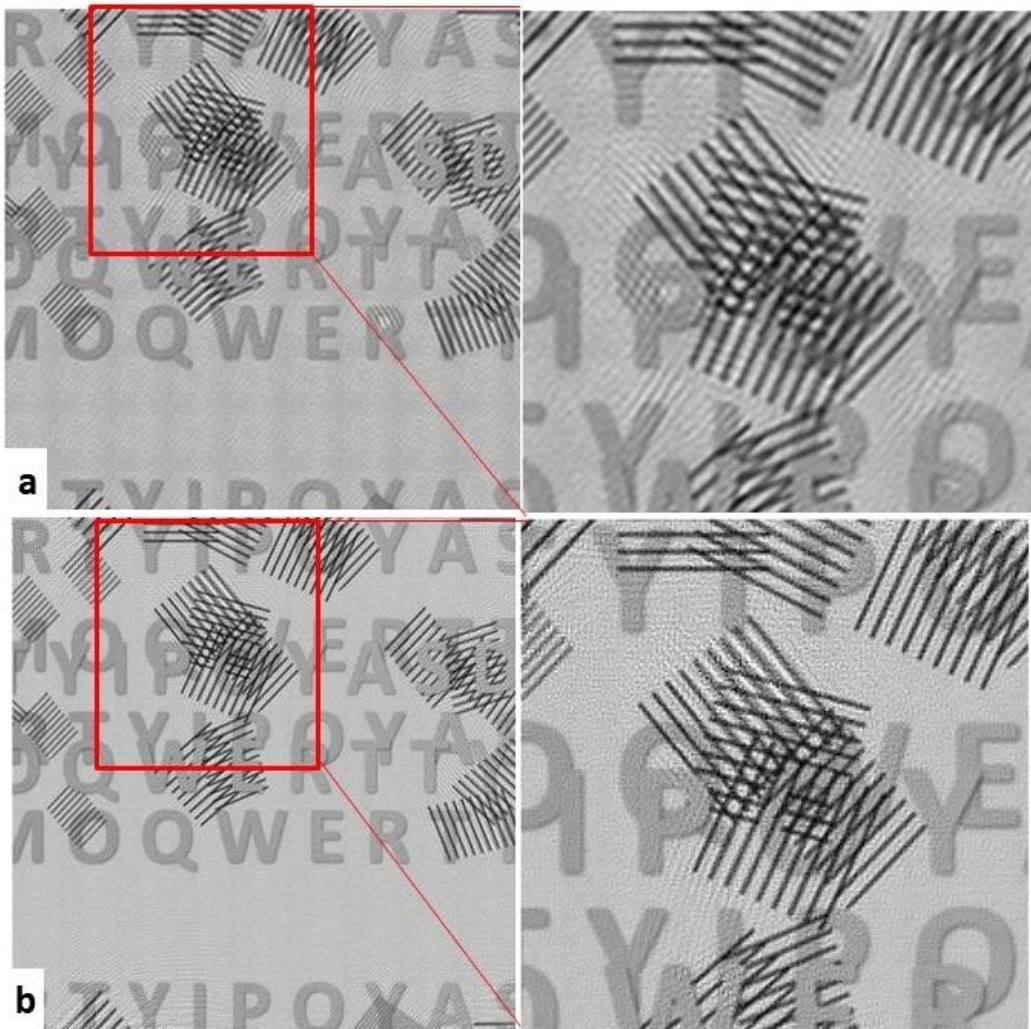


Figure 4.15 shows the reconstructions from the $\frac{1}{2}$ truncated dataset. (a) shows the result with the conventional ePIE; the resolution of the features is low, and there are some artificial fringes. (b) shows the reconstruction with the super resolution method; the features are sharper and most of the artificial fringes disappeared except that, there are still some slight fringes at the boundaries area.

4.2.5 Positions scaling error

The accuracy of the scanning positions is a critical factor that decides the quality of ptychographic reconstruction. Some investigations about the effect of scanning positions error on the ptychographic reconstruction have been done in the papers (Hue et al., 2011; Maiden et al., 2012; Zhang et al., 2013). We can categorise the scanning positions errors as two categories: random positions errors and global scaling error. Several methods which

work efficiently to correct the random positions errors have been proposed, for example, the annealing method jiggle (Maiden et al., 2012), and the exit wave cross correlation method (Zhang et al., 2013). The global positions error will be introduced when there is an error in the calibration of the shifting pitch of the sample stage or the illumination. For example, if a shifting pitch of 1nm is calibrated as 1.1nm, then when we command the stage to shift ten pitches, instead of shifting $10 \times 1 = 10$ nm, it will shift $10 \times 1.1 = 11$ nm. Thus, a 10% global scaling error will be introduced. In this section, we try to find out that, how the positions global scaling error will affect ptychographic reconstruction.

We reconstructed a set of ptychographic data with the scanning positions scaled by 1%, 3% and 10% respectively. Figure 4.16 plots the scaled scanning positions map. Figure 4.17 shows the reconstructions of the sample from the scaled positions with ePIE as well as the jiggle method. We can see that when there was 1% global scaling on the positions, the reconstruction lost the resolution slightly; when the error was 3% and 10%, the reconstructed samples look out of focus; the features were blurring and delocalized. Jiggle did not work well to correct the scaling error in the positions.

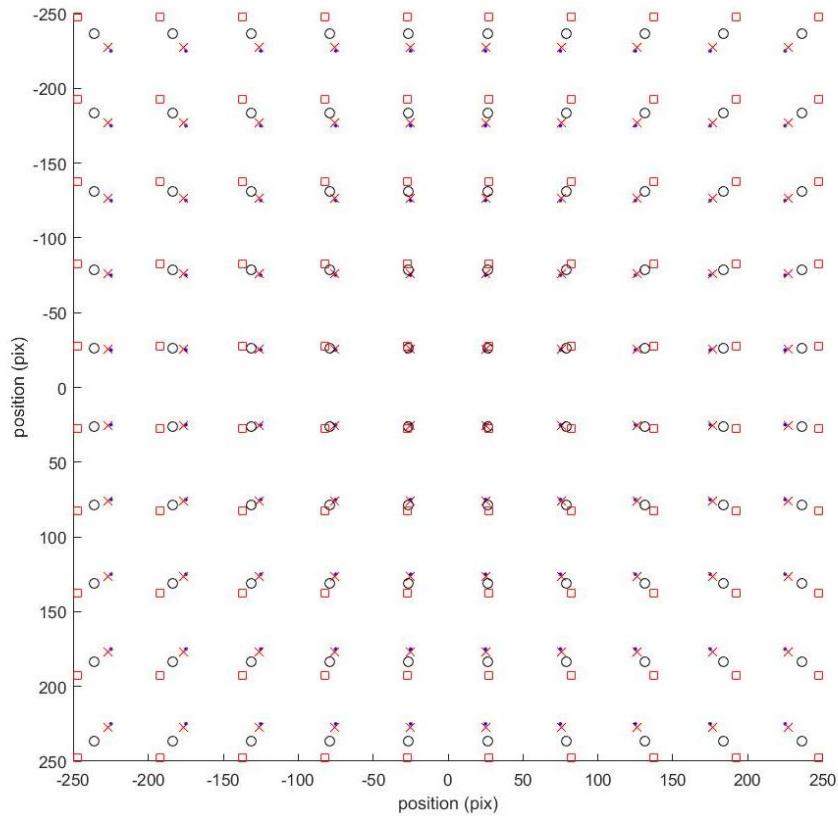


Figure 4.16 shows the scaled scanning positions. The black dots show the accurate positions; the red crosses show the positions when stretched by 1%; the black circles show the positions when stretched by 5%; the red squares show the positions when stretched by 10%.

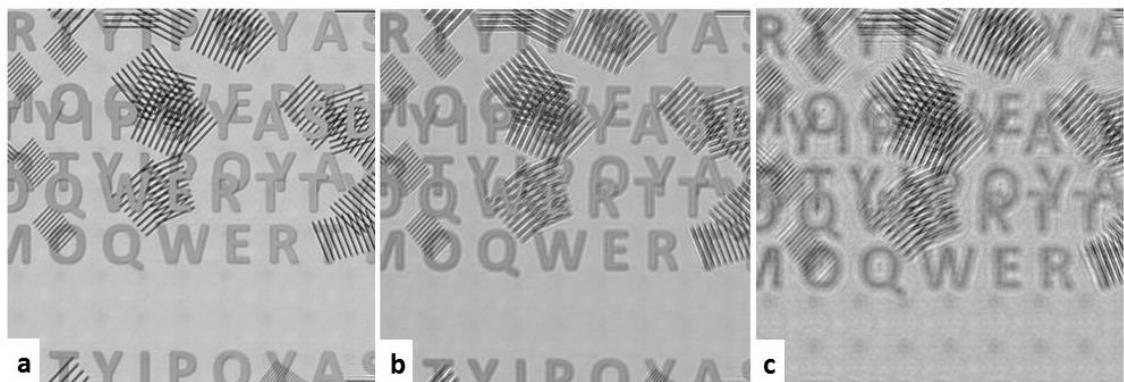


Figure 4.17 shows the reconstructions of the sample with the scaled positions. (a) shows the reconstruction when the positions are stretched by 1%; the features of the sample lost the resolution slightly. (b) shows the reconstruction when the positions are stretched by 3%; the features are very blurring. (c) shows the reconstruction when the positions are stretched by 10%; it appears like an out of focus image.

In conclusion, the global scaling error of the scanning positions will make the reconstructed sample look like an out of focus image; the features will be blurring and there will be no clear boundaries. Jiggle cannot correct the global scaling error in the positions.

4.3 Experimental result

In this section, we reprocessed the data from the paper (Humphry et al., 2012) with the algorithm, which is based on ePIE, having been improved with the modal decomposition method to decoherence and remove the detector point spread effect, detector pedestal corrector and super resolution. The contrast of the reconstruction has been improved, and the fringes are more localized.

The data was collected on an FEI Quanta 600 SEM with a thermally assisted Schottky field emission gun accelerated at 30 KeV (Humphry et al., 2012). The experimental setup can be simplified as Figure 4.4: the condenser system formed a defocused probe with a diameter about 15nm on the sample, which was gold particles on holly carbon film; a CCD detector was placed in the far field (at a distance about 0.127m), with a dimension of 2048*2048 pixels and each pixel 7.4um. The detector was binned by 2 to collect the diffraction patterns, so the frame size of the diffraction patterns is 1024*1024, which is truncated as 768*768 for saving the data because of the limit of the memory. 20*20 diffraction patterns were collected by shifting the illumination with a step size about 5nm. Figure 4.18 shows one of the diffraction patterns, which was truncated as 768*768 and embedded in a 1024*1024 box.

For the modal decomposition on this experimental data, we used 16 probe functions and one object function. 30 iterations of ePIE were run firstly to give the initial object function and probe function; then 16 probes were generated from the ePIE reconstructed probe, to continue the 800 iteration modal decomposition reconstruction. To retrieve the detector pedestal, we applied the extra mode method. Figure 4.19 compares the reconstruction with the improved algorithm with the previously published result; the contrast of the carbon film gets better, and the gold fringes are more localized. Figure 4.20 shows the decompositions of the partially coherent illumination involving the effect of the detector point spread; the modes were

propagated to the condenser aperture; The first mode takes a contribution of about 19%.Figure 4.21 shows the retrieved detector pedestal at one scanning position.

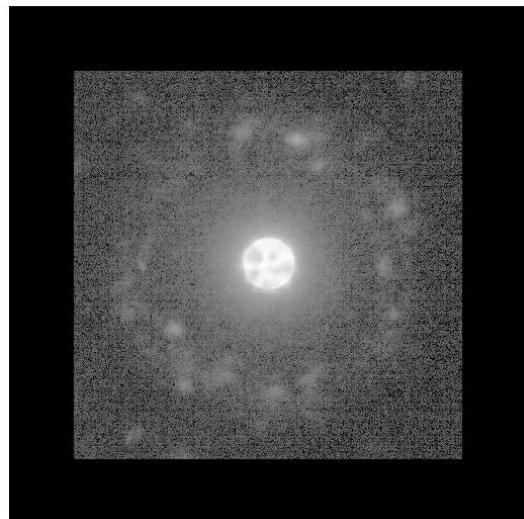


Figure 4.18 shows one diffraction pattern of the experimental data, which is plotted in logarithm scale; the data was truncated as 768*768; in the super resolution reconstruction, the truncated data is embedded in the frame size of 1024*1024.

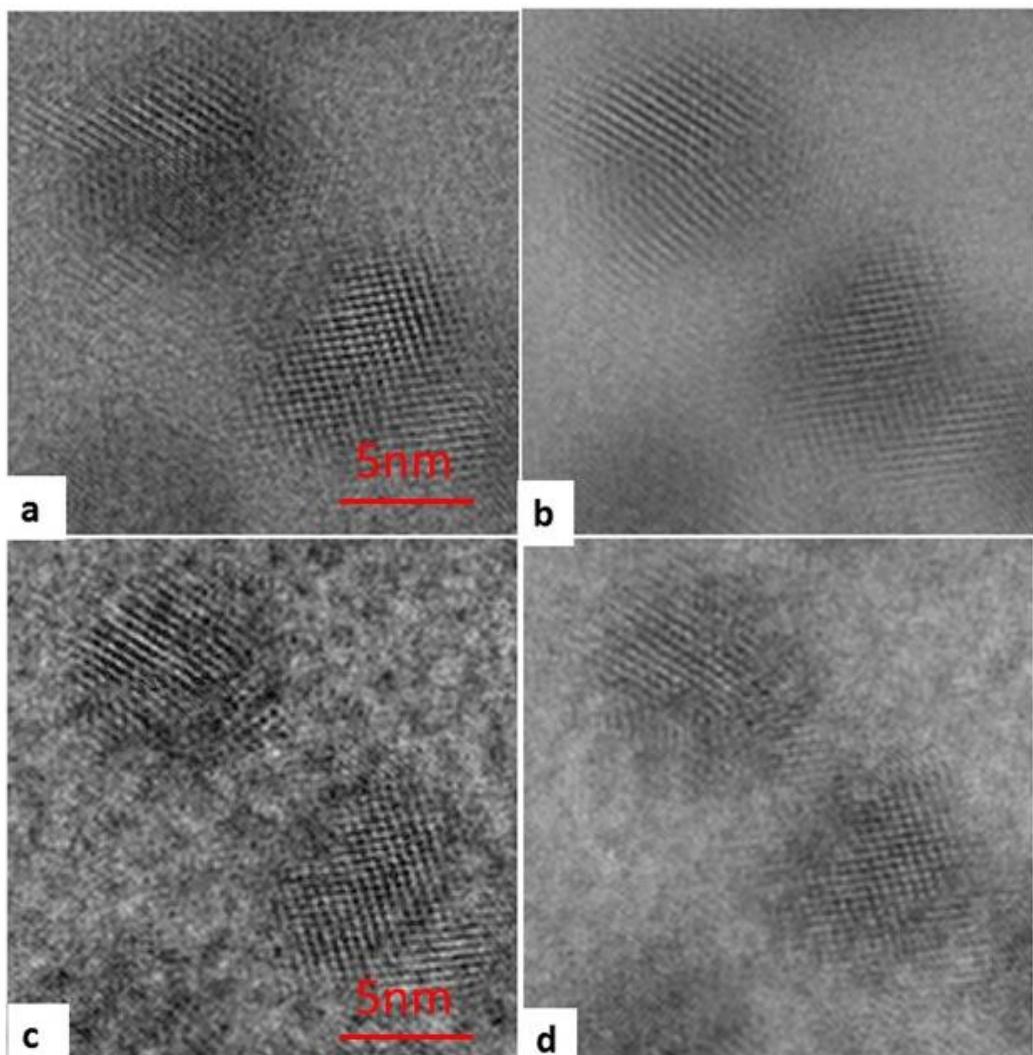


Figure 4.19 compares the reconstruction result from the improved algorithm with the previously published result. (a) and (b) show the modulus and phase of the reconstruction with ePIE; there are no clear boundaries of the gold particles. (c) and (d) show the modulus and phase with the improved algorithm; the contrast of the carbon film gets better, and the fringes are more localized inside the gold particles.

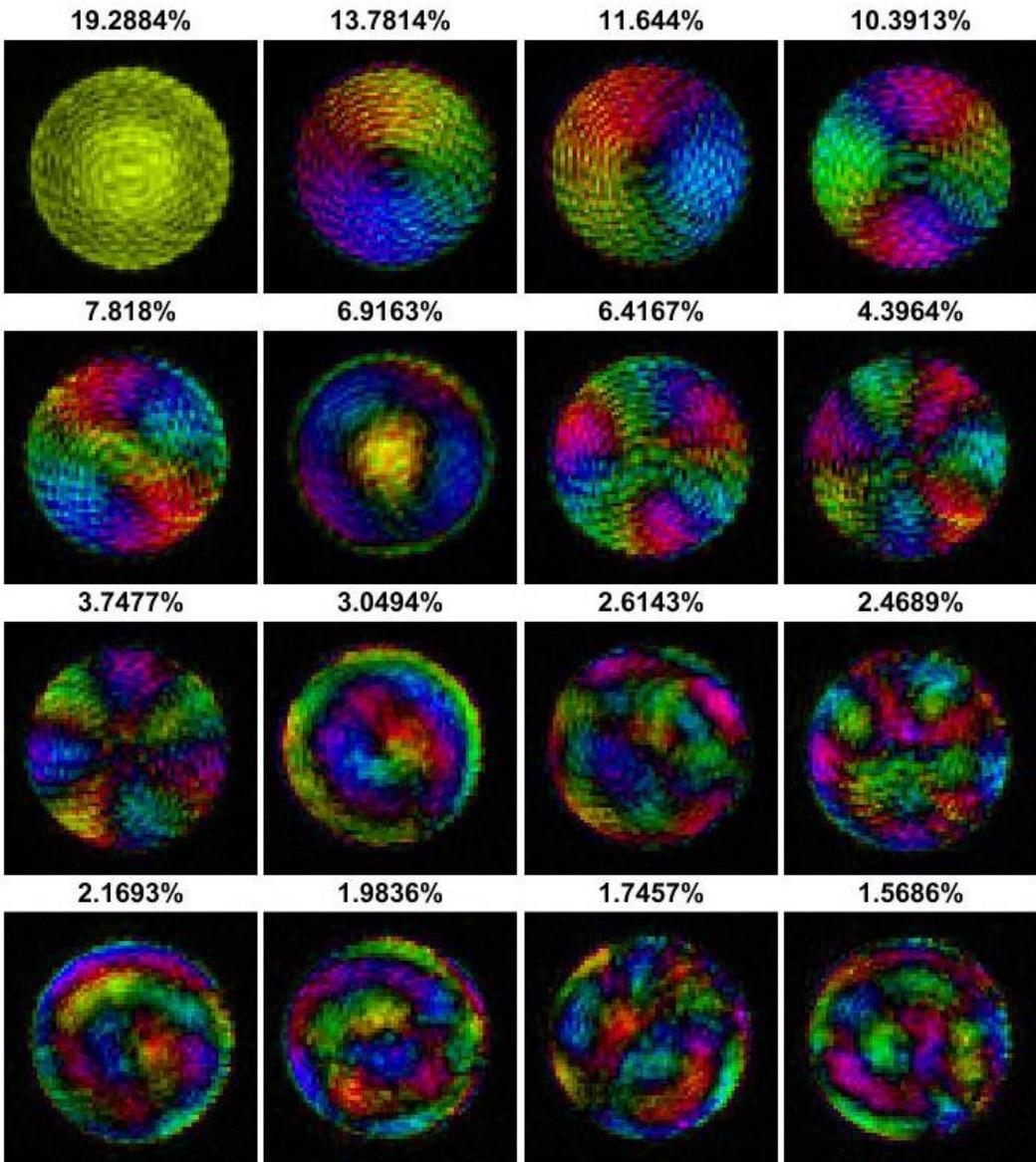


Figure 4.20 shows the retrieved modes of the effective electron source with the modal decomposition method; the modes were presented at the condenser aperture plane; the effect of the detector point spread was also included in these modes. The first mode takes about 19% contribution, for comparison, the first mode of a laser source takes about 90% contribution.

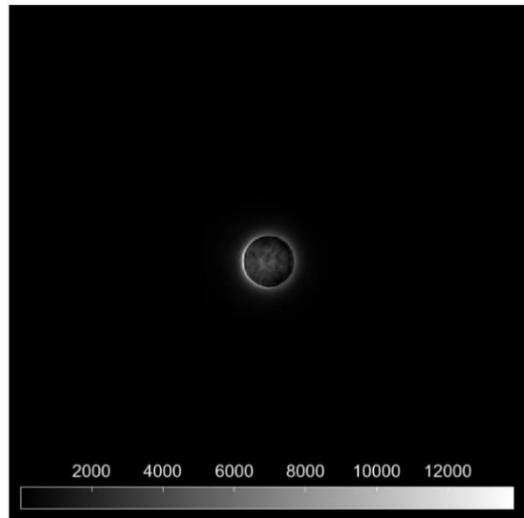


Figure 4.21 shows the retrieved detector pedestal at one scanning position from the experimental data with the extra mode method.

4.4 Conclusion

In this chapter, we diagnosed the delocalization problem existing in the previous reconstruction with the conventional ePIE algorithm. We found out that the partial coherence of the source, the detector point spread, the detector response pedestal, the data truncation, and the scanning positions global scaling error, will all result in the delocalization and the loss of resolution at a certain degree.

The scanning positions are the key factor that decides the resolution of the specimen reconstruction. Even a slight global scaling error in the scanning positions will result in the reconstructed object appearing out of focus. The transverse partial coherence of the source, as well as the effect of the detector point spread, is the second factor that affects the quality of the reconstruction. Both of them will result in the degrading of the contrast, the features blurring, and the fringes delocalization. The modal decomposition method works efficiently to decoherence and remove the effect of detector point spread, by decomposing the partially coherent illumination as well as the detector point spread effect into several modes. The next factor that affects the reconstruction quality is the detector response pedestal. The reconstruction from the dataset with detector pedestal will be very noisy; some features may completely disappear. There are several methods to reduce the effect of the detector pedestal before the processing, for example, removing a reference from each diffraction patterns when collecting the data, or taking off a constant pedestal from the whole dataset manually.

However, even though these methods do improve the reconstruction quality significantly, they are still not accurate enough, because the pedestal may vary between different pixels and it may change nonlinearly according to the counts of the electrons or photons. It is also possible to retrieve the pedestal from the reconstruction, since the rich redundancy in the ptychographic dataset. Two efficient methods, pedestal matching, and extra mode have been

proposed to retrieve the pedestal. We applied the extra mode method to reconstruct the pedestal in this chapter. Another factor that results in the delocalization in the result of Humphry et al. (2012) is the data truncation. Because of the limit of the memory, the original data, which was collected as 1024*1024, was truncated as 768*768 for storage. We investigated with the simulation and found out that, the data truncation will result in some artificial fringes in the reconstruction and the loss of the resolution. Super-resolution, which is a method to extract resolution beyond the diffraction limit of the data, has been proved that work well to reduce the effect of the truncation. But the efficiency of the super resolution method depends on the redundancy in the dataset, such as the truncated area, the overlap in the real space, the sample and illumination structure and so on.

We reprocessed the data in the paper (Humphry et al., 2012), with the ePIE algorithm improved with the modal decomposition method, super-resolution, and an extra mode. The reconstruction has been improved with better phase contrast of the light atoms (the carbon film), and the fringes are more localized inside the gold nanoparticles. Our improved algorithm relaxes the requirement for accurate knowledge of the experiment parameters, such as the illumination coherency, and the property of the detector. Further improvement on the algorithm to correct the global scaling error of the scanning positions is expected in the future work.

Chapter 5 Ptychography on a STEM

5.1 Introduction

In Chapter 3 and Chapter 4 of this thesis, we discussed the implementations of large probe ptychography in the selected area mode of a TEM (300 keV) and a SEM (30 keV). In this chapter, we will discuss the exploration of large probe ptychography on a STEM, which is accelerated by a 300 keV voltage and has condenser lens aberration corrector.

In the case of SAD ptychography, since the selected area aperture has to be used to localize the parallel illumination on the specimen, there is a limit to the minimum probe size; the minimum diameter of an electron aperture on a modern electron microscope is 10um.

Moreover, considering the short wavelength of electron waves and the setup of the TEM, it is impossible to place an aperture close to the specimen with a negligible propagation distance to localize the electron beams on the sample directly. In practice, we use the objective lens to form a magnified image of the specimen; the aperture is inserted at this first image plane to localize an area which is projected to the detector. In this case, what ptychography directly measures is not the specimen but the image of the specimen, which is the convolution of the specimen with the transfer function of the objective lens. With these limits, SAD ptychography is not a technique that aims for atomic resolution imaging.

The setup of the SEM in transmission mode, where a detector is installed behind the specimen, is ideal to perform ptychography – the condenser and objective lenses focus the electron beam on the specimen; there is no intermediate lens between the specimen and the detector. Thus, the quality of the diffraction patterns will not be affected by the transfer function of the imaging lenses. However, the accelerating voltage of a SEM is low, usually between 1KeV and 30 KeV. To implement ptychography on a STEM has its advantages. Firstly, by using the objective lens to form a defocused probe on the sample, it is easier to control the probe size. A modern STEM can form a finer probe than the SEM. The accelerating voltage of a STEM is usually between 80KeV and 300KeV.

There is some work that has already been done with ptychography on the STEM. We classify STEM ptychography as two categories. One is focused probe ptychography; the objective focuses the electron beams on the sample and scans across the sample densely; diffraction patterns are collected at each scanning position. Then the Wigner distribution deconvolution method is applied on the 4-dimensional dataset (the diffraction patterns have the information of the 2-dimensional real space and the 2-dimensional reciprocal space) to retrieve the complex object transmission function. Here we describe some of the focused probe STEM ptychographic experimental results as examples to compare with defocused probe STEM ptychography.

Resolution beyond the information limit was first obtained via focused probe ptychography on a STEM in 1995 (Nellist and Rodenburg, 1995). The specimen tested in ultra-thin crystal silicon (about 4.5 nm). The STEM was accelerated by 100KeV voltage; it had a conventional point resolution of 0.42 nm and conventional information limit of 0.33nm. The electron beams were focused on the specimen by the objective lens at a diameter of about 0.5nm. The focused probe was scanned across the specimen with a standard STEM raster scanning. At each scanning position, a diffraction pattern was collected with a CCD detector. A large

condenser aperture was inserted to make sure that the diffracted discs on the CBED patterns were overlapped with each other. Figure 5.1.a shows one of the diffraction patterns (Nellist and Rodenburg, 1995). The Wigner Distribution Deconvolution (WDD) method was used to reconstruct the dataset. 0.136 nm resolution was achieved in the result as shown in Figure 5.1.b.

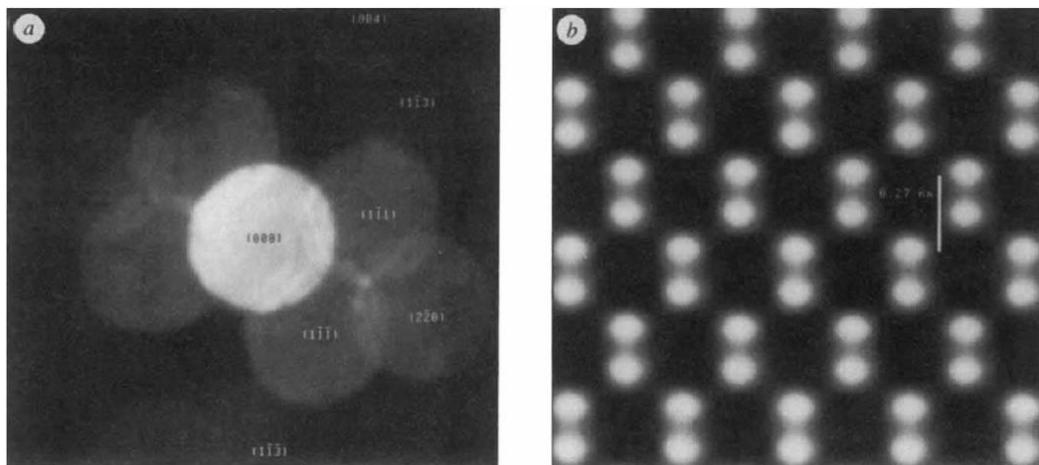


Figure 5.1 shows the focused probe ptychographic reconstruction by Nellist and Rodenburg in 1995. The specimen was crystalline silicon oriented on $<110>$. (a) is one example of a diffraction pattern; the bright central disc illustrates the size of the condenser aperture, which made sure the diffracted discs of different frequencies overlapped; (b) is the ptychographic reconstructed phase with Wigner distribution deconvolution method (WDD); 0.27nm resolution was achieved (Nellist, McCallum and Rodenburg, 1995).

Figure is reproduced with the permission of the publisher.

The application of focused probe STEM ptychography was limited then by the memory and computing capability of the computer. Recently, with a more powerful computer and a direct detector, rapid progress of focused probe STEM ptychography has been achieved. The main results and improvements are listed here. A recent paper by Yang et al. (2015) showed the atomic reconstruction of gold particles with an aberration-corrected WDD algorithm compared with the interference quantitative phase measurement (Figure 5.4); another (MacLaren et al., 2015) showed the reconstruction of a sample with both heavy and light

atoms; the reconstruction algorithm was aberration-corrected WDD; compared with a HAADF image, the advantage of ptychography is obvious: the light and heavy atoms were both reconstructed with ptychography, whereas HAADF can only image the heavy atoms (figure 5.5). Another paper by Yang et al. (2016) showed the Z-contrast reconstruction with the atomic resolution of carbon nanotube sample. The reconstruction algorithm was optical sectional WDD (Figure 5.6). Figure 5.7 shows a focused probe electron ptychographic reconstruction of a separated gold nanoparticle on a thin carbon support; the reconstructed resolution was even better than the ADF and ABF image (Yang et al., 2017).

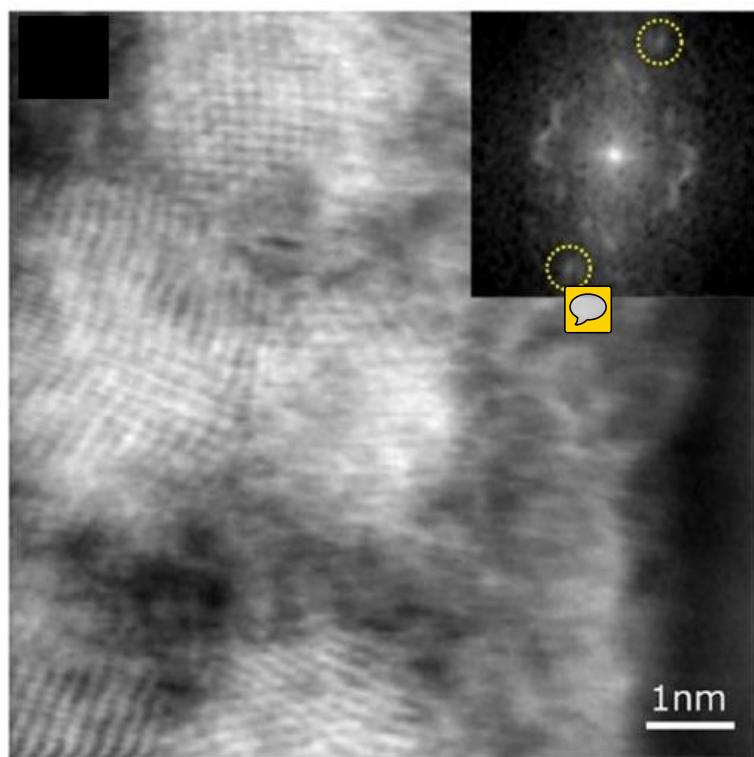


Figure 5.4 shows the reconstructed phase with focused probe ptychography (Yang et al., 2015). Figure is reproduced with the permission of the publisher.

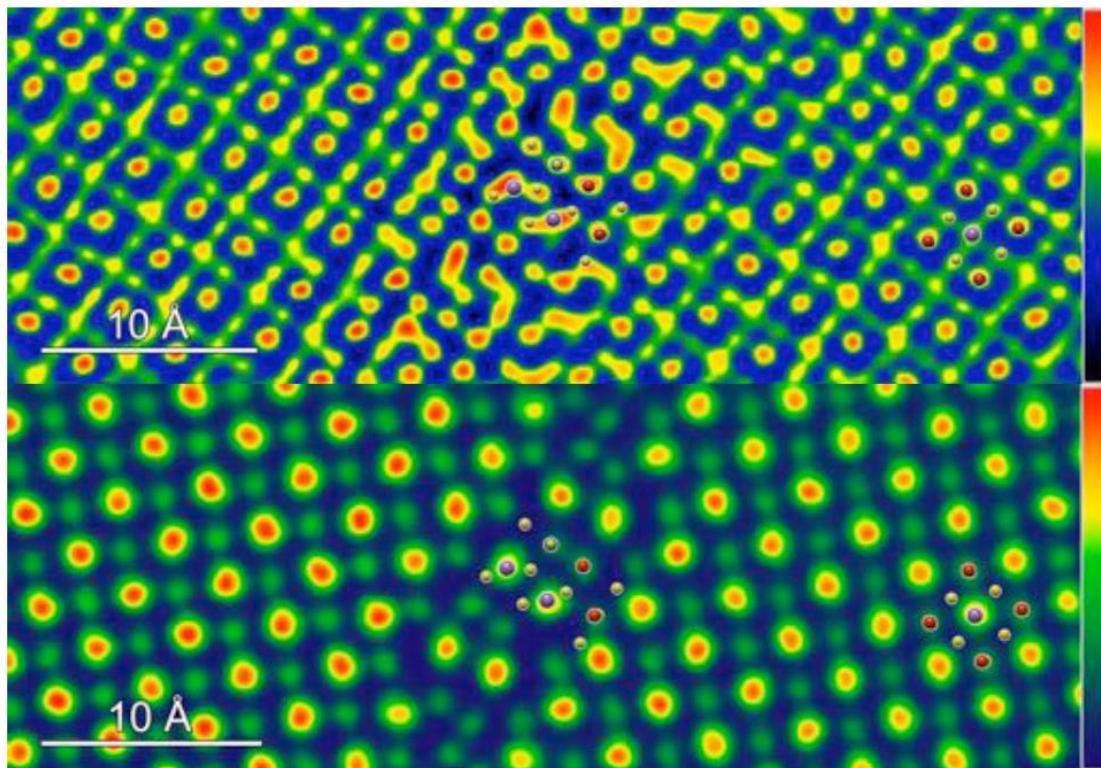


Figure 5.5 compares the focused probe electron ptychographic reconstruction (upper) with the HAADF image (bottom). The sample contained both light atoms and heavy atoms: the red shows iron atoms; the purple shows the bismuth atoms; the blue shows the titanium atoms and the yellow shows the oxygen atoms. Compared with the HAADF image, which only imaged the heavy atoms and missed the oxygen atoms, ptychographic reconstruction imaged both the heavy and the light atoms (McLaren et al., 2015).

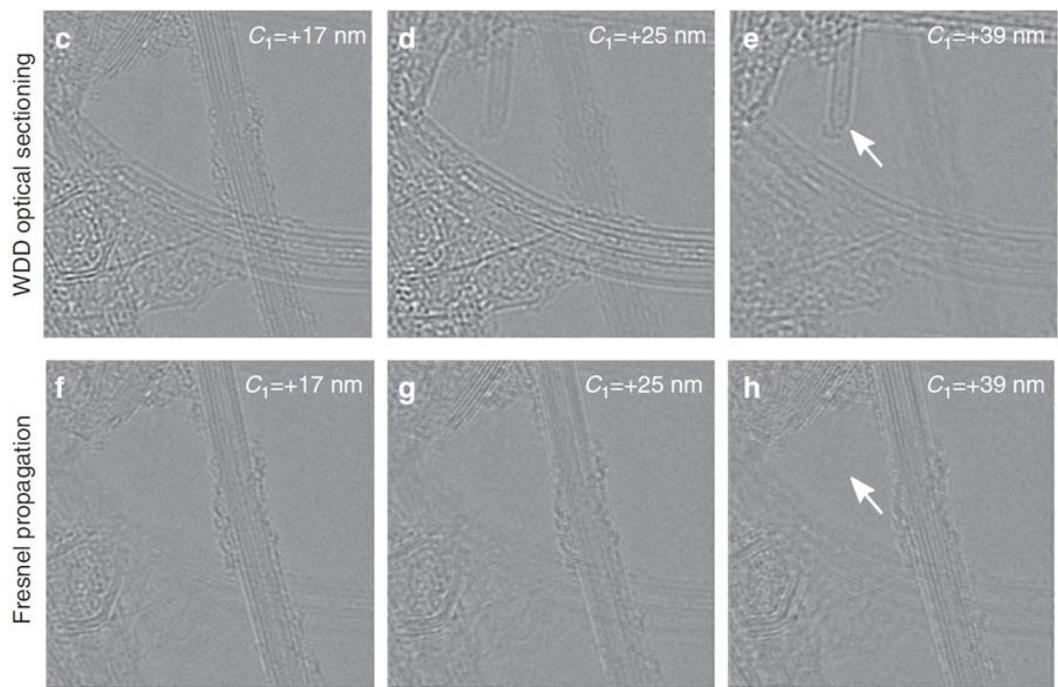


Figure 5.6. The focused probe electron ptychographic reconstruction with Z depth contrast (upper). With the optical sectional WDD algorithm, when focusing the probe at different focal depth (+17nm, +25nm, and +39nm), ptychography reconstructed the sample at different focuses. These optical sectional reconstructions are intrinsically different with the Fresnel propagation series (below) (Yang et al., 2016).

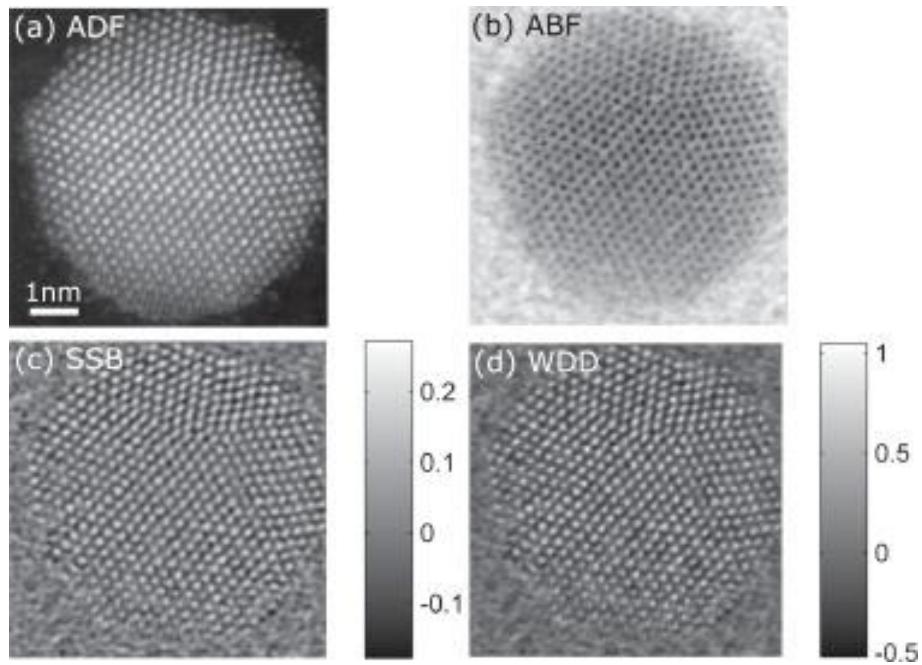


Figure 5.7 compares the focused probe electron ptychographic reconstruction of a gold nanoparticle on thin carbon support with its ADF and ABF images. The resolution of the reconstructions with SBB algorithm and WDD algorithm was even better than the ADF and ABF images (Yang et al., 2017). Figure is reproduced with the permission of the publisher.

Even though some impressive results have been achieved with focused probe STEM ptychography, there are some critical limits of its application. Firstly, the field of view of a focused-probe scan is small. In the STEM, when the electron beam is focused, the probe size is about several angstroms. The focused probe has to scan across the sample densely during a ptychographic data collection; this will be a huge dataset. For example, the reconstruction which in Figure 5.7 shows an area less than 5nm*5nm (a gold nanoparticle), but required 256*256 scanning positions. Secondly, when the electron beam is focused, the dose is high. Thus, there will be more damage to the sample. The high-intensity beam radiation limits focused probe ptychography only to be implemented with specific conditions, for example, extremely short exposure time, and samples able to stand the high electron dose without significant damage. The direct detector brings great convenience to focused probe ptychography. Its fast exposure frequency allows the STEM to collect the diffraction pattern

at each scanning position when we collect the HAADF image of the sample. Recent experiments (Yang et al., 2016 and 2017) were executed with a direct detector, which has an acquisition frequency of 20 kfps. Even so, the dose on the sample was still as high as $10^4/\text{\AA}^2$.

So far, there are only two publications of defocused probe STEM ptychographic results. The first atomic resolution obtained on the STEM (FEI 300KeV) by defocused probe ptychography was in 2011 (Putkunz et al., 2011). A Boron Nitride cones specimen was placed at 160 nm defocus to the beam focal plane; the probe diameter on the specimen was about 3.2nm. 5*4 CBED patterns were collected when the probe scanned over the specimen using a raster positions map and step size of 1 nm. Figure 5.2.b shows one of the diffraction patterns. The iterative algorithm DM (Thibault et al., 2009) was applied for the reconstruction, atomic resolution was achieved as shown in Figure 5.2.a.

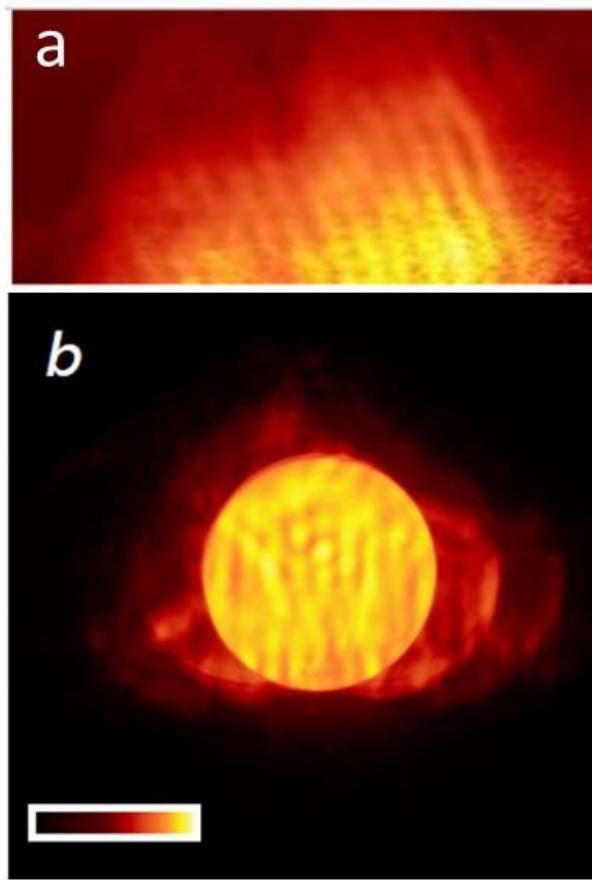


Figure 5.2 shows the defocused probe ptychographic reconstruction by Putkunz and Allen in 2011. The specimen was the tip area of BN cone, which has a lattice spacing about 0.2 nm. (a) shows the reconstructed phase. (b) shows the example of the diffraction patterns; the illuminated area was about 3.2 nm. The claimed resolution was 0.1 nm (Putkunz et al., 2011). Figure is reproduced with the permission of the publisher.

The next atomic resolution obtained through defocused probe ptychography on the STEM (JEOL 300KeV) was in 2014 (D'Alfonso et al., 2014). Cerium dioxide crystal sample was placed about 91 nm from the beam focal plane; the diameter of the probe was about 5 nm. 7*7 CBED patterns were collected in the far field behind the specimen by shifting the probe across the specimen with a raster scan, while making sure that the adjacent illuminated area has about 90% overlap. 

Figure 5.3.a shows one of the diffraction patterns. Atomic resolution was achieved with the iterative algorithm ePIE (Maiden et al., 2009), as well as an annealing method (Maiden et al.,

2012) to correct the random errors in the positions. Figure 5.3.b shows the reconstructed phase.

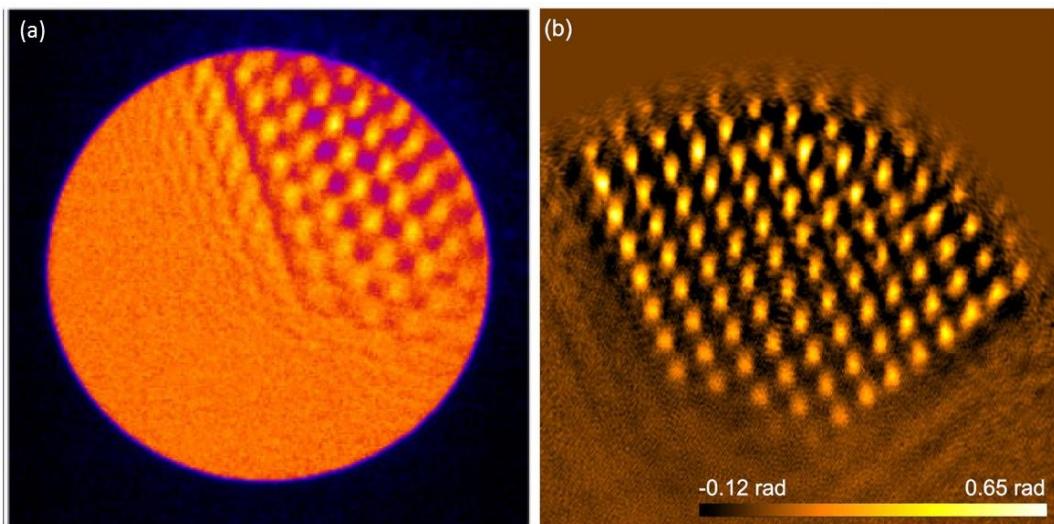


Figure 5.3 shows the defocused probe ptychographic reconstruction by Alfonso and Allen in 2014. The specimen was cerium dioxide orientated along the <110> direction. The lattice spacing between oxygen and cerium is 0.135 nm; the spacings between cerium atoms are 0.382 nm and 0.541 nm. (a) is one example of the diffraction patterns. (b) shows the reconstructed phase (D'Alfonso et al., 2014).

Even though reconstruction with atomic resolution has been reported in these two papers, we think that the results can be further improved from several aspects. Firstly, the reconstruction results in those two papers only used 7*7 diffraction patterns (Putkunz et al., 2011; D'Alfonso et al., 2014). The consequences of too few scanning positions for defocused probe STEM ptychography are a small field of view and a large error in the solution. If the measurements are fewer, for example, we only scan 2*2 positions, some part of the sample will only be measured once, there will not be enough knowns to solve the unknowns. Secondly, the reconstruction supposed the experimental conditions were ideal; for example, the electron source was assumed as fully spatially and temporally coherent; the scanning positions were supposed to be accurately measured, and the detector was assumed to give the true signal response.

In this project, we demonstrate defocused probe ptychography on the STEM (JEOL R005).

Compared with the previous two results, we make some improvements to both the experiment and the reconstruction algorithm. Experimentally, we shift the sample with a piezo stage; 15*15 diffraction patterns for each set of data. We also investigate the parameters of the setups to find out the optimal configurations. For the algorithm, we improve the conventional ePIE by considering the source partial coherence, the scanning positions error, and the detector response property. We also attempt to improve the algorithm for the reconstruction of noisy data.

In this chapter, I present the results that we have achieved in this project. The author has collected hundreds sets of data in this project; some of the datasets cannot reconstruct reasonable results because of the scanning positions were not correct. Of some of the datasets, we have successfully reconstructed both modulus and phase images of gold nanoparticles; the effective source of the STEM has been decomposed into a set of modes of the minimum representation. In this chapter, we will present two sets of reconstructions, which are the best reconstructions the author thinks she can get from different illumination conditions; one set of modes was decomposed from the electron beam with slight astigmatism; the other set of modes was decomposed from the electron beam with coma aberration. This is the first time that, the aberrated matter waves have been decomposed experimentally.

5.2 Experimental details

Two sets of data will be demonstrated in this chapter. The data was collected on a STEM (JEOL R005, accelerated by 300 KeV), which has a cold field emission gun as well as probe and image aberration correctors. Before the microscope was switched into STEM mode, it was aligned in TEM mode and several sets of bright field images were collected by shifting the specimen with a step size of 2-3 nm. By cross correlating these bright field images, we obtained a reference scanning positions map to test the step size accuracy and repeatability of the piezo stage. Figure 5.16 shows the scanning positions map. The microscope was then changed into STEM mode; the probe aberration corrector was used to correct the aberrations in the illumination. Figure 5.8 shows a CBED pattern that was used for correcting the aberrations. After the alignment, we refined the defocus of the objective lens manually to make sure that the illumination was focused on the specimen; then we collected the HAADF images of the sample at a range of magnifications. Figure 5.9 shows the HAADF images of the sample, which shows some gold particles on an amorphous carbon film. A condenser aperture with the diameter of 50um was then inserted to remove some high-order aberrations, producing a shadow image of the aperture on the detector. We change the strength of the objective lens to make the electron beam either defocused or over-focused. In theory, it is better to place the specimen at the defocused plane; at the defocused plane, the electrons beam is less spread than the over-focused plane because of the aberrations. This means that the beam is more localized on the sample at the defocused plane. There is no way to know the exact defocus distance in the experiment.

Figure 5.10 shows a defocused CBED pattern. The diameter of the illuminated area can be estimated from the structures inside the central disc of the CBED pattern; for example, in Figure 5.10, there are about eight gold fringes inside the disc. We know the fringe spacing is 0.204 nm, then the illumination diameter was about 2 nm. The specimen was shifted on the

piezo stage by 15*15 raster scan according to the reference map that had already been tested previously in TEM mode. The camera length could be changed in the range 8cm – 30cm. Considering the counts of the diffraction patterns and the sampling conditions in the Fourier domain, we chose a camera length of 12 cm, which is the readout value of the microscope and needs to be calibrated for the reconstruction. The detector used for collecting the diffraction patterns was the Gatan Ultrascan 1000, which has a dimension of 2048*2048 pixels, with each pixel 7.4um; the detector was binned by two during the acquisition.

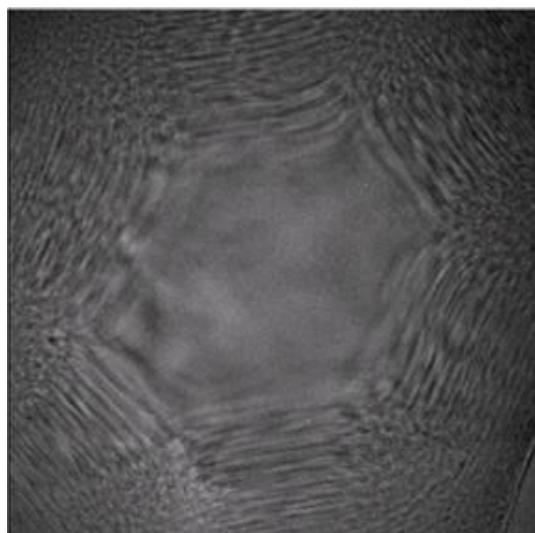


Figure 5.8 shows one CBED pattern that was used to correct the aberrations by the probe aberration corrector.

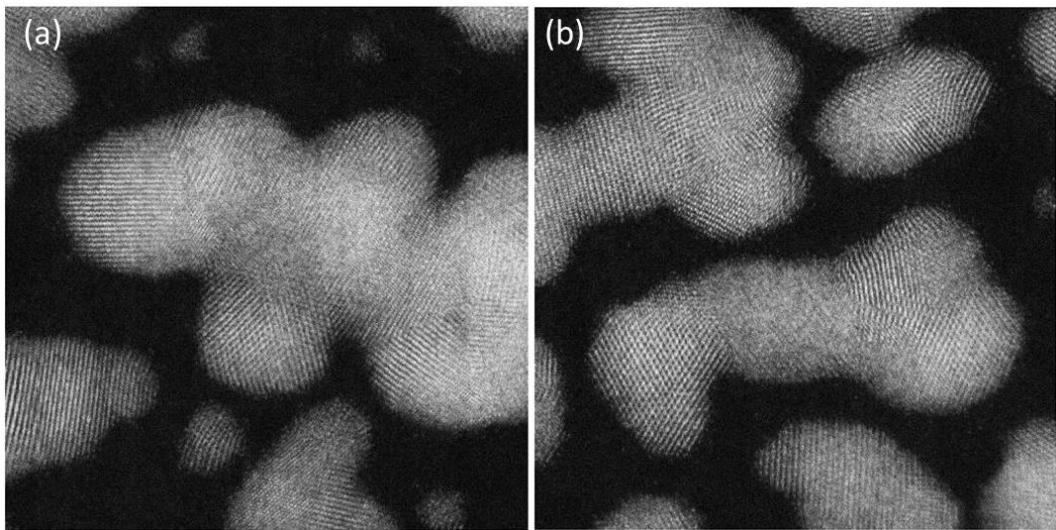


Figure 5.9 shows the HAADF image of the specimen. There are some gold nanoparticles stacked together in this area. The lattice spacings of gold particles are 0.204 nm, 0.143 nm and 0.102 nm. (a) is at higher magnification and (b) is at slightly lower magnification.

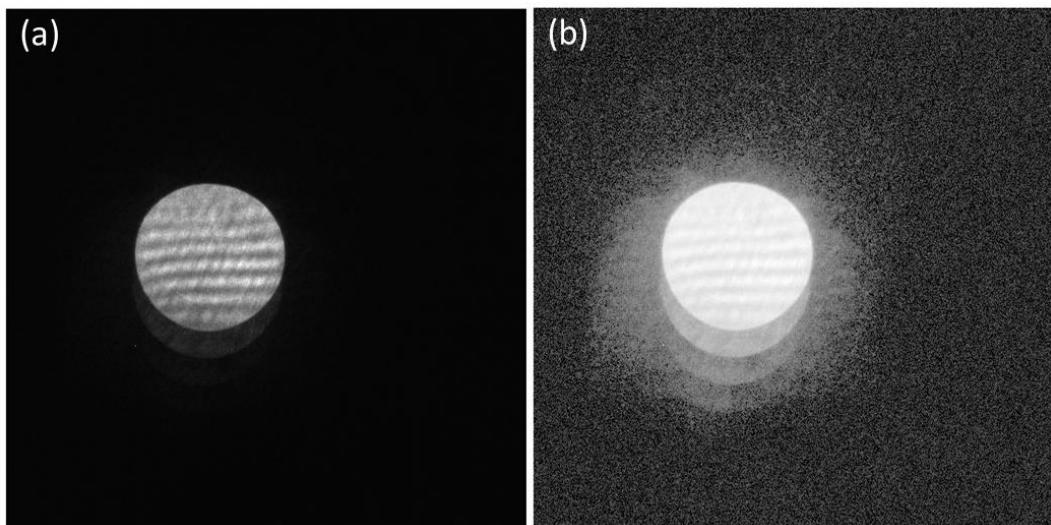


Figure 5.10 shows one example of the CBED patterns, which were collected at a defocused plane. (a) is the original diffraction pattern and (b) is plotted in logarithmic scale. We can tell that there are about eight fringes inside the central disc; we know the spacing is 0.204 nm; we can estimate that the diameter of the illumination area on the specimen was about 2 nm.

There are several parameters on the STEM that need to be investigated. By using different parameters, we collect CBED patterns with different configurations. One parameter is the condenser aperture size. The same probe size can be formed by either using a large condenser

aperture at a closer defocused plane, or using a smaller condenser aperture at a further defocused plane. Figure 5.12 shows different kinds of CBED patterns collected on the R005, of which, dp1, dp3, and dp4 were collected with a larger condenser aperture (dp1 was closer to the focal plane, and dp4 was further from the focal plane); dp2 and dp5 were collected with a smaller condenser aperture (dp2 was further to the focal plane, and dp5 was closer to the focal plane). The illuminated area of dp2 and dp4 was about the same.

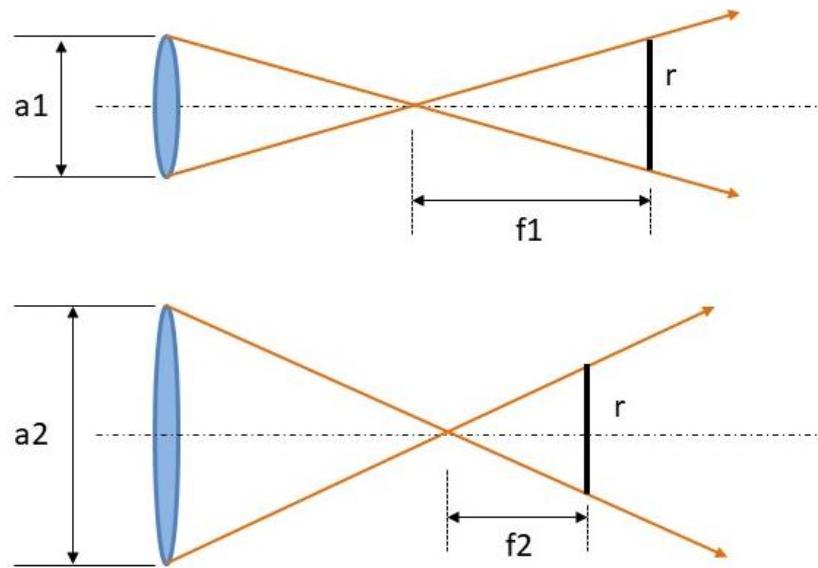


Figure 5.11 shows two configurations, in both illumination of diameter of r is formed. However, the top configuration has a smaller convergence angle a_1 and a large defocus f_1 , while the bottom configuration has a larger convergence angle a_2 and a shorter defocus f_2 .

Figure 5.13 shows the examples of the two sets of diffraction patterns that are processed in this section. Both sets of the data were collected with a small condenser aperture but at different defocused planes. Dataset 1 (5.13.a) was collected at a further defocused plane; the illumination size is relatively larger (about 10 nm). The counts of this dataset are higher – the highest count in a pixel is about 6000. Dataset 2 (5.13.b) was collected at a closer defocused plane; the illumination size is smaller (about 5 nm). The counts of this dataset are lower; the highest count in a pixel is about 700.

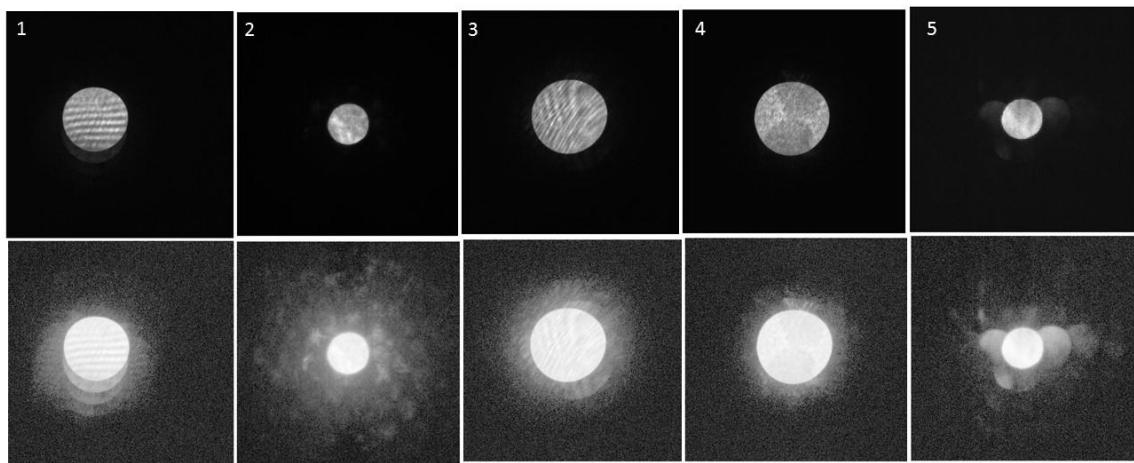


Figure 5.12 shows various convergent beam electron diffraction patterns that can be used for ptychographic reconstruction. dp1, dp3, and dp4 were collected with a large condenser aperture, but at different defocused planes. dp2 and dp5 were collected with a small condenser aperture at different defocused planes. The illuminated area in dp2 and dp4 was about the same.

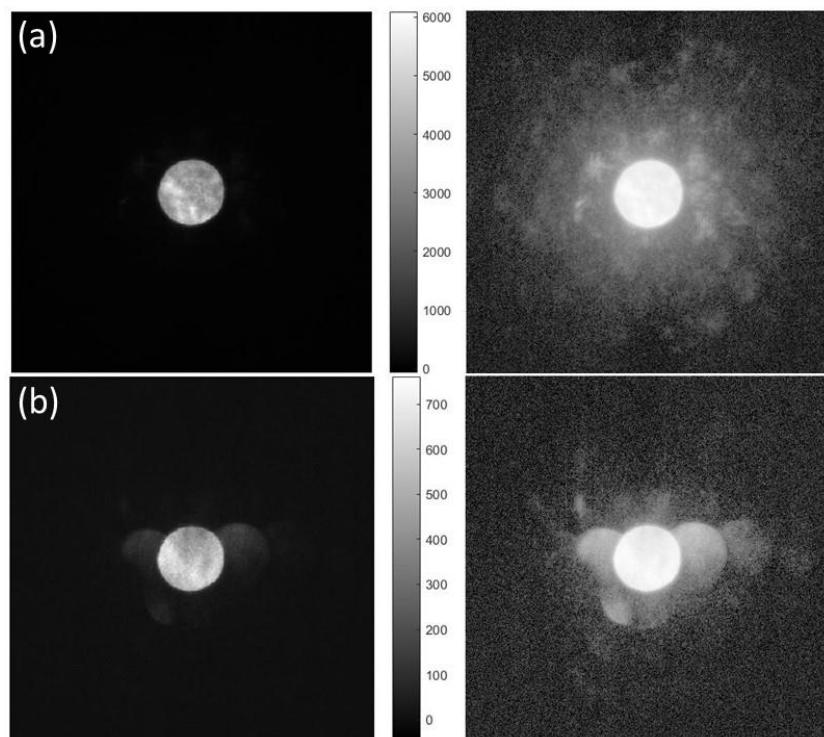


Figure 5.13 shows the examples of the two sets of diffraction patterns that are processed in this chapter. Both datasets were collected with a small condenser aperture. (a) was collected at a further defocused plane; the illumination diameter was about 10 nm; (b) was the at a closer defocus plane; the illumination diameter was about 5nm. (a) had more counts than (b).

5.3 Reconstruction and results

Before the reconstruction, we calibrate the camera length and the scanning positions.

5.3.1 Camera length calibration

The method to calibrate the camera length in the STEM is similar to the method of calibrating the camera length in SAD mode (Chapter 3). Both are based on the equation that

$$L = Rd/\lambda, \quad (5.1)$$

where λ is the wavelength; d is the lattice spacing of the sample; L is the calibrated camera length; R is the diameter of the diffraction rings of a SAD pattern. In the convergent beam electron diffraction (CBED) pattern, R is the diameter that is measured from the centre of the diffracted discs, as shown in Figure 5.14 and Figure 5.15.

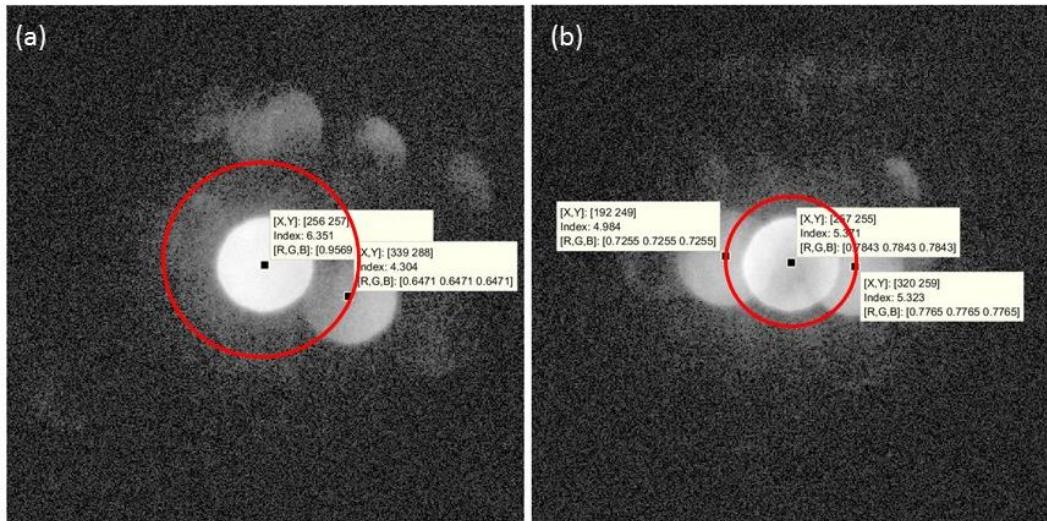


Figure 5.14 shows the examples of the diffraction patterns that were used to calibrate the camera length of dataset 1.

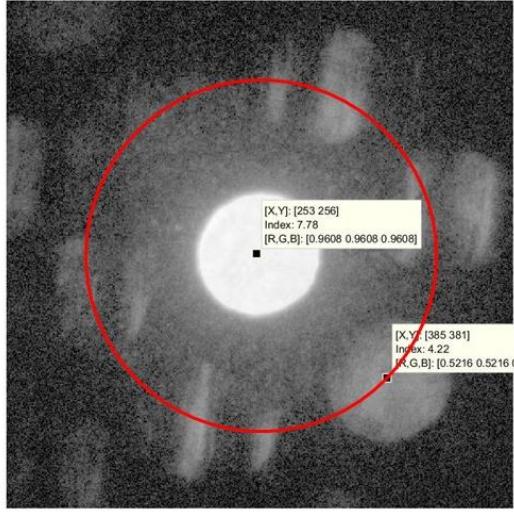


Figure 5.15 shows one example of the diffraction patterns that were used to calibrate the camera length of dataset 2.

The diameter of the diffracted rings from (a) is that

$$R_1 = 2\sqrt{(256 - 339)^2 + (257 - 288)^2} \cdot \Delta u, \quad (5.2)$$

$\Delta u = 7.4 * 2 = 14.8 \text{ } \mu\text{m}$, which is the pixel size of the detector when binned by two.

The calibrated camera length from diffraction pattern (a) is that

$$L1 = \frac{R_1 d}{\lambda} = 18.87 \text{ cm}. \quad (5.3)$$

The diameter of the diffracted rings from (b) is that

$$R_2 = 2\sqrt{(257 - 320)^2 + (255 - 259)^2} \cdot \Delta u; \quad (5.4)$$

$$R_3 = 2\sqrt{(257 - 192)^2 + (255 - 249)^2} \cdot \Delta u; \quad (5.5)$$

The calibrated camera length from diffraction pattern (b) is that

$$L2 = \frac{(R_2 + R_3)d}{2\lambda} = 19.68 \text{ cm}. \quad (5.6)$$

The camera length we applied in the reconstruction of the dataset 1 was 19.68cm.

The way to calibrate the camera length of dataset 2 is similar, as below

$$R = 2\sqrt{(253 - 385)^2 + (256 - 381)^2} \cdot \Delta u, \quad (5.7)$$

$$L = \frac{Rd}{\lambda} = 23.3 \text{ cm}. \quad (5.8)$$

5.3.2 Scanning positions examination

A set of bright field images were collected in TEM mode at the magnification of 100k when the probe scanned across the sample using the same positions map designed for collecting the diffraction patterns. Each pixel of the bright field image has a physical dimension of 0.2 nm.

We can work out the shifting positions map in *nm* by

$$P_{nm} = P_{pixel} * 0.2 \text{ nm}. \quad (5.9)$$

The other way to find out the scanning positions map is to cut out a small central area in the central disc of each CBED pattern of the dataset (Hurst et al., 2009), then cross correlate these central cut-out parts to find out their relative movements. One of the disadvantages of this method is that, to make sure there are enough structures inside the central disc of the CBED pattern, a large condenser aperture has to be inserted, and the specimen has to be placed at a plane with large defocus. The other disadvantage is that, even though we can work out the scanning positions map accurately by cross correlation the central discs of CBED patterns, it is difficult to decide the pixel dimension at the diffraction pattern plane.

5.3.3 Initial probe generation

Even though ptychography can retrieve both the complex specimen transmission function and the probe function, a relatively accurate initial probe function is essential for the success of defocused probe ptychographic reconstruction. The effects of an initial guessed probe

function at the wrong defocused plane, as well as the effects of the global scaling error in the positions on ptychographic reconstruction, will be discussed later in this chapter. Here we give the process on how to generate the initial probe function before the reconstruction:

- 1) From the CBED pattern, the subtended angle of the condenser aperture on the sample can be worked out, which is the diameter of the central disc;
- 2) From the structures inside the central disc of the CBED pattern, the size of the illuminated area can be estimated;
- 3) A top-hat profile of the same diameter (or slightly larger) as the central disc can be generated (Figure 5.16.a); a specific phase curvature is then added the top-hat function, to make sure that the Fourier transform of it has the similar size with the illuminated area. Usually, there will be an ambiguity when choosing the phase curvature if we do not know whether the illumination is at a defocused plane or over focused plane; thus, we have to try both directions in the reconstruction. Usually, after several iterations, we can see there is a big difference in the quality of the object reconstruction between the two probes with the different propagating directions.

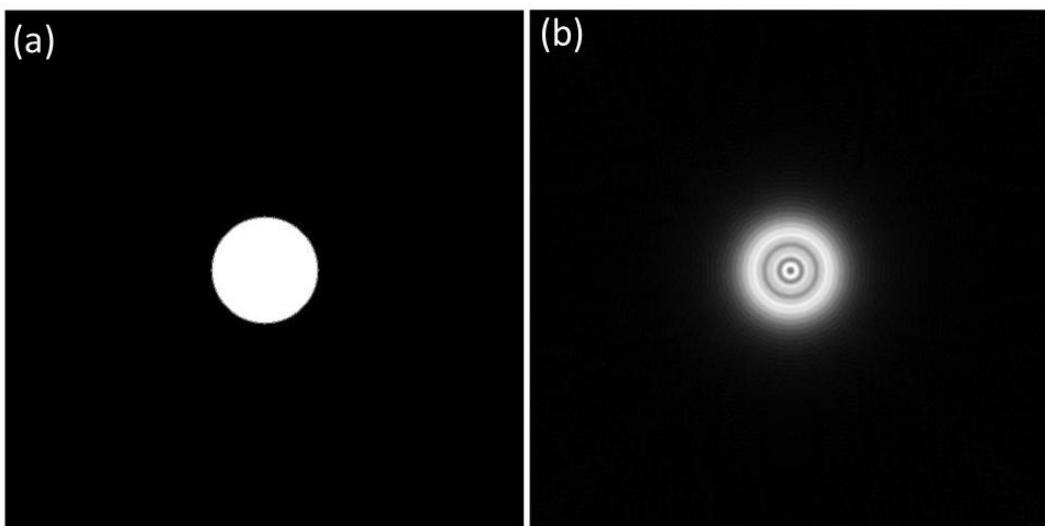


Figure 5.16 shows the initial aperture (a) and probe (b) that were generated before the reconstruction of dataset 2. The aperture has a similar size with the central disc of the CBED pattern. The illumination has a diameter of about 8 nm.

5.3.4 Algorithms and reconstruction

The algorithm applied for the reconstruction is based on the iterative algorithm ePIE (Maiden and Rodenburg, 2009), with the following improvements:

- 1) An annealing method ('jiggle') to correct the random errors in the individual positions (Maiden et al., 2012). The idea and calculation procedure of 'jiggle' was described in Chapter 2.7. Figure 5.17 shows the corrected positions map, which is marked as the blue dots; the original positions map is marked as the red crosses; at each position, another three random positions were generated within the initial radius of two pixels; annealing search began at the 20th iteration and ended at 20 iterations before the reconstruction finished. In Figure 5.17 we can see that quite a few original positions with minor errors were corrected.

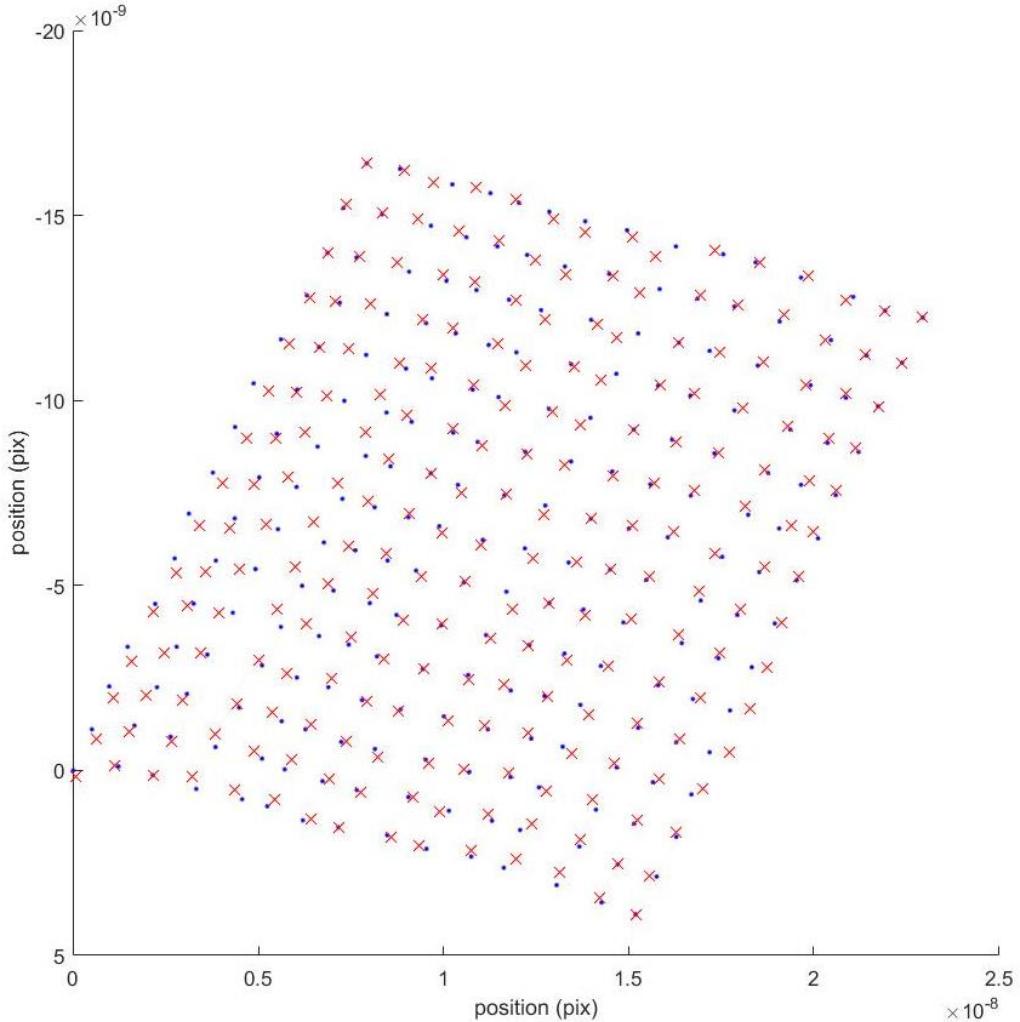


Figure 5.17 The red crosses plot the initial scanning positions used at the beginning of the reconstruction.

The blue spots plot the positions corrected by jiggle at the end of the reconstruction.

- 2) A modal decomposition method (Thibault et al., 2014; Li et al., 2016) was used to retrieve the partially coherent illumination function as well as the detector point spread. We applied the conventional ePIE algorithm first, which used a single probe function and a single object function for 50 iterations to generate a probe function and an object function; then we generate 16 probe functions from this probe function to do the modal decompositions reconstruction. These 16 probe functions can be the same

as the ePIE reconstructed probe function, or can be the linear transforms of the ePIE reconstructed probe function.

- 3) A detector pedestal (Maiden et al., 2015; Li et al., 2016); in the reconstruction, we used an extra mode method (Chapter 2.7 and Chapter 3) to retrieve the detector pedestal from each diffraction pattern.
- 4) Diffraction pattern drift was corrected by cross correlating the detected pattern with the calculated pattern at each position (Chapter 2.7).

Figure 5.18 shows the specimen reconstruction of the dataset 1; we can see the gold lattice fringes. Figure 5.19 shows the decomposed eigenstates of the partially coherent illumination function when propagated to the condenser aperture plane. The contribution of the first eigenstate is about 20%. During the propagation of the probe function, we found astigmatism in the illumination function. Figure 5.20 shows the propagated series of the first mode, which shows astigmatism in the illumination. Figure 5.21 shows the effective source profile by adding up the intensities of the 16 eigenstates when propagated to the focal plane; the effective source profile is the convolution of the demagnified source with the condenser aperture; it has a size of about 1.3 nm in one direction and 0.97 nm in the other direction.

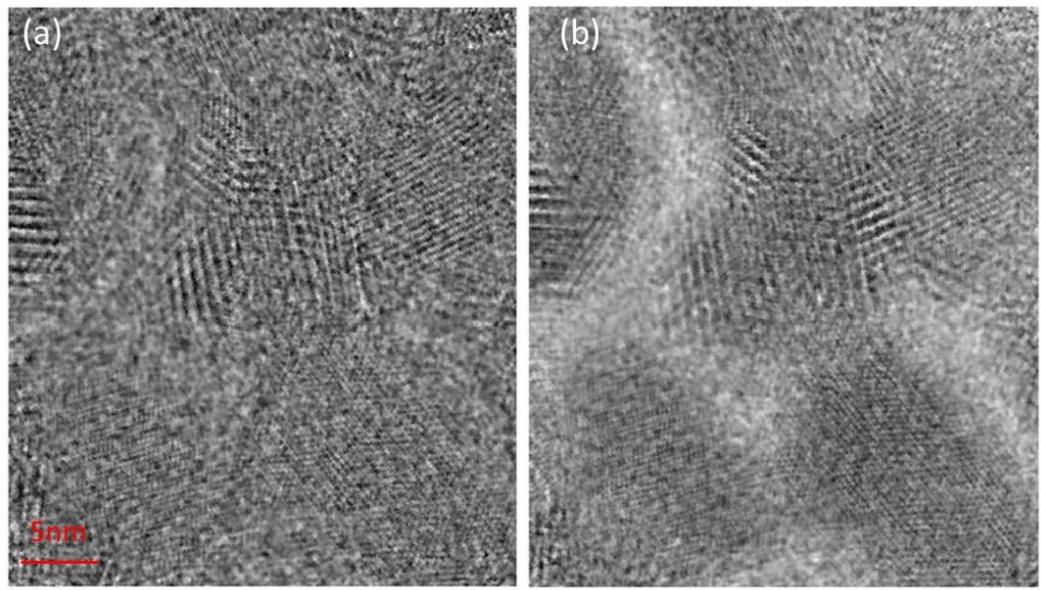


Figure 5.18 shows the reconstructed modulus (a) and phase (b) of dataset 1. The sample is gold nanoparticles on an amorphous carbon film. The gold lattice fringes can be seen in the reconstruction.

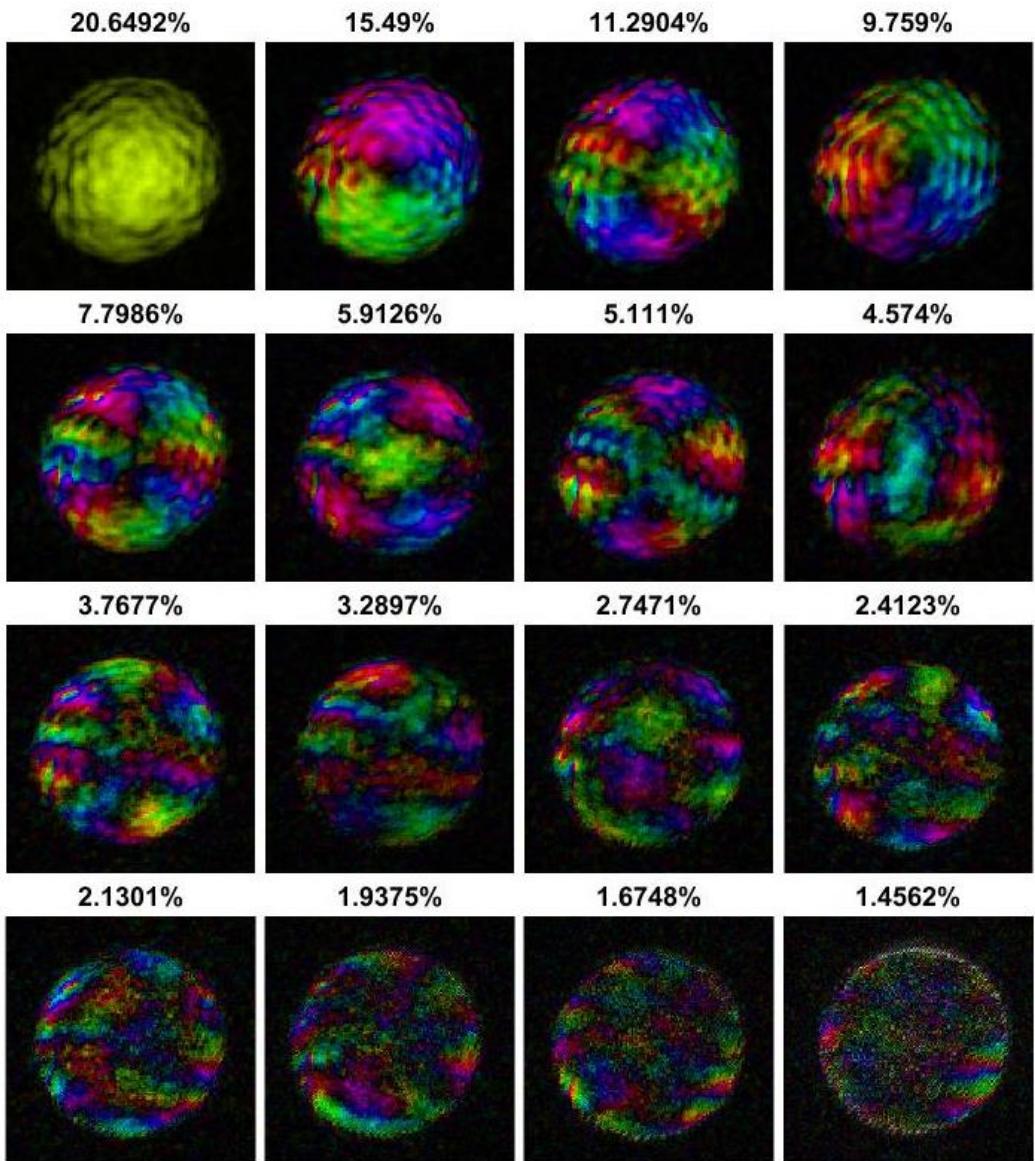


Figure 5.19 shows the reconstructed modes that are decomposed from the partially coherent STEM source via the modal decomposition ptychography. The modes were propagated and presented at the condenser aperture plane. The percentage above each mode shows the contribution of the corresponding mode to the effective source.

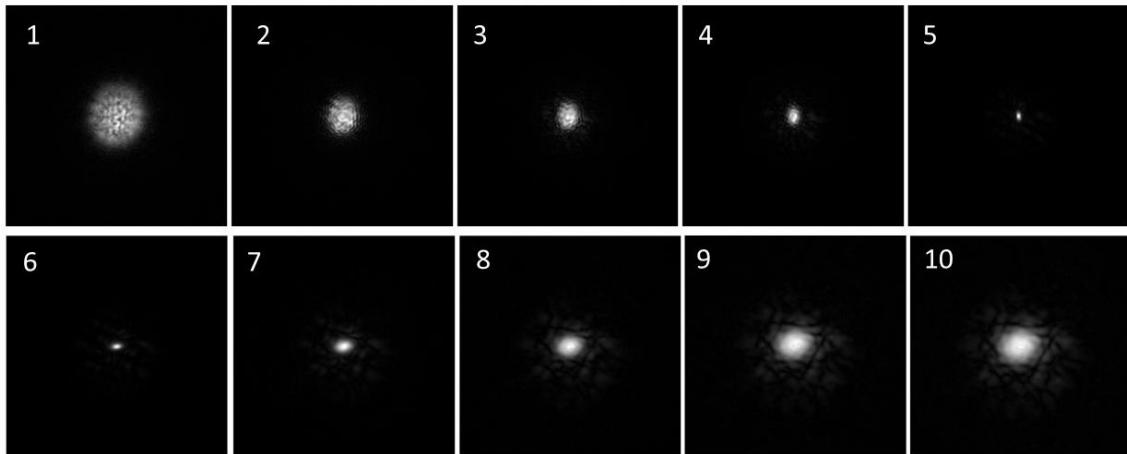


Figure 5.20 shows the propagating series of the first mode shown in Figure 5.19. (5) is at the focal plane.

The propagating series shows clearly the astigmatism in the illumination.

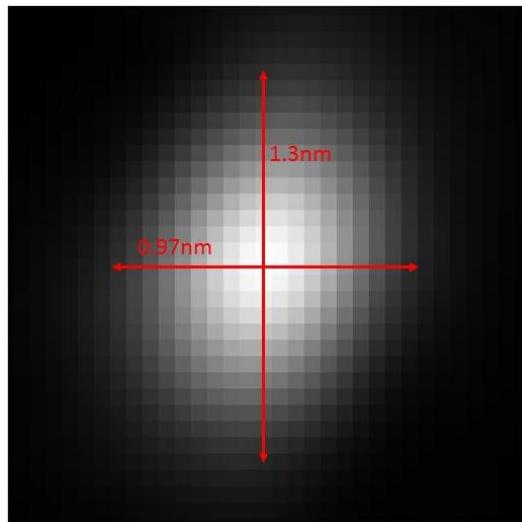


Figure 5.21 shows the reconstructed effective source profile, which is the superposition of all the modes shown in Figure 5.18 propagated to the focal plane. The physical dimension of the effective source is about 0.97 nm in one direction and 1.3 nm in the other direction. This is not the dimension of the real source, but the dimension of the convolution of the real source with the condenser aperture.

Figure 5.22 shows the specimen reconstruction of dataset 1 with a different Fourier constraint from the reconstruction result presented in Figure 5.18. The Fourier constraint we applied for the reconstruction in Figure 5.18 was exactly the intensities that the CCD detected, as

$$\Psi' = \sqrt{I} \frac{\Psi}{|\Psi|}; \quad (5.10)$$

while the Fourier constraint we applied for the reconstruction in Figure 5.22 was that

$$\Psi' = I^k \frac{\Psi}{|\Psi|}, \quad (5.11)$$

where k was set between 0.5 and 1 along the iterations.

The reason we replaced the Fourier constraint with Eq. 5.11 is that, when the data is noisy, the detected intensities of the diffraction patterns are not accurate constraints; if we still use the square root of them to replace the modulus of the calculated diffraction spectrum, we are forcing the algorithm to search for a wrong solution. So we need to modify the Fourier constraint. A more accurate model to modify the Fourier constraint when the count is low is Poisson distribution (Thibault and Guizar-Sicairos, 2012). The model we gave here is a simple model; it is not exactly accurate but it improved the reconstruction.

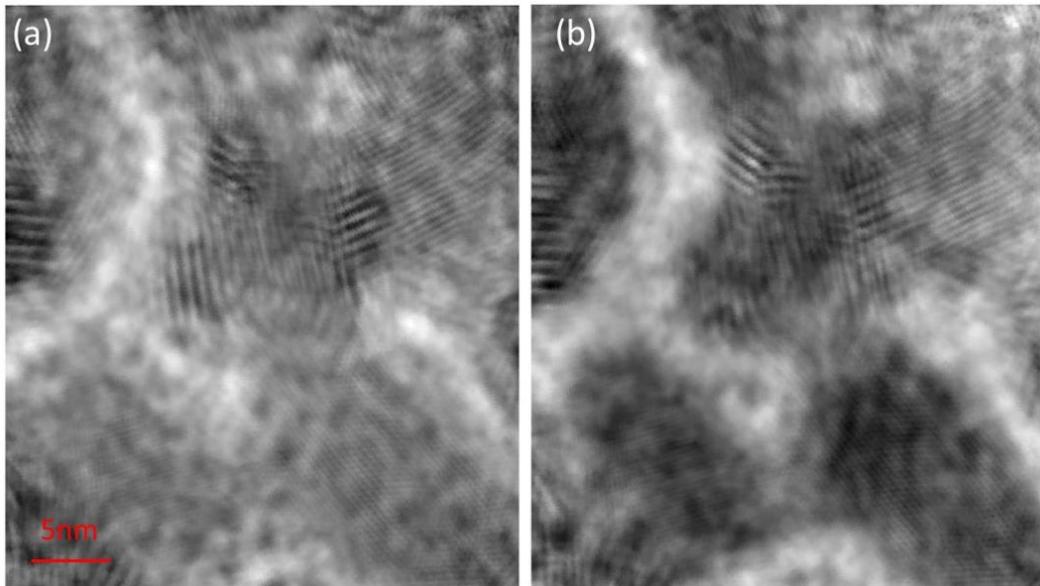


Figure 5.22 shows the modulus (a) and phase (b) of the reconstruction with the exponent of the diffraction patterns. The phase contrast was stronger compared with the result in Figure 5.17.

The reconstruction results of dataset 2 were shown in the following figures. Figure 5.23 shows the reconstructed phase and modulus of the sample. Some fringes were reconstructed

but the particles were very thick. Figure 5.24 shows the 16 modes of the decomposed aberrated illumination. The modes were propagated to the condenser aperture plane. When there is no aberrations, the first mode at the condenser aperture plane should be flat; however, in the case we show here, there is coma in the illumination; this is why we can see part of the first mode is flat and the other part of it shows some diffraction rings. Figure 5.25 shows the propagation series of the first mode. Figure 5.25 (1) and (10) are about at symmetric planes by the focal plane; that is why the diffraction rings appear upside down.

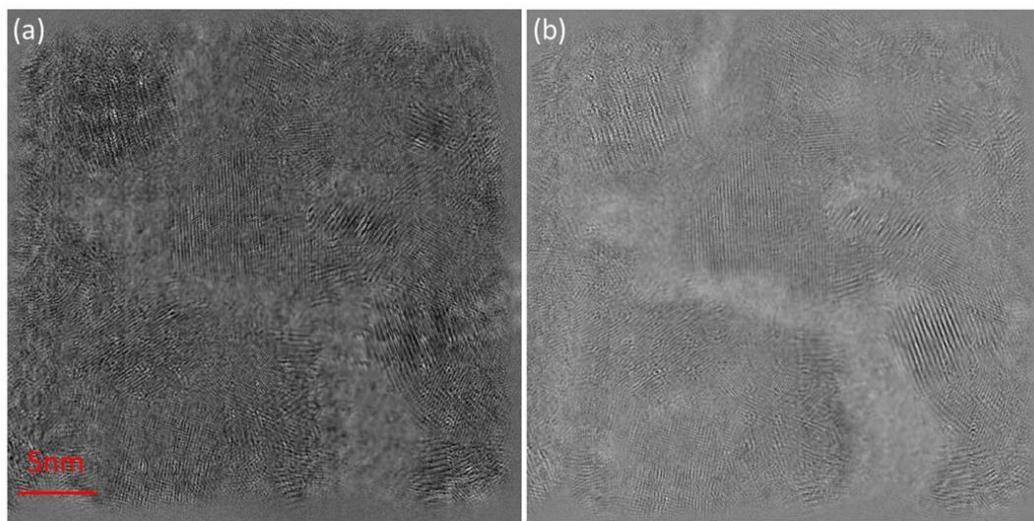


Figure 5.23 shows the reconstructed modulus (a) and phase (b) of dataset 2. It was the same sample as dataset 1, but a different area of it. There are thicker gold nanoparticles in this area; we can see some gold lattice fringes from the reconstruction.

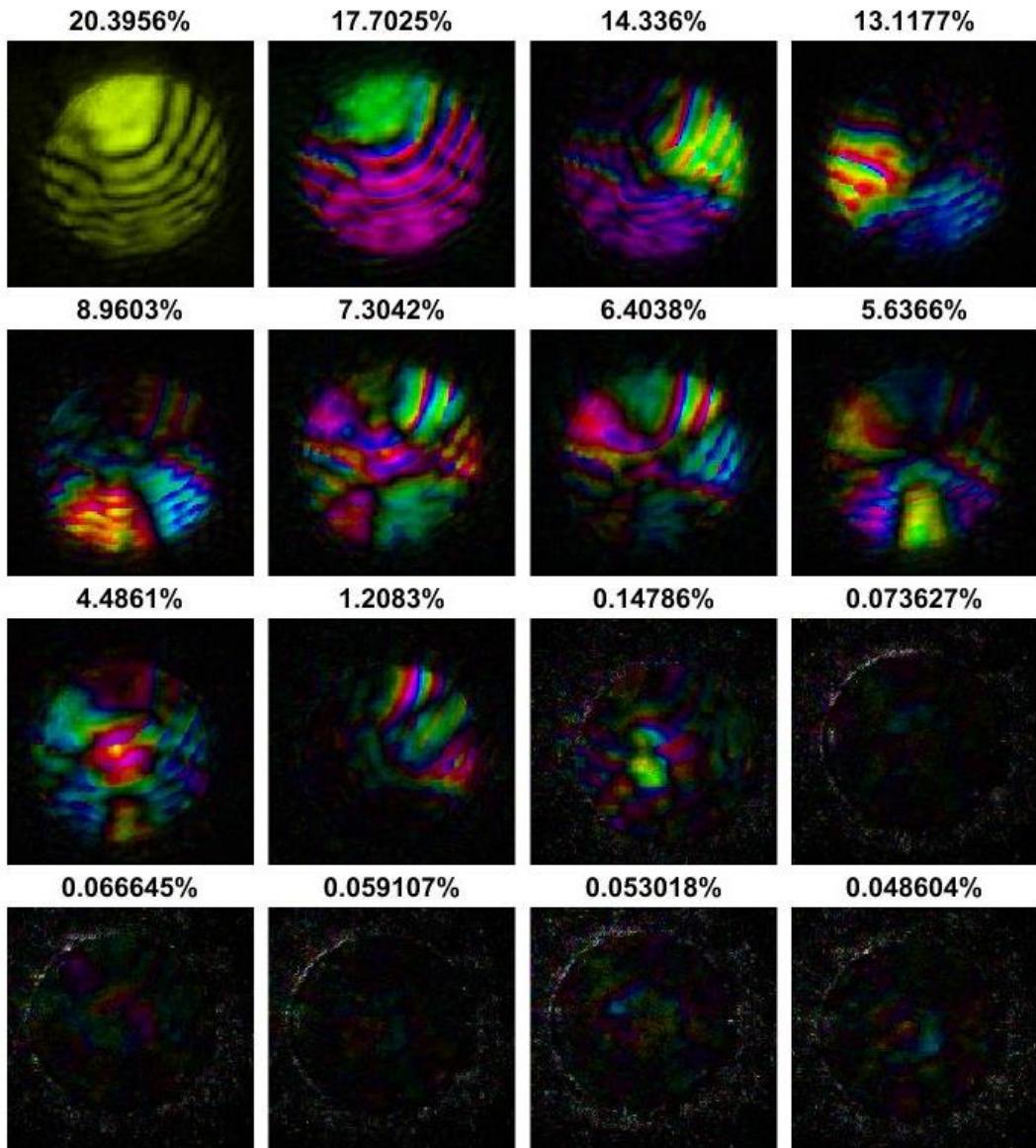


Figure 5.24 shows the reconstructed modes decomposed from the partially coherent illumination. The modes were propagated and presented at the condenser aperture plane. The percentage above each mode shows its contribution to the effective source.

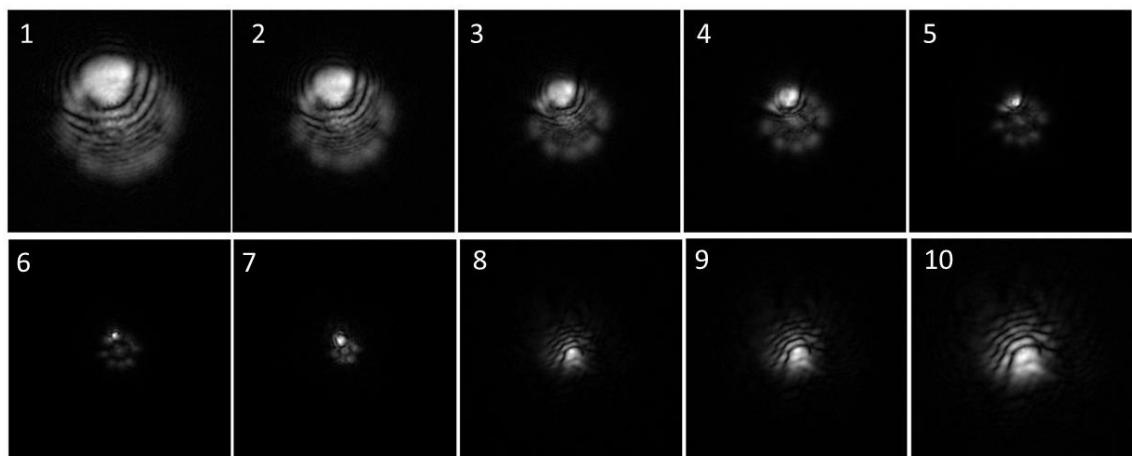


Figure 5.25 shows the propagating series of the first mode in Figure 5.24. (6) shows the mode at the focal plane. From the propagating series, we can see the coma aberration in the illumination was reconstructed by ptychography.

5.4 Discussion

In this chapter, we compared three kinds of images at the atomic scale of the gold nanoparticles. These are the reconstruction of defocused probe STEM ptychography, the reconstruction of focused probe STEM ptychography, and the HAADF image. Of these three methods, each has its pros and cons. The HAADF image is a direct imaging method, which is easy to obtain from a STEM, but can only image the heavy atoms. Currently, the focused probe STEM ptychography gives the best results, which has comparable resolution with the HAADF image. However, in the long run, focused probe ptychography is limited by the small field of view and sample radiation damage. We have presented the reconstruction of the defocused probe STEM ptychography in this chapter. Even though the lattice fringes of the gold nanoparticles have been reconstructed, the result is far from what we expected. We analyse that there are two critical issues in our experiment that hampered the reconstruction; one is the thickness of the specimen; the thickness of the sample that was applied in the focused probe STEM ptychography was about 5 nm; however, the sample we tested in this experiment was a standard TEM test sample, which is much thicker. When the electron beam goes through the thick gold particles, multiple-layer scattering and inelastic scattering have to be considered. In our reconstruction, the weak phase object approximation is not accurate.

The other issues in defocused probe ptychography are the scanning positions, the probe defocus, and the camera length coupling issue, which will be simplified as the positions-defocus coupling issue in the following text. The positions-defocus coupling issue occurs when, in the iterative ptychographic reconstruction, any two of the three parameters (the camera length, the scanning positions, and the probe defocus) are not determined. In this case, the algorithm searching can be easily trapped in a local minimum, where a set of wrong (but matched) solutions – wrong defocused probe function, wrong scanning positions with a scaling factor, and the object function at a wrong defocused plane – will be obtained. Here we

used two sets of data to demonstrate the positions-defocus coupling issue: the first dataset is dataset 1 that has been used in the previous section; the reconstruction process is the same as the previous reconstruction, except that the scanning positions were scaled with a factor of 110% in both scanning directions.

$$Pos_{stretch} = Pos \cdot 110\%. \quad (5.12)$$

Thus, of the three parameters: the camera length has been calibrated, the probe defocus, and the scanning positions are undetermined, and they are supposed to be refined during the reconstruction with the probe update function and the ‘jiggle’ method.

Figure 5.26.b shows the phase of one reconstructed gold nanoparticle when given the initial positions that were stretched by 10%; the lattices fringes of the gold nanoparticle in the reconstruction are still quite clear; however, the fringes; which were localized inside the gold nanoparticle in the previous reconstruction when the positions were corrected, now spread outside the particle; there are no clear boundaries to the particle. Figure 5.26.c and 5.26.d show the first mode of the reconstructed probe function at the sample plane; the size of the first mode with the correct scanning positions was about 84 pixels; the size of the first mode with the 10% stretched scanning positions was about 94 pixels; the pixel size of the two reconstructions is the same, which is obtained as

$$dxy = \frac{\lambda L}{D}, \quad (5.13)$$

where L is the camera length, and D is the detector dimension. So, the reconstructed probe with the 10% stretched scanning positions was also stretched by about 10%. Figure 5.27 shows the scanning positions. The red dots plot the correct positions. The blue crosses plot the ‘jiggle’ corrected positions when the algorithm does the reconstruction with the 10% stretched

positions. The annealing method lost its efficiency to correct the positions when there is a global scaled error in the scanning positions.

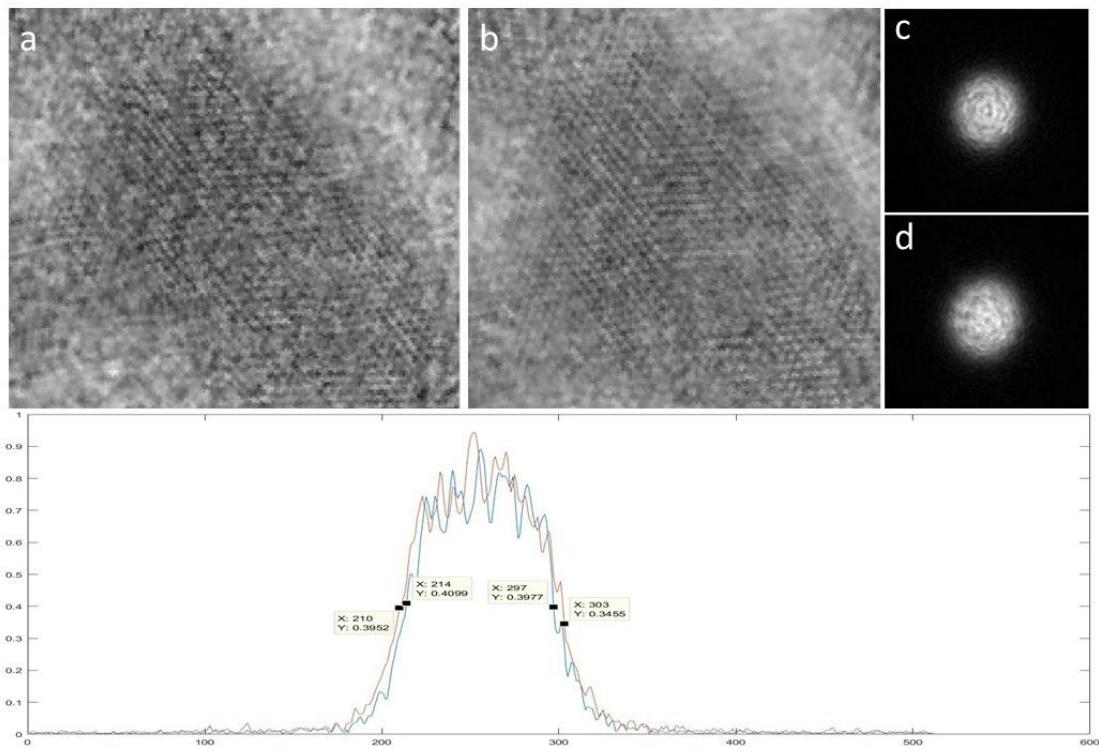


Figure 5.26 compares the specimen and probe reconstruction with the corrected positions and the 10% stretched positions. (a) shows one gold nanoparticle reconstructed with the corrected positions; (b) shows the same particle when reconstructed with the 10% stretched positions. The lattice fringes in (b) spread outside the particle. (c) shows the first mode of the reconstructed probes with the corrected positions; the diameter is about 84 pixels; (d) shows the first mode of the reconstructed probe with the 10% stretched positions; the diameter is about 94 pixels; which is also stretched by about 10% compared with (c).

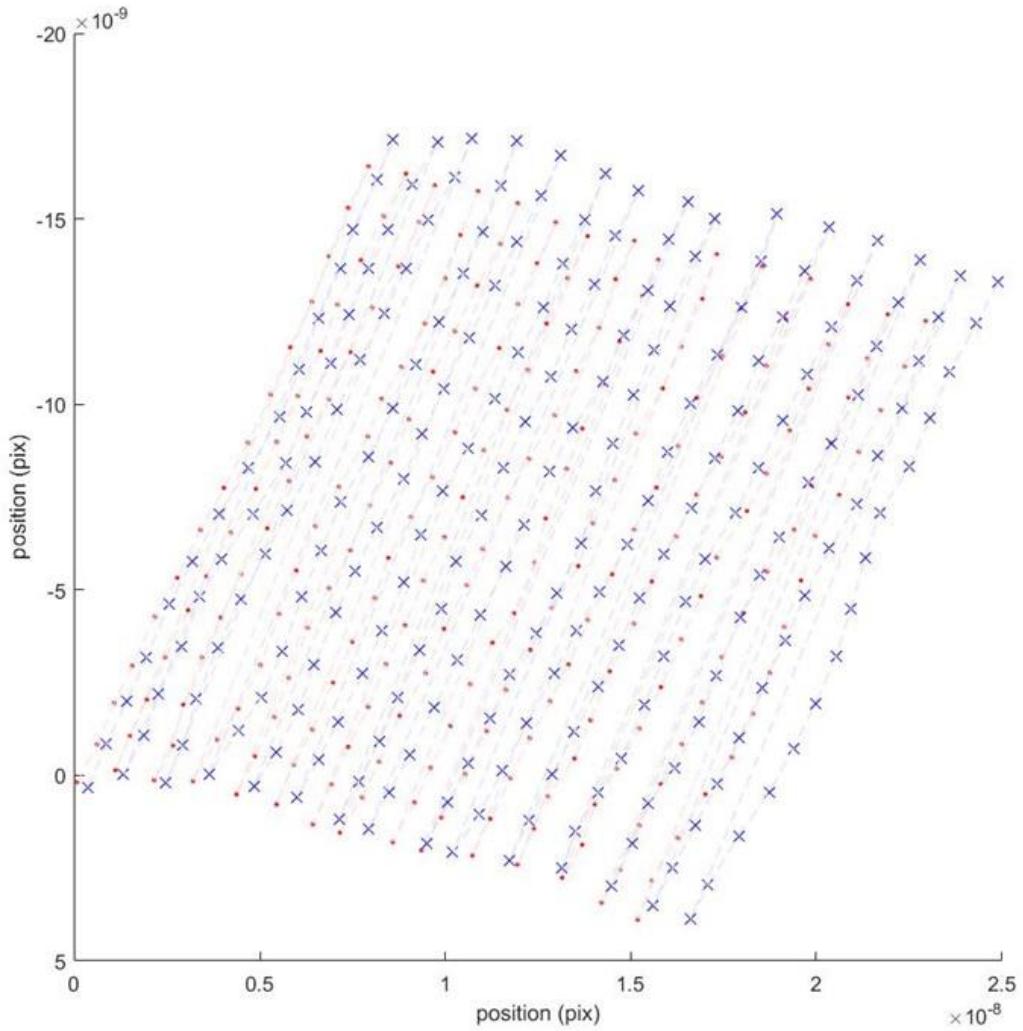


Figure 5.27 plots the probe scanning positions: the red dots plot the corrected positions; the blue crosses plot the ‘jiggle’ corrected positions when the algorithm reconstructs with the 10% stretched positions.

Jiggle failed to correct the stretching error in the scanning positions.

The other set of data used to demonstrate the positions stretching and camera length error coupling in the reconstruction was collected in the selected area diffraction mode (SAD). We try to show that the coupling problem also exists for the plane wave illumination ptychographic reconstruction when the camera length and the scanning positions are undetermined. The algorithm used for the reconstruction was the same with the previous reconstruction, except that blind Gaussian convolution method (Burdet et al., 2015; Maiden et al., 2015) was applied to remove the partial coherence of the source and the detector point

spread; more details of the method has been discussed in Chapter 2.7 and Chapter 3. Three reconstructions were compared: in the first reconstruction, all the parameters were corrected; in the second reconstruction, the camera length was scaled by 1/1.1; in the third reconstruction, the scanning positions were stretched by 10%.

Figure 5.28 shows the modulus and phase of the specimen from the three reconstructions.

Figure 5.28.a is the reconstruction with the correct parameters; Figure 5.28.b is the reconstruction with the scaled camera length; Figure 5.27.c is the reconstruction with the 10% stretched scanning positions. In the SAD ptychographic setup, the diffraction patterns are collected in the near field; when there is an error in the camera length, the sample will be reconstructed at a defocused plane, as shown in Figure 5.28.b, the modulus reconstruction of the specimen appears as though out of focus, while the phase reconstruction still has a relatively good resolution. When there is a stretching error in the scanning positions, as shown in Figure 5.28.c, the modulus reconstruction of the specimen also appears out of focus, but the resolution of the phase reconstruction is bad.

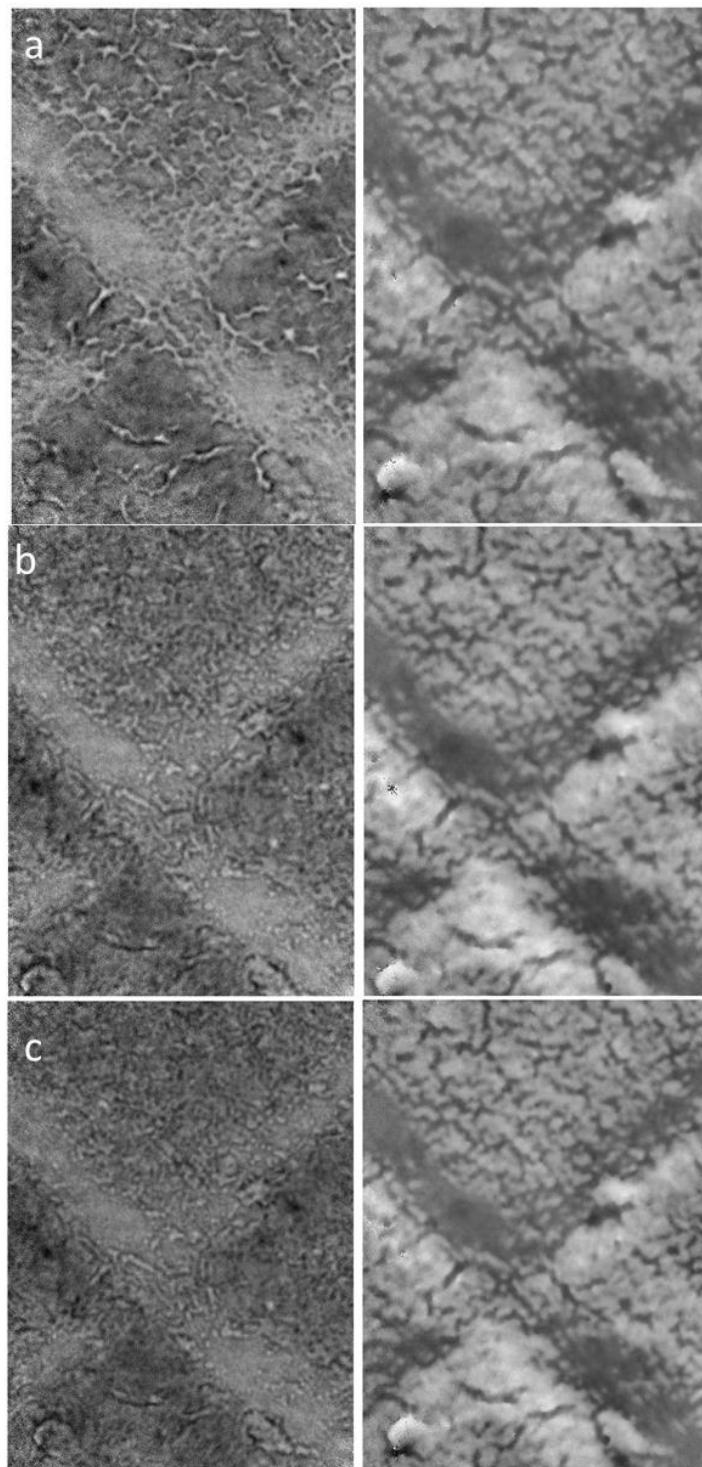


Figure 5.28 compares the modulus (left) and phase (right) reconstruction of the sample with (a) the correct parameters, (b) the 1/110% scaled camera length, and (c) the 10% stretched scanning positions. When the camera length is slightly scaled (b) or when the scanning positions are stretched (c), the modulus reconstructions have obvious differences – they look like out of focus images of the sample. The resolution and the contrast of the phase reconstructions are slightly affected.

Now let us see how ‘jiggle’ corrected the scanning positions when the camera length is slightly wrong or when the scanning positions are stretched. Figure 5.29 compares the ‘jiggle’ corrected scanning positions when applying the scaled camera length and when applying the corrected camera length; the crosses plot the ‘jiggle’ corrected positions when the camera length was wrong in the reconstruction; the circles plot the ‘jiggle’ corrected positions when the correct camera length was applied in the reconstruction, while the red spots plot the original positions. We can see that the crosses and the circles were completely overlapped with each other, which means that, when the camera length in the reconstruction has a slight error, jiggle still works well to correct the random errors in the scanning positions and the rotation.

Figure 5.30 compares the scanning positions before and after jiggle correction when there is a 10% stretched error in the original assigned scanning positions to the reconstruction; the red dots, which are the same with the red dots in Figure 5.29, plot the corrected positions; the circles plot the original positions assigned to the algorithm, which were stretched by 10% compared with the recorded positions; the crosses plot the ‘jiggle’ corrected scanning positions when reconstructed with the 10% stretched positions. We can see that, jiggle cannot correct the global stretching error in the scanning positions. Figure 5.31 compares the reconstructed probes from the correct scanning positions with the reconstructed probe from the 10% stretched scanning positions; The diameter of the reconstructed probe with the correct positions is about 198 pixels; the diameter of the probe reconstructed with the 10% stretched positions is about 217 pixels; So, when the scanning positions are stretched by 10%, the corresponding reconstructed probe size will also be stretched by about 10%.

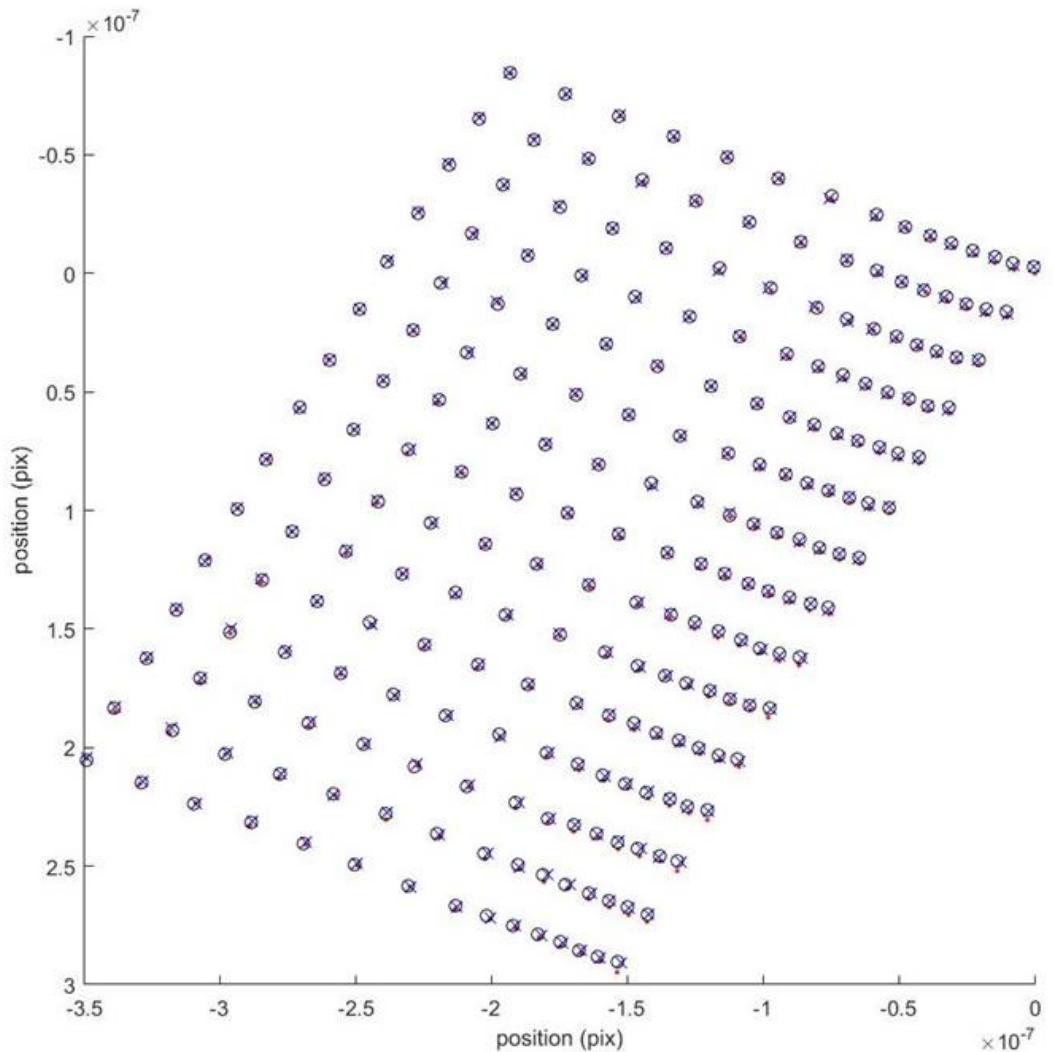


Figure 5.29 compares the ‘jiggle’ corrected positions when using the corrected camera length (plotted as circles) and when using a slightly wrong camera length (plotted as crosses). In both cases, jiggle successfully corrected the random positions errors and the rotation error. The red dots plot the originally recorded positions.

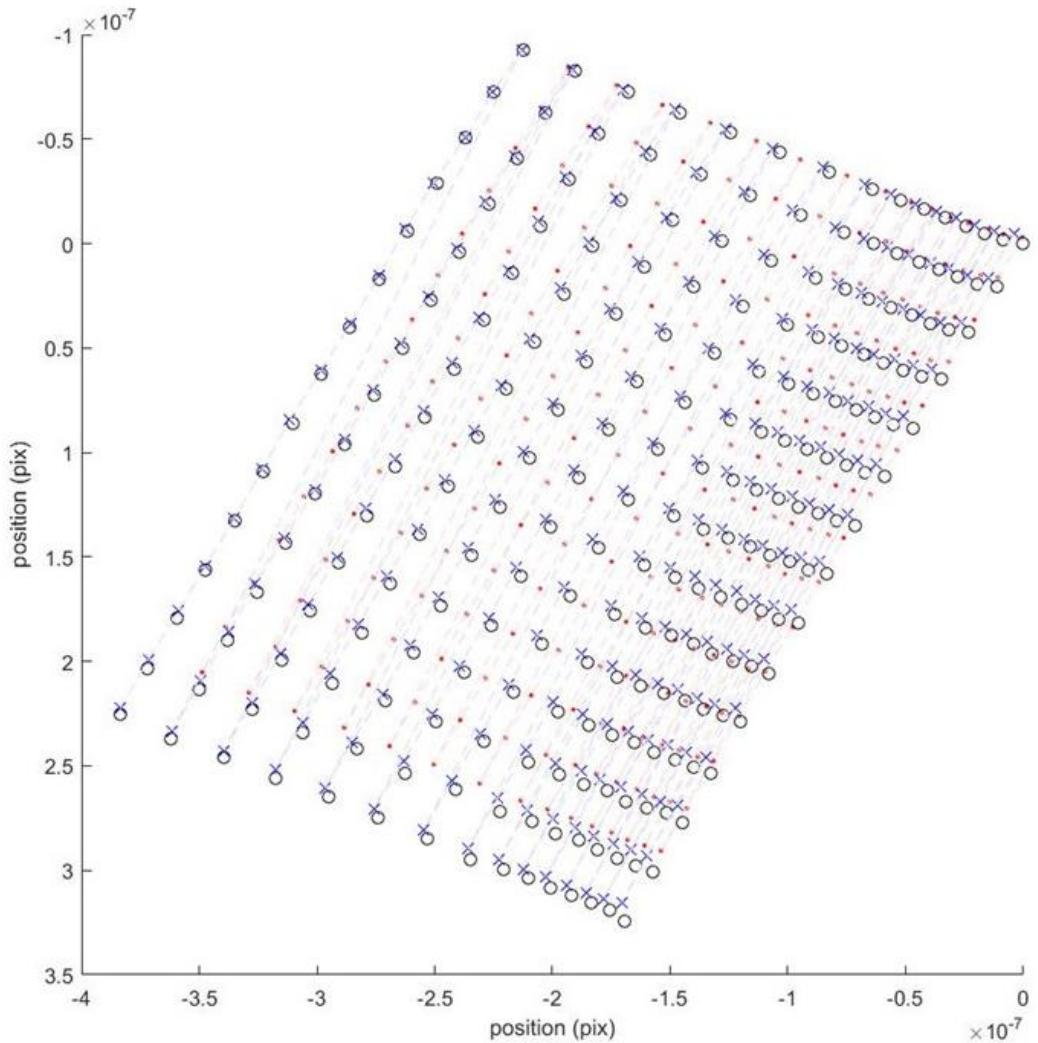


Figure 5.30 shows how ‘jiggle’ corrected the scanning positions when they have a 10% stretched error.
The red dots plot the original recorded positions; the circles plot the 10% stretched positions assigned to the algorithm at the beginning of the reconstruction; the crosses plot the ‘jiggle’ corrected positions at the end of the reconstruction. Jiggle corrected the rotation error and began to correct the stretching error,
but it does not work well.

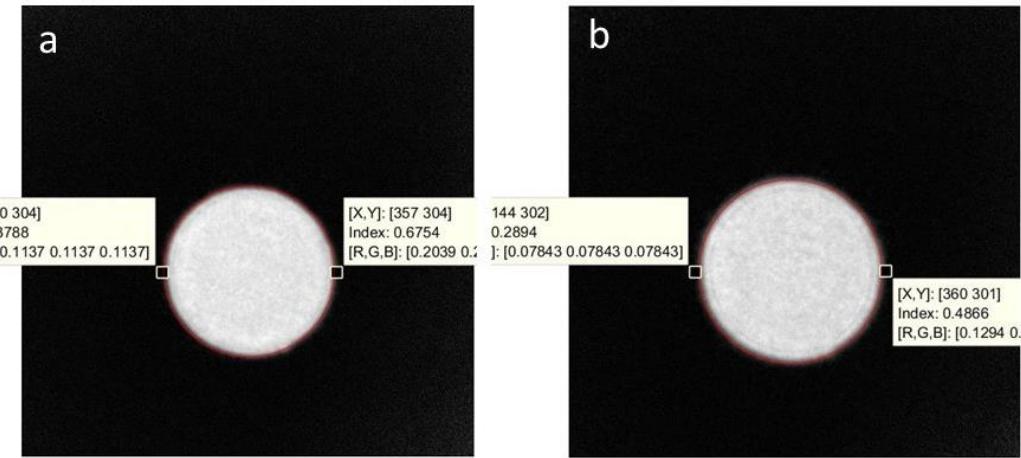


Figure 5.31 compares the reconstructed probes when using (a) the correct scanning positions and (b) the 10% stretched scanning positions. The diameter of the reconstructed probe from the correct positions is about 198 pixels; the diameter of the reconstructed probe from the 10% stretched positions is about 217 pixels, which is also stretched by $(217 - 198)/198 \approx 10\%$.

In conclusion, a correct camera length and accurate scanning positions are essential for ptychographic reconstruction. Random errors in the scanning positions can be corrected by either the annealing method (Maiden et al., 2012) or the cross correlation method (Zhang et al., 2013). However, when the camera length is wrong or when there is a global scaling error in the scanning positions, no matter if the illumination is planar or curved, the effect is to make the ptychographic reconstruction unrecoverable with the current algorithms. The wrong camera length or a scaled scanning positions map will likely give an out-of-focus modulus reconstruction and degrade the resolution and contrast in the phase reconstruction. When the positions are stretched (or squashed) in one or both directions, the probe reconstruction will also show stretching (squashing) in one direction or both directions correspondingly.

5.5 Conclusion

In this Chapter, we have calibrated a STEM to do large ptychography experiment. We have successfully collected good data and achieved reconstruction of some gold particles. For the first time, we decomposed aberrated electron illumination and showed the modal decomposition of astigmatism and coma. The factors that hampered the quality of the reconstruction have also been discussed; corresponding method that can improve the result has been proposed.



Chapter 6 The convolution in ptychography

6.1 Introduction

Hoppe (1969) discussed for the first time the concept of ptychography; the configuration is as shown in Figure 2.7.3. He also stated that ptychography can solve the object function from the diffraction pattern because of the convolution, which is the interference between the diffraction discs. ‘Ptycho’ means ‘fold’, which is like a convolution. Is the convolution necessary for large probe ptychography? In the paper (Humphry, 2012), the diffraction discs in the diffraction patterns (Figure 4.16) were not interfered. The reconstruction result has the fringe delocalization issue. Does the delocalization come from the lack of convolution? In this chapter, we are going to answer these questions.

The convolution in a diffraction pattern depends on both the structure of the sample and the structure of the illumination (Hoppe, 1969; Rodenburg, 2008). When we collect the data with a specific sample, the size of the condenser aperture will decide whether the diffraction discs overlap with each other or not. For example, Figure 6.1 shows some convergent beam diffraction patterns of a single crystal specimen when different size of the condenser apertures are inserted; when the condenser aperture is small enough, the diffraction discs are not overlapped; the larger the condenser aperture is, the more area the diffraction discs overlap with each other.

In the experiment, some factors may affect the convolution in the data, which is the connection between the diffraction discs of different orders. For example, the partial coherence of the illumination, the detector response property (point spread function, nonlinearity, pedestal and dead pixels), and the counts in the data. These factors may cut the connection (block the overlapping area) of the diffraction discs, or submerge some information in the connection area; thus, it may result some problems in the ptychographic reconstruction.

It is difficult to know how to measure convolution in the data in reality. Firstly, the convergent illumination has no clear boundary. The diffraction rings from a sharp aperture extend far away. So, there is always overlap in real space. Secondly, if the sample has amorphous structure or if the crystal has a boundary, there will always be convolution in reciprocal space. In this chapter, in order to clearly indicate the convolution in the data, we simulate a sample of a single crystal and without a boundary.

In section 6.2, we investigate how the convergence angle of the illumination and the structure of the sample affect ptychographic reconstruction. Two groups of simulations will be done; one is that we compare the ptychographic reconstruction from the diffraction patterns collected with different size of condenser apertures; the other is that we compare the reconstructions of the data collected from samples with structures from single crystal to nanoparticles on amorphous carbon film with different thickness. In section 6.3, we investigate how it affects the ptychographic reconstruction when the interference between the diffraction discs are interrupted. In section 6.4, we investigate the tolerance of iterative ptychographic reconstruction in the presence of Poisson noise.

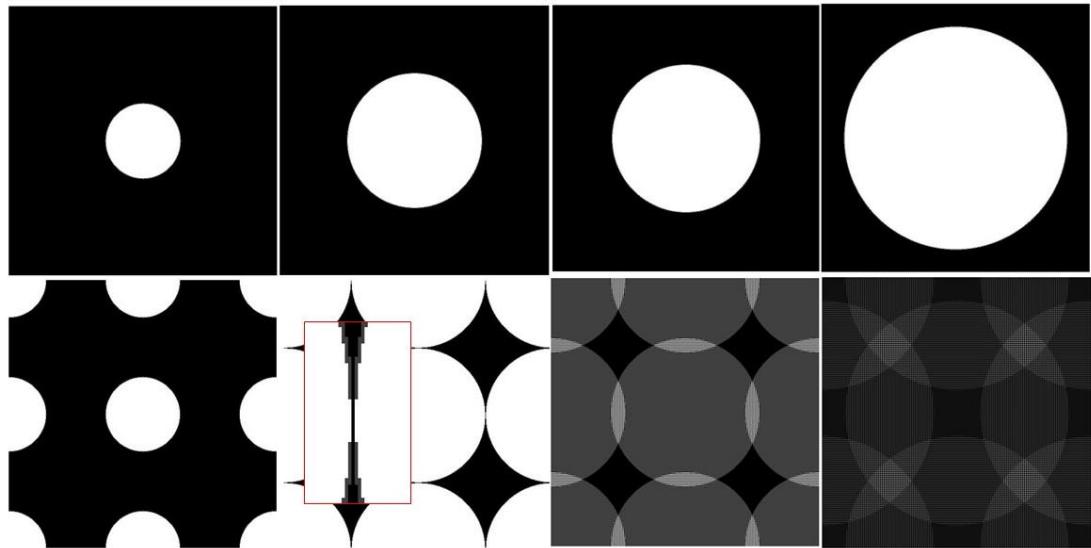


Figure 6.1 compares the diffraction patterns of a single crystal sample when different size of condenser apertures are inserted. (a) shows when a small condenser aperture is inserted, the diffraction discs are not overlapped with each other; (b) shows that the diffraction discs are just connected with each other; (c) and (d) show when larger condenser apertures are inserted, the diffraction discs are overlapped.

6.2 The convolution in ptychography

To quantify the overlap of the diffraction discs in reciprocal space, we use a single crystal specimen for the simulation (Figure 6.2). The Fourier transform of a single crystal sample is a series of delta functions, as shown in Figure 6.3. The diffraction pattern of it in reciprocal space is the convolution of the Fourier transform of the sample, which is the series of delta functions, with the Fourier transform of the illumination function, which is the profile of the condenser aperture (Goodman, 2004).

$$\mathcal{F}\{O \times P\} = \mathcal{F}\{O\} \otimes \mathcal{F}\{P\}, \quad (6.1)$$

where \mathcal{F} denotes the Fourier transform operator; O is the object function and P is the illumination function.

The convolution of any function with a delta function will be the function itself. So, the diffraction pattern of the single crystal sample will be a series of discs, where the delta functions decide the location of these discs, and the condenser aperture decides the size of the discs. Thus, it is easy to characterise how much overlap between the diffraction discs in the simulation.

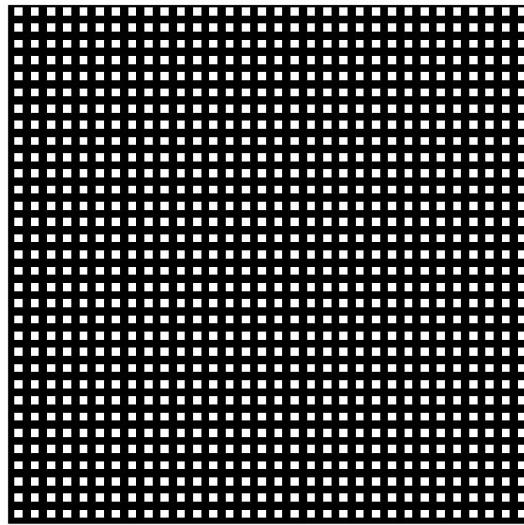


Figure 6.2 shows the sample applied for the simulation. It is a pure phase sample with single frequency; there are only two values of the phase; 0 and 1.

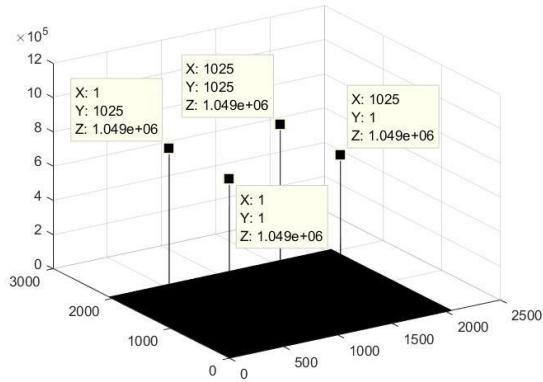


Figure 6.3 shows the intensities of the Fourier transform of the simulated sample. It is consisted of a series of delta functions.

We simulated 4 sets of data; the illumination function is fully coherent and there is no point spread effect or noise of the detector response; each of the dataset contains 10×10 diffraction patterns with a raster scan; the overlap of the illumination between each two adjacent positions is around 85%. Figure 6.4 shows the examples of the diffraction patterns from the 4 sets of data respectively. The only difference between these 4 sets of data is the size of the condenser aperture: the diameter of the condenser aperture inserted for collecting dataset 1 is

253 pixels; there is no overlap between the diffraction discs. The diameter of the condenser aperture inserted for collecting dataset 2 is 261 pixels; there is about 2% overlap compared with the area of the diffraction disc. The overlap of the diffraction discs in dataset 3 is about 5%. The overlap of the diffraction discs in dataset 4 is about 30%.

The iterative algorithm we apply to reconstruct the data is PIE (Rodenburg and Faulkner, 2004); we set the illumination function as known; the scanning positions are accurate; the initial object function is blank; during the reconstruction, only the object function is updated. The reconstructions of the sample are shown in Figure 6.5. When there is no connection between the diffraction discs in the data, the algorithm has not got enough information to retrieve the object function accurately; with a little bit connection between the diffraction discs (3%), the reconstruction has been improved a lot, however, the phase contrast is still not quite right; when the overlap between the diffraction discs increases to 5%, the reconstruction is correct.

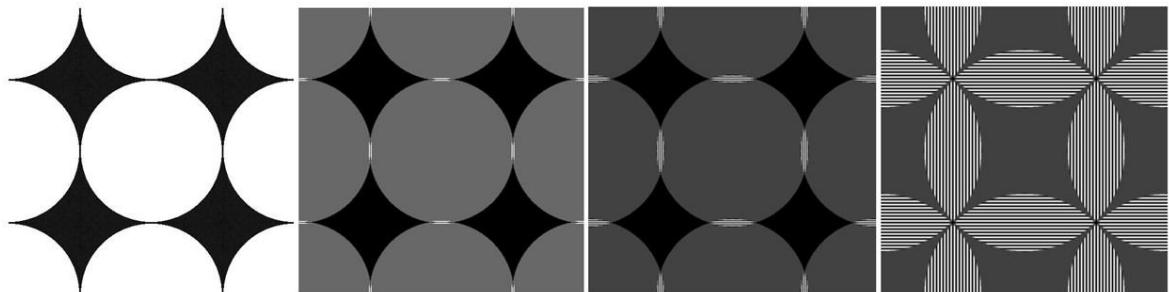


Figure 6.4 shows the examples of the 4 sets of diffraction patterns. The only difference between the 4 sets of data is the size of the condenser aperture: the diameter of the condenser aperture for dataset 1 is 253 pixels; there is no overlap between the diffraction discs; the diameter of the condenser aperture for dataset 2 is 261 pixels; the diffraction discs are just connected with each other; there is slightly more overlap of the diffraction discs in dataset 3, which is about 5%; the overlap of the diffraction discs in dataset 4 is about 30%.

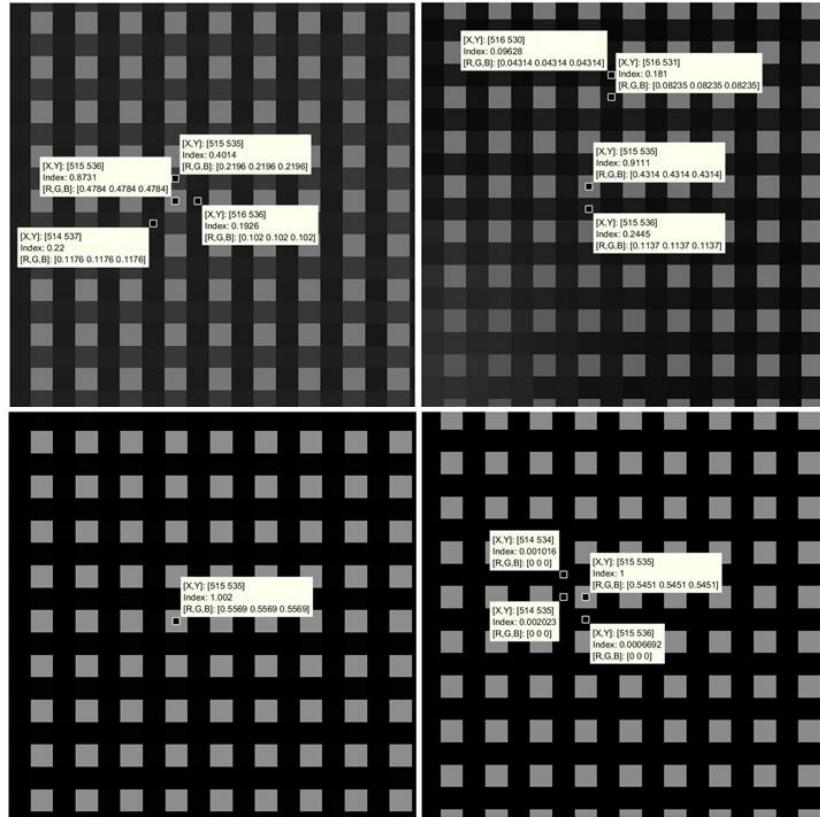


Figure 6.5 shows the reconstructions from the 4 sets of data: from the top left to the bottom right, the reconstructions from dataset with diffraction discs overlap of 0%, 3%, 5% and 30%. When there is no connection between the diffraction discs, the reconstruction is wrong; the algorithm failed to differ the contrast of the horizontal grids. When there is a slightly connection (3%), the reconstruction quality has been improved a lot; the contrast of the grids is clear, but the values are not quite right. When the overlap increases to 5%, both the contrast and the values of the grids is reconstructed correctly.

For a single crystal sample, we can conclude that ptychography reconstructs the sample from the interference between the diffraction discs. The interference is affected by both the structure of the sample and the structure of the illumination. In the simulation, when the structure of the sample is single crystal, the interference between the scattering discs only depends on the aperture size.

In the following part, we investigate when the structure of the sample becomes more complicate, whether it will relax the request for the size of the condenser aperture. The samples we simulated are: 1) sample 1 is made of some periodic particles with boundaries; 2)

sample 2 is made of some localized periodic particles on carbon film of 1nm thickness; 3)

sample 3 is made of some localized periodic particles on carbon film of 2nm thickness.

Figure 6.6 shows some areas of the three samples and their corresponding diffraction patterns.

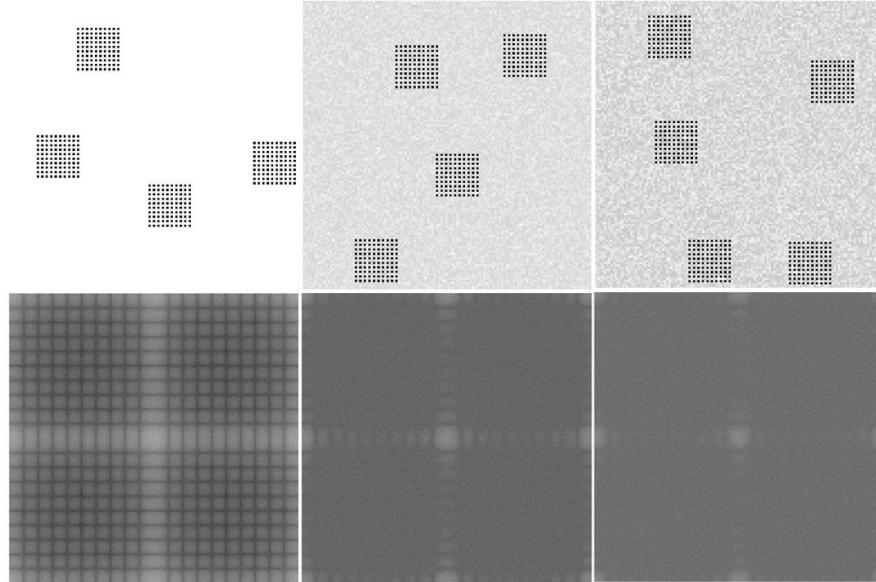


Figure 6.6 shows some area of the three simulated samples: (a) the sample is made of localized periodic particles; (b) the sample is consisted of periodic particles on carbon film with thickness of 1nm; (c) the sample is consisted of periodic particles on carbon film with thickness of 2nm. The figures below show the corresponding diffraction patterns.

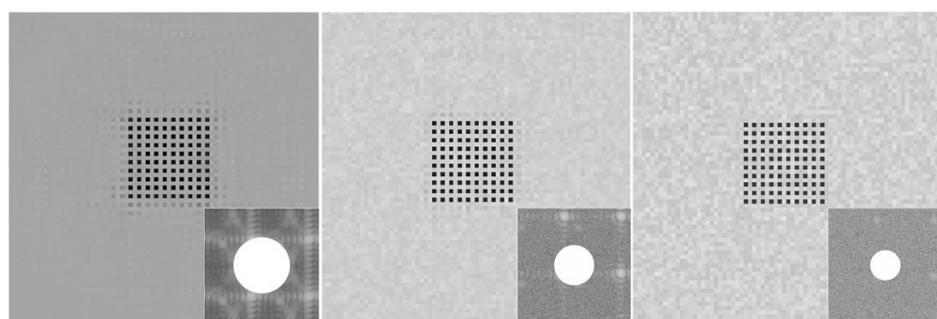


Figure 6.7 shows the ptychographic reconstruction of the data collected from the three samples. With the sample of localized particles, the smallest aperture needed for a reasonable ptychographic reconstruction has diameter of 254 pixels; with the sample of particles on 1nm thick amorphous carbon film, the smallest condenser aperture for a good ptychographic reconstruction has diameter of 180 pixels; with the sample

of particles on 2nm thick amorphous carbon film, the smallest condenser aperture needed for a good ptychographic reconstruction has diameter of 136 pixels.

For each sample, we simulated several sets of diffraction patterns with different size of the condenser apertures; each set of data was collected while the beam scanning across the sample with about 85% overlap over 10*10 grid positions. Then we do reconstructions on each set of data, until we find the smallest convergence angle where the aperture is just large enough to reconstruct the sample. Here are the results: the smallest condenser aperture needed for ptychographic reconstruction with the first sample is of dimeter 254 pixels; the smallest condenser aperture needed for ptychographic reconstruction with the second sample is of diameter 180 pixels; the smallest condenser aperture needed for ptychographic reconstruction with the third sample is of diameter 136 pixels. As there is more structure in the sample, a relatively smaller aperture is needed for ptychographic reconstruction. Moreover, as the sample has more structures, the reconstruction is more accurate; the boundaries of the particles are more localized (Figure 6.7). Table 6.1 concludes the critical condition of the condenser aperture size needed for ptychographic reconstruction with different samples.

Specimen	Single crystal	Local crystal	Local crystal with 1nm carbon	Local crystal with 2nm carbon
Convergence Angle	1.02	0.99	0.70	0.53

Table 6.1. the smallest convergence angle of condenser aperture size required for a good ptychographic reconstruction with different samples.

6.3 Tolerance of Ptychography on breaking the convolution

As we have explored in section 6.2, ptychography works relying on both the overlap of the illumination in real space, and the interference of the diffraction discs in reciprocal space.

The size of the condenser aperture and the structure of the sample are the direct factors that decide the interference of the diffraction discs.

In the ideal case, as long as the shifting step size in real space is reasonable, which is about 10% - 20% of the diameter of the probe (for a soft probe, the diameter here is defined at the full width at half maximum), and a large condenser aperture is selected respect to the structure (the thickness, the elements and the scattering property) of the sample, the iterative algorithm should be able to search for the correct solution of the illumination function and the object function. However, with the experimental data, because of some factors that will relax the constraints in real and reciprocal space, the algorithms may result in some ambiguous solutions even though we make sure there is convolution in the data (Faulkner and Rodenburg, 2005).

The factors that may relax the constraint in real space come from: 1) specimen shifting random error, which is introduced by the stage moving instability and hysteresis; 2) step scaling error, which is introduced by the calibration of the stepping pitch; 3) specimen drifting, which is caused by the temperature changing of the field; 4) illumination drift, which is introduced by the instability of the support current and temperature change. Some research has been done to improve the real space constraints accuracy from the algorithm (Maiden et al., 2012; Zhang et al., 2013; Maiden et al., 2015). More details about these methods have been discussed in Chapter 2.7.

The factors that may result the relaxation of the constraint in reciprocal space can be: 1) interference gap; this is the phenomenon that when there are some dead pixels on the detector

or when there are gaps between the pixels, some interference information may be lost. 2) detector noise; of the diffraction pattern, the central disc has the most counts, the information of which will be least influenced by the noise; while the counts in the interference area and some higher-order diffraction discs usually are much lower. The noise may submerge the interference and diffraction information in these areas. 3) source partial coherence, detector point spread and detector response nonlinearity, which will cause distortion of the information in the diffraction pattern. There is some research has been done to improve the influence of the detector point spread and detector response property, for example, modal decomposition and Gaussian blind convolution (Chapter 2.7 and Chapter 3). However, there is not any systemic investigation on how the interference gap and noise of the diffraction patterns affects the ptychographic reconstruction.

In this section, we investigate how much information in the diffraction pattern is necessary for ptychographic reconstruction. The simulation investigates how the partial dataset affects ptychographic reconstruction.

6.3.1 Partial diffraction patterns

The specimen used for this simulation was a pure phase single crystal specimen with the phase range between 0 to 1. The source was supposed to be fully coherent. A large condenser aperture was inserted to make sure that there was enough overlap between the scattering discs; the illumination was focused; the specimen was placed at a slightly defocused plane. 10*10 diffraction patterns were collected in the far field. Figure 6.8 shows one example of the diffraction patterns; Figure 6.9 plots the diffraction pattern through the centre along horizontal direction and along the diagonal direction respectively, to show the count distribution of the diffraction pattern.

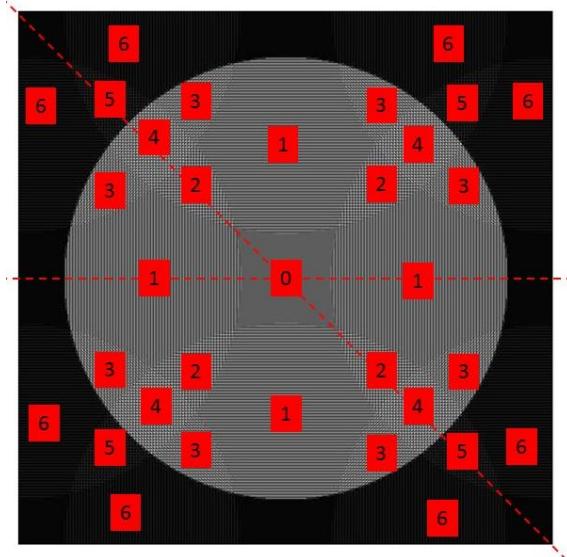


Figure 6.8 shows one example of the simulated diffraction patterns used for the calculations in this section.

The number labels the orders of the interference: 0th order shows the central disc, which has no interference.

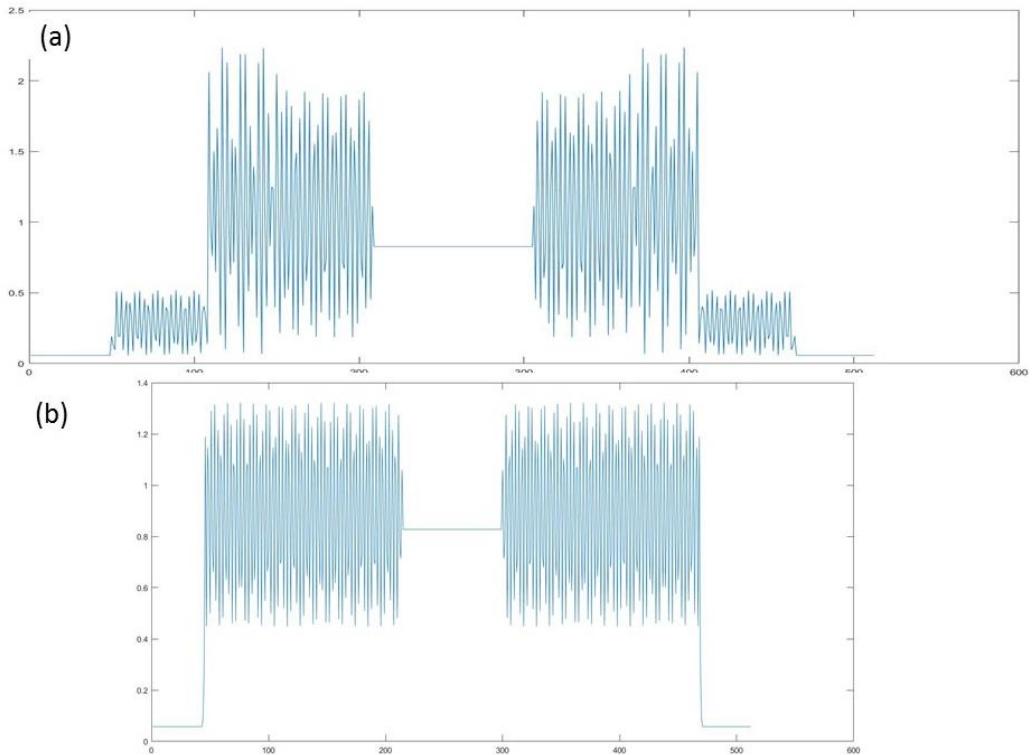


Figure 6.9 plots two cross sections of the diffraction pattern showed in **Figure 6.8**; (a) plots along the diagonal direction through the centre; (b) plots along the horizontal direction through the centre.

6.3.1.1 Simulation and results

To filter the information of different area in the diffraction patterns, several kinds of masks were generated. The filtered diffraction patterns are that the original diffraction patterns product with the mask:

$$I_m = I \times M, \quad (6.2)$$

where I is the original diffraction patterns and M is the mask.

The algorithm applied for the reconstruction was PIE (Rodenburg and Faulkner, 2004), which only reconstructed the object and supposed that the probe function was known. We also supposed that the mask function on the diffraction patterns was known in the reconstruction, so in reciprocal space, we replaced the modulus of the Fourier transform of the exit wave Ψ with the square root of the detected diffraction pattern by

$$\Psi' = \frac{\Psi}{|\Psi|} \cdot M \cdot \sqrt{I_m} + \Psi - \Psi \cdot M, \quad (6.3)$$

which means that, in the area where the information was detected, we replaced the modulus with the square root of the detected values; in the area where the information was lost, the values were floated as the calculated values. In other words, we used the redundancy in the dataset to retrieve the object transmission function as well as the complete diffraction patterns.

We explored the ptychographic reconstruction from these nine kinds of ‘information mask’:

- (1) Reconstruction with the central disc, which contain 0th – 4th order interference (Figure 6.10.a);
- (2) Reconstruction with part of the central disc which contain 0th, 1st, and 2nd orders interference (Figure 6.10.b);

- (3) Reconstruction with 1/16 corner of the whole patterns, which contain 4th – 6th complete order interference (Figure 6.12.a);
- (4) Reconstruction with ¼ corner of the whole patterns, which contain 0th – 6th order interference (Figure 6.12.b);
- (5) Reconstruction with 1/16 corner and central non-interference part (0th order) of the whole patterns, which contain 0th and 4th – 6th order (Figure 6.14.a);
- (6) Reconstruction with 1/16 corner and central 0th and little bit 1st order of the whole patterns: 0th, 1st, 4th – 6th order (Figure 6.14.b);
- (7) Reconstruction with the edge part (blocked the central disc) of the whole patterns, which contain 5th -6th order (Figure 6.16.a);
- (8) Reconstruction with the patterns which were blocked 1st order interference (Figure 6.16.b);
- (9) Reconstruction with patterns which were consisted of the 0th and 1st order (Figure 6.18);

Figures below show the reconstruction results with the corresponding masked diffraction patterns.

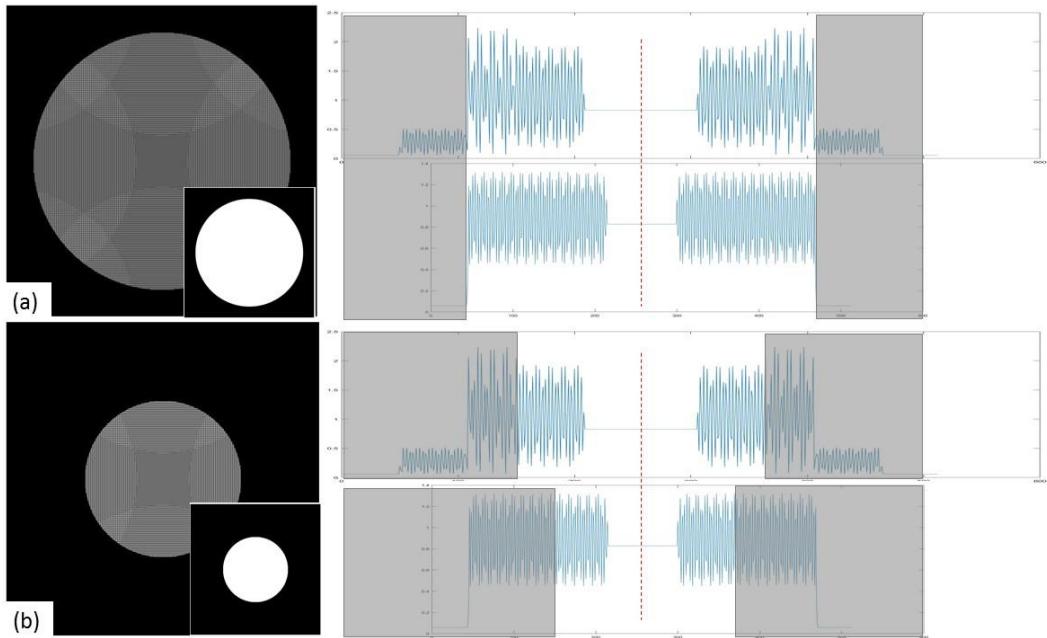


Figure 6.10 shows the masked diffraction patterns examples: (a) is from data 1; the information inside the central disc was kept and outside the central disc was blocked. (b) is from data 2, only part of the information inside the central disc (0^{th} and 1^{st} order) was kept. The figures at the right plot the corresponding filtered diffraction pattern along the diagonal and horizontal directions through the centre.

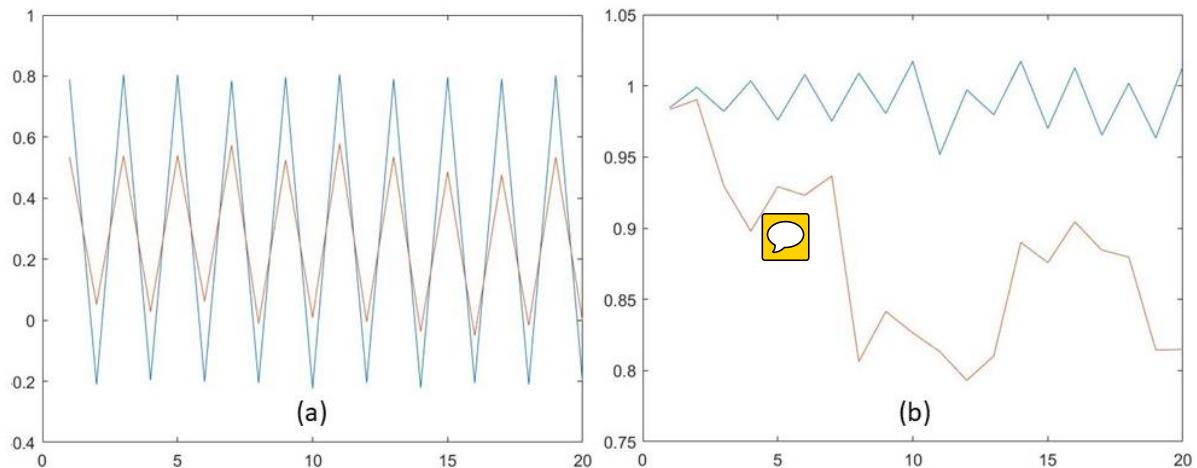


Figure 6.11 plots the reconstructed phase and modulus contrast of the two datasets with partially low frequencies. (a) Blue line plots the phase contrast reconstructed from the dataset which has the whole central disc; the phase contrast is 0, which is correct. Red line plot the phase contrast reconstructed from the dataset which has the central half of the central disc; the phase contrast is about 0.4, which is incorrectly. (b) the blue line plots the modulus contrast reconstructed from the dataset which has the

whole central disc, the contrast is nearly 1, which is correct; the red line plots the modulus contrast from the dataset which has the central half of the central disc; the modulus contrast is about 0.2.

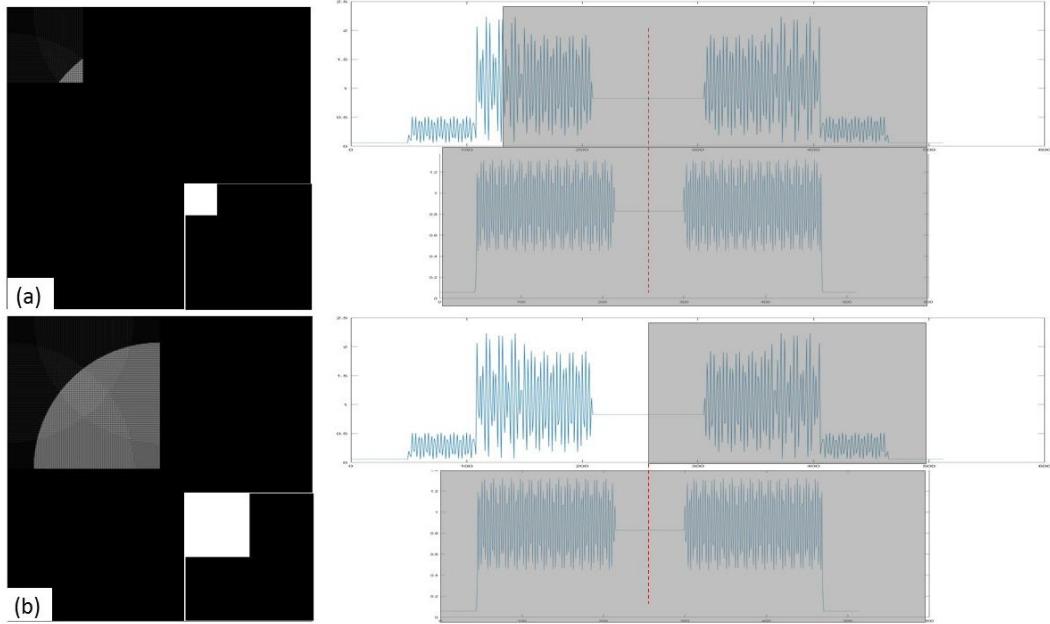


Figure 6.12 (a) shows one masked diffraction from data 3, only 1/16 of the top left corner information was kept. (b) shows one masked diffraction from data 4, 1/4 of the top left corner information was kept.

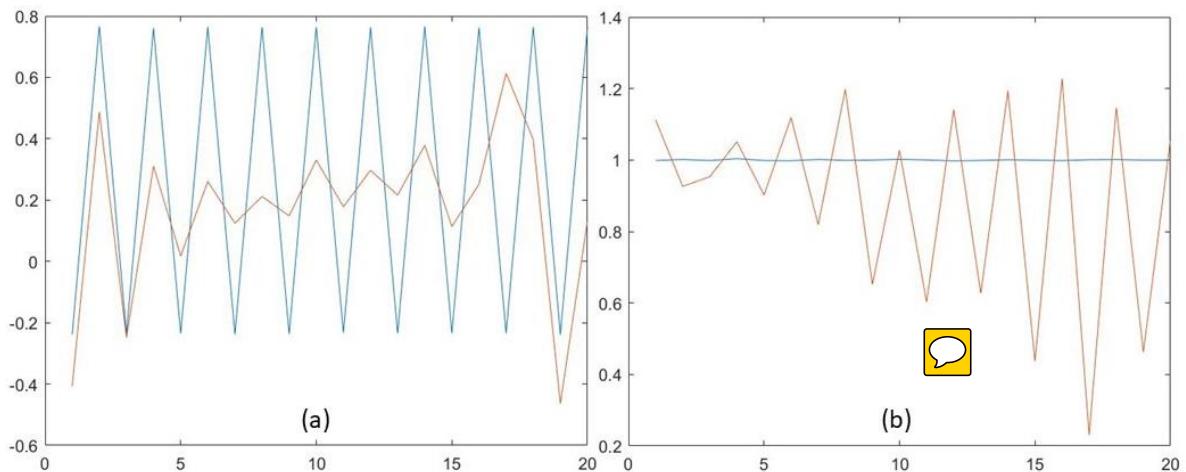


Figure 6.13 plots the reconstructed phase and modulus contrast from the two datasets with partially high frequencies. (a) the blue line plot the phase contrast reconstructed from one quarter of the diffraction pattern; the phase contrast is 1, which is correct. The red line plots the phase contrast reconstructed from one eighth of the diffraction pattern. The phase contrast is wrong. (b) The blue line plots the modulus contrast reconstructed from the dataset which has a quarter of the diffraction pattern; the contrast is 0,

which is correct. The red line plots the modulus contrast reconstructed from the one eighth dataset; the contrast is wrong.

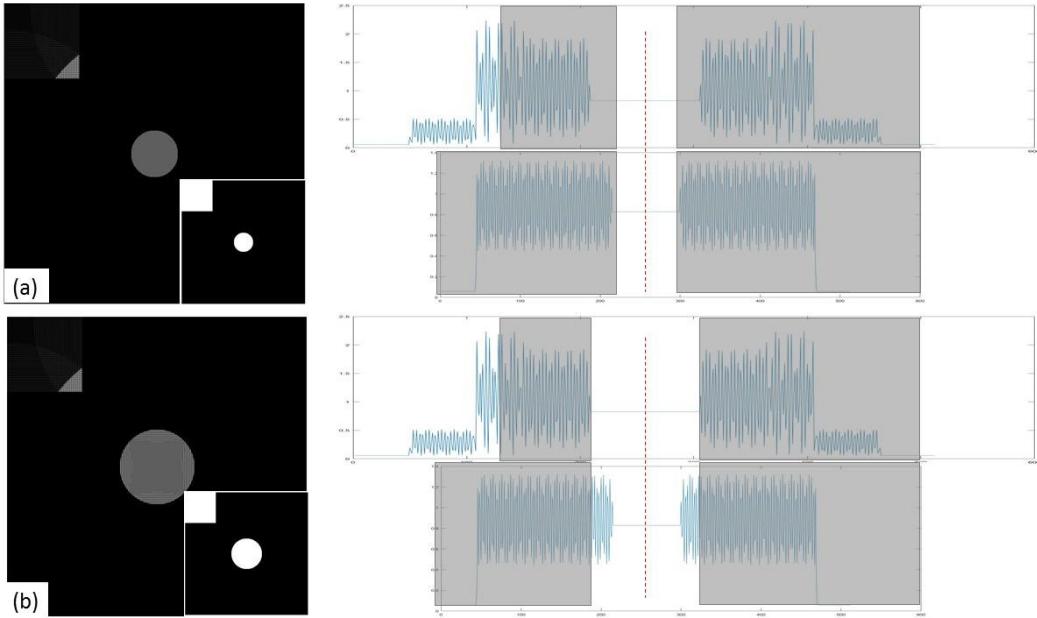


Figure 6.14 (a) shows one masked diffraction pattern from data 5, which had the information of the top left 1/16 and some 0th order information. (b) shows one masked diffraction pattern from data 6, compared with data 5, only slight the 1st order information in the central was added.

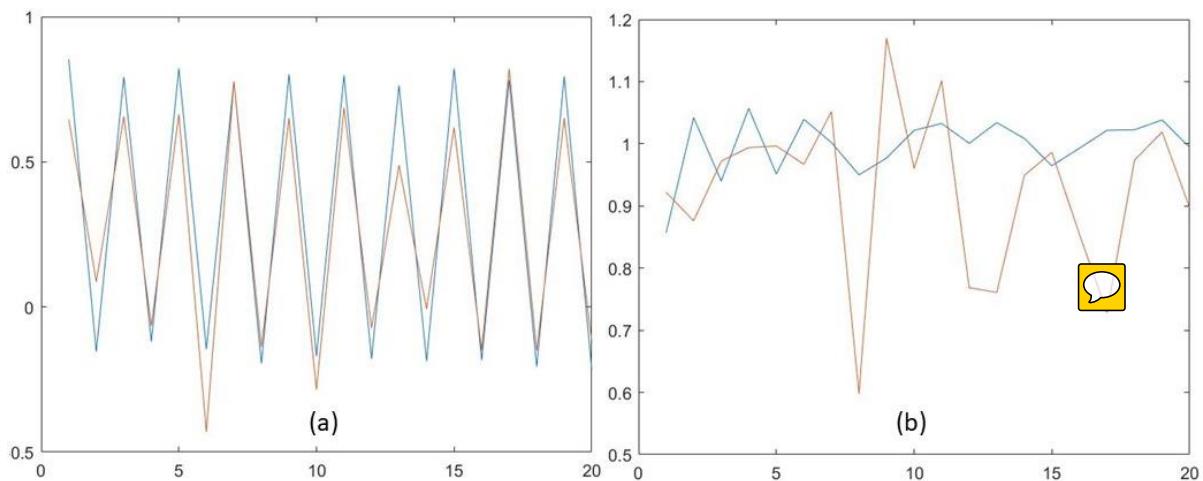


Figure 6.15 plots the reconstructed phase and modulus contrast from the datasets with partially low and partially high frequencies. (a) the blue line plots the phase contrast reconstructed from the dataset which has one eights of the diffraction pattern in the top left and a larger area inside the central disc; the contrast is 1, which is correct. The red line plots the phase contrast reconstructed from the dataset which

includes one eighth of the diffraction pattern in the top left and a smaller area inside the central disc; the phase contrast is not quite right. (b) the blue line plot the modulus contrast reconstructed from the dataset (1/8+larger central area); the modulus contrast is about right. The red line plot the modulus contrast reconstructed from the dataset (1/8+smaller central area), the contrast has larger error.

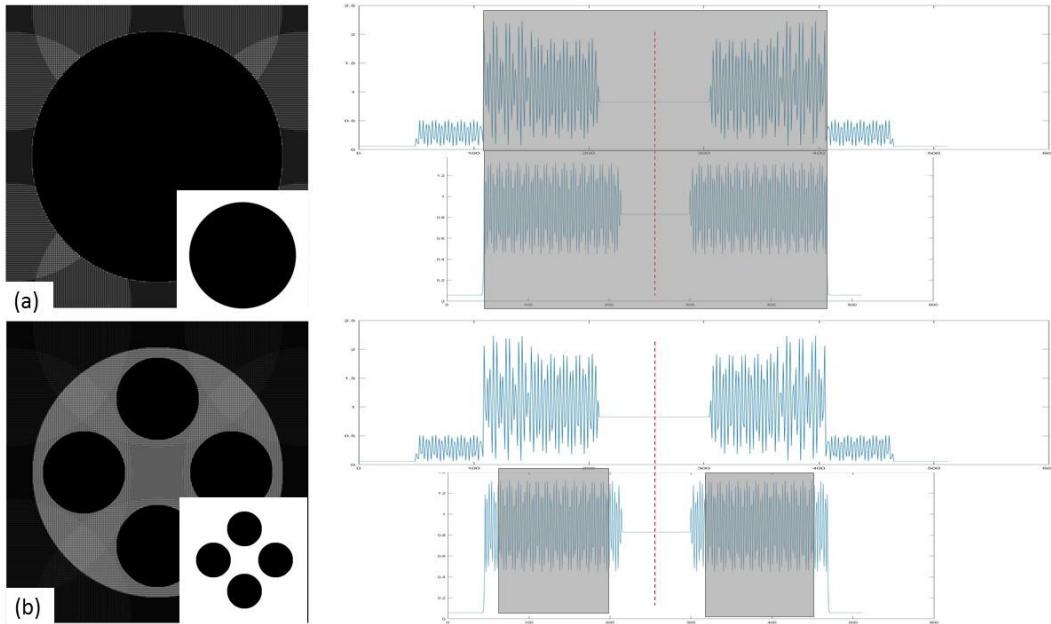


Figure 6.16 (a) shows one masked diffraction pattern from data 7; the information inside the central disc was blocked and the information outside the central disc was kept. (b) shows ones masked diffraction pattern from data 8, most of the 2nd order interference information was blocked.

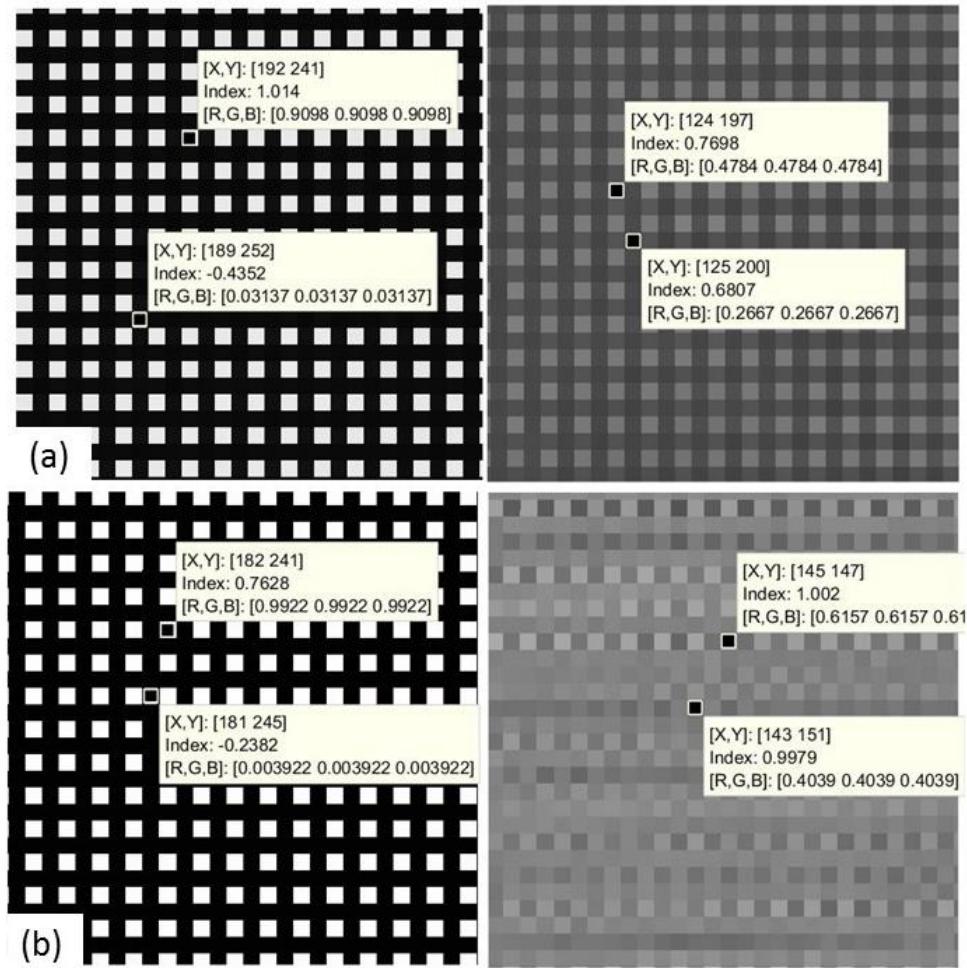


Figure 6.17 the phase reconstruction (left) and modulus reconstruction (right) from data 7 (a) and data 8 (b) in Figure 6.16. The phase contrast of both the two reconstructions were good, while the modulus reconstructions from both data had low contrast. Since data 7 had more information lost than data 8, the contrast in the modulus reconstruction from data 7 was stronger than that from data 8.

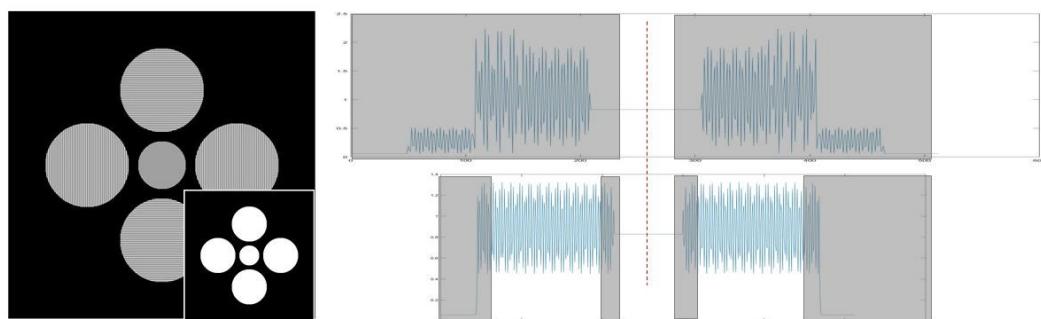


Figure 6.18 shows one masked diffraction pattern from data 9, only some of the 0th order and 1st order information was kept in data 9.

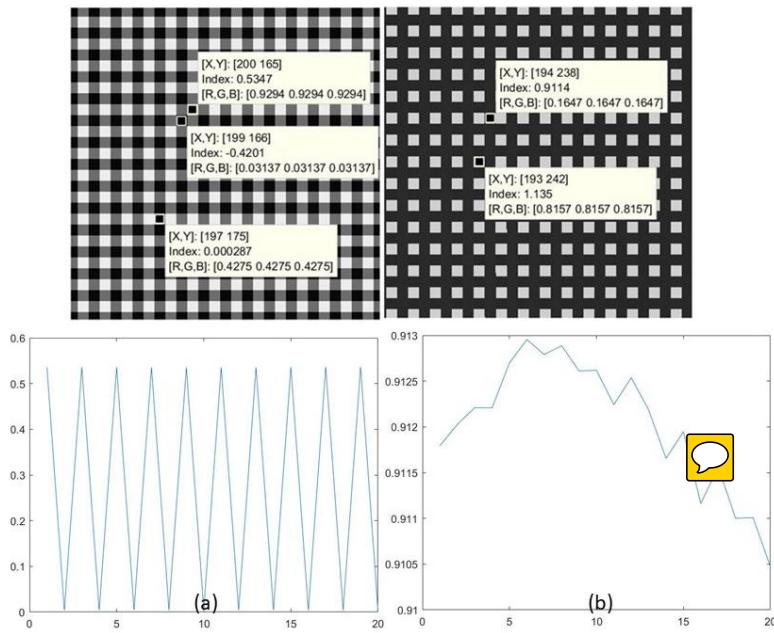


Figure 6.19 shows the phase reconstruction (top left) and the modulus reconstruction (top right) from data 9 in Figure 6.18. Neither the reconstructed phase contrast or the reconstructed modulus contrast was right. This reconstruction result is similar to the reconstruction form data 2. Both these two datasets only have the primary orders information 0th – 3rd. The figures in the bottom plot the reconstructed phase and modulus respectively.

6.3.1.2 Discussion

From the reconstruction results with the 9 sets of data, we can get the following conclusions:

- 1) The reconstructions of data 1, 4, 6, 7, and 8 have retrieved the phase contrast accurately. In this simulation, we do not need the whole diffraction patterns to solve the object, because respect to this very simple sample, the redundancy in the dataset, which comes from the oversampling in the real space, was enough to solve the unknowns.
- 2) Data 2, 3, 5 and 9 failed to reconstruct. When we keep reducing the information in the data by employing smaller masks, the reconstruction will fail at a point where the number of effective knowns in the dataset is less than the unknowns.

- 3) As shown in table 6.2, the datasets which succeeded to retrieve the phase contained at least 5 orders of interference information, except of data 7, which only had 2 orders of interference information.
- 4) Data 4 was the only dataset that reconstructed the correct modulus contrast, which was also the only dataset that has the complete interference orders from 0th to 6th. So, the missing orders of the interference in the dataset will be projected to the wrong contrast of the modulus reconstruction.
- 5) Data 2 and 9 were the datasets that only contain the primary orders interference (0th, 1st and 2nd); both the phase reconstruction and the modulus reconstruction from these two sets of data were wrong.

Data	Orders	Orders No.	Modulus	Phase
1	0 1 2 3 4	5	0	1
2	0 1 2	3	0*	0
3	4 5 6	3	0	0
4	0 1 2 3 4 5 6	7	1	1
5	0 4 5 6	4	0	0
6	0 1 4 5 6	5	0	1
7	5 6		0	1
8	0 2 3 4 5 6	6	0	1
9	0 1	2	0*	0

Table 6.2 summarizes the information of the 9 datasets. The Orders No. calculates the total orders each dataset has; in the modulus and phase column, 1 denotes good reconstruction and 0 denotes failed reconstruction.

6.4 conclusion



We give the conclusion that, large probe ptychography works relying on the overlap both in real space and in reciprocal space. When there is no overlap in real space, large probe ptychography will not work; when there is no convolution in reciprocal space, large probe ptychography will not work either.

How much overlap in real space does ptychography need? It depends on the structure of the illumination and the structure of the sample. Usually overlap of 70% - 90% is a safe choice for all kinds of setups and samples.

The convolution in the reciprocal space also depends on the structure of the sample and the structure of the illumination (the aperture size in the simulation of this chapter). When the sample is highly structured, the illumination can be simple to satisfy the convolution condition. As investigated in section 6.2, a single crystal specimen needs larger size condenser aperture to fulfil the reconstruction, while a sample has complicated structure needs a smaller condenser aperture. In an experiment, if the sample is weakly scattering, a structured illumination, which may be generated by a diffuser or by the aberration correctors (instead of reducing the aberrations, but increasing the aberrations), may improve the reconstruction.

Because of the huge information redundancy in the ptychographic dataset, which comes from the oversampling in real space and reciprocal space, we can not only retrieve the object function and the probe function via iterative ptychographic algorithms, but also can retrieve the detector response function and the coherence function with some methods like background matching (Chapter 2.7) and modal decomposition (Thibault and Menzel, 2014; Chapter 2.7). With an accurate model, for example, a known mask, it is also possible that to do the reconstruction of a partial dataset with the super-resolution algorithm (Maiden et al.,

2011). In section 6.3, we explored the convergence angle conditions of PIE on the partial datasets. PIE/ePIE has a high tolerance on data truncation and detector pedestal.

Chapter 7 Conclusion and future work

This thesis includes the exploration and application of large probe ptychography with the electron waves on a TEM, SEM, and STEM. We measured the coherence function of the electrons source completely via ptychography in the SAD mode; the experimental and computational details are presented in Chapter 3. We diagnosed and improved the delocalization issue in the reconstruction from the data collected on a SEM; the problems and the corresponding methods are included in Chapter 4. We tried to achieve atomic resolution via ptychography on a STEM; the problems and the results are included in Chapter 5. In the following, we summarize the characterizations, key points, and the future developments of SAD ptychography, SEM/STEM ptychography, and the algorithms.

In the SAD ptychography, the illumination can be considered as plane wave; thus, the electron dose on the sample per area is much lower than the focused or defocused setups. As investigated in Chapter 3, the counts can be as low as $5 e^-/A^2$ when reconstructed with ePIE. This makes the SAD ptychography a promising setup to measure the biological or organic samples. There are two main limits of the SAD ptychography; one is that the sample is exposed to the electrons beam continuously; thus, the electrons dose has to be very low for some beam-sensitive samples; the other is that the reconstructed resolution is low. As it has been investigated in Chapter 3, a structured beam makes the algorithms easier to retrieve the

illumination imperfection. The structured beam will also give more informational diffraction patterns. So, in the future, using a structured beam will be one method to improve the resolution of the SAD ptychography.

Atomic resolution can be achieved via SEM/STEM defocused probe ptychography; this has been proved in Chapter 4 and Chapter 5. The illumination imperfection (aberrations and partial coherence) will not affect the resolution of the reconstruction since they can be retrieved by the algorithmic methods that were proposed in the thesis. The main limit of the defocused probe ptychography at the current is to obtain the reliable scanning positions. Even though the errors in the scanning positions can be corrected by the annealing method, it may introduce the positions-defocus coupling issue during the reconstruction. This issue has been discussed in Chapter 5. Structured illumination that completely breaks the phase curvature of the wave front will be free of the positions-defocus coupling problem. So, in the near future, defocused probe ptychography on a designed STEM/SEM setup with the structured illumination will be promising to achieve reliable atomic resolution.

The algorithms we have used for the reconstruction in the thesis are mainly developed from the iterative algorithms PIE and ePIE. For the reconstruction of the conventional ptychographic data either collected from the SAD mode or SEM/STEM, we applied the annealing method Jiggle to correct the random scanning positions errors, and we applied either the extra mode method or the background blind matching method to remove the detector pedestal. To eliminate the influence of the electron source transverse partial coherence, for the SAD data, we applied either the Gaussian blind convolution method or the modal decomposition method; for the SEM/STEM data, we applied and only can apply the modal decomposition method. If the data is truncated, as the data we have processed in Chapter 4, we apply the super-resolution method to recover the truncated part of the diffraction patterns. As we have explored in Chapter 6, the ePIE/PIE algorithm has a low

tolerance for the Poisson noise. When the counts of the data are low, we need to modify the constraint in the Fourier domain. In Chapter 6, we tried to amend the Fourier constraint with the combination of the detected values and the calculated values, instead of completely with the detected values. The reconstruction got better but not perfect. As electron ptychography becomes a technique to measure the biological and organic samples, much lower dose has to be used to avoid damage the sample (lower than $1 e^-/A^2$); thus, more efficient methods to deal with the noisy data need to be developed in the future.

For all the calculations in this thesis, we supposed that the samples as two-dimensional objects. However, this approximation is not suitable for the sample we used in Chapter 4 and Chapter 5 when attempting to achieve atomic resolution. When the electrons beam goes through the gold particles of several tens nanometers, the electrons will be scattered by multiply, in this case, dynamical scattering has to be considered in our calculation model. The multislice algorithm is a method to model the multiple elastical scattering inside the sample. In the following research, we will apply the multislice ePIE with the electron data to reconstruct the three-dimensional structure of both the crystal and noncrystal samples. The other issue we have not considered in the calculations of this thesis is the inelastic scattering phenomenon. We assumed that the energy of the electrons kept constant when traveling through the sample. In fact, there will be a certain amount of energy loss when the electrons interact with the sample. We need to build the accurate electrons and atoms interaction models to reconstruct the resolution and the location of the atoms precisely.

In recent years, some progress has been achieved in both the two ways to implement electron ptychography. Compared with the large probe electron ptychography, the focused probe ptychography has the advantages of simultaneous collection of HAADF image, EELs or EDX spectrum. However, the large probe electron ptychography can image a much larger

field of view and has a much lower dose. This is very important when finally electron ptychography comes to the three-dimensional imaging like tomo-ptychography.

Since the first application of ptychography in the X-Ray imaging field, it only took about ten years for ptychography to become a standard imaging method. While about electron ptychography, the study still stays at the stage of the technique development. In this thesis, we have done plenty investigation about both the experiments and the algorithms of large probe electron ptychography. We have made big improvements to solve the existing problems (Chapter 4), and we have investigated some issues in large probe ptychography that have not been proposed before (Chapter 3, 5 and 6). We believe that the research in this thesis is a necessary step before large probe electron ptychography is used a standard imaging technique. The most important result in this thesis is the modal decomposition of matter waves, which has never been done before.

References

- Bates, R. H. T. and Rodenburg, J. M. (1989) ‘Sub-Angstrom transmission microscopy: A fourier transform algorithm for microdiffraction plane intensity information’, *Ultramicroscopy*, 31(3), pp. 303–307. doi: 10.1016/0304-3991(89)90052-1.
- Batey, D. J. *et al.* (2014) ‘Reciprocal-space up-sampling from real-space oversampling in x-ray ptychography’, *Physical Review A - Atomic, Molecular, and Optical Physics*, 89(4). doi: 10.1103/PhysRevA.89.043812.
- Batey, D. J., Claus, D. and Rodenburg, J. M. (2014) ‘Information multiplexing in ptychography’, *Ultramicroscopy*. Elsevier, 138, pp. 13–21. doi: 10.1016/j.ultramic.2013.12.003.
- Beleggia, M. *et al.* (2004) ‘On the transport of intensity technique for phase retrieval’, *Ultramicroscopy*, 102, p. 37.
- Bian, L. *et al.* (2014) ‘Content adaptive illumination for Fourier ptychography’, *Optics Letters*, 39(23), p. 6648. doi: 10.1364/OL.39.006648.
- Bian, Z., Dong, S. and Zheng, G. (2013) ‘Adaptive system correction for robust Fourier ptychographic imaging.’, *Optics express*, 21(26), pp. 32400–10. doi: 10.1364/OE.21.032400.
- Born, M. and Wolf, E. (1999) *Principles of Optics*. Cambridge: Cambridge University Press.

Bostan, Emrah et al. (2014) ‘Phase Retrieval by using transport-of-intensity equation and differential interference contrast microscopy’, *IEEE*, p. 3939.

Buban, J. P. et al. (2010) ‘high-resolution low-dose scanning transmission electron microscopy’, *Journal of Electron Microscopy*, 59(2), pp. 103.

Burdet, N. et al. (2015) ‘Evaluation of partial coherence correction in X-ray ptychography’, *Optics Express*, 23(5), p. 5452.

Cao, S. et al. (2016) ‘Modal decomposition of a propagating matter wave via electron ptychography’, *Physical Review A*, 94(6), pp. 1–8. doi: 10.1103/PhysRevA.94.063621.

Carlson, D. B and Evans, J. E. (2012) ‘Low dose imaging techniques for Transmission Electron Microscopy’, *The transmission electron microscope*, Chapter 5.

Clare, R. M. et al. (2015) ‘Characterization of near-field ptychography’, *Optics Express*, 23(15), p. 19728. doi: 10.1364/OE.23.019728.

Clark, J. N. et al. (2012) ‘High-resolution three-dimensional partially coherent diffraction imaging’, *Nature Communications*, 3(5), p. 993. doi: 10.1038/ncomms1994.

Clark, J. N. and Peele, A. G. (2011) ‘Simultaneous sample and spatial coherence characterisation using diffractive imaging’, *Applied Physics Letters*, 99(15), pp. 2013–2016. doi: 10.1063/1.3650265.

D’Alfonso, A. J. et al. (2014) ‘Deterministic electron ptychography at atomic resolution’, *Phys. Rev. B*, 89, p. 064101-1.

Danev, R. and Nagayama, K. (2001) ‘Transmission electron microscopy with Zernike phase plate’, 88(September 2000), pp. 243–252. doi: 10.1007/978-3-540-85156-1.

Dierolf, M. et al. (2010) ‘Ptychographic X-ray computed tomography at the nanoscale.’, *Nature*. Nature Publishing Group, 467(7314), pp. 436–9. doi: 10.1038/nature09419.

Donnelly, C. et al. (2017) ‘Three-dimensional magnetization structures revealed with X-ray vector nanotomography’, *Nature*, 547, p. 328.

Downing, R. M. (1986) ‘Improvement in high resolution image quality of radiation-sensitive specimens achieved with reduced spot size of the electron beam’, *Ultramicroscopy*, 20, pp. 269.

Van Dyck, D. and Chen, F.-R. (2012) ““Big Bang” tomography as a new route to atomic-resolution electron tomography.’, *Nature*, 486(7402), pp. 243–246. doi: 10.1038/nature11074.

Van Dyck, D. and De Jong, A. F. (1992) ‘Ultimate resolution and information in electron-microscopy - General-principles’, *Ultramicroscopy*, 47(1–3), pp. 266–281.

Edo, T. B. et al. (2013) ‘Sampling in x-ray ptychography’, *Physical Review A - Atomic, Molecular, and Optical Physics*, 87(5), pp. 1–8. doi: 10.1103/PhysRevA.87.053850.

Edo, T. B., Zhang, F. and Rodenburg, J. M. (2010) ‘Resolution improvement in coherent diffractive imaging (ptychography)’, *Spie* 7729, p. 77291H–77291H–7. doi: 10.1117/12.853346.

Egerton, R. F., Li, P. and Malac, M. (2004) ‘Radiation damage in the TEM and SEM’, *Micron*, 35(6), pp. 399–409. doi: 10.1016/j.micron.2004.02.003.

Elser, V. (2003) ‘phase retrieval by iterated projections’, *Journal of the Optical Society of America A*, 20, p. 40.

Enders, B. et al. (2014) ‘Ptychography with broad-bandwidth radiation’, *Applied Physics Letters*, 104(17). doi: 10.1063/1.4874304.

Faulkner, H. M. L. and Rodenburg, J. M. (2004) ‘Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm’, *Physical Review Letters*, 93(2), pp. 23903–1. doi: 10.1103/PhysRevLett.93.023903.

Faulkner, H. M. L. and Rodenburg, J. M. (2005) ‘Error tolerance of an iterative phase retrieval algorithm for moveable illumination microscopy’, *Ultramicroscopy*, 103(2), pp. 153–164. doi: 10.1016/j.ultramic.2004.11.006.

Fienup, J. R. (1978) ‘Reconstruction of an object from the modulus of its Fourier transform’, *Optics Lett.*, 3(1), p. 27.

Fienup, J. R. and Kowalczyk, A. M. (1990) ‘Phase retrieval for a complex-valued object by using a low-resolution image’, *Journal of the Optical Society of America A*, 7(3), p. 450. doi: 10.1364/JOSAA.7.000450.

Fienup, J. and Wackerman, C. C. (1986) ‘Phase retrieval stagnation problems and solutions’, *Journal of Optical Society of America A*, 3(11), pp. 1897–1907. doi: 10.1364/JOSAA.3.001897.

Fienup, M., Guizar-Sicairos, M. and Fienup, J. R. (2009) ‘Measurement of coherent x-ray focused beams by phase retrieval with transverse translation diversity’, *Optics Express*, 17(4), pp. 2670–2685. doi: 10.1364/OE.17.002670.

Frank, J. (1973) ‘The envelop of electron microscopic transfer functions for partially coherent illumination’, *Optik*, 38, pp. 519–536.

Frank, J. (2006) *Three-Dimensional Electron Microscopy of Macro-molecular Assemblies, in Visualization of Biological Molecules in Their Native State*. Oxford: Oxford University Press.

Gabor, D. (1948) ‘A new microscopic principle’, *Nature*, 161, p. 777.

Gabor, D., Kock, W. E. and Stroke, G. W. (1971) 'Holography', *Science*, 173, p. 11.

Gallagher, W. J. (1997) 'Magnetic memory array using magnetic tunnel junction devices in the memory cells', *United States Patent*, 5,640,343.

Gerchberg, R. W. (1974) 'Super-resolution through Error Energy Reduction', *Optica Acta: International Journal of Optics*, 21(9), pp. 709–720. doi: 10.1080/713818946.

Goodman, J. W. (2004), *Introduction to Fourier Optics*.

Gerchberg, R. W. and Saxton, W. O. (1972) 'A practical algorithm for the determination of phase from image and diffraction plane pictures', *Optik*, 35(2), pp. 237–246. doi: 10.1070/QE2009v039n06ABEH013642.

Gravel, S. and Elser, V. (2008) 'Divide and concur: A general approach to constraint satisfaction', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 78(3), pp. 1–5. doi: 10.1103/PhysRevE.78.036706.

Greenbaum, A. et al. (2013) 'Increased space-bandwidth product in pixel super-resolved lensfree on-chip microscopy', *Scientific Report*, 3, pp. 1.

Guizar-Sicairos, M. and Fienup, J. R. (2008) 'Image reconstruction by phase retrieval with transverse translation diversity', *Optics Express*, 16(10), p. 70760A. doi: 10.1117/12.792622.

Guizar-Sicairos, M., Thurman, S. T. and Fienup, J. R. (2008) 'Efficient subpixel image registration algorithms.', *Optics letters*, 33(2), pp. 156–158. doi: 10.1364/OL.33.000156.

Hegerl, R. and Hoppe, W. (1970) 'Dynamische Theorie der Kristallstrukturanalyse durch Elektronenbeugung im inhomogenen Prismaerstrahlwellenfeld', *Berichte der Bunsengesellschaft fur physikalische Chemie*, 74, p. 1089.

Holler, M. et al. (2017) 'High-resolution non-destructive three-dimensional imaging of

integrated circuits', *Nature*, 543(7645), pp. 402–406. doi: 10.1038/nature21698.

Hoppe, W. (1969a) 'Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen', *Acta Crystallographica Section A*, 25(4), pp. 495–501. doi: 10.1107/S0567739469001045.

Hoppe, W. (1969b) 'Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und Phasenbestimmung bei unperiodischen Objekten', *Acta Crystallographica Section A*, 25(4), pp. 508–514. doi: 10.1107/S0567739469001069.

Hoppe, W. and Strube, G. (1969) 'Beugung in inhomogenen Primärstrahlenwellenfeld. II. Lichtoptische Analogieversuche zur Phasenmessung von Gitterinterferenzen', *Acta Crystallographica Section A*. International Union of Crystallography, 25(4), pp. 502–507. doi: 10.1107/S0567739469001057.

Horstmeyer, R. and Yang, C. (2014) 'A phase space model of Fourier ptychographic microscopy', *Optics Express*, 22(1), p. 338. doi: 10.1364/OE.22.000338.

Huang, X. et al. (2015) 'Fly-scan ptychography', *Scientific Report*, 5, p. 1.

Hüe, F. et al. (2010) 'Wave-front phase retrieval in transmission electron microscopy via ptychography', *Physical Review B - Condensed Matter and Materials Physics*, 82(12), pp. 1–4. doi: 10.1103/PhysRevB.82.121415.

Hüe, F. et al. (2011) 'Extended ptychography in the transmission electron microscope: Possibilities and limitations', *Ultramicroscopy*, 111(8), pp. 1117–1123. doi: 10.1016/j.ultramic.2011.02.005.

Humphry, M. J. et al. (2012) 'Ptychographic electron microscopy using high-angle dark-field scattering for sub-nanometre resolution imaging', *Nature Communications*. Nature

Publishing Group, 3, p. 730. doi: 10.1038/ncomms1733.

Humphry, M. J. *et al.* (2012) ‘Ptychographic transmission microscopy in three dimensions using a multi-slice approach’, *Journal of the Optical Society of America A*, 29(8), p. 1606. doi: 10.1364/JOSAA.29.001606.

Hurst, a C. *et al.* (2010) ‘Probe position recovery for ptychographical imaging’, *Journal of Physics: Conference Series*, 241, p. 12004. doi: 10.1088/1742-6596/241/1/012004.

Hytch, M. *et al.* (2008) ‘Nanoscale holographic interferometry for strain measurements in electronic devices’, *Nature*, 453(19), p. 1086.

Ishizuka, K. and Allman, B. (2005) ‘Phase measurement in electron microscopy using transport of intensity equation’, *Microscopy Today*, p. 22.

James, E. M. and Rodenburg, J. M. (1997) ‘A method for measuring the effective source coherence in a field emission transmission electron microscope’, *Applied Surface Science*, 111, pp. 174–179. doi: 10.1016/S0169-4332(96)00697-6.

de Jong, A. F. and Van Dyck, D. (1993) ‘Ultimate resolution and information in electron microscopy II. The information limit of transmission electron microscopes’, *Ultramicroscopy*, 49(1–4), pp. 66–80. doi: 10.1016/0304-3991(93)90213-H.

Kirz, J. and Attwood, D. (2015), *X-ray Data Booklet*, Section 4.4.

Kirkpatrick, P. and Baez, A. V. (1948), ‘Formation of optical images by X-rays’, *Journal of the Optical Society of America*, 38(9), p. 1606.

Lebeau, J. M. *et al.* (2008) ‘Quantitative atomic resolution scanning transmission electron microscopy’, *Physical Review Letters*, 100(20), pp. 1–4. doi: 10.1103/PhysRevLett.100.206101.

Li, P. *et al.* (2015) ‘Separation of three-dimensional scattering effects in tilt-series Fourier ptychography’, *Ultramicroscopy*. Elsevier, 158, pp. 1–7. doi: 10.1016/j.ultramic.2015.06.010.

Li, P., Edo, T., *et al.* (2016) ‘Breaking ambiguities in mixed state ptychography’, *Optics Express*, 24(8), p. 9038. doi: 10.1364/OE.24.009038.

Li, P., Batey, D. J., *et al.* (2016) ‘Multiple mode x-ray ptychography using a lens and a fixed diffuser optic’, *Journal of Optics*. IOP Publishing, 18(5), p. 54008. doi: 10.1088/2040-8978/18/5/054008.

Li, P., Edo, T. B. and Rodenburg, J. M. (2014) ‘Ptychographic inversion via Wigner distribution deconvolution: Noise suppression and probe design’, *Ultramicroscopy*. Elsevier, 147, pp. 106–113. doi: 10.1016/j.ultramic.2014.07.004.

Li, P. and Maiden, A. M. (2017) ‘Ten implementations of ptychography’, *Journal of Microscopy*, 0(0), pp. 1–8. doi: 10.1111/jmi.12614.

Lin, J. A. and Cowley, J. M. (1986) ‘Reconstruction from in-line electron holograms by digital processing’, *Ultramicroscopy*, 19, p. 179.

Lubk, A. and Zweck, J. (2015) ‘Differential phase contrast: An integral perspective’, *Physical Review A - Atomic, Molecular, and Optical Physics*, 91(2), pp. 1–6. doi: 10.1103/PhysRevA.91.023805.

Lichte, H. (1986) ‘Electron holography approaching atomic resolution’, *Ultramicroscopy*. Elsevier, 20(3), pp. 293-304.

Marchesini, S. et al. (2003) ‘X-ray image reconstruction from a diffraction pattern alone’, *Phys. Rev. B*, 68, p. 140101-1.

Marchesini, S. and Schirotzek, A. (2012) ‘Augmented projections for ptychographic imaging’,

Inverse Problem, 29, p. 1.

Maiden, A. M. et al. (2011) ‘Superresolution imaging via ptychography.’, *Journal of the Optical Society of America A*, 28(4), pp. 604–612. doi: 10.1364/JOSAA.28.000604.

Maiden, A. M. et al. (2012) ‘An annealing algorithm to correct positioning errors in ptychography’, *Ultramicroscopy*, 120, pp. 64–72. doi: 10.1016/j.ultramic.2012.06.001.

Maiden, A. M. et al. (2013) ‘Soft X-ray spectromicroscopy using ptychography with randomly phased illumination.’, *Nature communications*. Nature Publishing Group, 4, p. 1669. doi: 10.1038/ncomms2640.

Maiden, A. M. et al. (2015) ‘Quantitative electron phase imaging with high sensitivity and an unlimited field of view’, *Sci Rep.* Nature Publishing Group, 5, p. 14690. doi: 10.1038/srep14690.

Maiden, A. M. and Rodenburg, J. M. (2009) ‘An improved ptychographical phase retrieval algorithm for diffractive imaging’, *Ultramicroscopy*. Elsevier, 109(10), pp. 1256–1262. doi: 10.1016/j.ultramic.2009.05.012.

Maiden, A. M., Rodenburg, J. M. and Humphry, M. J. (2010) ‘A new method of high resolution, quantitative phase scanning microscopy’, *Proceedings of SPIE - The International Society for Optical Engineering*, 7729(0). doi: 10.1117/12.853339.

Maiden, A. M., Rodenburg, J. M. and Humphry, M. J. (2010) ‘Optical ptychography: a practical implementation with useful resolution.’, *Optics letters*, 35(15), pp. 2585–2587. doi: 10.1364/OL.35.002585.

McBride, W., O’Leary, N. L. and Allen, L. J. (2004) ‘retrieval of a complex-valued object from its diffraction pattern’, *Phys. Rev. Lett.*, 93, p. 233902-1.

McCallum, B. C. and Rodenburg, J. M. (1993) ‘Simultaneous reconstruction of object and aperture functions from multiple far-field intensity measurements’, *Journal of the Optical Society of America A*, 10(2), p. 231. doi: 10.1364/JOSAA.10.000231.

McCartney, M. R. and Smith, D. J. (2007) ‘Electron holography: Phase imaging with nanometer resolution’, *Annual Review of Materials Research*, 37, pp. 729–767. doi: 10.1146/annurev.matsci.37.052506.084219.

McVitie, S. and Ngo, D. T. (2008) ‘Quantitative measurements of phase using the transport of intensity equation’, *Journal of Physics: Conference Series*, 126, p. 12041. doi: 10.1088/1742-6596/126/1/012041.

Miao, J., Charalambous, P., Kirz, J. and Sayre, D. (1999) ‘Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens’, *Nature*, 400, p. 342-344.

Midgley., P. A. (2001) ‘An introduction to off-axis electron holography’, *Micron*, 32, p. 167.

Morgan, A. J. et al. (2013) ‘fast deterministic single-exposure coherent diffractive imaging at sub-angstrom resolution’, *Phys. Rev. B*, 87, p. 094115-1.

Nellist, P. D., McCallum, B. C. and Rodenburg, J. M. (1995) ‘Resolution beyond the information limit in transmission electron microscopy’, *Nature*, 374, p. 630.

Nellist, P. D. and Rodenburg, J. M. (1993) ‘Image resolution improvement using coherent microdiffraction in STEM’, *Inst Phys Conf Ser*, 138, p. 239.

Ou, X. et al. (2013) ‘Quantitative phase imaging via Fourier ptychographic microscopy.’, *Optics letters*, 38(22), pp. 4845–8. doi: 10.1364/OL.38.004845.

Ou, X. et al. (2015) ‘High numerical aperture Fourier ptychography : principle ,

implementation and characterization', *Optics express*, 23(3), pp. 5473–5480. doi: 10.1364/OE.23.003472.

Pelz, P. M. et al. (2017) 'Low dose cryo electron ptychography via non-convex Bayesian optimization', *Scientific Reports*, pp. 1–8.

Pennycook, T. J. et al. (2015) 'Efficient phase contrast imaging in STEM using a pixelated detector. Part 1: Experimental demonstration at atomic resolution', *Ultramicroscopy*, 151, pp. 160–167. doi: 10.1016/j.ultramic.2014.09.013.

Putkunz, C. T. et al. (2012) 'Atom-scale ptychographic electron diffractive imaging of boron nitride cones', *Physical Review Letters*, 108(7), pp. 1–4. doi: 10.1103/PhysRevLett.108.073901.

Robisch, A. L. et al. (2015) 'Near-field ptychography using lateral and longitudinal shifts', *New Journal of Physics*. IOP Publishing, 17(7), p. 73033. doi: 10.1088/1367-2630/17/7/073033.

Rodenburg, J. (2011) 'New microscopic-imaging method delivers novel capabilities', *SPIE Newsroom*, pp. 2–4. doi: 10.1117/2.1201012.003414.

Rodenburg, J. M. (1989) 'The phase problem, microdiffraction and wavelength-limited resolution - a discussion', *Ultramicroscopy*, 27(4), pp. 413–422. doi: 10.1016/0304-3991(89)90009-0.

Rodenburg, J. M. and Faulkner, H. M. L. (2004) 'A phase retrieval algorithm for shifting illumination', *Applied Physics Letters*, 85(20), pp. 4795–4797. doi: 10.1063/1.1823034.

Rodenburg, J. M. et al. (2007) 'Hard-X-ray lensless imaging of extended objects', *Physical Review Letters*, 98(3), pp. 1–4. doi: 10.1103/PhysRevLett.98.034801.

Rodenburg, J. M. (2008) ‘Ptychography and related diffractive imaging methods’, *Advances in Imaging and Electron Physics*, 150(7), pp. 87–184. doi: 10.1016/S1076-5670(07)00003-1.

Rodenburg, J. M. and Faulkner, H. M. L. (2004) ‘A phase retrieval algorithm for shifting illumination’, *Applied Physics Letters*, 85(20), pp. 4795–4797. doi: 10.1063/1.1823034.

Rodenburg, J. M., Hurst, A. C. and Cullis, A. G. (2007) ‘Transmission microscopy without lenses for objects of unlimited size’, *Ultramicroscopy*, 107(2–3), pp. 227–231. doi: 10.1016/j.ultramic.2006.07.007.

Ross, F. M. and Stobbs, W. M. (1991) ‘The characterisation of GaAs/(Al,Ga)As heterostructure interface roughness using Fresnel analysis’, *Ultramicroscopy*, 36(4), pp. 331–354. doi: 10.1016/0304-3991(91)90125-P.

Roy, S. *et al.* (2011) ‘Lensless X-ray imaging in reflection geometry’, *Nature Photonics*, 5(4), pp. 243–245. doi: 10.1038/nphoton.2011.11.

Sakulich, A. R. and Li, V. C. (2011) ‘Microscopic Characterization of Autogenous Healing Products in Engineered Cementitious Composites (Ecc)’, *33rd Int'l Conf on Cement Microscopy, San Francisco, CA USA*, 16(Suppl 2), pp. 1–10. doi: 10.1017/S14319276100.

Seldin, J. H. and Fienup, J. R. (1990) ‘Numerical investigation of the uniqueness of phase retrieval’, *Journal of the Optical Society of America A*, 7(3), p. 412. doi: 10.1364/JOSAA.7.000412.

Shibata, N. *et al.* (2012) ‘Differential phase-contrast microscopy at atomic resolution’, *Nature Physics*. Nature Publishing Group, 8(8), pp. 611–615. doi: 10.1038/nphys2337.

Spence, J. C. H. (2003) *High-Resolution Electron Microscopy*. Oxford University Press.

Stockmar, M. et al. (2012) ‘Near-field ptychography; phase retrieval for inline holography

using a structured illumination', *Scientific Report*, 3, p. 1.

Suvorov, A. A. (2010) 'On formation of a partially coherent beam in a stable-resonator laser', *Quantum Electronics*, 40(3), pp. 210.

Taylor, B. et al. (1994) 'The van Cittert-Zernike theorem in atom optics', *Optics Communications*, 110(5–6), pp. 569–575. doi: 10.1016/0030-4018(94)90252-6.

Thibault, P. et al. (2006) 'Reconstruction of a yeast cell from X-ray diffraction data', *Acta Crystallographica Section A: Foundations of Crystallography*, 62(4), pp. 248–261. doi: 10.1107/S0108767306016515.

Thibault, P. (2008) 'High-Resolution Scanning X-ray', *Science*, 379(July), pp. 379–383. doi: 10.1126/science.1158573.

Thibault, P. et al. (2009) 'Probe retrieval in ptychographic coherent diffractive imaging', *Ultramicroscopy*. Elsevier, 109(4), pp. 338–343. doi: 10.1016/j.ultramic.2008.12.011.

Thibault, P. and Elser, V. (2010) 'X-Ray Diffraction Microscopy', *Annual Review of Condensed Matter Physics*, 1(1), pp. 237–255. doi: 10.1146/annurev-conmatphys-070909-104034.

Thibault, P. and Guizar-Sicairos, M. (2012) 'Maximum-likelihood refinement for coherent diffractive imaging', *New Journal of Physics*, 14. doi: 10.1088/1367-2630/14/6/063004.

Thibault, P., Guizar-Sicairos, M. and Menzel, A. (2014) 'Coherent imaging at the diffraction limit', *Journal of Synchrotron Radiation*. International Union of Crystallography, 21(5), pp. 1011–1018. doi: 10.1107/S1600577514015343.

Thibault, P. and Menzel, A. (2013) 'Reconstructing state mixtures from diffraction measurements', *Nature*. Nature Publishing Group, 494(7435), pp. 68–71. doi:

10.1038/nature11806.

Tripathi, A., McNulty, I. and Shpyrko, O. (2014) ‘Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods’, *Optics Express*, 22(2), pp. 1452–1466. doi: 10.1364/OE.22.001452.

Vandewalle, P., Süssstrunk, S. and Vetterli, M. (2003) ‘Superresolution images reconstructed from aliased images’, *Proceedings of SPIE - The International Society for Optical Engineering*, 5150 II, pp. 1398–1405. doi: 10.1117/12.506874.

Verbeeck, J., Béché, A. and Van den Broek, W. (2012) ‘A holographic method to measure the source size broadening in STEM’, *Ultramicroscopy*, 120, pp. 35–40. doi: 10.1016/j.ultramic.2012.05.007.

Vine, D. J. *et al.* (2009) ‘Ptychographic fresnel coherent diffractive imaging’, *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(6), pp. 1–5. doi: 10.1103/PhysRevA.80.063823.

Voelz, D. (2011) *Computational Fourier Optics*, SPIE press.

Volkov, V. V., Wall, J. and Zhu, Y. (2008) ‘Position-sensitive diffractive imaging in STEM by an automated chaining diffraction algorithm’, *Ultramicroscopy*, 108(8), pp. 741–749. doi: 10.1016/j.ultramic.2007.11.007.

Whitehead, L. W. *et al.* (2009) ‘Diffractive imaging using partially coherent X rays’, *Physical Review Letters*, 103(24), pp. 1–4. doi: 10.1103/PhysRevLett.103.243902.

Wilke, R. N., Vassholz, M. and Salditt, T. (2013) ‘Semi-transparent central stop om high-resolution X-ray ptychography using Kirkpatrick-Baez focusing’, *Acta Cryst. A*, 69, p. 1.

Williams, D. B. and Carter, C. B. (2009) *Transmission Electron Microscopy*. Springer.

Wolf, E. (1982) 'New theory of partial coherence in the space-frequency domain. Part I: spectra and cross spectra of steady-state sources', *Journal of the Optical Society of America*, 72(3), pp. 343–351. doi: 10.1364/JOSA.72.000343.

Wolf, E. (2011) 'What kind of phases does one measure in usual interference experiments?', *Optics Communications*. Elsevier B.V., 284(19), pp. 4235–4236. doi: 10.1016/j.optcom.2011.04.064.

Yang, H. *et al.* (2016) 'Enhanced phase contrast transfer using ptychography combined with a pre-specimen phase plate in a scanning transmission electron microscope', *Ultramicroscopy*. Elsevier, 171, pp. 117–125. doi: 10.1016/j.ultramic.2016.09.002.

Yang, H. *et al.* (2016) 'Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures', *Nature Communications*. Nature Publishing Group, 7, p. 12532. doi: 10.1038/ncomms12532.

Yang, H. *et al.* (2017) 'Electron ptychographic phase imaging of light elements in crystalline materials using Wigner distribution deconvolution', *Ultramicroscopy*. Elsevier B.V., 180, pp. 173–179. doi: 10.1016/j.ultramic.2017.02.006.

Yang, H., Pennycook, T. J. and Nellist, P. D. (2015) 'Efficient phase contrast imaging in STEM using a pixelated detector. Part II: Optimisation of imaging conditions', *Ultramicroscopy*. Elsevier, 151, pp. 232–239. doi: 10.1016/j.ultramic.2014.10.013.

Zernike, F. (1938) 'The concept of degree of coherence and its application to optical problems', *Physica*, 5(8), pp. 785–795. doi: 10.1016/S0031-8914(38)80203-2.

Zhang, F. *et al.* (2013) 'Translation position determination in ptychographic coherent diffraction imaging', *Optics Express*, 21(11), p. 13592. doi: 10.1364/OE.21.013592.

Zheng, G. *et al.* (2013) ‘Wide-field, high-resolution Fourier ptychographic microscopy’, *Nature Photonics*, 7(9), pp. 739–745. doi: 10.1038/nphoton.2013.187.

Zuo, J. M. and Huang, W. (2012) ‘Coherent electron diffraction imaging’, *Handbook of Nanoscopy*, 13, p. 438.

Appendix

TEM Image alignment

The purpose of the alignment in TEM image mode is to make sure that the source, the lenses and the apertures are lined on the optic axis, and to correct the aberrations in the lenses. The procedures to align TEM image mode are:

1. Preparation.

- 1) Check whether the microscope runs normally;
- 2) Check whether the microscope is in the TEM image mode;
- 3) Turn on the field cancellation, which will monitor the environment of the microscope room;
- 4) Top up the liquid Nitrogen, which is used for cooling up the microscope column;
- 5) Insert the specimen;
- 6) Low or high flash the electron emission tip, to clean the contamination of the tip (do this step when necessary);
- 7) Add up voltage on the anode *A2* to 4.6 ~ 4.7 kv, which is the first anode to pull the electron beam down.
- 8) Add up voltage on the anode *A1* to 1.8~2.4 kv, which is the second anode to pull the electron beam down, to make sure that the illumination emission is about 2-3 kv;
- 9) Turn on the beam.

2. Align the condenser aperture.

- 1) Set magnification to 30~40K;

- 2) Spread the illumination on the specimen by turning up the strength of the condenser lens, which is the knob on the operation panel that controls the brightness;
 - 3) Insert a condenser aperture; shift the aperture in x and y directions to the centre of the fluorescent screen;
 - 4) Focus the illumination by turning down the condenser lens;
 - 5) Shift the illumination in x and y directions to the centre of the fluorescent screen;
 - 6) Repeat step 1) - 4) until the illumination stays in the centre of the screen when changing the strength of the condenser lens.
3. Align gun tilt (A2 wobbler).
 - 1) Use the ‘A2 wobbler’ to slightly wobble the voltage on the anode 2; if the gun is tilted, the illumination on the fluorescent screen will wobble asymmetrically;
 - 2) Tilt the gun with the ‘Gun Tilt’ on the operation panel in x and y directions while keeping A2 wobbling, until the illumination wobbles symmetrically; we may also need to shift the beam to make sure the illumination is not out of the screen;
 - 3) Turn off the A2 wobbler.
 4. Align gun shift.
 - 1) Take the source further from the specimen by turning up the ‘spot size’; there are 5 spot sizes on the R005; if the gun is not aligned on axis, the illumination on the fluorescent screen will have a big shift; if the gun is aligned on axis, the illumination will stay still;
 - 2) If the illumination shifts, use the beam shift knobs on the operation panel in x and y directions to bring it back to the centre on the screen;
 - 3) Bring the source closer to the specimen by turning down the ‘spot size’;
 - 4) Repeat step 2) to make the illumination stay in the centre of the screen;

- 5) Repeat step 1) – 4), until the illumination keeps still when changing the spot size.
5. Align beam tilt.
 - 1) Turn up the magnification to about 100K;
 - 2) Move the specimen in z direction to focus the specimen in focus;
 - 3) Tilt the beam in x direction; if the beam is not aligned, the illumination will shake around;
 - 4) Use the ‘Beam Tilt X’ knob to align the beam in x direction, until the illumination stops shaking;
 - 5) Do step 4) and 5) in y direction;
 - 6) Turn up to higher magnification ~600K, to repeat step 3) – 5).
6. High-tension (HT) wobbler.
 - 1) Set magnification to 600K;
 - 2) When we press the ‘HT wobbler’ button, it means that we are adding slight voltage fluctuations ($\sim \pm 250V$) to the accelerating voltage of 300KeV; if the optic is aligned on axis, the illumination will wobble symmetrically, and the features of the sample image will not shift; otherwise, the structures on the sample image will wobble asymmetrically.
 - 3) We use the ‘beam Tilt X’ and ‘Beam Tilt Y’ buttons to align the beam, until the features on the image of the sample wobble in centre.
7. Check the alignment from step 2 to 6.

These are the procedures of the manual alignment. After the manual alignment, then we refine the alignment with the image aberration corrector. This will be done by the software; we do not view the image on the fluorescent screen, but view the image on the detector. The basic procedures are as below.

- 1) Set the magnification to 600K; shift the specimen to make the beam illuminate a flat area on the specimen (e.g. a clean carbon support area);
- 2) In the DigitalMicrograph software window, turn on the ‘Live FFT’, which will show the live Fourier transform of the illuminated area;
- 3) Change the location of the specimen in z direction, to focus the specimen by observing the live Fourier transform: when the first diffraction ring has the largest diameter, the specimen is in focus;
- 4) Correct the astigmatism in the projective lenses manually, to make sure that the diffraction rings are round.
- 5) Change the location of the specimen in z direction, to slightly defocus the specimen (about -260nm); then use the image aberration corrector to correct the aberrations; Figure 3.2.4 shows the image aberration corrector standards we used in the experiments.

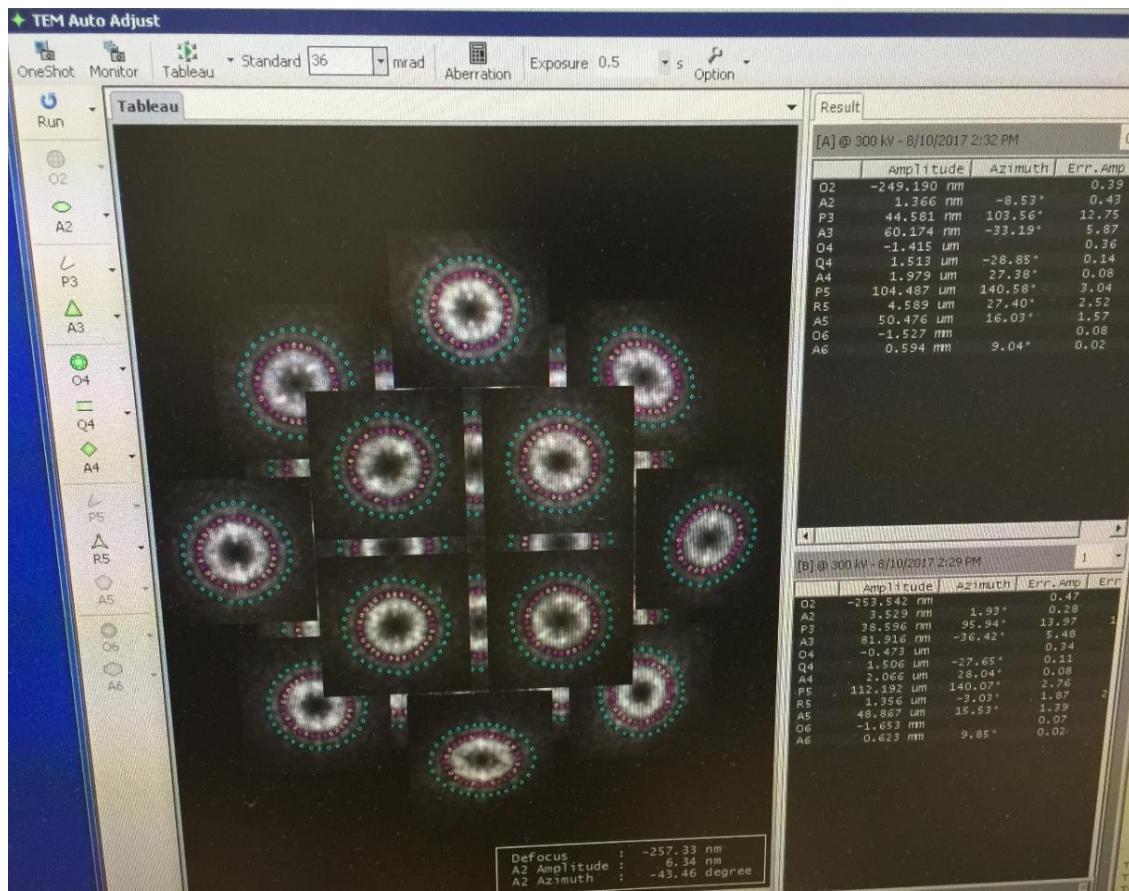


Figure 3.2.4 shows the TEM image aberration corrector standards that we have applied in the experiments.

SAD diffraction alignment

The purpose of the alignment in SAD mode is to make the detector is at the conjugate plane of the back focal plane of the objective lens, and to correct the aberrations in the projective lenses. The procedures to align the diffraction patterns are as below.

- 1) Change the microscope into diffraction mode from image mode; spread the illumination on the sample; now on the detector we should be able to see a shadow image of the sample;
- 2) Check the camera length; usually we align the diffraction pattern at a camera length of 15 cm;
- 3) Insert a large selected area aperture (the smallest one has a diameter of 10 um);
- 4) Align the selected area aperture; change the view to the fluorescent screen from the detector; shift the diffraction pattern to the centre of the screen;
- 5) Change the strength of the projective lens, to project the back focal plane onto the screen; on the screen, we will see a bright spot; insert the beam stop to block the very bright spot in the centre; now we should be able to see some diffraction rings; the diffraction rings may be blurred;
- 6) Insert a large ‘High Contrast’ aperture; move it to make sure we can see the edge of it on the detector (Orius SC200D).
- 7) Change the convergence angle of the condenser lens ‘C2’ and the strength of the projective lens at the same time to make sure that the diffraction rings and the edge of the HC aperture are as sharp as possible;
- 8) Use the image astigmatism corrector on the operation panel to correct the astigmatism in the projective lenses; make sure that the diffraction rings are round.

collect data

The procedures to collect ptychographic data are as below.

- 1) Set the microscope to the image mode, spot size 3, condenser aperture 2 (diameter of 100um); align the image; set magnification to 30-40K (Orius SC200D CCD, at magnification 30K, pixel pitch 0.35nm); Collect at least two sets of bright field images with the positions map that is designed for the ptychographic data collection, to test the step size and piezo stage shifting repeatability;
- 2) Change the microscope into the selected area diffraction mode; align the diffraction pattern; insert the smallest selected area aperture (10um); change the strength of the projective lens to make the detector project a defocused plane near the back focal plane of the objective lens; change the convergence angle of the condenser lens, to make sure that 1) the diffraction disc is inside the detector (binned 2); and 2) with an exposure time 0.1s – 3.0s, the counts of the diffraction pattern are 500 – 6000;
- 3) Collect the diffraction patterns with the calibrated scanning positions;
- 4) Remove the selected area aperture; collect a set of shadow images with the same scanning positions, cross correlated the shadow images to get the reference positions.

Here are some tips to collect various SAD ptychographic data.

1. When we defocus the projective lens,
 - 1) we can collect diffraction patterns at different defocus planes (any plane between the back focal plane and the image plane), by changing the strength of the projective lens with the knob ‘DIFF DEF’;
 - 2) we can collect data at either side of the projective lens – defocused plane or over focused plane;

2. when changing the convergence angle of the condenser lens with C2 knob, we will also change the coherence of the illumination;
3. when changing the spot size, the higher spot size means more coherent illumination;
4. we can use various diffraction planes for the ptychographic reconstruction, as long as the diffraction disc is located inside the detector;
5. Another option to change the coherence of the illumination and the radiation dose is to change the voltage of the anodes $A1$ and $A2$.