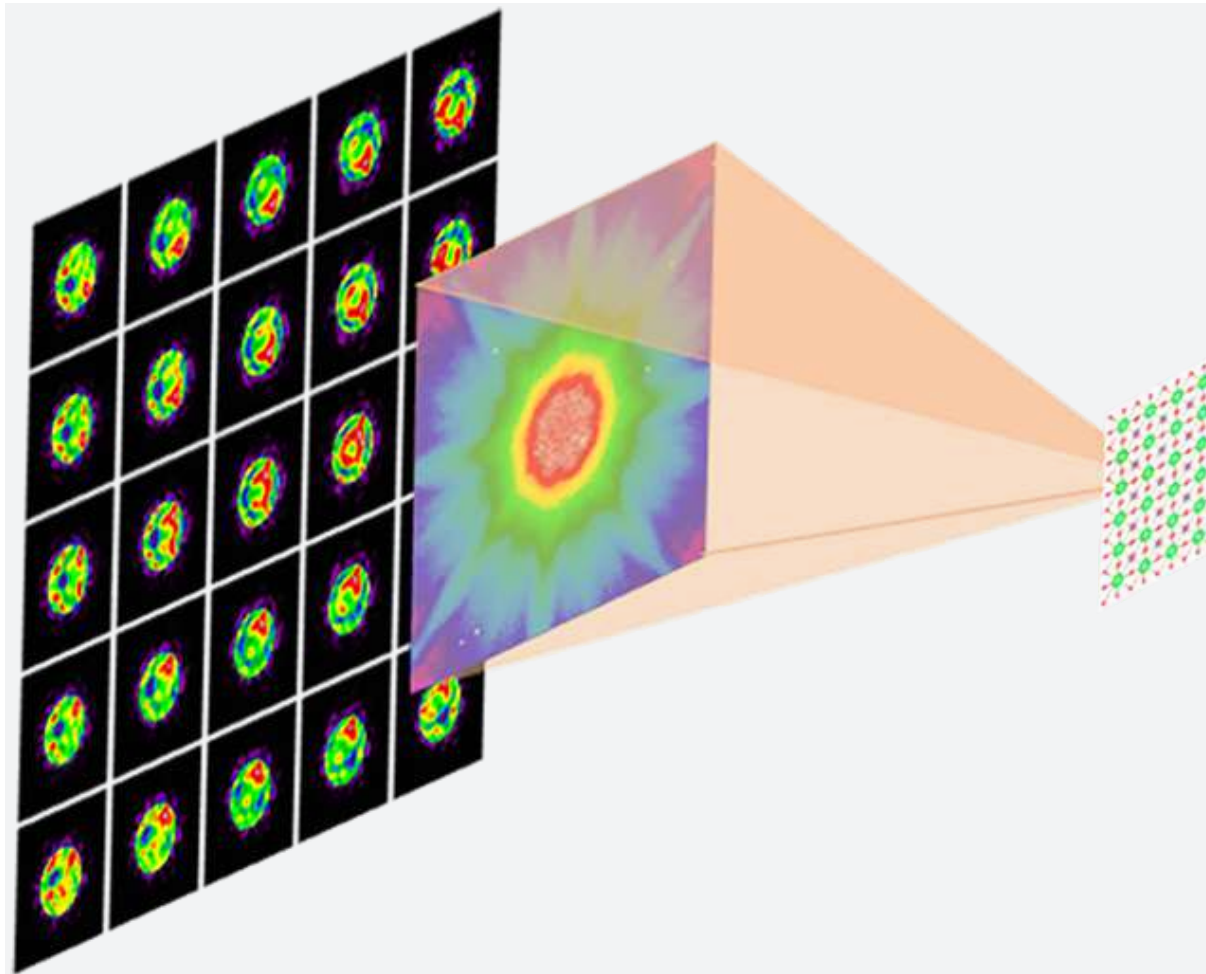


Exploring electron ptychography for low dose imaging

Chuang. Gao



Supervisor **Prof. Dr. Timothy. J. Pennycook**

Thesis submitted for the degree of Doctor of Science: Physics
Faculty of Science | Antwerpen, 2024



University
of Antwerp



Faculty of Science

Exploring electron ptychography for low dose imaging

Thesis submitted for the degree of
Doctor of Science: Physics
at the University of Antwerp

Chuang. Gao

Antwerpen, 2024

Supervisor
Prof. Dr. Timothy. J. Pennycook

Jury

Chairman

Prof. Dr. Sofie Cambré , University of Antwerp, Belgium

Supervisor

Prof. Dr. Timothy. J. Pennycook, University of Antwerp, Belgium

Members

Prof. Dr. Jan De Beenhouwer, University of Antwerp, Belgium

Prof. Dr. Peter Nellist, University of Oxford, United Kingdom

Prof. Dr. Maria Varela, Universidad Complutense de Madrid, Spain

Contact

Chuang. Gao

University of Antwerp

Faculty of Science

EMAT

Groenenborgerlaan 171, 2020 Antwerpen, België

M: chuang.gao@uantwerpen.be

T: +32 0484061838

© 2024 Chuang. Gao

All rights reserved.

ISBN 987-90-57285-34-7

Wettelijk depot D/2022/12.293/03



9 879057 285347

Dutch title:

Onderzoek naar
elektronenptychografie voor
beeldvorming met lage dosis

Preface

Transmission electron microscopy (TEM) is an important technique in the exploration of materials' structures. This is especially true since the development of electron optical aberration correctors greatly facilitated atomic resolution imaging. Generally speaking, there are two kinds of transmission electron microscopes: TEM, in which parallel beams of electrons are used to illuminate samples and form coherent images via the interference of transmitted and scattered electrons. The other general type of transmission electron microscope is a scanning transmission electron microscope (STEM), using a convergent beam to scan across samples and form coherent or incoherent images of samples depending on the geometrical structure of the detectors used. Many vital results have been achieved in materials science using electron microscopy. Nevertheless, there are still many challenges in this field due to the limitations of the imaging methods. For example, in TEM phase contrast images, the contrast of different spatial frequencies is determined by the combination of aberrations used. Unless a phase plate is used, conventional high resolution TEM (HRTEM) images show next to no contrast and thus the images generally depend on the use defocus and spherical aberration to produce contrast. Depending on the aberrations used, the same frequencies can appear bright or dark in HRTEM images. These and other effects lead to HRTEM phase contrast images generally being quite difficult to reliably interpret, generally requiring image simulations to ensure one sees what one thinks one sees in the images.

Among the imaging methods of the STEM, annular dark field imaging (ADF) is a powerful imaging method. ADF is also called Z-contrast imaging because it produces images with an intensity roughly proportional to the square of the atomic number, Z . This makes ADF images generally relatively easy to interpret, and often significantly more so than other imaging techniques, especially conventional TEM. However, one drawback of ADF, especially the high angle ADF (HAADF) with the purest Z-contrast, is that light atomic columns can be hidden by nearby heavy elements. Therefore imaging light atomic columns can be a significant challenge in ADF imaging in many materials. In addition, conventional STEM imaging methods are often not as dose efficient as TEM methods. This means fragile beam sensitive materials can be more easily damaged when striving to obtain sufficient signal. Although under ideal conditions the theoretical spatial resolution of the STEM is determined by accelerating voltage of the electron beam and the aperture size of the probe forming lens, in practice electron microscopists are limited by the fact that the electron beam damages many materials. This means that for many materials the maximum spatial resolution is not only limited by factors such as lens aberrations but the amount of signal that can be obtained before damage destroys the sample. Thus microscopists must seek to balance the conditions such that the most information is obtained for the dose budget of a given sample. This means that the theoretical electron optical spatial resolution is not always achievable. Therefore exploring new strategies to overcome beam damage is a hot topic in electron microscopy

research. These new strategies have to be highly dose efficient to provide improved imaging clarity in extremely low dose conditions.

In recent years, dose efficient electron ptychography has become an important cutting edge research topic due to the rapid development of direct electron detectors (DED) with faster recording speeds and higher dynamic ranges than before aided by improving computational power. Ptychography is a computational imaging method that can retrieve the phase changes imposed by an object on the electron wave from a set of recorded diffraction patterns. Compared to conventional TEM and STEM imaging methods, it has significant advantages such as:

1. Superresolution, in which image resolution goes beyond the conventional limit set by the convergence angle. This means that even with electron microscopes without an aberration corrector, high spatial resolution can still be achieved.
2. Very high dose efficiency, which can exceed that of HRTEM and other STEM methods. Researchers have successfully applied ptychography to beam-sensitive materials such as halide perovskites (HPs) and biological samples.
3. The residual aberrations of the probe can be removed via post collection aberration correction. This means even if the electron microscope is not well aligned, a high quality image can still be achieved with ptychography.
4. As a phase imaging method, it is highly sensitive to the electromagnetic fields of materials, and therefore also the charge density of a material. Changes in the charge density from charge transfer due to bonding have been detected with ptychography. Magnetic fields should also be detectable.

Currently, there are two approaches to ptychography. One is called direct inversion ptychography, requiring a focused probe to scan across the sample in a densely spaced grid. Then the wave function of the object can be derived from the recorded diffraction patterns with the phase or weak phase approximation. The other is called iterative ptychography, by using iterative optimization algorithms, the phase information of the object can be retrieved from the diffraction patterns. One thing that should be noted is that the probe for iterative ptychography can be either focused or highly defocused.

For focused probe ptychography, although DED cameras have improved greatly, relatively slow scan speeds have remained a substantial bottleneck for practical implementation with most cameras. Slow speeds generally result in data affected by sample drift and also tend to introduce contamination, substantially reducing the quality of the images one can obtain. In this thesis, by utilizing the event driven Timepix3 camera, the scan speed bottleneck is completely removed. The dwell time of the probe can be less than $1 \mu\text{s}$. We initially tested the camera with the two dimensional material WS_2 and zeolite samples at relatively low doses. The experimental results show, as expected, that the ptychographic method is more dose efficient than the HAADF, annular bright field imaging (ABF) and integrated center of mass (iCoM) methods with higher signal to noise ratios.

Most ptychographic algorithms are based on the phase approximation or weak phase approximation. These approximations assume that the samples are relatively thin. However most samples are thick enough to break these approximations, and sometimes

contrast reversals are observed for thicker materials. To overcome this barrier, multislice ptychography has been used, in which the dynamic scattering mechanism is considered explicitly. However, at low doses multislice ptychography sometimes does not converge. In this thesis, we propose methods that overcome contrast reversals for computationally simpler and faster single slice ptychographies, including physical defocusing, post collection defocus and what we call our phase offset method of contrast reversal correction. These generally allow single slice ptychography to perform remarkably well for thick samples despite the approximations used to motivate the techniques mathematically being violated.

Chapter 1 introduces the basic theoretical background of electron microscopy. From the source to the detector plane, a detailed derivation of how a diffraction pattern is formed is given. Then the concept of 4D STEM is introduced.

Chapter 2 reviews different kinds of ptychographic imaging methods. For direct inversion ptychography, single sideband (SSB) and Wigner distribution deconvolution (WDD) methods are introduced. For iterative ptychography, the extended ptychographic iterative engine (ePIE) and multislice ePIE methods are introduced and fundamental questions about the importance of phase retrievals are answered.

Chapter 3 delves into the application of conventional HAADF, coupled with electron energy loss spectroscopy (EELS), to unravel the layered structure of LaVO_3 and LaTiO_3 . The proximity of the atomic numbers of Ti and V poses a challenge for HAADF, as it cannot distinguish them based on the intensity of scattering electrons alone. EELS provides the complementary information for the Ti and V elements. The combination of these techniques proves to be successful in resolving the layered structure of the materials, as this material was sufficiently beam robust.

Chapter 4 introduces the use of the Timepix3 detector and the use of electron ptychography for low-dose operation. The Timepix3 is characterized by its distinctive event-driven mechanism, which enables a dwell time of less than $1 \mu\text{s}$. This is particularly advantageous for exploring beam-sensitive materials.

Chapter 5 discusses the application of electron ptychography in thicker materials typical for materials science, such as complex oxides like SrTiO_3 . A prevalent issue encountered in ptychography with thicker samples is contrast reversals. However, through careful experimentation and simulations, we discovered that by tuning the defocus of the probe we can often overcome these contrast reversals. Often what turns out to be best is what we call central focusing, in which the probe is physically focused to the central plane of the sample. However we can also use post-collection defocus, ptychographically changing the focus after collecting the data, which is particularly advantageous if we want to record an ADF image using probe at the optimal focus for ADF imaging. Our findings offer valuable insight into optimizing imaging conditions for thick specimens.

In Chapter 6, an integrated contrast transfer function (CTF) concept is employed to explain how central focusing can overcome contrast reversals. This conceptual framework provides an understanding of how tuning the defocus of the probe to the middle plane of the sample influences the contrast. We find this concept can not work for the strong phase objects.

In chapter 7, a further investigation to overcome contrast reversal is presented. We

find that phase wrap asymptotes in the frequency give rise to contrast reversal in SSB ptychography and by shifting the asymptotes to higher frequencies, the contrast reversal can be counteracted. With this knowledge we developed what we call the phase offset method of contrast reversal correction. This method is examined with simulated and experimental data, as is found to generally be superior to the application of a post-collection defocus.

In Chapter 8, a summary of the key findings and important conclusions derived from the present work. Proposed avenues for future research are outlined by identifying areas that warrant further exploration and development.

Samenvatting

Transmission electron microscopy (TEM) is een belangrijke techniek geworden in de verkenning van materialenstructuren met atomaire resolutie, vooral sinds de ontwikkeling van elektronenoptische aberratiecorrectoren. Over het algemeen zijn er twee soorten transmissie-elektronenmicroscopen; TEM, waarbij parallelle bundels van elektronen worden gebruikt om monsters te verlichten en coherente beelden te vormen via de interferentie van doorgelaten en verstrooide elektronen. Het andere algemene type transmissie-elektronenmicroscop is een scanningtransmissie elektronenmicroscop (STEM), die een convergente bundel gebruikt om over monsters te scannen en coherente of incoherente beelden van monsters te vormen, afhankelijk van de geometrische structuur van de detectoren. Een groot aantal belangrijke resultaten de materiaalwetenschap door elektronenmicroscopie. Desalniettemin zijn er nog steeds veel uitdagingen op dit gebied vanwege de beperkingen van de beeldvormingsmethoden. Bijvoorbeeld, in TEM wordt het fasecontrastbeeld sterk beïnvloed door de defocus en andere aberratiecoëfficiënten, wat kan leiden tot een moeilijke interpretatie van de beelden, over het algemeen vereist dit beeldsimulaties voor een betrouwbare interpretatie. Onder de beeldvormingsmethoden van de STEM is high-angle annular dark field imaging (HAADF) een krachtige beeldvormingsmethode die ook wel Z-contrast beeldvorming wordt genoemd omdat de intensiteit ruwweg evenredig is met het kwadraat van het atoomnummer, Z. Dit maakt HAADF-beelden gemakkelijker te interpreteren dan veel andere beeldvormingsmodaliteiten, vooral conventionele TEM. Echter, een nadeel van HAADF is dat lichte atomen verborgen kunnen worden door nabijgelegen zware atoomkolommen. Daarom is het afbeelden van lichte atoomkolommen een aanzienlijke uitdaging in HAADF. Daarnaast zijn conventionele STEM-beeldvormingsmethoden vaak niet zo dosisefficiënt als TEM-methoden. Dit betekent dat fragiele stralingsgevoelige materialen gemakkelijker beschadigd kunnen worden wanneer men streeft naar voldoende signaal. Hoewel de theoretische ruimtelijke resolutie van STEM alleen afhangt van de grootte van de opening van de convergerende lens die de probe vormt, moeten elektronenmicroscopisten in de praktijk een balans zoeken tussen de ruimtelijke resolutie en de kwaliteit van de beelden voor stralingsgevoelige materialen, omdat veel materialen slechts een bepaalde dosis kunnen verdragen voordat ze beschadigd of vernietigd worden. Dit betekent dat de theoretische elektronenoptische ruimtelijke resolutie niet altijd haalbaar is. Daarom is het verkennen van nieuwe strategieën om stralingsschade te overwinnen een hot topic in het onderzoek naar elektronenmicroscopie. Deze nieuwe methoden moeten dosisefficiënt zijn in extreem lage dosisomstandigheden.

In recente jaren is dosisefficiënte elektronenptychografie daarom een belangrijk vooruitstrevend onderzoeksonderwerp geworden door de snelle ontwikkeling van directe elektronendetectors (DED) met een snellere opnamesnelheid en een hoger dynamisch bereik dan voorheen, geholpen door verbeterde rekenkracht. Als een computationele beeldvormingsmethode kan ptychografie de fase- en amplitude-informatie van een object

halen uit de opgenomen diffractiepatronen. In vergelijking met conventionele TEM- en STEM-beeldvormingsmethoden heeft het significante voordelen zoals:

1. Het kan superresolutie bereiken die verder gaat dan de conventionele limiet die wordt bepaald door de convergentiehoek. Dit betekent dat voor elektronenmicroscopen zonder moderne aberratiecorrectoren een hoge ruimtelijke resolutie nog steeds mogelijk is.
2. Ptychografie is dosisefficiënter dan high-resolution TEM en andere STEM-methoden. Onderzoekers hebben ptychografie met succes toegepast op stralingsgevoelige materialen zoals biologische monsters en batterijmaterialen.
3. De resterende aberraties van de probe kunnen worden verwijderd via postverzamelings aberratiescorrectie. Dit betekent dat zelfs als de elektronenmicroscop niet goed is uitgelijnd, er nog steeds een hoogwaardig beeld kan worden bereikt met computationele methoden.
4. Als een fasebeeldvormingsmethode is het zeer gevoelig voor de elektromagnetische velden van materialen, en daarom kunnen ook ladingsdichtheid en zelfs ladingsverplaatsing worden gedetecteerd met ptychografie.

Momenteel zijn er twee benaderingen voor ptychografie. De ene maakt gebruik van een gefocuste probe om over het monster te scannen in een dicht opeengeplaatst raster, wat gefocuste probe ptychografie wordt genoemd. De andere maakt gebruik van een gedefocuste probe om met grote stapgrootte over het monster te scannen, wat gedefocuste probe ptychografie wordt genoemd. Het voordeel van gefocuste ptychografie is dat men tegelijkertijd nuttige HAADF-beelden uit dezelfde datasets kan verkrijgen. Een reden waarom het HAADF-beeld belangrijk is, is dat de fasebeelden minder gevoelig zijn voor het elementnummer in vergelijking met HAADF-beelden. Het voordeel van gedefocuste ptychografie is dat het het volume van de datasets kan verminderen met minder scanposities voor hetzelfde scanveld van zicht vergeleken met gefocuste probe ptychografie. In deze thesis gebruik ik voornamelijk gefocuste ptychografie.

Voor gefocuste probe ptychografie, hoewel DED-camera's sterk zijn verbeterd, zijn relatief trage scansnelheden een aanzienlijke knelpunt gebleven voor praktische implementatie met de meeste camera's. Trage snelheden resulteren over het algemeen in data die beïnvloed wordt door monsterdrift en introduceren contaminatie, wat de kwaliteit van de beelden die men kan verkrijgen aanzienlijk vermindert. In deze thesis, door gebruik te maken van de event-gedreven Timepix3-camera, wordt het knelpunt van de scansnelheid volledig overwonnen. De verblijftijd van de probe kan minder zijn dan $1 \mu\text{s}$. We hebben de camera getest met het tweedimensionale materiaal WS2 en zeolietmonsters onder lage dosisomstandigheden. De experimentele resultaten tonen aan dat de ptychografische methode dosisefficiënter is dan de HAADF, annular bright field imaging (ABF) en integrated center of mass (iCoM) methoden in lage dosisomstandigheden met hogere signaal-ruisverhoudingen.

De meeste ptychografische algoritmen zijn gebaseerd op de fasebenadering of zwakke fasebenadering. Deze benaderingen gaan ervan uit dat de monsters dun zijn. Echter, de meeste monsters zijn dik genoeg om deze benaderingen te doorbreken, en het verschijnen van faseomslag is gerapporteerd voor dikke monsters. Om deze barrière te overwinnen, is multislice ptychografie voorgesteld, die rekening houdt met het mechanisme

van dynamische verstrooiing. Terwijl in de situatie van lage dosis, het gereconstrueerde resultaat soms niet convergeert.

In deze scriptie zal ik het gebruik van elektronenptychografie in combinatie met de event-gedreven Timepix3-detector onderzoeken om stralingsgevoelige materialen te beelden. Verder onderzoek ik ook hoe het probleem van faseomslag voor dikke monsters kan worden overwonnen, inclusief door het aanpassen van de defocus van de probe. Ik heb ook een verklaring kunnen geven waarom centrale focussing vaak optimaal is met behulp van het concept van een geïntegreerde contrastoverdrachtsfunctie. Ik geef ook een meer algemene uitleg van de contrastomkeringen met behulp van faseomslag in de reciproke ruimte.

Hoofdstuk 1 introduceert de basis theoretische achtergrond van elektronenmicroscopie. Van de bron tot het detectievlak wordt een gedetailleerde afleiding gegeven van hoe een diffractiepatroon wordt gevormd. Vervolgens wordt het concept van 4D STEM geïntroduceerd.

Hoofdstuk 2 behandelt verschillende soorten ptychografische beeldvormingsmethoden. Voor directe inversie ptychografie worden de single side band (SSB) en Wigner distributie deconvolutie (WDD) methoden geïntroduceerd. Voor iteratieve ptychografie worden de uitgebreide ptychografische iteratieve engine (ePIE) en multislice ePIE methoden geïntroduceerd. In dit hoofdstuk zal ik voornamelijk introduceren waarom het ophalen van de fase van het object belangrijk is en hoe de fase met deze methoden te verkrijgen.

Hoofdstuk 3 gaat dieper in op de toepassing van conventionele HAADF, gekoppeld aan het elektronenenergieverliesspectrum (EELS), om de gelaagde structuur van LaVO_3 en LaTiO_3 te ontrafelen. De nabije atoomnummers van Ti en V vormen een uitdaging voor HAADF, aangezien het ze niet kan onderscheiden op basis van alleen de intensiteit van verstrooiende elektronen. EELS biedt de aanvullende informatie voor de Ti- en V-elementen. Deze strategische integratie van technieken blijkt succesvol in het oplossen van de gelaagde structuur van de materialen.

Hoofdstuk 4 benadrukt het gebruik van elektronenptychografie in omstandigheden met een lage dosis met de Timepix3-detector. De Timepix3-detector, gekenmerkt door zijn onderscheidende event-gedreven mechanisme, maakt een verblijftijd van minder dan $1 \mu\text{s}$ onder lage dosisomstandigheden mogelijk. Deze capaciteit is bijzonder voordelig voor het verkennen van stralingsgevoelige materialen.

Hoofdstuk 5 gaat over de toepassing van elektronenptychografie op monsters dikker dan 2D-materialen, zoals SrTiO_3 . Een veelvoorkomend probleem bij ptychografie met dikkere monsters is contrastomkering. Echter, door zorgvuldige experimenten en simulaties ontdekten we dat door het afstellen van de defocus van de probe we vaak contrastomkeringen kunnen overwinnen. Vaak blijkt wat het beste werkt wat wij centrale focussing noemen, waarbij de probe wordt gefocust op het centrale vlak van het monster, maar we kunnen ook defocus gebruiken na het nemen van gegevens met de probe elders gefocust. Onze bevindingen bieden waardevol inzicht in het optimaliseren van beeldvormingscondities voor dikke specimens.

In Hoofdstuk 6 wordt een geïntegreerde contrastoverdrachtsfunctie (CTF) gebruikt om uit te leggen hoe centrale scherpstelling contrastomkeringen kan overwinnen. Dit conceptuele kader biedt inzicht in hoe het afstemmen van de onscherpte van de probe op het

middenvlak van het monster de contrasten beïnvloedt. We ontdekken dat dit concept niet werkt voor sterk fase-objecten.

In hoofdstuk 7 wordt verder onderzoek gedaan om contrastomkering te overwinnen. We vinden dat faseomslag-asymptoten in de frequentie aanleiding geven tot contrastomkering in SSB ptychografie en door de asymptoten naar hogere frequenties te verschuiven, kan de contrastomkering worden tegengegaan. Deze methode wordt onderzocht met de simulatie- en experimentele gegevens, die superieur blijken te zijn aan post verzamelsapplicatie van defocus.

In Hoofdstuk 8 bied ik een samenvatting van de belangrijkste bevindingen en conclusies die voortvloeien uit het gepresenteerde werk. Daarnaast schets ik voorgestelde wegen voor toekomstig onderzoek, waarbij ik gebieden identificeer die verdere verkenning en ontwikkeling rechtvaardigen.

Contents

1	Background	1
1.1	Why electron microscopy?	1
1.2	STEM introduction (from source to detector)	2
1.2.1	Components of STEM	2
1.2.2	Electron gun system	3
1.2.3	Condenser lenses system	5
1.2.4	Interaction of the electrons with the specimen	7
1.2.5	Theory of kinetic and dynamic scattering in the specimen	11
1.2.6	STEM imaging modes in the detector plane	14
1.3	Convergent Beam Electron Diffraction	16
1.3.1	Interference of electron beams	16
1.3.2	Bragg's law	16
1.3.3	Diffraction modes in TEM and STEM	17
1.3.4	Selected area diffraction	19
1.3.5	CBED in STEM	19
1.3.6	4D STEM	21
1.3.7	Direct electron detectors	22
2	Literature review	25
2.1	Phase contrast imaging in optical microscopy	25
2.2	Ptychography	25
2.3	The phase problem	26
2.4	Why is phase information of the object important?	26

2.5	Introduction to ptychographic algorithms	27
2.5.1	Direct computational ptychography	28
2.5.2	Wigner distribution deconvolution method	32
2.5.3	Iterative ptychography	38
2.5.4	Conclusion	41
3	Characterizing oxide samples with HAADF and EELS	43
3.1	Abstract	43
3.2	The structure of LaTiO ₃ and LaVO ₃	43
3.3	Characterization of LaTiO ₃ /LaVO ₃ with HAADF	44
3.4	Characterization with EELS	45
3.5	Conclusion	46
4	Dose efficiency in low dose condition with Timepix3 detector	49
4.1	Introduction	50
4.2	Introduction to the Timepix3 camera	51
4.3	Processing the raw data	51
4.4	Experimental Results	52
4.4.1	2D material: WS ₂	52
4.4.2	Zeolite	56
4.5	Aberration correction	56
4.6	Conclusion	58
5	Overcoming contrast reversals in focused probe ptychography of thick materials	61
5.1	Introduction	62
5.2	Results	63
5.3	Conclusion	68
5.4	Supplementary information	69

CONTENTS

6	On central focusing for contrast optimization in direct electron ptychography of thick samples	77
6.1	Introduction	78
6.2	Theory	78
6.2.1	Conventional phase object approximation theory.	79
6.2.2	Thick weak phase object model	79
6.3	Results and Discussion	84
6.3.1	Exploring the iCTF concept in relation to contrast reversals	86
6.3.2	Influence of the convergence angle	86
6.4	Conclusion	90
6.5	Supplementary information	91
7	Phase offset method of ptychographic contrast reversal correction	97
7.1	Introduction	97
7.2	Theory	99
7.2.1	Conventional Zernike phase contrast imaging theory.	99
7.2.2	Phase offset method	100
7.3	Results and Discussion	100
7.4	Conclusion	115
8	Conclusion and perspective	117
8.1	Conclusion	117
8.2	Perspective	118
A	Symbols	121
B	Outputs	123
B.1	Publications	123
B.2	Conferences	123
C	Acknowledgement	137

CONTENTS

Chapter 1

Background

1.1 Why electron microscopy?

Understanding the structure of materials with higher spatial resolution is important in materials science, physics, and chemistry due to the many physical and chemical properties determined by the structures of materials. Many modern technologies rely on manipulating the atomic structure of materials and thus, understanding and acquiring more information about the structure of materials has become increasingly vital. As a consequence, the more clearly we see, the more we understand materials. Different kinds of microscopes have been invented and are useful for imaging at different spatial resolutions. For example, optical microscopy is widely used in many fields, however, the Abbe limit of spatial resolution for optical microscopy is about 0.22 micrometers, depending on the light wavelength, which is far from atomic resolution. However super-resolution methods circumvent the resolution limit and provide spatial resolutions in the range 1–100 nm in optical microscopy [1,2]. To achieve atomic resolution is still not possible with such resolution.

Electron waves, on the other hand, can reach shorter wavelengths when accelerated to the speeds used in electron microscopes, and as a consequence, higher spatial resolution can be reached in electron microscopes. Electron microscopes are in many ways analogous to optical microscopes with electrons as the source instead of photons. However, unlike the glass based lenses used for light optical microscopy, the trajectory of the electrons is generally manipulated by controlling the current of copper coils inside electromagnetic lenses. De Broglie in 1927 proposed that the electron has wavelike properties [3]. The De Broglie wavelength of electrons at high speeds is much shorter than that of visible photons. According to the classic Rayleigh criteria, the smallest distinguishable distance δ between two objects is defined as:

$$\delta = \frac{0.61\lambda}{\mu \sin \alpha}, \quad (1.1)$$

where λ is the wavelength, μ is the refractive index of the medium, and α is the convergence angle. From the above equation, we can see shorter wavelengths of beams correspond to higher spatial resolution for the microscopes. For example, for electrons accelerated by a 200 kV voltage difference, the wavelength is about 0.025 Å, assuming

the numerical aperture $\mu \sin \alpha$ is unity, the highest spatial resolution can reach 0.015 \AA . The distance between atomic columns in materials is usually angstrom or sub-angstrom scale, which is much larger than the spatial resolution limit of electrons. This is the principle which allows electron microscopes to achieve atomic resolution.

The first TEM was demonstrated by Max Knoll and Ernst Ruska in 1931, based on the principle of light microscopes. In 1986, Ernst Ruska was awarded the Nobel prize alone [4], after Max Knoll passed away in 1961. The first STEM was designed and constructed by Manfred von Ardenne in 1937-1938 [5], which was motivated by the development of camera tubes for television when Ardenne developed scanning electron microscopy (SEM). Since then both of these kinds of electron microscopes have become widely used in science. However the operation of TEM and STEM have important fundamental differences.

TEM: Conventional TEM uses a parallel beam to illuminate the sample and image contrast is produced by fundamentally different mechanisms. For HRTEM, the interference of the scattered and transmitted beams are used to form images at the image plane, as controlled by intentional use of aberrations, in particular defocus and spherical aberration.

STEM: In STEM, a convergent beam is formed by the illuminating system. The beam focuses on the sample and images are formed when the beam scans across the field of view driven by the scan coils. Various imaging modes exist, depending on the region of electrons gathered by the detectors. In my study, I mainly focused on STEM, therefore I will mainly introduce the principles and current development status of STEM in the following.

1.2 STEM introduction (from source to detector)

STEM is a powerful tool in the characterization and analysis of samples with nanometer even atomic resolution. It has been widely used in subjects including material science, physics and chemistry. In this section, the basic principles of STEM will be introduced. The image formation process includes the following parts:

1. The probe formation
2. Interaction between the electrons and specimen
3. Diffraction pattern formation in the back focal plane
4. Electron collection with different geometrical detectors
5. Image formation

1.2.1 Components of STEM

The STEM consists of the main parts shown in Figure 1.1, the electron gun, condenser lens system, sample stages, and the detector systems.

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

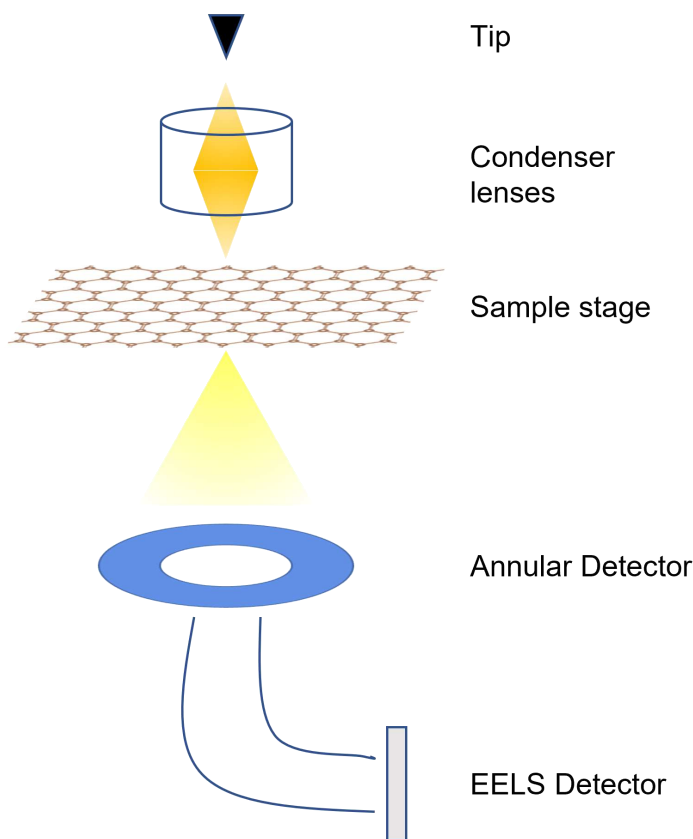


Figure 1.1: Schematic of essential elements of STEM

1.2.2 Electron gun system

The electron gun provides the electron beam. The properties of the electron source, from which the electrons are emitted, are important for the performance of the STEM system. Currently there are three kinds of electron guns, (1) those based on thermionic sources, (2) field emission guns (FEG) and (3) cold field emission guns (CFEGs). Thermionic sources are composed of LaB_6 or tungsten, which are then heated to emit electrons. Field emission gun tips are composed of tungsten needles which are also heated but about a 1 kV potential is applied on the tips of the tungsten needles which can lower the work function barrier and force electrons to tunnel out from the tungsten. CFEGs also use tungsten as tips but at ambient temperature. For a thermionic source, the energy spread of the emitted electrons is higher than in a FEG and especially a CFEG, but the vacuum requirement is not as high as FEGs and especially CFEGs. However, the source brightness is not as high as FEGs and CFEGs. For CFEGs, the extremely small source size means the beam has high spatial coherence and the energy spread is smaller compared with FEGs. However, CFEGs require ultra high vacuum to decrease the contamination and prevent degradation or destruction from ionized gas particles. Therefore FEGs and especially CFEGs are mostly only used in high end STEMs, as they command much higher prices than the thermionic sources used in basic TEMs. The choice of electron

guns depends on the needs of the electron microscopists.

There are three main characteristics to describe the performance of the electron gun, (1) brightness (2) temporal coherence (3) spatial coherence.

1.2.2.1 Brightness

Brightness is defined as the current density per solid angle of the source. Assuming the diameter of the probe cross-over is d_0 , the cathode emission current is i_e and the divergence angle is α . The brightness β is defined as:

$$\beta = \frac{4i_e}{\pi^2 d_0^2 \alpha^2}. \quad (1.2)$$

Brightness is important because a higher brightness means more electrons put into a given probe size and more information about the sample can be collected, which is beneficial for increasing the signal to noise ratio in imaging and spectra analysis. At the same time, the damage for sensitive materials also increases with a higher dose.

Coherency is also an important property of the electron beam. coherency refers to the correlation between waves at different points in space and time. In electron microscopy it can be classified into two main types: spatial coherence and temporal coherence.

1.2.2.2 Temporal coherence

In theory, the ideal coherent electron beam should be monochromatic, which means all of the emission electrons should be of the same energy and therefore frequency. We refer to this aspect of coherence as temporal coherence. The coherence length concept defines the length scale at which the coherent electron waves maintain a high degree of coherence. The definition of coherence length is

$$\lambda_t = \frac{vh}{\Delta E}, \quad (1.3)$$

where v is the electron velocity, h is the Plank constant, ΔE is the energy spread of the beam. For typical electron sources, ΔE is from 0.3 eV to 3 eV. The energy spread of the field emission gun, especially a CFEG, is much less than the thermionic gun. Energy spread influences the spatial resolution and energy resolution, and monochromators are sometimes used to narrow the energy spread of the beam significantly.

1.2.2.3 Spatial coherence

Spatial coherence refers to the property of a wave, where the phase of the wave is correlated over a certain spatial distance. Perfect spatial coherence means that the electron beam acts as if all the electrons were emitted from the same point. This means every point at the electron wavefront maintains the same phase. In practice the source

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

size is not infinitely small, and the effective source size d_c for coherent illumination is described by

$$d_c = \frac{\lambda}{2\alpha}, \quad (1.4)$$

where λ is the wavelength and α is the angle subtended by the source at the specimen. To maximize the spatial coherence, the tip of the source should be as small as possible. The field emission gun has a narrower tip and thus is a better choice for spatial coherence compared with thermionic guns.

1.2.2.4 Wavelength of the electrons

When the electrons are emitted from the source, they undergo acceleration by the high voltage to very high velocity, comparable to the speed of light. For example, the velocity of electrons accelerated with 100 kV can reach half of the speed of the light. Higher voltages of course produce even faster more relativistic speeds. Therefore the wavelength of the electrons should be calculated relativistically,

$$\lambda = \frac{hc}{\sqrt{eV(2m_0c^2 + eV)}}, \quad (1.5)$$

where e is the electron charge, m_0 is the rest mass of electron, c is the speed of light, h is the Planck constant, and V is the accelerating voltage. For instance, at 100 kV, the wavelength is 0.037 Å. According to the Rayleigh criteria, increasing the acceleration voltage can decrease the wavelength of the electrons and therefore increase the spatial resolution. Of course, this assumes idealised conditions and in reality aberrations limit the achievable spatial resolution, but nonetheless higher voltages do produce higher resolutions.

1.2.3 Condenser lenses system

The accelerated electrons will pass through the condenser lens system and be demagnified to produce as small a spot as possible. During this process, a circular aperture is used to block the higher angle aberrated electrons and form the probe. This is illustrated in figure 1.2. Mathematically this can be described with the an aperture function $A(\mathbf{k})$ which takes the value one for angles \mathbf{k} inside the aperture and zero for angles outside the aperture. The intensity distribution of the probe in the real space can then be expressed as the modulus square of Fourier transform of the aperture function,

$$P(\mathbf{r}) = |\mathcal{F}(A(\mathbf{k}))|^2 = I_0 \left| \frac{J_1(\mathbf{r})}{\mathbf{r}} \right|^2, \quad (1.6)$$

where J_1 is the Bessel function of first order of the first kind, \mathbf{r} is the two dimensional coordinates in real space, I_0 is the maximum intensity of the pattern at the Airy disc center, and \mathcal{F} corresponds to the Fourier transform. The distance between the central peak of the intensity and the first minimum point of an Airy disc is $\frac{0.61\lambda}{\alpha}$. This is defined as the Rayleigh resolution criterion. This formula implies that the wavelength and the aperture size determine the image resolution. According to this criterion, shorter

wavelengths and larger apertures are desirable for higher spatial resolution. While in practice a larger aperture would introduce more geometric aberrations introduced by the condenser lens system, especially the objective lens. Shorter wavelengths mean higher electron energies, which could potentially damage the sample if the sample is susceptible to knock on damage. Therefore the choice of acceleration voltage and aperture size depend on the condition of the electron microscope and samples.

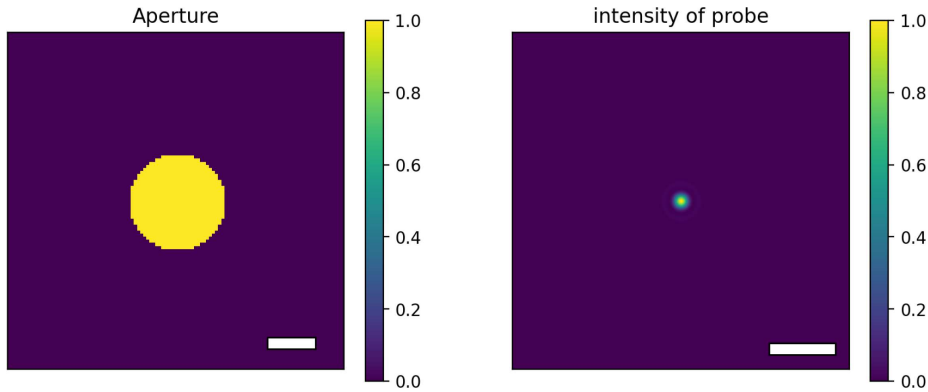


Figure 1.2: The aperture function in reciprocal space and the corresponding probe in real space in the ideal condition. The scale bar if the left figure is 20 mrad, the scale bar in the right figure is 2 Å.

1.2.3.1 Geometric aberrations

Electromagnetic lenses are far from perfect. They introduce geometrical aberrations to the waves due to the imperfect lenses. As described above, an aperture is usually used to cut off the higher angles when forming the probe. A larger aperture means more geometric aberrations would be involved in the probe and aberration correctors only correct so far out in angle. In this section, I will introduce some basic concepts of geometrical aberrations. As shown in figure 1.3.1, ideally a lens would converge the spherical wave emitted from a point from the left side to a point as shown on the right side labeled by dashed lines. While due to the residual aberrations of the lenses, the lens cannot perfectly focus the emitted wave into a point, but only a disk. The aberrations cause the wave to deviate from a spherical wave with a distance δ as shown in figure 1.3.2, or a phase shift $\chi = \frac{2\pi\delta}{\lambda}$. Currently there are two main notations used to define the aberrations, one is proposed by Haider [6] and the other is by Krivanek [7]. With Krivanek's notation, the aberration function can be written as

$$\chi(\alpha, \varphi) = \frac{2\pi}{\lambda} \sum_{mn} \frac{\alpha^{n+1}}{n+1} [C_{nma} \cos(m\varphi) + C_{nmb} \sin(m\varphi)], \quad (1.7)$$

where φ is the azimuthal angle, α is the angle between the sample and ray in the specimen plane. n is the order of aberrations and m is the multiplicity. C_{nm} denotes the aberration coefficients with the unit of length. a describes aberrations whose effects are mirror-symmetric about the principal longitudinal plane of the system, while b describes antisymmetric aberrations.

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

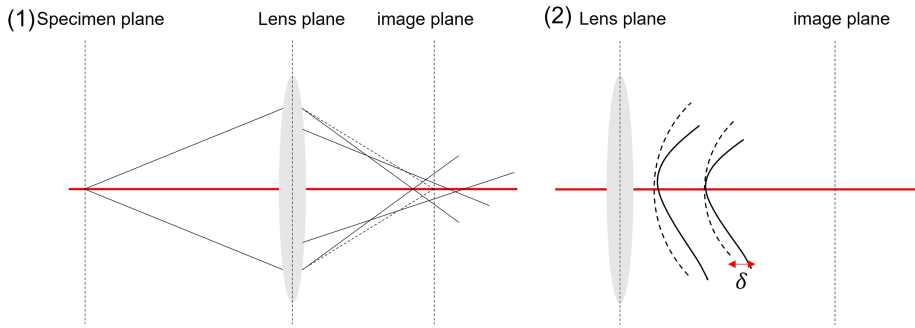


Figure 1.3: Schematic of geometric aberrations for the condenser lens. (1) The dashed lines indicate the electron trajectories for an ideal lens, the solid lines are the electrons trajectories for an imperfect lens with aberrations. (2) The definition of the aberration function. Dashed lines indicate the ideal spherical wave front for the convergent electron beam, the solid lines indicate the real wave front due to geometric aberrations, where δ is the distance between the ideal and real electron wave front.

Some axial aberration symbols through fifth order in two different systems of notation are listed in the figure 1.4 .

1.2.3.2 The probe function

The condenser lens focuses the electrons to the specimen plane and forms the probe. Again, this process can be described as the Fourier transform of the aperture function when considering the aberration function. $P(\mathbf{r}) = \mathcal{F}(A(\mathbf{k})e^{-i\chi(\mathbf{k})})$ where \mathbf{r} is the real space vector, \mathbf{k} is the reciprocal space vector, χ is the aberration function. In STEM, the quality of the images are highly influenced by the geometrical aberrations. As shown in figure 1.5, a set of probes with different aberrations are plotted. The voltage is 200 kV, and the convergence angle is 20 mrad. The aberration coefficients are listed below. We can see with different aberrations, the shape of the probe changes substantially. Modern STEM instruments are usually equipped with aberration correctors to correct the geometrical aberrations. Through minimizing the aberrations, the contrast transfer function can be modified to deliver the best spatial resolution. In HRTEM, usually the defocus is used to compensate the spherical aberration to give the best spatial resolution. Therefore, for different imaging methods, the question how to treat the aberrations depends on the imaging methods we use.

1.2.4 Interaction of the electrons with the specimen

When the electrons pass through the illumination system and the specimen, many different kinds of signals can be generated by different mechanisms as shown in figure 1.6. Usually for the specimen in the microscope, it is thin enough to allow the incident electrons to be transmitted. Here we just consider the transmitted beams which are mostly used in STEM. Other signals such as back scattered electrons, and x-rays will not be considered in this chapter.

Coefficients (Krivaneck)	Coefficients (Haider)	Description
C_{10}	C_1	Defocus
$C_{12}(a, b)$	A_1	Twofold astigmatism
$C_{21}(a, b)$	B_2	Axial coma
$C_{23}(a, b)$	A_2	Threefold astigmatism
$C_{30}(C_s)$	C_3	Third-order spherical
$C_{32}(a, b)$	S_3	Axial star aberration
$C_{34}(a, b)$	A_3	Fourfold astigmatism
$C_{41}(a, b)$	B_4	Fourth-order axial coma
$C_{43}(a, b)$	D_4	Three-lobe aberration
$C_{45}(a, b)$	A_4	Fivefold astigmatism
C_{50}	C_5	Fifth-order spherical
$C_{52}(a, b)$	S_5	Fifth-order axial star
$C_{54}(a, b)$	C_5	Fifth-order rosette
$C_{56}(a, b)$	A_5	Sixfold astigmatism

Figure 1.4: Aberration coefficients with different notations.

Assuming the incident beam is coherent and monochromatic, after the interaction with the specimen, we usually separate the transmitted scattered beam into coherent and incoherent parts. For the transmitted coherent beams, their phase still remains in phase, such as a conventional bright field image. While for the transmitted incoherent beams, they have no phase relationships, typical for the high angle elastic scattered electrons and the inelastic electrons. These coherent and incoherent beams carry the structure and spectroscopic information of the specimen. By collecting these signals with specific detectors, we can retrieve the related information about specimens.

To better understand the nature of the coherent and incoherent beams transmitted by the specimen, we have to understand the interactions between the incident electrons and the specimen. Generally speaking, there are two kinds of scattering mechanisms: elastic scattering and inelastic scattering, which are related to whether the beam is coherent or incoherent. We will introduce the scattering mechanisms below.

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

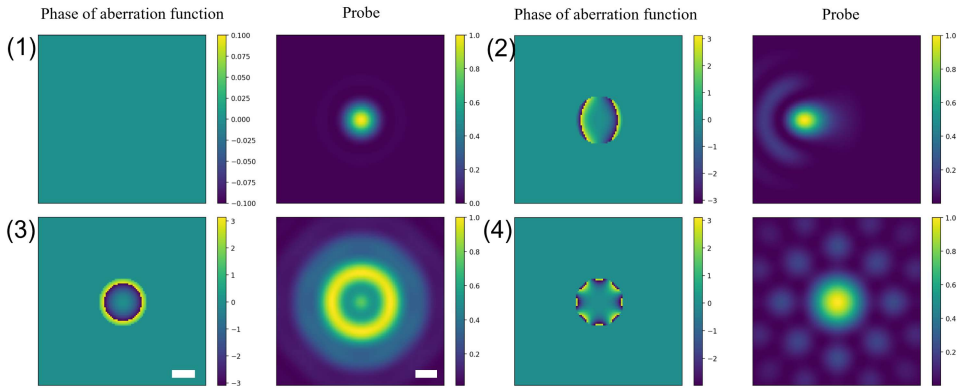


Figure 1.5: Probe shapes with different single aberrations. The voltage is 200 kV and the convergence angle is 20 mrad. (1) Aberration free. (2) Coma with $C_{21} = 1000$ nm. (3) defocus=10 nm. The scale bar in left figure is 20 mrad and the scale bar in the right figure is 2 Å. (4) $C_{34a} = 50000$ nm.

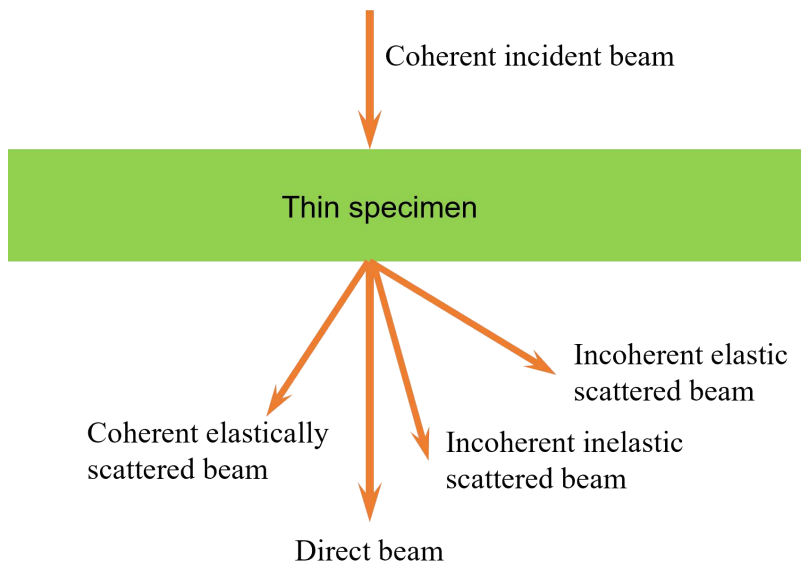


Figure 1.6: Interaction of high energy electrons with a thin specimen.

1.2.4.1 Elastic scattering

The electron is a low mass, negatively charged particle. The Coulombic interaction of the electron and electron cloud or nucleus causes scattering. (1) If the specimen is thin enough, most electrons do not change their path very much or lose significant energy, we usually group these electrons into what we call the direct beam as shown in figure 1.6. At the same time, some electrons interact more strongly with the atoms. (2) Weak Coulombic interaction between the incident electrons and electron cloud causes the incident beam to have a small deviation from the original incident direction (1

to 10 degrees) as shown in figure 1.7. We call these electrons low angle scattered electrons. These low angle scattered electrons retain their coherence. (3) When the incident electrons pass near the nuclear cores, the strong Coulombic interaction scatters the electrons to high angles (>10 degree). Due to thermal vibrations and the dominance of thermal diffuse scattering, these electrons generally have incoherent properties, which is a significant difference compared to low angle scattered electrons.

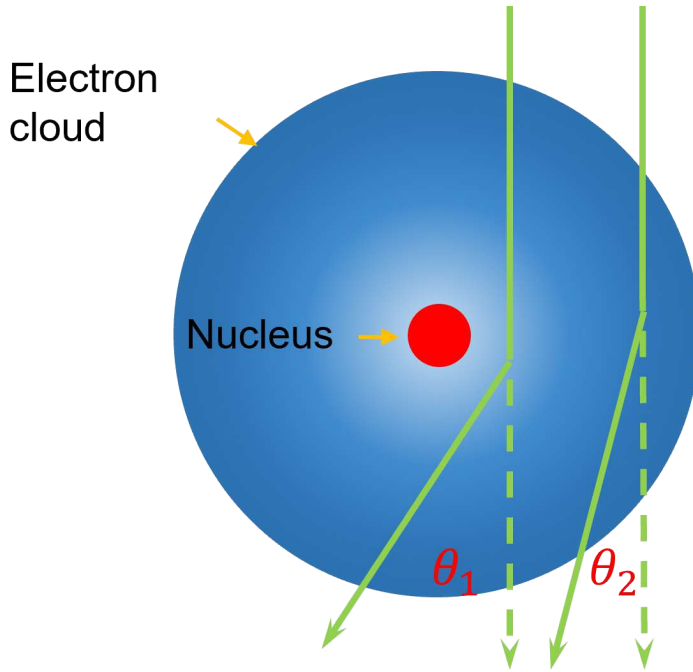


Figure 1.7: The two mechanisms of the isolated atom scattering incident electrons. Coulombic interaction within the electron cloud results in low angle scattering. Coulombic interaction with the nucleus results in high angle scattering.

1.2.4.2 Inelastic scattering

In addition to elastic scattering, the incident electrons may also suffer inelastic collisions and lose energy by a variety of processes, like phonon excitations, plasma excitations, and inner shell ionizations. We can obtain tremendous information about the chemistry and the electronic structure of the specimen using these energy loss electrons, which are hard to obtain from the elastic electrons. The basic principle is shown in figure 1.8. The incident electrons excite the inner shell electrons to the conduction band and lose specific energies according to quantum mechanics, composing what is called the high energy loss spectrum. These characteristic features are called ionization edges, which contain the electronic structure information about the atoms. Another mechanism is exciting valence electrons to the conduction band, which due to the low energies required for these transitions contribute to what is called the loss electron spectrum. The low loss also reflects the dielectric response of the specimen. Phonons and plasmons are collective

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

excitations which also contribute to the low loss region of the spectra. Although I will not introduce them here in detail, all these signals can be detected with spectrometers. Currently, most modern electron microscopes are equipped with spectrometers to detect the energy loss spectrum. This method is called EELS.

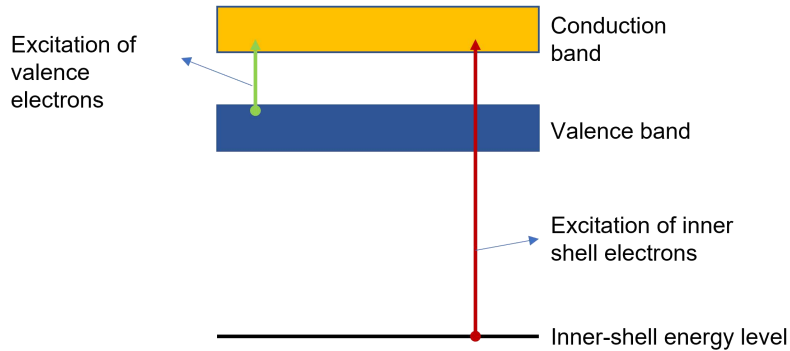


Figure 1.8: Schematic EELS with different excitation mechanisms.

1.2.5 Theory of kinetic and dynamic scattering in the specimen

When the electrons pass through the specimen, the elastic scattering of electrons within the specimen is not actually as simple as the single atom model shown in figure 1.7. The electrons can undergo single or multiple scattering processes depending on the thickness of the specimen. Different models can be used to approximate the scattering depending on the thicknesses involved and the required level of sophistication. For specimens which are quite thin, we can use the phase or weak phase approximation to describe the object function ignoring the thickness of the specimen and multiple scattering effects. For thicker specimens, a multislice model is generally much more accurate. In the following section, these models will be introduced.

1.2.5.1 Projected potential approximation

For elastic scattering, the primary interaction of incident electrons and the specimen is with the electrostatic potential of the specimen and the charge of electrons. Compared to the electrostatic potential, the magnetic potential is usually small, especially for non-magnetic samples, and thus will not be discussed here. For TEM, the incident electrons can be treated as plane waves. For STEM, the convergent probe is the superposition of many plane waves. For simplicity, we only consider the incident electrons as plane waves here. Next I use a specimen function $f(x, y)$ to represent the specimen. A general model is

$$f(x, y) = A(x, y) \exp(-i\Phi_t(x, y)), \quad (1.8)$$

where $A(x, y)$ is the amplitude and Φ_t is the phase which depends on the thickness of the specimen. To further simplify the model, the amplitude is set to unity, which means the amplitude of the incident beam amplitude does not change when passing through

the specimen. With this model, the phase change of the beam depends on the potential $V(x, y, z)$. In practice, the 3D potential can be quite complex for thick specimens. If we assume the specimen is thin, then we can use a projected potential $V_t(x, y)$ to represent $V(x, y, z)$:

$$V_t(x, y) = \int_0^t V(x, y, z) dz. \quad (1.9)$$

With this model, we essentially squeeze a 3D specimen to a 2D specimen. That is why it is called a projected potential model. If the incident wave is a plane wave, then the exit wave function can be expressed as:

$$\Psi(x, y) \approx \exp(2\pi i k_z z) \exp(i\sigma V_t(x, y)). \quad (1.10)$$

Here the interaction parameter σ is

$$\sigma = \frac{2\pi m e \lambda}{h^2}. \quad (1.11)$$

We can see for a specimen function $f(x, y)$, the specimen is a pure phase object. That is why it is also called the phase object approximation. This approximation only holds true for thin specimen.

Actually, we can continue simplifying the model for very thin specimens, for which the projected potential $V_t(x, y) \ll 1$. Then we can expand the exponential function and ignore the higher order items, the new object function $f(x, y)$ is

$$f(x, y) = 1 - i\sigma V_t(x, y). \quad (1.12)$$

This approximation is called the weak phase object approximation (WPOA). For a very thin specimen, the amplitude of the exit wave function is linearly related to the projected potential. It is clear that the projected potential $V_t(x, y)$ varies with the position of the incident beams. For example, in the center of an atom and further away from the atom, the amplitudes have strong differences.

Although the POA and WPOA are generally only valid for thin specimens, it is still sufficient to provide insights to understand the structure of thick specimens. For HRTEM and STEM imaging methods, these assumptions are widely used to analyse materials. Besides most of the ptychographic algorithms are also based on phase or weak phase approximation. I will introduce these algorithms in detail later.

1.2.5.2 Theory of dynamical scattering in the specimen

The electrons can interact with the sample strongly and scatter more than once when the sample is as thin as 10-40 Å depending on the elements it contains. If the electrons only scatter once when passing through the sample, we call it a kinetic scattering process. Otherwise, the scattering process more than once, we call it dynamic scattering process. The best calculation method depends on the thickness of the specimens and the desired level of accuracy. For very thin specimens, as we have introduced before, the projected potential approximation is often sufficient. In practice many specimens are thick enough that the multiple scattering plays a significant role. The theory of dynamical scattering

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

has been investigated for many decades. The multislice method [8], which can handle such scattering, will be introduced below.

When the incident electrons waves $\Psi_f(x, y, z)$ interact with the specimen, they can be considered governed by the time independent Schrodinger equation:

$$\left[-\frac{\hbar}{2m}\nabla^2 - eV(x, y, z)\right]\Psi_f(x, y, z) = E\Psi_f(x, y, z), \quad (1.13)$$

where \hbar is called the Planck constant reduced and is the Plank constant divided by 2π . m is the relativistic mass of the electron. e is the magnitude of the charge of the electron. V is the electromagnetic potential of the electron. E is the kinetic energy of the electron. Due to the energy of incident electrons being much higher than the potential of the sample, the sample has only a small perturbation to the transmitted electrons when electrons pass through. We can divide the $\Psi_f(x, y, z)$ into two parts, a plane wave function traveling along the z axis and another term varying slowly along the z axis. Then

$$\Psi_f(x, y, z) = \Psi(x, y, z) \exp\left(\frac{2\pi iz}{\lambda}\right). \quad (1.14)$$

Now substitute this function into the Schrodinger function. Then it can be written as

$$\left[\nabla_{xy}^2 + \frac{\partial^2}{\partial z^2} + \frac{4\pi i}{\lambda} \frac{\partial}{\partial z} + \frac{2meV(x, y, z)}{\hbar^2}\right]\Psi(x, y, z) = 0. \quad (1.15)$$

Due to λ being quite small and $\Psi(x, y, z)$ varying slowly along z axis,

$$\left|\frac{\partial^2 \Psi}{\partial z^2}\right| \ll \left|\frac{1}{\lambda} \frac{\partial \Psi}{\partial z}\right|. \quad (1.16)$$

The new function can be approximated as

$$\left[\nabla_{xy}^2 + \frac{4\pi i}{\lambda} \frac{\partial}{\partial z} + \frac{2meV(x, y, z)}{\hbar^2}\right]\Psi(x, y, z) = 0. \quad (1.17)$$

This function can be rewritten as a first order differential function as follow

$$\begin{aligned} \frac{\partial \Psi}{\partial z} &= \left[\frac{i\lambda}{4\pi} \nabla_{xy}^2 + \frac{2mei\lambda}{4\pi\hbar^2} V(x, y, z)\right]\Psi(x, y, z) \\ &= \left[\frac{i\lambda}{4\pi} \nabla_{xy}^2 + i\delta V(x, y, z)\right]\Psi(x, y, z), \end{aligned} \quad (1.18)$$

where δ is the interaction parameter and can be written as $\frac{2\pi me\lambda}{\hbar^2}$. This equation has a formal operator solution, assuming the incident wave function is $\Psi(x, y, z)$. Integrating from z to Δz yields:

$$\Psi(x, y, \Delta z + z) = \exp \int_z^{z+\Delta z} \left(\frac{i\lambda}{4\pi} \nabla_{xy}^2 + i\delta V(x, y, z')\right) \Psi(x, y, z), \quad (1.19)$$

where Δz is a small piece of specimen. Then $\Psi(x, y, z + \Delta z)$ can be approximated as

$$\Psi(x, y, \Delta z + z) = \exp\left(\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2 + i\delta V_{\Delta z}(x, y, z)\right) \Psi(x, y, z), \quad (1.20)$$

where $V_{\Delta z} = \int_z^{z+\Delta z} V(x, y, z) dz$ is the projected potential of specimen between z and $z + \Delta z$. Because Δz is small, we can simplify this equation to

$$\begin{aligned}\Psi(x, y, \Delta z + z) &= \exp\left(\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2\right) \exp(i\delta V_{\Delta z}(x, y, z)) \Psi(x, y, z) + \mathcal{O}(\Delta z^2) \\ &= \exp\left(\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2\right) t(x, y, z) \Psi(x, y, z) + \mathcal{O}(\Delta z^2).\end{aligned}\quad (1.21)$$

If the propagation distance Δz is small enough, $\Psi(x, y, z)$ varies slowly, and then $\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2$ can be approximated as unity, and this can be further simplified to

$$\Psi_e(\vec{r}) = \Psi_i(\vec{r}) t(\vec{r}), \quad (1.22)$$

where $\Psi_e(\vec{r})$ is the exit wave function, $\Psi_i(\vec{r})$ is the incident wave function and $t(\vec{r}) = \exp(i\delta V(\vec{r}))$ is the transmission function. As introduced in the section on phase objects, the multiplicative relationship is often called the strong phase approximation. If we Fourier transform the function:

$$\begin{aligned}\mathcal{F}\left[\exp\left(\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2\right) (t\Psi)\right] &= \int \left[\exp\left(\frac{i\lambda}{4\pi} \Delta z \nabla_{xy}^2\right) (t\Psi)\right] \exp[2\pi i(k_x x + k_y y)] dx dy \\ &= \exp[-i\pi\lambda\Delta z(K_x^2 + k_y^2)] \mathcal{F}[(t\Psi)] = P(k, \Delta z) \mathcal{F}[t\Psi].\end{aligned}\quad (1.23)$$

In real space, this can be written as

$$\Psi(x, y, z + \Delta z) = P(x, y, \Delta z) \otimes \Psi(x, y, z). \quad (1.24)$$

$P(x, y, z + \Delta z)$ can be viewed as a propagator. $\Psi(x, y, z)$ is the exit wave function from the previous slice, \otimes denotes the convolution. Figure 1.9 illustrates the procedure of the multislice simulation process. We first slice the thick sample into n slices with a thickness Δt , in each slice, phase object approximation holds true. Then we use the exit wave function propagated from one slice to the next as the incident wave function of the next slice until the electron waves pass through the entire sample.

1.2.6 STEM imaging modes in the detector plane

After the electrons pass through the specimen, they form diffraction patterns in the diffraction plane and different types of detectors are placed to collect these electrons. In STEM, there are different kinds of imaging methods based on the geometry of the detectors. The detectors determine which angles of electron trajectories contribute to the final images. The principles of these methods differ substantially. In STEM, the most commonly used imaging mode is the ADF, often complemented with a STEM bright field (BF), ABF or increasingly a differential phase contrast (DPC) type imaging mode.

1. **BF:** A conventional STEM bright field detector is a small disk detector located at the center of the direct transmitted beam, otherwise known as the bright field disk. According to the principle of reciprocity, the BF is equivalent to the HRTEM imaging method, although generally less efficient due to only using a small fraction of the direct beam. The small size is to maintain coherence. Usually BF

1.2. STEM INTRODUCTION (FROM SOURCE TO DETECTOR)

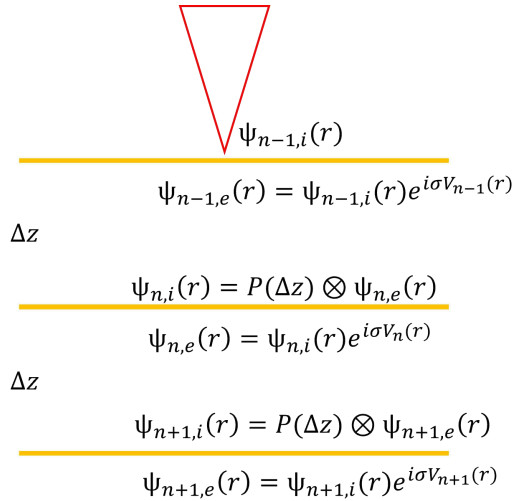


Figure 1.9: Multislice decomposition of a thick sample. The sample is sliced into multiple slices, with each slice approximated as its projected potential. The incident wave function at each slice can be viewed as the convolution of the transmission function and the propagator function.

- is a complementary method to the ADF, it can detect light atoms in materials. As for HRTEM, image simulations are generally required to help interpret the experimental images.
2. **ABF:** ABF is also a kind of bright field imaging, integrating an annular ring in the outer part of the BF disk. The contrast is a combination of phase contrast and channelling contrast. ABF has become a popular method for imaging light atomic columns, however it must be noted that this does not always work as a focused probe interacting with a weak phase object such as graphene produces very little contrast unless aberrations are also introduced. This is because the ABF method relies on channeling contrast to reveal light atomic columns with an aberration corrected focused probe. This is why light atomic columns often appear with negative contrast in ABF. It has been widely used in imaging oxygen atoms in oxide samples, lithium atoms in battery materials.
 3. **ADF:** ADF uses an annular detector with an inner radius wide enough to exclude the BF disk, thus collecting the scattered electrons in just the dark field region. Higher atomic number atoms scatter more electrons to high angle due to Rutherford scattering from the nucleus of atoms. Therefore the intensity of the images is proportional to roughly the square of the atomic number Z of the atoms encountered by the probe. This is why ADF imaging is also called Z -contrast imaging, and is very easily interpreted in comparison to other imaging modes especially phase contrast imaging. This is also due to the incoherent nature of Z -contrast signal, in which atoms are always bright on a dark background. Higher intensity means heavier atomic columns or simply greater sample thickness. Coherent phase contrast images on the other hand are far more complicated due to how many variables effect the interference forming the images. The inner angle of the ADF detector is varied according to the needs of the experiment. High angle ADF

(HAADF), in which the annular detector collects only the electrons scattered to a range of very high angles, produces the most pure atomic number or Z-contrast. However when imaging weakly scattering materials a smaller inner angle can be beneficial. For example when imaging light 2D materials such as graphene or BN a medium angle ADF (MAADF) is commonly used to obtain the best signal from the weak scattering from C, B and N and other similarly light elements [9]. However the disadvantage of ADF is that light atomic columns can be hidden from view when they neighbor heavy atomic columns. This is one reason supplementing ADF images with other imaging modalities such as ABF or indeed EELS can be important. More recently the quest to locate light elements near heavy elements has seen the uptake of methods such as iDPC and as we will see, ptychography.

1.3 Convergent Beam Electron Diffraction

Convergent beam electron diffraction (CBED) is the type of diffraction one obtains in STEM. CBED can be used to measure crystal structural parameters such as lattice parameters, strain information, defects, crystal symmetry, and orientations of the crystal samples at a very fine scale. As a diffraction technique, CBED is widely used in the STEM to analyze the crystallographic properties of materials with much higher spatial resolution than conventional TEM diffraction technique due to the small size of the STEM probe. In the next section, I will introduce the basic principles of CBED.

1.3.1 Interference of electron beams

Diffraction is a behaviour of all kinds of waves, in which they spread out around an obstacle. According to De Broglie's theory, electrons have a wave-particle duality, behaving both as particles and waves. When the electrons diffract, they will form diffraction patterns from constructive and destructive interference introduced by the sample. As we introduced in the elastic scattering section, electrons which are elastically scattered to small angles (less than three degree) are coherent. The scattered electrons are mainly affected by the arrangement of the atoms. For a crystal sample, all kinds of atomic planes can reflect the electrons which can result in diffraction patterns in the diffraction plane. The diffraction is determined by the angles of incident beams to the atomic planes in the specimen, the spacing between atomic planes, and the inter atomic distance between these planes. We can obtain these information about a crystal from the diffraction patterns. Therefore electron diffraction is very valuable for crystallography and characterization generally.

1.3.2 Bragg's law

Bragg's law is a fundamental principle in crystallography, which describes the relationship between the angles at which electrons (photons, neutrons) are scattered by a crystal lattice and the spacing between the atomic planes in that crystal lattice. It can be crucial to understanding the structure of a crystalline sample. The mathematical relationship

1.3. CONVERGENT BEAM ELECTRON DIFFRACTION

is illustrated in figure 1.10. If the two adjacent scattered waves remain in phase, the difference in the path ($AB+BC$) must be the integers of the wavelength λ . Mathematically, it can be expressed as

$$n\lambda = 2d \sin(\theta), \quad (1.25)$$

where n is an integer (known as the order of the diffraction), λ is the wavelength of the incident waves, d is the spacing between crystal lattice planes, also known as the interplanar spacing, θ is the angle of incidence, the angle between the incident beam and the lattice plane.

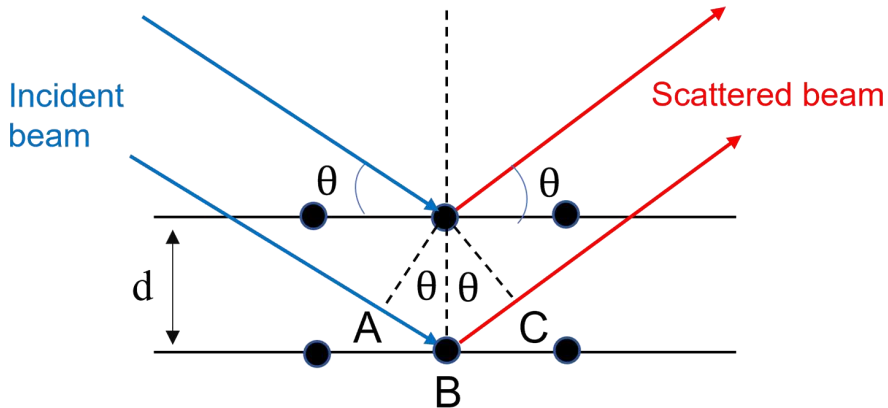


Figure 1.10: Schematic of Bragg's law. A plane wave with the wavelength of λ incident on atomic planes with a spacing of d at an angle of θ . The path difference of the scattered plane waves are $AB+BC$.

1.3.3 Diffraction modes in TEM and STEM

In this section, we will discuss diffraction in TEM and STEM, in particular selected area diffraction (SAD) in TEM and CBED in STEM.

1.3.3.1 Diffraction formation in TEM

Figure 1.11 shows a diagram of the imaging system of a TEM. The objective lens below the specimen can collect the electrons exiting the sample and form diffraction patterns in the back focal plane or form images in the image plane. The intermediate lens projects the beams diffracted by the specimen collected by the objective lens to form diffraction pattern or image depending on the strength of the lens to the image plane of the projector lens. Then the projector lens projects the patterns in the object plane to the detector plane and forms the final image. Thus the intermediate lens plays an important role in transitioning between imaging and diffraction modes in TEM.

Diffraction: As shown in figure 1.11A. The back focal plane of the objective lens acts as the object plane for the intermediate lens, then the intermediate lens projects the DP to the image plane of the projector lens and finally forms a DP on the detector.

Imaging: We just need to adjust the intermediate lens to make the image plane of the objective lens as the object plane of the projector lens. Then the images will be projected on the screen as shown in figure 1.11B.

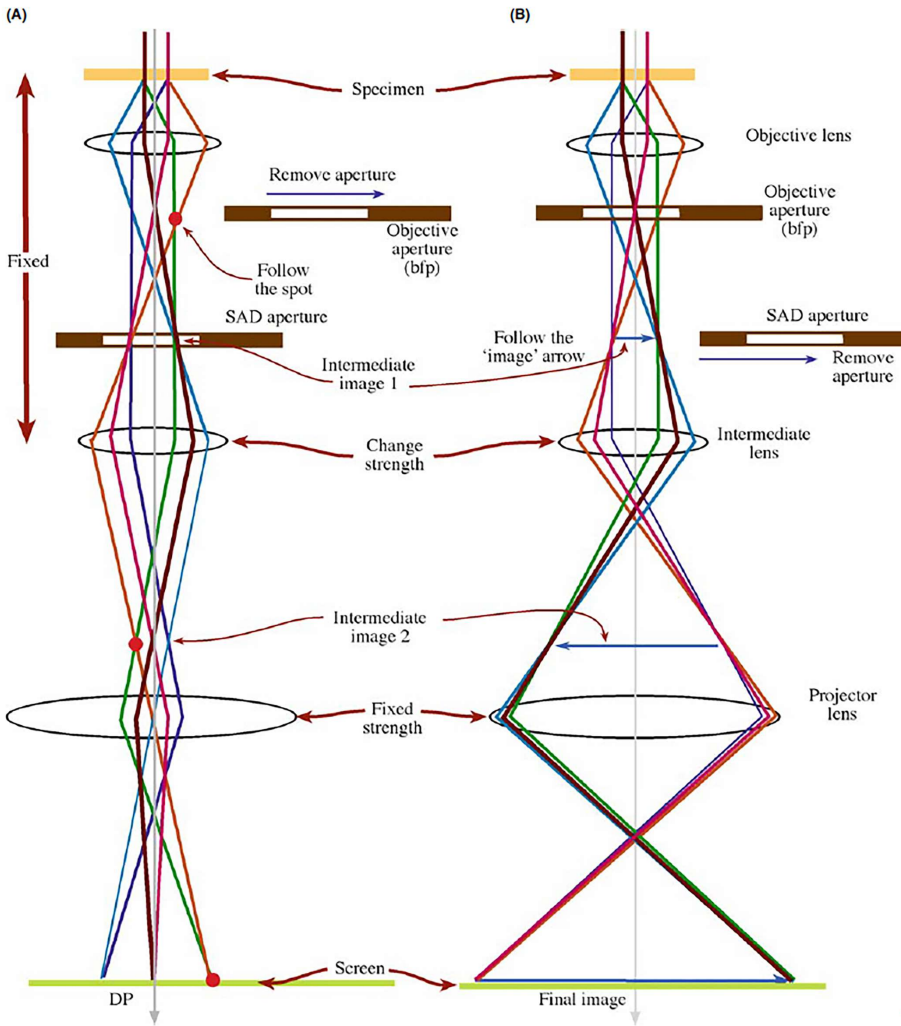


Figure 1.11: Imaging and diffraction modes in TEM. In figure A, The projector lens projects the diffraction patterns onto the screens. In figure B, the projector lens projects the image onto the screens. This figure is adapted from reference [10].

1.3. CONVERGENT BEAM ELECTRON DIFFRACTION

1.3.4 Selected area diffraction

The diffraction patterns contains information from the entire area of specimen that we illuminate with the electrons unless we use a selected area diffraction (SAD) aperture. Sometimes we just need the diffraction information from the specific areas and use a SAD aperture in the image plane of the objective lens to select the area of interest.

1.3.5 CBED in STEM

Instead of using a parallel beam in TEM, for STEM, a convergent beam is used to illuminate the specimen.

1.3.5.1 Why use a convergent beam?

For SAD in TEM, although it can give us a great deal of information about the specimen, there are still some disadvantages to the technique. For example, due to the limitation of the size of the selected area aperture, the smallest area that one can selectively sample is about 100 nm in ideal conditions. This is still too large for most nanoparticles, especially with defects or a secondary phase for example. In addition, the SADPs provide inaccurate crystallographic information about the specimen because the Bragg condition is relaxed for the thin or small particles. That is because the diffraction spots are only infinitely sharp if the specimen is perfect and infinite in all directions. For the small particles, the diffraction spots are not ideal points anymore, therefore obtaining precise crystallographic information become more difficult.

One way to overcome the limitation of SAD in TEM is using a convergent beam to illuminate the specimen and sample regions significantly smaller than than 100 nm. In addition to improving spatial resolution, a great deal of crystallographic information is contained in the convergent beam electron diffraction (CBED), which I will introduce next.

1.3.5.2 Setup in STEM

As shown in figure 1.12, the condenser lens focuses the probe at the specimen plane. The illuminated area is far less than with a parallel beam, which is how CBED can achieve higher spatial resolution than TEM diffraction. A convergent beam contains the electrons with a range of angles, and therefore, the diffraction patterns in the detector plane are not tiny spots as in TEM but instead are discs. The drawbacks of CBED compared to conventional diffraction include (1) the convergent beam may damage the sample because of the high intensity if the dose is not managed sufficiently well and (2) a focused probe is more likely to induce contamination.

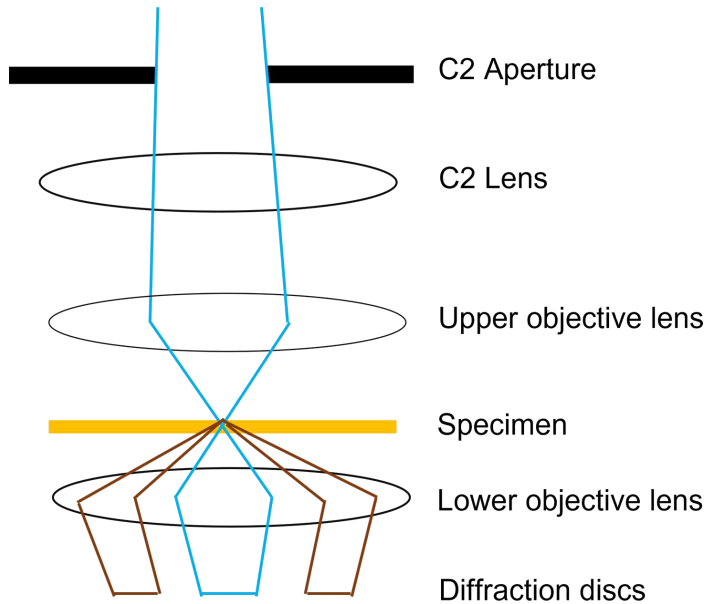


Figure 1.12: Schematic diagram of CBED. The condenser and objective lenses focus the probe onto the specimen with a very tiny area. In the back focal plane, it would form diffracted discs.

1.3.5.3 Experimental variables for CBED

To form CBEDs in the experiment, one must consider the (1) the beam convergence angle (2) camera length (3) the size of the probe (4) the thickness of the specimen and (5) the focus of the probe.

1. Beam convergence angle

For a focused probe, the convergence angle in combination with the electron wavelength, determines the range of spatial frequencies which can be detected. It is also what determines the maximum conventional spatial resolution in STEM images. In the diffraction plane, a larger convergence angle translates into broader transmitted and diffracted discs. There are 3 kinds of CBED based on whether the discs overlap or not, as shown in figure 1.13. (a) The pattern of non-overlapping discs is a Kossel-Mollenstedt (K-M) pattern. When the convergence angle is small enough, the radius of the discs is smaller than the spacing of disks (the spacing is determined by the Bragg angle θ_B). Then each disc is separated. This condition is widely used in mapping the orientations of crystals. (b) If the convergence angle α is large enough, the discs will overlap, then the individual diffraction discs are not discernible. These are called 'Kossel patterns', and are useful for ptychography as I will discuss in detail later.

2. Camera length

The camera length L controls the magnification of the patterns. A large L gives a high-magnification pattern with a narrow angular range in reciprocal space.

1.3. CONVERGENT BEAM ELECTRON DIFFRACTION

The choice of camera length depends on the conditions used. For example, for ptychography, to achieve super resolution beyond the limit of the aperture, we should choose a small camera length to obtain high scattering angle electrons.

3. The size of the probe

The size of the probe determines the spatial resolution. Compared with SAD technique in TEM, the size of the convergent beam can be much smaller than the selected area in TEM. This is the advantage of CBED.

4. Effect of sample thickness

If the sample is sufficiently thin, kinematical scattering dominates. The intensity of the diffracted discs will be uniform. While for a thick sample in which multiple scattering dominates, dynamical scattering starts to produce contrast in the diffracted discs. The contrast in the diffraction pattern can be more complex. That is because the diffuse, inelastic scattering can raise the background intensity to levels that mask the useful contrast in the pattern.

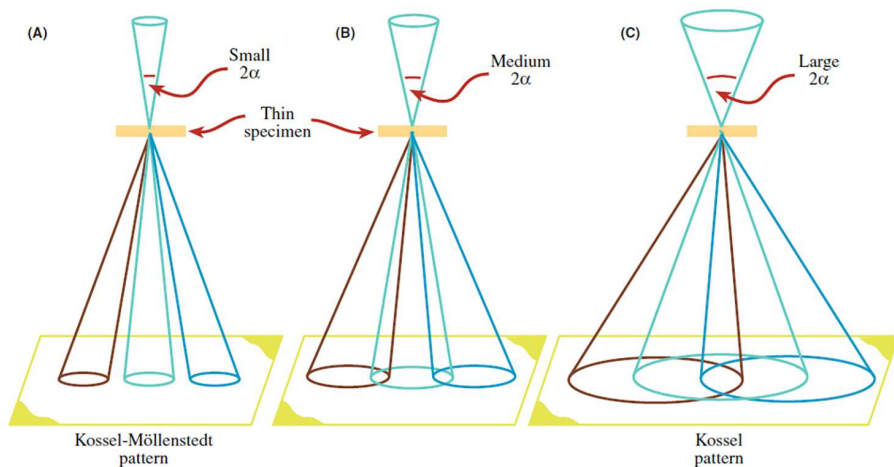


Figure 1.13: The diagram shows how the convergence angle influences the CBED patterns. In the case transmitted and diffracted discs are not overlapped, it is called a K-M mode. If the discs are overlapped, it is called a Kossel pattern. This figure is adapted from reference [10].

1.3.6 4D STEM

4D STEM [11] refers to using a convergent probe to scan across the sample while recording the diffraction patterns with a pixelated detector. The '4D' aspect comes from the 4 dimensional datasets, 2 dimensions refers to the coordinates in real space and 2 dimensions refer to the coordinates in reciprocal space as shown in figure 1.14. Due to the slow recording time of the cameras used for 4D STEM previously, data acquisition was extremely slow and therefore not widely used in the electron microscopy community. With the development of high performance cameras, 4D STEM is becoming increasingly

popular. Not only conventional BF, ABF and ADF images can be reconstructed from the datasets with the variety of the “virtual” detectors, new algorithms which have been developed in recent years can provide unique advantages compared with conventional imaging methods. For example, electron ptychography is a phase imaging method which has advantages such as high dose efficiency and super resolution. The phase of the object can be reconstructed from the recorded 4D datasets with iterative or direct computation algorithms.

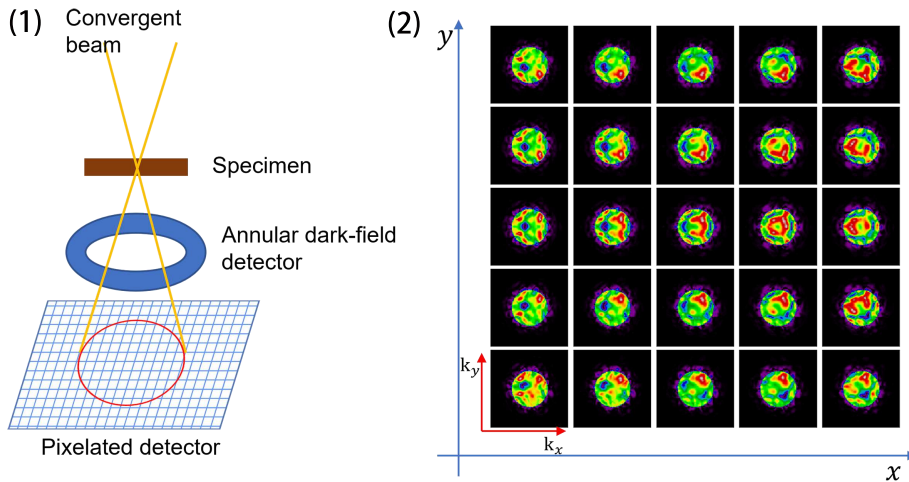


Figure 1.14: Schematic diagram of 4D STEM. (1) The convergent probe raster scans across the specimen while a pixelated camera simultaneously records the diffraction patterns in reciprocal space. (2) The 4D datasets in the 4D coordinates. To index a diffraction pattern, the scan position (x, y) and diffraction pattern in the detector (k_x, k_y) have to be known.

1.3.7 Direct electron detectors

Direct electron cameras are an outstanding technical advance in electron microscopy in recent years. Direct electron detection cameras have been widely used in materials and biological electron microscopy in recent years due to their ability to record high signal to noise ratio images compared with scintillator based detectors. For conventional STEM microscopes, the charge-coupled-device (CCD) cameras and CMOS are equipped as a standard camera to monitor the images or Ronchigrams. CCD cameras have the disadvantages of the relatively slow read out speed and low signal to noise ratio. The low signal to noise ratio is because the fast electrons are converted into light first, then generate stored charge in the CCD for later readout and amplification or conversion into a digital signal. While for direct electron detectors, the incident electrons illuminate the CMOS sensor directly and the generated electrons within the pixel will be transferred and converted to digital signal. The fast electrons can be directly recorded within the device, which makes it sensitive to single electrons and demonstrates much better SNR than CCD cameras. For both CCD and direct electron cameras, most of them are frame-based, which means the export speed is still a bottleneck for practical use. In this thesis,

1.3. CONVERGENT BEAM ELECTRON DIFFRACTION

we used an event-driven camera Timepix3 to overcome the bottleneck of slow scan speed.

1.3.7.1 Merlin detector

MerlinEM [12] is a pixelated detector developed by Quantum Detectors based on the Medipix3 [13] chip with 256x256 pixels. As a detector that is sensitive to even single electrons, it can operate in several counting modes with different dynamic ranges. The maximum dynamic range of one pixel is 24 bit, corresponding to 16.7 million electrons. While in this mode however, the scan speed is limited to about 1.7 *ms* per frame, which results in very slow scans. The detailed information about the MerlinEM detector is shown in figure 1.15. The fastest speed for the MerlinEM detector is in 1-bit mode, which can only really be used with an extremely low probe current because otherwise signal will be lost due to the pixels being saturated after a single hit. The dwell time is about 53 μ s per scan position, which is almost comparable to most conventional STEM scans [14]. In practice, the most widely used mode for 4D-STEM is the 6 bit mode, which balances the scan speed and the dynamic range. However this scan speed is still very slow compared to a normal STEM scan.

Counting mode	Max. value	Speed (fps)
24-bit	16,777,215	600
12-bit	4095	1600
6-bit	63	3200
1-bit	1	18800

Figure 1.15: Counting modes of the latest Merlin Medipix3 camera.

1.3.7.2 Timepix3 detector

Unlike the frame-based Medipix3 camera, Timepix3 detector [15] is an event-driven detector. At low doses, the diffraction patterns are usually very sparse, which means most of the pixels only contain zeros. In the figure below, I give a brief explanation of why the event-driven detector is much faster than the frame-based detector. Figure 1.16(1) shows a sparse diffraction pattern with 256 by 256 pixels acquired at a low dose of 2000 $e^-/\text{\AA}^2$. The bright points in the diffraction pattern are the positions where electrons hit the camera. The dark points are where no electron hits occurred. For such sparse data, the nonzero signal only occupies a small part of the whole signal. While the frame-based detector needs to transmit the entire signal including a very large amount of data only containing zero, the event driven data saves all that data transfer of zeros

and in so doing obtains a huge speedup. The principle of the event-driven detector is shown in figure 1.16 (2), the detector only records the information of arrival time and positions of each hit. This means only the nonzero signal will be transmitted. Compared with cameras that operate frame-by-frame, it only needs to export a very small volume of data, for low dose data. Therefore the Timepix3 detector can provide far faster scans than the Merlin detector.

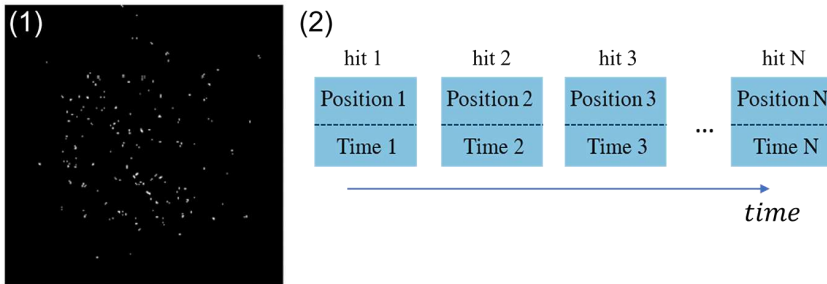


Figure 1.16: The principle of Timepix3. (1) An example of a sparse diffraction pattern acquired with the Timepix3 camera. (2) Illustration of how the Timepix3 reads out a stream of packets registering each hit with a timestamp and position on the camera.

1.3.7.3 Comparison of the speed of commercial pixelated detectors

In figure 1.17, we list some widely used commercial pixelated detectors and compare the readout speed. From the figure, we can see that in low dose mode, Timepix3 detector is much faster than other frame-based cameras.

Detector type	Pixel size	Dwell time (μs)
Merlin	(256,256)	313 (6 bits mode) 53 (1 bit mode)
EMPAD	(128,128)	~900
pnCCD	(264,264)	~500
Timepix3	(256,256)	1 (low dose mode)

Figure 1.17: Comparison of different kinds of direct electron detector.

Literature review

2.1 Phase contrast imaging in optical microscopy

Phase-contrast microscopy is a sophisticated optical technique that enhances the contrast of transparent and colorless specimens, making it an invaluable tool in biological and medical research. This method, developed by the Dutch physicist Frits Zernike in the early 1930s [16], addresses the limitations of traditional brightfield microscopy, where transparent specimens often appear nearly invisible due to the lack of inherent contrast.

The basic principle of phase-contrast microscopy is the ability to convert phase shifts in light passing through a transparent specimen into visible intensity changes. These phase shifts occur because different parts of the specimen have varying refractive indices and thicknesses, which cause light waves to slow down and change phase. While these phase differences are not directly detectable by the human eye or standard cameras, phase-contrast microscopy translates them into variations in light intensity, thereby making the structures within the specimen visible.

Zernike phase contrast method is using a phase plate which is positioned in the back focal plane of the objective lens, the phase plate introduces a phase shift (typically a quarter wavelength) to the undiffracted (background) light. This phase shift enhances the contrast between the background light and the light diffracted by the specimen.

Phase contrast microscopy provides high-contrast images of transparent and unstained specimens, allowing detailed observation of live cells, organelles, and microorganisms. This technique is a cornerstone in biological and medical research, offering unparalleled insights into the intricate world of living organisms.

Zernike phase contrast imaging method is not only used in optical microscopy but also in the electron microscopy [17].

2.2 Ptychography

Ptychography is a phase retrieval method that can reconstruct the phase change imposed by the object on the beam electrons from the diffraction patterns. Ptychography has

proven to be a powerful technique in various scientific and industrial applications [18], enabling researchers to achieve image resolutions beyond the capabilities of conventional microscopy methods. It has been applied in areas such as materials science, biology, and nanotechnology, where the ability to visualize structures at the nanoscale and even the atomic scale is crucial. In this chapter, a detailed review of the principles of ptychography will be given.

2.3 The phase problem

A monochromatic electromagnetic wave can be characterized by the complex function $\Psi(\mathbf{r}) = |\Psi(\mathbf{r})| \exp(i\Phi(\mathbf{r}))$, where $|\Psi(\mathbf{r})|$ is the modulus of the wave, and $\Phi(\mathbf{r})$ is the phase. \mathbf{r} denotes the vector positions of the wave field. In ideal condition, we can predict the information of the wave function at any position if we know the complex wave function. In reality, the detector can only record the intensity of the wave field, $I(\mathbf{r}) = |\Psi(\mathbf{r})|^2$. This means the phase part of $\Phi(\mathbf{r})$ is lost. This is what is referred to as the phase problem.

In ptychography, as introduced in section 3 in this chapter, in the phase object approximation, the wave function in the diffraction plane can be expressed as the convolution between the probe function $P(\mathbf{R}_p, \mathbf{r})$ and object function $O(\mathbf{r})$ in real space, $\Psi(\mathbf{k}_f, \mathbf{R}_p) = P(\mathbf{R}_p, \mathbf{r}) \otimes O(\mathbf{r})$, where \mathbf{R}_p is the scan position of the probe, \mathbf{r} is the extent of the wave in real space, \mathbf{k}_f is the reciprocal vector corresponding to \mathbf{r} in real space. We should note that the wave function consists of complex values. While in the experiment, only the intensity of the diffraction pattern can be recorded with the detector. The intensity is the modulus square of the wave function. This means the phase information is lost in the experiment. Without the phase information, the exit wave function cannot be obtained by simply using the inverse Fourier transform, hence the object function cannot be deduced. To retrieve the phase information about the object is desirable to obtain more information about the specimen.

Ptychography is a computational algorithm which aims to retrieve the phase information from the diffraction patterns.

2.4 Why is phase information of the object important?

The atomic potential is an important property of a specimen. When electron waves pass through the specimen, the interaction between the electrostatic potential of the specimen and the charge on the incident electrons changes the phase of the electron waves. For thin specimens, the kinematical approximation assumes that each electron scatters only one time in the specimen. Mathematically, the plane wave passing through the specimen can be expressed as

$$\Psi(x, y) = \exp(2\pi i k_z z) \exp(\delta V_z(x, y)z), \quad (2.1)$$

where $V_z(x, y, z) = \int V_s(x, y, z) dz$, k_z is the wave vector along the z axis, δ is the interaction parameter, and $V_z(x, y)$ is the projected atomic potential.

For a thin sample, the phase of the exit wave is proportional to the electrostatic potential.

2.5. INTRODUCTION TO PTYCHOGRAPHIC ALGORITHMS

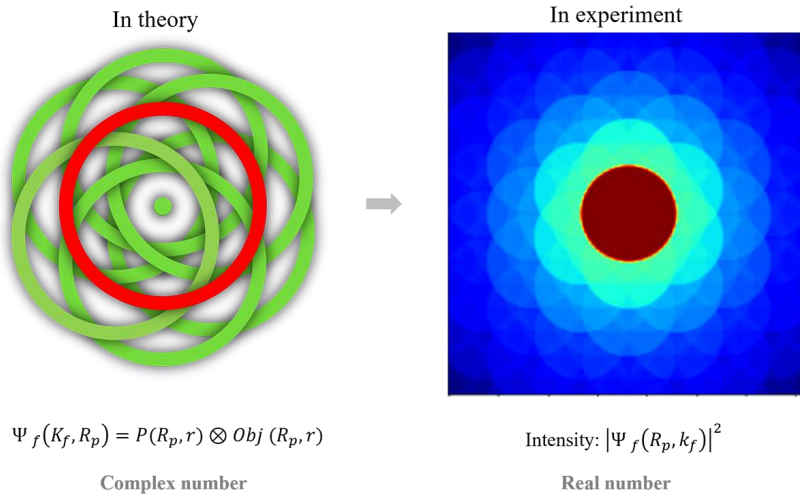


Figure 2.1: The mathematical description of wave function in the diffraction plane in theory and diffraction intensity detected in the experiment. In theory, the diffraction wave function is the convolution between probe function $P(\mathbf{R}_p, \mathbf{r})$ and object function $O(\mathbf{R}_p, \mathbf{r})$, resulting in a complex valued function. While in the experiment, only the intensity of the diffraction patterns is recorded, providing only real images rather than the full complex values. A red circle denotes the transmitted direct beam in the left figure.

In figure 2.2, I present a simple example to show the relationship between the electrostatic potential and the phase reconstructed with ptychographic methods. Figure 2.2(1) is the single atom model used for simulations with the atomic numbers from 1 to 56. Figure 2.2(2) shows an example reconstructed phase image. Figure 2.2(3) shows the phase of the normalized projected potential of each element and the intensity of the normalized phase of each element by the direct ptychographic methods WDD and SSB. We can see the intensity of the reconstructed phases closely matches the strength of the potential. The iterative ePIE algorithm produces a very similar phase vs atomic number plot, as has been demonstrated previously by [19]. Therefore, by retrieving the phase information of the specimen, we can understand more about the electrostatic potential of the specimen as well as its structure.

2.5 Introduction to ptychographic algorithms

The concept of ptychography was first proposed by Hoppe in 1969 [20]. The essential idea of ptychography is the intensity variation in the overlap area of the transmitted and diffracted disks can be used to calculate the relative phases of the frequencies composing the exit wave.

As a computational imaging method, several theoretical algorithms have been proposed. Generally, we can sort them into two classes. (1) Direct inversion algorithms, like WDD and SSB (2) and iterative algorithms, like ePIE. Iterative algorithms rely on the redun-

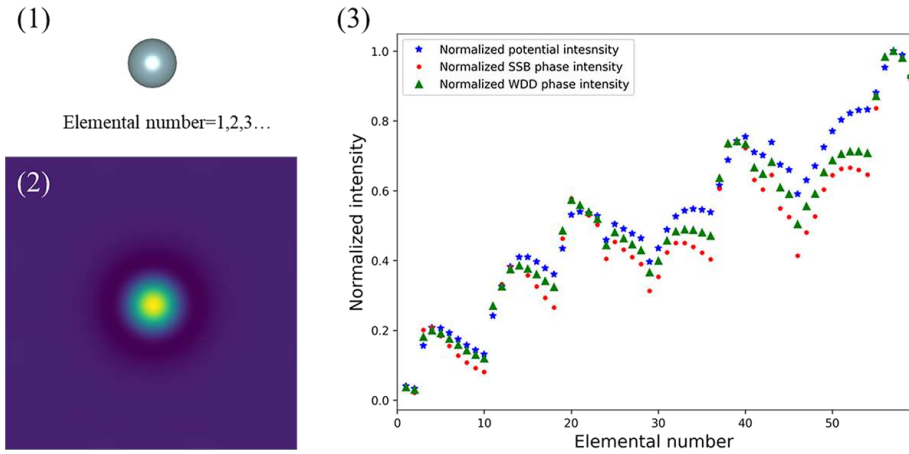


Figure 2.2: Comparison of normalized elemental potential intensity and reconstructed Ptychographic phase intensity. (1) the single atom model that is used for 4D dataset simulation. (2) an example of the reconstructed Ptychographic images. (3) The normalized projected potential strength versus the Ptychographic phase.

dancy provided by probe illumination overlap between two adjacent probe positions to solve for the relative phases. In this section, we will review both direct and iterative Ptychography methods.

2.5.1 Direct computational Ptychography

2.5.1.1 4D datasets

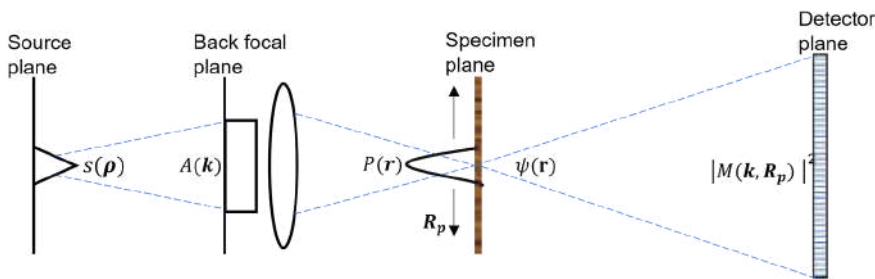


Figure 2.3: The coordinates definition in the Ptychography setup.

Direct computation Ptychographic methods are based on the projected potential assumption. As introduced earlier, this approximation means that for a sample with a thickness of t , the total potential through the whole sample is projected into a single 2D projected potential. Therefore, while thickness is considered in the sense of the total potential encountered by the probe at each scan position, the details of how the probe interacts with the potential as it traverses the sample thickness are largely ignored, such

2.5. INTRODUCTION TO PTYCHOGRAPHIC ALGORITHMS

as channelling and multiple scattering. Using the projected potential approximation, the diffraction pattern in the detector plane can be expressed as the Fourier transform of the product of the probe function $P(\mathbf{r} - \mathbf{R}_p)$ and the object transmission function $\Psi(\mathbf{r})$:

$$M(\mathbf{K}_f, \mathbf{R}_p) = \int P(\mathbf{r} - \mathbf{R}_p)\Psi(\mathbf{r}) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}) d\mathbf{r}, \quad (2.2)$$

where \mathbf{R}_p denotes the scan position of the probe and \mathbf{K}_f is the scattering vector. The intensity at the detector plane is $|M(\mathbf{K}_f, \mathbf{R}_p)|^2$. Both \mathbf{K}_f and \mathbf{R}_p are 2D vectors. An example of simulated 4D dataset is shown in figure 2.4.

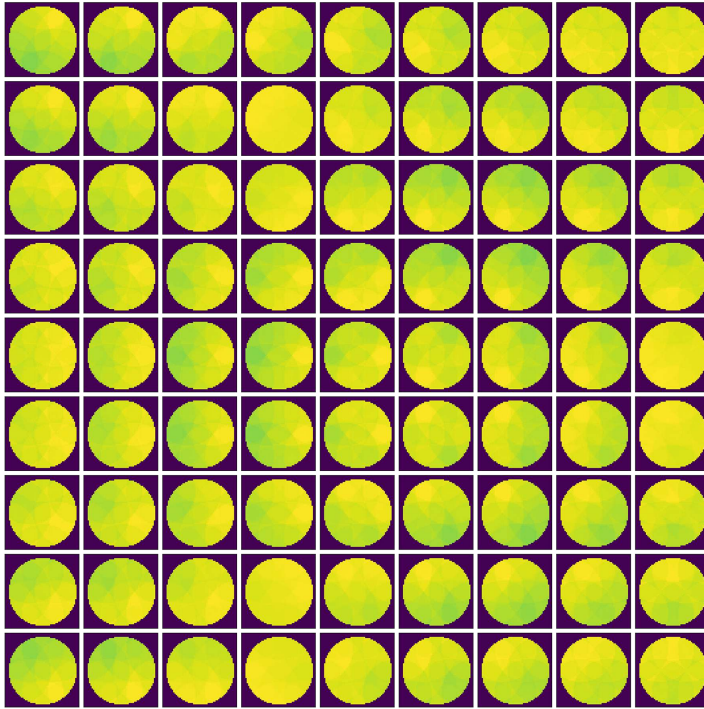


Figure 2.4: Simulated 4D dataset of graphene showing how the intensities of the diffraction patterns vary with respect to probe position.

2.5.1.2 Mathematics of direct ptychography methods

Within the phase object approximation, the exit wave function $\varphi(\mathbf{r})$ can be expressed as

$$\varphi(\mathbf{r}) = P(\mathbf{r} - \mathbf{R}_p) \cdot \Psi(\mathbf{r}). \quad (2.3)$$

Then in the detector plane, the complex wave function is the Fourier transform of the exit wave function $\varphi(\mathbf{r})$

$$M(\mathbf{K}_f, \mathbf{R}_p) = \mathcal{F}(\varphi(\mathbf{r})) = \int \varphi(\mathbf{r}) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}) d\mathbf{r}, \quad (2.4)$$

and the intensity of the diffraction pattern can be expressed as

$$I(\mathbf{K}_f, \mathbf{R}_p) = |M(\mathbf{K}_f, \mathbf{R}_p)|^2 = |\mathcal{F}(P(\mathbf{r} - \mathbf{R}_p) \cdot \Psi(\mathbf{r}))|^2. \quad (2.5)$$

The Fourier transform of $\varphi(\mathbf{r})$ can be written as the convolution of the Fourier transforms of $P(\mathbf{r} - \mathbf{R}_p)$ and $\Psi(\mathbf{r})$, which can be expressed as $A(\mathbf{K}_f) \cdot \exp(i2\pi\mathbf{K}_f \cdot \mathbf{R}_p)$, using the Fourier shift theorem, and $\Psi(\mathbf{K}_f)$ respectively, and therefore

$$I(\mathbf{K}_f, \mathbf{R}_p) = |(A(\mathbf{K}_f) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{R}_p)) \otimes \Psi(\mathbf{K}_f)|^2. \quad (2.6)$$

Next we perform the Fourier transform with respect to the probe scan position \mathbf{R}_p ,

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \int I(\mathbf{K}_f, \mathbf{R}_p) \exp(i2\pi\mathbf{R}_p \cdot \mathbf{Q}_p) d\mathbf{R}_p, \quad (2.7)$$

where \mathbf{Q}_p is the reciprocal vector corresponding to the \mathbf{R}_p , i.e. the spatial frequency. If we write the diffraction intensity expression explicitly,

$$I(\mathbf{K}_f, \mathbf{R}_p) = \iint A(\mathbf{K}_{fa}) \Psi(\mathbf{K}_f - \mathbf{K}_{fa}) A^*(\mathbf{K}_{fb}) \Psi^*(\mathbf{K}_f - \mathbf{K}_{fb}) \exp(i2\pi\mathbf{R}_p \cdot (\mathbf{K}_{fa} - \mathbf{K}_{fb})) d\mathbf{K}_{fa} d\mathbf{K}_{fb}, \quad (2.8)$$

where \mathbf{K}_{fa} and \mathbf{K}_{fb} are dummy variables, which do not influence the final integral results. Now $G(\mathbf{K}_f, \mathbf{Q}_p)$ can be expressed as,

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \iiint A(\mathbf{K}_{fa}) \Psi(\mathbf{K}_f - \mathbf{K}_{fa}) A^*(\mathbf{K}_{fb}) \Psi^*(\mathbf{K}_f - \mathbf{K}_{fb}) \cdot \exp(i2\pi\mathbf{R}_p \cdot (\mathbf{K}_{fa} - \mathbf{K}_{fb} + \mathbf{Q}_p)) d\mathbf{K}_{fa} d\mathbf{K}_{fb} d\mathbf{R}_p, \quad (2.9)$$

with A and Ψ independent of \mathbf{R}_p , because the probe and object do not depend on the scan position. Integrating over \mathbf{R}_p ,

$$\int \exp(i2\pi\mathbf{R}_p \cdot (\mathbf{K}_{fa} - \mathbf{K}_{fb} + \mathbf{Q}_p)) d\mathbf{R}_p = \delta(\mathbf{K}_{fa} - \mathbf{K}_{fb} + \mathbf{Q}_p), \quad (2.10)$$

where δ is the Dirac delta function, therefore

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \iint A(\mathbf{K}_{fa}) \Psi(\mathbf{K}_f - \mathbf{K}_{fa}) A^*(\mathbf{K}_{fb}) \Psi^*(\mathbf{K}_f - \mathbf{K}_{fb}) \delta(\mathbf{K}_{fa} - \mathbf{K}_{fb} + \mathbf{Q}_p) d\mathbf{K}_{fa} d\mathbf{K}_{fb}. \quad (2.11)$$

Only when $\mathbf{K}_{fb} = \mathbf{K}_{fa} + \mathbf{Q}_p$ the result is nonzero. We use $\mathbf{K}_{fa} + \mathbf{Q}_p$ to substitute \mathbf{K}_{fb} , giving

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \int A(\mathbf{K}_{fa}) \Psi(\mathbf{K}_f - \mathbf{K}_{fa}) A^*(\mathbf{K}_{fa} + \mathbf{Q}_p) \Psi^*(\mathbf{K}_f - \mathbf{K}_{fa} + \mathbf{Q}_p) d\mathbf{K}_{fa}. \quad (2.12)$$

Then according to the definition of a convolution

$$G(\mathbf{K}_f, \mathbf{Q}_p) = A(\mathbf{K}_f) A^*(\mathbf{K}_f + \mathbf{Q}_p) \otimes_{\mathbf{K}_f} \Psi_s(\mathbf{K}_f) \Psi_s^*(\mathbf{K}_f - \mathbf{Q}_p). \quad (2.13)$$

2.5.1.3 Single side band ptychography

The WPA can simplify equation 2.13 to

$$\begin{aligned} G(\mathbf{K}_f, \mathbf{Q}_p) &= |A(\mathbf{K}_f)|^2 \delta(\mathbf{Q}_p) \\ &\quad + A(\mathbf{K}_f) A^*(\mathbf{K}_f + \mathbf{Q}_p) \Psi_s^*(-\mathbf{Q}_p) \\ &\quad + A^*(\mathbf{K}_f) A(\mathbf{K}_f - \mathbf{Q}_p) \Psi_s(\mathbf{Q}_p), \end{aligned} \quad (2.14)$$

2.5. INTRODUCTION TO PTYCHOGRAPHIC ALGORITHMS

which is the equation at the heart of SSB ptychography. Now $G(\mathbf{K}_f, \mathbf{Q}_p)$ consists of three terms. The first term is the direct beam (the zero frequency, or DC term). The second and third terms represent the interference between the direct beam and diffracted beams a spatial frequency of \mathbf{Q}_p . Assuming an ideal probe without aberrations, the aperture function A is a top hat function, being one inside the circle defined by the aperture and zero outside. This can be seen in figure 2.5 below. In the double overlap areas (the red and blue areas in the figure), the object transmission functions $\Psi_s^*(\mathbf{Q}_p)$ and $\Psi_s(\mathbf{Q}_p)$ are transferred. Because $\Psi_s^*(\mathbf{Q}_p)$ and $\Psi_s(\mathbf{Q}_p)$ are conjugate to each other, the phase in the triple overlapped area cancels out. Therefore we can obtain the object function $\Psi_s(\mathbf{Q}_p)$ by integrating the values in a single double overlap area at the spatial frequency \mathbf{Q}_p . Collecting the phase and amplitude in this way for every spatial frequency transferred, the total object function $\Psi_s(\mathbf{Q}_p)$ can be obtained. Then one performs the inverse FFT on $\Psi_s(\mathbf{Q}_p)$ to obtain the complex transmission function $\Psi_s(\mathbf{r})$ in real space.

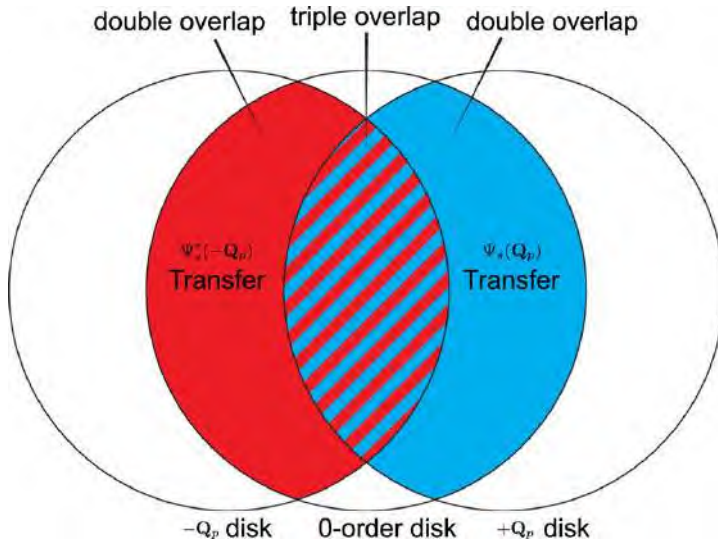


Figure 2.5: Schematic of $G(\mathbf{k}_f, \mathbf{Q}_p)$ at the spatial frequency \mathbf{Q}_p . The left circle corresponds to frequency $-\mathbf{Q}_p$, the middle circle to 0 frequency, and the right circle to frequency \mathbf{Q}_p . In the left double overlapped area (red area), $\Psi_s^*(-\mathbf{Q}_p)$ is transferred. In the right double overlapped area (blue area) $\Psi_s(\mathbf{Q}_p)$ is transferred. In the triple overlap area, $\Psi_s(\mathbf{Q}_p)$ and $\Psi_s^*(\mathbf{Q}_p)$ destructively interfere and the phase of them cancels out. This figure is reproduced from reference [21]

2.5.1.4 Workflow of SSB ptychography

In this section, a simulated 4D dataset of graphene will be used to demonstrate the workflow of SSB ptychography. The simulation parameters are a 60 kV accelerating voltage, a 30 mrad convergence angle, and the scan step size is 0.2 Å. Thermal vibrations were not included.

1. Load the 4D dataset as shown in 2.4

2. Perform the FFT with respect to probe position \mathbf{R}_p and obtain the complex 4D quantity $G(\mathbf{K}_f, \mathbf{Q}_p)$, an example of the phase and amplitude of which is shown in figure 2.6

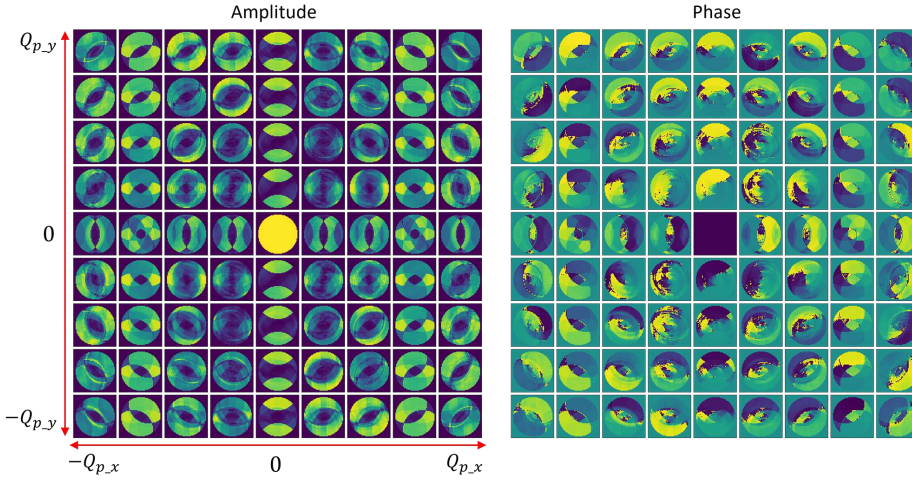


Figure 2.6: Amplitude and phase of $G(\mathbf{K}_f, \mathbf{Q}_p)$ of different spatial frequency from the graphene simulation.

An example of a trotter at a specific spatial frequency is shown in figure 2.7

3. Integrate one side of the trotter at each spatial frequency \mathbf{Q}_p and obtain the object function $\Psi_s(\mathbf{Q}_p)$ in reciprocal space as shown in figure 2.8
4. Perform the inverse FFT on the object function $\Psi_s(\mathbf{Q}_p)$ and obtain $\Psi_s(\mathbf{r})$ in real space as shown in figure 2.9.

2.5.2 Wigner distribution deconvolution method

For WDD ptychography, observe that G in equation 2.13 is a convolution between the aperture function \mathbf{A} and the probe function \mathbf{P} in reciprocal space. If we know the probe function, the object function can be obtained by a deconvolution. WDD method does not invoke the WPA.

After obtaining the expression for G , the next step is to Fourier transform with respect to \mathbf{K}_f ,

$$H(\mathbf{R}, \mathbf{Q}_p) = \mathcal{F}_{\mathbf{K}_f}^{-1}[G(\mathbf{K}_f, \mathbf{Q}_p)] = \chi_A(\mathbf{R}, -\mathbf{Q}_p) \cdot \chi_\Psi(\mathbf{R}, \mathbf{Q}_p), \quad (2.15)$$

where $\chi_A(\mathbf{R}, -\mathbf{Q}_p)$ can be expressed as $\int A^*(b)A(b + \mathbf{R}) \exp(-2\pi\mathbf{Q}_p \cdot b) db$, $\chi_\Psi(\mathbf{R}, \mathbf{Q}_p)$ equals to $\int \Psi^*(c)\Psi(c + \mathbf{R}) \exp(i2\pi\mathbf{Q}_p \cdot c) dc$. H consists of the probe function $\chi_A(\mathbf{R}, -\mathbf{Q}_p)$ and object function $\chi_\Psi(\mathbf{R}, \mathbf{Q}_p)$ Wigner distributions. The object function part can in principle be separated from H by dividing out the probe part, but in practice it is performed using a Wiener filter to avoid to divide by zero issues

$$\chi_\Psi(\mathbf{R}, \mathbf{Q}_p) = \frac{\chi_A^*(\mathbf{R}, -\mathbf{Q}_p)H(\mathbf{R}, \mathbf{Q}_p)}{|\chi_A(\mathbf{R}, -\mathbf{Q}_p)|^2 + \epsilon}, \quad (2.16)$$

2.5. INTRODUCTION TO PTYCHOGRAPHIC ALGORITHMS

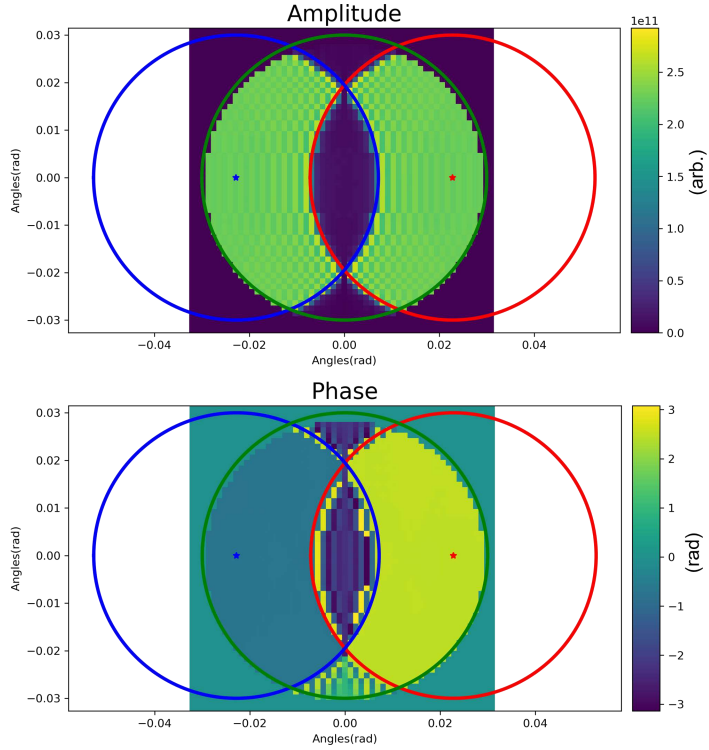


Figure 2.7: An example of a trotter from a single spatial frequency. The green circle indicates the location of the BF disk, and the red and blue circles, the locations of the shifted diffracted disks. Double disk overlap occurs where the red and blue disks overlap just the green disk. Triple overlap occurs when the red and blue disks also overlap each other and the amplitude can be seen to zero in this region.

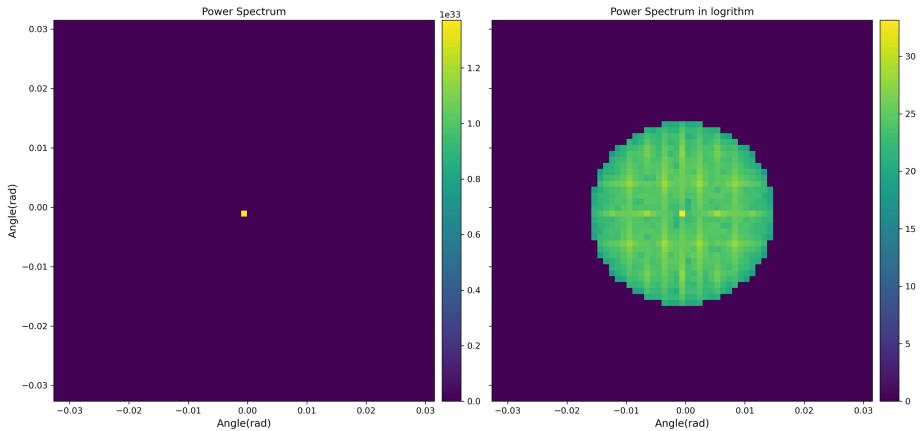


Figure 2.8: $\Psi_s(\mathbf{Q}_p)$ in reciprocal space.

where ϵ is a very small value used to ensure the denominator is not zero. Fourier

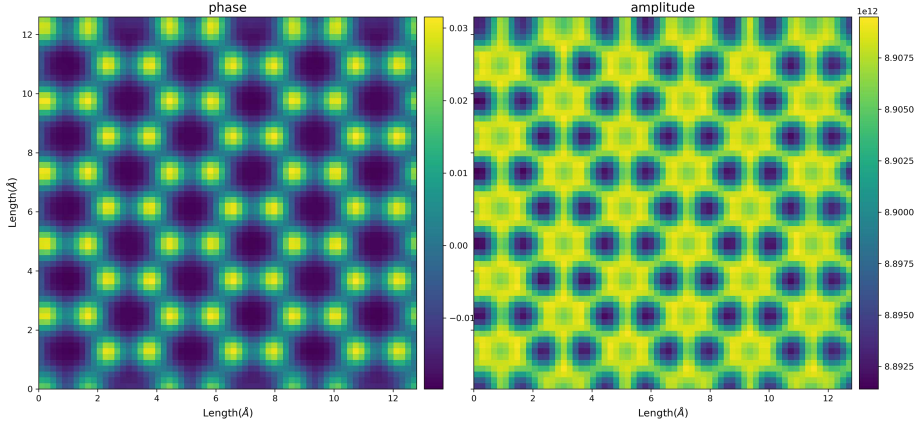


Figure 2.9: The final reconstruction result showing the graphene lattice.

transforming $\chi_\Psi(\mathbf{R}, \mathbf{Q}_p)$ with respect to \mathbf{Q}_p one obtains

$$D(\mathbf{K}_f, \mathbf{Q}_p) = \mathcal{F}_{\mathbf{K}_f}[\chi_\Psi(\mathbf{R}, \mathbf{Q}_p)] = \Psi(\mathbf{K}_f)\Psi^*(\mathbf{K}_f - \mathbf{Q}_p). \quad (2.17)$$

A simple way to obtain $\Psi(\mathbf{Q}_p)$ is slicing the 4D datasets along $\mathbf{K}_f = 0$, then

$$\Psi(\mathbf{Q}_p) = \frac{D^*(0, \mathbf{Q}_p)}{\sqrt{D(0, 0)}}. \quad (2.18)$$

The last step is to perform the inverse Fourier transform with $\Psi(\mathbf{Q}_p)$ and obtain the object function $\Psi(\mathbf{r})$.

2.5.2.1 Contrast transfer function for SSB ptychography

The contrast transfer function is a concept which is used to describe the modulation of contrast in an electron microscopy image as a function of spatial frequency and is determined by the combination of the electron wavelength and the convergence angle. According to the definition [22], the CTF for SSB is

$$|A^*(\mathbf{K}_f)||A(\mathbf{K}_f - \mathbf{Q}_p)|. \quad (2.19)$$

Without any aberrations, the CTF is determined by the overlapped area of the aperture function. Here we should note that when the spatial frequency \mathbf{Q}_p is less than the convergence angle α , the triple overlapped area, in which destructive interference between overlapped double disk overlaps, should be excluded as shown in figure 2.10(a). In the case that \mathbf{Q}_p is larger than α as shown in figure 2.10(a), we only need to consider the area of the double disk overlaps. Hao Yang et al [22] derived the following expression for the CTF of SSB geometrically,

$$\begin{aligned} CTF(\omega) &= \frac{4}{\pi}[\cos^{-1}(\frac{\omega}{2}) - \cos^{-1}(\omega) + \omega\sqrt{1-\omega^2} - \frac{\omega}{2}\sqrt{1-(\frac{\omega}{2})^2}], 0 \leq \omega \leq 1 \\ &= \frac{4}{\pi}[\cos^{-1}(\frac{\omega}{2}) - \frac{\omega}{2}\sqrt{1-(\frac{\omega}{2})^2}], 1 \leq \omega \leq 2, \end{aligned} \quad (2.20)$$

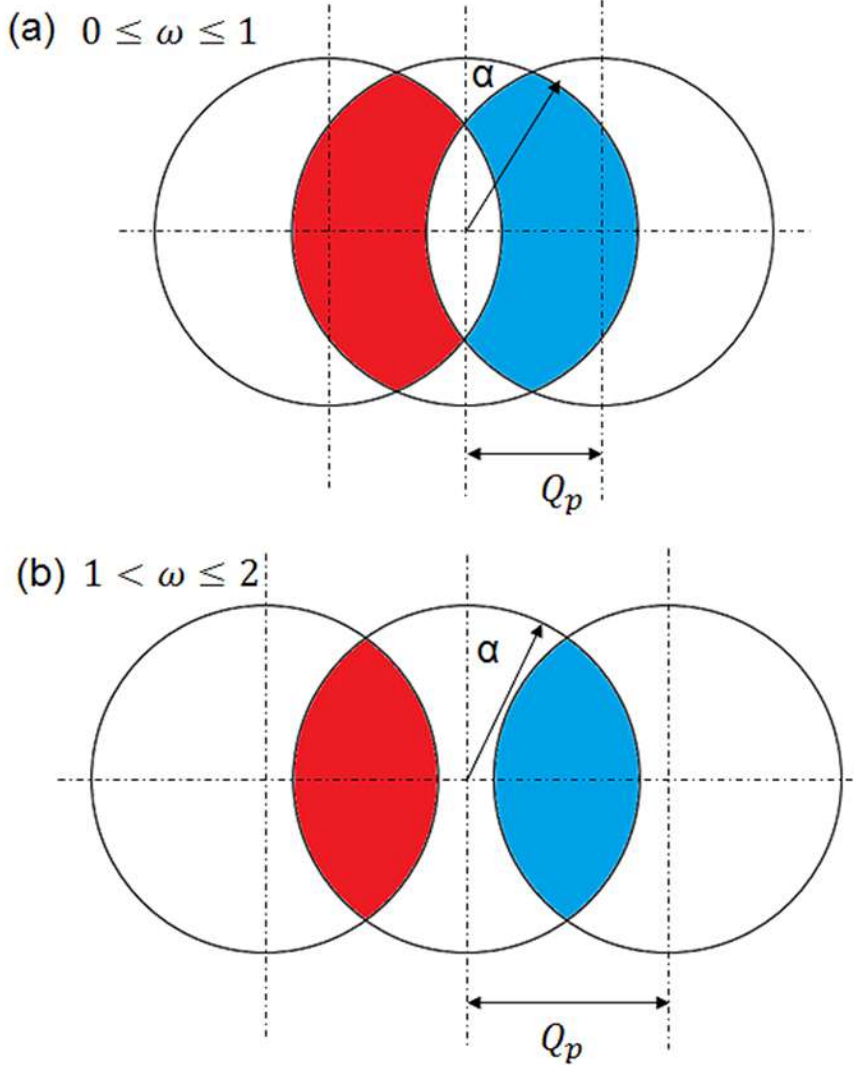


Figure 2.10: Schematic of the useful coherent area with different conditions. This figure is adapted from reference [22].

where ω is the ratio of spatial frequency \mathbf{Q}_p and convergence angle α . The CTF is shown in figure 2.11 below, If we wish to consider the aberrations of the lens as well, we can define a new concept, the complex CTF

$$A^*(\mathbf{K}_f)A(\mathbf{K}_f - \mathbf{Q}_p) = |A(\mathbf{K}_f)||A(\mathbf{K}_f - \mathbf{Q}_p)| * \exp(i(\chi(\mathbf{K}_f - \mathbf{Q}_p) - \chi(\mathbf{K}_f))), \quad (2.21)$$

which is no longer a purely real function but a complex function. Taking one specific frequency \mathbf{Q}_p as an example, when there are no lens aberrations, the phase of the double disk overlaps (colloquially called the trotters) should be zero. However when aberrations are introduced, for instance with a nonzero defocus, the phase is not zero anymore as shown in figure 2.12, which can have a significant influence on the final reconstructed phase image. We can see the phase induced by the aberrations in the double overlapped

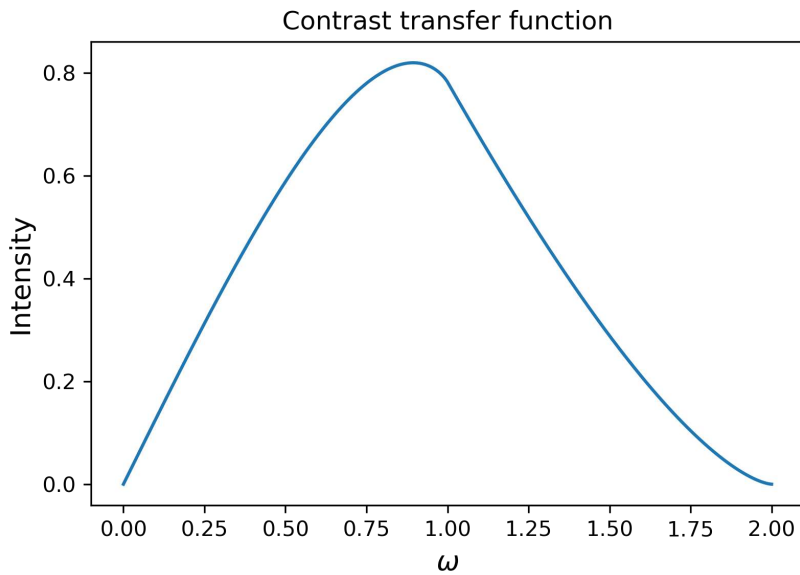


Figure 2.11: CTF of SSB ptychography.

area is not zero anymore. Therefore the aperture function will not just influence the modulus but also directly influence the phases of the object function. This inevitably introduces aberrations in the final reconstructed phase image.

This can be seen clearly in figure 2.13, the left image is a phase image of graphene with the aberration coefficient $C_{32}=250$ nm. One can see the three fold aberrations in the phase image.

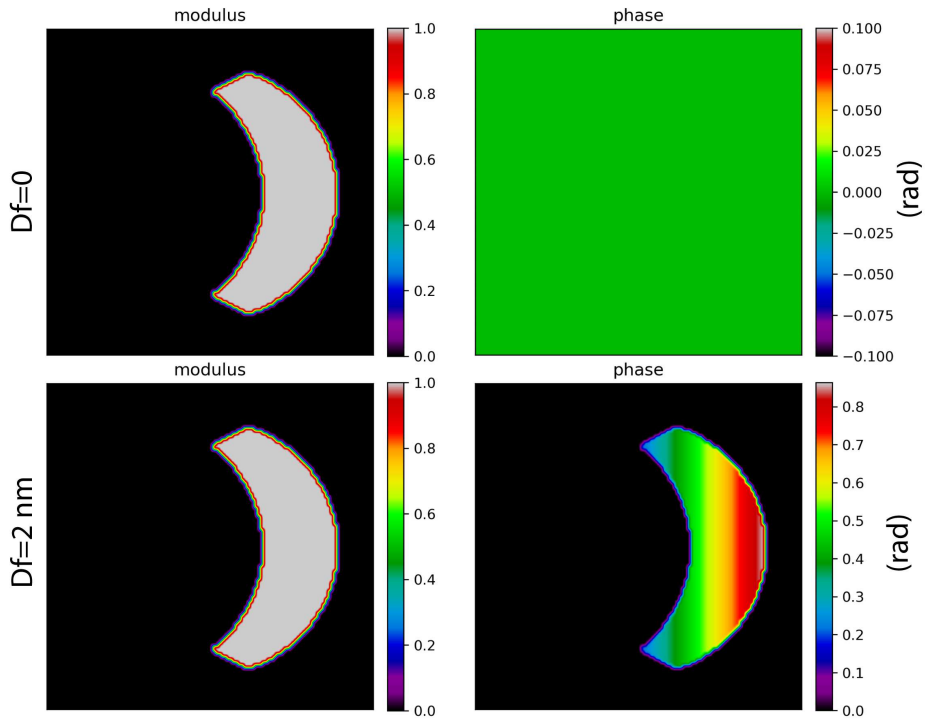


Figure 2.12: The trotter at one specific spatial frequency \mathbf{Q}_p . The top row shows the modulus and phase information without any aberrations. The phase is zero in the double overlap area. The bottom row shows the modulus and phase information with a defocus of 2 nm. The phase now varies with the scattering angle.

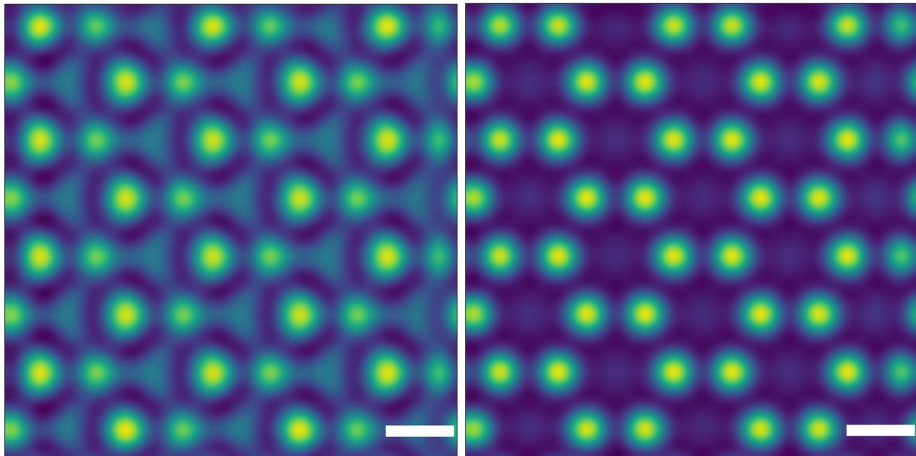


Figure 2.13: Reconstructed SSB phase images of graphene for the simulated 4D datasets with and without aberrations. The left phase image is reconstructed from the simulated 4D dataset with $C_{32}=250$ nm. We can clearly see the influence of the 3 fold astigmatism in the phase image. The right phase image is the equivalent aberration free image. The scale bar is 2 Å.

2.5.3 Iterative ptychography

2.5.3.1 ePIE

The PIE algorithm was first proposed by Faulkner and Rodenburg with a movable aperture function as the illumination function [23]. This method is based on coherent diffraction imaging (CDI), which can also be called scanning CDI. While an aperture function without considering about the geometrical aberrations can not precisely describe the real probe in the experiment. For this reason, a further improvement on PIE algorithm was proposed by Andrew Maiden and Rodenburg. This new method is called ePIE [24]. The basic principle of ePIE is based on optimization algorithms. Using an overlap constraint in real space and an amplitude constraint in Fourier space to find a unique solution for both the probe and object functions. The ePIE algorithm is illustrated in figure 2.14.

The probe function is denoted as $P(\mathbf{r})$, the object function as $O(\mathbf{r})$, the exit wave function as $\Psi(\mathbf{r})$, the wave function in the diffraction plane as $\psi(\mathbf{k})$, and the intensity of experimental intensity as $M(\mathbf{k})$. \mathbf{r} is the two dimensional vector in real space. \mathbf{k} is the two dimensional vector in reciprocal space.

1. First a guess for the probe function $P(\mathbf{r})$ and object function $O(\mathbf{r})$ is provided. This can be a two dimensional array with arbitrary values. At the probe position \mathbf{r}_i , the exit wave function $\Psi(\mathbf{r} - \mathbf{r}_i)$ can be expressed as:

$$\Psi(\mathbf{r} - \mathbf{r}_i) = P(\mathbf{r}) \cdot O(\mathbf{r} - \mathbf{r}_i). \quad (2.22)$$

2. Then in the diffraction plane, the initial diffraction pattern $\Psi(\mathbf{k})$ can be written as the Fourier transform of exit wave function $\Psi(\mathbf{r} - \mathbf{r}_i)$:

$$\Psi(\mathbf{k}) = \mathcal{F}(\Psi(\mathbf{r} - \mathbf{r}_i)). \quad (2.23)$$

3. The third step is the modulus constraint. Using the modulus of the measured diffraction pattern $M(\mathbf{k})$ at the probe position \mathbf{r}_i to replace the amplitude of the diffraction wave function guess, $\Psi(\mathbf{k})$, and keeping the phase of the guessed diffraction pattern. Then one performs the inverse Fourier transform to back propagate to real space,

$$\Psi_{new}(\mathbf{k}) = \sqrt{M(\mathbf{k})} \frac{\Psi(\mathbf{k})}{|\Psi(\mathbf{k})|}. \quad (2.24)$$

4. Next one updates the object function $O(\mathbf{r})$ and probe function $P(\mathbf{r} - \mathbf{r}_i)$. The coefficients α and β are the learning ratios, usually we choose α as 0.1 and β as 1 from experience:

$$P_{update}(\mathbf{r}) = P(\mathbf{r}) + \alpha \frac{O^*(\mathbf{r})[\Psi(\mathbf{r}) - \Psi_{new}(\mathbf{r})]}{(|O(\mathbf{r})|^2)_{max}}, \quad (2.25)$$

$$O_{update}(\mathbf{r}) = O(\mathbf{r}) + \beta \frac{P^*(\mathbf{r})[\Psi(\mathbf{r}) - \Psi_{new}(\mathbf{r})]}{(|P(\mathbf{r})|^2)_{max}}. \quad (2.26)$$

5. After updating the probe function and object function at the position \mathbf{r}_i , one proceeds to the next probe position and repeats the procedure of (1) to (4) until all probe positions have been updated. Then one proceeds to the next iteration.

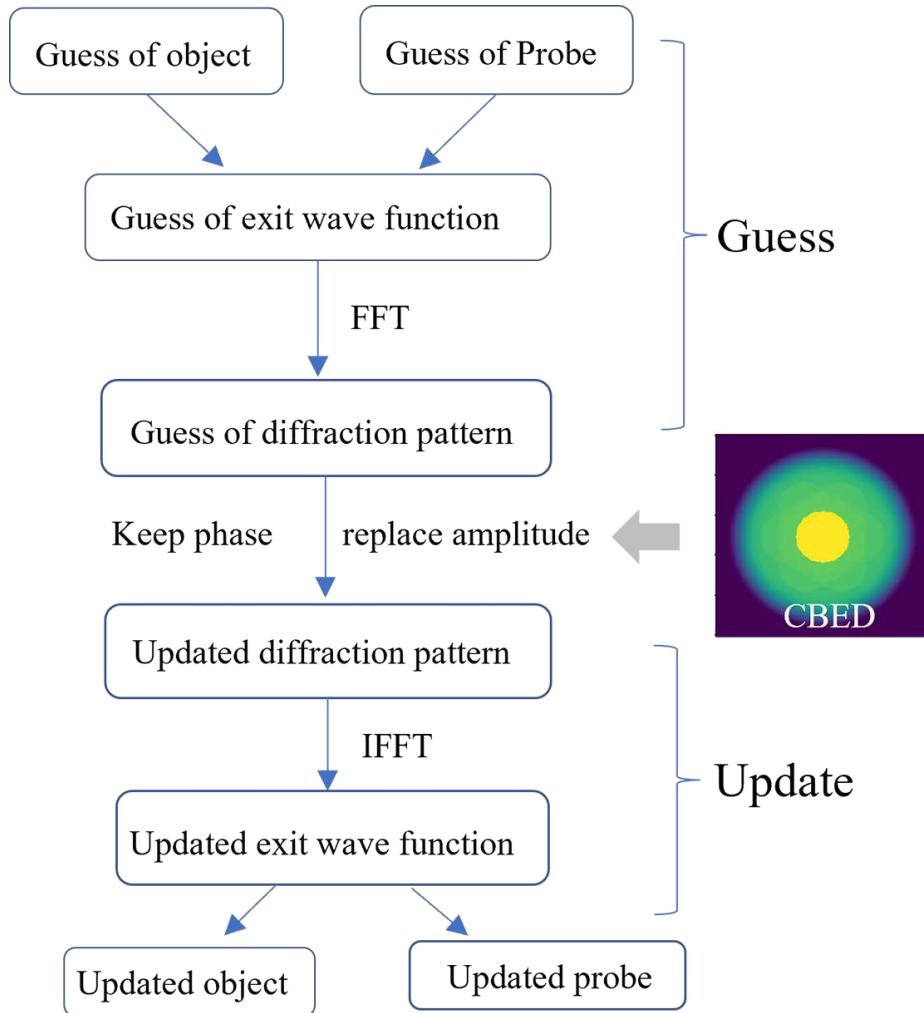


Figure 2.14: Workflow of ePIE algorithm.

2.5.3.2 Multislice ptychography

In this thesis so far, we deal with the samples using the POA model, which is applicable to 2D materials comprising a few layers. However, as detailed in Chapter 2, the POA model becomes inadequate for thicker samples where dynamic scattering mechanisms predominate. Consequently, phase images reconstructed using ptychographic algorithms that rely on POA or the WPOA can exhibit artifacts. To circumvent this limitation, we can introduce a methodology that incorporates multiple scattering mechanisms into the reconstruction process, known as multislice ptychography.

Taking multislice ePIE as an example, this algorithm integrates the dynamic scattering process within the reconstruction framework. The multislice model, as depicted in figure 1.9, offers a simplified explanation of this technique. Here, a sample of thickness T is sliced into n slices at intervals of Δt , within which the POA assumptions are once

again valid. Similar to ePIE, multislice ePIE processes each diffraction pattern within a single iteration cycle, consisting of two primary steps: (1) forward propagation through each slice and (2) backward propagation with concurrent updates. This approach can overcome the limitations of the POA and WPOA and the complexities of thick sample analysis, and enhance the accuracy of the phase image reconstructions.

1. Calculate the exit wave function $\Psi_{1e}(\mathbf{r} - \mathbf{r}_p)$ with the multiplication of object function in the first layer $t_1(\mathbf{r})$ and probe function $P(\mathbf{r} - \mathbf{r}_p)$:

$$\Psi_{1e}(\mathbf{r} - \mathbf{r}_p) = t_1(\mathbf{r})P(\mathbf{r} - \mathbf{r}_p), \quad (2.27)$$

where \mathbf{r}_p denotes the scan probe position.

2. Propagate the exit wave function $\Psi_{1e}(\mathbf{r} - \mathbf{r}_p)$ to the second slice with Fresnel propagator $P(\mathbf{r}, \Delta z)$, which can be viewed as the incident wave for the second slice:

$$\Psi_{2i}(\mathbf{r} - \mathbf{r}_p) = P(\mathbf{r}, \Delta z) \otimes \Psi_{1e}(\mathbf{r} - \mathbf{r}_p), \quad (2.28)$$

where $P(\mathbf{r}, \Delta z) = \frac{1}{i\lambda\Delta z} \exp\left(\frac{i\pi r^2}{\lambda\Delta z}\right)$, \otimes denotes convolution.

3. Calculate the exit wave function from the second slice with the multiplication of object function in the second slice $t_2(\mathbf{r})$ and incident wave function $\Psi_{2i}(\mathbf{r} - \mathbf{r}_p)$:

$$\Psi_{2e}(\mathbf{r} - \mathbf{r}_p) = t_2(\mathbf{r})\Psi_{2i}(\mathbf{r} - \mathbf{r}_p). \quad (2.29)$$

4. Then the incident wave function in the third slice can be $\Psi_{3i}(\mathbf{r} - \mathbf{r}_p) = P(\mathbf{r}, \Delta z) \otimes \Psi_{2e}(\mathbf{r} - \mathbf{r}_p)$ and thus the exit wave function in the third slice can be expressed as

$$\Psi_{3e}(\mathbf{r} - \mathbf{r}_p) = t_3(\mathbf{r})\Psi_{3i}(\mathbf{r} - \mathbf{r}_p). \quad (2.30)$$

5. Repeat these procedures 2 to 4 until obtain the exit wave function from the n th layer $\Psi_{ne}(\mathbf{r} - \mathbf{r}_p)$.

6. The wave function in the detector plane is $\Psi_{ne}(\mathbf{k})$ is the Fourier transform of $\Psi_{ne}(\mathbf{r} - \mathbf{r}_p)$, then using the intensity detected by the detector to update the amplitude of $\Psi_{ne}(\mathbf{k})$:

$$\Psi_{nupdate}(\mathbf{k}) = \frac{\Psi_{ne}(\mathbf{k})}{|\Psi_{ne}(\mathbf{k})|} * \sqrt{I(\mathbf{k})}, \quad (2.31)$$

where $I(\mathbf{k})$ is the diffraction intensity at the probe position \mathbf{r}_p detected by the detector.

7. Transform the updated wave function $\Psi_{nupdate}(\mathbf{k})$ into the real space by inverse FFT:

$$\Psi_{ne}^*(\mathbf{r} - \mathbf{r}_p) = IFFT(\Psi_{nupdate}(\mathbf{k})). \quad (2.32)$$

8. Update the object function $t_n(\mathbf{r})$ and incident function $\Psi_{ni}(\mathbf{r} - \mathbf{r}_p)$ in the n th slice:

$$t_n^*(\mathbf{r}) = t_n(\mathbf{r}) + \alpha \frac{\Psi_{ni}^*(\mathbf{r} - \mathbf{r}_p)}{|\Psi_{ni}(\mathbf{r} - \mathbf{r}_p)|_{max}^2} [\Psi_{ne}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{ne}(\mathbf{r} - \mathbf{r}_p)], \quad (2.33)$$

$$\Psi_{ni}^*(\mathbf{r} - \mathbf{r}_p) = \Psi_{ni}(\mathbf{r} - \mathbf{r}_p) + \beta \frac{t_n^*(\mathbf{r})}{|t_n(\mathbf{r})|_{max}^2} [\Psi_{ne}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{ne}(\mathbf{r} - \mathbf{r}_p)], \quad (2.34)$$

where α and β are update parameters.

2.5. INTRODUCTION TO PTYCHOGRAPHIC ALGORITHMS

9. Back propagate the updated incident wave function $\Psi_{ni}^*(\mathbf{r} - \mathbf{r}_p)$ into the $n - 1$ slice and yield the exit function from $n - 1$ th slice:

$$\Psi_{n-1e}^*(\mathbf{r} - \mathbf{r}_p) = P(\mathbf{r}, -\Delta z) \otimes \Psi_{ni}^*(\mathbf{r} - \mathbf{r}_p). \quad (2.35)$$

10. Update the object function $t_{n-1}(\mathbf{r})$ and incident function $\Psi_{n-1i}(\mathbf{r} - \mathbf{r}_p)$ in the $n - 1$ slice and obtain:

$$t_{n-1}^*(\mathbf{r}) = t_{n-1}(\mathbf{r}) + \alpha \frac{\Psi_{n-1i}^*(\mathbf{r} - \mathbf{r}_p)}{|\Psi_{n-1i}(\mathbf{r} - \mathbf{r}_p)|_{max}^2} [\Psi_{n-1e}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{n-1e}(\mathbf{r} - \mathbf{r}_p)], \quad (2.36)$$

$$\Psi_{n-1i}^*(\mathbf{r} - \mathbf{r}_p) = \Psi_{n-1i}(\mathbf{r} - \mathbf{r}_p) + \beta \frac{t_{n-1}^*(\mathbf{r})}{|t_{n-1}(\mathbf{r})|_{max}^2} [\Psi_{n-1e}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{n-1e}(\mathbf{r} - \mathbf{r}_p)]. \quad (2.37)$$

11. Repeat the procedures in 8-10 until reach the first slice. And update the object function in the first slice $t_1(\mathbf{r})$ and probe function $\Psi_{1i}(\mathbf{r} - \mathbf{r}_p)$:

$$t_1^*(\mathbf{r}) = t_1(\mathbf{r}) + \alpha \frac{\Psi_{1i}^*(\mathbf{r} - \mathbf{r}_p)}{|\Psi_{1i}(\mathbf{r} - \mathbf{r}_p)|_{max}^2} [\Psi_{1e}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{1e}(\mathbf{r} - \mathbf{r}_p)], \quad (2.38)$$

$$\Psi_{1i}^*(\mathbf{r} - \mathbf{r}_p) = \Psi_{1i}(\mathbf{r} - \mathbf{r}_p) + \beta \frac{t_1^*(\mathbf{r})}{|t_1(\mathbf{r})|_{max}^2} [\Psi_{1e}^*(\mathbf{r} - \mathbf{r}_p) - \Psi_{1e}(\mathbf{r} - \mathbf{r}_p)]. \quad (2.39)$$

12. Then go to the next iteration with the updated probe function and object function in each slice as the input functions.

2.5.4 Conclusion

This concludes my review of ptychographic algorithms. In this chapter, I have explained why the phase information is important and introduced the basic concepts of ptychography. Ptychography is not yet widely used in electron microscopy, but can be expected to become increasingly popular as faster detectors facilitate experiments.

CHAPTER 2. LITERATURE REVIEW

Characterizing oxide samples with HAADF and EELS

3.1 Abstract

In this chapter, I use conventional HAADF combined with EELS to characterize the mixed layered structures of oxide samples. As previously explained, intensity in HAADF images is proportional to roughly the square of the atomic numbers of the elements encountered. However, if the atomic numbers of two elements are close, the contrast of two atomic columns can be indistinguishable. EELS can provide complementary information about which elements are where using their spectroscopic signatures. Therefore, HAADF combined with EELS can enable us to solve this problem.

In this study, I employed HAADF and EELS to characterize the layered structures of LaTiO_3 and LaVO_3 with atomic resolution. Importantly for the EELS, the core loss spectra of Ti and V are distinguishable, compared to many EELS edges which overlap.

3.2 The structure of LaTiO_3 and LaVO_3 .

The sample was grown by the molecular beam epitaxy (MBE) technique. As illustrated in in figure 3.1 , the LaVO_3 and LaTiO_3 were grown layer by layer on an SrTiO_3 substrate. The remaining questions are:

- (1) How to identify a mixed layered structure with electron microscopy?
- (2) Does the valence of vanadium and titanium atoms change at the boundary? To answer these questions, I used the Titan1 electron microscope in EMAT to study this oxide sample. The Titan1 electron microscope is an aberration corrected instrument equipped with a HAADF detector and an EELS spectrometer, which is suitable for our study.

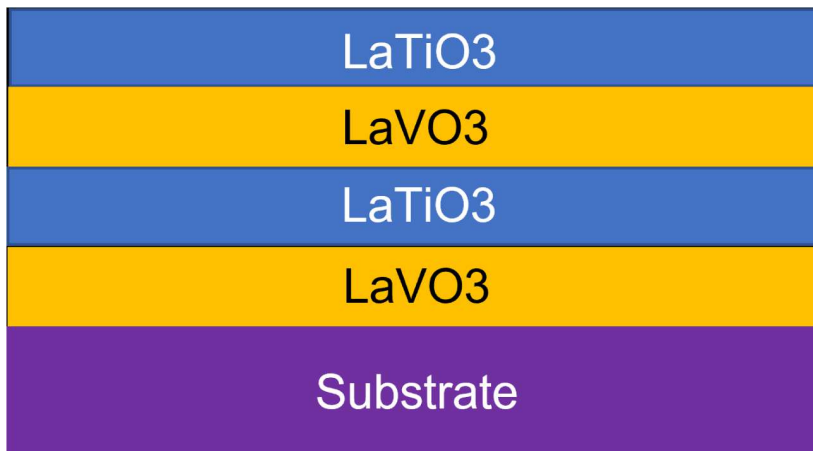


Figure 3.1: Schematic diagram of the LaVO_3 and LaTiO_3 layered structures.

3.3 Characterization of $\text{LaTiO}_3/\text{LaVO}_3$ with HAADF

To prepare the sample we first used focused ion beam (FIB) to thin the sample to several tens of nanometers of thickness, sufficiently thin to allow the incident electrons to transmit through the sample without too large a background signal. A HAADF image of the sample is shown in figure 3.2. We used a 300 kV accelerating voltage and a 20 mrad convergence angle. From the HAADF image, we can distinguish four different layers, as labeled in figure 3.2. They are the Pt and carbon layers which are deposited on the sample to protect it from being damaged by the incident ions of the FIB. The $\text{LaVO}_3/\text{LaTiO}_3$ and SrTiO_3 layers can also be distinguished not only by the contrast but also their lattice spacings. However for the $\text{LaVO}_3/\text{LaTiO}_3$ layers, it is hard to distinguish the layered structure within these layers because the atomic number of V is 23 and that of Ti is 22. They are too close for the HAADF image to easily distinguish. I will investigate further in figure 3.3.

Figure 3.3 (a) shows an atomic resolution HAADF image of the $\text{LaVO}_3/\text{LaTiO}_3$. From the image, we can see the perovskite structure of the $\text{LaVO}_3/\text{LaTiO}_3$. The brightest spots in the image are the heavy La atomic columns while the lighter V and Ti columns are the columns with lower intensity among the La atomic columns as labeled in figure 3.3 (b). In this HAADF image, we cannot distinguish the layered structure of LaVO_3 and LaTiO_3 by comparing the intensity of the atomic columns. Figure 3.3 (b) gives an enlarged image of the sample as labeled in figure 3.3 (a). In this figure, we can distinguish La atomic columns by the contrast easily. While for the adjacent atomic columns, the contrast of Ti or V are almost equal. We plot the line profile of the La, Ti (V) atomic columns along the red arrow in figure 3.3 (b) as shown in figure 3.3 (c). It is hard for us to say which columns are Ti and which columns are V. One thing we should note is that the oxygen atomic columns are hidden in the image because the contrast of oxygen atoms is too weak and hidden by the strong scattering of La and Ti columns.

3.4. CHARACTERIZATION WITH EELS

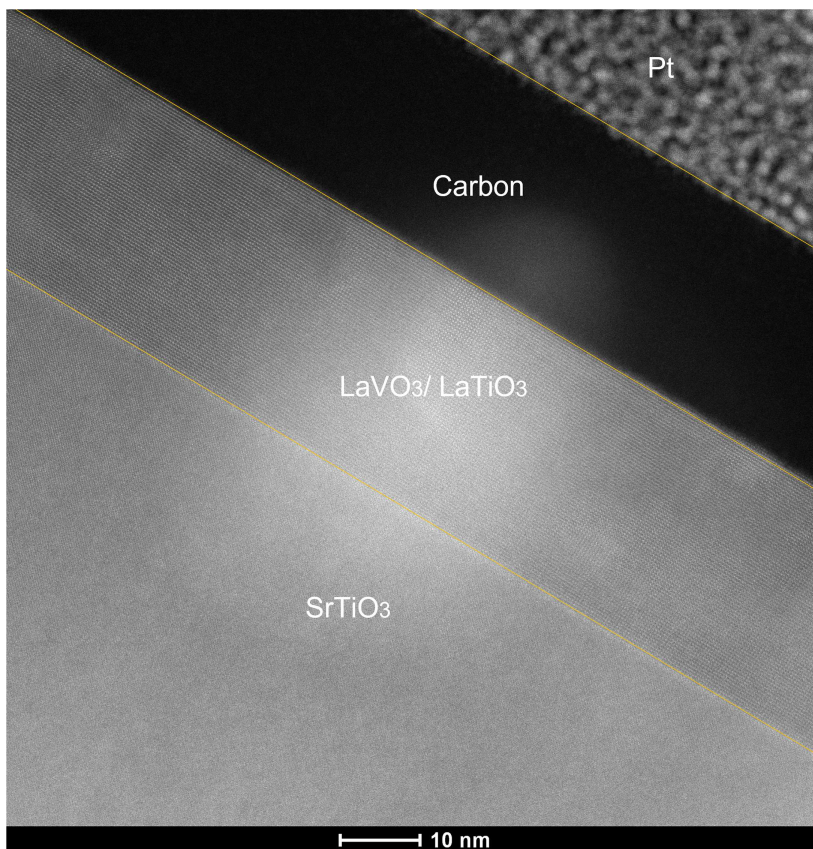


Figure 3.2: HAADF image of LaVO₃/LaTiO₃. The carbon and Pt layers are deposited on the sample to protect the sample during preparation by FIB. The white area in the image is the contamination of carbon.

3.4 Characterization with EELS

From the HAADF image, it is difficult to distinguish between the Ti and V atomic columns. However, the spectroscopic signature of Ti and V elements is quite different, which provides a way to separate these two elements with EELS. This procedure is demonstrated in figure 3.4. We took a two dimensional EELS spectrum in the specific area. As shown in figure 3.4(a), the area labeled by the red rectangle is where we took the EELS spectrum image. Then we integrated the core loss signal from the entire area producing the spectrum shown in figure 3.4 (b). The averaging provides a high signal to noise ratio. In the figure 3.4, I have labeled the L edges of Ti and V. Using the integrated intensity of these edges we can in principle distinguish Ti and V layers.

We also took 2D spectrum imaging of the sample as shown in figure 3.6. Figure 3.6(a) is the HAADF survey image, we can see the distorted atomic columns in the image. This is because of the sample drift due to the very long acquisition time. Figure 3.6 (b), (c) and (d) show the spectrum images of V, O and Ti. We can distinguish the V layers and Ti layers by the intensity of the contrast in (b) and (d). Even though the sample drift

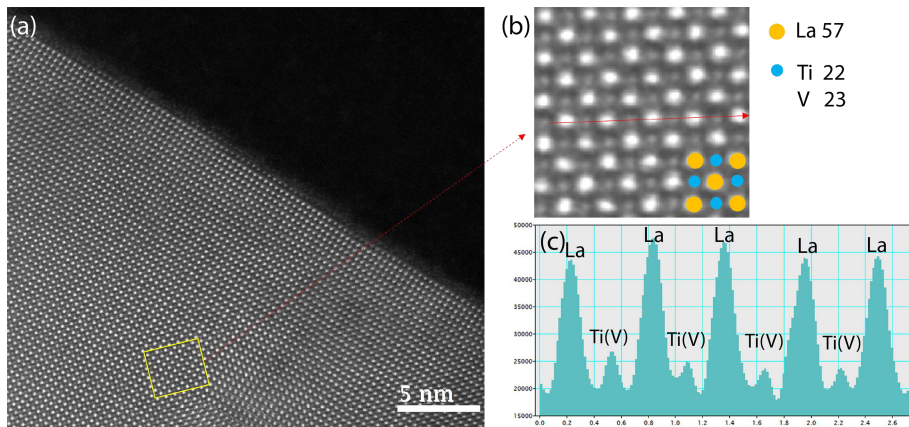


Figure 3.3: Quantitative analysis of $\text{LaVO}_3/\text{LaTiO}_3$. (a) Atomic resolution HAADF image of the sample. (b) Enlargement of the area in the yellow rectangle in (a). Overlaid are orange disks to represent the La columns with atomic number 57, and blue disks representing Ti or V atoms, the corresponding atomic numbers of which are 22 and 23. (c) Line profile of atomic columns along the red arrow overlaid on the image in (b). The La atomic columns can easily be distinguished from the Ti and V columns, but it is much harder to distinguish between the Ti and V atomic columns.

cannot be avoided, we still obtain 2D spectrum images with atomic resolution.

For EELS mapping, in order to obtain core loss spectra with high signal to noise ratio (SNR), the dwell time generally has to be tens of milliseconds. To avoid sample drift due to the long dwell time, we chose to take line scan spectra. The results are shown in figure 3.5. Figure 3.5(a) shows the HAADF survey image before we took the line scan. The red arrow indicates the line scan trace of the probe. Figure 3.5(b) is the raw spectrum image of the line scan as captured by the EELS camera. In this figure, the x axis denotes the energy loss of the spectrum, the y axis denotes the scan of the probe along the sample. We can see the L edges of Ti and V and the K edge of O as labeled in the figure (note there is no background subtraction used here). Then I subtracted the background and integrated the signature peaks of Ti, V and O elements. The final result is shown in figure 3.5(c). The x axis denotes the scan distance in the real space and the y axis denotes the intensity of the integrated spectra. We can see the peaks of Ti and V oscillate and are perfectly mismatched, showing the layering of the LaVO_3 and LaTiO_3 .

3.5 Conclusion

To solve the $\text{LaVO}_3/\text{LaTiO}_3$ layered structure, we combined HAADF and EELS to successfully show this sample has the expected layered structure. One thing we should note is that the HAADF technique is not sensitive to the light atoms which are adjacent to the heavy atomic columns. Therefore, either EELS or more advanced imaging techniques must be utilized. This is exactly what we will explore in the next chapter.

3.5. CONCLUSION

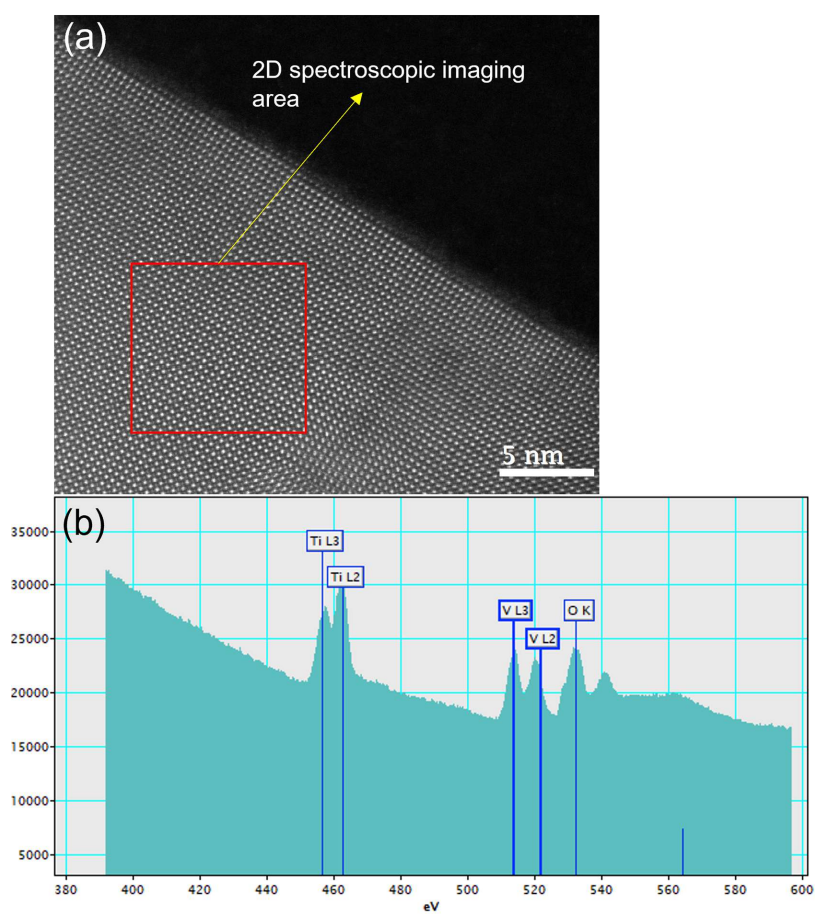


Figure 3.4: Spectrum imaging. (a) HAADF image showing the region used for spectrum imaging in a red rectangle. (b) Integrating the core loss spectra over the region inside the red rectangle area in (a) clearly shows the Ti L edge and V L edge sufficiently separated for elemental analysis.

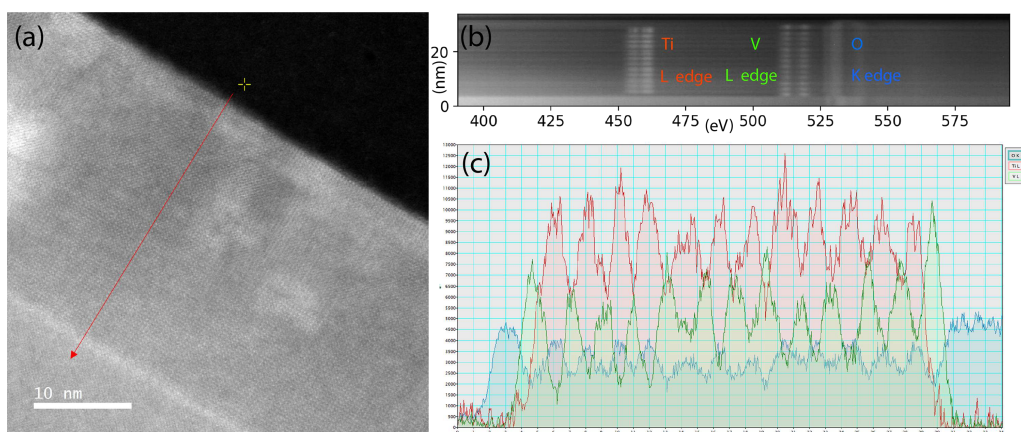


Figure 3.5: EELS line scan over the LaVO_3 and LaTiO_3 . (a) survey ADF image of the sample, the red line labels the scan trace of the probe. (b) Image of the line scan core loss spectrum image obtained from the trace as shown in the red arrow in (a). The x axis is energy loss. The y axis is the distance from the scan start point to the end point. From the figure, we can see the L edge of Ti and V and the K edge of the O clearly. (c) Plot of the line trace integrated across the EELS camera. From the figure, we can see the intensity of the Ti L edge and V L edge oscillate and the peaks of each two elements are perfectly out of phase, as expected.

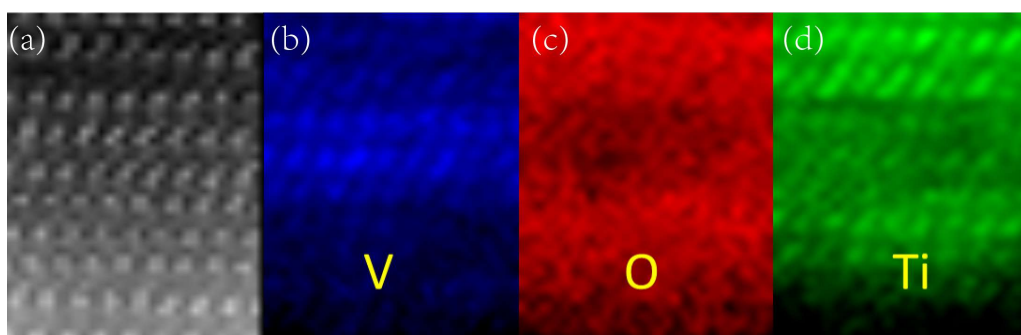


Figure 3.6: 2D spectrum imaging of $\text{LaVO}_3/\text{LaTiO}_3$. (a) HAADF survey image. The spectrum image of V (b), O (c), and Ti (d). From figure (a), (b), (c) and (d), we can distinguish the Ti and V elements with atomic resolution, but the sample drift is significant.

Dose efficiency in low dose condition with Timepix3 detector

4D STEM can provide many advantages compared to conventional STEM methods, such as mapping the distribution of the electromagnetic field and high dose efficiency. Although many advances have been achieved in direct electron detectors previously, the recording speed of most cameras is still relatively slow. Slow scans are a major problem because they may result in sample drift and contamination. In this chapter, an event driven detector Timepix3 is used to overcome the bottleneck of speed with the dwell time reaching as short as 100 ns. In the low dose condition, the Timepix3 detector is far faster than frame based detectors. I have also tested various ptychographic methods with the data collected from the Timepix3 detector. The results show that ptychographic methods are highly dose efficient.

This chapter is based on the following paper:

D. Jannis, C. Hofer, C. Gao, X. Xie, A. Béch e, T.J. Pennycook, J. Verbeeck, Event driven 4D STEM acquisition with a Timepix3 detector: Microsecond dwell time and faster scans for high precision and low dose applications, Ultramicroscopy 233 (2022) 113423

My contribution includes writing the ptychographic codes to analyse the experimental results.

4.1 Introduction

Material science has witnessed a growing preference for STEM, primarily due to the capability of STEM to yield easily interpretable images. In ADF STEM, image contrast is obtained from the scattered electrons captured by an annular detector. High-resolution images, typically 1024 by 1024 pixels, are captured during experiments, with a relatively fast dwell times, for instance with a 10 μ s dwell time resulting in a total acquisition time of approximately 10 seconds. This time is often acceptable for preventing sample drift and contamination by mobile carbon on the sample, although faster scans are also often performed and multiple fast scans can be aligned to obtain high signal to noise images free of drift .

To retrieve more information, segmented detectors were invented to detect local electromagnetic fields by measuring the center of mass of the electron scattering [25]. This method not only has advantages in terms of dose efficiency compared to ADF, but also excels in imaging light atoms within heavy atomic columns. Recognizing the importance of precise information about scattered electron distribution, pixelated detectors have been introduced, marking a revolutionary advancement in the ability of electron microscopy to retrieve information about specimens.

This revolution is known as 4D STEM. As previously mentioned, despite the ptychographic methods and other advanced algorithms having been proposed in 1960s [20] and first demonstrated in the 1990's [26] , their adoption in the electron microscopy community has been limited, largely due to the speed constraints of cameras and the computational requirements for the method. The advent of fast cameras and more powerful computers is now changing this. Fast cameras enable the collection of diffraction patterns with a significantly faster timeframe compared to previous camera technologies, with direct electron detection providing much higher detection efficiency. More powerful computers make the data volumes and processing much more manageable than in the 1990's. However, the frame rate of most cameras still limits the scan speed of 4D STEM quite severally compared to the conventional scan dwell time of STEM, posing challenges for all applications with drift but is also a particular challenge for beam-sensitive materials that demand a considerably lower dose with a fast scan speed as the probe current can only be reduced so much practically.

To address this limitation, one option is to employ a defocused probe to scan across the sample, thereby minimizing the scan positions in real space. Nevertheless, this strategy comes with a trade-off, as it compromises the ADF – a heavily defocused ADF image will typically be blurred beyond usefulness, aside perhaps from some very low resolution Z-contrast. Furthermore, iterative ptychography is required in such a defocused probe configuration, which generally involves a longer processing time compared to direct ptychography.

In this chapter, we use an event driven detector called the Timepix3 instead of frame-based detectors to record the diffraction patterns. We can easily achieve a dwell time of 1 μ s which is already comparable with the scan speeds used in normal high scan speed STEM. Significantly faster scan speeds can also be employed if desired. What is more, we can reconstruct higher signal to noise ratio images with ptychography compared with other methods. Furthermore, the aberrations of the probe can be removed after taking the data using ptychography.

4.2 Introduction to the Timepix3 camera

The camera utilized in this experiment is the AdvAPIX TPX3, a Timepix3-based detector with a sensitive silicon layer thickness of approximately $300\ \mu\text{m}$ which we adapted for use in the electron microscope with a custom designed and in house built vacuum tight housing with a retractable mount. Each individual Timepix3 chip contains 256 by 256 pixels. The size of each pixel is $56\ \mu\text{m} \times 56\ \mu\text{m}$. Unlike conventional cameras that solely record the intensity of incident electrons on the sensor of incident electrons accurately within a given frame, the Timepix3 can also provide an unprecedented temporal resolution of $1.6\ \text{ns}$. This means not only can it capture the positions, but it can also record the time of arrival for each incident electron with very high time resolution. The maximum readout speed is 40 million hits per second for a single Timepix3 chip, which means the maximum beam current to deal with is equivalent to $6.4\ \text{pA}$. The innovative approach of combining spatial and temporal information proves to be invaluable in ultra-fast 4D-STEM experiments, as illustrated in the accompanying figure 4.1.

As one must capture an image from the camera at each probe position in 4D STEM, the scan speed cannot exceed that of the camera. Therefore in contrast to conventional pixelated detectors the capabilities of the Timepix3 are exceptional. For instance, with cameras like the Merlin detector operating in 6-bit mode, the frame rates reach 2400 fps. However, the corresponding time per frame is approximately $417\ \mu\text{s}$ — far longer than the typical dwell time in conventional STEM mode. Even in 1-bit mode, where the dwell time is increased to 14400 fps, the corresponding time per frame remains at $69\ \mu\text{s}$ [14]. Although faster than the 6-bit mode, it is still relatively slow when dealing with large datasets, such as 1k by 1k probe position grids. Furthermore the 1-bit mode will not count more than one electron hit in each pixel before saturating, which is a far greater problem when the scan speed is low.

Frame-based detectors exhibit a readout speed that is independent of the dose itself. As depicted in the accompanying figure 4.3, even in low-dose mode, where most pixels register as zeros, an entire 2D array of whichever bit depth is used must be transferred. This can be a very significant amount of data which takes time to transfer but which does not add information if one realizes that if a hit does not occur it should be a zero count, and that these zeros do not need to be described in detail. On the other hand, event-driven detectors, such as the Timepix3, only record information about the pixels that are hit. Consequently, at low-doses, the volume of data generated is dramatically smaller compared to frame-based detectors. This allows the Timepix3 detector to outpace frame-based detectors in terms of speed for low probe currents. Higher probe currents could be achieved using a quad chip setup and a faster connection than the USB 3.0 connection used with the Advapix readout device.

4.3 Processing the raw data

In Antwerp a Timepix3 camera was mounted on an aberration-corrected FEI Themis-Z, with synchronization between the camera and scan engine facilitated by a 10 MHz reference clock. To transform the raw data into a new 4D dataset, a Python script was written to convert the raw data into npy-formatted 4D datasets by constructing a list

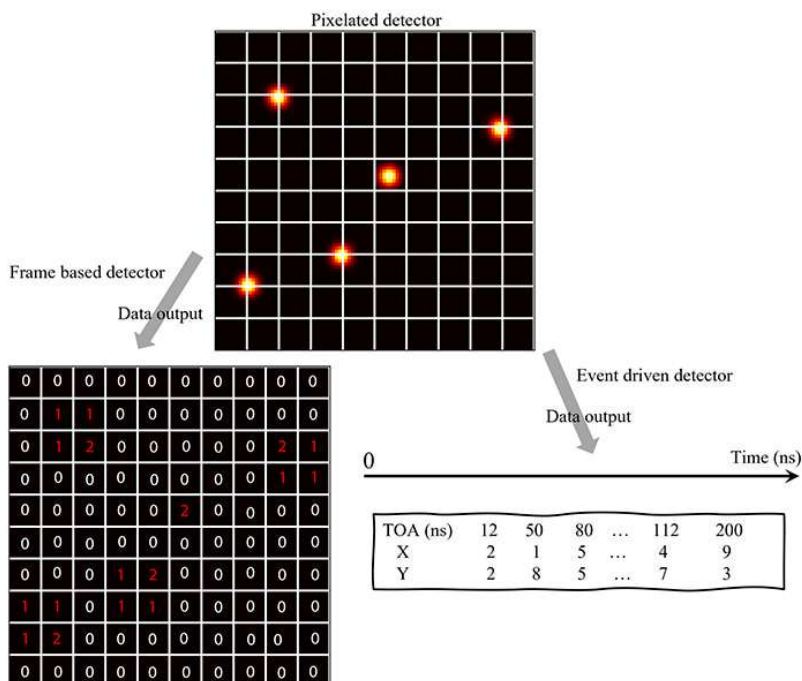


Figure 4.1: Diagram illustrating the output distinctions between frame-based and event-driven detectors. In frame-based detectors, the output manifests as a 2D array, regardless of if most of the pixels contain only zero counts, whereas for event-driven detectors, the output takes the form of a list of numbers detailing just the position on the detector and the time of each hit in nanoseconds.



of bins with intervals corresponding to the probe dwell time, followed by organizing the time sequence and positions in the pixelated detector of the raw data into these bins. A visual representation is provided in figure 4.2. Operating at a voltage of 200 kV and using a convergence angle of 20 mrad, with a 2048x2048 pixel scan, the raw data occupies a modest 2.3 GB using the relatively low dose of $2000 e^-/\text{\AA}^2$. However, upon conversion to a 4D dataset, including the zeros, the space requirement surges to 64 GB. Consequently, a substantial amount of memory is needed on the computer for efficient handling of the data. Obviously it would also be better if we could work directly with the sparse data, but so far we have not found a way to do so in a practical way with even SSB after the FFT step. Although sparse FFT algorithms exist, we did not find a significant benefit in speed over the highly optimized non-sparse FFT libraries.









4.4 Experimental Results

4.4.1 2D material: WS_2

We initially tested the Timepix3 camera using 2D WS_2 . The experimental parameters included an accelerating voltage of 60 kV and a 25 mrad convergence angle. The probe

4.4. EXPERIMENTAL RESULTS

Original data collected from Timepix3				2.3 GB
	Fri_Dec_4_14_35_52_2020_STEM_200kV...	12/4/2020 2:36 PM	T3P File	2,549,334 KB
	Fri_Dec_4_14_35_52_2020_STEM_200kV...	12/4/2020 2:36 PM	INFO File	1 KB

4D datasets remade from the original data				64 GB
	MOF(19)_0.npy	1/21/2021 8:13 AM	NPY File	8,388,609 KB
	MOF(19)_1.npy	1/21/2021 8:28 AM	NPY File	8,388,609 KB
	MOF(19)_2.npy	1/21/2021 8:48 AM	NPY File	8,388,609 KB
	MOF(19)_3.npy	1/21/2021 9:10 AM	NPY File	8,388,609 KB
	MOF(19)_4.npy	1/21/2021 9:21 AM	NPY File	8,388,609 KB
	MOF(19)_5.npy	1/21/2021 9:30 AM	NPY File	8,388,609 KB
	MOF(19)_6.npy	1/21/2021 9:40 AM	NPY File	8,388,609 KB
	MOF(19)_7.npy	1/21/2021 9:49 AM	NPY File	8,388,609 KB

2048x2048x128x128 (datatype=uint8)

Figure 4.2: Raw data collected with the Timepix3 camera and the 4D dataset as converted by the python script.

was scanned over 1024×1024 positions, with dwell time of $6 \mu\text{s}$. The estimated dose amounted to $6000 e^-/\text{\AA}^2$. Both the CBED and position-averaged CBED results are illustrated in figure 4.3. Upon processing the 4D dataset, BF, ABF, and ADF images

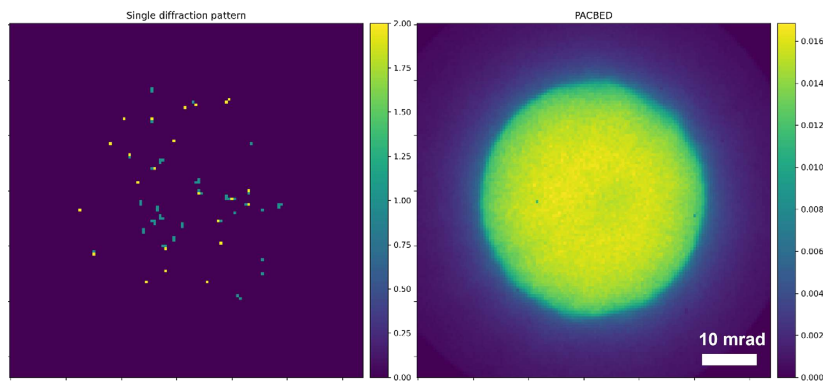


Figure 4.3: Example of a single CBED pattern from the Timepix3 at the dose of $6000 e^-/\text{\AA}^2$ (left). The image on the right is the position averaged CBED pattern from the same dataset.

were reconstructed, as shown in figure 4.4. Under these low-dose conditions, all these conventional images exhibit a poor signal-to-noise ratio as one would expect. In the FFTs, the spots of the second ring of the hexagonal lattice are barely discernible. The ABF image, lacking the channeling effect that provides contrast for the modality in thicker materials is not present for a single-layer material, and unsurprisingly displays a low SNR. As for the ADF image, the scarcity of incident electrons resulted in a minimal number of scattered electrons at higher angles, resulting in the comparatively weak contrast seen in the ADF image.

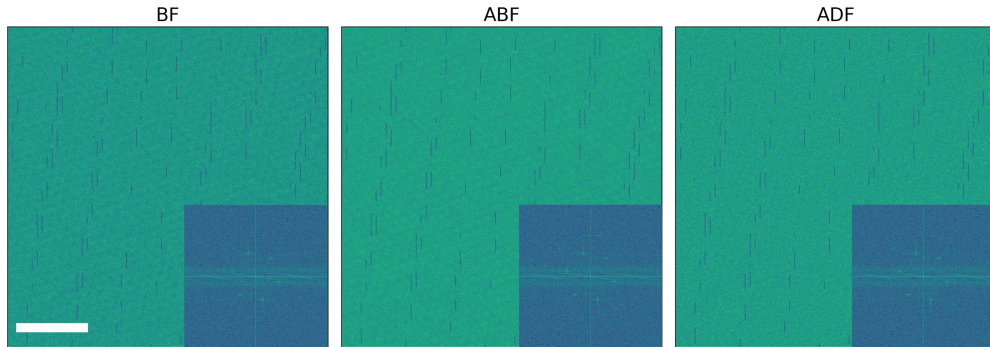


Figure 4.4: BF, ABF and ADF images reconstructed from the collected 4D dataset. In each image, the corresponding FFT is shown in an inset. Streaks appear in the BF and ABF images due to temporary saturation of the timepix3 data stream due to the current used. Scale bar is 2 nm.

Our next step for ptychographic reconstruction involves performing an FFT with respect to probe positions on the 4D dataset. Through this process, we transform the 4D dataset $M(\mathbf{r}_p, \mathbf{k}_f)$ into the G dataset, as detailed in Chapter 3. Despite the sparsity of the individual diffraction patterns, the new dataset G exhibits a high SNR after the FFT is performed. In the accompanying figure 4.5, I have selected frequencies with high intensity and presented the amplitude and phase information of the double disk overlaps. The white circle designates the bright field edges, while the red and blue circles label the left and right scattered electrons. Ideally, in the double-overlap areas, the phase should be uniform. However, in figure 4.5, we observe phase variations due to residual aberrations in the probe.

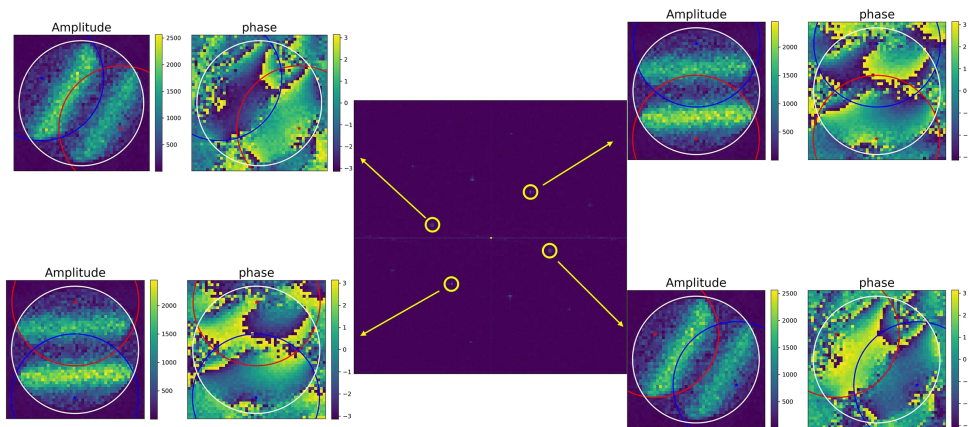


Figure 4.5: The double disk overlaps, also known as the trotters colloquially, for different spatial frequencies.

The final reconstructed phase image is shown in figure 4.6. From the phase image, we can see clearly that the contrast of WS_2 is much better than the conventional BF, ABF and ADF images. From the inserted FFT of figure 4.6, we can observe that the third ring of spots from the hexagonal lattice are present, which are not present in figure 4.4 using

4.4. EXPERIMENTAL RESULTS

the other signals. These results show that SSB ptychography is significantly more dose efficient.

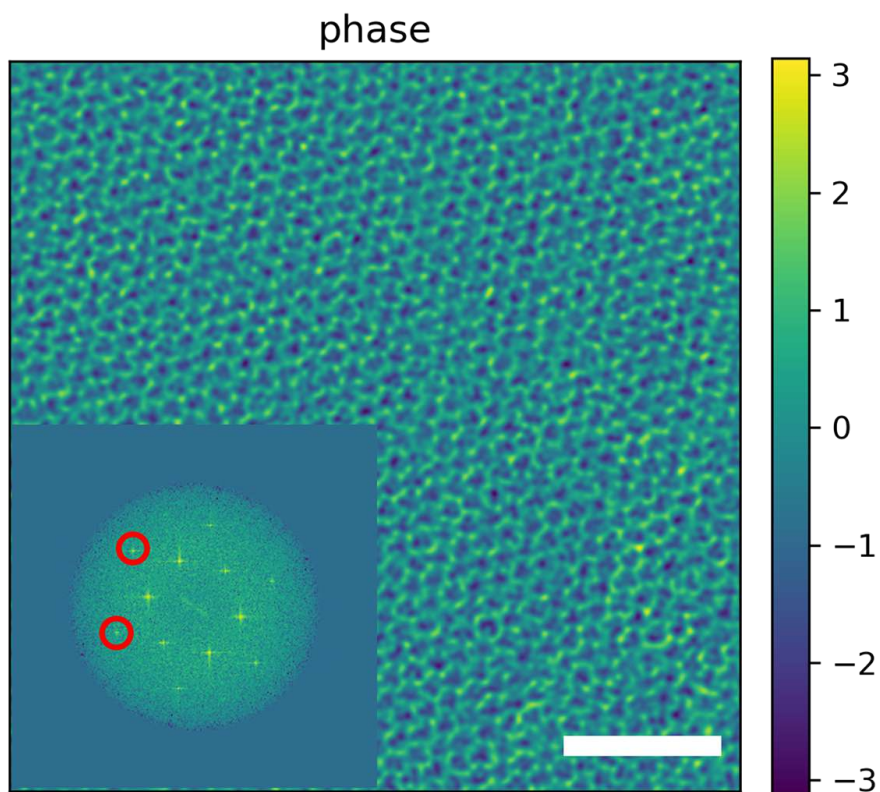


Figure 4.6: The phase information of WS₂ reconstructed from the 4D dataset. The FFT image is inserted in the phase image. The red circles labels the third ring of spots. The scale bar is 2 nm.

4.4.2 Zeolite

We also tested the camera with a silicalite-1 zeolite sample. The voltage we used is 200 kV, a convergence angle of 20 mrad, and the dwell time is $1 \mu\text{s}$ with the dose we estimated to be $300e^-/\text{\AA}^2$. Under such low dose conditions, we can observe that the contrast of the reconstructed BF, ABF and ADF in figure 4.7 is quite poor, similar to the results of WS_2 in figure 4.4, although here we are at a significantly lower dose with a thicker sample.

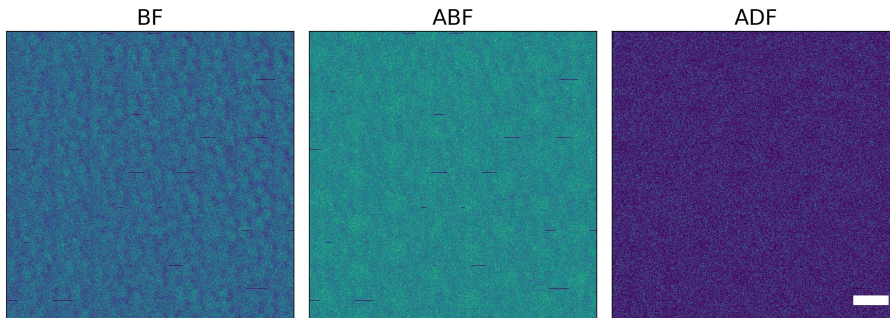


Figure 4.7: The reconstructed BF, ABF and ADF images from a silicalite-1 zeolite. The scale bar is 1 nm.

Then we performed iCoM and SSB reconstructions on the recorded 4D dataset. The results are shown in the figure 4.8. Obviously both iCoM and SSB results show much better signal to noise ratios compared with figure 4.7. Furthermore from the FFT image, we can see clearly that the SSB algorithm works better than iCoM, providing a stronger signal at this low dose.

4.5 Aberration correction

Figure 4.5 shows experimental double disk overlaps, in which the presence of residual aberrations can be seen. Such residual aberrations introduce artifacts in the final reconstructed phase image. These artifacts pose a challenge in accurately interpreting the data. The sensitivity of the phase images means that the polarity of a lattice can actually be reversed by such residual aberrations, even when the ADF looks perfectly fine [27]. Obviously, if the polarity of a lattice can be reversed, this will entirely prevent precise analysis such as required for detection of the subtle changes in charge distribution due to bonding. Therefore, it is very useful to employ a method capable of measuring and eliminating these artifacts, as can be done in ptychography. To identify aberrations, the singular value decomposition (SVD) based method, proposed in reference [28], often proves to be an effective solution. In essence, this method enables the calculation and removal of aberration coefficients associated with the lenses. For a detailed understanding of this technique, please refer to the aforementioned paper. To facilitate its application, I have developed a Python script based on the SVD method for calculating and removing residual aberrations as part of the pyPtychoSTEM package [29].

For aberration-corrected SSB ptychography, accurate calculation of aberration coeffi-

4.5. ABERRATION CORRECTION

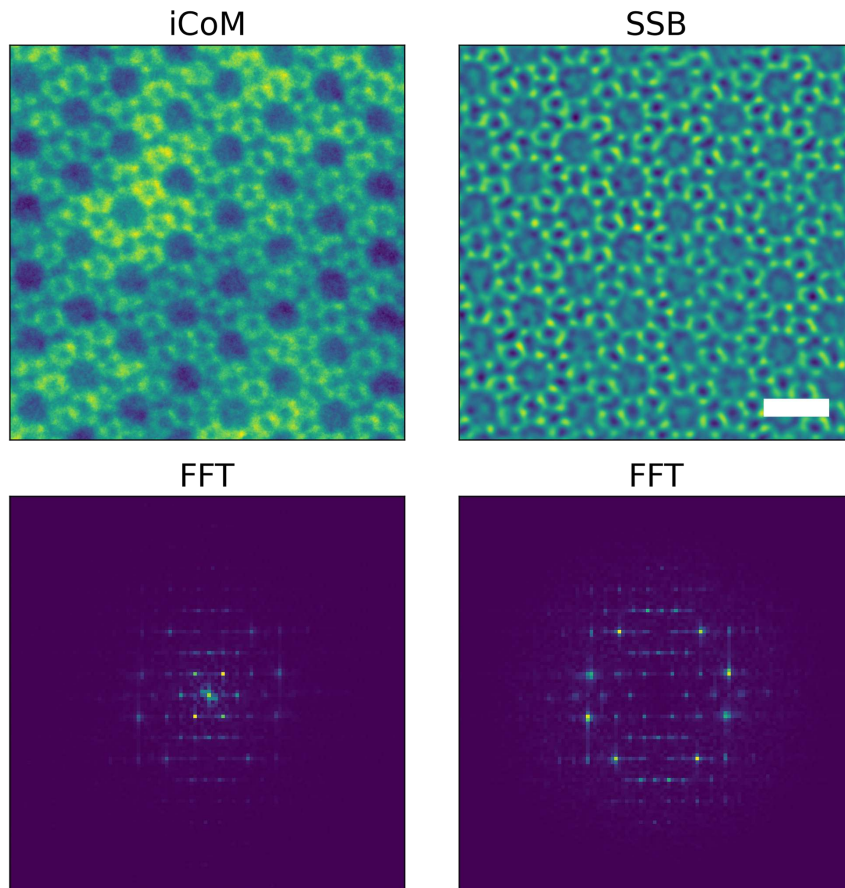


Figure 4.8: The reconstructed iCoM and SSB images of the silicalite-1 zeolite. The scale bar is 1 nm.

icients is necessary. In figure 4.9, I show an example showing an experimental pair of double disk overlaps with the artifacts, alongside the double disk overlaps showing just the double disk overlaps due to aberrations as calculated from the aberration coefficients identified by the SVD. By subtracting the phase due to the aberrations calculated from the SVD from the experimental double disk overlaps, a smoother phase is obtained as shown in the right side of the figure. Sometimes additional phase artifacts remain in experimental data due to for example charging on the aperture.

The final ptychographic phase image with the aberrations removed with the above described process with the SVD identification is shown in figure 4.10. In the phase image without aberration correction, distinguishing the two adjacent atoms of the WS_2 is challenging, despite the evident lattice structure in the figure, due to the presence of residual aberrations. Examination of the two FFTs below reveals weakened points on the third ring, as highlighted by the yellow circles. In contrast, in the post collection ptychographically aberration-corrected result, the two adjacent atoms are distinctly discernible, which can be seen in the inserted images cut down from the original images

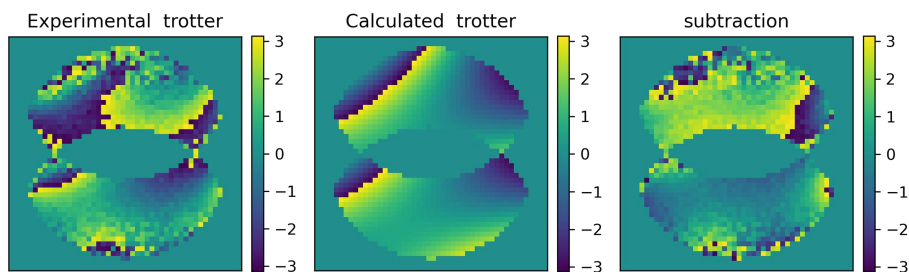


Figure 4.9: Aberration Removal Process: On the left is the experimental double disk overlap phase image containing residential aberrations. In the middle, the phases computed by the aberration coefficients calculated using the SVD method. On the right is the result obtained by subtracting the calculated phases due to aberrations from the experimental phases.

at the same palaces. The corresponding FFT displays enhanced clarity in the third ring of spots. These results affirm that, despite the presence of residual aberrations, effective post-processing enables their removal.

4.6 Conclusion

In this chapter, I have provided an overview of the fundamental principles of the Timepix3 detector and elucidated why this camera proves exceptionally well-suited for low-dose 4D-STEM experiments. Through experiments conducted on WS_2 and zeolite samples, we demonstrated the superior dose efficiency of the advanced imaging methods of ptychography, when compared to conventional BF, ABF, and ADF imaging methods in low-dose scenarios. Furthermore, employing aberration correction through SSB ptychography allows for the retrieval of high-quality phase images, even in the presence of residual aberrations. The results affirm the efficacy of the Timepix3 camera for beam-sensitive materials under low-dose conditions.

4.6. CONCLUSION

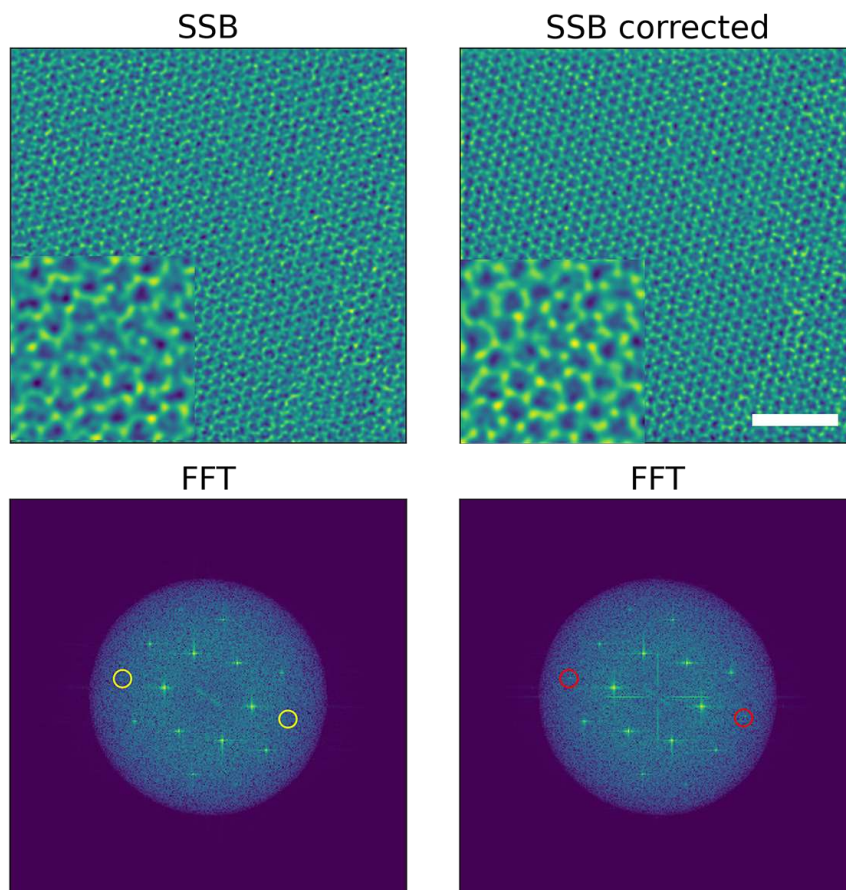


Figure 4.10: Comparison of normal SSB and aberration corrected SSB ptychography. The scale bar is 2 nm.

*CHAPTER 4. DOSE EFFICIENCY IN LOW DOSE CONDITION WITH TIMEPIX3
DETECTOR*

Overcoming contrast reversals in focused probe ptychography of thick materials

This chapter discusses how to overcome the contrast reversals that can occur for relatively thick samples.

In the projected potential approximation, as the sample increases in thickness the phase of the object increases. As there is only a 2π range of phases, a phase wrap will eventually occur as an increasing potential increases the phase response. Such a phase wrap can produce a contrast reversal. In practice most atomic resolution ptychography images do not contain a range of phases in the final image that actually spans 2π but only a much lower range of phase values. However the phases of the individual frequencies can exceed the 2π range.

We found that by tuning the defocus value to the middle plane of a sample, the contrast reversal can often be removed, as I will discuss here. Additionally, in this chapter, I will compare results from the Merlin camera and Timepix3 detector. The Timepix3 camera can provide much faster speed than the Merlin camera, which is very useful for avoiding drift and achieving low doses.

This chapter is based on the following paper:

C. Gao, C. Hofer, D. Jannis, A. Beche, J. Verbeeck, T. J. Pennycook. *Overcoming contrast reversals in focused probe ptychography of thick materials: An optimal pipeline for efficiently determining local atomic structure in materials science. Appl. Phys. Lett.* 121, 081906 (2022)

My contribution includes analysing the experimental results, as well as writing the codes needed for this and performing the simulations.

5.1 Introduction

The ability of STEM to determine structure and composition has made it essential to materials science. Typically this has been via Z-contrast ADF imaging [9, 30, 31], with simultaneous spectroscopies providing greater compositional sensitivity at the cost of far higher doses and drift [32, 33]. One particularly prominent deficiency of ADF imaging is the difficulty the modality faces in resolving light atoms near heavy elements. The precise locations of such light elements can dramatically alter the properties of materials containing them. Such is the case in many oxides where tiny changes in bond angles or lengths can completely change their magnetic or electronic properties [34–36]. Given sufficient stability of the stage and the sample, electron energy loss spectroscopy (EELS) can reveal the locations of such light elements [32, 33]. However, drift remains a challenge for precise measurements of atomic locations in EELS, and for many materials systems the damage from the doses required for spectroscopy preclude its use at atomic resolution. In many materials, the propensity to damage under the electron beam also precludes atomic resolution ADF imaging [37]. Alternative means of imaging with both greater sensitivity to light elements and higher overall dose efficiency have thus been sought to complement ADF imaging [38, 39].

Annular bright field (ABF) became popular because it can often determine the locations of light elements in heavy lattices at higher precision and at lower doses than spectroscopic elemental mapping [40–43]. Recently methods based on tracking the center of mass (CoM) of the electron scattering have begun to take over from ABF, as they are able to provide a higher signal to noise ratio. Integrated differential phase contrast (iDPC) [44] and integrated CoM (iCoM) [45, 46] have shown particular promise in this regard. iDPC can be performed using a set of conventional detectors arranged into quadrants or segments [44, 47], while iCoM is performed using four dimensional STEM (4D STEM) [11] data. Thus while iCoM utilizes a more accurate measure of the CoM, most cameras used for 4D STEM have made it significantly slower than iDPC. Now however developments in camera technology have made 4D STEM possible without any decrease in scan speed compared to even the most rapid ADF imaging [15]. This technological advance also accelerates the acquisition speed for highly efficient ptychography [21, 26, 48, 49]. Direct focused probe forms of ptychography are compatible with simultaneous ADF imaging, but their applicability to stronger objects such as thicker samples containing heavier elements have remained in question due to the use of the weak phase object (WPO) or multiplicative approximations in the theory underlying these ptychographic methods [21, 22, 28, 50, 51].

Here we demonstrate the significant advantages of combining focused probe ptychography with conventional rapid scan ADF workflows for the thicker samples typical of materials science. Both the single side band (SSB) and Wigner distribution deconvolution (WDD) methods are capable of providing significantly clearer images of thick structures at lower doses than ABF and iCoM. This is true well beyond the approximations used in the theoretical description of the methods. Although complex thickness induced contrast reversals can appear in the SSB and WDD phase images, as has previously been shown [50], we show how defocus adjustment can remove these effects and provide contrast reversal free images that can clearly show the positions of all the atomic columns in thick structures. The amount of defocus required can be within the range of defocus that is set naturally when optimizing ADF images manually in conventional

5.2. RESULTS

imaging, but it can also be applied and optimized after acquiring the data due to the ability of ptychography to perform post collection aberration correction.

In SSB and WDD ptychography the phases of diffracted convergent beam electron diffraction (CBED) disks are solved via their mutual interference with the direct beam in probe position reciprocal space. Experiments proceed by recording the intensity distribution of the scattering as a function of probe position with a camera. The Fourier transform of this 4D STEM data with respect to probe position yields, in the multiplicative approximation [28],

$$G(\mathbf{K}_f, \mathbf{Q}_p) = A(\mathbf{K}_f)A^*(\mathbf{K}_f + \mathbf{Q}_p) \otimes_{\mathbf{K}_f} \Psi_s(\mathbf{K}_f)\Psi_s^*(\mathbf{K}_f - \mathbf{Q}_p), \quad (5.1)$$

in which \mathbf{K}_f is the scattering vector leading to a location on the camera, \mathbf{Q}_p is spatial frequency, A is an aperture function expressing both the effect of the aperture size and the aberrations that appear within it, and Ψ_s represents the Fourier transform of the specimen transmission function. This represents the convolution between the region of mutual overlap of the shifted and unshifted aperture functions and the interference of the diffracted beams differing in frequency by \mathbf{Q}_p . In SSB ptychography the WPO approximation is invoked to further simplify this to [26]

$$\begin{aligned} G(\mathbf{K}_f, \mathbf{Q}_p) = & |A(\mathbf{K}_f)|^2\delta(\mathbf{Q}_p) \\ & + A(\mathbf{K}_f)A^*(\mathbf{K}_f + \mathbf{Q}_p)\Psi_s^*(-\mathbf{Q}_p) \\ & + A^*(\mathbf{K}_f)A(\mathbf{K}_f - \mathbf{Q}_p)\Psi_s(\mathbf{Q}_p), \end{aligned} \quad (5.2)$$

from which it can be understood how to obtain the phase and amplitude of each frequency of the specimen transmission function from the double disk overlap regions where the shifted and unshifted aperture functions coincide [21]. In WDD the WPO approximation is avoided and deconvolution is used to separate the aperture functions from the specimen transmission function [28]. However the principle remains essentially the same in that the phase and amplitude of each spatial frequency relative to the direct beam is determined from the double overlap regions.

5.2 Results

Although the details of CBED patterns are dependent on thickness, both the WDD and SSB can in practice be seen to provide clear images of the structures of materials containing heavy elements that are likely tens of nanometers thick. An experimental example using a Medipix3 camera [12] is shown with SrTiO₃ in Figure 5.1 comparing simultaneous ADF and 4D STEM based ABEF, iCoM and the focused probe SSB and WDD ptychography signals. The thickness is most likely on the order of tens of nanometers from the position averaged CBED [52] data (see supplementary information) which puts this sample well beyond both the WPO and multiplicative approximations. The conditions were optimized for the ADF image during the acquisition at 200 kV with a 20 mrad convergence angle. Distortions due to drift are apparent due to the very slow speed of the scan compared to conventional ADF imaging. The 2400 frames per second of the 6 bit mode of the Medipix3 camera used here is a typical speed for 4D STEM cameras over the past five years, and with a typical probe current of a few tens of picoamps results in a dose on the order of $10^6 e^-/\text{\AA}^2$. However it is clear that both SSB

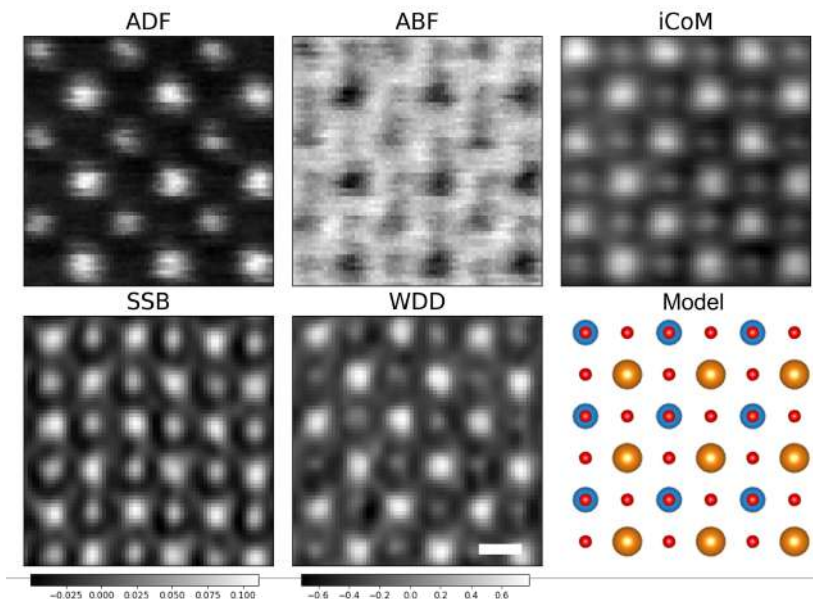


Figure 5.1: ADF and Medipix3 based 4D STEM data from SrTiO_3 . iCoM and ptychography show the O columns more clearly than the ABF. However this data shows some drift from the slow scan imposed by the camera. Sr, Ti and O columns are indicated by orange, blue and red respectively in the overlays. The scale bar shows 2 \AA and the colorbars are in radians.

and WDD provide atomic resolution images that, along with the iCoM, are far clearer than that of the ABF, despite the strength of the scattering here. A degree of filtering is apparent, but most of the clarity is attributed to the dose efficiency of the iCoM and ptychographic methods. Note that the ADF remains useful for clearly distinguishing the Sr and Ti columns, but the O atoms are not visible without the additional signals.

Frame based direct detection 4D STEM cameras such as the Medipix3 and its competitors are faster and more efficient than previous generation cameras, and are continuing to improve, but the susceptibility to instabilities such as drift and jitter resulting from their relatively low speeds remains a major problem for precision measurements. Furthermore the slow scan speeds they impose also make it more difficult to reach low doses. These two issues, instabilities and dose, have thus motivated continued use of methods based on conventional STEM detector setups such as ABF and iDPC which are in many ways inferior to 4D STEM methods but which have so far been more practical [21, 22, 25, 45].

Now we can remove the camera as a bottleneck to the speed at which 4D STEM can be performed. With event based detectors such as the Timepix3 [15], we can easily reach the equivalent of not just a few thousand frames per second, but millions of frames per second. We refer to the frames per second equivalent here in relation to the event driven camera only for the ease of comparison it can provide, as event driven operation does not utilize camera frames. Instead, each electron detected by the camera is labeled with a pixel location and time, and read out directly after the hit occurs rather than waiting for a full camera frame to be exposed and read out together. This provides a more efficient

5.2. RESULTS

readout that enables us to easily perform 4D STEM at the single to few microsecond dwell times used in rapid scan ADF imaging. Faster scans are also possible, such as the 100 ns dwell time we demonstrated previously [15], and may become increasingly desirable with the enhanced dose efficiency available with CoM and especially ptychography. Indeed we can also perform the same type of drift corrected multiple scan based imaging as has become common in rapid scan ADF imaging, allowing us to use data up until the point damage sets in. Thus with event driven cameras 4D STEM can now be acquired as rapidly as any other STEM modality.

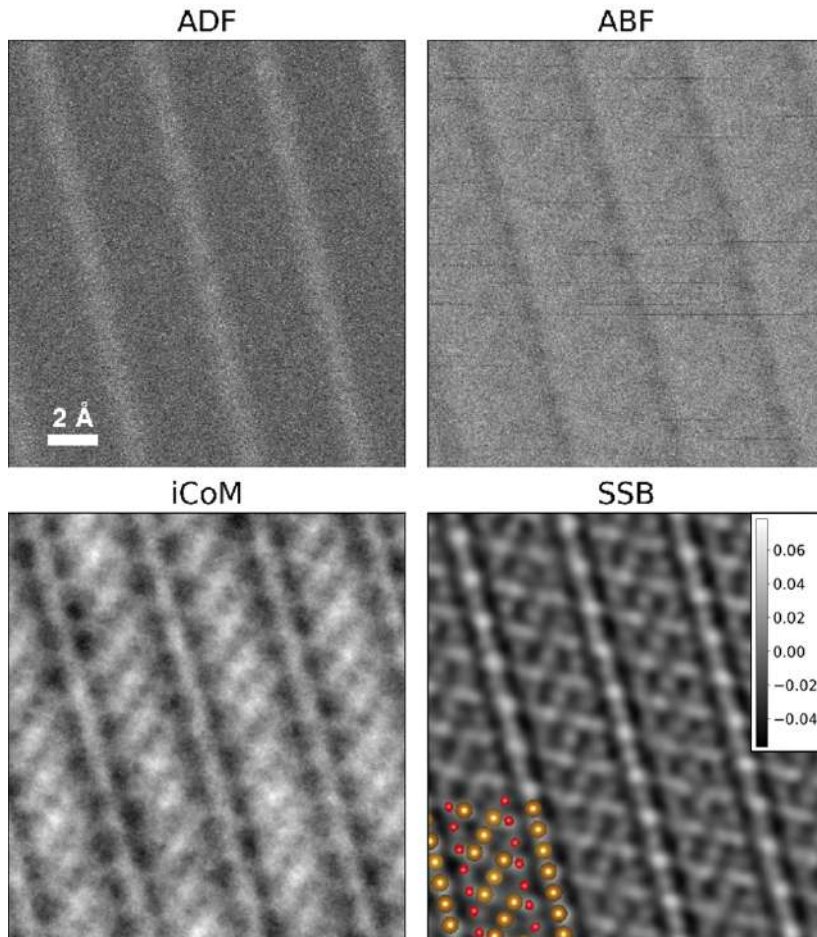


Figure 5.2: ADF and 4D STEM ABF, iCoM and SSB ptychography images of Fe_3O_4 from the same $2 \mu\text{s}$ scan enabled by a Timepix3 camera. Despite the low dose, the SSB image clearly shows the oxygen columns. The crystallinity and absence of distortions allowed the full image to be divided into 4 parts and aligned, based on the SSB signal, and averaged to produce the higher signal-to-noise ratio images shown. O and Fe atoms are shown in red and brown respectively in the model. The projection vector of the plane is $(2,3,1)$ and the colorbar for the SSB is in radians.

We illustrate the step change in the quality of images microsecond dwell time 4D STEM provides in Figure 5.2 with data acquired on a Timepix3 complementing ADF data

collected simultaneously from Fe_3O_4 , again using 200kV and a 20 mrad convergence angle. The electron dose was $10000 e^-/\text{\AA}^2$ and the dwell time was 2 microseconds. The ADF signal was used to optimize the imaging conditions, just as in a standard high end ADF STEM workflow. From both the Ronchigram and ADF imaging the sample is qualitatively of a thickness typical of 3D materials science samples for STEM. As seen in Figure 5.1 the clarity of the iCoM and focused probe ptychography far exceed that of the ABF image. However unlike Figure 5.1, this scan is essentially free of drift and other instabilities and at approximately two orders of magnitude lower dose. Because of the relatively low dose employed, the ADF is extremely noisy, as is the ABF signal which provides little advantage over the ADF. With the lower dose used here, the strength of the focused probe ptychography over that of iCoM also becomes much more apparent. Essentially every single atomic column is visible in the focused probe ptychography, but they are not in the iCoM.

Although the contrast transfer functions of SSB and WDD ptychography are single signed and provide easily interpretable images in which the contrast of the atoms is consistent against the background for thin weak specimens [22, 53], the contrast becomes more complex when handling thicker samples. Complex changes in contrast can appear as the projected potential of the sample becomes stronger as evidenced previously in GaN [50]. Atomic columns can be seen to change from bright to dark as the sample thickens, generally beginning in the center of the atomic columns where the projected potential is strongest. Thus one can encounter images in which the center of atomic columns have become dark while further out they remain bright. Such complex contrast reversals can not only impede the ease of interpretability but also degrade the dose efficiency by reducing the available contrast in low dose images.

An example of such contrast reversals is shown with simulations in Figure 5.3 (b) and (c). The 4D datasets are simulated using the multislice method with abTEM [54]. The voltage and convergence angle are the same as the experimental results shown in Figure 5.1. With the probe focused to the entrance surface of the 16 nm thick STO (zero defocus), both the SSB and WDD produce complex phase images in which the centers of both the Ti and Sr columns are dark but the light O columns remain bright. Surrounding the dark centers of the Ti and Sr columns are regions of higher phase values before these reduce to an intermediate phase value in the background. It is still possible to interpret which columns are where, particularly with the ADF signal present and indeed, the O columns already appear bright. Thus it is already feasible to locate the O atoms with this combination. However, these complex contrast reversals are clearly often undesirable, and by applying a small amount of defocus it is possible to remove them as illustrated in Figure 5.3 for SSB (e) and WDD (f) with a defocus of 8 nm, which has a minimal effect on the ADF image. With this defocus all the atomic columns appear bright on a dark background in the ptychographic images. Note that we define positive defocus as underfocus, moving the focal point into the sample from the entrance surface, and negative values as overfocus, moving the focal point before the sample. Focal series for thicknesses of 16, 20, 28 and 50 nm are shown in the supplementary materials. The optimal defocus appears to be close to the middle of the sample for the thinner examples, but for the 50 nm case defoci corresponding to a few nm before the sample or near the exit surface appear preferable.

In Figure 5.3 we also show the effect of adding 10 nm of defocus post collection to the SSB (g) and WDD (h) for the 16 nm thick STO simulation. 10 nm was chosen as it

5.2. RESULTS

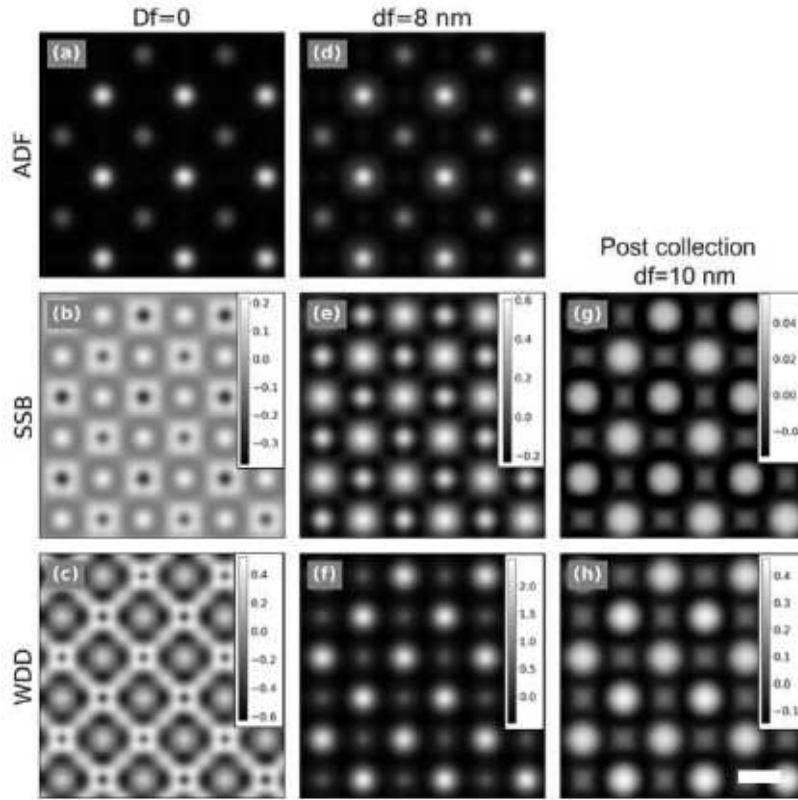


Figure 5.3: ADF, SSB and WDD images simulated for 16 nm thick STO. At zero defocus complex contrast reversals appear in the SSB and WDD images, but when the probe is focused 8 nm into the material the contrast reversals are removed. Importantly however the contrast reversals can also be removed by inserting defocus post collection, meaning the experiment can proceed with the focus optimized for the ADF. The scale bar indicates 2 Å and the colorbar is in radians.

optimizes the removal of the contrast reversals. This also works for considerably thicker materials as we show in the supplemental materials with a focal series for the 50 nm thick STO simulation. Using post collection SSB defocus adjustment we are able to remove the complex contrast reversals with the probe focused to the entrance surface during the simulated scan (with the CBED patterns simulated with defocus equal to zero). Although not identical to performing the defocus adjustment on the probe directly, this illustrates that post acquisition adjustment can also be used to remove the contrast reversals. The difference between applying the defocus during and after acquisition could be explained by channeling effects which are not accounted for with the post acquisition adjustment of the defocus.

The availability of the simultaneous ADF signal is thus of two fold importance. Its atomic number contrast provides greater compositional discernment without resorting to the greater doses required for atomic resolution spectrum imaging, while providing a reference image which is not susceptible to contrast reversals. Atoms are always bright

on a dark background in ADF images. This advantage is not available in defocused probe methods of ptychography because the ADF image is blurred beyond usability.

Although elucidating the physics of the removal of contrast reversals via defocus in detail is reserved for future studies, we believe the effect can be understood as rolling the phase around across the field of view such that the centers of the atomic columns are maximum and the background at a minimum. This works in practice because the ptychographic phase images generally only use a small fraction of the 2π radian range of values available for phase value images, and thus the phase rolling does not bring other features into non intuitive ranges of contrast.

In our experimental examples, the ptychographic data did not need additional defocus to produce clear images. This could be because they were obtained at a focus that produced clear ADF and ptychographic images. The defocus applied to provide contrast reversal free ptychographic images in Figure 5.3 only slightly reduced the quality of the ADF image. However when the defocus is not optimal for ptychography, it can be adjusted after taking the data. The injection of defocus and post collection aberration correction has been demonstrated previously [28, 53] with focused probe ptychography, but here provides an unanticipated benefit. Post collection aberration correction can also be combined with such defocus optimization.

In comparison to direct ptychography, iterative ptychography is not guaranteed to converge to the correct result or indeed to converge at all, but they are guaranteed to take a longer time to produce a useful image. We timed serial SSB and extended ptychographic iterative engine (ePIE) [24] calculations from the same focused probe dataset using an i7-10700K processor. The entire SSB calculation took 2.4 s regardless of dose. The ePIE calculation took 3.5 s (1 α limit) to 24.4 s (3 α limit) per iteration so that 10 ePIE iterations takes between 15 and 100 times as long as the SSB result. Convergence depends on the dose, and at $1.2 \times 10^5 e^-/\text{\AA}^2$ ePIE did not converge whereas the SSB produces clear images. Parallelism can speed up both algorithms, but this shows the far higher computational costs of iterative vs direct methods. See the supplementary information for further details. Live iCoM [55] and SSB [56] have already been reported, but live imaging with iterative methods would be at least far more difficult. Detailed information about simulation methods and the related parameters can be found in the supplementary information.

5.3 Conclusion

In conclusion, we have demonstrated the utility of performing focused probe ptychography within existing rapid scan ADF workflows for general materials science samples. Simultaneous ADF and ptychography provide a robust and interpretable method of efficiently imaging all the atomic columns in such samples. In particular, when combined with the ability to apply corrective defocus, the method works well beyond the approximations which have been used to motivate the SSB and WDD methods. The clarity of the ptychographic images, provides the best possible combination of the general efficiency of structural imaging, visibility of light elements and interpretability when combined with simultaneous Z-contrast ADF imaging. With a detector such as the Timepix3, drift is no longer a significant issue with 4D STEM and there is indeed very little reason

5.4. SUPPLEMENTARY INFORMATION

not to perform simultaneous ADF and 4D STEM imaging. We therefore anticipate the widespread adoption of this imaging pipeline in the future.

5.4 Supplementary information

Additional simulation and reconstruction details

The abTEM simulations used the Kirkland potential parameterization [57]. The convergence angle, probe position spacing, and rotation angles used in the ptychographic reconstructions matched the data input to them. No additional parameters are needed for the SSB algorithm. For WDD we additionally used an epsilon value of 0.01 and for ePIE values of $\alpha = 0.1$ and $\beta = 1$ were used.

The SSB and WDD reconstructions were performed using the pyPtychoSTEM code which we provide on gitlab at <https://gitlab.com/pyptychostem/pyptychostem>.

For ePIE we used the code available from Cornell at https://github.com/paradimdata/Cornell_EM_SummerSchool_2021/tree/main/Tutorial%206%20-%20Ptychography.

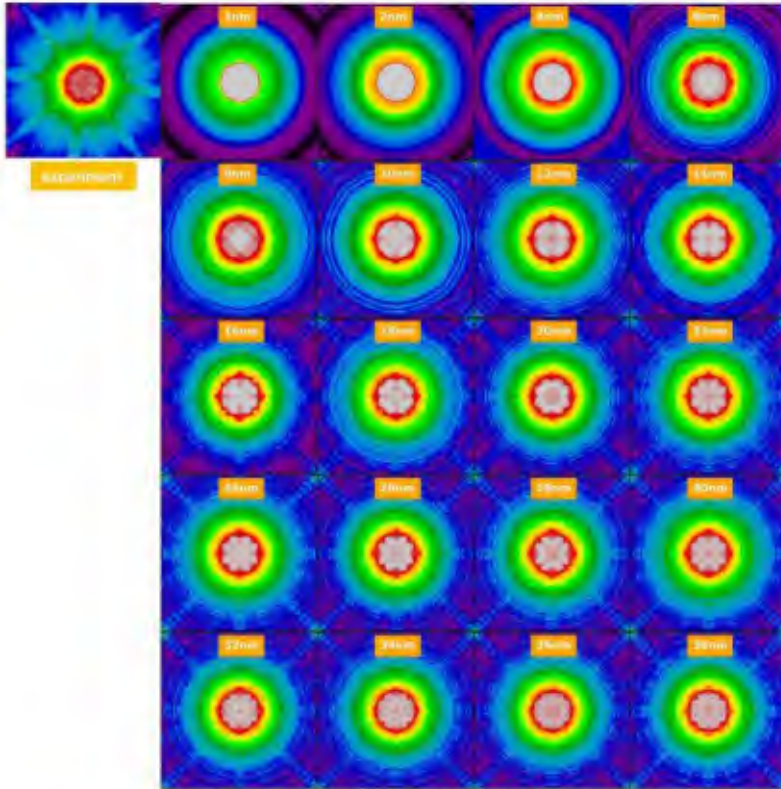


Figure 5.4: Comparison of the position averaged convergent beam electron diffraction (PACBED) [52] from the experimental data used for Figure 5.1 of the main text with PACBED simulated for different STO sample thicknesses using the same 200 kV accelerating voltage and 20 mrad convergence angle used in the experiment. While no single thickness provides a perfect match, it is clear the experimental PACBED is consistent with the sample tens of nanometers thick. Looking at the symmetries present, we estimate the experiment was performed in a region that is at least approximately 20 nm thick. The scale bar is 2 Å.

5.4. SUPPLEMENTARY INFORMATION

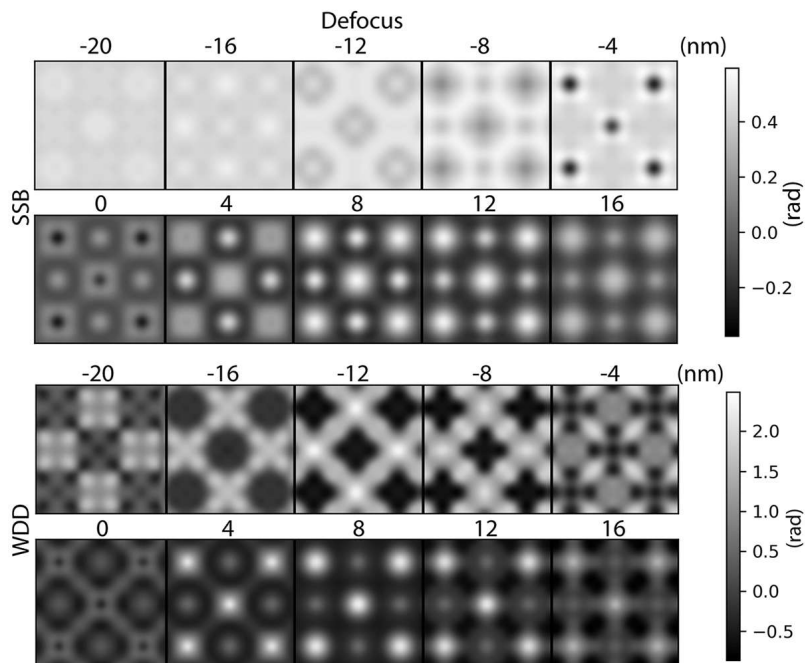


Figure 5.5: Simulated SSB and WDD reconstructed phase results for 16 nm thick STO with varying probe defocus.

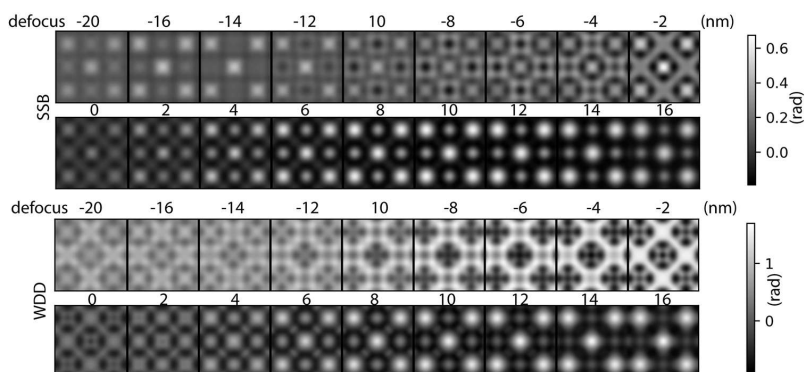


Figure 5.6: Simulated SSB and WDD reconstructed phase results for 20 nm thick STO with varying probe defocus.

CHAPTER 5. OVERCOMING CONTRAST REVERSALS IN FOCUSED PROBE
PTYCHOGRAPHY OF THICK MATERIALS

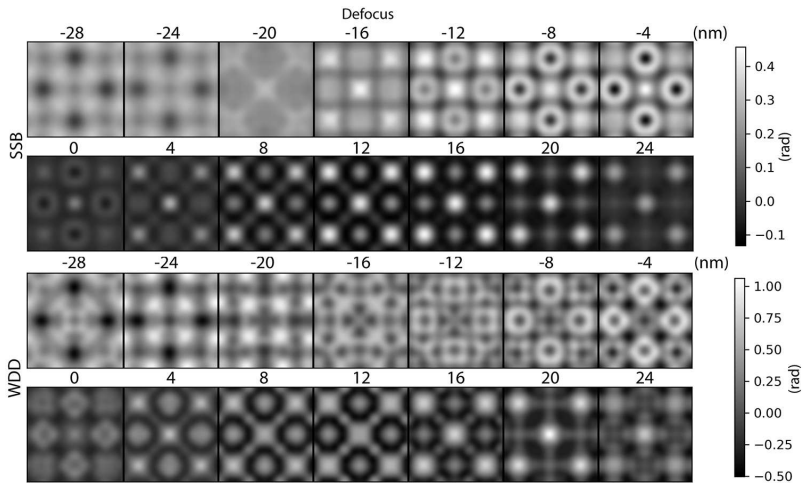


Figure 5.7: Simulated SSB and WDD reconstructed phase results for 28 nm thick STO with varying probe defocus.

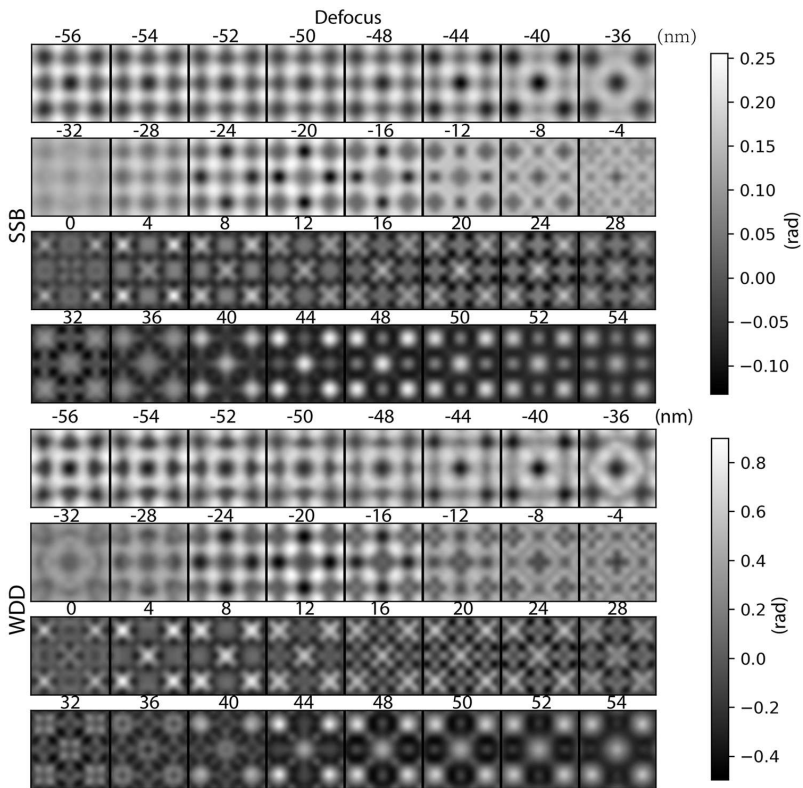


Figure 5.8: Simulated SSB and WDD reconstructed phase results for 50 nm thick STO with varying probe defocus.

5.4. SUPPLEMENTARY INFORMATION

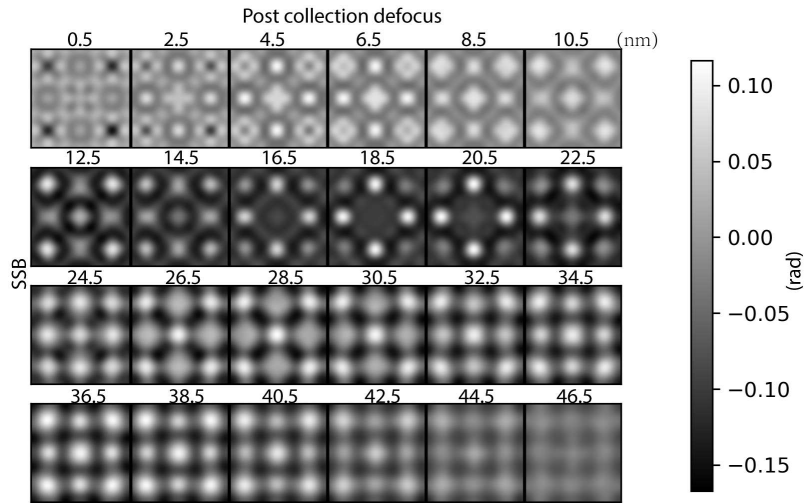


Figure 5.9: Simulated post collection SSB focal series for 50 nm thick STO, demonstrating that the contrast reversals can also be removed for this thickness when the probe is focused to the entrance surface during the scan with post collection defocus adjustment.

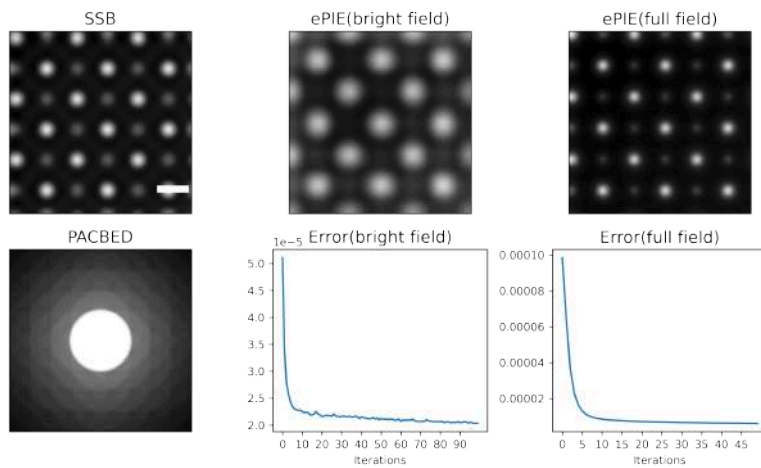


Figure 5.10: ePIE and SSB results computed using a single core of an i7-10700K operating at 3.8 GHz using data simulated for a single unit cell of STO at 200 KV, with a 20 mrad convergence angle, zero defocus and a 59 by 59 probe position grid. The 47 by 47 pixels of the bright field disk are used for bright field ePIE, while the 137 by 137 pixels of the full field CBED patterns are used for the full field ePIE. The SSB image was computed in a total time of 2.4 s. 100 iterations were used in computing the bright field ePIE result taking a total time of 347 s, for an average of 3.47 seconds per iteration. The full field ePIE result and the error per iteration are also provided. The PACBED of the STO dataset is also shown with a log scale.

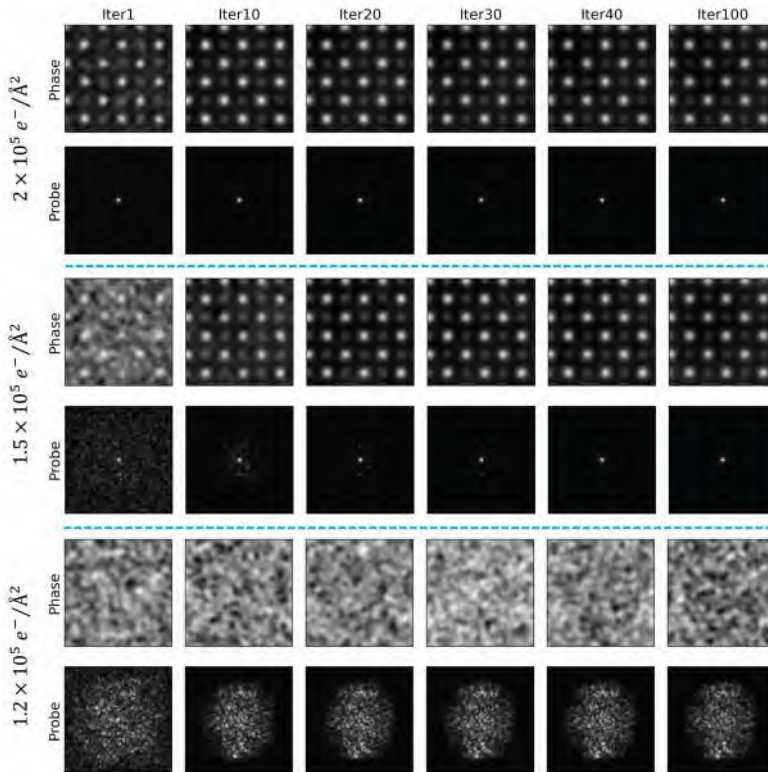


Figure 5.11: Full field single unit cell thick STO ePIE reconstruction as a function of iteration number at different doses. At $2 \times 10^5 e^-/\text{\AA}^2$, after 10 to 20 iterations very little apparent change occurs in the phase and probe. At $1.5 \times 10^5 e^-/\text{\AA}^2$ some artefacts are visible at 20 iterations, but disappear between iterations 30 to 40. At $1.2 \times 10^5 e^-/\text{\AA}^2$ the algorithm does not converge even after 100 iterations, reflecting the fact that the convergence of ePIE depends critically on the amount of signal in the CBED patterns. We note that using a significantly defocused probe increases the signal per CBED pattern, but a strongly defocused probe precludes useful simultaneous Z-contrast ADF imaging.

5.4. SUPPLEMENTARY INFORMATION

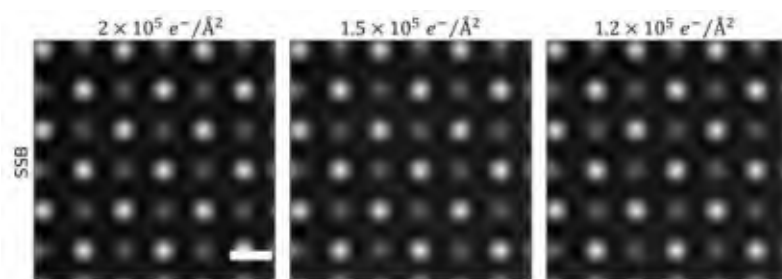


Figure 5.12: SSB reconstructions at the same doses as the ePIE results shown in Figure S8. As a direct non-iterative method, SSB ptychography does not depend on the dose to converge. Of course, the clarity of the resulting SSB images do depend on the dose, but it is a consistent and immediate result, and the above doses pose no problem at all for the SSB method. Scale bar is 2 Å.

*CHAPTER 5. OVERCOMING CONTRAST REVERSALS IN FOCUSED PROBE
PTYCHOGRAPHY OF THICK MATERIALS*

On central focusing for contrast optimization in direct electron ptychography of thick samples

Ptychography provides high dose efficiency images that can reveal light elements next to heavy atoms. However, despite ptychography having an otherwise single signed contrast transfer function, contrast reversals can occur when the projected potential becomes strong for both direct and iterative inversion ptychography methods. It has recently been shown that these reversals can often be counteracted in direct ptychography methods by adapting the focus. We explore the use of an integrated contrast transfer function concept to explain this phenomenon, but find it can not work for the strong phase object. In addition we show that the convergence angle can be an important consideration for removal of contrast reversals in relatively thin samples.

This chapter is based on the following paper:

C. Gao, C. Hofer, T.J. Pennycook, On central focusing for contrast optimization in direct electron ptychography of thick samples. Ultramicroscopy 256 (2024) 113879

My contribution is proposing the iCTF concept, performing the simulations and analysing the data.

6.1 Introduction

As introduced in chapter 5, contrast reversal in electron ptychography is a phenomenon where the reconstructed image displays an inverted contrast relative to the actual sample. For instance, regions that should appear bright (high projected potential) appear dark, and regions that should be dark (lower projected potential) appear bright. Contrast reversals can appear with both direct and iterative methods, often with the columns taking on what can be described as a donut or caldera shape [50, 58, 59]. This can occur even though the range of phase values in the final phase images is significantly smaller than the 2π range available to pure phase images. Such contrast reversals have also recently been observed in integrated Center of Mass (iCoM) images [59]. Recent investigations have revealed that for direct ptychography, as in iCoM [45], optimal contrast often occurs with the probe focused to the center of the sample for quite a range of thicknesses [58, 59].

In chapter 2, we know the Contrast Transfer Function (CTF) describes how different spatial frequencies are transferred from the sample to the image by the electron microscope. In practice, the CTF is affected by several factors such as: (1) Defocus: The distance between the focal plane and the sample. (2) Spherical Aberration: Distortion caused by the spherical shape of lenses. Astigmatism: Imperfections in the lens causing different focal points in different planes. These factors cause artifacts in the CTF, leading to regions where certain spatial frequencies are not transferred correctly,

The integrated Contrast Transfer Function (iCTF) is a concept that builds upon the traditional CTF by integrating the effects of various imaging conditions over multiple images or defocus values. iCTF has been successfully used to explain the several imaging characteristics of ABF and DPC imaging [60, 61]. This approach aims to provide a more comprehensive understanding of how contrast is transferred in an electron microscope.

Here we borrow this iCTF idea and try to explain how the equal and opposite phase contributions of the defocus at sample slices above and below the central slice effectively cancels out when the probe is focused to the central slice of the sample. This explains why focusing to the middle of a sample often provides the best contrast, at least in samples in which strong dynamical scattering does not overly unbalance the contributions from above and below the central plane. We also show how the convergence angle can be an important consideration for avoiding contrast reversals, particularly in thinner samples with significant projected potentials.

6.2 Theory

In this document, I would like to demonstrate how to derive the integrated contrast transfer function concept with the model which is called thick weak phase approximation [60, 61].

Different with the the phase object approximation, the thick weak object approximation model can be demonstrated in the figure 6.1, which is quoted from the paper [61]. Thick weak object approximation has the assumptions that for the crystalline with the thickness of t , it can be sliced into n layers with the interval of Δt as the conventional multislice

6.2. THEORY

method. Each layer can be treated as the weak phase object. One thing that different with the multislice method is we just consider about single scattering happens in each layer. This means for the incident beam in each layer, we ignore the multiply scattering beams. Therefore with this assumption, the incident beams for each layer is the probe freely propagates to that layer.

6.2.1 Conventional phase object approximation theory.

With the concept of phase object function, assume the object transmission function is $O(\mathbf{r})$, the probe function is $P(\mathbf{r} - \mathbf{r}_p)$, where \mathbf{r}_p denotes the probe scan position. Then the exit wave function $\Psi(\mathbf{r})$ is :

$$\Psi(\mathbf{r}) = P(\mathbf{r} - \mathbf{r}_p) \cdot O(\mathbf{r}) \quad (6.1)$$

For the probe function \mathbf{r}_p in real space, when considering the influence of the defocus Δf of the probe. it can be expressed as the Fourier transform of the aperture function $A(\mathbf{K}_f) = a(\mathbf{K}_f) \exp[-i\chi(\mathbf{K}_f)]$ describing both the extent of the aperture via the top hat function $a(\mathbf{K}_f)$ and any lens aberrations present via the aberration function $\chi(\mathbf{K}_f)$. if we just consider the influence of defocus Δf in $\chi(\mathbf{K}_f)$, we can write the new aperture function as $A(\mathbf{K}_f, \Delta f) = a(\mathbf{K}_f) \exp[-i\chi(\mathbf{K}_f)]$:

$$P(\mathbf{r} - \mathbf{r}_p) = \mathcal{F}^{-1}(A(\mathbf{K}_f, \Delta f) \cdot \exp(-i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)) \quad (6.2)$$

In the diffraction plane, the wave function $M(\mathbf{K}_f, \mathbf{R}_p)$ can be expressed as the Fourier transform of the exit wave function.

$$M(\mathbf{K}_f, \mathbf{r}_p) = \mathcal{F}(\Psi(\mathbf{r})) = \int \Psi(\mathbf{r}) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}) d\mathbf{r}, \quad (6.3)$$

and the intensity of the diffraction pattern can be expressed as

$$I(\mathbf{K}_f, \mathbf{r}_p) = |M(\mathbf{K}_f, \mathbf{r}_p)|^2 = |\mathcal{F}(P(\mathbf{r} - \mathbf{r}_p) \cdot O(\mathbf{r}))|^2 = |(A(\mathbf{K}_f) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)) \otimes O(\mathbf{K}_f)|^2. \quad (6.4)$$

6.2.2 Thick weak phase object model

With the thick weak phase object model, the object function in the j th slice can be denoted as $O_j(\mathbf{r})$, the probe function in this slice is written as $P_j(\mathbf{r} - \mathbf{r}_p, z_j)$, where z_j is the defocus value at the slice j . Therefore, the exit wave function can be written as :

$$\Psi_j(\mathbf{r} - \mathbf{r}_p) = O_j(\mathbf{r}) \cdot P_j(\mathbf{r} - \mathbf{r}_p, z_j) \quad (6.5)$$

For the $j + 1$ th layer, with the multislice theorem, the incident beam is the exit wave function $\Psi_j(\mathbf{r})$ at the slice j propagates to this layer with a propagation function. While here we treat the incident beam as $P_{j+1}(\mathbf{r} - \mathbf{r}_p, z_{j+1})$, where we ignore the scattered beams in the slice j for the incident beam.

Then for the projected potential in slice j , with the weak phase object approximation, the object function $O_j(\mathbf{r})$ can be written as:

$$O_j(\mathbf{r}) = 1 - i\delta V_j(\mathbf{r}), \quad (6.6)$$

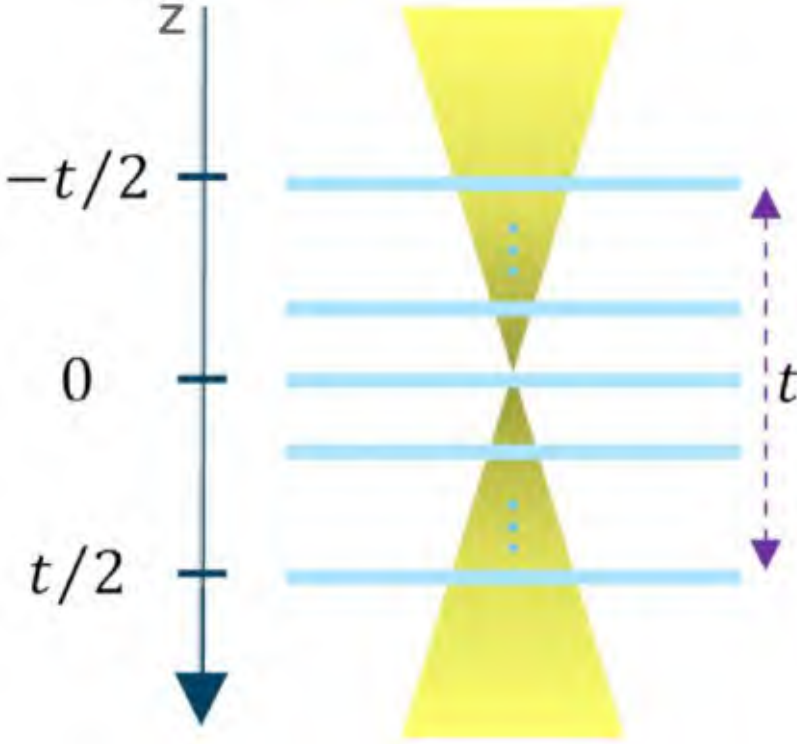


Figure 6.1: Schematic defining the z -coordinates.

where δ is the interaction parameter and $V_j(\mathbf{r})$.

Then the transmission function $\Psi_j(\mathbf{r} - \mathbf{r}_p)$ is:

$$\Psi_j(\mathbf{r} - \mathbf{r}_p) = P_j(\mathbf{r} - \mathbf{r}_p, z_j) - i\delta V_j(\mathbf{r}) \cdot P_j(\mathbf{r} - \mathbf{r}_p, z_j) \quad (6.7)$$

For the incident beam to the $j + 1$ th slice,

$$P_{j+1}(\mathbf{r} - \mathbf{r}_p) = \hat{p}(\mathbf{r}, \Delta z) \otimes \Psi_j(\mathbf{r} - \mathbf{r}_p) \quad (6.8)$$

where $\hat{p}(\mathbf{r}, \Delta z) = \frac{1}{i\lambda\Delta z} \exp\left(\frac{i\pi\mathbf{r}^2}{\lambda\Delta z}\right)$ is the Fresnel propagator. \otimes denotes convolution. Fourier transform to $\hat{p}(\mathbf{r}, \Delta z)$, $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_t}(\hat{p}(\mathbf{r}, \Delta z)) = \exp(-i\pi\lambda\mathbf{K}_f^2\Delta z)$.

With the thick weak phase object approximation, we have the assumptions that the incident beam for the j th layer is the probe freely propagating to this slice, not influenced by the multiscattering process and the scattered beams at the j th layer only scattered once. Then the exit wave function $\Psi_n^{out}(\mathbf{r} - \mathbf{r}_p)$ at the final layer can be expressed as:

$$\Psi_n^{out}(\mathbf{r} - \mathbf{r}_p) = P(\mathbf{r} - \mathbf{r}_p, z_n) + \sum_{j=1}^n \hat{p}^{n-j}(\mathbf{r}) \otimes (v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)) \quad (6.9)$$

Here the notation \hat{p} is a little different with the derivation in the paper [61], in paper [61], it defines \hat{p} as an operator which contains convolution operation. Here to make it

6.2. THEORY

simpler, I put the convolution in the equation 6.9. I also combine $i\delta V_j(\mathbf{r})$ into $v_j(\mathbf{r})$. In this equation, the first term describes the freely propagating probe. \mathbf{r}_p is the scan probe position, z_n denotes the defocus value of the probe at the slice n . With the equation 6.2,

$$P(\mathbf{r} - \mathbf{r}_p, z_n) = \mathcal{F}_{\mathbf{K}_f \rightarrow \mathbf{r}}^{-1}[A(\mathbf{K}_f, z_n) \cdot \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)] \quad (6.10)$$

where $A(\mathbf{K}_f, z_n) = a(\mathbf{K}_f) \exp(-i\pi\lambda\mathbf{K}_f^2 z_n)$.

For the 2nd term, $a_j = v(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)$ is the scattered beam from the slice j . $\hat{p}^{n-j}(\mathbf{r}) \otimes$ means the scattered beam a_j at the slice j freely propagates to the slice n . $\sum_{j=1}^n$ represents we sum up all the scattered beams propagating to the slice n .

In the diffraction plane, the wave function $M(\mathbf{K}_f, \mathbf{r}_p)$ is the Fourier transform of $\Psi_n^{out}(\mathbf{r} - \mathbf{r}_p)$ with respect to \mathbf{r} .

$$\begin{aligned} M(\mathbf{K}_f, \mathbf{r}_p) &= \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(\Psi_n^{out}(\mathbf{r} - \mathbf{r}_p)) \\ &= \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(P(\mathbf{r} - \mathbf{r}_p, z_n)) + \sum_{j=1}^n \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[\hat{p}^{n-j}(\mathbf{r}) \otimes (v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j))] \end{aligned} \quad (6.11)$$

For the first terms $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(P(\mathbf{r} - \mathbf{r}_p, z_n))$, it is the inverse process of equation 6.10 and can be written as $A(\mathbf{K}_f, z_n) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)$. It can be interpreted as the transmitted convergent beam in the diffraction plane.

For the second term,

$$\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[\hat{p}^{n-j}(\mathbf{r}) \otimes (v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j))] = \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(\hat{p}^{n-j}(\mathbf{r})) \cdot \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)] \quad (6.12)$$

In this equation $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(\hat{p}^{n-j}(\mathbf{r}))$ denotes how the beam propagates from layer j to layer n and can be obtained by $\exp(-i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z)$. $(n-j)\Delta z$ means the scattered beam at the slice j propagates to layer n with a distance of $(n-j)\Delta z$. For $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)]$, we just transform the scattered beam at the slice j from \mathbf{r} space to \mathbf{K}_f space.

$$\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)] = \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(v_j(\mathbf{r})) \otimes \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(P(\mathbf{r} - \mathbf{r}_p, z_j)) \quad (6.13)$$

where $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(v_j(\mathbf{r}))$ can be written as $v_j(\mathbf{K}_f)$, the weak phase function in the reciprocal space. $\mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}(P(\mathbf{r} - \mathbf{r}_p, z_j))$ is the inverse process of equation 6.10 therefore can be expressed as $A(\mathbf{K}_f, z_j) = a(\mathbf{K}_f) \exp(-i\pi\lambda\mathbf{K}_f^2 z_j) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)$. Next we can express the equation 6.13 as:

$$\begin{aligned} \mathcal{F}_{\mathbf{r} \rightarrow \mathbf{K}_f}[v_j(\mathbf{r}) \cdot P(\mathbf{r} - \mathbf{r}_p, z_j)] &= [a(\mathbf{K}_f) \exp(-i\pi\lambda\mathbf{K}_f^2 z_j) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)] \otimes v_j(\mathbf{K}_f) \\ &= \int d\mathbf{K}_{f1} a(\mathbf{K}_{f1}) \exp(-i\pi\lambda\mathbf{K}_{f1}^2 z_j) \exp(i2\pi\mathbf{K}_{f1} \cdot \mathbf{r}_p) v_j(\mathbf{K}_f - \mathbf{K}_{f1}) \end{aligned} \quad (6.14)$$

where \mathbf{K}_{f1} is dummy variable.

Then for the scattered beam at the slice j to propagate to the slice n , it can be expressed as:

$$\exp(-i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z) \cdot \int d\mathbf{K}_{f1} a(\mathbf{K}_{f1}) \exp(-i\pi\lambda\mathbf{K}_{f1}^2 z_j) \exp(i2\pi\mathbf{K}_{f1} \cdot \mathbf{r}_p) v_j(\mathbf{K}_f - \mathbf{K}_{f1}) \quad (6.15)$$

Then the diffraction beam $M(\mathbf{K}_f, \mathbf{r}_p) = a + b$, $a = A(\mathbf{K}_f, z_n) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p)$, the transmitted beam. and b is the summation part of the diffracted beams as shown in equation 6.11.

The intensity of the diffraction pattern $I(\mathbf{K}_f, \mathbf{r}_p) = |M(\mathbf{K}_f, \mathbf{r}_p)|^2$, therefore:

$$I(\mathbf{K}_f, \mathbf{r}_p) = M(\mathbf{K}_f, \mathbf{r}_p) \cdot M^*(\mathbf{K}_f, \mathbf{r}_p) = aa^* + ab^* + a^*b + bb^* \quad (6.16)$$

Owing to the bb^* being the multiplication of the two weak phase functions, we ignore this part and then

$$I(\mathbf{K}_f, \mathbf{r}_p) = M(\mathbf{K}_f, \mathbf{r}_p) \cdot M^*(\mathbf{K}_f, \mathbf{r}_p) = aa^* + ab^* + a^*b \quad (6.17)$$

Now aa^* can be expressed as:

$$aa^* = A(\mathbf{K}_f, z_n)A^*(\mathbf{K}_f, z_n) = a(\mathbf{K}_f)^2 \quad (6.18)$$

where $a(\mathbf{K}_f)$ is the top hat function.

For the second term ab^* , it can be expressed as:

$$ab^* = A(\mathbf{K}_f, z_n) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p) \cdot \left(\sum_{j=1}^n \exp(i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z) \right. \\ \left. \cdot \int d\mathbf{K}_{f1} a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j) \exp(-i2\pi\mathbf{K}_{f1} \cdot \mathbf{r}_p) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \right) \quad (6.19)$$

for the 3rd term, a^*b :

$$a^*b = A^*(\mathbf{K}_f, z_n) \exp(-i2\pi\mathbf{K}_f \cdot \mathbf{r}_p) \cdot \left(\sum_{j=1}^n \exp(-i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z) \right. \\ \left. \cdot \int d\mathbf{K}_{f2} a(\mathbf{K}_{f2}) \exp(-i\pi\lambda\mathbf{K}_{f2}^2 z_j) \exp(i2\pi\mathbf{K}_{f2} \cdot \mathbf{r}_p) v_j(\mathbf{K}_f - \mathbf{K}_{f2}) \right) \quad (6.20)$$

As the procedure in single side band ptychography, we take the Fourier transform of the diffraction pattern $I(\mathbf{K}_f, \mathbf{r}_p) = |M(\mathbf{K}_f, \mathbf{r}_p)|^2$ with respect to the scan probe position \mathbf{r}_p and the $G(\mathbf{K}_f, \mathbf{Q}_p)$ can be expressed as:

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(I(\mathbf{K}_f, \mathbf{r}_p)) \quad (6.21)$$

We use the approximation as shown in equation 6.17.

$$G(\mathbf{K}_f, \mathbf{Q}_p) = \mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(aa^* + ab^* + a^*b) \\ = \mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(aa^*) + \mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(ab^*) + \mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(a^*b) \quad (6.22)$$

For the first part $\mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(aa^*)$, as shown in equation 6.18, $aa^* = a(\mathbf{K}_f)^2$, then

$$\mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(aa^*) = a(\mathbf{K}_f)^2 \delta(\mathbf{Q}_p) \quad (6.23)$$

6.2. THEORY

For the 2nd part, $\mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p}(ab^*)$, because $\sum_{j=1}^n$ is a linear calculation process, to simply the equation, here I just pick the j th scattered beam as an example:

$$\mathcal{F}_{\mathbf{r}_p \rightarrow \mathbf{Q}_p} \left(A(\mathbf{K}_f, z_n) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p) \cdot \left(\exp(i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z) \cdot \int d\mathbf{K}_{f1} a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j) \exp(-i2\pi\mathbf{K}_{f1} \cdot \mathbf{r}_p) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \right) \right) \quad (6.24)$$

For this equation, $A(\mathbf{K}_f, z_n) \exp(i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z)$ is $a(\mathbf{K}_f) \exp(-i\pi\lambda\mathbf{K}_f^2 z_n)$
 $\exp(i\pi\lambda\mathbf{K}_f^2(n-j)\Delta z) = a(\mathbf{K}_f) \exp(-i\pi\lambda\mathbf{K}_f^2 z_j) = A(\mathbf{K}_f, z_j)$. It can be expressed as:

$$\int d\mathbf{r}_p \exp(i2\pi\mathbf{r}_p \cdot \mathbf{Q}_p) \left\{ A(\mathbf{K}_f, z_j) \exp(i2\pi\mathbf{K}_f \cdot \mathbf{r}_p) \cdot \int d\mathbf{K}_{f1} a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j) \exp(-i2\pi\mathbf{K}_{f1} \cdot \mathbf{r}_p) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \right\} \quad (6.25)$$

Then we put the terms with \mathbf{r}_p together and the new formula can be expressed as:

$$\int d\mathbf{K}_{f1} A(\mathbf{K}_f, z_j) a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \int d\mathbf{r}_p \exp(i2\pi(\mathbf{K}_f + \mathbf{Q}_p - \mathbf{K}_{f1}) \cdot \mathbf{r}_p) \quad (6.26)$$

where $\int d\mathbf{r}_p \exp(i2\pi(\mathbf{K}_f + \mathbf{Q}_p - \mathbf{K}_{f1}) \cdot \mathbf{r}_p) = \delta(\mathbf{K}_f + \mathbf{Q}_p - \mathbf{K}_{f1})$. then

$$\int d\mathbf{K}_{f1} A(\mathbf{K}_f, z_j) a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \cdot \delta(\mathbf{K}_f + \mathbf{Q}_p - \mathbf{K}_{f1}) \quad (6.27)$$

And for $a(\mathbf{K}_{f1}) \exp(i\pi\lambda\mathbf{K}_{f1}^2 z_j)$, it equals to $A^*(\mathbf{K}_{f1}, z_j)$ and we can write the formula as:

$$\int d\mathbf{K}_{f1} A(\mathbf{K}_f, z_j) A^*(\mathbf{K}_{f1}, z_j) v_j^*(\mathbf{K}_f - \mathbf{K}_{f1}) \cdot \delta(\mathbf{K}_f + \mathbf{Q}_p - \mathbf{K}_{f1}) \quad (6.28)$$

Use $\mathbf{K}_{f1} = \mathbf{K}_f + \mathbf{Q}_p$ to substitute \mathbf{K}_{f1} , then we obtain

$$A(\mathbf{K}_f, z_j) A^*(\mathbf{K}_f + \mathbf{Q}_p, z_j) v_j^*(-\mathbf{Q}_p) \quad (6.29)$$

Then we consider about the linear summation, the 2nd term in equation 6.22 can be written as

$$\sum_{j=1}^n A(\mathbf{K}_f, z_j) A^*(\mathbf{K}_f + \mathbf{Q}_p, z_j) v_j^*(-\mathbf{Q}_p) \quad (6.30)$$

With the same procedure, the 3rd term in equation 6.22 can be written as:

$$\sum_{j=1}^n A^*(\mathbf{K}_f, z_j) A(\mathbf{K}_f - \mathbf{Q}_p, z_j) v_j(\mathbf{Q}_p) \quad (6.31)$$

In the last, we obtain the expression of $G(\mathbf{K}_f, \mathbf{Q}_p)$:

$$G(\mathbf{K}_f, \mathbf{Q}_p) = a(\mathbf{K}_f)^2 \delta(\mathbf{Q}_p) + \sum_{j=1}^n A(\mathbf{K}_f, z_j) A^*(\mathbf{K}_f + \mathbf{Q}_p, z_j) v_j^*(-\mathbf{Q}_p) \\ + \sum_{j=1}^n A^*(\mathbf{K}_f, z_j) A(\mathbf{K}_f - \mathbf{Q}_p, z_j) v_j(\mathbf{Q}_p) \quad (6.32)$$

If we assume the sample is uniform, $v_j(\mathbf{Q}_p)$ should be the same for each layer and we can rewrite them as:

$$G(\mathbf{K}_f, \mathbf{Q}_p) = a(\mathbf{K}_f)^2 \delta(\mathbf{Q}_p) + \left(\sum_{j=1}^n A(\mathbf{K}_f, z_j) A^*(\mathbf{K}_f + \mathbf{Q}_p, z_j) \right) v_0^*(-\mathbf{Q}_p) + \\ \left(\sum_{j=1}^n A^*(\mathbf{K}_f, z_j) A(\mathbf{K}_f - \mathbf{Q}_p, z_j) \right) v_0(\mathbf{Q}_p) \quad (6.33)$$

v_0 is just the object function in each slice in the reciprocal space. The contrast transfer function then can be written as the summation of the double A function. This is the idea of iCTF.

6.3 Results and Discussion

Figure 6.2 shows noise free SSB ptychography images of 24 nm thick (100) oriented STO simulated using the multislice algorithm with the probe focused to different distances from the entrance surface. The simulations were performed using abTEM [54] in combination with pyPtychoSTEM [29]. A 200 kV accelerating voltage, a 20 mrad convergence angle, and a scan step size of 0.2 Å were used. The entrance surface of the sample is generally the optimal focal plane for ADF imaging, but this focus produces contrast reversals in every atomic column of the STO in the SSB image with this thickness of STO. Furthermore, the overall contrast in the SSB image is much lower compared to focusing the probe more into the middle of the sample. A matching iCoM simulation is shown in the supplementary information.

Consistent with our previous simulations of other STO thicknesses up to 28 nm [58], the optimum contrast occurs with the focal point of the probe midway through the sample. Furthermore, from the images shown using the same grayscale rendering, this central focal point clearly produces both contrast reversal free images and the strongest contrast overall.

Figure 6.3 compares SSB, iCoM, ABF and ADF images simulated for the same STO projection as figure 6.2, again as a function of how far within the sample the probe is focused, but now with the finite dose of $2000 \text{ e}^-/\text{Å}^2$. The ABF, iCoM and SSB ptychography signals are all sensitive to light atoms, with the iCoM and SSB signals capable of providing far stronger signal than the ABF. However it is only with the probe focused more towards the center of the sample that iCoM and SSB provide the strongest contrast

6.3. RESULTS AND DISCUSSION

(we provide the iCoM results with varying defocus and thickness of STO in SI, the SSB results can be found in the SI in paper [58]). With the probe focused to the entrance surface, and this particular sample configuration, the ABF actually provides stronger contrast. It can also be seen here how the ADF and ABF contrast reduces as the focus is moved further into the sample from the entrance surface. While this focal dependence has been explained for ABF [60] and iCoM [45], so far relatively little has been published on this for direct ptychography beyond observation [58,59].

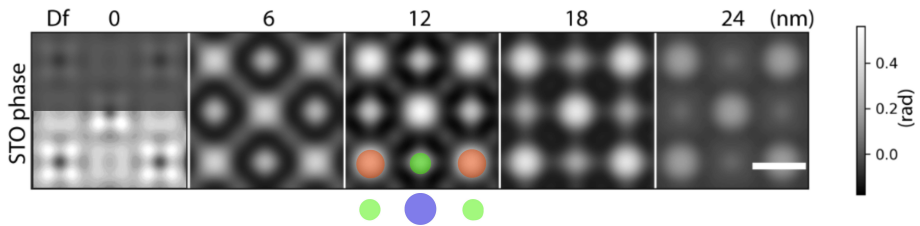


Figure 6.2: SSB ptychography images of 24 nm thick (100) STO simulated noise free with the probe focused to different depths of the sample from the entrance surface (0 nm of defocus) to the exit surface (24 nm of defocus). Overlaid on the bottom of the zero defocus image is a contrast boosted version to facilitate seeing the contrast reversals. A half unit cell model is also overlaid on the bottom of the 12 nm defocus image to aid identification of the columns with Ti shown in orange, O in green columns and Sr in purple. The scale bar is 2 Å.

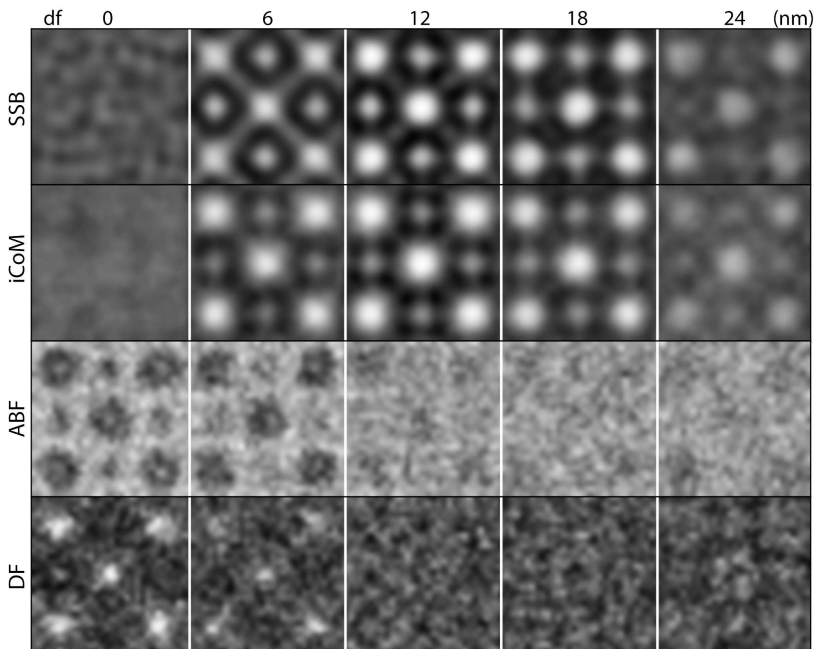


Figure 6.3: Probe focal point dependence (with zero defocus (df) defined here as the entrance surface of the sample) of the contrast of the SSB, iCoM, ABF, and ADF signals simulated at a dose of $2000 \text{ e}^- / \text{Å}^2$ for 24 nm thick STO.

6.3.1 Exploring the iCTF concept in relation to contrast reversals

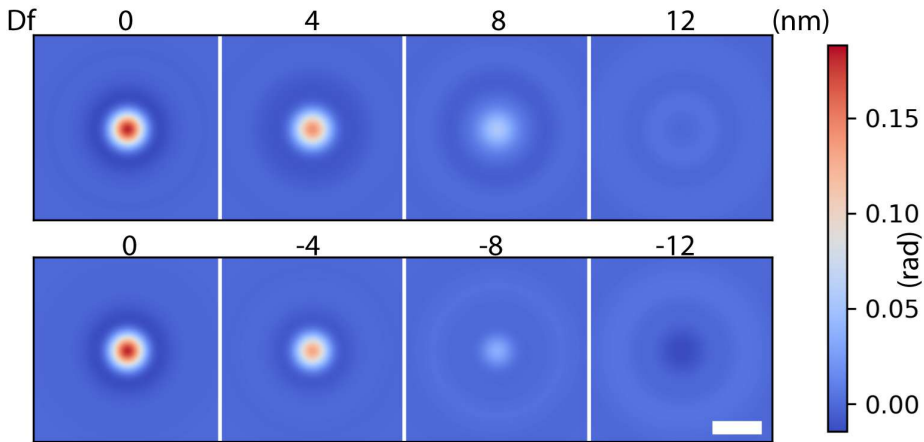


Figure 6.4: U phase images with different defocus values. Negative values denote overfocus condition. Positive values denote underfocus condition. The scale bar is Å.

With the iCTF concept, In the absence of other aberrations, in the plane at which the probe is focused there is no defocus contribution and phase only comes from the object function. For a thin weak object the phase is flat across the double disk overlap regions. However defocus induces phase variation within the double disk overlaps such as that illustrated in figure 2.12. With the probe focused to the central plane of the sample, the defocus of two slices equidistant from the center of the sample is equal and opposite. Therefore, if we neglect multiple scattering, and the sample is the same in these two slices, the contribution of defocus in these two planes will be equal and opposite, and might be expected to cancel out. It seems iCTF theory could explain why central focusing works well for overcoming contrast reversals.

To check if this theory is correct, I simulated 4D datasets of a single uranium atom at 200 kV with a 20 mrad convergence angle with defocus values of +12 and -12 nm. The simulation results are shown in figure 6.4. For the slices with opposite defocus values, if the iCTF concept is correct, the summation of the two object functions in reciprocal space can cancel out the influence of the defocus. The results of summing the +12 and -12 nm defocus slices is shown in figure 6.5. In this step, the influence of the defocus should be cancelled out according to the iCTF. However the third column of figure 6.5 shows the amplitude and phase images from the new combined object function in the real space has very significant artifacts, which shows the iCTF concept does not work well even for a single U atom. STO is a stronger object and therefore the observation of central focusing being optimal can not be understood with the iCTF theoretical model.

6.3.2 Influence of the convergence angle

Central focusing has been observed to be optimal for SSB imaging of STO thicknesses of 16, 20, 24 and 28 nm by Gao et al. [58]. For reference we provide in the supplementary information iCoM versions of the defocus series published in Gao et al. of these thick-

6.3. RESULTS AND DISCUSSION

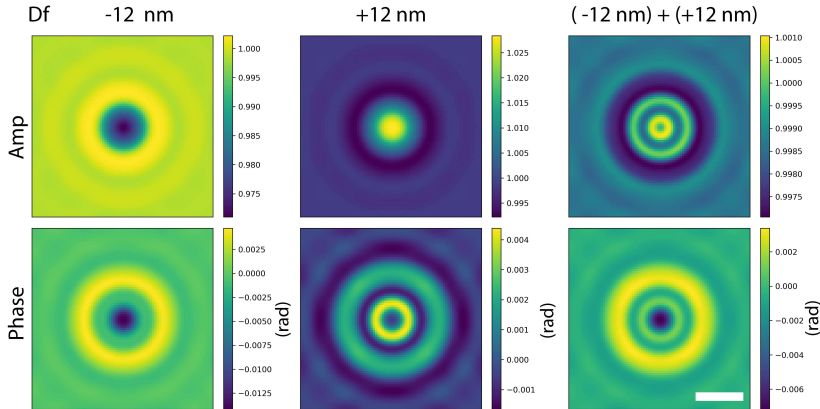


Figure 6.5: Amplitude and phase images of a single U atom. The first two columns show the images from -12 and +12 nm of defocus. The third column shows the reconstruction resulting from summing the two object functions from the two defocus values in reciprocal space. We can see significant artefacts remain and the cancellation of defocus does not occur as would have been expected using the iCTF concept. The scale bar is Å.

nesses of STO, and these also show central focusing to be optimal. Interestingly, Clark et al. [59] observed a band of thicknesses for GaN in which central focusing did not remove the caldera shaped contrast reversals [59]. It seems noteworthy that this band of thicknesses was within the range one would consider quite thin samples and are within the 17 nm depth of field for the 14.4 mrad convergence angle and 300 kV conditions they used. This is in contrast to the results here and all the thick samples examined in Gao et al. [58], where the samples are all thicker than the depth of field. We therefore now test what happens with thin 5 nm thick STO as a function of both physical probe focus and convergence angle in figure 6.6 for SSB and for iCoM in supplementary figure 6.16. For comparison to the GaN results, Ga is atomic number 31 while Sr is 38 and Ti 22, and on the lighter side N is atomic number 7 while O is 8. This thickness of STO therefore contains relatively similar projected potentials as the range of thickness of GaN which Clark et al. found continued to exhibit caldera with central focusing. We also find that central focusing makes relatively little difference with such a thin sample, when smaller convergence angles are used such that the depth of field is significantly larger than the sample thickness. Donuts appear on the heavier Sr and Ti columns regardless of which plane within the sample the probe is focused to.

However, when the convergence angle is increased the donuts close up and we again find central focusing becomes optimal. At the 200 kV we used, a 30 mrad convergence angle puts the depth of field at just under the 5 nm thickness and at 40 mrad the depth of field is significantly less than the thickness of the sample. At 30 mrad there remains a very slight dip in the heaviest columns with the probe at the entrance surface. At 40 mrad this dip has disappeared, but for both 30 mrad and 40 mrad the strongest contrast is still seen with the probe focused to the center of the sample.

With the whole sample within the depth of focus, the effect of defocus within the sample seems natural to make less difference than when the sample thickness exceeds the depth of field. Even though the defocus values are not by themselves different at each

CHAPTER 6. ON CENTRAL FOCUSING FOR CONTRAST OPTIMIZATION IN DIRECT ELECTRON PTYCHOGRAPHY OF THICK SAMPLES

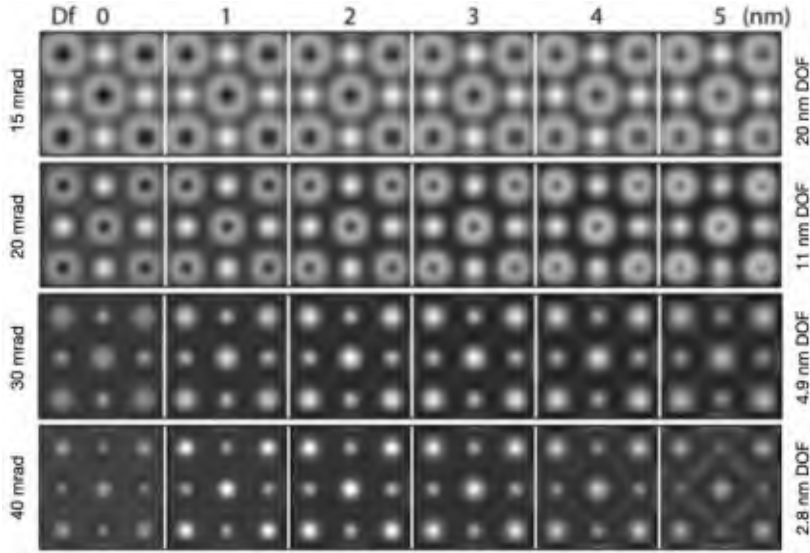


Figure 6.6: SSB images of very thin 5 nm thick STO as a function of physical probe focus within the sample depth (zero defocus (df) is again the entrance surface) as a function of the convergence angle. As the convergence angle increases such that depth of field is on par or shorter than the thickness the contrast reversals are overall diminished or absent and central focusing again provides the best contrast.

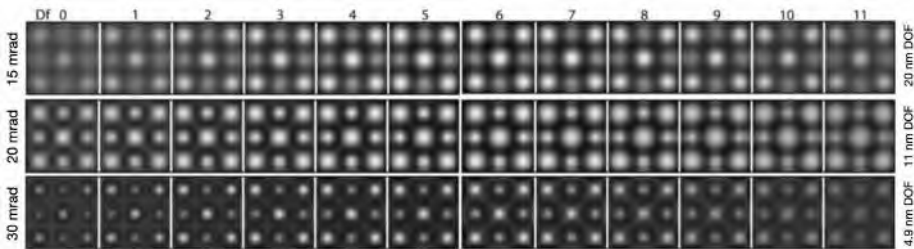


Figure 6.7: SSB images of 12 nm thick STO as a function of physical probe focus within the sample depth (zero defocus (df) is again the entrance surface) as a function of the convergence angle. For this thickness the central focusing again seems to work regardless of the convergence angle.

plane, when the probe is focused to different points within the sample with different convergence angles, the effect on the double disk overlaps as a whole should be reduced with a smaller range of angles included inside each double disk overlap as with the smaller convergence angles which have the sample entirely within the depth of field, as can be understood in terms of the optical sectioning effect.

At 200 kV the 20 mrad convergence angle we used earlier, the depth of field is 11 nm. So one might think that 5 nm of STO, with less sample potential within the depth of field, would show less contrast reversals than 16 nm or more of STO with the same conditions. As this is, interestingly, not what is observed we turned to 12 nm thick STO simulations. As shown in figure 6.7, regardless of if the convergence angle produces a depth of field

6.3. RESULTS AND DISCUSSION

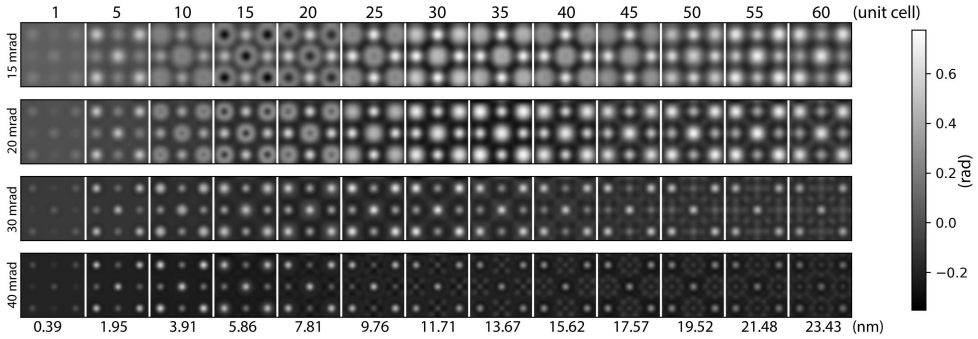


Figure 6.8: SSB images of 1 to 60 unit cells of STO using central focusing with 15, 20, 30 and 40 mrad convergence angles. Contrast reversals appear for the smaller convergence angles around 15 unit cells of thickness. The 15 mrad images continue to show a dip at the center of the heavier columns to quite high numbers of unit cells before disappearing, but contrast reversals do not appear on the atomic columns in this range of thickness for the higher convergence angles.

similar to, much greater than or much less than the 11 nm thickness of the material, central focusing is again optimal.

To further explore the thickness dependence with different convergence angles figure 6.8 displays the results of central focusing for 15, 20, 30 and 40 mrad convergence angles in 5 unit cell thickness steps from 1 to 60 unit cells. As can be seen there is significant variation of the contrast with respect to the convergence angle, but the donut contrast reversals do not reappear on the atomic sites at higher thickness, although they do persist at 15 mrad to the thicker end of this range of thicknesses.

Dynamical effects are generally more associated with thicker samples, and 5 nm is certainly on the thin side, so one might expect dynamical scattering to effect the contrast of samples thicker than 5 nm more. Indeed, the divergence from central focusing being optimal seen at 50 nm of STO, where focusing to the exit surface was instead found to be optimal previously by Gao et al. [58], seems more naturally a result of strong dynamical effects. A focal plane near the exit surface also turns out to be optimal for iCoM imaging, again behaving similarly to the SSB as shown in supplementary figure 6.13.

The convergence angle of course weights the frequencies differently [22,53], as is already usually considered when optimizing the contrast transfer function without considering contrast reversals. However different frequency weightings will also naturally change where asymptotes occur in the frequency response, due to phase wrapping in probe reciprocal space, which in turn seems to effect the appearance of contrast reversals. Further complicating the appearance of ptychographic images is the nonlinear effect of neighboring atomic columns [62]. Phase images which have a low response at low frequencies, including high pass filtered iCoM or iDPC images, will cause a negative halo around a positively phased atomic column or atom. This results in additional contrast between the columns even in the case of 2D materials and are therefore not by themselves an argument to move to an alternative probe focal plane.

Given that optimal contrast is generally obtained for ADF with the probe focused to the

entrance surface and for both iCoM and direct ptychography methods generally with the probe deeper within the sample, the optimal focus for ADF and iCoM and direct ptychography methods such as SSB can increasingly deviate as the sample increases in thickness. Different strategies can therefore be employed depending on the needs of the sample being investigated. One can optimize for the iCoM and ptychography by aiming to put the probe deep within the sample at the cost of contrast in the ADF signal, or one can put the focus to the entrance surface as is best for the ADF. In practice the second option has been easier to achieve experimentally, using the ADF to guide selecting a region of interest and final tuning, including the focus.

Live processing options are becoming available for both iCoM [55] and SSB [56,63] and beginning to enable the focus to be optimized for these imaging modes while scanning. However so far live imaging has not been able to keep up with microsecond level dwell times, and it is important to remember that the ptychographic contrast can also be adjusted significantly post collection [58] either by post collection defocus or other schemes to adjust the phases. This is equally true for WDD, the other direct method of ptychography, as for SSB.

Single slice iterative ptychography methods, such as ePIE [24], do not explicitly consider the effect of each plane of the sample having a different defocus. Single slice iterative methods which attempt to deconvolve the probe and remove the effects of aberrations as ePIE does, will also only solve for the defocus as it effects the whole sample, and indeed one also sees contrast reversals in ePIE data [59]. However it is possible to break the sample into multiple slices and treat the propagation of the probe explicitly using multislice ptychography [49,64] at the cost of considerable additional computational complexity. However for iterative methods the choice of defocus depends on various other considerations, and a large defocus is commonly employed to aid convergence. However the optimization of defocus for iterative methods is a highly complex non-trivial topic, and not within the scope of the present work.

The choice of focusing strategy will depend on the dose the sample can handle and the importance of the Z-contrast ADF signal, which remains very useful for correct interpretation when sufficient Z-contrast signal is available, but which again is lost when large a defocus is employed such as is common in iterative ptychography. The Z-contrast signal is all the more important when contrast reversals can occur with phase imaging, or indeed apparently different structures result using very slightly different conditions in iterative methods of ptychography. Thus it is very useful to understand the overall contrast variation of ptychography methods with defocus and thickness as well as the convergence angle.

6.4 Conclusion

I have shown further examples of the overall strength of direct SSB electron ptychography being at its maximum when the probe is focused midway through the sample. I explored the iCTF concept as a means of explaining this observation. Within the iCTF, most of the defocus contribution from sample planes away from the focal point of the probe would cancel out when the probe is focused to the center of a homogenous sample. However the iCTF requires that only linear terms in the potential are retained assuming

6.5. SUPPLEMENTARY INFORMATION

the sample is weak. This is quite an unrealistic approximation for STO. Furthermore I demonstrated that even with a single uranium atom this cancellation predicted by the iCTF concept does not occur. Summing the contributions from two equal and opposite defocuses produces an image of U with significant rather than a clear image of the atom. Our results show that for strong phase objects, the iCTF concept can not explain why central focusing is often optimal. In addition we have demonstrated the convergence angle can be a crucial factor in determining if contrast reversals will be present at all and if central focusing will remove them. This knowledge will therefore help inform the design of focused probe ptychography experiments of thicker samples in determining where exactly to focus the probe.

6.5 Supplementary information

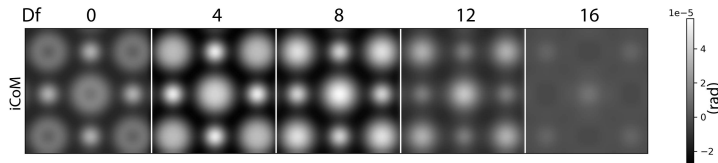


Figure 6.9: Simulated iCoM reconstructed phase results for 16 nm thick STO with varying probe defocus at infinite dose. The voltage is 200 kV, convergence angle is 20 mrad, scan step size is 0.2 Å. As we can see, when the probe focuses around the central plane, the phase images gives the best contrast.

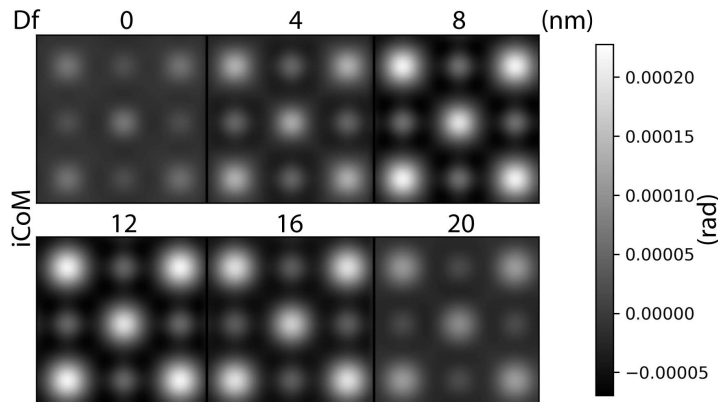


Figure 6.10: Simulated iCoM reconstructed phase results for 20 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.

CHAPTER 6. ON CENTRAL FOCUSING FOR CONTRAST OPTIMIZATION IN
DIRECT ELECTRON PTYCHOGRAPHY OF THICK SAMPLES

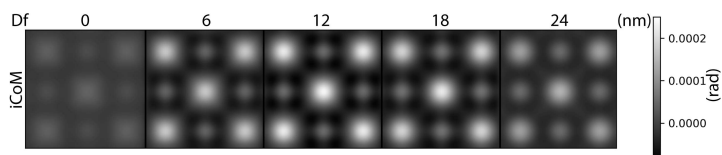


Figure 6.11: Simulated iCoM results for 24 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.

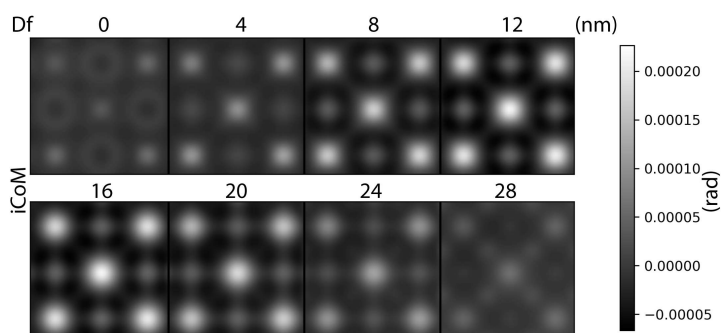


Figure 6.12: Simulated iCoM reconstructed phase results for 28 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.

6.5. SUPPLEMENTARY INFORMATION

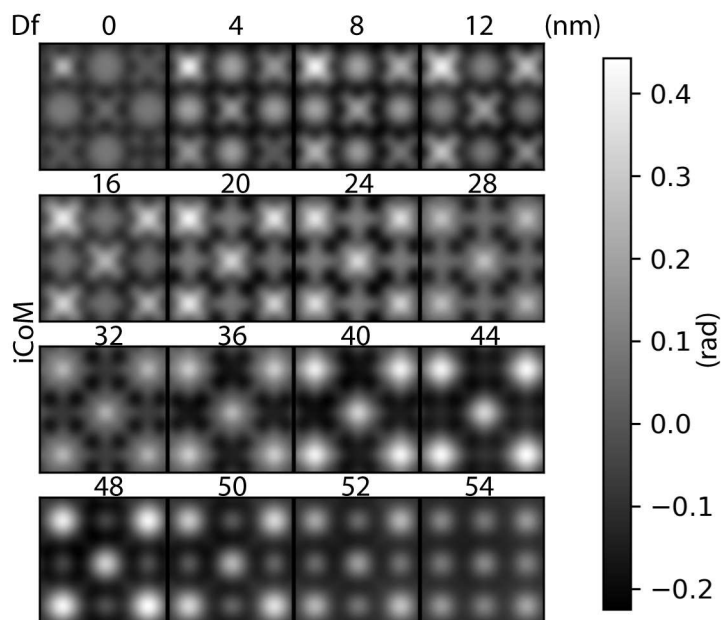


Figure 6.13: Simulated iCoM reconstructed phase results for 50 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast. At this thickness central focusing is no longer optimal. Instead focusing the probe near the exit surface of the sample (defocus values from 48 nm to 54 nm) gives the best contrast.

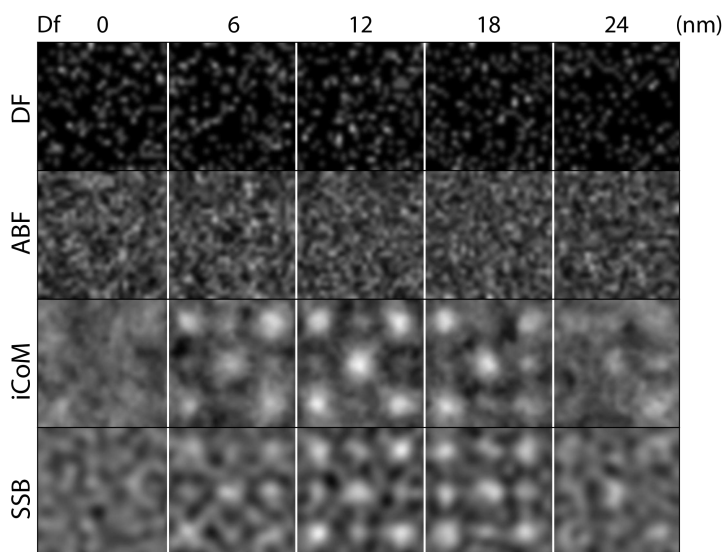


Figure 6.14: ADF, ABF, iCoM and SSB results simulated at the dose of $100 \text{ e}^-/\text{\AA}^2$ for 24 nm thick STO, with the other simulation parameters the same as for figure 2 of the main text. From the figure, we can see that when the dose is extremely low, no structure is visible in either the ADF or ABF images at any defocus conditions. However with iCoM and SSB, central focusing gives sufficient signal to noise to resolve all the atomic columns.

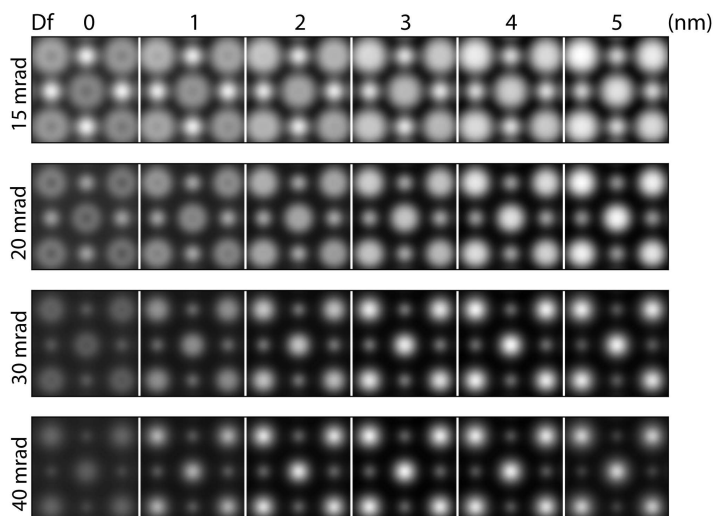


Figure 6.15: iCoM version of figure 6.6 of the main text, showing simulate infinite dose images of 5nm thick STO using different defocus values, covering the entire thickness of the sample, with different convergence angles. Conditions match those of figure 7.

6.5. SUPPLEMENTARY INFORMATION

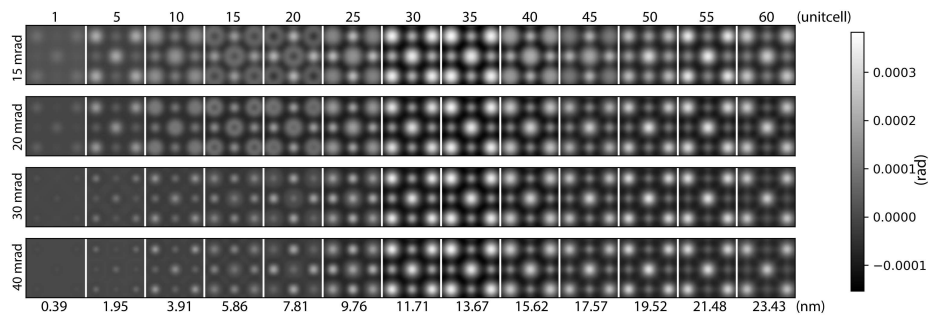


Figure 6.16: iCoM version of figure 9 of the main text showing simulated ininite dose images of STO with different thicknesses of STO and different convergence angles. Similar with the SSB results, when the convergence angle is 15 or 20 mrad, contrast reversals appear around the 5 nm thickness. With the increasing of convergence angle, the contrast reversals also disappear with the iCoM.

*CHAPTER 6. ON CENTRAL FOCUSING FOR CONTRAST OPTIMIZATION IN
DIRECT ELECTRON PTYCHOGRAPHY OF THICK SAMPLES*

Phase offset method of ptychographic contrast reversal correction

7.1 Introduction

Electron ptychography offers very high dose efficiency [51,65,66], the ability to reveal the locations of light elements neighboring heavy atoms [28,58], post collection aberration correction and superresolution [48,49]. These advantages make the method a very attractive complement to Z-contrast annular dark field (ADF) workflows [9], with the phase images providing stronger images of the structure and the Z-contrast stronger sensitivity to composition. With advances in cameras having greatly reduced or completely removed the problem of slow scans with 4D STEM [15,67], there is now relatively little drawback to collecting data for ptychography.

Compared to phase contrast imaging with conventional high resolution transmission electron microscopy (HRTEM), the contrast transfer function (CTF) of direct focused probe methods such as single side band (SSB) [21,22,53] and Wigner distribution deconvolution [50] are very simple, requiring no aberrations to form contrast and exhibiting no zero crossings. This makes these ptychographic methods much easier to interpret than HRTEM, at least for potentials that are not overly strong. The phase is related to the strength of the potential encountered by the beam electrons and thus as the strength of the potential increases eventually the phase can exceed the limit imposed by the 2π range of values available to phase and wrap around. This means that as phase increases it can suddenly go from being maximally positive at π to maximally negative at $-\pi$, causing a very large change in contrast from only a small change in the sample. Therefore, even though the SSB ptychography CTF, derived using the same weak phase approximation as HRTEM CTFs, shows all frequencies being passed with the same sign, contrast reversals can occur because of wrap around.

One of the most observed contrast reversal behaviour in atomic resolution imaging is a dip in the phase at the center of the atomic columns [50,58,59]. These often appear as donut shapes in the images and like a volcano with a caldera in line profiles, and represents a reversal from the centrally peaked probe shaped atoms one observes when the potential is weaker. This makes some intuitive sense as the center of an atomic column is the location of the strongest potential, and thus it is natural to expect that

this will be where wrap around will occur first as the potential increases. This intuitive expectation is also in accord with the fact that it is the heavier atomic columns that exhibit donuts first as thickness increases [58]. Such contrast reversals have also been observed in iCoM and iterative ptychography [59, 68] as well as S-Matrix inversion [69–71]. This again makes intuitive sense as all these methods are attempting to retrieve the same phase shift induced on the beam electrons by the sample. Perhaps less intuitive is the fact that the range of phase values in the final images generally remains much less than 2π in atomic resolution imaging, at least in single slice ptychography, but it is not just the phases in the final image that can phase wrap – the individual frequency components can also phase wrap.

If the goal is to locate the light elements hidden in the ADF signal by strong scattering of nearby heavy elements, the appearance of donut contrast on the heavier columns is often not a significant impediment. However as the thickness increases the contrast can become more complex [58]. Furthermore, it is often preferable for images of atomic structure to resemble as much as possible to the relatively simple probe shaped spikes in intensity occurring in ADF imaging, even if simply for ease of interpretation. However the contrast reversals can also degrade overall contrast and reduce the visibility of the structure at lower doses [72], as well as complicate quantification.

For quite a range of thicknesses, central focusing of the beam offers phase images free from contrast reversals and with the strongest contrast overall [58, 59, 72]. However this conflicts with the optimal focus for ADF imaging, the entrance surface. Thus optimizing the probe focus for the phase images during acquisition can significantly degrade the quality of simultaneously acquired ADF images, especially as the sample thickness increases and the distance between the entrance surface and optimal focus for ptychography widens. Fortunately post collection adjustments can be applied. The ability of ptychography to adjust aberrations post collection can be leveraged to apply a post collection defocus which can often remove the contrast reversals [58]. However the application of post collection defocus also often reduces the overall contrast, even if the atoms all appear “atom like” after the contrast reversal correction. Furthermore, in some cases, post collection defocus does not remove contrast reversals with satisfying results, and indeed in other cases neither does physically focusing the probe during data acquisition [59, 72]. Another interesting approach to overcome contrast reversals is multislice ptychography [64]. Here, the specimen is divided into multiple slices and the phase is solved in each slice separately. Crucially, each slice is thin enough so that phase wraps are avoided within the slices.

Here we delve deeper into the phase wrapping process causing the contrast reversals in direct ptychography, and demonstrate a superior way to counteract them than post collection defocus, which does not rely on focus. As the potential of a single atom is increased, the phases of the spatial frequencies change nonlinearly, with the higher frequencies changing more quickly than the lower frequencies. This means that as the potential is increased, the higher spatial frequencies eventually hit the limit imposed by the 2π range of phases available and wrap around. Once a frequency wraps around, its phase tends to contrast very strongly with the frequencies that have not yet wrapped around. Thus these asymptotes in the phase response produce contrast reversals, and they can do so without any dynamical scattering. Applying defocus can roll the phases back around or reduce the amplitude of the wrapped around frequencies sufficiently that the contrast reversals can be removed. However, we find directly adjusting the phases

7.2. THEORY

with an offset applied to the zero frequency (DC) phase is generally a superior method. We show that this method can robustly counteract contrast reversals regardless of the thickness or initial focus of the probe. Although central focusing remains preferable for the absolute best phase contrast in many cases, the phase offset method provides significantly improved contrast compared to defocus correction when a post collection solution is required. Furthermore, the phase offset method can correct contrast reversals in cases where physical defocus cannot satisfactorily do so. The ability of the phase offset to retain contrast is especially important when the sample is fragile and one has a low dose budget. We demonstrate this experimentally at a dose of at $50 \text{ e}^-/\text{\AA}^2$ with a thin highly beam sensitive methylammonium (MA)- PbI_3 perovskite solar cell material [73,74] which exhibits contrast reversals that cannot be corrected by defocus at all, whether applied during data collection or after. Furthermore, because the phase offset does not rely on changing focus, it is a useful tool to apply in the context of optical sectioning in which the focus is vital to determining the 3D location of objects. The phase offset allows one to better interpret high resolution ptychographic focal series where the effects of strong potentials and the focal dependence of the contrast reversals are intertwined. The phase offset method can provide a means of disentangling these effects and seeing the objects as clearly as possible in each slice.

7.2 Theory

For the phase offset method, it is very similar to the Zernike phase contrast method [16]. Zernike phase contrast imaging is a powerful technique in optical microscopy that enhances the contrast of transparent, unstained specimens, making it an invaluable tool in fields such as cell biology and microbiology. Developed by Frits Zernike in 1934, this technique exploits the phase differences in light waves passing through a transparent specimen. These phase differences arise due to variations in the refractive index and thickness of different parts of the specimen.

7.2.1 Conventional Zernike phase contrast imaging theory.

For a thin sample with phase object approximation, the transmission function $t(x, y)$ can be written as

$$t(x, y) = \exp(i\phi(x, y)) \quad (7.1)$$

where $\phi(x, y)$ is the phase shift of the incident beam induced by the sample. If $\phi(x, y)$ is small enough, with Taylor expansion and the first order approximation, it can be written as

$$t(x, y) = \exp(i\phi(x, y)) \approx 1 + i\phi(x, y) \quad (7.2)$$

For Zernike phase contrast imaging method, it introduces a phase shift of $\pi/2$ to the unscattered beam by placing a phase plate in the back focal plane. The phase plate is typically a thin, transparent disk with ring structure. The central area (transmitted light) of the phase plate introduces the phase shift, while the surrounding area (diffracted light) usually remains unaffected. Then the intensity of the beam $I(x, y)$ in the real space can be expressed as:

$$I(x, y) = |i + i\phi(x, y)|^2 \approx 1 + 2\phi(x, y) \quad (7.3)$$

By this way, the phase shift can be visualized from the intensity of the image.

7.2.2 Phase offset method

We still use the weak phase object approximation as in equation 7.2 and do Fourier transform on the transmission function $t(x, y)$:

$$\mathcal{F}(t(x, y)) \approx \mathcal{F}(1) + i\mathcal{F}(\phi(x, y)) \approx \delta(k_x, k_y) + i\Phi(k_x, k_y) \quad (7.4)$$

For the phase offset method, we add a phase offset Δ to the non zero frequencies of the object function. It equals to:

$$\delta(k_x, k_y) + [i\Phi(k_x, k_y)] * \exp(i\Delta) \quad (7.5)$$

Next we perform the inverse Fourier transform and obtain the object function in the real space:

$$\begin{aligned} \mathcal{F}^{-1}(\delta(k_x, k_y)) + i \exp(i\Delta) * \mathcal{F}^{-1}(\Phi(k_x, k_y)) &= 1 + i \exp(i\Delta) * \phi(x, y) \\ &= 1 - \sin(\Delta)\phi(x, y) + i \cos(\Delta)\phi(x, y) \end{aligned} \quad (7.6)$$

From this equation, we can see that the phase offset does transfer phase information into the real part $1 - \sin(\Delta)\phi(x, y)$ with a phase offset Δ , that information from the phase domain appears in the real part. The amplitude f of the new object function can be expressed as:

$$f = \sqrt{(1 - \sin(\Delta)\phi(x, y))^2 + (\cos(\Delta)\phi(x, y))^2} = \sqrt{1 - 2 \sin(\Delta)\phi(x, y) + \phi(x, y)^2} \quad (7.7)$$

With weak phase approximation, if $\phi(x, y)$ is small, we can ignore the high order term of $\phi(x, y)$ and then the new amplitude f information can be expressed as:

$$f = \sqrt{1 - 2 \sin(\Delta)\phi(x, y)} \quad (7.8)$$

Then the intensity of the object function can be expressed as:

$$I(x, y) = |f f^*| = 1 - 2 \sin(\Delta)\phi(x, y) \quad (7.9)$$

which is quite similar with the Zernike phase contrast imaging method as shown in equation 7.3. In fact one obtains the Zernike result using $\Delta = -\pi/2$. However it is important to note that it is found that other phase offset values can produce better results than the Zernike offset, depending on the sample.

7.3 Results and Discussion

Figure 7.1 displays single atom SSB images starting with a U atom and then increasing its potential incrementally up to a factor of 20. The multiplied potentials correspond

7.3. RESULTS AND DISCUSSION

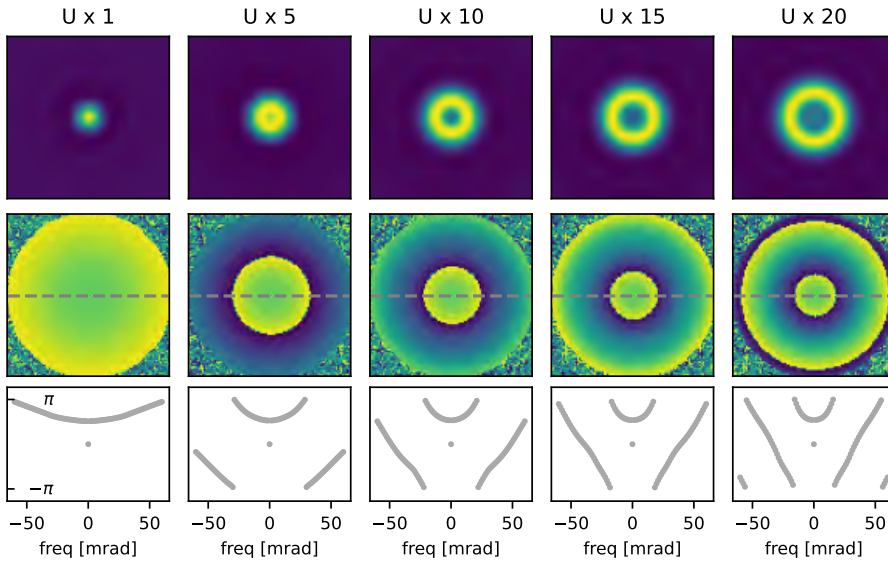


Figure 7.1: Single atom SSB simulations using potentials ranging from a U atom potential to 20 times that potential, showing how the potential strength itself causes contrast reversals, which manifest here as donut shaped atom contrast. The 2nd row shows the phase vs frequency in 2D, and the 3rd row line profiles of the rotationally symmetric phase response. As the potential increases the curvature of the phase increases, the phase hits the top limit and wraps around resulting in contrast reversals at the center of the atom. As the wrap around shifts to lower frequencies the donut hole expands. Scale bar is 2 \AA .

to atoms heavier than any known, but these super heavy atoms allows us to probe the effect of the potential without dynamical scattering complicating matters. For SSB reconstructions, we use the PyPtychoSTEM package [29] with the 4D data simulated with abTEM [54] at 200 kV with a 30 mrad convergence angle.

Below the images we plot the phase vs frequency first in 2D, from which it can be seen that phases are continuously rotational symmetric, and then as line plots. The line plots better allow one to see the limits of the 2π phase range available. With the already very heavy U atom, the phase curves significantly upwards as a function of the magnitude of the spatial frequency, but remains entirely within the 2π phase range without any phase wrap. Multiplying the U potential by a factor of 5, the phase increases more rapidly with respect to frequency, and a phase wrap occurs as the phase extends beyond π . This results in a significant proportion of the higher spatial frequencies switching from being strongly positive to strongly negative, creating a strong contrast between frequencies lower than and higher than the asymptote, and resulting in a contrast reversal in the form of a small donut hole appearing in the center of the image of the atom.

As the potential is increased to $U \times 10$, the phase vs frequency curvature further increases. The first wrap around asymptote shifts to lower frequency, a second wrap around point appears and the donut hole increases in size.

As the potential is further increased the curvature increases further still, continuing to alter the balance between positive and negative phase frequencies, and further increasing the width of the donut hole. The trend continues as we increase to $U \times 20$, but now the wrapped phase increases in curvature sufficiently to itself surpass the upper limit and itself wrap around.

Figure 7.2 shows how the different ranges of frequencies influence the phase image for the $U \times 20$ single atom potential using masking in probe reciprocal space. Before applying any masking, we see the two phase wraps occurring within the 2α range of frequencies passed using our 30 mrad convergence angle (α), with the last wrap occurring almost at the 2α limit of the SSB contrast transfer function [22, 53]. These two phase wraps define three distinct frequency ranges with strongly contrasting phases. Starting with all spatial frequencies included (Fig. 7.2a), the single atom appears as a thin ring of strongly positive phase with a lower phase region in the center. The central region has a small peak of slightly more positive phase within the central donut hole with this potential.

As we mask out the higher frequencies (Fig. 7.2b) we exclude both the strongly negative frequencies higher than the second phase wrap and the positively phased before it, leaving the negatively phased frequencies after the first phase wrap. The results in the ring of positive phase in the image filling inwards, and the small positive peak in phase at the center of the atom disappears and instead becomes a minimum. As we mask more of the frequencies between the first asymptote and second asymptote, the donut shape fills in more, with the central hole becoming more positive overall but still dipping significantly compared to the phase further from the center of the atom (Fig. 7.2c). In this step we have further reduced the number of strongly negatively phased frequencies above the first phase wrap asymptote contrasting with the strongly positively phased frequencies below the asymptote. When we remove all the frequencies above the first wrap point, we remove these final strongly contrasting negatively phased frequencies and the atom turns “atom shaped” (Fig. 7.2d). We note that this phenomenon could also be due to an image convolution with a larger point spread function as the mask size is decreased.

As has been shown previously, defocus can often be used to counteract contrast reversals [58, 59, 72]. For many samples, physically placing the probe focus in the centre of the sample is often found to be optimal. With a single atom, however, there is no difference between the entrance surface and the center of the sample. Thus there is no significant variation of defocus through the sample, balancing of which being the reason central focusing is optimal in many cases [72]. However we can still counteract the contrast reversals of a single atom with physical defocus as shown in figure 7.3 using a 5 nm defocus with the $U \times 20$ potential. While physical defocus often produces the highest contrast, this is not always the case and the option to counteract the reversals with the probe focused elsewhere can also be a significant benefit. This can often be achieved with the ability of ptychography to alter aberrations post collection [58]. Figure 7.3 illustrates this for the $U \times 20$ potential using a 3 nm defocus applied post collection.

Physical and post collection defocus are generally found to behave somewhat differently, as is also the case here despite this being a single atom with presumably insignificant dynamical effects. Although both physical and post collection defocus remove the donut shape in the SSB image, the behavior further from the atom is different with a ring of

7.3. RESULTS AND DISCUSSION

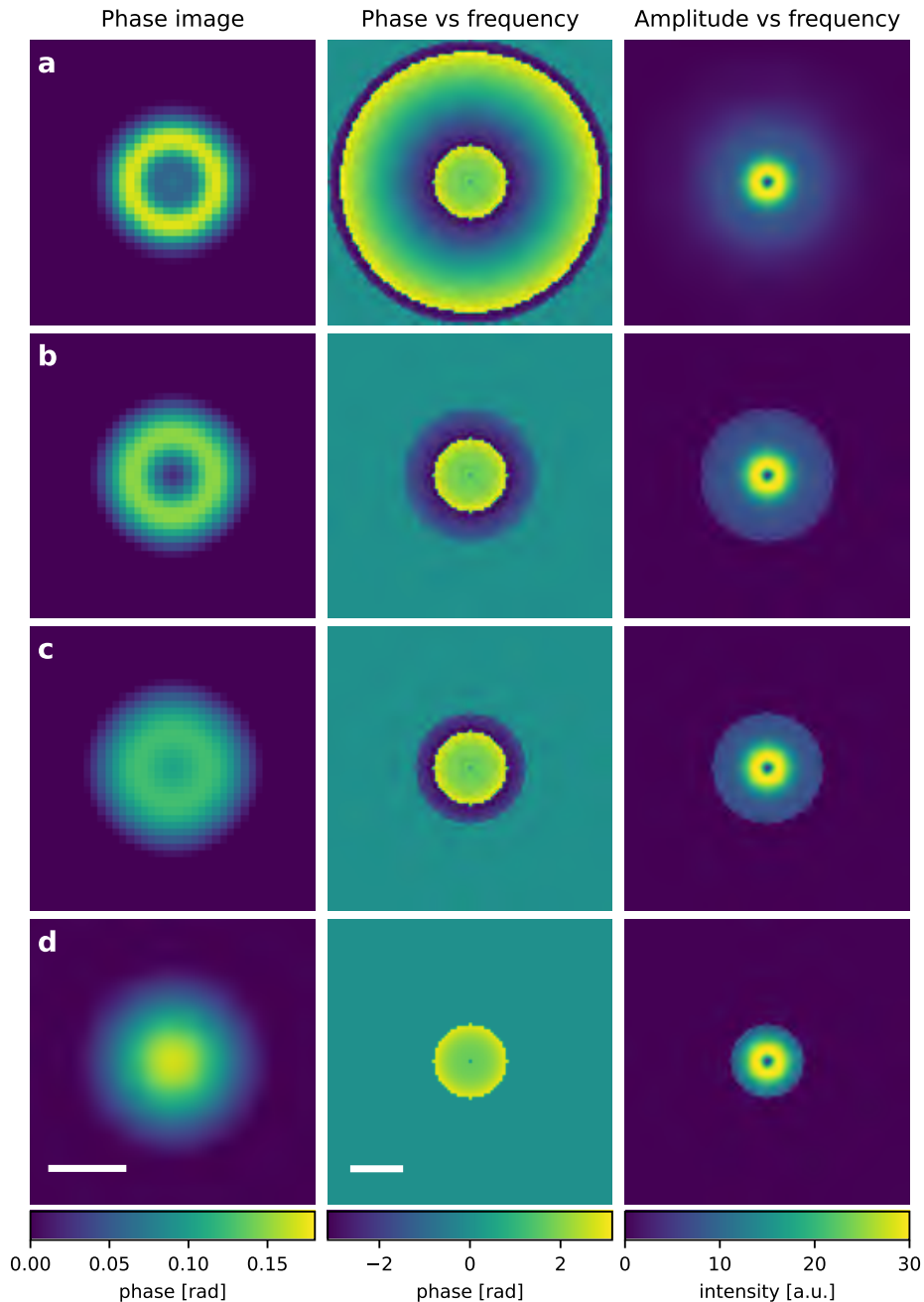


Figure 7.2: Illustration of the effects of the different frequency ranges on the simulated SSB image of the $U \times 20$ potential using masking in Fourier space. With the full range of frequencies out to 2α included (a), the atom appears as a thin ring with a slight peak in its center. As we progressively mask out the frequencies after each phase wrap around asymptote (b–d), the positive ring of phase progressively fills more of the central region of the atom, until after removing the contributions from all frequencies above the first wrap around it becomes atom like again, with a peak at the center of the atom. Of course, by limiting the contribution to lower spatial frequencies the image is also limited in resolution. Scale bar is 1 \AA for the phase image and 20 mrad for the phase and amplitude vs. frequency plot, respectively.¹⁰³

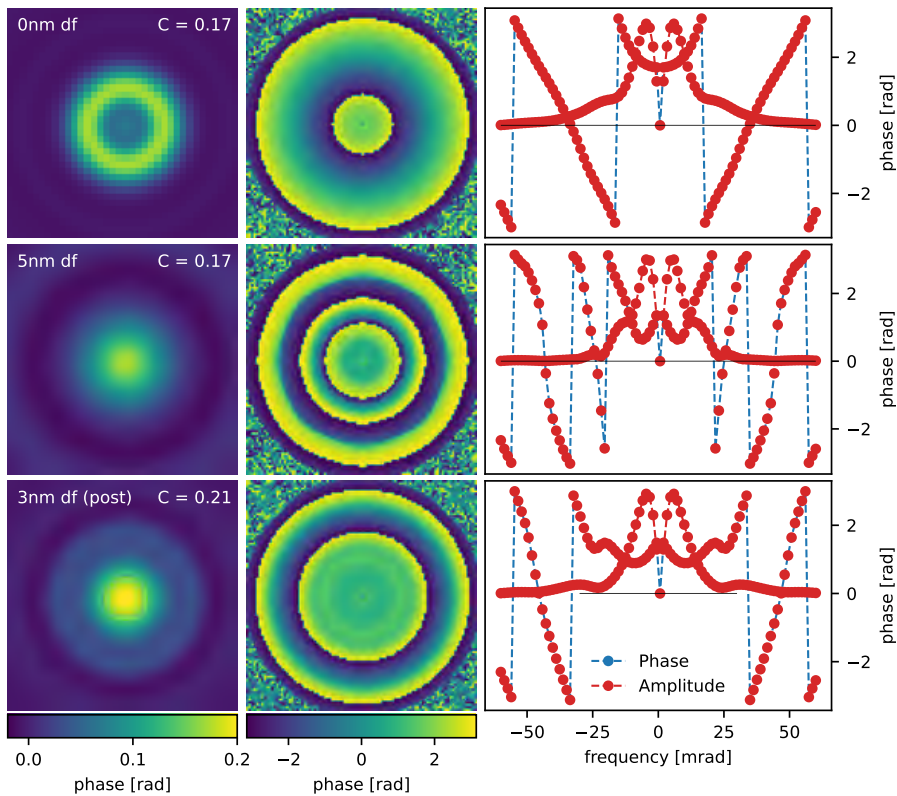


Figure 7.3: Illustration of the influence of post collection and physical defocus on SSB ptychography with simulations using the $U \times 20$ potential. The SSB images are displayed on the same intensity scale, with the contrast ratio (C) of the maximum phase to background phase indicated. The full frequency response is displayed in the 2D phase vs frequency plots. From the line profiles it is apparent that the physical defocus brings the amplitudes close to zero after the first wrap around, whereas the post collection defocus pushes the first wrap point out to higher spatial frequencies, resulting in stronger contrast in this case.

7.3. RESULTS AND DISCUSSION

higher phase appearing in the post collection defocus results closer to its center than in the physically defocused case. The 2D plot of phase vs frequency is more complex in the physical defocus case here, with the post collection appearing to nonlinearly push the phase wraps out to higher frequency. There are still two phase wraps in the post collection case, but they are concentrated closer to the 2α transfer limit, leaving a broader range of unwrapped lower frequencies. In the physically defocused case, it appears more that it is the suppression of the amplitudes of the phases after the first wrap that results in the contrast reversal removal. We observe here three phase wraps, but with low amplitudes the strongly contrasting phases of the wrapped higher frequencies do not contribute significantly to the image.

However, the reduction of contrast that results from post collection defocus motivated us to search for an alternative strategy to counteract contrast reversals. We present here what we call the phase offset method.

Given that the DC term provides the baseline phase which all other frequencies interfere with when transforming from probe reciprocal space into real space to form an image, by altering the relative phase of the DC term and all the other phases with a rigid offset, we can manipulate the phase wrap point and move it to higher frequencies without otherwise altering the overall shape of the phase vs frequency plot. In practice one can simply shift the DC term itself, although for the purposes of illustration here we instead shift the phases of the other frequencies while keeping the DC term phase constant in our plots of phase vs frequency as this better shows the effect on the phase wraps.

Figure 7.4 illustrates the use of a phase offset with the $U \times 10$ potential. Without any correction the phase vs frequency curve displays a sharp jump from the DC term to the first nonzero frequency. Applying an offset of -1.8 radians to the nonzero frequency components rigidly shifts the phase curve down such that there is no discontinuity moving from the DC term to the higher spatial frequencies until the positive π upper limit is hit and the phase wrap occurs as shown in Figure 7.4e. Importantly, with this offset the phase wrap occurs at a significantly higher spatial frequency than without the offset, and as can be seen from the figure the resulting image, Figure 7.4c, is donut free and much more closely resembles the shape of a lighter single atom that does not cause wrap around. There is a negative halo, but this is normal for a single atom in ptychographic images [53,62]. If we increase the phase offset to -2.7 radians the phases instead obtain a negative sign in the low frequency region which contrasts with the now positive higher frequencies and magnifies the negative halo leading to an “inverted donut” (cf. Figure 7.4d). This implies that the “best” offset value is the one which provides as much of a single signed phase curve as possible.

While a single atom is a rather simple system, the phase offset method also works well with crystals. As a first example, Fig. 7.5 examines the use of the phase offset with 16 nm thick SrTiO_3 (STO). In panel a the probe is focused to the entrance surface, which we emphasize is the best condition for ADF imaging. However, this leads to contrast reversals at the heavy Sr and Ti sites in the SSB image. Physically focusing to the middle of the specimen, as in Fig. 7.5b, removes the contrast reversals as a result of the defocus phase compensation of different layers [72]. To correct the contrast reversals using post collection defocus applied to data taken with the probe physically focused to the entrance surface during acquisition requires, in this case, a significantly larger defocus which, as we will show, results in a significant contrast reduction of the phase image.

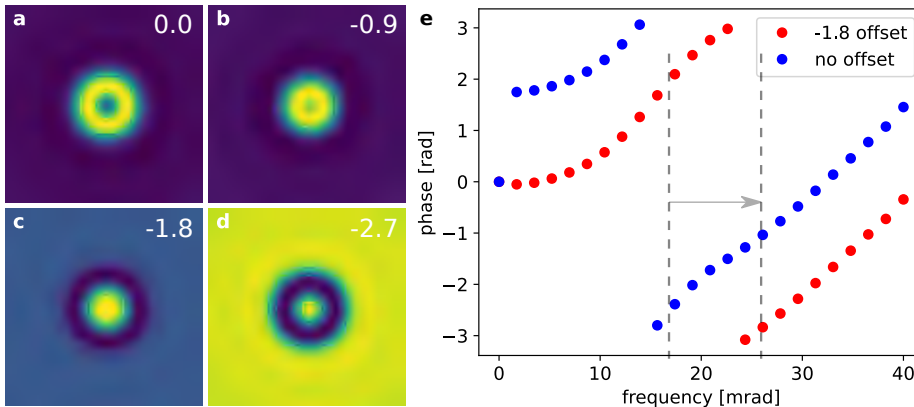


Figure 7.4: Illustration of the use of a phase offset on strong phase objects. (a–d) SSB images simulated with the $U \times 10$ potential using phase offsets of 0.0, -0.9, -1.8 and -2.7 mrad on the nonzero spatial frequencies with the DC values set to zero. (e) Line profiles of the phases with no offset and the optimal -1.8 rad offset, again with only the upper phase vs frequency curve shown for simplicity.

Post collection defocus adjustment often leads to sufficiently large contrast reduction that the atoms are not visible at low doses such as the $500 \text{ e}^-/\text{\AA}^2$ used in the bottom row of Fig. 7.5. As seen in the figure, the image in which the contrast reversals have been corrected using physical defocus remains quite clear at this dose. Compared to post collection defocus, the phase offset does not reduce the contrast nearly as much, providing an image in which all the locations of the atoms are easily identified also at the lower dose.

We note that as the ptychographic contrast is not as high here with the focus set at the entrance surface even with the phase offset method, compared to physically focusing to the center of the sample, one must choose to prioritize either optimal ptychographic contrast at the expense of the ADF or having a better ADF contrast by focusing to the entrance surface and compensating the ptychography with a phase offset. Many materials science samples can handle many orders of magnitude higher doses than $500 \text{ e}^-/\text{\AA}^2$, and for these one may wish to optimize the ADF by focusing to the entrance surface while still obtaining a high quality contrast reversal free ptychographic image via the phase offset method. On the other hand, if the dose budget for a given sample is very low, one might consider that one might not obtain useful information from the far less dose efficient ADF signal even with the focus at the entrance surface, and choose to physically focus to the center of the sample. Of course, optimizing the focus under low dose conditions is also very difficult and thus there likely remains benefit to optimizing via post collection adjustments such as the offset method at low doses as well, even if a central focus was the aim.

Since ptychographic contrast is quite sensitive to the sample thickness, we now demonstrate that the offset correction can be successfully applied to a large variety of thicknesses. Fig. 7.6 shows STO SSB images noise free and with a dose of $500 \text{ e}^-/\text{\AA}^2$ at thicknesses of 16 nm, 20 nm, 24 nm and 50 nm. This covers a range of specimen thicknesses typical of atomic resolution electron microscopy in materials science. Focusing

7.3. RESULTS AND DISCUSSION

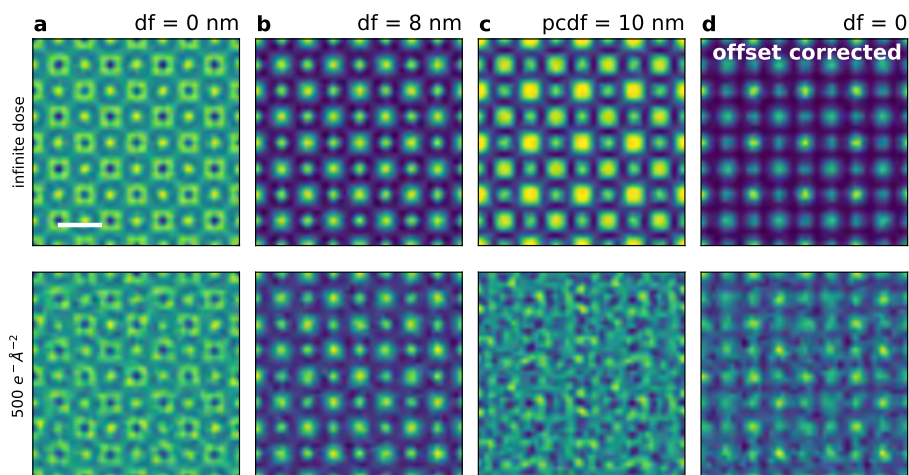


Figure 7.5: Comparison of SSB images simulated for 16 nm thick STO with the probe focused to the entrance surface ($df = 0$), the central slice ($df = 8$ nm) without further correction, and focused to the entrance surface with a post-collection defocus of 10 nm and using the phase offset method. The top row of images is noise free, while the bottom row of images uses a dose of $500 \text{ e}^-/\text{\AA}^2$. Here central focusing is optimal, correcting the reversals with strong contrast. The post collection defocus correction is very noisy in the low dose simulation, but the phase offset method retains sufficient contrast to locate all the columns at low dose while retaining the optimal probe focus for the ADF. Scale bar is 3 \AA .

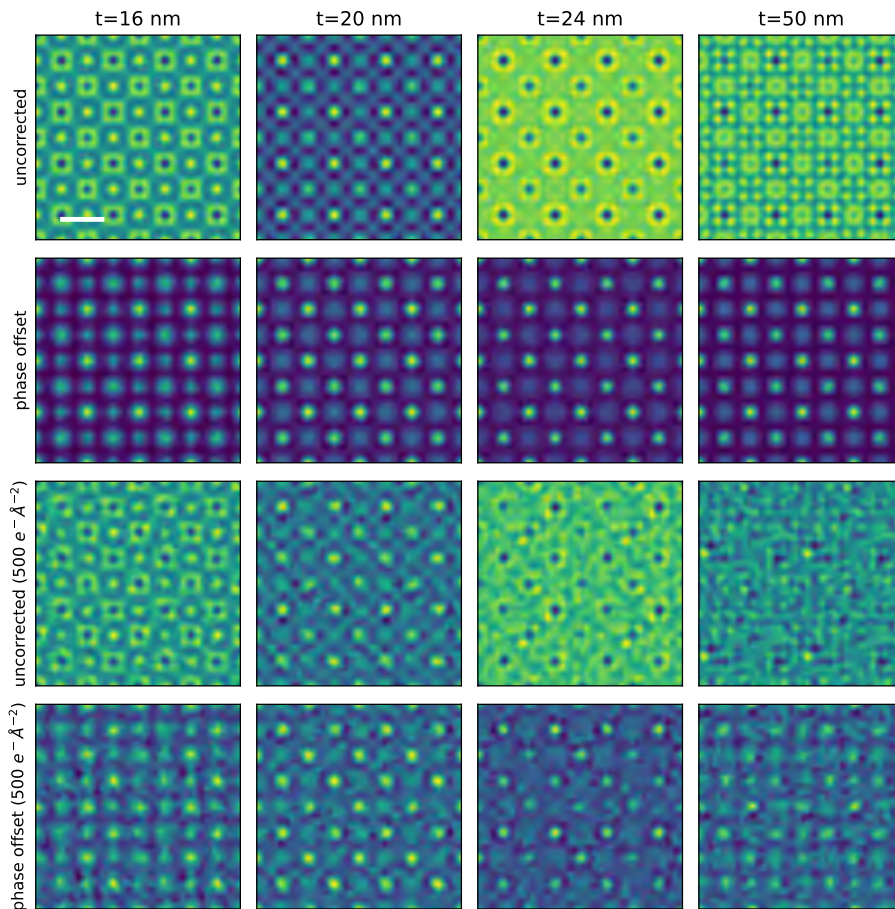


Figure 7.6: Simulated SSB imaging of STO as a function of thickness (t , indicated above each column) comparing the uncorrected and phase offset results with the probe focused to the entrance surface. The top half of the figure is with infinite dose, and the bottom with a dose of $500 \text{ e}^-/\text{\AA}^2$. These results show the robustness of the phase offset method across a wide range of thickness. Scale bar is 3 \AA .

7.3. RESULTS AND DISCUSSION

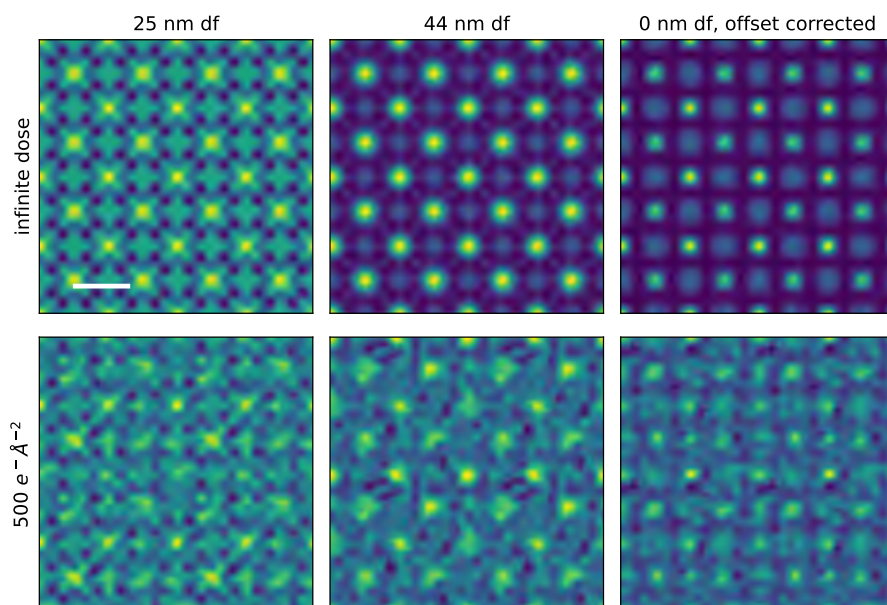


Figure 7.7: Comparison of physical defocus and the phase offset method for contrast reversal removal with 50 nm thick STO simulations. Central focusing (25 nm defocus) does not remove the contrast reversals in this case. Instead, close to the exit surface, using 44 nm of defocus, was found to be optimal using a focal series. However, physically focusing just 6 nm from the exit surface results in additional atom like features appearing where there are no atoms. This does not occur using the phase offset on data taken with the probe focused to the entrance surface. Furthermore the O columns are more visible at the low 500 e⁻/Å² dose using the phase offset method than physical focusing. Scale bar is 3 Å.

to the entrance surface leads to contrast reversals as seen in the first and third rows of Fig. 7.6. The phase offset correction leads to a reasonable contrast with all thicknesses. In the 24 nm case, the oxygen columns have a much weaker contrast in both the uncorrected and the corrected phase images. This can be improved by physically focusing to the center of the sample thickness, as we showed previously [72], however, this can be difficult in experiments in practice without live ptychography.

For 50 nm of STO, the contrast reversals are sufficiently complex in the uncorrected image that it is practically uninterpretable at $500 \text{ e}^-/\text{\AA}^2$. This is a very low dose for STO, but it is nevertheless informative regarding the contrast generally as well for samples that handle only very low doses.

50 nm is also an interesting case because the center of the sample is not the focal plane exhibiting optimal contrast with physical focusing, as we found previously by performing a simulated focal series [58]. Figure 7.7 shows that the central focal plane exhibits quite strong contrast complexity that is not “atom like”. Instead, it was found that physically focusing to near the exit surface provides much better correction of contrast reversals. An example of this is shown in figure 7.7 using a 44 nm defocus from the entrance surface, just 6 nm from the exit surface of the sample. Here the contrast is much better, appearing more atom like but also clearer, including at $500 \text{ e}^-/\text{\AA}^2$. However some artifacts remain in the form of atom like spots in between the actual atoms, as is seen in the noise free image. These are not present in the phase offset images, which not only show no contrast reversals or artifacts but actually show more visible contrast on the O sites at $500 \text{ e}^-/\text{\AA}^2$. Given the difficulty of optimizing focus during low dose work, the performance of the phase offset here is encouraging.

To confirm the theory that the phase information can be transferred to the amplitude domain as shown in equation 7.8, we use the 50 nm STO case shown in figure 7.6 as an example. We can observe that with a phase offset, the phase image in the first row is quite similar with the amplitude figure in the second row. This shows the phase offset method can transfer phase information into amplitude domain, and demonstrate the phase offset theory is correct.

Furthermore, focus adjustment of the beam cannot not always remove contrast reversals. While one might expect that contrast reversals arise only in relatively thick materials, surprisingly thin materials can also exhibit contrast reversals, and these can be impossible to remove with defocus. Clark et al. showed this for very thin GaN [59]. We have explored this for 5 nm thick STO [72], which we find also exhibits contrast reversals which cannot be counteracted with central focusing. This is perhaps in some sense intuitive given the small range of defocus that exists within the sample. However, the reversals also can not be corrected with any focal point within the sample, or even within a useful range beyond as shown in Fig. 7.9. While perhaps the contrast reversal begins to be counteracted far beyond the exit surface, the contrast has reduced to the point that the O sites are almost invisible. However, if we instead apply a phase offset to the entrance surface focused data the contrast reversals are completely removed. This shows that the phase offset method can in some cases work better than any type of defocus adjustment.

Extending the complexity of the system beyond a bulk crystal, we also tested an oxide heterostructure. Here we again used STO but now interfaced with ZrO_2 in the cubic fluorite structure epitaxially lattice matched to the STO. The structure is 16 nm thick in

7.3. RESULTS AND DISCUSSION

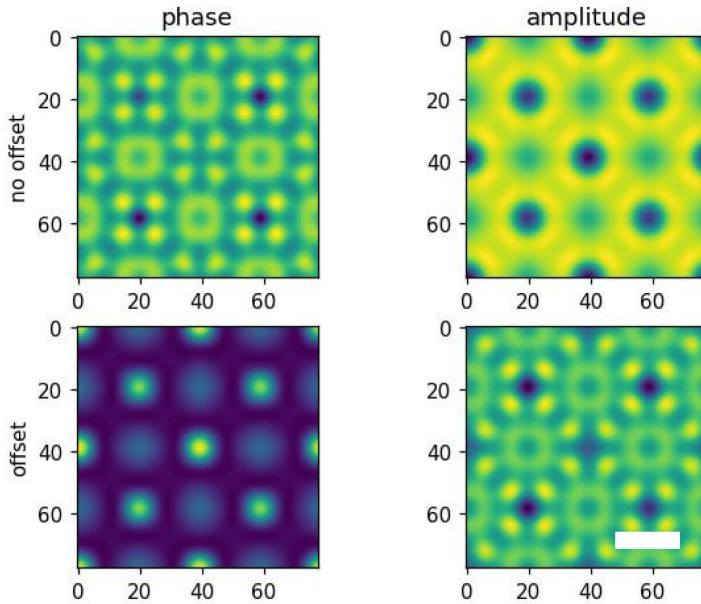


Figure 7.8: An example showing how the phase information can transfer to the amplitude domain with the phase offset method using a 50 nm thick STO simulation. Scale bar is 2 Å.

the beam direction. Interestingly, the uncorrected image in Fig. 7.10 (left) shows strong contrast reversals at all STO sites. The zirconia also shows virtual atoms as indicated by the red arrows. All these artifacts are removed by applying an offset. Therefore the phase offset method appears robust to handling more complex structures than pristine bulk crystal structures. The results show again how such contrast reversal correction can be vital to providing meaningful and interpretable images of atomic structures. We therefore conclude that the phase offset correction seems very useful for ptychography of a wide range of materials.

We now demonstrate the phase offset method with experimental data. 4DSTEM data of a methylammonium(MA)-PbI₃ perovskite was acquired using our Timepix3 event driven camera to easily achieve very low doses and avoid drift [15]. A 13 mrad convergence angle was used to optimize for contrast in the frequency range of interest. Due to the extreme beam sensitivity of the material, we use a dose of just 50 e⁻/Å². We note that this is in the dose regime used in cryo electron microscopy of proteins. Although the event driven camera makes such low dose scans easy to achieve, the very low dose still makes it very difficult to find the best focus during the experiment, especially as one wishes to spend all the dose budget on imaging the regions of interest, not on adjusting the focus. In practice, at present focusing is performed by optimizing the ADF image, which again is often not the optimal defocus for the contrast of the ptychography.

An SSB image of the MAPbI₃ is shown without correction in Fig. 7.11a. The heavy columns, which include Pb, tend to be donut shaped, despite the very low thickness of

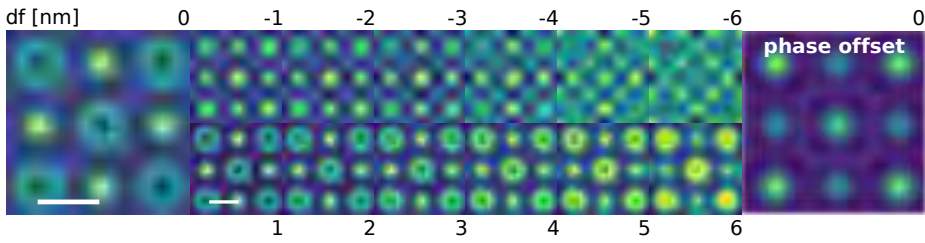


Figure 7.9: Simulated focal series for thin 5 nm thick STO showing that no defocus value can counteract the contrast reversals within a range that does not overly distort the images. Importantly, the phase offset method counteracts the contrast reversals with the probe focused to the entrance surface and retains good contrast. Scale bar is 2 Å.

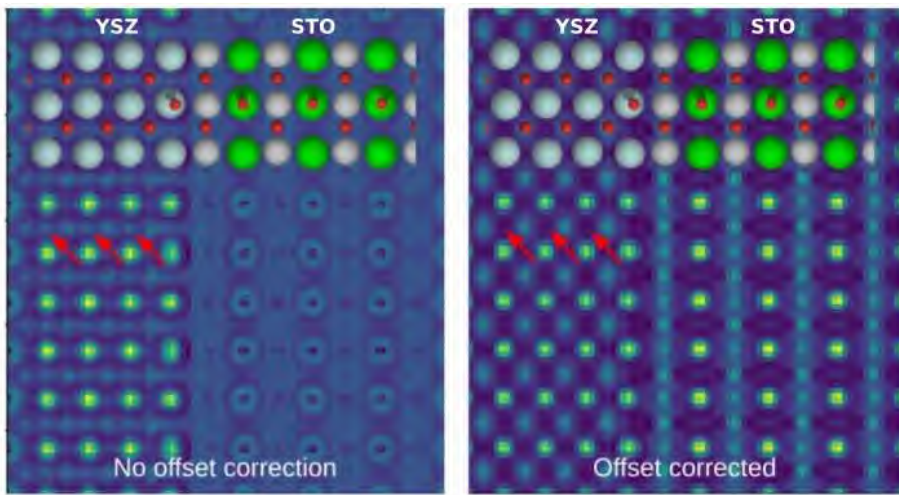


Figure 7.10: Simulated SSB image of a STO/zirconia interface before and after phase offset correction. Contrast reversals as well as the “virtual atoms” indicated by the red arrows are removed in the phase offset corrected image.

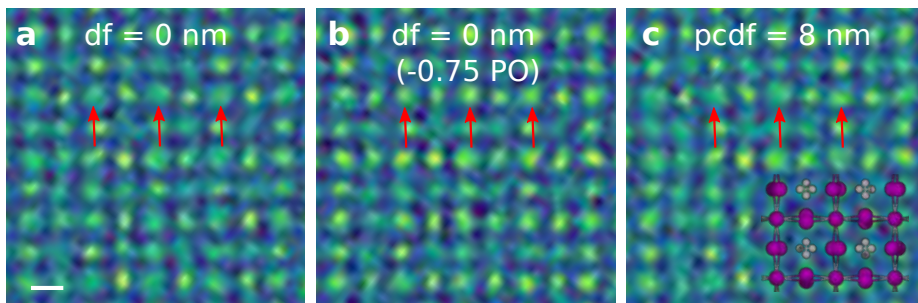


Figure 7.11: Experimental SSB images of thin ca. 2–4 nm thick MAPbI₃. (a) Uncorrected SSB image taken with the focus optimized for the ADF (not shown). Contrast reversals are clearly visible on the heavy columns as highlighted by the red arrows. The phase offset corrected SSB image (b) removes the contrast reversals while the post collection defocus (pcdf) cannot completely remove the contrast reversals. Scale bar is 3 Å.

7.3. RESULTS AND DISCUSSION

approximately 4 nm as indicated by electron energy loss spectroscopy. This thickness falls within the regime where the contrast reversals cannot be corrected by defocusing, assuming the STO results are representative as we expect. Indeed a post acquisition defocus series does not remove the contrast reversals without losing so much contrast as to make the atoms essentially invisible. In Fig. 7.11c shows the result of using an 8 nm post collection defocus which retains sufficient contrast to resolve the atoms but only reduces the contrast reversals rather than removing them. Applying a phase offset, however, completely removes the contrast reversals without any obvious compromise of the contrast, as shown in Fig. 7.11b.

The fact that increasing the post collection defocus reduces the contrast to the point of losing the lattice contrast completely is in agreement with the earlier discussion regarding the thin STO and the defocus series shown in Fig. 7.9, where a high defocus only corrects the reversals to a small extent. For low dose data, such as that of the MAPbI₃, such high defocus values lead to a complete loss of the atomic resolution signal as a result of the contrast reduction associated with defocusing. In this case, the phase offset is the only method that can practically be used to obtain an easily interpretable image without contrast reversals.

Since contrast reversals have also been observed in iterative ptychography reconstructions [59], it is interesting to see if the phase offset can also be used for these methods as one would expect. For this reason, we simulated a 4D data set of MAPbI₃ with a focused probe and processed it with the regularized iterative ptychography (rPIE) algorithm as implemented in abTEM [54]. Donuts appear at the Pb sites, similar to the SSB case, as shown in Fig. 7.12a. Applying the phase offset to the rPIE result indeed removes the contrast reversals as shown in panel b. A 1000 e⁻/Å² dose using a 50 nm defocus version of the phase offset corrected rPIE reconstruction is shown in Fig. 7.12c.

We note that another possible way to avoid contrast reversals due to phase wrapping in reciprocal space in iterative ptychography is to solve for the potential directly rather than the complex object as proposed by [75]. We have tested this with the gradient descent algorithm as implemented in the py4DSTEM code, and find that in the case of the thin MAPbI₃ it does indeed also remove the contrast reversals seen with rPIE as shown in Fig. 7.13. However, we note that the images also appear considerably less sharp.

Optical sectioning makes use of finite depth of field to locate objects in 3D by observing at which focus objects appear most sharp. Optical sectioning with direct ptychography was demonstrated at relatively low resolution by Yang et al. [28], using a single scan and altering the defocus post-collection. This allowed the 3D locations of nanotubes to be discerned, and it was shown that the method provides a true optical sectioning effect rather than a Fresnel propagated version of the exit wave. Defocus away from the plane at which an object is located acts to introduce a variation in the phase in the double overlap regions that diminishes the overall amplitude the more an object is out of focus. This is the reason why using defocus to correct contrast reversals overall tends to result in lower contrast as opposed to using the phase offset which does not introduce phase oscillations that diminish the amplitudes.

For optical sectioning we can therefore use the phase offset without reducing the optical sectioning effect, and by correcting contrast reversals independently of the defocus significantly clean up our view of a sample at different depths, aiding interpretation. Introducing defocus can itself introduce contrast reversals [58], therefore it is valuable

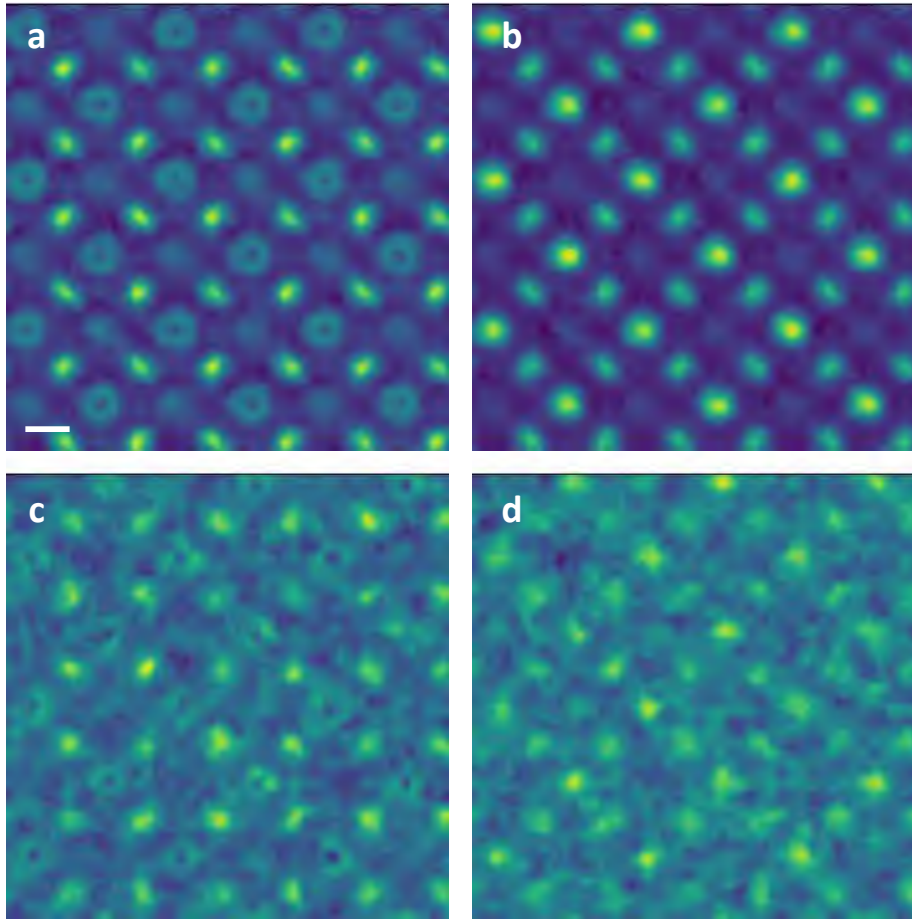


Figure 7.12: Simulated rPIE images of thin 2.4 nm thick MAPbI₃. (a) Noise-free uncorrected rPIE image taken with a focused probe. Donuts are visible on the Pb sites, similar to the SSB image. (b) Phase offset corrected version of (a). (c) is the same as (b) but with an electron dose of 1000 e⁻/Å² and using a defocus of 50 nm. (d) is the phase offset corrected version of (c) using the same offset as (b). The scale bar is 3 Å.

7.4. CONCLUSION

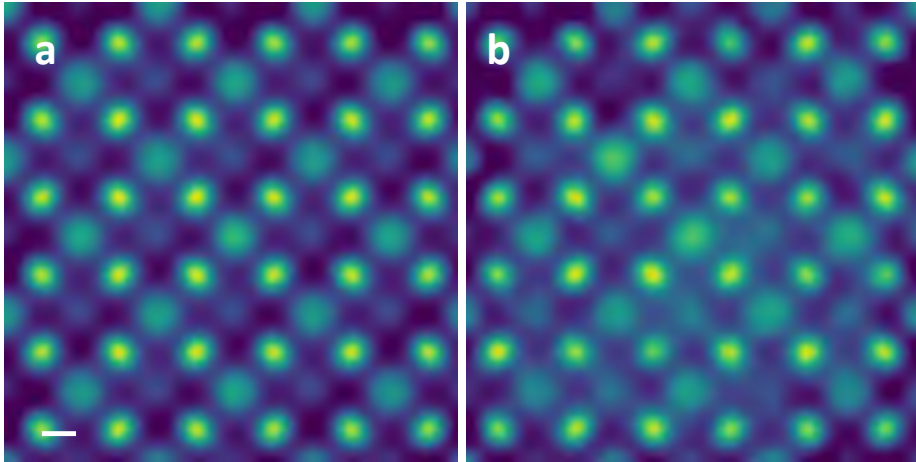


Figure 7.13: Simulated gradient descent iterative ptychography images of 2.4 nm thick MAPbI_3 . (a) Noise-free reconstruction using a focused probe, with no donuts visible on the Pb sites. (b) is the same as (a) but using a dose of $1000 \text{ e}^-/\text{\AA}^2$ no contrast reversals are visible on the reconstruction. Scale bar is 3 \AA .

to be able to counteract these with another independent method such as the phase offset. This is important because interpreting atomic resolution optical sectioning in the presence of contrast reversals can be rather confusing.

To illustrate this we show simulated optical sectioning of 5 nm of graphene placed on top of STO in figure 7.14 using a 30 mrad probe convergence angle. We can optimize the phase offset for viewing either lattice as clearly as possible within the bounds of the optical sectioning effect and otherwise available contrast. Thus for example we can use the phase offset to see the STO go out of focus much more clearly as the probe moves into the graphene out of the STO using offset number one compared to using no offset. Offset number two optimizes the contrast of the graphene layers. The C potential is much weaker than the far heavier Sr and Ti columns which are also much denser than the C columns in the widely spaced van der Waals layers of graphene. Therefore it is harder to see the graphene over the out of focus STO, but offset number two still mitigates the dark dips in phase due to the STO inside the graphene.

7.4 Conclusion

In conclusion, the phase offset method offers a significant boost to our ability to counteract contrast reversals in electron ptychography. Optimizing the focus used with the data acquisition often provides the best ptychographic contrast, such as central focusing with intermediate thicknesses. However there are many situations where using a focus optimal for the ptychographic contrast is not practical, and indeed cases where defocus cannot be used to counteract contrast reversals at all. Often one may prefer to optimize defocus for the ADF, as this cannot be corrected post collection. At doses sufficiently high that the less efficient ADF signal shows good contrast, the contrast reduction from

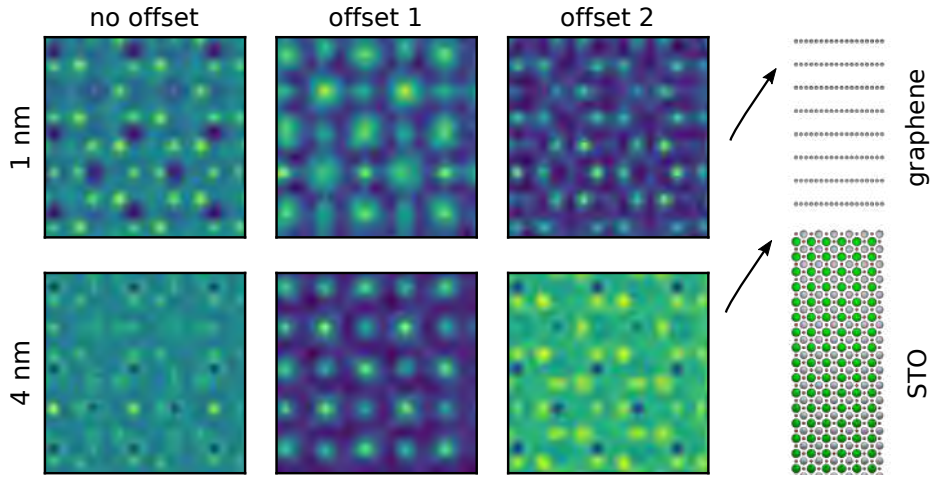


Figure 7.14: Simulated SSB phase images with the probe focused to two different points in a graphene/STO heterostructure. Top: Focal point 1 nm below the top graphene layer. Bottom: Focal point close to the STO interface. In both cases, the uncorrected image shows a mixture of both structures as well as contrast reversals on the heavy sites of the STO. Offset one maximizes the contrast of the STO while offset two maximizes the contrast of the graphene.

using the offset method vs a physical defocus optimised for the ptychography will often not be so significant as to matter for locating atoms. At very low doses where the ADF provides very poor contrast, one may choose to abandon the ADF and prioritize focusing for the ptychography. However, the ADF can provide very informative information at surprisingly low doses, even if exceedingly noisy, and in practice accurate focusing is particularly challenging at extremely low doses. Furthermore, defocus cannot always correct contrast reversals as is the case for the thin few nm thick samples we discussed. Thus, as we see with the experimental example with MAPbI_3 here, the phase offset can be an important tool to remove contrast reversals, even at the extremely low doses used in cryo electron microscopy of proteins. Overall, we find the phase offset method reliably overcomes contrast reversals with minimal contrast reduction, or even improved contrast, compared to a defocus optimized for ptychography at the time of acquisition and thus we expect it to become a standard tool in the use of direct electron ptychography. Finally, we find that the offset method can be a useful tool for optical sectioning with ptychography due to its ability to improve contrast and interpretability independent of the focus.

Conclusion and perspective

8.1 Conclusion

We are currently experiencing an ongoing revolution in electron microscopy with the widespread adoption of DEDs. Scientists have reported a lot of key scientific findings facilitated by DEDs. One particular research domain is electron ptychography, which holds promise for unraveling the intricate structures of highly beam-sensitive materials like bio samples and achieving super-resolution without the limitation of aperture in the condenser lens system. Nevertheless, challenges persist both in experimental setups and algorithmic processes. Issues such as the comparatively sluggish scanning speed of cameras and contrast reversals of the reconstructed phase for relatively thick specimens, disrupting phase or weak phase approximations, remain noteworthy limitations.

This thesis addresses these challenges by the event-driven Timepix3 detector, presenting a viable solution to the speed bottleneck. Moreover, innovative approaches for applying electron ptychography to relatively thick samples, employing a middle focusing strategy, are proposed. This research aims to push the boundaries of electron microscopy, offering solutions to existing limitations and advancing the field towards more efficient and accurate imaging techniques.

Chapter 3 showcases the successful resolution of the mixed layered structures of LaVO_3 and LaTiO_3 through the combined application of HAADF imaging and EELS. Additionally, utilizing the Merlin detector, a comprehensive 4D dataset of SrRuO_3 was acquired. The outcomes of this investigation show the high sensitivity of electron ptychography to light atoms, further emphasizing its efficacy in exploring and characterizing intricate material compositions.

In Chapter 4, the implementation of the event-driven Timepix3 detector is highlighted as a pivotal strategy for overcoming speed limitations. A distinctive advantage of Timepix3, particularly in low-dose conditions, is underscored when compared to frame-based detectors, owing to its event-driven mode. Using examples of a 2D sample, WS_2 , and a bulk sample, zeolite, I demonstrate the remarkable dose efficiency of electron ptychography as opposed to conventional BF, ABF and ADF imaging, including the iCoM method. This chapter underscores the immense potential of electron ptychography in effectively imaging beam-sensitive materials.

In Chapter 5, a noteworthy discovery emerges regarding the STO sample with an estimated thickness of 20 nm. The reconstructed phase through ptychography exhibits normal phase contrast, deviating from the simulation outcomes. Through an extensive series of simulations, it becomes evident that focusing the probe on the middle plane of the crystal model produces results consistent with experimental findings. Subsequent simulations reveal that, for relatively thick samples, adopting a middle focusing approach proves to be the optimal choice. However, it is essential to note that this conclusion may not hold true as the sample thickness increases to 50 nm. This chapter proposes a solution to counteract contrast reversal in relatively thick samples by strategically tuning the focus of the probe.

In Chapter 6, an explanation is provided regarding the overcome of contrast reversal achieved in ptychography when focusing on the middle plane, utilizing the integrated contrast transfer function concept. Without considering the impact of dynamic scattering in the relatively thin samples, this method works for weak phase object while we find it can not well explain the thick sample with strong phase objects. We will explore using Bloch wave theory to explain the behaviours in the future. Additionally, the convergence angle proves to be a factor influencing the optimal defocus condition.

In Chapter 7, we delve deeper into strategies for mitigating contrast reversal issues. With the “fake” single atom models with strong potential, our research reveals that the occurrence of phase wrap asymptotes within the frequency domain is a primary cause of contrast reversal in SSB ptychography. By adjusting these asymptotes to higher frequencies, we can effectively neutralize the contrast reversal effect. This approach has been rigorously tested through both simulation and experimental data, demonstrating its superiority over the application of post collection defocus correction.

8.2 Perspective

Electron ptychography is a dynamic and rapidly advancing field, marked by continuous research and evolving developments. The growing recognition of its immense potential within the electron microscopy community is evident; however, the intricate nature of the algorithms has limited its widespread adoption among scientists unfamiliar with these complexities. Despite the publication of some open-source codes, their adaptation for general research remains a challenge. Furthermore, electron ptychography has yet to be established as a standard method in electron microscopes due to practical issues. Notably, processing large datasets tends to be time-consuming, posing a significant hurdle for the processing speed of computer systems installed on electron microscopes. While some attempts at live reconstruction have been reported, widespread application remains limited. The further works in this field:

1. High speed ptychography: Achieving high-speed ptychographic imaging necessitates a camera with a sufficiently fast export of data. Timepix3 camera is fast enough in low dose condition, while the bandwidth limits its application on large dose condition. Besides, real-time or near-real-time imaging additionally requires the adaptation of algorithms for swift processing.
2. Low dose imaging: In the low-dose imaging, the presence of noise in ptychographic

8.2. PERSPECTIVE

images can impede their interpretation. Developing robust algorithms to address noise is critical, particularly for beam-sensitive materials.

3. Machine learning applications: Recognizing the transformative role of machine learning across various scientific domains, its application in assisting the reconstruction and analysis of ptychographic data becomes essential. Implementing machine learning algorithms has the potential to expedite processes and enhance the quality of results.

In conclusion, electron ptychography stands as a cutting-edge imaging technique with widespread applications across scientific disciplines. The realization of its full potential hinges on ongoing advancements in instrumentation, methodologies, and collaborative interdisciplinary efforts, paving the way for groundbreaking discoveries in the microscopic realm.

Appendix **A**

Symbols

TEM	Transmission electron microscopy
STEM	Scanning Transmission Electron Microscopy
HAADF	High Angle Annular Dark Field Imaging
DED	Direct Electron Detectors
ABF	Annular Bright Field imaging
BF	Bright Field imaging
iCoM	Integrated Center of Mass
SSB	Single Side Band
WDD	Wigner Distribution Deconvolution
ePIE	Extended Ptychographical Iterative Engine
CTF	Contrast Transfer Function
SEM	Scanning Electron Microscopy
FEG	Field Emission Guns
EELS	Electron Energy Loss Spectroscopy
WPOA	Weak Phase Object Approximation
CBED	Convergent Beam Electron Diffraction
SAD	Selected Area Diffraction
CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
FFT	Fast Fourier Transforms
SVD	Singular value decomposition
CDI	Coherent Diffraction Imaging
FIB	Focused Ion Beam

B.1 Publications

1. D. Jannis, C. Hofer, C. Gao, X.B. Xie, A. Béch , T.J. Pennycook, J. Verbeeck, *Event driven 4D STEM acquisition with a Timepix3 detector: Microsecond dwell time and faster scans for high precision and low dose applications*. *Ultramicroscopy*. 233, 113423 (2021);
2. C. Gao, C. Hofer, T.J. Pennycook. *Overcoming contrast reversals in focused probe ptychography of thick materials: an optimal pipeline for efficiently determining local atomic structure in materials science*. *Appl. Phys. Lett.* 121, 081906 (2022);
3. C. Gao, C. Hofer, T.J. Pennycook. *On central focusing for contrast optimization in direct electron ptychography of thick samples*. *Ultramicroscopy* 256, 113879 (2024);
4. C. Hofer, C. Gao, T.J. Pennycook. *Phase offset method of ptychographic contrast reversal correction*. *Ultramicroscopy*. 258, 113922 (2024);

B.2 Conferences

1. C. Gao, C. Hofer, T.j. Pennycook. *Removing Thickness Induced Contrast Reversals in Direct Electron Ptychography*. PICO Meeting, 8-12 May 2022. *Poster presentation*
2. C. Gao, C. Hofer, T.j. Pennycook. *Addressing Thickness Induced Contrast Reversals in Focused Probe Ptychography*. *Microscopy and Microanalysis*, 31 July-4 August, 2022. *Poster presentation*
3. C. Gao, C. Hofer, T.j. Pennycook. *Optimizing low dose structural imaging for materials science with fast focused probe ptychography and contrast reversal correction*. EUROMAT 2023, 3 September-7 September. *Oral presentation*

List of Figures

1.1	Schematic of essential elements of STEM	3
1.2	The aperture function in reciprocal space and the corresponding probe in real space in the ideal condition. The scale bar of the left figure is 20 mrad, the scale bar in the right figure is 2 Å.	6
1.3	Schematic of geometric aberrations for the condenser lens. (1) The dashed lines indicate the electron trajectories for an ideal lens, the solid lines are the electrons trajectories for an imperfect lens with aberrations. (2) The definition of the aberration function. Dashed lines indicate the ideal spherical wave front for the convergent electron beam, the solid lines indicate the real wave front due to geometric aberrations, where δ is the distance between the ideal and real electron wave front.	7
1.4	Aberration coefficients with different notations.	8
1.5	Probe shapes with different single aberrations. The voltage is 200 kV and the convergence angle is 20 mrad. (1) Aberration free. (2) Coma with $C_{21} = 1000$ nm. (3) defocus=10 nm. The scale bar in left figure is 20 mrad and the scale bar in the right figure is 2 Å. (4) $C_{34a} = 50000$ nm.	9
1.6	Interaction of high energy electrons with a thin specimen.	9
1.7	The two mechanisms of the isolated atom scattering incident electrons. Coulombic interaction within the electron cloud results in low angle scattering. Coulombic interaction with the nucleus results in high angle scattering.	10
1.8	Schematic EELS with different excitation mechanisms.	11
1.9	Multislice decomposition of a thick sample. The sample is sliced into multiple slices, with each slice is approximated as its projected potential. The incident wave function at each slice can be viewed as the convolution of the transmission function and the propagator function.	15
1.10	Schematic of Bragg's law. A plane wave with the wavelength of λ incident on atomic planes with a spacing of d at an angle of θ . The path difference of the scattered plane waves are $AB+BC$	17

1.11	Imaging and diffraction modes in TEM. In figure A, The projector lens projects the diffraction patterns onto the screens. In figure B, the projector lens projects the image onto the screens. This figure is adapted from reference [10].	18
1.12	Schematic diagram of CBED. The condenser and objective lenses focus the probe onto the specimen with a very tiny area. In the back focal plane, it would form diffracted discs.	20
1.13	The diagram shows how the convergence angle influences the CBED patterns. In the case transmitted and diffracted discs are not overlapped, it is call a K-M mode. If the discs are overlapped, it is called a Kossel pattern. This figure is adapted from reference [10].	21
1.14	Schematic diagram of 4D STEM. (1) The convergent probe raster scans across the specimen while a pixelated camera simultaneously records the diffraction patterns in reciprocal space. (2) The 4D datasets in the 4D coordinates. To index a diffraction pattern, the scan position (x, y) and diffraction pattern in the detector (k_x, k_y) have to be known.	22
1.15	Counting modes of the latest Merlin Medipix3 camera.	23
1.16	The principle of Timepix3. (1) An example of a sparse diffraction pattern acquired with the Timepix3 camera. (2) Illustration of how the Timepix3 reads out a stream of packets registering each hit with a timestamp and position on the camera.	24
1.17	Comparison of different kinds of direct electron detector.	24
2.1	The mathematical description of wave function in the diffraction plane in theory and diffraction intensity detected in the experiment. In theory, the diffraction wave function is the convolution between probe function $P(\mathbf{R}_p, \mathbf{r})$ and object function $O(\mathbf{R}_p, \mathbf{r})$, resulting in a complex valued function. While in the experiment, only the intensity of the diffraction patterns is recorded, providing only real images rather than the full complex values. A red circle denotes the transmitted direct beam in the left figure.	27
2.2	Comparison of normalized elemental potential intensity and reconstructed ptychographic phase intensity. (1) the single atom model that is used for 4D dataset simulation. (2) an example of the reconstructed ptychographic images. (3) The normalized projected potential strength versus the ptychographic phase.	28
2.3	The coordinates definition in the ptychography setup.	28
2.4	Simulated 4D dataset of graphene showing how the intensities of the diffraction patterns vary with respect to probe position.	29

LIST OF FIGURES

2.5 Schematic of $G(\mathbf{k}_f, \mathbf{Q}_p)$ at the spatial frequency \mathbf{Q}_p . The left circle is corresponds to frequency $-\mathbf{Q}_p$, the middle circle to 0 frequency, and the right circle to frequency \mathbf{Q}_p . In the left double overlapped area (red area), $\Psi_s^*(-\mathbf{Q}_p)$. In the right double area (blue area) $\Psi_s(\mathbf{Q}_p)$ is transferred. In the triple overlap area, $\Psi_s(\mathbf{Q}_p)$ and $\Psi_s^*(\mathbf{Q}_p)$ destructively interfere and the phase of them cancels out. This figure is reproduced from reference [21] . 31

2.6 Amplitude and phase of $G(\mathbf{K}_f, \mathbf{Q}_p)$ of different spatial frequency from the graphene simulation. 32

2.7 An example of a trotter from a single spatial frequency. The green circle indicates the location of the BF disk, and the red and blue circles, the locations of the shifted diffracted disks. Double disk overlap occurs where the red and blue disks overlap just the green disk. Triple overlap occurs when the red and blue disks also overlap each other and the amplitude can be seen to zero in this region. 33

2.8 $\Psi_s(\mathbf{Q}_p)$ in reciprocal space. 33

2.9 The final reconstruction result showing the graphene lattice. 34

2.10 Schematic of the useful coherent area with different conditions. This figure is adapted from reference [22]. 35

2.11 CTF of SSB Ptychography. 36

2.12 The trotter at one specific spatial frequency \mathbf{Q}_p . The top row shows the modulus and phase information without any aberrations. The phase is zero in the double overlap area. The bottom row shows the modulus and phase information with a defocus of 2 nm. The phase now varies with the scattering angle. 37

2.13 Reconstructed SSB phase images of graphene for the simulated 4D datasets with and without aberrations. The left phase image is reconstructed from the simulated 4D dataset with $C_{32}=250$ nm. We can clearly see the influence of the 3 fold astigmatism in the phase image. The right phase image is the equivalent aberration free image. The scale bar is 2 Å. 37

2.14 Workflow of ePIE algorithm. 39

3.1 Schematic diagram of the LaVO_3 and LaTiO_3 layered structures. 44

3.2 HAADF image of $\text{LaVO}_3/\text{LaTiO}_3$. The carbon and Pt layers are deposited on the sample to protect the sample during preparation by FIB. The white area in the image is the contamination of carbon. 45

3.3	Quantitative analysis of $\text{LaVO}_3/\text{LaTiO}_3$. (a) Atomic resolution HAADF image of the sample. (b) Enlargement of the area in the yellow rectangle in (a). Overlaid are orange disks to represent the La columns with atomic number 57, and blue disks representing Ti or V atoms, the corresponding atomic numbers of which are 22 and 23. (c) Line profile of atomic columns along the red arrow overlaid on the image in (b). The La atomic columns can easily be distinguished from the Ti and V columns, but it is much harder to distinguish between the Ti and V atomic columns.	46
3.4	Spectrum imaging. (a) HAADF image showing the region used for spectrum imaging in a red rectangle. (b) Integrating the core loss spectra over the region inside the red rectangle area in (a) clearly shows the Ti L edge and V L edge sufficiently separated for elemental analysis.	47
3.5	EELS line scan over the LaVO_3 and LaTiO_3 . (a) survey ADF image of the sample, the red line labels the scan trace of the probe. (b) Image of the line scan core loss spectrum image obtained from the trace as shown in the red arrow in (a). The x axis is energy loss. The y axis is the distance from the scan start point to the end point. From the figure, we can see the L edge of Ti and V and the K edge of the O clearly. (c) Plot of the line trace integrated across the EELS camera. From the figure, we can see the intensity of the Ti L edge and V L edge oscillate and the peaks of each two elements are perfectly out of phase, as expected.	48
3.6	2D spectrum imaging of $\text{LaVO}_3/\text{LaTiO}_3$. (a) HAADF survey image. The spectrum image of V (b), O (c), and Ti (d). From figure (a), (b), (c) and (d), we can distinguish the Ti and V elements with atomic resolution, but the sample drift is significant.	48
4.1	Diagram illustrating the output distinctions between frame-based and event-driven detectors. In frame-based detectors, the output manifests as a 2D array, regardless of if most of the pixels contain only zero counts, whereas for event-driven detectors, the output takes the form of a list of numbers detailing just the position on the detector and the time of each hit in nanoseconds.	52
4.2	Raw data collected with the Timepix3 camera and the 4D dataset as converted by the python script.	53
4.3	Example of a single CBED pattern from the Timepix3 at the dose of $6000e^-/\text{\AA}^2$ (left). The image on the right is the position averaged CBED pattern from the same dataset.	53
4.4	BF, ABF and ADF images reconstructed from the collected 4D dataset. In each image, the corresponding FFT is shown in an inset. Streaks appear in the BF and ABF images due to temporary saturation of the timepix3 data stream due to the current used. Scale bar is 2 nm.	54
4.5	The double disk overlaps, also known as the trotters colloquially, for different spatial frequencies.	54

LIST OF FIGURES

4.6 The phase information of WS_2 reconstructed from the 4D dataset. The FFT image is inserted in the phase image. The red circles labels the third ring of spots. The scale bar is 2 nm. 55

4.7 The reconstructed BF, ABF and ADF images from a silicalite-1 zeolite. The scale bar is 1 nm. 56

4.8 The reconstructed iCoM and SSB images of the silicalite-1 zeolite. The scale bar is 1 nm. 57

4.9 Aberration Removal Process: On the left is the experimental double disk overlap phase image containing residential aberrations. In the middle, the phases computed by the aberration coefficients calculated using the SVD method. On the right is the result obtained by subtracting the calculated phases due to aberrations from the experimental phases. 58

4.10 Comparison of normal SSB and aberration corrected SSB ptychography. The scale bar is 2 nm. 59

5.1 ADF and Medipix3 based 4D STEM data from $SrTiO_3$. iCoM and ptychography show the O columns more clearly than the ABF. However this data shows some drift from the slow scan imposed by the camera. Sr, Ti and O columns are indicated by orange, blue and red respectively in the overlays. The scale bar shows 2 Å and the colorbars are in radians. 64

5.2 ADF and 4D STEM ABF, iCoM and SSB ptychography images of Fe_3O_4 from the same 2 μs scan enabled by a Timepix3 camera. Despite the low dose, the SSB image clearly shows the oxygen columns. The crystallinity and absence of distortions allowed the full image to be divided into 4 parts and aligned, based on the SSB signal, and averaged to produce the higher signal-to-noise ratio images shown. O and Fe atoms are shown in red and brown respectively in the model. The projection vector of the plane is (2,3,1) and the colorbar for the SSB is in radians. 65

5.3 ADF, SSB and WDD images simulated for 16 nm thick STO. At zero defocus complex contrast reversals appear in the SSB and WDD images, but when the probe is focused 8 nm into the material the contrast reversals are removed. Importantly however the contrast reversals can also be removed by inserting defocus post collection, meaning the experiment can proceed with the focus optimized for the ADF. The scale bar indicates 2 Å and the colorbar is in radians. 67

5.4	Comparison of the position averaged convergent beam electron diffraction (PACBED) [52] from the experimental data used for Figure 5.1 of the main text with PACBED simulated for different STO sample thicknesses using the same 200 kV accelerating voltage and 20 mrad convergence angle used in the experiment. While no single thickness provides a perfect match, it is clear the experimental PACBED is consistent with the sample tens of nanometers thick. Looking at the symmetries present, we estimate the experiment was performed in a region that is a least approximately 20 nm thick. The scale bar is 2 Å.	70
5.5	Simulated SSB and WDD reconstructed phase results for 16 nm thick STO with varying probe defocus.	71
5.6	Simulated SSB and WDD reconstructed phase results for 20 nm thick STO with varying probe defocus.	71
5.7	Simulated SSB and WDD reconstructed phase results for 28 nm thick STO with varying probe defocus.	72
5.8	Simulated SSB and WDD reconstructed phase results for 50 nm thick STO with varying probe defocus.	72
5.9	Simulated post collection SSB focal series for 50 nm thick STO, demonstrating that the contrast reversals can also be removed for this thickness when the probe is focused to the entrance surface during the scan with post collection defocus adjustment.	73
5.10	ePIE and SSB results computed using a single core of an i7-10700K operating at 3.8 GHz using data simulated for a single unit cell of STO at 200 KV, with a 20 mrad convergence angle, zero defocus and a 59 by 59 probe position grid. The 47 by 47 pixels of the bright field disk are used for bright field ePIE, while the 137 by 137 pixels of the full field CBED patterns are used for the full field ePIE. The SSB image was computed in a total time of 2.4 s. 100 iterations were used in computing the bright field ePIE result taking a total time of 347 s, for an average of 3.47 seconds per iteration. The full field ePIE result and the error per iteration are also provided. The PACBED of the STO dataset is also shown with a log scale.	73
5.11	Full field single unit cell thick STO ePIE reconstruction as a function of iteration number at different doses. At $2 \times 10^5 e^- / \text{\AA}^2$, after 10 to 20 iterations very little apparent change occurs in the phase and probe. At $1.5 \times 10^5 e^- / \text{\AA}^2$ some artefacts are visible at 20 iterations, but disappear between iterations 30 to 40. At $1.2 \times 10^5 e^- / \text{\AA}^2$ the algorithm does not converge even after 100 iterations, reflecting the fact that the convergence of ePIE depends critically on the amount of signal in the CBED patterns. We note that using a significantly defocused probe increases the signal per CBED pattern, but a strongly defocused probe precludes useful simultaneous Z-contrast ADF imaging.	74

LIST OF FIGURES

5.12 SSB reconstructions at the same doses as the ePIE results shown in Figure S8. As a direct non-iterative method, SSB ptychography does not depend on the dose to converge. Of course, the clarity of the resulting SSB images do depend on the dose, but it is a consistent and immediate result, and the above doses pose no problem at all for the SSB method. Scale bar is 2 Å. 75

6.1 Schematic defining the z-coordinates. 80

6.2 SSB ptychography images of 24 nm thick (100) STO simulated noise free with the probe focused to different depths of the sample from the entrance surface (0 nm of defocus) to the exit surface (24 nm of defocus). Overlaid on the bottom of the zero defocus image is a contrast boosted version to facilitate seeing the contrast reversals. A half unit cell model is also overlaid on the bottom of the 12 nm defocus image to aid identification of the columns with Ti shown in orange, O in green columns and Sr in purple. The scale bar is 2 Å. 85

6.3 Probe focal point dependence (with zero defocus (df) defined here as the entrance surface of the sample) of the contrast of the SSB, iCoM, ABF, and ADF signals simulated at a dose of 2000 e⁻/Å² for 24 nm thick STO. . . . 85

6.4 U phase images with different defocus values. Negative values denote overfocus condition. Positive values denote underfocus condition. The scale bar is Å. 86

6.5 Amplitude and phase images of a single U atom. The first two columns show the images from -12 and +12 nm of defocus. The third column shows the reconstruction resulting from summing the two object functions from the two defocus values in reciprocal space. We can see significant artefacts remain and the cancellation of defocus does not occur as would have been expected using the iCTF concept. The scale bar is Å. 87

6.6 SSB images of very thin 5 nm thick STO as a function of physical probe focus within the sample depth (zero defocus (df) is again the entrance surface) as a function of the convergence angle. As the convergence angle increases such that depth of field is on par or shorter than the thickness the contrast reversals are overall diminished or absent and central focusing again provides the best contrast. 88

6.7 SSB images of 12 nm thick STO as a function of physical probe focus within the sample depth (zero defocus (df) is again the entrance surface) as a function of the convergence angle. For this thickness the central focusing again seems to work regardless of the convergence angle. 88

6.8	SSB images of 1 to 60 unit cells of STO using central focusing with 15, 20, 30 and 40 mrad convergence angles. Contrast reversals appear for the smaller convergence angles around 15 unit cells of thickness. The 15 mrad images continue to show a dip at the center of the heavier columns to quite high numbers of unit cells before disappearing, but contrast reversals do not appear on the atomic columns in this range of thickness for the higher convergence angles.	89
6.9	Simulated iCoM reconstructed phase results for 16 nm thick STO with varying probe defocus at infinite dose. The voltage is 200 kV, convergence angle is 20 mrad, scan step size is 0.2 Å. As we can see, when the probe focuses around the central plane, the phase images gives the best contrast.	91
6.10	Simulated iCoM reconstructed phase results for 20 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.	91
6.11	Simulated iCoM results for 24 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.	92
6.12	Simulated iCoM reconstructed phase results for 28 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast.	92
6.13	Simulated iCoM reconstructed phase results for 50 nm thick STO with varying probe defocus at infinite dose, 200 kV, a 20 mrad convergence angle, and a scan step size of 0.2 Å. Central focusing gives the best contrast. At this thickness central focusing is no longer optimal. Instead focusing the probe near the exit surface of the sample (defocus values from 48 nm to 54 nm) gives the best contrast.	93
6.14	ADF, ABF, iCoM and SSB results simulated at the dose of $100 \text{ e}^-/\text{Å}^2$ for 24 nm thick STO, with the other simulation parameters the same as for figure 2 of the main text. From the figure, we can see that when the dose is extremely low, no structure is visible in either the ADF or ABF images at any defocus conditions. However with iCoM and SSB, central focusing gives sufficient signal to noise to resolve all the atomic columns.	94
6.15	iCoM version of figure 6.6 of the main text, showing simulate infinite dose images of 5nm thick STO using different defocus values, covering the entire thickness of the sample, with different convergence angles. Conditions match those of figure 7.	94
6.16	iCoM version of figure 9 of the main text showing simulated infinite dose images of STO with different thicknesses of STO and different convergence angles. Similar with the SSB results, when the convergence angle is 15 or 20 mrad, contrast reversals appear around the 5 nm thickness. With the increasing of convergence angle, the contrast reversals also disappear with the iCoM.	95

LIST OF FIGURES

- 7.1 Single atom SSB simulations using potentials ranging from a U atom potential to 20 times that potential, showing how the potential strength itself causes contrast reversals, which manifest here as donut shaped atom contrast. The 2nd row shows the phase vs frequency in 2D, and the 3rd row line profiles of the rotationally symmetric phase response. As the potential increases the curvature of the phase increases, the phase hits the top limit and wraps around resulting in contrast reversals at the center of the atom. As the wrap around shifts to lower frequencies the donut hole expands. Scale bar is 2 Å. 101
- 7.2 Illustration of the effects of the different frequency ranges on the simulated SSB image of the U×20 potential using masking in Fourier space. With the full range of frequencies out to 2α included (a), the atom appears as a thin ring with a slight peak in its center. As we progressively mask out the frequencies after each phase wrap around asymptote (b–d), the positive ring of phase progressively fills more of the central region of the atom, until after removing the contributions from all frequencies above the first wrap around it becomes atom like again, with a peak at the center of the atom. Of course, by limiting the contribution to lower spatial frequencies the image is also limited in resolution. Scale bar is 1 Å for the phase image and 20 mrad for the phase and amplitude vs. frequency plot, respectively. 103
- 7.3 Illustration of the influence of post collection and physical defocus on SSB ptychography with simulations using the U×20 potential. The SSB images are displayed on the same intensity scale, with the contrast ratio (C) of the maximum phase to background phase indicated. The full frequency response is displayed in the 2D phase vs frequency plots. From the line profiles it is apparent that the physical defocus brings the amplitudes close to zero after the first wrap around, whereas the post collection defocus pushes the first wrap point out to higher spatial frequencies, resulting in stronger contrast in this case. 104
- 7.4 Illustration of the use of a phase offset on strong phase objects. (a–d) SSB images simulated with the U×10 potential using phase offsets of 0.0, -0.9, -1.8 and -2.7 mrad on the nonzero spatial frequencies with the DC values set to zero. (e) Line profiles of the phases with no offset and the optimal -1.8 rad offset, again with only the upper phase vs frequency curve shown for simplicity. 106
- 7.5 Comparison of SSB images simulated for 16 nm thick STO with the probe focused to the entrance surface ($df = 0$), the central slice ($df = 8$ nm) without further correction, and focused to the entrance surface with a post-collection defocus of 10 nm and using the phase offset method. The top row of images is noise free, while the bottom row of images uses a dose of $500 e^-/\text{Å}^2$. Here central focusing is optimal, correcting the reversals with strong contrast. The post collection defocus correction is very noisy in the low dose simulation, but the phase offset method retains sufficient contrast to locate all the columns at low dose while retaining the optimal probe focus for the ADF. Scale bar is 3 Å. 107

- 7.6 Simulated SSB imaging of STO as a function of thickness (t , indicated above each column) comparing the uncorrected and phase offset results with the probe focused to the entrance surface. The top half of the figure is with infinite dose, and the bottom with a dose of $500 \text{ e}^-/\text{\AA}^2$. These results show the robustness of the phase offset method across a wide range of thickness. Scale bar is 3 \AA 108
- 7.7 Comparison of physical defocus and the phase offset method for contrast reversal removal with 50 nm thick STO simulations. Central focusing (25 nm defocus) does not remove the contrast reversals in this case. Instead, close to the exit surface, using 44 nm of defocus, was found to be optimal using a focal series. However, physically focusing just 6 nm from the exit surface results in additional atom like features appearing where there are no atoms. This does not occur using the phase offset on data taken with the probe focused to the entrance surface. Furthermore the O columns are more visible at the low $500 \text{ e}^-/\text{\AA}^2$ dose using the phase offset method than physical focusing. Scale bar is 3 \AA 109
- 7.8 An example showing how the phase information can transfer to the amplitude domain with the phase offset method using a 50 nm thick STO simulation. Scale bar is 2 \AA 111
- 7.9 Simulated focal series for thin 5 nm thick STO showing that no defocus value can counteract the contrast reversals within a range that does not overly distort the images. Importantly, the phase offset method counteracts the contrast reversals with the probe focused to the entrance surface and retains good contrast. Scale bar is 2 \AA 112
- 7.10 Simulated SSB image of a STO/zirconia interface before and after phase offset correction. Contrast reversals as well as the "virtual atoms" indicated by the red arrows are removed in the phase offset corrected image. . 112
- 7.11 Experimental SSB images of thin ca. 2–4 nm thick MAPbI_3 . (a) Uncorrected SSB image taken with the focus optimized for the ADF (not shown). Contrast reversals are clearly visible on the heavy columns as highlighted by the red arrows. The phase offset corrected SSB image (b) removes the contrast reversals while the post collection defocus (pcdf) cannot completely remove the contrast reversals. Scale bar is 3 \AA 112
- 7.12 Simulated rPIE images of thin 2.4 nm thick MAPbI_3 . (a) Noise-free uncorrected rPIE image taken with a focused probe. Donuts are visible on the Pb sites, similar to the SSB image. (b) Phase offset corrected version of (a). (c) is the same as (b) but with an electron dose of $1000 \text{ e}^-/\text{\AA}^2$ and using a defocus of 50 nm. (d) is the phase offset corrected version of (c) using the same offset as (b). The scale bar is 3 \AA 114

LIST OF FIGURES

- 7.13 Simulated gradient descent iterative ptychography images of 2.4 nm thick MAPbI₃. (a) Noise-free reconstruction using a focused probe, with no donuts visible on the Pb sites. (b) is the same as (a) but using a dose of 1000 e⁻/Å² no contrast reversals are visible on the reconstruction. Scale bar is 3 Å. 115
- 7.14 Simulated SSB phase images with the probe focused to two different points in a graphene/STO heterostructure. Top: Focal point 1 nm below the top graphene layer. Bottom: Focal point close to the STO interface. In both cases, the uncorrected image shows a mixture of both structures as well as contrast reversals on the heavy sites of the STO. Offset one maximizes the contrast of the STO while offset two maximizes the contrast of the graphene. 116

LIST OF FIGURES

Acknowledgement

My PhD life is nearing the end. This is an exciting journey blended with some kind of bitter feelings. Being the first to discover something new in science is the most exciting thing in the world. However the path to the hall of science is always full of thorns. When I come to review this journey, I am really grateful for all of these successes and setbacks. They have shaped me into someone more resilient and prepared to tackle future challenges. Thanks to the people who assisted and supported me both in life and work on this journey.

Here I would like to express my deepest gratitude to my supervisor Prof. Timothy. J. Pennycook for his unwavering support, insightful feedback, and invaluable guidance throughout my PhD life. Thanks Tim for giving me the opportunity to study electron microscopy in EMAT. Here I not only learned the advanced technique of electron ptychography but also made some friends from all over the world.

I am also thankful to the members of my thesis committee, Prof. Dr. Sofie Cambré, Prof. Dr. Jan De Beenhouwer, Prof. Dr. Peter Nellist and Prof. Dr. Maria Varela for their constructive critiques and encouragement.

I am thankful to current and past colleagues and friends in the EMAT: Dr. Armand. Beche, Dr. Da Wang, Dr. Daen Jannis, Dr. Pei Liu, Dr. Chu-ping Yu, Dr. Zezhong Zhang, Dr. Min Tang, Qiongyang Chen, Qiang Lu. I also would like to thank my group members: Dr. Christopher Hofer, Dr. Xiaobin Xie, Dr. Biao Yuan, Tamazouzt Chennit, Songge Li.

From my personal perspective, I would like to dedicate my special thanks to my father Lizhi Gao and mother Junying Shen. My parents gave birth to me and my twin brother Yuan Gao in a modest village in the He Bei province, North of China. Despite their lack of university education, my parents were convinced that their children would pursue higher education in the future. This belief has been motivating them to work hard to support me and my brother for better education. Therefore this degree is not just my achievement but a tribute to my parents' sacrifices and belief in education. Thank you for everything.

APPENDIX C. ACKNOWLEDGEMENT

Bibliography

- [1] W. E. (William E.) Moerner. Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy (nobel lecture). *ANGEWANDTE CHEMIE-INTERNATIONAL EDITION*, 54(28):8067–8093, JUL 6 2015. 1
- [2] W. E. (William E.) Moerner. Nobel lecture: Single-molecule spectroscopy, imaging, and photocontrol: Foundations for super-resolution microscopy. *REVIEWS OF MODERN PHYSICS*, 87(4), OCT 21 2015. 1
- [3] L. de Broglie. *Researches on the quantum theory*. PhD thesis, University of Paris, Paris, 1925. 1
- [4] Ernst Ruska. The development of the electron microscope and of electron microscopy. *Rev. Mod. Phys.*, 59:627–638, Jul 1987. 2
- [5] D. McMullan. SEM—past, present and future. *Journal of Microscopy*, 155, 1989. 2
- [6] Max Haider, Harald Rose, Stephan Uhlemann, Eugen Schwan, Bernd Kabius, and Knut Urban. A spherical-aberration-corrected 200kv transmission electron microscope. *Ultramicroscopy*, 75(1):53–60, 1998. 6
- [7] O.L. Krivanek, N. Dellby, and A.R. Lupini. Towards sub-Å electron beams. *Ultramicroscopy*, 78(1):1–11, 1999. 6
- [8] J. M. Cowley and A. F. Moodie. The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Crystallographica*, 10(10):609–619, Oct 1957. 13
- [9] Ondrej L. Krivanek, Matthew F. Chisholm, Valeria Nicolosi, Timothy J. Pennycook, George J. Corbin, Niklas Dellby, Matthew F. Murfitt, Christopher S. Own, Zoltan S. Szilagy, Mark P. Oxley, Sokrates T. Pantelides, and Stephen J. Pennycook. Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy. *Nature*, 464(7288):571–574, MAR 25 2010. 16, 62, 97
- [10] C. Barry Carter David B. Williams. *Transmission Electron Microscopy: A Textbook for Materials Science*. Springer, New york, 2nd edition, 2009. 18, 21, 126
- [11] Colin Ophus. Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microscopy and Microanalysis*, 25(3):563–582, 2019. 21, 62
- [12] J.A. Mir, R. Clough, R. MacInnes, C. Gough, R. Plackett, I. Shipsey, H. Sawada, I. MacLaren, R. Ballabriga, D. Maneuski, V. O’Shea, D. McGrouther, and A.I. Kirkland. Characterisation of the Medipix3 detector for 60 and 80keV electrons. *Ultramicroscopy*, 182:44–53, 2017. 23, 63

- [13] R Plackett, I Horswell, E N Gimenez, J Marchal, D Omar, and N Tartoni. Merlin: a fast versatile readout system for Medipix3. *Journal of Instrumentation*, 8(01):C01038–C01038, jan 2013. 23
- [14] C. M. O’Leary, C. S. Allen, C. Huang, J. S. Kim, E. Liberti, P. D. Nellist, and A. I. Kirkland. Phase reconstruction using fast binary 4d stem data. *APPLIED PHYSICS LETTERS*, 116(12), MAR 23 2020. 23, 51
- [15] D. Jannis, C. Hofer, C. Gao, X. Xie, A. Béché, T.J. Pennycook, and J. Verbeeck. Event driven 4D STEM acquisition with a Timepix3 detector: Microsecond dwell time and faster scans for high precision and low dose applications. *Ultramicroscopy*, 233:113423, 2022. 23, 62, 64, 65, 97, 111
- [16] F ZERNIKE. How i discovered phase contrast. *SCIENCE*, 121(3141):345–349, 1955. 25, 99
- [17] Radostin Danev and Kuniaki Nagayama. Transmission electron microscopy with zernike phase plate. *Ultramicroscopy*, 88(4):243–252, 2001. 25
- [18] Mirko Holler, Manuel Guizar-Sicairos, Esther H. R. Tsai, Roberto Dinapoli, Elisabeth Muller, Oliver Bunk, Jorg Raabe, and Gabriel Aeppli. High-resolution non-destructive three-dimensional imaging of integrated circuits. *NATURE*, 543(7645):402+, MAR 16 2017. 26
- [19] Dillan J. Chang, Dennis S. Kim, Arjun Rana, Xuezheng Tian, Jihan Zhou, Peter Ercius, and Jianwei Miao. Ptychographic atomic electron tomography: Towards three-dimensional imaging of individual light atoms in materials. *Phys. Rev. B*, 102:174101, Nov 2020. 27
- [20] W. Hoppe. Diffraction in inhomogeneous primary wave fields: 1. principle of phase determination from, electron diffraction interference. *Acta Crystallogr, A* 25, 1969. 27, 50
- [21] Timothy J. Pennycook, Andrew R. Lupini, Hao Yang, Matthew F. Murfitt, Lewys Jones, and Peter D. Nellist. Efficient phase contrast imaging in stem using a pixelated detector. Part 1: Experimental demonstration at atomic resolution. *Ultramicroscopy*, 151:160–167, APR 2015. 31, 62, 63, 64, 97, 127
- [22] Hao Yang, Timothy J. Pennycook, and Peter D. Nellist. Efficient phase contrast imaging in stem using a pixelated detector. Part II: Optimisation of imaging conditions. *Ultramicroscopy*, 151:232–239, APR 2015. 34, 35, 62, 64, 66, 89, 97, 102, 127
- [23] H. M. L. Faulkner and J. M. Rodenburg. Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm. *Phys. Rev. Lett.*, 93:023903, Jul 2004. 38
- [24] Andrew M. Maiden and John M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10):1256–1262, 2009. 38, 68, 90
- [25] Knut Mueller, Florian F. Krause, Armand Beche, Marco Schowalter, Vincent Galioit, Stefan Loeffler, Johan Verbeeck, Josef Zweck, Peter Schattschneider, and Andreas Rosenauer. Atomic electric fields revealed by a quantum mechanical approach to electron picodiffraction. *Nature communication*, 5:5653, DEC 2014. 50, 64

BIBLIOGRAPHY

- [26] J.M. Rodenburg, B.C. McCallum, and P.D. Nellist. Experimental tests on double-resolution coherent imaging via STEM. *Ultramicroscopy*, 48(3):304–314, 1993. 50, 62, 63
- [27] Christoph Hofer, Jacob Madsen, Toma Susi, and Timothy J. Pennycook. Detecting charge transfer at defects in 2d materials with electron ptychography, 2024. 56
- [28] H. Yang, R. N. Rutte, L. Jones, M. Simson, R. Sagawa, H. Ryll, M. Huth, T. J. Pennycook, M. L. H. Green, H. Soltau, Y. Kondo, B. G. Davis, and P. D. Nellist. Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures. *Nature communications*, 7, AUG 2016. 56, 62, 63, 68, 97, 113
- [29] C. Gao, C. Hofer, and T. J. Pennycook. Pyptychostem. <https://gitlab.com/pyptychostem/pyptychostem>, 2021. 56, 84, 101
- [30] S. J. Pennycook. Z-contrast STEM for materials science. *Ultramicroscopy*, 30(1):58–69, 1989. 62
- [31] S. J. Pennycook and D.E. Jesson. High-resolution Z-contrast imaging of crystals. *Ultramicroscopy*, 37(1):14–38, 1991. 62
- [32] P. Perna, D. Maccariello, M. Radovic, U. Scotti di Uccio, I. Pallecchi, M. Codda, D. Marre, C. Cantoni, J. Gazquez, M. Varela, S. J. Pennycook, and F. Miletto Granozio. Conducting interfaces between band insulating oxides: The LaGaO₃/SrTiO₃ heterostructure. *Applied physics letters*, 97(15), OCT 11 2010. 62
- [33] Wu Zhou, Myron D. Kapetanakis, Micah P. Prange, Sokrates T. Pantelides, Stephen J. Pennycook, and Juan-Carlos Idrobo. Direct determination of the chemical bonding of individual impurities in graphene. *Phys. Rev. Lett.*, 109:206803, Nov 2012. 62
- [34] Chen Ge, Kui-Juan Jin, Qing-Hua Zhang, Jian-Yu Du, Lin Gu, Hai-Zhong Guo, Jing-Ting Yang, Jun-Xing Gu, Meng He, Jie Xing, Can Wang, Hui-Bin Lu, and Guo-Zhen Yang. Toward switchable photovoltaic effect via tailoring mobile oxygen vacancies in perovskite oxide films. *ACS applied materials & Interfaces*, 8(50):34590–34597, DEC 21 2016. 62
- [35] Gang Ou, Yushuai Xu, Bo Wen, Rui Lin, Binghui Ge, Yan Tang, Yuwei Liang, Cheng Yang, Kai Huang, Di Zu, Rong Yu, Wenxing Chen, Jun Li, Hui Wu, Li-Min Liu, and Yadong Li. Tuning defects in oxides at room temperature by lithium reduction. *Nature communications*, 9(1302), 2018. 62
- [36] Mudi Wu, Shiyi Chen, and Wenguo Xiang. Oxygen vacancy induced performance enhancement of toluene catalytic oxidation using LaFeO₃ perovskite oxides. *Chemical engineer journal*, 387, MAY 1 2020. 62
- [37] Ray Egerton. Radiation damage and nanofabrication in TEM and STEM. *Microscopy Today*, 29(3):56–59, 2021. 62
- [38] Karen C. Bustillo, Steven E. Zeltmann, Min Chen, Jennifer Donohue, Jim Ciston, Colin Ophus, and Andrew M. Minor. 4D-STEM of beam-sensitive materials. *Accounts of chemical research*, 54(11):2543–2551, JUN 1 2021. 62

- [39] A. Velazco, A. Béch e, D. Jannis, and J. Verbeeck. Reducing electron beam damage through alternative STEM scanning strategies, Part I: Experimental findings. *Ultramicroscopy*, 232:113398, 2022. 62
- [40] S. D. Findlay, N. Shibata, H. Sawada, E. Okunishi, Y. Kondo, T. Yamamoto, and Y. Ikuhara. Robust atomic resolution imaging of light elements using scanning transmission electron microscopy. *Applied physics letters*, 95(19), NOV 9 2009. 62
- [41] Ryo Ishikawa, Eiji Okunishi, Hidetaka Sawada, Yukihito Kondo, Fumio Hosokawa, and Eiji Abe. Direct imaging of hydrogen-atom columns in a crystal by annular bright-field electron microscopy. *Nature materials*, 10(4):278–281, APR 2011. 62
- [42] S. D. Findlay, N. Shibata, H. Sawada, E. Okunishi, Y. Kondo, and Y. Ikuhara. Dynamics of annular bright field imaging in scanning transmission electron microscopy. *Ultramicroscopy*, 110(7):903–923, JUN 2010. 62
- [43] Young-Min Kim, Stephen J. Pennycook, and Albina Y. Borisevich. Quantitative comparison of bright field and annular bright field imaging modes for characterization of oxygen octahedral tilts. *Ultramicroscopy*, 181:1–7, 2017. 62
- [44] Emrah Yucelen, Ivan Lazic, and Eric G. T. Bosch. Phase contrast scanning transmission electron microscopy imaging of light and heavy atoms at the limit of contrast and resolution. *Scientific reports*, 8:2676, FEB 8 2018. 62
- [45] Ivan Lazić, Eric G.T. Bosch, and Sorin Lazar. Phase contrast STEM for thin samples: Integrated differential phase contrast. *Ultramicroscopy*, 160:265–280, 2016. 62, 64, 78, 85
- [46] Jordan A. Hachtel, Juan Carlos Idrobo, and Miaofang Chi. Sub-Angstrom electric field measurements on a universal detector in a scanning transmission electron microscope. *Advanced structural and chemical imaging*, 4:10, AUG 24 2018. 62
- [47] Zhongbo Li, Johannes Biskupek, Ute Kaiser, and Harald Rose. Integrated differential phase contrast (IDPC)-STEM utilizing a multi-sector detector for imaging thick samples. *Microscopy and Microanalysis*, pages 1–11, 2022. 62
- [48] Yi Jiang, Zhen Chen, Yimo Hang, Pratiti Deb, Hui Gao, Saien Xie, Prafull Purohit, Mark W. Tate, Jiwoong Park, Sol M. Gruner, Veit Elser, and David A. Muller. Electron ptychography of 2D materials to deep sub-Angstrom resolution. *Nature*, 559(7714):343–349, JUL 19 2018. 62, 97
- [49] Zhen Chen, Yi Jiang, Yu-Tsun Shao, Megan E. Holtz, Michal Odstrčil, Manuel Guizar-Sicairos, Isabelle Hanke, Steffen Ganschow, Darrell G. Schlom, and David A. Muller. Electron ptychography achieves atomic-resolution limits set by lattice vibrations. *Science*, 372(6544):826–831, 2021. 62, 90, 97
- [50] Hao Yang, Ian MacLaren, Lewys Jones, Gerardo T. Martinez, Martin Simson, Martin Huth, Henning Ryll, Heike Soltau, Ryusuke Sagawa, Yukihito Kondo, Colin Ophus, Peter Ercius, Lei Jin, Andr as Kov acs, and Peter D. Nellist. Electron ptychographic phase imaging of light elements in crystalline materials using wigner distribution deconvolution. *Ultramicroscopy*, 180:173–179, 2017. 62, 66, 78, 97
- [51] Timothy J. Pennycook, Gerardo T. Martinez, Peter D. Nellist, and Jannik C. Meyer. High dose efficiency atomic resolution imaging via electron ptychography. *Ultramicroscopy*, 196:131–135, JAN 2019. 62, 97

BIBLIOGRAPHY

- [52] James M. LeBeau, Scott D. Findlay, Leslie J. Allen, and Susanne Stemmer. Position averaged convergent beam electron diffraction: Theory and applications. *Ultramicroscopy*, 110(2):118–125, 2010. 63, 70, 130
- [53] Colum M. O’Leary, Gerardo T. Martinez, Emanuela Liberti, Martin J. Humphry, Angus I. Kirkland, and Peter D. Nellist. Contrast transfer and noise considerations in focused-probe electron ptychography. *Ultramicroscopy*, 221:113189, 2021. 66, 68, 89, 97, 102, 105
- [54] Jacob Madsen, Timothy J. Pennycook, and Toma Susi. ab initio description of bonding for transmission electron microscopy. *Ultramicroscopy*, 231:113253, 2021. 66, 84, 101, 113
- [55] Chu-Ping Yu, Thomas Friedrich, Daen Jannis, Sandra Van Aert, and Johan Verbeeck. Real-time integration center of mass (ricom) reconstruction for 4d stem. *Microscopy and Microanalysis*, 28(5):1526–1537, 2022. 68, 90
- [56] Achim Strauch, Dieter Weber, Alexander Clausen, Anastasiia Lesnichaia, Arya Bangun, Benjamin März, Feng Jiao Lyu, Qing Chen, Andreas Rosenauer, Rafal Dunin-Borkowski, and et al. Live processing of momentum-resolved STEM data for first moment imaging and ptychography. *Microscopy and Microanalysis*, 27(5):1078–1092, 2021. 68, 90
- [57] Earl J. Kirkland. *Advanced computing in electron microscopy*. Springer, 2 edition, 2010. 69
- [58] C. Gao, C. Hofer, D. Jannis, A. Beche, J. Verbeeck, and T. J. Pennycook. Overcoming contrast reversals in focused probe ptychography of thick materials: An optimal pipeline for efficiently determining local atomic structure in materials science. *Applied Physics Letters*, 121(8):081906, 2022. 78, 84, 85, 86, 87, 89, 90, 97, 98, 102, 110, 113
- [59] Laura Clark, Gerardo T Martinez, Colum M O’Leary, Hao Yang, Zhiyuan Ding, Timothy C Petersen, Scott D Findlay, and Peter D Nellist. The Effect of dynamical scattering on single plane phase retrieval in electron ptychography. *Microscopy and Microanalysis*, 29(1):384–394, 2022. 78, 85, 87, 90, 97, 98, 102, 110, 113
- [60] Takehito Seki, Naoto Takanashi, and Eiji Abe. Integrated contrast transfer function for aberration corrected phase contrast STEM. *Ultramicroscopy*, 194:193–198, 2018. 78, 85
- [61] Takehito Seki, Kushagra Khare, Yoshiki O. Murakami, Satoko Toyama, Gabriel Sánchez-Santolino, Hirokazu Sasaki, Scott D. Findlay, Timothy C. Petersen, Yuichi Ikuhara, and Naoya Shibata. Linear imaging theory for differential phase contrast and other phase imaging modes in scanning transmission electron microscopy. *Ultramicroscopy*, 240:113580, 2022. 78, 80
- [62] Christoph Hofer and Timothy J. Pennycook. Reliable phase quantification in focused probe electron ptychography of thin materials. *Ultramicroscopy*, 254:113829, 2023. 89, 105
- [63] Philipp M. Pelz, Ian Johnson, Colin Ophus, Peter Ercius, and Mary C. Scott. Real-time interactive 4D-STEM phase-contrast imaging from electron event representation data: Less computation with the right representation. *IEEE Signal Processing Magazine*, 39(1):25–31, 2022. 90

- [64] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg. Ptychographic transmission microscopy in three dimensions using a multi-slice approach. *J. Opt. Soc. Am. A*, 29(8):1606–1614, Aug 2012. 90, 98
- [65] Liqi Zhou, Jingdong Song, Judy S Kim, Xudong Pei, Chen Huang, Mark Boyce, Luiza Mendonça, Daniel Clare, Alistair Siebert, Christopher S Allen, et al. Low-dose phase retrieval of biological specimens using cryo-electron ptychography. *Nature communications*, 11(1):2773, 2020. 97
- [66] Xudong Pei, Liqi Zhou, Chen Huang, Mark Boyce, Judy S Kim, Emanuela Liberti, Yiming Hu, Takeo Sasaki, Peter D Nellist, Peijun Zhang, et al. Cryogenic electron ptychographic single particle analysis with wide bandwidth information transfer. *Nature Communications*, 14(1):3027, 2023. 97
- [67] Daniel G Stroppa, Matthias Meffert, Christoph Hoermann, Pietro Zambon, Darya Bachevskaya, Hervé Remigy, Clemens Schulze-Briese, and Luca Piazza. From STEM to 4D STEM: Ultrafast Diffraction Mapping with a Hybrid-Pixel Detector. *Microscopy Today*, 31(2):10–14, 04 2023. 97
- [68] Frederick Allars, Peng-Han Lu, Maximilian Kruth, Rafal E. Dunin-Borkowski, John M. Rodenburg, and Andrew M. Maiden. Efficient large field of view electron phase imaging using near-field electron ptychography with a diffuser. *Ultramicroscopy*, 231:113257, 2021. 98
- [69] Hamish G. Brown, Philipp M. Pelz, Shang-Lin Hsu, Zimeng Zhang, Ramamoorthy Ramesh, Katherine Inzani, Evan Sheridan, Sinéad M. Griffin, Marcel Schloz, Thomas C. Pekin, and et al. A three-dimensional reconstruction algorithm for scanning transmission electron microscopy data from a single sample orientation. *Microscopy and Microanalysis*, 28(5):1632–1640, 2022. 98
- [70] Philipp M. Pelz, Hamish G. Brown, Scott Stonemeyer, Scott D. Findlay, Alex Zettl, Peter Ercius, Yaqian Zhang, Jim Ciston, M. C. Scott, and Colin Ophus. Phase-contrast imaging of multiply-scattering extended objects at atomic resolution by reconstruction of the scattering matrix. *Phys. Rev. Res.*, 3:023159, May 2021. 98
- [71] E.W.C. Terzoudis-Lumsden, T C Petersen, Hamish G. Brown, Philipp Pelz, Colin Ophus, and Scott D. Findlay. Resolution of virtual depth sectioning from four-dimensional scanning transmission electron microscopy. *Microscopy and microanalysis : the official journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada*, 29 4:1409–1421, 2023. 98
- [72] C. Gao, C. Hofer, and T.J. Pennycook. On central focusing for contrast optimization in direct electron ptychography of thick samples. *Ultramicroscopy*, 256:113879, 2024. 98, 102, 105, 110
- [73] Yanbin Li, Weijiang Zhou, Yuzhang Li, Wenxiao Huang, Zewen Zhang, Guangxu Chen, Hansen Wang, Gong-Her Wu, Nicholas Rolston, Rafael Vila, et al. Unravelling degradation mechanisms and atomic structure of organic-inorganic halide perovskites by cryo-EM. *Joule*, 3(11):2854–2866, 2019. 99
- [74] Junhui Ran, Ondrej Dyck, Xiaozheng Wang, Bin Yang, David B. Geohegan, and Kai Xiao. Electron-beam-related studies of halide perovskites: Challenges and opportunities. *Advanced Energy Materials*, 10(26):1903191, 2020. 99

BIBLIOGRAPHY

- [75] Felix Wittwer, Johannes Hagemann, Dennis Brückner, Silja Flenner, and Christian G. Schroer. Phase retrieval framework for direct reconstruction of the projected refractive index applied to ptychography and holography. *Optica*, 9:295–302, 2022. 113

