The University Of Sheffield.

Near Field Electron Ptychography

By Frederick William Allars

Supervised by:

Dr Andrew Maiden

Dr Gavin Williams

A Thesis Submitted for the degree of Doctor of Philosophy

The University of Sheffield

Department of Electronic and Electrical Engineering

January 2023

Submitted by a student with dyslexia

Abstract

Phase imaging in the Transmission Electron Microscope (TEM) has a long history, from the implementation of off-axis holography in TEM to Differential Phase Contrast (DPC) on the Scanning Transmission Electron Microscopy (STEM). The advent of modern computing has enabled the development of iterative algorithms which attempt to recover a phase image of a specimen from measurements of the way it diffracts an incident electron beam. One of the most successful of these iterative methods is focused probe ptychography, which relies on far field diffraction pattern measurements recorded as the incident beam is scanned through a grid of locations across the specimen. Focused probe ptychography implemented in the STEM has provided the highest resolution images available to date, allows for lens-less setups avoiding the aberrations typical in older STEMs and allows for simultaneous reconstruction of the illumination and specimen. Ptychography is computationally flexible (highly constrained), allowing for additional unknowns other than the phase of the specimen to be recovered, for example positions can be refined during reconstruction.

Near field ptychography is a recent variation on ptychography that replaces the far-field diffraction data with diffraction patterns recorded in the near field, or Fresnel, region. It promises to obtain a much larger field of view with fewer diffraction patterns than focused probe ptychography. The main contribution of this thesis is the implementation of a new form of near field ptychography on the Transmission Electron Microscope (TEM), using an etched silicon nitride window to structure the electron beam. Proof-of-concept results show the method quantitatively recovers megapixel phase images from as few as 9 recorded diffraction patterns, compared to many hundreds of diffraction patterns required for focused probe ptychography. Additional sets of results show how near-field ptychography can recover extremely large fields of view, deal effectively with inelastic scattering, and accommodate several sources of uncertainty in the experimental process.

Further contributions in the thesis include: experiments and results from visible-light versions of near field ptychography, which explain its limitations and practical application; a description and code for analysis tools that are used to assess phase imaging performance; DigitalMicrograph (DM) code and a data collection workflow to realise TEM-based near-field ptychography; details of the design, realisation and performance of the etched silicon nitride windows; and simulation studies aimed at furthering understanding of the frequency response of the technique. Future work is outlined, focusing on potential applications in a wide range of real-world specimens and improved TEM setups to implement near field ptychography.

Acknowledgements

I would like to acknowledge those who have helped me through my academic studies. Firstly, I would like to acknowledge my supervisor, Dr Andrew Maiden, who has kept a cool head whilst dealing with my questions and faults. I would have not got this far without Dr Maiden's help and appreciate that he was willing to take a chance on me when no one else would and believed in me to the end. I would like to thank my second supervisor Dr Gavin Williams for being understanding and calm when I was panicking.

Special thanks to Peng-Han Lu for his assistance on section 5 and being helpful during my week at Julich, Germany.

I would like to thank Professor John M Rodenburg, who was always willing to answer any questions I had and explain any misunderstood details.

For making the office an excellent research environment, I would like to thank: Dr Yangyang Mu, Dr Zhuoqun Zhang, Dr Samuel McDermott, Dr Peng Li, Yiqian Zhang, Shengbo You and Ziyang Hu.

Finally, I would like to thank my family as whole, as they have supported me through this process, special thanks go to Dad, Mum, and Max for always believing in me.

This thesis is dedicated to the memory of Doren Allars

Contents

Publications

Parts of the work published in this thesis were previously published as the following journal paper publications:

Frederick Allars, Peng-Han Lu, Maximilian Kruth, Rafal E. Dunin-Borkowski, John M. Rodenburg, Andrew M. Maiden, Efficient large field of view electron phase imaging using near-field electron ptychography with a diffuser, Ultramicroscopy, 2021,113257, ISSN 0304-3991

My contribution included: highly engaged in the experimental process such as determining the diffraction conditions and designing the experiments, reconstructed the collected data, designed the diffusers and writing the draft paper.

Parts of the work published in this thesis were previously presented at a conference:

Near-field Electron Ptychography using a silicon nitride diffuser – Virtual Early Career European Microscopy Congress 2020 – (November. 2020) – Online

Chapter 1

1 Literature Review

1.1 Introduction

In this literature review there will be a discussion of the necessary subjects to contextualise near field electron ptychography as a phase retrieval technique. First there will be a discussion of the electron lens and its aberrations. The Transmission Electron Microscope (TEM) and various techniques will be outlined. The TEM is particularly important to various fields of science and engineering, not only for high resolution imaging, but for the richness of the information it can provide. The TEM can provide information on the chemical composition of the specimen imaged [1], the strain of a semiconductor device [2], the strength of magnetic field produced by magnetic thin film [3] and can produce 3D images at nanoscale [4]. The focus of this literature review will then shift to various phase retrieval methods used in TEM/STEM (Scanning Transmission Electron Microscope). Most importantly, the concepts of ptychography and near field ptychography will be covered. At the end of the literature review there will be a short review of propagation of complex fields, coherence and the Fourier Ring Correlation. These three concepts are important to the work shown in this thesis and are therefore covered to demonstrate that the author has the required knowledge to tackle these issues in real world experiments.

One of the key parameters in determining the resolution of an imaging system is the wavelength of the illumination used, this is highlighted by the Rayleigh diffraction criteria [5]. De Broglie was the first to describe the wavelength of a matter wave of the electron, and relate the wavelength of the electron to the electron's momentum (or energy) [6]. By connecting the resolution of a microscope to the wavelength of the electron matter wave, the motivation behind developing the electron microscope becomes clear, because of the potential resolution that they can achieve. It was then proved by Scherzer that the design of the electron lens itself would present a barrier in maximising the resolution of the TEM (Transmission Electron Microscope). Scherzer identified that the lens used to focus the electrons suffered from spherical aberration due to the shape of the electron lens [7]. There will now be a description of the aberrations of the electron lens, as they are one of the key factors in limiting the resolution of the TEM.

The performance of an electron lens is hampered by its aberrations, not only the spherical aberrations but also chromatic aberrations. Furthermore, there are a number of other aberrations to be aware of, which have been defined by a set of basis functions often described as Zernike polynomials [8, 9]. The theory of Zernike polynomials describes aberrations in polar coordinates (radius and angle) and makes categorising aberrations more systematic. For example, spherical aberrations are described as a fourth order aberration in the Zernike system, as the highest power of radius in the function describing spherical aberration is 4. The second descriptor used in Zernike polynomials is called the frequency, which describes the multiplier applied to the polar angle. Frequency in the context of Zernike polynomials can be negative or positive in value, and it indicates whether a sine or cosine was used (-1 for sine). A simplified version of the first four orders of the Zernike polynomials is shown in Figure 1 (order is indicated by r^n, where n is the order and frequency is indicated by multiples of θ). As an aside, the tilt aberrations effectively describe the action of the deflector coils of the electron microscope which are described in subsection 1.2 and they cause the beam to be tilted by a certain angle.

Spherical aberration as mentioned earlier is the greatest cause of resolution loss in the electron microscope, and can be described physically as the focal point of the lens being a function of the radius of the lens [6]. An example phase map of spherical aberration can be seen in Figure 1 i). The effect of spherical aberration is illustrated by Figure 2a. Figure 2a also demonstrates how spherical aberration limits the resolution of the electron microscope, as the image is blurred at the detector plane by spherical aberration (the rays are not focused), this can be somewhat mitigated by refocusing the lens to make use of the disc of least confusion [6]. The disc of least confusion in Figure 2a is the cross-section where the rays are closest together (marked by the arrows). The disc of least confusion effectively describes the best resolution that can be obtained from an aberrated lens, this is achieved by using a defocused condition.

Another limitation that affects electron microscopes is called chromatic aberration. Chromatic aberration describes the fact that when an electron lens interacts with electrons with varying wavelengths, those electrons will be focused to different positions along the optical axis. This is illustrated in Figure 2, where the red dotted line shows electrons being focused to a different plane as they have different energy to the acceleration voltage. Technically, Figure 2 should show a spectrum of differing focal points, but one focal point is shown here for clarity. Figure 2 shows the effect of the chromatic aberration on STEM (Scanning Transmission Electron Microscopy) operation. However, the effect on the TEM operation would be that electrons which have lost energy would be deflected by differing amounts by the lens, additionally these images formed from inelastic electrons would be defocused compared to the intended image plane. The differences in the focal plane and the magnification then cause the images formed by the aberrated lens to be blurred (different aberrations also have a similar effect on TEM image formation, but the fundamental cause of the defocus and magnification is different). Chromatic aberration is detrimental to the operation of the electron microscope, as the mere presence of a specimen is sufficient to cause inelastic scattering, and the spread of the electron wavelengths resulting from the specimen interaction would cause an image formed by the objective lens to suffer from chromatic aberration. It is not only the inelastic scattering caused by the specimen which contributes to the energy spread which causes the blurring associated with chromatic aberration, but also the instabilities of the microscope itself, as the variation in the acceleration voltage of the microscope can cause the same blurring effect (another similar blurring effect is the variance of the objective lens strength/current) [10]. There are two possible methods to address chromatic aberrations; fix the aberration or reduce the energy spread of the electrons in the column. This is why many high performance microscopes have pre specimen energy filters to reduce the energy spread of the illumination system. Chromatic aberration correction is particularly important for low energy imaging of sensitive samples because as the acceleration voltage decreases, the instability of the acceleration voltage remains the same and therefore the blurring caused by chromatic aberration increases as the acceleration voltage decreases [11].

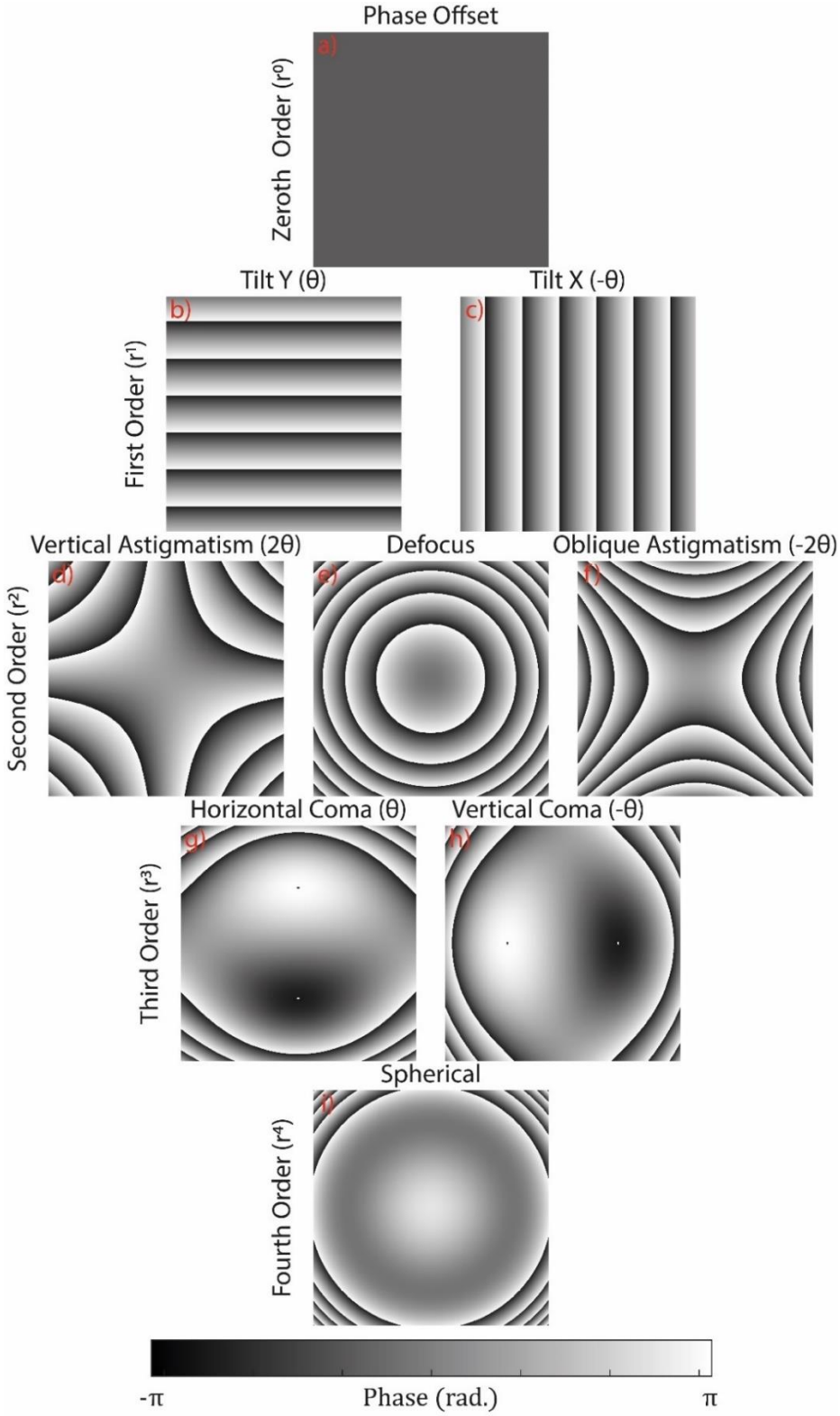

Figure 1: A figure shows a simplified version of the first four orders of the Zernike polynomials. a) displays the zero order Zernike polynomial of phase offset, b-c) shows the first order Zernike polynomial of phase ramps along the y and x axes respectively, d) shows second order Zernike polynomial of vertical astigmatism. e) shows the Zernike polynomial of defocus. f) shows the second order Zernike polynomial of oblique astigmatism. g-h) show the third order Zernike polynomials of horizontal and vertical coma respectively. i) displays the fourth order Zernike polynomial associated with spherical aberrations.

The astigmatism of the electron lens effectively means that it is impossible to focus a beam such that the width of the beam in the x and y axis is simultaneously minimized. Examples of phase maps of astigmatism can be seen in Figure 1 d) and Figure 1 f). The effect of astigmatism is further illustrated by Figure 2 c) and Figure 2 d). This is particularly important to STEM as this would prevent the microscope forming the smallest possible probe. STEM images produced under astigmatic conditions suffer from streaking in a particular direction, extending that feature over a longer distance than what is true to the actual specimen. In TEM images the effect of astigmatism is that the degree of focus of a feature will depend on the orientation of that feature. Astigmatism is typically identified in the TEM context by taking the Fourier transform of a defocused TEM image, which will display elliptical ring shape in the case of astigmatism and ring like shape in its absence.

Figure 1 e) shows the aberration of defocus, which is an interesting aberration, as the cause of defocus is not limited to the lens itself. For example, if the specimen is not in the eucentric plane of the TEM this effectively causes a defocus aberration. The eucentric plane is the plane where the tilting of the specimen around the tilt axis causes the minimum amount of apparent orthogonal movement with respect to the optical axis, x/y [6]. Furthermore, a defocus aberration is exactly the same as propagating the specimen a short distance with the angular spectrum method (in the case of collimated illumination, the angular spectrum method is discussed in Section 1.28). The defocus aberration is of importance to near field ptychography and inline holography, as it is a method of forming interference fringes which carry information on the phase of the specimen.

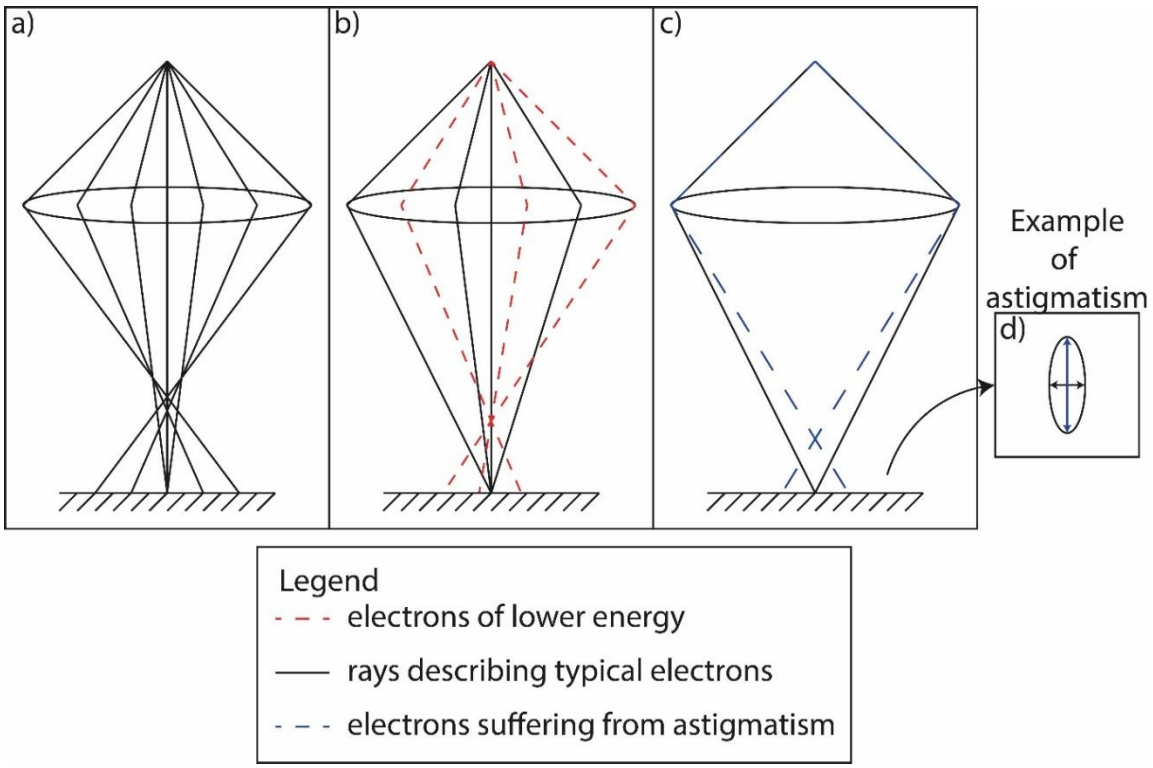

Figure 2: A figure showing the effect of various aberrations on the ability of an electron lens to focus to a point, and therefore the effect of various aberrations on STEM operation. a) shows how a spherical aberration causes rays which interact at different radii of lens are focused to a different point along optical axis. b) shows how a lens suffering from chromatic aberrations focuses electrons with lower energies more strongly than typical electrons. c) and d) shows how astigmatism causes electrons which interact with different sides of the electron lens to have different focal points [6]. The dotted line in c shows the different focal point of electrons that interact with the top and bottom of the lens as opposed to the left and right which is represented by the solid black line.

1.2 Scan coils

Scan coils are an important component of the electron microscope and other technologies. For example, scan coils were a vital component of the television until the late 1990s, as they were used to scan the electron beam from a cathode ray tube on to a phosphor screen producing an image. This application required that the scan coils operated at a high frequency (such that the whole screen could be scanned before the next frame of the program), similar to the scanning speed required by the STEM. The main contemporary applications of scan soils are electron microscopes and lithography. Other names for scan coils include magnetic yokes and deflection coils.

The scan coil in the electron microscope is the simplest electro-magnetic component of the microscope. It is comprised of two electromagnets with their north and south poles aligned such that magnetic flux flows between the two poles. Special care is taken to reduce the inductance of the wires of the electromagnets so that they can raster the beam across the specimen in the shortest possible time frame. Typically, they are driven with a triangle wave signal. The constant flux between the two poles ensures that there is phase ramp applied to the electron beam [12]. Typically, the application of the phase ramp is instead described as a deflection of the electron by the Lorentz force, deflecting the electron beam by a particular angle ϵ. William's and Carter's work [6] on TEM describes the deflection angles as follows:

$$\epsilon = \frac{eLB}{mv} \tag{1}$$

The parameters of equation 1 are as follows:

- ϵ is the deflection angle
- e is the charge of an electron
- L is the length/distance over which magnetic force is applied
- B is the strength of the magnetic field produced by the scan coils in Teslas
- m is the mass of the electron
- v is the electron current velocity

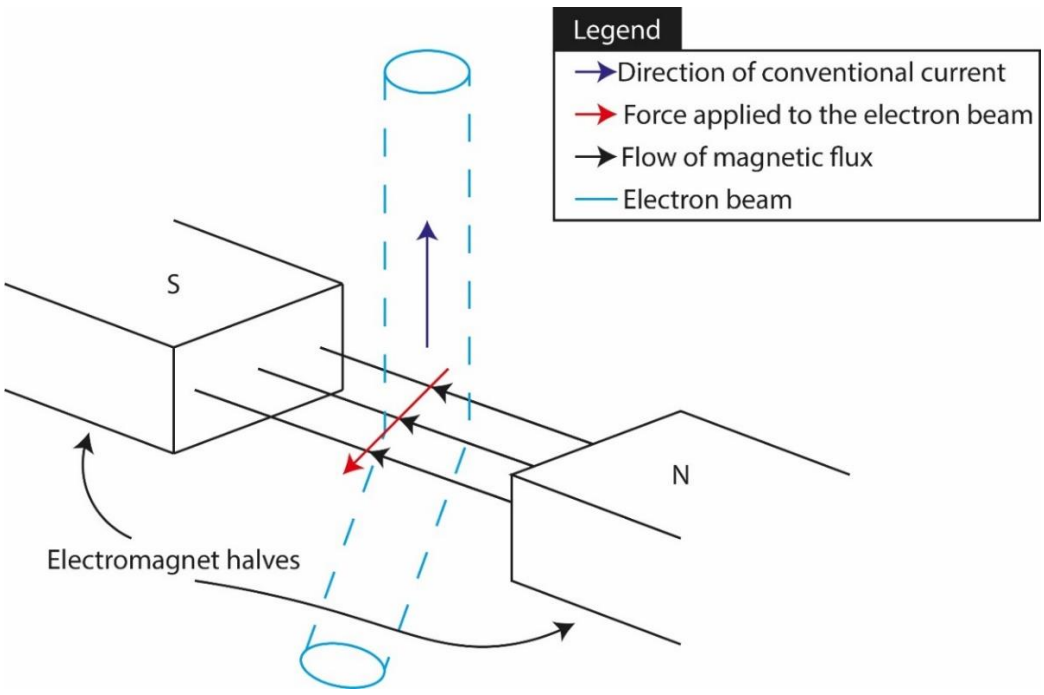

Figure 3: A simple diagram outlining the principle of deflection coil in S/TEM.

Two scan coils form a double deflection pair as one deflects the electron beam by an angle ϵ, the other deflects the beam by $-\epsilon$, but due to the displacement of the deflection coils this causes the electron beam to traverse a distance perpendicular to the optical axis (an effective shift). This is used primarily in STEM to move the focused electron probe in a raster like fashion to form STEM images (this is expanded upon in section 1.5). STEM primarily makes use of the condenser lens scan coils, which are typically found between the first and second condenser lenses. There are other scan coils throughout the electron microscope, for example: there are post specimen scan coils after the objective lens but before the selected area aperture. Typically, these are used to tilt the electron beam for dark field imaging of crystalline specimens (i.e. when only scattered/highly scattered electrons are imaged). The post objective scan coils have also been used to descan the beam in confocal scanning transmission electron microscopy. Descanning is the process of correcting a shift from the condenser scan coils such that the electron beam is centred on the optical axis. If the electron beam is parallel to the optical axis but not directly on it, it will suffer from aberrations to a greater degree, as the nature of spherical aberration correction assumes that you are using the centre of the objective lens.

The descan coils' primary importance is during STEM operation whereby they ensure the diffraction patterns are centred on the detector, regardless of the position of the probe on the specimen. This is particularly important in procession electron diffraction, where the beam is tilted at various angles with respect to the specimen to average out the effect of dynamical scattering. The descan coils are also import to differential phase contrast (DPC) imaging as they ensure the collected diffraction data are centred correctly. This is important to DPC as it ensures that the obtained differential phase image is accurate, as the differential phase images are proportional to the movement of the central beam with respect to its original position where the specimen is absent [13].

Scan coils are particularly interesting to near field electron ptychography's future as they can be used as part of a hybrid stage and scan system, which combines the long-range movements of the specimen stage with the quick short distance movements of the scan coils.

1.3 Electron source

The electron gun provides a well-defined source of electrons. There are two types of electron gun and then a composite of the two types: the thermionic emission gun, the field emission gun (FEG), and the field enhanced thermionic emission gun (or Schottky gun). The type of electron gun used will be application dependant. For example, in near field electron ptychography the chosen gun will likely be the field emission gun, as the FEG meets the requirement of having higher brightness and good spatial coherence.

Richardson's law succinctly outlines thermionic emission with respect to the electron gun [6]:

$$J = AT^2 \exp\left(\frac{-\phi}{kT}\right) \tag{2}$$

The parameters of equation 2 are as follows:

- J is the current density of the emitted electron beam
- A is Richardson's constant for a given material
- T is the temperature of the material/electron gun
- ϕ is the work function of the material often in terms of electron volts
- k is the Boltzmann constant (1.381×10^{-23})

Richardson's law states that electrons have the potential to escape from their binding nucleus at any given time, but the probability of escape is dependant on the temperature of the material and its work function [6]. Therefore, if the correct material is heated, a high current density will be obtained, and this can then be used to form an electron beam, which is the aim of the thermionic electron gun. In terms of designing the value of *J* there are two approaches: firstly, aim for a material with a high melting temperature such that the exponential of Richardsons's law can approach unity through intense heating and a reasonable current is obtained. Secondly, start with a reasonably low work function which does not require much heating in order to obtain an appropriate beam current. The results of these aims are: the tungsten thermionic electron gun which operates at high temperatures, and the Lanthanum Hexaboride (LaB_6) thermionic electron gun which is known for its low work function. Currently tungsten is becoming less important in the context of thermionic emission due to its lower brightness and larger source size (lower spatial coherence) relative to LaB_6.

Another way of forming an electron beam is via field emission in the form of an electron field emission gun or FEG [6, 14]. The FEG operates on the principle of stripping a tungsten cathode of its electrons through the application of a strong electric field, which enables the electrons of the tungsten cathode to tunnel through their band structure into the vacuum of the electron microscope. The electric field reduces the binding energy between the electrons and their atoms such that the inherit energy of the electron gun's temperature is enough to enable the electrons to escape, as seen in equation 3. The effect of the electric field is maximised by making the tungsten cathode into a very fine tip (high surface area to volume ratio) which has the additional benefit of creating a very small probe size and therefore highly spatially coherent electron beam. Furthermore, FEGs emit the largest density of electrons of any conventional electron gun due to their emission mechanism.

$$J = AT^2 \exp\left(\frac{-\phi + \phi_{electric\ field}}{kT}\right) \tag{3}$$

The $\phi_{electric\ field}$ parameter in equation 3 represents the effect of the electric field on the work function of the material.

A combination of the principles of the field emission gun and the thermionic emission together created the Schottky emission gun [6]. This gun is effectively the middle point between thermionic emission and field emission, as it uses both mechanisms to produce an electron beam current. One of the advantages of the Schottky gun is the relatively good performance without the requirement of a high quality vacuum.

1.4 Lenses and apertures of the electron microscope

Electron microscope lenses are magnetic lenses formed of a soft iron pole piece in the shape of a ring [6, 14]. A copper wire is wrapped around the pole piece to form an electromagnet. When current flows through the copper wire a magnetic field is induced which acts as our electron lens. If designed appropriately, electron lenses can be described in the same geometric manner as visible light microscopy lenses, where they have focal points. The categories of electron lenses in the TEM are: the condenser lenses, the objective lens, and the projector lenses. The positions of these lenses and other components discussed in this section can be seen in Figure 4.

The action of the condenser lens system is to form a probe with the electrons from the electron gun crossover [6]. After the electrons are collected by the first condenser lens, the second condenser lens alters the size of the electron beam on the sample through altering it by its own magnification

value. It should be noted that there are many ways in which the condenser lens system and the upper objective pole piece can change the probe on the sample, but they are beyond the scope of this introduction. The condenser lenses like other electron lens (such as the objective lens) suffers from spherical aberrations which presents difficulties in forming small high current probes. The condenser lens system is of great importance to STEM, as the size of the illumination on the specimen determines the ultimate resolution of the STEM images produced.

The objective lens magnifies the exit wave of the specimen. The objective lens has the greatest magnifying power of all of the lenses in the microscope [6]. The objective lens suffers from spherical aberration, which limits the resolution of the entire system. As the magnification value of the object lens is so high, it is often a requirement that the sample be effectively submerged in the objective lens magnetic field, which is achieved by splitting upper and lower pole pieces. Magnetic samples require that they are not submerged in a magnetic field to avoid their magnetic fields being realigned to match that of the objective lens. So when imaging magnetic samples the low pole piece is turned off; this form of microscopy is called Lorentz TEM [15].

There are many apertures in the electron microscope, typically located in the image and back focal planes of the various lenses. The first apertures are the electron gun or the condenser lens apertures. These apertures limit the area of the electron gun which contributes to the electron beam, therefore varying the spatial incoherence of the electron beam accordingly. The C2 aperture is used to limit the spatial extent of the electron beam while operating in TEM mode.

The condenser aperture improves the spatial coherence of the probe at the cost of counts (i.e. beam current). Therefore, there is a balance between spatial coherence and the signal to noise ratio needed to form a reliable analysis after the image is taken. Like the condenser lenses, the apertures in the condenser lens system are critical to the performance of STEM, as they effectively determine the kind of STEM experiment which is performed [6]. As the C2 aperture becomes larger, the diffraction spots on the detector become larger and begin to interfere with one another.

The C2 aperture is effectively placed in the Fourier domain with respect to the illumination at the specimen. As will be shown later when discussing the propagation of waves through space (see section 1.28), the Fourier space effectively describes the angles at which plane waves move with respect to the optical axis. In a sense the C2 aperture determines the set of angles (or spatial frequencies) that are allowed in the illumination function. Furthermore, the aberrations of the electron lens become more prominent the further away a particular ray interacts with a lens from its optical axis (as was demonstrated earlier by spherical aberrations). Therefore, the C2 aperture places a limit on the angular extent of the lens.

The largest possible width of the C2 aperture is determined by the aberration correction available to the electron microscope, because if the C2 aperture is made larger than the aberration corrected portion of the condenser lens, then the probes it forms will be aberrated and blurred, increasing their width and reducing the resolution in the case of Scanning Transmission Electron Microscopy (STEM).

The Ronchigram is a method by which to visualise the aberrations and quality of a STEM probe; forming a convergent beam diffraction pattern with an amorphous material as the specimen allows the visualisation of the aberrations of the condenser lens. The Ronchigram is used in order to select the correct aperture size such that the aberrations are not included, whilst minimising the size of the probe on the specimen [16, 17].

The action of the apertures after the sample is to select an area either in real space or Fourier space. Selecting an area in Fourier space with an objective aperture (an aperture in the back focal plane of the objective lens) in its most basic form is a form of filtering (low pass, band pass, high pass) [18]. Selecting an area in Fourier space with an aperture also provides other benefits, such as enabling the detector to take longer exposures of less intense parts of the exit wave (dark field imaging), but again at the cost of counts and signal to noise ratio. Another post-objective aperture is the selected area aperture in the image plane of the objective lens. The selected area aperture enables the selection of real space areas, which enables the collection of diffraction patterns from specific parts of the specimen. This enables the diffraction pattern of crystalline structures to be separated from the diffraction pattern of the amorphous carbon film which supports it in the electron microscope. Furthermore, the ability to apply an aperture in real space is important for ptychography in its historical context, but this will be discussed in section 1.15.

The projector lens system further magnifies the exit wave from the objective lens, such that it is appropriately sampled by the detector [6, 14]. The detector is typically a charged coupled device which has the benefit of a high frame rate and decent detection quantum efficiency (i.e. the number of electrons which are counted with respect to actual number of electrons.) Now that the components of the TEM have been discussed, this section will now move on to the state of art techniques used to obtain phase images.

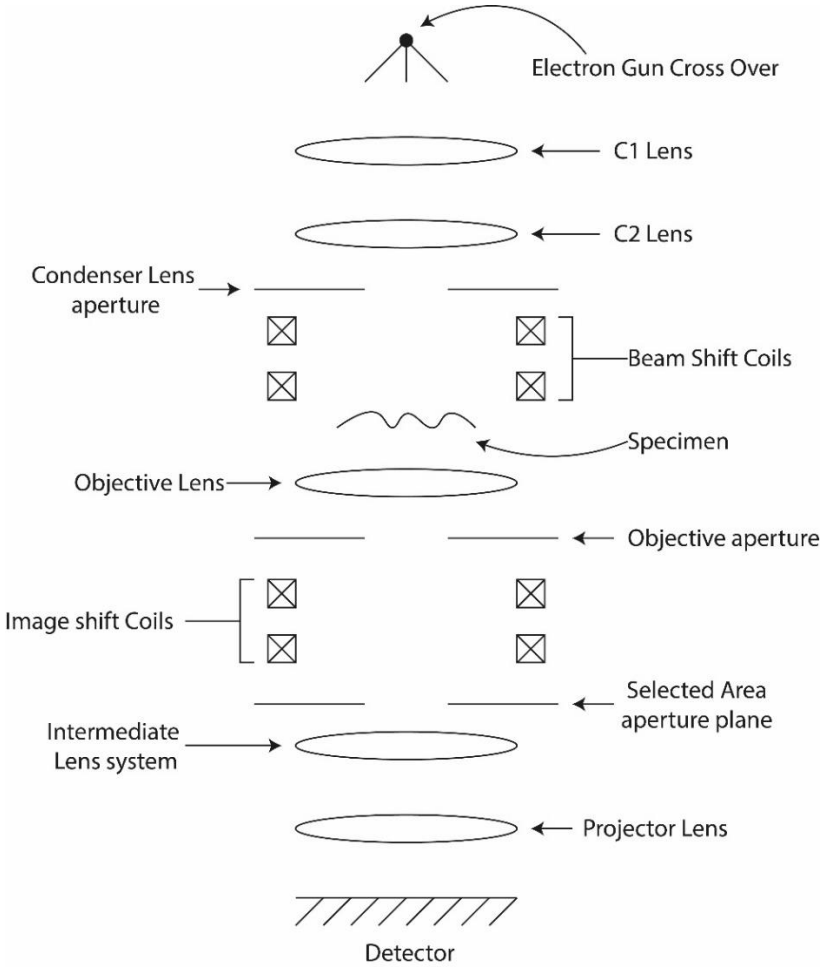

Figure 4: A diagram of some of the important components of TEM and relative positions in the column [19]

1.5 Scanning transmission electron microscopy:

STEM was developed as an alternative to transmission electron microscopy to overcome the aberrations of the objective lens, particularly the chromatic aberrations of the lens, which limit the spatial extent of the lens in Fourier space [20, 21]. In theory diffraction patterns produced by STEM do not require post specimen lenses only a detector, therefore avoiding the aberrations of such lenses. STEM is now more commonly used to obtain images which are a strong function of the specimen's atomic number, whilst also providing additional information at an atomic scale via analysis of emitted X-rays (STEM EDX – Energy Dispersive X-ray spectrometry) and the energy loss of the electrons (STEM EELS - electron energy loss spectrometry) [22]. These additional signals can map the elements of the sample at the atomic scale. The original concept of scanning transmission electron microscopy was to reduce the importance of the objective lens by reducing/removing its role in providing magnification in the optical setup. Instead, the condenser lens system was used to create the smallest possible probe, which was then scanned across the specimen. In this setup it was the size of the illumination on the specimen which determined the resolution of the final image. With the advent of aberration correction, both STEM and TEM experienced a boost in performance, as the aberrations of electron lens had prevented sub-angstrom probes from being realised [23].

The first STEM was developed by Manfred von Ardenne in parallel with the scanning electron microscope [24]. It was not until much later with the development of the cold field emission gun that STEM was further developed with the next major work being reported by Crewe et al. [25]. Crewe et al. also demonstrated an important milestone of visualising single atoms via the use of STEM with an annular detector in [26] . Pennycook and Jesson [27] later demonstrated high resolution Z-contrast incoherent imaging with a high angle annular detector (Z-contrast means a strong function of the specimen's atomic number).

1.6 Optical setup for scanning transmission electron microscopy

In this introduction to the STEM setup we will follow the electron beam from the source to the detector. STEM performs best when the electron source is cold-FEG (Field Emission Gun) due to its improved brightness and high spatial coherence. After the electrons are emitted from the electron gun they are collected by the C1 lens, then the image of the electron source is demagnified by the C2 lens. Depending on the microscope, a condenser mini-lens may be utilised to further demagnify the image of the electron source [6]. During the demagnification process the C2 aperture will be used to define the semi-angle of the illumination (α). The C2 aperture, regardless of its actual physical position, can be thought as being in the back focal plane of the C2 lens, so the C2 aperture directly determines the maximum spatial frequencies (convergence angle) of the illumination. Between the back focal plane and specimen plane the illumination undergoes a Fourier transform becoming a fine point or probe on the specimen (this is why the reconstructed illumination in ptychography is called a probe, as it originates from STEM). Once the illumination interacts with the specimen the objective and projector lens systems perform a further Fourier transform before the exit wave reaches the detector. As the illumination undergoes two Fourier transforms, the diffraction pattern at the detector has a bright disc at its centre, an image of the C2 aperture. Provided the sample is sufficiently thin, the diffraction pattern at the detector can be described as a convolution of the C2 aperture and the Fourier transform of the specimen subsection under the probe [28, 29]. STEM diffraction patterns are commonly known as Convergent Beam Electron Diffraction (CBED) patterns, due to the illumination setup.

1.7 Imaging in STEM

Now that the optical setup of STEM has been described, the discussion can move on to how STEM images are formed with this setup. Instead of a typical imaging setup inspired by optical

microscopes, scanning transmission electron microscopy makes use of the advent of the digital age. This is achieved by using a single pixel detector unit, capable of counting the number of electrons that have interacted with the detector, and by changing the position of the illumination through the use of a pair of scan coils controlled by a digital to analog converter. In this set up, a single exposure of the detector collects one pixel of the STEM image, after which the scan coils of the microscope are used to move the focused illumination a small distance to the right of the previous position. At this new position another exposure is taken. Typically, the distance between exposures is smaller than the width of the illumination, such that there is overlap at each position. The process of moving the illumination and exposing the detector is repeated until all of the pixels of the requested image are filled, as illustrated in Figure 5. The brightness value of each pixel of the STEM image is determined by the total counts that reach the detector at a particular illumination position. The number of counts that reach the detector is determined by the material and its scattering properties, which then determine the spread of counts in frequency space. The number of counts collected by the detector also depends on the magnification of frequency space by the post specimen lens (camera length control), which are used to appropriately scale the physical (and constant) size of the detector (more on this in section 1.8).

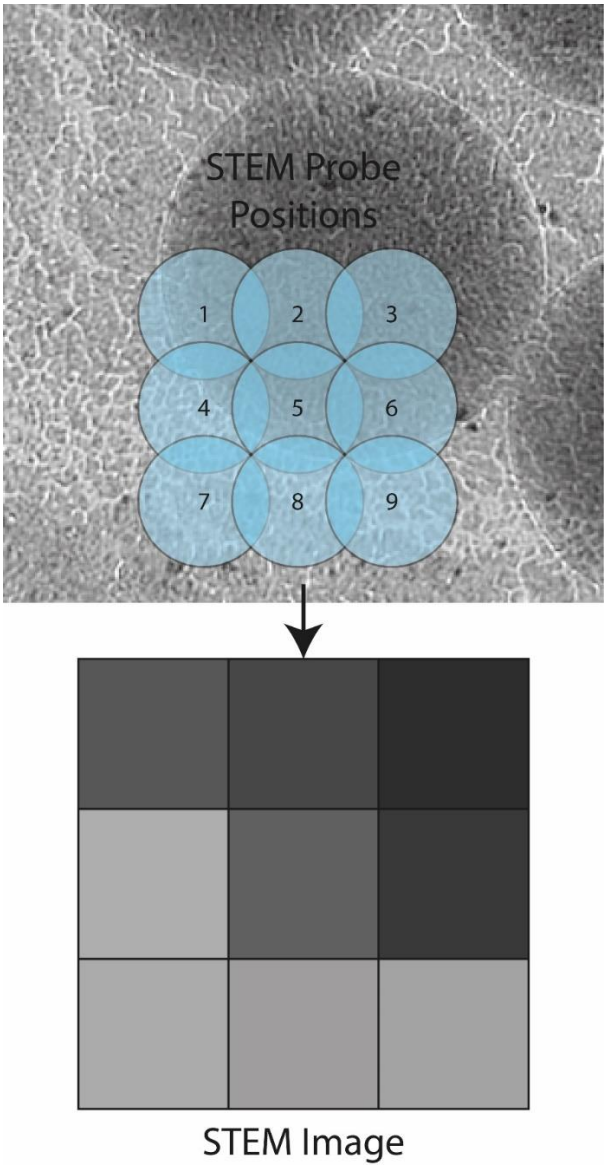

Figure 5: A figure showing how STEM scanning positions correlate to the pixels of STEM images

1.8 Scanning transmission electron microscopy detectors

In TEM, historically the detector has either been a square sheet of photographic paper or a square pixelated CCD detector coupled with a scintillator. This is not the case in STEM, due to the method through which STEM forms images (pixel by pixel). In STEM the detectors are circular or annular in nature and consist of a single detector unit. As each pixel in a STEM image is a summation of a convergent beam electron diffraction (CBED) pattern over a particular part of frequency space determined by the detector design, the detector design can be used to filter the information projected by the sample. This gives greater flexibility to STEM in terms of how images can be filtered, and allows experiments to collect more than one filtered image in a single experiment by using multiple detectors, i.e. dark field and bright field detectors. STEM experiments are often defined by how the detector was used as a filter in that particular experiment, for example if the detector was used to collect data at "high spatial frequencies" (elastically scattered high angle deflection electrons) the experiment would be called "high angle annular dark field" imaging or HAADF imaging (annular describes the fact the detector is a ring shape). The difference between HAADF and ADF (or Annular Dark Field imaging as labelled in Figure 6) is that the data collected by a HAADF detector is mostly elastic (it has not lost energy interacting with the specimen) but is spatially incoherent (due to the large radius of the annular detector [30]), while ADF images have both inelastic and elastic scatter (and are slightly more spatially coherent due to the smaller radius of the annular detector). The signal obtained by HAADF imaging is formed by electrons that directly interact with the nucleus of the atom, and therefore this scattering is a strong function of the atomic number of the specimen. This is also why the signal is elastic, as there is little momentum transfer between the electron and the nucleus. ADF images can be considered spatially incoherent in nature, and therefore their contrast does not change as a function of defocus whilst bright field images are spatially coherent and hence experience contrast change as a function of defocus [31].

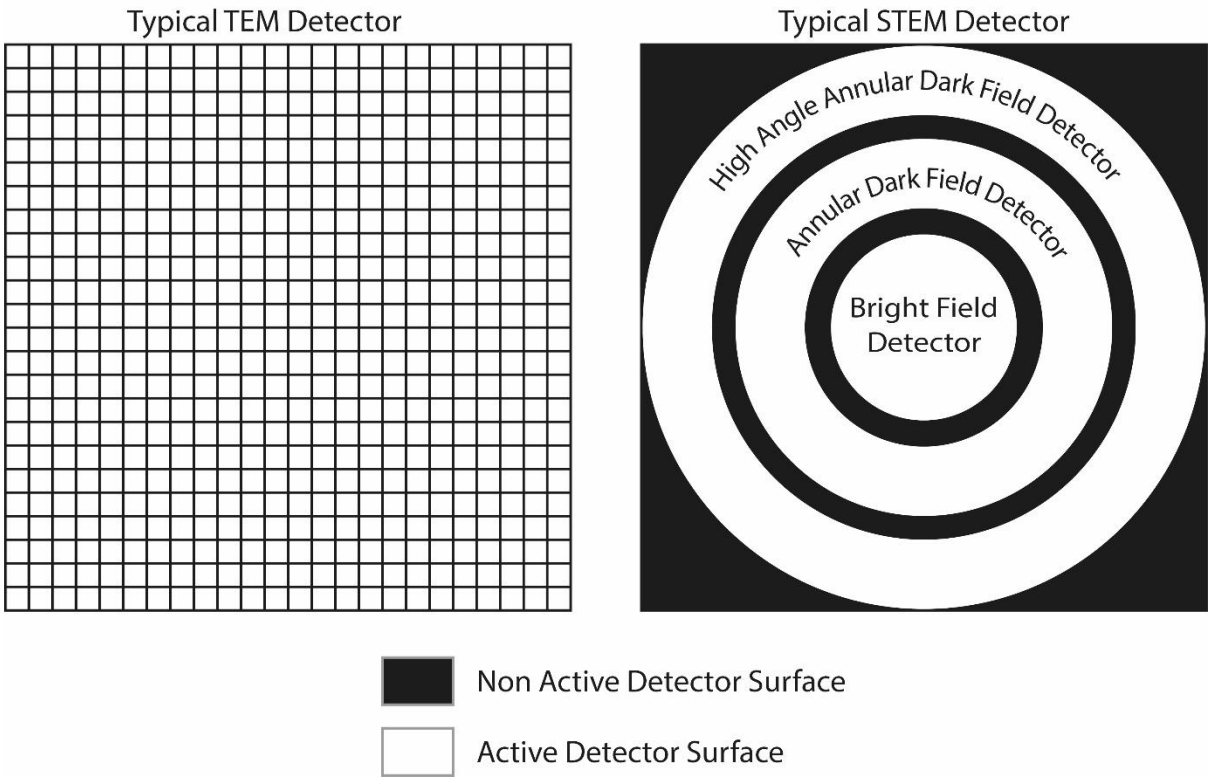

Figure 6: a figure comparing a pixeled TEM detector against annular STEM detectors

A historical reason for the STEM detectors being annular and effectively single pixel is the computational limits on memory space and processing power at the time STEM was first developed (1966 onwards). This is because a STEM image formed from a pixelated detector requires orders of magnitude more memory space than an image from a single pixel detector. If every pixel of a 512 by 512 image required a 256 by 256 pixelated detector image, the total number of pixels used by the final image would be around 17 billion pixels. if one uses an annular single pixel detector however, the number of pixels remains 512 by 512, which most early computers would have had a better chance of dealing with. With modern computing power however, the benefits of using a pixelated detector become obvious, as it enables the collection of BF, ADF, and HAADF data in a single dataset and also enables custom post processing filtering of the STEM dataset (the data can be post-processed an infinite number of times). Pixelated detectors became more attractive for STEM experiments with the development of direct electron detectors, as these have improved resolution as counts were not collected through an intermediate scintillator which applies its own blurring function to the image. A scintillator has an additional blurring function, as when an electron impacts with it, the scintillator produces additional photons which can spread out from the original impact point of the electron affecting multiple pixels of the detector [32]. In some direct electron detectors, the scintillator is replaced by a thick layer of depleted silicon which helps to ensure that the electron is detected, as it generates a large number of secondary electrons boosting the signal [33, 34]. Secondary electrons can still be generated in neighbouring pixels (similar to the scintillator) and are dealt with by post-detection circuitry, which only counts a detection event if it is above a particular number of secondary electrons, therefore reducing the noise in the collected diffraction pattern [35]. Direct electron detectors are also attractive as they have improved read out times, meaning increased frame rates (1100 frames per second for EMPAD [35] compared to 30 frames per second from CCDs [34]) this can be used to prevent the saturation of the detector, as multiple exposures can be combined to obtain the same effect as a longer exposure. There are many more benefits to direct detection pixelated detectors which can be found in the references of this paragraph, but the discussion will now move on to a STEM method called Differential Phase Contrast (DPC).

1.9 Differential phase contrast

Differential phase contrast (DPC) is an established phase imaging method in the STEM [36]. The basic principle of DPC is to direct a focused beam of electrons at a small part of the sample. The sample will then deflect the electron beam depending on its electrical and magnetic properties. DPC shares the same optical setup as STEM, and therefore the C2 aperture defines the electron beam at the detector. The effective image of the C2 aperture at the centre of the CBED pattern moves depending on the deflection strength of the sample. The deflection strength is measured by comparing the shift of the electron beam at the detector to the case where there is no sample. The shift of the C2 aperture is measured by a quadrant detector (see Figure 7), and algorithms calculate the difference between the quadrants and integrate the function to obtain an image whose contrast is proportional to the phase of the specimen [36]. The measured shift is used by a specialized algorithm which can calculate the phase of the sample at the particular beam sample cross-section. In DPC the focused electron beam is scanned across a whole region of interest in a grid like fashion, and by using the calculated phase at each scanning position a whole phase image of the specimen can be obtained. By summing the intensity at each scan point a magnitude image can also be obtained, in common with STEM.

The detector used in DPC has historically been an annular quadrant detector, a detector split into 4 photosensitive detectors; see Figure 7. DPC is very similar to other forms of STEM imaging such as

HAADF imaging, and in some cases uses the same detectors. In most forms of STEM imaging the probe is focused to a very small point compared to the sample, which is due to the size of the beam at the specimen determining the pixel size of the final STEM image. This means that DPC imaging can achieve sub-nanometre resolution when imaging electron fields and light elements. However, its objective lens magnification is limited when the sample is magnetic in nature, as the objective lens is only weakly on or completely off for such a sample [37].

Differential phase contrast historically was the hypothesis of Rose and then was later implemented by Dekkers and de Lang in 1974, where a split detector was used in a STEM to obtain a phase sensitive image of a Siemens star specimen (the split detector was two halves instead of the typical quarters used today) [28]. Differential phase contrast found its main application in examining magnetic materials [13, 28], and DPC has also been used in many situations such as mapping electric fields [38, 39].

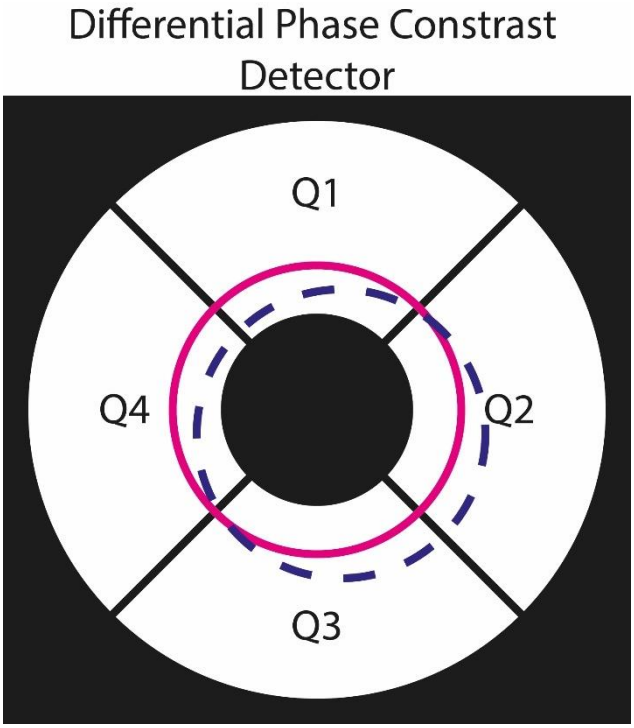

Figure 7: A figure showing a quadrant detector used for DPC. The pink circle indicates an aligned STEM diffraction pattern in a case where there is no specimen in the beam. The blue dotted line indicates where the specimen in the beam has caused the diffraction pattern to shift because of its electric field/magnetic field/ changing thickness.

In order to explain differential phase contrast phase recovering ability, let us consider a hypothetical sample, the thickness of which is constant in one section, and a second section where the thickness changes in a linear manner. The phase change an electron wave would experience when passing through the second section would be akin to a phase ramp.

Now let us consider the STEM operation: first the upper objective/C2 lens images the condenser lens system output [29]. In the back focal plane of the upper objective lens there is an aperture that limits convergence of the illumination (alpha). This aperture is effectively Fourier transformed on to the specimen, the probe and specimen multiply together, the post specimen lens system then forms the diffraction pattern of the sample and probe on the detector by performing another Fourier transform, the detector takes the exit wave and collects its intensity. So if there was no sample in the microscope the diffraction pattern on the detector would be an image of the aperture used to limit the convergence angle.

Now let us consider the same operation in the presence of a half constant – half wedge sample [36]. In the constant part of the sample there would be no change to position of the un-scattered central disc as the beam would only undergo a constant phase delay when it interacted with the sample (there may be inelastically scattered electrons added). However, when the scan coils make the beam interact with the wedge part of the sample, the central disc starts to shift. The sample's wedge subsection is an effective phase ramp, therefore the Fourier transform of the aperture is multiplied by a phase ramp and basic Fourier theory states that any signal multiplied by a phase ramp which is later Fourier transformed will have its Fourier transform shifted relative to the case where no phase ramp was involved. Therefore, all of the diffraction patterns collected in the wedge section of the sample will have the un-scattered central disc shifted compared to the part of the sample which has a constant thickness. Therefore, DPC STEM is sensitive to the change in thickness of the sample. As established earlier electric field and magnetic fields can cause phase changes in electron waves as well, meaning that they can also deflect the un-scattered central disc, and this in turn means that they can be quantitatively measured.

Therefore, using a basic understanding that a phase ramp in real space is a shift in Fourier space, it is clear that a sample that is effectively a phase ramp causes a well defined shift of the aperture function at the detector.

1.10 The phase problem

The phase problem describes the fact that the phase information of a diffraction pattern cannot be directly measured by a detector [5]. This is because the signal measured by detector pixels per capture period is solely determined by the power of photons/electrons that hit them, which is proportional to the electric field squared of the wave to be considered, and so has no phase dependence. This is typically stated in the literature of ptychography as equation 4:

$$I \propto \Psi e^{j\theta}\left(\Psi e^{j\theta}\right)^{*} = \Psi e^{j\theta}\Psi e^{-j\theta} \tag{4}$$

The parameters of equation 4 are the following:

- I is the intensity measured at the detector
- Ψ is the amplitude of the propagated exit wave
- * indicates the conjugate of a complex variable
- θ is the phase of the exit wave.

The problem is that the phase of the exit wave is a crucial component of the wave front, and it must be known precisely in order to obtain an image of the sample that generated the recorded data. The limitation of the detector means that recovering the phase requires the use of some form of phase retrieval mechanism. The word 'mechanism' is used due to the fact that the phase can be obtained by first changing the experimental setup (off-axis holography and inline holography) or/and implementing software specifically designed to work out the phase (ptychography, Gerchberg-Saxton, Coherent Diffractive Imaging (CDI)).

1.11 Off-axis holography

Gabor was a major contributor to the progress of electron microscopy, and was responsible for the idea of holography. He defined holography as the process of capturing the whole of a specimen (both phase and magnitude) from only an intensity image [12]. The first implementation of holography was inline holography, where a defocused specimen wave was collected by photographic paper from an electron microscope and Gabor attempted to reconstruct it using an optical setup, which then corrected for the aberrations and defocus to obtain the true specimen. Unfortunately,

this did not succeed due to the twin image problem of phase retrieval, which is discussed in section 1.18 [12]. It was not until later that off-axis holography was developed by Leith and Upatnieks in the optical domain. By placing a prism before the specimen, they managed to create a spatially separate reference wave which contained no specimen [40]. This reference wave was then forced to interact with the specimen wave at the detector via the action of the prism (as shown in Figure 8). By causing a difference in the approach angle of the reference and the specimen wave, Leith and Upatnieks managed to "disperse" the twin image problem experienced by Gabor. As will be shown in Section 1.28, the angle of the wave with respect to the optical axis is related to a position in frequency space, so by making the reference wave approach the detector in an oblique fashion, the twin image problem is "dispersed" in frequency space [40]. Furthermore, due to the angle between the waves at the detector, the characteristic interference lines which are synonymous with off-axis holography are obtained (a simulated hologram can be seen in Figure 11). Off-axis holography was then later implemented on the electron microscope by using Möllenstedt's work on the electron biprism, in which an electron biprism was used to split the electron beam into the specimen wave and the reference wave in a TEM context (See Figure 9) [41-44]. The biprism is commonly located at the selected area aperture plane (the first image plane after the specimen plane). Off-axis holography's mainstream acceptance was only achieved with the development of more coherent electron sources [42, 45].

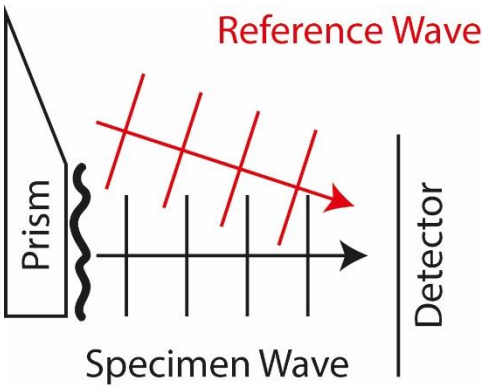

Figure 8: A simple diagram outlining creation of the reference wave and subsequent interaction with the specimen wave used by Leith and Upatnieks [40].

Off-axis holography in TEM works by splitting the electron beam into an object wave which interacts with the specimen, and a reference wave which is left unaltered from the collimated electron source. These waves are then allowed to propagate through the objective lens until they reach the electric bi-prism. Once they reach the bi-prism, the independent waves are then deflected so that they interact with each other on the detector, and the area where the object wave and reference wave interfere forms the recorded hologram. A simplified diagram outlining the setup of off-axis holography can be seen in Figure 9.

A key requirement of electron holography is the use of a highly coherent source of electrons, hence most progress with off-axis holography came in the 90s, when the CFE (Cold-field emission) guns became widely available [41]. The coherent electron sources mainly improve the number and quality of interference fringes in the collected holograms. The coherence of the beam is critical as it is proportional to the sensitivity of the provided phase images, as outlined by equation 5 from [44].

$$\delta\Phi_{min} = \frac{SNR}{\mu}\sqrt{\frac{2}{N \times DQE}} \tag{5}$$

The parameters of equation 5 are the following:

- $\delta\Phi_{min}$ is the minimum difference in phase which can be detected in the reconstructed hologram
- *SNR* is the signal to noise ratio of the collected hologram
- μ is a value describing the degree of coherence of the source
- N is the number of electrons per pixel of the reconstructed hologram
- DQE is detection quantum efficiency of the detector

The field of view of a hologram is related to the charging of the biprism and the effective separation of the electron "sources", meaning that as the desired field of view increases, the damping effect on the fringes from the incoherence of the electron source increases (See Section 1.35 for more on incoherence) [12]. This relationship between the field of view and phase sensitivity (phase sensitivity is proportional to the fringe contrast) means that it is often better to image a smaller field view to obtain the best phase images.

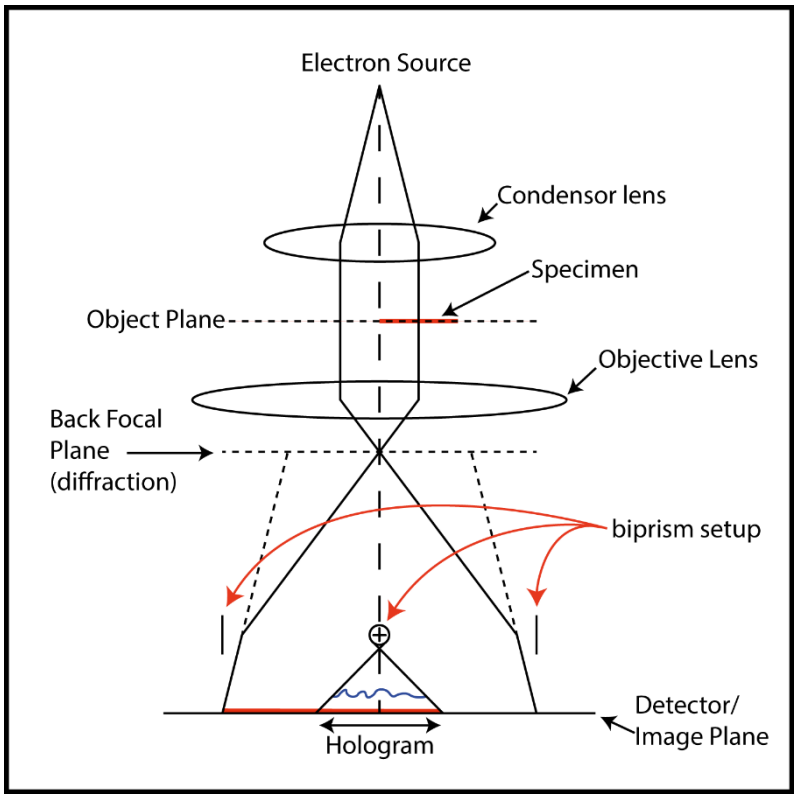

Figure 9: A diagram showing the essential components of off-axis holography implemented in the TEM [44].

The recorded hologram is processed by performing a Fourier transform and cutting out one of the sidebands of the hologram (a side band is one of the dispersed twin images). This process is analogous to decoding in FM radio, where the reference wave acts as a carrier wave, which displaces the important information in the frequency space and this encoding acts to separate the object wave from its complex conjugate (the twin image problem). After cutting out one of the sidebands the aberrations are corrected and the specimen image is refocused via back propagation. Sometimes the reconstructed phase from the hologram contains phase ramps and distortions due to incorrect selection of the subsection (the brightest pixel is not at the centre of the subsection) and aberrations of the projector lens system [12]. Typically, a reference hologram is collected for every hologram in order to correct subsection misalignment and the projector lens aberrations, since reconstructing

the phase of the reference hologram will obtain the reference phase, and the reference will contain these inaccuracies. Therefore, by dividing the reconstruction phase by the reference a clear phase image is obtained. A general outline of the holographic reconstruction process is illustrated in Figure 10. Due to the reconstructed phase coming from a subsection of a larger image, it is typically of lower resolution than that desired by the microscopist, so typically collected holograms are purposely collected at higher resolutions in order to account for this loss of resolution [46]. The downside of this resolution mitigation is that the field of view is reduced and is smaller compared to other techniques under similar experimental conditions.

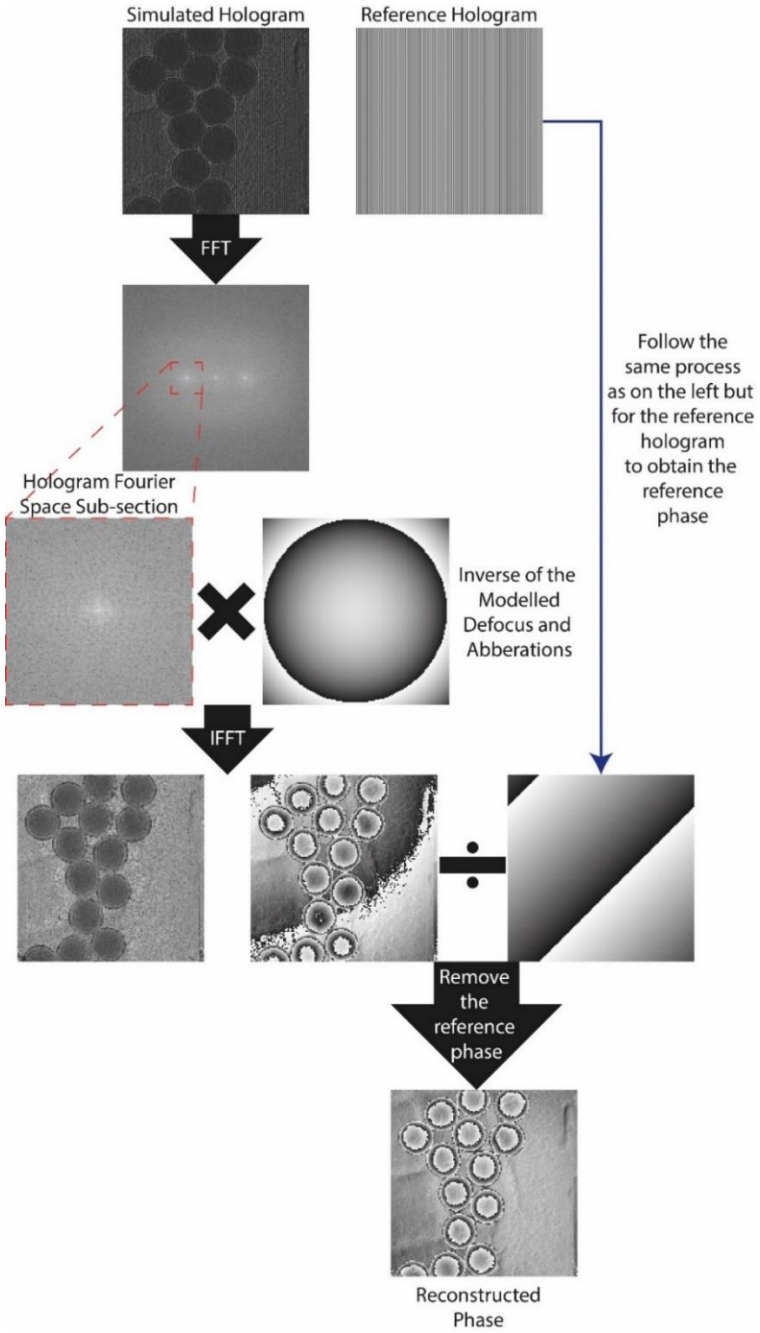

Figure 10: An outline of the reconstruction process of off-axis holography, starting with example holograms and ending with the reconstructed phase of the specimen. Note that data and reconstruction are obtained from a simulation.

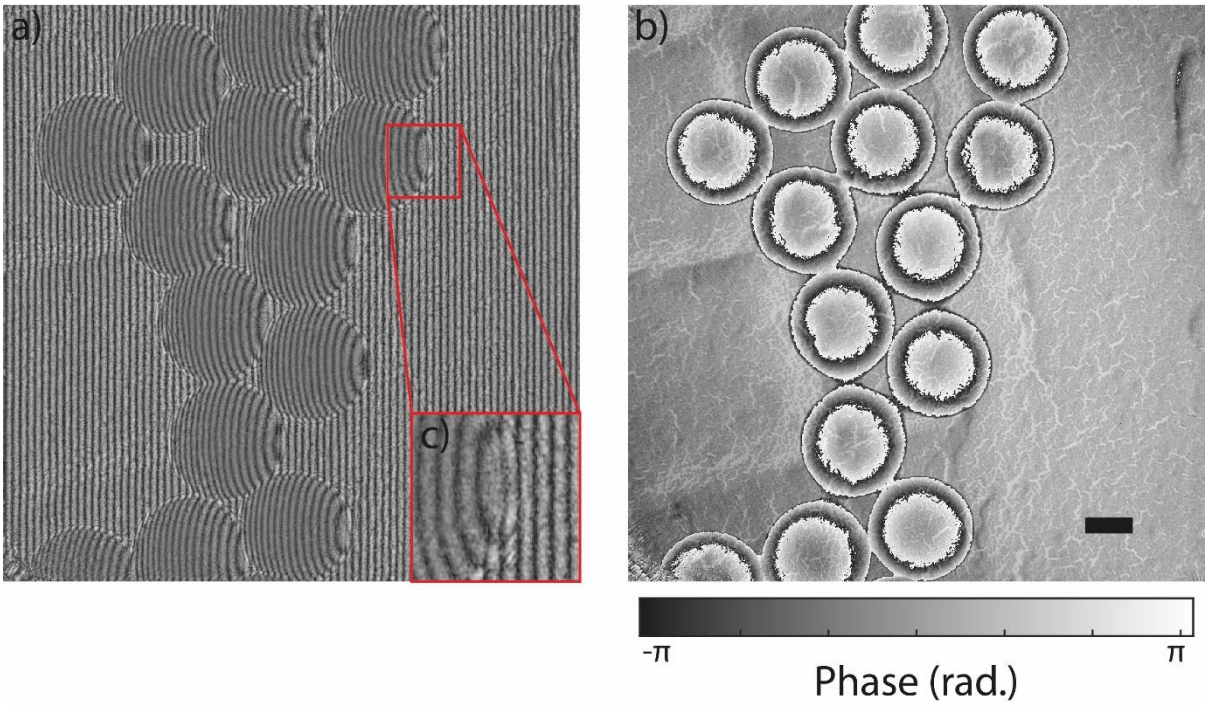

Figure 11: a) A simulated hologram of latex spheres on top of diffraction grating replica. b) the reconstructed phase from a near field electron ptychography experiment used to generate the hologram in a). c) is a subsection of a) showing the bending of the holograms fringe's as the specimen changes from diffraction grating to latex sphere. The scale bar in b) is 100nm and the colour bar coverts the greyscale of b) into radians.

Figure 11 a) shows a simulated hologram. Through inspection of the hologram and comparison of the phase object imaged by the hologram, it is clear in the spatial domain how off axis holography encodes the phase of the specimen. Looking at Figure 11 c) it can be seen that hologram fringes bend when they interact with the strong phase delay caused by the latex spheres, while there is no bending of the fringes in the diffraction grating replica next to the latex spheres. Even though Figure 11 a) is an intensity image, phase changes have affected the collected intensity obtained by the detector. An aspect common to off-axis holography, inline holography, ptychography and near field ptychography, is that phase changes of the specimen have changed the magnitude at the detector in order that the phase of the specimen can be reconstructed.

The complex-valued information contained in the hologram results from the object of interest and the transfer function of the lens, so if the lens transfer function is known the aberrations of the lens can be suppressed in post processing.

1.12 Inline holography

Another method of phase retrieval in electron microscopy is inline holography, also known as through focal series reconstruction, and the transport of intensity equation. Electron inline holography was first described by M. E. Haine and T. Mulvey in a paper called "The Formation of the Diffraction Image with Electrons in the Gabor Diffraction Microscope" [46, 47]. This method is simpler to implement than off axis holography, as it does not require any additional components, only that the in-focus image from the microscope can be defocused in some way, usually by adjusting the objective lens strength. This alteration of the lens function enables access to different parts of the spatial frequencies of the specimen function which is unique to that defocus. In principle, by collecting multiple defocused images at different de-foci, a stack of images is obtained which as a whole contains all the phase information of the sample (or as much information, as the aberrations of the microscope allow, primarily chromatic aberrations). This stack of images is then

used in a reconstruction algorithm to form a more complete description of the object, which is unaffected by the aberrations of the electron lens. There are a few algorithms by which the object can be reconstructed, but only two will be mentioned here: The Transport of Intensity Equations (TIE) [48-51] and the Maximum Likelihood (MAL) algorithm [52, 53]. Further references to different algorithms can be found in [54].

Although inline holography is less difficult to implement in the microscope, it has disadvantages. For example, the collected data requires more pre-processing before being input to the TIE and MAL algorithms, such as corrections for magnification and rotation, which all change as a function of defocus in the electron microscope. Furthermore, as each diffraction pattern emphasises different spatial frequencies the phase reconstructed from each diffraction pattern must be filtered such that optimal spatial frequencies contribute to the final phase and any noise is minimised. Some implementations of inline holography are limited to applications which are weak phase objects, which is not the case for ptychography [55-58] . However, inline holography can image absorptive specimens under the slowly varying object approximation [55].

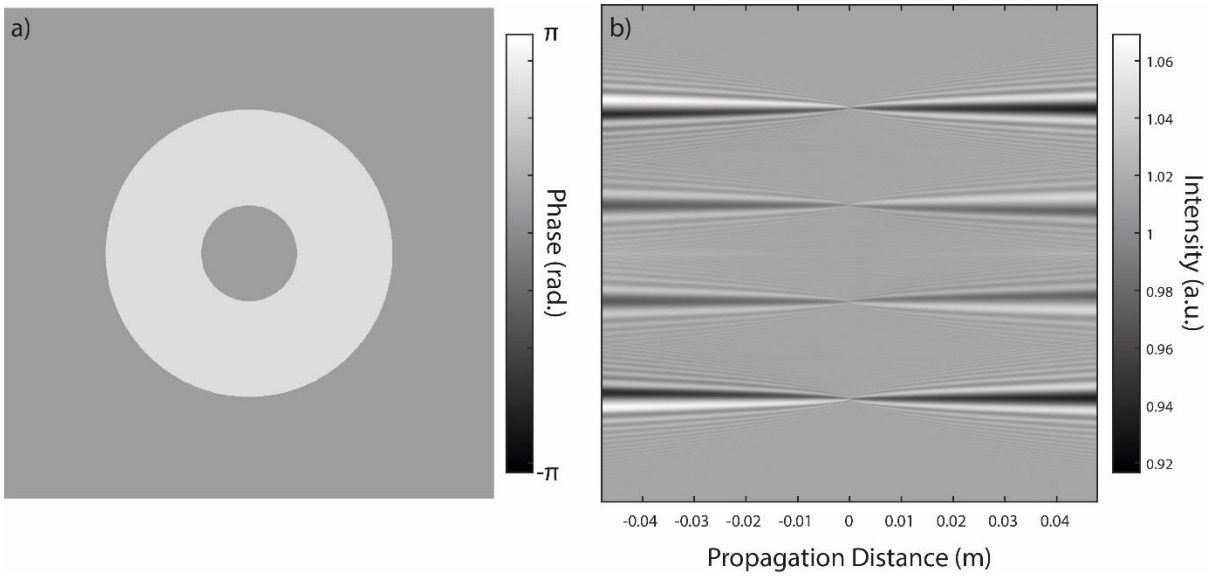

Figure 12: a figure demonstrating how the propagation/defocus distance between a phase only specimen and the detector causes the collected intensity to change in a predictable fashion. a) is the simulated phase specimen used to generate b). b) shows how the collected intensity at a detector changes as a function of the propagation distance between the specimen and the detector. The colour bar to the right of a) converts the greyscale of a) into radians. The colour bar to the right of b) converts the greyscale of b) into intensity with arbitrary units.[55]

Figure 12 a) shows a phase only object. In order to form the illustration seen in Figure 12 b), Figure 12 a) was propagated various distances, and the intensity was collected at each distance and averaged along the y-axis [55]. The average intensities were taken and placed side by side, resulting in Figure 12. Figure 12 b) illustrates how the intensity varies as a function of propagation distance. For reference, the wavelength used in the simulation shown in Figure 12 was 635nm. The transport of intensity equations effectively uses the fact the intensity at a particular defocus can be traced back to the specimen to recover the phase of the specimen. Figure 12 a) and b) show that only changes in phase can influence the collected intensity of the defocused image of a specimen, as where there is no phase change the intensity does not vary. Figure 12 also illustrates why the simplest method to obtain contrast from phase only objects has always been to operate the microscope with a slight defocus.

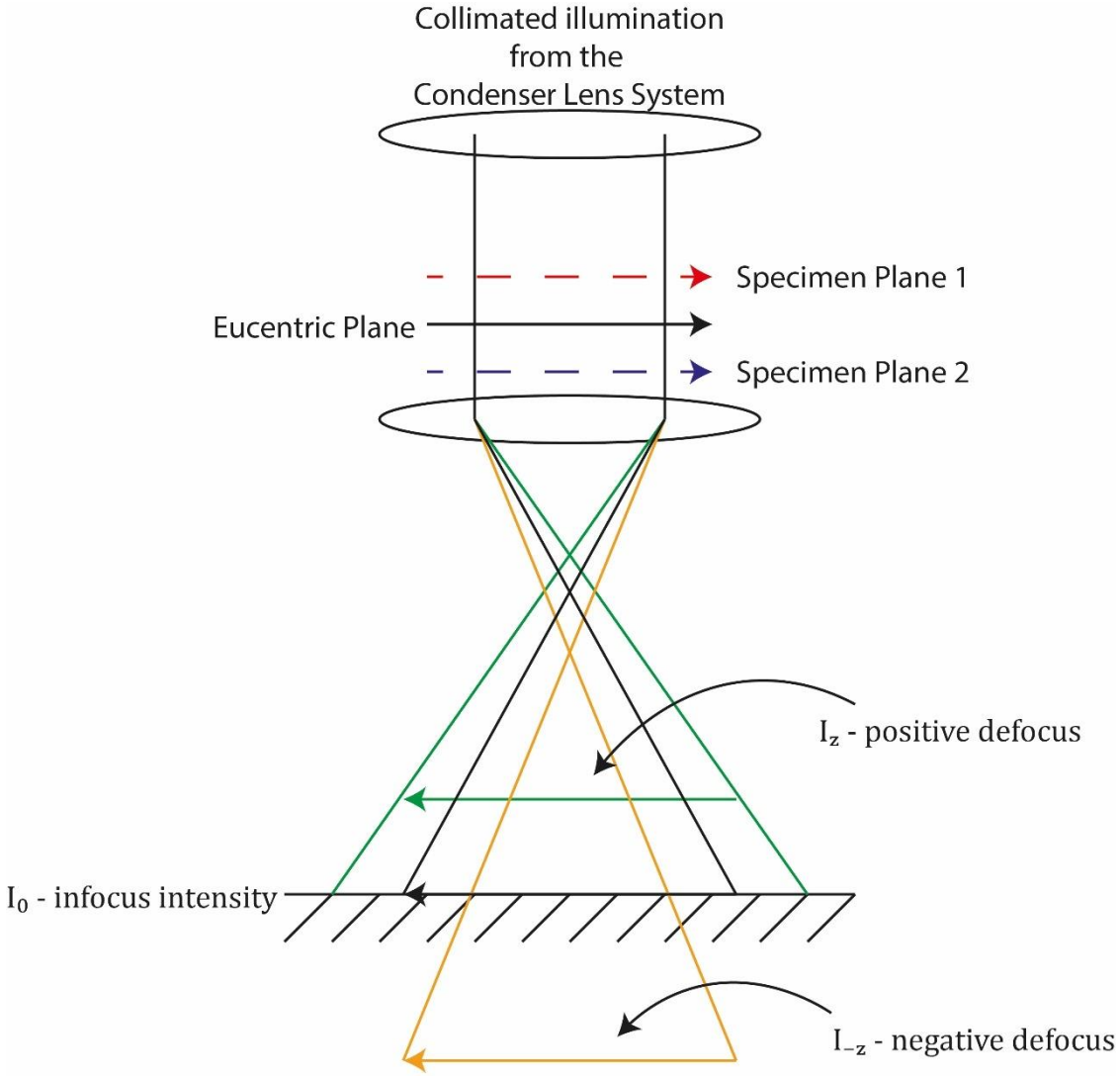

Figure 13: An illustration showing how the defocus condition is realised in TEM for inline holography data collection [46]

Figure 13 shows a typical setup used to perform inline holography. The defocus required by inline holography can be achieved by either moving the specimen with respect to the eucentric plane as indicated by specimen planes 1 and 2 in Figure 13 (red arrow negative defocus, blue arrow positive defocus) [46]. Alternatively, the positive and negative defocus can be implemented by changing the strength of the objective lens; the positive defocus is denoted by green rays (stronger objective lens) and negative defocus is denoted by the orange rays (weaker objective lens). Both the illumination and the output of the objective lens need to be collimated so that either the detector or the sample can be moved to create the desired defocus condition without suffering additional magnification from the curvature of the beam. If the beam has curvature at the specimen or the detector, the effective magnification will change as a function of defocus, adding an additional step to the phase retrieval algorithm.

While the setup of inline holography is similar to typical microscope setups making it easier to implement, the reconstruction process is more involved than that of off axis holography. The transport of intensity equation solves the phase problem by solving partial differential equations. The phase problem can be considered as such since the rate of change of the collected intensity with

respect to defocus can be considered proportional to the first and second differential of the phase of the sample, as shown in equation 6 [55].

$$-k\frac{dI}{dz} = \nabla \cdot (I_0 \nabla \varphi) \tag{6}$$

The parameters of equation 6 are the following:

- k is the wave number, $\frac{2\pi}{\lambda}$
- $\frac{dI}{dz}$ is the gradient of the intensity with respect to change in distance along the optical axis
- I_0 is the intensity of the in-focus specimen
- ∇ is the gradient operator
- φ is the phase of the specimen

However, the simpler situation is obtained when the specimen is a phase only specimen; see equation 7:

$$-k\frac{dI}{dz} = \nabla^2 \varphi \tag{7}$$

The mathematical description shown in equation 6 has been interpreted such that it can be implemented by various software packages, so the focus of this introduction to TIE will be in describing how the problem outlined in equation 7 is solved using a computer. The second order differential in equation 7 can be inversed via several Fourier transforms and division by the radius of the Fourier space of the collected data. The equation for the inverse second order differential is shown in equation 8:

$$\nabla^{-2}\left(\frac{dI}{dz}\right) = IFFT\left(\frac{FFT\left(\frac{dI}{dz}\right)}{4\pi^2|u|^2}\right) \tag{8}$$

The parameters of equation 8 are the following:

- ∇^{-2} is inverse Laplacian operation
- FFT is the fast Fourier transform
- $IFFT$ is the inverse fast Fourier transform
- u is coordinate in frequency space related to effective detector dimensions used to collect the data

Furthermore, the rate of change of the intensity with respect to the propagation distance is approximated by subtracting a negative defocus intensity from positive defocus counterpart (see equation 9). Although there are other ways of doing this, they are beyond the scope of this introductory section [59].

$$\frac{dI}{dz} \approx \frac{I_Z - I_{-Z}}{2\Delta z} \tag{9}$$

The parameters of equation 9 are the following:

- I_Z is the collected intensity at an arbitrary positive propagation distance (defocus)
- I_{-Z} is the collected intensity at a negative propagation distance equal to the positively propagated intensity (defocus)
- z is the distance between the defocused image plane of I_z and the in-focus specimen plane

Using equations 8 and 9 together, the complex field of a phase only specimen can be recovered from several oppositely defocused inline holograms, as shown in equation 10.

$$\varphi = -k\nabla^{-2}\left(\frac{dI}{dz}\right) \tag{10}$$

Both inline holography and High Resolution Transmission Electron Microscopy (HRTEM) are related to the concept of the phase contrast transfer function (PTF). The PTF describes what spatial frequencies are captured from a defocused image of the sample. The simplified PTF which will be described in this section relies on various assumptions, specifically the phase only assumption and weak phase assumption [6]. The phase only assumption assumes that the specimen is so thin that the sample does not significantly attenuate the electron beam. The weak phase approximation assumes that the specimen is so thin that if the specimen was described by a complex exponential, the phase delay it would cause is such that the power of the complex number is equal to an imaginary number if it was converted to Cartesian coordinates (see equation 11).

$$\mathrm{e}^{j\theta} \approx 1 + \theta j \text{ if } \theta < 1 \tag{11}$$

PTF functions are important to HRTEM as they help the microscopist determine whether a particular defocus will enable them to observe the atomic columns of their specimen, and whether the atomic columns are bright or dark spots, since the only contrast that is available when looking at the atomic level is phase contrast (assuming the atomic number of the specimen is not too high).

Jingshan et al. state that the Fourier transform of an inline hologram can be described by the following equation [50]:

$$I(f_x, f_y, z) = \delta - 2U(f_x, f_y)\cos\left[\pi\lambda\left({f_x}^2 + {f_y}^2\right)z\right] - 2P(f_x, f_y)\sin\left[\pi\lambda\left({f_x}^2 + {f_y}^2\right)z\right] \tag{12}$$

The parameters of equation 12 are the following:

- f_x and f_y are the positions along the x-axis and y-axis of frequency space respectively
- z is the defocus to consider or the distance along the optical axis
- δ is a function which has value of one at origin and zero everywhere else
- λ is the wavelength of the accelerated electrons
- U is the function which describes the attenuation of the specimen
- P is the function which describes the phase of the specimen

If the specimen is assumed to be a phase only object, then the equation can be further simplified to:

$$I(f_x, f_y, z) = \delta - 2P(f_x, f_y)\sin\left[\pi\lambda\left({f_x}^2 + {f_y}^2\right)z\right] \tag{13}$$

This suggests that the intensity varies sinusoidally in frequency space and the rate of change of the sinusoid is determined in part by the propagation distance between the specimen and the detector, as the terms within the brackets of the sin term is the defocus propagator (See Section 1.28). Therefore, the propagation distance can be used to vary the spatial frequency content in the hologram. However, as the PTF is determined by a sinusoid there are zero crossings where that spatial frequency will not be transferred to the inline hologram, meaning the inline hologram has no information on that frequency. Therefore, in inline holography experiments defocus conditions are typically chosen such that the contrast transfer function never crosses through zero. Examples of how the intensity of inline holograms can vary in frequency space are shown in Figure 14. The disadvantage of not having a zero crossing in inline holography is poor low frequency response in the

recovered phase, as can be seen from Figure 14, as the sinusoid varies so slowly that the contrast transfer function stays close to zero for a significant amount of frequency space compared to larger defocus conditions. This means that the signal to noise ratio for this phase is reduced. This is one of the reasons why multi-focal transport of intensity equations were developed, as additional defocused conditions allow not only for the low frequencies to be recovered better by collecting an inline hologram with a large defocus [60], but de-foci between these extremes ensure all spatial frequencies can be recovered [51] (within the limits allowed by the chromatic aberrations).

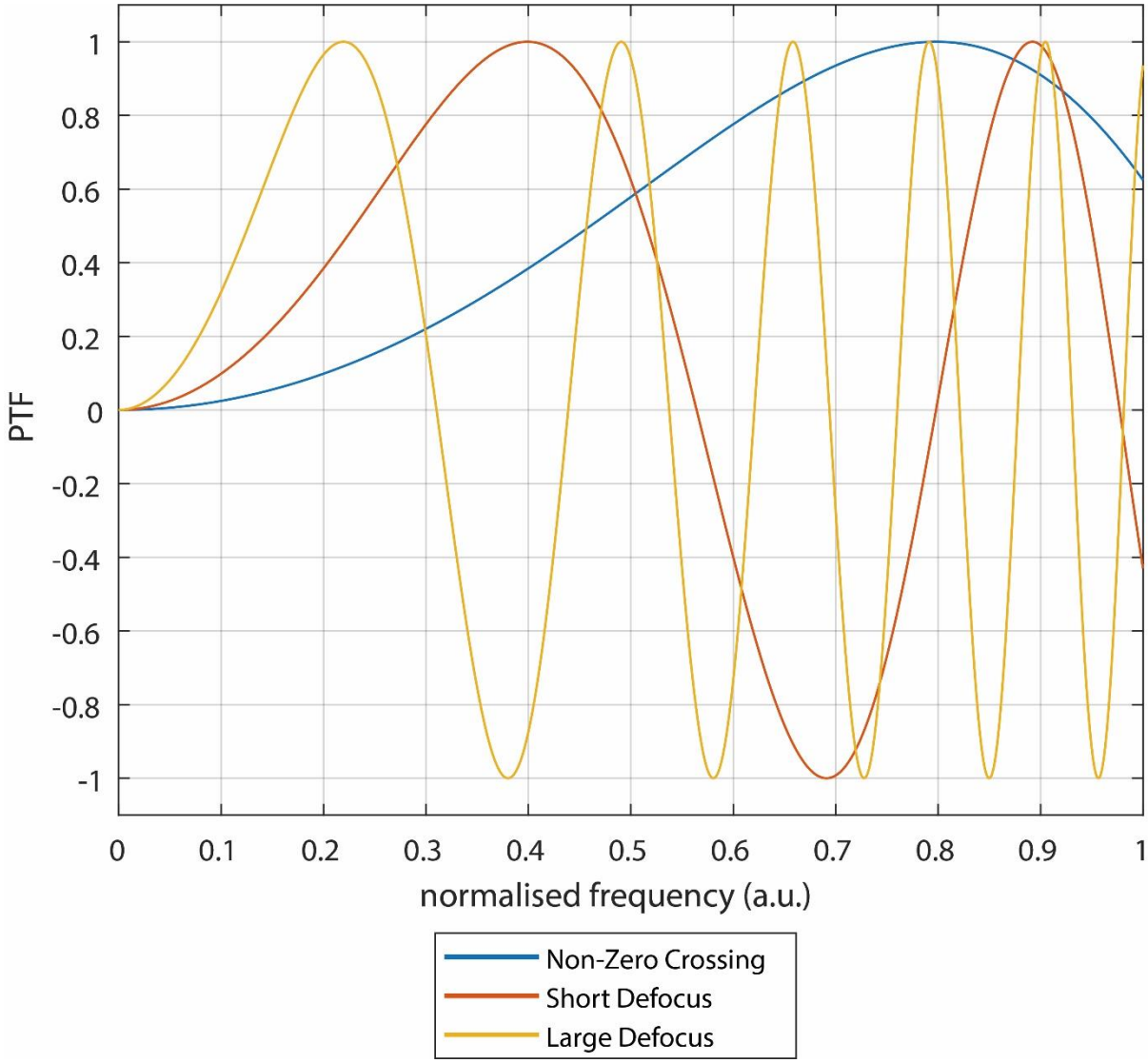

Figure 14: an illustration showing how inline holography constrast transfer function can vary as a function of defocus.

Figure 15 shows a more realistic phase transfer function, as this includes the effects of spherical aberrations and the chromatic aberration envelope which both limit and change the PTF. The chromatic aberration (and its envelope) in

Figure 15 is defined by the focal spread of a lens at its focal point which is due to many factors such as: the variance of the high tension voltage supply accelerating the electrons from the electron gun, the variance of the objective lens current and variance of electron energies emitted from the electron gun as described by Freitag et al. [10]. The equations used in the work by Freitag et al. are repeated here for convenience (in the one dimensional form, see equations 14 through 16). Current literature states that the chromatic aberrations imposes an information limit on the maximum

resolution obtainable; this information limit is traditionally defined as the point by which the chromatic envelope reaches e^{-2} or 0.136 [61].

$$PTF = E_{cc}\sin\left(\pi z\lambda {f_x}^2 + 0.5\pi C_s\lambda^2 {f_x}^4\right) \tag{14}$$

$$FocalSpread = \Delta F = C_c\sqrt{\left(\frac{\Delta E}{E}\right)^2 + \left(\frac{\Delta V}{V}\right)^2 + \left(\frac{2\Delta I}{I}\right)^2} \tag{15}$$

$$E_{cc} = \exp\left(\frac{\lambda^2\pi^2\Delta F^2 {f_x}^4}{2}\right) \tag{16}$$

The parameters of equations 14-16 are the following:

- E_{cc} is the effective envelope function of the chromatic aberration
- C_s is the spherical aberration constant of the objective lens
- C_c is the chromatic aberration constant of the objective lens
- ΔF is the focal spread caused by the chromatic aberration in combination with the energy spread $\frac{\Delta E}{E}$, the acceleration voltage instability $\frac{\Delta V}{V}$ and objective lens current instability $\frac{\Delta I}{I}$.

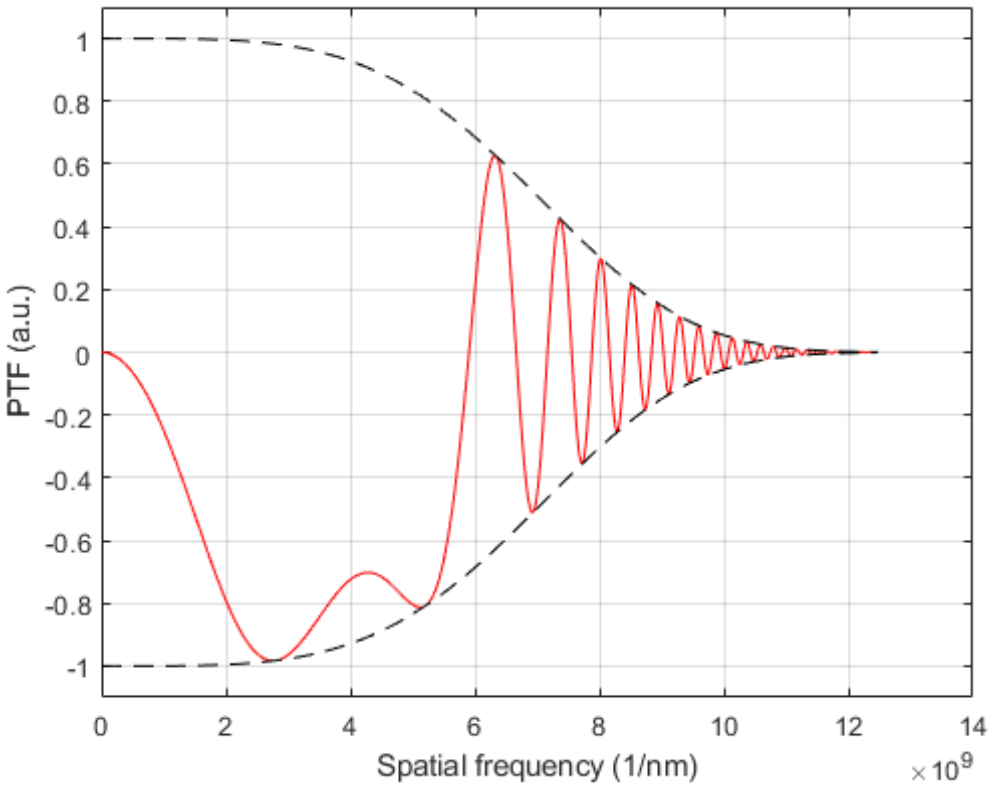

Figure 15: shows a PTF whilst operating in the Scherzer defocus condition with added consideration of the chromatic aberration envelope. The PTF is represented by the red line in the plot while the black dotted line represents the chromatic aberration envelope.

The parameters used in the simulation which generated

Figure 15 are the following: the acceleration was voltage 200keV, the chromatic aberration constant was 1.2×10^{-3}, the spherical aberration constant was 0.3×10^{-3}, the defocus between the specimen and in-focus plane was 33nm, the simulated pixel size was 0.04nm, the energy spread instability was

1.5×10^{-6}, the voltage instability was 0.5×10^{-6} and the lens current instability was 1×10^{-6}. Note that is an example rather that a realistic simulation.

1.13 A short history of coherent diffractive imaging

In this section there will be a general discussion on the constraints of phase retrieval systems throughout history, and related algorithms. The constraints are the diffraction pattern constraint, the support constraint and the overlap constraint (although this will be discussed in the next section on the theory of ptychography). The constraints ground the simulations required to solve the phase problem in reality, and ultimately enable the phase problem to be solved. As coherent diffractive imaging (CDI) has progressed, more constraints have been defined and applied, improving the reconstruction quality.

For coherent diffractive imaging (CDI), the software is used to effectively simulate the experiment again without the limitations of a real detector, which cannot collect phase information. Such a simulation requires diffraction patterns (propagated exit waves) and experimental parameters. Examples of experimental parameters are the distances between the various optical components, and the size of the detector and illumination wavelength. Solving the phase problem with a single diffraction pattern is a difficult task due to the many possible ambiguous solutions. Therefore, it is a common practice among phase retrieval methods to impose as many constraints as possible in order to reduce the number of possible solutions and to increase the quality of the final reconstructed image. The following sections describe the constraints and algorithms that have been investigated.

1.14 The Gerchberg-Saxton algorithm

Firstly, the diffraction intensities impose a constraint as the simulated exit waves must always have the same magnitude as the square root of the collected diffraction intensities [62]. To impose this constraint, the magnitude of the simulated exit wave is removed and replaced with the diffraction intensity, as shown in equation 17.

$$\varphi(x_1, y_1) = {D_z}^{-1}\left(\frac{\sqrt{I(x_2, y_2)}}{|D_z[\psi(x_1, y_1)]|} \times D_z[\psi(x_1, y_1)]\right) \tag{17}$$

The parameters of equation 17 are the following:

- $\varphi(x_1, y_1)$ is the adjusted exit wave
- $\psi(x_1, y_1)$ is the guess at the exit wave
- D_z is the propagation method which best fits the experiment and ${D_z}^{-1}$ is the inverse propagation action, although $\mathcal{F}$ is used when the propagation method can be modelled by the Fourier transform
- $I(x_2, y_2)$ is the collected diffraction pattern from the experiment
- x_1 and y_1 describe x/y axis in the specimen plane
- x_2 and y_2 describe x/y axis in the detector plane

Furthermore, the collected diffraction pattern enables the simplest form of progress tracking of the phase retrieval software. The progress tracking takes the form of a comparison between the intensity of the simulated propagated adjusted exit wave and the measured diffraction pattern. The equation for the diffraction intensity error metric is given below in equation 18:

$$\varepsilon = \sqrt{\frac{\sum_x \sum_y ||\Phi(x_2, y_2)|^2 - I(x_2, y_2)|}{\sum_x \sum_y I(x_2, y_2)}} \tag{18}$$

The parameters of equation 18 are the following:

- ε is the diffraction intensity error metric
- $\Phi(x_2, y_2)$ is the adjusted exit wave in the same domain as the measured intensity data
- Σ_x and Σ_y is shorthand for summing across the x and y dimensions

The diffraction pattern constraint is an important part of every iterative phase retrieval algorithm. In their paper "A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures", Gerchberg and Saxton were the first to describe the interplay of the Fourier transform and the diffraction pattern constraint [63]. Summarising their points: whenever a complex field's magnitude or phase changes then both the phase and the magnitude of the Fourier transform will change. Therefore, the diffraction pattern constraint is the main mechanism via which new guesses at the object's complex field is improved upon, as the applied magnitude data in the Fourier domain causes a new guess at the exit wave in the spatial domain and vice versa. The first iterative phase-resolving algorithm was the Gerchberg-Saxton (G&S) algorithm, which required two measurements: a diffraction pattern recorded in the Fourier domain and an image recorded in the spatial domain. These measurements impose two corresponding constraints that are applied iteratively by the G&S algorithm. The algorithm also requires an arbitrary starting phase (marked initial guess in Figure 16). Through reapplying the same constraints over and over again the G&S algorithm iteratively approaches the correct solution for the value of phase in the spatial and Fourier domains. This is the case, because when the simulated exit waves intensity is removed and replaced with a measured value, the incorrect magnitude values are corrected, which are partly the result of incorrect phase terms in the Fourier domain. Therefore, through magnitude correction in the spatial domain there is an improvement in Fourier domain phase approximation and this is also true of frequency domain magnitude correction and approximation of phase in the spatial domain. A general outline of the algorithm is given below in Figure 16.

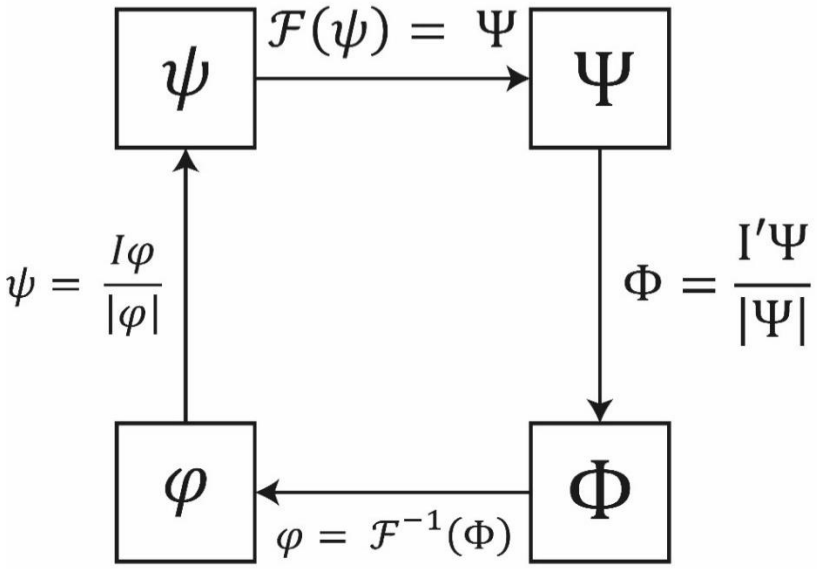

Figure 16: outline of the Gerchberg and Saxton algorithm – where $\mathcal{F}$ is the Fourier transform

In *Figure 16* parameters used are the following:

- ψ is the modelled exit wave in the spatial domain
- Ψ is the modelled exit wave in the frequency domain
- I' is the collected intensity in the frequency domain
- Φ is the updated model of the exit wave in the frequency domain
- φ is the updated model of the exit wave in the spatial domain
- I is the collected intensity in the spatial domain
- $\mathcal{F}$ is the Fast Fourier Transform

Now that the Gerchberg-Saxton (G&S) algorithm and diffraction pattern constraint (information constraint) has been understood, its positive and negative aspects will be put under the spotlight. The G&S algorithm was one of the first iterative algorithms to successfully solve the phase problem, and therefore influenced all proceeding algorithms. However, the G&S algorithm has many negative aspects, for example, it can take thousands of iterations to come to a useful solution, it often gets stuck in local minima, and requires an image of the sample with sufficient contrast, so it cannot be applied to X-ray or electron domains easily. However in the context of the 1970s when it was developed, computing power was very much lower than it is today, so it did well to get a valid reconstruction at all.

1.15 The support constraint and the hybrid input output algorithm

The next development in phase retrieval techniques was applying the support constraint, which removed the requirement from the G&S algorithm of having spatial domain data of the object being imaged [64-66]. The support constraint is particularly useful in the cases of electron and X-ray microscopy, where the transmission coefficients of objects are high, which make it difficult to obtain data on what the transmission value is. Therefore, it is difficult to collect I in Figure 16, which makes the Gerchberg and Saxton algorithm inapplicable to X-ray and electron domains. The support constraint works on the idea of predefined knowledge of the value of the exit wave of the object at particular locations. For example, the values can be predefined by an aperture, in which case parts of the exit wave outside of the aperture are defined to have zero value, as shown in Figure 17 b). Incorrect guesses of the phase in the Fourier domain cause there to be intensity in the spatial domain in the region outside of the applied support, but because we have knowledge of the support we can remove these values, as we know they are physically impossible. This works in tandem with the findings from G&S on the properties of the Fourier domain constraint. When values which are in violation of the support constraint are removed, it causes a change in the magnitude in the spatial domain which then leads to a change in the magnitude and phase in the Fourier domain. As incorrect values are removed it implies that the new phase approximation in the Fourier domain should be improved compared to the previous iteration. This development was made by Fienup, and can be found in his paper published in 1978, where the objective was high resolution imaging of the sun despite the effects of changing atmospheric conditions [65]. The sun had an effective support constraint, in that there was a region around the sun where no intensity was collected. In terms of software implementation, this constraint is applied through inputting the diffraction patterns and guessing the initial size of the aperture.

This was not the only development Fienup provided. He also improved the algorithm by which the phase problem is solved, in the form of the hybrid input and output (HIO) algorithm [66]. The update function HIO is formed from the updated version of the object like the G&S algorithm, but the area outside of the support constraint was also updated by a form of negative feedback. The negative feedback applied was such that pixels which did not conform to the support constraint were

changed such that they did not break it again. The algorithm can be seen in Figure 17. Although the feedback was negative and applied to the aperture region instead of the object region, it ensured all future iterative algorithms would implement feedback as well.

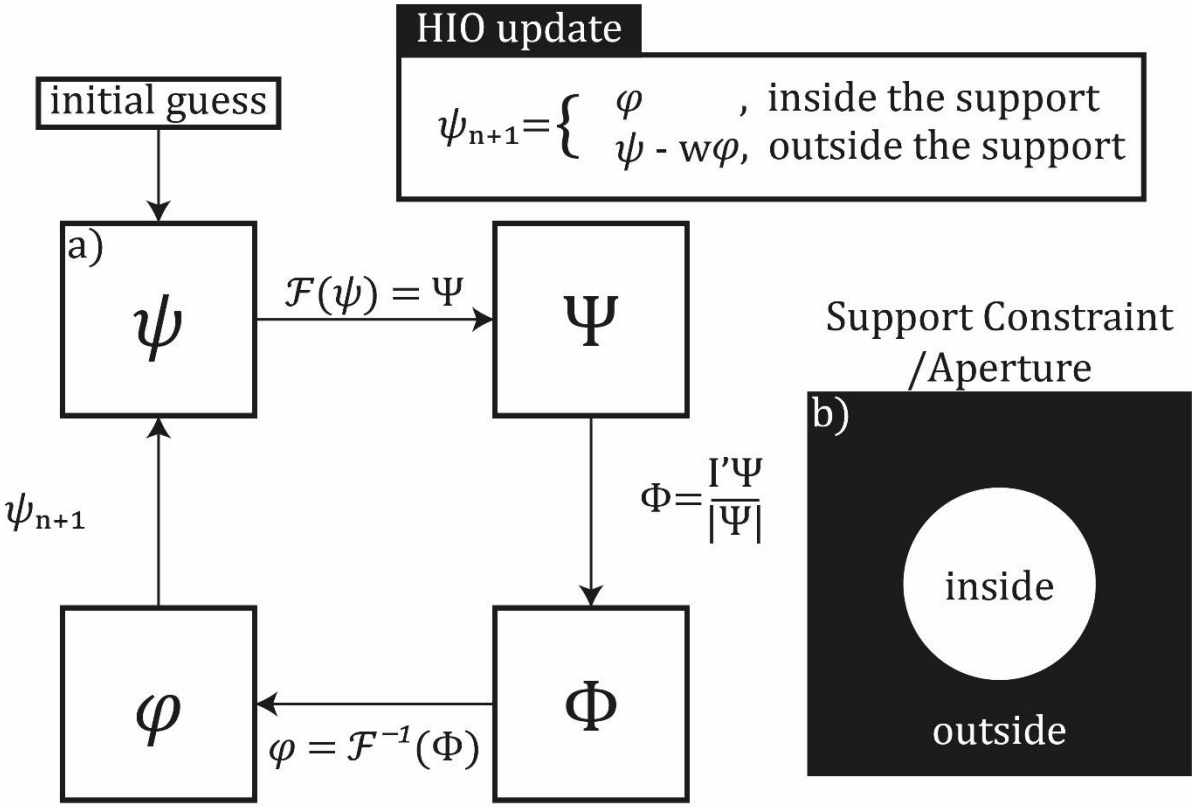

Figure 17: a figure illustrating the hybrid input and output algorithm; mathematical terms are shared with Figure 16 with the addition of w which is used to modulate the strength of the feedback. The HIO update is given at the top of the figure showing how to obtain an updated ψ from φ.

The use of the support constraint in the HIO algorithm meant it improved upon the G&S algorithm by removing the necessity to have a spatial domain image of the sample, which allowed phase retrieval via iterative algorithm to be applicable to the X-ray and electron domains of microscopy. Additionally, the negative feedback of the HIO algorithm reduced both the likelihood of stagnation of the algorithm in incorrect local minima and the total number of iterations required to solve the phase problem satisfactorily [67]. Nevertheless, the HIO algorithm was still not ideal, because the cost of the support constraint's additional stability and success was that the reconstruction had to have a limited field of view, as seen in Figure 17 b). Another difficulty of the support constraint is the required knowledge of the size of the support, such that the experimental support matches the support simulated in the software, incurring additional experiment and program run time.

1.16 Ptychography

Ptychography was the next major development after the support constraint and Fienup's HIO algorithm. Ptychography uses all previous constraints and a new constraint, the overlap constraint. But first there will be a description of ptychography, as a physical setup, a description of the overlap constraint, and then later a description of the Ptychographic Iterative Engine (rPIE algorithm).

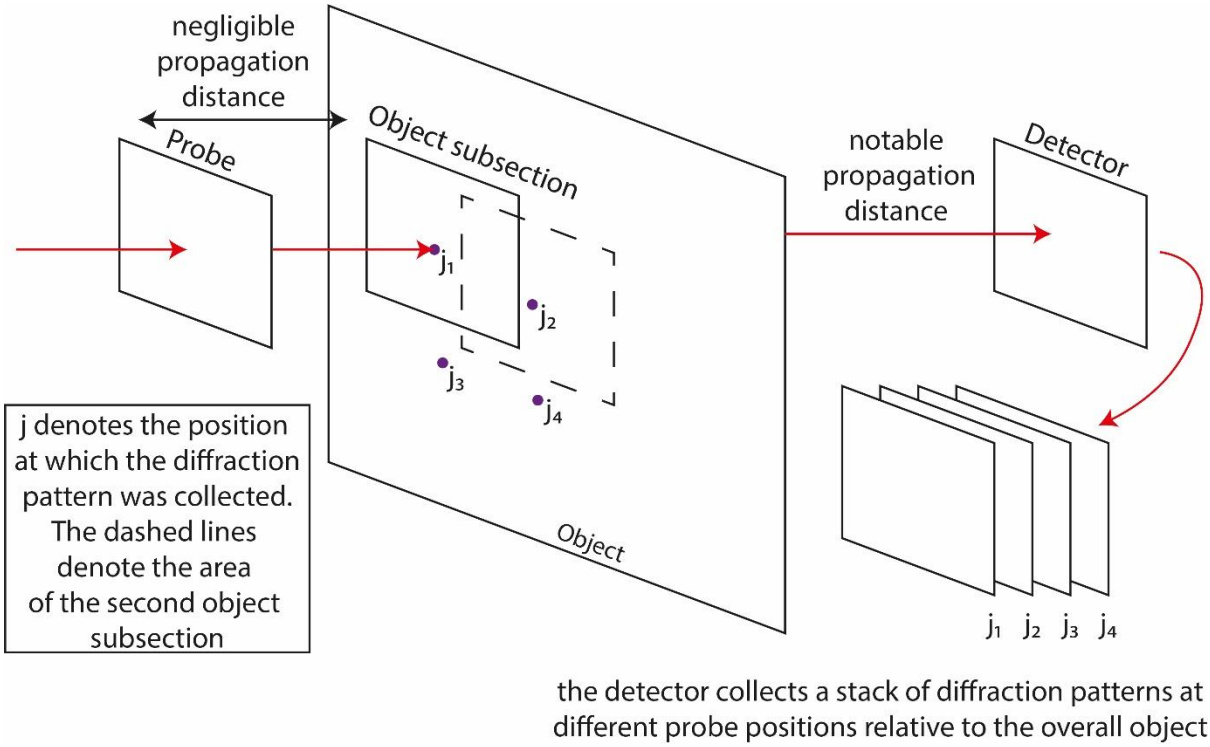

Figure 18: A diagram outlining the data collection process of a ptychographic experiment

In Figure 18 the probe is assumed to be illuminated by a source of coherent and monochromatic light or electrons. The probe in Figure 18 can be the following: an aperture in order to enforce the support constraint, as discussed in the previous section 1.15, or a diffuser in the case of near field ptychography (this will be discussed in section 1.19). The light/electrons which have interacted with the probe then propagate further to the object illuminating a specific object subsection. This gives a specific exit wave after the object which, provided the sample is sufficiently thin (the projection approximation) [68], is a multiplication of the probe's complex field and the object's complex transmission. This exit wave is then propagated to the detector using one of the propagation methods discussed in section 1.25 - traditionally a Fraunhofer propagation to form a diffraction pattern at the detector. Ptychography in effect is multiple Hybrid Input Output algorithms all occurring over the same object, with overlaps in the regions of operations of the algorithms (note the object subsection and dashed lined box overlapping it) [69]. Either the object or the illumination can be translated to realise ptychography. Translation of the object is done with a piezoelectric stage/stepper motor/ linear motor, which also records the position values for each collected diffraction pattern, which is another requirement of ptychography. The probe is assumed to remain the same for all collected diffraction patterns. Furthermore, remembering that any movement of the object can be modelled as a reciprocal movement of the entire optical setup, collected positional data may be referred to as the probe position in the literature. The key aspect which improves ptychography compared to the previously discussed techniques is the overlap in area from which diffraction data is collected, as this forms an additional constraint. Such is the strength of the overlap constraint that it is no longer a requirement to have knowledge of the probe/support before beginning the reconstruction process [62]. This means that more complex probes can be used. The probe, like the object, is recovered by a reconstruction algorithm [62].

1.17 Experimental setup of ptychography

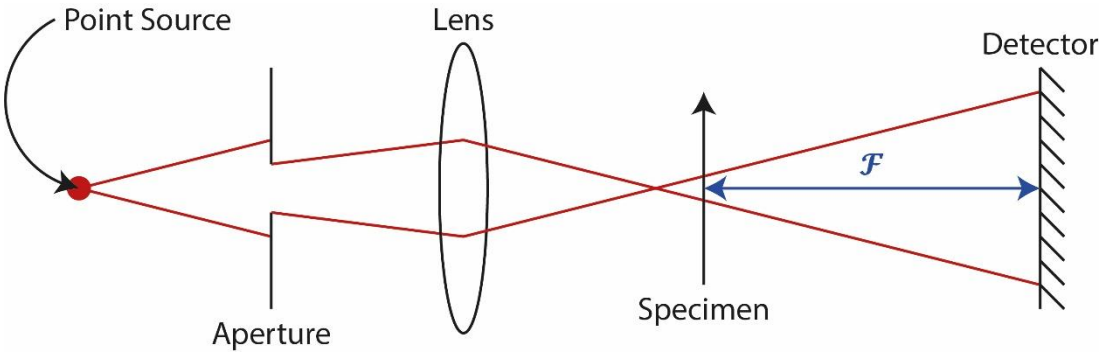

Figure 19: a diagram of simplified ptychographic setup. (Defocused probe ptychography) [70]

There are many possible ptychography setups (defocused probe), but in order to keep the introduction brief, this subsection will focus on the setup shown in Figure 19. Starting at the left of the figure, a point source illuminates an aperture. The point source in a physical experiment would typically be the focal point of a previous lens (or other component such as a mirror in X-ray ptychography). The job of the aperture is to add diversity (diversity in this case means the interaction of the diffraction fringes of the specimen and the aperture where they spatially overlap, see section 1.20 for more on diversity and its definition within this thesis) to the beam when it reaches the specimen and to limit the spatial extent of the exit wave (formed by the sample and the probe), such that when the exit wave is Fourier transformed it is correctly sampled by the detector. The lens after the aperture can demagnify the aperture onto the specimen. The Fourier transform and propagation distance associated with it is indicated by the blue arrow and $\mathcal{F}$ symbol. An unincluded element in the setup shown in Figure 19 is the specimen stage, which moves the specimen perpendicular to the optical axis, however this is covered in Figure 18. The propagation distance between the specimen and detector determines pixel size of the reconstruction, so must be chosen with care.

1.18 Ptychography and the overlap constraint

Like the diffraction pattern constraint and the support constraint, the overlap constraint reduces the number of possible solutions to the phase problem [69]. The overlap constraint is best defined by Figure 20, in that the image on the left shows an approximate overlap of twenty percent of the exit waves, while the image on the right shows no overlap at all. The overlap is defined by the area covered by more than one diffraction pattern reconstruction, and is therefore defined by the size of the support's transmission area and the relative movement of the object, as they determine the information held within the collected diffraction pattern. So the image on the right of Figure 20 shows the case where the movement of the object is greater than the radius of the support constraint, a situation to be avoided for a successful reconstruction.

The overlap constraint works because individual solutions for individual diffraction patterns must match the overall object solution and constant probe/support solution, at least to the extent that noise allows [69]. Furthermore, the overlap/support constraint is also the reason that the object and probe can be separated and put in their appropriate places, since the probe/support is common to all collected data, while the object is not. This was initially the greatest difference between ptychography and other forms of coherent diffractive imaging (CDI), and caused the greatest improvement in the quality of reconstructions obtained as shown by Figure 20. Figure 20 a) shows ptychographic reconstruction with overlapping probe positions, comparing this to Figure 20 c) where

the probe positions do not overlap [69]. The importance of overlap is demonstrated as the specimen can only just be made out in c) whilst it is very clear in a).

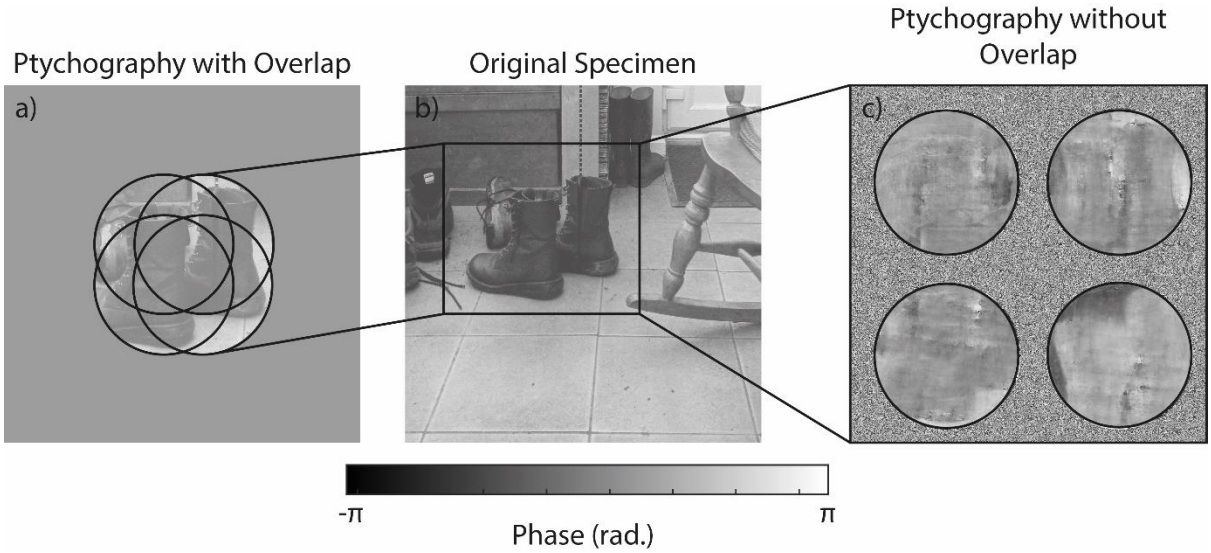

Figure 20: A figure to illustrate the difference in reconstruction quality of ptychographic reconstructions when there is or is not overlap of probe positions. a) shows a phase reconstruction when there is overlap in the dataset. b) shows the specimen which was used to create the datasets that are used in a) and c). c) shows a phase reconstruction produced by the same data as a) but with incorrect positions/camera length, such that probe positions do not overlap. The black circles indicate the size of the probe. This is a reproduction of a figure in [69].

The best way to visualise how ptychography and overlap improve reconstruction quality is to look at an example [65]. The example is an ambiguity of Fienup's HIO algorithm called the complex conjugate ambiguity/twin image problem (see Figure 20) and Fienup stated it as such (see equation 19) [71]:

$$|\mathcal{F}[U(x,y)]| = |\mathcal{F}[U^*(-x,-y)]| \tag{19}$$

The parameters of equation 19 are the following:

- $U^*(-x,-y)$ is the complex conjugate of U which has also had its positional coordinates flipped (an example can be seen on the bottom right hand side of Figure 20 c)).
- $\mathcal{F}$ is the Fourier transform

Equation $(19$ states that the magnitude constraint cannot tell the difference between the sides of the equation, thereby both solutions meet the magnitude constraint and the support constraint, therefore this is an ambiguity of the Fienup approach. This ambiguity cannot occur in ptychography while probe positions overlap, for the reasons previously stated above.

In ptychography's current state the overlap constraints have been optimised, as it has been found that the optimal overlap amount is 80%, although this will be somewhat object dependent [72]. Another aspect of note is the scan pattern, as this can affect the reconstruction quality [73]. The most notable is the raster pathology, which is where there are visible artefacts in the reconstruction of the object, due to the fact that the specimen stage was moved purely in a square grid pattern when collecting the diffraction data. Ptychography has benefited from the development of position correction [74] and recently from the ability to use continuous stage movement in order to reduce the time it takes to collect a dataset for reconstruction [75, 76]. Although the above topics will not be discussed in depth they are representative of the fact that improvements in ptychography are not solely in the reconstruction algorithm but also rely on making the most of the setup.

This section on ptychography concludes with a discussion of its advantages and disadvantages. The great advantage of ptychography is the unlimited size of the reconstructed object, which means for example that whole cell cultures can be imaged without losing the ability to zoom in to a singular cell level and still be able to resolve fine detail. Another major advantage is the robustness of the reconstruction process, since the conjugate image ambiguity is removed. Ptychography also enables the collection and removal of the illumination function from the exit wave, such that it is no longer a requirement for the support constraint shape or size to be known beforehand, making ptychography simple to apply practically [62]. A disadvantage arises from the large amounts of data that are required to apply this constraint to the phase problem for a significant field of view (specifically focused probe ptychography). This leads to a corresponding increase in time taken to collect the data in the experiment and which has further negative feedback, such as increasing probe drift.

1.19 Near field ptychography

Before describing near field ptychography there will be a discussion on ptychography with a diffuser and speckle tracking in X-ray microscopy. These two subjects are precursors to near field ptychography and therefore are covered to give context of the 'how and the why' of near field ptychography development.

In advance of the development of near field ptychography, diffusers were already playing a role in focused/defocused probe ptychography, as due to the Fraunhofer diffraction condition typically applied in traditional ptychography the centre of the detector would quickly become saturated. The centre of the diffraction pattern would saturate as the majority of the beam is not strongly scattered by the specimen or the aperture. Therefore through implementing a diffuser, the beam is effectively scattered twice, once by the diffuser and then by the specimen, which leads to a greater degree of counts in the high spatial frequencies (see Figure 21). Figure 21 a) shows a diffraction pattern from a simulated ptychographic experiment with a diffuser, and Figure 21 b) shows a simulated diffraction pattern from an experiment without a diffuser. Several papers have reported on implementing a diffuser in traditional ptychography [77, 78]. Although the diffusers increased the spread of counts across frequency domain in traditional ptychography, the diffraction pattern would still quickly saturate the detector. This is one of the historical reasons why near field ptychography was developed, to make efficient use of the detector's dynamic range, as near field ptychography diffraction patterns have reduced standard deviation in the number of counts across the whole of the detector.

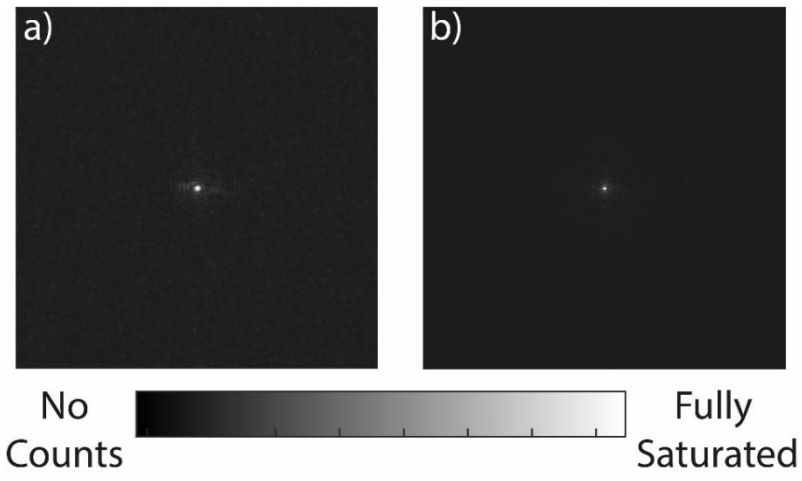

Figure 21: a) a ptychographic diffraction pattern where the probe was an aperture and a diffuser. b) a ptychographic diffraction pattern where the probe was only an aperture. The colour bar at the bottom of the figure converts the greyscale of a) and b) into counts between no counts and a fully saturated detector [77].

Diffusers were also being used in a different form of phase retrieval, that of X-ray speckle tracking. X-ray speckle tracking helped inspire the development of near field ptychography, so will be briefly discussed in this paragraph [79]. X-ray speckle tracking consists of disrupting an X-ray beam with a diffuser (often a piece of paper). The image of the paper is out of focus on the detector, such that diffraction of the paper causes a speckle pattern (an inline hologram) [80]. X-ray speckle tracking collects a diffraction pattern without a specimen as a reference. Once a specimen is placed in the beam, the diffraction pattern on the detector changes with respect to the reference. An example of this phenomenon can be seen in Figure 22, in which Figure 22 a) shows the reference diffraction pattern and Figure 22 b) shows the diffraction when the specimen is introduced into the beam. The specimen used in this simulated example is shown in Figure 22 c). As can be seen by looking at Figure 22, the diffraction pattern has only changed in the x-axis and the speckles have been stretched compared to the reference. This stretching is the result of the changing phase of the specimen (or gradient), and therefore the phase gradient can be recovered by measuring the shift between Figure 22 a) and Figure 22 b). Near field ptychography effectively uses the same effect as speckle tracking to determine the phase, as through moving the specimen with respect to the diffuser the speckles are displaced uniquely as the specimen is scanned [81]. There is sufficient diversity to constrain the phase retrieval algorithms used by traditional ptychography (see section 1.20 for more details on diversity).

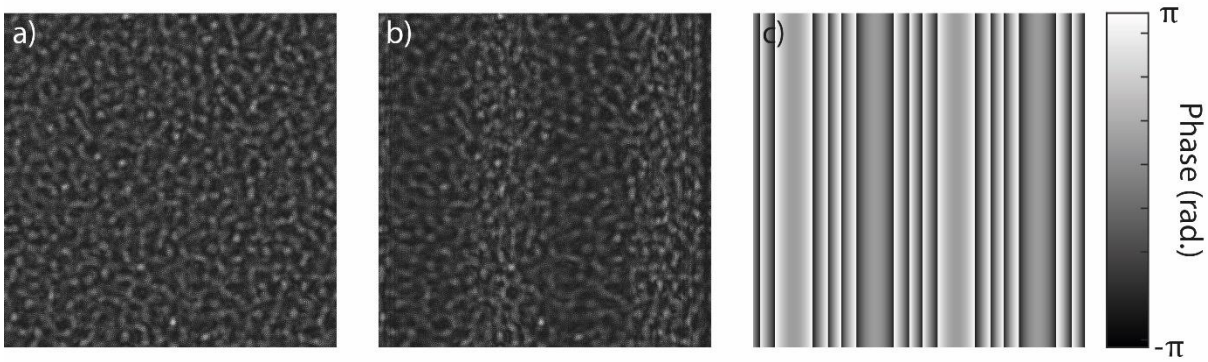

Figure 22: a) inline hologram of simulated diffuser. b) inline hologram of combined diffuser and specimen, where specimen has caused speckles of the hologram to shift, stretch and compress. c) the phase of specimen used in b). the colour bar to the right of the figure converts the greyscale of c) into phase in terms of radians.

Now that ptychography with a diffuser and X-ray speckle tracking has been discussed, it is time to look at the specific variant of ptychography that is the focus of this thesis; near field ptychography. Near field ptychography differs from previous forms of ptychography in its removal of the support constraint, or equivalently the localisation of the probe. This has many benefits which will be discussed in section 1.21. Near field ptychography also makes use of an altered setup to encode phase diversity into collected interference patterns (which are spatial domain images, as opposed to diffraction patterns in the Fourier domain) which assist the overlap constraint in resolving the phase problem, as will be seen in section 1.20. To begin there will be a description of the setup of near field ptychography.

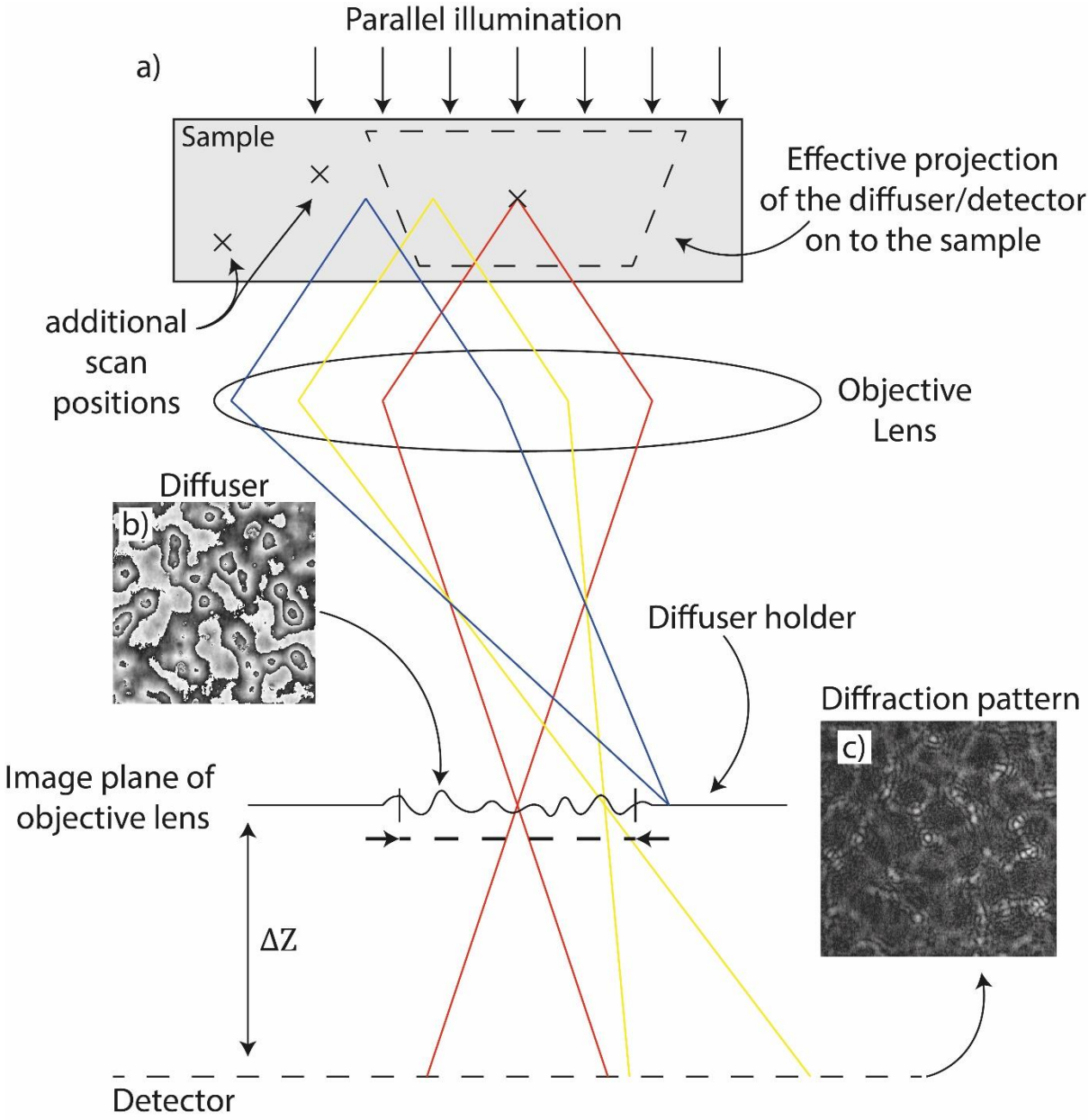

Figure 23: a) A diagram of near field ptychography where the crosses represent the centres of other sample subsections, where near field interference patterns are collected from the sample (the same as Figure 18). The dotted lines in a) represent the demagnified area of the detector at positions in the setup (see diffuser holder and sample). The sample is then magnified on to a diffuser located in the selected area plane where it is modulated before it is propagated on to the detector. b) is the phase of the typical diffuser used in this setup. c) is an example of a diffraction pattern this setup produces.

In near field ptychography a collimated coherent light source illuminates a sample which is then magnified/demagnified by a lens onto a diffuser, such that the image of the specimen is disrupted by the structured and random pattern of the diffuser. The structured illumination and sample subsection then multiply together to form an exit wave which is propagated to the detector. Typically the propagation distance is such that the propagation is modelled by the angular spectrum method or the Fresnel propagation method (high Fresnel number – i.e. the when the multiplication of the propagation distance and wavelength is small or equivalent to the feature size of the specimen, see Section 1.30 for more details on the Fresnel number). As with previous forms of ptychography (Figure 18) the sample is moved by a translation stage, such that different subsections

of the sample can be illuminated by the probe, and a stack of interference patterns is collected, encoding the overlap constraint in the collected data.

The main difference between focused/defocused probe ptychography and near field ptychography are:

1. The removal of an explicit support constraint and replacing it with structured illumination (diffuser and lens). This enables each diffraction pattern to contribute more pixels to the final reconstructed image.
2. A change in the propagation distance, such that the collected intensities resemble the exit wave.
3. Through changing the propagation distance and diffusing the incident beam, the detector dynamic range has increased utility.

Disadvantages of near field ptychography:

1. near field ptychography in the TEM is not yet optimised for high throughput data collection, so while near field ptychography requires less data to reconstruct a given field of view, ptychography in STEM collects a diffraction pattern at a higher rate.
2. The resolution of the technique is determined by the demagnified pixel size.
3. It requires a diffuser to constrain the phase problem (or some other method to structure the illumination), limiting the number of TEMs which can currently implement near field electron ptychography.
4. Its contrast transfer function is not fully understood.

The question then becomes, how is near field ptychography constrained without the support constraint of the previous forms of phase retrieval, and this will be discussed in the next section.

Near field ptychography has undergone significant developments since its first implementation, therefore significant papers and their results are outlined in this paragraph. Firstly, near field ptychography has been shown to be able to image thick specimens in the X-ray domain [56]. This was done by M. Stockmar et al. in order to demonstrate that near field ptychography had the potential to perform nano-tomography. This was followed up by the paper “X-ray tomography using near-field ptychography” [82]. In [82] it was shown that near field ptychography could produce 3D images of specimens with as few as 16 diffraction patterns per projection. For context, focused/defocused probe ptychography typically requires hundreds of diffraction patterns per projection in the X-ray domain; take for example [83] where 213 diffraction patterns were required per projection. Although it should be noted that these two different experiments cannot be directly compared fairly, the difference in number of diffraction patterns per projection does give a general sense of how near field ptychography has reduced data demand compared to ptychography. However, in current practical implementations ptychography achieves higher resolutions. Furthermore, near field ptychography in the X-ray domain has been shown to provide quantitative tomographic reconstruction, as shown by I. Zanette et al. in [84]. Near field ptychography has important developments in the optical domain, for example it was shown that with as few as 6 diffraction patterns, near field optical ptychography can obtain phase images of blood cells and near field optical ptychography was capable of imaging thick specimens in the optical domain given the correct setup (the diffuser interacts with image of the specimen). Another optical development can be found in [85], where a lens-less form of optical near field ptychography was developed, which makes use of up-sampling to provide magnification. Furthermore, Jiang et al. [85] showed that near field ptychography with a sufficiently high Fresnel number (or when multiplication of the

propagation distance with wavelength is sufficiently small compared to the dimensions of the specimens features - see section 1.30) does not need to collect the probe positions, and instead they obtained the positional data from the collected diffraction patterns, which is found to be useful in section 3.3.

1.20 How near field ptychography is constrained

Near field ptychography drops the support constraint for reasons outlined in section 1.21, but continues to prove valid reconstructions of both the probe and object for setups like the one described in Figure 23. This poses the question, why is this the case?

To start with, the effect of the diffuser on the phase problem will be considered. The diffuser structures the illumination to add localised random phase disturbances, leading to different localised intensities in different interference patterns due to movement of the sample via the stage. This can be seen in Figure 24 c) and d) which show the intensity of two diffraction patterns created from two different sample-probe interactions. Figure 22 a) and b) show the phase of the object subsections used in these interactions. Figure 22 e) and f) show a subsection of the diffraction patterns shown in Figure 24 c) and d), the change in intensity shows how a simple change in the phase of the specimen causes the collected intensity to change, therefore demonstrating how the phase is encoded. In effect the probe is applying localised phase plates to the sample, such that some of the diffracted beam then interacts and undergoes a phase shift and the deflected beam can interfere with the unaltered beam to produce positive (constructive interference) or negative contrast (destructive interference) on the detector in a localised area. The method of using phase to produce a diverse set of data in order to solve the phase problem is called phase retrieval via phase diversity [86, 87]. In fact phase diversity and the overlap constraint are very similar, in that both use the diversity of collected interference patterns to constrain the phase problem to a workable solution, by having many simultaneous equations (collected intensities) which have one solution: the correct phase [86]. Therefore, using the two constraints in parallel may be a reason why an explicit support constraint is not needed.

Diversity is an important term for ptychography, and especially important to near field ptychography as it is a way to describe how easy the dataset is to reconstruct. But what diversity actually means may vary slightly from author to author. The aim of this paragraph is not to define diversity in the context of near field ptychography, but to describe how it is used throughout this thesis. Let us start off with an example, say there is a near field diffraction pattern named c) (shown in Figure 24 below) in which certain aspects of the diffraction pattern are strong, for example the darker bars. Now if we obtain a second diffraction pattern named d) which is another diffraction pattern in the same scan pattern, there is typically an 80% overlap between the specimen in c) and d). The large amount of overlap theoretically means that c) and d) should be very similar to each other, as essentially they are displaying the same information apart from the fact that the specimen is shifted in one diffraction pattern with respect to the other (See Figure 24 a) and b)). This is why the diffuser is important, essentially to disrupt the specimen's features or interference fringes. The diffuser disrupts the interference fringes of the specimen by creating its own interference fringes, and where the interference fringes of the specimen and diffuser overlap, they combine into a unique set of interference fringes for that diffuser specimen subsection combination. As near field ptychography relies on scanning of the specimen with respect to the diffuser, there are many different interactions between their diffraction fringes which cause a unique interference pattern to that particular scan position. So a diverse dataset would be one where there is plenty of overlap in the diffraction fringes of the specimen and probe, but where diffraction patterns differ greatly from each other, even when they share the same specimen area.

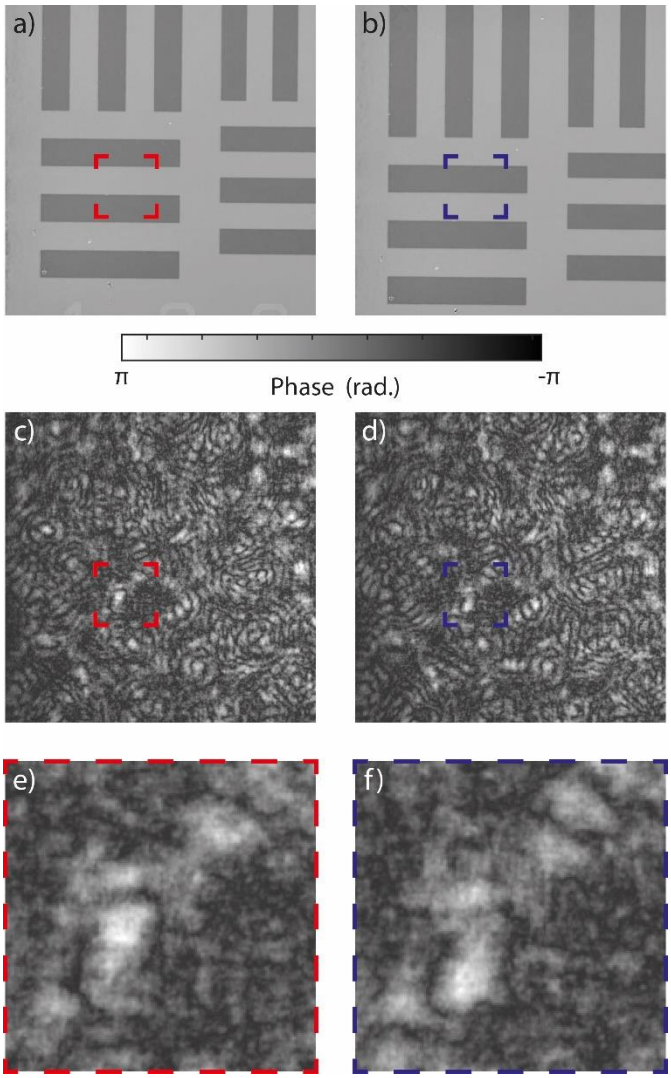

Figure 24: a) and b) show the phase of two sample subsections. c) shows a near field ptychography diffraction pattern which was collected when a) was the specimen subsection being imaged. d) shows a near field ptychography diffraction pattern which was collected when b) was the specimen subsection being imaged. e) and f) are subsections associated with the area indicated by red and blue dotted lines in a)-d) respectively.

1.21 The benefits of near field ptychography

Historically one of the benefits of near field ptychography is that it utilises more of the detector's dynamic range, as in previous forms of ptychography diffraction patterns were collected, while near field ptychography collects interference patterns [88]. The centre peak of a diffraction pattern represents the unscattered photon/electron beam which contains most of the power, and therefore the centre peak has the largest count rate in the frequency domain and quickly saturates the detector. Meanwhile, high frequencies in the Fourier domain accumulate counts far more slowly, and consequently if saturation of the detector is to be avoided then the signal to noise ratio of high frequency components will be reduced due to the exposure time the centre peak demands. Operating in the near field leads to the case where the collected intensity is more evenly distributed across the detector, enabling more time for high frequency information to be collected by the detector relative to collecting far field diffraction patterns. Examples of interference patterns can be seen in Figure 24, where it can be seen from inspection that intensity values are almost constant across the detector and that high frequency information is captured. However, direct electron detectors with high read out speeds can collect far field diffraction patterns consisting of a small

number of counts per diffraction pattern, preventing saturation, whilst the diffraction patterns retain a high signal to noise ratio [89]. Furthermore, the experimental parameters of focused probe ptychography's (or STEM ptychography) can be altered in order to capture a particular spatial frequency of the specimen. By varying the convergence angle of the illumination size of the C2 aperture, spatial frequencies up to twice the convergence angle can be accessed. The PTF of focused probe ptychography is affected by the amount of overlap/interference in the Fourier domain of nearby diffraction discs, which is directly affected by the choice of the C2 aperture [89, 90] (see references for further details). However, as direct electron detectors are not implemented on every microscope, near field electron ptychography may be more useful on lower specification microscopes.

When using a CCD (charge coupled device) the spreading of counts across the detector is also helped through the use of a lens and a diffuser which act to disperse and scatter the beam, such that it occupies a greater range of frequencies before it reaches the detector and object [88]. The diffuser also reduces the likelihood that there is a single high transmission and low scattering path through both the diffuser and sample.

Another benefit of near field ptychography is the extended field of view it enables per collected detector frame, compared to previous ptychography methods, as there is no aperture acting as a support constraint [79]. The reconstructed exit wave of near field ptychography resembles the whole wave front, instead of a circle half the size of the sampled region (see Figure 24). This extended field of view leads to a reduction in the amount of data required to solve the phase problem for a given field of view, which in turn reduces the time required to collect a dataset, a major problem experienced in focused probe ptychography. However, it is important to mention that ptychography in general does not require an aperture to operate correctly since the illumination function is recovered alongside the specimen, but the limited field of view of the aperture is still effectively enforced by the Nyquist sampling requirements of the Fourier diffraction patterns. The correct sampling of a Fourier diffraction pattern requires that its real space equivalent is half the field of view of the detector; this is so that the Nyquist sampling condition can be satisfied [70]. The factor of a half comes from the actuality that intensities have twice the rate of the change of the magnitude or phase of the specimen. The traditional sampling conditions of CDI can be overcome by ptychography via increased real space sampling (smaller step size) as shown by Edo and Batey, yet this does not result in a greater field of view per diffraction pattern but instead a greater calculation window [91]. The limited field of view per diffraction pattern is also evidenced by the probe reconstructions of defocused probe ptychography (often quoted as requiring fewer diffraction patterns per given field of view than focused probe ptychography), whereby reconstructed probes only have a significant magnitude for a limited part of the calculation window; see supplementary figure 13 of [92] and supplementary figure 1 of [93]. However, near field ptychography's probe has a significant magnitude for the entire calculation window, see Figure 49 and Figure 102. Near field ptychography provides a benefit compared to off-axis holography, in that it takes account of imperfections of the illumination function automatically [56, 79]. So near field ptychography does not need to obtain flat field images to correct non-uniform illumination functions, which can cause further ambiguities [56, 79]. This is achieved through obtaining the probe (separability constraint) function encoded in the dataset, and therefore the imperfections and errors are collected in the probe function if they are independent of the probe position on the sample. Furthermore, prior knowledge of the probe/support constraint helps with the convergence of the reconstruction to a high quality solution in fewer iterations, which can be implemented by using the same diffuser for multiple experiments, enabling prior knowledge of the probe.

1.22 Selected area ptychography

Selected Area Ptychography (SAP) is a specific form of near field ptychography, in which the diffuser is replaced by an aperture. SAP is primarily of importance in the context of TEM, as implementing a diffuser requires that the TEM is capable of holding custom apertures either in the selected area aperture plane or condenser lens system. While the aperture is far less diverse than a diffuser and has a greatly reduced field of view, it has the advantage of being readily available in all TEMs, whereas custom aperture holders are not always available. SAP can only operate when the aperture is sufficiently diffracted by the defocus between the specimen (or selected area aperture plane) and the detector plane of the microscope, such that the ringing/diffraction of the aperture fills the opening of the aperture. An example of a SAP diffraction pattern can be seen in Figure 25. Although SAP was developed after near field ptychography, it was a precursor of the work done in this thesis. SAP was originally described in the paper "Quantitative electron phase imaging with high sensitivity and an unlimited field of view" [94]. As can be seen from Figure 25, the field of view per diffraction pattern can be severely restricted by the requirement that the aperture is sufficiently diffracted. If larger apertures are used in SAP, this increases the required coherence of the source, in order to obtain the required diffraction condition. However it was shown in [94], that SAP can obtain quantitative phase images of latex spheres and fields of view which were un-restricted (in terms of an upper bound – theoretically).

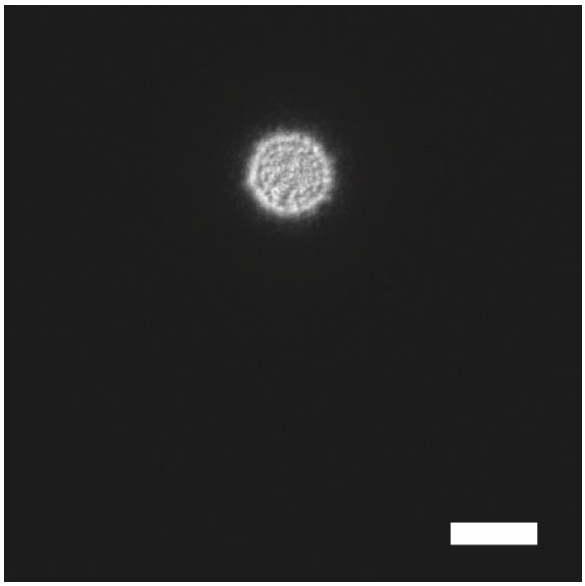

Figure 25: An example of a SAP diffraction pattern produced by a 10μm diameter aperture from the dataset used in [94]. The scale bar indicates 10μm in selected area aperture plane.

1.23 Discussion on phase retrieval

So in review, there has been a discussion of how holography (inline and off-axis) can be used to retrieve phase information. However off-axis and inline holography have imperfections. For example, electron off-axis holography requires an electric bi-prism which is not standard in most TEM around the world, therefore limiting the number of TEMs on which it is applicable. Off-axis holography has limited field of view compared to inline holography due to the method by which the phase is reconstructed, and higher requirements on the coherence of the source [46]. Inline holography has a simpler setup than off-axis holography, but it has problems with the following: it requires additional steps to ensure low frequency noise is eliminated [51, 58], has a limited field of view compared to near field ptychography, requires extensive software correction to make collected data usable in electron microscope applications [56], and inline holography also requires multiple

exposures during which the specimen can change [95]. This weakness is shared by ptychography, however it has been shown that ptychography can be robust to dynamic movements of objects through the use of object modes and perhaps can be robust to changes in the specimen with further development [96-98].

However, it should not be understated that off-axis holography and inline holography have successfully obtained large amount of quantitative data in the electron domain. Off-axis holography has generated many review articles providing updates on recent experiments and future steps [99-101], whilst inline holography has provided quantitative information on semiconductor devices [102] and magnetic materials [103-105]. Near field ptychography in theory shares the quantitative applications in the references above, so future work will include identifying where near field electron ptychography can improve on previous quantitative measurements in the electron domain.

Then development of CDI was discussed. CDI was developed to address the phase problem from a different angle compared to holography, that of iterative phase retrieval algorithms. The first phase retrieval algorithm was the Gerchberg-Saxton (G&S) algorithm, which applied the constraint that simulated waves must have the same magnitude as the collected intensity in both the spatial domain and in the Fourier domain. This approach was successful but had a large number of problems, practically and algorithmically. The major problem was that it was limited to the case where samples could attenuate the incident beam strongly enough to produce contrast in the spatial domain and where optics produced high quality images. These conditions ruled out the majority of the X-ray and electron microscope phase imaging applications. In response to this, Fienup developed the Hybrid Input Output (HIO) algorithm, which applied a support constraint of a circular aperture and feedback from previous values. This greatly improved the quality and applicability of phase retrieval.

However, the HIO algorithm could still be improved upon. This was addressed by ptychography which extends the field of view of the recovered sample and removes the requirement to explicitly know the support. Ptychography's success was due to the application of the overlap constraint where there are spatial overlaps in collected data, so phase solutions had to agree across multiple instances of the phase problem. Then near field ptychography improved upon focused/defocused probe ptychography, by removing the support totally and in principle enabling faster data collection for a given field of view. Furthermore, the propagation distance and setup enables improved utilisation of the detector, improving signal to noise ratio for high frequency components [56] (in the case of older CCDs).

This concludes the discussion on phase retrieval. There will now be a shift in subject matter to describe some of the tools and maths that are required in order to retrieve phase practically with near field ptychography and then assess the quality of the retrieved phase in an objective and quantifiable manner. The tools which will be discussed are the rPIE algorithm, propagation techniques (for example near field propagation and Fresnel propagation), the effect of incoherence, and then the Fourier Ring Correlation, as a method to compare reconstructions from different phase retrieval techniques.

1.24 rPIE

The regularized Ptychographical Iterative Engine (rPIE) is the latest iteration of 'PIE' algorithms that were developed in Sheffield to solve the ptychographic phase problem [106, 107]. The rPIE method is used throughout this thesis to solve the phase problem, therefore it is natural that there should be a discussion on how it works. An outline of the algorithm can be seen in Figure 26.

The first action of the algorithm is to 'cut out' a section of the initialised object. The position of the cut out is based on information fed in by the experiment. The cut out is then multiplied by the initialised probe in order to give a guess at the exit wave:

$$\psi_{j,n} = P_{j,n} O_{j,n} \tag{20}$$

The parameters of equation 20 are the following:

- P is the probe
- O is the object cut out
- *j* denotes both the position and number of the cut out and which diffraction pattern it is in the experimental dataset.
- *n* denotes the iteration number
- $\psi_{j,n}$ is the guessed exit wave of the cut out j and the iteration number n

This guess at the exit wave is then forward propagated, such that it is in the same physical space as the collected data from the experiment. This forward propagation requires knowledge of the propagation distance, correct propagation method, detector size, number of pixels and pixel pitch, in order for the overall simulation to be accurate:

$$\Psi_{j,n} = D_z\left[\psi_{j,n}\right] \tag{21}$$

The parameters of equation 21 are the following:

- D_z is the propagation method most accurate to the experiment, for example Fraunhofer propagation
- Ψ is the forward propagated guessed exit wave

The next step is to apply the diffraction intensity constraint: this amounts to removing the absolute magnitude of the current guess of the propagated exit wave and replacing it with the measured diffraction data from the experiment corresponding to the initial cut out of number j:

$$\Phi_{j,n} = \frac{\sqrt{I_j}}{\left|D_z\left[\psi_{j,n}\right]\right|} \times D_z\left[\psi_{j,n}\right] \tag{22}$$

The parameters of equation 22 are the following:

- Φ is the forward propagated adjusted exit wave
- $\sqrt{I_j}$ is the diffraction data which corresponds to the object cut out j

The forward propagated adjusted exit wave is then back propagated such that it is in the same domain as the probe and object:

$$\varphi_{j,n} = {D_z}^{-1}\left[\Phi_{j,n}\right] \tag{23}$$

The parameters of equation 23 are the following:

- φ is the adjusted exit wave
- ${D_z}^{-1}$ is the inverse propagation action

After the constraint of the collected data has been applied and the adjusted exit wave has been back propagated, we then use the adjusted exit wave to update our guesses at the object and probe, using the following update functions:

$$O_{j,n+1} = O_{j,n} + \frac{P_{j,n}{}^{*} \times \left(\varphi_{j,n} - \psi_{j,n}\right)}{\left((1-\alpha) \times \left|P_{j,n}\right|^{2}\right) + \left(\alpha \times \left|P_{j,n}\right|^{2}_{max}\right)} \tag{24}$$

$$P_{n+1} = P_{j,n} + \frac{O_{j,n}{}^{*} \times \left(\varphi_{j,n} - \psi_{j,n}\right)}{\left((1-\beta) \times \left|O_{j,n}\right|^{2}\right) + \left(\beta \times \left|O_{j,n}\right|^{2}_{max}\right)} \tag{25}$$

The parameters of equation 24 and 25 are the following:

- α is the tuning parameter which determines the rate at which the object is updated. (The ratio of the old object to the new object in the cut out, which is placed back into the overall object, has a typical value of one)
- β is the tuning parameter which decides the update rate of the probe (typically set to one)
- Again the * denotes the complex conjugate

The updated probe and object are then used for the next cut out j+1. Once the algorithm has cycled through all of the cut outs of the experiment, the iteration number is incremented and cut outs are then processed again. It is, at this point, important to remember what was previously said in section 1.14: changing the magnitude or phase of either side of the Fourier transform (propagation technique) causes a corresponding change in the other side. Therefore, the probe and object updates act to ensure that a new guess is generated in the next iteration of the algorithm. α and β then determine the extent of the update, but it is the subtraction between the adjusted and unadjusted exit wave which forms the main driver of the reconstruction improvement, as it is the feedback mechanism of ptychography, much like Fienup's HIO algorithm.

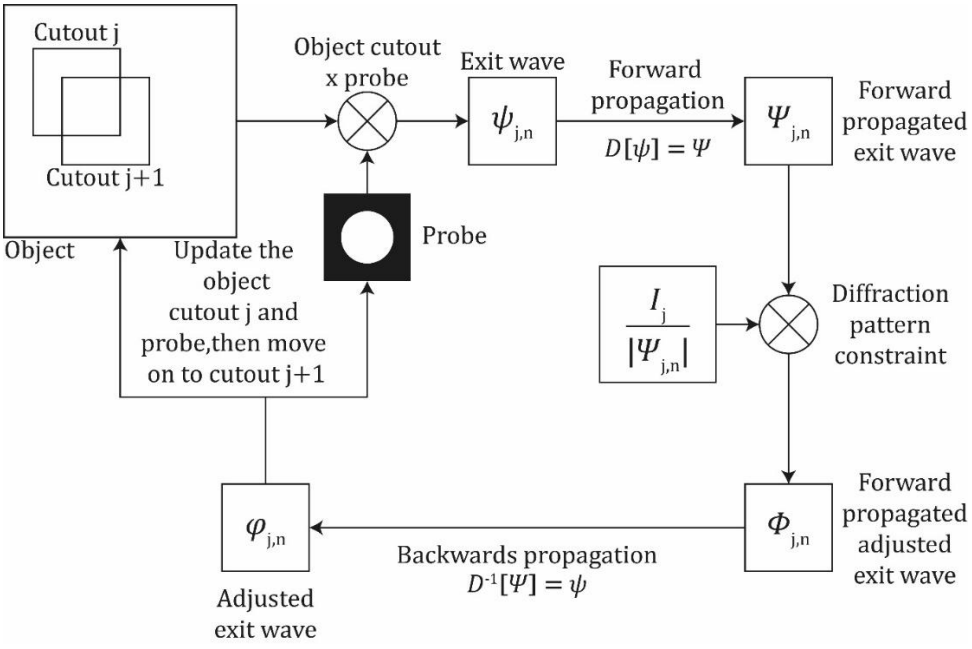

Figure 26: outline of the rPIE algorithm

In this section, general propagation terms have been used to show that rPIE in theory can be applied to either general ptychography or near field ptychography. In the next section there will be an

investigation into three specific forms of propagation: angular spectrum and Fresnel propagation, which are relevant for near field ptychography, and Fraunhofer propagation, which is required for traditional ptychography. Understanding the propagation method is critical to ptychography, as will be shown by this section. This is because phase retrieval requires that the correct propagator is used to transform the collected data into the exit wave formed by the probe and the object at the object position, and no other position.

1.25 Propagation of light

In this section there will be a summary of the mathematical modelling tools that simulate the propagation of light. The tools are: the angular spectrum, the Fresnel propagator, and the Fraunhofer propagator. Being able to mathematically describe how light propagates is a requirement to retrieve phase information using ptychography, as it enables the correct conversion between the spatial domain description of the object and the spatial frequency domain of the diffraction pattern.

First, there will be an outline of a general example, which will be the same for the angular spectrum, Fresnel and Fraunhofer propagation. The purpose of the example is to diagrammatically illustrate how light can propagate and define common terms for the mathematics. The rest of the section will be presented in the following order: angular spectrum; the effect of propagation distance on the angular spectrum, Fresnel propagation and consideration of its validity with respect to propagation distance, and Fraunhofer propagation.

1.26 Plane waves

To understand propagation requires an understanding of some fundamental aspects of optics and waves, starting with the plane wave [5, 108]. A 3D plane wave is a wave whose amplitude and phase do not vary in the two dimensions which are orthogonal to the direction of propagation, so the plane wave is only a function of the distance it propagates. Like all waves, the plane wave can be described as a complex field.

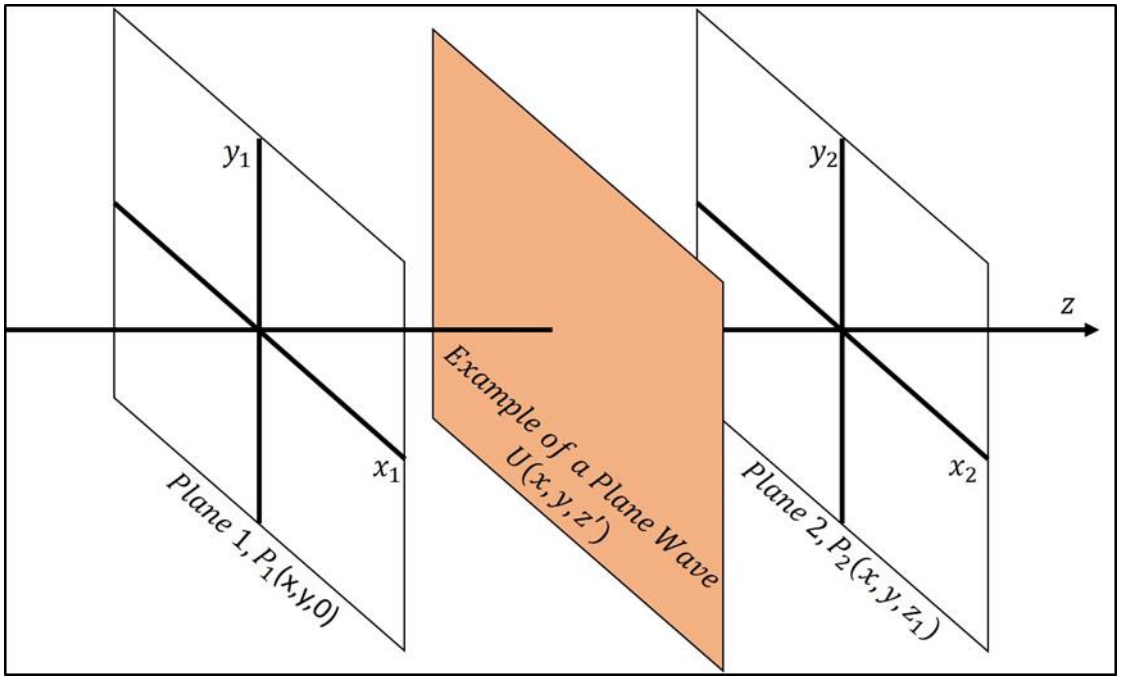

Figure 27: A general example of light propagation

Figure 27 illustrates a plane wave traveling in the direction *z*, plane 1 was a previous cross-section of the plane wave and plane 2 is a cross-section further along in the *z* direction. Moreover, in Figure 27 there is a cross section of a plane wave (*U*(*x,y,z'*)) at position *z'* which is an arbitrary distance between the two planes. To complete this consideration of a plane wave, its equation is stated below in equation 26:

$$U(x_1, y_1, 0) = U_0 \exp\left(j\left(-\vec{\boldsymbol{k}} \cdot \overrightarrow{\boldsymbol{r}(x_1, y_1)} + \omega t + \varphi\right)\right) \tag{26}$$

U_0 is the magnitude of the plane wave, $\vec{\boldsymbol{k}}$ is the wave vector (typically a unit vector) which describes the direction of the wave in three dimensional space, and $\vec{\boldsymbol{r}}$ is the spatial position vector which determines where you want to sample the plane wave. The ω term is the temporal angular frequency, t is time and φ is a general phase offset: in the following material the temporal aspects will be dropped from equation 26, as we make the assumption that all waves are temporally and energetically coherent (i.e. emitted at the same time and of the same wavelength).

1.27 Spherical waves

It is typical in optics literature to categorize waves in terms of shape, which brings us to spherical waves which have applications in the Fresnel and the Fraunhofer propagators [5]. However, ultimately spherical and plane waves are very similar mathematically, as can be seen from inspection of equation 27, which describes a spherical wave:

$$U(x_1, y_1, 0) = \frac{U_0}{|\vec{r}|} \exp\left(j\left(-\vec{\boldsymbol{k}} \cdot \overrightarrow{\boldsymbol{r}(x_1, y_1)} + \omega t + \varphi\right)\right) \tag{27}$$

Equation 27 also relives the truth of A-level physics (specifically the class on the Earth's gravitational field) that a wave/field shape is mainly a function of the distance away from the source and area/size of the detector observing the wave/field. In the case of the spherical wave there are radii of constant phase, and also spherical waves are a good descriptor of another optics fundamental, Huygens's wavelets. This concludes this short discussion on basic types of waves. Plane waves form a fundamental part of the next section on the angular spectrum.

1.28 The angular spectrum

In terms of understanding light propagation, the angular spectrum is the best starting point from an engineer's perspective, as this approach is similar to the linear systems analysis studied at undergraduate level [109]. Referring to Figure 27, our objective is to obtain the relationship between U at $z' = 0$ and U at $z' = z$. This idea is best tackled using the tools of the Fourier transform. The two dimensional Fourier transform of $U(x, y, 0)$ is:

$$A(f_x, f_y, 0) = \iint_{-\infty}^{\infty} U(x, y, 0) \exp\left[-j2\pi\left(xf_x + yf_y\right)\right] dxdy \tag{28}$$

The parameters of equation 28 are the following:

- A is the Fourier transform of U
- x and y are the coordinates in the spatial domain
- f_x and f_y are the coordinates in the frequency domain

A is also called the angular spectrum of *U*. Remembering the one dimensional Fourier transform – that all one dimensional signals can be described as a summation of sine waves – the same can be said for the two dimensional case, except the sine waves are now 2D complex exponentials.

Furthermore, this Fourier transform can be inverted such that *U* can be expressed as a function of *A*, as follows in equation 29:

$$U(x,y,0) = \iint_{-\infty}^{\infty} A(f_x, f_y, 0)\exp\left[j2\pi\left(xf_x + yf_y\right)\right] df_x df_y \tag{29}$$

The idea of the angular spectrum is that any wave can be described as a summation of plane waves of different spatial frequencies. This is similar to the Fourier series for electrical signals, except that plane waves are used instead of sinusoids [108]. Looking at equation 26 and equation 29 one could interpret the exponential of equation 28 to be a plane wave where $\vec{\boldsymbol{k}} \cdot \vec{\boldsymbol{r}}$ is equivalent to xf_x meaning that there is literally no difference between the summation of plane waves and the Fourier transform. In order to give more weight to the previous statement there must be further investigation into the physical aspect of summing plane waves, how this links to the name of this section and the Fourier transform. To start, let us take an arbitrary example of a plane wave (equations 31 and 32) which has a magnitude of one (U_0), a wavelength of 1 millimetre and the following $\vec{\boldsymbol{k}}$ and $\vec{\boldsymbol{r}}$ terms (see equation 30):

$$\vec{\boldsymbol{k}} = \frac{2\pi}{\lambda}\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \vec{\boldsymbol{r}} = \begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix}, \text{so } \vec{\boldsymbol{k}} \cdot \vec{\boldsymbol{r}} = \frac{2\pi}{1\times 10^{-3}}\left(\frac{1}{\sqrt{3}}x_1 + \frac{1}{\sqrt{3}}y_1\right) \tag{30}$$

$$U(x_1, y_1) = U_0 \exp\left(j\left(\vec{\boldsymbol{k}} \cdot \overrightarrow{\boldsymbol{r}(x_1, y_1)}\right)\right) \tag{31}$$

Therefore, the equation of the plane wave equation $U(x_1, y_1)$ becomes:

$$U(x_1, y_1) = 1\exp\left(j2\pi(557.4x_1 + 577.4y_1)\right) \tag{32}$$

Equation 32 suggests that a plane wave describes exactly one spatial frequency over the x-y plane; the frequency defined by the wave vector $\vec{\boldsymbol{k}}$. The wave vector $\vec{\boldsymbol{k}}$ describes the direction of propagation of the wave, but it also describes the frequency of the pattern imposed on a plane orthogonal to the *z* axis, as shown by Figure 28 and described by equation 32.

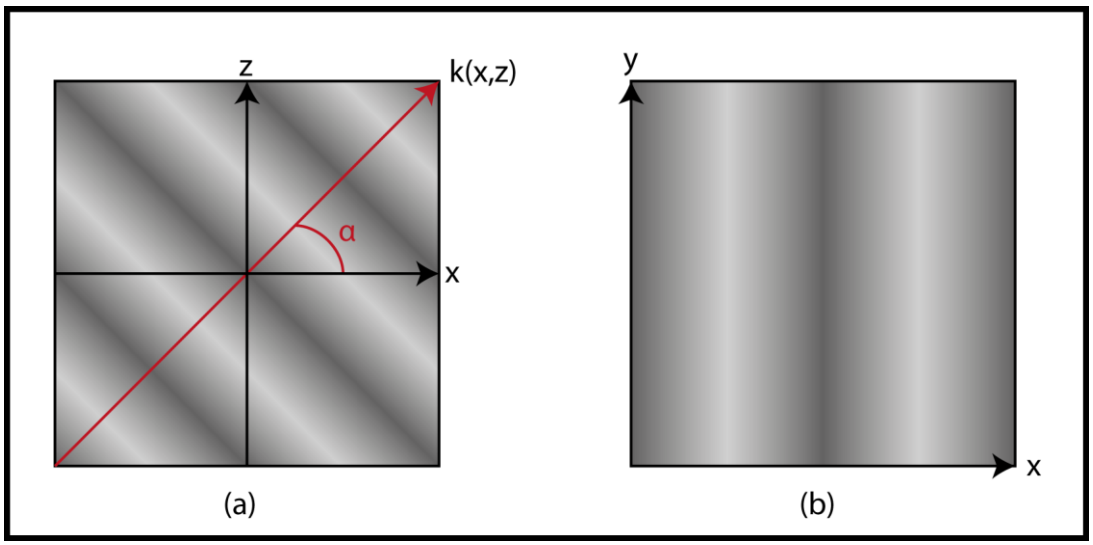

Figure 28: (a) shows an example of a plane wave from above with propagation direction defined by $\vec{\boldsymbol{k}}$(x,z) which can be equally described by α, and (b) shows the phase of a cross section of the plane wave at the x-axis of figure (a).

Furthermore, a more useful way of describing the wave vector $\vec{\boldsymbol{k}}$, would be as the frequency position vector, as it is the Fourier space complement of the spatial position vector $\vec{\boldsymbol{r}}$. The link between angles and frequencies is contained in the unit vector of $\vec{\boldsymbol{k}}$, because a unit vector $\vec{\boldsymbol{k}}$ can equally be described by the three angles between the coordinate axis and the direction of propagation, as seen in Figure 29 and equation 33. The unit vector of $\vec{\boldsymbol{k}}$ is formed of 3 direction cosines, one for each axis, as seen by equations 34, 35 and 36, but in effect the unit vector and direction cosines are the same thing. The fact that cosines of real space angles relate to positions in the frequency space is in part why this propagation algorithm is called the angular spectrum, because through scanning different angles in real space we are sampling in the Fourier space.

$$\vec{\boldsymbol{k}} = \frac{2\pi}{\lambda}\begin{bmatrix}\alpha\\ \beta\\ \gamma\end{bmatrix} = \begin{bmatrix}k_x\\ k_y\\ k_z\end{bmatrix} = 2\pi\begin{bmatrix}f_x\\ f_y\\ f_z\end{bmatrix} \tag{33}$$

$$\alpha = \cos(a) = \frac{k_x}{\sqrt{{k_x}^2 + {k_y}^2 + {k_z}^2}} = \frac{\text{adjacent}}{\text{hypotenuse}} \tag{34}$$

$$\beta = \cos(b) = \frac{k_y}{\sqrt{{k_x}^2 + {k_y}^2 + {k_z}^2}} \tag{35}$$

$$\gamma = \cos(c) = \frac{k_z}{\sqrt{{k_x}^2 + {k_y}^2 + {k_z}^2}} \tag{36}$$

$$\gamma = \sqrt{1 - \alpha^2 - \beta^2} \tag{37}$$

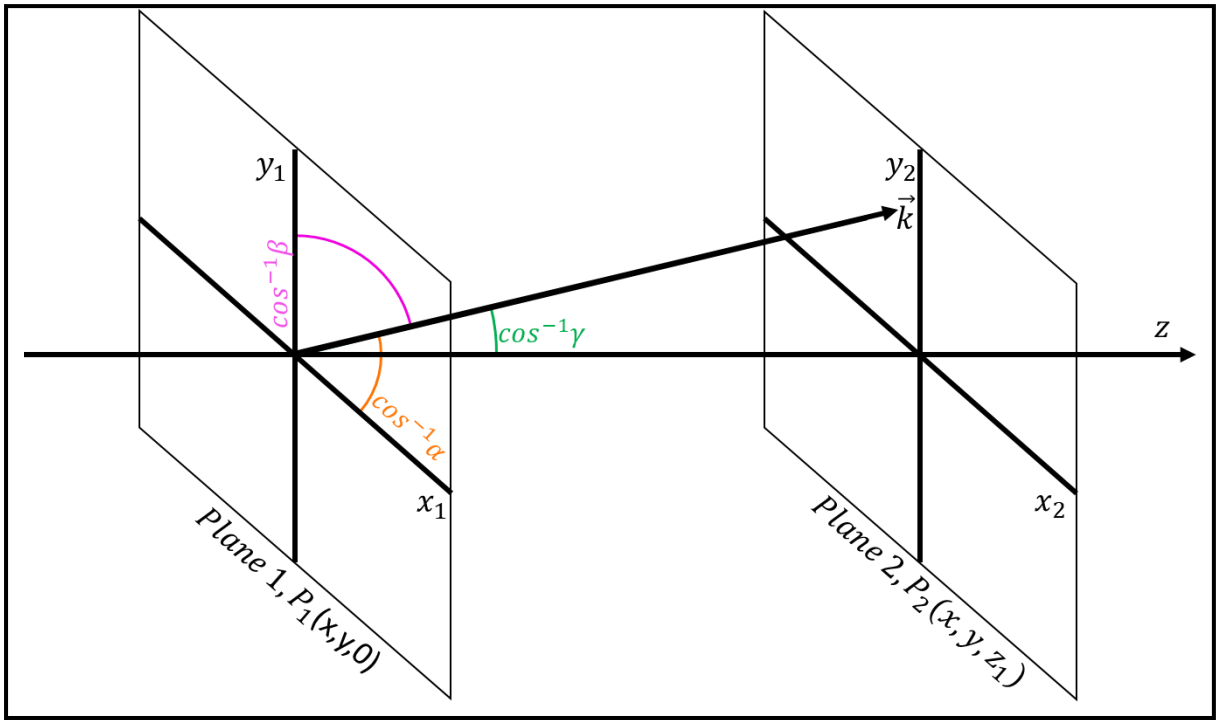

Figure 29: A diagram representing direction cosines [109]

equation 29 can be rearranged into the angular spectrum form using equations 33 to 37 and the discussion above, as the relationship between the direction cosines and the frequency domain can be seen in equation 38:

$$f_x = \frac{\alpha}{\lambda}, f_y = \frac{\beta}{\lambda} \tag{38}$$

By substituting equation 38 into equation 29:

$$U(x,y,0) = \iint_{-\infty}^{\infty} A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda},0\right)\exp\left[j2\pi\left(x\frac{\alpha}{\lambda}+y\frac{\beta}{\lambda}\right)\right]d\frac{\alpha}{\lambda}d\frac{\beta}{\lambda} \tag{39}$$

Equation 39 captures the main sentiment of the angular spectrum, that there is no difference between the two dimensional Fourier transform and the summation of multiple plane waves at different angles of propagation with respect to the optical axis.

1.29 Angular spectrum when *z* does not equal 0

Now that the angular spectrum has been explained, it is time to describe how the angular spectrum helps with the modelling of light propagation [109]. Considering the case of a plane wave with a position vector where *z* is not equal to zero (but still an arbitrary constant z'), the plane wave equation can be written as follows:

$$U(x,y,z') = U_0\exp\left(j\frac{2\pi}{\lambda}\left(\alpha x+\beta y+\sqrt{1-\alpha^2-\beta^2}z'\right)\right) \tag{40}$$

For the previous case where the position vector had a *z* term equal to zero, the plane wave equation could be written as:

$$U(x,y,0) = U_0\exp\left(j\frac{2\pi}{\lambda}(\alpha x+\beta y)\right) \tag{41}$$

In order to see how the complex field of the angular spectrum should be propagated we will divide the propagated plane wave by the unpropagated plane wave, in order see the propagation term which drops out:

$$H(0\rightarrow z') = \frac{U_0\exp\left(j\frac{2\pi}{\lambda}\left(\alpha x+\beta y+\sqrt{1-\alpha^2-\beta^2}z'\right)\right)}{U_0\exp\left(j\frac{2\pi}{\lambda}(\alpha x+\beta y)\right)} = \frac{U(x,y,z')}{U(x,y,0)} = \exp\left(j\frac{2\pi}{\lambda}z'\sqrt{1-\alpha^2-\beta^2}\right) \tag{42}$$

As previously defined, the angular spectrum is a summation of plane waves of different angles, where the summation is signified by a double integral. As we now understand how one plane wave propagates, it is a simple case of supplementing our previous angular spectrum equation with this propagation term. So this propagation term now enables the entire complex field to be propagated from *z* equals nought to an arbitrary constant. Using this propagation term in equation 39 the following result is obtained:

$$U(x,y,z) = \iint_{-\infty}^{\infty} A\left(\frac{\alpha}{\lambda},\frac{\beta}{\lambda},0\right)\exp\left[j\frac{2\pi}{\lambda}(\alpha x+\beta y)\right]\exp\left[j\frac{2\pi}{\lambda}z\sqrt{1-\alpha^2-\beta^2}\right]d\frac{\alpha}{\lambda}d\frac{\beta}{\lambda} \tag{43}$$

In order to see the effect of this propagator on the frequency domain there will be a conversion using equation 38:

$$U(x,y,z) = \iint_{-\infty}^{\infty} A(f_x, f_y, 0)\exp[j2\pi(f_x x + f_y y)]\exp\left[j\frac{2\pi}{\lambda}z\sqrt{1-(f_x\lambda)^2-(f_y\lambda)^2}\right]df_x df_y \quad (44)$$

Moreover, equation 44 suggests the angular spectrum propagator can be described as follows:

$$H\left(f_x, f_y, z\right) = \exp\left[j\frac{2\pi}{\lambda}z\left(\sqrt{1-(f_x\lambda)^2-(f_y\lambda)^2}\right)\right] \quad (45)$$

$$A(f_x, f_y, z) = A(f_x, f_y, 0)H(f_x, f_y, z) \quad (46)$$

So the wave which has propagated a distance *z* can be modelled by taking its Fourier transform at its previous position, multiplying by a propagation kernel, then taking the result through an inverse Fourier transform.

The final point to make about the angular spectrum is that there is no theoretical limit to its application, which is not true for Fresnel or Fraunhofer propagation. However it is limited in terms of implementation in software, due to the fact that the angular spectrum has a physical size in the spatial domain which is not inherently clear, as we only use the angular spectrum in the frequency domain. So by defining the angular spectrum in the frequency domain, its size in the spatial domain is defined and is proportional to the propagation distance *z*. When a large propagation distance is chosen, the spatial size of the propagation term may be larger than the area defined in the software being considered, which causes errors in the propagated object due to the assumption of periodicity implicit in the discrete Fourier transform. This means that features near the edge of a propagated image will 'wrap around' as the propagation distance increases and the features spread out. The mix of the Fourier repeat and the central image would cause an error, as the simulation would introduce data not possible in actuality. Although the application range can be increased by padding the data with a ring of zeros, this padding greatly increases the time the software takes to process the data, which is not practical when this propagator can be called upon hundreds to thousands of times to solve the phase problem. This error therefore limits the application of the angular spectrum propagator to short distances where Fourier repeats have little effect.

1.30 Fresnel and Fraunhofer propagation

Fresnel and Fraunhofer diffraction have roughly the same set-up and ideas as the angular spectrum case stated above [109]. The easiest way to distinguish when to use one or the other is by considering the Fresnel number of the experiment. Equation 47 describes how this Fresnel number can be calculated:

$$F = \frac{A^2}{\lambda Z} \quad (47)$$

The parameters of equation 47 are the following:

- A is the radius or half length of the object to be considered
- λ is the wavelength of light used
- Z is the propagation distance from the object of interest to the detector

Figure 30 then gives an example in which particular propagation techniques are accurate for a range of propagation distances and object sizes at a wavelength of 700nm. Outlining the differences between these propagation models is important to enable them to be used appropriately when an experimenter supplies a set of diffraction patterns. Furthermore, obtaining a path through which the final equations are obtained reveals the assumptions made and gives some insight into how all the

propagation models are connected. Explicit approximations of where the Fresnel and Fraunhofer propagation methods are appropriate will be given later.

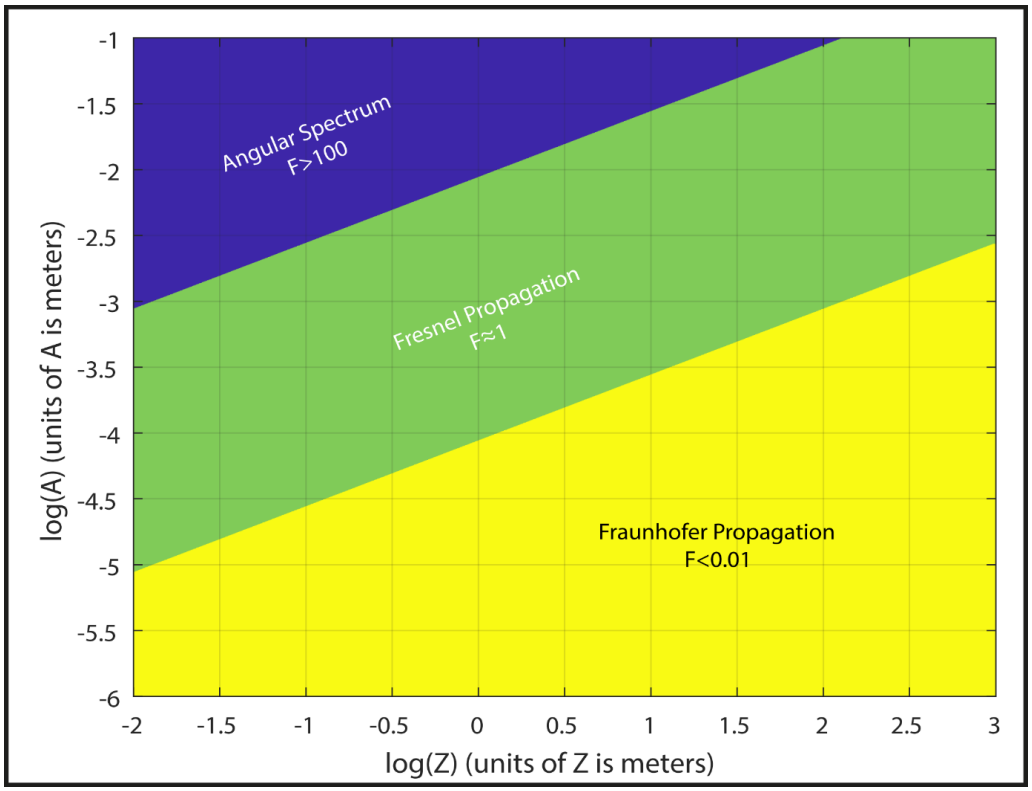

Figure 30: A figure showing appropriate propagation technique for a wavelength of 700nm for a given object size (A) and propagation distance(Z)

The Fraunhofer and Fresnel propagators both have their basis in the Rayleigh-Sommerfeld diffraction integral, which is mathematically described below:

$$P_2(x_2, y_2) = \frac{1}{j\lambda} \iint_{-\infty}^{\infty} P_1(x_1, y_1) \frac{\exp(-jkR)}{R} dx_1 dy_1 \tag{48}$$

$$\text{where } R = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + z^2} \tag{49}$$

Using the terms defined in Figure 27, where: x_2 and y_2 are the x-axis and y-axis in the detector plane, x_1 and y_1 are the x-axis and y-axis in the object plane, P_2 is the diffraction pattern at the detector plane and P_1 is the light at the object plane. Finally, *R* is the distance from any point in the object plane to any point in the detector plane. Another way to interpret the Rayleigh-Sommerfeld diffraction integral is to think of it as a discrete summation of spherical waves defined in section 1.27. So, P_1 determines the magnitude and position of these spherical waves. This means that the Rayleigh-Sommerfeld diffraction integral is very similar to the angular spectrum, in that the propagation of the complex field can be described as a summation of a number of simpler waves, in this case spherical waves.

1.31 Fresnel Propagation

The binomial expansion can be employed to simplify the integral in equation 48 [109]. The binomial expansion states that:

$$(p+q)^r = p^r + rp^{r-1}q + \frac{r(r-1)}{2!}p^{r-2}q^2 + \cdots \tag{50}$$

This can be applied to the equation for *R* above, as follows:

$$R = z\sqrt{1 + \left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2} \tag{51}$$

$$r = \frac{1}{2}, \quad p = 1 \text{ and } q = \left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2 \tag{52}$$

$$\sqrt{1 + \left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2} \cong 1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right] \tag{53}$$

$$\therefore R \cong z\left(1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]\right) \tag{54}$$

The assumption here is that *q* is small, due to the large size of *z* compared to the size of the object and detector plane.

Substituting equation 54 into equation 48:

$$P_2(x_2, y_2) \approx \frac{1}{j\lambda}\iint_{-\infty}^{\infty} P_1(x_1, y_1) \frac{\exp\left(-jkz\left(1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]\right)\right)}{z\left(1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]\right)} dx_1 dy_1 \tag{55}$$

For the denominator inside the integral, the assumption that the '1' dominates and the other terms can be ignored will be assumed to be valid. This assumption will be explored more in section 1.32:

$$P_2(x_2, y_2) \approx \frac{\exp(-jkz)}{j\lambda z}\iint_{-\infty}^{\infty} P_1(x_1, y_1) \exp\left(-\frac{jk}{2z}[(x_2 - x_1)^2 + (y_2 - y_1)^2]\right) dx_1 dy_1 \tag{56}$$

Now to expand the squared terms to see if it can be further rearranged:

$$P_2(x_2, y_2) \approx$$

$$\frac{\exp(-jkz)}{j\lambda z}\iint_{-\infty}^{\infty} P_1(x_1, y_1) \exp\left(-\frac{jk}{2z}[{x_2}^2 + {y_2}^2 + {x_1}^2 + {y_1}^2 - 2x_2x_1 - 2y_2y_1]\right) dx_1 dy_1 \tag{57}$$

The ${x_2}^2$ and ${y_2}^2$ terms can be moved outside the integral, so that $P_2(x_2, y_2)$ can be described as:

$$P_2(x_2, y_2) \approx \frac{\exp(-jkz)}{j\lambda z}\exp\left(-\frac{jk}{2z}({x_2}^2 + {y_2}^2)\right) \ldots$$

$$\ldots \iint_{-\infty}^{\infty} P_1(x_1, y_1) \exp\left(\frac{jk}{2z}[{x_1}^2 + {y_1}^2]\right) \exp\left(-\frac{jk}{z}(x_2x_1 + y_2y_1)\right) dx_1 dy_1 \tag{58}$$

Remembering that:

1. $P_2(x_2, y_2)$ is the image at the detector plane. This is the desired outcome because we want a relationship between the light at the object plane and the light collected at the detector plane.

2. $P_1(x_1, y_1)$ represents the object that diffracts light.
3. $\frac{\exp(-jkz)}{j\lambda z}$ is a term which describes the phase change as the light propagates in the *z* direction.
4. $\exp\left(-\frac{jk}{2z}({x_2}^2 + {y_2}^2)\right)$ is a term which describes an approximate phase curvature in the detector plane.
5. $\exp\left(\frac{jk}{2z}[{x_1}^2 + {y_1}^2]\right)$ is a term which describes the phase curvature in the object plane
6. $\iint_{-\infty}^{\infty} P_1(x_1, y_1) \exp\left(-\frac{jk}{z}(x_2 x_1 + y_2 y_1)\right) dx_1 dy_1$ can be described as the Fourier transform of the complex field at the object plane $P_1(x_1, y_1)$

Parts 3 and 4 make up what is defined in the literature as the Fresnel's propagator or convolution kernel (equation 59) [109]. This propagator has two software implementations, but these will not be discussed in this report:

$$h(x, y) = \frac{\exp(-jkz)}{j\lambda z} \exp\left(-\frac{jk}{2z}({x_2}^2 + {y_2}^2)\right) \tag{59}$$

Now there will be more in depth exploration of statement 6:

Firstly, using the assumption that *z* is very large with respect both to x_1 and x_2 [110], the following equations can be obtained:

$$\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{x_2}{z} \rightarrow \lim_{\theta \to 0} \tan\theta = \theta \tag{60}$$

$$\frac{x_2}{z} = \theta \tag{61}$$

Remembering equation 38 and the wave vector *k*:

$$f_x = \frac{\alpha}{\lambda}, f_y = \frac{\beta}{\lambda}, k = \frac{2\pi}{\lambda} \tag{62}$$

The parameters α and β are direction cosines. If these cosines can be approximated by their angles due to the small angle approximation, then the following is true:

$$\alpha = \theta \rightarrow f_x = \frac{x_2}{z\lambda}, f_y = \frac{y_2}{z\lambda} \tag{63}$$

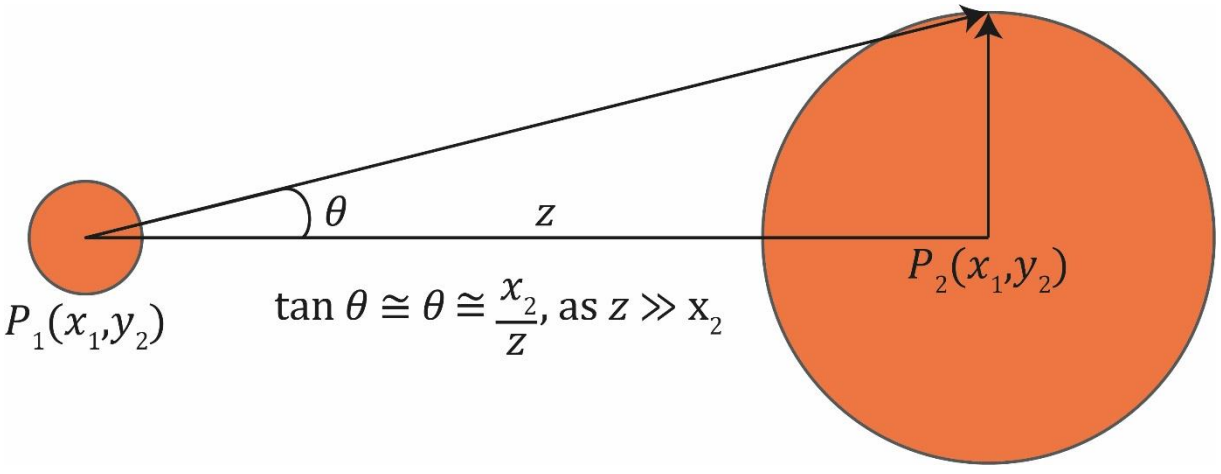

Figure 31: The mapping of spatial frequencies to physical distances

Substituting equation 63 into equation 58:

$$P_2(f_x, f_y) = \frac{\exp(-jkz)}{j\lambda z}\exp\left(-j\pi z\lambda\left[{f_x}^2 + {f_y}^2\right]\right) \ldots$$

$$\ldots \iint_{-\infty}^{\infty} P_1(x_1, y_1)\exp\left(\frac{jk}{2z}\left[{x_1}^2 + {y_1}^2\right]\right)\exp\left(-j2\pi\left[f_x x_1 + f_y y_1\right]\right)dx_1 dy_1 \quad (64)$$

Equation 64 is important for the physical aspect of the experiment. Because x_2 and *z* are physical distances which are orthogonal to each other, one can draw a right-angled triangle between planes 1 and 2. This triangle can be seen in *Figure 31* and is used to consider the angle subtended by the hypotenuse of the triangle. Using the small angle approximation as seen in equation 60, the direction cosine can be directly calculated, and therefore taking the *k* term of the exponential to convert to radians, the spatial frequency and the coordinates in plane 2 are shown to be directly related.

The above general analysis of the equation suggests that the propagation via the Fresnel approximation can be implemented using a modified Fourier transform of the original object, where the object is multiplied by the radial quadratic phase factor before the transform. After the Fourier transformation, a second quadratic phase factor is applied. The assumptions made during the conversion between the Rayleigh-Sommerfeld diffraction integral and the Fresnel diffraction will be explored in the next section.

1.32 For what range of *z* is the Fresnel description of diffraction valid?

In formulating the Fresnel diffraction integral (equation 58), a key assumption was the use of the binomial expansion [109]. The question is, whether ignoring the additional terms of the expansion (equation 50) has had a significant effect on the exp*(-jkR)* term of equation 48 , and therefore an effect on the overall result. Looking at the binomial expansion again and looking up to the third term:

$$\sqrt{1 + \left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2} \cong 1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right] - \frac{1}{8}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]^2 \quad (65)$$

Remembering the original context (equation 48) in which the assumption was made, the result of the difference between the approximation and the actual value of *R* is a difference in phase at plane 2. Typically, the approximation is seen to be appropriate, if the difference is much less than one radian between the approximation with two terms (R_{2nd}) and the approximation with three terms (R_{3rd}) as stated by Goodman in [109]. As the much less symbol (<<) is a non-explicit value, it is substituted with a factor of 1/100 within this thesis (see equation 69). The phase shift between the terms is defined in equation 66 and then is used to obtain equation 71, which is a useful 'back-of-the-envelope tool' to work out whether the Fresnel propagator is the correct propagator for the given constraints. The constraints being the propagation distance *z* and the maximum translation in the x-y axis between plane 1 and plane 2, this means that for a given propagation distance, the translation between planes can cause a violation of earlier assumptions.

$$\varphi = \angle\left[\frac{\exp(-jkR_1)}{\exp(-jkR_2)}\right] = \angle\left[\exp\left(-jk(R_{3rd} - R_{2nd})\right)\right] \quad (66)$$

$$R_2 - R_1 = z\left(1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right] - \frac{1}{8}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]^2\right) \ldots$$

$$-z\left(1 + \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]\right) \quad (67)$$

$$\angle\left[\exp\left(-jk(R_{3rd} - R_{2nd})\right)\right] = \frac{kz}{8}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right]^2 \quad (68)$$

$$k = \frac{2\pi}{\lambda}, \ll = \frac{1}{100} \quad (69)$$

$$\frac{1}{100} \geq \frac{\pi}{4\lambda z^3}[(x_2 - x_1)^2 + (y_2 - y_1)^2]^2 \quad (70)$$

$$z^3 \geq \frac{100\pi}{4\lambda}[(x_2 - x_1)^2 + (y_2 - y_1)^2]_{max}{}^2 \quad (71)$$

Let us take an example where $\lambda = 700 \times 10^{-9}\text{m}, x_1 = y_1 = 300 \times 10^{-6}\text{m}, x_2 = y_2 = 500 \times 10^{-6}$ m. An appropriate *z* would be greater than the following value:

$$z^3 \geq \frac{100\pi}{4 \times (700 \times 10^{-9})}[(200 \times 10^{-6})^2 + (200 \times 10^{-6})^2]^2 \quad (72)$$

$$z \geq 8.96\text{mm} \quad (73)$$

Also, in the previous section we made the assumption that:

$$1 \gg \frac{1}{2}\left[\left(\frac{x_2 - x_1}{z}\right)^2 + \left(\frac{y_2 - y_1}{z}\right)^2\right] \quad (74)$$

Equation 74 can be rearranged into equation 75:

$$z^2 \gg \frac{1}{2}[(x_2 - x_1)^2 + (y_2 - y_1)^2]_{max} \quad (75)$$

on inspection it can be seen that equation 71 is a tighter constraint on the value of *z*, therefore if the criteria of equation 71 is met, so is the above assumption. This can also be shown by our example:

$$(8.955 \times 10^{-3})^2 = 80.19 \times 10^{-6} \gg (200 \times 10^{-6})^2 = (40 \times 10^{-9}) \quad (76)$$

It can be seen from the example above that the approximation gives a value for the propagation distance that is plausible for the near field condition. In the next section on Fraunhofer diffraction the required *z* value will be increased, as further assumptions are made based on the size of *z*.

1.33 Fraunhofer propagation

Fraunhofer diffraction is a special case of the Rayleigh-Sommerfeld diffraction formula, where *z* is so large that the phasor term $\exp\left(-\frac{jk}{2z}[{x_1}^2 + {y_1}^2]\right) \to 1$ [109, 110]. Using the starting point of equation 58 and removing the previously mentioned phase term, equation 77 is obtained:

$$P_2(x_2, y_2) =$$

$$\frac{\exp(-jkz)}{j\lambda z}\exp\left(-\frac{jk}{2z}({x_2}^2 + {y_2}^2)\right)\iint_{-\infty}^{\infty} P_1(x_1, y_1)\exp\left(-\frac{jk}{z}(x_2x_1 + y_2y_1)\right)dx_1dy_1 \quad (77)$$

So we can see from equation 77 that P_2 (or diffraction pattern) is equal to:

1. A phasor term which is proportional to *z*

2. A phasor term which is dependent on the final position of the light rays themselves in the detector/frequency plane
3. A Fourier transform of the object in the object plane.

A benefit of Fraunhofer propagation is how easy the propagator is to implement, as it involves only simple terms in front of the Fourier transform. As the Fraunhofer model of diffraction has used the assumption that *z* is sufficiently large to further simplify the propagation equation, the conditions under which it is valid must be reassessed. Re-evaluating the distance at which equation 77 is appropriate to use is relatively simple, as all that is required is to find a *z* value which causes the phase curvature in the specimen plane not to change significantly over P_1. Goodman suggests that a phase difference much less than 1 rad between the case where the phase curvature at the specimen is included, compared to its absence, is sufficient justification for simplification of a propagation function [109]. The phase curvature can be simplified, assuming the phase curvature varies less than 10 mrad over the specimen plane, then it is simply the case that *z* has to be a hundred times greater than the rest of the phase curvature terms:

$$\frac{k}{2z}[{x_1}^2 + {y_1}^2] \cong \frac{1}{100} \tag{78}$$

$$z \geq \frac{100\pi}{\lambda}[{x_1}^2 + {y_1}^2] \tag{79}$$

Let us take an example where $\lambda = 700 \times 10^{-9}\text{m}, x_1 = y_1 = 300 \times 10^{-6}\text{m}$. Therefore, an appropriate *z* would be greater than the following value:

$$z \geq \frac{100\pi}{700 \times 10^{-9}}\left[300 \times 10^{-6^2}\right] \times 2 \tag{80}$$

$$z \geq 80.784\text{m} \tag{81}$$

This value is far greater that any previously required value *z* for the Fresnel diffraction condition (this number is larger than expected due to the large pixel size used in this example).

1.34 Discussion of the different propagation methods

As *z* has increased, the propagation method has become simpler and closer to the Fourier transform. The discussion started with the angular spectrum method where the diffraction pattern at the second plane was a blurred version of the original object. Then the propagated fields of the object become completely unrecognisable in the Fresnel and Fraunhofer domain, as only a single Fourier transform occurs. It has also been shown that the angle at which a plane wave of light propagates away from an object corresponds to a single location in the Fourier domain. This core idea of mapping spatial frequencies to physical space was also covered by Figure 31 and equation 63. It is also important to remember the key message of the angular spectrum, that all complex fields can be described by a summation of plane waves. Furthermore, Fresnel and Fraunhofer propagation have both been derived from the Rayleigh-Sommerfeld diffraction integral, and estimates of where each method is appropriate have been provided.

Now with an understanding of coherent complex fields and how they propagate, this literature review turns its attention to incoherence. Incoherence affects the diffraction patterns collected in experiments by blurring the collected data.

1.35 Incoherence

Now that we have a clear understanding of how coherent waves propagate through space and how various diffraction patterns are formed, it is time to consider when optical conditions are less ideal.

As electron sources such as cold field emission guns and Schottky emission guns have finite sizes, their spatial incoherence must be considered. Furthermore, laser beams and electron sources emit waves with varying energy, therefore the illumination on the sample is a mix of wavelengths which causes blurring of the diffraction patterns. This is termed temporal incoherence.

Temporal incoherence can be illustrated with an example: imagine a laser that emits a laser beam with two discrete wavelengths shining on a plate which has two small slits etched into its surface. After the slits is a wall/detector which will be used to observe the diffraction pattern formed (this is Young's double slit experiment [5]). The distances between the laser light source and the slits can be considered sufficiently long to be modelled as a Fourier transform, and this is also true of the distance been the slits and the wall. The positions of the minima and the maxima of the Young's double slit experiment is dependent on the wavelength of the illumination used [5]. If the wall was capable of filtering out one wavelength, we would see a perfect diffraction pattern of the slits. However, if the wall was a realistic detector, it would have to collect data over a period of time, during which it would collect the average counts from both wavelengths. The effective displacement between the two experiments would cause the collected diffraction pattern to be blurred compared to our ideal case. This blurring of the collected data is the effect of incoherence and the cause of this incoherence is temporal, as the underlying cause is the different wavelengths that are being emitted by the source.

Spatial incoherence can be considered using the same Young's double slit experiment, except that the light source is now composed of multiple light sources placed next to each other, and the light sources emit only one wavelength. This effectively means the illumination coming from the laser can no longer be considered a pure plane wave. Instead, it should be considered as many plane waves of varying angles incident on the slits. As some of the illumination approaches the slits at an angle, a phase ramp is formed across the slits, which in turn causes the diffraction pattern of the slits on the wall to shift. Therefore, the diffraction pattern on the wall is formed of many shifted versions of the same diffraction pattern, causing the data collected by a detector to be blurred. The blurring this time is caused by the extended nature of the source, and is termed spatial incoherence.

Spatial and temporal incoherence demonstrate the general principle of incoherence: incoherence describes how the imperfect nature of the source can reduce the resolution of the experiment. The imperfection of the source is modelled by considering the source as a sum of perfect sources (or ensemble) displaced in terms of space (angle) or wavelength (energy).

Incoherence can be measured and experienced in the real world by various experiments like Young's double slit experiment. The electron implementation of the Young's double slit experiment is off-axis electron holography (see section 1.11), where the visibility of the interference fringes of the hologram inform the experimenter of the coherence of the electron source and can be calculated by the simple equation below from [12] (see equation 82):

$$\eta = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \propto \mu \tag{82}$$

The parameters of equation 82 are the following:

- η is the measured visibility of the off-axis hologram
- I_{max} is the maximum intensity of the hologram
- I_{min} is the minimum intensity of the hologram
- μ is the term which describes the damping factor of the incoherence (i.e. how much the fringe contrast is reduced compared to the ideal case)

This is also considered in the equation describing simulated holograms, as they include a coherence damping function [111]. This is shown in equation 83 for clarity.

$$I = A_1 + A_2 + 2\mu A_1 A_2 \cos(2\Delta\phi) \tag{83}$$

The parameters of equation 83 are the following:

- A_1 is the intensity of the reference wave
- A_2 is the intensity of the sample wave
- μ is the coherence term
- $\Delta\phi$ describes the phase difference between the two waves; this represents the interference fringes of the hologram.

The transverse coherence length of a source is often described as the distance between two points (in the specimen plane) which still can interfere at the detector/wall [112, 113]. The Young's double slit experiment is often used to demonstrate the transverse coherence length as the spacing between the slits can be varied in order to observe the effect of coherence on the diffraction of the slits. Figure 32 shows how the Young's double slit experiment can be used to define the transverse coherence length. Figure 32 illustrates that by equating the angle subtended by the outermost part of the incoherent source and the angle subtended by the 1st minima on the detector, a basic equation for the transverse coherence length can be obtained. Essentially by matching the angle of the incoherent source to the slit width which causes destructive interference, the mathematical description of the transverse coherence length is obtained (see equations -86). It is assumed throughout that the distances D_1 and D_2 are sufficiently long that the angles considered here are appropriately small to meet the small angle approximation.

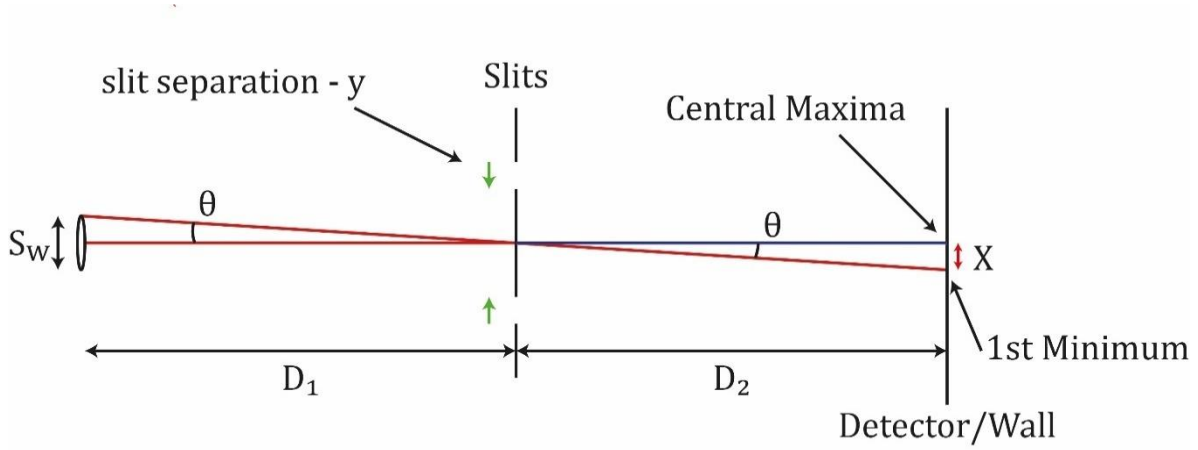

Figure 32: A diagram showing how equating the angle subtend by an incoherent source to the destructive interference of Young's double slit enables the calculation of the transverse coherence length [114].

$$\theta = \frac{S_w}{2D_1} \tag{84}$$

$$\theta = \frac{\lambda}{2y} \tag{85}$$

$$\therefore y = \frac{\lambda D_1}{S_w} \tag{86}$$

The parameters of equations 84-86 are the following:

- S_w is the physical width of the incoherent source
- D_1 is the distance between the source and slits along the optical axis
- λ is the wavelength of the source
- y is the width/distance between the slits. In the case of equation 86, y is also the transverse coherence of the source.
- θ is both: the angle subtended by furthest source element (see equation) and the angle subtended by the first local minima at the detector (See equation 85).

The transverse coherence length is important to near field ptychography, as in effect it limits the available useful propagation distance which can be implemented [115]. This is because if the first Fresnel zone [116] is larger than that of the transverse coherence length then the collected diffraction data will be highly attenuated and diffraction fringes of the diffracted specimen will not be able to move far beyond the edge of the specimen, as illustrated by Figure 33. The intensities shown in Figure 33 a) through c) are the result of the simulated glass cylinder surrounded by an opaque ring shown in d) and e) being propagated a distance *z* but illuminated with sources of varying incoherence. Sources with short transverse coherence length make it difficult for near field ptychography to achieve the condition where the diffraction fringes of the probe and specimen interact. An equation calculating the size of the first Fresnel zone can be seen in equation 87.

$$FFZ = (\lambda z)^{-\frac{1}{2}} \tag{87}$$

The parameters of equation 87 are the following:

- FFZ is the radius of the first Fresnel Zone
- λ is the wavelength of the illumination
- z is the propagation distance between the specimen and the detector

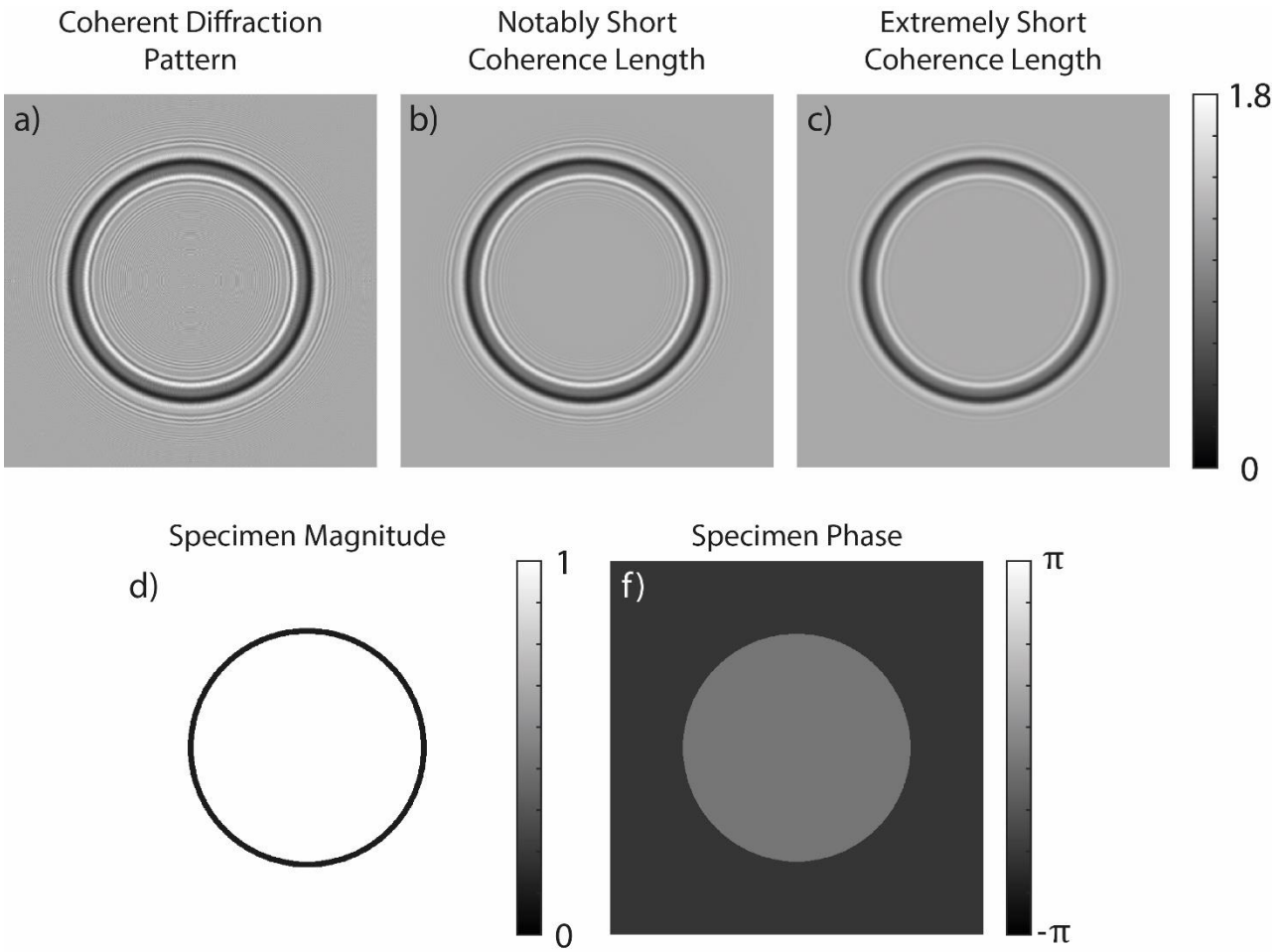

Figure 33: a figure to illustrate the effect of spatial incoherence (a short transverse coherence length) on collected intensities. a) a coherent diffraction pattern produced by propagating the specimen shown in d) and e) an arbitrary distance z. b) a collected intensity affected by spatial incoherence. c) a collected intensity severely affected by spatial incoherence. d) and e) show the magnitude and phase of the specimen used to simulate a)-c). the colour bar to the right of a)-c) converts the greyscale into normalized intensity. The colour bar to the right of d) converts the greyscale of d) into normalized magnitude. The colour bar to the right of f) converts the greyscale of f) into phase in terms of radians

In this section a basic understanding of coherence has been discussed, the two types of coherence have been outlined, and the effect of varying transverse coherence length has been illustrated. Although temporal coherence has been discussed, spatial coherence will be a bigger factor in the results section of this thesis, as will be discussed in section 4.12.

Now that propagation, incoherence and ptychography have been explored, it is a case of understanding Fourier Ring Correlation (FRC), a relatively old idea conceptually, but one which is finding new applications in the area of phase retrieval techniques, as the Fourier Ring Correlation is a method to measure the acceptable resolution of a reconstruction, as well as assessing the reconstruction quality as a function of frequency.

As an aside, if a near field ptychography dataset is affected by incoherence, it can be dealt with in an addition to the rPIE algorithm. This thesis makes use of the method laid out in the supplement material of [94] to account for the effect of incoherence on the reconstruction. Detailing the algorithms used in [94] are beyond the scope of this thesis.

1.36 Fourier Ring Correlation

Having introduced phase retrieval, and specifically near field ptychography, the objective of this section is to set out ways of being able to compare multiple complex images to each other, to evaluate their consistency (when using the same phase retrieval technique) and their differences when using different techniques.

The Fourier Ring Correlation (FRC) is set out in equation 88. The FRC is a method by which to compare two different images of the same specimen with each other, measuring the agreement in frequency space. In the case of ptychography in general two different images can be obtained of the same specimen, if a dataset is split in half and the halves are then independently reconstructed by the rPIE algorithm. This process is outlined in Figure 34. The splitting of the dataset is typically executed such that one half of the dataset contains only odd diffraction patterns, and the other half consists of only even diffraction patterns. This of course requires that both halves of the dataset have sufficient redundancy in order to constrain the phase problem when reconstructed by the rPIE algorithm. It is also important to note the effective step size is doubled when splitting a dataset in this way. Before performing the FRC, pre-processing steps are required to remove any ambiguities between the two reconstructions, such as displacements in real space and phase ramps. Once the ambiguities are removed, each reconstruction from each dataset half is Fourier transformed, and once it has been Fourier transformed the odd Fourier transform is multiplied by the complex conjugate of the even Fourier transform (see equation 88, remember this is equivalent to a cross-correlation in the spatial domain, hence the correlation in FRC). Now consider the phase of a single pixel of this multiplication; if the two reconstructions are identical, the result of the multiplication is a purely real number with zero phase, but if reconstructions are not identical the result of the multiplication will be a complex number with a non-zero phase. As the reconstructions are now in Fourier space this gives us additional vectors through which the reconstructions can be considered, most importantly that a ring in Fourier space centred around the origin contains information on a particular spatial frequency. Taking a sum of the multiplication along a particular set of frequencies (or a ring), a value is obtained proportional to the agreement of the reconstructions for that particular spatial frequency. Consider again the case of identical reconstructions: the sum along a particular frequency ring would be of purely real numbers, meaning the magnitude is equal to the real component (see Figure 35 a)). If the reconstructions were not identical, the sum would consist of a sum of complex numbers with varying phases, meaning the magnitude of this sum of complex numbers can vary greatly depending on the complex numbers themselves. This is illustrated by Figure 35 b), where the magnitude of the sum of the real numbers is compared to a possible magnitude of a sum of complex numbers. By the mechanism of complex addition, the agreement of a particular spatial frequency between two images can be assessed. Only the numerator of equation 88 has been considered, and the denominator simply acts as a normalising factor ensuring the end result is between zero and one. The denominator achieves this goal by having the same magnitude as if the two images were considered identical.

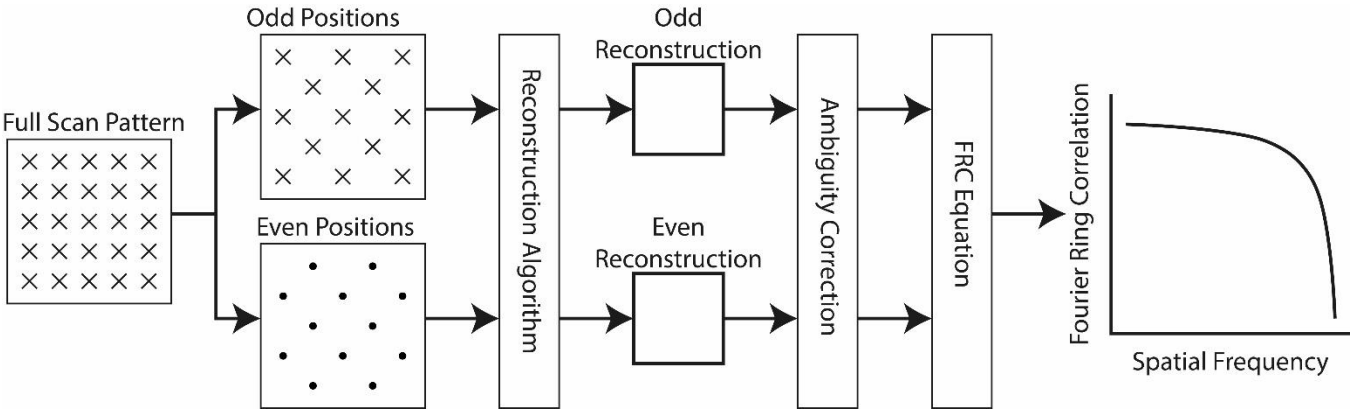

Figure 34: an illustration of how FRC analysis can be applied to ptychographic dataset. First the dataset is split into odd and even diffraction patterns and then each dataset is reconstructed independently. The odd and even reconstructions are then compared to each other such that any ambiguities are removed, then the two reconstructions are compared via FRC obtaining an FRC plot with respect to the spatial frequencies of the specimen.

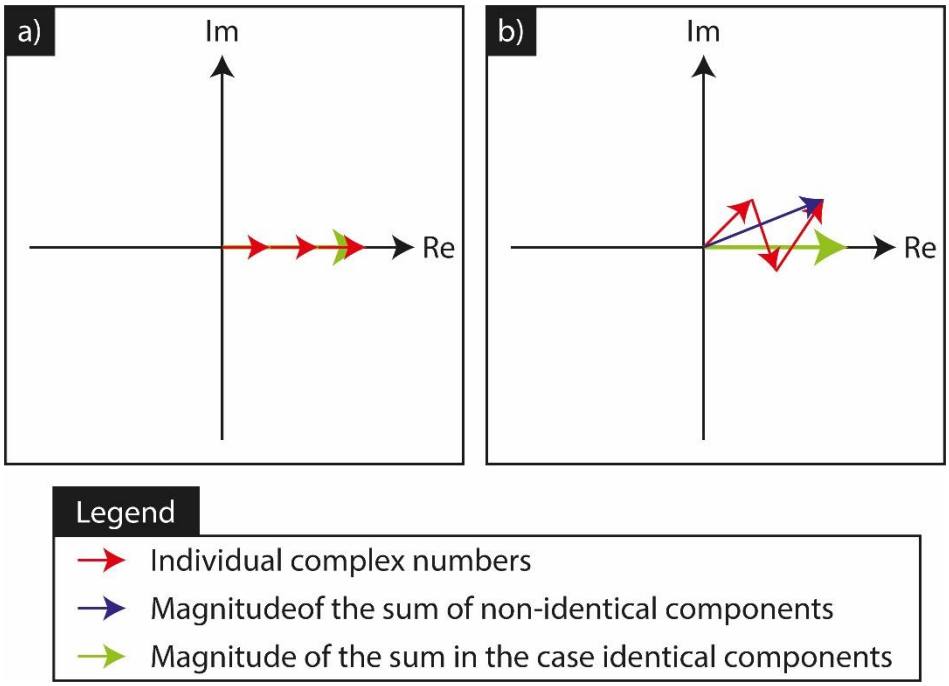

Figure 35: a) an illustration of the sum of real numbers along a ring in Fourier space when the FRC process is performed on identical images. b) an illustration of the sum of complex numbers along a ring in Fourier space when the FRC process is performed on non-identical images.

1.37 The Importance and history of the Fourier Ring Correlation

The Fourier Ring Correlation (FRC) has its origin in a paper published by W.O. Saxton and W. Baumeister, who investigated bacterial proteins and how their regular structure can be used to form an averaged image [117]. The Fourier Ring Correlation was first applied to ptychography in "The Role of the illumination spatial frequency spectrum for ptychography", by Manuel Guizar-Sicairos et al. [78]. The FRC has also been applied to through focal series phase retrieval (TIE) in "Multi-focus TIE algorithm including partial spatial coherence and overlapping filters", where it was used to compare different reconstruction methods [51]. The Fourier Ring Correlation method has the following aspects:

1. It requires two sets of similar reconstruction data which have gone through the same process (self-referencing).

2. It involves splitting of the Fourier transform of reconstructions into rings with the width of a single pixel.

The FRC is defined as follows:

$$FRC(a) = \frac{\sum_{a \in a_i} F_1(a) {F_2}^*(a)}{\sqrt{\sum_{a \in a_i} |F_1(a)|^2 \sum_{a \in a_i} |F_2(a)|^2}} \tag{88}$$

Here, a designates a ring in Fourier space where all pixels in that ring have the same spatial frequency. F_1 is the Fourier transform of a first image reconstruction and ${F_2}^*$ is the complex conjugate of the Fourier transform of a second reconstruction. The FRC is effectively the normalised cross-correlation of the spatial frequency content of the two images [118].

One of the most important aspects of the FRC is the idea that a ring in Fourier space can be linked together by one property; that they all belong to a single frequency [119]. Using this aspect, we can increase the reliability of the analysis we perform. The reliability is increased by taking into consideration every pixel in the ring when calculating the FRC, and like all forms of statistical analysis the conclusion of analysis is stronger when it is formed with more data points. However, this leads to the fact that the reliability of the FRC is proportional to the frequency considered. This will be discussed more in section 1.38.

The FRC also gives some insight into the resolution of the images, as the Fourier Ring Correlation goes to zero at higher frequencies [118] as this is where the signal to noise ratio is lower and the number of data points is large, so the chances of the two datasets correlating is reduced. It is useful to understand the power spectral density of the image, before performing an FRC on it, as the information content changes with frequency, the image may not have significant information in certain frequencies, meaning it will not correlate well with similar images due to the domination of noise at those frequencies.

Fourier Ring Correlation is appropriate for ptychography and other software-based phase retrieval systems, because unlike traditional imaging there is effectively no system or optics that can cause systematic aberrations in the traditional manner (there is no spherical aberration, chromatic aberration or astigmatism in a lens-less system) [119]. Moreover, the system can only fail in terms of how we constrain the system: poor constraints do not cause the system to blur, assuming probe positions are correct or can be corrected by appropriate position correcting software [66]. Rather, the effect of poor constraints is that the reconstruction becomes noisy or to is trapped by incorrect local minima. Therefore, FRC's ability to test the consistency of reconstructions is more important for ptychography than for traditional resolution definitions.

1.38 The effect of noise and the requirement for thresholds in the FRC

As stated before, when noise dominates a signal it is often an assumption that two such signals would not correlate with each other. This is why high frequencies tend to correlate less, as they do not correspond to the majority of information/power within the image, and therefore the noise dominates. The subject of noise and the FRC/FSC (Fourier Shell Correlation) has been covered by Marin van Heel and Michael Schatz, and the author would encourage the reader to refer to their paper [118]. However, this thesis will cover some of the same material, hopefully giving insight when considering ptychographic reconstructions.

As discussed in reference [118], the correlation of any given ring is affected by the number of pixels within that ring. Therefore, low frequency rings which have fewer pixels are more susceptible to

noise as there are fewer opportunities/pixels which can cause the noise to decorrelate. The importance and success of the Fourier Ring Correlation comes from the fact that it relates directly to the spectral signal to noise ratio, a noise measurement often used by electron microscopy communities in biology and chemistry [118, 119]. The spectral signal to noise ratio is effectively the signal to noise ratio of a given spatial frequency. The simplest conversion method between the two is described by the following equations:

$$FRC = \frac{SSNR}{SSNR + 1} \tag{89}$$

$$SSNR = \frac{FRC}{1 - FRC} \tag{90}$$

The parameters of equations 89 and 90 are the following:

- SSNR is the spectral signal to noise ratio
- FRC is the Fourier Ring Correlation

These equations for the conversion between FRC and SSNR is an effective simplification of the equation for noise threshold of FRC reliability [119] under the assumption that the number of pixels under consideration is large with respect to signal to noise ratio or unity, whichever is bigger. Equation 91 is typically used to calculate thresholds by which data are deemed either reliable or unreliable, and where the FRC crosses this threshold curve is often stated as the resolution of the image which is useful for analysis:

$$FRC(f_r) = \frac{SNR(f_r) + \frac{2\sqrt{SNR(f_r)}}{\sqrt{n(f_r)}} + \frac{1}{\sqrt{n(f_r)}}}{SNR(f_r) + \frac{2\sqrt{SNR(f_r)}}{\sqrt{n(f_r)}} + 1} \tag{91}$$

$$\text{large number of pixels assumption } \sqrt{n(f_r)} \gg 2\sqrt{SNR(f_r)} \text{ or } 1$$

The parameters of equation 91 are the following:

- *SNR* is the signal to noise ratio
- f_r is frequency (radius) under consideration
- $n(f_r)$ is the number of pixels in a particular frequency ring

1.39 The effect of the ambiguities of ptychography on the FRC

As seen in the sections on ptychography and constraints of phase retrieval systems (section 1.18), phase retrieval systems can form a solution which is not necessarily wrong, but is instead the correct solution flipped and with inverse phase. The complex conjugate ambiguity is one example of an ambiguity in phase retrieval systems. Ambiguities in the reconstructions could cause the FRC to give a misleading result in terms of how well two images correlate to one another. For example, two images which are the same in every respect apart from a relative shift could cause them not to correlate. Therefore, there should be some investigation into what ambiguities can occur in ptychography specifically, and whether they affect the result of our comparison method (FRC). To start this investigation the ambiguities of ptychography will be stated, starting with equation 92 from [107].

$$P_c O_c = \left(u \exp(jv)\exp(jsc)\hat{P}_{c+t}\right)\left(u^{-1}\exp(-jv)\exp(-jsc)\hat{O}_{c+t}\right) \tag{92}$$

The components of equation 92 are as follows:

- u is an ambiguous magnitude error which occurs due to the fact that it is the modelled exit wave which has to match the collected intensity, not the probe and object individually. For example, the probe can accumulate a large magnitude due to u whilst the magnitude of the specimen is diminished by u^{-1}.
- The ambiguous phase offset of $\exp(jv)$ is similar to u in that the exit wave has to have a certain phase to have a low error, but the way in which that phase is achieved is arbitrarily shared between the probe and object.
- The hatted values $\hat{P}$ and $\hat{O}$ are the unambiguous probe and object.
- The ambiguous phase ramp $\exp(jsc)$ applied to the specimen is a result of a coordinate shift in the Fourier domain, and the opposite phase ramp affects the probe ensuring the simulated intensity has low error with respect to the collected data.
- s denotes the magnitude and direction of the phase ramp.
- Finally, the object itself could be shifted from its true position (c) by a number of pixels denoted by t in the spatial domain, which is due to an ambiguous phase ramp in the Fourier domain. Note that a phase ramp in the Fourier domain causes no change in the calculated error, as only intensities are compared.

As we are considering the FRC there are two objects, so their ambiguities will combine together, the two magnitude errors will multiply with each other, phase shifts will add together, phase ramps will combine into a single different phase ramp, and the shifts in the Fourier domain will also add together into an overall shift between the two images, as stated in equation 95 (the subscripts 1 and 2 are used to denote different numerical values of the same ambiguity). The spatial ambiguities convert into the same ambiguities in the Fourier domain, except that the phase ramps and shifts are swapped as shown by Fourier shift theory (equations 93 and 94):

$$\mathcal{F}[f(x-p,y-q)] = F(f_x,f_y)\exp\left(-j2\pi\left(pf_x+qf_y\right)\right) \tag{93}$$

$$\mathcal{F}^{-1}\left[F\left(f_x-p,f_y-q\right)\right] = f(x,y)\exp\left(-j2\pi(px+qy)\right) \tag{94}$$

The parameters of equations 93 and 94 are the following:

- p and q are coefficients of shifts/ramps in real space (units of meters/pixels) or Fourier space (units of m^{-1}).
- p corresponds specifically to either the x or f_x axis, while q corresponds specifically to either the y or f_y axis.
- $\mathcal{F}$ denotes the fast Fourier transform.

$$\begin{aligned} OO^* &= \left(u_1\exp(jv_1)\exp(js_1c)\hat{O}_{c+t_1}\right)\times\left(u_2\exp(jv_2)\exp(js_2c)\hat{O}^*{}_{c+t_2}\right) \\ &= u_1u_2\exp(j[v_1+v_2])\exp(jc[s_1+s_2])\,\hat{O}_{c+t_1}\hat{O}^*{}_{c+t_2} \end{aligned} \tag{95}$$

Now that the ambiguities of ptychography have been understood it is a matter of putting them through the FRC to see which ones have an effect on the outcome. Phase offsets cannot be easily removed mathematically, but also have no meaningful effect on the FRC, as the meaningful part of the FRC is the magnitude. Magnitude errors are dealt with through the normalisation process of the FRC and can be directly cancelled out. The FRC is strongly affected by the phase ramps and shifts in the data in the Fourier domain. Other errors include the fact that the FRC must be performed with

the same data in both images, as any data in one image which are not in the other is a form of noise when performing the FRC. Using this understanding, the equation for FRC with ambiguous reconstructions can be simplified to the version seen in equation 97 from the previous version seen in equation 96.

$$FRC(a) = \frac{\sum_{a\in a_i} \mathcal{F}\left[u_1 e^{jv_1} e^{js_1 c} \hat{O}_{c+t_1}\right] \mathcal{F}\left[u_2 e^{jv_2} e^{js_2 c} \hat{O}_{c+t_2}\right]}{\sqrt{\sum_{a\in a_i} \left|\mathcal{F}\left[u_1 e^{jv_1} e^{js_1 c} \hat{O}_{c+t_1}\right]\right|^2 \sum_{a\in a_i} \left|\mathcal{F}\left[u_2 e^{jv_2} e^{js_2 c} \hat{O}_{c+t_2}\right]\right|^2}} \tag{96}$$

$$FRC(a) = \frac{\sum_{a\in a_i} (e^{jfct_1} F_1(a)_{fc+s_1})(e^{jfct_2} {F_2}^*(a)_{fc+s_2})}{\sqrt{\sum_{a\in a_i} \left|e^{jfct_1} F_1(a)_{fc+s_1}\right|^2 \sum_{a\in a_i} \left|e^{jfct_2} {F_2}^*(a)_{fc+s_2}\right|^2}} \tag{97}$$

The parameters of equation 97 are the following:

- $e^{jf_c t}$ is a phase ramp in the Fourier domain due to a shift in the spatial domain of size *t*
- F_{f_c+s} denotes the Fourier transform of *O* which has been shifted from the origin of the Fourier domain by *s* due to a phase ramp in the spatial domain. (Again, subscripts indicate different numerical values)
- a denotes a particular ring in frequency space over which FRC analysis is being performed.

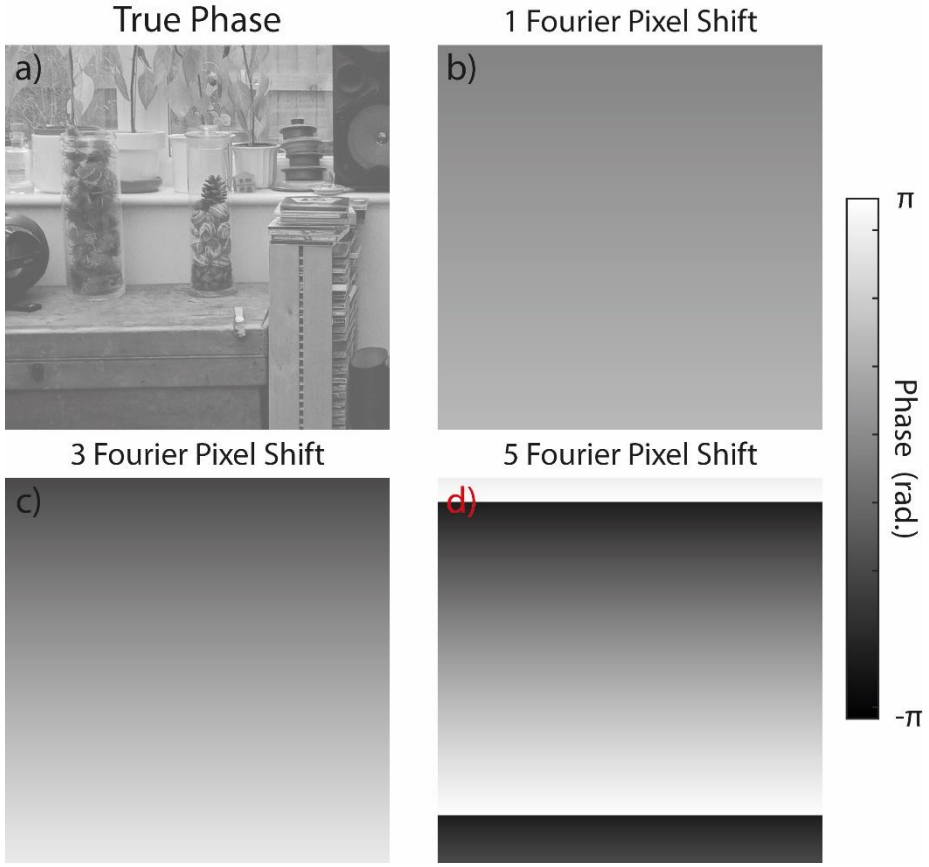

Figure 36: a) the true phase of the simulated specimen used to investigate the effect of real space phase ramps on the Fourier Ring Correlation of an otherwise identical specimen. b) the phase ramp in real space caused by moving the magnitude of an image in Fourier space by one pixel with respect to the centre. c) the phase ramp in real space caused by moving the magnitude of an image in Fourier space by three pixels with respect to the centre. d) the phase ramp in real space caused by moving the magnitude of an image in Fourier space by five pixels with respect to the centre. The colour bar to the right coverts the greyscale of a) to d) in phase in terms of radians.

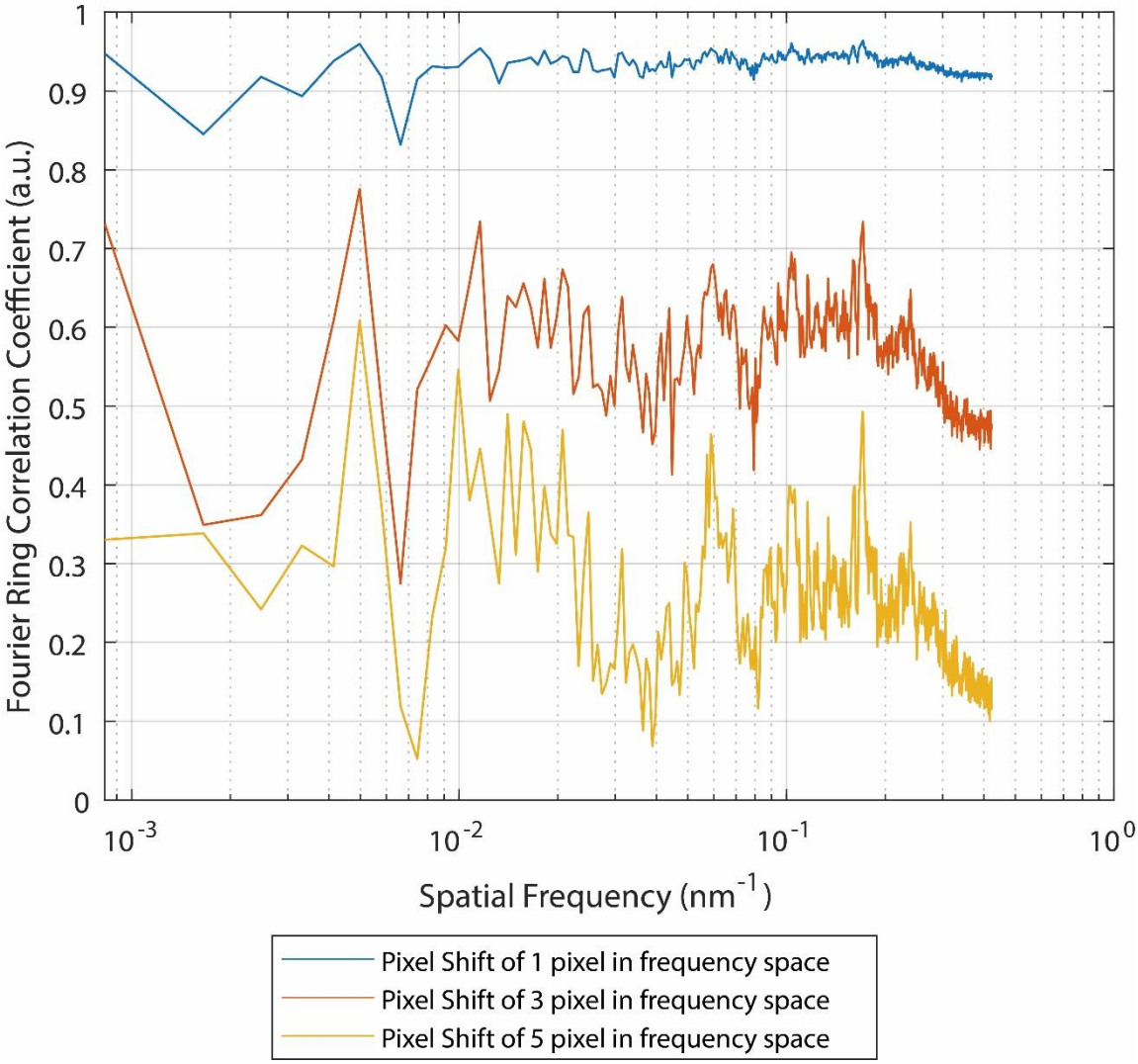

Figure 37: A figure to show the effect of pixel shifts in frequency space on the outcome of the FRC of two otherwise identical images. The shifts are described in term of pixels. The spatial frequency sampling in this figure is based on the diffraction conditions laid out in Section 4.5.

The effect of a misalignment between the reconstructions in the Fourier domain can be simulated in software by multiplying an image by a phase ramp in the spatial domain and performing an FRC with the unaltered image. This is done to obtain Figure 37. The true phase of the simulated specimen is shown in Figure 36 a) while the phase ramps applied are shown in Figure 36 b)-d). Figure 37 suggests that in general it is the lower frequencies which are most effected by a pixel shift in the Fourier domain, as shown by the steep drops in the FRC coefficient in all of the shifts considered. Further to this there is an almost constant reduction which is applied to all frequencies equally, which seems proportional to the magnitude of the pixel shift. Furthermore, the result does not have a strong dependence on the direction of the pixel shift, as the same outcome was obtained regardless of the direction (this is omitted from the figure for simplicity), although this will be object dependent.

Phase ramps in the Fourier domain are due to shifts in the spatial domain, as stated by Fourier shift theorem as shown by equations 93 and 94.

Phase ramps in the frequency domain can have two different effects on the FRC; one is reducing the FRC by applying different phase shifts along the ring, and the other is introducing new information into the calculation (the shift in the spatial domain). If the phase ramp goes through 2π within the

extent of the image, the value of the FRC will oscillate. Equation 98 gives an approximation of the FRC value in the case of two images which should perfectly correlate, except that one has had a phase ramp applied to it. The reliability of the approximation degrades as the correlation between the two images becomes less ideal. The importance of phase ramps in FRC comes from the necessity to align data before the FRC is processed (reduce the spatial shift to zero). If the data are not aligned there will be a phase ramp and different data.

$$FRC(a) = \frac{\sum_{a\in a_i} \exp(-j2\pi ac)}{\sqrt{\sum_{a\in a_i}|\exp(-j2\pi ac)|^2 \sum_{a\in a_i}|\exp(-j2\pi ac)|^2}} = \frac{\sum_{a\in a_i} \exp(-j2\pi ac)}{n(a)} \tag{98}$$

The parameters of equation 98 are the following:

- a is the ring in Fourier space
- c is the direction/size of the phase ramp in the Fourier space
- $n(a)$ is the number of pixels in ring a

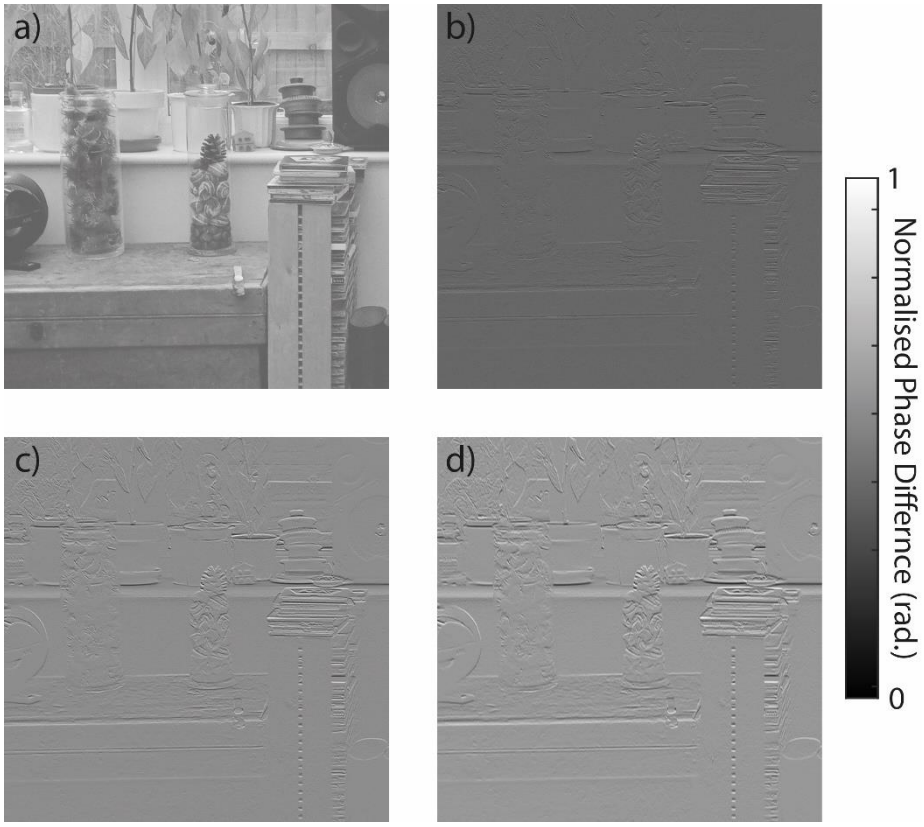

Figure 38: a) the true phase of the specimen used to investigate the effect of pixel shifts in the spatial domain on the calculated Fourier Ring Correlation coefficient. b) the phase difference between a) and itself but shifted downwards by one pixel. c) the phase difference between a) and itself but shifted downwards by three pixels. d) the phase difference between a) and itself but shifted downwards by five pixels. The colour bar to the right converts the greyscale of b-d) into normalised phase difference. This was achieved by dividing b) - d) by the largest phase difference present.

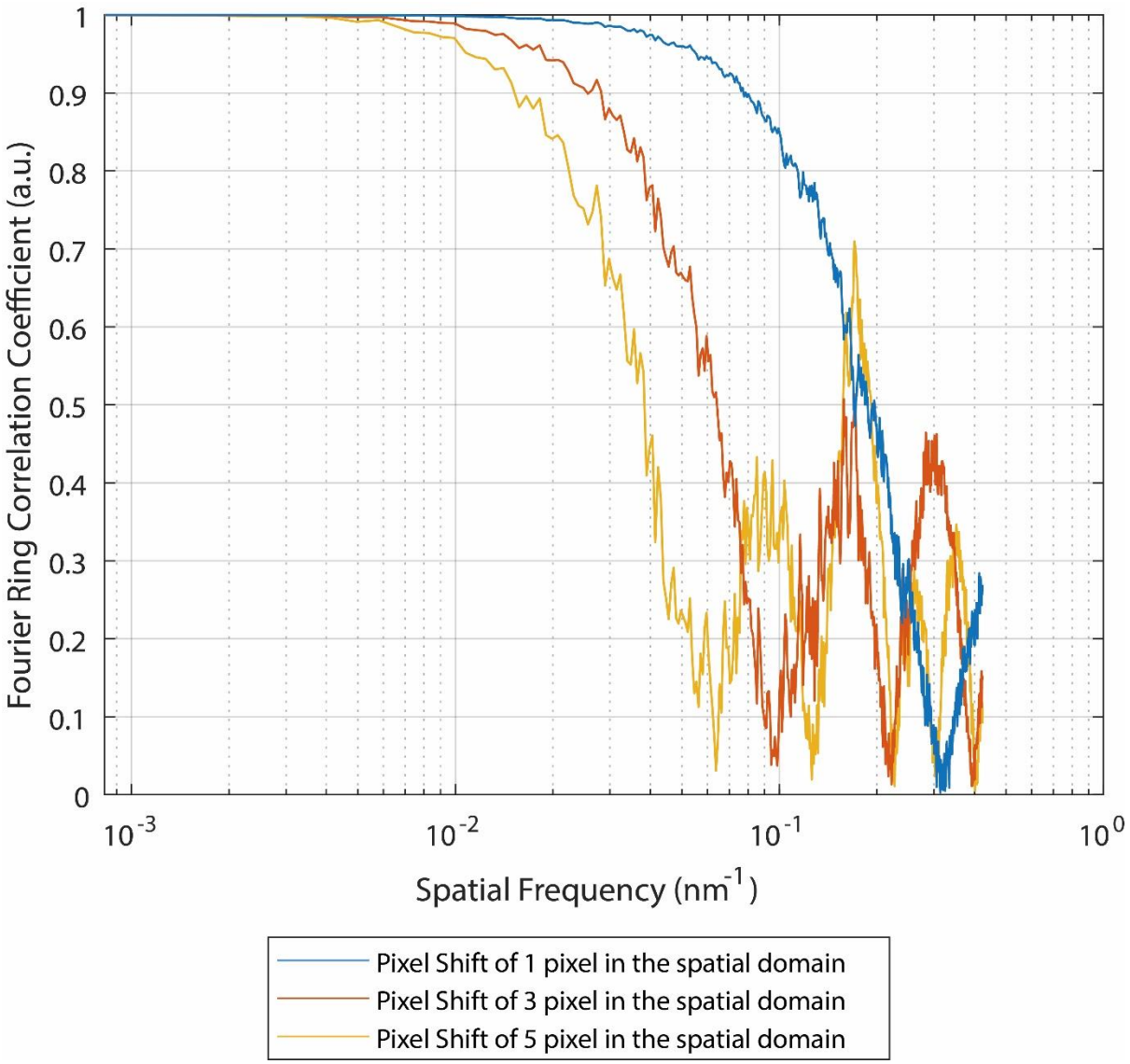

Figure 39: A figure to show the effect of different pixel shifts in the spatial domain on the outcome of an FRC of two otherwise identical images.

Figure 39 suggests that we can predict the effect of a phase ramp/misalignment when the signal to noise ratio is high and the effect of a phase ramp is a Fourier Ring Coefficient that oscillates as a function of frequency. Figure 38 shows the true phase of the specimen used in the FRC calculation shown in Figure 39, as well as the difference between the ideal specimen and the misaligned specimens. As the spatial misalignments between the ideal specimen and misaligned specimen are so small that they would be indistinguishable in Figure 38, their phase difference is shown instead. The phase difference is calculated by subtracting the misaligned phase from the true phase. The effect of a phase ramp on the Fourier Ring Correlation between two images does not have a strong dependency on the direction of the phase ramp, only its magnitude. It should be stated that the effects of the above ambiguities on the FRC may have object dependencies that are currently unexplored. It is important to note that the effect of misalignment on the FRC does not cause a shift in all values of FRC, but only those of a high frequency; the low frequency components are in almost total agreement.

In traditional ptychography, the method of assessing image quality is normally the error function defined in the constraints section above (equation 18), where lower error equates to a better

reconstruction. However, the FRC has the potential to granulate this number across spatial frequency, which offers a greater amount of information on the improvements made as different experiment setups and algorithms are developed. Furthermore, comparison between ptychography and other phase imaging modalities will require frequency analysis, as holography and through focal series reconstructions often perform poorly at certain spatial frequencies [51].

1.40 The half bit threshold resolution measure

The Fourier Ring Correlation is the method by which to measure the underlying signal to noise ratio of the dataset used to create two independent reconstructions. The half bit threshold is a predefined signal to noise ratio that the FRC coefficient must be greater than in order for the reconstructions to be valid at that particular spatial frequency. Therefore, the first crossing of the half bit threshold is often taken as the resolution of the tested image. The half bit threshold is unique compared to other resolution tests, in that it takes account of the possible correlation between the signal of the image and the noise of the image [118]. Any form of threshold test can be fabricated with the equation below:

$$\text{Threshold}(n) = \frac{SNR + \frac{2\sqrt{SNR}+1}{\sqrt{n}}}{SNR + 1 + \frac{2\sqrt{SNR}}{\sqrt{n}}} \tag{99}$$

The parameters of equation 99 are the following:

- SNR – is the signal to noise ratio of half the dataset
- n is the number of pixels considered in shell/ring which covers one spatial frequency.

The signal to noise ratio of a half dataset must be 0.2071 in order to reach the half bit threshold [118]. However, the half bit threshold is a method in which the number of sampling points is taken into account before determining whether the image is valid or not, as shown by the inclusion of n in equation 99. This is required because as mentioned previously in subsection 1.38, the FRC has a higher probability of high correlation at low frequencies even in the presence of noise due to the lower number of pixels at these frequencies, so the low frequencies have to pass a more stringent test than the high spatial frequencies.

1.41 Introduction summary

Now that the Fourier Ring Correlation has been discussed, the literature review of this thesis is complete. In the first part of this literature review the basics of the Transmission Electron Microscope were detailed, followed by a discussion of various TEM and STEM methods used to extract useful information on specimens, primarily focusing on phase imaging. This then transitioned into a discussion on ptychography, its history, and how the phase problem is constrained by collecting multiple diffraction patterns from overlapping areas. After this, near field ptychography was described, which is the focus of the thesis. The differences between near field and focused probe ptychography were highlighted, particularly the greater field of view per diffraction pattern which near field ptychography offers. Near field ptychography also has the advantage over off-axis holography in that it does not require a reference wave to obtain quantitative phase images. Furthermore, near field ptychography is consistent over a range of spatial frequencies using a single defocus condition, while inline holography requires multiple diffraction conditions to achieve such consistency.

The second half of the literature review focused on outlining the appropriate tools required by near field electron ptychography, the basics of the reconstruction algorithm, propagation methods, and an understanding of incoherence and Fourier Ring Correlation. Furthermore, for the first time it was demonstrated how the various ambiguities of ptychography can affect the FRC coefficient, highlighting the importance of removing the ambiguities before performing the FRC. Some tools developed in the second half of the literature review will now be used in the following section, where a direct comparison between the phase sensitivity of near field ptychography and focused probe ptychography is attempted. Phase sensitivity is a particularly important parameter in the electron domain. This interest is highlighted by several papers comparing the phase sensitivity of off-axis holography and DPC [120, 121], so the next section can aid in near field electron ptychography's development in providing an approximation of its phase sensitivity.

Chapter 2

2 Optical Experiments

2.1 Introduction

Now that the necessary literature has been reviewed it is important to demonstrate near field ptychography in a practical context. But before near field ptychography is implemented on TEM it is important to show that near field optical ptychography can be implemented competently. Furthermore, by investigating near field ptychography and SAP in an optical context, practical experience can be gained for what is required in the TEM context. This section will also include a comparison of the phase sensitivity of near field ptychography and SAP.

Phase sensitivity defines the minimum phase step that can be differentiated by a particular technique, off-axis holography using multiple exposures offers a phase sensitivity of 2π/1000 or 0.0063 radians. Techniques with high phase sensitivities are important as they enable differentiation of semiconductors doped to different levels [122] and enable high precision measurement of the strain between different materials [123]. The doping and strain measurements are important to the semiconductor industry as they inform designers whether their intentions/designs have been correctly implemented in the final product, thereby allowing them to iterate on their design. Therefore, in this section phase sensitivity of near field optical ptychography will be measured and compared to that of focused probe optical ptychography, via the investigation of a standard test specimen described in [124]. Investigating optical near field ptychography's phase sensitivity will give an approximation of the phase sensitivity of near field electron ptychography and comparing the results to focused probe optical ptychography may reveal any fundamental causes for their different phase sensitivities. Not only will the phase sensitivity of near field optical ptychography be assessed, but also that of selected area optical ptychography, as this will suggest whether the implementation of a diffuser increases the phase sensitivity of near field ptychography. Selected area ptychography has been implemented on the transmission electron microscope and has demonstrated to have a phase sensitivity of at least $2\pi/100$ [94], therefore if optical near field ptychography shows a greater phase sensitivity than its selected area counterpart, it suggests that near field electron ptychography should have a greater phase sensitivity than 0.063 radians. In this section the FRC analysis tools (see Section 1.36) will be used to assess the consistency of the frequency response of near field optical and optical selected area ptychography. This will be achieved by splitting an exemplar dataset into two smaller datasets which are then reconstructed independently by ptychographic algorithms and compared via the FRC method. This will be an interesting experiment, as it will suggest whether the diffuser has a role in encoding any particular spatial frequency into the dataset or simply improves the signal to noise ratio of the dataset as a whole.

It will also be shown that SAP can be performed with a larger aperture than used in its first implementation [94]. Apart from being an analogue for near field electron ptychography, near field optical ptychography has potential of its own. For example, optical near field ptychography has been shown to be quantitative and able to image a large range of specimens, including thick specimens [125]. This makes near field optical ptychography interesting for biological applications. This section begins with the experimental setup used throughout these optical experiments.

2.2 Experimental setup of near field ptychography

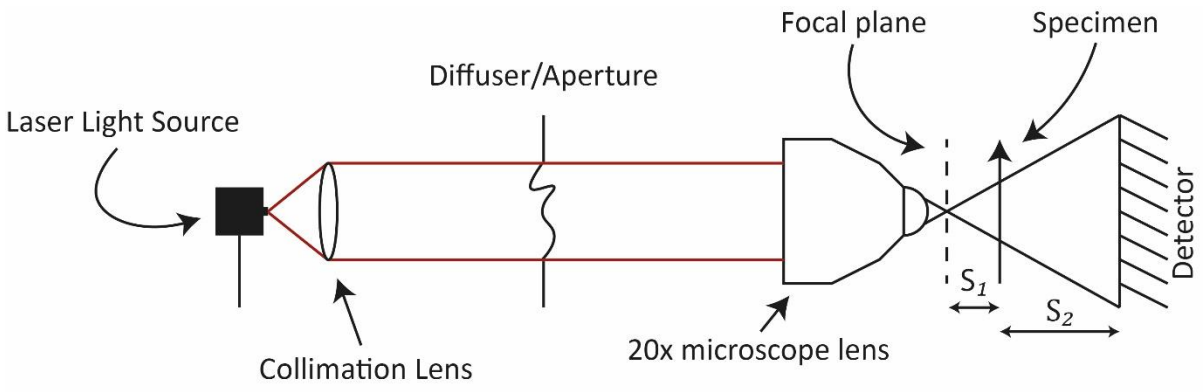

Figure 40: A simplified digram of the setup used to perform optical near field ptychography

The setup used throughout this section is illustrated in Figure 40. The setup will now be described from left to right. The laser light source is a Thorlabs MCLS1, and the 635nm wavelength output was used. The laser is collimated by a Thorlabs AC254-075-A-ML convex lens. This lens is placed one focal length (75mm) away from the light source. The laser was attenuated, so as not to saturate the detector, by adding the following absorptive filters: NE05A, NE04A and NE02A (all Thorlabs). A diffuser was placed roughly 100mm away from the entrance of the 20x lens. The 20x lens in this setup is an Olympus Plan N 20x, which has a numerical aperture of 0.40. The 20x microscope lens is used to set up a cone beam geometry imaging system [126]. The detector used was a PCO.edge sCMOS detector, which has a pixel size of 6.5µm but was used with a binning value of 2. It should be noted that the diffuser is not in focus in the specimen plane; this was done to demonstrate that the diffuser and specimen do not have to share the same in focus plane in order to reconstruct the specimen successfully. None of the above components are vital to the implementation of near field optical ptychography due to its robust nature; a stripped down version would simply consist of a coherent light source, a diffuser to disrupt the diffraction fringes of the specimen, a lens to set up the cone beam geometry (maybe the divergent illumination from the light source might be sufficient), a detector to collect defocused images of the specimen and diffuser, and a ruler to measure the distances.

Typically, in published works on near field ptychography the diffuser and specimen share the same image plane, but ptychography has the capability to recover the illumination function alongside the specimen function, which means the diffuser can be implemented in an arbitrary position along the optical axis although this is typically avoided for sake of simplicity. This has some interesting implications in terms of the phase transfer function of the collected data, as it means that the diffuser and specimen can have different transfer functions, suggesting that spatial frequencies could be captured depending on whether one is looking at the diffuser or the specimen. This has interesting implications in terms of diversity as the diffraction fringes can be quite strong for either the specimen or the diffuser. However, the important point being made here is to demonstrate that near field ptychography can be implemented with its diffuser not sharing the same image plane as the specimen and will still function adequately. This is important for future work in the electron domain where the diffuser will be implemented in the condenser lens aperture, as in such an implementation it may be difficult to ensure the alignment of the specimens and diffusers image planes along the optical axis, hence the results here suggest that this will not be an issue going forward.

The cone beam geometry has been used in near field ptychography before, by Stockmar et al. [79] and [127] in which the authors moved the specimen along the *z* axis in addition to the typical *x*-*y* movements, to increase the diversity of the dataset. The cone beam geometry relies on the divergence of the illumination to provide magnification. The 20x microscope lens focal point acts as a point source for the cone beam geometry, and the distance between the focal point of the 20x lens and the detector determine practical limits of the propagation distance and magnification. The magnification of the cone beam geometry is determined by the following equation:

$$M = \frac{S_1 + S_2}{S_1} \tag{100}$$

In theory, the propagation between specimen plane and the detector plane could be modelled by propagating divergent waves (phase curvature), but the simpler approach taken by [127], [126] is also used here, whereby the divergence propagation is converted to an equivalent plane wave propagation (Fresnel scaling theorem). This conversion is performed by demagnifying the detector and simultaneously the distance between the specimen and the detector.

$$Z_{\text{eff}} = \frac{S_2}{M} = \frac{S_1 S_2}{S_1 + S_2} \tag{101}$$

$$PixSize_{\text{eff}} = \frac{PixSize_{\text{True}}}{M} \tag{102}$$

The parameters of equations 100-102 are the following:

- S_1 is the distance between the focal plane of the objective lens and the specimen
- S_2 is the distance between the specimen and the detector
- Z_{eff} is the effective propagation distance between the specimen and the detector when the diffraction condition has been converted into the plane wave geometry
- M is the magnification provided by the cone beam geometry
- $PixSize_{\text{eff}}$ is the effective pixel size once the diffraction condition has been converted into the plane wave geometry
- $PixSize_{\text{True}}$ is the physical pixel size of the detector.

2.3 Optical near field ptychography

In the experiment described below, near field ptychography was used to image a phase resolution chart which can simultaneously give the pixel size, resolution and test the quantitative nature of the phase obtained. Furthermore, the collected data was designed such that it could be split into two separate datasets whose independent reconstructions could be compared via Fourier Ring Correlation to assess performance as a function of spatial frequency. The data collected in this experiment consisted of a 30 by 25 raster scan (*y* by *x*) with a step size of 120μm plus ±20% random offset to avoid the raster grid pathology [128]. The physical propagation distance between the detector and the sample was 5.3cm (the camera length will be discussed at the end of this section). The exposure time of the PCO CMOS detector was 500μs, 8 exposures were taken for each probe position and were then averaged to reduce noise. The detector was binned down to 1024 by 1024 pixels from 2048 by 2048.

The specimen used in this experiment is a phase resolution chart (Phase Focus Ltd), which was used previously to demonstrate the phase accuracy of traditional optical ptychography [124]. The widths and lengths of the elements have a documented value, which can be used to determine the pixel size of the image (note this is not the resolution). For this experiment, group three, element three

has been used to measure the pixel size. The elements used have a width of 55μm and a length of 275μm. Looking at Figure 41 and Table 1, the average length of an element in pixels was 207.31pixels. Therefore, the estimated pixel size is 1.33μm (275/ 207.31). Figure 41 shows that the entire specimen has been imaged, and the field of view has been calculated to be equal to 4mm by 4mm (this is a 9 Megapixel image, 3000 by 3000 pixels). The measured magnification is 9.77 times ($13 \times 10^{-6}/1.33 \times 10^{-6}$).

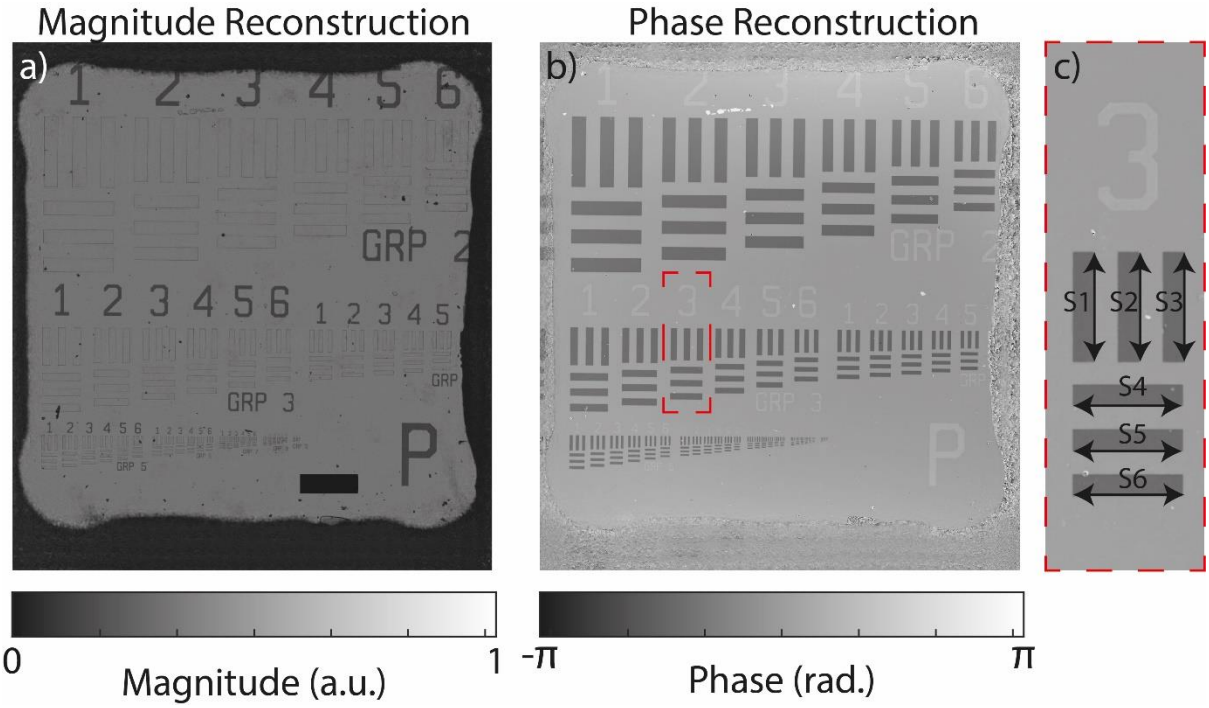

Figure 41: a figure showing an optical near field ptychography reconstruction, where a) is the magnitude of the reconstruction and b) is the phase. c) shows the subsection used to measure the pixel size. The scale bar in a) indicates 500 micrometres.

Element label	Measured length in pixels
S1	205.69
S2	206.77
S3	206.77
S4	207.85
S5	207.85
S6	208.93

Table 1: A table showing the measured feature lengths from Figure 41.

Now the pixel size has been established, the Fourier Ring Correlation can be performed (see Section 1.36). First, the dataset used to produce Figure 41 is split into odd and even datasets; the positions of the odd and even datasets can be seen in Figure 42.

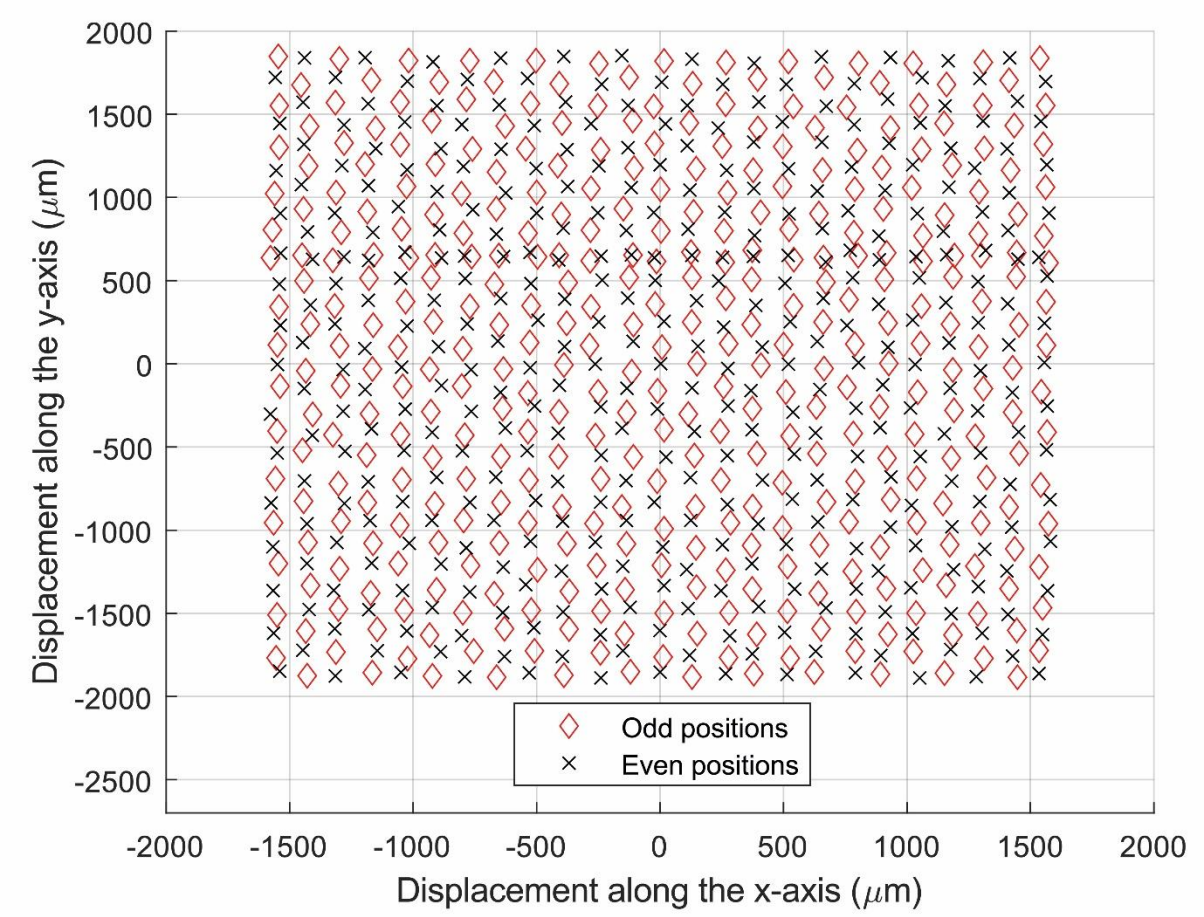

Figure 42: A figure showing the odd and even positions which correspond to Figure 43. The even positions are indicated by diamonds and the odd positions are indicated by crosses. All positions were used to obtain the reconstructions shown in Figure 41. The reason for splitting the dataset into odd and even diffraction patterns is outlined in section 1.36.

An additional step taken was to initialise the reconstruction with a random and slowly varying phase. This ensured that the recovery of low spatial frequencies was appropriately tested, since if the reconstruction had been initialised with a flat phase, it may have been difficult to tell when the low frequencies had been recovered, as the underlying phase of the specimen might have been flat. Whilst initialising with a random phase, the change in low frequencies can be explicitly seen. Figure 43 shows the initialised phase and reconstructed phase for both the odd and even datasets. The change from Figure 43 a) to c) shows that during the reconstruction process the initial phase has been completely removed. This change suggests that near field ptychography can recover low spatial frequency information.

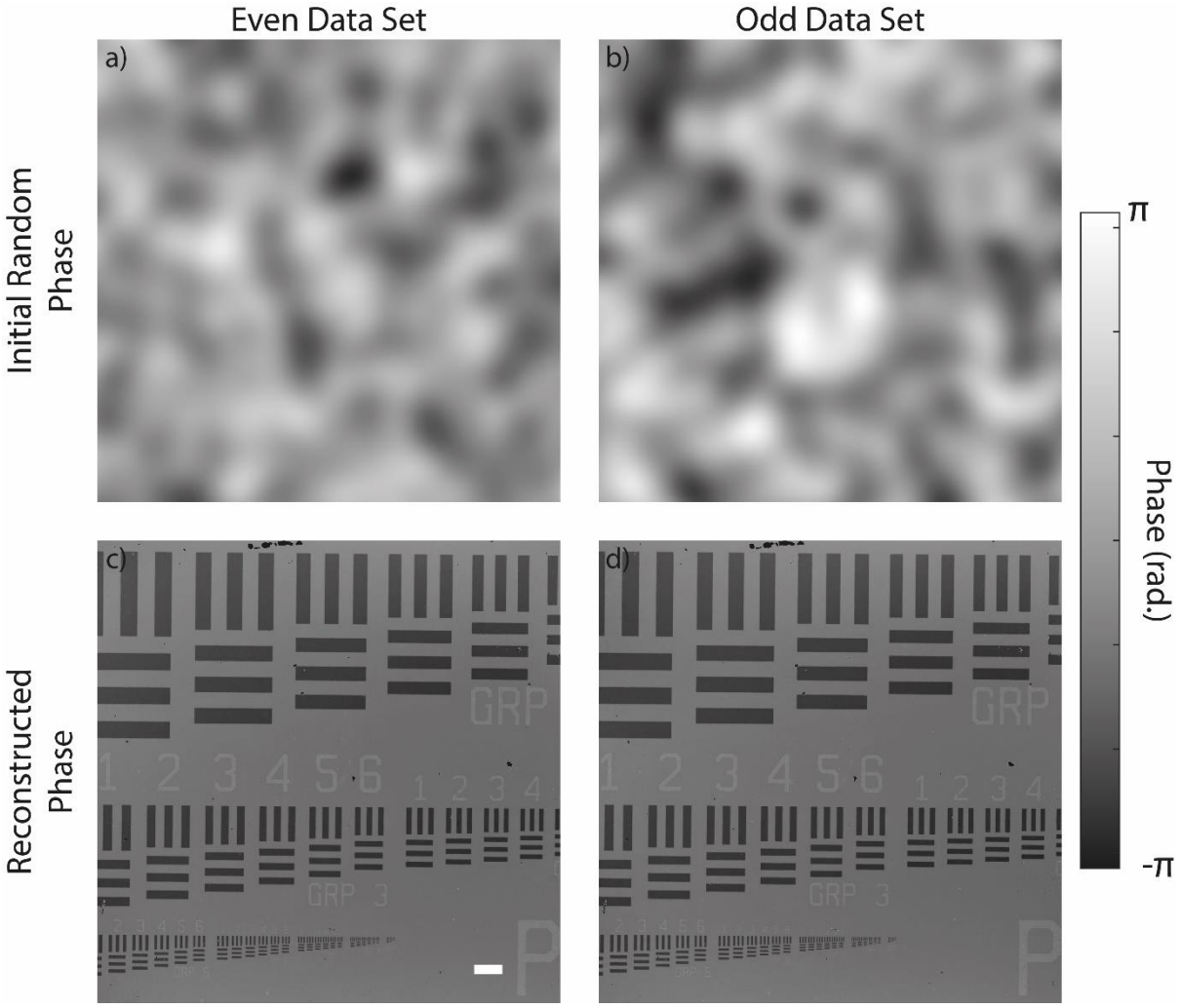

Figure 43: a) and b) show the initial phase the rPIE algorithm was provided with as input for two separate reconstructions. c) shows the reconstructed phase obtained when using only the data indicated by diamonds in Figure 42 (the initial phase of this reconstruction was a)). d) shows the reconstructed phase when using only the data indicated by crosses in Figure 42 (the initial phase of this reconstruction was b)). The white scale bar in c) indicates 200µm. The colour bar at the side of the figure converts the greyscale of all images in radians.

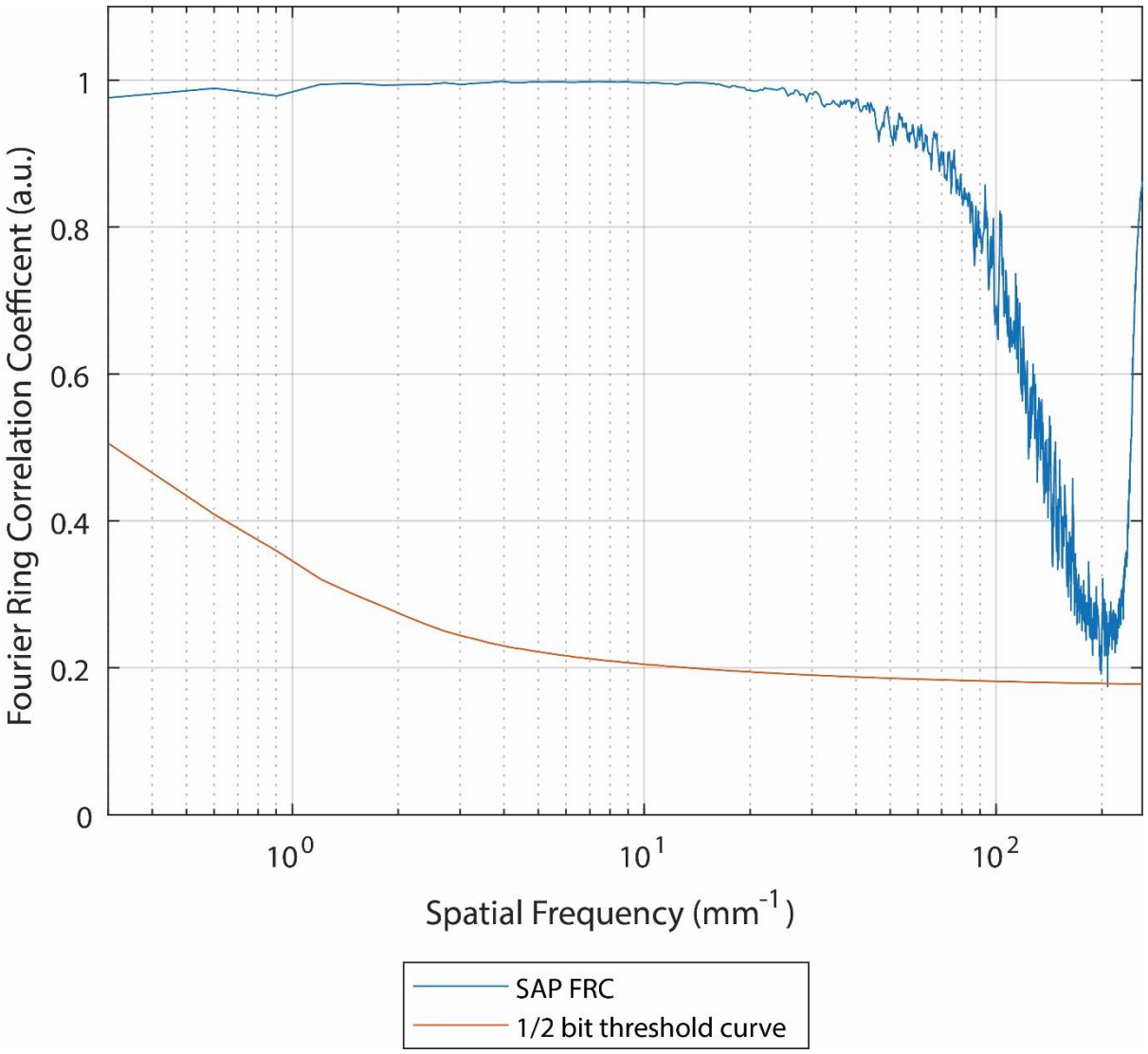

Figure 44: The Fourier Ring Correlation result produced by comparing the reconstructions shown in Figure 43 c) and Figure 43 d). The half bit threshold is also plotted for reference and to enable resolution measurement.

The ability of near field ptychography to recover low spatial frequencies is confirmed by the FRC of the two independent reconstructions, which show an almost perfect correlation at low spatial frequencies (see Figure 44). However, as the spatial frequencies increase the FRC coefficient decreases, suggesting that the signal to noise ratio of the underlining data is decreasing. The increase in the FRC coefficient after $2 \times 10^2 \mathrm{mm}^{-1}$ is the result of aliasing of the propagation function; the propagation function can be seen in Figure 45. The propagation function seen in Figure 45 is limited by a mask at high frequencies to prevent incorrect propagation of those frequencies, the mask is calculated using equation 103 (further details can be found in [129]). Due to the implementation of the mask, it is impossible to have a resolution greater than that allowed by the mask, hence the lowest point of the curve is considered to be the resolution of this reconstruction, and therefore this reconstruction has a resolution of 4.8µm, even though it barely crosses the half bit threshold. The aliasing of the propagation function was due to the fact that the minimum camera length was limited by the casing of the detector (effectively S2's minimum value in equation 101 was limited by the physical dimensions of the detector). This frequency response suggests that near field ptychography is capable of recovering all spatial frequencies which have propagated correctly. However, the nature of this FRC analysis should be considered, as the near field ptychography result

was compared to another near field ptychography reconstruction, and therefore the calculated FRC result measures the consistency of the reconstructions, not the phase transfer function of near field ptychography. This will be expanded upon in Section 5.2.3.

The field of view considered in this FRC calculation was 2500 pixels by 2500 pixels, which is 2.5 times the size of the reconstructed probe. Only a FOV which is significantly larger than the size of the probe is able to fully test the behaviour of near field ptychography at extremely low spatial frequencies; more on this in section 4.12.

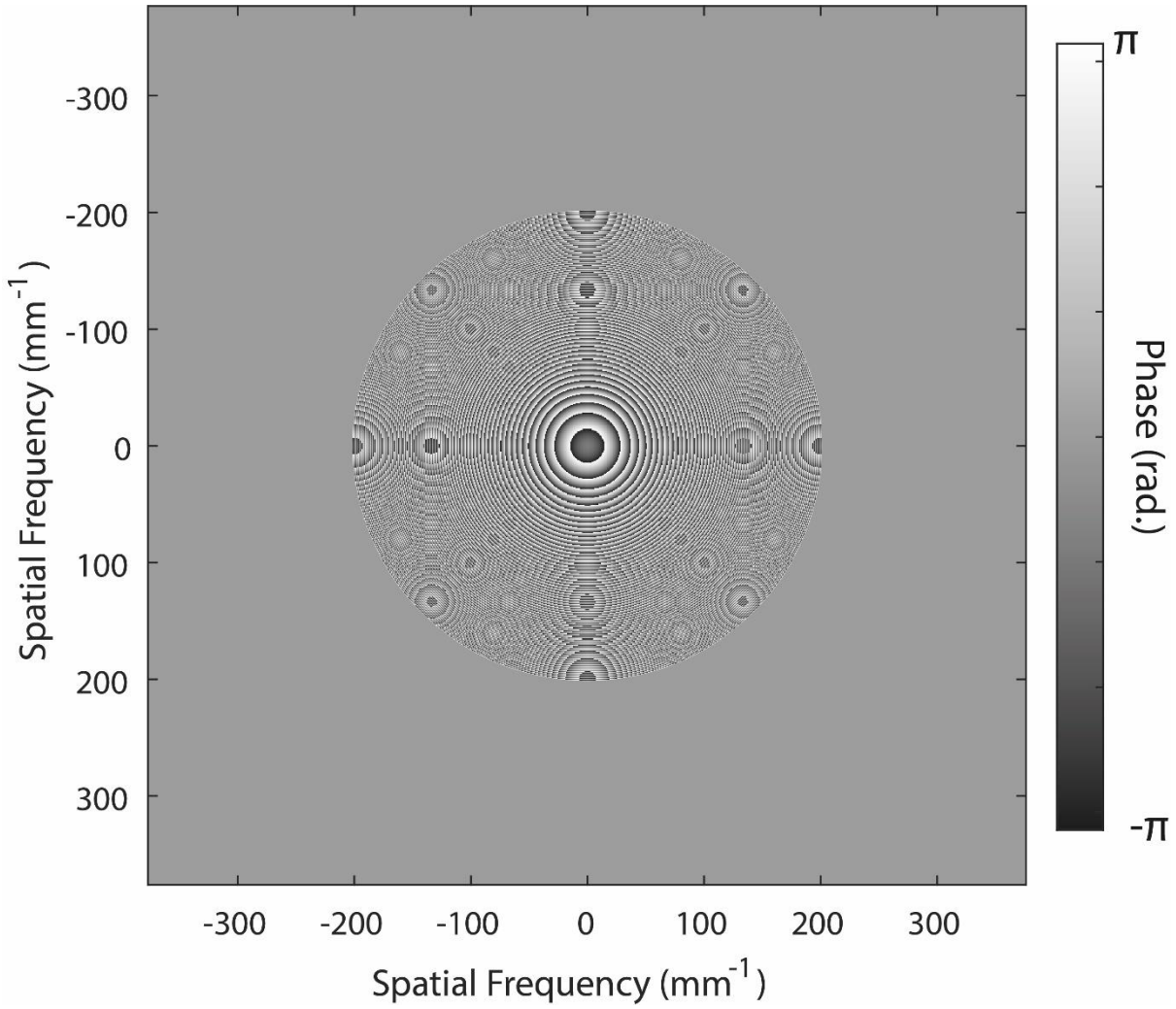

Figure 45: the propagation kernel used in the reconstruction of Figure 41, Figure 43 and Figure 48. The colour bar at the right of the figure converts the greyscale of image into radians. The flat grey area around the propagator indicates the part of the propagator set to zero in order to avoid the aliasing of the propagation function.

$$\sqrt{{f_x}^2 + {f_y}^2} < \frac{W_D}{2\lambda C_l} \tag{103}$$

Equation 103 determines when a propagator becomes aliased as a function of frequency space [129]. The parameters of equation 103 are the following:

- f_x and f_y is the spatial frequency along the *x* and *y* axes respectively
- W_D is the effective demagnified detector width
- λ is the wavelength of the illumination used
- C_l is the propagation distance between the specimen and the detector [129]

Following on from FRC analysis, the reconstructions in Figure 43 will be used to test whether near field ptychography is quantitative. The phase sample used in Figure 43 is produced by etching a material with a known refractive index to known specific depths, causing the phase contrast seen in Figure 44 c) and d). Equation 104 shows how the phase delay of the sample can be converted into a depth map of the specimen[124]:

$$d = \frac{\lambda\theta}{2\pi\Delta n} \tag{104}$$

The parameters of equation 104 are the following:

- d is the depth
- λ is the wavelength (635nm)
- θ is the phase image of the specimen
- Δn is the difference in the refractive index between the specimen and air (0.4665)

A reconstruction obtained from the whole dataset is used to measure the depth. The image is rotated and then groups of elements are selected from the specimen. In this test, groups 2, 3, and 4 were analysed, matching results in [124]. The measured phase, θ, from the reconstruction is used to calculate the depth via equation 104. The depth images of the groups considered are shown in Figure 47. Thresholding was used to create Figure *46* to remove extreme values outside the designed depth of the specimen (303nm), as these extreme values obscured smaller variances from the human eye. These extreme values were often caused by the presence of dust particles, as the extreme values were found to move between experiments (compare Figure *46* to Figure 53) and dust particles were also found on calibration samples. The thresholding values used in Figure *46* are between 100nm and -400nm. In [124] the authors describe their reconstruction as containing "very low frequency background phase curvature", and they then proceed to remove the background phase curvature via second order polynomial fit. The reconstructions here also experienced a low frequency phase component, the low frequency component may be due to the poor low frequency response of near field ptychography (see Section 5.2.3). However, there seems to be no other features associated with poor low frequency response similar to the original paper [124], for example the cloudy phase which can be seen in subsection 4.12 or ringing at hard edges [130]. Furthermore, the propagation function used in this experiment (see Figure 45) would suggest a reasonable frequency response for the lower spatial frequencies in the case of inline holography, in turn suggesting the low frequency phase feature is a real component of the specimen. This low frequency component is dealt with in this thesis by the removal of the phase ramps over the sample subsections used to calculate mean etched depth (see Figure *46* and Table 2). The depth values from each group were then plotted in histograms, such that the distribution of the depths could be visualised. The histograms for each group can be seen in Figure 47. The mean and standard deviation of the etched distributions is tabulated in Table 2. The black dotted lines in Figure 47 indicate part of the distribution used to calculate the mean and standard deviation; everything within the dotted lines was used. The red line indicates the calculated mean of the etched surface, and the solid black line indicates the average depth of the unetched surface. The expected value is 303nm; near field ptychography obtained 294-306nm, so in the worst case there is a 3% offset between our measured and expected values. Looking at the histograms, this offset is partly due to wide spread of the depth measurements of the etched and unetched surfaces. Figure 48 shows that the reconstruction of the etched surface is noisy, which may correlate with the FRC result in Figure 44, in which a reduced high frequency response was shown. This result again highlights that the FRC performed with images collected by identical techniques and under identical conditions test consistency more than quality.

If the FRC was performed with a near field ptychography reconstruction, an inline holography reconstruction may give a less positive picture about the agreement of the high and low frequencies than Figure 44 suggests.

In [124] a 100 by 100 micrometre area is used to determine the phase sensitivity of optical ptychography. Due to the lower magnification of this experiment, the largest element of the group 2 elements was used instead. The area used to measure the phase sensitivity was an 80 by 404 pixel array, giving 35200 sampling points. The standard deviation of the depth in this area was calculated to be 6.03nm, and multiplying by 2.355 gives the FWHM (Full Width Half Maximum) of the depth as 14.2nm. The FWHM is then converted back into its equivalent phase and the result is converted into a fraction of 2π, obtaining a phase sensitivity of $\frac{2\pi}{96}$. The depth map of the element used in the phase sensitivity calculation can be seen in Figure 48. Figure 48 and Figure 41 b) show a practical difficulty in this measurement, in that the sample has clearly been affected by the dusty conditions under which has it been imaged, and the dust has affected the results of this section. The blue arrows in Figure 48 indicate possible depth change caused by dust. However, the larger contribution is the reduced high frequency response of the reconstruction which is clear in Figure 48, which shows the reason for the poor high frequency response is the aliased propagation kernel. Comparing these results to those found in [124] there is at least a threefold reduction in the phase sensitivity of near field ptychography compared to traditional ptychography, although further experiments need to be performed to confirm these results. A hypothesis for this reduction in phase sensitivity is that it is due to the aliasing of the propagation function which allows high frequency information to effectively be poorly defined. This effect can be seen in figure 46 where a flat surface is seen to vary by at least 20nm. The aliasing was resolved by setting a large region of the propagator's magnitude to zero, such that aliased features are not propagated, however this had the unintended consequence of allowing these high frequency components to have a floating value in the spatial domain instead of the intended value of zero. Another possible factor is the increased number of sampling points used, for example 23409 (100µm by 100µm area with a pixel size of 653nm) pixels were used in the original work, while 35200 pixels were used here. Another possible reason could be that the depth measurement was taken from the etched area instead of the background. As reported in the original work, the etched area experienced a greater variance in its measured phase value. These factors therefore prevent near field ptychography operating at its full potential in terms of phase sensitivity. The author believes that further experiments are required to confirm that near field ptychography has an inherently lower phase sensitivity than focused probe ptychography, as these results are from a suboptimal experimental setup.

Another hypothesis for the difference in phase sensitivity assumes that near field optical ptychography and optical ptychography have similar properties to off-axis holography, in that their phase sensitivity is proportional to the number of counts, the visibility of speckles (or fringes in the case of off-axis holography) and the coherence of the beam [131]. Then, the difference in phase sensitivity may be due to one of these parameters not being fully optimised in this experiment. For example, ptychography's greater scan density may be advantageous in terms of phase sensitivity, as there was a greater number of photons per pixel in the final reconstructed image compared to the near field experiment presented here. Therefore, in this implementation of near field optical ptychography the phase sensitivity could be increased by increasing the number of multiple exposures for a given scanning position and combining the diffraction patterns post-acquisition in order to boost the signal to noise ratio of the dataset similar to off-axis holography [131]. Unfortunately, Godden et al. do not explicitly give exposure time or average counts per pixel [124]. The average counts per pixels for a near field optical ptychography diffraction pattern was 3479

counts per pixel and 4760 counts per pixel for SAP diffraction pattern (inside the bright disc). Interesting further work would involve confirming this hypothesis as to whether near field ptychography has a similar equation relating the phase sensitivity to the flux per pixel and the visibility/contrast of the diffraction pattern.

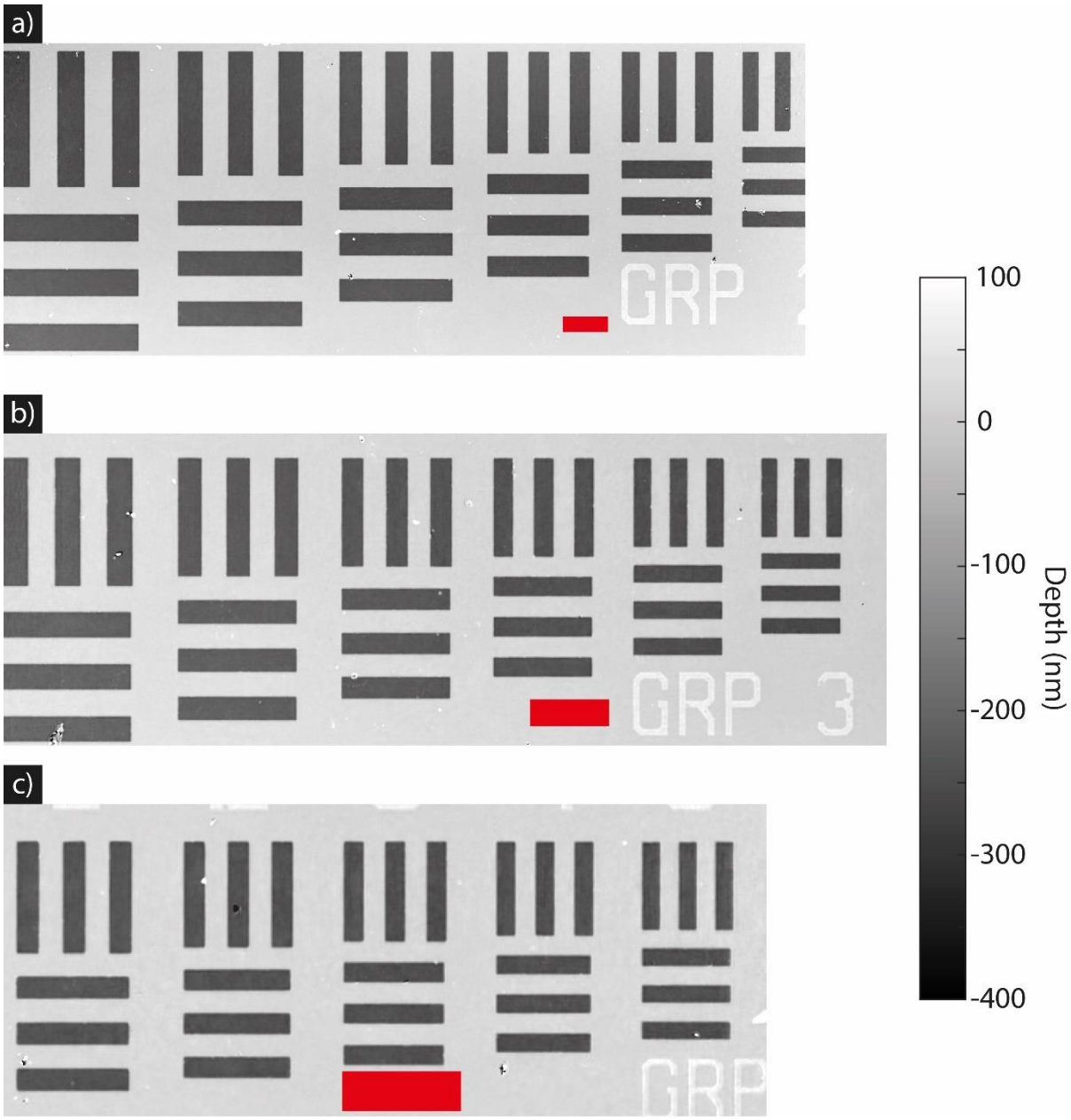

Figure 46: a) depth map of the group 2 elements. b) depth maps of the group 3 elements. c) depth maps of group 4 elements. Note that this figure deploys contrast enhancement; the maximum depth of this figure is artificially limited to 100nm, and the minimum depth is artificially limited to -400nm. The red scale bar indicates 200µm. The colour bar to the right of the figure converts the greyscale of the images to depth in terms of nm.

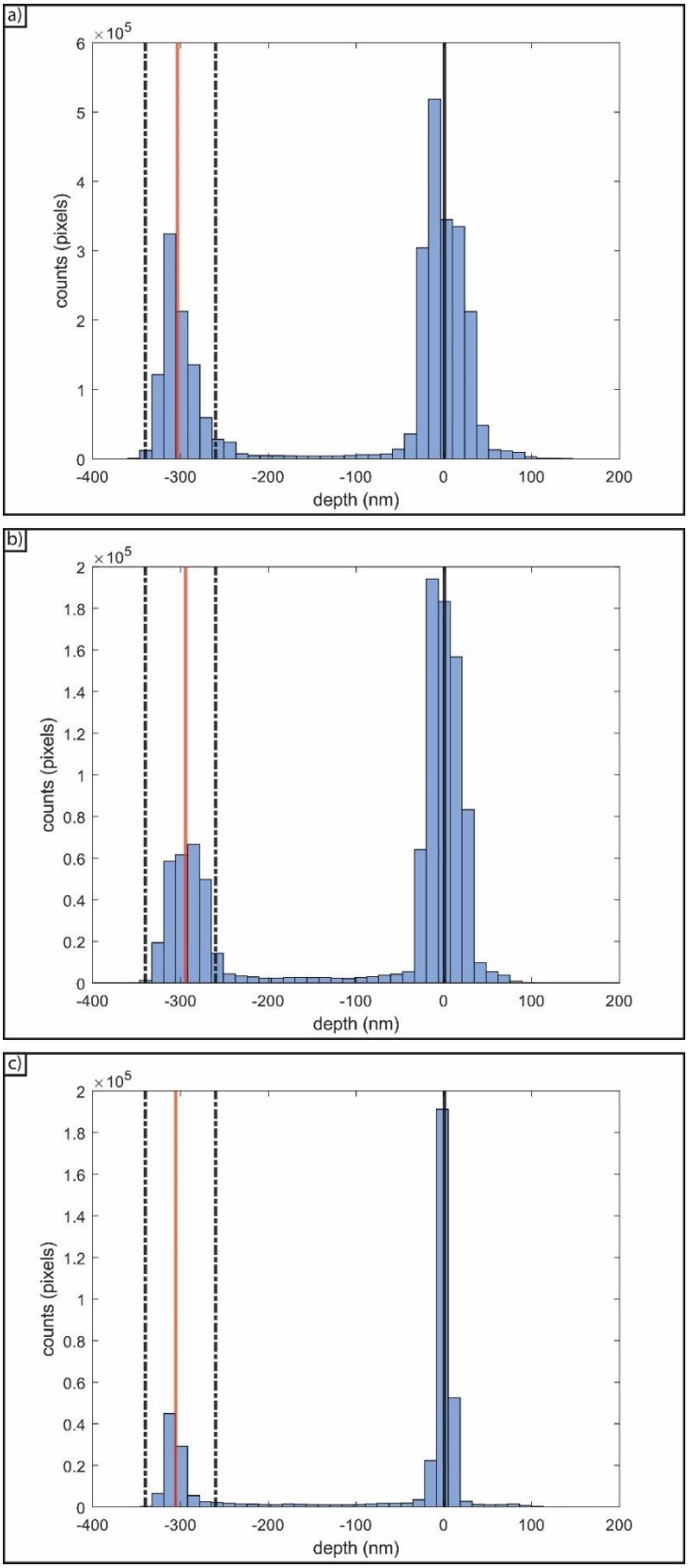

Figure 47: a) histogram of the Group 2 depth values which can be seen in Figure 46 a), b) a histogram of the Group 3 depth values which can be seen in Figure 46 b). c) a histogram of the Group 4 depth values which can be seen in Figure 46 c). The red lines in a)-c) indicate the average depth of the etched surface which was calculated using the data within the bounds indicated by dotted black lines. The solid black lines indicate the average of the unetched surface.

Group number	Mean etched depth (nm)	Standard deviation of the etched depth (nm)
2	303.3	15.7
3	294.2	17.2
4	306.6	11.6

Table 2: a table showing how the measured mean and standard deviation of the etched depth from a near field ptychographic image varies with group number. These values were calculated using the values within the black dotted lines seen in Figure 47.

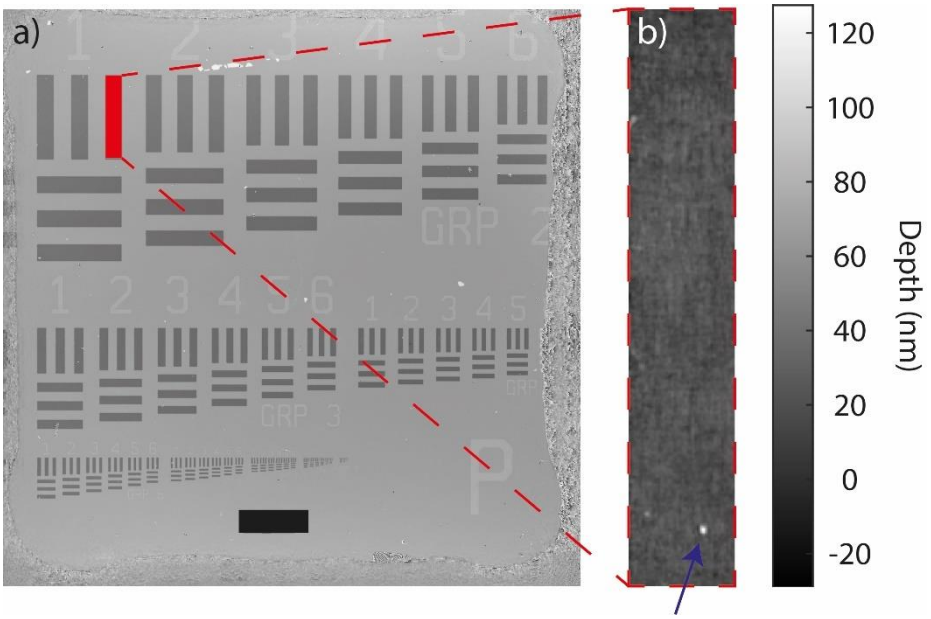

Figure 48: a) the total reconstructed phase of the specimen, where a red rectangle is used to indicate the subsection used in the depth map shown in b). b) the depth map subsection which was used to calculate the phase sensitivity of near field optical ptychography. The blue arrow in b) indicates a depth change possibly caused by dust. Note that the colour-bar only applies to b) and converts the greyscale of b) into depth in terms of nm. The scale bar in a) is 500µm.

As an aside, it was mentioned in Section 2.2 that the diffuser and specimen did not share the same in-focus plane, this was primarily done to show that it had no adverse effect on the reconstruction. The ability to have the diffuser out of focus when it interacts with the specimen may make the implementation of the diffuser in the condenser lens aperture in a TEM easier. If the diffuser was implemented in the condenser aperture of the TEM, depending on the condenser lens setup, the diffuser may or may not share the same in-focus plane (see Section 5.2.2). The results of this section suggest that in the case of a thin specimen, the diffuser does not need to share the same in-focus plane. The reconstructed diffuser obtained from the dataset discussed in this section can be found in Figure 49. Figure 49 a-c) show the probe at the specimen plane, which is then back propagated until the Sellotape which is used as a diffuser is in focus, as confirmed by the specks of dust/glue which are seen Figure 49 b).

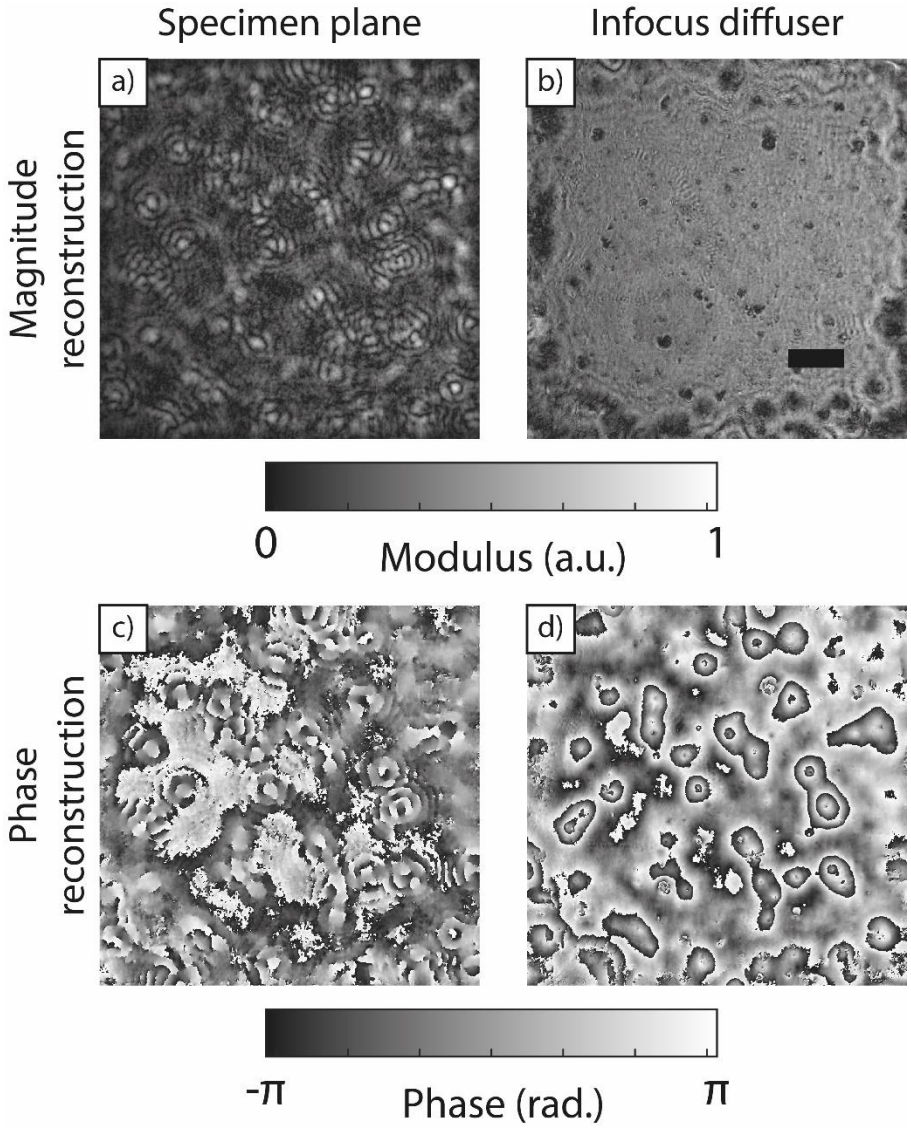

Figure 49: a) the reconstructed magnitude of the probe at the specimen plane. b) the magnitude of the in-focus diffuser used to create the probe seen in a). c) the reconstructed phase of the probe at the specimen plane. d) the phase of the in-focus diffuser used to create the probe seen in c). The scale bar in b) indicates 200μm. The colorbar under a) and b) converts the greyscale into an arbitrary scaled modulus. The colour bar under c) and d) converts the greyscale of c and d) in phase in terms of radians. Note that this probe was obtained from the same reconstruction as Figure 41.

2.4 Optical selected area ptychography

SAP was the precursor to near field electron ptychography on the TEM (see Section 1.22). The purpose of this experiment is to demonstrate that SAP can be performed with a larger aperture than used in its initial publication [94]. This experiment is designed to investigate whether near field ptychography with an aperture acting as a diffuser may be a valid option for an electron microscopist, who wants to perform near field electron ptychography but does not want to have to alter their microscope, such that it holds custom made apertures/diffusers. The following experiment is to confirm the simulation results through an optical experiment with a quantitative sample.

The dataset consisted of a 35 by 40 (*x* by *y*) rectangular scan with a step size of 120μm. The specimen is the PFPT01 - Phase Calibration Target (Phase Focus Ltd) which was used in the optical implementation of near field ptychography; see section 2.3 for further details on the sample. The laser light source used had a wavelength of 635nm. The exposure time of the PCO CMOS detector

was 1000µs, 8 exposures were taken per probe position and then averaged to reduce noise. The PCO CMOS detector has a pixel size of 6.5µm, the detector was binned such that the effective pixel size was 13µm (binning 2). The magnification of the cone beam geometry setup was 10.4 times, meaning the effective pixel size at the specimen plane was 1.3µm. The effective propagation distance between the specimen and detector was 3.93mm. Further details about the experimental setup can be found in section 2.2. To test the frequency response of this dataset, the reconstruction was initialised with slowly varying random phase. The dataset was split into two independent datasets for the sake of FRC analysis as explained in section 1.36 (even and odd).

The similarity between the final reconstructions of the even and odd datasets, and the previous reconstructions of this specimen, suggest that these datasets have been successfully reconstructed. Furthermore, comparing the initial phase to their final phases suggests that the low frequency components of the specimens have been sufficiently encoded in the underlining datasets.

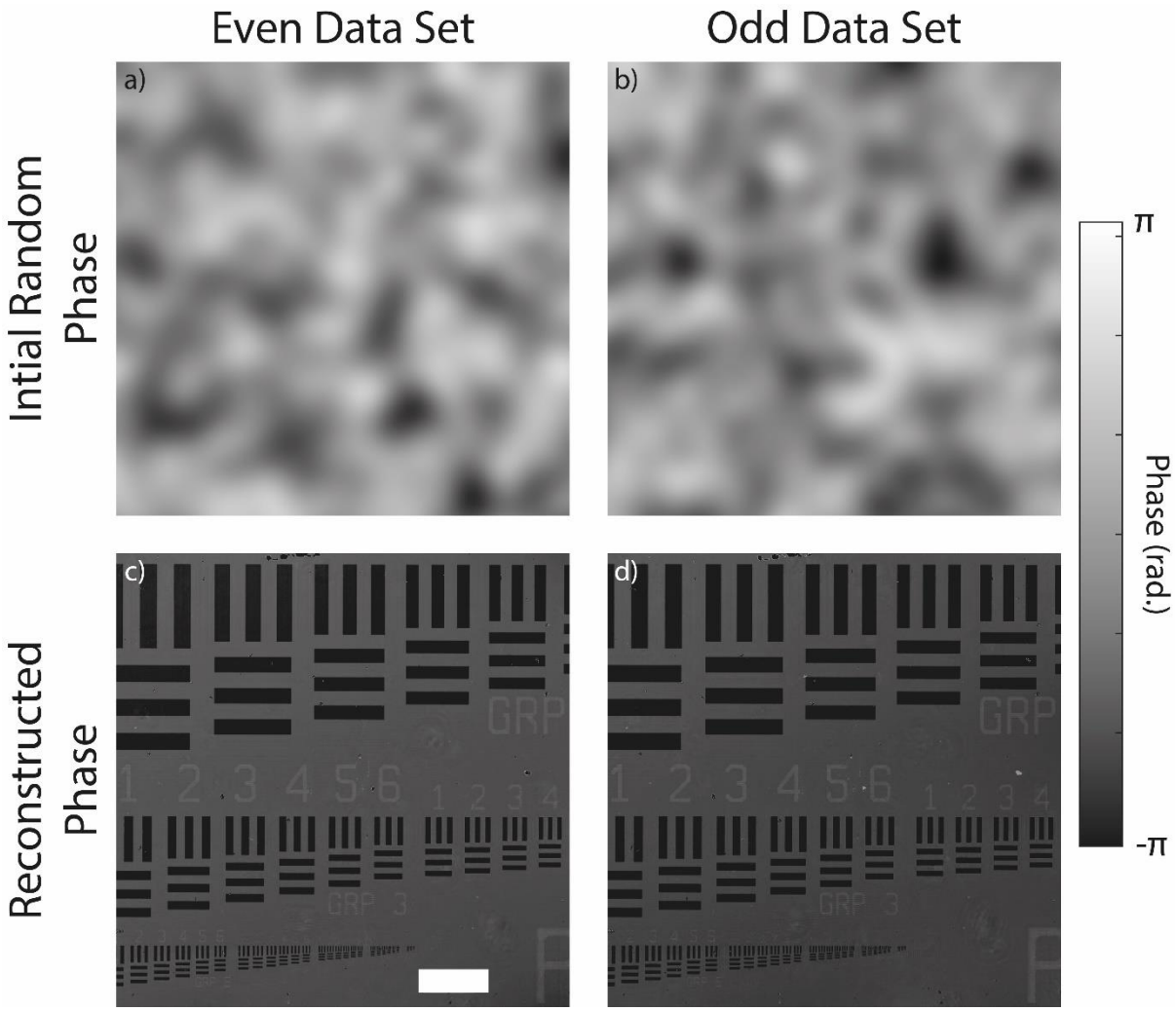

Figure 50: a) and b) show the initial phase the rPIE algorithm was provided with as an input for two separate reconstructions. c) shows the reconstructed phase of the even SAP dataset. d) shows the reconstructed phase of the odd SAP dataset. The white scale bar in c) indicates 500µm. The colour bar at the side of the figure converts the greyscale of all images in radians.

The two independent reconstructions from Figure 50 c) and d) were used to perform an FRC, the result of which is shown in Figure 51. As the FRC curve never crosses the ½ bit threshold, the resolution of this dataset has a resolution of 2.6µm, which is twice the effective pixel pitch. However, the resolution of this SAP dataset is again limited by the aliasing of the propagation kernel,

which is slightly better than in section 2.3 due to the shorter camera length. The propagation limited resolution is found to be 3.9μm, 1.3μm larger than the resolution allowed by the pixel size. Figure 52 shows the propagation function used during the reconstruction of Figure 50, and confirms that resolution of the SAP reconstruction is limited by the aliasing of the propagation function, as after around $250 mm^{-1}$ the phase of Figure 52 is cut off by the aliasing prevention mask described in equation 103.

Additionally, the FRC curve shown in Figure 51 shows a clear dip at low frequencies, although looking at Figure 50 c) and d) there seems to be no clear difference between the two. The subsection used to perform the FRC analysis was a 2500 x 2500 pixel area ($10.6 mm^2$). The focused probe had a diameter of 820 pixels, which should be sufficient to test near field ptychography's low frequency response, as the area imaged is approximately 9 times the area of the probe.

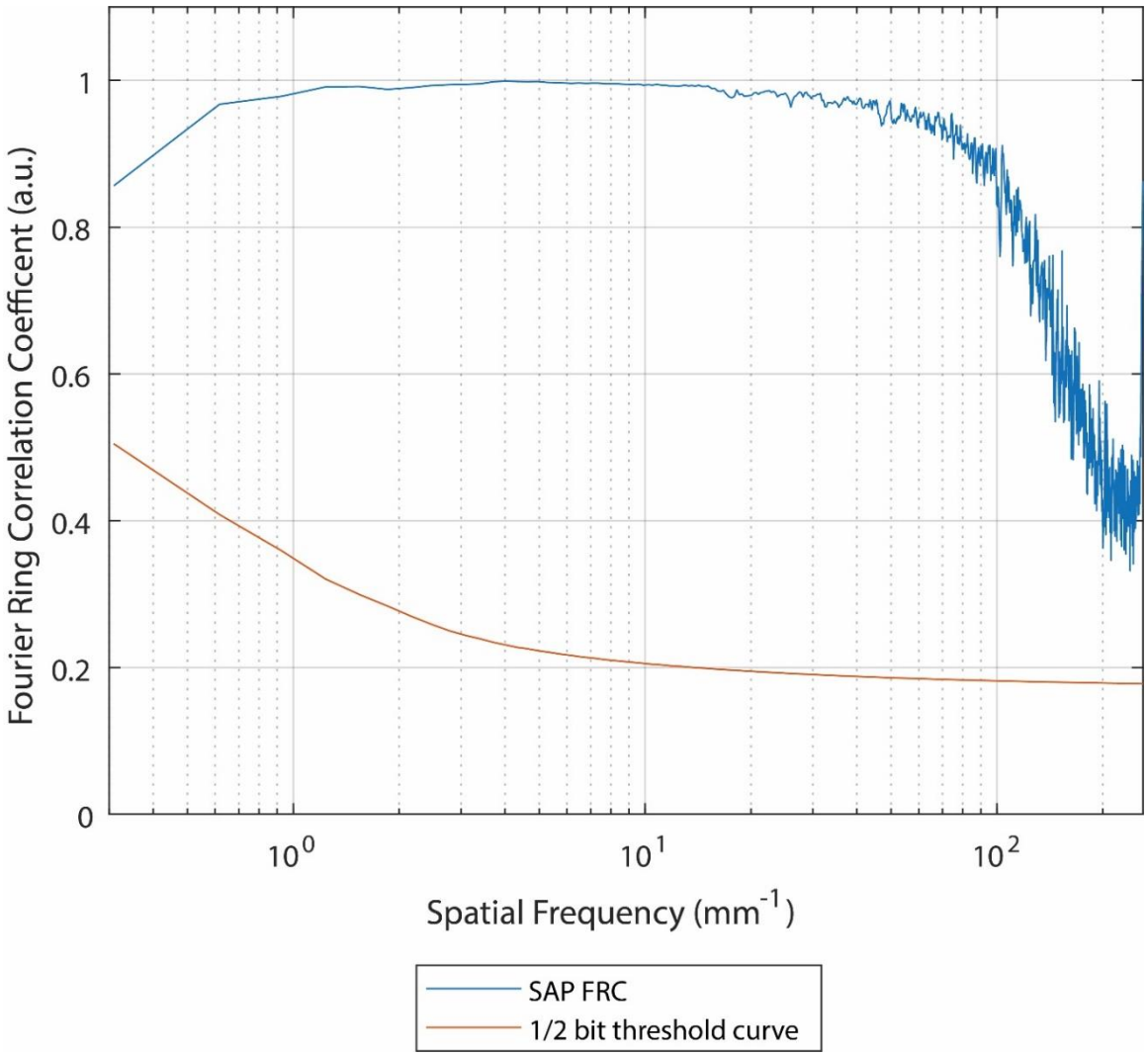

Figure 51: The Fourier Ring Correlation result produced by comparing the reconstructions shown in Figure 50 c) and Figure 50 d). The half bit threshold is also plotted for reference and to enable resolution measurement. Note that the FRC is cut off early as the propagation kernel aliased beyond the range plotted here.

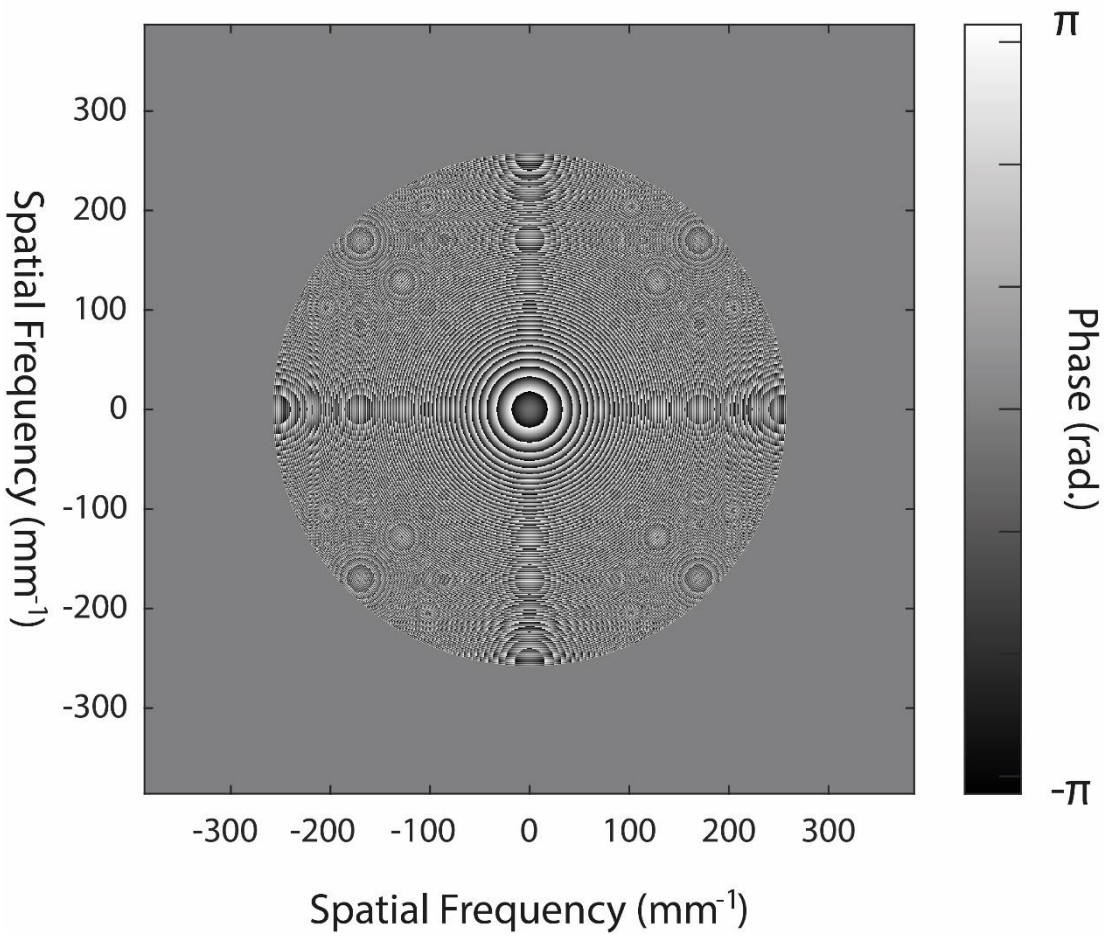

Figure 52: the propagation kernel used in the reconstruction of Figure 50. The colour bar at the right of the figure converts the greyscale of image into radians. The flat grey area around the propagator indicates the part of the propagator set to zero in order to avoid the aliasing of the propagation function.

Following the same process which was used to assess the phase sensitivity of near field optical ptychography to determine the phase sensitivity of optical SAP, groups 2-4 of elements were considered (see section 2.3). Their phase was converted to depth using equation 104. Depth maps of groups 2 to 4 can be seen in Figure 53, and the corresponding histograms can be seen in Figure *54*. Table 3 contains the mean and standard deviation of the etched depth for these particular depth subsections. Looking at Table 3, it can be seen that all the mean depths are very close to the specified depth of 303nm, showing that the phase recovered is quantitative. However, SAP seems less precise at measuring the phase, as the standard deviation of the etched depth has increased for all groups compared to the near field ptychography reconstructions. A possible reason for this increase can be seen when looking at the magnitude reconstruction of the specimen: high frequency artefacts cover the specimen phase and magnitude reconstruction (see Figure 55 b) compared to Figure 55 a)). These high frequency artefacts take the form of lines with equivalent lengths and widths to the elements of the phase targets, and are likely due to high frequency components of the collected diffraction pattern being incorrectly propagated due to aliasing of the propagation kernel. Several can be seen in Figure 56. Figure 56 shows the depth map of a single element; by singling out an element these artefacts are not hidden by wide range depths/phase imaged in the whole reconstruction. Following the same process as laid out in section 2.3, the phase sensitivity of SAP is found to be $\frac{2\pi}{74}$. The phase sensitivity was calculated using the subsection shown in Figure 56 b). The reduced phase sensitivity of optical SAP is likely because the aliasing of the propagation function has a greater effect on the SAP reconstruction than on near field ptychography reconstruction. This

seems to suggest that it is not solely the interference of diffraction fringes which determines whether the phase is recoverable, but also the degree to which they are changed by the interaction. Comparing diffraction patterns from optical near field ptychography and optical SAP datasets, it can be seen that the diffuser has caused a greater degree of variance in the intensity values than the aperture, and this greater variance may have helped the reconstruction process. This likely relates to the findings of Clare et al. [132] who investigate the effect of speckle and variance on reconstruction error of near field ptychography. Although this result is as expected, it is important to have an explicit visual example. It is also important that SAP reconstructions were successful and as can be seen from comparing Figure 57 with Figure 25, have made better use of the detector field of view compared to previous SAP experiments. This suggests it is possible to perform SAP experiments on TEM with an improved field of view per diffraction pattern. TEM results are required to confirm this, with the biggest practical issues likely to be the transverse coherence (see section 1.35) of the illumination system. It is important to confirm these results in the electron domain with a similar direct comparison.

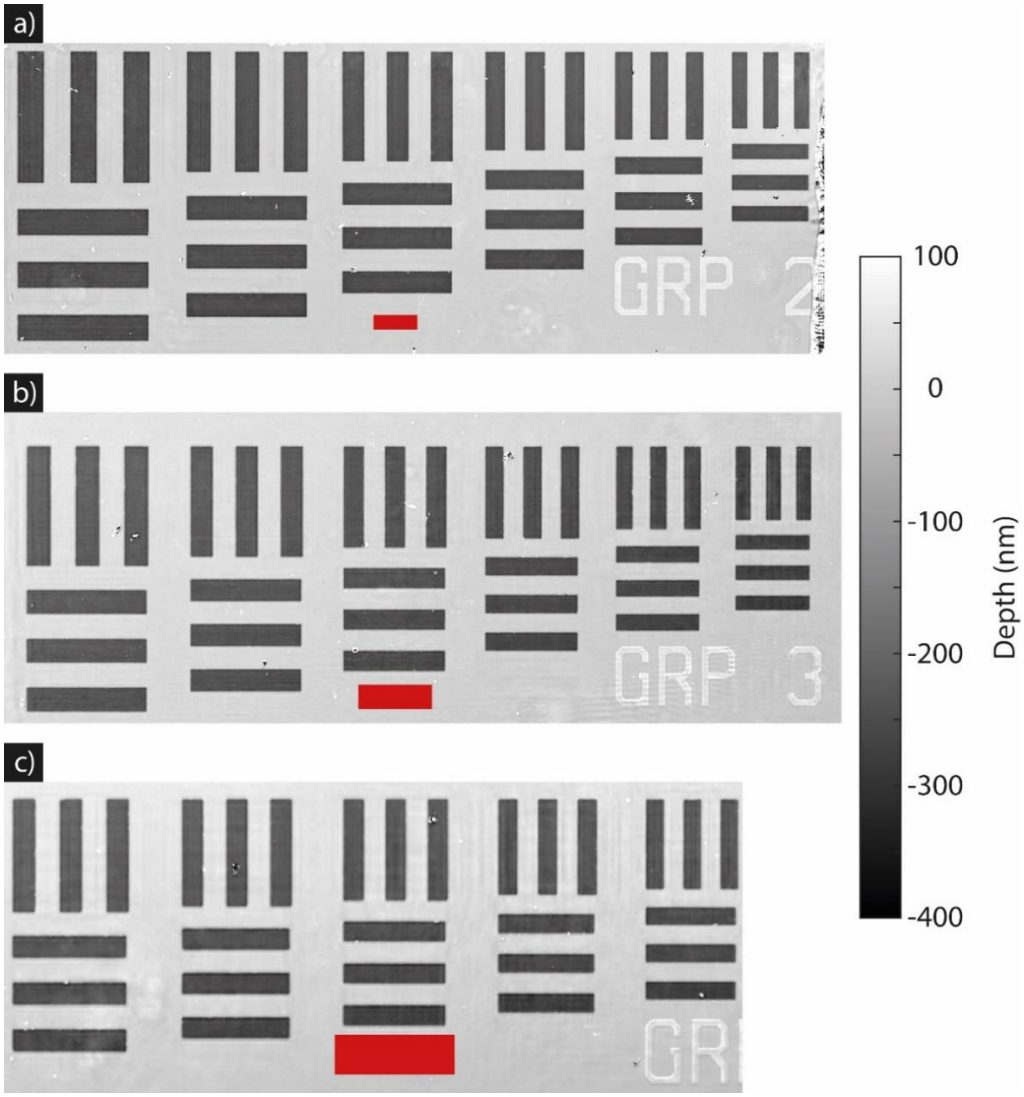

Figure 53: a) depth map of the group 2 elements produced by a SAP dataset. b) depth maps of the group 3 elements produced by a SAP dataset. c) depth maps of group 4 elements produced by a SAP dataset. Note that this figure deploys contrast enhancement, the maximum depth of this figure is artificially limited to 100nm. The minimum depth is artificially limited to -400nm. The red scale bar indicates 200µm. The colour bar to the right of the figure converts the greyscale of the images to depth in terms of nanometres.

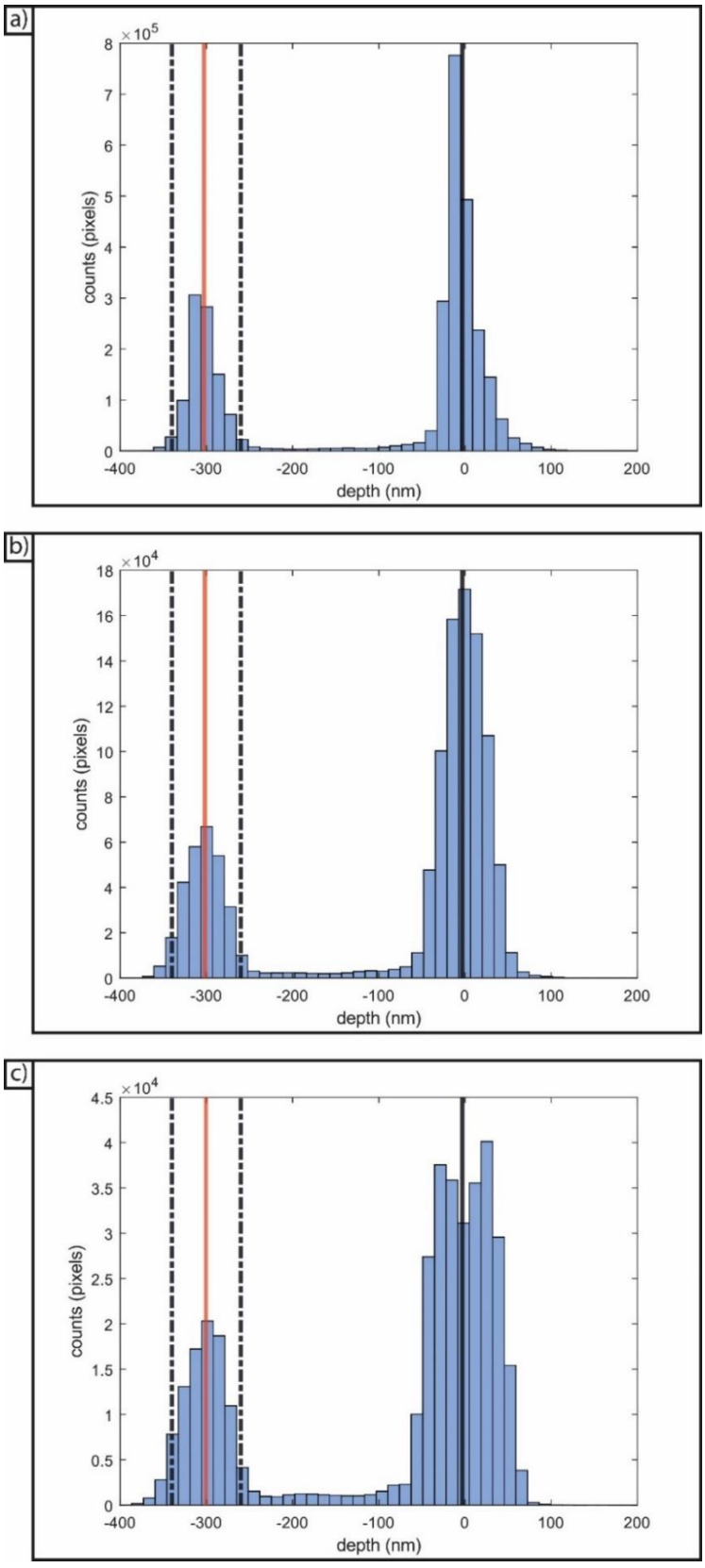

Figure 54: a) histogram of the Group 2 depth values which can be seen in Figure 53 a), b) a histogram of the Group 3 depth values which can be seen in Figure 53 b). c) a histogram of the Group 4 depth values which can be seen in Figure 53 c). The red lines in a)-c) indicate the average depth of the etched surface which was calculated using the data within the bounds indicated by dotted black lines. The solid black lines indicate the average of the unetched surface.

Group number	Mean depth (nm)	Standard deviation of the etched depth (nm)
2	302.8	15.8
3	301.7	19.0
4	300.5	19.6

Table 3: a table showing how the measured mean and standard deviation of the etched depth of the SAP image varies with group number. These values were calculated using the values within the black dotted lines seen in Figure 54.

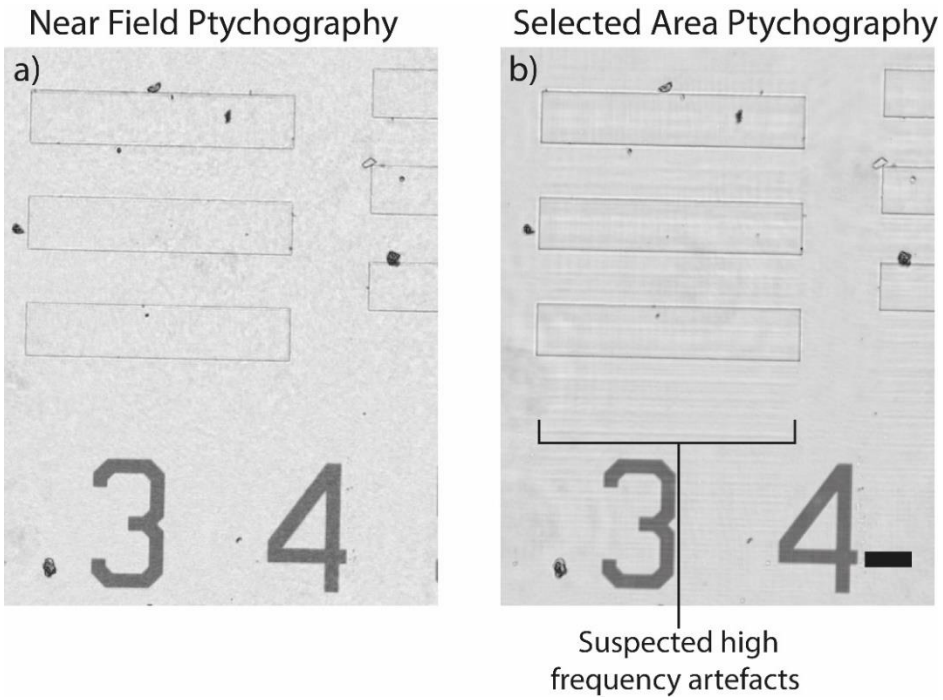

Figure 55: a) a subsection of the reconstructed magnitude of optical near field ptychography. b) a subsection of the reconstructed magnitude of an optical SAP dataset displaying a high frequency artifact. The scale bar in this figure indicates 100µm.

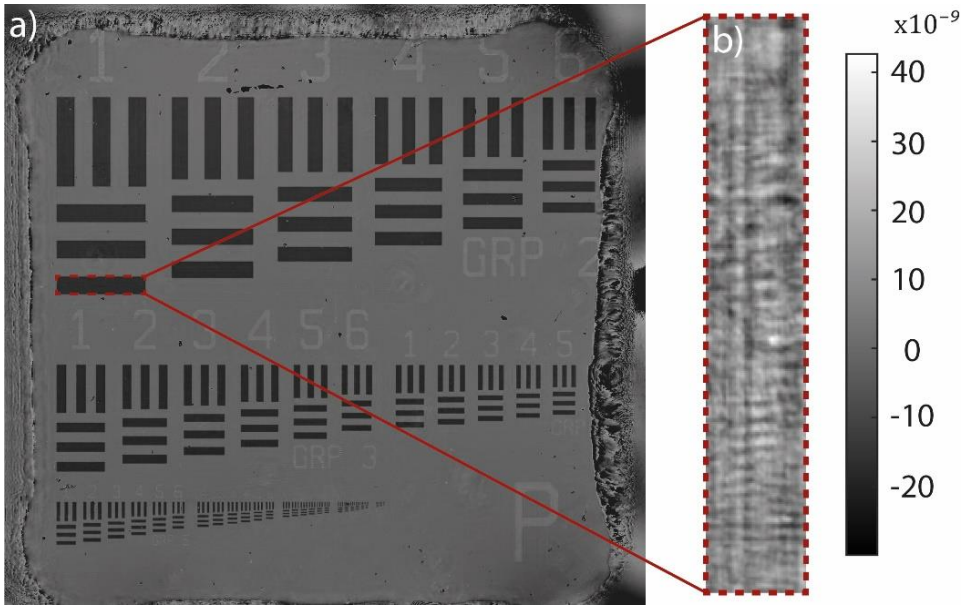

Figure 56: a) the total reconstructed phase of the specimen obtained by SAP, where a red dotted line is used to indicate the subsection used in the depth map shown in b). b) the depth map subsection which was used to calculate the phase sensitivity of optical SAP. Note that the colour-bar only applies to b) and converts the greyscale of b) into depth in terms of nanometres.

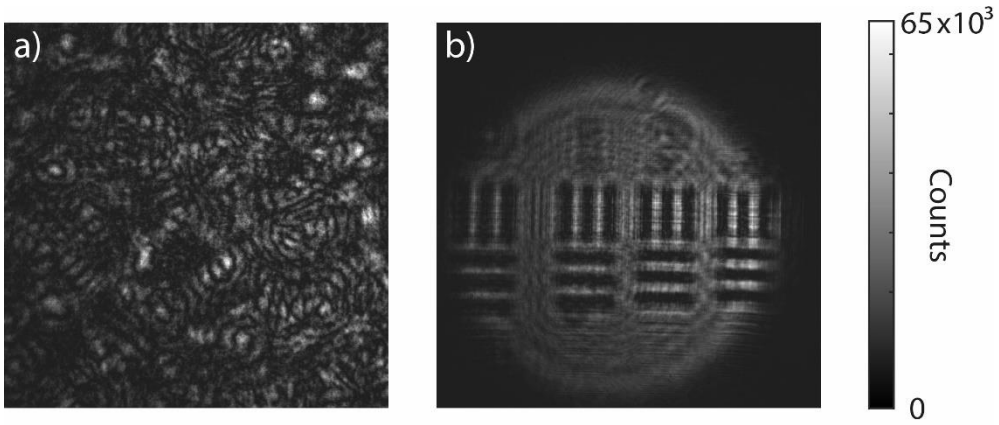

Figure 57: a) an example of near field optical ptychography diffraction pattern. b) an example of a SAP diffraction pattern.

2.5 Conclusions on optical near field ptychography

In this section near field optical ptychography and optical selected area ptychography have been deployed using the cone beam geometry. The phase sensitivity of near field optical ptychography and selected area optical ptychography was investigated, as this would give an approximation of the phase sensitivity of near field electron ptychography before further experiments. This allowed the application of tools described in the second half of section 1, mainly the FRC. Phase sensitivity is important as the potential difference between a p-type doped semiconductor and n-type semiconductor can be less than a volt, or the electric field between the two can be less than 1 MV per cm [120] . With this context the importance of measuring the phase sensitivity of near field optical ptychography is highlighted, in that if near field electron ptychography is going to be a useful technique in the TEM it will need to achieve a high phase sensitivity in order that it can address the problems currently being addressed by off-axis holography and DPC [120]. It was also shown here that near field optical ptychography's phase sensitivity is almost the same order as focused probe optical ptychography [124]; if this is true in the electron domain then near field electron ptychography should have the required phase sensitivity for many applications in the TEM. Therefore, the results presented here suggest near field optical ptychography has a phase sensitivity of 0.0654 radians, suggesting that near field electron ptychography may have the same potential. This claim is supported in later sections in which the mean inner potential of latex spheres is determined with a standard deviation of < 1V. This phase sensitivity was achieved in sub optimal conditions such as a dusty environment and an aliased propagator, suggesting that near field optical ptychography could have greater phase sensitivity than the values shown here, but would require further experiments to confirm. The results presented here align with previous near field optical experiments [125], as the phase images collected by both near field optical ptychography and optical selected area ptychography are quantitative as they match well with the design of the specimen. Not only are the results here quantitative, but are shown to be self-consistent over a range of spatial frequencies, which is important as it highlights the possibility that near field ptychography can encode multiple spatial frequencies in a single defocused diffraction condition, which is a significant advantage over methods like inline holography. Interestingly, SAP performed less well at lower frequencies than near field ptychography, which could be due to the smaller probe size with respect to the specimen (requiring more up sampling of the lower frequencies), or the reduced diversity of the dataset. In this section it was also shown that the SAP could be implemented with a larger aperture than previously implemented in the electron domain, suggesting that SAP can collect larger fields of view more efficiently for those who do not wish to implement a diffuser into their TEM. However, effects like the required coherence of the electron beam were not considered here, which may limit the implementation of a large aperture SAP in the TEM. Near field optical ptychography

still does provide a larger field of view per diffraction pattern, as near field ptychography required only 750 diffraction patterns, whilst SAP requires 1400 for the same field of view in the above experiments. Near field optical ptychography has shown it is robust to the form of the illumination, even allowing for the illumination to have a different image plane to the specimen. This is important for the electron implementation, as will be set out in detail in section 5.2.1. It has now been shown that near field optical ptychography, and by extension near field electron ptychography, has sufficient phase sensitivity, and it is therefore time to investigate the challenges of implementing near field electron ptychography in the TEM before it is used to investigate samples and its ability to obtain quantitative images is tested.

Chapter 3

3 How to collect near field electron ptychography data

3.1 Introduction

Now that near field ptychography has been demonstrated on the optical bench where components and distances can easily be calibrated and measured, it is time to consider the difficulties that result in applying near field ptychography to/in the TEM (Transmission Electron Microscope). As discussed previously, near field ptychography requires that the image of the specimen and the diffuser is out of focus (or defocused) at the detector. The TEM often operates in defocused modes for various types of experiments such as high resolution imaging of single particles, where the defocus is used to enhance the contrast by using the Scherzer defocus condition [133]. Inline holography requires that either the specimen is moved away from the image plane to create the required defocus condition, or the objective lens strength is purposely weakened [95]. However, the requirement that the diffuser is also defocused and ideally shares the same plane as the specimen, or its image, means that near field electron ptychography's defocus has to be implemented in a different way compared to these other methods mentioned before. The defocus condition is achieved in the same fashion as set out by Maiden et al. [94], the desired microscope condition is achieved by placing the TEM into diffraction mode and then changing the strength of the diffraction/intermediate lens, such that a defocused image of the specimen and diffuser is obtained. This defocus is then fine-tuned by further adjustment of the diffraction/intermediate lens and the camera length control of the TEM in diffraction mode. A disadvantage of this implementation of the diffraction condition is that the exact distance between the image plane (of the specimen/diffuser) and the detector is unknown, and therefore additional reference data are required, as will be discussed later in this chapter. Figure 23 gives a good visual representation of how near field electron ptychography is set up in the TEM, which is notably different to the optical implementation in the previous section and the original implementation in the X-ray domain [79] where the cone beam is utilised. This is due to the advantage of imaging with an electron lens, which can directly shape the electron beam, as opposed to X-ray optics which typically rely on reflection and diffraction to shape the beam.

However, the diffraction condition is not the only critical component of near field ptychography; the other critical component is the scanning of the diffuser with respect to the illumination. For many years (focused and defocused probe) ptychography has made use of the scanning system of the STEM to perform this element of ptychography, for example the Gatan DigiScan system used to synchronise the scanning of the beam with detector captures [134]. However, due to the diffraction and illumination conditions (parallel illumination) of near field ptychography, such a scanning system is not easily applicable to near field electron ptychography. Therefore, in these proof of principle experiments, the specimen stage will be used to scan the specimen with respect to the illumination. However, near field electron ptychography does not yet have a dedicated system which automates and synchronises the movement of the specimen stage with the detector captures, so in this section the DigitalMicrograph code used to collect near field electron ptychography data will be discussed.

This chapter will describe the following:

- DigitalMicrograph code used to collect near field electron ptychography data
- How the positions recorded by the microscope can be processed in MATLAB to enable successful reconstructions
- How the camera length of the experiment is found with reference diffraction data of an aperture with a known diameter
- The problems of diffraction pattern drift and Fresnel scaling

Block diagrams are provided, describing code in two different software packages: DigitalMicrograph (for data collection code) and MATLAB (for processing the data such that is it suitable for reconstruction).

3.2 The data collection process for near field electron ptychography

The initial step in the data collection code was to generate a grid of positions. The scan pattern positions were stored in a DigitalMicrograph image format, as these were easier to transfer to the MATLAB software system than the tags system typically used in DigitalMicrograph. Raster scan pattern generation requires a few pieces of specific DigitalMicrograph code to determine the initial position of the specimen stage. As a mechanical stage was used in the data collection and there was a requirement for the data collection to be as quick as possible, alternative scan patterns were not tested. The original code could only produce square scan patterns and not rectangular scan patterns (where the number of steps in the x and y directions are unequal), and therefore additional code has been provided in the appendix A (see Section 7.1) which allows for rectangular and square scan patterns. The block diagram of the code is shown in Figure 58 and describes both the improved code and the original code.

The algorithm described in Figure 58 starts by obtaining the initial positions of the stage and using this as the starting point of the scan (the EMGetStageX/Y function obtains the position of the stage from the microscope). This means that the specimen should be moved so that the sample of interest should be offset from the starting field of view, so that it is in the centre of the scan pattern. This is altered by pre-calculating the size of the scan and offsetting the DigitalMicrograph image by this amount. The initial position generated by the code is an overscan position, as is every first position in each column (when Counter2 == 1). The camera test (CT) value, which determines whether a camera exposure is required, is set to zero during these overscan steps, in order to reduce the time the experiment takes. As there is an overscan, the input StepY is increased by one in order to obtain the expected number of diffraction patterns. The code has 4 outputs: the *x* positions, the *y* positions, the CT vector, and the sequencer vector. The sequencer vector is used to debug the timing of events in the code. The term ‘overscan’ in this thesis means an additional step in the scan pattern implemented in the mechanical and piezo systems, with the intention of mitigating a hysteresis-like effect or cogging, which may prevent the desired step size from being fully implemented when the direction of the scan pattern suddenly changes, for example at the beginning of a new row.

The scan parameters generated by the code described by Figure 58 are used by the data collection code, which is illustrated by Figure 59. Once the positions are generated it is a matter of executing the generated positions in a sequential order and checking the CT value to determine whether a camera acquisition should take place or not. When the microscope is given the command EMSetStageX(oX+3), where oX is the original position, the microscope action will be to move the specimen 3 micrometres to the right of the current position. A secondary objective for the data collection code is to collect information on the reported position of the microscope. This is done using the EMGetStageX\Y functions.

Saving diffraction patterns to disk is an important step in the data collection code, as it forms a significant proportion of the data collection time. Whether or not this is necessary depends on the available RAM of the computer accompanying the TEM. If the dataset is small and the RAM large enough, it may be worth storing the data in the RAM until the end of the experiment, at which point the data can be transferred to storage, reducing the total data collection time. For reference, a single 1024 by 1024 diffraction pattern has a memory consumption of 5131 kilobytes of RAM. Lastly, the time it takes to save a diffraction pattern can be approximated pre-experiment, by commenting out the microscope commands and saving a blank image to the hard drive instead of activating the detector. In order to obtain the average save time, start a stop watch (or use DigitalMicrograph functions) and measure the time this debug code takes, and then divide by the number of diffraction patterns.

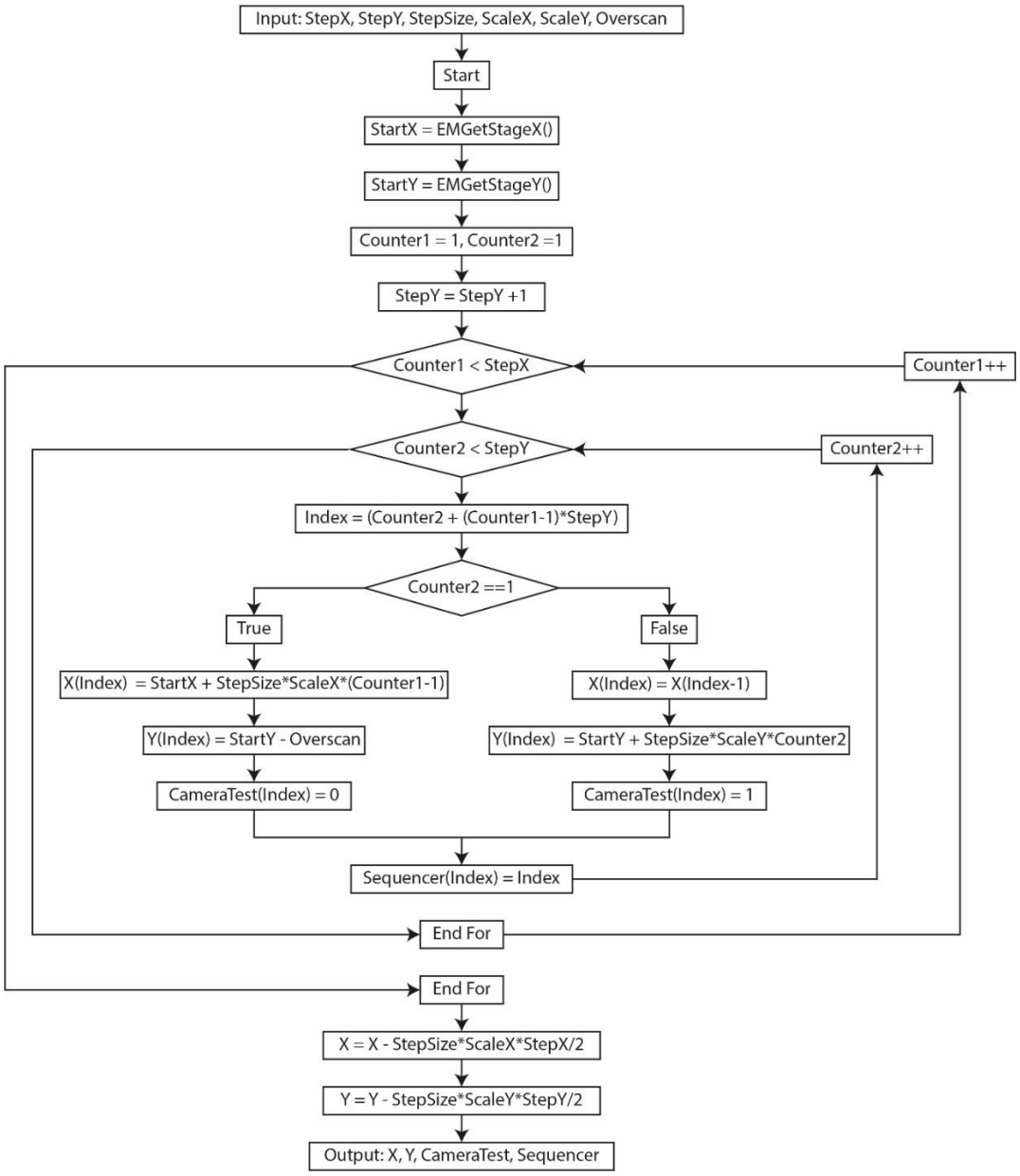

Figure 58: block diagram outlining the algorithm for creating a scan pattern for a mechanical stage with overscan after every column. This is for DigitalMicrograph.

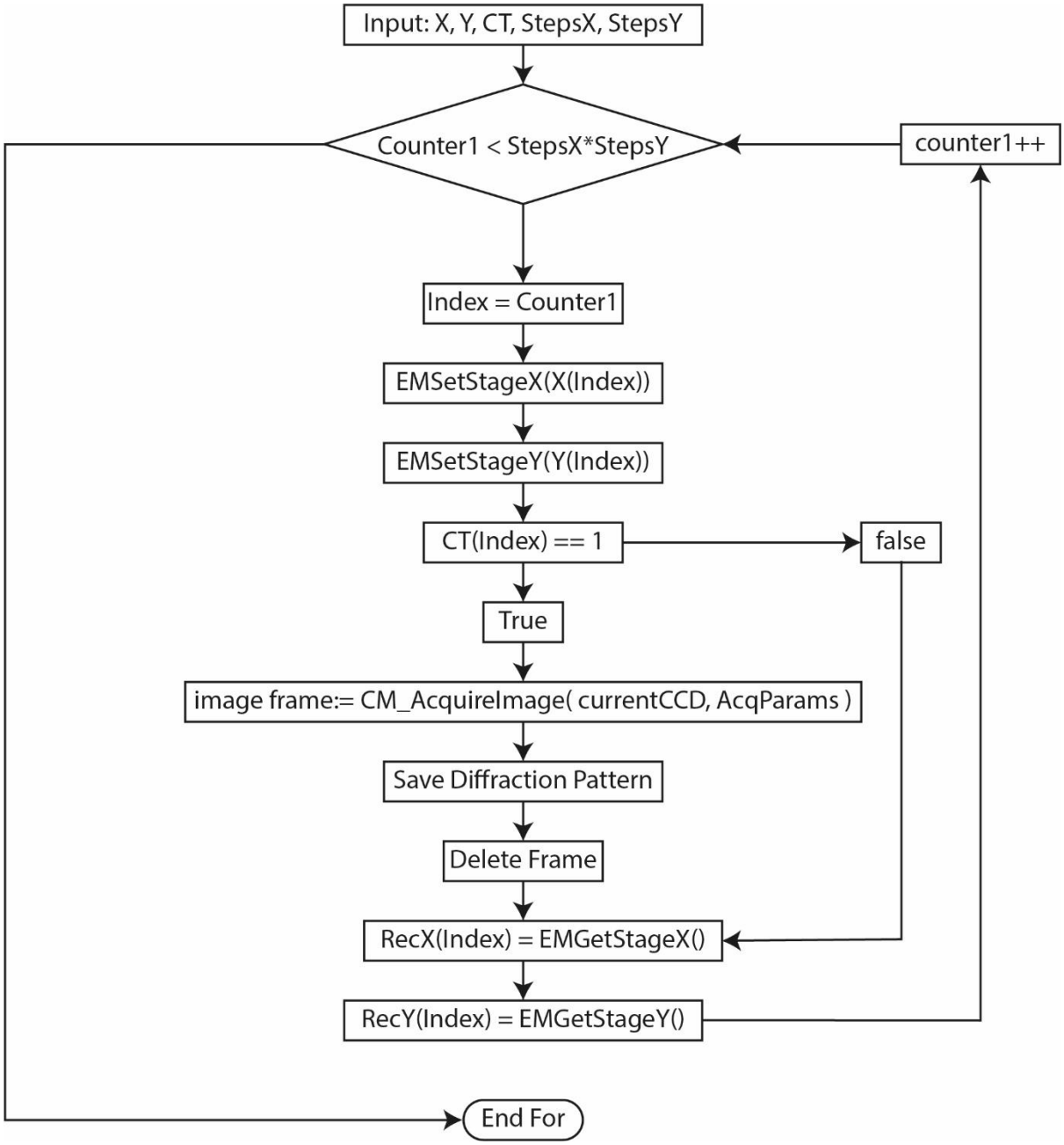

Figure 59: A flow chart laying out the process of the data collection code. The exact code this describes can be found in the appendices (see 7.2-7.4). This is for DigitalMicrograph.

An alternative to using the microscope specimen stage to translate the sample is to scan the image produced by the objective lens across the selected area aperture. This can be accomplished using the shift coils, by converting the part of the code in Figure 59 which moves the specimen stage. Prototype code can be found in the appendix B (see section 8). Scanning via the scan coils would enable different scan patterns, as there will be negligible cogging or hysteresis, although to date, the only work that has to be done for this code to work is to calibrate the EMSetImageShift function.

Another important next step will be to combine the action of the mechanical stage and the action of the scan coil together, to form a hybrid scanning system. Hypothetically, a hybrid system will overcome the faults in both individual methods, as the scan coils can move the image of the specimen quickly and reliably, while the mechanical stage can move the specimen long distances without adding aberrations to the electron beam [135, 136]. The last step to mention is beam

blanking. This is especially important if biological samples are to be considered, as most biological samples are highly beam sensitive. Beam blanking can be achieved using the pre-specimen scan coils and the C2 aperture, as if the scan coils displace the beam outside of the aperture the beam is effectively stopped. This simple form of beam blanking can be implemented via DigitalMicrograph code, but other methods should also be investigated [137].

Once the data collection code of Figure 59 has been run and the diffraction data collected, the DigitalMicrograph files (.dm3 and .dm4) are converted into MATLAB format using the "dmread" function written by Andreas Korinek and found in the MATLAB file exchange (https://uk.mathworks.com/matlabcentral/fileexchange/45933-gatan-digital-micrograph-file-reader).

3.3 Processing position data from the TEM into usable positions for near field electron ptychography

Now that the data collection code has been discussed, we must analyse aspects of its performance, starting with reported positions. In this section the FEI Titan 80-300 STEM operating at 300keV was used to test the near field ptychographic data collection code described in the previous subsection, whilst in the altered uncorrected diffraction mode described at the start of this section. Further details on the microscope conditions used during this section and the final experiments can be found in subsection 4.4. Near field electron diffraction patterns were collected by the Gatan UltraScan CCD (2048 by 2048 pixels, with binning value of two). All specimen movements were achieved via movement of the mechanical specimen stage in order to achieve ptychographic overlap (the specimen stage was a standard mechanical stage of the FEI Titan range). Generally, in normal TEM mode, the positions reported by the specimen stage (goniometer is alternative name) are of sufficient accuracy for ptychographic reconstruction when combined with position correction and rotation correction as is demonstrated in this section. The FEI Titan microscope used in this thesis has a minimum step size of 1-2nm, although this is based on personal experience rather than thorough analysis. However, a previous ptychographic experiment on another microscope (JEOL R005 HRTEM) experienced significant hysteresis when using the specimen stage attempting a step size between 10nm and 20nm [14], therefore the typical step size used in this thesis was around 100-200nm. Due to the unique diffraction condition of near field electron ptychography, the positions reported by the microscope need to be processed before they can give the best reconstruction quality. This subsection will explain why this is the case, and describe the tools developed to partially correct the positions so that the reconstruction is improved.

The best way to describe the position processing method is to go through an example. The example that will be examined here is a 20 by 20 square scan. The step size was 0.2µm and an over-scan of 0.5µm was included when starting a new column. Figure 60 shows the ideal positions generated by the scan pattern generation code.

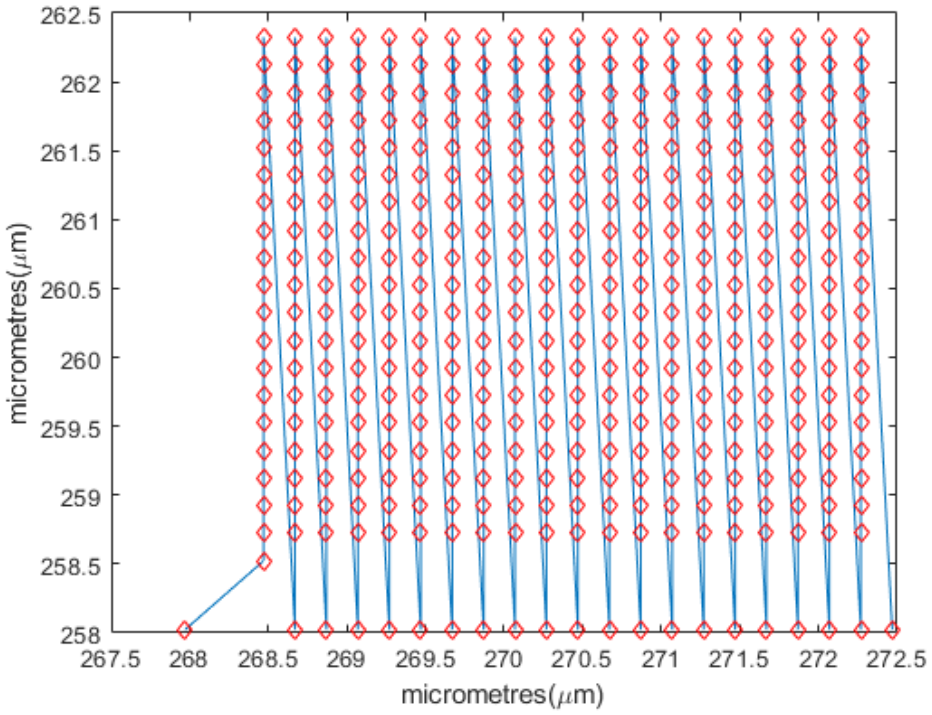

Figure 60: Positions generated by position generation code. The red diamonds outline the positions at which the stage will stop to collect diffraction data, while the blue line outlines the overall movement of the stage. There are two additional points at the start compared to the recorded data, which represent the initialling of the stage to start movement without cogging.

As can be seen by looking at Figure 60 there is an additional position which is not described by Figure 59, as it is not technically part of the data collection. This diamond in the bottom left corner is the position used to prepare the stage for the scan and to prevent cogging of the stage. The first step in the position analysis should be to compare the generated positions to the positions recorded by the microscope itself; Figure 61 displays the positions recorded by the microscope via the EMGetStage function.

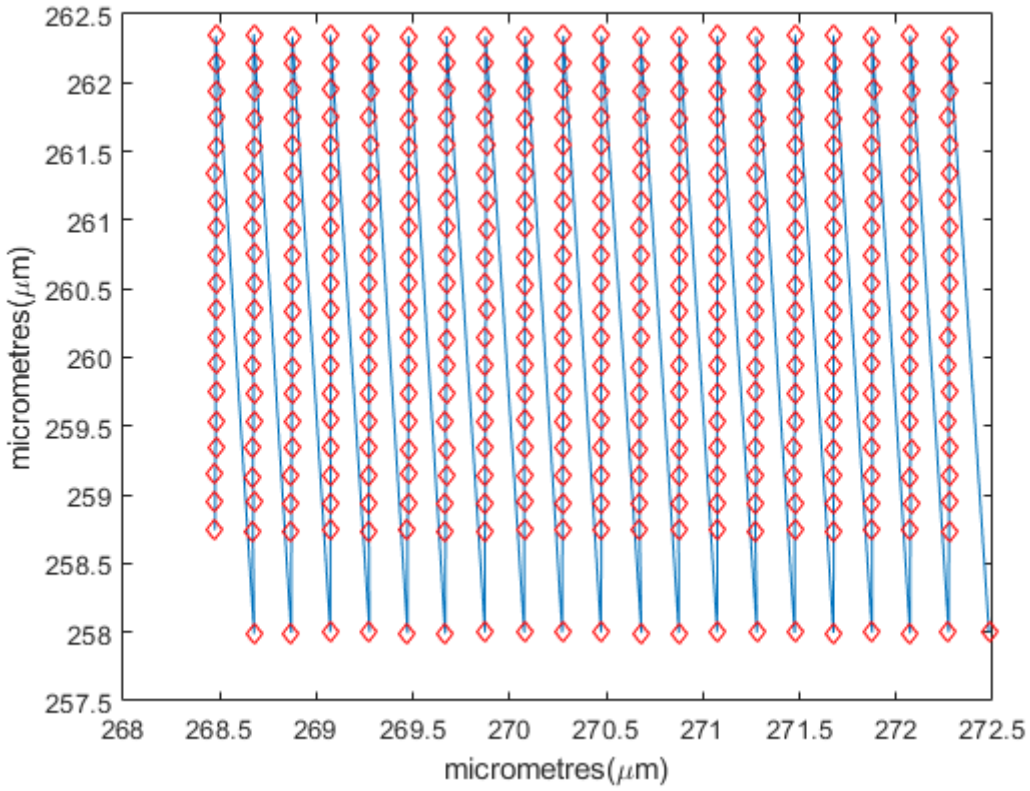

Figure 61: positions recorded by the electron microscope. The red diamonds outline the positions at which the stage will stop to collect diffraction data, while the blue line outlines the overall movement of the stage.

As can be seen by inspection, the only difference between the generated positions and the recorded positions is the exclusion of the initial positions used to prepare the stage for the scan pattern. After discarding the bottom row of overscan positions, the remaining positions and their corresponding diffraction patterns were fed into the reconstruction algorithm, leading to the reconstruction seen in Figure 62. Whilst features are visible in the image, suggesting that the step size was consistent and the overall form of the positions is correct, resolution is low and there are clear artefacts.

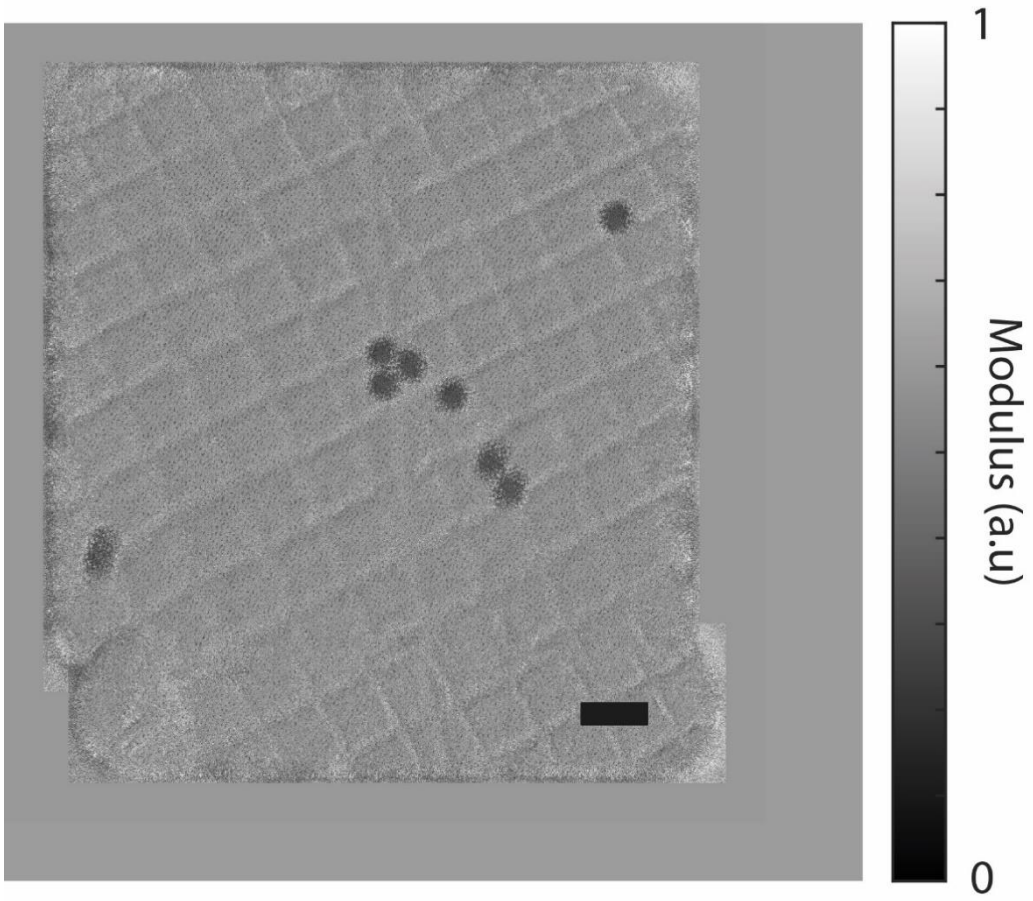

Figure 62: Magnitude reconstruction of the dataset using unrotated positions after ten iterations. The scale bar is 500nm.

The first step in debugging the positions should be to consider potential sources of global positioning errors. Global positioning errors are a combination of rotation and scaling of the scan pattern in *x* and *y* directions [74]. Rotation is likely to be the greatest form of global error due to the nature of defocusing electron beams relative to their image plane, as when the electrons interact with the magnetic lens they experience a force perpendicular to their direction of travel and therefore obtain angular momentum around the optical axis due to the Lorentz force [6]. The most useful way to iteratively test for the global error of rotation is to observe the error of the reconstruction after a few iterations at a particular rotation. If the reconstruction error is less than the error of the last reconstruction, that is often the correct direction.

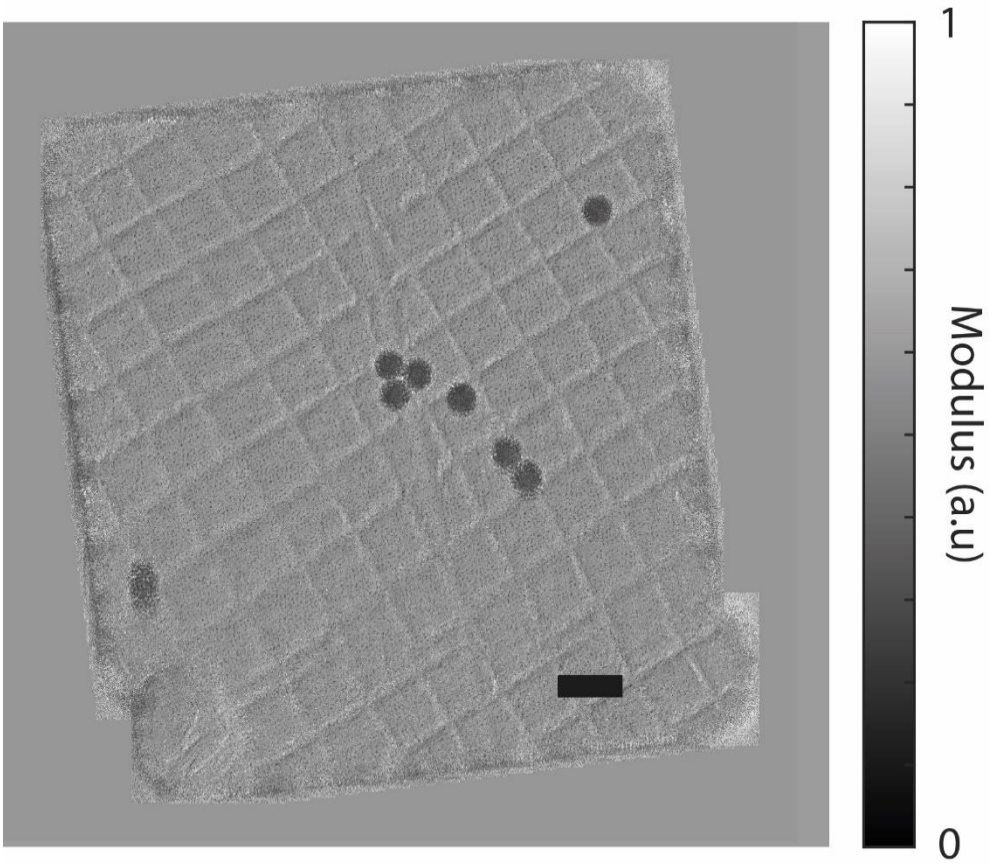

Figure 63: Magnitude reconstruction of the dataset using rotated positions after ten iterations. The scale bar indicates 500nm.

Figure 63 shows a reconstruction attempted with rotated positions such that the error was notably reduced compared to the unrotated case (this was the end result of a period of trial and error using the mPIE reconstruction algorithm). It can be seen from the reconstruction that just rotating the data was insufficent to optimise the positions fully, as some of the latex spheres in the field of view are still blurred and their edges are not clear. The rotation has improved the reconstruction quality compared to the unrotated reconstruction, as the grid lines are clearer than before, but the fine detail of the grid cannot be seen. However, this is only after ten iterations, and it is possible that the best positions could now be found by a dedicated position correction algorithm. This is somewhat true of the unrotated positions too if the position correction algorithm can take account of global changes in positions, such as rotation and stretching.

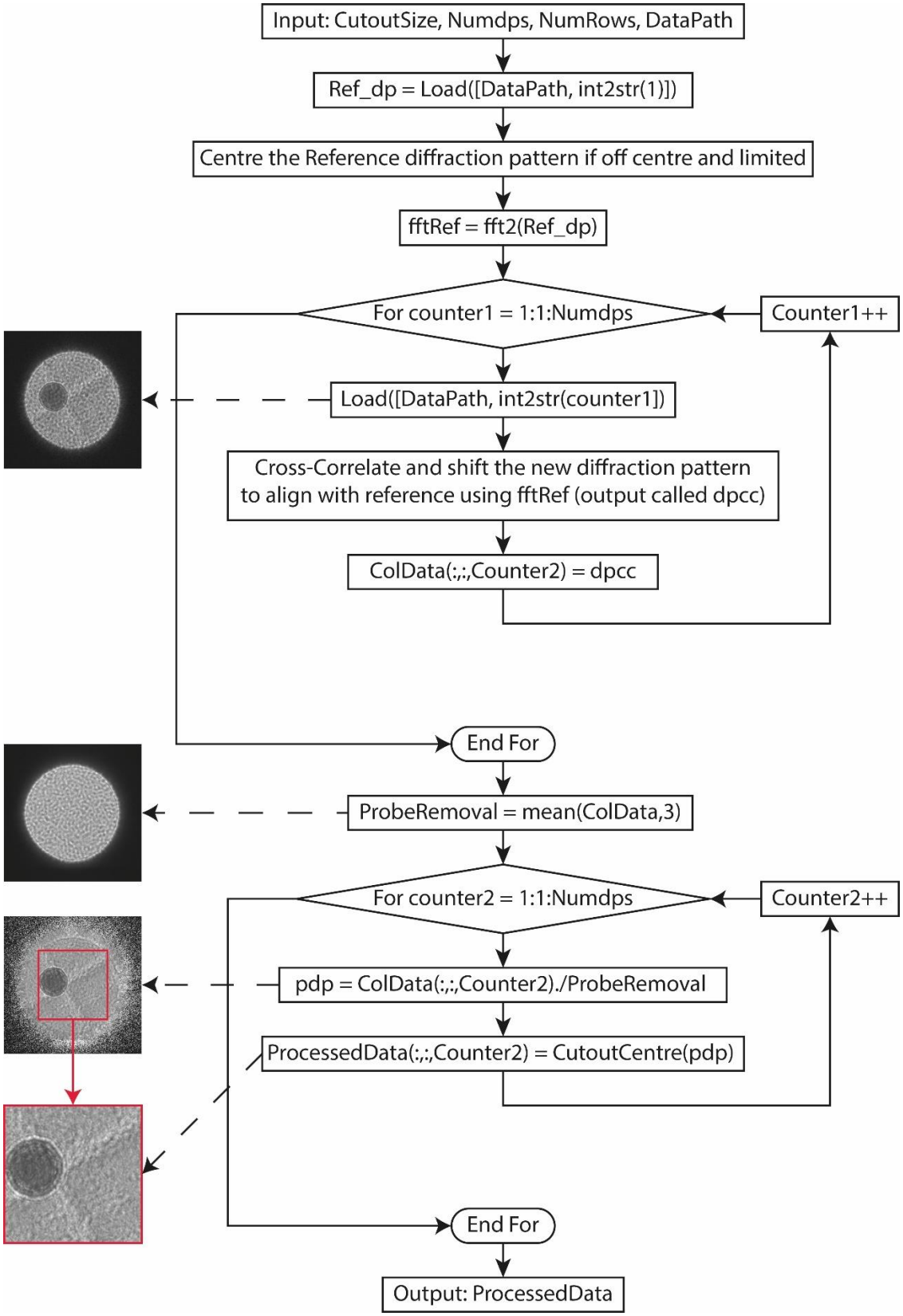

Figure 64: A block diagram showing how diffraction patterns are pre-processed before they are cross-correlated. First the diffraction pattern drift is corrected via cross-correlation with a reference diffraction pattern (typically the first diffraction pattern in the dataset). Then, once the stack of diffraction patterns is aligned, take the average of the stack to obtain a model of the probe at the detector. Then an attempt is made to remove the probe from all diffraction patterns, so the specimen is clear, by dividing each diffraction pattern by the probe model. Once this is done the ProcessedData (MATLAB variable) can be cross-correlated to find the positions of the dataset. This is for MATLAB.

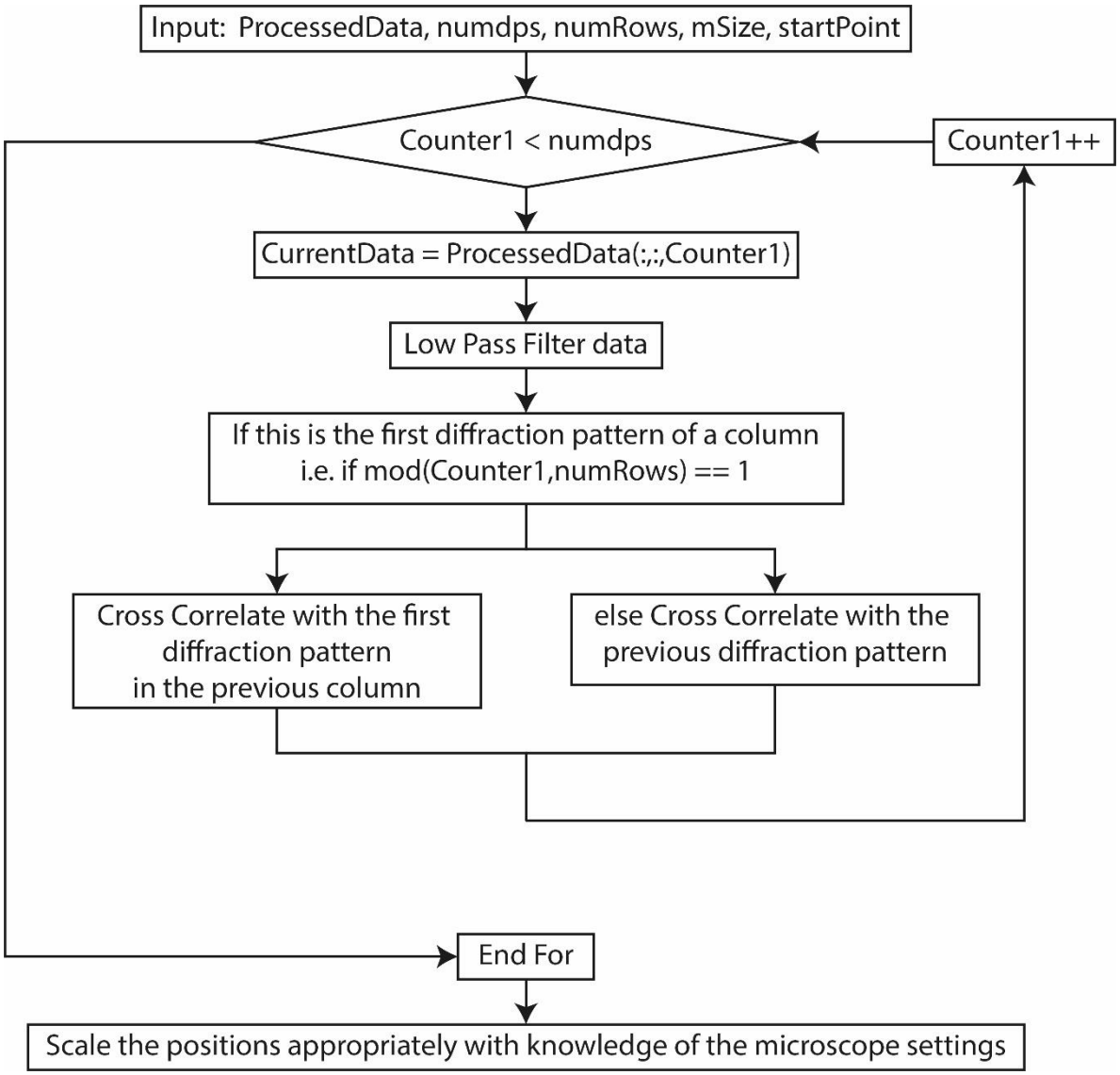

Figure 65: A block diagram showing the process by which positions can be found be cross-correlation of pre-processed diffraction patterns. This is for MATLAB.

To improve on the basic trial-and-error approach to correct positioning errors, Figure 65 displays the process by which the scan positions can be found via cross-correlation: this is a unique benefit of operating in the near-field, where the sample features are visible in the recorded diffraction data. The first step in using cross-correlation is to pre-process the data such that cross-correlation can be effective. The pre-processing (shown in Figure 64) begins with correcting the drift of the diffraction patterns (drift is explained later in Section 3.6), so that the cutting out of the centre of the diffraction pattern is viable. This is required if the diffraction pattern does not cover the entire field of view. The diffraction patterns are aligned by cross correlating each diffraction pattern with a reference diffraction pattern, then circularly shifting the diffraction pattern to the correct position. Circular shift is typically used in binary logic, where the bits of a variable are shifted along memory in such a way that the values at the end are moved to the start (wrap around). Here that same principle is applied in two dimensions. This works as the diffraction patterns are bordered by zeros, which would not be the case if the whole of the detector was utilised. In such a situation it would be recommended to use the average intensity of the diffraction pattern for any new values generated

by the shift. A reference diffraction pattern is ideally a diffraction pattern containing no specimen, or an arbitrary chosen diffraction pattern from the dataset. Once the diffraction patterns are aligned, all of the diffraction patterns are averaged together, such that an average diffuser function is obtained. Using the averaged diffuser to remove the diffuser from all collected diffraction patterns enables the cross-correlation, as the specimen would not be clear if the diffuser was not removed. After that, the centre of each of the processed diffraction patterns is cut out, and the dataset is now ready for cross-correlation.

The optimal way of cross correlating this dataset is to find the shift between the current diffraction pattern and the previous diffraction pattern. However, at the start of the next column (see Figure 66), to compare with the previous diffraction pattern would be incorrect, since the stage returns to the first row in a raster fashion. Instead, by taking the number of diffraction patterns in a column (the height) from the current index, the diffraction pattern to the left or right is accessed, which is appropriate if the shift in the *x*-direction is to be measured. Thereby undertaking the process highlighted in Figure 65, the positions are found and are displayed in Figure 66.

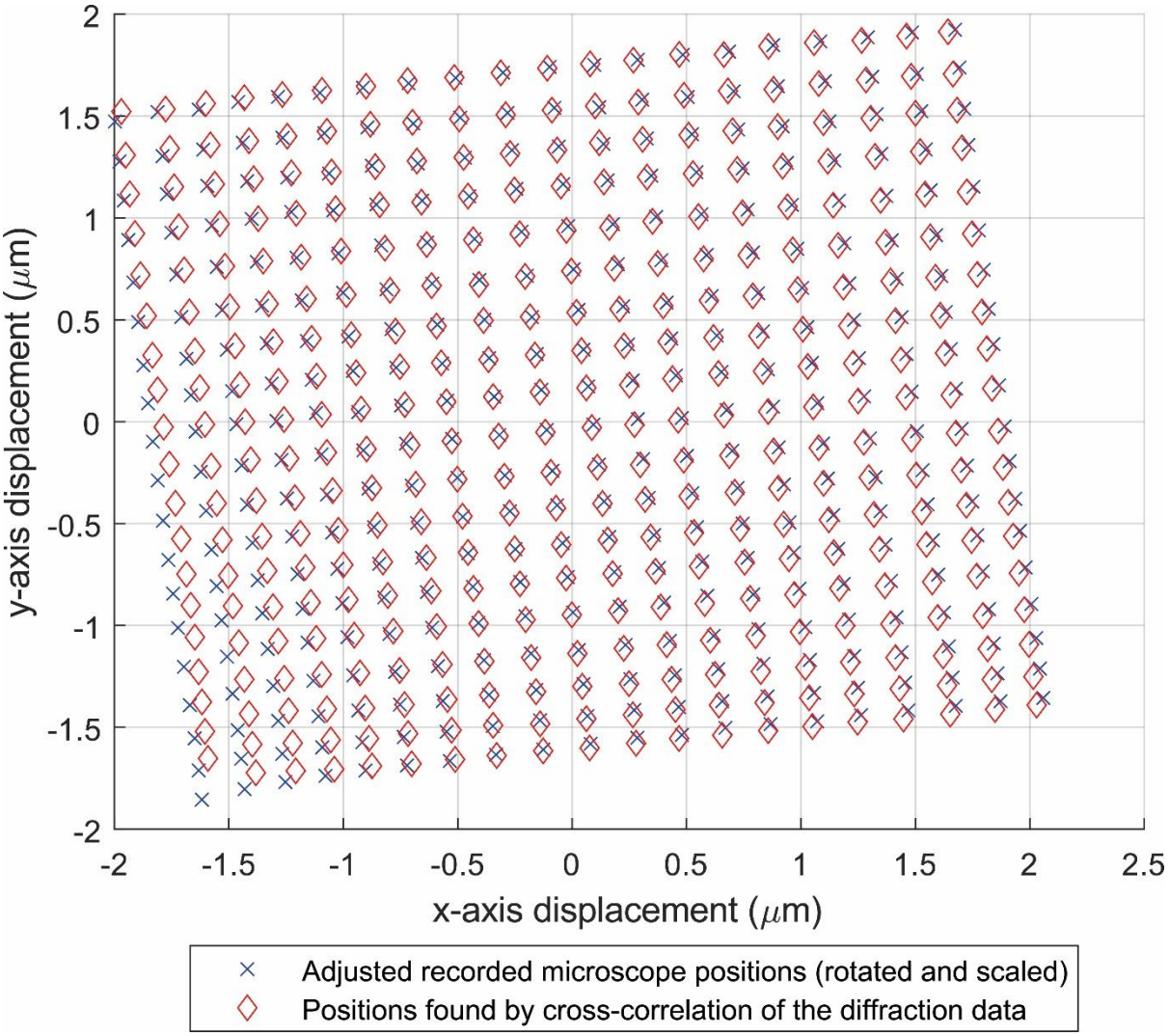

Figure 66: Comparison of the positions found by cross-correlation and the adjusted positions (rotated and scaled).

Also, in Figure 66 are the rotated positions for comparison. Despite the rotated positions matching closely to the positions found by the cross-correlation (CC), the reconstructions differ in quality significantly, as can be seen by comparing Figure 63 to Figure 67. The difference between adjusted

positions and the cross-correlation positions is due primarily to drift in the mechanical stage, particularly in the *x*-direction.

Being able to reliably find the positions helps with finding the other parameters such as the camera length, and is therefore an important step in the experiment. If the object is not directly visible in the diffraction data (like electric fields), the cross-correlation method of finding the positions breaks down. There are therefore specific limitations to the cross-correlation method. If the positions cannot be found via cross-correlation, or the diffuser cannot be reliably removed from the diffraction patterns, a secondary scan immediately after or before the desired dataset without the diffuser can be collected, which can be cross correlated in order to generate positions which are typically reliable. For instance, most of the reconstructions in this thesis had their positions generated by cross-correlation of a diffuser-less dataset collected after the diffraction pattern dataset. It is also important to note that the cross-correlation process discussed here only works for raster scan patterns. For different scan patterns, the cross-correlation process will need to be altered.

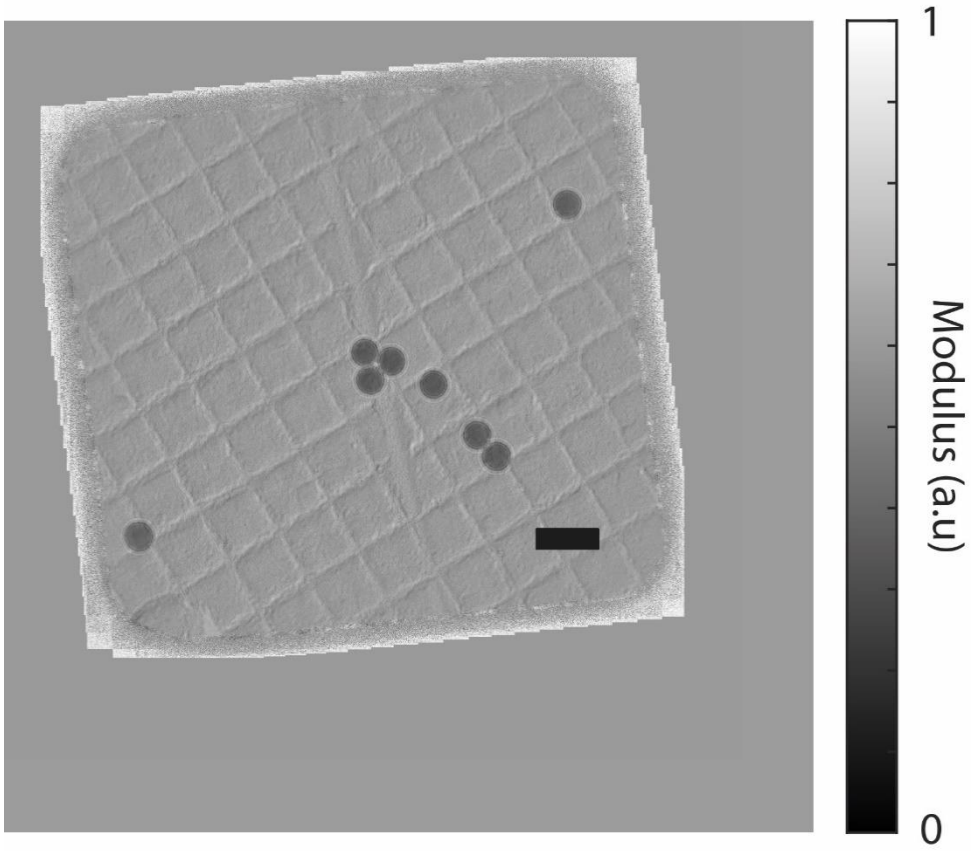

Figure 67: A magnitude reconstruction after ten iterations with positions found via cross-correlation of the diffraction data (the cross-correlation process is outlined in Figure 64 and Figure 65). Scale bar is 500nm.

A final point on the positions: they need to be appropriately scaled to the operational conditions of the microscope (or the diffraction condition). This means that although the positions provided by the microscope are in terms of nanometres, we also need them in terms of the pixels of the demagnified detector. This is also true of the positions found by cross-correlation; the effective pixel size at the specimen plane needs to be known, to scale the cross-correlated positions, as positions found by cross-correlation are in terms of pixels not nanometres. Therefore, the total magnification of the diffraction condition must be known, as well as the magnification of the intermediate image planes of the microscope.

3.4 How was the camera length found?

Near field electron ptychography requires a diffraction condition where a defocused image of the specimen multiplied by the diffuser is incident at the detector plane. This cannot be achieved through moving the object out of the eucentric plane, as it is necessary that the diffuser is also out of focus at the detector. The easiest way of achieving this condition is to set the microscope into diffraction mode, then adjust the diffraction lens strength until its image plane is just out of focus compared to the image plane of the objective lens. In diffraction mode there is direct control of the diffraction condition, as the projector lens system can be controlled by changing the camera length setting. To calibrate the defocus introduced in this way, a diffraction pattern of an aperture of known diameter was collected under the same diffraction condition as the experiment itself. A ten-micron diameter circular aperture was used as a reference in these experiments, but any aperture can be used (the author recommends a 50µm diameter aperture). Figure 68 shows a ten-micron aperture in four different diffraction conditions. These four diffraction conditions relate to an experiment discussed in section 4.13.

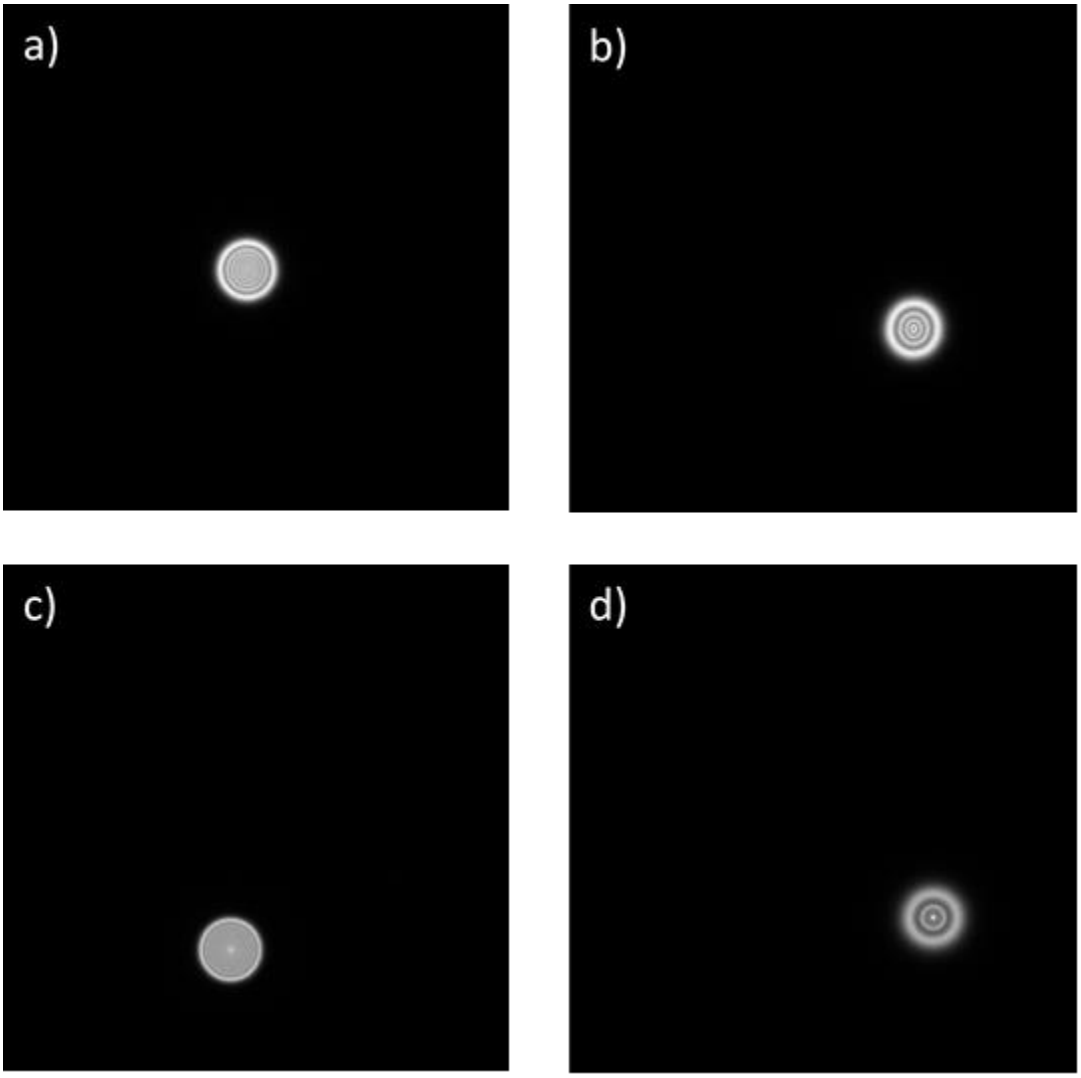

Figure 68: a) shows a 10µm diameter aperture which is 0.77m out of focus on the detector. b) shows another 10µm aperture with a different defocus, where calibrated microscope camera length has been doubled. c) same as b) but the camera length has been halved instead. d) same as b) but the camera length has been increased further.

The data displayed in Figure 68 are processed according to the equation found in [94] by Maiden et al. who worked on a precursor of near field electron ptychography; the equation is found below in equation 105:

$$Z = \frac{D}{d} M C_L \tag{105}$$

The parameters of equation 105 are the following:

- *Z* is the propagation distance between the diffuser and the detector (the specimen image plane also intersects with the diffuser)
- *M* is the magnification of the objective lens
- *d* is the measured diameter of the bright field disc at the detector plane.
- *D* is the known diameter of the aperture (10µm)

Equation 105 describes the process of finding the camera length. The first step is to find the total magnification of the post objective lens system. This is performed by comparing the known diameter of the aperture and the approximation of the diameter of the diffracted aperture. This approximation can be performed in MATLAB by testing the collected intensity to see where it is greater than the minimum of the bright diffraction disc. This test is used to produce a mask that approximates the radius of the diffracted aperture on the detector. Using the known size of the detector pixels and the mask radius, the magnification of the post objective lens system of the TEM can be estimated. It should also be noted that any divergence of the beam (an additional source of magnification) on the specimen is accounted for as well (ideally the illumination should be parallel). The magnification of the objective lens should be measured before the experiment takes place, and can be done with a standard calibration specimen (Ted Pella Product No. 603). Using the numbers obtained, an approximation of the camera length can be found via equation 105. A basic check, which is typically done before using a camera length in a reconstruction, is to take the mask used to measure the diameter of bright disc and propagate it by the estimated camera length. The simulated bright disc is then compared to the reference data. If they are similar the estimated camera length should be close to the actual or a functional camera length. Visual signs that the simulated and reference data are close include a similar number of rings and indistinguishable ring thickness. As noted earlier at the end of section 3.3, the positions provided by the microscope and cross-correlation algorithm need an estimated pixel size in order to be functional in the reconstruction process, and therefore it is helpful that the reference diffraction pattern provides an estimate for the overall magnification of the microscope.

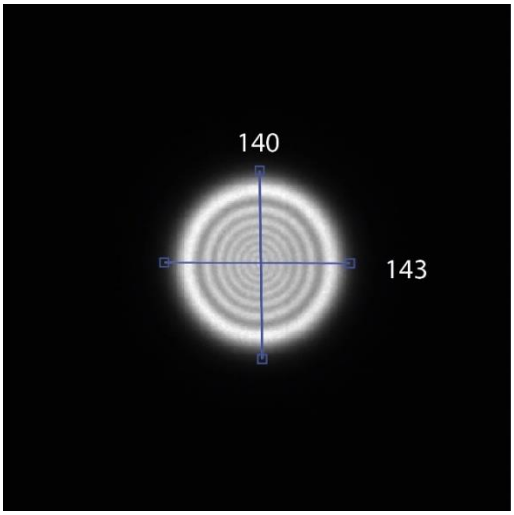

Figure 69: a figure showing how the diameter of a diffracted aperture was measured in terms of pixels using the imtool function of MATLAB.

Figure 69 shows the measured diameter of calibration diffraction pattern in pixels. Using the known pixel size at the selected area plane, which is 28nm, the magnification provided by the intermediate lens system is found to be around 392 times as follows (see equation 106):

$$M_I = \frac{10 \times 10^{-6}}{140 \times (28 \times 10^{-9})} = 392 \quad (106)$$

This means that the total magnification of the system is around 23,000 times, or the pixel size is around 1.21nm.

3.5 Fresnel scaling theorem

A possible problem that may occur during a reconstruction is that the probe, or sometimes the object will develop a phase curvature. This is in part due to Fresnel scaling theorem, mentioned earlier in section 2.2 [126]. Fresnel scaling theorem connects the parameters of camera length and magnification together; however, it also effects the scaled positions in the case of near field ptychography. As the pixel size is effectively demagnified by the Fresnel scaling theorem, scan positions must also be scaled to ensure successful reconstruction. However, the Fresnel scaling theorem can occur unintentionally, due to the fact that near field electron ptychography currently determines the camera length and the propagation distance from reference data. If any of the parameters obtained from the reference data are incorrect due to human error or experimental error, the reconstruction algorithm may cause the phase curvature in the probe reconstruction, as the reconstruction algorithm will attempt to mitigate for these incorrect parameters by converting the incorrect plane wave propagation into a correct cone beam geometry. This is not necessarily a bad development if the reconstruction is of good quality, as the probe curvature can be measured (or matched to a simulation) in order to work out the equivalent plane wave implementation of the propagation. This primarily occurs when a position correction algorithm is working alongside the rPIE algorithm, as in this situation there are two free parameters which are interlinked by the Fresnel scaling theorem; namely the positions (position correction algorithm), and the phase of the probe determined by rPIE which can convert from a plane wave geometry to a cone beam geometry.

Since detectors collect only intensity information, diffraction patterns produced by plane waves can look exactly the same as diffraction patterns produced by spherical waves, in terms of intensity, which causes great difficulty to near field electron ptychography. For example, in Figure 70 both

apertures can produce the diffraction pattern in Figure 70 a), but only one of the apertures will be physically true to the experiment. Figure 70 b) can only create Figure 70 a) if it is illuminated by a spherical wave; see the phase in Figure 70 c). Meanwhile Figure 70 d) can only create the diffraction pattern in Figure 70 a) in the case of plane wave illumination; see the flat phase in Figure 70 e). This can be identified by observing that Figure 70 b) is larger than its diffraction pattern, since during the propagation the exit wave leaving the aperture must reduce. This can only happen if its illumination is curved/spherical. This magnification/demagnification via propagation is also important to the positions in near field ptychography. If the object is de-magnified by back propagation, the positions must also be de-magnified, such that the features of the specimen correlate and do not incur a large degree of reconstruction error. Furthermore, the specimen of the experiment may have experienced curved illumination, if efforts were not made to ensure the illumination provided by the condenser lenses was parallel. More details on Fresnel scaling theorem can be found in section 2.2.

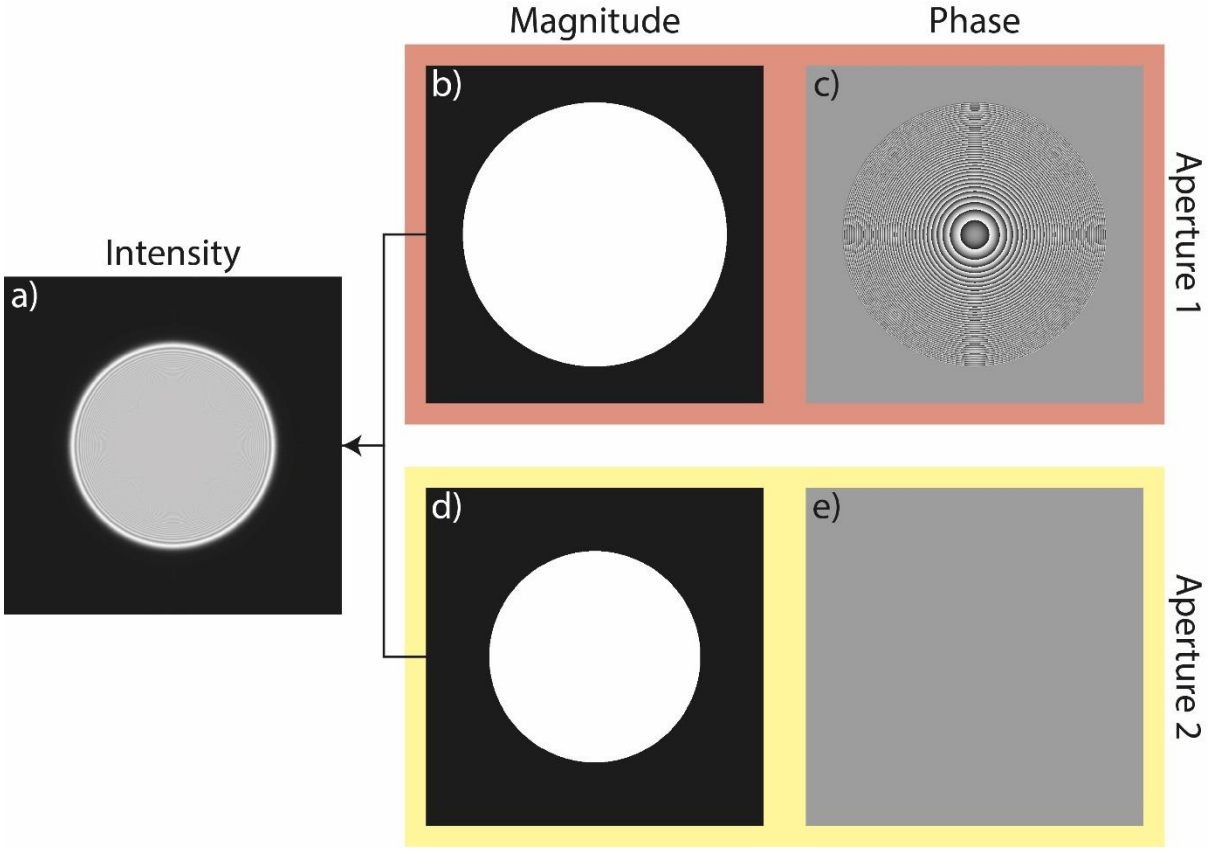

Figure 70: a) a diffraction pattern of an aperture, b) an aperture that can cause the diffraction pattern seen in a) if the illumination of the aperture was planar in nature. c) an aperture that can cause the diffraction pattern seen in a) if the illumination is spherical in nature.

The Fresnel scaling theorem can also present opportunities for near field electron ptychography, since in near field X-ray ptychography it is used to combine diffraction patterns from different diffraction conditions into one reconstruction [127]. The results of this paper showed that a specific diffuser was not required to create a diverse dataset, as diversity came from the multiple diffraction conditions. If near field electron ptychography was implemented in this way, it would become very similar to inline holography, where multiple diffractions conditions have been used before to constrain the phase problem [54].

3.6 Diffraction pattern drift during data collection

During long experiments, one additional factor that has affected our data collection is a slow drift of the diffraction pattern across the detector. An example of diffraction pattern drift is shown in Figure 71, which shows the first diffraction pattern of a dataset and the last diffraction pattern of the same dataset. The diffraction pattern in Figure 71 b) has drifted 113 pixels downwards and 79 pixels to the left with respect to the diffraction pattern shown in Figure 71 a), over a period of 20 minutes. This means that the average effective drift rate of the optical axis for this dataset is 0.1112 pixels per second (combining the *x* and *y* drifts). Although the drift rate is not constant throughout the experiment, by measuring the drift of every diffraction pattern via cross-correlation, the drift can be plotted as a function of time (as it takes 3 seconds to collect each diffraction pattern). This is shown in Figure 72, and this figure suggests that the rate of drift decreases as the experiment progresses (note that the pixel size is 1.18nm).

It is typical that the microscope specimen stage has a certain degree of drift, as it is difficult to create mechanical and electrical systems which are stable to a sub-nanometre degree. The typical drift of a specimen stage of S/TEM would be around 0.5-1nm min^{-1} [138-141]; the values of drift experienced here are above what would be expected due to the stage alone. In the STEM context, work has been done in order to reduce the problems caused by drift with postprocessing steps such as those outlined by Ophus et al. [142] and Jones et al. [143]. There are many causes of optical axis drift in the S/TEM [144], for example the specimen could experience uneven heating or charging by the beam, leading the specimen to move in a particular direction. The room the TEM is placed in could be a second factor, for example the temperature generated by the operator could be a factor as it would increase the temperature of the room and therefore alter the current of the objective lens, changing its strength. There is also the important point that near field electron ptychography operates in a non-standard configuration, meaning that instabilities might be magnified by this particular setup. Furthermore, the typical methods used to ensure that the microscope is stable may need to be altered for this particular mode of operation. All of the above factors should be considered in future work, when trying to reduce the drift of the diffraction patterns to standard levels.

Optical axis drift has been experienced before in electron implementations of ptychography [145] and selected area ptychography [94]. Maiden et al. suggested using different intermediate lens settings to explore whether a particular setting exacerbated the optical axis drift. It is also worth noting that drift is problematic for other forms of microscopy in the TEM, for example energy filtered transmission electron microscopy, where the purpose is to collect different energy filtered images of the specimen, to capture the variation of energy loss over an area, which enables the identification (and density measurement) of particular elements [146]. Due to the low number of counts per second of EFTEM, extended exposure times are required, allowing for significant drift of the optical axis. EFTEM overcomes the "spatial drift" of the specimen by dividing a long exposure into many smaller ones and cross-correlating the data post experiment [147]. A similar method of multiple exposures is used in off-axis holography in order to obtain high phase sensitivity phase images [148].

Diffraction pattern drift is a limiting factor on the experiment time and therefore the FOV (field of view) of near field electron ptychography, as the experiment must be stopped once the diffraction pattern drifts out of the field of view of the detector. Furthermore, diffraction pattern drift alone can make a dataset "unreconstructable". This is illustrated by Figure 73, where a) and c) show a case where drift has not been accounted for while, in b) and d) the reconstruction code has been altered to counter the drift of the dataset. Figure 73 only shows the reconstructions after ten iterations, as

the reconstruction which has its drift unaccounted for would be completely unrecognisable after a few more iterations.

This subsection will present the current hypothesis on the cause of diffraction pattern drift, a method to reduce the diffraction pattern drift, and an outline of how it is accounted for in the reconstruction code.

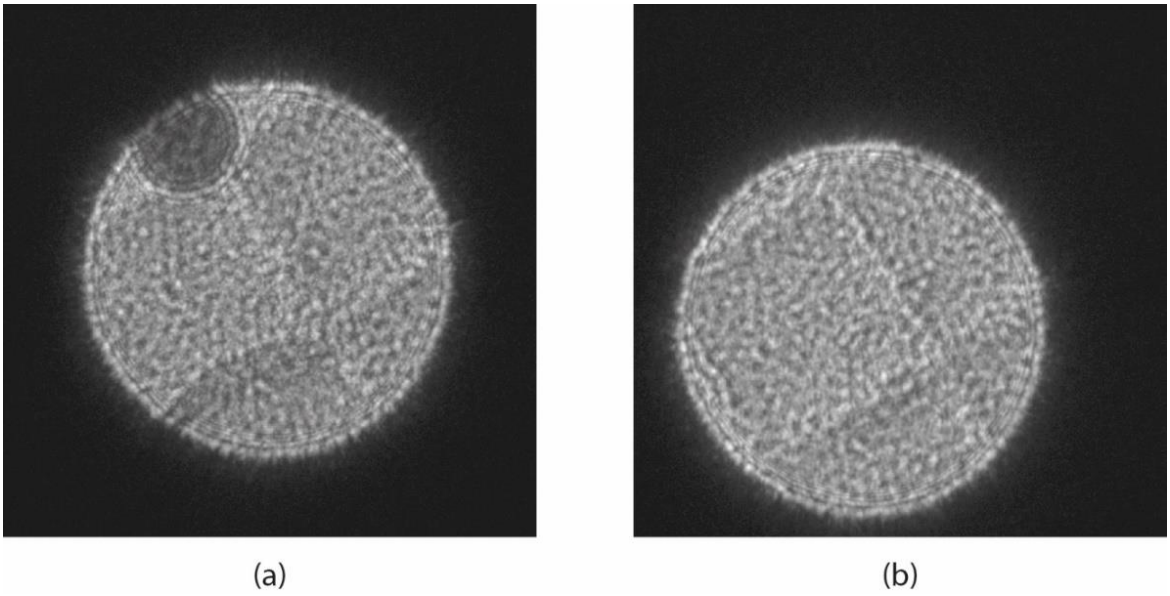

Figure 71: An example of diffraction pattern drift, where a) is the first diffraction pattern in a dataset and b) is the last diffraction pattern in the same dataset.

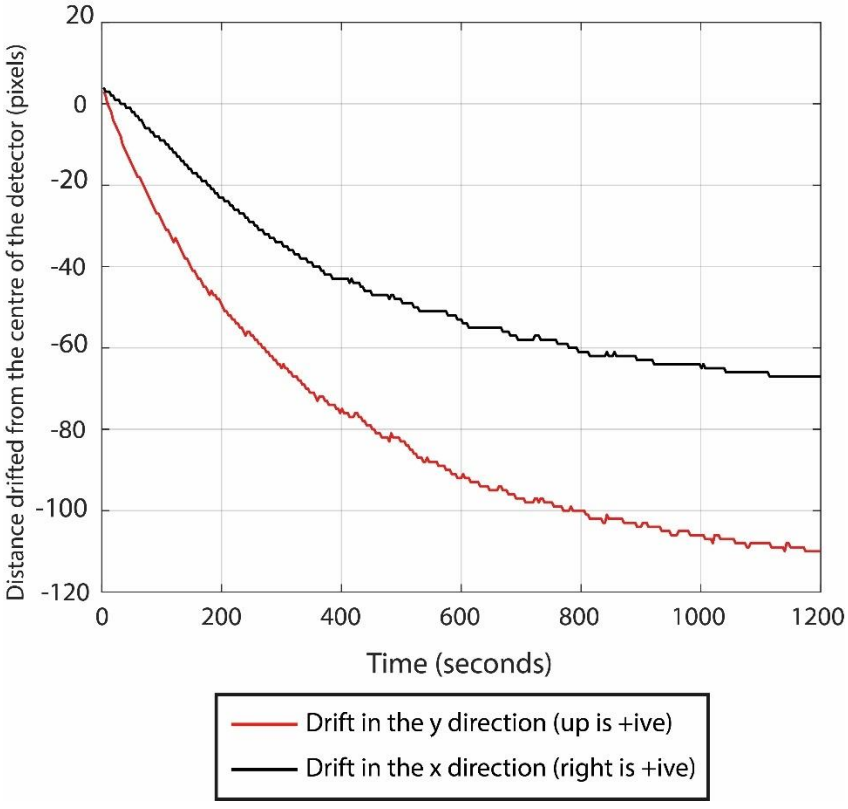

Figure 72: A plot of diffraction pattern drift as function of time of an example near field ptychographic dataset collected on a FEI Titan 80-300 STEM operating in uncorrected TEM diffraction mode (see section 4.4 for more details).

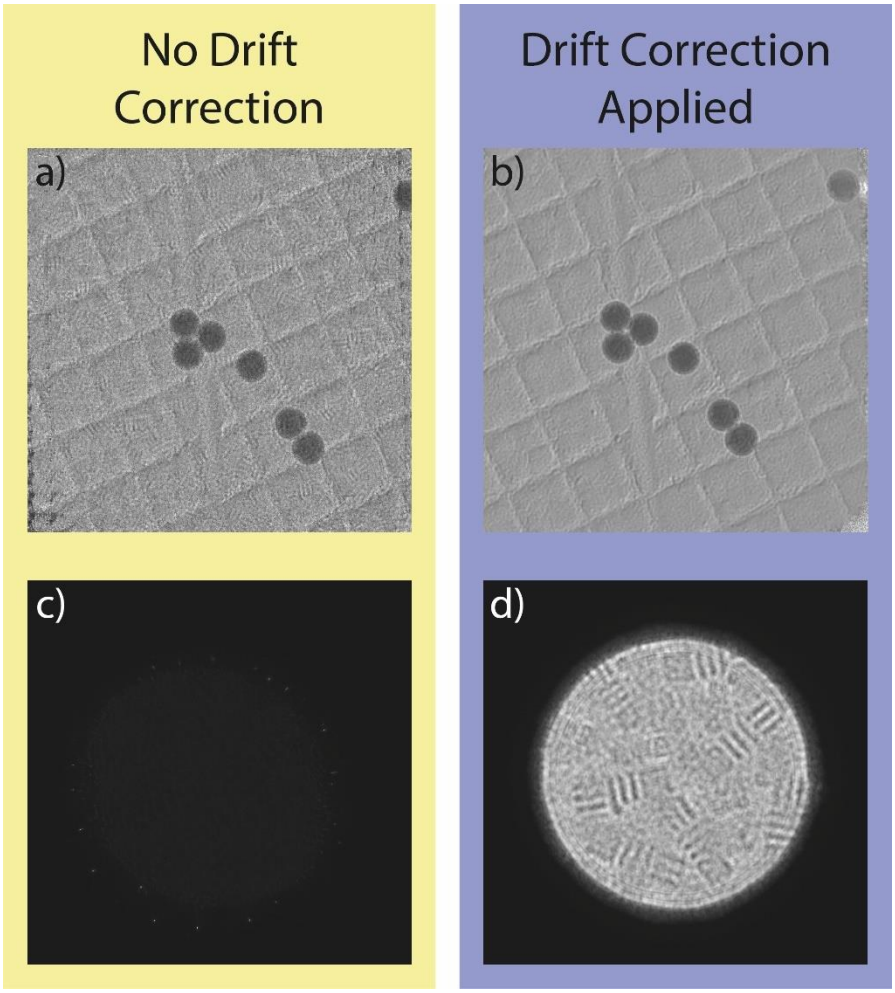

Figure 73: a) magnitude reconstruction of the specimen without drift correction, b) magnitude reconstruction of the specimen with drift correction. c) probe reconstruction without drift correction, d) probe reconstruction with drift correction. All of the reconstructions shown here are from ten iterations of rPIE code.

The practical solution to the drift was to give the microscope time to settle. Once the diffraction condition required by near field electron ptychography was achieved, the diffract pattern on the detector would drift significantly after it was centred and it would leave the side of the detector after 20 mins. However, by using the image scan coils to repeatedly correct the diffraction pattern drift, the drift rate would be reduced but not eliminated after 2 minutes of continuous correction. The reduction in drift rate is shown in Figure 74, which displays the measured diffraction pattern drift of the 240 diffraction pattern dataset after taking drift mitigating measures (described above). The drift was only 27 pixels to the right and 23 pixels downwards in the 240 diffraction pattern dataset, and correspondingly the total drift rate was 0.0462 pixels per second, suggesting the drift mitigating measures are having the desired effect. However, it could be suggested that the drift decays naturally, as shown by both Figure 72 and Figure 74, so further work includes finding a more active measure by which to reduce the drift rate. The fact that drift has not currently been eliminated possibly limits the ability of near field electron ptychography to retrieve quantitative phase information, as shifts of the diffraction pattern can also be the result of a low frequency component of the specimen whose phase we wish to image. The term “shift” is used to delineate the difference between a diffraction pattern which has moved due to an imperfect setup, and a diffraction pattern which has moved due to the phase of the specimen, as ultimately, they both cause the diffraction pattern to move and therefore cannot be separated.

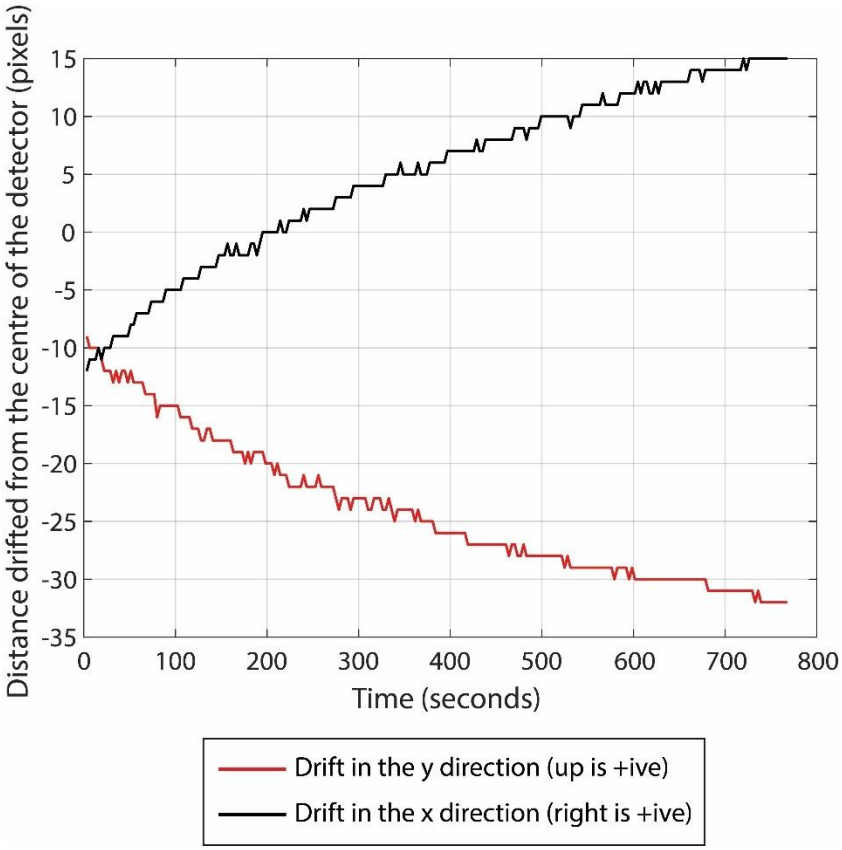

Figure 74: a figure showing the diffraction pattern drift as a function of time after repeatedly using the scan coils to centre the diffraction pattern. This figure can be compared to Figure 72 where no drift mitigation steps were taken. The data used in this plot were collected on a FEI Titan 80-300 STEM operating in uncorrected TEM diffraction mode (see section 4.4 for more details).

In the following section, an algorithmic method to correct drift is described. This correction step only works if the diffraction pattern remains within the field of view of the detector, so an important discovery was finding out that the image scan coils can be used mid-experiment to recentre the diffraction pattern. Although this solution is not ideal, it has not caused an obvious degradation of the reconstruction quality when used. For example, Figure 75 shows a reconstruction where the diffraction pattern drift was changed mid-experiment, and Figure 76 displays the measured drift of the diffraction patterns in the dataset used in Figure 75. This enables near field electron ptychography to cover an extremely large FOV, where data collection can take over an hour (for example the large field of view dataset in Section 4.8 could have not been collected without this discovery). Figure 75 shows the phase of the latex spheres and also shows the first instance of phase wrapping; phase unwrapping is discussed later in subsection 3.7.

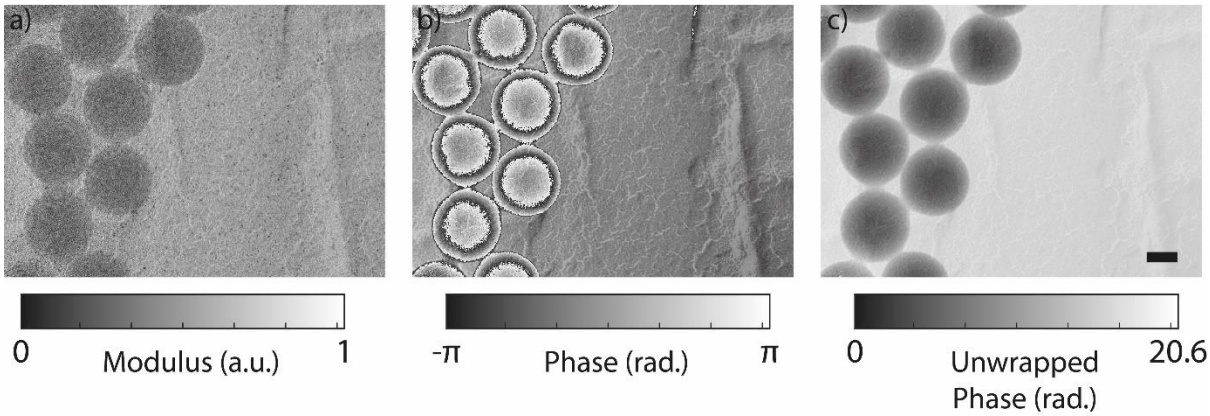

Figure 75: Reconstruction results where the image scan coils were used to shift the diffraction pattern mid-way through an experiment to demonstrate that this has no adverse effect on the reconstruction quality. a) is the magnitude reconstruction and b) phase reconstruction. c) Unwrapped version of the phase displayed in b). The scale bar in c) is 200nm long (applies to all panels). The colour bars under the respective figures convert the greyscale into modulus, wrapped phase and unwrapped phase accordingly.

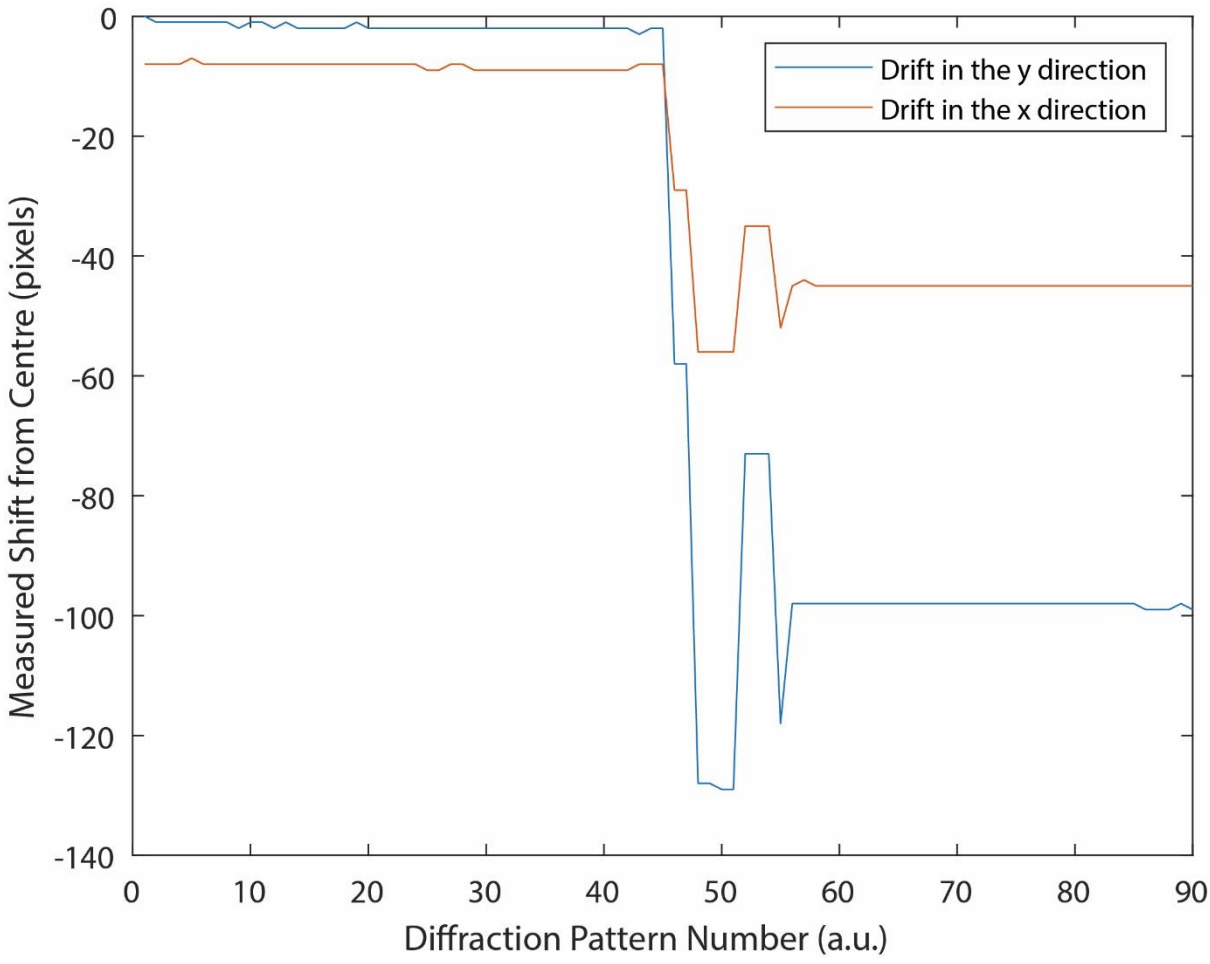

Figure 76:A plot showing how the diffraction pattern drift varied with time. Each diffraction pattern on average took three seconds to collect.

Now the practical steps to mitigate the diffraction pattern drift have been discussed, the drift must also be dealt with algorithmically, as the drift can have a significant effect on the forward model of near field ptychography, leading to incorrect spatial assignation of the phase and magnitude in the probe. This incorrect assignment itself causes the probe reconstruction to collapse and to be a

function of the average position of the diffraction pattern. Figure 77 illustrates the issue that a diffraction pattern drift causes to the reconstruction algorithm, starting from the right of Figure 77 c) which shows the modelled diffraction pattern obtained from the reconstruction algorithm (note it is in the centre of the calculation window and is denoted by $\Psi_{j,n}$ in Figure 77 d)). Figure 77 b) shows the collected diffraction pattern which has drifted from the optical axis denoted by I_j in Figure 77 d). As part of the intensity update of the rPIE algorithm (see section 1.14), the modelled intensity is replaced by the measured intensity as displayed in Figure 77 d). The misalignment between b) and c) causes the incorrect assignment of phase, and also causes the updated magnitude to be highly erroneous, as displayed by Figure 77 a). Figure 77 a) has two clearly incorrect aspects: 1. features of the diffuser now occur twice, as can be seen by inspecting the bar-like features of a), 2. the updated intensity (Φ) is no longer circular like the actual diffuser but is instead oval. The oval error is due to the multiplication of the modelled and measured intensities, which both include zeros (or values far less than 0.1). As the zeros are misaligned, an oval shape occurs and the updated model is dissimilar to the resemblance of the collected or modelled diffraction pattern. Therefore, Figure 77 illustrates why the reconstruction process breaks down (see Figure 70) when the drift of the diffraction patterns is not addressed algorithmically.

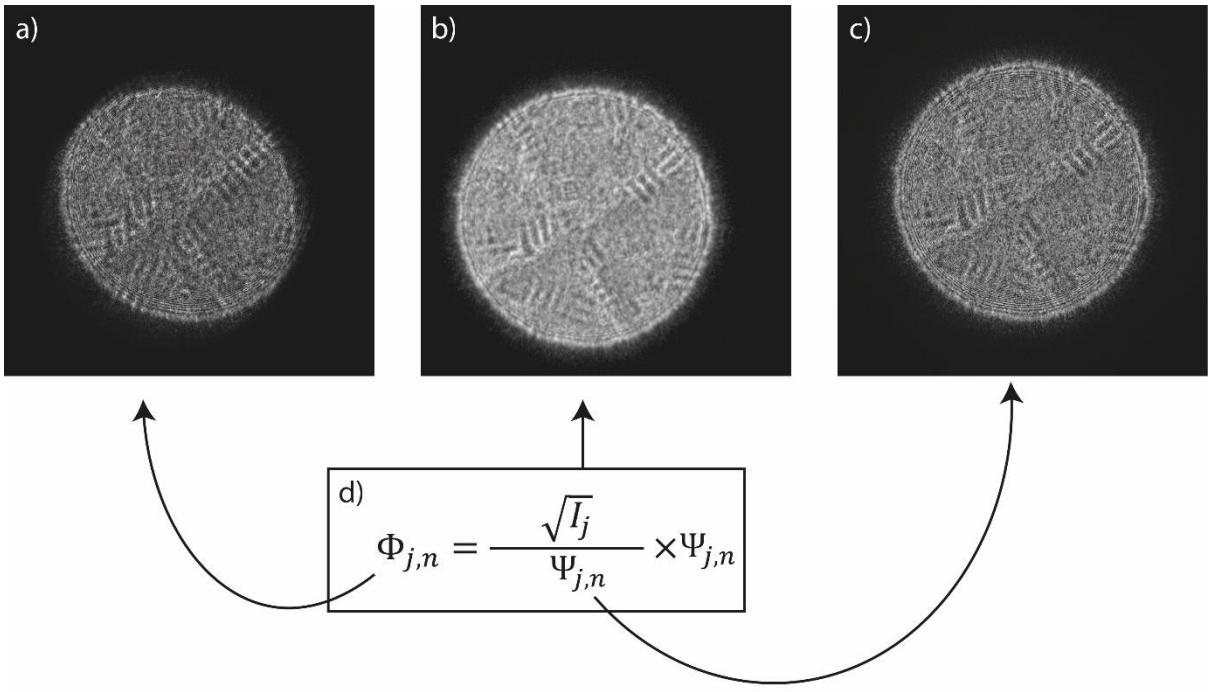

Figure 77: A figure showing how the drift of the diffraction pattern causes intensity update to have an adverse effect on the reconstruction algorithm. a) the magnitude of the adjusted exit wave obtained by replacing the modelled magnitude by the square root of the collected intensity. b) the square root of the collected intensity. c) the magnitude of the guess at the exit wave before it was updated.

The parameters of Figure 77 d) are the following:

- φ is the adjusted exit wave
- Ψ is the guess at the exit wave
- D_z is the propagation method which best fits the experiment
- I is the collected diffraction pattern from the experiment

This problem is largely solved through the deployment of a cross-correlation between the simulated diffraction patterns and the off-centre data to measure the shift between the two, as outlined in supplementary information of [94] by Maiden et al. Given this measurement, the collected data are circularly shifted to align with the simulation and the shift value is stored for later review. Subpixel shifting is typically not deployed in the reconstructions mentioned in this paper, as subpixel alignment is an intensive GPU/CPU calculation, while the current alignment technique can be performed in every iteration for every diffraction pattern without a significant effect on the reconstruction time. Furthermore, performing a cross-correlation in every iteration enables the cross-correlation to be improved if the initial model of the diffraction pattern is not closely matched to the collected data. When processing diffraction patterns which have drifted an additional change has to be made to the reconstruction code, both the simulated diffraction patterns and collected data should be Fourier transformed before the reconstruction error is measured. This alteration prevents the error being unrepresentative because the diffraction patterns are not aligned perfectly with the simulated diffraction patterns. Measuring the error in Fourier space enables the position correction code to function correctly when the dataset suffers from diffraction pattern drift.

3.7 Measurement of mean inner potential of latex spheres

Throughout the results section (see section 4) the mean inner potential of latex spheres will be used to verify that near field electron ptychography is a quantitative technique. This verification process involves taking the phase image of the latex spheres, converting it to mean inner potential, and comparing the results to previous experiments. The process followed here is based on the method of obtaining the mean inner potential used in the supplementary material [94]. The least squares method is used to fit the measured phase delay to the thickness of the latex sphere; the process by which the data is prepared for the fitting process is described, but not the least squares fitting process. Once a good fit is found, the gradient of the fitted curve gives the estimated mean inner potential of a latex sphere.

The specimen used in section 4 was a carbon diffraction grating replica, populated with latex spheres (Ted Pella Product No. 603). This sample is primarily used in experiments to measure the magnification of the TEM at particular settings. The latex spheres have a nominal diameter of 261nm. The measurement of the mean inner potential of latex spheres has been performed before, and results can be found in the following references [149, 150]. These prior works suggest the expected value for the mean inner potential is 8.5V with a variance of 0.7V. Now that the specimen and the expected value have been presented, this leads to the question of how near field electron ptychography images are processed to measure the mean inner potential of the latex spheres.

To measure the mean inner potential of the latex spheres imaged by near field electron ptychography, the equation which relates the thickness of the specimen to its imparted phase delay must be considered (see equation 107):

$$\Delta\varphi = tV_{MIP}C_E \tag{107}$$

The parameters of equation 107 are the following:

- $\Delta\varphi$ is the change in phase of the latex sphere in the reconstruction
- t is the modelled thickness of the latex sphere
- V_{MIP} is the measured mean inner potential in volts
- C_E is the interaction constant which depends on the wavelength of the illumination used [111]. The equation for calculating C_E is equation 108 (6.52×10^6 rad eV^{-1} m^{-1} at $300\mathrm{keV}$).

Below is the equation for calculating the interaction constant C_E:

$$C_E = \frac{2\pi e}{\lambda E} \times \frac{E_0 + E}{2E_0 + E} = \frac{2\pi(n-1)}{V_{MIP}\lambda} \tag{108}$$

The parameters of equation 108 are the following:

- E_0 is the rest energy of an electron
- E is the kinetic energy of the incident electrons, also $E = eU_{acc}$ (U_{acc} is the acceleration voltage 300 keV)
- e is the charge of an electron $(-1.602 \times 10^{-19}\ \mathrm{C})$
- λ is the wavelength of the electron wave
- n is the index of refraction of the material (in this case $\mathrm{Si_3N_4}$)

It is clear from looking at equation 107 that the thickness needs to be obtained before the mean inner potential is calculated. All the processing above is dealt with using a MATLAB home-written algorithm which is described by Figure 78 and Figure 79. To be clear, Figure 78 and Figure 79 describe the same function; the two figures are intended to make it easier to read.

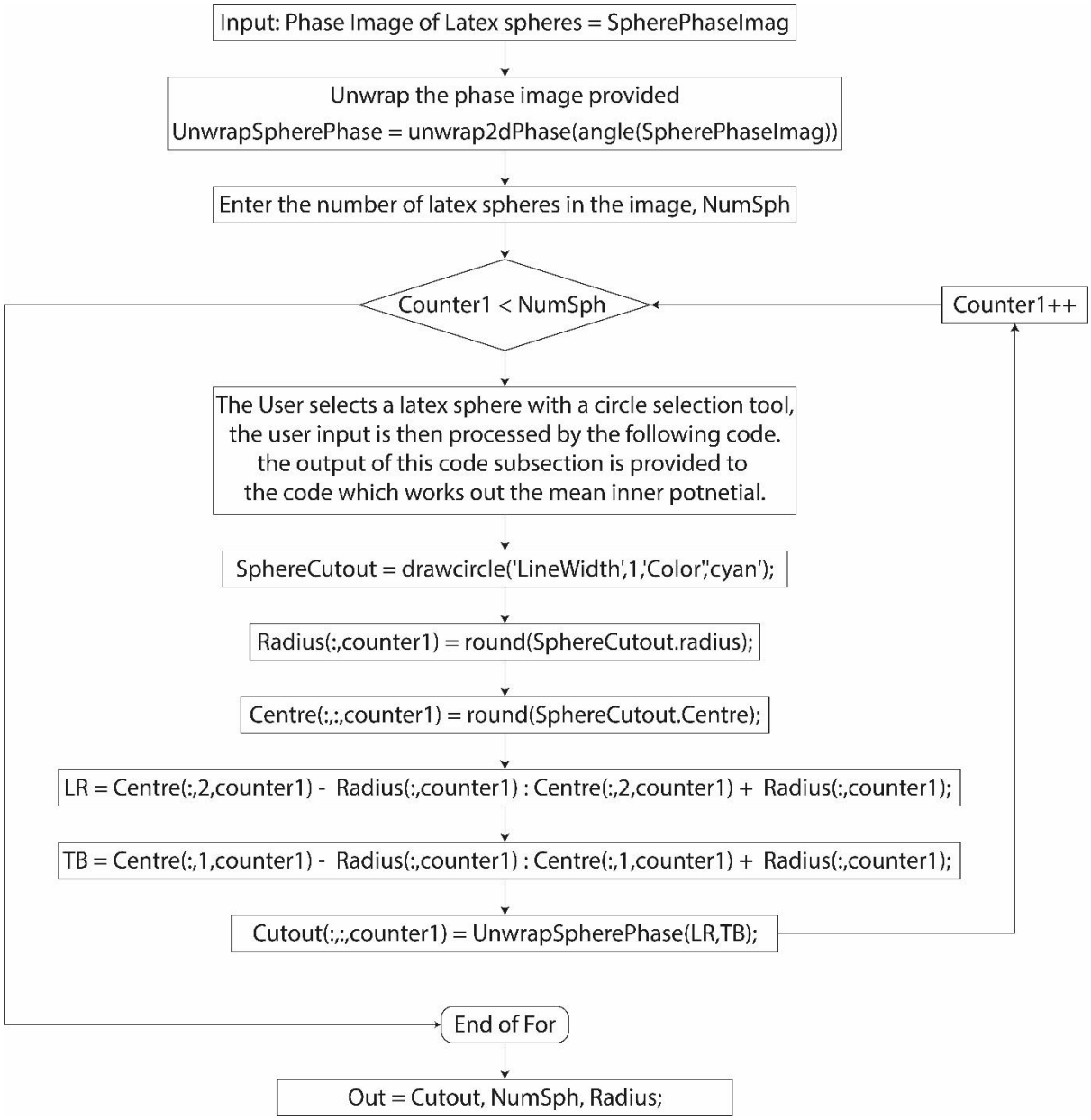

Figure 78: a flow diagram describing the process by which the user takes their phase image of latex spheres, enters the number of spheres they wish to measure and then selects them individually with the draw circle function. The output of the draw circle function is then used to cut out subsections of the image to be used in Figure 79. During this process the radius of the latex spheres is measured in terms of pixels such that it can be used to model the thickness of the latex sphere later in Figure 79.

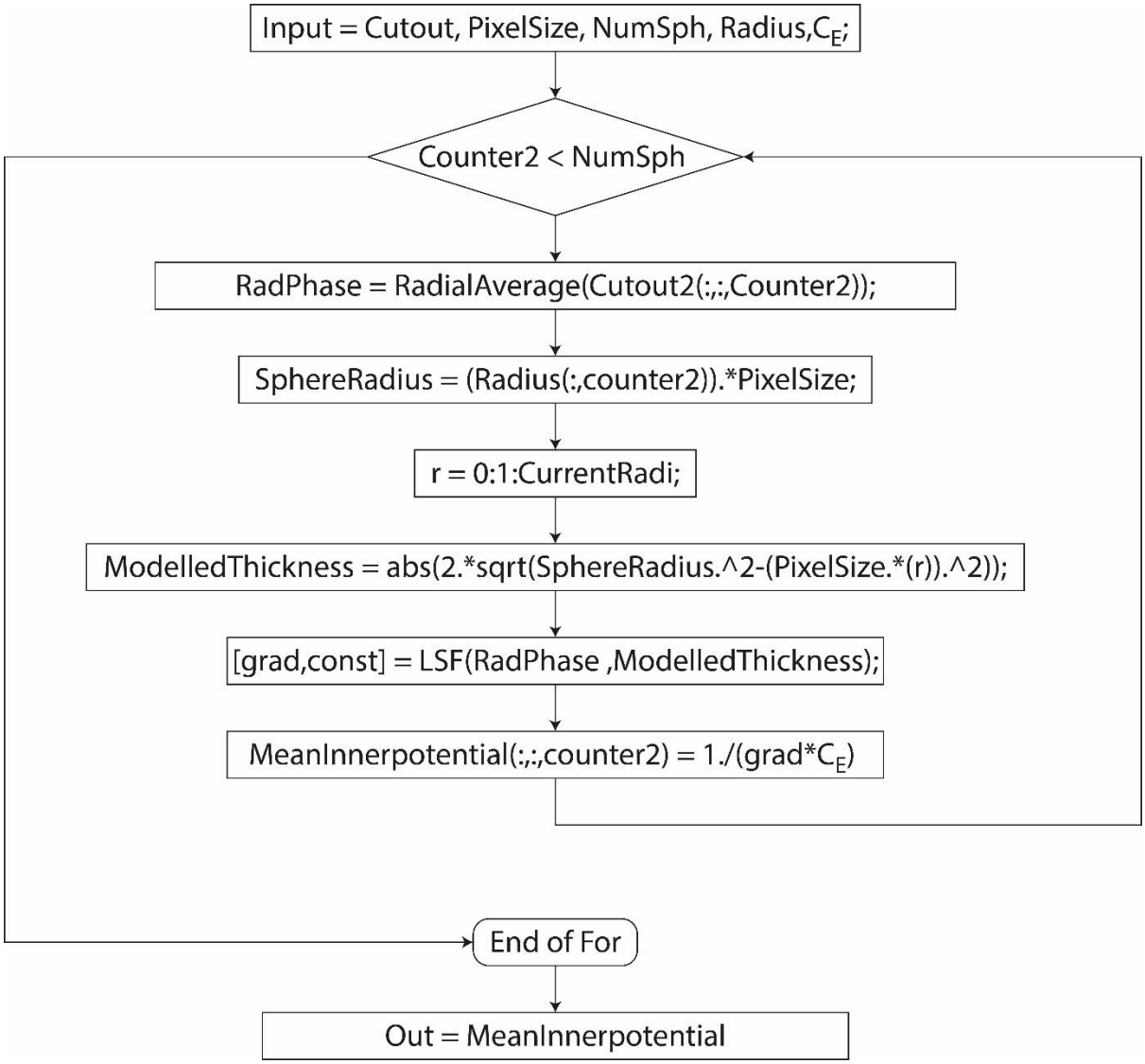

Figure 79: a flow diagram where the latex sphere subsections from Figure 78 are further processed. This starts with taking the radial average of the phase of the subsections/cutouts. Then the measured latex sphere radius is converted from pixels to nanometres by multiplying by the user's inputted pixel size. The latex sphere radius is then used to model the thickness of the sphere. The modelled sphere thickness and radially averaged phase is then inputted into the least squares function. The least squares function finds the gradient which best fits the radial averaged phase to the modelled thickness, and this gradient is then converted into the mean inner potential of latex.

The first step taken in Figure 78 is to unwrap the 2D phase of the reconstruction via the Goldstein method as will be described in section 3.8; this is important for a future step in Figure 79. The second step is for the user to enter the number of latex spheres to be considered, such that the "for" loops in the function can be initialised correctly. Within the first "for" loop of the function the radii of the latex spheres are measured such that a correct model of their thickness can be obtained.

The thickness of the latex spheres is measured in terms of pixels by the user drawing a circle (a built-in MATLAB function: drawcircle) around one of the latex spheres. The diameter of the circle is given as part of the structure called SphereCutout, which also contains the centre of the latex sphere in the image array. The diameters and centres are stored such that they can be further processed, as in Figure 79. The radius and centre values are also used to select a square subsection around the latex sphere to reduce the calculation time of later functions (as they do not have to process the entire image). These subsections are stored in Cutout.

The input to the second part of the function (Figure 79) includes: cut out of the latex spheres, pixel size of the demagnified detector, the number of latex spheres to be considered, the radius of each latex sphere in pixels, and the interaction constant of electrons. The first step of the process described by Figure 79 is to calculate the radial average of the latex sphere cutouts, which reduces the phase delay from being a function of *x* and *y* to being a function of radius. In contrast, previous papers have considered each pixel of the image of the latex sphere individually, to calculate the mean inner potential [149]. An advantage of radial average method is that it acts to average out any noise the phase delay of the sphere has. Radial averaging requires that the centre is chosen correctly, so that the radial average is reported without error to the rest of the algorithm. Most importantly, the measured phase delay is now simply a function of the distance from the centre of the latex sphere (s_c). The next step is to process the thickness of the latex sphere into the same form.

The reported radius (obtained in Figure 78) in pixels is now used in tandem with the known pixel size at the specimen plane to measure the radius of the latex spheres in nanometres. It is assumed that the latex sphere has the same radius in all dimensions, so its maximum thickness is inferred to be twice its measured radius. Using this assumption, the thickness of the latex sphere is modelled using equation 109:

$$t(r) = 2\sqrt{R^2 - r^2} \tag{109}$$

The parameters of equation 109 are the following:

- $t(r)$ is the modelled thickness as a function of distance from the centre of the latex sphere
- R is the measured radius of the sphere
- r is the distance from the point where the sphere has its maximum thickness

Now that a model of the thickness as a function of the distance from the centre has been obtained, this can be used as reference for finding the mean inner potential of the latex sphere. Rearranging equation 107 such that the thickness of the latex spheres is the output of the equation:

$$t(r) = \frac{\Delta\varphi(r)}{C_E V_{MIP}} \tag{110}$$

The parameters of equation 110 are the following:

- $t(r)$ is the modelled thickness as a function of distance from the centre of the latex sphere
- V_{MIP} is the measured mean inner potential in volts
- C_E is interaction constant which depends on the wavelength of the illumination used [111]
- $\Delta\varphi(r)$ is the radially averaged phase of the latex sphere being considered

Now there are essentially two ways of obtaining the thickness: one from measuring the latex sphere, another from guessing the mean inner potential of the latex sphere. This is the ideal situation in which to deploy least squares fitting, as by measuring the error between the values of thickness we can improve our guesses of the mean inner potential of the latex spheres. Inputting the measured radial phase and measured latex sphere radius into the least squares fitting, the gradient and offset of the fitted curve is obtained. The gradient can then be processed to obtain the mean inner potential by multiplying the gradient by the interaction constant and then taking the inverse of the result.

There are several things to note: one is to check that the specimen is in focus before attempting to measure the mean inner potential of latex spheres, as the diameter of the latex spheres varies as the

specimen is propagated in and out of focus. Even small changes in the propagation distance (0.05m) can cause the radius to change by several pixels. Secondly, there is an assumption that needs to be discussed: despite any changes to the size of the latex spheres due to the beam, it is assumed that the latex spheres are reduced in size equally in all dimensions (i.e. their roundness is unchanged). This assumption is made such that the thickness profile is easier to model, and currently no papers suggest that latex spheres reduce in a non-uniform manner. Lastly it should be noted that the method by which the latex sphere radius is measured adds human error in the calculation. Therefore, future work regarding the latex sphere measurement includes improving MIP measurement software to include automatic sphere detection and radius measurement (or something to reduce the amount of human error). The mean inner potential recovered by the process described above is negative, therefore the unwrapped phase image is sometimes inverted to obtain a positive mean inner potential, see Figure 85 and Figure 97.

3.8 Phase unwrapping - Goldstein method

Phase unwrapping is the process by which the signal that is limited in value between set boundaries is released from those boundaries via analysis of the signal's rate of change. In the situation encountered in this thesis, the phase is effectively set within the boundaries of zero and two π (or $\pm\pi$). There is a rich literature describing methods to solve the phase unwrapping problem, however only a few examples will be mentioned: for example the Transport of Intensity phase unwrapping algorithm [151], which uses some of the concepts highlighted in Section 1.12, there is the least squares method [152], and the branch and cut algorithms [153].

The choice of phase unwrapping algorithm is especially important in the context of assessing the quantitative phase of near field electron ptychography, as phase unwrapping methods may incur side effects in the unwrapping process. For example, the least squares phase unwrapping algorithms can cause the dynamic range of the unwrapped phase to be reduced [154], whilst others effectively apply a low pass filter to the unwrapped phase. These side effects could then cause the fitted phase curves in the case of the latex spheres to be a closer or poorer match to the spherical model than they actually are. Therefore, in this thesis the Goldstein phase unwrapping algorithm, an early branch and cut algorithm, was chosen due to its simplicity and its ability to recover the original phase without alteration [155]. The theory of the Goldstein phase unwrapping algorithm will now be briefly outlined.

Goldstein's work is primarily based on the work of Itoh [156] who developed methods to unwrap the phase in the one-dimensional case. Goldstein then extended this to the two dimensional case. Itoh's process consists of comparing the current phase value to its previous value and determining whether the change is greater than π in magnitude, then the sign of the change is used to determine whether 2π should be subtracted or added to the present value of phase (see equations 111 to 113). Through this process an unwrapped one dimensional signal is obtained.

$$\Delta\theta = \theta(x) - \theta(x-1) \tag{111}$$

$$\text{if } \Delta\theta > \pi, \text{ then subtract } 2\pi \text{ to } \theta(x) \tag{112}$$

$$\text{if } \Delta\theta < -\pi, \text{ then add } 2\pi \text{ to } \theta(x) \tag{113}$$

However, Goldstein identified a possible error which could occur in the two dimensional case whereby a phase residue could cause the unwrapped phase to be inconsistent, depending on the order in which the phase is unwrapped. A phase residue in the phase unwrapping context can be understood as four values in a square. An example can be seen in Figure 80 a) by applying Itoh's test in a clockwise direction (starting at the bottom right) the phase is unwrapped: -0.5π to -1.2π there is

no phase wrapping, -1.2π to π a phase wrapping has occurred according to Itoh, therefore we subtract 2 π from π to get -π, with the newly adjusted -π to 0.3π another phase wrap has occurred, 0.3π is then updated to -1.7π via subtraction of 2π. However, this now means that a phase wrap is occurring between the starting and ending point, and following the same process in this loop an infinite phase would be obtained. Figure 80 a) represents a negative phase residue while b) represents a successful phase unwrapping. Goldstein's discovery was that phase could be unwrapped successfully if there was an equal number of positive and negative phase residues within the integral area; the phase unwrapping would be consistent no matter where it started. A second condition was that positive and negative singularities had to be connected by so-called branches (hence the 'branch' in branch cut algorithm), these branches could not be crossed by the unwrapping algorithm in order that the unwrapped phase was kept consistent, regardless of the starting point.

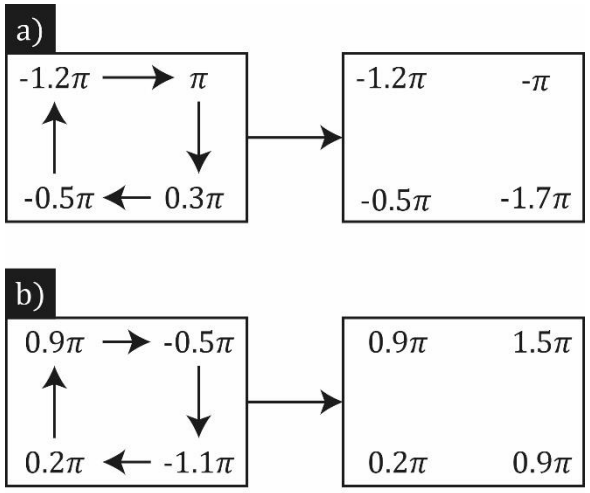

Figure 80: A figure illustrating the process by which to identify phase singularities in two dimensional phase images. a) shows a negative phase singularity before and after it is unwrapped. b) shows an example of non-singularity phase unwrapping. The arrows between the boxes denote the action of the phase unwrapping algorithm.

Goldstein's phase unwrapping algorithm consists of two main steps: identifying and connecting phase residues via the shortest possible path, and unwrapping the phase via Itoh's method in a way which avoids crossing of the branches connecting the phase residues. Phase residues can be found by searching a nine-pixel square and checking each combination of four adjacent pixels for phase residues, as was achieved in Figure 80. Once a phase residue is found, the search algorithm then searches for residues of the opposite sign in neighbouring pixels, and if none are found in the immediate vicinity the search radius is increased until one is found or the edge of the image is reached. The edge of the image acts as a residue of both signs for the sake of residue balancing.

The Goldstein phase unwrapping algorithm can experience problems if the branch cuts isolate a particular region of the reconstruction. This then means the isolated area has to be unwrapped on its own and the result will likely be disconnected from the rest of the image. Furthermore, Goldstein's method struggles to operate successfully when operating with images with low signal to noise ratios, as the number of branch cuts will be high, leaving most of the phase wrapped. The version of the Goldstein algorithm used in this thesis can be found on the MATLAB file exchange (https://ww2.mathworks.cn/matlabcentral/fileexchange/22504-2d-phase-unwrapping-algorithms, written by Bruce Spottiswoode).

3.9 Summary of the data collection process of near field electron ptychography

The main contribution of this section is setting up the code and the tools required in order to collect near field electron ptychography data using the TEM. The data collection code was set out in flow diagrams enabling the code to be rewritten in any other programming language and the code used to collect quantitative phase images in Section 4 is provided in the appendix (see section 7), therefore near field electron ptychography is open access to anyone with this thesis and a TEM. Furthermore, this code and post processing steps work for selected area ptychography as well as near field electron ptychography, meaning that there is no barrier to implementing SAP in the TEM (although a diffuser is still required for near field electron ptychography). The post processing steps were outlined: focusing on the accurate measurement of scan positions via cross-correlation of the diffraction patterns and measurement of the propagation distance was achieved by matching a simulated propagated aperture to collected reference data. Possible issues were addressed, such as drift of the diffraction pattern as the experiment progresses, this was discussed both in terms of practical reduction and negative algorithmic effects. The Fresnel scaling theorem which can lead to reconstructions with ambiguous phase curvature in the case of incorrectly scaled positions and incorrect propagation distance was outlined. The work presented also shows that near field electron ptychography is undemanding in terms of its implementation, only requiring the synchronisation of the movement of the specimen and detector captures. This is encouraging for future work, in which it would be desirable to implement tomographic near field electron ptychography. Tomographic near field ptychography has already been performed in the X-ray domain [82, 157, 158] where it has demonstrated the required sensitivity and fields of view to provide useful results.

Without discovering the methods in this section to mitigate the drift of the diffraction patterns, it would not be possible to implement the extreme fields of view experiment in Section 4.8. Therefore, the work presented here is important as it helps demonstrate near field electron ptychography's greater field of view per diffraction pattern. Moreover, without the ability to use the collected diffraction patterns to determine the scan positions, the high magnification dataset (see Section 4.10) could not be correctly processed. The high magnification section also demonstrates the full utilisation of the detector, a key goal in achieving the greatest fields of view per diffraction pattern (for a set magnification). The start of the next section will begin with some discussion of the diffusers used in the TEM to perform near field electron ptychography.

Chapter 4

4 Near field electron ptychography results

In this section near field electron ptychography will be performed for the first time. The motivation for implementing near field ptychography in the electron domain comes not from obtaining the highest resolution images such as its focused probe counterpart [159], but instead from a desire to obtain megapixel images (2D or 3D) with nanometre resolutions in an as efficient method as possible. Efficiency for near field electron ptychography could mean the least amount of experiment time or the fewest number of diffraction patterns, or possibly the smallest total dose. This efficiency has already been demonstrated in the X-ray domain where near field ptychography was initially developed, as Zhang et al. [160] suggest that near field X-ray ptychography occupies a particular niche in terms of being able to image 100µm² field of view in the shortest time, whilst retaining a resolution in the tens of nanometres. There are many applications that sit above sub nanometre resolution to which near field electron ptychography can theoretically be applied to, for example: quantitative tomographic phase imaging [82, 161], biological samples [92, 162] and investigation of semiconductor devices [163]. However, as near field electron ptychography is being implemented for the first time in this work, it has not been sufficiently developed such that it can demonstrate its full potential in its current form. Therefore, in this section a balance is struck between illustrating near field electron ptychography's strengths and highlighting challenges to be addressed in future work.

Now that near field electron ptychography data collection code has been discussed it will be utilised in this section. For example, the data collection code will be tested in terms of how little data is required to constrain the phase of both the specimen and diffuser, and it will be tested in terms of what is the largest dataset that can be reconstructed successfully. These two tests will demonstrate that near field electron ptychography can obtain reasonable fields of view in 27s seconds of experiment time and obtain mega-pixel images of the specimen with approximately a 100µm² field of view. Furthermore, the quantitative nature of near field electron ptychography will be tested via the comparison of the recovered phase against the modelled thickness of latex spheres (the modelled thickness is based on their measured diameter). This result is vital as it will show that near field electron ptychography is quantitative and therefore in principle can be used in the application of other phase sensitive techniques such as off-axis holography, to fill the same role or provide a better alternative. Furthermore, the importance of the diffuser will be demonstrated in this section, since it is the most efficient manner by which the interference fringes of the specimen and illumination can be made to interact with one another. The importance of the diffuser will be highlighted by its replacement with an aperture in otherwise identical diffraction conditions, and then comparisons will be made between diffuser and aperture reconstructions. Then the effect of energy filtering on the reconstructions will be assessed by varying the width of the energy filtering slit and comparison to a non-energy filtered dataset. The effect of energy filtering will be particularly interesting as there are two sources of inelastic scatter, that of the specimen and the silicon nitride diffuser, so it is suspected that inclusion of energy filtering would be vital to the successful operation of near field electron ptychography. However, the result presented in this thesis seems to contradict this expectation. It will also be shown that non-energy filtered near field electron ptychography can

achieve visually similar results to its energy filtered counterpart via the inclusion of background modelling. Then there will be an assessment of the consistency of near field electron ptychography with respect to the spatial frequencies of the specimen, using the same method as was used in the optical section (see section 2.3). The FRC results will show that near field electron ptychography is consistent over a range of spatial frequencies and achieves a resolution of under 4nm. Another subsection will discuss near field electron ptychography imaging an MgO cube. This experiment will show that near field electron ptychography has the potential to investigate electric/magnetic fields, but this experiment also highlights the importance of the specimen's diversity in recovering its phase. Then near field electron ptychography magnification will be increased, demonstrating that full detector utilisation is feasible (as typical of near field ptychography), meaning that efficiency of near field electron ptychography can be increased compared to previous subsections. Then there will a reflection on the diffuser reconstructions and how well they compare to the diffuser designs. After that there will be a more critical assessment of the frequency response of near field electron ptychography by initialising the specimen reconstruction with a low frequency random phase noise. This will demonstrate that near field electron ptychography experiences challenges when recovering low frequency phase when initialized in such a way, especially when considering extreme fields of view. The last subsection of this section will demonstrate near field electron ptychography with different propagation distances between the specimen and the detector. These different propagation distances will show that near field electron ptychography is flexible in its implementation and may provide an avenue to investigate near field electron ptychography's frequency response. But this section will begin with a discussion of the specimen used throughout these experiments. Then the thesis will move on to discuss how the diffuser was implemented into the TEM via insertion of Focus Ion Beam (FIB) etched silicon nitride membranes into the selected area plane and the design choices which determined the form of the diffusers.

The specimen used throughout most of the experiments was a diffraction grating replica populated with latex spheres (Ted Pella Product No. 603). One reason for using this specimen is that it has two elements of known sizes, the latex spheres (261nm in diameter) themselves and the diffraction grating replica (2160 lines per millimetre) that they are placed on. These features make determining the resolution (and magnification) convenient (the magnification is unknown initially, due to the conditions under which near field ptychography operates in TEM). The latex spheres also allow for the determination of the quantitative nature of near field ptychography, as latex has a known mean inner potential which can also be measured via phase images produced by near field electron ptychography and compared to other published results [149]. The cracks in the diffraction grating replica help ensure there is sufficient diversity from the specimen. Any specimen would be sufficient in theory to test the quantitative nature of near field ptychography, as long as it has a well-defined thickness rate of change such as cleaved wedges used to measure the mean inner potential in off-axis holography experiments [164]. Cleaved wedges and well-defined shapes are preferable in quantitate measurements as they increase the accuracy, by reducing uncertainty in the thickness measurement [165, 166]. This is why latex spheres are used here, as in theory their thickness can be measured directly by measuring their diameter and assuming that the latex sphere maintains a constant shape during exposure to the electron beam. This particular specimen cannot test whether near field electron ptychography is sensitive to magnetic fields, as there is no magnetic component to this specimen.

4.1 Near field electron ptychography diffuser parameters

Near field electron ptychography relies on a diffuser to constrain the phase problem of the detector. As near field electron ptychography is entirely new, this section begins with details of the diffuser

design in the electron microscope context. These details include a description of how the diffuser is made, the trade-off between the phase delay of the diffuser and the inelastic scatter it generates (and other design decisions), the diffuser designs used in this experiment, and why these designs were used.

Firstly, natural or pre-existing diffusers were not investigated. An example of a pre-existing diffuser would be sandpaper in X-ray near field ptychography or Sellotape in optical near field ptychography. For electron ptychography, a 'natural' diffuser could be an amorphous thin film, e.g. a standard calibration test specimen. Instead, here an etched Si_3N_4(silicon nitride) thin film was used as a diffuser, similar to previous electron beam vortex shaping experiments [167] and a previous implementation of far field electron ptychography [168]. Figure 81 outlines the basics of the diffusers used in the following experiments. The silicon nitride thin film was etched with custom-designed patterns that varied the thickness of the silicon nitride such that the phase delay it imparts varies across the thin film. The Si_3N_4 diffuser is surrounded by a gold support/charge sink and the diffuser pattern is on top of a Si_3N_4 base of constant thickness.

As this was the first attempt at near field electron ptychography with a diffuser, a careful approach was taken when considering the diffuser properties. Examples of design choices include:

- restricting the phase change imparted by the diffuser to $<2\pi$ radians
- minimising the inelastic counts added by the diffuser
- avoiding abrupt changes in depth, as they would require at least 20nm in a perpendicular direction to transition from the lowest value to the highest value (this is to ensure FIB fabrication is accurate)
- restricting the Si_3N_4 window inside the gold aperture to a maximum diameter of $50\mu m$, since if the aperture increased in size, the charging of the Si_3N_4 window would become an issue.

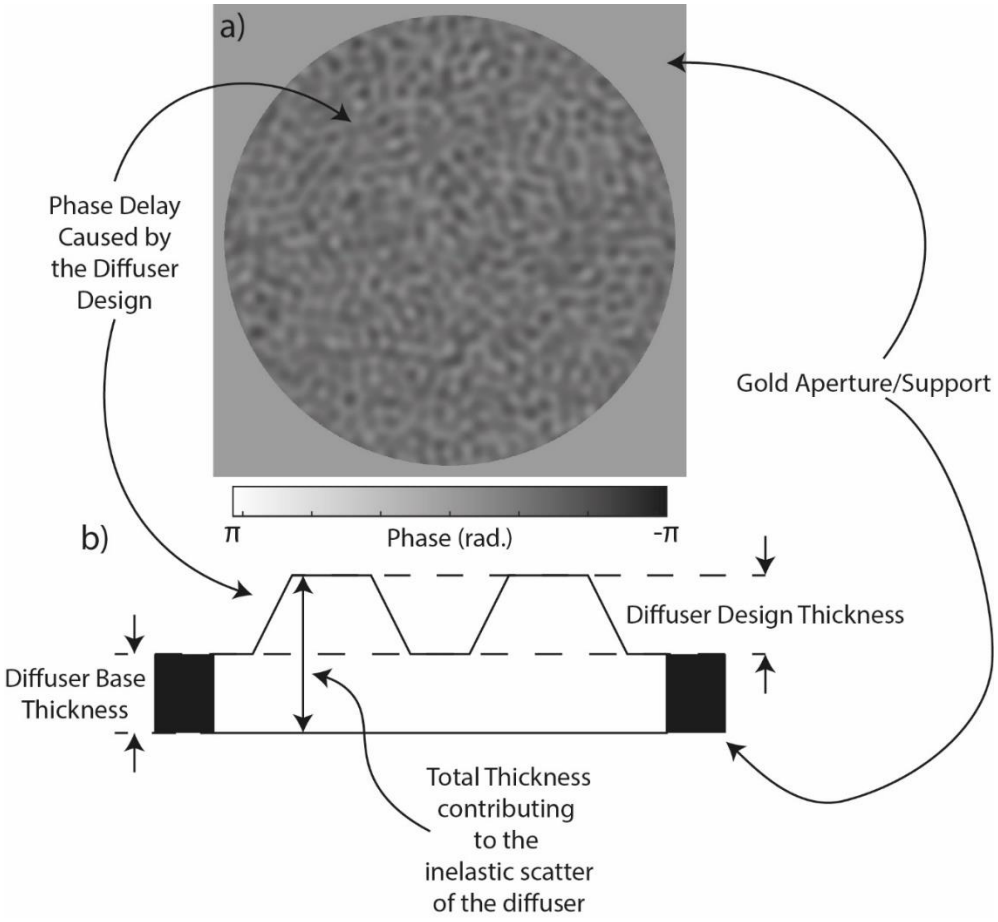

Figure 81: a) phase delay caused by an example of a diffuser design. b) cross section of the diffuser design. The colour bar under a) transforms the greyscale of a) into radians.

These design decisions were taken to minimise the risk of an error occurring when reconstructing the data formed by these diffusers. A diffuser or object which exhibits strong phase will often have its phase wrapped (as it cannot be contained within 2π), which causes problems in the reconstruction, as the transition between π and -π is often difficult to resolve and may incur phase vortices [84]. These phase vortices will make other problems in the data collection harder to identify and will require more iterations of the ptychographic algorithm to solve.

It should be noted that these are extreme limits applied to diffuser design to avoid any problems and should be questioned and subject to further investigation. To satisfy these self-imposed limits on the diffuser design, beginning with the equation describing the relationship between the diffuser thickness and the inelastic counts it incurs, shown in equation 114 [169]:

$$\frac{t}{\lambda_{mfp}} = \ln\left(\frac{I_T}{I_0}\right) \tag{114}$$

The parameters of equation 114 are the following:

- t is the thickness of the silicon nitride
- λ_{mfp} is the mean free path of silicon nitride at 300 keV (166nm at a large collection angle <100mrads, see equations 117-121
- I_T is the total counts on the detector both elastic and inelastic
- I_0 is the number zero loss counts.

As the Si_3N_4 diffuser has two effective layers; the support layer and patterned layer, the inelastic contribution of the support layer should be considered first, as there cannot be a patterned layer without it. As the diffuser was created via Focused Ion Beam etching (FIB etching), its support layer thickness was 50nm [170], which alone incurs significant inelastic scatter. Putting the thickness of the support layer into equation 114 suggests that only 74% of the counts after traveling through the support layer would not have lost significant energy. However as the support is of uniform thickness it does not contribute to the phase delay caused by the diffuser.

To ensure that any datasets collected using a diffuser are reconstructable, the percentage of elastic counts in the total counts should be as high as possible (considering the diffuser individually and not the contribution of the specimen). Using equation 107, the thickness that causes a 2π phase change can be calculated.

$$t_{12} = \frac{2\pi}{12 \times 6.52 \times 10^6} = 80.3\text{nm} \tag{115}$$

$$t_{15} = \frac{2\pi}{15 \times 6.52 \times 10^6} = 64.2\text{nm} \tag{116}$$

It is found that a thickness 64-80nm of Si_3N_4 is sufficient to cause a 2π phase change in the electron beam. A range of thickness values is stated (corresponding to t_{12} and t_{15}) as the mean inner potential FIB etched Si_3N_4 can have varying mean inner potential depending on the effect of Ga ion implantation and on the original quality of the Si_3N_4 membrane, so the mean inner potential of the final diffuser can vary significantly from expected values. A possible worst case is that the Si_3N_4 membrane is not significantly different from pure silicon, corresponding to a mean inner potential of around 12V [171] and therefore produces the greatest amount of inelastic scatter for a given phase delay (compared to any higher value). A more standard value of 15V has also been used, corresponding to a more likely scenario as this is the upper range of the mean inner potential for Si_3N_4 found by Grillo et al. [172]. However, the FIB is calibrated before etching the Si_3N_4 diffusers to ensure they cause the desired phase delay regardless of the mean inner potential. Further details on the FIB process can be found in subsection 4.2. Adding the thicknesses, the 2π phase thickness on to the base thickness means the total thickness of the diffuser is 114-130nm, which in turn means that the percentage of elastic counts in the total counts is reduced to a range of 46-50% . This breaks the second self-imposed constraint of keeping the inelastic counts to a minimum. Therefore, it is clear that the phase change imposed by the diffuser should be limited to less than 2π. A diffuser which causes a 2.35 radian phase change can be implemented with a total thickness of 74 – 80nm, and the electron beam after interacting with this diffuser will have 62-64% elastic counts with respect to the total. With this information it was decided that the phase imparted by the diffusers should be 2.35 radians. It is important to note that the degree of inelastic scatter is related to the acceleration voltage of the electrons, meaning that this diffuser would cause a greater degree of inelastic scatter as the acceleration voltage is reduced. The equation for the constant of proportionality that converts thickness and mean inner potential into phase has been given for reference in equation 107. The proportion of elastic counts was estimated assuming inelastic mean free path of 166nm. The way in which this value is obtained will be discussed in the following paragraph.

Determining the inelastic mean free path for a given material in the TEM or STEM is a challenging task which has been the focus of many researchers examining the way in which various experimental parameters can affect the measured inelastic mean free path, such as: multiple scattering within the specimen [173] and collection angles [174]. In this thesis the model formulated by Iakoubovskii et al. to map inelastic mean free path as a function of acceleration voltage is used. Using the formula from

their paper entitled “Thickness measurements with electron energy loss spectroscopy” [174], this formula is used primarily for the reasons set out in Egerton’s book [169], in that Iakoubovskii’s method does not require the collection semi angle or incident semi angle of the illumination in the case of large collection angles. The collection semi-angle of near field electron ptychography is currently unknown, but is likely to be greater than 100 mrads, as no objective aperture was utilised in our experiments [175]. The equation for calculating the inelastic mean free path is given in equation 117, and a simplified version where the collection semi-angle is large is given in equation 118.

$$\frac{1}{\lambda_{imp}} = \frac{11\rho^{0.3}}{200FE_0}\ln\left(\frac{\alpha^2+\beta^2+2{\theta_E}^2+|\alpha^2-\beta^2|}{\alpha^2+\beta^2+2{\theta_C}^2+|\alpha^2-\beta^2|}\times\frac{{\theta_C}^2}{{\theta_E}^2}\right) \tag{117}$$

If the collection angle $\beta \gg \theta_E, \theta_C$ then the equation can be simplified to the form seen below in equation 118:

$$\frac{1}{\lambda_{imp}} = \frac{11\rho^{0.3}}{200FE_0}\ln\left(\frac{{\theta_C}^2}{{\theta_E}^2}\right) \tag{118}$$

The equation used to determine the relativistic factor F is given below in equation 119:

$$F = \frac{1+\frac{E_0}{1{,}022{,}000}}{\left(1+\frac{E_0}{511{,}000}\right)^2} \tag{119}$$

The parameters of equations 117-119 are the following:

- θ_C is the cut-off angle [169]
- θ_E is the characteristic scattering angle of the material being considered
- ρ is the density of the material being considered, in this case 3.2 g/cm³
- E_0 is the acceleration voltage of the electron beam in keV (kilo electron Volts)
- F is a relativistic factor to take account of the effects that occur when the electron is moving close to the speed of light computed via equation 119.
- β collection angle semi-angle, in this case assumed to 150 milliradians
- α incident converge semi-angle, in this case assumed to be 20 milliradians

Below a filled out version of equation 118 is provided in order to show the values used to calculate the inelastic mean free path used in this thesis (see equations 120 and 121):

$$\frac{1}{\lambda_{imp}} = \frac{11\times 3.17^{0.3}}{200\times 0.5135\times 300}\ln\left(\frac{0.020^2}{(50.5\times 10^{-6})^2}\right) \tag{120}$$

$$\lambda_{imp} = 166\text{nm} \tag{121}$$

The value of 166nm for the inelastic mean free path of silicon nitride at an acceleration voltage of 300 keV seems reasonable when compared to the simulated value of 179nm obtained for silicon when using Iakoubovskii’s method at the same acceleration voltage obtained by Potapov [175]. Due to the similar densities and atomic numbers of these materials, it makes sense that they have similar inelastic mean free paths, suggesting that the mean free path of Si_3N_4 has been correctly calculated.

Returning to discussion of the diffuser design, the requirement that the diffuser features change at a limited rate poses a constraint on the effective magnification of near field electron ptychography. It is advantageous if the diffuser can produce a speckle pattern which is only slightly larger than a pixel, as this has been found to reduce the reconstruction error [132] compared to when the speckles are larger. Moreover, the diffusers would be non-functional in the case where the diffusers are magnified to such a level that they produce no diversity/speckle at the detector (i.e. when a single feature covers the field of view of the detector). Furthermore, as the features of the diffusers are magnified, they require longer propagation distances in order to generate equivalent diffraction fringes. If the FIB etching could produce Si_3N_4 diffusers with finer features it may even enable near field ptychography to image atomic columns.

4.2 FIB details

To further detail the FIB work required by near field electron ptychography, it is important to state that any FIB work, apart from the design of the diffuser and writing of the stream file, is due to the work of Peng-Han Lu and Maximilian Kruth of Julich Forschungszentrum. The base material for the diffuser was a silicon wafer sandwiched between two layers of silicon nitride. This base was then processed to remove one of the silicon nitride layers and the silicon wafer, this was achieved via anisotropic wet etching which produced a series of trenches with the thinnest part being the silicon nitride (an effective silicon nitride window) [176, 177]. This process was used to create 9 silicon nitride windows, 4 of which were used in this thesis as will be detailed later in this section (Si_3N_4 grids are also available from Ted Pella and Norcada). Before the diffusers were etched, gold was used to cover the surface of the Si_3N_4 grid (approximately 200nm thickness) to create a conductive surface preventing the non-etched surface contributing to the electron beam. To regain access to the Si_3N_4windows the gold layer was milled out in 50µm diameter circles. The silicon nitride windows were then processed into diffusers via Focused Ion Beam etching which was performed using the FEI Helios NanoLab 460F1 (which also removed the gold layer). The FEI Helios NanoLab 460F1 thinned the Si_3N_4 window inside the gold aperture via ion milling (using gallium ions), such that it incurred the desired phase map with a maximum delay of 2.35 radians whilst retaining a base thickness of 50nm. The ion beam energy was 30keV. The ion beam current was 2.5nA for milling the gold and 0.40nA for stream file patterning. The smaller ion beam current is used to produce a finer probe size to improve the fabrication precision for the diffuser. The diffuser designs were given to the FEI Helios NanoLab 460F1 via a stream file which detailed the dwell time as a function of the position on the window. The required milling time to achieve a set thickness/phase delay of Si_3N_4 was calibrated before the fabrication began, by milling calibration films using different ion beam settings. A hole was then milled in the calibration films such that their phase delay could be measured by off-axis holography. With these results the ion beam settings could be fine-tuned to achieve the desired results. Details on how the FIB is calibrated to cause a set phase delay will be reported in [178].

While the details of the FIB process given above are important, it is vital to state that the near field electron ptychography diffusers can in principle be created by any FIB capable of etching silicon nitride windows with a feature size less than 5µm. For example, if a FIB system can etch its associated university or business logo in an area of 50µm in silicon nitride, that logo could then be used as a diffuser for near field electron ptychography and still perform adequately for the purpose of collecting phase images (for example in Robisch et al. [127] use a logo like design and achieve reasonable phase images).

4.3 Designs for the etched thickness profiles

There will now be a discussion on the form of the phase patterns used in these experiments. Near field ptychographic diffusers are generally robust to the exact form of illumination, as can be seen by comparing Figure 105 and Figure 49, which were reconstructed successfully along with their corresponding specimens. Diffusers can also be highly structured, like the one used in figure 2 of [127]. However, work by others on X-ray near-field ptychography has suggested that smaller speckle sizes and high visibility in the collected diffraction patterns can lead to reconstructions with reduced error values [132].

The diffusers designs used are shown in Figure 82. Many of the diffusers have sharp features such as lettering and rectangles, which is primarily to act as a guide in the reconstruction process, since these sharp features should be in focus and legible when the reconstruction parameters are accurate. As discussed in the data collection section, the propagation distance of near field electron ptychography diffraction condition is unknown until after the experiment. Where reference data are used to estimate the propagation distance, this estimated propagation distance can then be quickly judged by the sharp features of these diffusers. Diffuser 1 (Figure 82 a)) was designed to be a combination of sharp features and large speckle-like features. Diffuser 2 (Figure 82 b)) was designed to test the phase sensitivity of near field electron ptychography to a basic degree, as essentially it is three separate designs layered on top of each other. Each layer was assigned 0.783 radians of phase delay, and where the layers overlap their phase delays combine to a maximum of 2.35 radians. Diffuser 3 (Figure 82 c)) was designed to mimic the kind of diffraction data that are obtained from X-ray and optical diffusers (Sellotape and pieces of paper). Diffuser 4 (Figure 82 d)) is basically a non-speckle diffuser, whose regular structure enables a simple test: confirming the maximum phase delay matches to the design specification. These are likely not the optimal diffusers for near field electron ptychography – which more likely contain a greater range of spatial frequencies - but they cover a sufficient range to start testing whether near field electron ptychography is robust to the form of illumination.

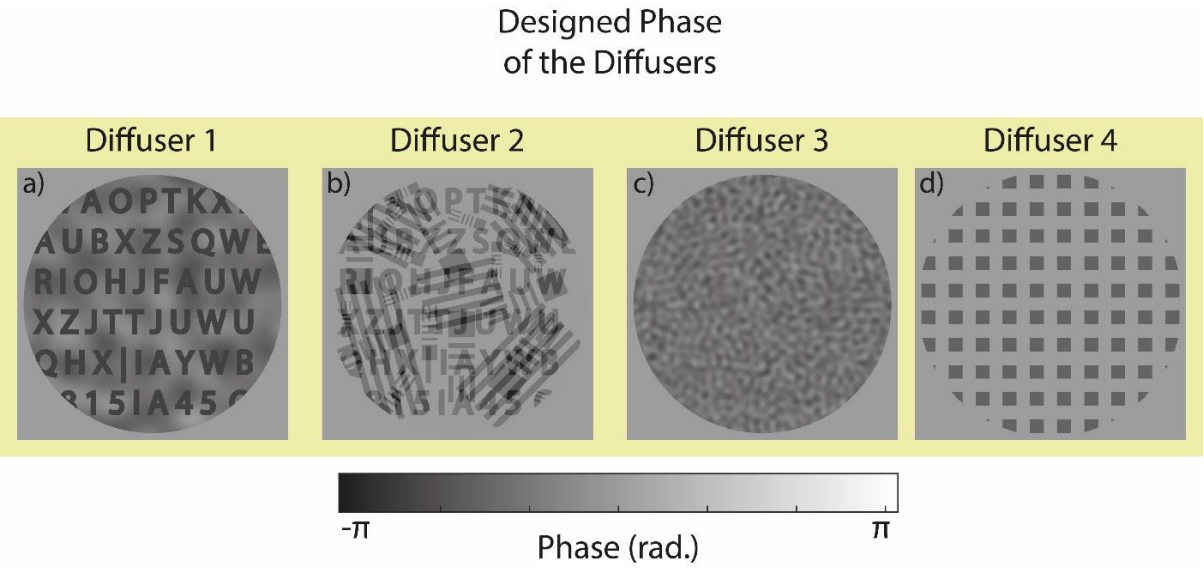

Figure 82: A figure showing the range of theoretical diffuser designs used to generate the stream files provided to the FIB system described in section 4.2. a) combination of phase which causes typical speckle patterns whilst containing highly structured features. b) a highly structured diffuser made up of three independent levels to enable basic test of phase sensitivity. c) a diffuser designed to create a speckle pattern similar to those found in other modalities. d) A structured diffuser with single repeating feature which enables comparison between the reconstructed and designed phase.

4.4 Experimental details and diffuser mounting

The details of the experiment parameters and instrumentation were originally published in [179] but are restated here. The microscope used to collect the near field electron ptychography data displayed in this section of the thesis is the FEI Titan 80–300 STEM (a link to further microscope details is provided: FEI Titan 80-300 STEM - er-c). It is also important to state that apart from setting up the diffraction condition for near field electron ptychography, calibration and operation of the TEM was performed by Peng-Han Lu of Julich Forschungszentrum. The microscope was operated as a TEM in an altered diffraction mode, this altered diffraction mode allowed for defocused images of both the diffuser and the specimen. This altered diffraction mode was achieved by changing the strength of the diffraction lens, such that the microscope provided defocused images of the selected area aperture plane instead of imaging the back focal plane of the objective lens. The acceleration voltage of the electron gun was 300keV and the spot size of the illumination system was 3. The convergence angle of the illumination was estimated to be of the order of 0.5mrad. The objective lens magnification was 59 times, and the magnification of post objective lenses was measured using the specimen (in this case the carbon grating replica) to be 402 times. A Gatan image filter (Tridiem 866 ER) was utilised to obtain the energy filtered datasets (Sections 4.5, 4.6, 4.7 and 4.11 have datasets which were energy filtered). Unless stated overwise, the energy filter slit width used when energy filtering was 20eV. All diffraction patterns were collected using the Gatan Ultrascan CCD, which has a pixel size of 14µm. The detector was set to binning 2 during the experiments and the exposure time for all experiments was 0.5 seconds unless stated otherwise. For all experiments apart from those documented in subsection 4.13, the camera length displayed by DigitalMicrograph was 0.48m. The specimen was scanned with respect to the diffuser via the use of the FEI CompuStage motorised positioning stage, and the stage was controlled by the scripts outlined in Section 3.

The selected area aperture holder of the FEI Titan 80–300 STEM has four individual exchangeable aperture holders, each is 3 mm in diameter. In order to insert the diffuser into the selected area aperture plane the column was vented so that the holder could be removed. Then one of the existing apertures was replaced with the fabricated diffusers, after which the aperture holder was inserted back into the microscope. In order to restore the vacuum to the column, the air was pump out of the column for a few hours. The microscope was then left overnight and the experiments began the next day. The diffuser which was used throughout this section was diffuser 3 in Figure 82 c), other diffusers were used during the initial calibration of near field electron ptychography as their strong features assisted in the determination of the propagation distance.

4.5 Minimum diffraction pattern experiment

The first experiment performed with near field electron ptychography was to explore the minimum number of diffraction patterns required to obtain accurate phase images. The minimum number of diffraction patterns required to constrain the phase problem in X-ray and optical near field ptychography has been found to be 16 (or even 6 diffraction patterns in the optical case [125]) [79]. Therefore, the results of this experiment can easily be compared to other forms of near field ptychography. This experiment is also designed to inform any future user of the minimum amount of time that is required to collect a reasonable dataset.

The experiment consisted of generating square scan patterns of increasing size, starting at a two by two scan and ending in a five by five scan. The step size used in this experiment was 100nm. As the effective diffuser diameter at the specimen was 870nm, the overlap between adjacent diffraction patterns was calculated to be 88.5% ((870 - 100 / 870)*100). The diffraction patterns in this experiment were energy filtered to reduce the noise from inelastic scattering, and the energy slit

used had a width of 20eV (electron volts). As an additional step in the investigation, the reconstruction process was carried out again, this time using the probe reconstruction from the largest dataset to initialise the reconstruction of all the other datasets. This was done to see if the minimum number of diffraction patterns could be reduced by knowing the probe reconstruction beforehand.

Other settings in these experiments were as follows: the exposure time of the camera was 0.5 seconds. The camera length displayed by DigitalMicrograph was 0.48m, and the propagation distance used in the reconstruction was found to be 0.815m. The de-magnified camera dimensions at the diffuser plane were 72.17µm by 72.17µm for a 1024 by 1024 (binning 2) pixel detector, and acceleration voltage was 300keV corresponding to a wavelength of 1.97pm. The objective lens magnification was 59 times. This leads to a pixel size of 1.19nm. The pixel size was confirmed to be 1.18nm when using the diffraction grating replica of the large field of view experiment (see section 4.8). The pixel size provided by the large field of view measurement is used through all experiments that share this diffraction condition. The datasets were processed with an altered version of the rPIE code mentioned in 1.24, until their error values stagnated at a minimum value. The alterations made to the rPIE code can be found in the supplemental material [94].

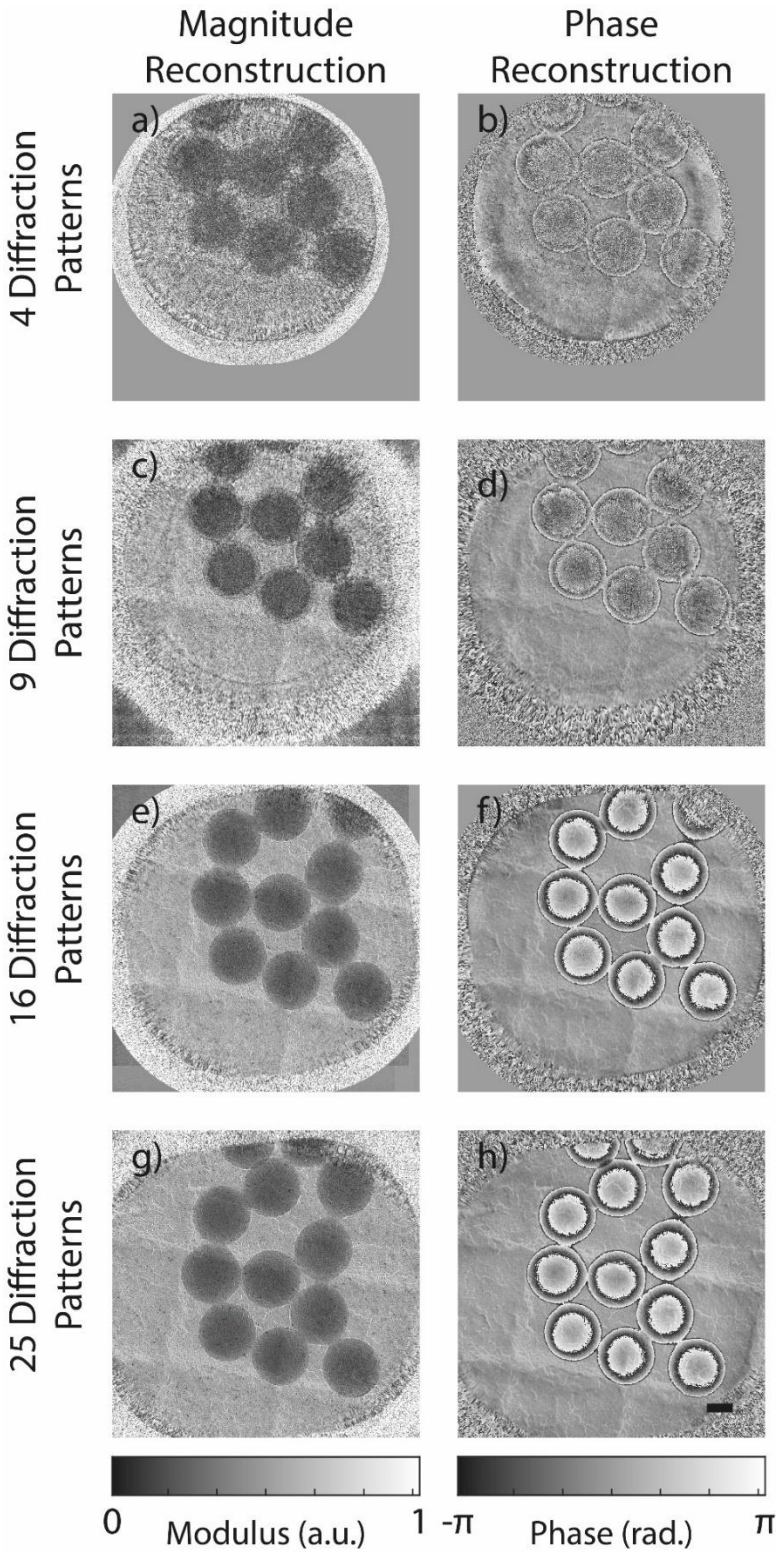

Figure 83: a figure showing how the reconstruction quality varies as the number of diffraction patterns is varied. a-b) 4 diffraction patterns, c-d) 9 diffraction patterns, e-f) 16 diffraction patterns, g-h) 25 diffraction patterns. The colour bars at the bottom of the columns convert the greyscale into an arbitrary modulus and radians respectively.

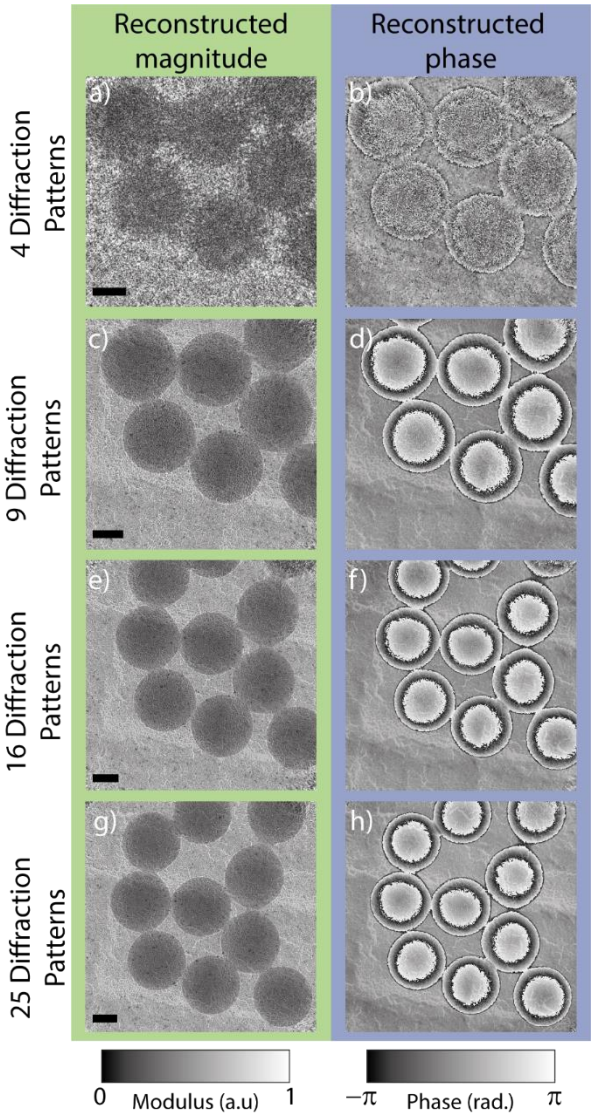

Figure 84: A figure showing how the reconstruction quality varies as a function of diffraction pattern number, in the case where a reconstructed probe was provided to the reconstruction algorithm, reducing the number of unknowns. a), c), e), and g) are the magnitude reconstructions as the diffraction pattern number varies. b), d), f), h) are the phase reconstructions as the diffraction number changes. The scale bars in the magnitude reconstructions incidate 100nm. The colour bars at the bottom of the colunms convert the greyscale into an arbitrary modulus and radians respectively.

Figure 83 shows the results from investigating the minimum number of diffraction patterns where the reconstruction algorithm had to recover both the specimen and the illumination function. The minimum number of diffraction patterns needed in order to constrain the phase problem was found to be 16 diffraction patterns, in line with results from previous X-ray and optical experiments [79, 125]. This can be seen when comparing Figure 83 d) to Figure 83 f), as the strong phase of the latex sphere in f) is comparatively weaker in d), such that it is similar to the carbon grating replica and the phase wraps are unresolved.

To assess the phase accuracy of these reconstructions in a more quantitative manner, the mean inner potential of each of the latex spheres was calculated in the way outlined in section 3.7. The values calculated by this process can be found in Table 4 through Table 6. On average the mean inner potential of the latex spheres was 8.45±0.39V for the 16 diffraction patterns and 8.54±0.35V for 25 diffraction patterns.

The latex sphere number in Table 4 through Table 6 refers to the numbered latex spheres presented in Figure 85 a). Due to the nature of the experiment not every latex sphere is imaged, as the field of view shrinks as the number of diffraction patterns is reduced. Figure 85 also includes the radial average of the latex sphere phase plotted with a grey area which represents a perfect sphere. The axes are normalised as the latex spheres vary in size. Figure 85 b) suggests that near field electron ptychography has obtained quantitative phase information, as the phase matches closely to the normalised perfect sphere for most of the plot. The largest deviations occur at the start and end of the plot shown in Figure 85 b). The deviations at the start of the plot are in part due to the low number of data points in the radial average and the diffraction grating replica influencing the phase changes as the phase of the latex spheres changes slowly near their centre. The deviations at the end of the plot occur due to the manually selected centre of the latex sphere being off centre from its true value, and therefore the final values may be affected by the phase of the carbon diffraction grating replica being included in the radial average.

Results from the second experiment suggest that using a probe reconstruction from another dataset to initialise the probe does reduce the number of diffraction patterns required to retrieve an accurate phase of the specimen. This can be seen by inspection of Figure 84, as the reconstruction using 9 diffraction patterns has improved in quality compared to the case where the probe is unknown. Furthermore, the mean inner potential of the latex spheres was measured and found to match other measured values. The measured mean inner potential of the 9 diffraction pattern reconstruction was found to be 8.43±0.39V. Apart from illustrating that a reduced number of diffraction patterns is required to recover the phase of the specimen, there are few differences between Figure 83 and Figure 84. This lack of difference suggests that the only benefit of providing the diffuser to the reconstruction algorithm is an increase in the rate at which the reconstruction converges to a solution, when there are enough diffraction patterns.

Latex Sphere Number	Mean Inner Potential	Latex Sphere Radius (nm)
1	8.56	114.35
2	8.70	117.60
3	8.37	115.18
4	9.22	115.89
5	8.61	114.61
6	8.16	115.64
7	8.02	115.64
8	8.60	115.75

Table 4: a table showing latex sphere number, mean inner potential and measured latex sphere radius from a near field electron ptychography dataset with 25 diffraction patterns. Note that the pixel size is 1.18nm for this experiment.

Latex Sphere Number	Mean Inner Potential	Latex Sphere Radius (nm)
1	8.78	115.16
2	8.42	119.07
3	9.05	118.11
4	9.41	118.90
5	8.47	116.76
6	8.16	118.10
7	8.04	117.18
8	9.06	118.56

Table 5: a table showing latex sphere number, mean inner potential and measured latex sphere radius from a near field electron ptychography dataset with 16 diffraction patterns

Latex Sphere Number	Mean Inner Potential	Latex Sphere Radius (nm)
2	8.66	121.18
3	8.60	117.19
4	8.71	116.32
5	8.68	116.88
7	7.74	116.18

Table 6: a table showing latex sphere number, mean inner potential and measured latex sphere radius from a near field electron ptychography dataset with 9 diffraction patterns

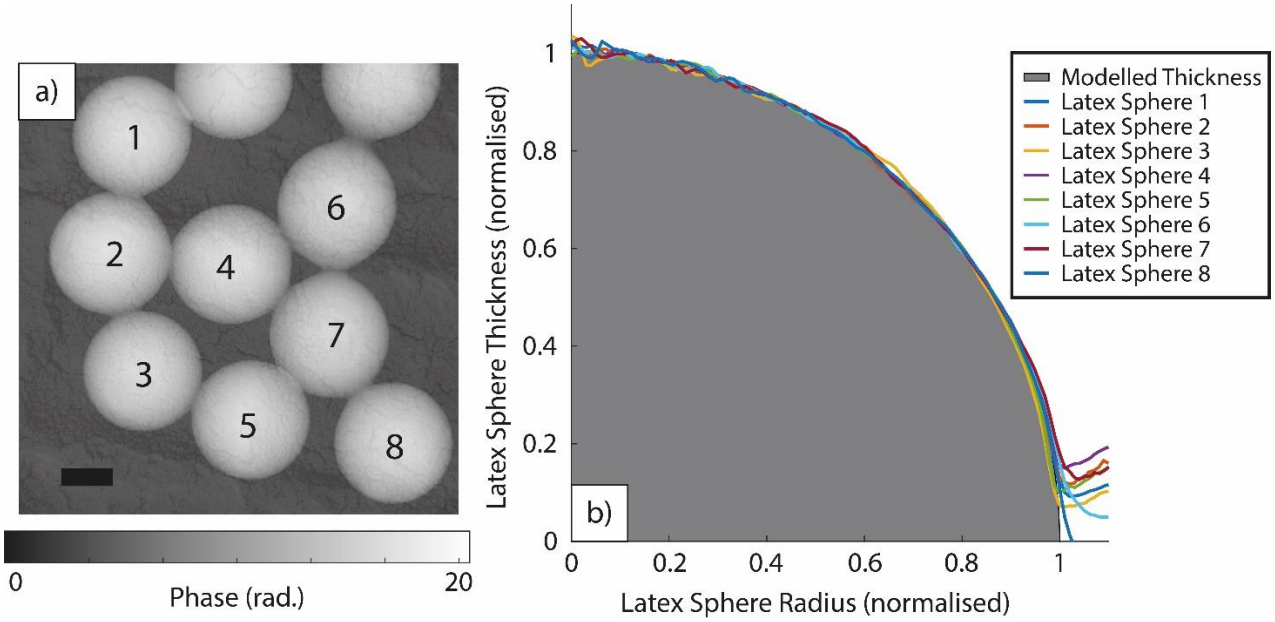

Figure 85: a) The unwrapped phase of Figure 84 h) with latex spheres numbered. The mean inner potential of the latex was measured for each of these spheres and the results averaged to give the statistics listed in the main text. The scale bar is 100 nm. The colour bar at the bottom of the figure converts the greyscale image into unwrapped phase in terms of radians. The sign of the unwrapped phase in a) was inverted to obtain a positive mean inner potential. b) shows the radially averaged phase of the numbered latex sphere fitted to a normalised modelled thickness in order to demonstrate the quality of the fitted curves. The data used in b) come from Figure 84 h) via the process described in Section 3.7 [179].

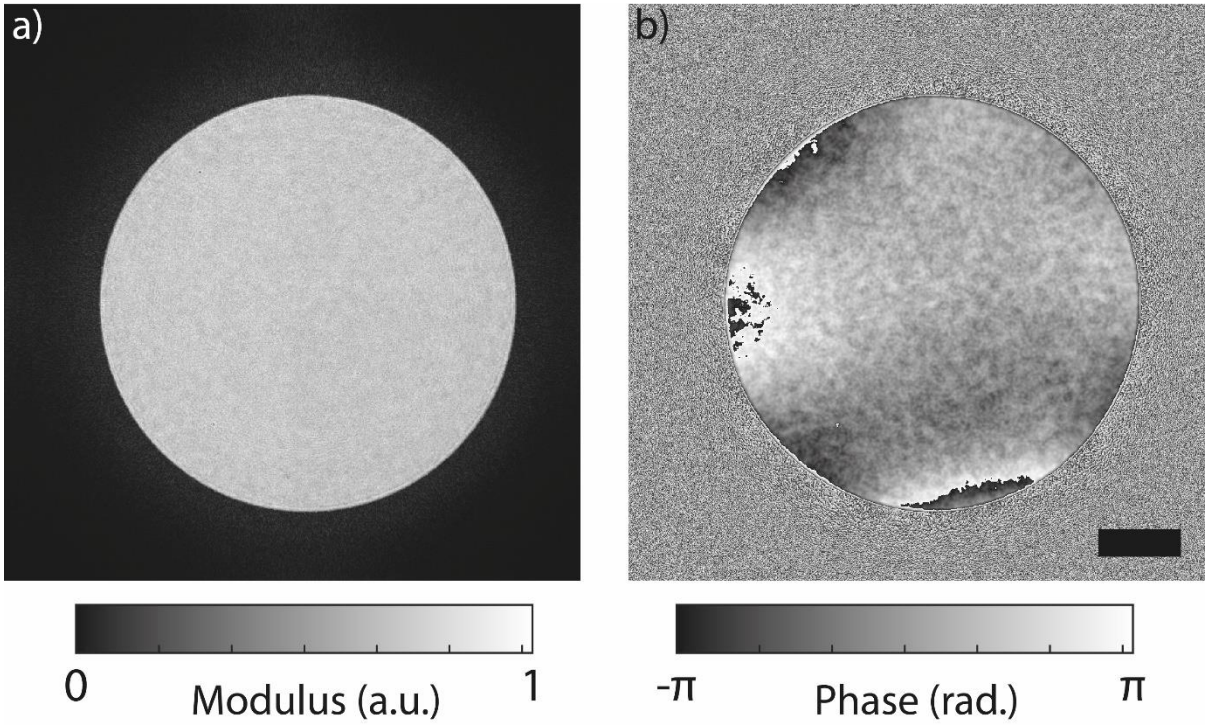

Figure 86: a figure showing the magnitude a) and phase b) of the reconstructed diffuser from the 25 diffraction pattern dataset. The diffuser shown here corresponds to the diffuser design shown in Figure 82 c). The colour bar under a) converts the greyscale of a) into a normalised unitless magnitude. The colour bar under b) converts the greyscale of b) into phase in terms of radians. The scale bar in b) indicates a distance of 10μm in the selected area aperture plane.

Figure 86 shows the diffuser reconstruction obtained from the 25 diffraction pattern dataset. The diffuser shown in Figure 86 was also used to help reduce the number of unknowns when reconstructing the 9 diffraction pattern dataset. The diffuser used in this experiment was diffuser 3 and this diffuser reconstruction closely matches the diffuser design shown in Figure 82 c), apart from the strong phase features at the edges of the aperture, which will be discussed in more detail in Section 4.11.

The fact that near field electron ptychography has been able to successfully measure the mean inner potential of latex spheres is an important first step, as in the past other phase imaging methods have struggled to obtain consistent mean inner potential values for various materials. The mean inner potential obtained via off-axis holography has on occasion been inconsistent, for example the measurement of the mean inner potential of zinc oxide [165]. The inconsistencies experienced are often due to the incorrect determination of the thickness of the specimen, which can be performed via EELS (Electron Energy Loss Spectroscopy), analysis of CBED patterns (Convergent Beam Electron Diffraction patterns), or tomography [165]. These thickness measurements can often be affected by contamination of the surface of the sample, either by sample preparation steps or contamination build-up during the experiment, so in theory this can occur in near field electron ptychography as well. The uncertainty of the thickness measurements can be circumvented by instead fitting a phase change to a thickness change; this is illustrated by [164], where Gajdardziska-Josifovska et al. used crystal wedges with well-defined thickness changes, which is now a well-established method to measure the mean inner potential of a specimen. The latex spheres investigated in this experiment also provide a well-defined thickness change profile which has likely contributed to the accurate measurement of the mean inner potential of latex. Further examples of inconsistencies in the measured and simulated values of the mean inner potential are documented in [171], where

Pennington et al. suggest that this discrepancy is due to the surfaces of the prepared samples. The variance in the measured mean inner potential has also been attributed to dynamical scattering, as discussed in [180]. If this is the case, near field electron ptychography may prove to be useful, as electron ptychography has been shown to take account of dynamic scattering via multi-slice reconstruction algorithms [161], which in theory can also be applied to near field electron ptychography. Near field electron ptychography has demonstrated here that it can obtain quantitative phase images of latex spheres, suggesting it may be helpful in determining the mean inner potential of further specimens, however the question of whether near field electron ptychography is more consistent than off-axis holography will have to wait until a direct comparison is performed with prepared samples of equal quality.

This multi-slice approach could possibly be used to mitigate the effect of contaminated surfaces. Imagine three sections of the sample: the upper surface, the central cut out and the bottom surface; by having an individual phase image just for the central section of the semiconductor sample, the additional thickness of the contamination is removed, increasing the certainty in the phase images and the thickness measurement. Although currently this is just a hypothesis.

To be balanced in this analysis, near field electron ptychography has only been used to investigate one type of surface/material with a well defined thickness profile, so an important next step will be investigating further materials to give greater credibility to the claim that near field electron ptychography is quantitative.

Another important aspect to discuss is the data collection time, as on average it takes 3 seconds to collect a near field electron diffraction pattern, meaning that the minimum time to collect a quantitative dataset is 27 seconds, assuming that a good model of the illumination/diffuser is available. Otherwise, the data collection time increases to 48 seconds to obtain quantitative specimen and diffuser images, if there is no illumination/diffuser model available. Comparing this to focused probe ptychography, the number of diffraction patterns collected per second is far lower, mainly due to the fact that near field electron ptychography has to wait for the mechanical stage to move and settle before taking an exposure. For example, focus probe ptychography with modern detectors operating in a binary counting mode can achieve data collection rates of 12.5kHz [181]. If near field electron ptychography can leverage improved data collection rates from scan coils and direct detection detectors, then quantitative phase datasets could be accessed in seconds as opposed to tens of seconds. Off-axis holography may have already provided a solution to the slow scan times of near field electron ptychography, in combining the small movements of scan coils with large movements provided by the specimen stage [135]. Further work will therefore include updating the data collection code of near field electron ptychography to use such a system.

The fact that near field electron ptychography has obtained quantitative phase images without the need of a reference wave is important to state, as the requirement that a vacuum is present has at times been detrimental to off-axis holography. For example, off-axis holography cannot directly investigate the fields around charged/magnetic tips as the electric/magnetic fields of these specimens are long range in nature, meaning that they often corrupt the reference wave. The corruption of the reference wave means that off-axis holography is often limited to measuring the phase difference opposed to the phase itself, meaning that additional computation models are required to confirm results [182]. This advantage of near field electron ptychography means that it has potential to investigate the application of charged/magnetic tips which can act to impart orbital angular momentum on electron beams [183] or charged nanowires, which may have applications as next generation electron guns [182, 184].

4.6 The 'no diffuser' experiment

In order to explicitly show the importance of a diffuser to the functional operation of near field electron ptychography, it was thought prudent to show the results of an experiment where no diffuser was used. In this experiment the specimen was scanned through a 9 by 10 rectangle scan pattern. The step size was 100nm and the propagation distance was 0.815m. Instead of a diffuser, a 45µm diameter aperture was used. The same diffraction conditions were used as for the minimum diffraction pattern experiment.

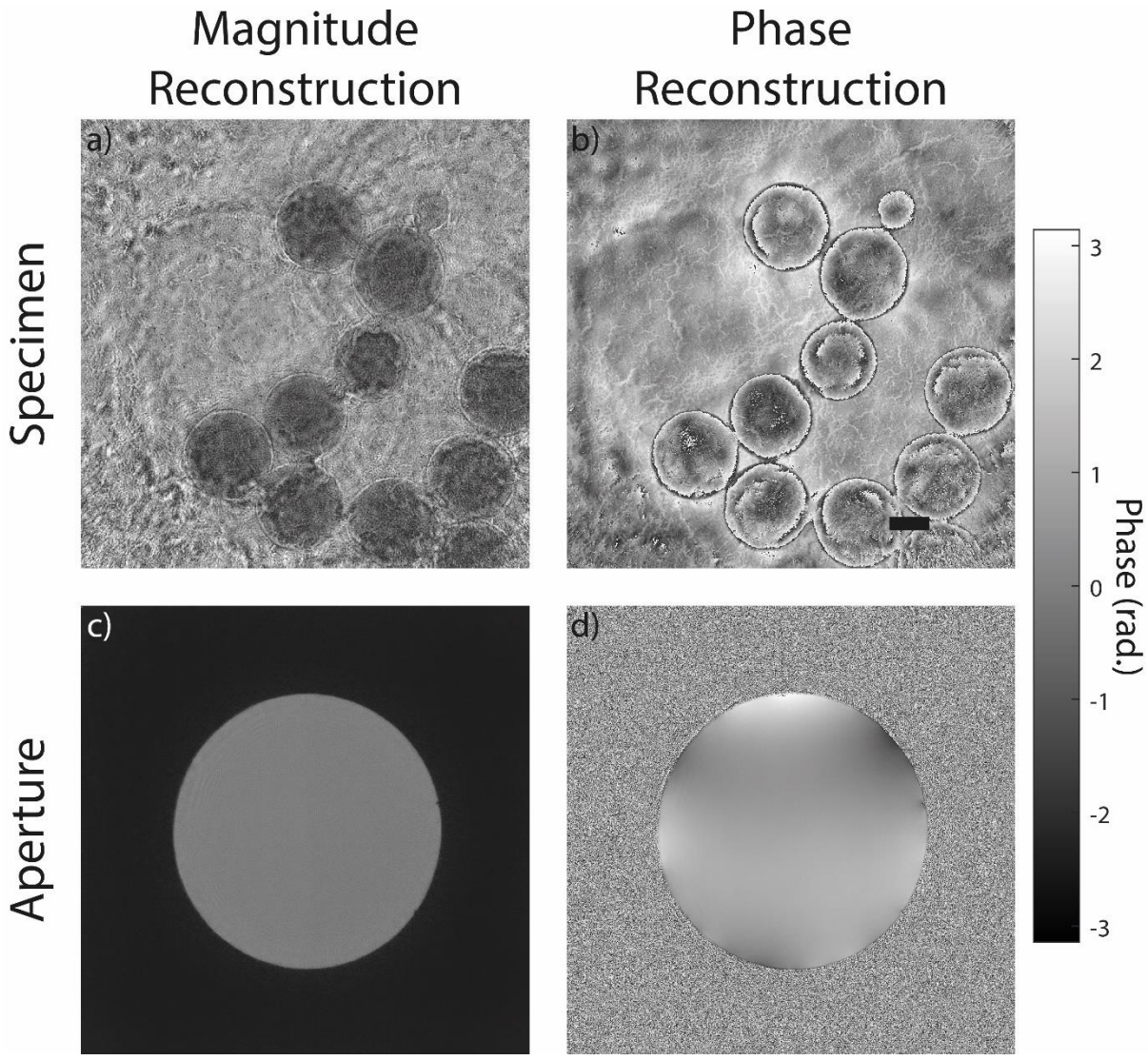

Figure 87: a) magnitude reconstruction of latex spheres on diffraction grating replica. b) phase reconstruction of latex spheres on a diffraction grating replica. c) magnitude reconstruction of a 45µm diameter aperture. d) phase reconstruction of a 45µm diameter aperture. The scale bar in b) indicates 100nm. The colour bar on the right of the figure coverts the phase of b) and d) into radians.

Reviewing Figure 87, it can be seen that reconstruction has been unsuccessful, especially comparing it to the reconstructions in Figure 84. In Figure 87 the phase wrapping of the latex spheres is unresolved, leading to phase vortexes and zeros in the magnitude reconstruction. The probe has fared better visually than the specimen but has only recovered significant phase close to the aperture edge.

In the original Selected Area Ptychography paper (SAP) [94], detailing a precursor of near field electron ptychography, a 10µm aperture was used. An aperture can function as an effective diffuser given the correct diffraction conditions, as the SAP paper proved and section 2.4 of this thesis. A visual check, showing when the diffraction condition is sufficient for an aperture to act as a diffuser, is that the ringing (diffraction) of the edge of the aperture has reached the centre of the aperture, so that no part of the propagated aperture is flat or undisturbed by the propagation distance. Current expectations are that most electron microscopes will struggle to make a 45 micrometre diameter aperture ring/diffract sufficiently to act as a diffuser, as the required propagation distances are too long. The required propagation distances may not be possible due to the limit of the camera length in diffraction mode, which is typically limited in the range of 3m to 4.5m. Another barrier may be the transverse coherence length of the illumination system, as the centre of the aperture may be too far to allow coherent diffraction in the case of apertures larger than 10µm. It should be stated that no such experiment has been attempted. An example of the diffraction condition where an aperture can act as a diffuser, and the diffraction pattern from a 45µm aperture used in the reconstruction of Figure 87, can be seen in Figure 88.

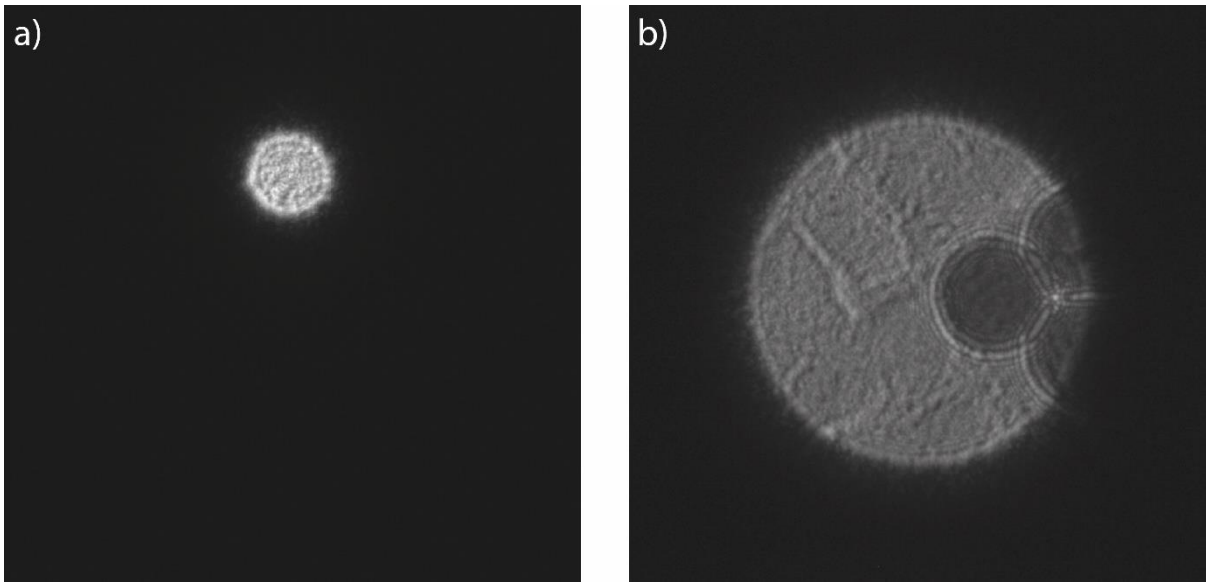

Figure 88: a) Selected area ptychography diffraction pattern, b) a near field ptychography diffraction pattern where 45 micrometre aperture has been used as a diffuser.

4.7 The effect of energy filtering on the reconstruction quality of near field electron ptychography

Our first experiment used an energy filter to reduce the inelastic background present in the diffraction data. Since the need for an energy filter reduces the number of microscopes that could use near field electron ptychography, our second experiment investigated the possibility of a filter-free implementation.

A known issue with energy filtering is the effective drift of the energy filtering slit or energy drift (i.e. instability in the acceleration voltage)[146]. The drift of the energy slit of the image filter is a major issue if the slit drifts significantly, such that the zero loss peak is filtered out of the image and plasmons are imaged instead. This can be seen in Figure 89. This is not an issue for single exposure experiments such as off-axis holography, or for near field electron ptychography with a small number of diffraction patterns. However, near field electron ptychography over highly extended fields of view can take a long time (up to an hour and half using specimen stage movement), so if the

energy filtering can be replaced by noise counter measures (or is not needed), this enables near field electron ptychography to image larger fields of view not accessible by other techniques.

Proposed and investigated here is an algorithmic correction of the inelastic background, where instead of energy-filtering, removal of an inelastic background is factored into the reconstruction process. In this section there will be:

- Reconstruction of energy filtered and non-energy filtered datasets without background modelling
- A demonstration of this algorithmic background removal using simulated datasets with added background noise
- The effect of background modelling will be shown on a non-energy filtered dataset
- Initial assessment of resolution and the effect of energy filtering on the resolution.

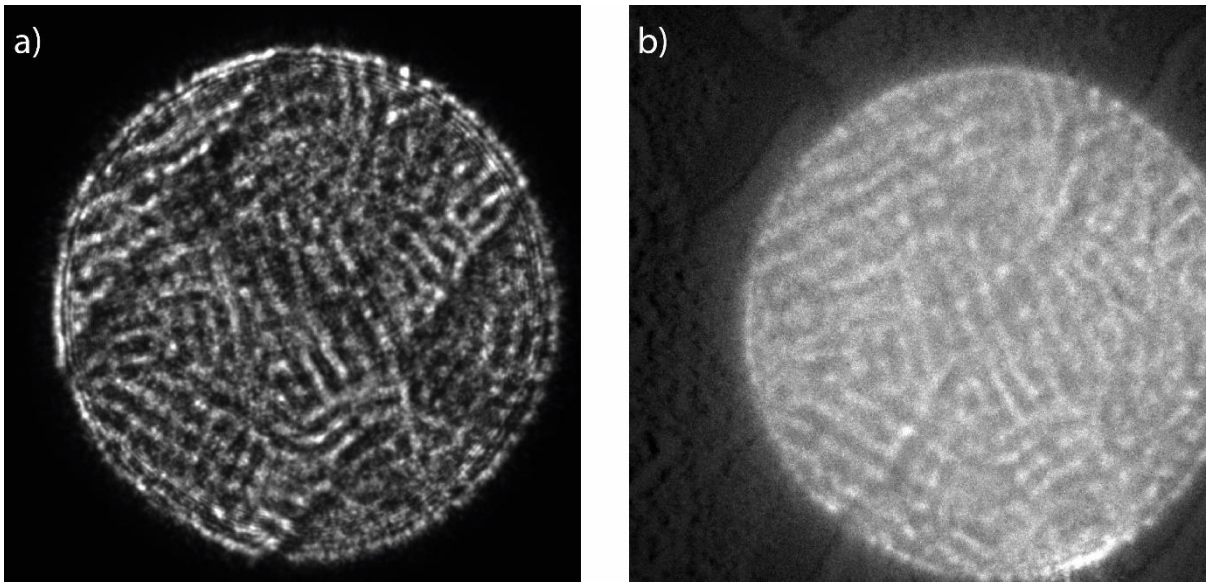

Figure 89: a) A typical near field electron ptychography diffraction pattern. b) a near field electron ptychography diffraction pattern which has been energy filterd incorrectly. The energy slit in b) has drifted such that zero loss peak has been filtered out and the electrons which have lost energy have imaged instead.

The effect of energy filtering on near field electron ptychography will now be demonstrated by the reconstruction of various datasets, where the width of the energy slit was varied and no background modelling was applied. The scan patterns in this experiment consisted of 90 diffraction patterns (in a 9 by 10 grid). Otherwise, all experimental parameters remain unchanged from the last experiment except for the energy slit used by the Gatan Image Filter. The following slit values were chosen: 20eV, 10eV and unfiltered.

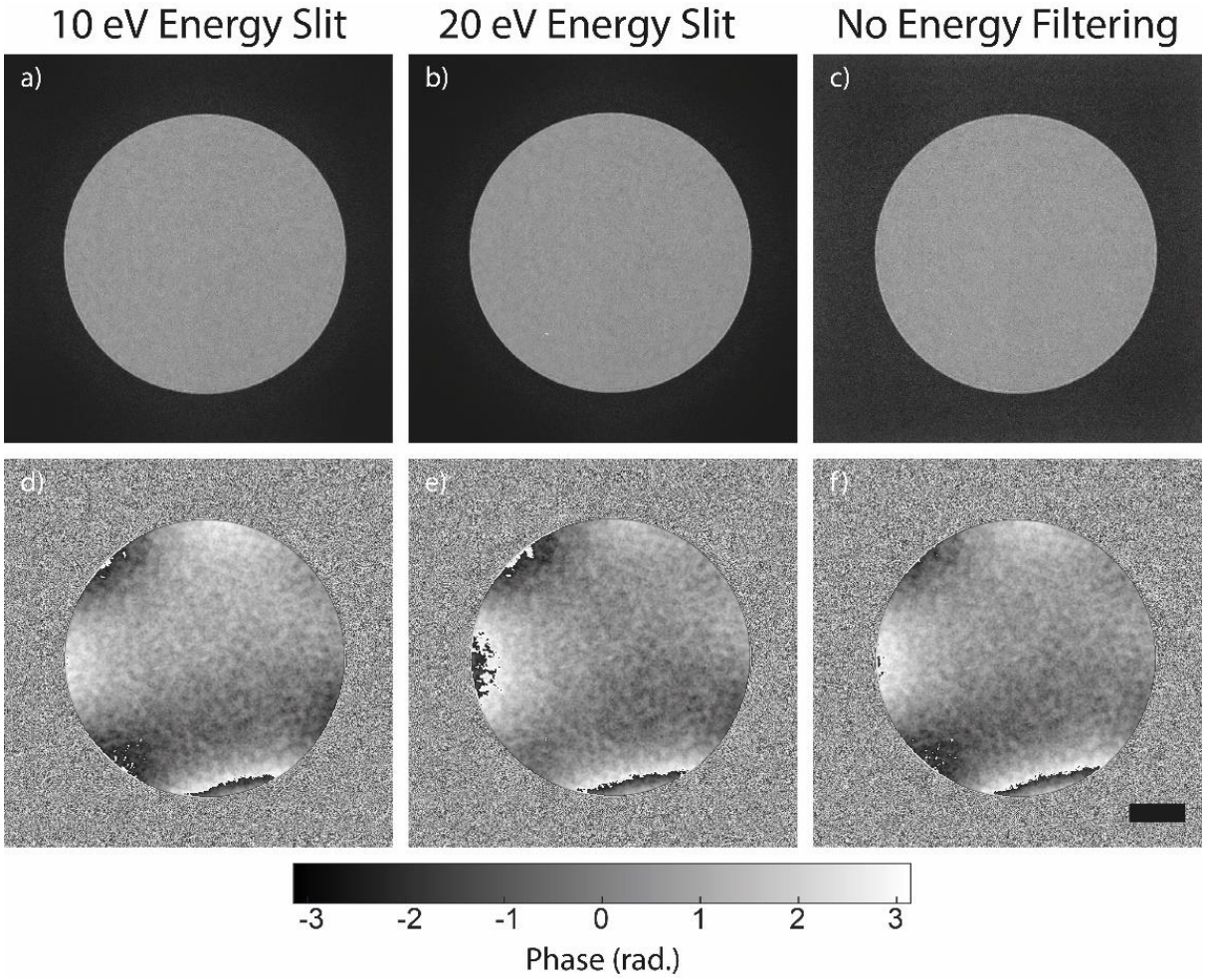

Figure 90: Reconstructed probes with differing levels of energy filtering. a) to c) magnitude reconstruction. d) to f) phase reconstruction. The colour bar at the bottom of the figure coverts the colour of d) to f) to radians. The scale bar denotes 10μm in the selected area aperture plane of the microscope and applies to all panels in the figure.

Figure 90 displays reconstructions of the diffuser shown in Figure 82 c) (diffuser 3), reconstructed from datasets with the different levels of energy filtering. Inspecting Figure 90 it is clear that energy filtering is not required in order to obtain a visually reasonable reconstruction. The difference between the reconstructed diffuser magnitudes in Figure 90 a) and b), corresponding to the 10eV and 20eV energy slit datasets, is insignificant. However, it is clear that Figure 90 c) differs notably from Figure 90 a) and Figure 90 b), as there is clear noise outside of the diffuser aperture. As this dataset was not energy filtered this noise is directly linked to the inelastic scattering of the specimen and the diffuser. In terms of phase of the probe, the reconstructions are close to identical, apart from the low frequency phase features at the edge of the aperture and additional phase wrapping. Other experiments have suggested that this variance in the probe phase occurs regardless of the energy filtering (see section 4.11). The work in Section 4.11 suggests that this phase wrap near the edge of the aperture (see Figure 90 d-f)) may be due to the diffraction patterns being non-circular, whilst the diffusers are circular, meaning that the reconstruction algorithm develops a strong phase to account for this mismatch. One hypothesis is that the aberrations/distortions in the projector lens system are causing these non-circular diffraction patterns. If this is true, it suggests rPIE is accounting for imperfections of the lenses.

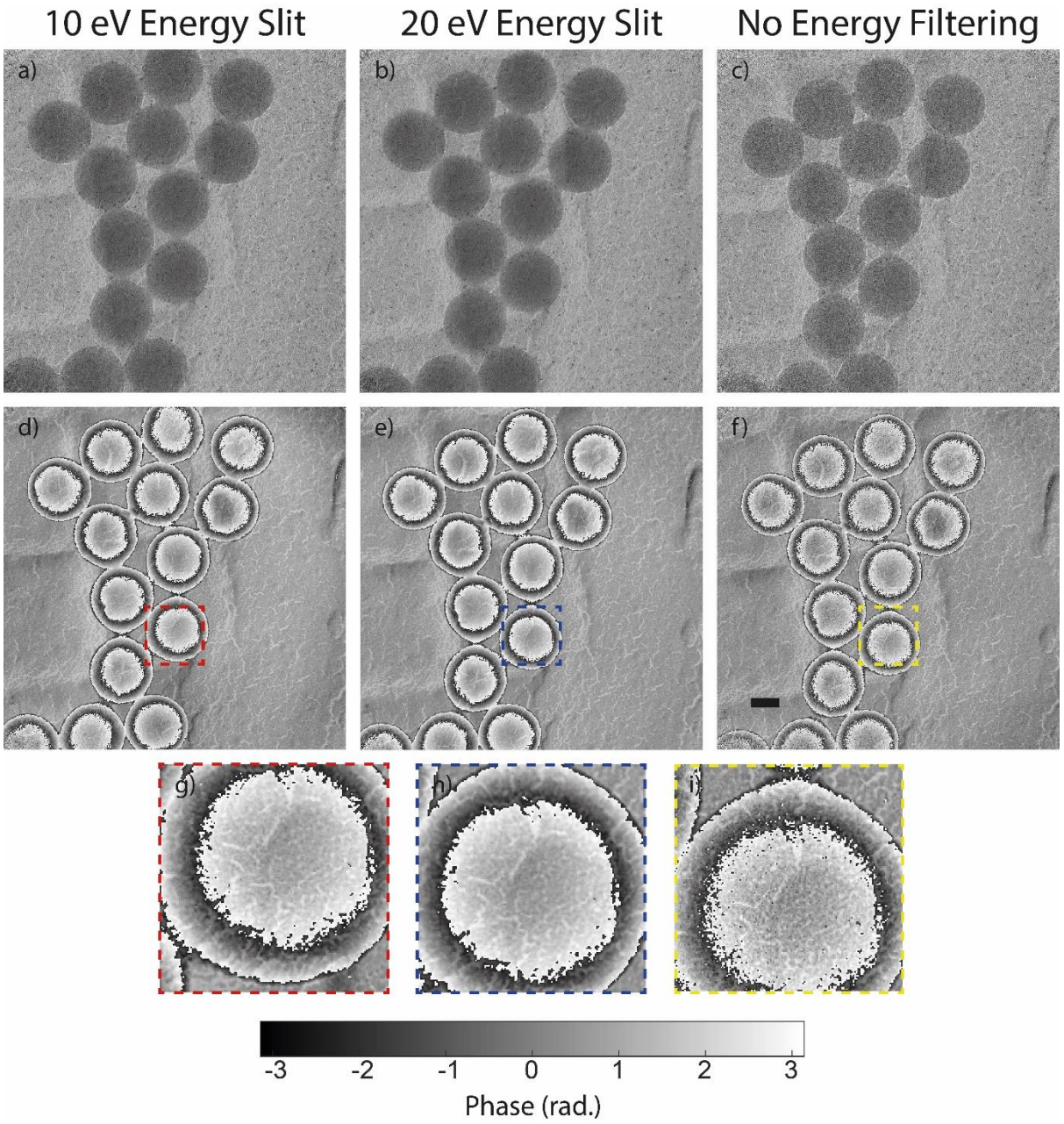

Figure 91: Reconstructed specimens with differing levels of energy filtering. a) to c) magnitude reconstruction. d) to f) phase reconstruction. The colour bar at the bottom of the figure converts the colour of d) to f) to radians. g) to i) show zoomed in subsections of individual latex spheres which correspond to dotted lines in d) to f) which share the same colour. The scale bar in f) represents 100nm and applies to a)-f) in this figure.

Figure 91 shows how the specimen reconstruction varies as the filter slit width increases. Again, the results suggest that energy filtering is not necessarily required, but its absence can be seen in the fine details of the reconstruction. For example, if an individual latex sphere is inspected, in the energy filtered case the carbon support can be seen through the latex sphere, whilst this is more difficult in the non-energy filter case (see Figure 91 g-i). Also, in the non-energy filtered case the phase wrapping is far noisier compared to the energy filtered case. While these are minor issues, they are worth noting due to the limited specimen selection of this experiment; it is not clear yet to what extent this problem is exacerbated as either the specimen or diffuser become thicker.

Now the effect of a lack of energy filtering has been seen, methods to model and remove the background noise can be considered. The first method that can be used to deal with the noise due to inelastic scatter from the specimen and diffuser is the thresholding method [185], where a fixed pedestal is removed from the recorded data. This method is described in equations 122 and 123.

$$dp_{NR} = dp - B \tag{122}$$

$$dp_{NR} = \begin{cases} dp_{NR}, dp_{NR} \geq 0 \\ 0, dp_{NR} < 0 \end{cases} \tag{123}$$

The parameters of equations 122 and 123 are the following:

- dp_{NR} is the thresholded diffraction pattern
- dp is the original diffraction data (typically an array of 1024 by 1024 intensity values)
- B is the estimated average intensity of the background noise (typically a set [single/non-matrix] value, for example 100, hence different to $Bkgd$ which is a matrix equal in size to the collected diffraction patterns)

A second method to account for the noise of inelastic scattering is background modelling. Background modelling is an additional step included in the reconstruction code, where before the error between the simulated intensity and collected intensity is measured, a modelled background is added to the simulated intensity [186]. This modelled background is obtained and updated by either taking the difference or the quotient of the simulated intensity compared to the measured intensity. This background modelling requires that the forward model closely matches the diffraction condition of the experiment or it will fail. In this experiment that means correct incoherence modelling, correct positions and correct propagation, so that background modelling can achieve its goal.

$$Bkgd_{n+1} = (1-\gamma)Bkgd_n + \gamma Bkgd_n \left(\frac{M}{I+\delta}\right)^2 \tag{124}$$

The parameters of equation 124 are the following:

- $Bkgd_{n+1}$ is new updated background model
- $Bkgd_n$ is the previous or initial background model
- γ is the tuning parameter that determines the update speed
- M is the intensity of the collected diffraction
- I is the simulated intensity
- δ is a small-value constant that prevents divisions by zero

equation 124 describes how the quotient obtained by dividing the measured intensity by the modelled intensity is used to update the current model of the background. The quotient effectively informs the algorithm where the strength of the background needs to be increased, because in these areas the quotient will be above 1 in value. γ in equation 124 is typically used to slow the update rate of the background, such that it does not account for the whole of the measured intensity (γ values used in this paper were normally 0.01).

Although the background is required to calculate the error of the simulation, this is typically updated at the end of every diffraction pattern update (typically after equation 21 but before equation 22 of Section 1.24). As at this point the diffraction pattern constraint has already been applied. Therefore, we have the current best model of the diffraction pattern which has all inaccuracies and alteration accounted for (incoherence, drift and positions), which makes it reasonable to assume that any

difference between the modelled diffraction pattern and collected diffraction pattern is due to unmodelled inelastic scatter or background noise, or any other aspect which has not been directly addressed. As this background modelling requires that all other possible issues are dealt with beforehand, the updating of the background is typically delayed by around 100 iterations in order that reconstruction is stable and the background does not account for something it was not designed to fix.

To test the background modelling code, an artificial background image was added to the dataset which was energy filtered by a 20eV energy slit (See Figure 92 a)). The background was added by generating a matrix where the intensity varied in a Poisson manner and the mean was equal to 10% of the total counts of the diffraction pattern being considered. This random matrix was then multiplied by the normalised artificial image shown in Figure 92 a) and then added to the diffraction pattern. This process was repeated for every diffraction pattern. Figure 92 c) shows the modelled background obtained from the reconstruction code. As can be seen via inspection of Figure 92, the background modelling code has managed to attribute the artificially added intensity to the background and the reconstructed specimen is unaffected (see Figure 92 b)), suggesting it is capable of doing the same with any inelastic counts from the specimen or diffuser. The artificial background seen in Figure 92 a) was chosen as it could be easily identified in either the specimen reconstruction or recovered background, allowing it to be quickly determined whether the background modelling is operating correctly.

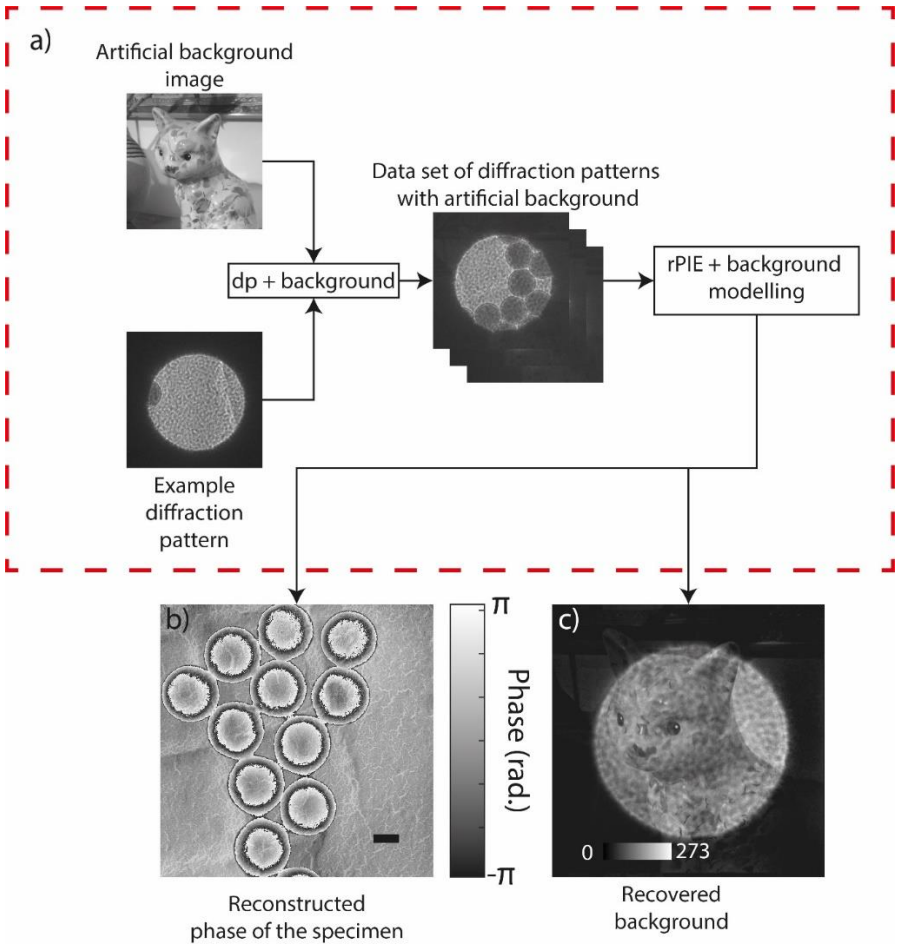

Figure 92: a) a block diagram illustrating the process by which the background correction algorithm was tested. In the top left of a) is the simulation image added to every diffraction pattern after it had been modulated by Poisson distribution with mean equal to ten percent of the diffraction pattern intensity. b) the reconstructed specimen phase obtained from the dataset which has artificial noise added to it. c) the recovered background when using equation 124.

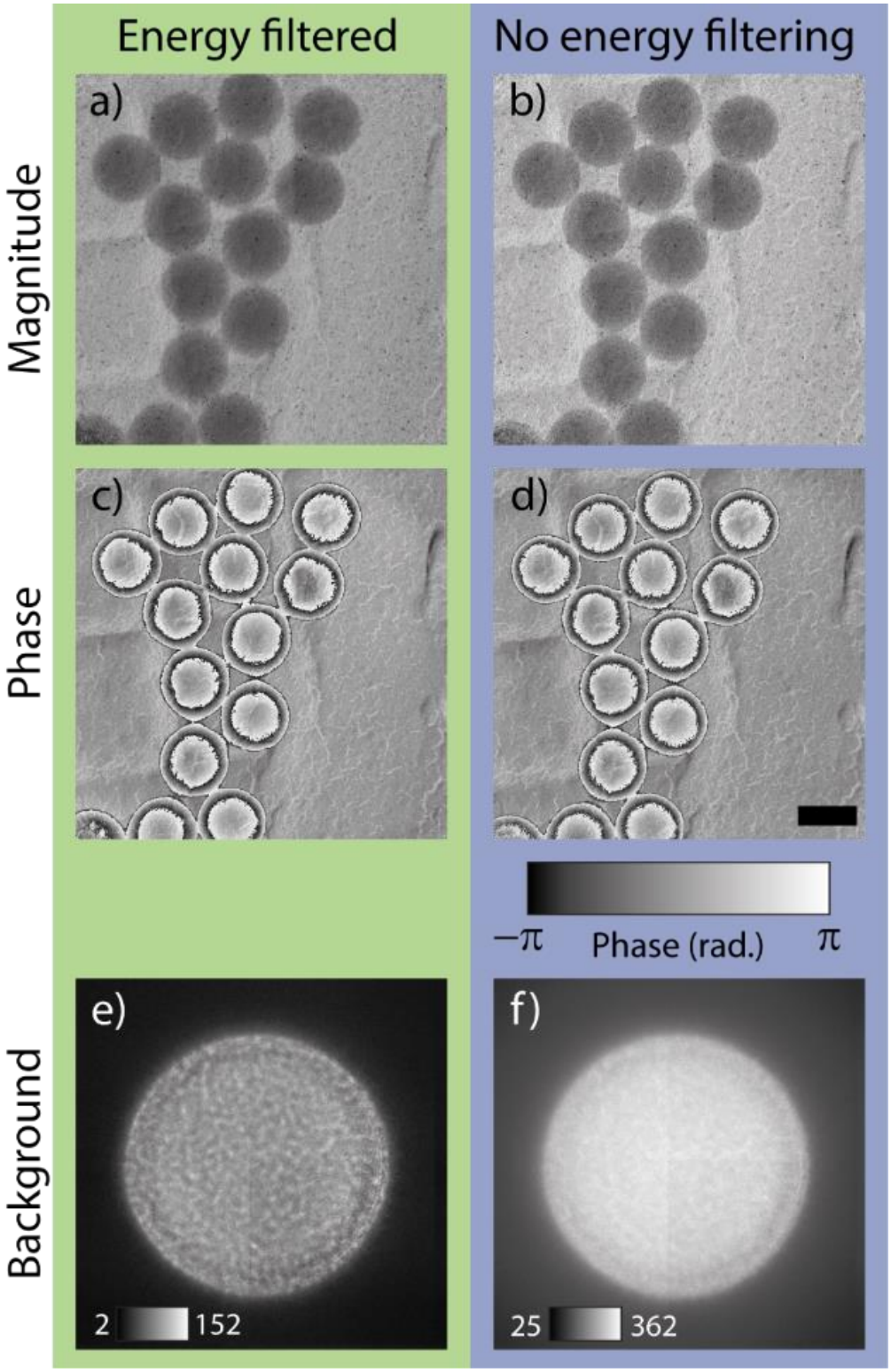

Figure 93: a) and c) magnitude and phase reconstruction respectively of the 261nm-diameter latex spheres on 463 nm-periodicity diffraction grating replica, where the dataset was both energy filtered and background compensation was applied. b) and d) magnitude and phase reconstruction of the sample from a dataset which was not energy filtered, but background compensation was applied. e) and f) modelled background obtained from the compensation algorithm. The scale bar in d) is 100 nm and applies to a) to d) [179].

Figure 93 shows both an energy filtered and non-energy filtered dataset reconstructed with both threshold removal and background modelling. As can be seen from inspection of Figure 93 a) through d) there is little difference between the non-energy filtered and the filtered reconstructions. Furthermore, the form of the modelled background in Figure 93 e) and f) suggests some degree of success, as they display quadrant variance similar to that found in gain reference images. In addition, the greater strength of the background in the non-energy filtered case suggests that background modelling had to account for additional noise because of the inelastic scatter.

In order to assess the effect of the energy filtering on the obtained resolution of the collected datasets, Fourier Ring Correlation (FRC) was deployed to measure the resolution of the reconstructions. Egerton has directly related inelastic scatter (a form of temporal incoherence) to resolution loss [169], therefore it will be useful to quantify the resolution loss for near field electron ptychography when the dataset is not energy filtered. Because the step size in the dataset used to reconstruct Figure 91 is small (100nm), it can be split into two separate datasets, one containing the odd diffraction patterns and the other containing the even diffraction patterns, and reconstructions from these two independent sets of data can be directly compared. The Fourier Ring Correlation result for the case where the non-energy filtered dataset is reconstructed with background modelling will be compared with reconstructions from data collected using a 20-eV energy slit.

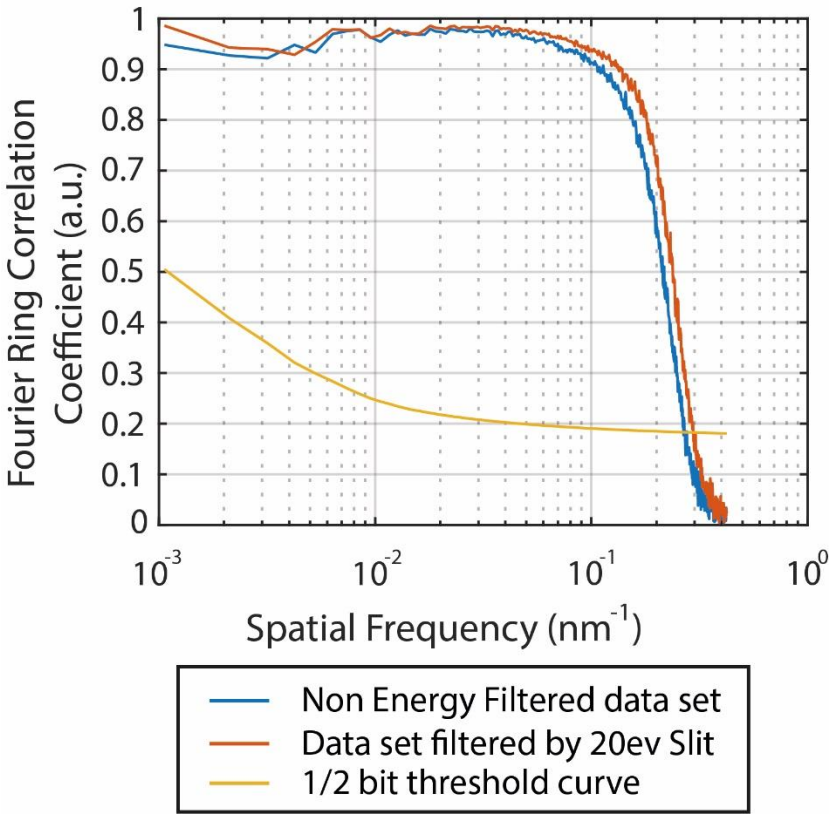

Figure 94: Fourier Ring Correlation analysis of the energy filtered and non-energy filtered datasets. An additional curve describes the Fourier Ring Correlation of unfiltered and non-background modelled reconstruction. A 1/2-bit threshold curve is also included as reference to display the difference in signal to noise ratios between the three correlations. The intersection of this curve with the correlation curves gives an indication of resolution in the reconstruction: for the reconstruction using unfiltered data this calculation gives a resolution of 3.73nm, and for the filtered data the resolution is estimated at 3.36nm [179].

Figure 94 displays the calculated Fourier Ring Correlation for the energy and non-energy filtered cases, where the background modelling was deployed. It shows that despite the background modelling, the energy filtered case has a slightly better resolution. The energy filtered case has a ½

bit threshold resolution of 3.36nm ($0.298nm^{-1}$), while the non energy filtered case has a resolution of 3.73nm ($0.268nm^{-1}$). The half bit threshold is used here, as it is a widely accepted signal to noise ratio which phase images have to reach in order to become valid. The half bit threshold has been described in section 1.40 [187, 188]. Figure 94 directly suggests that inelastic scattering places some limit on the obtainable resolution of near field electron ptychography. Additionally, it can be seen from Figure 94 that even in the energy filtered case the resolution is limited, therefore maximum resolution is not just determined by the inelastic scatter. Consequently, finding the resolution limiting aspects of the diffraction condition is an important next step.

To investigate the quantitative nature of non-energy filtered near field electron ptychography reconstructions, the latex spheres in Figure 93 d) were measured using the code described in Section 3.7. Both the energy filtered and non- energy filtered reconstructions of Figure 93 give values for the mean inner potential of the latex spheres within one standard deviation of previous results [179]. The mean inner potential for the energy filtered reconstruction was 8.37 V with a standard deviation of 0.43 V, and for the non-energy filtered case the mean inner potential was 8.57 V with a standard deviation of 0.34 V. This suggests that near field electron ptychography is quantitative regardless of energy filtering. Figure 95 a) shows the unwrapped phase of Figure 93 d), each of the latex spheres are numbered in Figure 95 a). The number on each of the latex sphere corresponds to the plotted unwrapped radially averaged phase, shown in Figure 95 b). The unwrapped phase seems to match well with the modelled thickness further suggesting that the phase recovered by near field electron ptychography is quantitative. Individual values of mean inner potential for each of the latex spheres can be seen in Table 7 and Table 8 for the non-energy filtered and energy filtered reconstructions respectively.

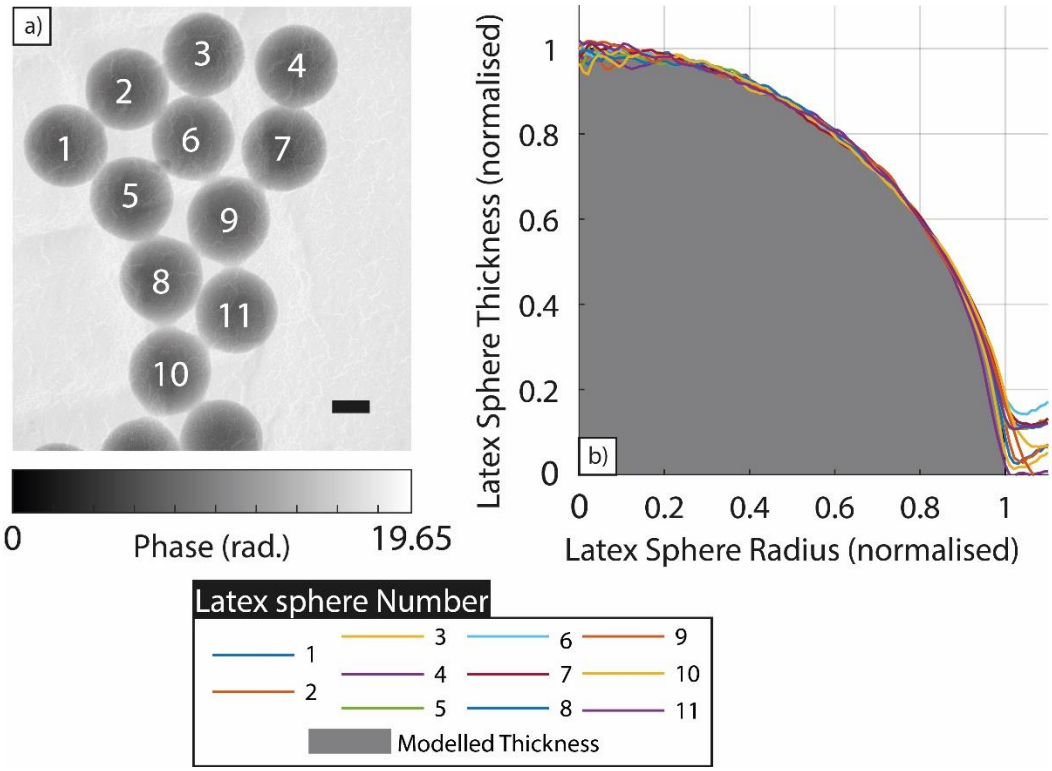

Figure 95: a) A figure showing the unwrapped version of the phase shown in Figure 93 d) (non-energy filtered). The colour bar under a) converts the greyscale of a) into unwrapped radians. b) A figure showing a series of plots of unwrapped radially averaged latex sphere phase profiles, for each of the latex spheres numbered in a). The grey area in b) indicates how the modelled thickness changes as a function of radius in the case of an ideal sphere. The scale bar in a) indicates 100nm.

Latex sphere number	Mean inner potential (volts)	Latex sphere radius (nm)
1	8.69	114.59
2	8.32	113.11
3	8.21	112.63
4	8.01	113.84
5	8.83	114.66
6	8.48	113.00
7	9.20	115.29
8	8.89	115.94
9	8.34	113.73
10	8.61	112.45
11	8.71	113.54

Table 7: a table showing latex sphere number, mean inner potential and measured latex sphere radius in nm obtained from the non-energy filtered phase image shown in Figure 93 d).

Latex sphere number	Mean inner potential (volts)	Latex sphere radius (nm)
1	8.7	118.81
2	8.19	115.41
3	8.39	116.29
4	7.49	113.59
5	8.72	115.86
6	8.22	115.41
7	9.03	117.93
8	8.53	115.85
9	7.85	114.42
10	8.55	116.04
11	8.42	114.59

Table 8: a table showing latex sphere number, mean inner potential and measured latex sphere radius in obtained from the energy filtered phase image shown in Figure 93 c).

4.8 Large Field of View experiment

Now that it has been shown that non-energy filtered datasets can be reconstructed to provide acceptable reconstructions, larger datasets can be collected as the method is no longer limited by the drift of the energy filtering slit. It is important to investigate the largest possible field of view that can be obtained using the near field electron ptychography method, since this would allow near field electron ptychography to demonstrate that it provides images that are difficult for other phase retrieval methods implemented in TEM.

For the large field of view experiment, the specimen was moved in a grid like pattern, with 39 positions in the *y*-direction and 40 in the *x*-direction. The step size was 250nm. This experiment had the same diffraction condition as the previous experiments, which means that the resolution of the large field of view dataset should have a similar resolution to that measured in the FRC (Fourier Ring Correlation) experiment, as the datasets in these experiments have similar step sizes of 250nm and 200nm respectively. This experiment inherently tests the reliability of the data collection code and

the specimen stage via the sheer scale of the dataset requested. Again, the mean inner potential of the latex spheres will be used to test whether the results of the reconstruction are quantitative.

The result shown in Figure 96 is a 66-megapixel image with a field of view of 93.4μm². This means that each diffraction pattern on average contributes 0.0413 megapixels and 0.0584μm² to the reconstruction. Looking at Figure 96 e) it can be seen that all latex spheres have reconstructed successfully, but inspecting Figure 96 c) it can be seen that one of the latex spheres has melted and the other two have dark halos. It is suspected that these dark halos are due to the charging of the latex spheres, but this is subject to further investigation as it beyond the scope of this experiment. The measured mean inner potential for each of the uncharged latex spheres is shown in Table 9. The average of the mean inner potential in the table is 8.73±0.696V. Inspecting the measured radii in Table 9 and understanding that the latex spheres were originally 131nm in radius, clearly the latex spheres have shrunk in size under exposure to the electron beam [149, 189, 190]. This effect has been documented in [149, 190] but does not seem to affect the measurement of the mean inner potential in this case. It can also be seen that the phase unwrapping of the latex spheres may have caused the mean inner potential to be calculated incorrectly. This is illustrated in the latex spheres numbered 4 and 5 in Figure 97. In common with previous sections, the radial average of the unwrapped phase of various latex spheres were plotted as a function of normalised radii and compared to a shaded area representing the modelled thickness of a normalised ideal latex sphere (see Figure 97 b)). Due to the large number of latex spheres, only a sub-set is plotted in Figure 97 b), the sub-set shown was chosen as they were shown in Figure 85, allowing for comparison. Compared to their previous plots some of the latex spheres do not match the modelled thickness as well for the last 10% of the normalised radius, as the unwrapped radially averaged phase is greater than that of the ideal modelled thickness. This could be due to the poor choice of the centre of the latex spheres or to the latex spheres changing shape during exposure under the electron beam, possibly becoming more lens like, but for most of the plot there is good agreement between the modelled and measured thickness.

The reason for the development of near field electron ptychography was to maximise the field of view per diffraction pattern, so that large field of view experiments could be performed in less time compared to the focused probe ptychographic implementation. However, examining the probe reconstruction of the large field of view dataset (see Figure 105 h)), it is still not taking up the whole field of view of the detector. The detector field of view could not be filled by the illumination due to the drift of the optical axis, as described in section 3.6. A comparison between the collected dataset and a hypothetical dataset, where the diffuser fills the field of view, now follows. The ratio of overlap to step size in the above experiment can be estimated by dividing the step size by the effective probe diameter at the specimen, as seen in equation 125. Applying that overlap ratio to the effective size of the detector at the specimen gives the equivalent step size, for the case where the diffuser covers the whole of the detector as calculated in equation 126. This will be called the detector step size to differentiate it from other parameters. Using the calculated detector step size the Detector Limited Field Of View (DLFOV) can be calculated, as shown in equation 127; the value of the DLFOV is purposely underestimated for ease of calculation. Comparing the result from equation 127 and to the 93.4μm² area obtained from the above experiment, it can be seen that the field of view would more than double if the whole detector was utilized. The benefit of using the whole of the detector could also be converted to improve the reconstruction (decreased step size and increased signal to noise ratio) or reduce the data collection time, through reducing the number of diffraction patterns.

$$Overlap\ Estimate = \frac{Step\ Size}{Probe\ Diameter} = 1 - \frac{250 \times 10^{-9}}{875 \times 10^{-9}} = 0.714\ or\ 71.4\% \quad (125)$$

$$\begin{aligned} \text{Detector Step Size} &= \text{Overlap Estimate} \times \text{Detector Width} \\ &= 0.286 \times 1024 \times 1.18 \times 10^{-9} \\ &= 345 \times 10^{-9} nm \end{aligned} \quad (126)$$

$$\begin{aligned} \text{Estimated DLFOV} &= (\text{Detector Step Size} \times \text{Number of steps})^2 \\ &= \frac{(345 \times 10^{-9} \times 40)^2}{1 \times 10^{-12}} \approx 191 \mu m^2 \end{aligned} \quad (127)$$

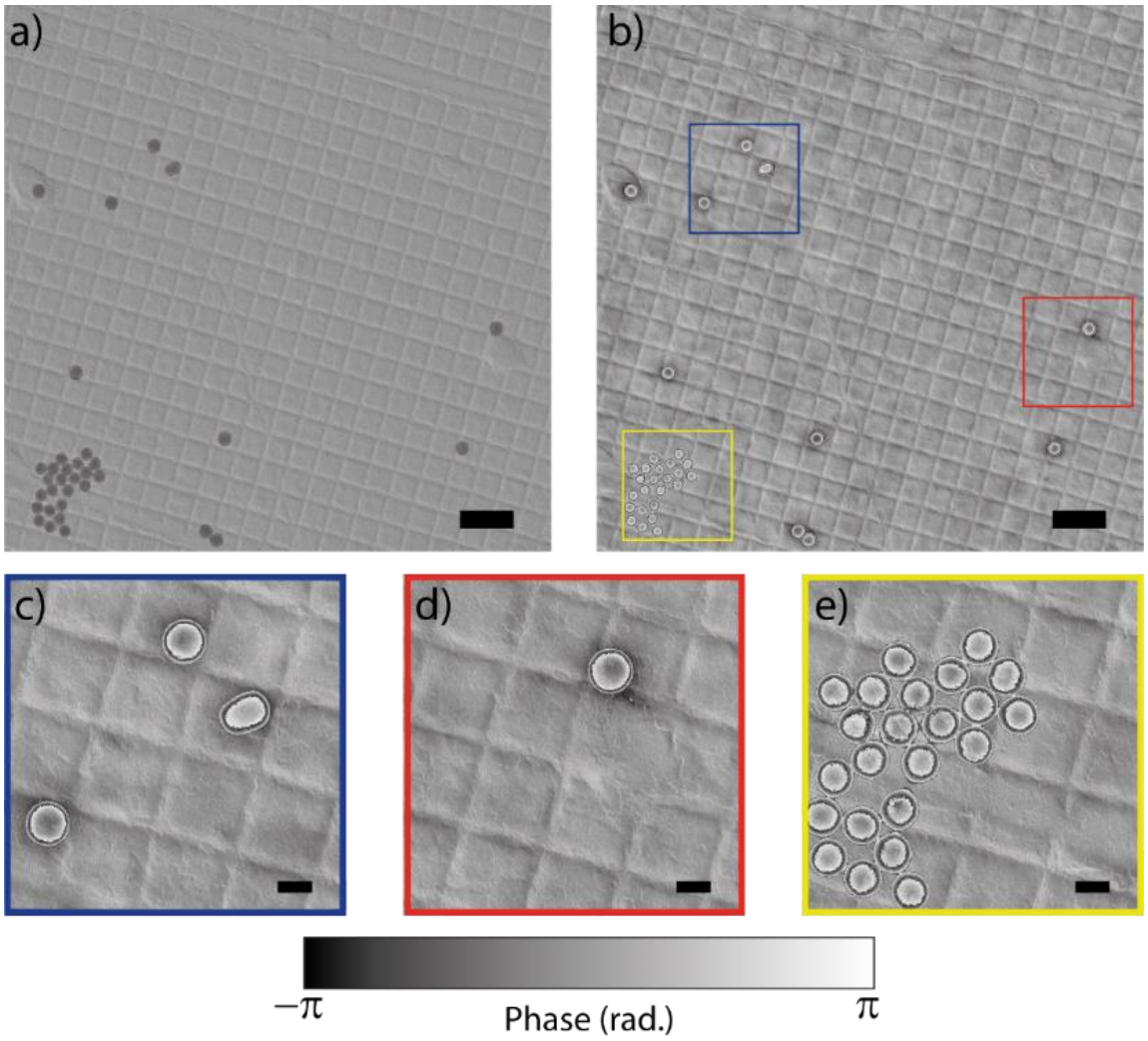

Figure 96: a) Magnitude and b) phase of the reconstructed sample. c)-e) Enlarged subsections of b), each corresponding to the boxes in b) with the same colour. Note the halos around the latex spheres in c) and d), which are not present around the cluster of spheres in e). Note also the melted sphere in c). Scale bars in a) and b), 1 µm. Scale bars in c)-e), 200 nm.

Figure 96 shows that datasets with up to 1600 diffraction patterns can successfully be collected and reconstructed by the data collection and reconstruction code, respectively. The diffraction pattern drift was manually corrected once through the data collection process. Manual correction of the diffraction pattern drift mid data collection has been shown not to compromise reconstruction quality; this was demonstrated in Section 3.6. As this dataset took 1 hour and 30 minutes to collect, in theory near field electron ptychography without energy filtering can be stable for multiple hours

with additional diffraction pattern drift corrections. It can also be assumed that this dataset has an approximate resolution of 3.73nm, as this dataset has a similar step size to the previous section where the effect of energy filtering on the resolution was assessed (see Section 4.7).

Now that near field electron ptychography has demonstrated its capabilities in terms of field of view, this will be compared to other methods in a crude way. Defocused probe ptychography considers a large field of view to be of the order of 1600 by 1600 pixels, as demonstrated by the work of Zhou et al. [92], while STEM considers an image of 2048 by 2048 pixels to be a large field of view as demonstrated in [163]. STEM scan generators can technically produce 8k by 8k scan patterns, but this capability is not often used due the long data collection time combined with drift of the specimen during this period leading to image being effected by scan distortions [191]. Finally, off-axis holography considers a 530 by 200 pixel phase image to be a reasonable field of view [120]. Whilst not attempting a quantitative comparison, as these methods did other things apart from providing a large field of view, these references provide the necessary context to show that near field electron ptychography does in fact provide large field of view images with quantitative information as is shown by the 66 megapixel image shown in Figure 96.

Figure 96 highlights near field ptychography's niche, as near field ptychography has the theoretical advantage of providing a greater field of view per diffraction pattern than other ptychographic counterparts. This has also been noted in the X-ray domain by Sala et al. [160]. Figure 2 of their paper shows that near field ptychography can collect a set field of view in the shortest time frame whilst retaining a reasonable resolution. Sala et al. also explained that their paper is meant for guidance, so further experiments are required to confirm that near field electron ptychography occupies this same niche in the electron domain. But once the niche is confirmed it will then be a case of applying near field electron ptychography to specimens where large field of view of phase images provide an advantage or change in the status quo.

One of the applications for micrometre fields of view is the measurement of strain fields in semiconductor devices [163], as strain fields help improve the performance via enhancing the mobility of the carriers, and therefore it is important to check that the strain has not been compromised by the inclusion of defects. While it is important to note that near field electron ptychography has not yet demonstrated that it is capable of dark field imaging which, is required by strain imaging, it does show great potential in this application of strain measurement via the field of views it can provide, as demonstrated in this section. However. traditional phase images provided by near field electron ptychography could possibly provide dopant concentrations over similar fields of view [122], as off-axis holography has already provided. This naturally leads to a speculative statement that near field electron ptychography might be capable of imaging tens of transistors at a time.

Furthermore, this extreme field of view may be helpful in performing single particle analysis [133], as single particle analysis relies on the capture of thousands of single particles which share the same orientation, such that they can be averaged to obtain an average image with greater signal to noise ratio. By repeating this process for all orientations, tomography can be performed by using different averaged orientations. Typically single particle analysis collects large field of view micrographs with typically tens to hundreds of particles present in each image (see figure 4 of [192] for an example). As near field electron ptychography and single particle analysis are similar, in that both require a defocus in order to capture phase information of the sample, this gives near field electron ptychography an advantage over other forms of ptychography, as it may be possible for near field electron ptychography to be incorporated into already existing TEMs specialised for single particle analysis (this of course requires confirmation via experiment). Ptychography in general offers many

advantages to single particle analysis, as outlined in [92] by Zhou et al, for example, reduced experiment times due to the greater signal to noise ratio of ptychography, meaning that fewer images are required for tomographic reconstruction.

Figure 96 is a demonstration of near field electron ptychography's ability to collect extreme fields of view, and illustrates the greatest advantage of near field electron ptychography over the most common quantitative phase imaging methods. For example, the field of view offered by off-axis holography is often severely limited compared to near field electron ptychography, as off-axis holography has to sacrifice its field of view in order to get the desired resolution. The resolution is limited by the process of selecting a side band in frequency space, so the resolution is increased during data collection in order to mitigate its later reduction, and the field view is then reduced to a ninth or quarter of a standard TEM image (this was highlighted in section 1.11) [46, 135]. Off-axis holography can collect multiple holograms then stitch phase images together to obtain a larger field of view, but this can be limited in turn by the sample, for example off-axis holography cannot operate with the sample presented in Figure 96, as there is no available vacuum area with which a reference wave can be formed, and milling a vacuum may incur specimen damage. However, it has been shown that off-axis holography can provide fields of view of at least a micrometre in one direction [193].

Latex Sphere Number	Mean Inner Potential (volts)	Latex Sphere Radius (nm)
1	9.81	94.08
2	9.03	93.17
3	8.38	91.45
4	7.96	96.13
5	10.00	94.71
6	9.77	91.48
7	7.26	92.42
8	7.36	89.82
9	8.99	95.84
10	9.30	94.73
11	8.79	96.5
12	8.49	97.65
13	8.23	89.96
14	8.71	95.99
15	8.44	95.29
16	8.87	91.17
17	8.82	95.02
18	9.47	92.4
19	8.83	92.69
20	8.50	91.22
21	9.07	94.01
22	8.11	91.33
23	8.56	94.87

Table 9: measured mean inner potential of latex spheres seen in Figure 96 e). The latex sphere number corresponds to the labeling of latex spheres in Figure 97. Additionally the measured radius of the latex sphere is given.

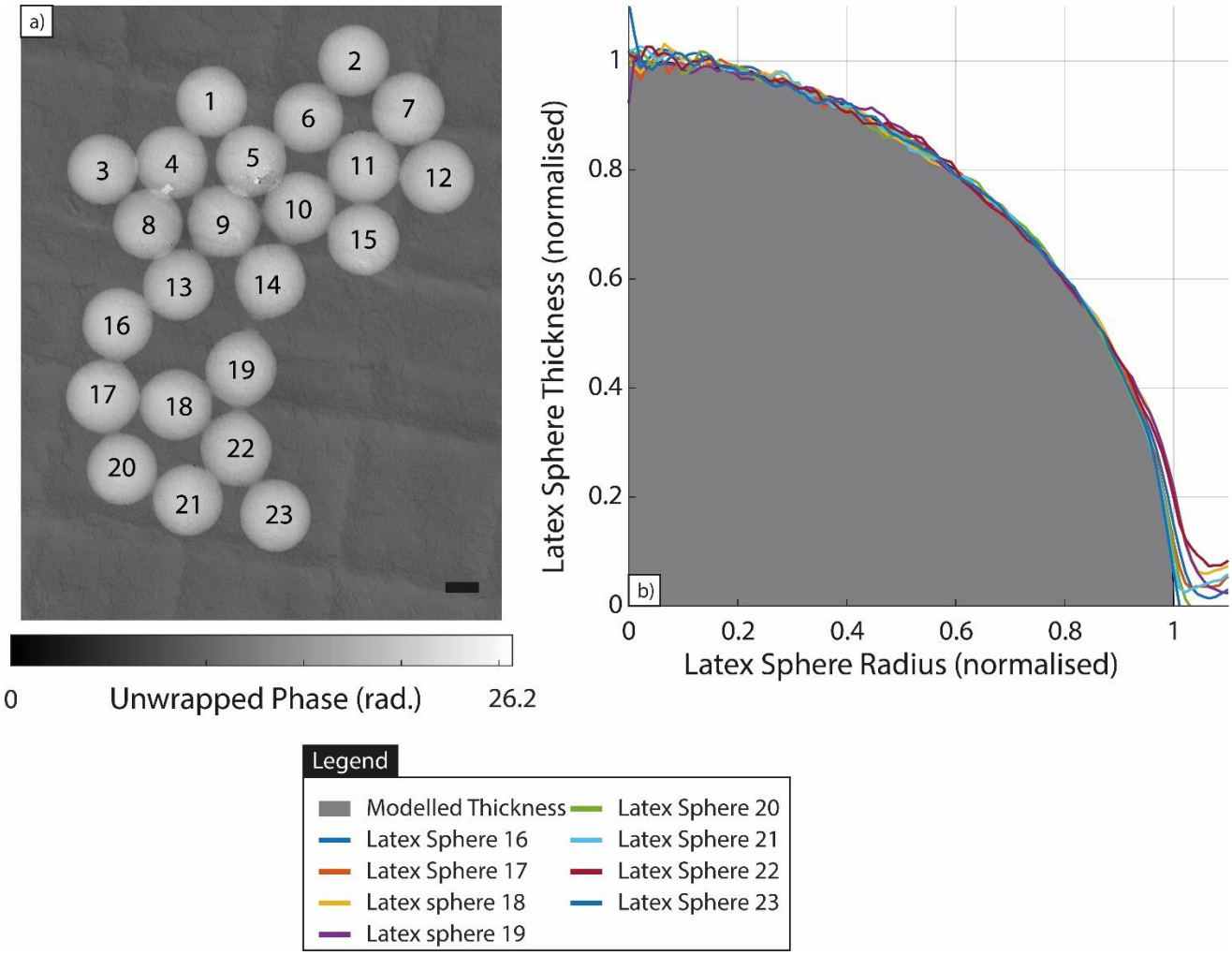

Figure 97: a) a figure showing the unwrapped phase of the latex spheres shown in Figure 96 e), the number over a latex sphere correspondes to the latex sphere number in Table 9. The sign of the unwrapped phase in a) was inverted to obtain a positive mean inner potential. b) a plot showing the radial average of the phase of varoius latex spheres with respect to the normalised radial average (the number in the legend corresponds to latex sphere number in a)). In b) a modelled thickness is plotted as well for reference (see Section 3.7 for more details). The colourbar at the bottom of the figure a) coverts the greyscale of the image to phase in unwrapped radians. The scale bar in a) is 100nm.

4.9 Magnesium oxide cube phase imaging

A further experiment was carried out to test near field electron ptychography with other specimens, to see if issues arise. One of the specimens chosen was magnesium oxide cubes (MgO cubes) as they have well defined dimensions allowing further quantitative analysis. Furthermore, MgO cubes are capable of obtaining significant charge such that an electric field can be measured in the vacuum around them. An electric field of a charged specimen would allow initial analysis of the sensitivity of near field electron ptychography. Furthermore, the MgO cubes imaged are mostly surrounded by a vacuum, a situation advantageous to off-axis holography but disadvantageous to near field electron ptychography, as will be explained later.

The data collected consisted of a 9 by 10 square scan with a step size of 100nm. The MgO cube is found at the lower left of the scan. The diffraction condition is the same as the original experiment investigating the minimum number of diffraction patterns. This dataset was not energy filtered. The

reconstructed specimen from this collected dataset is seen in Figure 98 and the reconstructed probe is seen in Figure 99.

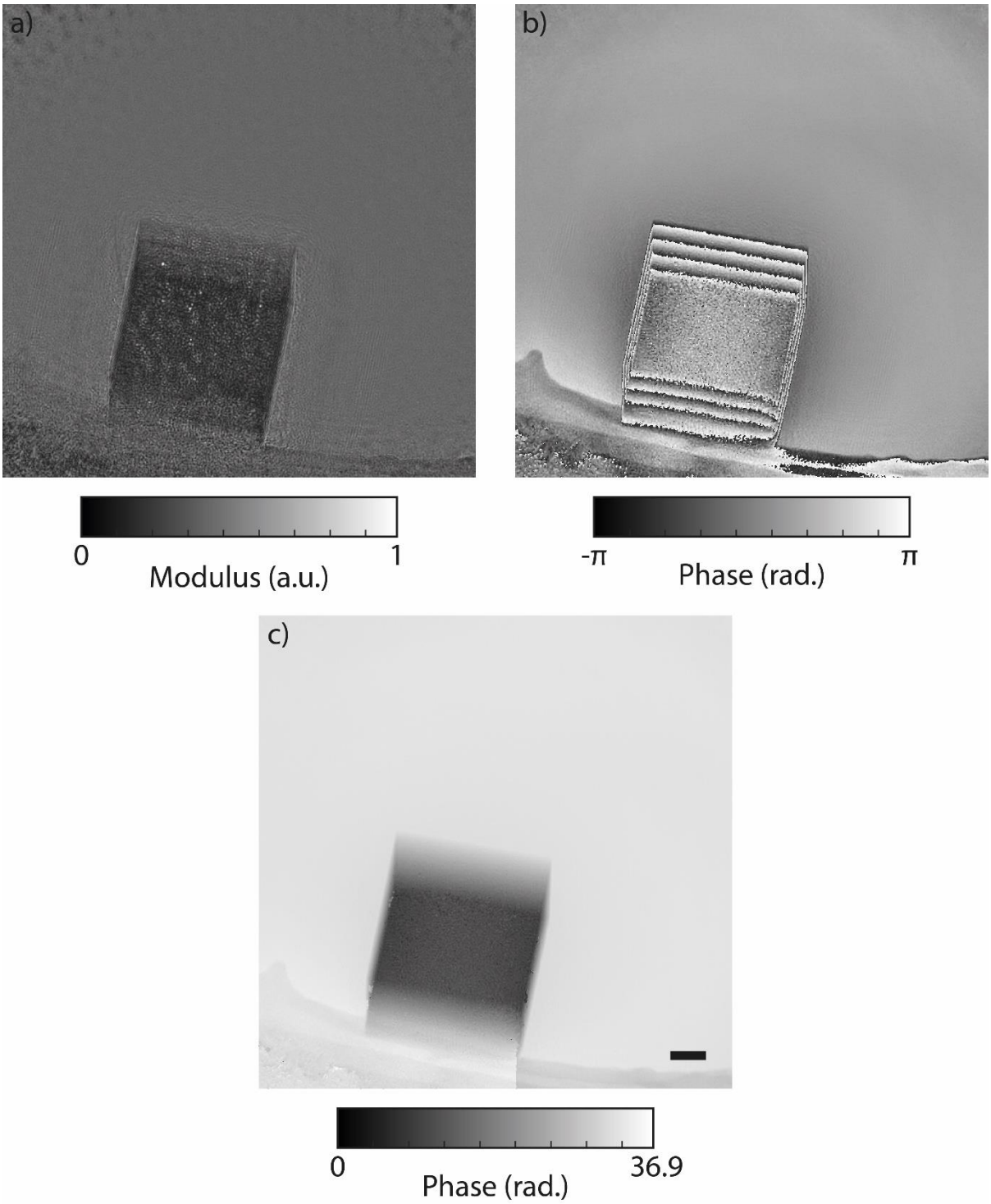

Figure 98: a) the magnitude reconstruction of an MgO cube. b) phase reconstruction of MgO cube. c) the unwrapped phase of the MgO cube. The colour bar under a) converts the greyscale of a) into a normalised unitless magnitude. The colour bar under b) converts the greyscale of b) into phase in terms of radians. The colour bar under c) converts the greyscale of c) into unwrapped phase in terms of radians. The scale bar in c) is 100nm in length.

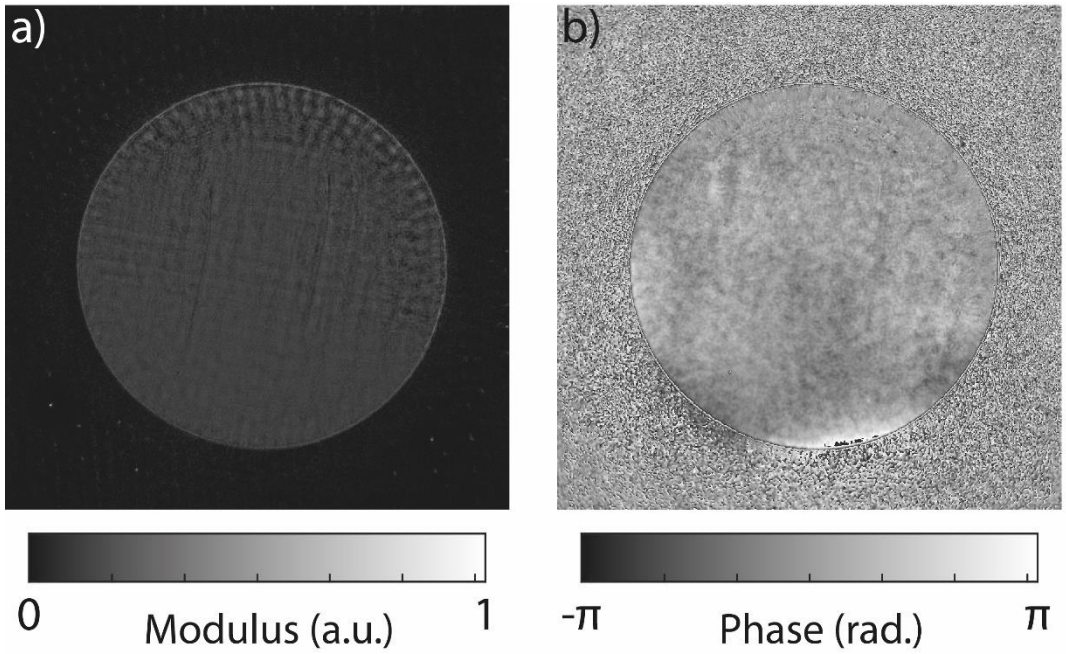

Figure 99: a) magnitude reconstruction of the diffuser obtained when processing the MgO cube dataset. b) phase reconstruction of the diffuser obtained when processing the MgO cube dataset. The colour bar on the right converts the greyscale of the right to phase in radians.

Analysis of the phase of MgO cubes has been performed before in [194]. The reconstruction seems to have gone reasonably well, as the edges of the cube can be seen and the phase of the reconstruction seems well defined. The centre of the MgO cube is quite noisy, which is likely due to a combination of factors, one of which is that the probe has not been completely reconstructed, but that will be discussed later. The high level of noise in the reconstruction is also indicated in the unwrapped phase image Figure 98 c), as there is a disconnect in the unwrapped phase at the bottom of c) between the left and right of the image. This disconnect in the unwrapped phase is due to the branch and cut algorithm resolving a phase residue with the edge of the image (see section 3.8 for details). Based on the relative darkness (absorption) of the specimen, it can be assumed that the MgO cube has caused a large degree of inelastic scatter. Looking at the phase image of the MgO cube it is clear this is a strong phase specimen, as there are several phase wraps within it. Inspecting the edge of the MgO cube there is some slight change in the phase, indicating that the MgO cube may have obtained some charge from the electron beam. The exact structure and orientation of the MgO cube is not obvious from its phase image (Figure 98) making direct analysis difficult.

Another difficulty experienced while imaging the MgO cube was that the dose on the MgO cube was sufficiently high to cause the MgO cube to move and morph during the imaging process. Considering this, the phase image reconstructed is surprisingly good.

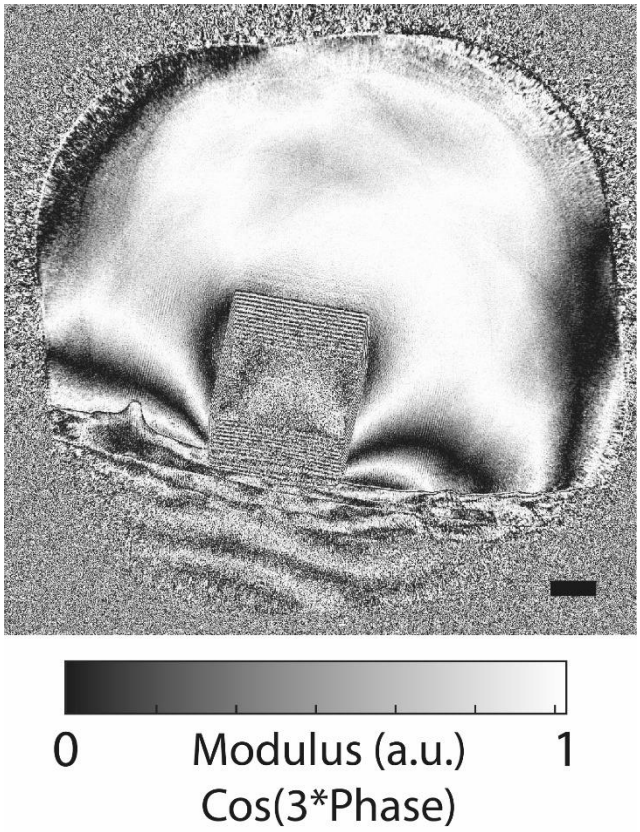

Figure 100: Phase reconstruction from Figure 98 b) multiplied by three then cosined. The scale bar indicates 100nm.

By calculating the cosine of 3 times the phase image of the specimen, the charging of the MgO cube (and the lacey carbon it is attached to) becomes clear. However, the reconstruction artefacts also become clearer. Examples of reconstruction artefacts are: the imprints of the probe edge in Figure 100, and the phase at the edge of the reconstruction which seems unattached to any underlying cause. This experiment is only an initial investigation and needs additional experiments to confirm the results.

This brings us to the discussion of the difficulty this specimen causes to near field electron ptychography. Firstly, as many of the diffraction patterns collected contained only vacuum, these diffraction patterns had to be removed from the dataset. This is because it is currently impossible to find the correct scan position via cross-correlation if the corresponding diffraction pattern has no visible specimen (i.e., only vacuum). Of the 90 diffraction patterns collected, only 30 were used. Furthermore, there was a lack of diversity in the collected datasets, as the top of the diffuser never interacted with the MgO cube or its carbon support. The effect of a lack of interference on the probe reconstruction can be seen in the top of the magnitude reconstruction of the probe in Figure 99, as the edge of the aperture is not sharp and the magnitude varies. This does not happen in any other reconstruction of this diffuser; for an example see Figure 90.

Whilst this dataset has problems that both inhibit quantitative analysis and reduce the reconstruction quality, there are some simple alterations that would improve this dataset. The first change would be to ensure that all parts of the diffuser interact with the MgO cube. This could be done by increasing the amount of magnification which is provided by the objective lens, so that the field of view at specimen plane is reduced but the area of the diffuser at the detector plane remains constant. A second change would be increasing the amount of lacey carbon in the scan field of view, which would also increase the reconstruction quality. Although the lacey carbon is not part of the

specimen of interest, it adds additional diversity to the dataset which can be used to interfere with the probe, making the probe function easier to retrieve. An ideal position of the MgO cube would be similar to that of the latex spheres in the previous experiments, placed on top of a carbon support. Although this position may prevent the charging of the MgO cube, or make it more difficult to directly analyse, such a specimen position would ensure that the positions given to the reconstruction code are correct. Another step to help the reconstruction process would be to initialize the reconstruction of the probe with a reconstructed probe from a different dataset, or collected diffraction pattern where the specimen was not present [195]. A final change would be to add beam blanking to the data collection process, such that the specimen structure is stable throughout the experiment.

4.10 High magnification and full field of view reconstruction

The reconstruction shown below in Figure 102 is from an experiment where the magnification provided by the intermediate lens was increased, such that the diffuser covered the whole field of view of the detector, meaning that the detector was fully utilized. Figure 101 illustrates the importance of detector utilisation. An increasing detector utilisation means more pixels are added to the reconstructed image per diffraction pattern. The increased magnification of the intermediate lens system resulted in a smaller overall field of view in this case. A greater field of view would be obtained if the objective lens magnification was reduced.

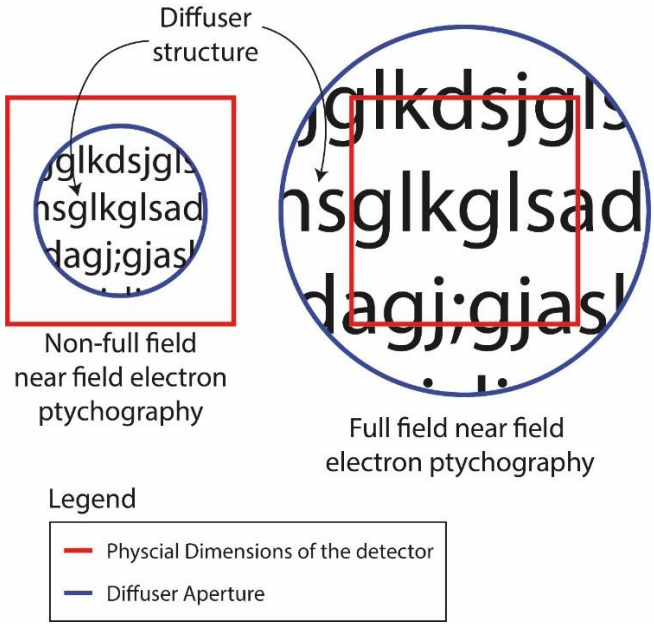

Figure 101: a figure illustrating how increasing the magnification of the intermediate lens system increases the utilisation of the detector when using a diffuser with fixed diameter. Note also the increased feature size, possibly decreasing the diversity of the underlying collected diffraction patterns.

The reconstruction in Figure 102 shows that near field electron ptychography in principle is capable of collecting and reconstructing data from a detector that is fully utilized, which is typically how near field ptychography operates in other modalities (see Section 2.3, where the whole detector is utilised). Unfortunately, there are no latex spheres to confirm that the reconstruction is quantitative. However, the cracks in the diffraction grating replica are clearly visible in the reconstruction, suggesting that the reconstruction was successful, as the cracks are similar to those obtained in previous experiments. A reasonable point to make would be to suggest that the carbon grating replica itself could be used to confirm the quantitative nature of near field electron ptychography at high magnifications. However, as there is no vacuum available the phase change caused by the

average thickness of the diffraction grating is inaccessible, and therefore only the thickness of the surface topography of the diffraction grating is measurable (assuming a mean inner potential of carbon of 9.09V [196]). As the mean inner potential of the latex spheres was only accurately measured due to the well-defined shape of the latex spheres, such a well-defined shape is missing in the case of the carbon grating replica. This highlights one of the additional benefits of the vacuum reference, in that it allows direct assessment of the thickness of the material under investigation, instead of the change in thickness.

The specimen magnitude reconstruction is noisier than the previous reconstructions where the diffuser did not fill the field of view. One reason why the specimen reconstruction was noisy was the fact that the illumination intensity was unchanged from the previous experiments, as the signal to noise ratio was reduced by the fact the total counts were spread across the whole detector and beyond. Figure 103 compares the average intensity of collected diffraction patterns in the non-energy filtered dataset from Section 4.7 against the high magnification dataset being considered here. These datasets were compared as neither was energy filtered. It should also be noted that the exposure time for the high magnification dataset was 1s, while for the non-energy filtered dataset it was 0.5s. The non-energy filtered dataset also has latex spheres in the field of view, hence variance in the intensity. In Figure 103 only the intensities within the Si_3N_4 window were used to calculate the average. Figure 103 illustrates that the intensity per pixel has significantly decreased by spreading the diffraction pattern across the whole of the detector, as even with twice the exposure time and lack of latex spheres the intensity of the high magnification dataset never reaches the same value as previous datasets. Future experiments should ensure that a reasonable number of counts are collected when utilizing all of the detector and further work should investigate the required/minimum dose for quantitative phase imaging when using near field electron ptychography.

Furthermore, through inspection of the probe the designed diffuser (Diffuser 3 was used, see Figure *82* c)) is difficult to confirm as it cannot be separated from the Focused Ion Beam (FIB) etch effects at this magnification. If a different diffuser design was used this may not be the case. This may suggest that the current concerns about the feature sizes of the diffusers being a limitation to obtaining higher magnification images from near field electron ptychography are unfounded; more details can be found in Section 4.1.

Without the discovery that the dataset itself could be used to find the positions required by the reconstruction, which is achieved via cross-correlation as explained in Section 3.3, the full field of view dataset was difficult to reconstruct. More positively, this result shows that the effect of diffraction pattern drift may be less severe than initially thought, as no additional processing to account for drift was used to get this reconstruction. This went against expectations, since it was suspected the diffraction pattern drift would have a severe effect in this case as it would change diffraction pattern content instead of simply shifting it. The diffraction pattern drift may have reduced the resolution by blurring the diffraction fringes. This reconstruction also suggests that a doubling of the pixels added per diffraction pattern, as discussed in the large field of view experiment section, is possible and can be even more plausible with the changes suggested at the end of this section.

The step size was 80nm and the scan pattern consisted of a 9 by 10 grid scan. An estimate of the pixel size based on the single square of the grating replica gives an approximate value of 0.329nm. This pixel size, along with constant magnification of the objective lens (59 times magnification), suggests that the post objective lens system magnified both the specimen and diffuser by roughly 1400 times. The step size as percentage of the detector field of view is estimated to be 24%

($100 \times \frac{80\times10^{-9}}{0.329\times10^{-9}\times1024}$). It is clear looking at Figure 102 that the resolution is lower than that permitted by the detector, suggesting that the resolution limiting factor must be addressed before near field electron ptychography can reach its full potential.

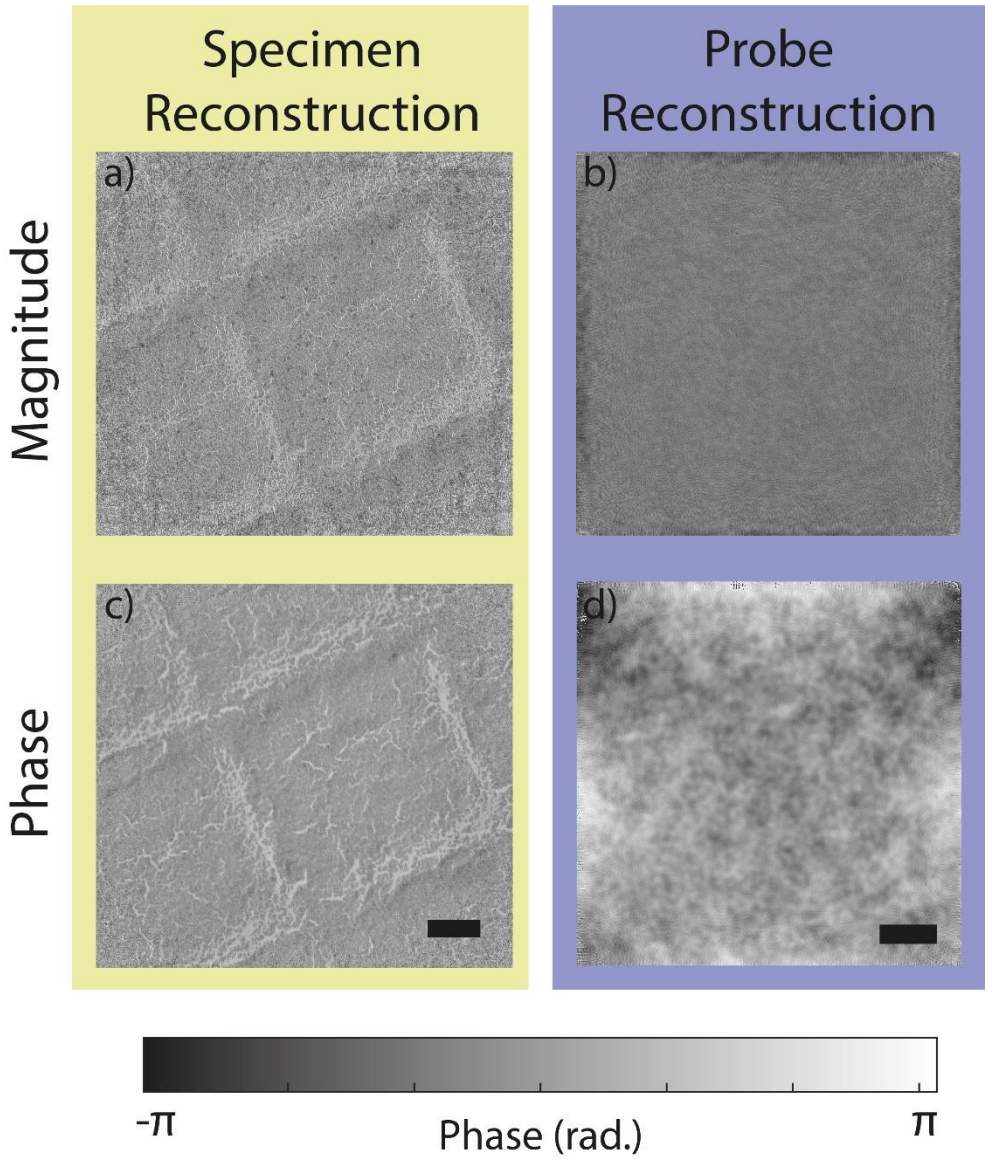

Figure 102: a) is the magnitude reconstruction of the diffraction grating replica. b) is the magnitude reconstruction of the diffuser shown in Figure 82 c). c) is the phase reconstruction of the diffraction grating replica. d) the phase reconstruction of the diffuser. The scale bar in c) indicates 100nm. The scale bar in d) indicates 50nm. The colour bar at the bottom of the figure converts the greyscale of c) and d) into radians.

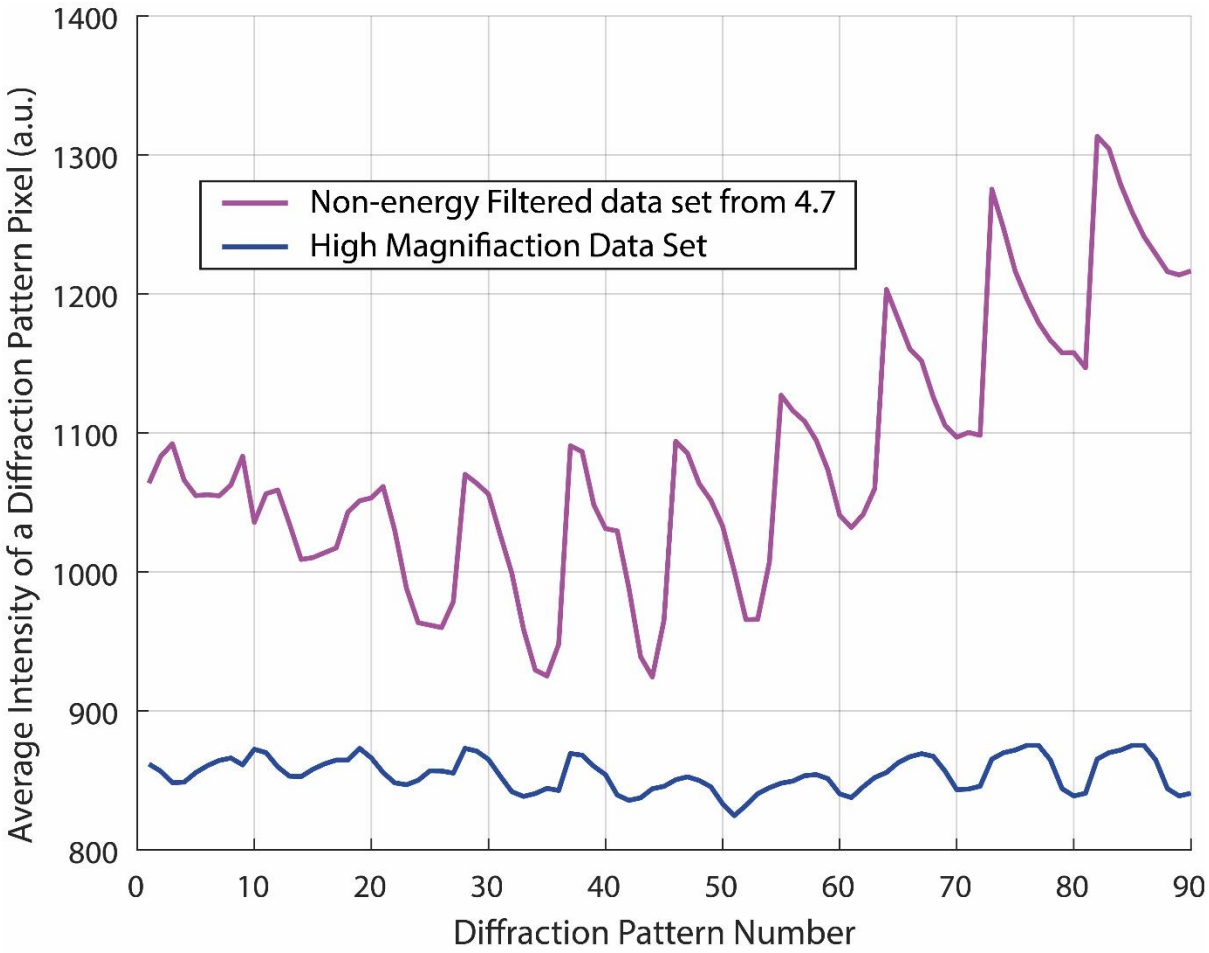

Figure 103: a figure comparing the average Intensity of diffraction pattern pixel as a function of the diffraction pattern number of the high magnification dataset and non-energy filtered dataset respectively. See Section 4.7 for more details on the non-energy filter dataset.

A better understanding of the full-field scan idea detailed in this section could be obtained by changing the following about the data collection: having a latex sphere in the field of view, changing the illumination such that there are more counts on the detector, and imaging a larger field of view. Another experiment which needs to be performed before this one is whether the coherence of the electron source (spot size) is impeding the resolution of the reconstructions.

4.11 Diffuser reconstructions and assessing diffuser choice

Several different diffuser designs were trialled during the calibration phase of our experiments (see Section 4.1 for details of these designs). As can be seen from inspection and comparison of the designed and reconstructed diffusers in Figure 105, they appear to match well, confirming that these designs can be etched reliably using an FIB system and have intended properties (speckle).

The strong phase effects seen at the edge of the probe reconstruction are a distinctive difference between the designs and the reconstructions shown in Figure 105. At the time of writing there is no concrete explanation for the cause of this phase change. Possible explanations include the charging of the gold that surrounds the diffuser (silicon nitride window), or aberrations in the projector lens system (and Gatan Image Filter [GIF]) being reconstructed in conjunction with the probe reconstruction, similar to that experienced by off axis holography in reference [12].

The effect of this strong phase can be seen more clearly if both the designed probe and the reconstructed probe are propagated a long distance. Both the designed and reconstructed probe are

propagated 5m in Figure 104 and this shows that the collected intensity for the reconstructed case is distorted compared to the design. This distortion is proportional to the propagation distance, but the simulation demonstrates that the distortion happens to the collected data to a lesser extent. This suggests that the collected data is not perfectly circular, while the diffuser design is. Therefore, the phase of the probe reconstruction has adopted the strong phase at the edge of the aperture to mitigate/emulate this effect.

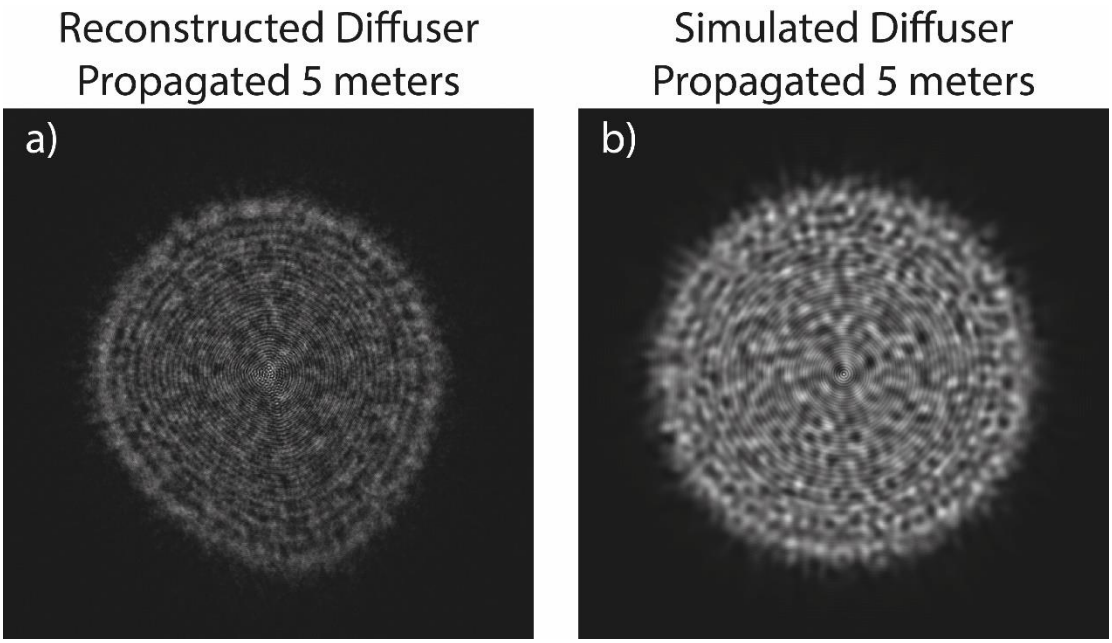

Figure 104: a) the simulated intensity from propagating the reconstructed phase of the diffuser seen in Figure 105 b) five metres forward. b) the simulated intensity from propagating the designed phase seen in Figure 105 c) forward 5m.

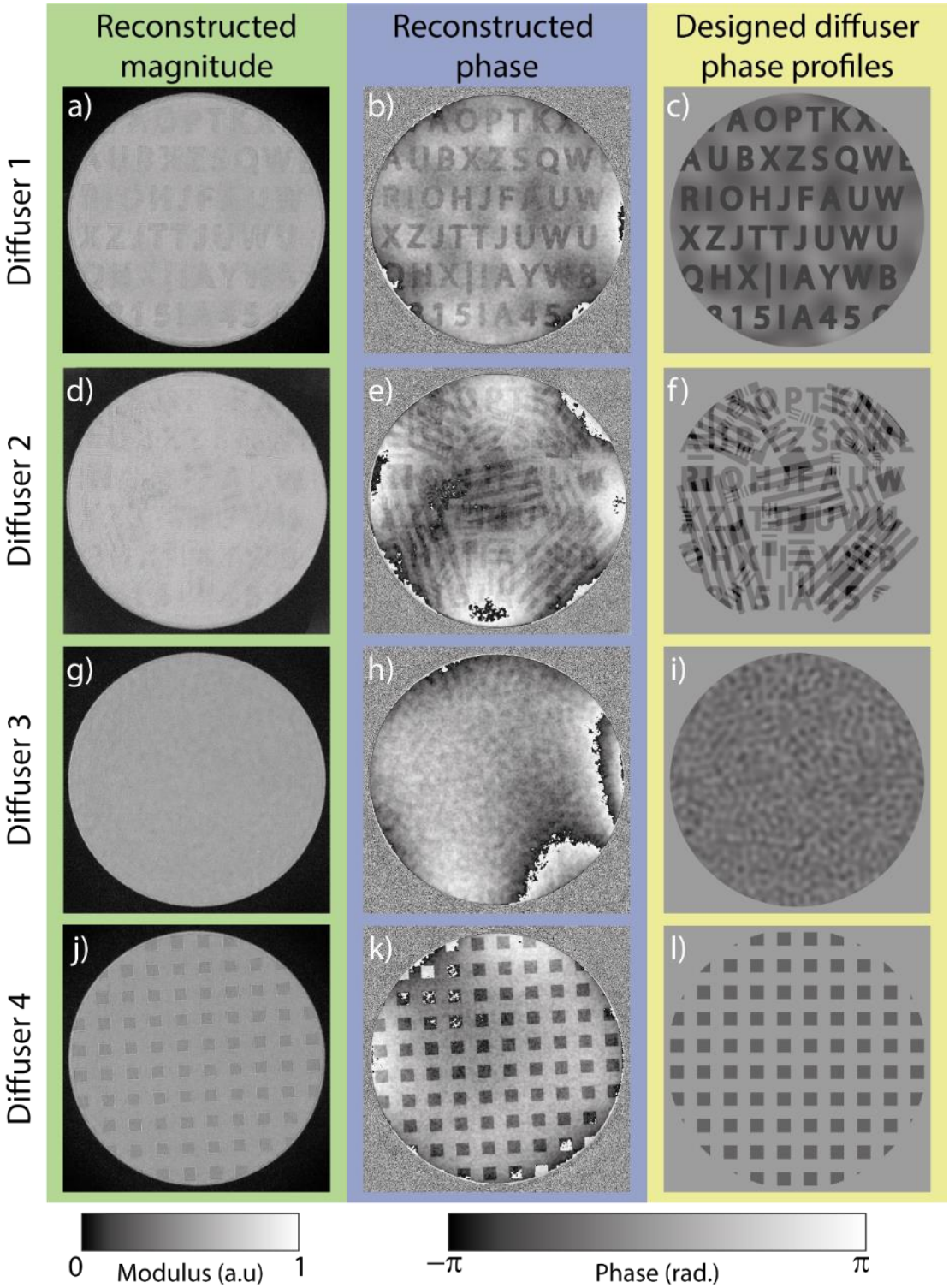

Figure 105: Comparison between reconstructed and designed diffusers after they have been propagated a long distance and simulated intensity has been collected [179].

The last focus of this diffuser section is the confirmation that the diffusers have their designated phase. This is difficult to confirm due to the strong phase at the edge of the reconstructions mentioned earlier. Focusing on the diffuser reconstruction shown in Figure 106, which has a well defined pattern, and then manually modelling for the localised phase effect, the average phase can then be measured. Figure 106 shows the subsection used in Figure 107, which is indicated by the red dotted line. The subsection was then processed by averaging along the *y*-axis obtaining the "Original

averaged cross-section" seen in Figure 107. As the "Original Averaged Cross-section" is affected by a low frequency phase variation, this phase variation was modelled by a manually optimised sinusoid (see the "modelled low frequency component" in Figure 107) in order to flatten the average phase of the specimen subsection such that the average phase difference between peaks and the valleys could be easily assessed (see the "Adjusted averaged cross-section" in Figure 107). The measured phase difference between the peaks and valleys of the "Adjusted averaged cross-section" in Figure 107 was found to be 2.18 radians, with a standard deviation of 0.07 radians [179].

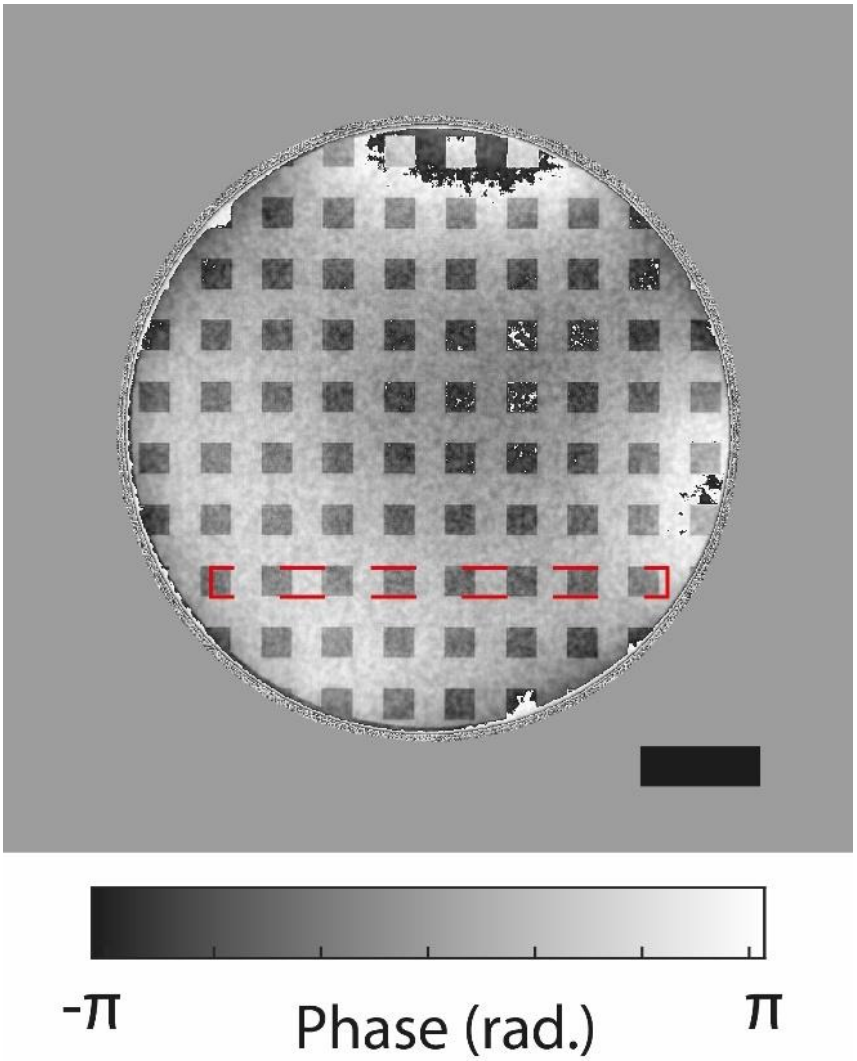

Figure 106: Reconstructed phase of the diffuser shown in Figure 82 d). The scale bar is 10µm. The area indicated by the red dotted line is the cross-section used to obtain the plots shown in Figure 107. The colour bar at the bottom of the figure converts the greyscale of the image into radians.

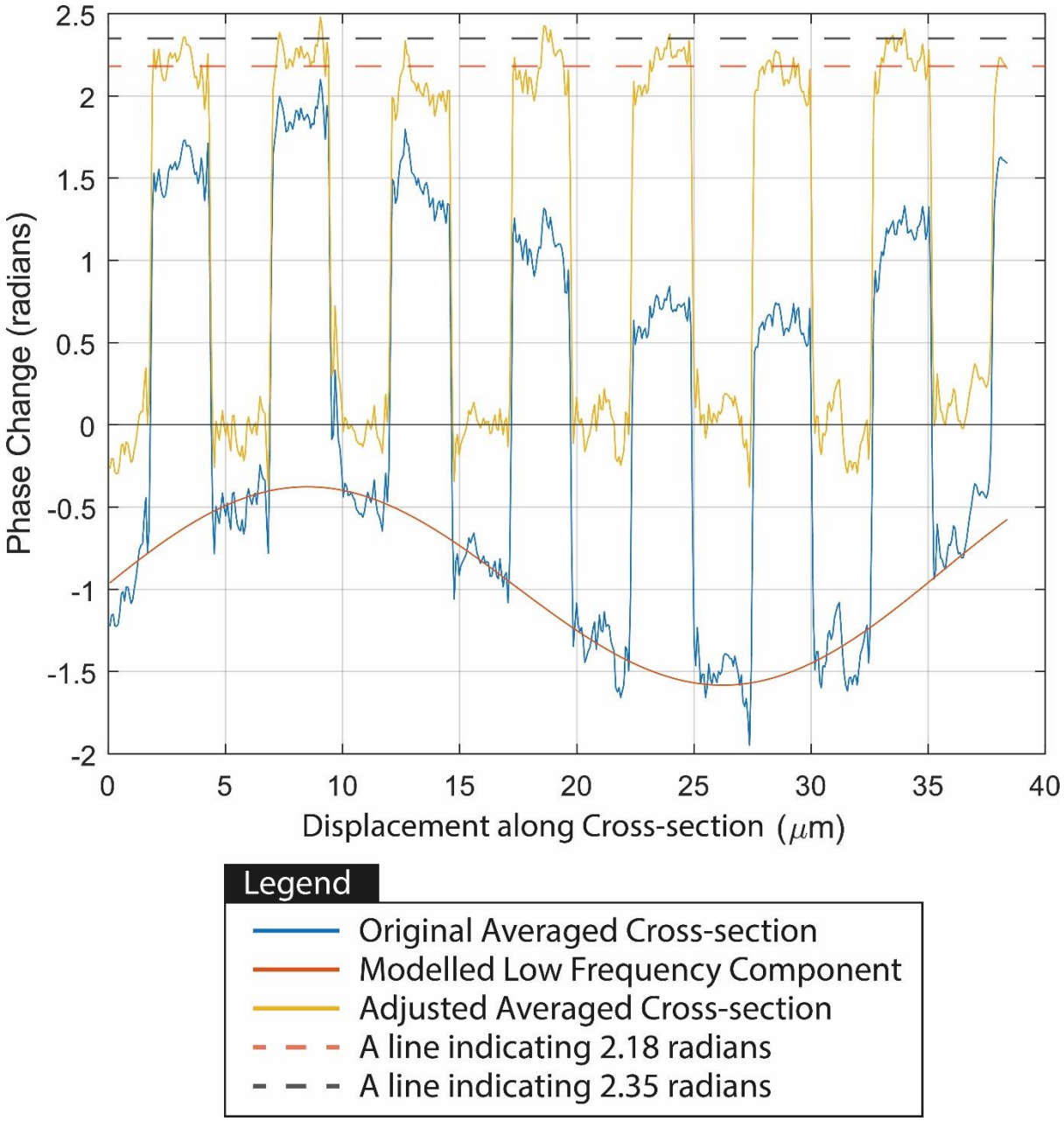

Figure 107: a figure showing the data used to calculate the phase delay caused by the diffuser after the removal of a low frequency component. The Original Averaged Cross-section phase was obtained by averaging along the y-axis of the subsection highlighted in Figure 106, (the blue plot of this figure). The modelled low frequency used to flatten the Original Averaged Cross-section Phase is also displayed as the orange plot. The yellow plot – the Adjusted Average Cross-section was used to calculate the average of the peaks and valleys of the square wave. The black dotted line indicates the intended phase of 2.35 radians. The red dotted line indicates the measured mean of the peaks of the square wave.

4.12 Low spatial frequency recovery

Figure 94 in Section 4.7 demonstrated strong correlation between independent phase images across a wide range of spatial frequencies. The fact that the lowest spatial frequencies correlate well could be explained by the initial conditions of the two reconstructions, which were both initialised with a flat phase [179]. To remove this potential bias, the two energy-filtered datasets were reconstructed again, this time with the strong, random low spatial frequency initialisations shown in Figure 108 a) and b). The reconstructed phase images in Figure 108 c)-f) demonstrate robust recovery from these random initialisations, except close to the edges of the reconstruction area. Phase recovery in these peripheral regions is less well conditioned, because they contribute to relatively few of the recorded

diffraction data. Comparison of the reconstructions via FRC proceeded as in Section 4.7, except that in this case a relative linear ramp between the two phase images was removed. This is a known ambiguity in ptychography [106]. Although it still shows good correlation at low spatial frequencies, the FRC plot in Figure 108 shows a slight dip in correlation across the full range of spatial frequencies, relative to Figure 94, which we attribute to imperfect removal of this phase ramp (see Figure 37). The intersection with the 1/2bit threshold is unaffected. Optimising the experimental setup to better recover low spatial frequencies is an interesting area for further exploration, for example by changing the camera length or increasing the spatial frequency bandwidth of the diffuser design. Algorithms adapted from inline holography may also assist in low spatial frequency recovery [51, 197].

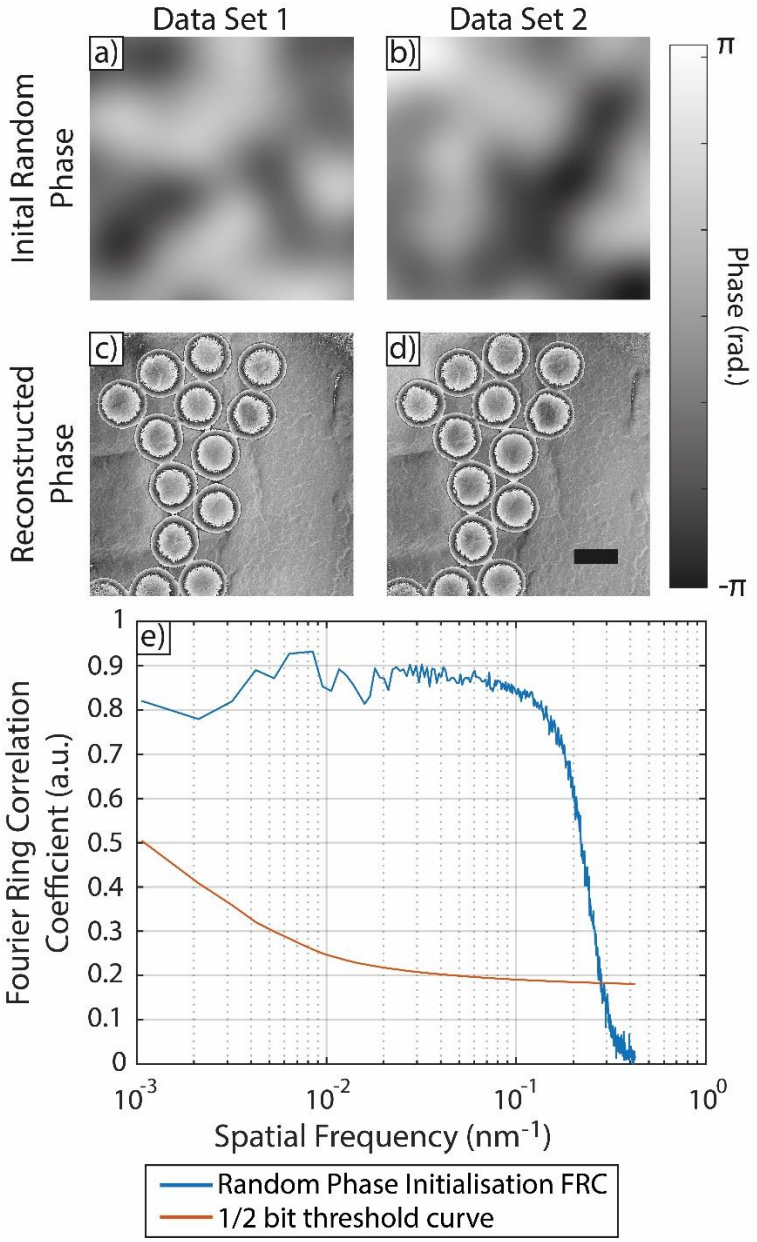

Figure 108: a figure illustrating how near field electron ptychography's low frequency response was tested by initialising two reconstructions with different and random phases. The two independent reconstructions were then compared via FRC analysis a) and b) the initial random phase of the reconstructions. c) and d) two independent reconstructed phases from two different datasets. e) the FRC coefficient as a function of frequency obtained by comparing c) and d). The ½ bit threshold is included in order to enable resolution measurement. The scale bar in d) indicates 200 nm and applies to a)-d). The colour bar to the right coverts the greyscale of a)-d) into radians [179].

Now a similar test will be performed on the large field reconstruction discussed in section 4.8 [179]. As the reconstructed field of view shown in Figure 96 is approximately 8 times the field of view of the detector, recovering the low spatial frequencies over such a large area from a random starting point undoubtedly poses a challenge. The random phase the reconstruction was initialised with is shown in Figure 109 a), and the final reconstructed phase is shown in Figure 109 b). There are many factors which could have affected the result shown in Figure 109 b), one example is the drift of the diffraction patterns during data collection; as a low frequency phase ramp may cause the collected intensity to shift in a similar fashion to the situation described in the DPC section (see section 1.9) which may be incorrectly accounted for by the drift correction. Alternatively, the drift could be interpreted by the rPIE algorithm to be due to a gradient in the phase of the specimen. The drift correction step also measures the drift of the diffraction patterns and a subsection of the measured drift is plotted in Figure 110. The erratic nature of the drift shown in Figure 110 is in part due to the manual correction of the diffraction pattern's drift during the experiment to prevent the diffraction pattern drifting off the detector. This is explained in greater detail in Section 3.6. A second reason for the erratic nature of Figure 110 is that the order of the diffraction patterns had to be changed due to a bug in the code, which saved the collected diffraction patterns to the hard drive but assigned the wrong diffraction pattern number.

Although the reconstruction shown in Figure 109 has not fully recovered from the random phase it began with, the final reconstruction has recovered quite well compared to its starting point. A possible reason may be the ambitious step size of 250nm (71.4 % overlap), which is the largest step size deployed in this thesis. As mentioned in section 4.8, if the whole of the detector was used this would enable a greater degree of overlap whilst maintaining the field of view and possibly reducing visibility of the low frequency noise, as suggested by section 5.2.3. An important question posed by this reconstruction is whether the size of the probe has any effect on the quality of recovered low frequencies. As can be seen by looking at Figure 109, the low frequency noise has a somewhat consistent size which is similar to that of the probe reconstruction shown in the top right of Figure 109 b). Although the exact reason for the low frequency noise is not fully understood, it is important that future work is directed at this problem, such that near field ptychography can operate at its full potential. One way to test whether the drift is affecting the low frequency response would be to perform multiple reconstructions initialised with random phase to see how the measured drift changes depending on the random phase. Another simple test would be designing diffusers with varying aperture sizes, which would then be used to collect large field of view datasets (under the constant diffraction conditions). These large field of view datasets could then be reconstructed as was demonstrated in this section, and then it would be possible to investigate whether there is a connection between the probe size and low frequency artefacts.

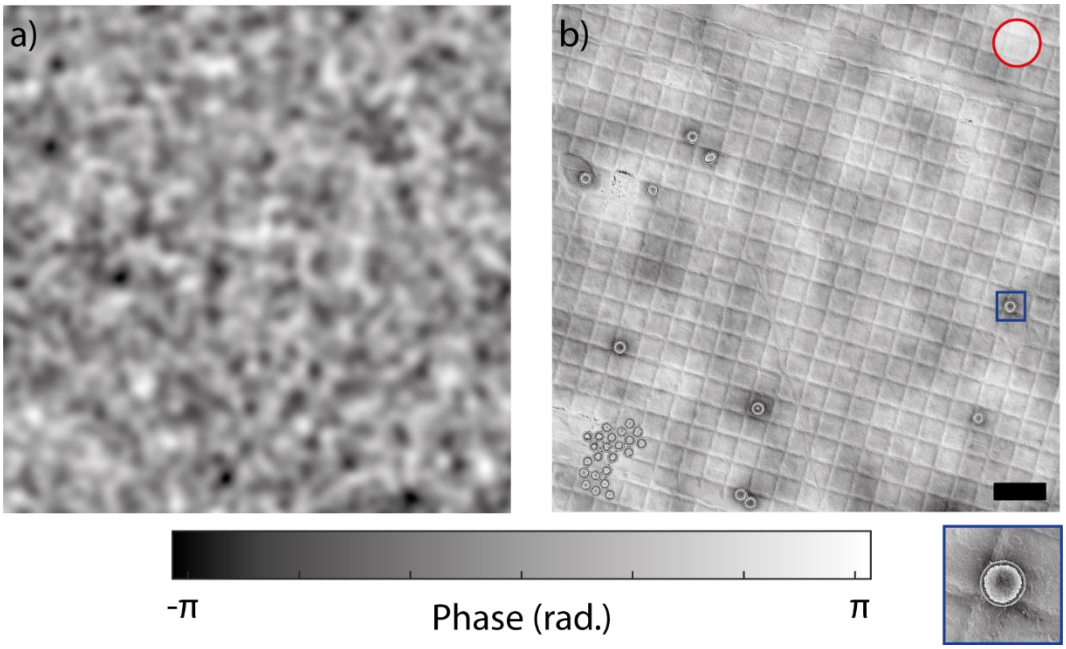

Figure 109: Large field of view phase reconstruction, with a random phase initialisation. a) shows the initial object phase fed to the reconstruction algorithm, b) shows the final phase image. The circle in the top right indicates the size of the diffuser aperture. The zoom shows the dim phase 'halo' still surrounding the boxed latex particle. Scale bar, 1 µm [179].

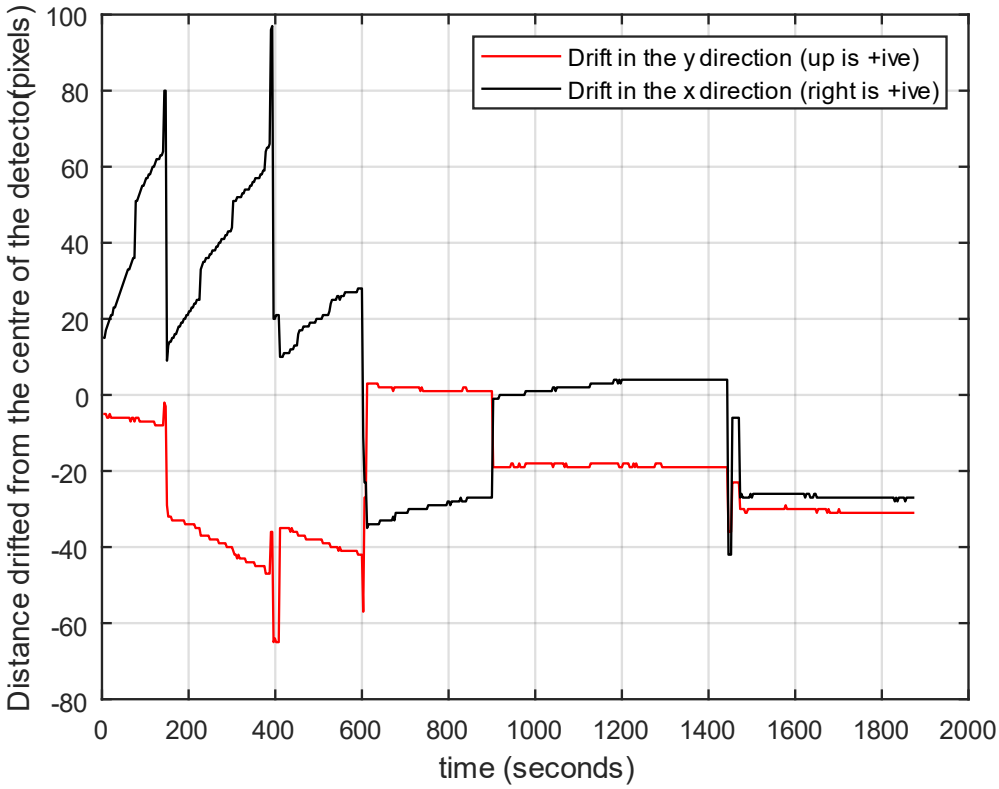

Figure 110: a figure showing the diffraction pattern drift for a sub-set of 625 diffraction patterns from the 1600 diffraction patterns in the large FOV dataset. The diffraction pattern drift is a problem as it cannot be separated from shifts caused by the specimen's phase; this is expanded upon in section 5.2.4.

4.13 Different diffraction conditions

Another set of experiments that took place was an investigation into the effect of the propagation distance (or defocus) between the image plane of the lens and the detector. Inline holography suggests that this distance is important in determining the spatial frequencies recovered in the phase reconstruction, as the defocus determines the spatial frequencies collected in the underlying data (the collected diffraction patterns). This can be seen if a collected diffraction pattern is Fourier transformed as shown in Figure 111.

The investigation consisted of four different camera lengths; 2.80 metres, 1.645 metres, 0.815 metres, and 0.38 metres. In this section:

- The Fourier transforms of the collected data will be assessed.
- Reconstructions at these different camera lengths will be visually compared.
- Any difference in the magnification will be assessed.
- The latex spheres in the reconstruction will be used to test whether the reconstructions are quantitative.
- Near field electron ptychography's frequency response and its reliance on spatial frequencies of the collected data will be discussed.

Just by inspecting Figure 111 the relationship linking the propagation distance between the specimen and the detector on the Fourier transform of the collected intensities can be seen. As the propagation distance decreases, the rings in the Fourier transforms of the diffraction patterns become larger and vary more slowly. Looking at the collected diffraction patterns, it can be seen that the specimen and the diffuser's diffraction fringes are interfering with one another, suggesting in principle that the phase can be obtained from all of these datasets. More details on how the different diffraction condition are obtained can be found in Section 3.4.

Using the imtool function in MATLAB to measure the diameters of the bright field disc of the collected diffraction patterns in terms of pixels and known diameter of the diffuser (50µm), the difference in magnification between the different diffraction conditions can be assessed. The diameters of the diffraction patterns are (710, 702 for 0.815m), (761,754 for 1.645m), (701,712 for 0.38m) and (770,794 for 2.8m). Using this information, the difference in magnification is generalised: the 1.645m and 2.8m propagation datasets experience an additional magnification factor in the order of 1.1 times. Due to the large field of view experiment, we have already obtained a reliable value for the pixel size when the propagation distance is 0.815 metres (i.e. the pixel size = 1.18nm). Therefore, using an estimated additional magnification factor of 1.1 for 1.645m and 2.8m diffraction conditions, the pixel sizes for these diffraction conditions is estimated to be 1.23nm. That fact that the difference in magnification between the different diffraction conditions is small is encouraging, as this means that it is very plausible to use the Fresnel scaling theorem to make the reconstructions the same size (see Section 3.5). Therefore, it would be valuable to perform direct FRC comparison if there is no sample deformation between experiments, and to measure whether near field electron ptychography with different diffraction conditions reconstruct spatial frequencies with differing degrees of signal to noise ratios, an important next step in near field electron ptychography.

It should be noted that the specimen itself could not be used to obtain the magnification, as the field of view is insufficient to capture a significant number of squares of the diffraction grating replica, such that the resolution measurement is reliable; it needs to be more than the two or three full squares currently imaged. The latex spheres could not be used as they may have shrunk between experiments [149]. It must be conceded that this makes any assessment of whether these

reconstructions are quantitative more questionable, and the author suggests that these experiments are repeated with a few changes which will be discussed at the end of the subsection.

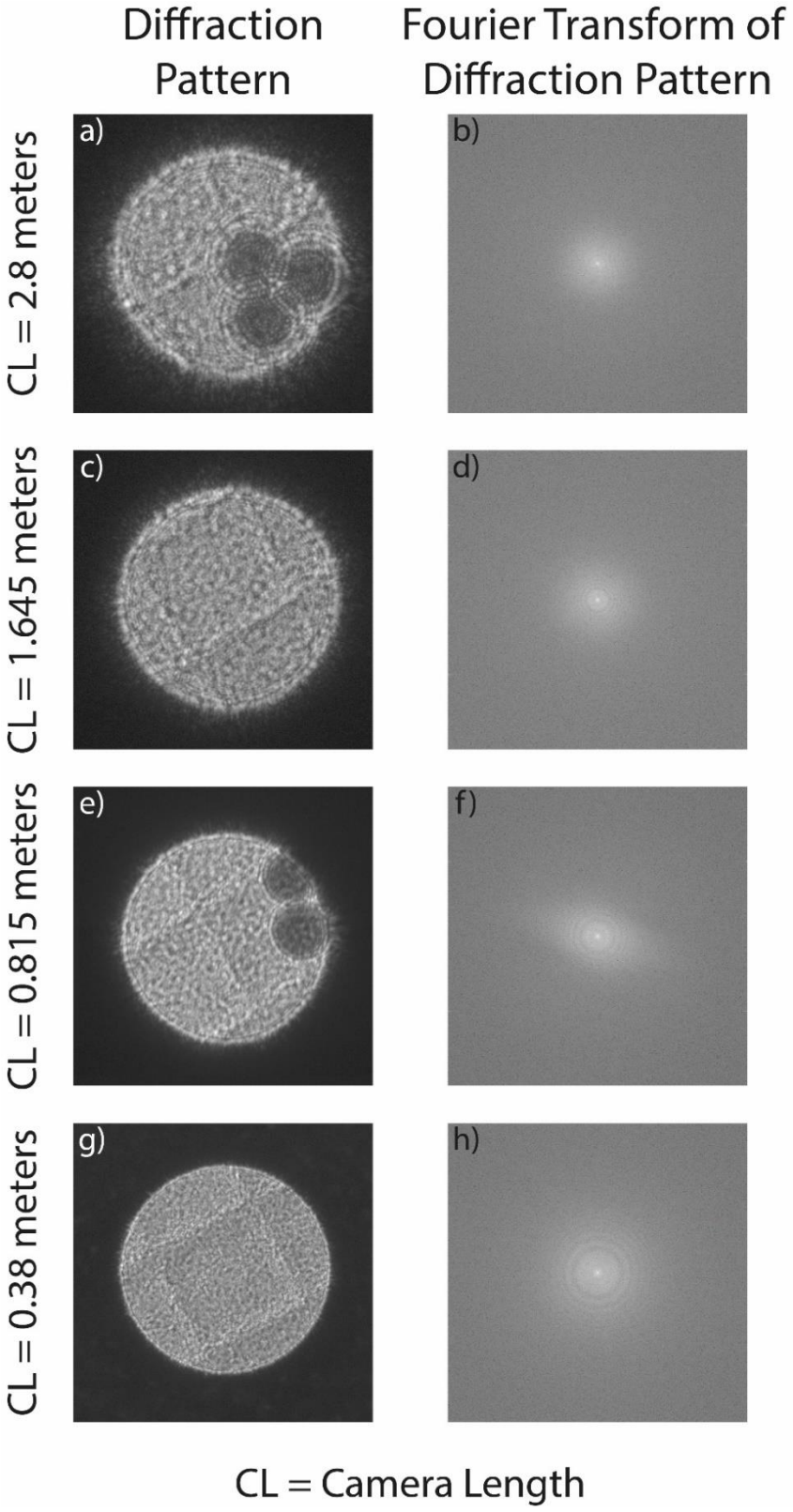

Figure 111: a), c), e) and g) an example diffraction patterns collected at 2.8m, 1.645m, 0.815m and 0.4m respectively. b), d), f) and h) are the Fourier transforms of a), c), e) and g) displaying modulated magnitude frequency response. This figure shows that even the near field electron ptychography diffraction patterns (in their current form) are affected by the contrast transfer function of the diffraction condition.

The experiment parameters are as follows: the step size all of the datasets was 100nm, the scan grid was a 10 by 10 and the exposure time of the detector was 0.5 seconds. The diffuser used during this experiment was diffuser 3, seen in Figure 82 c).

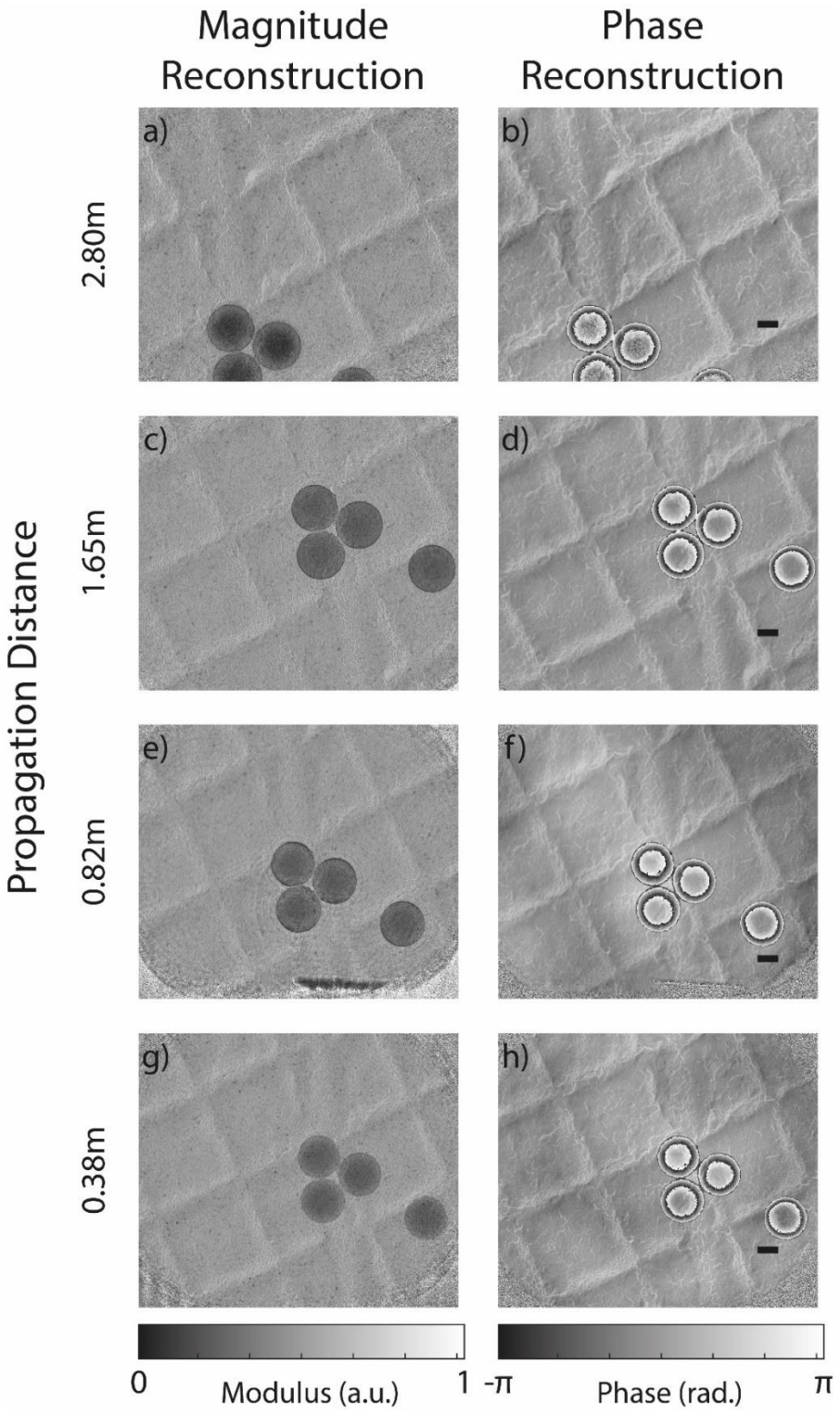

Figure 112: a figure showing how the recovered specimen function changes as the propagation distance between the detector and specimen is changed. a), c), e) and g) show how the magnitude of the specimen reconstruction changes as the camera length is altered. b), d), f) and h) show the how the phase of the specimen reconstruction changes as the camera length is altered. All scale bars indicate 100nm in respective images. The colour bar under a), c), e) and g) converts the greyscale of these images into modulus. The colour bar under b), d), f) and h) converts the greyscale of these image into radians.

Inspecting Figure 112, it can be seen that all of the datasets exhibit good phase recovery, and the reconstructions are visually similar to figures shown in previous sections. This confirms that near field electron ptychography can be performed at a variety of propagation distances. A greater degree of variance can be observed in the magnitude reconstructions as propagation distance changes, such as the visibility of dark flecks which changes depending on the reconstruction, opacity of the latex spheres, and the contrast of the cracks in the diffraction grating replica. However, none of these variances seem to suggest that any reconstructions have significantly different spatial frequency content than the others. A possible cause is reconstructions being slightly defocused from one another. The phase reconstructions for the most part are identical, apart from some low spatial frequency noise seen in the reconstructions with shorter camera lengths, and this matches well with expectations set out by inline holography.

Maiden et al. [94] has previously used magnitude images provided by selected area ptychography to determine the inelastic mean free path of latex. It is perfectly possible to do the same here with the magnitude images obtained via near field electron ptychography, although there are no available reference data for the inelastic mean free path of latex at 300 keV. Therefore, future work should include repeating this experiment at a lower acceleration voltage (200 keV), such that the magnitude of near field electron ptychography reconstructions can be proven to be quantitative like its SAP counterpart.

Figure 112 shows that the phase has been retrieved successfully, and now the latex spheres will be examined to confirm that the propagation distance does not affect whether the reconstructions are quantitative. The latex spheres of the 0.38 dataset have an average mean inner potential of 9.49±0.28V. This is unusually high in comparison to the literature and to other results in this thesis. The latex spheres of the 0.815 dataset have an average mean inner potential of 9.11±0.13V. Again, this is higher than previous results, but only slightly beyond a single standard deviation of previous results. The latex spheres of the 1.645 dataset have average mean inner potential of 8.54±0.21V. Due to the latex spheres in Figure 112 b) being so close to the edge of the reconstruction, there was an increased amount of noise affecting their phase which prevented the correct unwrapping of the phase via the Goldstein method (see section 3.8), due to the large number of branch cuts occurring in the latex spheres themselves. This means that there is no value for the mean inner potential for the 2.80m dataset. The individual mean inner potentials of each of the latex spheres under each diffraction condition can be seen in Table 10 through Table 12. An unwrapped version of the phase image shown in Figure 112 d) is shown in Figure 113 a), the numbers on the latex spheres in Figure 113 a) correspond to the latex sphere number in the tables below. Figure 113 b) shows the plot of the radial average unwrapped phase of the latex spheres shown in Figure 112 d) on normalised axes, the plots match well with the normalised modelled thickness. Figure 113 b) like previous figures suggests that the phase recovered by near field electron ptychography is quantitative, however similar to previous plots the fit is poorest at the start and end (reasons for this are suggested in section 4.5). The current hypothesis of the author is that the larger mean inner potential values for the 0.38 dataset are likely due to the uncertainty in the pixel size rather than incorrect reconstruction of the specimen. This links back to the discussion of the consistency of mean inner potential measurements at the end of subsection 4.5, in which it was discussed that off-axis holography experienced problems determining the mean inner potential of ZnO. The inconsistencies in the mean inner potential measurements are mainly due to uncertainty in the measured thickness provided by tomography and CBED patterns. Here the same problem could have occurred to near field electron ptychography, as once the measured thickness is uncertain (due to the unknown pixel size) the mean inner potential seems inconsistent with previous results. It is also possible that

frequency response may be a factor in the different mean inner potential values, but this would require further experiments to confirm.

The main result of experimenting with the propagation distance is that near field electron ptychography can obtain quantitative phase images at many different propagation distances. Before moving on to the conclusion, suggestions for further experiments are given that would further confirm and quantify this observation: one suggestion would be to increase the field of view such that the ability to recover low spatial frequencies can be appropriately assessed. A second suggestion would be to increase the amount of latex spheres in the field of view, making the claim that near field electron ptychography is quantitative regardless of the camera length more plausible. Finally, it would also be interesting to see if these diffraction conditions are affected by diffuser choice, as only one diffuser was used under these different diffraction conditions.

0.38m		
Latex sphere number	Mean inner potential (volts)	Latex sphere radius (nm)
1	9.4	104.83
2	9.17	104.76
3	9.54	105.26
4	9.83	103.83
Average MIP	9.49	

Table 10: a table showing the latex sphere number, the mean inner potential and the measured latex sphere radius in terms of nm from a near field electron ptychography dataset with a propagation distance of 0.38m. The latex sphere number corresponds to the latex spheres shown in Figure 113.

0.82m		
Latex sphere number	Mean inner potential (volts)	Latex sphere radius (nm)
1	9.19	107.91
2	9.23	107.69
3	8.94	109.58
4	9.06	108.60
Average MIP	9.11	

Table 11: a table showing the latex sphere number, the mean inner potential and the measured latex sphere radius in terms of nm from a near field electron ptychography dataset with a propagation distance of 0.82m. The latex sphere number corresponds to the latex spheres shown in Figure 113.

1.65m		
Latex sphere number	Mean inner potential (volts)	Latex sphere radius (nm)
1	8.37	116.80
2	8.79	119.76
3	8.63	121.76
4	8.37	118.34
Average MIP	8.54	

Table 12: a table showing the latex sphere number, the mean inner potential and the measured latex sphere radius in terms of nm from a near field electron ptychography dataset with a propagation distance of 1.65m. The latex sphere number corresponds to the latex spheres shown in Figure 113.

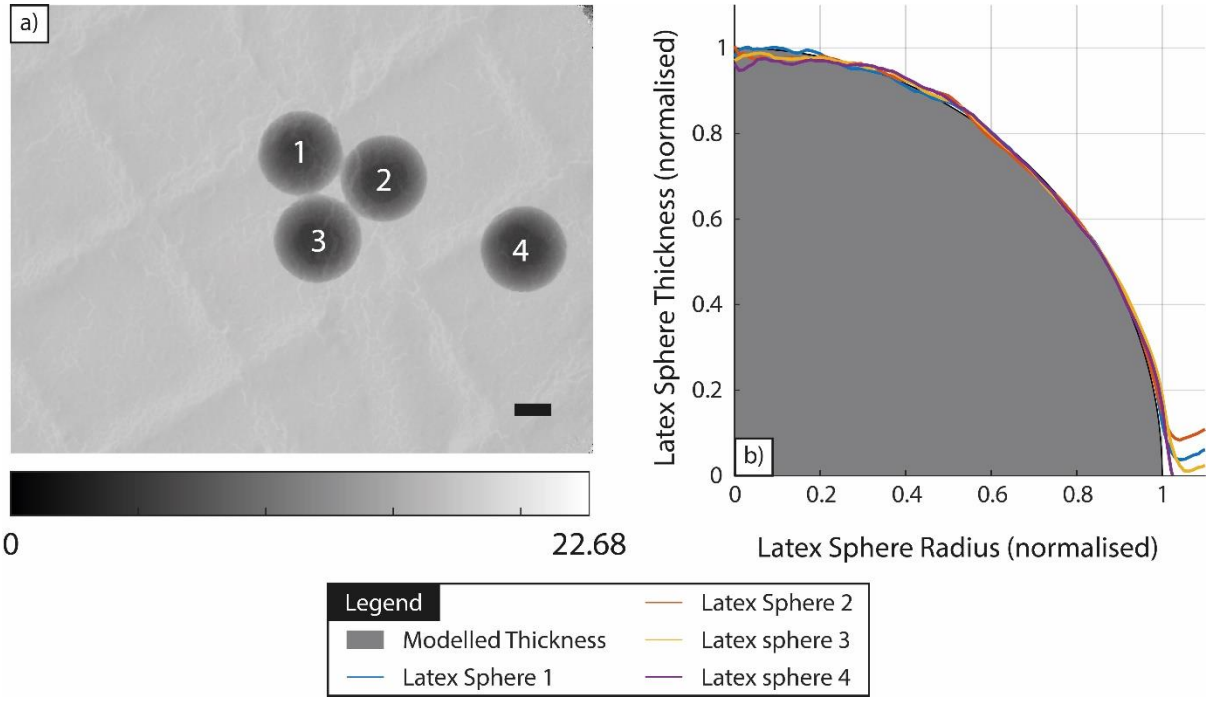

Figure 113: a) The unwrapped phase of Figure 112 d) with latex spheres numbered. The mean inner potential of the latex was measured for each of these spheres and the results averaged to give the statistics listed in the main text and the values tabulated in Table 11. The scale bar is 100 nm. The colour bar at the bottom of the figure converts the greyscale of the image into unwrapped phase in terms of radians. b) shows the radially averaged unwrapped phase of the numbered latex spheres shown in a) fitted to a normalised modelled thickness in order to demonstrate the quality of the fitted curves. The process of measuring the mean inner potential is described in Section 3.7.

4.14 Conclusion on the near field electron ptychography results

This section began with the description of the silicon nitride diffuser which is the core component of the implementation of near field electron ptychography in the TEM. The design choices which determined the phase delay of the diffuser were then discussed and the design choices ensured that the diffusers did not incur unreasonable amount of inelastic scatter. Then details of the fabrication process of silicon nitride diffusers were outlined, followed by details of the microscope and the conditions under which it operated.

The results shown in this section of the thesis demonstrate a new phase revival method in the transmission electron microscope, that of near field electron ptychography. In this section it has been shown that near field electron ptychography is a quantitative method with as few as 9 diffraction patterns (see section 4.5), as shown by the investigation into the mean inner potential of latex spheres where the data collection time was 27 seconds. With further development, near field electron ptychography could possibly be developed into a high throughput or real time technique, where the reconstruction code is running in parallel with the data collection [195], and researchers are provided with phase images of their sample almost immediately. Near field electron ptychography has been implemented with and without energy filtering and it has been shown that there is enough redundancy in the data to separate the background noise in the collected diffraction patterns from the elastic counts in line with previous work [94]. The modelled background consisted mostly of the inelastic scatter caused by the diffuser. This background modelling enabled non-energy filtered reconstruction to look visually similar to energy filtered datasets despite the inelastic scattered added by the diffuser. This is important, as it suggest that algorithmic steps can be used to mitigate one of the current disadvantages of near field electron ptychography compared to other forms of ptychography, and possibly to allow greater flexibility in the diffuser design (such as thicker

diffusers with greater phase changes). Then near field electron ptychography's greatest benefit was exhibited, in that it can provide mega-pixel fields of view via improved detector utilisation compared to its focused and defocused probe counterparts. The example given in this thesis was a 65 mega-pixel image covering an area of almost 100µm squared (8000 by 8000 pixels), which is particularly impressive considering that large fields of view provided by defocused probe ptychography have consisted of 1600 by 1600 pixel images [92]. This kind of direct comparison is unfair on defocused probe ptychography, which performs very well with low doses and is only used to highlight differences in the field of view of these two methods. Next in this section, the frequency response of near field electron ptychography was assessed via the use of the FRC and initialising the reconstruction with randomised low frequency phase. It was found that near field electron ptychography is consistent over a range of spatial frequencies despite the fact only one defocus condition was used, consistent with the optical results of Section 2. However, it was found that the low frequency information, especially in the case of extreme fields of view, may be a challenge for near field electron ptychography to recover.

On the subject of large fields of view it was found that near field electron ptychography can be more efficient in terms of field of view per diffraction pattern than the results presented in subsection 4.8, as the results of subsection 4.10 show. In subsection 4.10 the diffuser/illumination function covered the whole of the detector therefore achieving full detector utilisation, typical of near field ptychography in other domains, whilst in previous sections the illumination had not covered the detector in order to mitigate the effects of diffraction pattern drift. It is also worth noting that the results of subsection 4.10 defied expectations that the diffraction pattern drift or diffuser feature size would impede the successful reconstruction of the high magnification near field electron ptychography experiment. This result is particularly important, as it suggests that near field electron ptychography can operate at full detector utilisation, although this result agrees with the analysis from section 4.7 that the resolution is limited. Whilst the pixel size was sub-nanometre, the resolution seemed unchanged compared to previous lower magnification experiments, and this will be addressed in future experiments by exploring the coherence of the illumination, the effective drift of the diffraction pattern and aberrations of the objective lens. This future work will be vital if near field electron ptychography is going to be able to resolve individual atomic columns like its focused probe counterpart [198].

Subsection 4.9 covered the topic of imaging an MgO cube with near field electron ptychography. This subsection highlighted the difficulties that can be encountered if the specimen is not diverse (mostly vacuum). The large amount of vacuum in the imaged field of view meant that positions were difficult to confirm and the dataset had reduced diversity. This meant the reconstructed probe was poorly constrained, leading to phase artefacts in the specimen reconstructions. Despite this, near field electron ptychography was able to obtain a reasonable reconstruction of the specimen. For example, near field electron ptychography was able to pick up on the phase changes outside of the lacey carbon support and the MgO cube, suggesting that they have accumulated charge and are producing an electric field. Although near field electron ptychography experienced problems with this specimen, simple changes to the experiment can be made to improve performance, such as changing the magnification and imaging the specimen on a lacey carbon background instead of a vacuum.

In this section it was also shown that near field electron ptychography can perform well with a variety of diffusers, given that the dataset it helps create is diverse and includes interference between diffraction fringes of the specimen and the diffuser. The probe/diffuser reconstructions shown in figure 97 show that near field electron ptychography can successfully operate with a range

of diffusers. Furthermore, figure 96 and figure 97 suggest that near field electron ptychography is capable of recovering (or accounting for) the illumination function, as not only do the recovered probe reconstructions match closely to the designed diffusers, but the probe reconstruction also adapted additional phase features to account for possible distortions in the collected data. Lastly, near field electron ptychography at different propagation distances was investigated, and in subsection 4.13 near field electron ptychography diffraction patterns and their Fourier transforms were displayed. The Fourier transforms seemed similar to Fourier transforms of inline holograms, suggesting information loss at particular frequencies. Regardless of different diffraction conditions all datasets were successfully reconstructed, to give similar phase images compared to previous sections. Then the mean inner potential of the latex spheres was assessed and demonstrated that near field electron ptychography at different diffraction conditions gave similar or higher values of mean inner potential compared to previous sections, but this was attributed to the uncertainty in the pixel size. This suggests that near field electron ptychography is robust to the form of the diffraction condition and in future it may be possible to implement multifocal near field electron ptychography, as has been implemented in other domains [127].

Chapter 5

5 Conclusion and Future Work

5.1 Conclusion

Phase recovery in the electron domain is particularly important due to the relative weakness of the absorption characteristics of typical specimens compared to the phase delay they incur, hence the many methods available in electron microscopy by which the phase can be retrieved, each of which have their own specializations. Therefore, this thesis began with a review of the TEM via its components and then moved on to describe established techniques in the electron domain. Off-axis holography, inline holography and DPC were discussed. Ptychography in recent years has challenged these established techniques by simultaneously being able to recover the phase of the specimen and obtain unparalleled resolutions. Near field ptychography later iterated on focused/defocused probe ptychography by requiring fewer diffraction patterns for a given field of view, making it attractive for applications like tomography in the X-ray domain [79, 82]. Later in the literature review, the tools to propagate fields and assess the agreement between reconstructions via the FRC method were outlined.

In section 2 of the thesis, near field optical ptychography was discussed and investigated. This section was primarily interested in comparing the phase sensitivity of optical near field ptychography and optical focused probe ptychography. Phase sensitivity is an important parameter for electron microscopy, as highlighted in the following papers [120] and [121], and therefore optical experiments presented here suggest that near field optical ptychography and focused probe optical ptychography occupy the same order of phase sensitivity (the difference is less than ten times). The experiments performed in this section will have to be repeated in the electron domain provide confirmation this, but this result is encouraging in that the phase sensitivity is not drastically different. This work followed on from the paper entitled “Near-field ptychographic microscope for quantitative phase imaging” [125], which demonstrated quantitative phase imaging with near field optical ptychography; the work presented here went beyond this by performing FRC analysis with two near field reconstructions showing that near field optical ptychography is consistent over a range of spatial frequencies. The optical version of SAP (Selected Area Ptychography) with a larger aperture diameter was investigated to see whether greater fields of view could be investigated by SAP in the electron microscope (or at least greater detector utilisation) which is confirmed by comparing figure 23 and figure 54 b), which shows that the optical SAP diffraction pattern presented here has a greater detector utilisation than in its first implementation in the TEM. The fact there is improved detector utilisation suggests reasonable fields of view can be obtained with SAP with even less data than the initial results shown in [94]. Whilst less efficient than near field optical ptychography, SAP with a larger aperture may prove attractive to microscopists who do not want to alter their microscopes with custom aperture holders.

Near field electron ptychography data collection has been discussed, including the home-made DigitalMicrograph scripts, the workflow from raw data to a reconstructable dataset, the measuring of the camera length via reference data, and an understanding of diffraction pattern drift. It is important to note that without the work detailed in this section, later work such as the extreme field of view experiment in section 4.8 and the high magnification experiment in section 4.10, could not have been performed. This section documents the data collection process of near field electron ptychography and in combination with the data collection code in the appendix (section 7), in theory

allows anyone to collect near field electron ptychographic data or even selected area aperture ptychographic data. Near field electron ptychography's quantitative nature was tested by matching the reconstructed phase profiles of latex spheres to modelled thickness based on their measured diameter. The process by which the test was performed has been outlined in section 3.7 for reference.

In this thesis near field electron ptychography, a new method to recover phase of the specimen and the diffuser function in the TEM is presented. Near field electron ptychography has been shown capable of obtaining quantitative megapixel phase images and quantitative reconstructions with as few as 9 diffraction patterns, only taking 27 seconds of experiment time. Furthermore, when compared to other methods such as defocused probe electron ptychography [92], the fields of view provided by near field electron ptychography can be considered large, as demonstrated by the 65 megapixel image shown in section 4.8. Moreover, near field electron ptychography recovered the phase of the specimen without requiring a reference wave, allowing for investigation of different specimens, like charged tips [184, 199]. In this thesis near field electron ptychography was achieved by implementing the silicon nitride diffuser in the selected area aperture plane of the microscope, and by operating in diffraction mode, with the diffraction lens weakened such that a defocused image of the specimen and diffuser is present on the detector. The intermixing of the diffraction fringes and the movement of diffraction speckles as the specimen is scanned encode the phase information in the collected intensities, such that it can be recovered by iterative algorithms.

The importance of the diffuser was highlighted in subsection 4.6 by using the same diffraction conditions as the minimum diffraction pattern experiment, but replacing the diffuser with a 45-micrometre diameter aperture. This illustrated the advantages of the diffuser in that it maximises the interactions between diffraction fringes whilst requiring only moderate propagation distance between the specimen and detector. The diffuser designs set out in Figure 82 were then compared to their respective reconstructions in Figure 105, where the capability of near field electron ptychography to recover the illumination function alongside the specimen function was confirmed. Additionally, diffuser 4's reconstruction (Figure 82 d)) had its phase delay measured, obtaining a value of 2.18 radians for the phase imparted by the pattern, compared to the designed value of 2.35 radians.

It was also shown in Section 4.7 that near field electron ptychography is similar to its optical equivalent as it is consistent over a range of spatial frequencies with a single defocus, giving it a possible advantage over inline holography, which typically requires multiple diffraction conditions to achieve this same consistency [51]. However, when initialised with a low frequency random phase, the consistency of the low frequencies was reduced and this was highlighted again when the extreme field of view was tested in the same way. The extreme field of view phase image contained cloudy low frequency phase noise, suggesting it was challenging to recover the low frequencies; this will be investigated in future work.

It was also demonstrated that near field electron ptychography has sufficient redundancy that consistent background effects can be separated from the reconstructed specimen and diffuser. In the case of near field electron ptychography this was used to mitigate the effect of inelastic scatter from the silicon nitride diffuser on the reconstructed phase image. This was illustrated in Figure 93, where both energy filtered and non-energy filtered datasets were visually similar after applying background modelling, suggesting that near field electron ptychography does not require energy filtering. Demonstrating that near field electron ptychography could perform well without energy filtering was important to the extreme field of view experiment (see section 4.8), as the effective drift of the energy slit would limit the experiment's duration to short time frames (approximately 20

minutes) after which the zero-loss peak was filtered out and plasmons were imaged instead (see section 4.7 for more details). Knowing that energy filtering was not required at this acceleration voltage enabled the extreme field of view experiment to be performed, which with the current data collection method required an hour and half of microscope time.

Near field electron ptychography's current limits were then explored, via investigation of an MgO cube. It was then demonstrated that near field electron ptychography struggles with isolated specimens surrounded by a vacuum, as there is a lack of diversity in the specimen. The high magnification dataset (see section 4.10) demonstrated that the field of view per diffraction pattern could be increased even further than was experienced in section 4.8, but also demonstrated that the resolution of near field electron ptychography is limited. Possible explanations include microscope instabilities, lens aberrations, and incoherence, all of which need to be explored in future work.

The successful implementation of near field electron ptychography is an important stepping stone in the advancement of the state-of-the-art of phase imaging in TEM.

Near field electron ptychography can obtain quantitative phase images with as few as 9 diffraction patterns, as demonstrated in section 4.5, where the experiment only took 27 seconds. This with further improvements to data collection rates, could allow ptychographic reconstruction algorithms to recover the phase of the specimen of interest within a few minutes or less, allowing high throughput phase imaging.

In this thesis, the capability of near field electron ptychography to collect large fields of view more efficiently in terms of required diffraction data than its defocused/focused counterparts is demonstrated via the 93.4 micrometre squared field of view image shown in section 4.8, which only required 1600 diffraction patterns. This result helps demonstrate near field electron ptychography's potential niche of being able to collect large field of view images with nanometre resolutions in the shortest possible time frame, if its data collection rate can be improved to be closer to its focused probe counterpart.

It has been demonstrated in section 4.7 of this thesis that background modelling can mitigate the effects of the inelastic scatter of the implemented silicon nitride diffuser, such that a non-energy filtered reconstruction looks visually similar to its energy filtered counterpart. This suggests that energy filtering is not required in order to perform near field electron ptychography at this acceleration voltage, as long as the silicon nitride diffuser is appropriately designed.

As near field electron ptychography is new, it is rich in terms of the future work it can provide. Next steps for near field electron ptychography include: the improvement of the data collection rate by looking at modern automation methods and making the most of improvements of TEM/STEM in recent years (aberration correction and modern detectors). Another important step will be questioning the diffuser implementation in TEM and whether it can be replaced by apertures or aberrated lenses, enabling easier implementation in TEMs. Although near field electron ptychography has demonstrated a one hundred micrometre square field of view, it has struggled when recovering low spatial frequencies when initialized with a random phase. Therefore, an important next step will be assessing the frequency response of near field electron ptychography and how the data collection/reconstruction algorithm can be optimised to ensure the best possible recovered phase. Once optimised near field electron ptychography, will have to demonstrate that it can obtain quantitative results from specimens of interest. Examples include electron beam vortices and biological samples.

5.2 Future work

In this final section the proposed next steps for near field electron ptychography will be laid out. They are:

- Improving the implementation of near field electron ptychography
- Further exploration of diffuser design choices
- Reviewing the frequency response of near field ptychography
- Investigating further specimens with near field electron ptychography

5.2.1 Improving the implementation of near field electron ptychography

Near field electron ptychography in its current form can collect diffraction patterns at an approximate rate of 1 diffraction pattern every 3 seconds [179]. Half a second of the total time is the camera exposure time, and a quarter of a second is taken saving the diffraction pattern to the hard drive (note that this is an approximation). Therefore, around 2.25 seconds is taken up in image readout, stage movement and stage settling. Future work will then focus on methods to improve the data throughput of near field electron ptychography. It has been reported that using a combination beam and image scan coils effective specimen shifts up to several micrometres perpendicular to the optical axis can be achieved, which suggests that mega pixel images are obtainable with current TEMs in reduced time frames [200, 201]. However, moving the specimen large distances with the beam image shifts may introduce possible coma in the collected diffraction patterns without appropriate aberration correction [202, 203]. Tomography automation and automated single particle analysis literature may provide useful references when implementing the image scan coils in the context of near field electron ptychography.

Beam image shifts are calibrated by taking three images: a reference image with no image shift applied, and then two further images with different image shift values [200, 204]. These images are then cross-correlated in order to calibrate the microscope. Shift calibration and specimen scanning via beam image shifts could be implemented in the custom DigitalMicrograph code in the appendix (section 7) via insertion of the EMImageShift and EMBeamShift (found at: http://www.dmscripting.com/tutorial_microscope_commands.html). However, as many beam image shift systems have already been implemented, it likely just a case of inserting a diffuser into the selected area aperture plane of these microscopes. This makes performing single particle analysis via near field electron ptychography a distinct possibility in the near future.

5.2.2 Further exploration of diffuser design choices

In this thesis a limited number of the variables that affect diffuser performance have been explored. For example: the diffusers have been limited to a maximum phase shift of 2.35 radians, they have been limited to 50µm in diameter, they have been limited in feature size, they have been limited in terms of materials and have also been limited to a particular degree of inelastic scatter. Therefore, although the diffusers used throughout this thesis have been successful, the important work remains of finding the limits of diffusers in the electron microscope: we have not yet implemented a diffuser which does not work. It would be useful to have a rough idea of the extremes that cannot be tolerated by near field electron ptychography and the TEM. For example, if the diffuser can be made larger than the 50µm in diameter, then this would make investigating larger fields of view easier at lower magnifications. If diffusers can be made with smaller features, then that would allow the possibility of investigating whether near field electron ptychography can be used for high magnifications. Investigating different materials may allow for a faster phase change for less inelastic scatter. Or maybe near field electron ptychography can deal with the high degree of inelastic scatter. All of these questions are currently unanswered. A further question is whether magnitude-only

diffusers can be used. For example, an aperture could be altered to have thin wires crossing its diameter, to ensure there are diffraction fringes over the field of view of the detector. Furthermore, SAP may be worth reviewing now that a better understanding of how to correct the drift has been obtained (see Section 3.6), as larger apertures may be used, increasing the field of view per diffraction pattern (but still less than that which can be obtained by near field electron ptychography). However, SAP with large apertures will require a long propagation distance in order to generate the required fringes, which in turn requires a highly coherent beam. SAP, unlike near field electron ptychography, can be applied to most TEMs, as it does not require a custom aperture holder in the selected area aperture plane.

All of the diffusers used in this thesis were realised using FIB ion etching, a form of subtractive milling. Alternative fabrication processes such as lithography require thinner base thicknesses (around 15nm), meaning that diffusers fabricated in this fashion could offer a greater phase delay whilst incurring the same degree of inelastic scatter. Future work will also include operating near field electron ptychography at lower acceleration voltages, to investigate how increasing inelastic scatter from the diffuser may affect reconstructions. This is because the degree of inelastic scatter increases as the wavelength of the illumination decreases [205]. This is particularly important for biological samples which are beam sensitive and cannot take high acceleration voltages.

A possible limitation of Si_3N_4 diffusers is that they may have a limited lifetime, similar to that of early Zernike phase plates [206]. In the experiments performed in Section 4, the same diffuser was used over 4 days for 8-hour sessions without any obvious degradation in its performance. Therefore, an important study would be to investigate over how many experiments a diffuser remains operational. It may be the case that as the diffuser is damaged its performance may improve or remain unchanged, as the only requirement of the diffuser is to causes diffraction fringes, which then interfere with the diffraction fringes of the specimen.

It was shown in section 4.13 that the propagation distance between the detector and the specimen can easily be varied without changing the magnification. However, near field electron ptychography has a trade-off regarding the operational magnification and maximising the field of view (the goal of near field electron ptychography). Since the diffuser has a set radius and feature size, there is a limited range in which it can operate. This could be extended by altering the magnification power of the objective lens as the change in magnification would only apply to the sample, assuming the diffuser is placed in the selected-area aperture plane. But due to the aberrations of the objective lens, it would become quite tiresome to find the optimal conditions for the aberration correctors each time the magnification needs to be changed. A possible solution would be placing the diffusers in the condenser aperture (C2 aperture). This would have several interesting implications: the diameter of the diffuser may affect the coherence of the illumination (the semi angle of the illumination), and it would also be possible that the diffuser and the specimen would have differing focal planes, meaning that reconstruction algorithms may need an additional propagation step to refocus the diffuser. However, successful optical reconstructions have been performed where the diffuser was out of focus see (section 2). However, the advantage of this approach is that the size of the features of the diffuser would have greater independence from the total magnification required by the specimen. Implementing the diffuser in the condenser aperture would however have the disadvantage that thick specimens may be difficult to image, as the diffuser may no longer be thin with respect to the specimen [125, 179].

Furthermore, investigating whether Si_3N_4 diffusers could be replaced by an electric/magnetic field will be important to near field electron ptychography. If physical diffusers can be replaced by fields, this will decrease the degree of inelastic scatter experienced by near field electron ptychography and

will increase the resolution where energy filters are not available. Aberrations of the condenser lens are possible candidates for field only diffusers, as are laser phase plates [206, 207].

5.2.3 Frequency response of near field ptychography

One of the most interesting questions raised by near field electron ptychography is the recovery of low spatial frequencies in the context of mega-pixel images (extremely large field of view). This is because from a sampling perspective, obtaining mega-pixel images from a detector which is far smaller than the desired image suggests that when near field ptychography extends the field of view it is effectively up-sampling in Fourier space. This happens because the detector combined with the experimental condition causes the collected diffraction pattern to have well defined frequency space sampling. However, by extending the field of view we are attempting to recover information not available in any collected data. For example, in section 4.8 (large field of view experiment), the frequency space sampling size was roughly 8 times finer in the final reconstructed image compared to the data used in the reconstruction process (See Figure 114). In addition, inline holography, a similar technique (in terms of diffraction conditions) to near field electron ptychography, has had historical problems with recovering low spatial frequencies [51, 58]. It will therefore be interesting to see the degree of Fourier up-sampling that near field electron ptychography is capable of.

The FRC analysis performed on the energy and non-energy filtered datasets suggested that energy filtering improved the resolution of the near field electron ptychography. It was also shown in subsection 4.7 that resolution in the energy filtered reconstruction was greater (inferior) than the demagnified pixel size, meaning that the incoherence of inelastic scatter is not the only source of resolution loss for near field electron ptychography. This suggests an important next step will be determining what is limiting near field electron ptychography's high frequency response. The TEM which was used in the experiment described in Section 4 was uncorrected, so near field electron ptychography operating with aberration correction will certainly yield improved resolution images [179]. Furthermore, modern STEM/TEM detectors, which have reported greater data collection rates and higher signal to noise ratios, were not used in any of the experiments described [208]. The most likely candidate as a cause of the reduction of the resolution of near field electron ptychography is the incoherence of the source, but contributors may be the diffuser design, the camera length (phase transfer function), and instabilities of the setup/diffraction condition.

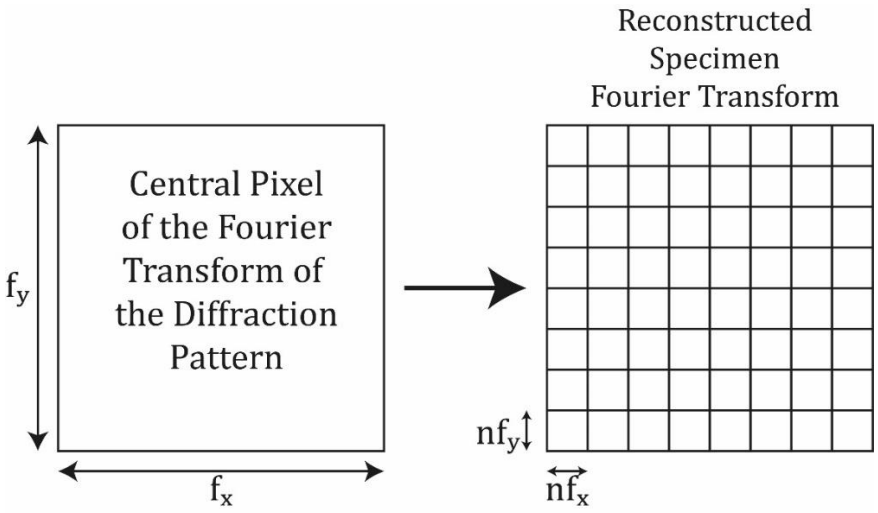

Figure 114: A figure illustrating the up-sampling in Fourier space caused by extending the field in near field ptychography. Where fx and fy indicates the pixel size in Fourier space in collected diffraction pattern. nfx and nfy indicate the pixel size in Fourier space of the extended field of view reconstruction.

It is also important to state that there have been no reports from X-ray near field ptychography in which a similar issue has been encountered. Therefore, it is a possibility that the problems described in this thesis are the result of improper implementation of various experimental factors. An example of such an experimental factor may be the diffuser choice, as depending on the diffuser implemented, the Fourier transform of the diffraction pattern may or may not display a contrast transfer function like variation.

Even if near field ptychography is limited in terms of the spatial frequencies available in the collected diffraction data, there are practical solutions, namely increasing the number of pixels of the detector. Near field electron ptychography has shown decent performance when the field of view is not sufficiently greater than the field of view of the detector (see 4.12). There are 4K detectors available today. Increasing the number of pixels of the detector, and therefore increasing the detector area/field of view while a native approach should still be stated. This approach would incur problems in the illumination system of the TEM, as the source would have to be sufficiently bright and coherent over a large enough area.

An important next step would be comparing an inline holography phase image with that of near field ptychography reconstruction via the FRC method. The resulting FRC would measure the contrast transfer function of near field ptychography instead of its consistency, as the phase obtained by inline holography will be independent of that obtained by near field ptychography. An important reference for this next step is “Near-field ptychography using lateral and longitudinal shifts” [127] as this paper outlines the practical setup which can collect both near field ptychography data and inline holograms.

5.2.4 Using near field electron ptychography to investigate further specimens

It has also been shown that near field electron ptychography has the potential to image strong phase specimens, as shown by the imaging of the MgO cube in section 4.9. Future work should include replicating the work performed in the X-ray and optical modalities of imaging thick and absorptive specimens, such as the equivalent of the uranium sphere used in [56]. A future objective will be repeating the MgO cube (Section 4.9) and full detector utilisation experiments (Section 4.10). Future work will also include imaging an extended electric field to ascertain whether the lack of positional data will cause near field ptychography the difficulties we expect. Investigating further specimens would be beneficial to the claim that near field electron ptychography is a quantitative technique beyond the scope of imaging latex spheres. Further experiments could also investigate whether the magnitude recovered by near field electron ptychography is quantitative, by investigating latex spheres at lower acceleration voltage (200 keV) and comparing the results to measured and estimated values of mean inner potential for latex found by other experiments.

One of the proposed benefits of near field ptychography over off-axis holography is that it does not require a reference wave. This advantage could be beneficial when imaging electrical and magnetic fields, as off axis holography has difficulty forming reference waves when imaging long range electrical and magnetic fields [199]. This problem is typically circumvented by using a computational model of the specimen which is then compared to the phase obtained from the off-axis hologram.

Off-axis holography has also been used to investigate nanotubes, which are an area of interest as they have shown potential as high performance electron sources due to their high surface area to volume ratio [209] and long life time [210]. An important step is the characterisation of the charge along the nanotube which can be performed by off-axis holography. However, due to the limited field of view of off-axis holography, this could only be performed for the tip and not the whole

nanowire [209]. Near field electron ptychography has the potential to provide quantitative phase information for the whole nanowire due to its unlimited field of view.

6 Bibliography

[1] L. F. Drummy, R. J. Davis, D. L. Moore, M. Durstock, R. A. Vaia, and J. W. Hsu, "Molecular-scale and nanoscale morphology of P3HT: PCBM bulk heterojunctions: energy-filtered TEM and low-dose HREM," *Chemistry of Materials,* vol. 23, no. 3, pp. 907-912, 2011.

[2] D. Cooper, T. Denneulin, N. Bernier, A. Béché, and J.-L. Rouviere, "Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope," *Micron,* vol. 80, pp. 145-165, 2016.

[3] R. E. Dunin-Borkowski *et al.*, "Off-axis electron holography of magnetic nanowires and chains, rings, and planar arrays of magnetic nanoparticles," *Microscopy research and technique,* vol. 64, no. 5-6, pp. 390-402, 2004.

[4] P. A. Midgley, E. P. Ward, A. B. Hungría, and J. M. Thomas, "Nanotomography in the chemical, biological and materials sciences," *Chemical Society Reviews,* vol. 36, no. 9, pp. 1477-1494, 2007.

[5] E. Hecht, "Optics, 4th," *International edition, Addison-Wesley, San Francisco,* vol. 3, p. 2, 2002.

[6] D. B. Williams and C. B. Carter, "The transmission electron microscope," in *Transmission electron microscopy*: Springer, 1996, pp. 3-17.

[7] O. Scherzer, "The theoretical resolution limit of the electron microscope," *Journal of Applied Physics,* vol. 20, no. 1, pp. 20-29, 1949.

[8] C. McAlinden, M. McCartney, and J. Moore, "Mathematics of Zernike polynomials: a review," *Clinical & experimental ophthalmology,* vol. 39, no. 8, pp. 820-827, 2011.

[9] V. Lakshminarayanan and A. Fleck, "Zernike polynomials: a guide," *Journal of Modern Optics,* vol. 58, no. 7, pp. 545-561, 2011.

[10] B. Freitag, S. Kujawa, P. Mul, J. Ringnalda, and P. Tiemeijer, "Breaking the spherical and chromatic aberration barrier in transmission electron microscopy," *Ultramicroscopy,* vol. 102, no. 3, pp. 209-214, 2005.

[11] M. Linck *et al.*, "Chromatic aberration correction for atomic resolution TEM imaging from 20 to 80 kV," *Physical review letters,* vol. 117, no. 7, p. 076101, 2016.

[12] P. A. Midgley, "An introduction to off-axis electron holography," *Micron,* vol. 32, no. 2, pp. 167-184, 2001.

[13] J. Chapman, "The investigation of magnetic domain structures in thin foils by electron microscopy," *Journal of Physics D: Applied Physics,* vol. 17, no. 4, p. 623, 1984.

[14] S. Cao, "Large Field of View Electron Ptychography," 2017.

[15] J. Chapman and M. Scheinfein, "Transmission electron microscopies of magnetic microstructures," *Journal of magnetism and magnetic materials,* vol. 200, no. 1-3, pp. 729-740, 1999.

[16] E. James and N. Browning, "Practical aspects of atomic resolution imaging and analysis in STEM," *Ultramicroscopy,* vol. 78, no. 1-4, pp. 125-139, 1999.

[17] M. Weyland and D. A. Muller, "Tuning the convergence angle for optimum STEM performance," *arXiv preprint arXiv:2008.12870,* 2020.

[18] S. Bals, B. Kabius, M. Haider, V. Radmilovic, and C. Kisielowski, "Annular dark field imaging in a TEM," *Solid state communications,* vol. 130, no. 10, pp. 675-680, 2004.

[19] A. Surrey, "Preparation and Characterization of Nanoscopic Solid State Hydrogen Storage Materials," 2016.

[20] S. J. Pennycook, "Seeing the atoms more clearly: STEM imaging from the Crewe era to today," *Ultramicroscopy,* vol. 123, pp. 28-37, 2012.

[21] S. J. Pennycook, "A scan through the history of STEM," in *Scanning transmission electron microscopy*: Springer, 2011, pp. 1-90.

[22] T. J. Slater, A. Janssen, P. H. Camargo, M. G. Burke, N. J. Zaluzec, and S. J. Haigh, "STEM-EDX tomography of bimetallic nanoparticles: A methodological investigation," *Ultramicroscopy,* vol. 162, pp. 61-73, 2016.
[23] P. E. Batson, N. Dellby, and O. L. Krivanek, "Sub-ångstrom resolution using aberration corrected electron optics," *Nature,* vol. 418, no. 6898, pp. 617-620, 2002.
[24] F. Haguenau, P. Hawkes, J. Hutchison, B. Satiat–Jeunemaître, G. Simon, and D. Williams, "Key events in the history of electron microscopy," *Microscopy and Microanalysis,* vol. 9, no. 2, pp. 96-138, 2003.
[25] A. Crewe, J. Wall, and L. Welter, "A high-resolution scanning transmission electron microscope," *Journal of Applied Physics,* vol. 39, no. 13, pp. 5861-5868, 1968.
[26] A. V. Crewe, J. Wall, and J. Langmore, "Visibility of single atoms," *science,* vol. 168, no. 3937, pp. 1338-1340, 1970.
[27] S. Pennycook and D. Jesson, "High-resolution Z-contrast imaging of crystals," *Ultramicroscopy,* vol. 37, no. 1-4, pp. 14-38, 1991.
[28] I. Lazić, E. G. Bosch, and S. Lazar, "Phase contrast STEM for thin samples: Integrated differential phase contrast," *Ultramicroscopy,* vol. 160, pp. 265-280, 2016.
[29] E. G. Bosch and I. Lazić, "Analysis of HR-STEM theory for thin specimen," *Ultramicroscopy,* vol. 156, pp. 59-72, 2015.
[30] P. Hartel, H. Rose, and C. Dinges, "Conditions and reasons for incoherent imaging in STEM," *Ultramicroscopy,* vol. 63, no. 2, pp. 93-114, 1996.
[31] Y.-M. Kim, S. J. Pennycook, and A. Y. Borisevich, "Quantitative comparison of bright field and annular bright field imaging modes for characterization of oxygen octahedral tilts," *Ultramicroscopy,* vol. 181, pp. 1-7, 2017.
[32] R. R. Meyer and A. Kirkland, "The effects of electron and photon scattering on signal and noise transfer properties of scintillators in CCD cameras used for electron detection," *Ultramicroscopy,* vol. 75, no. 1, pp. 23-33, 1998.
[33] M. W. Tate *et al.*, "High dynamic range pixel array detector for scanning transmission electron microscopy," *Microscopy and Microanalysis,* vol. 22, no. 1, pp. 237-249, 2016.
[34] I. MacLaren, T. A. Macgregor, C. S. Allen, and A. I. Kirkland, "Detectors—The ongoing revolution in scanning transmission electron microscopy and why this important to material characterization," *APL Materials,* vol. 8, no. 11, p. 110901, 2020.
[35] A. Faruqi and G. McMullan, "Direct imaging detectors for electron microscopy," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 878, pp. 180-190, 2018.
[36] M. Lohr *et al.*, "Differential phase contrast 2.0—Opening new “fields” for an established technique," *Ultramicroscopy,* vol. 117, pp. 7-14, 2012.
[37] M. Krajnak, D. McGrouther, D. Maneuski, V. O'Shea, and S. McVitie, "Pixelated detectors and improved efficiency for magnetic imaging in STEM differential phase contrast," *Ultramicroscopy,* vol. 165, pp. 42-50, 2016.
[38] S. Toyama *et al.*, "Quantitative electric field mapping of ap–n junction by DPC STEM," *Ultramicroscopy,* vol. 216, p. 113033, 2020.
[39] N. Shibata, S. D. Findlay, Y. Kohno, H. Sawada, Y. Kondo, and Y. Ikuhara, "Differential phase-contrast microscopy at atomic resolution," *Nature Physics,* vol. 8, no. 8, pp. 611-615, 2012.
[40] E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *JOSA,* vol. 52, no. 10, pp. 1123-1130, 1962.
[41] J. Cowley and J. Spence, "Principles and theory of electron holography," in *Introduction to Electron Holography*: Springer, 1999, pp. 17-56.
[42] J. Cowley, "Twenty forms of electron holography," *Ultramicroscopy,* vol. 41, no. 4, pp. 335-348, 1992.
[43] H. Lichte, D. Geiger, and M. Linck, "Off-axis electron holography in an aberration-corrected transmission electron microscope," *Philosophical Transactions of the Royal Society of London*

A: Mathematical, Physical and Engineering Sciences, vol. 367, no. 1903, pp. 3773-3793, 2009.
[44] M. Lehmann and H. Lichte, "Tutorial on off-axis electron holography," *Microscopy and Microanalysis,* vol. 8, no. 6, pp. 447-466, 2002.
[45] A. Tonomura, "Applications of electron holography using a field-emission electron microscope," *Microscopy,* vol. 33, no. 2, pp. 101-115, 1984.
[46] C. T. Koch and A. Lubk, "Off-axis and inline electron holography: A quantitative comparison," *Ultramicroscopy,* vol. 110, no. 5, pp. 460-471, 2010.
[47] M. E. Haine and T. Mulvey, "The formation of the diffraction image with electrons in the Gabor diffraction microscope," *JOSA,* vol. 42, no. 10, pp. 763-773, 1952.
[48] C. T. Koch, "A flux-preserving non-linear inline holography reconstruction algorithm for partially coherent electrons," *Ultramicroscopy,* vol. 108, no. 2, pp. 141-150, 2008.
[49] M. R. Teague, "Deterministic phase retrieval: a Green's function solution," *JOSA,* vol. 73, no. 11, pp. 1434-1441, 1983.
[50] Z. Jingshan, R. A. Claus, J. Dauwels, L. Tian, and L. Waller, "Transport of intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes," *Optics express,* vol. 22, no. 9, pp. 10661-10674, 2014.
[51] A. Eljarrat, J. Müller, M. R. S. Huang, and C. T. Koch, "Multi-focus TIE algorithm including partial spatial coherence and overlapping filters," *Optics Express,* vol. 26, no. 9, pp. 11819-11833, 2018/04/30 2018, doi: 10.1364/OE.26.011819.
[52] W. M. J. Coene, A. Thust, M. O. De Beeck, and D. Van Dyck, "Maximum-likelihood method for focus-variation image reconstruction in high resolution transmission electron microscopy," *Ultramicroscopy,* vol. 64, no. 1-4, pp. 109-135, 1996.
[53] A. Thust, W. M. J. Coene, M. O. De Beeck, and D. Van Dyck, "Focal-series reconstruction in HRTEM: Simulation studies on non-periodic objects," *Ultramicroscopy,* vol. 64, no. 1-4, pp. 211-230, 1996.
[54] C. T. Koch, "Towards full-resolution inline electron holography," *Micron,* vol. 63, pp. 69-75, 2014.
[55] C. Zuo *et al.*, "Transport of intensity equation: a tutorial," *Optics and Lasers in Engineering,* p. 106187, 2020.
[56] M. Stockmar *et al.*, "X-ray near-field ptychography for optically thick specimens," *Physical Review Applied,* vol. 3, no. 1, p. 014005, 2015.
[57] P. Cloetens *et al.*, "Holotomography: Quantitative phase tomography with micrometer resolution using hard synchrotron radiation x rays," *Applied physics letters,* vol. 75, no. 19, pp. 2912-2914, 1999.
[58] M. R. Huang, A. Eljarrat, and C. T. Koch, "Quantifying the data quality of focal series for inline electron holography," *Ultramicroscopy,* p. 113264, 2021.
[59] L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Optics express,* vol. 18, no. 12, pp. 12552-12561, 2010.
[60] D. Paganin, A. Barty, P. J. McMahon, and K. A. Nugent, "Quantitative phase-amplitude microscopy. III. The effects of noise," *Journal of microscopy,* vol. 214, no. 1, pp. 51-61, 2004.
[61] P. Tiemeijer, M. Bischoff, B. Freitag, and C. Kisielowski, "Using a monochromator to improve the resolution in TEM to below 0.5 Å. Part I: Creating highly coherent monochromated illumination," *Ultramicroscopy,* vol. 114, pp. 72-81, 2012.
[62] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy,* vol. 109, no. 10, pp. 1256-1262, 2009.
[63] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik,* vol. 35, pp. 237-246, 1972.
[64] S. Marchesini, "Invited article: A unified evaluation of iterative projection algorithms for phase retrieval," *Review of scientific instruments,* vol. 78, no. 1, p. 011301, 2007.

[65] J. R. Fienup, "Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint," *JOSA A,* vol. 4, no. 1, pp. 118-123, 1987.
[66] J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Optics letters,* vol. 3, no. 1, pp. 27-29, 1978.
[67] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied optics,* vol. 21, no. 15, pp. 2758-2769, 1982.
[68] K. A. Nugent, "Coherent methods in the X-ray sciences," *Advances in Physics,* vol. 59, no. 1, pp. 1-99, 2010-01-01 2010, doi: 10.1080/00018730903270926.
[69] H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm," *Physical review letters,* vol. 93, no. 2, p. 023903, 2004.
[70] J. Rodenburg and A. Maiden, "Ptychography," in *Springer Handbook of Microscopy*: Springer, 2019, pp. 2-2.
[71] M. Guizar-Sicairos and J. R. Fienup, "Understanding the twin-image problem in phase retrieval," *JOSA A,* vol. 29, no. 11, pp. 2367-2375, 2012.
[72] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, "Influence of the overlap parameter on the convergence of the ptychographical iterative engine," *Ultramicroscopy,* vol. 108, no. 5, pp. 481-487, 2008.
[73] X. Huang, H. Yan, R. Harder, Y. Hwu, I. K. Robinson, and Y. S. Chu, "Optimization of overlap uniformness for ptychography," *Optics Express,* vol. 22, no. 10, pp. 12634-12644, 2014.
[74] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus, and J. M. Rodenburg, "An annealing algorithm to correct positioning errors in ptychography," *Ultramicroscopy,* vol. 120, pp. 64-72, 2012/09/01/ 2012, doi: https://doi.org/10.1016/j.ultramic.2012.06.001.
[75] M. Odstrčil, M. Holler, and M. Guizar-Sicairos, "Arbitrary-path fly-scan ptychography," *Optics express,* vol. 26, no. 10, pp. 12585-12593, 2018.
[76] X. Huang *et al.*, "Fly-scan ptychography," *Scientific reports,* vol. 5, p. 9074, 2015.
[77] A. Maiden, G. Morrison, B. Kaulich, A. Gianoncelli, and J. Rodenburg, "Soft X-ray spectromicroscopy using ptychography with randomly phased illumination," *Nature communications,* vol. 4, no. 1, pp. 1-6, 2013.
[78] M. Guizar-Sicairos, M. Holler, A. Diaz, J. Vila-Comamala, O. Bunk, and A. Menzel, "Role of the illumination spatial-frequency spectrum for ptychography," *Physical Review B,* vol. 86, no. 10, p. 100103, 2012.
[79] M. Stockmar *et al.*, "Near-field ptychography: phase retrieval for inline holography using a structured illumination," *Scientific reports,* vol. 3, p. 1927, 2013.
[80] K. S. Morgan, D. M. Paganin, and K. K. Siu, "X-ray phase imaging with a paper analyzer," *Applied Physics Letters,* vol. 100, no. 12, p. 124102, 2012.
[81] A. J. Morgan, H. M. Quiney, S. Bajt, and H. N. Chapman, "Ptychographic X-ray speckle tracking," *Journal of applied crystallography,* vol. 53, no. 3, pp. 760-780, 2020.
[82] M. Stockmar *et al.*, "X-ray nanotomography using near-field ptychography," *Optics express,* vol. 23, no. 10, pp. 12720-12731, 2015.
[83] A. Diaz, P. Trtik, M. Guizar-Sicairos, A. Menzel, P. Thibault, and O. Bunk, "Quantitative x-ray phase nanotomography," *Physical Review B,* vol. 85, no. 2, p. 020104, 2012.
[84] I. Zanette *et al.*, "Phase-Vortex Removal for Quantitative X-Ray Nanotomography with Near-Field Ptychography," *Physical Review Applied,* vol. 14, no. 6, p. 064078, 2020.
[85] S. Jiang *et al.*, "Wide-field, high-resolution lensless on-chip microscopy via near-field blind ptychographic modulation," *Lab on a Chip,* vol. 20, no. 6, pp. 1058-1065, 2020.
[86] R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Optical Engineering,* vol. 21, no. 5, p. 215829, 1982.
[87] R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," *JOSA A,* vol. 9, no. 7, pp. 1072-1085, 1992.

[88] P. Li, D. J. Batey, T. B. Edo, A. D. Parsons, C. Rau, and J. M. Rodenburg, "Multiple mode x-ray ptychography using a lens and a fixed diffuser optic," *Journal of Optics,* vol. 18, no. 5, p. 054008, 2016.
[89] H. Yang, T. J. Pennycook, and P. D. Nellist, "Efficient phase contrast imaging in STEM using a pixelated detector. Part II: Optimisation of imaging conditions," *Ultramicroscopy,* vol. 151, pp. 232-239, 2015.
[90] C. M. O'Leary, G. T. Martinez, E. Liberti, M. J. Humphry, A. I. Kirkland, and P. D. Nellist, "Contrast transfer and noise considerations in focused-probe electron ptychography," *Ultramicroscopy,* vol. 221, p. 113189, 2021.
[91] D. J. Batey *et al.*, "Reciprocal-space up-sampling from real-space oversampling in x-ray ptychography," *Physical Review A,* vol. 89, no. 4, p. 043812, 2014.
[92] L. Zhou *et al.*, "Low-dose phase retrieval of biological specimens using cryo-electron ptychography," *Nature communications,* vol. 11, no. 1, pp. 1-9, 2020.
[93] J. Song *et al.*, "Atomic resolution defocused electron ptychography at low dose with a fast, direct electron detector," *Scientific reports,* vol. 9, no. 1, pp. 1-8, 2019.
[94] A. M. Maiden, M. C. Sarahan, M. D. Stagg, S. M. Schramm, and M. J. Humphry, "Quantitative electron phase imaging with high sensitivity and an unlimited field of view," *Scientific reports,* vol. 5, p. 14690, 2015.
[95] T. Latychevskaia, P. Formanek, C. T. Koch, and A. Lubk, "Off-axis and inline electron holography: Experimental comparison," *Ultramicroscopy,* vol. 110, no. 5, pp. 472-482, 2010.
[96] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson, "Dynamic imaging using ptychography," *Physical review letters,* vol. 112, no. 11, p. 113901, 2014.
[97] X. Shi, N. Burdet, D. Batey, and I. Robinson, "Multi-modal ptychography: Recent developments and applications," *Applied Sciences,* vol. 8, no. 7, p. 1054, 2018.
[98] P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," *Nature,* vol. 494, no. 7435, pp. 68-71, 2013.
[99] D. Shindo and Z. Akase, "Direct observation of electric and magnetic fields of functional materials," *Materials Science and Engineering: R: Reports,* vol. 142, p. 100564, 2020.
[100] M. R. McCartney, R. E. Dunin-Borkowski, and D. J. Smith, "Quantitative measurement of nanoscale electrostatic potentials and charges using off-axis electron holography: Developments and opportunities," *Ultramicroscopy,* vol. 203, pp. 105-118, 2019.
[101] P. A. Midgley and R. E. Dunin-Borkowski, "Electron tomography and holography in materials science," *Nature materials,* vol. 8, no. 4, p. 271, 2009.
[102] V. Özdöl, D. Tyutyunnikov, C. Koch, and P. Van Aken, "Strain mapping for advanced CMOS technologies," *Crystal Research and Technology,* vol. 49, no. 1, pp. 38-42, 2014.
[103] M. Uchida *et al.*, "Topological spin textures in the helimagnet FeGe," *Physical Review B,* vol. 77, no. 18, p. 184402, 2008.
[104] S. Bajt, A. Barty, K. A. Nugent, M. McCartney, M. Wall, and D. Paganin, "Quantitative phase-sensitive imaging in a transmission electron microscope," *Ultramicroscopy,* vol. 83, no. 1-2, pp. 67-73, 2000.
[105] V. Volkov and Y. Zhu, "Lorentz phase microscopy of magnetic materials," *Ultramicroscopy,* vol. 98, no. 2-4, pp. 271-281, 2004.
[106] A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine," *Optica,* vol. 4, no. 7, pp. 736-745, 2017/07/20 2017, doi: 10.1364/OPTICA.4.000736.
[107] A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine: supplementary material."
[108] D. Fleisch and L. Kinnaman, *A student's guide to waves*. Cambridge University Press, 2015.
[109] J. Goodman, *Introduction to Fourier optics*, 3rd edition

ed. Englewood, Colorado

Roberts & Company, 2005.

[110] D. G. Voelz, *Computational fourier optics: a MATLAB tutorial*. SPIE press Bellingham, WA, 2011.
[111] M. R. McCartney and D. J. Smith, "Electron holography: phase imaging with nanometer resolution," *Annu. Rev. Mater. Res.,* vol. 37, pp. 729-767, 2007.
[112] F. Van der Veen and F. Pfeiffer, "Coherent x-ray scattering," *Journal of Physics: Condensed Matter,* vol. 16, no. 28, p. 5003, 2004.
[113] P. Li, "Investigations and improvements in ptychographic imaging," University of Sheffield, 2016.
[114] L. Mandel and E. Wolf, "Coherence properties of optical fields," *Reviews of modern physics,* vol. 37, no. 2, p. 231, 1965.
[115] Y. Kashyap, H. Wang, and K. Sawhney, "Two-dimensional transverse coherence measurement of hard-X-ray beams using near-field speckle," *Physical Review A,* vol. 92, no. 3, p. 033842, 2015.
[116] P. Cloetens, R. Barrett, J. Baruchel, J.-P. Guigay, and M. Schlenker, "Phase objects in synchrotron radiation hard x-ray imaging," *Journal of Physics D: Applied Physics,* vol. 29, no. 1, p. 133, 1996.
[117] W. O. Saxton and W. Baumeister, "The correlation averaging of a regularly arranged bacterial cell envelope protein," *Journal of microscopy,* vol. 127, no. 2, pp. 127-138, 1982.
[118] M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *Journal of structural biology,* vol. 151, no. 3, pp. 250-262, 2005.
[119] P. A. Penczek, "Resolution measures in molecular electron microscopy," in *Methods in enzymology*, vol. 482: Elsevier, 2010, pp. 73-100.
[120] B. Haas, J.-L. Rouviere, V. Boureau, R. Berthier, and D. Cooper, "Direct comparison of off-axis holography and differential phase contrast for the mapping of electric fields in semiconductors by transmission electron microscopy," *Ultramicroscopy,* vol. 198, pp. 58-72, 2019.
[121] V. Boureau, M. Staňo, J.-L. Rouvière, J.-C. Toussaint, O. Fruchart, and D. Cooper, "High-sensitivity mapping of magnetic induction fields with nanometer-scale resolution: comparison of off-axis electron holography and pixelated differential phase contrast," *Journal of Physics D: Applied Physics,* vol. 54, no. 8, p. 085001, 2020.
[122] J. Li, M. R. McCartney, and D. J. Smith, "Semiconductor dopant profiling by off-axis electron holography," *Ultramicroscopy,* vol. 94, no. 2, pp. 149-161, 2003.
[123] A. Béché, J. Rouvière, J. Barnes, and D. Cooper, "Dark field electron holography for strain measurement," *Ultramicroscopy,* vol. 111, no. 3, pp. 227-238, 2011.
[124] T. M. Godden, A. Muñiz-Piniella, J. D. Claverley, A. Yacoot, and M. J. Humphry, "Phase calibration target for quantitative phase imaging with ptychography," *Optics express,* vol. 24, no. 7, pp. 7679-7692, 2016.
[125] S. McDermott and A. Maiden, "Near-field ptychographic microscope for quantitative phase imaging," *Optics express,* vol. 26, no. 19, pp. 25471-25480, 2018.
[126] D. Paganin, *Coherent X-ray optics*. Oxford University Press on Demand, 2006.
[127] A. L. Robisch, K. Kröger, A. Rack, and T. Salditt, "Near-field ptychography using lateral and longitudinal shifts," *New Journal of Physics,* vol. 17, no. 7, p. 073033, 2015.
[128] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," *Ultramicroscopy,* vol. 109, no. 4, pp. 338-343, 2009.
[129] K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," *Optics express,* vol. 17, no. 22, pp. 19662-19673, 2009.
[130] C. Ozsoy-Keskinbora, C. B. Boothroyd, R. E. Dunin-Borkowski, P. A. Van Aken, and C. T. Koch, "Hybridization approach to in-line and off-axis (electron) holography for superior resolution and phase sensitivity," *Scientific reports,* vol. 4, no. 1, pp. 1-10, 2014.

[131] E. Voelkl and D. Tang, "Approaching routine 2π/1000 phase resolution for off-axis type holography," *Ultramicroscopy,* vol. 110, no. 5, pp. 447-459, 2010.
[132] R. M. Clare, M. Stockmar, M. Dierolf, I. Zanette, and F. Pfeiffer, "Characterization of near-field ptychography," *Optics express,* vol. 23, no. 15, pp. 19728-19742, 2015.
[133] M. Rames, Y. Yu, and G. Ren, "Optimized negative staining: a high-throughput protocol for examining small and asymmetric protein structure by electron microscopy," *JoVE (Journal of Visualized Experiments),* no. 90, p. e51087, 2014.
[134] C. Ophus *et al.*, "Efficient linear phase contrast in scanning transmission electron microscopy with matched illumination and detector interferometry," *Nature communications,* vol. 7, no. 1, pp. 1-7, 2016.
[135] Y. Takahashi, T. Akashi, A. Sato, T. Tanigaki, H. Shinada, and Y. Murakami, "Automated acquisition of vast numbers of electron holograms with atomic-scale phase information," *Microscopy,* vol. 69, no. 2, pp. 132-139, 2020.
[136] S. Konings, M. Kuijper, J. Keizer, F. Grollios, T. Spanjer, and P. Tiemeijer, "Advances in Single Particle Analysis Data Acquisition," *Microscopy and Microanalysis,* vol. 25, no. S2, pp. 1012-1013, 2019.
[137] H. Fujioka and K. Ura, "Electron beam blanking systems," *Scanning,* vol. 5, no. 1, pp. 3-13, 1983.
[138] H. Stegmann and E. Zschech, "Advanced EELS Applications in Process Development," in *AIP Conference Proceedings*, 2005, vol. 788, no. 1: American Institute of Physics, pp. 549-557.
[139] E. Tyukalova *et al.*, "Challenges and applications to operando and in situ TEM imaging and spectroscopic capabilities in a cryogenic temperature range," *Accounts of Chemical Research,* vol. 54, no. 16, pp. 3125-3135, 2021.
[140] E. Bianco and L. F. Kourkoutis, "Atomic-Resolution Cryogenic Scanning Transmission Electron Microscopy for Quantum Materials," *Accounts of Chemical Research,* vol. 54, no. 17, pp. 3277-3287, 2021.
[141] B. H. Goodge, E. Bianco, N. Schnitzer, H. W. Zandbergen, and L. F. Kourkoutis, "Atomic-resolution cryo-STEM across continuously variable temperatures," *Microscopy and Microanalysis,* vol. 26, no. 3, pp. 439-446, 2020.
[142] C. Ophus, J. Ciston, and C. T. Nelson, "Correcting nonlinear drift distortion of scanning probe and scanning transmission electron microscopies from image pairs with orthogonal scan directions," *Ultramicroscopy,* vol. 162, pp. 1-9, 2016.
[143] L. Jones *et al.*, "Smart Align—a new tool for robust non-rigid registration of scanning microscope data," *Advanced Structural and Chemical Imaging,* vol. 1, no. 1, pp. 1-16, 2015.
[144] H. Von Harrach, "Instrumental factors in high-resolution FEG STEM," *Ultramicroscopy,* vol. 58, no. 1, pp. 1-5, 1995.
[145] F. Hüe, J. Rodenburg, A. Maiden, F. Sweeney, and P. Midgley, "Wave-front phase retrieval in transmission electron microscopy via ptychography," *Physical Review B,* vol. 82, no. 12, p. 121415, 2010.
[146] B. Schaffer, G. Kothleitner, and W. Grogger, "EFTEM spectrum imaging at high-energy resolution," *Ultramicroscopy,* vol. 106, no. 11-12, pp. 1129-1138, 2006.
[147] T. Heil and H. Kohl, "Optimization of EFTEM image acquisition by using elastically filtered images for drift correction," *Ultramicroscopy,* vol. 110, no. 7, pp. 745-750, 2010.
[148] T. Niermann and M. Lehmann, "Averaging scheme for atomic resolution off-axis electron holograms," *Micron,* vol. 63, pp. 28-34, 2014.
[149] Y. C. Wang, T. M. Chou, M. Libera, E. Voelkl, and B. G. Frost, "Measurement of polystyrene mean inner potential by transmission electron holography of latex spheres," *Microscopy and Microanalysis,* vol. 4, no. 2, pp. 146-157, 1998.
[150] D. Wolf, A. Lubk, H. Lichte, and H. Friedrich, "Towards automated electron holographic tomography for 3D mapping of electrostatic potentials," *Ultramicroscopy,* vol. 110, no. 5, pp. 390-399, 2010.

[151] J. Martinez-Carranza, K. Falaggis, and T. Kozacki, "Fast and accurate phase-unwrapping algorithm based on the transport of intensity equation," *Applied Optics,* vol. 56, no. 25, pp. 7079-7088, 2017.
[152] Y. Guo, X. Chen, and T. Zhang, "Robust phase unwrapping algorithm based on least squares," *Optics and Lasers in Engineering,* vol. 63, pp. 25-29, 2014.
[153] I. Herszterg, M. Poggi, and T. Vidal, "Two-dimensional phase unwrapping via balanced spanning forests," *INFORMS Journal on Computing,* vol. 31, no. 3, pp. 527-543, 2019.
[154] J. C. Estrada, M. Servin, and J. Vargas, "2D simultaneous phase unwrapping and filtering: A review and comparison," *Optics and Lasers in Engineering,* vol. 50, no. 8, pp. 1026-1029, 2012.
[155] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio science,* vol. 23, no. 4, pp. 713-720, 1988.
[156] K. Itoh, "Analysis of the phase unwrapping algorithm," *Applied optics,* vol. 21, no. 14, pp. 2470-2470, 1982.
[157] F. Monaco *et al.*, "A comparison between holographic and near-field ptychographic X-ray tomography for solid oxide cell materials," *Materials Characterization,* vol. 187, p. 111834, 2022.
[158] J. Neubauer *et al.*, "Evolution of the particle size distribution of tricalcium silicate during hydration by synchrotron X-ray nano-tomography," *Cement and Concrete Research,* vol. 156, p. 106769, 2022.
[159] Y. Jiang *et al.*, "Electron ptychography of 2D materials to deep sub-ångström resolution," *Nature,* vol. 559, no. 7714, pp. 343-349, 2018.
[160] S. Sala *et al.*, "Multiscale X-ray imaging using ptychography," *Journal of Synchrotron Radiation,* vol. 25, no. 4, pp. 1214-1221, 2018.
[161] S. Gao *et al.*, "Electron ptychographic microscopy for three-dimensional imaging," *Nature communications,* vol. 8, no. 1, pp. 1-8, 2017.
[162] Y. Zhang *et al.*, "Single-particle cryo-EM: alternative schemes to improve dose efficiency," *Journal of Synchrotron Radiation,* vol. 28, no. 5, 2021.
[163] M. Couillard, "Micrometre-scale strain mapping of transistor arrays extracted from undersampled atomic-resolution images," *Micron,* vol. 148, p. 103100, 2021.
[164] M. Gajdardziska-Josifovska, M. R. McCartney, W. J. De Ruijter, D. J. Smith, J. K. Weiss, and J.-M. Zuo, "Accurate measurements of mean inner potential of crystal wedges using digital electron holograms," *Ultramicroscopy,* vol. 50, no. 3, pp. 285-299, 1993.
[165] Y. Ding, Y. Liu, K. C. Pradel, Y. Bando, N. Fukata, and Z. L. Wang, "Quantifying mean inner potential of ZnO nanowires by off-axis electron holography," *Micron,* vol. 78, pp. 67-72, 2015.
[166] A. Auslender, M. Halabi, G. Levi, O. Diéguez, and A. Kohn, "Measuring the mean inner potential of Al2O3 sapphire using off-axis electron holography," *Ultramicroscopy,* vol. 198, pp. 18-25, 2019.
[167] R. Shiloh, Y. Lereah, Y. Lilach, and A. Arie, "Sculpturing the electron wave function using nanoscale phase masks," *Ultramicroscopy,* vol. 144, pp. 26-31, 2014.
[168] W. Van den Broek *et al.*, "Towards Ptychography with Structured Illumination, and a Derivative-Based Reconstruction Algorithm," *Microscopy and Microanalysis,* vol. 25, no. S2, pp. 58-59, 2019, doi: 10.1017/S1431927619001028.
[169] R. F. Egerton, "TEM Applications of EELS," in *Electron Energy-Loss Spectroscopy in the Electron Microscope*. Boston, MA: Springer US, 2011, pp. 293-397.
[170] R. Shiloh *et al.*, "Spherical aberration correction in a scanning transmission electron microscope using a sculpted thin film," *Ultramicroscopy,* vol. 189, pp. 46-53, 2018.
[171] R. S. Pennington, C. B. Boothroyd, and R. E. Dunin-Borkowski, "Surface effects on mean inner potentials studied using density functional theory," *Ultramicroscopy,* vol. 159, pp. 34-45, 2015.

[172] V. Grillo, G. Carlo Gazzadi, E. Karimi, E. Mafakheri, R. W. Boyd, and S. Frabboni, "Highly efficient electron vortex beams generated by nanofabricated phase holograms," *Applied Physics Letters,* vol. 104, no. 4, p. 043109, 2014.
[173] H. Meltzman *et al.*, "An experimental method for calibration of the plasmon mean free path," *Journal of microscopy,* vol. 236, no. 3, pp. 165-173, 2009.
[174] K. Iakoubovskii, K. Mitsuishi, Y. Nakayama, and K. Furuya, "Thickness measurements with electron energy loss spectroscopy," *Microscopy research and technique,* vol. 71, no. 8, pp. 626-631, 2008.
[175] P. Potapov, "The experimental electron mean-free-path in Si under typical (S) TEM conditions," *Ultramicroscopy,* vol. 147, pp. 21-24, 2014.
[176] P. Thakkar, V. A. Guzenko, P.-H. Lu, R. E. Dunin-Borkowski, J. P. Abrahams, and S. Tsujino, "Fabrication of low aspect ratio three-element Boersch phase shifters for voltage-controlled three electron beam interference," *Journal of Applied Physics,* vol. 128, no. 13, p. 134502, 2020.
[177] L. Grünewald, D. Gerthsen, and S. Hettler, "Fabrication of phase masks from amorphous carbon thin films for electron-beam shaping," *Beilstein journal of nanotechnology,* vol. 10, no. 1, pp. 1290-1302, 2019.
[178] P.-H. Lu and M. Kruth, "manuscript in preparation."
[179] F. Allars, P.-H. Lu, M. Kruth, R. E. Dunin-Borkowski, J. M. Rodenburg, and A. M. Maiden, "Efficient large field of view electron phase imaging using near-field electron ptychography with a diffuser," *Ultramicroscopy,* p. 113257, 2021.
[180] A. Lubk, D. Wolf, and H. Lichte, "The effect of dynamical scattering in off-axis holographic mean inner potential and inelastic mean free path measurements," *Ultramicroscopy,* vol. 110, no. 5, pp. 438-446, 2010.
[181] C. O'Leary *et al.*, "Phase reconstruction using fast binary 4D STEM data," *Applied Physics Letters,* vol. 116, no. 12, p. 124101, 2020.
[182] M. Wu, A. Tafel, P. Hommelhoff, and E. Spiecker, "Determination of 3D electrostatic field at an electron nano-emitter," *Applied Physics Letters,* vol. 114, no. 1, p. 013101, 2019.
[183] A. H. Tavabi *et al.*, "Experimental demonstration of an electrostatic orbital angular momentum sorter for electron beams," *Physical Review Letters,* vol. 126, no. 9, p. 094802, 2021.
[184] L. de Knoop, F. Houdellier, C. Gatel, A. Masseboeuf, M. Monthioux, and M. Hÿtch, "Determining the work function of a carbon-cone cold-field emitter by in situ electron holography," *Micron,* vol. 63, pp. 2-8, 2014.
[185] L. Hou, H. Wang, J. Wang, and M. Xu, "Background-noise Reduction for Fourier Ptychographic Microscopy Based on an Improved Thresholding Method," *Current Optics and Photonics,* vol. 2, no. 2, pp. 165-171, 2018.
[186] S. McDermott, P. Li, G. Williams, and A. Maiden, "Characterizing a spatial light modulator using ptychography," *Optics Letters,* vol. 42, no. 3, pp. 371-374, 2017/02/01 2017, doi: 10.1364/OL.42.000371.
[187] M. Holler *et al.*, "X-ray ptychographic computed tomography at 16 nm isotropic 3D resolution," *Scientific reports,* vol. 4, no. 1, pp. 1-5, 2014.
[188] D. J. Batey *et al.*, "X-ray Ptychography with a Laboratory Source," *Physical Review Letters,* vol. 126, no. 19, p. 193902, 2021.
[189] Y. Yamada, K. Miyamoto, and A. Koizumi, "Size determination of latex particles by electron microscopy," *Aerosol science and technology,* vol. 4, no. 2, pp. 227-232, 1985.
[190] M. Cermola and W. H. Schreil, "Size changes of polystyrene latex particles in the electron microscope under controlled physical conditions," *Journal of Electron Microscopy Technique,* vol. 5, no. 2, pp. 171-179, 1987.

[191] V. Prabhakara, D. Jannis, G. Guzzinati, A. Béché, H. Bender, and J. Verbeeck, "HAADF-STEM block-scanning strategy for local measurement of strain at the nanoscale," *Ultramicroscopy,* vol. 219, p. 113099, 2020.
[192] R. Danev and W. Baumeister, "Cryo-EM single particle analysis with the Volta phase plate," *Elife,* vol. 5, p. e13046, 2016.
[193] A. Béché, J. Rouvière, J. Barnes, and D. Cooper, "Strain measurement at the nanoscale: Comparison between convergent beam electron diffraction, nano-beam electron diffraction, high resolution imaging and dark field electron holography," *Ultramicroscopy,* vol. 131, pp. 10-23, 2013.
[194] T. C. Petersen, V. J. Keast, and D. M. Paganin, "Quantitative TEM-based phase retrieval of MgO nano-cubes using the transport of intensity equation," *Ultramicroscopy,* vol. 108, no. 9, pp. 805-815, 2008.
[195] S. Marchesini *et al.*, "SHARP: a distributed GPU-based ptychographic solver," *Journal of applied crystallography,* vol. 49, no. 4, pp. 1245-1252, 2016.
[196] M. Wanner, D. Bach, D. Gerthsen, R. Werner, and B. Tesche, "Electron holography of thin amorphous carbon films: measurement of the mean inner potential and a thickness-independent phase shift," *Ultramicroscopy,* vol. 106, no. 4-5, pp. 341-345, 2006.
[197] A. Parvizi, W. V. d. Broek, and C. T. Koch, "Recovering low spatial frequencies in wavefront sensing based on intensity measurements," *Advanced Structural and Chemical Imaging,* vol. 2, no. 1, p. 3, 2016/03/22 2016, doi: 10.1186/s40679-016-0017-y.
[198] Z. Chen *et al.*, "Electron ptychography achieves atomic-resolution limits set by lattice vibrations," *Science,* vol. 372, no. 6544, pp. 826-831, 2021.
[199] M. Beleggia, T. Kasama, D. J. Larson, T. F. Kelly, R. E. Dunin-Borkowski, and G. Pozzi, "Towards quantitative off-axis electron holographic mapping of the electric field around the tip of a sharp biased metallic needle," *Journal of Applied Physics,* vol. 116, no. 2, p. 024305, 2014.
[200] C. Suloway *et al.*, "Automated molecular microscopy: the new Leginon system," *Journal of structural biology,* vol. 151, no. 1, pp. 41-60, 2005.
[201] D. N. Mastronarde, "Automated electron microscope tomography using robust prediction of specimen movements," *Journal of structural biology,* vol. 152, no. 1, pp. 36-51, 2005.
[202] R. M. Glaeser, D. Typke, P. C. Tiemeijer, J. Pulokas, and A. Cheng, "Precise beam-tilt alignment and collimation are required to minimize the phase error associated with coma in high-resolution cryo-EM," *Journal of structural biology,* vol. 174, no. 1, pp. 1-10, 2011.
[203] M. Linck, H. Müller, P. Hartel, S. Perl, S. Uhlemann, and M. Haider, "On the benefit of aberration correction in cryo electron microscopy," *Microscopy and Microanalysis,* vol. 26, no. S2, pp. 2156-2157, 2020.
[204] A. Koster, H. Chen, J. Sedat, and D. Agard, "Automated microscopy for electron tomography," *Ultramicroscopy,* vol. 46, no. 1-4, pp. 207-227, 1992.
[205] H. Shinotsuka, S. Tanuma, C. Powell, and D. Penn, "Calculations of electron inelastic mean free paths. X. Data for 41 elemental solids over the 50 eV to 200 keV range with the relativistic full Penn algorithm," *Surface and Interface Analysis,* vol. 47, no. 9, pp. 871-888, 2015.
[206] R. Danev and K. Nagayama, "Phase plates for transmission electron microscopy," *Methods in enzymology,* vol. 481, pp. 343-369, 2010.
[207] O. Schwartz, J. J. Axelrod, S. L. Campbell, C. Turnbaugh, R. M. Glaeser, and H. Müller, "Laser phase plate for transmission electron microscopy," *Nature methods,* vol. 16, no. 10, pp. 1016-1020, 2019.
[208] H. Ryll *et al.*, "A pnCCD-based, fast direct single electron imaging camera for TEM and STEM," *Journal of Instrumentation,* vol. 11, no. 04, p. P04006, 2016.

[209] F. Zheng *et al.*, "Quantitative measurement of charge accumulation along a quasi-one-dimensional W 5 O 14 nanowire during electron field emission," *Nanoscale,* vol. 12, no. 19, pp. 10559-10564, 2020.
[210] N. De Jonge, Y. Lamy, K. Schoots, and T. H. Oosterkamp, "High brightness electron beam from a multi-walled carbon nanotube," *Nature,* vol. 420, no. 6914, pp. 393-395, 2002.

7 Appendix – A

Near field Ptychography Data Collection Code – DigitalMicrograph code

In this part of the appendix the data collection code used to collect near field electron ptychography data is included. Note that although the functions are void type, the “&” symbol in the inputs of a function enable input to be changed by the function. Therefore, all inputs with “&” symbol in front of them are combined inputs and outputs. This section starts with the code to generate the scan pattern, execute the scan, a function to save the collected frames mid experiment and the main function which calls all previous functions. The main function should always be at the bottom, so that when the execute button is pressed errors are avoided. Due to the length of the code it will be spread across multiple pages, code fragments are numbered so that the order is explicit.

For more information and microscope command functions visit:
http://www.dmscripting.com/tutorial_microscope_commands.html

7.1 GenerateRasterScanPattern

```
void GenerateRasterScanPattern(number steps_x,number steps_y, number step_size, image &x_img, image &y_img, image
&sequencer, image &CameraTest, number ScaleX, number ScaleY, number OverStep)
{
	number counter_1
	number counter_2
	number sign = 1
	number index_1
	number index_2
	number index_3
	number ScanTestLower
	number ScanTestUpper
	//Genrate the scan pattern
	x_img[0,0,1,1] = EMGetStageX() - OverStep
	y_img[0,0,1,1] = EMGetStageY() - OverStep
	x_img[1,0,2,1] = EMGetStageX()
	y_img[1,0,2,1] = EMGetStageY()
	for (counter_1 = 1; counter_1<=steps_x; counter_1++)
	{
		for (counter_2 = 1;counter_2<=steps_y; counter_2++)
		{
			index_1 = counter_2+(counter_1-1)*steps_y-1 +1
			index_2 = counter_2+(counter_1-1)*steps_y+1
			index_3 = counter_2+(counter_1-1)*steps_y+1 +1

			if (counter_2 == steps_y)
			{
				//result("something has happened! \n") //debug statement
				x_img[index_2,0,index_3,1] = x_img[1,0,2,1] + step_size*ScaleX*counter_1;
				y_img[index_2,0,index_3,1] = y_img[1,0,2,1] - OverStep;
				CameraTest[index_2,0,index_3,1] = 0
			}
			else
			{
				x_img[index_2,0,index_3,1] = x_img[index_1,0,index_2,1];
				y_img[index_2,0,index_3,1] = y_img[1,0,2,1] + step_size*ScaleY*counter_2;
				CameraTest[index_2,0,index_3,1] = 1
			}
			sequencer[index_2,0,index_3,1] = counter_2+(counter_1-1)*steps_y;
		}
	}
	ShowImage(x_img)
	ShowImage(y_img)
	ShowImage(sequencer)
	showImage(CameraTest)
}
```

7.2 ScanStage - 1

```
void ScanStage(number steps_x, number steps_y, image x_img, image y_img, number DelayLength, image CameraTest, image &Data, image &rec_x, image &rec_y, number processing, number exposure, number binning, number dimension, string &currentpath, string &currentdirectory)
{
	//Camera objects
	object currentCCD
	object AcqParams
	//Camera variables
	number no_frames  = 1
	number bit_depth  = 8
	number Acquisition_test, getpixel_x, getpixel_y, number counter_1, number counter_2, number index_1, number index_2
	number index_3, CameraBinary
	number ImageCounter = 0
	//talk to the current camera
	currentCCD=CM_GetCurrentCamera()
	//Set acquisition Parameters
	AcqParams = CM_CreateAcquisitionParameters_FullCCD(currentCCD, processing, exposure, binning, binning)
	Acquisition_test = CM_IsValid_AcquisitionParameters(currentCCD,AcqParams)
	if(Acquisition_test == 1)
	{
		//Go to the overstep Position
		getpixel_x = getpixel(x_img,0,0)
		getpixel_y = getpixel(y_img,0,0)
		EMSetStageX(getpixel_x)//result("X = " + getpixel_x )
		EMSetStageY(getpixel_y)//result("Y = " + getpixel_y )
		sleep(DelayLength)
		//set the stage back to its initial position in the case that it has dirfted away
		getpixel_x = getpixel(x_img,0,1)
		getpixel_y = getpixel(y_img,0,1)
		EMSetStageX(getpixel_x)//result("X = " + getpixel_x )
		EMSetStageY(getpixel_y)//result("Y = " + getpixel_y )
		sleep(DelayLength)
```

7.3 ScanStage - 2

```
for (counter_1 = 1; counter_1<=steps_x; counter_1++)//10
{
	for (counter_2 = 1;counter_2<=steps_y; counter_2++)//11
	{
		//use the same indexing to cetrain of correct indexing
		index_1 = counter_2+(counter_1-1)*steps_x-1 +1 //the final +1 is for the addtional OverStep
		index_2 = counter_2+(counter_1-1)*steps_x  +1 //the final +1 is for the addtional OverStep
		index_3 = counter_2+(counter_1-1)*steps_x+1 +1 //the final +1 is for the addtional OverStep

		//move the stage to the next position
		getpixel_x = getpixel(x_img,0,index_2)
		getpixel_y = getpixel(y_img,0,index_2)
		EMSetStageX(getpixel_x)//result("X = " + getpixel_x + "\n\n")
		EMSetStageY(getpixel_y)//result("Y = " + getpixel_y + "\n\n")
		CameraBinary = getpixel(CameraTest,0,index_2)
		//delay the stage so that it can move in to position
		sleep(DelayLength)
		//Collect a Frame
		if (CameraBinary == 1)
		{
			image frame:= CM_AcquireImage( currentCCD, AcqParams )
			image DisplayUP := frame
			if(ImageCounter == 1)
			{
				showimage(DisplayUP)
			}
			showimage(DisplayUP)
			updateimage(DisplayUP)
			SaveDiffractionPattern(frame,Imagecounter, currentpath,currentdirectory)
			deleteimage(frame)
			ImageCounter = ImageCounter + 1;
			result("Image Counter = " + ImageCounter + "\n\n")
		}
		//once the stage is in position record the position data from the Microscope
		rec_x[index_2,0,index_3,1] = EMGetStageX()//getpixel_x//EMGetStageX()
		rec_y[index_2,0,index_3,1] = EMGetStageY()//getpixel_y//EMGetStageY()
	}
}
```

7.4 ScanStage - 3

```
                    //set the stage back to its initial position in the case that it has dirfted away
                    getpixel_x = getpixel(x_img,0,1)
                    getpixel_y = getpixel(y_img,0,1)
                    EMSetStageX(getpixel_x)//result("X = " + getpixel_x )
                    EMSetStageY(getpixel_y)//result("Y = " + getpixel_y )
                    sleep(DelayLength)
          }
          else
          {
                    OkDialog("\n there is a problem with Camera Acquistion Parameters!\n Please change them and try
again!\n")
          }
          showImage(rec_x)
          showImage(rec_y)
}
```

7.5 SaveDiffractionPattern

```
void SaveDiffractionPattern(image frame, number Imagecounter, string &currentpath, string currentdirectory)
{
          result("stuff is happening \n\n")
          number savecounter = Imagecounter + 1
          string defaultname = "dp" + savecounter
          string thispath
          if (Imagecounter == 0)
          {
                    if(!SaveAsDialog("Save as", defaultname, currentpath))exit(0)
                    currentdirectory = pathextractdirectory(currentpath,2)
          }
          thispath = pathconcatenate(currentdirectory,defaultname)
          saveasgatan(frame, thispath)
}
```

7.6 Main

```
//Main

number steps_x = 40

number steps_y = 40

number step_size = 0.25

number ScanPadding = 0;

number TrueScanSize

number bits = 8

number dimension = 1024

number DelayLength = 0

number processing = 3

number exposure = 0.5

number binning = 2

number RemoveApertureTest = 0

number PiezoStageTest = 0

string currentpath

string currentdirectory

number OverStep = 0.5 //5 times the normal step

number ScaleX = 1

number ScaleY = 1

//define images (images in this cause are used as vectors to hold data not as actual images)

image x_img := RealImage("x positions", bits, 1, steps_x*steps_y+4)

image y_img := RealImage("y positions", bits, 1, steps_x*steps_y+4)

image sequencer := RealImage("sequencer", bits, 1, steps_x*steps_y+4)

image CameraTest := RealImage("Camera Test", bits, 1, steps_x*steps_y+4)

GenerateRasterScanPattern( steps_x, steps_y,  step_size,  x_img,  y_img, sequencer, CameraTest, ScaleX,  ScaleY, OverStep)

RealImage ApertureData := realimage("Aperture Data",bits,dimension,dimension,1)

image rec_x := RealImage("recorded x positions", bits, 1, steps_x*steps_y+4)

image rec_y := RealImage("recorded y positions", bits, 1, steps_x*steps_y+4)

RealImage NonData := realimage("Non Aperture Data",bits,dimension,dimension,1)

image nonrec_x := RealImage("recorded x positions non aperture", bits, 1, steps_x*steps_y+4)

image nonrec_y := RealImage("recorded y positions non aperture", bits, 1, steps_x*steps_y+4)

ScanStage(steps_x, steps_y, x_img, y_img, DelayLength, CameraTest, ApertureData, rec_x, rec_y, processing, exposure, binning,
dimension, currentpath, currentdirectory);

RemoveApertureTest = ContinueCancelDialog("Please remove the selected Area Aperture. \nOnce this has been done press
Continue \nto Finsh the Experiment. Otherwise press \ncancel and the experiment will end now!")

if (RemoveApertureTest == 1)

{

    ScanStage(steps_x, steps_y, x_img, y_img, DelayLength, CameraTest, NonData, nonrec_x, nonrec_y, processing, exposure,
    binning, dimension, currentpath, currentdirectory);

}
```

8 Appendix – B

Scan Coils Testing Code – DigitalMicrograph code

Due to the limited amount of information available on the functions which directly control the microscope scan coils, this script's purpose is to get a feeling for how the scan coils operate and what are their desired inputs and outputs. Note that this code should not be executed all at once but instead gradually uncommented as information is gained. For example, do not attempt to change the scan coils until you know what a reasonable shift value is. Although all of the code here is written for image shift, beam shift functions are provided and can be used to replace the image shift as the same testing process applies. This code is untested, therefore run it at your own/microscope's risk – the author is not accountable for any damage this code causes.

8.1 Scan Coils Testing Code - 1

```
//Scan Coil Code
//void EMPrepareImageShift( )
//void EMPrepareShift( )
//void EMImageShift(number x,number y)          //Shift the image by x,y.
//void EMBeamShift( number x,number y)          //Shift the beam by x, y
//void EMSetImageShift( long shiftX, long shiftY )
//void EMGetImageShift( long *shiftX, long *shiftY )
//void EMSetBeamShift( long shiftX, long shiftY )
//void EMGetBeamShift( long *shiftX, long *shiftY )
// For a Start we try And measure the image shift of microscope as it
// currently is
number shiftX = 8
number shiftY = 9
EMGetImageShift(shiftX,shiftY)
result("\nshiftX = " + shiftX + "\n" + "shiftY = " + shiftY)
//once the image shift has been measured it will be the case of moving the
//curent object image with the scan coils manuelly with the physical
//microscope controls and then measuring the scan coil values again with
//the function again to see want a resonable input to the EMSetImageShift
//code would be if we wanted to move the object by say 100 nanometres
number shiftX2 = 8
number shiftY2 = 9
EMGetImageShift(shiftX2,shiftY2)
result("\nshiftX2 = " + shiftX2 + "\n" + "shiftY2 = " + shiftY2)
deltaShiftX = shiftX - shiftX2
deltaShiftY = shiftY - shiftY2
```

8.2 Scan Coils Testing Code -2

```
//Now that we Have an idea of resonable shift distance we can use this
//to perform a simple scan and collect some diffraction to cross
//correlate in matlab so a good sample will be the waffle grid so we
//roughly know what change in shift values correpond to a phyiscal
//distance
object currentCCD
object AcqParams
number bits = 8
number dimension = 1024
number DelayLength = 1
number processing = 3
number exposure = 1
number binning = 2
currentCCD = CM_GetCurrentCamera();
RealImage ApertureData := realimage("Cross-correlation Data",bits,dimension,dimension,5)
RealImage frame := realimage("frame",8,1024,1024,1)
//Set acquisition Parameters
AcqParams = CM_CreateAcquisitionParameters_FullCCD(currentCCD, processing, exposure, binning, binning)
Acquisition_test = CM_IsValid_AcquistionParameters(currentCCD,AcqParams)
void EMSetImageShift(shiftX,shiftY)
image frame:= CM_AcquireImage( currentCCD, AcqParams )
Data[0, 0, 0, dimension, dimension, 1]=frame
deleteimage(frame)
sleep(1)
void EMSetImageShift(shiftX + deltaShiftX,shiftY + deltaShiftY)
image frame:= CM_AcquireImage( currentCCD, AcqParams )
Data[0, 0, 1, dimension, dimension, 2]=frame
deleteimage(frame)
sleep(1)
void EMSetImageShift(shiftX - deltaShiftX,shiftY + deltaShiftY)
Data[0, 0, 2, dimension, dimension, 3]=frame
deleteimage(frame)
sleep(1)
void EMSetImageShift(shiftX + deltaShiftX,shiftY - deltaShiftY)
Data[0, 0, 3, dimension, dimension, 4]=frame
deleteimage(frame)
sleep(1)
void EMSetImageShift(shiftX - deltaShiftX,shiftY - deltaShiftY)
Data[0, 0, 4, dimension, dimension, 5]=frame
deleteimage(frame)
sleep(1)
```