Advances in Ptychography

Tega Edo

A Thesis submitted for the degree of

Doctor of Philosophy

Department of Electronics and Electrical Engineering

University of Sheffield

England

September 2011

Abstract

Ptychography aims to completely revolutionise imaging in visible light, X-rays and Electron wavelengths by providing a robust platform for sub-Nyquist high-resolution real-time imaging. This thesis explores the framework of the very promising implementation of ptychography called the Ptychographic Iterative Engine (PIE). The PIE algorithm provides an elegant solution to the phase problem that facilitates recovery of complex representations of both the illuminating wave and the object of interest. The aim of this thesis is to present work done on the study of the machinery behind the PIE algorithm. This thesis formulates the solution provided by the PIE algorithm in reciprocal space; this shows the exact minimisation routine implemented by the PIE update function and provides a unified framework for quantifying the performance of the PIE algorithm.

This work is timely because it highlights aspects of the PIE algorithm that permits practical implementation of bandwidth extrapolation of a specimen from a small detector and demonstrates the uniqueness of the corresponding solution provided by the PIE algorithm. This thesis also presents a viable scheme that utilises the redundancy of the ptychographic dataset to greatly reduce the sampling requirement on the detector; thus optimising the dataset size employed in real-time high-resolution reconstruction of the specimen over a wide field of view.

Declaration

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

Tega Edo

31st of October 2011

Acknowledgements

I would to express my thanks to my supervisor, Prof. John Rodenburg, for his guidance and support throughout the course my research. I am very grateful for the opportunity to work in the frontier of diffractive imaging and feel very privileged to be have been exposed to his research and thinking style. I would also like to take this opportunity to thank my group manager, Grainne Riley, for all the invaluable assistance and advice offered during the course of my research. I also want to thank the staff of Phase Focus Ltd for providing the experimental equipment used in investigating the resolution improvement scheme of ptychography and I am grateful to Andy Maiden and Martin Humphry for their assistance in these experiments.

I also want to express my gratitude to the members of my research group for the stimulating conversations that impacted on the direction of my research and for the support of Andy Maiden, Aaron Hurley, especially Darren Batey and John Rodenburg during the write up stage. I am indebted to the late Andy Hurst for his invaluable assistance, both academic and non-academic, during the course of my research. He is greatly missed.

Publications

Edo, T.B., Zhang, F. and Rodenburg, J.M. (2010) 'Resolution improvement in coherent diffractive imaging (ptychography)', Proc. SPIE 7729, 77291H

Edo, T.B., Sweeney, F., Lui, C. and Rodenburg, J.M. (2010) 'Noise limit on practical electron ptychography', *J. Phys.: Conf. Ser*. **241** 012065

Contents

List of Figures

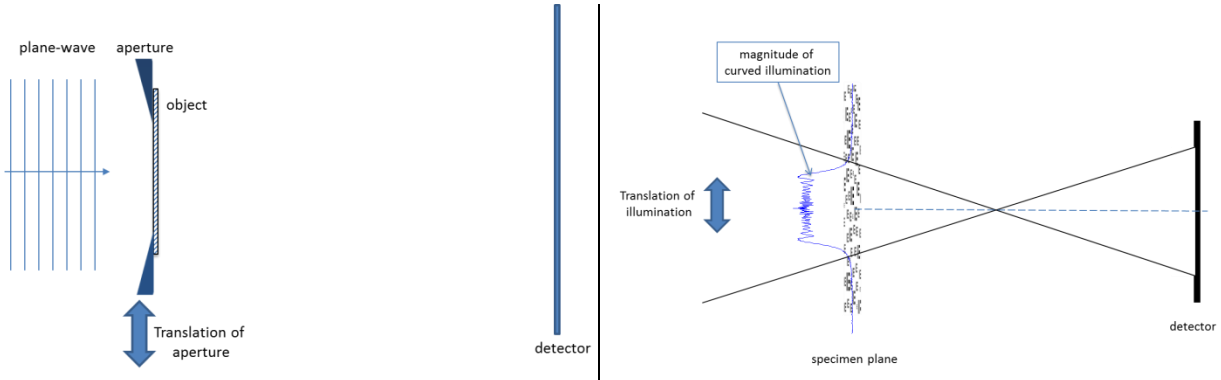
Chapter 1

1 Introduction

This thesis presents a novel theoretical analysis, backed up by computational modelling, of an imaging technique called ptychography, which in the last few years has been the subject of intense research. This approach does not require the use of a lens to form the image and as a result circumvents the limitations of lenses such as aberrations, which restrict the maximum resolution that can be achieved in conventional imaging systems. The technique works at visible light wavelengths, for both transmission and reflection geometry. Although resolution is not an issue at visible light wavelengths, where very good lenses are available, ptychography nevertheless can benefit light optics. In particular, it provides a very accurate measure of the phase of the exit wave from the specimen. Because there is no need for a lens, it can also allow for very long working distances to be used, say for the imaging of live cells through thick vessels such as petri dishes.

However, the most important application of ptychography is at X-ray and electron wavelengths where it promises to revolutionise the quality and resolution of atomic-scale transmission microscopy. Specifically the resolution of a ptychographic imaging system is

only limited by the size of the detector, which contrasts with the conventional imaging methods that require aberration correction over the angle of the wavefield processed by the lens. Moreover, ptychography provides the means of solving for the complex optical potential of the specimen and as a result eliminates the need for different contrast mechanisms such as bright field Zernicke phase contrast or diffraction contrast, which introduce serious artefacts into the image. The basic experimental setup for the technique comprises a detector placed downstream from the specimen. The specimen is illuminated by a localised wavefield that is translated across it, as shown in Fig. 1.1. In Fig 1.1a the aperture is translated across the specimen in order to illuminate different parts from which the ptychographic diffraction patterns are collected. The configuration of Fig. 1.1b, where a curved wave illuminates the specimen, is the main thrust of this thesis. A typical diffraction pattern from this setup is composed of a Gabor Hologram at the central part and intensities outside this region. The central bright disc region is called the bright field and the outside is called the dark field. Whilst Gabor holography requires a highly transparent specimen, this is not a requirement of ptychography and the dark field data provides the means for increasing the resolution, without using a larger lens.



(a) (b)

FIGURE 1.1 – Model setup for ptychography, which employs (a) an aperture illuminated by a plane wave (b) a curved wave illumination, made by a lens that is not shown here.

Ptychography utilises a set of diffraction patterns generated by moving a finite illumination beam over the specimen, with a stepping size that is equal to a substantial fraction of the illumination size, as illustrated in Fig. 1.1. The stepping size and number of diffraction patterns determines the redundancy of the ptychographic dataset, when compared to other imaging diffractive imaging methods that employ only a single diffraction pattern. This allows robust recovery of complex objects using iterative algorithms that compute the phase of the diffraction patterns via a deconvolution process. A highly successful deconvolution algorithm used to retrieve specimen information from such diffraction patterns is called the Ptychographic Iterative engine (PIE). The PIE algorithm has proved to be a very robust, especially in the presence of noise. Furthermore, an extension of the fundamental algorithm called the extended-PIE (ePIE) provides the means for simultaneously solving for the object and illuminating wave using the ptychographic dataset. The ePIE algorithm has been widely adopted and is currently used at five X-ray synchrotrons worldwide. To date, these algorithms are used because they work, however the underlying mechanism that underpins their operation is not well understood.

This thesis undertakes the first detailed mathematical analysis of the PIE algorithm in order to provide a better understanding of its operation. It explains the exact mechanism that the PIE algorithm employed in object recovery and also provides the framework for investigating further extensions to both the PIE and ePIE algorithm, especially when count rates are low so that noise is a significant issue in the effectiveness of the technique. It is shown that even though the PIE algorithm was invented for the inverse solution of ptychographic data sets, a key part of the algorithm, called the 'update function' can be used on single diffraction patterns with great effect, especially in the case when the illumination at the specimen is curved. The thesis also investigates various improvements and extensions of the technique, including resolution improvement via an algorithm which can recover not only the phase of

the diffraction pattern, but also the phase and modulus of parts of the diffraction pattern that have not actually been measured. It also considers sub-Nyquist sampling of the diffraction pattern: to date, an absolutely key underlying assumption of all so-called diffractive imaging methods (coherent diffractive imaging, or CDI) is that the diffraction pattern must be sampled at twice the Nyquist sampling of the underlying complex wave. Here it is shown for the first time that in the case of ptychography, redundancy in the data can reduce this required sampling by a staggering factor of 5.

Although at the time of writing electron ptychography is not well developed (though it is rapidly being developed), the results presented here have some very significant implications for the future of electron imaging. In particular, the curved illumination configuration, which for experimental reasons is the only practical implementation of electron ptychography, offers an easy route to double the resolution of a microscope run in the scanning transmission mode. Because in this mode only a bright central region of the diffraction pattern needs to be processed, the dynamic range of the detector does not need to be very high (as is the case if the dark-field data is processed simultaneously). Given the extreme difficulty of improving the resolution of electron imaging by further development of aberration-corrected lens technology, a possible improvement by a factor of two using the work presented here should be very significant.

Chapter 2 presents the theory of conventional imaging and discusses some of its limitations. This chapter also reviews the literature of diffractive imaging methods, which provide viable alternatives to conventional imaging methods and it also presents a detailed description of the PIE and ePIE algorithms.

Chapter 3 investigates the impact of the illumination type on the quality of the reconstructions when the algorithm processes a single diffraction pattern. This allows the

optimisation of diffractive imaging setup and highlights the need to discuss solubility of the phase problem in terms of interference rather than sampling. This is because interference provides a general explanation that cuts across different experimental setups. Chapter 3 also formulates the PIE algorithm in reciprocal space to provide a framework for discussing the behaviour of the algorithm in later chapters.

Chapter 4 investigates the impact of experimental parameters on the PIE algorithm, in calculations that apply the algorithm to a single diffraction pattern for the setup shown in Fig. 1.1b. These calculations use a single diffraction pattern in order to consider the algorithm in isolation, so that the deconvolution process provided by the algorithm is decoupled from the benefits of the multiple diffraction patterns available in ptychography. This work is also applicable to diffractive imaging experiments that use a single diffraction pattern. The parameters examined include the number of counts in the diffraction pattern, the condenser aperture and the defocus. The defocus calculations also explore the issue of sampling on this setup. This investigation provides additional insight into the inner workings of the deconvolution process, provided by the PIE algorithm.

In Chapter 5, the error tolerance of the ePIE algorithm is investigated in calculations that use the model setup of Fig. 1.1b. These calculations examine several error sources including the number of counts in ptychographic diffraction patterns and incorrect estimation of the illuminating wavefield. The wavefield is characterised by the defocus and astigmatism parameters. Furthermore, the distribution of counts information in a set of diffraction patterns is also investigated, in order to determine the optimum data collection strategy for ptychographic experiments.

Chapter 6 explores the expression of information in a single diffraction pattern and the diversity that is introduced by moving from single diffraction pattern calculations to multiple

diffraction pattern calculations. The investigation was undertaken for the setup shown in Fig. 1.1b using electron parameters. This analysis provides insights that facilitate bandwidth extrapolation or resolution improvement, which has significant implications in electron microscope imaging because it could potentially double the resolution of even the latest aberration corrected microscope. It also improves the resolution of uncorrected microscopes by a factor of two without the need for aberration correctors. Chapter 6 also explored sub-Nyquist sampling of ptychographic diffraction patterns and the findings of this work radically call into question the conventional wisdom of the sampling requirement for diffractive imaging.

Chapter 2

2 Theory of Imaging

Section 2.1 formulates the theory of light propagation and provides the Fourier transform representation that is crucial for diffractive imaging methods. Section 2.2 presents conventional imaging using Abbé's theory to highlight the connection between diffractive imaging and conventional imaging. In Section 2.3, the difficulties of conventional image formation using short wavelength radiation and works that have been done to overcome some of these limitations are discussed. Section 2.4 introduces the theory of lens aberration and discusses the condenser lens aberrations that occur when forming a finite illuminating wavefield at the specimen in short wavelength diffractive imaging experiment using electrons. Diffractive imaging methods provide a robust imaging alternative to conventional imaging methods; these are reviewed in Section 2.6. A description of early phase retrieval algorithms and a detailed outline of the PIE algorithm are given in Section 2.6.3 and Section 2.6.4. Section 2.7 discusses the sampling requirement in diffractive imaging and Section 2.8 covers the resolution of diffractive imaging methods. In Section 2.9 resolution improvement using ptychography is discussed.

2.1 Wave Propagation

This section provides an explanation for the use of the Fourier transform as a propagator in diffractive imaging. The Fourier propagator provides an efficient means for computing the scalar wave function, which encodes the probability distribution of photons or electrons over a plane given prescribed initial conditions at a different plane with a separation distance (z). The scalar wave approximation of the electromagnetic field is discussed in Section 2.1.1 where the Helmholtz equation is introduced. The Helmholtz equation is a partial differential equation derived from Maxwell's equations that describes the propagation of light in free space. The Helmholtz equation is equivalent to the time independent Klein-Gordon equation, which adequately describes the electron scattering process (Fujiwara, 1961), thus the succeeding derivations also apply to electrons. The scalar wave function is a solution to the Helmholtz equation. The Rayleigh-Sommerfeld diffraction integrals provide a solution to the Helmholtz equation in situations where boundary conditions are imposed in one of the planes. The Rayleigh-Sommerfeld diffraction integral is a convolution integral and can thus be expressed as a product using the Fourier convolution theorem. In this formulation, the propagation of waves in free space is described by a transfer function. In experiments where the scattering angles are small, the transfer function is approximated by the Fresnel transfer function. Furthermore, increasing the distance between the planes, as in the case of far field diffractive imaging experiments, gives the propagation between planes in terms of the Fourier transform.

2.1.1 Scalar wave approximation of light propagation

In conventional imaging, structural details of an object/specimen are revealed via interactions with light or electrons. Light is an electromagnetic (EM) radiation that satisfies Maxwell's equations. In free space Maxwell's equations can be expressed as separate wave equations for the electric ($\mathcal{E}$) vector of the EM wave (Goodman, 2005), given by

$$\nabla^2 \mathcal{E}_j - \frac{1}{c^2}\frac{\partial^2 \mathcal{E}_j}{\partial t^2} = 0, \tag{2.1b}$$

and the magnetic ($\mathcal{H}$) vectors of the EM wave, given by

$$\nabla^2 \mathcal{H}_j - \frac{1}{c^2}\frac{\partial^2 \mathcal{H}_j}{\partial t^2} = 0. \tag{2.1a}$$

These equations show that all the components of the electric ($\mathcal{E}_i$) and magnetic ($\mathcal{H}_i$) fields of propagating EM waves satisfy an identical wave equation and can thus be summarised by a single scalar equation given by

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0, \tag{2.1c}$$

where c is the velocity of propagation in a vacuum. The subscript j identifies the x, y and z components of the electric and magnetic fields. If the scalar field is strictly monochromatic, we get $\phi(x, y, z) = \psi(x, y, z)\exp[i\omega t]$, where the space-dependent part ($\psi(x, y, z)$) satisfies the time-independent Helmholtz equation given by

$$\nabla^2 \psi - k^2 \psi = 0, \tag{2.2}$$

where $k \equiv \sqrt{k_X^2 + k_Y^2 + k_z^2} = \omega/c = 2\pi/\lambda$.

A solution to Helmholtz equation for the case where the aperture ($\mathcal{A}$) of an opaque screen is illuminated by a monochromatic wave is given by the Rayleigh-Sommerfeld diffraction integrals in Eq. (2.3) (Bouwkamp, 1954).

$$\psi(u,v,z) = \frac{1}{2\pi}\iint_{\mathcal{A}} \psi(x,y,z=0)\frac{\partial}{\partial z}\left(\frac{\exp[ikR]}{R}\right)dxdy \tag{2.3a}$$

$$\psi(u,v,z) = -\frac{1}{2\pi}\iint_{\mathcal{A}} \frac{\partial}{\partial z}\left(\psi(x,y,z=0)\right)\frac{\exp[ikR]}{R}dxdy \tag{2.3b}$$

Here the distance from a point (x,y) in the aperture plane to another point (u,v) in the detector plane (See Fig. 2.1) is given by $R = \sqrt{(u-x)^2 + (v-y)^2 + z^2}$.

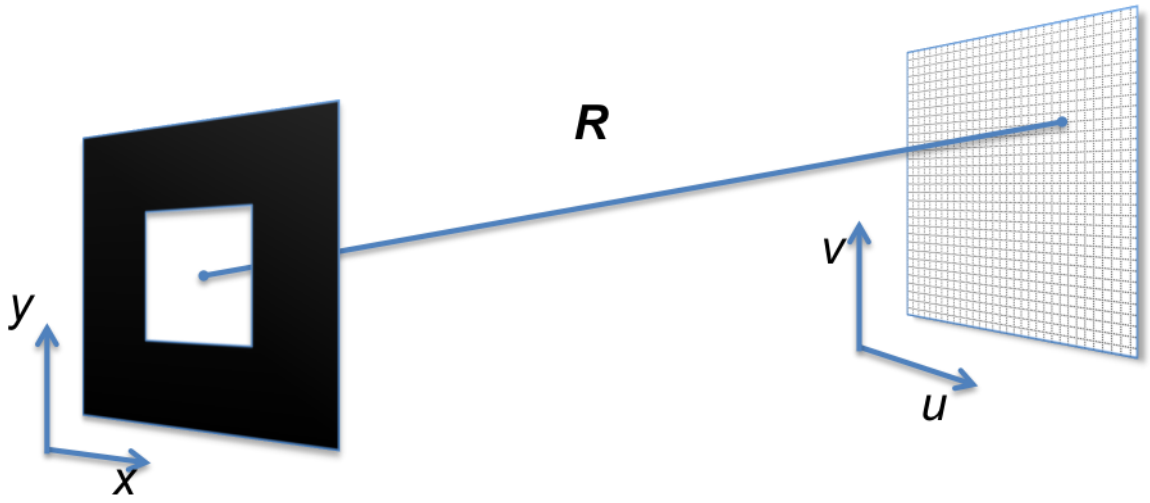


FIGURE 2.1 – illustration of aperture and detector plane coordinates

Eq. (2.3a) can be can be rewritten as Eq. (2.3b) using integration by parts together with the boundary condition ($\psi(x,y,0) = 0$ outside the aperture). Eq. (2.3a) gives the propagated wavefield in the region $z > 0$ when the values of the wavefield $\psi(x,y,0)$ are known *a priori*, while Eq. (2.3b) provides a similar solution when the derivatives of the wavefield

$\frac{\partial}{\partial z}\big(\psi(x,y,0)\big)$ take on prescribed values (Born and Wolf, 1999). In this thesis we employ Eq. (2.3a) where the complex wave distribution at the aperture plane is used to compute the propagated wavefield at the detector plane.

Eq. (2.3a) can be viewed as a convolution integral that describes a linear system with input $\psi(x,y)$ and output $\psi(u,v)$. In this view, the convolution kernel $h(x,y;z) = -\frac{1}{2\pi}\frac{\partial}{\partial z}\left(\frac{\exp[ikr]}{r}\right)$ defines the impulse response of the system with $r = \sqrt{x^2+y^2+z^2}$. It should be noted that z is not a variable in the equations of this section but a parameter of the convolution kernel that encodes the distance between the planes of propagation in Fig. 2.1.

In order to express Eq. (2.3) in the frequency domain, we review the Fourier convolution theorem. The Fourier convolution theorem enables the convolution integral of two functions $O(x,y)$ and $P(x,y)$ to be expressed as a multiplication in the frequency domain (k_X, k_Y). This facilitates efficient computation of convolution integrals via

$$\varphi(u,v) = \int O(x,y)P(u-x,v-y)dxdy, \tag{2.4a}$$

$$\varphi(k_X,k_Y) = O(k_X,k_Y)P(k_X,k_Y). \tag{2.4b}$$

The relationships between the real space representation and the Fourier space representation of these functions are given by

$$\varphi(k_X,k_Y) = \mathcal{F}\{\varphi(u,v)\} = \iint \varphi(u,v)\exp[i(uk_X+vk_Y)]dudv, \tag{2.4c}$$

$$O(k_X,k_Y) = \mathcal{F}\{O(x,y)\} = \iint O(x,y)\exp[i(xk_X+yk_Y)]dxdy, \tag{2.4d}$$

$$P(k_X,k_Y) = \mathcal{F}\{P(x,y)\} = \iint P(x,y)\exp[i(xk_X+yk_Y)]dxdy, \tag{2.4e}$$

where $\mathcal{F}$ is the Fourier transform operator. The reader should note that this thesis employs the same symbol to represent a function in real space and Fourier space, where the distinction is encoded by explicit dependence of the function on their corresponding coordinates; in the

two-dimensional representation the real space coordinate is given by (x, y) and the Fourier space coordinate given by (k_X, k_Y). Furthermore, the parameter z quantifies the distance between the aperture plane and the detector; it is not a variable but a parameter in the convolution kernel.

Using the Fourier convolution theorem of Eq. (2.4b), the frequency domain representation of the convolution integral in Eq. (2.3a) is given by

$$\psi(k_X, k_Y; z) = \psi(k_X, k_Y, 0)\exp\left[ikz\sqrt{1 - \lambda^2(k_X^2 + k_Y^2)}\right], \tag{2.5}$$

where the Fourier transforms of the wavefields and the impulse response are given by

$$\psi(k_X, k_Y; z) = \mathcal{F}\{\psi(u, v; z)\} = \iint \psi(u, v; z)\exp[i(uk_X + vk_Y)], \tag{2.6a}$$

$$\psi(k_X, k_Y; 0) = \mathcal{F}\{\psi(x, y; 0)\} = \iint \psi(x, y; 0)\exp[i(xk_X + yk_Y)], \tag{2.6b}$$

$$h(k_X, k_Y; z) = \mathcal{F}h(r) = \begin{cases} \exp\left[ikz\sqrt{1 - \lambda^2(k_X^2 + k_Y^2)/4\pi^2}\right], & \theta < 1, \\ 0, & \theta \geq 1. \end{cases} \tag{2.6c}$$

Here $\theta = \sqrt{\lambda^2(k_X^2 + k_Y^2)/4\pi^2}$ maps angles at the aperture plane and $h(k_X, k_Y; z)$ is the transfer function of free space propagation over a distance z. Eq. (2.5) is the Angular Spectrum formulation of scalar wave propagation (Sherman, 1967). It expresses the spectrum of the propagated field in terms of the spectral decomposition of the field at the aperture plane and the transfer function of free space. The propagated wavefield at the detector plane is the inverse Fourier transform of Eq. (2.5), given by

$$\psi(u, v; z) = \mathcal{F}^{-1}\psi(k_X, k_Y; z) = \iint \psi(k_X, k_Y; z)\exp(-i(uk_X + vk_Y))dk_X dk_Y, \tag{2.7}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform operator.

Most diffractive imaging experiments are performed in the paraxial regime, where diffraction angles $(\theta_x = \lambda k_x/2\pi)$ and $(\theta_y = \lambda k_y/2\pi)$ are small. The Fresnel transfer function $h_f(k_X, k_Y; z)$ in Eq. (2.8d) approximates the more general transfer function $h(k_X, k_Y; z)$ of Eq. (2.8b), when small angle approximations are made. The Fresnel transfer function h_f is given by the first two terms of the binomial expansion of the general transfer function h. The corresponding impulse response or convolution kernel is the inverse Fourier transform of the Fresnel transfer function, given by

$$h_f(x, y; z) = \mathcal{F}^{-1} h_f(k_X, k_Y; z) = (i\lambda z)^{-1} \exp[ikz] \exp\left[i\frac{k}{2z}(x^2 + y^2)\right], \tag{2.8a}$$

where the Fresnel transfer function derivation is given by

$$h(k_X, k_Y; z) = \exp\left[ikz\sqrt{1 - \lambda^2(k_X^2 + k_Y^2)}\right], \tag{2.8b}$$

$$= \exp\left[ikz\left(1 - \tfrac{1}{2}\lambda^2(k_X^2 + k_Y^2) + \cdots\right)\right], \tag{2.8c}$$

$$h_f(k_X, k_Y; z) = \exp\left[ikz\left(1 - \tfrac{1}{2}\lambda^2(k_X^2 + k_Y^2)\right)\right]. \tag{2.8d}$$

Using the Fresnel kernel $h_f(x, y; z)$, the propagation equation in the Fresnel regime is given by

$$\psi(u, v; z) = (i\lambda z)^{-1} \exp[ikz] \iint_{\mathcal{A}} \psi(x, y, 0) \exp\left[i\frac{k}{2z}((u - x)^2 + (v - y)^2)\right] dxdy, \tag{2.9a}$$

$$= Q \iint_{\mathcal{A}} \psi(x, y, 0) \exp\left[-i\frac{k}{2z}(x^2 + y^2)\right] \exp\left[-i\frac{k}{z}(ux + vy)\right] dxdy, \tag{2.9b}$$

where $Q = (i\lambda z)^{-1} \exp[ikz] \exp\left[i\frac{k}{2z}(u^2 + v^2)\right]$ is a parabolic phase distribution over the detector plane that can be moved out of the integral. The second quadratic term in the integral that varies over the aperture plane, depending on (x, y), and cannot be taken out of the integral.

In the far field (large values of z), the Fraunhofer approximation of the Fresnel integral in Eq. (2.9) gives the wavefield in the detector plane as the Fourier transform of the wavefield at the aperture plane. The Fourier relationship holds because in the Fraunhofer regime, the quadratic phase term $\left(\exp\left[-i\frac{k}{2z}(x^2+y^2)\right]\right)$ inside the integral of Eq. (2.9b) varies very slowly over the diffracting region of the specimen (defined by $\mathcal{A}$) and can be considered to be constant; this condition is satisfied for propagation distances where $z \gg k(x^2+y^2)_{\max}/2$. Furthermore, the size of the detector is larger than the specimen so that the propagated wavefield at points $u > x_{\max}$ and $v > y_{\max}$ results from interference conditions with appreciable linear phase variations over the specimen. These phase variations come from the term $\exp\left[i\frac{k}{z}(ux+vy)\right]$ in Eq. (2.9b), which introduces linear phase ramps over the specimen, whose gradient depends on the values of u and v. The term $\exp\left[i\frac{k}{z}(ux+vy)\right]$ can also be viewed as the Fourier basis with spatial frequencies of $k_X = ku/z$ and $k_Y = kv/z$. With these considerations, the wavefield in the detector plane becomes the Fourier transform of the wavefield at the specimen plane and is given by

$$\psi(u,v,z) = Q\iint_{\mathcal{A}} \psi(x,y,0)\exp\left[i\left(\left(\frac{ku}{z}\right)x+\left(\frac{kv}{z}\right)y\right)\right]dxdy, \tag{2.10a}$$

$$= Q\iint_{\mathcal{A}} \psi(x,y,0)\exp[i(k_X x + k_Y y)]dxdy. \tag{2.10b}$$

By dropping the explicit dependence on z, the diffraction pattern $I_m(u,v)$, measured by the detector is given by

$$I_m(u,v) = \left|Q\iint_{\mathcal{A}} \psi(x,y)\exp[i(k_X x + k_Y y)]dxdy\right|^2, \tag{2.11a}$$

$$= \left|\iint_{\mathcal{A}} \psi(x,y)\exp[i(k_X x + k_Y y)]dxdy\right|^2, \tag{2.11b}$$

where $|Q| = 1/\lambda z$, because it is a pure phase term.

2.2 Abbe theory of imaging

Abbe's theory of imaging describes image formation as a two-step process consisting of the analysis and synthesis of the exit wave emanating from the specimen/object (Abbe, 1873). In the analysis stage, visible light rays emanating from the object at a given angle are all focused to a point in the back focal plane of the converging lens, shown in Fig. 2.2. This process of mapping angles (θ) to a unique point in the back-focal plane is identical to the Fourier transform in Eq. (2.10) and corresponds to the 'Fourier analysis' step of image formation. The Fourier transforming property of a converging lens comes from the parabolic phase change introduced by the lens thickness function over the lens aperture (Goodman, 2005).

In the case where the object is placed against the lens, Eq. (2.10) gives the complex wavefield at the back focal plane of the lens, in which the propagation distance z equals the focal length of the lens (Goodman, 2005). The intensity pattern of the wavefield at the back-focal plane is a scaled down version of the Fraunhofer diffraction pattern that would be measured at large distances from the object.

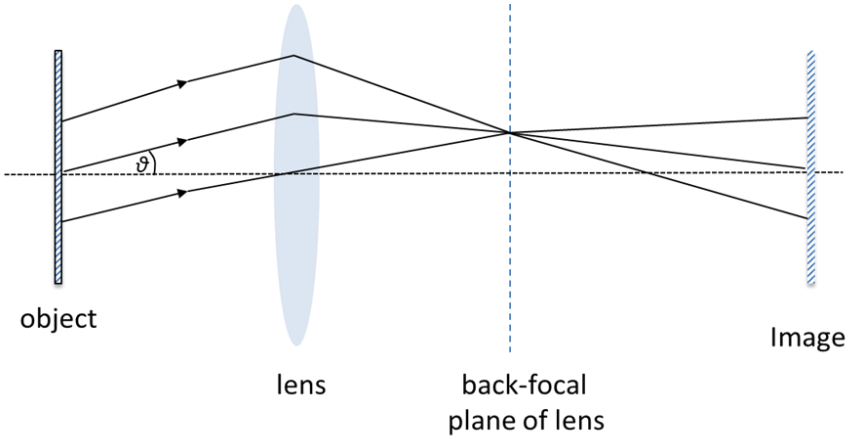


FIGURE 2.2 - Image analysis step of Abbe's theory. The analysis stage of Abbe's theory is an angular spectrum decomposition of the object transmission function. The process terminates at the back focal plane of the objective lens where linear displacement from the optical axis relates to a corresponding angular displacement about the optical axis in the object plane. This is equivalent to a Fourier transform.

In the synthesis stage, the wavefield at the back-focal plane of the lens propagates onwards to form an image of the object at the image plane. This propagation step is given by a single Fourier transform. Abbe's theory allows us to view the converging lens as a linear system that imparts phase changes to the propagated wavefield from the object in such a way as to allow information that emerges from a point in the object plane to converge to a conjugate point in the image plane.

Abbe's theory shows that knowledge of the complex distribution in the back focal plane together with Fourier processing of this information provides the complex transmission function of the object. However, measurement of the wavefield at the back focal plane using a detector gives only intensity information because detectors can only measure intensity data. The corresponding phase information (over the back focal plane) is therefore lost during the measurement process. This inability to capture the phase of complex signals occurs in many areas in science and is called the phase problem.

The analysis above assumes that the lens is perfect, with an identical frequency response over a wide range of angles. However, in practice deviations from the perfect lens occur due to a finite lens thickness, which gives rise to an angular dependent frequency response. Furthermore, fabrication errors result in a lens that generates a distorted image of the specimen. The next section discusses some of the limitations of the lens.

2.3 Limitations of the lens

Imperfections in the simple lens results in distortions of the image, as the lens no longer maps points from the object plane to corresponding points in the image plane. These distortions are characterised by the aberrations of the lens, such as astigmatism and spherical aberration.

Astigmatism arises when the lens possesses different focusing powers in the horizontal and vertical planes. As a result on axis image points are blurred to an oval shape. The effect gets worse for off-axis points where coma dominates and the oval morphs into a conical shape that points toward the optical axis. In the case of spherical aberration, the lens focussing strength increases with distance from the optic axis. This means that rays that traverse regions of the lens further from the optic axis get bent more and thus focus at planes other than the image plane, resulting in image blur. In addition to monochromatic aberrations, chromatic aberrations also arise in cases where the incident beam has an energy spread. This is because the refractive index of the lens varies with wavelength and as a result, different wavelengths focus on different planes. In visible light optics, doublets which are composed of two simple lenses paired together to function as a single lens are used to reduce the effects of aberrations. Doublets are employed because they provide additional optical surfaces that can be engineered to reduce optical aberrations. However, a good camera requires up to 10 lens or more in order to completely minimise the effects of aberrations.

The requirements of short wavelength lenses (such as X-ray and electron) are very stringent. In the case of X-rays the predominant lens used is the zone plate, which requires manufacturing precision of the outer circular zones to be the same order as the required resolution. This is impractical for high resolution imaging with hard X-rays, where a zone plate with a thickness of $\sim 8\mu m$ and outer zone lateral dimensions of $\sim 2nm$ would need to be manufactured (Schroer, 2006).

In the case of electrons, the potential used to accelerate the electrons needs to be stable to within 1 part of 10^7 in order to minimise the chromatic aberrations resulting from a spread of the beam energy. The current through the objective lens coil also needs to provide a stable enough magnetic field that preserves the path difference (less than the electron wavelength) of scattered beams so they properly reinterfere to form an image at the image plane, where

the path difference defines the relative phase values of the scattered beams. In 1936 Scherzer showed that spherical aberration term of rotationally symmetric static lenses is always non-negative and as a result cannot be eliminated by skilful design, which employs a system of round electron lenses (Scherzer, 1936). In practice spherical aberration limits the resolution of electron images to about 50 electron wavelengths (Hawkes, 2001). Scherzer (1947) and (1949) proposed the use of a non-symmetric lens to cancel the effects of aberrations but the complexity of implementing non-symmetric lens aberration correctors delayed its realisation until Zach and Haider (1995) demonstrated a correction system in the low voltage Scanning Electron Microscope (SEM). In subsequent years, Krivanek et al (1997) and Haider et al (1998) successfully implemented aberration correctors in the Scanning Transmission Electron Microscope (STEM) and Transmission Electron Microscope (TEM) configurations. Although these lens systems allow atomic resolution imaging, they require precision alignment, otherwise the image is highly distorted by the aberration of the corrector lenses and the alignment procedure is quite complicated. Furthermore, aberration correctors are very costly, where a typical corrector costs about $1M and high resolution electron microscopes (HREM) only image thin samples in order to produce interpretable images using the projection approximation.

2.4 Aberrations in imaging systems

The theory of aberrations is further explored in this section in order to provide background material to be used in subsequent chapters of this thesis. In direct imaging methods, the quality of an imaging system is determined by its ability to map points from the specimen plane on to points in the image plane. Any imperfections in the imaging system that prevent such ideal mappings are called aberrations. While this is not a problem in diffractive

imaging, it will be useful to understand the formulation of aberrations because the illumination forming optics in the case of X-ray and electron microscopy makes use of aberration parameters in order to characterise the form of the illumination function.

Since propagation from the source plane to the specimen plane can be computed via convolution integrals, this means that any departure of the radiation from the ideal path is captured by an appropriate convolution equation. Consequently, its effect can be incorporated via multiplication in the frequency domain. In the case where the illuminating radiation traverses a lens the appropriate frequency domain is the back focal plane of the lens. It is important to note that the lens has angular dependent frequency response and thus requires different transfer functions for beams incident at different angles to the optical axis (i.e. plane waves with different $\boldsymbol{k}$-vectors incident on the lens requires different transfer functions, where $\boldsymbol{k} := (k_X, k_Y)$). This section considers the case where the incident plane wave propagates along the optic axis so that a perfect lens forms a single bandlimited spot at the centre of the specimen plane. The wavefront at the back focal plane of a perfect lens is illustrated by the Gaussian reference in Fig. 2.3.

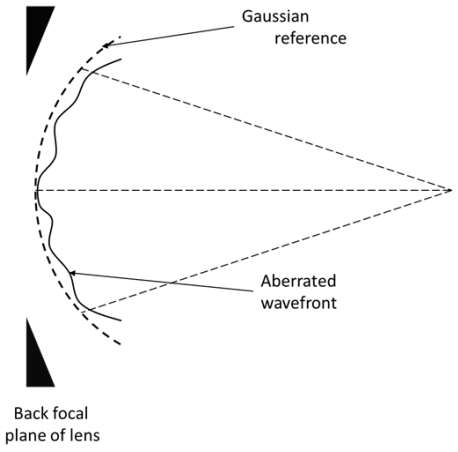


FIGURE 2.3 – illustration of aberration at the back focal plane of the lens

In a real lens, departure of the converging aberrated wave front from the Gaussian reference (see Fig. 2.3) is encoded by the phase of the aberration function (A) in Eq. (2.12) (Kirkland,

2010). The terms of the aberration function that are considered in this thesis are those with coefficients corresponding to defocus (Δz) and two-fold astigmatism (C_{12a}, C_{12b}). The aberration function $A(\theta, \phi)$ is given by

$$A(\theta,\phi) = \exp\left(\frac{2\pi i}{\lambda}\left(\frac{1}{2}\Delta z \cdot \theta^2 + \frac{1}{2}C_{12a} \cdot \theta^2 \cos(2\phi) + \frac{1}{2}C_{12b} \cdot \theta^2 \sin(2\phi)\right)\right) \tag{2.12}$$

where θ and ϕ are polar coordinate variables in the back focal plane of the condenser lens and with the change of variable given by $\theta_x = \theta\cos(\phi)$ and $\theta_y = \theta\sin(\phi)$.

Defocus

The first term in Eq. (2.12) models the case where the focus plane of the condenser lens is different from that of the specimen plane. The illumination at the specimen plane is spread over an area determined by the defocus value Δz. This is particularly important for diffractive imaging experiments where the size of the diffracting region, required for adequate sampling at the detector plane, cannot be made with a pinhole aperture. This is especially true for high resolution X-ray and electron experiments.

Two-fold Astigmatism

The last two terms in Eq. (2.12) model the effect of two-fold astigmatism. Two-fold astigmatism is also a common aberration introduced by imperfections in electron imaging optics and is the primary parameter that defines the shape of the illumination function. In traditional bright field transmission imaging, the presence of two-fold astigmatism introduces distortions into the imaging system in which points are mapped on to ovals resulting in degradation of the image. In STEM mode, this determines the minimum size of the focused

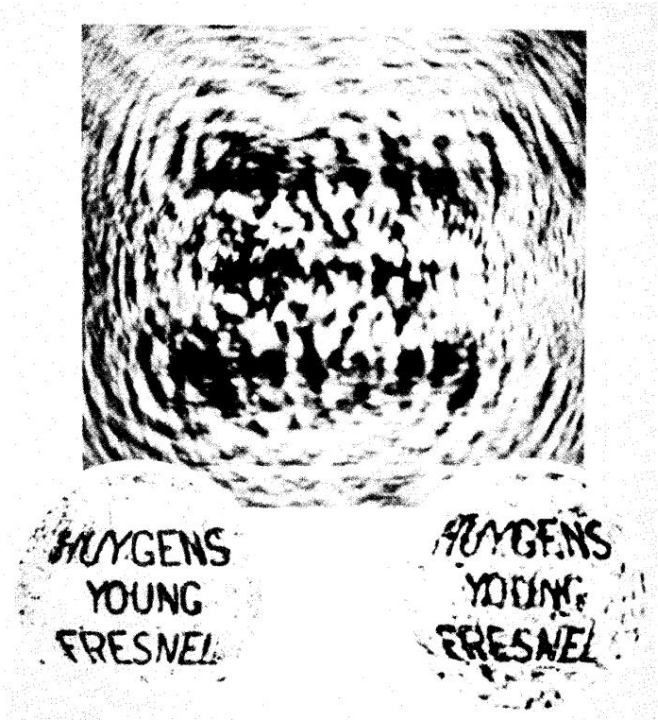
probe and thus limits the effective resolution of the system. Astigmatism does not limit the resolution in diffractive imaging because the objective lens is not required to form an image of the illuminated object. Furthermore, in the case where the incident illumination has some astigmatism, the situation is not detrimental because diffractive imaging methods can cope with structured illumination.

2.5 Overcoming the limitations of the lens

Several routes have been suggested to liberate the process of image formation from the limitations of the lens. For example, Gabor's invention of holography was motivated by the need to provide a method that circumvents the theoretical limitations of the round electron lens (Gabor, 1948). Holography is an imaging method that encodes the phase information of the propagating wavefield as intensity variations in the recording plane and thus provides a solution to the phase problem. In holography the propagated wavefield interferes with a reference wave. The reference wave comes from a part of the incident beam that does not interact with the specimen.

For a highly transparent specimen most of the incident radiation traverses the specimen without interaction and forms a strong reference at the recording plane. This setup generates an inline hologram or Gabor hologram. The hologram is developed and printed on a transparent material to give a transparency.

(a)



(b) (c)

FIGURE 2.4 – (a) Recorded hologram (b) image of micrograph (c) reconstruction. The discs that define the size of the original and reconstructed images have a diameter of 1mm. These images were reproduced from the original paper by Gabor (1949)

Fig. 2.4a shows the hologram recorded by Gabor (1949) and Fig. 2.4c shows the corresponding reconstruction from the hologram. By employing a strong curved reference wave Gabor provided the means to encode most of the phase information of the wavefield in the recorded hologram. Gabor's formulation is particularly efficient, as it requires no mathematical calculations to recover the phase of the wavefield in the recording plane. This is because the phase of the wavefield in the recording plane is almost equal to the phase of the reference beam, due to the strong reference wave and because both waves interfere at the recoding plane. That is to say adding the phasors of the strong reference wave to low magnitude diffraction components affects the phase values slightly and also results in intensity variations over the recording plane. Subsequent illumination of the transparency

with a model of the reference wave, recreates a complex wavefield at the transparency that closely approximates the original wavefield at the recording plane.

As a general imaging method, Gabor holography has the limitation that the specimen needs to be substantially transparent in order to encode phase information. In addition the Gabor holographic reconstruction is plagued by a diffuse twin image when illuminated by a curved wave. The presence of the twin image arises from the measurement process of squaring a complex function. Leith and Upatnieks (1962) devised a method that eliminates the twin image present in the Gabor holographic reconstruction. This formulation provides the means of separating the twin images by employing a tilted reference beam; the encoded phase information separates the conjugate solutions at the image plane and is called off-axis holography. Off-axis holography requires stability in the reference beam, which preserves the path difference between the reference and scattered beams in order to correctly encode phase information into the intensity variations. It also places a stringent sampling requirement on the detector because of the large phase gradient (over the recording plane) needed to separate the true image from the conjugate image.

An alternative method that provides a solution to the phase problem is X-ray crystallography. In X-ray crystallography the relative magnitudes of the Bragg peaks together with *a priori* information about the finite size of atoms, crystallographic space groups and knowledge about the constituent atoms are used to reduce the number of candidate solutions. Diffraction patterns are computed using crystallographic models of the remaining candidates and the closest to the recorded diffraction pattern gives the best guess of the crystal structure. Although crystallographic methods employ finite crystals (as opposed to perfect crystals of infinite extent), the strengths of signals at Bragg peaks are very high because the Bragg intensities are proportional to the sixth power of the linear dimension of the unit cell (in the case of a cubic unit cell). As a result, finite crystals with several unit cells diffract most of

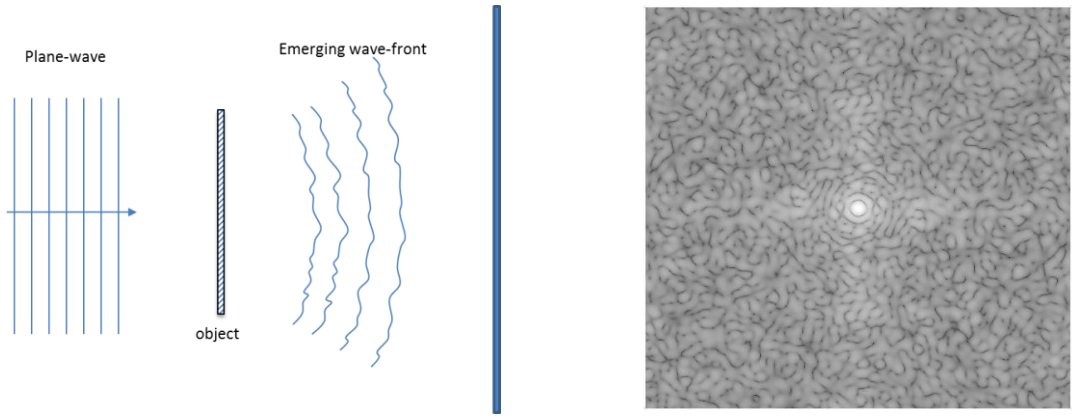
the incident radiation to Bragg peaks and thus allow crystallography with finite samples because of the high signal to noise ratio (SNR) at Bragg peaks. Additional information can also be incorporated by taking the Fourier transform of the diffraction pattern to give the Patterson function where analysis of the peaks provides information about distribution of interatomic distances in the crystal (Patterson, 1934). As an imaging technique, X-ray crystallography has the drawback of imaging only crystalline materials whereas most materials of interest are not crystalline.

A technique that combines the merits of these methods while avoiding their limitations at the expense of computational cost is diffractive imaging. In diffractive imaging, a diffraction pattern is measured, just like crystallography, from a specimen that can be either amorphous or crystalline. In order to encode phase information in the diffraction intensity, like holography, a structured beam illuminates the specimen where the convolution integral encodes the necessary phase information. In the case of curved beam illumination, the reference beam need not be as strong as that of Gabor holography.

2.6 Diffractive Imaging

Diffractive imaging is a lensless imaging method that computes the optical potential of a scattering medium from a diffraction pattern or set of diffraction patterns. Unlike conventional imaging methods, the process of image formation uses only the analysis stage of Abbé's theory, where a specimen is illuminated by coherent radiation and a Fresnel or Fraunhofer diffraction pattern is measured. Diffractive imaging provides the complex representation of the object (crystalline or amorphous) that is free from aberration distortions,

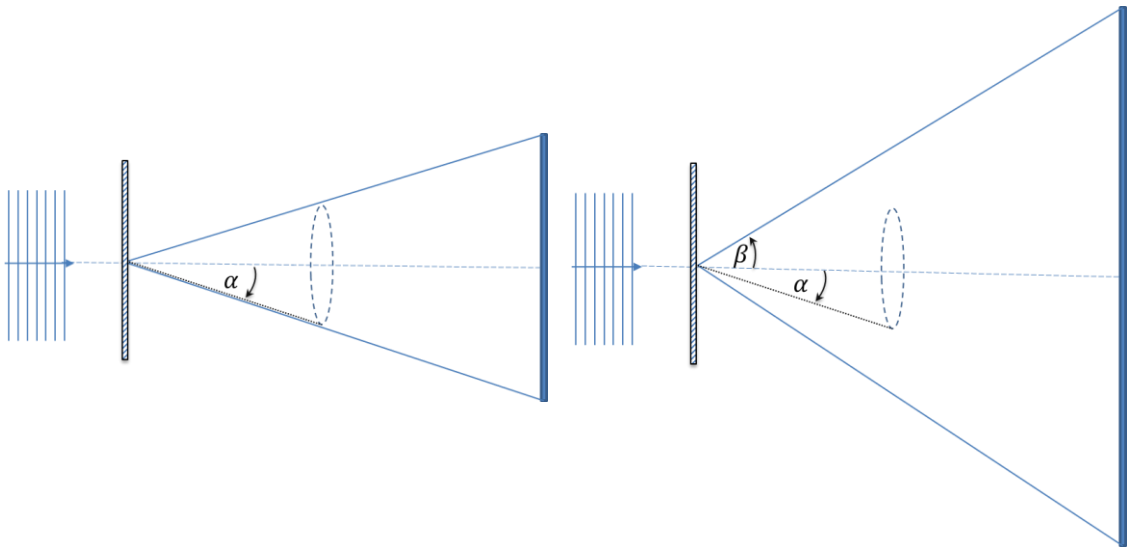
which plagues conventional imaging systems. It is rapidly becoming an invaluable imaging tool in short wavelength microscopy because it circumvents the limitations of X-ray and electron lenses. Fig. 2.5a shows a typical diffractive imaging setup that generates a diffraction pattern from a finite specimen illuminated by a plane-wave, the corresponding diffraction pattern measured in this setup is shown in Fig. 2.5b.



(a) (b)

FIGURE 2.5 – (a) illustration of diffractive imaging setup where a plane wave illuminates a finite amorphous specimen. (b) Typical far-field diffraction pattern from a finite amorphous object illuminated by a plane-wave illumination

An additional benefit of diffractive imaging is the ability to easily improve resolution by increasing the size of the detector as illustrated in Fig. 2.6, where the detector records intensities at scattering angles greater than the processing range of the lens. This is not easily achieved with conventional imaging methods because the imaging system has to correct lens aberrations to very large angles, which is very difficult to achieve, especially for electrons and X-rays.



(a) (b)

FIGURE 2.6 – (a) Lens and camera in reciprocal space. The maximum spatial frequency processed by the lens in angle space is mapped to the size of the detector, where the detector is used to capture diffraction patterns. (b) Increasing the angular size of the detector extends the range of spatial frequencies captured by the detector. This means that image resolution can be improved by simply increasing the detector size.

Furthermore, increasing the size of the detector allows strong expression of three-dimensional (3D) information of the specimen in the diffraction pattern (Rodenburg, 2008), although the sampling requirement of the detector is more stringent at high angles because speckle sizes become smaller (Rodenburg, 1988). Expression of 3D information is analogous to parallax where the top surface of the specimen moves with respect to the bottom surface. As a result the interference condition changes with viewing angle (off-centre detector pixels), thus a large detector improves our ability to extract 3D information about the specimen. By similar reasoning, structured illumination that comprises several plane-waves move 3D information to the central part of the detector. This means that diffractive imaging has the potential to recover 3D information from a diffraction pattern without tilting the object as done in tomography. Maiden, Humphry and Rodenburg (2011) recently developed a method for extracting 3D information from a set of ptychographic diffraction patterns.

2.6.1 Ptychography

Ptychography is a diffractive imaging method that employs multiple diffraction patterns to provide a unique solution to the phase problem. The concept of ptychography was developed by Walter Hoppe in a series of papers (Hoppe, 1969a), (Hoppe, 1969b) and (Hegerl and Hoppe, 1972). Hoppe's approach to the phase problem begins with the realisation that the encoding of phase information in the diffraction pattern gives rise to a complex conjugate ambiguity that is similar to the twin image problem in holography. Consequently any method devised to retrieve the phase information needs to resolve such ambiguities. He examined the phase problem in the case of crystals illuminated by a STEM probe. In this situation the diffraction pattern comprises a set of discs, which can be made to overlap by increasing the size of the lens aperture that generates the beam. Hoppe showed that shifting the illumination to generate two diffraction patterns in this configuration provided a way to decode phase information (Hoppe, 1969a) and the technique was called ptychography by Hegerl and Hoppe (1972).

In 1969, Hoppe and Strube (1969) demonstrated the validity of ptychography with periodic objects at visible light wavelengths and Hoppe (1969b) attempted to extend the idea of ptychography to non-periodic objects but without much success. The lack of success was because a large number of interfering beams made it difficult to devise a simple algorithm to decode the phase information from the diffraction patterns. As a result, Hoppe (1982) made a distinction between the applications of ptychography to crystalline and aperiodic objects because he assumed that ptychography might not provide a solution in the aperiodic case.

Further work on visible light ptychography extended classical ptychography from two diffraction patterns to a multiple diffraction pattern dataset generated by scanning the illumination across the specimen (with a step size equal to the sampling interval) and

collecting a diffraction pattern at each scan point. These experiments generated a dense ptychographic dataset that was processed using the Wigner Distribution Deconvolution method (WDDC) to generate a 1-D plot of the specimen (Friedman and Rodenburg, 1992) and later work by McCallum and Rodenburg (1992) demonstrated two-dimensional imaging of the specimen.

McCallum and Rodenburg (1993) showed that the illuminating wave and specimen information could be simultaneously retrieved from this dataset although with great difficulty. Further investigation of probe position errors by McCallum and Rodenburg (1993b) showed that translation errors from half to twice the step size can be accommodated by the WDDC method. In the case of crystalline samples, Fourier processing of a subset of the dense ptychographic dataset provides a direct route to recover the complex representation of the specimen function and was demonstrated with electrons by Nellist, McCallum and Rodenburg (1994). This work showed that a dense ptychographic dataset was not limited by the coherence function of the electron microscope. Chapman (1996) carried out similar work at the X-ray wavelengths.

The dense ptychographic dataset comprises a set of STEM diffraction patterns that have large-scale features and can thus be sampled using segmented detectors. Landauer, McCallum and Rodenburg (1995) used a quadrant detector to recover the image of a weak phase object and McCallum, Landauer and Rodenburg (1995) employed a three-sector detector for complex image reconstruction. These works employed systematic addition and subtraction of the intensities from the different sectors of the detector to infer amplitude and phase contrast information.

Although these works demonstrate the principle of diffractive imaging, the reconstruction methods lacked speed needed for routine imaging. For example, the WDDC method needs

all the data to be recorded before it can start processing. The direction of diffractive imaging research shifted to the iterative framework when Gerchberg and Saxton (1972) demonstrated an algorithm that improved on the speed of image recovery. This framework is especially robust when compared to the existing methods, at that time, because it does not require a deconvolution operation. Subsequent works along this line of research now constitutes the field of iterative phase retrieval.

2.6.2 Iterative phase retrieval

Iterative phase retrieval is a diffractive imaging framework, where the transmission function of an object is recovered using algorithms that propagate a scalar wavefield (ψ) between the specimen and detector planes. The first iterative phase retrieval algorithm was implemented by Gerchberg and Saxton (1972) where two intensity measurements (an image of the specimen and a diffraction pattern) are used to constrain the algorithm. The Gerchberg Saxton (GS) algorithm was first applied to modelled electron data comprising an electron image and a Fraunhofer diffraction pattern. Propagation between the specimen and diffraction plane was implemented with the fast Fourier transform (FFT) algorithm. At each stage of the iteration the wavefield from the specimen is propagated to the detector plane where the phase of the propagated wavefield is retained and the magnitude replaced by the squared root of the modelled diffraction intensities; the reverse process is computed and the specimen image is used to constrain the algorithm at the specimen plane and the corresponding phase distribution is retained for the next iteration. The motivation behind this scheme was the realisation that the measured intensities derive from coherent processes that use phase information. Consequently, enforcing the magnitude signal $\sqrt{I_M}$ at the detector

plane constrains the phase distribution at the specimen plane and vice versa. Chapman (1975) applied the GS algorithm to the experimental electron diffraction patterns from periodic magnetic structures and discussed some of the practical issues that surround the implementation of the GS method in the electron microscope.

The GS algorithm can be described in a projection framework. In this formulation, the operator (π_M) projects the estimated wavefield at the detector plane to the closest point that satisfies the diffraction modulus constraint. It is useful to refer the projection operator (π_M) from the detector plane to the specimen plane for analysis (Elser, 2003). The real-space view of imposing the diffraction constraint in iterative phase retrieval is given by

$$\pi_M\{\psi_n(\boldsymbol{r})\} = \mathcal{F}^{-1}\{\sqrt{I_M} \cdot \exp[i \cdot \arg(\mathcal{F}\{\psi_n(\boldsymbol{r})\})]\}. \tag{2.13}$$

Employing Eq. (2.13) in the GS algorithm gives the estimated complex function ($\tilde{\psi}_n(\boldsymbol{r})$) at the specimen plane as

$$\tilde{\psi}_n(\boldsymbol{r}) = \pi_M\{\psi_n(\boldsymbol{r})\}. \tag{2.14}$$

The complex function in Eq. (2.14) is then updated with the specimen image constraint (I_S) so that the updated object is given by

$$\psi_{n+1}(\boldsymbol{r}) = \sqrt{I_S} \cdot \exp\left[i \cdot \arg\left(\tilde{\psi}_n(\boldsymbol{r})\right)\right]. \tag{2.15}$$

In some situations in astronomy and X-ray crystallography experiments, only measurements corresponding to the Fourier transform of the object of interest are available. Fienup (1978) extended the GS algorithm to work with a single intensity measurement and the resulting algorithm was called the error-reduction (ER) algorithm. The ER algorithm utilises a single diffraction pattern and *a priori* knowledge of the object size in real space. As with the GS algorithm, the ER algorithm imposes two constraints (at the object and detector planes) in each iteration step. However, the constraint in the specimen plane takes the form of a binary mask ($A(\boldsymbol{r})$) rather than information from measured image intensity. This mask is called the

support mask and takes a value of zero in regions that do not contribute to the diffraction pattern and unity in regions that may contribute to the diffraction pattern, i.e. it defines an upper bound for the object size in real space. The running estimate of the object in the ER algorithm is given by

$$\psi_{n+1}(\boldsymbol{r}) = \begin{cases} \pi_M\{\psi_n(\boldsymbol{r})\}, & A(\boldsymbol{r}) = 1, \\ 0, & A(\boldsymbol{r}) = 0. \end{cases} \tag{2.16}$$

The input-output approach was incorporated into the ER method for cases in which the running estimate of the object function does not satisfy the size constraint in real space (Fienup, 1978). This was done to avoid stagnation problems that plagued the ER algorithm. The running estimate of the input-output method is given by

$$\psi_{n+1}(\boldsymbol{r}) = \begin{cases} \psi_n(\boldsymbol{r}), & A(\boldsymbol{r}) = 1, \\ \psi_n(\boldsymbol{r}) - \beta \cdot \pi_M\{\psi_n(\boldsymbol{r})\}, & A(\boldsymbol{r}) = 0. \end{cases} \tag{2.17}$$

This particular implementation does not update the object in the region where $A(\boldsymbol{r}) = 1$, thus it was augmented with the ER algorithm in Eq. (2.16). In this approach, Eq. (2.16) and Eq. (2.17) were used to update the object at different iterations. Fienup (1982) combined the interchanging algorithm into a single algorithm called the Hybrid-Input-Output (HIO) given by

$$\psi_{n+1}(\boldsymbol{r}) = \begin{cases} \pi_M\{\psi_n(\boldsymbol{r})\}, & A(\boldsymbol{r}) = 1, \\ \psi_n(\boldsymbol{r}) - \beta\pi_M\{\psi_n(\boldsymbol{r})\}, & A(\boldsymbol{r}) = 0. \end{cases} \tag{2.18}$$

The HIO algorithm was shown to be a special case of a general class of iterative algorithms that employ the Difference Map method (Elser, 2003). The Difference Map method updates the object with average contributions from different constraint sets and as a result does not readily get stuck in local minima of the search space. Several algorithms such as Solvent Flip

(Oszlányi and Süto, 2004), Average Successive Reflection (Bauschke, Combettes and Luke, 2002), Hybrid Projection Reflection (Bauschke, Combettes and Luke, 2003), Relaxed Average Alternating Reflection (Luke, 2005) have found applications in iterative phase retrieval. A review on the connections between the Difference Map method and these algorithms, together with visual representations of their behaviour was given by Marchesini (2007).

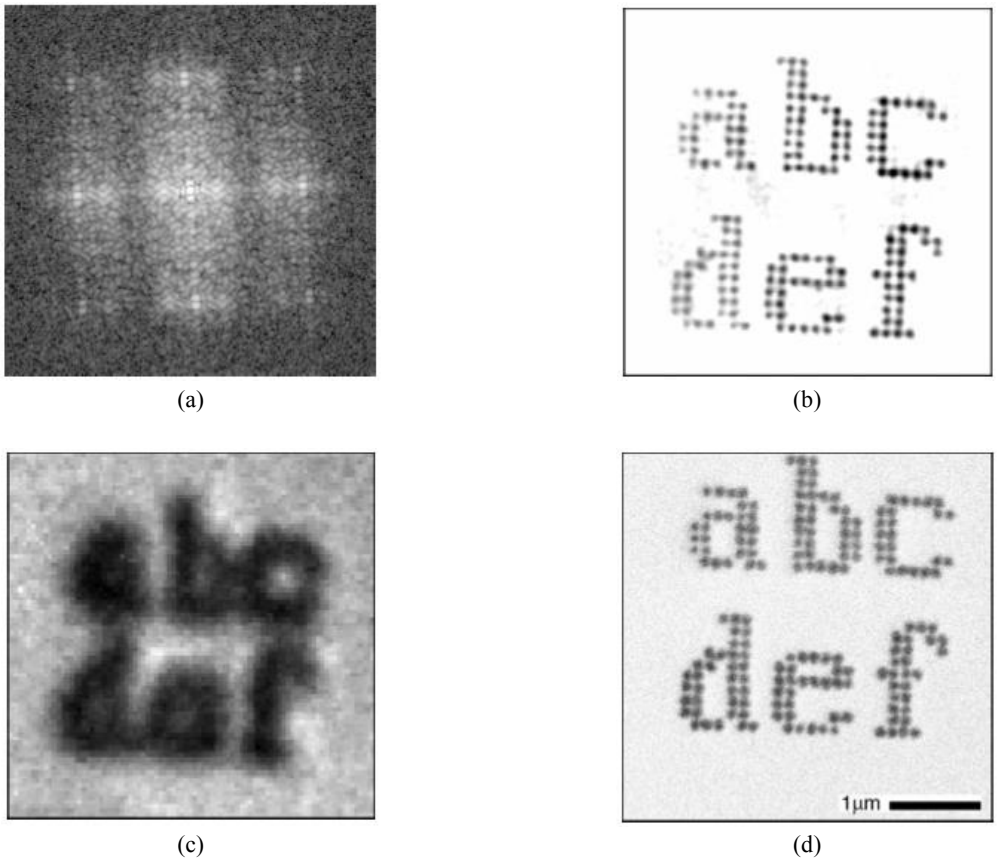


FIGURE 2.7 – (a) Recorded diffraction pattern. (b) Diffractive imaging reconstruction from the diffraction pattern. (c) visible light image of specimen (d) scanning electron microscope image of specimen. These images are reproduced from the paper by Miao et al (1999).

Miao et al. (1999) demonstrated the first experimental reconstruction at X-ray wavelengths (see Fig. 2.7b). This work was done in transmission geometry where the reconstruction algorithm employed the finite support constraint together with the positivity constraint.

Robinson et al. (2001) also imaged gold nano-crystals using hard X-rays in the reflection geometry. Improving the convergence of X-ray diffraction pattern calculations with curved illumination was explored by Nugent et al. (2005), Abbey et al. (2008) with an extension to complex constraints by Clark et al (2010). Diffractive imaging at X-ray wavelengths has been extensively investigated; a review of the literature is given in a recent paper by Nugent (2010).

Error Metric

The sum squared error between the measured intensity and the modulus square of the Fourier transform provides a suitable metric for investigating trends of the iterative algorithms. The error metric (E) measures the error between the estimated object's diffraction pattern and the measured diffraction pattern. This is given by

$$E = \sum (I_M - |\mathcal{F}\{\psi_n(\boldsymbol{r})\}|^2) \tag{2.19}$$

Gerchberg and Saxton (1978) showed that in the case of two intensity constraints (i.e. when one has a diffraction pattern and an image of the object), the error metric either decreases with iterations or stays the same, based on Parseval's theorem. Fienup (1982) showed that the case of two intensity constraints (image and diffraction pattern) and the case of a single intensity constraint (diffraction pattern) together with a support in real space are both error reduction methods, because, in both cases, the projection operations move the object guess to a point closer to the correct object function in the region outside of the support. This thesis employs the mean squared error (MSE) metric to investigate the convergence of the calculations in subsequent chapters. The MSE is given by the error in Eq. (2.19) divided by

the power of the measured diffraction pattern and provides identical convergence properties as Eq. (2.19).

2.6.3 Benefits of diversity in iterative phase retrieval

Ptychography is one of several ways of expressing diversity in a set of diffraction patterns. A technique that employs wavefront modulation of the specimen exit wave was shown to also overcome the stagnation problems of single diffraction pattern iterative calculations (Zhang, Pedrini and Osten, 2007). Diversity can also be added to the diffraction pattern via wavelength variation for non-dispersive samples (Boa et al., 2008). In relation to ptychography, Guizar-Sicairos and Fienup (2008) employed a non-linear optimisation approach on ptychographic datasets to minimise the impact of small errors in the illumination positions and small changes in the illumination size. Thibault et al (2008) used the difference map method together with the diversity of ptychographic diffraction patterns to simultaneously refine the initial estimate of the illumination in ptychographic calculations. The recovered transmission profiles of a set of 2D slices from ptychographic reconstructions using the difference map method have been successfully combined with tomographic techniques to produce 3D representations of biological specimens (Dierolf et al., 2010). Maiden and Rodenburg (2009) showed that serial processing of the ptychographic dataset using the ePIE algorithm, which is described in the next section, provides a better reconstruction than the difference map method when the diffraction patterns contain high noise.

2.6.4 The Ptychographic Iterative Engine (PIE)

In this thesis, it is important for the reader to understand the difference between two types of iterative algorithms employed called the PIE and ePIE algorithm. These algorithms are both iterative implementation of ptychography that employs a set of diffraction patterns generated from overlapping illuminated regions of the specimen. These diffraction patterns are used to solve for the specimen optical potential, where calculations that employ the PIE algorithm require *a priori* knowledge of the illumination, whereas the ePIE algorithm also solves for the illumination function, by refining the initial illumination guess. This section presents the developments that resulted in the current form of these algorithms, together with a detailed description of their implementations.

The PIE algorithm was original developed for use with short wavelength imaging but its usefulness extends beyond the realm of X-ray and electron imaging. This is because it can also be used to calculate the complex transmission/reflection functions of objects at visible light wavelengths. There are several benefits of having a complex representation of an object when compared with the conventional image. Amongst these is the ability to provide quantitative phase imaging, perform offline refocusing, long working distance and improvement of image contrast using a combination of magnitude and phase information.

Some of the benefits of ptychography for iterative phase retrieval were first demonstrated computationally with hard aperture (a top hat with sharp edges) illumination by Faulkner and Rodenburg (2004), where a precursor to the PIE algorithm extended the ER method to multiple diffraction pattern measurements and was shown to eliminate stagnation problems that plagued the ER algorithm. The algorithm was also shown to accurately recover the transmission function of complex objects without the complex conjugate ambiguity, which plagued solutions from single diffraction patterns that employed circular supports.

In the electron microscope, the diffracting region of the specimen cannot be selected with a pinhole because it is very difficult to fabricate; there is also the problem of charging around the pinhole from the electron beam so that the aperture quickly contaminates (Rodenburg, 2008). The only viable way to provide a localised spot at the specimen plane is with a focused beam made by the condenser lens of the electron microscope. At focus the illumination takes the form of the Airy function, with size inversely proportional to the angular span of the condenser aperture. Moving the specimen slightly away from the focus of the lens along the z- axis changes the beam size on the specimen. The same effect can be accomplished by adjusting the current in the condenser lens to either overfocus or underfocus the illumination at the specimen plane. This produces a spot at the specimen plane with an effective size given by the defocus. However, the illumination is extended at the specimen plane because the condenser aperture, which is at the Fourier plane of the specimen, has a finite size, leading to Fresnel type ringing effects at the edge of the illumination (see Fig. 3.2b). Consequently, the illumination does not have a finite boundary (support constraint) that ER and HIO methods need to recover the specimen function.

This difficulty was solved by Rodenburg and Faulkner (2004) with the creation of the PIE algorithm, given by

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + U_{s(n)}(\boldsymbol{r})\{\pi_M\{\psi_{s(n)}(\boldsymbol{r})\} - \psi_{s(n)}(\boldsymbol{r})\}, \tag{2.20}$$

where $O_n(\boldsymbol{r})$ represents the running estimate of the object after n iterations of the algorithm, $\pi_M\{\ldots\}$ again represents the real space view of enforcing the diffraction pattern constraint, $\psi_{s(n)}(\boldsymbol{r})$ is the exit wave immediately downstream of the specimen at the nth iteration, and $U_{s(n)}(\boldsymbol{r})$ is called the update function. The PIE algorithm facilitates soft illumination iterative phase retrieval via the introduction of the update function, which combines weighted

contributions from a set of diffraction patterns into a single running estimate of the object. The update function is given by

$$U_{s(n)}(\boldsymbol{r}) = \beta \frac{\left|P\left(\boldsymbol{r}+\boldsymbol{R}_{s(n)}\right)\right|^{l} P^{*}\left(\boldsymbol{r}+\boldsymbol{R}_{s(n)}\right)}{|P(\boldsymbol{r})|^{2}_{\max}\left(\left|P\left(\boldsymbol{r}+\boldsymbol{R}_{s(n)}\right)\right|^{2}+\delta\right)}, \tag{2.21}$$

where β acts as a feedback parameter that determines the amount of correction added to the running estimate of the object in Eq. (2.20), $|P(\boldsymbol{r})|^{2}_{\max}$ represents the maximum value of the illumination intensity, the parameter δ is a small non-zero value that prevents division by zeros in the deconvolution process and l maximises the update effect of the weighting scheme (Rodenburg and Faulkner, 2005). $P\left(\boldsymbol{r}+\boldsymbol{R}_{s(n)}\right)$ is the illumination/probe function translated by the probe position vector $\boldsymbol{R}_{s(n)}$ and $s(n)$ is used to randomly address the diffraction patterns during iterative calculations. This improves the robustness of the algorithm and was introduced by Maiden and Rodenburg (2009). In one ptychographic iteration cycle, the algorithm processes each diffraction pattern of the dataset once in a serial fashion.

The PIE algorithm has been demonstrated at visible light wavelengths by Rodenburg, Hurst and and Cullis (2007) and using hard X-ray by Rodenburg et al (2007). The PIE algorithm was also used by Hue et al (2010) to recover the phase shift introduced by ferromagnetic nano particles when probed by an electron beam; this information is not available in conventional bright field electron imaging. A flow chart of a single iteration step of the PIE algorithm is illustrated in Fig. 2.8

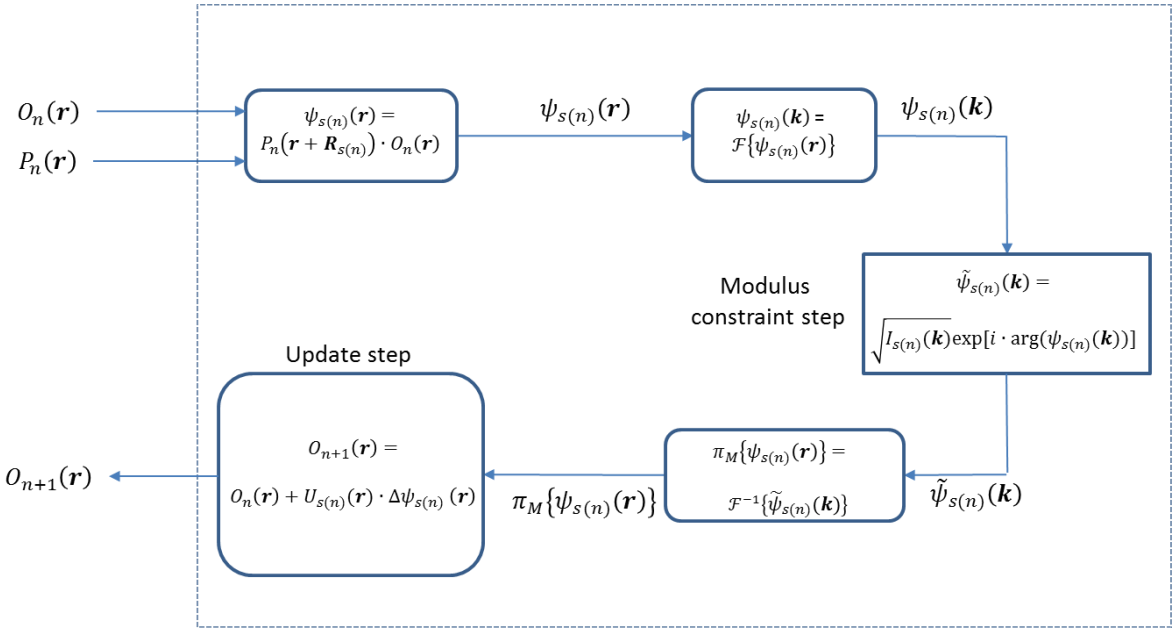


FIGURE 2.8 – illustration of a single iteration step of the PIE algorithm, where $\Delta\psi_{s(n)}(\boldsymbol{r}) = \pi_M\{\psi_{s(n)}(\boldsymbol{r})\} - \psi_{s(n)}(\boldsymbol{r})$. The variable $\boldsymbol{k}$ represents the Fourier transform basis for expanding the propagated exit wave. It corresponds to the detector coordinate in diffractive imaging experiments.

The strength of the PIE algorithm comes from the overlap of shifted versions of illumination at different scan positions. This is characterised by the overlap parameter, which measures the step size in units of the effective illumination diameter. A percentage overlap of 50% corresponds to a step size equal to half of the illumination diameter. The percentage overlap parameter was investigated by Bunk et al (2008) where it was shown that the PIE algorithm provides robust reconstruction of complex objects for an overlap parameter greater than 60%. By optimising the percentage overlap with respect to radiation dose (total exposure time) an optimum overlap range below 60% was deduced. It was also reported that a percentage overlap greater than 90% does not provide an optimal dataset because the diversity of the dataset drops since all the diffraction patterns in the ptychographic dataset have similar intensity distributions.

The PIE algorithm was extended to simultaneously calculate the object transmission function

and the illumination function by Maiden and Rodenburg (2009). The new algorithm was called the extended-PIE (ePIE). The ePIE algorithm can be compared with the blind deconvolution approach, where two unknown signals are extracted from their convolution (Ayers and Dainty, 1988). However, in the case of ptychography the redundancy of the dataset constrains the calculation enough for the ePIE algorithm to explore the solution space of complex functions. Simultaneous recovery of the illuminating wave in ptychography was first demonstrated by McCallum and Rodenburg (1993) but using the rather ill-conditioned WDDC deconvolution method. The update function in Eq. 2.21 provides the stability of the ePIE algorithm via its robust iterative deconvolution process and is further explored in Chapter 3 and Chapter 4.

In the ePIE algorithm, a value ($l = 2$) is used for the magnitude weight parameter in order to incorporate radiation flux information at the specimen plane. The radiation flux at the specimen plane is given by the intensity of the illumination. This weighting scheme assumes that regions of the specimen with high radiation flux are more strongly expressed in the diffraction pattern than regions with low radiation flux. Consequently, the ePIE algorithm employs the illumination intensity in the object update step of the algorithm. Furthermore, in the limit where the multiplicative approximation holds, the exit wave is given by the product of the illumination and the specimen function. The product symmetry between the specimen and the illumination function means that the probe update step employs the intensity of the specimen function as the corresponding update function, see Eq. (2.22). A typical ptychographic calculation that employs 64 diffraction patterns, each of which comprises 512x512 pixels, takes approximately two minutes to run 200 iterations of the ePIE algorithm. A flow chart of a single iteration step of the ePIE algorithm is illustrated in Fig. 2.9.

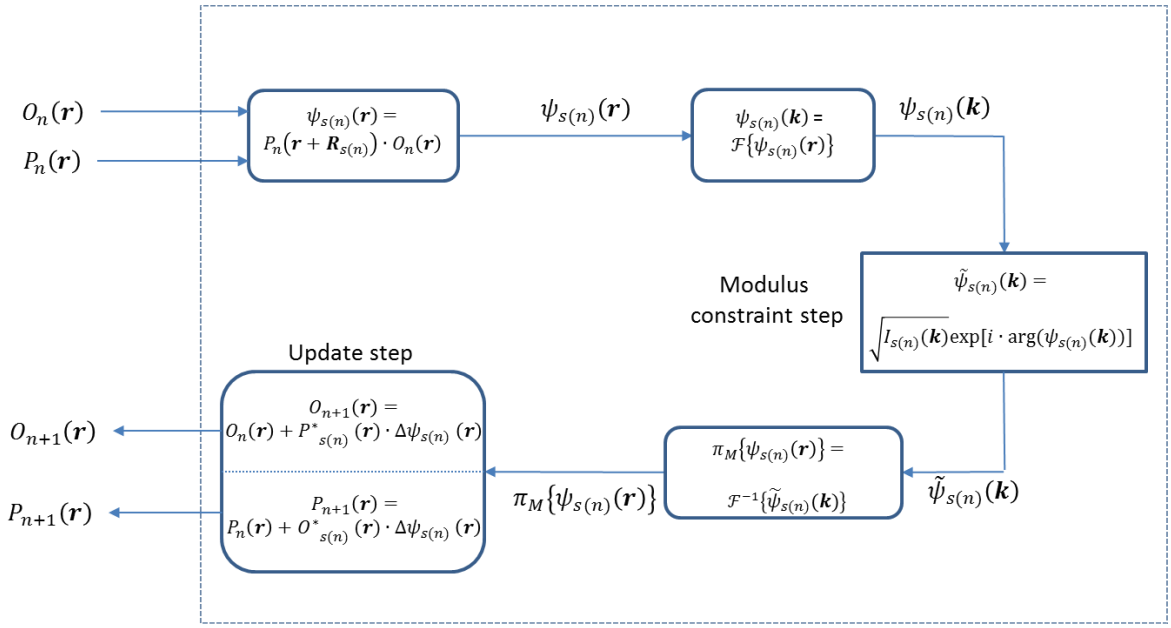


FIGURE 2.9 – illustration of a single iteration step of the ePIE algorithm, where $\Delta\psi_{s(n)}(\boldsymbol{r}) = \pi_M\{\psi_{s(n)}(\boldsymbol{r})\} - \psi_{s(n)}(\boldsymbol{r})$. The variable $\boldsymbol{k}$ represents the Fourier transform basis for expanding the propagated exit wave. It corresponds to the detector coordinate in diffractive imaging experiments. The illumination and object functions have been normalised in the update function of this figure.

In this framework the parameter δ is no longer required since the magnitude squared in the numerator cancels with the illumination intensity in the denominator when $\delta = 0$, so that the ePIE algorithm avoids the problem of dividing by zeros. For the parameter configuration $(l = 2, \delta = 0)$, the ePIE algorithm is given by

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + \beta \frac{P_n^*(\boldsymbol{r} + \boldsymbol{R}_{s(n)})}{\left|P_n(\boldsymbol{r} + \boldsymbol{R}_{s(n)})\right|^2_{\max}} \{\pi_M\{\psi_{s(n)}(\boldsymbol{r})\} - \psi_{s(n)}(\boldsymbol{r})\},$$

$$P_{n+1}(\boldsymbol{r}) = P_n(\boldsymbol{r}) + \beta \frac{O_n^*(\boldsymbol{r} + \boldsymbol{R}_{s(n)})}{\left|O_n(\boldsymbol{r} + \boldsymbol{R}_{s(n)})\right|^2_{\max}} \{\pi_M\{\psi_{s(n)}(\boldsymbol{r})\} - \psi_{s(n)}(\boldsymbol{r})\}. \quad (2.22)$$

By setting the feedback parameter to unity ($\beta = 1$) and scaling the illumination function so that $\left|P_n(\boldsymbol{r} + \boldsymbol{R}_{s(n)})\right|^2_{\max} = 1$, the update function becomes the conjugate of the illumination function (i.e. $U\left(\boldsymbol{r} + \boldsymbol{R}_{s(n)}\right) = P^*(\boldsymbol{r} + \boldsymbol{R}_{s(n)})$) during the object update step. By repeating above procedures for the probe update step, the ePIE algorithm takes the form given by

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + P_n^*\left(\boldsymbol{r} + \boldsymbol{R}_{s(n)}\right)\left\{\pi_M\left\{\psi_{s(n)}(\boldsymbol{r})\right\} - \psi_{s(n)}(\boldsymbol{r})\right\}, \tag{2.23}$$

$$P_{n+1}(\boldsymbol{r}) = P_n(\boldsymbol{r}) + O_n^*\left(\boldsymbol{r} + \boldsymbol{R}_{s(n)}\right)\left\{\pi_M\left\{\psi_{s(n)}(\boldsymbol{r})\right\} - \psi_{s(n)}(\boldsymbol{r})\right\}.$$

This form of the ePIE algorithm is convenient for further analysis of the update function, presented in Chapter 3. The ePIE algorithm has been used to obtain the complex transmission profile of the specimen and illuminating wave at visible light wavelengths (Maiden, Rodenburg and Humphry, 2010b). It has also been employed in reflection geometry to characterize the surface profile of materials (Maiden, Rodenburg and Humphry, 2010a). The ePIE algorithm has also found applications in the characterisation of X-ray beam parameters (Honig et al., 2011).

2.7 Sampling

Shannon (1949) showed that a bandlimited continuous signal is completely determined by a discrete set of sample points with a sampling interval equal to the reciprocal of twice of the signal's bandwidth. This is particularly important because it allows digitisation of continuous signals, which facilitates signal processing and image processing in the case of 2D functions. The fundamental importance of the minimum sampling interval called the Nyquist interval

was pointed out by Harry Nyquist in relation to telegraphy (Nyquist, 1928). In the present context, Shannon's sampling theorem allows a distortion-free discrete representation of either the exit wave at the specimen plane, or the propagated wavefield at the detector plane, on the condition the corresponding spectrum is band limited; here the propagated wavefield at the detector is given by the Fourier transform of the exit wave at the specimen plane.

An efficient algorithm for computing the discrete Fourier transform is the Fast Fourier transform (FFT) algorithm by Cooley and Turkey (1965). To avoid aliasing in the real space window, the maximum spatial frequencies $(|k_x|_{max}, |k_y|_{max})$ and the real space sampling intervals $(\Delta x, \Delta y)$ in a calculation that uses a square grid of NxN points must satisfy the Shannon sampling theorem. These relations are given in Eq. (2.22a) and Eq. (2.22b) (Kirkland, 2010).

$$|k_x|_{max} < \frac{1}{2\Delta x}, \tag{2.22a}$$

$$|k_y|_{max} < \frac{1}{2\Delta y}, \tag{2.22b}$$

where $|k_x|_{max} = N\Delta k_x$ and $|k_y|_{max} = N\Delta k_y$.

For the inverse Fourier transform calculation, the maximum window in the specimen plane $(|x|_{max}, |y|_{max})$ and the sampling of reciprocal space (spatial frequency domain) automatically satisfy the relationships,

$$|x|_{max} < \frac{1}{2\Delta k_x}, \tag{2.23a}$$

$$|y|_{max} < \frac{1}{2\Delta k_y}, \tag{2.23b}$$

where $|x|_{max} = N\Delta x$ and $|y|_{max} = N\Delta y$.

Eq. (2.22) requires that the specimen's spectrum be band limited (extended in real space) and Eq. (2.23) requires a finite size specimen in real space. These conflicting requirements imply that the discrete representation always has some element of aliasing. Gerchberg and Saxton (1972) showed that the effect of aliasing on diffractive imaging calculations can be minimised by judicious selection of which plane to enforce the sampling criteria. For the case of single intensity measurements, the intensity of the diffraction pattern is used to enforce the correct magnitude constraint (that is not aliased) in reciprocal space during iterative calculations. Consequently, the sampling criterion is enforced in reciprocal space so that the wavefield in the specimen plane falls to a value very close to zero in the real space calculation window.

In the case where a finite aperture is employed in real space to limit the size of the diffracting region at the specimen plane, aliasing is only present at the detector plane. The magnitude of the frequency components that wrap into the diffraction pattern window are weighted by the Airy function (Fourier transform of the aperture) and so their effects are minimal because of the rapid fall-off of the Airy function. On the other hand, when the illumination spot at the specimen plane is formed with a lens that has a finite aperture, the illumination function is not finite in real space; thus the illumination function has some aliasing. Gerchberg and Saxton (1972) also showed that sampling a chirp at the Nyquist interval where 99% of the signal is contained within the calculation window still results in failure of their algorithm and further commented that such behaviour was not characteristic of most signals.

2.8 Resolution in diffractive imaging

The Abbé's resolution (Born and Wolf, 1999) for an imaging system that illuminates the specimen with a collimated coherent beam is given by

$$\Delta r_a = 0.82\frac{\lambda}{\mathrm{NA}}, \tag{2.24a}$$

which corresponds to the smallest periodicity of a sinusoid that is correctly reproduced by the imaging system. The numerical aperture $\mathrm{NA} = \sin(\theta)$ in free space, θ is the one-half of the angle that the objective lens subtends on the specimen, λ is the wavelength of the illuminating radiation. An alternative way of quantifying the resolution of a coherent system is the Sparrow's criterion (DeViels and Thompson, 1998) given by

$$\Delta r_s = 0.47\frac{\lambda}{NA}. \tag{2.24b}$$

The Sparrow's criterion measures the distance Δr_s between two points that results in a uniform intensity gradient of sinc^2 functions centred at these points.

In the case of incoherent systems, the Rayleigh's criterion gives the resolution as

$$\Delta r_R = 0.61\frac{\lambda}{NA}. \tag{2.24c}$$

In all of these cases, the NA of the imaging system determines the resolution. In diffractive imaging the size of the detector corresponds to the NA of the imaging system. In ptychography, the iterative calculations produce a virtual detector with a larger NA (see Chapter 6).

2.9 Resolution improvement with ptychographic diffraction patterns

Rodenburg, McCallum and Nellist (1993) showed that the dense ptychographic dataset, where a focused spot illumination, in the STEM configuration, is moved to every position in the image, contains double resolution information about the image. A method for recovering double resolution information in the iterative formulation of ptychography (PIE) using a sparse ptychographic dataset, where a defocused illumination is moved by a substantial fraction of its effective size, was presented by Edo, Zhang and Rodenburg (2010). This work was further extended to structured illuminating wavefields by Maiden et al (2011) using the ePIE algorithm. These are important developments in the implementation of ptychography because they show that more information about the object can be extracted from the ptychographic dataset if the correct framework is employed. For example, information pertaining to 3D structure of a specimen is embedded in the highly redundant ptychographic dataset; and should be recoverable with the right algorithmic framework.

An analogous way of looking at the resolution enhancement of this method is in the form of aperture synthesis via tilt series reconstruction (Kirkland et al., 1995). In a tilt series reconstruction, the specimen is illuminated by a plane-wave at different incident angles and an image is recorded for each illumination tilt. This allows different parts of the specimen's Fourier transform to pass through the transfer function envelope of the objective. As a result this set of recorded images contains a range of the specimen's spatial frequencies greater than the limit imposed by the transfer function envelop on the corresponding bright field image. These images are then used to reconstruct a single image with a spatial frequency range equal to the union of the spatial frequencies of the recorded images. This procedure synthesises a larger aperture in the back focal plane of the objective lens. In a similar way, the STEM

probe is composed of a coherent superposition of several tilted plane-waves, where the maximum tilt angle (α) is determined by the size of the condenser aperture. This means that a set of bright field diffraction patterns contains a redundant expression of double resolution (2α) information about the specimen, as illustrated in Fig 2.10. Thus applying the PIE algorithm to a set of bright field diffraction patterns (Ronchigrams) that have a semi-angle (α) recovers an object with double resolution information.

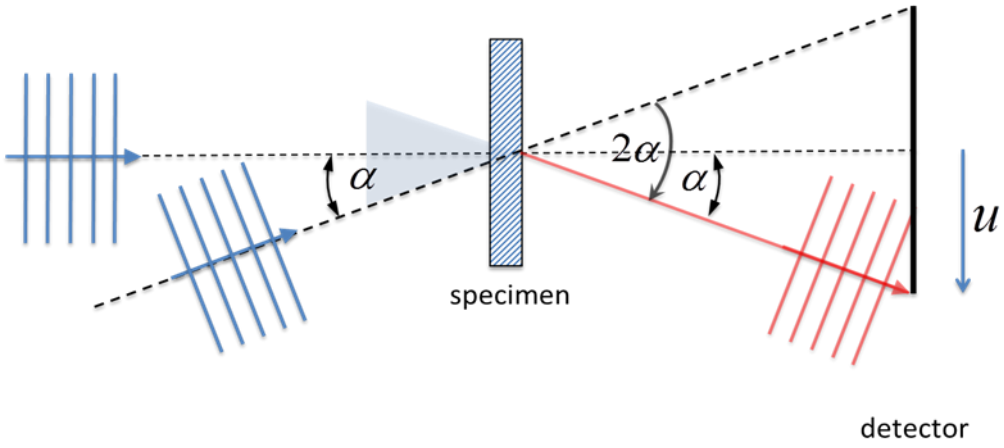


FIGURE 2.10 – illustration of double resolution information expression in the bright field of a STEM diffraction pattern.

Gerchberg (1974) and Papoulis (1975) implemented early iterative methods for extrapolating the frequency spectrum of finite 1D signals. These methods rely on the fact that a finite object has an extended frequency spectrum; this means that imposing the specimen plane boundary conditions at every stage of the iteration introduces spectra components outside the known frequency range. This is equivalent to performing a convolution of the estimated signal with a sinc function in the Fourier domain, in the case where the real space boundary constraint is a top hat function. Since the sinc function oscillates about zero for large frequencies, the convolution operation generates complex values at regions outside the

known spectra. However, the rapid fall-off of the sinc profile means that the extent of bandwidth extrapolation is very small especially in the presence of noise; so that the effective amount of bandwidth extrapolation is determined by the size of the central lobe of the sinc function.

2.10 Summary

In this chapter we reviewed the theory of imaging and wave propagation. The limitations of the lens in conventional imaging were discussed and diffractive imaging was shown to provide a means to liberate image formation from the limitations of the lens. The PIE and ePIE algorithms that implement diffractive imaging via ptychography were also presented. In the next chapter, the PIE update function is investigated in more detail using simulations and mathematical analysis in reciprocal space.

Chapter 3

3 The PIE update function

Section 3.1 investigates the update function of the PIE algorithm for two different model setups. These setups employ both plane-wave and curved-wave illumination. The following calculations are used to investigate the influence of the illumination curvature on the update function. This should also highlight the connection between information expression in the diffraction pattern and the quality of the solution provided by the update function. Section 3.2 reviews the interference framework of the phase problem based on the ptychographic principle. Although this framework traditionally deals with multiple diffraction patterns, its applicability to single diffraction pattern calculations is investigated in order to provide an explanation of the results from Section 3.1. This is important because it provides a general framework for discussing the recovered solutions from a wide range of illumination functions. Section 3.3 derives a general form of the update function using information mapping between the object and diffraction pattern. This approach provides a means for understanding the impact of the current update function on the spatial frequencies of the recovered object and thus determines whether it provides the optimum deconvolution

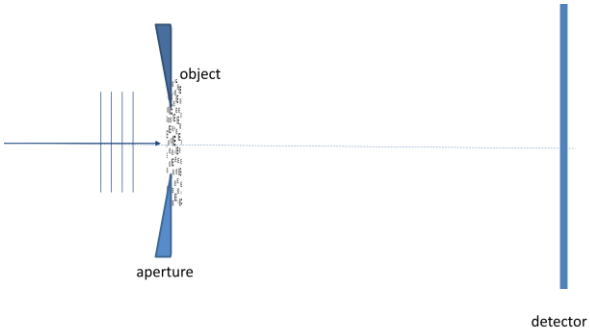
scheme. The derivations in Section 3.3 show the connection between the illumination intensity and the quality of the recovered object. Section 3.4 discusses the consequences of these connections in terms of coupling amongst spatial frequencies of the recovered object.

3.1 PIE update function with different experimental setups

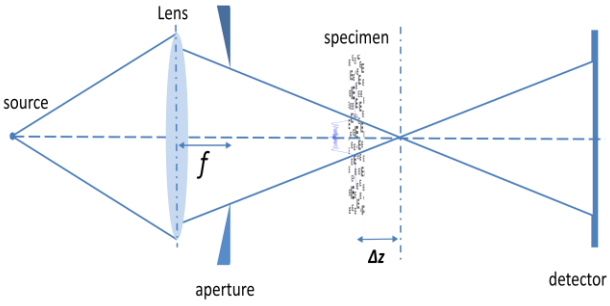
In the PIE algorithm, the update function has two main functions; it weights information from each diffraction pattern of the ptychographic dataset and provides a bridge that couples the contributions from several diffraction patterns into a single image. This produces a very robust framework that guarantees the unique recovery of complex objects in most experimental setups; however it does conceal the difficulties inherent in solving the phase problem in different setups. Unfortunately, these difficulties highlight aspects of the phase problem that should, in turn, influence the form of the update function. Consequently, the contribution from the ptychographic dataset is removed in this chapter for the first time and the update function is investigated for single diffraction pattern calculations.

In this section, the PIE update function is applied to two modelled configurations, employing plane-wave and curved-wave illuminations. Plane-wave illumination is predominantly used in X-ray diffractive imaging; this is because the illumination profile from a highly coherent synchrotron source takes the form of a plane-wave, that is to say it has a flat phase profile over the diffracting region of a finite specimen. In such an experiment, an isolated specimen is introduced into the beam in order to satisfy the sampling requirement at the detector plane, this isolation allows the application of the support constraint in iterative algorithms that solve for the specimen transmission function. This experimental configuration is modelled with the setup shown in Fig. 3.1a, where a pinhole is used to limit the size of an extended specimen.

The diffracting region of the specimen is thus limited to the size of the pinhole aperture. This model will be called the Type-I setup and is defined by the flat phase profile of the illumination over the diffracting region of the specimen. In electron experiments, the diffracting region of an extended object is selected with an illumination spot made by a converging lens. This experimental configuration is modelled by the Type-II setup, shown in Fig. 3.1b. The effective size of the illumination spot formed by the lens is determined by the defocus value Δz as illustrated in Fig. 3.1b (using a geometrical optics argument). It is worth noting that because of diffraction broadening the illumination generated by the lens in Fig. 3.1b does not have a finite boundary as with the case of the pinhole aperture in Fig. 3.1a, although most of the incident beam energy traverses the central spot defined by the defocus. Furthermore, using a defocused beam introduces curvature into the illumination at the specimen plane, where the local curvature over any part of the specimen decreases with increasing defocus. The curvature of the illumination over the specimen is the defining property of the Type-II setup. The diffraction patterns in both the Type-I and the Type-II setups are measured in the far-field.



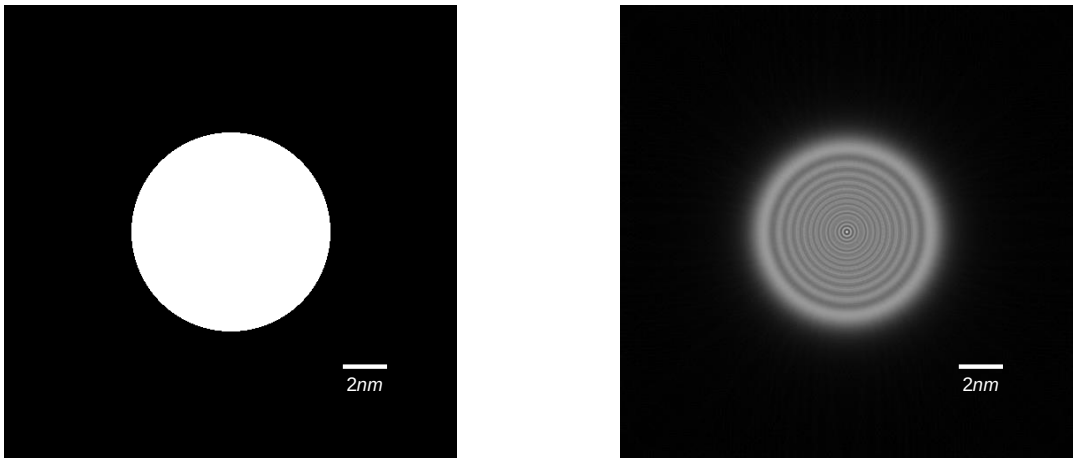
(a) The Type-I setup. Experimental configuration of diffractive imaging in which the finite extent of the diffracting region takes the shape of a circular pinhole aperture.



(b) The Type-II setup. Experimental configuration of diffractive imaging in which the finite extent of the diffracting region is facilitated by the defocused spot of a lens. This setup is approximately like Gabor holography when the specimen is weak.

FIGURE 3.1 - Schematic of the experimental setups modelled with the plane and curved illumination functions

Diffractive imaging setups that employ plane-wave and curved-wave illuminations were reported to have different convergence properties by Nugent et al (2005) and should thus serve as ideal candidates for investigating the properties of the update function. In this investigation, the size of the diffracting region for the Type-II setup is engineered to be approximately equal to the size of the aperture opening in the Type-I setup. This is done to test whether the size of the diffracting region alone gives a good handle on the convergence of the update function. This is important because the size of the diffracting region in these calculations directly relates to the sampling of the diffraction pattern. Thus, the results from this investigation can be used to investigate the connection between the sampling of the diffraction pattern and the convergence of the update function in different experimental setups.



(a) Magnitude of plane-wave illumination that is restricted by a disc-like mask.

(b) Magnitude of curved illumination at the specimen plane

FIGURE 3.2 - Structure of the illumination at the specimen plane for plane-wave and curve-wave diffractive imaging setups

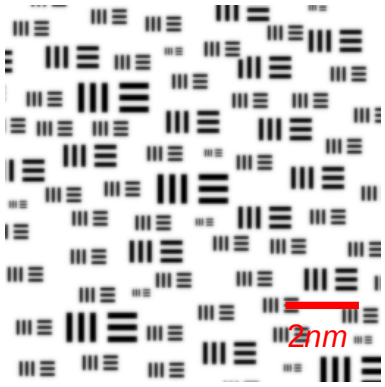
The magnitude of the illumination incident on the specimen for the Type-I and the Type-II setups are modelled and shown in Fig. 3.2a and Fig. 3.2b respectively. The illumination

shown in Fig. 3.2b is generated using parameters that model a 200 keV electron beam that is focused with an electromagnetic lens via a 15 mrad condenser aperture. As a result, the dimensions of these calculations correspond to those of electron diffractive imaging experiments. In this case, the Type-I setup models a finite size nano-particles experiment performed in the transmission electron microscope (TEM) mode; this setup is not prone to drift of the nano-particles within the illumination because of the shift invariance property of the resulting plane-wave diffraction patterns.

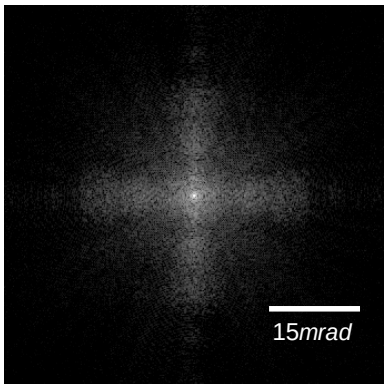
The total number of counts in the diffraction patterns from both model setups (see Fig. 3.3b and Fig. 3.3c) is normalised to have the same value (10^8). This is done because count values of 10^8 and above do not substantially affect the quality of the recovered object in region where the illumination has high intensity, for calculations that employ the test object shown in Fig 3.3a. This minimum count value varies with specimen type, i.e. specimens with a textured background require a lower number of counts in the diffraction pattern for the same exposure time because most of the incident radiation is diffracted from the zero order peak of the diffraction pattern and results in a higher signal to noise ratio (SNR). The total number of counts in these calculations is also used to incorporate Poisson noise into the diffraction patterns in order to account for the quantum noise inherent in the statistical arrival of electrons; the intensity values of the calculated diffraction patterns are used as mean values of the Poisson distribution, from which the noise is generated. For high count values, the Poisson distribution approaches the Gaussian distribution. This is a crucial input to these model calculations because the measured diffraction patterns in any experiment do not have the exact intensity values that are calculated with the FFT algorithm.

Employing one diffraction pattern in these iterative calculations removes the benefit of being able to routinely recover a complex object using the update function with probe overlap. Thus, the object used in this investigation is a resolution target with a positive transmission

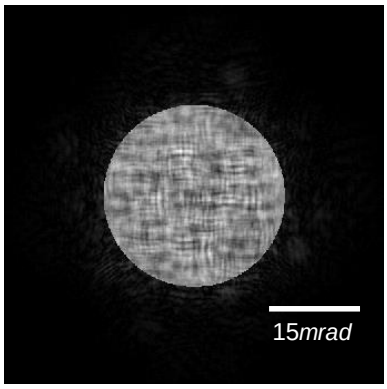
magnitude, shown in Fig. 3.3a. It is worth noting this absorbing object is more appropriate at X-ray and visible light wavelengths, where the dimensions of the calculations need to be rescaled. In electron experiments, the object would usually be a phase object (see Section 4.1). However, this image type suffices to investigate the convergence properties of the update function for the Type-I and the Type-II setups. The background is transparent with transmission coefficient of unity and the bars of the object have transmission values that represent absorption strength. The sharpness of the diffracting bar edges in the object was removed in order to minimise the effect of aliasing in the forward calculation of the diffraction pattern during iterative calculations because the detector plane is multiplied by the atomic scattering factor.



(a) Magnitude of transmission function of test object. White corresponds to maximum transmission value of 1 and black correspond to minimum transmission value of 0.1.



(b) Diffraction pattern from an object in the Type-I setup.



(c) Diffraction pattern from an object in the Type-II setup.

FIGURE 3.3 – The transmission function of the test object and the corresponding diffraction patterns for plane and curved wave illumination

Fig. 3.3b and Fig. 3.3c show the diffraction patterns from the Type-I and the Type-II setups respectively. The diffraction patterns are calculated by taking the Fast Fourier Transform (FFT) of the exit wave at the specimen plane. The detectors in both setups comprise 512x512 pixels and span the same angular range of 62.7 mrad. This gives a sampling pitch of 0.04 nm

at the specimen plane. The forward calculation in the generation of the test data is performed with a sampling pitch of 0.01nm and the central part of the resulting diffraction pattern, spanning 62.7 mrad was extracted. This was done to ensure that the edges of the diffraction patterns result from linear convolution rather than cyclic convolution, which results from the discrete implementation of the FFT algorithm. It is worth noting that the four-fold oversampling of the real space coordinate was only used to generate the test data and does not correspond to the real space sampling pitch in subsequent iterative calculations. The update step of the algorithm that employs the original update function is given by

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + P^*(\boldsymbol{r})\{\pi_M\{\psi_n(\boldsymbol{r})\} - \psi_n(\boldsymbol{r})\}. \tag{3.1a}$$

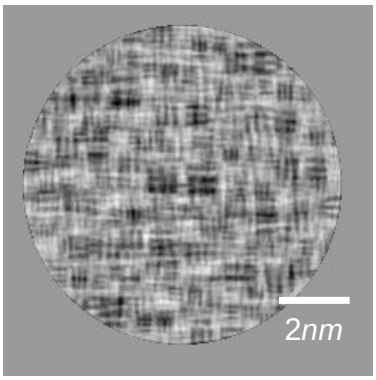
The algorithm is then modified to take advantage of the non-complex nature of the test object, which improves the quality of the reconstructions (see Fig. 3.4). This was implemented by taking the real part of the exit wave difference. The modified algorithm with normalised illumination magnitude is given by

$$O_{n+1}(\boldsymbol{r}) = O_n(\boldsymbol{r}) + \text{real}(P^*(\boldsymbol{r})\{\pi_M\{\psi_n(\boldsymbol{r})\} - \psi_n(\boldsymbol{r})\}) \tag{3.1b}$$

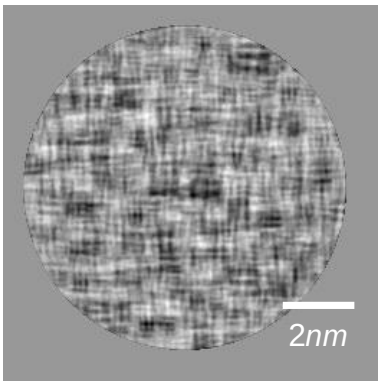
3.1.1 Results

The update schemes in Eq. (3.1) were applied to the diffraction patterns of Fig. 3.3b and Fig. 3.3c. These algorithms were run for 10000 iterations, which corresponds to the number of times the update function was applied at the specimen plane. Fig. 3.4a and Fig. 3.4b show the recovered objects from calculations that applied the original and modified versions of the algorithm to the Type-I diffraction pattern. Fig. 3.4c and Fig. 3.4d show the results for the

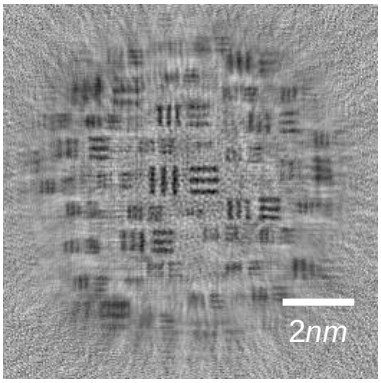
same calculations using the Type-II diffraction pattern. The evolution of the error metric over 500 iterations is shown in Fig. 3.5.



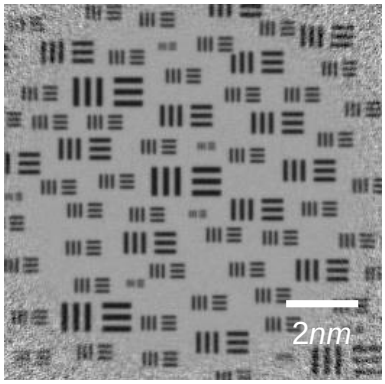
(a) Recovered object from the Type-I experiment using original update function



(b) Recovered object from the Type-I experiment using update function with non-complex constraint

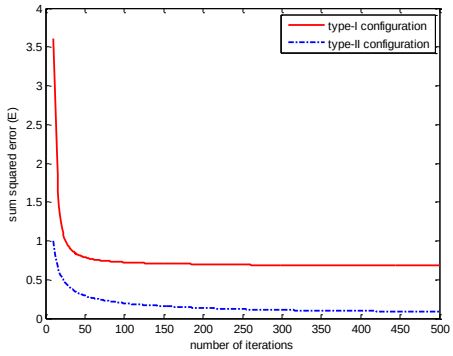


(c) Recovered object from the Type-II experiment using original update function

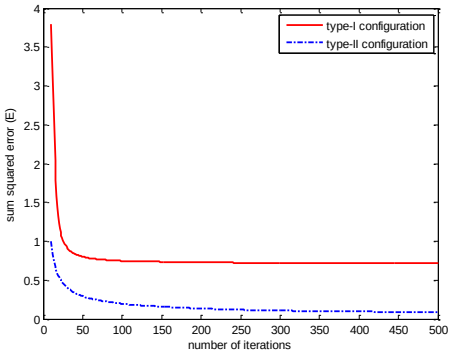


(d) Recovered object from the Type-II experiment using update function with non-complex constraint

FIGURE 3.4 – Recovered transmission function of the model object using diffraction patterns from the Type-I and the Type-II setups.



(a) Calculations that employ the original update function



(b) Calculations that employ the update function with non-complex constraint

FIGURE 3.5 - Trend of mean squared error with number of iterations for the Type-I and the Type-II experiments.

3.1.2 Discussion

The quality of the recovered objects in Fig. 3.4 shows that the algorithms converge to a better estimate of the test object for the case of the Type-II setup than they do in the case of the Type-I setup. This result is interesting because the diffracting regions have the same size in both the Type-I and the Type-II calculations. In fact, the effective size of the diffracting region is larger from the Type-II setup since the illumination does not have a finite boundary. This suggests that the size of the diffracting region does not explain the quality of the recovered object in calculations that employ different forms of illumination. One explanation for the quality of the recovered objects in Figs. 3.4c and 3.4d may be the close correspondence between the central bright field region of the diffraction pattern with curved-illumination and the Gabor hologram (Gabor, 1948). In the Type-II calculations, the estimate of the object after the first iteration gets very close to the original object because lower spatial frequency information (up to the aperture size) of the object is encoded in intensity variations of the hologram. Fig. 3.4d shows that the recovered object improves significantly with the application of the non-complex constraint in the Type-II calculations however similar improvements do not occur in the Type-I calculations (see Fig. 3.4b). This means that the expression of the object information in the Type-II setup provides a better conditioned phase problem for the update function.

The algorithms defined by Eq. 3.1 uses a feedback (β) of unity (see Eq. (2.18)). As a result, the algorithm reduces to the Error Reduction method for the Type-I calculations. This is because the weighting function is a binary mask that has a normalised value of unity inside the aperture and a value of zero outside. Consequently, the update function enforces a boundary constraint (just like ER) since the illumination is constant inside the aperture. The ER method was shown to have stagnation problems for the Type-I setup (Fienup, 1982) when

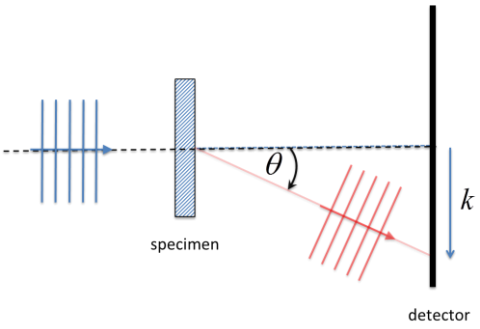
applied to a single diffraction pattern. A better estimate of the object in the Type-I setup can be obtained with the HIO method. However the HIO method is only applicable to simple illumination functions with boundaries defined by an aperture, as in the case of the Type-I setup. Thus it does not provide a general solution that can be applied to the case of structured illumination.

The error plots in Figs. 3.5a and 3.5b have similar error magnitude suggesting that the error metric does not accurately quantify the quality of the recovered objects. These results mirror the analysis by Fienup (1982) and Guizar-Sicairos and Fienup (2008) that show these algorithms minimise the error in the diffraction pattern, however this does not guarantee the correct solution from the minimisation routine.

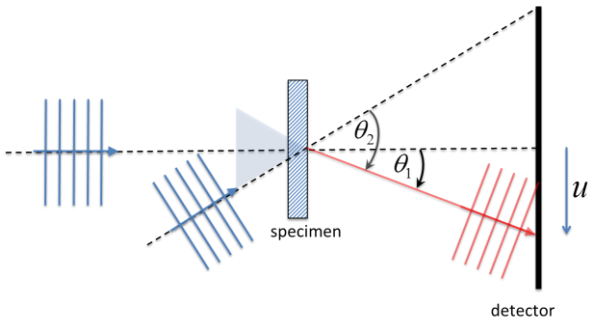
The above explanations draw from the insights of holography and from iterative phase retrieval techniques with plane-wave illumination. This suggests that the behaviour of the update function and the quality of the recovered object for experimental setups that employ other forms of illumination may only rely on insights particular to that setup. However, this results in a multitude of explanations that obscure the fundamental reasons for the solubility of the phase problem in different experimental setups. To provide a coherent explanation for the results in a general framework, diffractive imaging experiments need to be viewed as concurrent plane-wave interference experiments, where the strength of each plane-wave is determined by the magnitude of the Fourier component that corresponds to its $\boldsymbol{k}$-vector. The next section introduces the relevant mathematics required for viewing diffractive imaging experiments in a framework where interference serves as the solubility metric for the phase problem. This will provide a unified way of characterising the behaviour of the update function and the quality of the recovered object.

3.2 Mathematical Concepts and Nomenclature

This section introduces some new and essential mathematical concepts that are required for an effective discussion on the solution (recovered object) provided by the update function. This section adopts a nomenclature in which the symbols that represent functions in real space, Fourier space and the detector plane (in the far-field) are the same except for the explicit dependence on different coordinates. The real space domain is spanned by the coordinate $\boldsymbol{r} =: (x, y)$, the Fourier transform of the object is spanned by the coordinate $\boldsymbol{k} =: (k_x, k_y)$ and the detector plane is spanned by the coordinate $\boldsymbol{u} =: (u, v)$. The Fourier transform coordinate, $\boldsymbol{k}$, points in the same direction as the detector coordinate, $\boldsymbol{u}$, but is used as a dummy variable in the convolution integral that follows.



(a) Specimen illuminated by plane-wave illumination. Unique mapping of spatial frequencies to angle ($k = \theta/\lambda$). The detector plane is accurately mapped by the $\boldsymbol{k}$-basis.



(b) Specimen illuminated by curved-wave illumination. Different spatial frequencies $k_1 = \theta_1/\lambda$ and $k_2 = \theta_2/\lambda$ are mapped to one point in the detector thus the detector plane is not uniquely mapped by the $\boldsymbol{k}$-coordinate but is instead spanned by the $\boldsymbol{u}$-coordinate.

FIGURE 3.6 – illustration of spatial frequencies mapping for plane-wave and curved wave illumination

The distinction between the Fourier domain and the detector plane in a far-field diffractive imaging setup is necessary in order to incorporate the effects of structured illumination. In this framework, the directions of $\boldsymbol{k}$-vectors capture scattering events when the specimen is

illuminated by a plane-wave (see Fig. 3.6a) however $\boldsymbol{u}$-vectors capture scattering events when the specimen is illuminated by a structured wave. Since a structured wave has several plane-wave components, $\boldsymbol{u}$-vectors account for scattering from different plane-wave directions into a given detector pixel as illustrated in Fig. 3.6b. The Fourier transform is used to calculate the diffraction pattern in far-field diffractive imaging experiments. The Fourier transform relationship between the specimen exit-wave and the complex wave at the detector plane is given by

$$\psi(\boldsymbol{u}) = \int \psi(\boldsymbol{r}) \exp(i\boldsymbol{u} \cdot \boldsymbol{r})\, d\boldsymbol{r}. \tag{3.2}$$

In practical experiments, the detector can only measure the intensities of the complex coefficients in the expansion; this results in loss of the phase information.

3.2.1 Solution to the phase problem via interference

The specimen function can be represented with the real space basis function $\delta(\boldsymbol{X} - \boldsymbol{r})$ or the Fourier basis $\exp(i\boldsymbol{k} \cdot \boldsymbol{X})$, where δ is the Dirac delta function. This is equivalent to viewing the specimen function as an image distribution in real space or an angular distribution in reciprocal space, such that

$$\begin{aligned} O(\boldsymbol{r}) &= \int O(\boldsymbol{X})\delta(\boldsymbol{X} - \boldsymbol{r})d\boldsymbol{X}, \\ O(\boldsymbol{k}) &= \int O(\boldsymbol{X}) \exp(i\boldsymbol{k} \cdot \boldsymbol{X})\, d\boldsymbol{X}. \end{aligned} \tag{3.3}$$

This enables the phase problem to be in two conjugate domains (real and Fourier space) and allows different aspects of the problem to be solved in the domain in which it takes a simple form. In diffractive imaging, the region of the specimen $O(\boldsymbol{r})$ that contributes to the diffraction pattern is selected with an illumination function $P(\boldsymbol{r})$. The exit wave from a

sufficiently thin specimen is given by the product of the specimen transmission function and the illumination function, $\psi(\boldsymbol{r}) = O(\boldsymbol{r}) \cdot P(\boldsymbol{r})$, where the term 'thin' has been used to exclude 3D geometric interference. The specimen can also introduce strong phase changes into the illuminating beam (i.e. scatters strongly). The complex wave distribution in the far-field is given by the Fourier transform of the exit wave $\psi(\boldsymbol{u}) = \mathcal{F}\psi(\boldsymbol{r})$, thus the convolution theorem allows one to also express the wavefield at the detector in terms of a convolution integral. These steps are outlined in Eq. (3.4), where $I_m(\boldsymbol{u})$ represents the intensity values measured by detector.

$$\psi(\boldsymbol{r}) = O(\boldsymbol{r}) \cdot P(\boldsymbol{r}) \tag{3.4a}$$

$$\psi(\boldsymbol{u}) = \mathcal{F}\psi(\boldsymbol{r}) \tag{3.4b}$$

$$\psi(\boldsymbol{u}) = O(\boldsymbol{k}) * P(\boldsymbol{k}) = \int O(\boldsymbol{k}) P(\boldsymbol{u} - \boldsymbol{k}) d\boldsymbol{k} \tag{3.4c}$$

$$I_M(\boldsymbol{u}) = |\psi(\boldsymbol{u})|^2 = \left| \iint O(\boldsymbol{k}) P(\boldsymbol{u} - \boldsymbol{k}) d\boldsymbol{k} \right|^2 \tag{3.4d}$$

A different way to arrive at Eq. (3.4d) is to decompose the incident beam into a set of plane-waves, where each plane-wave illuminates the specimen at an angle proportional to its $\boldsymbol{k}$-vector. The complex wavefield at the detector plane resulting from each plane-wave component is laterally shifted by an amount proportional to the $\boldsymbol{k}$-vector of the incident plane-wave (using the Fourier shift theorem). Furthermore, the distribution of all the complex wavefields is identical to the Fourier transform of the specimen function except for lateral displacements of these wavefields at the detector plane. The recorded diffraction pattern is given by the intensities of the sum of all the complex wavefields. This corresponds to an addition of shifted versions of the Fourier transform of the specimen (i.e. convolution operation). This means that the measured intensity value at a given pixel ($\boldsymbol{u}$) of the detector results from adding different spatial frequencies of the object. This results in interference amongst various spatial frequencies of the specimen and is expressed mathematically by Eq.

(3.4d), where the contribution from each spatial frequency is given by the $\boldsymbol{k}$th component of the convolution kernel, $P(\boldsymbol{u}-\boldsymbol{k})$.

The preceding argument suggests that the solubility of the phase problem (in far-field geometry) may be viewed in a framework where convolution in Fourier space provides a metric for solubility. Consider the case where the diffracting region of the specimen has dimensions that correspond to the sampling pitch of the detector (i.e. Nyquist sampling) and is illuminated by a plane-wave $P_{\boldsymbol{a}}(\boldsymbol{r})$; In such a situation the exit wave immediately downstream of the specimen is given by $\psi(\boldsymbol{r}) = O(\boldsymbol{r})$, where the illumination function is unity i.e. $(P_{\boldsymbol{a}}(\boldsymbol{r}) = 1)$. By applying Eq. (3.4d) to generate the diffraction pattern, the spatial frequency or $\boldsymbol{k}$-space representation of the specimen $O(\boldsymbol{k})$ is convolved with the Dirac delta function as shown in Eq. (3.5).

$$I_M(\boldsymbol{u}) = |O(\boldsymbol{u})|^2 = \left| \int O(\boldsymbol{k})\delta(\boldsymbol{u}-\boldsymbol{k})d\boldsymbol{k} \right|^2 \tag{3.5}$$

Eq. (3.5) reveals that there is no room for interference amongst the object spatial frequencies due to the sifting property of the delta function. The phase problem is therefore insoluble because phase information is not coded in the intensities of the diffraction pattern. The diffraction pattern from Eq. (3.5) gives an accurate representation of the power spectrum of the specimen but contains no information about the relative phase difference amongst different spatial frequencies.

3.2.2 Diffractive imaging with plane-wave and curved-wave illuminations

It is important to note that the plane-wave is an idealisation since the wavefront from a source such as a laser is usually finite in extent. In diffractive imaging, the wavefield is considered

to be a plane-wave if it is covers the region of interest with a reasonably flat phase profile. In the Type-I setup, the size of the diffracting region is made small enough to allow the central lobe of the Airy function (which corresponds to the Fourier transform of the circular aperture) to span more than one pixel in the detector plane. This means that size limitation of the diffracting region of an extended object provides room for interference amongst different spatial frequencies of the object. In the case of a finite object, the propagated wave at the detector plane automatically results from an interference, which is governed by the Fourier transform of the particle's shape. Hegerl and Hoppe (1972) suggested the connection between the finite size of the diffracting region and the phenomenon of interference, where they showed that interference provides the means for solving the phase problem. In the X-ray field size limitation of the diffracting region, which results in a spread of the central lobe of the Airy function over several pixels of the detector is called oversampling.

By making the diffracting region smaller, the interference strength amongst the object's spatial frequencies increases because the central lobe of the Airy function spreads over a larger area of the detector. An alternative way of increasing the strength of interference is to introduce structure into the illumination. Structured illumination enables a spread of the illumination intensity over a larger area of the detector in the absence of any specimen. Curved illumination is one of the simplest forms of structured illumination and can be easily made with a lens whose quality need not be as good as that required for normal imaging. This is because the lens is used only to form an illumination spot on the specimen, in order to limits the size of the diffracting region for an extended specimen and the aberrations of the lens do not affect the quality of the recovered object. It also facilitates interference conditions that do not have the rapid fall-off profile of the Airy function, without the need to reduce the size of the diffracting region. This means that structured illumination can provide the means for large field of view (FoV) diffractive imaging, possibly at sub-Nyquist sampling

of the diffraction pattern. This is investigated in Sections 4.5 for the case of single diffraction pattern calculations; an extension to the case of multiple diffraction patterns is also investigated Section 6.7.

The illumination in the Type-II setup has a large range of plane-waves with uniform strength. This allows maximum interference amongst the spatial frequencies of the object when compared to the Type-I setup where the interference strength falls-off with the magnitude of the Airy function. In addition, *a priori* knowledge of the phase structure of the illumination enables the decoupling of the contributions from each spatial frequency and is further explored in Section 3.3. From an interference point of view, the strength of interfering components of the object's spatial frequency provides a possible explanation for why the calculations of the Type-II setup gives a better quality reconstruction than the case of the Type-I setup.

3.2.3 Diffractive imaging and sampling

This section considers the issue of sampling in diffractive imaging. The Shannon sampling theorem states that in order to capture the entire content of a bandlimited signal, the signal should be sampled at a rate that correspond to twice the bandwidth of the signal. This ensures that the discrete representation of the signal does not suffer from the effect of aliasing. In relation to image processing, this means that the window in the specimen plane forms an independent unit cell that does not overlap with adjacent cells (that results from the periodic framework of the discrete Fourier transform or DFT). The symmetry between the Fourier transform and its inverse implies that this window can represent either real space or reciprocal space. Since the image cannot be bandlimited in both domains, there is always

some element of aliasing in diffractive imaging calculations. The effect of aliasing is minimised by judicious sampling of the diffraction pattern. In forward calculations where one computes the FFT of an image that is completely contained within the real space window, the diffraction pattern is aliased because adjacent windows in the detector plane overlap. The strength of the aliased components at atomic wavelengths is, however, small because the diffracting region consists of finite size of atoms and thus the diffraction window is multiplied by the atomic scattering factor. Sayre (1952) suggested that if one assumes a Gaussian model for atoms, then the Fourier spectrum is multiplied by a Gaussian profile so that the effect of aliasing is minimal in the detector plane. Furthermore, enforcing the correct magnitude distribution in the detector plane ensures that the magnitude profile within the diffraction window is correct; although this does not guarantee that the associated phase distribution is also correct. This is because the finite window in diffraction space results in a convolution of real-space sampling points with a 2D sinc function as opposed to the Gaussian. This effect can only be minimised by increasing the angular size of the detector.

So the question arises, is oversampling the key requirement for the solubility of the phase problem? It was shown earlier that oversampling could be viewed as a way of creating interference that encodes phase information. In order to connect solubility via interference and the Shannon sampling criterion, we need to think about the representation of the information content of a signal. Firstly, it is important to note that the Shannon sampling criterion assumes that the signal of interest is represented in an orthogonal basis so that the sampled points are independent. In the case of diffractive imaging, diffraction patterns are measured in non-orthogonal bases so that the measured components interfere. Furthermore, we can factorise the exit wave into two components, the object and the illumination, for which the illumination function is sometimes known *a priori*. As a result, the diffraction pattern provides a redundant representation of the specimen. Furthermore, knowledge of the

illumination function provides the structure for the redundancy. This suggests that a possibility is to sample the diffraction pattern at the Nyquist rate of the underlying complex distribution and get a reasonable reconstruction from the undersampled diffraction pattern. The diffraction pattern is undersampled because its bandwidth is larger than the bandwidth of the complex wavefield from which it is derived. The calculation still needs to be performed in a window that satisfies Shannon sampling in order to avoid aliasing. Performing the calculation with a Nyquist window requires upsampling of the diffraction pattern. When the diffraction pattern intensity values are not interpolated, parts of the detector plane lack intensity information that needs to be solved for. Undersampling the diffraction pattern facilitates large FoV experiments and is further explored in Chapter 6, where the ptychographic dataset provides the necessary diversity to constrain the calculations.

3.3 Derivation of the PIE iterative deconvolution kernel (IDK)

In this section the iterative deconvolution process in Fourier space is explored, where the machinery of the PIE update function takes a less compact form that lends itself to analysis. This work is timely because the update function provides the framework for bandwidth extrapolation of the recovered object in resolution-improvement ptychographic calculations. It is therefore paramount to establish that the current weighting scheme of the update function provides the right framework and a robust platform for bandwidth extrapolation calculations, which are further discussed in Chapter 6. Consequently, this section explores the structure of the update function and provides an alternative description for the optimum weighting ($l = 2$) of the update function used in the PIE algorithm. The analysis presented in this section lays

the foundation for further discussions on the behaviour of the update function that are presented in Chapter 4.

The PIE algorithm retrieves the specimen function using an iterative deconvolution process that is facilitated by an update function. The update function constrains the deconvolution process by weighting the feedback errors or corrections that are added to the object during iterative calculations. This weighting scheme can be viewed in real space where it has the effect of preventing division by zeros during the deconvolution process. Since this analysis is performed in reciprocal space, we need a representation of the update function in reciprocal space. The update function has a reciprocal space representation called the iterative deconvolution kernel (IDK) and is equal to the Fourier transform of the update function i.e. $IDK[\boldsymbol{u}] = \mathcal{F}\{U[\boldsymbol{r}]\}$.

During iterative calculations, the PIE algorithm processes the input (guess of object function) and outputs a new estimate of the object. The algorithm adds corrections calculated from the measured diffraction pattern, in order to ensure that the iterative calculations move the running estimate of the object towards the correct solution. These corrections are codified by what might be called the feedback error function $\varepsilon_n[\boldsymbol{u}]$. This error function needs to be transformed from the detector plane ($\boldsymbol{u}$-space) to the Fourier domain of the estimated object ($\boldsymbol{k}$-space). Since the diffraction pattern is calculated from the convolution operation in Eq. (3.4c), the forward calculation suggests that the reverse process requires the convolution inverse. Implementation of the convolution inverse can be accomplished with a division in the Fourier domain of the detector (i.e. real-space). However the diffracting region in real-space is usually limited to ensure adequate sampling of the diffraction intensities by the detector. This means the illumination function, which defines the boundary of the diffracting region, contains zeros. As a result, the deconvolution operation contains divisions by zero

and is, therefore, ill-conditioned. Moreover, the diffraction pattern contains no phase information; so direct deconvolution does not provide the required solution.

Iterative deconvolution using the PIE algorithm resolves this problem by assigning phase values to the diffraction pattern whilst solving for the object. In the PIE algorithm, the running estimate of the object function is updated using feedback errors from the diffraction pattern; where the error function, $|\varepsilon_n[\boldsymbol{u}]|^2$, measures the departure of the calculated diffraction pattern from the measured diffraction pattern in the n^{th} iteration. The calculated diffraction pattern is given by the intensity of the FFT of the exit wave.

A general form of the update function is given by

$$O_{n+1}[\boldsymbol{k}] = O_n[\boldsymbol{k}] + \varepsilon_n[\boldsymbol{k}], \tag{3.6a}$$

$$O_{n+1}[\boldsymbol{k}] = O_n[\boldsymbol{k}] + \sum IDK[\boldsymbol{k} - \boldsymbol{u}] \cdot \varepsilon_n[\boldsymbol{u}], \tag{3.6b}$$

where the feedback error function $\varepsilon_n[\boldsymbol{k}]$ has been appropriately transformed from the detector plane to the Fourier domain of the object using the iterative deconvolution kernel, $IDK[\boldsymbol{u}] = \mathcal{F}\{U[\boldsymbol{r}]\}$. Here square brackets are used to emphasize the fact that the analyses in this thesis only consider the discrete case. The rest of this section investigates the form of the IDK that is required to correctly transform the feedback error ($\varepsilon_n[\boldsymbol{u}]$) from the diffraction pattern to the feedback error ($\varepsilon_n[\boldsymbol{k}]$) in the object.

Eq. (3.6b) shows that the IDK provides the means for mapping errors from the diffraction pattern to errors in the object. In order to derive the form of the IDK, we first need to consider the influence of the object on the detector wavefield in order to get a handle on the flow of information from the object to the diffraction pattern. Here 'flow of information' means identifying the set/group of the object's spatial frequencies that contribute to a given diffraction pattern intensity. This information flow can then be backtracked from the detector

plane to the object. This allows the reverse calculation (deconvolution) to be constrained in such a way that the diffraction pattern intensity values only influence those spatial frequencies from which they were derived. This approach is essential because it provides a generic means of updating the object components (spatial frequencies) based on information expressed in the diffraction pattern.

Consider the case where the object function is changed by a small amount (ΔO), the corresponding change in the complex wave at the detector plane is given by

$$\Delta\psi = \sum \frac{\partial\psi}{\partial O}\Delta O,$$

$$\Delta\psi[\boldsymbol{u}] = \sum_{S(\boldsymbol{u};\boldsymbol{k})} \frac{\partial\psi[\boldsymbol{u}]}{\partial O[\boldsymbol{k}]}\Delta O[\boldsymbol{k}], \tag{3.7}$$

where $S(\boldsymbol{u};\boldsymbol{k})$ defines the components of the object function ($O[\boldsymbol{k}]$) that influence the wave distribution (ψ) at the detector pixel ($\boldsymbol{u}$). Fig. 3.7 shows a pictorial representation of the transformation indicated in Eq. (3.7) for the Type-II setup, where the region enclosed by the white circle on the FFT of the object corresponds to $S(\boldsymbol{u};\boldsymbol{k})$. The disc shape of $S(\boldsymbol{u};\boldsymbol{k})$ in the Type-II setup is inferred from Eq. (3.8).

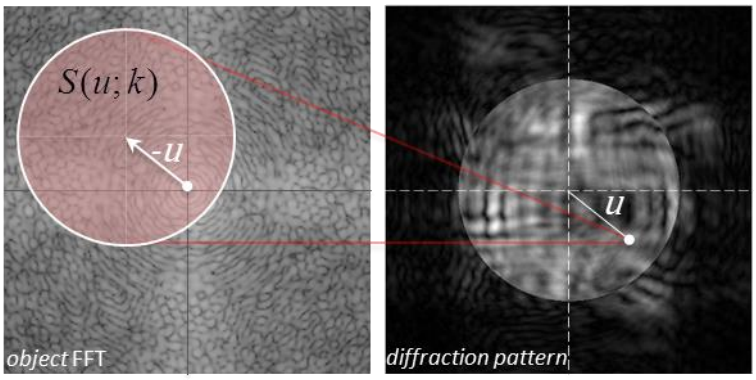


FIGURE 3.7 - illustration of spatial frequencies of the object that contributes to a pixel in the diffraction pattern. The circular region is centre on $k = -u$ because the convolution operation flips the Fourier transform of the object so that integration is done over a region centred on $u = -k$, but is illustrated here as $k = -u$. This is valid for a circular region.

Since the forward calculation corresponds to a convolution, the change of the wave distribution at the detector plane $\Delta\psi$ can be calculated directly from Eq. (3.4c). This is

$$\begin{aligned}\Delta\psi[\boldsymbol{u}] &= \psi_{\Delta O}[\boldsymbol{u}] - \psi[\boldsymbol{u}], \\ &= \Big(\sum P[\boldsymbol{u}-\boldsymbol{k}](O[\boldsymbol{k}] + \Delta O[\boldsymbol{k}])\Big) - \Big(\sum P[\boldsymbol{u}-\boldsymbol{k}]O[\boldsymbol{k}]\Big), \\ &= \sum_{S(\boldsymbol{u};\boldsymbol{k})} P[\boldsymbol{u}-\boldsymbol{k}]\Delta O[\boldsymbol{k}].\end{aligned} \tag{3.8}$$

Equating the terms of Eq. (3.7) and Eq. (3.8), gives

$$\frac{\partial\psi[\boldsymbol{u}]}{\partial O[\boldsymbol{k}]} = P[\boldsymbol{u}-\boldsymbol{k}] \tag{3.9}$$

Eq. (3.9) captures the flow of information from the object to the diffraction pattern, because it identifies the correct set of object spatial frequencies that contributes to a given diffraction pattern intensity value; i.e. $S(\boldsymbol{u};\boldsymbol{k}) = (|P[\boldsymbol{u}-\boldsymbol{k}]| > 0)$. This means that the diffraction pattern intensity value ($I_M[\boldsymbol{u}]$) has a contribution from a set of spatial frequencies defined by $S(\boldsymbol{u};\boldsymbol{k})$. In the Type-II configuration, the magnitude of the convolution kernel is constant and can be normalised so that $|P|_{max} = 1$. Thus, $S(\boldsymbol{u};\boldsymbol{k})$ corresponds to the disc (encircled patch) shown in Fig. 3.7. The shape of $S(\boldsymbol{u};\boldsymbol{k})$ is different for different illumination functions, also the magnitude profile of the convolution kernel is not constant for other types of illumination. An extension of the patch idea to other structured illumination indicates that the general case should have $S(\boldsymbol{u};\boldsymbol{k}) = |P[\boldsymbol{u}-\boldsymbol{k}]|$, where the contributions to a given diffraction pattern intensity value are weighted by the magnitude of the convolution kernel. It is important to note that other power functions, i.e. $S(\boldsymbol{u};\boldsymbol{k}) = |P[\boldsymbol{u}-\boldsymbol{k}]|^n$, with $n \neq 1$, do not correspond to the correct solution, because the convolution integral encodes the interference information using a linear power of the object spatial frequencies, i.e. $n = 1$.

The remaining analysis is undertaken with the convolution kernel of the Type-II setup and the result applies to a general illumination function.

In the present context, the transformation (convolution) equation depends only on the difference between the $\boldsymbol{u}$ and $\boldsymbol{k}$ coordinates, therefore the patch is best described as $S(\boldsymbol{u};\boldsymbol{k}) = S(\boldsymbol{u}-\boldsymbol{k})$. Furthermore, since we only consider a single diffraction intensity at a time (i.e. the value of $\boldsymbol{u}$ is fixed), we can drop the $\boldsymbol{u}$ dependence so that $S(\boldsymbol{u}-\boldsymbol{k})$ goes to $S(-\boldsymbol{k})$, which will be called $S(\boldsymbol{k})$ because of the symmetry of the illumination FT in the Type-II setup. The negative sign indicates that the transformation between the diffraction pattern and the object is a convolution instead of a correlation, so that the coordinates are flipped.

3.3.1 IDK magnitude calculation

To investigate the inverse case, consider perturbations of the object that are restricted to a single spatial frequency (i.e. $\varepsilon[\boldsymbol{k}] = \Delta O[\boldsymbol{k}] = \delta[\boldsymbol{k}'-\boldsymbol{k}]$); Eq. (3.4c) indicates that the resulting effect in the detector plane is proportional to the Fourier transform of the illumination function centred on that frequency i.e. $\Delta\psi(\boldsymbol{u}) = \int \delta(\boldsymbol{k}'-\boldsymbol{k})P(\boldsymbol{u}-\boldsymbol{k}')d\boldsymbol{k}' = P(\boldsymbol{u}-\boldsymbol{k})$. This conclusion can also be drawn from Eq. (3.9) by keeping $\boldsymbol{k}$ fixed and allowing $\boldsymbol{u}$ to vary. This gives the set of $\boldsymbol{u}$-components in the detector plane required for the inverse operation via information backtracking. These $\boldsymbol{u}$-components define the inverse domain ($S^{-1}(\boldsymbol{k};\boldsymbol{u})$) which is now referred to as $S^{-1}(\boldsymbol{u})$ in order to emphasize that all contributions come from the detector domain. Put differently, in the Type-II setup, information from the object component ($O[\boldsymbol{k}]$) can only contribute to a finite set of diffraction intensities defined by $S^{-1}(\boldsymbol{u}) = (|P[\boldsymbol{k}-\boldsymbol{u}]|) > 0$ and, conversely, only feedback errors from this region should contribute to the running estimate of the object component

($O_n[\boldsymbol{k}]$) during iterative calculations. This piece of information defines the magnitude of the iterative deconvolution kernel and is given by

$$|IDK| = |P[\boldsymbol{k} - \boldsymbol{u}]| \tag{3.10}$$

where the magnitude of the IDK is again a linear function of the kernel's magnitude, because the object spatial frequencies contribute linearly towards the interference pattern in the detector plane

3.3.2 IDK phase calculation

The most efficacious way to infer the phase components of the IDK is to find a phase configuration in the detector plane that transforms the contribution from the object (which is $\Delta O[\boldsymbol{k}]P[\boldsymbol{u} - \boldsymbol{k}]$) back to the object ($O[\boldsymbol{k}]$) with a minimal change in the relative phase values of the object components. This is important because we need to preserve the relative phase difference amongst all the object spatial frequencies. One strategy would be to remove the phase of the illumination from the diffraction pattern contribution ($\Delta O[\boldsymbol{k}]P[\boldsymbol{u} - \boldsymbol{k}]$). The function that provides the required phase configuration over the diffraction pattern is the conjugate of the convolution kernel, because $\Delta O[\boldsymbol{k}] \propto \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{u} - \boldsymbol{k}](\Delta O[\boldsymbol{k}]P[\boldsymbol{u} - \boldsymbol{k}])$. This suggests that the IDK could simply be the conjugate of the convolution kernel. To confirm this hypothesis, the rest of this section explores an alternative route to find the required phase configuration.

To find the IDK phase configurations the convolution inverse is required, this is derived by finding the change of the object function due to changes in the complex wave at the detector plane. This is given by

$$\Delta O = \sum \frac{\partial O}{\partial \psi} \Delta \psi,$$
$$\Delta O[\boldsymbol{k}] = \sum_{S^{-1}(\boldsymbol{u})} \frac{\partial O[\boldsymbol{k}]}{\partial \psi[\boldsymbol{u}]} \Delta \psi[\boldsymbol{u}]. \tag{3.11}$$

The error in the detector plane is given by $\varepsilon[\boldsymbol{u}] = \Delta\psi[\boldsymbol{u}]$ and $S^{-1}(\boldsymbol{u})$ defines the region in the detector plane that is influenced by information from $O[\boldsymbol{k}]$ in the forward convolution calculation. It is important to note that Eq. (3.11) does not uniquely map information back to a single object component because it is not possible to independently vary points in the detector plane since they are linked via the convolution relation in Eq. (3.4c). Nevertheless, such an approach does give a good handle on the phase components, provided that only the correct patch of the diffraction pattern is used. The key to preventing the inverse operation from being ill-conditioned is to correctly identify $S^{-1}(\boldsymbol{u})$, which was identified in the previous section as $S^{-1}(\boldsymbol{u}) = (|P[\boldsymbol{k} - \boldsymbol{u}]|) > 0$. So far this shows the set of $\boldsymbol{u}$-components that should contribute to the inverse transformation (i.e. magnitude information) but not information on how they should contribute (i.e. phase information). To resolve this issue, the inverse term $\frac{\partial O[\boldsymbol{k}]}{\partial \psi[\boldsymbol{u}]}$ of Eq. (3.11), which quantifies their contributions, needs to be calculated so that the phase information can be extracted.

For this task, we calculate the change of the object function in $\boldsymbol{k}$-space due to a change in a single component of the diffraction pattern. This is given by

$$\frac{\partial O[\boldsymbol{k}]}{\partial \psi[\boldsymbol{u}]} = \left(\frac{\partial \psi[\boldsymbol{u}]}{\partial O[\boldsymbol{k}]}\right)^{-1} \approx \left(\frac{P^*[\boldsymbol{k} - \boldsymbol{u}]}{|P[\boldsymbol{k} - \boldsymbol{u}]|^2}\right)\Big|_{S^{-1}(\boldsymbol{u})} \tag{3.12}$$

where the kernel of the convolution inverse in Eq. (3.12) is approximated by ignoring the regions where $S^{-1}(\boldsymbol{u}) = 0$. As far as information flow is concerned, this approximation is

valid in a general setup because we know the regions where $S^{-1}(\boldsymbol{u}) = 0$ do not have contributions from the object component $O[\boldsymbol{k}]$; consequently, any correction of the object component $O[\boldsymbol{k}]$ should not be influenced by regions where $S^{-1}(\boldsymbol{u}) = 0$. Fig. 3.8 illustrates the inverse mapping from the diffraction pattern to a spatial frequency of the object.

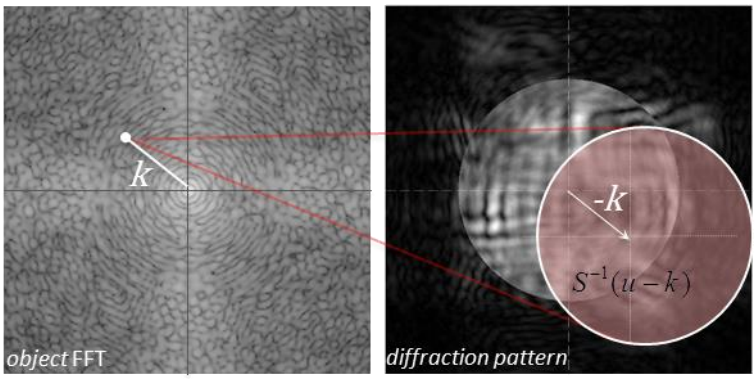


FIGURE 3.8 – Illustration of the convolution inverse domain

The iterative deconvolution kernel is built from the magnitude information of Eq. (3.10) and the phase information of Eq. (3.12). Since division by zero is avoided in Eq. (3.12) via the reverse of information flow constraint, the phase value of the approximated inverse kernel equals the phase of the conjugate kernel. This substitution transforms Eq. (3.13b) to Eq. (3.13c) in

$$IDK[\boldsymbol{k}-\boldsymbol{u}] = |IDK|\exp\left[i \cdot \arg\left(\frac{\partial O[\boldsymbol{k}]}{\partial \psi[\boldsymbol{u}]}\right)\right] \tag{3.13a}$$

$$= |P[\boldsymbol{k}-\boldsymbol{u}]|\exp\left[i \cdot \arg\left(\frac{P^*[\boldsymbol{k}-\boldsymbol{u}]}{|P[\boldsymbol{k}-\boldsymbol{u}]|^2}\Big|_{S^{-1}(\boldsymbol{u})}\right)\right] \tag{3.13b}$$

$$|P[\boldsymbol{k}-\boldsymbol{u}]|\exp[i \cdot \arg(P^*[\boldsymbol{k}-\boldsymbol{u}])] \tag{3.13c}$$

$$= P^*[\boldsymbol{k}-\boldsymbol{u}] \tag{3.13d}$$

This gives the IDK as the conjugate of the convolution kernel. The first paragraph of this section showed that the conjugate kernel acting as an IDK has the effect of minimising phase distortions in the spatial frequencies of the recovered object. Taking the Fourier transform of the IDK gives the corresponding update function as the conjugate of the illumination function, which is the PIE update function with the parameter configuration ($l = 2, \delta = 0$). Thus, the preceding analysis indicates that the PIE algorithm extracts specimen information in a way that minimises phase distortions in the spatial frequencies of the recovered object. Using the IDK in Eq. (3.13d), the procedure for mapping the feedback error function from the diffraction pattern to the object spatial frequency is

$$\varepsilon[\boldsymbol{k}] \propto \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{k}-\boldsymbol{u}]\varepsilon[\boldsymbol{u}] = \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{k}-\boldsymbol{u}]\Delta\psi[\boldsymbol{u}] \tag{3.14a}$$

$$\varepsilon_n[\boldsymbol{k}] \propto \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{k}-\boldsymbol{u}]\,\Delta\psi_n[\boldsymbol{u}] \tag{3.14b}$$

$$O_{n+1}[\boldsymbol{k}] = O_n[\boldsymbol{k}] + \varepsilon_n[\boldsymbol{k}] = O_n[\boldsymbol{k}] + \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{u}-\boldsymbol{k}]\Delta\psi_n[\boldsymbol{u}] \tag{3.14c}$$

$$O_{n+1}[\boldsymbol{r}] = O_n[\boldsymbol{r}] + P^*[\boldsymbol{r}]\,\Delta\psi_n[\boldsymbol{r}] \tag{3.14d}$$

Here $\Delta\psi_n[\boldsymbol{r}] = \pi_M\{\psi_n[\boldsymbol{r}]\} - \psi_n[\boldsymbol{r}]$ defines the exit wave difference in the nth iteration.

Eq. (3.14a) transforms the error from the detector plane to the $\boldsymbol{k}$-space representation via the conjugate kernel. Eq. (3.14b) expresses the error in Eq. (3.14a) as the feedback error in the nth iteration of the algorithm. In Eq. (3.14c), the feedback error is added to the $\boldsymbol{k}$-space running estimate of object, this is the correction step of the algorithm. Eq. (3.14d) is the Fourier transform of Eq. (3.14c), which transforms the solution to the specimen plane in real space.

These arguments provide a first proof for the intensity weighting of the PIE update function, which can now be viewed as a scheme that minimizes phase distortions of the recovered object's spatial frequencies.

3.4 The coupling function of the PIE algorithm

This section investigates the impact of the update function on the quality of the reconstructions from the PIE algorithm. The illumination intensity is used as a weight in the update function to provide an effective and practical way to extract the object information. This was shown to be equivalent to mapping errors from the correct patch of the diffraction pattern to the object using the conjugate of the convolution kernel. The derivation of the IDK in Section 3.3 shows that the conjugate kernel minimises the phase distortions of the recovered object's spatial frequencies but does not completely eliminate them. This is because the convolution kernel that makes the diffraction pattern (see Eq. (3.4)) is not orthonormal to the conjugate kernel. As a result, the error correction for a given spatial frequency inevitably gets contributions from adjacent spatial frequencies. The remaining phase and amplitude distortions are captured by the coupling function, which is given by

$$\eta[\boldsymbol{k}] = \mathcal{F}|P[\boldsymbol{r}]|^2 \tag{3.15}$$

Eq. (3.15) quantifies the coupling amongst spatial frequencies of the recovered object. This representation highlights the impact of magnitude and phase structure of the illuminating wave, which is not easily discerned in the real space representation. It also provides the means of quantifying the effect of the illumination structure on the quality of the recovered object, which is explored in Chapter 4. In the ideal case, the coupling function should be a

delta function so that all the object's spatial frequencies are accurately recovered. This requires the conjugate kernel to be the exact inverse of the convolution kernel. Since this is not the case, it is important to understand the implications of using the conjugate kernel as the IDK in the update equation.

The finite FoV recovered by the update equation (via the illumination intensity distribution) is directly connected to coupling errors amongst adjacent spatial frequencies of the recovered object and is encoded in Eq. (3.15). As the illumination size increases, $\eta[\boldsymbol{k}]$ approaches a delta function, because of the scaling property of the Fourier transform. Consequently, the reconstruction from the algorithm approximates the object over a larger FoV. To illustrate the coupling of the recovered object's spatial frequencies, consider the situation where the running estimate of the object contains an error in a single spatial frequency, i.e. $\Delta O_n[\boldsymbol{k}'] = \delta[\boldsymbol{k}']$. The update function spreads the error fed back to adjacent spatial frequencies during the update step,

$$
\begin{aligned}
\Delta O_{n+1}[\boldsymbol{k}] &= \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{u}-\boldsymbol{k}]\,\Delta\psi_n[\boldsymbol{u}], \\
&= \sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{u}-\boldsymbol{k}]\left(\sum_{S(\boldsymbol{k}')} P[\boldsymbol{u}-\boldsymbol{k}']\Delta O_n[\boldsymbol{k}']\right), \\
&= \sum_{S^{-1}(\boldsymbol{u})}\sum_{S(\boldsymbol{k}')} P[\boldsymbol{u}-\boldsymbol{k}']\Delta O_n[\boldsymbol{k}']\,P^*[\boldsymbol{u}-\boldsymbol{k}], \\
&= \sum_{S(\boldsymbol{k}')}\left(\sum_{S^{-1}(\boldsymbol{u})} P^*[\boldsymbol{k}-\boldsymbol{u}]P[\boldsymbol{u}-\boldsymbol{k}']\right)\Delta O_n[\boldsymbol{k}'], \\
&= \sum_{S(\boldsymbol{k}')} \eta[\boldsymbol{k}-\boldsymbol{k}']\Delta O_n[\boldsymbol{k}'], \\
&= \sum_{S(\boldsymbol{k}')} \eta[\boldsymbol{k}-\boldsymbol{k}']\delta[\boldsymbol{k}'] = \eta[\boldsymbol{k}],
\end{aligned}
\tag{3.16}
$$

where $\eta[\boldsymbol{k}] = \mathcal{F}|P[\boldsymbol{r}]|^2$, using the Fourier convolution theorem. The second line of Eq. (3.16) assumes that the estimated diffraction pattern is close to the measured diffraction pattern so that the difference can be expanded using the FFT of the illumination as a basis. This is possible because the diffraction pattern is built out of shifted versions of the illumination FFT via the convolution operation in Eq. (3.4).

Phrased differently, updating the running estimate of the object with the weighted intensity of the illumination is equivalent to obtaining an estimate of the underlying object with spatial frequency errors determined by the coupling function ($\eta[\boldsymbol{k}]$). This analysis suggests that any departure of $\eta[\boldsymbol{k}]$ from a delta function should show up in the recovered image. The case where the specimen is illuminated by a localised spot gives rise to an $\eta[\boldsymbol{k}]$ with low spatial frequencies departure from a delta function profile, this limits the FoV of the recovered object. Thus, a finite FoV in this framework is equivalent to a form of structured low frequency noise in real space, where the variation of the noise is of the order of the size of the illumination. By the same logic, non-zero values of $\eta[\boldsymbol{k}]$ at high spatial frequencies should result in high frequency distortions in the recovered object. Thus highly structured illumination that gives rise to peaks in $\eta[\boldsymbol{k}]$ at locations other than the central point should result in high frequency distortions in the recovered image. The effects of spatial frequency coupling in single diffraction pattern calculations are demonstrated in Chapter 4. An interesting property of the ptychographic dataset is the extension of the recovered FoV, which in a ptychographic reconstruction is, in principle, unlimited. Thus, extension from a single diffraction pattern to a ptychographic dataset must consequently provide the means for reducing the low frequency distortions in the recovered object.

3.5 Summary

This chapter shows that characterising the convergence property of the update function in different experimental setups requires that the solution provided by the update function be viewed in a framework where interference provides a metric for solubility of the phase problem. The derivation of the form of the update function in Section 3.3 shows that the update function extracts spatial frequency information of the specimen from the correct patch of the detector plane in a way that minimises relative phase distortions in those spatial frequencies. Section 3.4 introduced the coupling function, which quantifies the effects of using the illumination intensity as a weight in the update function.

Chapter 4

4 Impact of experimental constraints on the PIE update function

This chapter further investigates the update function for the Type-II setup to get a handle on how it extracts object information from diffraction patterns. In the Type-II setup the illumination is formed by a lens and is consequently curved. The calculations of Section 4.1 are used to investigate the performance of the update function for different types of specimen functions, such as transparent and diffuse objects. In Section 4.2 the number of counts in the diffraction patterns, processed by the update function, is reduced to a value that gives a poor representation of the test object. Analysis of the resulting trend (quality of the recovered objects) provides insight into signal to noise (SNR) properties of the Type-II diffraction patterns. In the Type-II setup, the main parameters that control the size of the illumination/probe at the specimen plane are the defocus and size of the lens aperture. Sections 4.3 to 4.5 investigate the trend of the error metric with each of these parameters in order to determine the optimum parameter configuration for the Type-II setup. Furthermore, it is useful to recover a large field of view (FoV) with high-resolution information; hence

Section 4.5 also investigates the largest illumination size that can be accommodated in Type-II diffractive imaging. In this investigation, the size of the illumination is increased using the defocus parameter to regimes that highlights the impact of sub-Nyquist sampling on the reconstruction. Section 4.6 investigates the impact of incorrect characterisation of the defocus parameter on the quality of the recovered object in the Type-II setup.

4.1 Impact of object type on the update function

This section aims to understand the influence of object type on the update function. For this investigation, the update function is applied to magnitude-only and phase-only objects derived from the images in Fig. 4.1 and Fig. 4.2b. The calculations with magnitude-only object model situations, where the object is absorbing, as in some classes of visible light and X-ray experiments. The phase-only object calculations model both visible light and electron experiments. The modified update function of Eq. (3.1b) is used in the magnitude-only object calculations, where the real part of the iterative calculations is added to the running estimate of the object. A similar routine was adapted to the phase-only object calculations, in which the magnitude of the running estimate of the object is set to unity at each iteration step of the algorithm. These provide additional projection constraint sets defined by the real line (in the case of magnitude-only objects) and the circle of unit radius centred at the origin (in the case of phase-only objects). The real line constraint reduces the number of unknown variables in the detector plane by a factor of 2. This is because the Fourier transform of a real object ($g(\boldsymbol{r})$) is the Hermitian function $g(\boldsymbol{k})$, that possesses the symmetry for which complex conjugation together with change of the variable sign leaves the function unchanged; i.e.

$g(\boldsymbol{k}) = g^*(-\boldsymbol{k})$, where $\boldsymbol{k}$ represent the Fourier coordinate and $g(\boldsymbol{k}) = \mathcal{F}g(\boldsymbol{r})$ gives the Fourier transform relationship. Consequently, only half of the Fourier plane has independent coefficients. The unit circle constraint has the same impact as the modulus constraint of the diffraction pattern with the exception that the intensity value in real space is equal to one for all the object pixels.

The results from calculations that use the transparent object in Fig. 4.1 are compared to the corresponding Gabor hologram reconstructions (See Fig. 4.3 and Fig. 4.4). The recovered objects from the Gabor holography method were calculated by multiplying the central bright field disc of the diffraction pattern by the reference wave (FFT of the illumination); the estimated exit wave at the specimen plane is the inverse FFT of this complex wave distribution. In the case where the test object comprises only real transmission coefficient values, the magnitude of the exit wave was extracted. The phase of the exit wave was multiplied by the conjugate of the illumination function for the case where the test object is a phase-only object; this was done to remove the curvature of the illumination. The results of these calculations are shown in Fig. 4.3b and Fig. 4.4b.

The illumination used for this investigation was modelled in terms of an electron experiment with a defocus of 300 nm and a convergence semi-angle of 15 mrad. Fig. 4.2a shows the magnitude of the illumination at the specimen plane. The initial estimate to the specimen was free-space with uniform transmission value of unity and the appropriate modified update functions were applied to the diffraction patterns over 10000 iterations. The results are presented in the next section.

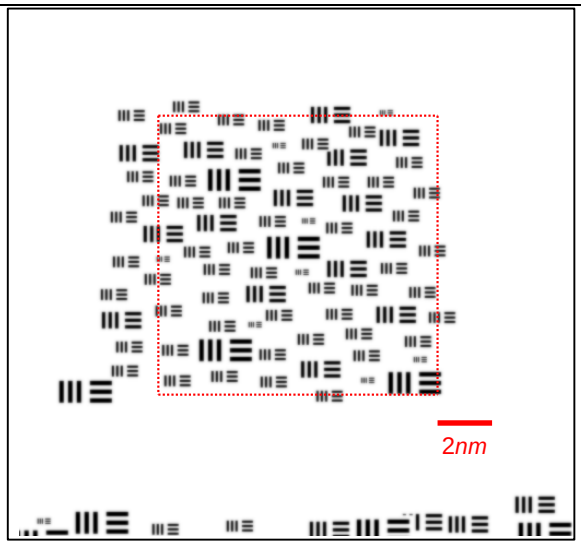
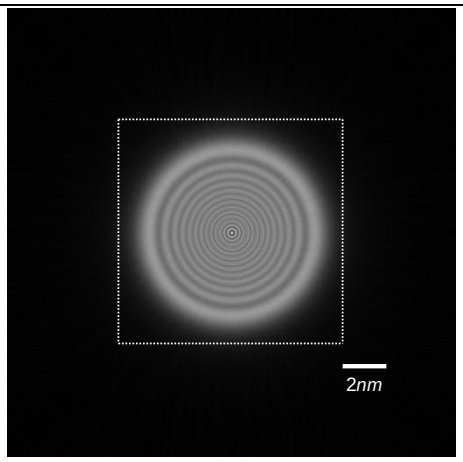
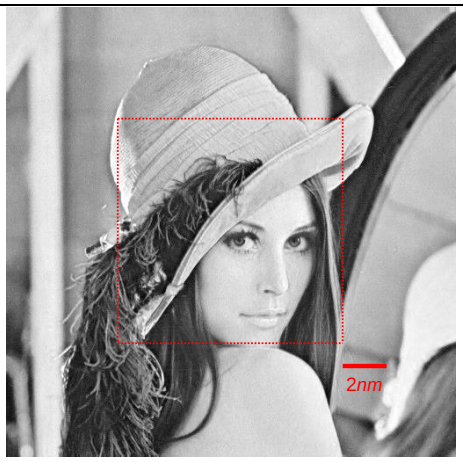


FIGURE 4.1 - Test specimen that models a resolution target. The calculations in this chapter employ the entire object but only the region enclosed by the rectangle is shown in the most of the result sections because it corresponds to the region where the illumination has high intensity values. For magnitude-only object calculations, the range of the specimen is $[0.1\ \ 1]$ and for phase-only object calculations, the range of the specimen is $[-0.9\pi\ \ 0]$.



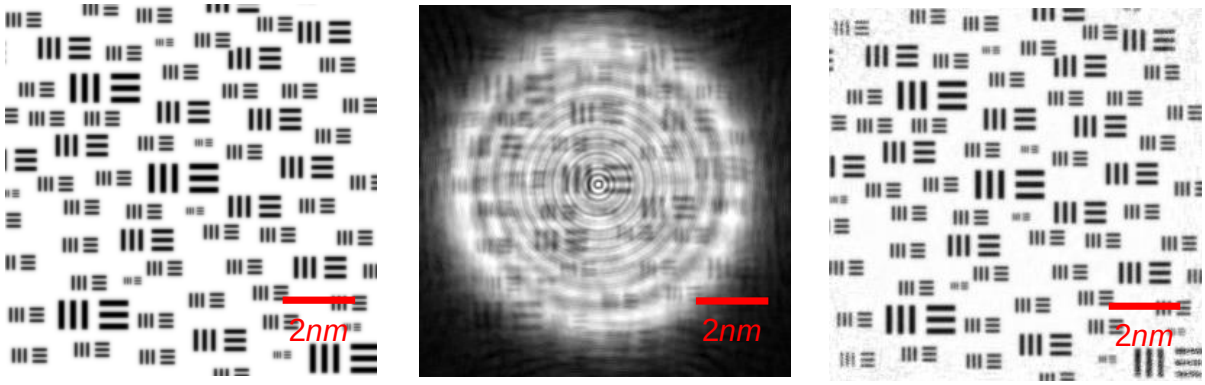
(a)



(b)

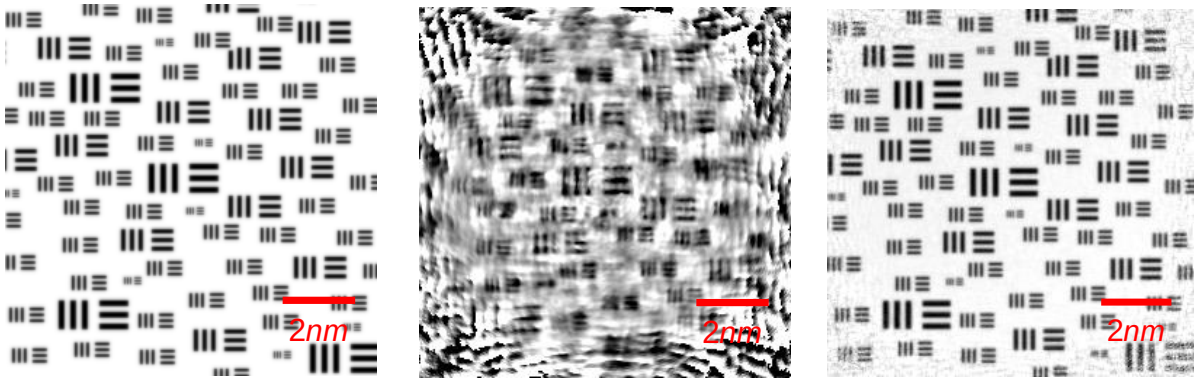
FIGURE 4.2 – (a) Magnitude of illumination function that models a 200 keV electron beam with defocus of 300nm. The aperture at the back focal plane of the illumination forming lens span a semi-angle of 15 mrad. (b) Test specimen used for textured object calculations. The dashed lines in these images delineate the region of the recovered objects in Fig. 4.3 to Fig 4.5.

4.1.1 Results



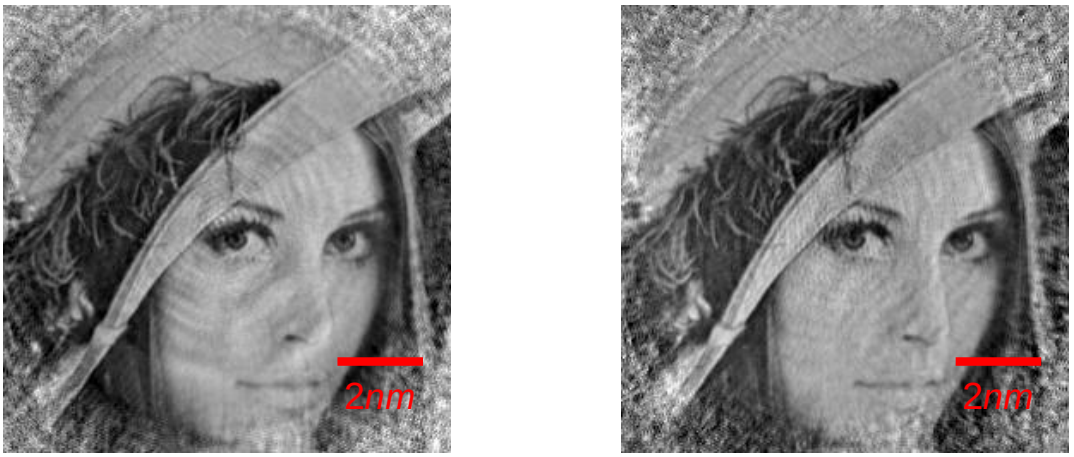
(a) (b) (c)

FIGURE 4.3 – (a) Test specimen comprising real-valued transmission coefficients. (b) The Gabor hologram reconstruction of the magnitude object in (a). (c) Magnitude of recovered object after 10000 iterations of the calculations that used the modified update function in Eq. (3.1b). The field of view of the images shown in (a), (b) and (c) correspond to the region enclosed by the rectangles in Fig. 4.1 and Fig. 4.2.



(a) (b) (c)

FIGURE 4.4 - (b) Test specimen used for phase object calculations. The range of the specimen phase is $[-0.9\pi \ \ 0]$. (b) The Gabor hologram reconstruction for the phase object in (a). (c) Phase of recovered object after 10000 iterations of the update function, with the additional constraint that sets the magnitude to the running guess of the object to unity at the start of each iteration. The field of view of the images shown in (a), (b) and (c) correspond to the region enclosed by the rectangle in Fig. 4.1.



(a) (b)

FIGURE 4.5 - (a) Magnitude of recovered object for magnitude-only diffuse object calculations. (b) Phase of recovered object for phase-only diffuse object calculations. The field of view of these images corresponds to the region enclosed by the rectangle in Fig. 4.2b. These results correspond to the output of the algorithms after 10000 iterations.

4.1.2 Discussion

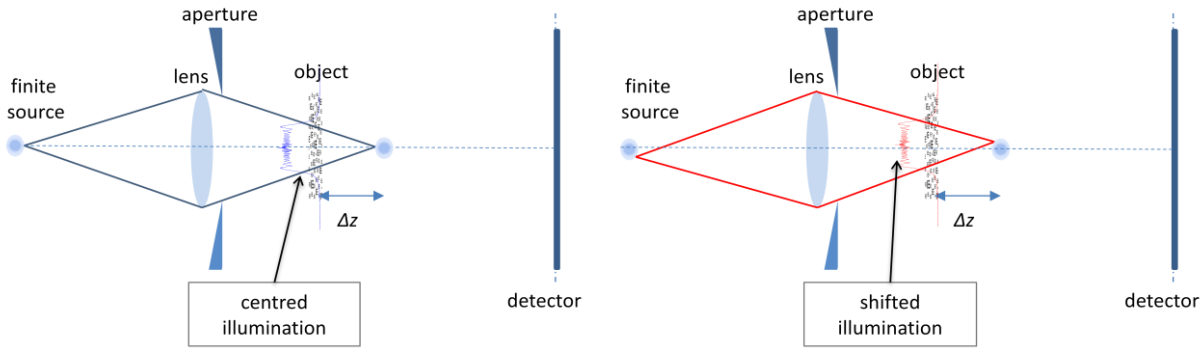
The recovered objects in Fig. 4.3c and Fig. 4.4c show that the algorithm recovers a better representation of the object compared to the corresponding reconstructions from the Gabor reconstructions in Fig. 4.3b and Fig. 4.4b. This is understandable because the Gabor reconstruction can be retrieved from the estimated exit wave of the algorithm after the first iteration. This estimated exit wave comprises the sum of the Gabor exit wave and contributions from points outside the bright field disc of the diffraction pattern. The algorithm refines this estimate in subsequent iterations using the appropriate modified update function, where the known phase curvature of the illumination discriminates against the twin image solution that distorts the Gabor reconstructions. The quality of the magnitude and phase reconstructions in Fig. 4.3c and Fig. 4.4c indicate that the update function can be tailored to provide similar performance in iterative algorithms that recover magnitude-only or

phase-only objects. The magnitude and phase reconstructions in the case of diffuse object (see Fig. 4.5) contain imprints of the illumination magnitude. This means that application of the update function to single diffraction patterns produces better quality reconstruction in the case where the object is sparse (like the object in Fig. 4.1). This is an interesting result because one would expect the converse to be the case, since a textured object samples the illumination at a larger number of points and should therefore provide a more constrained calculation. Subsequent calculations in this chapter employ the magnitude-only version of the resolution target in Fig. 4.1 to further investigate the properties of the update function.

4.2 Impact of Counts on the update function

In diffractive imaging, the number of counts in the diffraction pattern needs to be high enough to encode information about the scattering potential of the illuminated object. The number of counts in the diffraction pattern of the Type-II experiments is governed by several experimental factors. For example, in the electron microscope implementation of the Type-II experiment, the specimen may drift during data collection. Specimen drift can arise from thermal gradient in the specimen holder or local charging of the specimen by the electron beam. The phase curvature of the illumination over the specimen in the Type-II setup means that lateral movement of the specimen during data collection affects the distribution of intensities in the detector plane. This compromises the quality of the diffraction pattern by blurring interference fringes. In the case where the specimen is illuminated by a defocused beam, translations in the specimen plane result in both translation and distortion of the shadow image within the central bright field region of the diffraction pattern. This has the

same effect as using an extended source that introduces spatial partial coherence into the experiment (see Fig. 4.6). Partially coherent diffraction patterns degrade the quality of the recovered object because diffractive imaging calculations employ wave propagation between the specimen and detector plane that requires coherence for meaningful interpretation of phase difference between different parts of the wave. In the type-II setup, iterative calculations that employ a real space sampling pitch, greater than the effective spot size at the lens focus, can correctly process the corresponding partially coherent diffraction pattern.



(a) Illumination profile from a point source that is located along the central axis of the extended source.

(b) Illumination profile from a point source that is displaced from the central axis of the extended source.

FIGURE 4.6 - Translation effect of illumination profile due to finite source size in the Type-II setup. This shows the correspondence between specimen drift during data collection and partial coherence in the Type-II setup.

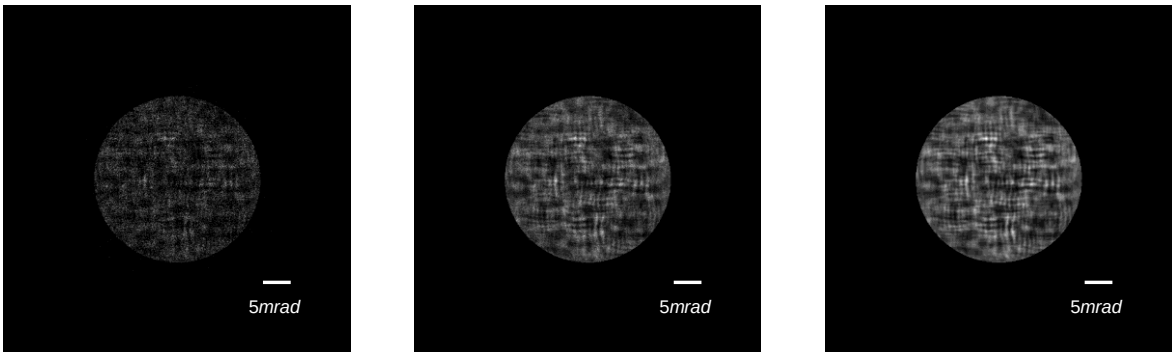
The effect of drift on the Type-II diffraction pattern can be minimised by reducing the exposure time of the diffractive imaging experiment, which reduces the counts in the diffraction patterns. Furthermore, the source size is demagnified to improve coherence of the source at the specimen plane and results in low beam intensity. There is also the issue of specimen damage from the beam that is especially true in the context of electron microscopy imaging of soft matter (e.g. biological specimens) where counts as low as 10 electrons/Å^2 are sometimes required. The effect of damage can also be minimised by using a short exposure

time. These constraints set an upper limit on the number of counts in diffractive imaging experiments and may well result in noisy diffraction patterns.

In low count electron experiments, the dominant source of noise is described by the statistical distribution of electron arrivals at a given pixel of the detector and the distribution is Poisson (Spence, 2003). In the case where N electrons are expected to arrive at a given detector pixel due to finite exposure time, the standard deviation of Poisson statistics gives the noise in that pixel as $\sqrt{N}$. In other words, the signal to noise ratio (SNR) for that pixel is given by the ratio of the mean signal (N) to the noise $\left(\sqrt{N}\right)$, shown in Eq. (4.1).

$$SNR = \frac{N}{\sqrt{N}} = \sqrt{N} \tag{4.1}$$

In the Type-II setup, the diffraction pattern is made up of two regions called the bright field and the dark field; the bright field corresponds to the central bright disc region of the diffraction patterns of Fig. 4.7 and the dark field corresponds to the outer region. In the absence of a specimen, all the counts are registered in the bright field. Thus the bright field act as a reference region in the Type-II setup. The resulting diffraction pattern is close to the Gabor hologram (Gabor, 1948) but we can have strong scattering specimens with textured backgrounds. For the specimen used in these calculations (see Fig. 3.3a), the measured intensities in the bright field result from the interference of diffracted beams with the reference beam. This means that in a low count experiment, the majority of the counts reside in the central bright field disc of the diffraction pattern. Eq. (4.1) suggests that the SNR of scattered signals that resides in the bright field region is greater than those in the dark field region, because the number of counts in the bright field disc is higher.



(a) Diffraction pattern from calculations using 10^5 counts

(b) Diffraction pattern from calculations using 10^6 counts

(c) Diffraction pattern from calculations using 10^8 counts

FIGURE 4.7 - Diffraction patterns from calculation using different count values. These diffraction patterns were used as inputs to the PIE update function for the generation of the results in the next section. The diffraction patterns in (a), (b) and (c) span a semi-angle of 62.3 mrad.

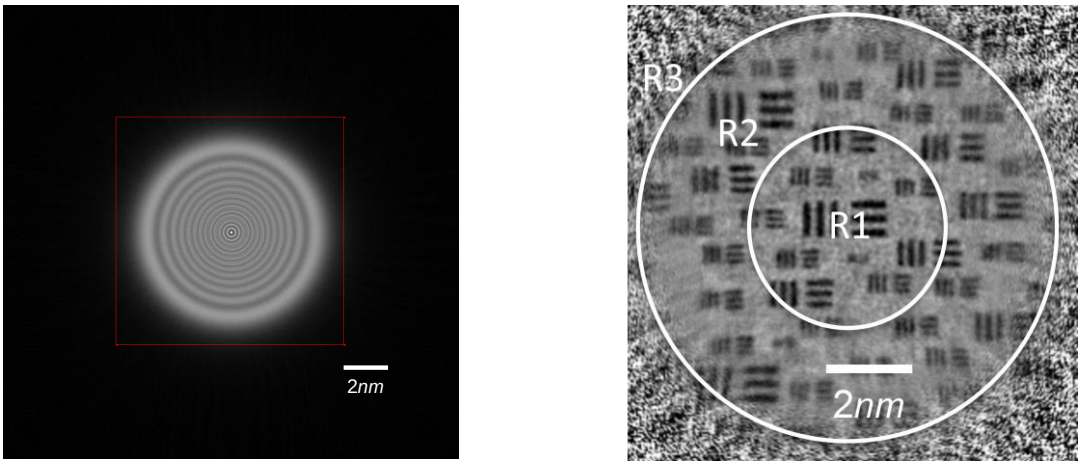
To model the effect of different exposure times, the power of the diffraction pattern was fixed by the total number of counts in the experiment. Poisson noise was added to the diffraction patterns, where the intensities of the calculated diffraction patterns were used as mean values of the Poisson distribution. The noisy diffraction patterns were used as model diffraction patterns for the associated count values. The diffraction patterns generated for counts values of 10^5, 10^6 and 10^8 are shown in Fig. 4.7. These diffraction patterns were processed by the algorithm in order to understand the impact of counts on the update function and to investigate how much specimen information the update function can extract from noisy diffraction patterns.

The illumination used for this investigation corresponds to a 200 keV defocused electron probe generated with a lens that has a convergence semi-angle of 15 mrad. The specimen was offset from the focus of the lens by a defocus distance of 300 nm. Fig. 4.8 shows the magnitude of the illumination at the specimen plane. This investigation employed the resolution target Fig. 4.1 as a test specimen and the initial estimate to the specimen was free-space with uniform transmission value of unity. The algorithm applied the update function to

the noisy diffraction patterns over 10000 iterations and results are presented in the next section.

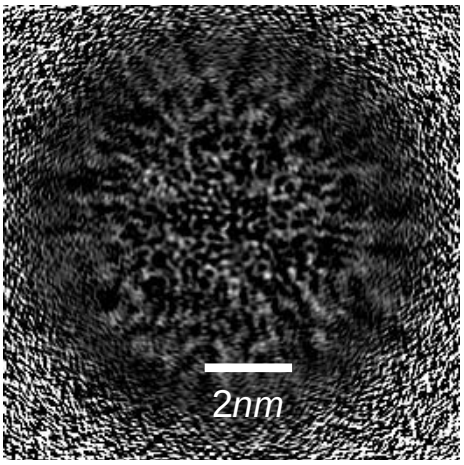
4.2.1 Results

Fig. 4.9 shows the middle part of the recovered object transmission functions for the cases with count values of 10^4, 10^5, 10^6 and 10^8. The region delineated by the dashed line in Fig. 4.8a corresponds to the part of the recovered object shown in Fig.4.8b and Fig. 4.9. To aid in the analysis of the qualities of the recovered images, it is useful to split the recovered objects into three regions; the inner region (R1), the middle region (R2) and the outer region (R3) as shown in Fig. 4.8b. The outer region R3 corresponds to parts of the recovered object where the illumination has low intensity values. The middle region R2 and the inner region R1 refer to the parts of the object, where the illumination has high intensity values and are distinguished by the average displacement of the object pixels from the centre of the illumination. The evolution of the error metric over 10000 iterations is shown in Fig. 4.10.

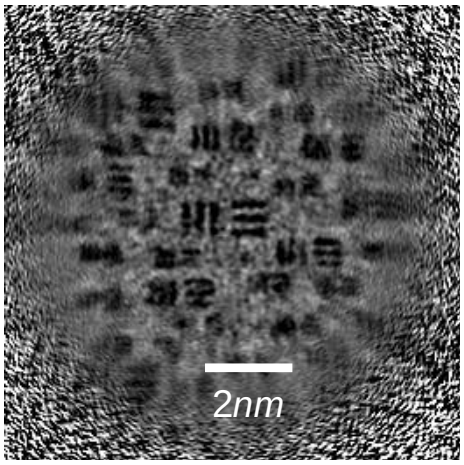


(a) (b)

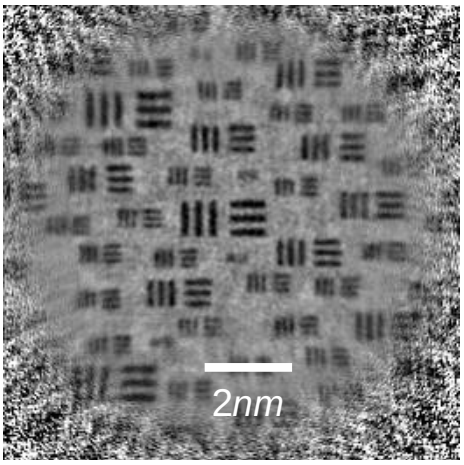
FIGURE 4.8 - (a) Magnitude of illumination function used in count calculations. The illumination models a 200 keV electron beam with a defocus of 300 nm and a convergence semi-angle of 15 mrad. The rectangle delineates the region of the recovered objects in (b) and Fig. 4.9. (b) Identification of three regions with different recovered properties. The sampling pitch of these calculations is 0.04 nm and the calculation size is 512x512 pixels.



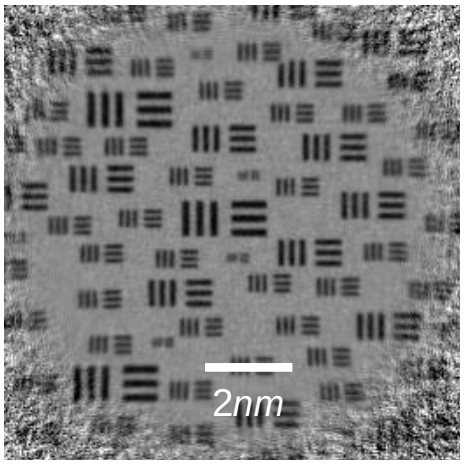
(a) Recovered object from calculations using 10^4 counts



(b) Recovered object from calculations using 10^5 counts



(c) Recovered object from calculations using 10^6 counts



(d) Recovered object from calculations using 10^8 counts

FIGURE 4.9 - Magnitude of recovered objects after 10000 iterations of the algorithm for calculation that use different count values. These results show that the recovered object from different count calculations contain graininess in region R3. The image in (a) shows that the update function fails to recover the object features in calculation where the number of count in the diffraction pattern equals 10^4. The sampling pitch of these calculations is 0.04 nm and corresponds to a semi-angle at the detector plane of 62.3 mrad.

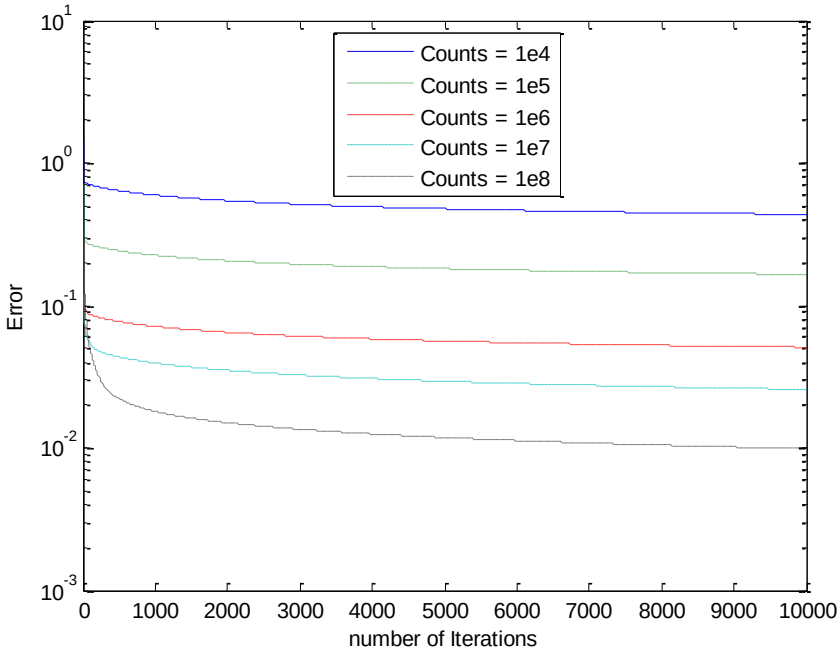


FIGURE 4.10 - Evolution of the error metric over 10000 iterations for calculations with different counts.

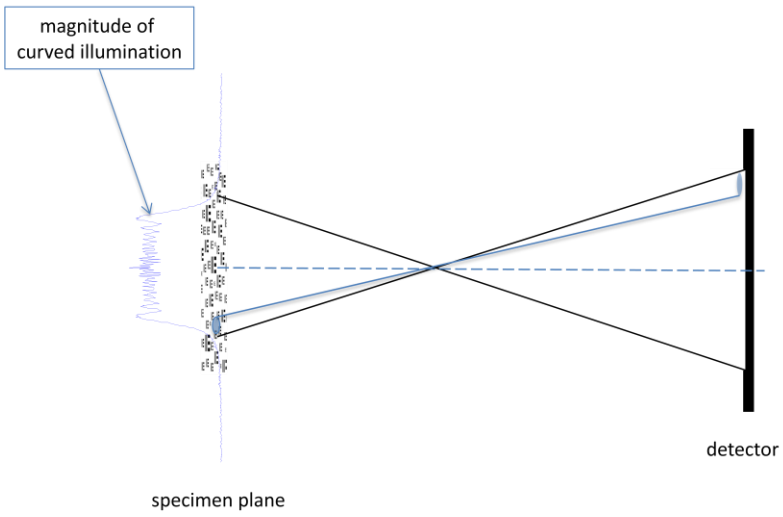


FIGURE 4.11 - illustration of mapping between points at the specimen plane and points in the central bright field disc in the case where the illumination is a defocused probe formed by a lens. Points inside the region defined by the effective illumination size are mapped to points in the bright field disc of the diffraction pattern while points outside this region are mapped to the dark field of the diffraction pattern.

4.2.2 Discussion

As we would expect, Fig. 4.9 shows that the update function provides the best representation of the test specimen in the calculations with the highest count value. Fig. 4.9a shows that the update function fails to recover the object features in the case where the number of counts in the diffraction pattern equals 10^4. The departure of the recovered objects from the test specimen in R3 appears as high frequency noise in the recovered objects. This noise is less pronounced in Fig. 4.9d where parts of the specimen are discernable but the same region in Fig. 4.9b and Fig. 4.9c are completely noisy. This suggests that the regions with low illumination intensity have the lowest SNR in the specimen plane. This is understandable because the number of electrons the traverse any part of the specimen is given by the intensity of the illumination, so that a point with a relatively higher illumination intensity value has a greater probability that its information is expressed in the diffraction pattern for the limited exposure time of the experiment.

To understand the origin of the high frequency noise in R3 of the recovered object, we need to examine the way that the update function incorporates information into the recovered object. During the forward calculations, the specimen transmission function is multiplied by the illumination function; this suppresses the influence from specimen points in R3 because their contributions are not strongly expressed in the calculated diffraction patterns. However as the iteration progresses, the algorithm slowly updates the information in R3 because the illumination is not zero in this region. As a result, information is introduced into this region by the update function. In the previous paragraph, we inferred that the SNR is small in this region so that for a short exposure experiment the information that pertains to this region may not be strongly expressed in the diffraction pattern. Thus attempts by the algorithm to extract information from the diffraction in relation to this region (R3) result in noise. This is a

compelling argument for the origin of high frequency noise because Fig. 4.9d indicates that the algorithm does incorporate some of the object information in R3 with less noise; thus the quality improves with the total counts in the diffraction pattern.

The degradation of the recovered object in R2 is manifested as loss of angular resolution as we move radially outwards in R2. Since the illumination intensity is high in R2, this suggests that the illumination intensity cannot completely account for the quality of the recovered object. To explain the quality of the reconstructions in R2, we need to view the expression of information in the diffraction pattern. Using geometric optics, we can show that the location of a shadow image within the bright field of the diffraction pattern has a one to one correspondence with part of the specimen in the illuminating beam, as illustrated in Fig. 4 11. This means that parts of the specimen that lie close to the edge of the illumination have low spatial frequency information residing close to the edge of the disc that defines the bright field. Consequently slightly higher spatial frequency information resides outside the bright field where the SNR is low. This means that as we move towards the edge of the illumination, the likelihood of the specimen scattering electrons outside the bright field increases; as a result, most of the information in R2 resides in the dark field. Since the dark field has a low SNR, the outer region of the image is recovered with low resolution compared to the inner region. By the same logic, the outer part of the illumination (R3) with low intensity values can be mapped to points in the dark field so that low SNR of the dark field translate to low SNR in R3. R2 appears smooth rather than noisy in the low counts calculation because in situations where most of the spatial frequency information reside in the dark field, the bright field still encodes the unscattered signal from specimen points in R2; thus the only information the algorithm can extract in these regions corresponds to absorption strength of the specimen features but not the detailed structure of the feature as evident in Fig 4.3b.

Fig. 4.10 shows that the recovered object generates a diffraction pattern that disagrees more with the measured data for low count experiments because the error is higher for low count diffraction pattern calculations. This is due to the fact that the diffraction pattern from a thin specimen builds up information about the convolution between the illumination and the specimen transmission function as the number of counts increases; consequently, diffraction patterns with small number of counts do not correctly represent the convolution integral and the estimated diffraction patterns from iterative calculations (that employ a convolution operation) deviate more from the diffraction data.

It is worth noting that in all of these count calculations (see Fig. 4.9), the recovered object did not possess high frequency noise in the region where the illumination had high intensity values. This means that a noisy diffraction pattern does not result in a high frequency noisy image in R1 and R2. This is an interesting result because it suggests that applying the correct intensity weight of the update function on noisy diffraction patterns injects additional information into the iterative calculations, since the illumination function encodes probabilistic information about the interaction between the specimen transmission function and the incident radiation. This additional information facilitates extraction the specimen information from noisy data in the bright field disc. This result is very important for low dose applications in biology where a 3Å resolution is good enough for most biological features of interest; in which case we only need to process the central bright field disc of the diffraction pattern, to generate an image of the specimen that is not hampered by the low SNR property of the dark field region.

4.3 Optimisation of illumination parameters

The previous section showed that for modest counts values above 10^6 counts the update function correctly extracts specimen information from the diffraction pattern in regions where the illumination has high intensity. This suggests that it may be beneficial to engineer the illumination to have high intensities over a larger area at the specimen plane and thus provide a large FoV representation of the specimen. In the Type-II setup, the major parameters that affect the size of the illumination are the defocus and the aperture size of the lens. Here we investigate the impact of these parameters on the update function where the performance was measured by the error metric.

The illumination size depends on the aperture size and the defocus via different mechanisms. In the case of defocus, the total phase curvature of the illumination over the back focal plane of the lens increases with defocus and the resulting change in interference conditions gives rise to a larger beam. However, in the case of varying aperture size, the total phase curvature of the illumination over the back focal plane of the lens is fixed, and a larger aperture size introduces additional higher angle plane-waves at the specimen plane. The resulting change in interference conditions also gives rise to a larger beam. To decouple the effects of these parameters, the calculations in Section 4.4 were performed with a fixed defocus value and the calculations in Section 4.5 were performed with fixed aperture size.

4.4 Impact of lens aperture size on the PIE update function

In Section 4.2, it was shown that the central bright field region of the diffraction pattern influences the quality of the reconstruction because of its high SNR property. The aperture size of the lens in the Type-II setup determines the size of the bright field of the diffraction pattern, thus a large lens aperture increases the region of the diffraction pattern with high SNR. The lens aperture also provides a means to control the radiation flux at the specimen because it acts as a Fourier filter. For example, in the electron microscope, the condenser aperture blocks electrons outside the aperture from reaching the specimen. Furthermore, the size of the lens aperture plays a crucial role in the Type-II diffractive-imaging setup because it defines the size of the convolution kernel that generates the diffraction pattern. This is because points in the aperture generate plane-waves that coherently illuminate the specimen and thus define the illuminating wave. Consequently, the number of object spatial frequencies that contribute to a single diffraction intensity is directly linked to the aperture size in the Type-II setup.

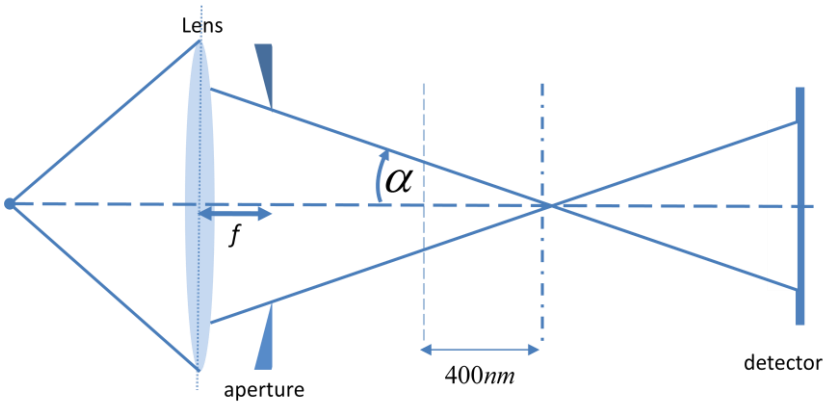


FIGURE 4.12 - Experimental setup modelled by calculations in this section. The angular size of the aperture at the plane of the specimen in given by α mrad. The calculations in this section used aperture sizes ranging from 5 mrad to 20 mrad.

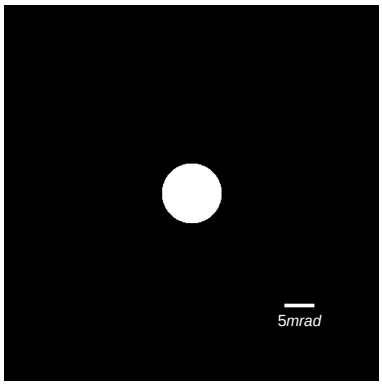
Since diffractive imaging is all about interference amongst the specimen spatial frequencies, we expect a large aperture to give a better-conditioned diffraction pattern. Moreover, shrinking the aperture size to a value that corresponds to a point in $\boldsymbol{k}$-space is equivalent to illuminating the specimen with a plane-wave and the resulting diffraction pattern for the case of an extended object does not encode phase information because the corresponding spatial frequencies are not given the opportunity to interfere.

As we increase of the aperture size in the back focal plane of the lens, the diffraction limited illumination spot formed at the lens focus decreases due to a reduction in diffraction limitation between these planes. The illumination size at planes close to focus actually increases in size and can be inferred from Fig. 4.12 using a geometric optics argument where the illumination size equals to $2 \cdot 400 \cdot \tan(\alpha)\ nm$. This shows that for a fixed defocus value ($400\ nm$ in this section calculations), the illumination size increases with aperture size. The relationship is linear for small values of the convergence semi-angle (α) since $\tan(\alpha) \approx \alpha$; here the convergence semi-angle relates directly to the aperture size as illustrated in Fig. 4.12.

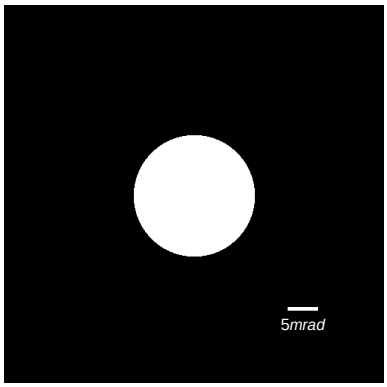
In Chapter 3, we showed that the coupling function ($\eta[\boldsymbol{k}]$) is given by the FFT of the illumination intensity in Eq. (3.15) and as a result it determines the FoV of the iterative calculations. The calculations in this section affect the illumination size (intensity distribution) at the specimen plane, which will show the impact of the coupling function on the convergence rate of the algorithm.

Fig. 4.13a to Fig. 4.13c show the back focal plane of the condenser lens for aperture sizes ranging from $\alpha = 5\,mrad$ to $\alpha = 15\,mrad$, and Fig. 4.13d to Fig. 4.13f show the corresponding diffraction patterns. The diffraction patterns used in all calculations in this section contain 10^8 counts. The illumination functions used in these calculations model a

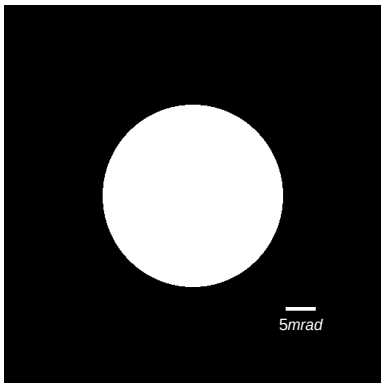
200 keV electron beam with a defocus of 400nm. This investigation employed the magnitude-only version of the resolution target in Fig. 4.1 as a test specimen and the initial estimate of the specimen was free-space with uniform transmission value of unity.



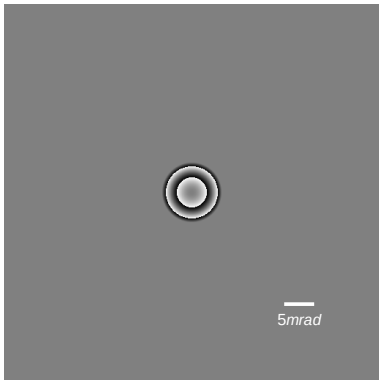
(a_1) 5 mrad aperture - Magnitude.



(b_1) 10 mrad aperture - Magnitude.

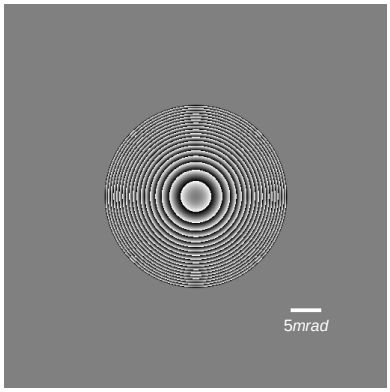


(c_1) 15 mrad aperture - Magnitude.

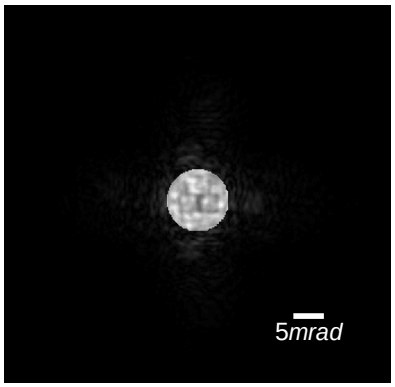


(a_2) 5 mrad aperture - Phase

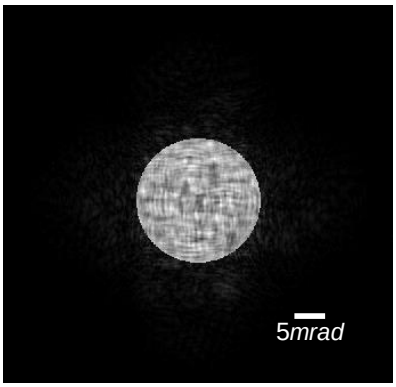
(b_2) 10 mrad aperture - Phase



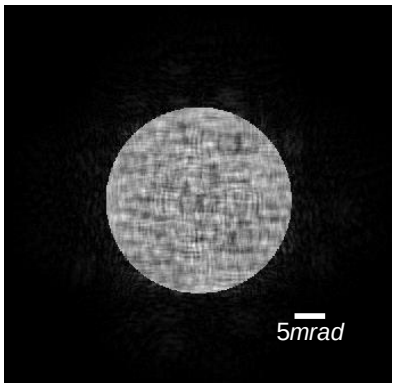
(c_2) 5 mrad aperture - Phase



(d) diffraction pattern from calculation using 5 mrad aperture



(e) diffraction pattern from calculation using 10 mrad aperture

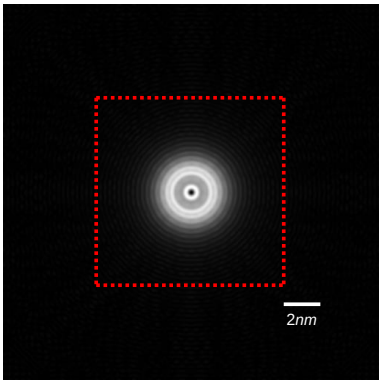


(f) diffraction pattern from calculation using 15 mrad aperture

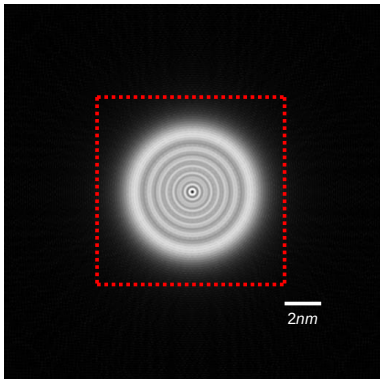
FIGURE 4.13 - Back focal plane of condenser lens and corresponding diffraction patterns using electron beam energy of 200 keV and defocus of 400nm. (a_1) to (c_1) The central region in white correspond to a transmission value of 1 and the opaque region outside has a transmission value of zero. (a_2) to (c_2) Phase curvature over the back focal plane of the condenser lens. This is also the phase distribution of the convolution kernel that generates the diffraction pattern. The quadratic phase curvature is governed by the defocus of the illumination. The aperture size limits the maximum phase gradient of the convolution kernel.

4.4.1 Results

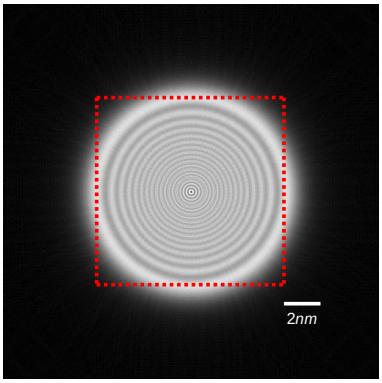
Fig. 4.14d to Fig. 4.14f show the recovered images from calculations employing apertures sizes of 5 mrad, 10 mrad and 15 mrad; where the regions delineated by the dashed lines in Fig. 4.14a to Fig. 4.14c corresponds to the part of the recovered object shown in the results. Fig. 4.15 shows the trend of the error metric with aperture size. The lines of the error plot connect several aperture size calculations for different numbers of iteration. Fig. 4.16 shows a line profile of the coupling function, which provides a tool to explain the results.



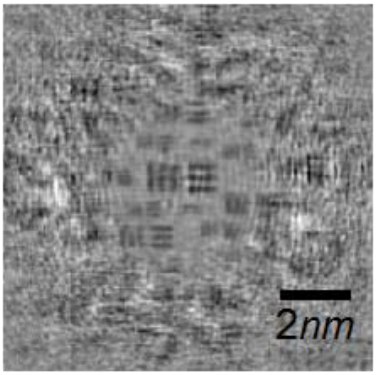
(a) Magnitude of Illumination function at the specimen plane made by a lens with 5 mrad aperture size



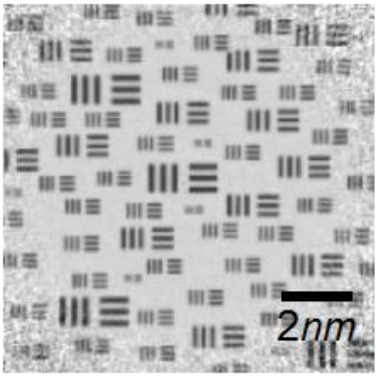
(b) Magnitude of Illumination function at the specimen plane made by a lens with 10 mrad aperture size



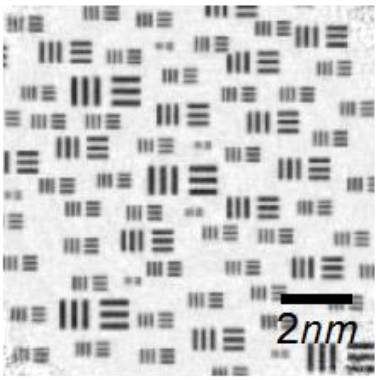
(c) Magnitude of Illumination function at the specimen plane made by a lens with 15 mrad aperture size



(d) Recovered image from 5mrad probe calculation



(e) Recovered image from 10mrad probe calculation



(f) Recovered image from 15mrad probe calculation

FIGURE 4.14 - (a) to (c) show the probe at the specimen plane for aperture sizes in the range of 5 mrad to 15mrad. (d) - (f) show the recovered object for the region enclosed by the respective rectangles of (a) to (c). All calculations employ the same total count value of 10^8.

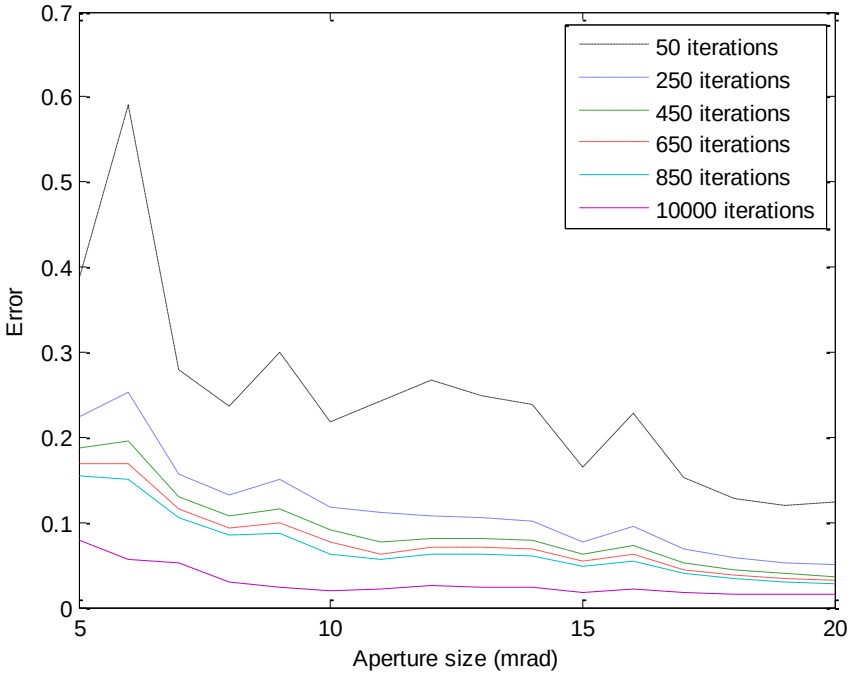


FIGURE 4.15 – Trend of the error metric with aperture size for different number of iterations of the update function.

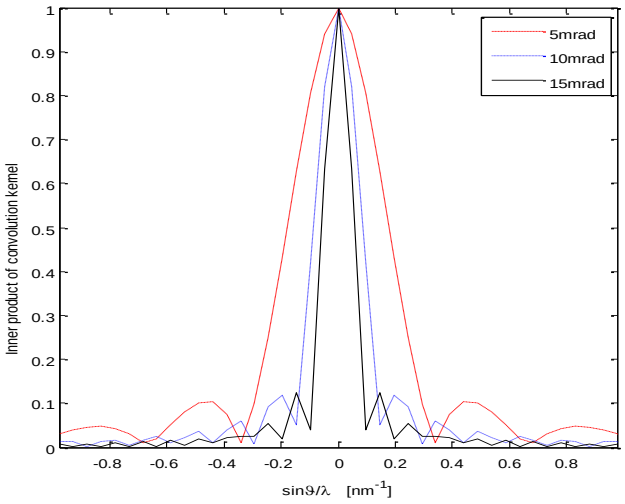


FIGURE 4.16 – Line profile of coupling function ($\eta(\boldsymbol{k})$) for 5 mrad, 10 mrad and 15 mrad aperture sizes.

4.4.2 Discussion

The recovered images in Fig. 4.14d to Fig. 4.14f show that the region, over which the update function recovers a good reconstruction of the object, increases as the aperture size increases. This correlates with the fact that the illumination size at the specimen plane increases with aperture size. As a result, the update function recovers a better representation of the specimen over a larger area in the case where a 15 mrad aperture is employed compared to the case where a 5 mrad aperture is used. The error plots in Fig. 4.15 show a general trend of decreasing error as the lens aperture size increases. Furthermore, it shows that after a fixed number of iterations, the average error per pixel is smaller for larger aperture sizes. This means that the convergence rate of the algorithm increases with aperture size despite the fact that the update function recovers more object information. This suggests that the number of interfering spatial frequencies in the Type-II setup provides a good handle on the convergence rate of the update function.

In Chapter 3.4, it was conjectured that the fidelity of the information recovered at any spatial frequency of the recovered object is directly linked to the coupling function of the algorithm, where the coupling function equals the FFT of the illumination intensity. The algorithm updates the specimen transmission function in real space by adding the feedback error $\varepsilon_n(\boldsymbol{r})$ to the object that is derived from the exit wave difference $\Delta\psi_n(\boldsymbol{r})$. Since the exit wave difference implicitly contains the illumination function as a factor, the ideal procedure would be to factor out this component by dividing by the illumination but this could result in division by zeros. So the route taken by the update function is suppression of the phase contribution from the illumination function by multiplying the exit wave difference with the conjugate of the illumination $P^*(\boldsymbol{r})$. This means the illumination intensity multiplies the error ($\Delta O_n(\boldsymbol{r})$) added to the object, i.e. $\varepsilon_n(\boldsymbol{r}) \approx |P(\boldsymbol{r})|^2 \Delta O_n(\boldsymbol{r})$. Since the multiplicative

update in real space is equivalent to a convolution update of the specimen spatial frequencies, the selectivity of the convolution process is governed by the width of the coupling function. In the case of a 5 mrad aperture calculation, the width of the coupling function is broader. This means that the update function simultaneously updates a wide range of adjacent points around the intended spatial frequency in the central point of the coupling function. The resulting coupling of adjacent spatial frequencies introduces low frequency noise in real space that is of the order of the illumination size. As a result, the recovered object in Fig. 4.14d contains a large area dominated by noise compared to Fig. 4.14f where the coupling function is more selective.

Furthermore, as the coupling function becomes delta-like in character, the update function should quickly arrive at the correct value of any spatial frequency of the recovered object; provided the corresponding convolution kernel does not simultaneously approach a delta function. Thus for the Type-II setup, as the coupling function tends to a delta function, the convergence rate of the algorithm should increase as well. Fig. 4.16 shows that the coupling function becomes more delta-like for a 15 mrad aperture illumination compared to a 5 mrad aperture illumination. This indicates that the general trend of the error metric in Fig. 4.15 can be explained with the framework of the coupling function. It is clear from Fig. 4.15 that the trend of the error metric with aperture size is not smooth. This is connected to the fact that the coupling function is structured. If the coupling were Gaussian in profile, the trend would be smoother. This is illustrated in the results of Section 4.5.

The information encoded in the intensities of the diffraction pattern comes from interference amongst several spatial frequencies of the specimen. In the Type-II setup, the interference strength amongst the specimen spatial frequencies is uniform because the convolution kernel (Fourier transform of the illumination) comprises uniform magnitude over the aperture (see Fig. 4.13a_1 to Fig. 4.13c_1). However the phase curvature over the convolution kernel

increases quadratically for a defocused illumination (see Fig. 4.13a_2 to Fig. 4.13c_2) because it is proportional to the defocus value as shown in Eq. (4.2). As a result, the contribution from spatial frequencies close to the edge of the aperture should eventually get masked as the aperture size increases. This can be related to the finite window constraint of the Fresnel propagator where regions outside the window with high quadratic phase gradient in the Fresnel kernel encode negligible information in the propagation integral. This means that the benefit of increasing aperture size should fall off and is evident by the convergence rate in Fig. 4.15, where the difference of error values after a fixed number of iterations approaches a constant value with increasing aperture size. Furthermore, the ability of the update function to extract object information should also map directly to the sampling rate of the phase curvature and thus the diffraction pattern.

An alternative explanation for the quality of the recovered images in Fig. 4.14 can be given by considering the diffraction pattern as a Gabor hologram. In this formalism, a transparent object, like the resolution target employed in these calculations, has a large zero-order component. Convolving the object spatial frequency representation with the Fourier transform of the illumination function generates the diffraction pattern. If we consider the generation of a single diffraction pattern intensity, the convolution process can be viewed as a summation of complex numbers that represent the object spatial frequencies. For diffraction intensities in the bright field, the convolution kernel encloses several object spatial frequencies including the zero-order component of the specimen. Since the zero-order for a highly transparent object is far higher than the remaining spatial frequencies, the resultant phasor in the bright field disc can be decomposed into two components, the zero-order phasor and a second phasor that represents the remaining spatial frequencies.

At different detector pixels within the bright field, the zero-order phasor scans the convolution kernel and as a result takes on the phase configuration of the convolution kernel.

Similarly, the second phasor encodes the inner product information between the kernel and the intersecting spatial frequencies of the object. The diffraction pattern over the bright field region is given by the squared magnitude of sum of these two phasors. Since the zero-order phasor has a higher magnitude, it dominates the summation so that the global magnitude and phase variation over the bright field take on values close to the convolution kernel. As a result, the illumination acts as a reference. The exact departure of magnitude and phase from the reference encodes information about the object's spatial frequencies. In the diffractive imaging experiment, we have the diffraction pattern intensities of the bright field disc and can thus infer the magnitude variations. As for the phase variations, knowledge of the illumination function provides an initial phase configuration for the detector plane guess wave distribution ($\psi_{n=1}(\boldsymbol{u})$) that is very close to the actual phase distribution within the bright field of the diffraction pattern.

Employing a large lens aperture generates a diffraction pattern with a large bright field, where the initial guess of the phase configuration is close to the correct values. Consequently, one would expect the reconstruction to approach the solution at a rate that is directly connected to the aperture size. The Gabor hologram explanation gives an insightful picture into the formation of the diffraction pattern and suggests the general trend of information expression in the diffraction pattern that maps to the convergence of the algorithm. It thus provides a suitable foundation for a detailed explanation of the trends described earlier in this section that are specific to the update function.

4.5 Impact of defocus on the update function

This section investigates the effect of increasing defocus where the sampling pitch of the detector is fixed over the entire calculations; this is useful for optimising the FoV in the Type-II diffractive imaging calculations. In the previous section, we saw that the quality of the reconstructions improve with a larger lens aperture because of an increase of the illumination spot size at the specimen plane. In the Type-II diffractive imaging experiments, the effective illumination/probe size at the specimen plane is also controlled by the defocus parameter. In the Type-II setup, having a large defocused beam over a finite sample generates a hologram of the sample within the bright field region of the diffraction pattern. Fig. 4.18a to Fig. 4.18c shows the calculated exit waves from a resolution target illuminated by beams with defoci ranging from 400 nm to 850 nm. The corresponding diffraction patterns are shown in Fig. 4.18d to Fig. 4.18f. The diffraction patterns show that the size of features in the hologram decrease with increasing defocus and as a result, the sampling requirement of the hologram becomes more stringent as the defocus value increases. In real space this corresponds to an increase of the illumination size with defocus.

The sampling requirement for the diffraction pattern is directly related to the sampling of the phase curvature of the convolution kernel (Fourier transform of the illumination) for the Type-II setup. As the defocus of the illumination increases, the local phase curvature of the convolution kernel also increases. The convolution kernel is given by

$$A(\theta) = \exp\left(\frac{2\pi i}{\lambda}\left(\frac{1}{2}\Delta z \cdot \theta^2\right)\right) \tag{4.2}$$

where the defocus parameter multiplies the quadratic phase of the wavefield at the back focal plane of the lens, θ measures the radial distance from the centre of the lens aperture and $\theta_{max} = 15\, mrad$ in these calculations. The back focal plane of the lens has a one to one

correspondence with the detector plane in the Type-II setup because both planes have a Fourier transform relationship with the specimen plane. As the defocus value increases, the size of the illumination increases and eventually results in undersampling of the diffraction pattern, given a fixed detector pixel pitch. The line of Fig. 4.17 that corresponds to the defocus value Δz_3 illustrates a defocus value that results in undersampling of the corresponding diffraction pattern.

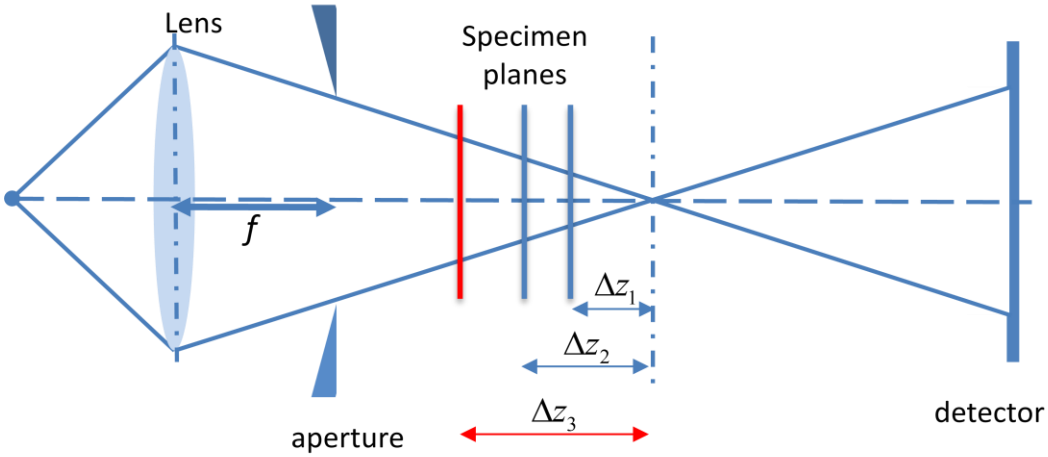


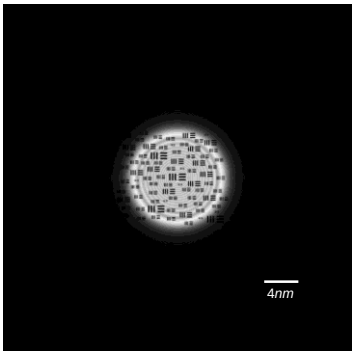
FIGURE 4.17 – illustration of several defoci planes. Defocus values greater than Δz_3 result in severe aliasing of the illumination where the high intensity values from adjacent Nyquist cells overlap with the model illumination function.

As the defocus increases, the local curvature of the illumination at the specimen plane decreases and the effective probe size at the specimen plane is determined by region of high illumination intensity. Using geometric optics, this is given by

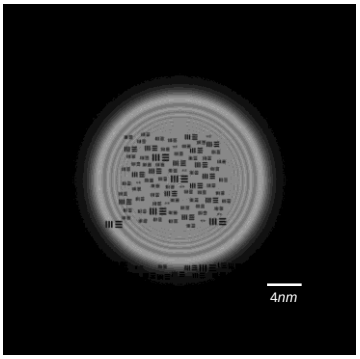
$$P_{size} = 2\alpha\Delta z \tag{4.3}$$

where Δz is the defocus of the illumination; α represents the aperture size of the lens and equals 15 mrad in the calculations of this section. The effective size of the probe defines a spot at the specimen plane that most of the beam energy traverses during the model

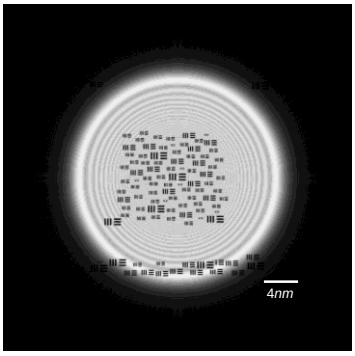
experiment and, thus, defines the diffracting region in these model experiments. This is just an estimate since we know the illumination function extends further than the effective probe size because a band limited aperture in Fourier space creates the illumination; thus the illumination has an infinite extent in the specimen plane. Consequently, there is always some element of aliasing in any calculation that models the Type-II setup.



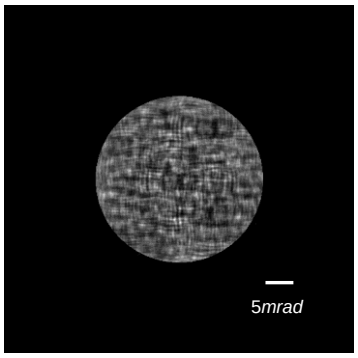
(a) Magnitude of exit wave from 400nm defocus probe.



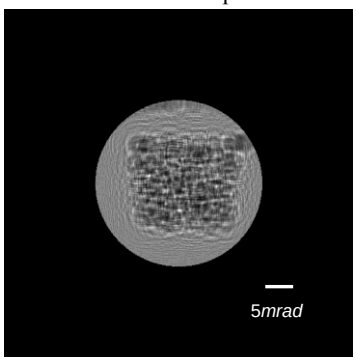
(b) Magnitude of exit wave from 700nm defocus probe.



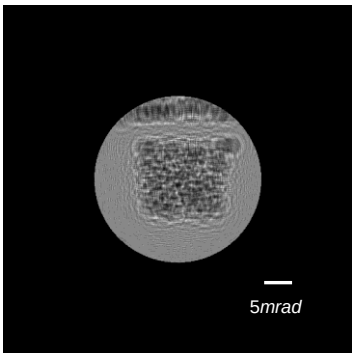
(c) Magnitude of exit wave from 850nm defocus probe.



(d) Diffraction pattern from calculation using 400nm defocus probe.



(e) Diffraction pattern from calculation using 700nm defocus probe.



(f) Diffraction pattern from calculation using 850nm defocus probe.

FIGURE 4.18 – (a) to(c) represent the Magnitude of exit waves from the specimen illuminated by a probe of defocus values ranging from 400 nm to 850 nm. (d) to (f) represents the corresponding diffraction patterns.

The effect of aliasing can be minimised in the forward calculations that generates the diffraction patterns, by employing a calculation window that is twice the size of the calculation window used in the iterative calculations. Furthermore, the sizes of the

diffraction patterns shown in Fig. 4.18d to Fig. 4.18f span one quarter of the respective calculated diffraction patterns. This is done to ensure that the diffraction patterns used in the calculations model a linear convolution system rather than cyclic convolution system that one naturally gets out of the FFT algorithm. The specimen transmission function is convolved with the atomic potential of carbon. Although this is unrealistic for a resolution target type specimen, it helps limit the FT of the object. The convolution is done with a sampling pitch of 0.01nm and sets the limit of the angular span of the calculated diffraction patterns. These procedures should also allow conclusions that are drawn from these calculations to be in close agreement with experimental diffraction patterns.

The detector modelled in these calculations has a pitch of 0.12 mrad and is made up of 512x512 pixels. This gives a sampling pitch of 0.04 nm at the specimen plane from the detector semi-angle of 62.3 mrad and a wavelength of 0.0025 nm (corresponding to a 200 keV electron beam). The resulting Nyquist window in real space has a linear dimension of $\sim$20.5 nm. Visual inspection of the illumination function used in the iterative calculations showed that the maximum defocus, above which the effect of aliasing becomes serious, corresponds to a defocus of 600 nm, giving an effective probe size of $\sim$18 nm. The calculations in this section were performed with defocus values ranging from 200 nm to 1200 nm. The top end of the defocus range results in undersampling of the diffraction pattern. This was done to model a set of experiments that investigates the connection between the sampling and the quality of the results provided by the update function. An extension of this investigation in the case of ptychography is presented in Section 6.7. The diffraction patterns used in all calculations in this section contain 10^8 counts.

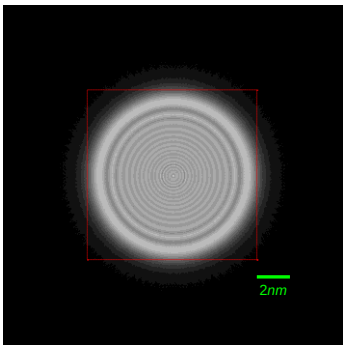
4.5.1 Results

This investigation employed the resolution target of Chapter 3 as a test specimen and the initial estimate to the specimen was free-space with uniform transmission value of unity. The illumination functions used in these calculations model a 200 keV electron beam with a lens aperture of 15 mrad. Fig. 4.19d to Fig. 4.19f show the recovered images for calculations using defocus values of 350nm, 500nm and 650nm respectively. The regions shown correspond to the areas delineated by the red lines in Fig. 4.19a to Fig. 4.19c. Fig. 4.20 shows the results for similar calculations using defocus values of 750nm, 900nm and 1200nm. The error plots for these calculations are shown in Fig. 4.21.

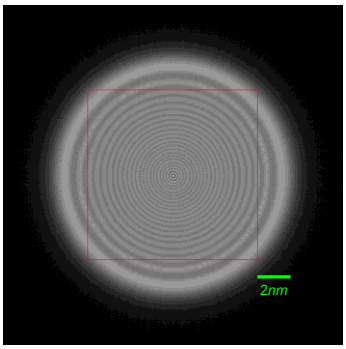
The results shown in Fig. 4.23 correspond to calculations where the diffraction patterns were artificially upsampled using nearest neighbour interpolation. In this interpolation scheme, the intensities of the original diffraction pattern are assigned to 3 adjacent pixels of the upsampled detector. This is illustrated in Fig. 4.22. This was done to prevent the severe aliasing effect in the illumination as evident in the images of Fig. 4.20a to Fig. 4.20c.

Fig. 4.23d to Fig. 4.23f show the recovered images using the interpolated diffraction patterns in the region delineated by the dashed lines in Fig. 4.23a to Fig. 4.23c. The images in Fig. 4.23g to Fig. 4.23i represent larger FoV cutouts of the reconstruction. The blue lines in Fig. 4.23a to Fig. 4.23c delineate these FoV. The error plots for the interpolated calculations are shown in Fig. 4.25a. The error plots shown in Fig 4.25b offer an alternative trend for the error metric (compared to Fig. 4.21) in which errors from interpolated calculations replace the calculations that were compromised by the effect of aliasing. Alternate interpolation schemes that average the diffraction intensities in the interpolated points result in blurring of the high-resolution information in the specimen as shown in Fig. 4.24a and Fig. 4.24b. The result in Fig. 4.24c was generated with the floating interpolation scheme, where the

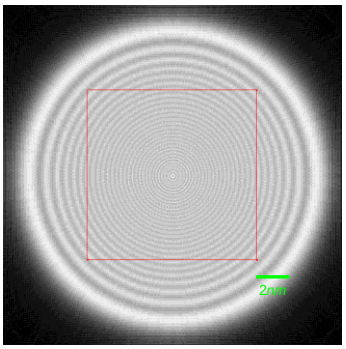
diffraction intensity in pixels represented by the black boxes in Fig. 4.22 is recovered by the algorithm during the iterative calculations; this calculation scheme is further examined in Chapter 6. The algorithm recovers the complex wavefield in these pixels by retaining the values calculated in the forward calculation while replacing the intensity values in pixel (represented by white box) by the data. This scheme is further investigated in Chapter 6. Fig. 4.26 and Fig. 4.27 show line profiles of the coupling function for these calculations. These correspond to the inner product of the convolution kernel and its conjugate along the k_x-direction.



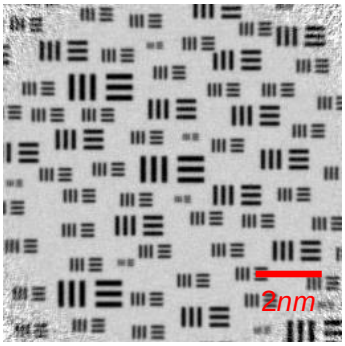
(a) Magnitude of 350 nm defocus illumination, calculated using the fixed sampling pitch of the detector



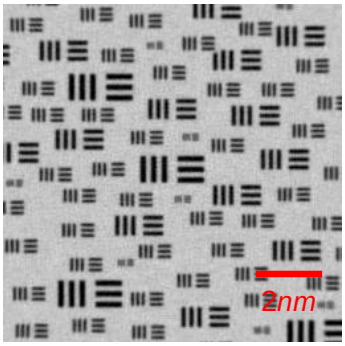
(b) Magnitude of 500 nm defocus illumination, calculated using the fixed sampling pitch of the detector



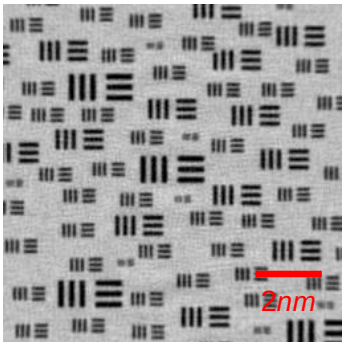
(c) Magnitude of 650 nm defocus illumination, calculated using the fixed sampling pitch of the detector



(d) Recovered image from 350 nm probe calculation

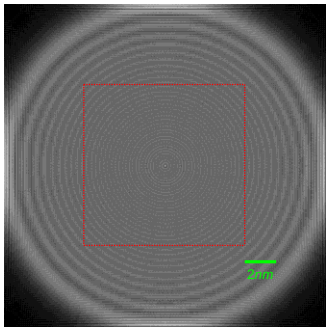


(e) Recovered image from 500 nm probe calculation

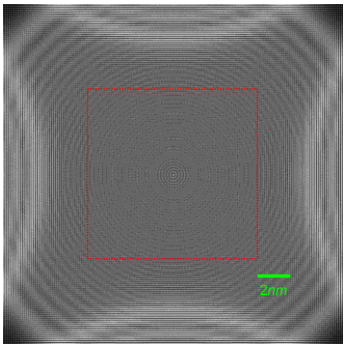


(f) Recovered image from 650 nm probe calculation

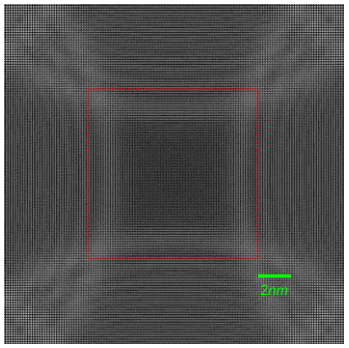
FIGURE 4.19 – (a) to (c) show the probe at the specimen plane for defocus values in the range of 350 nm to 650 nm. (d) - (f) show the recovered object for the region enclosed by the respective red boxes of (a) to (c).



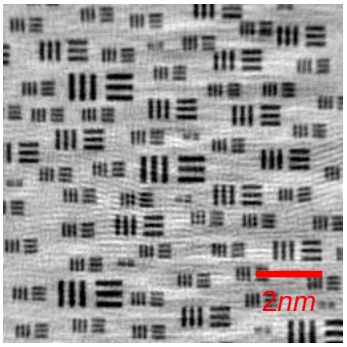
(a) Magnitude of 750 nm defocus illumination, calculated using limited sampling in diffraction space



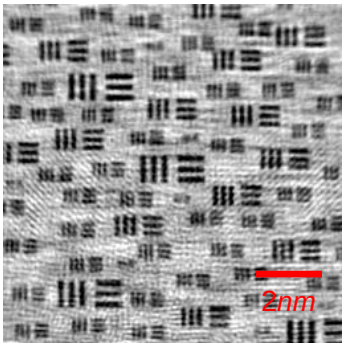
(b) Magnitude of 900 nm defocus illumination, calculated using limited sampling in diffraction space



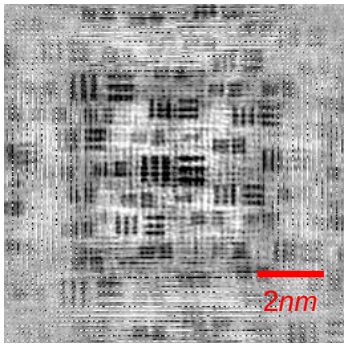
(c) Magnitude of 1200 nm defocus illumination, calculated using limited sampling in diffraction space



(d) Recovered image from 750 nm probe calculation



(e) Recovered image from 900 nm probe calculation



(f) Recovered image from 1200 nm probe calculation

FIGURE 4.20 – (a) to (c) show the magnitude of the illumination at the specimen plane for defocus values in the range of 750 nm to 1200 nm, where the window size is determined by the sampling of the diffraction pattern (d)- (f) show the recovered object for the region enclosed by the respective red boxes of (a) to (c).

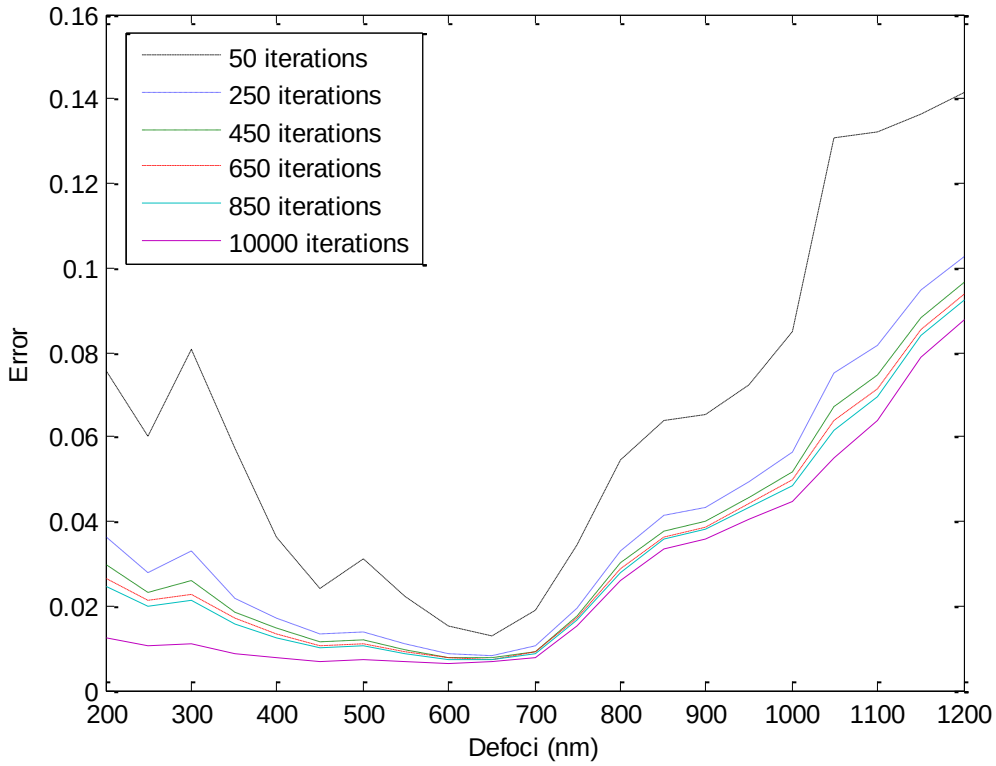


FIGURE 4.21 – Trend of the error metric with defoci. Each line plot connects the error values of different defocus calculations after a fixed number of iterations.

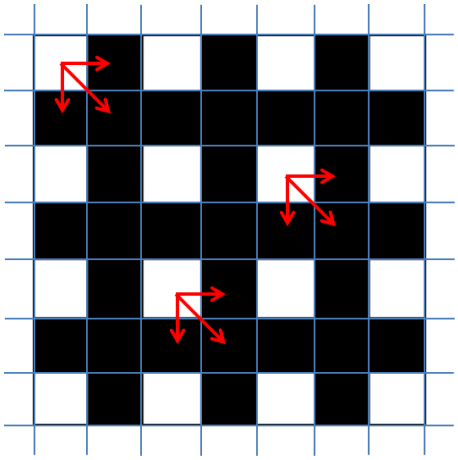
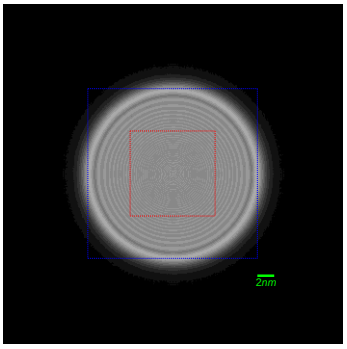
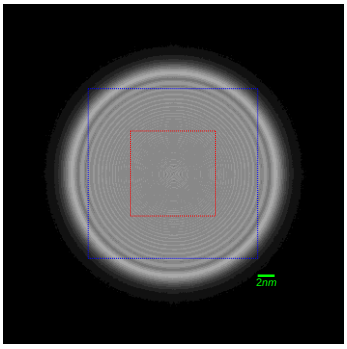


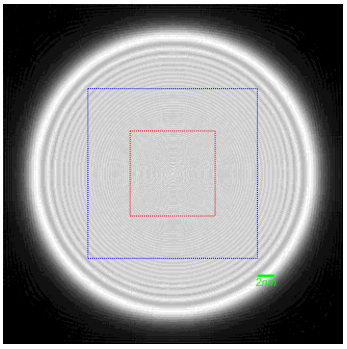
FIGURE 4.22 – illustration of nearest neighbour interpolation scheme. The white boxes correspond to pixels with the correct sampled diffraction pattern intensity values and the black boxes correspond to pixels in between the sampling points. These regions are updated with measured values of the diffraction intensities. The red arrows illustrate the mapping process where the intensity in one pixel is assigned to three neighbouring pixels.



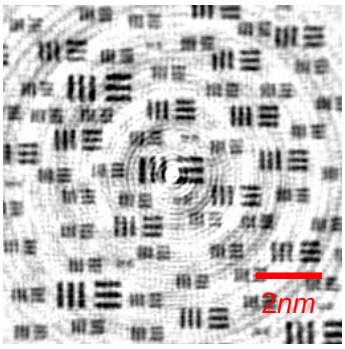
(a) Magnitude of 750 nm defocus illumination, calculated using increased sampling of diffraction space



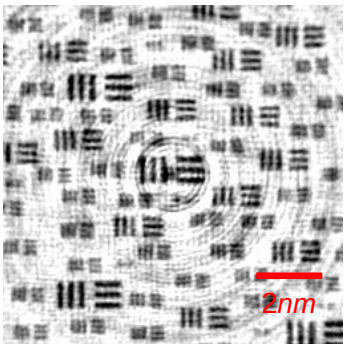
(b) Magnitude of 900 nm defocus illumination, calculated using increased sampling of diffraction space



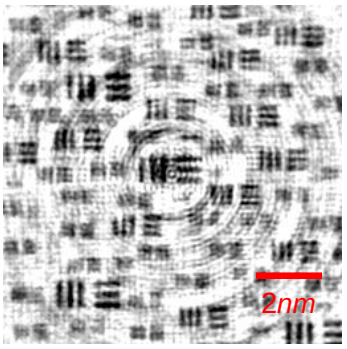
(c) Magnitude of 1200 nm defocus illumination, calculated using increased sampling of diffraction space



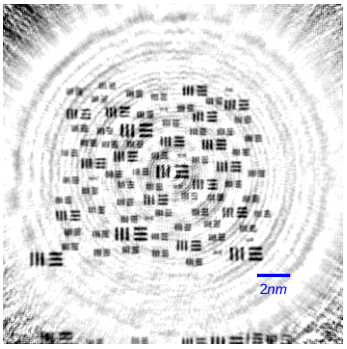
(d) Recovered image from 750 nm probe calculation



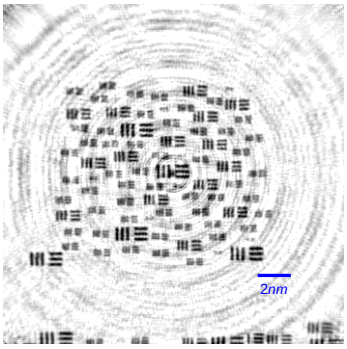
(e) Recovered image from 900 nm probe calculation



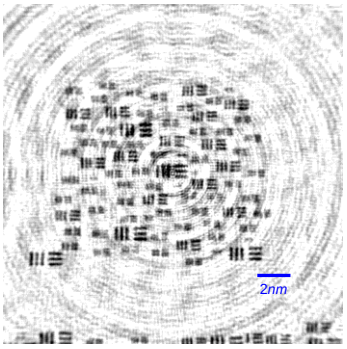
(f) Recovered image from 1200 nm probe calculation



(g) Larger FoV of recovered image from 750 nm probe calculation

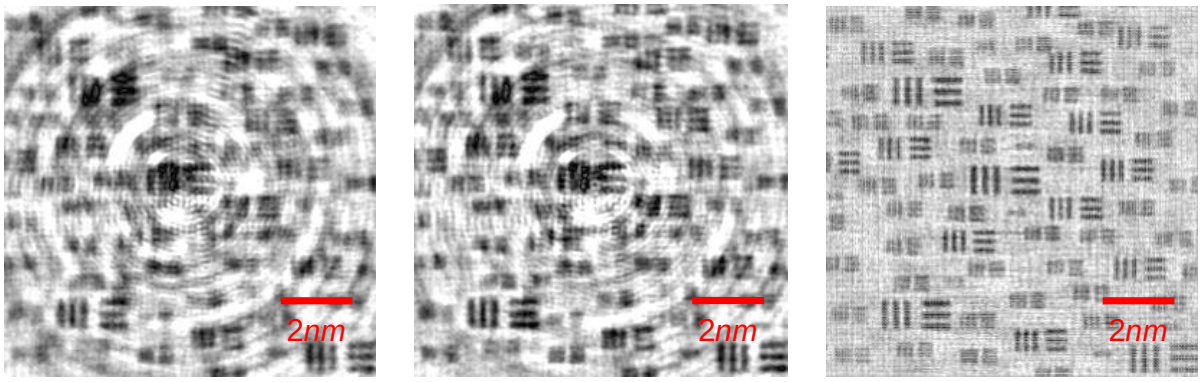


(h) Larger FoV of recovered image from 900 nm probe calculation



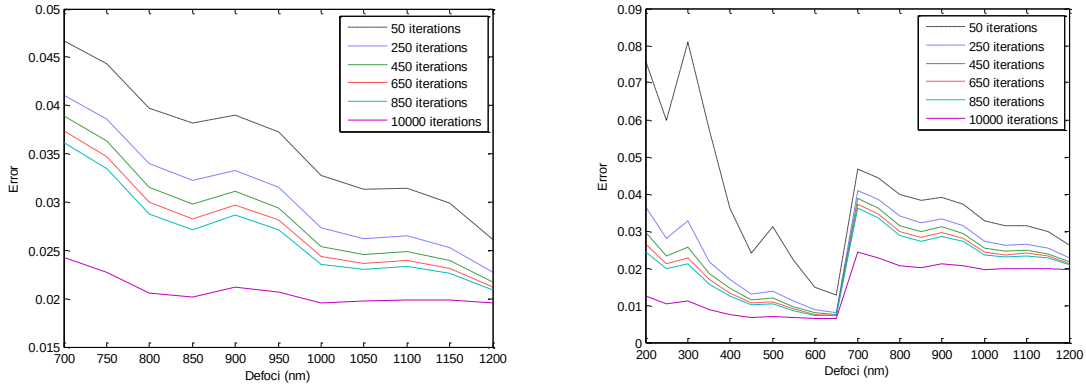
(i) Larger FoV of recovered image from 1200 nm probe calculation

FIGURE 4.23 - (a) - (c) show the magnitude of the illumination used in the calculations where the diffraction pattern sampling was artificially increased to avoid aliasing of the illumination. (d) – (f) show the recovered transmission function in the region delineated by the inner red dashed lines of (a) - (c). The images in (g) – (i) represent the recovered transmission function in the region delineated by the outer blue dashed lines of (a) - (c).



(a) (b) (c)

FIGURE 4.24 – Transmission function of recovered objects in calculation that use a 1200 nm defocus illumination and different interpolation schemes for the diffraction patterns (a) Bilinear interpolation scheme (b) Bicubic interpolation scheme (c) Result calculated using the floating interpolation scheme where the algorithm solves for the missing intensity values in pixels represented by the black boxes in Fig. 4.22. This calculation extends the phase recovery of the diffraction pattern to include region of missing intensities; it does so by retaining the complex values in regions of missing intensity values. These results show that the bars of the recovered object are not well resolved in (a) and (b) compared with the result in Fig. 4.23f.



(a) (b)

FIGURE 4.25 - (a) Error trend with defoci for calculations where the diffraction pattern is interpolated to avoid aliasing. (b) Error trend with defoci where the error values in defoci range of 200 nm to 650 nm correspond to direct calculation and the error values for defoci greater than 650 nm correspond to the case where the diffraction pattern is interpolated as in (a). The jump in the error plot at 700 nm corresponds to the case where the diffraction patterns are upsampled using nearest neighbour interpolation.

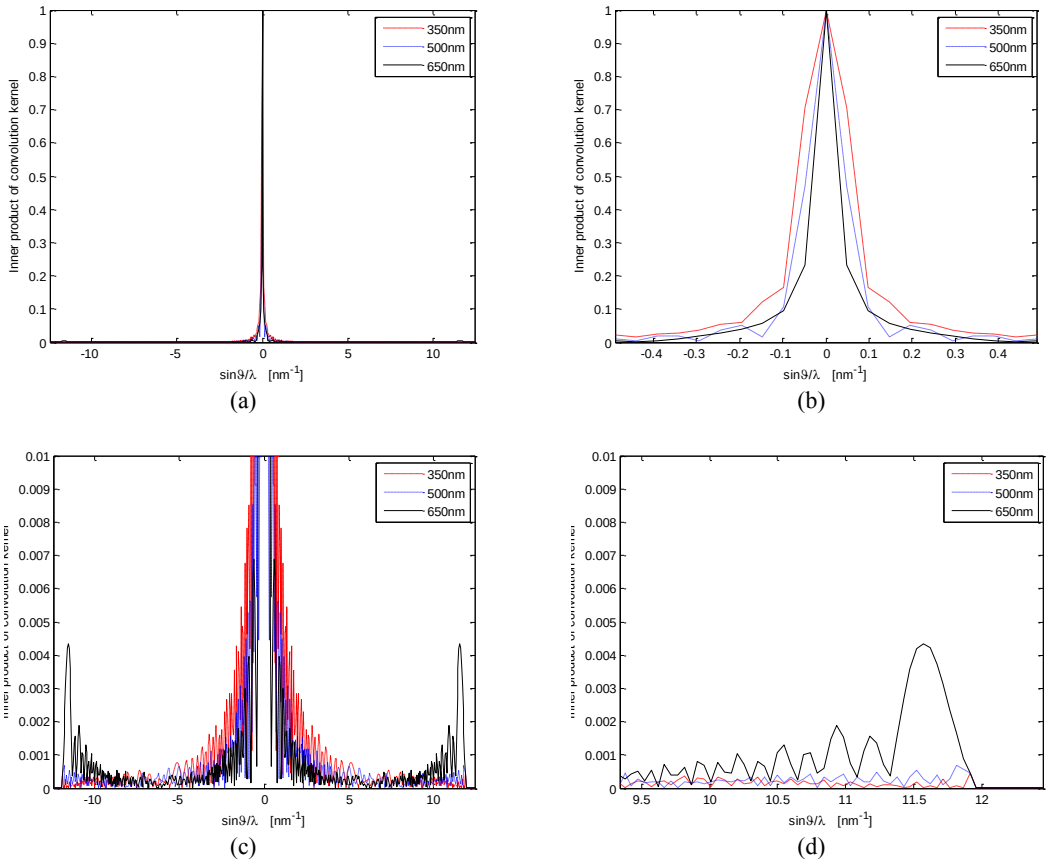


FIGURE 4.26 – Line profile of the coupling function ($\eta(\boldsymbol{k})$) for defocus values of 350nm, 500nm and 650nm. (a) Plot represents the full range of the cyclic convolution that corresponds to k_x-direction slice of the FFT of the illumination intensity. (b) Central region of the coupling function illustrates the departure of inner product from a delta function for small spatial frequencies separations. (c) Height range limited to illustrate the presence of multiple peaks at large spatial frequencies separations due to aliasing. (d) Zoomed view of interference amongst large spatial frequencies separations to illustrate that the magnitude of the peaks increases with increasing defoci as the illumination size crosses the Nyquist limit.

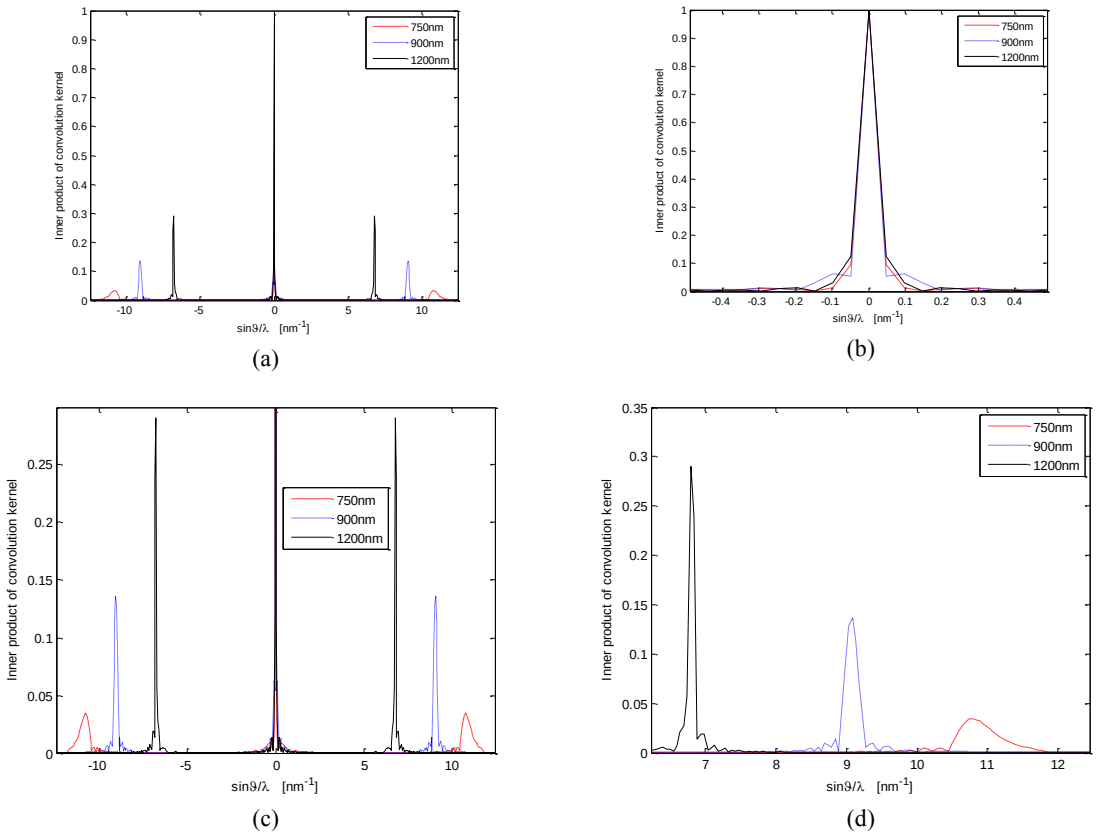


FIGURE 4.27 – Line profile of the coupling function ($\eta(\boldsymbol{k})$) for defoci values of 750nm, 900nm and 1200nm. (a) Plot represents the full range of the cyclic convolution that corresponds to k_x-direction slice of the FFT of the illumination intensity. (b) Central region of the coupling function illustrates the departure of inner product from a delta function for small spatial frequencies separations. (c) Height range of the coupling function limited to emphasise the relative magnitude of the peaks for different defoci. The plot shows that the off-axis vertices of different defocused illuminations peak at different spatial frequency separations due to aliasing. (d) Zoomed view to illustrate that the peaks move from the large spatial frequencies separations to low frequencies separations with a corresponding increase in height.

4.5.2 Discussion

The trend of the error metric in Fig. 4.21 indicates that the optimum defocus value in the Type-II setup corresponds to the case where the diffraction pattern is sampled at the Nyquist pitch. This corresponds to a value of ~600 nm for the calculations performed in this section. The convergence rate of the error metric increases with defocus because the coupling function tends to a delta-like function with increasing defocus (see Fig. 4.26b). Hence the feedback errors added to any spatial frequency quickly approach the limiting value. Fig. 4.26b shows that the coupling function is relatively smooth for 350 nm and 650 nm but contains additional structure at 500 nm defocus, thus explaining the jump in the error trend at 500 nm, i.e. the trend would be smooth if the coupling function smoothly becomes delta-like without additional structural changes like the dips at defocus of 500 nm.

The images in Fig. 4.19a and Fig. 4.19b show that the update function faithfully recovers the specimen function in regions where the illumination has high intensities for calculations where the effective size of the illumination is smaller than that defined by the Nyquist limit. However, the recovered specimen functions for defoci greater than 600nm (see Fig. 4.19c and Fig. 4.20d to Fig. 4.20f) contain speckle-like variations in the background. These are artefacts in the reconstruction. The speckles sizes increase and become more pronounced as the illumination size increases above 650nm (see Fig. 4.19f and Fig. 4.20d) and becomes more aliased.

The line profile of the coupling functions in Fig. 4.26 and Fig. 4.27 provides a measure for the amount of aliasing present in these calculations, where the amount of aliasing is quantified by the magnitude of additional peaks. Analysis of the coupling function for defocus values across the Nyquist boundary reveal that as the illumination size increases

beyond the Nyquist limit, additional peaks arise at high angles and move towards smaller angles with a corresponding rise in magnitude. This is illustrated in Fig. 4.27d.

The emergence of background speckle in these calculations correlates with the magnitude and positions of the additional peaks of the coupling function, i.e. smaller and less pronounced speckle size corresponds to situations where additional peaks of low magnitude in the coupling function occur at large angles (see Fig. 4.26d) while bigger and more pronounced speckles correspond the situations where additional peaks of coupling function occur at low angles with high magnitude (see Fig. 4.27d).

To explain the origin of the speckles, recall that in Section 3.4 it was mentioned that the selectivity of the update function depends on the form of the coupling function ($\eta(\boldsymbol{k})$). Thus as the algorithm updates a given spatial frequency of the specimen (corresponding to the central peak of $\eta(\boldsymbol{k})$), it simultaneously but incorrectly updates other spatial frequencies where $\eta(\boldsymbol{k})$ contains addition peaks. This suggests that the distance between the central peak and additional peaks should relate to the scale of inconsistencies (speckles) introduced into the iterative calculations.

The outer region of the image in Fig. 4.20f shows that the update function fails to recover the specimen function in places where the intensities of the illumination contain a large contribution from parts of the probe that wrap around the calculation window due to aliasing. This effect of aliasing is minimised by artificially upsampling the diffraction pattern to increase the physical dimensions of the calculation window in the specimen plane.

Fig. 4.23f shows that the recovered image from the interpolated diffraction pattern gives a better representation of the specimen (for a defocus value of 1200 nm) compared to the calculation where the aliased illumination was used directly as shown in Fig. 4.20. Fig. 4.23 show that the update function recovers most of the specimen features without similar

speckles associated with the reconstructions of Fig. 4.20d to Fig. 4.20f. This supports the assertion that the speckles present in the recovered objects in Fig. 4.20d to Fig. 4.20f result from the way the diffraction patterns were processed (i.e. calculation schemes where $\eta(\boldsymbol{k})$ has a single peak avoid the emergence of speckle). It is worth noting that the recovered objects in Fig. 4.23d to Fig. 4.23f contain imprints of the illumination. These imprints may come from the interpolation scheme used in the iterative calculations because iterative calculations that use adequately sampled diffraction patterns do not have these imprints. The recovered objects in Fig. 4.24a and Fig. 4.24b show that the interpolation scheme in Fig. 4.22, where the interpolated intensity values do not come from averaging of adjacent data points, provides a better representation of the test object. Also the reconstruction quality in Fig. 4.24c suggests that the scheme where the algorithm recovers the complex values provides a good reconstruction of the object without imprints of the illumination. This scheme is further explored in Chapter 6 where the redundancy of ptychography is used to improve the quality of the reconstruction.

Fig. 4.25a shows that the convergence rate of the error metric in the upsampled diffraction pattern calculation increases with defocus despite the fact that the diffraction pattern gets highly undersampled. This trend is similar to the case of defocus values below the Nyquist limit and supports the conjecture that the convergence rate is directly connected to the delta-like response of the coupling function. The jump in the error plot of Fig. 4.25b arises because diffraction patterns from defoci values $\geq$ 700nm were upsampled to minimise the effect of aliasing in the iterative calculations. As a result, the update function processed diffraction patterns that contain three-quarters incorrect information.

4.6 Impact of defocus error on the PIE update function

The previous section showed that accurate knowledge of the illumination function facilitates diffractive imaging from a region as large as the Nyquist limit in the Type-II setup. In this section we investigate the effect of utilising an incorrect illumination in the iterative calculations. Estimating the correct illumination profile at the specimen plane is a crucial part of diffractive imaging experiments because the type of interference fringes generated in the diffraction pattern depends on the illumination. Unfortunately, the parameters needed for an accurate model of the illumination are not readily available. Even in cases where these parameters are available, for example electron experiments done in a Transmission Electron Microscope (TEM), the nominal defocus value given by the instrument differs (slightly or sometimes by a great amount) from the correct value that represents the illumination distribution at the specimen plane. Since the main parameter that controls the size of the illumination in the Type-II setup is the defocus, this section investigates the impact of using an illumination with incorrect defocus value on the quality of the result recovered by the update function.

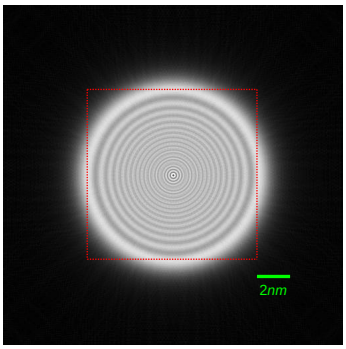
The algorithm requires the correct illumination function for two crucial steps in the update function. In the first step, the illumination generates an exit wave that provides an estimate of the diffraction patterns in the forward calculations; these estimates are used to compute the appropriate error function in the detector plane. In the second step, the update function utilises the conjugate of the illumination to map the feedback error function to the appropriate object domain function, in order to correct the running estimate of the object. This suggests that using an incorrect illumination function in the algorithm may completely ruin the quality of the reconstruction. To investigate the extent that the update equation can cope with this effect, the calculations in this section were carried out using nominal defocus values ranging

from 300 nm to 550 nm. In each of these calculations, the algorithm use illumination with incorrect defocus values that span an error range of ± 10 % of the correct defocus value.

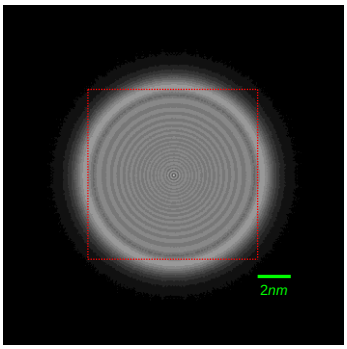
4.6.1 Results

This investigation employed the resolution target of Chapter 3 as a test specimen and the initial estimate to the specimen is free-space with uniform transmission value of unity. The illumination functions used in these calculations model a 200 keV electron focus beam made by a lens with a 15 mrad semi-angle aperture. The detector modelled in these calculations has a pitch of 0.12mrad and was made up of 512x512 pixels. This gives a sampling of 0.04 nm at the specimen plane from the detector semi-angle of 62.3 mrad and a wavelength of 0.0025 nm corresponding to a 200 keV electron beam.

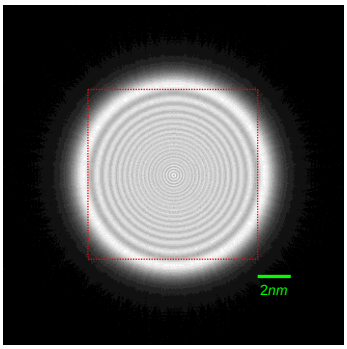
The results shown in this section correspond to the case where a defocus value of 400 nm is used to generate the diffraction patterns. Fig. 4.28d to Fig. 4.28f show the recovered objects from calculations that employ illuminations with defocus values of 400 nm, 412 nm and 420 nm. These correspond to defoci error of 0%, 3 % and 5%. Fig. 4.29 shows the trend of the error metric after 10000 iterations with defoci error ranging from -10% to +10%.



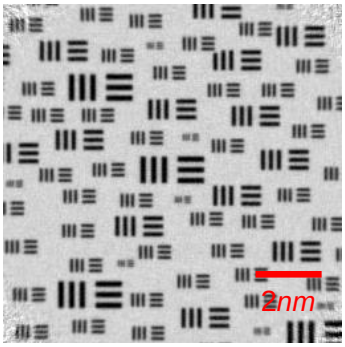
(a) Magnitude of Illumination function at the specimen plane made by a lens with a defocus error of 0%. This corresponds to a defocus of 400 nm.



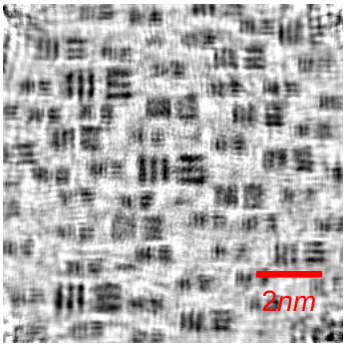
(b) Magnitude of Illumination function at the specimen plane made by a lens with a defocus error of 3%. This corresponds to a defocus of 412 nm.



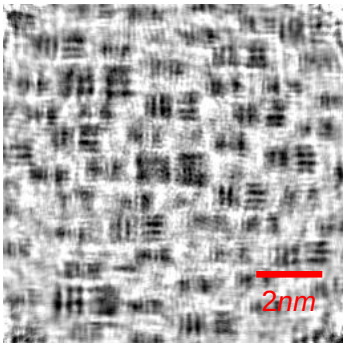
(c) Magnitude of Illumination function at the specimen plane made by a lens with a defocus error of 5%. This corresponds to a defocus of 420 nm.



(d) Recovered image from calculations using 0% defocus error



(e) Recovered image from calculations using 3% defocus error



(f) Recovered image from calculations using 5% defocus error

FIGURE 4.28 – (a) to (c) show the magnitude of the probes at the specimen plane that correspond to errors of 0%, 3% and 5% of the correct defocus value. This gives a range of 400nm to 420nm for the case where the correct defocus value equals 400 nm. (d) - (f) show the recovered object for the region enclosed by the respective red boxes of (a) to (c).

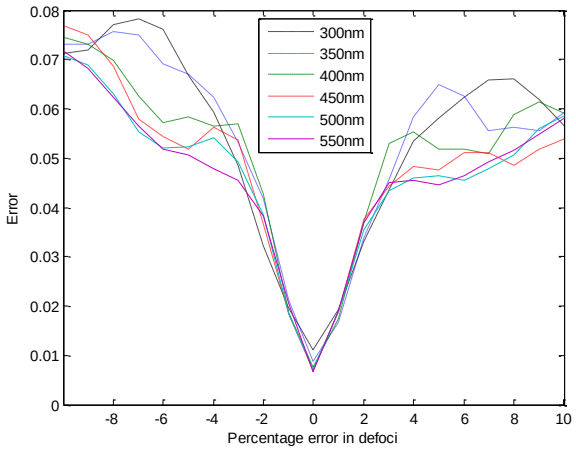


FIGURE 4.29 - Trend of error metric for with percentage error in defoci value. Each line plot connects error values for different defoci calculations after 10000 iterations.

4.6.2 Discussion

The results show that a slight change in defoci severely degrades the quality of the recovered object. The trend of the error metric in Fig. 4.29 suggests that error increases rapidly with small change in defoci until about 3% after which the error fails to accurately quantify the distortions in the recovered image.

In the Type-II experiment, the size of features in the diffraction patterns decrease with increasing defoci. Therefore, calculations that employ a probe with defocus of 412 nm and 420 nm generate interference fringes that are smaller than those of the data (diffraction patterns generated by a 400 nm defocused illumination). As a result, using an incorrect defocus value in the algorithm results in incompatible feature sizes between the estimated diffraction patterns and the data. In real space, this corresponds to incorrectly inferring the area of the specimen that contributes strongly to the diffraction pattern.

4.7 Summary

In this chapter, it is shown that applying the update function to a noisy diffraction pattern in the case where the illumination profile is known produces good quality reconstructions for modest count values, because knowledge of the illumination introduces information into the iterative calculations. The behaviour of the update function was also investigated with illumination formed by lens of different aperture sizes with the conclusion that large aperture illuminations produce better quality reconstructions and also have improved convergence properties. It was also shown that in the Type-II setup, the update function can reliably recover the transmission function of real objects for cases where the illumination fills the

calculation window. The coupling function was shown to explain the convergence property of the update function as well as the emergence of background variations in undersampled diffraction patterns calculations. Finally, it was shown that small error in the defocus of the illumination (less than 3%) resulted in a significant distortion of the recovered object.

Chapter 5

5 Impact of Experimental parameters on the ePIE Algorithm

In Chapter 4 it was shown that the quality of the recovered object degrades with even a small defocus error (~3%) when the update function processes a single diffraction pattern in the Type-II setup. In this Chapter, Section 5.2 and Section 5.3 present the investigation of the impact of incorrect characterisation of the defocus and astigmatism parameters on the ePIE algorithm. The ePIE algorithm applies the update function to multiple diffraction patterns and simultaneously solves for the object and illumination functions (see Eq. 2.23). In Section 4.2, it was shown that the update function recovered the object with varying resolution (over regions with high illumination intensities) in the case of low count single diffraction pattern calculations. Section 5.4 extends this investigation to the ePIE algorithm to see if a similar phenomenon occurs in multiple diffraction patterns calculations. Should the ePIE algorithm recover a uniform resolution reconstruction over the central part of the ptychographic calculations (that comprises overlapping probes), analysis of the result will clarify the means via which the ePIE algorithm extracts resolution information from multiple diffraction

patterns and consequently highlight the appropriate framework for investigating resolution improvement calculations with the ePIE algorithm, which is done in Chapter 6. Section 5.5 investigates the optimal data collection strategy in the Type-II setup, where either a large set of noisy diffraction patterns or a small set of high quality diffraction patterns make up the ptychographic dataset. This is particularly important in electron ptychographic experiments where drift of the specimen during data collection compromises the quality of the recorded diffraction patterns.

5.1 Impact of illumination parameters on ePIE

Work done at visible light wavelengths shows that structured illumination can be recovered from the ptychographic dataset when processed by the ePIE algorithm. The types of illumination employed in these experiments include wavefields generated from a pinhole placed in front of the specimen (but not coincident with the specimen) and wavefields where the aperture of the pinhole contains a diffuser, such as thin plastic films or spatial light modulator (SLM) encoded with a random phase pattern; these setups are shown in Fig. 5.1. In relation to the setup of Fig. 5.1a, Maiden and Rodenburg (2009) showed that the ePIE algorithm can recover the Fresnel fringes of the propagated aperture that result from the distance between the pinhole and the specimen planes. The ePIE algorithm has also been shown to recover the structured wavefield from diffusers in the setup shown in Fig. 5.1b, where the initial guess of the illumination is a pinhole aperture (Maiden, Rodenburg and Humphry, 2010b).

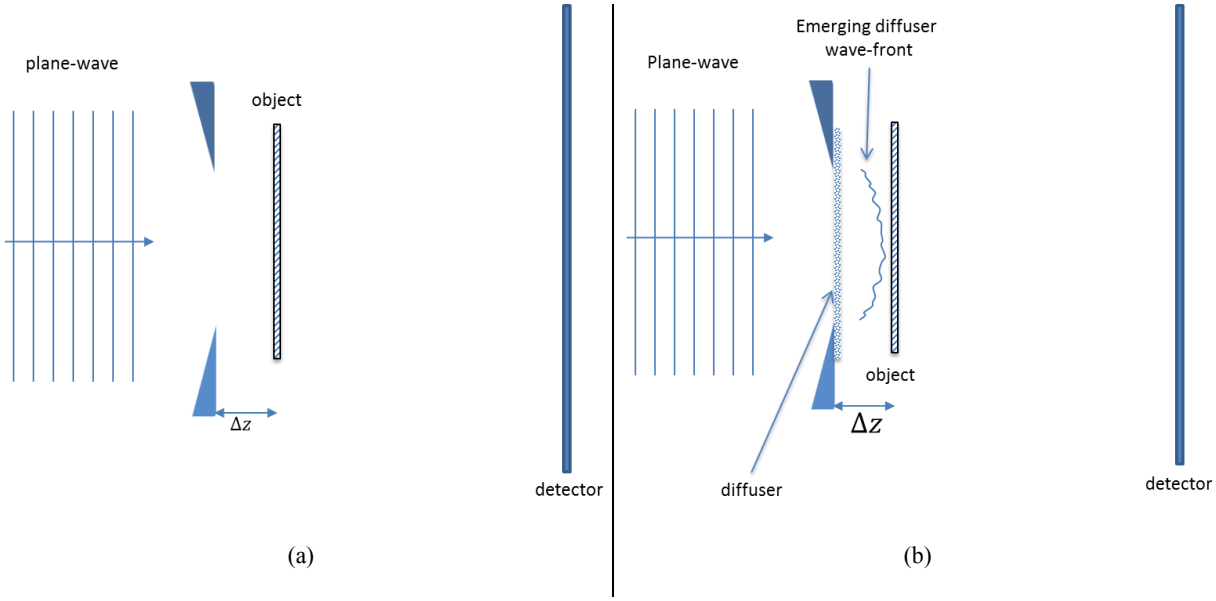


FIGURE 5.1 – (a) Experimental setup that employs the propagated wavefield from a pinhole aperture to limit the size of the diffracting region of the specimen. The ePIE algorithm recovers the specimen and illumination function at the plane of the specimen. The recovered illumination function contains Fresnel fringes due to the propagation distance (Δz). (b) Experimental setup that employs a diffuser placed in front of a pinhole. The ePIE algorithm also recovers the structure of the diffuser at the specimen plane together with the specimen function.

The ePIE algorithm should have a greater impact in short wavelength microscopy, X-rays and electrons, where accurate characterisation of the illumination function at the specimen plane is difficult. For example, in the case of the electron microscope the illumination can be characterized by several parameters, such as defocus, astigmatism and higher order aberrations; while the defocus parameter can be read off the instrument, the nominal value is not always accurate. Other aberrations can be measured using tilt-series tableau in TEM (Meyer, Kirkland and Saxton, 2004) or a set of Ronchigrams in STEM. The ePIE algorithm provides the advantage of automatically refining the illumination guess during the iterative calculations. This recovers an accurate representation of the illuminating wave, which subsequently improves the fidelity of high resolution information within the reconstructions. Section 5.2 and Section 5.3 investigate the amount of defocus and astigmatism error that the ePIE algorithm can robustly deal with in electron ptychographic experiments.

5.2 Impact of defocus error on ePIE

This section investigates the amount of error that the ePIE algorithm can resolve in relation to the defocus parameter of the Type-II illumination. The defocus error results from an incorrect estimation of the defocus parameter, which is used to characterise the illumination function employed by the ePIE algorithm. The defocus error of the illumination in the n^{th} iteration results in a mismatch between feature sizes of the n^{th} estimated diffraction pattern from the algorithm and the corresponding ptychographic data. It is expected that these discrepancies provide the ePIE algorithm with the necessary information needed to update the illumination.

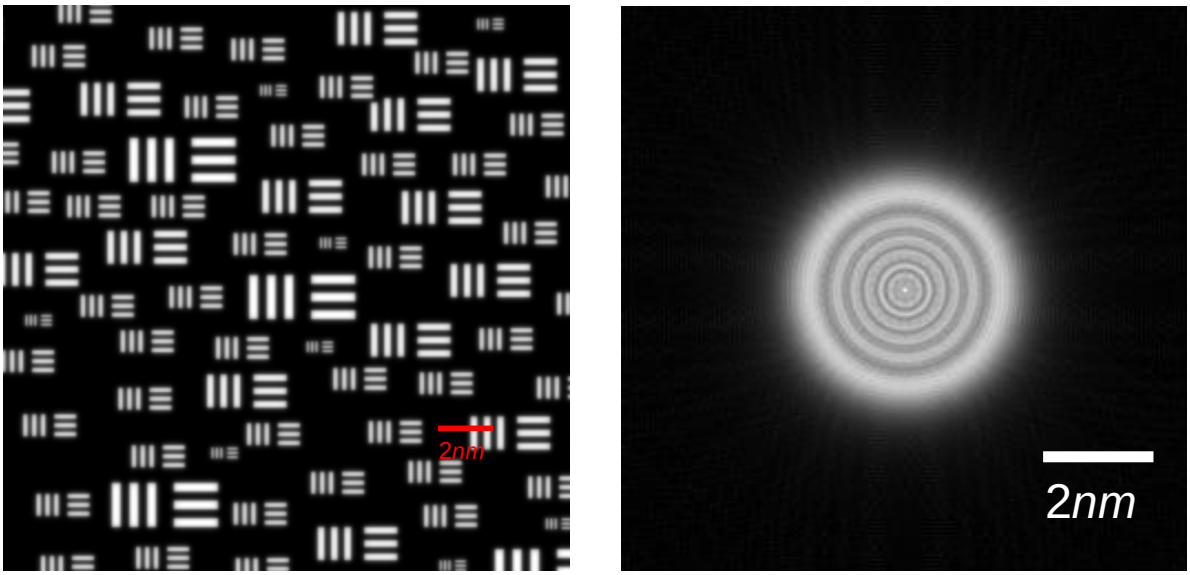
In electron diffractive imaging experiments, a curved illumination is used to limit the region of the object that contributes to a diffraction pattern (see the Type-II setup). This implementation is favoured over placing an aperture coincident with the specimen because the diffracting region required is too small for physical stops to be practical due to stringent requirements on fabrication (high aspect ratio pinhole) and also contamination at the edge of the pinhole, as pointed out earlier. The size of the illumination spot at the specimen plane is controlled by the defocus parameter and the condenser aperture. Here we consider only the defocus parameter, which measures the distance from the focus of the lens to the specimen plane, along the z-direction. In this case, the FoV of single diffraction pattern calculations is determined by the defocus, which defines the effective size of the illumination. However this is not the case in ptychographic calculations because the illumination is shifted to several positions. Consequently, the regions where the shifted illumination has high intensities over the entire object determines the FoV of the reconstructions for ptychographic calculations.

The test object used in these calculations introduce a maximum phase change of 0.9π into the electron beam at the specimen plane and has no magnitude variations. Fig. 5.3a shows the phase distribution of the test object and Fig. 5.3b shows magnitude of the illumination function at the specimen plane. The illumination models a 200 keV electron beam with a defocus value of 150nm and a convergence semi-angle of 15 mrad.

(a) (b)

FIGURE 5.2 – (a) illustration of probe translation at the specimen plane. The effective size of the probe is given by the variable (d). The average translation vector ($\vec{s}$) used in this calculation is equal to 30% of the effective probe size (d). (b) The probe positions were randomised by adding a random vector ($\vec{e}$) to the average translation vector. This gives $\vec{r_s}$ as the correct translation vector for the probe positions, with $|\vec{e}|_{max} = 0.2|\vec{s}|$.

The ptychographic diffraction patterns are generated with illumination positions close to the intersection points of an 8x8 square grid. The average step size ($|\vec{s}|$) is 30% of the effective diameter (d) of the illumination function (see Fig 5.2). The illumination positions differ from the intersection points of the 8x8 square grid by a maximum distance of 20% of the average step size. This is done to prevent the emergence of periodic artefacts in the recovered image. Periodic artefacts arise because the symmetry of a square grid of probe positions results in degeneracy of the solution space; in this case, any function with the correct grid periodicity can be factored out of the recovered object provided its reciprocal is factored out of the recovered probe (Thibault et al., 2009).



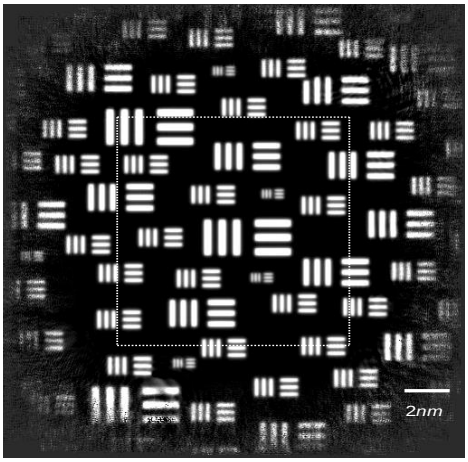
(a) (b)

FIGURE 5.3 – (a) Phase of test object. The background in black represent a phase value of zero and the bars introduce a maximum phase change of 0.9π. (b) The illumination function at the specimen plane, which models a 200 keV electron beam that was generated with a 15 mrad aperture and a defocus of 150 nm.

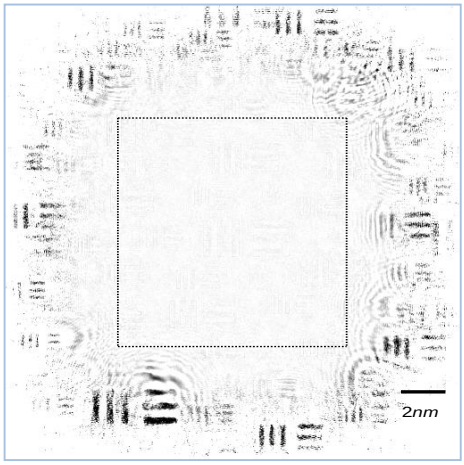
Each diffraction pattern has a count of 10^8 electrons, which incorporates the effect of Poisson noise that is necessary to emulate the fundamental shot noise limitations of ptychographic experiments. The range of incorrect defocus values employed in these calculations change the probe size at the specimen plane by approximately a factor of two in both directions, where the boundary values of the defocus range correspond defocus values at which the ePIE algorithm can no longer recover the illumination. All of these calculations use the correct probe positions.

5.2.1 Results

To provide a reference, the ePIE algorithm was seeded with the correct illumination as the initial guess and the corresponding magnitude and phase of the recovered object is shown in Fig. 5.4. This investigation is important because should the algorithm fail in such a situation, then it will most likely fail in cases that use incorrect initial estimates of the illumination. Further calculations, which used incorrect estimates of the illumination defocus, were run with defocus values that span the range [75% - 160%] of the correct defocus value. The defocus error range is identified by the defocus-ratio, which is defined as the value of the fraction that comprises the correct defocus value in the denominator and the initial estimate of the defocus value in the numerator. These ptychographic calculations use illumination functions with defocus values of 150 nm, 170nm, 190 nm and 210 nm. The trends of mean squared error (MSE) for these different calculations are shown in Fig. 5.5, where each plot corresponds to the MSE after 1000 iterations of the ePIE algorithm. A fewer number of iterations are required in ptychography calculations compared to single diffraction pattern calculations in Chapter 4, because of the redundancy and diversity of the dataset. The images in Fig. 5.6 and Fig. 5.7 show results from calculations that employ diffraction patterns generated from a 150 nm defocus probe. Fig. 5.6a to Fig. 5.6c show initial estimates of the illumination functions that were used as inputs to the ePIE algorithm and Fig. 5.6d to Fig. 5.6f show the corresponding recovered probe functions. Fig. 5.7 shows the corresponding objects recovered from these calculations. In Chapter 2, it was shown that the machinery of the update function corresponds to a convolution in reciprocal space. In the discrete implementation of the convolution, the FFT results in a cyclic convolution. To prevent the algorithm from incorrectly coupling different edges of the diffraction pattern, we insert the diffraction patterns in a larger array of zeros, whose size is larger by the diameter of the aperture.



(a)



(b)

FIGURE 5.4 – (a) Recovered phase of the test object from the ePIE algorithm using the correct illumination function with defocus of 150 nm. The recovered phase has the correct range of 0.9π. (b) Recovered magnitude of test object. The background in white has a transmission coefficient of unity and the bars in black have a minimum transmission coefficient value of 0.3. The dashed lines delineate the region of the recovered objects shown in Fig. 5.6 for calculations that use incorrect defocus values.

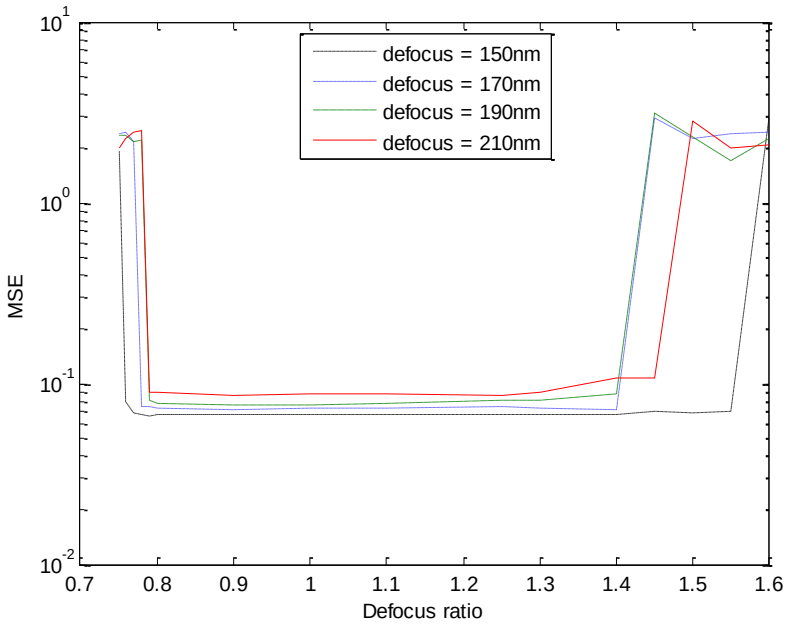
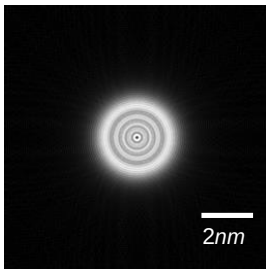
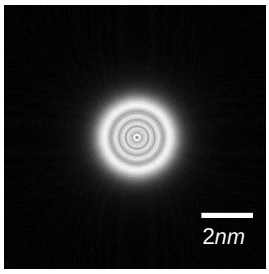


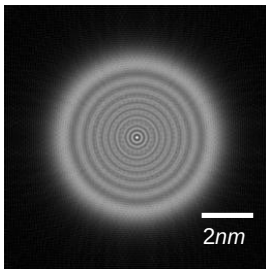
FIGURE 5.5 – Error plot from calculations that employ illumination functions generated from several estimates of the defocus parameter. The lines in the plot correspond to calculations that employ ptychographic diffraction patterns generated from different defocus values (150nm, 170nm, 190 nm and 210 nm). This was done to investigate whether the trend is independent of the defocus of the illumination. These lines comprise error points that correspond to MSE from the ePIE algorithm after 1000 iterations for defocus ratio ranging from 0.75 to 1.6. The defocus-ratio is defined as the value of the fraction that comprises the correct defocus values (150 nm, 170nm, 190 nm and 210 nm) in the denominator and the initial estimate of the defocus (employed by the ePIE algorithm) in the numerator.



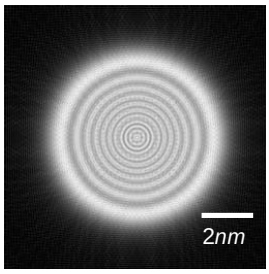
(a) Magnitude of initial probe guess with defocus of 112.5 nm. This corresponds to a defocus-ratio of 0.75.



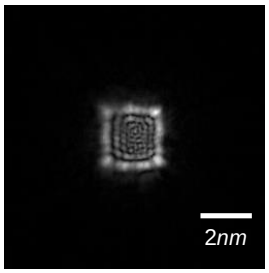
(b) Magnitude of initial probe guess with defocus of 114 nm. This corresponds to a defocus-ratio of 0.76.



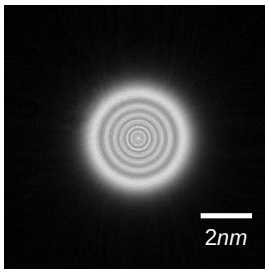
(c) Magnitude of initial probe guess with defocus of 232.5 nm. This corresponds to a defocus-ratio of 1.55.



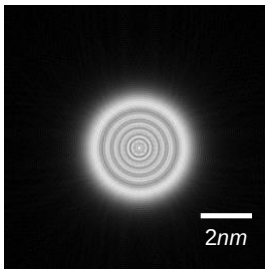
(d) Magnitude of initial probe guess with defocus of 240 nm. This corresponds to a defocus-ratio of 1.60.



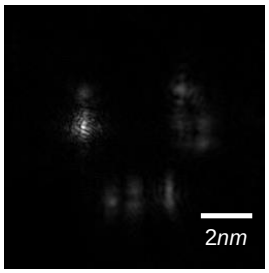
(e) Magnitude of recovered probe from ePIE where 112.5 nm was used as the initial defocus value.



(f) Magnitude of recovered probe from ePIE where 114 nm was used as the initial defocus value.

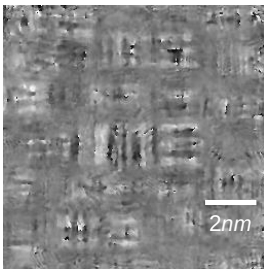


(g) Magnitude of recovered probe from ePIE where 232.5 nm was used as the initial defocus value.

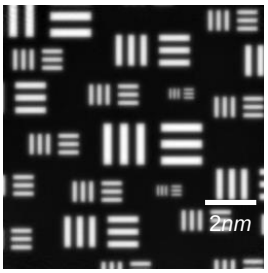


(h) Magnitude of recovered probe from ePIE where 240 nm was used as the initial defocus value.

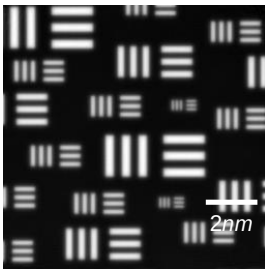
FIGURE 5.6 – Magnitude of initial estimates of the probe are shown in (a) to (d) and corresponding probes recovered by the ePIE algorithm are shown in (e) to (h). These calculations use ptychographic diffraction patterns generated with the probe shown in Fig. 5.1b, which has a defocus value of 150 nm. The defocus-ratio is the value of the fraction that comprises the correct defocus value (150 nm) in the denominator and the initial estimate of the defocus (employed by the ePIE algorithm) in the numerator. These images show that the ePIE algorithm breaks down when the defocus-ratio decreases below 0.76 and also for the case of where the defocus-ratio increases above 1.55. Although only the magnitude images of the recovered illumination are shown here, it is worth noting that the ePIE algorithm also solves for the phase of the illumination.



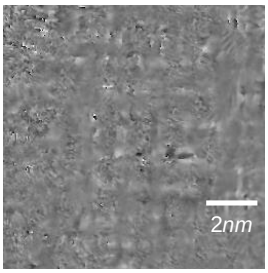
(a) Phase of the recovered object from ePIE where 112.5 nm was used as the initial defocus value of the illumination.



(b) Phase of the recovered object from ePIE where 114 nm was used as the initial defocus value of the illumination.



(c) Phase of the recovered object from ePIE where 232.5 nm was used as the initial defocus value of the illumination.



(d) Phase of the recovered object from ePIE where 240 nm was used as the initial defocus value of the illumination.

FIGURE 5.7 – (a) Phase of the recovered object from the ePIE algorithm for calculations that employ the illumination functions in Fig. 5.6a to Fig. 5.6d. These results also show that the recovered objects departs from the test object in calculation where the ePIE algorithm fails to correctly recover the illumination function (see Fig 5.6e and Fig. 5.6h).

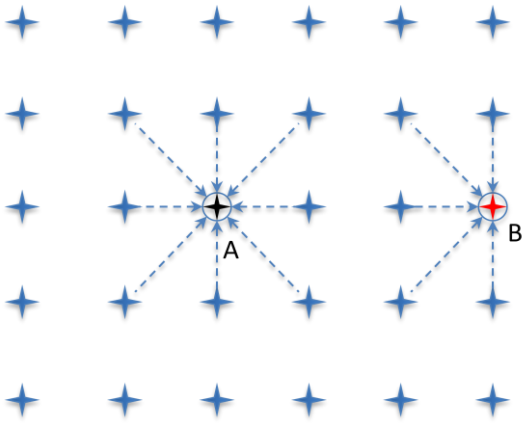


FIGURE 5.8 - illustration of probe position distribution about a typical point (A) within the ptychographic FoV and a typical point (B) located at the edge of the FoV. The grid of 4-point stars represent a set of probe positions in a ptychographic experimental setup. The arrows that point towards point A and point B identify several unique local wave curvatures that illuminate these points when the probe is centred at the origin of these arrows. Here local wave curvature is equivalent to local phase gradient of the illumination function. The local phase gradients at any point displaced from the centre of a quadratic varying phase distribution is determined by the distance from the centre to that point and the direction. This means that the set of local phase gradients at point A span an angle of 360° while the set of local phase gradients at point B span an angle of 180°.

5.2.2 Discussion

Fig. 5.4 shows the results from ePIE calculations where the initial estimate of the illumination is correct but is still updated by the algorithm. At the beginning of the iterative calculation, the ePIE algorithm distorts the form of the illumination but restores the correct structure after a few iterations. This calculation shows that the ePIE algorithm is stable because it does not diverge from the correct solution when the correct illumination is employed. The outer edge of the reconstructions in Fig. 5.4b shows magnitude variations in the recovered object contrary to the constant magnitude of the underlying object, where the magnitude variations appear as cross talk between the magnitude and phase of the recovered object. This compares with diffraction contrast in TEM imaging where regions that scatter the incident radiation to large angles are blocked by the objective aperture and consequently appear dark in the image. In a similar fashion, we can view the magnitude contrast of the recovered object as resulting from an incomplete set of scattered intensities processed by the algorithm. In this framework, the flat magnitude profile of the central part of the recovered object suggests that the missing intensities in one diffraction pattern are expressed in other diffraction patterns and that the update function correctly incorporates the complementary pieces of information into the reconstruction. Object points in the inner region are illuminated by local phase gradients that result from translated versions of the probe, as illustrated in Fig. 5.8. In the Type-II setup, each ptychographic diffraction pattern corresponds to a unique local phase gradients at any point of the specimen. For a typical point in the ptychographic FoV (see point A in Fig. 5.8), intensities scattered outside one diffraction pattern can be found in adjacent diffraction patterns in the dataset because the set of local phase gradients span ~360°; however object points in the outer region (see point B in Fig. 5.8) are illuminated by phase gradients that span an angle of ~180°, so only half of the dataset diversity is available. Thus,

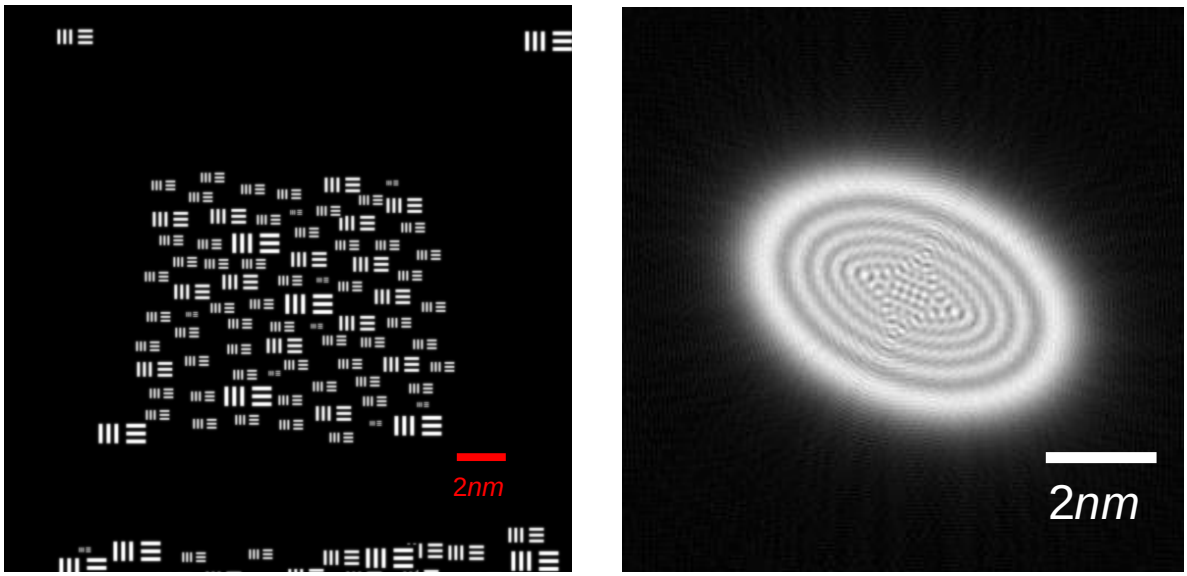
ptychographic diversity can be defined at any point in the recovered object, as a function of the set of phase gradients that the illumination presents at that point from different probe positions. In this framework, the reason for the incorrect magnitude contrast in the outer regions results from the absence of scattering information within the entire dataset because the angular span of the set of phase gradients at these points, is less than 180°.

Although only the magnitude images of the recovered illumination are shown in Fig. 5.6, it is worth noting that the ePIE algorithm also solves for the phase of the illumination. The images in Fig. 5.6 and Fig. 5.7 indicate that the ePIE algorithm can cope with defocus errors that change the illumination size by a factor of two in either direction, in the case of a 150 nm defocused probe. Furthermore, the error plots for several defoci (150 nm, 170 nm, 190 nm, and 210 nm) in Fig. 5.5 show that the ePIE algorithm can cope with a larger range of defocus error in calculations where the defocus-ratio is greater than one compared to calculations where the defocus-ratio is less than one. This behaviour can be attributed to the fact that the percentage overlap parameter in the ptychographic calculations changes with illumination size, provided the relative positions of the illumination remains the same, which is the case these calculations. The percentage overlap parameter increases in ePIE calculations where the defocus-ratio is an improper fraction but decrease for ePIE calculations where the defocus-ratio is a proper fraction. As a result, the ePIE algorithm performs better in the case where the defocus-ratio is greater than one because the algorithm is presented with a large overlapping area, through which it enforce the consistency constraint. Conversely, the overlap area is small in ePIE calculations with the defocus-ratio less than one. An important conclusion is therefore that it is better to overestimate, rather underestimate, the initial defocus value of the probe used in the ePIE algorithm.

5.3 Impact of Two-fold astigmatism error on ePIE

This section investigates the implications of providing the ePIE algorithm with a probe whose two-fold astigmatism coefficient is incorrectly estimated. These calculations also consider the case where the ePIE algorithm processes ptychographic diffraction patterns from the electron microscope realisation of the Type-II setup. The results of this investigation should provide a handle on the amount of astigmatism that the ePIE algorithm can cope with. While the effective size of the illumination is controlled by the defocus parameter, the shape depends on other aberration coefficients. The presence of two-fold astigmatism breaks the cylindrical symmetry of the illumination along the optic axis and thus affects the overlap area of the illumination in ptychographic experiments, where the overlap of the illumination varies with angle. In the Type-II setup, two-fold astigmatism is incorporated into the illumination by adding phase changes, given by the final two terms of Eq. (2.24), at the back focal plane of the condenser lens. This means that astigmatism affects the phase of the convolution kernel and thus influences the expression of information in the diffraction pattern.

The test object used in this investigation is a phase object that introduces a maximum phase change of 0.9π at the specimen plane when illuminated by a 200 keV electron beam. Fig. 5.9 shows the phase of the test object and Fig. 5.9b shows the magnitude of the illumination employed in these calculations. The illumination models a 200 keV electron beam with a 15mrad convergence semi angle, defocus of 190 nm and astigmatism coefficient of ($C_a = [25\ 25]nm$). To generate the ptychographic dataset, the illumination is shifted to positions close to the intersection points of an 8x8 grid. The step sizes span a range of 28% to 32% of the illumination diameter, calculated using only the defocus (190 nm). The results shown in the next section are generated from calculations that use initial illumination guesses with astigmatism coefficients that vary from zero to twice of the correct value.

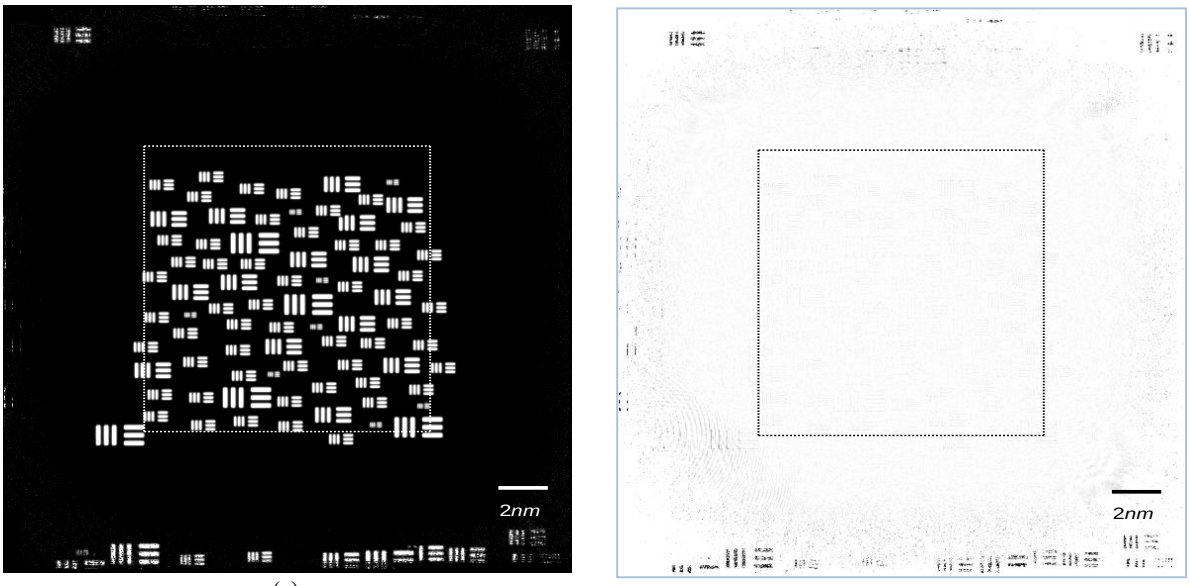


(a) (b)

FIGURE 5.9 – (a) Phase of test object. The background in black represent a phase value of zero and the bars introduce a maximum phase change of 0.9π. (b) The illumination function at the specimen plane, which models a 200 keV electron beam that was generated with a 15 mrad aperture, a defocus of 190 nm and astigmatism coefficient ($C_a = [25\ 25]nm$).

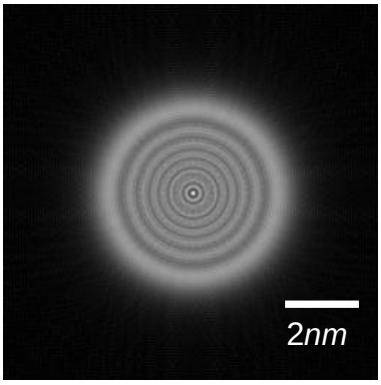
5.3.1 Results

Fig. 5.10 shows the magnitude and phase of the recovered object in calculations that employed the correct illumination as the initial guess. The magnitudes of the initial probe guesses used in these calculations are shown in Fig. 5.11a to Fig. 5.11c and the corresponding probes recovered with the ePIE algorithm are shown in Fig. 5.11d to Fig. 5.11f. The phase of the reconstructions are shown in Fig. 5.12. Fig. 5.13 shows the evolution of the error metric over 1000 iterations.

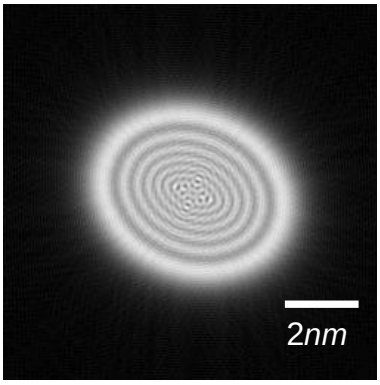


(a) (b)

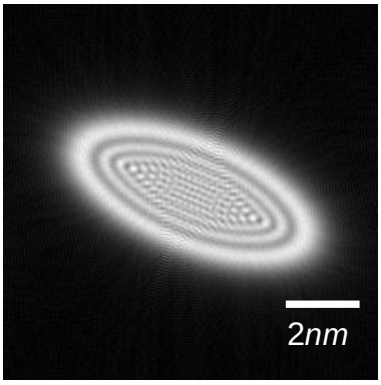
FIGURE 5.10 – (a) Recovered phase of the test object from the ePIE algorithm using the correct illumination function with astigmatism coefficient $(C_a = [25\ 25]\ nm)$. The recovered phase has the correct range of 0.9π. (b) Recovered magnitude of test object. The background in white has a transmission coefficient of unity and the bars in black have a minimum transmission coefficient value of 0.4. The dashed lines delineate the region shown in Fig. 5.12 for calculations that used incorrect astigmatism coefficients.



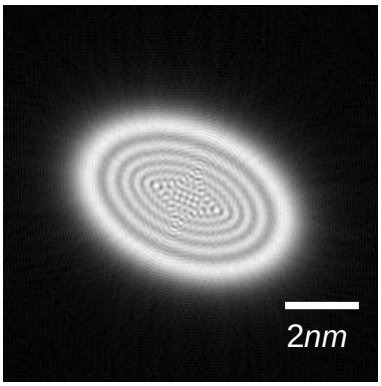
(a) Magnitude of initial probe guess with defocus of 190 nm and astigmatism coefficient ($C_a = [0\ 0]\ nm$). This corresponds to an astigmatism error of -100% from ($C_a = [25\ 25]\ nm$)



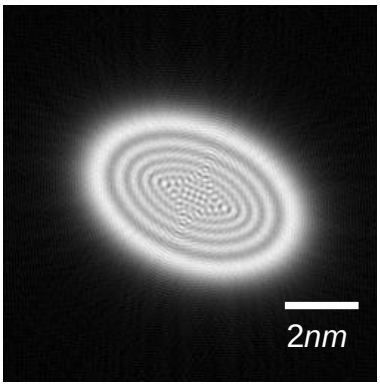
(b) Magnitude of initial probe guess with defocus of 190 nm and astigmatism coefficient ($C_a = [12.5\ 12.5]\ nm$). This corresponds to an astigmatism error of -50% from ($C_a = [25\ 25]\ nm$)



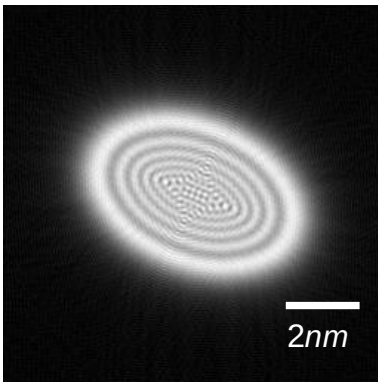
(c) Magnitude of initial probe guess with defocus of 190 nm and astigmatism coefficient ($C_a = [50\ 50]\ nm$). This corresponds to an astigmatism error of +100% from ($C_a = [25\ 25]\ nm$)



(d) Magnitude of recovered probe from calculation in (a).

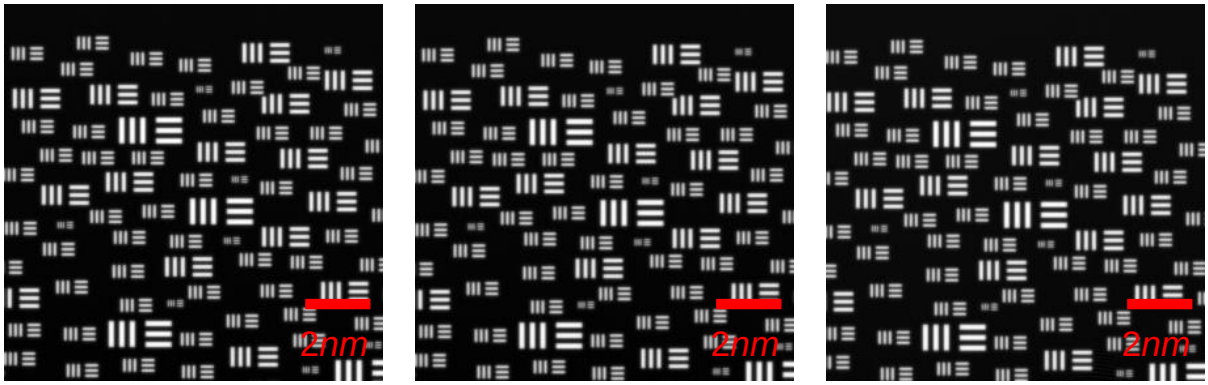


(e) Magnitude of recovered probe from calculation in (b).



(f) Magnitude of recovered probe from calculation in (c).

FIGURE 5.11 – The images in (a) to (c) represent the magnitude of the initial estimates of the probe and the images in (d) to (f) represent the corresponding probe recovered from the ePIE algorithm.



(a) (b) (c)

FIGURE 5.12 – (a) Recovered phase of the test object from the ePIE algorithm using the correct illumination function with $(C_a = [0\ 0]\ nm)$. The images in (b) and (c) represent the phase of the recovered object from the ePIE calculations that use initial illumination estimates shown in Fig. 5.11b and Fig. 5.11c respectively. The FoV of these images correspond to the region delineated by the dashed lines in Fig. 5.10.

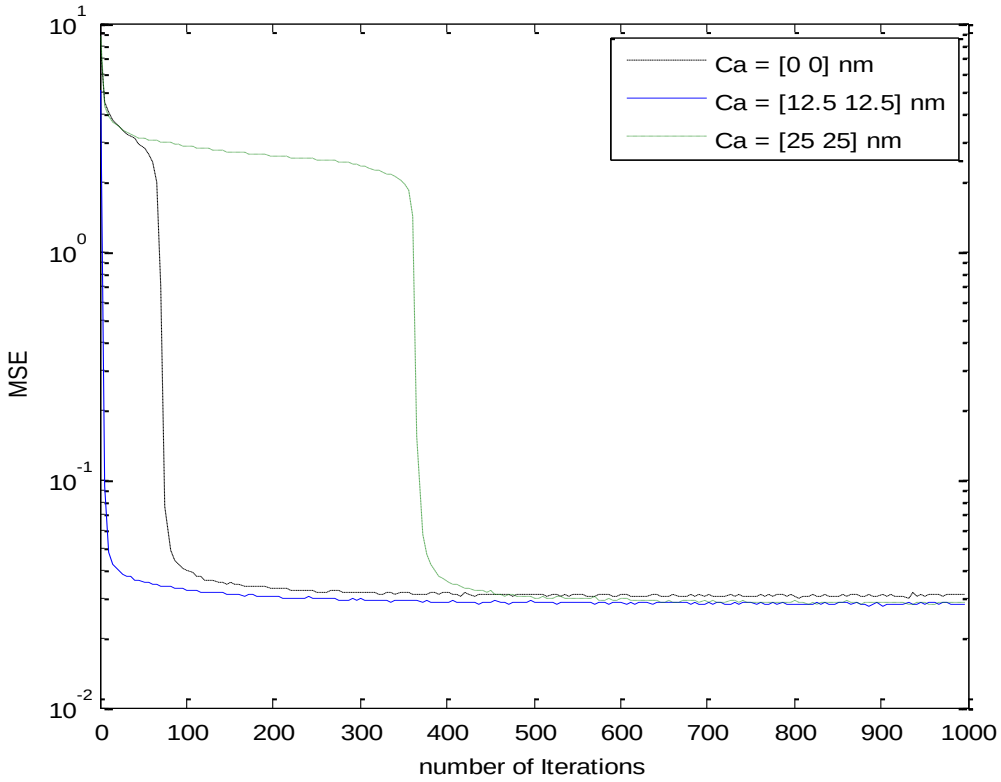


FIGURE 5.13 – Evolution of the error metric over 1000 iterations for ePIE calculations shown in Fig. 5.12. The correct illumination function used to generate the diffraction patterns has an astigmatism value of $(C_a = [25\ 25]\ nm)$. The plot shows that the error of the recovered object drops abruptly after 100 iterations in calculation that use a symmetric probe with aberration coefficient $(C_a = [0\ 0]\ nm)$ as the initial estimate. A similar drop in error occurs after 380 iterations in the case where the astigmatism value of the initial estimate of the probe was overestimated by a factor of 2, i.e $(C_a = [50\ 50]\ nm)$.

5.3.2 Discussion

The quality of the phase reconstructions in Fig. 5.12 and the recovered illumination functions in Figs. 5.11d to 5.11f show that the ePIE algorithm is robust to astigmatic errors in the illumination function. The rapid drop of error values in Fig. 5.13 correlates with the number of iterations required to correctly recover the illumination profile. Although this correlation was established by visual inspection of the changes in the illumination profile as the iteration progresses, it can also be inferred from the error reduction property of the ePIE algorithm; i.e. the algorithm initially minimises the error metric with respect to the guess illumination at the start of the iterative calculation, which gives rise to the rapid fall-off at the start of the iteration. However, subsequent iterations result in a tail-off of the error metric as expected. In this phase of the iterative calculations, the algorithm explores the object and illumination space and slowly corrects the illumination function; as the illumination approaches the correct form, the algorithm goes into another error reduction phase that moves the running estimates of both the object and illumination to their correct profiles.

Thus, when the algorithm correctly recovers a good estimate of the illumination, subsequent calculations are more able to correct for the error in the object, which results in the corresponding rapid decrease of the MSE in Fig. 5.13. Furthermore, Fig. 5.13 shows that the algorithm recovered the illumination function in a fewer number of iteration steps for the case where the initial estimate of the astigmatism coefficient was zero. This is a very important result because it suggests that in ePIE calculations, the astigmatism of the illumination function does not need to be characterised. The trend of the error plots in Fig. 5.13 can be explained by the varying overlap parameter of the illumination with different astigmatism values. The case where the astigmatism coefficient is zero provides a circular shape illumination function, which results in a greater overlap between calculations that the

algorithm performs at several probe positions. This gives the algorithm a larger coupling area to correct the illumination, using the consistency constraint. Conversely, the case with an overestimated value for astigmatism for the illumination corresponds to calculations with a small average percentage overlap because the boundary of the illumination changes from a circular to an oval shape. (See Fig. 5.11a to Fig. 5.11c). In conclusion, these results suggest that the astigmatism coefficient of the illumination function in ePIE calculations do not need to be characterised.

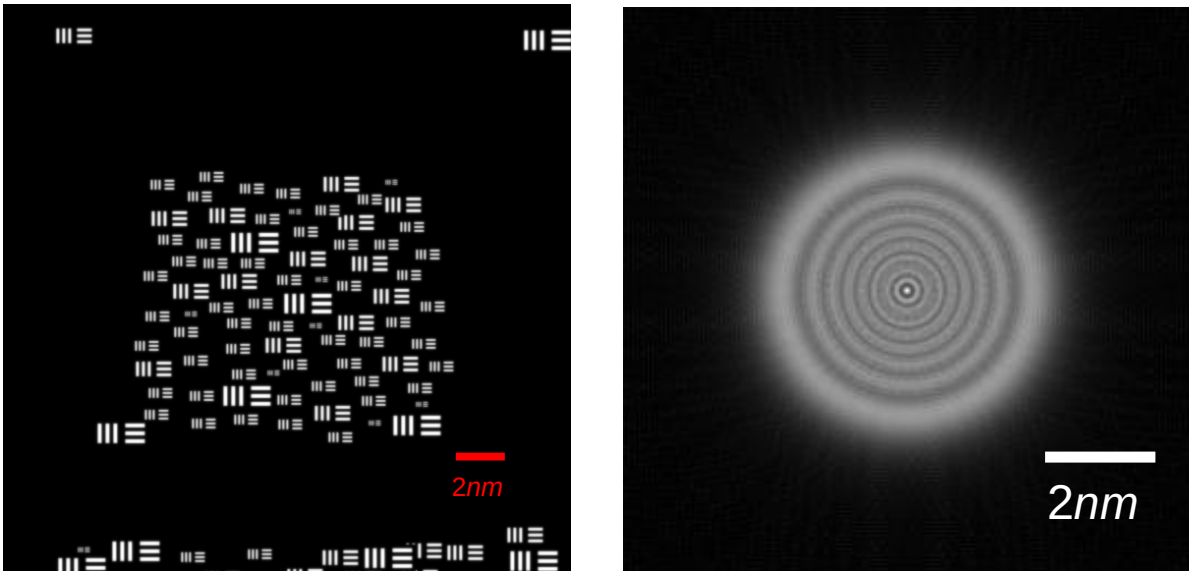
5.4 Impact of count variations on ePIE

The impact of noise on the PIE and ePIE algorithms was first investigated crudely by Faulkner and Rodenburg (2005) and Maiden and Rodenburg (2009) where it was shown that these algorithms are robust to both random noise and Poisson noise. This section investigates the ePIE algorithm in respect to the conclusions reached in Section 4.2, where it is shown that the effective resolution of the recovered objects in the Type-II experiment drops off towards the edge of the FoV, the FoV is the illumination size in single diffraction pattern calculations. In ptychography, the effective FoV depends not only on the illumination size but also on the number of illumination positions spanned by the ptychographic dataset. This section investigates whether this phenomena extends to ptychography; For example, do the recovered images from ptychography contain resolution variation over the FoV, with the highest resolution at locations that correlate with the centre of the illumination positions, or does the ptychographic dataset contain enough information to provide a uniform resolution over the entire FoV? Providing an answer to this question is important because it relates to the mechanism that the update function employs in order to extract information from multiple

diffraction patterns and consequently describes the resolution improvement framework of the PIE and ePIE algorithms, which is discussed in Chapter 6.

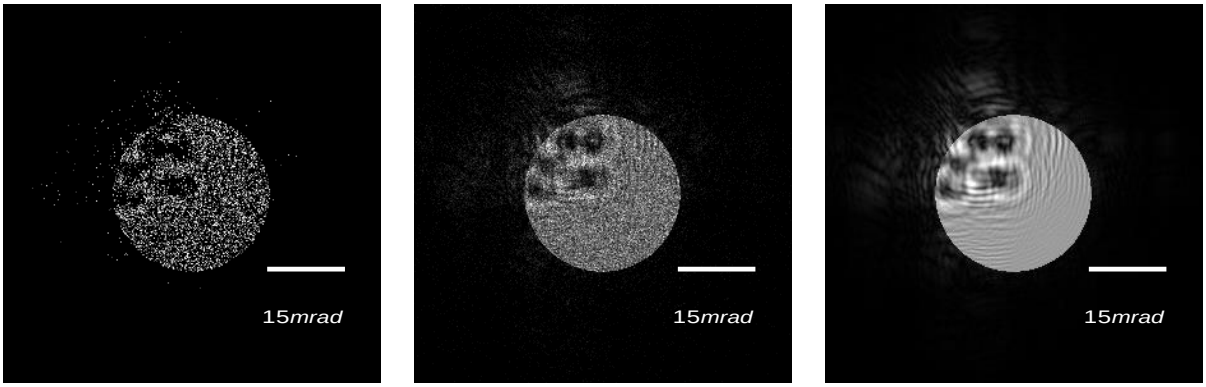
Furthermore, Section 4.2 shows that the update function injects information into the iterative calculations by using the correct illumination. A diffraction pattern with low counts may not contain enough information about the convolution process between the Fourier transform of the illumination and the specimen spatial frequencies. Applying the ePIE algorithm to diffraction patterns with varying degrees of count, in which the algorithm updates both the object and the illumination function, should produce figures for the minimum number of counts required in the ptychographic dataset in order to encode the convolution. The preceding arguments suggest that ePIE calculations may not be properly constrained when the diffraction patterns lack the required convolution information, even when the correct illumination is used as the initial estimate for the ePIE algorithm.

Fig. 5.14a shows the test object used in this investigation. It is a phase object with a constant magnitude of unity, as with the object used in previous sections. Fig. 5.14b shows the magnitude of the illumination used in this investigation. Fig. 5.15 shows three diffraction patterns with different count values, which are typical of the ptychographic diffraction patterns processed by the ePIE algorithm in this section. The ptychographic dataset employed in this section comprise 64 diffraction patterns, collected from a 8x8 grid of 64 illumination positions with a 70% overlap parameter. The illumination models a 200 keV electron beam with a convergence semi-angle of 15 mrad and a defocus of 190 nm.



(a) (b)

FIGURE 5.14 – (a) Phase of test object. The background in black represent a phase value of zero and the bars introduce a maximum phase change of 0.9π. (b) The illumination function at the specimen plane, which models a 200 keV electron beam that was generated with a 15 mrad aperture and a defocus of 190 nm.

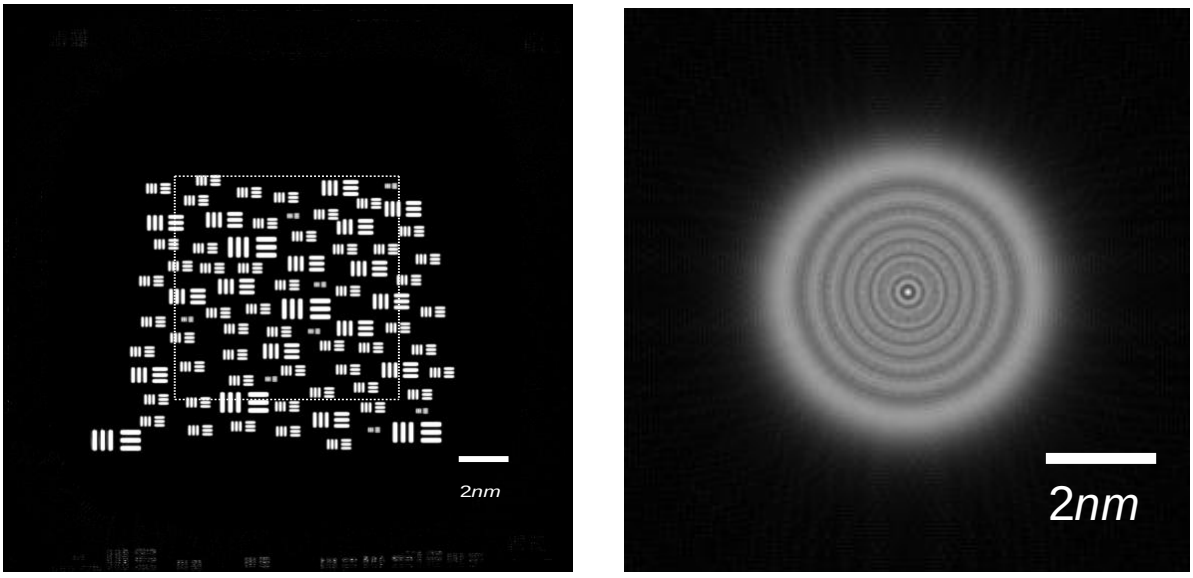


(a) (b) (c)

FIGURE 5.15 – (a) A typical diffraction pattern a ptychographic dataset that containing 10^4 counts. (b) Typical diffraction pattern with 10^5 counts. (c) Typical diffraction pattern with 10^8 counts.

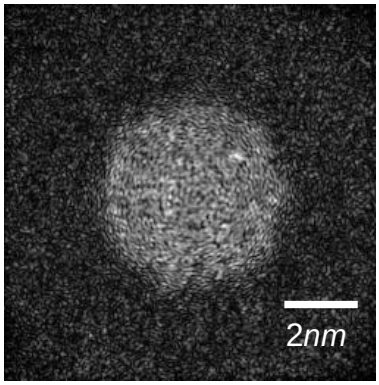
5.4.1 Results

To provide a reference from which to access these calculations, 10^8 count per diffraction pattern were used as a benchmark. This is shown to sufficiently constrain the algorithm in the calculations in Sections 5.2 and 5.3. All calculations in this section used the correct illumination as the initial guess to the ePIE algorithm and were run for 200 iterations. Visual inspection of the recovered object shows that the quality of the recovered object does not improve after 200 iterations. Fig. 5.16 shows the recovered object and illumination function from ePIE for the case of 10^8 counts per diffraction pattern. Fig. 5.17 shows the magnitude of recovered probes from calculations with varying counts after 200 iterations and Fig. 5.18 shows the phase of the corresponding recovered objects. Fig. 5.19 shows the MSE plots over 200 iterations from these calculations.

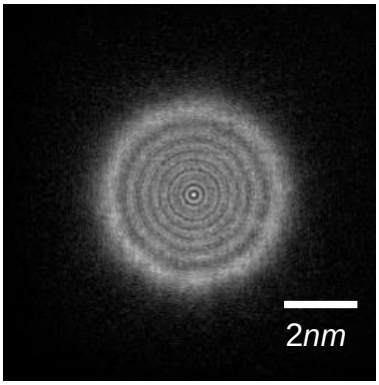


(a) (b)

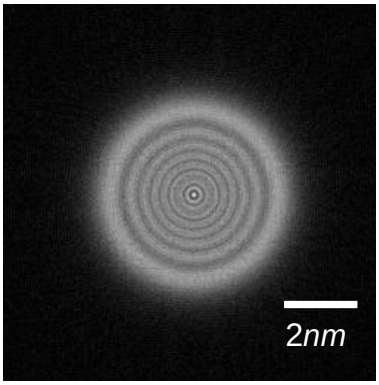
FIGURE 5.16 – (a) Recovered phase of the test object from ePIE, where the counts in each diffraction pattern was normalised to 10^8. The dashed lines delineate the region shown in Fig. 5.18 for calculations that used lower counts per diffraction pattern. (b) Magnitude of recovered probe from ePIE where the correct illumination function with defocus of 190 nm was used as an initial estimate to ePIE.



(a) Magnitude of recovered probe from ePIE where the counts in each diffraction pattern was normalised to 10^4.

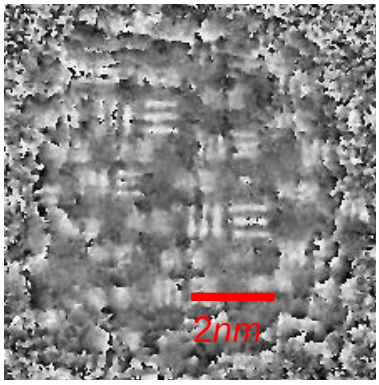


(c) Magnitude of recovered probe from ePIE where the counts in each diffraction pattern was normalised to 10^5.

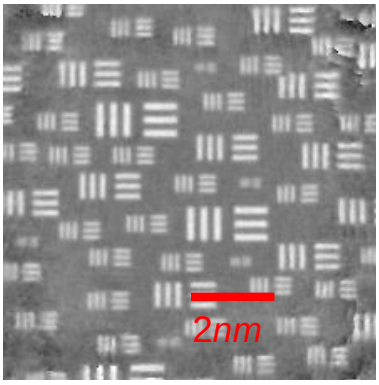


(c) Magnitude of recovered probe from ePIE where the counts in each diffraction pattern was normalised to 10^6.

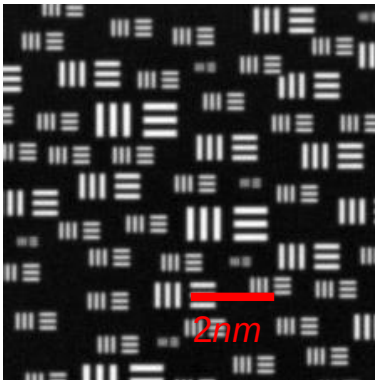
FIGURE 5.17 – Recovered magnitude of the illumination from ePIE calculations that used different counts. The correct illumination function with defocus of 190 nm was used as an initial estimate to ePIE in all three cases.



(a) Phase of recovered object from ePIE where the counts in each diffraction pattern was normalised to 10^4.



(b) Phase of recovered object from ePIE where the counts in each diffraction pattern was normalised to 10^5.



(c) Phase of recovered object from ePIE where the counts in each diffraction pattern was normalised to 10^6.

FIGURE 5.18 – Recovered phase of object from ePIE calculations that used different counts. The correct illumination function with defocus of 190 nm was used as an initial estimate to ePIE in all three cases.

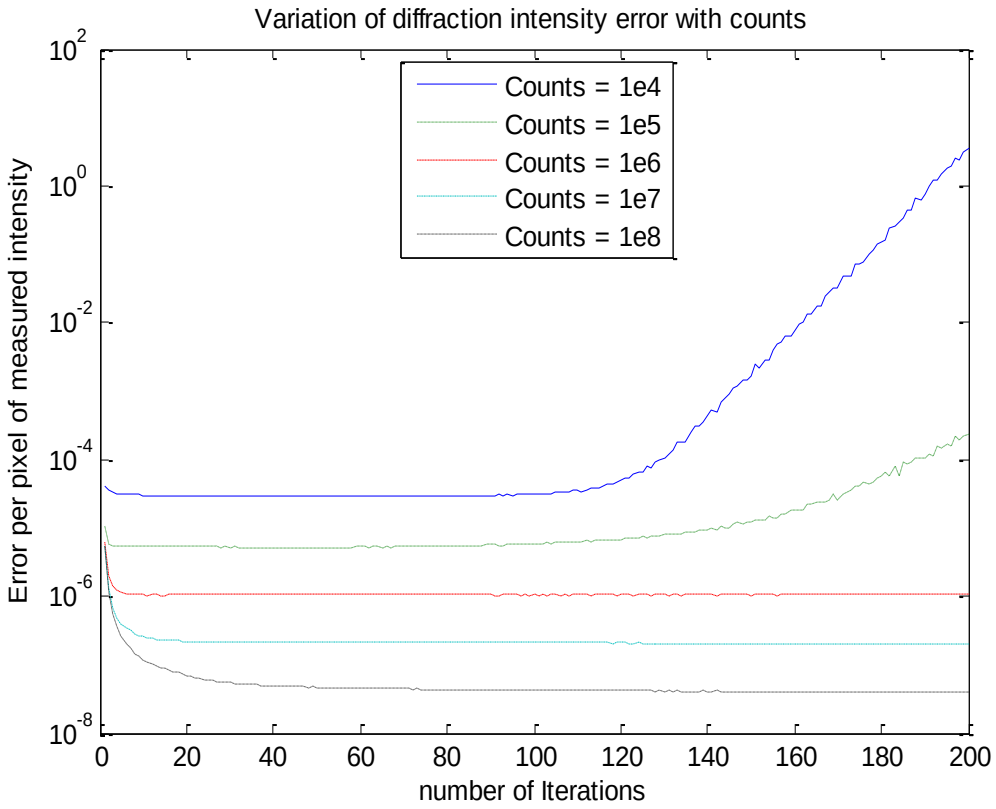


FIGURE 5.19 – Evolution of the error metric over 200 iterations for ePIE calculations where the number of count per diffraction pattern varied from 10^4 counts to 10^8 counts. The plots show that the error increase rapidly after 120 iterations in calculations where the count value is less than 10^6. This suggests that the ePIE algorithm is not sufficiently constrained in calculations with very low count values.

5.4.2 Discussion

Fig. 5.18 shows that the ePIE algorithm is sufficiently constrained with a count of 10^6 per diffraction pattern and that it provides a uniform resolution reconstruction over the ptychographic FoV. This count value gives a very good estimate of the minimum number of counts required to constrain the ePIE algorithm using the illumination parameters employed in this investigation, because the specimen is substantially transparent. Specimens with

structured background scatter more of the incident radiation from the reference beam and consequently provide better signal to noise ratio for the same number of counts.

Fig. 5.18a shows that the algorithm is able to extract some information from the ptychographic diffraction patterns in the case where each one of the 64 diffraction patterns contain 10^4 counts. Fig. 5.17a and Fig. 5.18a confirm the conjecture that diffraction patterns with low counts do not contain enough information about the convolution integral, which is implicit in the update function. This also explains why the results of Section 4.2 are better than expected for low counts calculations. In the calculations of Section 4.2, the update function employs the correct convolution kernel during the iterative update of the object, which serves as an additional source of information for the algorithm. The calculations in this section however do not have this additional constraint and so reducing the counts severely impacts on the quality of the recovered object as is evident in the form of the recovered probe in Fig. 5.17a. It is worth mentioning that the recovered object's phase does not possess the non-uniform resolution reconstruction reminiscent of the results of Section 4.2. Fig. 5.18b shows that the ePIE algorithm recovers uniform resolution information in the inner part of the reconstruction for a count value of 10^5 even though the algorithm struggles to constrain the calculations.

Fig. 5.18 indicates that ePIE extracts most of the information about the object when the number of counts per diffraction pattern is increased by one order of magnitude from 10^4 to 10^5. However, the magnitude of the recovered probe in Fig. 5.17b suggests that the algorithm struggles to constrain the calculations for the case of 10^5 counts per diffraction pattern. The increasing error in Fig. 5.19 confirms that the calculations are not properly constrained in the case of 10^4 to 10^5 counts per diffraction pattern. This is a very interesting result because it suggests that the ePIE algorithm is very different from its predecessor, PIE,

which has been shown to be an error reduction method. This is evident in the calculations of Section 4.2, since the error never increases for cases where the count values are low. It is worth noting that the quality of the reconstructions do not necessarily get worse with iteration as indicated by the increasing MSE after 100 iterations for counts values below 10^5. Instead what happens is that the outer regions of the illumination contain high intensities from incorrect updates by the ePIE algorithm, which subsequently increases the power of the estimated diffraction pattern in the forward calculations and results in high error values.

In the PIE algorithm framework, the calculations are constrained in two planes, with *a priori* knowledge of the illumination in the specimen plane and the diffraction intensity values in the Fourier plane. This is not the case for ePIE because the illumination changes with iterations, so the only constraint (holding things together) is the convolution information embedded in the measured intensity values. Removing this information, as in the case of low counts, results in an unconstrained problem. Finally, Fig. 5.18c and the error trend of the calculations in Fig. 5.19 indicate that the algorithm is better constrained with counts greater than 10^6 for the object used in this investigation.

5.5 Noise and Counting Statistics in practical electron ptychography

This section investigates two different strategies for collecting ptychographic diffraction patterns in the Type-II setup in order to optimise the counts distribution in the ptychographic dataset. Section 5.4 confirmed that the ptychographic dataset contains a partition of resolution information that helps recover a large FoV with uniform resolution. This suggests

that the counting statistics per resolution pixel for ptychographic reconstructions may depend only on the total number of counts within the ptychographic dataset. This is similar to the dose-partition theory of Hegerl and Hoppe (1976), where the number of counts required for a tomographic reconstruction can be distributed over a large number of noisy two-dimensional projections (McEwen, Downing and Glaeser, 1995). This suggests that for a given resolution, the count statistics of ptychographic reconstructions may depend only on the dose at the specimen plane so that the counts can be distributed into a large number of diffraction patterns. This has the advantage of minimizing the effect of specimen drift during the data collection process. The effect of drift during ptychographic data collection can be minimised by distributing the counts into a large number of noisy diffraction patterns, where each diffraction pattern is recorded during a short exposure time. For this data collection scheme, the effective specimen drift vector is mapped to deviations of probe positions from their correct values, which can be accounted for at least approximately during the iterative calculations (Hurst et al., 2010). However, a long exposure time collection scheme is plagued by specimen drift during data collection because each diffraction pattern results from an averaging of the diffraction pattern intensity values from a range of specimen positions. This blurs out the spatial frequency information in the diffraction pattern and reduces the maximum achievable resolution. In the Type-II setup, this is equivalent to spatial partial coherence, which is illustrated in Fig. 4.6. Section 5.5.1 compares results from calculations that employ a small probe to generate a large set of noisy diffraction patterns dataset with calculations that employ a large probe to produce fewer less noisy diffraction patterns.

5.5.1 Optimisation of data collection scheme

To investigate the effect of the count distribution over the ptychographic dataset, the model setup shown in Fig. 5.20 is employed. The number of probe positions required to cover a given FoV by a large defocused probe is fewer compared to that required for a small defocused probe, as illustrated in Fig. 5.21 and Fig. 5.22a. To maximise the radiation flux traversing the specimen in the short exposure calculation the probe size at the specimen is made small so that counts in the corresponding diffraction pattern comes mainly from a small region of the specimen. This collection scheme generates a large ptychographic dataset that takes longer to process but should improve the coherence property of the noisy ptychographic diffraction patterns.

The alternative collection scheme employed generates a ptychographic dataset with fewer, less noisy diffraction patterns from a large probe. The exposure time at each position is longer and the illumination covers a larger area of the specimen, such that only a few probe positions are required to cover the same FoV. To model a fixed exposure time in these calculations, the total number of counts over the entire dataset in both setups was fixed at the same value. This count value is used to incorporate Poisson noise in the diffraction patterns. Consequently, the dataset created with a large defocused probe would have more counts per diffraction pattern but fewer diffraction patterns and vice versa, this is shown in Fig. 5.21b and Fig. 5.22b. As a result some of the detector pixels have values equal to zero representing the absence of electrons at the corresponding scattering angle in a given diffraction pattern. Although this suggests a loss of information regarding several spatial frequencies, it is not necessarily the case because the same information may reside in a different diffraction pattern. The expression of spatial frequency information across the entire ptychographic dataset is one of several advantages of ptychography and is further explored in Chapter 6.

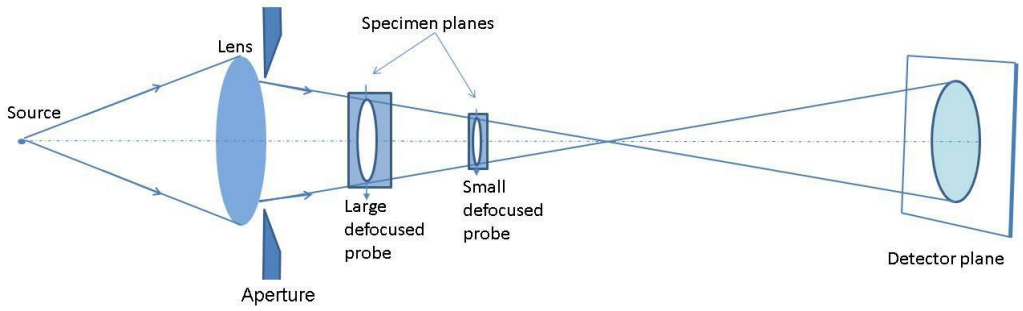
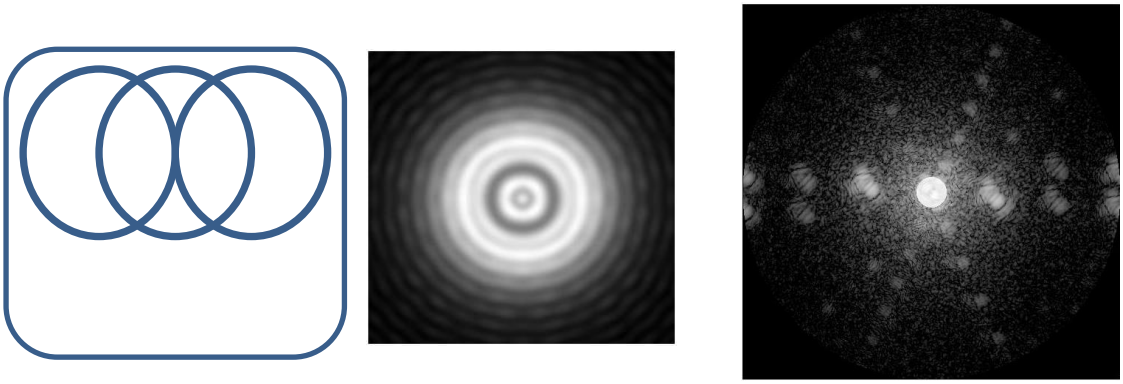
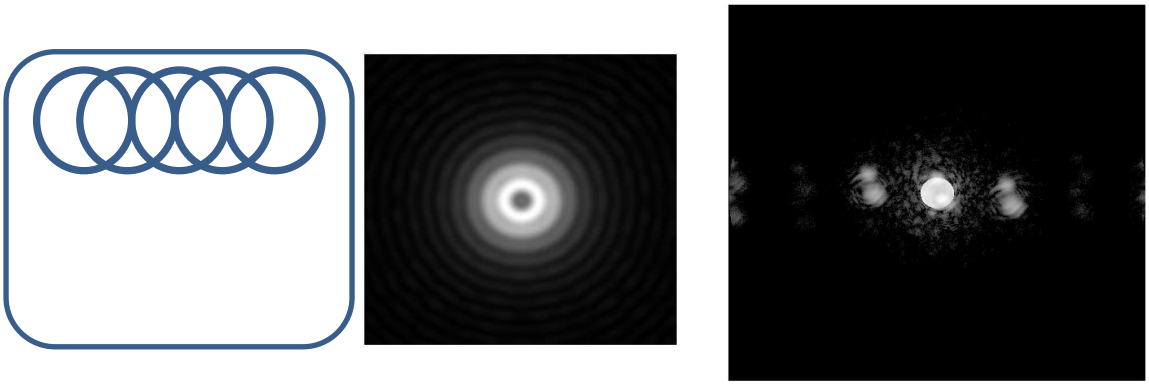


FIGURE 5.20 - Modelled experimental setup used to investigate the dependence of counting statistics on dataset size due to variation in probe defoci in Ptychography.



(a) (b)

FIGURE 5.21 – (a) illustration of the number of probe positions required to span a given FoV for 90 nm defocus probe for a 50% overlap parameter. (b) Typical diffraction pattern from 90 nm defocus probe. The central bright field disc of the diffraction pattern spans a semi-angle of 10 mrad.



(a) (b)

FIGURE 5.22 – (a) illustration of the number of probe positions required to span a given FoV for 45 nm defocus probe for a 50% overlap parameter. (b) Typical diffraction pattern from 45 nm defocus probe. The central bright field disc of the diffraction pattern spans a semi-angle of 10 mrad.

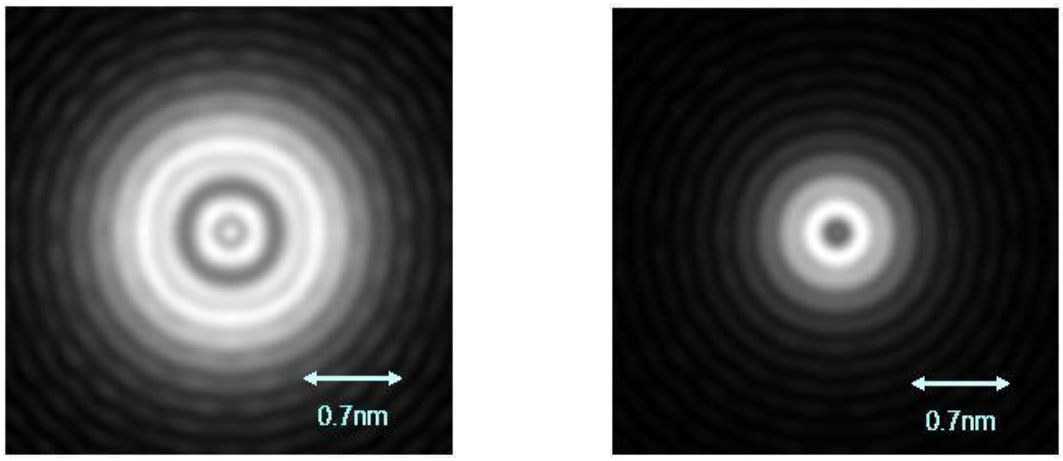
The minimum number of counts required for a good reconstruction depends on the specimen used and is higher for free standing, transparent specimens because most of the incident radiation propagates through the specimen with little interaction. In most cases, the specimen under investigation requires a support film such as amorphous carbon film. In a diffraction experiment, the support film contributes coherently to the signal in the diffraction pattern. As a result, the recovered function from the ptychographic dataset gives an approximation of the projected atomic potential of the specimen and the support film. The calculations in this section include an amorphous carbon support film with a thickness of 4 nm. Atomic coordinate distributions in amorphous carbon are from Mountjoy (1999), which were used in the investigation of atomic structure probing of amorphous structures in high-resolution electron microscopy (Mountjoy, 2001). Test features analogous to small gold particles were placed on the support. The periodicities of these particles model different particle orientations, in which the smallest periodicity represents atomic planes that are seen in projection to be separated by 0.06 nm. The exit wave immediately downstream of the 4 nm amorphous support is modelled with the *multislice* method (Kirkland, 2010). The ptychographic diffraction patterns are calculated by propagating the illumination through the amorphous support at each probe position and taking the FFT of the corresponding exit wave fields.

Figs. 5.23a and 5.23b show the magnitude of the probes used in the ptychographic reconstruction. These probes model a 200 keV electron beam with a defocus of 90 nm and 45 nm respectively and represent the probe profile at the top surface of the amorphous carbon film. The calculations in this section employ a percentage overlap of 87% of the respective probe diameters in order to generate both ptychographic datasets. The FoV in these calculations is spanned by 5x5 probe positions of a 90 nm defocus probe and 16x16 probe positions of a 45 nm probe. Thus the 45 nm defocused probe dataset is about 10 times larger

than the dataset of 90 nm defocused probe. The maximum scattering angle captured by the detector in these calculations is 128 mrad and the detector comprises 512x512 pixels.

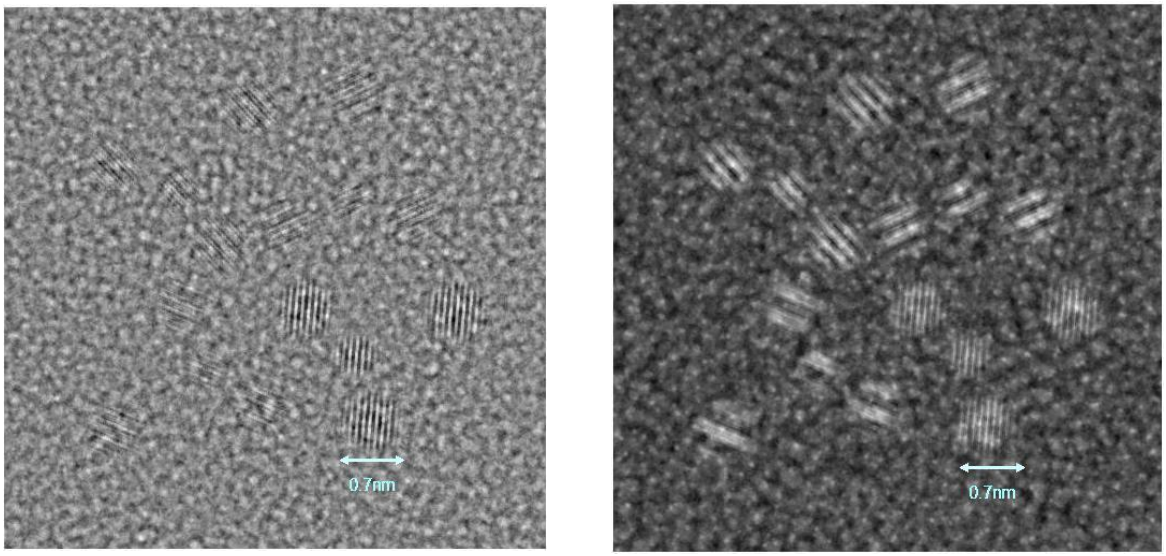
5.5.2 Results

Figs. 5.24a and 5.24b show the amplitude and phase of the exit wave emanating from the test specimen as it would appear under a plane-wave illumination. The diffraction patterns were processed with the PIE algorithm so that only the object function was solved for during the iterative calculations. The calculations were performed with different count values and the count threshold where one of the calculations gave a better reconstruction was selected. This corresponded to a count value of 10^7. Fig. 5.25 shows the phase of the recovered objects in the calculations that employed a total of 10^7 counts. Count values above 10^7 produced similar quality reconstructions for both collection schemes.



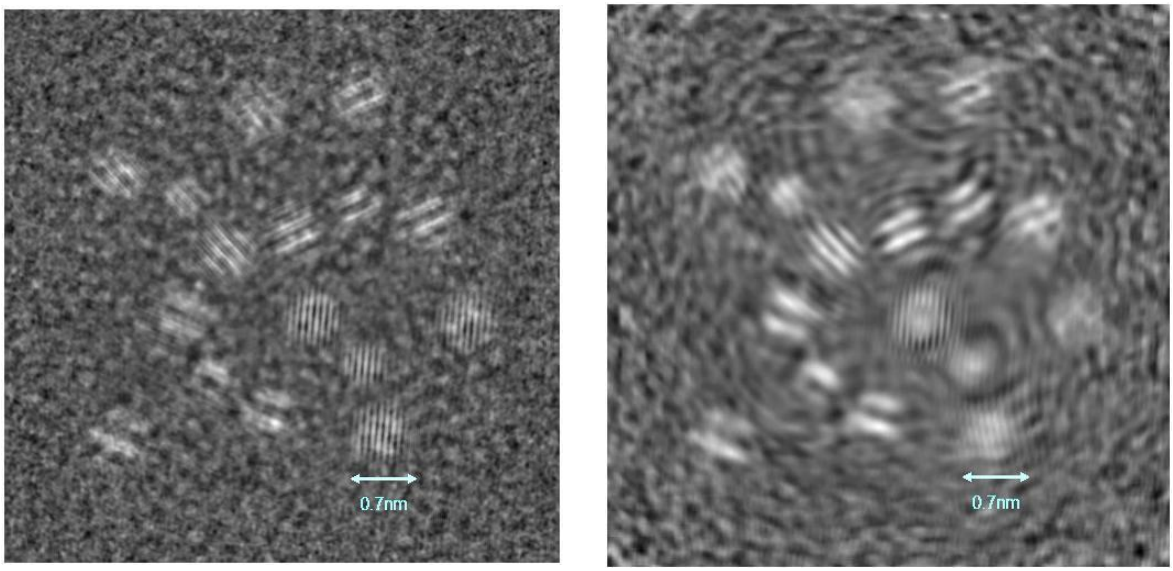
(a) (b)

FIGURE 5.23 – (a) Magnitude of probe generated with a defocus of 90 nm, convergence semi angle of 10 mrad, a 200 kV acceleration voltage. (b) Magnitude of probe generated with a defocus of 45 nm, convergence semi angle of 10 mrad and a 200 kV acceleration voltage. The FoV of these images represents the central half of the real space window used in the iterative calculations.



(a) (b)

FIGURE 5.24 – (a) Magnitude of complex exit wave of specimen on a 4 nm amorphous carbon support, together with synthetic test particles. (b) Phase of complex exit wave of specimen on a 4 nm amorphous carbon support, together with synthetic test particles



(a) (b)

FIGURE 5.25 – (a) Phase of reconstruction for 90 nm defocus probe. (b) Phase of reconstruction for 45 nm defocus probe. These calculation used a total of 10^7 counts in the ptychographic dataset. This count value corresponds to calculations for which the quality of the reconstruction recovered by the PIE algorithm differs; that is to say, number of counts an order of magnitude above this value gave similar quality phase reconstruction from the 45 nm and 90 nm probe calculations.

5.5.3 Discussion

The phase of the recovered objects in Fig. 5.25 demonstrates that the PIE algorithm can recover a fairly good representation of the test object, when placed on a three-dimensional amorphous carbon structure with a thickness of 4 nm, even though the algorithm assumes a multiplicative approximation at the specimen plane. This means that an amorphous carbon supports does not severely degrade the quality of the object of interest, which is an important finding because most specimens under investigation require a support film to hold them in place during data collection. Fig. 5.25b shows that the particles at the outer region of the reconstructions were recovered with lower resolution than those in the middle of reconstructions. This confirms earlier findings in Section 4.5, which showed that the central region of the reconstruction has the most redundancy in the dataset, because its information is expressed in more diffraction patterns compared to the outer region of the recovered object. The reconstruction quality of the central region in both phase images of Fig. 5.25 demonstrates that spatial frequency information is distributed throughout the ptychographic dataset. However, the recovered particles shapes are poorly constrained in the calculations that employ a 45 nm defocused probe, where a typical diffraction pattern contains an average dark field count of 1 electron per detector pixel. This suggests that the shape of recovered particles is more affected by Poisson noise in the diffraction pattern compared to the high-resolution information within the particles. This result can be explained by considering the properties of the diffraction pattern intensities that encode particle shape information and periodicity information.

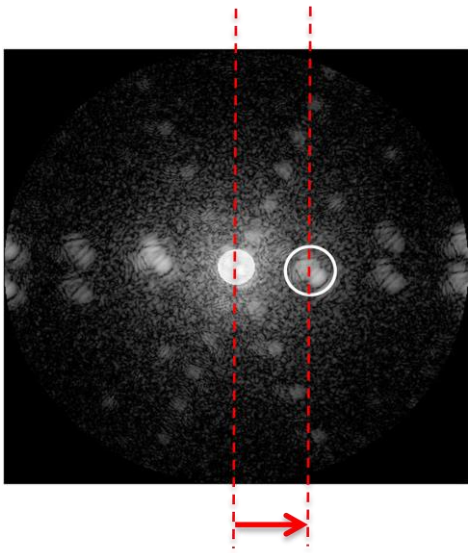


FIGURE 5.26 - illustration of patches of intensity scattered outside the bright field disc. This diffraction pattern results from a region of the specimen where two closely spaced particles contribute to the off-centred intensity patches. The intensity distribution within the white circle is determined by the shape of the illuminated particles.

Periodicity information corresponds to the presence of off-centred patches of intensity in the diffraction patterns while the shape of the particles correspond to the intensity configuration within these patches as illustrated in Fig. 5.26. These mappings arise because each particle can be considered to be a composite function that is composed of the product of a periodic function (which has the correct particle periodicity) and a mask (which defines the shape of the particle). This gives a convolution in Fourier space that encodes the shape information around the delta peaks of the periodic function. A further convolution of this composite function with the Fourier transform of the illumination gives the measured diffraction pattern, where it is assumed that the exit wave can be separated into the illumination and the projected atomic potential of both the particles together and the support film. This means that intensity values located at the correct region of the diffraction patterns allow the recovery of high-resolution information within the particles. However, intensity configurations within any patch of the diffraction pattern are distorted by Poisson noise for very low count diffraction patterns, which is the case for the 45 nm defocused probe dataset. Noise distortion of the

intensity configurations in the diffraction pattern results in delocalisation of the fringes in the recovered object as shown in Fig. 5.25b. It is worth noting that the redundancy and diversity of the ptychographic dataset is the only reason why such a low count value, an average of 1 electron per detector pixel in the case of 45 nm probe, gives a reasonably good reconstruction. This is because the corresponding intensity configuration in the diffraction pattern is highly distorted. The results in Fig. 5.25 suggest that the optimum ptychographic configuration should contain fewer diffraction patterns in order to build up enough counts per diffraction pattern to improve reconstruction of particle shapes. In fact, this is a low count for any type of experiment.

Experimentally, a highly coherent configuration with a beam current of 10pA produces a count of 10^8 electrons for an exposure time of 1 second (as measured on a JEOL 2010F, courtesy of Dr. Francis Sweeney). This count value is a factor of 10 greater than the count value used in these calculations, which is 10^7. Therefore, by employing a count value of 10^8, the calculation that employs a 45 nm defocused probe improves to give a similar quality reconstruction to that shown in Fig. 5.25a, because the number of counts per diffraction pattern for a 45 nm defocused probe dataset is a factor of 10 lower that of a 90 nm defocused probe dataset. This means that we can overcome the effect of specimen drift during the collection of experimental diffraction patterns. Furthermore, a small probe size gives rise to large features in the diffraction pattern, which can be binned to improve the counting statistics at the detector plane, provided that the size of the resulting window in realspace is greater than the Nyquist cell size in order to minimise the effect of aliasing.

5.6 Summary

Section 5.1 shows that the ePIE algorithm can recover the correct illumination function from estimates that are within a factor of two of the correct illumination size. The ePIE algorithm is also shown to completely recover the astigmatism of the illumination in the Type-II setup. A connection between the phase gradient of the illumination and the structure of the ptychographic diversity was made by analysing the cross-talk between the magnitude and phase of the recovered object. Section 5.4 shows that the ptychographic dataset contains the required redundancy and diversity to provide a reconstruction with uniform resolution over the central part of the FoV of the recovered object. This diversity is further explored in Chapter 6. Section 5.5 shows that the effects of drift can be minimised in the electron microscope implementation of the Type-II setup by collecting a large set of noisy ptychographic diffraction patterns.

Chapter 6

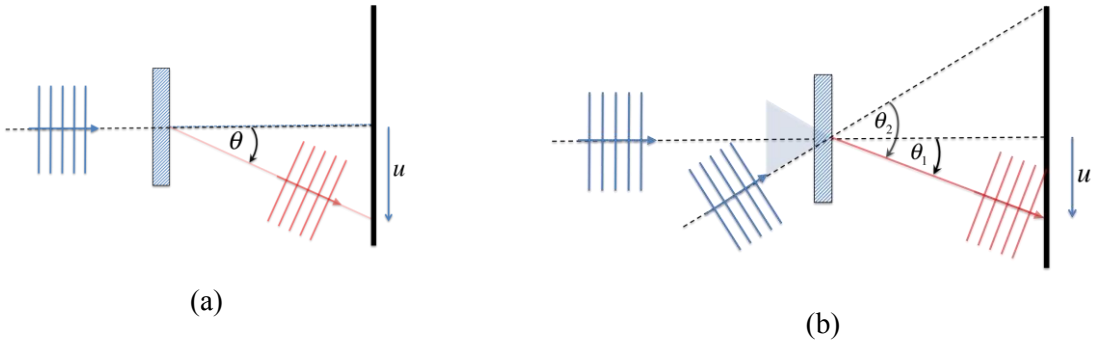
6 Bandwidth extrapolation in Ptychography

Resolution improvement of the recovered object using the PIE algorithm has been demonstrated with numerical simulations by Edo, Zhang and Rodenburg (2010). This concept has also been demonstrated experimentally at visible light wavelengths by Maiden et al (2011), where the technique was extended to the ePIE algorithm. These resolution improvement methods work by inserting the measured diffraction pattern into a larger array of zeros and letting the PIE and ePIE algorithms solve for the complex values outside the measured data. In order to explain the reason why these algorithms can recover scattered intensity values outside the detector, Section 6.1 discusses specimen scattering in the case of plane-wave and curved wave illumination, where it is shown that the diffraction pattern recorded by a detector encodes spatial frequency information greater than the semi-angle of the detector, when the specimen is illuminated by structured illumination. Section 6.2 investigates the extent of the specimen spatial frequencies encoded in the ptychographic dataset for the Type-II setup. This setup was used because it allows analysis in a relatively

simple framework that is easily extended to other forms of structured illumination. Section 6.3 discusses the issues that need to be taken into consideration when implementing the aforementioned resolution improvement scheme using the PIE algorithm. The results of implementing this scheme using the PIE algorithm are presented in Section 6.4. The pitfalls associated with the implementation of the resolution improvement scheme, in the case of the ePIE algorithm, are discussed in section 6.5. A novel method, developed based on a feedback system that addresses these issues is also presented. Section 6.6 presents experimental ePIE reconstructions, which demonstrate that the ePIE algorithm is sufficiently constrained by the aforementioned feedback scheme. Section 6.7 investigates the impact of sub-Nyquist sampling of the ptychographic diffraction patterns, where the PIE algorithm uses the redundancy and diversity of the ptychographic dataset to help alleviate the effect of aliasing.

6.1 Theory of resolution improvement with curved illumination

In diffraction experiments, where a plane-wave is incident on the specimen, the coordinates of the detector plane in the far field can be mapped to angular coordinates relative to the specimen plane as illustrated in Fig 6.1a. In this case, the counts measured at off-centred pixels correspond to radiation flux that is diffracted by the specimen. i.e. in the absence of the specimen, all the counts are registered at the central pixel of the detector. In this configuration, the pixels of the detector have one to one correspondence with the spatial frequencies of the specimen. This is the setup used to measure the power spectrum of a specimen.



(a) (b)

FIGURE 6.1 - (a) Unique mapping of detector pixel coordinate to spatial frequency component of the object. (b) Non - unique mapping of detector pixel coordinate to spatial frequency component of the object

In the Type-II setup, the illumination is not a plane-wave but a curved wave. In this setup, off-axis pixels in the detector no longer possess a one to one correspondence with the spatial frequencies of the specimen, as illustrated in Fig. 6.1b. This is because beams scattered by different angles can arrive at the same detector pixel. As a result, resolution is not determined by the angular size of the detector. In the Type-II setup, one detector pixel contains information from several spatial frequencies. Conversely, information about a single spatial frequency spreads over several pixels of the detector. This non-unique mapping of spatial frequency information comes from the folding of diffraction space and results from the use of structured illumination. This folding of diffraction space is closely linked to the ideas of Hoppe (1969a) in his formulation of a solution to the phase problem and resonates in the name of the resulting solution (ptychography: where 'ptycho' means to fold (Hegerl and Hoppe, 1970)). The exact nature of this folding for structured illumination is expressed mathematically by the convolution integral of Eq. (6.1b).

$$\psi(\boldsymbol{r}) = O(\boldsymbol{r}) \cdot P(\boldsymbol{r}) \tag{6.1a}$$

$$\psi(\boldsymbol{u}) = \mathcal{F}\psi(\boldsymbol{r}) = \int O(\boldsymbol{k})P(\boldsymbol{u}-\boldsymbol{k})d\boldsymbol{k} \tag{6.1b}$$

$$I_m(\boldsymbol{u}) = |\psi(\boldsymbol{u})|^2 = \left|\int O(\boldsymbol{k})P(\boldsymbol{u}-\boldsymbol{k})d\boldsymbol{k}\right|^2 \tag{6.1c}$$

Here $\psi(\boldsymbol{u})$ represents the complex wave distribution at the detector plane, $\psi(\boldsymbol{r})$ is the exit wave immediately downstream of the specimen, $O(\boldsymbol{r})$ corresponds to the object transmission function and $P(\boldsymbol{r})$ is the illumination function; again the coordinates $\boldsymbol{r} := (x, y)$ and $\boldsymbol{u} := (u, v)$ differentiate between the exit wave at the specimen plane and the propagated wave at the detector plane respectively, also the coordinate $\boldsymbol{k} := (k_x, k_y)$, spans the Fourier transform of the object and is used as a dummy variable in Eq. 6.1b and 6.1c. The integral in Eq. (6.1b) is a consequence of the Fourier convolution theorem where $P(\boldsymbol{k}) = \mathcal{F}P(\boldsymbol{r})$ and $O(\boldsymbol{k}) = \mathcal{F}O(\boldsymbol{r})$.

This chapter deals with the case of two-dimensional specimens to ensure direct correspondence between the Fourier transform of the specimen ($O(\boldsymbol{k})$) and spatial frequency content of the specimen. Apart from this approximation, the specimen can be either crystalline or textured as in the case of amorphous materials. Consequently, the exit wave from the specimen is a product of the specimen transmission function and the illumination function as shown in Eq. (6.1a).

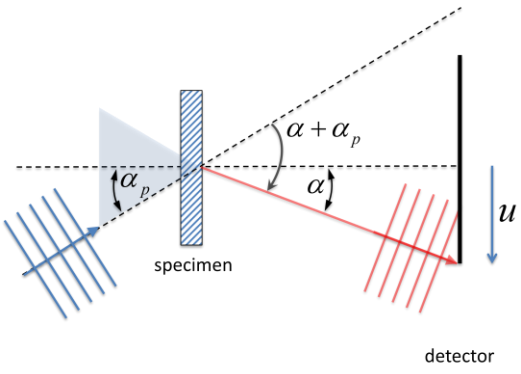


FIGURE 6.2 - The detector of angular size α mrad captures scattering information up to $(\alpha + \alpha_p)$ mrad, when the specimen is illuminated by a curved-wave composed of a range of plane waves that span an angle of α_p mrad.

Eq. (6.1c) suggests that any point in the detector captures information relating to several spatial frequencies of the object. Put differently, the detector measures the interference

amongst the spatial frequencies of the specimen. This means that a detector with a semi-angle of α can capture frequencies up to $\alpha + \alpha_p$ in the detector plane, where α_p represents the range of angles that make up the illumination in mrad, as shown in Fig. 6.2. This suggests that in principle, a finite-size illumination function at the specimen plane should give wavelength limited resolution because of the extended nature of its frequency spectrum. For example, in the case where the illumination has the form of a top hat function at the specimen plane, the corresponding frequency spectrum is the Airy function, which is extended in the frequency domain. Consequently, the diffraction pattern contains information for very high spatial frequencies that is weighted by the magnitude of the Airy function. However, the maximum resolution for diffractive imaging experiments is limited by the wavelength λ; thus, this setup could in principle give a wavelength limited resolution.

In the Type-II setup, the condenser aperture determines the relative strength of the plane-wave components that make up the illumination and is given by the top hat function. Since the intensity is uniform over the aperture, the effective range of angles (α_p) that make up the Type-II curved illumination (see Fig. 6.2), equals the semi-angle of the condenser aperture. The finite extent α_p/λ of the condenser aperture in frequency space means that it convolves only a finite number of spatial frequencies into a pixel of the detector. This means that resolution improvement with this setup can only extend the bandwidth of the image by a finite amount, α_p/λ. In the case of the electron microscope realisation of the Type-II setup, the resolution of the microscope can be doubled, when only the central bright field of the diffraction patterns or Ronchigrams are processed by the ePIE or ePIE algorithms. This has profound implications for electron diffractive imaging because it significantly reduces the dynamic range requirement of the detector and allows doubling of the resolution for both aberration-corrected and uncorrected electron microscopes.

Deconvolution of the complex wave field at the detector plane provides the means for extracting spatial frequencies that are greater than the bandwidth α/λ of the detector, where the detector subtends a semi-angle of α on the specimen. The detector only measures the intensity of the complex wave field and as a result direct deconvolution is not possible. It is worth noting that direct deconvolution cannot be used in the Type-II setup even when the complex wave distribution at the detector plane is known. This is because the direct deconvolution problem is ill-conditioned; i.e. implementing a direct deconvolution (through the Fourier division) results in division by zeros at real space points where the illumination intensity values equal zero. Iterative deconvolution techniques are better suited to deal with this problem. Specifically, the PIE algorithm performs iterative deconvolution via a weighted convolution operation and thus avoids the problem of dividing by zeros. Iterative calculations either require complete knowledge of the illumination as in the case of PIE, or a reasonable estimate, as in the case of ePIE, which is demonstrated in Chapter 5.

6.2 Information expression of ptychographical diffraction patterns

In this section, we consider the expression of resolution information in the diffraction pattern of the Type-II setup. In this setup a lens generates a defocused illumination spot by introducing a quadratic phase shift at its back focal plane. This is illustrated in Fig. 6.3. The complex wave profile $P(\boldsymbol{k})$ at the back focal plane of the lens is described by the Fourier transform of the illumination function and is given by

$$P(\boldsymbol{k}) = A(\boldsymbol{k}) \exp\left(i\pi\Delta z\lambda\left(k_x^2 + k_y^2\right)\right), \tag{6.2}$$

where $A(\boldsymbol{k})$ represents the aperture function in the back focal plane, $k_x = X/\lambda d$ and $k_y = Y/\lambda d$ are spatial frequency coordinates in x and y directions (see Fig. 6.3).

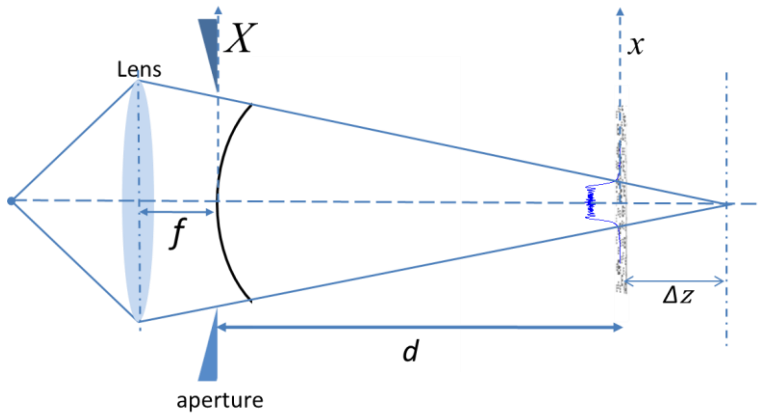


FIGURE 6.3 - illustration of the wave curvature at the condenser aperture plane due to defocus of the illumination in the Type-II setup.

Translation of the illumination relative to the object generates the diffraction patterns of the ptychographic dataset. Using the Fourier shift theorem, the illumination wave distribution at the back focal plane of the lens is modified by multiplication with an exponential phase ramp $\exp(i\pi \boldsymbol{R} \cdot \boldsymbol{k})$, which gives the convolution kernel as

$$P(\boldsymbol{k}) = \exp\left(i\pi\left(m\Delta x k_x + n\Delta y k_y\right)\right) \cdot A(\boldsymbol{k}) \exp\left(i\pi\Delta z\lambda\left(k_x^2 + k_y^2\right)\right), \tag{6.3}$$

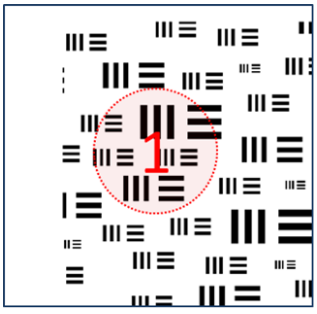
where Δx and Δy are sampling intervals along the x and y directions in the object plane and $\boldsymbol{R} = (m\Delta x, n\Delta y)$ is the translation vector of the illumination relative to the specimen. Grouping all the phase terms in a single exponent and dropping a constant phase term gives

$$P(\boldsymbol{k}) = A(\boldsymbol{k}) \exp\left(i\pi\Delta z\lambda\left((k_x + U_x)^2 + \left(k_y + U_y\right)^2\right)\right), \tag{6.4}$$

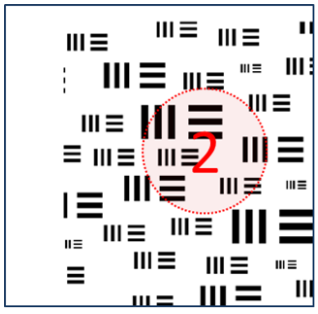
where $\boldsymbol{U} =: (U_x, U_y)$ represent the lateral translation vector for the phase of the kernel. Its components $U_x = \frac{m\Delta x}{2\Delta z\lambda}$ and $U_y = \frac{m\Delta y}{2\Delta z\lambda}$ are scaled values of the translation vector in real space.

The convolution kernel in Eq. (6.2) generates the diffraction pattern in the Type-II diffractive imaging experiment. In the case of ptychography, where a set of ptychographic diffraction patterns are measured from different probe positions, the kernel is modified by the probe position vectors (which quantifies the shift of the illumination), as shown in Eq (6.3). Furthermore, Eq. (6.4) indicates that this modification, in the case where the curvature is quadratic, can be viewed as a shift of the phase distribution of the convolution kernel, where the method of completing the square was used to transform Eq. (6.3) into Eq. (6.4). The magnitude of the kernel $A(\boldsymbol{k})$ represents the transmission profile of the aperture, which is a positive non-zero constant inside the aperture and zero outside.

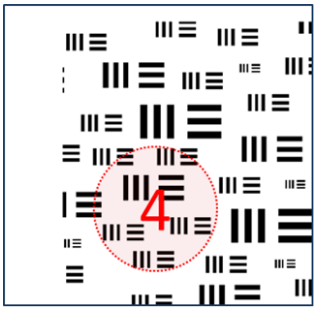
To visualise the effect of Eq. (6.4), the transmission function of a resolution target is illuminated at four different positions as shown in Fig. 6.4. The phase profiles of the convolution kernel at these positions are laterally displaced with respect to one another as shown in Fig. 6.5. The displacement vector is given by the variable ($\boldsymbol{U}$) in Eq. (6.4).



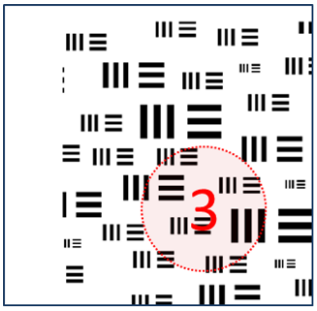
(a) illumination at position-1



(b) illumination at position-2

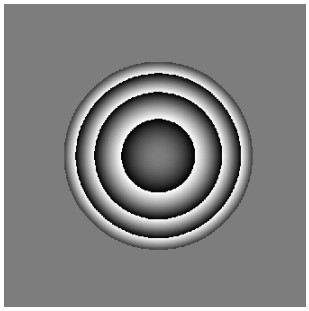


(d) illumination at position-4

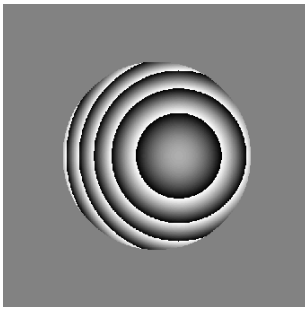


(c) illumination at position-3

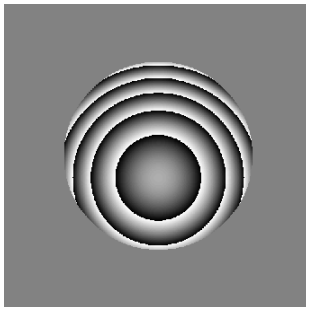
FIGURE 6.4 - Test object illuminated at various positions from which the ptychographic diffraction patterns of Fig. 6.6 were generated.



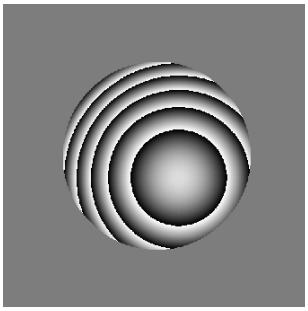
(a) Phase of convolution kernel from position-1



(b) Phase of convolution kernel from position-2



(d) Phase of convolution kernel from position-4



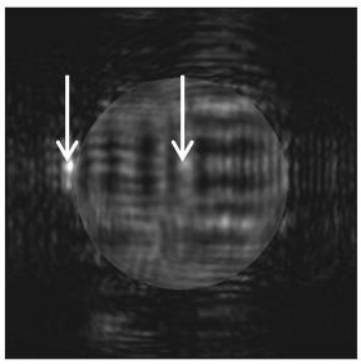
(c) Phase of convolution kernel from position-3

FIGURE 6.5 - Phase of convolution kernel from different illumination positions. The central disc spans a semi angle of 15 mrad.

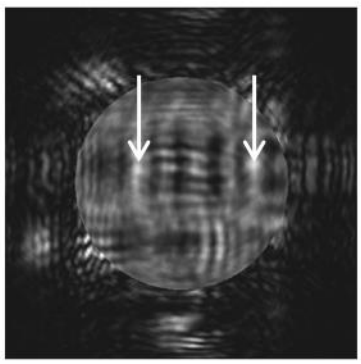
By analyzing the interaction amongst spatial frequencies whose separations are far smaller than the aperture size, the effect of the aperture can be ignored because the aperture is unity for all such interactions. In this case, the expression of information in the diffraction pattern is mainly governed by the phase distribution of Eq (6.4). Thus a lateral shift of the kernel's phase distribution, due to shift of the illumination, results in a lateral shift of the expression of

information. This means that the diffraction pattern in this idealised ptychographic dataset comprises intensities that are shifted relative to one another. However, the finite size of the aperture cannot be completely ignored, because the global structure of the diffraction pattern is determined by contributions from all the spatial frequencies within the aperture. As a result, information expression (in the form of diffraction intensities) has an additional facet together with the translation effect. The diffraction patterns in the ptychographic dataset encode translation of resolution information as well as a contribution from the finite size of the aperture. Hence, the diffraction patterns appear as distorted translated versions of one another. The clarity of the translation in the diffraction intensities from one diffraction pattern to another is greater when the translation vector of the illumination is small, because the effect of the aperture is minimal.

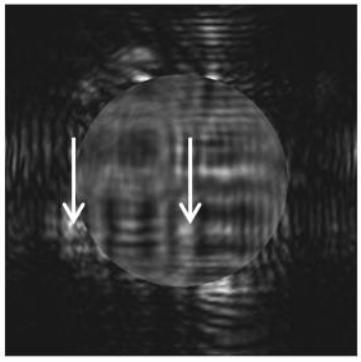
Fig. 6.6. shows the resulting diffraction patterns that are calculated with the illumination at the positions illustrated in Fig. 6.4. The motion of the large features within the central bright field region of the diffraction pattern is similar to the motion of the shadow image in the Gabor hologram, when the illuminated particle is moved laterally within the beam.



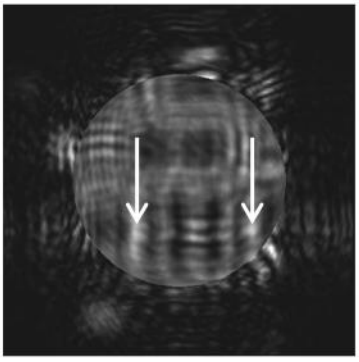
(a) diffraction pattern from position-1



(b) diffraction pattern from position-2



(d) diffraction pattern from position-4



(c) diffraction pattern from position-3

FIGURE 6.6 - Diffraction patterns generated from test object at corresponding positions from Fig. 6.4. The arrows show points in the diffraction pattern that encodes similar information about the object in the ptychographic dataset. The intensities in the dark field was increased by a factor of 10 relative to the bright field to illustrate movement of information between these regions

The arrows in Fig. 6.6 show points in the diffraction pattern that encodes similar information about the object in the ptychographic dataset. It also indicates that information can move from the bright field region of one diffraction pattern to the dark field region of another diffraction pattern in the ptychographic dataset where the magnitude of the translation is given by the vector $\boldsymbol{U}$ in Eq. (6.4). Section 6.4 explores the means for incorporating this translation of intensity values into resolution improvement for the PIE algorithm reconstructions, by processing a set of modelled bright field diffraction patterns.

6.3 Considerations for implementing resolution improvement scheme with the PIE algorithm

The translation of intensity values in Fig. 6.6 suggests that the size of the iterative calculation window should be larger than the size of any diffraction pattern of the ptychographic dataset, because data missing from one diffraction pattern is located in a different diffraction pattern

via the translation of information expression. Thus the PIE algorithm is extended to incorporate resolution improvement by using a larger calculation window in the iterative calculations. It is worth noting that in general the size of the iterative calculation window (in number of pixels) is matched to the size of the measured data. So in the case where one measures a diffraction pattern with a 1024x1024 detector, the size of the iterative phase retrieval calculation is matched to that of the detector (1024x1024). It is sometimes useful to bin the diffraction pattern to improve SNR of the measured signal or downsample the diffraction pattern to decrease processing time of the iterative algorithm. In either of these cases the resulting diffraction pattern usually satisfies the Nyquist sampling criterion. The analysis of information expression in ptychographic diffraction patterns presented in Section 6.2 suggests that the measured data size (in pixels) need not determine the size of the calculation window provided that a set of ptychographic diffraction patterns are recorded. This is because information pertaining to the region outside a given diffraction pattern resides in other measured diffraction patterns of the ptychographic dataset. This means that each measured diffraction pattern can be embedded in a larger of array of zeros, where the angular span of this combination decreases the sampling pitch at the specimen plane. The resolution of the recovered image does not automatically increase because the PIE algorithm needs to correctly recover the object representation on this finer sampling grid. The PIE algorithm can incorporate high spatial frequencies by letting the pixels outside the measured region of the diffraction pattern take on complex values estimated by the algorithm. These values change as the iterative calculations progress and approach a configuration consistent with the ptychographic dataset.

It is worth noting that the intensities outside the measured data should differ slightly from the intensities measured in an experiment. This is because the dataset in the Type-II setup contains spatial frequency information about a finite number of spatial frequencies of the

object; hence the final object solution has a finite bandwidth whereas in the actual experimental situation the bandwidth of the underlying object can be greater. This means that the diffraction patterns generated from both the recovered object, at the virtual part of the detector, should differ slightly from the intensities that would be measured in an experiment and the deviation should increase as one approaches the edge of the virtual detector. This is a very important point because if the diffraction patterns were identical, the recovered diffraction pattern could, in principle, be used to recover an even larger virtual detector. This procedure can be repeated to get an even larger virtual detector and the process continues without limit, which presents an anomaly that allows the creation of information from nothing. i.e. it suggests that in the Type-II experiment one can keep extending the diffraction pattern, but as mentioned earlier, the number of spatial frequencies encoded by the ptychographic dataset is limited by the semi-angle of the lens aperture.

This means that the algorithm needs to have a mechanism to ensure that the calculated diffraction patterns deviate from the measured diffraction patterns in just the right way to allow for consistency between the recovered object and the underlying object in the spatial frequency range captured by the detector. This departure should also take a configuration that avoids the anomaly of the process that allows unlimited extension of the diffraction pattern through hierarchical increase of the virtual detector size.

It turns out that such a mechanism exists in the PIE algorithm. To understand its origin, we need to take a closer look at the PIE update function, which is reprinted in Eq. (6.5) for convenience.

$$\begin{aligned} O_{n+1}[\boldsymbol{r}] &= O_n[\boldsymbol{r}] + P^*[\boldsymbol{r}]\left(\pi_M\{\psi_{s(n)}[\boldsymbol{r}]\} - \psi_{s(n)}[\boldsymbol{r}]\right) \\ O_{n+1}[\boldsymbol{k}] &= O_n[\boldsymbol{k}] + \sum_{S^{-1}(\boldsymbol{k};\boldsymbol{u})} P^*[\boldsymbol{u}-\boldsymbol{k}]\left(\pi_M\{\psi_{s(n)}[\boldsymbol{u}]\} - \psi_{s(n)}[\boldsymbol{u}]\right) \end{aligned} \tag{6.5}$$

To picture the self-limiting mechanism of the update function, consider the feedback error added to the object in Eq. (6.5). The region R1 of Fig. 6.7 defines the detector and corresponds to the part of the calculation array that is populated with measured data. Region R2 and R3 define the parts of the calculation array that are set to zero at the start of the iterative calculation. Bandwidth extrapolation is incorporated into the PIE algorithm by setting the values outside the measured data (R2 &R3) to the current estimate of the wavefield at the detector plane during the iteration process. This is equivalent to setting the values of $\pi_M\{\psi_{s(n)}[\boldsymbol{u}]\}$ in R2 and R3 to equal that of $\psi_{s(n)}[\boldsymbol{u}]$, since $\psi_{s(n)}[\boldsymbol{u}]$ is the current estimate of the wave field at the detector and the measured intensity updates the magnitude of the wave field at the central part, R1.

Put differently, to generate $\pi_M\{\psi_n[\boldsymbol{u}]\}$ only the central part of $(\psi_n[\boldsymbol{u}])$ is updated with measured data while the outer parts, R2 and R3, remain unchanged. This calculation scheme is called floating-dark-field (FDF) because the region outside of the measure data is allowed to float.

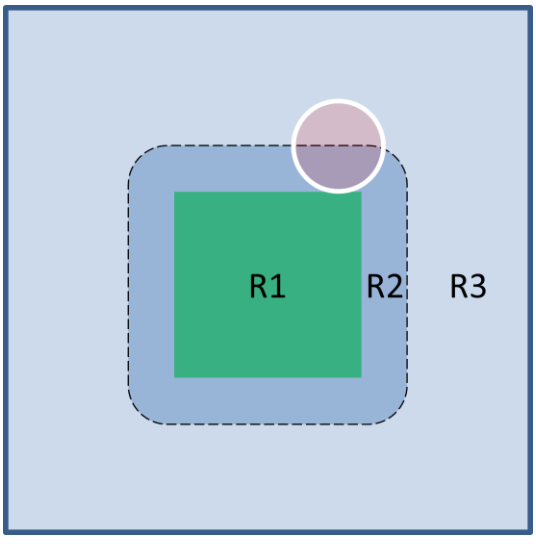


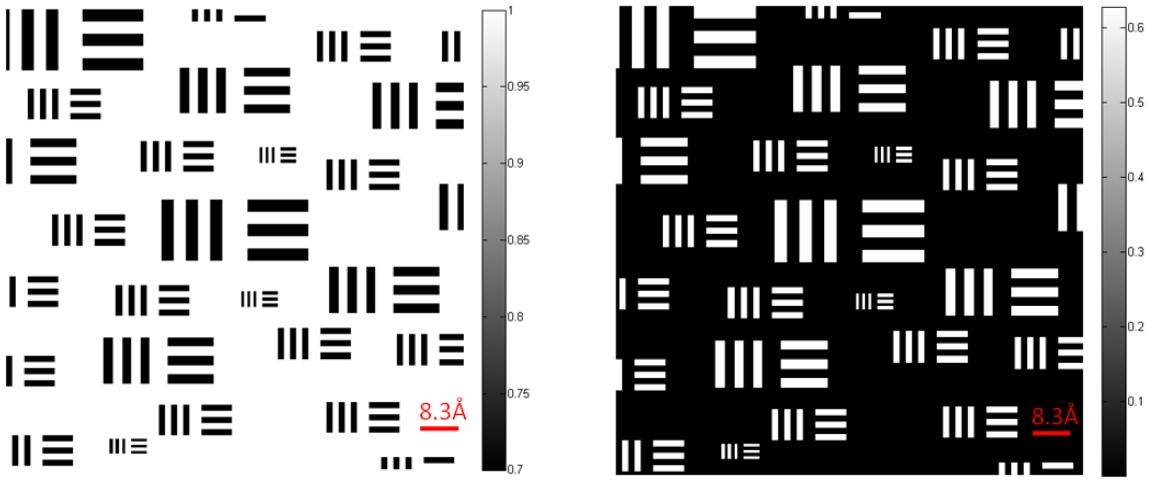
FIGURE 6.7 - Illustration of the convolution operation in the detector plane. The inner region (R1) marks the area of the detector. R2 correspond to intensities that can be inferred from R1. R3 refer to scattered intensities that are not encoded in the diffraction pattern in the Type-II setup.

The convolution operation corresponds to the region enclosed by the white circle in Fig. 6.7 ($S^{-1}(\boldsymbol{u})$) moving to every point in the detector plane where the spatial frequency of the recovered object represents the position vector of the centre of this region. The auxiliary variable $\boldsymbol{u}$, sums over all the points in the diffraction pattern that the object component ($\boldsymbol{k}$) contributes towards in the forward calculation; these points contain the necessary information to constrain that spatial frequency ($\boldsymbol{k}$). In the FDF scheme, the difference $\Delta\psi_{s(n)}[\boldsymbol{u}] = \pi_M\{\tilde{\psi}_{s(n)}[\boldsymbol{u}]\} - \psi_{s(n)}[\boldsymbol{u}]$, is non-zero only in R1. The feedback error ($\varepsilon[\boldsymbol{k}] = \sum_{S^{-1}(\boldsymbol{u})} P^*_{s(n)}[\boldsymbol{u}-\boldsymbol{k}]\Delta\psi_{s(n)}[\boldsymbol{u}]$) added to the object in Eq. (6.5) falls off to zero at the interface between R2 and R3 because the difference ($\Delta\psi_{s(n)}[\boldsymbol{u}]$) has non-zero values only in the region defined by R1. This means that the PIE algorithm automatically band limits the recovered object at the correct spatial frequency encoded in the intensity values measured by the detector and as a result avoids the problem associated with the unlimited but incorrect extension of the virtual detector. This also shows that the object recovered from the PIE algorithm is a band limited function whose size is equal to the size of the detector plus the full angle of the aperture as suggested in Section 6.1. This is an important insight because it suggests that floating values of $\tilde{\psi}_{s(n)}[\boldsymbol{u}]$ are equivalent to updating the object directly with feedback errors derived only from the measured data, as should be the case.

6.4 Implementation of bandwidth extrapolation with PIE

This section demonstrates the effectiveness of the FDF scheme. The object used for the computational investigation of bandwidth extrapolation with the PIE algorithm is shown in

Fig. 6.8. It is the complex transmission function of a resolution target. A notional spacing of 1.5Å separates the smallest bars in the object. This spacing scatters a 200 keV electron beam by an angle of 17 mrad. The detector was engineered to have the same angular size of 17 mrad.

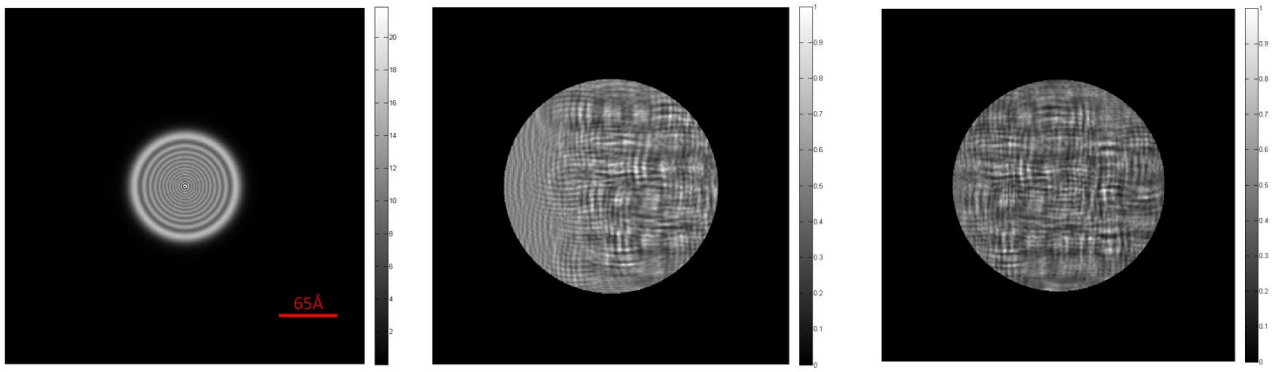


(a) Object magnitude (b) Object phase

FIGURE 6.8 - Complex transmission function representation of a resolution target.

Fig. 6.9b and Fig. 6.9c show the diffraction patterns calculated from the object at two illumination positions. The effective signal from the object was confined to the bright field disc by setting the values outside the Ronchigram to zero to model an experiment where the signal from the dark field of the diffraction pattern is masked by noise. The noise source can be a combination of detector shot noise, readout noise, Poisson noise etc. Poisson noise becomes critical in situations where the number of counts in each diffraction pattern need to be small in order to minimize the effect of specimen drift, damage or contamination during data collection. In all of these cases the signal to noise ratio (SNR) in the dark field is very much lower than that of the bright field. Moreover, using the poor/compromised dark field

signal result in distortions in the recovered object as incorrect a priori information is fed to the algorithm during the ptychographic reconstruction.



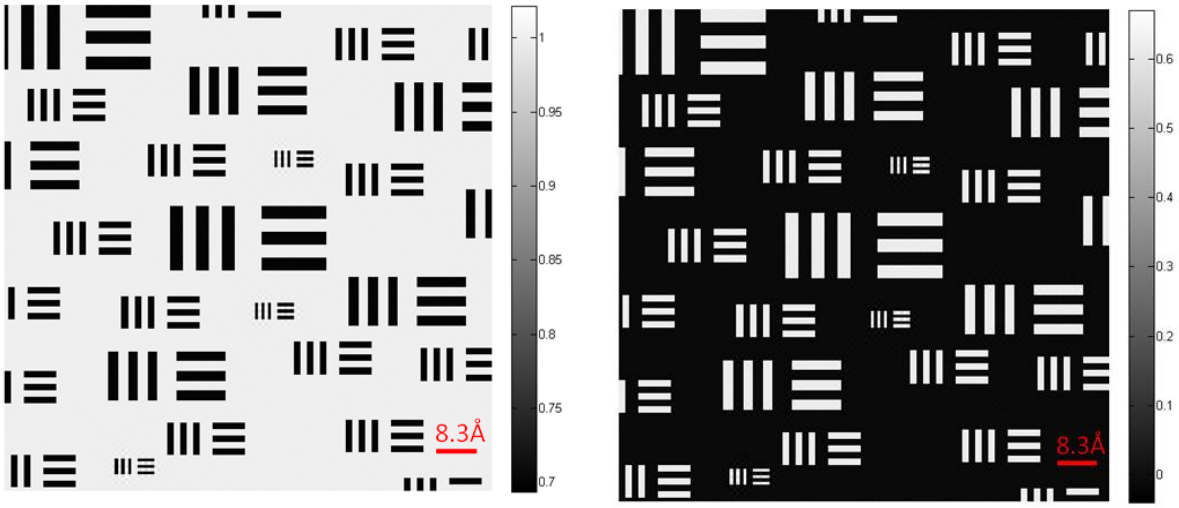
(a) Probe intensity at the specimen plane (b) Typical diffraction pattern at position 1 (c) Typical diffraction pattern at position 2

FIGURE 6.9 - (a) Probe intensity generated with a 200kV electron accelerating voltage, semi angle of convergence =10 mrad, defocus = -650nm. The bright field disc radius in (b) and (c) corresponds to a semi angle of 10 mrad.

The ptychographic dataset employed in this investigation comprises 81 diffraction patterns generated from illumination positions that correspond to the intersection points of a 9x9 square grid. The distance between adjoining points of the grid is 30% of the effective diameter of the illumination function. This gives a 70% overlap of the illumination positioned at these intersection points. In addition the illumination positions were slightly offset from the regular intersection points of the grid to avoid the emergence of periodic artefacts in the reconstruction. The illumination parameters model a 200 keV electron beam having a convergence semi-angle of 10 mrad and a defocus of -650 nm.

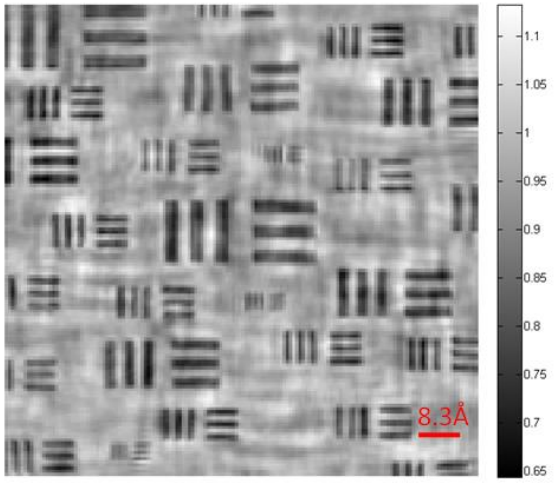
6.4.1 Results

Running the PIE algorithm for 1000 iterations generated the results shown in Fig. 6.10 to Fig. 6.12. Fig. 6.10 shows the magnitude and phase of the recovered object in the case where the dark field values in each diffraction pattern were updated by the PIE algorithm. The corresponding results from calculations where the dark-field values were not updated by the algorithm are shown in Fig. 6.11. For comparison, the dark field values of one of the original diffraction patterns, the corresponding recovered diffraction pattern using the FDF scheme and the error between them are shown in Fig. 6.12.

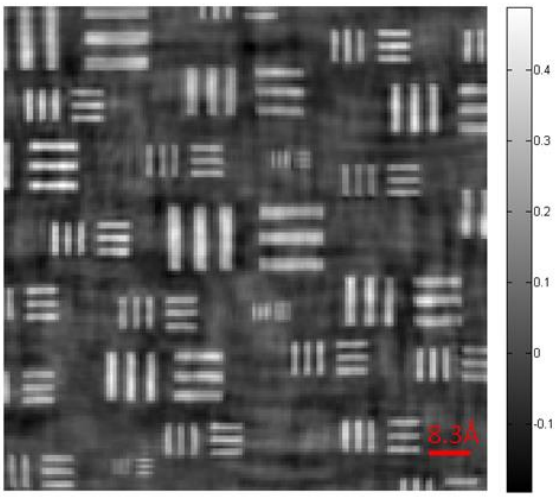


(a) Recovered object magnitude using FDF scheme (b) Recovered object phase using FDF scheme

FIGURE 6.10 - Recovered complex object in the case where the dark field values are updated during the ptychographic reconstruction.

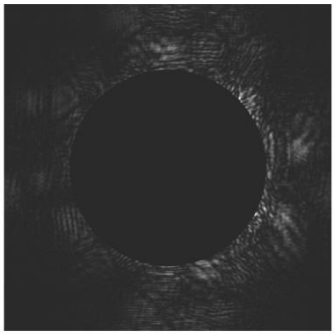


(a) Recovered object magnitude from set of Ronchigrams calculations

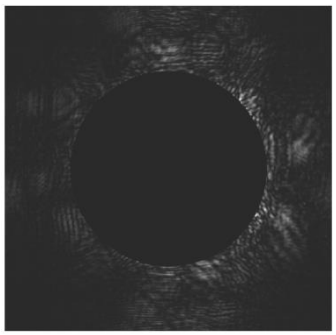


(b) Recovered object phase from set of Ronchigrams calculations

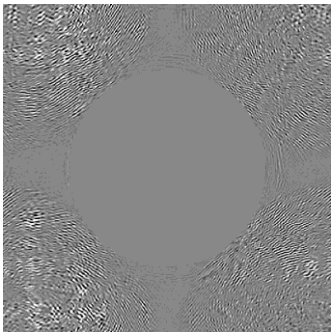
FIGURE 6.11 - Recovered complex object in the case where the dark field values are suppressed during the ptychographic reconstruction



(a) Original dark field



(b) Recovered Dark field



(c) Error in dark field

FIGURE 6.12 - (a) and (b) show a comparison of an original and recovered dark field region of a diffraction pattern. (c) shows the resulting difference map between (a) and (b) after 1000 iterations. The maximum of error value in (c), shown as white in the image, corresponds to 0.01% of the intensity value at that pixel. This means that the error in the recovered region is less than 0.01%.

6.4.2 Discussion

It is clear that the magnitude and phase reconstructions give an accurate representation of the specimen in the PIE iterative calculations where the dark field is allowed to float (see Fig. 6.11). The algorithm correctly recovers the finest details in the specimen corresponding to a periodicity of 1.5Å. This periodicity maps to a scattering semi-angle of 17 mrad that is greater than the size of the Rochigram used in the calculation. Since the Ronchigram spans a semi angle of 10 mrad, recovery of features finer than 2.6Å (corresponding to 10 mrad) indicates that the recovered object is a bandwidth extrapolated reconstruction from the set of Ronchigrams used in the calculations. Furthermore, the recovered object magnitude is in close agreement with the test specimen transmission magnitude and the phase has a constant offset, which was expected because only relative phase values are recoverable.

Fig. 6.12a and Fig. 6.12b show a comparison between one of the original and corresponding recovered dark field diffraction patterns in the ptychographic dataset. These two diffraction patterns are visually indistinguishable and have a similar intensity range. Fig. 6.12c shows the difference between these diffraction patterns and indicates a discrepancy of less than 1% error.

The quality of the reconstruction in Fig. 6.11 is worse than expected. It shows ripples in the magnitude and phase of the recovered object. This suggests that the PIE algorithm incorrectly recovers low-spatial frequencies of the object. This is due to inconsistency in the PIE reconstruction in the case where the zero values in the darkfield of the ptychographic diffraction patterns enforce incorrect *a priori* information during iterative calculations. This is a very important result. All ptychography calculations where the intensity has not fallen to zero at the edge of the diffraction pattern must be embedded in a larger array with floating pixels. An insight into the mechanism that results in this poor reconstruction is presented in the next section after the introduction of a feedback term that is required for ePIE reconstructions.

6.5 Considerations for implementing bandwidth extrapolation with the ePIE algorithm

In Section 6.3, it was shown that the PIE algorithm automatically band limits the spatial frequency content of the recovered object at the maximum frequency encoded in the diffraction pattern. This conclusion was reached by assuming the one has *a priori* information about the illumination function (thus the convolution kernel) and that the kernel does not change during iterative calculations. However the ePIE algorithm does not require complete a priori knowledge of the illumination as demonstrated in the Chapter 5. Moreover, the illumination function changes as the iteration progresses. i.e. $P_{n+1}[\boldsymbol{k}] = P_n[\boldsymbol{k}] + \sum_{\boldsymbol{u}} O^*[\boldsymbol{u}-\boldsymbol{k}]\Delta\psi_{s(n)}[\boldsymbol{u}]$. This means the convolution kernel ($P^*_{n+1}[\boldsymbol{u}-\boldsymbol{k}]$) also changes with iteration. As a result, automatic band limitation of the feedback error function ($\varepsilon[\boldsymbol{k}]$) at the correct spatial frequency cannot be guaranteed. This aspect of bandwidth extrapolation with the ePIE algorithm was first reported by Maiden et al (2011), where the edge of the diffraction pattern was clamped to constrain the calculation. This section introduces a feedback framework encompassing the method implemented by Maiden et al (2011). It also provides a means to incorporate other classes of solutions for the ePIE algorithm.

The ePIE algorithm updates the object and the illumination with the same feedback error function; as a result the running estimate of the object can contain non-zero values at spatial frequencies higher than the spatial frequency information encoded in the measured ptychographic dataset. Furthermore, the illumination may not converge to the right form because ePIE uses the running estimate of the object that contains high frequency noise;

where high frequency noise represent complex values of the recovered object or illumination functions at spatial frequencies outside the limits constrained by the measured data and arise because the illumination changes with iteration. This is in contrast with the stability of the PIE algorithm presented in Section 6.3

Region R3 in Fig. 6.7 represents spatial frequencies that are not constrained by the measured data. This suggests that engineering the calculation size to encompass region R1 and R2 only, may help constrain the algorithm. However, such a scheme requires exact knowledge of the aperture size. Moreover, the convolution is cyclic because the calculations are implemented with the FFT algorithm. This means that R3 serves as a buffer zone to prevent the artefact of cyclic convolution from affecting PIE and ePIE reconstructions.

Constraining the calculation size to R1 and R2 is also not a practical solution in situations where the convolution kernel does not have a finite boundary, as with other forms of structured illumination. In such situations, the full extent of the bandwidth extrapolated image is not known thus embedding the diffraction pattern in an array of zeros smaller than the effective spatial frequency of the recovered object will result in aliasing in reciprocal space because the calculation no longer has the buffer zone R3. The preceding discussion suggests that the ePIE algorithm needs an additional structure to cope with bandwidth extrapolation reconstructions. The most logical place to start is the update function in Eq. (6.5). For simplicity, we now limit the discussion to the Type-II setup and later generalise. Consider the form of the feedback error term of the update equation ($\varepsilon[\boldsymbol{k}] = \sum_{S^{-1}(\boldsymbol{k};\boldsymbol{u})} P^*[\boldsymbol{u} - \boldsymbol{k}]\Delta\psi_{s(n)}[\boldsymbol{u}]$). We require that this term be appropriately limited by the finite angular size of the convolution kernel, in the case of PIE calculations. However, this requirement cannot be guaranteed by the finite size of $\Delta\psi_{s(n)}[\boldsymbol{u}]$, because the convolution kernel changes with iteration, thus, a different way of enforcing the appropriate band limit is required. A crucial

feature of any candidate solution to this problem is a mechanism that provides a means to suppress high frequency noise in the object or illumination. As mentioned earlier, appropriate band limitation of the object should suffice. However, rather than explicitly imposing a band-limiting filter on the object, we need a different way of band limiting the object that works with convolution kernels of different aperture sizes or a generally structured kernel.

The solution lies with the form of the update function in real-space, i.e. the object needs to be updated with a reciprocal space weighting function (RSWF) that highly weights the low spatial frequencies relative to the high spatial frequencies and also curbs any divergence of high-spatial frequencies that are not constrained by the measured data. In this fashion all the spatial frequencies present in the measured data are expressed in the object, where low spatial frequencies information emerges very early in the iterative calculations and high spatial frequencies emerges as the iteration progresses.

The number of iterations required for the highest spatial frequencies to fully express their correct contribution in the recovered object should depend on the form of the RSWF. The RSWF needs to also provide a negative feedback mechanism that suppresses high frequency noise for frequencies not constrained by the measured data. The noise arises because of the departure of earlier estimates of the convolution kernel from the correct value.

The ePIE algorithm incorporates the aforementioned properties with a single feedback parameter β, where the updated estimate of the complex wave field at the detector plane $\psi_{s(n)}[\boldsymbol{u}]$ is multiplied by $1-\beta$ in R2 and R3. Note that setting $\beta = 0$ reverts the algorithm to the original FDF ePIE structure. To illustrate the effect of the feedback term on the object function in R3, consider the feedback error term given by

$$
\begin{aligned}
\varepsilon_{s(n)}[\boldsymbol{k}] &= \sum_{\boldsymbol{u}\in \boldsymbol{R3}} P^{*}_{s(n)}[\boldsymbol{k}-\boldsymbol{u}]\left(\tilde{\psi}_{s(n)}[\boldsymbol{u}]-\psi_{s(n)}[\boldsymbol{u}]\right), \\
&= \sum_{\boldsymbol{u}\in \boldsymbol{R3}} P^{*}_{s(n)}[\boldsymbol{k}-\boldsymbol{u}]\left(\psi_{s(n)}[\boldsymbol{u}](1-\beta)-\psi_{s(n)}[\boldsymbol{u}]\right), \\
&= -\beta\cdot\left(\sum_{\boldsymbol{u}\in \boldsymbol{R3}} P^{*}_{s(n)}[\boldsymbol{k}-\boldsymbol{u}]\psi_{s(n)}[\boldsymbol{u}]\right), \\
&= -\beta\cdot\left(\sum_{\boldsymbol{u}\in \boldsymbol{R3}} P^{*}_{s(n)}[\boldsymbol{k}-\boldsymbol{u}]\left(\sum_{\boldsymbol{k}'} P_{s(n)}[\boldsymbol{u}-\boldsymbol{k}']O_n[\boldsymbol{k}']\right)\right), \\
&= -\beta\cdot\left(\sum_{\boldsymbol{k}'}\left(\sum_{\boldsymbol{u}\in \boldsymbol{R3}} P^{*}_{s(n)}[\boldsymbol{k}-\boldsymbol{u}]P_{s(n)}[\boldsymbol{u}-\boldsymbol{k}']\right)O_n[\boldsymbol{k}']\right), \\
&= -\beta\cdot\left(\sum_{\boldsymbol{k}'}\eta[\boldsymbol{k}-\boldsymbol{k}']O_n[\boldsymbol{k}']\right), \\
&= -\beta O_n[\boldsymbol{k}].
\end{aligned}
\tag{6.6}
$$

In the second step, we make the substitution $\tilde{\psi}_{s(n)}[\boldsymbol{u}] = \psi_{s(n)}[\boldsymbol{u}](1-\beta)$. The fourth step expresses the current estimate of the wave field at the detector plane in terms of the object function and the final step substitutes the inner product of the convolution kernel with its conjugate.

In most cases, the resulting coupling function approaches a delta function ($\eta[\boldsymbol{k}-\boldsymbol{k}'] \approx \delta[\eta[\boldsymbol{k}-\boldsymbol{k}']]$) as explained in Chapter 2. This indicates that the error added to the object is a negative proportion of its previous value weighted by the feedback parameter β. As the iteration progresses, erroneous spatial frequencies components in R3 get suppressed by the algorithm through the mechanism outline in Eq. (6.6). The erroneous components are initially added to the object at the start of the iteration process when the illumination function is further from the correct function. These errors are not reinforced by measured data and are subsequently removed by the feedback mechanism.

This feedback process also occurs in the lower spatial frequency region (R1 & R2) but with positive contributions from the measured data in R1. In this case, the contribution from the measured data adds information about the object spatial frequencies in R2 and the feedback mechanism removes an amount proportional to β. As a result the resolution information in R2 does not get fully expressed in the reconstruction for cases where β is very large. A side effect of a large feedback parameter ($\beta = 1$) can be discerned from the results in Fig 6.11, which is further explained in the next section.

6.5.1 Explanation of poor quality reconstructions in Fig. 6.11.

Consider a single diffraction pattern of the ptychographic dataset employed in Section 6.4, where each diffraction pattern is a Ronchigram, so that all the scattered intensity values are confined to the central bright field disc. The PIE algorithm retrieves object information by adding the running estimate of the feedback error function given by Eq. (6.6). The iterative calculations where the dark-field values of the diffraction patterns are not updated using the FDF scheme is equivalent to having a feedback term of $\beta = 1$ for the entire diffraction pattern in the PIE algorithm. This is because the difference between the updated wave distribution ($\tilde{\psi}_{s(n)}[\boldsymbol{u}]$) and the guessed wave distribution ($\psi_{s(n)}[\boldsymbol{u}]$) in the detector plane takes the form outlined in Eq. (6.7) where the values of the updated wave are zeros outside the Ronchigram (i.e. in region labelled R_2 in Fig. 6.13).

$$\begin{aligned} \Delta\psi_{s(n)}[\boldsymbol{u}] &= \tilde{\psi}_{s(n)}[\boldsymbol{u}] - \psi_{s(n)}[\boldsymbol{u}] \\ &= 0 - \psi_{s(n)}[\boldsymbol{u}] \end{aligned} \tag{6.7}$$

$$= -\psi_{s(n)}[\boldsymbol{u}]$$

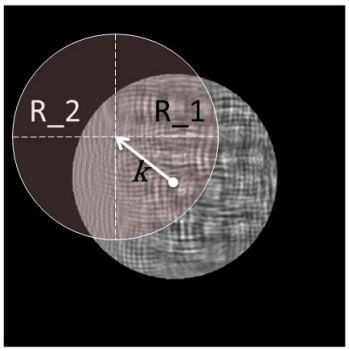


FIGURE 6.13 - The Ronchigram corresponds to the region labelled R_1 and the outside of the Ronchigram corresponds to the region labelled R_2. The circle represents the boundary of convolution kernel. The displacement from the centre of the diffraction pattern represents the spatial frequency ($\boldsymbol{k}$) that is updated by the convolution operation shown in the figure.

Fig. 6.13 shows that for all displacement vectors ($\boldsymbol{k}$) in the diffraction plane, the convolution kernel intersects both R_1 and R_2. Since contributions from R_2 can be interpreted, via Eq. (6.6) and (6.7), as feedback suppression by the algorithm, this suggests that the algorithm suppresses all spatial frequencies of the object except the DC term during the iterative calculations. This means that Fig. 6.10 illustrates the effect of spatial frequencies suppression by the algorithm. As expected, the feedback suppression does not completely remove the spatial frequencies information of the recovered object. However, it does appear to distort spatial frequencies expression in the recovered object for a large value of the feedback parameter.

6.6 Experimental curved illumination ePIE calculations with resolution improvement

The analysis presented in Section 6.3 shows that the resolution improvement scheme is an essential part of the PIE algorithm because information outside the detector aperture cannot be ignored during the iterative calculations. This section investigates the influence of updating the illumination while solving for a bandwidth extrapolated object representation.

The experimental setup employed in this investigation is shown in Fig. 6.14. A resolution target is used as the specimen in the ptychographic experiment. An iris is used to create an aperture that is imaged on the detector. A total of 225 diffraction patterns are collected from 225 overlapping regions of the specimen. The translation vector of the illumination at adjacent position generates a 70% overlap dataset. The initial estimate of the object is an image array of ones, corresponding to free space.

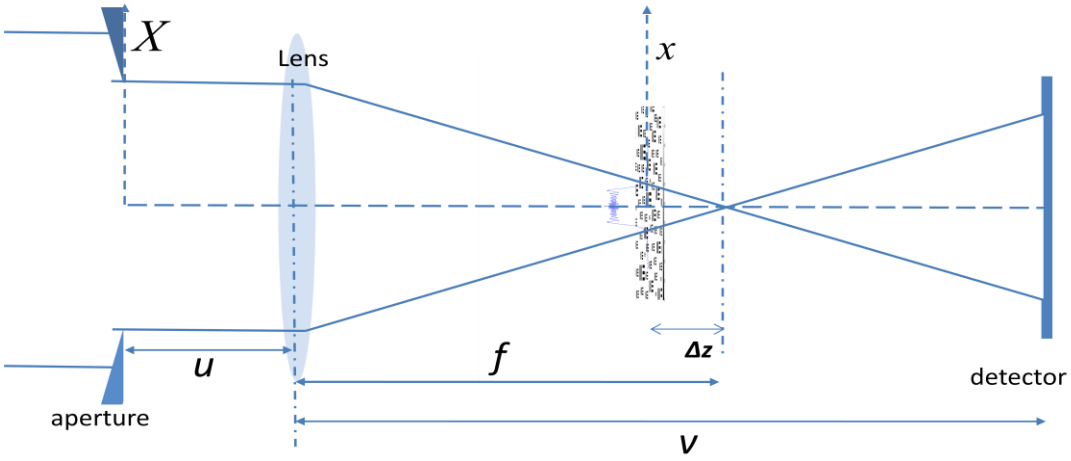


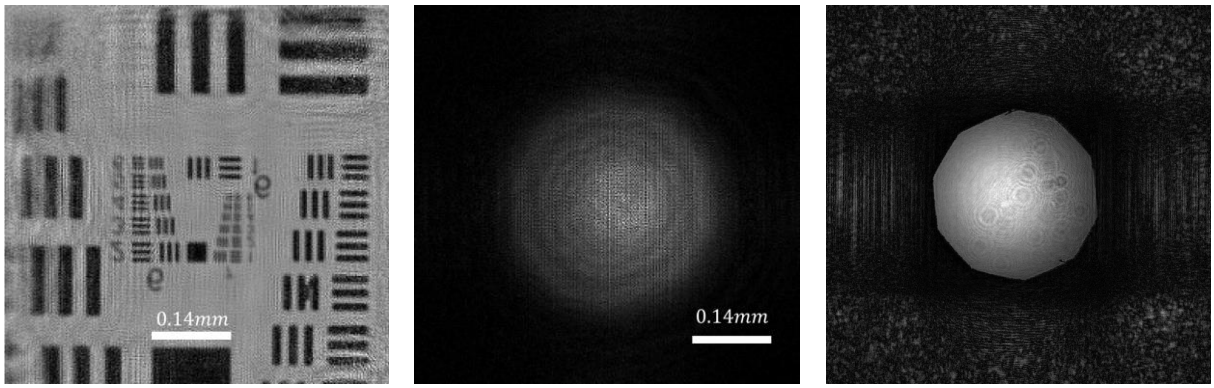
FIGURE 6.14 – The aperture and the detector planes satisfy the imaging condition $\left(\frac{1}{u}+\frac{1}{v}=\frac{1}{f}\right)$ for the lens. The aperture is illuminated by a collimated beam from a laser source with a wavelength of 675 nm. $v = 17.7\ cm$, $u = 4.4\ cm$, $f = 3.5\ cm$, $\Delta z = 1.6cm$, detector pitch = 7.4 μm. The detector comprises 1024x1024 pixels

The ePIE algorithm was run for three different cases. The first case corresponds to the FDF scheme, that is to say the values outside the measured data are solved by the ePIE algorithm. This is done to test whether the algorithm could accurately constrain the recovered

illumination and object functions for spatial frequencies that are not encoded in the diffraction pattern. The second case investigated the impact of explicit band limitation on the recovered object. It was suggested in Section 6.5 that this should also provide a means of constraining the algorithm. In the final case the feedback parameter is used as a constraint for the algorithm.

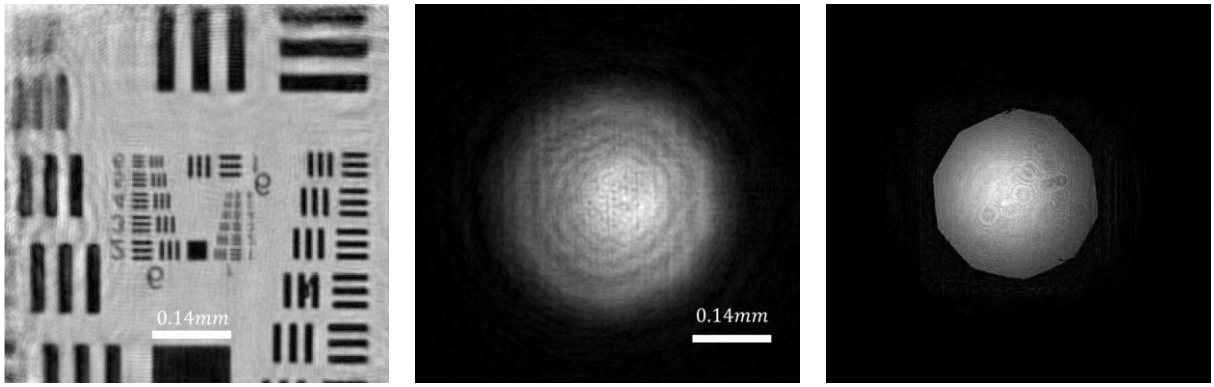
6.6.1 Results

Fig. 6.15a and Fig. 6.15b show the recovered magnitudes of the resolution target and illumination function respectively for the case where the dark field is allowed to float. Fig. 6.15c is a magnitude image of the FFT of the recovered illumination. Fig. 6.16 and Fig. 6.17 show corresponding results for band limiting the object and the use of the feedback framework.



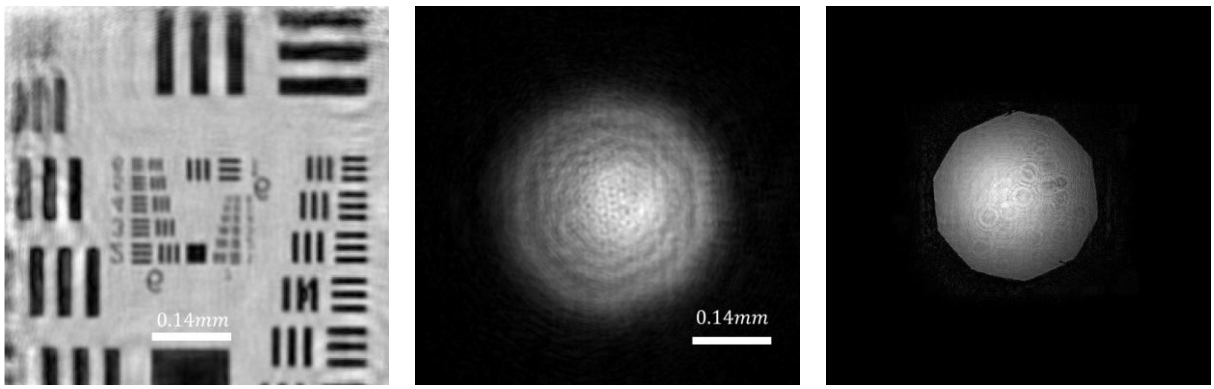
(a) (b) (c)

FIGURE 6.15 - (a) shows the recovered object using ePIE where the dark field is allowed to float. The resolution of the recovered object is highest in this calculation, but the recovered object and illumination function contain graininess that results from the unconstrained nature of the iterative calculations (b) shows the magnitude of the recovered illumination function, which is also plagued by high frequency noise. (c) shows the magnitude of the Fourier transform of the recovered illumination 1000 iterations. The high frequency noise is clearly shown as prominent intensity values in the outer rectangular region because this region has a transmission coefficient of zero in the aperture plane.



(a) (b) (c)

FIGURE 6.16 - (a) shows the recovered object using ePIE where a band limiting filter is applied to the FFT of the object to prevent the illumination from diverging at high frequencies. (b) shows the corresponding recovered illumination and (c) shows the magnitude of the Fourier transform of the recovered illumination after 1000 iterations. Low intensity values in the outer region suggests that band limitation does constrain the ePIE algorithm



(a) (b) (c)

FIGURE 6.17 - (a) shows the recovered object using ePIE where the feedback parameter β is set to unity to prevent the illumination from diverging at high frequencies. The high resolution details are not as sharp at those of Fig. 6.16, where the calculation is band limited at a spatial frequency determined by the sum of the detector size and the semi-angle of the aperture. (b) shows the corresponding recovered illumination and (c) shows the magnitude of the Fourier transform of the recovered illumination after 1000 iterations. Low intensity values in the outer region suggests that the feedback parameter does constrain the ePIE algorithm.

6.6.2 Discussion

In Section 6.3, it was suggested that an ePIE calculation with bandwidth extrapolation is unconstrained at very high spatial frequencies because the measured data contains no information about these frequencies. As a result updating the illumination with the running estimate of the object (which is unconstrained at high frequencies) introduces artefacts into the recovered illumination function. The manifestation of this phenomenon is apparent in the recovered object and probe functions in Fig. 6.15. These artefacts appear as graininess in the recovered image. This is a very interesting result, since it clearly indicates that the artefacts are restricted to only high spatial frequencies. One could have expected the convolution process to eventually move the artefacts into lower spatial frequencies as the iteration progresses, but it seems the measured data sufficiently constrains the spatial frequency range spanned by the detector.

Fig. 6.16 shows that band limiting the recovered object at the right spatial frequency ($\alpha + \alpha_p$) determined by the detector semi-angle and the semi-angle of the illumination prevents the emergence of artefacts in the reconstruction. Since knowledge of the wave vectors forming the illumination may be unavailable it is necessary to provide a mechanism that automatically constrains the ePIE algorithm and the results in Fig. 6.17 represent such a case. The feedback parameter with value $\beta = 1$ is used to test the effectiveness of the feedback mechanism for appropriately constraining the algorithm. The absence of high frequency artefacts in the result of Fig. 6.17 confirms that the feedback mechanism, which is implicit in the structure of PIE, does indeed provide the necessary constraints in regions where the measured data fails to constrain the calculations. However, the resolution of the recovered object is slightly reduced by the feedback mechanism. The feedback scheme employed here uses only one parameter, which means that the algorithm is not optimised. For example, an alternative

scheme such as clamping of the edge to the diffraction pattern as implemented by Maiden et al (2011) corresponds to a two parameter feedback configuration with $\beta = 0$ in R2 and $\beta = 1$ in R3, where R3 corresponds to the border of the virtual detector. It is expected that feedback scheme may be tailored to the detail of a particular experimental setup. These results also suggest that the departure of the illumination function at high $\boldsymbol{k}$-vectors can also be used as a metric for the convergence of the ePIE algorithm, which should complement the usual mean squared error metric.

6.7 Sub-Nyquist ptychographical calculations

This section investigates the limit of the PIE algorithm in the case where the diffraction patterns of the ptychographic dataset are under-sampled. One of the motivations for this investigation is to further explore the connection between sampling (or oversampling) and solubility of the phase problem. For example, the calculations in Section 4.4 showed that the algorithm could extract some information about the object from a single under-sampled diffraction pattern. Furthermore, the error metric and the quality of the reconstructions for several defoci values in Section 4.5, suggests that the algorithm performs better in cases where the illumination size is large before the effect of aliasing reduces the quality of the reconstruction. Thus, if the effects of aliasing are suppressed, can the undersampled diffraction pattern provide the necessary information required for a good reconstruction of the underlying object?

This section aims to use the diversity and redundancy of the ptychographic dataset to help alleviate the effect of aliasing in a set of under-sampled diffraction patterns. To accomplish

this task, we need to explore ways to minimize the effect of aliasing. Aliasing of the illumination function occurs because the calculation window used to represent the update region in the specimen plane is smaller than the illumination size. Thus to avoid aliasing, the detector plane needs to be resampled with a finer sampling interval that gives a corresponding real space window larger than the effective size of the illumination function.

To investigate the effect of undersampling, this section samples the detector plane at a rate greater than the Nyquist rate to minimise the effect of aliasing; however the diffraction intensities only fills a subset of the detector plane as illustrated in Fig 6.18. A down-sampling-factor (DSF) of 2 corresponds to the illustration in Fig. 6.18a where diffraction intensities maps to a diffraction pattern sampled at twice the sampling interval of the detector plane. In the ptychographic calculations, the PIE algorithm solves for the complex values in regions that do not have intensities using the same method that was applied to resolution improvement reconstructions, the FDF scheme. Here we use the FDF scheme to recover both the magnitude and phase information at pixels represented by black boxes in Fig. 6.18.

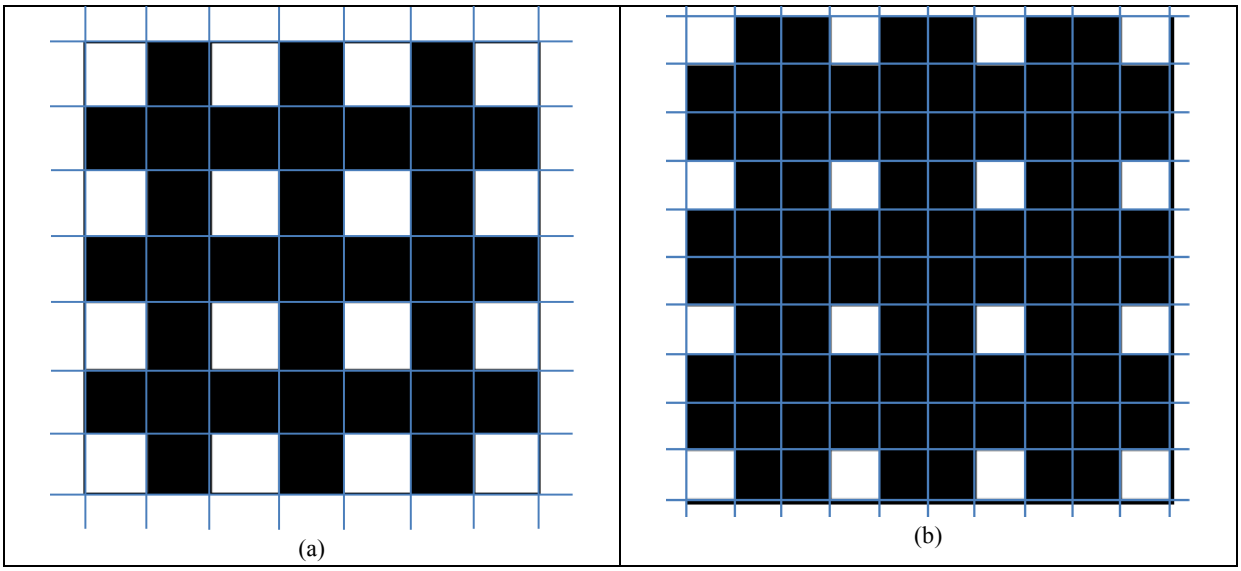


FIGURE 6.18 – illustration for downsampling of the detector plane. The white boxes represent pixels with measured intensities and black boxes represent pixels with no information. The algorithm recovers both the magnitude and phase information in pixels represented by black boxes (a) Diffraction pattern is downsampled by a factor of 2. (b) Diffraction pattern downsampled by a factor of 3.

The reconstruction strategy illustrated in Fig. 6.18 corresponds to the case where the detector has a fill factor of 0%, so that the pixels of the detector correspond to delta sampling functions. The case where the detector pixels bin the data by integrating over a range of angles to give a 100% fill factor, has been investigated by Batey, Edo and Rodenburg (2011). In this scheme, the updated wavefield in the detector plane takes the form

$$\tilde{\psi}(\boldsymbol{u}) = \psi(\boldsymbol{u}) \left(\sum\nolimits_{u'} |\psi(\boldsymbol{u}')|^2 \right)^{-1/2} \sqrt{I(\boldsymbol{u})}, \tag{6.8}$$

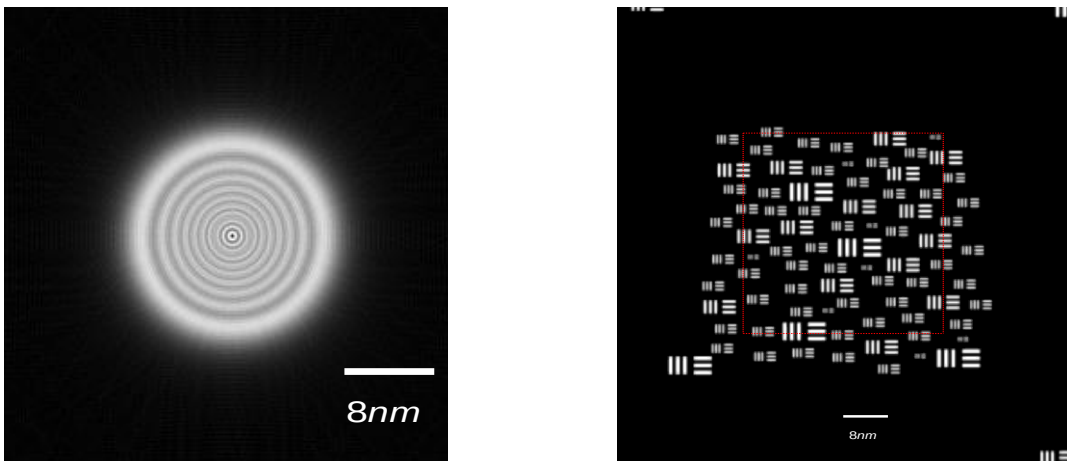
where the diffraction patterns are upsampled using nearest neighbour interpolation that defines a block/subset of intensity values spanned by the coordinate $\boldsymbol{u}'$, in which a single diffraction pattern intensity value is replicated. Thus, a single diffraction pattern intensity value updates a subset of the running estimate of the propagated wavefield at the detector plane; this imposes the correct power constraint for each block of the propagated wavefield and can be compared to the method proposed by Clark (2011) to account to partial coherence in the Type-I setup. This comparison is valid only for the Type-I setup that employs a source with a top-hat profile; in this configuration the binned data corresponds to the sub-sampled partial coherent diffraction pattern, with source size equal to the pixel pitch of the detector, where both are measured by the angle they subtend at the specimen plane. The remainder of this section presents calculations that investigate the case where the detector has a 0% fill factor and the reconstruction strategy corresponds to that illustrated in Fig. 6.18.

Fig. 6.19a shows the illumination function employed in this calculation. It models a large defocussed 200 keV electron beam generated with a 5 mrad aperture. The distance from the specimen plane to the lens focus gives a defocus of 2000 nm. The test object used in these calculations introduced a maximum phase change of 0.9π into the electron beam at the specimen plane and has no magnitude variations. Fig. 6.19b shows the phase distribution of

the test object. The ptychographic dataset had a 70% overlap parameter and is composed of 81 diffraction patterns.

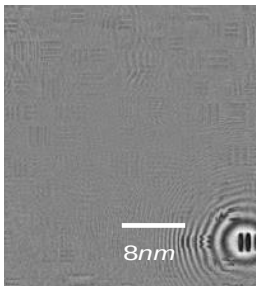
6.7.1 Results

Fig. 6.20 shows the recovered objects for cases where the diffraction pattern was downsampled by factor of 2 to 5. In the case where the DSF equals 2, the number of unknown pixels in the diffraction pattern is increased by a factor of 4 and for the case of DSF equal 5, the number of pixels in the diffraction pattern is increased by a factor of 25. The normalised mean squared error of the diffraction pattern at measured points is shown in Fig. 6.21a and the normalised mean squared error of the diffraction pattern at floating points is shown in Fig. 6.21b

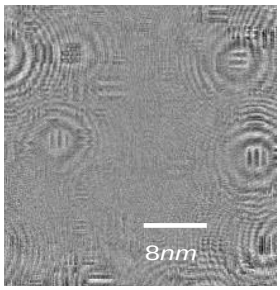


(a) (b)

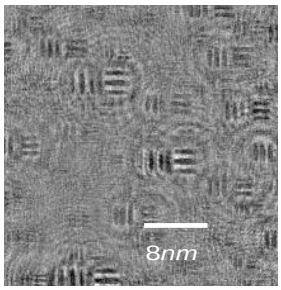
FIGURE 6.19 - (a) Magnitude of illumination function that models a 200 keV electron beam with defocus of 2000 nm and a convergence semi-angle of 5 mrad. (b) Phase of test object. The background in black represent a phase value of zero and the bars introduce a maximum phase change of 0.9π.



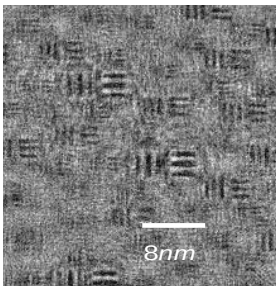
(a) Magnitude of recovered object from diffraction patterns downsampled by a factor of 2. Minimum transmission coefficient of recovered object = 0.9.



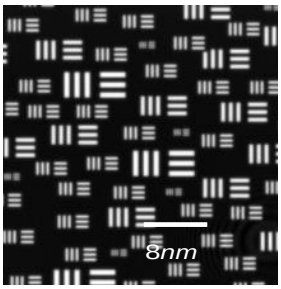
(b) Magnitude of recovered object from diffraction patterns downsampled by a factor of 3. Minimum transmission coefficient of recovered object = 0.7.



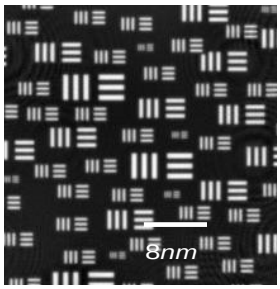
(c) Magnitude of recovered object from diffraction patterns downsampled by a factor of 4. Minimum transmission coefficient of recovered object = 0.5.



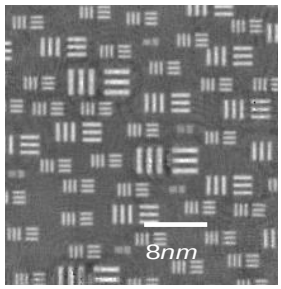
(d) Magnitude of recovered object from diffraction patterns downsampled by a factor of 5. Minimum transmission coefficient of recovered object = 0.2.



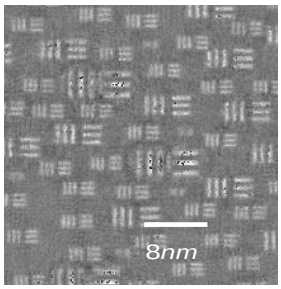
(e) Phase of recovered object from diffraction pattern downsampled by a factor of 2



(f) Phase of recovered object from diffraction pattern downsampled by a factor of 3

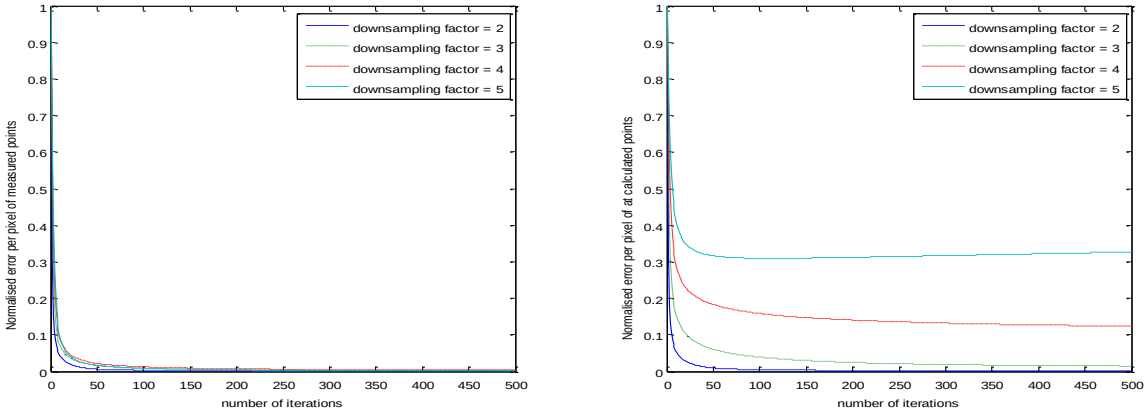


(g) Phase of recovered object from diffraction pattern downsampled by a factor of 4



(h) Phase of recovered object from diffraction pattern downsampled by a factor of 5

FIGURE 6.20 - Recovered transmission function of phase object from diffraction pattern downsampled by different factors. (a) – (d) Magnitude of recovered objects using DSF of 2 to 5. (e)-(f) Phase of recovered object corresponding to the calculations in (a)-(d). These results in (f) and (g) show that the PIE algorithm can recover specimen information from a set of undersampled ptychographic diffraction patterns, which suggests that oversampling of the diffraction pattern is not a requirement for the solubility of the phase problem in ptychography. The phase reconstruction in (h) shows that the PIE algorithm can still extract some of the specimen information from iterative calculations where the number of unknown pixels in each diffraction pattern is a factor of 24 of the diffraction data.



(a) (b)

FIGURE 6.21 – (a) The normalised MSE between the calculated intensities and the original data for detector pixels that were used to constrain the calculation. It is very interesting to see that the MSE is constrained in the region where the intensity values are known, as it suggests that the number of unknown intensity values in the diffraction pattern does not affect the performance of the PIE algorithm in this region of the detector plane. (b) The normalised MSE between the calculated intensities and the original data detector pixels that were set to zero at the start of the iteration. The plots show that the PIE algorithm constrains the calculation in all the detector pixels for calculations where the downsampling factor of the detector plane has a value less the 5.

6.7.2 Discussion

The downsampled diffraction pattern with a DSF of 2 corresponds to the case where the diffraction pattern is sampled at the Nyquist rate of the complex wave. This means that the diffraction intensity is undersampled by a factor of 2. The results in Fig. 6.20a and Fig. 6.20e show that ptychographic calculations can be performed at Nyquist sampling of the complex wavefield. The error plots in Fig. 6.21 show that the calculations for DSF of 2 are properly constrained.

The results in Fig. 6.20b and Fig. 6.20f correspond to sub-Nyquist sampled diffraction patterns. The recovered magnitude contains more variation in the magnitude compared to result in Fig. 6.20a; however the phase part of the reconstruction gives a good representation of the object. This means that the PIE algorithm can recover the object from sub-Nyquist

sampled diffraction patterns using the FDF scheme. It is important to note that these calculations are only possible because of the diversity of the ptychographic dataset that constrains the calculations at pixels in between the sampling points of the known data.

The final cases where the diffraction patterns are undersampled by a factor of 4 and 5, the PIE algorithm recovers features of the object, which shows that there is no sharp sampling pitch for which the PIE algorithm fails. The corresponding error plots in Fig. 5.21a show that the MSE values after 500 iterations are higher in these cases. Furthermore, the error plot in Fig. 5.21b for the case with a DSF value of 5, shows that the error for the recovered points increases after 150 iterations; this can be used as an indicator of the largest value of DSF that should be used in the PIE algorithm. These calculations demonstrate that the PIE algorithm together with the FDF scheme provides a practical means for processing undersampled diffraction patterns. This is amazing, since it shows that oversampling is not the main reason for the solubility of the phase problem.

6.8 Summary

This chapter examined a practical implementation of a resolution improvement scheme using the PIE and ePIE algorithms. The PIE algorithm is shown to have the correct framework required for bandwidth extrapolation calculations because of the feedback mechanism inherent in the PIE algorithm. This is extended to constrain the ePIE bandwidth extrapolation calculations, where the feedback parameter of unity is employed. The last section of this chapter demonstrates that sub-Nyquist ptychographic diffraction patterns provide good

quality reconstruction of complex objects, when these diffraction patterns are processed using the FDF scheme, in which the algorithm recovers the missing values of the diffraction patterns. It is extraordinary that this finding completely supersedes all previous discussions of oversampling in coherent diffractive imaging.

Chapter 7

7 Conclusion

This thesis has undertaken the first detailed analysis of the phase recovery algorithm (PIE), which has been widely adopted and is currently used at five X-ray synchrotrons worldwide. Its experimental development is being actively pursued in the case of electron imaging. Most of this thesis has been directed at understanding the specific embodiment of the technique required in the electron application. This research is timely because the PIE algorithm and its extension, the ePIE algorithm, have been widely adopted solely because they work, but the underlying mechanism that underpins their operation is not understood. The analysis presented here shows that these algorithms extract the object information from the diffraction pattern in a way that minimizes the phase distortion of the recovered object's spatial frequencies. It also provides an explanation for the process through which these algorithms solve the phase problem in a general object and demonstrates that the current update scheme provides a robust foundation on which advanced methods such as bandwidth extrapolation and 3D object reconstruction can build.

The results for low counts calculations show that experimental uncertainties such as specimen drift, contamination, and damage can, in principle, be accommodated in electron ptychographic experiments, because the PIE and ePIE algorithms provide good quality reconstructions from a large set of noisy diffraction patterns, which are collected with reduced dwell time of the illumination at any part of the specimen. This thesis also showed that the ePIE algorithm can completely characterise two-fold astigmatism in the electron microscope probe and can accurately recover the complex profile of the illuminating wave from guess illumination functions that are within a factor of two of the correct probe size. This is a significant finding that highlights the strength of this imaging method and it also provides a bound for the accuracy of the illumination parameters needed in the Type-II ptychographic calculations.

One facet of the diversity of the ptychographic dataset is shown to correspond to the angular span of the set of phase gradients illuminating any given part of the object, from the illumination located at different probe positions. This provides an invaluable tool for further analysis of results from ePIE calculations. This thesis also demonstrates that the PIE algorithm has the required framework to correctly extrapolate the bandwidth of the recovered object function in resolution improvement calculations which facilitates double resolution improvement even in the very best electron microscope. A feedback method was developed that incorporates the general class of techniques that are used to constrain resolution improvement calculations with the ePIE algorithm. Finally, the PIE algorithm is shown to correctly recover a complex object from sub-Nyquist sampling of diffraction patterns. It is extraordinary that this finding completely supersedes all previous discussions of oversampling in coherent diffractive imaging.

7.1 Future Work

Additional research into different class of techniques that use the feedback framework presented in this thesis, to constrain the ePIE algorithm, in resolution improvement calculation, is required. This will determine whether a unique class of feedback technique can be used for a generalised illumination type, or whether the feedback method needs to be optimised for different experimental configurations.

The work investigated in this thesis considers the case where the detector comprises pixels with 0% fill factor, in which the algorithm solves for the complex wavefield at points with missing intensity values. It was also shown that an alternative strategy could be employed for a detector with 100% fill factor, in which case, the algorithm employs the measured diffraction pattern to constrain the power within a block of pixels, in the upsampled calculations. Since most experimental detectors have fill factors between these extremes, further work needs to be done to deal with the case of a detector with an arbitrary fill factor. It is expected that for a small fill factor, the floating scheme will provide a better approximation of the ptychographic calculations, whilst for the case of high detector fill factor, the power constraint strategy will provide a better approximation. Hence it is likely that the general case may comprise an amalgamation of both of these techniques.

Furthermore, research into the connection between binned coherent diffraction patterns and partially coherent diffraction patterns, in different experimental setup, may provide the means for high fidelity specimen reconstruction from a set of ptychographic partial coherent diffraction patterns.

8 Bibliography

Abbe, E. (1873) 'Contributions to the theory of the microscope and microscopic perception', *Arch. Mikrosk. Anat.*, December, pp. 413-468.

Abbey, B., Nugent, K.A., Williams, G.J., Clark, J.N., Peele, A.G., Pfeifer, M.A., Jonge, M.D. and McNulty, I. (2008) 'Keyhole coherent diffractive imaging', *Nature Physics*, vol. 4, May, pp. 394-398.

Ayers, G.R. and Dainty, J.C. (1988) 'Iterative blind deconvolution and its application', *Optics Letters*, vol. 13, no. 7, July, pp. 547-549.

Batey, D., Edo, T.B. and Rodenburg, J.M. (2011) 'Private communication - results pending publication'.

Bauschke, H.H., Combettes, P.L. and Luke, D.R. (2002) 'Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization', *Journal of the Optical Society of America A*, vol. 19, no. 7, pp. 1334-1345.

Bauschke, H.H., Combettes, P.L. and Luke, D.R. (2003) 'Hybrid projection–reflection method for phase retrieval', *Journal of the Optical Society of America A*, vol. 20, no. 6, pp. 1025-1034.

Boa, P., Zhang, F., G, P. and Osten, W. (2008) 'Phase retrieval using multiple illumination wavelengths', *Optics Letters*, vol. 33, no. 4, pp. 309-311.

Born, M. and Wolf, E. (1999) *Principles of Optics*, 7th edition, Cambridge: Cambridge University Press, pp. 514, 471.

Bouwkamp, C.J. (1954) 'Diffraction Theory', *Reports on Progress in Physics*, vol. 17, no. 1, pp. 35-100.

Bunk, O., Dierolf, M., Kynde, S., Johnson, I. and Marti, O. (2008) 'Influence of the overlap parameter on the convergence of the ptychographical iterative engine', *Ultramicroscopy*, vol. 108, no. 5, pp. 481-487.

Chapman, J.N. (1975) 'The application of iterative techniques to the investigation of strong phase objects in the electron microscope', *Philosophical Magazine*, vol. 32, no. 3, pp. 541-552.

Chapman, H. (1996) 'Phase-retrieval X-ray microscopy by Wigner-distribution deconvolution', *Ultramicroscopy*, vol. 66, no. 3-4, pp. 153-172.

Clark, J.N. (2011) 'Advances in Fresnel Coherent Diffractive Imaging', Thesis, (PhD), La Trobe University, Victoria, Australia. pp. 135-137

Clark, J.N., T, P.C., Pfeifer, M.A., Peele, A.G., Williams, G.J., Chen, B., Nugent, K., Hall, C., Fullagar, W., Kim, S. and McNulty, I. (2010) 'Use of a complex constraint in coherent in coherent diffractive imaging', vol. 18, no. 3, pp. 1981-1993.

Cooley, J.W. and Tukey, J.W. (1965) 'An Algorithm for the Machine Computation of the Complex Fourier Series', *Mathematics of Computation*, vol. 19, pp. 297-301.

DeViels, R. and Thompson, P. (1998) *The new physical optics notebook: Tutorial in Fourier optics*, 1st edition, SPIE, pp. 41.

Dierolf, M., Menzel, A., Thibault, P., Schneider, P., Kewish, C.M., Wepf, R., Bunk, O. and Pfeiffer, F. (2010) 'Ptychographic X-ray computed tomography at the nanoscale', *NATURE*, vol. 467, Septembe, pp. 436-440.

Edo, T.B., Sweeney, F., Lui, C. and Rodenburg, J.M. (2010) 'Noise limit on practical electron ptychography', *J. Phys.: Conf. Ser*. **241** 012065

Edo, T.B., Zhang, F. and Rodenburg, J.M. (2010) 'Resolution improvement in coherent diffractive imaging (ptychography)', Proc. SPIE 7729, 77291H; doi:10.1117/12.853346

Elser, V. (2003) 'Phase retrieval by iterated projections', *Journal of the Optical Society of America A*, vol. 20, no. 1, January, pp. 40–55.

Faulkner, H.M.L. and Rodenburg, J.M. (2004) 'Movable Aperture Lensless Transmission Microscopy: A Novel Phase Retrieval Algorithm', *Physical Review Letters*, vol. 93, no. 2, pp. 1-4, Available: 023903.

Faulkner, H.M.L. and Rodenburg, J.M. (2005) 'Error tolerance of an iterative phase retrieval algorithm for moveable illumination microscopy', *Ultramicroscopy*, vol. 103, pp. 153-164.

Fienup, J.R. (1978) 'Reconstruction of an object from the modulus of its Fourier transform', *Optics Letters*, vol. 3, no. 1, pp. 27-29.

Fienup, J.R. (1982) 'Phase retrieval algorithms: a comparison', *Applied Optics*, vol. 21, no. 15, August, pp. 2758-2769.

Friedman, S.L. and Rodenburg, J.M. (1992) 'Optical Demonstration of a New Principle of Far-Field Microscopy ', *Journal of Physics D: Applied Physics*, vol. 25, pp. 147-154.

Fujiwara, K. (1961) 'Relativistic Dynamical Theory of Electron Diffraction', *J. Phys. Soc. Jpn.*, vol. 16, pp. 2226-2238.

Gabor, D. (1948) 'A new microscopic principle', *Nature*, May, pp. 777-778.

Gabor, D. (1949) 'Microscopy by Reconstructed Wave-Fronts', *Proc. R. Soc. Lond. A*, vol. 197, pp. 454-487.

Gerchberg, R.W. (1974) 'Super-resolution through error energy reduction', *Optica Acta*, vol. 21, pp. 709-720.

Gerchberg, R.W. and Saxton, W.O. (1972) 'A practical algorithm for the determination of phase from image and diffraction plane pictures', *Optik*, vol. 35, no. 2, pp. 237-246.

Goodman, J.W. (2005) *Introduction to Fourier Optics*, 3rd edition, Roberts & Company Publishers, pp. 38, 100, 104.

Guizar-Sicairos, M. and Fienup, J.R. (2008) 'Phase retrieval with transverse translation diversity: a nonlinear optimization approach', *Optics Express*, vol. 16, no. 10, pp. 7264-7278.

Haider, M., Uhlemann, S., Schwan, E., Rose, H., Kabius, B. and Urban, K. (1998) 'Electron microscopy image enhanced', *Nature*, vol. 392, no. 6678, pp. 768-769.

Hawkes, P.W. (2001) 'The long road to spherical aberration correction', *Biology of the Cell*, vol. 93, no. 6, pp. 432-433.

Hegerl, R. and Hoppe, W. (1970) 'Dynamical theory of crytal structure analysis by electron in inhomogeneous primary beam wavefield', *Ber. Bunsenges. Physik. Chemie*, vol. 74, pp. 1148-1154.

Hegerl, R. and Hoppe, W. (1972) 'Phase evaluation in generalised diffraction (Ptychography)', Proc. fifth European Congress on Electron MIcroscopy, 628-629.

Hegerl, R. and Hoppe, W. (1976) 'Influence of electronic noise on three-dimensional image reconstruction', *Z. Naturforsch*, vol. 31a, pp. 1717-1721.

Honig, S., Hoppe, R., Patommel, J., Schropp, A., Stephan, S., Schoder, S., Burghammer, M. and Schroer, C.G. (2011) 'Full optical characterization of coherent x-ray nanobeams by ptychographic imaging', *Optics Express*, vol. 19, no. 17, August, pp. 16324-16329.

Hoppe, W. (1969a) 'Diffraction in inhomogeneous primary wave fields.1. Principle of phase determination from electron diffraction interference', *Acta Crystallographica*, vol. A 25, July, pp. 495-501.

Hoppe, W. (1969b) 'Diffraction in inhomogeneous primary wave fields.III. Amplitude and phase determination for nonperiodic objects', *Acta Crystallographica*, vol. A 25, July, pp. 508-515.

Hoppe, W. (1982) 'Trace structure analysis, ptychography, phase tomography', *Ultramicroscopy*, vol. 10, pp. 187-198.

Hoppe, W. and Strube, G. (1969) 'Diffraction in inhomogeneous primary wave fields.II. Optical experiments for phase determination of lattice interferences', *Acta Crystallographica*, vol. A 25, July, pp. 502-507.

Hue, F., Rodenburg, J.M., Maiden, A.M., Sweeney, F. and Midgley, P.A. (2010) 'Wave-front phase retrieval in transmission electron microscopy via ptychography', *PHYSICAL REVIEW B*, vol. 82, no. 12, September, pp. 1214151-1214154.

Hurst, A.C., Edo, T.B., Walther, T., Sweeney, F. and Rodenburg, J.M. (2010) 'Probe position recovery for ptychographical imaging', *Journal of Physics: Conference Series*, vol. 241, no. 1, p. 012004.

Kirkland, E.J. (2010) *Advanced Computing in Electron Microscopy*, 2nd edition, Springer Science+Business Media, pp. 18, 63.

Kirkland, A.I., Saxton, W.O., Chau, K.L. and Tsuno, K.K.M. (1995) 'Super-resolution by aperture synthesis: tilt series reconstruction in CTEM', *Ultramicroscopy*, vol. 57, no. 4, March, pp. 355-374.

Krivanek, O.L., Dellby, N., Spence, A.J., Camps, R.A. and Brown, L.M. (1997) 'Aberration correction in the STEM', Institute of Physics Conference Series (Proceedings 1997 EMAG Meeting), ed. Rodenburg J M, 35-39.

Landauer, M.N., McCallum, B.C. and Rodenburg, J.M. (1995) 'Double resolution imaging of weak phase specimens with quadrant detectors in the STEM', *Optik*, vol. 100, no. 1, pp. 37-46.

Leith, E.N. and Upatnieks, J. (1962) 'Reconstructed Wavefronts and Communication Theory', vol. 52, no. 10, pp. 1123-1130.

Luke, D.R. (2005) 'Relaxed Averaged Alternating Reflections for Diffraction Imaging', *Inverse Problem*, vol. 21, no. 1, February, pp. 37-50.

Maiden, A.M., Humphry, M.J. and Rodenburg, J.M. (2011) 'Private communication - results pending publication'.

Maiden, A.M., Humphry, M.J., Zhang, F. and Rodenburg, J.M. (2011) 'Super-resolution imaging via ptychography', *Journal of the Optical Society of America A*, vol. 28, no. 4, pp. 604-612.

Maiden, A.M. and Rodenburg, J.M. (2009) 'An improved ptychographical phase retrieval algorithm for diffractive imaging', *Ultramicroscopy*, vol. 109, no. 10, pp. 1256-1262.

Maiden, A.M. and Rodenburg, J.M. (2011) 'Private communication - results pending publication'.

Maiden, A.M., Rodenburg, J.M. and Humphry, M.J. (2010a) 'A new method of high resolution, quantitative phase scanning microscopy', Proc. SPIE, 7729, 77291I; doi:10.1117/12.853339

Maiden, A.M., Rodenburg, J.M. and Humphry, M.J. (2010b) 'Optical ptychography: a practical implementation with useful resolution', *Optic Letters*, vol. 35, no. 15, August, pp. 2585-2587.

Marchesini (2007) 'A unified evaluation of iterative projection algorithms for phase retrieval', *Review of Scientific Instruments*, vol. 78, no. 1, Available: 011301.

McCallum, B.C., Landauer, M.N. and Rodenburg, J.M. (1995) 'Complex image reconstruction of weak specimens from a three-sector detector in the STEM', *Optik*, vol. 101, no. 2, pp. 53-62.

McCallum, B.C. and Rodenburg, J.M. (1992) '2-Dimensional Optical Demonstration of Wigner phase-retrieval microscopy in the STEM configuration', *Ultramicroscopy*, vol. 45, pp. 371-380.

McCallum, B.C. and Rodenburg, J.M. (1993) 'Simultaneous Reconstruction of Object and Aperture Functions from Multiple Far field Intensity Measurements', *Journal of the Optical Society of America A*, vol. 10, no. 2, August, pp. 231-239.

McCallum, B.C. and Rodenburg, J.M. (1993b) 'An error analysis of crystalline ptychography in the STEM mode', *Ultramicroscopy*, vol. 52, no. 1, September, pp. 85-99.

McEwen, B.F., Downing, K.H. and Glaeser, R.M. (1995) 'The relevance of dose-fractionation in tomography of radiation-sensitive specimens', *Ultramicroscopy*, vol. 60, pp. 357-373.

Meyer, R.R., Kirkland, A.I. and Saxton, W.O. (2004) 'A new method for the determination of the wave aberration function for high-resolution TEM.: 2. Measurement of the antisymmetric aberrations', *Ultramicroscopy*, vol. 99, no. 2-3, pp. 115-123.

Miao, J., Charalambous, P., Kirz, J. and Sayre, D. (1999) 'Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens', *Nature*, vol. 400, pp. 342-344.

Mountjoy, G. (1999) 'Order in two-dimensional projections of thin amorphous three-dimensional structures ', *J. Phys.: Condens. Matter*, vol. 11, pp. 2319-2336.

Mountjoy, G. (2001) 'Atomic structure of amorphous solids from high resolution electron microscopy – a technique for the new millennium?', *Journal of Non-Crystalline Solids*, vol. 293-295, pp. 458-463.

Nellist, P.D., McCallum, B.C. and Rodenburg, J.M. (1994) 'Resolution beyond the 'information limit' in transmission electron microscopy', *Nature*, vol. 374, pp. 630-632.

Nugent, K.A. (2010) 'Coherent methods in the X-ray sciences', *Advances in Physics*, vol. 59, no. 1, pp. 1-99.

Nugent, K.A., Peele, A.G., Quiney, H.M. and Chapman, H.N. (2005) 'Diffraction with wavefront curvature: a path to unique phase recovery', *Acta Crystallographica*, vol. A, no. 61, April, pp. 373-378.

Nyquist, H. (1928) 'Certain topics in telegraph transmission theory', *Transactions of the A. I. E. E*, pp. 617-644.

Oszlányi, G. and Süto, A. (2004) 'Ab initio structure solution by charge flipping', *Acta Crystallographica Section A*, vol. 60, no. 2, pp. 134-141.

Papoulis, A. (1975) 'A new algorithm in spectral analysis and band-limited extrapolation', *IEEE Transactions on Circuits and Systems* , vol. CAS-22, September, pp. 735-742.

Patterson, A.L. (1934) 'A Fourier Series Method for the Determination of the Components of Interatomic Distances in Crystals', *Phys. Rev.*, vol. 46, no. 5, pp. 372-376.

Robinson, I.K., Vartanyants, I.A., Williams, G.J., Pfeifer, M.A. and Pitney, J.A. (2001) 'Reconstruction of the Shapes of Gold Nanocrystals Using Coherent X-Ray Diffraction', *Physical Review Letters*, vol. 87, no. 19, pp. 195505 : 1-4.

Rodenburg, J.M. (1988) 'Properties of Electron Microdiffraction Patterns from Amorphous Materials', *Ultramicroscopy*, vol. 25, pp. 329-343.

Rodenburg, J.M. (2008) 'Ptychography and related diffractive imaging methods', *Advances in Imaging and Electron Physics*, pp. 87-180.

Rodenburg, J.M. and Faulkner, H.M. (2004) 'A phase retrieval algorithm for shifting illumination', *Applied Physics Letters*, vol. 85, no. 20, pp. 4795-4797.

Rodenburg, J.M. and Faulkner, H.M.L. (2005) , Available: PCT/GB2005/001464.

Rodenburg, J.M., Hurst, A.C. and and Cullis, A.G. (2007) 'Transmission microscopy without lenses for objects of unlimited size', *Ultramicroscopy*, vol. 107, pp. 227-231.

Rodenburg, J.M., Hurst, A.C., Cullis, A.G., Dobson, B.R., Pfeiffer, F., Bunk, O., David, C., Jefimovs, K. and Johnson, I. (2007) 'Hard X-ray lensless imaging of extended objects', *Phys. Rev. Lett.*, vol. 98, no. 3, p. 034801.

Rodenburg, J.M., McCallum, B.C. and Nellist, P.D. (1993) 'Experimental tests on double-resolution coherent imaging via STEM', *Ultramicroscopy*, vol. 48, no. 3, pp. 304-314.

Sayre, D. (1952) 'Some implications of a theorem due to Shannon', *Acta Cryst.*, vol. 5, p. 843.

Scherzer, O. (1936) 'Some defects of electron lenses', *Z. Physik*, vol. 101, pp. 593-603.

Scherzer, O. (1947) 'Spherical and chromatic correction of electron lenses', *Optik*, vol. 2, pp. 114-132.

Scherzer, O. (1949) 'The theoretical resolution limit of the electron microscope', *Journal of Applied Physics*, vol. 20, January, pp. 20-29.

Schroer, C.G. (2006) 'Focusing hard x rays to nanometer dimensions using Fresnel zone plates', *PHYSICAL REVIEW B*, vol. 74, no. 3, p. 033405.

Shannon, C.E. (1949) 'Communication in the Presence of Noise', vol. 37, no. 1, January, pp. 10-21.

Sherman, G.C. (1967) 'Application of the Convolution Theorem to Rayleigh's Integral Formulas', *Journal of the Optical Society of America*, no. 4, 57, pp. 546-547.

Spence, J.C.H. (2003) *High-Resolution Electron Microscopy*, 3rd edition, Oxford: Oxford University Press.

Thibault, P., Dierolf, M., Bunk, O., Menzel, A. and Pfeiffer, F. (2009) 'Probe retrieval in ptychographic coherent diffractive imaging', vol. 109, no. 4, pp. 338-343.

Thibault, P., Dierolf, M., Menzel, A., Bunk, O., David, C. and Pfeiffer, F. (2008) 'High-resolution scanning X-ray diffraction microscopy ', *Science*, vol. 321, p. 379–382.

Zach, J. and Haider, M. (1995) 'Aberration correction in a low voltage SEM by a multipole corrector', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 363, no. 1-2, pp. 316-325.

Zhang, F., Pedrini, G. and Osten, W. (2007) 'Phase retrieval of arbitrary complex-valued fields through aperture-plane modulation ', *Physical Review A*, vol. 75, p. 043805.