

UNIVERSITY OF ALBERTA

QUANTITATIVE PHASE RETRIEVAL IN  
TRANSMISSION ELECTRON MICROSCOPY

BY

ROBERT ALEXANDER MCLEOD

A thesis submitted to the Faculty of Graduate Studies and Research in partial  
fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Physics

©Robert Alexander McLeod

Fall 2013

Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

# ABSTRACT

Phase retrieval in the transmission electron microscope offers the unique potential to collect quantitative data regarding the electric and magnetic properties of materials at the nanoscale. Substantial progress in the field of quantitative phase imaging was made by improvements to the technique of off-axis electron holography. In this thesis, several breakthroughs have been achieved that improve the quantitative analysis of phase retrieval. An accurate means of measuring the electron wavefront coherence in two-dimensions was developed and practical applications demonstrated. The detector modulation-transfer function (MTF) was assessed by slanted-edge, noise, and the novel holographic techniques. It was shown the traditional slanted-edge technique underestimates the MTF. In addition, progress was made in dark and gain reference normalization of images, and it was shown that incomplete read-out is a concern for slow-scan CCD detectors. Last, the phase error due to electron shot noise was reduced by the technique of summation of hologram series. The phase error, which limits the finest electric and magnetic phenomena which can be investigated, was reduced by over 900 % with no loss of spatial resolution. Quantitative agreement between the experimental root-mean-square phase error and the analytical prediction of phase error was achieved.

# ACKNOWLEDGEMENTS

I would like to thank my supervisors Marek, Ray, and Mark for their ongoing support and mentorship.

Thanks go to fellow collaborator Michael Bergen made important and useful contributions to the hologram series and Matlab instrument control, as well as being helpful in discussions on mathematical and technical details.

Thanks also go to the technical staff at NINT for maintenance of Hitachi HF-3300 and numerous discussions on microscopy over the years.

The financial support of NINT and NSERC are acknowledged.

I would also like to thank Yoshifumi Taniguchi of Hitachi High Technologies Corporation for frequent advice. The ongoing support of Hitachi High Technologies Canada is gratefully acknowledged.

# TABLE OF CONTENTS

1	Introduction .....	1
1.1	Phase in the TEM .....	2
1.2	Historical Introduction to Electron Holography .....	3
1.3	Overview of Off-axis Electron Holography .....	4
1.4	Quantitative Electron Phase Microscopy.....	10
1.4.1	Optical-Transfer Function Cascade .....	12
1.4.2	Automated TEM Control and Processing.....	13
1.5	References .....	14
2	Determination of Localized Visibility in Off-axis Electron Holography .....	17
2.1	Introduction and Motivation .....	17
2.1.1	Measurement and Physical Origin of the Visibility.....	20
2.2	Statistical Holographic Visibility Metric .....	21
2.2.1	Amplitude in Off-axis Electron Holography.....	23
2.3	An Algorithm for Local Visibility Map Generation.....	24
2.4	Bias in Visibility Metrics .....	29
2.5	Example Experimental Results .....	33
2.6	Conclusions .....	36
2.7	References .....	37
3	Characterization of Detector Modulation-transfer Function with Noise, Edge, and Holographic Methods .....	40
3.1	Introduction.....	40
3.1.1	Optical Transfer Function and Detector Quantum Efficiency .....	41
3.1.2	Analytical Form of the MTF and Normalization.....	42
3.1.3	Impact of the Detector on Microscope Performance .....	43
3.2	Methods for Characterization of Modulation-transfer Function .....	44
3.2.1	The Noise Transfer Function .....	44
3.2.2	The Slanted-Edge Modulation Transfer Function .....	51
3.2.3	The Holographic Modulation Transfer Function.....	52
3.3	Experimental Methods And Results .....	53
3.3.1	Experimental Methods to Measure MTF.....	53
3.3.2	MTF Results And Discussion .....	55

3.3.3	Practical Experimental Technique.....	60
3.4	Conclusions.....	61
3.5	References.....	62
4	Phase Measurement Error in Summation of Electron Holography Series.....	67
4.1	Introduction.....	67
4.2	Optimization of Electro-Optical Configuration.....	68
4.3	Sources of Error in Hologram Series Summation.....	73
4.3.1	Image Drift.....	74
4.3.2	Phase Drift.....	77
4.3.3	Registration Errors.....	80
4.3.4	Combined Phase Error and Spatial Resolution for Hologram Series..	82
4.4	Experimental hologram series.....	84
4.4.1	Experimental Results.....	85
4.4.2	Experimental Phase Error Metric.....	88
4.4.3	Radiation Damage Concerns.....	90
4.5	Conclusions.....	90
4.6	References.....	92
5	General Discussion and Conclusions.....	95
5.1	Future Work.....	96
5.2	References.....	98
A1	Appendix 1 – Supplementary Material for Chapter 3.....	100
A1.1	High Fidelity Dark Reference.....	100
A1.2	Digital Micrograph Dark Reference Sum Script.....	103
A1.3	Voight Profile Best-Fit Parameters.....	104
A2	Appendix 2 – Supplementary Material for Chapter 4.....	106
A2.1	Description of Hologram Series Summation Algorithm.....	106
A2.2	Appendix Videos.....	108

# LIST OF FIGURES

Figure 1.1 Schematic representation of off-axis electron holography .....	5
Figure 1.2 Example hologram of Ir nanoparticles.....	6
Figure 1.3 Fourier transform of Ir nanoparticles hologram.....	8
Figure 1.4 Holographic reconstruction of Ir nanoparticles .....	9
Figure 1.5 An example OTF hierarchy for the Hitachi HF-3300 TEM .....	13
Figure 2.1 Schematic representation of off-axis electron holography .....	19
Figure 2.2 Sideband and centerband amplitude for Ni/NiO particles .....	24
Figure 2.3 Example of Fourier-ratio visibility metric for a reference hologram .....	29
Figure 2.4 Comparison of three visibility metrics .....	32
Figure 2.5 Empirical normalization of Fourier-ratio visibility metric .....	32
Figure 2.6 Example of visibility image processing on Ni/NiO core-shell particles .....	35
Figure 3.1 Flow-chart depicting scattering and amplification of CCD detector .....	42
Figure 3.2 1-D Noise-transfer function .....	46
Figure 3.3 Evaluation of incomplete read-out and effect on gain references .....	50
Figure 3.4 MTF for all three methods compared by accelerating voltage .....	57
Figure 3.5 Comparison of MTF methods for each accelerating voltage .....	59
Figure 3.6 Comparison of holographic and edge-based DQE .....	60
Figure 4.1: Visibility as a function of biprism bias .....	72
Figure 4.2 Phase error as a function of condenser excitation .....	73
Figure 4.3 Comparison of numerical Monte Carlo random walk to analytic formula.....	76
Figure 4.4 Example random-walk MTFs for selected exposure times .....	76
Figure 4.5 Histogram of extended image drift series .....	77
Figure 4.6 Histogram of extended phase wavefront drift series.....	79
Figure 4.7 Visibility loss as a function of exposure time .....	80
Figure 4.8 Image registration error for simulated holograms .....	81
Figure 4.9 Phase registration error for simulated holograms .....	82
Figure 4.10 Optimization of summed phase error .....	83
Figure 4.11 Experimental hologram of WS <sub>2</sub> and registered, summed centerband.....	86
Figure 4.12 Registered and summed visibility map .....	87
Figure 4.13 Comparison of unregistered and registered phase summation.....	87
Figure 4.14 Estimates of image and phase wavefront drift.....	88
Figure 4.15 2-D map of even-odd RMS phase difference.....	89
Figure 4.16 Convergence of even-odd phase error.....	90
Figure 5.1 Projected phase shift of a dipole.....	97
Figure A1.1 Dark current mean and standard deviation for Gatan US-1000.....	101
Figure A1.2 1,000-frame dark reference.....	101
Figure A1.3 120,000-frame dark reference.....	102
Figure A1.4 Difference between two 17,000-frame dark references recoded a year apart.....	102

# LIST OF TABLES

Table 2.1: Results for mean phase shifts within contours of Fig. 2.6(c) .....	36
Table A1.1: Least-squares best-fit parameters for MTF and NTF results.....	104

# LIST OF VIDEO STILLS

Video A2.1: Verification of image drift registration.....	108
Video A2.2: Cross-section of registered cumulative phase compared to unregistered cumulative phase.....	109

# List of Abbreviations

$BP_1$  – Upper biprism (at image plane)

$BP_2$  – Lower biprism (at reciprocal plane)

CB – Centerband (of the Fourier transform of a hologram)

CCD – Charged-coupled device

DQE – Detector quantum efficiency

EELS – Electron energy-loss spectroscopy

EH – Electron holography

FoV – Field-of-view

HRTEM – High-resolution transmission electron microscopy

MTF – Modulation-transfer function

NPS – Noise-power spectrum (square of NTF)

NTF – Noise-transfer function

OTF – Optical-transfer function

PSF – Point-spread function

PSD – Power spectral density (distribution of image frequencies)

SB – Sideband (of the Fourier transform of a hologram)

SNR – Signal-to-noise ratio

TEM – Transmission electron microscope

2-D – Two-dimensional

# List of Symbols

**Note on subscripts:** The object and reference portions of the electron wavefront are indicated by the respective subscripts  $1$  and  $2$ . When separate object and reference holograms are acquired, they are indicated by the respective subscripts  $o$  and  $r$ .

$a$  – Apodization window radius [ $\text{nm}^{-1}$ ]

$A$  – Electron wavefront amplitude

$\alpha_o$  – Convergence half-angle of illumination [mrad]

$\delta_i$  – Registration error, image [pix]

$\delta_\phi$  – Registration error, wavefront phase [rad]

$D_i$  – Image diffusivity coefficient [ $\text{pm}^2 \text{s}^{-1}$ ]

$D_\phi$  – Wavefront phase diffusivity coefficient [ $\text{pm}^2 \text{s}^{-1}$ ]

$\gamma$  – Zernike's complex degree of coherence (OTF of the electron source)

$I_e$  – Electron current density (at sample plane) [ $\text{e}^- \text{nm}^{-2} \text{s}^{-1}$ ]

$\lambda$  – Electron relativistic wavelength [pm]

$m$  – Number of frames (for hologram series)

$\mu$  – Mean electron dose [ $\text{e}^- \text{nm}^{-2}$ ]

$\phi$  – Electron wavefront phase shift [rad]

$\psi$  – Electron wavefunction

$q$  – Spatial frequency [ $\text{nm}^{-1}$ ]

$q_c$  – Carrier frequency [ $\text{nm}^{-1}$ ]

$\sigma_\phi$  – Standard deviation of electron wavefront phase [rad]

$\sigma_{\Sigma-\phi}$  – Standard deviation of hologram *series* phase [rad]

$\sigma_V$  – Standard deviation of electron wavefront visibility

$t_x$  – Exposure time [s]

$v_i$  – Image constant-velocity drift velocity [ $\text{pm s}^{-1}$ ]

$v_\phi$  – Wavefront phase constant-velocity drift velocity [ $\text{pm s}^{-1}$ ]

$U$  – Electrostatic potential

$V$  – Holographic visibility (or fringe contrast)

$V_n$  – Normalized holographic visibility (by detector MTF)

$V^*$  – Reduced holographic visibility (lowered by wavefront drift)

# 1 INTRODUCTION

This thesis concerns progress towards quantitative phase retrieval of specimens in the *transmission electron microscope* (TEM). To provide broad context for the thesis, the introduction will first define the phase as it relates to an electron wavefront in section 1.1. In section 1.2 a historical overview of the field of electron holography is provided. In section 1.3, an overview of the phase retrieval method of choice, off-axis electron holography, is provided. In section 1.4, requirements for electron microscopy in general and electron holography in particular to be considered quantitative is discussed.

In Chapter 2, a third component signal in electron holography, the visibility or contrast of the interference fringes is examined. The visibility is related to the coherence of the electron wavefront. Historically, coherence has been evaluated by measurement based on the minimum and maximum intensity values. A new method is presented, based on statistical moments that allow the visibility to be measured in a deterministic and reproducible fashion suitable for quantitative analysis. An algorithm that allows the visibility to be resolved in two-dimensions is provided, which is termed the *local visibility*. Thus the local visibility may be used to evaluate the loss of coherence due to electron scattering within a specimen, or as an aid in image analysis and processing. All visibility metrics are biased at low-dose conditions by the presence of shot-noise, which can be empirically normalized to achieve linear response.

In Chapter 3, a new method for characterization of detector performance in the TEM is developed, based on the measured contrast of holographic fringes. The new method changes spatial frequency of the measured holographic fringes, generated by an electrostatic biprism and Schottky or cold field-emission gun, to sample the *modulation-transfer function* (MTF) of the detector. The MTF of a Gatan Ultrascan™ 1000 *charged-coupled detector* (CCD) is evaluated using the new method and the results are compared to the established noise and slanted-edge method results. Requirements for accuracy of the edge and noise MTF methods are discussed. We consider issues surrounding incomplete read-out and how it affects the gain reference normalization of the detector. We evaluate how the MTF affects optimization of experimental parameters in the TEM.

In Chapter 4, the concept of improving the phase resolution by summation of hologram series is introduced. Optimization of the double biprism holography configuration is provided. An analytical model of image and phase drift, composed of a combination of linear drift and Brownian random-walk, is derived and experimentally verified. The accuracy of image registration via cross-correlation and phase registration is characterized by simulated hologram series. The model of series summation errors allows the optimization of phase error as a function of

exposure time and fringe carrier frequency for a target spatial resolution. An experimental example of hologram series summation is provided on WS<sub>2</sub> fullerenes. A metric is provided to characterize the object phase error from experimental results. The ultimate object root-mean-square phase error is 0.0085 rad (2π/740) at a spatial resolution less than 0.615 nm and a total exposure time of 900 s. The ultimate vacuum phase error is 0.0052 rad (2π/1200). The analytical prediction of phase error agrees with the experimental metric to +6.2 % inside the object and -2.1 % in the vacuum.

In Chapter 5, how the three topics in Chapters 2-4 interrelate are discussed. The contributions towards the topic of quantitative phase retrieval are outlined and conclusions and directions for future work are offered.

## 1.1 PHASE IN THE TEM

Typically in the TEM one uses illumination that may be approximated as a plane wave. For the sake of discussion the electron wavefunction at the exiting surface of the specimen  $\psi(\mathbf{r})$ , known as the *exit wavefunction*, will also be assumed to be a plane-wave,

$$\psi(\mathbf{r}) = A(\mathbf{r})e^{i\phi(\mathbf{r})} \quad (1.1)$$

where  $A(\mathbf{r})$  is the amplitude modulation and  $\phi(\mathbf{r})$  the phase modulation of the electron wavefront as a function of position  $\mathbf{r}$ . The phase represents the sum contributions from a number of physical phenomena [1].

$$\phi(\mathbf{r}) = \phi_E(\mathbf{r}) + \phi_M(\mathbf{r}) \quad (1.2)$$

where E denotes the phase shift arising from electric fields, and M the phase shift arising from magnetic fields. The electrostatic phase shift is,

$$\phi_E(\mathbf{r}) = 2\pi \frac{e}{h\nu} \int U(\mathbf{r})dz \quad (1.3)$$

where  $\nu$  is the electron velocity,  $e$  is the electron charge,  $h$  is Planck's constant, and  $U$  is the electrostatic potential integrated along the electron path  $z$ . The magnetic-vector potential  $\phi_M$  is,

$$\phi_M(\mathbf{r}) = -\frac{e}{\hbar} \int \mathbf{A}(\mathbf{r})dz \quad (1.4)$$

where  $\mathbf{A}$  is the magnetic potential in the plane defined by the position vector  $\mathbf{r}$ .

Electron holography has been applied to many material science applications [1-4]. A motivation that is germane to Alberta, Canada is the study of charge-transfer from a substrate to a catalyst nanoparticle [5]. In such a case, the goal is to assess the total charge on the nanoparticle from the strength of the fringe field in vacuum.

There are numerous issues that this simple electromagnetic model for phase shift, as outlined by Eqn. 1.3 and 1.4, does not address. The model for the phase does not address cross-talk between the amplitude and phase caused by the contrast-transfer

function of the objective lens [6]. Nor does it incorporate dynamic diffraction [7]. When Bragg diffraction is included in the analysis, the electric phase shift  $\phi_E$  is further sub-divided to include crystal and geometric phase shift [8].

## 1.2 HISTORICAL INTRODUCTION TO ELECTRON HOLOGRAPHY

The technique of electron holography was originally envisioned by Denis Gabor in 1947 as a means to correct the wavefront aberration of the electron microscope [8,9]. The magnetic lenses typically used in electron microscopy are all optically convex. Therefore the means of correcting wavefront aberration used in photon optics, by mixing convex and concave optical elements, is not available in electron microscopy. Gabor's scheme involved recording an electron interference pattern in a film emulsion. The hologram could then be reconstructed optically, by illuminating the emulsion and using optical components to correct the wavefront aberration. Gabor was awarded the Nobel Prize in Physics 1971 for his work in holography [10].

Electron holography did not enjoy initial success due to the limited coherence of Tungsten hairpin electron sources available at the time. Holography was first demonstrated in 1962 in photon optics, with the development of the laser [11]. Progress towards electron holography was made throughout the 1960s and 1970s. The electron biprism was invented by Mollenstedt in 1956 [12]. This enabled electron interference to be demonstrated for the first time. Cylindrical lenses were used to create highly elliptical illumination, which in turn extended the source transverse to the hologram enough to detect faint interference fringes.

The preferred method of holography in transmission electron microscopy is the off-axis method. In the off-axis method, the Mollenstedt biprism deflects two halves of the electron wavefront together. The degree and angle of overlap is determined by the electrical bias on the biprism. For rotationally symmetrical illumination, the shadow of the limiting aperture is bisected into two 'D'-shaped illumination patterns which slightly overlap. The specimen (or object) is placed within one of the 'D's, which forms the object wave, while the other wave passes through the adjacent vacuum. Within the overlap region of the two 'D's the two waves form a sinusoidal interferogram on the detector. The phase difference between the object and reference waves is encoded in the periodicity (or frequency) of the fringe pattern. Thus off-axis holography is essentially an interferometry method that is called 'holography' for historical reasons.

The practical employment of off-axis electron holography had to await the introduction of the Schottky and field-effect source [13,14]. This early work was reconstructed optically with a Mach-Zender interferometer, as originally envisioned by Gabor. Tonomura et al. focused on the fact that electron's are phase shifted by passing through electrostatic and magnetic fields. An example of early work in

electron holography was reconstructing the magnetic field in triangular Cobalt particles [15].

A particularly famous experiment conducted by the Tonomura wavefront project was the direct observation of the Aharonov-Bohm effect [16-18]. Aharonov and Bohm claimed that as a consequence of gauge-theory, with relevance to the standard quantum physics model, that electron phase should be retarded in the presence of electrostatic and magnetic vector potentials. A microscopic superconducting toroid was fabricated to create a pure magnetic potential in the bore, with the magnetic field completely expelled by the superconducting coating. Off-axis electron holography was used to show both that there was little leakage of magnetic field outside of the superconducting coating. The result showed a clear phase shift of the holographic fringes inside the bore relative to outside the bore.

The other primary line of research, first employed by Lichte in 1985, was the use of sub-Angstrom fringes for atomic-scale electron holography [19]. In this case, Lichte's goal was the original proposal of Gabor to correct the lens aberrations post-collection of the hologram [20], which continues to be an ongoing challenge.

Lichte performed the first numerical reconstruction of an electron hologram (i.e. with Fourier transforms in a computer). The photographic emulsion was scanned into a computer at a resolution of 2048 x 2048 pixels. With the advent of greater computer memory, the large area of the film plates allowed 4096 x 4096 pixel holograms to be reconstructed. The later introduction of the CCD detector to electron microscopy allowed holograms to be recorded electronically without the need for scanning. Lichte has also made many important contributions by defining the theoretical limits of off-axis electron holography, for example the phase error equation [21]. Another example is optimization of the TEM for holography such as optimum focus to minimize cross-talk between amplitude at phase [22].

Off-axis electron holography has found numerous applications in the characterization of nanoscale electric and magnetic behaviour of materials. For a complete introduction, the reader is referred to the reviews by Lichte et al. [1][2]. Examples of electron holography electric applications include: dopant profiling in semiconductors, measurement of electric dipoles and polarization in ferroelectric materials, measurement of strain in crystalline materials, and measurement of field around field-effect tips. Magnetic applications include domain mapping, magnetization of thin-film, and measurement of magnetic moments in ferromagnetic nanoparticles. Electron holography has also been used to measure the mean inner potential of materials, for composition mapping, and for thickness mapping.

### 1.3 OVERVIEW OF OFF-AXIS ELECTRON HOLOGRAPHY

Off-axis electron holography is an interferometric method known as 'wavefront-splitting' in the literature, whereby a Fresnel biprism splits the source into two

virtual sources. The electro-optical configuration uses an electrostatic biprism to separate the electron wavefront into an object and reference half. The object wave passes through a specimen, while the reference passes through vacuum adjacent to the specimen, as shown in Figure 1.1.

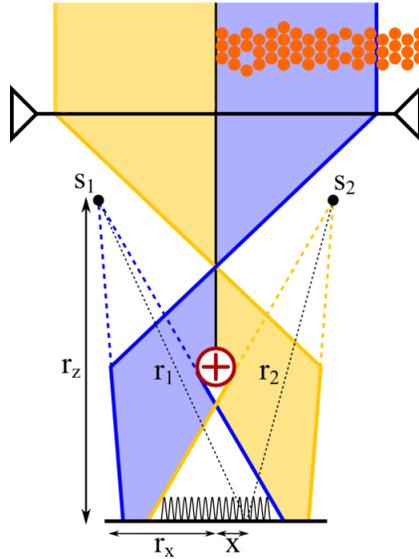


Figure 1.1 (reprinted from Chapter 2): Schematic representation of off-axis electron holography. The electron wavefront is divided into two halves: the object wave (blue) passes through the specimen (orange spheres) while the reference wave (yellow) passes through adjacent vacuum. The two plane waves converge on the detector, producing a parallel interference pattern known as an off-axis hologram.

The biprism is typically a fine silica filament coated with a layer gold to make it electrically conductive. When biased, it creates an electrostatic field between it and grounding plates (not shown) which deflects the two halves of the electron wavefront. The result is periodic plane-wave interference pattern known as a hologram which is observed on a suitable detector. An example hologram of Iridium nanoparticles on a carbon substrate is shown in Figure 1.2.

The wavefront deflection induced by the biprism creates two virtual sources, denoted  $S_1$  and  $S_2$  in Figure 1.1. As biprism bias increases, the separation between the virtual sources increases. For the single biprism configuration shown in Fig. 1.1, both the hologram width and carrier frequency increase with increasing biprism bias. However, the separation of virtual sources puts strong requirements on the coherence of the source. The complex degree of coherence of the source,  $\gamma(q)$ , describes the maximum fringe contrast that can be observed at the carrier frequency  $q_c$  [23]. In electron microscopy, it is often more convenient to describe functions in terms of the spatial frequency  $q$  instead of position  $r$ , to be discussed further in section 1.3.1. The function  $\gamma(q)$  can be conceptually approximated as 'Gaussian-like' [24]. For low carrier frequency there is little loss of coherence but then a rapid fall-off with higher carrier frequencies. As a result, the contrast of the

holographic fringes, which is known as the holographic visibility  $V$ , decreases with increasing carrier frequency.

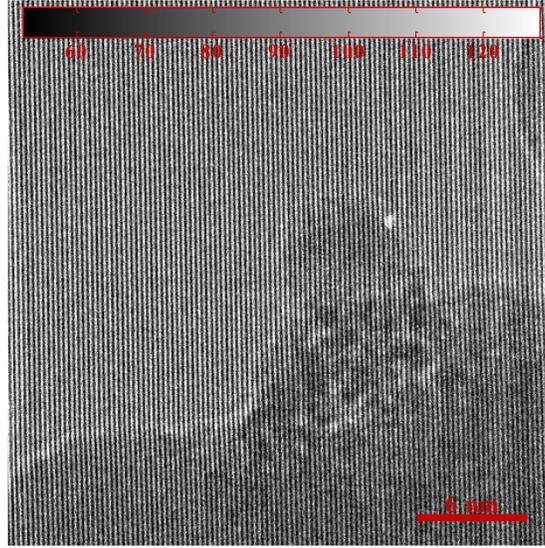


Figure 1.2: An example hologram of Ir nanoparticles (center) on an amorphous carbon substrate (bottom). The interference fringes can be observed as the fine vertical lines, with the image of the specimen superposed. The upper half of the image is vacuum.

Electron holography is well-described by two-beam interference. The hologram is modelled as two plane waves, the object wave  $\psi_1$  and reference wave  $\psi_2$ , converging at a semi-angle  $\theta$ ,

$$\psi_1(\mathbf{r}) = A_1(\mathbf{r})\exp\left(i(2\pi\mathbf{q}_c \cdot \mathbf{r} + \phi_1(\mathbf{r}))\right) \quad (1.5.a)$$

$$\psi_2(\mathbf{r}) = A_2(\mathbf{r})\exp\left(i(2\pi\mathbf{q}_c \cdot \mathbf{r} + \phi_2(\mathbf{r}))\right) \quad (1.5.b)$$

Here  $A_1$  is the object wave amplitude,  $A_2$  is the reference wave amplitude,  $\phi_1$  and  $\phi_2$  are the object and reference wave phase shift respectively. The carrier frequency is  $\mathbf{q}_c = 2\sin(\theta)/\lambda_e$  where  $\lambda_e$  is the de Broglie wavelength of the primary electron. The interference of the object and reference wave produces the intensity pattern

$$\psi_{12}^2(\mathbf{r}) = A_1^2(\mathbf{r}) + A_2^2(\mathbf{r}) + 2V(\mathbf{r}, \mathbf{q}_c, \alpha_o)A_1(\mathbf{r})A_2(\mathbf{r})\cos(2\pi\mathbf{q}_c \cdot \mathbf{r} + \phi_1(\mathbf{r}) - \phi_2(\mathbf{r})) \quad (1.6)$$

where  $V$  is the holographic visibility and reflects the partial coherence of the electron wavefront. The visibility varies across the field of view, due to source properties, optical aberrations of the TEM lenses, and incoherent scattering within the specimen. The visibility is discussed further in section 1.3.1 and in depth in chapter 2.

The primary advantage of off-axis holography, as compared to in-line holography, is that interpretation is straightforward. In off-axis holography, the phase shift of the electron wavefront is encoded at the carrier fringe frequency  $\mathbf{q}_c$ . In addition, a

complex conjugate is found at the negative pole,  $-q_c$ . In-line holography has no such separation of the image from its complex conjugate which compromises reconstruction of the complex wave.

The complex wave amplitude and phase is retrieved from the hologram intensity pattern by means of a digital Fourier reconstruction algorithm [25,26]. After post-processing to remove detector artifacts, a fast-Fourier transform is applied to the hologram. In Fourier-space, it is straight-forward to separate the centerband  $CB$  (about  $q = 0$ ) from the sideband  $SB$  (about  $q = q_c$ ) by an apodization function. The Fourier transform of Figure 1.2 is shown in Figure 1.3 with the centerband and sideband highlighted. An apodization function acts as a digital window (or low-pass spatial filter), that allows the sideband (or centerband) signal to pass through. Generally the maximum observed spatial frequency in the reconstructed is half the carrier frequency. Apodization functions are typically based on cosine or Gaussian functions to minimize Fourier ringing artifacts. Spatial frequencies outside the apodization are forced to zero. After apodization, the sideband is digitally shifted from being centered at  $q = q_c$  to being centered at  $q = 0$ . The sideband and centerband are then inverse Fourier transformed to give the complex sideband amplitude and phase and the centerband amplitude in real-space. After inverse Fourier transformation, the centerband and sideband have the analytic form,

$$CB = A_1^2 + A_2^2 \quad (1.7.a)$$

and

$$SB = 2A_1A_2V e^{i(\phi_1 - \phi_2)} \quad (1.7.b)$$

From Eqn.s (1.7.a) and (1.7.b) we can solve for the object amplitude  $A_1$ , the visibility  $V$ , and the phase difference between the object and reference wave,  $\phi_1 - \phi_2$ . As such, the phase shift is always a relative measure. The reconstruction of the object wave amplitude and phase shift is shown in Figure 1.4

All experiments in this thesis were performed on the Hitachi HF-3300 TEM located at the National Institute for Nanotechnology (NINT). The Hitachi TEM is equipped with a double biprism configuration [27]. In the double biprism configuration, the upper biprism  $BP_1$  is placed at an image plane (i.e. the selected area aperture plane) between the objective lens and first intermediate lens. As  $BP_1$  is at an image plane, the electric field creates a pure wavefront tilt with no shift. The second biprism  $BP_2$  is placed near the cross-over between the first and second intermediate lenses. In practice, the lower biprism is not directly at a cross-over as the focused beam could damage it so it contributes both wavefront tilt and shift. As the upper biprism shadows the lower one, the characteristic Fresnel fringes that superpose on the interference fringes are almost entirely absent in double biprism electron holography. In Figure 1.2, the upper biprism bias is low, so the lower biprism is not completely occluded. As such, faint Fresnel fringes are barely visible.

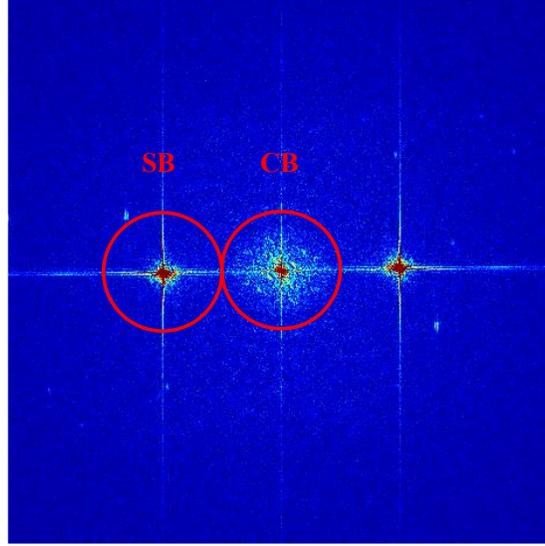


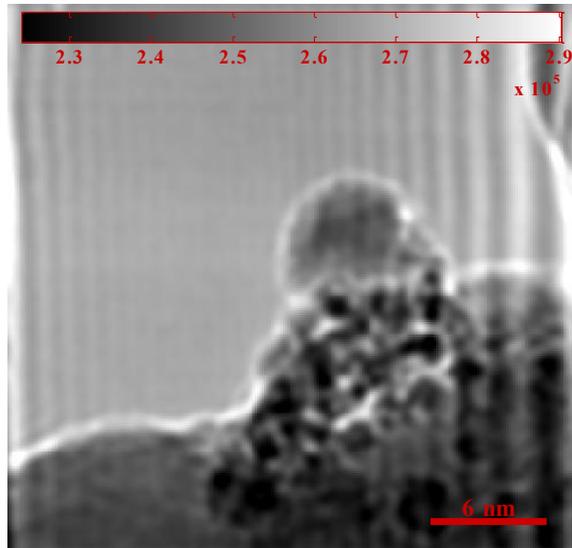
Figure 1.3: Fourier transform of the hologram with the sideband SB and centerband CB indicated. Vertical and horizontal streaks are due to the finite size of the fast-Fourier transform. The right-hand peak is the complex conjugate of the sideband.

At the image-plane, the upper biprism bias primarily controls the beam tilt which determines the carrier frequency of the fringes,  $q_c$ , while the lower biprism bias controls the width of the hologram fringe field of view,  $w$ . For the single biprism case, the carrier frequency is optimized at one-third the maximum frequency that observed in the specimen at high spatial resolution [28]. For the double biprism configuration, the decoupling of carrier frequency and hologram width relaxes this relationship as discussed in Chapter 4. Furthermore, single biprism configurations often have difficulty generating very fine fringes ( $\lambda_c < 50 \text{ pm}$ ) as the biprism is not stable at the required bias of  $> 1000 \text{ V}$  [29]. With the double biprism configuration, such fringes can be obtained at  $450 \text{ V}$  bias on  $BP_1$  at  $300 \text{ keV}$  and considerably less at lower primary electron energies.

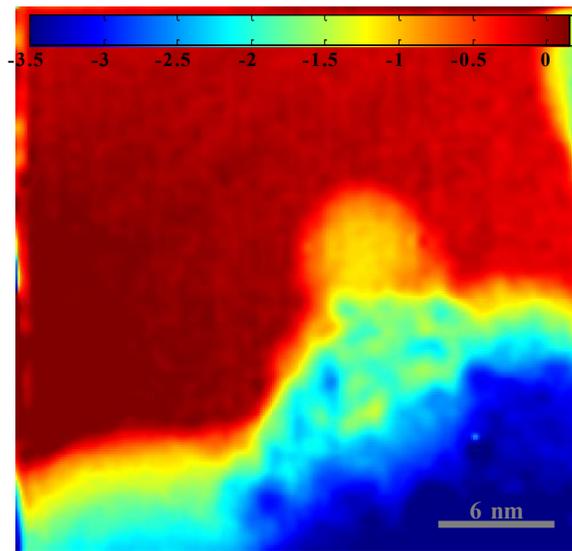
Off-axis electron holography suffers from a number of drawbacks which will be outlined as follows: (1.) The principle figure of merit for electron holograms is the phase error, sometimes referred to as the phase resolution. The phase error was derived by Lenz using a method analogous to the formalism of complex circular random variables [21,30,31]. The standard deviation of the phase has the form,

$$\sigma_\phi = \sqrt{\frac{2}{\mu V^2}} \quad (1.8)$$

where  $\mu$  is the mean dose per reconstructed pixel and  $V$  is the holographic visibility.



(a)



(b)

Figure 1.4: Holographic reconstruction of the (a) intensity (amplitude squared), and (b) phase. The intensity and phase are from a 50-hologram series that was summed using the technique described in chapter 4. Due to the low bias on the upper biprism  $BP_1$ , Fresnel fringes are evident in the intensity. They were removed from the phase by means of a reference hologram series. Color bars give the dose in (a) electrons per reconstructed pixel and (b) phase shift in radians.

From Eqn. 1.8 it can be seen that the phase error is minimized by maximizing the current density and the visibility. In order to improve the effective coherence of the source, the image of the source on the specimen plane is typically highly magnified (i.e. defocused). The source is magnified by widely-spreading the illumination with the condenser lenses. Widely-spread illumination has reasonable spatial coherence, in that the trajectories of the population of electrons that compose the 'beam' are reasonably parallel. However, spreading the illumination results in a corresponding

reduction in the beam current density, lowering  $\mu$ . At high magnification ( $q_c = 10 \text{ nm}^{-1}$ ) the typical phase error on the order of 0.06 rad ( $2\pi/100$ ) is achieved [29]. As discussed in section 5.1, this phase error is often insufficient to sample electric and magnetic phase shift on the nanometer scale. In chapter 4 we show that summation of hologram series may be used to improve the phase error by an order of magnitude.

(2.) The sampling requirement for holographic fringes and finite number of detector pixels limits the field of view that may be surveyed. Based on the detector modulation-transfer function results discussed in chapter 3, we generally aim for 10 pixels per fringe to avoid excessive damping of fringe contrast. In turn if atomic resolution is desired, that implies that approximately 30 pixels per inter-atomic distance are required. On a  $2048 \times 2048$  pixel detector, that implies a field of view only 70 inter-atomic distances across. In medium-resolution holography, spatial resolution is traded for a larger field of view.

(3.) The reference wave requires a field-free vacuum region adjacent to the specimen. This often makes specimen preparation challenging and laborious. The problem is illustrated with focused ion beam (FIB) prepared cross-sections. The portion of the specimen adjacent to vacuum typically suffers from radiation damage and Ga implantation. The separation between the object and reference waves can be increased by higher biprism bias, but this negatively impacts the visibility. The limitation may be overcome by using a flat and featureless portion of the specimen for the reference wave [8,32], at the expense of loss of coherence due to incoherent scattering within the reference area of the specimen.

## 1.4 QUANTITATIVE ELECTRON PHASE MICROSCOPY

Transmission electron microscopy suffers from the fact that many, often unknown, factors contribute to the formation of the electro-optical image on the detector plane. A discussion of some of the challenges of quantitative electron microscopy can be found in [20] and references therein. Quantitative phase microscopy, in the form of off-axis holography, has its own particular challenges as discussed in [2-4,33,34] and references therein.

Rotational symmetric illumination was used in this thesis rather than the typical elliptical illumination of electron holography to ensure isoplanar illumination [35]. Rotational symmetric illumination has a significant performance penalty in terms of electron beam current density and a small penalty in terms of holographic visibility compared to elliptical illumination. When aberration-corrected electron microscopy and electron holography are combined reasonable isoplanicity may be obtained with elliptical illumination [28,36].

When discussing quantitative measures, we are generally concerned with its precision, repeatability, and accuracy. Precision is typically the easiest criteria to satisfy by accumulating sufficient statistics. An appropriate example is the

measurement of hologram visibility. The standard deviation of the visibility was derived by Lenz [21,31], and has the form,

$$\sigma_V = \sqrt{\frac{2 - V^2}{\mu}} \quad (1.9)$$

where  $\mu$  is the electron dose. A typical electron hologram may have  $V = 0.25$  and  $\mu = 150 \text{ e}^-/\text{pix}$ . If we want to compute the average visibility across the entire hologram, then the dose is integrated across  $2048 \times 2048 \text{ pix}$ , so the precision of the visibility is  $\sigma_V \sim 0.001$ . In this sense, we are referring to the precision as the resolution of the standard error of an individual measurement.

Accuracy is harder to ascertain because of the potential for unaccounted biases to act upon the measurement [38]. Holographic visibility is biased towards higher values in the presence of shot noise. In the case of computing the average visibility across the hologram, the impact of shot noise should be minimal, but if it is experimentally desirable to measure it over a smaller area, an empirical correction should be made. Assessment of accuracy in science is often done by comparing the consistency of results from different methods. In TEM, there is often the difficulty that other nanoscale characterization tools, such as atom probe and x-ray microscopy, may not collect equivalent information.

Repeatability is the variation in measured values for the same specimen under the same microscope parameters [39]. For example, in chapter 3 the visibility of reference holograms is used to sample the detector MTF over a range of magnifications. Several holograms are taken at each spatial frequency to sample the repeatability of the holographic visibility. We find the repeatability is approximately  $\Delta V \sim 0.025$ , which is an order of magnitude higher than the precision of the visibility measurement. The reason for the discrepancy is that the TEM drifts out of alignment over time, and individual holograms are disturbed by transient events.

A common approach to verification in electron microscopy is to match experiment with simulation (i.e. computational microscopy). An outstanding accuracy problem in the computational microscopy counterpart to high-resolution imaging is the Stobb's factor [40,41]. The Stobb's factor is a long-standing discrepancy between the observed contrast of crystal lattice fringes and the simulated contrast. As is discussed in section 1.3.1, the large number of factors that make TEM a complicated problem which is prohibitively difficult to match experiment with simulation.

In electron holography, the visibility and phase shift of holographic fringes in vacuum may be measured. This implies that the bias from all environmental and instrument effects can be measured from the vacuum reference. The ratio for visibility of a vacuum reference hologram to an object hologram gives the loss of visibility (or coherence) due to the specimen alone, within the repeatability of the

measurement. In this sense, it is often more straight-forward to achieve quantitative results for electron holography compared to high-resolution imaging.

Continuing with the example case, as Eqn. 1.9 is overprecise, it should be corrected. We expect that if enough of the environmental factors that affect the repeatability were included in the expression we would find agreement between experiment and analysis. This is the approach taken in chapter 4, where the analytical and experimental phase error are compared and reasonable agreement is found.

#### 1.4.1 OPTICAL-TRANSFER FUNCTION CASCADE

In an optical system, it is often helpful to utilize the formalization of linear transfer theory to model the transfer of spatial frequencies for each component of the TEM. The optical transfer function is defined as

$$\text{OTF}(q) = \text{MTF}(q)e^{i\text{PTF}(q)} \quad (1.10)$$

Often the OTF of each component in the TEM has different names in the literature. Zernike terms the OTF of the source the complex degree of coherence while Frank combines it with the illumination lenses to create the *phase-coherence transfer function* (PCTF) [9, 42]. For the objective lens alone, the MTF is the Airy function of the limiting aperture, i.e. the bore of the lens, while the PTF is the *contrast-transfer function* (CTF). For the projector lenses, the OTF is expressed by the barrel and pincushion distortion of each lens [43]. In electron holography, each OTF may be called a visibility  $V_x$  whereas in high-resolution imaging simulation the OTFs may be represented as loss of contrast  $C_x$ .

Modeling the microscope as a cascade of transfer functions is an approach to achieve quantitative results from electron holography that has been taken in this thesis. The frequency-space representation of OTF is chosen over the image-space representation of *point-spread function* (PSF) to render analysis straight-forward. If the overall PSF of the TEM is broken-down into its constituents, it becomes a convolution cascade. If the OTF of the TEM is broken-down into its constituents, it becomes a multiplicative cascade, which is both simpler to conceptualize and calculate. I.e.,

$$\begin{aligned} \text{PSF}_{TEM} &= \text{PSF}_1 \otimes \text{PSF}_2 \otimes \dots \\ \text{OTF}_{TEM} &= \text{OTF}_1 \cdot \text{OTF}_2 \cdot \dots \end{aligned} \quad (1.11)$$

In this thesis portions of the TEM have been modeled as an OTF cascade as a means to achieve quantitative results. The approach was inspired by the description of holographic visibility by Lehmann [12],

$$V = \gamma_{source} V_{object} V_{instability} \text{MTF}_{ccd} \quad (1.12)$$

Each individual OTF in Eqn. 1.11 reflects numerous individual components and/or physical phenomena. As such, the OTF can be thought of as not just a cascade, but also a hierarchy. An example OTF hierarchy is shown in Figure 1.5, based on the Hitachi HF-3300 located at the National Institute for Nanotechnology. Groups of

OTFs are usually organized by how they are experimentally measured. In many cases, the individual OTFs cannot be separated experimentally, but the conceptual distinction can be valuable to understand the behavior of the microscope.

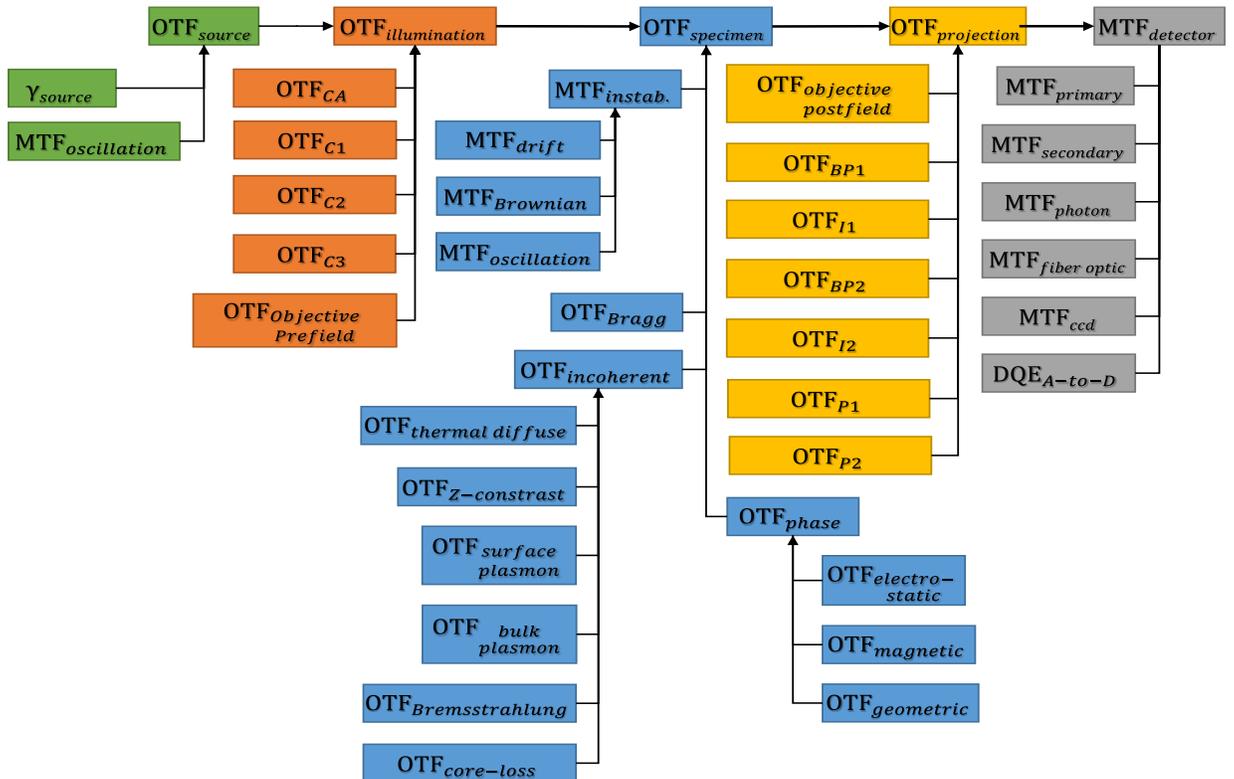


Figure 1.5: An example model of Hitachi HF-3300 TEM for electron holography mode. Experimental conditions can be varied to measure individual optical-transfer functions, but more often the functions are grouped, as shown by the hierarchical structure of the cascades. The figure is not intended to be exhaustive, but rather illustrate the complexity of quantitative experiments in transmission electron microscopy.

Experimentally, it is often possible, with careful consideration, to vary one parameter of the microscope and record a series of measurements. In this way, individual OTFs can often be examined. Electron holography is invaluable for experimental determination of OTFs because of the ability to create sinusoidal patterns, i.e. reference holograms where there is no specimen in the field of view. For example, in section 4.2 the biprism bias is adjusted over a wide range for reference holograms. The experiment samples the product of the source OTF, the detector OTF, and the instabilities of the biprisms themselves. The detector MTF is deconvolved using the holographic MTF found in section 3.2.3, leaving the product of the complex degree of coherence of the source and source and biprism instabilities.

#### 1.4.2 AUTOMATED TEM CONTROL AND PROCESSING

One of the major themes in the research content of this thesis is the use of automated control of the instrument and its peripherals. Historically, electron microscopy has been manpower-intensive, with human inputs at every step of the experiment and the processing of data afterwards. Automation of the TEM is hardly new [44] but it has been applied herein in a systemic fashion.

Here we use automation in the sense that it applies to the expert user. The expert not only uses automation but develops it and adjusts it to the particular experiment. Essentially then, the software is just another peripheral device of the instrument. There are numerous examples of the efficiency of the automated approach in collecting and analyzing data throughout this thesis. In chapter 2 the Fourier-ratio algorithm is developed to calculate the holographic visibility automatically. In addition, the bias in visibility due to shot noise was also evaluated with autonomous Matlab scripts. In chapter 3, long series of holograms at varied magnification were used as a stand-alone method to characterize the MTF of the detector. In chapter 4, extensive use of automated reconstruction and registration techniques are applied to successfully sum hologram series. Additionally, many of the techniques used to characterize drift in the TEM were performed with both automated experiments as well as post-experiment processing.

## 1.5 REFERENCES

- [1] H. Lichte, M. Lehmann, Electron holography: basics and applications, Reports on Progress in Physics, 71 (2008) 016102.
- [2] H. Lichte, P. Formanek, A. Lenk, M. Linck, C. Matzeck, M. Lehmann, P. Simon, Electron Holography: Applications to Materials Questions, Annual Review of Materials Research, 37 (2007) 539-588.
- [3] M.R. McCartney, D.J. Smith, Electron holography: Phase imaging with nanometer resolution, in: Annual Review of Materials Research, 2007, pp. 729-767.
- [4] M.R. McCartney, N. Agarwal, S. Chung, D.A. Cullen, M.-G. Han, K. He, L. Li, H. Wang, L. Zhou, D.J. Smith, Quantitative phase imaging of nanoscale electrostatic and magnetic fields using off-axis electron holography, Ultramicroscopy, 110 (2010) 375-382.
- [5] D.M. Adams, L. Brus, C.E.D. Chidsey, S. Creager, C. Creutz, C.R. Kagan, P.V. Kamat, M. Lieberman, S. Lindsay, R.A. Marcus, R.M. Metzger, M.E. Michel-Beyerle, J.R. Miller, M.D. Newton, D.R. Rolison, O. Sankey, K.S. Schanze, J. Yardley, X. Zhu, Charge transfer on the nanoscale: Current status, Journal of Physical Chemistry B, 107 (2003) 6668-6697.
- [6] H. Lichte, Optimum focus for taking electron holograms, Ultramicroscopy, 38 (1991) 13-22.
- [7] A. Lubk, D. Wolf, H. Lichte, The effect of dynamical scattering in off-axis holographic mean inner potential and inelastic mean free path measurements, Ultramicroscopy, 110 (2010) 438-446.
- [8] M.J. Hytch, F. Houdellier, F. Hue, E. Snoeck, Dark-field electron holography for the measurement of geometric phase, Ultramicroscopy, 111 (2011) 1328-1337.

- [9] D. Gabor, Microscopy by Reconstructed Wave-Fronts, Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 197 (1949) 454-487.
- [10] D. Gabor, Microscopy by Reconstructed Wave Fronts .2, Proceedings of the Physical Society of London Section B, 64 (1951) 449-&.
- [11] D. Gabor, Holography, 1948-1971, Science, 177 (1972) 299-&.
- [12] G. Mollenstedt, H. Duker, Beobachtungen Und Messungen an Biprisma-Interferenzen Mit Elektronenwellen, Zeitschrift Fur Physik, 145 (1956) 377-397.
- [13] A. Tonomura, J. Endo, T. Matsuda, Application of Electron Holography to Interference Microscopy, Optik, 53 (1979) 143-146.
- [14] J. Endo, T. Matsuda, A. Tonomura, Interference Electron-Microscopy by Means of Holography, Japanese Journal of Applied Physics, 18 (1979) 2291-2294.
- [15] A. Tonomura, T. Matsuda, J. Endo, T. Arie, K. Mihama, Direct Observation of Fine Structure of Magnetic Domain Walls by Electron Holography, Phys. Rev. Lett., 44 (1980) 1430-1433.
- [16] Y. Aharonov, D. Bohm, Significance of Electromagnetic Potentials in the Quantum Theory, Physical Review, 115 (1959) 485-491.
- [17] A. Tonomura, T. Matsuda, R. Suzuki, A. Fukuhara, N. Osakabe, H. Umezaki, J. Endo, K. Shinagawa, Y. Sugita, H. Fujiwara, Observation of Aharonov-Bohm Effect by Electron Holography, Phys. Rev. Lett., 48 (1982) 1443-1446.
- [18] A. Tonomura, N. Osakabe, T. Matsuda, T. Kawasaki, J. Endo, S. Yano, H. Yamada, Evidence for Aharonov-Bohm Effect with Magnetic-Field Completely Shielded from Electron Wave, Phys. Rev. Lett., 56 (1986) 792-795.
- [19] H. Lichte, Electron Holography Approaching Atomic Resolution, Ultramicroscopy, 20 (1986) 293-304.
- [20] Q. Fu, H. Lichte, E. Volkl, Correction of Aberrations of an Electron-Microscope by Means of Electron Holography, Phys. Rev. Lett., 67 (1991) 2319-2322.
- [21] H. Lichte, K.-H. Herrmann, F. Lenz, Electron noise in off-axis image plane holography, Optik, 77 (1987) 135-140.
- [22] H. Lichte, Optimum focus for taking electron holograms, Ultramicroscopy, 38 (1991) 13-22.
- [23] F. Zernike, The concept of degree of coherence and its application to optical problems, Physica V, 8 (1938) 785-795.
- [24] H. Shimoyama, S. Maruse, Theoretical considerations on electron optical brightness for thermionic, field and T-F emissions, Ultramicroscopy, 15 (1984) 239-254.
- [25] P.A. Midgley, An introduction to off-axis electron holography, Micron, 32 (2001) 167-184.
- [26] M. Lehmann, H. Lichte, Tutorial on Off-Axis Electron Holography, Microscopy and Microanalysis, 8 (2002) 447-466.
- [27] K. Harada, A. Tonomura, Y. Togawa, T. Akashi, T. Matsuda, Double-biprism electron interferometry, Applied Physics Letters, 84 (2004) 3229-3231.
- [28] D. Geiger, H. Lichte, M. Linck, M. Lehmann, Electron Holography with a Cs-Corrected Transmission Electron Microscope, Microscopy and Microanalysis, 14 (2008) 68-81.

- [29] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, T. Niermann, State of the art in atomic resolution off-axis electron holography, *Ultramicroscopy*, 116 (2012) 13-23.
- [30] N.R. Goodman, Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction), *Annals of Mathematical Statistics*, 34 (1963) 152-177.
- [31] F. Lenz, Statistics of phase and contrast determination in electron holograms, *Optik*, 79 (1988) 13-14.
- [32] M. Hytch, F. Houdellier, F. Hue, E. Snoeck, Nanoscale holographic interferometry for strain measurements in electronic devices, *Nature*, 453 (2008) 1086-1089.
- [33] D.J. Smith, Progress and problems for atomic-resolution electron microscopy, *Micron*, 43 (2012) 504-508.
- [34] R.E. Dunin-Borkowski, M.R. McCartney, D.J. Smith, S.S.P. Parkin, Towards quantitative electron holography of magnetic thin films using in situ magnetization reversal, *Ultramicroscopy*, 74 (1998) 61-73.
- [35] D.J. Smith, W.J. De Ruijter, J.K. Weiss, M.R. McCartney, Quantitative Electron Holography, in: E. Voelkl, L.F. Allard, D.C. Joy (Eds.) *Introduction to Electron Holography*, Kluwer Academic, New York, 1999.
- [36] M. Lehmann, Influence of the elliptical illumination on acquisition and correction of coherent aberrations in high-resolution electron holography, *Ultramicroscopy*, 100 (2004) 9-23.
- [37] H. Lichte, M. Linck, D. Geiger, M. Lehmann, Aberration Correction and Electron Holography, *Microscopy and Microanalysis*, 16 (2010) 434-440.
- [38] D.L. Sackett, Bias in analytic research, *Journal of Chronic Diseases*, 32 (1979) 51-63.
- [39] M.J. Bland, D.G. Altman, Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement, *The Lancet*, 327 (1986) 307-310.
- [40] M.J. Hytch, W.M. Stobbs, Quantitative comparison of high resolution TEM images with image simulations, *Ultramicroscopy*, 53 (1994) 191-203.
- [41] A. Howie, Hunting the Stobbs factor, *Ultramicroscopy*, 98 (2004) 73-79.
- [42] J. Frank, The Envelope of Electron Microscopic Transfer Functions for Partially Coherent Illumination, *Optik*, 38 (1973) 519-536.
- [43] F. Hue, C.L. Johnson, S. Lartigue-Korinek, G. Wang, P.R. Buseck, M.J. Hytch, Calibration of projector lens distortions, *Journal of Electron Microscopy*, 54 (2005) 181-190.
- [44] A.J. Koster, W.J. de Ruijter, Practical autoalignment of transmission electron microscopes, *Ultramicroscopy*, 40 (1992) 89-107.

## 2 DETERMINATION OF LOCALIZED VISIBILITY IN OFF-AXIS ELECTRON HOLOGRAPHY

The contents of Chapter 2 were submitted for publication to *Ultramicroscopy* by R.A. McLeod and M. Malac on December 17<sup>th</sup>, 2013. Contributing author M. Malac edited the paper and provided valuable discussion throughout the development process.

### *ABSTRACT*

Off-axis electron holography is an interferometric method known as ‘wavefront-splitting’ in the literature, whereby a Fresnel biprism splits the source into two virtual sources. Electron holography allows the phase shift and amplitude of the electron wavefront to be separated and quantitatively measured. An additional component of the holographic signal is the coherence of the electron wavefront. Historically, coherence has been evaluated by measurement of the holographic fringe visibility (or contrast) based on the minimum and maximum intensity values. A new method is presented here, based on statistical moments that allow the visibility to be measured in a deterministic and reproducible fashion suitable for quantitative analysis. An algorithm that allows the visibility to be resolved in two-dimensions is provided, which we term the *local visibility*. Thus the local visibility constitutes a previously unused *third component signal* in an electron hologram, which may be used to evaluate the loss of coherence due to electron scattering within a specimen, or as an aid in image analysis and processing. All visibility metrics are biased at low-dose conditions by the presence of shot-noise, which can be empirically normalized to achieve linear response.

### 2.1 INTRODUCTION AND MOTIVATION

We introduce a new method to evaluate the visibility (or contrast) of holographic fringes in the *transmission electron microscope* (TEM), which represents the coherence of the electron wavefront. The method uses statistical moments to generate a *two-dimensional* (2-D) map of the local visibility. The existing visibility measure uses minimum and maximum values which, as explained in Section 2.2, does not have an associated statistical uncertainty and cannot be used to map the visibility in 2-D. In Section 2.2 the new statistical method is introduced. In Section 2.3 we provide an algorithm to compute the localized visibility in 2-D. In Section 2.4 we compare the performance of the visibility metrics in the presence of shot noise in low dose conditions and explain how to empirically correct for the bias from

noise. In Section 2.5 we provide examples for how the localized visibility may be applied for experimental characterization with an example specimen of Ni-NiO core-shell particles.

*High-resolution electron microscopy* (HRTEM) is a conventional TEM technique that is sometimes considered qualitative because the amplitude and phase shift of the electron wavefront is convolved with the contrast transfer function of the optics and contrast reversal makes the amplitude and phase inseparable [1]. *Electron holography* (EH), and in particular the off-axis method, has been used by the electron microscopy community to quantitatively measure electron phase shift in TEM for characterization of sample properties [2]. In addition, the degree of partial coherence, known as holographic visibility, can be measured from an electron hologram but is not typically utilized. The coherence of an electron wavefront is a function of the ensemble electron population, [3,4], that is, the electron population's angular and energy distributions. An interference pattern is typically recorded from many millions of electrons, so the visibility is a stochastic metric which provides insight on electron scattering in materials [5,6].

In the off-axis method, one or two electrostatic biprisms separates the electron wavefront into two halves [7]. The biprism, which is directly analogous to a Fresnel biprism in light optics, is a very thin conducting wire that is electrically biased [8]. As shown in Figure 2.1, one half of the wavefront passes through the specimen, known as the object wave while the other half-wavefront passes through vacuum adjacent to the specimen, known as the reference wave. The potential on the biprism deflects the two halves of the wavefront together at a fixed angle such that the two waves overlap on a detector where a hologram (i.e. interference pattern) is recorded. When there is no specimen in the beam, both the object and reference waves pass through vacuum and the interference generates a regular set of periodic sinusoidal interference fringes. When a specimen is introduced, fringe modulus represents the wavefront amplitude and the deviation of the fringes from their expected periodicity encodes the wavefront phase shift, while the fringe visibility (or contrast) represents the wavefront coherence. The electron wavefront is phased shifted by the scalar electric and magnetic vector potentials of the specimen, known as the Aharonov-Bohm effect [9-11]. This implies that the electron wavefront phase shift is a quantitative measure of the electric and magnetic fields/potentials of a specimen integrated along the path of the beam.

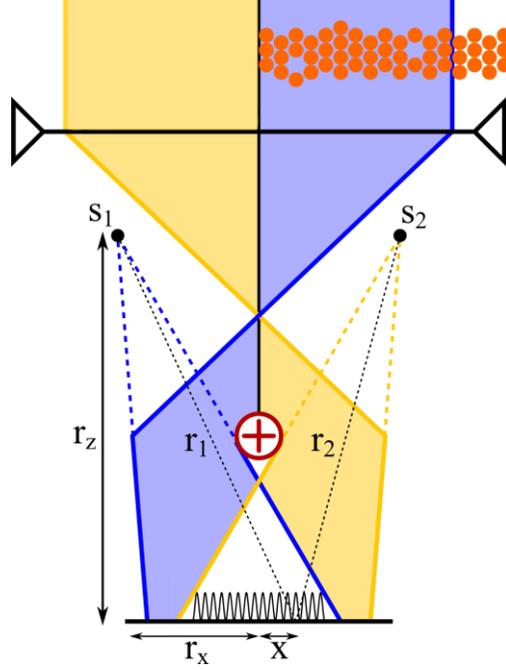


Figure 2.1: A schematic representation of off-axis electron holography configuration. The object wavefront (blue) travels through the specimen and the reference wavefront (yellow) through the adjacent vacuum, providing a phase reference. An electrostatic biprism below the specimen deflects the two half-beams together, which generates two virtual sources  $s_1$  and  $s_2$  that are separated by a distance  $2r_x$ , at a height  $r_z$  from the detector. The vector  $x$  indicates the pixel position on the detector. For purposes of this paper we will assume that the virtual sources produce planar waves, such that  $r_z \gg r_x$ .

The general form of a planar-wave hologram, not including the envelope function that determines the extent of the interference region and Fresnel fringes, is represented as two planar waves  $\psi_1$  and  $\psi_2$  that converge at an semi-angle  $\theta$ ,

$$\psi_1(\mathbf{x}) = A_1(\mathbf{x}) \exp \left( i(\mathbf{k} \cdot \mathbf{r}_1 + \omega t + \phi_1(\mathbf{x})) \right)$$

$$\psi_2(\mathbf{x}) = A_2(\mathbf{x}) \exp \left( i(\mathbf{k} \cdot \mathbf{r}_2 + \omega t + \phi_2(\mathbf{x})) \right)$$

where  $\mathbf{r}$  is the displacement from the virtual source as a function of pixel position on the detector  $\mathbf{x}$ ,  $A$  is the wavefront amplitude,  $\mathbf{k}$  is the electron wavevector,  $\omega$  is the relativistic electron frequency,  $t$  is time, and  $\phi$  is the wavefront phase shift. For the steady-state (time integrated) equation the time-variable term  $\omega t$  is omitted. The object and reference halves of the electron wavefront are denoted by subscript 1 and 2, respectively. Often a reference hologram is recorded with the specimen removed from the field of view to normalize the phase-shift of the electron optics, in which case the object and reference holograms are denoted by  $o$  and  $r$  subscripts respectively.

Plane waves with a small convergence angle  $\theta$ , typically a few microradians in EH, interfere on the detector to generate plane-parallel fringes with carrier frequency  $\mathbf{q}_c = 2\mathbf{k}\sin(\theta)$ ,

$$\begin{aligned} \psi_{12}^2(\mathbf{x}) = & A_1^2(\mathbf{x}) + A_2^2(\mathbf{x}) \\ & + 2V(\mathbf{x}, \mathbf{q}_c, \alpha_o)A_1(\mathbf{x})A_2(\mathbf{x})\cos(2k\sin(\theta)\mathbf{x} + \phi_1(\mathbf{x}) \\ & - \phi_2(\mathbf{x})) \end{aligned} \quad (2.1)$$

where  $V(\mathbf{x}, \mathbf{q}_c, \alpha_o)$  is the holographic visibility, which is a function of position  $\mathbf{x}$ , the separation of the virtual sources as determined by the carrier frequency  $\mathbf{q}_c$ , and the half-angle of the illumination  $\alpha_o$ .

### 2.1.1 MEASUREMENT AND PHYSICAL ORIGIN OF THE VISIBILITY

Visibility measures the degree of coherence of an interference condition and is a measure of practical hologram quality. The visibility impacts the signal-to-noise ratio and achievable phase resolution [12-14]. The concept of visibility was first discussed by Michelson [15], who calculated the visibility in terms of the observed intensities of interference fringes,

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (2.2)$$

with  $V$  being the visibility and  $I_{max}$  and  $I_{min}$  the measured maximum and minimum intensity, respectively. Visibility spans the range [0,1] with the limits being perfectly incoherent and perfectly coherent conditions, respectively. Zernike later developed a more rigorous formalism for partial coherence, called the complex degree of coherence [16].

In a TEM, low beam current and high electron velocity ensures that interference is a single-particle phenomenon; that is, each electron interferes with itself. Over the ensemble of many electrons, the electron probability density functions (PDFs) overlap stochastically, which produces interference pattern [17]. Coherence in TEM is a measure of how similar the PDFs of the ensemble electron population are to each other [3]. The ensemble PDF is determined by the distribution in momentum  $\Delta\mathbf{k}$ , corresponding to spatial coherence, and energy  $\Delta E$ , corresponding to temporal coherence, of the electron population.

The coherence is modeled in optics in terms of spatial frequency using *optical-transfer functions*,  $OTF(\mathbf{q}) = MTF(\mathbf{q})\exp(-i \cdot PTF(\mathbf{q}))$ , where PTF is the phase-transfer function and  $\mathbf{q}$  is the spatial frequency coordinate. Each component involved in image formation in the TEM has an associated OTF, and by linear systems theory [18], functions are cascaded, such that

$$\begin{aligned} V(\mathbf{q}, q_c, \alpha_o) = & |\gamma_{source}(\mathbf{q})OTF_{optics}(\mathbf{q}, \alpha_o)OTF_{specimen}(\mathbf{q}, \alpha_o)| \\ & \cdot MTF_{instability}(\mathbf{q})MTF_{detector}(\mathbf{q}) \end{aligned} \quad (2.3)$$

where  $\gamma$  is the complex degree of coherence of the source and  $\alpha_o$  is the half-angle of illumination [3,19]. The dependence of visibility on spatial frequency implicitly

includes the separation of virtual sources seen in off-axis holography, making the substitution  $\mathbf{q} \rightarrow \mathbf{q}_c + \mathbf{q}$ , where  $\mathbf{q}_c$  is the carrier frequency of the fringe pattern into  $\gamma_{\text{source}}$ . The complex degree of coherence of the source can be measured by different means [20,21], and has been modeled as a *phase-coherence transfer function* (PCTF) [22,23]. The OTF of the optics is termed the *contrast-transfer function* (CTF) and/or its damping envelope [22,24]. The MTF due to instabilities can in principle be calculated from the measured resonances affecting the TEM [25]. Long-period transient events and drift of the specimen are more problematic as they are not deterministic. The detector MTF is measured by a variety of means [26-28].

The goal is to isolate OTF of the specimen in Eqn. 2.3 from the other terms to retrieve the coherence of electron scattering information from a specimen. The uncertainties in each OTF in the cascade make measurement of the relative visibility a more practical goal. The visibility of holographic fringes inside a specimen compared to those in vacuum (such as in a reference hologram) represents the relative loss of coherence due to scattering in the specimen. In relative measure the complex component of the OTF is lost, giving the specimen MTF.

## 2.2 STATISTICAL HOLOGRAPHIC VISIBILITY METRIC

The standard min-max metric for visibility devised by Michelson (Eqn. 2.2) has several practical problems. The use of maximum and minimum implies that the visibility covers the entire sampled range of intensity values, and hence is not robust in the presence of noise, as discussed in Section 2.4. Furthermore, the metric inherently cannot quantify the uncertainty in the visibility measurement. In addition, the min-max metric is not typically applied in a deterministic fashion but rather evaluated manually with an intensity profile tool in image processing software to generate a cross-section of the holographic fringes before application of Eqn. 2.2. The human input, i.e. pointing and clicking with a mouse, has the potential to introduce experimenter's bias [29].

A metric which addresses the drawbacks of the min-max method will improve quantitative measurements by electron holography. We propose the following criteria for a new visibility metric:

1. It should be statistically quantifiable, such that statements may be made with regards to the significance of the visibility measurement.
2. It should be deterministic and reproducible.
3. It should be capable of determining the visibility and its variance for an individual hologram at an arbitrary position within the interference region (i.e. localized in 2-D).

The simplest analysis is to apply the substitutions  $I_{max} = \mu + c\sigma$  and  $I_{min} = \mu - c\sigma$  to Eqn. 2.2,

$$V = c \frac{\sigma}{\mu}$$

where  $\mu$  is the analytic mean and  $\sigma$  the standard derivation of the intensity, and  $c$  is a confidence interval [30-33]. The confidence interval that most accurately determines the visibility is found from the statistical moments of Eqn. 2.1,

$$\mu(\mathbf{r}) = \frac{\int_{-\mathbf{k}/\pi}^{\mathbf{k}/\pi} \psi_{12}^2(\mathbf{x}) d\mathbf{x}}{\int_{-\mathbf{k}/\pi}^{\mathbf{k}/\pi} d\mathbf{x}} = A_1(\mathbf{x})^2 + A_2(\mathbf{x})^2$$

$$\sigma^2(\mathbf{x}) = \frac{\int_{-\mathbf{k}/\pi}^{\mathbf{k}/\pi} (\psi_{12}^2(\mathbf{x}) - \mu(\mathbf{x}))^2 d\mathbf{x}}{\int_{-\mathbf{k}/\pi}^{\mathbf{k}/\pi} d\mathbf{x}} = 2V(\mathbf{x})^2 A_1(\mathbf{x})^2 A_2(\mathbf{x})^2$$

Here we find the analytic statistical moments; in practice the sample mean and variance are substituted. The integration requires that the mean and variance be calculated over an integer number of fringe periods, but this requirement is relaxed when several fringe periods are averaged.

To derive analytical expressions for object and reference hologram visibility, one makes the approximation that the object-hologram reference-wave amplitude, reference-hologram 'object'-wave amplitude, and reference-hologram reference-wave amplitude are all equivalent, i.e.  $A_{o2} = A_{r1} = A_{r2}$ . Equating the object- and reference-hologram reference-wave amplitude are equivalent is reasonable. Notwithstanding electron source drift between acquisition of the object and reference hologram, both waves pass through vacuum and identical microscope optics. The equivalence approximation between the reference-hologram, object-wave ( $r1$ ) and the reference-wave ( $r2$ ), amplitude assumes that the illumination is flat-field and has uniform intensity which is very reasonable for typical spread-illumination conditions used in EH. With these approximations, object visibility for flat-field illumination has the form,

$$V_o^2(\mathbf{x}) = \frac{\sigma_o^2(\mathbf{x})}{\mu_r(\mathbf{x}) \left( \mu_o(\mathbf{x}) - \frac{\mu_r(\mathbf{x})}{2} \right)} \quad (2.4.a)$$

For a reference hologram, with no specimen and hence no amplitude scattering,

$$V_r^2(\mathbf{x}) = 2 \frac{\sigma_r^2(\mathbf{x})}{\mu_r^2(\mathbf{x})} \quad (2.4.b)$$

Eqn.s 2.4.a and 2.4.b are the equations that fulfill the three criteria for a new visibility metric presented in this paper. The standard error of the visibility can be

estimated from the statistical moments [12-14,34],  $\sigma_V = \sqrt{(2 - V^2)/\mu}$ , which gives the relations for the object and reference visibility variance as

$$\sigma_{V_o}^2 = \frac{2}{\mu_o(\mathbf{x})} - \frac{\sigma_o^2(\mathbf{x})}{\mu_r(\mathbf{x})\mu_o(\mathbf{x})\left(\mu_o(\mathbf{x}) - \frac{\mu_r(\mathbf{x})}{2}\right)}$$

and

$$\sigma_{V_r}^2 = \frac{2}{\mu_r(\mathbf{x})} - 2\frac{\sigma_r^2(\mathbf{x})}{\mu_r^3(\mathbf{x})}$$

Thus we have expressions for both the visibility and its variance in terms of the sample mean and variance of an area of pixels, for both object holograms and reference holograms.

### 2.2.1 AMPLITUDE IN OFF-AXIS ELECTRON HOLOGRAPHY

In EH, the ratio of the magnitude of the object hologram sideband over the reference hologram sideband is considered to be the *amplitude* signal. However, analytically the sideband represents product of the visibility and amplitude and hence is a product of the two different modes of electron scattering.

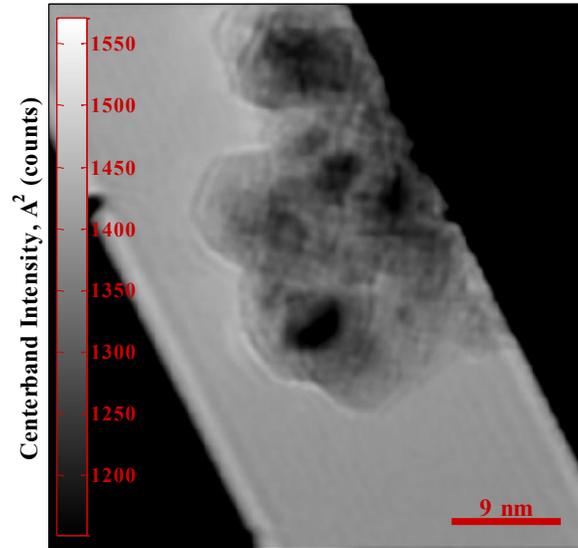
$$\text{holographic amplitude} = \frac{V_o A_{o1} A_{o2}}{V_r A_{r1} A_{r2}} \cong \frac{V_o A_{o1}}{V_r A_{r1}}$$

If the amplitude represents angular-transfer-of-momentum scattering and the visibility loss-of-coherence scattering, then for quantitative measurement of either, the two signals must be deconvolved. The amplitude can be isolated in a similar manner to the visibility,

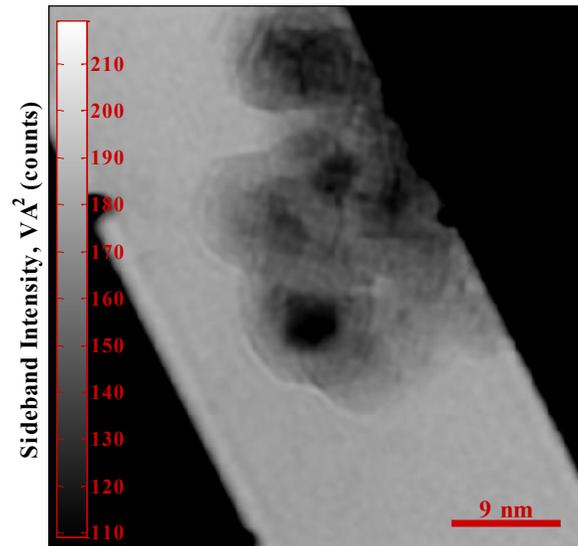
$$A_o^2 = \mu_o - \frac{\mu_r}{2} \tag{2.5.a}$$

$$A_r^2 = \frac{\mu_r}{2} \tag{2.5.b}$$

The centerband signal is stronger than the sideband, as shown in Figure 2.2, because the visibility is always less than one and the DQE of the detector decreases with increasing spatial frequency.



(a)



(b)

Figure 2.2: The intensity (amplitude squared) signal computed from the (a) centerband as compared to the (b) sideband for Ni-NiO core-shell particles. Experimental details discussed in Section 2.5.

### 2.3 AN ALGORITHM FOR LOCAL VISIBILITY MAP GENERATION

We previously reported on a method where the statistical moments sampled over small sub-areas and image transforms were used to generate a two-dimensional map of the local visibility using Eqn. 2.4.a and 2.4.b [35]. Here we present an alternative method based on the 2-D Fourier ratio of sideband-to-centerband signal and show that it is mathematically equivalent to the statistical metrics (Eqn. 2.4.a and 2.4.b). The centerband-to-sideband Fourier ratio method, presented below, is more intuitive, easier to implement, and less prone to generate image artifacts than directly using the statistical moments.

When a 2-D discrete Fourier transform is applied to a hologram, there are two obvious peaks, 1.) a centerband (or autocorrelation) at the origin, and 2.) a sideband and conjugate sideband, located at the spatial carrier frequency  $\mathbf{q}_c$  of the holographic fringes, as shown in Figure 2.3. The ratio of the sideband maximal pixel over the centerband maximal pixel provides the mean visibility over the entire field of view [19,36]. Often the hologram does not extend over the entire field of view, so the region without interference fringes contributes to the average visibility with a value of approximately zero when evaluated with the maximal pixel method. A 2-D map of the visibility is required to compute the average visibility in most cases.

Analytic representations of the centerband (CB) and sideband (SB) are based on Eqn. 2.1,

$$CB(\mathbf{x}) = A_1(\mathbf{x})^2 + A_2(\mathbf{x})^2 = \mu$$

$$SB(\mathbf{x}) = 2V(\mathbf{x})A_1(\mathbf{x})A_2(\mathbf{x}) \cos(2kx\sin(\theta) + \phi_1(\mathbf{x}) - \phi_2(\mathbf{x}))$$

The sideband term is averaged over an integer number of fringe periods such that the cosine term vanishes,

$$\langle SB(\mathbf{x}) \rangle = 2V(\mathbf{x})A_1(\mathbf{x})A_2(\mathbf{x}) = \sqrt{2}\sigma$$

The sideband is averaged by isolation with an apodization function and shifting it to zero-frequency in exactly the same manner as hologram reconstruction [37]. Thus the sideband and centerband are equivalent to the above results derived through the statistical moments. The formulas for the object and reference hologram visibility then are,

$$V_o^2(\mathbf{x}) = \frac{\langle SB_o(\mathbf{x}) \rangle^2}{2 \cdot CB_r(\mathbf{x}) \left( CB_o(\mathbf{x}) - \frac{1}{2} CB_r(\mathbf{x}) \right)} \quad (2.6.a)$$

$$V_r^2(\mathbf{x}) = \frac{\langle SB_r(\mathbf{x}) \rangle^2}{CB_r(\mathbf{x})^2} \quad (2.6.b)$$

An algorithm to compute the results from Eqn.s 2.6.a and 2.6.b was developed in MATLAB. Holograms are first imported into MATLAB from Gatan Digital Micrograph [38]. The steps are illustrated in Figure 2.3 by means of a reference hologram taken on a Hitachi HF3300 TEM at 100 keV equipped with the double biprism electron holography configuration [7]. The algorithm has also been tested on single biprism holograms and works as expected.

The acquired data are processed by application of the following steps:

1. **Pre-process** the image by application of high-quality dark and gain references. Remove artifacts, such as saturated pixels generated by x-rays, by application of a confidence interval median filter, as shown in Figure 2.3(a).

2. **Determine the sideband wavevector** by application of a 2-D discrete Fourier transform to the hologram (typically the reference hologram if one is available). Then,
  - a. Fit and subtract the centerband background by taking the rotational average of the Fourier amplitude and fitting a power-law.
  - b. Sub-divide the Fourier transform into quadrants, which isolates the sideband from its complex conjugate.
  - c. Pick the maximum values as the pixel position corresponding to the sideband wave-vector in each quadrant.

With the sideband position determined, we return to the original hologram.

3. **Apply a rectangular Hamming window** over the entire hologram, as shown in Figure 2.3(b), and then take the Fourier transform. This reduces the streaking normally seen in reciprocal-space due to the finite image size so that the counts from the sideband and centerband do not cross-contaminate. The Hamming window has the form,

$$A_{Hamming} = \left( \frac{27}{50} + \frac{23}{50} \cos\left(\frac{2\pi x}{a}\right) \right) \cdot \left( \frac{27}{50} + \frac{23}{50} \cos\left(\frac{2\pi y}{a}\right) \right)$$

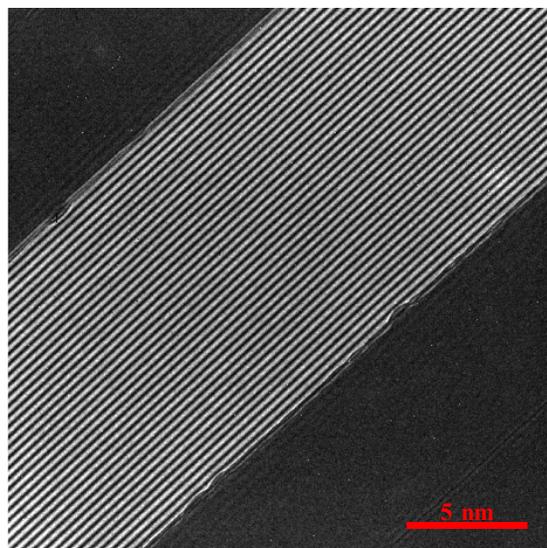
where  $x$  and  $y$  are the associated pixel coordinates and  $a$  is the length of the detector in pixels. This step is essential for consistent results as otherwise counts are lost to the streak artifacts. The Hamming window is the best choice for apodization as it does not force the edges to zero, and hence can be deconvolved after band separation.

4. **Isolate the sideband and centerband** separately through the use of apodization functions applied in the reciprocal-space representation of the hologram computed in step 2, as shown in Figure 2.3(c). A truncated Gaussian apodization function provides good performance tests of simulated holograms [35],

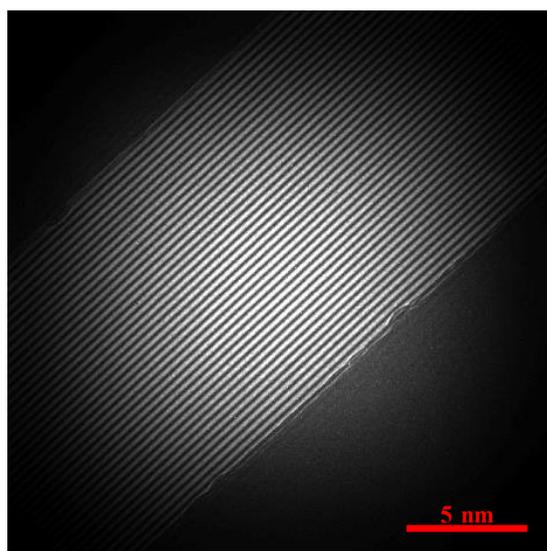
$$A_{Gauss} = (\mathbf{q}(u, v) < a) \exp(-(\mathbf{q}(u, v)/a)^2)$$

where  $\mathbf{q}$  is the reciprocal position vector formed by the pixel coordinates  $u$  and  $v$  and  $a$  is the radius of the window in pixels. The von Hann window is also popular and has similar performance. An apodization down-samples the hologram considerably, since the maximum radius  $a$  is half the sideband wavevector.

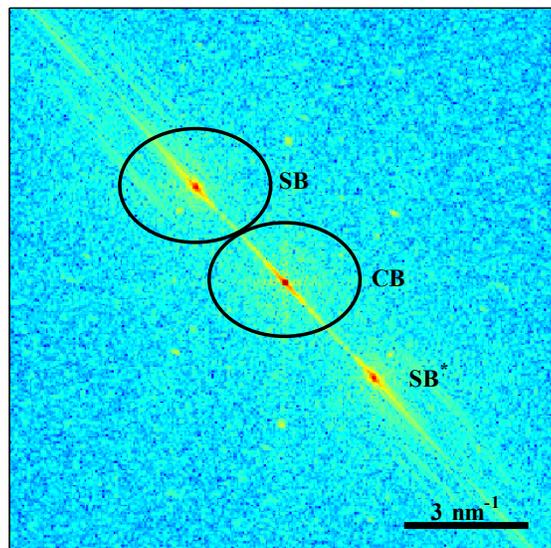
5. **Apply an inverse 2-D discrete Fourier transform** to the filtered sideband and center-band, separately. Both should have dimensions of  $(2a, 2a)$ .
6. **Calculate the visibility** per pixel using the appropriate ratio Eqn. 2.6.a or 2.6.b. This step also deconvolves the Hamming window applied in step 3, resulting in a 2-D map of the visibility, as shown in Figure 2.3.d.



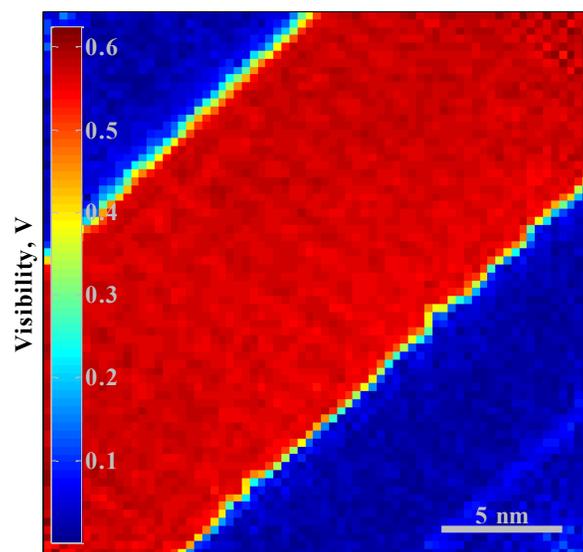
(a)



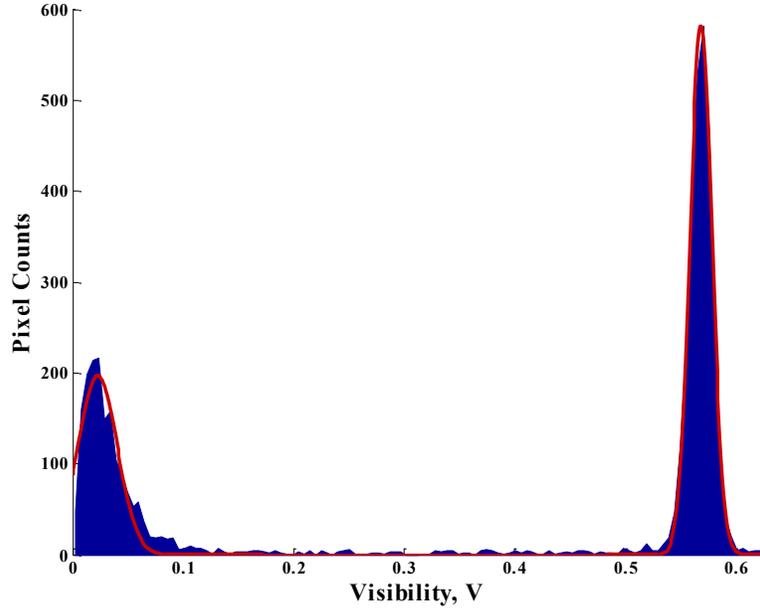
(b)



(c)



(d)



(e)

Figure 2.3: (a) An example reference hologram. The visibility was manually evaluated as  $V_{min-max} = 0.58$ . (b) The reference hologram with a rectangular Hamming window applied. (c) The Fourier representation of the Hamming-apodized hologram, with the centerband (CB) and sideband (SB) indicated. Note the absence of vertical and horizontal streaks due to the application of the Hamming apodization in the previous step. (d) The resulting visibility map for a reference hologram, showing the visibility across the field of view, with the interference region having nearly constant visibility and the region outside of it nearly zero visibility ( $\sim 0.02$ ), as expected. (e) A histogram of the visibility map in Fig. 3.4. From the histogram, the mean visibility inside the interference region is  $0.569 \pm 0.018$ . The visibility outside the interference pattern is not zero due to standard deviation from shot noise and faint fringes in the dark-field.

## 2.4 BIAS IN VISIBILITY METRICS

Here we examine the experimental factors that bias visibility measurements described in Sections 2.2 and 2.3. In Section 2.2, the sample mean  $\mu$  and variance  $\sigma^2$  are substituted for the analytic mean and variance. In practice, the substitution implies the metric is biased by shot noise and by detector noise such as dark counts. The bias may be estimated and removed from the visibility metric [30,31],

$$V_o^2(\mathbf{x}) = \frac{\sigma_o^2(\mathbf{x}) - \sigma_N^2}{(\mu_r(\mathbf{x}) - \mu_d) \left( \mu_o(\mathbf{x}) - \frac{\mu_r(\mathbf{x})}{2} - \frac{3\mu_d}{2} \right)} \quad (2.7.a)$$

$$V_r^2(\mathbf{x}) = 2 \frac{\sigma^2(\mathbf{x}) - \sigma_N^2}{(\mu(\mathbf{x}) - \mu_d)^2} \quad (2.7.b)$$

where  $\sigma_N^2$  is the variance due to shot noise and dark current, and  $\mu_d$  the background due to dark counts. The dark noise includes both thermal noise and detector artifacts. A GATAN Ultrascan 1000 detector operating at -25.4 °C has a mean dark current of 5.6 counts and dark variance of  $\pm 6.9$  counts<sup>2</sup> for 1 s exposures. The bulk of the dark variance does not result from thermal noise but from an exponential drop-off at the edges of the detector, orthogonal to the read-out axis, from  $\sim 993$  counts at the edge to  $\sim 1001$  counts in the center of the detector (the default pixel register for a Gatan Ultrascan CCD detector is 1000) [28].

Normalizing for low-dose using Eqn.s 2.7.a and 2.7.b requires characterizing the shot noise, which is complicated by the detector quantum efficiency (DQE). The DQE characterizes the damping of signal and noise as a function of spatial frequency,

$$DQE(q, N) = \frac{SNR_{out}^2}{SNR_{in}^2} = N \frac{MTF^2(q)}{NPS(q, N)}$$

where, N is the mean dose per pixel, MTF is the detector modulation transfer function, and NPS is the noise-power spectrum (which is the square of the noise-transfer function) [26-28].

The complication of the DQE makes empirical normalization of shot noise bias preferable. A large series of holograms ( $\sim 800$ ) was acquired over a range of exposure times designed to give mean counts in the range of 10-500 counts/pix. To minimize the addition of noise from reference images, we made use of high quality reference images, consisting of a  $10^5$ -frame dark reference (with each frame passed through a  $10\text{-}\sigma$  median filter to filter x-rays) and a 300-frame low-mean-dose gain reference. Adjustment of exposure time was judged to be the best way of holding visibility constant [28].

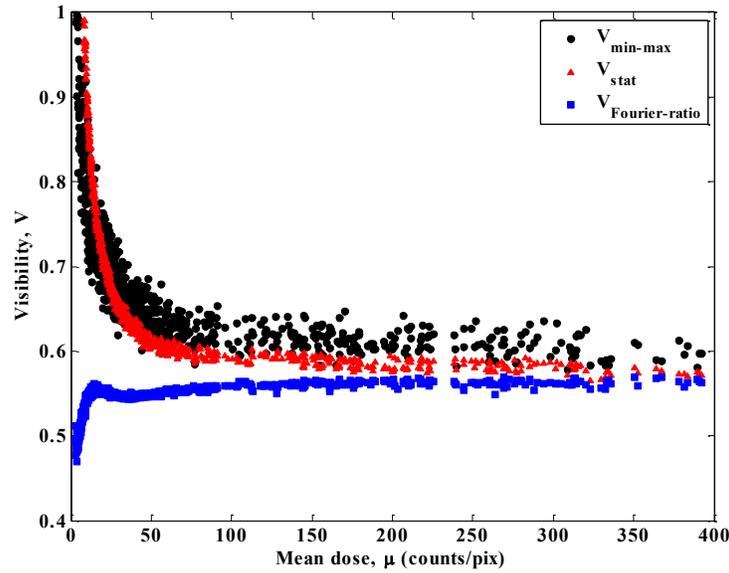
When the statistical moments, are used over a fixed area (Eqn. 2.4.b, Figure 2.4.a red triangles), the standard deviation from shot noise dominates the standard deviation from interference fringes at lower counts, resulting in a power-law increase. In contrast, the sideband-to-centerband Fourier ratio algorithm (Eqn. 2.6.b, Figure 2.4.a blue squares) is largely self-correcting but does break down at very low counts. The Fourier-ratio is self-correcting because the shot-noise in each band is proportional to the signal in each band. The Michelson min-max metric using cross-sectional averaging (Eqn. 2.2, Figure 2.4.a black circles), shows that it is a more heavily biased and less precise version of the statistical method when applied over an identical number of pixels. At typical hologram dose levels (200-500 counts/pix) the statistical and Fourier-ratio methods converge to the same ultimate visibility, but the min-max method overestimates the visibility. The min-max method naturally overestimates the visibility because it considers the visibility to span the entire sampled range, and hence is not robust in the presence of noise. Simulated holograms with shot noise but no dark counts have similar characteristic curves, as shown in Figure 2.4.b.

The Fourier-ratio metric can be corrected by use of a scale reference that normalizes the visibility as a function of dose, similar to gain linearization references used by CCD detectors, shown in Figure 2.5. The visibility-dose relationship is a function of DQE and must be done separately for each accelerating voltage, acquired with the same methodology described above. The visibility-dose reference is fit using a 3<sup>rd</sup>-order rational function by least squares,

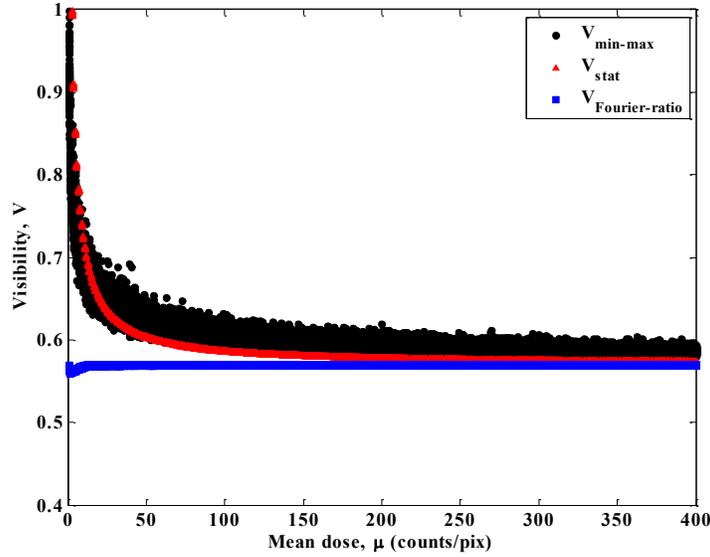
$$V_{fit}(\mu) = \frac{p_1\mu^3 + p_2\mu^3 + p_2\mu + p_4}{\mu^3 + q_3\mu^3 + q_3\mu + q_3}$$

where  $p_m$  and  $q_n$  are the fit parameters. The fit is asymptotic, such that an ultimate visibility can be defined as  $V_{ult} = V_{fit}(\mu \rightarrow \infty)$ . The corrected visibility is then given by,

$$V_{corr}(\mu) = V_{raw} V_{ult} / V_{fit}(\mu)$$



(a)



(b)

Figure 2.4: (a) Scatter plot of measured visibility by all three metrics as a function of dose for a series of 810 holograms at 200 keV. The ultimate visibility was found to be  $V_{ult} = 0.569 \pm 0.003$ . (b) An equivalent scatter plot of visibility metrics for simulated holograms,  $V = 0.569$ , with per-pixel computed Poisson shot noise applied and no dark noise or detector DQE applied.

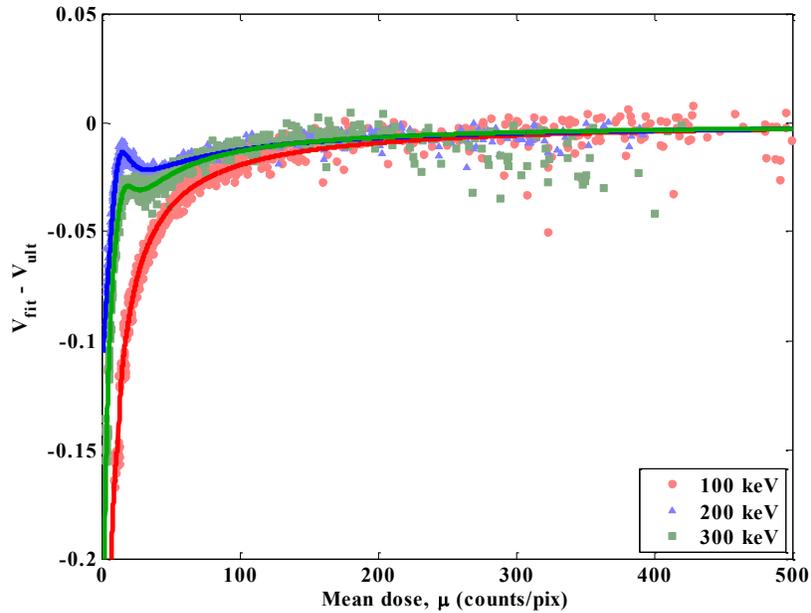


Figure 2.5: The visibility-dose scale references for the Fourier-ratio metric at 100, 200, and 300 keV for the GATAN USC1000 detector on a Hitachi HF3300 TEM (data – shaded dots, best fits – solid lines). There are clear characteristics that change with the accelerating voltage, which indicates signal- and noise-transfer properties of the detector as the source.

## 2.5 EXAMPLE EXPERIMENTAL RESULTS

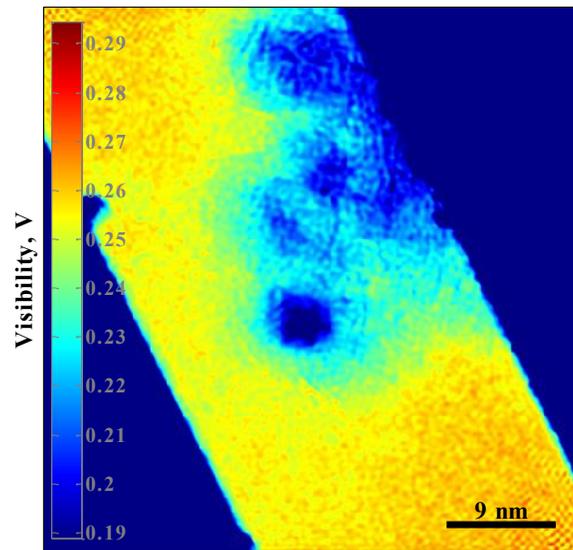
Many TEM methods are an under-constrained problem, i.e. there are too many unknowns and too few signals for quantitative interpretation. The local visibility is valuable in itself as a third signal in electron holography, but also presents unique opportunities to explore the physics and coherence of inelastic scattering. In this section potential applications of the coherence signal obtained from localized visibility are explored for segmenting chemically different areas of a Ni-NiO core-shell nanoparticles, sourced from Quantum Sphere Inc. (qsinano.com), on a lacey Carbon support. All experiments were conducted on a Hitachi HF-3300 which is equipped with a cold-field electron gun, operated at 300 keV, and the double bi-prism configuration, which suppresses the Fresnel fringes seen in single bi-prism holograms [7].

For batch acquisitions, for example for the purpose of summing reconstructed holograms [39,40], the visibility of a hologram is an effective method for quality control. In most experimental holograms the interference fringes do not fill the entire camera, leaving a dark zone of lower counts and very low calculated visibility (typically  $\sim 0.02$ ). The interference and dark regions may be easily identified via segmentation by thresholding a histogram of the hologram. The visibility histogram of object holograms may also be used for segmentation of specimens that consist of distinct compositions, for example locating a metal nanoparticle on a carbon substrate. We start with the visibility map of Ni-NiO core-shell nanoparticles, as shown in Figure 2.6.a. The metallic Nickel cores (dark blue intensity) exhibit large loss of visibility due to strong incoherent scattering. The outer Nickel Oxide and Carbon contamination layers (cyan and green intensity) reduce the measured visibility (i.e. coherence) of the electron wavefront less than the Ni core. The vacuum around the particles (yellow intensity) has nearly constant visibility, within the limits determined by shot noise. The mottling seen in the vacuum visibility at the corners are Fourier ringing artifacts.

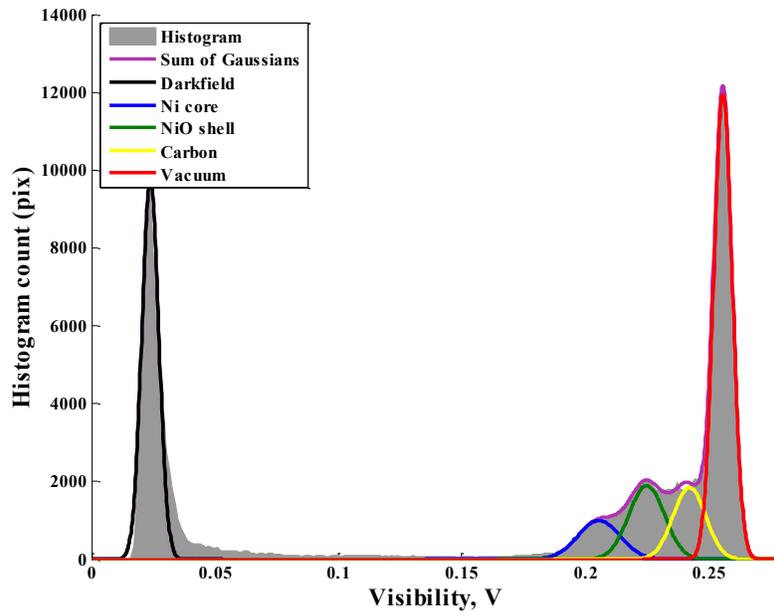
Segmentation is accomplished by fitting a sum of Gaussian functions to the visibility histogram, as shown in Figure 2.6.b. One Gaussian function is used each for the vacuum and dark portions of the image, and one for each expected region of constant composition and thickness. This method does not account for variation in mass thickness, but still achieves reasonable segmentation of the regions into the metallic Ni cores (blue line), the oxide shell (green line), and the oxide and carbon contamination layer (yellow line). Vacuum is indicated by the major peak (red line) while the region outside of the holographic fringes has nearly zero visibility (black line). Segmentation is accomplished by thresholding at the intersections of the Gaussian functions. In order to smooth the interfaces, we apply a standard image opening (dilation followed by erosion) to eliminate small features and variations in the segmentation maps. For the segmentation of the histogram to be successful, the overlap of the peaks must be reasonably low such that separation by least square fit

is manageable. Segmentation is accomplished more easily with the visibility than for the amplitude because the darkfield region is well-separated in the histogram.

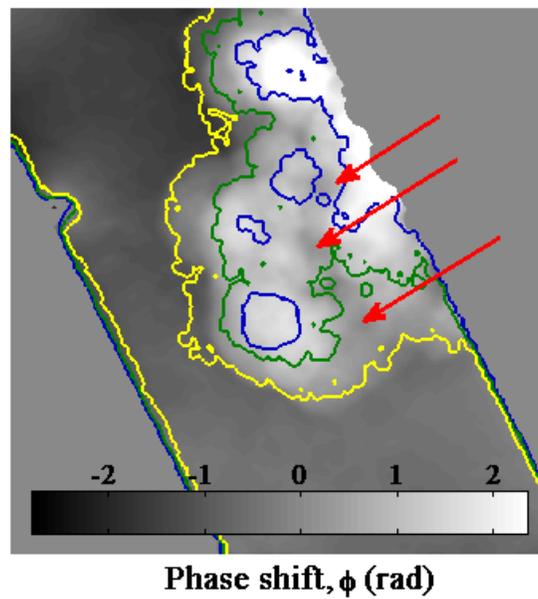
The resulting segmentation contours, shown in Figure 2.6.c with the unwrapped phase intensity value, does not separate mass thickness and composition. To separate thickness and composition would require additional information, generally in the form of energy filtered images or *a priori* information such as the characteristic scattering length of the material [41]. The segmentation map can then be applied to calculate average values for the amplitude, visibility, and phase shift for each region, as detailed in Table 2.1. Some anomalous phase holes are observed where the visibility (and amplitude) indicates no local variation in thickness or composition, as indicated by red arrows. The estimated deviation is approximately -1 radian, a possible indication of local electric (i.e. electrostatic charging) or magnetic potential centers.



(a)



(b)



(c)

Figure 2.6: (a) A holographic visibility map of core-shell, Ni-NiO nanoparticles. (b) A histogram of the visibility map in (a). The histogram may be segmented by a least squares fit of a series of Gaussian functions. (c) Segmentation applied to the unwrapped phase shift of the hologram. The area outside of the holographic fringes (top right and bottom left) has been masked.

Table 2.1: Results for the mean phase shift within the contours of Figure 2.6(c) and the associated mean amplitude and visibility.

<b>Region</b>	<b>Mean Phase Shift –relative to vacuum (rad)</b>	<b>Mean Amplitude (counts)</b>	<b>Scattering Factor (<math>A/A_{\text{vacuum}}</math>)</b>	<b>Mean Visibility (V)</b>	<b>Specimen MTF (<math>V/V_{\text{vacuum}}</math>)</b>
<b>Metallic Ni (blue contours)</b>	1.45	2470	0.879	0.208	0.812
<b>NiO (green contours)</b>	0.87	2600	0.925	0.226	0.883
<b>NiO and carbon contamination (yellow contours)</b>	0.22	2750	0.978	0.243	0.948
<b>Vacuum</b>	-	2810	-	0.256	-

## 2.6 CONCLUSIONS

Holographic visibility is a measure of the ensemble coherence of electrons. Partial decoherence of the electrons is affected by the components of the TEM, including the specimen. Thus the visibility is of potential interest for characterization of materials. The original metric for fringe visibility defined by Michelson over a century ago is not suitable for quantitative analysis as it lacks a defined variance. The Michelson metric is a 100 % confidence interval that overestimates the visibility in the presence of noise. Its lack of precision also makes it impractical for evaluating the visibility over a small area, which is necessary if generating a 2-D localized visibility map is desired.

We propose a new statistical metric based on the mean and variance that is statistically quantifiable, deterministic, and reproducible. We also developed an algorithm based on the Fourier ratio of the sideband to the centerband (or autocorrelation), based on the statistical metric, that calculates the local visibility in 2-D, thus making the visibility a third signal in electron holography. The application of a real-space Hamming apodization to the hologram before application of the Fourier transform is critical to suppress Fourier artifacts and accurately measure the visibility.

The visibility may also be used as a component that is independent of the amplitude and phase shift, for the purpose of image analysis. The visibility finds application in hologram segmentation and separating the interference region from the non-interference region in off-axis holograms. Initial experiments indicate that the visibility is more sensitive to changes in thickness or composition than the amplitude signal (Table 1). The visibility varies with specimen mass thickness and composition, similar to amplitude.

The main source of experimental error in evaluating the visibility comes from shot noise when the dose is  $< 200$  counts/pix. The detector DQE renders the noise spectrum non-uniform, so empirical corrections are the most effective means to remove the influence of shot noise. It is possible to remove the influence of shot noise from visibility measurement even at extremely low counts ( $< 20$ ) for both the statistical and Fourier-ratio methods.

## 2.7 REFERENCES

- [1] L. Reimer, H. Kohl, *Transmission Electron Microscopy*, 5th ed., Springer Science, New York, 2008.
- [2] E. Voelkl, L.F. Allard, D.C. Joy, *Introduction to Electron Holography*, in, Kluwer Academic, New York, 1999.
- [3] D. Geiger, H. Lichte, M. Linck, M. Lehmann, *Electron Holography with a Cs-Corrected Transmission Electron Microscope*, *Microscopy and Microanalysis*, 14 (2008) 68-81.
- [4] P.L. Potapov, H. Lichte, J. Verbeeck, D. van Dyck, *Experiments on inelastic electron holography*, *Ultramicroscopy*, 106 (2006) 1012-1018.
- [5] P.L. Potapov, J. Verbeeck, P. Schattschneider, H. Lichte, D. van Dyck, *Inelastic electron holography as a variant of the Feynman thought experiment*, *Ultramicroscopy*, 107 (2007) 559-567.
- [6] F. Roder, H. Lichte, *Inelastic electron holography - first results with surface plasmons*, *Eur. Phys. J.-Appl. Phys*, 54 (2011).
- [7] K. Harada, A. Tonomura, Y. Togawa, T. Akashi, T. Matsuda, *Double-biprism electron interferometry*, *Applied Physics Letters*, 84 (2004) 3229-3231.
- [8] G. Mollenstedt, *The History of the Electron Biprism*, in: E. Voelkl, L.F. Allard, D.C. Joy (Eds.) *Introduction to Electron Holography*, Kluwer Academic, New York, 1999.
- [9] Y. Aharonov, D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, *Physical Review*, 115 (1959) 485-491.
- [10] T. Boyer, *Semiclassical Explanation of the Matteuccia-Pozzi and Aharonov-Bohm Phase Shifts*, *Foundations of Physics*, 32 (2002) 41-49.
- [11] A. Caprez, B. Barwick, H. Batelaan, *Macroscopic test of the Aharonov-Bohm effect*, *Phys. Rev. Lett.*, 99 (2007).
- [12] H. Lichte, K.-H. Herrmann, F. Lenz, *Electron noise in off-axis image plane holography*, *Optik*, 77 (1987) 135-140.
- [13] F. Lenz, *Statistics of phase and contrast determination in electron holograms*, *Optik*, 79 (1988) 13-14.
- [14] A. Harscher, H. Lichte, *Experimental study of amplitude and phase detection limits in electron holography*, *Ultramicroscopy*, 64 (1996) 57-66.
- [15] A.A. Michelson, *I. On the application of interference methods to astronomical measurements*, *Philosophical Magazine Series 5*, 30 (1890) 1-21.
- [16] F. Zernike, *The concept of degree of coherence and its application to optical problems*, *Physica V*, 8 (1938) 785-795.
- [17] A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, H. Ezawa, *Demonstration of single-electron buildup of an interference pattern*, *American Journal of Physics*, 57 (1989) 117-120.
- [18] J. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, John Wiley & Sons, New York, 1978.

- [19] F. Zernike, Phase Contrast, a New Method for the Microscopic Observation of Transparent Objects, *Physica IX*, 7 (1942) 686-698.
- [20] J. Carrasquilla-Alvarez, R. Castaneda, J. Garcia-Sucerquia, M.A. Schofield, M. Beleggia, Y. Zhu, G. Matteucci, Retrieving the complex degree of spatial coherence of electron beams, *Optik - International Journal for Light and Electron Optics*, 119 (2008) 127-133.
- [21] C. Maunders, C. Dwyer, P.C. Tiemeijer, J. Etheridge, Practical methods for the measurement of spatial coherence: A comparative study, *Ultramicroscopy*, 111 (2011) 1437-1446.
- [22] J. Frank, The Envelope of Electron Microscopic Transfer Functions for Partially Coherent Illumination, *Optik*, 38 (1973) 519-536.
- [23] F. Hosokawa, M. Suzuki, K. Ibe, Determination of the effective source from its image in the backfocal plane of the objective lens, *Ultramicroscopy*, 36 (1991) 367-373.
- [24] J.-J. Fernandez, J. Sanjurjo, J.-M. Carazo, A spectral estimation approach to contrast transfer function detection in electron microscopy, *Ultramicroscopy*, 68 (1997) 267-295.
- [25] O. Hadar, I. Dror, N.S. Kopeika, Image resolution limits resulting from mechanical vibrations. Part IV: real-time numerical calculation of optical transfer functions and experimental verification, *Optical Engineering*, 33 (1994) 566-578.
- [26] R.R. Meyer, A.I. Kirkland, Characterisation of the signal and noise transfer of CCD cameras for electron detection, *Microscopy Research and Technique*, 49 (2000) 269-280.
- [27] M. Vulovic, B. Rieger, L.J. van Vliet, A.J. Koster, R.B.G. Ravelli, A toolkit for the characterization of CCD cameras for transmission electron microscopy, *Acta Crystallographica Section D*, 66 (2010) 97-109.
- [28] R.A. McLeod, M. Malac, Characterization of Detector Modulation-transfer Function with Noise, Edge, and Holographic Methods, Submitted, (2012).
- [29] D.L. Sackett, Bias in analytic research, *Journal of Chronic Diseases*, 32 (1979) 51-63.
- [30] M. Shao, M.M. Colavita, B.E. Hines, D.H. Staelin, D.J. Hutter, K.J. Johnston, D. Mozurkewich, R.S. Simon, J.L. Hershey, J.A. Hughes, G.H. Kaplan, The Mark-II Stellar Interferometer, *Astron. Astrophys.*, 193 (1988) 357-371.
- [31] J.A. Benson, D. Mozurkewich, S.M. Jefferies, Active optical fringe tracking at the NPOI, in: *Astronomical Interferometry*, SPIE, Kona, HI, USA, 1998, pp. 493-496.
- [32] M.M. Colavita, Fringe visibility estimators for the palomar testbed interferometer, *Publ. Astron. Soc. Pac.*, 111 (1999) 111-117.
- [33] M.M. Colavita, A.J. Booth, J.I. Garcia-Gathright, G. Vasisht, R.L. Johnson, K.R. Summers, Fringe Measurement and Control for the Keck Interferometer, *Publ. Astron. Soc. Pac.*, 122 (2010) 795-807.
- [34] A. Stuart, K.J. Ord, *Kendall's Advanced Theory of Statistics*, 6th ed., Hodder Education, London, 1994.
- [35] R.A. McLeod, M. Malac, Evaluating Visibility and Spatial Resolution in Electron Holography, *Microscopy and Microanalysis*, 14 (2008) 854-855.
- [36] E. Voelkl, L.F. Allard, A. Datye, B. Frost, Advanced electron holography: a new algorithm for image processing and a standardized quality test for the FEG electron microscope, *Ultramicroscopy*, 58 (1995) 97-103.

- [37] P.A. Midgley, An introduction to off-axis electron holography, *Micron*, 32 (2001) 167-184.
- [38] R.A. McLeod, DM3 Import for Gatan Digital Micrograph, in, <http://www.mathworks.com/matlabcentral/fileexchange/29351>, 2010.
- [39] E. Voelkl, D. Tang, Approaching routine  $\pi/1000$  phase resolution for off-axis type holography, *Ultramicroscopy*, 110 (2010) 447-459.
- [40] R.A. McLeod, M. Bergen, M. Malac, Technique for Complex Averaging of Electron Holograms, *Microscopy and Microanalysis*, 17 (2011) 918-919.
- [41] M.R. McCartney, M. Gajdardziska-Josifovska, Absolute measurement of normalized thickness,  $t/\lambda_i$ , from off-axis electron holography, *Ultramicroscopy*, 53 (1994) 283-289.

### 3 CHARACTERIZATION OF DETECTOR MODULATION-TRANSFER FUNCTION WITH NOISE, EDGE, AND HOLOGRAPHIC METHODS

The contents Chapter 3 were accepted for publication in the journal *Ultramicroscopy* by R.A. McLeod and M. Malac on February 22<sup>nd</sup>, 2013. Contributing author M. Malac edited the paper and provided valuable discussion throughout the development process.

#### *ABSTRACT*

We developed a new method for characterization of detector performance used in the transmission electron microscope (TEM) based on the measured contrast of holographic fringes. The new method changes spatial frequency of the measured holographic fringes, generated by an electrostatic biprism and Schottky or cold field-emission gun, to sample the modulation-transfer function (MTF) of the detector. The MTF of a Gatan Ultrascan™ 1000 charged-coupled detector (CCD) is evaluated using the new method and the results are compared to the established noise and slanted-edge method results. Requirements for accuracy of the edge and noise MTF methods are discussed. We consider issues surrounding incomplete read-out and how it affects the gain reference normalization of the detector. We evaluate how the MTF affects optimization of experimental parameters in the TEM.

#### 3.1 INTRODUCTION

We present a new method that employs holographic fringes to accurately and reliably characterize the *modulation-transfer function* (MTF) of *charge-coupled detectors* (CCDs) used in the *transmission electron microscope* (TEM). Off-axis electron holography is typically used to investigate electric and magnetic properties of specimens in TEM [1]. The method uses a deterministic algorithm to evaluate the contrast, or visibility, of the holographic fringes. The measured contrast of holographic fringes at constant illumination conditions but varied magnification samples the detector response as a function of spatial frequency, i.e. the MTF. The holographic method complements the existing noise and edge MTF methods. Although the noise and edge methods are reproducible, i.e. precise, they are not necessarily accurate measure of the true detector MTF. The holographic method does not introduce obvious biases that distort the measured MTF.

Two-dimensional (2-D) image recording in transmission electron microscopy (TEM) traditionally used film. In the 1990s there was a shift towards the use of CCDs to record electron intensity [2-4]. Slow-scan CCDs suffer radiation damage if exposed to electrons at typical TEM energies [5]. A scintillator, composed of a doped single-crystal YAG or a powder phosphor, acts as a radiation shield and generates optical photons from the incident high-energy electrons by photoluminescence. Photons are coupled from the scintillator with a bundled fiber-optic waveguide, or an optical lens, to the CCD which detects the photons.

The objective in characterization of CCDs is to find the signal transfer from the distribution of electrons impinging on the scintillator to the distribution of digital counts in the electronic read-outs. In TEM, each primary electron that impacts the scintillator can be regarded as a signal impulse. CCD/scintillator (and film) broadens the signal impulse because high-energy electrons have a stopping range on the order of 100  $\mu\text{m}$  in scintillator materials which is an order of magnitude higher than typical pixel dimensions (5-15  $\mu\text{m}$ ) [6]. Consequently, the detector heavily oversamples the image. The stochastic average energy deposition as a function of position for a single incident electron is known as the impulse response or the *point-spread function* (PSF). In TEM, it is usually preferred to characterize the detector response in terms of spatial frequency,  $q = (u, v)$ , rather than spatial position,  $r = (x, y)$ . The Fourier transform of the PSF is the MTF, and gives the image contrast as a function of spatial frequency [7,8].

The standard approach to measure the MTF is the slanted-edge method, discussed in section 3.2.2. In addition to the signal-transfer, we also characterize the noise-transfer, which is characterized by the *noise-power spectrum* (NPS) or similarly the noise-based MTF, known as the *noise-transfer function* (NTF), described in section 3.2.1. Both the MTF and NPS are used to calculate the *detector quantum efficiency* (DQE), which relates the signal input to output ratio of the detector as a function of spatial frequency and dose. We devised a new method to measure the MTF that uses off-axis holographic fringes [9], as discussed in section 3.2.3. We discuss the experimental results from a Gatan Ultrascan™ 1000 CCD detector system for all three methods in section 3.3 and practical guidelines for microscope conditions in section 3.3.3.

### 3.1.1 OPTICAL TRANSFER FUNCTION AND DETECTOR QUANTUM EFFICIENCY

The performance of an (electro-)optical system is described using two-dimensional linear system theory [10]. Transfer functions known as the *optical-transfer function* (OTF) are used to describe the relationship of the complex input to output signal, with the form,  $\text{OTF}(q) = \text{MTF}(q)\exp(-i \cdot \text{PTF}(q))$  where PTF is the *phase-transfer function*, and reflects the phase shift of the wavefront. The OTF is presented in terms of spatial frequency so the overall OTF of the microscope and detector is the multiplicative cascade of the individual subsystem OTFs. The entire instrument OTF cascade includes the coherence of the source, the electron-optics (referred to as the

*contrast-transfer function* (CTF)), the specimen itself, and the detector. The spectrum of spatial frequencies that falls on the detector is magnification dependent and dampened by the microscope source and electro-optical OTFs. The preceding OTFs in the cascade affect how the detector fits into the overall optimization of the microscope. For example, if the CTF limits the incident spatial frequencies to low frequencies, then the high-frequency damping of the detector is not observed. A different example is parallel-beam diffraction where the high-frequency damping of the detector broadens the sharp diffraction peaks.

The detector system consists of a series of amplification and scattering subsystems [11,12], shown in Figure 3.1, that are represented by a single MTF. The scintillator integrates incident electron intensity and hence is incoherent and is represented by an MTF and not an OTF. The detector gain amplifies the signal but also injects additional noise into the spectrum which renders the MTF and NTF non-equivalent [13-15]. The addition of noise from the amplification steps is characterized by the detector quantum efficiency (DQE) [11,16-18],

$$DQE(q, G, N) = \frac{SNR_{out}^2}{SNR_{in}^2} = N \frac{MTF^2(q)}{NPS(q, N)} \quad (3.1)$$

where  $N$  is the dose in counts per pixel, and  $NPS$  is the output *noise power spectrum*. Note that we do not include the factor of gain-squared in front of the  $MTF^2$  found in some references as we present the MTF as normalized to unity at zero-frequency.

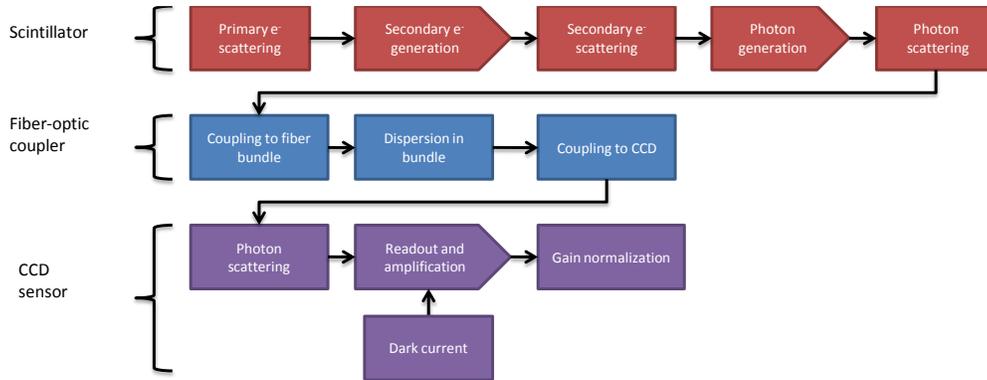


Figure 3.1: A flow chart of the scattering (square boxes) and amplification (arrow boxes) processes in a typical TEM detector stack. In practice the individual components cannot be separated, with the exception of dark current and gain normalization, so the entire system is modeled by a single MTF.

### 3.1.2 ANALYTICAL FORM OF THE MTF AND NORMALIZATION

A primary electron interacts with the scintillator through a combination of Gaussian and Lorentzian broadening processes. The general form of the convolution of Lorentzian and Gaussian processes is known as a *Voight profile* [19]. For example, photon-scattering is a result of random thermal motion of atoms and is Gaussian whereas the double differential scattering cross-section of core electrons has a

Lorentzian distribution [6,20,21]. Note this implies that the PSF has the form of a Gaussian and Laplacian product.

The Voigt profile is challenging to fit by least squares, but it can be approximated by a sum of Lorentzian and Gaussian profiles [22]. The form of the least squares function fit to the rotational average of the MTF is

$$\text{MTF}(q) = \sum_i a_i \exp\left(-\left(\frac{q-b}{c_i}\right)^2\right) + \sum_j \frac{a_j}{1 + \left(\frac{q-b}{c_j}\right)^2} \quad (3.2)$$

where  $a_i$ ,  $b_i$ , and  $c_i$  are best-fit parameters and  $q$  is pixel frequency. For most cases a single Gaussian and double Lorentzian sum obtains a satisfactory correlation to the data with no discernible residual pattern, as is the case for all results presented in this paper.

A key challenge to accurately determine the MTF is normalization. The true zero-frequency response of the detector cannot be measured and an error in the normalization will bias the result for all spatial frequencies. When the MTF is measured near zero frequency it may be erroneous due to poor signal-to-noise ratio or low-frequency artifacts, as is the case for both the noise and slanted-edge methods. We argue that the MTF should have mirror symmetry about  $q = 0$  as a consequence of the symmetry of the Fourier transform. If the MTF slope is non-zero at  $q = 0$ , the derivative would have discontinuity not readily explained by the physics of electron scattering in the scintillator or photon propagation through the detector stack. A homogenous material will result in a radially symmetric scattering profile but NTF measurements show an anisotropic artifact peak that appears along read-out axis of the detector and is absent in the orthogonal direction, as discussed in section 3.2.1.2. Low frequency peaks are not found in the MTF of theoretical models of CCD sensors [23,24]. Therefore we conclude that the Voigt-style least-squares best-fit is the most practical solution for determination of the zero-frequency response and hence normalization of the MTF for all methods.

### 3.1.3 IMPACT OF THE DETECTOR ON MICROSCOPE PERFORMANCE

The detector directly impacts the overall TEM instrument performance in three ways: 1.) signal and shot noise are both dampened by the detector, but noise is injected at high frequencies, which worsens the signal-to-noise ratio (DQE); 2.) the information limit of the instrument is reduced by the detector MTF through its contribution to the instrument OTF cascade that gives the ultimate information resolution. However, the detector MTF has the distinction that its effect varies with TEM magnification. Hence, there is an optimization between the field of view and the information limit discussed further in Section 3.3.3, and; 3.) the decrease in contrast due to the detector MTF requires increased electron irradiation dose to achieve a desired signal level. For radiation sensitive specimens this causes difficulties [25,26]. Even for radiation-stable specimens, the increased dose

requires longer exposures and introduces additional motion blur from drift. Accurate characterization of detector performance is important for optimization of the overall performance of the transmission electron microscope and provides directions in which future high-energy electron detection technology may be improved.

### 3.2 METHODS FOR CHARACTERIZATION OF MODULATION-TRANSFER FUNCTION

For microscope optimization, image processing, and comparison of simulation to data, the MTF of a detector must be accurately measured. There are two well-established methods used in transmission electron microscopy to determine the MTF: the noise method, which determines the NTF, and the slanted-edge method, which is historically thought to represent the true detector MTF [13,14,27]. In addition there also exists methods that use amorphous thin films [28], and subpixel sampling of a finely focused diffraction-mode spot [29,30]. In this section we review the noise and edge methods and introduce a new method where the visibility, or contrast, of off-axis holographic fringes at variable magnification is used to measure the MTF.

#### 3.2.1 THE NOISE TRANSFER FUNCTION

For the noise method, the MTF is derived from the input and output *power spectral density* (PSD) obtained by uniformly illuminating the detector and acquiring noise images. In the electron microscopy and optics communities the noise-transfer is referred to as the noise-based MTF or NTF, a convention we will follow here [13,27,31,32],

$$PSD_{out}(\mathbf{q}) = NTF^2(\mathbf{q})PSD_{in}(\mathbf{q}) = \langle |\mathcal{F}(I(\mathbf{r}) - \bar{I})|^2 \rangle PSD_{in}(\mathbf{q}) \quad (3.3)$$

In the astronomy and medical imaging communities the noise transfer of the detector is more commonly known as the *noise power spectrum* (NPS) [33,34] which is defined as,

$$NPS(\mathbf{q}) = \langle |\mathcal{F}(I(\mathbf{r}) - \bar{I})|^2 \rangle$$

The actual methods used to calculate the NPS vary from that of the NTF and often use a moving synthetic aperture or other sampling method [34], which produce marginally different results. Despite the different methods used for calculation, analytically the  $NPS = NTF^2$ .

There are a number of artifacts which can potentially make the NTF different from the signal MTF including but not limited to: the input power spectral density of the illumination, aliasing of high-frequency noise above the Nyquist limit into lower frequencies, electronics noise injected by analog-to-digital conversion and amplification of the signal, incomplete read-out amongst adjacent rows, afterglow of the scintillator, and convolved artifacts and correlated noise introduced by the use of reference images.

Modern large format CCD detectors are segmented into quadrants, each with its own readout electronics such that a noise image has discontinuities at the quadrant borders. In order to calculate the NTF correctly, each quadrant's NTF must be calculated separately and then the quadrants may be averaged to estimate the detector-wide NTF. Failure to separate the quadrants introduces box-car functions into the NTF, producing a strong low-frequency peak near  $0.0 \text{ pix}^{-1}$ .

### **3.2.1.1 Aliasing of noise above Nyquist limit in noise Transfer Function**

A noise image will contain spatial frequencies above the Nyquist limit which is imposed by the finite number of detector pixels [35]. As the 2-D FFT used to calculate the NTF conserves the total electron counts, the counts associated with these frequencies will then alias back into lower frequencies with the relation,  $q_{\text{nyquist}} + \Delta q \rightarrow q_{\text{nyquist}} - \Delta q$ .

This implies that high frequencies (near to the Nyquist limit) will have a larger proportion of extraneous counts from aliasing than near zero-frequency. As such, the NTF typically overestimates high-frequency detector performance. The effect of aliasing may be estimated by fitting a twin Voight curve, one centered at zero-frequency and the other at twice the Nyquist limit (i.e.  $1.0 \text{ pix}^{-1}$ ), and subtracting the post-Nyquist peak from the results. This is a first-order correction that does not remove the next-nearest-neighbor peak, but is sufficient for most purposes.

### **3.2.1.2 Detector Read-Out Errors**

Read-out errors are artifacts associated with the camera either while the shutter is in the process of closing or when the charge on the CCD detector is read-out. These artifacts are notable because they are anisotropic, present along the axis the CCD detector is read-out (vertical, Y-axis) and absent in the orthogonal direction (horizontal, X-axis). There are three potential sources of read-out error: 1.) incomplete read-out, caused by the CCD wells being read-out too quickly; 2.) afterglow, caused by long-time-constant-decay states in the scintillator, and; 3.) motion-blur caused by slow shutter close times. We only see evidence of incomplete read-out, as explained below. These artifacts are different from the residual images (ghosting) sometimes seen in the dark reference after a high-dose exposure, which is thought to be due to long-lifetime defect states in the CCD detector [36].

The typical algorithm to calculate the NTF can be modified to investigate read-out artifacts. 1-D Fourier transforms of each individual row or column of noise images are applied and summed accordingly. The resulting singleton dimensional noise transfer function is not comparable to the 2-D NTF as it is heavily affected by the finite size of the FFT, but it allows for convenient comparison between the read-out and orthogonal axes. Our results for the NTF and slanted-edge MTF both exhibit a low frequency artifact. We argue in Section 3.1.2 that this was likely non-physical and should not influence MTF normalization. Here we show conclusively in Figure

3.2 that the low-frequency artifact appears exclusively in the read-out axis,  $NTF_y$ , and is absent in the orthogonal direction,  $NTF_x$ . The effect worsens with total dose (percent saturation) on the detector, as expected for incomplete read-out. Afterglow is discounted as introducing a one-second delay between shutter closure and CCD read-out has no effect. Shutter motion blur is also disregarded as the artifact worsens with increased exposure time/dose. Shutter artifacts are eliminated by installing an ultra-fast mechanical shutter driven by a stepper motor that operates on a millisecond time scale, used for all our measurements [37], or a fast electrostatic shutter [38].

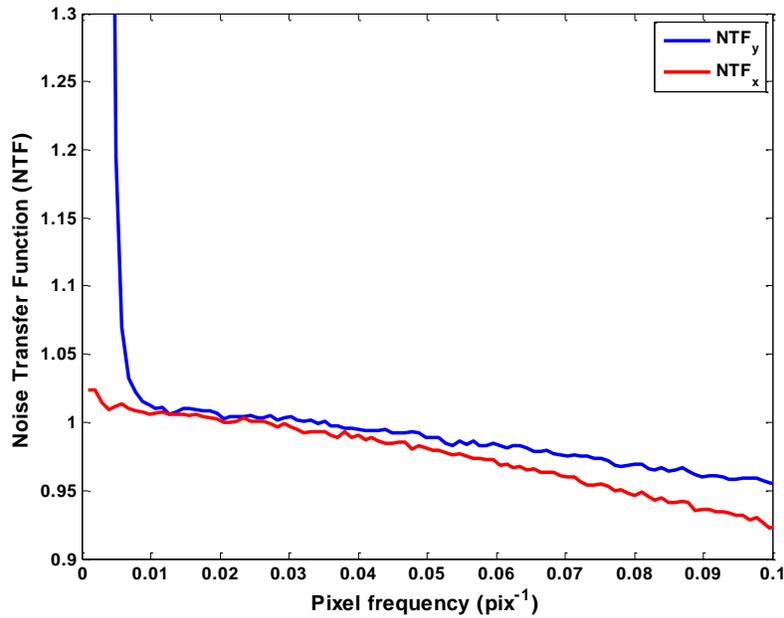


Figure 3.2: The 1-D noise transfer function at 300 keV in the read-out (Y, blue solid) axis,  $NTF_y$ , and orthogonal (X, red dashed) axis,  $NTF_x$  for a mean dose of 2500 *counts/pixel*. The strong low-frequency read-out error peak has a maximum value of approximately 13.0 at the lowest measured frequency and extends out to 0.01  $\text{pix}^{-1}$ , compared to a maximum of only 1.03 for  $NTF_x$ .

Incomplete read-out is well-known issue with CCD detectors [39, 40]. The charge on a CCD is read-out along its edge, and each subsequent row is shifted towards to the edge. If the CCD is read-out too quickly, residual charge may remain in the well and contribute to subsequent rows. The read-out period of our GATAN Ultrascan™ is approximately 1 s, whereas a more typical slow-scan CCD read-out period for noise sensitive applications is 20-45 s [41]. The slower the read-out time of a CCD device, the less noise is injected by the analog-to-digital convertor and the better the DQE. With regards to normalization of the MTF curve, it is clear that read-out error represents amplification or transfer of low-frequency noise and not dampening of the observed signal. Experiments show that the read-out error increases substantially as the detector approaches saturation, but measurements of the contrast of off-axis holograms and lattice fringes show only trivial saturation

dependence. Consequently, the slanted-edge MTF or NTF cannot be normalized by the lowest-observed frequency raw data, and normalization by a best-fit Voigt profile is required.

### 3.2.1.3 Power Spectral Density and reference normalization

Both gain references and NTF are acquired under parallel, uniform illumination with the assumption that the input noise is uniform across all frequencies, i.e. white, which may not be the case [31]. By the Wiener-Khinchin theorem the input power spectral density ( $PSD_{in}$ ) of Eqn. 3.3 is given by the autocorrelation function of the limiting aperture [42], which for a circular aperture is similar to a triangle function (also known as a Bartlett window) and has the analytic form,

$$PSD_{in}(q) = \frac{2}{\pi} a^2 \cos^{-1} \left( \frac{q}{2a} \right) - \frac{q}{2\pi} \sqrt{4a^2 - q^2}$$

where  $q$  is spatial frequency and  $a$  is the radius of the aperture. Given the small numerical aperture values used in TEM, assuming the input noise spectrum is white is not valid and can bias results for the NTF. For a magnified aperture diameter of  $\times 8$  the detector width,  $PSD_{in}(q = 0.5 \text{ pix}^{-1}) = 0.92$ , and for  $\times 16$ ,  $PSD_{in}(q = 0.5 \text{ pix}^{-1}) = 0.96$ . Also, the illumination should be incoherent, such that diffraction (Fresnel fringes) does not influence the result. In practice, the  $PSD_{in}$  has a limited impact on the measured NTF. This is because the  $PSD_{in}$  is a function of the angular-distribution of electrons onto the detector, and the thick scintillator found on fiber-optic coupled CCD systems acts as an angle-diffusing, intensity-integrating device similar to an integrating sphere in photon optics. The  $PSD_{in}$  may be significant for thinner scintillator devices that are not as heavily oversampled, such as lens-coupled and direct-exposure detector systems.

Even the most basic image processing makes use of gain and dark reference images to remove artifacts such as the periodicity introduced by the fiber-optic coupling, non-uniform thickness of the scintillator, and bad pixels on the CCD detector. The normalization has the form,  $I_{proc}(\mathbf{r}) = (I_{raw}(\mathbf{r}) - I_{dark}(\mathbf{r})) / I_{gain}(\mathbf{r})$ . The use of references and the noise in the reference images can bias the resulting images, and it is important to be aware of how those biases may lead to erroneous results [43]. Furthermore, the gain reference alters the apparent  $PSD_{in}$  as a function of percent saturation and illumination angle.

Normally a single frame with the shutter closed is used as a dark reference. This is a sub-optimal approach, as the standard deviation of a dark reference is higher than the mean, measured at  $0.6 \pm 4 \text{ counts}$  at  $-25.4 \text{ }^\circ\text{C}$  for our Gatan US 1000 detector. The dark current varies exponentially with temperature  $T$ , in Celsius,  $\mu_{dark} = 4.24 \cdot \exp(0.106T)$  and  $\sigma_{dark} = 17.25 \cdot \exp(0.077T) + 0.288$ . The dark reference does vary with exposure time, as thermally-generated electron-hole pairs are generated, but is insignificant at exposure times less than a minute and concentrated at the

corners of the detector for our Gatan US 1000. The long-period temperature stability of detectors from other manufacturers is expected to vary.

When the same dark reference is subtracted from every image in a series, it introduces correlated noise which can affect frame-averaging techniques where the median counts is near the dark counts, such as core-loss EELS [43]. We have taken the approach that the dark reference remains constant as long as the CCD remains at constant temperature, and average  $> 10000$  dark references to remove the systemic dark current. Comparison of two 17000-frame dark references, acquired sixteen months apart, shows that 4084 (0.1 %) pixels differ by more than one count, illustrating excellent long-term temperature stability. In comparison, in single frame dark references  $\sim 70$  % of pixels will vary by more than one count due to non-systemic, thermal dark noise. Our Hitachi HF-3300 is baked weekly to maintain high vacuum ( $1.8 \cdot 10^{-8}$  Torr), during which the Pelletier coolers on the detectors are switched off. Hence an overnight acquisition of a dark reference once per year is generally sufficient for the Gatan US 1000. Further discussion and acquisition script are found in appendix A1.

Averaging many dark frames encounters problems with so-called *zingers*, typically cosmic ray events that saturate pixels [44]. Zingers accumulate relatively quickly because saturated pixel in a single frame (65000 counts) is equivalent to  $\sim 12000$  dark reference frames. A solution is to apply a median-threshold filter to dark references so that only pixels above a threshold of 50 counts are modified. Pixels that exceed a 10-sigma filter are median filtered which effectively strips out zingers from a long-run dark reference.

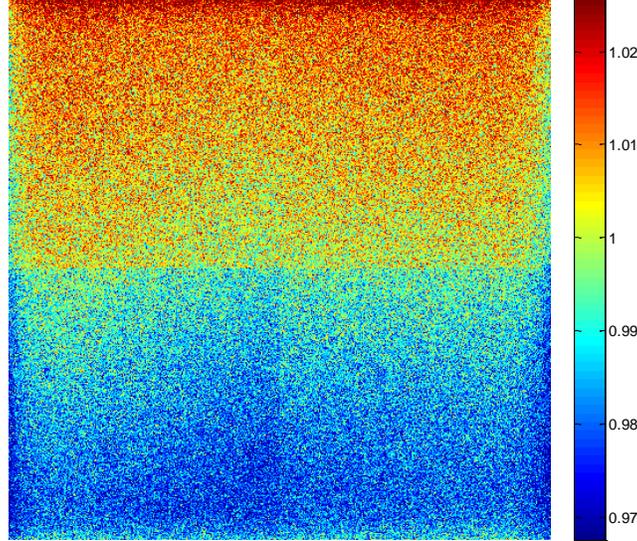
The gain normalization successfully removes gain-related artifacts because they have fixed positions on the camera, i.e. they have fixed phase in frequency space. Normalization of the gain reference itself is challenging to do accurately and avoid bias in the output dose. We take a histogram of the intensity values in the gain reference, fit a sum of 3 to 4 Gaussian functions, and normalize by the centroid of the Gaussian sum. As the NTF depends on primary electron energy, separate gain references must be prepared for each accelerating voltage used.

The gain reference is typically recorded by summing many frames with uniform illumination to a significant fraction of saturation [45]. Gain references normalization is known to be not wholly satisfactory [46]. For quantitative purposes, the gain reference changes with dose per frame (i.e. saturation fraction) and illumination convergence angle. The application of a gain reference can potentially modify the spatial frequency distribution of an image, which is clear in the frequency-space representation,  $I_{raw} = \mathcal{F}^{-1}\{NTF^2 PSD_{raw} A_{grad} A_{foc} A_{r-o}(N)\}$ , where  $A_{grad}$  is the low-frequency Fourier peak from non-uniform thickness of the scintillator and/or non-uniform illumination,  $A_{foc}$  is periodic artifacts from the fiber-optic coupling, and  $A_{r-o}$  is the artifact due to incomplete read-out, which is a

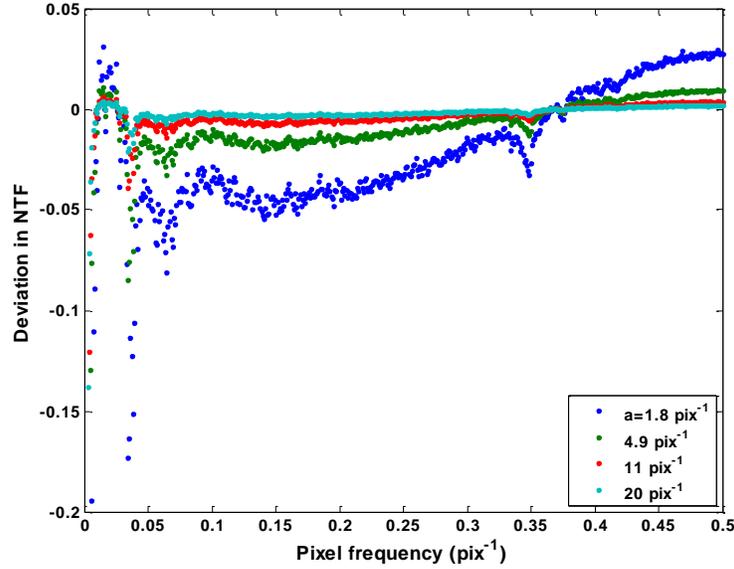
function of dose per frame. Then the Fourier representation of the gain-normalized image, if we omit the dark reference for simplicity, is,

$$\begin{aligned}
\mathcal{F} \left\{ I_{raw} \cdot \frac{1}{I_{gain}} \right\} &= \text{NTF}^2 \text{PSD}_{raw} A_{grad} A_{foc}(N_{raw}) A_{r-o}(N_{raw}) \\
&\otimes^{-1} \text{NTF}^2 \text{PSD}_{gain} A_{grad} A_{foc}(N_{raw}) A_{r-o}(N_{gain}) \\
&\cong \text{NTF}^2 \left( \text{PSD}_{raw} A_{r-o}(N_{raw}) \otimes^{-1} \text{PSD}_{gain} A_{r-o}(N_{gain}) \right)
\end{aligned}$$

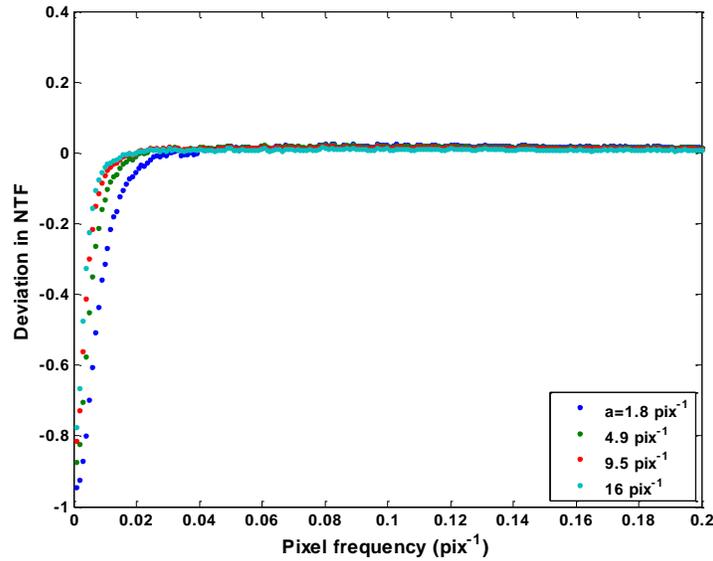
The effect is seen clearly from the ratio of a high-dose gain reference to a low-dose gain reference, which produces a ramp due to incomplete read-out, as shown in Figure 3.3.a. To understand the relation among NTF, saturation fraction, and illumination angle we acquired a series of noise images taken at variable illumination angle but constant dose obtained by adjustment of exposure time. We found that when normalized by a gain reference of higher dose, as shown in Figure 3.3.b, some of the periodic artifacts are not completely removed. For the converse case, as shown in Figure 3.3.c, the incomplete read-out, which is partially deconvolved, is more apparent than usual but there are no periodic artifacts injected into the image. The change in the periodic fiber-optic coupling artifacts with illumination angle may reflect a shift in the fiber cross-talk [47].



(a)



(b)



(c)

Figure 3.3: (a) The effect of incomplete read-out can be observed by taking the ratio of two gain references taken at different average electron dose. Here the deviation is approximately 5 % over the field of view for gain references at 10000 counts/pix and 500 counts/pix. (b) Consequences of using a high-dose, well-spread gain reference (10000 counts/pix and  $a = 16 \text{ pix}^{-1}$ ) to normalize a moderate-dose image (500 counts/pix) as shown by the ratio of the NTF under the two conditions. The fiber-optic coupling and incomplete read-out artifacts are not properly removed when the dose and illumination angle of the gain reference are significantly different from those used for acquisition. (c) The converse case of using a low-dose gain reference (2500 counts/pix and  $a = 20 \text{ pix}^{-1}$ ) to normalize a high-dose image (10000 counts/pix). This effect is anisotropic (not shown) and dominated by the read-out axis, so likely arises from the partial deconvolution of incomplete read-out.

The results indicate is that it is preferable to acquire a gain reference at the expected experimental dose with a well-spread beam. When that is not practical, we suggest gain references should be composed of low saturation fraction images, such that periodic artifacts, which can disrupt the success of cross-correlation alignment methods, are minimized. The potential error introduced by incorrect-dose gain normalization is small and should not affect qualitative imaging applications.

### 3.2.2 THE SLANTED-EDGE MODULATION TRANSFER FUNCTION

In this section we review the established method of MTF measurement that places a electron-opaque knife-edge over the detector to measure the broadening of the edge, known as the *edge-spread function* (ESF) [7,48]. An ideal step-function contains all spatial frequencies and as such can be used to sample the detector MTF. Unless the edge is moved across the detector, the slanted-edge MTF samples only a limited number of pixels in one direction. In the resulting 2-D image of the knife-edge, the position of the edge in each row or column is registered with subpixel accuracy. The edge profiles are then aligned and averaged, to find the ESF of the detector. Approximately fifty images are summed to reduce shot noise in the bright-field half of the detector and dark variance in the dark-field half.

A step function contains all spatial frequencies; therefore the spatial frequencies above the Nyquist limit ( $> 0.5 \text{ pix}^{-1}$ ) can alias into the lower frequencies due to the finite sampling of the CCD pixels. This is a serious problem in photon optics where detector systems are designed to slightly undersample each incident photon. The solution is to oversample the line profile two or four-fold and create a subpixel offset in each row by setting the edge at a slant angle ( $< 10^\circ$ ) [49]. This mitigates the insufficient sampling and anisotropic sampling of pixels when a straight edge is used. However, TEM detector systems are typically already heavily oversampled, such that the MTF is typically approximately 0.1 at the Nyquist frequency, so digital oversampling is not necessary.

The measured ESF is transformed to a *line-spread function* (LSF) via a numerical derivative. The LSF is in turn a slice through the two-dimensional PSF, perpendicular to the knife-edge. Numerical differentiation uses 2–5-points to estimate the derivative. The choice of numerical derivative (typically 2-point difference or 3-point derivative formulae) method introduces a low-pass filter of the form [50]:

$$MTF_{\text{derivative}}(q) = \text{sinc}(\pi/\delta (k - 1)q) \quad (3.4)$$

where  $\delta$  is the oversampling rate, and  $k$  is the number of points in the numerical derivative. The numerical derivative also worsens the signal-to-noise ratio of the data, countered by averaging numerous images of the edge.

The slanted edge method suffers from several sources of systemic error as the experimental knife-edge used is not an ideal sharp and smooth edge [51,52]. The potential sources of systemic error are: 1.) defocus between the edge-plane and the

detector-plane which results in blurring of the edge and potentially Fresnel fringes; 2.) surface roughness of the edge on the order of pixel lateral dimension (15  $\mu\text{m}$ ); 3.) translucency and diffraction at the edge; 4.) uncertainty in edge position and angle registration in the oversampling calculation, and; 5.) uneven illumination of the detector between bright and dark portions; the character of the bright-field noise is dominated by shot noise, and the dark-field noise by read-out noise. Each of these potential errors broadens the measured ESF, hence the slanted-edge MTF will trend towards being an underestimate of the true MTF in TEM.

The edge method is not affected by many of the difficulties encountered with the NTF. For example, using a gain reference compared to unprocessed images does not change the overall line-spread function but does make it less noisy. Similarly, alignment of the slanted-edge with the read-out or orthogonal axes of the detector does not change the shape of the low-frequency peak. This is likely because incomplete read-out should effectively act upon the edge as a sub-pixel shift, rather than blurring it.

### *3.2.3 THE HOLOGRAPHIC MODULATION TRANSFER FUNCTION*

Historically, one of the first techniques for evaluation of MTF in photon optics was the contrast (or visibility) of a target composed of sinusoidal fringes, with each fringe frequency sampling the MTF at one point [53]. In TEM, lattice fringes are a common source of sinusoidal fringes and unlike photon optics, the magnification of magnetic lenses in TEM are easily altered, allowing many spatial frequencies to be measured using a single sinusoidal target. Electromagnetic lenses rotate the image with increasing optical power, so the sinusoidal fringes produce a rotational average of the 2-D MTF. Measurement of MTF from sinusoidal fringes is a relative measurement. Therefore, most of the artifacts that affect the edge and noise methods discussed above have no impact.

In principle, lattice fringes from a crystalline sample can produce sinusoidal fringes but we have chosen to use an off-axis holography configuration with an electrostatic bi-prism to generate a holographic sinusoidal target [1]. This technique was previously used to support the NTF, but without any means of determining the zero-frequency visibility of the holograms so the method was only relative [45]. We show in section 3.2.3 that the holographic method with a sufficiently high number of sampled fringe frequencies reliably determines the MTF via a fit to a Voigt profile [9]. Normalization of the holographic-derived MTF is done via extrapolation, using the zero of the best-fit Voigt profile, as is the case with the NTF and slanted-edge MTF.

The advantages of using an electrostatic bi-prism compared to crystal lattice fringes are: 1.) the lower magnification of the biprism compared to the specimen on the detector plane, so the impact of mechanical drift and vibration of the biprism is reduced; 2.) there are no sample interactions that reduce the observed contrast,

such as inelastic scattering, variations in thickness, contamination layers, or generation of charge defects (known as the ‘bee-swarm’ effect [54]), and; 3.) there is only one set of parallel fringes, which provides the best possible signal-to-noise ratio and simplifies data processing. A similar condition can be obtained with lattice fringes in the two-beam condition.

The holographic method does have drawbacks compared to the edge method. First is the scale and complexity of the experiments required to collect the data. Large data sets are required to achieve consistent Voigt profile least-squares fits. With long experiments, instrument drift causes misalignments that gradually reduce the observed contrast. We address this issue in section 3.2.3. Another potential drawback is the absence of very low frequency data. However, the holographic MTF does provide data up to the limit where the NTF and edge MTF become contaminated by read-out error. The quality of the Voigt profile best-fit is very reliable; we have conducted holographic MTF experiments 4–6 times and found repeatable results. Last, the holographic MTF implicitly measures the product of the detector system MTF with the OTF of the instrument. If the OTF changed with magnification, the result could be biased. We discount this as microscope OTF is dominated by the objective lens, which is held constant. Projector lens distortion is significant only at very low excitation, which we avoid in our experimental configuration. Furthermore, distortion results in a phase shift of the fringes, and not a change in observed contrast [55].

### 3.3 EXPERIMENTAL METHODS AND RESULTS

#### 3.3.1 EXPERIMENTAL METHODS TO MEASURE MTF

The experiments were performed at 100, 200, and 300 keV on a Hitachi HF-3300 TEM equipped with a cold field electron gun (CFEG) and the double bi-prism off-axis electron holography configuration [56]. We present results for a GATAN Ultrascan™ 1000 detector system (2048 x 2048 pixels) but the methods presented here are general for other slow-scan CCD detectors or film. Automation of the acquisition process is required to reduce the total experiment time and limit the impact of any long-period instrument instabilities so the HF3300 was controlled via MATLAB scripts over TCP/IP. Similarly, a plug-in was written for Gatan Digital Micrograph to allow control of the CCD acquisition over TCP/IP<sup>1</sup>.

First a 120,000 frame dark reference was acquired as described in section 3.2.1.3. Then the instrument was placed in low-magnification mode, with the objective lens turned-off, and the illumination well-spread and uniform,  $a \approx 30 \text{ pix}^{-1}$ . A gain reference was acquired for each accelerating voltage, each composed of 500 frames with a mean electron dose of  $\sim 350$  counts/pixel. For the NTF, one-hundred noise images were acquired. After reference normalization the Fourier transform of each

---

<sup>1</sup> To be released by Hitachi as the MAESTRO package.

quadrant was applied at the magnitude summed. A 2-D rotational average was applied to each individual quadrant NTF, and the four NTFs were then averaged. Then dealiasing, as described in section 3.2.1.1, was applied to arrive at a detector NTF estimate.

For the edge MTF the same optical configuration is used as for the NTF, but a carefully polished copper plate was placed over the detector in the film chamber, at a slant angle measured as  $1.2^\circ$ , such that the edge was over the two right-most detector quadrants. Fifty images were taken at each accelerating voltage, reference normalized, and then summed to reduce shot and dark noise in the bright and dark fields respectively. Each row was normalized and fitted with a sum of error function and arctangent functions to locate the edge position that provides subpixel accuracy in the registration of each row without needing interpolation. The ensemble of edge positions was then fitted to find the slant-edge angle and position across the field of view. Each row was then bicubically sampled and shifted by the slant-edge function so that each row could be summed to calculate the overall ESF. We use the slant-edge function rather than each individual row fit so that detector artifacts do not randomly displace rows. A 5-point numerical derivative was then applied, and the expected damping function, Eqn. 3.4., due to the derivative, was deconvolved.

For the holographic MTF with the double bi-prism setup, the lower bi-prism is shadowed by the upper which suppresses the Fresnel fringes seen in single bi-prism off-axis holograms [56]. This renders evaluation of mean fringe visibility simple [57,58]. Typical experimental conditions cover the range from  $0.025$  to  $0.3 \text{ pix}^{-1}$ , with approximately 30-35 spatial frequencies chosen and 8 holograms taken at each spatial frequency of holographic fringes to establish reasonable statistics. The entire series of  $\sim 250$  holograms requires  $\sim 90$  minutes to acquire. Holographic fringe spatial frequency is adjusted by changing the projector lens excitation which sets the hologram magnification at the CCD plane. All other parameters that would affect the measured coherence at the detector plane, such as biprism bias, objective and condenser lens excitation, are fixed. Electron intensity on the detector is targeted to be approximately 300 counts/pix and the appropriate exposure time is estimated by short control exposures. Fringe spacing should be sufficiently large ( $> 0.5 \text{ nm}$ ) such that fringe-walk does not decrease the observed visibility as a function of exposure time. The acquisition order for each spatial frequency is randomized so that there is no systemic error introduced as the instrument is affected by environmental instabilities over the course of the experiment.

To avoid potential bias introduced by manual determination of fringe contrast, the data was processed in MATLAB via the method of the sideband to centerband ratio to find the average visibility inside the hologram region [57-59]. Individual holograms may be affected by transients, such as doors being opened in other parts of the laboratory, which produces an outlier datum. Analysis of holographic visibility measurements finds heavy skew of outliers below the mean. Outliers more

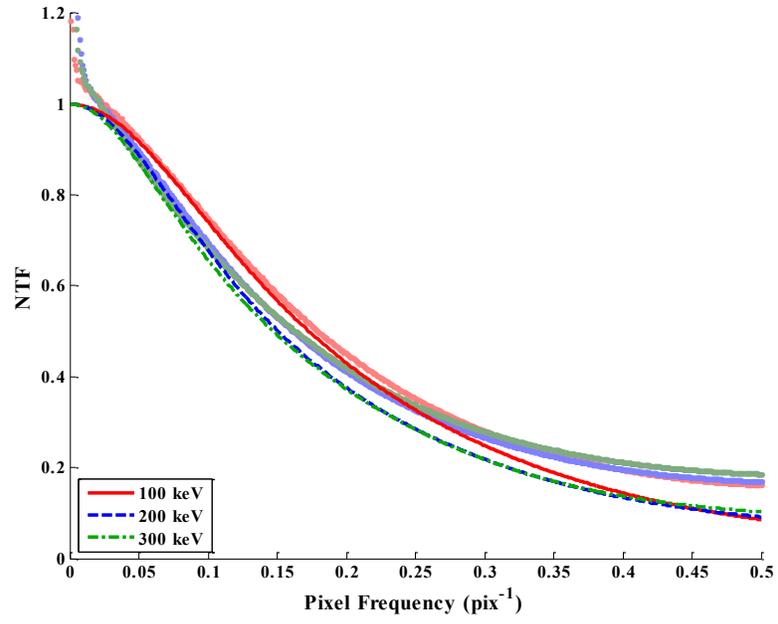
than 2.0 standard deviations outside the mean are discarded. From the data, a Voight profile is generated by least-squares best-fit, and the normalized best-fit is the holographic MTF.

### 3.3.2 MTF RESULTS AND DISCUSSION

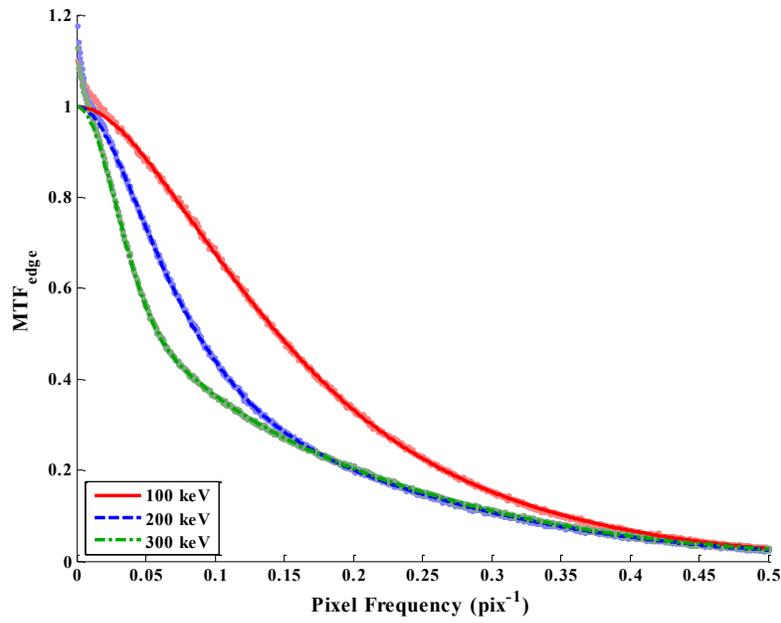
Here we compare the results obtained from all three methods. Best-fit parameters for the Voight profile defined by Eqn. 3.2, may be found in Table A1.1 in the appendices. The variation with primary electron energy is shown in Figure 3.4 for the holographic MTF, slanted-edge MTF and NTF. Lower accelerating voltage results in improved detector performance, but the MTF at 200 keV at spatial frequencies  $> 0.17 \text{ pix}^{-1}$  is comparable to 300 keV. In real-space, this implies the waist of the 200 and 300 keV point spread functions are similar, but the 300 keV PSF has longer tails. This occurs because 300 keV electron has a higher probability to penetrate the bottom face of the scintillator, thereby reducing the interaction volume of the most forwardly-scattered electrons [60]. Electrons exiting the bottom surface of the scintillator would have a cone-like interaction volume, as opposed to the more spherical volume seen from Monte Carlo simulations for infinitely thick targets. Higher energy electrons that undergo high-angle scattering and do not exit the bottom face of the scintillator could travel very long transverse distances, and contribute to the 200 and 300 keV interaction volumes having longer tails than at 100 keV. Additionally, higher energy electrons may also backscatter back into the scintillator from the fiber-optic coupling, further increasing the interaction volume.

Figure 3.5 compares the three methods at each primary electron energy. The NTF is consistently an overestimate of the MTF, and the slanted-edge MTF is consistently an underestimate compared to the holographic MTF. Since the holographic method is directly analogous to lattice fringes, and the sources of error for the edge MTF bias it to underestimate, this suggests the holographic MTF is a more accurate estimate of the detector MTF. The relationship among the NTF, slanted-edge MTF, and holographic MTF is consistent at all measured primary electron energies. The consistency supports the conclusion that the slanted-edge MTF is an underestimate compared to the holographic MTF, and reinforces that the NTF is not a measure of signal MTF.

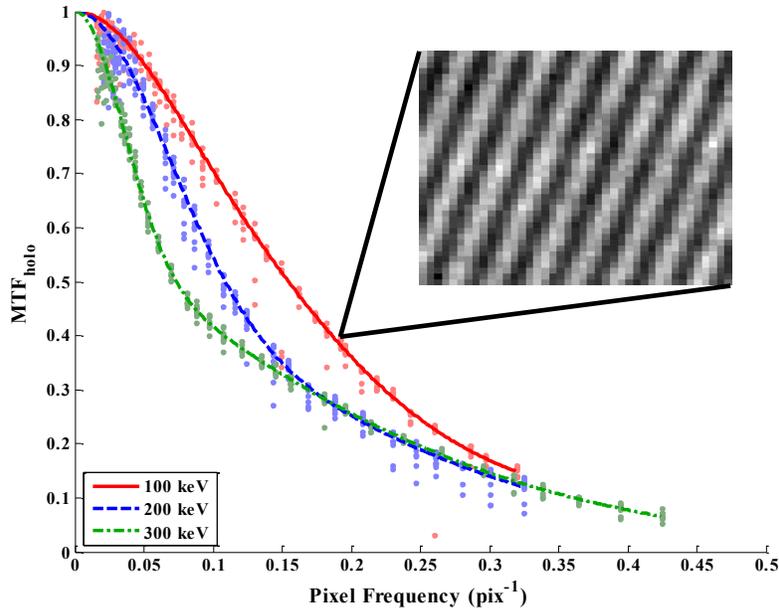
As shown by Figure 3.6, the detector DQE sharply reduces the improvement in signal-to-noise ratio arising from increased source brightness at 300 keV as opposed to 100 keV. The DQE for 100 keV decreases with a monotonic character. In comparison, both the 200 keV and 300 keV DQE show a strong shoulder, with an onset at about  $0.22 \text{ pix}^{-1}$  for 200 keV and  $0.12 \text{ pix}^{-1}$  for 300 keV. It is conceivable that the shoulders arise from the primary electrons having a greater penetration range in the scintillator than the scintillator thickness, as the shoulder originates from the MTF part of the ratio in Eqn. 3.1.



(a)

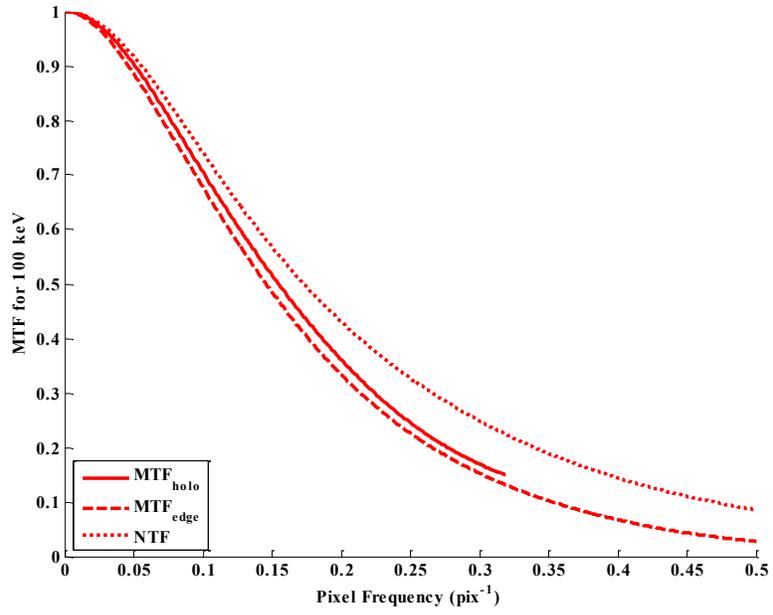


(b)

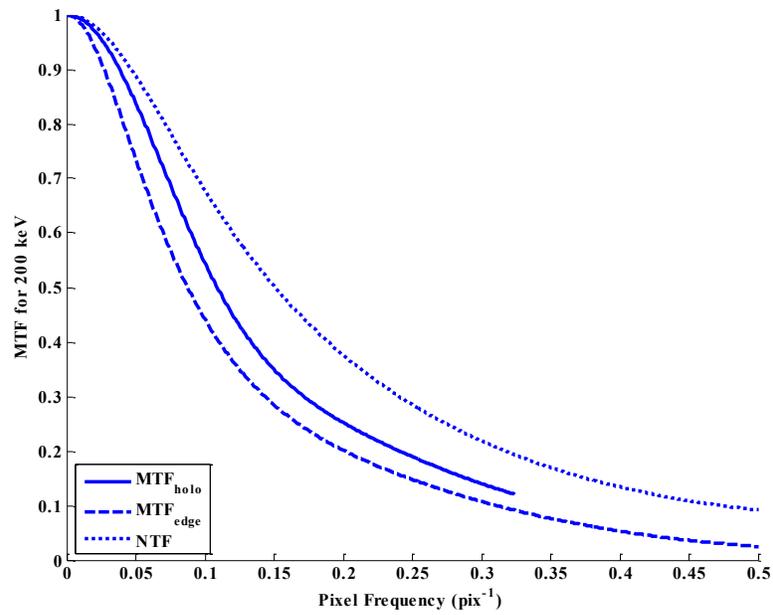


(c)

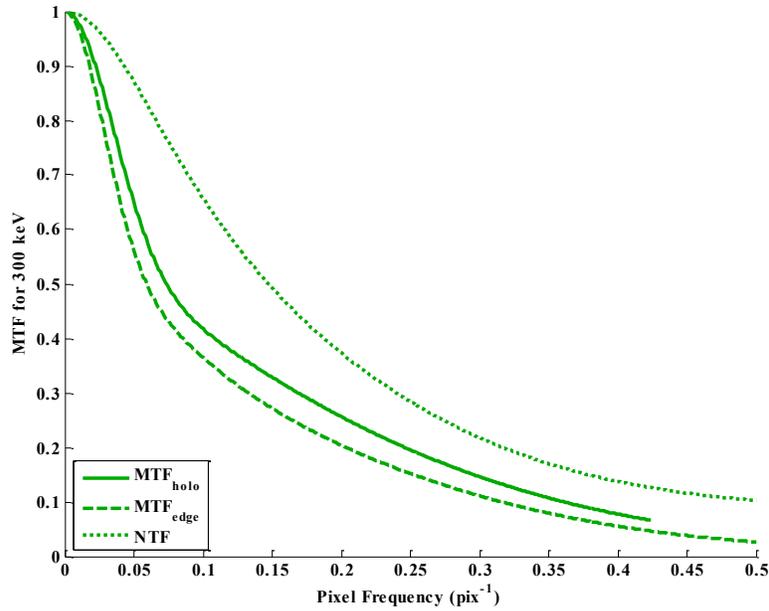
Figure 3.4: (a) NTF for 100, 200 and 300 keV primary electrons for a Gatan Ultrascan 1000. Note that the data (shaded dots) is aliased, with frequencies above  $0.5 \text{ pix}^{-1}$  contributing to frequencies below  $0.5 \text{ pix}^{-1}$ . Dealiasing of fit curves (100 keV – solid red, 200 keV – dashed blue, 300 keV – dotted green) is accomplished as outlined in section 3.2.1.1. (b) Slanted-edge MTF for 100, 200, and 300 keV primary electrons. The 200 keV curve has superior performance to the 300 keV curve up to  $0.16 \text{ pix}^{-1}$ , in agreement with the holographic results. The low frequency peaks are artifacts due to camera read-out, as described in Section 3.2.1.2. (c) Holographic MTF for 100, 200, and 300 keV primary electrons. The data in (c) has been used to generate least-squares best fits. Each curve is composed from over 240 holograms with outliers at each pixel frequency dewighted from the fit. The inset shows an example of holographic fringes ( $40 \times 40$  pixel sample) at  $0.19 \text{ pix}^{-1}$  taken at 100 keV.



(a)



(b)



(c)

Figure 3.5: Comparison of the best-fit results for noise, holographic, and slanted-edge methods at (a) 100 keV, (b) 200 keV, and (c) at 300 keV primary electron energy for a Gatan Ultrascan 1000 detector system. All three accelerating voltages show similar results: the noise power spectrum indicates that noise is transferred at a higher efficiency than signal, with obvious consequences for the DQE. The holographic MTF is consistently higher than the slanted-edge MTF at all frequencies. Best-fits for the holographic method were not extrapolated beyond frequencies than were measured.

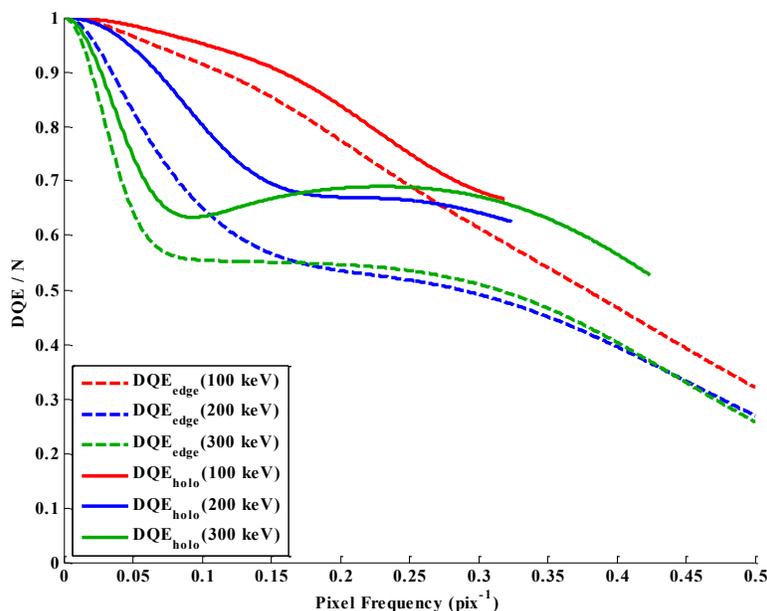


Figure 3.6: Comparison of the DQE normalized by gain and dose for both the edge method (dashed lines) and holographic method (solid lines) at 100 (red), 200 (blue), and 300 (green) keV. This plot is derived by Eqn. 3.1, using the best-fit Voigt profiles for the appropriate MTF and the  $NPS/NTF^2$ .

### 3.3.3 PRACTICAL EXPERIMENTAL TECHNIQUE

In this section we briefly summarize how to optimize the experimental parameters in TEM to account for detector characteristics. Source brightness improves at higher accelerating voltages [61], but as shown in Figure 3.4, the higher the primary electron energy, the worse the detector MTF, and in Figure 3.6, the worse the DQE. Lower energies are valuable if an increase in the elastic interaction constant is required, or if radiation damage is a concern [25,26,62].

The slope of the DQE shown in Figure 3.6 is relevant to the trade-off between contrast and field of view in HRTEM, as determined by microscope magnification. At 300 keV, the DQE shoulder extends from approximately  $0.08 - 0.35 \text{ pix}^{-1}$ , so there is little to no advantage to operating the instrument at 12.5 pix/fringe as opposed to 3 pix/fringe. For the example of crystalline Si 110 with  $d = 0.313 \text{ nm}$ , this implies performance at a magnification of x135k will be identical to x560k. However, magnification higher than x560k where the DQE slope steepens (below  $0.08 \text{ pix}^{-1}$ ) will markedly improve the observed contrast and the resultant signal-to-noise ratio. For the inter-dumbbell spacing of  $d = 0.136 \text{ nm}$ , the contrast will not start to improve until above x1300k magnification, which is higher than typically used in TEM.

For electron energy-loss spectroscopy (EELS) and diffraction, which have peak and edge-type features, the dynamic range of the detector is often the limiting factor. For EELS, there is no hard-and-fast rule on which dispersion to employ because different EELS features have varying degrees of sharpness. In the low-loss, high

dispersion (i.e. small eV/pix) is often used to resolve weak transitions near the strong zero-loss peak, which minimizes the impact of the detector MTF. In the core-loss region a similar argument applies to the near edge fine structure, although the dynamic range typically not a concern. For diffraction, oversaturation and incomplete read-out are typically the main concern, and can result in peak shifting in the direction of detector read-out. The MTF will significantly broaden diffraction peaks at typical camera lengths (30-200 cm), and especially so the diffraction tails at higher accelerating voltages.

Post-processing of images may be applied with characterization of the detector. Proper use of dark and gain references is discussed in section 3.2.1.3. The detector MTF may also be deconvolved from the final image, which increases the sharpness of the image but not the signal-to-noise ratio [40]. Also, deconvolution methods are known to introduce ringing artifacts (a.k.a. Gibb's phenomenon). A solution is to under-deconvolve by the noise floor of the image, i.e. the NTF, which makes the noise in an image 'white' and is less likely to generate ringing artifacts.

### 3.4 CONCLUSIONS

Complete characterization of the instrument is an important aspect for quantitative imaging, such that experiment and simulation are comparable and for optimization of experimental parameters. The long-standing discrepancy between high-resolution TEM simulations and experiments, known as the Stobb's factor [63], is thought to be partially due to the detector MTF [64]. However, precise measurement of the MTF does not ensure accurate measure of the MTF.

We have examined the three MTF measurement methods and refined the details of their implementation. We have expanded the holographic MTF to be a valid method. The slanted-edge MTF suffers from several systematic errors that bias it to underestimate the true MTF. In comparison, the holographic method is directly analogous to imaging (sinusoidal) lattice fringes which suggests that it is an accurate estimate of MTF in a TEM. The slanted-edge method is the standard in photon optics, but nature of high-energy electrons renders it difficult to fabricate a perfect knife-edge for a TEM. In comparison, the fixed magnification of photon optics makes the holographic method experimentally prohibitive, a problem not encountered in TEM. Further work is needed to elucidate whether the discrepancy between the observed contrast of lattice and holographic fringes and the slanted edge MTF at high spatial frequency ( $> 0.2 \text{ pix}^{-1}$ ) is a result of shortcomings of the edge method with high energy electrons or due to higher order transfer effects, as discussed by Niermann et al. [29]. The NTF measures noise-transfer, not signal-transfer, and hence is not equivalent to the MTF but is required to calculate the DQE.

The holographic method as presented struggles to characterize the MTF up to and past the Nyquist limit due to frequency aliasing, which is not in general a limitation of the oversampled slanted edge method [45]. In the future, systemic subpixel

displacements by computer-controlled projector shifts could mitigate this limitation, in the same manner that the slant of the edge does.

For all three methods, accurate normalization of the MTF by the response at zero frequency is critical to obtain an accurate result. Both the slanted-edge and noise methods are affected by read-out artifacts at very low-frequencies, and it is impractical to collect holographic data at very low-frequencies. We have shown that an approximation to a Voigt-profile fit by least-squares provides a reliable estimate of the zero-frequency response. The MTF of thick-scintillator detectors strongly depends on the accelerating voltage. Incomplete read-out of the CCD caused by the short read-out period is potential problem, especially in concert with gain references. In general, restricting the dose to < 20 % of saturation limits incomplete read-out. Gain references should be recorded at the same expected dose to minimize the impact of incomplete read-out on results. Manufacturers could provide the option of different amplification circuits with longer read-out times for users that desire quantitative results.

There are several implications for detector development. The optimal pixel size is determined by the interaction volume of the fast electron in the scintillator: the higher the kinetic energy of the electron, the larger the required pixel size. The DQE results suggest that the pixel size on modern CCD detectors for 300 keV instruments is too small and would benefit from a pixel length on the order of approximately 4x larger than is typical, corresponding to a pixel size of 60  $\mu\text{m}$ . Such large-format scintillators could require an additional projector lens to achieve the desired magnification. Binning large format detectors (> 4x4k) is a potential solution although hardware binning does worsen the MTF due to increased pixelization error, and decreases the detector's dynamic range.

Alternatively, the scintillator may be made thinner as electron scattering is highly anisotropic, so that the interaction volume resembles a narrow cone. This is the approach of direct exposure [65,66] and by lens-coupled detectors [60,67]. Based on published specifications, direct-direction detectors significantly outperform fiber-optic coupled CCDs for low-dose applications [68]. Lens-coupled systems generally have lower collection efficiency than fiber-optic coupled systems. Lens-coupled systems can be designed to demagnify the scintillator image, thereby reducing the oversampling problem, at the cost of further reduced collection efficiency [67].

### 3.5 REFERENCES

- [1] P.A. Midgley, An introduction to off-axis electron holography, *Micron*, 32 (2001) 167-184.
- [2] J.C.H. Spence, J.M. Zuo, Large dynamic range, parallel detection system for electron diffraction and imaging, *Review of Scientific Instruments*, 59 (1988) 2102-2105.
- [3] J.N. Chapman, A.J. Craven, C.P. Scott, Electron detection in the analytical electron microscope, *Ultramicroscopy*, 28 (1989) 108-117.

- [4] O.L. Krivanek, P.E. Mooney, Applications of slow-scan CCD cameras in transmission electron microscopy, *Ultramicroscopy*, 49 (1993) 95-108.
- [5] R.F. Egerton, S.C. Cheng, The Use of Photodiode Arrays to Record 100 keV Electrons, *Journal of Microscopy*, 127 (1982) RP3-RP4.
- [6] D. Drouin, A.R. Couture, D. Joly, X. Tastet, V. Aimez, R. Gauvin, CASINO V2.42—A Fast and Easy-to-use Modeling Tool for Scanning Electron Microscopy and Microanalysis Users, *Scanning*, 29 (2007) 92-101.
- [7] G.D. Boreman, *Modulation Transfer Function in Optical and Electro-Optical Systems*, SPIE Press, Bellingham, 2001.
- [8] L.M. Peng, S.L. Dudarev, M.J. Whelan, *High Energy Electron Diffraction and Microscopy*, Oxford University Press, New York, 2004.
- [9] R.A. McLeod, M. Malac, Determination of Camera Modulation-Transfer Function by Electron Holography, *Microscopy and Microanalysis*, 16 (2010) 140-141.
- [10] J. Gaskill, *Linear Systems, Fourier Transforms, and Optics*, John Wiley & Sons, New York, 1978.
- [11] M. Rabbani, R. Shaw, R. Van Metter, Detective quantum efficiency of imaging systems with amplifying and scattering mechanisms, *J. Opt. Soc. Am. A*, 4 (1987) 895-901.
- [12] M. Rabbani, R. Van Metter, Analysis of signal and noise propagation for several imaging mechanisms, *J. Opt. Soc. Am. A*, 6 (1989) 1156-1164.
- [13] R.R. Meyer, A.I. Kirkland, Characterisation of the signal and noise transfer of CCD cameras for electron detection, *Microscopy Research and Technique*, 49 (2000) 269-280.
- [14] R.R. Meyer, A.I. Kirkland, R.E. Dunin-Borkowski, J.L. Hutchison, Experimental characterisation of CCD cameras for HREM at 300 kV, *Ultramicroscopy*, 85 (2000) 9-13.
- [15] G.Y. Fan, M.H. Ellisman, Digital imaging in transmission electron microscopy, *Journal of Microscopy*, 200 (2000) 1-13.
- [16] R.C. Jones, On the Quantum Efficiency of Photographic Negatives, *Photographic Science and Engineering*, 2 (1958) 57-65.
- [17] P. Fellgett, Equivalent Quantum-Efficiencies of Photographic Emulsions, *Royal Astronomical Society Monthly Notices*, 118 (1958) 224-233.
- [18] R. Shaw, Evaluating the efficiency of imaging processes, *Reports on the Progress of Physics*, 41 (1978) 1105-1150.
- [19] B.H. Armstrong, Spectrum line profiles: The Voigt function, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 7 (1967) 61-88.
- [20] D.C. Joy, *Monte Carlo modeling for electron microscopy and microanalysis*, Oxford University Press, New York, 1995.
- [21] R.F. Egerton, *Electron Energy-loss Spectroscopy in the Electron Microscope*, 3rd ed., Springer, New York, 2011.
- [22] Y. Liu, J. Lin, G. Huang, Y. Guo, C. Duan, Simple empirical analytical approximation to the Voigt profile, *J. Opt. Soc. Am. B*, 18 (2001) 666-672.
- [23] E.G. Stevens, J.P. Lavine, An analytical, aperture, and two-layer carrier diffusion MTF and quantum efficiency model for solid-state image sensors, *Electron Devices, IEEE Transactions on*, 41 (1994) 1753-1760.
- [24] I. Djite, M. Estribeau, P. Magnan, G. Rolland, S. Petit, O. Saint-Pe, Theoretical Models of Modulation Transfer Function, Quantum Efficiency, and Crosstalk for CCD and CMOS Image Sensors, *Electron Devices, IEEE Transactions on*, 59 (2012) 729-737.

- [25] R. Henderson, R.M. Glaeser, Quantitative-Analysis of Image-Contrast in Electron-Micrographs of Beam-Sensitive Crystals, *Ultramicroscopy*, 16 (1985) 139-150.
- [26] R.F. Egerton, P. Li, M. Malac, Radiation damage in the TEM and SEM, *Micron*, 35 (2004) 399-409.
- [27] W.J. de Ruijter, J.K. Weiss, Methods to measure properties of slow-scan CCD cameras for electron detection, *Review of Scientific Instruments*, 63 (1992) 4314-4321.
- [28] E.J. van Zwet, H.W. Zandbergen, Measurement of the modulation transfer function of a slow-scan CCD camera on a TEM using a thin amorphous film as test signal, *Ultramicroscopy*, 64 (1996) 49-55.
- [29] T. Niermann, A. Lubk, F. Roder, A new linear transfer theory and characterization method for image detectors. Part I: Theory, *Ultramicroscopy*, 115 (2012) 68-77.
- [30] A. Lubk, F. Roder, T. Niermann, C. Gatel, S. Joulie, F. Houdellier, C. Magen, M.J. Hytch, A new linear transfer theory and characterization method for image detectors. Part II: Experiment, *Ultramicroscopy*, 115 (2012) 78-87.
- [31] A. Daniels, G.D. Boreman, A.D. Ducharme, E. Sapir, Random transparency targets for modulation transfer function measurement in the visible and infrared regions, *Optical Engineering*, 34 (1995) 860-868.
- [32] M. Vulovic, B. Rieger, L.J. van Vliet, A.J. Koster, R.B.G. Ravelli, A toolkit for the characterization of CCD cameras for transmission electron microscopy, *Acta Crystallographica Section D*, 66 (2010) 97-109.
- [33] J.T. Dobbins, E. Samei, N.T. Ranger, Y. Chen, Intercomparison of methods for image quality characterization. II. Noise power spectrum, *Medical Physics*, 33 (2006) 1466-1475.
- [34] T.-y. Fu, H. Roehrig, Noise power spectrum, MTF, and DQE of photoelectric radiographic systems, in: *Applications of Optical Instrumentation in Medicine XII*, Society of Photo-optical Instrumentation Engineers, San Diego, CA, 1984, pp. 377-386.
- [35] R.J.I. Marks, *Introduction to Shannon Sampling and Interpolation Theory*, Springer-Verlag, New York, 1991.
- [36] A. Rest, L. Mundermann, R. Widenhorn, E. Bodegom, T.C. McGlenn, Residual images in charged-coupled device detectors, *Review of Scientific Instruments*, 73 (2002) 2028-2032.
- [37] M. Malac, R.A. McLeod, P. Li, W.C. Bigelow, J.Y. Howe, L.F. Allard, Y. Taniguchi, M.S. Moreno, Fast mechanical shutter in Hitachi HF 3300, a 60 kV to 300 kV TEM, *Microscopy and Microanalysis*, 16 338-339.
- [38] M. Hayashida, Y. Kimura, Y. Taniguchi, M. Otsuka, Y. Takai, Installation of electric field electron beam blanker in high-resolution transmission electron microscopy, *Review of Scientific Instruments*, 77 (2006) 116114-116113.
- [39] A.R. Faruqi, H.N. Andrews, C. Raeburn, A large area cooled-CCD detector for electron microscopy, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 348 (1994) 659-663.
- [40] J.M. Zuo, Electron detection characteristics of a slow-scan CCD camera, imaging plates and film, and electron image restoration, *Microscopy Research and Technique*, 49 (2000) 245-268.
- [41] S. Miyazaki, Y. Komiyama, M. Sekiguchi, S. Okamura, M. Doi, H. Furusawa, M. Hamabe, K. Imi, M. Kimura, F. Nakata, N. Okada, M. Ouchi, K. Shimasaku, M.

- Yagi, N. Yasuda, Subaru Prime Focus Camera - Suprime-Cam, *Publ. Astron. Soc. Jpn.*, 54 (2002) 833-853.
- [42] S. Engelberg, *Random signals and noise: a mathematical introduction*, CRC Press, 2007.
- [43] M. Bosman, V.J. Keast, Optimizing EELS acquisition, *Ultramicroscopy*, 108 (2008) 837-846.
- [44] S.L. Barna, M.W. Tate, S.M. Gruner, E.F. Eikenberry, Calibration procedures for charge-coupled device x-ray detectors, *Review of Scientific Instruments*, 70 (1999) 2927-2934.
- [45] W.J. de Ruijter, Imaging properties and applications of slow-scan charge-coupled device cameras suitable for electron microscopy, *Micron*, 26 (1995) 247-275.
- [46] P. Mooney, J.R. McIntosh, Optimization of Image Collection for Cellular Electron Microscopy, in: *Methods in Cell Biology*, Academic Press, 2007, pp. 661-719.
- [47] L. Huang, U.L. Oesterberg, Cross talk study of flexible multimode fibers, in: *Fiber Optic Sensors V*, SPIE, Beijing, China, 1996, pp. 311-320.
- [48] B. Tatian, Method for Obtaining the Transfer Function from the Edge Response Function, *J. Opt. Soc. Am.*, 55 (1965) 1014-1019.
- [49] S.E. Reichenbach, S.K. Park, R. Narayanswamy, Characterizing digital image acquisition devices, *Optical Engineering*, 30 (1991) 170-177.
- [50] P.D. Burns, Slanted-Edge MTF for Digital Camera and Scanner Analysis, in: *Proc. IS&T 2000 PICS Conference*, 2000, pp. 135-138.
- [51] E. Samei, M.J. Flynn, D.A. Reimann, A method for measuring the presampled MTF of digital radiographic systems using an edge test device, *Medical Physics*, 25 (1998) 102-113.
- [52] E. Samei, N.T. Ranger, J.T. Dobbins, I.I.Y. Chen, Intercomparison of methods for image quality characterization. I. Modulation transfer function, *Medical Physics*, 33 (2006) 1454-1465.
- [53] J.W. Coltman, The Specification of Imaging Properties by Response to a Sine Wave Input, *J. Opt. Soc. Am.*, 44 (1954) 468-469.
- [54] D.B. Dove, Image Contrasts in Thin Carbon Films Observed by Shadow Electron Microscopy, *Journal of Applied Physics*, 35 (1964) 1652-1653.
- [55] F. Hue, C.L. Johnson, S. Lartigue-Korinek, G. Wang, P.R. Buseck, M.J. Hytch, Calibration of projector lens distortions, *Journal of Electron Microscopy*, 54 (2005) 181-190.
- [56] K. Harada, A. Tonomura, Y. Togawa, T. Akashi, T. Matsuda, Double-biprism electron interferometry, *Applied Physics Letters*, 84 (2004) 3229-3231.
- [57] R.A. McLeod, M. Malac, Evaluating Visibility and Spatial Resolution in Electron Holography, *Microscopy and Microanalysis*, 14 (2008) 854-855.
- [58] R.A. McLeod, M. Malac, Determination of Localized Visibility in Off-axis Electron Holography, Submitted (2012).
- [59] E. Voelkl, L.F. Allard, A. Datye, B. Frost, Advanced electron holography: a new algorithm for image processing and a standardized quality test for the FEG electron microscope, *Ultramicroscopy*, 58 (1995) 97-103.
- [60] B. Bailey, D.A. Agard, J.W. Sedat, M. Barfels, H. Chao, P.E. Mooney, Performance Evaluation of a Transmission Scintillator-Based Lens-Coupled 4K CCD Camera for Use in Low-Dose Electron Cryo-Tomography, *Microscopy and Microanalysis*, 10 (2004) 1204-1205.

- [61] D.B. Langmuir, Theoretical Limitations of Cathode-Ray Tubes, Proceedings of the Institute of Radio Engineers, 25 (1937) 977-991.
- [62] R.F. Egerton, R. McLeod, F. Wang, M. Malac, Basic questions related to electron-induced sputtering in the TEM, Ultramicroscopy, 110 (2010) 991-997.
- [63] M.J. Hytch, W.M. Stobbs, Quantitative comparison of high resolution TEM images with image simulations, Ultramicroscopy, 53 (1994) 191-203.
- [64] A. Thust, High-Resolution Transmission Electron Microscopy on an Absolute Contrast Scale, Physical Review Letters, 102 (2009) 220801-220804.
- [65] L. Jin, Direction electron detection in transmission electron microscopy, in: Physics, UC San Diego, San Diego, 2009.
- [66] P. Mooney, D. Contarato, P. Denes, A. Gubbens, B. Lee, M. Lent, D. Agard, A High-Speed Electron-Counting Direct Detection Camera for TEM, Microscopy and Microanalysis, 17 (2011) 1004-1005.
- [67] G.Y. Fan, M.H. Ellisman, High-sensitivity lens-coupled slow-scan CCD camera for transmission electron microscopy, Ultramicroscopy, 52 (1993) 21-29.
- [68] B.E. Bammes, R.H. Rochat, J. Jakana, D.H. Chen, W. Chiu, Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency, J. Struct. Biol., 177 (2012) 589-601.

## 4 PHASE MEASUREMENT ERROR IN SUMMATION OF ELECTRON HOLOGRAPHY SERIES

The contents of Chapter 4 were submitted for publication to *Ultramicroscopy* by R.A. McLeod, M. Bergen, and M. Malac on March 26<sup>th</sup>, 2013. Contributing author M. Bergen wrote the original version of the hologram series summation software, provided valuable insight on how to average complex images, and provided valuable feedback throughout the development process. M. Malac edited the paper and provided valuable discussion throughout the development process.

### *ABSTRACT*

Off-axis electron holography is a method for the transmission electron microscope (TEM) that measures the electric and magnetic properties of a specimen. The electrostatic and magnetic potentials modulate the electron wavefront phase. The error in measurement of the phase therefore determines the smallest observable changes in electric and magnetic properties. Here we explore the summation of a hologram series to reduce the phase error and thereby improve the sensitivity of electron holography. Summation of hologram series requires independent registration and correction of image drift and phase wavefront drift, the consequences of which are discussed. Optimization of the electro-optical configuration of the TEM for the double biprism configuration is examined. An analytical model of image and phase drift, composed of a combination of linear drift and Brownian random-walk, is derived and experimentally verified. The accuracy of image registration via cross-correlation and phase registration is characterized by simulated hologram series. The model of series summation errors allows the optimization of phase error as a function of exposure time and fringe carrier frequency for a target spatial resolution. An experimental example of hologram series summation is provided on  $WS_2$  fullerenes. A metric is provided to measure the object phase error from experimental results and compared to analytical predictions. The ultimate experimental object root-mean-square phase error is 0.006 rad ( $2\pi/1050$ ) at a spatial resolution less than 0.615 nm and a total exposure time of 900 s. The ultimate phase error in vacuum adjacent to the specimen is 0.0037 rad ( $2\pi/1700$ ). The analytical prediction of phase error differs with the experimental metrics by +7 % inside the object and -5 % in the vacuum, indicating that the model can provide reliable quantitative predictions.

### 4.1 INTRODUCTION

Off-axis *electron holography* (EH) in the *transmission electron microscope* (TEM) is a technique for the characterization of electrostatic and magnetic properties of a specimen [1][2]. A population of electrons, each being a wave-particle, has a distribution of amplitude and phase shift in both time and space. The electron is phase shifted by the electric and magnetic potentials integrated along its path [3]. The most principal electromagnetic phase shift is that of the object/specimen, analogous to the index of refraction of a material phase shifting a photon.

An attractive strategy for improving the *signal-to-noise ratio* (SNR) of electron holograms is summation of a hologram series [4][5][6][7][8][9]. An image series can be used to break a long exposure into frames with the image drift for each frame corrected by cross correlation. In an electron hologram the strongest feature is the fringes of the interference pattern, which shift with drift of the electron phase. As a result, application of cross-correlation alignment cannot be applied to electron holograms directly to correct the image drift. Registration of the image and phase drift must be performed separately, which requires operating on the complex data of the reconstructed hologram. In this paper we provide methods to optimize the phase error of hologram series as a function of the targeted spatial resolution.

In Section 4.2, we introduce the dependence of phase error on spatial resolution and how to optimize the microscope column. In Section 4.3 we extend the optimization for a hologram series, taking into account image and wavefront phase drift. In Section 4.4 we provide examples of high-resolution hologram series results on a specimen of inorganic fullerene WS<sub>2</sub>. In Section 4.4.2 we develop a metric to measure phase error experimentally and compare it to the estimates derived in Section 4.3.4. A description of the algorithm used to align and sum a hologram series is found in appendix A2.

## 4.2 OPTIMIZATION OF ELECTRO-OPTICAL CONFIGURATION

In off-axis electron holography, the electron wavefront is split by an electrostatic biprism. The object wave (subscript 1) passes through the specimen, while the reference wave (subscript 2) passes through vacuum adjacent to the specimen. The two waves converge on a detector at semi-angle  $\theta$ , forming an interference fringe pattern, or hologram. The interference pattern intensity,  $\psi_{12}^2(\mathbf{r})$ , as a function of position  $\mathbf{r}$ , is given by,

$$\psi_{12}^2(\mathbf{r}) = A_1^2(\mathbf{r}) + A_2^2(\mathbf{r}) + 2V(\mathbf{r}, q_c, \alpha_o)A_1(\mathbf{r})A_2(\mathbf{r})\cos(2\pi\mathbf{q}_c\mathbf{r} + \phi_1(\mathbf{r}) - \phi_2(\mathbf{r})) \quad (4.1)$$

where  $A_1$  and  $A_2$  are the object and reference wave amplitude, respectively,  $\phi_1$  and  $\phi_2$  are the object and reference wave phase shift, respectively,  $\mathbf{q}_c = 2\sin(\theta)/\lambda$  is the carrier frequency of the fringe pattern,  $V$  is the holographic visibility (or fringe contrast) which depends on the illumination angle  $\alpha_o$  and the separation of virtual sources determined by  $q_c$ .

The path length, i.e. phase, difference between the reference and object waves measures both the electrostatic potential background of the object, and the electrostatic and magnetic potentials/fields [10][11][12]. As with all measurements, there is an associated error that limits the minimum variation in specimen thickness, compositional variation, and electric and magnetic field that may be measured [13]. The standard error of the phase has historically been expressed by the estimate of its variance [14][15][16][17],

$$\sigma_{\phi}^2(\mu, V) = 2/\mu V^2 \quad (4.2)$$

where,  $\mu$  is the number of electrons *per reconstructed pixel* and  $V$  is the holographic visibility, which may be calculated by various means [18]. Minimization of phase error requires maximization of both current density and holographic visibility. Holographic visibility is largely dependent on the high wavefront coherence which results from parallel, widely-spread illumination. However, the more widely spread the illumination, the lower the current density incident on the specimen. Long exposure times may be used to increase the electron dose, but image drift, shifts in specimen position, and phase drift, changes in the electron path-length, blur the hologram in space and phase. Thus it is necessary to optimize the electro-optical configuration in order to minimize the phase error [19].

The estimator of Eqn. 4.2 is limited in that it does not relate the spatial resolution to the phase error, although the two are linked [20]. Phase error is related to spatial resolution, both due to the *optical-transfer function* (OTF) of the TEM [21] and *modulation-transfer function* (MTF) of the detector [22]. The effective binning of the detector from the holographic reconstruction process also has a major impact on the mean dose per pixel. For example, an estimated phase error of 0.001 rad ( $2\pi/6300$ ) with a measured phase noise of 0.02 rad ( $2\pi/300$ ) has been demonstrated at 12 nm spatial resolution [23], compared to an estimated phase error of  $\sim 0.06$  ( $2\pi/100$ ) at 0.1 nm spatial resolution [24]. As we show in Section 4.3, the drift of the image and holographic fringes also affects the phase error, especially for prolonged exposures at high spatial resolutions.

Reconstruction of electron holograms is typically done using the Fourier method [2][20]. In Fourier-space, a hologram consists of a central band at zero frequency (i.e. the autocorrelation) and two sidebands, one at the spatial frequency of the fringe pattern and the other at its complex conjugate. The phase shift is encoded in the positions of the fringe pattern, therefore reconstruction of the complex (amplitude and phase) electron wave-function is achieved by isolating the sideband with an apodization function, such as a von Hann window, translation to zero-frequency, and then an inverse Fourier transform is applied.

To introduce the spatial frequency dependence, the electron counts is defined as [20],

$$\mu = t_x I_e R \cdot \text{DQE}_{cca}(\mathbf{q}) \quad (4.3)$$

where  $t_x$  is exposure time,  $I_e$  is the electron flux at the object plane,  $q_c$  is the carrier frequency of the fringes, and  $R$  is the resampling factor of the apodization, which represents the pixel binning that occurs during the Fourier reconstruction [25][26]. For a circular ‘hard’ von Hann apodization, used throughout this paper,

$$R = (2\sqrt{2}/a)^2$$

where  $a$  is the radius of the von Hann window (in reciprocal space units) and the additional factor of  $\sqrt{2}$  is the Ishizuka resampling factor, due to the apodization being circular and the detector is assumed to be square. In general, the apodization radius is substituted by half the carrier frequency  $a = q_c/2$ . In some cases, a more restrictive apodization is necessary to prevent cross-talk between the centerband (autocorrelation) and sideband, but for the remainder of the paper we will use the substitution. The effective dose is reduced by the *detector quantum efficiency* (DQE) [27]. We use a simplified version of the  $DQE_{ccd}(\mathbf{q}) = MTF_{ccd}^2(\mathbf{q})/NTF_{ccd}^2(\mathbf{q})$ , where  $MTF_{ccd}$  is the holographic MTF and  $NTF_{ccd}$  is the *noise-transfer function* [22]. This is not a complete treatment of the DQE as it does not account for the variation in DQE with dose, but it does effectively estimate the increase in shot noise over the expected Poisson value.

The observed visibility a combination of many factors [20],

$$V = \gamma(q_c, \alpha_o) MTF_{osc} V_{obj} MTF_{drift}(q_c, t_x) MTF_{ccd}(q_c - pix) \quad (4.4)$$

where  $\gamma$  is the complex degree of coherence of the source which is demagnified by the illumination optics  $\alpha_o$ ,  $MTF_{osc}$  represents the loss of contrast from high-frequency oscillations (both of the source and biprisms),  $V_{obj}$  is the loss of contrast from incoherent scattering in the object,  $MTF_{drift}$  is the visibility loss due to drift of the holographic fringes over the exposure time, and  $MTF_{ccd}$  is the holographic-based detector MTF [22]. The complex degree of coherence [28][29] and high-frequency oscillation MTF are combined into  $V_n(q_c, \alpha_o) = \gamma(q_c, \alpha_o) MTF_{osc}$  as shown in Figure 4.1. For the moment, we will omit the visibility of the object/specimen, which must be measured experimentally [18], and the MTF due to fringe drift, which we address in section 4.3. Then substituting into Eqn. 4.3 and 4.4 into 4.2 yields,

$$\sigma_\phi(\mathbf{q}, t_x, I_e, q_c, \alpha_o) = \sqrt{\frac{q_c^2}{t_x I_e V_n(q_c, \alpha_o)^2}} \cdot MTF_{ccd}^{-1}(q_c - pix) DQE_{ccd}^{-0.5}(\mathbf{q}_{pix}) \quad (4.5)$$

The detector DQE and MTF are explicitly separated because they depend on TEM magnification (i.e. they are functions of pixel spatial frequency). The impact of Eqn. 4.5 on optimization can be best understood in terms of operational parameters of the microscope: 1.) for the illumination, increased condenser excitation spreads the beam and makes the illumination more parallel. Visibility is increased but current density is decreased; 2.) for the biprism, the carrier frequency  $q_c$  and interference-field width increase with bias; 3.) the measured visibility of fine fringes is decreased by the detector MTF and the effective shot noise increased due to DQE.

Furthermore, detector MTF and DQE are a function of primary electron energy. Lower primary electron energy improves the detector MTF and DQE for slow-scan CCD detectors, increases the interaction constant with electrostatic fields [30], but decreases brightness [31] and penetration range.

The choice of specimen determines the desired *field of view* (FoV) which in turn determines the hologram width and magnification. The specimen also determines the target spatial frequency  $q_t$ , in the holographic reconstruction, which could be lattice fringes or some other feature to be investigated. The carrier frequency of the holographic fringes must be  $q_c \geq q_t/2$  due to the Fourier reconstruction process, and for the single biprism configuration is optimally  $q_c = q_t/3$  [21].

For the double biprism configuration [32], the hologram width is controlled by the bias of the lower biprism,  $BP_2$ . As such,  $BP_2$  is fixed based on the size of the object. Then the relationship between visibility  $V$  and carrier frequency  $q_c$  may be defined in terms of the upper biprism bias,  $BP_1$ . The relationship is affected both by source coherence and detector MTF. As such, the detector MTF should be measured and an analytic best-fit calculated [22], such that the visibility may be calculated independent of magnification (denoted  $V_n$  for normalized by detector MTF). The  $V_n$  for our TEM at 200 keV is shown in Figure 4.1.

$$q_c = 0.060 BP_1 + 0.045 BP_2 - 0.038 [nm^{-1}] \quad (4.6)$$

$$V_n = 0.94 - 0.0023 BP_1 - 0.0065 BP_2 - 6.8 \cdot 10^{-6} BP_2^2 + 2.7 \cdot 10^{-5} BP_1 BP_2 \quad (4.7)$$

Then we can fix  $BP_2$  to generate the desired hologram width  $w$  and find an expression for  $V_n(q_c, w)$ . We omit this step to be concise; it is accomplished by solving Eqn. 4.6 and 4.7. Extrapolation of Eqn. 4.7 to zero biprism bias,  $V_{00} = 0.94 \pm 0.025$ , provides an estimate for the loss of visibility due to instrument instabilities that are fast relative to the exposure time. This is expected to be oscillations of biprisms and potentially the electron gun by 60 Hz disturbances as well as higher frequency components such as vibrations arising from turbomolecular pumps.

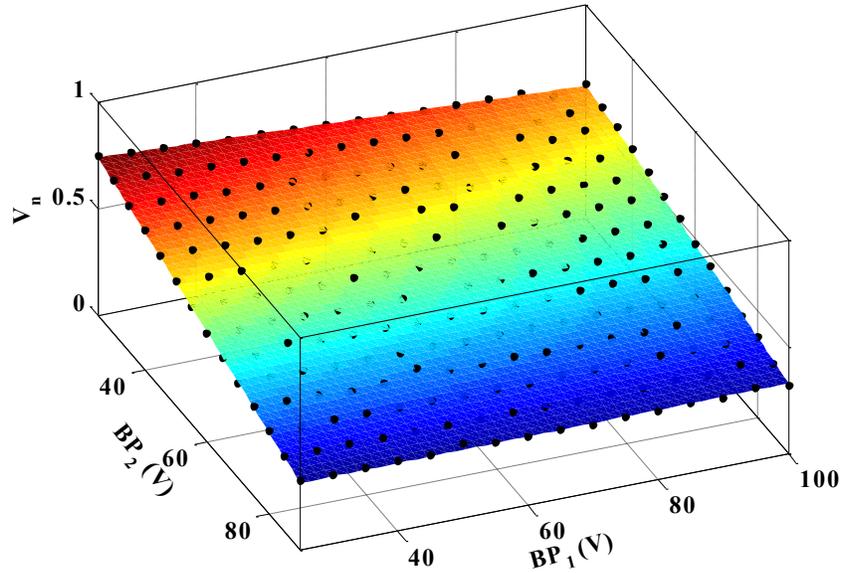


Figure 4.1: Visibility as a function of upper biprism bias  $BP_1$  and lower biprism bias  $BP_2$  at 200 keV at an optimized condenser lens  $C_1$  excitation. Blue dots are data and the colored surface is a polynomial fit of with a correlation coefficient  $R^2 = 0.97$ . The MTF of the detector has been removed, so the visibility is independent of magnification. The carrier frequency  $q_c$  covers the range  $2.6-10 \text{ nm}^{-1}$  ( $0.1 - 0.38 \text{ nm}$  wavelength). The fringe frequency increases with biprism bias.

The illumination convergence,  $\alpha_o$ , is optimized based on a series of reference holograms of over a range of condenser lens excitations and plotting the phase error estimated by Eqn. 1. As shown in Figure 4.2, there is a minimum phase error at specific condenser excitation. A condenser excitation optimized to minimize the phase error effectively fixes the electron current density  $I_e$ , and removes a degree of freedom from the optimization.

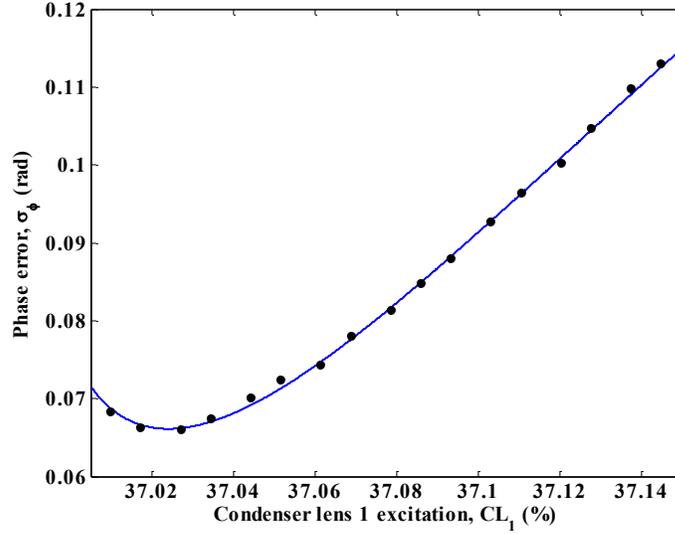


Figure 4.2: Phase error as a function of condenser excitation, normalized to an exposure time of 1 s for  $BP_1 = 40$  V,  $BP_2 = 30$  V. The condenser excitation that corresponds to the minimum phase error is easily selected. The image-plane cross-over is located at 36.97 % excitation.

### 4.3 SOURCES OF ERROR IN HOLOGRAM SERIES SUMMATION

In section 4.2, we described how to optimize the microscope configuration ignoring sample and wavefront phase drift, the later originating from biprism and illumination tilt drift. In this section, we expand our optimization to practical holographic series on the parameters of  $m$  frames each with an exposure time of  $t_x$  and carrier frequency  $q_c$ . Ideally, the addition of  $m$  frames should result in a linear increase in the counts, such that the sum phase error would be

$$\sigma_{\Sigma-\phi} = m^{-0.5} \sigma_{\phi}(\mathbf{q}, t_x, I_e, q_c, \alpha_o) \quad (4.8)$$

Each pixel in a hologram reconstruction may be considered to be a *complex circular random variables* (CCRV) [33][34]. As with normally distributed random variables, the mean and variances of CCRVs sum along the real and imaginary axes. As such, the phase of CCRVs may be averaged by summing their real and imaginary components, forming the basis for improved phase resolution through hologram series [7]. First image drift is registered via cross-correlation [35]. The reconstructed holograms are then shifted. With the two holograms aligned in space, the mean phase difference [36] between the two reconstructed holograms is estimated and subtracted to remove wavefront phase drift. Extended discussion of the algorithm details is provided in appendix A2.

In practice the summation of a series of holograms suffers from drift of the image (object) and phase (fringes). Drift causes both shifts in mean position, which can be estimated with some uncertainty and removed, and motion blur integrated over the duration of the exposure, which acts as a MTF in frequency space or *point-spread*

*function* (PSF) in image space. To gain insight into the effect of drift on holograms, drift is modeled as a linear combination of linear, constant velocity drift and Brownian motion random-walk.

We take this approach because while techniques have been developed to extract the complete drift kernel they generally require a high signal-to-noise ratio beyond what is achieved in electron holography due to the use of blind deconvolution methods [37][38][39] or use a sensor to track and estimate the motion [40]. A third component of motion in the TEM is oscillations with high-frequency relative to the exposure time, such as the ubiquitous 60 Hz disturbance [41][42]. We do not account for the high-frequency MTF because it does not worsen with increasing exposure time.

Gaussian random walks are those where a particle have a normal probability of moving in any direction. For the TEM, we generalize this to the motion of the specimen observed at an image plane. In general, for a Gaussian random walk the total displacement scales with the square root of the number of steps. For Brownian motion, which is a Gaussian random walk with an infinitesimal step size, the mean square displacement,  $\langle r^2 \rangle$ , as a function of time is

$$\langle r^2 \rangle = 2nDt_x \quad (4.9)$$

where  $n$  is the number of dimensions,  $D$  is the diffusivity constant, and  $t_x$  is the exposure time [43][44].

Thus there are six sources of error that impact the phase error of hologram series:

1. Image drift velocity,  $v_i$ : constant velocity image drift, typically caused by environmental gradients, such as temperature or pressure, acting on the TEM.
2. Phase drift velocity,  $v_\phi$ : constant velocity phase drift, typically caused by gradients acting on the biprism holders or tilt of the beam by lens and deflector drift.
3. Image diffusivity,  $D_i$ : random-walk of the object.
4. Phase diffusivity,  $D_\phi$ : random-walk of the fringes from the biprism.
5. Image registration error,  $\delta_i$ : the accuracy of the cross-correlation based registration is limited by shot noise, which is modulated by the detector.
6. Phase registration error,  $\delta_\phi$ : the accuracy of the scalar phase drift registration is limited by shot noise, image drift, and potentially specimen charging during acquisition of the series.

The first four errors affect individual holograms while the registration errors, 5 and 6, are particular to the summation of hologram series.

### 4.3.1 IMAGE DRIFT

Image drift acts upon the image as a motion blur that is modeled as an MTF. Conceptually if one breaks down the image drift over the course of the exposure time into infinitesimal velocity slices  $dv$  and time slices  $dt$ , the resulting MTF is [45],

$$\text{MTF}_i(q, dv, dt) = \int \text{sinc}(\pi dv dt q) dt$$

Constant-direction linear drift is typically caused by thermal gradients affecting the TEM. The MTF due to linear drift component is the Fourier transform of a rectangular PSF, sometimes referred to as a Heaviside- $\pi$  function, which has a cardinal sine form,

$$\text{MTF}_{v_i}(q, t_x) = |\text{sinc}(\pi \mathbf{v}_i t_x q)| \quad (4.10)$$

Theoretically the linear component of the image drift can be deconvolved after  $\mathbf{v}_i$  is estimated by cross-correlation. However, the cardinal sine function has zeros if the drift exceeds two pixels, so the deconvolution is ill-defined due to division by zero.

The *probability-distribution function* (PDF) of a Gaussian random-walk for 2-D isotropic diffusion with diffusion coefficient  $D_i$  at time  $t$  is [46][47],

$$\text{PDF}_{D_i}(\mathbf{r}, t) = \frac{1}{4\pi D_i t} \exp\left(\frac{-r^2}{4D_i t}\right) \quad (4.11)$$

where  $\mathbf{r}$  is the spatial image coordinate. Then the PSF is the PDF integrated over  $[0, t_x]$ ,

$$\text{PSF}_{D_i}(\mathbf{r}, t_x) = \frac{1}{t_x} \int_0^{t_x} \frac{1}{4\pi D_i t} \exp\left(\frac{-r^2}{4D_i t}\right) dt = \frac{1}{4\pi D_i t_x} \Gamma\left(0, \frac{-r^2}{4D_i t_x}\right)$$

where  $\Gamma$  is the Euler gamma function. Then the MTF is the Hankel transform  $r \rightarrow 2\pi q$ ,

$$\text{MTF}_{D_i}(q, m, t_x) = 2\pi \int_0^{\infty} \text{PSF}_{D_i} J_0(2\pi q r) r dr = \frac{(1 - \exp(-4\pi^2 D_i t_x q^2))}{4\pi^2 D_i t_x q^2} \quad (4.12)$$

where  $J_0$  is a Bessel function of the first kind. Eqn. 4.12 is plotted in Figure 4.3 and compared to Monte Carlo simulation of the MTF. Discretization error in the Monte Carlo simulation was minimized by interpolating the pixel position. The analytical solution to Eqn. 4.12 (solid blue line) matches the Monte Carlo simulation well, although residual discretization error is observed. In Figure 4.4 an estimate of the random-walk  $\text{MTF}_{D_i}$  for a range of exposure times is provided. The damping from the random-walk blur is very significant at high spatial resolutions and long exposure times. The linear combination of constant-velocity and random-walk drift  $\text{MTF}_i$  has the form,

$$\text{MTF}_i(q, m, t_x) = |\text{sinc}(\pi \mathbf{v}_i t_x q)| \frac{1 - \exp(-4\pi^2 D_i t_x q^2)}{4\pi^2 D_i t_x q^2} \quad (4.13)$$

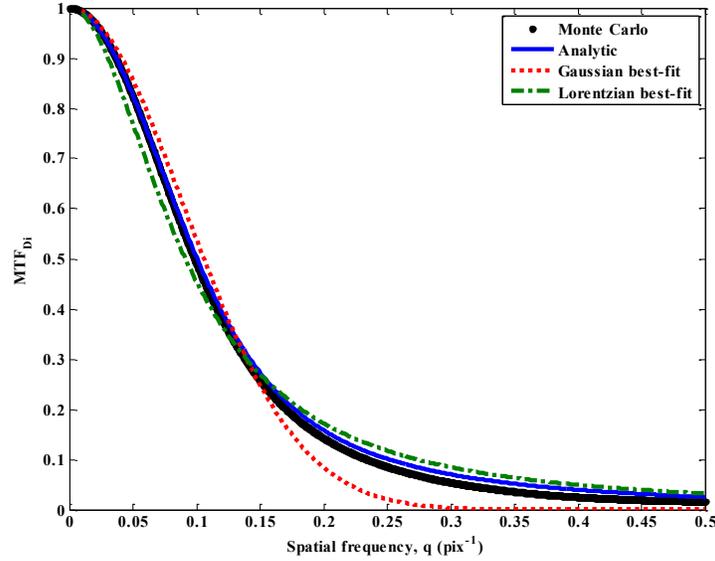


Figure 4.3: (a) Example of modulation-transfer function  $MTF_{D_i}$  from Gaussian random-walk in 2-D when simulated by Monte Carlo (black dots) for  $D_i = 0.8 \text{ pix}^2/\text{s}$  and  $t_x = 5 \text{ s}$ . The exposure time was broken up into 100 000 steps and the MTF was then averaged over 2000 Monte Carlo runs. The analytical prediction from Eqn. 4.12 is presented as the solid blue line. For comparison, best-fits of a Gaussian (dotted red) and Lorentzian (green dot-dash) to the Monte Carlo results are shown.

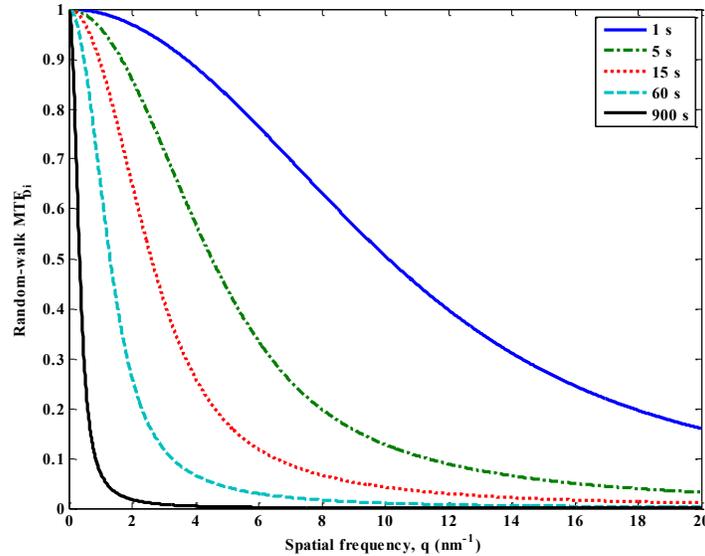


Figure 4.4: Examples of random-walk  $MTF_{D_i}$  from Eqn. 4.12 for an image diffusivity of  $D_i = 395 \text{ pm}^2/\text{s}$  for selected exposure times.

Accurate optimization requires extensive characterization of the drift instability of the TEM in its quiescent laboratory environment. To establish that the object drift in our Hitachi HF-3300 TEM can be expressed as a superposition of constant-velocity drift and a random walk, a 1500-image drift series at x560k was acquired over  $\sim 6.5$  hours. Each frame was cross-correlated to the previous frame to measure the

displacement  $\mathbf{r}$  with the time between frames  $dt = 15.07$  s. Then the constant-velocity drift component can be found from the mean,  $\langle \mathbf{r} \rangle = \mathbf{v}_i dt$ , and the random-walk component from the variance  $\langle (\mathbf{r} - \langle \mathbf{r} \rangle)^2 \rangle = 4D_i dt$ . The histogram shown in Figure 4.5, shows the drift velocity has a Gaussian distribution. The mean of the histogram corresponds to an image drift velocity of  $\mathbf{v}_i = -20 \hat{\mathbf{x}} - 41 \hat{\mathbf{y}} = 45$  pm/min. The diffusivity coefficient is  $D_i = 385 \hat{\mathbf{x}} + 405 \hat{\mathbf{y}} = 395$  pm<sup>2</sup>/s or  $\sim 0.8$  pix<sup>2</sup>/s at typical holography magnification ( $\sim \times 600k$ ). Based on the data, random-walk is the dominate component of image drift in our HF-3300 TEM. If the random-walk drift is not isotropic, then Eqn. 4.15 derived for 1-D random-walks in Section 4.3.1 may be applied independently for each axis.

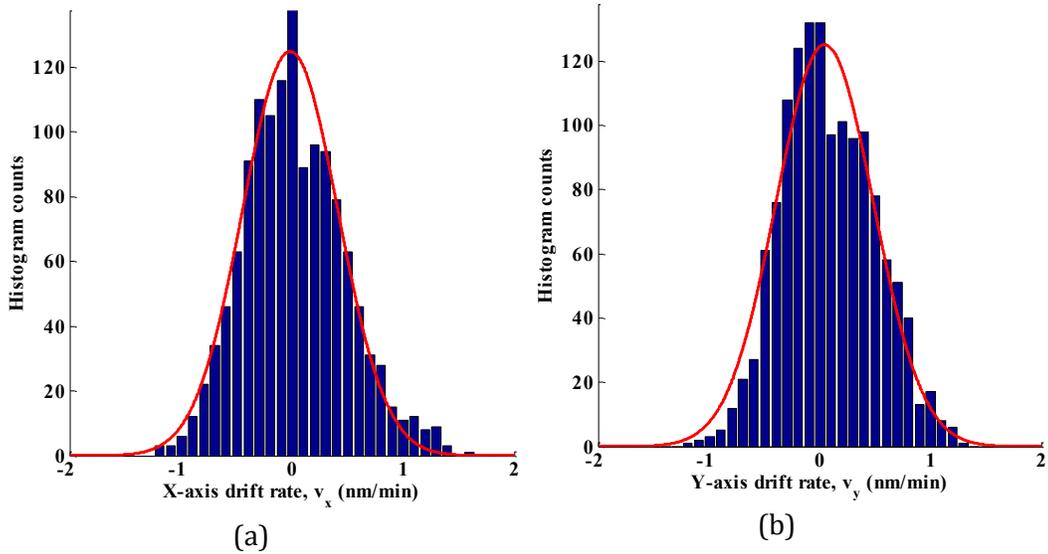


Figure 4.5: Histogram of the image drift registered by cross-correlation for 1500 successive images along the (a) x-axis and (b) y-axis. The specimen was Au nanoparticles on amorphous Ge substrate at a magnification of  $\times 560k$ . Average time between exposures was 15.07 s. This shows that the image drift is a Gaussian distribution, as expected from the probability distribution function of Eqn . 4.10.

The object drift characterization is dependent on the composition and morphology of the specimen. The reference sample for this case was Au nanoparticles (3-5 nm diameter) on amorphous Ge substrate. It is evident that the Au nanoparticles move relative to each other as well as moving as an ensemble from video of the drift series. There are a number of potential causes of specimen random-walk in the TEM including thermal forces, secondary electron emission charge forces, and plasmonic forces [48]. The analysis does not separate the motion of the nanoparticles from that of the entire sample at present.

#### 4.3.2 PHASE DRIFT

In this subsection we discuss wavefront drift, which is observed as the holographic fringes walking back-and-forth on the detector. Phase drift arises due to a

combination of biprism drift and tilt of the beam through the column changing the relative path length of the object and reference wavefronts. We conducted experiments with artificially induced beam tilt that showed only a scalar phase shift. Consequently we conclude the phase drift can be modeled as a scalar. This implies that the mean phase shift can be subtracted from each frame in a series so that only the phase drift that occurs during an exposure impacts the phase error.

Phase drift causes motion blur of the sinusoidal interference fringes which worsens the observed visibility and consequently the phase error. The random-walk MTF of the phase is found analogous to that of the image, except in 1-D. The PDF is defined as

$$\text{PDF}_{D_\phi}(x, t) = \frac{1}{\sqrt{2\pi D_\phi t}} \exp\left(\frac{-x^2}{2D_\phi t}\right) \quad (4.14)$$

where  $D_\phi$  the phase diffusivity in  $\text{pm}^2/\text{s}$ . Then integrating  $t$  over  $[0, t_x]$ ,

$$\text{PSF}_{D_\phi}(x, t_x) = \sqrt{\frac{2t_x}{\pi D_\phi}} \exp\left(\frac{-x^2}{2D_\phi t_x}\right) - \frac{|x|}{D_\phi} \left[1 - \text{erf}\left(\frac{|x|}{\sqrt{2D_\phi t_x}}\right)\right]$$

The MTF is found by taking the Fourier transform  $x \rightarrow 2\pi q$  and normalizing by the limit  $t_x$  as  $q \rightarrow 0$ ,

$$\text{MTF}_{D_\phi}(q, t_x) = \frac{1 - \exp(-2\pi^2 D_\phi t_x q^2)}{2\pi^2 D_\phi t_x q^2} \quad (4.15)$$

Thus MTF for a random-walk in 1-D is the same as 2-D except for the change in dimension in Eqn. 4.9. On the basis of 1-D diffusive motion of the fringes, the visibility as a function of exposure time,  $V^*$ , can be found from associated MTFs at the carrier frequency  $q_c$ ,

$$V^*(q_c, v_\phi, D_\phi, t_x) = |\text{sinc}(\pi q_c v_\phi t_x)| \frac{1 - \exp(-2\pi^2 D_\phi t_x q_c^2)}{2\pi^2 D_\phi t_x q_c^2} V_n \quad (4.16)$$

where  $q_c$  the fringe frequency in  $\text{pm}^{-1}$ .

Similar to image drift, the phase drift was characterized by a 1500 hologram series with a fringe carrier wavelength,  $\lambda_c = 336 \text{ pm}$  and a mean time between exposures of 4.0 s. Based on the histogram of the phase drift, shown in Figure 4.6, the fringe drift velocity is  $v_\phi = -0.0005 \text{ rad/s} = -0.025 \text{ pm/s}$ , which is negligible. The phase diffusivity coefficient is  $D_\phi = 0.003 \text{ rad}^2/\text{s} = 9.6 \text{ pm}^2/\text{s}$  at the specimen plane. The histogram should not be significantly broadened from the measurement error of the average phase, which is estimated at  $0.005 \text{ rad}$  from Eqn. 4.2.

In order to verify the behavior of fringe drift, a second experiment was performed of holograms recorded at different exposure times. Experimental results are shown in Figure 4.7 for 0.1 nm fringe spacing holograms taken at exposure times ranging

from 1 to 45 s, alongside analytical predictions for  $D_\phi = 10 \text{ pm}^2/\text{s}$ . The results show that there is excellent agreement between the predicted and measured visibility  $V^*$ . Data for Figure 4.6 and Figure 4.7 were taken ten months apart, so we consider the characterization of  $D_\phi$  quite stable for a TEM in its quiescent laboratory environment. We cannot experimentally measure a loss of visibility for broad fringes (0.5 nm) with exposure times up to 120 s.

It is necessary to state the phase drift in terms of picometers of fringe displacement relative to the object plane because the loss of visibility changes with fringe carrier frequency. For the double biprism configuration [32] the demagnification relative to the object plane is fixed in order to place the top biprism at an image plane. However, for single biprism configurations, where changing the defocus of the biprism to alter relationship between carrier frequency and hologram width also changes the demagnification of the biprism relative to the object plane, characterization of phase drift is more onerous.

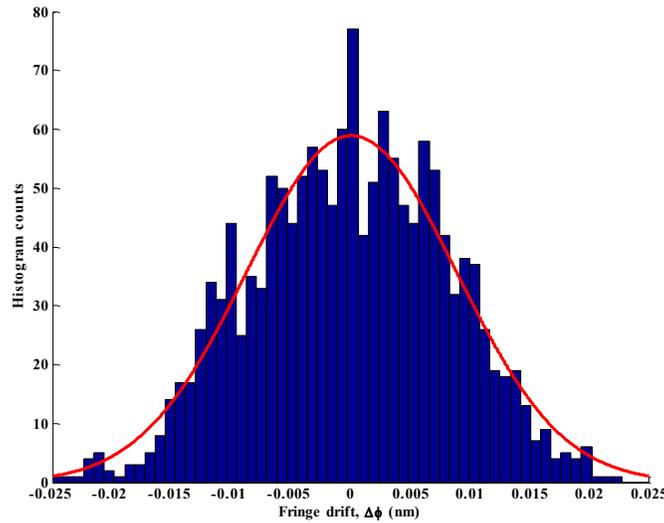


Figure 4.6: Histogram of phase drift for 0.336 nm fringes the Hitachi HF-3300 TEM for a 1500 hologram series, with an average time between exposures of 4.0 s. The histogram shows that the drift follows the probability distribution function defined by Eqn. 4.14.

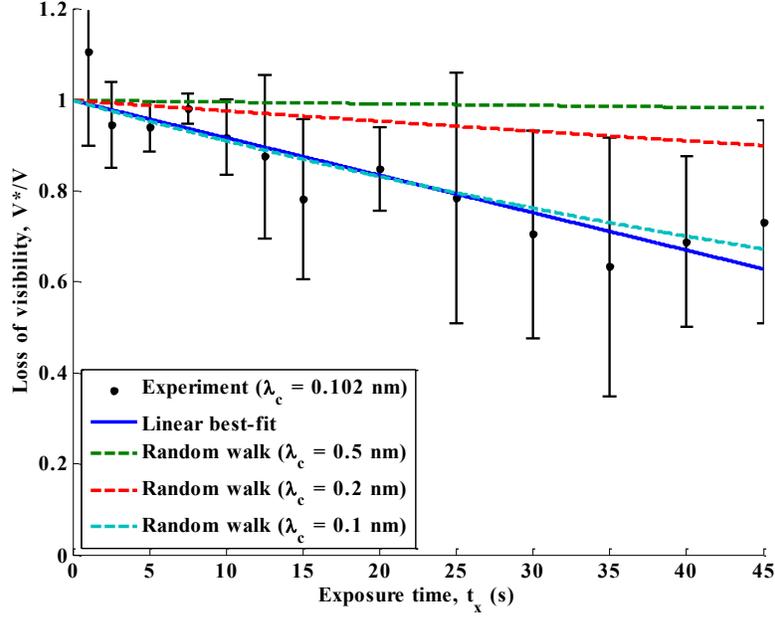


Figure 4.7: Visibility loss (normalized to 1 at 0 s by the linear best fit to the data) as a function of exposure time for selected fringe wavelengths. Twelve holograms were taken at each exposure time (black dots with error bars). The analytical solution to Eqn. 4.16 (dashed cyan line) shows good agreement to a linear best-fit (dark blue line) of the experimental 0.1 nm holographic fringes. The analytical solution shows that for large fringe spacing (green and red dashed), the effect of phase random-walk is minimal.

### 4.3.3 REGISTRATION ERRORS

In this subsection the errors specific to the summation of hologram series are investigated. In the presence of noise and other image artifacts such as motion blur, the cross-correlation of the image drift and of the estimation of the mean phase drift are not exact. This results in both a drift registration error and phase registration error.

The drift registration error is composed of a sum of  $m$  Dirac functions with standard deviation  $\delta_i$ . As  $m \rightarrow \infty$ , the PSF becomes a Gaussian function with standard deviation  $\delta_i$ .

$$\text{PSF}_{\delta_i}(\mu, V, r) = \exp\left(-r^2/2\delta_i^2(\mu, V)\right) \quad (4.17)$$

and

$$\text{MTF}_{\delta_i}(\mu, V, q) = \exp(-2\pi^2\delta_i^2(\mu, V)q^2) \quad (4.18)$$

The phase registration error is the error in estimating the mean over the *region-of-interest* (RoI), which ideally is

$$\delta_\phi = \sqrt{2/hw\mu V^2} \quad (4.19)$$

where  $h$  is the height and  $w$  the width of the RoI. If the pixels of reconstructed holograms are treated as complex circular random variables (CCRVs) [33][34], then

total phase variance is  $\sigma_\phi^2 + \delta_\phi^2$ . As discussed below in the results for simulated holograms, the phase registration error is much higher in the presence of image drift. As such, Eqn. 4.19 is in general only valid for reference holograms.

The primary challenge for estimation of registration errors is that the true drift must be known. As the large number of practical factors renders analytic approaches intractable, we have resorted to simulations to estimate the registration errors. The process is outlined as follows: a hologram object is generated. For  $m = 25$  hologram series, random object drift and phase drift are generated and applied to each frame. Motion blur is applied to each frame. The carrier frequency is applied to form a hologram and Poisson shot noise is individually calculated for each pixel based on expected dose. Detector MTF is applied to the image. The resulting holograms and resulting references are then processed by our hologram averaging algorithm identically to experimental data. This process was repeated  $\sim 3 \cdot 10^4$  times to generate reasonable statistics. The resulting empirical relationships are,

$$\delta_i = 8.3\mu^{-1} + 0.3 \text{ pix} = 0.34t_x^{-1} + 0.3 \text{ pix} \quad (4.20)$$

$$\delta_\phi = 1.6 \cdot 10^{-3} \sqrt{\mu} + 2 \cdot 10^{-3} \text{ rad} = 7.9 \cdot 10^{-3} \sqrt{t_x} + 2 \cdot 10^{-3} \text{ rad} \quad (4.21)$$

Essentially no dependence of registration errors on holographic visibility is found. This is expected for cross-correlation performed on the centerband as it does not vary with visibility. The phase registration error is an order of magnitude higher than expected by Eqn. 4.19, and appears to be governed by the motion blur from image drift,  $\text{MTF}_i$ . High phase registration error is not observed with simulated reference holograms, for which  $\delta_{\phi-ref} \sim 0.001 \text{ rad}$ .

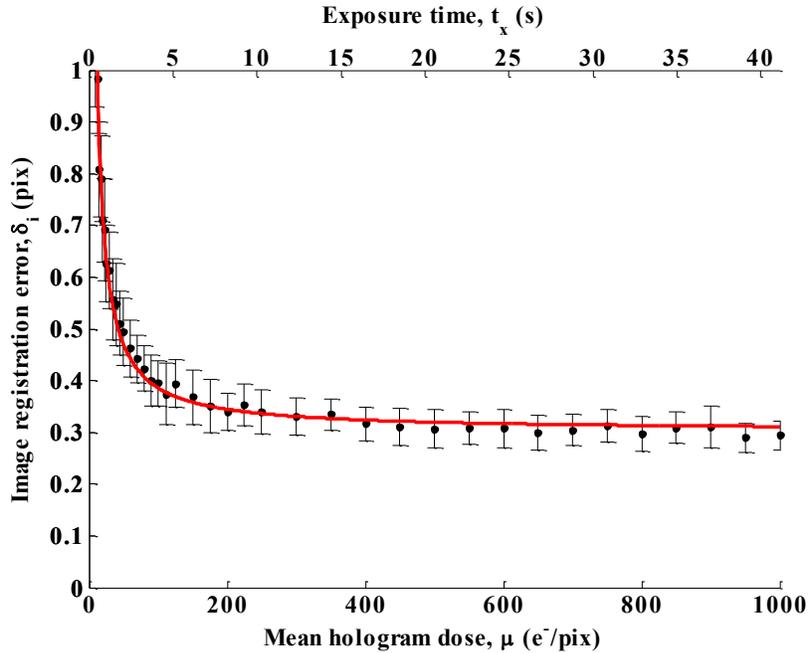


Figure 4.8: Image registration error for simulated holograms,  $q_c = 0.1 \text{ pix}^{-1}$ .

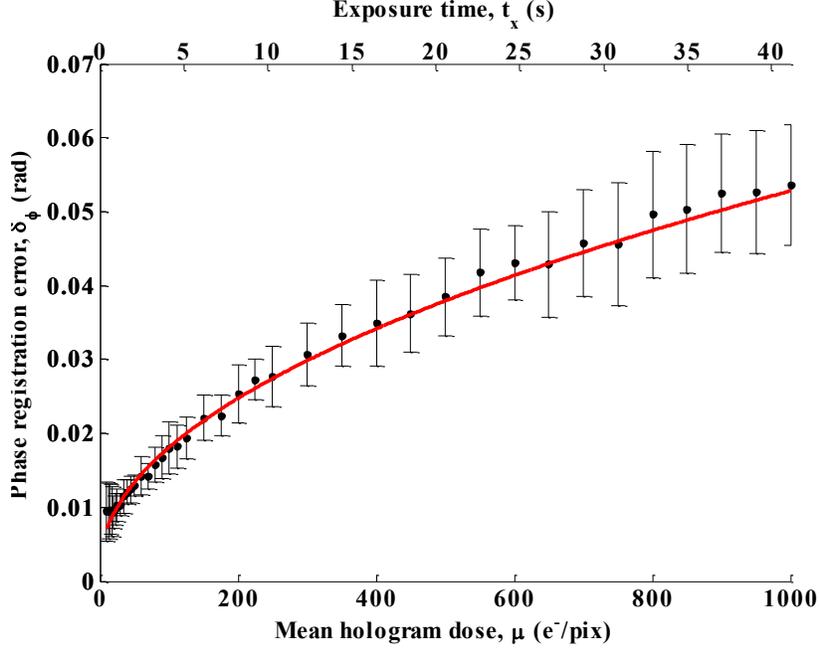


Figure 4.9: Phase registration error for simulated holograms,  $q_c = 0.1 \text{ pix}^{-1}$ .

#### 4.3.4 COMBINED PHASE ERROR AND SPATIAL RESOLUTION FOR HOLOGRAM SERIES

With expressions for the drift and registration errors derived above, the phase resolution for a hologram series may be defined as,

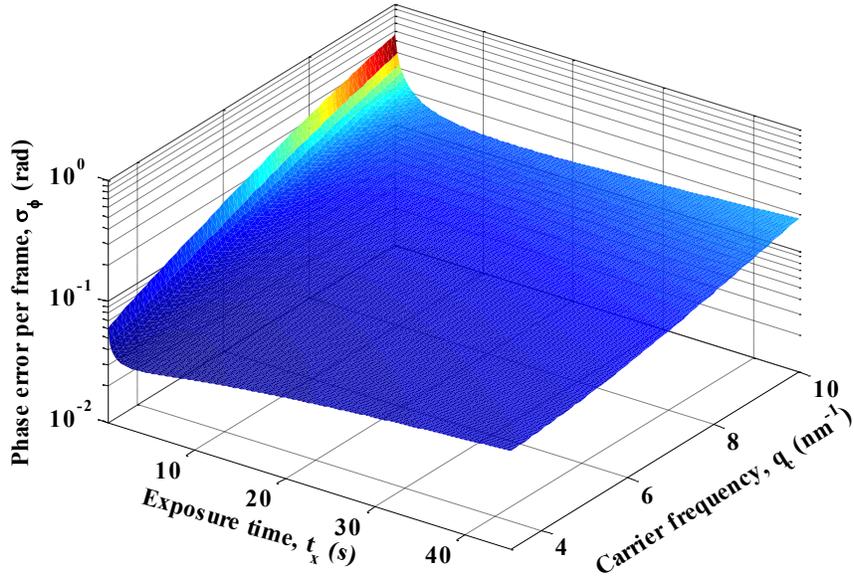
$$\sigma_{\Sigma-\phi}(\mathbf{q}, m, t_x, q_c) = \frac{1}{\sqrt{m}} \left( \sigma_{\phi}(\mathbf{q}, q_c, t_x) + \delta_{\phi}(\mathbf{q}, q_c, t_x) \right) \quad (4.22)$$

where

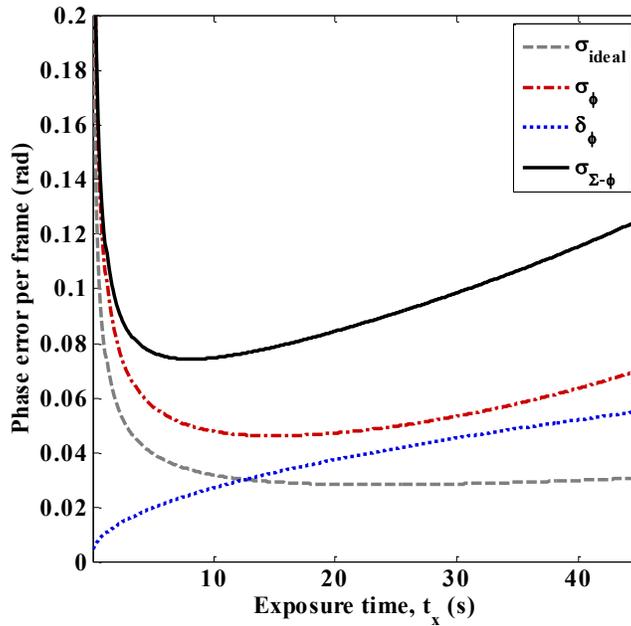
$$\sigma_{\phi} = \frac{q_c^2}{\sqrt{t_x I_e V^*(q_c, w, v_{\phi}, D_{\phi}, t_x)^2 \cdot \text{MTF}_{ccd}^{-1}(q_c) \text{DQE}_{ccd}^{-0.5}(q) \text{MTF}_{\delta_i}^{-1}(q, t_x) \text{MTF}_i^{-1}(q, \mathbf{v}_i, D_i, t_x)}}$$

This equation includes terms from Eqns. 4.13, 4.16, 4.18, 4.20, and 4.21. The phase registration error, presented in section 4.3.3, includes the inverse MTF terms in the simulation. The main factors to optimize for any given spatial frequency  $q$  is the exposure time  $t_x$ , and carrier frequency  $q_c < q_{\max}/2$ . Other parameters, such as the drift, are characteristic of the instrument whereas illumination and biprism bias optimization was previously discussed in section 4.2. A surface plot of sum phase error versus exposure time and carrier frequency for the target frequency,  $q = 1.62 \text{ nm}^{-1}$  is shown in Figure 4.10.a. The components of the phase error are well-behaved and generally change monotonically, so no local minima are observed. A cross-section through  $q_c = 10 \text{ nm}^{-1}$  is provided in Figure 4.10.b which provides

also the components of the phase error to show how they inter-relate with exposure time.



(a)



(b)

Figure 4.10: (a) Optimization of phase error  $\sigma_{\Sigma-\phi}$  for exposure time and carrier frequency. (b) A cross-section through  $q_c = 0.1 \text{ pix}^{-1}$ , components of the phase error. In (b) we use experimentally observed (see section 4.4)  $V = 0.163$  for  $\sigma_\phi$  inside the object and  $V = 0.224$  for  $\sigma_{ideal}$ , which forms the estimate for the vacuum phase error.

For most experiments it is desirable to optimize for the total exposure time,  $t_{tot} = mt_x$ , which is equivalent to finding the minimum of  $\sigma_{\Sigma-\phi}\sqrt{t_x}$ . Optimizing for  $t_{tot}$

leads to short exposure times,  $t_x \sim 0.5$  s. Based on practical experience, the image registration error can be considerably worse at low doses compared to what is predicted in Figure 4.8, possibly due to correlated noise from detector artifacts or another factor unaccounted for in the hologram simulation. It is generally advisable to operate in regimes where the slope of the image registration error flattens,  $> 100 e^-/pix$  from Figure 4.8. This corresponds to a threshold of  $\sim 10\,000 e^-/pix$  in the reconstruction for the example case of  $q_c = 0.1 pix^{-1}$ . Care should be taken to acquire high-quality dark and gain references for the detector to avoid the build-up of correlated artifacts [22].

To predict  $\sigma_{\Sigma-\phi}$  for an experiment, the distribution of spatial frequencies in a hologram must be known. The spatial frequency distribution in an object hologram is given by the rotational average of the *power spectral density* (PSD), such that the average phase error is,

$$\langle \sigma_{\Sigma-\phi} \rangle = \frac{2}{q_c} \int_0^{q_c/2} PSD_{holo}(q) \sigma_{\Sigma-\phi}(q, m, t_x, q_c) dq \quad (4.23)$$

where  $\langle \sigma_{\Sigma-\phi} \rangle$  is the average phase error over the entire hologram. The measured  $PSD_{holo}$  includes the detector MTF, and we explicitly include the detector MTF in  $\sigma_{\Sigma-\phi}$ . As such, this estimate is biased towards lower spatial frequency response. We justify this as  $\sigma_{\Sigma-\phi}$  varies slowly with  $q$ .

#### 4.4 EXPERIMENTAL HOLOGRAM SERIES

In order to establish the quantitative nature of the phase error derived in section 4.3, it is desirable to compare it to results from an experimental hologram series. The experiment was performed on a Hitachi HF-3300 TEM equipped with a cold-field emission gun, double biprism configuration [32], and Gatan US1000 detector. The biprisms were biased to generate fringes of  $\lambda_c = 0.16 nm = 9.8 pix$  and a hologram width of  $30.6 nm$ . The magnification of the object was  $\times 860k$ , corresponding to a detector pixel size of  $16 pm$ . The object was inorganic-fullerene Tungsten Disulphide ( $WS_2$ ) from Nanomaterials (apnano.com) [49], which has an SWS—SWS interlayer spacing of  $0.615 nm$ , suspended on holey Carbon film. A total of 150 object holograms were recorded with an exposure time of  $5 s$  each and readout time of  $0.8 s$  each using MATLAB control software. Although the Hitachi HF-3300 TEM normally operates at  $300 keV$ , the experiment was performed at  $200 keV$  to reduce the damping of holographic visibility by the detector and limit radiation damage via sputtering of Sulfur atoms [50]. At a carrier frequency  $q_c = 0.1 pix^{-1}$ , the detector MTF is  $0.42$  at  $300 keV$  and  $0.54$  at  $200 keV$  [22]. Similarly, at a lattice frequency of  $q = 0.03 pix^{-1}$ , the detector DQE is  $0.87$  at  $300 keV$  and  $0.98$  at  $200 keV$ . Spherical rather than elliptical illumination was used for this experiment to ensure isoplanicity of the electron wavefront [51].

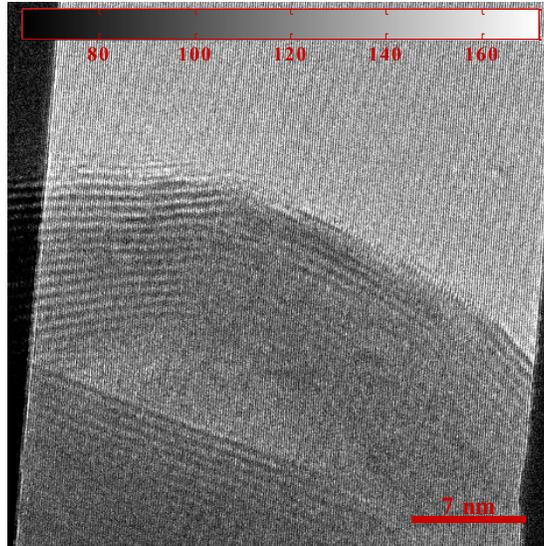
#### 4.4.1 EXPERIMENTAL RESULTS

An example experimental hologram is shown in Figure 4.11(a) beside the (b) reconstructed cumulative centerband. The layers of W atoms are evident as dark bands. In the central portion of the image is a WS<sub>2</sub> fullerene, while a second fullerene is superposed on the bottom quarter of the image. The local visibility is shown in Figure 4.12, from which the mean visibility in the vacuum region is  $V_{vac} = 0.224 \pm 0.003$  while  $V_{obj} = 0.163 \pm 0.014$  was observed inside the central WS<sub>2</sub> fullerene.

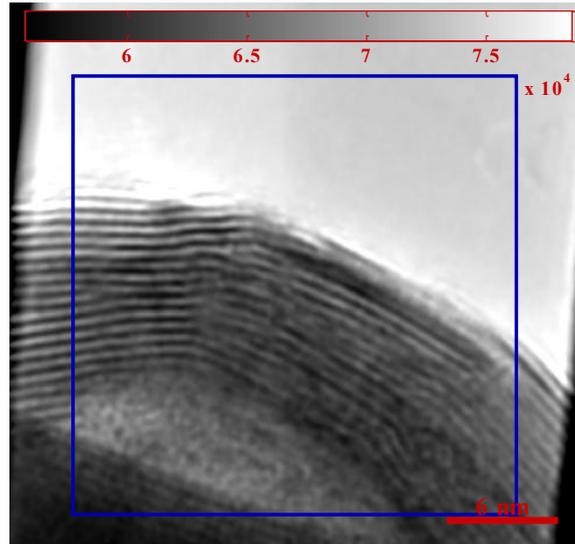
A side-by-side comparison of the unregistered cumulative phase and the registered cumulative phase is shown in Figure 4.13. Both phase maps have been unwrapped using Schofield's Laplacian method [52], but the unregistered phase features numerous unwrapping errors that originate on the interior surface and kinks of the WS<sub>2</sub> fullerene. Strong interlayer contrast is not observed in the phase, compared to the amplitude or visibility. Instead, object thickness dominates. This implies the difference in mean inner potential between the tungsten and sulfur layers is smaller than the effect of channeling from the high-Z tungsten layers [24].

The experimental values for drift is shown in Figure 4.14 for (a) image and (b) phase. Mean values for image drift are  $v_i = -4 \hat{x} + 1 \hat{y} \text{ pm/s}$  and  $D_i = 160 \text{ pm/s}^2$ , and the phase drift is  $v_\phi = -0.02 \text{ pm/s}$  and  $D_\phi = 65 \text{ pm/s}^2$ .

Appendix Video A2.1 shows the unregistered centerband and registered and centerband side-by-side. From the absence of observed motion in the registered centerband, we can infer the success of the image cross-correlation technique. Appendix Video A2.2 shows the cumulative phase and a cross-section the across the WS<sub>2</sub> lattice. It can be seen in the cross-section of the registered phase that there is no loss of contrast in the lattice fringes with frame number. As such, we have not significantly worsened the spatial resolution. The holograms were apodized to  $2a \times 2a = 210 \times 210 \text{ pix}$  in the Fourier reconstruction series, but the reconstructions are zero-padded to  $1024 \times 1024 \text{ pix}$ . The combination of subpixel registration and shifting and zero-padding reduces aliasing concerns for lattice fringes.



(a)



(b)

Figure 4.11: An (a) experimental hologram. Shown alongside is (b) the registered cumulative centerband, the signal used for image correlation. The blue box indicates the region of interest used for image and phase registration. Color bars give electron dose per pixel. The counts per pixel in (b) is much higher than (a) due to both the summation of 150 holograms and effective binning in the reconstruction process.

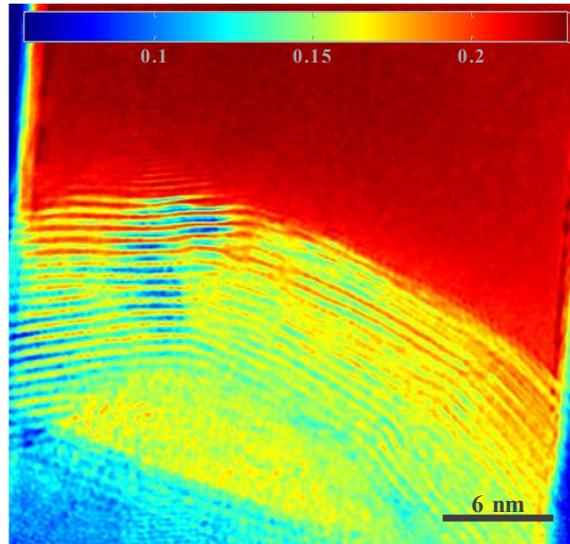


Figure 4.12: Registered cumulative visibility map [18]. Color bar gives measured visibility obtained by the Fourier-ratio method.

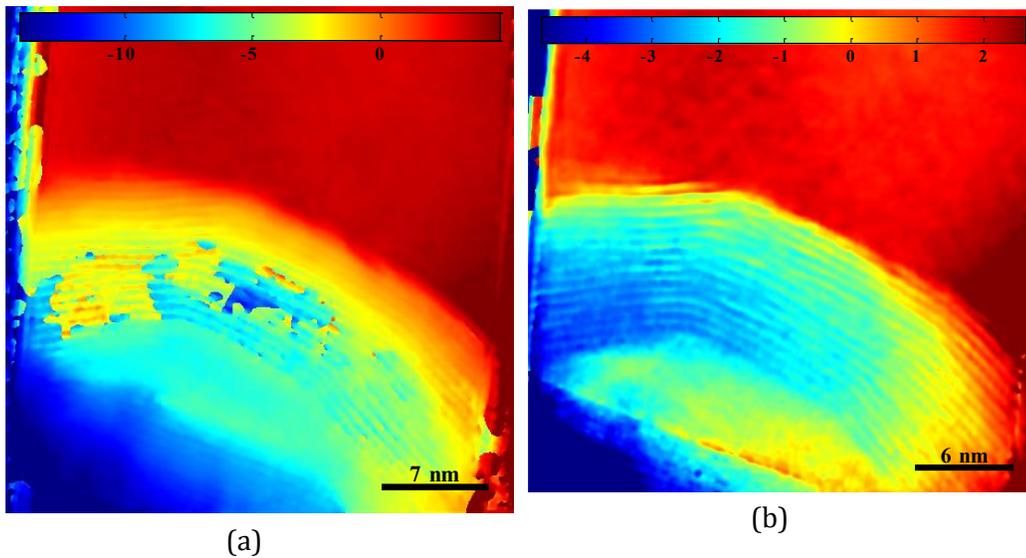


Figure 4.13: (a) Unregistered, unwrapped cumulative phase compared to (b) registered, unwrapped cumulative phase of the hologram series. Color bars give phase shift in radians.

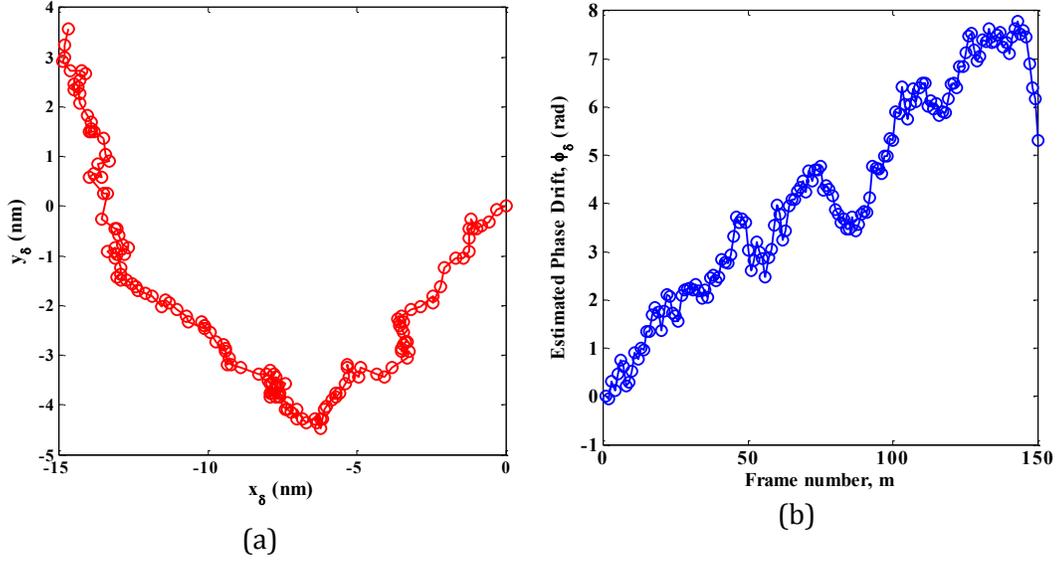


Figure 4.14: Estimated drift path of the object in nanometers and (b) estimated phase shift of the electron-optics, as a function of frame number. The time between frames was 5.8 s.

#### 4.4.2 EXPERIMENTAL PHASE ERROR METRIC

Previous hologram series methods have provided phase error estimates based on the standard deviation of phase in vacuum [7][9]. To track the success of image and phase registration, it is necessary to define the phase error within the object of an object hologram. For this, we have developed a metric based on the work of Fienup in optical diffractive imaging [53]. For two images with wavefront phase  $\phi_1(\mathbf{r}) \pm \sigma_\phi$  and  $\phi_2(\mathbf{r}) \pm \sigma_\phi$ , then from the root-mean-squared difference,

$$\Delta\phi_{rms} = \sqrt{\frac{1}{2} \langle \text{angle}(\exp(i\phi_1(\mathbf{r}) \pm i\sigma_\phi) / \exp(i\phi_2(\mathbf{r}) \pm i\sigma_\phi))^2 \rangle} \approx \sigma_\phi(n) \quad (4.24)$$

where diagonal brackets indicate the mean over all pixels  $\mathbf{r}$  in the region of interest, and *angle* computes the phase of a complex number. Essentially Eqn. 4.24 uses two object holograms as references for each other. This metric averages over all component spatial frequencies, and does not measure loss of spatial resolution.

A simple and robust means to evaluate the experimental phase error is to reconstruct two sub-series, composed of even-numbered and odd-numbered frames respectively. The difference of the two sub-series registered cumulative phase gives the RMS even-odd phase error,

$$\Delta\phi_{e-o}(m) = \frac{1}{2} \sqrt{\langle \text{angle} \left( \frac{\sum_{j=1}^{m/2} \exp(i\phi_{2j}(\mathbf{r}))}{\sum_{j=1}^{m/2} \exp(i\phi_{2j+1}(\mathbf{r}))} \right)^2 \rangle} \quad (4.25)$$

where  $m$  is the number of frames to average over. A 2-D map of the phase error can be generated by omitting the mean over all pixels  $\langle \dots \rangle$ . An example 2-D phase error map is shown for  $\Delta\phi_{e-o}(m = 150)$  in Figure 4.15. As the sub-series are composed of completely different data,  $\Delta\phi_{e-o}$  measures the repeatability of the phase error and should be considered more robust than pure precision estimates such as Eqn. 4.2.

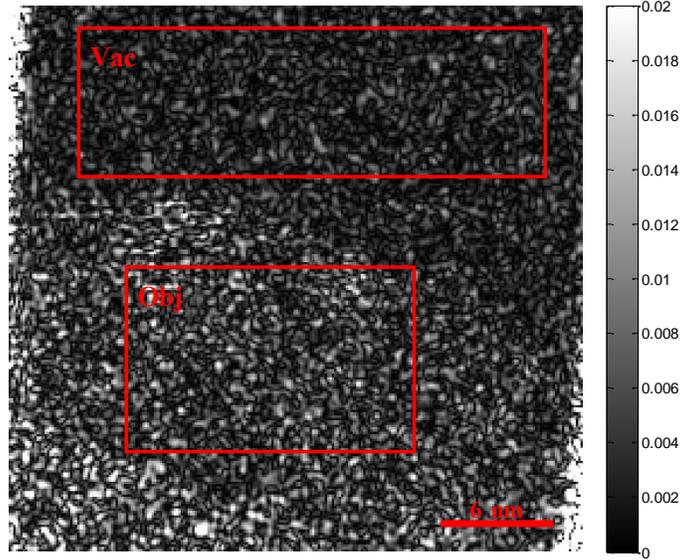


Figure 4.15: 2-D mapping of the RMS phase difference between even and odd-frame series. It can be observed that phase error scales with object visibility from the grayscale values. The areas chosen to calculate vacuum and object phase error are indicated with the respectively labeled boxes.

The convergence of the phase error is shown in Figure 4.16. The mean phase error between individual holograms without summation (red) is 0.056 rad ( $2\pi/110$ ). This provides a baseline for the improvement offered by summation. The averaged data without the benefit of image or drift registration (green) shows improvement over the first 50 frames but then diverges. The phase error is plotted for the registered series in both the object (purple) and vacuum (black) regions (as indicated in Figure 4.15). The associated analytical prediction for object phase error (Eqn. 4.23, blue dot-dashed), and the analytic vacuum phase error (Eqn. 4.22 with object terms omitted, gray dot-dashed) both agree well with the experimental metric. The mean phase error after summation inside the object is 0.0060 rad ( $2\pi/1050$ ). The mean phase error after summation in the vacuum region adjacent to the specimen is 0.0037 rad ( $2\pi/1700$ ). The difference between the analytical prediction and experimental result is +6.9 % for the object and -4.5 % for the vacuum. The object phase error improved by 930 % compared to the individual hologram average; the theoretical improvement for a 150-hologram series is  $\sqrt{150} = 1220$  %.

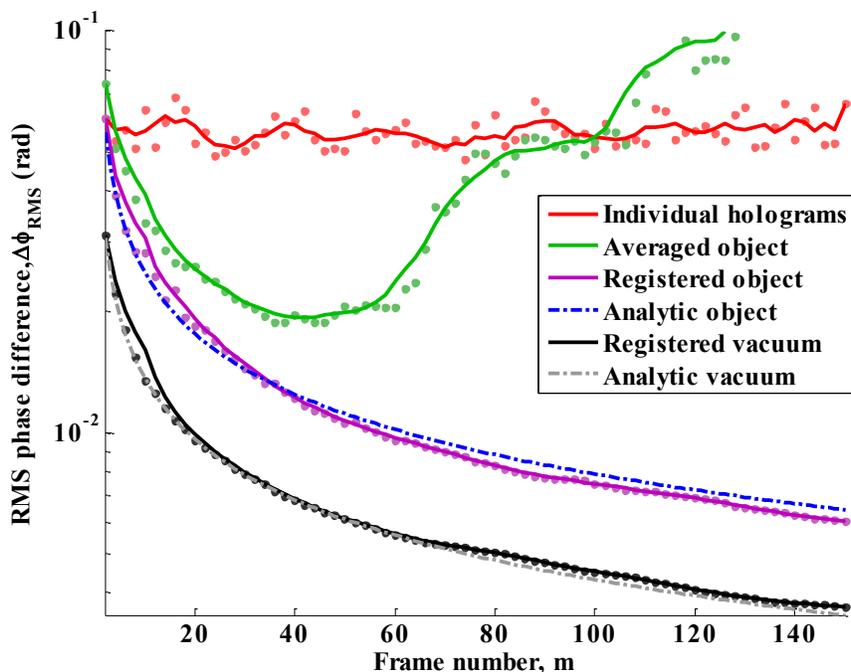


Figure 4.16: Phase error as a function of frame number, with phase error indicated by dots and a 10-frame moving average by solid lines. Lower values of RMS phase difference reflect the ability to more precisely detect smaller variations in the phase shift.

#### 4.4.3 RADIATION DAMAGE CONCERNS

The cumulative radiation dose is considerable, being  $2200 e^-/\text{\AA}^2$  or  $3.5 C/cm^2$  per frame ( $325\,000 e^-/\text{\AA}^2$  or  $520 C/cm^2$  total in the series) from the object wave. However, as can be seen from Figure 4.16, the scaling with  $\sqrt{m}$  limits the value of very high frame counts. While the specimen may suffer considerable damage during acquisition of the hologram series, it is possible to cut-off the radiation damage effects in the data by discarding all frames past an arbitrary dose limit. Last, if phase resolution per unit dose is the target goal, the illumination should be further spread to enhance the visibility at the cost of reduced dose rate.

## 4.5 CONCLUSIONS

Electron holography is a TEM technique that has long held promise to conduct novel characterization of materials. However, historically the relatively high phase error has limited electron holography to within the constraints of shot noise and holographic visibility. Therefore, applications requiring high sensitivity have been inaccessible. Here we demonstrate that it is practical to achieve an order of magnitude improvement in the precision of the phase error by the method of summation of hologram series. The summation of hologram series to achieve improved phase error was demonstrated experimentally on  $WS_2$  nanofullerenes. The experimental phase error inside the  $WS_2$  nanofullerene object was found to be  $0.0060$  rad ( $2\pi/1050$ ) after summation of 150 holograms. In comparison, individual

holograms were found to have an object phase error of 0.056 rad ( $2\pi/110$ ). The RMS phase error metric developed in section 4.4.2 is, to our knowledge, the first method for experimental verification of phase error in electron holography.

The analytical prediction for the phase error is in quantitative agreement with the analytical prediction derived in section 4.3.4, with a difference of 7 % after summation of 150 holograms. Thus it is practical to estimate the phase error in electron holograms for a given carrier frequency and exposure time if the instability of the TEM is adequately characterized. The only source of *a priori* information used for the analysis is the loss of visibility due to incoherent scattering within the specimen, which is a characteristic of the specimen.

The order of magnitude improvement in phase error between the sum and individual holograms (930 %) has important consequences for the scale of physical phenomena which may be investigated by electron holography. For example, Lichte calculates that a phase shift of  $2\pi/50$  is produced by the magnetization of a cluster of 420 Cobalt unit cells (or 1680 Co atoms) [13]. A linear improvement in phase error results in a corresponding reduction in the cluster size within which magnetization observed. As such, at a  $SNR = 2$ , the hologram series results presented here are equivalent to resolving about 40 unit cells (or 160 Co atoms). This is insufficient to detect a single Bohr magneton, but is of practical interest for the study of magnetism in small particles or grains [54][55]. The improvement in phase error is achieved without a corresponding loss of spatial resolution, due to the success of the experimental technique.

The results presented for high-resolution hologram series on  $WS_2$  do not represent the ultimate limits of our instrument, a Hitachi HF-3300 with a cold field-emission gun but no aberration correction. Instruments equipped with the latest generation of higher-brightness cold field-emission gun, aberration correction with isoplanar elliptical illumination, and direct detection devices will see further improved performance.

Adapted techniques for optimization of the electro-optical configuration were presented for the double biprism electron holography configuration. The analytical estimate for phase error has been improved by modeling the instabilities of the TEM. The presence of these instabilities has long been known but not previously characterized in a systemic way that permits predictions. A new model for drift in the TEM has been introduced that includes a random-walk component. It was shown experimentally that the random-walk component of the drift dominates over the gradient-driven linear velocity component for both the image drift and wavefront phase drift in a Hitachi HF-3300. The provided model for drift due to Gaussian random walk and the associated MTF is potentially valuable for other microscopy modalities where long exposure times are common, such as core-loss EELS. The image drift MTF also will affect the information limit of a TEM and is especially significant for corrected TEMs with 100 – 50 pm spatial resolution. The

results show that the double biprism configuration is very stable and shows minimal loss of visibility with exposure time.

#### 4.6 REFERENCES

- [1] E. Voelkl, L.F. Allard, D.C. Joy, Introduction to Electron Holography, in, Kluwer Academic, New York, 1999.
- [2] P.A. Midgley, An introduction to off-axis electron holography, *Micron*, 32 (2001) 167-184.
- [3] J.E. Bonevich, G. Pozzi, A. Tonomura, Electron Holography of Electromagnetic Fields, in: E. Voelkl, L.F. Allard, D.C. Joy (Eds.) Introduction to Electron Holography, Kluwer Academic, New York, 1999.
- [4] Q. Ru, J. Endo, T. Tanji, A. Tonomura, Phase-Shifting Electron Holography by Beam Tilting, *Applied Physics Letters*, 59 (1991) 2372-2374.
- [5] Q. Ru, G. Lai, K. Aoyama, J. Endo, A. Tonomura, Principle and Application of Phase-Shifting Electron Holography, *Ultramicroscopy*, 55 (1994) 209-220.
- [6] K. Yamamoto, I. Kawajiri, T. Tanji, M. Hibino, T. Hirayama, High precision phase-shifting electron holography, *J. Electron Microsc.*, 49 (2000) 31-39.
- [7] E. Voelkl, D. Tang, Approaching routine phase resolution for off-axis type holography, *Ultramicroscopy*, 110 (2010) 447-459.
- [8] R. McLeod, M. Bergen, M. Malac, Technique for Complex Averaging of Electron Holograms, *Microscopy and Microanalysis*, 17 (2011) 918-919.
- [9] T. Suzuki, S. Aizawa, T. Tanigaki, K. Ota, T. Matsuda, A. Tonomura, Improvement of the accuracy of phase observation by modification of phase-shifting electron holography, *Ultramicroscopy*, 118 (2012) 21-25.
- [10] Y. Aharonov, D. Bohm, Significance of Electromagnetic Potentials in the Quantum Theory, *Phys. Rev.*, 115 (1959) 485-491.
- [11] T. Boyer, Semiclassical Explanation of the Matteuccia-Pozzi and Aharonov-Bohm Phase Shifts, *Foundations of Physics*, 32 (2002) 41-49.
- [12] A. Caprez, B. Barwick, H. Batelaan, Macroscopic test of the Aharonov-Bohm effect, *Phys. Rev. Lett.*, 99 (2007).
- [13] H. Lichte, M. Lehmann, Electron holography: basics and applications, *Reports on Progress in Physics*, 71 (2008) 016102.
- [14] H. Lichte, K.-H. Herrmann, F. Lenz, Electron noise in off-axis image plane holography, *Optik*, 77 (1987) 135-140.
- [15] F. Lenz, Statistics of phase and contrast determination in electron holograms, *Optik*, 79 (1988) 13-14.
- [16] W.J. de Ruijter, J.K. Weiss, Detection limits in quantitative off-axis electron holography, *Ultramicroscopy*, 50 (1993) 269-283.
- [17] A. Harscher, H. Lichte, Experimental study of amplitude and phase detection limits in electron holography, *Ultramicroscopy*, 64 (1996) 57-66.
- [18] R.A. McLeod, M. Malac, Determination of Localized Visibility in Off-axis Electron Holography, Submitted.
- [19] D.J. Smith, W.J. De Ruijter, J.K. Weiss, M.R. McCartney, Quantitative Electron Holography, in: E. Voelkl, L.F. Allard, D.C. Joy (Eds.) Introduction to Electron Holography, Kluwer Academic, New York, 1999.
- [20] M. Lehmann, H. Lichte, Tutorial on Off-Axis Electron Holography, *Microscopy and Microanalysis*, 8 (2002) 447-466.

- [21] D. Geiger, H. Lichte, M. Linck, M. Lehmann, Electron Holography with a Cs-Corrected Transmission Electron Microscope, *Microscopy and Microanalysis*, 14 (2008) 68-81.
- [22] R.A. McLeod, M. Malac, Characterization of detector modulation-transfer function with noise, edge, and holographic methods, *Ultramicroscopy*.
- [23] D. Cooper, R. Truche, P. Rivallin, J.-M. Hartmann, F. Laugier, F. Bertin, A. Chabli, J.-L. Rouviere, Medium resolution off-axis electron holography with millivolt sensitivity, *Applied Physics Letters*, 91 (2007) 143501-143503.
- [24] M. Linck, B. Freitag, S. Kujawa, M. Lehmann, T. Niermann, State of the art in atomic resolution off-axis electron holography, *Ultramicroscopy*, 116 (2012) 13-23.
- [25] K. Ishizuka, Optimized sampling schemes for off-axis holography, *Ultramicroscopy*, 52 (1993) 1-5.
- [26] E. Voelkl, Noise in off-axis type holograms including reconstruction and CCD camera parameters, *Ultramicroscopy*, 110 (2010) 199-210.
- [27] R. Shaw, Evaluating the efficient of imaging processes, *Reports on Progress in Physics*, 41 (1978) 1103.
- [28] F. Zernike, The concept of degree of coherence and its application to optical problems, *Physica V*, 8 (1938) 785-795.
- [29] J. Verbeeck, A. Beche, W. Van den Broek, A holographic method to measure the source size broadening in STEM, *Ultramicroscopy*, 120 (2012) 35-40.
- [30] M. Gajdardziska-Josifovska, A.H. Carim, Applications of Electron Holography, in: E. Voelkl, L.F. Allard, D.C. Joy (Eds.) *Introduction to Electron Holography*, Kluwer Academic, New York, 1999.
- [31] D.B. Langmuir, Theoretical Limitations of Cathode-Ray Tubes, *Proceedings of the Institute of Radio Engineers*, 25 (1937) 977-991.
- [32] K. Harada, A. Tonomura, Y. Togawa, T. Akashi, T. Matsuda, Double-biprism electron interferometry, *Applied Physics Letters*, 84 (2004) 3229-3231.
- [33] N.R. Goodman, Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction), *Annals of Mathematical Statistics*, 34 (1963) 152-177.
- [34] P.O. Amblard, M. Gaeta, J.L. Lacoume, Statistics for complex variables and signals Part I: Variables, *Signal Processing*, 53 (1996) 1-13.
- [35] D. Padfield, Masked Object Registration in the Fourier Domain, *Image Processing, IEEE Transactions on*, 21 (2011) 2706-2718.
- [36] M. Guizar-Sicairos, S.T. Thurman, J.R. Fienup, Efficient subpixel image registration algorithms, *Opt. Lett.*, 33 (2008) 156-158.
- [37] M. Elad, A. Feuer, Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images, *Image Processing, IEEE Transactions on*, 6 (1997) 1646-1658.
- [38] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, W.T. Freeman, Removing camera shake from a single photograph, *ACM Trans. Graph.*, 25 (2006) 787-794.
- [39] S.H. Lim, A. Silverstein, Estimation and Removal of Motion Blur by Capturing Two Images with Different Exposures, in, *HP Laboratories*, 2008.
- [40] M. Ben-Ezra, S.K. Nayar, Motion-based motion deblurring, *IEEE Trans. Pattern Anal. Mach. Intell.*, 26 (2004) 689-698.
- [41] O. Hadar, I. Dror, N.S. Kopeika, Image resolution limits resulting from mechanical vibrations. Part IV: real-time numerical calculation of optical

- transfer functions and experimental verification, *Optical Engineering*, 33 (1994) 566-578.
- [42] R.M.J. Bokel, J. Jansen, D. van Dyck, H.W. Zandbergen, The effect of mechanical vibration and drift on the reconstruction of exit waves from a through focus series of HREM images, *Ultramicroscopy*, 80 (1999) 255-269.
- [43] A. Einstein, The theory of the Brownian Motion, *Ann. Phys.-Berlin*, 19 (1906) 371-381.
- [44] G.E. Uhlenbeck, L.S. Ornstein, On the theory of the Brownian motion, *Phys. Rev.*, 36 (1930) 0823-0841.
- [45] A. Einstein, *Investigation on the Theory of the Brownian Movement*, Dover Publications, New York, 1956.
- [46] W. Jost, *Diffusion in solids, liquids, gases*, Academic Press, New York, 1960.
- [47] P.E. Batson, A. Reyes-Coronado, R.G. Barrera, A. Rivacoba, P.M. Echenique, J. Aizpurua, Nanoparticle movement: Plasmonic forces and physical constraints, *Ultramicroscopy*, 123 (2012) 50-58.
- [48] R. Tenne, L. Margulis, M. Genut, G. Hodes, Polyhedral and cylindrical structures of tungsten disulphide, *Nature*, 360 (1992) 444-446.
- [49] R.F. Egerton, R. McLeod, F. Wang, M. Malac, Basic questions related to electron-induced sputtering in the TEM, *Ultramicroscopy*, 110 (2010) 991-997.
- [50] M. Lehmann, Influence of the elliptical illumination on acquisition and correction of coherent aberrations in high-resolution electron holography, *Ultramicroscopy*, 100 (2004) 9-23.
- [51] M.A. Schofield, Y. Zhu, Fast phase unwrapping algorithm for interferometric applications, *Opt. Lett.*, 28 (2003) 1194-1196.
- [52] J.R. Fienup, Invariant error metrics for image reconstruction, *Appl. Opt.*, 36 (1997) 8352-8357.
- [53] M. Beleggia, M.A. Schofield, Y. Zhu, G. Pozzi, Quantitative domain wall width measurement with coherent electrons, *Journal of Magnetism and Magnetic Materials*, 310 (2007) 2696-2698.
- [54] M. Beleggia, T. Kasama, R.E. Dunin-Borkowski, The quantitative measurement of magnetic moments from phase images of nanoparticles and nanostructures-I. Fundamentals, *Ultramicroscopy*, 110 (2010) 425-432.

## 5 GENERAL DISCUSSION AND CONCLUSIONS

In this section we discuss how the novel results in chapters 2-4 inter-relate and the conclusions that can be drawn from the body of research in the thesis. The local visibility and detector characterization methods described in chapters 2 and 3 were developed in parallel and results from each influenced the development of the other. In particular, the holographic MTF was dependant on the Fourier-ratio method for measurement of holographic visibility. The holographic MTF at each accelerating voltage was composed of several hundred holograms. Measurement of the visibility by hand would have been onerous and fraught with the potential for human error.

In turn, difficulties encountered in measurement of the holographic MTF illustrated short-comings of the visibility metrics. For example, the initial results for the holographic MTF were clearly biased due to uneven dose at the different spatial frequencies sampled. The holographic MTF method relies on changing the magnification by changing the excitation of the projector lenses. Beam current density decreases with increasing magnification, so this lead to the high-spatial frequency (low-magnification) holograms having much higher dose than those at low frequency. As a result, low-spatial frequency results were biased and the MTF in that regime was overestimated. Two corrections were employed to remove the bias: first, empirical corrections for the bias in the visibility due to shot noise were developed as discussed in section 2.4. Second, the exposure time for each spatial frequency was adjusted (via the Matlab control system) to try and collect uniform dose so that the bias due to shot noise would be weak.

Both the local visibility and the holographic MTF were required to achieve quantitative agreement between the experimental and analytic phase error estimates for the method of hologram series. The ability to measure contrast accurately is invaluable for determining the numerous OTFs that affect TEM micrographs. Both the detector MTF and DQE are factors in predicting the phase error, as discussed in chapter 4. The high-quality dark and gain reference developed in chapter 3 were invaluable for suppressing detector artifacts that can break the registration of image drift by cross-correlation. As a result reference holograms do not show the build-up of the honeycomb fibre-optical coupling seen in other publications [1].

Similarly, the solitary piece of *a priori* information included in the analytical prediction for phase error in hologram series is the loss of visibility caused by incoherent scattering in the specimen. The object wave visibility changes

significantly across the field of view, and rapidly in the presence of lattice fringes. Therefore, it is not clear how the historical Michelson min-max metric can be used to robustly and reliably measure the visibility inside the specimen. The Fourier-ratio visibility metric bypasses the problem by measurement of the phase error at any position in the holographic reconstruction. Essentially, by localizing the visibility to 2-D, the phase error can also be localized to 2-D.

In conclusion, substantial progress on quantitative phase imaging in the form of off-axis electron holography has been made. The methods presented herein should provide opportunities to conduct novel electron holography measurements in the future. The understanding of a number of effects, such as coherence of the electron wavefront and the nature of drift in the transmission electron microscope, have been improved.

## 5.1 FUTURE WORK

The most obvious research direction to take in the future is to apply both the local visibility and hologram series techniques to traditional electron holography applications. For the case of local visibility, it is not clear where the differences between the amplitude and coherence (visibility) signals originate. At present there is not a strong intuitive understanding of the difference between amplitude and visibility. Future work should examine the correlation among the amplitude, visibility, and phase shift for a variety of material compositions and morphologies.

The visibility is also invaluable for quantitatively assessing loss of coherence. In the future, the local visibility can be used to further understanding of the OTF hierarchy, both of instrument parameters and the nature of incoherent scattering in materials. For example, the local visibility can establish the loss of coherence due to relatively high-frequency oscillations. Examples include the 60 Hz from electrical mains and vibrations from turbomolecular pumps.

High phase error has been historically the limiting factor in electron holography [2][3]. As discussed in the conclusion of chapter 4, the smallest magnetization that could previously be resolved was a cluster on the order of  $\sim 10$  nm in diameter (1700 atoms). The limitation is not from the spatial resolution of the Lorentz lens, but rather the weak magnetic phase shift requires averaging over a large volume. Lichte provides the following estimate for magnetic phase shift,

$$\phi_m = n_{atom} n_B \mu_o \frac{e^2}{2m_o a} \quad (5.1)$$

where  $n_{atom}$  is the number of atoms per unit cell (= 4 for Co),  $n_B$  the number of Bohr magnetons (= 1.7 for Co),  $\mu_o$  the permeability of free space,  $e$  is the unit charge,  $m_o$  is electron rest mass, and  $a$  is the interatomic distance (= 0.4 nm for Co). With  $2\pi/1000$  phase error and a signal-to-noise ratio (SNR) of 2.0, the associated cluster size is  $\sim 1.7$  nm (80 atoms) which is less than the typical spatial resolution of

uncorrected Lorentz imaging. The low phase error achieved by summation of hologram series makes investigation of weak electric and magnetic potentials possible.

The phase error floor can be compared to how close an electron travelling through vacuum would have to pass to a dipole to see a significant phase shift. An electric dipole, with the dipole moment oriented orthogonally to the electron trajectory, has a projected potential,

$$\phi = C_E U_{proj} = C_E k_c q \int_{-\infty}^{\infty} \frac{1}{\sqrt{(x - d/2)^2 + y^2 + z^2}} - \frac{1}{\sqrt{(x + d/2)^2 + y^2 + z^2}} dz \quad (5.2)$$

where  $C$  is the electrostatic interaction constant ( $C_E = 6.66 \cdot 10^6 \text{ Cm}^{-1}\text{J}^{-1}$  at 200 keV),  $k_c$  is Coulomb's constant,  $q$  is the screened dipole charge, and  $d$  is the distance between the two charges. Then at an SNR of 2.0, the minimum projected potential that the hologram series could measure would be  $U_{proj} = 1.6 \text{ V nm}$ . A dipole with  $q = e/20$  and  $d = 0.4 \text{ nm}$ , which has a dipole moment of  $\sim 1 \text{ Debye}$ , can be integrated numerically over  $z$ . A rendered phase shift map is shown in Figure 5.1 with contour lines indicating a phase shift of  $2\pi/500$ , which represents the phase detection limit with  $SNR = 2$ . From this, the  $1 \text{ Debye}$  dipole can be observed  $1.3 \text{ nm}$  away from the core. As such, electron holography could potentially be used to investigate the dipole moments of quantum dots and other nanostructures [4]. For reference, the dipole moment of water is  $1.84 \text{ Debye}$ .

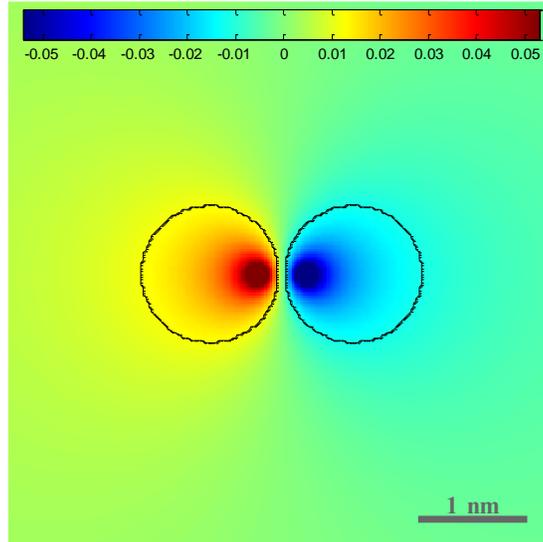


Figure 5.1: The projected phase shift for an electric dipole with  $q = e/20$  and  $d = 0.4 \text{ nm}$  for an accelerating voltage of 200 keV. The black contours indicate the approximate detect limit for the hologram series results in chapter 4. The colorbar gives the phase shift in radians.

A similar argument can be made for atomic compositional analysis based on phase. Based on published simulations for the phase-shift of single atoms [5] the phase shift of a single substitutional Phosphorous atom in a column of Silicon atoms is  $\sim 0.006$  rad ( $2\pi/1000$ ). Since that is equivalent to the noise floor of the results published in chapter 4, a single substitutional P atom cannot currently be measured. However, there is a path to achieving that goal by aberration correction in conjunction with elliptical illumination. Therefore the sensitivity of electron holography for dopant profiling in semiconductors is potentially perfect.

In practice, contamination, electron-beam induced sputtering, and charging of the specimen by secondary electron emission are more likely to limit compositional mapping than precision of the measurement. With regards to phase imaging, it is necessary to expand the scope of what is considered 'radiation damage.' For example, the bee-swarm effect is the random phase fluctuations caused by long-lived holes generated by secondary electron emission [6]. Accumulation of charge on the specimen or on imperfections in the biprism wire (i.e. 'dirt') can cause changes in the mean phase shift of an area over a hologram series. Reconfiguration of the crystal lattice, such as crystallites flipping back and forth between FCC and BCC lattices, is another potential concern. We have also in the past observed 'weather veining' where amorphous carbon under extended illumination begins to form graphite layers parallel to the electron beam.

Another concern for summation of hologram series is management of the reference hologram series. In conventional electron holography, the reference hologram is recorded immediately after the object hologram, so correspondence of the projector wavefront phase is assured. Hologram series take many minute to acquire, so one-to-one correspondence is a greater concern after many minutes of biprism drift. At lower magnifications than used in chapter 4, distortion of the projector lenses should be more significant. In that case, modeling the wavefront drift as a scalar may not be strictly valid. There is also concern for 'broken symmetry' caused by debris on the biprism wire resulting in wavefront drift being non-scalar in nature. Displacement over time of the envelope function caused by the biprism shadow is also a potential concern.

Furthermore, with a reference hologram, the biprism is incoherently illuminated. Therefore the reference hologram normalization only removes the wavefront aberrations electro-optical components below the biprism plane. The most significant electro-optical aberrations, from the objective lens contrast-transfer function, are not characterized. It is known that aberration correctors suffer from wavefront drive over time [7]. A superior methodology for reference subtraction should attempt to use the information redundancy present in hologram series to track the evolution of the wavefront shift over the series.

## 5.2 REFERENCES

- [1] E. Voelkl, D. Tang, Approaching routine phase resolution for off-axis type holography, *Ultramicroscopy*, 110 (2010) 447-459.
- [2] H. Lichte, P. Formanek, A. Lenk, M. Linck, C. Matzeck, M. Lehmann, P. Simon, Electron Holography: Applications to Materials Questions, *Annual Review of Materials Research*, 37 (2007) 539-588.
- [3] M.R. McCartney, D.J. Smith, Electron holography: Phase imaging with nanometer resolution, in: *Annual Review of Materials Research*, 2007, pp. 729-767.
- [4] R.J. Warburton, C. Schulhauser, D. Haft, C. Schaflein, K. Karrai, J.M. Garcia, W. Schoenfeld, P.M. Petroff, Giant permanent dipole moments of excitons in semiconductor nanostructures, *Phys. Rev. B*, 65 (2002).
- [5] M. Linck, H. Lichte, M. Lehmann, Off-axis electron holography: Materials analysis at atomic resolution, *Int. J. Mater. Res.*, 97 (2006) 890-898.
- [6] D.B. Dove, Image Contrasts in Thin Carbon Films Observed by Shadow Electron Microscopy, *Journal of Applied Physics*, 35 (1964) 1652.
- [7] D. Geiger, H. Lichte, M. Linck, M. Lehmann, Electron Holography with a Cs-Corrected Transmission Electron Microscope, *Microscopy and Microanalysis*, 14 (2008) 68-81.

# A1 APPENDIX 1 – SUPPLEMENTARY MATERIAL FOR CHAPTER 3

## A1.1 HIGH FIDELITY DARK REFERENCE

We examine the relationship between detector temperature, measured from the Gatan First Light™ controller unit. Temperature was varied by switching the Pelletier cooler off while an automated script took dark references at specified intervals. Figure A1.1 shows the dark current mean and standard deviation. As can be seen, the relationship is exponential with temperature and the detector is maintained at a temperature where the slope is low. Consequently, we observe excellent dark current stability.

As described in the manuscript, the dark reference is best composed from a sum of many dark references that have been individually filtered for ‘zingers’ from cosmic rays. An example Digital Micrograph script for overnight collection of dark references is provided in section S1.2. Ten-thousand dark references is generally a sufficient number to remove all correlated noise from the reference. Example of a 1000 frame dark reference is provided in Figure A1.2. Example of a 120 000 frame dark reference is provided in Figure A1.3.

The long-term stability of the dark reference sum is very good, as shown in Figure A1.4, with hot pixels gradually appearing over the lifetime of the detector. We recommend acquisition of a > 10 000 frame dark reference once per year, which can typically be performed overnight while the instrument is not in use.

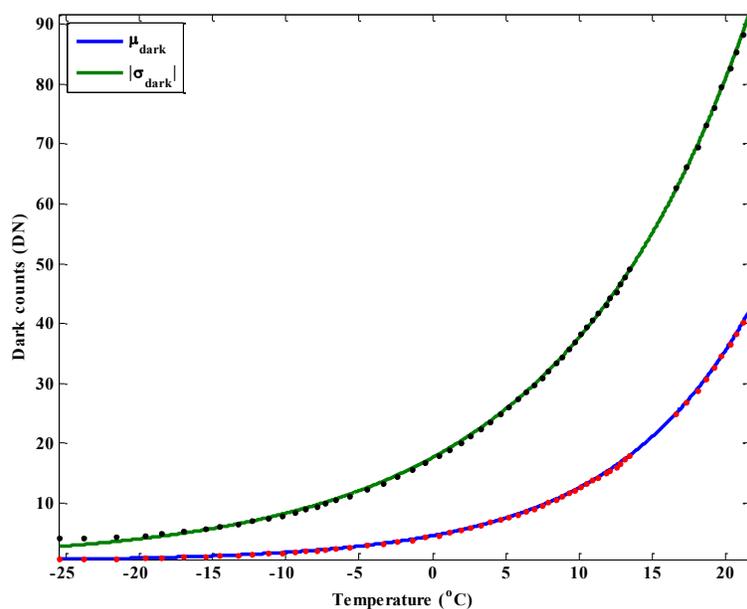


Figure A1.1: The dark current mean and standard deviation (dots) varies exponentially with temperature of the detector (best fits – solid lines).

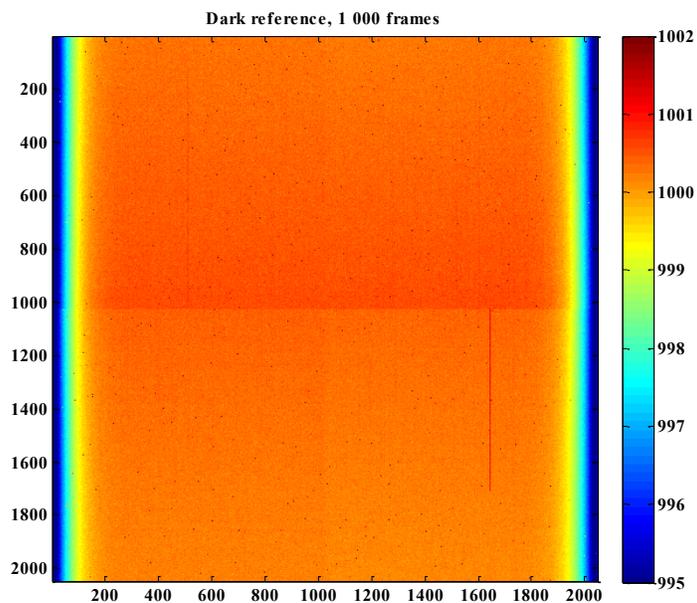


Figure A1.2: A dark reference composed of 1000 frames. Speckle noise is still apparent in the reference. Sidebar gives intensity scale in counts.

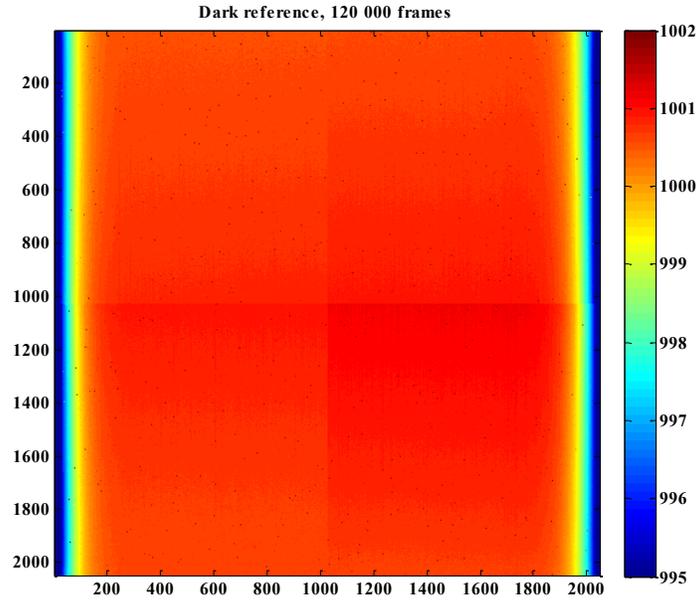


Figure A1.3: A dark reference composed of 120 000 frames. Compared to a typical single-frame reference, the amount of speckle noise is extremely small ( $\ll 1$  count). Note the roll-off on the horizontal axes of the detector. Detector read-out is performed in the vertical direction. Sidebar gives intensity scale in counts.

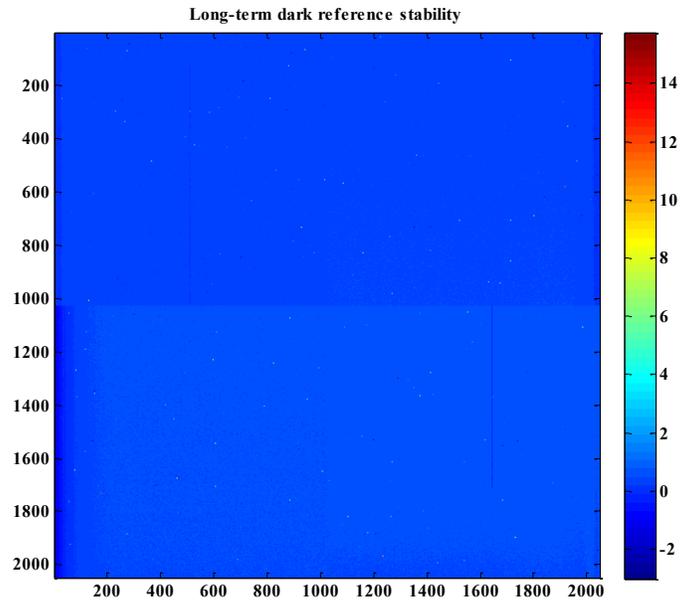


Figure A1.4: The difference between two 17 000 frame dark references, acquired on 11 September 2011 and 22 January 2013, or sixteen months apart. A scattered number of new hot pixels have appeared over 16 months of usage. As noted in the

manuscript, only 4084 pixels (0.1 % of the total) differ by more than one count while 531 pixels differ by more than 5 counts.

## A1.2 DIGITAL MICROGRAPH DARK REFERENCE SUM SCRIPT

Code is designed to save a copy of the dark reference every 1000 dark reference images. After the summed image has been saved to disk, it is necessary to normalize it by the exposure count, m. Note that the background level of the Gatan US1000 is 1000 counts. I.e.

```
m= 17000;
darksum = ((darksum - m*1000)/m) + 1000
```

Script as follows:

```
// Dark reference sum script
// 22 Dec 2010
// Robert A. McLeod
//
// Designed to be run over a weekend. Takes a huge series of images and sums
// them into a single image.

// In: input image
// MTyp: 3 for 2-d image, 0 for 1-D graphs
// MSize: Size of the median filter.
// Sigma: Multiplier for standard deviation. Standard deviation is a 68% confidence interval
// 1    0.6826895
// 2    0.9544997
// 3    0.9973002
// 4    0.9999366
// 5    0.9999994
// OPTcount: count the number of x-rays? 1 Yes, 0 No
// FOUND: number of x-rays found
image RemoveXrayFilter( image IN, number MTyp, number MSize, number Sigma, number
OPTcount, number &FOUND)
{
    image median = MedianFilter( in, MTyp, MSize)
    image diff = abs( in - median )
    number meandiff = Sigma * sqrt( variance(diff) )
    image out = tert( diff > meanDiff, median, in)
    If ( OPTcount ) FOUND = sum(!(out-in))
    return out
}

number imagecount, top, left, height, width, exposure
image darksum, temp
string imageName,path, imagetitle
number xraycount

top = 0
left = 0
height = 2048
width = 2048
exposure = 1
```

```

sigma = 10

if( ! GetDirectoryDialog(path) )
{
    // Failed, exit script
    Result( "No directory chosen, exiting script\n" )
    exit( 0 )
}

Result( "Hold SPACEBAR to halt execution\n" )
darksum = exprsize( height, width, 0.0 )
imagecount = 0;
while( 1 )
{
    temp = sscunprocessedbinnedacquire(exposure, 1, top, left, top+height, left+width)

    // Apply x-ray filter
    temp = RemoveXrayFilter( temp, 3, 1, sigma, 1, xraycount )

    darksum = darksum + temp
    temp.closeimage()

    imagecount = imagecount + 1
    //Result( "Summed image #" + imagecount + "\n" )
    if( getkey() == 32 ) // spacebar
    {
        break
    }
    // Save a copy every 1000 images
    if( mod(imagecount, 1000) == 0 )
    {
        Result( "X-ray count: " + xraycount );
        imagetitle = path + "DarkRefSum" + imagecount + ".dm3"
        saveimage( darksum, imagetitle )
    }
}
Result( "X-ray count: " + xraycount );
imagetitle = path + "DarkRefSum" + imagecount + ".dm3"
saveimage( darksum, imagetitle )

```

### A1.3 VOIGHT PROFILE BEST-FIT PARAMETERS

Table A1.1: Least-squares best-fit parameters and standard deviation for Gaussian and Lorentzian sum of Voight profile of Eqn. 3.2. Note that the 3<sup>rd</sup> term in the sum is generally not required for the NTF but is necessary for the MTF best-fits. Note that the NTF best-fits are not dealiased, but rather the actual fits to the recorded Fourier spectra.

Profile		Primary Electron Energy		
		100 keV	200 keV	300 keV
Lorentzian	a <sub>1</sub>	0.7924	0.7431	0.6099
	b <sub>1</sub>	0.2076	0.1225	0.1021

Gaussian	a <sub>2</sub>	0.1725	0.2180	0.3196
	b <sub>2</sub>	0.3226	0.2876	0.2596
Lorentzian	a <sub>3</sub>	0.0	0.0389	0.0705
	b <sub>3</sub>	1.0	10	10
<b>MTF<sub>edge</sub></b>		<b>100 keV</b>	<b>200 keV</b>	<b>300 keV</b>
Lorentzian	a <sub>1</sub>	0.2826	0.6449	0.3544
	b <sub>1</sub>	0.08896	0.06848	0.1023
Gaussian	a <sub>2</sub>	0.3451	0.2051	0.2048
	b <sub>2</sub>	0.2931	0.3001	0.2979
Gaussian	a <sub>3</sub>	0.3723	0.1500	0.4408
	b <sub>3</sub>	0.1549	0.09696	0.03771
<b>MTF<sub>holo</sub></b>		<b>100 keV</b>	<b>200 keV</b>	<b>300 keV</b>
Lorentzian	a <sub>1</sub>	0.1649	0.0476	0.2982
	b <sub>1</sub>	0.4433	1.0	0.1556
Gaussian	a <sub>2</sub>	0.6723	0.3594	0.2228
	b <sub>2</sub>	0.1905	0.2616	0.3019
Gaussian	a <sub>3</sub>	0.1628	0.5930	0.4789
	b <sub>3</sub>	0.0827	0.093	0.04803

## A2 APPENDIX 2 – SUPPLEMENTARY MATERIAL FOR CHAPTER 4

### A2.1 DESCRIPTION OF HOLOGRAM SERIES SUMMATION ALGORITHM

Here we discuss the algorithm, implemented in Matlab, used for the results presented in the manuscript. The algorithm consists of three basic code blocks: 1.) input and reconstruction of each individual hologram in the series; 2.) registration and correction of the image and phase drift; and 3.) plotting and movie display functions of the hologram series and summed reconstruction. Hologram series are acquired by automated TEM control via Matlab script over TCP/IP with the Maestro package. Holograms are acquired in ‘unprocessed’ mode which has the shortest read-out time and allows customized normalization routines to be applied [22].

In conventional off-axis holography, reference holograms are acquired immediately after an object hologram. For hologram series it is most practical to acquire all of the object holograms in series, and then acquire an identical number of reference holograms with the object and any associated long-range fields removed from the field of view. As the biprisms drift during the object exposure, the hologram envelope will be in a different position for the reference series compared to the object series. Therefore, we apply the heuristic strategy of separately summing the hologram series and the reference series, and then aligning the fringe fields prior to subtraction.

Hologram reconstruction is performed in a manner described in the literature [2][20]. Extra attention is put on pre-processing to reduce artifacts in the result, as discussed in the steps below. Artifacts can be problematic in the registration stage, producing anomalous cross-correlation and phase offset results. The reconstruction of every hologram sideband and centerband is stored in computer memory, which provides maximum flexibility. The reconstruction steps are as follows:

1. Import Gatan Digital Micrograph-format holograms, into Matlab.
2. Determine sideband position, for example from a representative reference hologram.
3. Apply high quality dark reference subtraction and gain reference normalization [22]. Due to the repeated normalization, there is the potential to introduce correlated, periodic noise artifacts which can affect the cross-correlation procedure. An ideal gain reference is approximately 300 frames recorded at similar dose to what is expected for an individual hologram.
4. Dampen zingers (cosmic rays) with a confidence interval median filter.
5. Reconstruct the sideband for each hologram in the series.
6. Extract the centerband for each hologram in the series.

As discussed in section 3 of the manuscript, the image and phase drifts over the course of the exposures so image and phase registration is required. Complex numbers are challenging to sum or average because small changes in phase can cause large changes along the real and imaginary axes. If the phase drift among all frames is estimated and subtracted, the real and imaginary components of each pixel can be summed (or averaged) arithmetically. Registration of hologram series consists of:

1. User selection of cross-correlation area. Display centerband of first frame and last frame side-by-side.
2. Cross-correlation to register image drift of each frame, relative to middle frame.
3. Shift and crop image to limits determined by maximum and minimum drift.
4. Calculate and subtract phase offset for each frame, relative to middle frame.
5. Sum hologram series.

Normalized cross-correlation is applied to a masked region, chosen by the operator, to estimate the image drift, typically with four-fold oversampling to give quarter-pixel precision [35]. Typically a corner of the object is the best region for the cross-correlation to operate on, as it provides an edge along both the horizontal and vertical axes to register. Care should be taken with lattice fringes inside the cross-correlation region. Digital apodization in the reconstruction process can alias the lattice fringes, which then acts as an artifact in the cross-correlation. Lowering the radius of the apodization function can remove high-frequency fringes.

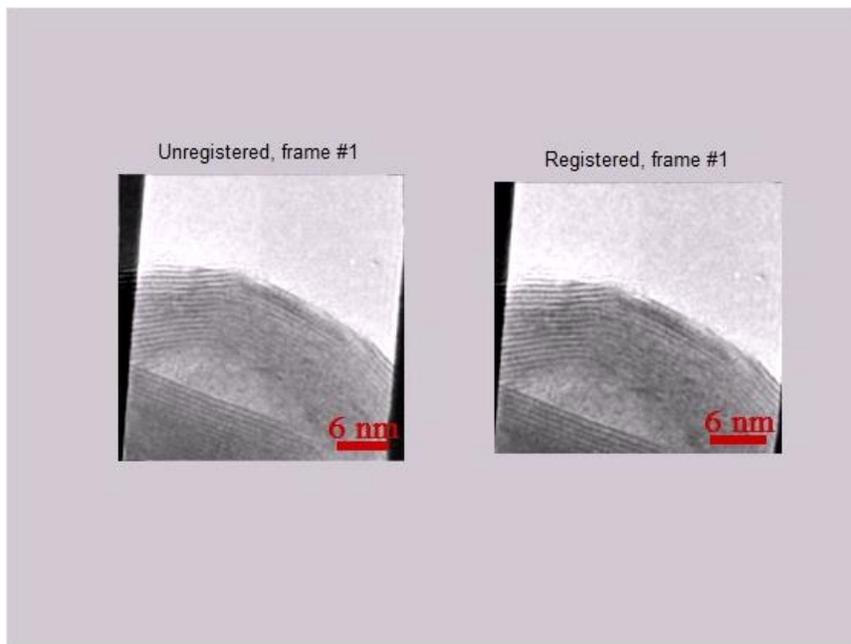
We have chosen to register the image drift by cross-correlation on the centerband (the autocorrelation) of the reconstructed hologram. Compared to the centerband, the sideband signal is lowered by the holographic visibility. Therefore the image registration error,  $\delta_i$ , and phase registration error,  $\delta_\phi$ , are lowest when the centerband is used. After the registration of each frame is completed the total drift from first to last frame in pixels is known, so each frame is shifted and cropped appropriately. The estimated subpixel shift is then applied to each frame via a cyclic phase shift in reciprocal space. Post image shift and crop, there is still a scalar phase offset amongst frames caused by wavefront phase drift. The wavefront phase is removed by a Fourier phase matching method [36].

Display of the results, particularly in movie format, is convenient to evaluate and validate that the hologram series was well behaved during the registration process. Periodic objects, such as lattice fringes, create local maxima in the cross-correlation. The local maxima in turn generate the potential for the registration to 'jump a period'. Fresnel fringes could also potentially cause registration difficulties, although in the double biprism configuration they are strongly suppressed. As such, the operator watches a movie of the unregistered centerband side-by-side with the registered and shifted centerband as it accumulates to check for anomalous jumps. The unregistered and registered phase can also be displayed to check the phase

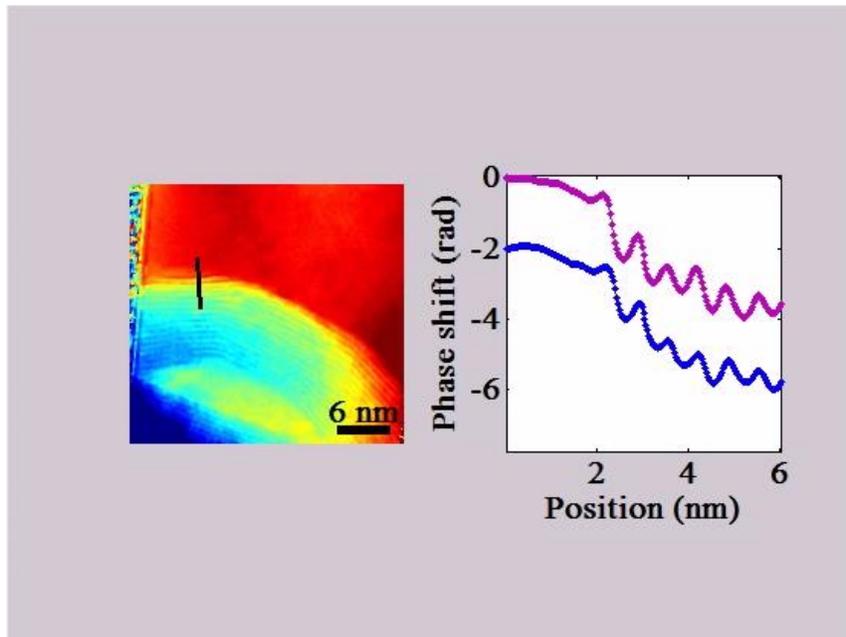
offset. When desired, the phase is unwrapped with a Fourier unwrapping technique [52]. The estimated image drift, phase drift, dose, and phase error metrics may also be plotted. Examples are shown in section 5 of the manuscript.

## A2.2 APPENDIX VIDEOS

*Author's note: video stills and captions are provided here. Videos are available through the provided links.*



Video A2.1: Verification of image drift registration is important to ensure accuracy of summation. The unregistered, unsummed centerband reconstruction (left) is shown alongside the registered, unsummed centerband (right). It is observed that the random-walk motion of the unregistered video is substantial, whereas there is no apparent residual drift after registration.



Video A2.2: The unwrapped, registered cumulative sum phase image (left) is shown alongside a selected cross-section of the WS<sub>2</sub> layers (right). The black line in the phase image shows the position and direction of the cross-section. The purple cross-section is that of the registered cumulative sum phase whereas the blue cross-section is the unregistered cumulative sum phase. The cross-sections origins have been offset for clarity. It is observed that there is no peak shortening or broadening in the registered phase (purple), whereas the unregistered phase flattens significantly (blue). Cross-sections were computed by bicubic interpolation.