Efficient Diffractive Phase Optics for Electrons

Tyler Harvey[1], Jordan Pierce[1], Amit Agrawal[2], Peter Ercius[3], Martin Linck[4] and Benjamin J. McMorran[1]

[1] Department of Physics, University of Oregon, Eugene, Oregon
[2] Department of Electrical Engineering, Syracuse University, Syracuse, New York
[3] National Center for Electron Microscopy, Lawrence Berkeley National Laboratory, Berkeley, California
[4] Corrected Electron Optical Systems GmbH, Englerstr. 28, D-69126 Heidelberg, Germany

E-mail: mcmorran@uoregon.edu

Abstract. Electron diffraction gratings can be used to imprint well-defined phase structure onto an electron beam. For example, diffraction gratings have been used to prepare electron beams with unique phase dislocations, such as electron vortex beams, which hold promise for the development of new imaging and spectroscopy techniques for the study of materials. However, beam intensity loss associated with absorption, scattering, and diffraction by a binary transmission grating drastically reduces the current in the beam, and thus the possible detected signal strength it may generate. Here we describe electron-transparent *phase* gratings that efficiently diffract transmitted electrons. These phase gratings produce electron beams with the high current necessary to generate detectable signal upon interaction with a material. The phase grating design detailed here allows for fabrication of much more complex grating structures with extremely fine features. The diffracted beams produced by these gratings are widely separated and carry the designed phase structure with high fidelity. In this work, we outline a fabrication method for high-efficiency electron diffraction gratings and present measurements of the performance of a set of simple prototypical gratings in a transmission electron microscope. We present a model for electron diffraction gratings that can be used to optimize the performance of diffractive electron optics. We also present several new holograms that utilize manipulation of the diffraction efficiency to produce new types of electron beams.

1. Introduction

Scanning transmission electron microscopy (STEM) has recently offered a large number of critical insights into the structure and behavior of materials at the atomic scale [1, 2, 3]. As a result of several decades of advancements in electron optics, modern STEM instruments use precisely controlled electric and magnetic fields to prepare Ångstrom-sized 60 to 300 keV electron probe beams with currents on the order of nano-Amperes [4, 5]. The focused probes have an approximately Gaussian intensity distribution and a flat phase profile.

Following the development of atomic-scale STEM, there has been a surge of interest in using shaped probe beams with carefully designed phase structure and non-Gaussian intensity distributions. The interaction of such a shaped probe beam with a material can offer more information than is available through the use of a traditional probe beam [6, 7]. For example, electron vortex beams [8, 9, 10] may soon be employed to accomplish atomic resolution spin imaging [9, 11] with STEM. Holographic diffraction gratings [9, 10, 12, 13] and phase plates – both material [8] and magnetic [14, 15] – have been successfully used to imprint the helical phase that defines a vortex beam.

Electron probe beams with well-defined phase structure can be produced via diffraction from a nanofabricated holographic diffraction grating placed in the probe-forming aperture of a STEM instrument (see Figure 1). A second aperture lower in the column can be used to isolate one of the diffracted probes and use it to form an image. One can, in general, define the surface shape $s(x, y)$ of a holographic diffraction grating by interference of a wavefunction of interest $\Psi(x, y)$ with a reference wave [16]:

$$s(x, y) = |\Psi(x, y) + \Psi_{\text{ref}}(x, y)|^2 \tag{1}$$

For example, to produce an electron vortex beam with $m\hbar$ orbital angular momentum in the first diffraction order of a linear grating periodic in x, one uses the shape

$$s(x, y) = \left| \frac{1}{\sqrt{2}} \left(e^{im\phi} + e^{ikx} \right) \right|^2 = 1 + \cos(m\phi + kx) \tag{2}$$

One can create such a diffraction grating from a material which only modulates the phase of the incoming wave by varying the thickness of the grating in proportion to the surface shape $s(x, y)$. We will return to the theory which guides the design of holographic diffraction gratings in Section 4.

Other methods exist for imprinting a spiral phase dislocation onto electron beams; both material phase plates [8] and magnetic nanowire phase plates [14, 15] add a spiral phase to an electron passed through them. However, production of pure spiral phase states with integer topological charge m is a significant challenge with these techniques. The topological charge imprinted depends on electron beam energy and is highly sensitive to fabrication errors and fringing fields, respectively. Diffractive electron optics offer the advantage that arbitrary phase structure can be imprinted on the electron beam with high fidelity and independent of beam energy. Furthermore, these devices can

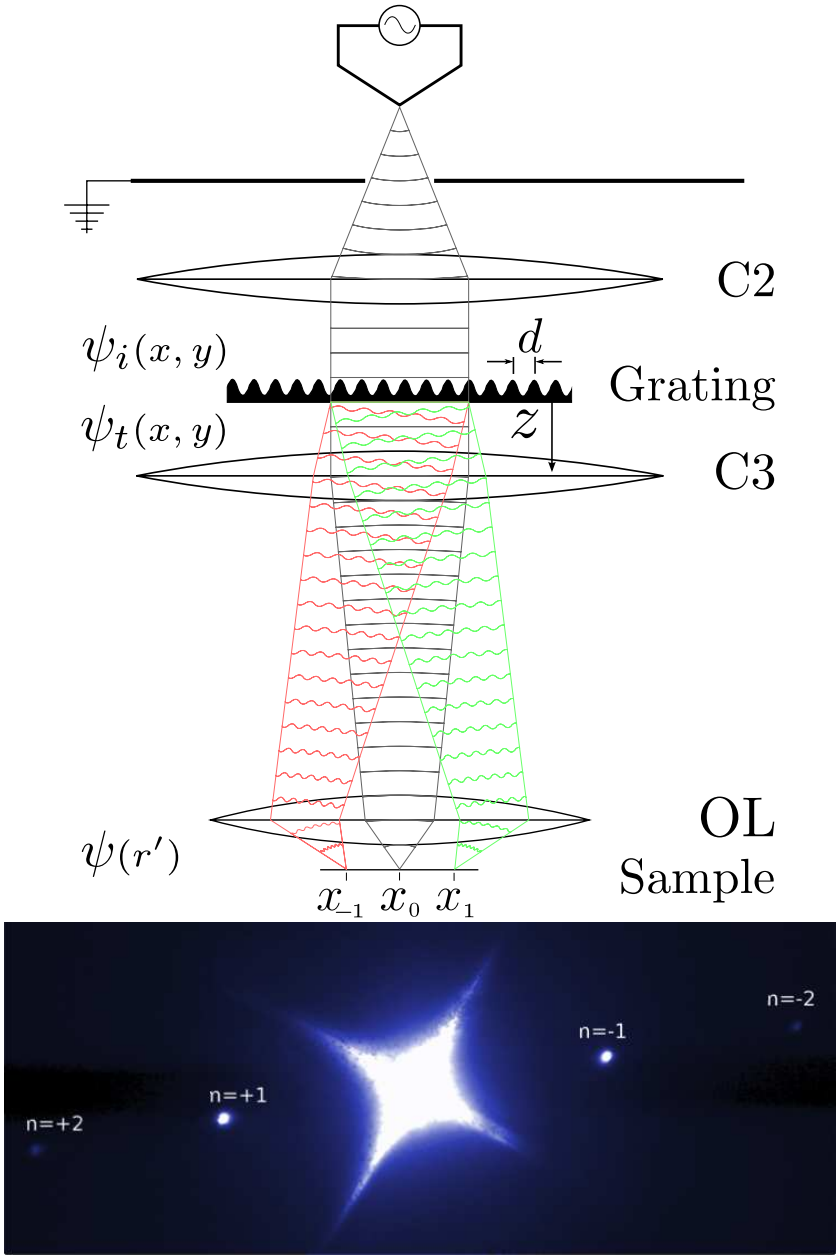


Figure 1. (Top image) Layout of the diffraction process in a STEM showing the grating in the condenser aperture producing separated STEM probes at the specimen plane. When placed at the condenser aperture plane, the patterned circular area of the transparent grating has the same beam-defining properties as a regular circular condenser lens aperture. The diffracted beams can either be isolated using a second aperture, or the entire set of beams can be scanned across a local sample feature, providing multiple images of the object each containing unique information. (Bottom image) TEM image of multiple diffracted STEM probes from a 50 μm-wide fork-dislocation grating at the specimen plane.

simultaneously produce multiple probe beams with complementary phase dislocations for dichroism techniques [17, 6].

There are two common challenges that must be addressed before diffractive electron

optics may be widely adopted for use in electron microscopy. First, diffraction gratings must produce sufficiently intense diffracted beams so that information carried by a diffracted probe is measurable. Second, the multiple diffracted probe beams must be sufficiently separated in the specimen plane for use in the study of materials with varied shapes and sizes. For example, several recent diffractive structures used to create electron vortex beams [9, 12, 13] are composed of patterns of slits milled all the way through a relatively thick foil that is opaque to electrons. These structures operate by selectively subtracting beam current through high-angle scattering in the material and transmitting the rest through the slits. Thus, these structures behave as binary amplitude gratings, which can only place a maximum of 10.1 % of the incident electrons into the first diffraction order. This is particularly troublesome for STEM applications, in which beam current must be maximized in order to generate a detectable signal of interest over background noise. These electron amplitude gratings furthermore have small separation between diffracted beams; it is thus difficult to isolate signal generated by a particular beam. An amplitude grating must be sufficiently thick so as to be opaque to electrons. This restricts the smallest lateral feature size that can be patterned, which consequently limits the diffraction angle. One recent demonstration of electron vortex production employed an amplitude diffraction grating which produced a beam separation of 3 μrad [9], or, equivalently, a real-space separation on the order of 5 nm in the specimen plane of a modern high-resolution STEM instrument.

To address these challenges, in our work developing electron vortex beams [18, 10] we have devoted considerable effort towards developing electron-transparent *phase* grating structures [19] that modulate the phase of the electron wave rather than the amplitude. Here we present a detailed study of these diffractive phase optics for electrons. We discuss the materials and nanofabrication method, performance measurements, and a model for these devices that incorporates the effects of both phase and amplitude modulation. Here we primarily discuss simple straight gratings periodic in one dimension, as shown in Figure 2, in order to elucidate the role of several basic grating structure parameters on diffraction efficiency. Optimization of the nanofabrication process for these simple gratings can then be applied to fabricate diffraction holograms which produce electron beams with non-trivial transverse wavefunctions $\Psi(x, y)$. In Figure 3, we demonstrate four gratings which produce diffracted beams with various different phase dislocations.

2. Electron Diffractive Phase Optics Design Goals

To be broadly useful for novel electron microscopy techniques, diffractive electron optics should efficiently diffract beam current into a desired diffraction order, and produce diffraction orders that are well-isolated from one another.

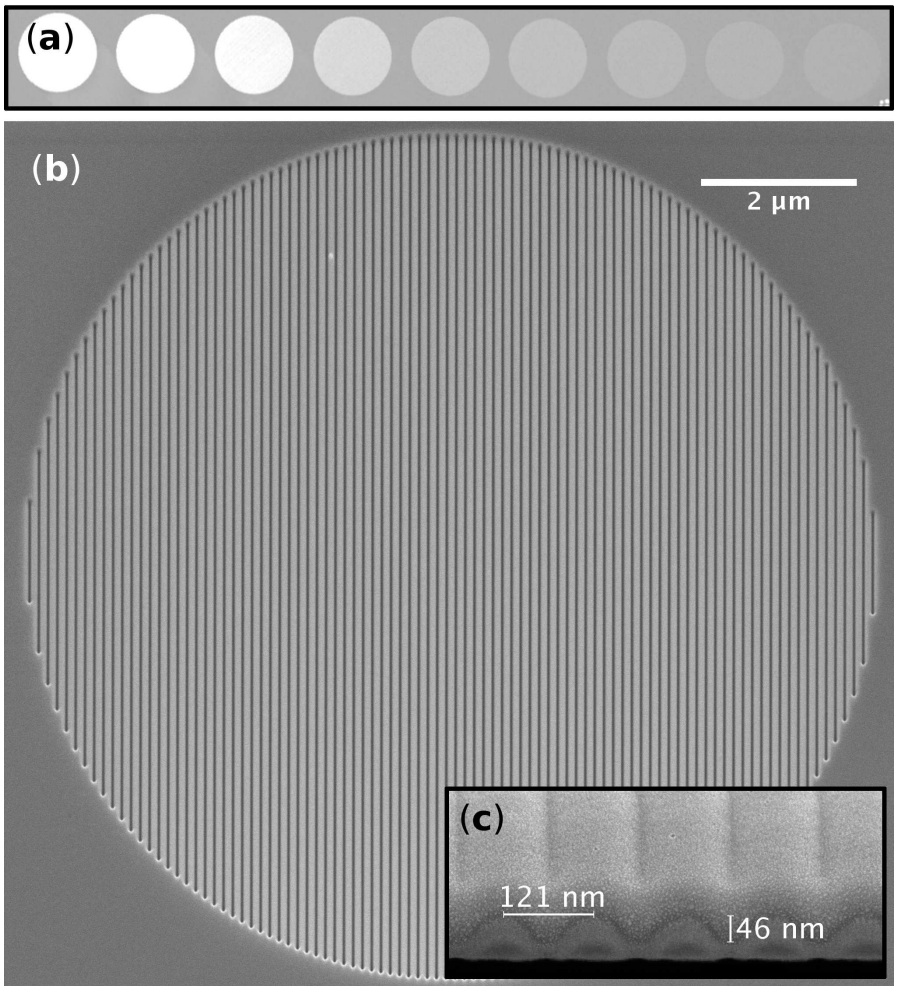


Figure 2. Uniform phase gratings for electrons, which serve as a simple prototype to measure performance. In a TEM operated at 300 keV these gratings are nearly transparent – all of these images were recorded in an SEM at 5 keV in order to show contrast. (a) An example of a FIB patterning dose array in which the ion beam dwell time was varied. (b) A smaller, 10 μm diameter grating with 120 nm pitch. (c) A perspective view of a FIB cross section of the phase grating shows the corrugated surface. Approximately 50 nm of Pt was deposited on the corrugated surface only so that the cross section could be prepared and imaged – this Pt coating is not normally included in phase gratings meant for TEM.

2.1. Diffraction Efficiency

In emerging STEM techniques using beams with engineered phase dislocations, it is important to maximize current in the probe beam. The efficiency of the diffraction hologram producing the beam thus becomes a primary concern. The diffraction hologram can be designed to generate just one electron probe beam, or simultaneously produce two or more probe beams with complementary phase dislocation properties. For example, a forked grating hologram with a symmetric thickness profile simultaneously produces pairs of electron probe beams with phase dislocations that are equal in magnitude but opposite in sign. It is expected that these two beams with opposite

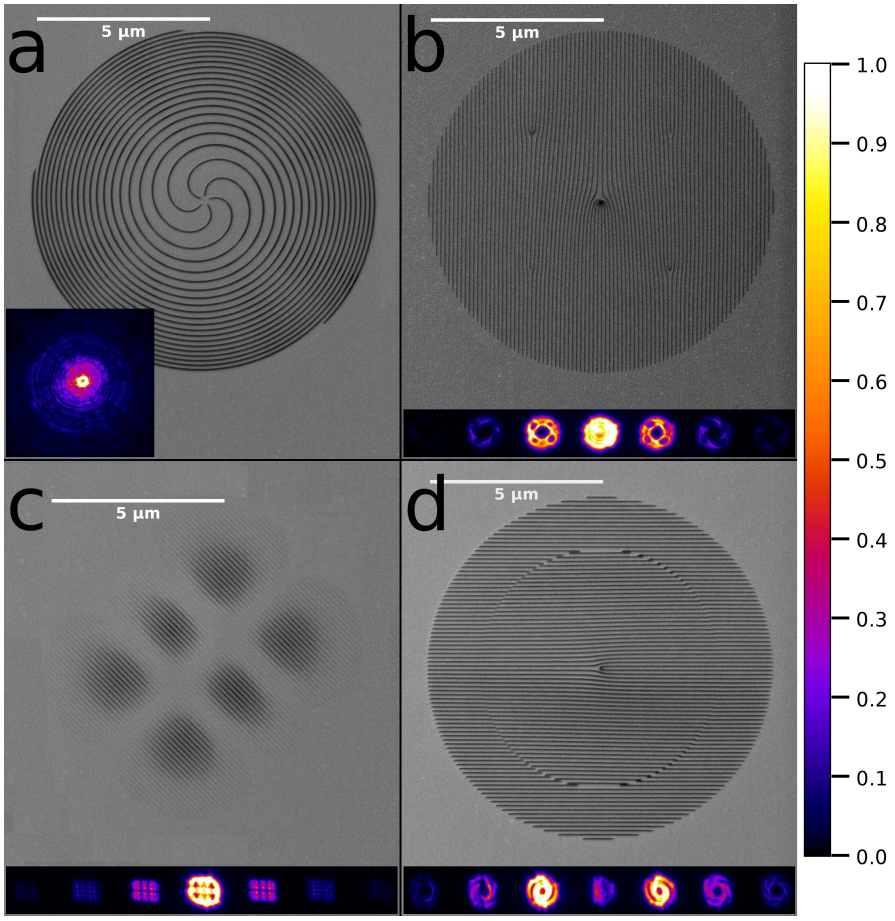


Figure 3. Examples of several phase grating designs that produce beams with non-trivial phase structure, together with the corresponding diffraction patterns. (a) A spiral phase dislocation grating produces beams with $L_z = 5n\hbar$ orbital angular momentum at a defocus $n\Delta f$, where n is the diffraction order number. The diffraction pattern is defocused so that the $n = +1$ order is in focus. (b) Multiple spatially separated fork dislocations produce diffracted beams with a net $\langle L_z \rangle = 0\hbar$ when in focus. (c) A Hermite Gaussian (1, 2) mode pattern. (d) An inner double fork dislocation grating and an outer zero-dislocation grating produces beams that are co-propagating superpositions of vortex beam and spherical wave. Note the suppressed 0-order intensity and enhanced ± 1-order intensities. In all of these examples, the beams have been defocused to reveal the details of their intensity and phase.

wavefront topologies can be used to provide dichroic image contrast [17, 6]. While in some applications it is desirable to produce only these two beams, in other instances it is desirable to also have a conventional electron probe beam with no phase dislocation to use for a reference signal. In all of these circumstances, the thickness profile of the hologram can be tailored to maximize the current in the electron beam and diffract it into one or more desired beams. It is also possible to form a blazed diffraction hologram that produces only one probe beam with a particular magnitude and/or sign of phase dislocation.

There are several important independent figures of merit for the diffraction

efficiency of a holographic grating. Absolute diffraction efficiency $\eta_n^{(i)}$, transmitted diffraction efficiency $\eta_n^{(t)}$, and relative diffraction efficiency $\eta_n^{(n')}$ are each useful for characterizing total detectable current, total detectable current as a fraction of the theoretical maximum for a given grating design, and detectable current relative to the noise produced by inelastic scattering in the grating, respectively.

The absolute diffraction efficiency is defined as

$$\eta_n^{(i)} \equiv \frac{I_n}{I_{\text{inc}}} \tag{3}$$

where I_n is the current in the n^{th} diffraction order, and I_{inc} is the total current incident on the diffraction grating. Potential electron microscopy applications require a maximal beam current in the $+1^{\text{st}}$ diffraction order, in which case $\eta_{+1}^{(i)}$ is the primary figure of merit. For a binary transmission grating, however, the maximum $\eta_{+1}^{(i)}$ is only 10.1 %. We demonstrate phase gratings that have twice this diffraction efficiency. When using blazed gratings, $\eta_{+1}^{(i)}$ can be made even larger at the expense of $\eta_{-1}^{(i)}$.

For material electron diffractive phase optics, some loss of electrons due to inelastic scattering is unavoidable. The total coherently transmitted beam current I_{trans}, equal to the sum of currents in all diffraction orders, is always less than the incident beam current, $I_{\text{trans}} = \sum_n I_n < I_{\text{inc}}$, and this is largely independent of the parameters of the particular groove shape. So to compare just the effects of groove shape and depth on diffraction efficiency independently from the effects of different thicknesses of material, it is then useful to define a *transmitted* diffraction efficiency describing the intensity in a particular diffraction order relative to the integrated intensity of all coherently transmitted electrons:

$$\eta_n^{(t)} \equiv \frac{I_n}{I_{\text{trans}}} \tag{4}$$

For STEM EELS with diffracted probes, quantitative analysis of an energy loss spectrum depends not just on the overall current in the beam, but also on the isolation of signal from a single diffracted probe. Adjacent diffraction orders may contribute to background signal. Thus, for such an application, the *relative* diffraction efficiency, defined as the ratio of intensity of the n^{th} order to the n'^{th} order, $\eta_n^{(n')}$, is a key figure of merit

$$\eta_n^{(n')} \equiv \frac{I_n}{I_{n'}}. \tag{5}$$

In most of the applications proposed to date for shape probe beams, the probe of interest is in the $n = \pm 1$ order, and the next most intense probe is the $n' = 0$ order. We will explicitly discuss the relative efficiency $\eta_1^{(0)}$ of gratings produced in this work.

Each of these figures of merit for efficiency represents a unique property of the STEM probes produced by a grating; absolute efficiency $\eta_n^{(i)}$ measures total detectable current in the n^{th} diffracted order, transmitted efficiency $\eta_n^{(t)}$ is a good measure of total detectable current as a function of theoretical maximum, and relative efficiency $\eta_n^{(n')}$ affects the signal-to-noise ratio for a measurement involving the n^{th} probe where noise from the n'^{th} is a concern. During review of this manuscript, Grillo et al. reported 25 %

efficiency for the first-order diffracted probe of their phase diffraction gratings [20]; we note that this was a *transmitted* efficiency. Using our model developed in Section 4, we estimate that the 120 nm thick excess silicon nitride material supporting the grating structure blocked or incoherently scattered at least 90 % of the incident intensity; after accounting for this intensity loss, the *absolute* efficiency could have been no larger than 2.5 %.

As we shall see in Section 4, the diffraction efficiency of a grating is a function of the shape and depth of grooves, the projected mean inner potential of the grating material, and the electron beam energy. The theoretical maximum efficiency of a sinusoidal pure phase grating is 33.9 % [21]. For electron sinusoidal phase gratings made of silicon nitride (Si_3N_4), our model predicts that a physical groove depth of about 33 nm achieves this maximum. Utilizing high resolution FIB milling we have consistently fabricated gratings with $\eta_1^{(i)} > 20$ %.

2.2. Diffraction Order Separation

A diffractive optical element in a STEM application produces in general multiple probe beams, and there must be sufficient angular separation between them such that the signal they each generate can be isolated. To meet this design goal and provide large free space diffraction angles, electron diffractive optical elements should be fabricated with as small a feature size as possible without sacrifing pattern fidelity.

For electrons of de Broglie wavelength λ transmitted through a grating-like diffraction hologram with pitch d, where typically $\lambda \ll d$, the angular separation $\Delta\theta$ between diffracted beams is

$$\Delta\theta = \frac{\lambda}{d}. \tag{6}$$

In the specimen plane of a STEM, the real-space physical separation Δx between diffracted probe spot centers at the specimen plane is

$$\Delta x = \frac{z\Delta\theta}{M} = \frac{z\lambda}{Md} \tag{7}$$

where M is the magnification of the lower probe-forming STEM optics (*not* the image magnification) and z is the physical distance between the diffraction hologram and the specimen plane. Alternatively, in terms of the effective camera length L of the lower probe-forming optics, the spot separation is

$$\Delta x = \frac{L\lambda}{d}. \tag{8}$$

In the TEAM 1 instrument at the National Center for Electron Microscopy, a grating with pitch $d = 83$ nm installed in the second condenser lens aperture produces diffracted 300 keV ($\lambda = 1.97$ pm) probe beams separated by $\Delta x = 43$ nm in the specimen plane (See Figure 1).

We note that it is easier to achieve larger angular separation between diffraction orders using diffractive phase optics. A distinct advantage of phase gratings over

amplitude gratings is that they can be fabricated with much finer feature sizes, and so can produce much wider separation between diffracted beams. Free-standing amplitude-type gratings must be sufficiently robust to support the mass of the relatively thick, electron-opaque material; this condition limits the minimum lateral feature size of such a structure. Amplitude-blocking diffractive optics demonstrated to date [9, 12, 13] have a minimum periodic feature size on the order of 1 μm. On the other hand, phase gratings can be much thinner than amplitude gratings, and can be fabricated on an electron-transparent supportive membrane [19]. In [10] we demonstrated phase gratings with 75 nm period, and have since fabricated gratings with periodic feature sizes down to 20 nm [22].

3. Nanofabricated Diffractive Electron Optics

To achieve our stated design goals for efficient electron diffractive optics, we explored a diverse array of nanofabrication techniques for imprinting the phase grating onto electron-transparent materials. High-quality gratings may be produced with high-resolution focused ion beam (FIB) milling, electron beam-induced deposition (EBID), and electron beam lithography (EBL); the choice of technique places some limitations on the structure of the gratings produced but is primarily a question of practical considerations. In this work, we consider prototypical electron diffractive optics produced by focused ion beam (FIB) milling, as FIB instruments are present in many TEM labs and can be used safely to produce an electron diffraction grating with minimal training. Some specific considerations for nanofabrication of gratings with FIB are detailed in the Appendix.

4. Phase Grating Model for Optimizing Diffraction Efficiency

To compare the measured diffraction efficiencies of our holographic phase gratings to theoretical limits, we developed a model for mixed phase/amplitude gratings for electrons. The structure of the grating is a thin membrane with a modulated thickness $h(x, y)$. Figure 4 shows the thickness patterns of several fabricated grating measured with atomic force microscopy (AFM). The transmission function describing the effects of this structure on transmitted electron wavefunctions is

$$t(x, y) = e^{-\alpha h(x,y)} \cdot e^{iCV_0 h(x,y)} = e^{i\tilde{k}h(x,y)} \tag{9}$$

where $\tilde{k} = CV_0 + i\alpha$ is the effective complex wavenumber of the electron within the material. The amplitude decay coefficient α describes the effects of inelastic and high angle scattering, and V_0 is the mean inner potential of the grating material. C depends only on the energy of the beam [23]. If such a grating is illuminated by an incident electron wave ψ_i, then the transmitted electron wavefunction immediately behind the grating is

$$\psi_t(x, y) = \psi_i(x, y)e^{i\tilde{k}h(x,y)}. \tag{10}$$

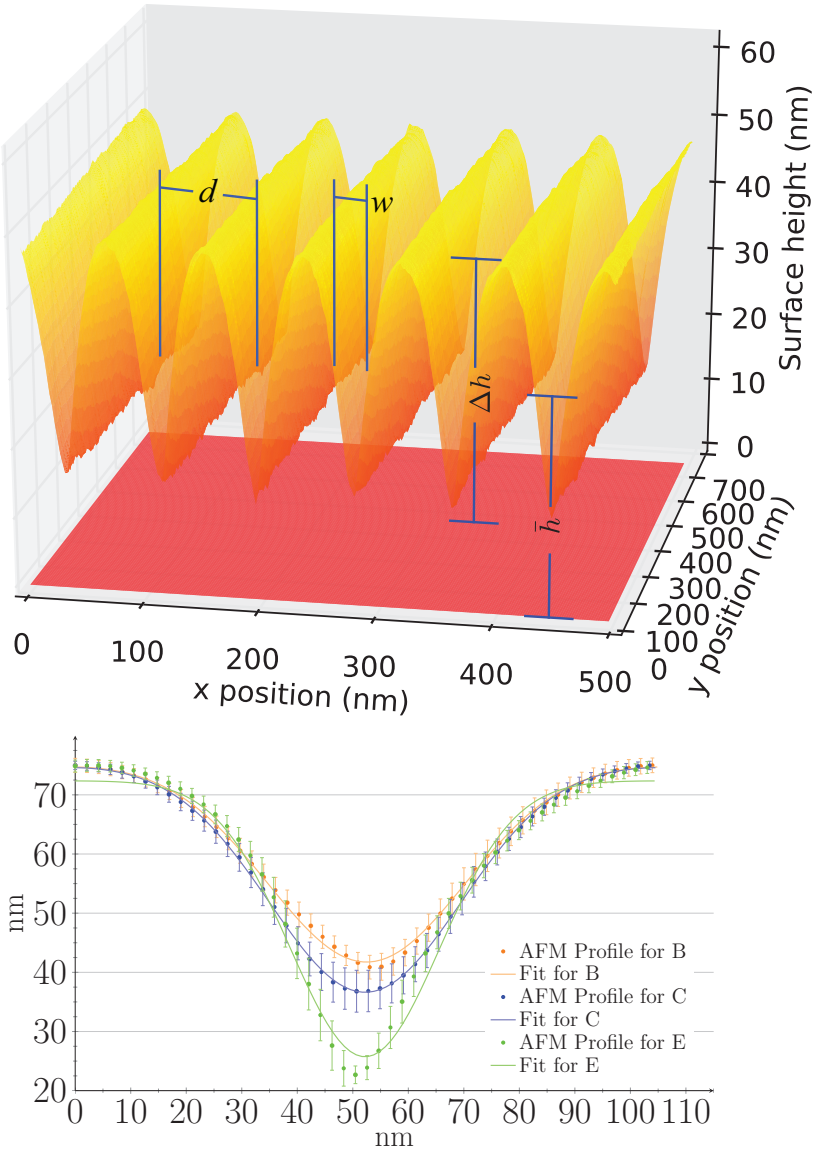


Figure 4. (top) Two-dimensional AFM surface profile of an 80 nm pitch grating with a trench depth of 30.6 nm. The grating was milled by focused ion beam on a 50 nm thick silicon nitride membrane. The average membrane thickness $\bar{h}$, groove depth Δh, groove spacing, or pitch, d, and relative groove width w are labeled for reference. (bottom) One-dimensional AFM profiles of several 100 nm pitch gratings. The error bars represent variation in the measured groove depth over the area of the grating. The solid lines represent a best fit Gaussian-shaped groove used in our model of phase gratings.

The far field diffracted wave can be computed analytically with the Fraunhofer approximation as shown in the Appendix (C.13). While the diffracted wave could also be computed numerically using an FFT, our analytical approximation can be evaluated much more quickly by a computer. This becomes a major consideration when evaluating

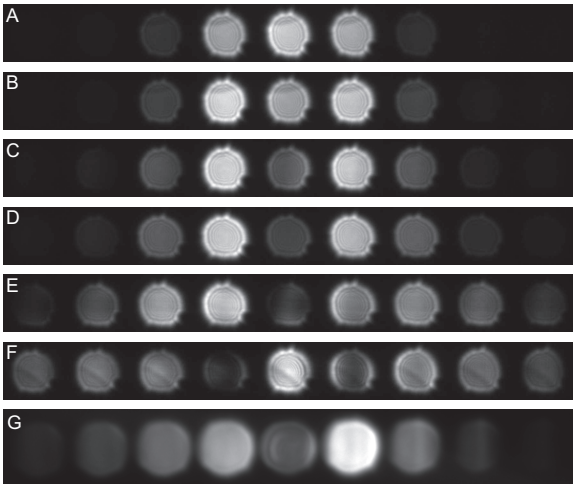


Figure 5. TEM diffraction images for gratings A - F. The diffracted beams are intentionally defocused in order to examine the uniform phase and amplitude in each beam. The grooves for each grating in the figure are successively deeper from top to bottom (associated AFM profiles are in Figure 4). (A-D) The 0^{th} order can be suppressed by deepening the grooves; here, $\eta_0^{(1)}$ can increase well beyond unity, and in fact reaches $\eta_0^{(1)} = 14$ here. (F) 0^{th} order intensity will increase again if the grooves are too deep. (G) Asymmetric grooves forming a blazed grating result in an asymmetric diffraction pattern, and a first-order diffracted beam with $\eta_1^{(t)} > 80$ %.

diffraction efficiency over a parameter space of grating groove geometries such as in Figure 6, which requires such a calculation for every pixel in the background.

In the case of simple linear phase gratings with uniform periodicity in only one direction, x, the modulated thickness of the material can be described as an array of grooves,

$$h(x,y) = \bar{h} + \sum_{n=-\infty}^{\infty} g(x - nd, y) - b \tag{11}$$

where $\bar{h}$ is the average thickness of the material grating, d is the pitch, or center-to-center groove spacing, and $g(x,y)$ describes the profile of the grooves. The offset b cancels any residual constant from the infinite sum. These parameters, as well as the groove depth Δh and width w, are shown in Figure 4. If the incoming wave ψ_i is a normally-incident plane wave (C.8), the diffracted wave amplitude from this linear grating may be written as

$$|\psi(\mathbf{r}')| = \frac{(2\pi)^2}{\lambda z' \sqrt{V}} e^{-\alpha \bar{h}} \sum_{n=-\infty}^{\infty} |c_n| \delta\left(\frac{kx'}{z'} - \frac{2\pi n}{d}\right) \delta\left(\frac{ky'}{z'}\right) \tag{12}$$

where λ is the de Broglie wavelength and V is a plane wave normalization factor. The relative amplitude of the nth diffraction order, c_n, depends upon the specific groove profile $g(x,y)$. A general form for computing c_n from any arbitrary periodic structure is described in detail in (C.22-C.24). The transmitted diffraction efficiency $\eta_n^{(t)}$ defined in (4) is calculated as

$$\eta_n^{(t)} = |c_n|^2. \tag{13}$$

Note that the absolute diffraction efficiency $\eta_n^{(i)}$, defined in (3) as a fraction of the *incident* beam, is less than (13) due to high-angle scattering

$$\eta_n^{(i)} = e^{-2\alpha\bar{h}}|c_n|^2. \tag{14}$$

where $\bar{h}$ is the average thickness of the grating over the back surface, as illustrated in Figure 4.

Cross sections (Figure 2) and AFM profiles (Figure 4) of the surface of our gratings indicate that individual grooves have a somewhat Gaussian shape. Earlier findings suggest that single-pixel-wide FIB-milled trenches have a remarkably Gaussian profile [24], so we model our gratings as a periodic sum of Gaussian grooves. The profile of an individual groove is

$$g(x) = \frac{A}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{15}$$

with amplitude A, characteristic width σ; the profile of the entire grating is described by the periodic sum in (11) with this grating shape $g(x)$. The groove depth Δh and groove width w are related to the parameters A, σ, and b in (C.35-C.36) and (C.39). Groove width is defined as the full width of the groove at half the maximum depth as a fraction of the pitch d.

With structure parameters describing the groove shape (depth Δh and width w) and material properties (absorption coefficient α and mean inner potential V_0), we can calculate the expected relative diffraction amplitudes c_n for a grating according to (C.22). We can thus reliably model the diffraction efficiency with a small number of measured parameters.

We paramaterized the average groove shape of every grating in all of our fabrication arrays using the Gaussian groove model to determine the width w and depth Δh. In Figure 6, we used these average groove dimensions to categorize each grating (location in figure) and express the measured electron diffraction efficiency as a color value. These measured values compare well to the theoretical diffraction efficiency calculated using our model (background color).

The variations in efficiency between gratings with nearly identical trench depth and width, seen most clearly in Figure 7, are primarily due to gallium implantation from the FIB. We have observed up to 10 atomic % gallium concentrations in milled areas of our diffraction gratings with energy-dispersive X-ray spectroscopy (EDX) composition analysis. We do not account for any modification of the mean inner potential V_0 of the material in the mill process. We expect that the lower-than-expected first-order transmitted efficiency $\eta_{\pm 1}^{(t)}$ and higher-than-expected zeroth-order transmitted efficiency $\eta_0^{(t)}$ observed for gratings with a trench depth $\Delta h = 20 \pm 5$ nm are also due to gallium implantation. However, as gallium concentration is not easy to measure, and as this variation in efficiency is specific to only small range of gratings fabricated with FIB and does not affect our prediction for peak efficiency, we have chosen not to include

	Predicted	Measured
$\eta_1^{(i)} \equiv I_1/I_{\text{inc}}$	0.136 ± 0.005	0.17 ± 0.04
$\eta_1^{(t)} \equiv I_1/I_{\text{tran}}$	0.33 ± 0.01	0.34 ± 0.07
$\eta_1^{(0)} \equiv I_1/I_0$	5 ± 3	3.8 ± 0.8

Table 1. Comparison of predicted and measured efficiencies for a grating with a width $w = 0.411 \pm 0.003$, trench depth $\Delta h = 35.7 \pm 0.5$ nm, and an average thickness $\bar{h} = 57.2 \pm 0.3$ nm.

extra parameters in our model to characterize this additional variation. We can use this simple model, which we have shown to predict efficiency with reasonable accuracy, to make prescriptions for the fabrication of highly efficient gratings.

The analysis summarized by Figure 6 provides a map that we used to explore a space of many interdependent parameters used for the nanofabrication process. We see in Figure 6 that any grating fabricated with a depth Δh between 27 nm and 40 nm and a full width at half max $w > 0.40$ will produce first-order diffracted beams with transmitted efficiency $\eta_1^{(t)} > 30$ %. The most efficient grating we fabricated indeed lies within this range. Table 1 offers a comparision of predicted and measured efficiencies for this grating. The predictions of the Gaussian groove model compare well with the diffraction efficiencies we measured from actual gratings.

This model provides a useful tool in the future design and fabrication of nanoscale electron phase gratings. For example, Figure 3 shows how the diffraction efficiency of the holograms can be manipulated to produce several different types of electron beams with engineered phase and intensity.

5. Conclusion

Here we described our studying and optimizing the efficiency of electron diffractive phase optics for use in high resolution electron microscopy. We described a process for fabricating spatially coherent nanoscale material phase gratings, and characterized the performance of simple prototypical phase gratings in a TEM. We developed a theoretical model for optimizing the design of the gratings for particular uses. We demonstrated diffractive phase gratings with 80 nm periodic feature sizes covering areas 50 μm in diameter, and demonstrated grating periods down to 20 nm. We demonstrated phase gratings that produce equally intense -1$^{\text{st}}$, 0$^{\text{th}}$, and +1$^{\text{st}}$ diffraction order beams with 25 % of the initial beam current each, gratings with suppressed 0$^{\text{th}}$ order beams (intensities in the 1$^{\text{st}}$ order up to 16 times greater) and blazed gratings that place up to 80 % of the incident beam current into the +1$^{\text{st}}$ diffraction order. When installed in a TEM operated at 300 keV, these gratings are capable of providing multiple Ångstrom-scale electron probes with precisely manufactured phase dislocations. We demonstrate example nano fabricated holograms in which this diffraction efficiency is controlled in order to efficiently convert an incident electron beam into various engineered electron

beams. This technology now makes possible shaped STEM probe techniques designed to produce more information about magnetically ordered, superconductor, semiconductor and biomolecular materials.

Acknowledgments

This work was supported by the Office of Science, U.S. Department of Energy under the Early Career Research Program Award No. DE-SC0010466 (B.M) and Contract No. DE-AC0205CH11231 (P.E. and M.L.). The authors wish to thank Jeff Ditto, Kurt Langworthy, and Josh Razink of the CAMCOR facility at University of Oregon, and Henri Lezec, Ian Anderson, and Jabez McClelland of NIST-Gaithersburg for their helpful discussions and equipment training for this study. The CAMCOR High-Resolution and Nanofabrication Facility (TEM, FIB and SEM) are supported by grants from the W.M Keck Foundation, the M.J. Murdock Charitable Trust, ONAMI, the Air Force Research Laboratory (agreement number FA8650-05-1-5041), NSF (award numbers 0923577, 0421086) and the University of Oregon.

6. References

[1] Nobuo Tanaka. Present status and future prospects of spherical aberration corrected TEM/STEM for study of nanomaterials. *Science and Technology of Advanced Materials*, 9(1):014111, 2008.

[2] Knut W. Urban. Studying atomic structures by aberration-corrected transmission electron microscopy. *Science*, 321(5888):506–510, 2008.

[3] S. J. Pennycook, M. F. Chisholm, A. R. Lupini, M. Varela, A. Y. Borisevich, M. P. Oxley, W. D. Luo, K. van Benthem, S.-H. Oh, D. L. Sales, S. I. Molina, J. Garca-Barriocanal, C. Leon, J. Santamara, S. N. Rashkeev, and S. T. Pantelides. Aberration-corrected scanning transmission electron microscopy: from atomic imaging and analysis to solving energy problems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1903):3709–3733, 2009.

[4] O.L. Krivanek, N. Dellby, and A.R. Lupini. Towards sub-a electron beams. *Ultramicroscopy*, 78:1–11, 1999.

[5] Heiko Müller, Stephan Uhlemann, Peter Hartel, and Maximilian Haider. Advancing the hexapole cs-corrector for the scanning transmission electron microscope. *Microscopy and Microanalysis*, 12(06):442–455, 2006.

[6] Benjamin McMorran. United states patent: 8680488 - system and method for producing and using multiple electron beams with quantized orbital angular momentum in an electron microscope, 2014.

[7] Jo Verbeeck, Giulio Guzzinati, Laura Clark, Roeland Juchtmans, Ruben Van Boxem, He Tian, Armand Bch, Axel Lubk, and Gustaaf Van Tendeloo. Shaping electron beams for the generation of innovative measurements in the (S)TEM. *Comptes Rendus Physique*, 15(23):190–199, February 2014.

[8] Masaya Uchida and Akira Tonomura. Generation of electron beams carrying orbital angular momentum. *Nature*, 464:737–739, 2010.

[9] J. Verbeeck, H. Tian, and P. Schattschneider. Production and application of electron vortex beams. *Nature*, 467:301–304, 2010.

[10] Benjamin J. Mcmorran, Amit Argawal, Ian M. Anderson, Andrew A. Herzing, Henri J. Lezec, Jabez J. McClelland, and John Unguris. Electron vortex beams with high quanta of orbital angular momentum. *Science*, 331:192–195, 2012.

[11] Juan C. Idrobo and Stephen J. Pennycook. Vortex beams for atomic resolution dichroism. *Journal of Electron Microscopy*, 60(5):295–300, 2011.
[12] Koh Saitoh, Yuya Hasegawa, Nobu Tanaka, and Masaya Uchida. Production of electron vortex beams carrying large orbital angular momentum using spiral zone plates. *Journal of Electron Microscopy*, 61(3):171–177, 2012.
[13] J. Verbeeck, H. Tian, and A. Béché. A new way of producing electron vortex probes for stem. *Ultramicroscopy*, 113:83–87, 2012.
[14] A. M. Blackburn and J. C. Loudon. Vortex beam production and contrast enhancement from a magnetic spiral phase plate. *Ultramicroscopy*, 136:127–143, 2013.
[15] A. Béché, R van Boxem, G. van Tendeloo, and J. Verbeeck. Magnetic monopole field exposed by electrons. *Nature Physics*, 10:26–29, 2013.
[16] Robert Collier. *Optical Holography*. Elsevier, June 2013.
[17] P. Schattschneider, S. Rubino, C. Hebert, J. Rusz, J. Kunes, P. Novak, E. Carlino, M. Fabrizioli, G. Panaccione, and G. Rossi. Detection of magnetic circular dichroism using a transmission electron microscope. *Nature*, 441:486–488, 2006.
[18] Benjamin J. McMorran. *Electron Diffraction and Interferometry Using Nanostructures*. Ph. d. dissertation, University of Arizona, University of Arizona, Tucson, AZ, January 2009.
[19] J.S. Pierce, T.R. Harvey, T.S. Yahn, and B.J. McMorran. High efficiency electron diffractive optics. *Microscopy and Microanalysis*, 19(Supplement S2):1188–1189, 2013.
[20] Vincenzo Grillo, Gian Carlo Gazzadi, Ebrahim Karimi, Erfan Mafakheri, Robert W. Boyd, and Stefano Frabboni. Highly efficient electron vortex beams generated by nanofabricated phase holograms. *Applied Physics Letters*, 104(4):043109, 2014.
[21] R. Magnusson and T. K. Gaylord. Diffraction efficiencies of thin phase gratings with arbitrary grating shape. *Journal of the Optical Society of America*, 68(6):806–809, 1978.
[22] Tyler R. Harvey, Gii Brougher, Kurt Langworthy, and Benjamin J. McMorran. Small-pitch electron diffraction holograms patterned on inorganic resist with electron beam lithography. Poster presented at EIPBN 2013, May 2013.
[23] Edgar Völkl, Lawrence F. Allard, and David C. Joy. *Introduction to Electron Holography*. Springer, 1999.
[24] Ampere A. Tseng. Recent developments in micromilling using focused ion beam technology. *Journal of Micromechanics and Microengineering*, 14(4):R15, 2004.
[25] S. Cruz-Arreola and O. Mata-Mendez. Diffraction of beams by infinite or finite amplitude-phase gratings. *Revista mexicana de fsica*, 57(1):6–16, February 2011.
[26] P. W. Hawkes and E. Kasper. *Principles of Electron Optics: Wave Optics*, volume 3. Academic Press, London, 1994.
[27] NIST digital library of mathematical functions. `http://dlmf.nist.gov/20`. Release 1.0.8 of 2014-04-24.

Appendix A. Diffraction Intensity Measurement

To determine the intensity of one diffraction order, we first measured noise in the image; we calculated the average background intensity I_b in a part of the micrograph where no signal intensity was present. We then subtracted this background intensity from the entire image, then set to zero all pixels which had an intensity less than this. Finally, to measure the intensity of the nth diffraction order, we centered a circular measurement region with a diameter equal to the spacing between orders over the nth diffraction spot. We then measured the total integrated intensity inside the circle.

We performed this measurement on an FEI 80-300 Titan TEM at 300 keV in Low-Angle Diffraction at a camera length of 104 m. We used a gun lens strength of 6

and a spot size of 8, a 150 um C2 aperture to limit the extent of the incoming wave, and an illuminated area of 104 μm. We then limited the outgoing wave with a 10 um objective aperture. The incident, transmitted and relative diffraction efficiencies of the beams formed by the grating are independent of aperture size and shape when incident efficiency is measured as a function of incident intensity after the aperture [25]. We used a defocus of $-21.37 \cdot 10^{-6}$ so that the spots did not saturate the detector but were still well spaced.

Appendix B. Nanofabrication Considerations

To meet our stated design goals for efficient electron diffractive optics, we explored a large number of nanofabrication techniques for imprinting the phase grating onto electron-transparent materials. We have found a combination of grating material, charge alleviation layer, and patterning technique that provides good results, and discuss each of these separately in the following sections.

Appendix B.1. Grating Material

While there are many electron-transparent materials that can be used for this purpose, we choose to use silicon nitride membranes because of its mechanical robustness, thermal stability under electron illumination, and availability. We used silicon nitride membranes from several manufacturers, ranging in thicknesses from 15 nm to 100 nm thick, suspended over windows from 80 um across to 2 mm. Low-stress 30 nm to 100 nm-thick silicon nitride membranes purchased from SPI Supplies, Inc. performed best under long-term mill processing. We note that in general, a membrane thickness on the order of twice the final mill depth minimizes beam amplitude loss due to high-angle inelastic scattering while maintaining mechanical stability. In particular, we found that while 30 nm and 50 nm thick membranes block roughly 20 % and 50 % of the incident beam intensity, respectively, these standard thickness membranes routinely provide the greatest absolute diffraction efficiency for a 300 keV beam.

Appendix B.2. Nanofabrication Optimization

We have applied several nanofabrication techniques for patterning electron diffractive optics, including high-resolution focused ion beam (FIB) milling, electron beam-induced deposition (EBID), and electron beam lithography (EBL). We will concentrate our description here on the FIB-milling technique, since FIBs are present in many TEM labs and this maskless fabrication technique provides a quick method for making electron gratings. In particular, we used an FEI Helios NanoLab FIB to mill all gratings described in this work.

Modern FIB instruments offer many different parameters that can be used for milling the electron diffraction hologram pattern. The ion beam dwell time, number of passes, milling scan direction, beam current, ion beam convergence angle, addressable

pixel spacing, total ion dose, and pattern complexity all play a large and interconnected role in the quality of the final grating. An in-depth discussion and review of these effects is provided in [24]. We performed a systematic exploration of this complex parameter space in order to find the best combination for manufacturing efficient gratings for electrons.

We created a series of arrays of linear diffraction gratings in which we systematically varied total ion dose, dwell time, number of passes, patterning order, and beginning membrane thickness. These dose-arrays were placed in a field emission TEM operated at 300 keV. Low angle electron diffraction patterns were recorded under identical illumination conditions for each grating. Examples of these diffraction patterns are shown in Figure 5. The diffraction spots were defocused in order to examine the uniformity of the intensity of each beam – darker areas indicate an unwanted variation in groove width and depth across the grating area. We then measured the surface topology of each grating using an atomic force microscope (AFM) with a small diameter tip (Figure 4).

Many FIBs provide the ability to mill a pattern using a bitmap image or by direct programmable control of the beam path. We find that when using the bitmap patterning method, the pixel spacing of the magnified bitmap image should be an integer multiple of the minimum pixels spacing [24] of the FIB in order to avoid artifacts in the final structure due to nonuniform ion dosing. The scan direction of the beam should be chosen such that the slow scan axis is perpendicular to the grooves. While we have made decent gratings using both raster and serpentine scans during milling, we find best results when using a "vector scan" technique [24]; taking full programmable control of the beam path such that one complete groove can be milled before moving on to the next element. For most purposes, we found that 10 passes, reversing the milling order each time, resulted in the highest quality gratings. However, when milling patterns covering large areas, settling and movement of the silicon nitride membrane reduces the quality of the pattern – in these cases it is best to reduce the number of passes. We find that ion beam currents below 10 pA produce the finest features, but larger currents on the order of 20 pA to 50 pA are necessary to complete a pattern larger than 20 μm diameter without major thermal drift over the course of the mill. For the pattern milled in Figure 2, the ion beam current was set to 10 pA and the number of passes was 80 with a pixel size of 3.3 nm.

Appendix B.3. Charge Alleviation Layer

The nitride gratings must be coated with a thin conductive layer to alleviate charge. We have experimented with using sputtered Ni, Ti, Cr, and Au, thermally evaporated C, and Pt from ion beam induced deposition (IBID). We find that 15 nm of amorphous carbon sufficiently minimizes charging and causes little absorption in the transmitted beam. However, when the carbon-coated grating is placed in an upper condenser lens of a TEM and exposed to a beam over the course of several weeks, the carbon can

migrate into the grating trenches and decrease the absolute diffraction efficiency. A 5 nm to 10 nm layer of Au deposited on a 1 nm Cr adhesion layer leads to a slightly lower absolute diffraction efficiency $\eta_1^{(i)}$, due to absorption and scattering, but the diffraction efficiency of such gratings remain stable for weeks under exposed conditions in the beam path of the TEM. We find that Pt deposited by IBID contains significant amounts of carbon which can migrate under electron beam exposure, but the grating diffraction efficiency can be renewed by regular plasma cleaning. In all cases where a metal film is used, enhanced electron scattering decreases the absolute diffraction efficiency $\eta_1^{(i)}$ and contributes to an unwanted background signal. However, a diffractive structure composed of silicon nitride grooves capped with a metal layer, produced either with FIB or IBID, modulates both the amplitude and phase of transmitted electrons. Such a mixed amplitude-phase grating can almost entirely suppress the 0^{th} (undiffracted) order (see Figure 5(d) and Figure 5(e)).

Appendix B.4. Hologram Uniformity, Spatial Coherence and Quality of Imprinted Phase

Spatial variations in the width and depth of grooves across the grating affect the amplitude uniformity inside each diffracted beam. These spatially dependent errors of FIB-milled gratings are primarily due to secondary dynamic processes such as redeposition, charging, heating, and membrane relaxation associated with the incidence of ions on the substrate surface. Large-area patterns with very fine pitch typically take several hours to mill, and changes in substrate tension or temperature can cause the substrate to drift at speeds on the order of nanometers per minute, which noticeably impacts the spatial coherence of the resulting structure. However, with an appropriate choice of a moderate ion beam current, a fewer number of patterning passes, and the application of a conductive layer to the membrane to prevent local charging, we have successfully produced spatially coherent gratings with 10^3 grooves over areas several tens of microns in diameter. We are in the process of developing a quantitative measure of spatial coherence.

Appendix C. Theoretical Diffraction Efficiency from a General Mixed Amplitude/Phase grating

The diffractive electron optical elements described in this work are thin membranes with a modulated thickness $h(x, y)$. As the both the de Broglie wavelength of electrons and the maximum thickness of an electron-transparent membrane in a transmission electron microscope are necessarily small relative to the grating period, the thin grating condition under which wave interference due to propagation inside the grating material is negligible [21]

$$\lambda h << d^2 \tag{C.1}$$

is necessarily satisfied. Therefore, if an electron diffractive grating is illuminated by an incident electron wave ψ_i, then the transmitted electron wavefunction immediately behind the grating is

$$\psi_t(x,y) = \psi_i(x,y)t(x,y). \tag{C.2}$$

where $t(x,y)$ is the transmission function describing the effects of a thin grating structure on transmitted electron wavefunctions. We note that if $\psi_i(x,y,z)$ is a normally plane wave, then $\psi_t(x,y) \propto t(x,y)$. We can express the transmission function as

$$t(x,y) = e^{-\alpha h(x,y)} \cdot e^{iCV_0 h(x,y)} = e^{i\tilde{k}h(x,y)} \tag{C.3}$$

where $\tilde{k} = CV_0 + i\alpha$ is the effective complex wavenumber of the electron within the material. The amplitude decay coefficient α describes the effects of inelastic and high angle scattering, and V_0 is the mean inner potential of the grating material. C depends only on the energy of the beam [23]:

$$C = \frac{2\pi}{\lambda V_a}\frac{eV_a + m_e c^2}{eV_a + 2m_e c^2} \tag{C.4}$$

where V_a is the accelerating voltage for the electron.

We can describe any surface $h(x,y)$ periodic in the x-direction as a Fourier series with period d:

$$h(x,y) = \frac{a_0}{2} + \sum_{m=1}^{\infty} \left[a_m \cos(k_m x) + b_m \sin(k_m x)\right] \tag{C.5}$$

where $k_m = \frac{2\pi m}{d}$ and the coefficients are defined by

$$a_m = \frac{2}{d}\int_{x_0}^{x_0+d} h(x,y)\cos(k_m x)\mathrm{d}x \tag{C.6}$$

$$b_m = \frac{2}{d}\int_{x_0}^{x_0+d} h(x,y)\sin(k)mx)\mathrm{d}x \tag{C.7}$$

If an electron plane wave normalized in a finite-size box with volume V,

$$\psi_k = \frac{1}{\sqrt{V}}e^{ikz} \tag{C.8}$$

is normally incident on a grating with a surface described by $h(x,y)$, so that $\psi_i(x,y) = \psi_k$, we see from (C.2) that if we place the back of the grating at $z = 0$, as shown in Figure 1, the wavefunction $\psi_t(x,y)$ at $z = 0$ is proportional to the transmission function $t(x,y)$ of the grating,

$$\psi_t(x,y) = \frac{1}{\sqrt{V}}t(x,y) \tag{C.9}$$

Let's examine how this wave propagates.

Far from the grating, the outgoing electron wavefunction behind a grating can be described by the Fraunhofer formula [26],

$$\psi(\mathbf{r}') = \frac{1}{i\lambda z'}e^{ikz'}\int \psi_t(x,y)e^{-i(xq_{x'}+yq_{y'})}\mathrm{d}x\ \mathrm{d}y \tag{C.10}$$

If we define the two-dimensional Fourier transform of a function $f(x, y)$ as

$$\tilde{f}(q_{x'}, q_{y'}) = \frac{1}{(2\pi)^2} \int f(x, y) e^{-i(xq_{x'} + yq_{y'})} \mathrm{d}x \ \mathrm{d}y \tag{C.11}$$

we can then rewrite $\psi(\mathbf{r}')$ simply in terms of the Fourier transformation of the transmission function

$$\psi(\mathbf{r}') = \frac{(2\pi)^2}{i\lambda z'} e^{ikz'} \tilde{\psi}_t(q_{x'}, q_{y'}) \tag{C.12}$$

$$\psi(\mathbf{r}') = \frac{(2\pi)^2}{i\lambda z' \sqrt{V}} e^{ikz'} \tilde{t}(q_{x'}, q_{y'}) \tag{C.13}$$

at coordinates (x', y', z'), where the spatial frequencies $q_{x'}$ and $q_{y'}$ are given by

$$q_{x'} = \frac{kx'}{z'} \qquad q_{y'} = \frac{ky'}{z'} \tag{C.14}$$

Equation (C.13) can be easily modified to take into account the shape of a finite-size aperture which limits the incident intensity; the diffraction efficiencies we will calculate, however, are unaffected by aperture shape and size [25].

Let's then rewrite $t(x, y)$ so that we can compute the Fourier transform easily. In terms of the Fourier series expansion of $h(x, y)$, we have

$$t(x, y) = \exp\left(i\tilde{k} \left[\frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos(k_m x) + b_m \sin(k_m x) \right] \right) \tag{C.15}$$

$$= e^{i\tilde{k}a_0/2} \prod_{m=1}^{\infty} e^{i\tilde{k}a_m \cos(k_m x)} e^{i\tilde{k}b_m \sin(k_m x)} \tag{C.16}$$

Using the Jacobi-Anger expansion,

$$e^{iz\cos\theta} = \sum_{\ell=-\infty}^{\infty} i^{\ell} \mathrm{J}_{\ell}(z) e^{i\ell\theta} \qquad e^{iz\sin\theta} = \sum_{\ell=-\infty}^{\infty} \mathrm{J}_{\ell}(z) e^{i\ell\theta} \tag{C.17}$$

we can rewrite $t(x, y)$ simply in the plane wave basis.

$$t(x, y) = e^{i\tilde{k}a_0/2} \prod_{m=1}^{\infty} \sum_{\ell=-\infty, \ell'=-\infty}^{\infty} i^{\ell} \mathrm{J}_{\ell}(\tilde{k}a_m) \mathrm{J}_{\ell'}(\tilde{k}b_m) e^{i(\ell+\ell')k_m x} \tag{C.18}$$

If we now change the variables in our double sum, defining $j = \ell + \ell'$ so that our plane wave term depends only on one index, j, we have

$$t(x, y) = e^{i\tilde{k}a_0/2} \prod_{m=1}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{\ell=-\infty}^{\infty} i^{\ell} \mathrm{J}_{\ell}(\tilde{k}a_m) \mathrm{J}_{j-\ell}(\tilde{k}b_m) e^{ijm\frac{2\pi x}{d}} \tag{C.19}$$

$$t(x, y) = e^{i\tilde{k}a_0/2} \prod_{m=1}^{\infty} \sum_{j=-\infty}^{\infty} \gamma_j(\tilde{k}, a_m, b_m) e^{ijm\frac{2\pi x}{d}} \tag{C.20}$$

where we've defined a coefficient

$$\gamma_j(\tilde{k}, a_m, b_m) = \sum_{\ell=-\infty}^{\infty} i^{\ell} \mathrm{J}_{\ell}(\tilde{k}a_m) \mathrm{J}_{j-\ell}(\tilde{k}b_m) \tag{C.21}$$

We now have $t(x, y)$ written in terms of plane waves, but not as a linear superposition of plane waves. We can write the transmission function more simply if we perform the product. As the product of two plane waves with wave vectors $\mathbf{k_1}$ and $\mathbf{k_2}$ is another plane wave whose wave vector is the sum of the first two, we can rewrite a product of a sum of plane waves as a sum of plane waves whose coefficients are products over all terms whose total wave vector is constant.

$$t(x,y) = e^{i\tilde{k}a_0/2} \sum_{n=-\infty}^{\infty} c_n e^{ik_n x} \tag{C.22}$$

$$c_n \quad = \sum_{s_n} \prod_{m=1}^{\infty} \gamma_{j_m}(\tilde{k}, a_m, b_m) \tag{C.23}$$

$$s_n \quad = \left\{ j_m : \sum_{m=1}^{\infty} j_m m = n \right\} \tag{C.24}$$

The calculation of c_n can be considered as a discrete path integral in momentum j and a time m to a final position

$$q_f = \sum_{m=1}^{\infty} j_m m = n$$

Each set s_n describes one path which terminates at q_f; c_n is calculated as the sum of products of the coefficient $\gamma_{j_m,m}$ over all such paths.

We see that only those paths which include a finite number of steps at non-zero j_m or rapidly oscillate in j as m approaches infinity can possibly terminate at finite q_f. Fortunately, the contribution of oscillatory paths to the integral is negligible, as, for any physically realistic grating surface $h(x, y)$, $a_m, b_m << 1$ and thus $\gamma_{j_m,m} < 1$ for $m >> 1$ and $|j_m| > 0$. Thus, we can approximate c_n by choosing a cutoff m_c based on parameters of the model $h(x, y)$ for the product.

Now that we have written $t(x, y)$ in the plane wave basis, we can quickly calculate the diffracted wavefunction far behind the grating. Plugging $t(x, y)$ from (C.22) into our calculation of the diffracted wavefunction, (C.13), we have

$$\psi(\mathbf{r}') = \frac{1}{i\lambda z'\sqrt{V}} e^{ikz'} \int e^{i\tilde{k}a_0/2} \sum_{n=-\infty}^{\infty} c_n e^{ik_n x} e^{-i(xq_{x'} + yq_{y'})} \mathrm{d}x \; \mathrm{d}y \tag{C.25}$$

$$\psi(\mathbf{r}') = \frac{(2\pi)^2}{i\lambda z'\sqrt{V}} e^{i(kz' + \tilde{k}a_0/2)} \sum_{n=-\infty}^{\infty} c_n \delta\left(q_{x'} - k_n\right) \delta\left(q_{y'}\right) \tag{C.26}$$

at coordinates (x', y', z'), where the spatial frequencies $q_{x'}$ and $q_{y'}$ are given by

$$q_{x'} = \frac{kx'}{z'} \qquad q_{y'} = \frac{ky'}{z'} \tag{C.27}$$

The amplitude of $\psi(\mathbf{r}')$ in (C.26) produces (12):

$$|\psi(\mathbf{r}')| = \frac{(2\pi)^2}{\lambda z'\sqrt{V}} e^{-\alpha\bar{h}} \sum_{n=-\infty}^{\infty} |c_n| \delta\left(\frac{kx'}{z'} - \frac{2\pi n}{d}\right) \delta\left(\frac{ky'}{z'}\right) \tag{C.28}$$

As noted in (11) and (15), we model our nanofabricated gratings as an array of Gaussian-shaped trenches. In general, for a grating with pitch d and characteristic width σ, the surface thickness $h(x, y)$ of a FIB-milled grating can be described as the periodic sum of Gaussians,

$$h(x,y) = \bar{h} + \sum_{n=-\infty}^{\infty} \left(\frac{A}{\sqrt{2\pi}\sigma} e^{-\frac{(x/d-n)^2}{2\sigma^2}} \right) - b \tag{C.29}$$

where $\bar{h}$ is the average grating height. The normalization A and offset b depend only on trench depth Δh and characteristic width σ. The Fourier coefficients of this model for $h(x, y)$ are

$$a_0 = \bar{h} + A - b \tag{C.30}$$

$$a_m = 2Ae^{-\frac{1}{2}(2\pi\sigma m)^2} \tag{C.31}$$

Therefore, written as a Fourier series, the height profile is

$$h(x,y) = A\left(1 + 2\sum_{m=1}^{\infty} e^{-\frac{1}{2}(2\pi\sigma m)^2} \cos\left(k_m x\right)\right) + \bar{h} - b \tag{C.32}$$

where $k_m = 2\pi m/d$ as usual.

We note that as the coefficients a_m in our model are Gaussian in m, they fall off quickly and the diffraction from such a model can be calculated with good precision by cutting off c_n at $m_c = \frac{\beta}{\sigma}$, where β is an $O(1)$ number.

In order to quickly numerically compute $h(x, y)$, we note that the Fourier series can be written in terms of the third Jacobi Theta function, ϑ_3 [27]. This function is defined as

$$\vartheta_3(z,\tau) = 1 + 2\sum_{m=1}^{\infty} \left(e^{i\pi\tau}\right)^{m^2} \cos\left(2\pi m z\right) \tag{C.33}$$

Thus, we see that we can rewrite $h(x, y)$ as

$$h(x,y) = A\vartheta_3\left(\frac{x}{d}, 2\pi i\sigma^2\right) + \bar{h} - b \tag{C.34}$$

Now, we can easily set the normalization and offset A and b in terms of Δh and $\bar{h}$. We choose Δh to be positive always by convention; A and b may change sign depending on groove shape. Profiles with narrow trenches have $A < 0$; FIB-milled gratings have a narrow-trench profile. Profiles with wide trenches, or, equivalently, narrow peaks have $A > 0$; deposited structures on a flat substrate have a narrow-peak profile. Let's define A and b for the case that $A < 0$.

$$A = -\frac{\Delta h}{\vartheta_l - \vartheta_r} \tag{C.35}$$

$$b = A \tag{C.36}$$

where we've used the shorthand

$$\vartheta_l = \vartheta_3\left(0, 2\pi i\sigma^2\right) \tag{C.37}$$

$$\vartheta_r = \vartheta_3\left(\frac{1}{2}, 2\pi i\sigma^2\right) \tag{C.38}$$

As the scaled full-width at half-maximum of a groove w is more intuitively meaningful, we use w in the main paper. In terms of the characteristic width σ, we see that

$$w \approx 2\sqrt{2\ln 2}\sigma \tag{C.39}$$

This approximately linear relationship breaks down as w approaches 0.5; at this point, the surface is only better-approximated as a single sinusoid with $w = 0.5$ as the characterstic width σ increases. To represent the very wide trenches of a low-density line array produced by electron beam lithography or ion beam induced-deposition, we need only to flip the physical interpretation of the model: we actually want very narrow peaks, so we may flip the sign of Δh and fit the peaks with a small value of σ.

If we furthermore want to calculate $\bar{h}$ from a known maximum grating thickness h_{max}, trench depth Δh and characteristic width σ, we see that

$$\bar{h} = h_{max} - \Delta h \frac{1 - \vartheta_r}{\vartheta_l - \vartheta_r} \tag{C.40}$$

To calculate diffraction coefficicents c_n in this model, we choose a cutoff

$$m_c = \frac{0.5}{\sigma} \tag{C.41}$$

With this cutoff, we can calculate c_n to 10^{-5} precision for any physically reasonable width or height.

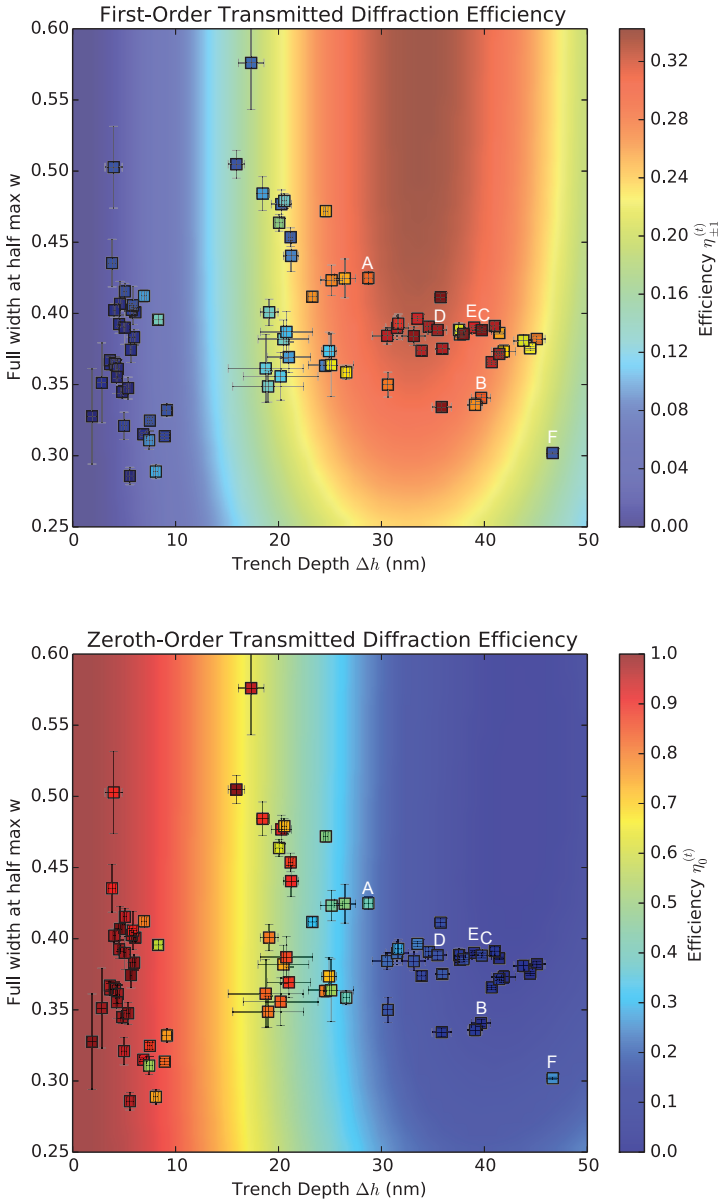


Figure 6. Parameter space map showing the predicted (background color) and measured (squares) first-order (top) and zeroth-order (bottom) transmitted diffraction efficiencies for electron phase gratings as a function of the depth Δh and width w of grooves. Each square data point represents a phase grating fabricated in a silicon nitride membrane using a unique combination of FIB milling parameters. Each square's position describes the associated grating's groove depth (horizontal axis) and width (vertical axis) measured by AFM scans, and the square color is the diffraction efficiency of the respective order measured by TEM diffraction. The labeled data points correspond to the gratings that produced the diffraction patterns in Figure 5. The background color map represents a model that assumes symmetric grooves with uniform Gaussian-shaped depth cross sections and plane wave illumination. Both measured and predicted efficiencies in each plot share the same scale denoted by the right-hand bar.

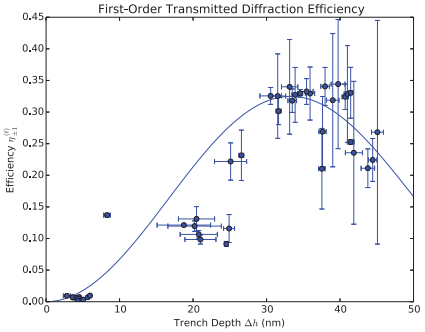


Figure 7. Measured transmitted diffraction efficiency of electron phase gratings as a function of groove depth Δh. The grooves of each grating plotted here all have the same width $w = 0.375 \pm 0.025$ relative to the center-to-center groove spacing. The solid line is a theoretical model that assumes Gaussian-shaped grooves, using only previously measured parameters.