Differential methods in phase imaging for optical lithography

Aamod Shanker
Antoine Wojdyla
Laura Waller, Ed.
Andrew R. Neureuther, Ed.

Electrical Engineering and Computer Sciences
University of California at Berkeley

December 1, 2018

Acknowledgement

Differential methods for phase imaging in optical lithography

by

Aamod Shanker

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Prof. Laura Waller, Chair
Prof. Andy Neureuther
Dr Patrick Naulleau
Prof Hayden Taylor

Spring 2018

Differential methods for phase imaging in optical lithography

To mum and dad

the mother that holds us, and the father that surrounds us

to Doctors Adam Warren and Benjamin Breyer, engineers of the human body

To Laura and Andy

Contents

List of Figures

Acknowledgments

Many have contributed to the formulations of the ideas in this work - first in the list is the ever benevolent and strikingly astute Laura Waller, who with a calm ingenuity has been the guiding hand in the trajectory of innovation presented here. Andy Neureuther, lithography guru who tirelessly encourages students to break through their inhibitions in pursuit of a great vision of the truth, has with his wealth of experience added many dimensions to the doctorate work, and with his eye for precision made it more readable than it might be. Antoine Wojdyla at LBNL and Patrick Naulleau, who in the last year of the dissertation with the passion of career scientists, have helped hone the infantile concepts to working ideas that can touch the world. Colleagues at the national lab, and students on campus, especially Margaret Cruz for her stockpile of apples on hungry afternoons, Gautam Gunjala who starting as a brilliant undergraduate is now on his way to grad school greatness for making life so much more dynamic by adding the flavor of mentorship, all the good people in the IMPACT/CDEN consortium, including Michael Lam as an amazing mentor at Mentor graphics. From faraway Dresden shores, Martin Sczyrba as the patient overseer of the many painstaking experiments, always helpful and with deep wisdom near deadlines, along with Brid Connolly, lending confidence during fast paced conferences. The computational imaging lab members for their unwavering passion and incisive discussions, each student, present and past, postdocs with their own unique strengths, and above all with their willingness to share, being in some way teachers and inspirations on this long winding academic road to salvation. The ALS for their doctoral fellowship, that enables this last contemplative year in the hills of berkeley, around very dedicated scientists pushing the frontiers of knowledge. The lab shuttle drivers for their easy banter. The hills of berkeley, a haven for the hermit, the rivers that yearn for the ocean, the skies that are constantly complementing the moods of the inhabitants, nature's constant inspiration. Finally also to the community in the valley, musicians philosophers and scientists, the larger human family of this nook in the west coast that has allowed a blossoming of the spirit as a scientist and as a human, to where the only

service is to the vision of ultimate unity, and an empathy for the world that reflects the universal, immutable as-it-is.

Abstract

Differential methods for phase imaging in optical lithography

by

Aamod Shanker

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Prof. Laura Waller, Chair

Differential methods are shown to reveal inherent phase in light, corresponding to either the structure of the scattering object, or to aberrations in the imaging medium. Intensity measurements of scattered electromagnetic fields are related to the phase structure of the underlying coherent wavefronts. For an ideal imaging system, intensity differentials with varying focus encode the phase and hence optical density distribution of a scattering object. This is used to characterize diffraction due to thick absorber topography in deep UV lithography (193nm wavelength) using a focus stack from a mask imaging tool. The imaging methods thus developed enable feature-specific process window calibration during lithography, while supplementing mask design capabilities from experimental data.

Switching from strongly absorbing , patterned lithography masks to weakly scattering phase diffusers allows probing the imaging system as well, much like audio system characterization using wide spectrum noise. This enables in-situ characterization of optical-system transfer functions using the speckle generated from the rough diffuser. Under a linear approximation between the diffuser phase and speckle intensity, system aberrations are shown to be encoded in the power spectrum of the speckle. This proves to be an effective way of probing the imaging system in extreme UV lithography (13.5nm wavelength), where mask

roughness naturally creates weak speckle in the image intensity. A new theoretical formulation demonstrates that the phase of the pupil can be completely recovered in a bandlimited imaging system by computing image intensity differentials created by tilting the illumination angle on the EUV mask. Experiments on a synchrotron EUV imaging tool validate the method, showing real time recovery of imaging system aberrations, simply from differentials of speckle images from a blank mask.

Chapter 1

Phase imaging for optical lithography

Optical lithography is the process of imaging patterns from a pre-inscribed mask to a photosensitive resist coating on wafer, etching which transcribes the patterns onto the wafer. Diffraction effects at mask edges influence the phase of the electromagnetic field and are of concern to lithographers, as optical phase defines the through-focus intensity behavior at the wafer plane, impacting the process window during layout and design. Additionally, printed feature sizes in modern chip manufacturing are at the tens of nanometers scale; hence undesired phase at the thick edges of absorber stacks in photomasks can have a major impact on manufacturability and throughput, especially as scaling continues to reduce node sizes and tighten process windows. We aim to measure and model these edge effects to better enable mask designers and OPC engineers to understand their behavior and mitigate undesirable consequences, using measurements from an industrial mask imaging tool from Toppan Photomask operating at deep ultra-violet wavelength ($\lambda = 193nm$), in chapters 1-5. Subsequently methods are developed to not just quantify the mask, but also the optical signature of the imaging tool that images the mask to the wafer, using optical speckle passing through the imaging system. The speckle metrology method is demonstrated experimentally in the soft X-Ray wavelength ($\lambda = 13.5nm$)on the SHARP aerial imaging system at the Advanced Light Source, Lawrence Berkeley National Laboratory. Data from Toppan and

LBNL have made much of the current work relevant to cutting edge lithography challenges in 2018.

1.1 Electromagnetic edge effects in photomasks

Traditional chrome masks are purely absorbing with no designed phase shifts; the absorber however diffracts light around feature edges causing stray light to fall onto the dark regions. As feature sizes have reduced with progressive scaling, the process has grown more sensitive to the diffracted light in the dark areas, and absorbing masks have been mostly replaced with phase shift masks around the 65 nm node size. Attenuating phase shift masks let through a small fraction of out-of-phase light through the absorber, which destructively interferes with diffracted light to print sharper features on the resist, as shown in Figure 1.1.

PSMs however have their own problems - the phase shifts are implemented by increasing the thickness of the absorber stack, increasing the severity of diffraction at feature boundaries. Hence, despite the improved contrast in the intensity images, the phase of the light near feature boundaries has undesirable variations due to edge effects. As the phase of the electromagnetic field impacts through-focus behavior of intensity, smaller nodes which have a smaller depth of focus during printing are more susceptible to edge effects. More recently, OMOG (Opaque MoSi on Glass) masks have replaced phase shift masks at 45nm and 22nm nodes[19], where the phase shift has been abandoned in favor of thinner, better attenuating absorber materials, as shown in the schematic in Figure 1.2.

The diffraction around thick edges is a significant effect, as suggested by the transition to a thinner absorber in OMOG, which mitigates phase effects at the edges but doesn't wholly eliminate them. The ability to measure these phase effects would enable mask designers and proximity correction engineers to compensate for them during mask layout, or process engineers to design the process window while accounting for edge effects.
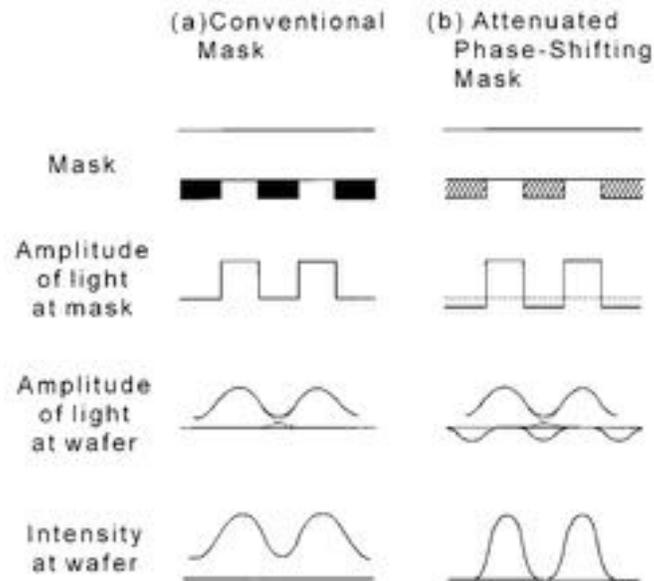
Figure 1.1: Principle of operation of a phase shift mask.a) Conventional absorber mask, with Molybdinum Silicide absorber. b) The MoSi absorber layer is thickened such that the relative phase shift with the clear area is 180°; hence the light through the absorber is out of phase with the light diffracted around the edges. The bottom plot shows the final image, with reduced net intensity in the dark regions, and higher contrast for lithography printing.[Kim et. al. AO 1997]

1.2   Measuring edge effects in a phase imaging system

Full wave simulations

Thick mask edge effects modify the best focus and also the through-focus behavior of the wafer plane images, and hence need to be modeled. Thin mask boundary layer models have been proposed[1, 84, 5] that can approximate diffraction along thick edges as effective boundary layers in simulations. Miller et al.[53] used FDTD simulations on the full 3D geometry of the mask to show that complex valued 2D boundary layers can approximate 3D effects, as shown in Fig 1.3 for an Opaque MoSi on Silica (OMOG) mask. The boundary layers are found to be complex valued, the real part of the boundary layer results in a shifted
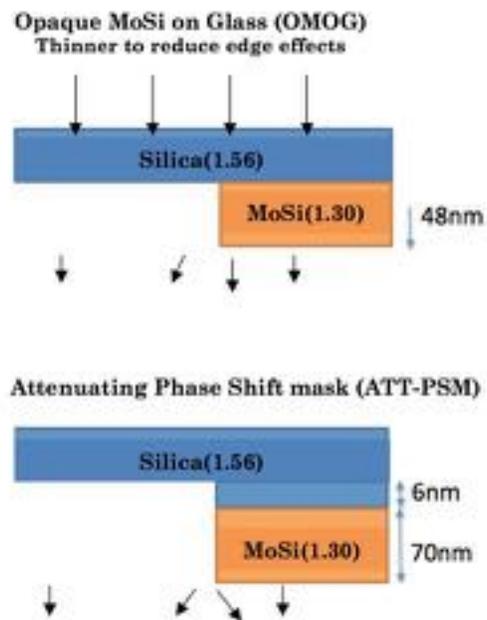
Figure 1.2: Schematic illustrating the difference between attenuating phase shift mask (ATT-PSM) and Opaque MoSI on Silicon (OMOG). ATT-PSM masks are π phase shifting at the absorber and hence have a thick absorber stack, causing diffraction at the edges. More efficient attenuating materials allow OMOG masks to have a thinner absorber stack at smaller phase shift and higher absorption, mitigating the diffraction induced phase effects at the edges of features.

edge position, and the imaginary part contributing to the phase at the wafer plane. The edge effects are primarily in the direction of the illumination polarization, the transverse magnetic (TM) component in polarization direction being about four times as pronounced as the transverse electric (TE) component [53].

Aerial image measurements

Despite being a workhorse for predicting electromagnetic effects, full wave simulations rely on forward models of the mask imaging process and hence have to make certain general assumptions about the imaging system, besides being computationally intensive for large
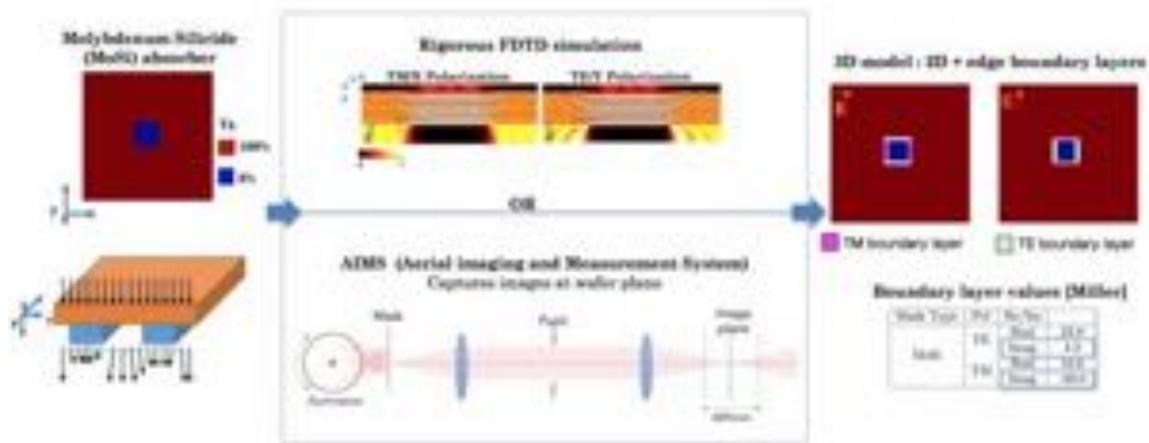
Figure 1.3: Modeling edge effects with simulation vs. experiment. Rigorous FDTD simulations[53] can be used to estimate the effective complex valued thin mask boundary layers due to edge effects. Here we propose the use of through-focus measurements from an aerial imaging tool (AIMS) to extract the phase at the wafer plane, from which the boundary layer values can be estimated directly.
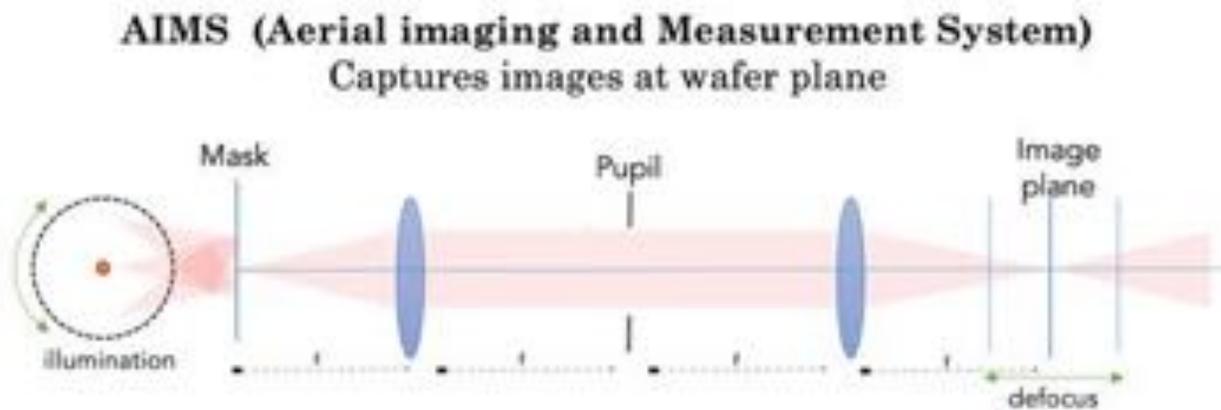
Figure 1.4: Schematic showing a '4f' optical imaging system, or equivalently a lithography Aerial Imaging and Measurement (AIMS) tool. The AIMS tool accepts an input mask and images it to the wafer plane with demagnification, similar to a projection printer. The image at the wafer plane is captured on a camera, which allows metrologists to measure the intensity at the wafer plane for a given mask.

intricate masks. Characterizing edge effects by inverting measurements from the imaging system, on the other hand, enables fast and direct computation of the system-specific phase effects at the wafer plane.

The experimental approach adopted here uses measurements from an aerial imaging tool to characterize the electromagnetic edge effects from intensity measurements at the wafer. As the through-focus intensity behavior at the image plane is dependent on the electric field phase at the wafer plane, the through focus behavior can be inverted to extract the phase from measurements, and subsequently quantify electromagnetic diffraction at feature edges. Additionally the aerial imaging tool may have aberrations.

Figure 1.4 shows the schematic of an Aerial Imaging and Measurement System (AIMS) tool, used for studying the fields at the wafer plane for a mask during inspection and metrology. The system replicates a projection printer in most aspects, with a camera at the back end replacing the silicon wafer. The field at the wafer plane is magnified back onto the camera, and enables inspection of the intensity that would have been printed on the silicon wafer. Since the AIMS tool has an adjustable focus to characterize the through-focus behavior of the wafer plane intensity, it is hence a suitable tool for quantifying the wafer plane phase using defocus measurements.

1.3 The phase of propagating light

Light transport is modeled as wave propagation by the Maxwell's equations, which solve for the complex valued electric and magnetic field as light travels through a dielectric medium. The intensity of the wave is a bilinear function of the field amplitude, and can be directly measured by a photon flux detecting device such as a CMOS camera. Optical phase, a wave property of the propagating light, cannot be directly measured due to the rapidity of the wave oscillations (on the order of THz for visible light). Relative phase shifts, however, are often of interest for imaging application since they correspond directly to the optical path length of a transmissive sample, as shown in Fig. 1.5. In thick photomasks, for

instance, deviations from the designed phase are an indicator of topographical diffraction effects that modify the optical path length of the imaged light. Biologists and doctors have found tremendous utility in examining quantitative phase images of microscopic samples and X-Ray scans of the human body respectively, and lithographers could do the same.

Relative Optical Phase

$$\varphi = \frac{2\pi}{\lambda} \Delta n \times h$$
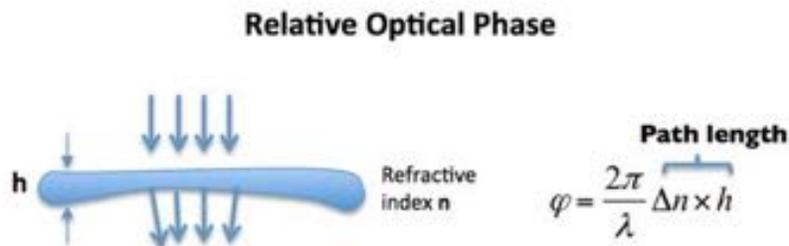
Figure 1.5: The relative optical path length through a transmitting object at different points across its surface depends on the refractive index contrast with the surrounding media and the object thickness at those points. Phase images can hence give information about the object's structure and shape.

The propagation of coherent light depends uniquely on both its amplitude and phase at a given plane. In lithographic setups, where a projection printer acts to image a transmissive mask onto the wafer substrate, variations in the phase of the mask will change the through-focus behavior of optical intensity, and hence put restrictions on the defocus tolerance of the system. Quantitative phase measurements that invert the phase from propagated intensity can be used to image phase defects, measure uniformity across the mask, and help tighten the process window by correcting for stray phase effects using proximity correction methods. The following sections introduce the Transport of Intensity phase imaging method used in this work for extracting phase from defocus intensity, delving into some of its properties and nuances.

1.4 Phase imaging using defocus measurements

Interferometric methods have typically been used for looking at phase variations in coherent light beams, and reveal the phase of an optical beam by comparing with a coherent reference. Due to strict conditions on coherence and alignment of the reference beam, however, these methods are limited in scope. Additionally the need for an extra reference beam makes them impractical for use with existing imaging systems without major hardware modifications, especially in the EUV and X-Ray regimes. Hence defocus based phase imaging methods are preferred in such cases.

In principle, phase recovery from intensity measurements can be performed with any transfer function which couples phase information of the complex valued field into intensity measurements. Iterative methods that use the Fourier transform as the transfer function, iterate between an intensity measurement and the corresponding intensity of the Fourier transform, imposing the measured intensity at each plane until the solution converges to a stable phase value[24, 22] at given intensities. A similar methodology can be used in the fractional Fourier domain, which is equivalent to free space propagation for an electromagnetic wave, or to defocusing for an imaging system. A typical iterative scheme for phase estimation is shown in Fig. 1.6.

Iterative methods are computationally intensive, and convergence properties are object and transfer function dependent. Although applicable to virtually any transfer function, the convergence to a solution and corresponding error bounds are difficult to predict in purely iterative methods, especially when the initial guess for the phase is arbitrary.

Deterministic phase retrieval methods, on the other hand, rely on a invertible transfer function from which the phase is directly solved, enabling real-time recovery of phase. The Transport of Intensity is one such deterministic method, which recovers phase starting from a small defocus approximation that linearizes the dependence of defocused intensity on phase. A stack of through focus intensity images, from a microscope in this case, can then be inverted to recover the phase as shown in Fig 1.7
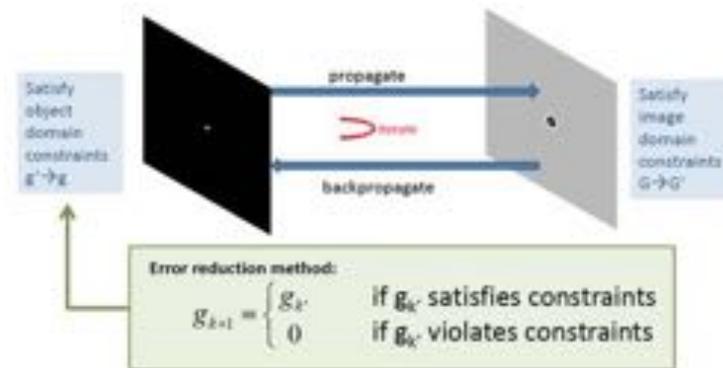
Figure 1.6: Iterative phase methods can recover phase from intensity measurements at two or more planes, given the transfer function between the planes. This schematic shows the Gerchberg-Saxton [31] method for recovering phase from intensities measured at two planes of propagation.
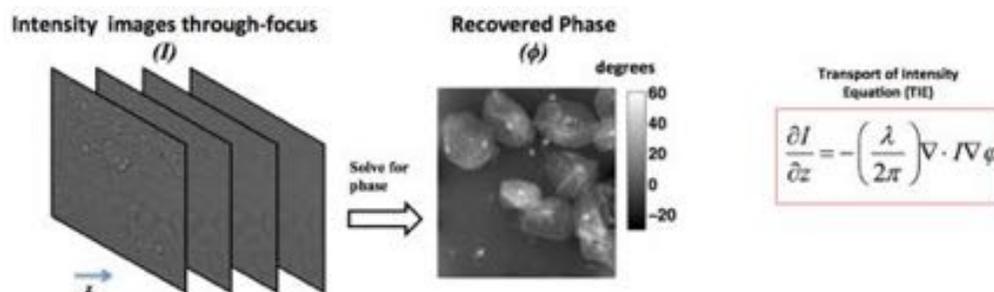
Figure 1.7: The Transport of Intensity Equation (TIE) can be used to recover phase from closely spaced defocus measurements in an optical microscope. a) Stack of images through-focus for a cheek cell sample. b) The intensity difference about focus is related to the phase by the TIE, which can be inverted to recover the phase profile.The recovered phase in radian shows the optical density of the cell.

The Transport of Intensity Equation (TIE) models the intensity difference on propagation as a second order differential equation in the wavefront intensity and phase, which is solved for the phase starting from intensity measurements. [81, 80]. The equation assumes a paraxial approximation, which is most valid for small defocus distances, as elaborated in section 1.5. The advantage of the method is in the fast, direct recovery of phase from defocus measurements and the simplicity of the underlying experimental apparatus. Most imaging systems, including projection printers used in wafer printing, come with a focusing knob, enabling TIE based phase imaging with existing hardware.

1.5 Transport of Intensity phase retrieval

In this section we derive the Transport of Intensity equation (TIE) starting from conservation principles for optical power flow. Most generally, the continuity equation describes a conserved vector field $\vec{D}$ in a N dimensional space as

$$\vec{\nabla}_{1,2,3..N} \cdot \vec{D} = 0$$

where $\vec{\nabla}_{1,2...N} = \{\partial x_1, \partial x_2, ..., \partial x_N\}$ are the partial derivatives. In the case of physical quantities conserved in space- time the equation is alternatively written as

$$\vec{\nabla}_{x,y,z} \cdot \vec{D} = -\frac{\partial D}{\partial t}$$

Physical systems are often understood by modeling the dynamics of conserved quantities, and hence the continuity equation arises in various disciplines, such as fluid mechanics (Navier-Stokes), heat flow, quantum mechanics (probability current) and electromagnetism (electric current). Propagating light in source-free space can be described as electromagnetic oscillations; the corresponding harmonic electric field has a stationary solution in three dimensional space (x,y,z) given by

$$E = \sqrt{I(x, y, z)}e^{i\phi(x,y,z)}$$

In this case the conserved vector field is the power vector, also known as the Poynting vector from the Poynting theorem [66]

$$\vec{P} = I\vec{\nabla}_{x,y,z}\phi$$

where the $\vec{\nabla}_{x,y,z}$ operator is the three dimensional gradient acting on the electric field phase ϕ. The magnitude of the Poynting vector is the intensity I(x,y,z). and the direction is normal to the wavefront, given by the Huygen's principle as $\vec{\nabla}\phi$. Hence the power conservation equation for propagating light in free space is given by the continuity equation for the Poynting vector,

$$\vec{\nabla}_{x,y,z} \cdot \vec{P} = 0$$

$$\Rightarrow \vec{\nabla}_{xyz} \cdot I\vec{\nabla}_{xyz}\phi = 0$$

A paraxial approximation assumes most of the wave energy to be traveling close to the axis of propagation, say z, in which case the wave is almost a plane wave in z,

$$E = \sqrt{I(x, y, z)}e^{i(kz+\phi(x,y))}$$

where $k = \frac{2\pi}{\lambda}$ is the propagation constant, and we make the approximation that $k_z \approx k$. The continuity equation then becomes [85]

$$\vec{\nabla}_{xyz} \cdot I\vec{\nabla}_{xyz}\{kz + \phi(x, y)\} = 0$$

$$\vec{\nabla}_\perp \cdot I\vec{\nabla}_\perp\phi(x, y) = -k\frac{\partial I}{\partial z} \tag{1.1}$$

where $\nabla_\perp = \nabla_{x,y}$ is the two dimensional gradient operator, represented henceforth as $\nabla = \nabla_\perp$ for compactness. Equation 8.3 is the Transport of Intensity Equation (TIE), which relates the through-focus (z axis) intensity gradient to the two dimensional (x-y) intensity

and phase distribution. The equation is a useful approximation of paraxial light propagation as it can be inverted to recover the phase using intensity measurements in the axial direction.

A typical result is shown in Figure 1.7 [44]. Intensity measurements at different defocus planes of a microscope are used to estimate dI/dz at the focal plane for a transparent biological sample with phase variations. Equation 8.3 can then be inverted to solve for phase ϕ, as described in the following section.

Solving the Transport of Intensity equation

The TIE is typically solved [81] by converting the divergence of the Poynting vector, $\vec{\nabla} \cdot I\vec{\nabla}\phi$, into a double derivative using the following substitution,

$$I\vec{\nabla}\phi = \vec{\nabla}\psi \tag{1.2}$$

in Eqn. 8.3, giving,

$$\frac{dI}{dz} = -\frac{\lambda}{2\pi}\nabla^2\psi \tag{1.3}$$

where ψ is an auxiliary variable. Equation 3.2 is a Poisson equation, and can be integrated in the Fourier domain [4, 38], as follows,

$$\psi = \mathscr{F}^{-1}[-\frac{2\pi}{\lambda}\frac{\mathscr{F}(\frac{dI}{dz})}{-4\pi^2 f^2 + \epsilon}] \tag{1.4}$$

where f is the spatial frequency variable, ϵ is a regularization parameter, and $\mathscr{F}$ represents a Fourier transform. The regularization avoids division by zero frequency, but degrades the reconstruction fidelity of low spatial frequencies. Periodic boundary conditions are implied due to the periodic property of the discrete Fourier transform, and a mirror flipping method is used to eliminate discontinuities at the boundary[86].

For a pure phase object, the intensity is constant in-plane, $I(x,y) = I_0$, and the phase is simply $\phi = \psi/I_0$. More generally though, once ψ is known, another Poisson equation needs to be solved to obtain the phase ϕ, as follows,

$$I\vec{\nabla}\phi = \vec{\nabla}\psi$$

$$\Rightarrow \vec{\nabla} \cdot \frac{\vec{\nabla}\psi}{I} = \nabla^2\phi \tag{1.5}$$

which is then similarly inverted by a Poisson solver to recover the final phase $\phi(x,y)$.

Inaccuracies in the inverted phase, ϕ arise due to approximations in estimating dI/dz from measurements, or due to the Poisson solver which is noise and regularization sensitive. Methods such as using more through-focus images[88] or estimating the derivative in the Fourier domain [44, 13] have been shown to improve the derivative estimation as well as noise tolerance of the solution.

The TIE has been commonly used for imaging phase in other applications such as optical microscopy[48], electron microscopy[41] and X-Ray imaging[57, 59]. As demonstrated in the next section, it is fairly robust under partially coherent illumination[64, 98] which makes it particularly suitable for looking at phase also in lithography.

Validity of the TIE

Applicability under varying degrees of coherence is an important requirement for any phase imaging method that is intended for lithographic applications, as sources used in projection printers tend to be partially coherent for maximizing resolution. In this section we derive the TIE from a Taylor series approximation of defocus, and show that the TIE stays valid for partially coherent illumination as long as the source is symmetric.

Validity of TIE for small defocus

The TIE of equation 8.3 can also be obtained as a first order approximation of defocused electric field. Defocus can be modeled as convolution with the paraxial point spread function (PSF), $h_z(x,y)$,

$$h_z(x,y) = \frac{e^{ikz}}{i\lambda z} exp(ik\frac{x^2+y^2}{2z})$$

where λ is the wavelength, and $k = \frac{2\pi}{\lambda}$ is the spatial frequency. Starting from an electric field $E_0 = \sqrt{I}e^{i\phi}$, the field at a defocus distance of z is given as a convolution by the PSF, $h_z(x, y)$,

$$E_z = E_0(x, y) \otimes exp(ik\frac{x^2 + y^2}{2z})$$

In the Fourier domain, $x \to u$ and $y \to v$, the convolution becomes a product,

$$\mathscr{E}_z(u, v) = \mathscr{E}_0(u, v) \times exp(-i\pi\lambda z(u^2 + v^2))$$

The exponential can be Taylor expanded as a polynomial,

$$\mathscr{E}_z(u, v) = \mathscr{E}_0(u, v) \times [1 - i\pi\lambda z(u^2 + v^2) + \pi\lambda^2 z^2(u^2 + v^2)^2 + O(\lambda^3 z^3)...]$$

On inverse Fourier transforming, the property $u^2 + v^2 \to -\nabla^2/4\pi^2$ can be applied, and the defocus field in the spatial domain is obtained as,

$$E_z(x, y) = [E_0(x, y) + i\frac{\lambda z}{4\pi}\nabla^2 E_0 + O(\lambda^2 z^2)]$$

The intensity is bilinear in electric field E, and hence we can write

$$I_z(x, y) = E_z E_z^* = I(x, y) - \frac{\lambda z}{2\pi}\nabla \cdot I\nabla\phi + O(\lambda^2 z^2) + ... \tag{1.6}$$

where the substitution $E_0(x, y) = \sqrt{I}e^{i\phi}$ has been used. A finite difference about z=0 will remove all terms with an even exponent in z, including the quadratic z^2 terms, yielding the TIE of equation 8.3

$$I_z(x, y) - I_{-z}(x, y) = -\frac{\lambda(2z)}{2\pi}\nabla \cdot I\nabla\phi + O(\lambda^3 z^3) \tag{1.7}$$

Hence the error term is of the order of z^3 (with third order derivatives and higher), strengthening the case for dropping the higher order terms within the small defocus regime.
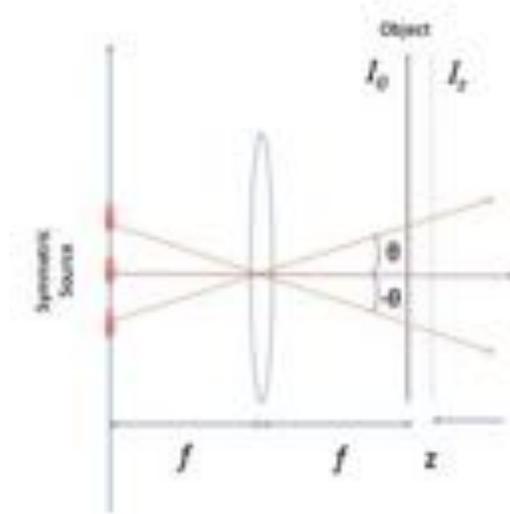
Figure 1.8: The Transport of Intensity equation remains valid for an extended Koehler source which is symmetrical about the optical axis. Contributions to the defocus plane intensity I_z due to source points bisected by the axis cancel out on incoherent addition, maintaining the applicability of the Transport of Intensity for small defocus distance z.

The finite difference can be approximated as a derivative to obtain the more familiar form of the Transport of Intensity as a first order derivative,

$$\frac{dI}{dz} \approx \frac{I_z(x,y) - I_{-z}(x,y)}{2z} = -\frac{1}{k}\nabla \cdot I\nabla\phi \tag{1.8}$$

Validity of the TIE for an extended source

In the presence of an off-axis source at an angle α, the equation 1.6 gets an extra term proportional to the gradient of object intensity, ∇I, and the illumination angle α,

$$I_z(x,y) = I - \frac{\lambda z}{2\pi}\{\nabla \cdot I\nabla\phi + \alpha\nabla I\} + O(\lambda^2 z^2) + O(\lambda^2\alpha^2 z^2) + ... \tag{1.9}$$

with extra higher order shift terms of $O((\lambda\alpha z)^n)$ also appearing due to the off-axis illumination. An extended source can be modeled as an incoherent sum of intensity over α_i, as shown in figure 1.8, and the defocus intensity becomes,

$$I_z(x,y) = I - \frac{\lambda z}{2\pi}\{\nabla \cdot I\nabla\phi + \frac{\sum_i \alpha_i \nabla I}{S}\} + O(\lambda^2 z^2) + O(\lambda^2 z^2 \sum_i \alpha_i^2) + ... \tag{1.10}$$

where S is the total number of points being summed over. Now in the presence of symmetry in the source, $\sum_i \alpha_i^n$ is zero for odd values of n as points at $\pm\alpha$ add to give a zero sum. Hence a symmetric extended source will remove error terms of the type $O((\lambda z\alpha)^n)$ for odd n, and a symmetric finite difference about $z = 0$ will eliminate terms of the form $O((\lambda z\alpha)^n)$ for even n, recovering the finite difference equation 1.7 for a coherent source,

$$I_z(x,y) - I_{-z}(x,y) = -\frac{\lambda(2z)}{2\pi}\nabla \cdot I\nabla\phi + O(\lambda^3 z^3) \tag{1.11}$$

and subsequently the Transport of Intensity equation as a partial derivative. Comparing equations 1.7 and 1.11 shows that the error for a symmetric extended source is invariant of the incoherence parameter α. Sources used in lithography are generally extended sources with symmetry about the center, examples being extended circular sources, dipole sources, and quadrupole sources. The analysis here does not account for finite numerical aperture of the imaging system, assuming that the features are low frequency and hence far from the bandlimit of the system. Nevertheless it goes to show that the Transport of Intensity is robust under an extension of the point source while maintaining symmetry[80] , and hence a suitable phase imaging method for lithographic systems.

1.6 Regimes of optical scattering

From DUV to EUV lithography

It takes light to see light - in the case of an imaging system, optical signals generated out of an electo-optic source, typically a lamp at DUV (193nm) or EUV (13.5nm), or a laser beam, are incident on a patterned material that has some refractive index $n + ik$. The

refractive index is an exponential modulation of the electric field, relating the output field to the input as a function of the propagation distance x through the sample,

$$E_{out} = E_{in}e^{i(n+ik)\frac{2\pi}{\lambda}x} \tag{1.12}$$

with the incident wave for a coherent source being a single harmonic frequency, $E_{in} = e^{i\frac{2\pi}{\lambda}x}$ where n is the real part of the index resulting in a phase delay in the incident light, and k is the absorption coefficient that damps the wave. While the wave propagates with dispersion coefficient $k = \frac{\omega}{c}$, implying a temporal frequency of ω. In this thesis, the temporal variation is always averaged out during measurements, which are typically many orders of magnitude smaller that the oscillation frequency ω of light. Hence the studies are made on spatial distribution of energy, or intensity, a conserved quantity, often related to the 'particle' nature of light.

$$I = E.E^* = |E|^2 \tag{1.13}$$

where the quantization of particles is given by the Plank's equation $I = h\omega$, and the field itself is treated like a complex valued quantity.

Photons, matter -like particles, are 'bosons', conserved, following exclusion principles . They are akin to matter in so far as they have exclusivity in space, one photon at one point, and distribute thereon. Hence the propagation of information is through the waves that constitute the electromagnetic field $E = ae^{(i\phi)}$, the phase ϕ being modulated by the 'clear' part of the sample indicated by the refractive index n. On the other hand, the amplitude a, the energetic term, is modulated by the 'dark' part, indicated by the complex refractive index k. Of course, contrast in the final pattern is written on the intensity a^2 - even though the phase modulation of the wave often is more instrumental in representing information as it appears in the final intensity pattern .

The energetic part of the wave, intensity a^2, is the only way that information in light waves can be accessed by an observer. Imaging is an interaction of an observer with an object

of interest, with the observer often being interested in properties of the object, as is the case for patterned mask lithography (first half of this thesis). Lithographers are concerned with an outcome - being able to create dense repeatable structures with high precision on material such as photosensitive resists and ultimately on silicon. Hence they operate on the 'k' regime, pushing the limits of absorptive materials to achieve higher contrast at smaller transistor nodes. Microscopists, on the other hand, are less materialistic (so to speak), and are used to looking at bodies of water, cells, that act on the refractive index, with an intent to study functional and structural properties of the objects of interest - with scope for mixing imagination with the observations to advance theories in biological sciences. Lithographers on the other hand are driving towards creating and perfecting patterns at nano-meter scales (7nm at the time of writing), with very little room for error, especially since these can transport upwards to the function levels where the computation occurs.

These regimes of imaging represent various polarities in thinking about information transmission and synthesis. In either case, however, there is first an object of interest, classified for this thesis as transmitting phase objects (biological cells, diffusers), or absorbing amplitude objects (lithography photomasks). Then there is the intervening medium, the projection optics, that brings the information of the optically active object, to the observer's perceptory and cognitive functionalities. Such is an 'imaging system', the simplest example is the classic 4f imaging system [34], the lateral-symmetry-preserving space-teleportation system, that can replicate photon patterns in space using a combination of two lenses placed symmetrically.

And finally the sensor, which is closest to captured reality for the microscopist, each photon coupling to an electron in matter to transmit its energy in quantized bits, smeared across sensor buckets arrange in square arrays that quantize the spatial resolution as pixels. The bit level quantization induces shot noise besides the inherent camera dark noise; hence a noisy approximation of the object of interest is captured by the digital camera onto an enlarged pixelized screen for the experimentalist to interpret. The pixelization allows the
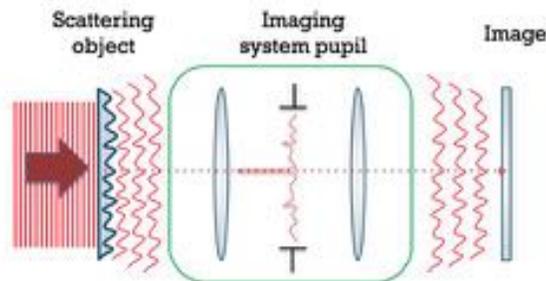
Figure 1.9: Schematic showing a 4f imaging system, where the fields at the object plane are mapped to the image plane, at angular resolution limited by the maximum collection angle, or equivalently pupil size.

advantage of digital electronics, such as scaling onto a screen, or storage onto hard drives, or most importantly logical operations on the received bits represented in matrix forms to implement computational algorithms. Hence the paradigm of computational imaging, where the digital camera acts as the portal for the computational algorithms to propagate out of camera backwards - influencing hardware design choices now which can serve algorithms better, as well as trading-off hardware complexity in lieu of computational effort.

The key contribution of this work to the technical fields lies in investigating phase effects in lithography mask diffraction, where the wave phase arising from mask shape are often overlooked in the quest for designing higher absorption materials. While strong absorption pattern can dominate the image projected by the medium (the imaging system), phase objects allow all light to pass through, accentuating effects of aberrations, or deviations in the imaging system from ideal, in the final image. Further, if the object chosen is a random field, as opposed to a patterned structure, all spatial frequency transmission modes of the imaging system can be probed in parallel. A truly random object loses all biases in its own characteristics, in so far as its spectrum is 'white', hence being able to probe the characteristics of the medium instead. From absorption to phase, from structure to randomness, the thesis transitions from studying 3D phase effects in photomasks (object of interest) to aberration retrieval using weak diffusers (medium of interest). This is represented
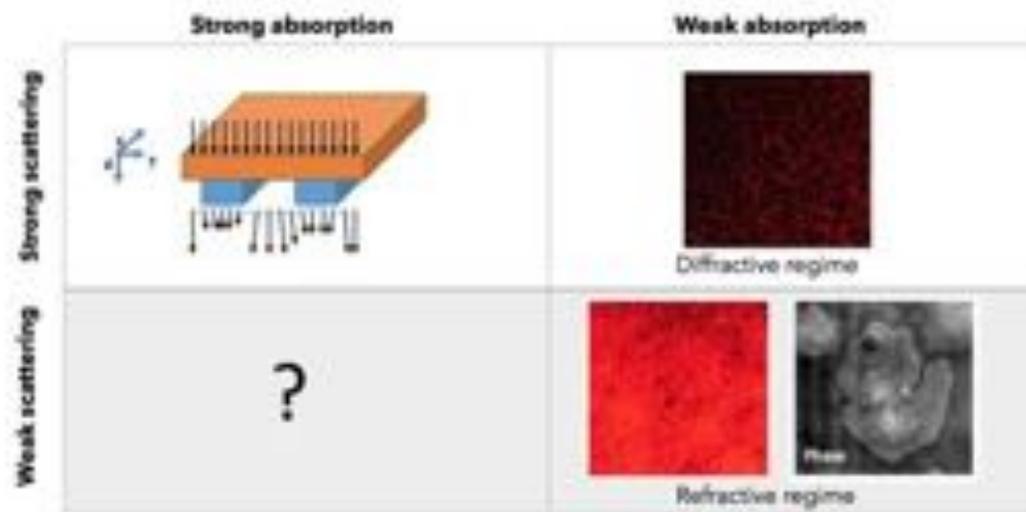
Figure 1.10: Regimes of coherent optical scattering studied in this thesis. The strongly absorbing photomasks (top left) have well defined scattered wavefronts amenable to phase imaging methods, assuming a perfect imaging system. The weakly scattered speckle (bottom right), on the other hand, is useful for probing the imaging system pupil aberrations due to its wide spectral coverage, despite having a random structure. It is shown in Ch. 6 how optical speckle can be used for imaging system aberration retrieval in the weakly scattering, or refractive regime. Strongly scattering speckle (top right) reveal various non-linearities in the phase image, described in Ch. 7 and Ch. 8. The bottom left is an empty block, since much like magnetic monopoles, we are yet to find objects in nature that have strong absorption structure but are weakly scattering.

in the tabular schematic in Fig. 1.10.

There is finally the nature of the observer himself, almost the most crucial step in the pipeline of assimilating understanding from the natural world [79] - left to another parallel work that complements the external, objective, logical nature of the current thesis. This relates closely to the role of the observer in Quantum mechanics, that the founders (Dirac, Von-Neumann etc.) reflected upon at length arising from their study of another wave equation - the Schrodinger's equation. The stationary solutions of both the Maxwell's wave equation for electromagnetic photons, that we base this thesis on, and the Schrodinger's equation for electron waves in the quantum regime, are almost identical. Hence insights gathered in studying coherent wave mechanics extends to the object-subject duality in the

wave equations, and density matrix formulation of quantum mechanics. Finally, the observer must be acknowledged, the researcher, scientist, sentient being, him who lends meaning to observations that are at best muddles of numbers, albeit rife with patterns, waiting to be unraveled by the dense networks of correlated memories centered in the shared scientific consciousness of humanity, channeled through the subjective experience of the researcher. [1]

1.7   Contributions

Edge diffraction metrology in DUV lithography

Phase imaging relies on multiplexing information about the wavefront shape (or phase) in space-time, into intensity measurements. This is typically achieved by the variation of some system parameter. These system parameters must necessarily intersperse phase and amplitude information; the number of parameter states would correspond to the degrees of freedom of the imaging modality, or the number of independent measurements from which the phase and amplitude of the coherent wave must be extracted. For partially coherent systems, the set of unknowns to be reconstructed from the measurements corresponds to the phase and amplitude of each coherent mode.

In a lithography imaging system, as shown in Fig. 1.4, the degrees of freedom in the imaging system are typically at the extrema or boundaries. The initial boundary is at the source, where shifting the source tilts the illumination angle on the scattering object or lithography mask. Alternatively the system can be modified at the detector, most simply by shifting the focal plane (defocus stack). Hence illumination modulation and defocus based phase recovery are some of the most common, minimally-invasive, phase imaging methods,

---

[1]It must be stated, however, that the most direct cognition of light, optics, is through the senses, instantly recognized as a primal communication of fundamental truth, for instance at the golden hours of dawn or dusk, when shifting light shifts moods, sleep cycles, decisions. However the ability to compute on curated environments, with specific objects, designed media, and subsequent numerical conversion using camera sensor pixels, allows for computational projection of reality onto mathematical formulations - exploiting the essential power of numerics as a way of studying and modeling nature with logic.

seen in developments in the fields of Fourier Ptycography [83, 93], and Fienup's popular iterative phase retrieval schemes utilizing multiple defocused intensity measurement [23].

In this work, the phase imaging modality for the strongly absorbing lithography masks is defocus based, allowing for in-situ phase imaging of the scattered light from the mask without modifications to the lithography imaging configuration, thus the configuration closest to in-fab lithography tools is maintained. Chapters 2-5 describe lithography phase retrieval, and its utility for process engineering in the semiconductor fabrication facility. Chapter 2 describes the impact of mask absorber thickness on image phase, and subsequently the utility of phase retrieval in lithography. Chapter 3 shows novel algorithms for phase retrieval in practical systems using intensity images captured through-focus, with sensitivity of upto 0.01 radian. Chapter 4 explores the effect of illumination variation on the phase at the image plane due to absorber thickness. It is shown that illumination tilting can be used as an effective phase retrieval method. Chapter 5 presents a new, simplified, quantitative phase contrast method that allows for effective phase information to be extracted from through-focus images with minimal numerical intervention. The only numerical operation utilized by this through-focus phase modality, called *Differential Intensity Imaging*, is algebraic differences between adjacent images in a focal stack.

Imaging system metrology for EUV lithography

Apart from the optical source and imaging plane, the other pieces of an imaging system are the scattering object, and the imaging pupil (Fig. 1.4). For the phase retrieval of absorbing masks, the object is usually an unknown amplitude object, and the pupil is assumed to be a perfect aperture. However if the object is made perfect in some stochastic sense, such as a diffuser with known first order statistics, it is possible to solve for imperfect pupil functions (aberrations) of the imaging system.

The last section of the thesis transitions from the strong scattering regime of lithography masks at wavelength λ of 193nm (Deep Ultra-Violet) regime, to the weak scattering seen

ubiquitously at soft X-ray wavelength of 13.5nm (Extreme Ultra-Violet), where most materials don't interact strongly with the electromagnetic field. The regimes could alternatively be called diffractive and refractive. The diffractive regime exhibits significant interference effects due to strong scattering, hence requiring a wave mechanical treatment of phase retrieval. The refractive regime, on the other hand, is more akin to ray optics, where energy propagation can be modeled locally with rays of light that add linearly in intensity. Fig. 1.11 shows far field scattered intensity from the weak and strongly scattering regions of an optical diffuser, engineered with index matching oil at visible (628nm) wavelength. Chapter 6 describes the transition between the ray and wave optics regimes in greater detail. By engineering the object, the scattering spans the refractive-diffractive, or linear - nonlinear regimes of optical interaction.
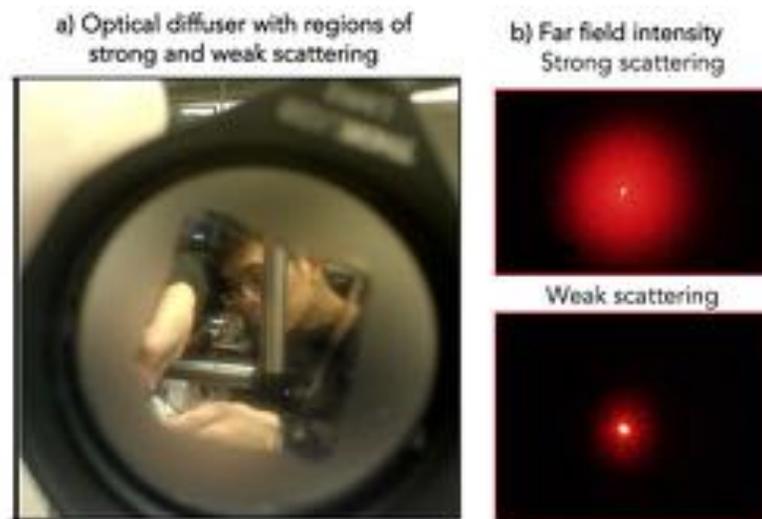
Figure 1.11: Phase diffuser in the visible regime engineered to have regions of strong and weak scattering for visible light. In the far-field, the speckle from the strong scattering region (outside the square) smears across a cross section determined by the propagation distance and diffuser roughness. The cross-section of the speckle from the weak scattering region (inside the square) is halved due to reduced second order interactions. The strong and weak scattering regimes are roughly analogous to deep ultra-violet and extreme-ultraviolet regimes of optical lithography, respectively.

The transition from strong scattering to weak scattering allows a linearization between

the phase of the scattering object and the intensity at the image plane. The linearization, specifically known as the *c*ontrast transfer function, allows solving for the last piece in the imaging system puzzle - the pupil. Chapter 6 derives the weak scattering approximation and its applicability in imaging system metrology using a weakly scattering diffuser. The applications, quite fortuitously, appear in imaging with soft X-rays in EUV lithography ($\lambda = 13.5nm$), where even blank objects have atomic scale roughness that makes them weakly scattering diffusers. A powerful and simple spectral domain method reveals the shape of these aberration functions of the pupil in the simple Fourier transform of the speckle image obtained from an extreme ultra-violet imaging ($\lambda = 13.5nm$) system. The entire aberrated pupil can be retrieved from differences in the speckle spectra on tilting the illumination, which we christen the *Differential Speckle Imaging* method, demonstrated for the first time in EUV imaging systems, with sensitivity to upto $\lambda/20$ wavefront aberrations in the imaging system. [2]

1.8 Epilogue: On the nature of light

Science as a way of investigating the natural world comes instinctively to us humans - the scientific method depending upon an observed quantification of physical parameters in matter, their interaction in time, and formulation of laws thereof to isolate relationships locally in systems interacting at many spatio-temporal scales. Access to the scientific method, then, often comes by observation of the natural world, followed by precise interactions with the systems under test, to deduce more generally applicable laws - it is indeed a miracle then that there is any hope of generality, but nature is kind enough to allow consistency across various contexts. The first of these natural systems to interact and generate laws for, is of course the mind itself, the outcome thereof being the laws of the mind - including logic, geometry, algebra, topology - or any of the many mathematical concepts that although

---

[2]Coincidentally, electron microscopists has used similar speckle based aberration retrieval methods since the 70's, but without the algebraic/computational rigor or tilted illumination scheme suggested in Ch 6.

discovered by the mind, are seated in the physical matter of the brain, the temple of the mind . However since all information exists in a 'relative' sense, the mind, must interact with its environment to generate a sense of identity, and subsequently meaning - this gives rise to the discipline of metaphysics, where minds interact with each other, somehow interfere, for the creation of content, meaning, information - the general 'something' out and instead of the 'nothing'.

For the scientifically inclined however, there is a primal need for experiments, for interacting with physical systems, to transform them, project them onto previously calibrated systems, called measurement devices. These measurements are observed with structures in the mind - theoretical frameworks. The most ancient of these physical experiments have been with light, chronologically relevant, since the first form of being rising out of the void of precreation is light itself, electromagnetic radiation so named to contrast it with the heavy state, matter. Since Galen from 2nd century Greece formulating theories of the body and mind based on the visual sensations, through Newton's color splitting experiments in post medieval England, to Einstein's insights into the relative nature of space-time starting with the fundamental manifestation of reality at the speed of electromagnetic propagation, to modern day post-silicon virtual reality paradigms that combine light field rendering holographic headsets, with visual perception principles to generate virtual realities - optical physics is the principle resource for progressing human understanding of nature, and subsequently, of their own minds and beings.

The wave mechanics that underlies the physics of light, first uncovered by Young in localized correlation experiments, later described with great elegance by Maxwell in coupled differential equations, is fundamental to all information propagation. Time can be argued to arise as a flow of this information, the speed of light 'c' being the rate at which cause and effect originate in each point at space, the baseline for information flux in the cosmos. Further, as these propagating waves, an outcome of elliptical differential equations, arise from the projection of circles onto straight lines, the harmonic basis for information flow is established. Complex exponentials define circles in a vector space, giving a primary language

to describe waves interacting, and their interference. What is even more surprising, and an extremely propitious natural law, is that harmonics, those sinusoids that change by their second derivative, are also eigenmodes of propagation in matter of energy. The speed of information flow in this case however, is many orders of magnitude lesser than of light, since the density of the medium is much larger. Localization in space, in this sense, can be thought of as waves beating against each other, interacting and exchanging energy over many scales, to create lower vibration modes, including those in water, in solids, in fluids, or in electrons at scales length approaching their coherence length. Information flow is slower than the speed of light in these cases, indicating why ultra-fast optical experiments are such a common way to investigate the properties of chemical reactions.

While these fields, the hidden information flow, are the waves that move and combine linearly, the measurements, or observable that they hence manifest, are in correlation space-typically second order correlations, as seen in the conservation principles of energy (bilinear in the field), and phase-space representations, such as Wigner distributions, that simplify the visualization of optical cause-effect simply by looking at the correlation of fields in space, and their marginal (one dimension at a time) projections. This allows visualizing the whole optical flow process, through space, or matter, or through optical elements such as lenses, as simple transformations on a conserved volume in correlation space, or Wigner space. This allows a holistic description of fundamental optical properties, such as propagation, diffraction, uncertainty principle, design of light fields - effectively all of holography converted to topographical manipulations of the conserved energy volume in six dimensional phase space.

All scientific theory, as demonstrated so well by group theory and its application in describing symmetries in particle physics, must be founded in finding repeating patterns in structures observed in nature. Hence correlation based methods, as powerful as they prove in light, also underlie all the other sciences. In fact, it is barely too speculative to claim that the mind operates in correlation spaces as well, following laws of symmetry and periodicity at various scales, spatial and temporal.

Studying the properties of light, as naturally as it comes to children blowing soap bubbles in the sun, allows an approach to phenomenology that transcends disciplines, being arguably crucial to address the growing challenges of human society. From Galileo's telescope that sparked many revolutions simply by putting in context the place of the earth in the grander scheme, to Moore's law where the fine structure of light is used to translate complexity to silicon, each chip then designing the parameters for the next generation allowing for exponential complexity in circuits; enabling technology that engages and employs people across the globe, keeping step with population growth, but eventually hitting diminishing returns as do most exponentials in real life. Progress must continue, however, and the answers are all always here - it simply takes looking at the light - fiat lux!

Chapter 2

Phase edge effects in DUV photomasks

2.1 Introduction

The lithography industry has long relied upon principles of optical imaging to push the limits of manufacturing at the nano-scale. Using various tricks in the imaging system (high numerical apertures, lower wavelengths), the process (double and triple patterning), and the material (high k materials), the resolution of imaging a pattern from the strongly absorbing photomasks down to the silicon wafer has been incessantly refined. The resolution scales according to Rayleigh's classical criteria as [51],

$$FWHM = \frac{k}{2}\frac{\lambda}{NA} \tag{2.1}$$

describing the full width at half maximum (FWHM) of the smallest point that can be projected by a optical tool with numerical aperture NA ($\in (0, n)$) in a medium with refractive index n, and wavelength λ. The next generation of optical lithography tools will rely on every smaller wavelengths. The current workhorse ($\lambda = 193m$, deep - UV) has sustained Moore's law upto the 10nm transistor size on silicon. To go beyond, next generation soft-

xray wavelengths ($\lambda = 13.5nm$, EUV), are eagerly anticipated by industry experts at the time of writing.

Thick mask effects, arising due to the failure of approximating a lithography mask as a two dimensional pattern, grow in significance as the wavelengths shrinks. This is because the 'negligible' height of the absorber becomes more significant as absorber area reduces compared to absorber height. Hence the mask behaves as a thick, 3D object, another area of close scrutiny in this era of rapidly emerging EUV technology, where shadowing effects are now significant contributors to the final image on wafer.

Thick mask effects cause phase variations across features in the aerial image at the wafer plane of a lithographic stepper. This wafer-plane phase will introduce asymmetry in the intensity through-focus, causing shrinkage of the process window (Fig 2.1). With the advent of more absorbing materials, industry has switched to thinner binary masks from the higher contrast phase shifting masks in an attempt to mitigate undesirable phase due to mask thickness. The attenuating OMOG (Opaque MoSi on Silica) masks designed by Shin-Etsu, for instance, use a high-k material to achieve large extinction with a thinner absorber [20]. However, phase effects at the wafer still persist for OMOG masks, causing defocused intensity to deviate from the symmetrical behavior expected for an ideal binary mask (Fig. 2.1).

Lithographic processes are often evaluated with focus exposure matrices (FEMs), where the critical dimension (CD) of a feature is plotted against defocus for various exposure levels (or conversely for various resist thresholds). These "Bossung" plots are a good indicator of process window with respect to focal budget and exposure latitude for a given CD tolerance [12]. Since the presence of phase at the mask causes intensity to be asymmetric through-focus, it causes the Bossung plot to tilt. The Bossung tilt will lead to a smaller focal budget, and can thus be used as a measure of the loss in process window due to topographic phase effects. In work preceding this thesis [74], using experimental images of an OMOG absorber taken with AIMS^{TM} (Aerial Imaging Measurement System[95]), the wafer phase across a feature was obtained from a through-focus intensity stack by solving the Transport of Intensity Equation (TIE) [81, 80]. It was observed that significant phase modulation does
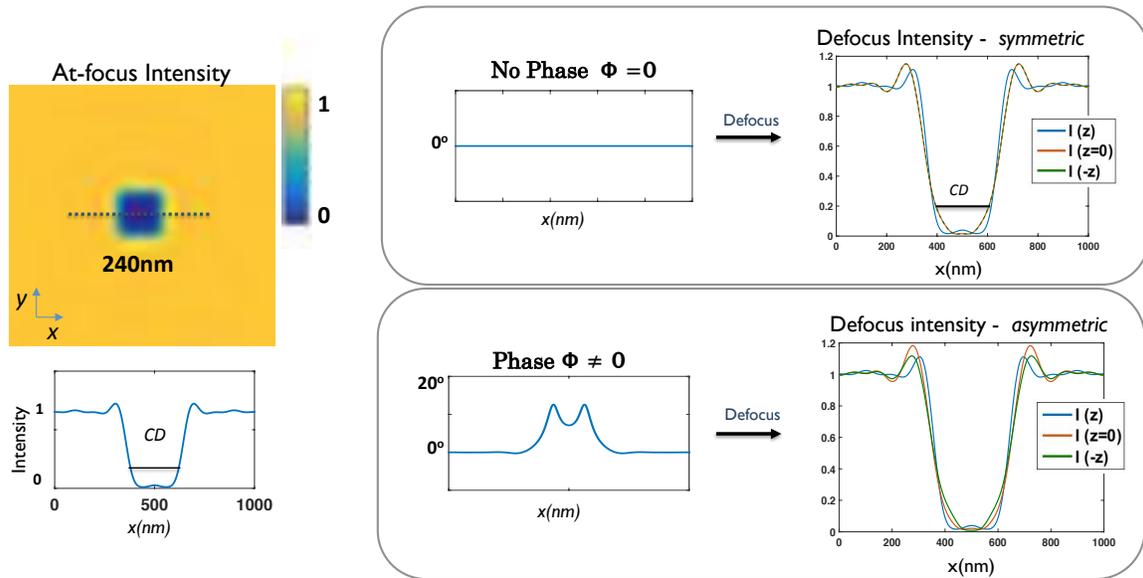
Figure 2.1: Intensity variations through-focus depend on both intensity and phase at the mask. For an ideal binary mask with no phase effects (simulation), positive and negative defocus produce the same intensity curves; hence, defocus is symmetric through-focus (top right). A real OMOG mask has phase modulation across the feature at the wafer plane, which causes symmetry-breaking such that intensity is different on either side of focus (bottom right).

exist across the absorber, especially when illumination polarization is perpendicular to the feature edge (TM), seen in Fig. 2.2. This undesired phase is a consequence of the mask topography, and must be modeled in order to be mitigated. A common approach to include phase introduced by mask 3D topography in thin mask models is to decorate feature edges with complex-valued boundary layers [1, 84, 53] that represent diffraction effects. Here we will use a similar boundary layer framework to identify phase due to topographical factors on the mask.

It has also been demonstrated that mask topography impact on aerial imaging can indeed be mitigated with absorber design; studies on mask thickness dependence of best focus[26, 14] show that a thinner absorber causes smaller phase effects at the wafer, reduces shifts in best focus, and lessens linewidth variation through-focus. In what way exactly the wafer

phase and Bossung plot relate to mask phase, however, still needs clarification. More specifically, what is the contribution of phase through the large area of the absorber stack (*bulk phase*) vs. contribution from diffraction at the sidewall (*edge phase*), and how do each influence wafer phase? Additionally, how does the wafer phase influence aerial image behavior through-focus? Even though the qualitative dependence of defocus behavior on wafer phase has been observed previously [40, 84], an explicit analytic relationship between linewidth variation through-focus and wafer phase, derived here, is essential to relate phase effects to the lithographic process window. It has also been demonstrated that mask topography impact on aerial image signal be mitigated orber mask thickness dependence [14] shows er absorber phase effects at the wafer s in best focus, and lessens line ough-focus. In what way fer phase and Bossung plot relate to mask phase, however, still needs clari specifically contri ugh the large area of the *bulk phase* ion fr sidewall (*edge phase*), and how do each influence wafer phase? Addition afer phase influence aerial image behavior through-focus? Even though the qualitative dependence of defocus behavior on wafer phase has been observed previously[40, 84], an explicit analytic relationship between linewidth variation through-focus and wafer phase, derived here, is es-

Figure 2.2: Topography of the absorber stack causes polarization-dependent phase modulation across the feature at the aerial image, as seen in this measurement of a 240nm absorbing OMOG contact, imaged with an aerial imaging (AIMS) tool. The phase is recovered by measuring a stack of through-focus intensity measurements and solving for phase at-focus.

sential to relate phase effects to the lithographic process window. This work analyzes mask topography effects with rigorous electromagnetic (EMF) simulations of the thick mask to investigate its influence on wafer phase and Bossung behavior. Studies are performed for TM polarization for a 240nm isolated feature (except in Sec. 2.6), using a threshold-based resist model to define CD in the aerial image with partial coherence of $\sigma = 0.3$. First, a small-defocus linearization is developed, which algebraically relates the wafer phase to Bossung tilt, and hence to process window. Next, as the topography varies, the contribution of *edge phase* and *bulk phase* at the mask to the phase at wafer are identified using comparisons with a thin mask model, and their individual impact on the Bossung plot is studied. Finally, our conclusions are verified with aerial image experiments, demonstrating the dependence of wafer phase on the mask parameters as discovered in simulations. The layout is as follows: Sections 2.2 and 2.3 explore the relationship between wafer phase and Bossung tilt, arriving at an algebraic expression connecting the two. Sections 2.4-2.6 study the impact of absorber height, sidewall angle and feature size on wafer phase and Bossung tilt. Subsequently, experiments on ATT-PSM masks are used to confirm predicted trends (Sec. 2.7). We find that even though absorber height and sidewall angle can affect the best-focus by shifting the Bossungs up or down (as concluded by Finders et al [26]), the Bossung tilt is tied to *edge phase* and stays mostly unaffected by fluctuations in mask topography. Hence, the process window, primarily determined by Bossung tilt, is robust to fluctuations in absorber height/sidewall angle. Having identified polarization dependent phase in experiments, here we hope to arrive at a breakdown of various topographical parameters that influence phase contribution due to absorber bulk vs absorber edge, and hence on the net phase at the wafer for TM polarization. Using a sharp threshold to represent the photoresist, the wafer phase will then be related to changes in the Bossung plot and impact on the process window.
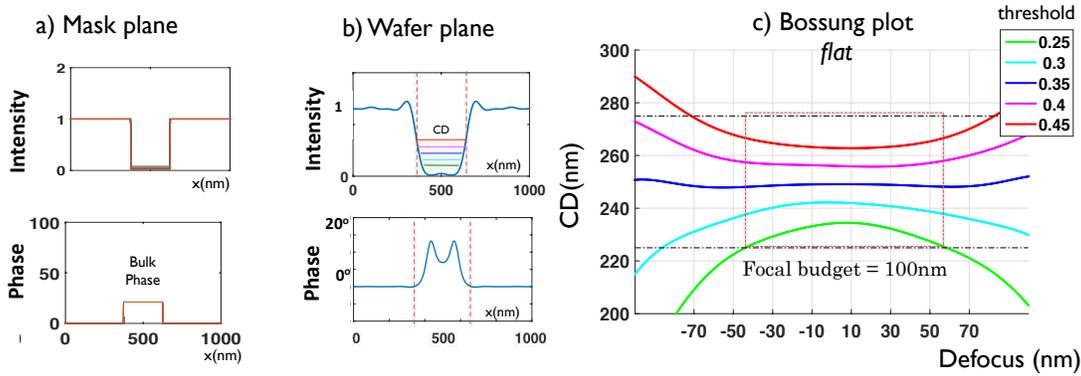
Figure 2.3: a) Thin mask simulations with a constant *bulk phase* across the absorber. b) Even though bulk phase modulates the wafer-plane phase, phase gradients still fall to zero at the position where CD is defined on the aerial image (red dots). c) As a result, Bossung plots are symmetric about focus, allowing a wide focal budget.

2.2 Bossung tilt dependence on wafer phase

To identify the contributions of the *edge phase* and *bulk phase* at the wafer, their individual impact on the Bossung plots and wafer-plane phase is studied. First, the relationship between Bossung plots (CD vs defocus) and wafer phase is demonstrated using through-focus
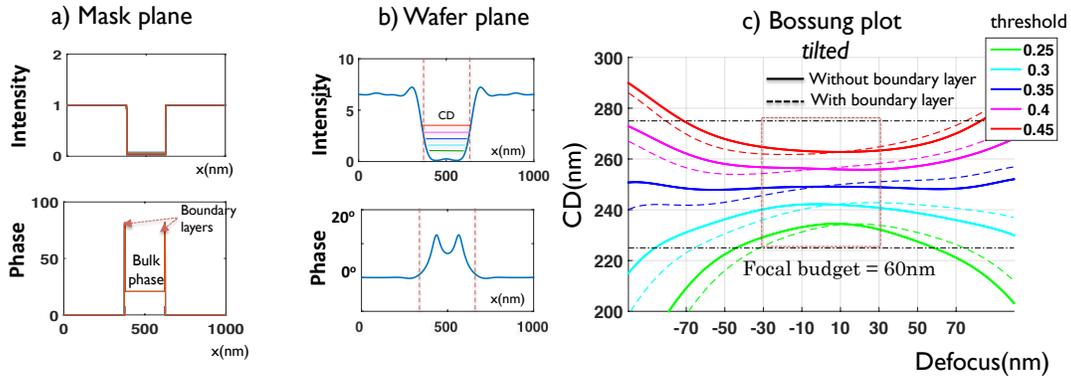
Figure 2.4: a) Boundary layers are added in the thin mask model, along with bulk phase. b) As a result phase gradients are no longer zero at the feature edge at the wafer plane. c) This causes the Bossung plots defined at various thresholds to tilt, forcing the focal budget centered about best-focus to shrink by about 40nm compared to Fig. 2.3c.

aerial image simulations of a thin mask.

Figure 2.3 simulates an absorbing contact in the ideal thin mask case - with a constant *bulk phase* across the absorber, but no boundary layers to model the *edge phase*. The Bossung plot is shown for five different thresholds, corresponding to decreasing exposure from green to red. Even though the wafer-plane phase shows up to 15° modulation across the feature, it is interesting to see that the Bossung plot is flat (i.e. CD is symmetric about defocus). Intensity is thus symmetric about focus at the feature edges, but not necessarily at other positions on the feature. As explained later in Sec. 2.3, Bossung tilt depends on wafer phase gradients at the position where the CD is defined (red dashed lines in Fig. 2.3b) ; in this case the phase gradient falls to zero at the edge, thus yielding a flat Bossung. The process window impact can further be quantified in terms of focal budget - for a CD tolerance of 10%, the focal budget is the smallest focus range that can contain the CD within tolerance for given range of exposure - in this case the Bossung is flat, and hence the focal budget is a healthy 100nm (Rayleigh depth of focus being about 140nm for given $\lambda = 193nm$, NA = 1.4).

The intensity cutlines on either side of focus are found to be symmetric, as expected (see Fig. 2.3b). The Bossung plot is shown for five different thresholds, corresponding to decreasing exposure from green to red. For each case, they are seen to be symmetric about focus, as expected. Now we introduce complex-valued additive boundary layers at the feature edge ($1\angle 90°$, 8nm wide) that represent *edge phase* at the mask. This causes the wafer phase across the feature to spread, such that phase gradients are no longer zero at the edge (Fig. 2.4b). This reflects as a tilt in the Bossung plot, shown as dotted curves in Fig. 2.4c. The tilt compromises the focal range that permits the CD to stay within tolerance for the given exposure range. In this example, the focal budget is narrowed to 60nm (from 100nm for a flat Bossung), shrinking the process window significantly. Note that the best focus (the peaks of the Bossung plots) may change for a tilted Bossung, even if it shifts up or down with no additional tilt, for instance on changing the absorber height as we shall see later. However, this is not a fundamental loss in process window, since a shift of the Bossung plot

can be compensated by redefining zero focus and nominal CD.

2.3 CD-TIE : Quantifying Bossung tilt vs. wafer phase

Clearly, Bossung tilt is directly related to wafer phase - the dependence is now quantified for the region near focus, based on the Transport of Intensity Equation (TIE) [81, 80]. The TIE expresses the changes of intensity through-focus (z) in terms of the phase (ϕ) derivative in the lateral dimensions (x, y) at the focal plane,

$$\frac{dI}{dz} = -\frac{\lambda}{2\pi}\nabla \cdot I\nabla\phi, \tag{2.2}$$

where $\nabla = \frac{d}{dx}\hat{x} + \frac{d}{dy}\hat{y}$ is the two-dimensional gradient operator. This linear formulation relies on a paraxial approximation and small defocus assumption. Originally, the TIE was derived to solve for phase from intensity images through-focus (which is the basis of our experimental phase retrieval methods). Here, we use the same equation to instead derive a quantitative relationship between phase effects and Bossung tilt, in terms of CD. The TIE also explains defocus behavior of intensity in Figure 2.1 - zero phase will cause the intensity gradient through-focus (z) to be zero (by Eqn. 8.3), and hence results in symmetric, overlapping intensity curves on either side of focus. This symmetry is broken in the presence of phase, and the intensity on either side of focus is no longer the same. Hence defocus intensity, and critical dimension (CD), defined here as the linewidth at a given resist threshold, depend on the phase and intensity of the electric fields at-focus. Alternatively, the TIE can be solved starting from the left hand side of Eqn. 8.3 - measurements of intensity change with defocus can be solved to recover the phase at-focus (e.g. Fig. 2.2), which corresponds to the shape of the wavefront[70], allowing potential application as a metrology/inspection method.

The Bossung tilt (CD vs. defocus) near focus can be related to the phase of the electric

field at-focus starting from the 1D TIE, since CD is defined for one dimension at a time,

$$\frac{dI}{dz} = -\frac{\lambda}{2\pi}\frac{d}{dx} \cdot I\frac{d\phi}{dx}. \tag{2.3}$$

Expanding the derivative on the right side of Eq. 2.3, and multiplying the left by dx/dx, we get,

$$\frac{dI}{dx}\frac{dx}{dz} = -\frac{\lambda}{2\pi}\left[\frac{dI}{dx} \cdot \frac{d\phi}{dx} - I\frac{d^2\phi}{dx^2}\right]. \tag{2.4}$$

Since a change in the x position of the aerial image (at resist threshold) with defocus z corresponds to a change in the critical dimension (CD), then for a laterally symmetric feature (which introduces a factor of 2) one can equate $2dx/dz = dCD/dz$. In this way, Eq. 2.4 can be re-written as

$$\frac{dCD}{dz} = -\frac{\lambda}{\pi}\left[\frac{d\phi}{dx} - \frac{I}{dI/dx}\frac{d^2\phi}{dx^2}\right], \tag{2.5}$$

which we call the CD-TIE. The CD-TIE relates the Bossung tilt around focus (dCD/dz) to the phase (ϕ) and intensity (I) derivatives where the resist threshold meets the aerial image (dotted lines in Fig. 2.4b). Hence, the Bossung tilt depends strongly on the first derivative of phase at the feature edge, but also on the second derivative of the phase weighted by the inverse Normalized Image Log Slope, NILS = $(dI/dx)/I$. Typically, the second derivative of the phase is much smaller than the first term, and hence is negligible for a high NILS. In the situation of Fig. 2.4, for instance, $d\phi/dx = 7.7 \times 10^{-4} m^{-1}$, while $\frac{1}{NILS}d^2\phi/dx^2 = 2.3 \times 10^{-8} m^{-1}$, so the second term is negligible. The first derivative of phase at the feature edge is therefore the main factor causing the Bossung to tilt by 0.04nm/nm in CD per defocus units.

An intuition for the effect of phase derivatives in the CD-TIE can be obtained by considering Huygen's principle, which states that light propagates normal to phase fronts. The first term on the right side of Eq. 5.4, containing the first derivative of phase, models intensity steering by the local slopes of phase front with propagation, while the second term represents focusing from the local curvatures in the phase front [70]. The left-hand side is

a finite difference of the CD on either side of focus. The end result is a quantification of how phase derivatives at the feature edge determine Bossung tilt, which in turn impacts the process window.

2.4 Impact of absorber height on aerial imaging

How do the phase at the wafer and Bossung plot respond to the fluctuations in mask topography? We start with an absorber modeled as a trapezoid (n =1.23, k =1.45), with topography variations being approximated by changes in absorber height and sidewall angle. This section studies the dependence of aerial imaging performance on absorber height variations, with the analysis for sidewall angle in the next section.

A rigorous EMF solver (Panoramic EM Suite [65]) is used to investigate the variation of aerial imaging performance with up to 10% fluctuations in the height of a simulated OMOG absorber, as shown in Fig. 2.2. The simulations use DUV light with wavelength $\lambda = 193$nm. The imaging is with a monopole source (incoherence parameter $\sigma = 0.3$) and numerical aperture NA=1.4. The imaging system simulates low-pass filtering by the NA and summing over each point of the partially coherent illumination to obtain wafer-plane fields.

Changing the absorber height modulates the *bulk phase* linearly, since phase through the large area absorber $(\frac{2\pi}{\lambda} h \times \delta n)$ depends linearly on absorber height h (dots in Fig. 2.5b). The phase at the aerial image has taller peaks as the absorber gets thicker (Fig. 2.5a), the phase swing following the increasing trend of *bulk phase* through the absorber. The sensitivity of the phase to absorber height is about 1°/nm - larger than the drop in absorption (0.13%/nm). Hence, phase at the wafer could be used as a sensitive metric for measuring absorber thickness.

Furthermore, changing absorber height shifts the Bossung plot upward, *without affecting its tilt* (Fig. 2.5c), since the phase gradient at the feature edge is about the same for all heights. This will also cause best focus (Bossung maxima) to walk linearly with absorber height. This is not a loss in the fundamental process window however, which can be regained
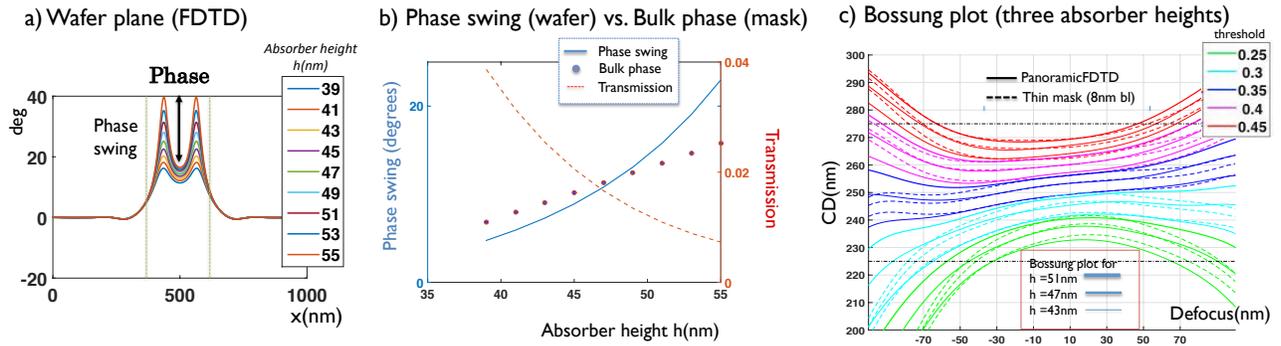
Figure 2.5: Variation in wafer-plane phase with 10% variation in absorber height. a) The wafer-plane phase swing increases with absorber height, although gradients at the feature edge stay about the same. b) The phase swing closely follows the calculated *bulk phase* with increasing absorber height, while transmission drops exponentially for given material. c) The Bossung plot shifts upward for increasing absorber height, with no change in the tilt (solid lines). The thin mask Bossung plot (dotted lines) needs to include the changing *bulk phase* and a constant *edge phase* to get a good fit.

simply by redefining optimal focus and nominal CD. As we saw in Sec. 2.2, the tilt is caused by edge diffraction at the mask sidewall (*edge phase*), which must hence also be insensitive to absorber height. Next we quantify the exact contribution using a thin mask model to fit to the FDTD wafer phase.

The contribution due to edge diffraction may also increase as the absorber gets thicker. How much of the phase effects of Fig. 2.5b are due to the *bulk phase* vs. *edge phase* from the thick mask sidewalls, and what is the impact on the Bossung plot and process window? We answer this next, using a thin mask model.

Thin mask modeling : *bulk phase* vs *edge phase* contributions of thick absorber

While the EMF simulations, followed by aerial imaging generates the net electric fields at the wafer plane, it is not trivial to isolate contributions from the bulk vs edge of the absorber, since the mask near-field has been low-pass filtered by the imaging system. A thin mask approximation that treats the phase in the near field as a sum of *bulk phase* and *edge*

phase will give insight into the amount of edge diffraction in play at the mask.
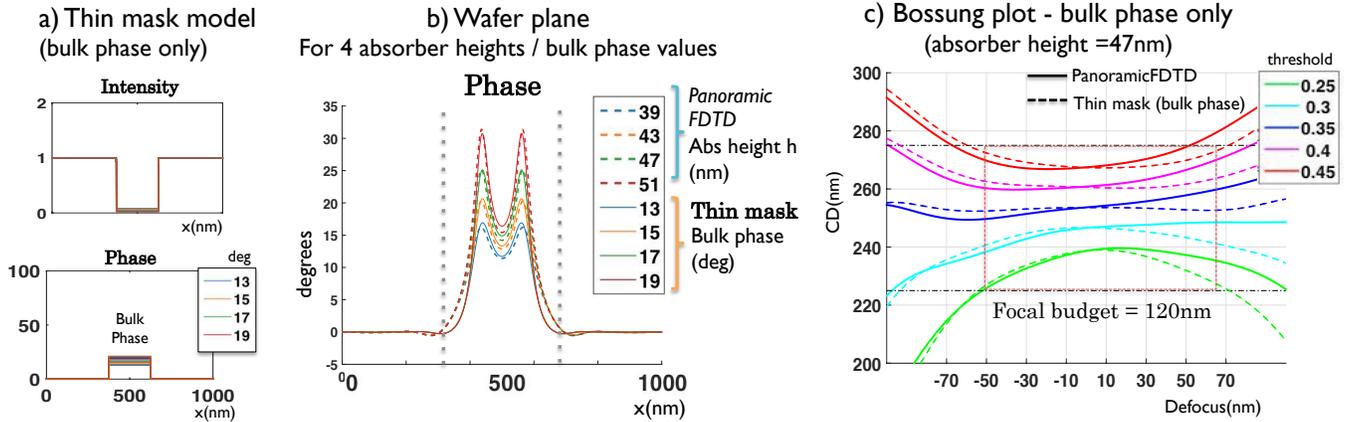
Figure 2.6: Thin mask simulations without *edge phase* modeled do not fit well to wafer phase calculated with rigorous simulations. a) The thin-mask *bulk phase* is varied linearly as the absorber gets thicker, ignoring edge phase effects. b) The wafer-plane phase matches the phase swing of rigorous simulations, but does not match at the feature edges (dotted lines). c) The Bossung tilt is not captured by the thin mask model (dashed curves) since *edge phase* has not been modeled.

Bulk phase is first modeled in the thin mask as a constant phase across the feature (Fig. 2.6a). After imaging, the phase at the wafer is compared to the EMF result in Fig. 2.6b. While the peak phase swing can be matched to rigorous simulation by changing the *bulk phase* linearly with absorber height, notice that at the feature boundary, the widths of the phase profiles do not match. In fact, there exists no value of *bulk phase* on the thin mask that can provide a good match. The addition of phase at the edges is necessary for a good fit. Additionally the Bossung tilt seen in rigorous simulation is not replicated by this thin mask when imaged to the wafer (Fig. 2.6c), unless *edge phase* is included in the model.

Next, we show that it is possible to obtain a good fit to EMF simulations if both *bulk phase* and *edge phase* are included on the thin mask. The *edge phase* is modeled as thin mask boundary layers at the edge of the absorbing feature. Figure 2.7 shows the same fitting as Fig. 2.6, but with an 8nm phase, imaginary-valued ($1\angle 90°$) boundary layer added to the edges of the feature at the thin mask. The boundary layer is of quadrature phase (90°)
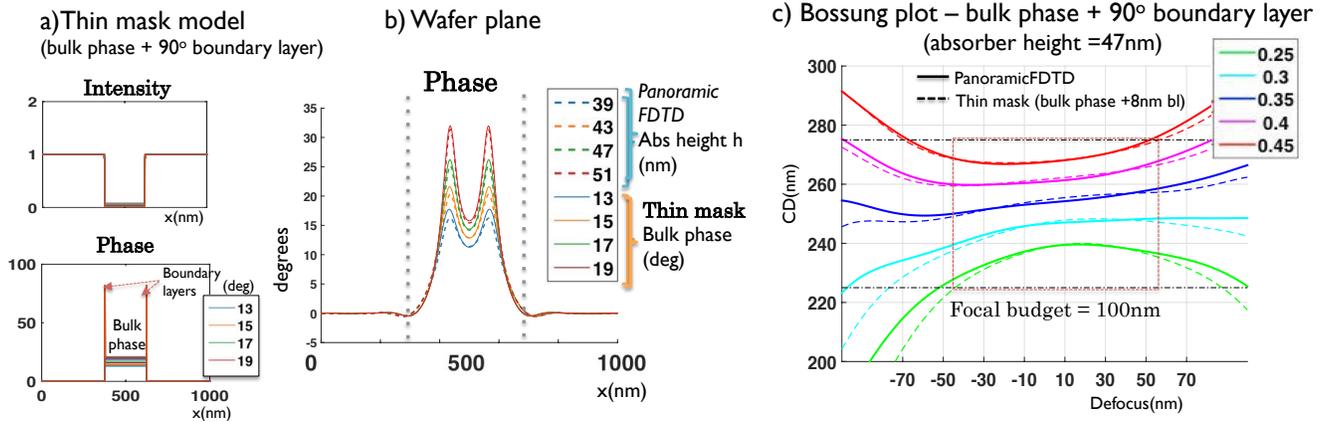
Figure 2.7: Thin mask simulations with both *bulk phase* and *edge phase* effects fit well to wafer-plane phase from rigorous simulations. a) Thin mask model with 8nm wide, imaginary-valued (90°) boundary layers added at the feature edge to represent *edge phase*. b) The phase profiles are now well replicated when compared to the rigorous result, accurately capturing the phase gradients at the feature edge. c) Bossung tilt depends on phase gradient at edges, and is now matched.

since both 0° and 180° correspond to a real-valued mask, with no Bossung tilt. Irrespective of absorber height, the same width and value of the additive boundary layer fits the phase profile, widening the thin mask wafer-plane phase to achieve an overlap with the rigorous result (Fig. 2.7b). Moreover, the same boundary layer is able to replicate the Bossung tilt as well (Fig. 2.7c), since the wafer phase gradients are now matched - in this example the tilt leads to a loss of about 20nm in the focal budget. Hence the *edge phase* model can be used to accurately fit to rigorous simulations, whereas the *bulk phase* only model cannot - we infer that phase due to thick mask edge effects (*edge phase*) is critical for modeling mask topography impact on process window.

Two key observations should be noted. First, the boundary layer, and hence edge diffraction, are independent of absorber height within the 10% range simulated here. This explains why the Bossung plot tilt is independent of absorber height fluctuations in EMF simulation (thick lines in Fig. 2.5c) , and is matched with a constant boundary layer for each case (dotted lines in Fig. 2.5c) . Second, the boundary layers representing edge diffraction influences

the width of the wafer phase profile at feature edges, *but not its peak swing.* It is the width of the phase bumps that matters for determining the process window, since they determine the phase gradients $(d\phi/dx)$ at the feature edge, which affect the Bossung tilt (dCD/dz), as is shown by the CD-TIE (Eq. 5.4).

$$\frac{dCD}{dz} = -\frac{\lambda}{2\pi}[\frac{d\phi}{dx} - \frac{I}{dI/dx}\frac{d^2\phi}{dx^2}] \tag{2.6}$$

The effect of increasing absorber height (and hence decreasing net absorption) is to move the CD upwards for the same threshold, and hence the whole Bossung plot moves upward with *no change in Bossung tilt.* Figure 2.5c plots the Bossungs for three different absorber heights on the same plot. The is also predicted by CD-TIE, which says that Bossung tilt depends on phase gradients at feature edge; the phase gradients at the feature edge are identical for the various absorber heights (see Fig. 2.5b). Note that this will move the best focus linearly with absorber height (as reported by Finders et. al. [26]), since the peaks also move laterally when the tilted curves move up. However, since the Bossung tilt is unaffected, a simple redefinition of the exposure/threshold range will recover the original Bossung plot and process conditions.

Since the process window depends on the Bossung tilt, as seen in Fig. 2.4, the phase at the edge of the absorber, when non-zero, affects the Bossung tilt, dCD/dz according to the CDTIE above. The Bossung tilt has a direct impact on the focus budget and the process window. Consider again the Bossung plots for two situations- with and without boundary layer - in Figs. 2.7c and 2.6c, for a fixed absorber thickness. The Bossung tilt predicted by rigorous Panoramic simulations is not achieved by the thin mask model with bulk phase effects only (Fig. 2.6b), but does fit nicely when the boundary layers are added (Fig. 2.7b). More specifically, the CD-TIE predicts this Bossung tilt as the first and second derivatives of the phase at the feature threshold.(Eqn. 5.4).

These findings can be further related to previous investigations by Finders[26] and Erdmann et. al.[17], which showed that the phase of the 0th order in the pupil affects the best

focus at the aerial image for dense line-space patterns. Although we are looking only at an isolated contact here, the phase of the 0th order, if thought of as the average phase across the near field, would nevertheless correspond one-to-one with the *bulk phase* through the contact. Larger *bulk phase*, which consequently means smaller absorption, will shift the CD up, and (if there is a Bossung tilt) will also cause best focus to walk for various thresholds. Another interesting conclusion is that in the hypothetical absence of edge diffraction, the Bossungs will not be tilted, and despite the CD curves moving up/down, best-focus would remain unaffected.

2.5 Impact of absorber sidewall angle on aerial imaging
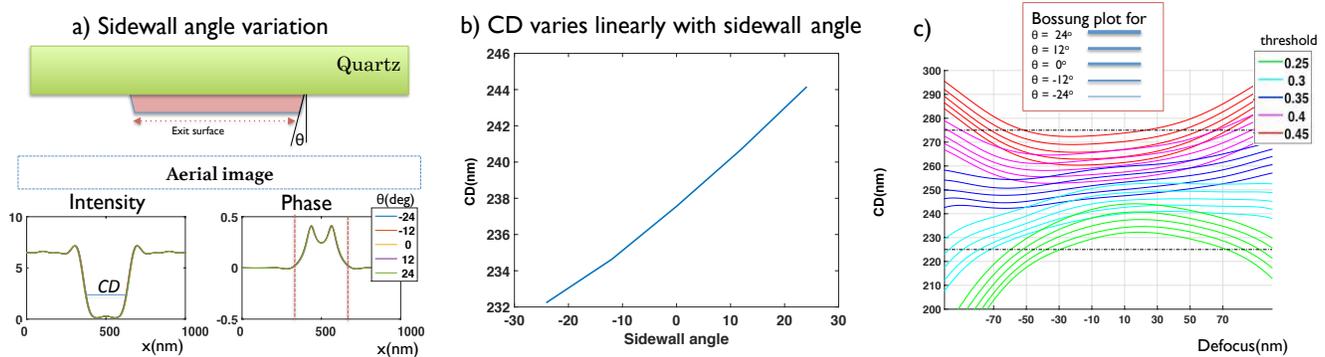
Figure 2.8: Sidewall angle changes the CD at the wafer, but leaves the phase and Bossung tilt unaffected. a) Sidewall angle change causes no variation in phase (overlapping curves). b) The CD depends on the size of the exit surface of the absorber, and hence gives a linear increase with sidewall angle. c) The Bossungs shift up due to the increased CD with increasing sidewall angle, which may affect Bossung maxima (best-focus) but not Bossung tilt (or process window).

The next aspect of mask topography studied is the sidewall angle. The sidewall angle is varied by modifying the width of the exit surface of the absorber stack (bottom surface in Fig. 2.8a), over a range representing sidewall angles from $-25°$ to $25°$. The aerial image
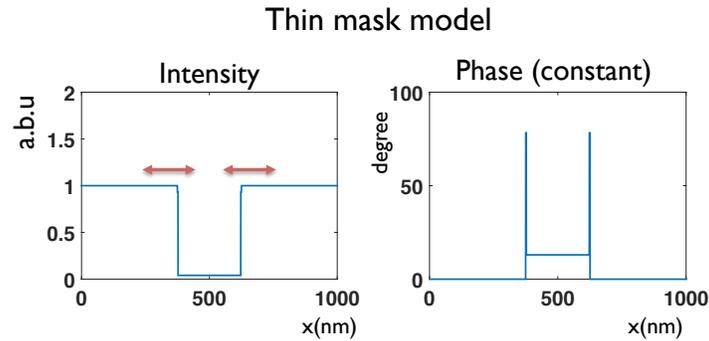
Figure 2.9: Thin mask model accounting for sidewall angle changes needs only the near-field absorption edges to be shifted to account for CD change with sidewall angle. The phase modeled as *bulk phase*+boundary layer, remains unaltered.

is then generated by low-pass filtering the near-fields output from the EMF simulator, as before. Results show that the wafer phase is unaffected by sidewall angle changes, with the phase curves in Fig. 2.8a overlapping.

The CD defined at a fixed threshold (0.35 here, chosen at the isofocal point), however, varies linearly with sidewall angle (Fig. 2.8b), indicating that the CD at the aerial image is a direct outcome of the width of the exit surface of the absorber stack. The Bossung plot of Fig. 2.8c shows a simple shift up as the sidewall goes from over-cut to undercut, a similar behavior to the Bossung shift with absorber height. This is a result of the CD increasing linearly with sidewall angle (similar to CD vs. absorber height, which also gives a linear relationship for a given threshold). However, since the wafer phase is unaffected, the thin mask model to account for sidewall angle needs only the absorption edge to be shifted on the mask in order to adjust the hole size and hence modulate the CD accordingly. The boundary layers (8nm @ $1\angle 90°$) serving to model the *edge phase*, as well as the *bulk phase*, stay constant in the thin mask. Hence, the sidewall angle, although changing the position of the feature edge, does not affect the wafer phase, Bossung tilt, or process window for an isolated feature. (Fig. 2.9).

2.6 Impact of feature size on aerial imaging

It is natural to ask: does the constancy of edge diffraction (*edge phase*) and the linear dependence of *bulk phase* on absorber height hold for a smaller feature size? We show here that it does. Figure 2.10 shows simulations with a smaller, resolution-limited feature (105nm contact, $\lambda = 193nm$, NA =1.4) than in previous plots. We again use a boundary layer model with 8nm imaginary-valued phase edges, which is able to successfully match the Bossung curve tilt for a given absorber height. This demonstration implies that edge diffraction is also independent of feature size, as predicted. Note that the Bossung tilt for this smaller feature is a smaller fraction of the Bossung curvature, so becomes less noticeable than that for the larger feature studied earlier.
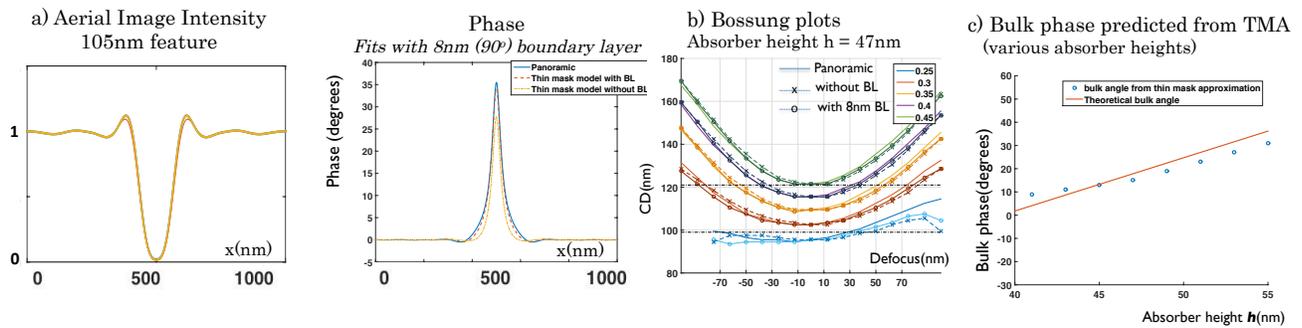
Figure 2.10: *Edge phase* is invariant with feature size and absorber height, although *bulk phase* still varies linearly with absorber height when looking at a smaller feature. For the smaller feature, the same boundary layer model is able to predict a) the phase at the wafer plane and b) the Bossung tilt for a given absorber height. c) The *bulk phase* needed in the thin mask model to fit to rigorous simulation matches the theoretical bulk phase calculated from the absorber height.

The *bulk phase* in the thin mask model that is required for achieving a good fit with the wafer phase (plotted in Fig. 2.10c) for different absorber heights shows a good agreement with the *bulk phase* calculated from the optical path for a single pass through the absorber ($\frac{2\pi}{\lambda} n \times h$). Hence, the invariance of edge diffraction, as well as the linear dependence of phase swing on *bulk phase* $\frac{dI}{dz}$, $\frac{1}{k}$, holds for features at least down to the resolution limit. This statement

may break down if illumination differs from the monopole used in the current study. Off-axis illumination will change the effective phase edge effects at the wafer plane, since each source point passes through the thick mask at a different angle. Thus, off-axis illumination with non-symmetric sources requires a modification of the left vs. right boundary layers, as analyzed in our previous work[73]. Quantifying the exact source dependence of the conclusions made here is left for future studies.

experiments on phase-shifting masks (ATT-PSM) with 90nm contact holes, each mask having a different absorber height and sidewall angle, which have been measured independently using Atomic Force Microscopy. All AIMS measurements were taken by collaborators Martin Scyzrba and Brid Connolly at Toppan Photomask, Dresden.

The topographical sidewall angle and absorber height are correlated with the phase recovered from the AIMS images in order to explore dependencies.
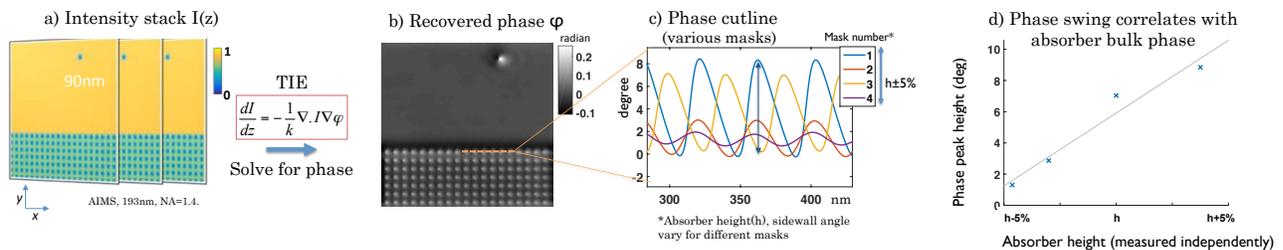
Figure 2.11: Validating dependence of measured phase swing on *bulk phase* for a set of phase-shifting masks with 90nm contact holes. a) Though-focus intensity images on an AIMS tool are used to calculate phase by solving the TIE. b) The recovered phase of the absorber at the wafer (modulo π). c) Cutlines of the recovered phase for various masks shows that the phase swings by different amounts for each case. d) The phase swing directly correlates with absorber height and hence *bulk phase* and no correlation is found with sidewall angle, confirming the dependences seen in simulation.

To recover the phase at the wafer, a stack of through-focus intensity measurements for each mask is captured in the Deep UV AIMS^{TM} tool, with quadrupole illumination but low sigma values (outer $\sigma < 0.5$). A modified version of the TIE (Eq. 8.3) is solved in order to recover the phase at-focus from this dataset (Fig. 2.11b) using an iterative algorithm[86, 75] that has the added advantage of being robust to reasonable amounts of partial coherence. Cutlines of the phase for each mask show a different peak phase swing in every case, despite the sidewall angle and absorber height (and hence *bulk phase*) varying between datasets. Simulations predict that the peak phase should increase only with the *bulk phase* (and hence absorber height), the *edge phase* being constant irrespective of absorber height. Furthermore, changes in the sidewall angle should not affect wafer-plane phase either. This in indeed confirmed by looking at the correlation of absorber height (and hence *bulk phase*) and peak phase swing in Fig. 2.11c. A good linear fit is seen across four different masks, confirming the trend predicted by simulation; no correlation is found with sidewall angle. Note that our experiments were done with phase-shifting masks, while studies in earlier sections used OMOG masks, confirming that the results hold for both mask types.

2.8 Discussion

Thick mask diffraction effects influence intensity and phase at the wafer, impacting the process window. The two major contributors to unwanted phase at the wafer - *bulk phase* and *edge phase* - affect aerial imaging differently. While the *bulk phase* changes the phase swing at the aerial image, it does not affect phase gradients at the feature edge, so will not affect Bossung tilt. *Bulk phase* increases linearly with absorber height, but does not change with sidewall angle. The *edge phase* due to diffraction from mask sidewall, on the other hand, is independent of absorber height or sidewall angle fluctuations, but is responsible for the phase gradients at feature edges and hence is the main cause of tilting in the Bossung plots. In each case, an 8nm wide imaginary-valued thin mask boundary layer was found to approximate edge diffraction well, independent of feature size (at least down to the resolution

limit).

It might be conjectured that the edge diffraction is insensitive to absorber thickness because the phase contribution from the edges is due to the break in the wavefront caused by the top and bottom corners of the absorber. Since the corners stay the same irrespective of absorber height, the edge diffraction would also be indifferent; the remaining change in sidewall length is already included in the *bulk phase* changes of the absorber.

The current study does have limitations. The topographical model used represents the absorber as a simple trapezoid with only two degrees of freedom, whereas actual thick mask profiles are more complex. Additionally, the effect of over-etch into the quartz, which would cause a two-level absorber stack (part quartz, part OMOG), has not been considered. Furthermore, the concept of *bulk phase*, i.e. phase due to a single pass through the absorber area, might need refining, since multiple reflections between the interfaces, which are ignored here, may cause multiple passes through the absorber before transmission.

Nevertheless, the broad principles found here could potentially be extrapolated to more sophisticated mask models, off-axis illumination and dense features. The sensitivity of the phase peaks at the wafer to absorber height at the mask opens up the potential for using phase imaging as a metrology method for investigating the phase of the absorber across a given mask, and for different masks. Another key finding is the constancy of edge diffraction (*edge phase*), which means that modeling edge effects in thick masks is as simple as adding phase corrections along the edges, which will be mostly independent of absorber topography fluctuations. This is instructive for modeling thick mask effects on dense masks with tall absorber stacks (such as in EUV masks), where fast simulation models are needed despite large topography variations.

To enable validation of the boundary layer model shown here in a experimental setting, the next chapter further develops the phase imaging algorithm for imaging the diffraction at the boundaries of an isolated OMOG feature, measured on an industrial aerial imaging tool.

2.9 Addendum: surface plasmons on mask sidewalls

Beyond the scope of this chapter describing phase arising at mask sidewalls, but of great relevance and interest from a phenomenological perspective, are surface modes that form at the air metal interface. This 'plasmon-polaritons' have high wavenumbers that's a combination of the electromagnetic wave number in air and metal (being greater than either one), spreading on the interface due to localization on the metal side by high attenuation, and evanescence in both the air and metal due to the larger wavenumber. The epistemology of plasmons is often labelled 'metal optics'.

Plasmons at metal-air interface are also observed for our OMOG absorber in DUV lithography, albeit the effect is somewhat mitigated since their resonance lengths are longer than the OMOG mask thickness. Rigorous EMF simulations attempt to visualize plasmonics at the absorber sidewall, and their polarization dependence (they exist only for the orthogonal polarization), in the following figures;
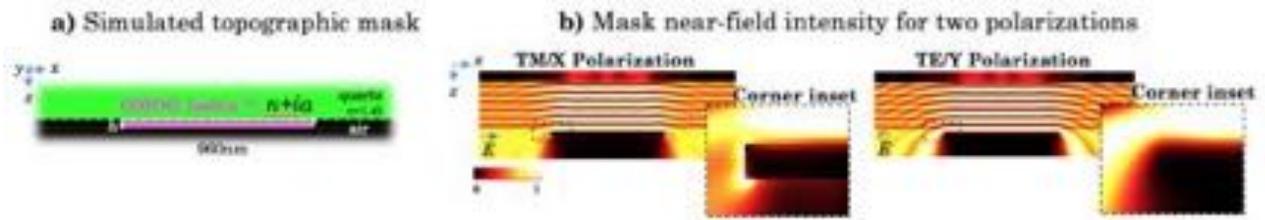
Figure Near-field rigorous simulations of a 960nm attenuating contact showing high absorption effects at the air-metal interface. **a)** Vertical cross section of the simulated mask; absorber has thickness h and complex refractive index n+ia, with n =$_{quartz}$ to minimize back-reflections. **b)** Instantaneous near-field intensities from rigorous simulation with a = 2n show the presence of surface fields at the corner of the test feature for TM polarization due to plasmonic effects arising for a > n .

Figure 2.12:
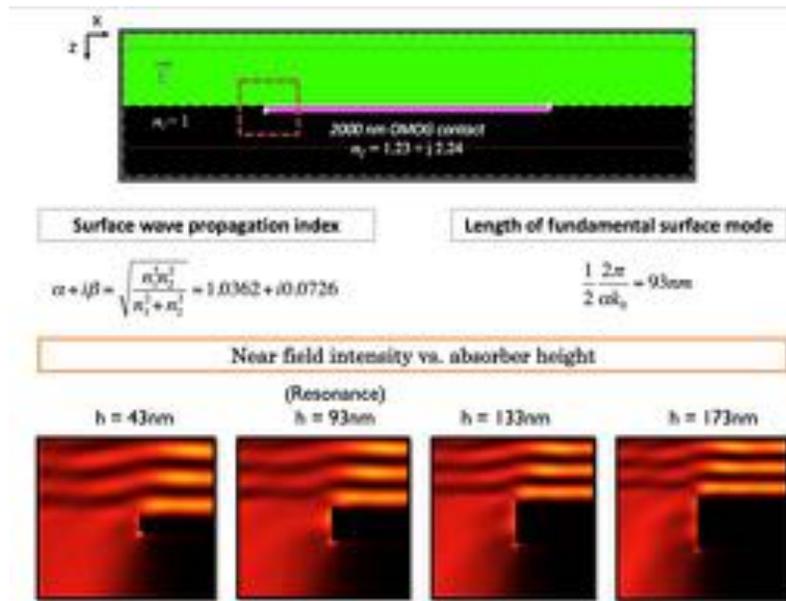
Figure 2.13: Absorber height dependence of surface plasmon resonance; depending on the height of the absorber sidewall, the first or second harmonic of the fundamental plasmon frequency can be seen on the sidewall for TM polarization.
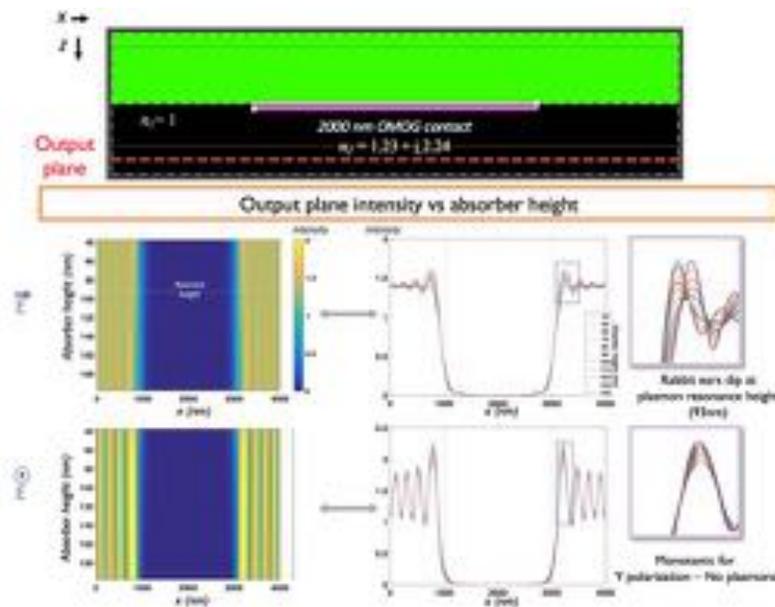
Figure 2.14: Near fields at various absorber heights show the presence of surface plasmon polaritons only for polarization normal to the mask sidewall
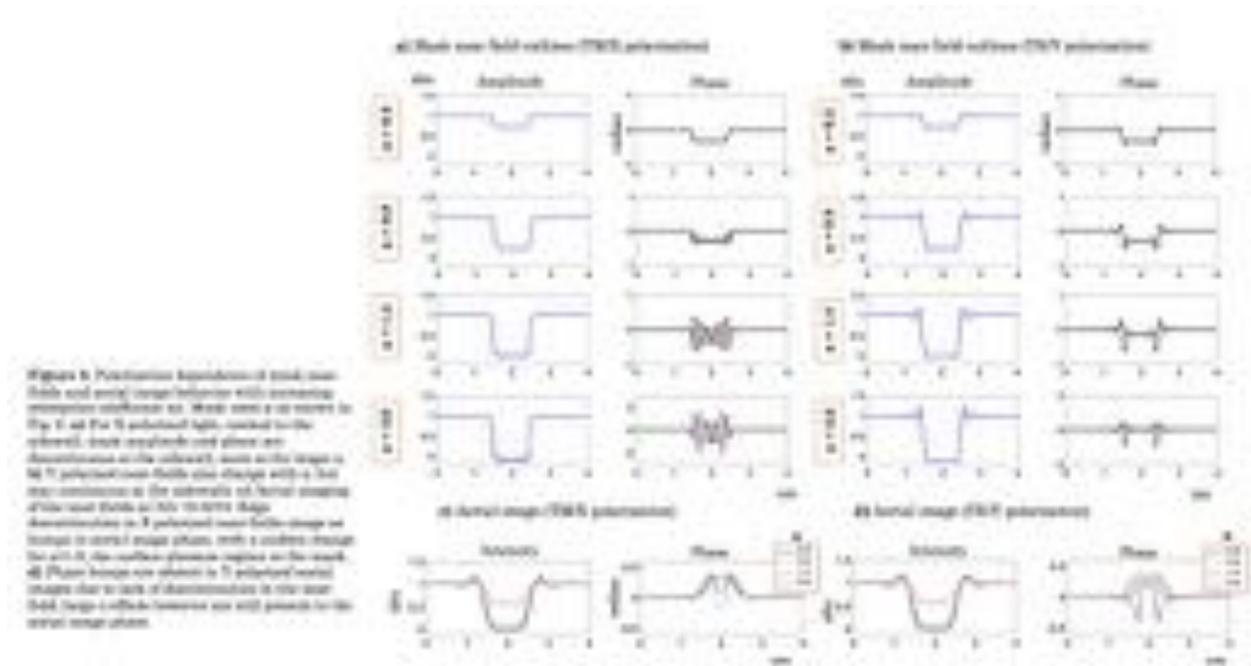
Figure 2.15: Absorption coefficient dependence of the phase due to the mask absorber, the plasmonic regime sets in only for high enough absorption $\alpha > 1.8$ and TM polarization.

Chapter 3

Optical angular momentum : Phase imaging curl at photomask corners

3.1 Introduction

Now we describe a Transport of Intensity based phase image technique for absorbing photomasks in the DUV regime, in order to measure the phase effects arising from mask topography as described in Chapter 2. TIE based phase imaging is easy to implement in litho aerial imaging systems since it only needs defocus measurements; however we ran into algorithmic challenges when imaging the phase near corners where light flow has angular momentum, requiring an extension of the TIE solver as described here.

The Transport of Intensity Equation (TIE) describes how phase can be recovered from intensity images captured at different focus positions. Its experimental simplicity makes it amenable to existing microscopes in optical [81, 80, 9, 8, 97, 47], X-ray [57, 56] and electron[10, 62, 42] imaging. Sub-wavelength phase accuracy and real-time processing [49] are routinely achieved and errors can be reduced with multiple images [78, 87, 97]. One particularly convenient advantage of the TIE method is that it is fairly robust under partially coherent illumination [60, 39], making it suitable for lithography aerial imaging tools, which

we use here.

The TIE is directly derived from the paraxial wave equation to relate intensity variations over small defocus distances to gradients of phase as light propagates [81]. Being a partial differential equation, solving the TIE involves inverting the equation to recover phase. The traditional TIE solver has an implicit assumption that any curl component in the power flow is not captured in the intensity data. Thus, the TIE is thought to only recover the 'scalar' part of the phase and fails for the 'rotational' component [60, 37]. In fact, the rotational (curl) component does affect the through-focus intensity and therefore causes phase artifacts in the traditional TIE solver. A standard example of a wave-field with a curl is a phase vortex. This class of curl components has been studied in detail, and it was shown empirically that phase vortices can be recovered by either an iterative algorithm that use many images through-focus [3] or by modifying the traditional TIE solver [50], assuming small intensity gradients [85].

Here, we discuss a different class of curl effects (distinct from phase vortices) which arise in our application and any situation where phase gradients are not collinear with intensity gradients. This case has been studied theoretically [68, 21], but a solution to the phase recovery problem was not presented. Here, we study the non-physical phase recovery artifacts resulting from curl components induced by strong absorption. We then propose an iterative wrapper for the traditional TIE solver that corrects such errors to produce a more accurate phase result. This method was first described in [74, 70] and later appeared in [96]. We further derive how the curl component is coupled into the intensity measurements, and show that our proposed method also recovers part of the missing curl.

This work is motivated by our studies of electromagnetic phase edges in optical lithography masks [74, 70]. Unlike many other applications, photomasks are designed specifically to have strong absorption. An ideal mask would be infinitesimally thin with no phase variations. Since material constraints mean that real-world masks are thick relative to the size of the feature, the optical field incurs unwanted 3D diffraction effects as it passes through the mask, termed 'electromagnetic edge effects' [1, 84, 5] (Fig. 3.1). The result is that the
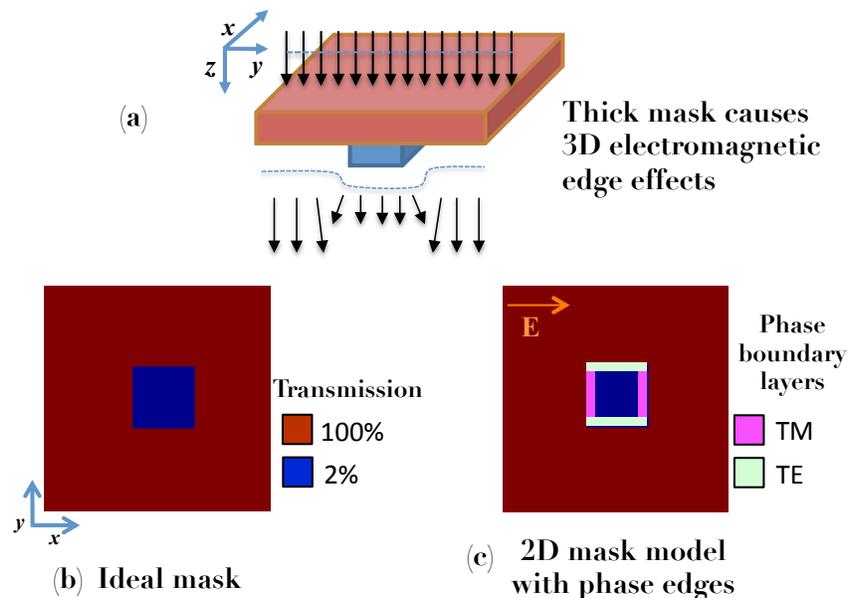
Figure 3.1: Photolithography masks incur polarization-dependent electromagnetic edge effects. (a) Because the mask is relatively thick, the electric field accumulates an unwanted phase delay at feature edges, due to diffraction. (b) An ideal mask has only absorption variations. (c) Phase edge effects can be modeled by phase strips at the feature edges, depending on polarization [1, 84, 5].

field exiting the mask has added complex-field variations near feature edges. The real component is easily measured as a line edge placement bias, but the phase component produces an asymmetrical, feature-dependent edge placement change through-focus which is difficult to measure. As node sizes shrink, these undesired phase effects become more prominent and also more problematic, reducing the process window.

Mask designers often account for electromagnetic effects using an equivalent thin mask model, which replaces the complicated 3D effects with a 2D complex field at the exit plane of the mask. For example, boundary layer models represent the added phase effects with quadrature (90°) phase strips along the feature edges [1, 84, 5], where the width of the strip depends on the mask shape and material. The phase edges are also polarization-dependent, being much stronger in the direction perpendicular to the electric field, so separate boundary

layers must be used for each polarization.

In this work, we aim to physically measure phase edge effects using the TIE method in an aerial inspection tool. However, phase variations always occur at the edges of features, where intensity changes rapidly. Since the gradients of the intensity and phase are not collinear, significant curl components result near feature corners, and we must correct the artifacts in order to recover phase accurately.

3.2 TIE solvers and curl effects

First, we describe the traditional TIE solver and derive how absorption-induced curl can produce errors in the phase result. For a 2D complex object $\sqrt{I}e^{i\phi}$ with intensity I and phase ϕ, the TIE describes the change of axial intensity as a divergence of the in-plane power flow [81, 80],

$$\frac{\mathrm{d}I}{\mathrm{d}z} = -\frac{\lambda}{2\pi}\vec{\nabla} \cdot I\vec{\nabla}\phi, \tag{3.1}$$

where λ is wavelength, $\vec{\nabla}$ is the lateral gradient and z is defocus distance. Thus, one can solve for phase after estimating the intensity derivative $\mathrm{d}I/\mathrm{d}z$ from two or more intensity images at different z.

Since $\vec{\nabla}$, I and ϕ are in-plane, $I\vec{\nabla}\phi$ is the in-plane Poynting vector [60]. To solve Eq. (3.1), Teague's solver [81] defines an auxiliary variable ψ such that $I\vec{\nabla}\phi = \vec{\nabla}\psi$, which converts the TIE into a Poisson equation,

$$\frac{\mathrm{d}I}{\mathrm{d}z} = -\frac{\lambda}{2\pi}\nabla^2\psi. \tag{3.2}$$

Equation (3.2) can use any Poisson solver (e.g. in Fourier domain [38]) to solve for the auxiliary variable ψ. Substituting this value back into its relation with phase gives $\vec{\nabla}\phi = \vec{\nabla}\psi/I$, and taking another divergence yields a Poisson equation in phase,

$$\vec{\nabla} \cdot \left(\vec{\nabla}\psi/I\right) = \nabla^2\phi, \tag{3.3}$$

from which a second Poisson solver can recover the final phase $\phi(x, y)$ to within an unimportant constant offset. This two-step solution is required for objects with non-uniform

absorption. For the case of constant intensity (e.g. a pure phase object), $I(x, y) = I_0$, and
the phase is recovered directly after solving the first Poisson equation, $\phi = \psi/I_0$.

It has been noted [68, 21] that substituting the Poynting vector with the gradient of a
scalar field [to obtain Eq. (3.2)] makes the implicit assumption that the Poynting vector is
curl-free, requiring collinearity of phase and intensity gradients:

$$\vec{\nabla} \times (I\vec{\nabla}\phi) = \vec{\nabla}I \times \vec{\nabla}\phi = 0. \tag{3.4}$$

As described previously, photomasks have both strong absorption and phase at the edges
of features, resulting in a significant curl component (i.e. non-collinear phase and intensity
gradients). This is illustrated in Fig. 3.2 for a simulated OMOG (Opaque MoSi on Silica)
type mask with a 2% transmitting block on a clear background. To model electromagnetic
edge effects for horizontally polarized illumination, we add a boundary layer of width = 20nm
with 90° quadrature phase along the vertical direction. In this simulation, parameters have
been chosen to match those of our experiment, described later. We use deep UV wavelength
$\lambda = 193nm$, NA 1.35 at the wafer (0.3375 at the mask) and illumination coherence $\sigma = 0.3$.
Since the phase strips are much smaller than the resolution of the optical system, they
become blurred and result in smaller peak phase values. Here, we directly observe that the
gradients of intensity and phase are non-collinear at the corners, leading to non-negligible
curl for the in-plane Poynting vector, $\vec{\nabla}I \times \vec{\nabla}\phi \neq 0$. The curl components for this simulated
mask are shown in Fig. 3.2, along with the phase recovered by Teague's solver (using the
Poisson solver in [38, 86]). This phase result incurs significant error, due to curl effects.

To calculate the phase error due to the missing curl, consider the Helmholtz decomposi-
tion of the Poynting vector, with curl-free and divergence-free source terms [60],

$$I\vec{\nabla}\phi = \vec{\nabla}\psi + \vec{\nabla} \times \vec{\mathbf{A}}_1, \tag{3.5}$$

where ψ and $\vec{\mathbf{A}}_1$ are the scalar and vector potentials of the power flow. Since the TIE
describes the divergence of the Poynting vector, the first Poisson equation [Eq. (3.2)] of
Teague's solution is uniquely and exactly solved for the scalar potential ψ, given appropriate

boundary conditions. In the presence of a vector potential for the Poynting vector [Eq. (3.5)],

however, the second Poisson equation in Teague's solution [Eq. (3.3)] has an extra term due

to the curl,
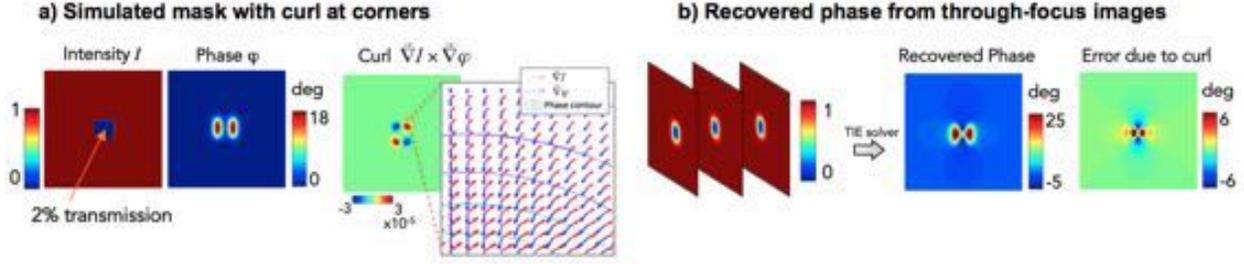
Figure 3.2: (Left) Simulation of a 240nm square absorbing feature on a photomask with
phase edges added along the vertical sides, causing non-zero curl in the Poynting vector
near feature corners, where phase gradient is tangential to intensity contours. (Right) When
through-focus images are simulated for this complex field and used as input to the traditional
TIE solver, the recovered phase suffers serious errors due to the curl effects.

$$\vec{\nabla} \cdot \frac{\vec{\nabla}\psi}{I} + \vec{\nabla} \cdot \frac{\vec{\nabla} \times \vec{\mathbf{A}}_1}{I} \quad = \nabla^2 \phi \tag{3.6}$$

$$\Rightarrow \nabla^2 \phi_{\text{TIE}} + \nabla^2 \phi_{\text{res}} \quad = \nabla^2 \phi, \tag{3.7}$$

where ϕ_{TIE} is the phase returned by Teague's solver and ϕ_{res} is the residual error that occurs

due to the curl component, shown in Fig. 3.2 to create a severe saddle-shaped artifact for

the square feature.

3.3   Recovering curl by iterative TIE

Next, we describe our algorithm and prove analytically that curl components of the

power flow are not entirely lost in through-focus measurements, and can thus be recovered

computationally. We demonstrate this on the simulated photomask described above, as well

as experimental measurements, showing its efficacy in the presence of strong absorption.

Our algorithm iterates back and forth through Teague's solver, estimating both the phase
and the curl component at each step. In the first step, we obtain an initial phase estimate
using Teague's method, ϕ_{TIE}, then plug it back into the TIE to estimate the axial intensity
derivative that would have been produced by ϕ_{TIE},

$$\left.\frac{\mathrm{d}I}{\mathrm{d}z}\right|_{\mathrm{est}} = -\frac{\lambda}{2\pi}\vec{\nabla}\cdot I\vec{\nabla}\phi_{\mathrm{TIE}}. \tag{3.8}$$

The residual intensity derivative is then obtained by calculating the difference between
the estimated and measured intensity derivatives,

$$\left.\frac{\mathrm{d}I}{\mathrm{d}z}\right|_{\mathrm{res}} = \frac{\mathrm{d}I}{\mathrm{d}z} - \left.\frac{\mathrm{d}I}{\mathrm{d}z}\right|_{\mathrm{est}}, \tag{3.9}$$

and is expected to be zero in the absence of curl, notwithstanding numerical errors and noise.
The residual intensity derivative is then used as input to Teague's solver for estimating the
phase residual,

$$\left.\frac{dI}{dz}\right|_{\mathrm{res}} = -\frac{\lambda}{2\pi}\vec{\nabla}\cdot I\vec{\nabla}\phi_{r1}, \tag{3.10}$$

where ϕ_{r1} is the current estimate of the phase residual [ϕ_{res} in Eq. (3.7)], after the first
iteration. The estimated phase residual is then subtracted from the previously estimated
phase ϕ_{TIE} to give an improved phase estimate. More iterations can then be used to further
refine the result.

To understand how the curl of the Poynting vector couples into the next iteration of
Teague's solver, we examine the intensity derivative residual in Eq. (3.9). Consider the
Helmholtz decomposition of the vector field $\vec{\nabla}\psi/I$,

$$\vec{\nabla}\psi/I = \vec{\nabla}\phi_{\mathrm{TIE}} + \vec{\nabla}\times\vec{\mathbf{A}}_2, \tag{3.11}$$

where the scalar potential is simply ϕ_{TIE} according to Eq. (3.3), and the $\vec{\mathbf{A}}_2$ denotes the
vector potential. By substituting Eqs. (3.2) and (3.8) into Eq. (3.9), and considering the
relation in Eq. (3.11), we obtain
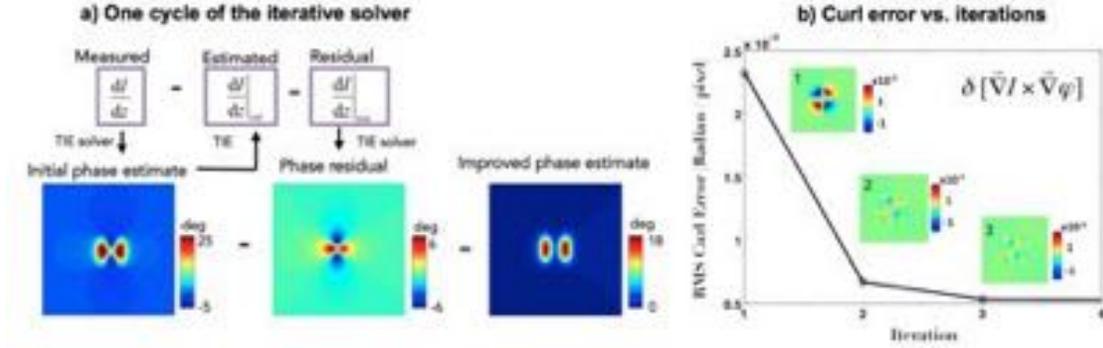
Figure 3.3: Simulation showing the first iteration of our algorithm. Teague's solver recovers
the initial phase estimate, then we plug that into the TIE to find the estimated intensity
derivative. The residual between the measured and estimated intensity derivatives is put into
Teague's solver a second time in order to estimate the phase residual, which is subtracted
from the recovered phase for an improved estimate.

$$\frac{dI}{dz}\bigg|_{\text{res}} = -\frac{\lambda}{2\pi}\{\nabla^2\psi_1 - \vec{\nabla}\cdot I\vec{\nabla}\phi_{\text{TIE}}\} \tag{3.12}$$

$$\Rightarrow \frac{dI}{dz}\bigg|_{\text{res}} = -\frac{\lambda}{2\pi}\vec{\nabla}\cdot I\vec{\nabla}\times\vec{A}_2. \tag{3.13}$$

The curl term $\vec{\nabla}\times\vec{A}_2$ is thus responsible for the derivative residual on plugging the solved

phase back into the TIE. The TIE solution of the residual intensity derivative [Eq. (3.10)]

will try to estimate the error arising due to this curl, which is in fact directly related to the

Poynting vector curl $\vec{\nabla}\times\vec{A}_1$ from Eqs. (3.5) and (3.11),

$$\nabla\times\nabla\times\vec{A}_2 = -\nabla\times\left(\frac{\nabla\times\vec{A}_1}{I}\right), \tag{3.14}$$

which holds also for vector potentials of any two successive iterations. In the absence of

curl in the power flow, $\vec{A}_1 = \vec{A}_2 = 0$, and hence the residual intensity derivative vanishes

such that the solution converges immediately. In the presence of curl, however, subsequent

iterations will recover some of the curl missed in the previous iteration, the solution reaching

convergence when the estimate of the phase gradient at the ith iteration, $\vec{\nabla}\psi_i/I$, approaches

zero curl, i.e. $\vec{A}_{i+1} \to 0$. A more rigorous formulation of the convergence criteria would have

to include the interplay of the object curl, numerical and focus sampling and the severity of the curl.

A simulation of the estimated phase residual from the first iteration is shown in Fig. 3.3. After only one iteration, the improved phase estimate is already very close to the true phase (shown in Fig. 3.2), with Root Mean Square (RMS) phase error having dropped by about 42%, from 0.0087 radian/pixel for Teague's solver to 0.005 radian/pixel for our iterative algorithm. Subsequent iterations further improve the estimate of the phase and its residual.

Since the reduction in the error is due to the recovery of the curl, our algorithm also produces an estimate of the curl components, which were previously considered unmeasurable. For the simulation case (where true curl components are known), we plot the error in our curl estimate as iterations progress (see Fig. 3.4). The plots compare the Poynting vector curl, $\vec{\nabla} \times I\vec{\nabla}\phi = \vec{\nabla}I \times \vec{\nabla}\phi$, for the true object with that recovered by the iterative method. As expected, the error in the curl is progressively reduced, with diminishing gains at each iteration. Notice that the error does not go to zero, since not all of the curl effects were transferred into intensity measurements. However, these unobservable areas of curl do not produce phase errors in our result. If the goal is to fully measure the curl terms, then systematic variation of the intensity would be needed [69, 94].

3.4   Experimental Results

Experiments were performed on an AIMS aerial imaging tool at AMTC/Toppan Photomasks at Dresden, Germany. The AIMS tool replicates the projection printing process, with demagnification to allow the wafer plane intensity to be captured by a camera. The experimental parameters and mask match the simulations described earlier (240nm square feature on an OMOG mask). Here, we use partially coherent illumination with $\sigma = 0.3$, which has been shown to produce accurate phase results [8], though larger or non-rotationally symmetric sources [80, 98] would require more sophisticated algorithms [63, 46, 43, 45]. Images were captured with 10nm defocus steps across a 200nm range, building up a through-focus
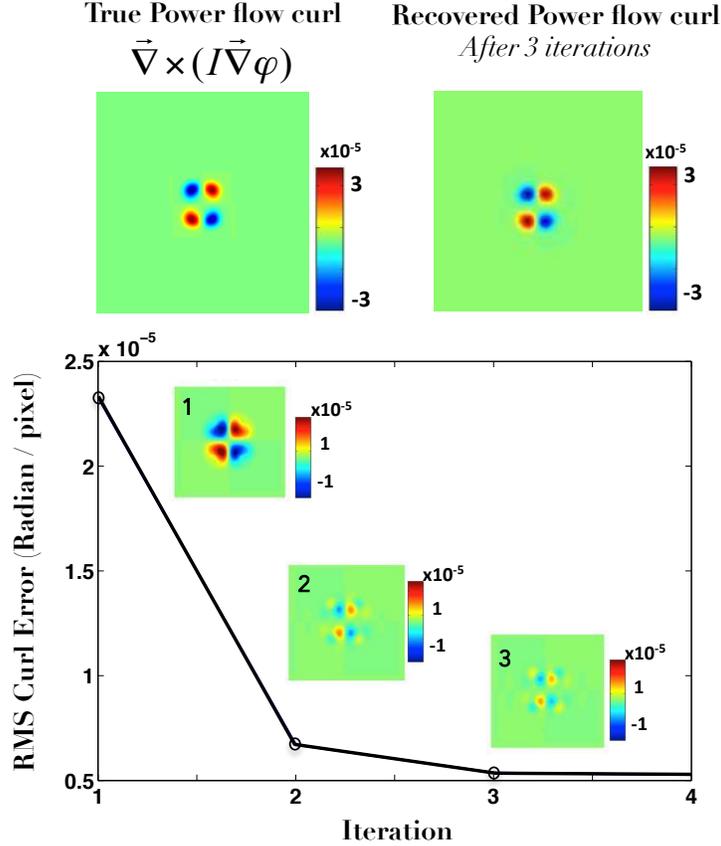
Figure 3.4: Estimating curl components with the iterative TIE. The top row shows the true curl for the simulated mask, and the corresponding curl recovered by our iterative algorithm. The bottom shows that the RMS error in our estimate of curl diminishes progressively as the algorithm iterates. The errors in the curl component estimation for the first 3 iterations of the algorithm are shown as insets.

stack. Then, the intensity derivative was calculated using the fitting methods described in [87, 47]. Note that, though the TIE equation is based on the paraxial approximation, it is justified for the mask-side NA of 0.3375 in the experiments here. More details on the experimental setup can be found in [74].

Images of the experimentally recovered phase are shown in Fig. 3.5. We clearly observe electromagnetic phase edge effects in all the results. In the experiments, we study two situations, one with horizontally polarized illumination and one with vertically polarized illumination. The recovered phase shows much stronger phase edge effects along the

direction perpendicular to the illumination polarization, as predicted by rigorous EMF simulations [53].

In comparing our iterative algorithm to the traditional TIE method, we see that the phase images recovered by Teague's solver have saddle-shaped artifacts which cause negative phase values. These clearly non-physical artifacts resemble the error seen in simulation results from earlier sections, indicating that they are indeed due to power flow curl near the corners of the square feature. With our iterative solver, however, they are removed and we get a much cleaner picture of the phase edge effects.

The results match well with rigorous simulation-based boundary layer model theory [53], which assumes that the peak phase value of 40 degrees in Fig. 3.5 is a convolution of the 90 degree boundary layer with the point spread function of the system. The smaller phase peaks for the edges parallel to the electric field in Fig. 3.5 likely indicate that OMOG is similar to ATT-PSM, in that diffraction at the primary edge reduces with the rotation of the polarization and the phase drops in magnitude by about a factor of five. A detailed analysis can be found in [74].

3.5 Conclusion

We have demonstrated both theoretically and experimentally an iterative extension to the Transport of Intensity method that provides both accurate phase recovery as well as an estimate of the Poynting vector curl. Our method is particularly useful for the situation of curl-induced artifacts due to strong absorption, for which we provide a motivating example in lithography. We show that phase edge effects due to 3D electromagnetic interactions break the curl-free assumption of the traditional TIE (Teague's method). However, by employing our iterative TIE solver, we can remove these artifacts and also solve for curl components in the process. The solution removes curl-induced phase errors with only a few iterations, providing significantly improved results without much computational overhead. Results were demonstrated for a square feature on an OMOG mask, with experimental data being
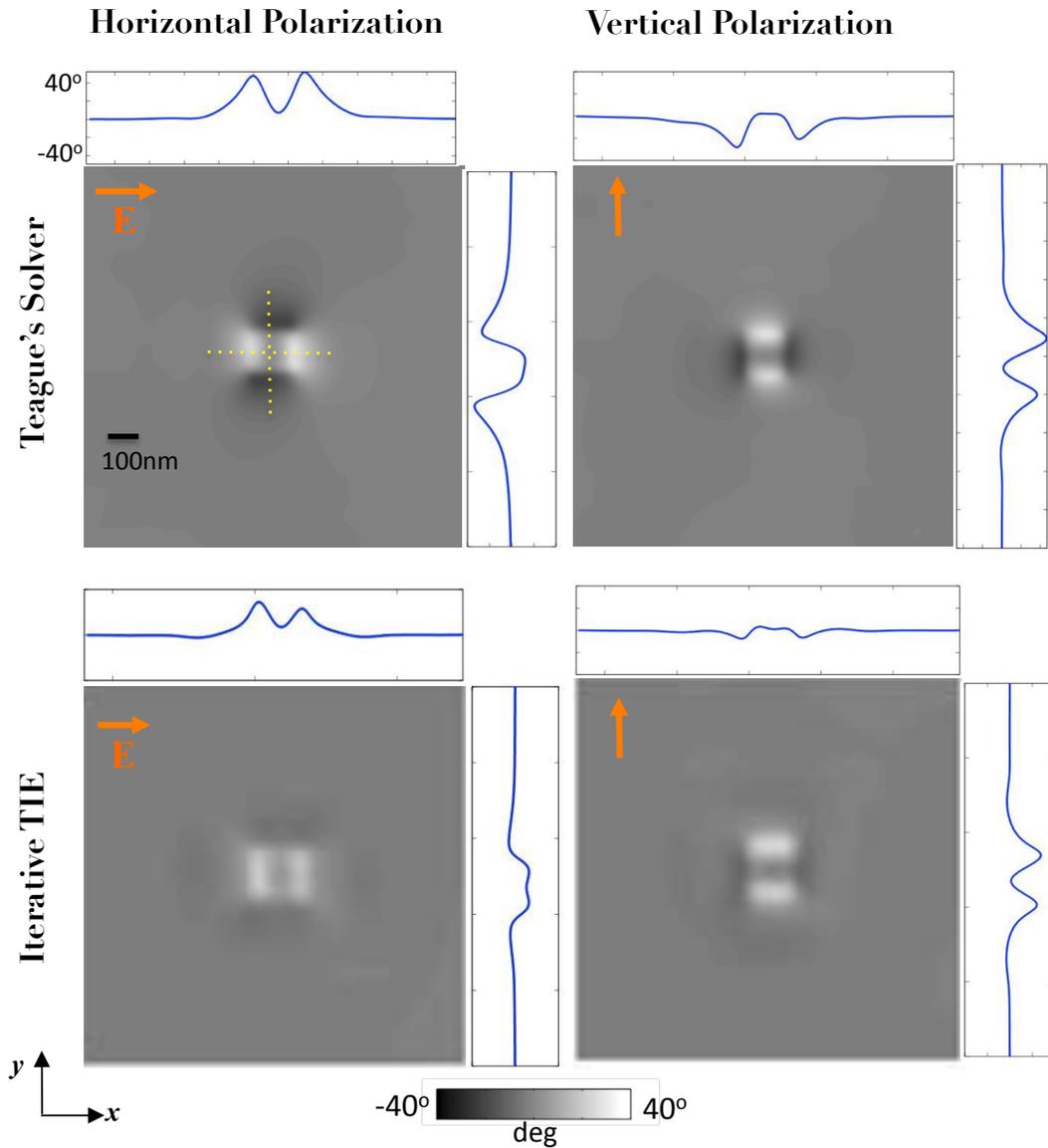
Figure 3.5: Experimental results for our iterative TIE method, as compared to the traditional Teague's solver, with a 240nm square feature on an OMOG mask. The top row shows the phase recovered by Teague's solver, with non-physical saddle artifacts due to the Poynting vector curl at the feature corners. The bottom row shows the phase recovered by the iterative solver, where artifacts have been corrected, clearly showing the presence of phase edges which match well with theoretical predictions. On the left is the result for illumination polarized in the horizontal direction, and on the right is that for the vertical direction, showing strong polarization dependence, as expected.

captured in an aerial imaging tool. The method serves to elucidate the influence of power-

flow curl on defocus based phase recovery, and should find general use in many applications,

particularly those with strong absorption at the sample.
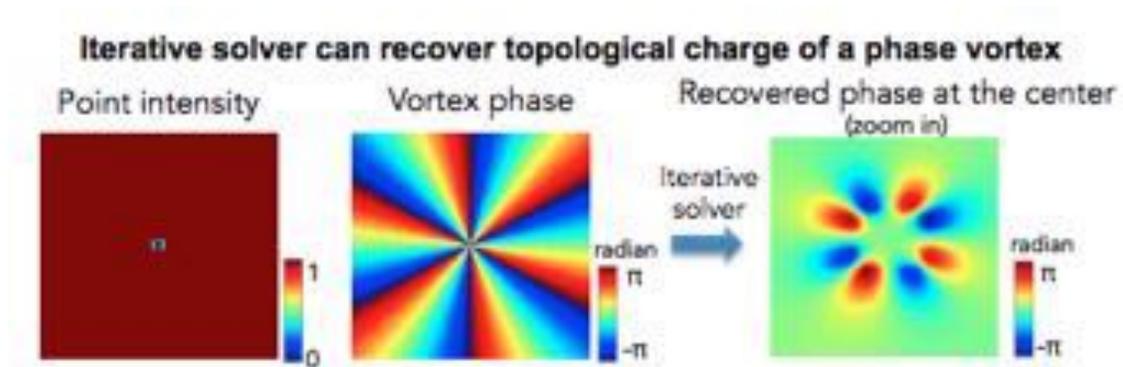
Figure 3.6: Optical power flow curl recovery using the iterative TIE, applied to simulation of
a phase vortex with topological charge 6. The iterative TIE recovers the curl from defocus
intensity measurements, but only near the position intensity variations at the center, since
the vector potential is non zero only where $\vec{\nabla} I \times \vec{\nabla} \phi \neq 0$.

Chapter 4

Illumination angle dependence of photomask edge diffraction

4.1   Introduction

Image irregularities introduced by mask topography are getting more attention as process window and critical dimension uniformity requirements get tighter at sub 30 nm lithography. [27]. While topographic edge effects are well known and typically quantified through rigorous simulations [89, 16], here we show that they can be directly deduced experimentally using through-focus intensity measurements in an aerial imaging measurement system (AIMS), even for off-axis illumination. The analysis used is quick to perform and could be integrated into the mask qualification process, enabling diagnosis of mask performance variability due to etch or material non-uniformity, or allowing comparison of topography effects between different absorber and mask types.

The binary Opaque MoSi on Silica (OMOG) mask used in this study is shown in Fig. 4.1a. It is expected to have reduced topographic effects compared to phase shifting masks and is hence a good control for testing the sensitivity of the proposed method. The square contact feature is chosen since it allows imaging four edge orientations in a single measurement.
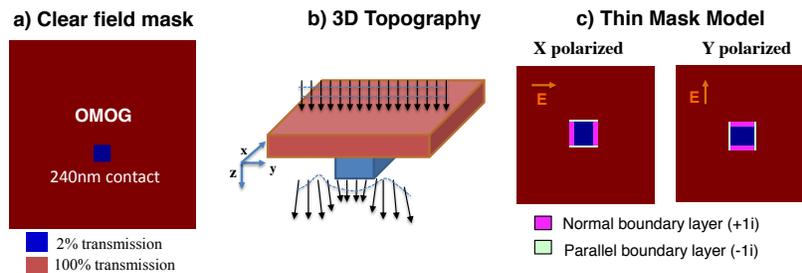
Figure 4.1: An OMOG contact on a clearfield mask is used to characterize edge effects experimentally. a) The 240nm contact is twice the size of the resolution limit, allowing each edge to be resolved at the wafer. b) The absorber has a finite height causing phase edge effects at the edges of the contact. c) Thin mask models account for edge diffraction with complex valued boundary layers, where the argument switches sign on switching polarization.

The MoSi absorber sidewall will typically alter mask near-fields at feature edges, and hence the aerial image intensity at wafer. 3D mask effects are commonly modeled using rigorous simulations by approximating edge diffraction as complex valued boundary layers on the mask (Fig. 4.1c) [53, 5], or Zernike aberrations [11] in the imaging pupil. [18, 27]. We have shown [75, 74] that edge effects can instead be visualized in experiments with an AIMS tool using phase imaging at the wafer plane (Fig. 4.2a). The tool operates at 193nm and a wafer NA of 1.4, with the option of X and Y polarizations at the source. Topography induced phase at the wafer can be recovered from through-focus measurements with a customized phase retrieval method. The edge phase was seen to be polarization dependent, being stronger for polarization normal to the sidewall (TM). Moreover, complex-valued boundary layers along feature edges were shown to replicate topographic diffraction in thin mask models, being identical for left and right-facing sidewalls.

Here, we extend the phase recovery algorithm and boundary layer modeling to off-axis illumination. Conventional lithographic processes will typically use off-axis illumination such as dipole or quadpole sources to enhance resolution and contrast for given feature pitch on mask [51]. We show that mask topography effects modify for an oblique trajectory through the MoSi stack, needing asymmetry in the thin mask model to approximate thick mask
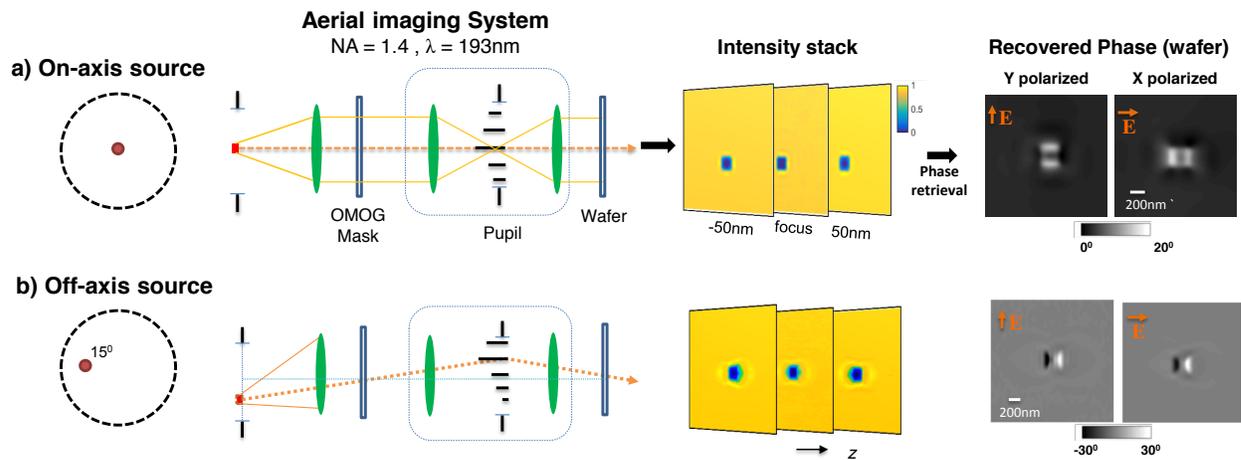
Figure 4.2: Measuring phase edge effects from through-focus AIMS images a) For on-axis illumination, a stack of through-focus intensity images is used to recover the phase at wafer plane[76]. Polarization dependence of topography induced phase is evident, in agreement with the boundary layer model of Fig 4.1c. b) AIMS tool configured for illumination with an off-axis monopole; the recovered phase is no longer directly indicative of mask diffraction since the pupil shift due to oblique incidence has a stronger phase signature than the thick mask edge phase.

diffraction. Additionally, the impact on aerial imaging will be found to be stronger for polarization parallel to the sidewall (Y/TE polarization), contrary to the normal incidence case where X/TM polarization has greater influence of aerial image quality. Interestingly, for the most general case of any non-symmetric source shape, phase edge effects will cause intensity variations not only in defocused images, but also at focus.

AIMS Measurements with an off-axis source

To experimentally investigate the impact of off-axis illumination on aerial imaging with a thick mask, the AIMS tool was configured for imaging with an shifted monopole (Fig. 4.2b). The angle of incidence at the mask was 15°, corresponding to dipole illumination optimized for 85nm pitch. Unlike on-axis illumination, mask diffraction effects for oblique incidence cannot be directly visualized from the recovered field at wafer (Fig. 4.2b). For the on-axis

case (Fig. 4.2a), the imaging system acts simply as a low-pass filter from mask to wafer. With oblique incidence, however, off-center propagation of the diffracted fields shifts their spectrum in the pupil creating an asymmetric bandpass filter (Fig. 4.4a). For a 15° angle of incidence, the shift corresponds to 0.7 of pupil bandwidth (with immersion index of 1.4), thus clipping 70% of the right half of the spectrum. Hence the phase at the wafer in Fig. 4.2b are a combination of mask diffraction and the pupil clipping, the latter being stronger.
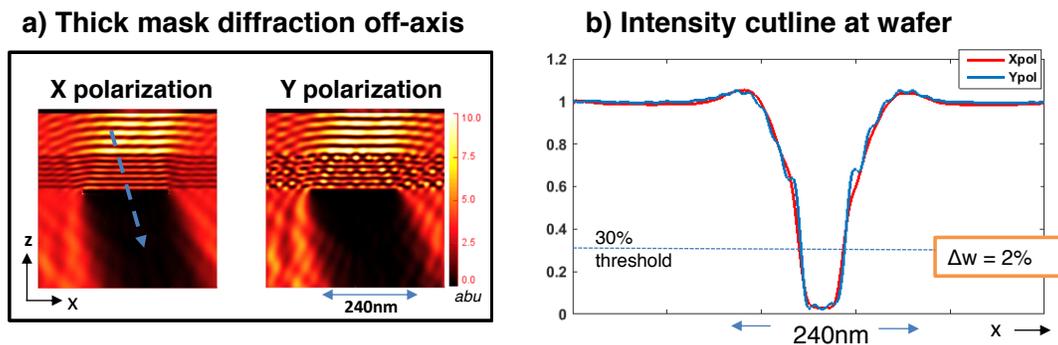
Figure 4.3: Oblique illumination at the mask modifies: a) the near-field diffraction pattern (hence the thin mask model needed to approximate the mask) as shown in a vertical slice of the mask near-field intensity computed from rigorous simulation. b) Measured intensity cutlines at focus shows stronger degradation for Y polarization due to mask diffraction.

Even though the recovered phase images cannot be directly interpreted, the at-focus intensity images already show stronger image degradation for Y-polarized light (horizontal cutline in Fig. 4.3). A photoresist with a 0.3 threshold for instance would see a 5nm (about 2%) width variation between the two polarizations. This is contrary to imaging on-axis, where mask phase effects do change through-focus behavior, but are invisible at the focal plane of the imaging system. Here the off-axis imaging introduces "phase contrast" also at-focus.

This difference in printing performance will be shown in the next few sections to arise from mask topography combined with pupil filtering in the imaging system. Edge diffraction will be found to be non-identical for left and right sidewalls under oblique illumination, along with a strong polarization dependence. Section 3 looks at the through-focus intensity

measurements for each polarization, followed by a description of the phase recovery scheme to obtain wafer and pupil fields. Finally, mask near-fields are modeled using a polarization-dependent boundary layer approximation in section 5.

4.2 Pupil induced asymmetries in through-focus intensity

Off-axis imaging can be considered in an equivalent on-axis scheme where the pupil is shifted instead of the source, as shown in Fig 4.4b.This setup is reminiscent of the Foucault knife edge test [35], where half the spectrum is blocked with a knife edge to ascertain the best focus in an imaging system. As only the positive spatial frequencies contribute to the image at focus, the defocus images have a lateral asymmetry about the optical axis which flips on either side of focus.

Since off-axis illumination shifts the diffracted spectrum in the pupil, AIMS intensity images show a defocus dependent lateral asymmetry as well (Fig. 4.5), which is absent for the on-axis case. The asymmetry flips on either side of focus for the off-axis case with both X and Y polarizations, as expected. Remarkably the lateral asymmetry is significantly more pronounced for Y polarization than for X, despite identical imaging conditions.
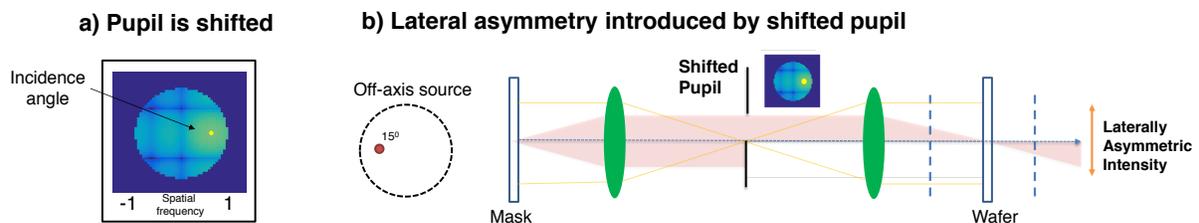
Figure 4.4: a) Tilting the illumination off-axis shifts the mask spectrum in the pupil . b) The effective pupil shift filters one side of the spectrum, causing the defocus intensity to acquire a lateral asymmetry about the optical axis that flips on either side of focus.

Hence looking at the defocus intensity it is already evident that polarization-dependent mask effects are influencing imaging at the wafer. The differences in intensity for the two polarizations are shown to originate at the mask and then amplified by the imaging system, more so for Y polarization than for X polarized illumination.
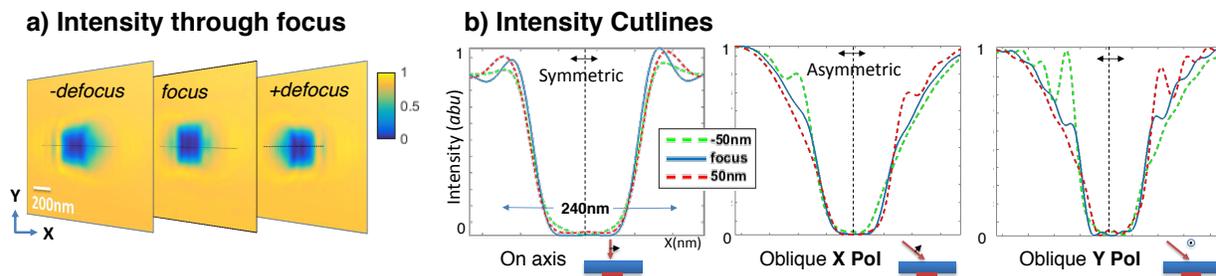
Figure 4.5: Through-focus intensity measurements of OMOG contact under oblique illumination. a) Intensity images show lateral asymmetry which flips on either side of focus (example with Y polarization) b) Overlapping through-focus cutlines are roughly symmetric for on-axis illumination, but more asymmetric for Y polarization than X polarization under oblique incidence. Since the imaging system is identical for both, this is indicative of polarization dependent effects at the mask.

4.3   Phase and pupil recovery

Phase retrieval algorithm

To investigate the electric fields being imaged to the wafer plane and how they differ for the two polarizations, the full field is first recovered with a phase retrieval algorithm. The Transport of Intensity Equation (TIE) [81, 80] relates the intensity stack at wafer to the phase at focus; here we use a modified solver for the TIE developed for use with strongly absorbing photomasks [75]. This gives a first estimate of the wafer plane phase (after removing the fixed phase ramp across the field of view due to the illumination angle).

The electric field at wafer could ideally be Fourier transformed to get the pupil spectrum in one step. To further refine the result, however, and impose the known pupil band-limit
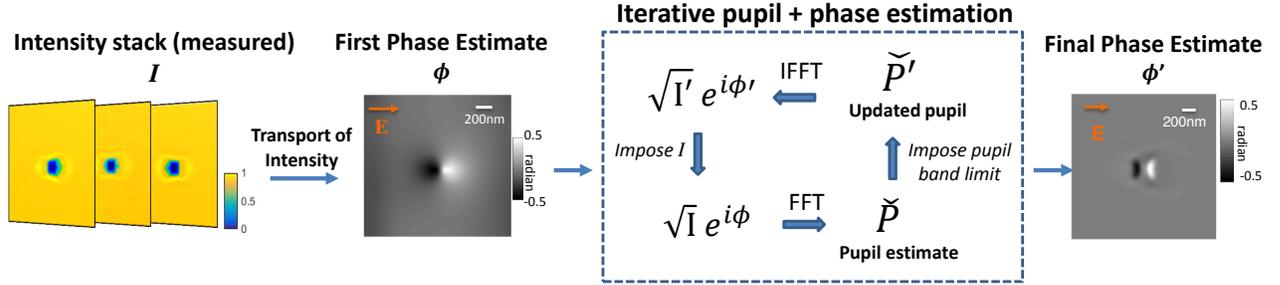
Figure 4.6: Schematic illustrating the phase recovery algorithm that iteratively reconstructs the electric fields at the wafer and the pupil. The defocus intensity stack is solved with TIE for an initial guess of the wafer phase ϕ, which is input into an iterative algorithm that imposes the pupil bandwidth and measured intensity as constraints to improve the estimate.

strictly, we employ an iterative update scheme, alternating between the pupil and wafer (Fig. 4.6). At each projection the known band-limit of the pupil (Fig. 4.4) or the measured intensity at the wafer are imposed as constraints, the process being repeated until convergence. This scheme resembles the method of alternating projections commonly employed in phase retrieval [32, 25], which uses measured intensity and a support constraint to solve for phase [29]. Here we combine it with a TIE based direct solver, which obtains a starting guess from the intensity stack to initialize the iterative method.

Recovered fields at wafer and pupil

The recovered phase at the wafer, and the corresponding pupil spectrum magnitudes are shown in Fig. 4.7. It is seen that the two polarizations have differences in the recovered fields at the wafer and the pupil, as expected from the disparity in the through-focus image behavior previously.

The recovered phase cannot however be directly interpreted as the mask phase due to contribution also from the shifted pupil. The off-center pupil filters the spectrum to create an antisymmetric phase signature, with mask topography adding an additional peak to peak variation. The phase antisymmetry across the feature can in fact be thought of as

being responsible for the lateral asymmetries in defocus intensity in Fig. 4.5, acting as a
linear phase ramp to guide intensity on either side of the optical axis with defocus. The Y
polarization has stronger antisymmetry in the phase at the wafer (Fig. 4.7b), the peak to
peak swing being 20% larger than X polarization, which must originate at the mask since
the imaging system is polarization independent.
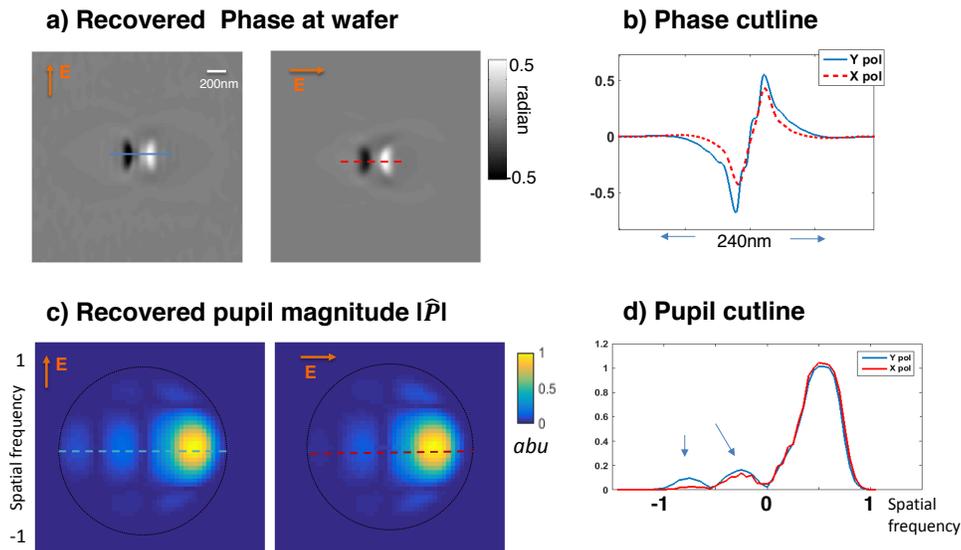
Figure 4.7: Recovered wafer phase and pupil spectrum a) Phase at the wafer has contribu-
tions from pupil filtering as well as mask topography, the strong antisymmetry mainly due
to clipping by the shifted pupil b) Cutlines show a difference in the peak to peak variation
of the two polarizations (about 20%) due to differences in mask near-fields. c) Fourier trans-
forming the wafer field gives the pupil field for the two polarizations d) Cutlines across the
pupil pointing to where Y pol has stronger high frequencies than X pol, adding up to 8%
more energy across the whole pupil.

Differences are also observed in the magnitude of the recovered pupil spectrum (Fig.
4.7c), which shows stronger high-frequency lobes in the diffracted spectrum for Y polarization
(the central frequency is shifted by the illumination to the right edge). This polarization
dependence of the pupil fields must arise at the mask, and is responsible for the degraded
aerial image with Y/TE polarization (Fig. 4.3b). Differences are also expected in the pupil

phase for two polarizations, which has been used in other work to approximate topography effects as pupil aberrations [67, 18, 27].

4.4 Boundary layer modeling of mask near-fields

Having observed considerable differences in the recovered electric fields at the wafer and pupil, the final step in characterizing the dependence of aerial image intensity on mask topography is to estimate the mask near-fields for the two polarizations.

Taking a cue from thin mask modeling using boundary layers for the on-axis case[53] (Fig. 4.2), we utilize a more general boundary layer model that assumes non-identical boundary layers on the two edges (Fig. 4.8). This assumption is necessary for oblique illumination since mask diffraction need not be identical for the left and right sidewalls of the square feature. In total the model has four parameters to be fit - the phase and width of the boundary layers on the each of the two sidewalls.

a) Boundary layer model

$Ae^{i\phi_1}$, width w_1
$Ae^{i\phi_2}$, width w_2

Intensity stack

Imaging system

-defocus   focus   +defocus

Find ϕ_1, w_1, ϕ_2, w_2 **that fit through-focus intensity measurements**

b) Boundary layer values

	ϕ_1	ϕ_2	w_1	w_2
Y Pol	-30^0	30^0	5nm	5nm
X Pol	30^0	-60^0	5nm	10nm
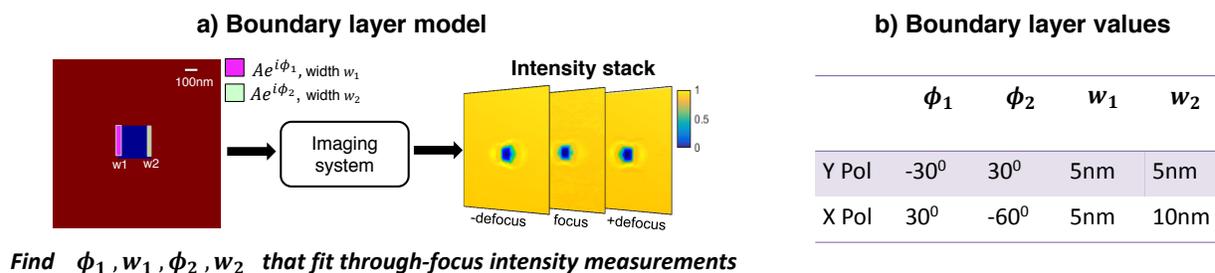
Figure 4.8: a) The mask near field can be modeled as complex valued boundary layers. For off-axis illumination, the left and right sidewalls are modeled as independent boundary layers, making the model asymmetric across the feature. b) The complex values and width of the boundary layers obtained by fitting to measured intensity. The phase is observed to switch sign on switching polarization at a given edge, or between edges for a given polarization.

It is found that a generalization of the simple but computationally fast boundary layer model is sufficiently accurate to predict thick-mask behavior for off-axis illumination in AIMS and hence for projection printers. The imaging model used is a simple low pass filtering of the

complex valued thin mask field (with boundary layers added) to obtain the image intensity at the wafer plane. The boundary layer width and angle are then adjusted manually using a global search over the width-angle $(w - \phi)$ parameter space, the merit function being an RMS difference between the estimated and fit output intensities.The values reported here have accuracy of about $\pm 1nm$ for the widths reported, with given angle.

The table of Fig. 4.8b shows the boundary layer values obtained from fitting to measured intensity (in Figure 4.9). The phase switches polarity left to right for the top row (Y polarization) and right-to-left for the bottom row (X polarization).

a) Boundary layer model **b) Through-focus intensity fit to experiment** (Y pol)
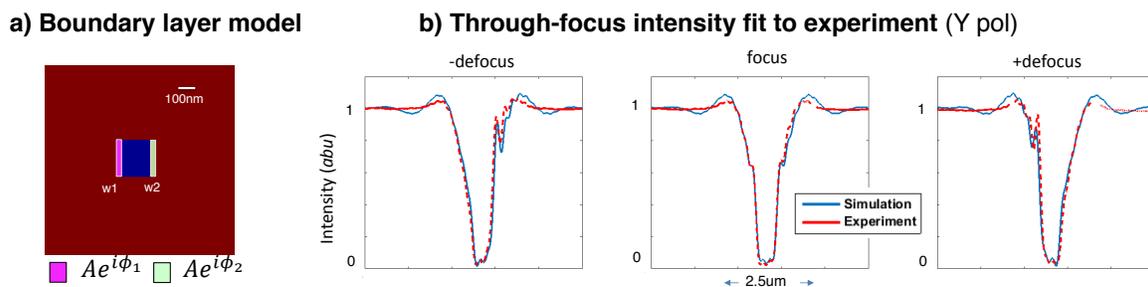
Figure 4.9: Boundary layer values fit to experimental data a) Complex-valued boundary layers used in the thin-mask model. b) Through-focus intensity cutlines at wafer showing the thin mask fit against the measured intensity from AIMS (example with Y polarized light).

For a given polarization, boundary layers at the two edges have opposite signs since the far edge (ϕ_2) picks up more optical phase for oblique illumination propagating through the thick mask, eventually switching sign after crossing 2π. The polarization dependence on the other hand is similar to on-axis boundary layer models [53, 74], where phase changes sign for X and Y polarized illumination (seen in Fig. 4.2). A similar polarization dependent polarity of edge phase is seen here for oblique incidence as well. The simplicity of the model makes it an appealing way to interpret thick mask near-fields, with the possibility of generalizing the framework to a more diverse set of feature sizes, illumination angles, and absorber types.

Finally, the near fields estimated by the thin mask model above can be used to explain

the differences in the pupil and wafer fields from Section 4.3. Figure 4.10a shows the mask phase predicted from the thin-mask boundary layer fitting for the two polarizations. As expected the near-field phase has opposite antisymmetry for Y polarized and X polarized illumination. Antisymmetry in the mask phase translates to asymmetry in the diffracted spectrum, fattening the left half of the spectrum for Y pol, and diminishing it for X pol. Since the shifted pupil filters more of the left spectrum, the combined effect of the mask phase and the shifted pupil is the additional 8% energy seen for Y polarized pupil (Fig. 4.10b). This is the same variation recovered in experiments (Fig. 4.7b), causing stronger image degradation for Y pol (Fig. 4.3b).
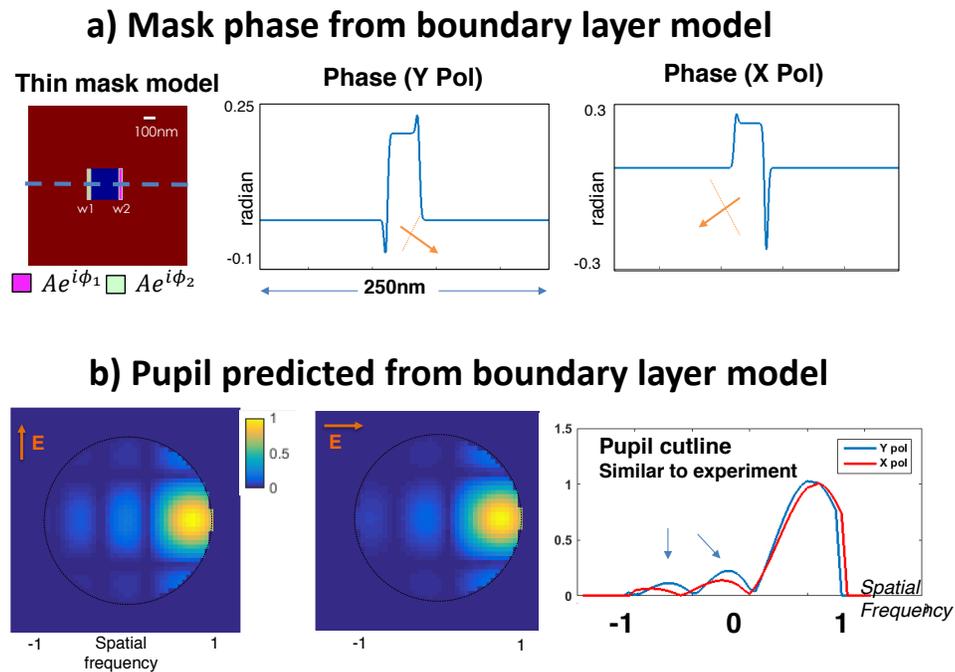
Figure 4.10: The boundary layer model predicts the effective complex valued near-field at the mask for the two polarizations, a) The mask phase is antisymmetric due to off-axis illumination, but has opposite polarity for Y and X polarizations. b) As a result the pupil allows more energy in the high frequencies for Y polarization than for X polarization, causing stronger aerial imaging variability in the former (as also seen in experiments in Fig. 4.7)

.

4.5 Conclusions

The preceding study has analyzed mask topography and imaging system effects for off-axis imaging of a binary mask using an AIMS tool at Toppan Photomask, Dresden with industrial collaborators Martin Scyzrba and Brid Connolly. The mask topographic phase affects not just through focus printing at the wafer, but also at-focus imaging as shown in Fig. 4.3b, due to the inherent pupil shift with oblique illumination. Additionally, topographic effects have a strong polarization dependence, being worse for the Y/TE case. This means there will be limitations when using pupil filters or mask optimization to simultaneously compensate for both X and Y polarized EM edge effects.

Mask diffraction for the given feature size can be modeled with a simple four parameter boundary layer model. Boundary layers with sub-resolution width (5-10nm) can approximate mask near-fields, showing opposite phase polarity for left and right sidewalls. Moreover, the polarity of mask phase switches with polarization due to electric-field boundary conditions at the sidewall. Hence, differences in the mask phase puts 8% more energy in spectrum for Y polarization, leading to the observed 5nm CD variation at wafer between polarizations.

The worse performing Y-polarized illumination in this setting would correspond to the typical TE illumination for printing with off-axis sources. Image quality will degrade further for higher angles of incidence, although using symmetric sources and printing close to focus can mitigate the impact. Since smaller feature sizes at future nodes will have larger illumination angles to push resolution, mask topographic effects can be expected to increase process sensitivity to errors in focus or asymmetries in illumination, hence requiring careful quantification and compensation.

Chapter 5

Differential intensity imaging of photomask edge diffraction

5.1 Introduction

Optical lithography for etching circuits on silicon, in both the 193nm (deep ultraviolet) and 13nm (extreme ultraviolet) wavelength regimes, uses absorbing materials at the mask with a specific thickness to attenuate light, as shown in Chapters 1 and 2. The mask is imaged from the mask to the wafer with a projection printer, hence modulating optical intensity at the photoresist coated silicon wafer to print the desired pattern. There is additionally a demagnification to scale the feature size from mask to wafer (typically a factor of 4). Ideally the pattern would be binary - etch where the mask is clear (when writing a space in a line-space pattern), and preserve the resist when the mask is absorbing (to write the corresponding line). But in the real situation, light scatters at the edge of the absorber, which being like a step diffracts light along its height. The wavefront at the edge of the absorber is hence distorted - by how much will depend on whether the light is polarized normal or parallel to the absorber edge, the absorption coefficient of the absorber, and the shape of the edge. Additionally, this diffraction effect is amplified by a factor of 4 at the

wafer due to the demagnification factor. Hence the wafer intensity near edges is no longer
binary.

These so called "mask 3D" effects [28, 2, 15], can be ignored in the ideal situation of
imaging at perfect focus. The wavefront distortion they cause directs how the desired inten-
sity pattern changes above and below focus ("through-focus"), since energy moves normal
to wavefronts of constant phase [6]. At focus, however, the intensity is still modulated in
the desired pattern by the absorber/clear parts of the mask. Except, in practice, the focus-
ing stage has some mechanical tolerance, typically of the order of a few nanometers. Also,
the photoresist has a finite thickness. Hence, how the intensity spreads "through-focus" is
suddenly important in determining the final critical dimension (CD) of the printed pattern.
This has to be accounted for in "process window" optimization while choosing the tool op-
erating conditions - the mask thickness will restrict how far the stage is allowed to drift, or
how thick the resist can be. [1] How can this through-focus sensitivity of the desired intensity
pattern at the wafer be measured and quantified? The measurement is performed in a Aerial
Image Measurement System (AIMS), a microscope that is like a projection printer, but a
microscope, i.e. it has a camera instead of photoresist to capture the projected image. It
mimics all properties of the projection printing system - numerical aperture (NA) is 1.4 at
wafer, demagnification by a factor of four, illumination is a monopole or multipole according
to the features on the mask . Hence the optical field at the output of the mask is filtered in
the spatial frequency domain by the NA band-limit of the system the image that would be
in the resist is then captured by the camera. A schematic is shown in Fig. 5.1.

There have been various attempts previously to describe "mask 3D" effects[28, 17], or
equivalently phase edge effects [53, 84] - and their impact on the intensity through-focus at
the wafer. The classic measurement scheme is the Focus-Exposure Matrix (FEM) - recording
the critical dimension that would print on the resist with varying defocus, performed for
various exposure levels. However, the CD vs defocus "Bossung plots" [12] thus obtained are
1D plots extracted from the 3D through-focus image stack by choosing an intensity threshold

---

[1]https://spie.org/samples/FG06.pdf
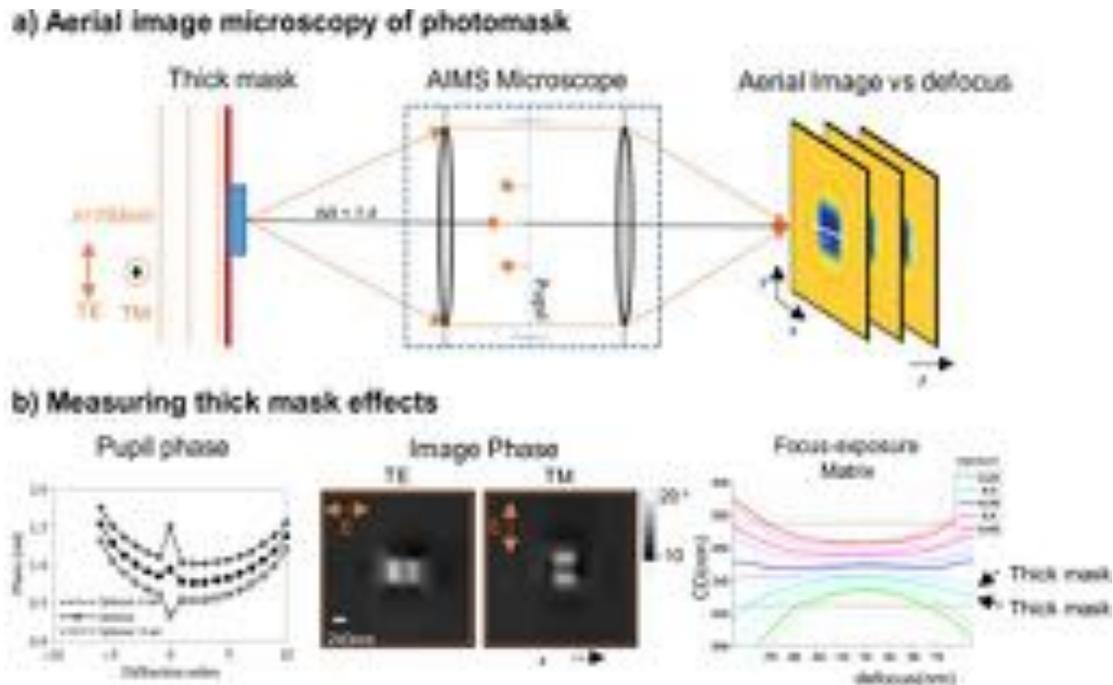
Figure 5.1: The AIMS microcope can measure light propagation properties at the image plane of a lithography stepper tool, which relate to diffraction from the thick mask. a) Schematic of the AIMS tool at DUV, NA =1.4, monopole illumination with source size 0.3x the NA b) Phase retrieval in the pupil or at the wafer can be used to study polarization-dependent thick mask diffraction, since phase captures field modulation due to diffraction from the mask edges.The phase at the wafer subsequently impacts the critical dimension through-focus revealed by the Bossung plot/Focus-Exposure Matrix. Presence of phase at the edges tilts the Bossung curve, shown by the dotted lines.

to represent the photoresist behavior - thus they throw away much of the raw information (Fig 5.1b right). Another method is to use phase retrieval [75, 90] that extracts the wavefront shape at focus using the through-focus image stack. Once the wavefront shape is solved at focus, all through-focus effects can be predicted digitally (similar to digital holography). Phase retrieval has shown promise in some of our previous studies (Chapters 3-4), since the data acquisition is simple, and the phase relates directly to the polarization dependent edge diffraction, seen in Fig. 5.1b center [75]. However this step is computationally intensive and ill-posed, and the algorithms need experience to regularize against numerical noise [82].

Additionally, once the optical fields are known at the wafer, they can be observed in the spatial frequency domain, or in the imaging system "pupil", as described by Finders et. al. in their studies.They have shown that the phase of the scattered orders in the pupil depends on mask thickness (Fig 5.1b left). Although useful in designing masks and visualizing the wavefront shape at focus, these methods are limited in quantitative precision due to the ambiguities in computing the wavefront from intensity propagation.

In this paper we attempt to show that the phase retrieval is not even necessary for extracting useful litho related metrics from the through-focus intensity stack. Instead, a simple algebraic difference of two adjacent defocus images gives the same information as more complicated phase retrieval methods. Similar to video compression algorithms, where only differences in subsequent frames are stored as information, in differential intensity imaging the diffraction effects are deduced from differences in neighboring defocus images. Additionally, this "Differential Intensity Imaging" methodology can be easily converted to defocus sensitivity of the critical dimension, or effectively the Bossung slope. Simple to implement, and working directly with raw images, the method shows surprisingly capability - it reveals polarization dependent symmetries in the underlying wavefront, intensity and CD sensitivity to defocus, and provides a metric to compare the quality of 2D thin mask models and how well they replicate defocus behavior. Section 2 will describe some of the underlying mathematical principles, section 3 shows experimental examples of DII with a square contact feature, and section 4 explores some further applications.

5.2 Theory of Differential Intensity

Differential intensity imaging is a phase contrast mechanism, shown in this section starting from the Transport of Intensity equation for light propagation in the small defocus limit.

Transport of Intensity Equation

Optical flow through-focus depends on the phase of light as it propagates, described by the classic Huygen's principle. Energy flows normal to surfaces of constant phase, or wavefronts, expressed for small distances by the Transport of Intensity Equation [81], essentially an energy conserving differential equation,

$$\frac{\Delta_z I}{\Delta z} = \frac{1}{k_0} \vec{\nabla}_x \cdot I \vec{\nabla}_x \phi \tag{5.1}$$

where the in-plane gradient of the phase $\vec{\nabla}_x \phi$, guides the energy flow at focus, to give the in-plane Poynting vector $I\vec{\nabla}_x \phi$ - its divergence is the TIE, and represents change of energy through-focus, or $\Delta_z I / \Delta z$.[2] This equation can be solved to recover the phase $\phi(x)$ at focus from through-focus measurements $I(x; z)$, as seen in Fig. 5.1b; the phase solved from the TIE is already unwrapped, since the intensity transport equation is a linearization in ϕ (normalized by k_0).

Differential Intensity Imaging

However, for the purposes of photomask metrology in lithography applications, it is sufficient to have phase contrast that is uniquely related to the underlying phase. Even though phase gives an intuitive picture of diffraction from thick edges, all the information is already present in the left hand side of Eqn. 8.3 - i.e. in the energy change through-focus. In fact the phase has to be processed out of this measurement, and is hence a derived quantity, suffering from noise due to inherent regularization ambiguities. Instead look at the differential intensity in z,

$$\Delta_z I(x; z) = I(x : z + \delta z) - I(x; z - \delta z) \tag{5.2}$$

---

[2]Note that ∇ refers to partial derivative while Δ refers to finite difference, which becomes a partial derivative in the infinitesimal limit. The partial derivative in the z direction is approximated in measurement as a finite difference, as shown by the left hand side of the TIE.

is a simple derivative in z of the through-focus intensity stack. It is related to phase - it follows the sign of the phase, and is zero when the phase is zero. It is hence is a simple contrast mechanism for visualizing diffraction effects through-focus. Diffraction is symmetric above and below a pure amplitude object; phase breaks this symmetry. The differential intensity is exactly this change in the intensity, hence quantifying phase or diffraction effects due to mask thickness or tool aberrations. It is rooted in the general idea of interferometry that phase effects in light can only be measured relative to light itself. Here two defocus images are compared to obtain phase contrast and show how energy flows through the stack (Fig. 5.2).

Critical dimension change through focus : CDTIE

The Transport of Intensity (TIE, Eqn. 8.3) inspires another equation for optical flow of particular relevance to lithography . This is the critical dimension (CD) change through-focus. The CD is chosen by defining a certain intensity threshold where the aerial image will etch the photoresist to define the pattern in the wafer. Hence, for a given threshold (say 0.3 of the clear-field value), the critical dimension change through-focus is described by the CD-TIE as [72],

$$\frac{\Delta_z I/\Delta z}{\Delta_x I/\Delta x} = \frac{\Delta CD}{\Delta z} \tag{5.3}$$

i.e. as a ratio of the intensity derivative in the z and the x directions at the position where the CD is being defined, i.e. $I = I(x = x_{CD}; z)$. In the limit, the CDTIE is a differential equation in I,

$$\frac{\partial CD}{\partial z} = \frac{\partial I/\partial z}{\partial I/\partial x} \tag{5.4}$$

This is a particularly useful way of computing CD sensitivity to defocus from a set of measurements in x-z (through-focus intensity measurement) - simply compute the local derivatives in the through-focus (out-of-plane) and in-plane directions - their ratios at the the threshold positions corresponds to CD sensitivity to defocus, or Bossung tilt. This also has a physical interpretation relating it to the phase and intensity of the field - the numerator $\partial I/\partial z$
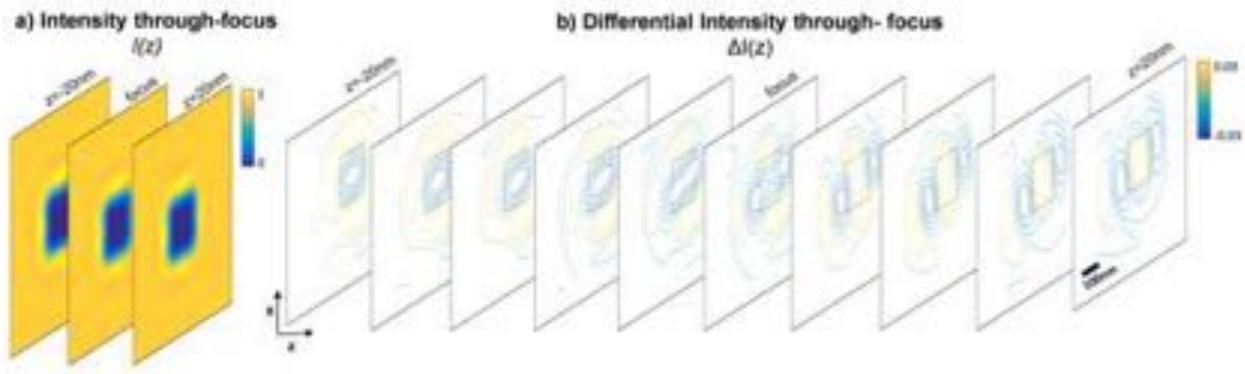
Figure 5.2: Differential intensity imaging corresponds to energy change with each defocus step, hence revealing optical dynamics through-focus. a) In this example with a 240nm feature and 10nm defocus steps, the raw intensity images through-focus look identical to the eye, since phase effects are hidden by the strong absorption of the square feature. b) With the differential intensity images, on the other hand, polarization dependent astigmatism is clearly observed - energy moves in the top and bottom on one side of focus, and moves out the side edges on the other side of focus.

represents phase modulations and the denominator $\partial I/\partial x$ captures intensity modulation by the absorber.

5.3   Differential intensity measurements of OMOG masks

The optical dynamics arising due to diffraction by the thick absorber on the mask is revealed by the differential intensity images through-focus, since the contrast extracted from the 3D image stack is directly related to the optical phase and hence to mask 3D effects.

Experiments are performed with a Opaque MoSi on glass (OMOG) absorber on a a glass substrate. with a wavelength of $\lambda = 193nm$, Three feature sizes are studied- 240nm (super λ), 120nm, and 50nm(sub λ). The images reveal astigmatism like behavior of the thick mask due to polarization dependent edge diffraction, corner effects, curl and divergence of the optical fields etc.

Astigmatism due to edge effects

Diffraction from mask edges depends on the orientation of the incident electric field relative to the mask sidewalls - hence horizontal and vertical edges have phase with opposite polarity [54]. This is due to light being pushed away from the boundary for electric field normal to the sidewall (TE or transverse electric), and being sucked into the absorber for electric field parallel to the sidewall (TM or transverse magnetic). Differential intensity images shown in contour form in Fig. 5.2 reveal how energy flows through-focus. The astigmatic nature is clearly observed in the at-focus differential image which switches sign with edge orientation. This also reverses the direction of energy flow in the two directions, seen in the shift behavior of the DII contours across the vertical and horizontal edge with defocus.

Feature size dependence of edge diffraction

The astigmatism due to the mask edges is seen for larger feature sizes, but is lost when the feature shrinks to sub resolution. In this case the edges merge, and hence polarization dependence of edge diffraction is lost. This is shown for three feature sizes in Fig. 5.3 - TE and TM polarizations for super resolution (240nm), at-resolution (120nm), and sub-resolution images (50nm) are compared, resolution of the AIMS tool being $\frac{\lambda}{NA} = 193/1.4 = 137nm$. The sign of the image difference (and hence mask phase) at the two edge orientations switches for the larger features, but is invariant for the sub-resolution feature - implying that for features smaller than the point spread function, the image is more like a circle than a square, and hence edge diffraction is insensitive to the relative orientation of the edge to light polarization.
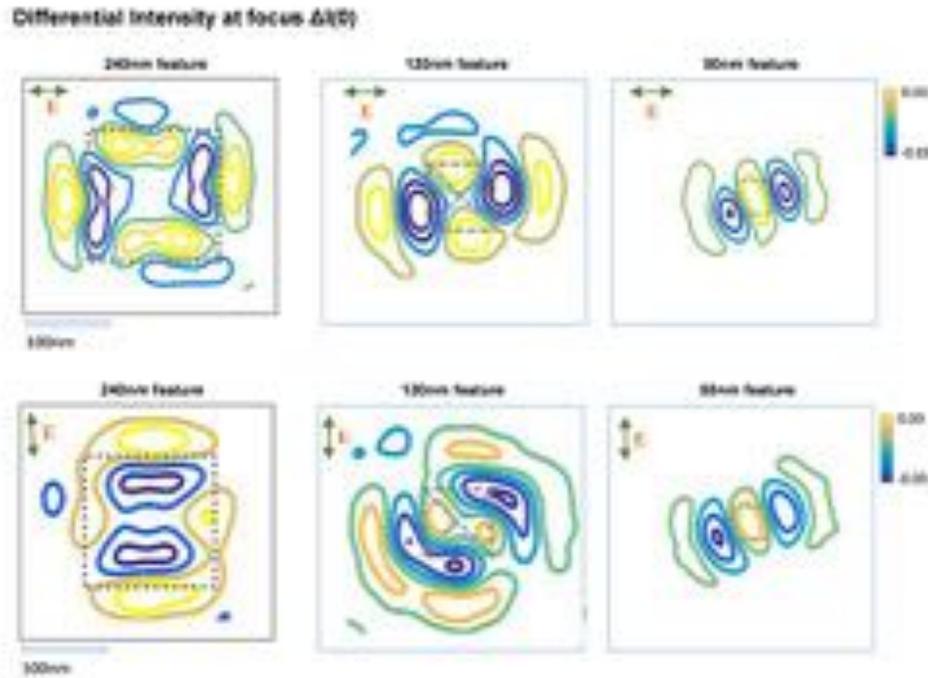
Figure 5.3: Differential Intensity Images about focus clearly reveal interaction of the mask sidewalls with the incident polarization. For the large feature (1st column), the horizontal polarization causes energy to flow out of the vertical edges (blue contours) but into the vertical edges (yellow contours) - the behavior is reversed on switching the polarization direction (2nd row). As the feature size reduces close to the resolution limit of the imaging system, the polarization discrimination of the edges diminishes - for sub resolution features (column 3), the differential image for both images in identical.

Curl from the mask corners

For smaller features, the edges shrink, and the contribution due to the corner starts becoming more significant, varying as the $\frac{area}{perimeter}$ ratio. The corners introduce orbital angular momentum into the beam, since the phase has opposite signs at the vertical and horizontal edges intersecting at the corner[58, 61, 77]. Hence there is a phase gradient along the polar angle around the corner, corresponding to angular momentum in the light propagation $(\vec{r} \times \vec{E} \times \vec{B})$[58], or equivalently curl in the Poynting vector $(\vec{\nabla} \times I\nabla\vec{\phi})$ [75]. The
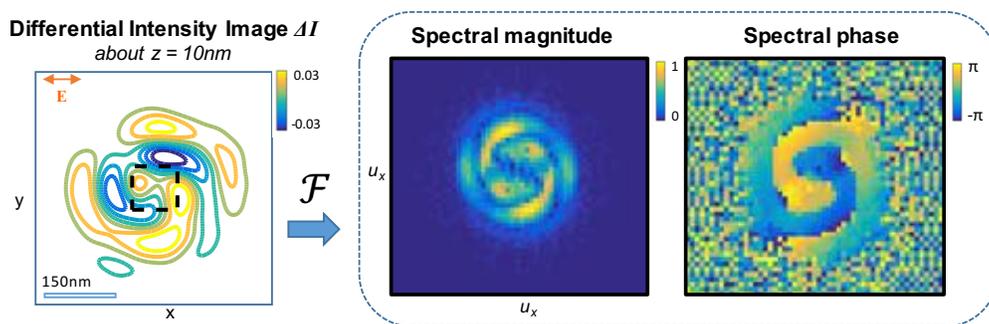
Figure 5.4: Novel physics in strongly absorbing photomasks is revealed in the differential image - light turns the corner downstream from the plane of focus, exhibiting a "curl" in the Poynting vector equivalently described as angular momentum in the energy propagation. Spiral structures in the corresponding spectral signature are typical of orbital angular momentum in light flow.

curl is invisible to differential images about the plane of the mask. However, it manifests in

differential images downstream from where the curl is introduced in the focal stack. This is

seen in Fig. 5.4, differential images about 10nm defocus revealing the curl present at focus.

A spiral pattern is seen in the corresponding Fourier Transform phase - characteristic of

orbital angular momentum in the diffracted light.

5.4   Applications in lithography

Differential intensity images reveal how light intensity moves through-focus, its interaction with the the absorber, and the optical dynamics arising thereof. This contrast mechanism has all the information present in the raw through-focus stack - it can hence be utilized during lithography to design the mask, resist or tool to meet the desired through-focus optical behavior. Three examples are demonstrated in the following.
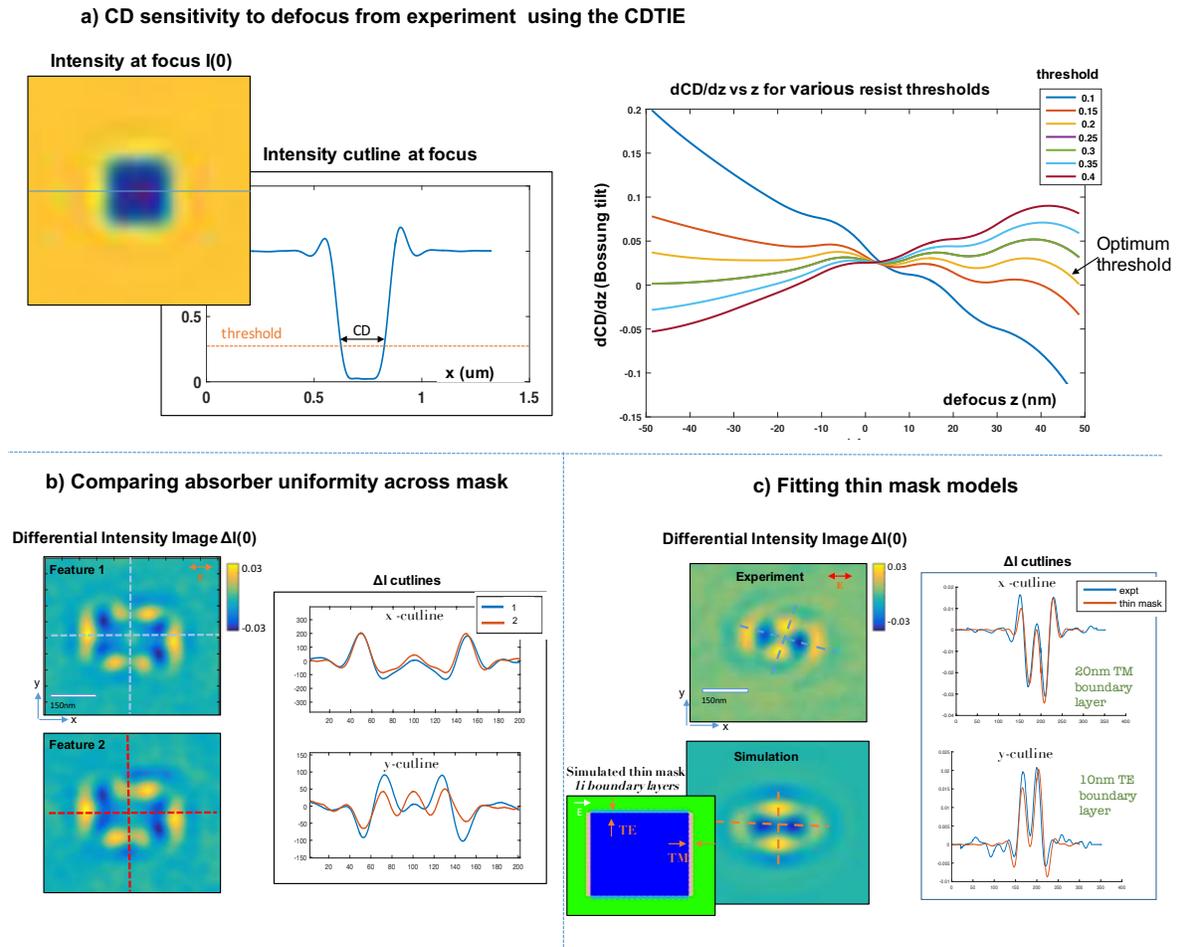
Figure 5.5: Since differential intensity images produce phase contrast through the aerial image stack, they can be used in lithography applications to a) compute CD sensitivity to defocus using the CDTIE of Eqn. 5.4, useful for choosing resist threshold that is least sensitive to defocus b) measure mask uniformity across the die, and c) fit 2D models to replicate the through-focus behavior of the 3D mask.

CDTIE : CD change through focus

The CD-TIE of Eqn. 5.4 can compute from the aerial image stack how sensitive the critical dimension is to defocus for a given photoresist sensitivity (with fixed exposure). The threshold is defined as the position in the intensity image where the feature will print in the resist to define the "critical dimension" on the wafer. It is typically assumed to be some fraction of the clear-field intensity.

Hence Bossung tilt is are obtained at each defocus z from the CDTIE,

$$\frac{\partial CD}{\partial z} = \frac{\partial I/\partial z|_{x_{CD},z}}{\partial I/\partial x|_{x_{CD},z}} \tag{5.5}$$

where all the derivatives are computed at $x = x_{CD}$, the position where the aerial image crosses the chosen resist threshold. For instance, with a resist threshold of 0.3, compute the CDTIE at the position where the at-focus aerial image crosses 0.3 to get the local Bossung tilt . Repeat for all z positions to directly get the differential CD change with defocus for all defocus positions (Fig. 5.5a). To compare with the classic Bossung plots, the CDTIE is the derivative of the Bossung plots with respect to defocus.

Mask uniformity

The differential intensity images are sensitive to anything that changes the phase of the image at-focus - mask thickness uniformity can hence be compared across different features across the die by comparing cutlines in the differential images, as shown in Fig 5.5b.

Thin mask model validation

Edge diffraction can be approximated by having thin mask models that represent 3D effects with complex value 2D features added to the ideal mask. This is similar to boundary layer models[53, 84] or domain decomposition method[1]. These models can be fit to real mask using the differential intensity images through-focus, which are sensitive to even slight

changes in the mask topography. Fig 5.5c shows a boundary layer model with edge diffraction modeled as imaginary valued (1i) additive boundary layers at the edge of the square feature. The thickness of the boundary layers represents strength of edge diffraction - the exact values are recovered by matching to differential intensity images from experiment . Clear polarization dependence is observed - TM boundary layers are twice as wide as TE, confirming values suggested in literature. [53, 75]

5.5 Conclusion

A method has been presented to analyze a measured defocus stack from an photomask microscope (AIMS) by simply looking at differences of adjacent images in the stack. This "Differential Intensity Imaging" modality reveals contrast in the strongly absorbing features that correspond to phase effects, including that induced by the thick mask due to diffraction by its sidewalls. The computed images present a intuitive picture of how light energy distributes with defocus, showing optical dynamics arising out of interaction of incident polarization and the photomask absorber. Furthermore, various lithography applications have been suggested, especially in cases where the defocus behavior is impacted, such as in designing the process window, detecting fluctuations in the mask thickness uniformity, fitting thin mask models to replicate the mask 3D effects through-focus, and for choosing exposure or resist sensitivity that is least sensitive through-focus. The method is much simpler than existing wavefront reconstruction techniques, require no hardware modification or sophisticated algorithms for phase retrieval, and shows promise as a phase metrology technique due to its high sensitivity to the phase and diffraction effects at the mask.

5.6 Outlook: Absorber characterization with through-focus images

Chapters 2-5 have described the impact of absorber topography, most significantly the height of the OMOG absorber, on the through-focus intensity, and the associated at-focus phase of the optical field at the wafer plane of the aerial imaging system. A generalized metrology method is now described for using through-focus aerial imaging data for absorber height measurement.

The process engineer can first apply the simple differential intensity imaging technique to compare the uniformity of the feature across the wafer, or between wafers, as shown in Fig. 5.6, from earlier in this chapter. The DII method is simple to implement, since it involves only subtracting intensity images on either side of focus, hence serving as a fast screening step for comparing uniformity across die.

Next, if the uniformity is out of tolerable specification, the transport of intensity based phase imaging techniques of Chapter 3 and the iterative phase retrieval scheme of Chapter 4 can then be applied for quantitative estimation of the mask uniformity variation. To estimate absorber height from the phase images, first the phase recovery has to be calibrated against known absorber heights (measured with AFM in Chapter 2, for instance), to get the phase variation per nm of absorber height variation. In Fig. 5.7, the phase is seen to vary as about $1°/nm$ for h = 40nm$\pm$ 5%. Once the phase variation with absorber height is calibrated for given mask, absorber type, and imaging conditions, the non-uniformity between features can be quantitatively recovered from the wafer plane phase images.
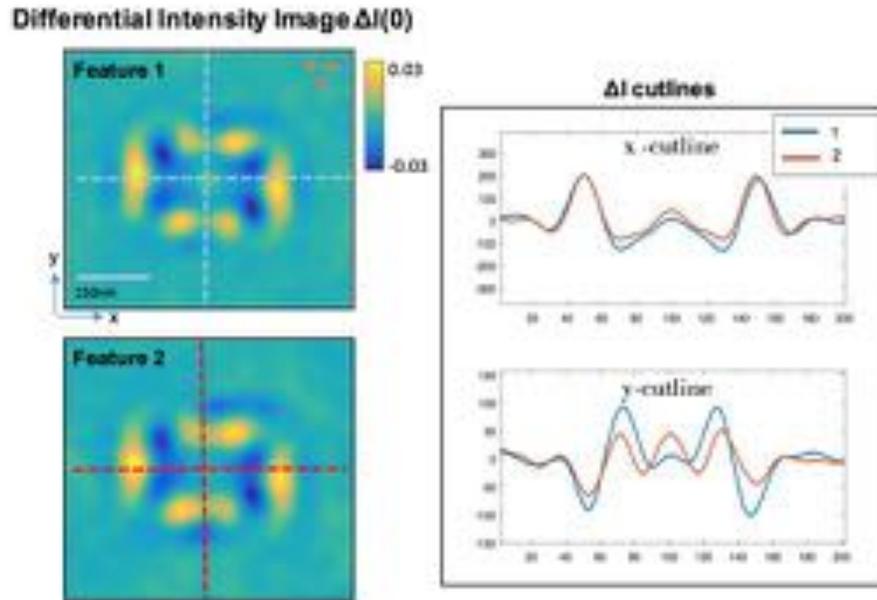
Figure 5.6: Differential intensity images of 40nm OMOG absorber at two different positions on a mask reveal polarization and edge dependent non-uniformity in the optical thickness of the absorber (from Chapter 5).
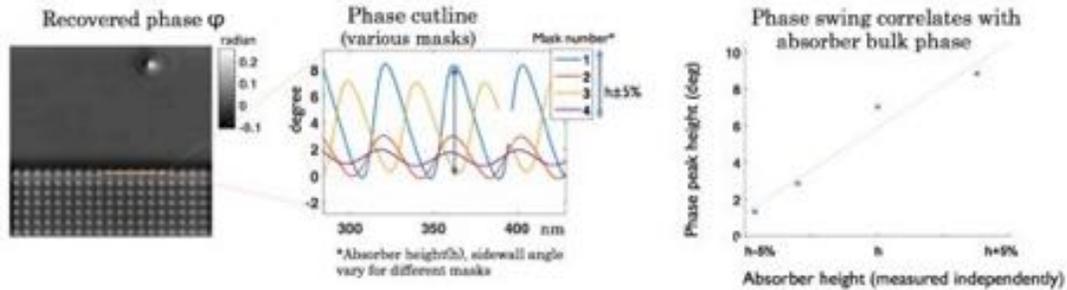
Figure 5.7: For a known set of absorbers with height varying from 38nm to 42nm, the associate phase variation at the wafer plane is about $1°/nm$. Once calibrated, the recovered phase in degrees of unknown absorbers can be used to exactly quantify absorber height fluctuation in nm (from Chapter 2)

Chapter 6

Differential speckle imaging of EUV aberrations

The preceding chapters have discussed methods of measuring the phase of the photomask, the object of interest, assuming an ideal imaging system. However aerial imaging systems used have aberrations in them. Since the final image at the camera is a coupling of the modes of the mask to the modes of the imaging system, to get a close-to-true representation of the photomask object one must disentangle the imaging system 'aberrations' from the structure of the mask as imaged at the camera plane. [1]

To recover the aberrations of the imaging system, the object needs to be generic, in the sense of filling all possible imaging modes, or spatial frequencies, such that the only contrast in the final image is due to the pupil aberration. Fortunately at Extreme Ultra Violet (13.5nm) wavelengths, most masks act as rough scattering surfaces, allowing for probing the aberrated pupil with the broadband scattering from any mask blank.

Coherent speckle produced by a rough surface fills the aberrated pupil of an optical system. Recombination with the on-axis background field interferometrically encodes the even part of the pupil aberration function in the speckle's power spectrum, under a weak object

---

[1]the terms pupil function, imaging system aberration, and pupil wave front are used interchangeably.

approximation. Additionally, the differential change in the speckle intensity with illumination angle is shown to encode the odd part derivative of the pupil aberration. The even and odd aberrations, multiplied by the mask roughness in the power spectrum, are extracted by projection onto a sparse polynomial basis. The method is demonstrated on the SHARP EUV microscope using simply the surface roughness of a blank mask and one illumination tilt along each lateral dimension for real-time computation of the system aberrations.

6.1 Introduction : speckle as a pupil probe

A coherent optical imaging system has a "point spread function" that maps a point in the object to a complex value at the image. For a linear, shift-invariant system, this "point spread function" convolves with the object field to give the image field

$$y(x) = o(x) * p(x) \tag{6.1}$$

where y is the image field o is the object field, and p is the point spread function ; in the frequency domain, the PSF becomes a multiplicative filter, called the "transfer function" of the imaging system.

$$\hat{Y}(u) = \hat{O}(u).\hat{P}(u) \tag{6.2}$$

Where $P(u)$ equivalently is the "pupil" of the imaging system, and u is the frequency domain variable. The pupil is typically related to the shape of the optical elements in the imaging system, that distort the incoming wavefront. Hence, for an aberrated system, this pupil is a pure phase map which distorts the point spread function of the system. Examples of pupil functions are defocus at distance z and wavelength λ, $\hat{P}(u) = exp(1i\frac{u^2}{\lambda z})$ which is an even function along any radius. The first odd aberration in the Zernike basis is coma, $\hat{P}(u) = exp(1i\frac{\alpha u^3}{\lambda^3})sin(\theta)$ which is odd function along any radius u. In general, the pupil phase W(u), $\hat{P}(u) = exp(1i.P(u))$, or the aberration function, can be separated into an even and odd part (Fig. 1b).

$$P(u) = P_{even}(u) + P_{odd}(u) \tag{6.3}$$

with the following decomposition,

$$P_{even}(u) = (P(u) + P(-u))/2$$
$$P_{odd}(u) = (P(u) - P(-u))/2 \tag{6.4}$$
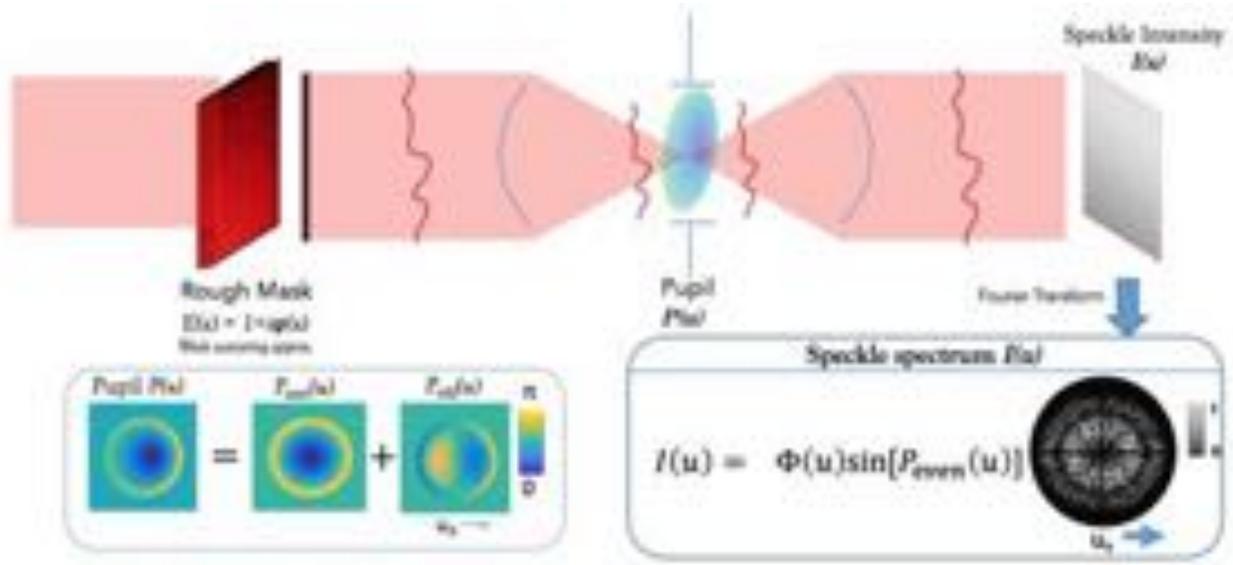
Figure 6.1: A rough mask acts as a weak phase object in the EUV regime, (top) An aberrated imaging system produces speckle modulated by the pupil function, (bottom left) any aberration can be decomposed as an odd and even function along each radial direction, (bottom right) the spectrum of the speckle encodes only the even part of the pupil aberration, which corresponds to the phase contrast transfer function.

Typical approaches to recovering the aberrations of an imaging system are of two kinds - using point objects such as a pinholes in real space, to obtain the point spread function (Green's function $\hat{P}(u)$ with $\hat{O}(u) = 1$); or alternatively, by using a series of periodic gratings $sin(u_{grating}x)$ in real space, to recover the transfer function one point at a time. $\hat{P}(u).\delta(u - u^{grating})$. The point object method suffers from low throughput, since all the
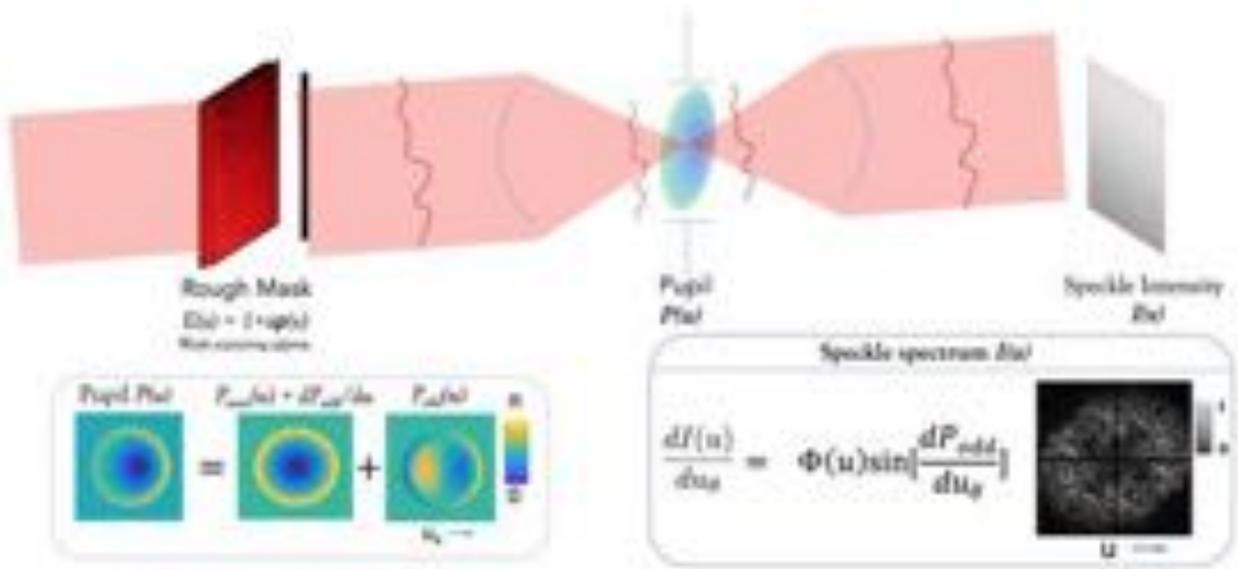
Figure 6.2: Tilting the illumination reveals the derivative of the odd aberration $\frac{dP_{odd}}{d\theta}$, encoded in the change in the speckle spectrum with the illumination tilt $\Delta_\theta I$.

light except from one point is rejected, while the swept frequency grating method needs multiple measurements over grating pitches $u^{grating}$ to reconstruct the entire pupil function $\hat{P}(u)$.

We propose combining the two methods for recovering the aberration, in such a way that the number of measurements is minimized, and the signal to noise ratio is maximized. This would amount to using a "phase-bump" as the object to probe the whole pupil simultaneously, but have a series of these phase bumps distributed randomly . Using a phase object ensures maximum light throughput, and the random distribution ensures that the entire pupil is sampled at the same time. This object amounts simply to a optical phase diffuser - in the ideal case, this diffuser has a white spectrum, hence filling up the pupil in a single shot. For EUV mask microscopy, this diffuser is the natural surface roughness of a blank mask.

The spectrum of the speckle intensity at the imaging plane encodes the "phase contrast transfer function" (PCTF or simply CTF), described in section 2. The CTF is the linear
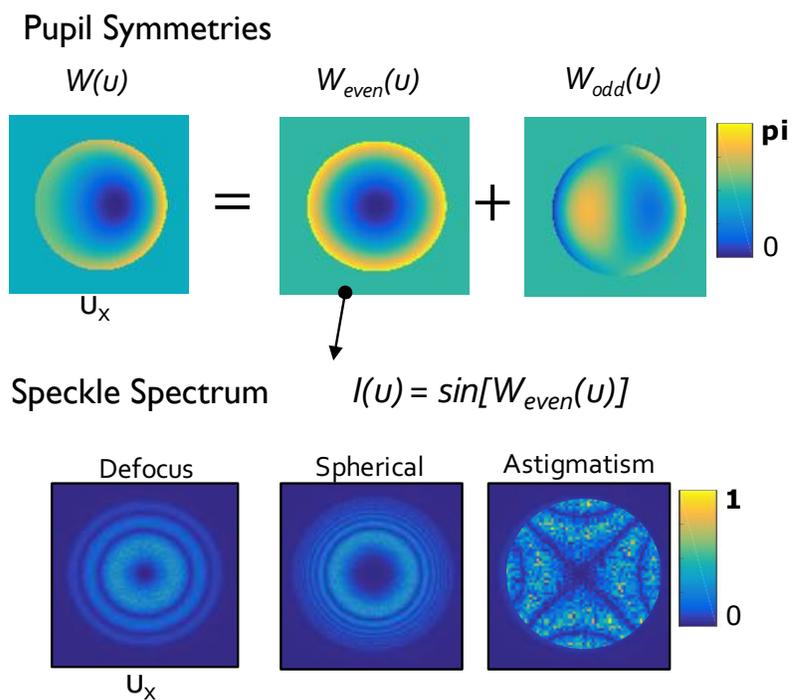
Figure 6.3: Speckle passing on-axis through an imaging system maps even aberrations of the imaging system in the intensity spectrum at the camera, albeit multiplied by the granular spectrum of the rough mask phase. Simulations here show the intensity spectrum for the first three even aberrations in the Zernike aberration basis - defocus (u^2), spherical u^4, and astigmatism $u_x^2 + u_y^2$.

transfer function between object phase and image intensity, valid when the object is a "weak" phase object. It is directly related to the coherent transfer function or point spread function of the imaging system. While in the visual wavelength regime the speckle created by coherent light is usually "strong", the weak object approximation naturally holds in EUV, X-Ray and electron microscopy, where the incident beam is only slightly perturbed by the scattering object. This preserves a strong "background" beam that creates the requisite phase contrast by interfering with the speckle phase at the image plane. The intensity pattern hence created has a spectrum that encodes the contrast transfer function of the system [92] [2]. In the visible regime where scattering is stronger, the "weakness" of the object has to be engineered - this is done using a phase diffuser that is "index matched" to make it "weak". [36]

A linear theory is presented that describes the symmetries of the coherent transfer function (or equivalently the point spread function) that interact with the phase diffuser's spectrum to create speckle at the image. This phase transfer function is contained in the intensity spectrum of the imaged speckle. It is hence centro-symmetric in the spectral domain, since the spectrum of the real valued intensity is even, imaging only even valued aberrations (Figs. 6.1,6.3). Further, changing the illumination direction on the object shifts the pupil, breaking the symmetries in the spectrum, allowing for recovering the odd aberration as well (Fig. 6.2). Subsequently a full reconstruction of the phase at the pupil of the imaging system is possible from purely looking at differences between speckle images for differential changes in the illumination angle, described in section 6.3 as "Differential speckle imaging".

The added advantage of this method is that the object and imaging system stay in place during the measurements, only the illumination direction needs to be shifted. Hence the mechanical complexity of the system is reduced. Also, since speckle occurs naturally at the blank parts of the mask, in-situ measurements of the aberrations are possible across the field of view, or even when measuring patterned mask, as long as there are blank areas.

The SHARP beamline at the Lawrence Berkeley National Lab [33], operating at the

---

[2]Contrast transfer function, also used interchangeably with the Phase transfer function or Coherent transfer function

extreme ultraviolet, is naturally suited for this application. Most speckle created by blank masks at extreme ultra violet (13.5nm) is within the "weak" scattering regime, since light at this frequency naturally has low index of refraction for most materials. This allows a linear phase contrast transfer function to exist, and its subsequent estimation from the speckle intensity created by a blank mask. Further, the SHARP uses a zone plate imaging lens designed to image perfectly only at the center of the field of view. Aberrations away from the center can be simply and easily characterized using the local speckle intensity in each region, shown in section 4.

6.2 Optical transfer function for weak phase objects

The phase contrast transfer function is a linear relationship between the object phase at the input $\phi(x)$, and the intensity at the output, $I(x)$, expressed as a convolution

$$I(x) = \phi(x) * k(x) \tag{6.5}$$

or equivalently as a point wise multiplication in the frequency domain,

$$I(u) = \phi(u).K(u) \tag{6.6}$$

Where K(u) is the "phase contrast transfer function", and u is the spatial frequency coordinate. $K(u)$ is related to the even part of the coherent transfer function of the pupil, as described shortly. The idea behind using a phase diffuser, is to have $\phi(u) \approx 1$ for $\forall u$ in Eqn. 6.6, so that $K(u)$ can be recovered directly from the spectrum of the image intensity, $I(u)$. Hence if the speckle from a rough object after passing through an imaging system is Fourier transformed, it instantly reveals patterns in the spectrum corresponding to $K(u)$. An example is shown in Fig 1, where the spectrum reveals defocus, spherical and astigmatic aberrations in the pupil.

Even pupil measurement with weak diffuser
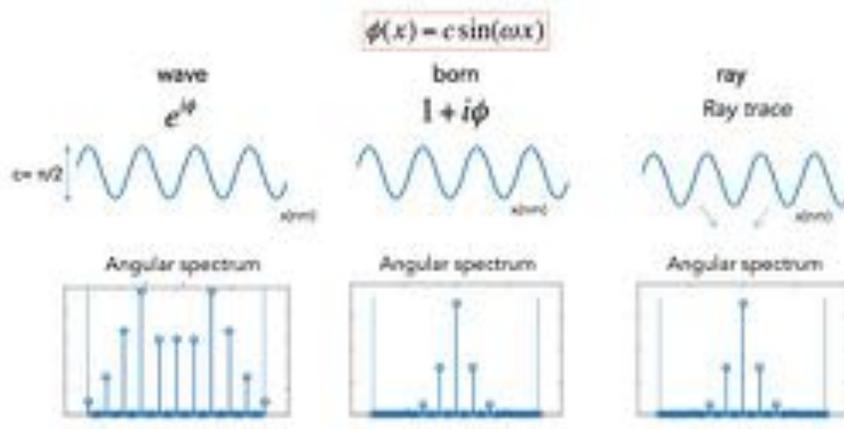
Figure 6.4: Regimes of strong and weak scattering of a sinusoid phase grating, $e^{i\gamma sin(kx)}$. In the refractive regime (Born approximation), the scattered orders are those of the sinusoidal grating ($\pm k$); as γ increases, higher harmonics appear in the Fourier domain, indicating scatter-scatter interactions and higher order correlations typical of wave interactions.

So how does this phase contrast transfer function, $K(u)$ depend on the coherent pupil phase $P(u)$ (where $\hat{P}(u) = \exp(iP(u))$? A linearization between the phase and intensity is needed, which depends upon the object being *weakly scattering*. The intensity is bilinear in the electric field to begin with,

$$I(x) = |y(x)|^2 = |o(x) * p(x)|^2 \tag{6.7}$$

where $o(x) = \exp(i\phi(x))$ for a pure phase object. Using the linearization $p(x) = 1 + i\phi(x)$ for a weak phase object effectively implies refractive scattering from the phase object, where the local change in propagation direction is determined by $\vec{\nabla}\phi(x)$. This holds if the phase surface is under a quarter wavelength in optical path length. Beyond this higher order derivatives of the phase also contribute, as the scattering becomes more diffractive (Fig.6.4). The corresponding experimental implementation is shown in Fig. 6.5, where an optical diffuser has been index matched by a film of oil to satisfy the weak object approximation. Under these conditions Eqn. 6.7 becomes
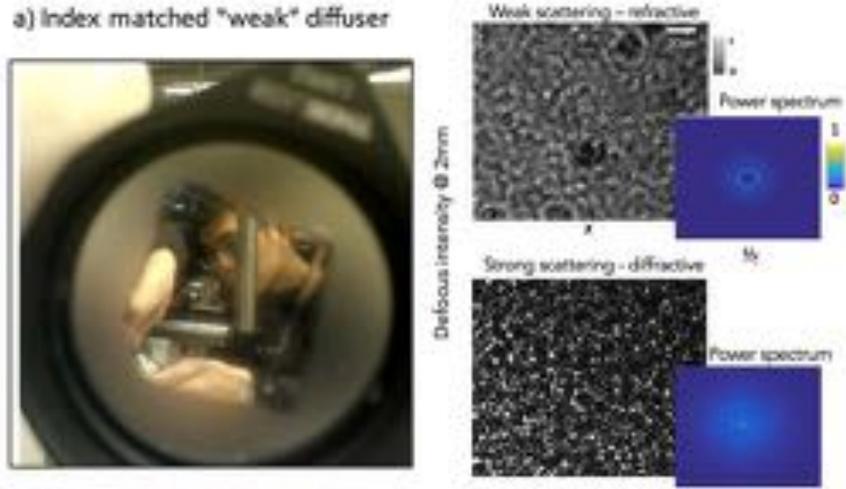
Figure 6.5: Phase diffuser (Newport 10 degree) index matched by applying a drop of oil on a coverslip to the rough side of the diffuser; a) the index matched area of the diffuser is almost clear, indicating a strong background term and the validity of the weak phase linearization. b) the speckle through the weak diffuser. c) The spectrum of the speckle encodes the contrast transfer function, in this case the sine of the defocus kernel $(\pi\lambda z u^2)$ between the diffuser the camera

$$I(x) = |(1 + i\phi(x)) * p(x)|^2 \tag{6.8}$$

where the 1 is interpreted as a background wave perturbed by the diffuser phase, $p(x)$ is the complex valued point spread function, and $\phi(x)$ is the real valued diffuser phase. This allows for the phase at the image plane, modulated by the pupil, to interfere with the background. Expressed mathematically,

$$I(x) = [(1 + i\phi) * p][(1 - i\phi) * p^*] \tag{6.9}$$

drop the second order scatter-scatter term assuming that all the contrast is due to interference with the background, obtaining the following linearization,

$$I(x) = 1 + \phi(x) * [p(x) - p^*(x)] \tag{6.10}$$

$$I(x) = 1 + 2\phi(x) * Im[p(x)] \tag{6.11}$$

which is the required linearization of the output intensity as a function of the input phase, such that $k(x) = 2Im[p(x)]$ is the phase contrast transfer function within the regime of linearization. Equation 6.11 states that the imaginary part of an object is imaged by the imaginary part of the point spread function. A good example is contrast due to defocus - the defocus kernel has an imaginary part $1i * sin(x^2/\lambda z)$, which produces phase contrast.

Additionally, the real part of the object is imaged by the real part of the point spread function, within the weak object regime. Hence coma, or other aberrations with odd pupil phase, which have a real valued point spread function, gives contrast only with real valued objects. In other words, phase objects are invisible to odd aberrations! This can be seen mathematically by considering the spectral form of Eqn. 6.11

$$I(u) = \delta(u) + 2\Phi(u)(e^{iP(u)} - e^{iP^*(-u)}) \tag{6.12}$$

where $\Phi(u)$ is the spectrum of the diffuser surface, $\hat{P}(u)$ is the aberated pupil phase function. Substitute $\hat{P}(u) = \exp(iP(u))$,

$$I(u) = \delta(u) + 2\Phi(u)sin(P_{even}(u)) \tag{6.13}$$

Hence the contrast transfer function $K(u) = sin(P_{even})$, is the sine of the even part of the pupil phase. In terms of Zernike coefficients, this implies that even aberrations such as defocus (u^2), spherical (u^4) etc are directly measured in the speckle spectrum, as shown in Fig 1c. This is hence a method for measuring the even part of the pupil - only half the information needed for reconstructing the entire pupil.

To intuitively explain the contribution of the even pupil to the speckle intensity, consider the interaction of the diffuser object with the pupil phase. A purely imaginary (weak phase object) has odd symmetry in the phase of its diffracted orders, and is invisible in the intensity
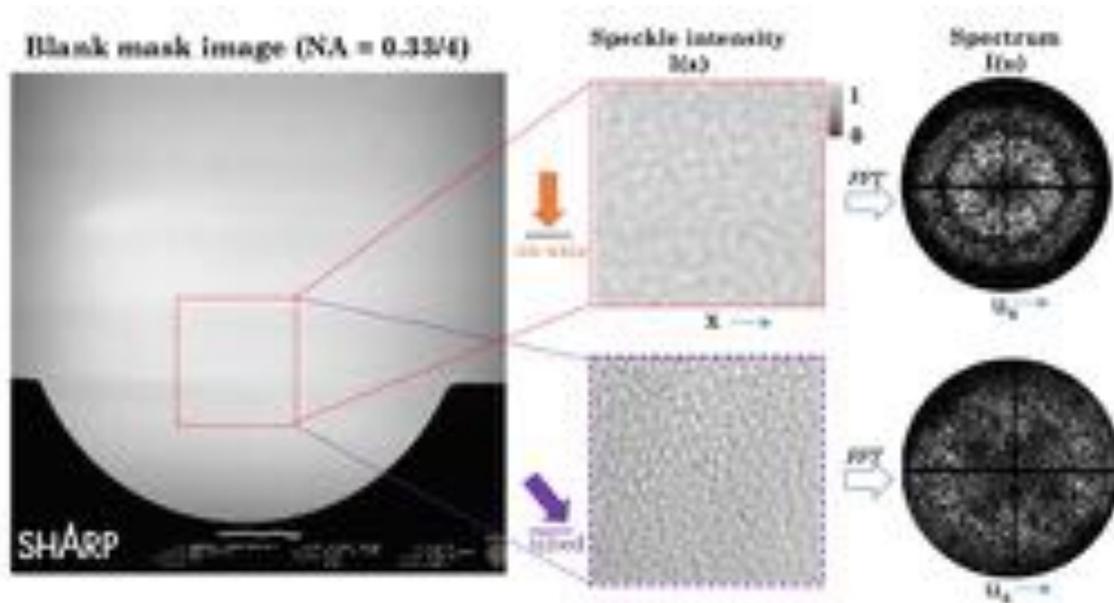
Figure 6.6: Speckle in EUV light can be used for estimating the net wavefront aberration of the imaging mirrors. The speckle in a blank mask imaged to an CCD camera from the SHARP aerial image tool (left); the speckle image spectrum with on-axis illumination gives contrast related the even part of the wavefront aberration pupil (right top); the image with off-axis illumination gives the odd part of the wavefront aberration pupil (right bottom). mapped directly in the spectrum of the speckle at the wafer plane.

at-focus .The even part of the pupil phase adds to the object's pupil phase, breaking this odd symmetry. Hence the complex electric field at the image now has a real part, causing intensity contrast. An odd function in the pupil, on the other hand, retains the odd symmetry, keeping the invisible phase object invisible (Fig 6.7. shows relationships between the symmetries of the pupil phase function and the illumination angle of the incoming plane wave on the speckle generating surface).

Odd pupil measurement using illumination tilt

The phase contrast transfer function, imaged in the spectrum of the speckle intensity, directly shows the even part of the pupil phase. These include aberrations such as defocus and spherical aberrations, where the aberration phase is even along any radial direction. To
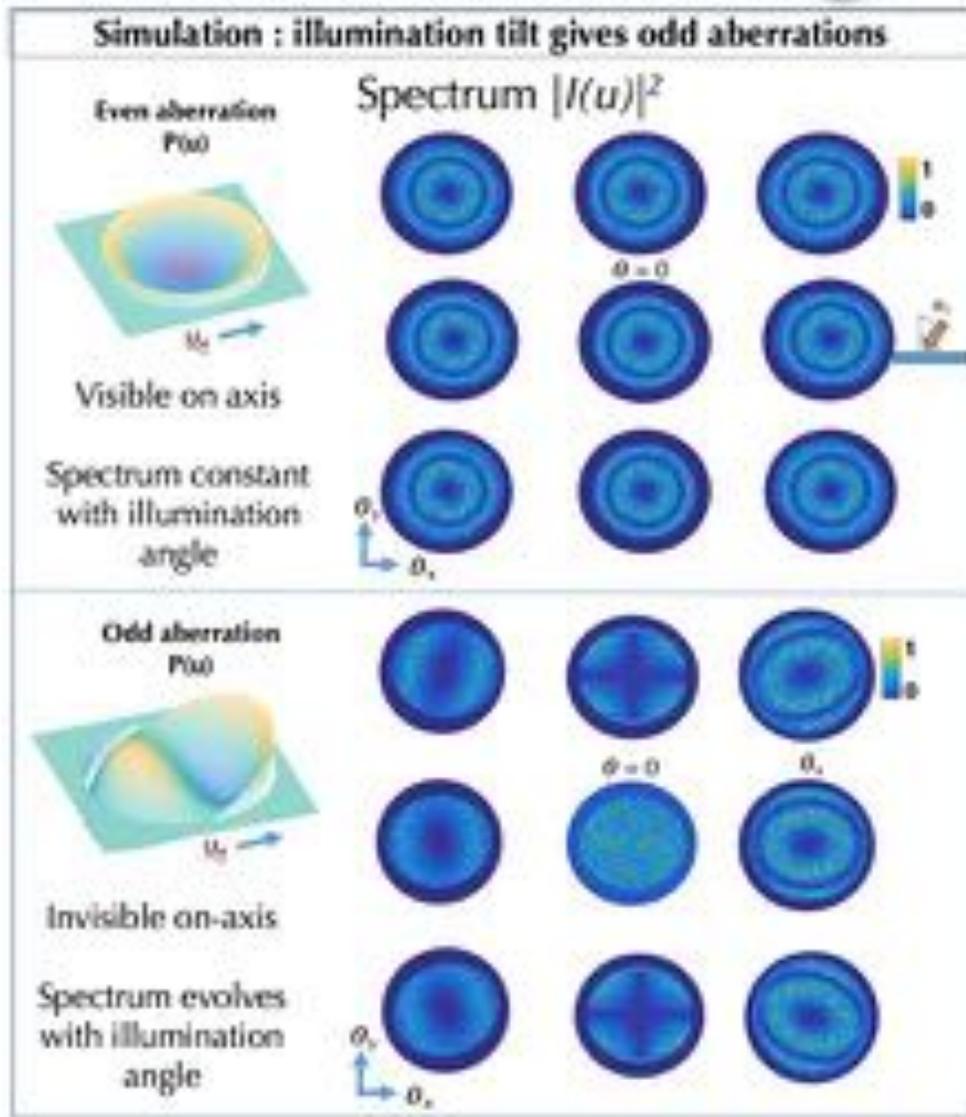
Figure 6.7: Simulations comparing even and odd aberrations with differential illumination tilts. Even aberrations(top) are visible on-axis, and invariant with pupil shift / illumination tilt, while odd aberrations(bottom) are invisible on-axis, showing contrast only on tilting the illumination. This indicates that the illumination dependence of the speckle spectrum relates directly to the imaging system symmetries.

obtain the odd aberrations and hence the whole pupil map, a shift operator is needed; this couples the odd part into the contrast transfer function. This is possible because shifting

an odd function gives it an even part corresponding to its derivative, as shown shortly. In a paraxial imaging system, the pupil can be shifted with respect to the object spectrum simply by changing the illumination angle, since a phase ramp across the object shifts its spectrum relative to the pupil of the imaging system. (refer Fig. 1). As an example of the underlying algebra, first consider the specific example of the simplest odd polynomial - the cubic function,

$$Even[\alpha u^3] = 0 \tag{6.14}$$

now with a shift by amount a, the odd function produces an even component proportional to its derivative multiplied by the shift amount a

$$Even[(u-a)^3] = 3a\alpha u^2 = a\frac{d\alpha u^3}{du} \tag{6.15}$$

In general, for any odd function, shifting it by a differential amount a couples its derivative into the even part. Here's the proof - any odd function $f_{odd}(x)$, by definition, is anti-symmetric about zero,

$$f_{odd}(x) = -f_{odd}(-x) \tag{6.16}$$

and so the even part of the function, symmetric about the origin, is zero,

$$Even[f_{odd}(x)] = \frac{f_{odd}(x) + f_{odd}(-x)}{2} = 0 \tag{6.17}$$

Now shift the function by a to get $f_{odd}(x-a)$ - the even part then becomes

$$Even[f_{odd}(x-a)] = \frac{f_{odd}(x-a) + f_{odd}(-x-a)}{2} \tag{6.18}$$

using the antisymmetry of odd functions (Eqn 6.16) , and the definition of an odd function, $f_{odd}(-x-a) = -f_{odd}(x+a)$, so the new even part of the shifted odd function becomes the following finite difference, which becomes a derivative for small values of shift,

$$Even[f_{odd}(x-a)] = \frac{f_{odd}(x-a) - f_{odd}(x+a)}{2} \approx a\frac{\partial f_{odd}(x)}{\partial x} \tag{6.19}$$

- a simple but remarkable result. Hence to obtain the derivative of an odd function, shift the function by a small amount a - the even part of the shifted function is a derivative of the original odd function, scaled by the shift amount.

The above result can be extended to any general function $f(x) = f_{even}(x) + f_{odd}(x)$, in which case shifting the function couples the odd derivative into the even part, and the even derivative into the odd part. Consider the shifted function,

$$f(x - a) = f_{even}(x; a) + f_{odd}(x; a) \tag{6.20}$$

where the relationship between the odd-even decomposition of the shifted and unshifted functions is,

$$\frac{df_{even}(x; a)}{da} = \frac{df_{odd}(x)}{dx} \tag{6.21}$$

which is the result of interest optically, since a phase object measures the even part of the contrast transfer function. Hence if the contrast transfer function can be shifted with respect to the object's diffraction spectrum (by tilting the illumination), then the odd derivative can be measured as a change in the even part.

For completion, the new odd part of a function f shifted by distance a is likewise the derivative of the even part of the unshifted function.

$$\frac{df_{odd}(x; a)}{da} = \frac{df_{even}(x)}{dx} \tag{6.22}$$

The next section applies these mathematical concepts to an imaging system, describing the aberration recovery process within the framework called differential speckle imaging - where both the even and odd aberrations are obtained from the power spectrum difference between speckle images.

6.3 Differential speckle imaging of the aberrated pupil

The mathematical ideas presented above are all implementable optically, of course. To recover, for instance,the even aberrations in the imaging system from the CTF of Eqn. 6.13, take the difference of the on-axis speckle intensity and the mean background intensity. Then the spectrum becomes

$$I(u) - \delta(u) = \Phi(u)sin(P_{even}(u)) \tag{6.23}$$

or in the image domain

$$I(x) - 1 = \phi(x) * (p(x) - p^*(-x)) \tag{6.24}$$

which sets the tone for the differential imaging method to be proposed, since the even aberrations are measured on-axis as the difference between the perfect mask, 1, and the speckle intensity from a rough mask, $I(x)$ as shown in Fig. 6.6a. Notice that the spectrum thus obtained contains the diffuser spectrum multiplied by the pupil function, signifying that the "random" diffuser phase is the carrier signal of the aberated pupil. The axially symmetric even aberrations can thus be measured on-axis, by comparing the contrast in the speckle with the ideal flat image expected for a perfect imaging system.

To optically perform the shift operation of Eqn. 6.20 and recover the odd pupil derivative, simply tilt the illumination at the object to shift the pupil relative to the object spectrum. The change in the even part of the shifted pupil must be the derivative of the odd-pupil, by Eqn 19. However since the measurement is inside a sine term in the CTF, the difference of the on-axis and off-axis speckle yields for θ angle of illumination,

$$I(u;0) - I(u;\theta) = \Phi(u)[sin\{P_{even}(u;0)\} - sin\{P_{even}(u;\theta)\}] \tag{6.25}$$

where a paraxial model allows the illumination tilt to be modeled with simply a shifted pupil - i.e $P(u;\theta) = P(u - u_\theta)$, where $u_\theta = \frac{sin(\theta)}{\lambda}$ is the pupil coordinate shift corresponding to illumination angle θ,

$$\Delta_\theta I(u) = \Phi(u) sin\{u_\theta \frac{dP_{even}}{du_\theta}\} cos\{\frac{P_{even}(u;\theta) + P_{even}(u;0)}{2}\} \tag{6.26}$$

Notice the assumption that $\Phi(u)$ or the diffuser spectrum, is the same for both on-axis and tilted illumination, true if the same area of the mask is illuminated by the tilted measurements. Now the change in the even pupil aberration with illumination tilt is the gradient of the odd pupil from Eqn. 6.21

$$dP_{even}/du_\theta = dP_{odd}/du \tag{6.27}$$

gives the following expression of the differential power spectrum in terms of the odd pupil derivative, assuming weak aberrations,

$$\Delta_\theta I(u) = \Phi(u) sin\{u_\theta \frac{dP_{odd}}{du}\} cos\{\frac{P_{even}(u;\theta) + P_{even}(u;0)}{2}\} \tag{6.28}$$

which is the rigorous form of the differential power spectrum. To obtain the linear form of the expression, the argument inside the sine and cosine term have to be small ($\rightarrow 0$). This is true for a small tilt angle and hence a small u_θ, for which the term inside the sine, $u_\theta dW_{odd}/du$ goes to zero. Or conversely, for lower order aberrations, especially if they are weak, the derivative dW_{odd}/du itself goes to zero . For imaging with a well calibrated system such as the SHARP EUV microscope, the aberrations are typically a fraction of the wavelength, hence allowing the linearization -

$$\Delta_\theta I(u) = u_\theta \Phi(u) \frac{dP_{odd}(u)}{du} \tag{6.29}$$

where $\Delta_\theta I(u) = \mathscr{F}[\Delta_\theta I(x)]$ is the spectrum of the difference of the speckle intensity. Hence the power spectrum of the speckle difference on shifting the illumination gives the derivative of the odd aberration in the pupil, albeit multiplied by the roughness spectrum $\Phi(u)$ (Fig. 6.6). Next, a sequence of measurements is used to show the extraction process of the aberrations from Eqn 6.29 and Eqn. 6.23, on the SHARP EUV imaging system using only speckle generated from a mask blank.
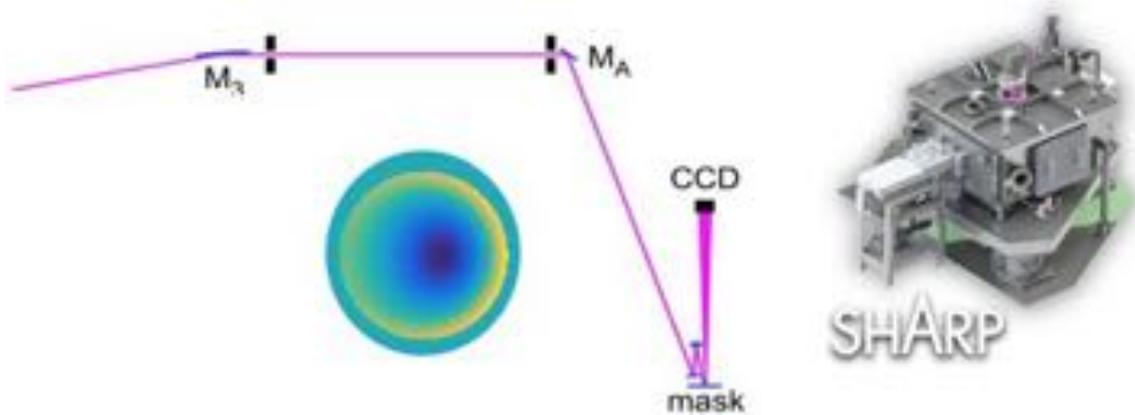
Figure 6.8: SHARP EUV microscope schematic, showing the reflective geometry of the EUV microscope. Since EUV light has vanishing refractive indices $n = \epsilon + i\delta$ in most materials, focusing surfaces have to be MoSi multilayer mirrors that Bragg reflect upto 70% of the light for a sufficient depth of multilayers. However due to the multilayer mirror losses multiplying, the source power is rapidly attenuated by the time it hits the wafer and reaches the CCD sensor, making source power the greatest limiting factor restricting manufacturability at production scale . The source is a coherent scanning pinhole at the output of the beamline, enabling illumination tilt or synthesizing complex source shapes. The focusing lens between the mask and the camera is a custom made diffractive zone-plate element at the Center of X-Ray Optics, Lawrence Berkeley National Laboratory.

6.4 Aberration extraction on the SHARP aerial imaging microscope

The aberrations across the field of view of the SHARP microscope can be measured from the speckle generated from a blank mask at the microscope object plane. As shown in the previous section, the even and odd parts of the pupil are buried in the speckle passing through the imaging system on-axis and off-axis respectively. Once the contrast transfer function $\Phi(u)sin(P_{even}(u))$ is known from the on-axis illumination spectrum, and $\Phi(u)\frac{dP_{odd}}{du}$ is obtained from the illumination tilt difference $\frac{dI(u;\theta)}{d\theta}$, the derivatives need to be combined to get the net pupil aberration $P(u) = \Phi(u)(P_{even}(u) + P_{odd}(u))/2$, where the speckle spectrum $\Phi(u)$ is still part of the solution - to get the pure aberration function back, blind deconvolution
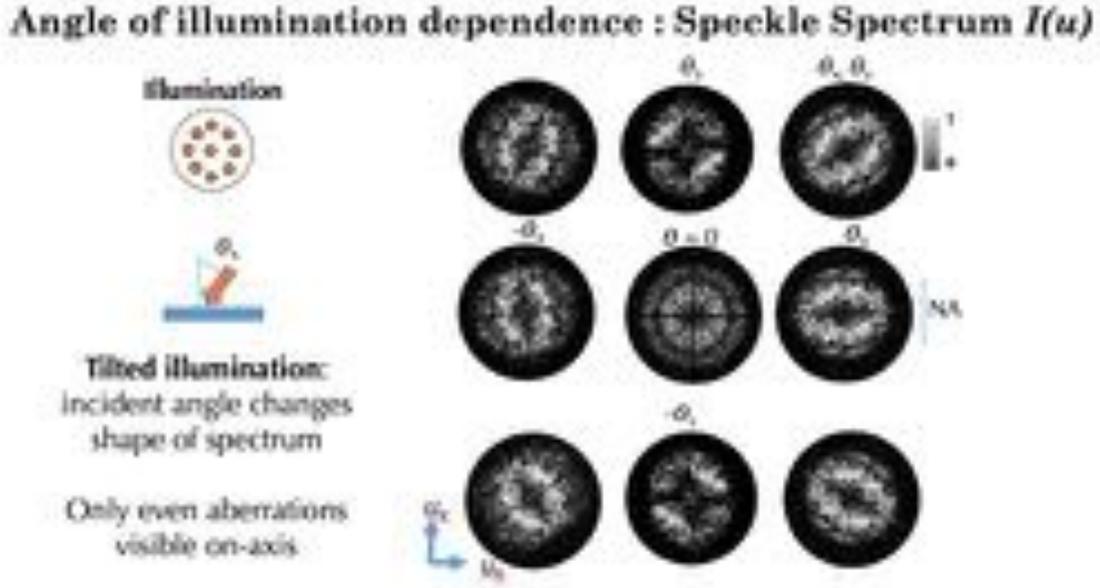
Figure 6.9: Tilting the illumination laterally(rows) or vertically (columns)changes the spectrum of the speckle, breaking the symmetry of the defocus rings at the center ($\theta = 0$). This indicates the presence of odd aberrations in the SHARP imaging system.

over a sparse basis function (like Zernike or Seidel) may be used as shown most recently by Gunjala et al. [30].

Here only $\Phi(u)P(u)$ is solved for, to retain maximum fidelity to the measured data, since more advanced numerical methods for extracting just the aberration pupil $P(u)$ will introduce numerical noise into the solution. The 2D pupil is inverted from its second derivative, since the laplacian has spherical $\nabla^2 P(u) = \nabla^2 P_{even}(u) + \nabla^2 P_{odd}(u)$, also called the Poisson equation (with many standard solvers). Solving the Poisson equation has the advantage of spherical symmetry in spectral space over both dimensions (u_x and u_y), additionally avoiding the phase unwrapping problem for extracting $P_{even}(u)$ from $sin(P_{even}(u))$.

On the SHARP tool since the aberrations are small ($< \lambda/10$), for a small aberration approximation, the following property is utilized to first arrive at the Poisson equation,

$$\frac{d^2 P_{even}}{du^2} \approx \frac{d^2 sin(P_{even})}{du^2} \tag{6.30}$$

or in terms of the measurements

$$\Phi(u)\frac{d^2 P_{even}}{du^2} \approx \frac{d^2 I(u)}{du^2} \tag{6.31}$$

and

$$\Phi(u)\frac{d^2 P_{odd}}{du^2} = \frac{d^2 I(u;\theta)}{dud\theta} \tag{6.32}$$

where x,u are space and reciprocal space (short hand for x-y and u_x, u_y, θ is the illumination tilt and $\Phi(u)$ is the omnipresent multiplicative speckle spectrum throughout these simulations, $I(u) = \mathbb{F}[I(x)]$; hence the final Poisson equation solved to arrive at the pupil function $P(u)$ from adding the above two equations for the laplacian (curvature) of the odd and even pupil

$$\Phi(u)\nabla^2 P(u) \approx \Phi(u)[\nabla^2 P_{even}(u) + \nabla^2 P_{odd}(u)] \tag{6.33}$$

as shown in Fig. 6.10. The Poisson solver used is the same as used for the optical diffusion equations described in Chapter 3, once solved, the aberrations of the SHARP microscope $P(u_x, u_y)$ for the chosen field of view $(\Delta x, \Delta y)\epsilon(x_{max} - x_{min}, y_{max} - y_{min})$ are solved with spectral resolution $(\frac{1}{\Delta x}, \frac{1}{\Delta y})$, with a sensitivity of $< \lambda/203$, or about 0.5nm at the EUV wavelength of 13.5nm, as shown in Fig. 6.10

Finally, since the differential method relies only on differences of the on-axis illuminated speckle intensity spectrum $I(u)$ with the tilt illuminated speckle spectrum $I(u;\theta)$ to solve for the odd aberration, and the difference of the on-axis spectrum $I(u)$ with the flat background wave $\delta(u)$ to solve for the even aberrations, only three measurements suffice to solve the aberration $P(u_x, u_y)$ for any part of the field of view $(\Delta x, \Delta y)$ (Fig. 6.12). Hence a phase space description (x, y, u_x, u_y) of the aberrations of the speckle captured from a rough mask surface can be extracted by extracting the aberrations $P(u_x, u_y)$ over local domains in the field of view $(\Delta x, \Delta y)$, with consideration to the usual spatio-spectral resolution limitations $\Delta x \Delta u_x > 1$, where NA is the numerical aperture of the imaging system defined as the spectral passband of the microscope objective. Hence fast recovery of aberrations across the field of view with only illumination tilt measurements is possible - no moving parts are
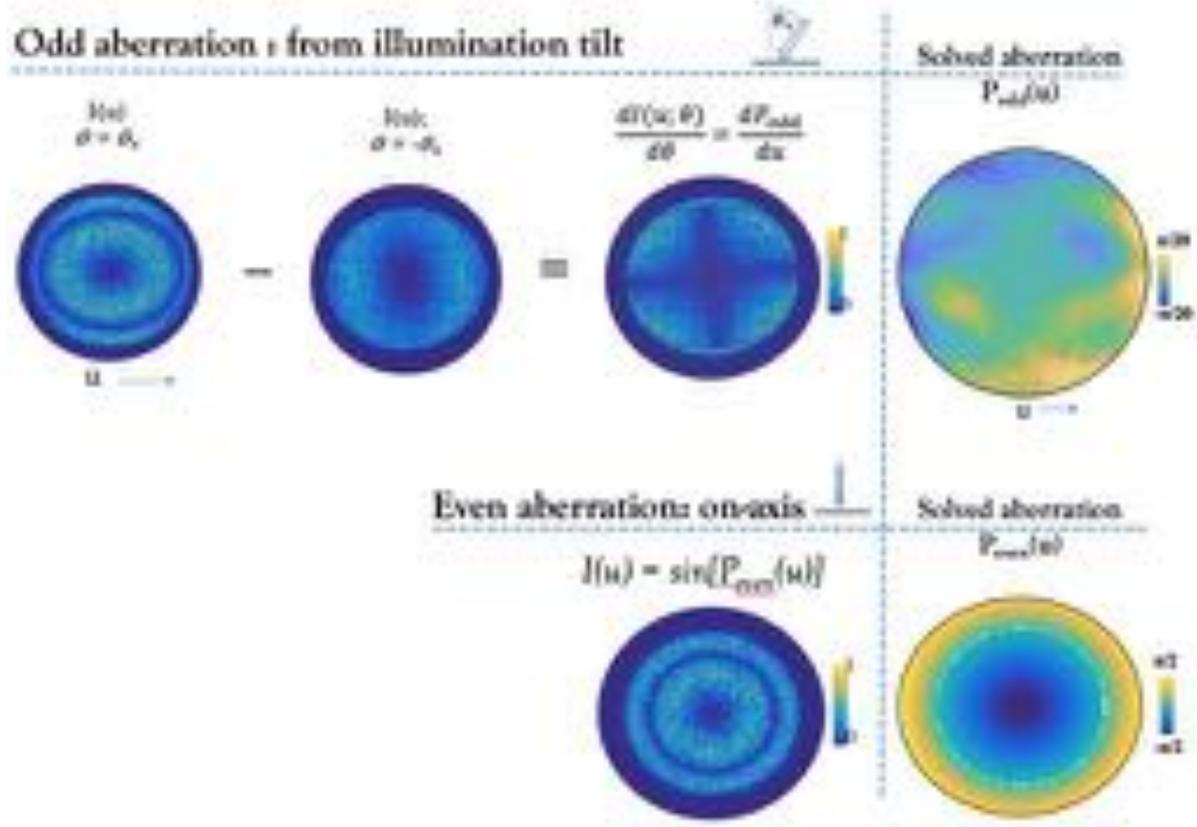
Figure 6.10: The solved aberration obtained by a minimum of two measurements - one on-axis to get the differential intensity against the background, and one at a slight illumination tilt to get the differential intensity with illumination. These are combined to obtain $\Phi(u)\nabla^2 P(u)$, speckle spectrum $\Phi(u)$ times the diffused pupil spectrum $\nabla^2 P(u)$, which is solved as a Poisson equation to obtain the aberration $P(u)$. Finally, the odd aberration is computed as $P_{odd}(u) = P(u) - P_{even}(u)$, still multiplied by the speckle spectrum $\Phi(u)$.

required except a scanning source, with sensitivity depending on the surface roughness and camera noise ($\lambda/20$ for the measurements shown in Fig. 6.10).

6.5 Conclusions

A differential method has been presented, that uses the speckle projected through an imaging system to estimate the transfer function of the imaging system. When the illumina-
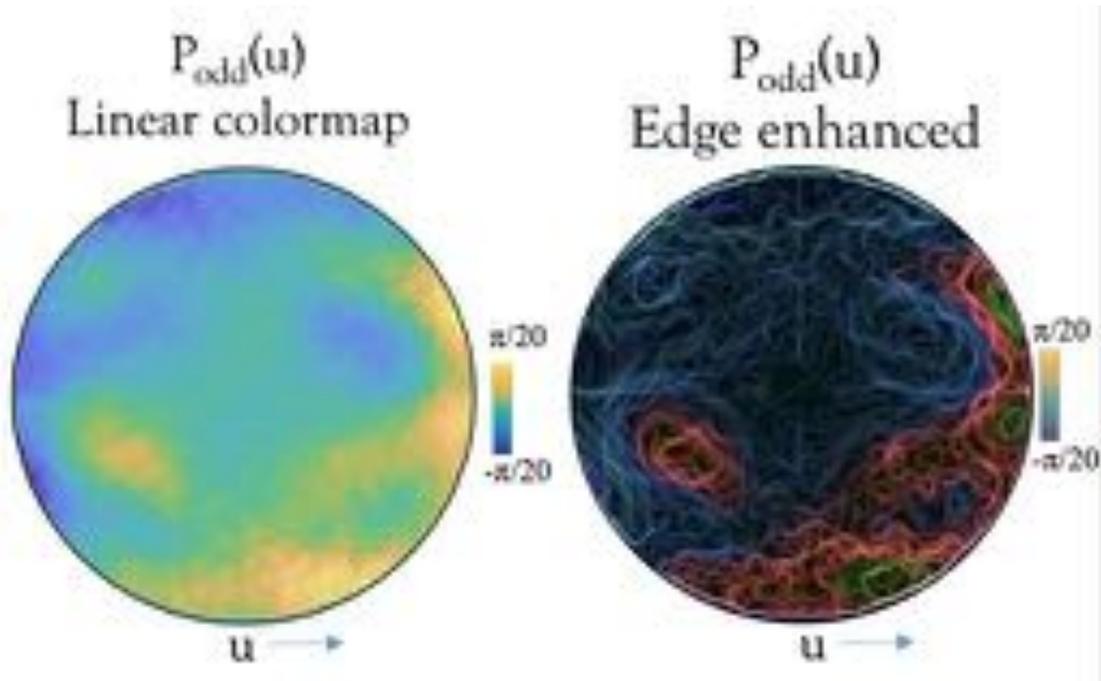
Figure 6.11: Visualization of the odd (centro-anti-symmetric) aberration from the center of the SHARP microscope's field of view. On the left odd pupil phase (in radians) in standard colormap -every part of the pupil has a negative counterpart in the opposite quadrant relative to the center. On the right, an edge enhanced visualization of $P_{odd}(u)$, showing that the aberrations peak near the edge of the pupil, where the beam from the synchotron makes the largest angles at the focusing zone-plate before reaching the camera. This implies an average lateral drift in the beam of $sin^{-1}(\frac{\pi}{10}) = 32\%$ across the diagonal direction with propagation.

tion is on-axis, only the centro-symmetric, or even aberrations are captured in the speckle at the camera plane. With a differential illumination tilt, the odd aberrations can be captured as well, enabling upto $\lambda/10$ sensitivity in measuring the imaging system's aberrated pupil. The method assumes weak aberrations and weakly scattering diffuser (strong background), assumptions that naturally hold in EUV lithography, making this a practical and fast method of in-situ aberration measurement in EUV litho imaging tools, using simply a mask blank.
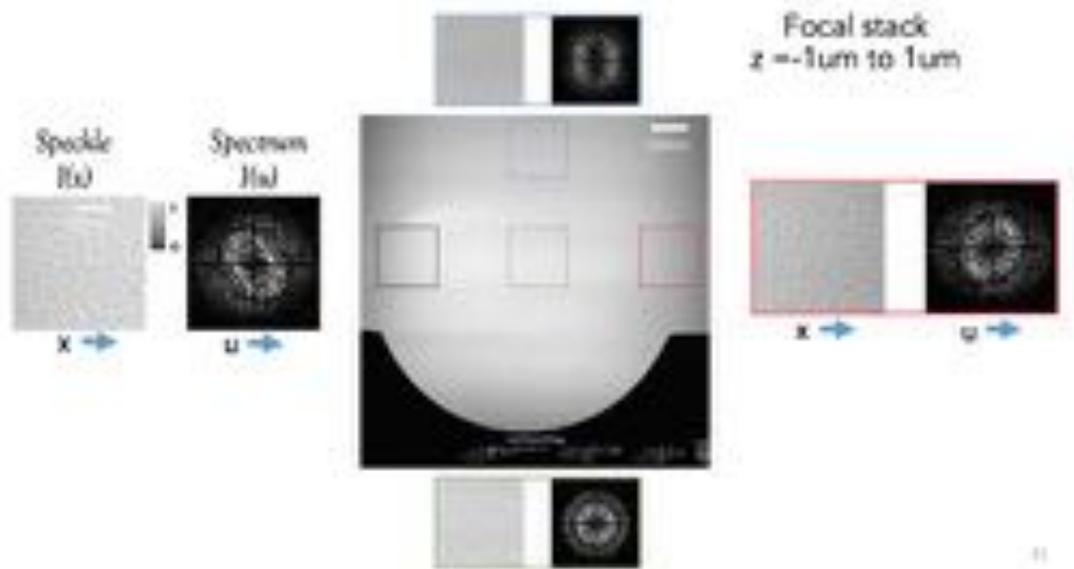
Figure 6.12: Phase space measurements of the field of view dependent aberrated point spread function of the SHARP EUV microscope. The even aberrations are encoded in the local speckle spectrum at each regional window of the field of view (δx), allowing for field of view dependent estimation of the even aberration P_{even}. The resolution trades off in the spectral domain with the size of the window chosen - large window size δx gives better spectral resolution for the recovered aberrated pupil $u_x = \frac{1}{\delta x}$, while a smaller window size estimates with better localization the aberrations in the chosen window within the field of view. Additionally, the odd aberrations are encoded in the change in the speckle spectrum with illumination angle $\delta \theta$ at each position (Fig. 6.9) , enabling a phase space description of the recovered pupil function $P(u; x)$ at each position of the field of view x for various illumination angles θ. $((c)A.Wojdyla)$

Chapter 7

Speckle capture with photoresist latent image

7.1 Introduction

Speckle in coherent light appears as a result of spontaneous symmetry breaking in a uniform coherent fields, due to roughness from mirrors, or from perturbations in the field that spread with propagation. In the extreme ultraviolet regime (13.5nm wavelength), since most objects are weakly scattering, the speckle formed has a strong unscattered component, also called the DC or background (Fig. 7.1). Imaging of the speckle on an aerial image sensor (for example, on a CCD camera) allows for direct measurement of the roughness in the light beam, which will then be translated onto the photoresist film (Fig. 7.2). Since line edge roughness engineering is possibly the most crucial challenge facing EUV lithography scaling in 2018, it can be argued that precise measurement and engineering of speckle is the first step in controlling roughness.

However aerial imaging sensors and cameras that measure the light falling onto silicon in a EUV printing tool, despite advantages of instant digitization of the data with fast analog to digital convertors, have the following shortcomings:

1. The point spread function of the resist material that gives rise to line edge roughness is inherently spherical, since the optical energy is distributed over the entire thickness of the resist, each point contributing photo-acid with a spherical Green's function. The final roughness after etch is a 2D cross-section of the 3D chemical process through the film; subsequently measuring the aerial image on a square array will exclude high frequencies, due limitations in sampling the transcendental π sphere with a square grid.

2. A CCD camera sensor is flat, measuring 2D sections of the 3D point spread function. Hence even with measurements with defocus [74], which measures the 2D image at various z-planes, there is inherent sampling ambiguity, since only a finite number of measurements can be made.

Hence here is proposed a new imaging modality - measure the aerial directly in the resist. However since the resist develop is inherently non-linear after etch (a thresholding process, in essence), we suggest measuring the aerial image directly on the resist before etch - i.e. the latent resist image of the EUV beam. It is shown that within a finite exposure region, the resist latent image, measured with an atomic force microscope (AFM), is inherently linear in the image intensity. Additionally, since the AFM has much better resolution than the CCD sensor, we also measure the aerial image with finer resolution, albeit with more uncertainty due to the resists blur function adding to the aerial image point spread function.

7.2 Pupil wavefront estimation from aerial image speckle

The weak speckle seen in EUV light can be used to estimate the optical transfer function, or Green's function, between the strongest speckle generating surface and the captured image, even before the resist is exposed, as shown in Ch. 6. Fig. 6.6 demonstrates the speckle and its spectrum for images of a blank mask measured in the SHARP [33] EUV microscope. Here
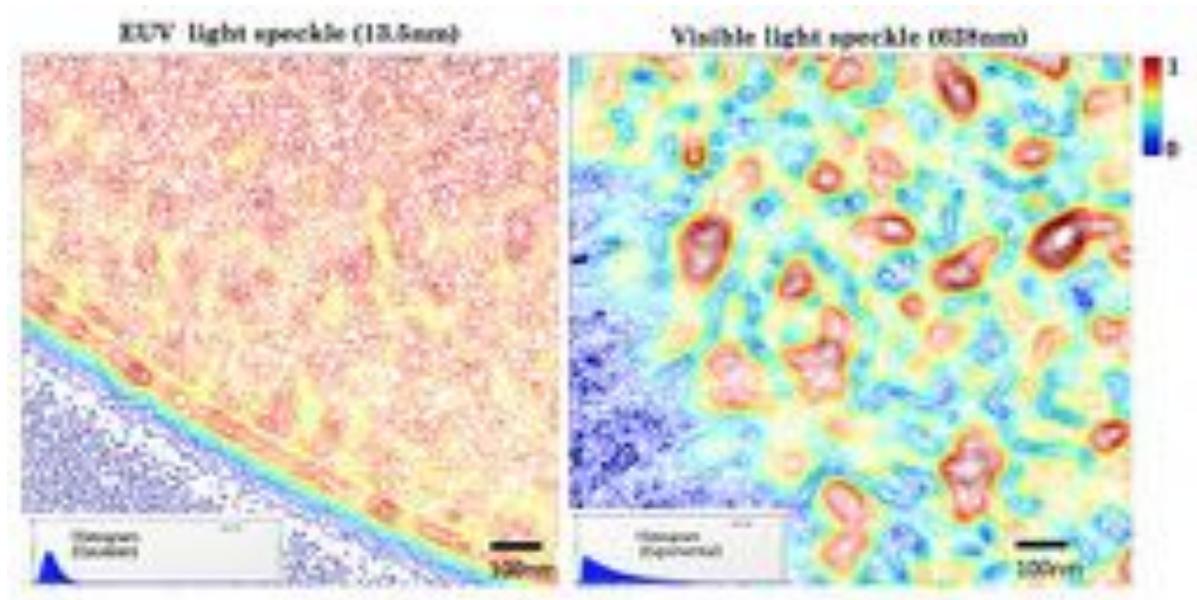
Figure 7.1: Electric field amplitude ($\sqrt{intensity}$) for coherent speckle - in the Extreme Ultra-Violet and optical regimes respectively, captured on a camera near the edge of a circular beam. Speckle arising due to stochastic effects, such as shot noise and roughness in the masks/optics spreads chaotically in a flat coherent light beam. The speckle in these images is purely in the aerial image, before photoresist exposure and develop, where substrate roughness and chemical noise further add to the final line edge roughness of the etched feature. In the optical regime, the speckle is exponentially distributed (fully developed), indicating the presence of null manifolds in the intensity distribution due to strong scattering, while in the EUV regime, the speckle has a gaussian distribution (partially developed), indicating a strong unscattered mean background with the scattering as an added perturbation. Despite larger contrast in the optical regime, speckle is a bigger challenge in EUV, due to the feature sizes ($< 20nm$) being comparable to the correlation length of the speckle blobs.
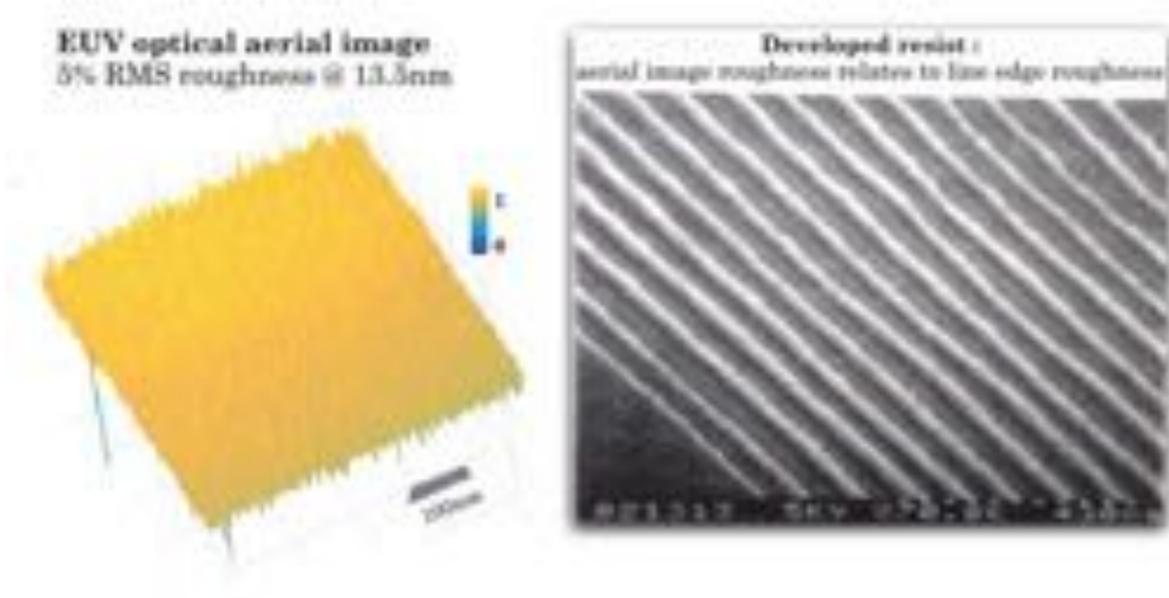
Figure 7.2: Roughness in the aerial image contributes to the final LER in the developed feature on wafer. Other contributors are the volumetric point spread function of the chemically amplified photoresist film, substrate roughness, and shot noise.

it is assumed that the aberration in reciprocal space P(u), is related to the Green's function or point spread function in real space, p(x) as the Fourier Transform ($\mathbb{F}$)

$$p(x) = \mathbb{F}[e^{iP(u)}] \tag{7.1}$$

According to the Born approximation for weak scattering, the mask roughness $\frac{\Phi(x)}{2\pi/\lambda}$ is approximated as the electric field $e^{i\Phi(x)} \approx 1 + i\Phi$, valid due to weak scattering in the EUV regime, with a strong unscattered background, or DC component ('1'). The imaginary valued mask roughness is imaged on to the wafer plane as a convolution with the imaginary part of the point spread function of the imaging system, p(x),

$$I(x) = 1 + i\Phi(x) \circledast i\mathbb{Im}[p(x)] \tag{7.2}$$

Where the '1' represents the unperturbed background; similar to phase contrast mi-

croscopy, the phase contrast in the speckle image $I(x)$ arising both due to the mask phase ($\Phi(x)$) and the imaginary part of the point spread function ($i\mathbb{I}m[p(x)]$). The power spectrum $I(u)$ reveals the imaging pupil spectrum $P(u)$ (equivalently the pupil wavefront aberration) as an interferogram in the Fourier domain (u), multiplied by the roughness spectrum $\Phi(u)$ as,[1]

$$I(u) - \delta(u) = \Phi(u)sin[P_{even}(u)] \tag{7.3}$$

hence the on-axis spectrum encodes the even part of the pupil wavefront aberration, where $P(u) = P_{even}(u) + P_{odd}(u)$. The odd part of the pupil wavefront[2] $P_{odd}(u)$ is obtained by tilting the illumination by angle θ corresponding to the spectral shift du_θ, to obtain

$$\frac{dI(u)}{du_\theta} - \delta(u) = \Phi(u)sin[\frac{P_{even}(u) - P_{even}(u + u_\theta)}{du_\theta}] = \Phi(u)sin[\frac{dP_{odd}(u)}{du}] \tag{7.4}$$

hence enabling recovery of odd valued pupil aberration functions, such as coma, trefoil etc., using only the differences between speckle intensity due to differential change in the illumination angle.[71] [3]

7.3 Aerial Imaging using latent resist

The previous section described aerial image measurements with a camera, artificially introduced into the image plane to capture the beam that would be incident on the photoresist. Could the aerial image be extracted directly from the photoresist in situ, in conditions identical to the projection printing process? Measuring the speckle in the resist can be utilized not just for inverse problems, such as imaging system point spread recovery, but also to understand the nature of the roughness transfer from the beam to the resist. Although here

---

[1]The derivation of Eqn. 7.3 uses the property $\mathbb{F}[\mathbb{I}m[p(x)]] = \mathbb{F}[p(x) - p^*(x)] = e^{iP(u)} - e^{-iP^*(-u)} = sin[P_{even}(u)]$.

[2]the imaging pupil $P(u)$ is called the pupil wavefront or pupil aberration interchangeably

[3]The derivation of Eqn. 7.4 uses the fact that even part of a shifted function is the derivative of its odd part.
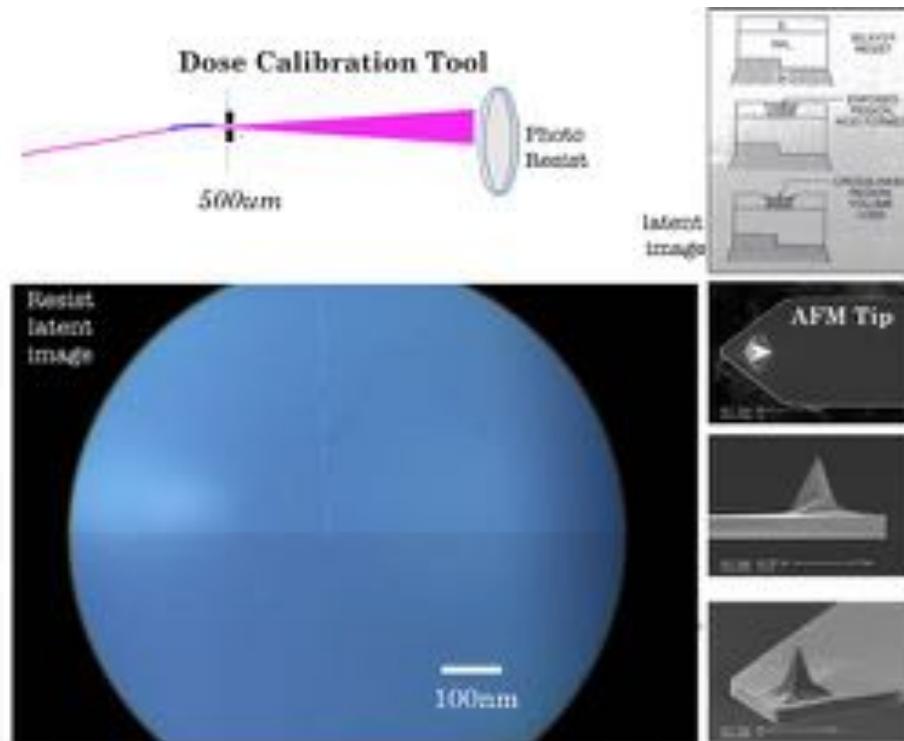
Figure 7.3: The aerial image from an exposure tool can be captured in the resist directly - analogous to film photography. The optical microscope image at 20x magnification (top left) shows contrast in the exposed region of the resist before develop. The process[7] (top right) is stopped before the non-linear develop, to preserve the latent resist image. The latent resist image shows slight contrast in the optical microscope (bottom left), with light blue in the exposed area, and dark blue in the unexposed area. The AFM tip (bottom right) is a vibrational tip which capacitively senses the surface of latent resist, with nanometer height resolution.

we measure and quantify the speckle only in the virgin latent image (before etch) as a first demonstration, the methodology can be extended for describing the roughness of the photoresist at various stages of development, enabling characterization of the combined spectral transfer function of the aerial imaging system and resist diffusion/develop chemistry, and enabling a new mathematical language of describing EUV stochastics during lithography process.

Measurement scheme

The dose calibration tool (DCT) at the Lawrence Berkeley National Lab is used to expose an industry standard photoresist film (EUVJ1915), followed by a post exposure bake. An AFM tip with sub-nm resolution (Bruker AFM, Santa Barbara, CA) is then used to extract the resist profile after the latent exposure (Fig. 7.3). The exposure is in the shape of a thumbprint (Fig. 7.4), capturing the stochastics of the EUV beam arriving at the resist from the synchotron beam. The atomic force microscope is a vibrating micro-electro-mechanical-sytem (MEMS), with an ultra sharp tip, with a resonance frequency that depends on the distance of the tip from the surface of the photoresist. The distance of the tip from the surface is modulated so as to keep the vibration amplitude constant at the resonant frequency - the modulation is read back as the surface height variation with sub-nm resolution, hence extracting an image of the optical pattern translated onto the photoresist film.

Measurements are performed in two modalities - the first at the edge of the exposure, in an attempt to detect diffraction rings (successfully detected in Fig. 7.5), with a high field of view (50um) and lower resolution ($\approx$ 10nm). The second measurement is at full resolution (sub 1nm) but for a lower field of view (2um) near the center of the exposure, measuring the fine structure in the beam roughness (Fig. 7.6). Multi-scale imaging is expected to be an important visualization method for photoresist roughness transfer functions as the exposure and develop process evolves during optical pattern transfer onto the photoresist, hence enabling a general vocabulary to quantify roughness during LER metrology.
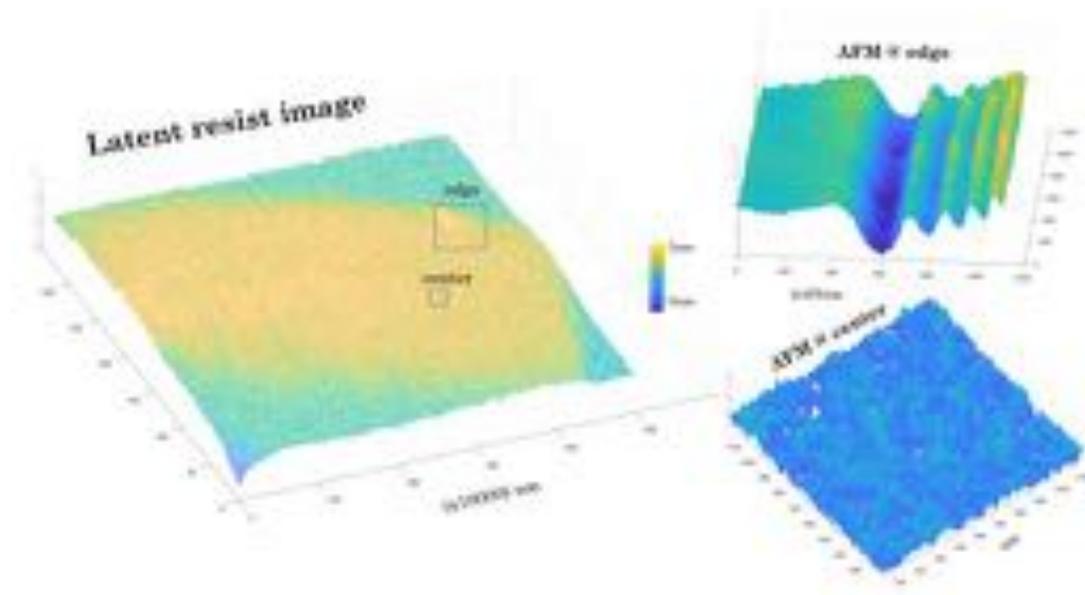
Figure 7.4: Thumbprint of the extreme ultraviolet light beam from the DCT exposure tool imprinted into the latent resist image, about 50um in length (left). Atomic force microscopy surface profiles of the latent resist image show optical signature of the beam in typical diffraction ripples(top right); beam speckle together with the resist blur function, are seen in the AFM surface profile near the center of the thumbprint (bottom right).

Viability as an aerial imaging method

Can the profile of the EUV optical beam roughness (speckle) imprinted on the resist be extracted from the resist height variation, as an in-situ aerial imaging scheme? To measure the EUV optical speckle over and above the intrinsic resist roughness seen in AFM measurements, the variance of the resist roughness measured with the flat EUV beam here, must be smaller than the roughness imparted to the resist from the EUV speckle arising in a lithography system; enabling extraction of the EUV optical speckle profile for implementing the spectral methods of Ch.6 for in-situ, photoresist based extraction of litho-tool aberrations.

The first AFM image of the resist at the edge of the beam 7.5 is used to extract the correspondence of EUV optical intensity to resist height. The resist contrast is measured as the dip in the ringing pattern for each row of the 2D ringing plot, and averaged over the
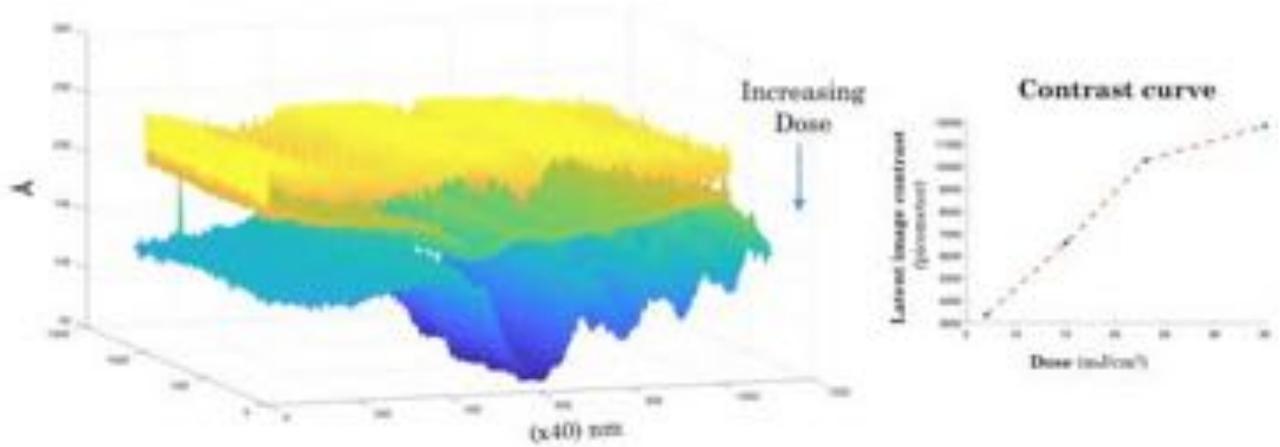
Figure 7.5: AFM scan of the latent resist image of and the EUV beam near the edge of the thumbprint, as a function of dose. The diffraction rings due to optical diffraction at the edge of the beam is seen to be linear in dose, as seen in the contrast curve, which is linear over the $5mJ/cm^2$ to $35mJ/cm^2$. This indicates that the latent image is a invertible function of the beam optical profile; hence the latent image serves as an analog volumetric camera for recording optical intensity, enabling in-situ imaging of exposure tool optics without a CCD camera sensor at the output.

columns. The contrast variation with exposure dose in this case shows the resist shrinkage dependence on the optical intensity to be $\approx 5Angstrom/mJ/cm^2$ of exposure.

The second AFM image of the rough resist near the center of the beam with varying exposures is plotted in the spectral domain as Power Spectrum Density ($|\mathbb{F}|^2$) on a log-log scale, showing the energy distribution of the rough speckle in the spectral domain (Fig. 7.6). Within the spectral passband of the SHARP tool (about $50lines/um$ at NA=0.33), the resist roughness PSD transmits 96% of the total energy, corresponding to 1.6nm (RMS) intrinsic roughness in the photoresist for flat beam exposure.

Finally we speculate on using the latent resist image for optical speckle metrology. The EUV optical speckle in an imaging system must have larger RMS roughness than the intrinsic resist roughness (1.6nm) to be extracted from the latent image. This implies that an EUV beam with 5% RMS speckle (Fig. 7.2), must have the minimum brightness $Dose_{min} = 1.6nm/[5Angstrom/mJ/cm^2]/[5\%] \approx 60mJ/cm^2$ to create latent resist contrast greater than
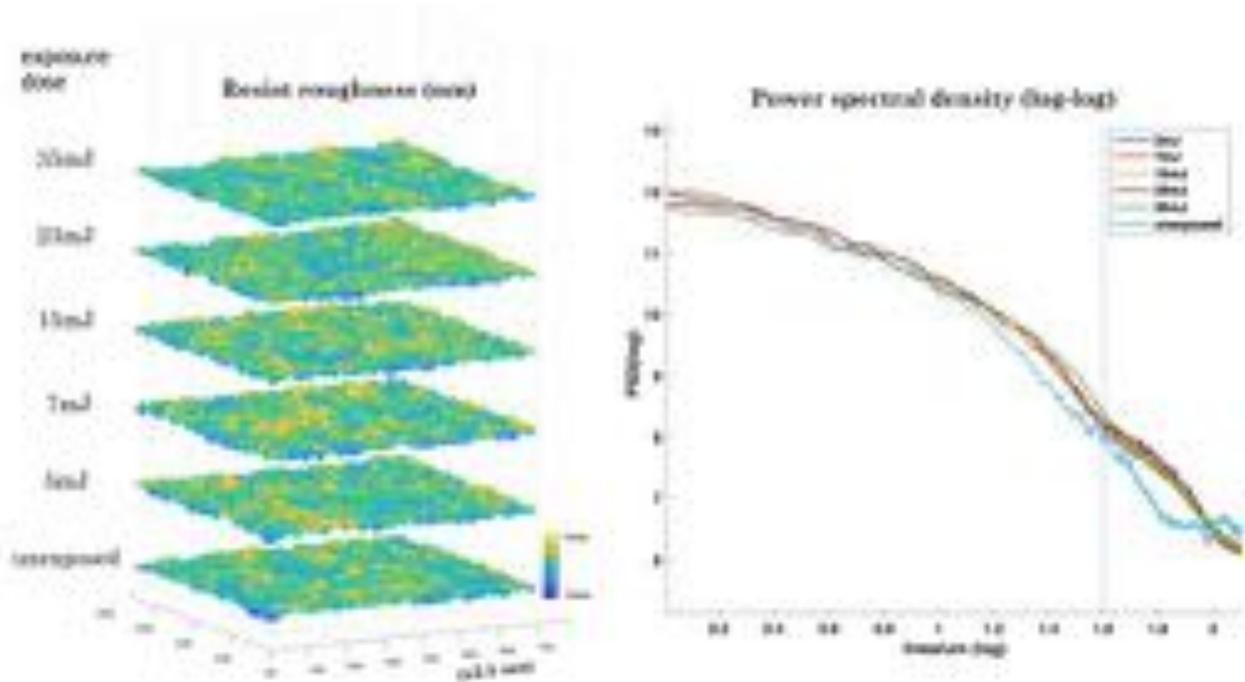
Figure 7.6: AFM measurement of surface roughness on the latent resist image (before etch) for various exposure dose values (left). The power spectral density after exposure for the roughness differs only at the higher frequencies, the vertical dotted line indicating the pass-band for SHARP aerial imaging tool (about 50lines/um).

the intrinsic roughness of 1.6nm rms. $60mJ/cm^2$ is at the upper limit of the exposure range used here, and about 5x the nominal dose of $\approx 10mJ/cm^2$ during typical EUV lithography exposure. The high dose requirements could however be engineered around by increasing the roughness in the EUV optical beam ($> 5\%$) by using a custom rough mask, or by improving resist sensitivity ($< 1.6nm$), to enable recovery of the aerial image speckle from the latent resist's surface roughness. However the weak scattering approximation is less valid for higher percentage roughness in the EUV optical beam - hence the utility of the captured speckle for aberration retrieval reduces as the speckle contrast increases. Thus the ability to record EUV optical speckle on resist and further use as an aberration retrieval method, depends largely on the properties of the photoresist used.
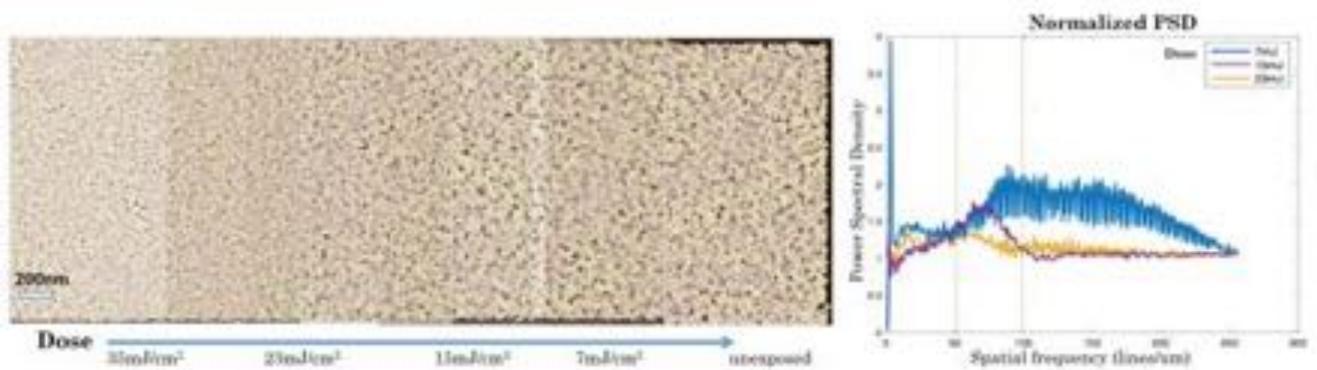
Figure 7.7: The atomic force microscopy (AFM) scan at sub-nm resolution showing the surface texture of the latent resist image as a function of increasing dose, shown here as a composite mosaic of four exposures (left)[91]. The characteristic grain texture changes smoothly as a function of increasing optical dose. The power spectrum density of the exposed surface roughness normalized by the power spectrum of the unexposed resist surface (right) for three exposures. The lines cross at about $50um^{-1}$, interpreted as the characteristic grain frequency of the unexposed resist. Effective imaging regime is potentially between the specific spectral range of $50um^{-1}$ to $100um^{-1}$ for this photoresist, where the power spectrum is monotonic in the dose, and the optical beam's contribution dominates the photoresist stochastics.

Quantifying resist roughness

To quantify the multi-scale stochastics of the rough resist, consider the fractal dimension [52] of an infinitely detailed but bounded structure - the fractal dimension for a fractal mass density function ρ_d as it scales along its linear dimension d is given as

$$\mathbb{F}_d = \frac{log(\rho_d)}{log(d)} \tag{7.5}$$

Hence fractional dimension $\mathbb{F}_d = 2$ for a square, $\mathbb{F}_d = 3$ for a cube, and so forth, where values of N are integers for regular polygons and equal to the number of corners on the N-D manifold. For spheres however, the fractal dimension is given as $\frac{log\pi + logr^2}{logr} = 2 + log_r\pi$, which is transcendental - illustrating for instance how a resist's topology after develop (with Green's functions that is spherical) cannot be perfectly measured with a square CCD

sensor pixel, since the fractal dimensions are not fractional multiples of each other, i.e. $\mathbb{F}_d[resist]/\mathbb{F}_d[camera]$ is transcendental (or irrational).

Measuring the aerial image on photoresist, despite being more challenging since the analog to digital conversion requires an AFM tip, promises to reveal fractal dimensions and recursive structures with greater fidelity. To visualize the image transfer into the resist, the latent resist image measurements are tiled to form a continuous mosaic, after exposure and bake, but before etch in resist develop process (Fig. 7.4a). Since, the AFM measurements of the larger ($50um$) region at the edge of the beam demonstrates the monotonicity of the resist shrink depth in the latent image with the exposed dose (Fig. 7.5), the inverse problem of aerial image estimation from latent image print is uniquely posed, or monotonically invertible, in the dose range of $5 - 35mJ/cm^2$ chosen here. Within a range of $50 - 100lines/um$, the surface roughness transfer function is defined as [55]

$$SRTF(\frac{2\pi}{x}, \frac{2\pi}{y}) = \frac{|\mathbb{F}[I_{dose}(x,y)]|^2}{|\mathbb{F}[I_{unexposed}(x,y)]|^2} \tag{7.6}$$

where $\mathbb{F}$ is the Fourier transform operation followed by a radial average. The radial average gives a one dimensional plot of the transfer function of the beam roughness (numerator) from the resist roughness (denominator), shown as the power spectral density (Fig. 7.7) as a function of spatial frequency. Within the $50 - 100lines/um$ spectral range indicated in Fig. 7.7, the PSD of the surface roughness transfer follows positive scaling with spatial frequency for the lowest dose (positive fractal dimension); for the higher dose, the scaling is negative, implying a switch in the fractal dimension around $15mJ/cm^2$.

Resist exposure is essentially a burning process, creating fractal structures in the 3D film volume, projected into 2D during etch. Plotting the power spectrum in the Fourier or spectral domain $x \to \frac{2\pi}{x}$ is equivalent to computing the negative fractal dimension for each spatial frequency since $log(\frac{2\pi}{x}) = -log(x)$(Eqn. 7.5).

Positive fractal dimension for higher doses ($> 15mJ/cm^2$) implies an increasing mass density scaling with spatial dimension, seen on the left most mosaic image of Fig. 7.7.

For lower doses the resist is more cracked; hence the mass density decreases with spatial dimension d, i.e. zooming in will reduce the total density of the area represented. since the dots dominate instead of the bright hills (Fig. 7.7).Hence it can be conjectured (without yet validating in this paper) that at a given exposure sweet spot the resist image has zero fractal dimension ($\mathbb{F}_d = 0$, when it switches from positive to negative fractal dimension, then the speckle in the latent image is linear in the aerial image intensity. In Fig. 7.7 this would be near the inflection point at 15mJ/cm^2 exposure (purple curve), at about 80 $lines/um$ pitch, when the imaging on latent resist is the most linear in the aerial image intensity. A similar spectral/fractal analysis can be extended to find the sweet spot in the dose range for any commercial photoresist, about which the latent image is linear and acts a linear analog camera for the optical intensity. However since the minimum dose requirement to over-ride resist intrinsic resist roughness has been shown to be $50mJ/cm^2$, the current photoresist may have diminished spectral sensitivity to aerial image fluctuations when there enough optical speckle to create contrast. In principle, an ideal photoresist for aerial imaging purposes, will have the exposure sweet spot above the minimum dose needed to overcome the intrinsic resist roughness.

7.4 Conclusions

We show that photoresist latent image could be viable as a EUV speckle measurement scheme in the $60mJ/cm^2$ dose range; the dose requirement can be further relaxed by having stronger speckle in the EUV tool, or by engineering more sensitive photoresist material. Analysis of photoresist roughness stochastics in terms of the fractal dimension ($\mathbb{F}_d$) and power spectrum density ($|\mathbb{F}_d|^2$) are powerful tools for describing stochastic process in EUV lithography, both in the aerial imaging and resist develop process. Here, a proof of principle method for imaging the aerial image optical profile on the photoresist latent image is shown, with fractional dimension and power spectrum analysis of the transferred pattern revealing regions of linearity between the optical and resist patterns . $\mathbb{F}_d$ and $|\mathbb{F}|^2$ as first order stochas-

tic quantization methods additionally enable an intuitive understanding of the resist develop process in terms of the proportion of hills, valleys and singular points in the topography. This paper barely begins to suggest, with a simple application of aerial image retrieval, the potential applicability of more sophisticated mathematical methods to describe and control stochastic processes in EUV lithography. As the industry moves forward, more attention needs to be given to detail at the sub-10nm litho nodes, to avoid endless trial and repetition in the fabrication facility or during process development, especially when encountering intractable stochastic methods. Hence careful metrology and analysis with mathematical methods that are already mature need to be introduced in the industry at this crucial time in the eve of Moore's law, enabling greater control and design of stochastic processes, line edge roughness, and eventually breaking through into quantum regime of lithography, where all processes are inherently stochastic.

Chapter 8

Learning from lithography : speckle as optical turbulence

The differential equations underlying the non-linear dynamics of turbulent fluid flow are extended to the transport of optical energy and optical phase in coherent laser speckle. Since light acts like a pure, incompressible, inviscid fluid, the momentum and mass transport of the Navier-Stokes equations translate directly to the intensity and phase transport equations in scalar diffraction theory, respectively. The non-linear term in the phase transport equation describes the emergence of turbulent dynamics near cusps and singularities in the optical field, enabling insights into the topology and dynamics of turbulence/speckle in 6D phase space. Turbulence is demonstrated on camera, speckle intensity measured through-volume.

i am an old man now, and
when i die and go to heaven
there are two matters on which
i hope for enlightenment. One
is relativity or quantum
electrodynamics, and the other
is turbulent motion of fluids.
About the former i am rather
optimistic.- Werner Heisenberg

Light is a super fluid, with no rest mass (latin root "light" for weightlessness), as well as no stickiness or viscosity characteristic of shear stress in fluid mechanics. Hence, the wave mechanics of light propagation encompasses only the inviscid aspect of fluid dynamics, allowing for a simplification of the laws of fluid mechanics in their differential form to describe the laws of optical propagation. Local descriptions of optical dynamics with partial differential equations (PDEs) parallel the Navier-Stokes equations, where the optical intensity plays the role of fluid mass, and the optical phase is the fluid momentum, guiding the propagation of mass. Additionally since there is no viscosity, the laws of optical propagation are isomorphic to the Euler equation, the inviscid form of the Navier-Stokes equation for flow.

Due to its low viscosity or high Reynolds number(R), coherent light enters turbulent regimes of flow rapidly due to even slight disturbances, causing laser beams to eventually develop granular stochastics called 'speckle' - essentially turbulent light. The mathematical model starts with the fundamental differential equation of electric field propagation (Helmholtz equation), which models light as a diffusion process in x-y, with propagation in the z direction with temporal frequency ω, such that $E_{x,y,z,t} = E_{x,y}e^{ikz+i\omega t}$, then

$$\nabla_{xy}^2 \vec{E}_{xy} + k^2 \vec{E}_{xy} = 0 \qquad (8.1)$$

is the effective diffusion Laplacian with a preferential direction of propagation (z); $k^2 = \frac{\omega^2}{c^2}$ is the dispersion relation, linear for light; c is the speed of light. The electric field is often represented as a complex valued vector at each (x,y), $E_{xy} = \sqrt{I(x,y)}e^{i\varphi(x,y)}$ - by splitting into two variables I and φ that are conjugates of each other, the dynamics of the equation alternates between the two at each propagation step. Here I is the optical intensity, analogous to the mass density of a fluid ρ; $\vec{\nabla}\varphi$ is analogous to the fluid velocity vector $\vec{v}$, energy flows in the direction of momentum, or equivalently of phase gradients (Huygen's principle). The differential equations of fluid flow consist of mass (ρ) and momentum($\rho\vec{v}$) transport; the mass transport equation

for fluid,

$$\frac{\delta\rho}{\delta t} = \vec{\nabla}.\rho\vec{v} \tag{8.2}$$

for light,

$$k\frac{\delta I}{\delta z} = \vec{\nabla}_{xy} \cdot I\vec{\nabla}_{xy}\varphi \tag{8.3}$$

The Transport of Intensity Equation, or TIE (Eqn. 8.3)is isomorphic with the continuity equation of fluid transport (Eqn. 8.2). It is derived from the Helmholtz equation (Eqn. 8.1) by convolving the eikonal form of the field $\vec{E} = \sqrt{I(x,y)}e^{i\varphi(x,y)}$ with the paraxial Green's function $G = e^{ikz}e^{ikz(x^2+y^2)/2}$ of Eqn.8.1.

$$I(x,y) = |E_{xy} * e^{ikz\frac{x^2+y^2}{2}}|^2 \approx |E_{xy} + i\lambda z\nabla^2_{xy}E_{xy}|^2 \tag{8.4}$$

The paraxial approximation assumes all third order derivatives, as well as non-linear terms in phase or intensity to be negligibly small, yielding the intensity transport equation 8.3. The transport of intensity can be expanded as $I\nabla^2\varphi + \vec{\nabla}I \cdot \vec{\nabla}\varphi$, linear in both intensity I and phase φ, the first and second terms corresponding to focusing by parabolic surfaces and bending by a prism, respectively. The non-linear, turbulent terms arise in the momentum ($\rho\vec{v}$) transport equation, or the Navier Stokes under pressure p,

$$\frac{\delta\vec{\nu}}{\delta t} = -(\vec{\nu}.\vec{\nabla})\vec{\nu} - \frac{\vec{\nabla}p}{\rho} \tag{8.5}$$

For irrotational flow, substitute $\vec{v} = \vec{\nabla}\varphi$, where φ is a scalar potential, to get the equation of a streamline (Eqn. 8.6). For light, the equivalent equation of momentum transport is the transport of phase equation obtained from the imaginary part of the Helmholtz eqn (Eqn. 8.1) with the eikonal form of $\vec{E} = \sqrt{I}e^{i\varphi}$, or by taking the argument of the Green's function convolution (Eqn. 8.4),

for fluid,

$$\frac{\delta\varphi}{\delta t} = -\frac{(\nabla\varphi)^2}{2} - \frac{p}{\rho} \tag{8.6}$$

for light,

$$\frac{\delta\varphi}{\delta z} = -\frac{(\nabla_{xy}\varphi)^2}{2} - \frac{\nabla_{xy}^2\sqrt{I}}{\sqrt{I}} \tag{8.7}$$

which describe the momentum or phase transport of fluid and light respectively. The first non-linear term $(\nabla\varphi)^2$ is the Bernoulli flow along a streamline, the second term is the pressure along the streamline. Interestingly, the curvature of the electric field amplitude $(\nabla^2\sqrt{I})$ acts as optical pressure, causing phase to evolve locally near regions of amplitude curvature. The level sets of this second non-linear term are the positions of the zero crossing of the electric field, where both $\nabla^2\sqrt{I}$ and $\frac{1}{\sqrt{I}}$ go to infinity (Fig. 8.1b). This suggests a Bernoulli's law for free space light propagation - the same way pressure is the inverse of flow speed, allowing airplanes to have lift with larger wing curvature on the top than the bottom, areas in an optical field with large phase flux (red lines in Fig. 8.1b) have a low pressure, causing energy to be suctioned towards cusps and singularities; much like the role of gravity in general relativity.

The phase φ can be retrieved from the coupled PDEs (Eqns.8.3,8.7 or 8.2,8.6) within a Hamilton-Jacobi variational formulation. The computational phase retrieval is posed as a minimization over the measured intensity transport with propagation z (Fig. 8.1a),with an additional constraint of the transport of phase.

$$\min_{\varphi} \left\| k\frac{\delta I}{\delta z} - \vec{\nabla}_{xy} \cdot I\vec{\nabla}_{xy}\varphi \right\|_2^2 \quad \text{s.t.} \quad \frac{\delta\varphi}{\delta z} = -\frac{(\nabla_{xy}\varphi)^2}{2} - \frac{\nabla_{xy}^2\sqrt{I}}{\sqrt{I}} \tag{8.8}$$

Once the phase recovery for turbulent optical flow is posed as coupled PDEs - a process equivalent to unfolding manifolds in phase-space - many of the numerical and analytical tools used in the turbulence theories of modern physics (e.g. Kolmogorov) are applicable, with new understanding gained from free space experiments with light, complemented with current computational modeling and learning approaches. The applications are wide ranging - imaging through turbulent media (such as the brain, retina, stars), cosmological & relativistic fluid mechanics, magnetohydrodynamics of plasmas among others.
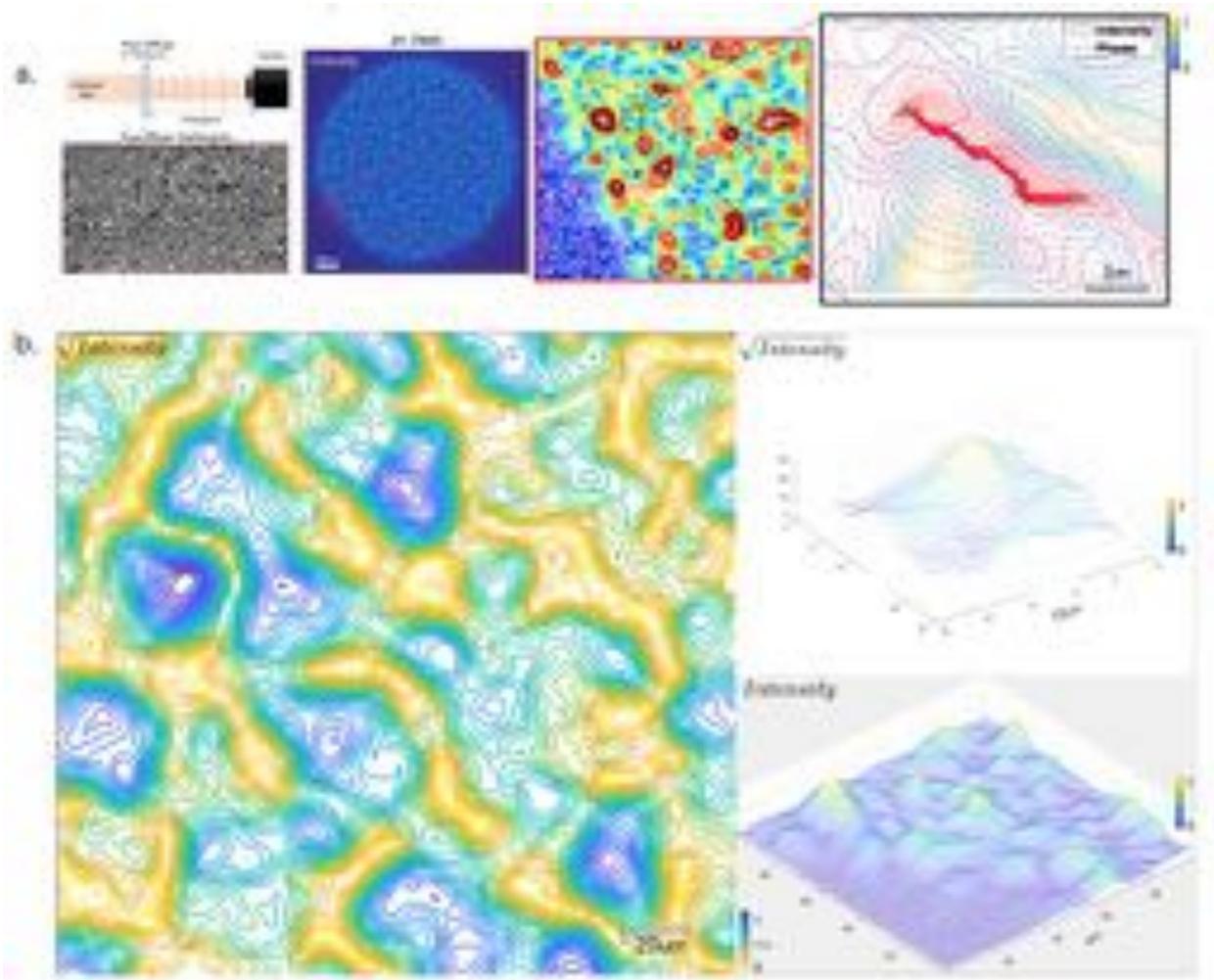
Figure 8.1: (a) Coherent speckle intensity measured as a function of propagation, the camera measuring a 3D volume by stepping through defocus or with propagation; turbulence is seen to appear as two dimensional singularity manifolds in the measured 3D speckle intensity. In the inset, this equates to dark rivers (blue) permeating the bright intensity hills (orange). In the in-inset, $\vec{\nabla} I \perp \vec{\nabla} \varphi$ indicate point singularities, with each line singularity terminating at point singularities with opposite topological charge at either end. (b) In negative contrast, the singularities pop (yellow) against intensity valleys (blue), a network of 1D lines in the 2D intensity terrain, the positions of singularity cusps indicating divergence or convergence of phase gradients or fluid momentum; measured intensity on the camera sensor is smooth, since cusps are swallowed when amplitude is squared.

Appendix A

Imaging with speckle: scattering in astronomy and microscopy

"I used to measure the heavens, now i shall measure the shadow of the earth, although my soul was from heaven, the shadow of my body lies here " – Johannes Kepler

Observing the primal light of the cosmos, arguably the first human science of astronomy, probes structure in the sky with optical instrumentation. The complementary science of microscopy has evolved more recently, concerning itself with detailing optically active materials in the earth - discerning structure in organisms, starting from the cellular scale down to molecular intracellular structures using shorter wavelengths (visible -> UV -> X-Ray). These are further assisted by wave-front imaging methods such as phase contrast microscopy for measuring optical density, and statistical methods for super-resolution. Additionally, microscopy techniques have birthed the art of lithography, printing circuits at the nanoscale on inorganic silicon, that has ushered in the computational age – opening digital storage and processing capabilities, further enabling numerical modeling that is the now the workhorse for solving physical systems across the imaging sciences.

The methods of electromagnetic inverse problem gleaned from the microscopy lab can be often translated directly to the astronomical sciences. Innovation often comes from cross-pollination between allied yet distinct disciplines, and navigating their underlying similarities and differences. Migrating from looking at scattering structures from the earth (microscopy) or imaging nanostructures on silicon (lithography), to looking at point objects in the sky (astronomy) retains the same underlying laws of electromagnetic physics. Yet it entails vast differences - a flipping from linear, bright field coherent imaging to non-linear, dark field incoherent imaging, from looking down at the details that support us to to looking up at the vastness that surrounds us.
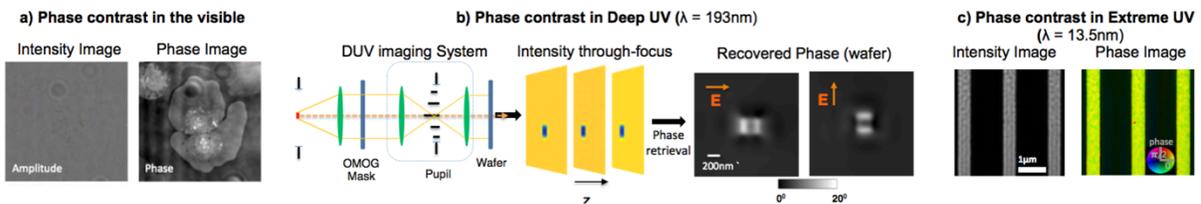
Fig 1. Wavefront recovery, or phase imaging, applied to microscopy and lithography. Fundamental tradeoffs between spatial and spectral resolution mean that any phase imaging method requires at least two measurements to recover the complex valued field at each position. Alternatively, as seen in Shack Hartmann style sensing, spatial resolution can be traded off with spectral resolution, depending upon the application.

The core of the scientific method lies in recreating at various spatial and temporal scales realities that can be repeated, first as scientific theories, but more definitively as empirical experiments. Hence spatio-temporal or spectral methods applied to astronomical imaging can be greatly assisted by tabletop methods, particularly in simplifying experiments (fewer optical components), envisioning algorithms for extracting phase information from fewer images and simpler hardware, and designing methods for faster, more reliable methods of solving the inverse problem – reconstruction of the source or object from the image. Hence it is crucial for understanding the cosmos to have numerical, or experimental models on the ground, where design choices can be made and tested.

Nature repeats itself at various scale, the world emerging out of a primitive blueprint that rearranges information telescopically through space and time. The structure of the sky is best reflected in the structures in the sand, the laws of coherent optics easily replicated in fluid dynamics of smoke or water. Starlight is typically an ensemble of point sources, each coherent with itself, but incoherent with each other, and in this way similar to coherent microscopy. Hence techniques such as defocus based phase imaging [Shanker 2014] for coherent light, or fluorescence microscopy (e.g. depth estimation using rotating PSF [Piestun 2008] can be directly extended to telescopes.

Kepler believed that the structure in the external world (cosmos) is best represented by geometric shapes (of the mind), math being the original divine language. In electromagnetic theory, the conserved quantity is information, typically represented as volume in phase-space (also called the space-bandwidth product in microscopy). It is the key design variable in an imaging system, restricted by uncertainty principles dependent of the wavelength of light while imaging, or the plank length in the quantum regime. This Wigner distribution (or phase-space representation), the spectral transform of the second order electric field,represents a variable and its conjugate variable simultaneously, phase imaging corresponding to the first order moment of phase space. The topology of the energy distribution in phase – space corresponds to all the information about the source (star), scattering medium (atmosphere) and imaging optic, in a 6 dimensional manifold (3D space x2).

Scattering from the atmosphere introduces wavefront errors that have to be corrected with AO devices, and typically considered a nuisance. However, as shown in our recent experiments coherent speckle engineering, scattering is advantageous in many ways – it maximizes information transfer if places between the source of interest and the receiver. A rough surface has a large absorption cross section, since the roughness encodes all spatial frequencies, and hence can receive the maximum angular spectrum. Hence wide field of view imaging is enabled in the presence of a strong scattering medium before the detector. This enables maximum information transfer (or Shannon entropy), although the increase in entropy also means that the inverse problem is harder, i.e. more information implies greater difficulty in decoding.

Hence a rough surface collects the most photons, but is the worst imager (since the information is maximally scrambled). However, it is possible to navigate this fine line between structure and randomness, or between most order and most information, by knowing the properties of the scattering medium. We've shown for instance, that weak speckle within the refractive regime in coherent EUV imaging allows for wavefront imaging in the pupil of the imaging system (Fig. 2) from a single image [Shanker 2016]. This could be directly applied to an adaptive optics system for instance, except that atmospheric turbulence tends to be strong, and cannot be directly engineered. Yet the converse is also possible (since astronomy is inverse microscopy) – the adaptive optic shape being known, the properties of the scattering medium can be modeled. Hence turbulence, despite being a nuisance in the information sense of adding entropy to the system, multiplexes more, enabling for instance an extended field of view when imaging with a telescope.
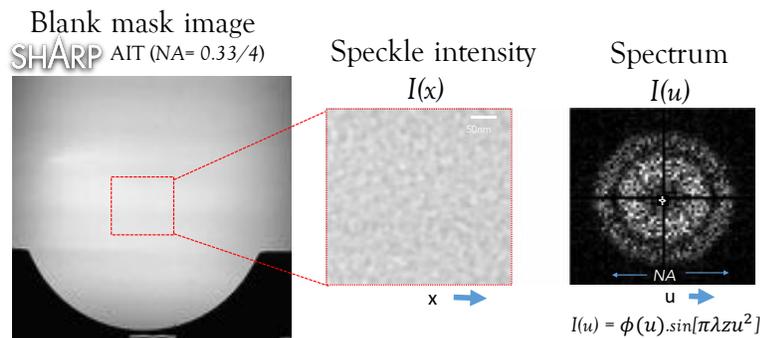
Fig 2. Mask roughness on an EUV mask images as speckle. The spectrum of the speckle gives the transfer function (in this case defocus zu^2) multiplied by the speckle spectrum $\emptyset(u)$, which holds under a linear scattering approximation. Conversely, knowing the shape of the aberration (or adaptive optic) can be used to estimate the properties of the scattering medium, or the speckle. The central order in the spectrum $I(u)$ would correspond to an image of the star.

Power spectrum ($I(u)$ in Fig 2) is really a bispectral analysis of the real space speckle [Wang 2011], since in Fourier domain the power spectrum corresponds to autocorrelation in real space. Higher order correlations can be expected to reveal properties corresponding to higher order scattering, moving gradually from the linearization

assumed for EUV speckle to strong scattering in highly turbulent media. Visualization of energy propagation in phase space allows methods such spatial frequency multiplexing of the scattered spectrum using point sources at multiple illumination angles (stars), also called Fourier Ptycography [Tian 2014], to be applied to astronomy, where the point sources are naturally diverse in angular content. Strong scattering further reveals vorticity in the optical flow, as seen in phase measurements of optical speckle fields showing spontaneous generation, sustenance and annihilation of phase vortices in strong scattering (Fig 3). This is related to communication is free space through turbulent media [Wang 2011], where orbital angular momentum is used to preserve holographic structure in the transmitting beam, another application of imaging through the atmosphere. Vorticity and strong scattering are closely related to each other, underlying the the dynamics of volumetric optical scattering.
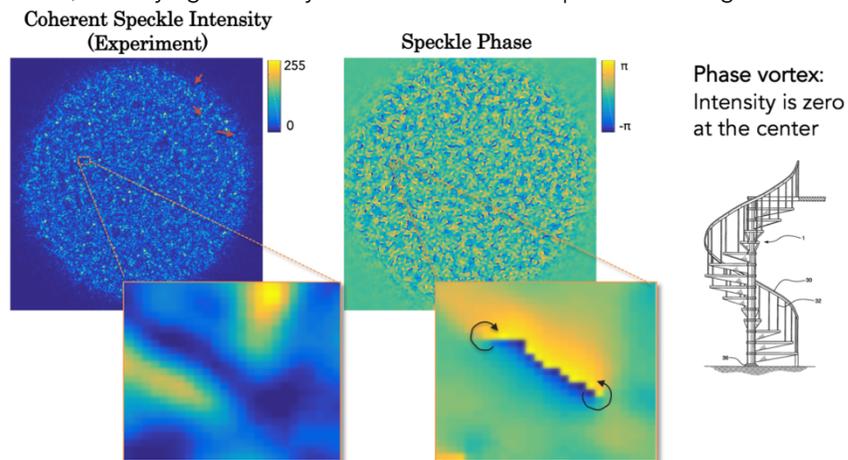
Fig 3. Strong optical speckle pattern where the dark singularities are the positions of vorticity in phase distribution. The intensity nulls thread through the optical volume, forming the backbone of the optical structure. Scattering in astronomy is coherent too, for each point source, allowing for characterizing the scattering medium by looking for intensity nulls.

Many of the methods used in astronomy are shared with microscopy – correlation based speckle analysis, point spread function engineering, adaptive optics for aberration removal etc. Stars are point objects, and easily lend themselves to methods in coherent optics. The underlying wave mechanics of optical imaging of the cosmos are much the same as that in fluid dynamics of liquids and gases and coherent sources closer to earth. As creatures tied to the the 2D dimensional spherical shell looking outward at our ancestry, the only way that we can understand ourselves, and our relationship to nature and its creative processes, is by replicating the machinery of the cosmos with the tools and methods accessible on the ground, conversely advancing our knowledge of immediate local reality from our relative history in context of the rest of the universe.

References :
[Shanker 2014] Shanker et al Appl. Opt. 53, J1-J6, 2014
[Piestum 2008] Opt. Express 16, 3484-3489, 2008
[Shanker 2016] Applied Optics 2016, ITh1F.2 (Optical Society of America, 2016)
[Lohmann 1983] Appl. Opt. 22, 4028-4037 (1983)
[Tian 2014] Biomed. Opt. Express 5, 2376-2389 (2014)
[Wang 2011] Opt. Express 19, 482-492 (2011)

Bibliography

[1] K. Adam and A.R. Neureuther. "Simplified models for edge transitions in rigorous mask modelling". In: *Proc. of the SPIE* 4346 (2001), pp. 331–344.

[2] Konstantinos Adam. "Modeling of electromagnetic effects from mask topography at full-chip scale". In: *Proc. SPIE* 5754 (2005), pp. 498–505. DOI: 10.1117/12.600139. URL: http://dx.doi.org/10.1117/12.600139.

[3] L. J. Allen et al. "Phase retrieval from images in the presence of first-order vortices". In: *Phys. Rev. E* 63.3 (2001), p. 037602.

[4] L.J. Allen and M.P. Oxley. "Phase retrieval from series of images obtained by defocus variation". In: *Optics Communications* 199.1–4 (2001), pp. 65–75. ISSN: 0030-4018.

[5] Jaione Tirapu Azpiroz et al. "Critical impact of mask electromagnetic effects on optical proximity corrections performance for 45 nm and beyond". In: *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures* 25.1 (2007), pp. 164–168.

[6] Bevan B Baker and Edward Thomas Copson. *The mathematical theory of Huygens' principle.* Vol. 329. American Mathematical Soc., 2003.

[7] V. Bakshi. *EUV Lithography.* Press Monograph Series. SPIE, 2008. ISBN: 9780819496409. URL: https://books.google.com/books?id=slQLtAEACAAJ.

[8]   E. D. Barone-Nugent, A. Barty, and K. A. Nugent. "Quantitative phase-amplitude microscopy I: optical microscopy". In: *Journal of Microscopy* 206.3 (2002), pp. 194–203. ISSN: 1365-2818.

[9]   A. Barty et al. "Quantitative optical phase microscopy". In: *Opt. Lett.* 23.11 (1998), pp. 817–819.

[10]   M. Beleggia et al. "On the transport of intensity technique for phase retrieval". In: *Ultramicroscopy* 102.1 (2004), pp. 37–49. ISSN: 0304-3991.

[11]   M. Born and E. Wolf. "The Diffraction Theory of Aberrations". In: *Ch. 9 in Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light* 9 (1989). Ed. by 6th, pp. 459–490.

[12]   John W Bossung. "Projection printing characterization". In: *Developments in Semiconductor Microlithography II*. International Society for Optics and Photonics. 1977, pp. 80–85.

[13]   E. Bostan et al. "Variational Phase Imaging Using the Transport-of-Intensity Equation". In: *IEEE Transactions on Image Processing* 25.2 (Feb. 2016), pp. 807–817. ISSN: 1057-7149. DOI: `10.1109/TIP.2015.2509249`.

[14]   Martin Burkhardt and Ananthan Raghunathan. "Best focus shift mechanism for thick masks". In: *Proc. SPIE* 9422 (2015).

[15]   Andreas Erdmann. "Mask modeling in the low k1 and ultrahigh NA regime: phase and polarization effects (Invited Paper)". In: *Proc. SPIE* 5835 (2005), pp. 69–81. DOI: `10.1117/12.637285`. URL: `http://dx.doi.org/10.1117/12.637285`.

[16]   Andreas Erdmann. "Topography effects and wave aberrations in advanced PSM technology". In: *Proc. SPIE* 4346 (2001), pp. 345–355. DOI: `10.1117/12.435734`. URL: `http://dx.doi.org/10.1117/12.435734`.

[17]   Andreas Erdmann et al. "Mask-induced best-focus-shifts in DUV and EUV lithography". In: *Proc. SPIE* 9426 (2015),

[18] A. Erdmann et al. "Mask-topography-induced phase effects and wave aberrations in optical and extreme ultraviolet lithography". In: *Journal of Vacuum Science and Technology B* 28.6 (2010), C6J1–C6J7.

[19] Tom Faure et al. In: *Photomask Technology 2010.* Ed. by M. Warren Montgomery and WilhelmEditors Maurer. SPIE-Intl Soc Optical Eng, Sept. 2010,

[20] Tom Faure et al. In: *Photomask Technology 2010.* Ed. by M. Warren Montgomery and WilhelmEditors Maurer. SPIE-Intl Soc Optical Eng, Sept. 2010,

[21] José A Ferrari et al. "Transport of intensity equation: Validity limits of the usually accepted solution". In: *Optics Communications* 318 (2014), pp. 133–136.

[22] J. Fienup. "Iterative method applied to image reconstruction and to computer-generated holograms". In: *Optical Engineering* 19.3 (1980), pp. 291–305.

[23] J. Fienup. "Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint". In: *Journal of the Optical Society of America A* 4 (1987), pp. 118–123.

[24] J. R. Fienup. "Phase retrieval algorithms: a comparison". In: *Appl. Opt.* 21.15 (1982), pp. 2758–2769.

[25] J. R. Fienup. "Phase retrieval algorithms: a comparison". In: *Appl. Opt.* 21.15 (Aug. 1982), pp. 2758–2769.

[26] Jo Finders. "The impact of Mask 3D and Resist 3D effects in optical lithography". In: *Proc. SPIE* 9052 (2014).

[27] Jo Finders. "The impact of Mask 3D and Resist 3D effects in optical lithography". In: *Proc. SPIE* 9052 (2014), DOI: 10.1117/12.2047279. URL: http://dx.doi.org/10.1117/12.2047279.

[28]   Jo Finders, Laurens de Winter, and Thorsten Last. "Mitigation of mask three-dimensional induced phase effects by absorber optimization in ArFi and extreme ultraviolet lithography". In: *Journal of Micro/Nanolithography, MEMS, and MOEMS* 15.2 (2016), p. 021408. DOI: 10.1117/1.JMM.15.2.021408. URL: http://dx.doi.org/10.1117/1.JMM.15.2.021408.

[29]   F. Fogel, I. Waldspurger, and A. d'Aspremont. "Phase retrieval for imaging problems". In: *ArXiv e-prints* 1304.7735 (Apr. 2013).

[30]   Gunjala G. et al. "Pupil function recovery using weak scattering media". In: *Opt. Exp.* (2018).

[31]   R. Gerchberg and W. Saxton. "A practical algorithm for the determination of phase from image and diffraction plane pictures". In: *Optik* 35 (1972), pp. 237–246.

[32]   R. Gerchberg and W. Saxton. "A practical algorithm for the determination of phase from image and diffraction plane pictures". In: *Optik* 35 (1972), pp. 237–246.

[33]   K. Goldberg et al. "New ways of looking at masks with the SHARP EUV microscope". In: *SPIE 9422* 4 (2015), pp. 482–489. DOI: 10.1117/12.2175553.

[34]   J. Goodman. *Introduction to Fourier optics*. Roberts & Co., 2005. ISBN: 0974707724.

[35]   E. P. Goodwin and J. C. Wyant. *Field Guide to Interferometric Optical Testing, SPIE Press, Bellingham.* 2006.

[36]   G. Gunjala et al. "Optical transfer function characterization using a weak diffuser". In: *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXIII.* Vol. 9713. Mar. 2016, p. 971315. DOI: 10.1117/12.2213271.

[37]   T. E. Gureyev, A. Roberts, and K. A. Nugent. "Partially coherent fields, the transport-of-intensity equation, and phase uniqueness". In: *J. Opt. Soc. Am. A* 12.9 (Sept. 1995), pp. 1942–1946.

[38] T.E Gureyev and K.A Nugent. "Rapid quantitative phase imaging using the transport of intensity equation". In: *Optics Communications* 133.1–6 (1997), pp. 339–346. ISSN: 0030-4018.

[39] T.E. Gureyev et al. "Linear algorithms for phase retrieval in the Fresnel region. 2. Partially coherent illumination". In: *Optics Communications* 259.2 (2006), pp. 569–580. ISSN: 0030-4018.

[40] Michael S. Hibbs and Timothy A. Brunner. *Phase calibration for attenuating phase-shift masks*. 2006.

[41] Kazuo Ishizuka and Brendan Allman. "Phase measurement of atomic resolution image using transport of intensity equation". In: *Journal of electron microscopy* 54.3 (2005), pp. 191–197.

[42] Kazuo Ishizuka and Brendan Allman. "Phase measurement of atomic resolution image using transport of intensity equation". In: *Journal of electron microscopy* 54.3 (2005), pp. 191–197.

[43] Micah H. Jenkins, Joshua M. Long, and Thomas K. Gaylord. "Multifilter phase imaging with partially coherent light". In: *Appl. Opt.* 53.16 (June 2014), pp. D29–D39.

[44] Zhong Jingshan et al. "Efficient Gaussian Inference Algorithms for Phase Imaging". In: *Proc. IEEE ICASSP* (2012), pp. 25–30.

[45] Zhong Jingshan et al. "Partially coherent phase microscopy with arbitrary illumination source shape". In: *Computational Optical Sensing and Imaging*. Optical Society of America. 2014, CTu1C–5.

[46] Zhong Jingshan et al. "Partially Coherent Phase Recovery by Kalman Filtering". In: *Frontiers in Optics 2013 Postdeadline*. Optical Society of America, 2013, FW6A.9.

[47] Zhong Jingshan et al. "Transport of Intensity phase imaging by intensity spectrum fitting of exponentially spaced defocus planes". In: *Opt. Express* 22.9 (May 2014), pp. 10661–10674. DOI: 10.1364/OE.22.010661. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-22-9-10661.

[48] Shan Shan Kou et al. "Transport-of-intensity approach to differential interference contrast (TI-DIC) microscopy for quantitative phase imaging". In: *Opt. Lett.* 35.3 (Feb. 2010), pp. 447–449. DOI: 10.1364/OL.35.000447. URL: http://ol.osa.org/abstract.cfm?URI=ol-35-3-447.

[49] Nick Loomis, Laura Waller, and George Barbastathis. "High-Speed Phase Recovery Using Chromatic Transport of Intensity Computation in Graphics Processing Units". In: *Biomedical Optics and 3-D Imaging*. Optical Society of America, 2010, JMA7.

[50] Axel Lubk et al. "Transport of Intensity Phase Retrieval of Arbitrary Wave Fields Including Vortices". In: *Phys. Rev. Lett.* 111 (17 Oct. 2013), p. 173902. DOI: 10.1103/PhysRevLett.111.173902. URL: http://link.aps.org/doi/10.1103/PhysRevLett.111.173902.

[51] Chris Mack. "Introduction to Semiconductor Lithography". In: *Fundamental Principles of Optical Lithography*. John Wiley and Sons, Ltd, 2007. Chap. 2. ISBN: 9780470723876.

[52] B. Mandelbrot. "How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension". In: *Science* 156 (May 1967), pp. 636–638. DOI: 10.1126/science.156.3775.636.

[53] Marshal Miller. "Mask Edge Effects in Optical Lithography and Chip Level Modeling Methods". PhD thesis. EECS Department, University of California, Berkeley, Dec. 2010.

[54] Marshal A. Miller and Andrew R. Neureuther. "Analysis and modeling of photomask edge effects for 3D geometries and the effect on process window". In: *Proc. SPIE* 7274 (2009), DOI: 10.1117/12.814250. URL: http://dx.doi.org/10.1117/12.814250.

[55] Patrick P. Naulleau and Gregg M. Gallatin. "Line-edge roughness transfer function and its application to determining mask effects in EUV resist characterization". In: *Appl. Opt.* 42.17 (June 2003), pp. 3390–3397. DOI: 10.1364/AO.42.003390. URL: http://ao.osa.org/abstract.cfm?URI=ao-42-17-3390.

[56] K. A. Nugent. "X-ray noninterferometric phase imaging: a unified picture". In: *J. Opt. Soc. Am. A* 24.2 (2007), pp. 536–547. URL: http://josaa.osa.org/abstract.cfm?URI=josaa-24-2-536.

[57] KA Nugent et al. "Quantitative phase imaging using hard x rays". In: *Physical review letters* 77.14 (1996), pp. 2961–2964.

[58] A. T. O'Neil et al. "Intrinsic and Extrinsic Nature of the Orbital Angular Momentum of a Light Beam". In: *Phys. Rev. Lett.* 88 (5 Jan. 2002), p. 053601. DOI: 10.1103/PhysRevLett.88.053601. URL: http://link.aps.org/doi/10.1103/PhysRevLett.88.053601.

[59] D. Paganin and K. A. Nugent. "Noninterferometric Phase Imaging with Partially Coherent Light". In: *Phys. Rev. Lett.* 80 (12 Mar. 1998), pp. 2586–2589. DOI: 10.1103/PhysRevLett.80.2586. URL: http://link.aps.org/doi/10.1103/PhysRevLett.80.2586.

[60] D. Paganin and K. A. Nugent. "Noninterferometric Phase Imaging with Partially Coherent Light". In: *Phys. Rev. Lett.* 80 (12 Mar. 1998), pp. 2586–2589.

[61] Sri Rama Prasanna Pavani et al. "Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function". In: *Proceedings of the National Academy of Sciences* 106.9 (2009), pp. 2995–2999. DOI: 10.1073/pnas.0900245106. eprint: http://www.pnas.org/content/106/9/2995.full.pdf. URL: http://www.pnas.org/content/106/9/2995.abstract.

[62] Tim C Petersen, Vicki J Keast, and David M Paganin. "Quantitative TEM-based phase retrieval of MgO nano-cubes using the transport of intensity equation". In: *Ultramicroscopy* 108.9 (2008), pp. 805–815.

[63] Jonathan C. Petruccelli, Lei Tian, and George Barbastathis. "The transport of intensity equation for optical path length recovery using partially coherent illumination". In: *Opt. Express* 21.12 (June 2013), pp. 14430–14441. DOI: 10.1364/OE.21.014430. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-21-12-14430.

[64] Jonathan Petruccelli, Lei Tian, and George Barbastathis. "The transport of intensity equation and partially coherent fields". In: *Biomedical Optics*. Optical Society of America. 2012.

[65] Tom Pistor. "EM Suite FDTD". In: *http://www.panoramictech.com/* ().

[66] J. H. Poynting. "On the Transfer of Energy in the Electromagnetic Field". In: *Philosophical Transactions of the Royal Society of London* 175 (1884), pp. 343–361. eprint: http://rstl.royalsocietypublishing.org/content/175/343.full.pdf+html.

[67] Johannes Ruoff et al. "Polarization-induced astigmatism caused by topographic masks". In: *Proc. SPIE* 6730 (2007),

[68] Jelena A. Schmalz et al. "Phase retrieval using radiation and matter-wave fields: Validity of Teague's method for solution of the transport-of-intensity equation". In: *Phys. Rev. A* 84 (2 Aug. 2011), p. 023808.

[69] Aamod Shanker, Lei Tian, and Laura Waller. "Defocus-based quantitative phase imaging by coded illumination". In: *SPIE BiOS*. International Society for Optics and Photonics. 2014, 89490R–89490R.

[70] Aamod Shanker, Laura Waller, and Andrew R Neureuther. "Defocus based phase imaging for quantifying electromagnetic edge effects in photomasks". MA thesis. University of California, Berkeley, May 2014.

[71] Aamod Shanker, Laura Waller, and Andrew R Neureuther. *Differential methods for phase imaging in optical lithography.* May 2018.

[72] Aamod Shanker et al. "Absorber topography dependence of phase edge effects". In: *Proc. SPIE* 9635 (2015), DOI: 10.1117/12.2197861. URL: http://dx.doi.org/10.1117/12.2197861.

[73] Aamod Shanker et al. "Characterizing the dependence of thick-mask edge effects on illumination angle using AIMS images". In: *Proc. SPIE* 9426 (2015).

[74] Aamod Shanker et al. "Critical assessment of the transport of intensity equation as a phase recovery technique in optical lithography". In: *Proc. SPIE* 9052 (2014),

[75] Aamod Shanker et al. "Transport of intensity phase imaging in the presence of curl effects induced by strongly absorbing photomasks". In: *Appl. Opt.* 53.34 (Dec. 2014), J1–J6. DOI: 10.1364/AO.53.0000J1. URL: http://ao.osa.org/abstract.cfm?URI=ao-53-34-J1.

[76] Aamod Shanker et al. "Transport of intensity phase imaging in the presence of curl effects induced by strongly absorbing photomasks". In: *Appl. Opt.* 53.34 (Dec. 2014), J1–J6. DOI: 10.1364/AO.53.0000J1. URL: http://ao.osa.org/abstract.cfm?URI=ao-53-34-J1.

[77] Yoav Shechtman et al. "Optimal Point Spread Function Design for 3D Imaging". In: *Phys. Rev. Lett.* 113 (13 Sept. 2014), p. 133902. DOI: 10.1103/PhysRevLett.113.133902. URL: http://link.aps.org/doi/10.1103/PhysRevLett.113.133902.

[78] Marcos Soto and Eva Acosta. "Improved phase imaging from intensity measurements in multiple planes". In: *Applied Optics* 46.33 (2007), pp. 7978–7981.

[79] Henry P. Stapp. "Quantum Theory and the Role of Mind in Nature". In: *Foundations of Physics* 31.10 (Oct. 2001), pp. 1465–1499. ISSN: 1572-9516.

[80] N. Streibl. "Phase imaging by the transport equation of intensity". In: *Optics Communications* 49 (Feb. 1984), pp. 6–10.

[81] Michael Reed Teague. "Deterministic phase retrieval: a Green's function solution". In: *J. Opt. Soc. Am.* 73.11 (Nov. 1983), pp. 1434–1441.

[82] Lei Tian, Jonathan C. Petruccelli, and George Barbastathis. "Nonlinear diffusion regularization for transport of intensity phase imaging". In: *Opt. Lett.* 37.19 (Oct. 2012), pp. 4131–4133. DOI: 10.1364/OL.37.004131. URL: http://ol.osa.org/abstract.cfm?URI=ol-37-19-4131.

[83] Lei Tian, Jingyan Wang, and Laura Waller. "3D differential phase-contrast microscopy with computational illumination using an LED array". In: *Opt. Lett.* 39.5 (Mar. 2014), pp. 1326–1329.

[84] J. Tirapu-Azpiroz, P. Burchard, and E. Yablonovitch. "Boundary layer model to account for thick mask effects in photolithography". In: *Proc. SPIE, Optical Microlithography XVI*. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series 5040 (June 2003). Ed. by A. Yen, pp. 1611–1619.

[85] V.V Volkov and Y Zhu. "Lorentz phase microscopy of magnetic materials". In: *Ultramicroscopy* 98 (2004), pp. 271–281. ISSN: 0304-3991.

[86] V.V. Volkov, Y. Zhu, and M. De Graef. "A new symmetrized solution for phase retrieval using the transport of intensity equation". In: *Micron* 33.5 (2002), pp. 411–416. ISSN: 0968-4328.

[87] L. Waller, L. Tian, and G. Barbastathis. "Transport of Intensity phase-amplitude imaging with higher order intensity derivatives". In: *Opt. Express* 18.12 (2010), pp. 12552–12561. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-12-12552.

[88] Laura Waller, Lei Tian, and George Barbastathis. "Transport of Intensity phase-amplitude imaging with higher order intensity derivatives". In: *Opt. Express* 18.12 (June 2010), pp. 12552–12561.

[89] A.K. Wong and Andrew R. Neureuther. "Mask topography effects in projection print-ing of phase-shifting masks". In: *Electron Devices, IEEE Transactions on* 41.6 (June 1994), pp. 895–902. ISSN: 0018-9383.

[90] Dongbo Xu, Peter Evanschitzky, and Andreas Erdmann. "Application of the transport of intensity equation to EUV multilayer defect analysis". In: *Proc. SPIE* 9422 (2015), DOI: 10.1117/12.2085468. URL: http://dx.doi.org/10.1117/12.2085468.

[91] Shazia Yasin et al. "Correlation of Surface Roughness with Edge Roughness in PMMA Resist". In: *Microelectron. Eng.* 78-79 (Mar. 2005), pp. 484–489. ISSN: 0167-9317. DOI: 10.1016/j.mee.2004.12.062. URL: http://dx.doi.org/10.1016/j.mee.2004.12.062.

[92] F. Zemlin et al. "Coma-free alignment of high resolution electron microscopes with the aid of optical diffractograms". In: *Ultramicroscopy* 3 (1978), pp. 49–60. ISSN: 0304-3991. DOI: https://doi.org/10.1016/S0304-3991(78)80006-0. URL: http://www.sciencedirect.com/science/article/pii/S0304399178800060.

[93] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang. "Wide-field, high-resolution Fourier ptychographic microscopy". In: *Nature Photonics* 7 (July 2013), URL: http://dx.doi.org/10.1038/nphoton.2013.187.

[94] Yunhui Zhu, Adam Pan, and George Barbastathis. "Low-noise TIE phase imaging by structured illumination". In: *Computational Optical Sensing and Imaging*. Optical Society of America. 2014, CTh3C–5.

[95] Axel M. Zibold et al. "Aerial image measurement technique for today's and future 193-nm lithography mask requirements". In: *Proc. SPIE* 5504 (2004), pp. 12–18.

[96] Chao Zuo et al. "Phase discrepancy analysis and compensation for fast Fourier trans-form based solution of the transport of intensity equation". In: *Opt. Express* 22.14 (July 2014), pp. 17172–17186.

[97]   Chao Zuo et al. "Transport-of-intensity phase imaging using Savitzky-Golay differen-
       tiation filter-theory and applications". In: *Optics express* 21.5 (2013), pp. 5346–5362.

[98]   Adam M Zysk et al. "Transport of intensity and spectrum for partially coherent fields".
       In: *Optics letters* 35.13 (2010), pp. 2239–2241.