

Adjoint based optimization to enable single mirror optic for extreme ultraviolet lithography

Greggory Scranton



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2015-131

<http://www.eecs.berkeley.edu/Pubs/TechRpts/2015/EECS-2015-131.html>

May 15, 2015

Copyright © 2015, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**Adjoint based optimization to enable single mirror optic for extreme
ultraviolet lithography**

by Gregg Scranton

Inverse Design Project

Submitted to the Department of Electrical Engineering and Computer Sciences,
University of California at Berkeley, in partial satisfaction of the requirements for the
degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Eli Yablonovitch
Research Advisor

5/14/15

* * * * *

Andrew Neureuther
Second Reader

Abstract: Traditionally, aberration correction in extreme ultraviolet (EUV) projection optics requires the use of multiple lossy mirrors, which results in prohibitively high source power requirements. We analyze a single spherical mirror projection optical system where aberration correction is built into the mask itself, through Inverse Lithography Technology (ILT). By having fewer mirrors, this would reduce the power requirements for EUV lithography. We model a single spherical mirror system with orders of magnitude more spherical aberration than would ever be tolerated in a traditional multiple mirror system. By using ILT, (implemented by an adjoint-based gradient descent optimization algorithm), we design photomasks that successfully print test patterns, in spite of these enormous aberrations. This mathematical method was tested with a 6 plane wave illumination source. Nonetheless, it would have poor power throughput from a totally incoherent source.

1. Introduction

Extreme ultraviolet (EUV) lithography is the leading contender to become the next industrial scale lithography technology in the semiconductor industry. Nonetheless, source power requirements are a major challenge that must be overcome [1]. In EUV lithography, multiple multilayer mirrors are used instead of lenses. Since the maximum reflectivity of a single mirror is 70% [1], projection optics systems employing 6 mirrors for aberration correction transmit less than 12% of the illumination power to the wafer. To address this problem, we consider a single mirror system in which the aberration correction is built in to the mask design. This could result in $(1-0.7^5)=83\%$ reduction in EUV source power required, but the mathematical procedure will constrain the source incoherence.

To design masks with built-in aberration correction, we employ the optimization approach called Inverse Lithography Technology (ILT), which was developed by Luminescent Inc. [2] and Intel [3-7], independently. This approach has the ability to explore a large design space and systematically find unintuitive, yet high-performing solutions to mask design that would not otherwise be found. We use the adjoint method, a gradient descent optimization algorithm that has great advantages over algorithms used previously for photomask design. For example, the use of gradient descent lets the algorithm converge orders of magnitude faster than non-gradient methods such as the binary search algorithm used in [8]. The adjoint method also provides more in-depth information than either the Gerchberg-Saxton algorithm in [9-11], or the back propagation technique in [12], allowing gradient descent to optimize more complex figures of merit.

In this report, we begin by describing the general form of the adjoint method, which has been used to successfully design of all manner of electromagnetic components [13-20]. We then present a specific way to apply the adjoint method to Inverse Lithography Technology. Finally, we apply this form of ILT to a single spherical mirror system with orders of magnitude greater aberrations than would ever be tolerated in a traditional multiple mirror system. The adjoint method allows us to design photomasks with non-intuitive shapes that nonetheless successfully print test patterns, in spite of these enormous aberrations.

2. The adjoint method for electromagnetic design

The adjoint method is a gradient descent optimization algorithm for designing the geometry of dielectric or metal electromagnetic devices under Maxwell's equations. Adjoint methods have been employed in the design of optical and photonic components [13-20] and mathematical derivations of the adjoint method are available in optimization textbooks [21, 22]. The adjoint method converges to an optimum much more rapidly than popular heuristic optimization methods such as genetic algorithms and particle swarm optimization, since it follows the gradient—the derivative of the Figure-of-Merit with respect to *all* geometric parameters.

The adjoint method calculates the gradient at all points in space within only 2 simulations, regardless of the size of the system. Absent the adjoint method, N simulations would be required to calculate the gradient using finite differences, where N is the number of geometrical parameters. For general geometry at all points in space, the adjoint method makes calculation of the gradient tractable when it would not be otherwise. For example, if a geometry is represented by a 1000×1000 pixel grid, and each pixel is a separate parameter, the adjoint method speeds up calculation of the gradient by $500,000\times$. A large number of parameters is desirable because this provides more degrees of freedom to the optimizer, and hence makes a better optimum achievable.

In our implementation, the adjoint inverse solver is a small subroutine that wraps around a forward solver. This means any existing commercial Maxwell forward solver can be used.

A flowchart describing the adjoint inverse solver is shown in Fig. 1. In a given iteration, the forward simulation provides the electromagnetic fields for the current geometry. Then the adjoint simulation calculates the gradient. In gradient descent, a local change in geometry is made, proportional to the calculated gradient, in preparation for the next iteration.

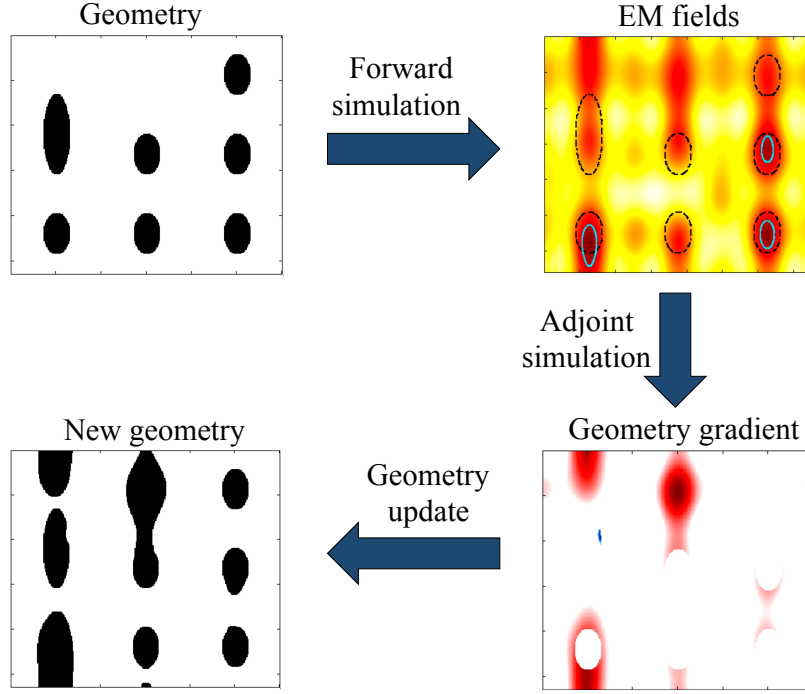


Fig. 1. A flowchart showing one iteration in the adjoint method. First, electric and/or magnetic fields are found for the current geometry through the forward simulation. Then, the geometry gradient is found through the adjoint simulation. The gradient is used to make an update to the geometry.

3. The adjoint method applied to ILT

This section describes our mathematical approach for applying the adjoint method to ILT for photomasks. We have adopted the mathematical formulation of the adjoint method previously presented in [18] and [29]. These references contain a more general form of the present derivation that accounts for vector forms of both electric and magnetic fields. Only scalar electric fields are considered here.

A reflective projection optics system with one mirror is depicted in Fig. 2(a). Equivalently, we model the system with a refractive lens as shown in Fig. 2(b). We will find the gradient of the Figure-of-Merit (the total image error) with respect to the mask transmission factor (which defines where the mask is opaque or transmissive). The mask transmission factor is in the mask plane, while the Figure-of-Merit is a function of the electric field in the wafer plane. To find the gradient of the Figure-of-Merit, with respect to the mask transmission factor, we apply the chain rule of calculus: First find the gradient with respect to the mask plane electric field, and then the derivative of electric field with respect to the transmission factor.

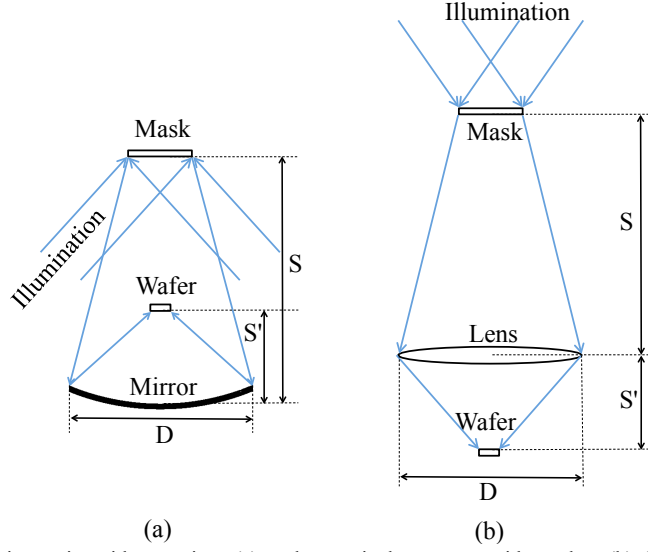


Fig. 2. Projection optics with one mirror (a), and an equivalent system with one lens (b). S is the distance from the mirror to the mask, and S' is the distance from the mirror to the wafer (not to scale). D is the diameter of the mirror/lens.

3.1. Gradient with Respect to Electric Field

The Figure-of-Merit is a sum of errors in the wafer plane image, and has the general form

$$FoM = \int_W f(E_W(\mathbf{r}_W)) d^2 r_W \quad (1)$$

where f represents a local error in the image at point \mathbf{r}_W , the subscript W denotes a variable in the wafer plane, E_W is the wafer plane electric field, \mathbf{r}_W is the two dimensional spatial position vector in the wafer plane, and bold face denotes a vector quantity. The local Figure-of-Merit f , is a step-like function of the local electric field E_W , which might be larger or smaller than a desired target electric field. Differentiating Eq. (1), with respect to the mask plane electric field E_M , we obtain

$$\frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} = \int_W \frac{\partial f}{\partial E_W(\mathbf{r}_W)} \frac{\partial E_W(\mathbf{r}_W)}{\partial E_M(\mathbf{r}_M)} d^2 r_W, \quad (2)$$

where the subscript M denotes a variable defined in the mask plane. During optimization, we adjust the mask to vary E_M to achieve the best possible image. To determine the partial derivative $\partial E_W(\mathbf{r}_W) / \partial E_M(\mathbf{r}_M)$, we must first express the wafer plane field in terms of the mask plane field.

$$E_W(\mathbf{r}_W) = \int_M E_M(\mathbf{r}_M) PSF_{M \rightarrow W}(\mathbf{r}_W - \mathbf{r}_M) d^2 r_M \quad (3)$$

Equation (3) is a convolution integral with the point spread function for propagation from the mask to the wafer plane, $PSF_{M \rightarrow W}$, which would generally require a solution of Maxwell's equations, but we use the paraxial and other approximations to determine $PSF_{M \rightarrow W}$. Substituting Eq. (3) into Eq. (2), we obtain

$$\frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} = \int_W \frac{\partial f}{\partial E_W(\mathbf{r}_W)} \frac{\partial}{\partial E_M(\mathbf{r}_M)} \left[\int_M E_M(\mathbf{r}'_M) PSF_{M \rightarrow W}(\mathbf{r}_W - \mathbf{r}'_M) d^2 r'_M \right] d^2 r_W, \quad (4)$$

where \mathbf{r}'_M is a dummy variable for convolution.

We are interested in the derivative of the term in square brackets with respect to the variable E_M at one particular position \mathbf{r}_M . Since $E_M(\mathbf{r}_M)$ and $E_M(\mathbf{r}'_M)$ are independently controlled variables, the derivative with respect to $E_M(\mathbf{r}_M)$ produces a delta function $\delta(\mathbf{r}'_M - \mathbf{r}_M)$ and Eq. (4) becomes

$$\frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} = \int_W \frac{\partial f}{\partial E_W(\mathbf{r}_W)} [PSF_{M \rightarrow W}(\mathbf{r}_W - \mathbf{r}_M)] d^2 r_W \quad (5)$$

Equation (5) nearly looks like a convolution integral, but $PSF_{M \rightarrow W}$ is an operator that only operates on functions defined in the mask plane, and $\partial f / \partial E_W(\mathbf{r}_W)$ is in the wafer plane. This can be resolved by the reciprocity of Maxwell's equations, which dictates the reciprocal relation:

$$PSF_{M \rightarrow W}(\mathbf{r}_W - \mathbf{r}_M) = PSF_{W \rightarrow M}(\mathbf{r}_M - \mathbf{r}_W) \quad (6)$$

where $PSF_{W \rightarrow M}$ is the point spread function for propagation from the wafer plane back to the mask plane. Plugging Eq. (6) into Eq. (5), we obtain

$$\frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} = \int_W \frac{\partial f}{\partial E_W(\mathbf{r}_W)} PSF_{W \rightarrow M}(\mathbf{r}_M - \mathbf{r}_W) d^2 r_W \quad (7)$$

Equation (7) is indeed a convolution integral, and it is the important result that we have been seeking. It states that $\partial f / \partial E_W(\mathbf{r}_W)$ can be treated as a source electric field and propagated from the wafer plane to the mask plane to obtain $\partial FoM / \partial E_M(\mathbf{r}_M)$. This is the adjoint simulation step shown on the right side of Fig. 1.

For a spatially incoherent system that is modeled as a sum of coherent systems, the above procedure must be executed for every angle of illumination. Gradients of the FoM with respect to the electric field, properly weighted over the angles of illumination must be considered.

3.2. Gradient with respect to mask transmission factor

In the geometry update, changes in mask geometry must be derived from changes in mask plane electric field. In our simple model, each pixel in the mask is either perfectly opaque or perfectly transmitting. Thus, the mask is represented by a transmission factor, T_M , which has values of either 0 or 1, and multiplies the incoming field. To include mask edge effects, a more complete electromagnetic model would be required. A method for including electromagnetic effects in the optimization is described in [18] and [29]. To update the mask geometry, represented by a transmission factor T_M , the Figure-of-Merit derivative with respect to local electric field, $\partial FoM / \partial E_M$ must be related to the derivative with respect to transmission factor $\partial FoM / \partial T_M$. The mask plane field is related to the mask transmission factor by

$$E_M(\mathbf{r}_M) = T_M(\mathbf{r}_M) E_o \exp[i\phi_{EM}(\mathbf{r}_M)] \quad (8)$$

Where E_o is the normalized incident electric field magnitude, ϕ_{EM} is the corresponding phase. $\partial FoM / \partial T_M$ can be found by using the *chain rule* on $\partial FoM / \partial E_M$. Care must be taken because E_M is generally complex. One could take derivatives with respect to the real and imaginary parts of E_M . An equivalent and more convenient method is to take derivatives with respect to E_M and its complex conjugate as follows

$$\frac{\partial FoM}{\partial T_M(\mathbf{r}_M)} = \frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} \frac{\partial E_M(\mathbf{r}_M)}{\partial T_M(\mathbf{r}_M)} + \frac{\partial FoM}{\partial E_M^*(\mathbf{r}_M)} \frac{\partial E_M^*(\mathbf{r}_M)}{\partial T_M(\mathbf{r}_M)} \quad (9)$$

where the asterisk * denotes complex conjugation. Since FoM and T_M must be real, the two terms on the right hand side of Eq. (9) are complex conjugates of each other. Thus, their imaginary parts cancel out, resulting in

$$\frac{\partial FoM}{\partial T_M(\mathbf{r}_M)} = 2 \text{Re} \left[\frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} \frac{\partial E_M(\mathbf{r}_M)}{\partial T_M(\mathbf{r}_M)} \right] \quad (10)$$

Plugging Eq. (8) into Eq. (10), we obtain

$$\frac{\partial FoM}{\partial T_M(\mathbf{r}_M)} = 2 \operatorname{Re} \left\{ \frac{\partial FoM}{\partial E_M(\mathbf{r}_M)} E_o \exp[i\phi_{EM}(\mathbf{r}_M)] \right\} \quad (11)$$

which translates from electric field gradient to the more operational mask transmission factor gradient.

For a spatially incoherent system modeled as a sum of coherent systems, $\partial FoM / \partial T_M$ can be expressed as the total derivative with respect to the electric field of equally weighted angles of illumination.

$$\frac{\partial FoM}{\partial T_M(\mathbf{r}_M)} = \sum_n 2 \operatorname{Re} \left\{ \frac{\partial FoM}{\partial E_{Mn}(\mathbf{r}_M)} E_{on} \exp[i\phi_{EMn}(\mathbf{r}_M)] \right\} \quad (12)$$

where E_{Mn} is the mask plane electric field for the angle of illumination indexed by the integer n .

$\partial FoM / \partial T_M$ is the gradient with respect to the operational mask design parameters, and provides information about how the Figure-of-Merit changes as the transmission factor T_M changes at each point in space. Gradient descent, as in Newton's method for solving polynomial equations, operates by changing the mask transmission proportional to the rate of increase in the Figure-of-Merit:

$$\Delta T_M \propto \frac{\partial FoM}{\partial T_M} \quad (13)$$

where ΔT_M is the change in T_M at a given iteration. As an optimum is approached, and the derivative approaches zero, the changes in T_M become smaller and smaller.

To model a binary amplitude mask, such as those used in EUV lithography, we constrain T_M to only take values of 0 or 1. Since the mask transmission is binary and does not take continuous values, the geometry update differs slightly from conventional gradient descent. Pixels in the mask are flipped only with the correct sign of $\partial FoM / \partial T_M$, and only when the gradient magnitude exceeds a threshold. The threshold is adjusted several times within each iteration to find the best improvement in the Figure-of-Merit. In this way, the iterative optimization procedure is well defined.

3.3. Figure-of-Merit

The Figure-of-Merit that we have preferred in these optimizations is the total error region area in which the printed pattern differs from the desired pattern. That area must be minimized. This Figure-of-Merit is illustrated by the grey region in Fig. 3. Thus, the error region is defined as

$$FoM = \int_w |P_d(\mathbf{r}_w) - P_a(E_w(\mathbf{r}_w))| d^2 r_w \quad (14)$$

where P_d and P_a are binary functions defining the desired and actual printed patterns, respectively. These are defined as

$$P_d(\mathbf{r}_w) = \begin{cases} 0 & \mathbf{r}_w \in \text{desired dark region} \\ 1 & \mathbf{r}_w \in \text{desired bright region} \end{cases} \quad P_a(E_w(\mathbf{r}_w)) = \begin{cases} 0 & |E_w(\mathbf{r}_w)|^2 < I_{th} \\ 1 & |E_w(\mathbf{r}_w)|^2 \geq I_{th} \end{cases} \quad (15)$$

where I_{th} is the exposure threshold for electric field intensity. Anywhere the intensity is greater than I_{th} , P_a is set to 1. Otherwise, its value is 0. Figure 3 shows an example of how the Figure-of-Merit is calculated.

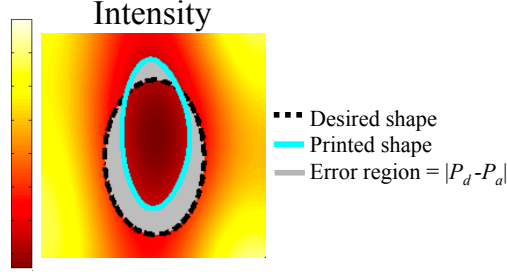


Fig. 3. An example Figure-of-Merit calculation at the wafer plane. The color map shows electric field intensity. The desired pattern, P_d , is outlined by the black dashed line. The actual printed pattern, P_a , is outlined in cyan. The “error region”, $|P_d - P_a|$, is shown in gray. This error region is integrated to obtain the Figure-of-Merit.

The integrand of Eq. (14), $|P_d(\mathbf{r}_w) - P_a(E_w(\mathbf{r}_w))| \equiv f$, must be differentiated to obtain the wafer plane gradient $\partial f / \partial E_w$. Unfortunately, P_a is not differentiable. Therefore, it is replaced by the continuous logistic function.

$$P_a(E_w(\mathbf{r}_w)) \approx \frac{1}{1 + \exp[-A(|E_w(\mathbf{r}_w)|^2 - I_{th})]} \equiv P'_a(E_w(\mathbf{r}_w)) \quad (16)$$

where A is a parameter defining the slope of the continuous differentiable function P'_a . To differentiate f , we replace the absolute magnitude with the square root of its square.

$$f = |P_d - P'_a| = \left[(P_d - P'_a)^2 \right]^{\frac{1}{2}} \quad (17)$$

Differentiating, we obtain

$$\frac{\partial f}{\partial E_w} = \frac{1}{2} \left[(P_d - P'_a)^2 \right]^{\frac{1}{2}} 2(P_d - P'_a) \left(-\frac{\partial P'_a}{\partial E_w} \right) \quad (18)$$

$$= \frac{P'_a - P_d}{f} \left(\frac{\partial P'_a}{\partial E_w} \right) \quad (19)$$

We then differentiate P'_a ,

$$\frac{\partial P'_a}{\partial E_w} = -\frac{1}{\left\{ 1 + \exp[-A(|E_w|^2 - I_{th})] \right\}^2} \frac{\partial}{\partial E_w} \left\{ \exp[-A(|E_w|^2 - I_{th})] \right\} \quad (20)$$

$$= -P_a'^2 \frac{\partial}{\partial E_w} \left\{ \exp[-A(|E_w|^2 - I_{th})] \right\} \quad (21)$$

$$= AP_a'^2 E_w^* \exp[-A(|E_w|^2 - I_{th})] \quad (22)$$

where we used $\partial |E_w|^2 / \partial E_w = E_w^*$ [23]. Plugging Eq. (22) into Eq. (19), we obtain the wafer plane gradient

$$\frac{\partial f}{\partial E_w} = \frac{P'_a - P_d}{f} AP_a'^2 E_w^* \exp[-A(|E_w|^2 - I_{th})] \quad (23)$$

We insert Eq.s. (23)&(7) into Eq. (11), and then proceed with Eq. (13) to complete one iteration.

For a spatially incoherent system modeled as a sum of coherent systems, $|E_W|^2$ in Eq. (15) is replaced by $\sum_n |E_{Wn}|^2$, where E_{Wn} is the wafer plane electric field resulting from one angle of illumination, indexed by the integer n . Differentiation with respect to E_{Wn} proceeds similarly to Eq. (16)-(23).

4. Results

To test the method outlined in the previous sections, we consider a single lens lithography system as shown in Fig. 2(b) that incorporates the aberrations to be expected in an equivalent single mirror EUV system as shown in Fig. 2(a). The magnification is 0.25, as is the convention in photolithography. The lens/mirror diameter is $D=30\text{cm}$, with a numerical aperture at the wafer plane, $NA_W=0.33$. This leads to a mirror surface-to-wafer distance $D/2\tan(\sin^{-1}NA_W)=S'=42.9082\text{ cm}$. The mirror focusing equation, $2/R=1/S'+1/S$ leads to a mirror surface-to-mask distance $S=171.6328\text{ cm}$, and a mirror radius of curvature $R=68.6532\text{cm}$. For these dimensions, a spherical mirror, relative to an ideal parabolic mirror, has aberrations amounting to $>10000\lambda$ for $\lambda=13.5\text{nm}$. We assign 6 significant figures to the mirror radius of curvature owing to the need to specify the mirror surface within $\sim 0.1\lambda$, as is common in high precision optics. Indeed we have found that even $\sim 0.1\lambda$ phase shift at the edge of the mirror produces $\sim 10\%$ errors in the test pattern features, unless the mask is redesigned to account for the newly shifted mirror surface.

Six discrete plane waves are used for illumination. These points were chosen to give the illumination some of the characteristics of an extended dipole source. The illumination pattern used is shown in Fig. 4. Our ILT mask solutions do correct aberrations very well, within the diffraction limit of the six selected illumination angles, but our solutions fail to accommodate the broad power from an extended incoherent source. We can model the incoherent source with more plane waves, but within the diffraction limit the number of plane waves would eventually equal the number of pixels. Each incident illumination angle imposes an additional constraint. For a totally incoherent source, the computation would not be manageable, nor would there be enough pixels in the mask to satisfy the multi-faceted constraints. This is discussed further in section 5. Thus, ILT for aberration correction is most suited to a partially coherent illumination source, like a laser.

Calculations of the basic Eq. (3) are executed in MATLAB, using fast Fourier transforms to compute the convolution with the point spread function. More mathematical details are included in Appendix A.

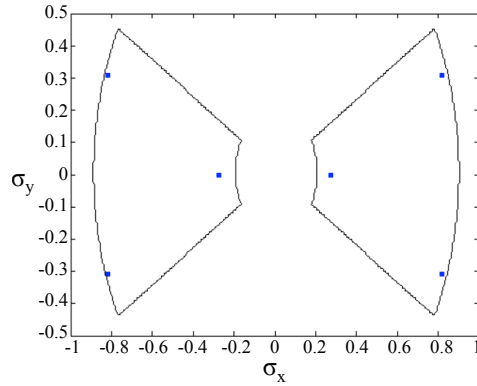


Fig. 4. Illumination pattern. $\sigma_x = \sin\theta_x/NA_W$ and $\sigma_y = \sin\theta_y/NA_W$. These six plane waves were chosen to give the illumination some of the characteristics of an extended dipole source, such as the one outlined in black. The four σ_x values are -0.8182, -0.2727, 0.2727, and 0.8182. The three σ_y values are -0.3099, 0, and 0.3099.

4.1. Correcting severe spherical aberration.

We use a test pattern from an industry presentation [24], which is shown by the dashed lines in Fig. 5(b). The pattern consists of six $14\text{nm} \times 22\text{nm}$ features and one $14\text{nm} \times 44\text{nm}$ feature. The features are placed 50nm apart in the x-direction, and 22nm apart in the y-direction. These dimensions should be compared

with a diffraction limit $\lambda/(4NA_H)=10\text{nm}$ for an EUV wavelength $\lambda=13.5\text{nm}$. The features are ellipsoidal to avoid sharp corners below the diffraction limit. The pattern in Fig. 5(a) & 6(a), is one unit cell of a periodic naïve mask, identical to the desired test pattern. The exposure threshold is taken to be half the clear field intensity.

For an un-aberrated case, the resulting wafer plane intensity and printed pattern are shown in Fig. 5(b). For the spherically aberrated case, (corresponding to a 30cm diameter focusing mirror) the wafer plane intensity and printed pattern are shown in Fig. 6(b). In the un-aberrated case, Fig. 5, all the features print. In Fig. 6, the high spherical aberration produces 4 missing features and 3 unacceptable features. This spherical aberration relative to a perfect parabolic reflector has a peak value of $>10000\lambda$ ($>140\mu\text{m}$) based on Eq. (33) in Appendix A. The wafer is readjusted to the plane of best focus for this level of spherical aberration, $\sim 2.6\text{mm}$ closer to the mirror than the best focus in the un-aberrated case.

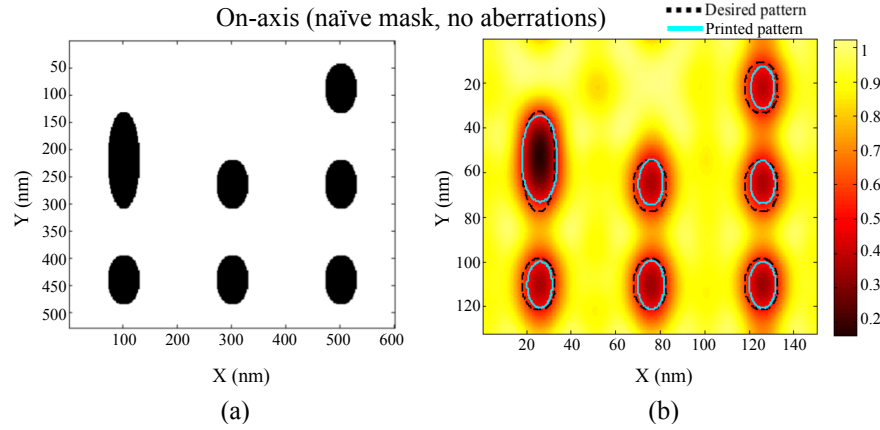


Fig. 5. (a) mask and (b) wafer plane intensity (normalized to clear field) for an optical system with a naïve mask and no aberrations in the on-axis position. The pattern is periodic, with one unit cell shown. The NA of the system is 0.33, the demagnification is 4, and the wavelength is 13.5 nm.

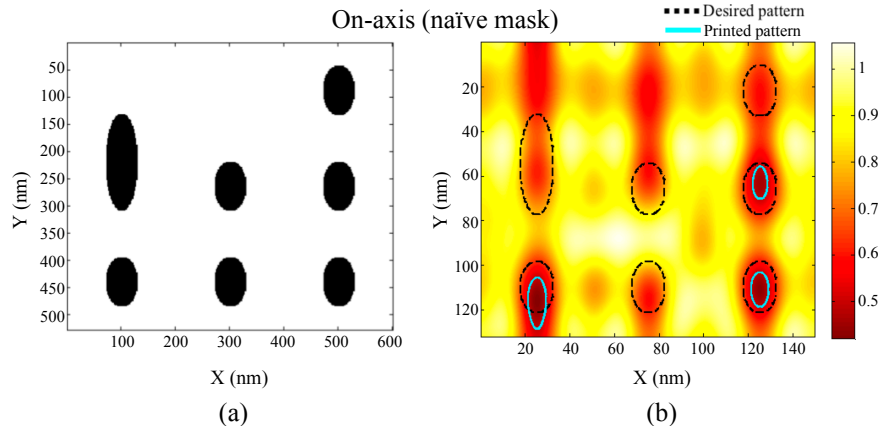


Fig. 6. (a) mask and (b) wafer plane intensity (normalized to clear field) for an optical system with a naïve mask in the on-axis position. The pattern is periodic, with one unit cell shown. The NA of the system is 0.33, the demagnification is 4, and the wavelength is 13.5 nm. The mirror radius is 15 cm. The image was taken at the center of the field. This naïve mask is used as the starting geometry for the optimization.

Using the test pattern in Fig. 6(a) as a starting point, we used our adjoint optimizer, Eq. (13), to optimize the mask. The pixel size for the simulation during optimization is 0.25 nm at the wafer plane and 1nm at the mask plane. After optimization, the mask solution is tested with a smaller pixel size=0.16nm at the mask plane, for validation. This change in pixel size is done to ensure the critical dimensions of the final shape are computed to within less than 1% accuracy during final analysis. This accuracy is not critically needed during optimization, but is important in validation. The optimized mask appearance is shown in Fig. 7(a). In the intensity profile of Fig. 7(b) all critical dimensions were achieved to within 5%. The optimization took 148 iterations to converge. Roughly 3 pixel-flip thresholds were compared per iteration, as discussed at the end of section 3.2. Additionally, a radius of curvature constraint $<12\text{nm}$ was imposed on the mask at each iteration. This radius of curvature constraint is applied by morphological opening as described in [25].

A coarser, pixelated, version of the optimized mask in Fig. 7 is shown in Fig. 8, with the mask constructed from $14\text{nm} \times 15\text{nm}$ rectangles to make the mask conform to Manhattan geometry. After pixelation, all critical dimensions are still achieved to within 8%. This demonstrates the robustness of the mask design.

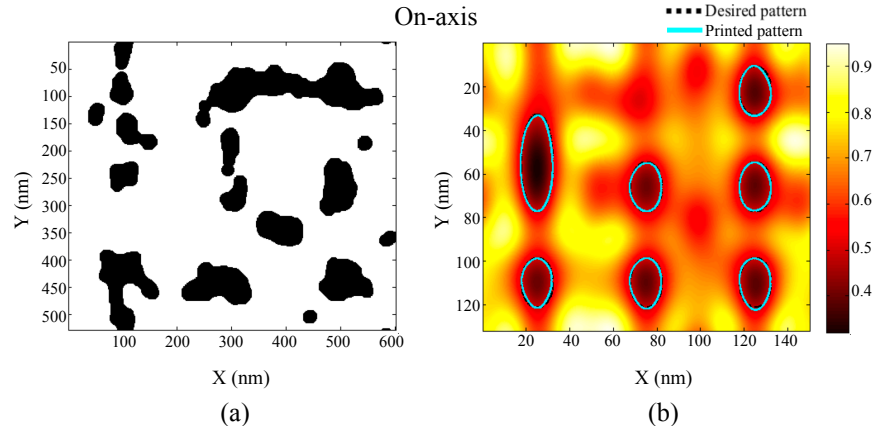


Fig. 7. (a) Mask and (b) wafer plane intensity (normalized to clear field) for an optical system with an optimized mask. The simulation conditions are the same as in Fig. 6. With this optimized mask, all critical dimensions are within 5% of their target.

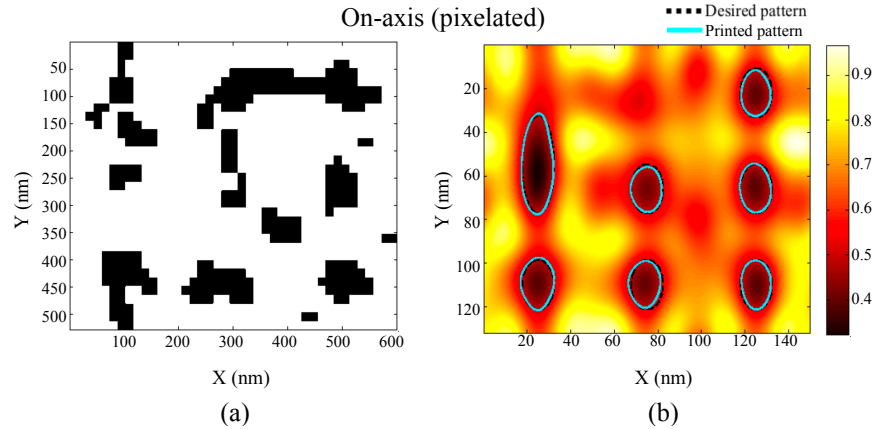


Fig. 8. The same optical system as in Fig. 7, with the mask pixelated. The pixels are $14\text{nm} \times 15\text{nm}$. All critical dimensions are within 8% of their target.

4.2. Depth-of-focus optimization

We have also performed mask optimization as a function of focal depth. To do this, we began with the optimal mask at focus, and then optimized the Figure-of-Merit at 4 planes: -50nm, -30nm, -10nm, and +10nm relative to the initial optimal plane, to investigate a 60 nm depth of focus. The Figure-of-Merit is the sum of area errors, as in Eq. (14), summed over all 4 image planes. Figure 9 shows the mask and wafer field at nominal focus resulting from this optimization. Roughly 5 pixel-flip thresholds were compared per iteration, as discussed at the end of section 3.2. This optimization required 339 iterations.

Figure 10(a) shows a Bossung plot for the worst performing feature for the mask optimized at focus, and Fig. 10(b) for the mask simultaneously optimized at the four different planes -50nm to +10nm. The sharp jumps seen in the plots correspond to changes in the location of the worst performing feature. For the mask optimized through focus, all critical dimensions remain within 11% for the full 60nm of defocus at nominal dose. The dose sensitivity, not optimized here, could be expected to improve if optimized.

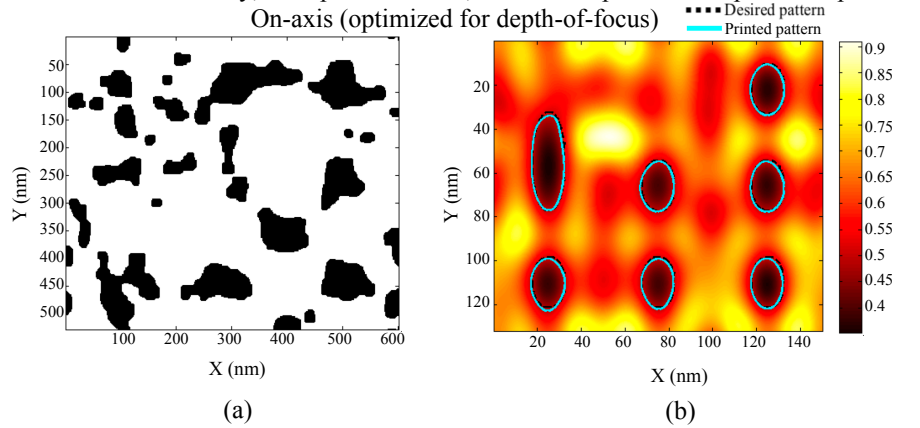


Fig. 9. (a) Mask and (b) wafer plane intensity (normalized to clear field) for an optical system with a mask optimized to perform through 60nm of defocus. The simulation conditions are the same as in Fig. 5. With this mask, all critical dimensions are within 7% of their target at focus, and remain within 11% through 60nm of defocus.

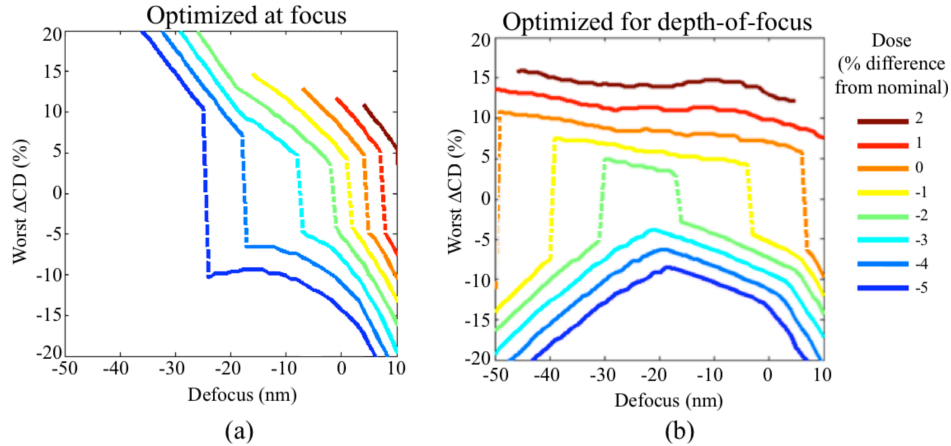


Fig. 10. Bossung plots for the worst performing feature for the masks optimized (a) at focus, and (b) for depth-of-focus. For the mask optimized through focus, all critical dimensions remain within 11% of their targets for 60nm of defocus at nominal dose. The sharp jumps seen in the plots correspond to changes in the location of the worst performing feature.

4.3. Off-axis aberration correction

We have considered severe spherical aberrations, combined with a depth of field requirement. Now we consider off-axis imaging, which includes, spherical aberration, coma, and astigmatism, all severe owing to the use of an uncorrected spherical optic. The additional aberrations change the point spread function used in simulation according to Eq. (27), (28), and (33) in Appendix A. The off-axis points are in a $33\text{mm} \times 26\text{mm}$ wafer and are shown in Fig. 11. The mid-field point 6.5mm off-center is $\sim 0.9^\circ$ off-axis, and the field edge-point 13mm off-center is $\sim 1.8^\circ$ off-axis. The mid-field point experiences $>4000\lambda$ of coma and $>230\lambda$ of astigmatism in addition to the spherical aberration present in the on-axis case. The field edge point experiences $>9000\lambda$ of coma, and $>900\lambda$ of astigmatism. For simplicity we don't account for aberration variation within the $150\text{nm} \times 132\text{nm}$ test pattern unit cell. This variation is relatively small, but must be taken into account in an industrial application.

The results for the mid-field optimization are shown in Fig. 12. The optimized mask for the on-axis case, Fig. 7(a) was used as the starting point for this optimization. After the final iteration, the critical dimensions are within 2% of the desired target. During optimization, the simulation pixel size was 2nm at the mask plane. As before, the pixel size was reduced to 0.16nm at the mask plane to accurately validate the critical dimensions after the last iteration. This optimization took 313 iterations with roughly 3 pixel-flip thresholds tested per iteration, as discussed at the end of section 3.2.

The results for the field edge case are shown in Fig. 13. The optimized mask for the mid-field case was used as the starting mask for this optimization. After the final iteration, all critical dimensions are within 3% of their desired target. During optimization, the simulation pixel size was 2nm at the mask plane. For validation after the last iteration, the pixel size was reduced to 0.16nm at the mask plane. This optimization took 185 iterations with roughly 3 pixel-flip thresholds tested per iteration, as discussed at the end of section 3.2.

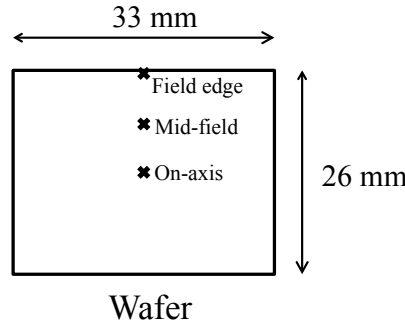


Fig. 11. A diagram showing the three points on the wafer we designed masks for. The wafer was assumed to be 33 by 26 mm. The mid-field point is displaced 6.5 mm from the optical axis and has >4000 wavelengths of coma and >230 wavelengths of astigmatism (peak value, using the convention in Eq. (33) in Appendix A). The field edge point is displaced 1.3 mm and has >9000 wavelengths of coma, and >900 wavelengths of astigmatism.

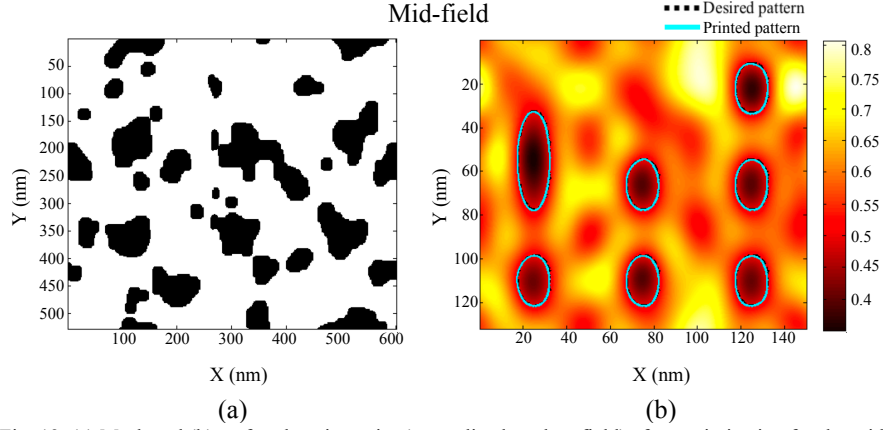


Fig. 12. (a) Mask and (b) wafer plane intensity (normalized to clear field) after optimization for the mid-field location 6.5mm off-axis. The mask resulting from the on-axis optimization was used as the starting mask for this optimization. All critical dimensions are within 2% of their target.

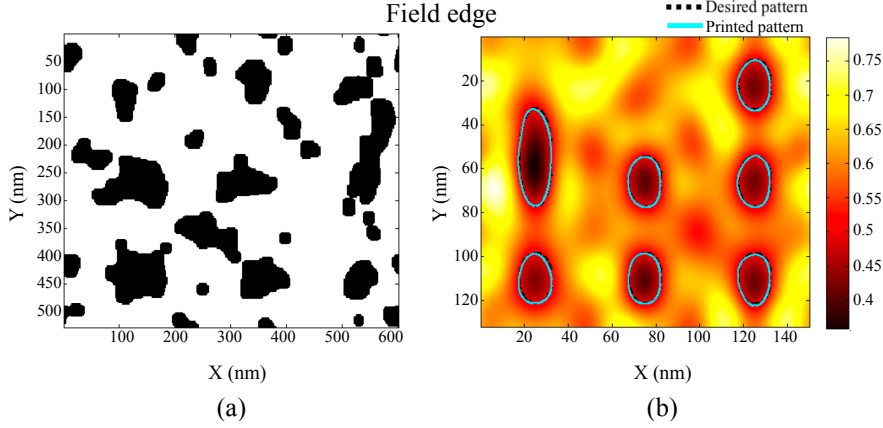


Fig. 13. (a) Mask and (b) wafer plane intensity (normalized to clear field) after optimization for the field edge location 13mm off-axis. The mask resulting from the mid-field optimization was used as the starting mask for this optimization. All critical dimensions are within 3% of their target.

A comparison of Fig. 7(a), on-axis, Fig. 12(a), 6.5mm off-axis, and Fig. 13(a), 13mm off-axis, show completely different mask solutions, even though the test pattern was identical. The mask solution is sensitive to the exact level of aberrations. A mask solution at the center of a chip would be different from a mask solution at the edge of a chip. Even with a repeating pattern as in a DRAM chip, the mask would be aperiodic, and computationally intensive to design.

We have considered spherical aberration, and off-axis aberrations coma and astigmatism. Additional aberrations can be trivially included in the current model by adding more terms to the phase shift at the mirror; therefore changing the point spread function. Additionally, in an industrial application, the Figure-of-Merit should include tolerances toward exposure dose and errors in photomask fabrication. Since off-axis aberrations vary across the chip-field, a global optimization across the whole chip would be required.

In addition, electromagnetic edge effects in the mask, and angle-dependent mirror reflectivity, must also be accounted for in the simulation. This does not pose problems to the optimization method, since our implementation of the adjoint method can wrap around any Maxwell solver.

5. Angular sensitivity analysis

Under severe aberrations, the printed pattern on the wafer for a coherent source becomes sensitive to small changes in the angle of illumination. This has implications for an incoherent source with a continuous spread of illumination angles. Since the incoherent source is represented as the sum of coherent sources, the angular precision of a coherent source determines how many coherent plane waves will need to be simulated to accurately represent the incoherent source. If this number is sufficiently large, there will not be enough design degrees of freedom to successfully design the mask.

The reason for the angular sensitivity from aberrations is illustrated in Figure 14. Figure 14 depicts a coherent system with one lens for the purposes of illustration. When the angle is changed by $\Delta\theta$, the position of a diffracted order on the lens shifts by Δr_L , where r_L is the lateral spatial coordinate on the lens in the direction of the angle change. This changes the aberration phase shift seen by the diffracted order by $\Delta\delta$. For different diffracted orders, $\Delta\delta$ will be different. This results in a different complex field on the lens, and hence a different field on the wafer.

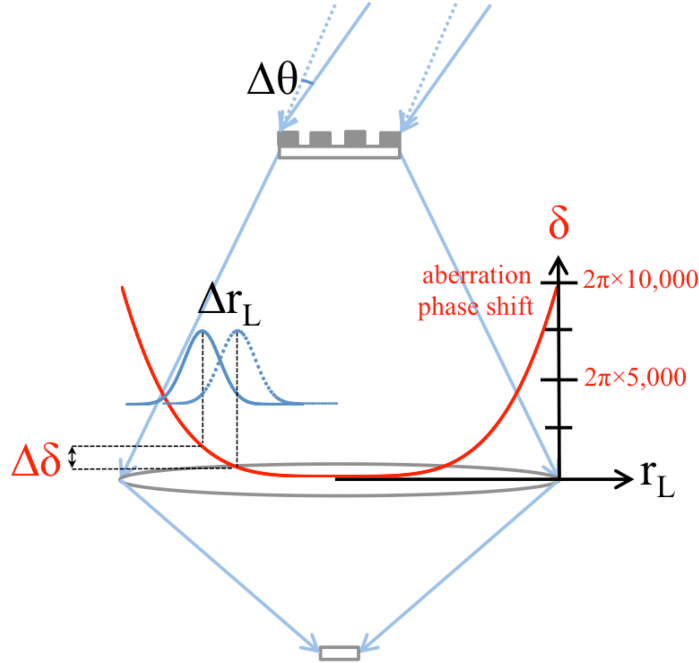


Fig. 14. A diagram showing how the angular sensitivity, $\Delta\theta$ of an illumination plane wave depends upon the magnitude of aberration phase shift on the lens of an optical system.

To calculate the tolerable $\Delta\theta$, we will assume a tolerable $\Delta\delta$ at the edge of the lens of 0.1λ . This is the $\Delta\delta$ that is likely to cause $\sim 10\%$ change in critical dimension in the wafer image. We note the relationship between $\Delta\theta$ and $\Delta\delta$

$$\Delta\delta = \frac{\partial\delta}{\partial r_L} \Delta r_L = \frac{\partial\delta}{\partial r_L} S \Delta\theta, \quad (24)$$

where S is the distance between the mask and the lens. From here, we note that if there is significant overlap between a diffracted order before and after an angle shift (the solid and dashed blue lines above the lens in Fig 14), the total electric field is affected less by the angle shift. To accommodate this, we make a modification to Eq. (24) if the angle shift $\Delta\theta$ is less than $\theta_w = \lambda/\Lambda$, the angular width of one diffracted order (where Λ is the width of the mask).

$$\Delta\delta_{eff} = \begin{cases} \frac{\partial\delta}{\partial r_L} S \Delta\theta & \Delta\theta > \theta_w \\ \frac{\partial\delta}{\partial r_L} S \Delta\theta \frac{\Delta\theta}{\theta_w} & \Delta\theta < \theta_w \end{cases}, \quad (25)$$

The second case in Eq. (25) multiplies Eq. (24) by the fraction $\Delta\theta/\theta_w$ to give an estimate of the effective $\Delta\delta$ from $\Delta\theta$.

We will assume spherical aberration, for which $\delta(r_L)$ has the form $\delta(r_L)=ar_L^4$. Plugging in the numbers $S=1.71\text{m}$, $\Delta\delta_{eff}=0.1\lambda$, $\lambda=13.5\text{nm}$, $\delta(r_{L,max})=10,000\lambda$, $r_{L,max}=15\text{cm}$, and $\Lambda=104\text{mm}$ into Eq. (25); we obtain $\Delta\theta=10^{-5}\text{°}$. With an illumination cone with a maximum angle of 5° , this means we would need to optimize for $\pi(5\text{°}/10^{-5}\text{°})=8\times 10^{11}$ separate coherent sources. This is much larger than the number of degrees of freedom in the mask ($\sim 3\times 10^5$). This implies that our method cannot be used with a coherent source and a large mask. Nonetheless, it would work well with a partially coherent source such as a laser. Thus, such a source for EUV warrants more scientific effort.

6. Conclusion

We have shown that, under a partially coherent source, like a laser, Inverse Lithography Technology can allow EUV Lithography to proceed in spite of severe aberrations, (as would be produced by a single-mirror imaging system). By reducing from 6 mirrors to 1 mirror, the power wasted by the projection optics would be reduced by $\sim 7\times$, owing to the diminished mirror losses with fewer mirrors. Since ILT is needed for mask optimization, the strategy of also using it for aberration correction seems well warranted. We have successfully designed photomasks to print test patterns in the presence of severe spherical aberration and including off-axis coma and astigmatism, and the requirement for 60nm depth of focus.

If we force current incoherent EUV sources to produce a six beam illumination pattern as in Fig. 4, the throughput would be very limited. Thus a partially coherent EUV source, like a laser should warrant more scientific and technological effort.

7. Appendix A: Linear system model of the lithographic imaging system

This section describes the linear system model based on the paraxial approximation used in our simulations. This model can be found in textbooks [1, 26, 27].

To simulate the projection optics, contributions to wafer intensity from different angles of illumination are considered. The electric field transmitted from the mask from one illuminating plane wave is

$$E_M(\mathbf{r}_M) = T_M(\mathbf{r}_M) E_o \exp\left[ik(x_M \sin\theta_x + y_M \sin\theta_y)\right] \quad (26)$$

where x_M and y_M are the vector components of \mathbf{r}_M . $k=2\pi/\lambda$ is the wave number of the light, θ_x and θ_y are each angles between the \mathbf{k} vector and a plane normal to the corresponding axis. T_M is the mask transmission factor, which is equal to 0 in absorbing regions, and 1 in transmitting regions.

To propagate this field to the wafer plane, it is convolved with the point spread function of the projection optics. The optical transfer function is the Fourier transform of the point spread function and is defined as

$$FT[PSF_{M \rightarrow W}(\mathbf{r}_M)] = OTF(\rho, \phi) = \begin{cases} \exp[-ik OPD(\rho, \phi)] & \rho \leq 1 \\ 0 & \rho > 1 \end{cases} \quad (27)$$

where

$$\rho = \sqrt{\left(\frac{f_x \lambda}{NA_w}\right)^2 + \left(\frac{f_y \lambda}{NA_w}\right)^2} \quad (28)$$

is the normalized radial coordinate, ϕ is the azimuthal angle, f_x and f_y are the spatial frequencies in the x and y directions, respectively, NA_w is the numerical aperture at the wafer plane, and OPD is the optical path difference defined by the aberrations present in the system.

The intensity at the wafer plane is

$$I_w(\mathbf{r}_w) = \sum_n \left| \int_M E_{Mn}(\mathbf{r}_M) PSF_{M \rightarrow W}(\mathbf{r}_w - \mathbf{r}_M) d^2 r_M \right|^2 \quad (29)$$

where $PSF_{M \rightarrow W}$ is the inverse Fourier transform of OTF in Cartesian coordinates. I_w is the sum of the intensities from each plane wave, E_{Mn} . Equation (29) is a modification of Eq. (3) for a spatially incoherent system.

7.1. Aberration Wavefronts

Figure 2(a) shows a one-mirror imaging system. The relationship between radius of curvature R , mirror-to-mask distance S , and mirror-to-wafer S' is

$$\frac{1}{S'} + \frac{1}{S} = \frac{2}{R} \quad (30)$$

If the imaging is on-axis and the height of the mirror at the edges is ignored, the numerical aperture at the wafer plane is

$$NA_w = \frac{a}{\sqrt{a^2 + S'^2}} \quad (31)$$

where NA_w is the numerical aperture at the wafer, and NA_M is the numerical aperture at the mask, and a is the lateral radius of the mirror ($a=D/2$ from Fig. 2). The magnification of the system is

$$m = \frac{S'}{S} \quad (32)$$

We use Eq. (30)-(32) to find the mirror dimensions R , S , and S' from values of NA_w , m , and a . The optical path difference due to the primary aberrations in such a system as given by [28, ch. 4] is

$$OPD(\rho, \phi; h') = \frac{a^4}{4R} \left(\frac{1}{R} - \frac{1}{S'} \right)^2 \rho^4 + \frac{S' - R}{R^2 S'^2} a^3 h' \rho^3 \cos \phi + \frac{a^2}{R S'^2} h'^2 \rho^2 \cos^2 \phi \quad (33)$$

where ρ as before, is the normalized radial coordinate, equal to 1 at the edge of the aperture, and h' is the radial height in the image (wafer) plane. The three terms correspond to spherical aberration, coma, and astigmatism, respectively.

References

1. H. J. Levinson, *Principles of Lithography*, Edition III, (SPIE, 2010).
2. L. Pang, Y. Liu, D. Abrams, "Inverse lithography technology (ILT), what is the impact to photomask industry?" Luminescent Technologies, Inc. (2006).
3. Y. Borodovsky, W. Cheng, R. Schenker, V. Singh, "Pixelated phase mask as novel lithography RET," Proc. SPIE **6924**, 69240E (2008).
4. P. S. Davids, S. B. Bollepalli, "Generalized inverse problem for partially coherent projection lithography," Proc. SPIE **6924**, 69240X (2008).
5. V. Singh, B. Hu, K. Toh, S. Bollepalli, S. Wagner, Y. Borodovsky, "Making a trillion pixels dance," Proc. SPIE **6924**, 69240S (2008).
6. W. Cheng, J. Farnsworth, W. Kwok, A. Jamieson, N. Wilcox, M. Vernon, K. Yung, Y. Liu, J. Kim, E. Frendberg, S. Chegwidan, R. Schenker, Y. Borodovsky, "Fabrication of defect-free full-field pixelated phase mask," Proc. SPIE **6924**,

- 69241G (2008).
7. R. Schenker, S. Bollepalli, B. Hu, K. Toh, V. Singh, K. Yung, W. Cheng, Y. Borodovsky, "Integration of pixelated phase masks for full-chip random logic layers," *Proc. SPIE* 6924, 69240I (2008).
8. G. Kim, J. A. Domínguez-Caballero, R. Menon, "Design and analysis of multi-wavelength diffractive optics," *Opt. Express* **20**(3), 2814-2823 (2012).
9. J. R. Fineup, "Iterative method applied to image reconstruction and to computer-generated holograms," *Opt. Eng.* **19**(3), 297-305 (1980).
10. C. Jacobsen, M. R. Howells, "A technique for projection x-ray lithography using computer-generated holograms," *J. Appl. Phys.* **71**(6), 2993-3001 (1992).
11. J. A. Domínguez-Caballero, S. Takahashi, S. J. Lee, G. Barbastathis, "Design and fabrication of computer generated holograms for Fresnel domain lithography" in *Digital Holography and Three-Dimensional Imaging*, Vancouver Canada April 26-30 (2009).
12. Y. Cheng, A. Isoyan, J. Wallace, M. Khan, F. Cerrina, "Extreme ultraviolet holographic lithography: initial results," *Appl. Phys. Lett.* **90**(023116), (2007).
13. J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," *Las. Photon. Rev.* **5**(2), 308-321 (2011).
14. P. Seliger, M. Mahvash, C. Wang, and A. F. J. Levi, "Optimization of aperiodic dielectric structures," *J. Appl. Phys.* **100**(3), 034310 (2006).
15. W. R. Frei, D. A. Tortorelli, and H. T. Johnson, "Geometry projection method for optimizing photonic nanostructures," *Opt. Lett.* **32**(1), 77-79 (2007).
16. V. Liu and S. Fan, "Compact bends for multi-mode photonic crystal waveguides with high transmission and suppressed modal crosstalk," *Opt. Express* **21**(7), 8069-8075 (2013).
17. G. Veronis, R. W. Dutton, and S. Fan, "Method for sensitivity analysis of photonic crystal devices," *Opt. Lett.* **29**(19), 2288-2290 (2004).
18. O. D. Miller, "Photonic design: from fundamental solar cell physics to computational inverse design," Ph.D. Thesis, EECS Department, Univ. of California, Berkeley (2012).
19. C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Opt. Express*, **21**(18), 21693-21701 (2013).
20. V. Ganapati, O. D. Miller, and E. Yablonovitch, "Light trapping textures designed by electromagnetic optimization for subwavelength thick solar cells," *IEEE J. Photovolt.* **4**(1), 175-182 (2014).
21. M. P. Bendsoe and O. Sigmund, *Topology Optimization Theory, Methods and Applications* (Springer, 2003).
22. G. Strang, *Computational Science and Engineering* (Wellesley-Cambridge, 2007).
23. S. Krantz, *A Guide to Complex Variables* (2007).
24. Y. Borodovsky, "EUV lithography at insertion and beyond," <http://www.euvlitho.com/2012/P1.pdf>.
25. C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: a Practical Approach with Examples in MATLAB*, (Wiley-Blackwell, 2011).
26. J.W. Goodman, *Introduction to Fourier Optics*, Edition II, (McGraw-Hill, 1996).
27. M. Born and E. Wolf, *Principles of Optics*, Edition VII, (Cambridge University, 1999).
28. V. N. Mahajan, *Aberration Theory Made Simple*, Edition II, (SPIE, 2011).
29. Samarth Bhargava and Eli Yablonovitch "Multi-objective inverse design of sub-wavelength optical focusing structures for heat assisted magnetic recording ", *Proc. SPIE* 9201, Optical Data Storage 2014, 92010M (September 5, 2014)