Optimization of Photomask Design for Reducing Aberration-Induced Placement Error

Giuseppe Y. Mak, *Member, IEEE*, Alfred K. Wong, *Senior Member, IEEE*, and Edmund Y. Lam, *Senior Member, IEEE*

***Abstract*—In semiconductor manufacturing, the accurate placement of circuit components ensures the proper functioning of microelectronic circuits. This is often subject to photolithography, an optical technique that transfers circuit patterns from photomasks to silicon wafers. Sources of placement error include aberration and misalignment between different levels, and we focus on the former. Aberration is an optical phenomenon that often degrades imaging system performance. Since aberration differs from one imaging system to another, a photomask design that minimizes the aberration-induced placement error is desired. In this paper, we discuss the optimization process of a general one-dimensional mask pattern under a general illumination condition. The constraint is a known population mean of the root mean square aberrations for the imaging systems under consideration. To apply the theory, we search for the optimal parameters for two common mask designs: alternating phase-shifting masks (PSMs) and attenuated PSMs. The theoretical results are compared with those from a Monte Carlo analysis on a large set of imaging systems. These results are indicative to mask manufacturers and circuit designers of increasing manufacturability of circuits.**

***Index Terms*—Aberration, alternating phase-shifting mask (PSM), attenuated PSM, Monte Carlo analysis, optimization, photomask, placement error.**

I. Introduction

PHOTOLITHOGRAPHY (also called optical lithography) has long been used to transfer circuit patterns from a template called photomask (or simply mask) to silicon wafers during integrated circuit (IC) fabrication. When a light source is used to project the mask image onto the wafer, the image quality is often affected by the performance of the imaging system (also called exposure system). One of the factors that limit its performance is aberration [1]. In terms of geometrical optics, when aberration is present, the one-to-one correspondence between object and image points no longer exists. This will lead to the decrease of depth of focus (DoF) [2], linewidth error [3], pattern asymmetry, and image placement error [4] (or simply called placement error in this paper). With the migration of lithographic limit towards the 65- and 45-nm nodes, aberration control becomes a paramount concern for robust circuit printing.

Besides modifying the exposure system to keep the aberration level as low as possible, is there anything we can do with the mask design so that the image is less vulnerable to aberration? This is what we are going to address in this paper. Among the imaging artifacts caused by aberration, we focus on placement error due to its importance in various processes such as DRAM production [5]. In reality, exposure systems with varying aberration levels exist. Each produces a different degree of placement error towards the same mask (Fig. 1). A mask optimized for a particular exposure system may not be optimal for another. Besides, the redesign and manufacturing cost for a new mask is also extremely high. Therefore, we seek to optimize the mask pattern so that the placement error, averaged over the set of exposure systems in our consideration, is minimized. It is assumed that the mean of the root mean square (rms) aberrations of those exposure systems is a known constant, which is the only constraint of this optimization problem. In our previous paper [6], we already considered the optimization of alternating phase-shifting mask (PSM) with symmetric 0° and 180° phase regions under coherent imaging. We extend the generality of the theory by considering a general mask pattern with one-dimensional (1-D) spatial variation under a general illumination.

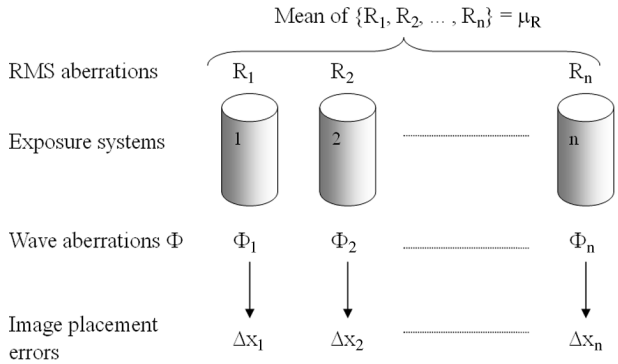


Fig. 1. Illustration of optimization problem.

II. Statistical Behavior of Aberration

Before optimization, we shall look at aberration in greater detail. Physically, aberration can be described by the optical path difference (OPD)[1] between Gaussian reference sphere S, the ideal image-forming wavefront, and the aberrated wavefront W (Fig. 2). Let Q_S and Q_W be the points where the ray $P'P$ intersects S and W, respectively. The OPD $\Phi = [Q_W Q_S]$ is called an aberration function (also known as wave aberration or wavefront error). It is regarded as positive when Q_S is situated between Q_W and P'.

[1]Optical path difference is equal to the path difference multiplied by the local refractive index.

Manuscript received November 4, 2005; revised March 23, 2006.

G. Y. Mak is with Brion Technologies Co., Ltd., Shenzhen 518057, China (e-mail: giuseppe@graduate.hku.hk).

A. K. Wong is with Magma Design Automation, Inc., Southborough, MA 01772 USA (e-mail: awong@magma-da.com).

E. Y. Lam is with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong (e-mail: elam@eee.hku.hk).

Digital Object Identifier 10.1109/TSM.2006.879412

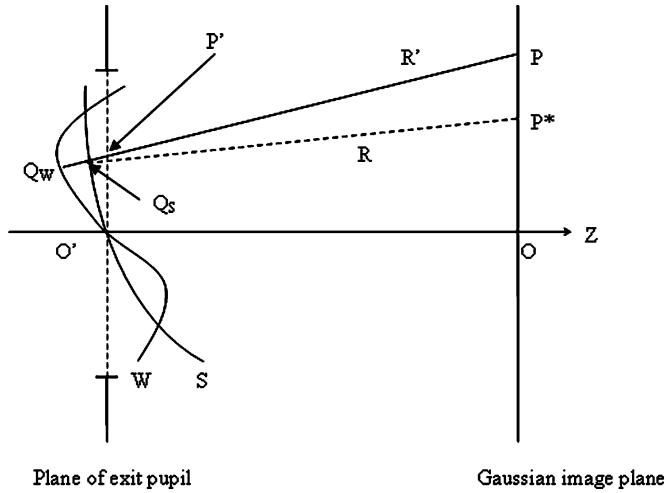


Fig. 2. Gaussian reference sphere and aberration wavefront.

Aberration function can be treated as a two-dimensional (2-D) function over the exit pupil plane. If the exit pupil is circular and the coordinates (ξ, η) on the pupil plane are normalized by the pupil radius a, the aberration function can be expanded in terms of a complete set of polynomials that are orthogonal over the interior of unit circle. One such candidate is the set of rms normalized Zernike polynomials [7]. If Z_j denotes the jth rms normalized Zernike polynomial and C_j denotes the corresponding coefficient, then the aberration function can be expressed as

$$\Phi(\xi, \eta) = \lambda\phi(\hat{f}, \hat{g}) = \lambda \sum_{j=1}^{\infty} C_j Z_j(\hat{f}, \hat{g}) \tag{1}$$

where $\hat{f} = \xi/a$, $\hat{g} = \eta/a$, λ is the wavelength of the light source, and $\phi(\hat{f}, \hat{g})$ is the normalized form of aberration function. In order to get a general impression of how large the aberration is, a parameter called rms aberration is defined as

$$\phi_{\text{rms}} = \sum_{j=1}^{\infty} C_j^2. \tag{2}$$

From the previous expressions, we see that the expansion is an infinite series. In lithography, our attention shall be limited to $\{Z_5, \ldots, Z_{37}\}$. This is because the piston term (Z_1) does not cause any image shift, x tilt (Z_2) and y tilt (Z_3) cause pattern-independent shift, and the defocus term (Z_4) causes intensity shift along the optical axis only [4]. Since we are only concerned with pattern-dependent placement error on the image plane, these Zernike terms are ignored. On the other hand, we assume that the lithographic systems are not well corrected. The higher order terms (those above 37) do not carry a sizable fraction of the rms aberration. This assumption only produces minor discrepancies between the simulated and the actual placement error values.

Going back to our original optimization problem, the exposure systems are black boxes to us now. We neither know the exact form of the aberration function, nor the value of rms aberration present in the exit pupil of each exposure system. Thus, the aberration function may be treated as a random quantity. To ask about the statistical behavior of aberration, we are actually asking how the Zernike coefficients are distributed over all interested exposure systems. Limited knowledge is available on the statistical behavior of the Zernike coefficients. As a reasonable assumption, each coefficient is modelled as an independent, normally distributed random variable with zero mean and identical nonzero variance. The justification is that the coefficients of real lithographic lenses are likely to be small in magnitude and equally likely to be positive or negative. For the sake of analysis, the variance is assumed to be the same for each coefficient. Symbolically, $C_j \sim \mathcal{N}(0, \sigma_c^2)$ for $5 \leq j \leq 37$.

Now let us determine the unknown variance σ_c^2. Referring to (2), with $C_j \sim \mathcal{N}(0, \sigma_c^2)$ and j running from 5 to 37, ϕ_{rms}^2 becomes a χ^2 random variable with $37 - 5 + 1 = 33$ degrees of freedom [8]. Let the population mean of ϕ_{rms} be μ_R. From [9], this is given by

$$\mu_R = \text{E}[\phi_{\text{rms}}] = \sqrt{2}\sigma_c \frac{\Gamma\left[\frac{1}{2}(33+1)\right]}{\Gamma\left(\frac{1}{2} \times 33\right)} \tag{3}$$

where $\text{E}[Z]$ denotes the expected value of Z and $\Gamma(p)$ is the gamma function. Hence

$$\sigma_c^2 = \left(\frac{\mu_R \Gamma(16.5)}{\sqrt{2}\Gamma(17)}\right)^2. \tag{4}$$

From these equations, we see that μ_R constraints the value of σ_c^2, which in turn governs the distribution of each Zernike coefficient.

III. Computation of Image Placement Error

After reviewing the random nature of aberration, we now need a formula to compute the image placement error based on the parameters of a lithographic system. Referring to the derivation in Appendix I, the placement error can be expressed in terms of three functions: effective light source, mask spectrum, and the aberration function. The effective light source is the image casted by an input light source on the exit pupil. It is denoted by $\hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s)$. Mask spectrum is the Fourier transform of mask transmission function. Mask transmission function represents the amplitude and phase of the transmitted electromagnetic wave at every point of a mask. Since we are only concerned with 1-D mask variation, the mask transmission function is also 1-D, denoted by $\hat{O}_x(\hat{x})$.[2] After Fourier transform, the mask spectrum $\hat{\mathcal{O}}_x(\hat{f})$ results. With these three functions in hand, the placement error can be computed by

$$\Delta\hat{x} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s) M(\hat{f}_s) \text{d}\hat{f}_s \text{d}\hat{g}_s}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s) N(\hat{f}_s) \text{d}\hat{f}_s \text{d}\hat{g}_s} \tag{5}$$

[2]In lithography, the spatial coordinates (x, y) are usually normalized by (λ/NA), where NA is the numerical aperture of the exit pupil. This is denoted by a caret $(\hat{})$ over the corresponding coordinate. The functions depending on those coordinates are also annotated by carets.

where $M(\hat{f}_s)$ and $N(\hat{f}_s)$ are two functions consisting of the mask spectrum $\hat{\mathcal{O}}_x(\hat{f})$ and the aberration function $\phi(\hat{f},\hat{g})$. For the details of M and N, please see Appendix I.

IV. Optimization Process

After introducing the image placement error formula, let us state the optimization problem more formally. Let $\{\theta_1,\theta_2,\cdots,\theta_n\}$ be a set of parameters that define a particular photomask design. These parameters can be the pitch of periodic polygons or the critical dimension (CD). For isolated features, CD is defined as the width of the smallest geometrical figure to be printed. For periodic features, however, it is defined as the smallest half pitch (i.e., half period) among the features. These parameters are wrapped inside the functions M and N in (5), because both functions contain the information about the photomask. The optimal parameter set is found by solving

$$\{\theta_1,\theta_2,\cdots,\theta_n\}_{opt}=\arg\min \mathrm{E}\left[\left(\Delta\hat{X}(\phi;\theta_1,\theta_2,\cdots,\theta_n)\right)^2\right] \quad \text{subject to } \mu_R=\mathrm{E}[\phi_{\mathrm{rms}}] \tag{6}$$

where

$$(\Delta\hat{X})^2=\left(\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)M(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)N(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s}\right)^2. \tag{7}$$

The randomness of image placement error is due to the random normalized aberration function ϕ. Since random variables are usually denoted by capital letters, we write $\Delta\hat{X}$ for image placement error hereafter. In this problem, we want to minimize the expected value of the squared placement error. The reason for taking the square is that we are only interested in the magnitude but not the direction of the image shift. To proceed, an analytical expression for $\mathrm{E}[(\Delta\hat{X})^2]$ is needed.

A. Computing Expected Squared Placement Error

The idea behind the derivation is to make use of the linear property of expected value operation. Since the derivation is somewhat tedious, we defer the details to Appendix II. The expected square placement error is given by

$$\mathrm{E}\left[(\Delta\hat{X})^2\right]=\frac{\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2A(\hat{f_{s1}},\hat{f_{s2}})\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}}{4\pi^2\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2B(\hat{f_{s1}},\hat{f_{s2}})\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}} \tag{8}$$

where A and B are two complicated functions consisting of the mask spectrum and aberration function (please see (25) for their explicit expressions). Equation (8) expresses the expected squared placement error as a function of illumination (represented by the effective light source $\hat{\mathcal{J}}$), 1-D mask spectrum (contained in A and B) and mean rms aberration μ_R [related to σ_c^2 in A and B by (4)]. The remaining task is to search for parameters that minimize the result of (8).

B. Verification of (8) and Optimal Parameter Search

To test the validity of (8), we compare the sample means of squared placement error obtained from the Monte Carlo analysis (random sampling) of (7) and the theoretical expected values from (8). We have two aims:

1) to show that the sample mean of $(\Delta\hat{X})^2$ and its expected value are approximately equal at different mask parameter combinations;
2) to show the optimal parameter set is likely to exist using Monte Carlo analysis and confirm its existence by theoretical computation.

In each trial of Monte Carlo analysis, a set of Zernike coefficients is randomly sampled according to their distributions. By substituting the coefficients into (7), a sample of $(\Delta\hat{X})^2$ is obtained. The sample mean is calculated after taking a large number of trials. For each choice on the values of the mask parameters $\{\theta_1,\theta_2,\cdots,\theta_n\}$, a different sample mean can be evaluated. Table I shows the parameter values for our lithographic system.

TABLE I
Parameters of Lithographic System

Wavelength (λ)	248 nm (KrF laser)
Numerical aperture (NA)	0.68
Mean RMS aberration μ_R	0.025λ

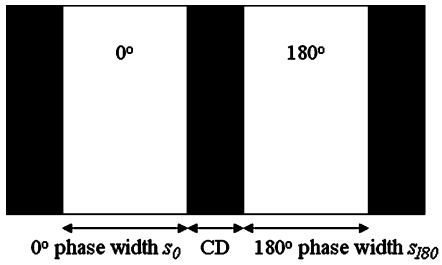


Fig. 3. Alternating PSM.

In the analysis, the mean rms aberration is taken to be 0.025λ, corresponding to a Strehl ratio of 0.975. This is often used as a guideline for lithographers to achieve the best wavefront control [10]. Using (4) and the data in Table I, the probability density function (PDF) of each Zernike coefficient is $\mathcal{N}(0, 1.9228\times 10^{-5})$.

First, we choose an alternating PSM with the structure shown in Fig. 3 for simulation. It is imaged under coherent illumination and partially coherent illumination (with partial coherence factor $\sigma = 0.3$). The CD and the widths of the 0° and 180° phase regions (called 0° and 180° phase widths hereafter) are the parameters to be optimized. By plotting the sample means of placement error against various parameter combinations, the location where minimum sample mean occurs can be found.

Let us look at an alternating PSM with symmetric phase regions (equal 0° and 180° phase widths) first. Fig. 4 shows the sample means of (7) plotted against phase width for various CDs. The number of trials for this Monte Carlo analysis is 5000. Note that the normalized placement error $\Delta\hat{X}$ is converted to the true placement error ΔX (without a hat) by multiplying the normalized one with (λ/NA). Hence, the vertical axis of each graph has nanometers squared as the unit. Now, if the value of CD is fixed (i.e., picking one of the curves), we see that each curve peaks at a phase width of approximately $1(\lambda/\mathrm{NA})$. For phase widths larger than zero, the global minimum for each curve occurs between $0.3(\lambda/\mathrm{NA})$ and $0.5(\lambda/\mathrm{NA})$. The optimal phase width is likely to lie in this range.

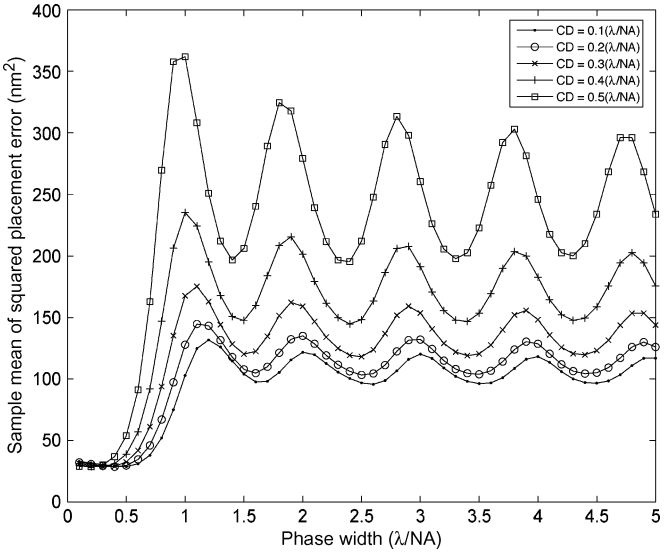


Fig. 4. Sample mean of $(\Delta X)^2$ as function of phase width under coherent imaging for various CD settings.

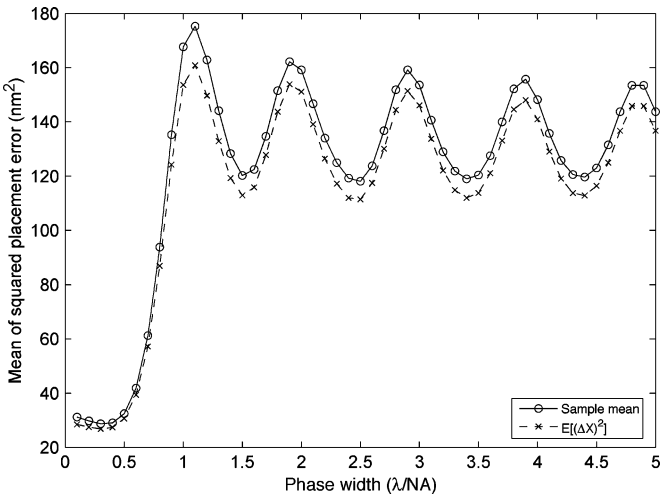


Fig. 5. Comparison between sample mean of $(\Delta X)^2$ and $\mathrm{E}[(\Delta X)^2]$ under coherent imaging.

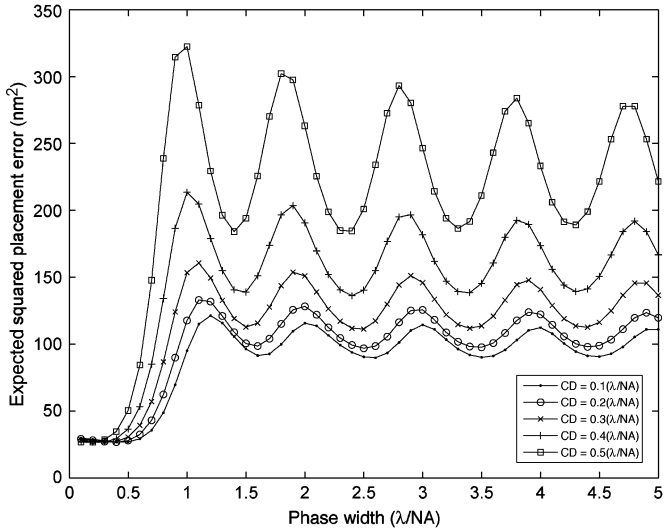


Fig. 6. $\mathrm{E}[(\Delta X)^2]$ as function of phase width under coherent imaging for various CD settings.

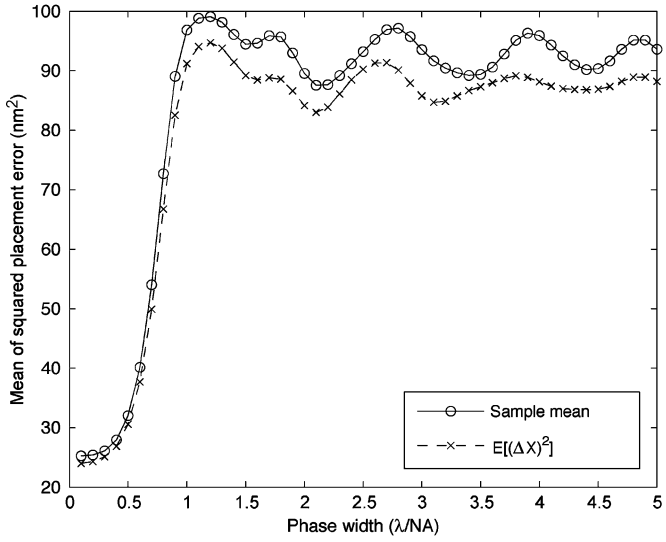


Fig. 7. Sample mean of $(\Delta X)^2$ as function of phase width under partially coherent imaging, $\sigma = 0.3$, $\mathrm{CD} = 0.3(\lambda/\mathrm{NA})$.

Now, we consider the effect when CD is varied. It is observed from Fig. 4 that the global minimum still occurs at a phase width ranging from $0.3(\lambda/\mathrm{NA})$ to $0.5(\lambda/\mathrm{NA})$ when CD varies from $0.1(\lambda/\mathrm{NA})$ to $0.5(\lambda/\mathrm{NA})$. This means that under coherent imaging, the optimal phase width is likely to be independent of the CD. These results are confirmed by the calculation using (8). Fig. 5 shows the good agreement between the Monte Carlo results and theoretical results, and Fig. 6 shows the theoretical results for different CDs.

Fig. 7 shows another plot for partially coherent imaging ($\sigma = 0.3$). The number of trials is also 5000. In this plot, no global minimum is observed for both the Monte Carlo curve and theoretical curve. The fluctuations of image placement error for phase widths greater than $1(\lambda/\mathrm{NA})$ are also smaller than that of coherent imaging. Therefore, to achieve low aberration sensitivity in partially coherent imaging, we should opt for a phase width below $1(\lambda/\mathrm{NA})$.

Now, we turn our attention to an alternating PSM with asymmetric phase regions. The 0° and 180° phase widths are allowed to be different. Fig. 8 shows the sample mean of $(\Delta X)^2$ plotting against the 0° and 180° phase widths under coherent imaging (number of trials $= 1000$) and Fig. 9 shows the mesh plot under partially coherent imaging ($\sigma = 0.3$, number of trials $= 100$). The value of CD is fixed at $0.3(\lambda/\mathrm{NA})$. The theoretical placement errors from (8) are also plotted in Figs. 10 and 11.

From these plots, we can observe that the sample mean of $(\Delta X)^2$ is likely to attain a low value when phase regions are symmetric (along the 0° phase width $=$ 180° phase width line) and when the phase widths are below $1(\lambda/\mathrm{NA})$. The placement error is also comparatively low when either of the phase widths is between $0.5(\lambda/\mathrm{NA})$ and $0.7(\lambda/\mathrm{NA})$. In contrast, a high placement error occurs when the phase regions are very narrow or highly asymmetric. For example, when 0° phase width $= 0.1(\lambda/\mathrm{NA})$ and

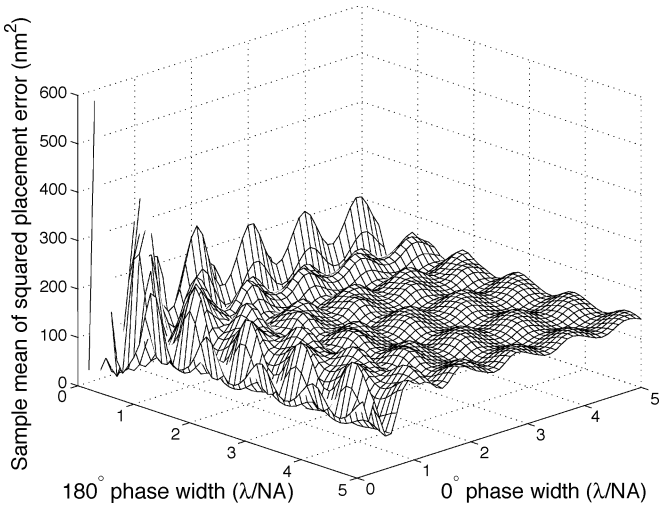


Fig. 8. Sample mean of $(\Delta X)^2$ as function of 0° and 180° phase widths under coherent imaging.

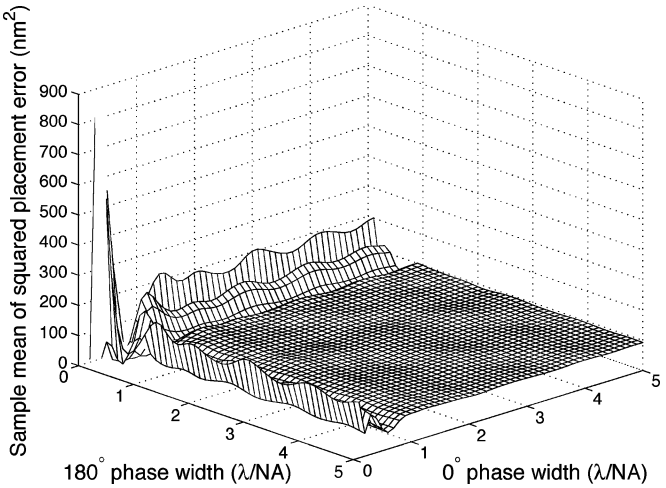


Fig. 9. Sample mean of $(\Delta X)^2$ as function of 0° and 180° phase widths under partially coherent imaging ($\sigma = 0.3$).

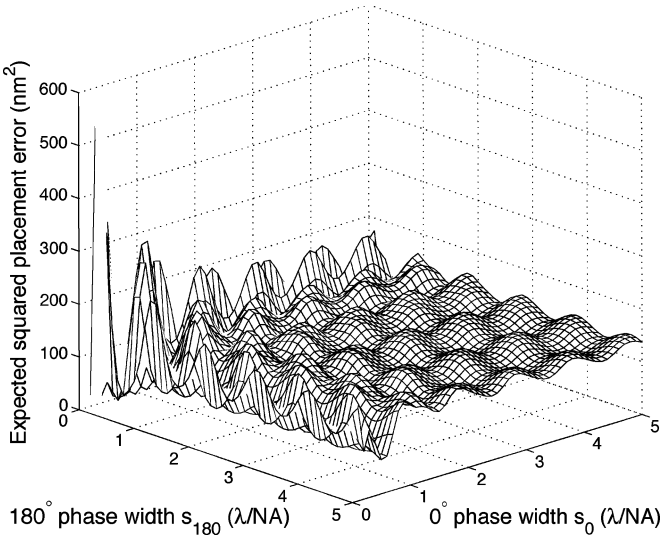


Fig. 10. $\mathrm{E}[(\Delta X)^2]$ as function of 0° and 180° phase widths under coherent imaging.

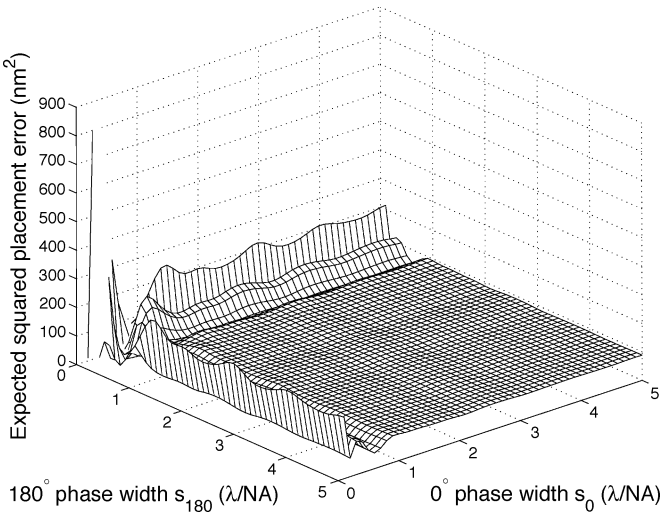


Fig. 11. $\mathrm{E}[(\Delta X)^2]$ as function of 0° and 180° phase widths under partially coherent imaging ($\sigma = 0.3$).

180° phase width $= 0.2(\lambda/\mathrm{NA})$, the sample mean of $(\Delta X)^2$ reaches 600 nm^2 under coherent imaging. For phase widths larger than $1(\lambda/\mathrm{NA})$, the placement error values fluctuates around 130 nm^2 under coherent imaging; whereas, the values stabilize around 90 nm^2 under partially coherent imaging. In general, the placement error under coherent imaging is higher than that under partially coherent imaging, except for a few points. The implications of these observations are that symmetric alternating PSMs with phase widths between $0.3(\lambda/\mathrm{NA})$ and $0.5(\lambda/\mathrm{NA})$ can achieve the lowest placement sensitivity. If space is not available on the mask so that the phase regions must be made asymmetric, one of them should be designed with a width between $0.5(\lambda/\mathrm{NA})$ and $0.7(\lambda/\mathrm{NA})$.

Having discussed the optimization of alternating PSM, our next target is attenuated PSM (Fig. 14). Since attenuated PSM is often exposed under annular illumination with subresolution assist features, we used an annular source as well as coherent source as a comparison. The effective light source of annular illumination has the shape of circular ring. It is defined by two parameters: radius of inner circle (σ_i) and radius of outer circle (σ_o) (also denoted by $\sigma = (\sigma_i, \sigma_o)$). We took $\sigma = (0.5, 0.8)$ in this case. Our target parameter is amplitude transmission factor t and we allowed it to run from 0 to 0.3 only. We chose this range because when t is large, sidelobe printing becomes so severe that the image is unacceptable. The number of trials for the Monte Carlo analysis is 5000 for both illuminations. The pitch of attenuated PSM is $10(\lambda/\mathrm{NA})$.

From simulation, it is found that under coherent imaging, both the sample mean (solid line) and expected value (dotted–solid line) of squared placement error are fairly constant for the concerned domain of t (Fig. 12). Since the difference between the Monte Carlo and theoretical results is only about 3 nm^2 on average, we may still accept the theoretical results as a close match with the Monte Carlo results. On the other hand, for the annular condition, the placement error is increasing over the whole domain (Fig. 13). The industrially accepted amplitude transmission level is approximately 0.24 to 0.26 (corresponding to an intensity transmittance from 6% to

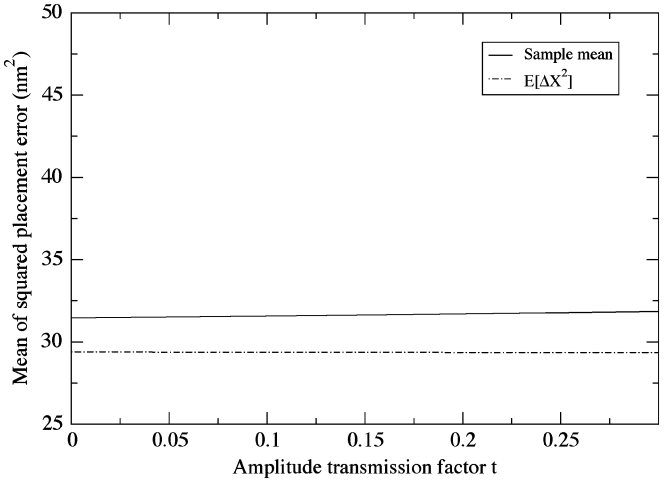


Fig. 12. Comparison between sample mean of $(\Delta X)^2$ and $\mathrm{E}[(\Delta X)^2]$ as function of amplitude transmission factor under coherent imaging.

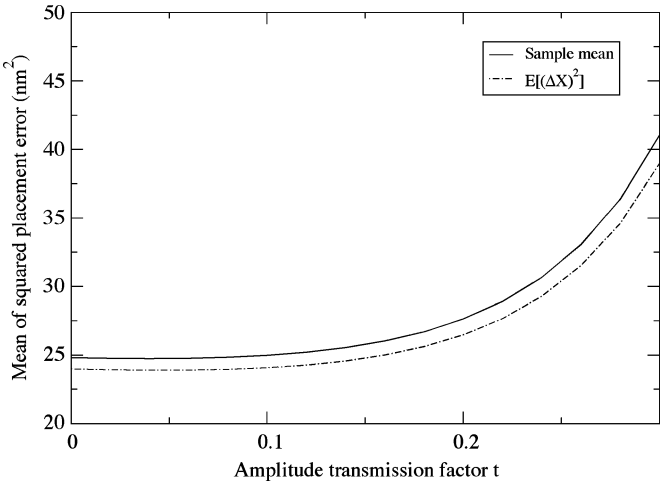


Fig. 13. Comparison between sample mean of $(\Delta X)^2$ and $\mathrm{E}[(\Delta X)^2]$ as function of amplitude tranmission factor under annular illumination ($\sigma_i = 0.5$, $\sigma_o = 0.8$).

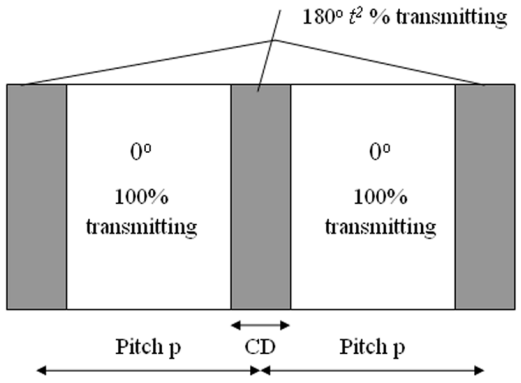


Fig. 14. Attenuated PSM.

7%) [11]. It is a balance between process latitude, undesirable side-lobe printing, and material availability. The previous results indicate that the industry can consider an attenuator with higher transmittance when σ is small and a lower transmittance when annular scheme is used.

V. Summary

In this paper, we have considered the optimization of a general photomask pattern under a general light condition and a set of exposure systems whose mean rms aberration is given. The aim is to lower the average image placement error. Equation (8) has been derived to express the optimization criterion in terms of effective light source, mask spectrum, and mean rms aberration. We demonstrate the applicability of the formula by considering the optimization of alternating PSM under coherent imaging and partially coherent imaging and attenuated PSM under coherent and annular illuminations. For alternating PSM, it is found that the 0° and 180° phase regions should be symmetric and they should be $0.3(\lambda/\mathrm{NA})$ to $0.5(\lambda/\mathrm{NA})$ wide in order to achieve the lowest placement sensitivity. Even if the phase regions have to be asymmetric, one of the phase widths should be set between $0.5(\lambda/\mathrm{NA})$ and $0.7(\lambda/\mathrm{NA})$. On the other hand, the placement error is independent of the amplitude tranmission factor of attenuated PSM under coherent imaging. However, under the more common annular illumination, the placement error increases with increasing transmission factor instead. This indicates that the industry can consider an intensity transmittance higher than the commonly accepted value of 7% when σ is small. However, a lower transmittance should be considered when annular source is utilized. Following similar procedures, it is hoped that other mask technologies can also be optimized to achieve low placement sensitivity towards aberration in the future.

Appendix I
Derivation of Image Placement Error Formula

The placement error of a feature image is computed by considering the extremum (maximum or minimum) shift of the image intensity profile. From diffractive optics, the intensity of an image due to a mask with features varying in one dimension only is given by

$$I(\hat{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s) I_s(\hat{f}_s; \hat{x}) \mathrm{d}\hat{f}_s \mathrm{d}\hat{g}_s \tag{9}$$

where

$$\hat{I}_s(\hat{x}; \hat{f}_s) = \int_{-1}^{1}\int_{-1}^{1} \hat{\mathcal{O}}_x(\hat{f}_{1s}) \hat{\mathcal{O}}_x^*(\hat{f}_{2s}) e^{-i2\pi[\hat{f}_{12}\hat{x} + \phi_{12}]} \mathrm{d}\hat{f}_1 \mathrm{d}\hat{f}_2$$

$$\hat{f}_{ij} = \hat{f}_i - \hat{f}_j, \text{ for any indices } i \text{ and } j$$

$$\phi_{12} = \phi(\hat{f}_1, 0) - \phi(\hat{f}_2, 0)$$

where $\hat{\mathcal{J}}$ is the effective light source and $\hat{\mathcal{O}}_x$ is the mask spectrum. The asterisk (*) denotes complex conjugation. The extremum position is found by differentiating (9) with respect to $\hat{x}$ and setting the derivative to zero. Since $\hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s)$ does not depend on $\hat{x}$, the differentiation is applied on $\hat{I}_s$ only

$$\frac{\mathrm{d}\hat{I}}{\mathrm{d}\hat{x}} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{\mathcal{J}}(\hat{f}_s, \hat{g}_s) \frac{\mathrm{d}\hat{I}_s(\hat{x}; \hat{f}_s)}{\mathrm{d}\hat{x}} \mathrm{d}\hat{f}_s \mathrm{d}\hat{g}_s = 0. \tag{10}$$

By evaluating the derivative $\mathrm{d}\hat{I}_s/\mathrm{d}\hat{x}$ and taking first-order approximation on the complex exponential factor appearing in the derivation, we get

$$\frac{\mathrm{d}\hat{I}_s(\hat{x};\hat{f}_s)}{\mathrm{d}\hat{x}} = 4\pi \int_{-1}^{1}\int_{\hat{f}_2}^{1} \hat{f}_{12}\zeta(\hat{x},\hat{f}_s;\hat{f}_1,\hat{f}_2)\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2 \qquad (11)$$

where

$$\begin{aligned}\zeta(\hat{x},\hat{f}_s;\hat{f}_1,\hat{f}_2)) = &\Big[D_{12}(\hat{f}_s)\cos(2\pi\phi_{12}) \\ &-S_{12}(\hat{f}_s)\sin(2\pi\phi_{12})\Big] \\ &-2\pi\hat{f}_{12}\Big[D_{12}(\hat{f}_s)\sin(2\pi\phi_{12}) \\ &+S_{12}(\hat{f}_s)\cos(2\pi\phi_{12})\Big]\hat{x} \\ D_{12}(\hat{f}_s) = &\operatorname{Im}\left[\hat{\mathcal{O}}_x(f_{1s})\right]\operatorname{Re}\left[\hat{\mathcal{O}}_x(f_{2s})\right] \\ &-\operatorname{Re}\left[\hat{\mathcal{O}}_x(f_{1s})\right]\operatorname{Im}\left[\hat{\mathcal{O}}_x(f_{2s})\right] \\ S_{12}(\hat{f}_s) = &\operatorname{Re}\left[\hat{\mathcal{O}}_x(f_{1s})\right]\operatorname{Re}\left[\hat{\mathcal{O}}_x(f_{2s})\right] \\ &+\operatorname{Im}\left[\hat{\mathcal{O}}_x(f_{1s})\right]\operatorname{Im}\left[\hat{\mathcal{O}}_x(f_{2s})\right].\end{aligned}$$

Note that ζ is linear in $\hat{x}$. By putting (11) into (10), we can get a linear equation in $\hat{x}$. In solving, we get the image placement error formula

$$\Delta\hat{x} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)M(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)N(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s} \qquad (12)$$

where

$$\begin{aligned}M(\hat{f}_s) = &\int_{-1}^{1}\int_{\hat{f}_2}^{1}\Big[D_{12}(\hat{f}_s)\cos(2\pi\phi_{12}) \\ &-S_{12}(\hat{f}_s)\sin(2\pi\phi_{12})\Big]\hat{f}_{12}\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2 \\ N(\hat{f}_s) = &2\pi\int_{-1}^{1}\int_{\hat{f}_2}^{1}\Big[D_{12}(\hat{f}_s)\sin(2\pi\phi_{12}) \\ &+S_{12}(\hat{f}_s)\cos(2\pi\phi_{12})\Big]\hat{f}_{12}^2\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2.\end{aligned}$$

APPENDIX II
DERIVATION OF EXPECTED SQUARED PLACEMENT ERROR

The expected squared placement error $\mathrm{E}[(\Delta\hat{X})^2]$ can be determined *without* knowing the PDF of $\Delta\hat{X}$. The idea is as follows.

Let U be the numerator and V be the denominator of (7). From (12), it is noted that both U and V depend on ϕ_{12}, which is the difference of the aberration function values evaluated at $(\hat{f}_1,0)$ and $(\hat{f}_2,0)$. As mentioned previously, since the Zernike coefficients C_j are random, ϕ_{12}, U, and V are also random in nature. However, we can still obtain the PDF and statistical parameters of ϕ_{12} (e.g., mean and variance), because all of them can be determined from the statistical behavior of the Zernike coefficients. If the expected value operation in our optimization problem can be performed on ϕ_{12} directly, our problem is mostly solved. The difficulty here is that both U and V are random, and there is no existing theorem in simplifying the expected value of a quotient of two random variables. In order to allow further operation on (7), we adopt

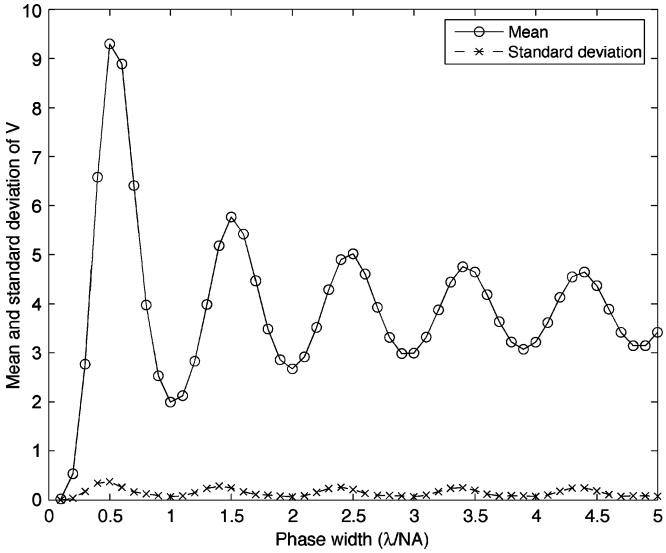


Fig. 15. Simulated mean and standard deviation of V.

$$\mathrm{E}\left[\frac{U}{V}\right] \approx \frac{\mathrm{E}[U]}{\mathrm{E}[V]}. \qquad (13)$$

The use of this approximation is to allow us to interchange the order of the two mathematical operations: expectation and double integral performed on U and V. It is justified because the standard deviation of V is usually very small as compared with the mean of V. It implies that most of the values of V are close to the mean and we may use the mean to approximate V. This can be observed from the Monte Carlo simulation of the mean and standard deviation of V using a symmetric alternating PSM under coherent imaging (Fig. 15). The wavelength of light source, the numerical aperture, and mean rms aberration of exposure system are the same as those in Table I. From Fig. 15, the standard deviation is very close to zero for every phase width, while the mean is much greater than zero, except for a phase width equal to $0.1(\lambda/\mathrm{NA})$. Nevertheless, this approximation yields a good agreement between the theoretical results and Monte Carlo results, as seen in Section IV-B.

Now, $\mathrm{E}[(\Delta\hat{X})^2]$ can be expanded as

$$\begin{aligned}&\mathrm{E}\left[(\Delta\hat{X})^2\right] \\ &\approx \frac{\mathrm{E}\left[\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)M(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s\right)^2\right]}{\mathrm{E}\left[\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{\mathcal{J}}(\hat{f}_s,\hat{g}_s)N(\hat{f}_s)\mathrm{d}\hat{f}_s\mathrm{d}\hat{g}_s\right)^2\right]} \\ &= \frac{\mathrm{E}\left[\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2M_1M_2\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}\right]}{\mathrm{E}\left[\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2N_1N_2\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}\right]} \\ &= \frac{\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2\mathrm{E}[M_1M_2]\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}}{\int_{-\infty}^{\infty}\cdots\int\hat{\mathcal{J}}_1\hat{\mathcal{J}}_2\mathrm{E}[N_1N_2]\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}}\end{aligned} \qquad (14)$$

where $\hat{\mathcal{J}}_1 = \hat{\mathcal{J}}(\hat{f}_{s1}, \hat{g}_{s1})$, $\hat{\mathcal{J}}_2 = \hat{\mathcal{J}}(\hat{f}_{s2}, \hat{g}_{s2})$, $M_1 = M(\hat{f}_{s1})$, $M_2 = M(\hat{f}_{s2})$, $N_1 = N(\hat{f}_{s1})$, $N_2 = N(\hat{f}_{s2})$. The problem turns to the evaluation of $\mathrm{E}[M_1M_2]$ and $\mathrm{E}[N_1N_2]$

$$\begin{aligned}\mathrm{E}[M_1M_2] &= \int_{-1}^{1}\int_{\hat{f}_4}^{1}\int_{-1}^{1}\int_{\hat{f}_2}^{1} \mathrm{E}\left[(D_{121}\cos(2\pi\phi_{12}) - S_{121}\sin(2\pi\phi_{12}))\right. \\ &\quad \left.\times (D_{342}\cos(2\pi\phi_{34}) - S_{342}\sin(2\pi\phi_{34}))\right] \\ &\quad \times \hat{f}_{12}\hat{f}_{34}\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2\mathrm{d}\hat{f}_3\mathrm{d}\hat{f}_4 \qquad (15a)\end{aligned}$$

$$\begin{aligned}\mathrm{E}[N_1N_2] &= \int_{-1}^{1}\int_{\hat{f}_4}^{1}\int_{-1}^{1}\int_{\hat{f}_2}^{1} \mathrm{E}\left[(D_{121}\sin(2\pi\phi_{12}) + S_{121}\cos(2\pi\phi_{12}))\right. \\ &\quad \left.\times (D_{342}\sin(2\pi\phi_{34}) + S_{342}\cos(2\pi\phi_{34}))\right] \\ &\quad \times \hat{f}_{12}^2\hat{f}_{34}^2\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2\mathrm{d}\hat{f}_3\mathrm{d}\hat{f}_4 \qquad (15b)\end{aligned}$$

where $D_{121} = D_{12}(\hat{f}_{s1})$, $S_{121} = S_{12}(\hat{f}_{s1})$, $D_{342} = D_{34}(\hat{f}_{s2})$, and $S_{342} = S_{34}(\hat{f}_{s2})$. The expected value in (15a) is expanded as

$$\begin{aligned}\mathrm{E}[\cdots] &= D_{121}D_{342}\mathrm{E}\left[\cos(2\pi\phi_{12})\cos(2\pi\phi_{34})\right] \\ &\quad - S_{121}D_{342}\mathrm{E}\left[\sin(2\pi\phi_{12})\cos(2\pi\phi_{34})\right] \\ &\quad - D_{121}S_{342}\mathrm{E}\left[\cos(2\pi\phi_{12})\sin(2\pi\phi_{34})\right] \\ &\quad + S_{121}S_{342}\mathrm{E}\left[\sin(2\pi\phi_{12})\sin(2\pi\phi_{34})\right] \qquad (16)\end{aligned}$$

and the expected value in (15b) similarly becomes

$$\begin{aligned}\mathrm{E}[\cdots] &= D_{121}D_{342}\mathrm{E}\left[\sin(2\pi\phi_{12})\sin(2\pi\phi_{34})\right] \\ &\quad + S_{121}D_{342}\mathrm{E}\left[\cos(2\pi\phi_{12})\sin(2\pi\phi_{34})\right] \\ &\quad + D_{121}S_{342}\mathrm{E}\left[\sin(2\pi\phi_{12})\cos(2\pi\phi_{34})\right] \\ &\quad + S_{121}S_{342}\mathrm{E}\left[\cos(2\pi\phi_{12})\cos(2\pi\phi_{34})\right]. \qquad (17)\end{aligned}$$

At this stage, the product-to-sum formulas in trigonometry can be used to evaluate each expected value, and we end up encountering the following four expressions involving sines and cosines:

$$\begin{aligned}&\mathrm{E}\left[\cos\left(2\pi(\phi_{12} - \phi_{34})\right)\right],\ \mathrm{E}\left[\cos\left(2\pi(\phi_{12} + \phi_{34})\right)\right] \\ &\mathrm{E}\left[\sin\left(2\pi(\phi_{12} - \phi_{34})\right)\right], \mathrm{E}\left[\sin\left(2\pi(\phi_{12} + \phi_{34})\right)\right].\end{aligned}$$

Since $(\phi_{12} \pm \phi_{34})$ can be treated as a linear combination of Zernike coefficients, i.e.,

$$\phi_{12} \pm \phi_{34} = \sum_i C_i\left(Z_{i1} - Z_{i2} \pm (Z_{i3} - Z_{i4})\right) \qquad (18)$$

and each coefficient is assumed to be independent and normal, $(\phi_{12} \pm \phi_{34})$ also becomes normal. With this observation, we can continue by considering the moment generating function of a normal random variable. Let $Y \sim \mathcal{N}(\mu, \sigma^2)$. The moment generating function of Y is [12]

$$\mathrm{E}[e^{tY}] = e^{\mu t + \frac{\sigma^2 t^2}{2}} \qquad (19)$$

for any complex number t. Then

$$\begin{aligned}\mathrm{E}[\cos aY] &= \mathrm{E}\left[\frac{e^{iaY} + e^{-iaY}}{2}\right] \\ &= \frac{1}{2}\left(\mathrm{E}[e^{iaY}] + \mathrm{E}[e^{-iaY}]\right) \\ &= \frac{1}{2}\left(e^{ia\mu - \frac{a^2\sigma^2}{2}} + e^{-ia\mu - \frac{a^2\sigma^2}{2}}\right) \\ &= e^{-\frac{a^2\sigma^2}{2}}\cos a\mu \qquad (20a)\\ \mathrm{E}[\sin aY] &= \mathrm{E}\left[\frac{e^{iaY} - e^{-iaY}}{2i}\right] \\ &= \frac{1}{2i}\left(\mathrm{E}[e^{iaY}] - \mathrm{E}[e^{-iaY}]\right) \\ &= \frac{1}{2i}\left(e^{ia\mu - \frac{a^2\sigma^2}{2}} - e^{-ia\mu - \frac{a^2\sigma^2}{2}}\right) \\ &= e^{-\frac{a^2\sigma^2}{2}}\sin a\mu \qquad (20b)\end{aligned}$$

where a is a constant. By substituting $a = 2\pi$, $Y = \phi_{12} - \phi_{34}$ (or $\phi_{12} + \phi_{34}$), we get

$$\begin{aligned}&\mathrm{E}\left[\cos\left(2\pi(\phi_{12} - \phi_{34})\right)\right] \\ &\quad = e^{-2\pi^2\mathrm{Var}[\phi_{12} - \phi_{34}]}\cos\left(2\pi\mathrm{E}[\phi_{12} - \phi_{34}]\right) \qquad (21a)\\ &\mathrm{E}\left[\cos\left(2\pi(\phi_{12} + \phi_{34})\right)\right] \\ &\quad = e^{-2\pi^2\mathrm{Var}[\phi_{12} + \phi_{34}]}\cos\left(2\pi\mathrm{E}[\phi_{12} + \phi_{34}]\right) \qquad (21b)\\ &\mathrm{E}\left[\sin\left(2\pi(\phi_{12} - \phi_{34})\right)\right] \\ &\quad = e^{-2\pi^2\mathrm{Var}[\phi_{12} - \phi_{34}]}\sin\left(2\pi\mathrm{E}[\phi_{12} - \phi_{34}]\right) \qquad (21c)\\ &\mathrm{E}\left[\sin\left(2\pi(\phi_{12} + \phi_{34})\right)\right] \\ &\quad = e^{-2\pi^2\mathrm{Var}[\phi_{12} + \phi_{34}]}\sin\left(2\pi\mathrm{E}[\phi_{12} + \phi_{34}]\right). \qquad (21d)\end{aligned}$$

The expected value operation is now directly applied to the aberration functions. The mean and variance of $(\phi_{12} - \phi_{34})$ and $(\phi_{12} + \phi_{34})$ are

$$\begin{aligned}\mathrm{E}[\phi_{12} - \phi_{34}] &= \mathrm{E}\left[\sum_i C_i(Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})\right] \\ &= \sum_i (Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})\mathrm{E}[C_i] \\ &= 0 \\ \mathrm{E}[\phi_{12} + \phi_{34}] &= \mathrm{E}\left[\sum_i C_i(Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})\right] \\ &= \sum_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})\mathrm{E}[C_i] \\ &= 0 \\ \mathrm{Var}[\phi_{12} - \phi_{34}] &= \mathrm{Var}\left[\sum_i C_i(Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})\right] \\ &= \sum_i (Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})^2\mathrm{Var}[C_i] \\ &= \sigma_c^2 P_{1234} \\ \mathrm{Var}[\phi_{12} + \phi_{34}] &= \mathrm{Var}\left[\sum_i C_i(Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})\right] \\ &= \sum_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})^2\mathrm{Var}[C_i] \\ &= \sigma_c^2 Q_{1234}\end{aligned}$$

where $Z_{ij} = Z_i(\hat{f}_j, 0)$ for $j = 1,2,3,4$, and

$$P_{1234} = \sum_i (Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})^2$$
$$Q_{1234} = \sum_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})^2$$

which are functions of Zernike polynomials only. By substituting these back to (21a)–(21d), we have

$$\mathrm{E}\left[\cos\left(2\pi(\phi_{12} - \phi_{34})\right)\right] = e^{-2\pi^2\sigma_c^2 P_{1234}} \tag{22a}$$
$$\mathrm{E}\left[\cos\left(2\pi(\phi_{12} + \phi_{34})\right)\right] = e^{-2\pi^2\sigma_c^2 Q_{1234}} \tag{22b}$$
$$\mathrm{E}\left[\sin\left(2\pi(\phi_{12} - \phi_{34})\right)\right] = 0 \tag{22c}$$
$$\mathrm{E}\left[\sin\left(2\pi(\phi_{12} + \phi_{34})\right)\right] = 0. \tag{22d}$$

Hence, (16) becomes

$$\mathrm{E}[\cdots] = \frac{1}{2}\left[D_{121}D_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} + e^{-2\pi^2\sigma_c^2 Q_{1234}}\right) + S_{121}S_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} - e^{-2\pi^2\sigma_c^2 Q_{1234}}\right)\right]. \tag{23}$$

Similarly, (17) becomes

$$\mathrm{E}[\cdots] = \frac{1}{2}\left[D_{121}D_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} - e^{-2\pi^2\sigma_c^2 Q_{1234}}\right) + S_{121}S_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} + e^{-2\pi^2\sigma_c^2 Q_{1234}}\right)\right]. \tag{24}$$

Putting (23) into (15a) and (24) into (15b), and putting the resulting equation into (14), $\mathrm{E}[(\Delta\hat{X})^2]$ is finally given by

$$\mathrm{E}\left[(\Delta\hat{X})^2\right] = \frac{\int_{-\infty}^{\infty}\cdots\int \hat{\mathcal{J}}_1\hat{\mathcal{J}}_2 A(\hat{f_{s1}}, \hat{f_{s2}})\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}}{4\pi^2\int_{-\infty}^{\infty}\cdots\int \hat{\mathcal{J}}_1\hat{\mathcal{J}}_2 B(\hat{f_{s1}}, \hat{f_{s2}})\mathrm{d}\hat{f_{s1}}\mathrm{d}\hat{g_{s1}}\mathrm{d}\hat{f_{s2}}\mathrm{d}\hat{g_{s2}}} \tag{25}$$

where

$$A(\hat{f_{s1}}, \hat{f_{s2}}) = \int_{-1}^{1}\int_{\hat{f}_4}^{1}\int_{-1}^{1}\int_{\hat{f}_2}^{1}\left[D_{121}D_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} + e^{-2\pi^2\sigma_c^2 Q_{1234}}\right) + S_{121}S_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} - e^{-2\pi^2\sigma_c^2 Q_{1234}}\right)\right] \times \hat{f}_{12}\hat{f}_{34}\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2\mathrm{d}\hat{f}_3\mathrm{d}\hat{f}_4$$

$$B(\hat{f_{s1}}, \hat{f_{s2}}) = \int_{-1}^{1}\int_{\hat{f}_4}^{1}\int_{-1}^{1}\int_{\hat{f}_2}^{1}\left[D_{121}D_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} - e^{-2\pi^2\sigma_c^2 Q_{1234}}\right) + S_{121}S_{342}\left(e^{-2\pi^2\sigma_c^2 P_{1234}} + e^{-2\pi^2\sigma_c^2 Q_{1234}}\right)\right] \times \hat{f}_{12}^2\hat{f}_{34}^2\mathrm{d}\hat{f}_1\mathrm{d}\hat{f}_2\mathrm{d}\hat{f}q_3\mathrm{d}\hat{f}_4.$$

REFERENCES

[1] M. Born and E. Wolf, *Principles of Optics*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[2] S. Nakao, A. Nakae, K. Tsujita, and W. Wakamiya, "Impact of spherical aberrations on printing characteristics of irregularly aligned patterns of alternating phase shift mask," *Japanese J. Appl. Phys.*, vol. 38(1999), no. 4A, pp. 1919–1926, Apr. 1999.

[3] R. T. Schmidt, C. A. Spence, L. Capodieci, Z. Krivokapic, B. Geh, and D. G. Flagello, "Impact of coma on CD control for multiphase PSM designs," in *Optical Microlithography XI* ser. Proc. SPIE, L. V. den Hove, Ed., vol. 3334, Jun. 1998, pp. 15–24.

[4] T. A. Brunner, "Impact of lens aberration on optical lithography," *IBM J. Res. Development*, vol. 41, no. 1/2, Jan. 1997.

[5] B.-H. Cho, D. Yim, C.-H. Park, S.-H. Lee, H.-J. Yang, J.-H. Choi, Y.-C. Shin, C.-D. Kim, J.-S. Choi, K.-O. Kang, S.-W. Kim, T.-H. Yu, J.-K. Hong, J.-C. Kim, M.-S. Han, H.-Y. Heo, Y.-D. Kim, D.-D. Lee, G.-H. Yoon, J. B. van Schoot, T. Theeuwes, and Y.-H. Min, "Aberration sensitivity control for the isolation layer in low-k_1 DRAM process," in *Optical Microlithography XV* ser. Proc. SPIE, A. Yen, Ed., vol. 4691, Jul. 2002, pp. 831–839.

[6] G. Y. Mak, A. K. Wong, and E. Y. Lam, "Alternating phase-shifting mask design for low aberration sensitivity," *J. Microlithography, Microfabrication, and Microsystems*, vol. 4, no. 1, p. 013008, 2005.

[7] M. Born and E. Wolf, *Principles of Optics*, 7th ed. London, U.K.: Cambridge Univ. Press, 1999, ch. 9.2, pp. 523–527.

[8] J. Goodman, *Statistical Optics*. New York: Wiley, 2000.

[9] J. G. Proakis, Ed., *Digital Communications*, 3rd ed. Englewood Cliffs, NJ: Prentice–Hall, 1995, pp. 46–47.

[10] C. J. Progler and D. C. Wheeler, "Optical lens specifications from the user's perspective," in *Optical Microlithography XI* ser. Proc. SPIE, L. V. den Hove, Ed., vol. 3334, Jun. 1998, pp. 256–268.

[11] A. K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. Bellingham, WA: SPIE Press, 2001, ch. 6, p. 147.

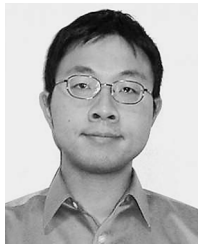
[12] R. J. Larsen and M. L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 2001, ch. 4, p. 281.

Giuseppe Y. Mak (M'05) received the B.Eng. and M.Phil. degrees in electrical and electronic engineering from the University of Hong Kong, Hong Kong, in 2002 and 2005, respectively, where he focused on the reduction of aberration sensitivity of photomasks.

He was a Research Assistant in the Department of Electrical and Electronic Engineering, University of Hong Kong, in 2004 and 2005, where he conducted research on the application of integrated computational imaging system concept. He is currently an Application Engineer at Brion Technologies Inc., Shenzhen, China. His research interests include microlithography and optical imaging system design.

Alfred K. Wong (M'98–SM'04) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1990, 1992, and 1994, respectively, where he focused on electromagnetic simulation using the time-domain finite-difference (TDFD) method on massively parallel computers.

He was affiliated with the Advanced Silicon Processing Division at IMEC, Belgium, in 1994 and 1995, the IBM Semiconductor Research and Development Center, Hopewell Junction, NY, from 1995 to 2000, and the Department of Electrical and Electronic Engineering, University of Hong Kong, from 2000 to 2003. His current research interests include resolution enhancement techniques and synergism between circuit design and fabrication. He is now with Magma Design Automation, Inc., Southborough, MA.



Edmund Y. Lam (S'97–M'00–SM'05) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1995, 1996, and 2000, respectively.

At Stanford, he conducted research for the Programmable Digital Camera project. He also consulted in industry in the areas of digital camera systems design and algorithm development. Before returning to academia, he was affiliated with the Reticle and Photomask Inspection Division (RAPID) of KLA-Tencor Corporation, San Jose, CA, as a Senior Engineer. He was involved in the design of defect detection tools for the core die-to-die and die-to-database inspections. He is now an Assistant Professor of electrical and electronic engineering at the University of Hong Kong, Hong Kong. His research interests include optics and imaging and their use in the semiconductor manufacturing process.