Alternating phase-shifting mask design for low aberration sensitivity

Giuseppe Y. Mak[a], Alfred K. Wong[b] and Edmund Y. Lam[a]

[a]Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong;
[b]Fortis Systems, Inc., U.S.A.

ABSTRACT

Theories are developed to optimize the mask structure of alternating phase-shifting masks (PSMs) to minimize the average image placement error towards aberration under coherent imaging. The constraint of the optimization is a given mean value of RMS aberration, which corresponds to infinitely many sets of random Zernike coefficients. To begin the analysis, the image placement error is expressed as a function of the mask spectrum and the wave aberration. Monte Carlo analysis on the Zernike coefficients is then performed, which assures us that a global minimum of average image placement error is likely to occur at low phase widths. This result is confirmed by analytically considering the expected value of the square of the image placement error. By Golden Section Search, the optimal phase width is found to be $0.3707(\lambda/\mathrm{NA})$ at 0.07λ RMS aberration. This methodology of finding the optimal phase width is applicable to the design of all alternating PSMs.

Keywords: aberration, alternating phase-shifting mask, image placement error, phase width, Monte Carlo analysis, Zernike coefficient

1. INTRODUCTION

In the production of integrated circuits, the miniaturization of devices and the rapid increase in integration density have led to the development of resolution enhancement techniques (RETs) in lithography.[1] Examples include modified illumination (such as annular, dipole and quadrupole illumination),[2,3] alternating phase-shifting mask (PSM)[4]and attenuated PSM.[5] Alternating PSM is one of the RETs that offers superior image quality for printing small and dark features. Making use of the destructive interference of light rays that are 180° out of phase, alternating PSMs are capable of both light-field and dark-field applications. A k_1 factor as low as 0.1 has been achieved by using alternating PSMs.[6]

Current issues such as image intensity imbalance, aberration sensitivity and mask defects have placed a challenge on the design of alternating PSMs. In this paper, we focus on aberration sensitivity. Aberration is the departure of light ray propagation from Gaussian or paraxial optics. There has been an extensive study on the relationship between aberration and PSMs. For example, spherical aberration is found to degrade the depth of focus (DOF) of semi-randomly aligned patterns printed by alternating PSMs.[7] Coma, on the other hand, causes critical dimension (CD) asymmetry in multiphase PSMs.[8] Although it has been demonstrated that alternating PSMs incur less intrafield linewidth variations than conventional chromium-on-glass (CoG) masks,[9] they generally result in higher image placement error (i.e. the lateral shift of printed features).[10] An effort to design an alternating PSM less susceptible to image placement error is thus necessary.

A linear combination of Zernike polynomials,[11,12] whose coefficients are called Zernike coefficients, is the usual candidate for quantifying aberration. During the fabrication of lenses, Zernike coefficients can be measured by wavefront sensors.[13] Adaptive techniques can be used to correct certain kinds of aberration on the lenses.[14] The goal of these efforts is to reduce the level of lens aberration in individual exposure systems. However,

Further author information: (Send correspondence to Giuseppe Y. Mak)
Giuseppe Y. Mak: E-mail: yhgmak@eee.hku.hk
Alfred K. Wong: E-mail: awong@fortis-systems.com
Edmund Y. Lam: E-mail: elam@eee.hku.hk

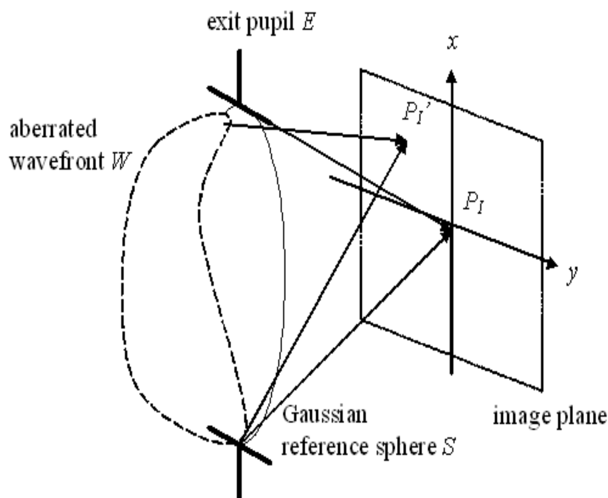

Figure 1. Image formation by light rays from Gaussian reference sphere S and aberrated wavefront W.

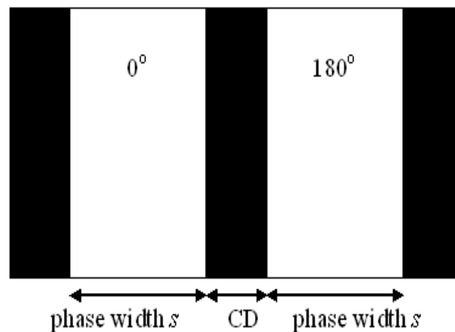

Figure 2. Model of an alternating PSM.

lithographers, as the lens users, are not able to tune the coefficients to suit every kind of circuit pattern. It is a common practice for them to analyze their lithographic requirements and generate some tolerance levels on the RMS aberration.[15] If we are given a certain value of RMS aberration, it remains a question of how the alternating PSM should be designed to minimize the image placment error. The following study aims at obtaining the theoretical optimal design. Coherent imaging is assumed here to simplify the analysis.

2. THEORY

2.1. Wave aberration and RMS aberration

To begin, we shall take a closer look on how aberration is expressed mathematically. Fig. 1 shows the formation of a point image. Without aberration, all transmitted light rays converge to the point image P_I. Taking P_I as the center, a spherical wavefront is formed at the exit pupil E. This wavefront is called Gaussian reference sphere S.

When aberration is present, the light rays no longer converge to a single point on the image plane. The aberrated wavefront W at the exit pupil deviates from the surface S. The optical path difference (OPD) between W and S is known as the wave aberration Φ.[16] Since Φ can be regarded as a surface over the exit pupil E, it is possible to use polynomials to fit this surface. The most common candidates are Zernike polynomials. If Z_i denotes the ith Zernike polynomial (RMS normalized) and C_i the ith Zernike coefficient, then the wave aberration at an arbitrary point (f, g) * on plane E is:[12]

$$\Phi(f, g) = \sum_{i=1}^{n} C_i Z_i(f, g). \tag{1}$$

In Fourier optics, (f, g) are called spatial frequencies and the points on the image plane are called spatial coordinates (x, y).[17] Following the convention of projection lithography, the spatial frequencies are normalized by (NA/λ), while the spatial coordinates are normalized by (λ/NA), where NA is the numerical aperture and λ is the wavelength. We shall denote this by a caret ˆ over the corresponding variables.

*(f, g) are normalized by r/k, where r is the distance from the center of the object to (f, g) and k is the propagation number of the light rays.

RMS aberration R is defined in terms of the Zernike coefficients as:

$$R = \sqrt{\sum_{i=1}^{n} C_i^2}. \tag{2}$$

Our optimization problem is now constraint by a single value of RMS aberration. From Eq. (2), we can see that each value of RMS aberration corresponds to infinitely many possible combinations of Zernike coefficients. Each coefficient is a random variable. They are dependent on one another and their probability density functions (PDFs) are unknown. All these complicate the analysis.

To simplify the problem, we relax the constraint a little bit. First, we assume that each Zernike coefficient C_i has a known PDF. Since there is not much statistical research on how the Zernike coefficients are distributed in different series of exposure systems, each coefficient is modelled as a normal random variable with zero mean and variance ${\sigma_c}^2$, i.e. $C_i \sim N(0, {\sigma_c}^2)$. It implies that the aberration present in various exposure systems is most likely very small. With this assumption, R^2 becomes a χ^2 random variable with n degrees of freedom.[18] Second, we assume that the given value of RMS aberration is the population mean of R (denoted as μ_R). Since μ_R is given by:

$$\mu_R = \sqrt{2}\sigma_c \frac{\Gamma[\frac{1}{2}(n+1)]}{\Gamma(\frac{1}{2}n)} \tag{3}$$

for all n, where $\Gamma(p)$ is the gamma function, σ_c is:

$$\sigma_c = \frac{\mu_R \Gamma(\frac{1}{2}n)}{\sqrt{2}\Gamma[\frac{1}{2}(n+1)]}. \tag{4}$$

2.2. Model of alternating PSM

A simple model of alternating PSM is depicted in Fig. 2. The 0° and 180° openings are called phase regions and their widths $\hat{s}$ are known as phase widths. Between them is a line with a width of critical dimension ($\hat{\text{CD}}$). Here we have chosen phase width as the object of our optimization, because it is an important parameter in optical proximity correction (OPC) and image intensity balance between the 0° and 180° openings.[19] As a simplication we only consider alternating PSMs having equal 0° and 180° phase widths and without phase or transmission error. Under this assumption, the mask spectrum $\tilde{\hat{O}}_x(\hat{f})$[†] of the alternating PSM is a purely imaginary function of $\hat{s}$ and $\hat{\text{CD}}$:

$$\tilde{\hat{O}}_x(\hat{f}) = \frac{i2}{\pi \hat{f}} \sin(\pi \hat{f} \hat{s}) \sin[\pi \hat{f}(\hat{\text{CD}} + \hat{s})]. \tag{5}$$

2.3. Image placement error formula

To facilitate the optimization, it is useful to express the image placement error as a function of the mask spectrum and the wave aberration. Let the placement error be $\Delta\hat{x}$. With reference to Ref. 20, the image shift for a general one-dimensional (1D) mask spectrum under coherent imaging is:

$$\Delta\hat{x} = \frac{\int_{-1}^{1}\int_{\hat{f}_2}^{1} [D_{12}\cos(2\pi\phi_{12}) - S_{12}\sin(2\pi\phi_{12})](\hat{f}_1 - \hat{f}_2)\, d\hat{f}_1\, d\hat{f}_2}{2\pi \int_{-1}^{1}\int_{\hat{f}_2}^{1} [D_{12}\sin(2\pi\phi_{12}) + S_{12}\cos(2\pi\phi_{12})](\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2}, \tag{6}$$

where

$$\begin{aligned}
D_{12} &= \text{Im}\left[\tilde{\hat{O}}_x(\hat{f}_1)\right]\text{Re}\left[\tilde{\hat{O}}_x(\hat{f}_2)\right] - \text{Re}\left[\tilde{\hat{O}}_x(\hat{f}_1)\right]\text{Im}\left[\tilde{\hat{O}}_x(\hat{f}_2)\right], \\
S_{12} &= \text{Re}\left[\tilde{\hat{O}}_x(\hat{f}_1)\right]\text{Re}\left[\tilde{\hat{O}}_x(\hat{f}_2)\right] + \text{Im}\left[\tilde{\hat{O}}_x(\hat{f}_1)\right]\text{Im}\left[\tilde{\hat{O}}_x(\hat{f}_2)\right], \\
\phi_{12} &= \Phi(\hat{f}_1, 0) - \Phi(\hat{f}_2, 0).
\end{aligned}$$

[†]A mask with mask features varying in the $\hat{x}$-direction only can be represented by the mask transmission function $\hat{O}_x(\hat{x})$. This function defines the amplitude and phase of the transmitted light wave at every point of $\hat{x}$. The mask spectrum $\tilde{\hat{O}}_x(\hat{f})$ is the Fourier transform of the mask transmission function (Tilde ˜ signifies a frequency-domain function)

Since the mask spectrum of our alternating PSM is a purely imaginary function, D_{12} in Eq. (6) is reduced to zero. Note that the value of $\Delta\hat{x}$ can be either positive or negative, which represents a shift towards the $+\hat{x}$ or $-\hat{x}$ direction respectively. Since we are only interested in the absolute amount of image placement error in our optimization, we ignore the sign of $\Delta\hat{x}$ and consider $|\Delta\hat{x}|$ instead:

$$|\Delta\hat{x}| = \left| \frac{\int_{-1}^{1}\int_{\hat{f}_2}^{1} S_{12}\sin(2\pi\phi_{12})(\hat{f}_1 - \hat{f}_2)\, d\hat{f}_1\, d\hat{f}_2}{2\pi\int_{-1}^{1}\int_{\hat{f}_2}^{1} S_{12}\cos(2\pi\phi_{12})(\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2} \right|. \tag{7}$$

Alternatively, we can also consider the quantity $(\Delta\hat{x})^2$:

$$(\Delta\hat{x})^2 = \left(\frac{\int_{-1}^{1}\int_{\hat{f}_2}^{1} S_{12}\sin(2\pi\phi_{12})(\hat{f}_1 - \hat{f}_2)\, d\hat{f}_1\, d\hat{f}_2}{2\pi\int_{-1}^{1}\int_{\hat{f}_2}^{1} S_{12}\cos(2\pi\phi_{12})(\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2} \right)^2. \tag{8}$$

This will be further explained in Section 4.

Equipped with Eq. (7) or Eq. (8), if we are given a deterministic wave aberration Φ and $\hat{\mathrm{CD}}$, we can determine the relationship between the image placement error and the phase width easily. However, due to the randomness of Zernike coefficients (as explained in Section 2.1), Φ is now a random function. This implies that the placement error is also a random quantity (hereafter as $\Delta\hat{X}$, with capital letter meaning random variable). A phase width optimized for a particular combination of Zernike coefficients may not be optimal for another combination. This calls for a need to perform the optimization on the average sense. Monte Carlo analysis is our first attempt to find out the relationship between the mean placement error and the phase width, which is described in the next section.

3. MONTE CARLO ANALYSIS ON ZERNIKE COEFFICIENTS

The aim of Monte Carlo analysis on Zernike coefficients is to obtain some preliminary evidence on the existence of global minimum of average placement error at certain phase width. The results underscore the possibility of theoretical analysis in the subsequent sections. Before going into the details, let us first examine how the Zernike coefficients are generated.

Sufficient for the current lithographic applications, we choose the Zernike coefficient set $\{C_5, \ldots, C_{37}\}$ to represent the wave aberration Φ. The mean RMS aberration is taken to be 0.07λ, because this value is the conventional definition of diffraction-limited optics.[15] Many state-of-the-art exposure systems have specifications tighter than this value.[15] Using Eq. (4), the PDF of each coefficient is found to be $N(0, 1.5075 \times 10^{-4}\lambda)$.

Table 1. Parameters in Monte Carlo analysis

wavelength (λ)	248 nm (KrF laser)
numerical aperture (NA)	0.68
CD	$0.3(\lambda/\mathrm{NA})$
mean RMS aberration	0.07λ
number of trials	10000

The parameters used in the simulation are listed in Table 1. In each trial of Monte Carlo analysis, a set of Zernike coefficients is randomly sampled according to the normal distribution just mentioned. By substituting the coefficients (i.e. the wave aberration Φ) into Eq. (7) and Eq. (8), we get the a sample of $|\Delta\hat{X}|$ and a sample of $(\Delta\hat{X})^2$ as functions of $\hat{s}$. After taking 10000 trials in the simulation, the sample means of $|\Delta X|$ and $(\Delta X)^2$ (without normalization by (λ/NA)) are plotted against the phase width in Fig. 3 and Fig. 4. Observing the two plots, we see that there is remarkable similarity in the shape of both plots. They both peak at $\hat{s} = 1(\lambda/\mathrm{NA})$. Besides, both plots have their global minima occurring at $\hat{s} = 0.3(\lambda/\mathrm{NA})$. These results assure us that the optimal phase width is likely to lie between $0.3(\lambda/\mathrm{NA})$ and $0.4(\lambda/\mathrm{NA})$. In Section 4, we return to our image placement error formulae and perform derivations to find the optimal phase width analytically.

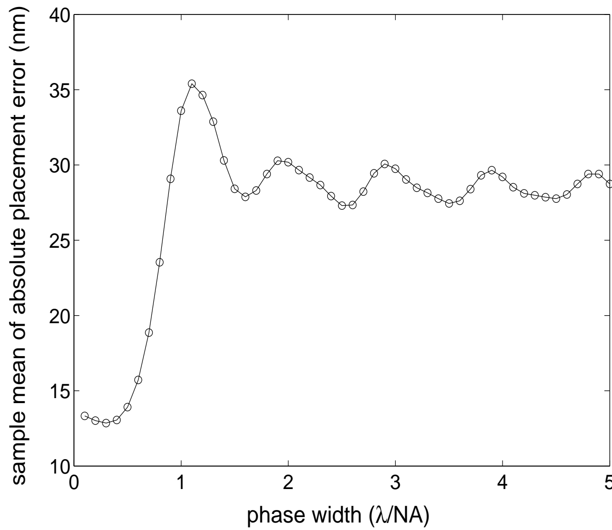

Figure 3. Sample mean of $|\Delta X|$ as a function of $\hat{s}$ for $0.1 \leq \hat{s} \leq 5$. The global minimum occurs at $\hat{s} = 0.3(\lambda/\text{NA})$.

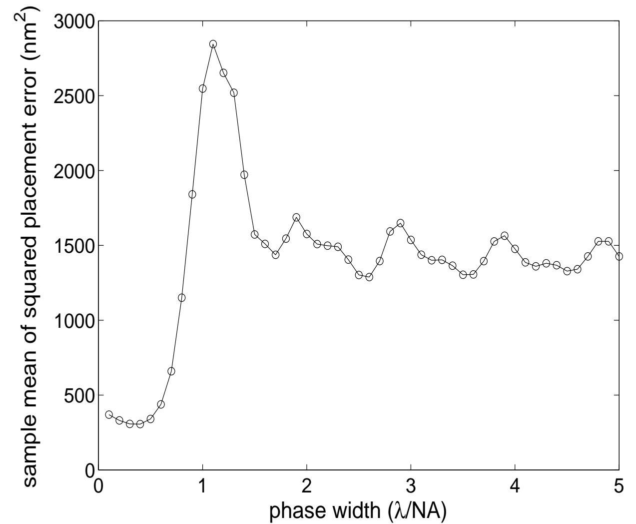

Figure 4. Sample mean of $(\Delta X)^2$ as a function of $\hat{s}$ for $0.1 \leq \hat{s} \leq 5$. The global minimum also occurs at $\hat{s} = 0.3(\lambda/\text{NA})$.

4. EXPECTED PLACEMENT ERROR AND OPTIMAL PHASE WIDTH

The encouraging results from Monte Carlo analysis prompt us to search for a theoretical way in obtaining the optimal phase width. To this end, it is natural for us to consider the expected value of $|\Delta\hat{X}|$ (denoted by $\text{E}[|\Delta\hat{X}|]$, where $\text{E}[\cdot]$ stands for expected value operation). However, without knowing the PDF of $\Delta\hat{X}$, it is a daunting task to determine its expected value. This is because from the mathematical point of view, expectation is an integral but absolute value operation is nonlinear. It is not possible to interchange the order of expectation and absolute value operation, i.e. $\text{E}[|\Delta\hat{X}|]$ is not identical to $|\text{E}[\Delta\hat{X}]|$.

To overcome this problem, we consider the expected value of $(\Delta\hat{X})^2$ instead. It is possible to determine $\text{E}[(\Delta\hat{X})^2]$ *without* knowing the PDF of $\Delta\hat{X}$. The idea is as follows. Referring to Eq. (8), it is noted that both the numerator and the denominator depend on ϕ_{12}, which is the difference of the wave aberration function values at $(\hat{f}_1, 0)$ and $(\hat{f}_2, 0)$. The statistical parameters of ϕ_{12} (e.g. mean, variance, etc.) can be determined from the statistical nature of the Zernike coefficients. If the expectation can be performed on ϕ_{12}, then our problem is mostly solved. The difficulty here is that Eq. (8) is a non-linear operation on ϕ_{12}. With ϕ_{12} in the denominator, the order of expectation and the double integral in the numerator cannot be interchanged. In order to linearize Eq. (8), i.e. to remove the dependence on ϕ_{12} in the denominator, the following empirical approximation has been made to $\cos(2\pi\phi_{12})$ in the denominator of Eq. (8):

$$\begin{aligned} \cos(2\pi\phi_{12}) &\approx \alpha(\hat{f}_1, \hat{f}_2) \\ &= \begin{cases} -0.3/1.3 & 0.95 \leq |\hat{f}_1| \leq 1 \text{ or } 0.95 \leq |\hat{f}_2| \leq 1 \\ 1/1.3 & \text{otherwise.} \end{cases} \end{aligned} \tag{9}$$

The numbers in Eq. (9) are estimated from the sample mean and sample variance of $\cos(2\pi\phi_{12})$ (number of samples = 50000). Their validity is verified by means of Monte Carlo analysis. All the parameters are the same as those in Table 1. We compare the sample mean of $|\Delta\hat{X}|$ obtained from Eq. (7) and its linearized version, as well as the sample mean of $(\Delta\hat{X})^2$ obtained from Eq. (8) and its linearized version. The results are shown in Fig. 5 and Fig. 6. In general, the plots from the linearized equations follow the trend of the plots from the original equations. The match is better for low phase widths ($0 \leq \hat{s} \leq 1$). Although the deviation is larger for higher phase widths, this is still acceptable because the optimal phase width is likely to occur at low values.

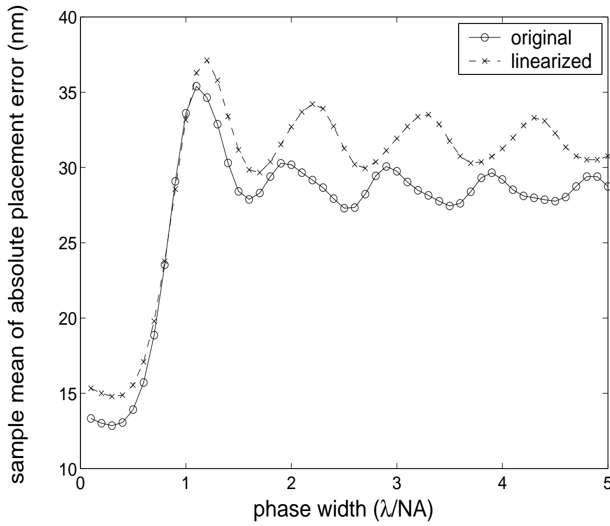

Figure 5. Comparison between the sample means of $|\Delta X|$ obtained from Eq. (7) and its linearized version.

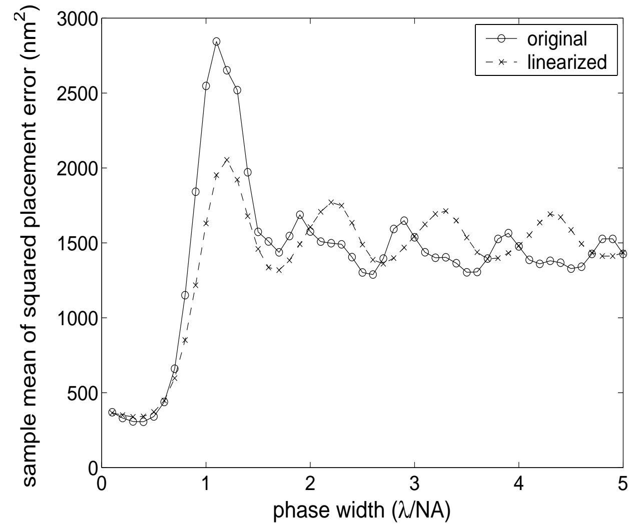

Figure 6. Comparison between the sample means of $(\Delta X)^2$ obtained from Eq. (8) and its linearized version.

With the linearization, $(\Delta\hat{X})^2$ can be expanded in the following manner:

$$\begin{aligned}(\Delta\hat{X})^2 &= \left(\frac{\int_{-1}^{1}\!\int_{\hat{f}_2}^{1} S_{12}\sin(2\pi\phi_{12})(\hat{f}_1-\hat{f}_2)\,d\hat{f}_1\,d\hat{f}_2}{2\pi\int_{-1}^{1}\!\int_{\hat{f}_2}^{1} S_{12}\alpha(\hat{f}_1,\hat{f}_2)(\hat{f}_1-\hat{f}_2)^2\,d\hat{f}_1\,d\hat{f}_2}\right)^2 \\ &= \frac{1}{K}\int_{-1}^{1}\!\int_{\hat{f}_2}^{1}\!\int_{-1}^{1}\!\int_{\hat{f}_4}^{1} S_{12}S_{34}\sin(2\pi\phi_{12})\sin(2\pi\phi_{34})(\hat{f}_1-\hat{f}_2)(\hat{f}_3-\hat{f}_4)\,d\hat{f}_1\,d\hat{f}_2\,d\hat{f}_3\,d\hat{f}_4 \\ &= \frac{1}{K}\int_{-1}^{1}\!\int_{\hat{f}_2}^{1} {S_{12}}^2\sin^2(2\pi\phi_{12})(\hat{f}_1-\hat{f}_2)^2\,d\hat{f}_1\,d\hat{f}_2 \\ &\quad+\frac{1}{2K}\iiiint_G S_{12}S_{34}\left[\cos(2\pi(\phi_{12}-\phi_{34}))-\cos(2\pi(\phi_{12}+\phi_{34}))\right] \\ &\quad\cdot(\hat{f}_1-\hat{f}_2)(\hat{f}_3-\hat{f}_4)\,d\hat{f}_1\,d\hat{f}_2\,d\hat{f}_3\,d\hat{f}_4, \end{aligned} \tag{10}$$

where K is the denominator of the first step of Eq. (10) and G is the region where $\hat{f}_1 \neq \hat{f}_3$ or $\hat{f}_2 \neq \hat{f}_4$. Note that K does not depend on the wave aberration Φ. Now Eq. (10) comprises of two integrals. The first integral is made up of the squared terms, whereas the second integral composes of the cross-multiplied terms during the squaring of $\Delta\hat{X}$.

Let us first consider the expected value of the first integral. Let

$$h(\hat{f}_1,\hat{f}_2,\hat{s};\{C_5,C_6,\ldots,C_{37}\}) = {S_{12}}^2\sin^2(2\pi\phi_{12})(\hat{f}_1-\hat{f}_2)^2 \tag{11}$$

represent the integrand of the first integral. Since the function h is continuous for any real values of $\hat{f}_1,\hat{f}_2,\hat{s}$ and $\{C_5,C_6,\ldots,C_{37}\}$, the order of expectation and the integration is allowed to be interchanged. This is shown as follows:

$$\begin{aligned}\mathrm{E}[\text{1st_integral}] &= \mathrm{E}\left[\int_{-1}^{1}\!\int_{\hat{f}_2}^{1} {S_{12}}^2\sin^2(2\pi\phi_{12})(\hat{f}_1-\hat{f}_2)^2\,d\hat{f}_1\,d\hat{f}_2\right] \\ &= \int_{-1}^{1}\!\int_{\hat{f}_2}^{1} {S_{12}}^2\mathrm{E}[\sin^2(2\pi\phi_{12})](\hat{f}_1-\hat{f}_2)^2\,d\hat{f}_1\,d\hat{f}_2 \\ &= \frac{1}{2}\int_{-1}^{1}\!\int_{\hat{f}_2}^{1} {S_{12}}^2(1-\mathrm{E}[\cos(4\pi\phi_{12})])(\hat{f}_1-\hat{f}_2)^2\,d\hat{f}_1\,d\hat{f}_2. \end{aligned} \tag{12}$$

The expected value of the cosine function in Eq. (12) is:

$$\mathrm{E}[\cos(4\pi\phi_{12})] = e^{-8\pi^2 \mathrm{Var}[\phi_{12}]} \cos(4\pi \mathrm{E}[\phi_{12}]). \tag{13}$$

Since

$$\begin{aligned}
\mathrm{E}[\phi_{12}] &= \mathrm{E}\left[\sum_i C_i(Z_{i1} - Z_{i2})\right] \\
&= \sum_i (Z_{i1} - Z_{i2})\mathrm{E}(C_i) \\
&= 0 \\
\mathrm{Var}[\phi_{12}] &= \mathrm{Var}\left[\sum_i C_i(Z_{i1} - Z_{i2})\right] \\
&= \sum_i (Z_{i1} - Z_{i2})^2 \mathrm{Var}(C_i) \\
&= {\sigma_c}^2 \sum_i (Z_{i1} - Z_{i2})^2,
\end{aligned}$$

where $Z_{ij} = Z_i(f_j, 0)$ for $j = 1, 2$, the expected value of the first integral becomes:

$$E[\text{1st_integral}] = \frac{1}{2}\int_{-1}^{1}\int_{\hat{f}_2}^{1} {S_{12}}^2 \left(1 - e^{-8\pi^2 {\sigma_c}^2 \sum_i (Z_{i1}-Z_{i2})^2}\right)(\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2. \tag{14}$$

Similar to the first integral, the integrand of the second integral is continuous for all independent variables present in it. The order of expectation and integration can thus be interchanged, which is shown as follows:

$$\begin{aligned}
& \mathrm{E}[\text{2nd_integral}] \\
=\ & \mathrm{E}\Big[\frac{1}{2}\iiiint_G S_{12}S_{34}\left[\cos(2\pi(\phi_{12} - \phi_{34})) - \cos(2\pi(\phi_{12} + \phi_{34}))\right](\hat{f}_1 - \hat{f}_2)(\hat{f}_3 - \hat{f}_4)\, d\hat{f}_1\, d\hat{f}_2\, d\hat{f}_3\, d\hat{f}_4\Big] \\
=\ & \frac{1}{2}\iiiint_G S_{12}S_{34}\mathrm{E}\big[\cos(2\pi(\phi_{12} - \phi_{34})) - \cos(2\pi(\phi_{12} + \phi_{34}))\big](\hat{f}_1 - \hat{f}_2)(\hat{f}_3 - \hat{f}_4)\, d\hat{f}_1\, d\hat{f}_2\, d\hat{f}_3\, d\hat{f}_4 \\
=\ & \frac{1}{2}\iiiint_G S_{12}S_{34}\Big\{\mathrm{E}\big[\cos(2\pi(\phi_{12} - \phi_{34}))\big] - \mathrm{E}\big[\cos(2\pi(\phi_{12} + \phi_{34}))\big]\Big\} \\
& \cdot(\hat{f}_1 - \hat{f}_2)(\hat{f}_3 - \hat{f}_4)\, d\hat{f}_1\, d\hat{f}_2\, d\hat{f}_3\, d\hat{f}_4.
\end{aligned} \tag{15}$$

The expected values of the two cosine functions in Eq. (15) are evaluated below:

$$\begin{aligned}
\mathrm{E}[\cos(2\pi(\phi_{12} - \phi_{34}))] &= e^{2\pi^2 \mathrm{Var}(\phi_{12}-\phi_{34})} \cos(2\pi \mathrm{E}[\phi_{12} - \phi_{34}]) \\
\mathrm{E}[\cos(2\pi(\phi_{12} + \phi_{34}))] &= e^{2\pi^2 \mathrm{Var}(\phi_{12}+\phi_{34})} \cos(2\pi \mathrm{E}[\phi_{12} + \phi_{34}]).
\end{aligned} \tag{16}$$

Since

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

$$
\begin{aligned}
&= \sum_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4}) \mathrm{E}(C_i) \\
&= 0 \\
\mathrm{Var}(\phi_{12} - \phi_{34}) &= \mathrm{Var}\left[\sum_i C_i (Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})\right] \\
&= \sum_i (Z_{i1} - Z_{i2} - Z_{i3} + Z_{i4})^2 \mathrm{Var}(C_i) \\
&= {\sigma_c}^2 P_{1234} \\
\mathrm{Var}(\phi_{12} + \phi_{34}) &= \mathrm{Var}\left[\sum_i C_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})\right] \\
&= \sum_i (Z_{i1} - Z_{i2} + Z_{i3} - Z_{i4})^2 \mathrm{Var}(C_i) \\
&= {\sigma_c}^2 Q_{1234}, \qquad (17)
\end{aligned}
$$

where

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

the expected value of the second integral becomes:

$$
E[\text{2nd_integral}] = \iiiint_G S_{12} S_{34} \left[e^{-2\pi^2 {\sigma_c}^2 P_{1234}} - e^{-2\pi^2 {\sigma_c}^2 Q_{1234}}\right] (\hat{f}_1 - \hat{f}_2)(\hat{f}_3 - \hat{f}_4)\, d\hat{f}_1\, d\hat{f}_2\, d\hat{f}_3\, d\hat{f}_4 \qquad (18)
$$

Combining Eqs. (10), (14) and (18), $\mathrm{E}[(\Delta\hat{X})^2]$ is finally given by:

$$
\mathrm{E}[(\Delta\hat{X})^2] = \frac{A + B}{2K}, \qquad (19)
$$

where

$$
\begin{aligned}
A &= \int_{-1}^{1}\int_{\hat{f}_2}^{1} {S_{12}}^2 \left(1 - e^{-8\pi^2 {\sigma_c}^2 \sum_i (Z_{i1} - Z_{i2})^2}\right) (\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2 \\
B &= \iiiint_G S_{12} S_{34} \left[e^{-2\pi^2 {\sigma_c}^2 P_{1234}} - e^{-2\pi^2 {\sigma_c}^2 Q_{1234}}\right] (\hat{f}_1 - \hat{f}_2)(\hat{f}_3 - \hat{f}_4)\, d\hat{f}_1\, d\hat{f}_2\, d\hat{f}_3\, d\hat{f}_4 \\
K &= 4\pi^2 \left[\int_{-1}^{1}\int_{\hat{f}_2}^{1} S_{12} \alpha(\hat{f}_1, \hat{f}_2)(\hat{f}_1 - \hat{f}_2)^2\, d\hat{f}_1\, d\hat{f}_2\right]^2.
\end{aligned}
$$

The essence of Eq. (19) is that $\mathrm{E}[(\Delta\hat{X})^2]$ is a function of two parameters only – the phase width $\hat{s}$ and the variance of the Zernike coefficients ${\sigma_c}^2$:

$$
\mathrm{E}[(\Delta\hat{X})^2] = u(\hat{s}, {\sigma_c}^2). \qquad (20)
$$

This is a very pleasant result because ${\sigma_c}^2$ is related to the mean RMS aberration. In other words, given a certain mean RMS aberration, $\mathrm{E}[(\Delta\hat{X})^2]$ depends on the phase width only.

In Fig. 7, the values of $\mathrm{E}[(\Delta\hat{X})^2]$ computed from the Monte Carlo analysis of Eq. (8) and those computed from Eq. (19) are plotted. A similar plot with the linearized version of Eq. (8) is shown in Fig. 8. This is to show the validity of Eq. (19). By comparing Fig. 7 and Fig. 8, we see that the agreement of the dotted curve and the solid curve in Fig. 8 is better. In both figures, the results show excellent agreement at low phase widths. Since we are only interested in low phase widths, Eq. (19) still maintains its validity.

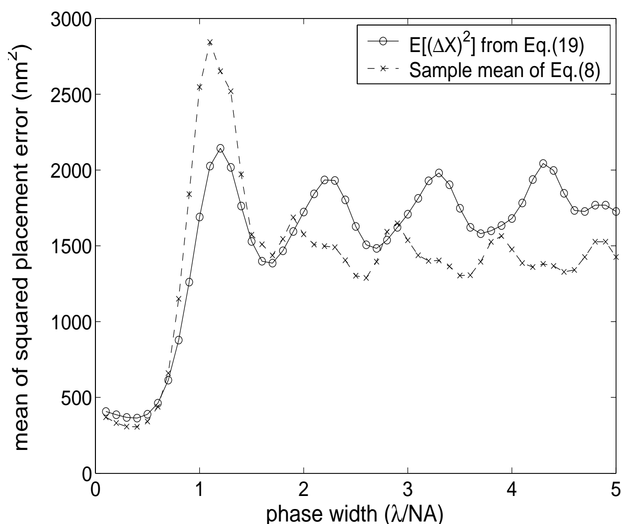

Figure 7. Comparison between the sample mean of $(\Delta X)^2$ obtained from the Monte Carlo analysis of Eq. (8) and $\mathrm{E}[(\Delta\hat{X})^2]$ from Eq. (19).

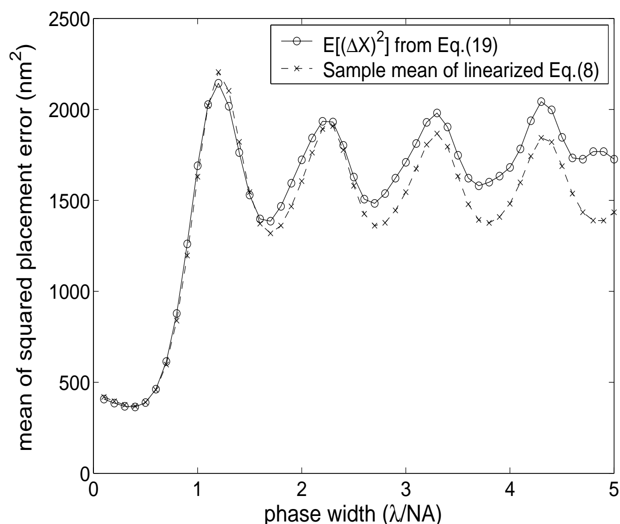

Figure 8. Comparison between the sample mean of $(\Delta X)^2$ obtained from the Monte Carlo analysis of the linearized Eq. (8) and $\mathrm{E}[(\Delta\hat{X})^2]$ from Eq. (19).

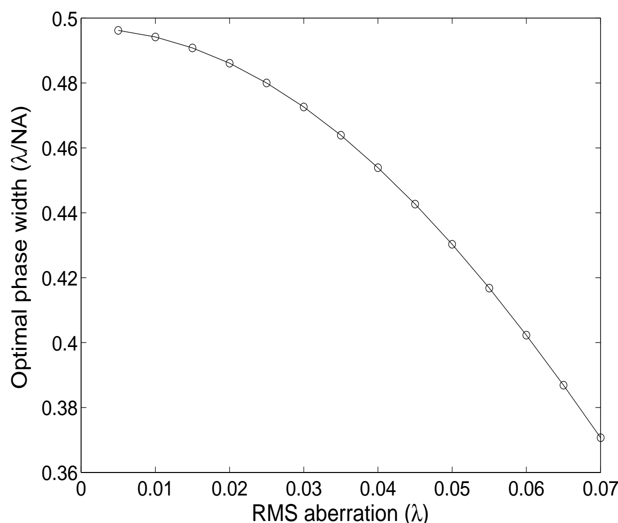

Figure 9. Optimal phase width as a function of RMS aberration. CD $= 0.3(\lambda/\mathrm{NA})$.

Equipped with Eq. (19), we can use numerical methods like method of bisection and Newton's method to determine the position of the global minimum. We do not proceed with the differentiation of Eq. (19) with respect to $\hat{s}$. It is because Eq. (19) is still a complicated function of $\hat{s}$. We have chosen the Golden Section Search method for its quick convergence. The range of search is limited to $0 \leq \hat{s} \leq 1$. After iterations, the optimal phase width is found to be $0.3707(\lambda/\mathrm{NA})$ (i.e. approximately 135 nm) at 0.07λ mean RMS aberration. The optimal phase width as a function of mean RMS aberration is also plotted in Fig. 9. This curve shows that the optimal phase width is between $0.3(\lambda/\mathrm{NA})$ and $0.5(\lambda/\mathrm{NA})$. The value tends to decrease with increasing aberration.

5. DISCUSSION

There are several points worth mentioning in the foregoing analysis. First of all, the applicability of Eq. (19) is not only limited to alternating PSMs. In fact, it can be applied to any masks whose spectra $\hat{\tilde{O}}_x(\hat{f})$ are purely imaginary functions. One example is the enhanced alternating PSM (without phase or transmission error).[19] The phase widths, linewidths and the number of phase regions per edge may be optimized for low aberration sensitivity using Eq. (19).

On the other hand, although coherent imaging is assumed throughout the analysis, Eq. (19) can be extended to partially coherent imaging, which is a more realistic imaging condition. This is possible by considering the image placement error under a general light source.[20] Other aspects that prompt for further study include: different phase widths for 0° and 180° phase regions, different variances for individual Zernike coefficients and the trade-off of other image quality metrics, e.g. exposure latitude. Nevertheless, the methodology detailed in this paper can be followed for further analysis.

6. SUMMARY

Optimization is performed on the phase widths of alternating PSMs. The aim is to minimize the mean image placement error towards aberration under coherent imaging. The constraint is a given value of mean RMS aberration. We first express the image placement error as a function of the mask spectrum and the wave aberration. By randomly generating wave aberrations that conform to our constraint, we perform Monte Carlo analysis to the absolute image placement error $|\Delta\hat{X}|$ and the square of placement error $(\Delta\hat{X})^2$. From the results in Monte Carlo analysis, on average, a global minimum of placement error is likely to occur at a phase width between $0.3(\lambda/\mathrm{NA})$ and $0.4(\lambda/\mathrm{NA})$. By the theoretical consideration of the expected value of $(\Delta\hat{X})^2$, the optimal phase width of alternating PSMs is obtained as a function of the mean RMS aberration. The results are generally applicable to the design of all alternating PSMs.

REFERENCES

1. A. K. Wong, *Resolution Enhancement Techniques in Optical Lithography*, SPIE Press, Bellingham, Washington, U.S.A., 2001.
2. A. Yen, W. N. Partlo, S. R. Palmer, M. A. Hanratty, and M. C. Tipton, “Quarter-micron lithography using a deep-UV stepper with modified illumination,” *Proc. SPIE* **1927**, pp. 158–166, 1993.
3. M. D. Himel, M. K. Poutous, J. D. Stack, and J. L. Leonard, “Microfabrication of controlled angle diffusers used for resolution enhancement in microlithography,” in *Micromachining Technology for Micro-Optics and Nano-Optics*, E. G. Johnson, ed., *Proc. SPIE* **4984**, pp. 230–233, Jan. 2003.
4. M. Levenson, N. Viswanathan, and R. Simpson, “Improving resolution in photolithography with a phase-shifting mask,” in *IEEE Trans. on Electron Devices*, **29**, pp. 1812–1846, Dec. 1982.
5. B. J. Lin, “The attenuated phase-shifting mask,” *Solid State Technology* **35**, pp. 43–47, Jan. 1992.
6. S. Nakao, J. Itoh, A. Nakae, I. Kanai, T. Saitoh, H. Matsubara, K. Tsujita, I. Arimoto, and W. Wakamiya, “Extension of KrF lithography to sub-50-nm pattern formation,” in *Optical Microlithography XIII*, C. J. Progler, ed., *Proc. SPIE* **4000**, pp. 358–365, Jul. 2000.
7. A. Nakae, S. Nakao, and Y. Matsui, “Proposal for pattern layout rule in application of alternating phase-shift mask,” in *Photomask and X-Ray Mask Technology IV*, N. Aizaki, ed., *Proc. SPIE* **3096**, pp. 362–374, Jul. 1997.
8. 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*, L. V. den Hove, ed., *Proc. SPIE* **3334**, pp. 15–24, Jun. 1998.
9. H. Y. Liu, “The application of alternating phase-shifting masks to 140 nm gate patterning: II. Mask design and manufacturing tolerances,” in *Optical Microlithography XI*, L. V. den Hove, ed., *Proc. SPIE* **3334**, pp. 2–14, Jun. 1998.
10. R. J. Socha, W. Conley, X. Shi, M. V. Dusa, J. S. Petersen, J. F. Chen, K. E. Wampler, T. L. Laidig, and R. F. Caldwell, “Resolution enhancement with high-transmission attenuating phase-shift masks,” in *Photomask and X-Ray Mask Technology VI*, H. Morimoto, ed., *Proc. SPIE* **3748**, pp. 290–314, Aug. 1999.

11. V. N. Mahajan, "Zernike circle polynomials and optical aberrations of systems with circular pupils," in *Engineering & Laboratory Notes*, pp. S–21–S–24, Aug. 1994.
12. V. N. Mahajan, "Zernike polynomials and aberration balancing," in *Current Developments in Lens Design and Optical Engineering IV*, P. Z. Mouroulis, W. J. Smith, and R. B. Johnson, eds., *Proc. SPIE* **5173**, pp. 1–17, Nov. 2003.
13. V. Y. Zavalova and A. V. Kudryashov, "Shack-Hartmann wavefront sensor for laser beam analyses," in *High-Resolution Wavefront Control: Methods, Devices, and Applications III*, M. T. G. John D. Gonglewski, Mikhail A. Vorontsov, ed., *Proc. SPIE* **4493**, pp. 277–284, Feb. 2002.
14. T. Kanda and T. Kato, "New-generation projection optics for microlithography," in *Optical Microlithography XIV*, C. J. Progler, ed., *Proc. SPIE* **4346**, pp. 595–605, Sep. 2001.
15. D. C. W. Christopher J. Progler, "Optical lens specifications from the user's perspective," in *Optical Microlithography XI*, L. V. den Hove, ed., *Proc. SPIE* **3334**, pp. 256–268, Jun. 1998.
16. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, New York, U.S.A., 1996.
17. E. Hecht, *Optics*, ch. 11. Addison Wesley, San Francisco, CA, U.S.A., 2002.
18. J. Goodman, *Statistical Optics*, Wiley, New York, U.S.A., 2000.
19. A. K. Wong, "Theoretical discussion on reduced aberration sensitivity of enhanced alternating phase-shifting masks," in *Optical Microlithography XV*, A. Yen, ed., *Proc. SPIE* **4691**, pp. 359–368, Jul. 2002.
20. G. Y. H. Mak, E. Y. Lam, and A. K. Wong, "Placement sensitivity to aberration in optical imaging," in *IEEE Conference on Electron Devices and Solid-State Circuits*, pp. 475–478, Dec. 2003.