

True Process Variation Aware Optical Proximity Correction with Variational Lithography Modeling and Model Calibration*

Peng Yu,[†] Sean X. Shi,[‡] and David Z. Pan[§]

Department of Electrical and Computer Engineering

The University of Texas, Austin, TX 78712

Abstract

Optical proximity correction (OPC) is one of the most widely used Resolution Enhancement Techniques (RET) in nanometer designs to improve subwavelength printability. Conventional model-based OPC assumes nominal process conditions without considering process variations because of the lack of variational lithography models. A simple method to improve OPC results under process variations is to sample multiple process conditions across the process window, which requires long runtime. We derive a variational lithography model (VLIM) which can simulate across the process window without much runtime overhead compared to the conventional lithography models. To match the model to experimental data, we demonstrate VLIM calibration method. The calibrated model has accuracy comparable to non-variational models, but it has the advantage of taking process variations into consideration. We introduce the variational edge placement error (V-EPE) metrics based on the model, a natural extension to the edge placement error (EPE) used in conventional OPC algorithms. A *true* process-variation aware OPC (PV-OPC) framework is proposed using the V-EPE metric. Due to the analytical nature of VLIM, our PV-OPC is only about 2-3× slower than the conventional OPC, but it *explicitly* considers the two main sources of process variations (exposure dose and focus variations) during OPC. Thus our post PV-OPC results are much more robust than the conventional OPC ones, in terms of both geometric printability and electrical characterization under process variations.

PACS numbers:

Keywords: Optical proximity correction, resolution enhancement technique, process variation, variational lithography modeling, exposure dose, defocus, variational edge placement error

* The preliminary version of this work appeared in [1].

[†]Electronic address: yupeng@cerc.utexas.edu

[‡]Electronic address: sean.shi@mail.utexas.edu

[§]Electronic address: dpan@ece.utexas.edu

I. INTRODUCTION

As the semiconductor process technology scales into nanometer dimension, the printability and process window of the finer lithographic patterns are significantly reduced due to the fundamental limit of the microlithography systems and process variations. As for now, leading IC fabs still use the 193nm lithography systems to print sub-wavelength feature size (e.g., 65nm or even 45nm), with the aid of various and sometimes exotic tricks so called resolution enhancement techniques (RET), such as optical proximity correction (OPC), phase shift mask (PSM), off-axis illumination (OAI) and sub-resolution assist feature (SRAF) insertion. These techniques modify illuminations, mask patterns, or transmissions. Since the 157nm lithography and other next generation lithography (NGL) systems are not likely to be in the mainstream in the near future [2–5], it is expected that more and more extensive RETs will be used to push the lithography systems to their limits.

The fundamental limits of optical lithography system, summarized in [3, 6–8], refer to the achievable lithography system bounds, e.g., resolution in terms of pitch and critical dimension - as shown in Figure 1, depth of focus. Figure 2 shows a sketch of a typical optical lithography system.

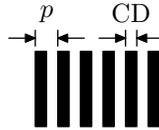


Figure 1: A periodic line/space pattern. The pitch p denotes its period in space. Critical Dimension (CD) refers to the line width in this case.

The mask is illuminated by the light source through the illumination lens. An image of the mask is formed in the photoresist through the projection lens. λ is the wavelength of the light source. n is the index of refraction of the medium between the lens and the photoresist. For 193nm water immersion lithography, $n = 1.44$; for air, $n \approx 1$. θ is the half-angle of the maximum cone of light that exits the lens.

The optimal illumination scheme for the line/space pattern is a dipole, a type of OAI. The minimum printable pitch using this illumination is given by

$$p_{\min} = \frac{\lambda}{2n \sin \theta}, \quad (1)$$

where $n \sin \theta$ is defined as the *numerical aperture* (NA). In general, for any pattern (including

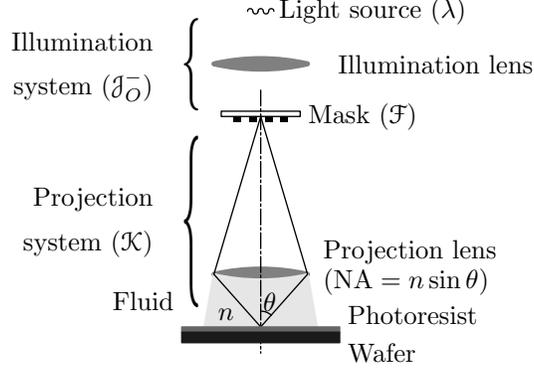


Figure 2: Optical lithography system. The mask is illuminated by a laser light source with wavelength λ . The photoresist is exposed through the projection system. The latent image is formed in the photoresist. The minimum pitch, in a broader sense, has the following form

$$p_{\min} = k_1 \frac{\lambda}{n \sin \theta}, \quad (2)$$

where k_1 is an OAI and pattern dependent factor, and its physical minimum is $1/2$.

Depth-of-Focus (DOF), indicating the pattern robustness with respect to focus variation, is defined as the range of focus that keeps the resulting printed feature within a variety of specifications (such as line width, sidewall angle, photoresist loss, and exposure latitude). It is estimated as

$$\text{DOF} = \frac{n\lambda}{2(1 - \cos \theta)}. \quad (3)$$

To print smaller pitch at a given wavelength λ , we can increase θ and n . However, unless higher n material is available to replace water, DOF will continue to decrease as θ goes to its physical limit 90° . Therefore, more CD variations are expected due to focus variations. Meanwhile, the exposure dose variation impact will also become more severe, since the defocused image qualities are degraded more at the same defocus value.

OPC is one of the most widely used RETs by simply modifying the mask patterns to improve the printability and image robustness. OPC algorithms can be classified as *sparse* OPC and *dense* OPC [9–11]. These two categories are mainly different in computing the photoresist profile either on sparsely sampled sites or on dense grids and in treating mask geometries as polygons or pixel-based images. Sparse OPC is the current dominant OPC methodology, while the dense OPC is gaining more interest. Cobb did pioneering work on sparse OPC [12–14]. Granik et al formulated OPC problem into a more rigorous framework based on Mask Error Enhancement Factor (MEEF) theory [15, 16]. Conventional model-based OPC assumes nominal process parameters [12–14].

As process variations become more important, OPC software should not disregard them any more. Some primitive attempts have been made in this direction. For example, it is pointed out that the expected contour should be on target [17]. However, no implementation details are provided in this paper. Defocus aerial images, instead of in-focus aerial images, can be used to improve process window robustness [18, 19]. But they rely on extensive lithography simulations to choose the appropriate defocus value, which is very expensive. It is shown in [20] how to modify the OPC algorithm to consider the expected defocus from CMP-induced wafer topography. But again, it is based on a certain defocus condition, without considering focus variations and dose variations. Image-log slope, as an indicator of process sensitivity to dose variations, has been used in [18, 21]. But this approach is incapable of handling focus variations. None of these attempts are aware of the entire process window during OPC. The reason is due to prohibitive runtimes of lithography simulations across the entire process window. Actually, even without considering process variations, it has been reported that model-based OPC software could run for days on multiple computers for a single design [22].

Ignoring OPC impacts or process variations could lead to erroneous timing, power and yield characterization analysis. For example, post-OPC silicon image based timing analysis is substantially different from that based on the drawn layout, e.g., with 36% increase in worst-case slack and significant critical path reordering [23]. Their analysis is based on OPC with the nominal process. It is expected that the difference in consideration of process variations would be even more [24]. Statistical simulation techniques are demonstrated to map the lithography variability to CD or chip timing [25, 26]. The awareness of across chip line width variations can account for as much as 40% tightening of the best-case to worst-case timing spread [27]. Post-OPC gate non-rectangularity should not be ignored to estimate timing and leakage more accurately [28]. Therefore, it is important to make the OPC aware of the process variations.

In this paper, we propose a *true process-variation aware OPC* (PV-OPC) framework. Our implementation is based on a sparse OPC algorithm, but the general principle can be applied to dense OPC as well. To the best of our knowledge, this is the first work to explicitly address the true process-variation awareness during OPC. Our PV-OPC is enabled by the *variational lithography modeling* and guided by the *variational edge placement error* (V-EPE) metrics. The main contributions of this paper are as follows.

- We derive a new analytical variational lithography model, which is generic to handle any focus variation and illumination scheme.

- We provide a variational model calibration method.
- We introduce the concept and obtain the close-form formulae for the variational EPE metrics, and use them to guide our PV-OPC algorithm with explicit consideration of the two main sources of process variations (exposure dose and focus variations).
- The robustness of the PV-OPC algorithm is demonstrated in terms of both the geometrical and the electrical characterizations compared to the conventional OPC.
- The runtime of the PV-OPC algorithm is only about $2\text{-}3\times$ that of the conventional OPC due to the analytical nature of our models, so it is feasible to be used in practice.

The rest of the paper is organized as follows. In Section II, we review the lithography fundamentals such as sources of variations and lithography modeling. In Section III, we derive the analytical variational lithography model and present the variational model calibration method, followed by the fast table lookup simulation method. In Section IV we derive the closed-form formulae for the variational EPE metrics. Section V presents the overall algorithm and implementation details of the PV-OPC algorithm. Section VI shows the experimental results, followed by the conclusions in section VII.

II. LITHOGRAPHY BACKGROUND

In this section, we first classify the variations in lithography systems. Then, we review the conventional non-variational lithography modeling, from which our new variation lithography model (VLIM) is derived (Section III).

A. Variations in Lithography System

The term “variation” can refer to the raw process variations, or the derived geometrical and electrical variations, i.e.,

- the distributions of the raw process parameters;
- how severe the print images or circuit parameters (e.g., power and frequency) change due to certain amounts of the process parameter changes.

The goal for PV-OPC is that based on the raw process variations (e.g., dosage and focus), the post PV-OPC image would have good property in terms of derived geometrical or electrical characteristics (e.g., less variations).

There are many manufacturing parameters, e.g., focus error, exposure dose, wavelength (λ), polarization. Great efforts have been made to control lithography system uniformity (over space) and stability (over time). Three kinds of lithography variation sources, dose, focus and mask variations are believed to be among most important [29]. Chen et al. rigorously related the Mask Error Enhancement Factor (MEEF) to the image log slope [30]. Because image log slope and MEEF indicate CD sensitivities to exposure dose variation and mask size variation respectively, mask size variation can be equivalently treated as exposure dose variation. Thus we only consider one of these two variations — exposure dose variation. Actually our experiments show that CD is approximately linearly related to exposure dose variation and quadratically related to focus variation. Assuming higher order terms can be ignored, other first and second order parameters can be made equivalent to either exposure dose and focus error. Therefore, we will focus on the exposure dose and focus variations in this paper.

B. Lithography Modeling Review

Various lithography models have been developed for lithography system simulation. Based on the details of the physics descriptions, these models in general can be classified into two categories — physics based models and phenomenological models [31–34].

Usually, the physics based models are slower but more accurate than the phenomenological ones. Because of the modeling of fundamental physics and chemistry in lithography systems, physics based models (eg. PROLITHTM and Solid-ETM) can tell the consequences of the process parameter alternations, which help the process engineers to fine tune the lithography processes. However, it is difficult to calibrate these models because of their complexities and the difficulties in measuring the model parameters. Phenomenological models, on the other hand, seek for simulation speed and reasonable accuracy. They do not model all the physics and chemistry in the lithography and only work in a very limited domain of the process parameter space. But, it is relatively easy to fit them to the experiments because of their simplicities.

OPC requires fast lithography simulation with reasonable accuracy, where the phenomenological models are the best candidates. We review the phenomenological lithography modeling, which

can be divided into three steps:

- Photomask patterns are transformed to chemical latent image in the photoresist bulk through the optical system by exposing the photomask.
- The chemical latent image is diffused in the post exposure bake (PEB) step. Chemical reactions are taken place in the photoresist.
- The photoresist development results in a 3-dimensional photoresist profiles. However, only 2-dimensional photoresist contour is of interest to the phenomenological models.

The details are discussed in the following.

1. Aerial Image to Latent Image

An aerial image, by definition, is a projected image which is "floating in air". In lithography, it usually refers to the image on top of the photoresist or in some plane in the photoresist bulk. It is described by *Hopkins Equation* [35], written as

$$\begin{aligned} \mathcal{J}_A(f, g) = & \iint \mathcal{T}(f' + f, g' + g; f', g') \\ & \times \mathcal{F}(f' + f, g' + g) \mathcal{F}^*(f', g') df' dg'. \end{aligned} \quad (4)$$

$\mathcal{F}(f, g)$ is the mask transmission function $F(x, y)$ in the Fourier domain. $\mathcal{J}_A(f, g)$ means the aerial image transformed into Fourier domain. (f, g) denotes a point in the Fourier domain. $\mathcal{T}(f, g; f', g')$ is called the *transmission cross coefficient* (TCC), given by

$$\begin{aligned} \mathcal{T}(f', g'; f'', g'') = & \iint \mathcal{J}_O^-(f, g) \mathcal{K}(f + f', g + g') \\ & \times \mathcal{K}^*(f + f'', g + g'') df dg, \end{aligned} \quad (5)$$

where $\mathcal{J}_O^-(f, g)$ and $\mathcal{K}(f, g)$ is the illumination function and the projection system transfer function, respectively. The superscript * means the complex conjugation.

$\mathcal{K}(f, g)$ can be written as

$$\mathcal{K}(f, g) = \mathcal{K}_0(f, g) e^{i2\pi z \phi(f, g)}, \quad (6)$$

where z denotes the focal error, which is dimensionless and normalized by the factor $\frac{\lambda}{\text{NA}^2}$, and

$$\phi(f, g) = \sqrt{1 - f^2 - g^2} \quad (7)$$

is the phase factor. $\mathcal{K}_0(f, g)$ is a factor related to the pupil. For a circle pupil,

$$\mathcal{K}_0(f, g) = \begin{cases} 1 & f^2 + g^2 < 1 \\ 0 & \text{otherwise} \end{cases}. \quad (8)$$

For the conventional illumination with *partially coherent factor* s , $\mathcal{J}_0^-(f, g)$ is written as

$$\mathcal{J}_0^-(f, g) = \begin{cases} \frac{1}{\pi s^2} & f^2 + g^2 < s^2 \\ 0 & \text{otherwise} \end{cases}. \quad (9)$$

Other illuminations can be described similarly.

Pre-PEB latent image is formed based on the aerial image. Some works have been done on the thin film effect [36, 37]. However, to the first order of accuracy, we assume the pre-PEB latent image is the same as the aerial image. High-NA [38] effects and lens aberrations [39] can also be included in the future.

2. Resist Blur

The diffusion process is modeled by convolving the pre-PEB latent image with a function, called the blur function. A number of blur functions have been proposed [40–45]. We use a Gaussian blur function

$$G(x, y) = \frac{1}{\sqrt{2\pi}d} e^{-\frac{x^2}{2d^2}}, \quad (10)$$

where d is the diffusion length. The Fourier transform of this function is

$$\mathcal{G}(f, g) = e^{-2\pi^2 d^2 (f^2 + g^2)}. \quad (11)$$

Because the diffusion operator is a multiplication operator in the Fourier domain, the Fourier transform of the diffused latent image can be written as

$$\mathcal{J}_G(f, g) = \mathcal{G}(f, g) \mathcal{J}_A(f, g). \quad (12)$$

Define the diffused TCC

$$\mathcal{T}_G(f', g'; f'', g'') = \mathcal{G}(f' - f'', g' - g'') \mathcal{T}(f', g'; f'', g''). \quad (13)$$

The post-PEB latent image equation (12), written as

$$\begin{aligned} \mathcal{J}_G(f, g) = & \iint \mathcal{T}_G(f' + f, g' + g; f', g') \\ & \times \mathcal{F}(f' + f, g' + g) \mathcal{F}^*(f', g') df' dg', \end{aligned} \quad (14)$$

is similar to that of (4).

3. Threshold Bias Model

Photoresist development model is to predict the print contour based on the latent image. We use the threshold bias photoresist model due to its simple analytical formulation. It has also been demonstrated in [44] that it predicts CD fairly accurately. This model assumes the printed contour can be computed by applying a constant bias (B) to the contour where the intensity is equal to an intensity threshold I_{th} (Figure 3). B and I_{th} are parameters subject to calibration (see Section III B).

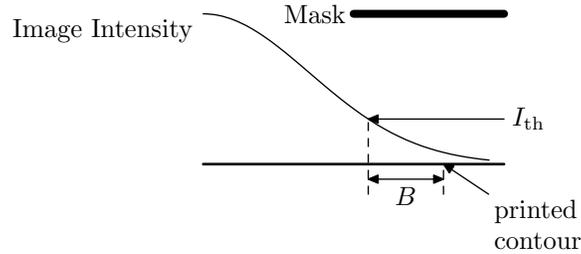


Figure 3: Threshold bias model. The image intensity is denoted by the curve. A constant bias B is applied the location where the intensity equals to the intensity threshold I_{th} to get the printed contour.

III. VARIATIONAL LITHOGRAPHY MODEL (VLIM)

In the previous section, we reviewed the conventional phenomenological lithography models, which require per focus error (z) simulation. To simulate through the range of focus variations, one naive solution is to simulate at many different z values. But there are the simulation speed problem and the model consistency issue.

- The simulation time is proportional to the number of discrete z values, which is unaffordable for full chip simulation when the number is big.

- Conventional model calibration methods only work for a single process condition. Models of different z values need separate calibrations. Since measurement errors are unavoidable, two models calibrated at slightly different process conditions might have dramatic difference on the simulation results.

We address the first problem by introducing a new variational lithography model (VLIM) in Section III A. The concepts of calibrating the models across the process window has been proposed [46, 47]. But no details on the calibration methods were revealed. To solve the second problem, Section III B provides a variational lithography model calibration method in details. Section III C discusses the fast image simulation for rectilinear polygons using the vertex-based table lookup method.

A. VLIM Derivation

We have mentioned exposure dose variation and focus variation are the two most important variations in lithography system. The exposure dose variation can be transformed to equivalent intensity threshold I_{th} variation, which is easy to be handled. However, the conventional model can not handle focus variation efficiently. We introduce VLIM to solve this problem. In particular, we derive an analytical formula for the defocus latent image for any illumination schemes, by adapting and extending the method used in [48] (which only handles the fully coherent illumination).

Expand e^x as $\sum_{n=0}^{\infty} \frac{x^n}{n!}$, plug it in (13) and use Binomial Expansion, we have (15)

$$\begin{aligned} \mathcal{T}_G(f', g'; f'', g'') &= \mathfrak{G}(f' - f'', g' - g'') \sum_{n=0}^{\infty} \frac{(i2\pi z)^n}{n!} \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \iint \mathcal{J}_O^-(f, g) \\ &\quad \times \left(\phi(f + f', g + g') \right)^k \mathcal{K}_0(f + f', g + g') \\ &\quad \times \left(\phi(f + f'', g + g'') \right)^{n-k} \mathcal{K}_0^*(f + f'', g + g'') df dg. \end{aligned} \quad (15)$$

That is, $\mathcal{T}_G(f', g'; f'', g'')$ can be expanded as

$$\mathcal{T}_G(f', g'; f'', g'') = \sum_{n=0}^{\infty} z^n \mathcal{T}_{G_n}(f', g'; f'', g''). \quad (16)$$

Plug (16) into (14), we end up with the following form

$$\mathcal{J}_G(f, g) = \sum_{n=0}^{\infty} z^n \mathcal{J}_{G_n}(f, g), \quad (17)$$

where

$$\begin{aligned} \mathcal{J}_{G_n}(f, g) &= \iint \mathcal{T}_{G_n}(f' + f, g' + g; f', g') \\ &\quad \times \mathcal{F}(f' + f, g' + g) \mathcal{F}^*(f', g') df' dg'. \end{aligned} \quad (18)$$

Fourier transform both sides of (17), we reach the expansion form of the latent image intensity

$$I_G(x, y) = \sum_{n=0}^{\infty} z^n I_{G_n}(x, y). \quad (19)$$

For binary mask or PSM with phase 0° and 180° (the mask transmission function $F(x, y)$ is always real), it can be proved that all the odd terms in (19) are equal to zeros based on the derivations similar to those in [49]. Then we have

$$I_G(x, y) = \sum_{n=0}^{\infty} z^{2n} I_{G_{2n}}(x, y). \quad (20)$$

We call the above equation the *defocus latent image expansion*. $I_{G_n}(x, y)$'s are called the variational latent images. It is easy to see I_0 is the in-focus ($z = 0$) latent image. The above equation tells us that the defocus latent image can be expressed as the in-focus image plus some correction terms. When z is small and z^n ($n \geq 4$) is much smaller than z^2 , the higher order terms can be ignored. We get the analytical formula,

$$I_G(x, y) \cong I_{G_0}(x, y) + z^2 I_{G_2}(x, y). \quad (21)$$

We simulated images with the conventional partially coherent illumination ($s = 0.7$), the wave length $\lambda = 193\text{nm}$ and the numerical aperture $\text{NA} = 0.8$ for a Five Bar pattern. Figure 4 shows I_G vs. z curves of five randomly chosen locations. We can see that the curves are approximately parabolas in $(-200\text{ nm}, 200\text{ nm})$. So (21) holds in this range.

A more rigorous way of deciding the range of z where the approximation (21) holds is to compare the magnitude z^2 term and the summation of all higher order terms. Figure 5 shows $I_G|_{z=z_0} - I_G|_{z=0}$, $I_{G_2}z_0^2$, $I_{G_4}z_0^4$, $I_{G_2}z_1^2$ and $I_{G_4}z_1^4$ of the same Five Bar pattern ($z_0 = 100\text{ nm}$, $z_1 = 200\text{ nm}$). $I_{G_4}z_0^4$ can be ignored because it is much smaller than $I_{G_2}z_0^2$ ($z_0 = 100\text{ nm}$). For $z_1 = 200\text{ nm}$, $I_{G_2}z_1^2$ is still about 5 times $I_{G_4}z_1^4$. Let us say the criterion is $I_{G_4}z_1^4$ can be ignore if it is smaller than one fifth of $I_{G_2}z_1^2$. Then the approximation in (21) is appropriate within $\pm 200\text{ nm}$. Typical lithography simulation shows that this property holds well in a few hundred nm (bigger than the typical defocus range in IC manufacturing).

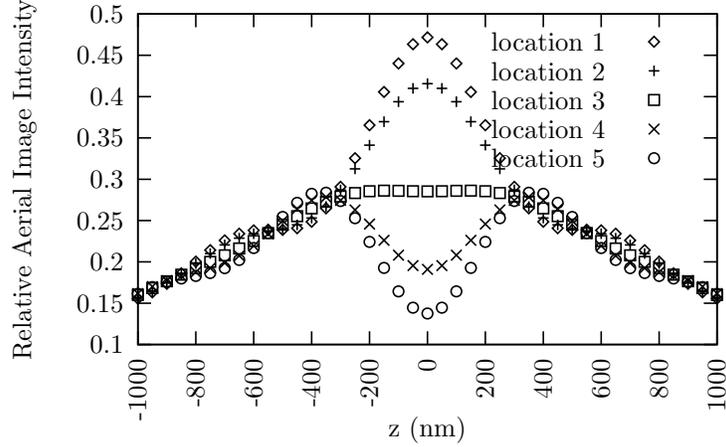


Figure 4: Aerial image intensity simulation results (PROLITH™) at 5 randomly chosen locations.

B. VLIM Calibration

Lithography systems are very complex. For example, PROLITH™ has pages of physics-based input parameters in its manual [50]. Some parameters in variable threshold resist (VTR) model family are merely fitting parameters and have no physical meaning [51, 52]. In practice, we only care about the consistencies between the model prediction and the experimental data, especially for OPC softwares. Therefore, instead of measuring each individual parameters, the model parameters are usually fitted to match the experiment. In this section, we show a VLIM calibration method.

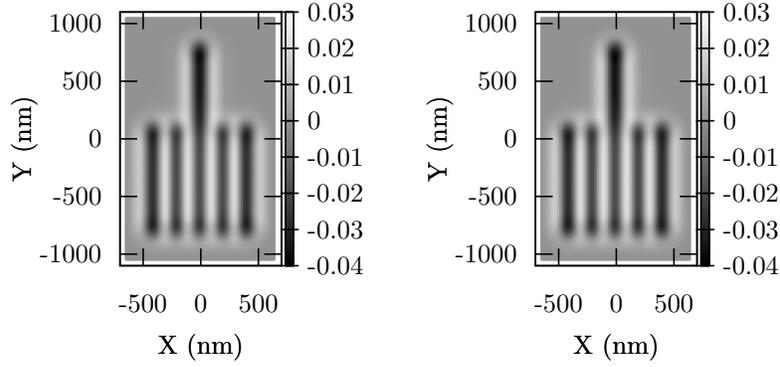
We denote the experimentally measured exposure dose and focus error as E and Z . We fit VLIM to the experimental data using nonlinear regression method. Hence, we can determine the four input parameters of VLIM, the intensity threshold I_{th} , the focus error z , the diffusion length d and the constant bias B .

Assuming the photoresist has a threshold behavior, the intensity threshold I_{th} is inversely proportional to exposure dose E [53]. Taking into account the offset in the exposure dose measurement,

$$I_{th} = \frac{\beta_E}{E - \alpha_E} \quad (22)$$

describes the relationship between E and I_{th} [43, 44, 54]. Because the refraction index of the photoresist film is not 1, the focus error should be scaled by a factor β_Z . α_Z represents the focus measurement offset. So we have

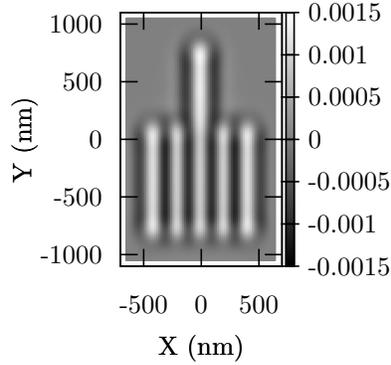
$$z = \beta_Z(Z - \alpha_Z). \quad (23)$$



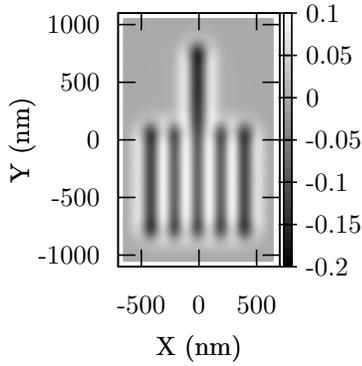
(a) $I_G|_{z=z_0} - I_G|_{z=0}$ image

(b) $I_{G_2 z_0^2}$ image intensity map

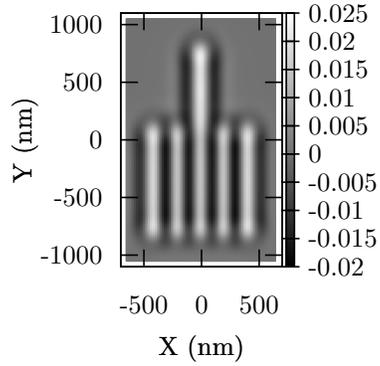
intensity map



(c) $I_{G_4 z_0^4}$ image intensity map



(d) $I_{G_2 z_1^2}$ image intensity map



(e) $I_{G_4 z_1^4}$ image intensity map

Figure 5: $I_G|_{z=z_0} - I_G|_{z=0}$ and $I_{G_2 z_0^2}$ are almost the same. $I_{G_2 z_0^2}$ is about 20 times $I_{G_4 z_0^4}$. $I_{G_2 z_1^2}$ is about 5 times $I_{G_4 z_1^4}$. ($z_0 = 100$ nm and $z_1 = 200$ nm)

Those relations are used in the model fitting method.

Since the CD responds differently to process variations for different mask pattern P , many patterns should be measured in the experiment. In order to predict the CD across process window, we need to take measurements at various process conditions as well. Suppose the $CD_{P,i}$ is the CD measure value at exposure dose E_i and focus error Z_i for mask pattern P . We can estimate the parameters $d, B, \alpha_E, \beta_E, \alpha_Z$ and β_Z by minimizing

$$\chi^2(d, B, \alpha_E, \beta_E, \alpha_Z, \beta_Z) \equiv \sum_P \sum_i \left(CD_{\text{VLIM}}^P \left(\frac{\beta_E}{E_i - \alpha_E}, \beta_Z(Z_i - \alpha_Z), d \right) - 2B - CD_{P,i} \right)^2, \quad (24)$$

where $CD_{\text{VLIM}}^P(I_{\text{th}}, z, d)$ is CD function based on VLIM for the mask pattern P .

To solve the minimization problem in (24), we require that CD_{VLIM}^P 's be analytical functions. Appendix shows how to generate the analytical functions based on VLIM simulation results.

We estimate the fitting error by computing the standard deviation

$$\sigma = \sqrt{\frac{\chi^2}{N - M}}, \quad (25)$$

where N is the number of data points and M is the number of fitting parameters (6 in this case). An example of VLIM calibration is shown Section VI.

C. Kernel Decomposition and Vertex Based Table-Lookup

In this section, we show how to compute I_{G0} and I_{G2} by the table lookup method.

Decomposing $\mathcal{T}_n(f', g'; f'', g'')$ (16) and ignoring the residue terms, we have

$$\mathcal{T}_n(f', g'; f'', g'') = \sum_{k=1}^p \sigma_{nk} Q_{nk}(f', g') Q_{nk}^*(f'', g''), \quad (26)$$

where σ_{nk} 's are real numbers. After some simple manipulations, we have

$$I_{Gn}(x, y) = \sum_{k=1}^p \sigma_{nk} |Q_{nk} ** F|^2, \quad (27)$$

where $**$ is the convolution operator and $Q_{nk}(x, y)$ (called the kernel) is the decomposition of \mathcal{T}_n in the spatial domain. Then each variational latent image can be computed by convoluting the mask transmission function with a few kernel functions. The table generation method is similar to [55]. Algorithm 1 shows how to generate the lookup tables.

Algorithm 1 Kernel decomposition and table generation

Input: Lithography Optics System Functions — the illumination function $\mathcal{J}_O^-(f, g)$, the projection system transfer function $\mathcal{K}(f, g)$ and the diffusion kernel $\mathcal{G}(f, g)$

Output: Generate the VLIM lookup tables

- 1: Compute \mathcal{T}_n ($n = 0, 2$) based on (15) and (16)
 - 2: Decompose \mathcal{T}_n into \mathcal{Q}_{nk} terms in (26)
 - 3: Fourier transform \mathcal{Q}_{nk} into Q_{nk} (27)
 - 4: Compute all the upper-right rectangle convolutions from Q_{nk} and write them into table format
-

We propose a vertex based convolution method instead of the rectangle based method used in [55], because it requires less number of table lookup times as shown in the example below. Algorithm 2 shows how to compute the variational latent images from the lookup tables. The region where $Q_{nk}(x, y)$ is non-zero is called the support region whose size is about a few times of $\frac{\lambda(1+s)}{\text{NA}}$.

Algorithm 2 Vertex based table lookup

Input: The tables generated in Algorithm 1, the decomposed mask in the form of polygons

Output: Compute the variational aerial images I_{G0} and I_{G2} at the point \mathcal{D}

- 1: Retrieve the polygons which intersect the \mathcal{D} 's support region \mathfrak{R}
 - 2: Compute the vertices of the polygons with non-zero convolution value
 - 3: Compute the sign of the convolution of each vertex
 - 4: Look up the tables to get the convolution values
 - 5: Compute I_{G0} and I_{G2} at point \mathcal{D} by summing up these values
-

To compute the convolution $Q_{nk} ** F$ in (27) at point \mathcal{D} , we need only the mask shapes overlapping Q_{nk} 's support region \mathfrak{R} , which in turn can be expressed as the summation of the convolution values of each mask shape (Figure 6).

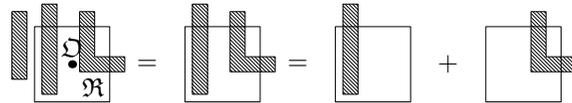


Figure 6: Mask truncation and decomposition.

For rectilinear polygons, we can compute the convolutions based on the vertices. As shown in Figure 7, any rectilinear polygon convolution can be decomposed into the summation of the

convolutions of the regions to the upper-right of each vertex. We store the convolutions of all the upper-right rectangles within the support region in a lookup table for each kernel.

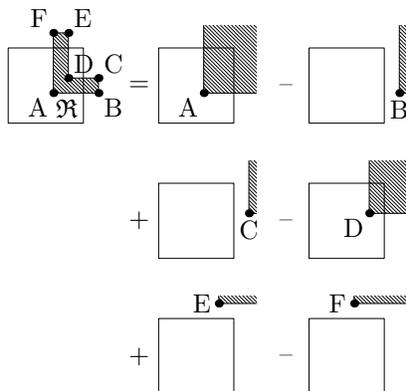


Figure 7: Vertex based rectilinear polygon convolution.

For the example in Figure 8, the contributions of B, C, E and F are zero. Only A's and D's convolutions are needed. If the method in [55] is used, four table lookups will be needed. It is clear vertex based convolution method is much better.

Both I_{G_0} and I_{G_2} can be computed by this method. So the runtime of computing variational latent images (I_{G_0} and I_{G_2}) will be in the same order as the runtime for the non-variational latent image.

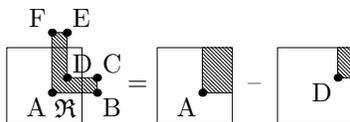


Figure 8: Look-up tables store the convolutions of all the upper-right rectangles within the support region. Convolutions with zero-contribution will not be stored.

IV. VARIATIONAL EPE (V-EPE) METRICS

For any given point on the target contour, we define its Edge Placement Error (EPE) as the displacement between that point and its nearest printed contour point (Figure 9). Note that the EPE defined here is a vector, which is slightly different from the conventional scalar EPE definition in [14] and our previous work [1]. We denoted the EPE for any point A on the target as E_A .

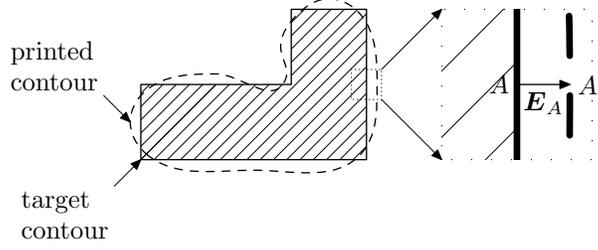


Figure 9: EPE concept. The target and printed contours are solid and dashed respectively. The right subfigure is a zoom-in subregion of the left one. A is a point on the target contour. A' is A 's closest point on the printed contour. The EPE for A is shown as $E_A = \overline{AA'}$.

The printed contour is uniquely determined if EPEs of all the target contour points are given.

Conventionally, people do not consider how EPE varies resulting from process variations. In this section, we derive the analytical variational EPE metrics (V-EPE) based on VLIM to describe this variation. Our process variation aware OPC will be based on the V-EPE metrics, while the conventional OPC is based on the nominal EPE.

A. Variational EPE Model

We set the bias $B = 0$ to simplify the V-EPE Model derivation, because it is easy to adjust E_A if $B \neq 0$. We drop the subscript A in E_A and the subscript G in I_G to simplify the notation in the following derivations. Since there is a one-to-one mapping between the intensity threshold and the exposure dose [43] (see (22) for an example), we substitute the intensity threshold I_{th} for the exposure dose.

At a certain intensity threshold, the printed contour is the least sensitive to the focus variations. This intensity threshold is called the *iso-focal threshold* denoted as I_{thiso} . The formal definition is

$$\frac{\partial E}{\partial z} \Big|_{I=I_{thiso}} = 0, \quad (28)$$

which means

$$\frac{\partial I}{\partial z} \Big|_{I=I_{thiso}} = 0. \quad (29)$$

Since $I = I_0 + z^2 I_2$ (see (21)), we have

$$I_2 \Big|_{I_0=I_{thiso}} = 0. \quad (30)$$

If we choose I_{thiso} as the intensity threshold I_{th} , we have

$$\mathbf{E}(I_{\text{thiso}}, 0) = \mathbf{E}(I_{\text{thiso}}, z) \quad (31)$$

We call the above quantity the *iso-focal EPE* and denote it as \mathbf{E}_{iso} . Negate both sides of (31) and add $\mathbf{E}(I_{\text{th}}, z)$ to both sides, we have

$$\mathbf{E}(I_{\text{th}}, z) - \mathbf{E}(I_{\text{thiso}}, 0) = \mathbf{E}(I_{\text{th}}, z) - \mathbf{E}(I_{\text{thiso}}, z) \quad (32)$$

Approximating $\mathbf{E}(I_{\text{th}}, z) - \mathbf{E}(I_{\text{thiso}}, z)$ as a separable function [43], we have

$$\mathbf{E}(I_{\text{th}}, z) - \mathbf{E}_{\text{iso}} = \mathbf{a}(z)b(I_{\text{th}} - I_{\text{thiso}}), \quad (33)$$

where $b(\cdot)$ satisfies $b(0) = 0$.

For small I_{th} variations (usually within 10% for modern lithography system), $b(I_{\text{th}} - I_{\text{thiso}})$ can be linearized. So we have

$$\mathbf{E}(I_{\text{th}}, z) - \mathbf{E}_{\text{iso}} = \mathbf{a}(z)(I_{\text{th}} - I_{\text{thiso}}). \quad (34)$$

Due to the $z \leftrightarrow -z$ symmetry of VLIM, $\mathbf{a}(z)$ can be expanded for small z 's as

$$\mathbf{a}(z) = \mathbf{a}_0 + \mathbf{a}_1 z^2. \quad (35)$$

We call the EPE as \mathbf{E}_t if the printed contour coincides the target contour. Plugging \mathbf{E}_t in (34) and using (21) and (35), we have

$$\begin{aligned} \mathbf{E}_t - \mathbf{E}_{\text{iso}} &= (\mathbf{a}_0 + \mathbf{a}_1 z^2) \\ &\times (I_0(\mathbf{E}_t) + z^2 I_2(\mathbf{E}_t) - I_{\text{thiso}}). \end{aligned} \quad (36)$$

Expanding (36) with respect to z , we have

$$\begin{aligned} \mathbf{E}_t - \mathbf{E}_{\text{iso}} &= \mathbf{a}_0 (I_0(\mathbf{E}_t) - I_{\text{thiso}}) \\ &+ \left(\mathbf{a}_1 (I_0(\mathbf{E}_t) - I_{\text{thiso}}) + \mathbf{a}_0 I_2(\mathbf{E}_t) \right) z^2 \\ &+ O(z^4) \end{aligned} \quad (37)$$

We ignore the highest order term of z (the z^4 term). Since the equality in (37) is independent of z , by setting the coefficients of z^0 and z^2 to zeros, we get the solutions for \mathbf{a}_0 and \mathbf{a}_1

$$\begin{cases} \mathbf{a}_0 &= \frac{\mathbf{E}_t - \mathbf{E}_{\text{iso}}}{I_0(\mathbf{E}_t) - I_{\text{thiso}}} \\ \mathbf{a}_1 &= -\mathbf{a}_0 \frac{I_2(\mathbf{E}_t)}{I_0(\mathbf{E}_t) - I_{\text{thiso}}} \end{cases}. \quad (38)$$

The vector \mathbf{a}_1 is proportional to the vector \mathbf{a}_0 as shown in the second equation of (38). We express the ratio between them as $a_1 = -\frac{I_2(\mathbf{E}_t)}{I_0(\mathbf{E}_t) - I_{\text{thiso}}}$. Then, variational EPE model under any intensity threshold and focus variation (34) can be written as

$$\mathbf{E}(I_{\text{th}}, z) = \mathbf{E}_{\text{iso}} + \mathbf{a}_0(1 + a_1 z^2)(I_{\text{th}} - I_{\text{thiso}}), \quad (39)$$

where

$$\begin{cases} \mathbf{a}_0 &= \frac{\mathbf{E}_t - \mathbf{E}_{\text{iso}}}{I_0(\mathbf{E}_t) - I_{\text{thiso}}} \\ a_1 &= -\frac{I_2(\mathbf{E}_t)}{I_0(\mathbf{E}_t) - I_{\text{thiso}}} \end{cases}. \quad (40)$$

B. Variational-EPE Metrics

From the variational EPE model (39), we can compute the V-EPE metrics of the interest to guide OPC. As an example, let us assume z and I_{th} are independent and normally distributed:

$$z \sim N(\mu_z, \sigma_z^2) \quad \text{and} \quad I_{\text{th}} \sim N(\mu_{I_{\text{th}}}, \sigma_{\text{th}}^2), \quad (41)$$

we can compute all the EPE moments easily. Assuming $\mu_z = 0$, from (39) and (41), we have the average EPE (the first moment) under the intensity threshold and focus variations

$$\begin{aligned} \text{V-EPE} = \langle \mathbf{E} \rangle &= \mathbf{E}_{\text{iso}} + \mathbf{a}_0(1 + a_1 \sigma_z^2)(\mu_{I_{\text{th}}} - I_{\text{thiso}}) \\ &= \underbrace{\mathbf{E}_{\text{iso}} + \mathbf{a}_0(\mu_{I_{\text{th}}} - I_{\text{thiso}})}_{\mathbf{E}_{\text{nom}}} + \underbrace{\mathbf{a}_0 a_1 \sigma_z^2(\mu_{I_{\text{th}}} - I_{\text{thiso}})}_{\boldsymbol{\mu}}, \end{aligned} \quad (42)$$

where \mathbf{E}_{nom} is the nominal EPE. It is clear that considering focus variation the average EPE $\langle \mathbf{E} \rangle$ will be always different from the nominal EPE \mathbf{E}_{nom} . Note that the definition of V-EPE is not limited to the average EPE. Other desirable quantities, such as the variance, can also be included.

For the real manufacturing process, as long as the joint distribution of measured exposure dose (E) and focus error (Z) is available, the joint distribution of I_{th} and z can be computed using (22) and (23). Hence, the average EPE $\langle \mathbf{E} \rangle$ can be computed without any difficulty.

V. PROCESS VARIATION-AWARE OPC ALGORITHM (PV-OPC)

Conventional OPC softwares try to reduce the nominal EPE. However, this would result in more average post-OPC EPE under process variations. Instead, our process variation aware OPC

(PV-OPC) algorithm is based on the V-EPE metric defined in the previous section to make the average on target. The metric is generic enough to apply to any sparse OPC algorithm. We show our implementation details in this section.

A. OPC Shape Engine

OPC shape engine refers to the representation, storage and lithography simulator interpretation of mask shapes. Although there are many OPC papers, most of them focused on OPC recipes for commercial OPC softwares, such as CalibreTM or ProteusTM. Only a few early papers [12–14] discussed the OPC shape engine data structure. In these papers, the original drawn shapes are represented as polygons, called *fixed mask objects*. Many so called *variable mask objects* are attached to the edges of each polygon (Figure 10). It can be seen that, to the first order, the simulation time is proportional to the number of vertices (Section III C). However, the vertices (e.g., v_0) in this method essentially present multiple times in both the fixed mask object and its variable mask objects, which results in inefficiency during computation.

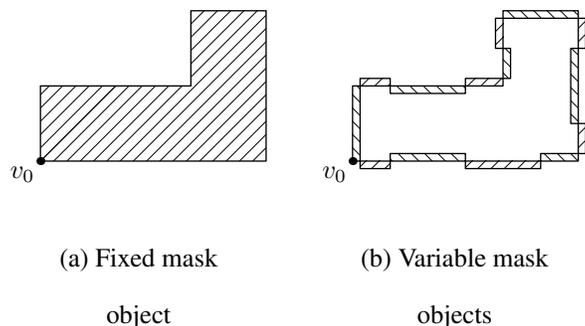
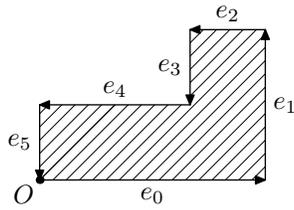


Figure 10: Previous mask shape representation example [12–14] : a fixed mask object and its variable mask objects. Vertices can present more than one times in the fixed mask object and its variable mask objects. The 45° shaded regions have positive convolution values (see (27)) and the 135° shaded regions have negative convolution values. It is semantically equivalent to that of Figure 12 which uses our new proposed method.

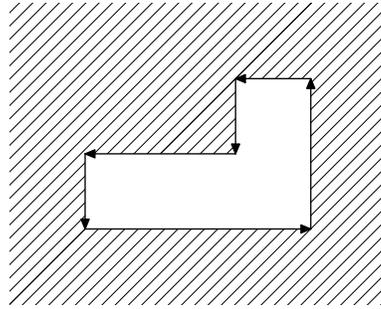
To get rid of the redundancies in the representation, we employ a similar idea to Chain Code [56–58]. Our proposed method parametrizes the polygon such that it can efficiently represent changes in the edge locations. We only discuss rectilinear polygons (Figure 11 and Figure 12). The method can be easily extended to polygons with 45° degree edges.

The rectilinear polygon in Figure 11 is composed of a series of directed edges. The head of

each edge e_i is connected to the tail of the next one ($e_{\text{mod}(i+1,N)}$), where N is the number of edges and mod denotes the modulo operation. Each edge can be specified by its length l and two Boolean variables h and p , where h indicates whether it is horizontal or vertical and p indicates whether it points to the positive direction (x or y) or the negative direction ($-x$ or $-y$). The per polygon Boolean variable c indicates whether the interior of the region is to the left or to the right of the edges of the polygon. O denotes the starting point of the first edge (e_0).



(a) The interior is to the left of the edges ($c = 1$).



(b) The interior is to the right of the edges

($c = 0$).

Figure 11: Rectilinear polygon representation. The interior region is shaded. Each edge is represented as a triplet $e_i = \{l_i, h_i, p_i\}$. The polygon is represented as $\{O, c, \{e_i | 0 \leq i < N\}\}$. In this example, $N = 6$.

Figure 12 shows that the polygon in Figure 11 is segmented into many small segments, denoted as dashed arrows. In addition to the parameters $\{l_i, h_i, p_i\}$, one more parameter d_i should be used to describe the displacement of the segment from its original location. By applying nonzero d 's to all the segments, they are shifted to the locations denoted as solid arrows. The vertices needed in lithography simulators can be computed based on this representation. Smaller number of vertices results in shorter lithography simulation runtime.

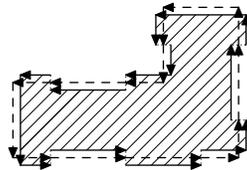


Figure 12: Segmented rectilinear polygon.

B. Segment Movement Scheme

Each segment can be moved based on the print contour information locally [14] or on the non local print contour information, e.g. the MEEF matrix based scheme [15].

We use the first approach. But our PV-OPC algorithm can be extended to use the second approach. We use the standard OPC segmentation and tagging strategy [14]. Each segment is moved based on the V-EPE metric $\langle E \rangle$ at its control point. However, we update all the segment displacements at the same time, instead of updating them one at a time [14]. The details of the PV-OPC algorithm are shown in Algorithm 3. It is an iterative algorithm, where the constant C controls the edge movement step.

Algorithm 3 PV-OPC algorithm

Input: Non-touching polygons decomposed from the original design

Output: PV-OPCed mask

```
1: Segment the polygons into movable edges and tag the middle points as their control points
2: repeat
3:   updated  $\leftarrow$  false
4:   for each control point do
5:     compute the maximum aerial gradient direction
6:     store  $\langle E \rangle$  along that direction
7:   for each edge do
8:     if  $|C\langle E \rangle \cdot \mathbf{n}| \geq$  manufacturing grid then
9:       move the edge by  $-C\langle E \rangle \cdot \mathbf{n}$  (rounding to a multiple of manufacturing grids)
10:   updated  $\leftarrow$  true
11: until updated = false
```

The main difference compared to the conventional OPC algorithm is the objective function $\langle E \rangle$ which incorporates the process-variation information. We could also use other variation-EPE metrics. Due to the analytical nature of our model and efficient table lookup, the complexity is the same as the conventional OPC, with just a slightly bigger constant (as we shall show in the experimental results).

VI. EXPERIMENTAL RESULTS

A. VLIM Calibration

We implemented VLIM in C++. We used PROLITHTM as our virtual fab and calibrated VLIM to the PROLITHTM simulation results. Note that the calibration method is generic enough to handle real fab data.

To calibrate the VLIM, we need to do PROLITHTM simulation on different patterns. In industrial lithography model test cases, there are many patterns, including lines, spaces and contacts, etc. For demonstration purpose, we only used four periodic line/space patterns, which have the same line width (65 nm), but different pitches (180, 300, 500 and 1000 nm).

For each pattern, we used VLIM to simulate at evenly sampled diffusion lengths d 's (0, 2, 4, ..., 20 nm). At each sampled d value, we also evenly sampled intensity threshold I_{th} and focus error z , such that $(\frac{I_{\text{th}}-I_{\text{th0}}}{\Delta I_{\text{th}}})^2 + (\frac{z}{\Delta z})^2 < 1$, where $I_{\text{th0}} = 0.17$, $\Delta I_{\text{th}} = 0.02$ and $\Delta z = 80$ nm. We computed the CD function $\text{CD}^{\mathbf{P}}(I_{\text{th}}, z, d)$ for pattern \mathbf{P} using the method in Appendix (A.1). The upper bound L , M and N are all set to 3 in our experiment.

We did PROLITHTM CD simulation on the four patterns at evenly sampled (E, Z) points in the region $(\frac{E-E_0}{\Delta E})^2 + (\frac{Z-Z_0}{\Delta Z})^2 < 1$, where $E_0 = 32.0$ mJ/cm², $\Delta E = 1.0$ mJ/cm², $Z_0 = 25$ nm and $\Delta Z = 75$ nm. The PROLITHTM input parameters are summarized in Table I. The fitting results are shown in Table II. The parameter σ denotes the fitting error. It is only 0.71 nm for all test patterns, thus our VLIM is fairly accurate under process variations. It should be noted that as more test patterns are present, the fitting error may grow, which is expected for any phenomenological model. The calibration with real experimental data will also probably result in bigger fitting error. However, to the first order, our VLIM is good enough to guide OPC or other mask/layout synthesis.

B. OPC results comparison

We implemented both the conventional OPC and the PV-OPC algorithms in C++. Our test layouts are the poly layers of a inverter and a NAND gate following 65 nm minimum and recommended design rules, named as minINV, recINV, minNAND, recNAND. We use the nominal condition $I_{\text{th}} = 0.15$ and $z = 0$ nm for the conventional OPC and the distribution parameters $\mu_{I_{\text{th}}} = 0.15$, $\sigma_{I_{\text{th}}} = 0.007$, $\mu_z = 0$ nm and $\sigma_z = 80$ nm for the PV-OPC. The bias B is set to zero

Table I: PROLITH™ parameter summary.

90 nm node line ArF example - FC	
Photoresist thickness	232 nm
Brewer DUV 42C BARC	
BARC thickness	50 nm
BARC refraction index	$1.48 + 0.41i$
Silicon substrate	
Substrate refraction index	$0.883143 + 2.77779i$
Ideal bake mode	
Bake time	60 sec
Bake temperature	95 °C
Kirchhoff mask simulation mode	
Conventional partially coherent illumination	
Partial coherence	0.7
Wavelength	193 nm
Numerical aperture	0.8
Focal position relative to middle of resist	
Offset from the top	-40 nm
Positive numbers shift up	
Ideal PEB model	
PEB time	90 sec
PEB temperature	120 °C
Base surface contamination	
Relative contaminant concentration	0.001
Contaminant diffusion length	50 nm
User defined developer	
Develop time	20 sec
Vector image calculation model	
Full physics resist model	
Number of Exposure passes	1
Main speed factor	2
Speed factor for XY step	4
Speed factor for Z step	4
Source grid step size	0.0322581
X step size	2.25
Z step size	2.32

Table II: Fitting results.

d	1.9484 nm
B	-0.666 471 nm
α_Z	22.7276 nm
β_Z	1.10935
α_E	11.5705 mJ/cm ²
β_E	3.301 49 mJ/cm ²
σ	0.71 nm

for both OPC algorithms.

We computed the CD mean $\langle \text{CD} \rangle$ and the CD variance $\text{Var}(\text{CD})$ at every 1 nm in the NMOS and PMOS active regions along the gate width direction. We then computed the averages of $\langle \text{CD} \rangle$ and $\text{Var}(\text{CD})$, denoted by $\overline{\langle \text{CD} \rangle}$ and $\overline{\text{Var}(\text{CD})}$, in each region. Table III shows the comparison between the results of the two OPC algorithms. $\overline{\langle \text{CD} \rangle}$'s from the PV-OPC are much closer to 65nm than those from the conventional OPC. $\overline{\text{Var}(\text{CD})}$ from the PV-OPC are comparable to those from the conventional OPC. Thus PV-OPC is more robust with respect to process variations.

Table III: Post-OPC CD mean and EPE variance comparison.

circuits	MOS type	average CD mean (nm)		average EPE variance (nm)	
		Conventional	PV-OPC	Conventional	PV-OPC
minINV	PMOS	58.83	66.28	3.56	3.77
	NMOS	58.89	66.75	3.27	3.47
recINV	PMOS	59.40	66.71	3.62	3.83
	NMOS	58.81	67.10	3.83	3.51
minNAND	PMOS	59.85	66.67	3.68	3.87
	NMOS	58.42	65.30	3.47	3.64
recNAND	PMOS	60.42	65.36	3.74	3.88
	NMOS	60.40	67.41	3.55	3.70

We show the results from both OPC algorithms for the poly layer layout of the inverter following the minimum design rules in Figure 13 and Figure 14. The NMOS region is indicated by the rectangles in Figure 13(a) and Figure 14(a). We also show the results from both OPC algorithms

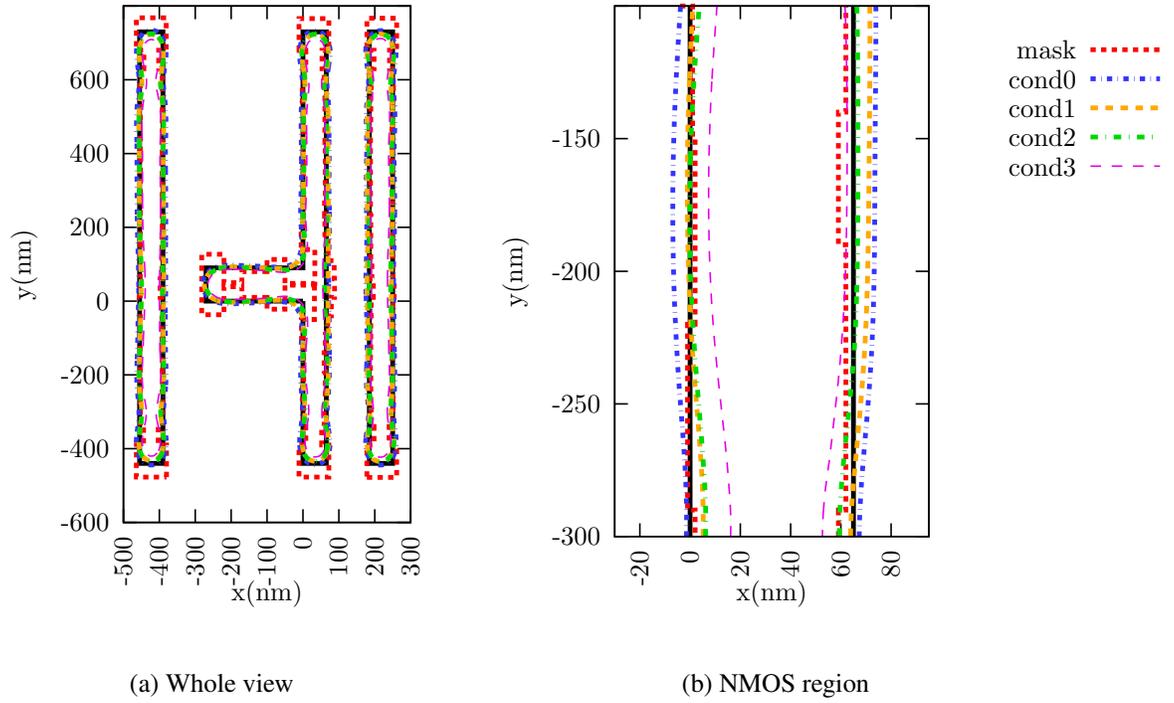


Figure 13: Conventional OPC (inverter following the minimum design rules): $\overline{\text{Var}(\text{CD})}$ CD error is -6.11 nm.

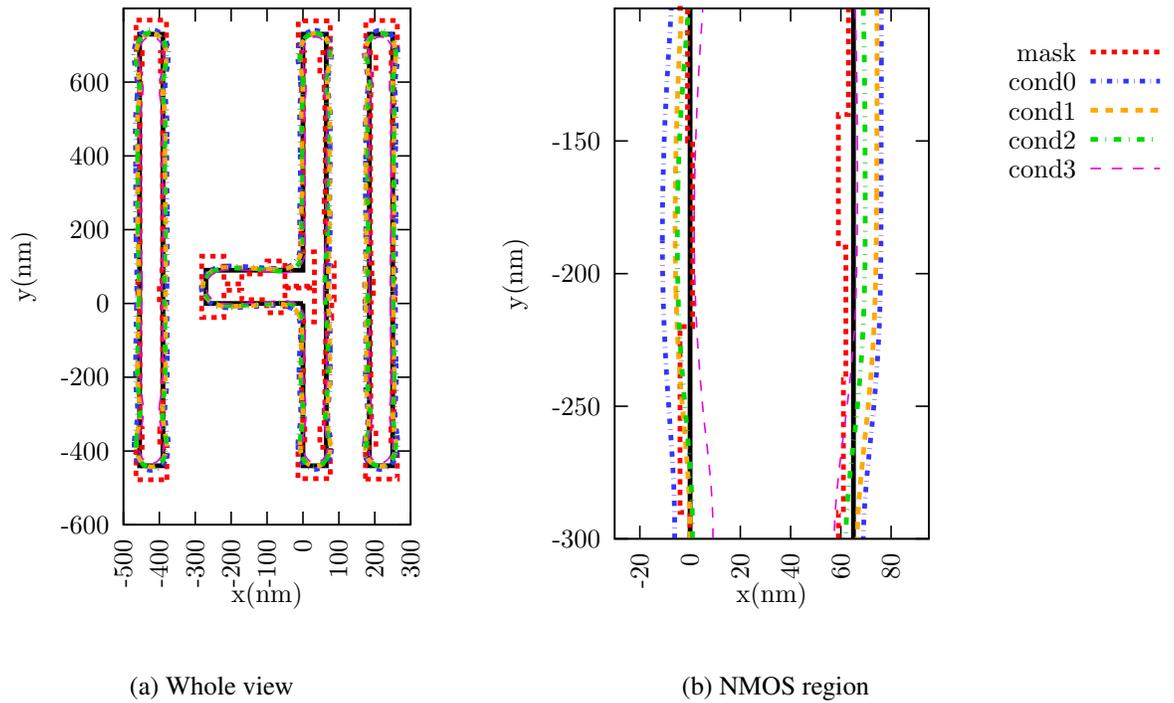


Figure 14: PV-OPC (inverter following the minimum design rules): $\overline{\text{Var}(\text{CD})}$ error is 1.75 nm.

for the poly layer layout of the NAND gate following the recommended design rules in Figure 15 and Figure 16. One of the PMOS regions is indicated by the rectangles in Figure 15(a) and Figure 16(a). The targets, OPCed mask shapes and the printed contours at four process conditions

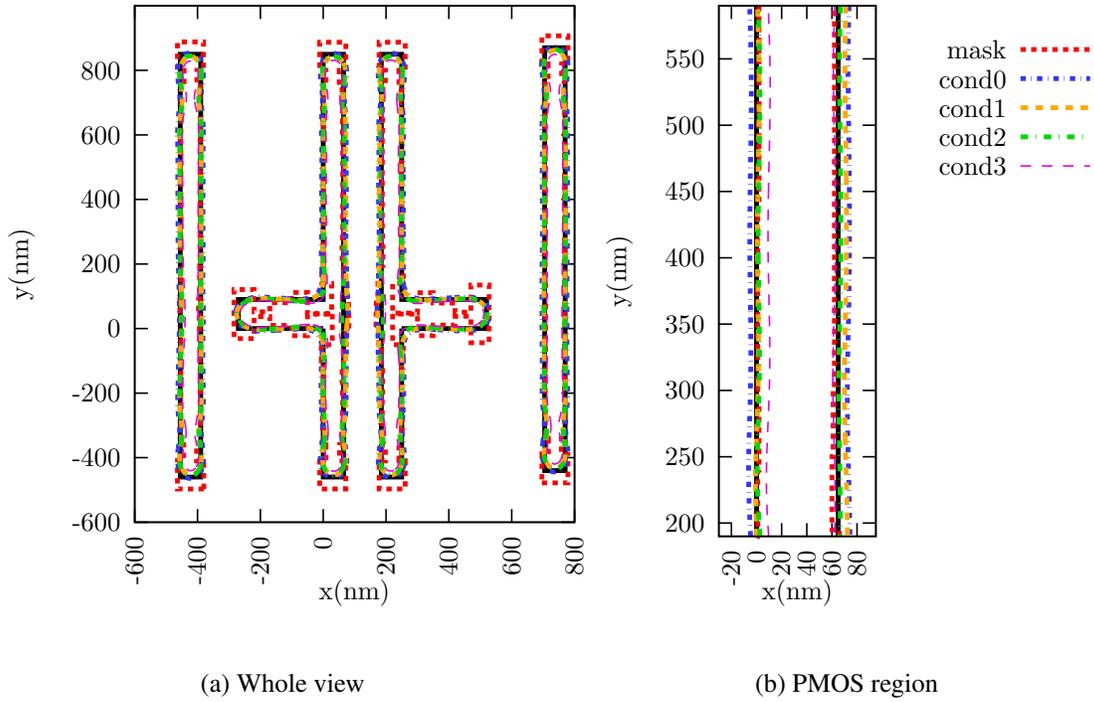


Figure 15: Conventional OPC (NAND following the recommended design rules): $\overline{\text{Var}(\text{CD})}$ error is -4.85 nm.

(Table IV) are shown in these figures.

Table IV: Four process parameters.

		I_{th}	
		0.143	0.157
z	0 nm	0	2
	80 nm	1	3

In these figures, the printed gate lengths tend to shrink due to focus variations¹. By comparing

¹ Depending on the processes, it could expand as well.

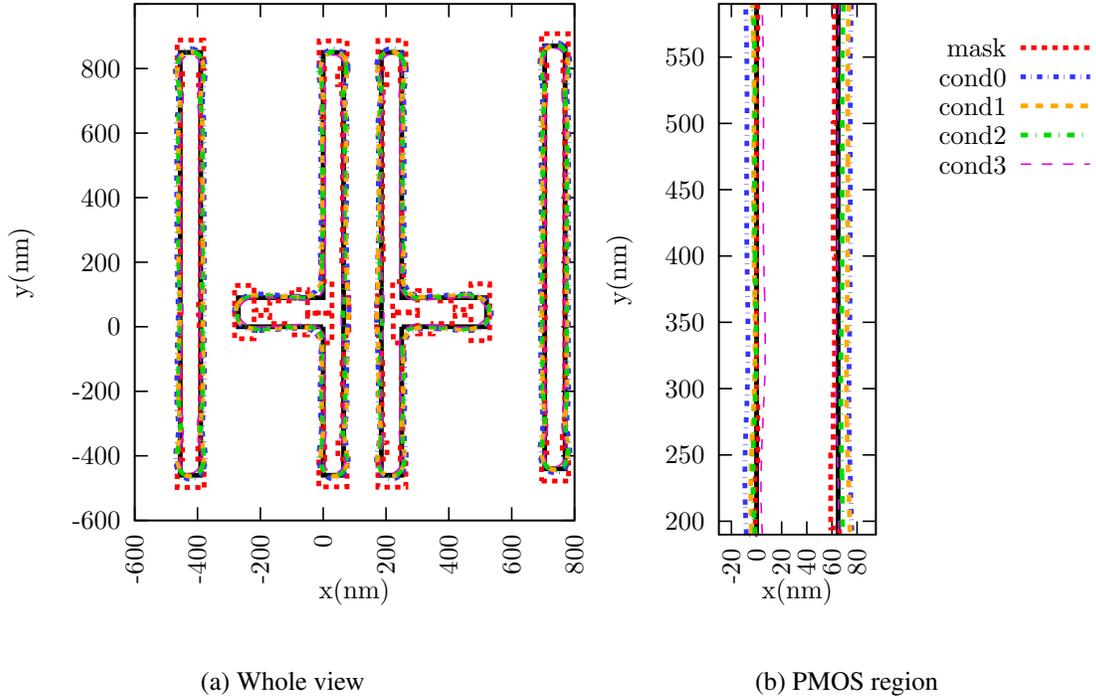
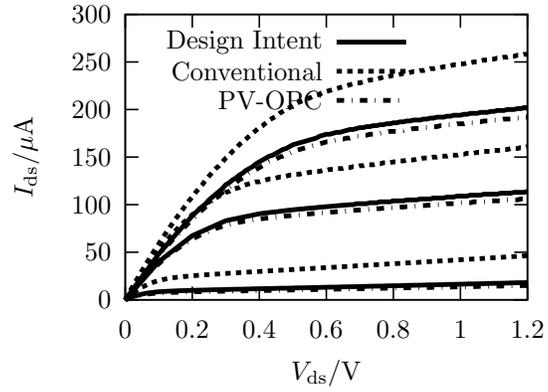


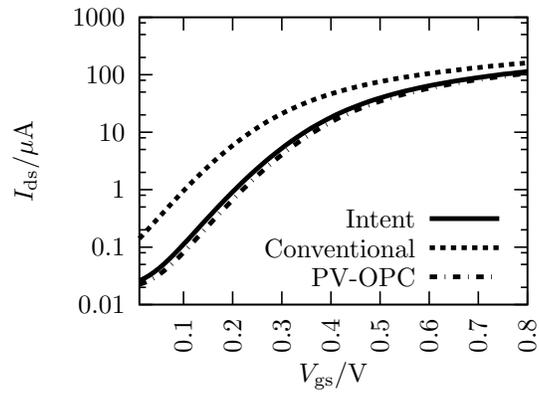
Figure 16: PV-OPC (NAND following the recommended design rules): $\overline{\text{Var}(\text{CD})}$ error is 0.36 nm. Figure 13 to Figure 14 and comparing Figure 15 to Figure 16, we can see that PV-OPC bias the edges toward the outside intelligently.

Figure 17 and Figure 18 show the robustness of PV-OPC with respect to the electrical characterizations. Figure 17(a) and Figure 18(a) show the NMOS and PMOS I - V curves at $V_{\text{gs}} = 0.4$ V, 0.8 V and 1.2 V, respectively. Figure 17(b) and Figure 18(b) show the NMOS leakage current as a function of V_{gs} at $V_{\text{ds}} = 1.2$ V. The solid curves represent the design intent of the drawn layout. The dash ones represent the post-OPC expectation of the conventional OPCed mask considering lithography variation. The dash-dot curves represent the expectation of the variation-aware OPCed mask. It is clear to see from Figure 17(a) and Figure 18(a), conventional OPC can not make the I - V curve expectations the same as the design intent. However, our variation-aware OPC does a good job. Figure 17(a) and Figure 18(a) show that the sub-threshold leakage expectations of the conventional OPC results can be $4\times$ bigger than the design intent. Our variation-aware OPC algorithm effectively reduces the gap.

Table V shows the runtime and the total number of iterations of the conventional OPC algorithm and the PV-OPC algorithm. The PV-OPC runtime is only about 2 - $3\times$ slower, which is very

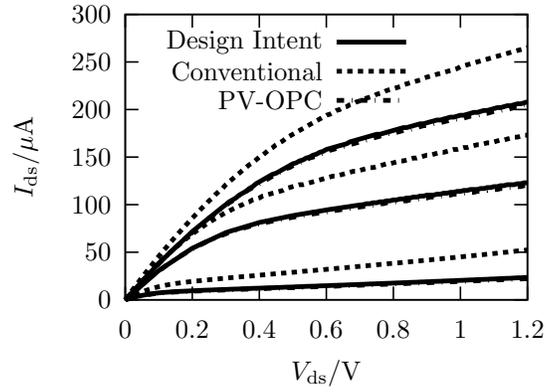


(a) I - V curves

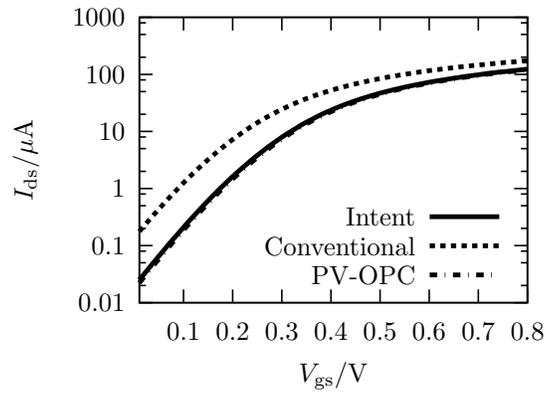


(b) Leakage (PV-OPC leakage curve is almost exact same as the design intent one)

Figure 17: NMOS electrical characterization of the inverter following 65 nm minimum design rules.



(a) I - V curves



(b) Leakage (PV-OPC leakage curve is almost exact same as the design intent one)

Figure 18: PMOS electrical characterization of the NAND gate following 65 nm recommended design rules.

impressive considering that it explicitly incorporates the entire process window information.

Table V: OPC Runtime Comparison

circuit	Conventional OPC		PV-OPC	
	Runtime	Iter#	Runtime	Iter#
minINV	3.74 sec	42	10.87 sec	47
recINV	4.14 sec	42	10.91 sec	42
minNAND	5.57 sec	41	14.99 sec	44
recNAND	7.84 sec	52	17.46 sec	44

VII. CONCLUSIONS

In this paper, a new variational lithography modeling is derived and calibrated to the industry standard lithography simulation software PROLITHTM. Base on this variational lithography model, a variational EPE metrics is presented. A variational aware OPC algorithm is proposed based on this new metrics. We show the implementation details of the PV-OPC algorithm. The PV-OPC algorithm is tried on some 65nm layouts. It obtains much more robust results than the conventional OPC in terms of both the geometric and electrical metrics. The runtime increases only 2-3 \times , which is very decent for a true process variation aware OPC.

Acknowledgments

This work is partially supported by SRC, IBM Faculty Award, Fujitsu, Sun, and Intel equipment and KLA-Tencor software donations. The authors would like to thank Dr. Chris Mack and Dr. Warren Grobman for helpful discussions.

Appendix: VLIM BASED BOSSUNG CURVES DEDUCTION

Focus-Exposure Matrix is defined as the variation of line width (and possibly other parameters) as a function of both focus and exposure energy. The data is typically plotted as line width versus

focus for different exposure energies and these plots are often referred to as smiley plots, spider plots, or Bossung curves [50].

Before we show how to reduce the model data into curves which have analytical formulas, we discuss the requirement on the production lithography systems — maturity. This means process condition variations should be small enough such that photoresist profile changes as a result should also be small. There should not be anything like photoresist collapse, etc. Otherwise, if the process is too sensitive to process condition variations or the process condition variations are too large, we should better improve the process first rather than using any tricks to survive in that process.

Having this assumption, the line width (or CD) can be expressed as polynomials of exposure dose and focus error. Based on this idea, various polynomial fitting functions have been proposed [43, 59, 60]. In VLIM, we have diffusion length parameter d . We assume CD can also be expressed as a polynomial of d . By taking the $z \leftrightarrow -z$ symmetry, we have only even order terms of z . The CD function can be expressed as

$$\text{CD}_{\text{VLIM}}^P(I_{\text{th}}, z, d) = \sum_{l=0}^L \sum_{m=0}^M \sum_{n=0}^N a_{lmn} I_{\text{th}}^l z^{2m} d^n, \quad (\text{A.1})$$

where the bias B is set as 0 in this case. a_{lmn} and $I_{\text{th}0}$ are fitting parameters. The superscript P denotes that the different CD function for different pattern P . Using non-linear least square regression, we can fit the CD function to VLIM to CD data generated at various I_{th} , z , d . The upper limits, L , M and N , of the summations are chosen heuristically. They should not too big or too small to overfit or underfit the curves.

-
- [1] P. Yu, S. X. Shi, and D. Z. Pan, “Process variation aware opc with variational lithography modeling,” in *Proc. DAC*, 2006, pp. 785–790.
 - [2] A. K.-K. Wong, *Resolution Enhancement Techniques in Optical Lithography*. SPIE, May 2001.
 - [3] A. K. Wong, “Microlithography: Trends, challenges, solutions, and their impact on design,” *IEEE Micro*, vol. 23, no. 2, pp. 12–21, 2003.
 - [4] L. Liebmann, “Layout impact of resolution enhancement techniques: impediment or opportunity?” in *Proc. International Symposium on Physical Design*, 2003, pp. 110–117.
 - [5] R. F. Pease, “Lithographic technologies that haven’t (yet) made it: lessons learned (Plenary Paper),” in *Proc. SPIE 5751*, 2005, pp. 15–25.

- [6] B. J. Lin, "Where is the lost resolution," in *Proc. SPIE 633*, 1986, p. 44.
- [7] —, "The k_3 coefficient in nonparaxial lambda/na scaling equations for resolution, depth of focus, and immersion lithography," *Journal of Microlithography, Microfabrication, and Microsystems*, vol. 1, no. 1, pp. 7–12, 2002.
- [8] T. A. Brunner, "Why optical lithography will live forever," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 21, no. 6, pp. 2632–2637, 2003.
- [9] N. Cobb, "Flexible sparse and dense OPC algorithms," in *Proc. SPIE 5853*, June 2005, pp. 693–702.
- [10] N. B. Cobb and Y. Granik, "Dense OPC for 65nm and below," in *Proc. SPIE 5992*, Nov. 2005, pp. 1521–1532.
- [11] N. Cobb and D. Dudau, "Dense OPC and verification for 45nm," in *Proc. SPIE 6154*, Apr. 2006, pp. 191–196.
- [12] N. B. Cobb and A. Zakhor, "Fast, low-complexity mask design," in *Proc. SPIE 2440*, May 1995, pp. 313–327.
- [13] N. B. Cobb, A. Zakhor, and E. Miloslavsky, "Mathematical and CAD framework for proximity correction," in *Proc. SPIE 2726*, June 1996, pp. 208–222.
- [14] N. B. Cobb, "Fast Optical and Process Proximity Correction Algorithms for Integrated Circuit Manufacturing," Ph.D. dissertation, University of California at Berkeley, 1998.
- [15] N. B. Cobb and Y. Granik, "Model-based opc using the meef matrix," in *Proc. SPIE 4889*, 2002, pp. 1281–1292.
- [16] Y. Granik, "Generalized mask error enhancement factor theory," *Journal of Microlithography, Microfabrication, and Microsystems*, vol. 4, no. 2, p. 023001, 2005.
- [17] A. Krasnoperova, J. A. Culp, I. Graur, S. Mansfield, M. Al-Imam, and H. Maaty, "Process window OPC for reduced process variability and enhanced yield," in *Proc. SPIE 6154*, Apr. 2006, pp. 1200–1211.
- [18] N. B. Cobb and Y. Granik, "OPC methods to improve image slope and process window," in *Proc. SPIE 5042*, 2003, pp. 116–125.
- [19] J. L. Sturtevant, J. A. Torres, J. Word, Y. Granik, and P. LaCour, "Considerations for the use of defocus models for OPC," in *Proc. SPIE 5756*, 2005, pp. 427–436.
- [20] P. Gupta, A. B. Kahng, C.-H. Park, K. Samadi, and X. Xu, "Wafer topography-aware optical proximity correction for better DOF margin and CD control," in *Proc. SPIE 5853*, 2005, pp. 844–854.
- [21] Y. Granik and N. B. Cobb, "MEEF as a matrix," in *Proc. SPIE 4562*, 2002, pp. 980–991.

- [22] N. Cobb and Y. Granik, “New concepts in OPC,” in *Proc. SPIE 5377*, 2004, pp. 680–690.
- [23] J. Yang, L. Capodiecici, and D. Sylvester, “Advanced timing analysis based on post-opc extraction of critical dimensions,” in *Proc. Design Automation Conference*, 2005, pp. 359–364.
- [24] J. L. Sturtevant, J. Word, P. LaCour, J. W. Park, and D. Smith, “Assessing the impact of real world manufacturing lithography variations on post-OPC CD control,” in *Proc. SPIE 5756*, 2005, pp. 240–254.
- [25] A. Borna, C. Progler, and D. Blaauw, “Correlation analysis of CD-variation and circuit performance under multiple sources of variability,” in *Proc. SPIE 5756*, May 2005, pp. 168–177.
- [26] C. J. Progler, A. Borna, D. Blaauw, and P. Sixt, “Impact of lithography variability on statistical timing behavior,” in *Proc. SPIE*, 2004, pp. 101–110.
- [27] P. Gupta and F.-L. Heng, “Toward a systematic-variation aware timing methodology,” in *Proc. Design Automation Conference*, 2004, pp. 321–326.
- [28] S. X. Shi, P. Yu, and D. Z. Pan, “A Unified Non-Rectangular Device and Circuit Simulation Model for Timing and Power,” in *Proc. ICCAD*, 2006.
- [29] L. Liebmann, S. Mansfield, G. Han, J. Culp, J. Hibbeler, and R. Tsai, “Reducing DfM to practice: the lithography manufacturability assessor,” in *Proc. SPIE 6156*, Apr. 2006, pp. 178–189.
- [30] C.-K. Chen, T.-S. Gau, J.-J. Shin, R.-G. Liu, S.-S. Yu, A. Yen, and B. J. Lin, “Mask error tensor and causality of mask error enhancement for low- k_1 imaging: theory and experiments,” *J. Microlithogr. Microfabrication, Microsyst.*, vol. 3, no. 2, pp. 269–275, Apr. 2004.
- [31] E. W. Conrad, D. C. Cole, D. P. Paul, and E. Barouch, “Model considerations, calibration issues, and metrology methods for resist-bias models,” in *Proc. SPIE 3677*, 1999, pp. 940–955.
- [32] D. C. Cole, E. Barouch, E. W. Conrad, and M. Yeung, “Using Advanced Simulation to Aid Microlithography Development,” *Proc. IEEE*, vol. 89, no. 8, pp. 1194–1213, 2001.
- [33] M. D. Smith, J. D. Byers, and C. A. Mack, “Comparison between the process windows calculated with full and simplified resist models,” in *Proc. SPIE 4691*, July 2002, pp. 1199–1210.
- [34] M. D. Smith and C. A. Mack, “Process sensitivity and optimization with full and simplified resist models,” in *Proc. SPIE 5040*, June 2003, pp. 1509–1520.
- [35] M. Born and E. Wolf, *Principles of Optics : Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. Cambridge University Press, Oct. 1999.
- [36] M. S. Yeung, D. Lee, R. Lee, and A. R. Neureuther, “Extension of the Hopkins theory of partially coherent imaging to include thin-film interference effects,” in *Proc. SPIE 1927*, Aug. 1993, pp. 452–

463.

- [37] D. G. Flagello, T. Milster, and A. E. Rosenbluth, "Theory of high-NA imaging in homogeneous thin films," *Journal of the Optical Society of America A: Optics, Image Science, and Vision, Volume 13, Issue 1, January 1996, pp.53-64*, vol. 13, pp. 53–64, Jan. 1996.
- [38] K. Adam, Y. Granik, A. Torres, and N. B. Cobb, "Improved modeling performance with an adapted vectorial formulation of the Hopkins imaging equation," in *Proc. SPIE 5040*, June 2003, pp. 78–91.
- [39] K. Toh, "Two-dimensional images with effects of lens aberrations in optical lithography," EECS Department, University of California, Berkeley, Tech. Rep. UCB/ERL M88/30, 1988. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/1988/1037.html>
- [40] J. G. Garofalo, J. Demarco, J. Bailey, J. Xiao, and S. Vaidya, "Reduction of ASIC gate-level line-end shortening by mask compensation," in *Proc. SPIE 2440*, May 1995, pp. 171–183.
- [41] C. Dolainsky and W. Maurer, "Application of a simple resist model to fast optical proximity correction," in *Proc. SPIE 3051*, July 1997, pp. 774–780.
- [42] C.-N. Ahn, H.-B. Kim, and K.-H. Baik, "Novel approximate model for resist process," in *Proc. SPIE 3334*, June 1998, pp. 752–763.
- [43] D. Fuard, M. Besacier, and P. Schiavone, "Assessment of different simplified resist models," in *Proc. SPIE 4691*, 2002, pp. 1266–1277.
- [44] ———, "Validity of the diffused aerial image model: an assessment based on multiple test cases," in *Proc. SPIE 5040*, 2003, pp. 1536–1543.
- [45] T. A. Brunner, C. Fonseca, N. Seong, and M. Burkhardt, "Impact of resist blur on MEF, OPC, and CD control," in *Proc. SPIE 5377*, May 2004, pp. 141–149.
- [46] J. Schacht, K. Herold, R. Zimmermann, J. A. Torres, W. Maurer, Y. Granik, C.-H. Chang, G. K.-C. Hung, and B. S.-M. Lin, "Calibration of OPC models for multiple focus conditions," in *Proc. SPIE 5377*, May 2004, pp. 691–702.
- [47] J. A. Torres, T. Roessler, and Y. Granik, "Process window modeling using compact models," in *Proc. SPIE 5567*, Oct. 2004, pp. 638–648.
- [48] Y.-T. Wang, Y. C. Pati, and T. Kailath, "Depth of focus and the moment expansion," *OPTICS LETTERS*, vol. 20, no. 18, pp. 1841–1843, Sept. 1995.
- [49] P. Yu, D. Z. Pan, and C. A. Mack, "Fast lithography simulation under focus variations for OPC and layout optimizations," in *Proc. SPIE 6156*, Apr. 2006, pp. 397–406.
- [50] C. A. Mack, *Inside PROLITH: A Comprehensive Guide to Optical Lithography Simulation*. Finle

Technologies Inc., 1997.

- [51] J. Randall, K. Ronse, T. Marschner, M. Goethals, and M. Ercken, "Variable-threshold resist models for lithography simulation," in *Proc. SPIE 3679*, July 1999, pp. 176–182.
- [52] Y. Granik, N. B. Cobb, and T. Do, "Universal process modeling with VTRE for OPC," in *Proc. SPIE 4691*, July 2002, pp. 377–394.
- [53] T. A. Brunner and R. A. Ferguson, "Approximate models for resist processing effects," in *Proc. SPIE 2726*, June 1996, pp. 198–207.
- [54] J.-Y. Yoo, Y.-K. Kwon, J.-T. Park, D.-S. Sohn, S.-G. Kim, Y.-S. Sohn, and H.-K. Oh, "CD prediction by threshold energy resist model (TERM)," in *Proc. SPIE 4691*, July 2002, pp. 1287–1295.
- [55] J. Mitra, P. Yu, and D. Z. Pan, "RADAR: RET-aware detailed routing using fast lithography simulations," in *Proc. DAC*, 2005, pp. 369–372.
- [56] T. Kaneko and M. Okudaira, "Encoding of Arbitrary Curves Based on the Chain Code Representation," *IEEE Transactions on Communications*, vol. 33, no. 7, pp. 697–707, July 1985.
- [57] B. Chaudhuri and S. Chandrashekhar, "Neighboring direction runlength coding: an efficient contour codingscheme," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 20, no. 4, pp. 916–921, 1990.
- [58] C.-C. Lu and J. Dunham, "Highly efficient coding schemes for contour lines based on chain code representations," *IEEE Transactions on Communications*, vol. 39, no. 10, pp. 1511–1514, Oct. 1991.
- [59] C. A. Mack and J. D. Byers, "New model for focus-exposure data analysis," in *Proc. SPIE 5038*, May 2003, pp. 396–405.
- [60] A. Bourov, S. A. Robertson, B. W. Smith, M. Slocum, and E. C. Piscani, "Experimental measurement of photoresist modulation curves," in *Proc. SPIE 6154*, Apr. 2006, pp. 1003–1008.