Genetic Algorithms to Improve Mask and Illumination Geometries in Lithographic Imaging Systems

Tim Fühner[1], Andreas Erdmann[1], Richárd Farkas[2], Bernd Tollkühn[1], and Gabriella Kókai[2]

[1] Fraunhofer Institute of Integrated Systems and Device Technology (FhG-IISB), Schottkystr. 10, 91058 Erlangen, Germany,
{fuehner, andreas.erdmann, tollkuehn}@iisb.fraunhofer.de
[2] University of Erlangen-Nuremberg, Department of Computer Science II, Martensstr. 3, 91058 Erlangen, Germany,
gabriella.kokai@informatik.uni-erlangen.de

Abstract. This paper proposes the use of a genetic algorithm to optimize mask and illumination geometries in optical projection lithography. A fitness function is introduced that evaluates the imaging quality of arbitrary line patterns in a specified focus range. As a second criterion the manufacturability and inspectability of the mask are taken into account. With this approach optimum imaging conditions can be identified without any additional a-priori knowledge of the lithographic process. Several examples demonstrate the successful application and further potentials of the proposed concept.

1 Introduction

Optical projection lithography transfers the layout of a mask into a photoresist on the top surface of a silicon wafer [1]. The resolution of a microlithographic process is defined in terms of the size Δx of the minimum half-pitch or half-period of dense pattern of lines and spaces which can be produced by this process. In the context of microelectronic processing the Rayleigh-criterion for the resolution capability of a certain process is written as $\Delta x = k_1 \frac{\lambda}{\mathrm{NA}}$, where λ specifies the wavelength of the used light and NA is the numerical aperture of the projection system. k_1 is usually referred to as the "k-factor" of a process, which depends on the spatial coherence, on the mask technology, and on the properties of the photoresist. This Rayleigh criterion dictates the fundamental trend of optical lithography towards smaller wavelengths and larger numerical apertures. Until mid of the nineties standard lithographic processes employed k-factors above 0.6. The mask was designed as a pattern of dark and bright features with spatial dimensions of the features to be printed on the wafer.

In contrast to this technologically rather simple imaging process, nowadays' k-factors of 0.3 to 0.4 require sophisticated techniques to improve resolution capabilities, some of which are shortly introduced in this paper.

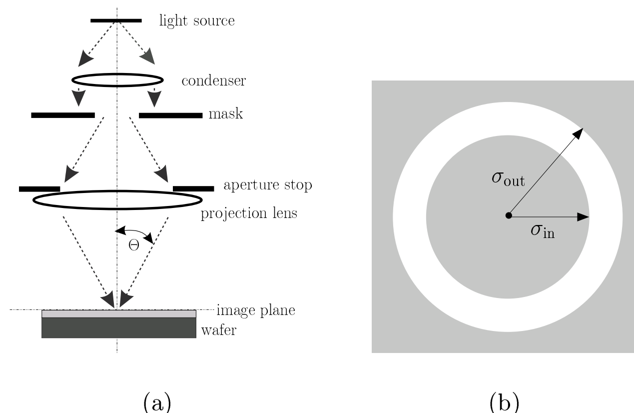
Furthermore, this paper proposes the utilization of a genetic algorithm (GA) to optimize mask and illumination source geometries, as the ever increasing

complexity in the imaging process often prevents conventional (analytical) approaches from being applicable.

The paper is organized as follows. Section 2 briefly reviews the basics of the lithographic projection system and of optical resolution enhancement techniques. Details of the optimization procedure, such as problem definition, data representation, fitness function, and a few details on the GA, are described in Section 3. Section 4 presents first results, obtained with the proposed method. This paper concludes with a short summary and an outlook on future work.

2 Introduction to Mask Layout and Illumination

Figure 1(a) shows a schematic drawing of a lithographic projection stepper/scanner. The imaging system consists of an illumination optics (light source, condenser) and a projection optics (projection lens, aperture stop). The condenser system is designed to ensure a homogeneous illumination of the mask. The projection system images the mask into the image plane close to the wafer surface. The projection lens transmits only a certain angular range of light which is diffracted by the mask, whose sine is bounded by the numerical aperture NA $= sin(\theta)$ of the projection system.



(a) (b)

Fig. 1. (a) Principle sketch of an optical projection system used in lithograhpy steppers and scanners; (b) Annular illumination source: σ_{out} denotes the radius of the illumination disk, σ_{in} the radius of the inner cover disk.

Standard lithographic projection equipment employs the Köhler illumination. In this specific geometry of the illumination system the light source is projected into the entrance pupil of the projection lens. The mask is considered to be illuminated by mutually incoherent plane waves which emerge from different points of the light source. The lithographic projection system is modeled with standard Fourier optics [2]. The spatial coherence of the system is taken into account by the Hopkins [3] theory. Polarization effects, which become important in high numerical aperture systems, require a vector extension of the scalar theory [4], [5]. The images which are shown in this paper were computed with the in-house lithography simulator of the Fraunhofer Institute IISB [6].

2.1 Optical Resolution Enhancement Techniques

The term “resolution enhancement techniques” (RET) covers modifications of the geometry of the illumination and of the mask which improve the resolution of lithographic processes. This subsection reviews the most important aspects of these RET, which are necessary to understand the optimization problems presented in the remaining part of this paper. For a more detailed understanding of RET the reader is referred to Alfred Wong’s excellent book [7].

Off-Axis Illumination (OAI). The geometry of illumination sources has a great impact on the imaging process. The investigations in this paper were performed for standard and annular illuminations Figure 1(b).

Optical Proximity Correction (OPC). Small sub-resolution assist features can strongly improve the performance of OAI for more isolated features. These assist feature are too small to be printed at the specified threshold intensity of the photoresist, but strongly improve the performance of the isolated feature. Additional assists at larger distances from the main feature can result in further improvements of the imaging performance.

Phase Shift Masks (PSM). In contrast to binary masks (BIM), which consist only of bright and dark areas, phase shift masks (PSM) modify the phase of the light by shifting the phase in its bright areas. This technique makes use of the fact that, with a phase-shift of $180\,^\circ$ of the adjoining features’ diffracted light, superposed fields are subtracted rather than added as in the case of unaltered phases. This leads to a significant enhancement of the resolution.

3 Optimization of the Lithographic Process

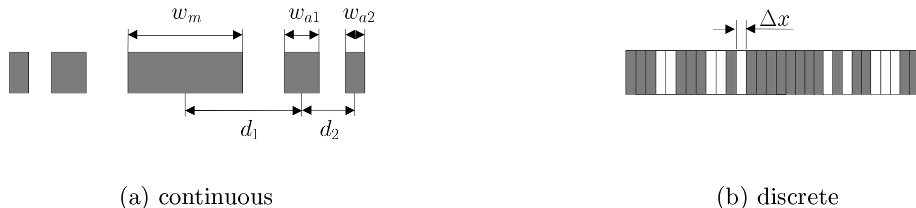
Implementation of the previously described techniques in real fabrication processes involves a number of problems. First of all, RET techniques are in most cases feature specific; e.g., OAI shows a good performance for dense features, but is only of limited use for isolated features. Moreover, there are many restrictions which result from manufacturability challenges. Thus, determination of the optimum mask design and source shape is a complex and demanding task.

3.1 Optimization Parameters

The optimization problem of one-dimensional mask with line/space patterns is specified in two different ways [8]. In the first case, the mask is represented by continuous variables which describe certain features on the mask.

Figure 2(a) depicts a main feature with sub-resolution assists, which are symmetrically positioned on the left and on the right side of the main feature. Simulations are performed using different numbers of assist features. Photomasks in optical lithography are commonly generated by so-called e-beam writers, that is, the desired structures are directly written onto the photoresist coated mask-substrate, which is subsequently processed. The electron beam can only be positioned on discrete grid point. Thus, the second, discrete representation of the

mask geometry (demonstrated in Figure 2(b)) resembles this method. The mask is divided into pixels with size Δx that is determined by the resolution of the mask writer. Each pixel takes pre-defined complex transmission values t_i (binary (BIM): $t_1 = 0.0 + 0.0j$, $t_2 = 1.0 + 0.0j$; phase-shift (PSM): $t_1 = 0.0 + 0.0j$, $t_2 = 1.0 + 0.0j$, $t_3 = -1.0 + 0.0j$).



(a) continuous (b) discrete

Fig. 2. Continuous and discrete representation of the mask geometry: gray areas symbolize intransparent chromium covered areas of the mask; w_m width of main feature, w_{ai} – width of assist feature i, d_i distance of assist feature i, Δx – grid resolution.

Additionally, the geometry of the illumination source is optimized. In this work the source is restricted to annular-type illumination. As shown in Figure 1(b), the source is presented by continuous values of the inner and outer radii of the annular illumination: σ_{in} and σ_{out}.

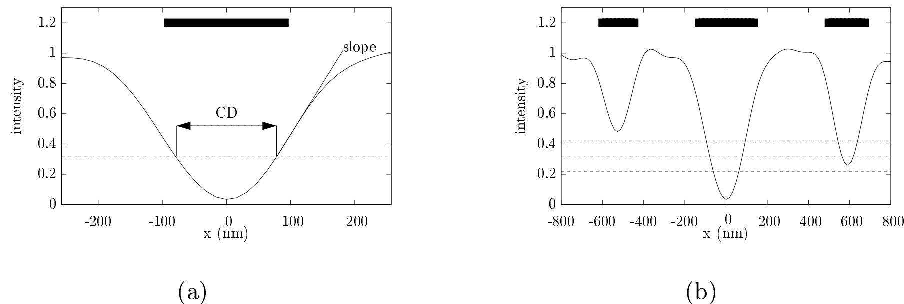
Thus, in the continuous case the optimizer typically has to deal with 3 (one main feature and one assist) to 5 (one main feature and two assists cf. Figure 2(a)) mask parameters and 2 illumination parameters, all of which are real numbers. In the discrete representation case the continuous mask parameters are replaced by more than 100 discrete parameters, taking 2 (BIM) or 3 (PSM) transmission states.

3.2 Fitness Function

For the purpose of this paper, the photoresist is assumed to operate as an intensity threshold detector. After development the photoresist is removed at all positions where the image intensity exceeds a certain threshold value. The threshold intensity can be determined for certain reference features or according to a specific resist process. In this paper, all intensity and threshold values are normalized with respect to the intensity which is obtained with a fully transparent mask. The evaluation of lithographic processes has to take several effects into account, these include a good imaging performance, but also manufacturability issues. Following criteria should be met (for a detailed discussion see also [8]):

CD criterion (ΔCD): The critical dimension (CD) specifies the size of the printed main feature. The deviation between the size of the printed feature at the threshold intensity CD_p and the target CD_t provides the first component of the fitness function.

Slope criterion (SC): This criterion ensures, that good solutions should be tolerant against variations in exposure doses. Thus, the slope of the normalized intensity curve at the edges of the target feature is to be maximized (cf. Figure 3(a)).



(a) (b)

Fig. 3. Evaluation of lithographic images: (a) main feature: critical dimension (CD), i.e., size of feature at a specified threshold intensity (dashed line), the (average) slope is evaluated at the nominal edges of the feature; (b) global evaluation: band criterion, which ensures that side lobes are only printed where required.

Band criterion (BC): In order to ensure that only the desired features are printed, those solutions whose side lobes exhibit low an intensity are punished. On the other hand, the feature's intensity should not exceed a certain value to avoid exposed spots within areas that are to be printed. Therefor a band around the threshold value (as shown in Figure 3(b)) has been introduced. Solutions that violate this criterion are punished. The band size in this work is set to $\pm 30\%$.

Mask manufacturability criterion (MC): For manufacturability reasons it is disadvantageous to have a large number of regions with alternating transmission. Therefore, for the discrete mask representation this criterion assigns a higher fitness to masks with less fragmented regions of the same transmission value.

In order to provide a stable process minor variations of the focus position should have little impact on the imaging performance. Therefore, the former criteria are evaluated for different focus settings.

All of these objectives are combined into one fitness function, yielding a scalar pay-off value:

$$\mu(\boldsymbol{x}) : \boldsymbol{x} \mapsto \frac{w_{\mathrm{MC}} \cdot \mu_{\mathrm{MC}}(\boldsymbol{x}) + \sum_{i=1}^{n_f} \frac{w_{\Delta\mathrm{CD}} \cdot \mu_{\Delta\mathrm{CD}}(\boldsymbol{x}) + w_{\mathrm{SC}} \cdot \mu_{\mathrm{SC}}(\boldsymbol{x}) + w_{\mathrm{BC}} \cdot \mu_{\mathrm{BC}}(\boldsymbol{x})}{n_f}}{w_{\mathrm{MC}} + w_{\Delta\mathrm{CD}} + w_{\mathrm{SC}} + w_{\mathrm{BC}}}, \quad (1)$$

where μ_{MC}, $\mu_{\Delta\mathrm{CD}}$, μ_{SC}, and μ_{BC} denote the results of the former criteria and w_{MC}, $w_{\Delta\mathrm{CD}}$, w_{SC}, and w_{BC} symbolize the corresponding weights. The number of focus settings, which each solution is evaluated for, is denoted by n_f.

The complete evaluation process (which takes a few milliseconds) works as follows: (1) The phenotypic representation is directly used for the manufacturability criterion, (2) the so-called aerial image is calculated [6], (3) slope and band criteria are evaluated, and (4) a threshold model is used to determine the resulting structure in the resist, which is used to evaluate the CD criterion.

4 Genetic Algorithm

The optimization routine used in this work is a simple generational genetic algorithm, using a single population throughout the entire evolution. The main

reasons for chosing a GA are: Placement of sub-resolution resists can rarely be done intuitively. Therefore, finding start values is very time consuming, rendering most (local) optimizers inapplicable. Furthermore, the large number of parameters (especially in the binary mask representation) proved to be hard for most analytic optimizers, since no properties of the search space can be utilized. Last but not least, parallelization of GAs is a relatively uncomplicated task, yielding a scalable procedure suited for cluster-computing. The GA was developed at Fraunhofer IISB and has been integrated into the in-house simulation toolkit [9]. Chromosomes in are coded as bit strings. The genetic operators are two-point-crossover, single-point mutation, and as selection operator fitness-proportional roulette wheel selection, binary tournament selection, and restricted tournament selection are applied. In order to maintain the currently best solution, elitism is performed, additionally.

Restricted Tournament Selection. In contrast to conventional selection operators, where each individual is competing against one another, with restricted tournament selection individuals only replace solutions which have a similar bit string [10]. Thus, this selection operator is not only well suited for multi-modal optimization tasks, but will also maintain a high level of diversity within the population. Two individuals A and B are randomly selected, recombination and mutation are performed, yielding offsprings A' and B'. For each of the both children, w individuals are randomly selected (window). First w individuals are compared to child A', the individual with the shortest distance to A' (A'') is compared with A''s fitness. If A' exhibits a higher fitness than A'' it replaces A''. The same procedure is applied to B'.

Chromosome Representation. Continuous values, such as the parameters for the continuous mask geometry representation and the inner and outer radii of the illumination system (cf. Section 3.1) are modeled as follows: The parameter's domain, that is, upper and lower bounds of the parameter (u and l) and the required number of decimal places (d), is specified in advance. The number of bits taken by this parameter in the chromosome can then be calculated by shifting the parameter's domain ($u-l$) by the required number of decimal places, and computing the required (up-rounded) exponent of radix two ($\ell := \lceil \log_2(u-l)10^d + 1)\rceil$).

Any real number x can now be encoded by subtracting the offset (lower bound l) and dividing it by the range of the parameter ($u-l$); yielding numbers ranging from 0 to 1. The result is then scaled by the maximum number that can be represented by this parameter's part of the bit string ($2^\ell - 1$): $a_{10}(x) := \frac{x-l}{u-l} \cdot (2^\ell - 1)$. The binary representation ($a_2(x) := \text{bin}(a_{10}(x))$), finally, yields the parameter's bit encoding in the chromosome. Thus, decoding of the parameter is performed as follows: $x := \frac{\text{dec}(a_2)}{2^\ell - 1}(u-l) + l$.

Parameters for the discrete mask representation are converted in a straightforward manner. For binary masks each grid point can only take two transmission values. Thus, only one bit is required. However, phase shift masks involve three values. In this case, two bits are required. Although this representation causes the occurrence of four values (where only three are valid, and consequently one

setting punished by the fitness function), it proved to be feasible. Any other encoding scheme, such as expanding the allele alphabet by an additional value, did not improve the convergence behavior.

5 Results

The optimization courses presented in this section have been conducted using a light's wavelength of $\lambda = 93\,\mathrm{nm}$ and a numerical aperture of $\mathrm{NA} = 0.7$. In a first experiment the objective of the optimization is to find the optimum mask for parameters for a lines/spaces pattern with a linewidth of 110 nm and a pitch between the lines of 760 nm. A reference process window (exposure dose vs. defocus) is defined by a dense lines and spaces pattern (same linewidth, pitch = 220 nm) with a threshold intensity of 0.314. This experiment was performed using the continuous mask representation: Three assist features have been placed between the two main features, one of which is placed in the center between both features, the remaining two assists have identical sizes and the same distance from the main feature. The goal was to find optimal width and distance parameters as depicted in Figure 4(a). The ranges for the four parameters are set as follows: width of main feature: $w_m \in [40\,\mathrm{nm}, 100\,\mathrm{nm}]$, width of center assist $w_c \in [40\,\mathrm{nm}, 80\,\mathrm{nm}]$, width of the other two assists $w_a \in [40\,\mathrm{nm}, 80\,\mathrm{nm}]$, and distance of theses assists $d_a \in [40\,\mathrm{nm}, 300\,\mathrm{nm}]$. For this example the illumination parameters are not varied but fixed to $\sigma_{\mathrm{in}} = 0.5$, $\sigma_{\mathrm{out}} = 0.7$.

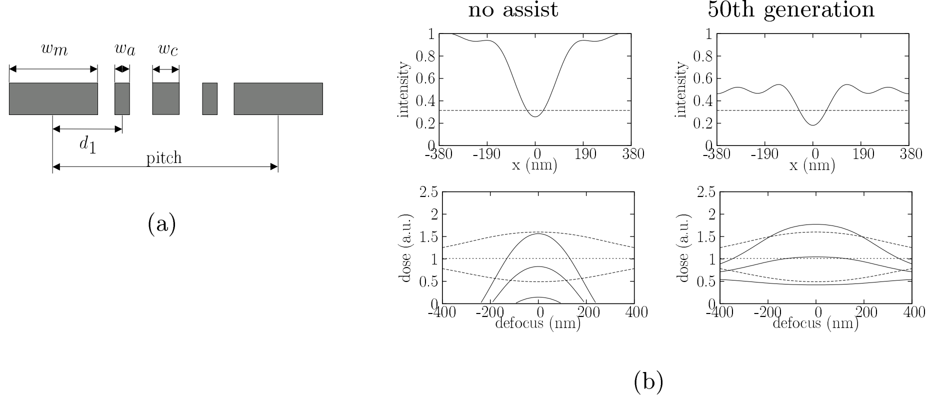


Fig. 4. (a) Mask geometry with four variables; (b) Improvement of the image performance obtained with the continuous mask representation, upper row: images at 250 nm lower row: process windows of the semidense feature (solid lines) compared to process widows of dense features (dashed lines), left column: isolated line without assist, right column: after 50 generations.

The number of individuals is 51; experiments with larger population numbers did not improve the convergence behavior – yet, in case of restricted tournament selection a remarkable slowdown could be observed. Crossover probability is set to 0.6, mutation probabilities are 0.001 and 0.005, and the number of generations is 500. As shown in Table 1 several settings are tested. In order to verify the

selection operator	window size	mutation rate	Continuous average fitness	Continuous lowest fitness	Continuous best fitness	Discrete average fitness	Discrete lowest fitness	Discrete best fitness
roulette wheel		0.001	55.58	55.33	56.42	28.24	26.94	28.88
roulette wheel		0.005	55.93	55.35	56.42	28.60	26.51	29.80
binary tournament		0.001	55.52	55.35	55.67	24.37	22.84	25.49
binary tournament		0.005	55.60	54.78	56.42	29.45	28.52	29.80
restricted tournament	20	0.001	56.16	55.60	56.42	28.91	28.10	29.80
restricted tournament	20	0.005	56.42	56.42	56.42	29.68	28.99	30.08
restricted tournament	50	0.001	56.42	56.42	56.42	29.64	29.01	29.80
restricted tournament	50	0.005	56.35	56.11	56.42	29.57	28.77	29.80

Table 1. GA settings for both the continuous and discrete mask representation. Five runs are performed for each setting. The window size specifies the number of compared individuals with restricted tournament selection (see Section 4). The maximum, minimum, and average fitness of the five runs' best solution are listed.

reproduceability of the results, five runs are conducted for each setting; each run takes about 5 – 6 minutes on a five node cluster with each machine having two 2.66 GHz Pentium 4 processors. Although with all settings the GA results do not differ significanctly, it is noticeable that restricted tournament selection proved to provide reproduceable results.

The improvement of the imaging performance after a certain number of generations is depicted in Figure 4(b). In the first row aerial images at a defocus of 250 nm are presented. The dashed line indicates the threshold intensity which is used for the image evaluation. The lower row shows process windows for the optimized mask (pitch = 760 nm, solid lines) compared to dense lines/spaces (dashed lines pitch= 220 nm). On the left, simulation results for the mask with no assist features are shown. The minimum intensity almost reaches the threshold value. Also, the overlap between the process windows of semidense and dense patterns is very small, i.e., semidense and dense patterns can only be printed simultaneously within a small range of dose and focus settings. After 50 generations the GA has already provided mask patterns that result in a highly improved imaging performance (result: $[w_m, w_a, w_c, d_1] = [125\,\text{nm},\ 50\,\text{nm},\ 55\,\text{nm},\ 215\,\text{nm}]$). Both quality and overlap are noticeably improved.

The same problem is approached with the discrete mask representation. As illustrated in Figure 5, in the first generations the GA is assembling clusters of lines. In the 10th generation the mask mainly consists of thin lines and spaces. However after a number of iterations the manufacturability criterion leads to larger regions. It appears, that in the following steps the GA is combining these clustered patterns. As the clustering phase requires some time, the number of generations has to be increased. In this example the GA was run with 2000 generations, computing time (using the same computer set-up as in the previous example) scaled up to about 45 minutes. After 800 generations the GA yields a mask geometry which guarantees a good imaging performance and which resembles the mask obtained from the continuous optimization approach. The different settings tested with this experiment are listed in Table 1.

The discrete representation of mask geometries can be used to compare different mask types and illumination options. This is demonstrated in Figure 6,

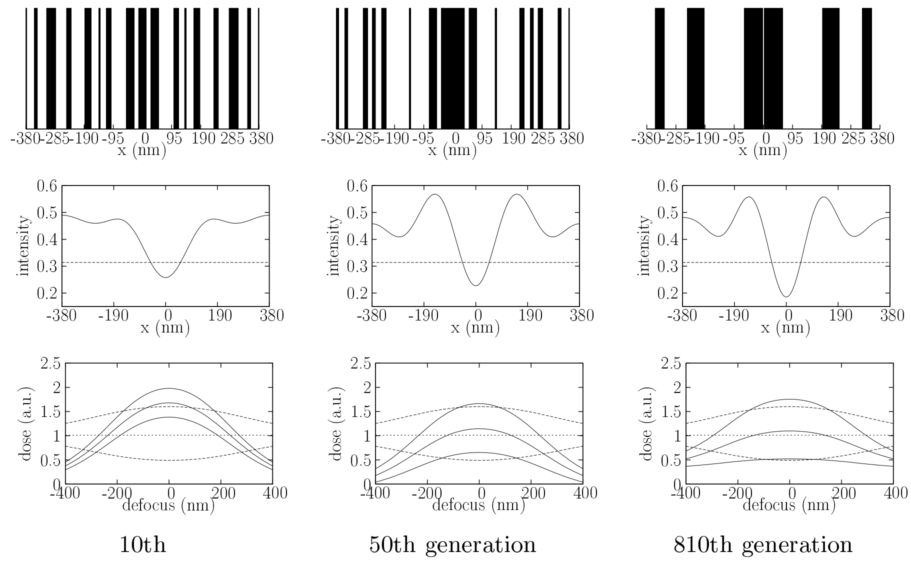


Fig. 5. Improvement of the image performance obtained with the GA for the discrete mask representation, upper row: mask geometry, center row: images at 250 nm defocus, lower row: process windows.

which shows the best mask layouts, annular illumination settings, and resulting process windows. The only information the GA was provided with, was the size of the addressable grid on the mask ($\Delta x = 5\,\text{nm}$), the complex transmission values of the mask (binary: $1+0j$, $0+0j$; PSM: $1+0j$, $-1+0j$, $0+0j$), and the type of illumination (annular). In case of binary masks the optimization of the fitness function results in a typical sub-resolution assisted mask configuration and strong off-axis illumination. For the PSM the algorithm suggests a standard phase edge design and coherent illumination. The process window for the PSM is considerably larger than for binary masks.

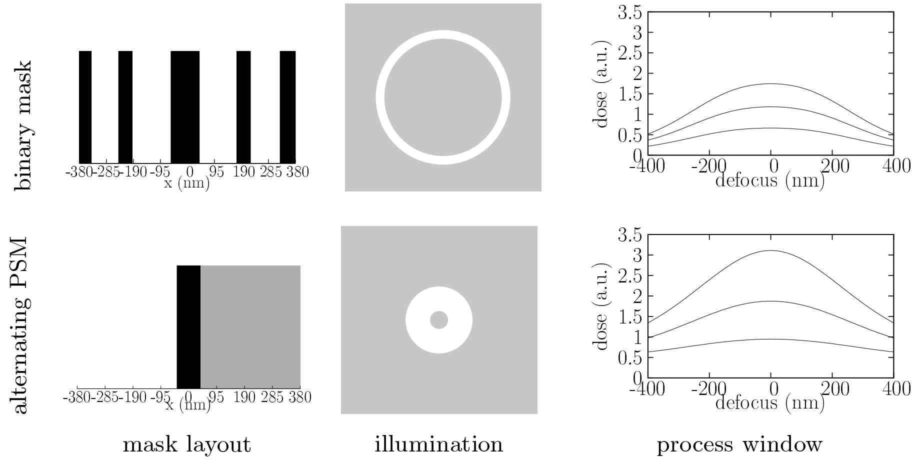


Fig. 6. Optimized imaging performance for 90 nm isolated lines for the discrete mask representation, upper row: binary mask (BIM), lower row: alternating phase shift mask (PSM).

6 Conclusions and Future Work

The proposed optimization procedure using a genetic algorithm identifies optimal imaging conditions without any additional a-priori knowledge about lithographic processes. It is applied to automatically place sub-resolution assists for lines/space patterns with a specified range of pitches. The procedure identifies the best mask layout and illumination conditions for the generation of isolated lines with different mask types. Especially for the binary mask representation utilization of the restricted tournament selection operator appears advantageous. Further development of the GA will aim at improving the convergence behavior by implementing so-called "competent GA" enhancement techniques [11]. This can be particularly useful to speed up the the first phase (finding manufacturable masks) of the binary representation case.

Further work will also be necessary to improve the description of the mask and source geometries. Additional imaging criteria such as the mask error enhancement factor (MEEF) and the aberration sensitivity have to be included in a generalized form of the merit function. Moreover, the proposed optimization procedure has to be combined with more advanced lithography simulation models.

References

1. H. Levinson. Optical Lithography. In P. Rai-Choudhury, editor, *Handbook of Microlithograph, Micromachining and Microfabrication*, volume 1. SPIE Press, 1997.
2. J.W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1992.
3. H.H. Hopkins. On the diffraction theory of optical images. In *Proc. Roy. Soc. A 217*, page 408, 1953.
4. M. Mansuripur. Distribution of light at and near the focus of high numerical aperture objectives. *J. Opt. Soc. Am.* ***A 3***, page 2086, 1986.
5. D.G. Flagello and T.D. Milster. High numerical aperture effects in photoresist. *Applied Optics 36*, page 8944, 1997.
6. A. Erdmann and W. Henke. Simulation of optical lithography. In *Optics and Optoelectronics – Theory, Devices and Applications, Proc. SPIE* ***3729***, page 480, 1999.
7. A. K. Wong. Resolution Enhancement Techniques in Optical Lithography. In *Tutorial Texts in Optical Engineering,* **TT47**. SPIE Press, 2001.
8. A. Erdmann, R. Farkas, T. Fühner, B. Tollkühn, and G. Kókai. Mask and Source Optimization for Lithographic Imaging Systems. In *Proc. SPIE 5182*, 2003. in print.
9. B. Tollkühn, T. Fühner, D. Matiut, A. Erdmann, G. Kókai, and A. Semmler. Will Darwin's Law Help Us to Improve Our Resist Models. In *Proc. SPIE 5039*, page 291, 2003.
10. G.R. Harik. Finding multimodal solutions using restricted tournament selection. In Larry Eshelman, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 24–31, San Francisco, CA, 1995. Morgan Kaufmann.
11. D.E. Goldberg. *Evolutionary design by computers*, chapter 4. The Race, the Hurdle, and the Sweet Spot: Lessons from Genetic Algorithms for the Automation of Design Innovation and Creativity. Morgan Kaufmann, San Francisco, CA, 1999.