Demonstrating the Benefits of Source-Mask Optimization and Enabling Technologies through Experiment and Simulations

David Melville[a], Alan E. Rosenbluth[a], Kehan Tian[b], Kafai Lai[b], Saeed Bagheri[a], Jaione Tirapu-Azpiroz[b], Jason Meiring[b], Scott Halle[e], Greg McIntyre[e], Tom Faure[f], Daniel Corliss[b], Azalia Krasnoperova[b], Lei Zhuang[b], Phil Strenski[a], Andreas Waechter[a], Laszlo Ladanyi[a], Francisco Barahona[a], Daniele Scarpazza[a], Jon Lee[a], Tadanobu Inoue[c], Masaharu Sakamoto[c], Hidemasa Muta[c], Alfred Wagner[a], Geoffrey Burr[d], Young Kim[d], Emily Gallagher[f], Mike Hibbs[f],

[a]IBM T.J. Watson Research Center, Yorktown Heights, New York, 10598;
[b]IBM Semiconductor Research and Development Center, Hopewell Junction, New York 12590
[c]IBM Research - Tokyo, Yamato, Kanagawa, Japan
[d]IBM Almaden Research Center, Almaden, CA
[e]IBM Research at Albany NanoTech, 255 Fuller Road, Suite 274, Albany NY 12203
[f]IBM Mask House, Burlington, VT

Alexander Tritchkov[h], Yuri Granik[g], Moutaz Fakhry[g], Kostas Adam[g], Gabriel Berger[g], Michael Lam[g], Aasutosh Dave[g], Nick Cobb[g].

[g]Mentor Graphics Corporation, San Jose, CA
[h]Mentor Graphics Corporation, Wilsonville, OR

ABSTRACT

In recent years the potential of Source-Mask Optimization (SMO) as an enabling technology for 22nm-and-beyond lithography has been explored and documented in the literature.[1–5] It has been shown that intensive optimization of the fundamental degrees of freedom in the optical system allows for the creation of non-intuitive solutions in both the mask and the source, which leads to improved lithographic performance. These efforts have driven the need for improved controllability in illumination[5–7] and have pushed the required optimization performance of mask design.[8,9] This paper will present recent experimental evidence of the performance advantage gained by intensive optimization, and enabling technologies like pixelated illumination. Controllable pixelated illumination opens up new regimes in control of proximity effects,[1,6,7] and we will show corresponding examples of improved through-pitch performance in 22nm Resolution Enhancement Technique (RET). Simulation results will back-up the experimental results and detail the ability of SMO to drive exposure-count reduction, as well as a reduction in process variation due to critical factors such as Line Edge Roughness (LER), Mask Error Enhancement Factor (MEEF), and the Electromagnetic Field (EMF) effect. The benefits of running intensive optimization with both source and mask variables jointly has been previously discussed.[1–3] This paper will build on these results by demonstrating large-scale jointly-optimized source/mask solutions and their impact on design-rule enumerated designs.

Keywords: SMO, Source-Mask Optimization, Source Optimization, Mask Optimization, off-axis illumination, pixelated illumination, programmable illumination reticle enhancement technology, RET, optical proximity correction, OPC, global optimization, linear program

1. INTRODUCTION

Source-Mask Optimization (SMO) as first described in ref.[2] is an intensive optimization methodology that seeks to determine the globally-optimal solution for the optical wavefront at the wafer given the limitations imposed by a conventional projection-lithography system. SMO achieves this by carefully breaking an intractable global problem into a series of global, local, and heuristic optimization steps. These steps, of global mask (with heuristics), global source, local source-mask (or joint), and wavefront engineering have been previously described[1,2,5,8] and additionally covered by other

Further author information: (Send correspondence to D.O.S.M)
D.O.S.M: E-mail: domelvil@us.ibm.com, Telephone: 1 914 945 1222

Optical Microlithography XXIII, edited by Mircea V. Dusa, Will Conley, Proc. of SPIE Vol. 7640, 764006 · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.846716

groups.[3,4] Figure 1 shows a pictorial view of the stages as 1) optimization of a source + optical-domain mask, followed by 2) a Wavefront Engineering (WE) step to produce a manufacturable mask that can reproduce the optical wavefront as optimized by the previous step. As no further improvement in the Numerical Apeture (NA) or optical wavelength is likely until the establishment of Extreme Ultra-Violet (EUV) as a production technology, Resolution Enhancement Technique (RET) solutions must push the limits of Deep Ultra-Violet (DUV) lithography systems through computationally intensive optimization techniques like SMO. The paradigm of SMO represents in some ways the next step in a decade long evolution of increasing reliance in computational solutions to the limitations in the optical system. However, in many ways SMO is a significant departure from the solutions of past nodes. Delivery of intensively optimized solutions requires support from a number of enabling technologies, such as configurable illumination, mask fabrication, and design rules. Now that the technology of SMO is maturing, the benefits can be experimentally explored and the advantages of SMO understood.

As lithography engineers explore potential patterning solutions at 22nm nodes and beyond, they are met with many trade-offs that can be complicated to understand, and often require numerous cycles of learning. Brute force data-collection efforts can be undertaken to try to explore the space, but visualizing multi-dimensional data to determine trade-offs can often be difficult. One of the key strengths of intensive optimization is that it provides automatic search through complicated optimization spaces. This can lead to faster learning in both design and lithography efforts.

This paper details the development efforts to deliver the benefits of SMO to the 22nm technology node and beyond. The paper begins in Section 2 by detailing experimentally achieved benefits of SMO with on-wafer demonstrations of SMO-derived Source and Mask solutions. Section 3 gives further exploration of the technology that produces the benefits. Finally, in Section 4, critical aspects for how SMO engages with designs of 22nm technologies.

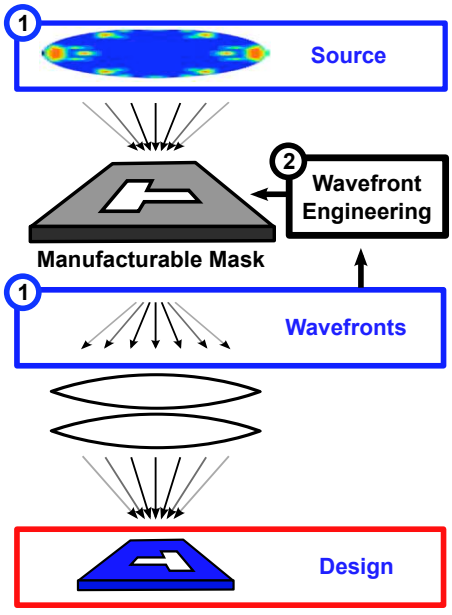


Figure 1: SMO optimization stages.

1.1 SMO Overview

SMO provides intensively optimized wave distributions that illuminate both the mask and the wafer, exploiting all possible benefit from the available degrees of freedom in the band limited exposure process. The illumination patterns ("source") that SMO generates will typically have a far more detailed structure than previous methods could consider, and the mask shapes that are designed by SMO to diffract the optimal imaging waves to the wafer are likewise highly non-intuitive, differing radically from conventional mask solutions.

Full SMO is arrived at by combining source optimization with optimization of the mask as filtered by the lens band limit, see Rosenbluth *et al.*[2] This reference shows that the SMO optimization problem is nonlinear and highly non-convex. To solve it, global (sequential mask source, with heuristics) and local (joint mask source) methods of mathematical optimization and high performance computing are employed. The objectives and constraints are grounded in image formation theory, as constrained by a broad set of engineering considerations involving mask/source manufacturability. (Figure 1 previously showed a pictorial representation of the algorithm).

In its simplest form, input to SMO is a set of target (circuit design) patterns to be imaged. The output of SMO is an illumination source pattern and a set of mask patterns which have been optimized to provide maximum optical performance for a set of input patterns. This differs from conventional mask design approaches such as model-based optical proximity correction which historically have been iterative feedback systems rather than true optimization solutions of the problem at hand.

In developing an SMO solution, one needs to take into account numerical processing methods that are embodied in software, the ability of the exposure tool to generate the requested source, the ability of the mask fabricator to generate complex mask shapes, and the application of the solution by the user. The approach we are pursuing relies on concurrent development of the software, the exposure tool, mask fabrication, and the semiconductor technology. The user is the

semiconductor technology development team. The results we are publishing represent output from different stages of the concurrent development.

2. EXPERIMENTAL BENEFITS

The following sections outline experimental efforts that demonstrate the ability of SMO to deliver lithographic benefit for 22nm technology. We cover three scenarios of SMO resolution enhancement where the benefit is derived from the illumination (Section 2.1), the mask(Section 2.2), and both in combination (Section 2.3).

2.1 Benefit from SMO Illumination Solutions

The process of designing a source for a lithographic process has evolved significantly in recent years[1,7,10] and as a result intuition for solutions that deliver critical advantages has been handed off to computational techniques. The design space for illumination has grown from a few potential degrees of freedom provided by traditional Diffractive-Optical-Element (DOE) to the hundred's, or even thousands, of degrees of freedom in the next generation of pixelated or programmable sources. Computational efforts, through either brute-force simulation or insensitive optimization, are often the only way to find solutions that deliver the best that the tooling can offer.

Here we demonstrate the benefit of an SMO-designed complex illumination source for a 22-nm metal layer. Previously, the practical feasibility of using pixelated DOE solutions was outlined[6] and demonstrated many of the critical aspects of freeform illumination considered. In this paper, Section 3.2 goes into some additional detail on recent developments in freeform source analysis.

The optimizations goal was to enable the printing of critical SRAMs along with a range of horizontal and vertical gratings. Of the gratings the 1x, 1.75x, and 2x horizontal, along with the 1x vertical periods were the most important to print and had been previously shown to lack printing fidelity with Optical Proximity Correction (OPC) and traditional illumination. The high-level goal was to avoid the process moving to a double-exposure requirement for enabling the minimum horizontal and vertical periods. The solution that was derived from SMO studies was carefully investigated for its potential benefit. Although intensive optimization of the mask was thoroughly studied, the benefit over a traditional RET was found to mainly come from the source.

Once the optimized illumination was developed the solution was delivered as a pixelated DOE. The following sections talk in detail about the delivery of the source hardware, experimental results, and simulation verification of the performance advantage.

2.1.1 Experimental Result

The robustness of freeform illumination manufacturability is critical to the success of intensively optimized source solutions. It would be counter-productive for SMO solutions to have more degrees of freedom than traditional DOE, but then give up robustness or stability due to their manufacturability. To investigate the quality of the produced source the source design was studied. Figure 2 shows pupilgrams of the metal layer source for the target, and the measured experiment DOE on linear and log scales. Using the measure pupilgram from the pixelated the simulated performance has been analysed and demonstrated as being acceptable.

Figure 3(a) shows the pupilgrams of the illuminations used to compare the traditional and pixelated experimental results. The details of the experimentally determined process windows are shown in Figure 3(b). The minimum "right-way" has slightly worsened in terms of exposure latitude, which is the SMO trade-off that is expected in this case. The trade-off has allow the improvement in the remainder of the critical features. In the case of the mid-range horizontal pitch, the traditional solution is unable to provide any process window while the SMO solution presents strong performance.

The experimental results are now provided to demonstrate the reality of the performance for a pixelated illumination and the delivery of benefit on wafer. Figure 4(a) shows traditional and pixelated-DOE comparison SEMs for best case exposure conditions of a difficult-to-print grating at two planes of de-focus. The benefit of the pixelated-DOE is calculated as a 65% gain in Depth-of-Field (DOF).

Figure 4(a) shows further detail of the simulated benefit for other critical features for the metal layer. For the traditional and pixelated-DOE the DOF and Mask Error Enhancement Factor (MEEF) performance is compared. The SMO solution has found a case where the minimum horizontal pitch has lost slight quality its DOF as a trade-off to bring up the performance off all other features. Most notable is the MEEF reduction of the 2x horizontal and minimum vertical.

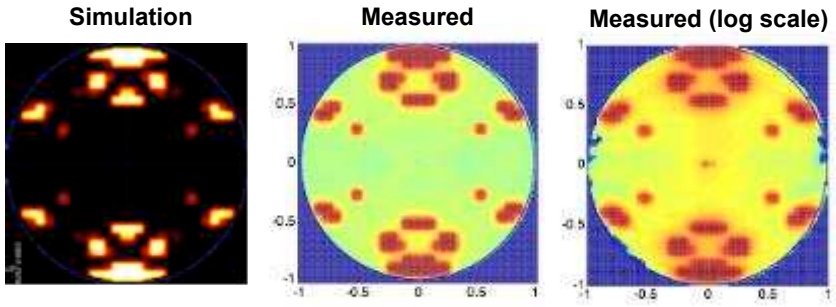
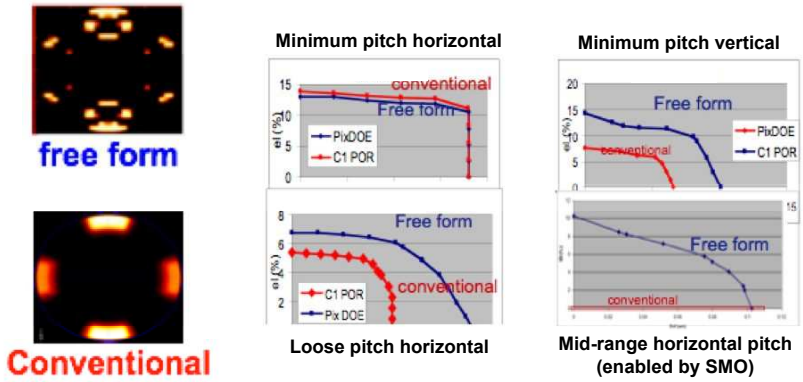
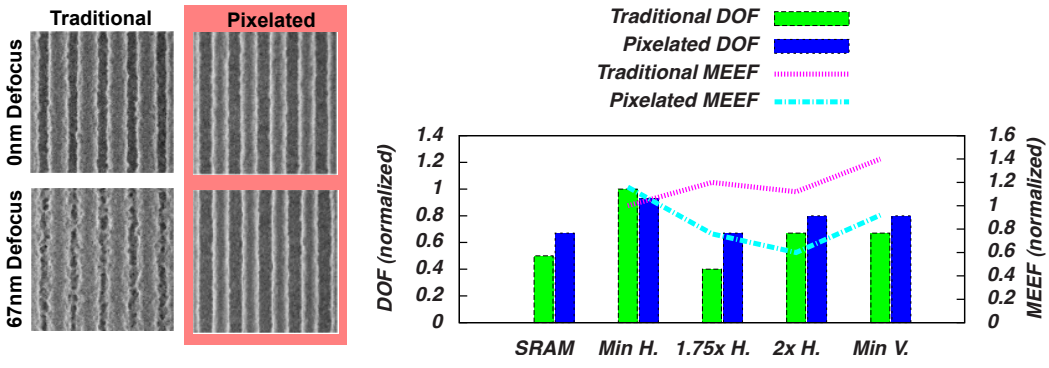


Figure 2: Pupilgrams of the metal layer source for the target, and the measured experiment DOE on linear and log scales.



(a) (b)

Figure 3: Experimental performance for metal source. (a) Traditional and freeform measured pupilgrams. (b) One-dimensional imaging shows minimal degradation of dense patterns in the preferred orientation (right-way), while showing significant improvement of loose and non-preferred orientation features.



(a) (b)

Figure 4: Experimental performance for metal source. (a) Experiment result for metal layer on 1.75x minimum horizontal grating. (b) The experimental performance comparison between the pixelated and traditional illuminations for a 1-X metal layer.

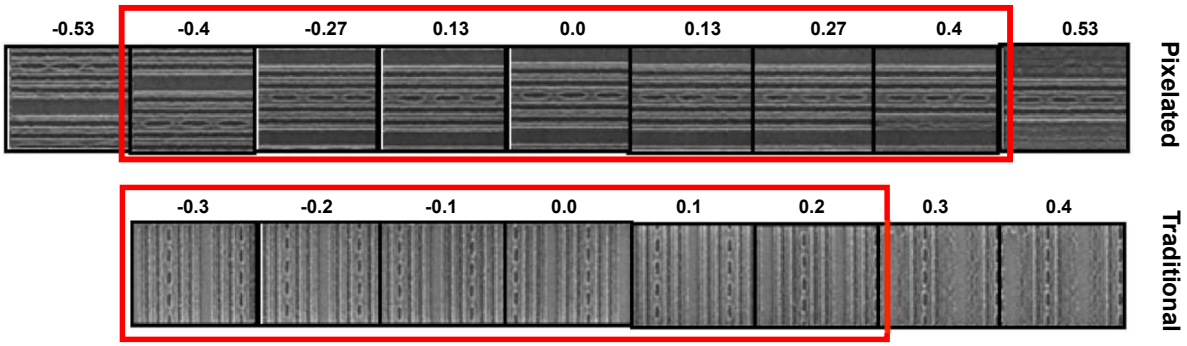


Figure 5: Performance of critical SRAM through focus for traditional and pixelated illumination. The de-focus values are normalized. In this focus measurement -0.5 to 0.5 is required for the process.

For this metal example, the performance of the SRAM is also improved with a 30% increase in the DOF. Figure 5 shows the images of the SRAM cell through focus. The traditional illumination suffers from a loss of printing fidelity at the line-ends along with lift-off of the horizontal lines.

2.1.2 Performance of Pixelated Illumination for Metal Layer

Figure 6 shows the simulated performance of the metal source through a range of critical horizontal and vertical gratings. Note that while the SMO masks were also optimized when the SMO source was being designed, the performance data refers to the solution quality obtained when ordinary OPC was used to design the masks employed with the various sources. The development of this source design, through SMO techniques, gives critical advantages in the ability to make trade-offs between the printing of required critical two-dimensional (2-D) structures (like SRAM) and critical one-dimensional (1-D) structures.

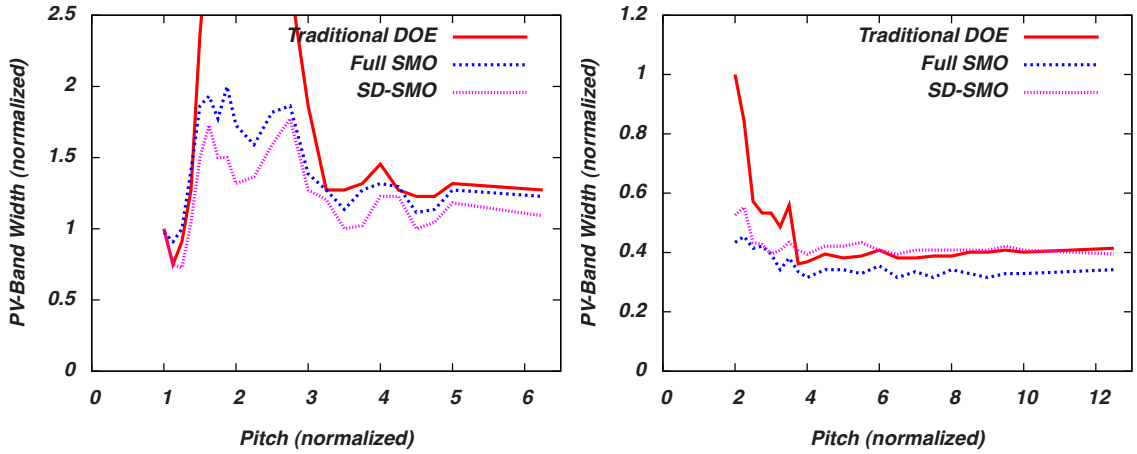


Figure 6: Through pitch performance data for critical (a) horizontal and (b) vertical gratings using the metal source with masks designed by conventional OPC. Data is shown for a source designed by traditional RET methods, and for sources designed by two forms of SMO. Although the performance data was based on OPC masks, the two SMO sources involved co-optimization of the clips that was carried out in two different ways; see text.

Figure 6 additionally shows the OPC performance difference arising from co-optimizing the mask clips in two different ways when designing the source (this source then being used by OPC to design the final masks, and to calculate the plotted

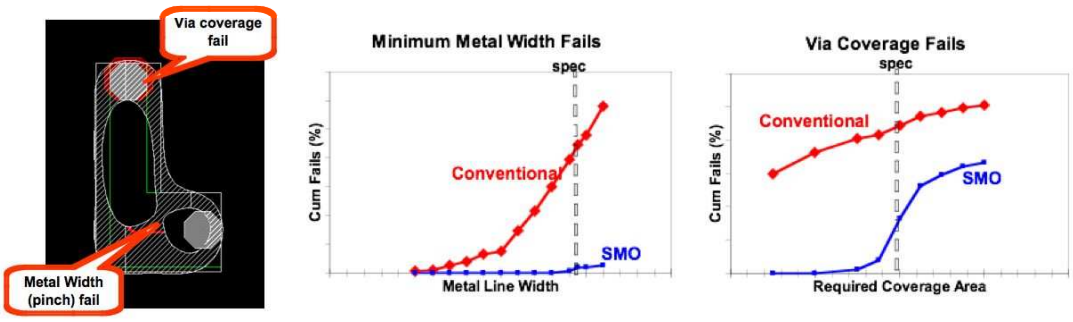


Figure 7: Expansion of metal-layer logic design space through SMO source optimization. Improvements in metal-width/pinch and via-coverage fails.

data). The full set of SMO modules are used for co-optimization in designing the source labeled Full-SMO, while the source labeled SD-SMO is obtained by using only a single module from the full SMO flow during co-optimization. For this optimization problem, the overall OPC performance with the SD-SMO-designed source is comparable to that with the Full-SMO-designed source. The similarity in performance illustrates that some optimizations are clearly dominated by the freedom in the source.

As a result of the improved lithography quality provided by the freeform source the performance of logic patterning also improved. Figure 7 shows the improvement in design space over a conventional solution by looking at critical failures for the metal layer across many configurations. A large set of metal and via clips were analyzed with patterning simulation with the current RET and again with metal source provided by SMO. The cumulative fails for cases at given value of via coverage, and number of widths below a given value were plotted before and after. The SMO source reduced both the number of fails at the proposed threshold, as well as the worst case failure value in both cases by significant amounts.

In conclusion, SMO was able to achieve 30% process window increase on critical features and represents a significant advancement over state-of-the-art. This is example of an intensively optimized pixelated illumination solution brought to reality.

2.1.3 Performance of Programmable Illumination Solution for Metal Layer

As an enabling technology for SMO, Programmable Illuminator (PI) solutions allow a significant reduce time-to-market versus DOE fabrication. They additionally remove the physical limit of concurrent DOE solutions in a scanner, which is a significant advantage over pixelated-DOE's, as this tooling consideration limits the diversity of pixelated-DOE's. Figure 8 demonstrates the experimental performance when the pixelated illumination is replaced by a programmable source illumination.

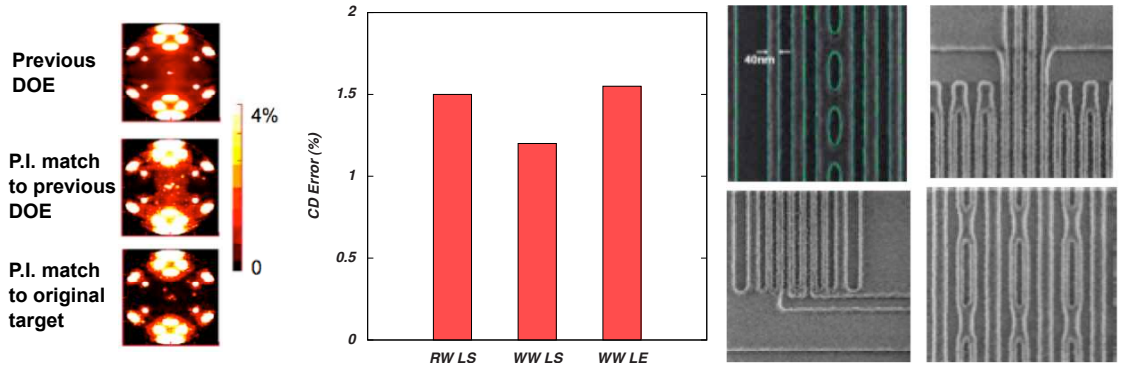
Figure 8(a) shows the measured pupil grams for the pixelated-DOE and a PI that has been setup to match. The match is done for both the DOE and to the original target. The main difference here is that the DOE has more background illumination so matching that low intensity region is more difficult for the PI. The solution without background illumination is a closer match to the desired SMO source solution. Figure 8(b) shows the simulated 3-sigma percent Critical-Dimension (CD) variation for 133 measured pupil grams from the PI. The variation across the multi-say measurement is less that 2%.

2.2 Benefit from SMO Mask Solutions

In the previous section we saw how the performance advantage achieved from optimization of the illumination led to a solution that could deliver acceptable quality in combination with traditional OPC mask optimization. In this section we briefly describe an example where mask optimization provides the benefit and conventional illumination provides an acceptable approximation to the optimal source

The experimental effort put together a wide range of solutions for various SRAMs with mask solutions taking numerous paths to explore the potential of the SMO performance*. The main comparision points are traditional solutions,

*Please note that the experimental results do suffer from some "off target" issues due to early cycle-of-learning for the resist process. This has been taken into account in the analysis.

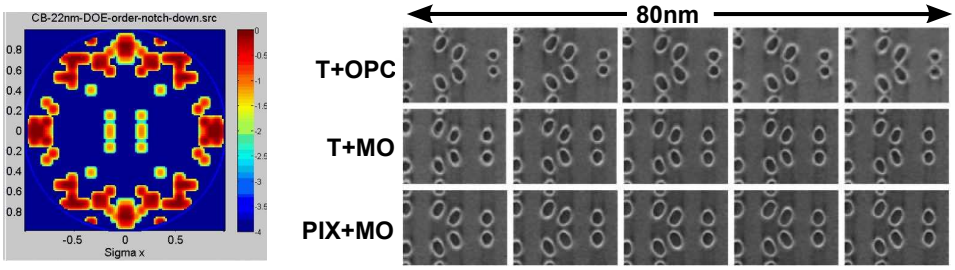


(a) (b) (c)

Figure 8: Experimental performance for metal source. (a) Freeform source has been realized with both DOEs and PI, showing how PI can either match background intensity of a DOE or create a more ideal pupil with minimal background. (b) Matching between measured DOE and PI pupils is less than 2% CD using a simulated CD-based approach. (c) Showing 2D on wafer results shows how source improves 2D image fidelity.

intensive Mask Optimization (MO) (i.e Frequency-Domain Mask-Optimization (FDMO), WE, then Spatial-Domain Mask-Optimization (SDMO)) with traditional sources, and full SMO solution with pixelated source.

Figure 9(a) shows the pixelated illumination on a log scale, while Figure 9(b) shows the through focus Scanning Electron Microscope (SEM) images for the traditional solution ("T+OPC"), SMO mask only ("T+MO"), and full SMO ("PIX+MO").



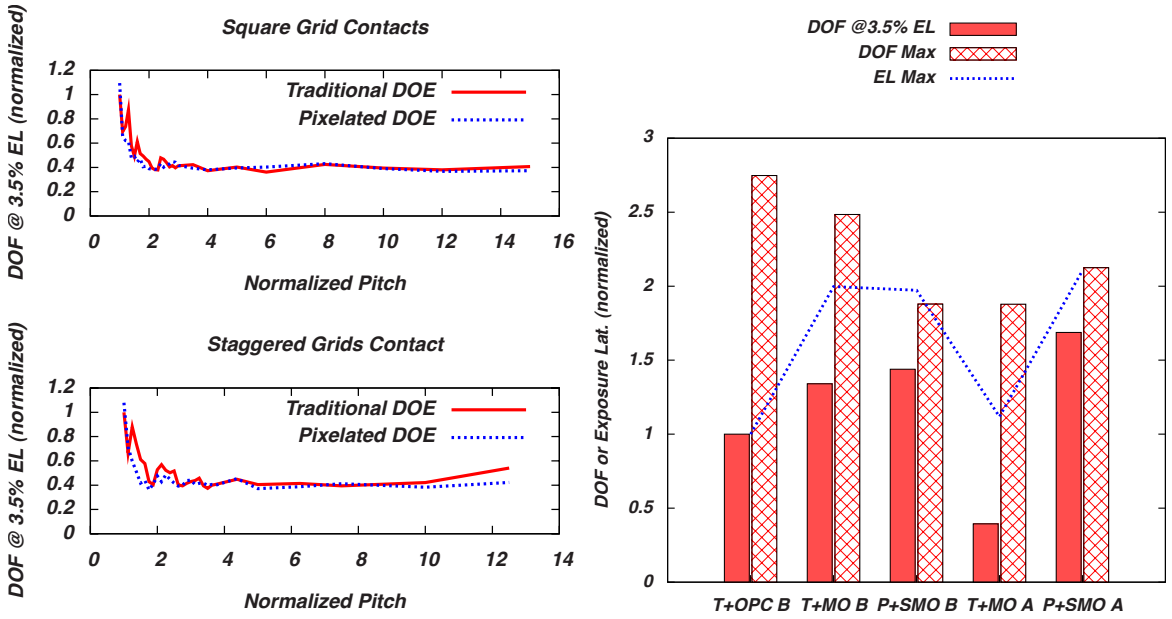
(a) (b)

Figure 9: Experimental performance for contact optimization. (a) Intensivelty optimized source for contact layer. (b) Experimental SEM images for SRAM solutions on a contact layer.

Figure 10(a) shows the comparative experimental performance of the traditional and pixelated illuminations for a range of critical square and staggered contact arrays. Although the performance of the SMO optimized pixelated illumination is better though some forbidden pitch ranges, the general performance is similar.

Figure 10(b) shows the experimental DOF performance of two 22nm technology SRAM's for a contact layer. Results for SRAM "A" and "B" are shown. SRAM "B" is a critical cell, and "A" is an example of a cell that increases the SRAM density limits to push the limits of the optimization. The plots show DOF for an acceptable CD variation with Exposure-Latitude (EL) of 3.5%, the maximum DOF for an individual feature, and the measured EL at the best point of focus. When comparing the performance of OPC versus MO with a traditional illumination we see a 35% increase in DOF performance. There is a decrease in the DOF of the best feature demonstrating that SMO is able to make the trade-off between individual feature quality and overall common process window. The full-SMO of SRAM-B, including the experimental use of pixelated illumination, demonstrates improvements above both OPC and MO, however advantage in this case over traditional illumination with intensive optimization is not particularly significant.

Although in this case the SMO mask was not used directly, the learning from the optimization was able to be transferred into traditional mask technology approaches. SMO-like MO performance cannot be fully acheived with traditional



(a) (b)

Figure 10: Experimental performance for contact optimization. (a) Experimental DOF and EL performance for contact layer. (b) Experimental through pitch performance for contact layer regular and staggered features.

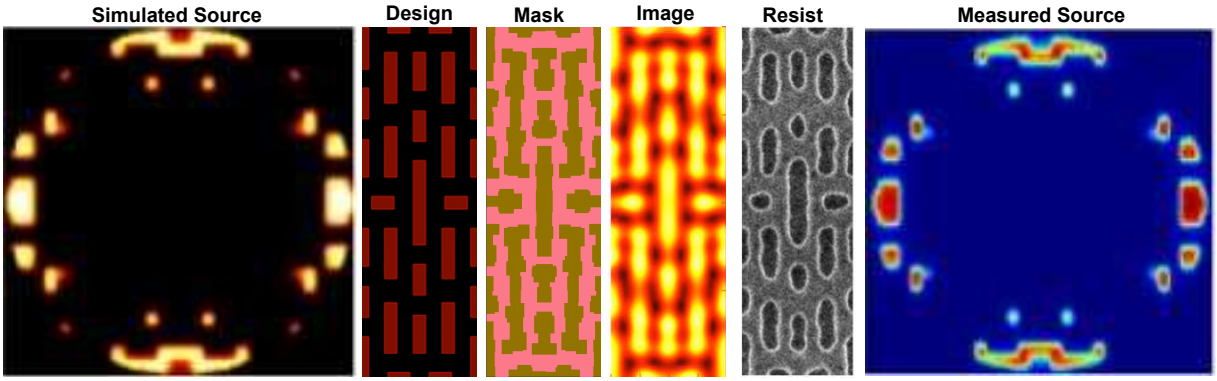


Figure 11: Optimization result through to measured wafer and pupil for programmable for source.

techniques due to the complex nature of the assist features, but the quality of the solution can be partially caputured through examination. In this situation, SMO has delivered a faster cycle of learning. The denser SRAM "A" did not follow this trend and the SMO solution was able to provide a significant boost with the inclusion of the pixelated illumination. The source of the SMO benefit for this layer's SRAMs is dependant on pattern density.

2.3 Benefit from Full SMO Solutions

The benefit from full SMO solutions is tightly coupled to freeform illumination. The introduction of programmable illumination provides a number of key benefits to process development. The two most important are the availability of freeform illumination and the rate at which this freeform illumination can be provided. With the benefit of SMO illumination solutions becoming well documented the rate at which these can be provided becomes very important for the rate of cycles-of-learning. Programmable illumination fulfills these requirements and enables the benefits of SMO techniques.

Figure 11 shows the result of an experiment from the optimized source and mask through to the measure illumination

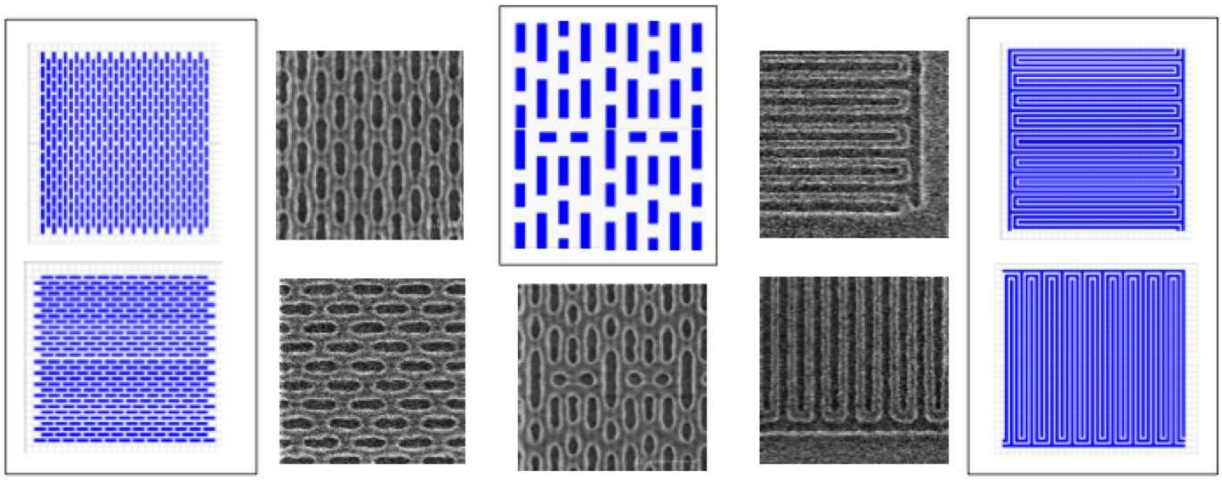


Figure 12: (Color) Concurrent imaging of SRAMs and minimum pitch features for metal layer single exposure solution. Shows both designs and SEM images for five critical patterns.

pattern with on wafer confirmation of the result. This represents early learning of programmable illuminator technology installed on IBM tools. This is the first time demonstration of a programmable illumination result from an optimized SMO source and mask combination. The source for this experiment has been optimized along with a number of critical patterns, but only the SRAM is shown in Figure 11.

Figure 12 experiment demonstrating a non-negative common process window across critical SRAM and minimum pitch features for single exposure metal layer. The solution uses a programmable illuminator with the source design shown in Figure 11. A Pixelated-DOE was also built to demonstrate and compare performance. Figure 13 shows an on wafer comparison for on-wafer results between pixelated-DOE and programmable illumination. The results are identical to within reasonable measurement, which is consistent with expectations.

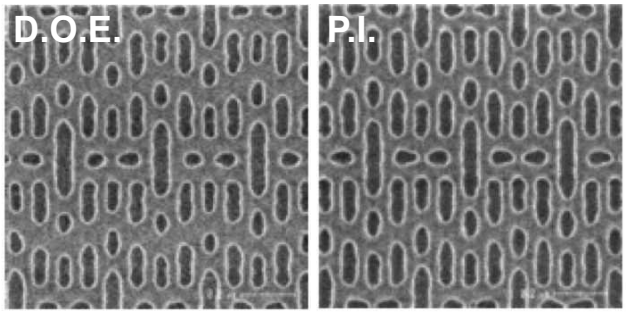


Figure 13: Comparison of on-wafer results for pixelated DOE (left) and programmable illumination (right) for the metal SRAM.

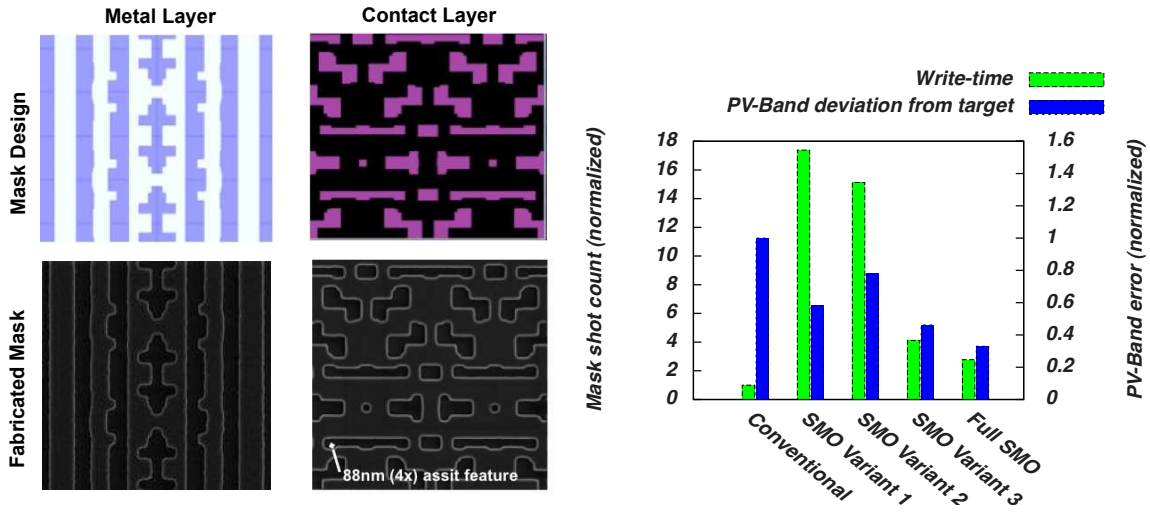
2.4 Mask Fabrication

By optimizing in the optical domain SMO can access the fundamental degrees of freedom in the optical system, thereby achieving maximum lithographic performance. However, as discussed elsewhere, this requires that mask manufacturability requirements be respected, particularly in the concluding "wavefront engineering" steps of SMO where feasible mask shapes are obtained.[1,2]

Mask complexity and associated shot count are an important component of mask feasibility when SMO is employed over large areas. The problem of maximizing attainable image quality is easier to solve if one can resort to a complex set of patterns that need only be fabricated for small-area experiments. However, as we will see, it is possible with sophisticated algorithms to contain pattern complexity to more practically accepted levels with little compromise in lithographic performance.

The masks in figure 14(a) were fabricated with the 32-nm node process for building critical level masks and used the industry standard OMOG (Opaque Mo-Si on Glass) binary mask blank. Figure 14(a) shows the mask designs in the SRAM on the contact and metal masks. The e-beam shot count for the SMO contact design was approximately 3.5x larger versus the conventional OPC design. The shot count for the SMO metal design was only about 10% higher than conventional OPC. Figure 14(a) also contains SEM images of the both masks which shows faithful replication of the SMO mask pattern

design including assist feature sizes as small as 88 nm (4x). E-beam print time of both masks was within acceptable manufacturing limits and was less than 12 hours.



(a) (b)

Figure 14: SEM images, write time, and lithography performance for SRAM masks. (a) SMO mask patterns. Showing mask design and fabricated mask and metal and contact SRAM cells. (b) Mask write times and lithographic performance versus mask optimization methods for contact layer.

Die to database defect inspection was performed using both the standard reticle plane inspection (RPI) mode and a newly available aerial plane inspection mode (API). No unique inspection difficulties were encountered with these masks. During the aerial plane defect inspection mode the pixelated exposure source map was directly input into the inspection tool and was used to successfully generate an aerial image comparison of the reference data and actual mask.

To investigate the impact of SMO technology on mask-writer times, an SRAM contact design was optimized as a test pattern using four SMO variants. An OPC-generated design was also created as a reference. The primary SMO solution (labeled "Full SMO") employed the full set of SMO modules, while the other SMO variants substituted various simplifications into the SMO polygon design process. All mask solutions use the same traditional source, and the optical-domain mask optimization solution is the same for all SMO designs.

Figure 14(b) compares the mask write times (green bars) and lithographic performance (blue bars) of the five mask solutions. The performance metric that is shown (blue bars) measures the maximum displacement of the print contour from target across all important process conditions (including local variation in mask bias). The OPC solution uses simple feature shapes, and employs model-based assist features that are low edge-count in their design, thus leading to very short mask write times. However, the OPC solution is not able to achieve acceptable lithographic performance for this problem.

The test clearly shows that the write-times of the SMO solutions are appreciably longer than that of the OPC design. However, the full SMO solution contains this increase to the range of about a 3X factor, which may still be considered feasible (arising mainly from an increased assist feature count). As a tradeoff, the full SMO solution also achieves the highest lithographic performance found, providing about 3X smaller overall process sensitivity than the OPC solution. The algorithmic simplifications in the SMO process considered here are found to decrease performance in terms of both write-time and image quality compared to the full SMO process; however the quantitative impact ranges from a relatively acceptable degradation of 40-50 percent (variant 3), to larger performance degradations and much longer write times (variants 1 and 2).

3. SIMULATION

The experimental evidence presented in Section 2 demonstrated that benefit can be derived from intensive optimization, but the process by which benefit is provided can range depending on the needs of the problem. A RET designer builds intuition

about RET tools over time, and becomes an expert driver. But this is only sensible if the tools are either simple or carefully represent the complexities of the optimization problem. In an optimization like SMO there has been a considerable jump in the complexity of the tool and algorithms that produce the solution, and as such it is the responsibility of the tool to abstract those complexities in such a way as to give the user intuition about how to deliver solutions. This is critical to allowing the fast uptake of complex technologies

SMO methodologies are by definition taking on significantly more degrees of freedom. At the same time that the optimization is taking on more degrees of freedom there is considerable increase in the coupling of variables in mask due to proximity and through the source. Giving all freedom to an intensive optimizer (as apposed to simple interactive feedback like OPC) leads to potential unwanted side effects if all qualities of a solution are not given attention. If the optimization is not constrained, either directly or through heuristics, then important dimensions to the problem that are not aligned run the risk of giving a self-determined trade-off. In a lithographic problem that nears the resolution of the optical system, leaving this trade-off to the will of the optimizer often leads to unusable solutions. MEEF, Electromagnetic Field (EMF) effects, Line Edge Roughness (LER), Process Window (PW), Process Variation Band (PV-Band), mask manufacturability, and source manufacturability are all examples of problem constraints that need to be taken into account for an advanced solution.

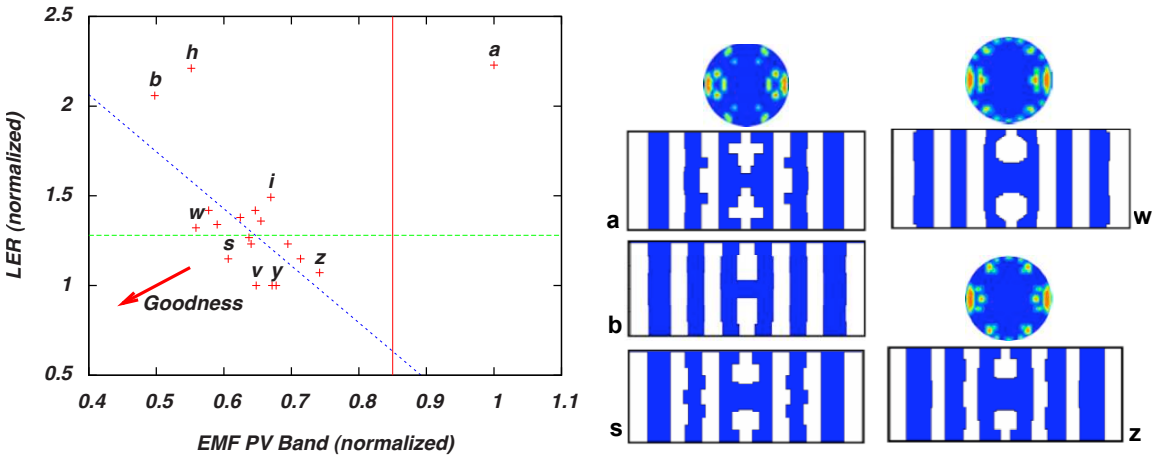
In the following sections, a detailed exploration in process trade-off is described before an investigation into optimization performance given different constraints on the source design.

3.1 Analyzing Process Trade-off

Mask topography effects are well known to introduce unwanted deviations of the amplitude and phase of the electromagnetic fields propagating through the reticle. Full characterization of these effects is often a function of the mask composition, profile and incident illumination. Alternatively, by trading contrast for better mask performance, a reduced impact of EMF effect was achieved through the migration to binary but thinner mask blanks from the previous mask technology of choice, the 6.5% attenuated phase shifting masks.[11] Still, given the sub-wavelength dimensions even at mask scale necessary to enable 32nm technology and beyond, residual mask EMF effects still introduce feature-dependent shifts on the plane of best focus, hence potentially degrading the common process window, and magnifying the impact of mask errors. Moreover, when combining mask errors with focus errors in the +/-3 sigma process variability range, the process variability band (pvband) on wafer tend to increase due to the coupling of mask errors with the drift in best focal plane induced by mask EMF effects.

We studied the case of a 22nm metal layer SRAM as illustrated in Figure 15(b), with deeply sub-wavelength dimensions and a critical line end structure with a tendency for shorting. A range of optimization control parameters were varied during optimization, which revealed a strong correlation between EMF and LER sensitivities. This required a fine tuned optimum solution that balanced the trade-off between both process components. Figure 15(a) is a plot of normalized LER vs pvband width, which has been normalized to initial SMO solution **a**, as illustrated in Figure 15(b), optimized with no LER or EMF constraints. As a result, the sensitivity to both LER and EMF of solution **a** turned relatively higher as shown the plot in Figure 15(b). Case **b**, on the other hand, resulted from applying traditional OPC to the mask after optimizing the source. The vertical solid line and the horizontal dashed line in the plot of figure 15(a) represent approximate cutoffs of solution quality in PV-band and LER metrics, respectively. The SMO solution of case **a** achieved superior Common Process Window (CPW) and smaller edge placement error than the OPC solution of case **b**, however, without proper EMF and LER control, the solution suffers larger degradation when these process components are factored in.

We evaluated the impact of applying four different constraints, one for LER control and three that minimize the sensitivity to EMF effects, in addition to optimization constraints common to all SMO runs. Details regarding these constraints will be discussed in future publications. Three of the solutions obtained with the new constraints are displayed in Figure 15(b), cases **s** which was obtained using mask optimization, and cases **w** and **z**, obtained through source and mask optimization. The resulting LER and EMF PV-Band width for all solutions are plotted in Figure 15(a), while the specific constraints used in each case are collected in Table 1. It is worth noting case **h**, which used one of the constraints to reduce EMF effects, but did not apply LER constraint, resulting in improved EMF pvband width but failed to reduce LER. Case **z**, on the other hand, only attempts to compensate for LER, hence resulting in a clear improvement in LER sensitivity (below the threshold), but at the expense of EMF pvband width when compared to case **h**. This shows the apparent trade-off between the factors contributing to LER and EMF sensitivity.



(a) (b)

Figure 15: (Color) The performance space for line-edge-roughness vs through process variation for metal layer SRAM. (a) Trade off of line-edge-roughness vs through process variation for metal layer SRAM. (b) Associated sources and masks for a number of critical optimizations.

Table 1: Optimization characterization for the solutions in Figure 15

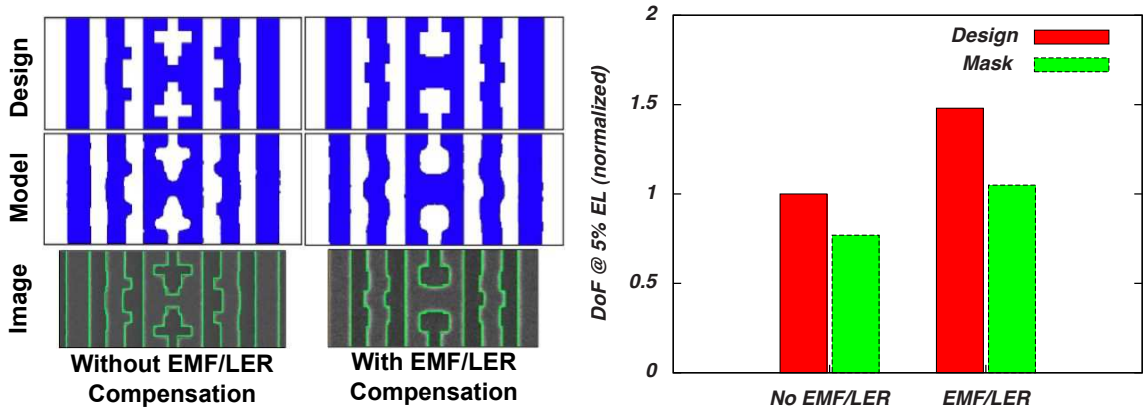
Case	a	b	h	z	w	v	i	y	s
Optimization	CPW	EPE	CPW	CPW	CPW	CPW	CPW	CPW	CPW
LER				x	x	x	x	x	x
EMF-1						x			x
EMF-2			x				x		x
EMF-3								x	

This counter-intuitive trade-off is consequence of a combination of factors which relate to the mechanisms that SMO uses to achieve high exposure latitude, by introducing extreme shapes on the mask that lead to higher EMF effects. An intensive optimization, however, is perfectly suited to address this balance as shown for instance in case **s**, which represents an SMO optimization where LER and two EMF constraints and corrections have been turned on. In this case the optimizer is given control over the needs in the optimization, which the user must provide, and is then able to make the suitable trade-off to deliver well balanced solution. Experimental wafer verification comparing cases **a** and **s** also confirmed improved performance, with nearly 20% improved depth of focus at 5% exposure latitude on the pattern line ends.

Some amount of corner rounding during mask writing is inevitable even when the impact appears small as illustrated in Figure 16(a) where the optimized manhattan mask shapes are shown in the top row. The overlay between manhattan mask shapes and Scanning Electron Microscope (SEM) images of the real mask shapes is displayed in the bottom row for solutions **a** (without EMF/LER compensation) and **s** (with EMF/LER compensation), respectively. The SEM images were used to extract the contours of the actual mask shapes, which are displayed in the middle row of Figure 16(a), and used to simulate the impact of the mask corner rounding on the performance of those two solutions. Simulations using the rounded mask shapes and assuming a thin mask approximation confirmed that mask making effects such as corner rounding can introduce certain amount of degradation in performance. The results from these simulations are displayed in Figure 16(b) and show a reduction of about 20% in the normalized value of depth of focus at 5% exposure latitude at the wafer plane when the effect of mask writing process is included, and this is consistent for both solution without and with EMF/LER optimization (solutions **a** and **s**, respectively).

3.2 Source Solution Robustness

To investigate free form source design robustness, grey level sources were generated via SMO for various staggered and orthogonal arrays of 22nm Contacts, and for SRAM and line/space patterns of Metal level. Primary illumination elements

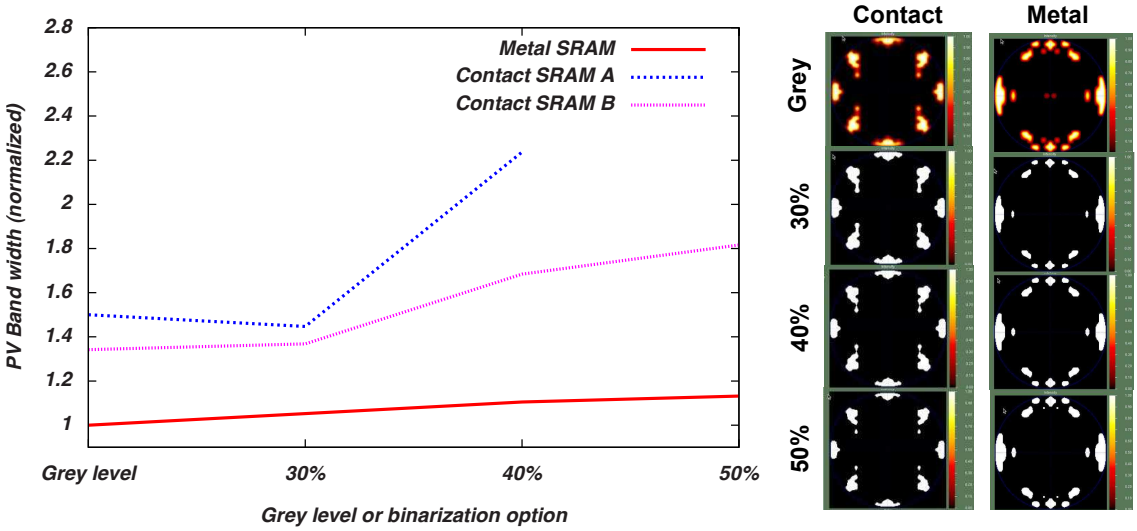


(a) (b)

Figure 16: (a) Top row: optimized manhattan mask shapes. Middle row: rounded mask shapes extracted from SEM images of the reticle. Bottom row: overlay between manhattan mask shapes and SEM images of solutions **a** (without EMF/LER compensation) and **s** (with EMF/LER compensation), respectively. (b) Simulations of the normalized depth of focus at 5% exposure latitude using the optimized manhattan mask shapes and the extracted rounded mask contours.

(pixels) need to attain certain minimum size in order to be manufacturable. We used pixels of 0.06 widths in units of sigma with spacing between pixels centers in sigma-space of 0.06, and minimum pupil fill of 10%. Binary sources were generated out of the grey level sources using binarization filters applied at various thresholds (30-50%). Lithographic performance was evaluated by means of process variation bands measured for SRAMS of Contact and Metal levels.

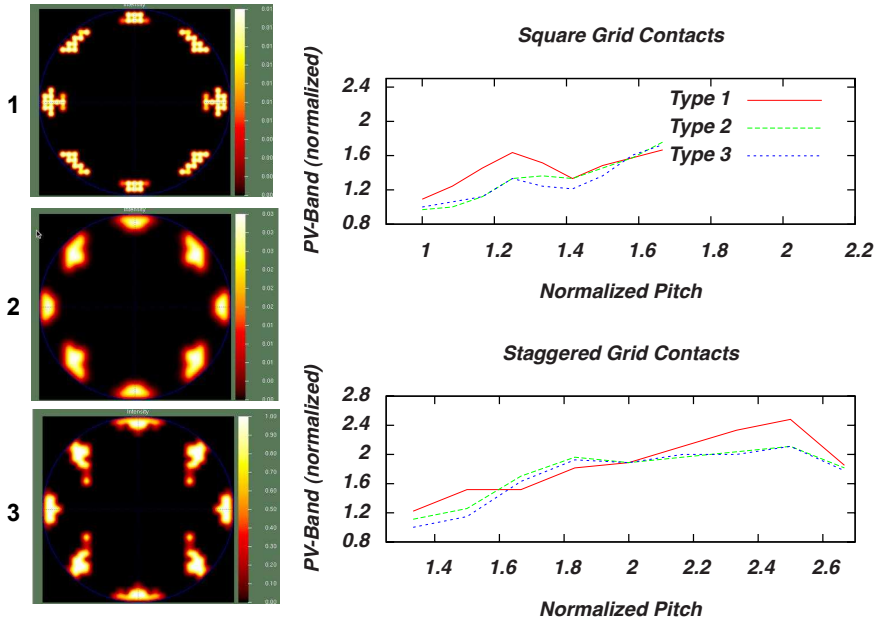
Figure 17(a) depicts the lithographic performance degradation with increased binarization filter threshold for both Contacts and Metal SRAMs. As shown on Figure 17(b) the higher the binarization filter threshold, the more critical for the lithographic performance source components get removed. As a result process variation bands values increase with threshold, which corresponds to degraded weakest process window for both Contacts and Metal level SRAMs.



(a) (b)

Figure 17: Lithographic performance of Grey level sources vs. sources put through binarization filters for Contact and Metal levels. (a) Simulated Process Variation Bands (+/- 57nm defocus, +/- 3.8% exposure variation, +/- 0.5nm mask error) as function of various binarization levels. (b) Output source maps of corresponding grey-level and various binarization level illumination solutions.

We investigated the lithographic performance of free form sources generated using different primary illumination components (pixels) types, which is demonstrated on Figure 18. As shown on Figure 18(a) grey level sources of three different pixel types were generated via SMO for various staggered and orthogonal arrays of 22nm Contacts.



(a) (b)

Figure 18: Lithographic performance through pitch of various types of illumination configuration primary elements for Contact level. (a) Output source maps of the corresponding solutions for 3 different primary illumination elements. (b) Simulated Process Variation Bands (+/- 57nm defocus, +/- 3.8% exposure variation, +/- 0.5nm mask error) as a function of pitch for each of the three 3 types of illumination primary elements.

The source of Type 1 was generated using circular pixels, where some of them are simply repeated duplicates of single circle on a grid, while the sources of type 2 and type 3 depart from the limitations of a simple grid and follow more sophisticated source representation strategy that obtains an extra benefit from the degrees of freedom allowed by manufacturability requirements. As demonstrated on Figure 18(b) the lithographic performance of source type 1 is noticeably degraded in comparison to the lithographic performance of the sources of type 2 and 3, measured through pitch for both orthogonal and staggered arrays of a 22nm contacts level. Process variation bands through pitch of sources of type 2 and 3 are very similar and outperform lithographically the above described source of type 1.

4. SMO WITH LARGE NUMBERS OF CLIPS

A basic requirement in solving inequality-constrained optimization problems is that active constraints be efficiently identified and handled; see e.g. the well-known textbook of Nocedal and Wright,[12] section 15.2. Since SMO can optimize against rigorous process window metrics that are defined in terms of gating critical features,[2] the solution obtained for the source will be determined by the limited set of critical clips which bind the process window, thereby contributing constraints which are active.[1,7] In the present work we are not concerned with the algorithmic identification of binding constraints, but will instead adopt a lithographic point of view in presenting examples of how active clips arise during SMO. We will illustrate this at two points in the SMO process: First, when initiating SMO optimization (section 4.1), where one might use pre-SMO dataprep tools in order to select the initial large population of clips from which the active constraints will be drawn, and second, at the completion of SMO optimization (section 4.2), where one can examine the clips that are binding in the solution.

4.1 Sampling design for Source-Mask Optimization

If the clip set used by SMO is selected from a large layout which has already been designed under a certain set of groundrules (perhaps developed for a conventional RET), it may be of interest to determine the optimum source for the layout on the basis of the selected sample clips. This sampling of the design reduces the size of the SMO problem, where by sampling we mean the combination of a method and a process of selecting non-overlapping layout regions. These clipping regions form a sample. Each clip in the sample is assumed to be of the same rectangular shape and size. The layout features in the clips are passed to SMO engine, then a solution obtained by solving this smaller-than-the-whole-chip SMO problem. The sample can be characterized in terms of its size (number of clips in the sample, or area coverage, which is ratio of total area of sampling clips to the area of chip) and quality (how well SMO solution for this particular sample performs on the whole chip.)

Three major sampling strategies are considered here: Litho-Aware, DRC-Aware, and Blind. The Litho-Aware strategy requires partial or full knowledge of the lithographical process conditions and design database. The hotspots and hard-to-print regions can be identified by running OPC Verification engine, or a combination of RET, OPC, and Verification. This strategy is expensive; it requires a lot of computational resources to be engaged and substantial setup time.

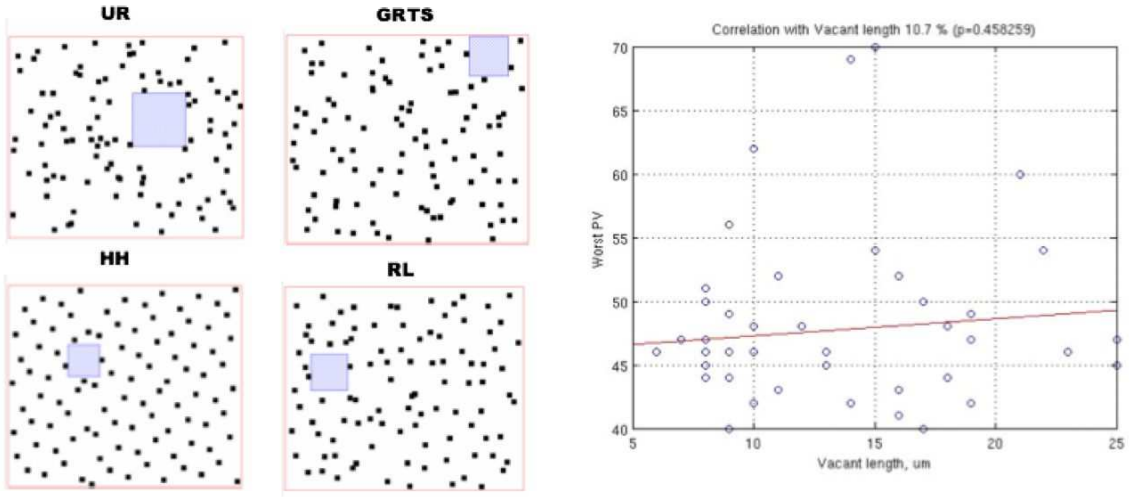
The DRC-Aware strategy can be conducted without knowing details of RET, OPC, and even processing conditions. Here one tries to clip hot spots using geometrical relationships between shapes of the target (design) layer, or conduct a pattern-matching.

The Blind strategy is the simplest, with the only input being the design extent coordinates. The sampling is conducted without paying attention to what is being clipped, only where it is located matters. Roughly speaking, one tries to position clipping rectangles evenly over the chip to reduce probability of missing content that is important for the SMO geometrical configuration. This is the simplest and the cheapest way to construct a sampling. In the practice of SMO application this strategy might naturally be combined with more sophisticated Litho- and DRC-Aware samplings. In this section we discuss types of Blind strategies and assess their quality.

The uniform random (UR) sampling is a straightforward type of Blind strategy. The coordinates of the clipping rectangles are randomly selected from a uniform distribution to cover design extent. The drawback of the random sampling is that it leaves large unattained regions. The size of these regions can be quantified as the largest square that is not covered by the sampling, or the largest void square. The size of a void square reduces very slowly with the sampling size. Another drawback is that the overlaps of clips are frequent, and have to be removed in a post-processing step. The next Blind strategy we consider is Generalized Random Tessellation Stratified (GRTS) sampling as described in.[13] it brings a dose of regularity into the sampling, as compared to the pure random selection, and tries to avoid large voids. even more regular sampling is generated using hammersley-halton (HH) algorithm. We also consider random lattice (RL) algorithm. Initially the clips are positioned at the vertices of an equally-spaced lattice, then random shifts are introduced to move points away from the original positions, but not too far, so they do not leave the vicinity of the lattice vertices they belong to. The differences between these sampling strategies can be seen in Figure 19(a).

In order to assess quality of these sampling strategies, we conducted the following numerical experiment. SMO is executed on N clip selected from a 100 by 100 um chunk of 22 nm a metal layer. The process conditions are: λ=193 nm, NA = 1.35, source is x/y polarized, OMOG mask, and single exposure. The resulting source is passed into MB SRAF module for RET, and then the mask is corrected in nmOPC. Finally we compare maximum process variability band within sampling regions to the maximum process variability band of the whole chunk using OPC Verify module. The process variability bands are obtained running full factorial combination of three defocus conditions, three dose conditions, and three mask biases (for MEF sensitivity), which adds up to 27 process conditions. This procedure of running SMO/MB SRAF/OPC/OPC Verify flow is repeated for different N from 10 to 100, and for UR, GRTS, HH, and RL sampling strategies.

As the result, we observed 5.4 % correlation between area coverage and maximum pv-band width, i.e. the maximum pv-band width for the whole chunk reduces with the increase in the sampling size. This is true for all sampling strategies. We also observed that the pv-band is getting better with the decrease in the size of the largest void square: this is shown in Figure 19(b). The best pv-bands were delivered by HH and RL methods, which is consistent with sizes of the largest void squares. These methods spread sampling clips more evenly than UR and GTRS methods. GRTS and UR methods trail HH by 3.4% and 8% correspondingly in pv-band metric. There was no discernable difference in the quality of pv-band between HH and RL strategies.



(a) (b)

Figure 19: Design sampling. (a) Examples of Blind samples with 100 clips for UR, GTRS, HH, and RL algorithms. The largest void square is calculated and is shown in gray. (b) We observed 10.7% correlation between size of the largest void square and worst pv-band, measured in nm/10.

The presented numerical experiment suggests that HH, RK, and GRTS sampling strategies are valid contenders for Blind sampling algorithms. Our limited experience does not show substantial benefits of using one over another, while the uniform random strategy performed the worst and is not recommended.

4.2 The Interaction of SMO with Design Rules

The experiment summarized in Figure 20 studies the critical clips in a large population which is designed to exhaustively enumerate the patterns allowed by a trial set of design rules, rather than being selected from a specific layout. The experiment tracks the change in the solution and constraining features as increasing numbers of strongly active clips are successively removed from the population.

The enumeration method used to generate the data set re-expresses conventional ground-rules as forbidden patterns in the context of regular griding. These pattern-based rules and coarse grids can be used to quickly generate legal layouts within a specified region, involving limited layers. The purpose of enumeration is to validate a given Restricted-Design-Rules (RDR) set with simulations of current technology assumptions to identify problems, or to confirm continued acceptability in the presence of changing assumptions.

In this case Joint Optimization (JO) is repeatedly carried out on the clip set using a relatively non-optimal set of fixed starting designs, with the optimization being re-run multiple times from this starting point after the most strongly active clips from the previous iteration have been removed. Strongly active constraints were identified in the standard way on the basis of multiplier values, with the cutoff being chosen somewhat arbitrarily.

Figure 20 outlines the evolution of the solution and clip set through the first 15 iterations, showing the source solution, objective value, and most strongly active clips. (The binding clips shown at each iteration are those which were removed before commencing the optimization, i.e. they were most strongly binding in the previous iteration.) The achieved process window shows a clear general trend in which greater improvement is obtained when the removed clips are those of the full population (initial iterations), with the improvement becoming smaller when the removed clips have only become binding after more difficult patterns were already absent. However, the improvement is not reliably monotonic due to the arbitrary starting solution used, and the non-convexity of the problem.

5. CONCLUSION

A number of the benefits of SMO have been reviewed. Evidence of performance advantaged through both simulation and experiment were presented. SMO as a RET technology is maturing and providing distinct benefit to critical layers. SMO

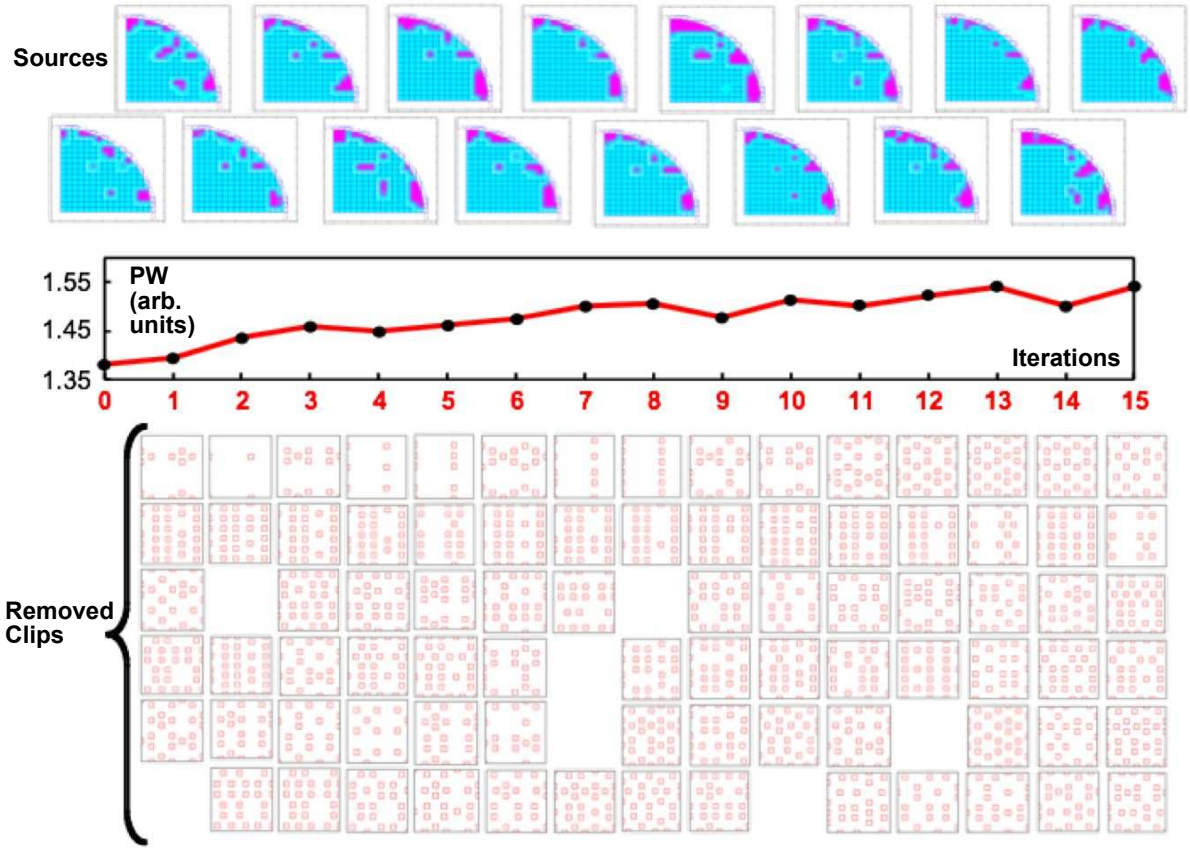


Figure 20: (Color) Evolution of common-window objective as lithographically binding clips are removed from optimization. (See text for details.) Vertical axis is Common Process Window. Horizontal is iteration number. Data on slide pertains to first 15 iterations (red curve). Source solution is shown for each iteration. (Iteration 0 includes all clips.)

optimization provides this benefit in a number of ways; through source design, mask design, RET cycle of learning, and through full source and mask optimization. SMO also enables designs to study and control important trade-offs during development.

ACKNOWLEDGMENTS

This work was performed by various International Business Machines (IBM) Research and Development Facilities, IBM's Semiconductor Research and Development Center, and at Mentor Graphics' Design to Silicon Center. Some layout content and wafer verification was supported by the independent alliance programs for SOI technology development and Bulk CMOS technology development.

REFERENCES

[1] Rosenbluth, A. E., Melville, D. O., Tian, K., Bagheri, S., Azpiroz, J. T., Lai, K., Waechter, A., Inoue, T., Ladanyi, L., Barahona, F., Scheinberg, K., Sakamoto, M., Muta, H., Gallagher, E., Faure, T., Hibbs, M., Tritchkov, A., and Granik, Y., "Intensive optimization of masks and sources for 22nm lithography," Optical Microlithography XXII **7274**(1), 727409+, SPIE (2009).

[2] Rosenbluth, A. E., Bukofsky, S., Fonseca, C., Hibbs, M., Lai, K., Molless, A. F., Singh, R. N., and Wong, A. K. K., "Optimum mask and source patterns to print a given shape," Journal of Microlithography, Microfabrication, and Microsystems **1**(1), 13–30 (2002).

[3] Socha, R., Shi, X., and Lehoty, D., "Simultaneous source mask optimization (smo)," in [Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series], Komuro, M., ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series **5853**, 180–193 (June 2005).

[4] Hsu, S., Hsu, H., Chen, L., Li, Z., Park, S., Gronlund, K., Liu, H.-Y., Callan, N., Socha, R., and Hansen, S., "An innovative source-mask co-optimization (smo) method for extending low k1 imaging," in [Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series], Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **7140** (November 2008).

[5] Rosenbluth, A. E. and Seong, N., "Global optimization of the illumination distribution to maximize integrated process window," Optical Microlithography XIX **6154**(1), SPIE (2006).

[6] Lai, K., Rosenbluth, A. E., Bagheri, S., Hoffnagle, J., Tian, K., Melville, D., Azpiroz, J. T., Fakhry, M., Kim, Y., Halle, S., McIntyre, G., Wagner, A., Burr, G., Burkhardt, M., Corliss, D., Gallagher, E., Faure, T., Hibbs, M., Flagello, D., Zimmermann, J., Kneer, B., Rohmund, F., Hartung, F., Hennerkes, C., Maul, M., Kazinczi, R., Engelen, A., Carpaij, R., Groenendijk, R., Hageman, J., and Russ, C., "Experimental result and simulation analysis for the use of pixelated illumination from source mask optimization for 22nm logic lithography process," Optical Microlithography XXII **7274**(1), 72740A+, SPIE (2009).

[7] Tian, K., Krasnoperova, A., Melville, D., Rosenbluth, A. E., Gil, D., Azpiroz, J. T., Lai, K., Bagheri, S., Chen, C. C., and Morgenfeld, B., "Benefits and trade-offs of global source optimization in optical lithography," Optical Microlithography XXII **7274**(1), 72740C+, SPIE (2009).

[8] Rosenbluth, A. E., Melville, D., Tian, K., Lai, K., Seong, N., Pfeiffer, D., and Colburn, M., "Global optimization of masks, including film stack design to restore tm contrast in high na tcc's," Optical Microlithography XX **6520**(1), SPIE (2007).

[9] Granik, Y., "Source optimization for image fidelity and throughput," Journal of Microlithography, Microfabrication, and Microsystems **3**(4), 509–522 (2004).

[10] Burkhardt, M., Yen, A., Progler, C., and Wells, G., "Illuminator design for the printing of regular contact patterns," Microelectron. Eng. **41-42**, 91–95 (1998).

[11] McIntyre, G., Hibbs, M., Tirapu-Azpiroz, J., Han, G., Halle, S., Faure, T., Deschner, R., Morgenfeld, B., Ramaswamy, S., Wagner, A., Brunner, T., and Kikuchi, Y., "Fast pixel-based mask optimization for inverse lithography," Journal of Microlithography, Microfabrication, and Microsystems **9**(1), 013010–1 (2010).

[12] Nocedal, J. and Wright, S. J., [Numerical Optimization], Springer (August 2000).

[13] Stevens, D. L. and Olsen, A. R., "Spatially restricted surveys over time for aquatic resources," Journal of Agricultural, Biological, and Environmental Statistics **4**(4), 415–428 (1999).