Statistical Timing Analysis based on Simulation of Lithographic Process

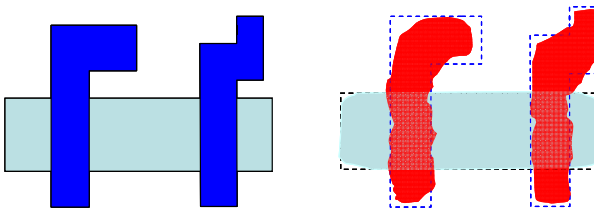
Aswin Sreedhar and Sandip Kundu
University of Massachusetts at Amherst
{*asreedha, kundu*}*@ecs.umass.edu*

***Abstract*—The length of poly-gate printed on silicon depends on exposure dose, depth of focus, photo-resist thickness and planarity of the surface. In sub-wavelength lithography, poly-gate length also varies with layout topology. Poly-gate length determines the effective channel length of a transistor, which determines its performance. Since the sources of error are hard to control, statistical analysis can be used to measure the impact on circuit timing characteristics. Typical lithography-aware methodologies consider only systematic variation such as across chip linewidth variation (ACLV). In this paper we propose a statistical technique for timing yield prediction, based on variational lithography modeling of physical circuit layout. By statistically varying lithographic process parameters we estimate the difference in timing yield estimation of a design. Our simulation results show that if manufacturing process parameters follow a Gaussian distribution, resulting transistors follow a skewed normal distribution, where a greater number of them will have shorter channel length. This led us to investigate whether Statistical Static Timing Analysis (SSTA) is overly pessimistic. The baseline delay model assumed for SSTA in out approach is a Gaussian delay model fitted to skew normal distribution data obtained from statistical litho simulation. Our experiments showed that even after re-centering Gaussian delay model to fit the channel length data with minimum error, it is still overly pessimistic and significantly underestimates circuit performance.**

I. Introduction

As the VLSI technology aggressively scales towards 45nm nodes and beyond, lithographic process variations has emerged as dominant concern. A basic problem in current manufacturing is that industry continues to use 193nm wavelength light source for photolithography, even though the transistors being printed are significantly smaller in dimension. This situation will still prevail in the coming technology generations until an alternative light source and focusing system is found. Lithographic process is sensitive to depth-of-focus, exposure dose; lens aberration and resist thickness variations. Exotic tricks such optical proximity correction (OPC) and other resolution enhancement techniques (RET) have been used to mitigate the effects of topography based printability issues, but gate shapes are still away from perfect rectangles on the resist [1][7].

Gate length variation is a significant contributor towards variability in chip leakage and timing estimation. This results in variation in timing characteristics of gates and wires, thus impacting timing yield. Sources of gate linewidth variation can be classified based on layout topology and process conditions. Layout topology induced proximity effects lead to gate linewidth variation [5][6]. Figure 1 shows the difference between design intent and wafer imprint of gate shapes.



(a) Drawn Layout (b) Printed Contour

Fig. 1. Difference between intent and imprint of gate shapes

Recent studies analyzing effects of lithography induced gate length variation have concentrated on only the systematic variation effects [11][12][14] and none on the random lithographic process effects. A recently proposed work addressed the post lithography analysis and optimization by proposing a timing flow based on residual OPC errors [15]. They consider only timing critical cells and do not take into account that the proximity effects are seen throughout the die. Only static timing analysis of post-OPC layouts are performed and statistical variation is not considered.

Traditional static timing analysis models consider process corners with gate and wire characteristics at nominal, best case and worst case delays. Standard cells are characterized at these process corners and the design is optimized to work for any combination of such delays. With increasing number of lithographic and electrical parameter variations, the number of process corners to be analyzed and characterized becomes computationally expensive. This leads to overly constraint and suboptimal designs. Since many process corners have very low probability of occurrence in actual conditions, this approach is pessimistic in predicting timing yield [1][3][4][7]. Statistical Static Timing Analysis (SSTA) is an alternative to Static Timing Analysis. By default, the underpinning delay model for SSTA is a Gaussian distribution. Gaussian distribution has several nice mathematical properties that have been used to speed-up SSTA [1][4].

The fundamental question raised in this paper is whether a Gaussian delay model for gates is appropriate. To answer that question, we start with the manufacturing process to study the impact of manufacturing parameters such as optical dosage, focus and resist thickness variation against

978-1-4244-5028-2/09/$25.00 ©2009 IEEE

various layout topographies. Since these are independent parameters, in absence of any further information, we assume that these parameters follow normal distribution. Then we study the impact of these parameters on transistor gate length. It was found that the transistor gate length does not follow normal distribution. We then fitted a normal distribution to delays that result from these gate length distributions obtained from statistical litho simulations and performed SSTA. The result from such SSTA was compared against Monte-Carlo simulation based on raw weighted gate length data. The results show that conventional SSTA is pessimistic. This is the central result of this paper. With this enhanced understanding, SSTA can be improved further. Rest of the paper is structured to explain the basis of our study and experimental methodology.

In this paper, we attempt to model the effect of random lithographic process parameter variations and estimate the timing yield of a design. Contributions of our work are:

- Comprehensive gate length variation analysis for both systematic (topology based) and random (process parameters) variations by performing lithography simulations
- Statistical standard cell characterization, circuit delay distribution using Stratified sampling based approach to model lithographic process parameter variation.

The rest of the paper is organized as follows. In section II we briefly describe some common systematic and random lithographic variations. We also describe our stratified sampling-based process parameter characterization method. In section III we present our novel layout-specific comprehensive lithographic variation aware timing analysis and standard cell characterization flow. Experimental results on benchmark circuits using our methodology have been shown in section III to quantify the impact of variational lithography on timing. And finally in section IV we conclude with inferences from our work and some pointers on ongoing research.

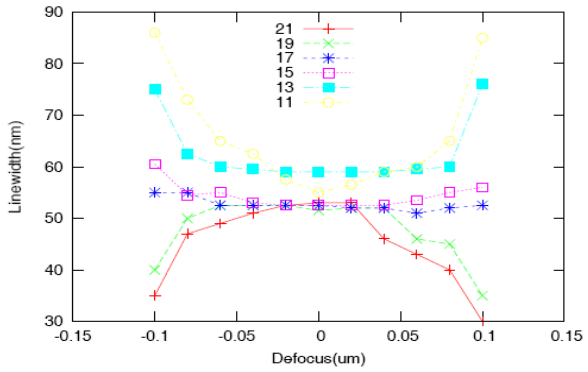


Fig. 2. Plot showing Linewidth vs Defocus for dense Metal-1 lines in 45nm technology under different exposure conditions calculated using PROLITH®.

II. MODELING PROCESS PARAMETER VARIATION

A. *Linewidth Estimation*

Across Chip Linewidth Variation (ACLV) is the variation of gate or interconnect linewidth with pitch between polygons across the chip. It is the most significant contributor towards chip leakage and timing variability. Although Optical Proximity Correction (OPC) techniques have been implemented in today's designs, the effects of these variations have not been completely eliminated [9]. ACLV due to pitch variation in 45nm process using λ=193nm illumination system for any layout is shown using Bossung plot. ACLV is primarily dependent on layout topology i.e. how polygons on the masks are spaced. Process parameter sources of such variation include defocus, exposure dose variation, lens aberration, wafer tilt and resist thickness variation.

Depth of Focus (DOF) can be defined as a point on the resist at which when the focal plane is placed, results in proper printability of the pattern on the mask. When the placement of the focal plane is improper, it leads to variation in linewidth across the wafer. This source of error is termed as defocus. Defocus is caused by several sources, such as lens aberration, resist thickness variation, chemical mechanical polishing and wafer misalignment. A well known illustration for the effect of focus variations on line width is the Bossung plot as shown in Figure 1. Another significant contributor for process variations is the exposure dose variations. Exposure dose variations are truly dependent on the quality of the light source used in the process. Dose variations also lead to linewidth variations. An example of dose variation may be due to flare of the light source. Dose variations can also be modeled as a normal distribution.

Aberrations are imperfections in the lens system leading to imperfect patterns in the printed wafer. Lens aberrations are major source of focus disturbances and effectively lead to linewidth variations. Though there have been recent studies on lens aberration that focus on the lens system and modifications that can compensate its effect, linewidth variation due this source will always be a factor to consider (Figure 2). Zernike's coefficients capture the deviation from ideal characteristics. The CD error due to lens aberration varies along both the horizontal and vertical direction. Zernike's coefficient models different types of lens aberration. Each type of aberration has a different set of coefficients. These different coefficients can be used in lithography simulation to estimate the impact on linewidth. Another interesting input error is the resist thickness variation. Resist thickness variation affects all the metal layers as it is a major source of focus variation. Linewidth changes due to resist thickness variation called swing curves are non-linear. Experiments have also been performed to estimate the effect of anti-reflection coating on these swing curves [16].

Focus and exposure dose variation are the two most important sources which are random in nature and are hard to control. All other source of errors can be lumped into these two sources equivalently [8]. Based on the above discussion, it is observed that CD variation due to

lithographic sensitivities is a significant concern. Such variations affect performance and power of a design and hence have to be estimated.

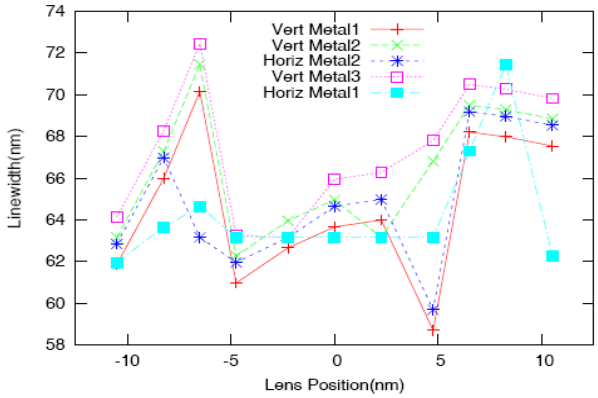


Fig. 3. Linewidth variation calculated due to lens aberration (using Zernike's coefficients) for Metal-1 lines in 45nm using PROLITH®.

B. Stratified/Zone Sampling

In order to better understand the effect of random manufacturing process variations on metal linewidth, a joint analysis is required. However, as line width is not a linear function of control variables, a multivariate analysis of variation of process parameters is required to obtain line width distribution. One way to perform this analysis is Monte Carlo simulation.

Full-chip lithography simulation or Monte Carlo based analysis can be used to deduce the effect of such variations in the lithographic process. However, it is well known that aerial imaging simulations are computationally expensive and the number of such simulations must be limited [8]. We have used a stratified sampling technique to simulate the range of the control parameters. Stratified sampling has been used in varied number of applications in statistical analysis domain. It has been proven that this technique does a simple, reasonably accurate and efficient sampling of the data such that entire distribution is represented by the samples from different regions. This method is used to create a pre-characterized look-up table for resist CD based on lithography simulation at specified process parameter sample points within each zone/strata. This technique reduces the number of simulations to obtain gate CD distribution, thus reducing the computational needs.

III. METHODOLOGY & GATE-CD CHARACTERIZATION

In this section we present our lithographic variation-aware gate linewidth-limited timing yield prediction methodology with an overview of the experimental approach & analysis.

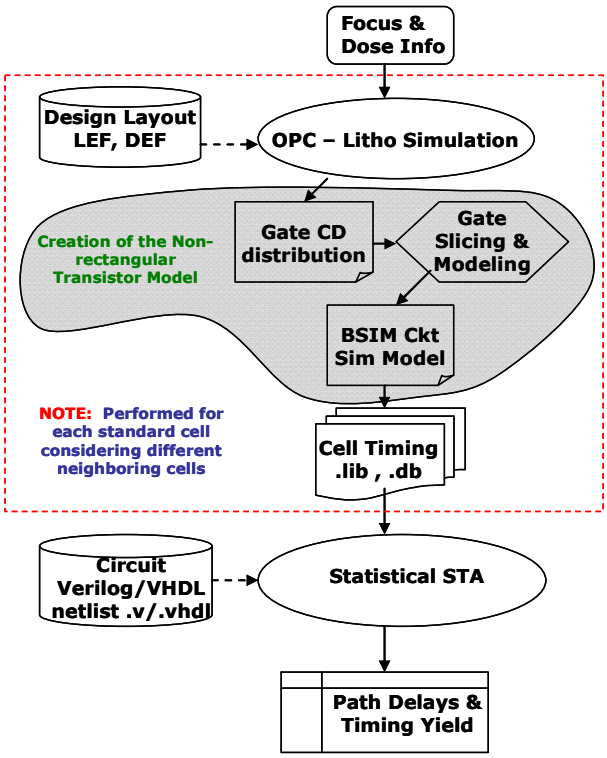


Fig. 4. Lithographic Variation-aware timing methodology

A. Methodology

Figure 4 shows our timing methodology with various sources of variations in today's lithographic manufacturing process. The initial processing steps are done using test layout structures to obtain the ideal printability conditions for each metal layer for different technology nodes. In this case we consider only 65nm and 45nm technology nodes. The nominal values for a process are based on data obtained from the focus exposure matrix (FEM) by running lithography simulation. The pitches in each case are kept at the DRC rule specified minimum.

Subsequently, optimal focus and exposure dose obtained in the pre-processing step is perturbed following a Gaussian distribution to model manufacturing process variations in lithography. Lithographic simulation of the layout with these perturbed parameters produces a CD distribution of the poly-gate. The gate length distribution obtained above is used to obtain a delay distribution of the cells as well as timing paths.

B. Gate CD Estimation & Cell Level Analysis

Masks for each cell of different drive strengths are drawn. The library contains cells that are commonly used in today's design and have typical drive strengths in ratios of 2 to 16. Library based OPC is run and CD resist profile is obtained. The gate CD varies for NMOS and PMOS gate masks as the spacing between them is different for different cells in the library.

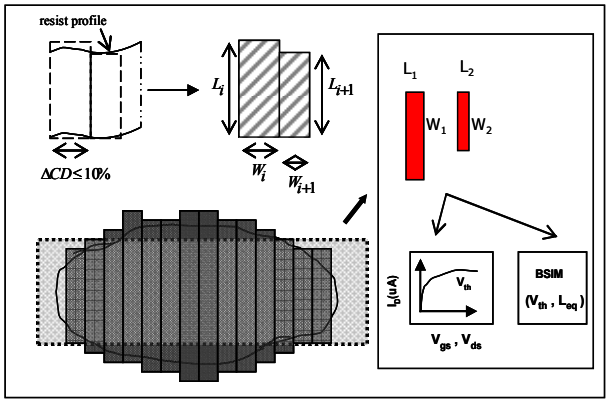


Fig. 5. Non-rectangular transistor modeling based Gate CD estimation

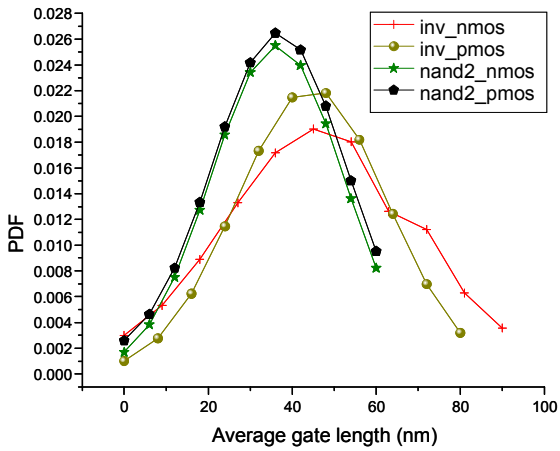


Fig. 6. Gate CD distribution for library cells

It is well known that rectangular structures such as transistors appear more elliptical in the silicon. Since printed transistors on silicon are not rectangular, an alternative approach is needed. In our methodology, poly-gate profile obtained from lithography simulation is used to derive a final gate CD distribution using a non-rectangular transistor model. We have used a model [13] based on gate slicing. In this model, the poly-gate is sliced over the width of the device. Device simulation is done on each slice and the currents at different regions of operation and the parasitic capacitances are estimated. Using these values of each slice, a BSIM model with set of transistors of different length and width is provided that models both static and dynamic behavior. It is also important to note that the final model for the gate is the same irrespective of the type of standard cell.

The placement of the cell layout defines the gate CD as it is dependent on the spacing between the mask patterns. Dimensions of the metrology window used in statistical litho simulation are equal to the optical range of influence ~ 1μ. Since this optical range of influence or optical diameter is large enough to contain more than one standard cell, the proximity influence of other neighboring cells are considered. Within the optical diameter region, the process parameters like exposure dose, depth of focus and resist thickness will tend to remain constant. Thus, there is correlation of process parameters between neighbors on boundary portions of most cells within this region. To reflect this in Monte Carlo Simulation, the process parameters such as focus, dose; resist thickness are selected first, then from lithography simulation, the transistor gate lengths follow. Subsequently, the gate length information is used for delay characterization. This is how the delay distributions are obtained. Hence the simulation estimates the CD variation close to the full-chip simulation. Based on the results obtained from statistical litho simulation, we observe that a great number of these poly-gate masks tend to have a shorter gate length. This behavior is captured more accurately by a skew-normal distribution than a Gaussian distribution used in traditional SSTA.

As shown in Figure 5, using the non-rectangular model, the transistor gate length distribution obtained from statistical litho simulation is used to calculate statistical delay distribution for library cells. A sample delay arc distribution of an INV and NAND2 cell obtained using spice non-rectangular transistor models is shown in Figure 6. More details on the actual model characterization can be seen in [13].

IV. EXPERIMENTATION AND RESULTS

In this section, we test our approach for various designs using leading-edge industry tools, and estimate the CD-limited timing yield.

A. Monte Carlo Simulations

PROLITH®, a commercial imaging simulator from KLA-Tencor was used for lithography simulation. The metal layer masks used for the simulation were binary masks (BIM) with alternate phase shift masking (Alt- PSM). The test cases used were poly-gate masks of widely available academic ISCAS85 benchmark circuits. The designs were synthesized using Synopsys Design Compiler vW-2004.12-SP3. HSPICE™ was used for Standard cell characterization and Synopsys PrimeTime® Suite was used in conjunction with Monte-Carlo simulation for statistical static timing analysis on designs synthesized with lithographic variation-aware standard cells to obtain circuit path delays. To compute mean path delays as shown in Table 1 , gate delay was based on the obtained skew-normal distribution and wire delays used were the nominal Elmore delay models.

The assumptions for input error distributions were based on ITRS specification that the acceptable variation of both dose and focus is ±10%. To include worst case variations, in our experimental approach we assume that the focus and exposure dose to be normally distributed and vary between ±25% of the nominal value. Any other type of error distribution can also be used for the input parameters.

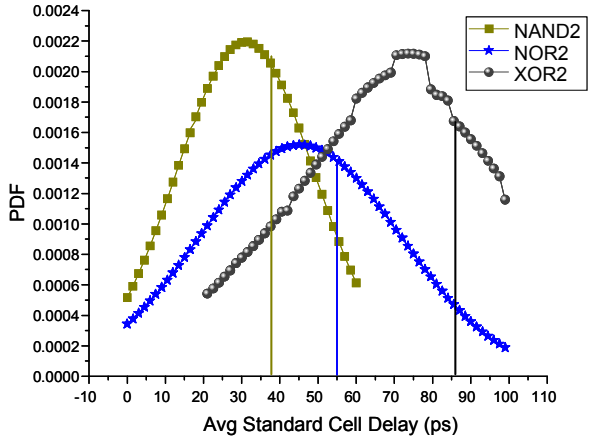


Fig. 7. Average Standard cell arrival time distribution

B. *Traditional SSTA*

In order to compare with traditional SSTA approach, we fitted gate delay data to a normal distribution and then used this distribution of delays to perform Monte-Carlo simulation using Primetime™. Table 1 quantifies the pessimism in traditional Gaussian distribution fitting-based SSTA for circuit path delays. It can be seen that there is inherent pessimism in the standard SSTA approach. Lithographic process variation increases the probability of having a shorter gate and consequently higher performance. It must also be noted that the increased performance is accompanied by drastically higher leakage.

Table 1. Pessimism comparison and timing trend

Design	65nm Mean delay			45nm Mean delay		
	Traditional	LV-A	Diff %	Traditional	LV-A	Diff%
C1355	464.0	432.6	7.2	459.4	405.5	12.3
C6288	1755.7	1473.6	16.1	1705.5	1394.9	19

In order to analyze the effect of lithographic sensitivity on the design at different technology nodes, a comparison was made. Table 1 shows the trend in design timing as we move from 65nm to 45nm technology. The deviation of mean delay from nominal timing has increased two fold at a smaller technology node. Due to skewed normal distribution of the channel length, the peak probability moves towards shorter channel. Consequently, the performance improves (shown in LV-A or Lithographic Variation-Aware column). This shows that the Gaussian delay assumption increases error in SSTA as we move towards smaller geometries. Figure 7 shows the lithographic variation aware setup/latest arrival time information for three standard cell designs that follow a skew-normal distribution. The lines that accompany each distribution represent mean arrival times of traditional SSTA approach. The same can be seen in the case of hold/earliest arrival time.

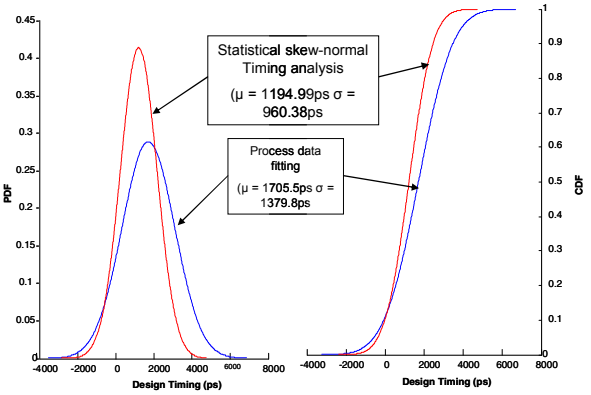


Fig. 8. Arrival time CDFs and PDFs for different approaches

The effect of lithographic process parameter variation on a large design was conducted and the delay distribution was measured. The input parametric error distribution was set to an industry defined FEM (Focus-Exposure Matrix) window range. The delay values were fitted to a Gaussian distribution and skew-normal distribution. Figure 8 shows the PDF and CDF of the distribution using two different approaches. The 99% confidence points or the peaks of the distribution occur at 1194ps for the skew-normal distribution based approach and 1705ps for the Gaussian fitting approach. For proper comparison, the skew-normal direct parameters such as location, scale and shape have been converted to centered parameters (mean μ and s.d σ).

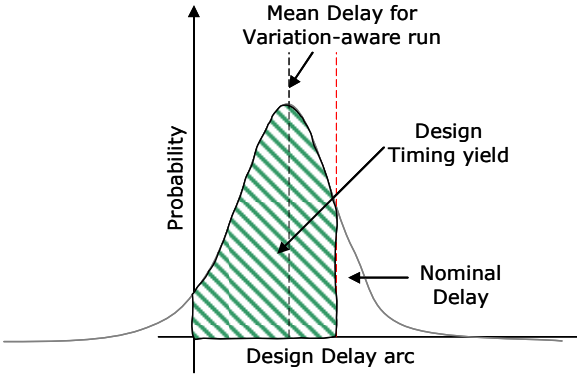


Fig. 9. Design Timing yield estimation

C. *Yield Estimation*

Timing yield is an important specification to be met in circuit design. The presence of lithographic process variation, timing yield will get affected. In our approach, we define timing yield as the probability of the design having an input-output delay less than the nominal delay when no variation is considered. This is shown pictorially in Figure 9. Table 2 shows the timing yield for different designs at 45nm technology. The timing yield using our stratified sampling based approach and the traditional SSTA approach is shown. It can be seen that the timing yield estimation is better since the statistical performance the design is better even in the presence of lithographic variation.

Table 2. Timing yield for ISCAS85 designs at 45nm

Design	Traditional SSTA	Stratified SN-based
c432	0.9941	0.9963
C2670	0.9836	0.9841
C3540	0.9715	0.9823
C5315	0.9445	0.9732
C6288	0.8568	0.9379
C7552	0.9214	0.9711

V. CONCLUSION

In this paper, we presented a technique to predict the probability distribution of transistor channel length under lithographic process variation. The process variation model incorporates the effects of exposure dose, depth of focus, photo-resist thickness variation on various layout topologies. The lithographic simulation results show that (i) process variation increases as we move towards smaller geometries and (ii) as a result of process variation, shorter gate lengths are more likely than longer gate lengths and also leads to the formation of non-rectangular gate structures. Due to this, (iii) traditional Gaussian delay distribution that is modeled based on rectangular gate structures introduces significant pessimism in SSTA, which (iv) increases with scaling. Thus, we predict, the timing yield to be better than one predicted by a simple Gaussian distribution of delays. On an average, we expect 14% better timing yield than what is predicted by a Gaussian delay model.

REFERENCES

[1] C Visweswariah, "Death, Taxes and Falling Chips", Proc. Design Automation Conference, 2003, pp343-347.

[2] 2006 International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net.

[3] M. Orshansky and K. Kuetzer, "A General Probabilistic Framework for Worst Case Timing Analysis," Proc. DAC, 2002, pp. 556-561.

[4] H. Chang, V. Zolotov, S. Narayan and C. Visweswariah, "Parameterized Block-based Statistical Timing Analysis with Non-Gaussian Parameters, Nonlinear Delay Functions," Proc. DAC, 2005, pp. 71-76.

[5] M. Orshansky, L. Milor, P. Chen, K. Keutzer, and C.Hu, "Impact of systematic spatial intra-chip gate length variability on performance of high-speed digital circuits," Proc. ICCAD, 2000,pp. 62-67

[6] Sandip Kundu, Aswin Sreedhar, Alodeep Sanyal, "Forbidden pitches in sub-wavelength lithography and their implications on design," J Computer-Aided Mater Des (2007)

[7] D. Blaauw, S. Nassif, L. Scheffer and A. Strojwas, "Design for Manufacturing in the Sub-100nm Era", Design Automation Conference, Tutorial.

[8] C. A. Mack, "Fundamental Principles of Optical Lithography," Wiley, 2008.

[9] F. M. Schellenberg. "Resolution enhancement technology: The past, the present and extension for the future," in SPIE Microlithography Symposium, 2004.

[10] A. B. Kahng and Y. C. Pati, "Subwavelength Lithography and its Potential Impact on Design and EDA", Proc. ACM/IEEE DAC, 1999, pp. 799-804.

[11] A. Balasinski, L. Karklin, V. Axelrad, "Impact of Subwavelength CD Tolerance on Device Performance," Proc. SPIE Vol. 4692, 2002.

[12] Puneet Gupta and Fook-Luen Heng, "Towards a Systematic-Variation Aware Timing Methodology", In Proc. Design Automation Conference, pp.321-326, June 2004.

[13] Aswin Sreedhar and Sandip Kundu, "On Modeling Impact of sub-wavelength Lithography on transistors," In Proc. International Conference on Computer Design, 2007.

[14] Ke. Cao, Sorin Dobre, Jiang Hu, "Standard Cell Characterization Considering Lithography Induced Variations," Proc. Design Automation Conference, 2006.

[15] Jie Yang, Luigi Capodieci, and D. M. Sylvester, "Advanced Timing Analysis Based on Post-OPC Extraction of Critical Dimensions," presented at DAC, 2005.

[16] E. Charrier, C. Mack, "Yield modeling and enhancement for optical lithography," Proc. SPIE, 2440,pp.435-447, 1995.