University of Massachusetts Amherst
ScholarWorks@UMass Amherst

Masters Theses 1911 - February 2014

2008

Automatic Techniques for Modeling Impact of Sub-wavelength Lithography on Transistors and Interconnects and Strategies for Testing Lithography Induced Defects

Aswin Sreedhar
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/theses

Sreedhar, Aswin, "Automatic Techniques for Modeling Impact of Sub-wavelength Lithography on Transistors and Interconnects and Strategies for Testing Lithography Induced Defects" (2008). *Masters Theses 1911 - February 2014*. 80.
Retrieved from https://scholarworks.umass.edu/theses/80

This thesis is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses 1911 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

AUTOMATIC TECHNIQUES FOR MODELING IMPACT OF SUB-WAVELENGTH LITHOGRAPHY ON TRANSISTORS AND INTERCONNECTS AND STRATEGIES FOR TESTING LITHOGRAPHY INDUCED DEFECTS

A Thesis Presented

by

ASWIN SREEDHAR

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

February 2008

Electrical and Computer Engineering

© Copyright by ASWIN SREEDHAR 2008

All Rights Reserved

AUTOMATIC TECHNIQUES FOR MODELING IMPACT OF SUB-WAVELENGTH LITHOGRAPHY ON TRANSISTORS AND INTERCONNECTS AND STRATEGIES FOR TESTING LITHOGRAPHY INDUCED DEFECTS

A Thesis Presented

by

ASWIN SREEDHAR

Approved as to style and content by:

Sandip Kundu, Chair

Israel Koren, Member

Massimo V. Fischetti, Member

C. V. Hollot, Department Head
Electrical and Computer Engineering

ACKNOWLEDGMENTS

I am deeply indebted to my advisor, Professor Sandip Kundu for his constant encouragement and support throughout my graduate studies. His vast technical expertise and insight has given me an excellent background in the field of vlsi design and test. I would like to thank Professor Israel Koren and Professor Max Fischetti for their guidance and valuable feedback thoughout my work. I would also like to thank KLA-Tencor for providing lithography tools.

I have been fortunate to have worked with knowledgeable and friendly people at the VLSI CAD laboratory at UMass. I am thankful to my parents and my brother for their constant support, encouragement and love in all aspects of my life.

ABSTRACT

AUTOMATIC TECHNIQUES FOR MODELING IMPACT OF SUB-WAVELENGTH LITHOGRAPHY ON TRANSISTORS AND INTERCONNECTS AND STRATEGIES FOR TESTING LITHOGRAPHY INDUCED DEFECTS

FEBRUARY 2008

ASWIN SREEDHAR

B.E, ANNA UNIVERSITY

M.S.E.C.E, UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Sandip Kundu

For the past four decades, Moore's law has been the most important benchmark in microelectronic circuits. Continuous improvement in lithographic technology has key enabler for growth in transistor density. In recent times, the wavelength of the light source has not kept its pace in scaling. Consequently, modern devices have feature sizes that are smaller than the wavelength of light source used currently in lithography. Printability in sub-wavelength lithography is one of the contemporary research issues. Some of the printability issues arise from optical defocus, lens aberration, wafer tilting, isotropic etching and resist thickness variation. Many of such sources lead to line width variation in today's layouts. In this work we propose to simulate such lithographic variation and estimate their impact on current devices and interconnects. We also propose to model such effects and aim to provide measures

at the design level to mitigate these problems. Variations arising out of lithography process also impact yield and performance. We plan to study the impact of sub-wavelength lithography on yield and provide solutions for its measure, and directed pattern developement and testing.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS . . . . **iv**

ABSTRACT . . . . **v**

LIST OF TABLES . . . . **x**

LIST OF FIGURES . . . . **xi**

CHAPTER

1. INTRODUCTION **1**

1.1 MOSFET Scaling - Benefits & Issues 1
1.2 Photolithography & Printability 2
1.3 Research Objectives & Thesis Outline 3

2. FORBIDDEN PITCHES AND THEIR IMPACT ON DESIGN **6**

2.1 Proximity Effects 6
2.2 Theory on Optical Interactions 7
2.3 Forbidden Pitches 9
2.4 Simulation Metrics 10
2.5 Design Impact 11
2.6 Summary 14

3. MODELING AND TESTING OF LITHOGRAPHY RELATED OPEN FAULTS **15**

3.1 Motivation 15
3.2 Related Work 16
3.3 Lithographic Defect Extraction 17

3.3.1 Methodology 18
3.3.2 Lithography Simulation 18
3.3.3 Defect characterization 19

3.3.4 Layout defect extraction 21
3.3.5 Logical Fault Translation 23
3.3.6 Experimental Setup 23

3.4 Fault Testability 24

3.4.1 Meta-stability of floating nets 24
3.4.2 Testing bi-stable floating nets 25

3.5 Results 26
3.6 Summary 28

4. MODELING IMPACT OF LITHOGRAPHY ON TRANSISTORS 30

4.1 Motivation 30
4.2 Previous Work 31
4.3 Proposed Modeling Methodology 33
4.4 Gate Characterization 33

4.4.1 Gate Slicing Technique 34
4.4.2 Current Modeling 35

4.5 Gate Length Computation 36

4.5.1 Modeling gate length variation 37
4.5.2 Model parameter extraction 39

4.6 Results 39

4.6.1 Transistor Model Validation 40

4.7 Standard Cell characterization & Circuit Simulation 42

4.7.1 Leakage Estimation 43
4.7.2 Timing Delay Estimation 43
4.7.3 Model Validation for Lithographic variation 44

4.8 Design Implementation 45
4.9 Summary 46

5. YIELD MODELING AT LITHOGRAPHIC PROCESS CORNERS 47

5.1 Motivation 47
5.2 Yield Loss Mechanisms 48

5.2.1 Random Yield Loss Mechanisms 49
5.2.2 Systematic Yield Loss Mechanisms 50
5.2.3 Parametric Yield Loss Mechanisms 50

5.3 Variational Lithography 51
5.4 Probabilistic Modeling & Yield estimation methodology 55

5.4.1 Methodology 55
5.4.2 Variation-aware CD-limited yield estimation technique 56

5.4.2.1 Line Fault Modeling 58
5.4.2.2 Half-width based Yield Estimation 59

5.5 Experimentation 62

5.5.1 Experimental setup 62
5.5.2 Probabilistic yield estimation results 63

5.6 Summary 65

6. CONCLUSION 66

6.1 Summary 66
6.2 Suggestions for Future Work 67

BIBLIOGRAPHY 68

LIST OF TABLES

Table **Page**

3.1 Critical lengths for different metal layers . . . 25

3.2 Lithography aware open fault statistics for ISCAS-85 benchmark circuits (METAL I layer) . . . 27

3.3 Lithography aware open fault statistics for ISCAS-85 benchmark circuits (METAL II layer) . . . 27

3.4 Comparison between lithography aware open fault count and stuck-at fault count for ISCAS-85 benchmark circuits . . . 28

4.1 Results from transistor modeling . . . 39

4.2 Leakage Power estimation . . . 43

4.3 Timing Delay Estimation . . . 44

4.4 Timing Delay Estimation with Variations . . . 45

4.5 Leakage Power estimation . . . 45

4.6 Design Leakage Estimation . . . 46

5.1 PROLITH^{TM}simulation parameters . . . 62

5.2 Probability at nominal spacing . . . 64

5.3 CD-limited design yield (Metal 1) . . . 65

LIST OF FIGURES

Figure		Page
1.1	Light source for various technology generations	3
2.1	Optical photolithography system	7
2.2	Binary mask and its transmission funtion	8
2.3	Forbidden pitches found at (a) 65nm and (b) 45nm technologies	11
2.4	Constriction in metal lines	12
2.5	Example metal layer	12
2.6	Case study I for diffraction problem in via placement: (a) the mask with four vias separated by more than the minimum separation; and (b) the photo-resist profile for the mask after development	13
2.7	Case study II for diffraction problem in via placement: (a) the mask with four vias separated by the minimum separation; and (b) the photo-resist profile for the mask after development	14
3.1	Particle driven Vs Systematic Defects [21]	16
3.2	Defect Extraction Methodology	18
3.3	Forbidden Pitch Rules	19
3.4	Processing Place & Route Data	20
3.5	Forbidden Pitch error locations	22
3.6	Open Fault nets	22
3.7	Example net location and its hierarchical name	23
3.8	Critical length and side channel excitation technique	26

4.1 Variation of Ion and Ioff with CD 31

4.2 Proposed post-litho transistor modeling technique 33

4.3 Gate Slicing Technique and Sliced poly-gate 34

4.4 Length and Width variation for Ion/Ioff 38

4.5 Transistor Leakage current and Drive current comparison 40

4.6 Leakage and Drive current analysis for different gate cases 41

4.7 Inverter circuit simulation model 42

4.8 NAND circuit simulation model 42

5.1 Sources of yield loss in today's lithographic manufacturing process 48

5.2 Parametric variation distribution 51

5.3 Bossung plot for dense Metal 1 lines in 45nm technology 52

5.4 Linewidth variation due to lens aberration (using *Zernike's coefficinets* for metal lines in 45nm 53

5.5 ACLV *Forbidden pitches*leading to yield loss 54

5.6 Variation-aware yield prediction methodology 55

5.7 Yield prediction methodology overview 56

5.8 Stratified sampling on bivariate normal input error distribution & output skew-normal distribution 57

5.9 Line Edge Roughness defined 58

5.10 Line Fault Modeling - (a) Metal line OPEN; (b) Metal line SHORT 60

5.11 Half-width based yield estimation 60

5.12 Pseudo-code for half-width yield estimation 61

5.13 CD distribution for input parameter variation at different spacing 63

CHAPTER 1

INTRODUCTION

1.1 MOSFET Scaling - Benefits & Issues

Moore's Law, which predicts that the number of transistors on a chip will double every two years, has been a driving force behind the semiconductor industry for four decades [40]. Continuation of Moore's law has been possible due to improvement in lithographic process, transistor and interconnect manufacturing technologies.

Increase in transistor density as predicted by Moore's law is due to constant reduction of physical MOS device dimensions that has helped improve design performance. Traditional benefits of scaling include (i) Integration : chip area decreases thus enabling higher transistor density and cheaper ICs (ii) Performance : chip frequency doubles with reduction in feature length every technology generation, and (iii) Power : chip supply voltage is reduced and hence power consumption decreases.

Today, the leading edge nano-scale semiconductor devices are manufactured using 90nm lithography process where the minimum size of drawn dimension corresponds to 90nm. The effective channel length of transistors manufactured in this technology is about 45nm that corresponds to a span of 137 or so Si-atoms. This technology allows manufacturing of chip containing 2B transistors in mere 580 mm^2 of silicon [22]. Transistors manufactured in 45nm technology are expected to have a channel length (L_{eff}) of 20nm which corresponds to a span of only 60 or so atoms [5]. As the channel length reduces, it becomes more comparable to the junction depths of the drain and source regions and hence will lead to short channel effects (SCE). This leads to increase in (a) Leakage power, (b) V_T roll off and (c) punch through. In nano-

scale CMOS technology that integrates billions of devices, sub-threshold leakage puts a floor on permissible transistor threshold voltage, which in turn limits scaling of power supply voltage. This has become a bane for power scaling.

While the relentless focus of scaling has been on devices and interconnects, the system level integration issues deserve no less attention. Even the benefit in terms of integration will come under strain as well. This strain comes from sub-wavelength lithography.

1.2 Photolithography & Printability

Integrated circuits are patterned using optical lithography. Photolithography is the art of printing patterns on semiconductor wafers to fabricate the devices and metal layers of an integrated circuit. Historically, continuation of transistor feature size scaling has been possible through concurrent improvement in lithographic resolution. The lithographic resolution was primarily improved by moving deeper into ultraviolet spectrum of light. However, the wavelength of the optical source used for lithography has not improved for nearly a decade Figure 1.1. The wavelength of the optical light source used for photolithography has not scaled at the same rate as that of the minimum feature size of the transistor. In fact, starting with 180nm devices, the wavelength of optical source has remained the same (at 193nm) due to difficulties in finding a flare-free, high energy, coherent light source with compatible improvement in lens material for focusing this light. Consequently the semiconductor industry will be fabricating 45nm CMOS devices using 193nm optical light.

This has lead to development of subwavelength lithography. A major issue in sub-wavelength lithography is printability. Printability issues encompass variations due to pitch, focus and light source exposure. The printability problems are related to optical diffraction and the resulting effects include (i) line pullback (ii) line width variation, (iii) corner rounding and (iii) pattern density issues. Today, the printabil-

ity problems are addressed by resolution enhancement techniques (RET) [50] which aim to reduce the effect of the above mentioned variation on printed wafers. Several

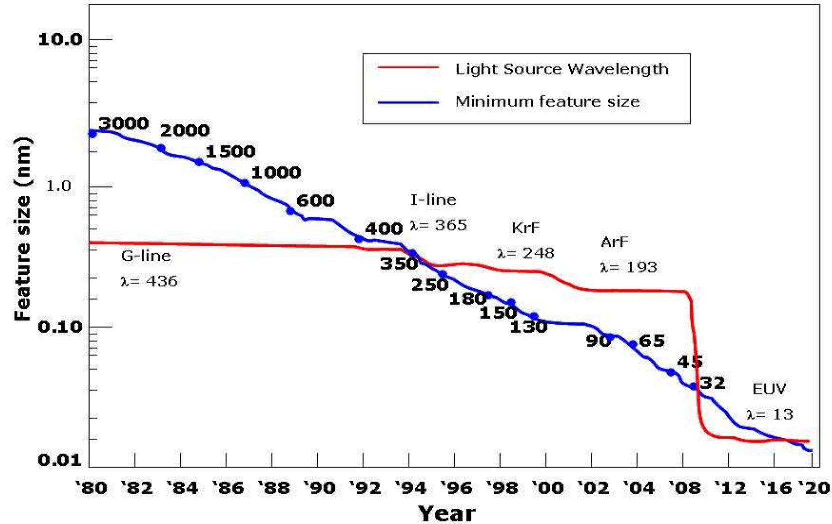

Figure 1.1. Light source for various technology generations

resolution enhancement techniques have made sub-wavelength lithography possible. Important ones include usage of Sub-Resolution Assist Features (SRAF) in the mask to control diffraction which is a subset of Optical Proximity Correction (OPC), Phase Shift Masking (PSM), Off-Axis Illumination (OAI) and Immersion lithography [51]. The above RETs have made sub-wavelength lithography possible. However, the improvement from each technique represents a one-time improvement. At 45nm, the benefits from known RETs will largely be exhausted.

1.3 Research Objectives & Thesis Outline

In this thesis work, a complete analysis of different types of lithographic printability issues have been studied by performing lithography simulations on different layout features. The important effects of lithography-related issues on design and test at 65nm and 45nm CMOS technology have been discussed. The key motivation for

this work is to tie manufacturability related issues to design and devise solutions to help designers overcome the issues.

In Chapter 2, an analysis of effects of mask metal layer pitch variation on line width has been performed. Based on lithography simulation it is shown that these effects can make certain types of metal features unprintable. These unprintable metal pitches were termed as Forbidden Pitches [53]. This observation shows that layout features have to regular to be able to print on the wafer. Further analysis on diffraction issues of regular arrays is also presented.

Line width variation is a subject of major concern as it affects the device performance in integrated circuits. Forbidden pitch is one form of line width variation, where due to destructive interference of light, the line widths shrink dramatically. The occurrence of forbidden pitches in 65 and 45 nm nodes lead to breaks in interconnect lines. Thus lithography related line width reduction is a major source of concern for open faults. Chapter 3 deals with identification and detection of these open faults.

In Chapter 4, another form of line width variation called poly-gate length variation is discussed. Poly-gate length varies due to printability issues. Due to sub-wavelength lithography, the shape of the transistor often differs from idealized rectangles on the wafer. Transistor gate length variation has the largest impact on circuit timing and power performance since it directly affects both transistor switching speed and leakage power. Even in the presence of RET techniques, in silicon, printed gate length varies across its width. The average channel length for ON current and leakage current is different and cannot be assumed the same to have an accurate estimation of both leakage and timing. This becomes a modeling problem to find gate length and width that satisfy both currents. In this work, we aim to model the transistor into a combination of two or more transistors, each having different gate lengths and widths that can accurately estimate both ON and OFF currents.

It has been observed that random defects are no longer the dominant yield loss mechanism for the sub-65nm technologies. Design features, systematic and parametric effects are the chief limitations for yield. Hence traditional random defect based yield estimation will not be sufficient. In Chapter 5, new systematic and parametric yield estimation procedures are analyzed. In this work, a methodology to model the yield for variational lithography based systematic errors in current designs is proposed.

CHAPTER 2

FORBIDDEN PITCHES AND THEIR IMPACT ON DESIGN

2.1 Proximity Effects

As semiconductor manufacturing technology pushes towards the limit of optical lithography, the current state of art processes produce geometries or critical dimensions that are well below the optical wavelength (λ) where, critical dimension (CD) being the smallest width of a feature or the smallest space between two features [34]. For features that are spaced or sized smaller than the optical wavelength, effects due to diffraction, also known as optical proximity effects (OPE) come into play.

Optical proximity effect is a well known phenomenon in the field of photolithography [51]. Specifically, these proximity effects occur when very closely spaced layout patterns are lithographically transferred to a photosensitive material (photoresist) on the wafer. This effect results from structural interaction of light waves of closely spaced main features with the neighboring features of a mask layout. This typically leads to undesirable variation in the critical dimension and also exposure latitude of the main feature. The magnitude of the proximity effect depends on the spacing between two features in a mask [61].

One source of such manufacturing variation is depth of focus problem. The image plane may not be in perfect focus leading to a slight blurring of features. Such defects are highly systematic.

It is generally understood from fundamental optical microlithography, that the process latitude for dense features in a layout is much better compared to an isolated

feature for conventional illumination techniques. Scatter Bar placement, a subset of OPC is used to make isolated features behave like dense features and thereby improve their printability. But as technology is scaled, more aggressive types of illumination techniques have emerged, such as annular, multi-pole and quasar illuminations. These new illuminations techniques are utilized to aid in the printability of smaller features in the design, but restrict printing of lines in certain directions and produce large errors at other angles [34].

2.2 Theory on Optical Interactions

The lens used in optical lithography, converts the geometrical information of different features stored in the mask to spatial frequency components with appropriate amplitudes. This information is displayed at the exit pupil of its imaging system. This spatial frequency information is in turn transferred onto the photosensitive material coated on the wafer. Figure 2.1 gives an example of an imaging system as depicted by Hopkins [20]. Consider a transmission function as shown below. This function rep-

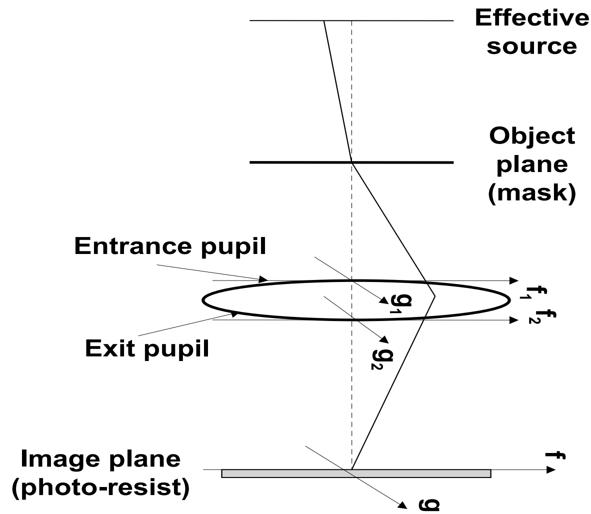

Figure 2.1. Optical photolithography system

resents the light intensity profile before the lens system based on the mask feature.

Parameter m is the width of the main feature, n is the width of the feature adjacent

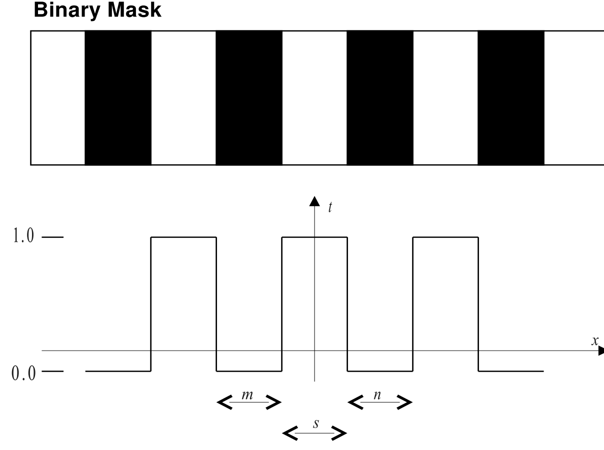

Figure 2.2. Binary mask and its transmission funtion

to the main feature and s is the edge to edge spacing between the main feature and the neighboring feature. α is the square root of transmittance. When α=0, the mask is a binary mask, if α=0.5, its a 25% attenuated phase shift mask (45° phase shift) and when α=1, it is a complete 180 phase shift mask also called as alternating phase shift masks.

The transmission function [53] of the mask at the entry pupil of the imaging system is given as,

$$f_1 = f(x_0) = 1 - \{1 - \alpha\}\{rect(\frac{x_0}{m}) + rect[\frac{x_0 - (\frac{m}{2} + s + \frac{a}{2})}{a}] + rect[\frac{x_0 + (\frac{m}{2} + s + \frac{a}{2})}{a}]\} \tag{2.1}$$

$$where, rect\left(\frac{x_0}{m}\right) = \begin{cases} 1; & \text{if } |x_0| < m/2 \\ 0; & \text{if } |x_0| > m/2 \end{cases} \tag{2.2}$$

At the exit pupil, the geometrical function is transformed into the spatial domain. Each point in the geometrical domain can be expressed in angular co-ordinates θ and ϕ. The corresponding frequency domain axes are.

$$a_x = \sin\theta\cos\phi; a_y = \sin\theta\sin\phi$$

The field F(a_x) at the exit pupil for the a_x axis is given by,

$$F(a_x) = FT\{f(x_0); \frac{a_x}{\lambda}\} \tag{2.3}$$

It is the Fourier spectrum of the object evaluated at the spatial frequency a_x/λ. And finally, the electric field at the image plane is given as

$$g(x) = \frac{1}{\lambda}\int_{-NA}^{+NA} F(a_x)e^{(\frac{2\pi i a_x x_i}{\lambda})}\,da_x \tag{2.4}$$

After integration of the function $F(a_x)$, it is found that the fields produced by the side features have a phase term associated with their intensity values. Thus it is this phase term of the neighboring features that play an important role in printability problems discussed here.

2.3 Forbidden Pitches

The fields produced by the light waves of the side features can be at the same phase or at 180 out of phase with that of the main feature. Same phase leads to constructive interference and hence leading to the main feature having better exposure latitude. Out of phase leads to destructive interference and thus reducing the exposure latitude. The fields produced by the light waves of the side and main features decide whether the existence of side features will increase or degrade the exposure latitude of the main feature.

Constructive interference was observed at certain pitches between metal lines and destructive interference at certain pitches. Pitches at which destructive interference take place and reduces the main feature width were termed as forbidden pitches. These are pitch ranges at which, the exposure latitude of dense features is found to

be worse than those of isolated features. Socha *et. al* [53] were the first to identify such pitches at 130nm technology. It was found that the pitches reduced by 10% of the critical dimension of the mask. Forbidden pitch phenomenon is a limiting factor for the current manufacturing technology . To estimate the effect such pitch ranges in the current 65nm and 45nm technologies, lithography simulation was performed on binary and alternating phase shift masks.

2.4 Simulation Metrics

Experiments were performed by varying the pattern density for a small window of a mask layer to find out the change in critical dimension of those metal lines. The patterns were (i) regular (ii) based on actual layout found in 90nm process and (iii) random. The metal line widths and pitches were chosen to be consistent with current practice [27].

A commercial aerial imaging simulator ($PROLITH^{\text{TM}}$) was used to find the forbidden pitches for the 65nm and 45nm nodes. Experiments were performed on 97.5 and 67.5nm metal lines to create a model of the photoresist after being developed. The metal lines used were of minimum width (3λ) based on design rules. The mask used for the simulation was a binary mask (BIM). The masks were presumed to be Alternating Phase Shift Masks (Alt-PSM).

A Quasar light source was used for this experiment. The wavelength of the light source was fixed at 193nm. The numerical aperture (NA) of the imaging system was 0.85. Figure 2.3 shows the simulation results for the occurrence of forbidden pitches in 65nm and 45nm technology nodes. The metal widths, the tuned pitch and the found forbidden pitch are also given. The occurrence of forbidden pitches has been found to be more prominent as we go down to the 65nm and 45nm regime. At the 65nm node, the minimum metal 1 line width (3λ) will be 97.5nm. This metal line was tuned to get the best possible resolution at a pitch of 174nm. The forbidden pitch

was found at 354nm. At the 45nm node, the minimum metal 1 line width (3λ) will be 67.5nm. This metal line was tuned to get the best possible resolution at a pitch of 135nm. The forbidden pitch was found at 165nm for which the critical dimension of the metal line goes to zero. Since forbidden pitch was found at 2X the separation for

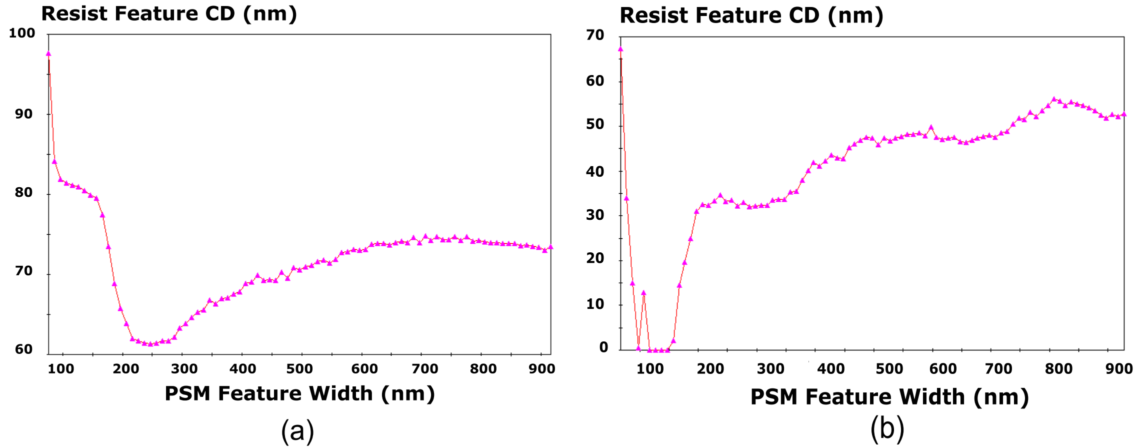

Figure 2.3. Forbidden pitches found at (a) 65nm and (b) 45nm technologies

45nm technology, aerial imaging simulation was performed on a mask with 3 times the regular gap. For the 65nm case, we found that the constriction at the forbidden pitch was 60% of the metal width. At this width, the metal lines are more resistive and cause a loss of performance. For 45nm technology, this resulted in fully open lines Figure 2.3. At this geometry lithography creates serious yield problems rather than traditional performance concerns [53].

This forbidden pitch issue will have to be addressed by design rules in 45nm technology such that gaps like are disallowed. However, such gaps are commonly found in current designs.

2.5 Design Impact

Optimal illumination and enhanced scatter bar techniques have been used effectively to suppress forbidden pitches in 130nm, 90nm and even in 65nm. However,

our previous study shows that this is not adequate at 45nm and below because they lead to unacceptable reduction in line width. Such gaps are a common feature in current layouts being done for 65nm and 45nm nodes. Simulation shows that such gaps result in constriction of metal lines as shown in Figure 2.4. Figure 2.5 shows

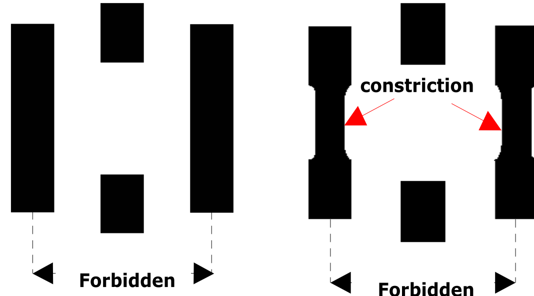

Figure 2.4. Constriction in metal lines

an example layout screen shot where a suspect forbidden pitch location can be seen. Such suspect features increase significantly in 45nm technology with a direct impact on yield. If these features are disallowed for lower metal layers, it forces the layout to

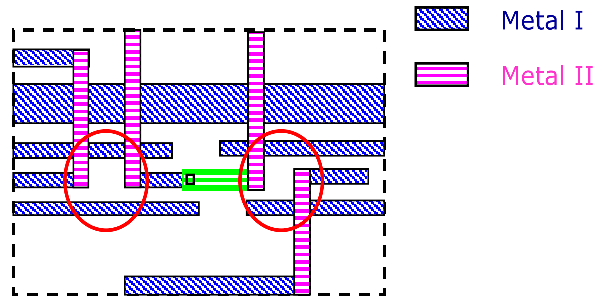

Figure 2.5. Example metal layer

be constrained to regular patterns in all lower metal layers. Pileggi *et al.* proposed a scheme called via-programmable gate array (VPGA)(figs), where, the metal lines have regular runs and the personalization for logic will be controlled by patterning the vias that connect one metal layer with another [42][57].

We investigated the viability of this technique as a potential 45nm technology solution by running aerial imaging simulation. In Figure 2.6 and Figure 2.7, we show results from aerial imaging simulation of two via patterns. In the simulation in Figure 2.6, all vias were separated by large distance while in Figure 2.7, they were separated by minimum spacing reflecting metal pitches. As the simulation results show, when the spacing between the vias follow the metal pitch, the vias may get fused together. However, when vias are placed significantly apart from each other with the same topology (Figure 2.6), it will yield normally. The results are also

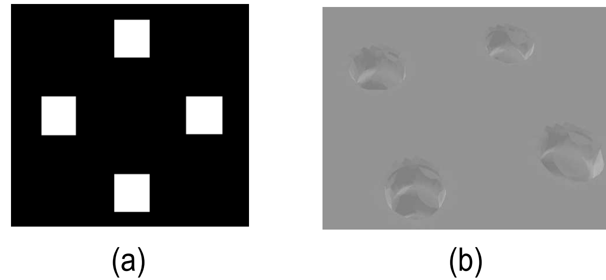

Figure 2.6. Case study I for diffraction problem in via placement: (a) the mask with four vias separated by more than the minimum separation; and (b) the photo-resist profile for the mask after development

supported by intuition. Interference effects from diffraction pattern are expected to diminish with distance. In terms of practical application, when vias are placed apart by significant distance, density is lost. We have not measured the loss of this density because logical physical design optimization is out of scope for this paper. When density is lost, average interconnect length increases and both power and performance gets worse. Previous studies on mapping a design to via programmable gate array did not consider such restrictions on via placement and was optimistic about power, performance and density potential of VPGA.

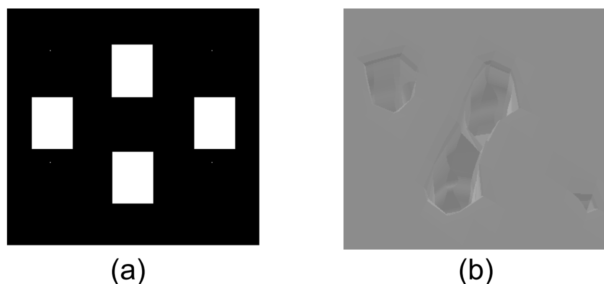

Figure 2.7. Case study II for diffraction problem in via placement: (a) the mask with four vias separated by the minimum separation; and (b) the photo-resist profile for the mask after development

2.6 Summary

Commonly found layout features are not regular in nature. We have found that sub-wavelength printability issues forbid many of these layout features. Aerial imaging simulation was used to study yieldability of various metal line patterns and via patterns. Since aerial imaging simulation for full chip is not practical, designs must be restricted to patterns that do not show yield problems under aerial imaging simulation. In this paper, we studied the printability issues for a selected set of patterns and inferred that in absence of any other resolution enhancement technique, many of the current layout patterns will be disallowed in future technology generations. Regular patterns are printable and have been hypothesized to form the basis for future designs. Our experiment on via placement shows that even the regular pattern designs such as Via Programmable Gate Arrays (VPGA) need to conform to an extended set of design rules, which when taken into consideration will further reduce layout density. Unfortunately, such restrictions may cost 25% in area and up to 3X in power-delay product. Our future research is focused on finding larger classes of permissible layouts that bring down the average interconnect length.

CHAPTER 3

MODELING AND TESTING OF LITHOGRAPHY RELATED OPEN FAULTS

3.1 Motivation

As the VLSI technology aggressively marches toward 45nm nodes and below, it is being greatly challenged not only by deep sub-micron design issues, but also by manufacturing and reliability issues. Current test strategies are mainly focused on random defects under the assumption that random particles based defects are the primary mechanisms for yield loss during manufacturing. But continuous scaling of devices has caused feature driven defects to play a major role in high volume manufacturing. It has been shown previously that systematic defects are on the rise compared to random defects for the current and future technology nodes [21]. Defects due to process-design interaction have a systematic nature. They therefore have a profound impact on yield. Hence the capability to detect and correct them is a requirement to continue to follow Moores law [32].

This paper deals with the extraction of such defect locations in a two-step process. In the first step a knowledge base of lithography related line width reduction/pullback is created by running aerial imaging simulation that simulates the lithography process. Such simulations are computationally expensive and can only be run on very small layouts. In the second phase, this library of knowledge is used to scan the physical design layout and identify possible open fault locations for different metal layers based on the line width variation of a specific layer. This is the main thrust of this paper. However, identification of potential open-fault locations is not enough. We need the

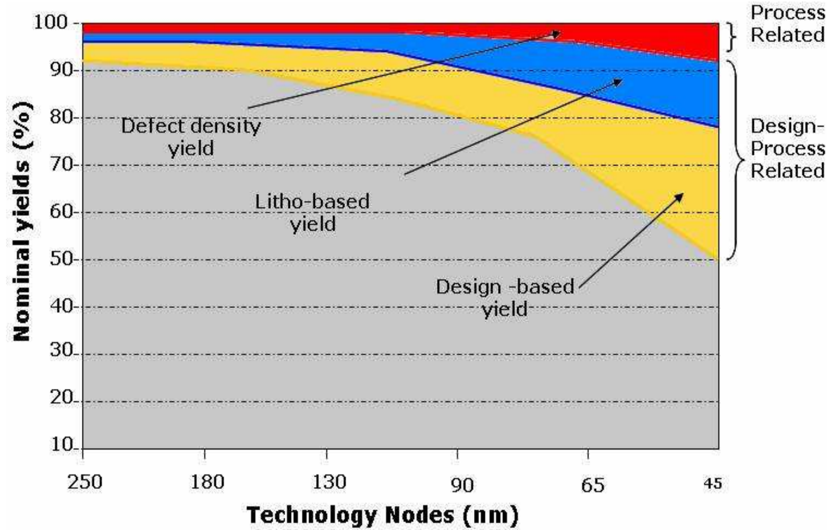

Figure 3.1. Particle driven Vs Systematic Defects [21]

ability to detect these faults. We present an important result showing that gate leakage in nano-CMOS technologies facilitate detection of open faults.

3.2 Related Work

Classically physical defects are modeled in the gatelevel representation as lines stuck at a specific logic value [1]. As the miniaturization and the density of devices on VLSI chips increases, bridging faults become more important and requires more refined fault models (such as transition fault model [1]) to improve the accuracy of the translation of physical defects into electrical faults.

With the consistentprogress of VLSI fabrication technology the problem of bridging faults became more prominent [39]. For the current generation VLSI circuits, the total number of bridging faults to be considered can easily become prohibitive if we restrict our fault modeling only in the gate level. This is so because in the gate level fault modeling a node can be short-circuited with every other node in the circuit, even if they are geometrically apart. This problem motivated the researchers to take the layout structure of the circuit into account while dealing with the fault

modeling and testing of bridging faults [35]-[25]. Ferguson et al. [12] extracted physical defects from the circuits layout and transformed them into transistor-level faults. They simulated the photolithography dependent defect generation mechanisms using Stappers [56] statistical defect models. Jacomet et al. [26] extended Fergusons basic idea to propose a technology and structure dependent fault extraction process. The extracted faults were weighted according to their likelihood of occurrence.

Other defect based tests have been conducted for process defects such as salicidation problems and high resistive shorts between metal layers [32]. IDDQ based testing was proposed for such defects and it was found that the tests werent sufficient enough to test all the faults.

As minimum feature sizes continue to shrink, the wavelength of light used in modern lithography systems becomes considerably larger than the minimum line dimensions to be printed. The fabrication industry approaches to tackle this problem by adding more and more stringent design rules. Kahng et al. [18] reported a detailed taxonomic description of these design rules and predicted the trend for future physical design.Such rules are termed as restricted design rules or guidelines. These guidelines are not binding, but if taken into consideration would make layouts printable.

Lithography related systematic defects in particular have not been analyzed in detail in the literature. Most of these defects tend to be temporarily solved by defining new design rules. As design rule explosion is a major concern, it is necessary to identify, model and test these defects.

3.3 Lithographic Defect Extraction

In this section we describe our defect extraction procedure and explain our experimental setup.

3.3.1 Methodology

Figure 3.2 shows the overall framework of our methodology. The framework consists of four phases viz. lithography simulation, defect characterization, layout fault extraction and logical fault translation. We now present a brief description of each of the four phases.

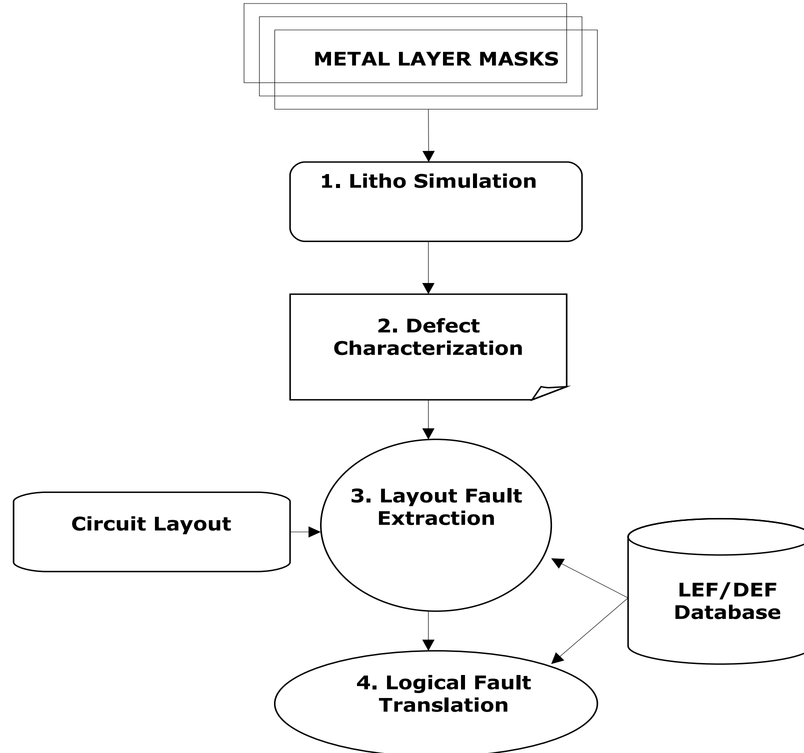

Figure 3.2. Defect Extraction Methodology

3.3.2 Lithography Simulation

In lithography simulation, metal layer masks are generated to find out the forbidden pitches. Experiments were performed by varying the pattern density for a small window of a mask layer to find out the change in critical dimension of those metal lines. We limit the simulation to the first two metal layers as we do not find the phenomenon prominent in other upper metal layers.

3.3.3 Defect characterization

Defect characterization involves the process of converting lithographic simulation data into models that can be used to find defect locations in layouts. A set of rules are defined for each metal layer taking pitch and simulation window. Forbidden pitches vary depending on the metal layer and the simulation region over which the pitches were found. The screen shot of a small are of layout and the defects occurring due to two forbidden pitch rules is shown . For 65nm process technology, metal 1 and metal

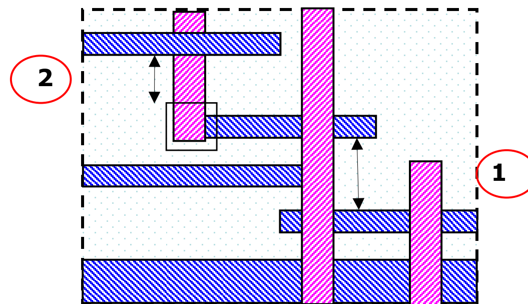

Figure 3.3. Forbidden Pitch Rules

2 layers had forbidden pitches at one particular pitch. The CD of these lines reduced to 30% of the actual tuned width. For the 45 nm process technology, the forbidden pitches were at a range of metal pitches for both metal 1 and metal 2 lines. Hence a defect type list with all the above conditions was made as a preprocessing step to facilitate the extraction process. One has to keep in mind that the list is device only for interconnect metal lines that connect between cells and not the metal/poly lines that are present within the cells.

The primary reason for this is that cell layouts are from standard cell libraries and are pre-characterized for a design technology. Pitch rules cannot be applied to intra-cell lines.

LEF/DEF Database

One significant problem with Layout data is that it usually does not have information about the design element names. Due to increasing sizes, layout databases (GDS2 files) do not have the hierarchical nets and cell names used in the design. Due to design complexity, design engineers now use automated tools to perform floor planning followed by place and route (P&R). Current P&R tools have the ability to generate certain formats of the processed layout such as LEF/DEF files. These files contain all connectivity information. The LEF (Library Exchange Format) files have definitions of standard cells used in the design. The standard cells are defined as macros in this file. The definitions include the size of the standard cell, its input and out pins, placement within the cell and the bounding box. LEF files can be treated as P&R based standard cell library database for a particular design technology. The

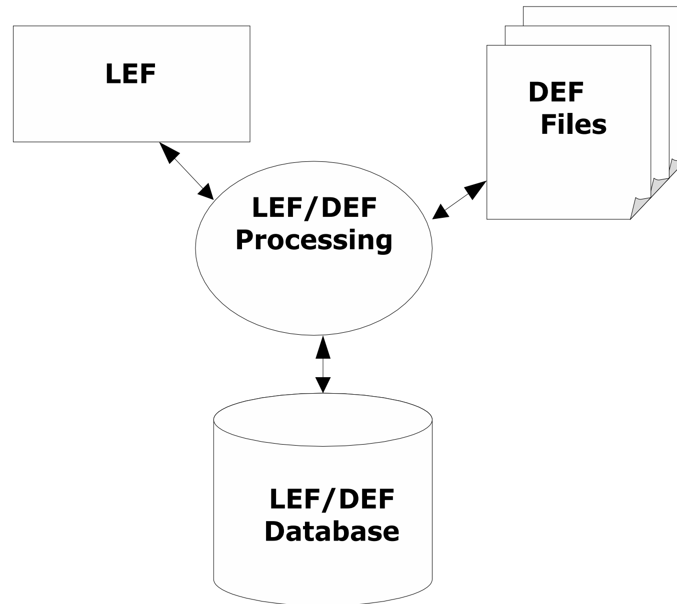

Figure 3.4. Processing Place & Route Data

DEF (Design Exchange Format) files have information about the placements of the standard cells defined by LEF files in the current layout. The contain information about the various sub-blocks in the design and their connectivity. It also list the place and route information i.e. placement and routing of cells using nets, and con-

nections of vias using segments. As another preprocessing step, shown in Figure 3.4, a LEF/DEF database is created by parsing the LEF/DEF files of the layout and storing them for further use in determining the location of the defects at the logical level.

3.3.4 Layout defect extraction

The layout defect extraction phase involves the layout parser and the fault extractor phases. The layout parser extracts the geometric features present in the design and stores them for future access by the fault extractor module. These are the polygons and shapes that fill the entire area which are used to create mask features.

Defects characterized based on lithography simulation have been transformed into layout rules for polygons. A Design Rule Check (DRC) based approach is followed [38]. This approach has been used to extract critical area in layouts but has not been used to extract layout defects caused by manufacturing issues. The screen shot of a small are of layout and the defects occurring due to two forbidden pitch rules is shown.

The error location indicated pointer 1 is caused by the placement of metal 1 lines spaced at the forbidden pitch. Due to this defect, the edges or the metal lines at that particular location shrink. They shrink to 30% in the case of 65nm node and to 0nm in the case of 45nm process technology. The defect type indicated by the pointer 2 is between the metal lines and the metal1-metal2 via. These defects are common in current layouts as they are highly dense. These via-metal line based defects can not only occur on interconnect lines between cells, but also within cell boundaries. Locations such as these may be the point at which a net fans out to two other nets. Defect at this via would cause both the nets to be at faulty.

Using the DRC rules written for such defects, the geometric features of the layout that were parsed are checked for such conditions. The area of the layout over which

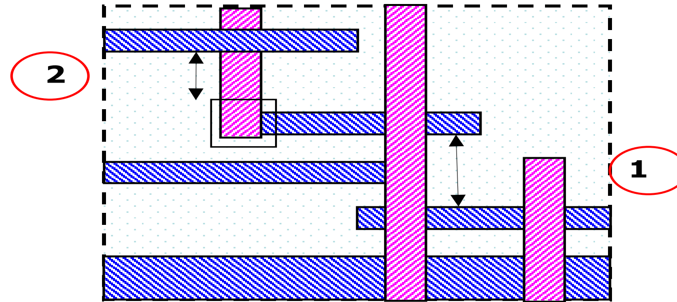

Figure 3.5. Forbidden Pitch error locations

these rules are checked depends on the simulation region over which the lithography simulation was conducted. If the condition is met, the error location is flagged and the coordinates are stored. The coordinates indicate the two lines that are at forbidden pitch. An example of such rule check is shown in Figure 3.6. The region marked by X is the forbidden pitch. The faulty nets on which constriction will happen are marked. We create a database of the LEF and DEF files at the beginning of our extraction

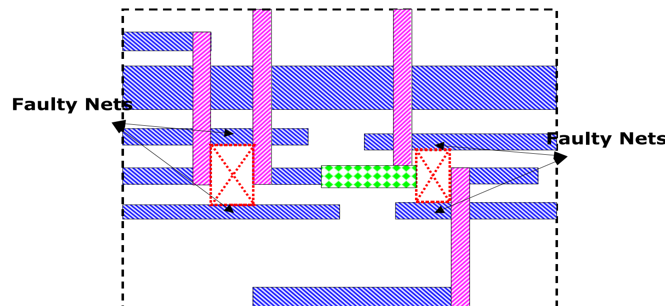

Figure 3.6. Open Fault nets

procedure. Using the database, we can perform queries to i) obtain the locations of bounding boxes of cell in the design, ii) hierarchical net, cell and pin names and iii) design elements. Once the error coordinates are obtained, database queries are done to find out what layout nets these fall under. Care must be taken that the complete hierarchical net names of the faulty nets are obtained. The above figure Figure 3.7

```
Rule Check:
Error Location => (4890,2178): (5994,2178)
Parser Output:
Faulty Net at → c880/N134
```

Figure 3.7. Example net location and its hierarchical name

gives example error locations and their corresponding net names and the metal layer over which it occurs. Connectivity information is also obtained for the net which will aid in obtaining the logic fault location.

3.3.5 Logical Fault Translation

The traditional approach in solving the problem of mapping layout level defects to logic faults has been solved using layout data (GDS2) and layout versus schematic information (LVS). A cross-mapping database is needed for processing the GDS2 and the LVS information. By using a physical design tool maintained by University of Michigan called Capo [22], we were able to the needed cross-referencing. Thus the physical faults are translated to logical faults.

3.3.6 Experimental Setup

We used a commercial imaging simulator called PROLITH^{TM}for all our lithography simulation. The metal layer masks used for the simulation were binary masks (BIM). The masks were presumed to be Alternating Phase Shift Masks (Alt-PSM). A Quasar light source was used for this experiment. The wavelength of the light source was fixed at 193nm. The numerical aperture (NA) of the imaging system was 0.93. Forbidden pitch simulations were performed on metal pattern masks for different simulation regions to get an optimal window over which the circuit layouts will be scanned for defects.

We used Cadence Composer v5141USR3 and Cadence Silicon Ensemble 5.3 to generate circuit layouts. Cadence SKILL language was used to define design rules and obtain the defect coordinates for all our designs. A LEF/DEF parser of Capo, a placement tool from University of Michigan was used to perform layout net name mapping from the defect polygons [47].

3.4 Fault Testability

As discussed earlier, open faults may be induced by lithography related issues. The location of open faults is dependent on the actual layout. However, identification of potential open-fault locations is not adequate. We need the capability to detect these faults.

3.4.1 Meta-stability of floating nets

Gate leakage (oxide leakage) is a significant concern for sub-90nm technology nodes. An open net connected to input of a gate is not a truly floating net. To illustrate this point, consider an inverter with an open input. Suppose, the initial voltage of this open net was 0V, the output of the inverter will be at VDD. This condition will lead to gate leakage current in the PFET which will raise the voltage of this open net. Similarly, if the input net started out at VDD, NFET will have gate leakage to bring down the input voltage. In order to determine, whether the floating input reaches a stability value, we ran simulation with different initial conditions for an inverter at 45nm with different initial conditions using BPTM models. It was observed that the input floating net stabilized at either 0.1735V or 0.579V, while the corresponding output voltages were at 0.8V or 0V respectively. This led us to observe that a floating net has bi-stable input states. Same observation is repeated with other gate types.

3.4.2 Testing bi-stable floating nets

Given that a floating input has two stable states, and the corresponding outputs are logical 0 or 1, there are two situations possible: (i) during test application, the floating input stays at a fixed state, i.e., it behaves as a stuck-at fault or (ii) it changes state. The latter can happen when there is charge flow from adjacent nets through capacitive X-talk. In this case a sufficient charge has to be applied to the net to change its value. The amount of the charge that can be injected by switching of a capacitively coupled net depends on the value of the coupling capacitance, which in turn depends on the coupling length. Another parameter that decides whether the output node will switch value or not, is the logic threshold voltage of the gate.

Critical coupling length is the length from the receiver over which the aggressor must be coupled to the victim floating net in order to induce a change in the output of the receiver (Figure 3.8). We used spice simulations to find the amount of charge required to be induced (coupled) on to the victim by adjacent nets in order for the voltage on the victim net to rise/fall to logic threshold. These capacitance values in turn are used to find the critical coupling length. Since coupling capacitance per unit length varies from layer to layer, critical coupling length will vary as well. In Table 3.1, we show critical coupling lengths by layer when the receiver gate is an inverter. The lengths for other gate types are quite similar.

Table 3.1. Critical lengths for different metal layers

	Critical length (microns)			
	65nm		45nm	
Metal Layer	Pull High	Pull Low	Pull High	Pull Low
M1	12.8	10	8	7
M2	15.2	13	11.5	10
M3	19	18	16.45	14
M4	20	18.5	16.5	14

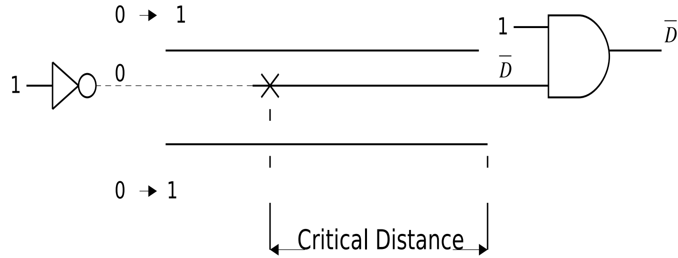

Figure 3.8. Critical length and side channel excitation technique

The lengths were obtained from Synopsys Raphael tool which was used to perform 3D field simulation to compute coupling capacitances. Table 3.1 shows the critical lengths for metal layers up to M4 in 65 and 45nm nodes. The critical length is different for transition from each meta-stable state of the input. Having obtained critical length for different metal lines, side channel excitation technique based approach can be used to test such lithography induced faults.

3.5 Results

In order to validate our methodology for open fault extraction, we performed the above mentioned extraction procedure on ISCAS-85 benchmark circuits. Table 3.2 shows the number of defect locations occurring in each of these circuits due to the forbidden pitch issue at metal layer 1. The total number of such defects for both 65nm and 45 nodes has been listed. We have also observed that the defects not only occur on interconnects between cells, but also within the standard cells. Table 3.3 shows the total defects that occur on metal layer 2. As metal 2 is not used for intra-cell routing, they are present only to connect different cells within the layout. In Table 3.4 we draw a comparison between our lithography aware open fault count and conventional stuck-at fault count for ISCAS-85 benchmark circuits. Different metal layers have different width and pitch. Consequently, lithography related problems

Table 3.2. Lithography aware open fault statistics for ISCAS-85 benchmark circuits (METAL I layer)

ISCAS 85 Circutis	Number of Defect Locations					
	65nm			45nm		
	Intra-cell	Inter cell	Total	Intra-cell	Inter cell	Total
c17	2	2	4	25	5	30
c432	34	6	40	370	27	197
c499	61	11	72	221	31	252
c880	8	8	16	177	73	250
c1355	60	10	70	241	84	325
c1908	15	15	30	265	91	356
c2670	25	23	48	336	45	381
c3540	34	26	60	361	134	495
c5315	91	65	156	315	231	546
c6288	62	22	84	344	272	616
c7552	126	41	167	698	153	851

Table 3.3. Lithography aware open fault statistics for ISCAS-85 benchmark circuits (METAL II layer)

ISCAS 85 Circutis	Number of Defect Locations	
	65nm	45nm
c17	0	3
c432	2	37
c499	2	72
c880	5	116
c1355	0	73
c1908	0	120
c2670	8	221
c3540	14	264
c5315	11	482
c6288	28	694
c7552	62	1073

Table 3.4. Comparison between lithography aware open fault count and stuck-at fault count for ISCAS-85 benchmark circuits

ISCAS 85 Circutis	Number of Defect Locations		Stuck-at fault count[48]
	65nm	45nm	
c432	42	200	524
c499	74	289	758
c880	21	322	942
c1355	70	441	1574
c1908	30	429	1879
c2670	56	602	2747
c3540	74	759	3428
c5315	167	1028	5350
c6288	112	1310	7744
c7552	229	1924	7550

affect them differently. Since, Metal layer 1 and 2 are usually the densest and by implication more prone to open faults, we present a comparison between the total number of possible open faults in these two layers with the total number of stuck-at faults. The point of this comparison is to show that lithography based technique narrows the fault list down and allows better targeting of test patterns. If it was the other way around, i.e., the total number of potential defects was very large then, there could be a potential issue of test data volume.

3.6 Summary

In this paper we made a case study for open faults induced by lithography related problems such as optical defocus. This has already been studied as a physical design problem. In this paper we examined the problem from a test aspect. We presented a methodology for extracting locations of such open faults. We showed that in general the number of such open faults is smaller than the number of stuck-at faults. We studied detection of such open faults and showed that due to presence of gate leakage, the floating inputs gravitate towards meta-stable voltage conditions. In this scenario a charge introduced by side channel excitation using capacitive cross-talk that can

exceed switching threshold of the receiver can lead to detection. We also computed minimum length of side-channels required for such open fault detection using side-channel excitation method. ATPG for side channel excitation is well-established in literature.

CHAPTER 4

MODELING IMPACT OF LITHOGRAPHY ON TRANSISTORS

4.1 Motivation

Line width variation is a subject of major concern as subthreshold leakage varies exponentially with change in line width (Figures 4.1, 4.2). Line width variation occurs due to focus variation, known otherwise as defocus. The sources of defocus include lens aberration, wafer tilting, resist thickness variation and other proximity effects [6]. Even with extensive RETs the gate shapes still vary from perfect rectangles which impact the leakage and timing of the whole chip. The channel region is the region under the gate poly that is characterized by the interface between the channel and the source/drain dopant profile. It has been shown in [14] that even under a proper rectangular gate poly; the channel region is still non-rectangular. This makes the poly gate, effectively non-rectangular. As the scaling of feature sizes continue, manufacturing issues and limitations cause non-rectangular transistors become unavoidable [49]. Non rectangular transistors have different gate lengths across their width. In todays advanced processes, they pose a big problem as current technologies have a compact transistor model that has only one length and width. Gate length variation directly affects the transistor ON and OFF state currents. This leads to large amounts of leakage and timing problems. With the constant scaling of technology, printing of non-rectangular devices cannot be avoided. Thus, a methodology to accurately model the effects the non-rectangular transistors has to be devised.

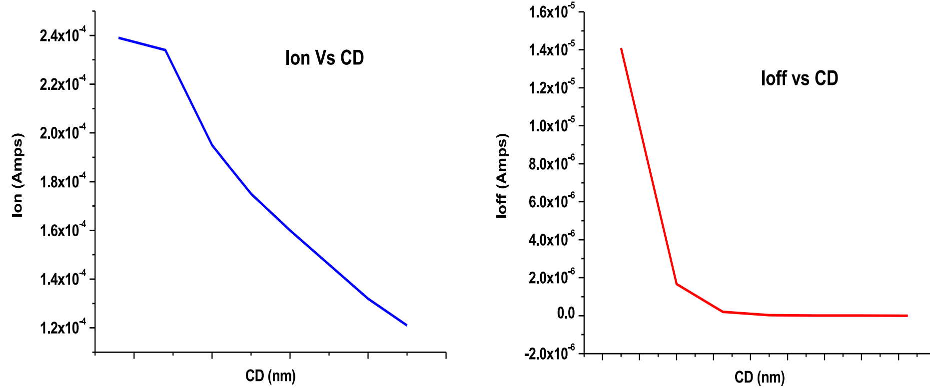

Figure 4.1. Variation of Ion and Ioff with CD

4.2 Previous Work

There have been several previous approaches to model the non-rectilinear property of gate SI for ultra deep sub-micron devices. Initial approaches attributed the problem to random variations at the gate edge also known as Line Edge Roughness or LER. LER is caused due to imperfections in the patterning process and leads to variation in gate length leading to irregularities in the doping profiles of the Source/Drain regions leading the change in threshold voltage [14, 8]. To model the effect of LER on circuit performance, the transistors were sliced based on the length across a constant width. Accurate simulations of these slices can be performed using 3D TCAD simulators. It is known that there exist non-uniform current flow effects at the abrupt junctions between the slices. Since 3D TCAD simulations take these parameters into consideration, these simulations take tremendous computation time [28].

A more feasible 2D TCAD based gate slicing and simulation approach is used for LER based analysis, where the non-uniform current flow effects were assumed to be negligible. A split-transistor model with n-number of parallel transistors in the place of one transistor was also proposed [2].

SPICE simulation based approach is much faster than the TCAD based approached. Equivalent gate length (EGL) is the method estimating the post-litho gate length using a lookup table based approach to find the slice currents and estimate an equivalent gate length. Two EGLs are defined to replace the original mask gate length; ON EGL when the device is on and for accurately estimating the delay in the circuit; OFF EGL when the device is off and to estimate the leakage [7, 44]. Though the EGLs can accurately model the device operating states(ON/OFF), they are not accurate in predicting in both the states. Since it is not possible to assign the gate length to each transistors during the operation, the EGL method is hard to be used in practice.

A unified non-rectangular device and circuit simulation model was proposed to achieve a single gate length for all states of operation of the transistor [49]. This scheme proposes to add modeling card to the circuit which changes the gate length based on the region of operation of the transistor. Though the model provides a good estimation of the timing and power, the modeling card cannot be added for each transistor in a large circuit. This would create a bloated model.

The main drawback of this and all other SPICE based approaches is that, they assume that the threshold voltage remains constant regardless of channel length of a slice. It is known that fringing effects due to LER, well proximity effects and dopant scattering due to shallow trench isolation are all factors that result in changing the threshold voltage of a device [41][43].

We summarize this section with the following observations:

- Poly-gates of transistors are non-rectangular.

- Non-rectangular transistors result in differing equivalent gate lengths for transistor ON current, OFF current and intermediate currents.

- If design objectives related to circuit performance, power and leakage currents have to be met we need better models for transistors.

Modeling non-rectangular transistors for spice simulation is the main objective of this work.

4.3 Proposed Modeling Methodology

Our proposed post litho-transistor modeling and standard cell characterization methodology is shown in Figure 5. OPC is run over the standard cell layout masks and lithography simulation is done to find the change in CD due to pitch, dose and focus variations. The post-litho transistor shapes are modeled using our technique and the cells are characterized.

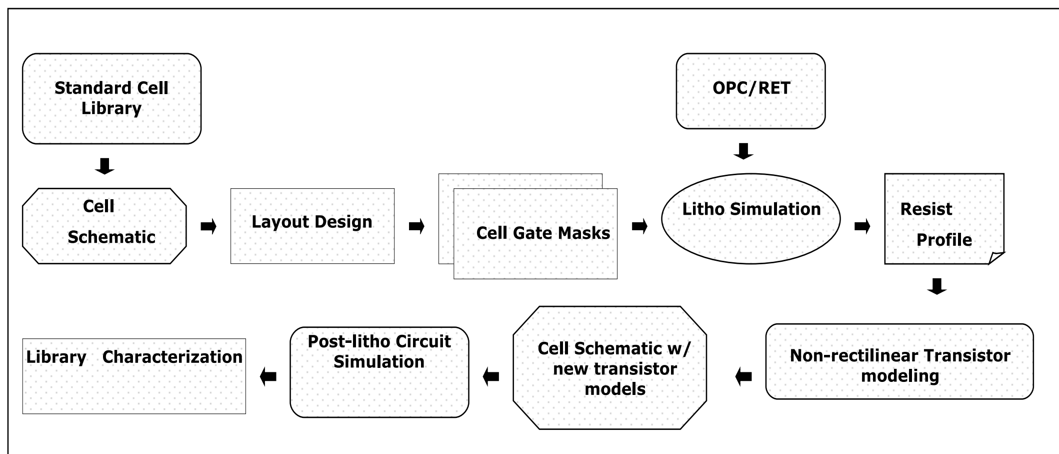

Figure 4.2. Proposed post-litho transistor modeling technique

4.4 Gate Characterization

Our transistor modeling approach is also based on poly-gate slicing. However, instead of defining an EGL for a given transistor width, we treat this as a modeling problem consisting of fewest parallel transistors (minimum of 2) with connected

sources, gates and drains, where each of these transistors have different widths and lengths. The overall objective is that, as a group these transistors should predict the ON and OFF currents of the corresponding non-rectangular transistor that it is trying to model. Next we present a detailed description of the gate slicing methodology and the equivalence steps. The non-rectangular poly-gate is obtained by from lithography simulation using a rectangular mask.

4.4.1 Gate Slicing Technique

The non-rectangular poly gate is scanned to obtain the variation of gate length across its width. Even when the simulation is done at the best process corner based on the FEM matrix for the mask, the variation between the center and edge of the gate is quite significant. Since the poly-gate is essentially not rectangular in geometry, a

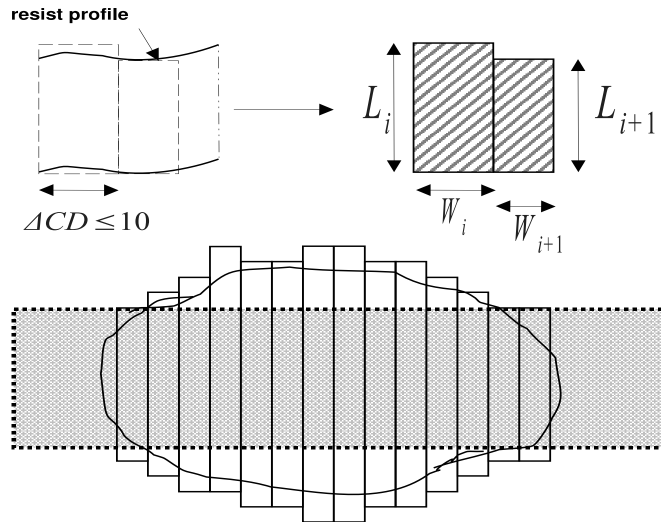

Figure 4.3. Gate Slicing Technique and Sliced poly-gate

constant length cannot be assumed over the entire width. Gate slicing methods have been used in various other transistor modeling approaches [46][8][44][7]. All these approaches approximate the gate length over a particular width. The gate width is

calculated by approximating a non-rectilinear contour with an equivalent rectangular active region with the same area [46].

In our methodology, we do not approximate the gate length over a region. We slice a particular width, by scanning through the resist profile. A gate width is determined as the region over which the variation of gate length is less than 10% of the critical dimension (CD). illustrates our gate slicing methodology. By this approach, the gate slices capture the variation in gate length across the width of the resist profile. Finer slices can be obtained by reducing the cut-off of variation below 10%. However, we are trading off accuracy at various levels in the subsequent steps and such margins should be weighed against levels of accuracies in all steps.

4.4.2 Current Modeling

Each gate slice is simulated and the drain currents at different values of gate-source (Vgs) and drain-source voltages (Vds) are measured. The first order drain current for each slice at different regions of operation is given as:

$$I_{DSi} = \frac{W_i}{L_i}\mu_{n,p}C_{ox}exp\left(\frac{V_{GSi} - V_{thi}}{\eta V_T}\right) \tag{4.1}$$

$$for V_{GSi} \leq V_{thi}(subthreshold)$$

$$I_{DSi} = \frac{W_i}{L_i}\mu_{n,p}C_{ox}\left(V_{GSi} - V_{thi}\right)V_{DS} \tag{4.2}$$

$$for V_{GSi} > V_{thi}, V_{DS} \leq V_{DSAT}(linear)$$

$$I_{DSi} = \frac{W_i}{L_i}\mu_{n,p}C_{ox}\left(V_{GSi} - V_{thi}\right)^2 \tag{4.3}$$

$$for V_{GSi} > V_{thi}, V_{DS} > V_{DSAT}(saturation)$$

The total current of the non-rectangular transistor is the sum of drain-source currents of each slice i of width Wi and length Li and is given by,

$$I_{DS} = \sum_{i=1}^{n} I_{DSi}\left(L_i, W_i, V_{TH}, V_{DS}, VGS\right) \tag{4.4}$$

In case of Equivalent Gate models (EGL), transistor currents are considered separately for two cases: (i) in saturation mode with full turn ON and in the (ii) OFF mode when the transistor is fully turned off . The OFF-EGL is calculated for the case when Vgs¡ Vth and the ON-EGL is calculated when Vgs ¿ Vth as shown.

$$I_{DS_{ON-EGL}} = I_{DS}\left(L, W_{ON}, V_{DS}, V_{GS}\right)$$

$$I_{DS_{OFF-EGL}} = I_{DS}\left(L, W_{OFF}, V_{DS}, V_{GS}\right)$$

One issue with this approach of modeling is that the issue of change in drain current as V_{DS} and V_{GS} change is not incorporated. Hence this is not an appropriate model for the non-rectangular transistors as it is valid in only two regions. In [49] a modeling card was used to modify the gate poly length as V_{DS} and V_{GS} change. A small δI is added to the existing device model to estimate the currents accurately. But it does not take the change in threshold voltage of the transistor across its width into consideration.

4.5 Gate Length Computation

In this section we propose to model the non-rectangular transistor into a pair of parallel transistors, each with differing lengths and widths. This model is based on 3D device simulation of gate slices. Since normal SPICE based techniques do not take the change in V_{th} along the width of the transistor into consideration, this model gives an accurate measurement of the drain currents for each slice.

4.5.1 Modeling gate length variation

I-V curves of the gate slices are generated using a commercial device simulator called *Davinci*™ from Synopsys Corporation. The gate slice length and width are given as input parameters to this simulator. The dopant concentrations, junction depths, oxide thickness and source/drain widths are specified and tuned to the 45nm technology node specifications. The initial values are taken from ITRS 45nm technology targets. The known drawback of all other previous SPICE based approaches is that, they do not consider the effect of threshold voltage variation along the width of the transistor. Through 3D device simulations, this effect is modeled accurately.

Let the total current as obtained by equation 4 be I_{DS-ON} and I_{DS-OFF}. Now consider a rectangular transistor whose length is given as L_x. For any given L_x, we can find a corresponding width W_x, such that this transistor has a drain current I_{DS-ON}. Similarly, we can also find a different W_x for which this transistor has OFF current that equals I_{DS-OFF}. By varying L_x over the channel length of various slices, we can obtain a plot as shown in 4.4. Our objective is to find a pair of transistors T1 and T2 with lengths L1 and L2 and corresponding widths W1 and W2 such that $I_{DS-ON}(T_1)$ + $I_{DS-ON}(T_2)$ is within 5% of I_{DS-ON} and $I_{DS-OFF}(T_1)$ + $I_{DS-OFF}(T_2)$ is within 5% of I_{DS-OFF} as shown by equations 2 and 3. Please note that we are targeting 5% accuracy from the model in this step. If the accuracy expectation is higher, the numbers should change correspondingly. In either case, it is not guaranteed that we can meet this goal with just T_1 and T_2 and additional transistors may be needed.

$$f_{ON,OFF}(T_1, T_2) = I_{DS-ON,OFF}(T_1) + I_{DS-ON,OFF}(T_2) \tag{4.5}$$

$$\Delta I_{ON,OFF} = |I_{DS-ON,OFF}(total) - f_{ON,OFF}(T_1, T_2)| \tag{4.6}$$

$$\Delta I_{ON,OFF} \leq 0.05 * I_{DS-ON,OFF}(total) \tag{4.7}$$

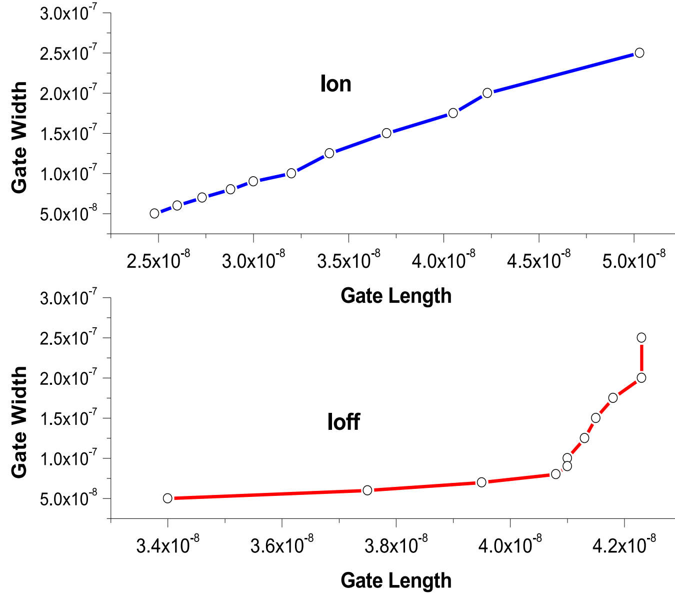

Figure 4.4. Length and Width variation for Ion/Ioff

In order to automate the process, we first constructed a piecewise linear (PWL) model of I_{DS-ON} and $\log(I_{DS-OFF})$. This is because, the OFF current rises exponentially with reduction in transistor channel length for a given transistor width. Such quantization also leads to small error that varies with step size, but lends itself to better automation.

The PWL current equations along with a set of constraints on lengths and widths are fed to a linear programming solver to obtain a suitable solution. In our experiments, we always succeed to satisfy these linear equations with just two transistors. The two transistors have different lengths (L_1,L_2) and widths (W_1,W_2). They can be incorporated as two parallel transistors to compensate the effect of the non-rectangular transistor. Table 4.1 shows the results obtained using our methodology. It can be seen that the transistor pair from our model satisfies both ON and OFF currents with negligible difference.

Table 4.1. Results from transistor modeling

	T_1		T_2			
	L_1(nm)	WL_1(nm)	L_2(nm)	W_2(nm)	I_{OFF}%diff	I_{ON}%diff
Drawn	45	250	—		88.5	12.5
New_Model	41.5	100	42.75	95.5	2.6	0.6
ON_EGL	50.5	250	—		96.2	0.8
OFF_EGL	42.25	250	—		2.2	22.4
Avg-GL	44.6	250	—		73.8	10.2

4.5.2 Model parameter extraction

For model parameter extraction, BSIM 4.3.0 is used with Synopsys Aurora™. The purpose of this tool is to match the current data obtained from modeling the gate lengths, to parameters to be used in circuit simulation with minimum RMS errors [54]. The tool takes in many AC and DC parameters, coefficients and also a few process parameters. Important process parameters that have to be provided before extraction include, gate oxide thickness, source/drain doping, S/D junction depth, poly-silicon gate doping concentration. The electrical gate oxide thickness is given to take the quantum effects into account and to estimate leakage currents [19].

It is important to note here that, by this extraction procedure, the two transistors that were modeled from the gate slices will have different BSIM models, each having its own, lengths, widths, threshold voltage, oxide thickness etc. Thus it can accurately estimate the leakage and timing delay of a circuit. This BSIM model can now be used in circuit simulation of standard cells and by implication on the entire design.

4.6 Results

The simulation results of our modeling approach are shown in this section. We validate our transistor modeling scheme at the device level and at the standard cell level. The drain currents at different regions of operation of the transistor form the

key metric to validate at the device level. At the standard cell level, the rise and fall delays and leakage power are the important factors.

As mentioned previously, the EGL based results are accurate for certain discrete values of V_{GS} and V_{DS}. ON-EGL accurately models the transistor at $V_{GS} = V_{DD}$, and OFF-EGL models at $V_{GS} = 0$. Hence the drain current values from our model are compared to the values generated by the EGL models at those discrete values. In standard cells, the delay and leakage power estimated by the EGL and average gate length methods are inaccurate by a large margin. The following subsections explain in detail about the simulation environments, results and inferences.

4.6.1 Transistor Model Validation

The transistor model generated after gate slicing and 3D device simulation is validated by comparing it to other methods that have been proposed previously. 4.5

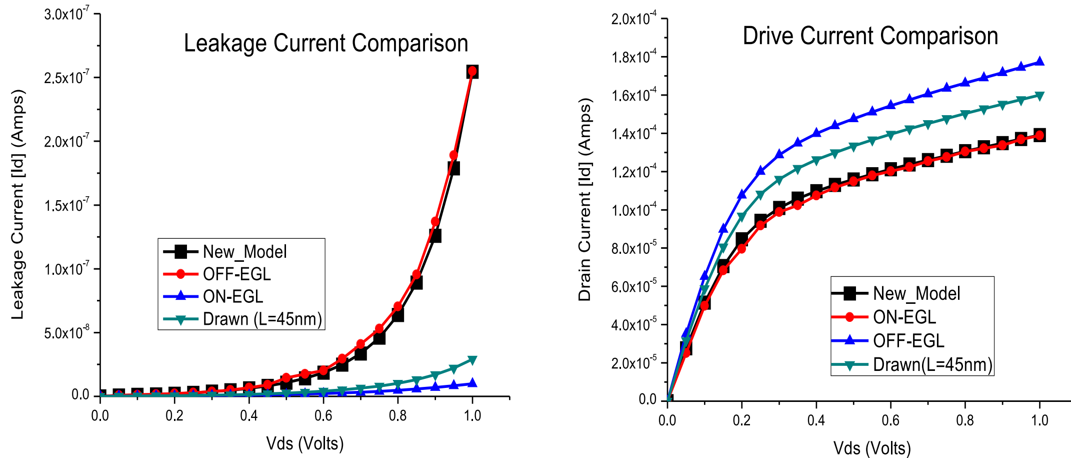

Figure 4.5. Transistor Leakage current and Drive current comparison

shows the ON and OFF current behavior for our new model and other known models. The figure shows the transistor behavior as a device when it is completely in the ON state. This happens when the gate-source voltage V_{GS} is equal to V_{DD}. It can be

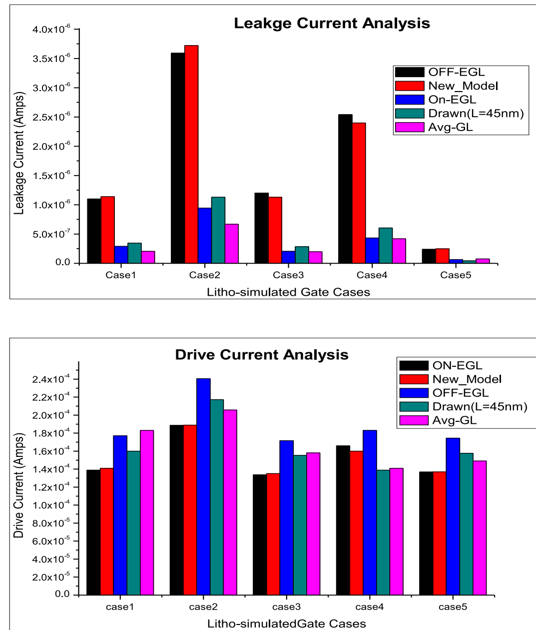

Figure 4.6. Leakage and Drive current analysis for different gate cases

seen the graph that the drive current using our new model is very close to the one predicted by the ON-EGL method.

Also, in figure the OFF-EGL and the drawn length currents are higher. shows the transistor behavior in its OFF state. This occurs when V_{GS} is equal to 0. This current, termed as leakage current has been accurately modeled by our new model and is compared with the OFF-EGL, ON-EGL and drawn gate models. It can be seen that our new model and the OFF-EGL are quite close to each other. The other two models do not accurately model the transistor behavior.

Graphs in Figure 4.6 show the variation in leakage power and propagation delay for different gate cases. The cases were obtained based on proximity effects, defocus and exposure dose variations. A comparison of different gate models for each is shown.

4.7 Standard Cell characterization & Circuit Simulation

The new post-litho transistor model is validated by circuit simulation after standard cell characterization. The device I-V characteristics and other process parameters are used to extract the circuit simulation model of the transistor. This procedure was explained in detail in section 4.2.

Complete extracted models are used to characterize the standard cells. For our experiment we used the INV and NAND standard cells. The circuit simulation model for these two cells is shown in Figure 4.7 and Figure 4.8.

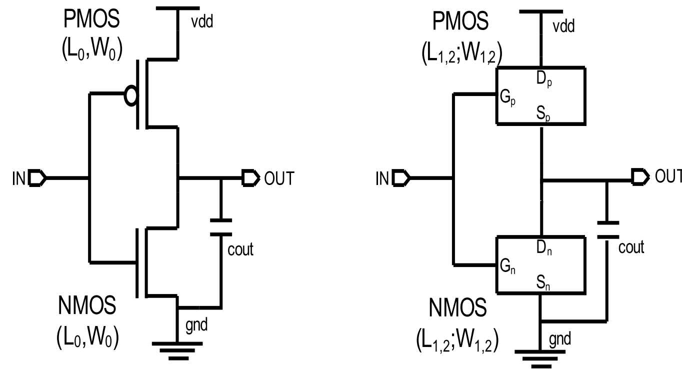

Figure 4.7. Inverter circuit simulation model

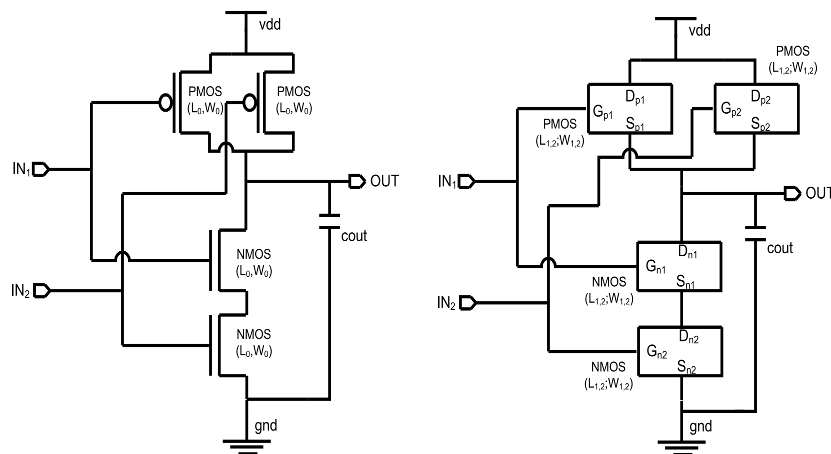

Figure 4.8. NAND circuit simulation model

4.7.1 Leakage Estimation

Leakage power has become an important component in the total power dissipated in a chip for sub-90nm technologies. And since the impact of leakage is also on thermal issues, it is important to analyze this for all standard cells in the design. Circuit simulation was performed on the INV and NAND standard cells. Leakage power dissipation for different models are tabulated in Table 4.2. In the table, all the leakage values are compared to the OFF-EGL model as it has been proven to give an accurate estimate of the leakage. Drawn indicates the mask gate length (L=45nm) in our case. Avg-GL model is where the algebraic average of the variation in gate length is taken as the approximate gate length for simulation. The ON-EGL model

Table 4.2. Leakage Power estimation

Std Cell	Gate Model	Leakage Power (nW)	% Diff
INV	OFF-EGL	21.18	-
	Drawn	18.13	-14
	Avg-GL	31.86	+50
	ON-EGL	17.86	-15.6
	New_Model	20.67	-2.40
NAND	OFF-EGL	23.40	-
	Drawn	19.11	-18.3
	Avg-GL	33.91	+40.60
	ON-EGL	20.51	+12.3
	New_Model	22.95	-1.90

based gate length is also simulated. It can be seen that, our new model, estimated the leakage power within 2.5% variation. All other models differ more than 10%.

4.7.2 Timing Delay Estimation

Overall timing of the chip is important for the logic to work. Hence timing analysis of the currently proposed model has to been done to validate it. As done previously for the leakage estimation section, various other models have also been used for circuit simulation. The timing of the circuit is compared for both rising and falling delays

as shown in Table 4.3. The timing of the standard cell is compared to the ON-

Table 4.3. Timing Delay Estimation

Gate Model	Rising Delay (ps)	% Diff	Falling Delay (ps)	% Diff
Inverter Std Cell				
ON-EGL	39.51	–	39.86	–
Drawn	44.52	12.6	43.52	9.20
Avg-GL	43.08	9.03	42.76	7.30
OFF-EGL	36.33	-8.05	33.06	-17.1
New_Model	38.83	-1.40	40.51	1.60
NAND Std Cell				
ON-EGL	17.31	–	68.36	–
Drawn	27.47	58	73.14	6.9
Avg-GL	25.32	46	70.47	3.80
OFF-EGL	22.21	28	65.04	4.80
New_Model	17.44	0.75	68.12	0.35

EGL model. It is seen that our model estimates the rising and falling delays for the simulated cells within 2% variation. All other models estimate inaccurately with a variation of more than 10%.

4.7.3 Model Validation for Lithographic variation

There are multiple sources of lithography related variation. Most common ones include proximity effects, depth of focus and exposure dose variations. For a post-litho transistor model to be valid, it must also be able to model depth of focus and exposure dose variations. We conducted experiments, where we took two cases of lithographic variations; 10% depth of focus and 10% exposure dose variation. The models were correspondingly tuned for these extremities. The results based on such variation are shown in Table 4.4 and 4.5.

DOF w/V is the new model with 10% variation in the depth of focus with respect to the ideal position of the focal plane. Dose w/V is the new model with 10% exposure dose variation with respect to the ideal exposure dose value.

The ideal values of DOF and exposure dose are determined based on the best process corner obtained from FEM matrix. It can be seen that the ON-EGL is

Table 4.4. Timing Delay Estimation with Variations

Gate Model	Rising Delay (ps)	% Diff	Falling Delay (ps)	% Diff
Inverter Std Cell				
New_Model	38.83	–	40.51	–
ON-EGL	39.51	+1.40	39.86	-1.60
DOF w/V	47.38	22.0	52.67	30.1
Exp Dose w/V	27.96	-27.9	35.42	-12.5
NAND Std Cell				
New_Model	17.44	–	68.12	–
ON-EGL	17.31	-0.75	68.36	-0.35
Drawn	26.43	51.7	74.92	9.82
Avg-GL	14.33	-17.8	62.17	-8.7

constant for all the cases of variation and cannot be used as an accurate model to estimate propagation delay and OFF-EGL cannot be used for leakage estimation taking variation into consideration.

Table 4.5. Leakage Power estimation

Std Cell	Gate Model	Leakage Power (nW)	% Diff
	New_Model	20.67	–
INV	OFF-EGL	21.18	+2.40
	DOF w/V	16.89	-18.2
	Exp Dose w/V	26.89	+30.1
	New_Model	22.95	–
NAND	OFF-EGL	23.40	+1.90
	DOF w/V	19.26	-16.1
	Exp Dose w/V	28,69	+20.2

4.8 Design Implementation

The newly characterized standard cell library is used in small designs and the leakage power is estimated. A subset of ISCAS circuits as shown in Table 4.6 is used to validate the model. The input patterns for the above experiment were based on

the paper by Ganeshpure et.al [15] and were obtained upon request from the authors.

Table 4.6. Design Leakage Estimation

Cell Gate	OFF-EGL (nW)	ON-EGL (mW)	Model-based (mW)
c17	0.0254	0.0118	0.0223
c499	0.0331	0.092	0.0291
c880	0.1973	0.0864	0.1535
c1355	0.4067	0.1894	0.299
c3540	1.4378	0.8931	1.278
c5315	3.513	1.309	2.987

4.9 Summary

Transistors on silicon are not rectangular. Transistor models based on traditional EGL cannot satisfy accuracy across delay, leakage and power estimations. Such models have a balloon squeeze problem, whereby fixing the model for accuracy in one parameter, causes large error in other measures. In this paper we proposed a new post-litho transistor modeling methodology to model non-rectangular transistors accurately for all measures, while also taking Leff dependent Vth variation into consideration. Our model was validated and compared with other gate length models. Our simulation results show that using the new model, the timing and leakage can be estimated within 5% of variation, which was our initial goal. The proposed methodology allows further improvement in accuracy by trading off accuracy vs. model size. As nanoscale transistors are no longer rectangular in silicon, this work signifies a major improvement over current practices.

CHAPTER 5

YIELD MODELING AT LITHOGRAPHIC PROCESS CORNERS

5.1 Motivation

Photolithography is at the heart of semiconductor manufacturing process. To support continued scaling of transistors, lithographic resolution must continue to improve. At todays volume manufacturing process, a light source of 193nm wavelength is used to print devices with 45nm feature size. To address subwavelength printability, a number of resolution enhancement techniques (RET) have been used. Although techniques such as Resolution Enhancement Techniques (RET) which involve optical proximity correction (OPC), phase shift masking (PSM), off-axis illumination (OAI) have been used to greatly improve the printability and better the manufacturing process window, they cannot perfectly compensate for these lithographic deficiencies [58][3][61]. While RET techniques allow printing of sub-wavelength features, the feature length itself becomes highly sensitive to process parameters, which in turn detracts from yield due to small perturbations in manufacturing parameters. Process variations in sub-wavelength lithography such as optical defocus, exposure and dose variations can be spatial and random. Wafer tilting can cause DOF variation, resist coat thickness varies (Figure 5.1). Apart from random or correlated random variations, systematic variations such as lens aberration and out-of-band radiation from optical sources also affect the final product yield.

Statistical approaches to yield prediction and yield modeling have been developed over a number of years [36][30]. These methods estimated functional and parametric

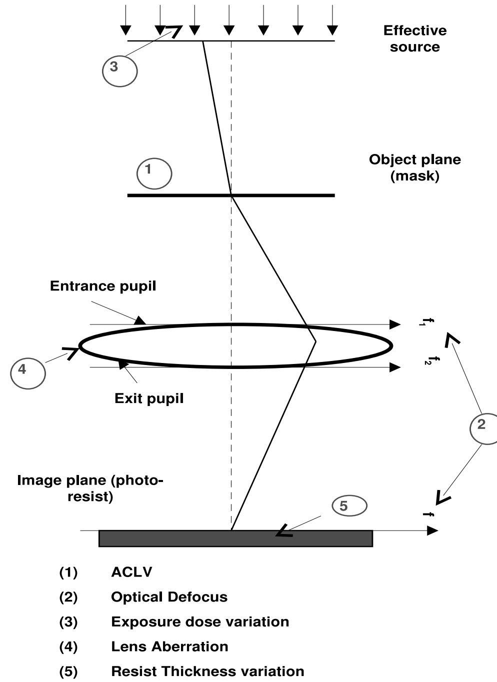

Figure 5.1. Sources of yield loss in today's lithographic manufacturing process

yield of a design based on random particulate defect densities in the manufacturing process. For many generations, yield loss due to random defects has been well understood and thoroughly researched [37][31][36][30]. But as technology scales, the influence of random defects from particulates has been in the decline compared to lithography and design related defects. Hence estimation of die yield based on lithographic sensitivities is a must for the current and future technology generations.

5.2 Yield Loss Mechanisms

Yield loss was primarily a problem solved by the manufacturing side until 180nm technology nodes. Design rules which define the window of variation of design variables such as width, spacing etc that would guarantee manufacturability were written. Designs were created using such golden rules and were assumed to be yieldable. This

assumption began to fail with the advent of sub-wavelength lithography. A major material shift from aluminum to copper also made many designs un-yieldable.

At the current 65nm node, the different yield loss mechanisms (YLM) are random, systematic and parametric YLMs.

5.2.1 Random Yield Loss Mechanisms

For many process generations, yield loss due to random particle defects have been well understood and thoroughly researched [10][11][36][30][31]. Random particles can attach to the wafer during semiconductor processing and may cause and unintended bridging/short between metal lines or an open in a single interconnect line. Such defects can destroy the functionality of the chip.

Random defects such as these vary from die to die and have very little correlation between wafers. The random defect based yield has been modeled as a Poisson equation based on Critical Area Analysis [11]. Based on Poisson distribution, it has been defined that the yield is a function of critical area and defect density. Where critical area (CA) is defined as the area over which when the center of the defect is placed, it can cause a bridging between two line or can cause a complete open on an interconnect line. The formula for yield is given as,

$$Y = e^{-(A_c * d)}$$

Where A_c is the critcal area and d is the defect density of the design considered. The designer controls the critical area and the foundry controls the defect density. The designer increases the yield by reducing the critical area of the die. Different test structures have been defined to characterize such particle defects in the past [59][31].The challenge here is to obtain a defect failure rate that is less 1 parts per billion.

5.2.2 Systematic Yield Loss Mechanisms

Systematic effects are those that are layout dependent and affect circuit functionality. Dopant fluctuations, lithographic sensitivity, poor wafer planarization and stress effects are some of the factors that lead to systematic defects.

Such defects are a strong function of IC layout [37]. They are highly spatially and temporally correlated [37]. Since they are strongly correlated to specific layout patterns, their failure rate can be orders of magnitude higher than that of random particle defects. The characterization of such defects is crucial for the overall yield of the chip. Some of the common systematic defects that can be seen in todays layouts and lead to yield loss include; (i) metal interconnect shorts/opens due to variation in width with respect to spacing, dose and focus variations, (ii) via stacking failure as a measure of interconnect length, (iii) metal line-end shortening due to misalignment, (iv) random dopant fluctuations, and (v) poor wafer planarization. Even with the currently available RETs, design patterns often experience such distortions. In such cases, this also affects transistor leakage, switching delay and power consumption .

There have been solution proposed to such problems, but all of these have been some kind of modification of design rules. Such design rules were termed as radically restricted design rules (RDRs). But since there are variety systematic defects, defining rules for each has lead to design rule explosion. Another issue is that most of these design rules cause contention and hence make the design difficult to manufacture causing more systematic failures [60].

5.2.3 Parametric Yield Loss Mechanisms

Parametric yield loss occurs due to variation of electrical performance of a design. Traditionally, variations in design parameters have been guard-banded over a 3σ distribution.When paramter variations go beyong the specified gaurd band, they are categorized as defects. It is assumed that such variations are between intra-die

and inter-die. But as process technologies improve, these variations violate the distribution and this leads to parametric loss 5.2. Such parametric loses lead to improper

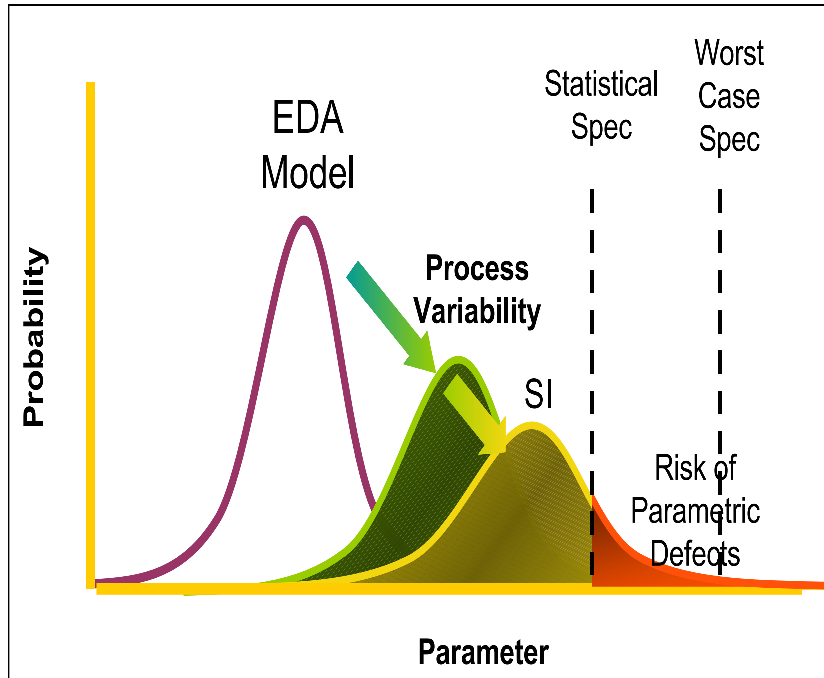

Figure 5.2. Parametric variation distribution

operation of the device and also do not lead to design closure. Characterization of parametric yield loss requires different test structures & methodologies, different mismatch modeling mechanisms and statistical static timing simulation techniques [60]

5.3 Variational Lithography

Across Chip Linewidth Variation (ACLV) is the most significant contributor towards chip leakage and timing variability. ACLV is dependent both on process and layout topology. Sources of ACLV include optical defocus, exposure dose variations, pitch variations, resist thickness, lens aberration etc. Although Optical Proximity Correction (OPC) techniques have been implemented in todays designs, the effects

these variations have not been completely eliminated [50]. The following sections include a more detailed description of the sources of variation we consider and their impacts on the design front. Depth of Focus (DOF) can be defined as a point on the

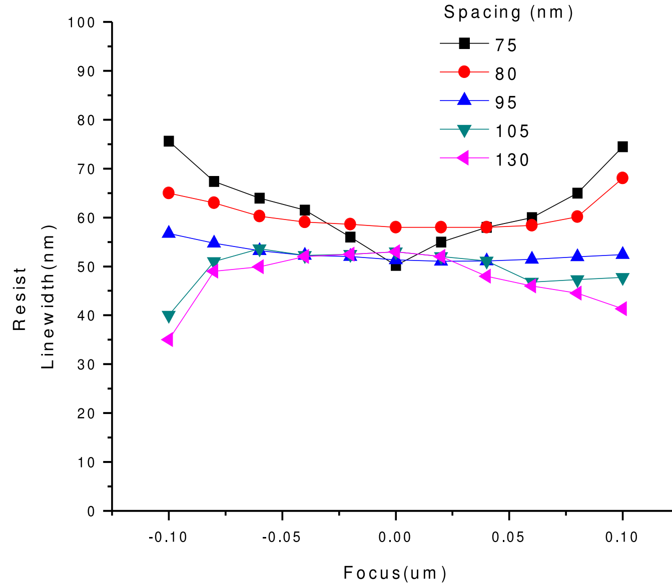

Figure 5.3. Bossung plot for dense Metal 1 lines in 45nm technology

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 5.3.

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 [9], linewidth variation due this source will always be a factor to consider Figure 5.4. Zernikes coefficients capture the deviation from ideal characteristics. The CD error due to lens aberration varies along both the horizontal and vertical direction. Zernikes 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

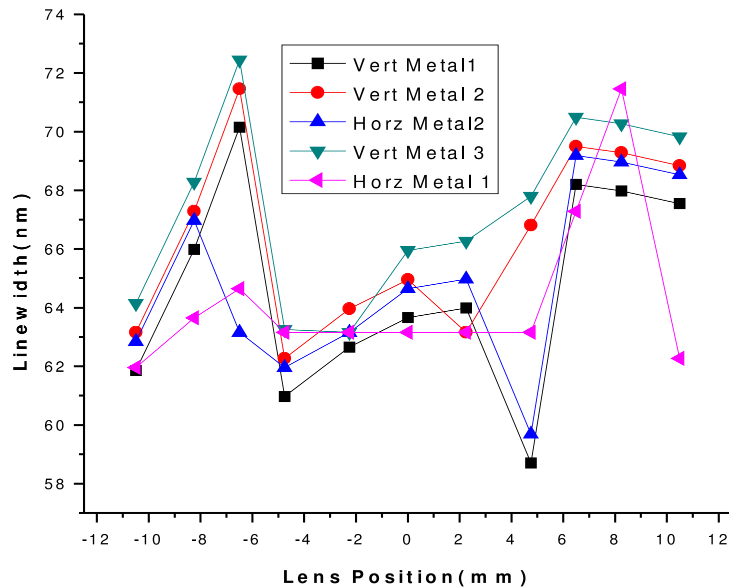

Figure 5.4. Linewidth variation due to lens aberration (using *Zernike's coefficinets* for metal lines in 45nm

input error is the resist thickness variation. Resist thickness variation affects all the

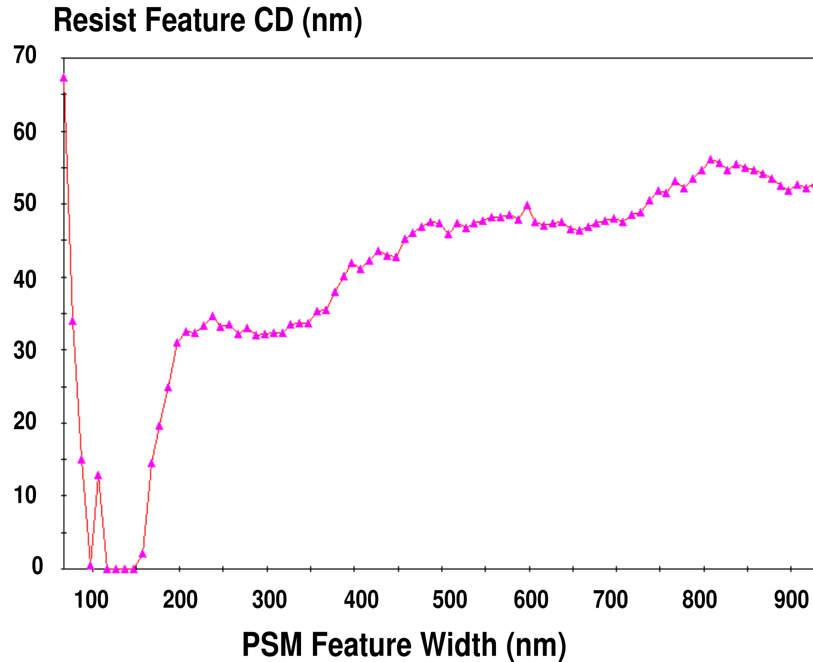

Figure 5.5. ACLV *Forbidden pitches*leading to yield loss

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 [45][4]. The above sources are dependent on the process and hence not in the hands of the designer. Topography dependent linewidth variations are due to pitch variations in the mask patterns for the design. Even when the process is accurate, the layout topology has its effects on the linewidth. Forbidden pitches are the best example for change in linewidth due to variation in spacing between mask features. Forbidden pitches were found at 130nm technology node by Robert Socha et.al [53] and are found to be much worse at 45nm technologies as shown in Figure 5.5.

5.4 Probabilistic Modeling & Yield estimation methodology

5.4.1 Methodology

Figure 5.6 shows our experimental methodology to estimate the yield due to various sources of variations in todays 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.

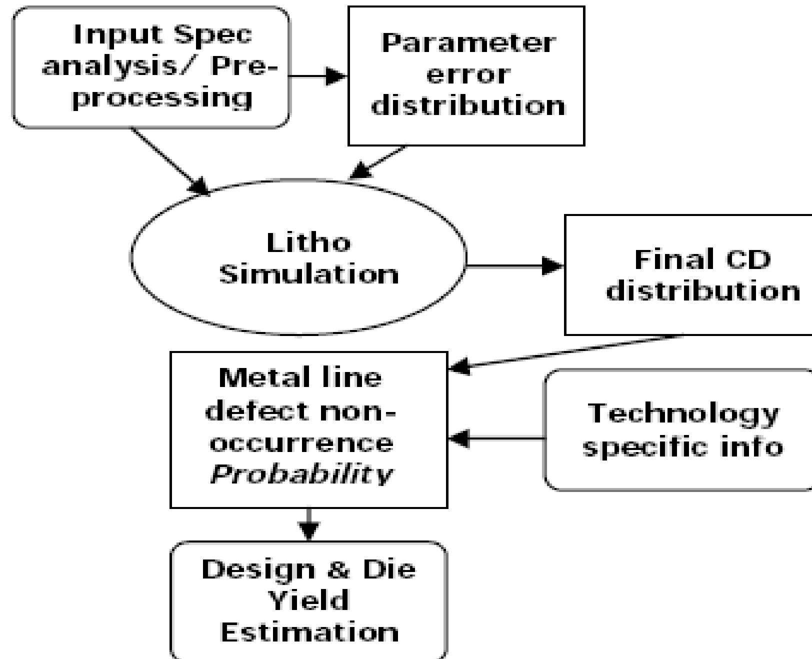

Figure 5.6. Variation-aware yield prediction methodology

Based on this preprocessing step a variation spread or error distribution of focus and exposure dose is chosen. This variation is also dependent on the metal layer of a particular technology. Lens aberration at different locations of the lens has been modeled by Zernike as coefficients of error. These capture the divergence of light from

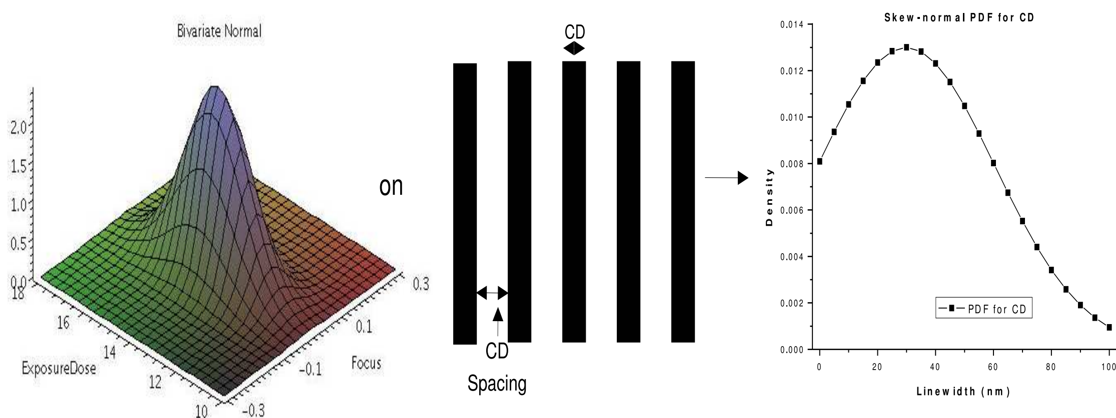

Figure 5.7. Yield prediction methodology overview

their ideal behavior and hence can be used as random variations in the simulation. The distribution of focus and dose variation is used in lithography simulation to obtain a distribution of CD. Lens aberration based on industry supplied Zernikes coefficients at multiple locations in the lens field are also used in the lithography simulation. CD distribution for all metal layers with different spacing is obtained. Figure 5.7 shows an example of the process. Once the CD distribution is obtained, CD-limited yield of the design can be estimated.

The analysis involves estimating the probability if a metal line with a particular spacing will not result in a open or a short. This probability can hence be used to predict the yield of that metal layer for the die.

5.4.2 Variation-aware CD-limited yield estimation technique

In order to better understand the effect of random manufacturing process variations on metal CD, an analysis may be done for two cases. One case is to consider that the process variation occurs only in one parameter. This is 1-D analysis and its simulation is simple as it involves only a few sampled intervals. The input variations of focus and exposure dose can be considered as being occurring independent

as single uniform distribution or together forming a bivariate normal distribution. A bivariate distribution as the input error distribution with lens aberration has more impact on the CD of the design (Figure 5.7, Figure 5.8). This analysis is more complicated but would mimic realistic variation of process parameters during lithography. Comprehensive lithography simulation or Monte Carlo based analysis can be used to deduce the effect of such variations in the lithographic process. It also well known that aerial imaging simulations on layout features consume high amounts of compute resources [61]. And hence both the above mentioned techniques used normally cannot be favored as todays layouts are highly dense and run up to 12 layers of metal. 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 technique reduces the number of simulations to obtain CD distribution, thus reducing the computational needs.

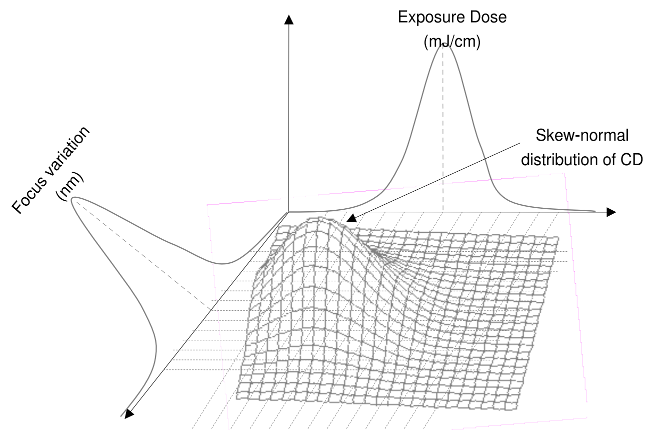

Figure 5.8. Stratified sampling on bivariate normal input error distribution & output skew-normal distribution

In our approach, we aim at dividing the input parameter distributions into different strata. The strata are regions over which the parameter variation is not more

than 0.25σ. A point is chosen from each strata and simulation is performed at these parametric values. We obtain a total of 24 simulation points from all the strata in the distribution. Lithography simulation is done at these simulation points to obtain the final CD distribution. We involve all the four sources of variation as mentioned on section 2 into our model. Since the effect of resist thickness variation will lead to a lesser than nominal CD, the final distributions are skewed as shown in Figure 5.12. Hence we fit the CD distribution to a skew-normal distribution to find the probability of not having an open or a short. To find this probability, the layout fault is modeled as follows.

5.4.2.1 Line Fault Modeling

As discussed previously, a line may shrink or expand in width due to variations in the lithography process. It is well known that the line etch process is not completely anisotropic, leading to random corrosion of a line along its edges. This is commonly referred to as Line Edge Roughness or LER. The amplitude of LER for a single edge can be found in the ITRS roadmap [23]. This phenomenon is used to model layout linewidth faults.

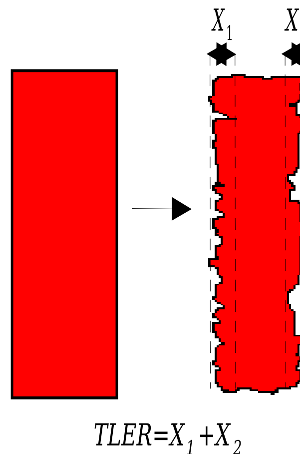

Figure 5.9. Line Edge Roughness defined

Consider the metal line shown in Figure 5.9. The post-litho linewidth is either smaller or larger than the expected CD as shown. Let the line edge roughness on the sides of the metal line be X1 and X2. The condition required for the shrinkage of linewidth leading to an open is given as follows,

$$CD_exp - CD_{post-litho} \leq TLER$$

$$TLER = X_1 + X_2$$

Similarly, the condition that the lithography process can lead to bridging of two adjacent lines at nominal spacing is give as,

$$CD_{ext1} + CD_{ext2} + TLER \geq CD_{spacing}$$

$$CD_{ext1,ext2} = CD_{post-litho} - CD_{exp}$$

CDexp is the expected CD for the lithography process, or otherwise termed as mask CD. CDpost-litho is the obtained CD after lithographic process with variations. CDexp is the expanded CD for two adjacent metal lines 1 and 2 as shown in Figure 5.10

5.4.2.2 Half-width based Yield Estimation

We can compute the probability of these events (line open and bridging) from the probability density function of CD obtained through the statistical process described in Section 5.4.2. The probability obtained here is for each metal layer and for different spacing values. This probability cannot be construed to be the probability of the entire metal line. The reason is that metal lines can have adjacent line running at different spacing on each of its sides. The impacts on of these adjacent lines and their probabilities will be different. The probability of the line as a whole will be a

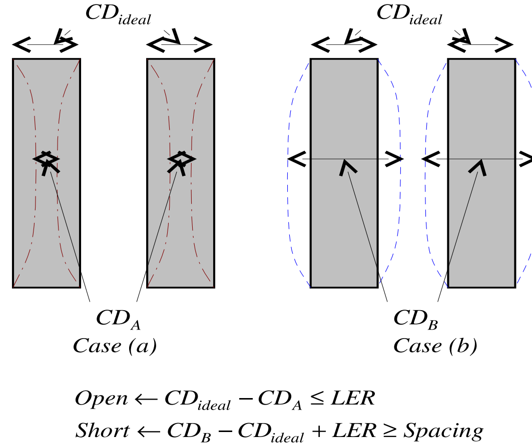

Figure 5.10. Line Fault Modeling - (a) Metal line OPEN; (b) Metal line SHORT

function of all the probabilities due to each of metal lines adjacent to it. In order take into account the different probabilities that a line can have, we perform a half-width based analysis.

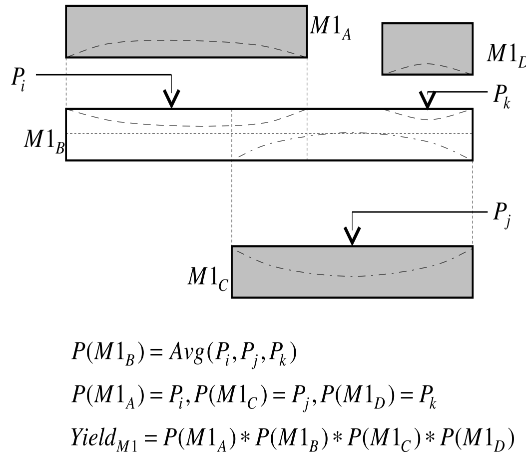

Figure 5.11. Half-width based yield estimation

For the half-width based analysis, each metal line is divided into two and the probability of not having an open or short due to the presence of adjacent metal lines is found. Figure 5.11 shows our scheme of estimating the yield of a metal line influenced by different adjacent line across its length. In this particular approach, we

use the average of the probabilities to simplify the yield estimation process. If a more accurate estimate is desired, a different function of the probabilities and the metal line length can also be used. A pseudo-code for our novel yield estimation technique is shown in Figure 5.12. Once the probability of failure for an individual line is obtained, it can be used in turn to compute design yield. $Designyield = \Pi p_i$ where p_i is the probability of yield of an individual line. pi in turn is simply 1- probability of failure of the line. Please note that this calculation does not consider the length of the line. It is implicit that a line is sufficiently long such that TLER from both edges can turn it into an open or short. Please note that this probability varies with line spacing and therefore related to actual layout. However, we run a simplified analysis of periodic lines indexed by various line-spacing to compute yield of a line indexed by its separation. Since the separation can be asymmetric in actual layout, we average probability of failure for each side.

YIELD (V_i)

{

foreach line (x_i ∈ Metal Layer (m))

Compute spacing s_i with adjacent metal line x_j

Calculate probability P_i of region which is f(s_i, V_i) based on half-width model

Total probability of line P_{line} is Avg(P_i, P_j, ...)

foreach Metal Layer (m_i ∈ design)

Metal yield = P_{l1} * P_{l2} * P_{l3} ...... * P_{ln} (l_1, l_2, l_3 ... l_n are lines in a metal layer)

Design Yield = Y_{M1} * Y_{M2} ..... * Y_{M12}

}

Figure 5.12. Pseudo-code for half-width yield estimation

Example: Let $p = p_i$, for all i, then $Yield = p^{noofmetallines}$ where, p is the probability of not having an open or short. Further, suppose there are a million such wires

of the same metal interconnect layer, then the yield is, $Yield = p^{10^6}$. Since yield is a multiplicative factor, each metal line should have a very high yield in order to have a good die yield. For instance, to have a yield of 92% from metal layer 4,

$$ln0.92 \leq 10^6 * lnp$$
$$lnp \geq \frac{ln0.92}{10^6}$$
$$p \geq e^{10^-6 * ln0.92} \approx 0.9999999916$$

Hence it can be seen that to obtain a very high yield, the probability of having an open or short should be very low.

5.5 Experimentation

5.5.1 Experimental setup

We used a commercial imaging simulator called $PROLITH^{TM}$ by KLA-Tencor for all our lithography simulation. The metal layer masks used for the simulation were binary masks (BIM). The masks were presumed to be Alternating Phase Shift Masks (Alt-PSM). The parameters used in the preprocessing step for a nominal process is listed in Table 5.1.

Table 5.1. PROLITH^{TM}simulation parameters

Paramter	Value	Mean	Std dev.
Depth of Focus(nm)	0.0	0.0	0.02
Exposure dose(mJ/cm^2)	14.5	14.5	0.462
Nominal Linewidth(nm)	70	70	-
Pitch(nm)	140	-	-
Wavelength (nm)	193	-	-
N.A	0.93	-	-
Resist Thickness(nm)	126	-	-
Refractive index	1.0	-	-

The test cases used are the widely available academic ISCAS85 benchmark circuits. The designs were synthesized using Synopsys Design Compiler vW-2004.12-SP3. Circuit layouts were generated using Cadence Composer v5141USR3 and Cadence Silicon Ensemble 5.3. Cadence SKILL language was used to perform layout parsing and obtain the number and length of metal lines at particular pitches.

5.5.2 Probabilistic yield estimation results

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. Obtained CD distributions for a metal layer at different spacing are shown in Figure 5.13. Please note that drawn length of CD was 70nm. It can be seen that the CD distributions at different spacing are skewed from their mean value of 70nm. ITRS reports that the

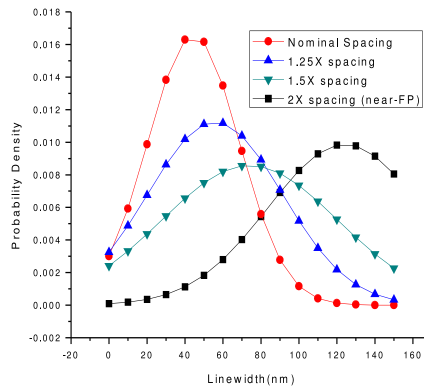

Figure 5.13. CD distribution for input parameter variation at different spacing

line edge roughness of a line is 5nm [23]. Then, for a reasonably long line of nominal width 10nm or below an open fault is very likely to occur (corresponding to erosion

from both edges of 5nm each). Similarly, suppose a line expands in width and if the expansion takes the line to reach 0.5*spacing TLER, then two adjacent lines could potentially bridge together. For the 45nm technology example above,nominal CD is 70nm, nominal spacing is 70nm. If linewidth becomes 70 (nominal width) + 0.5*70 (spacing) 5nm (TLER) = 100nm, then it is highly likely that this line will bridge with the adjacent line (assuming same effect) and result in an error.

Table 5.2 and Table 5.3 summarize the results of our experiment on different metal layers and for two technologies 65nm and 45nm. The probability of not having an open or short for each metal line kept at nominal spacing is listed in Table 5.2 . The scheme used to estimate this value was discussed in Section 5.4.2.1. It can be seen that the yield is greater for higher metal layers. The principal reason for such an increase is that fact that the proximity effects due to layout topography variations do not have any impact for these metal layers. As noted earlier, a tighter distribution of CD shows that it has less sensitivity to process parameter variation and leads to better yield.

Table 5.2. Probability at nominal spacing

	65nm (nm)		45nm (nm)			
Metal Layer	LW	SP	LW	SP	65nm lines	45nm lines
M1-M3	100	100	70	70	0.99835	0.9981
M4-M5	135	135	100	90	0.9968	0.9986
M6-M7	210	200	230	220	0.99954	0.99895
M8-M10	400	400	485	470	0.9999	0.9997

Table 5.3 shows the CD-limited yield for different ISCAS85 circuits. The yield is a joint function of the line spacing and variations. It is observed that the yield for comparable metal layers goes down as the circuit size increases. This follows an intuitive pattern. Table 5.3 shows both the nominal and worst case yields. The nominal case is when the litho-process variations are assumed to be within 10% off its mean value. This is the case when the process is within the FEM window. The

worst case shows how the yield differs when we consider the variation outside the FEM window which being 25% off the mean value.

Table 5.3. CD-limited design yield (Metal 1)

ISCAS 85 design	Nominal yield	Worst case yield
c17	0.992	0.772
c432	0.9658	0.698
c499	0.9236	0.682
c880	0.91	0.629
c1355	0.905	0.58
c1908	0.9012	0.574
c2670	0.90	0.5014
c3540	0.8945	0.4
c5315	0.87	0.358
c6288	0.8698	0.287
c7550	0.8358	0.2385

The yield values in Table 5.3 are lower than expected because the circuits were laid out according to design rules from 0.25m technology and then scaled for 45nm technology. Yield numbers will improve when actual 45nm layout rules are used. However, such rules are proprietary information.

5.6 Summary

A novel variation-aware CD-limited yield estimation technique based on statistical process modeling was proposed. Lithography simulation tools were used to estimate the CD distribution for multi-dimensional input parameter variations. Stratified sampling was used to reduce the number of samples to be simulated to obtain tail of the distribution. CD distribution was used to compute probability of failure of a line, which then was used to compute yield of a metal layer and finally the chip. This technique is simple, based on commercial litho simulation tools, yet effective.

CHAPTER 6

CONCLUSION

6.1 Summary

Scaling of transistor feature size over time has been facilitated by corresponding improvement in lithography technology. However, in recent times the wavelength of the optical light source used for photolithography has not scaled in the same rate as that of the minimum feature size of the transistor. In fact, starting with 180nm devices, the wavelength of optical source has remained the same (at 193nm) due to difficulties in finding a flicker-free, high energy, coherent light source with compatible improvement in lens material for focusing this light. Consequently, upcoming technology nodes (65nm, 45nm, 32nm and 22nm) will be using a light source with wavelength much greater than the feature size.

Process variations in the lithographic process have also added to the parametric variability of the design. Exotic tricks such optical proximity correction (OPC), Off-Axis Illumination (OAI), Phase-shifting Masks (PSM) and other resolution enhancement techniques (RET) have been used to mitigate the effects of topography based printability issues, but interconnect and gate shapes are still away from perfect rectangles on the resist. Interconnect varitions embedd themselves as defects on the die and gate linewidth variations impact circuit leakage and performance. Design yield also directly impacted due to non-linear effects caused by lithographic sensitivities.

Today's nanometer-scale design flows must comprehend with such increasingly critical yield and reliability concerns.

6.2 Suggestions for Future Work

Non-rectangular model suggested in chapter 4 attempts to model the static characteristics of the transistor i.e the ON and OFF currents. But the transistor transcient characteristics have not been implemented. As part of a future work the transcient behaviour of the device can be incorporated into the model. The important aspect to investigate would be whether the composite model obtained can fit the non-linear effects into a maximum of 3 transistors.

Another interesting approach is to investigate the effects of variational lithography on standard cell design. Standard cells can be designed for different variational lithographic process corners and static timing analysis can be done based on non-rectangular transistor models.

BIBLIOGRAPHY

[1] Abromovici, M., Breuer, M. A., and Fiedman, A. D., "Digital systems testing and testable design, "*IEEE Press, Piscataway, NJ, 1990*

[2] Artur Balasinki,"A Methodology to Analyze Circuit Impact of Process Related MOSFET Geometry,"*Proc.SPIE, Vol. 5378,2004*.

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

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

[5] Chau, R., Doyle, B., Doczy, B., Datta, S., Hareland, S., Jin, B., Kavalieros, J., and Metz, M., "Silicon Nano-Transistors and Breaking the 10nm Physical Gate Length Barrier,"*Intel Research Technical Report, IR-TR-2003-7, 2003*

[6] Choudhury, R., "Handbook of Microlithography and Micromachining "*Volume 1,In SPIE Press Monograph, 2002*.

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

[8] Croon J. A, G. Storms, S. Winkelmeier, I. Pollentier, M. Ercken, S. Decoutere, W. Sansen, and H. E. Maes, "Line edge roughness: characterization, modeling and impact on device behaviour,"*presented at Electronic Devices Meeting 2002*.

[9] Farrar, N., A. Smith, D. Busath, and D. Taitano, "In-Situ Measurement of Lens Aberrations,"*Proc. SPIE 4343 (2001), pp. 18-29*.

[10] Ferris-Prabhu, A. V., "A Cluster-Modfied Poison Model for Estimating Defect Density and Yield,"*IEEE Transactions on Semiconductor Manufacturing, Vol.3, May 1990*.

[11] Ferris-Prabhu, A. V., "Introduction to Semiconductor Device Yield Modeling,"*Artech House, 1992*.

[12] Ferguson, F. J., and Shen, J. P., "A CMOS fault extractor for inductive fault analysis,"*IEEE Trans. Computer Aided Design , Vol.CAD-7,pp. 1181-1194, Nov.1988*.

[13] Flagello, D. G., H. Laan, J. Schoot, I. Bouchoms, and B. Geha, "Understanding Systematic and Random CD variations using Predictive Modeling Techniques, "*Proc. SPIE 3679 (1999), pp. 162- 175.*

[14] Fukutome *et al*, "Direct evaluation of gate line edge roughness impact on extension profiles in sub-50nm N-MOSFETs,"*presented at Electronic Devices Meeting 2002.*

[15] Ganeshpure, K. P., A. Sanyal, S. Kundu, "A Pattern generation technique for Maximizing Power Supply Current,"*Proc. Intl. Conference on Computer Design, 2006.*

[16] Granik, Yuri; Cobb, Nicolas B.; Sahouria, Emile Y.; Toublan, Olivier; Capodieci, Luigi; Socha, Robert J.; "Subresolution process windows and yield estimation technique based on detailed full-chip CD simulation,"*Proc. SPIE, 4182, pp.335-341, 2000.*

[17] Guardiani, Carlo *et. al*, "An Effective DFM Strategy Requires Accurate Process and IP Pre-Characterization,"*Proc. Design Automation Conference, 2005.*

[18] Gupta, P., and Kahng, A. B., "Manufacturing-aware physical design,"*IEEE IC-CAD, pp.681-687, Nov.2003.*

[19] Henson, W. K. *et. al*, "Analysis of Leakage Currents and Impacts on Off-State Power Consumption for CMOS Technology in the 100-nm Regime,"*IEEE Trans. On Electron Devices, Vol.47, 2000.*

[20] Hopkins, H.H, "On the diffraction theory of optical images,"*Proc. R.Soc of London. A Math. Phys. Sci. vol.217(1130), pp408-432, 1953.*

[21] IBS Report

[22] Information available at: http://www.hpcwire.com/

[23] 2006 International Technology Roadmap for Semiconductors (ITRS), http://public.itrs.net

[24] Jacomet, M., "FANTESTIC: towards a powerful fault analysis and test pattern generator for integrated circuits,"*IEEE International Test Conference, pp. 633-642, 1989.*

[25] Jacomet, M., "Stuck-at faults vs. layout dependent CMOS faults," *IEEE Workshop Defect and Fault Tolerance in VLSI Systems, pp. 62-69, 1990.*

[26] Jacomet, M., and Guggenbuhl, W., "Layout-dependent fault analysis and test synthesis for CMOS circuits,"*IEEE Transactions on Comuter Aided Design and Systems, Vol.12, 1993.*

[27] James, Dick "2004- The Year of 90-nm: A review of 90nm Devices, "*Proc. IEEE Adv Semiconductor Manufacturing Conference, 2005.*

[28] Kim, Seong-Dong , H Wada, J.C.S Woo, "TCAD-based statistical analysis and modeling of gate line-edge roughness effect on nanoscale MOS transistor performance and scaling, "*Transactions on Semiconductor Manufacturing, vol. 17, pp.192, 2004*.

[29] Konuk, H., Ferguson, F. J., and Larrabee, T., "Charge-based fault simulation for CMOS network breaks,"*IEEE Transactions on Computer Aided Design and Systems, Vol.15, 1996*.

[30] Koren, I., and Stapper, C. H., "Yield Models for Defect Tolerant VLSI Circuits: A Review,"*Defect and Fault Tolerance in VLSI Systems, Vol.1, 1989*.

[31] Koren, I., Koren, Z., C. H. Stapper, "A Unified Negative Binomial Distribution for Yield Analysis of Defect Tolerant Circuits,"*IEEE Trans. On Computers, Vol.42, 1993*.

[32] Kruseman, Bram *et. al*, "Systematic Defects in Deep Sub-Micron Technologies,"*Proc. International Test Conference, pp. 290-299, 2004*.

[33] Kundu, Sandip; Aswin Sreedhar, Alodeep Sanyal, "Forbidden pitches in sub-wavelength lithography and their implications on design,"*J.Computer-Aided Materials Design, 2007, pp.79-89*.

[34] Mack, Chris A., "Understanding focus effects in submicron optical lithography, "*in Optical/Laser Microlithography, Proc. SPIE 922, 135-148 (1998); and Opt. Eng. 27(12), 1093-1100, 1998*

[35] Maly, W., Ferguson, F. J., and Shen, J. P., "Systematic characterization of physical defects for faults analysis of MOS IC cells, "*IEEE International Test Conference, pp. 309-399, 1984*

[36] Maly, W., *et. al*, "VLSI Yield Prediction and Estimation: A Unified Framework,"*IEEE Trans. Computer Aided Design, 1986*.

[37] Maly, W., "Modeling of Lithography Related Yield Losses for CAD of VLSI Circuits,"*IEEE Trans. On CAD, Vol. 4, 1985*.

[38] Maly, W., "A DRC-Based Algorithm for Extraction of Critical Areas for Opens in Large VLSI Circuits,"*IEEE Transactions on Computer Aided Design and Systems, 1999*.

[39] Mei, K. C. Y., "Bridging and stuck-at faults, "*IEEE Trans. Computers, Vol. C-23, pp. 720-727, July 1974*

[40] Moore, G. E., "Cramming More Components onto Integrated Circuits,"*Electron, 38, pp.114-117, 1965*.

[41] Pacha, C., *et. al*, "Impact of STI-Induced Stress, Inverse Narrow Width Effect, and Statistical Vth Variations on Leakage Current in 120nm CMOS,"*Solid-State Device Research conference, Set 2004, pp. 397-400*.

[42] Pillegi, L., Schmit, H., Strojwas, A. J., Gopalakrishnan, P., Kheterpal, V., Koorapaty, A., Patel, C., Rovner, V., Tong, K. Y., "Exploring regular fabrics to optimize the performance-cost trade-off," *Proc. DAC, 782-787, 2003.*

[43] Polishchik, I., N Mathur, C. Sandstrom, P. Manos, O. Pohland, "CMOS Vt-control improvement through implant lateral scatter elimination, " *ISSM, Sept 2005, pp. 193-196*

[44] Poppe, Wojtex. J. *et. al*, "From Poly Line to Transistor: Building BSIM Models for Non-Rectangular Transistors," *Proc. SPIE, Vol. 6156,2006.*

[45] Postnikov, S., et.al, "Monte carlo method for highly efficient and accurate statistical lithography simulation," *Proc. SPIE, Vol. 4691, pp. 1118-1126, 2002.*

[46] Puneet Gupta and Fook-Luen Heng, "Towards a Systematic-Variation Aware Timing Methodology," *Proc. Design Automation Conference, 2004.*

[47] Roy, J. A., *et. al*, "Capo: Robust and Scalable Open-Source Min-cut Floorplacer," *in Proceedings of International Conference on Physical Design, 2005.*

[48] Saab, D. G., Saab, Y. G., and Abraham, J. A., "CRIS: a test cultivation program for sequential VLSI circuits," *IEEE/ACM International Conference on Computer Aided Design, pp.216-219, 1992.*

[49] Sean X. Shi, Peng Yu, David Z.Pan, "A Unified Non-Rectangular Device and Circuit Simulation Model for Timing and Power," *Proc. Design Automation Conference, 2006.*

[50] Schellenberg, F. M., "Resolution enhancement technology The past, the present and extensions for the future," *in SPIE Microlithography Symposium, 2004.*

[51] Schellenberg, F. M., "A Little light magic, " *IEEE Spectrum, pp. 34-39, September 2003.*

[52] Socha, Robert *et al.*, "Understanding the Forbidden Pitch Phenomenon and Assist Feature Placement, " *Microlithography XVI, Proc. SPIE Vol.4689, 2000.*

[53] Socha, R., Dusa, M., Capodieci, L., Finders, Jo., Chen, F., Christopher J. Progler, Donis Flagello, Kevin Cummings, "Forbidden pitches for 130nm lithograph and below," *Optical Microlithography XIII*, *Proceedings of SPIE, Vol. 4000, pp. 1140-1155, 2000*,

[54] Souil, D.,*et. al*, "BSIM 4.1 DC Parameter Extraction on 50nm n-p MOSFETs," *Proc. IEEE Intl. Conference on Microelectronic Test Structures, Vol.15, April 2002.*

[55] Stamper, A. K., McDevitt, T. L., Luce, S. L., "Sub-0.25-micron interconnection scaling: damascene copper versus subtractive aluminium, " *Proc.Adv.Semiconductor Manufacturing Conference and Workshop, 1998*,

[56] Stapper, C. H., “Modeling of integrated circuit defect sensitivities, ”*IBM J. Research and Development, Vol.27, pp.549-557, Nov.1983.*

[57] Tong, K. Y., *et. al*, “Regular logic fabrics for a Via patterned gate array (VPGA),”*Proc.CICC, pp. 53-56, 2003.*

[58] Visweswariah, C., “Death, Taxes and Failing Chips,”*Proc. Design Automation Conference, 2005, pp343-347.*

[59] Wagner, I. A., and Koren, I., “An Interactive VLSI CAD Tool for Yield Estimation,”*IEEE Trans. Semiconductor Manufacturing, Vol.8,1995.*

[60] White, Dan, “A New Approach to Higher Yielding Solution,”*EE Times, China.*

[61] Wong, Alfred K., “Microlithography: Trends, challenges, solutions, and their impact on design ”*IEEE Micro vol23, pp.12-21, 2003.*

[62] Wong, A. K., *et al.*, “Forbidden area avoidance with spacing technique for layout optimization, ”*Proc. SPIE Vol.5379*

[63] Xiong, S. and J. Bokor, “A simulation study of gate line edge roughness effects on doping profiles of short-channel MOSFET devices,”*Transactions on Electron Devices, vol. 51, pp. 228, 2004.*