Electrical Design for Manufacturability Solutions: Fast Systematic Variation Analysis and Design Enhancement Techniques

by

Rami Fathy Amin Gomaa Salem

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2011

© Rami Fathy Amin Gomaa Salem 2011

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Since the scaling of physical dimensions has been faster than the optical wavelengths or equipment tolerances in the manufacturing line, process variability has significantly increased. This, in turn, has led to unreliable design, unpredictable manufacturing, and low yields. The results of these physical variations are circuit metric variations such as performance and power, known as the parametric yield.

Source of variations are either systematic (for example, metal dishing and lithographic proximity effects) or random (for example, material variations and dopant fluctuations). The former can be modeled and predicted, whereas the random variations are inherently unpredictable. There are several pattern-dependent process effects which are systematic in nature. These can be detected and compensated for during the physical design to aid manufacturability, and hence improve yield. This thesis focuses on ways to mitigate the impact of systematic variations on design and manufacturing by establishing a bidirectional link between the two. The motivation for doing so is to reduce design guardband and cost, and improve manufacturability and yield.

The primary objectives in this research are to develop computer-aided design (CAD) tools for Design for Manufacturability (DFM) solutions that enable designers to conduct more rapid and more accurate systematic variation analysis, with different design enhancement techniques. Four main CAD tools are developed throughout my thesis. The first CAD tool facilitates a quantitative study of the impact of systematic variations for different circuits' electrical and geometrical behavior. This is accomplished by automatically performing an extensive analysis of different process variations (lithography and stress) and their dependency on the design context. Such a tool helps to explore and evaluate the systematic variation impact on any type of design (analog or digital, very dense or less dense, large chips or small ones, and technology "A" versus technology "B").

Secondly, solutions in the industry focus on the "design and then fix philosophy", or "fix during design philosophy", whereas the next CAD tool involves the "fix before design philosophy". Here, the standard cell library is characterized in different design contexts, different resolution enhancement techniques, and different process conditions, generating a fully DFM-aware standard cell library using a newly developed methodology that dramatically reduce the required number of silicon simulations. Several experiments are conducted on 65nm and 45nm designs, and demonstrate more robust and manufacturable designs that can be implemented by using the DFM-aware standard cell library.

Thirdly, a novel electrical-aware hotspot detection solution, is developed by using a device parameter-based matching technique since the state-of-the-art hotspot detection

solutions are all geometrical based. This CAD tool proposes a new philosophy by detecting yield limiters, also known as hotspots, through the model parameters of the device, presented in the SPICE netlist. The device model parameters contain different abstracts of information such as the layout geometry, design context, and proximity effect on the process variability. This novel hotspot detection methodology is tested and delivers extraordinary fast and accurate results.

Finally, the existing DFM solutions that are focused on analyzing and detecting the electrical variations, mainly address the digital designs. Process variations play an increasingly important role in the success of analog circuits. Knowledge of the parameter variances and their contribution patterns is crucial for a successful design process. This information is valuable to find solutions for many problems in design, design automation, testing, and fault tolerance. The fourth CAD solution, proposed in this thesis, introduces a variability-aware DFM solution that detects, analyze, and automatically correct hotspots for analog circuits.

This thesis presents high performance and novel design-aware process variation classification and detection solutions along with automated design enhancement and correction techniques.

Acknowledgements

I would like to give my sincere appreciation to my academic and professional ancestry Prof. Mohab Anis for his assistance, support, vision, and strong encouragement throughout my research. His insights and leadership in the semiconductor manufacturing academia have led me to the interesting research topics. I would also like to thank Prof. David Nairn for accepting to be my co-supervisor and providing the tremendous support in facilitating my PhD. In addition to his valuable technical feedback on my research.

I would like to thank Professor Azadeh Davoodi for accepting to be the external examiner in my PhD defense and I appreciate her time, and effort for traveling and her valuable comments on this research . Also I would like to thank Professor Karim Karim, Professor Kumaraswamy Ponnambalam and Professor Selvakumar Chettypalayam for serving on my PhD defense committee and their valuable comments on this research.

I am very appreciative of the support from all of my colleagues at Mentor Graphics. In particular, I would like to express my gratitude to Haitham Eissa, Mohamed Al-Imam, Ahmed Arafa, Hesham Abdulghany, Tamer AbdelRahim, Sherif Hammouda, Nader Hindawy, Sherif Hany, and Abdel Rahman El-Mously for their technical leadership and support.

My stay at University of Waterloo was stimulating and entertaining, thanks to the friendship of Hassan Hassan, Mohamed AbuRahma, Ahmed Youssef, Muhammad Nummer, Aymen Ismail, Noman Hai, Hazem Shehata and many other friends.

All of this was only made possible by the encouragement and love of my family. My deepest gratitude goes to my mother and father for their never ending support, and for remembering me in their prayers. No words of appreciation could ever reward them for all they have done for me. I am, and will ever be, indebted to them for all the achievements in my life. I am also thankful to my siblings Tamer,and Sally for their endless support. My loving wife, Marwa, shared with me every moment of my Ph.D. She supported me with her unconditional love and care during all the ups and downs of my research, and she was always there to motivate me. I thank her for everything she have done.

Finally, I express my gratitude to Allah (God) for providing me the blessings and strength to complete this work.

Rami Fathy Amin Gomaa Salem
July, 2011

Dedication

*To my **Mother** and **Father**, for their invaluable prayers. To my wife **Marwa**, for her boundless love that inspires me to become a better man, for supporting my completion of this thesis, for being my emotional rock in times of doubt. To my beloved children **Zeina** and **Marwan***

Contents

List of Tables xi

List of Figures xii

1 Introduction and Motivation 1

1.1 Introduction . . . 1
1.2 Motivation: DFM Solutions for Parametric Yield . . . 2
1.3 Thesis Outline . . . 3
1.4 Thesis Contributions . . . 4

2 Background: Process Variations and DFM Solutions 7

2.1 Introduction . . . 7
2.2 Sources of Process Variations . . . 8
2.2.1 Sources of Systematic Variations: Optical Lithography . . . 9
2.2.2 Sources of Systematic Variations: Chemical Mechanical Polishing . 14
2.2.3 Sources of Systematic Variations: Mechanical Stress . . . 16
2.2.4 Sources of Systematic Variations: Material Dopping . . . 17
2.3 Yield Issues Due to Process Variations . . . 18
2.3.1 Impact of Process Variations on Parametric yield . . . 19
2.4 Design for Manufacturing . . . 26
2.4.1 Traditional DFM Methods . . . 26

2.4.2 Taxonomy 28
2.5 Summary 36

3 A Methodology for Analyzing Process Variation Dependency on Design Context and the Impact on Circuit Analysis 37

3.1 Introduction 37
3.2 Background 39
3.2.1 Corners Simulation 39
3.2.2 Monte Carlo Analysis 40
3.2.3 Response Surface Modeling 40
3.2.4 Issues with the Previous Process Analysis Tools 41
3.3 Design Context Impact on Process Variations 42
3.3.1 Design Context Impact on Lithography Variations 42
3.3.2 Design Context Impact on Stress Variations 43
3.4 Characterization Flow: Layout Design Context Impact on Lithography and Stress Variations 44
3.4.1 Context Generator Engine 44
3.5 Experiments and Results 48
3.5.1 Layout Design Context Effect on Standard Cells' Geometrical Level 48
3.5.2 Layout Design Context Effect on Standard Cells' Electrical Level . 49
3.5.3 Systematic Variations Impact on Ring Oscillator designed at 65nm 55
3.6 Summary 59

4 Fix Before Design CAD Solution: DFM-aware Standard Cell Re-characterization 60

4.1 Introduction 60
4.2 Background 61
4.3 DFM-aware standard cell re-characterization flow 62
4.3.1 Isolated Process Point Engine (IPPE) 64
4.4 Experiments and Results 72

4.4.1 Standard cell library re-characterization 72
4.4.2 DFM-aware Finite Impulse Response (FIR) filter designed at 45nm 76
4.5 Summary 80

5 High Performance Electrical Driven Hotspot Detection CAD Solution for Full Chip Design using a Novel Device Parameter Matching Technique 82
5.1 Introduction 82
5.2 Background 83
5.2.1 Rule-based Hotspot Detection 83
5.2.2 Model/Simulation-based Hotspot Detection 84
5.2.3 Pattern Matching-based Hotspot Detection 84
5.2.4 Electrical-driven Hotspot Detection 85
5.3 Proposed Flow Overview: Design and Electrical Driven Hotspot Detection solution 87
5.3.1 Smart Device Parameters Matching Engine 91
5.3.2 Intent Driven Design Engine 94
5.4 Experiments and Results 96
5.4.1 Accuracy of the proposed solution 96
5.4.2 Runtime Analysis 100
5.5 Summary 101

6 A Parametric DFM CAD Solution for Analog Circuits: Electrical Driven Hot Spot Detection, Analysis and Correction Flow 102
6.1 Introduction 102
6.2 Background 103
6.3 Proposed Flow Overview: Electrical Hot Spot Detection, Analysis and Correction Flow 105
6.3.1 Intent Driven Design Engine 109
6.3.2 E-hotspot Detection Engine 110

6.3.3 Sensitivity Analysis Algorithm 113
6.3.4 E-Hotspot Correction Engine 116
6.4 Experiments and Results 120
6.4.1 Silicon Wafer Full Speed Transmitter chip designed at 130nm Technology 120
6.4.2 Voltage Control Oscillator designed at 45nm Technology 125
6.4.3 OPAMP designed at 65nm Technology 128
6.5 Summary 129

7 Summary and Conclusions 130
7.1 Conclusion 130
7.2 Summary of Contributions 131
7.3 Future Research Directions 132

A Appendix: Pending Patents from this Work 133

References 146

List of Tables

3.1 Example of the Electrical Simulation Results for NAND gate 52

4.1 DFM-properties for different Standard cells 76

4.2 Timing reports simulating standard netlist versus process-aware netlists . 78

5.1 Grouping the devices as an output from the Device Parameter Matching Engine- Group 1: contains 4 matched transistors 98

5.2 Grouping the devices as an output from the Device Parameter Matching Engine- Group 2: contains 8 matched transistors 98

5.3 Top five device groups experiencing variation in the DC current values . . 98

5.4 Result comparison between previous hotspot detection methods and our method 100

List of Figures

1.1 Defect categories by yield and process nodes [1] 2
1.2 Systematic Variation Aware Analysis and Design Enhancement Techniques 5

2.1 Schematic of a step-and-scan wafer stepper. [2] 11
2.2 Features on the mask cannot be exactly reconstructed due to diffraction. [2] 13
2.3 Equipment used for CMP. [3] 15
2.4 Dishing and Erosion. [4] 15
2.5 A depiction of the STI stress effect.[5] 16
2.6 A depiction of the WPE.[5] 17
2.7 The limits of lithography techniques and the changes in printed results across process nodes [6] 20
2.8 Measured Vth versus channel length L for a 90nm which shows strong short channel effects causing sharp roll-off for Vth for shorter L.[4] 21
2.9 Ion and Ioff variations due to transistor length [7] 22
2.10 Interconnect resistance variation due to width. [8] 24
2.11 MOSFET device geometry using a shallow trench isolation scheme. [9] . . 24
2.12 Vth versus well-edge distance for 3.3V nMOS device.[5] 25
2.13 Systematic Variation Aware Analysis and Design Enhancement 30
2.14 Flow diagram of a lithography model [10] 33

3.1 Experimental data for pitch curve with feature type classification. [11] . . . 42
3.2 Stress effect on Ion for simple layout implementations.[12] 43

3.3 Flow used for Extracting Context-aware metrics 45
3.4 Pseudo Code for the Context Generator Engine 46
3.5 Different contexts surrounding an example of cell under test 47
3.6 Gate length variation in different design contexts 49
3.7 Standard cell with (a) nominal device parameters and wire parasitics exhibiting nominal delay, td and slew, ts and (b) modified device parameters and wire parasitics exhibiting modified delay, td1 and slew, ts1, due to manufacturing variations.[13] 50
3.8 Standard cell with (a) nominal device parameters drawing leakage current, Ioff1 and (b) modified device considering manufacturing variations drawing leakage current, Ioff2. [13] 51
3.9 Low-to-high propagation delay (Tplh) 53
3.10 High-to-low propagation delay (Tphl) 53
3.11 Inverter layout (Left) and flip flop layout (Right) without context 54
3.12 Inverter layout (Left) and flip flop layout (Right) within context 54
3.13 Schematic of a ring oscillator with enable pin: (i) Logic level. (ii) Transistor level. [14] 56
3.14 Frequency response curve for a 90nm five stage NAND gate ring oscillator. 56
3.15 Frequency and Power analysis for the 65nm ring oscillator using drawn device parameters vs. process-aware device parameters. 57
3.16 Frequency Response Curve for Poly Lithography Effects. 58
3.17 Gate Delay Analysis based on the effects of dense vs. isolated gate placement. 58

4.1 Proposed Flow for Implementing DFM-Aware Standard Cell Library . . . 63
4.2 Pseudo Code for the DFM-Aware Standard Cell Library Engine 64
4.3 Pseudo Code for the IPPE engine 66
4.4 Different cases for contour comparison results using the big window knob operation 66
4.5 Micro window scanning square 68
4.6 Flow for comparison of similarity between two process window conditions . 68

4.7 Inputs and outputs to and from IPPE engine 69
4.8 Left: Matrix of input conditions. Right: Selected process conditions shaded in grey 70
4.9 Left: Test polygon. Middle: PV bands of input matrix. Right: PV bands of output reduced matrix 71
4.10 Comaprison between bands from process window corners and reduced conditions to the bands obtained from the input matrix 71
4.11 EPE histogram for different RET aggressiveness 73
4.12 Design Sensitivity for AOI cell under different RET aggressiveness 75
4.13 Process Sensitivity for AOI cell under different RET aggressiveness 75
4.14 FIR Layout after Placement and Routing 77
4.15 Generating 1- Default 2- Litho-only 3-Stress-only 4-Litho and Stress 5- Real Design Context -aware critical path netlists 79
4.16 Generating the real design context critical path netlist 79

5.1 From Layout to Spice Instance Parameters. [12] 86
5.2 Proposed Design Context Aware and Electrical Driven DFM Solution . . . 88
5.3 Stress Effect related SPICE Parameters. [9] 90
5.4 Pseudo Code for the Device Parameters Matching Algorithm 92
5.5 Device Parameters Matching Flow 93
5.6 Intent Driven Design Flow 94
5.7 Pseudo Code for the Intent Driven Design Engine 97
5.8 FIR layout highlighting the potential electrical hotspots on the critical paths 99

6.1 Layout proximity effects on multi-finger devices can not be predicted in schematics phase [15] 104
6.2 Proposed Flow Overview: Complete Electrical DFM Solution 107
6.3 Intent Driven Design Engine 110
6.4 Ids vs. Distance from STI boundary Typical 0.13um process [16] 111
6.5 Incremental change in diffusion edges 116

6.6 E-hotspot Correction Engine 118
6.7 Pseudo Code for the E-hotspot Correction Engine 119
6.8 FullSpeed Transmitter testbench including supply/ground/DP/DM bonding and cable model. Cable end loaded with 50pF as for Full Speed specs. . . . 120
6.9 Lab Measurement: Full Speed Signals after 5m cable. DP(Yellow) DM (Blue) at end of cable. Also DP(Violet) before cable. DM is totally distorted by end of cable. 121
6.10 Simulation results without stress effects: a- Levelshifter DATAP/DATAM outputs b- Far end DP/DM 122
6.11 Simulation results with stress effects: a- Levelshifter DATAP/DATAM outputs b- Far end DP/DM 123
6.12 VCO schematic design 125
6.13 VCO layout after correction 126
6.14 The VCO frequency response chart for different cases: a-No layout context effects b-Before e-DFM correction c-After e-DFM correction 127
6.15 The VCO output swing chart after e-DFM correction 127
6.16 Histogram for worst case changes in saturation current I_{dsat} and pointing out most affected transistors in the tested layout design. [15] 128
6.17 Simulation results for the Opamp analog circuit showing curve close to actual curve after fixing stress effects. [15] 129

Chapter 1

Introduction and Motivation

1.1 Introduction

CMOS device scaling has outpaced advancements in manufacturing technologies. Thus, process variability compared to the feature size, continues to increase. Many designers have started to recognize that this push into advanced nodes has exposed a hitherto insignificant set of yield problems [1] [17]. Physical yield problems are known to cause catastrophic failures, such as bridging faults, have always been the primary focus of verification efforts. At sub-90nm technologies, integrated circuit (IC) designs are also exhibiting parametric yield failures [18]. Parametric yield issues arise when the process variations have not been sufficiently characterized such that a circuit might have achieved design closure through standard methodologies, but the silicon performance does not match the simulation results.

At sub-90nm technologies, yields drastically drop-off, as designs fail to consistently achieve their technical and competitive objectives [19]. As more and more features are placed into smaller and smaller spaces, the unintended effects of this crowding are creating havoc with yield and performance. Designers need to be aware that device behavior depends not only on traditional geometric parameters such as channel length and width, but also on layout implementation details of the device and its surrounding neighborhood. Figure 1.1 illustrates that within the category of systematic yield defects, there are two subcategories: 1- lithography/new materials-based and 2- design-based [1]. At the 130 nm node, the systematic defects are roughly, an equal split between the two subcategories. A dramatic increase of defects is obvious at the 90nm node, where the ratio is approximately 2.5:1 for design-based versus lithography/new materials yield defects [20].

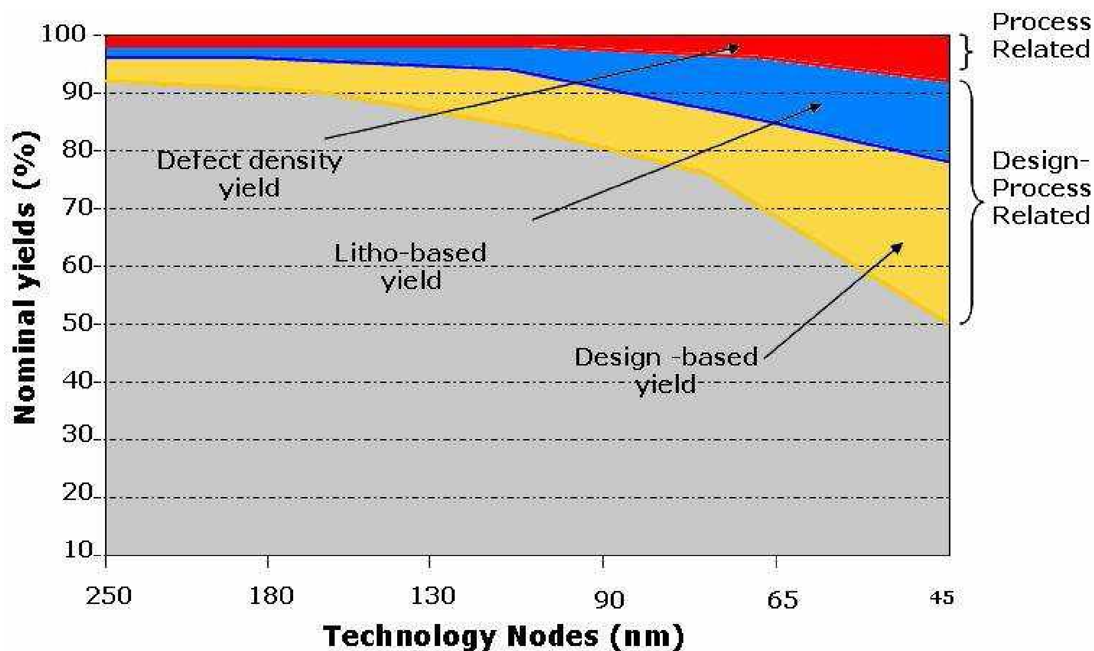


Figure 1.1: Defect categories by yield and process nodes [1]

1.2 Motivation: DFM Solutions for Parametric Yield

Physical and parametric yield failures have similar negative business implications, in that new circuit designs will fail to meet performance expectations. To combat yield failures, the semiconductor industry has started to deploy new tools and methodologies, commonly referred to as design for manufacturing (DFM) [21].

Traditionally, design and manufacturing have been conveniently kept separated, with only minimal information exchange. From the manufacturing side, SPICE models, technology file, and design rules are supplied to estimate the performance and power, and to establish manufacturing limitations. However, in today's era of such a large process variability, traditional corner-based analyses can be overly pessimistic, causing valuable performance to be left on the table. Design rules have also become extremely complex, substantially reducing productivity. From the design aspect, the layout is transferred to manufacturing as a set of shapes to be printed on silicon. To achieve the high fidelity of silicon shapes in the "drawn" shapes, in the manufacturing process applies several resolution enhancement techniques (RETs) are applied to the entire design. Unfortunately, RETs significantly increase the mask writing cost and multi-million dollar mask sets are now

common. To reduce the mask cost, it is crucial to additionally convey the design intent to manufacturing so that high fidelity is attempted only for selected features in the design that require accurate manufacturing. DFM techniques essentially address the questions related to the exchange of information across design and manufacturing, and the use of this information for yield enhancement.

Most efforts have been concentrated on catastrophic failures, or physical DFM problems. Recently, there has been an increased emphasis on parametric yield issues. A comprehensive DFM methodology reduces the parametric yield loss and leads to maximized utilization of the process. More sophisticated parametric-based DFM solutions are the ones that can perform different tasks:

- **Predict**: Designers can improve the parametric yield and chip performance by accurately determining the impact of systematic variations during the design stages.
- **Analyze**: Designers can quickly analyze various types of manufacturing non-idealities.
- **Enhance**: Designers can reduce sensitivity to manufacturing variations to achieve the desired predictability.

The goal of this thesis is on developing parametric-based DFM computer-aided design (CAD) tools for systematic variation analysis and design enhancement techniques.

1.3 Thesis Outline

In the next chapter, the relevant literature is reviewed to demonstrate the need to propose new ways to augment the yield.

Four CAD tools are proposed and developed in this thesis, as denoted in Figure 1.2. The first CAD tool, presented in Chapter 3, to facilitate a quantitative analysis of the systematic variation impact on different circuits in terms of electrical and geometrical behavior. Extensive lithography and stress analysis is performed under different design context.

In Chapter 4, a second CAD tool is proposed to manifest the "fix before design philosophy", avoiding long design cycles and re-spins. This is implemented by re-characterizing the standard cell library comprising different design context, different resolution enhancement techniques and different process conditions using a newly developed methodology that

dramatically reduces the required number of silicon simulations. Experiments prove that the DFM-aware standard cell library results in more robust and manufacturable designs.

Chapter 5 presents a novel electrical-aware hotspot detection solution by using a device parameter-based matching technique. While the state-of-the-art hotspot detection solutions are all geometrical based, this hotspot detection methodology experimentally demonstrates extraordinary fast and accurate results.

The existing DFM solutions that are focused on analyzing and detecting the electrical variations, principally address digital designs. On the other hand, process variations play an increasingly important role in the success of analog circuits. The fourth CAD solution, proposed in this thesis and presented in Chapter 6, introduces a variability-aware DFM solution that detects and automatically corrects hotspots for analog circuits.

Finally the contributions, conclusion and suggestions for future work are summarized in Chapter 7.

1.4 Thesis Contributions

- Implement four CAD solutions that tackle parametric and physical yield issues. These solutions enable the designers to achieve faster and accurate systematic variation analyses. In addition, designers are provided with different design enhancement and correction techniques.

- Novel CAD engines are introduced to solve key challenges that present state-of-the-art DFM solutions can not achieve. For example, systematic variation analysis flows for full chip digital designs face several challenges, mainly the high computational silicon simulation time. To address this issue, a novel solution is introduced to highly reduce the lithography simulation runtime (nearly 90%) compared to running lithography simulation using the complete set of different process conditions, with a minimum impact on accuracy. Another proposed engine uses SPICE parameter matching techniques for hotspot detection, avoid the need for silicon simulations and can identify the hotspots in matter of seconds. Another challenge with the current DFM solutions is the lack of design awareness. Most, if not all, hotspot detection solutions generate thousands of hotspots which challenges designers as to which hotspots are critical from the design point of view. Therefore a design intent driven engine that adds design awareness to the hotspot detection engines is introduced. It is worth mentioning that this engine is currently implemented within industrial CAD tools: Calibre

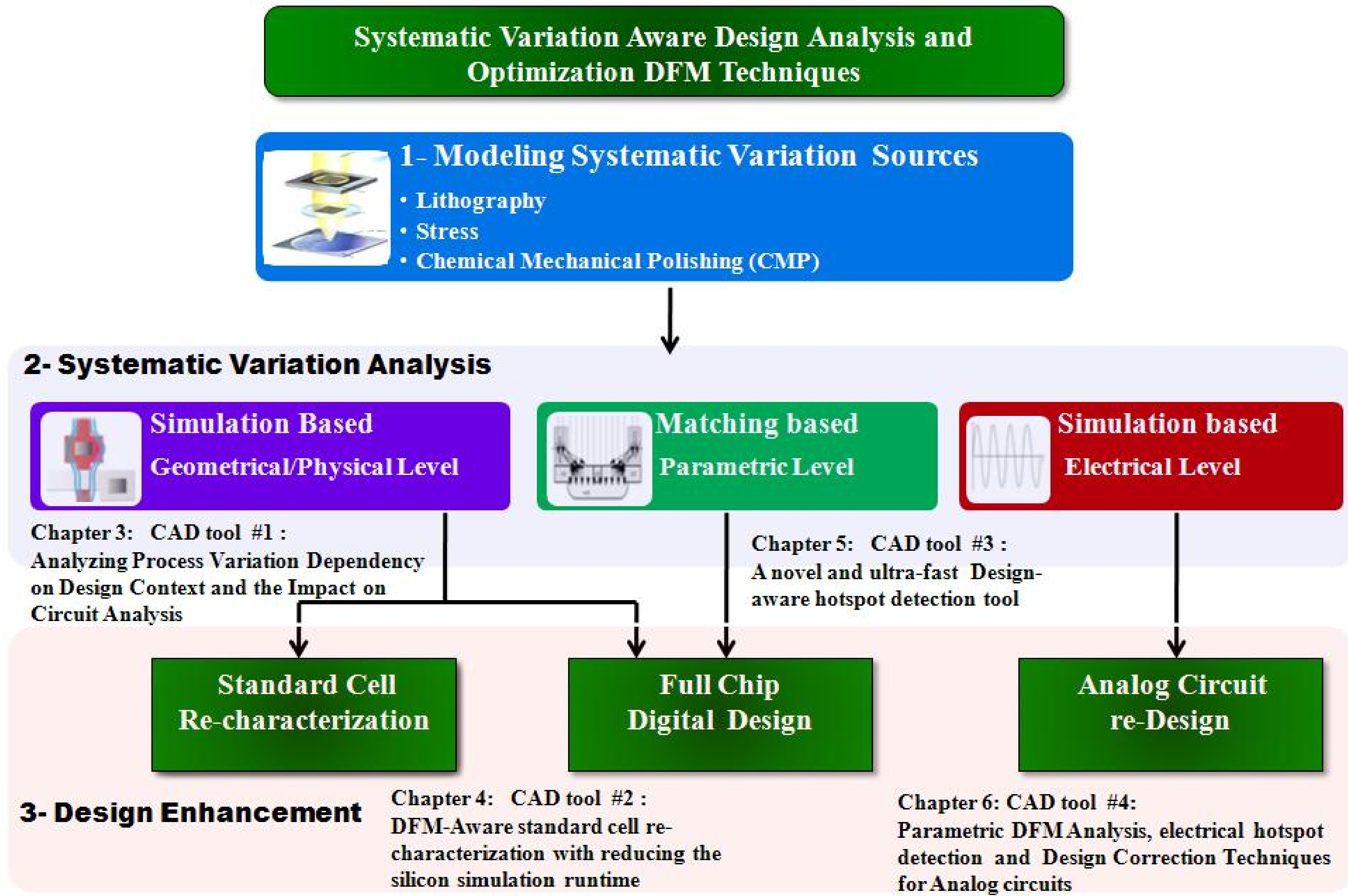


Figure 1.2: Systematic Variation Aware Analysis and Design Enhancement Techniques

PERC LDL Mentor Graphics, and tested in the industry at **ON-Semiconductor Corp.** design team.

- New philosophies are adopted in these newly designed DFM CAD solutions. For example, the current DFM solutions are mostly geometrical-based where a process variation analysis is run on the physical layout level. Also introduced is an electrical driven hotspot detection solution that stems from a SPICE parameter matching technique. Beside that solution, there is another CAD solution that supports the "fixing before design" philosophy which provides a more robust and manufacture-able physical layout designs that will reduce design re-spins and also reduce the number of hotspots.

- The proposed CAD solutions target different design types. One of the proposed CAD solutions is considered one of the first complete DFM solutions (Detection-Analysis-Correction), which is dedicated to solve parametric yield issues for analog circuits.

- These CAD solutions are technology independent and tested on industrial testcases designed in 65nm and 45nm technologies. Silicon-tested circuits are also used to verify these solutions.

Chapter 2

Background: Process Variations and DFM Solutions

2.1 Introduction

Process variation has long been a concern in integrated circuit design, manufacture, and operation. In recent years, with continued scaling and demand for higher performance and higher yield, the need and interest in techniques and tools that address variation has increased. The impact of process variations on circuit power and performance is exacerbated by the superlinear dependence of several electrical metrics on feature size (e.g., subthreshold leakage on gate length, and gate tunneling leakage on gate oxide thickness) [22]. Power, and especially leakage power, is another major challenge faced by designers today. Lowering of supply voltage to reduce dynamic power necessitates lowering of threshold voltage to sustain high-performance and adequate noise margins. Unfortunately, lowering threshold voltage causes a near-exponential increase in leakage power, and a larger ratio of static ("wasted") power to total power. Leakage variability, which is increasingly a determinant of parametric yield, is another important problem that must be addressed for continued CMOS scaling.

Traditionally, design and manufacturing have been conveniently kept separated, with only minimal information exchange. From the manufacturing side, SPICE models, technology file, and design rules are supplied for performance and power estimation, and to convey manufacturing limitations. However, in today's era of large process variability, traditional corner-based analyses can be overly pessimistic, causing valuable performance to be left on the table. Design rules also become extremely complex, substantially reducing

productivity. From the design side, the layout is transferred to manufacturing as a set of shapes to be printed on silicon. To achieve high fidelity of silicon shapes to "drawn" shapes, the manufacturing side applies several resolution enhancement techniques (RETs) to the entire design. Unfortunately, RETs significantly increase the mask writing cost and multi-million dollar mask sets are now common. To reduce mask cost, it is important to additionally convey the design intent to manufacturing so that high fidelity is attempted only for selected features in the design that require accurate manufacturing. Design for manufacturing (DFM) techniques essentially address the questions related to the exchange of information across design and manufacturing, and the use of this information for yield enhancement.

To better understand DFM, we first present a brief overview of major sources of process variations.

2.2 Sources of Process Variations

Variation is the deviation from intended values for structure or a parameter of concern [4]. The electrical performance of modern IC is subject to different sources of variations that affect both the device (transistor) and the interconnects. For the purposes of circuit design, the sources of variation can broadly be categorized into two classes [23], [24], [25]:

1. Inter-Die variations: also called global variations, are variations from die to die, wafer to wafer, and lot to lot. These variations affect all devices on the same chip in the same way (e.g., they may cause all the transistors' gate lengths' to be larger than a nominal value).
2. Within-Die (WID): also called local or intra-die variations, correspond to variability within a single chip, and may affect different devices differently on the same chip. In nanoscale regime, dimensions are small enough that device behavior is largely dependent of the neighborhood of the device. (e.g., some devices on the same die may have larger channel length L than the rest of the devices).

Inter-Die variations have been a longstanding design issue, and are typically dealt with using corner models [23]. These corners are chosen to account for the circuit behavior under worst-case variation and were considered efficient in older technologies where the major sources of variation were inter-die variations. However, in nanometer technologies, WID variations have become significant and can no longer be ignored [26], [27], [28]. As a result,

process corners based design methodologies, where verification is performed at a small number of design corners, are currently insufficient. WID variations can be subdivided into two classes [23], [24], [25]:

- Random variations: As the name implies, are sources that show random behavior, and can be characterized using their statistical distribution. Random variations are inherent fluctuations in process parameters, such as random dopant fluctuations from die-to-die, wafer-to-wafer and lot-to-lot.
- Systematic variations: These types of variations depend on the layout pattern and are therefore predictable. Systematic variations are highly dependent on the physical layout context (i.e. the systematic variations are dependent on the design). Examples of systematic variations are topography variations due to Chemical Mechanical Polishing (CMP) dishing, linewidth variation due to lithography defocus and exposure, and stress due to shallow trench isolation (STI). These effects have a substantial but deterministic impact on the critical dimension (CD) of a transistor gate or the width and thickness of an interconnect wire.

Historically, global random variations have been higher up the agenda than local systematic variations. But below 90 nm, local systematic variations have become the foremost concern [2]. The primary sources of manufacturing systematic variation will be discussed in more details.

2.2.1 Sources of Systematic Variations: Optical Lithography

Optical lithography, or simply lithography, is the mainstream technique to create patterns on silicon wafers. While conceptually simple, lithography has evolved into a highly sophisticated process due to precision requirements that are unmatched anywhere in modern manufacturing. Lithography involves several steps which can be simplistically grouped into photoresist deposition, exposure, and etching. The process begins with deposition of a thin layer that is intended to be patterned on the wafer. The thin layer is sacrificial and is used to selectively etch, dope, oxidize or deposit the underlying material. The pattern on the mask is first transferred to the photoresist that is deposited over the thin layer. An etchant is then used to remove the thin layer from where it is not protected by the photoresist. We now briefly describe the major lithography steps, further details of which can be found in [29].

Photoresist and its Deposition

Photoresists are materials that when exposed to light undergo a photochemical reaction that changes their solubility properties to a developer chemical. Positive photoresists become soluble in the regions that are exposed to light while negative photoresists become soluble in the regions occluded from light. Prior to deposition of the photoresist, the wafer may optionally be treated with a chemical that promotes adhesion between the thin layer and the photoresist.

The standard method of depositing the photoresist onto the wafer is resist spinning. In this method, a small amount of the photoresist in liquid form is dispensed onto the center of the wafer, and the wafer is then rotated about its center at a high rate. As the wafer spins, the resist spreads radially and solidifies into a uniform solid layer over the wafer. A baking step, known as soft bake, in which the wafer is heated to relatively low temperature for a short period of time, is then optionally performed to further densify the photoresist. Another optional step of coating the wafer with an anti-reflective coating (ARC) is then performed to suppress the light reflections in the succeeding exposure steps.

Exposure

By selectively exposing the photoresist to light, a pattern can be transferred to the photoresist. This process is accomplished in lithography by imaging of the mask to transfer patterns on it to the photoresist. The mask is a thin piece of a high-quality transparent material, typically quartz, partially covered with an opaque material, typically chromium, that has been removed according to the circuit pattern using an electron-beam mask writer.

Over the years, mask writing technology has improved but has failed to keep pace with the shrinkage of feature sizes. Thus, projection printing, in which projection optics (sometimes simply known as the lens) are used to reduce the mask image by a reduction factor (N), is now mainstream. The projection optics are typically an array of high-quality lenses cascaded to realize the reduction factor with minimal image distortion. The reduction factor in modern optics is most commonly equal to four or five. Larger reduction factors relax the precision requirements on the mask and reduce the linewidth variations due to mask errors. However, larger reduction factors increase the size of the mask and decrease the throughput in terms of the wafer area exposed under the mask. We note that the mask is also referred to as the reticle in the exposure context.

The equipment used to expose the photoresist-coated wafer is known as a wafer stepper. In a wafer stepper, a small portion of the wafer, known as the exposure field or simply the

field, is exposed under the reticle through the projection optics. The illumination is then turned off and the wafer is displaced so that a different portion of the wafer is exposed in the next step. Modern wafer steppers are extremely sophisticated, with very high stepping precision. Additionally, steppers also align the wafer to the proper position so that the projected image will precisely overlay the patterns already on the wafer from previous lithography steps.

Modern wafer steppers are of the step-and-scan type in which the field is partially exposed through a slit [29], [30]. The lens and the wafer are translated synchronously such that the illumination through the slit scans the field from side to side. Due to the image reduction by the projection optics, the lens must be translated N (i.e., the reduction factor) times faster than the wafer. Illumination through a small slit restricts the area of the projection optics that is utilized, which simplifies the projection optics and reduces their distortions. A schematic of the step-and-scan system is shown in Figure 2.1.

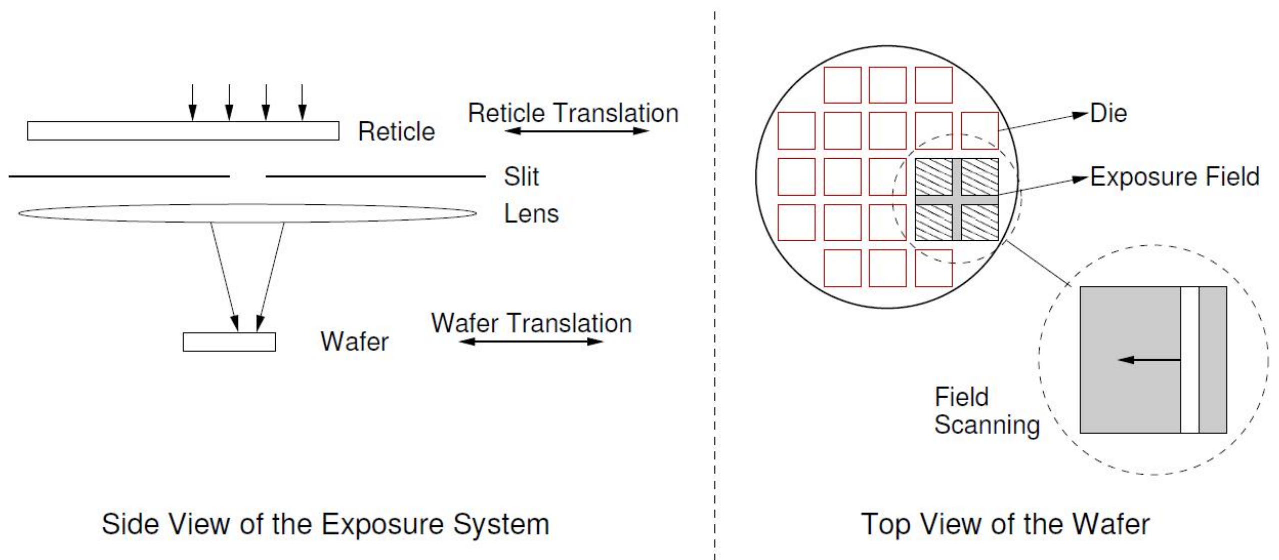


Figure 2.1: Schematic of a step-and-scan wafer stepper. [2]

After the patterning process completes, the photoresist undergoes post-exposure bake, which entails heating at a higher temperature than soft bake. The purpose of post-exposure bake is to further drive off low molecular-weight materials that may contaminate the post-lithographic equipment. Post-exposure bake also smoothes out the resist line profiles.

Then, a developer solution washes away the soluble parts of the resist and the pattern has been transferred from the mask to the photoresist.

While the patterning process is highly sophisticated, the image on the mask undergoes significant distortion as it is transferred to the photoresist. Due to the extremely small sizes of the mask features, diffraction effects, inherent to the wave nature of light, become considerable. Unfortunately, a finite-sized lens is not capable of collecting all the diffraction orders as shown in Figure 2.2, and the mask image cannot be completely reconstructed. This fundamentally limits the resolving capability of lithography, which is given by the following well-known Raleigh's equation:

$$L_{min} = k1 * \frac{\lambda}{NA} \tag{2.1}$$

- L_{min} is the minimum feature size that can be resolved.
- λ is the wavelength of the illumination source. An ArF plasma source with a wavelength of 193nm is used in modern lithography, and is projected to remain in use at least through the 45nm node.
- NA is the numerical aperture of the lens and is the sine of the maximum half-angle of light that can make it through a lens to the wafer, multiplied by the index of refraction of the medium (1.0 for air). The NA of a lens is a measure of its ability to capture the diffraction orders of light across a wide range of incidence angles.
- $k1$ is known as the k-factor and captures the capability of the lithography process; it has a fundamental lower limit of 0.25. For modern processes, k1 is around 0.3.

In addition to the minimum resolvable size, the depth of focus (DOF) is an important parameter of a patterning system. Ideally, the wafer should be placed at the focal plane of the lens. This, in practice, is infeasible and the wafer, or certain parts of it, may be positioned at a small distance, known as the defocus, from the focal plane. DOF captures the tolerance of a exposure system to defocus. DOF is given by

$$DOF = k2 * \frac{\lambda}{(NA)^2} \tag{2.2}$$

where k2 is a constant. Similar to DOF, exposure latitude quantifies the tolerance to exposure dose variations. Together with DOF, exposure latitude gives the lithography process window.

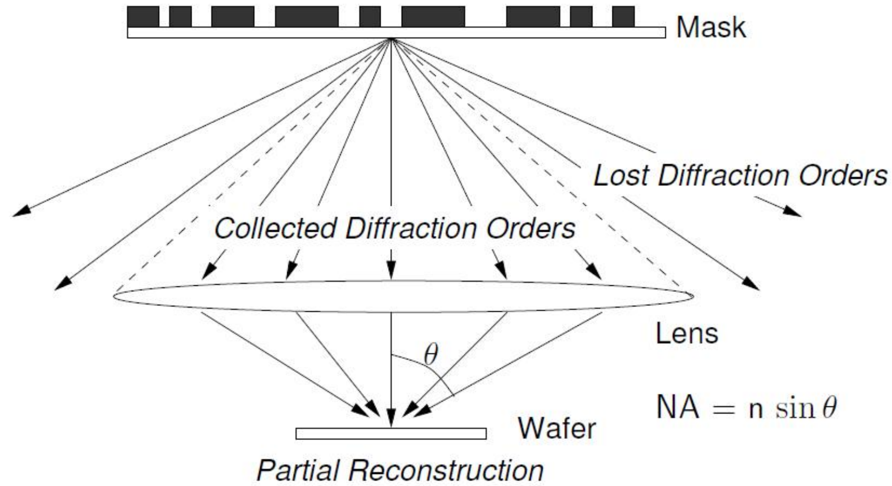


Figure 2.2: Features on the mask cannot be exactly reconstructed due to diffraction. [2]

Improvements in lithography equipment and resist technology, along with resolution enhancement techniques (RETs), reduce the k-factor and consequently the minimum resolvable size. RETs are methods used in lithography to enhance the printability of mask features. RETs are typically applied after signoff and before or during the mask data preparation stage. Commonly used RETs are as follows.

- *Optical proximity correction (OPC)* selectively alters the shapes of the mask patterns to compensate for patterning imperfections. OPC can be rule-based, which uses rules defined for different layout configurations, or model-based, which uses a lithography simulator. While OPC is very effective at reducing patterning variation, it requires a large runtime and significantly increases the mask complexity.

- *Off-axis illumination (OAI)* refers to illumination which has no on-axis component, i.e., which has no light that is normally incident on the mask. Examples of off-axis illumination include annular and quadrupole illumination. OAI improves the DOF for certain pitches while worsening it for others that are known as forbidden pitches. Fortunately, sub-resolution assist features can be inserted to eliminate or reduce the impact of the forbidden pitches.

- *Sub-resolution assist features (SRAFs)* or scattering bars are layout features that are inserted between layout features to improve their printability. SRAFs have very narrow widths and do not print on the wafer.
- *Phase shift mask (PSM)* adds transparent layers to the mask in certain locations to induce destructive interference at feature edges, which enhances pattern contrast and improves the k-factor.

Etching

Etching is used to transfer the pattern from the photoresist to the underlying thin layer. The chemical used in etching is known as the etchant; it selectively reacts with the underlying thin layer only in the areas that are not protected by the photoresist, while leaving the photoresist intact. The most common etching technique is reactive ion etching in which chemically reactive plasma is used to remove the thin layer in regions not protected by photoresist. After etching, the photoresist is completely removed by a variety of methods (e.g., dry etching [31]).

2.2.2 Sources of Systematic Variations: Chemical Mechanical Polishing

Chemical Mechanical Polishing (CMP) is the mainstream planarization technique used to remove excess deposited material and to attain wafer planarity over short and long ranges [32]. CMP involves use of chemicals to soften the material to be removed, and mechanical abrasion to polish away the material. Rotary CMP tools are the most prevalent and primarily consist of a rotating carrier on which the wafer is mounted, and a large polishing pad that rotates in the same direction. The wafer is held face down and pressed against the pad. To assist polishing, slurry, which is a mixture of abrasive particles and chemicals that soften the material to be polished, is fed onto the pad. CMP continues until the desired thickness is attained. A common method for endpoint detection (i.e., when to stop polishing), is the use of etch-stop materials which cause the motors to draw detectably more current when the desired thickness is attained. The basic setup of rotary CMP equipment is illustrated in Figure 2.3. CMP is used to planarize bare wafers, in front-end-of-line (FEOL) to remove and planarize overburden oxide, and in back-end-of-line (BEOL) to remove excess copper and barrier, and to planarize inter-level dielectric.

While several advancements have been made in CMP technology, imperfections remain and have always been a concern due to rapidly shrinking topography variation tolerances.

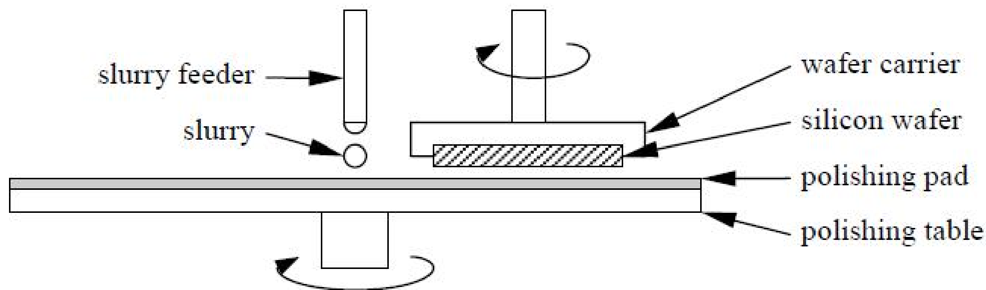


Figure 2.3: Equipment used for CMP. [3]

CMP is known to suffer from pattern-dependent problems known as dishing and erosion [33] (Figure 2.4). These two effects arise because of the existence of multiple materials of different softness that get polished simultaneously. Dishing quantifies the height difference seen in one material, while erosion captures the height loss of the harder material while polishing [34]. Two methods to reduce pattern-dependent effects are filling and slotting [35]. In fill insertion, non-functional or dummy geometries are added to increase the density of a material. A common objective is to make the material density over the chip uniform by adding fill to regions that have less material. Slotting works in the opposite way by removing material from large features without compromising their electrical functionality.

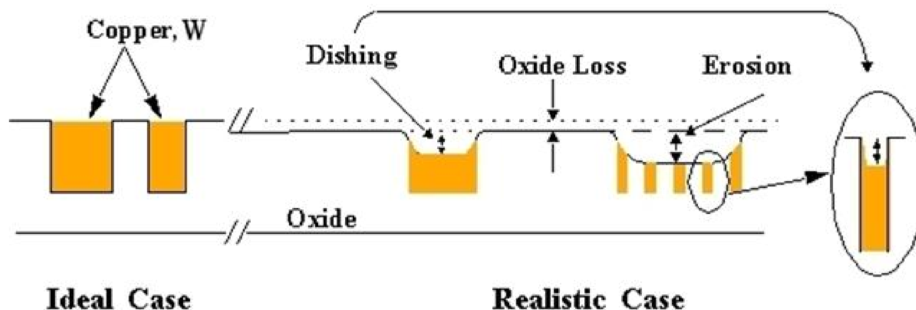


Figure 2.4: Dishing and Erosion. [4]

CMP imperfections manifest themselves into electrical variations in several ways. In FEOL, oxide dishing in STI wells and nitride erosion can cause poor isolation between devices and increase inter-device parasitics. Excessive nitride erosion into the underlying

silicon, and failure to completely remove oxide from over the nitride can cause device failure. In BEOL, copper dishing and dielectric erosion affect the interconnect resistance and capacitance, and consequently the interconnect delay. Poor planarity also poses difficulty in patterning the layers above and can cause large defocus during exposure. Planarization non-idealities also compound for higher metal layers due to the non-planarity of the underlying layer.

2.2.3 Sources of Systematic Variations: Mechanical Stress

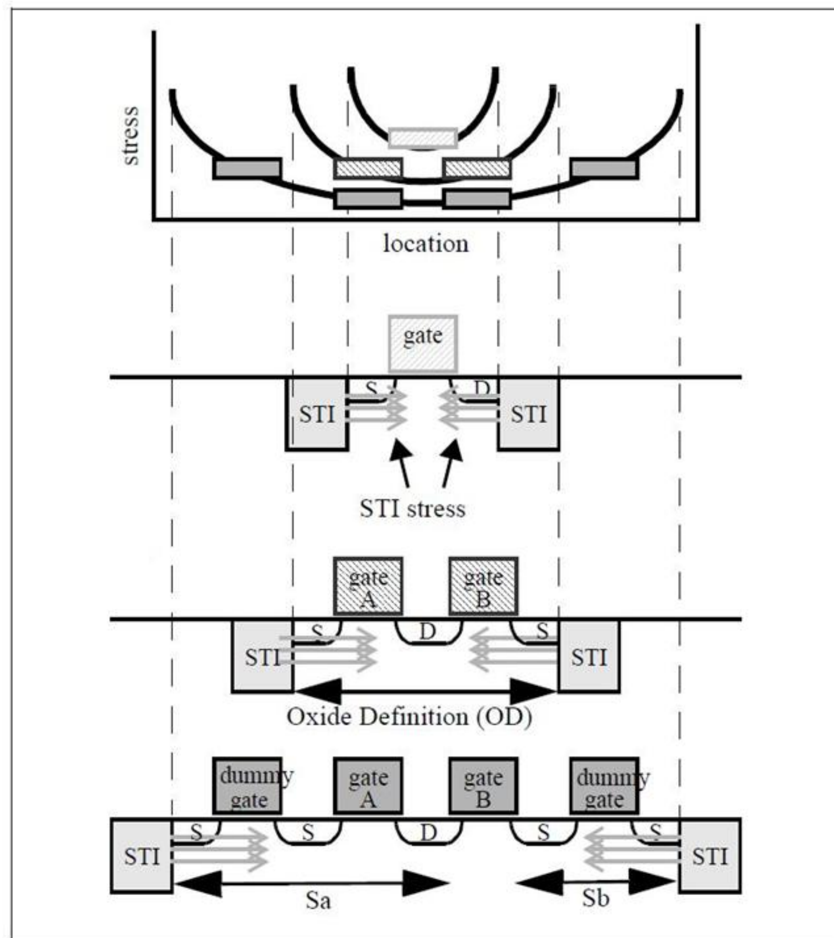


Figure 2.5: A depiction of the STI stress effect.[5]

For sub-0.25um CMOS technologies, the most prevalent isolation scheme is shallow trench isolation (STI). Mechanical stress on active regions of devices arising due to the proximity and width of STI wells are significant in existing technologies. The STI process leaves behind a silicon island that is in a non-uniform state of bi-axial compressive stress [5], [36], [37]. STI induced stress has been shown to have an impact on device performance [38], [36],[39],[37],[40],[41], introducing both Idsat and Vth offsets. These effects are significant and must be included when modeling the performance of a transistor. The stress state within an active opening is both non-uniform and dependent on the overall size of the active opening, meaning that MOSFET characteristics are once again a strong function of layout. The STI stress effect is depicted in Figure 2.5

2.2.4 Sources of Systematic Variations: Material Dopping

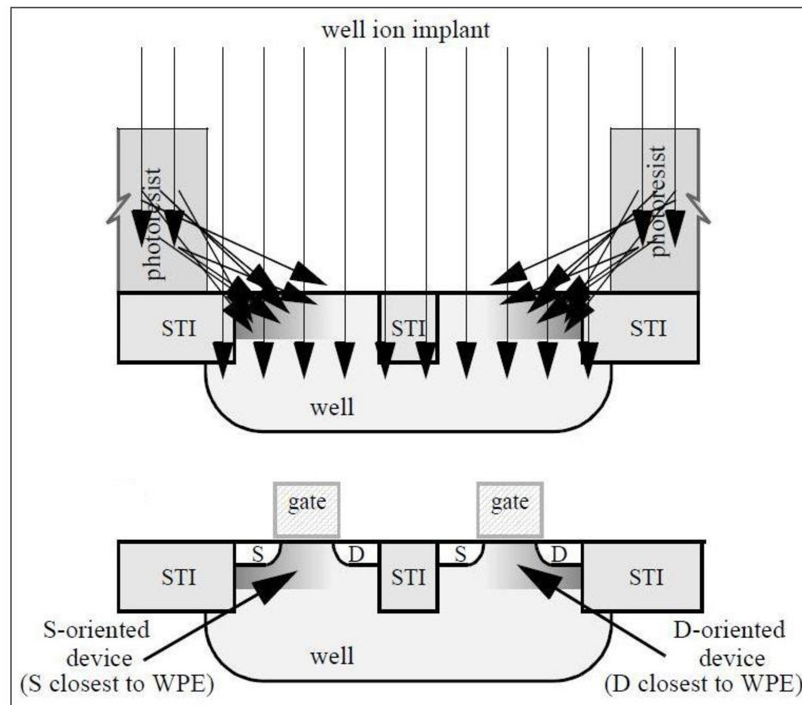


Figure 2.6: A depiction of the WPE.[5]

Highly scaled bulk CMOS technologies make use of high energy implants to form the deep retrograde well profiles needed for latch-up protection and suppression of lateral

punch-through [42]. During the implant process, atoms can scatter laterally from the edge of the photoresist mask and become embedded in the silicon surface in the vicinity of the well edge [40], [43], as illustrated in Figure 2.6 The result is a well surface concentration that changes with lateral distance from the mask edge, over the range of 1um or more. This lateral non-uniformity in well doping causes the MOSFET threshold voltages and other electrical characteristics to vary with the distance of the transistor to the well-edge. This phenomenon is commonly known as the well proximity effect (WPE).

2.3 Yield Issues Due to Process Variations

Yield is defined as the number of chips that function and meet delay and power specifications, expressed as a percentage of the total number of chips manufactured. For a mature process, yield of over 90% is typical. However, during process development and ramp-up, the yield can be much less. Yield is commonly classified into the following two categories.

1. Functional yield or catastrophic yield is the percentage of chips that are functional. Examples of functional failures that limit functional yield are shorts and opens in wires, open vias, line-end shortening, etc.

2. Parametric yield is the number of chips that meet delay and power specifications, as a percentage of the functionally-correct chips. Parametric yield loss is due to chips that are functional but cannot be sold because they fail to meet the delay and power specifications.

A variety of process variations and defects cause yield loss. Functional yield loss is usually caused by misprocessing and random contaminant-related defects. Parametric yield loss is typically due to process variations. However, process variations can also cause functional failures (e.g., line-end shortening leading to an always-on device) and defects can cause parametric yield loss (e.g., particle contamination that causes interconnect thinning but not a complete open).

While yield loss due to functional failures is significant, parametric failures have gained significance and now dominate functional failures. Arguably, measures to improve parametric yield are more challenging to develop and adopt. While most functional yield-enhancing methods are geometric and applied after signoff, parametric yield-enhancing methods often require understanding of the nature of process variations and modeling of their electrical effects. In this thesis, we focus on techniques that address parametric yield loss.

2.3.1 Impact of Process Variations on Parametric yield

Process variations, which are the primary cause of parametric yield loss, manifest themselves as circuit metric (power and delay) variations in the following ways.

Lateral dimension variations

As design scales down to more advanced nodes, the more difficult it becomes to accurately print the desired shapes on the wafer. As illustrated in Figure 2.7, at the start of 45 nm, traditional lithography techniques were not sufficient to print difficult minimum feature structures, referred to as the critical dimension (CD). Immersion lithography improved the k1 factor and restored stability all the way though 32 nm. However, at 22 nm, it may become very difficult to print shapes accurately even with immersion lithography, and without immersion, impossible. Due to the absence of lithography tool improvements, printing any random set of shapes at 22 nm with the current sources of light and lithography techniques may both be infeasible and unnecessary for design.

On the polysilicon layer, CD refers to the linewidth of the gate poly, which is equivalent to the gate length or channel length; on metal layers, CD is the wire width. Process variations affect the CD the most, and are manifested as delay and power variations of the circuit. For example, decrease in the gate length will decrease the device delay and capacitance, but dramatically increase subthreshold leakage. Decrease in the wire width will increase resistance but decrease capacitance.

Significant sources of CD variation are exposure and etching variations in lithography [44]. During exposure, CD variation is due to mask errors [45], resist thickness variation [29], exposure dose variations [29], defocus [29], lens aberration [46], [47], etc. Microloading effects during etching also cause CD variation [48], as microloading results in differing etch depths for different pattern resolution and densities. A substantial fraction of CD variation arising due to the these exposure and etching variations is considered systematic.

The variation in transistor's channel length has direct a impact on several electrical properties of a transistor, however, the most affected parameters are the transistor's drive current ($I_D \propto 1/L$) and V_{th} [18, 24]. The variation in V_{th} arises due to the exponential dependence of V_{th} on channel length L for short channel devices, mainly due to drain induced barrier lowering (DIBL) effect [9]. DIBL causes V_{th} to be strongly dependent on the channel length L as shown in Figure 2.8. V_{th} reduction due to DIBL can be modeled as [9]:

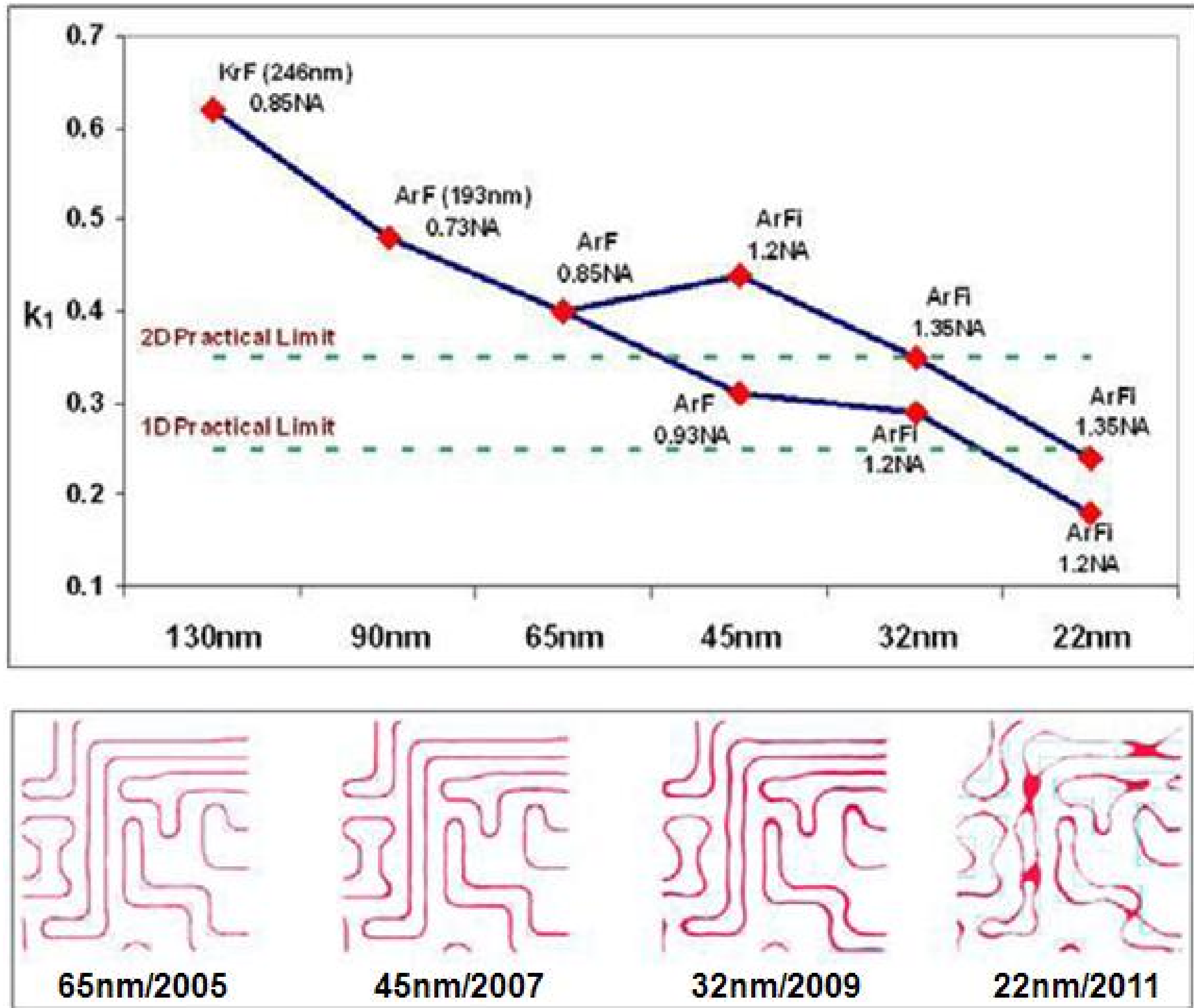


Figure 2.7: The limits of lithography techniques and the changes in printed results across process nodes [6]

$$V_{th} \approx V_{th0} - (\zeta + \eta V_{DS})e^{-}L/\lambda \tag{2.3}$$

where η is the DIBL effect coefficient, and V_{th0} is the long channel threshold voltage. Therefore, a slight variation in channel length will introduce large variation in V_{th},as shown in Figure 2.8.

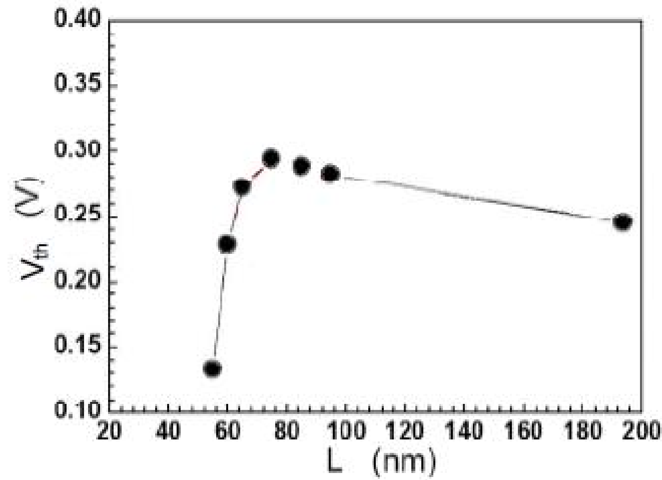


Figure 2.8: Measured Vth versus channel length L for a 90nm which shows strong short channel effects causing sharp roll-off for Vth for shorter L.[4]

Transistor performance depends heavily on gate dimension. A small gate variation changes the channel length, creating a variation in I_{on} and I_{off}. Dependence of transistor current is increasingly non-linear to channel length. As a result the variability in current I_{on} and I_{off} has been increasing with process node size, as shown in Figure 2.9. 10% transistor gate variation can translate to as much as a variation of -15% to +25% in gate delay, according to the left hand graph in Figure 2.9 [7]. Still, ones worse variations are seen for I_{off} on the right. The impact of variability on leakage power has been reported that in at least one instance, as little as 6% of CD variations produce enough leakage to create a chip failure. Shape variations on transistors directly and disproportionably translate into performance variation. It is currently impossible to accurately predict IC design performance without modeling final on-chip device shapes and process variations,

especially from 90nm technologies downwards. However, for predictable silicon success, these shape and process variations must be predicted and taken into account early in the design process [49].

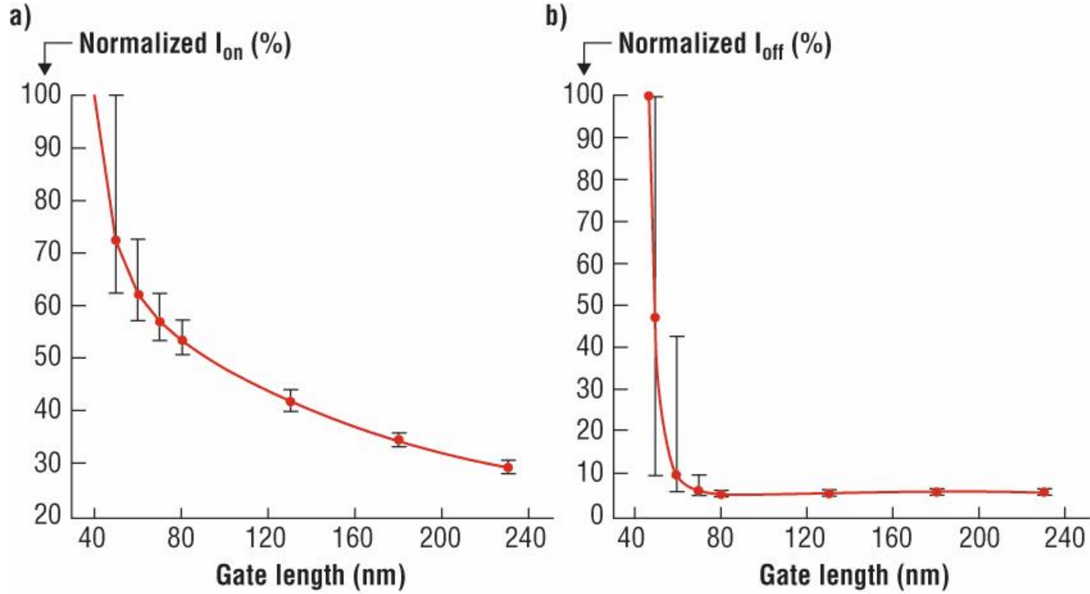


Figure 2.9: Ion and Ioff variations due to transistor length [7]

Topography variations

Chemical mechanical polishing (CMP) is performed between lithography steps to attain the designed layer height and to planarize the layer for successive process steps. Unfortunately, CMP is imperfect and cannot eliminate topography variation. Topography variation changes the metal height in back-end of the line (BEOL) which affects the wire resistance and capacitance. CMP for front-end of the line (FEOL) is used to planarize the oxide that is deposited for STI. Imperfect FEOL CMP leads to defocus during polysilicon patterning and poor inter-device isolation.

Topography variation is understood to be partly systematic for both FEOL [3] and BEOL [34]. Another example is gate oxide thickness variation which affects gate-tunneling leakage and device subthreshold slope. Gate oxide is manufactured by light oxidation and its thickness variation, though small, is considered random.

Interconnect parasitics are significant and complex components of circuit performance, signal integrity and reliability in IC design. Copper/low-k process effects are becoming increasingly important to accurately model interconnect parasitics. Sub-90nm interconnects with their narrow and tall (Z-plane) configurations create significant variations in the Z-plane, which, when added to the X-Y variations due to RET, have a large impact on parasitics.

These variations are amplified because metal sheet resistance varies as a nonlinear function of line width, as shown in Figure 2.10. Copper resistivity rises dramatically below 90nm due to increased electron scattering on grain boundaries and interfaces. This phenomenon causes significant systematic changes in sheet resistance as a function of line width. If line width cannot be predicted accurately, the result is inaccurate resistance and delay prediction.

According to ITRS roadmap, copper dishing/erosion after Chemical-Mechanical Polishing (CMP) and scattering effect may increase resistance significantly. Also, coupling capacitance between wires becomes dominant over ground and fringing capacitance at 0.18um technology (over 60% of the total capacitance), and increases rapidly with higher wire aspect ratio of the advanced technologies [50], [51],[52]. Therefore, interconnect delay will suffer from the increased resistance and coupling capacitance more seriously in the future [53].

Material variations

The STI stress effect is depicted in Figure 2.5. It has been shown that the residual stress and corresponding shift in electrical performance can be qualitatively described by two geometric parameters, SA and SB [36], [37] (Figure 2.11). These represent the distance from the gate to the edge of the OD on either side of the device. MOSFET parameters such as Vth, peak gm and Idsat have been shown to vary linearly with the following function,

$$Stress = \frac{1}{S_A + \frac{L}{2}} + \frac{1}{S_B + \frac{L}{2}} \tag{2.4}$$

This relationship has been incorporated into existing SPICE models [9] to allow this phenomenon to be included in circuit simulations.

Stress due to STI is compressive and typically enhances the mobility of PMOS while degrading the mobility of NMOS. Consequently, delay and leakage increase for PMOS while decreasing for NMOS. Several techniques have been proposed to reduce STI stress-induced

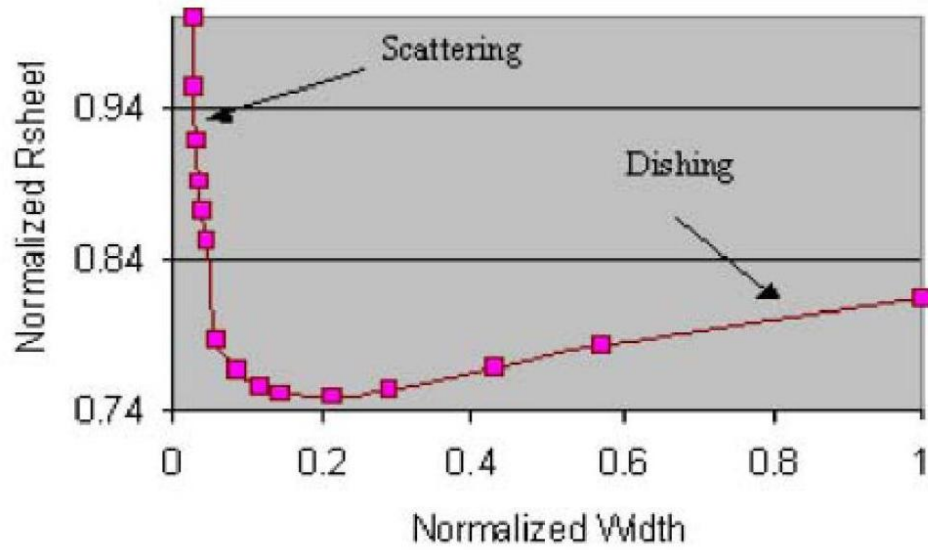


Figure 2.10: Interconnect resistance variation due to width. [8]

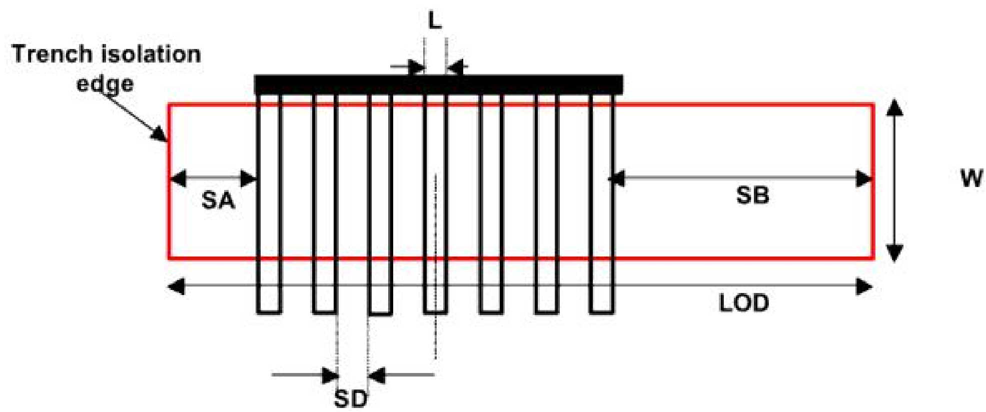


Figure 2.11: MOSFET device geometry using a shallow trench isolation scheme. [9]

variation [54]. STI stress is highly systematic and is partly modeled in today's design flows. Recent works have proposed modeling the residual STI stress effects [55].

In addition, material variations cause variations in carrier mobility, polysilicon resistance, etc. Dopant concentration in the device channels affects the threshold voltage and consequently subthreshold leakage and device delay. Due to the small number of dopant atoms in the channel in modern devices, dopant density varies significantly and randomly as a percentage and induces substantial random variation in threshold voltage. The impact of well proximity on the threshold voltage (Vth) of the nMOS device is shown in Figure 2.12. The Vth is found to increase by as much as 50mV as the device moves closer to the well edge which is consistent with previous observations [40],[43].

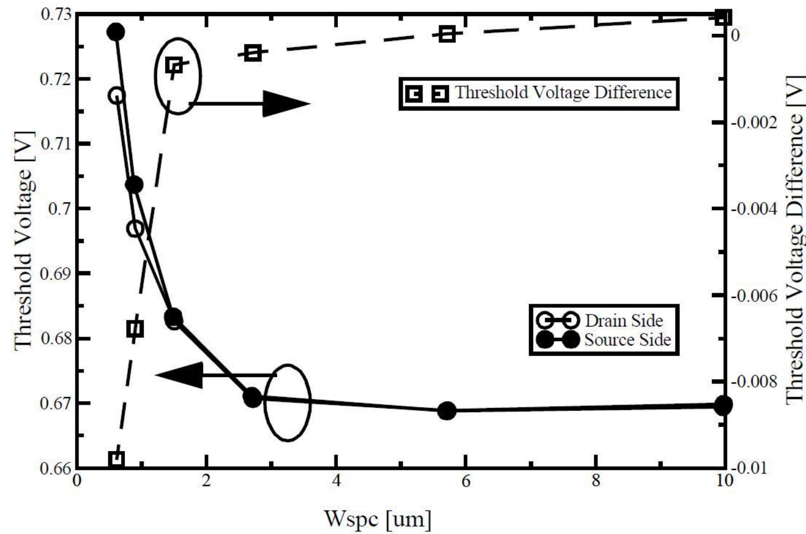


Figure 2.12: Vth versus well-edge distance for 3.3V nMOS device.[5]

As a consequence of these manifestations, a significant variation is seen in circuit delay and leakage. With technology scaling, process variations are increasing as a percentage, and consequently the delay and leakage variability is increasing. There is considerable parametric yield loss today especially during yield ramp-up phase causing substantial value loss.

2.4 Design for Manufacturing

While DFM has attracted great deal of attention recently from industry and academia, several techniques that can be arguably be considered DFM techniques have been in use for several years.

Design for Manufacturing (DFM) refers to measures taken during the design process to enhance yield. Parametric yield enhancement facilitated by DFM can contribute to improvement of design performance and/or power, and/or designer productivity. DFM techniques can compensate for, reduce, or make the design more robust to various types of manufacturing non-idealities.

2.4.1 Traditional DFM Methods

Traditionally, design and manufacturing have been conveniently kept separated, with only minimal information exchange. One traditional DFM category is a technique that controls the design implementation to make it more robust to various types of manufacturing non-idealities. Typically, these techniques are on the physical layout level. Examples of these techniques are:

- Design rule checking (DRC). Design rules have been the primary method for the foundry to convey manufacturing limitations to design. Design rule checking, verifies adherence to these rules, and a design that is design rule correct is expected to have a high functional and parametric yield. Simple examples of DRCs are minimum spacing, minimum and maximum dimension or area, and minimum and maximum metal density.
- Restricted design rules (RDR). The basic concept behind restricted layout is to limit what the designer is allowed to do. Designers want a wide assortment of possible features or shapes or constructs that they are allowed to use. Conversely, the fab would like a very limited set of what the designer is allowed to useideally, everything would be perfectly regular and repeatable, making designs much more simple to process and much more robust against manufacturing variability. For example, lithographic rounding of both the active and the contact in a source or drain connection can reduce the alignment marginality, creating the potential for a resistive contact. In a gate construct, horizontal bends in the field poly near the gate can induce an inherent systematic variation in the L-effective on the corner of the gate. With misalignment, that variation can be quite dramatic, causing the device in this

transistor to have more variation in terms of drive current, leakage, etc. Similarly, because of a horizontal-to-vertical transition in the active layer, this curvature can cause variation in the W-effective, affecting the drive strength, among other aspects. With alignment variability, this effect will vary dramatically as well.

The restricted design compromise is to have some assortment of allowable features or constructs, but a much smaller list and a more controlled list than what has been allowed in the past. With restrictive design, future designs will be much more regular than current designs, and therefore actually manufacturable.

Another category, are DFM techniques that account for various types of manufacturing non-idealities during the front end design phase. Typically, these techniques are on the electrical level of abstraction. Examples of these techniques are:

- Guardbanding. Considerable margin is allocated during design to account for process variations. Today's timing and power analysis flows are corner-based, i.e., a set of conservative process, voltage, and temperature (PVT) settings are assumed in analysis. With respect to process variations, hold and setup time checks are performed at fast and slow process corners respectively. Leakage is typically highest at the fast process corner, but the use of typical process corner to reduce pessimism in analysis is common. The premise behind corner-based flows is that if the design meets its specifications at conservative PVT settings, it will meet them at all other conditions. Unfortunately, this premise is not true and is now breaking down due to complex dependence of electrical metrics on variations. For example, shorter gate lengths do not necessarily have higher leakage (due to reverse short channel effect), and wider wires are not necessarily faster. The above techniques were relatively easy to adopt and served well until the 130nm node. Since then, as the complexity and extent of process variations has increased, these techniques, while remaining necessary, are no longer sufficient.

Several problems stem from the inadequacy of these techniques and call for novel DFM techniques that explicitly target yield enhancement.

- With scaling, as process variations have become complex and large, design rules are no longer able to capture the variations completely and precisely. In modern technologies, layout regions that do not meet design rules may yield well, while those that meet them may not. Thus, design rules have become extremely complex in an attempt to capture process variations that arise due to complex layout configurations.

Recommended design rules, which are preferably but not necessarily required to be met, have also been introduced. A large set of design rules poses maintainability problems, and limits the freedom of optimization algorithms and tools in physical design.

- The use of restricted design rules (RDRs) [56], for example those that enforce regularity by allowing only one or two pitches, increases the chip area.
- Corner-based analysis assumes conservative process conditions; this is overly pessimistic since all parameters have an extremely small likelihood of being at their conservatively assumed values at the same time. Moreover, the design metric under analysis may have a non-monotonic dependence on process parameters, in which case worst-casing the process parameter will not result in worst-casing of the design metric. To reduce pessimism and improve worst-casing of design metrics, analysis is performed at a large number of corners. Unfortunately, the number of corners can grow rapidly with process parameters and the analysis can be both pessimistic and risky at the same time [57]. Furthermore, corner-based methods cannot account adequately for inter-die variations since all components are assumed to be at the same process corner. A notable exception is on-chip variation analysis which allows clock and data path components to be at opposite corners.
- As guardbanding increases and compromises the advantages from scaling, designers are under tremendous pressure as they seek to meet market expectations. To improve delay, power, and area of the design, considerably more time must be spent on iterations and fixing violations. This reduces productivity.
- Design rules and guardbanding can no longer be sufficiently pessimistic to ensure high parametric yield. Unexpectedly large variations and failures can cause intolerable yield loss, and require costly design re-spins.

2.4.2 Taxonomy

To really bridge the gap between design and manufacturing, it is important to **model** and feed proper manufacturing metrics and cost functions upstream to the design side. Model based DFM techniques can be broadly classified into the following two categories depending on the yield loss component that they address.

1. **Functional yield enhancing.** Process-based, or sometimes referred as physical DFM technology, is concerned with identifying and correcting particle defects or

process variations that can lead to functional failures like shorts and opens. Several model-based techniques have been proposed to make the design robust to random and systematic variations. As for random contamination-caused defects and large process variations, critical area analysis [58] finds the chip areas that have a high chance of causing functional failures under an assumed contaminant particle size distribution. For systematic variations, hotspot detection [59] flags chip areas that are vulnerable to large variations due to lithography non-idealities.

Examples of corresponding design enhancements include wire spreading, wire widening, and via doubling. Physical DFM also implements geometric yield improvement with recommended rules. Functional yield enhancement techniques are simpler and easier to adopt because they are primarily shape-centric and have limited or no interactions with electrical metrics such as delay and power. Physical DFM is an important technology, especially during the early stages of new process development, when low functional yield is the primary obstacle to process qualification. The next critical milestone in nanometer design is the creation and validation of transistor and interconnect models that are accurate enough to ensure that predicted versus actual circuit performance supports design objectives, reduced costs, and increased product functionality.

2. **Parametric yield enhancing.** Parametric-related yield losses are caused by variations in electrical properties. As the physical DFM technologies are familiar to manufacturing and CAD personnel responsible for ensuring designs meet the basic requirements for manufacturability. In slight contrast, parametric-based, or sometimes referred as electrically-based DFM, is the knowledge that can be gained from the manufacture and critical to designers who must ensure that designs meet datasheet performance specifications and competitive yield targets. These techniques have attracted great interest recently as they address an ever-increasing and now dominant yield loss component. The objective of these techniques is to contain the variability in delay and leakage. This thesis focuses on such DFM techniques.

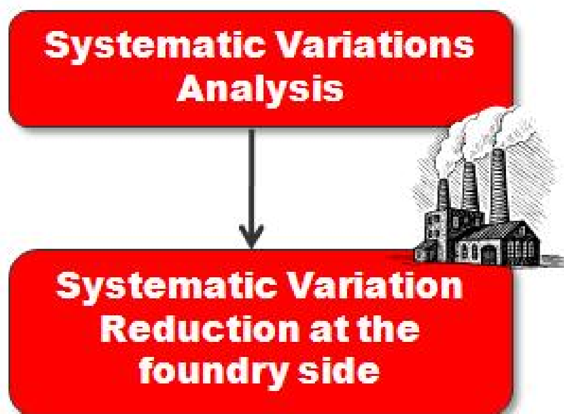
As illustrated in Figure 2.13, parametric yield enhancing techniques can be summarized as follows: starting with modeling and accounting for systematic variations followed by DFM techniques either for:

- Process Variation Reduction.

or

- Systematic Variation-Aware Analysis and Design Enhancement.

Systematic Variation Reduction DFM Flow



- RET: Resolution Enhancement Techniques to compensate for litho effects
- Dummy fill insertion to compensate for CMP effects
- Via redundancy & Wire widening to avoid random effects

☒ Higher mask cost
☒ Designers are still sending non-friendly designs

Systematic Variation Aware Analysis and Design Optimization DFM Flow

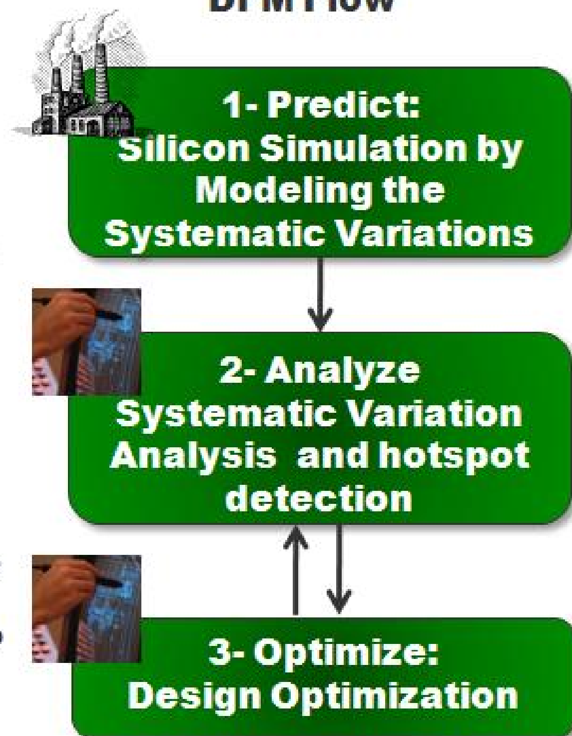


Figure 2.13: Systematic Variation Aware Analysis and Design Enhancement

Process Variation Reduction

1. **Reducing Random Particle Defects:** Random particle defects are a fact of IC manufacturing, but the number and severity of these defects increases as layout features and the space between them continue to shrink. Recommended design rules use manufacturing information to suggest spacing's that reduce the overall chances of random particle defects. However, recommended rules can generate thousands of rule violations, with no information to help you decide which ones are most critical. Critical area analysis (CAA) uses manufacturing information about the process particle sizes and probability distribution to identify specific areas of an integrated circuit layout with a higher than average vulnerability to random particle defects.

2. **Reducing Lithography Variation:** After modeling the lithography process, RETs are used as the primary methods to reduce lithography variations. As explained earlier, the purpose of these techniques is to minimize the lateral distortion between the drawn and the on-silicon shapes. Typically, RETs are transparent to the design phase and are performed after signoff. However, modifications can be made to circuits that make them more amenable to RETs, such that the RETs achieve stronger reduction of lithography variation.

3. **Reducing Planarity Variation:** The impact of chemical mechanical polishing (CMP) is inherently pattern-dependent. Whether a feature is isolated or located in a dense array affects both polishing time and results. Likewise, circuit performance is also affected by how the physical layout reacts to CMP. Controlling both thickness and capacitance variation is a balancing act. Several industrial tools can model and simulate the CMP process, based on that dummy fill insertion and slotting techniques are introduced. Dummy fill insertion and slotting are the primary design techniques used today to aid planarity by altering the density. For signal wires that are routed by gridded routers, metal density typically does not exceed nearly 50% because inter-wire spacing is nearly equal to the wire width; for these wires, slotting is not required. Slotting is done for special wires such as power/ground rails and is less desirable than fill insertion [60]. Fill insertion is the mainstream technique to increase density both for FEOL CMP [3] and BEOL CMP [34].

Systematic Variation Aware Analysis and Design Enhancements

To really bridge the gap between design and manufacturing, DFM must due the following two jobs: First, bring manufacturing awareness up into the design flow. Second, communi-

cate the intentions of the designer to the manufacturing flow (Figure 2.13). In general, the objective of DFM is to improve IC yield and cost by increasing manufacturing-awareness in the design phase, as well as design-awareness in the manufacturing phase. Therefore new design for manufacturability (DFM) paradigm has emerged in the recent past. To achieve this dual objective of DFM, design must be driven by **models** of variation in the manufacturing process and the manufacturing process, must be made aware of the design intent. The DFM paradigm encompasses a set of **design methodologies that address manufacturing and process non-idealities at the design level** to make ICs more robust to variations. DFM is also interpreted as a set of post-layout design **fixing** techniques that enhance and ease manufacturability.

Modeling Sytematic Varations: Physics-based models perform modeling of fundamental physics and chemistry of processes in a lithography system. Similarly for the other processes, these models allow process engineers to simulate the impact of process parameters and material chemistry changes, and thereby tune the process. Physics-based modeling is inherently complex because of the difficulty of capturing inputs and computing model parameters. Phenomenological models do not model detailed physics and chemistry of optics and materials involved in pattern transfer, and work with only a limited set of process (e.g., resist sensitivity) and equipment (e.g., numerical aperture, partial coherence factors) parameters. Fitting phenomenological models to experimental data using a limited set of parameters is less complex than fitting physics-based models.

Lithography models, mathematically represent the distinct steps in the 3D patterning sequence [61]. Figure 2.14 shows a basic schematic of the calculation steps required for lithography modeling [10]. A brief overview of the physical models found in most lithography simulation programs is provided below.

- **Aerial Image:** The extended source method is used to predict the aerial image of a partially coherent diffraction limited or aberrated projection system based on scalar and/or vector diffraction theory. Single wavelength or broadband illumination can be used. The standard image model accounts for the important effect of image defocus through the resist film. Mask patterns can be one-dimensional lines and spaces or small two dimensional contacts and islands. Phase-shifting masks and off-axis illumination can be simulated and pupil filters can be defined.

- **Standing Waves:** An analytical expression is used to calculate the standing wave intensity as a function of depth into the resist, including the effects of resist bleaching, on planar substrates. Contrast enhancement layers or top-layer anti-reflection coatings can also be included. High numerical aperture models include the effects of non-vertical light propagation.

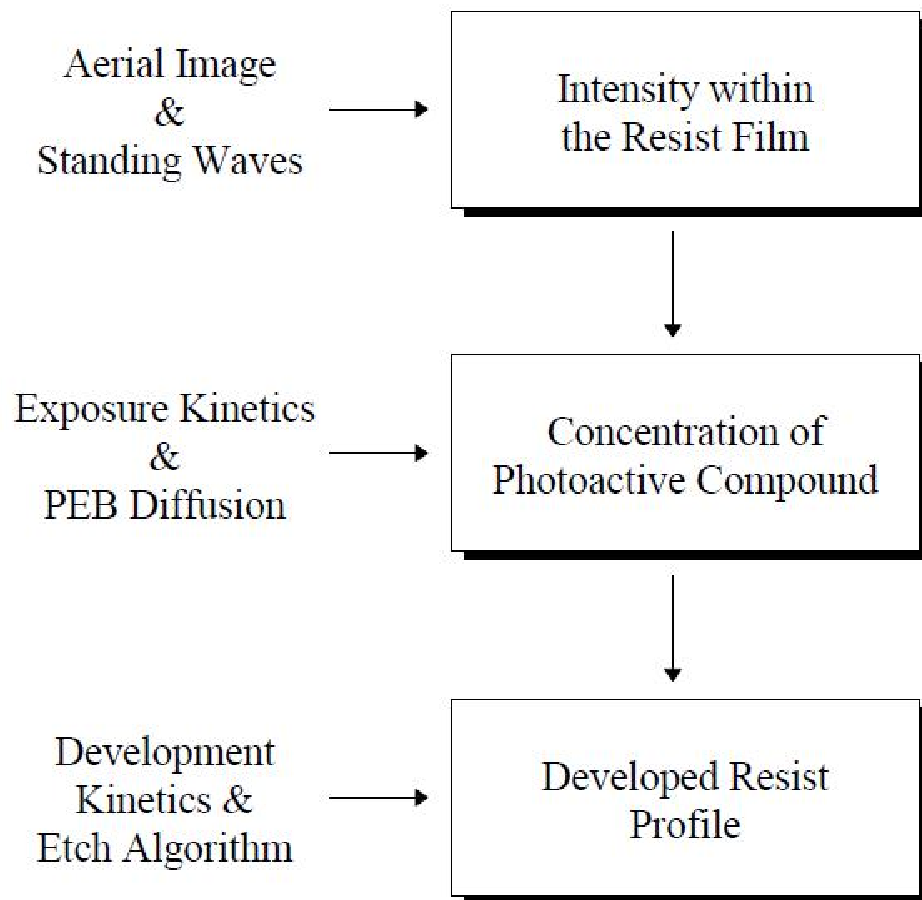


Figure 2.14: Flow diagram of a lithography model [10]

- **Prebake:** Thermal decomposition of the photoresist photoactive compound during prebake (also called post apply bake) is modeled using first order kinetics resulting in a change in the resist's optical properties. Many important bake effects, however, are not yet well understood.

- **Exposure:** First order kinetics are used to model the chemistry of exposure. Both positive and negative resists can be simulated.

- **Post-Exposure Bake:** A two-dimensional or three-dimensional diffusion calculation allows the post-exposure bake to reduce the effects of standing waves. For chemically amplified resists, this diffusion is accompanied by an amplification reaction which accounts for crosslinking, blocking, or deblocking in an acid catalyzed reaction. Acid loss mechanisms and non-constant diffusivity can also be simulated.

- **Development:** Kinetic models for resist dissolution are used in conjunction with an etching algorithm to determine the resist profile. Surface inhibition or enhancement can also be taken into account. The combination of the models described above provides a complete mathematical description of the optical lithography process. Use of the models incorporated in a full lithography simulation package allows the user to investigate many interesting and important aspects of optical lithography.

- **Etching:** To cope with the effect of edge topography on device characteristics, due to the high integration of VLSI's, accurate three-dimensional (3D) topography simulation models are required [61].

Analyzing systematic variations and detecting hotspot: Process variations can be observed at the topological or the electrical level. The topological level refers to the shape and physical attributes of the devices, including critical dimension (CD), minimum pitch and pattern density. Depending on the topological change, the electrical behavior can be affected if the variations occur in electrically sensitive areas of the design (e.g., polysilicon gate width differences between a pair of matched transistors, or narrowing of long metal lines that subsequently increase resistance). At the same time, topological changes induced by some types of dummy fill, short interconnect paths or non-critical corners of diffusion layers will not present any observable differences in the electrical behavior of the devices. Until now, only shape-centric parameters such as fidelity (achieved by aggressive RET) and pattern robustness (achieved by manufacturing-aware design) have been considered in the manufacturability of designs. However, other processing effects that are less dependent on shape or structure can also contribute to the final operation of the electronic devices. For example, even when poly gate widths are perfectly matched and robustly

built, processing steps such as ion implantation, diffusion or material selection can make an otherwise correct design fail.

With the former paragraph in mind, the scope of this work is to apply relevant DFM principles to quantify some electrical variables that shape and topology have on the electrical behavior of devices.

Design Enhancement: DFM is an art of balance to accomplish multiple goals such as high die yield, aggressive product performance, and zero reliability fallouts, which are often loosely tied or even incompatible. In that role, DFM has to rely on cross-disciplinary enhancement techniques, which deal with multiple design, manufacturing and product issues. Therefore, it requires a multi-objective optimization approach [62]. At the same time, DFM has limited resource base, as it should operate within the predetermined die area, product schedule, circuit timing budget, and design effort and yet make a positive difference. DFM enhancements need a lot of heavily processed information, e.g., in the form of litho, yield, or device models and the results of their implementation hinge on the accuracy of these models. From this perspective, one can divide recent DFM enhancement techniques into three groups aligned with the stage of their introduction in design cycle:

1. Design and process definition stage. At the earliest design definition stage, DFM aligns with the generic methodology of IC design rules and guidelines. Here, new DFM concepts emerge as correct-by-construction architecture at single cell level [63]. To assist the process, layout tolerance bands are introduced to help preserve design intent, i.e., ensure the expected electrical performance [64].

2. Design execution stage, correct-by-construction is also proposed, this time for multi-cell IP (e.g., modular design of multi-port cells [65]), die level routing [66], and then, at mask level [67]. For the complete IC layout, the new concepts on how to validate design data, mask data, or both [68], emphasize process proximity correction for single or multiple layers [69] and [70].

3. Finally, at the latest DFM verification stage, enhancement techniques are used to enhance manufacturing feedback to design/layout by correcting local yield limiters (hot spots, [67], [68], [69], [70], [71]). DFM verification techniques based on inspection or simulation, involve methods, systems, EDA algorithms, computer programs, and design kits [69], [70], [71], [72], [73], to improve circuit performance before more mask sets are built for the manufacturing line. After the improved design is placed in the fab, data collection systems need to provide feedback e.g., from defect detection [74].

2.5 Summary

Variations in the IC manufacturing process are manifested as

1. Deviation from the intended shapes of IC geometries, and
2. Variations in impurity (i.e., dopant) concentrations.

These variations are composed of systematic and random components. The systematic component of variation can be attributed to specific sources in the manufacturing process, while the random component is usually a result of confounding of several sources of variation and cannot be attributed to specific sources. A significant fraction of the total variation in shapes of IC geometries is systematic in sources such as focus, exposure dose, lens aberrations, during the lithography process. Mechanical stress and CMP processes are additional sources of process systematic variations. The primary goal in the associated research is to develop CAD solutions that can drive systematic variation-aware design analysis and enhancement.

Chapter 3

A Methodology for Analyzing Process Variation Dependency on Design Context and the Impact on Circuit Analysis

3.1 Introduction

As VLSI technology scales down towards 65nm and below, performance of integrated circuits is increasingly affected by fabrication variations. Moreover the fabrication variations at these advanced IC technologies are highly dependent on the physical layout and recently more dependent on proximity effects of this layout from neighboring cells. In a normal design process; designers only deal with schematics design entry or front end High Description Languages (HDL) digital codes. In this design entry phase, no layout information is available for the front end designers. This is acceptable in the IC technologies above 65nm; where most of the required layout or back end information is captured in the front end entry, so the layout shape has minimal impact on the original design. However; this is no more acceptable in the new advanced IC nanometer technologies; more and more layout parameters are now impacting directly the transistors performance. This makes the final fabricated IC performs differently from the original design, unless these layout effects and variations are to be considered earlier in the design cycle.

As an example, mechanical stresses induced by Shallow-Trench Isolation (STI) on the gate channel can alter the drive current of NMOS and PMOS transistors up to 20% de-

pending on the length of diffusion (LOD) [37] [75]. As a consequence, the drive current of a transistor is not only related to the parameters of the gate such as the gate length L and gate width W but also is related to the exact layout shape of the individual transistor. Similarly the lithography printability is strongly layout dependent [76]. As the density of VLSI circuit increases the proximity effects are becoming a significant consideration to design layout linewidth control in range of nano-meters using current state-of-the-art photolithography tools that operate with deep ultraviolet (DUV) light wavelengths of 193nm. Optical light rays in the photolithography process diffract and pass through the reticle in a different matter based on the layout context and mask pattern density, and hence impact the final Critical Dimension (CD) printability on wafer.

Therefore, when designing a standard cell or characterizing its electrical behavior prior to use in a chip implementation flow, it is important to know the impact of layout context-dependent manufacturing variations on the cells parametric views. Traditional standard cell characterization treats the cell in complete isolation from its context. This is somehow accepted for older technologies. However in advanced nano-IC technologies at 65nm and below this is not the case. During standard cell design, it is recommended to have a context-dependent parametric analysis that can be used to improve the cell layout and architecture so that variability is minimized. The proposed CAD tool presents a methodology to predict the impact of context-dependent systematic manufacturing variations on the parametric behavior of standard cells in an integrated circuit. Such a methodology can be applied to the analysis of a full chip composed of standard cell components, and reports layout context-dependent changes in chip timing and power. For lithography and stress variability, a study on 65nm, 45nm standard cell libraries and mid-size designs is done to examine the influence of cell context when looking at gate CD and the design performance. A context generator engine is implemented to study the effect of the design context variation. This engine has several user defined parameters to enable the user to simulate various context conditions.

The rest of this chapter is organized as follows: Section 3.2 presents the available solutions that analyze process variation and studies the impact on circuit performance. Section 3.3 discusses the stress and lithography variation dependence on design context. Section 3.4 depicts the proposed characterization methodology entailed in implementing the engines used in our flow. In order to verify the usefulness of our proposed methodology, industrial standard cell context-aware characterization is performed. Impact on electrical parameters will be examined on the context-aware standard cells and results are highlighted in Section 3.5.

3.2 Background

One of the major challenges facing semiconductor companies today is how to increase yield. Since yield is directly related to profitability, by predicting and improving yield before tapeout, an IC designer can have a direct dollar impact on the success of his design. The ability to predict and improve yield becomes even more vital as processes move to geometries under 100 nanometers. In fact, there are discouraging predictions that yields for ICs with geometries below 100 nanometers may not exceed 50 or 60 percent [77]. To account for process variations, an IC designer not only has to design for good electrical performance, but also for high manufacturing yield [78]. Electronic Design Automation (EDA) companies need to develop tools and methodologies that designers can easily incorporate into their flows to meet this challenge. There are many factors that affect yield. Physical problems that are highly dependent on layout can effect yield. Manufacturing issues such as defect density on the silicon, maturity of the process, and effectiveness of design rules also affect yield. Another factor, and the main topic of this chapter, is how the design will react to process variations, mainly speaking about systematic variations such as lithography, stress and CMP. The designs sensitivity to these variations is called parametric yield. Increasing parametric yield is a challenge within a designers realm. There are currently several tools, or analyses, that a designer may use to check the design for process variations. This section provides a brief description of the commonly used methods, including corners simulation, Monte Carlo analysis, response surface modeling.

3.2.1 Corners Simulation

Corners simulation is perhaps the most widely used method to test for process variations. With this method, a designer determines the worst case corners, or conditions, under which the design will be expected to function. The modeling group may provide the corners for the process. Next, each corner is simulated, and the output of each corner is examined to determine whether or not the design performs as required under each of the specified conditions. There are a number of issues with corners analysis. Perhaps the most troublesome of these is if the corners are not provided, then the designer may not know what the corners actually are, leading him to frequently make a best guess. Guessing worst case corners that will never occur in reality not only wastes design time, but also can result in a design that takes more area or consumes more power than needed, or even becomes impossible to design. Not guessing corners that will occur in reality may result in lower yield. The advantage of corners analysis is its relative simplicity. Corners analysis is neither compute-intensive nor time consuming (as compared to Monte Carlo

analysis, for example). To correct failed corners, a designer can run a sensitivity analysis or optimization on the failed corners, hoping that the other corners continue to pass.

3.2.2 Monte Carlo Analysis

Although the guess work in determining the worst case corners is removed with Monte Carlo analysis, the distributions of the parameters are still required. Whereas corners analysis produces a binary outcome (either the corner passes or fails), Monte Carlo analysis produces samples of a continuous function that can be used to estimate yield. In addition, Monte Carlo deals with the distributions of the process parameters among lots, wafers, and dies, and allows the designer to study the effect of parameter variations among devices on the same chip. With Monte Carlo analysis, hundreds of simulations are typically run with a variety of variations introduced randomly. However, the stdev (confidence interval) of a Monte Carlo yield estimate is proportional to 1/sqrt(n) where n is number of simulations. A Monte Carlo run of 100 to 200 simulations is very time consuming. Most of the time, it is not feasible to even run this many simulations, thereby introducing a sampling error. This means that one Monte Carlo run may generate a yield of 82 percent while another run may generate a yield of 85 percent, with no change to the design. Even worse, a designer may see an initial yield of 82 percent, tweak some parameters, rerun, and see a yield of 85 percent. At this point, the designer may not know if the increased yield was due to the change in the design parameter or because of the Monte Carlo sampling error. Another disadvantage of Monte Carlo is that all changes to design variables intended to improve yield will result in a complete rerun of the entire Monte Carlo analysis. So, although Monte Carlo eliminates the guesswork found in corners simulation, Monte Carlo simulations are very time consuming and do not run enough simulations to avoid sampling errors.

3.2.3 Response Surface Modeling

Response surface modeling (RSM) is the answer to the time-consuming sampling errors presented by Monte Carlo. RSM reduces the 1/sqrt(n) error from Monte Carlo and allows the user to reduce the sampling error to nearly zero. With RSM, a set of polynomial models are created from about 20 to 50 simulations (design of experiments) that approximate the original design. These models are run tens of thousands of times with the same random variations used during Monte Carlo analysis. Since a significantly larger number of simulations are run, the sampling error is reduced to nearly zero. With the design replaced by polynomial models, the designer can run the thousands of simulations much faster the

time is reduced to a matter of seconds as opposed to days. RSM is not, however, without its problems. The error introduced with this technique is clearly the difference between the polynomial models and the original design. Additionally, running the simulations on the original design (design of experiments) can be time consuming, because the number of simulations required is highly dependent on the number of statistical parameters (which is a major downfall compared to Monte Carlo analysis).

3.2.4 Issues with the Previous Process Analysis Tools

Despite their problematic aspects, Monte Carlo analysis and RSM are a big step forward from corners simulation in predicting parametric yield. However, most designers continue to use corners analysis most of the time. The main reasons are lack of statistical models (needed for Monte Carlo analysis and RSM), lack of tools to create the models for RSM, time constraints, and, most often, company policies. Beside the above issues, underlying the use of corner-based methods is the assumption that sufficient guard-bands or pessimism in the modeling and the timing analysis stages will 'guarantee' a working design. While over-design is an option, it ignores the competitive pressures to build a faster, lower-power and smaller silicon chip than another IC vendor, and essentially takes a new design back to an earlier generation silicon node. And finally, these pessimistic, over-constrained designs, surprisingly can still fail in production silicon because the outliers have not been analyzed. On top of that, all these analysis are done before the physical layout is implemented.

Layout-dependent variations, can cause two different layouts of the same device to have different characteristics even when the two instances are located close to each other. Layout-dependent variations are different from random ones because they are predictable and can be modeled as a function of deterministic factors such as layout structure and the surrounding topology. Process variation can be caused by many sources with each exhibiting a different dependence on layout and design parameters. One must identify, characterize, and model all sources of variation for better model-to-hardware correlation.

The proposed CAD solution presents an automated methodology to predict the actual impact of physical layout-dependent systematic manufacturing variations on the parametric behavior. Before presenting this methodology, the stress and lithography variation dependence on design context will be discussed first.

3.3 Design Context Impact on Process Variations

3.3.1 Design Context Impact on Lithography Variations

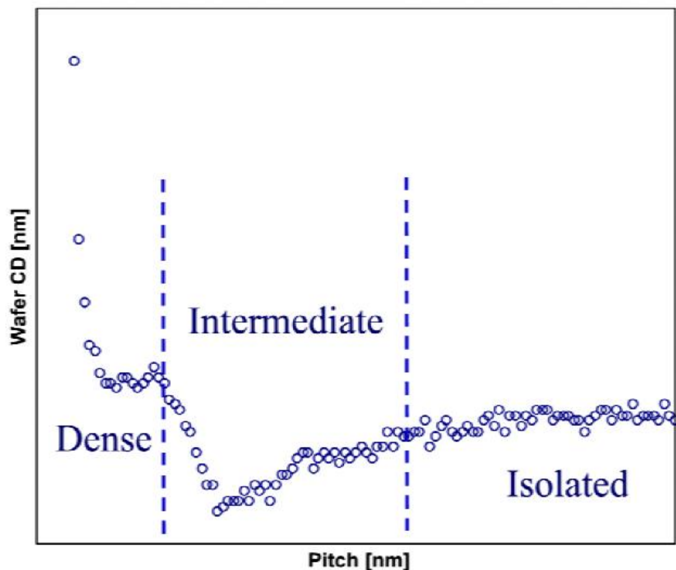


Figure 3.1: Experimental data for pitch curve with feature type classification. [11]

Fabrication of integrated circuitry typically involves a lithographic process. This process transfers a pattern which is disposed on a mask onto a layer of material such as photoresist received over substrate. Current photolithography tools operate with deep ultra-violet (DUV) light wavelengths of 193 nm. Optical light rays in the photolithography process diffract and pass through the reticle in different ways, based on the layout context and mask pattern density. The pattern on the mask generally defines integrated circuitry patterns and alignment patterns. It has been observed that differences in pattern development of circuit features can depend upon the proximity of the features relative to fone another. So-called "proximity effects" [79] [76] in a lithographic process can show up during imaging, resist pattern formation, and subsequent pattern transfer steps such as etching. The magnitude of the proximity effects depends on the proximity or closeness of the two features present on the masking pattern. Proximity effects result from the optical diffraction in the projection system used to form the pattern over the substrate. This diffraction and pass-through impact the final critical dimension (CD) printability on a wafer, which directly affects the circuit performance. The optical diffraction causes adja-

cent features to interact with one another in such a way as to produce pattern-dependent variations. These variations can affect the integrity of the final shape of the layout circuit devices. Figure 3.1, courtesy of [11], shows wafer CD measurements for same pattern but with different contexts.

3.3.2 Design Context Impact on Stress Variations

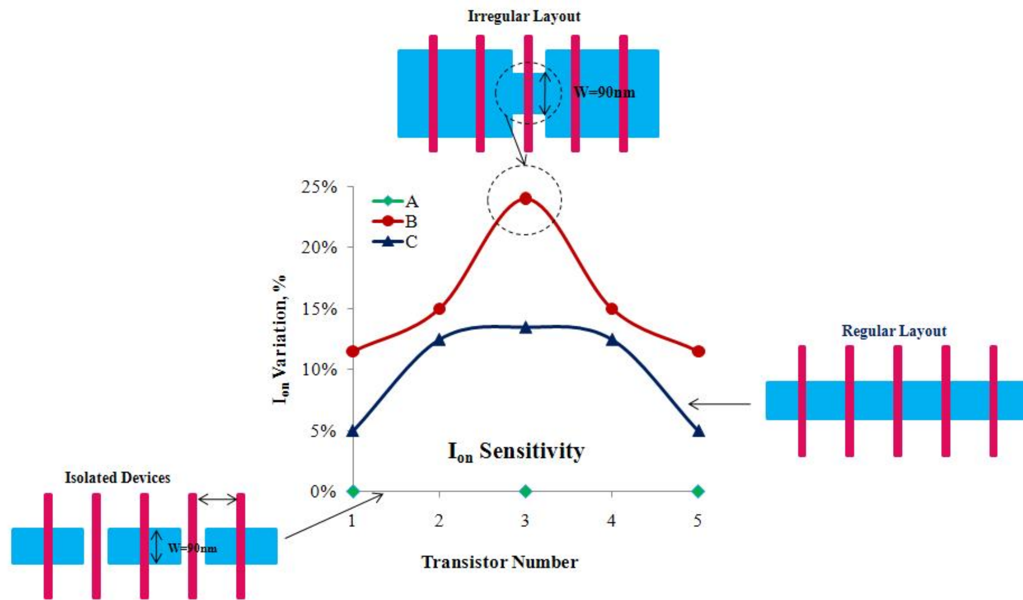


Figure 3.2: Stress effect on Ion for simple layout implementations.[12]

In 65nm IC technologies and above, the only stress effects considered are the Length Of Diffusion (LOD) effects. The LOD effects are currently included in the standard BSIM4 models as the gate to nearby diffusion spacing (SA) and gate to far diffusion edge (SB) parameters (average distance between the gate edge and the diffusion edge, as measured from left and right sides). Below 65nm, IC foundries have started to implement other stress parameters into the simulation models. These parameters are primarily used to account for the proximity of the other nearby transistors in the layout design. For example now other physical layout parameters need to be considered such as poly-to-poly spacing effect and STI-to-STI spacing effect [80]. Figure 3.2 illustrates the current variation for different layouts due to stress effects.

3.4 Characterization Flow: Layout Design Context Impact on Lithography and Stress Variations

Figure 3.3 illustrates the proposed methodology that characterizes the effect of the layout context on process variations and how that impacts the physical and electrical variations of standard cells [81]. The standard cells are placed in different contexts using the density based context generator engine, afterwards lithography and stress simulations are performed to identify lithography and stress hotspot layout topologies that are most sensitive to process variations. This is then followed by an analysis to come up with a relation between hotspots and their dependency on context effect. In addition timing analysis is also performed to come up with a relation between timing and context. These context-aware metrics will be used to guide front end, P&R engineers and back end designers to consider context impact during their design process.

3.4.1 Context Generator Engine

Performing silicon simulation and verification on standard cells will guarantee that the cell is lithographically clean within the proximity conditions used when doing the simulations. The proximity or the lithographical context surrounding the cell might change from one design to another according to the placement of the cell in that design. This will impact the lithographical behavior of the contours at nominal conditions and through process variations, which in turns suggests that the silicon simulation and verifications performed earlier for this cell might not still hold at these placement locations.

Therefore an automated script is implemented to generate the cell's surrounding proximity contexts with different densities. To generate a practical and predictable context, instead of generating randomly as used in [13], some user-defined parameters are given to this engine to design different context options.

The context generator then defines four regions around the chosen standard cell: top, bottom, right and left. These regions have enough width to cover the lithographical optical region. This is then followed by filling each region with the proper orientation and number of standard cells that meets the pre-defined density values inside each region. Figure 3.4 illustrate the pseudo code of the context generator engine.

The user-defined parameters are used in the script:

1. **Context**: It is the number of cells, placed in each region surrounds by the cell, parallel to the cell under test.

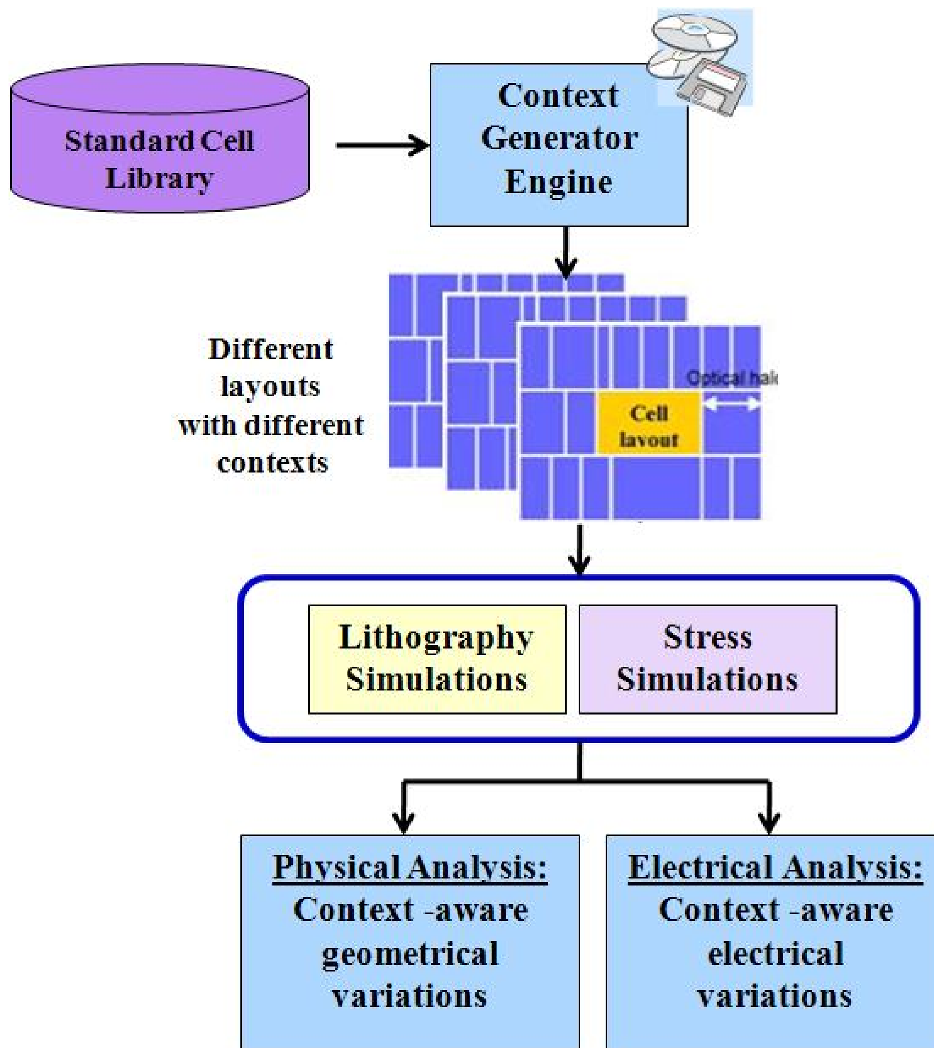


Figure 3.3: Flow used for Extracting Context-aware metrics

1) **Define user input data:**
   a. Name of cell under test
   b. Path of the standard cell library
   c. Optical diameter
   d. Number of adjacent cells
   e. Define name of gds layers
   f. Output database destination
   g. Context density range
   h. Surrounding cells separation distance from cell under test
2) **Start Placing Cell under test: Run PlaceCell module on cell undertest**
   i. Extract width (W) and height (H) of cell
   ii. Calculate the bottom left coordinates of the cell under test (X,Y)
      a) If angle =0; then place cell in (X,Y), with angle 0
      b) elseif angle =90; then place cell (X,Y), with angle 90
      c) elseif angle =270; then (X,Y), with angle 270
3) **Start Creating regions: Run CreateRegions module**
   i. Extract width (W) and height (H) of cell under test
   ii. Calculate the bottom left coordinates of the cell under test (X,Y)
   iii. Region 1 bottom left coordinates : (X+W +Separation, Y-Optical Diameter)
   iv. Region 1 upper right coordinates: (X+W +Optical Diameter, Y +H +Optical Diameter)
   v. Region 2 bottom left coordinates : (X- Optical Diameter, Y-Optical Diameter)
   vi. Region 2 upper right coordinates: (X-Separation, Y+H+Optical Diameter)
   vii. Region 3 bottom left coordinates : (X, Y-Optical Diameter)
   viii. Region 3 upper right coordinates: (X+W, Y-Separation)
   ix. Region 4 bottom left coordinates : (X, Y+H+ Separation)
   x. Region 4 upper right coordinates: (X+W, Y+H+Optical Diameter)
   xi. Generate layers for each regions
4) **Start Filling regions: Run FillRegions module**
   I. Fill region number 1
      i. Extract dimensions of region 1
      ii. Calculate the required cell width=
      iii. Region Width – {(Number of cells-1) *(Context separation)/ Number of cells }
      iv. Calculate the required cell height = Region Height
      v. Pick a cell from library
      vi. If the picked cell's dimension does not meet required cell and width ; then choose another cell
      vii. If the picked cell's dimension does meet required cell and width ; then add the cell to a "Good Dimension List" library
      viii. From each cell in the Good Dimension List library:
         a) Calculate the cell separations and number of cells to meet the required density range
         b) If the cells meet required density then add to "Good Density List" library
      ix. Pick first cell in the list
      x. Calculate coordinates values
      xi. Place cell-a in xa1,ya1,xa2,ya2
      xii. Calculate the adjacent coordinates values
      xiii. Place adjacent cell-b in xb1,yb1,xb2,yb2
      xiv. Repeat until filling region number 1
      xv. Report error message if there is no good cells combinations to fill in the region
   II. Repeat for region number 2
   III. Repeat for region number 3
   IV. Repeat for region number 4
5) **Save output layout in results database**

Figure 3.4: Pseudo Code for the Context Generator Engine

2. **Similar Context Cells**: It represents the number of different types of cells to fill each region.

3. **Optical diameter (OD)**: It denotes the max filling distance away from the cell under test.

4. **Context Separation**: It represents the separation distance between cell under test and the surrounding cells.

5. **Context Density** : It represents the summation of the areas of the cells, filling each region, to the total region area.

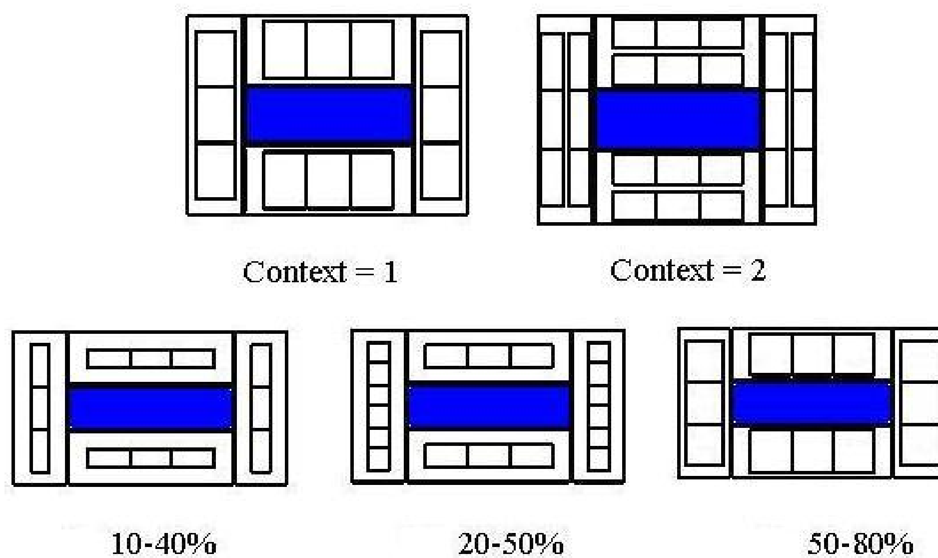


Figure 3.5: Different contexts surrounding an example of cell under test

Figure 3.5 denotes some snapshots of the dense context used in the case study. In the top left cell , shown in Figure 3.5, the user defines only one cell to be surrounded by the cell tested (Context =1), whereas the top right figure; two cells are used to surround the cell tested. The bottom figures show the cell tested in different context densities, i.e., using different (Context Density) values.

3.5 Experiments and Results

To verify the advantage of extracting context aware design netlists, lithography and stress variations are being characterized on different standard cells in 65nm and 45nm technologies. Context impact on the physical and electrical parameters has been studied on individual cells. Afterwards a 65nm five gate ring oscillator design constructed to study the effect of process variations on the oscillator's timing and power characteristics.

3.5.1 Layout Design Context Effect on Standard Cells' Geometrical Level

As a first step in our experiments the context generator is used to generate multiple layout contexts. As lithography variations are sensitive to the density of the same polygons, a steps of 5% increments in the *Context Density* parameter is used to generate different densities for poly and active layers. When studying stress variations, distances like active to active and also poly to poly are the main parameters that impact stress variations. Therefore the context generator is set to generate different locations of poly and active layers by using steps of 5% increments in the *Context Separation* parameter. Also a comparison study is performed between the stress variations that occurs when standard cells is characterized in isolation versus standard cells located in a design-like context. After generating these different contexts, lithography and stress simulations are performed.

In this work, the context variation has been characterized by defining: Length Safety Zone (LSZ) metric. LSZ is defined as the acceptable context density to keep the gate length under acceptable threshold of variation, depicted in Equation 3.1. The acceptable variation is the standard deviation value of the different gate lengths for the different occupancy rates.

$$LSZ = |L_{Density} - L_{Average}| < \sigma \tag{3.1}$$

where $L_{Density}$ is the gate length at a certain density, and $L_{Average}$ is the average gate length for all different context densities, and σ is the standard deviation of the gate lengths variations under all these contexts. Figure 3.6 represents the gate length variations in the inverter topology under different contexts. It is worth mentioning that the standard cells used in our test case had dummy poly similar to [82], confirming that the impact of manufacturing variations on devices is layout context dependent.

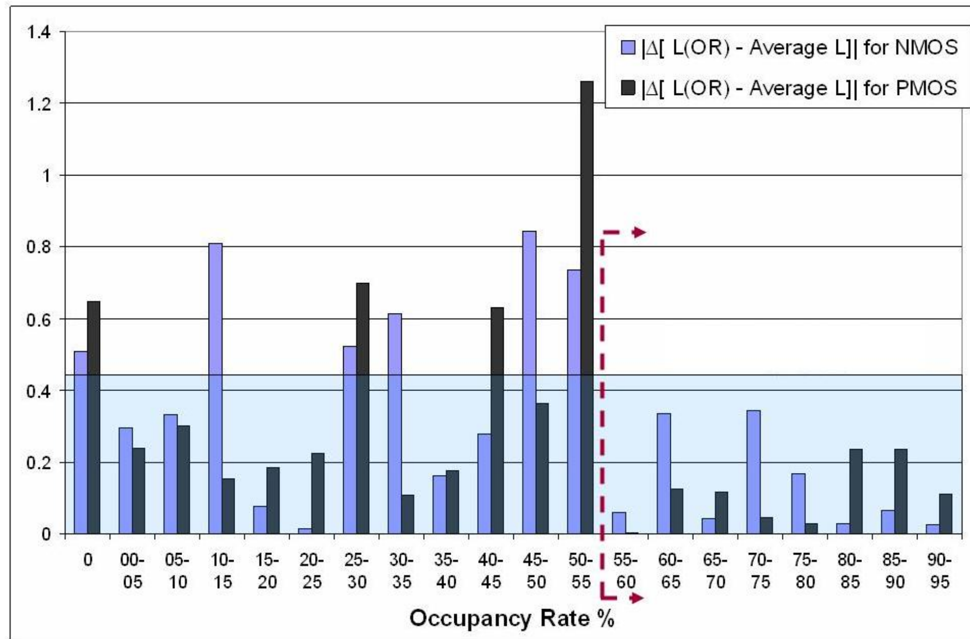


Figure 3.6: Gate length variation in different design contexts

3.5.2 Layout Design Context Effect on Standard Cells' Electrical Level

Prediction the impact of manufacturing variations on each unique instance of the standard cell without resorting to pre-characterization of that unique instantiation. Rather, in the context of the design, the timing behavior of each instance of a standard cell, given its input waveforms and wire load, is determined using a fast simulation of that cells transistors with device parameters appropriately modified to reflect manufacturing variations [13]. The design is traversed in a topological breadth-first order, starting from the primary inputs of the design where the input waveforms are known and given by the user. Each instance, as it is traversed, will have its input waveforms predetermined. The traversed instance is simulated with the cell transistors (using modified parameters) driving the wire load. The resulting delay across the cell instance and the waveform at its output are registered. The output waveforms are used to drive the next instance in the topological traversal while the delay is stored for subsequent timing check calculation. The change in timing due to manufacturing variation is calculated as a change in delay, between the cell with nominal device parameters and wire parasitics and the cell with modified device parameters and

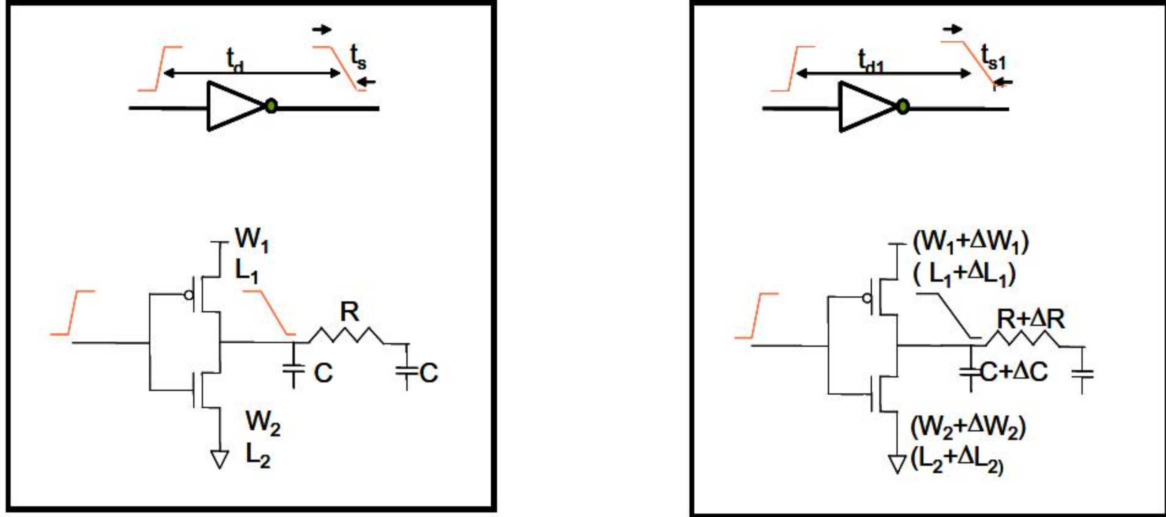


Figure 3.7: Standard cell with (a) nominal device parameters and wire parasitics exhibiting nominal delay, td and slew, ts and (b) modified device parameters and wire parasitics exhibiting modified delay, td1 and slew, ts1, due to manufacturing variations.[13]

wire parasitics. This is shown in Figure 3.7 where is calculated as:

$$\Delta\tau = t_{d1} - t_d \tag{3.2}$$

For chip-power analysis, change in off-state current drawn by that instance due to shape variation is also calculated. This change in leakage power due to manufacturing variation is calculated as a change in power, ΔP, between the cell with nominal device parameters and the cell with modified device parameters. This is shown in Fig 6 where ΔP is calculated as:

$$\Delta P = I_{off1} - I_{off2} \tag{3.3}$$

The leakage power for each instance of a cell is determined with and without shape distortions using SPICE simulation of that cells transistors with nominal device parameters and modified device parameters respectively.

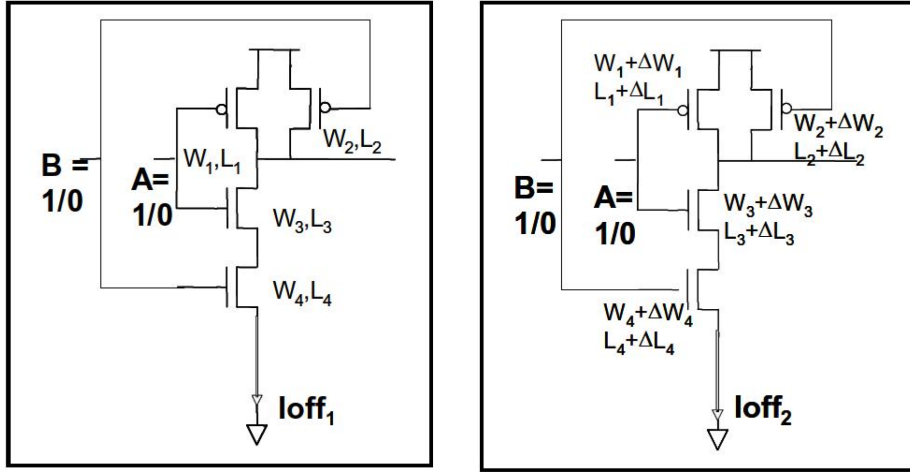


Figure 3.8: Standard cell with (a) nominal device parameters drawing leakage current, Ioff1 and (b) modified device considering manufacturing variations drawing leakage current, Ioff2. [13]

Lithography variations impact on Standard Cells' Electrical Level

After generating contexts around the cell under test, we run Calibre Litho Friendly Design (LFD) tool to create process contours for poly and active. Adjusted values for length and width of a transistor are calculated. Multiple styles of formulas are being proposed by EDA [83] . Some are based on creating an L/W value based on either I_{sat} or I_{off} [84]. Others are based on creating a new L_{adj} and W_{adj}.

Additionally, with more advanced tools, contours based on process variation can be created [85]. With the same technique used for nominal contour based extraction, a more thorough circuit analysis can be accomplished for process window. Performance based simulation of circuits using process variation information has become critical at the nanometer technologies. In our flow we used the Mentor Contour Simplified Gate (CSG) flow.

After running LFD and CSG, the netlist of the cell's under test netlist is extracted. The L_{drawn} and W_{drawn} are replaced by L_{adj} and W_{adj} for each transistor. Spice commands that are used to calculate timing and power values are automatically added in the extracted netlist. The proposed flow generates automatically a table that consists of the simulation results. Table 3.1 presents an example of the simulated results for a NAND gate.

Simulating the cell's timing performance through the high-to-low propagation delay (Tp_{HL}) and low-to-high propagation delay (Tp_{LH}). Figures 3.10 and 3.9 show that be-

Table 3.1: Example of the Electrical Simulation Results for NAND gate

Context Density%	Tplh (ns)	Tphl (ns)	P(0,0) (nW)	P(0,1) (nW)	P(1,0) (nW)	P(1,1) (nW)
Isolated	0.15265	0.13458	7.7515	3.77	4.515	0.638
Context 0-30%	0.15389	0.13548	7.6767	3.7333	4.472	0.632
Context 40-70%	0.148	0.132	7.43	3.61	4.36	0.634
Context 70-100%	0.14467	0.12822	7.7219	3.7518	4.5325	0.65676

havior of (Tp_{HL}) and (Tp_{LH}) are similar to Figure 3.6. As for typical inverters the relation between the high-to-low propagation delay (Tp_{HL}) and low-to-high propagation delay (Tp_{LH}) is defined by equations 3.4 and 3.5, respectively [86]. If the output load capacitance (C_L) is independent of the inverter's size (i.e., the inverter's intrinsic capacitance is neglected relative to the fanout and wiring capacitance, which was our case) then (Tp_{HL}) is inversely proportional to (W/L)n. Similarly, there is an inverse proportional relation between (Tp_{LH}) and (W/L)p , It is worth mentioning that the NMOS and PMOS gate width variations were also monitored with context variation, and compared to the standard deviation. However, the effect of context variation on the transistor gate widths was minor. This concludes that the variation of (Tp_{HL}) and (Tp_{LH}) are expected to follow the same pattern as gate lengths Ln and Lp respectively.

$$T_{pHL} = 0.52 * \frac{C_L}{(\frac{W_n}{L_n}) * K_n * V_{DSATn}} \tag{3.4}$$

$$T_{pLH} = 0.52 * \frac{C_L}{(\frac{W_p}{L_p}) * K_p * V_{DSATp}} \tag{3.5}$$

Stress variations impact on Standard Cells' Electrical Level

To investigate the effect of stress due to context on the behavior of the standard cell, a test case was developed using 45nm technology [87]. A standard cell is extracted then simulated without context first, then for the same standard cell, a context is generated around it then it is extracted using some modified extraction file (extract design only under a marker layer), and then simulated. After that results between the two cases are compared to find out how the stress due to context is really effective. Results from two cells will be presented; first, is a basic inverter representing a combinational cell, the other one is a flip-flop, representing a sequential cell .

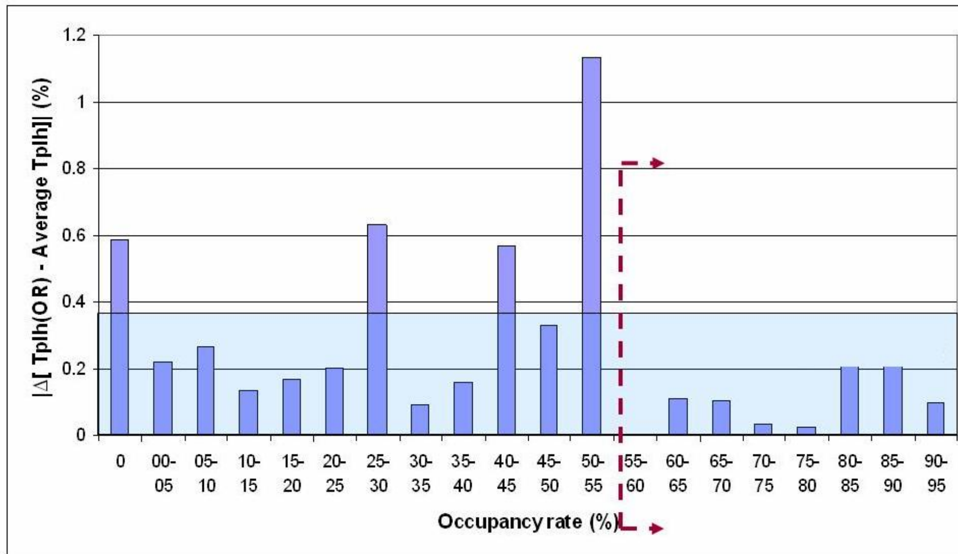


Figure 3.9: Low-to-high propagation delay (Tplh)

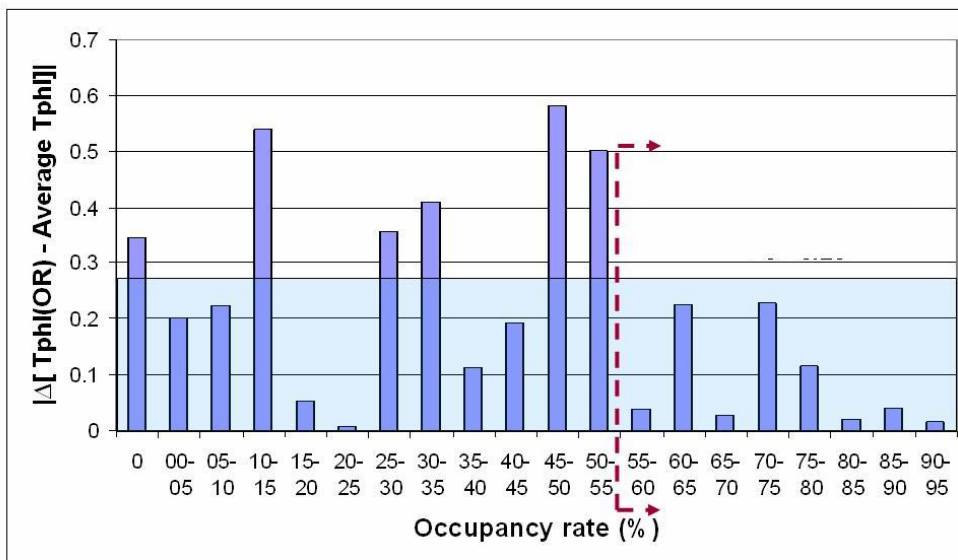


Figure 3.10: High-to-low propagation delay (Tphl)

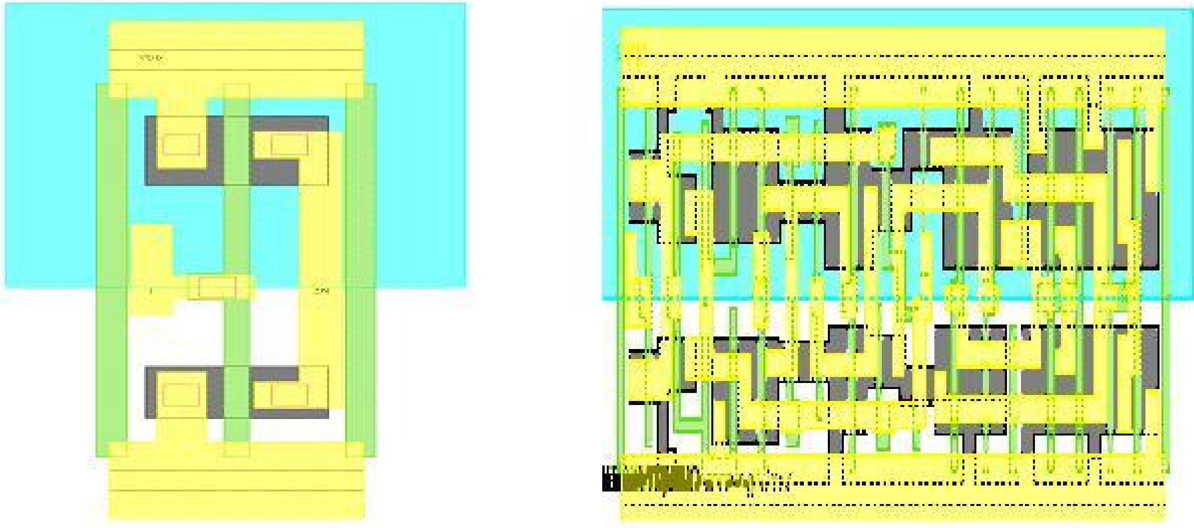


Figure 3.11: Inverter layout (Left) and flip flop layout (Right) without context

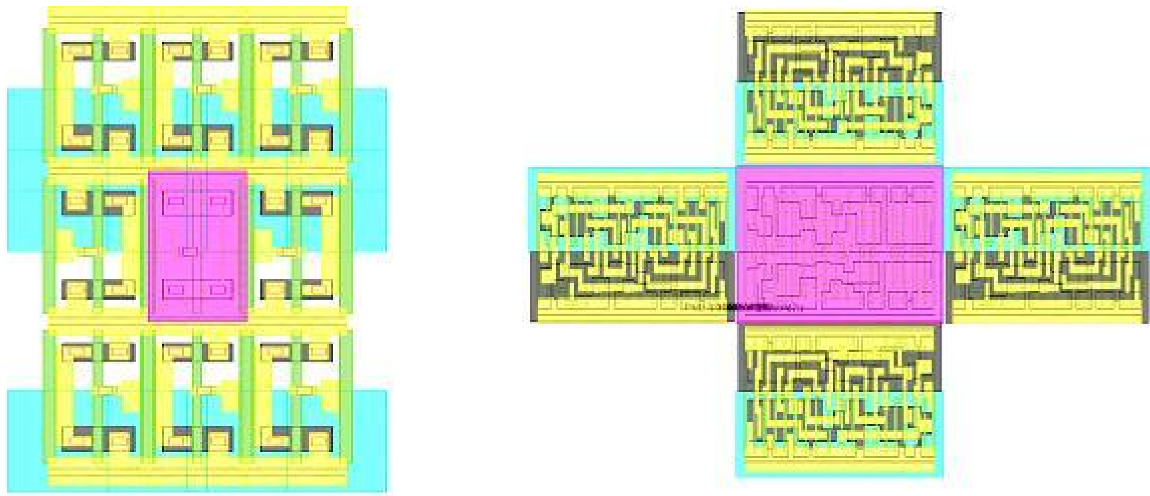


Figure 3.12: Inverter layout (Left) and flip flop layout (Right) within context

Figure 3.11, is the layout of the inverter and flip flop standard cells without context, while in Figure 3.12 shows the inverter and flip flop cells under context. The results that were obtained after the simulation of the inverter test case, show that just different context decreased the delay of the inverter by 3.5% in the rising edge, and 6.3% in the falling edge. This change is only due to context. For the flip-flop test case, the delay of the cell is decreased by 4.3% in the rising edge, where it decreased in the falling edge by 2.84%.

The above results show that only due to the design context the stress effect has changed the delay of the cell by values by percentages reaching up to 6%. This means that the stress in the cell will change according to the placement of the cell in different context, in consequence changes the performance of the same cell if they are in different context. This change expects to cause problems in the overall performance of the block.

3.5.3 Systematic Variations Impact on Ring Oscillator designed at 65nm

Questions can be formulated as to the pertinence of doing this type of analysis. First and foremost, is a time consuming analysis to adjust the length and width values of core transistors necessary? To answer this question it is possible to look at the effect of gate delay for a ring oscillator design. In [88] a five gate ring oscillator design using 90nm design rules was constructed (Figure 3.13). By creating a spice netlist for frequency analysis, and varying the L and W values in 1nm increments, a surface response curve can be created to determine the significance of this amount of variation for a ring oscillator (see Figure 3.14). Therefore a five gate ring oscillator is designed using the 65nm standard cell library. Analysis must be done with varying gate dimensions in silicon so that correlation of gate performance to the extraction software can be validated. This can also be done by designing ring oscillators with varying gate layouts to stress the variety of shapes that lithography effects can create. Using a common gate shape for each ring oscillator, one can then determine contour based prediction to silicon measurement for many varieties of active and poly gate shape combinations.

Spice model constants must be adjusted to accommodate the new style metric that is being used. This must be done in coordination with the extraction formula being used by the contour based extraction tool. This is also part of the correlation process to silicon. The standard constants for both the BSIM and PSP1 models must be adjusted to accommodate the fact that lithography and etch effects are being determined by an alternative tool. This can be accomplished by zeroing out the factors that control these offsets in the model.

For the case of this experiment, seven different process conditions to span the process

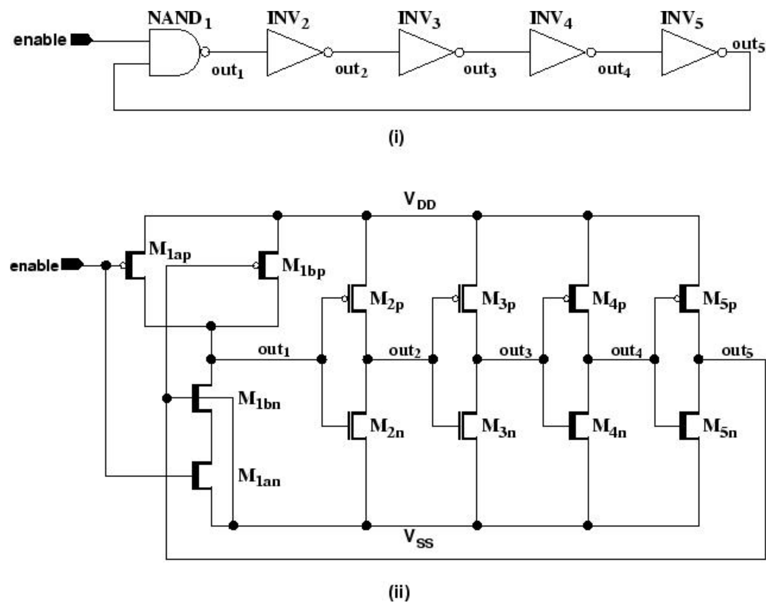


Figure 3.13: Schematic of a ring oscillator with enable pin: (i) Logic level. (ii) Transistor level. [14]

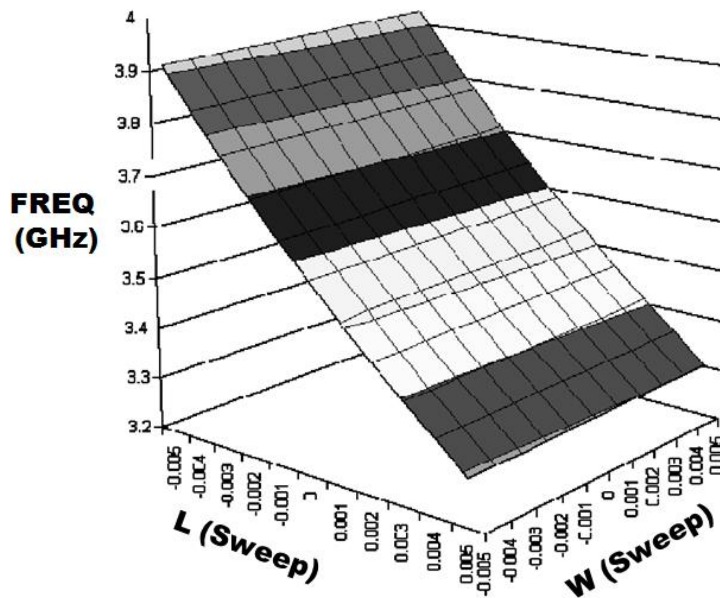


Figure 3.14: Frequency response curve for a 90nm five stage NAND gate ring oscillator.

window were used for both poly and active layers. Process mathematical models are used to generate process simulated silicon contours for the design. Since poly and active levels are independent (from manufacturing process point of view), 49 process variations for the gate level where generated based on seven active and seven poly contours. Calibre LFD tool was used to extract an adjusted gate dimensions L and W for each transistor in the oscillator. Devices were placed both in dense array layout and isolated to compare for contextual dependencies. Using the model based extraction software, 49 spice netlists are created to model each of the process variation combinations.

Using a SPICE simulator, a frequency and power analysis can be performed for each case. Testbench conditions were set to standard 1V 65nm conditions. By plotting power vs. frequency, one can observe both the variation from just doing drawn only analysis, and also see the process window for the circuit based on the 2-sigma models (Figure 3.15).

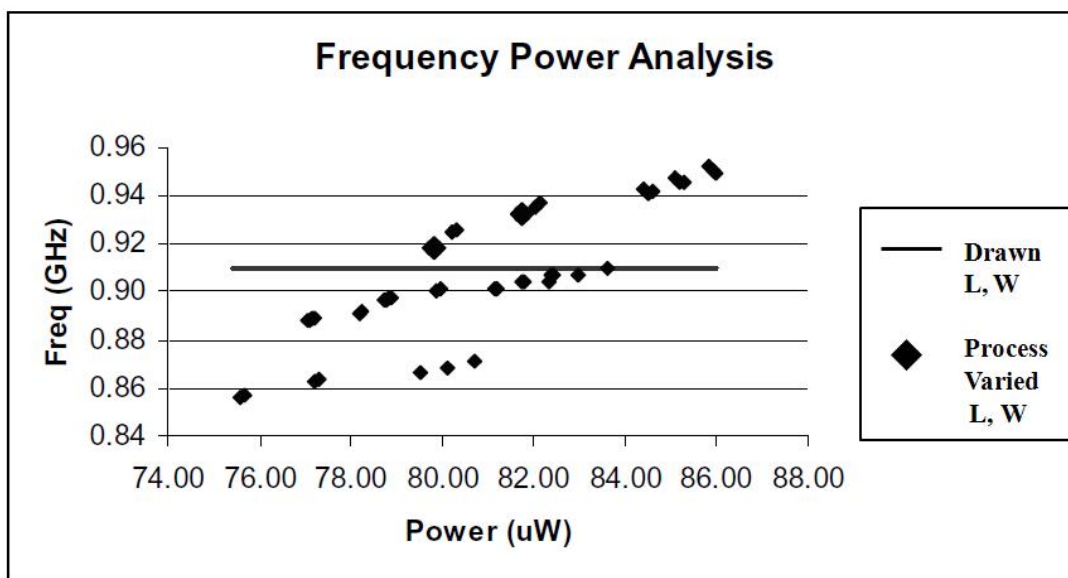


Figure 3.15: Frequency and Power analysis for the 65nm ring oscillator using drawn device parameters vs. process-aware device parameters.

In this experiment, power is varying 10% and frequency is varying 12% over process window. There is a large variation from the current methodology of just using the drawn layout to calculate performance. By plotting the surface response curve of process variation versus performance, the effect can be visualized. Figure 3.16 shows the effect of lithography focus and dose variation on the poly layer is plotted vs. frequency performance. Looking at

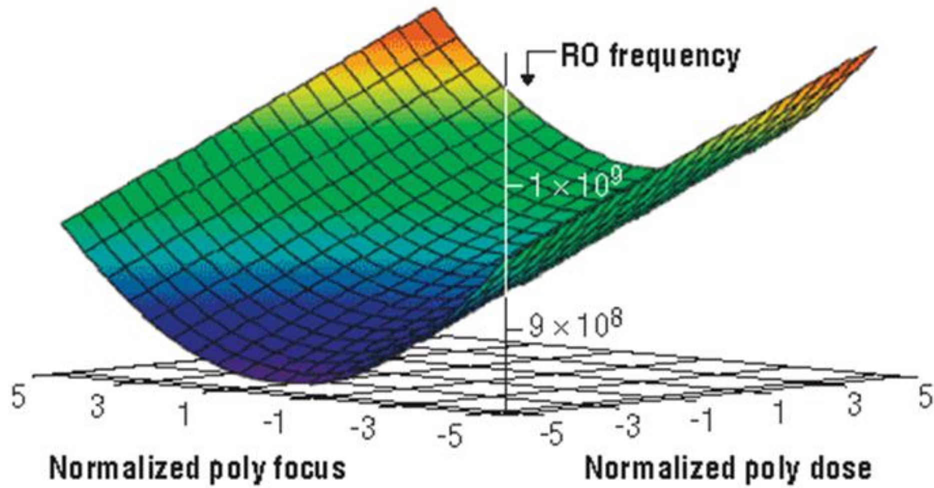


Figure 3.16: Frequency Response Curve for Poly Lithography Effects.

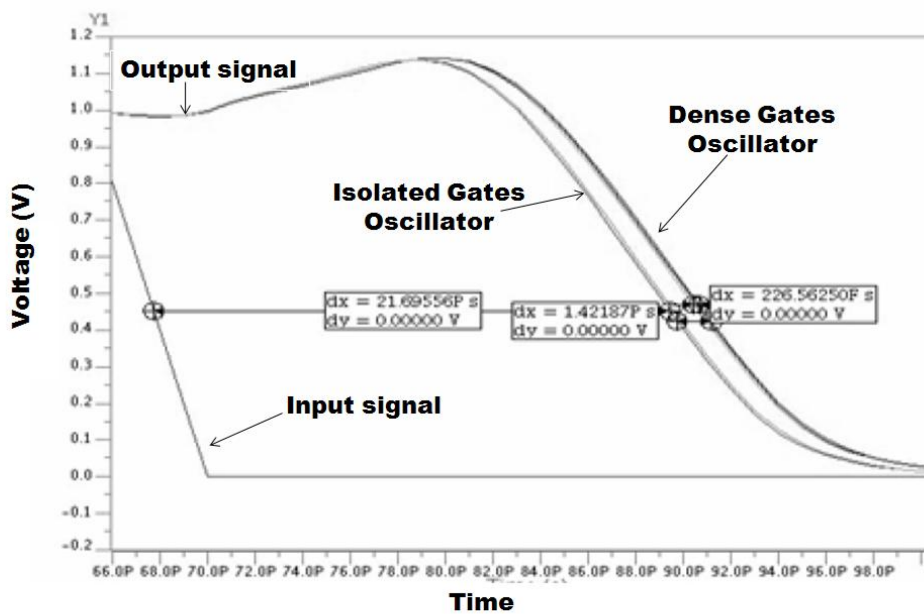


Figure 3.17: Gate Delay Analysis based on the effects of dense vs. isolated gate placement.

the amount of gate variation that is created at various process conditions vs. isolated and dense gate layout, one can calculate the timing differences in gate delay based on layout context dependency (Figure 3.17).

3.6 Summary

Lithography and stress variations are two dominant effects impacting the functionality and performance of designs at 65nm and below. In addition; proximity effects from neighboring cells, significantly influence the lithography process and stress variations values. Therefore studying the design context has to be considered in any variability-aware circuit analysis. In understanding the accuracy demands of nanometer design simulation, there is a need to accurately characterize the process variations. It has been proved that characterizing the standard cells in isolation and lack of awareness of the design context would lead to an inaccurate circuit analysis. This chapter proposed a methodology to characterize the design context impact on the process variations and hence on the cell's electrical parameters. The main contributions of the proposed methodology are listed as follows:

- **Design context awareness**: Context generator engine was implemented to simulate the design context effect.
- Study the effects of **different systematic variations** such as lithography and stress effects.
- A quantitative study of the impact of systematic variations for different circuits' **electrical and geometrical** behavior

The design context-aware process variations were examined on 65nm and 45nm industrial standard cell libraries and a 65nm five stage NAND ring oscillator. Results showed that power and frequency depends on the systematic process variations.

Chapter 4

Fix Before Design CAD Solution: DFM-aware Standard Cell Re-characterization

4.1 Introduction

In modern semiconductor industry, simulations of manufacturing processes are required to ensure circuit manufacturability. Fast and accurate lithography simulation is a key enabling technology [89] in the design-to-manufacturing flow. These computational lithography applications [90], [91] have recently received many interests [92], [93]. The major challenges in analyzing and enhancing digital designs for process variations, are the huge number of devices that highly impact the simulation runtime. In addition, most of the state-of-the-art design enhancement techniques are based on: **"Fixing during Design"** philosophy. As the simulations will predict the expected hotspot areas in the designs, afterwards these hotspot areas are sent back for physical layout modification or for DFM-aware placement and routing [94].

This work presents a **"Fix before Design"** philosophy through a software framework aimed at improving the manufacturability of a given standard cell library. Our proposed systematic variation-aware tool is a simulation-based flow, however there are techniques that are introduced to avoid the expensive full process window lithography simulation timing. In addition the flow incorporates some key features that make the re-characterization fully DFM-aware, by analyzing the gate length variations under:

- Different process conditions throughout the complete process window.
- Different design contexts using density based context variation.
- Different RET aggressiveness.

DFM metrics are extracted from this flow, such as the optimum placement locations of standard cells, which can guide the designer for a more robust and manufacture-aware placement and routing design. This flow is verified by re-designing the critical paths of 45nm Finite Impulse Response (FIR) filter circuit using the DFM-aware library.

4.2 Background

Recently, other solutions perform timing analysis to standard cells, such as [13] and [82], however the proposed fully automated methodology presented in this chapter, differs from the other work done such by the following:

1. As the lithographical simulations that include different process window conditions are crucial for verification of designs for technology nodes of deep sub-wavelength critical dimension, the runtime penalty of simulating at different process condition is large. Analytical methods that accounts for different lithographical simulation are limited to few process window conditions such as dose, focus and overlay shift. New technique is used to filter redundant combination of process conditions, pick specific conditions that need to be passed to electrical extraction, and mark placements of cells that will result in different lithographical simulation results. The suggested method named Isolated Process Point Engine (IPPE) can account for any changing parameter that need to be considered in the process window simulations, and provide a systematic way to filter redundant process conditions through contours comparisons. Moreover this analysis is done once per technology and no need to re-run this engine for each design. (IPPE) identifies the optimum and unique process conditions that represent the total process variation. This engine helps to reduces computation time by **avoiding the simulation of all possible process and lithography conditions and instead use optimized number of process conditions that is proven to represent the complete process window** while not compromising the quality of simulation.

2. The proposed flow examines the effect of different Resolution Enhancement Techniques (RET) recipes on timing and power analysis. Since lithographic contours are highly dependent on the RET techniques applied to a design [95], having the flexibility to examine different RET recipes is key for proper standard cell characterization. Besides considering different RET aggressiveness into the re-characterization flow, several new RET design concepts can be concluded from this feature as well, such as selective RET concept, where the RET engineer can define smartly the RET aggressiveness needed based on the type of the cell used and its location in the chip.

3. Several manufacturability aware metrics are proposed to be added to the standard cell library in a format that is understood by place and route tools. These metrics will better quantify the robustness of standard cell to layout context variations, and also guide the placement tool where would be the optimum location to place the standard cell. In addition a design-aware metric is also proposed for the RET engineers during their post layout operations aiding them to smartly selective the proper and optimum RET aggressiveness for such design.

4.3 DFM-aware standard cell re-characterization flow

Figure 4.1 illustrates the proposed flow for litho-aware standard cells characterization [96], while figure 4.2 shows the pseudo code used to implement the flow. The first step in this methodology,is placing the standard cells in different contexts using the density based context generator engine. Secondly, RET is applied on these cells, with different aggressiveness. In the third step, the lithography contours are simulated at a specific process condition, to identify layout topologies that are sensitive to this process condition, known as litho-hotspots. Lithography simulations are repeated at different process conditions to cover the complete lithography process variations. This step is done using Calibre™LFD for predicting the manufacturing variability through process variation. By accurately simulating the different effects of the lithographic process this enables designers to make trade-off decisions early, resulting in a design that is more robust and less sensitive to the lithographic process window. Calibre LFD use a built-in function known as Process Variability bands (PV-bands) [97] [98]. To generate the PV-bands, all different process corners conditions need to be fed to the simulator. However to simulate all expected process conditions, a major runtime issue will show up. To understand the issue, let us assume the variation in the lithography process is caused by variation in the dose of the light that varies, for example, from 90% to 110% from its nominal value. To cover this range, steps of 10 dose conditions are considered. Another source of lithography variation is the variation

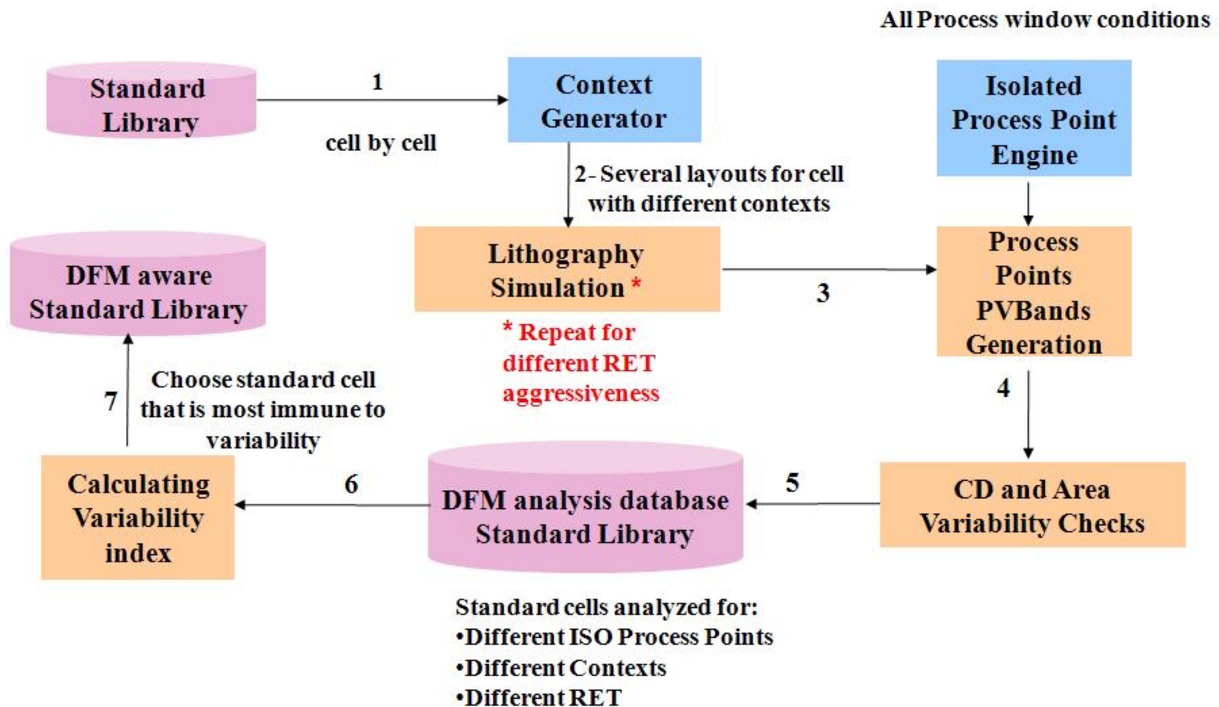


Figure 4.1: Proposed Flow for Implementing DFM-Aware Standard Cell Library

in the wafer position which can get out of focus +/- 100nm from its best focus, studied by steps of 10 conditions as well. This means 100 different cases for optical lithography dose and focus variations. These 100 simulations runs need to be repeated for at least 3 different RET aggressiveness. In addition, one set of simulations need to be applied for the poly layer and another set for active layer. To study the design context effect at least 10 different layout contexts are used for each standard cell. All these simulations need to be repeated for each standard cell in the library, which can result into nearly 2 million runs assuming we have an average of 300 std cells. However to overcome this issue, the Isolated Process Point engine, which will be described in more details in subsection 4.3.1, is used to supply the Calibre LFD tool with the optimum process conditions that accurately represents the complete set of process conditions and reduce the number of required simulations by nearly 90%.

The fourth step in the proposed flow, is to analyze and check the variations induced in each standard cell under different conditions of RET aggressiveness, process conditions and context densities. A database is generated for each standard cell simulating the vari-

1) Setting Environment Variables: Setenv function
- a)Defining the path of input/output database: Library, Cell to be tested, Results
- b)Defining the user-defined parameters for Context generator
- c)Defining degrees of RET aggressiveness
- d)Defining litho variation steps
- e)Defining ELDO settings
- f)Defining output tables/excel sheet settings

2) Creating database directories Mkdir function for temp output, for each cell there is directories for each layout context and under that there are directories for each RET aggressiveness where final results will be stored: Cell>Layout Context> RET

3) Run Context Generator Context engine - refer to Context Generator Pseudo code

4) Loop function: for each cell in the library
- I. **Loop function:** for layout context in each cell
  - A. **Loop function:** for each RET degree:
    - i. Call Calibre to run litho simulation on process conditions output from IPPE engine
    - ii. Copy outputs to results database
    - iii. Repeat on another RET degree
  - B. Repeat for another layout context
- II. Repeat for another cell

5) Call Calibre LVS: to run LVS netlist extraction on each silicon simulated layout (L,W after litho simulation) (SA,SB for stress)

6) Call Calibre LVS: to run LVS netlist extraction on original layout – with no litho simulation (Original L, W and stressvalues)

7) Call ELDO: to run SPICE simulation on each extracted netlist: per RET per layout context per cell

8) Generate tables for each standard cell: CD variation, Timing variation for each context and each RET.

Figure 4.2: Pseudo Code for the DFM-Aware Standard Cell Library Engine

ation effect under different conditions. Afterwards DFM-aware metrics is calculated to indicate the best conditions were standard cell which will be mostly immune towards process variations. Finally this standard cell, with its DFM metrics, will be updated into the DFM-aware library.

It is important not to forget that the above standard cells' electrical characteristics are highly dependent on the resolution enhancement techniques (RETs) used in the analysis. Therefore in our flow, we demonstrate the impact of different RET settings on the standard cells' behavior.

In the following subsection the Isolated Process Point engine will be discussed in detail.

4.3.1 Isolated Process Point Engine (IPPE)

Technology nodes that have critical dimensions of sub-wavelengths have shown limitations of purely geometric DRC verification techniques, and many cases have been reported where through process lithographical simulations have shown inadequate design structures that were not captured by DRC [99]. Given the fact that different process window conditions

should be taken into account while performing the layout verification, many questions come into the picture. What are the parameters that will be changing during the simulations? What is the range of values to be considered for each parameter? Taking into account the extremes for each range, will this be sufficient? Or do we need to take into account intermediate conditions? If we plan to do electrical extractions from the litho contours, how to pick the conditions for extraction? What is the runtime penalty? Would process window simulations be done once at the standard cell level? Or would we still need full chip verification?

The idea behind the IPPE is to reduce the number of process window conditions that need to be considered in a process window simulation while not sacrificing the accuracy of the simulation. This can significantly cut computation time. This process is done once per technology. Examples of process window conditions include the lithographic image focus, exposure dose, mask overlay shift, mask size, or any other parameters that might appeal to the user. IPPE picks one process condition from a set of process window conditions that result in similar contours. Isolated Process Point Engine takes as an input all the lithographic parameters that need to be considered in the process window simulation; in our experiments lithographic image focus, exposure dose are used. The practical range for each of these parameters is provided as well as the step size. Afterwards, the lithographic contours from all combinations of parameters are generated. Each contour resulting from a specific process condition is checked for similarity against all the other generated contours.

Figure 4.3 illustrates the pseudo code for the IPPE engine. The basic approach when comparing different polygons using a DRC tool is to perform an XOR operation. This is the first step used in the suggested engine. Later steps will use the results from this XOR operation and perform another two operations in order to decide if these contours are similar or not. These operations are called, macro window sample and the micro window sample respectively.

Macro Window Sample

The area of the XOR results will give an indication of how similar these two contours are, if the two contours are identical then the area of the XOR result should be equal to zero, higher values for the area consequently means bigger difference between the contours. The absolute value of the area by itself is not enough to judge whether the contours are similar or not. Normalization of this value to the value of the layout polygon producing these contours would be more meaningful. For example if the normalized result is 50% this means the contours are very different, while values like 0.02% means they are very similar, and so on.

```
1) Setting Environment Variables:
    a) Define the different ranges and steps of the process conditions. p1 to pn
    b) Define the GDS layout path
2) Loop function: For each process condition:
    I.  Call Calibre LFD: run lithography simulations on the complete layout
    II. Save the results in database
3) Loop function: For each process condition: sweep x= p1 to pn: start x=p1
    I.  Loop function: for each sweep y= p1 to pn: start y=p1
        I.  Run p1 XOR p2
            i.   Calculate Area (p1 XOR p2)
            ii.  Calculate Area of (Polygon)
            iii. If [Area (p1 XOR p2)/Area of (Polygon)] <= Threshold1
            iv.  Then Calculate Area (Square)
            v.   If [Area (p1 XOR p2)/Area of (Square)] <= Threshold2
            vi.  Then add to Database1: p1, p2
            vii. Else do nothing
        II. Loop on y: p1 XOR p3. Finish loop
    II. Loop on x. Finish loop
4) Output database: List of process conditions in each database
```

Figure 4.3: Pseudo Code for the IPPE engine

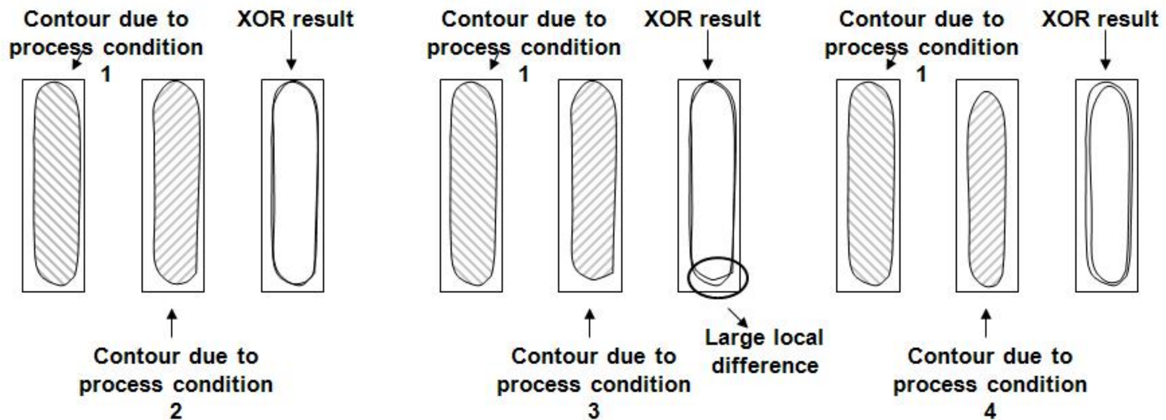


Figure 4.4: Different cases for contour comparison results using the big window knob operation

This operation will catch the general similarity between the two contours under investigation, however some case might be encountered where the value of the macro window sample is small, but there are some high local area differences. To deal with local differences the micro window sample operation is used.

Figure 4.4 shows three different cases for XOR of different contours. Case 1 is the left most picture where the contours are nearly identical and the XOR area is very small, Case 2 is the middle picture that shows low value for the macro window sample but there exists a local high difference between the contours that is not captured by this operation, Case 3 is the right most picture that shows high value for the macro window sample and two different contours.

However, the macro window operation by itself is not sufficient to decide if the contours are similar or not. There can be a situation where two contours have only one large local difference, but not large enough to exceed the pre-defined threshold, causing the contours to be treated similar. Therefore the macro window operation is followed by a micro window operation.

Micro Window Sample

The micro sample operation is a scanning operation. A small square of few nano meters side length is swept over the XOR result from the two contours under test, the area of the XOR result inside the square is reported for different locations as the scanning progresses. If there is a location where the value of the area of the XOR result is larger than a given threshold, then this means there is a local difference at this spot, and the contours can not be considered similar. Figure 4.5 shows the small scanning window, the window runs over the XOR results in all directions, the area of the XOR result is checked every time the square steps in positions.

The micro window operation by itself is not sufficient to decide if the contours are similar or not. There can be a situation where the area inside the scanning square is always less than threshold but when all areas are summed, the overall difference is large enough to consider the contours different.

The complete flow for comparing the similarity between to process conditions is shown in figure 4.6.

Having a systematic methodology for deciding if two contours are similar or not, we can reduce the number of process window conditions that need to be considered in a process window simulation by picking one process window condition from a set of process window

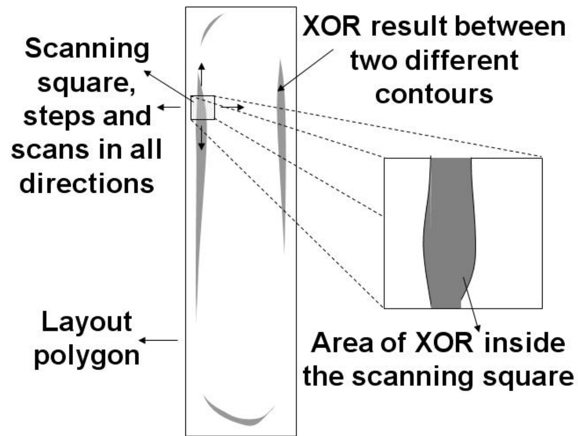


Figure 4.5: Micro window scanning square

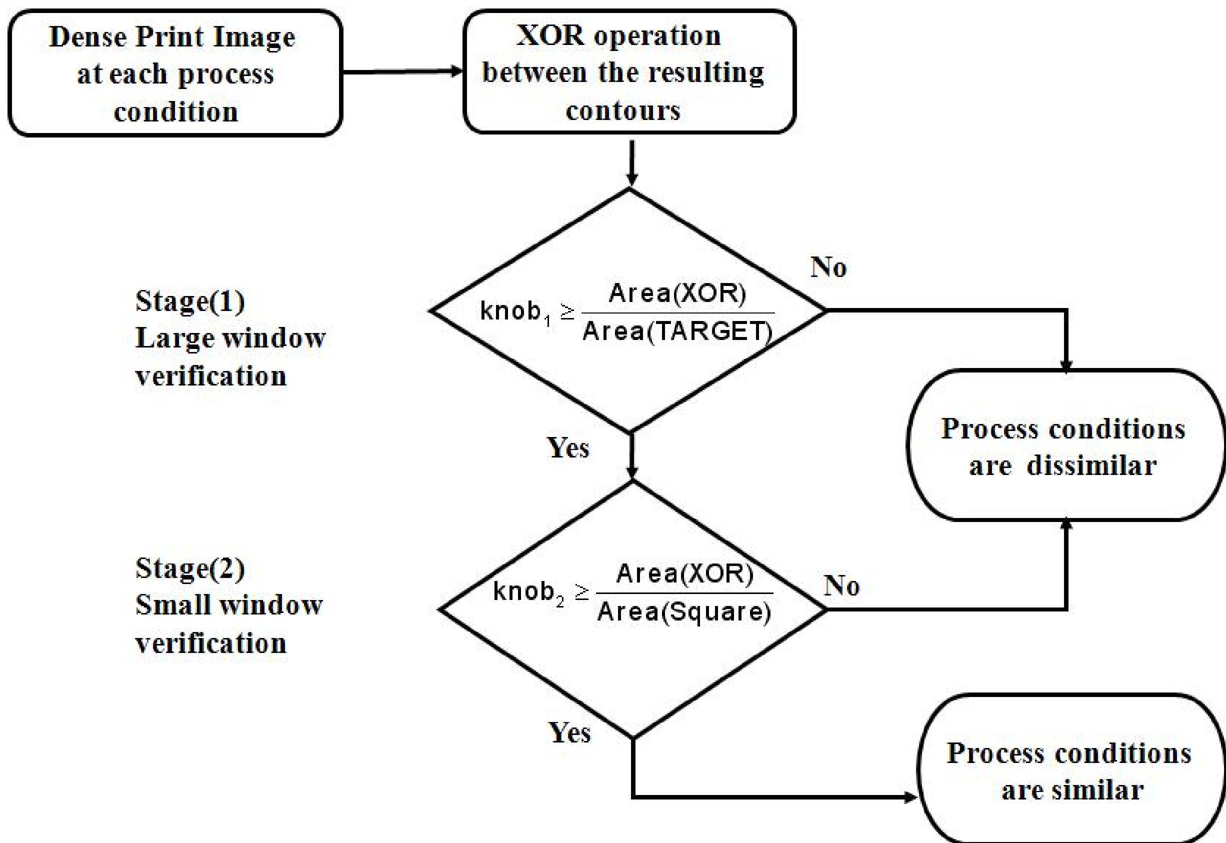


Figure 4.6: Flow for comparison of similarity between two process window conditions

conditions that result in similar contours. The IPPE takes as an input all the parameters that need to be considered in the process window simulation, example: focus, dose, overlay shift, size, and any other parameters that might appeal to the user of this application. The range and the grid of sampling for each parameter is given to the tool, contours from all different combinations of parameters are generated. Each contour resulting from a specific process condition is checked for similarity against all the other generated contours. Comparisons are made using the macro and micro window operations, and finally one contour is picked out of each similar group of contours.

Figure 4.7 shows quantitatively the input/output of the IPPE. From a matrix of (m*n) combinations, where m and n are the number of steps that represents the image focus, and exposure dose ranges the IPPE was able to define only 3 combinations that optimally can represent the (m*n) combinations. A thorough simulation for the input matrix can take considerable runtime, but the merit here is it is only needed once. The reduced output matrix can then be used whenever process window simulations are needed.

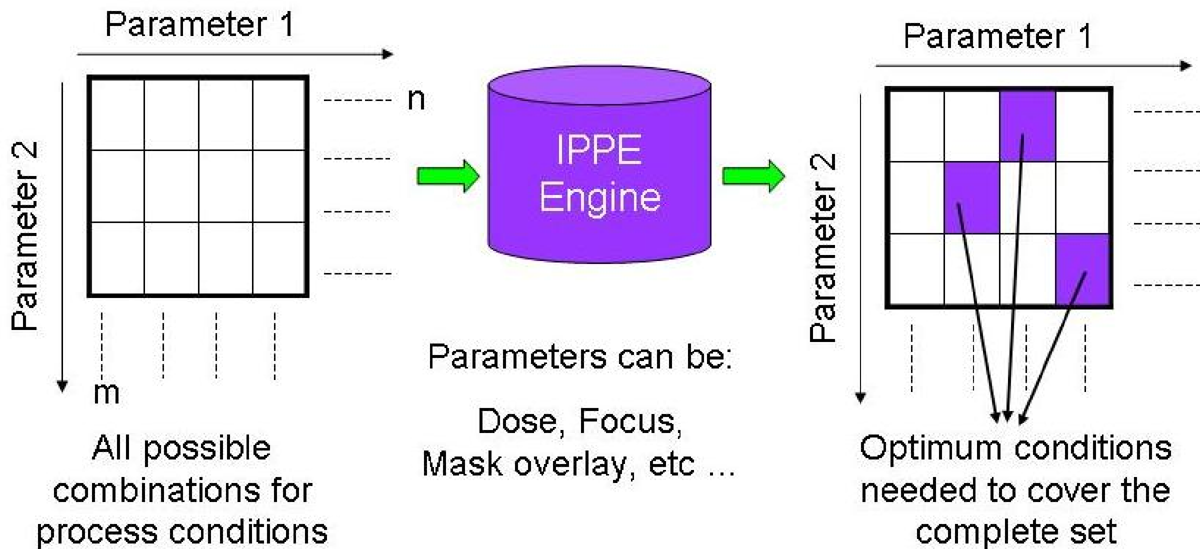


Figure 4.7: Inputs and outputs to and from IPPE engine

The picked contours are the ones really needed to describe the process variations, any other contour will have a contour representing it in this set. Electrical properties extractions through process window can rely on this set of selected contours only. There will be no need to consider every other condition, this makes circuit simulations through process

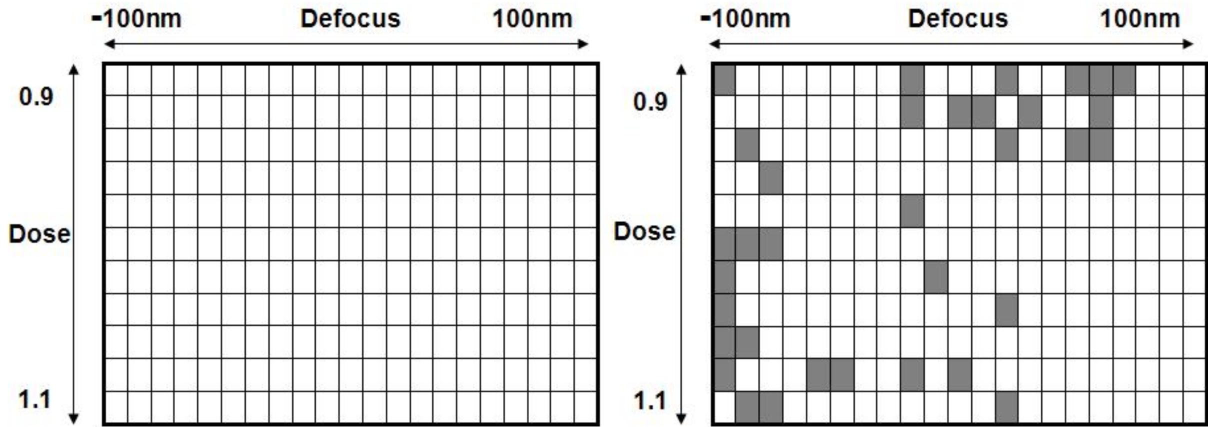


Figure 4.8: Left: Matrix of input conditions. Right: Selected process conditions shaded in grey

window more feasible and less time consuming concerning the runtime.

A standard cell of 65nm and 45nm poly technologies are used as the test vehicle for the tool, only focus and dose parameters are considered as the process variable parameters, a total of 21 focus conditions spanning the range from -100 nm to +100nm defocus, and 11 dose conditions spanning the range from 0.9 to 1.1 dose values normalized to the nominal dose value. The input matrix to IPPE is a 21 x 11 matrix of 231 elements, after running IPPE only 34 conditions are picked. Figure 4.8 shows the input and output matrix of conditions.

The process variability bands due to the input matrix are calculated, and also the process variability bands due to the output matrix of selected process conditions are simulated. Figure 4.9 shows nearly identical bands. This means that we can safely use the reduced matrix without losing any of the contours that might occur during the process variations.

Previous studies [99], [100], [101], [102], show that some analytical techniques can help to minimize the process window conditions used in lithographical simulations, these methods usually are constrained to dose, focus and overlay shift. However, this method relies directly on the generated contours from each process condition, any parameter can be tested for variations. The input matrix to IPPE can be multi-dimensional not limited to two or three, but can be more according to the users needs. Accuracy of process window models through different process window conditions must be verified in order to get accurate results from the IPPE tool [103] .

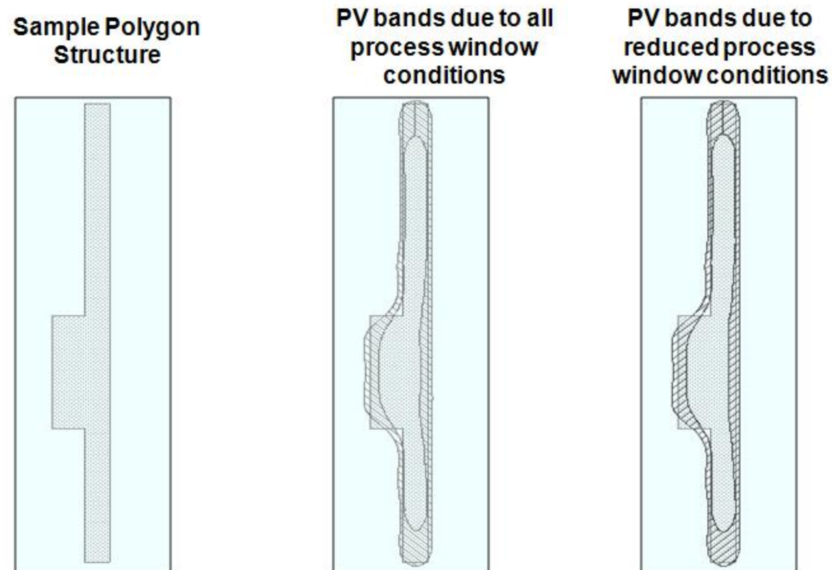


Figure 4.9: Left: Test polygon. Middle: PV bands of input matrix. Right: PV bands of output reduced matrix

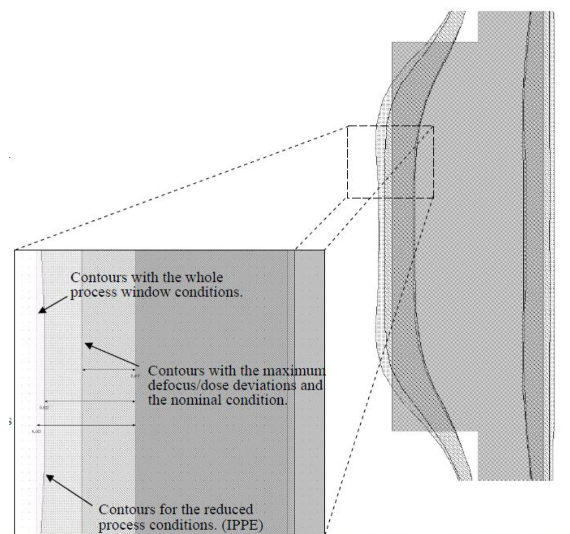


Figure 4.10: Comaprison between bands from process window corners and reduced conditions to the bands obtained from the input matrix

A thorough simulation for the input matrix can take considerable runtime, but the merit here is it is only needed once. The reduced output matrix can then be used whenever process window simulations are needed.

Comparisons between process variability bands obtained from extreme conditions of the process and the bands from the reduced conditions to the original bands from all process conditions are shown in figure 4.10, it is seen that the extreme values for process window does not always result in the worst contours, some intermediate conditions can cause contours that are worst than contours from extreme conditions. This proves the value of using reduction techniques rather than considering the process corners, because this captures the actual limits for the contours through process more accurately.

4.4 Experiments and Results

4.4.1 Standard cell library re-characterization

In order to verify the advantage of the proposed flow, two standard cell libraries at 65nm and 45nm technologies, are re-characterized. The variations in each standard cell are examined and monitored. As a first step, the context generator is set to generate multiple surrounding proximity contexts using steps of 5% increments in the *context density*.

Major knobs in the RET settings impacts the RET aggressiveness and hence its accuracy. Optical Proximity Correction (OPC), one of the RET techniques, is often run on the entire chip at once. The OPC algorithm involves subdividing polygons into smaller shapes or edge segments (fragmentation), moving or adding to the shapes, performing a fast simulation to determine if the new locations are better, moving them somewhere else, and iteratively repeating this process. Simulation-based OPC is complex and involves simulation of various process effects, which may be accomplished by computing a weighted sum of pre-simulated results for simple edges and corners that are stored in a library. In our example a set of seven RET recipes are generated with different aggressiveness, i.e. different set of fragmentation and iterations. One set for *active* layer and another set for *poly* layer. Afterwards the seven RET recipes of each layer are applied on a layout that contains the complete standard cell library. For each RET run, the Edge Placement Errors (EPEs) for all fragments in this layout is calculated and illustrated in a Gaussian-like histogram (Figure 4.11). From each of these histograms the mean, standard deviation and percentage number of fragments that have zero EPEs (peak %) are extracted. To categorize the RET recipes from less aggressive (weak) up to most aggressive, the following paramters are checked:

- Mean value: The most optimized recipe must have smallest mean value, i.e. closest to zero nm EPE.
- Standard deviation: The most optimized recipe must have smallest standard deviation value.
- Peak %: The most optimized recipe must have the largest peak percentage value.

Three selected RET recipes as weak, moderate and aggressive RET recipes. These three recipes are used in our analysis.

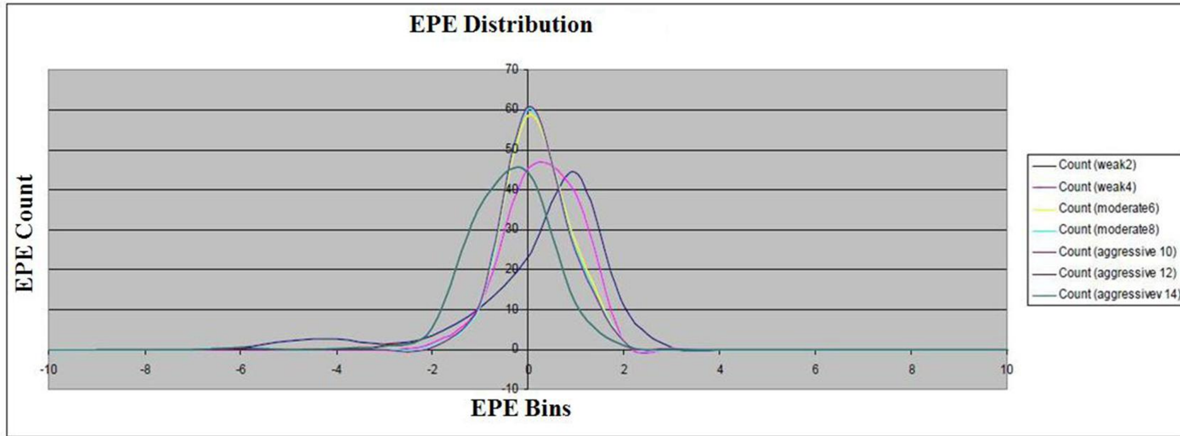


Figure 4.11: EPE histogram for different RET aggressiveness

In this work, the context variation has been characterized by monitoring two built in metrics in Calibre^{TM}tool, Design Variability Index (DVI) representing the design sensitivity and Process Variability Index (PVI) representing the process sensitivity [97]. The DVI score indicates how likely the variations in the simulated CD on wafer will negatively impact yield. DVI represents proportion of the area under investigation that is problematic (the total area on the error layers divided by the total area). This index is best used by design teams to locate and assign priority to errors for correction.

$$DVI = \frac{AREA(LFDerrors)}{AREA(window)} \tag{4.1}$$

While the PVI score indicates the degree to which printing is impacted by changes in the process. It represents the ratio of PV-band layer data to target layer data. When you

perform multiple process variation experiments on the same layer, all the experiments are factored into the scores.

$$PVI = \frac{AREA(pvband(layer))}{AREA(layer)} \tag{4.2}$$

After calculating the CD Variations, PVI and DVI values for a standard cell under different contexts and RET aggressiveness, two proposed metrics are extracted: Context Safety Zone (CSZ) and Context Robustness (CR). CSZ is defined as the maximum acceptable context variation, depicted and described as follows:

$$CSZ = \text{ Density at the maximum of } (PVI_{minimum} + \sigma_{PVI}) \text{ or } (DVI_{minimum} + \sigma_{DVI})$$

where σ is the standard deviation of the process and design variations under all possible context densities.

The context safety zone is a useful metric that can guide the way a circuit is synthesized and/or place and routed. At the synthesis phase, a library of standard cells is placed in context that is considered immune to variability. Figures 4.12 and 4.13 represents the process variations in the AOI topology under different contexts. In this example, the PVI standard deviation under all possible context densities was equal to 2.1 and for DVI standard deviation equal to 0.75. It can be seen, in Figures 4.12 and 4.13, that this AOI topology will experience high process variations on the transistors dimensions when used in less dense areas <55%, while this AOI will be nearly immune to process variations when its context has an Context Density >55%. It is worth mentioning that the standard cells used in our test case had dummy poly similar to what was recommended in [82], confirming that the impact of manufacturing variations on devices is layout context dependent.

Another metric we defined in our methodology is the Context Robustness (CR). CR is a statistical parameter that represents the robustness of a cell to the context variation (defined using equation 4.3) . This metric is useful to give the designer an indication how the standard cells are immune to context variation under different RETs conditions.

$$CR = \frac{(Mean - CV)}{Mean} \tag{4.3}$$

where the mean is the average value of DVI through context variations and the coefficient of variation (CV) is defined as the ratio of the standard deviation to the mean . CR

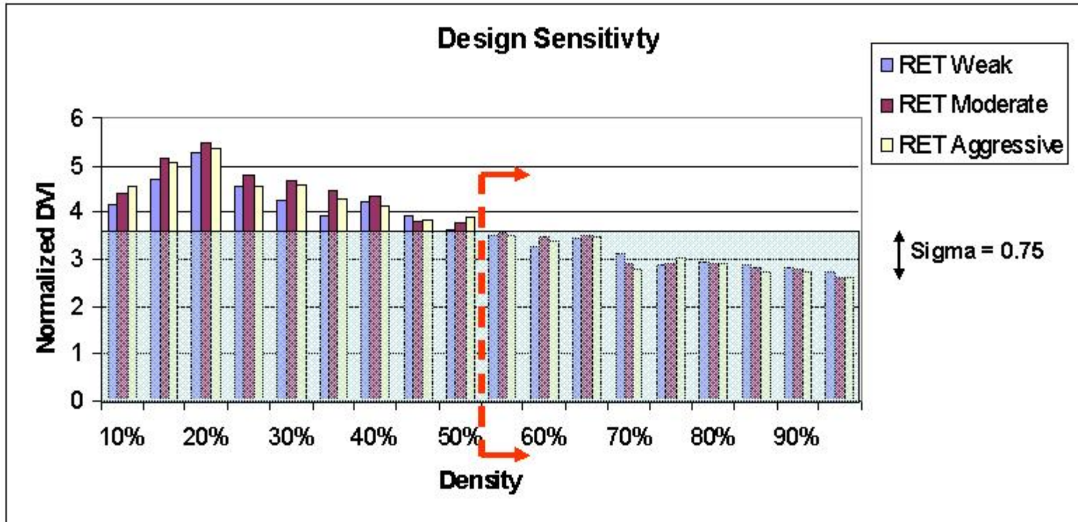


Figure 4.12: Design Sensitivity for AOI cell under different RET aggressiveness

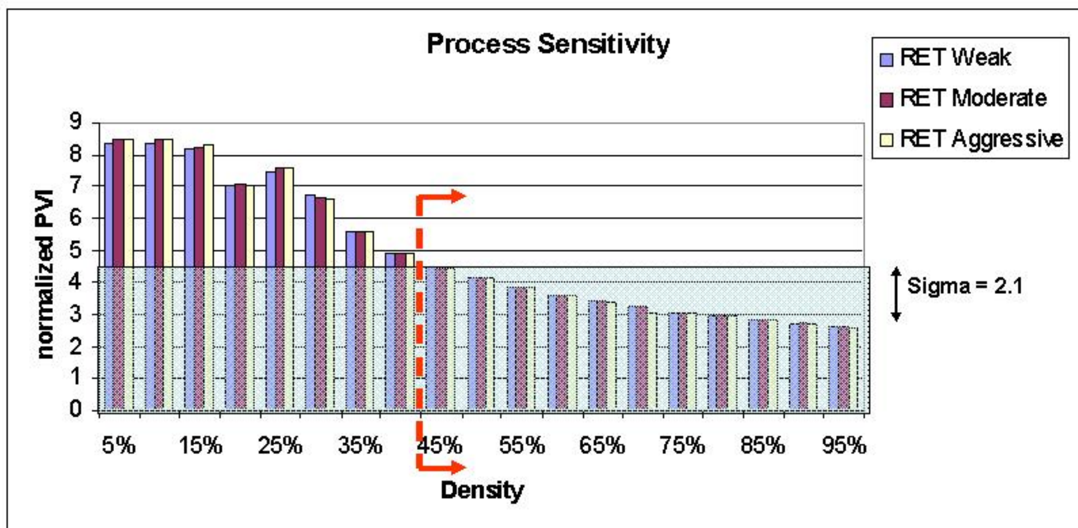


Figure 4.13: Process Sensitivity for AOI cell under different RET aggressiveness

Table 4.1: DFM-properties for different Standard cells

	CSZ	CR	RET agg
AOI211X4MTH	$> 55\%$	79.3%	mod
DFFNRHX1MTH	$> 80\%$	92.2%	mod
CLKINVX1MTH	$> 45\%$	90.1%	weak
NAND3X2MTH	$> 50\%$	93.4%	high

is checked for different RET aggressiveness. For the AOI example, the CR is 79.3% robust against context variation.

RET Aggressiveness Guidance is a third metric that can give the RET designers a clue on the RET aggressiveness needed on the different standard cells placed in the design so the fabricated design meets layout target. Based on the total Edge Placement Errors (EPEs) resulted from each RET runs, in addition checking which RET that gives the most context robustness (CR); we proposed the proper aggressiveness needed for each standard cell. It was concluded that moderate RET could be good enough to have an AOI standard cell that is immune to process variations.

Finally the conclusions mentioned above and the extracted metrics are attached to the AOI cell while developing the DFM-aware library. This work has been extended to cover other standard cells. Table 4.1 summarizes the DFM properties for randomly selected standard cells such as combinational cells, sequential cells and clocks.

4.4.2 DFM-aware Finite Impulse Response (FIR) filter designed at 45nm

To verify the accuracy of the DFM-aware standard cell library, the critical paths of an industrial 45nm FIR circuit (Figure 4.14), is designed with the DFM-aware standard cells.

Figure 4.15 illustrates the proposed methodology for analyzing the process variations on the critical path in a digital design. The first step is defining the critical paths in the design by using the P&R tools. Then five different netlists for the same critical paths are generated for delay calculations:

1. Netlist using the default litho and stress values from the original standard cell characterization. In other words, using the layout original values for the litho-dependent parameters (L, W) and the stress-dependent parameters (SA and SB) without context awareness.

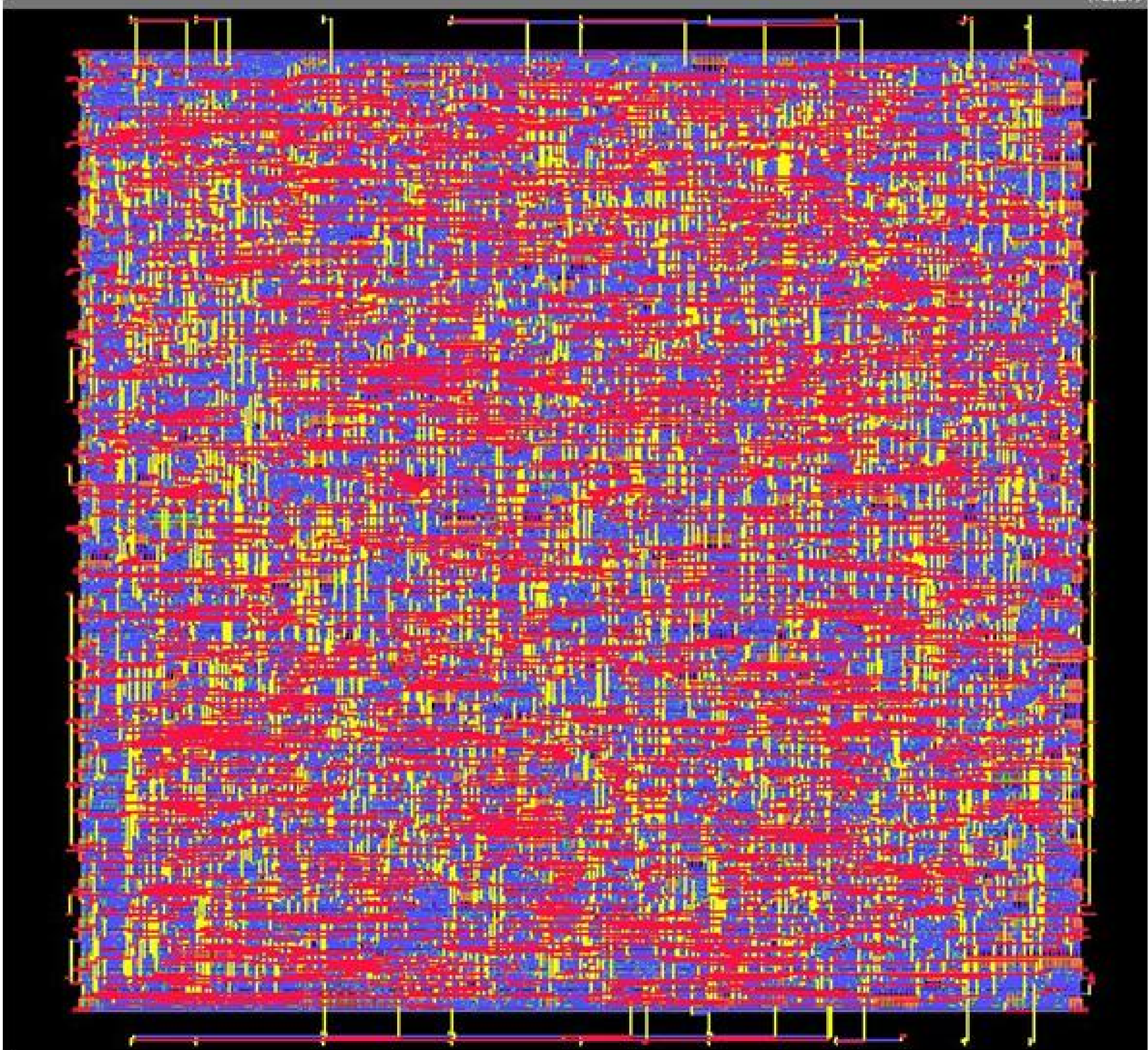


Figure 4.14: FIR Layout after Placement and Routing

2. Netlist using DFM-aware standard cells that were only characterized for lithography effects. *This netlist is used only for analysis purpose.*

3. Netlist using DFM-aware standard cells that were only characterized for stress effects. *This netlist is used only for analysis purpose.*

4. Netlist using DFM-aware standard cells that were characterized for lithography and stress effects.

5. Netlist using the extracted litho and stress parameters from the actual and full FIR layout using silicon simulation (Figure 4.16). For each standard cell in the critical path, the design-aware stress parameters such as the SA and SB parameters are extracted. In other words, the transistor stress parameters are re-evaluated by taking into account the real design context, instead of the parameters that are pre-defined in the standard cell netlist. These proximity parameters are extracted by LVS and updated in each standard cell transistor level netlist. Regarding the lithography effects, an industrial design kit containing look up tables for the equivalent L and W values according to their context, is used instead of full chip lithography simulation.

Table 4.2: Timing reports simulating standard netlist versus process-aware netlists

	Path 1 (ps)	Path 2 (ps)
Netlist 1	3368.834	3509.546
Netlist 2	3361.22	3491.88
Netlist 3	3100.966	3293.622
Netlist 4	3082.414	3256.776
Netlist 5	3035.22	3204.632

Table 4.2 lists the path delay calculated by different netlists. Several interesting observations can be drawn from our study:

1. **Context awareness characterization is necessary**: There is nearly a deviation of 250-300 ps, equivalent to 10% error in timing calculations, between the delay simulated using context-aware netlists (Netlists 2,3,4,5) versus delays calculated from Netlist 1, where litho and stress dependent parameters characterized without context information.

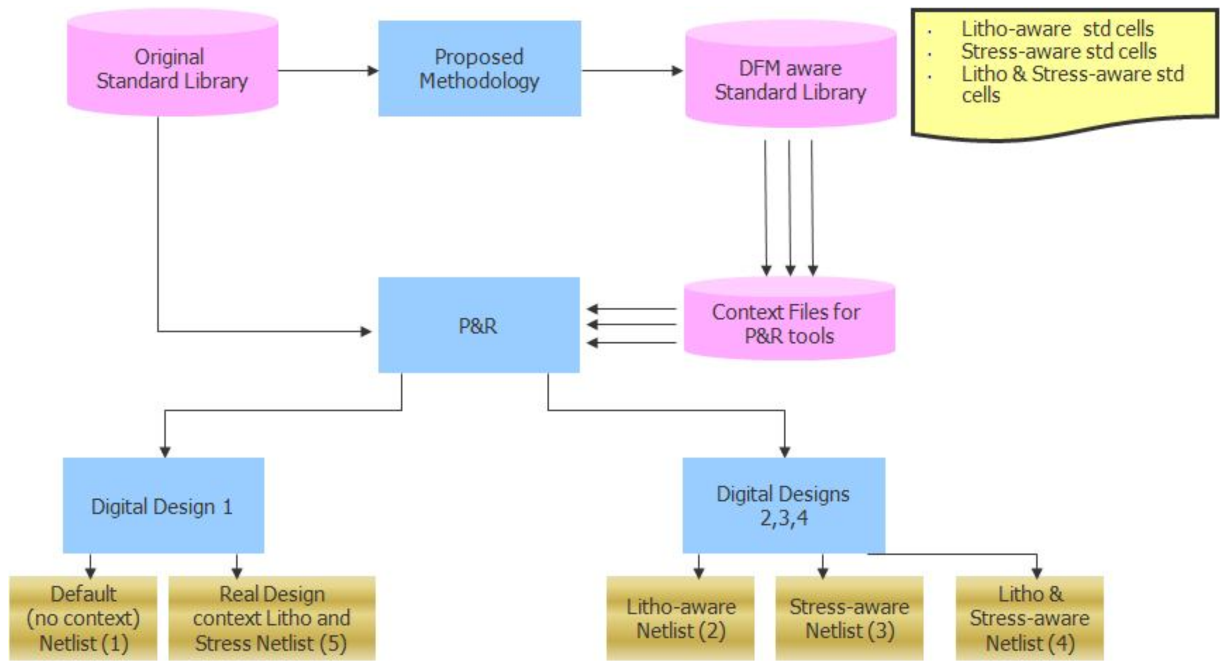


Figure 4.15: Generating 1- Default 2- Litho-only 3-Stress-only 4-Litho and Stress 5- Real Design Context -aware critical path netlists

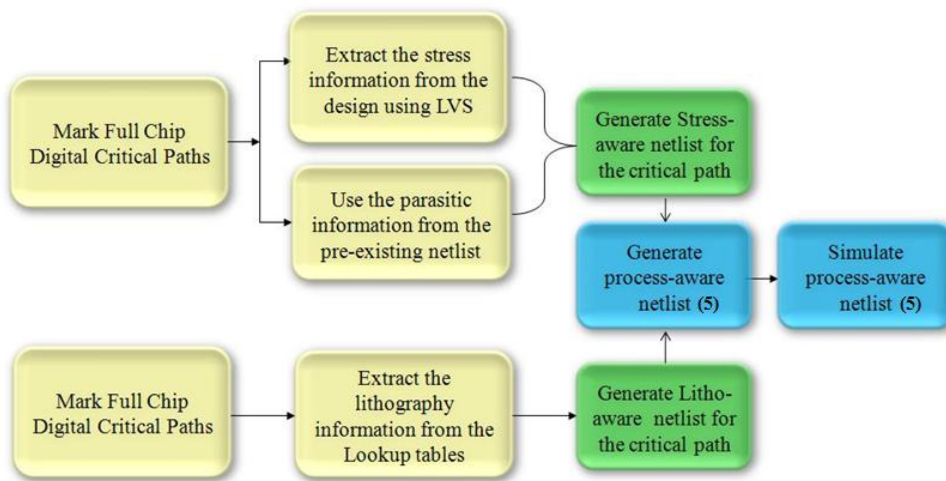


Figure 4.16: Generating the real design context critical path netlist

2. **Accuracy of the DFM tool**: Comparing Netlist 4 with Netlist 5, indicates that the characterization of the standard cells using the proposed solution gives a good predication of real design context (around 1% error in timing calculations).

3. **Lithography vs. Stress**: Comparing Netlist 2 with Netlist 3, indicates that lithography effects contribute to the timing variations in a less matter than stress effects. However this can not be a general conclusion to any other designs.

4. **Qualifying the Standard Cell Library**: Even though an industrial standard cell library is used in these experiments, it is worth mentioning that using a 45nm open source standard cell library where context awareness characterization was not performed during library characterization, can substantially impact the results. The FreePDK45 design kit [104] shows that the delay response due to the lithography variations only exhibits a behavior with a variation as high as 20ps per cell. It is evident that the timing variation is highly dependent on the characterization methodology. This conclusion supports the importance of implementing a DFM-aware standard cell library and using the updated library in the P&R. Lately, the high-tech industrial foundries have recognized the importance of DFM-aware standard cell re-characterization, and they started to characterize their advanced node standard cell libraries taking into considerations the lithography, stress and context impact.

4.5 Summary

In understanding the accuracy demands of nanometer design simulation, there is a need to add the capability of using silicon simulation based DFM solution to calculate standard cell characteristics. In this chapter we introduced a **Fix before Design** software framework aimed at improving the manufacturability of a given standard cell library. The main contributions of the proposed CAD solution are listed as follows:

- **Reduce the design development cycles** and the multiple design respins that arise from other "Fixing during Design" DFM solutions by providing a robust designs from the first design cycles.

- The proposed CAD solution re-characterizes the standard cell **under different** design contexts and Resolution Enhancement Techniques (RET).

- **Dramatically reduce the lithography simulation runtime** by introducing a novel engine that intelligently groups similar lithography effects within different process conditions.

To validate this work, we have re-designed the critical paths of FIR filter circuit using the DFM-aware library. An error of 1% was the difference when comparing the timing of the critical paths using full silicon simulations versus the critical paths designed from the DFM-aware library.

Chapter 5

High Performance Electrical Driven Hotspot Detection CAD Solution for Full Chip Design using a Novel Device Parameter Matching Technique

5.1 Introduction

With the continuous development of today's technology, IC design becomes a more complex process. The designer now not only takes care of the normal design and layout parameters as usual, but also needs to consider the process variation impact on the design to preserve the same chip functionality with no failure during fabrication. In the current process, schematic designers go through extensive simulations to cover all the possible variations of their design parameters and hence of the design functionality. At the same time, layout designers perform time-consuming process-aware simulations (such as lithography simulations) on the full chip layout, which impacts the design turn-around time. In this chapter, we present a fast physical and electrical-aware Design-For-Manufacturability (DFM) solution that detects hotspot areas in the full chip design without requiring extensive electrical and process simulations. Novel algorithms are proposed to implement the engines that are used to develop this solution. Our proposed flow is examined on a 45 nm industrial Finite Impulse Response (FIR) full chip filter. The proposed methodology is able to define a list

of electrical hotspot devices located on the FIR critical path that experience up to 17% variation in their DC current values due to the effect of process and design context. The total runtime needed to identify and detect these electrical hotspots on the full chip takes only minutes, compared to hours and days when using conventional electrical and process simulations.

The rest of this chapter is organized as follows: State-of-the-art hotspot detection solutions and their limitations are discussed in section 5.2. Section 5.3 depicts the proposed flow and algorithms entailed in implementing the engines used in the flow. To verify its usefulness, the proposed methodology is examined on a FIR filter, and the results are highlighted in Section 5.4. Finally, our concluding remarks are given in Section 5.5.

5.2 Background

As technology migrates from 90nm down to 45nm, it is increasingly difficult to achieve fast yield ramp due to random defects, process variations, systematic yield problems, and other limitations referred to as design for manufacturing (DFM) issues. At 90nm and below, these problems often appear as layout hotspots. To avoid downstream yield and manufacturing problems relating to layout hotspots, it is imperative that hotspots are addressed by different DFM techniques, as will be discussed shortly. Successful DFM techniques ensure high fabrication yield by incorporating manufacturability-aware models into the design stage to identify and remove potentially problematic process hotspots.

5.2.1 Rule-based Hotspot Detection

Hotspots have to be detected before they can be corrected. Typically, fabs use design rules to represent hotspots [105], however, this representation is inadequate. Some manufacturability issues are not local, for instance, lithography effects involve interactions over longer distances than typical minimum spacing rules. Complex conditional rules and recommended rules have been added to compensate some of this inadequacy; however, entirely using design rules for hotspot representation is resulting in an explosion of the design rule library.

5.2.2 Model/Simulation-based Hotspot Detection

To address the limitations of design rules some process models have been incorporated to analyze and drive corrections during the design stage. For instance, some design flows include a full chip simulator to verify the design and identify potential hotspots [106], [107], [108]. However, these approaches have two major limitations [105]:

1. Lack of knowledge of downstream steps. It is impossible to accurately model certain downstream processing steps such as OPC recipes. On the other hand, simple aerial image-based lithography simulations often tag regions that can be easily corrected using mask synthesis techniques. This over-estimation of hotspots is wasteful and can produce an unnecessary burden on the designers.

2. Computational burden. Certain process models are computationally expensive and hard to incorporate during physical design.

A currently prevailing approach to detect hotspots is to predict the resist pattern by applying a rigorous full chip post-OPC simulation and customer-defined checks [109], [110]. It is flagged as a hotspot if a certain configuration does not pass the checks, and the detected hotspots are corrected accordingly. A full chip simulation is usually computationally extensive, thus increasing the design-to-production time.

5.2.3 Pattern Matching-based Hotspot Detection

A pattern matching methodology has been developed to locate hotspots by scanning a layout using sample hotspots [111]. Pattern Matching is another approach that allows design, manufacturing, and failure analysis teams to identify, isolate, and define specific geometric configurations (patterns) directly from a design layout. Once recognized and defined, these patterns can be added to a pattern library that can be used by different pattern matching engines to automatically scan designs for matching patterns. Pattern matching itself is a very fast and efficient tool to identify hotspots, however, the population of sample hotspots is usually huge if they directly come from an off-line full chip post-OPC simulation. Therefore it is preferable if all the sample hotspots can be classified into groups and the representative hotspot of each group is then used for pattern matching. Thus when all the hotspots have been located, they are already classified into groups. Hotspots found in one trial of pattern matching belong to the same class as the sample hotspot used for this pattern matching trial. Furthermore, if a fixing solution is developed for the sample

hotspot, it can likely be applied to most of other hotspots in the same class. However such techniques experience from the following two major drawbacks:

1. **Accuracy vs. Runtime tradeoff**. The accuracy of the pattern matching technique depends on the quality of the pattern library (i.e. the number of patterns identified and added into the library). On the other hand too many patterns will lead to high over-estimate rate, and will directly increase the runtime of the flow and more false alarm rate.

2. **Purely geometrical-based**. Today, many of the hotspot detection approaches that use pattern matching techniques identify critical areas in the chip that are especially susceptible to defects. Additionally, they identify the proximity effects caused by the lithography process. However, these pattern matching tools are purely "geometric," without any knowledge of the electrical characteristics of the shapes that are manufactured in silicon [112]. Even other methods that incorporate modern data mining and machine learning methods [113] into hotspot detection tasks still deal with hotspots as purely geometric.

5.2.4 Electrical-driven Hotspot Detection

While the previous techniques may be useful in identifying/fixing functional failures, and increasing overall yield by a few percentage points, they completely ignore the more important category of parametric failures. At 130nm and above, parametric failures are negligible compared to functional failures. At 90nm, they become significant, and at 65nm, parametric failures become the single most critical yield-limiting factor [114] [115].

Electrical-driven DFM solutions have been proposed [95], [116] and [117] showing fair results and good performance, however these solutions were directed to reduce only lithography variations by proposing performance-driven Optical Proximity Correction (OPC) solutions and standard cell re-characterization.

In this solution, a novel electrically-aware device parameter-based matching technique is proposed for hotspot detection. The device parameters represented in the SPICE models contain different abstracts of information, such as the layout geometry, the design context and proximity effect on process variation, and the electrical information (Figure 5.1). Our technique enables an electrically and manufacturability-aware solution that tackles parametric yield issues. The fast and fully automated proposed CAD flow incorporates some key features, as follows:

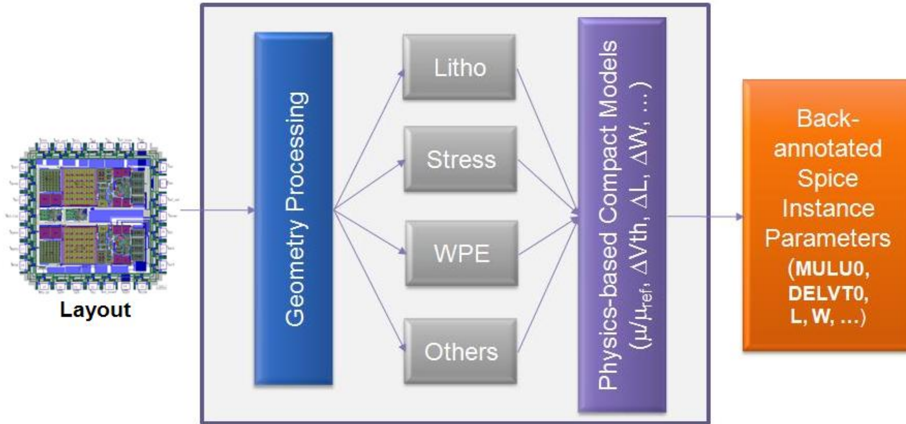


Figure 5.1: From Layout to Spice Instance Parameters. [12]

- **Process and electrically-aware hotspot analysis**

  The proposed solution presented in this chapter specifically addresses the parametric performance modeling problems encountered at smaller geometries. As this solution drives design requirements into physical layout design, and moves layout awareness upstream into design and SPICE models, a lot of useful information about the design (on the physical and electrical level) is captured, analyzed and simulated. Afterwards, this information is used during the physical layout verification stage to identify the parametric failures known as electrical hotspots.

- **Ultra fast electrical-DFM (e-DFM) solution, eliminating the need for full chip simulation**

  The electrical-DFM (e-DFM) solutions currently available perform extremely complex simulations on literally tens of millions of transistors. The commonly used methods on the front-end design stage, including corners simulation, Monte Carlo analysis, and others, are time-consuming tasks. At the same time, layout designers perform time-consuming process-aware simulations such as lithography, chemical mechanical polishing (CMP), and critical area analysis (CAA), and then perform verification checks such as Litho-Friendly Design (LFD) and stress checks. In the proposed method, we introduce a smart device matching technique, where the devices whose SPICE parameter values are within a given tolerance are smartly grouped together. The electrical behavior is very similar for all the devices in the same group. Unique devices are then selected from each set. Simulation time can be greatly reduced by simulating only these unique devices, and then mapping the electrical variation to the

rest of the devices in same group. At the same time, the process-aware simulations that are usually performed on the full chip layout are instead performed on areas where the devices that are considered critical or sensitive to variations are located.

- **Automated intent driven design solution**

  Few standard practices have been established among designers, especially regarding communication of the design's intent to those who layout the chip [118]. This is true from company to company, but also between design groups within the same organization. Existing methods, such as placement of text notes on schematics, manual annotation of schematic printouts, and even verbal instruction, have proven to be inadequate. Improving this communication avoids major layout rework that adds significant time to the design cycle. Important design requirements that are essential for parametric performing silicon must be properly captured and communicated to the layout designers. In the proposed flow, we present a fully automated engine that not only captures electrical variation constraints resulting from the physical and parametric yield analysis (such as current density, Vth, and mobility variations), but also captures the designer's intent regarding proper layout crafting recommendations to better manage the different electrical variation issues (such as device symmetry and orientation, device width segment matching, net balancing, devices to be isolated by guard rings, etc.). These constraints are propagated from the schematic level into layout rules to eliminate discrepancies between schematic design and layout. The physical and process verification step is now performed directly on the hotspot areas. In addition, the proposed design methodology guarantees that all design constraints/recommendations are met (when feasible), thus providing a robust and efficient design environment that reduces design costs and time-to- market.

- **Ability to address different types of process variations:** Lithography effects, Chemical Mechanical Polishing (CMP) effects, stress effects, and more.

5.3 Proposed Flow Overview: Design and Electrical Driven Hotspot Detection solution

Figure 5.2 illustrates the proposed DFM solution used for physical and electrical hotspot detection on a full chip design [119]. The flow can be divided into the following steps:

1. Extracting advanced device SPICE parameters from the layout extracted netlist,

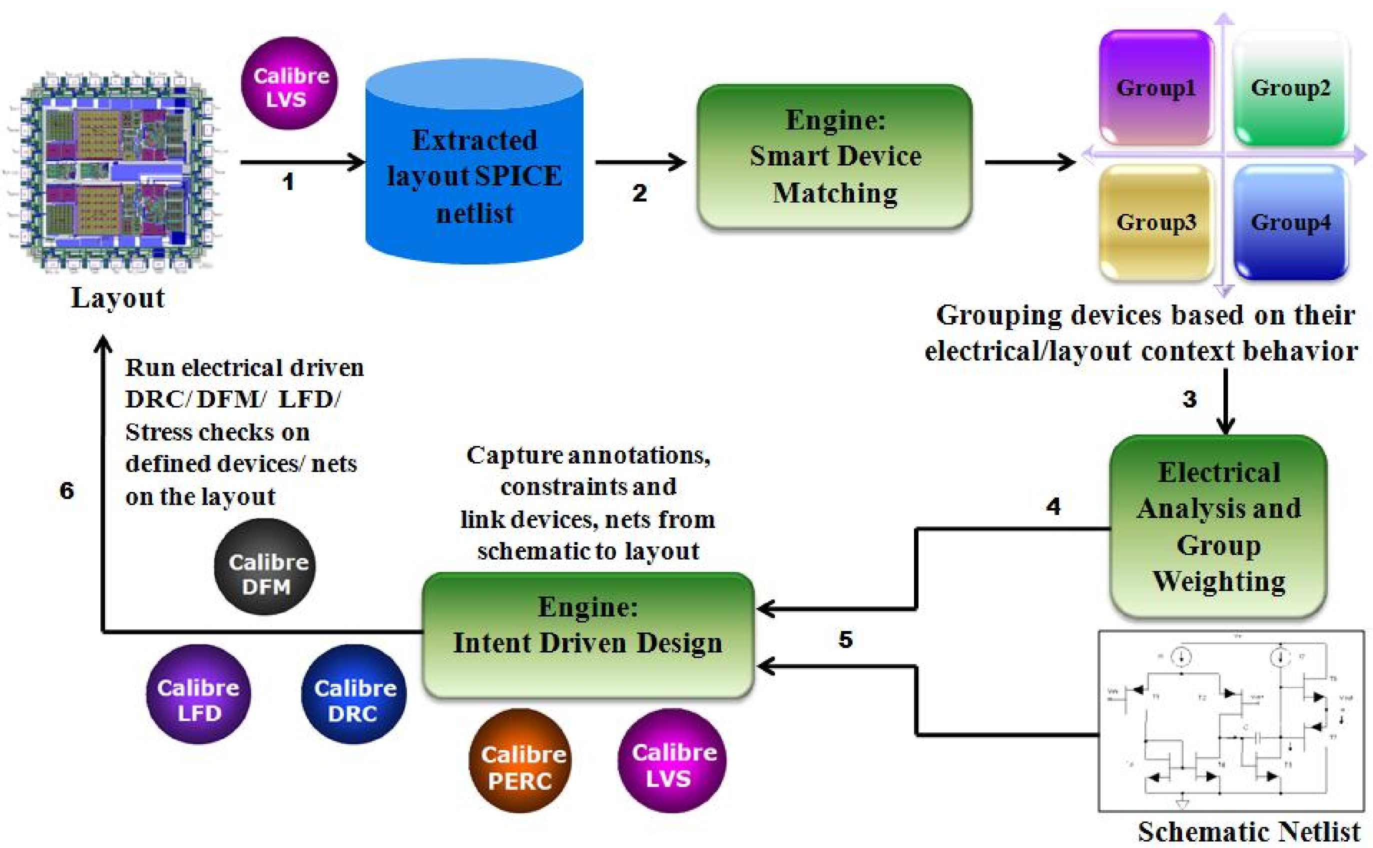


Figure 5.2: Proposed Design Context Aware and Electrical Driven DFM Solution

2. Grouping the transistors based on their device parameters,
3. Electrical and process analysis performed on samples from each group,
4. Identifying electrical/process hotspots and defining design recommendations on the schematic level,
5. Linking the schematic database to the layout database,
6. Generating and running local physical verification rules checks,
7. Fixing the electrical hotspots on the physical layout.

The first step in this methodology is extracting the SPICE netlist from the layout level. This netlist not only has the circuit information, but also the real-design context and the parasitic information. Theoretically, the netlist can be extracted from the schematic design and the electrical hotspots can still be identified. However, using the SPICE netlist that is extracted from layout provides insight into layout context that in turn impacts the electrical behavior of the circuit. This step is followed by using the newly developed Device Parameters Matching engine. The user has the option to define the list of different parameters to match. For example, Figure 5.3 shows a sample of the stress related parameters inside SPICE models [9]. These parameters gets different values based on the design context. The Device Parameters Matching engine quickly identifies similar devices that have similar device parameters values within a given tolerance. The similar devices are then grouped together. This engine will be described more later in section 5.3.1. The third step in this flow is picking a sample from each group and simulate the different samples, eliminating the need for full-chip electrical simulation. As will be shown in section 5.4, our full chip simulation results prove that all devices in the same group behave the same electrically, with negligible differences. Based on the electrical behavior of each group, we can identify and prioritize the groups based on their electrical variations and how these variations would impact the design specifications. This analysis gives us an idea of the critical devices that are sensitive to process variations. In addition, this information poses several physical design information, such as the different critical paths, pair of devices to be symmetry, net matching, and more. Once these sensitive devices, with their electrical constraints and physical design recommendations, are identified, another developed tool is used to capture this information and highlight the electrical hotspots on the layout database. In addition, physical layout rules and lithography-aware checks are generated to be used in the physical verification step.

Parameter	Description	Default Value	Unit
	Stress Effect related Parameters		
SA	Instance parameter: Distance between OD edge to poly Si from one side If not given or , stress effect will be turned off!	0.0	m
SB	Instance parameter: Distance between OD edge to poly Si from the other side If not given or , stress effect will be turned off!	0.0	m
SD	Instance parameter: Distance between neighboring fingers For NF > 1: if not given or , stress effect will be turned off!	0.0	m
SAREF	Reference distance between OD edge to poly Si from one side	1E-6	m
SBREF	Reference distance between OD edge to poly Si from the other side	1E-6	m
WLOD	Width parameter for stress effect	0.0	m
KU0	stress effect mobility degradation/enhancement coefficient	0.0	1/m
KVSAT	Stress effect saturation velocity degradation/enhancement parameter	0.0	

Figure 5.3: Stress Effect related SPICE Parameters. [9]

Litho analysis and physical verification are performed using different Calibre™ verification tools [120]. The physical verification step is performed not on the full-chip data, but on the specific devices that have been identified as electrical hotspots. Finally, different fixing algorithms can be applied on the hotspot devices.

In the following subsections, the Device Parameter Matching engine and the Intent Driven Design engine are discussed in more depth.

5.3.1 Smart Device Parameters Matching Engine

Analyzing circuit behavior after extracting lithography and stress effects from the layout requires simulating the electrical circuit with the lithography and stress parameters added to the SPICE model. For digital circuits, the goal is to find transistors whose behavior were affected by the lithography process and silicon stress such that their output current can no longer drive their load. In other words, we are looking for critical transistors whose delta current (i.e., difference in current before and after lithography and stress effects are applied), is larger than a certain value. The process of re-simulating the electrical circuit adds a new stage for designers to consider, which consequently affects the total budget for the tapeout flow. Optimizing simulation time for this stage is required to minimize its impact on the total design budget. The Device Parameters Matching engine is used to optimize the electrical simulation time. In section 5.4 we illustrate the simulation time and how it dramatically reduces the runtime using the conventional simulation methods. Comparisons of the results are also presented, showing no loss in accuracy. The methodology, as described in the pseudo code (Figure 5.4), relies on grouping the circuit transistors into groups of similar transistors, where similarity is defined as matching the values of their SPICE model parameters (Figure 5.5). With these groups, we only need to understand the delta current that occurs in one of the transistor groups to be able to later map this delta to all the other transistors in the same group. Matching of each individual parameter does not have to be exact; there can be some tolerance in each parameter that depends on the sensitivity of the electrical behavior of the transistor with respect to that parameter. Adding the tolerances allows for higher groupings or segmentation of the layout into a manageable set of groups of transistors.

Extending the idea of grouping, other models can be applied that further optimize the simulation time for this stage. For example, one model could store the results of the simulation into a library where representative transistors from each group are mapped to their corresponding delta currents. The library could be used to benefit from simulations of a prior design when creating a new design, or it could be utilized when there are iterative layout corrections and tests of the same design.

1) **User define:**
    I. The SPICE parameters of interest. Example p1,p2,p3,p4
    II. Extracted layout Spice netlist path
    III. Define the tolerance range of each parameter: dp1, dp2,dp3, dp4
2) **Loop function:** For each transistor inside the netlist M1 to Mn; start by transistor M1
    I. Read all the SPICE parameters in M1
    II. Write each parameter in an array by building a tree of nodes and pointers as follows:
        i. Node11: p1 +/- dp1 will point to Node12: p2 +/- dp2
        ii. Node12: p2 +/- dp2 will point to Node13: p3 +/- dp3
        iii. Node13: p3 +/- dp3 will point to Node14: p4 +/- dp4
    III. For the following transistor M2 run the following steps
        i. IF M2 had p1 value that lies in p1 +/- dp1;
        THEN don't build a new node, and use Node11 and write in a database that M2 intersects with M1 in p1;
        ELSE build a new pointer Node21.
        i. IF M2 had p2 value that lies in p2 +/- dp2;
        THEN don't build a new node, and use Node12 and write in a database that M2 intersects with M1 in p2;
        ELSE build a new pointer Node22.
        i. Repeat for p3
        ii. Repeat for p4
    IV. Repeat for M3,…Mn
3) **Loop function:** for all nodes N1x to Nnx:
    I. Loop function: for N1x, start x=1 to x=4
        i. Write in database all transistors grouped with M1in p1 (defined in N11)
        ii. Write in database all transistors grouped with M1in p1 and p2 (defined in N12)
        iii. Write in database all transistors grouped with M1in p1 and p2 and p3 (defined in N13)
        iv. Write in database all transistors grouped with M1in p1 and p2 and p3 and p4 (defined in N14)
    II. Repeat on N2x
4) **Output** all the groups of transistor that have complete device parameter matching to the results database

Figure 5.4: Pseudo Code for the Device Parameters Matching Algorithm

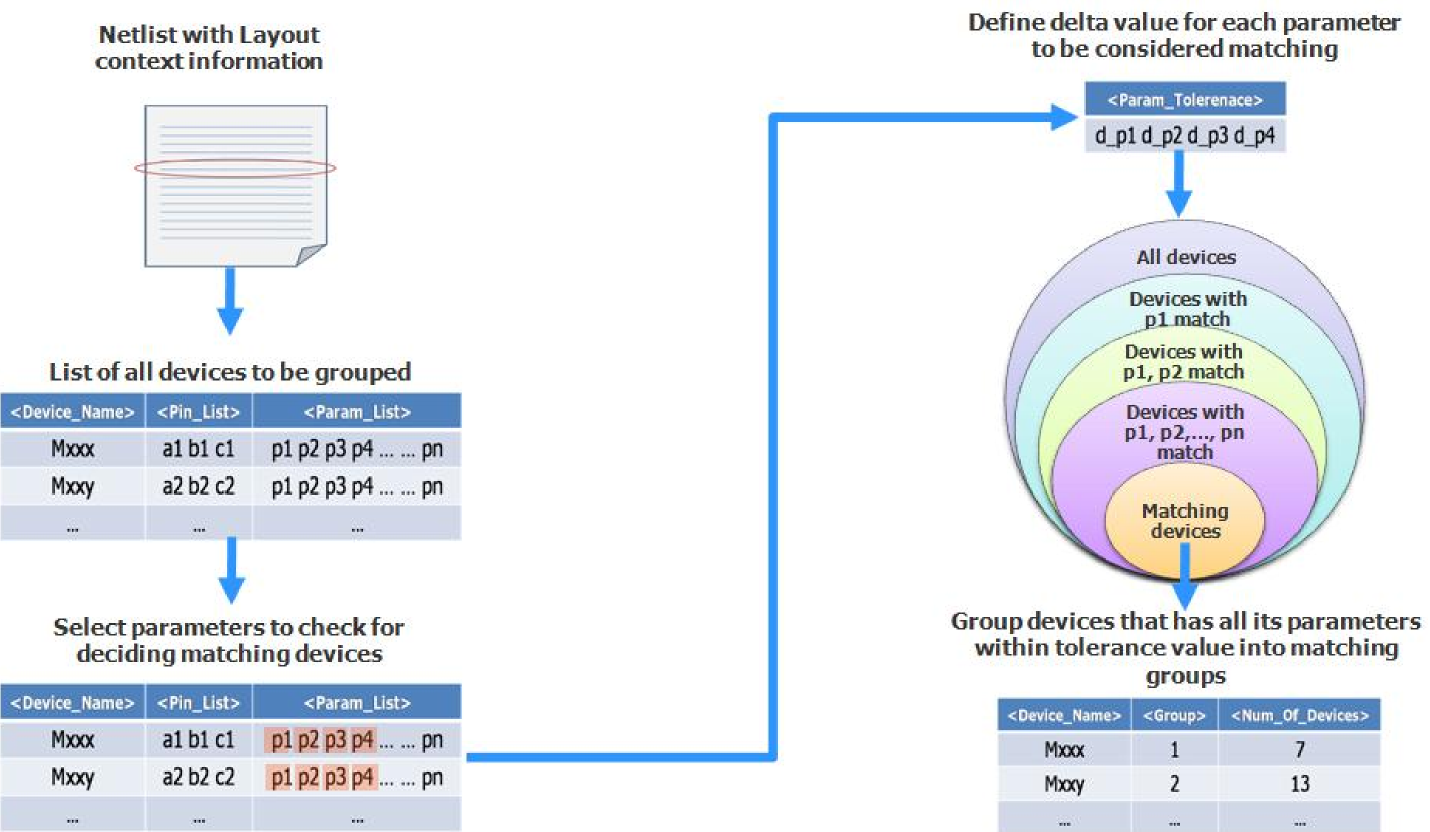


Figure 5.5: Device Parameters Matching Flow

5.3.2 Intent Driven Design Engine

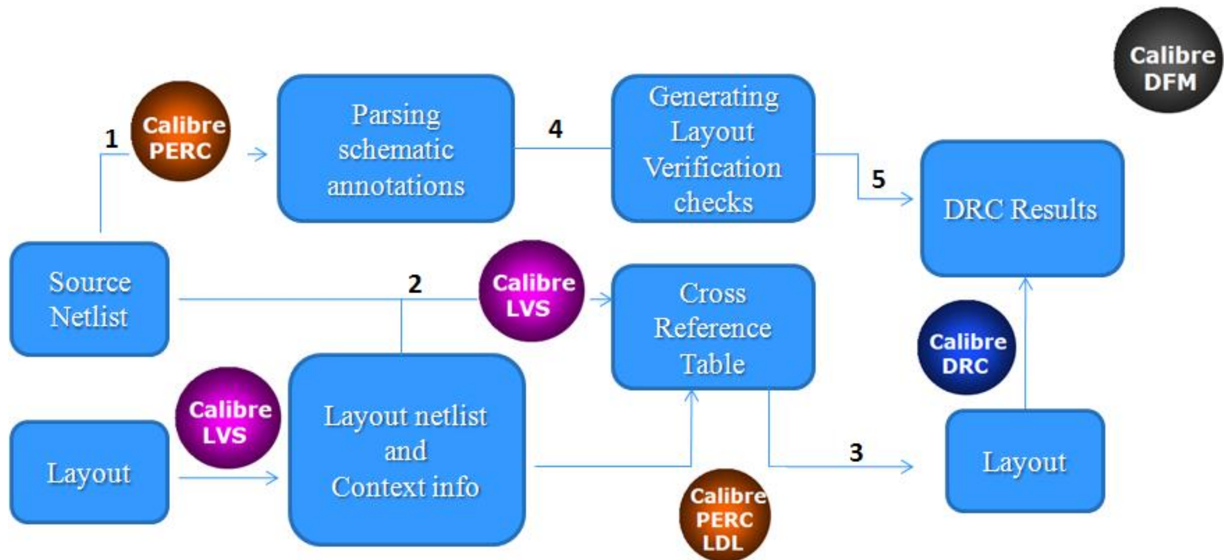


Figure 5.6: Intent Driven Design Flow

The proposed flow (Figure 5.6) automatically gathers annotation and device/net information from the schematic netlist. This information is then processed to generate text, marker layers, or other geometry identification on the layout to mark the annotated devices/nets. The annotation type is combined with the text and marker layer numbers to generate specific rules based on the assigned annotations. These rules are then tested using Calibre physical verification tools to determine if the intent was correctly interpreted and properly implemented on the physical layout design. Finally, the results are reported. An additional option uses defined analog structures (such as differential pairs) with Calibre PERC for topological matching. The flow is based on Calibre verification tools [121], [120] and TCL scripts to link these tools together. The main features of this engine are:

- Schematic-aware physical layout flow.
- Full automation, from schematic parsing through physical layout checking.
- Automated generation of complex DRC/LFD rules.
- Can be used in various applications (for example: analog layout physical verification, electrical-aware hotspot detection).

The engine flow can be divided into the following steps [122], as shown in Figure 5.6:

1. **Parsing the schematic annotations.** The flow assumes that the front-end designer places annotations on the schematic netlist in a certain format to inform the physical layout engineers that certain devices and nets have certain electrical constraints or should have special layout treatments (e.g., device symmetry and orientation, device width segment matching, net balancing, devices placed in cross-quad arrangement, or devices not isolated by guard rings). This step parses the annotations placed in the schematic netlist and identifies which device or net is involved in which device or net check. .

2. **Linking the schematic database to the layout database.** The second step links the parsed device/net to its corresponding mate in the layout. The main link between the schematic netlist and the layout is the Layout vs. Schematic (LVS) rule deck. Running LVS on a design generates the cross-referencing database that links the device/net on the schematic to its corresponding mate on the layout and provides the layout coordinates. To ensure the generated cross-referencing files correctly map the equivalent devices, the design must be LVS-clean.

3. **Marking the annotated devices and nets on the layout database.** For device checks, the coordinates obtained from the cross-referencing files place a marking text layer on each appropriate device. However, because a complete net may be composed from many layers, one pair of coordinates is not enough to mark the net for net checks, so all net coordinates are extracted from the LVS database. A script is then used to organize the data from the first two steps and generate the required text and marker layers. If there are general checks that must be applied on predefined analog topological structures, the first two steps are bypassed, and a library of these sets of structures is built to mark these devices on the layout.

4. **Generating rules based on the annotation readout.** When the device/net annotation calls for a certain procedure corresponding to a certain type of DRC check, the associated text and marker layers are passed to this procedure to automatically generate the required customized rule. The number of generated rules and the associated devices and nets are identified by the annotations.

5. **Running the generated rule files and viewing the results.** The generated rule files are executed using the Calibre tool suite [120] [121]. Violations are highlighted on the layout.

6. **Flow Integration.** The previous five steps are integrated into a single user interface, which enables the designer to have a push button solution.

The pseudo code description of the engine is shown in Figure 5.7.

5.4 Experiments and Results

In our experiment, we used a 45nm industrial Finite Impulse Response filter (FIR) full chip (75um x 65um) to detect electrical hotspots. The total number of transistors in the FIR full chip is nearly 23,000 transistors.

5.4.1 Accuracy of the proposed solution

Detecting electrical hotspots using our proposed methodology starts with running the device parameters matching engine. With this first step, all transistors based on their lithography and stress-related devices parameters are matched, such as width, length, and SA and SB parameters (average distance between the gate edge to diffusion edges measured from left and right sides). The device parameters matching engine categorized the transistors into 2153 groups. All transistors inside each group were proved to have similar electrical behavior. For example, the maximum mismatch error in DC current among all transistors in any one group is within the range of 0.001%-0.9%.

Tables 5.1 and 5.2 show examples of two groups as the output from the device parameter matching engine containing a list of the matched devices and the value of their DC current. It was confirmed visually that the SPICE stress parameters for all the devices in one group were matched. For example, devices M1, M3, M6, and M7 had exactly the same SPICE stress parameters (SA=1.1e-07, SB=2.08356e-07, SCA=28.0577, SCB=0.0260972, and SCC=0.003727), while nearby transistors, such as M2, had different SPICE stress parameters (SA=1.1e-07, SB=1.93467e-07, SCA=30.7143, SCB=0.0272008, and SCC=0.00423936).

The flow then picks one transistor from each group as a reference, to calculate the variation of the electrical specifications due to real physical design context for each transistor in that group. When the schematic netlist is first simulated (i.e., no physical design-aware information) these stress parameters are not loaded; in this case, the SPICE model card loads the default stress parameters. In this experiment, the DC current variation between

1. **Parsing The Schematic Annotations:**
   - The schematic netlist is parsed searching for the lines starting with "*@**device_ann**" which indicates the device annotations or "*@**net_ann**" which indicates the net annotations. These annotations are not standard, it can be customized.
   - After these lines are indicated, they are analyzed one by one to produce a new schematic netlist where the annotated devices/nets are re-named to be device/net_annotation1_annotation2_.....
   - For example devices **M1** & **M2** which have annotations "**match**" for matching are re-named to be **M1_match_M2** & **M2_match_M1**
   - And the same is applied for the nets, for example the nets **IN1** & **IN2** which have annotations "**bl**" for balancing are re-named **IN1_bl** & **IN2_bl**
   - The reason of this step is to pass the new adjusted schematic netlist to Calibre PERC which would be able to read the device/net names and collect the information about the device/net annotations in a separate file
   - The following step is to run Calibre PERC with a simple PERC rule deck that analyzes the name of each device/net in the schematic netlist and produces two files that contain the collected annotation information
   - The first file is "**.dev_name**" which has the device annotations for **M1_match_M2** as follows:

     M1_match_M2 match
   - The second file is "**.net_names**" which has the net annotations for net **IN1_bl** and **IN2_b1** for example as follows:

     IN1 bl
IN2 bl
2. **Linking the schematic database to the layout database:**
   - An LVS run is performed for the schematic netlist vs. the layout to get the required information to link the device/net annotations in the schematic to their corresponding ones in the layout
   - The two options **–ixf** and **–nxf** should be used with the Calibre LVS execution command to get the required information as follows

     calibre –lvs –ixf –nxf rule_file

     where ixf stands for "instance cross referencing" while nxf stands for "net cross referencing"
   - For correct mapping the layout should be LVS clean
   - Going to the resulted LVS database "**SVDB**" two files could be found name "***.ixf**" & "***.nxf**", these two files tell which device/net in the schematic netlist corresponds to which device/net in the layout, it also gives some info about the device/net lower left coordinates and the device/net id in the data base
3. **Marking the annotated devices and nets on the layout database:**
   - To mark the annotated device in the layout the information found in the **.dev_names** file together with the ***.ixf** to place the associated annotations as a text on the lower left most of each finger of the device
   - A similar approach is done with the annotated nets using the **.net_names** and the ***.nxf** files, but in this case a text is not enough on the lower left corner to identify the complete net, so the net id is used to access the LVS QUERY database with Calibre QUERY server "**calibre –query database_name**" to get all the coordinates of this net, then a script is used to place a marker with a certain gds number to cover this net
   - If there are general checks that must be applied on predefined analog topological structures, the first two steps are bypassed, and a library of these sets of structures is built (using Calibre PERC) to mark these devices on the layout.
4. **Generating rules based on the annotation readout**
   - The generic rules associated with each annotation is coded in a TVF format in an external file, these rules are coded as being TCL procedures that are called only when a certain annotation is detected
   - So for example if a "**match**" annotation is detected in step1, the procedure which defines the symmetry rule in TVF is called and the information collected in steps 2 & 3 (device name, text layer number,……) are passed to this TCL procedure
   - Finally this TVF file will be compiled producing the SVRF rule, this file is a customized DRC file based on the annotations found in the schematic netlist
5. **Running the generated rule files and viewing the results**
   - Finally the schematic driven rule file is texted on the layout under test and the results are normally viewed in Calibre RVE like any other DRC rule file and the Violations are highlighted on the layout for the designer
6. **Flow Integration**
   - The previous five steps are integrated into a single shell script which automatically runs these steps successively; this enables the designer to have a push button solution.

Figure 5.7: Pseudo Code for the Intent Driven Design Engine

Table 5.1: Grouping the devices as an output from the Device Parameter Matching Engine-Group 1: contains 4 matched transistors

Device Name	X	Y	DC Current	Error %
M1 (ref)	1160	6310	129.1432	0%
M3	1160	11350	129.1483	0.00395%
M6	1160	18910	129.143	0.0041%
M7	1160	21430	129.1496	0.0051%

Table 5.2: Grouping the devices as an output from the Device Parameter Matching Engine-Group 2: contains 8 matched transistors

Device Name	X	Y	DC Current	Error %
M13723 (ref)	10400	52815	81.8841	0%
M14199	12780	57390	81.8005	-0.1021%
M14311	13340	32190	81.505	-0.3612%
M14344	13480	52815	81.5493	0.0543%
M15024	17400	57390	81.5193	-0.03678%
M16104	22860	57855	81.4788	-0.0496%
M16131	23000	30135	81.759	0.34389%
M16182	23280	54870	81.8042	0.05528%

Table 5.3: Top five device groups experiencing variation in the DC current values

Group Index	Number of devices	DC Current simulated from design aware device parameters netlist	DC Current simulated from default device parameters netlist	Error %
2049	7	247.406	212.5301	16.4%
2153	6	27.2417	32.2432	15.52%
1307	5	95.3815	110.862	13.96%
12	2	121.1794	140.7287	13.8%
1310	7	95.86	110.6609	13.3%

devices with design-aware stress parameters is compared against the default stress parameters (no real design awareness). The groups were then reorganized based on their electrical variations. From this experiment, the variation in the DC current values due to process effect can reach up to 17%. Table 5.3 lists the top five groups that are experiencing electrical variations due to stress effect. These devices are considered electrical hotspots.

From the synthesis or schematic level, the critical paths are identified. Then the intent driven design engine highlighted these groups of transistors on the layout. High priority was given to devices located on the FIR's critical paths (Figure 5.8). LFD and stress simulations can now be checked locally on the highlighted hotspots.

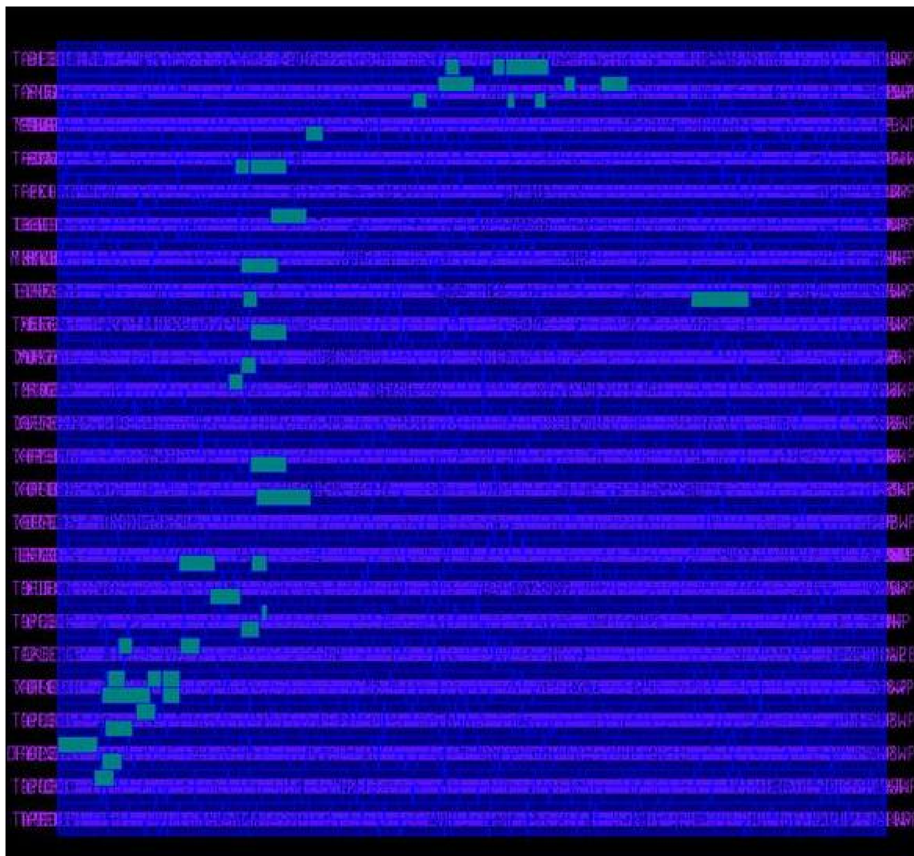


Figure 5.8: FIR layout highlighting the potential electrical hotspots on the critical paths

5.4.2 Runtime Analysis

Current state-of-the-art solutions use geometrically-based hotspot detection that depends on identifying problematic structures either by running lithography and stress simulations on full chip, or by using fast techniques such as pattern matching or machine learning. However, because our solution uses electrically-aware hotspot detection, it provides a major advantage over these solutions. To illustrate this advantage, we compared the runtime required by existing hotspot detections solutions to identify problematic structures against the runtime in our proposed solution for grouping the devices. Running the device matching engine on the FIR (75um x 65um) using one CPU took few seconds to complete. Runtime of existing solutions were studied and summarized in [123] for $1mm^2$ chip. Comparing solutions, as shown in Table 5.4, we can conclude that our grouping solution is extremely fast compared to existing geometrical-based solutions.

To enable the conventional flows to be electrically-aware, existing solutions must first extract the electrical netlist and parasitic information from the layout. An electrical simulation is then performed, followed by a calculation of the variation of the electrical specifications for each transistor due to real physical design context. The electrical DC current simulations of the FIR full chip alone required about one hour. However, using our grouping methodology, we were able to run electrical simulations on the complete set of devices in the FIR full chip in nearly two minutes. As mentioned, running the device matching engine on the FIR took few seconds to complete. This means that the complete flow took **nearly 3 minutes** to group, simulate, identify, and locate the electrical hotspots on the FIR full chip layout. Our proposed solution is an encouraging tool that provides a quick start to locating physical and electrical hotspots on full chip designs. Additional analytical simulations can be performed afterwards.

Table 5.4: Result comparison between previous hotspot detection methods and our method

	[124]	[125]	[113]	[123] A	[123] B	Our smart device matching method
Avg real-time run-time per mm^2 [1]	100	8.5	2.8	0.52	0.4	0.17
runtime	slow	medium	medium	fast	fast	very fast

Run-time calibrated in the unit of CPU hour/mm^2 on Linux station with 2.8GHz quad-core processors.[1]

5.5 Summary

In this chapter, a novel methodology is introduced that smartly categorizes all devices in the full chip by similar electrical and layout context behaviors. Afterwards, electrical and process variations analysis is performed on one device in each group, and prioritizes the hotspot groups based on designer's electrical and process constraints. The main contributions of the proposed CAD solution are listed as follows:

- Developed an **electrical-aware** and **design-driven** DFM solution that detects hotspot areas in a full chip design.
- The proposed CAD solution is using a **novel** device parameter matching engine
- This CAD is an **ultra-fast** electrical hotspot detector that avoids the time-consuming full chip simulations previously used for hotspot detection.
- This flow plugs directly into the designer's existing flows and **avoids** going through extensive post layout runtime simulations.

To validate this work, the proposed solution quickly identified electrical hotspot devices on a 45nm industrial FIR circuit in less than three minutes compared to days and weeks using the conventional silicon simulation methods.

Chapter 6

A Parametric DFM CAD Solution for Analog Circuits: Electrical Driven Hot Spot Detection, Analysis and Correction Flow

6.1 Introduction

Many designers have started to recognize that this push into advanced nodes has exposed a hitherto insignificant set of yield problems. Companies are searching diligently for real-life solutions that address the negative effects of process variation in manufacturing as it applies to total manufacturing costs. At 130nm and above, physical problems that cause catastrophic failures, such as bridging faults, were always the primary focus of verification efforts. At those nodes, parametric failures were negligible compared to physical failures. However, at 90nm, parametric failures become significant, and at 65nm, parametric failures become the single most critical yield-limiting factor [114] [115] especially for analog circuits.

Parametric yield issues arise when process variation has not been sufficiently characterized, such that a circuit may have achieved design closure through standard methodologies, but the silicon performance does not match the simulation results. Both physical and parametric yield failures have similar negative business implications, in that new circuit designs will fail to meet performance expectations [17].

In this chapter, we present a complete electrical-aware design for manufacturing solution that detects, analyzes, and fixes electrical hotspots (e-hotspots) caused by different process

variations within the analog circuit design. Novel algorithms are proposed to implement the engines that are used to develop this solution. Our proposed flow is examined on a 65nm OPAMP and 45nm industrial voltage control oscillator (VCO). In addition the proposed e-hotspot detection engine is verified against silicon wafer data for a level shifter circuit designed at 130nm. The e-hotspot devices with high variation in DC current are identified, and after fixing these e-hotspots, the DC current variation in these devices is reduced, while saving the original circuit specifications.

The rest of this chapter is organized as follows: Challenges in the current analog design causing parametric yield issues are discussed in section 6.2. Section 6.3 depicts the proposed flow, and the algorithms entailed in implementing the engines used in the flow. To verify its usefulness, the proposed methodology is examined on level shifter, VCO and OPAMP circuits, and the results are highlighted in Section 6.4. Finally, our concluding remarks are given in Section 6.5.

6.2 Background

Challenges in the current analog design flows: A typical analog design flow is divided into two phases; Front End Phase (FEP) and Back End Phase (BEP). In the FEP, designers deal with the schematics entry to build the circuit. Circuit designers will place transistors with certain width and length in the schematics. Spice transistor simulation tools are used in this phase to verify circuit performance. Extensive simulations runs are needed to check the design for process variations. Then schematics designer passes this design to layout designers in the BEP with some layout constrains to be considered. Constrains are such as; which transistors need to be matched; where are the differential pairs and other constrains as well. These schematics will be translated in the layout stage or BEP into actual physical layout devices. Then using physical verification tools, layout designers can verify that the physical design is following the design technology rules for fabrication. Also these physical verification tools are used to verify that the layout is matching the schematics. A parasitic extraction is back annotated again to the netlist in the FEP to run post layout simulations.

So, current design methodology is a kind of trial and error cycles. These cycles are highly computational expensive; especially in the analog and mixed signal flow with different design corners to be covered.

Analog design sensitivity to systematic variations: Lithography variation impact the devices dimensions (L,W,AD,AS). In addition, Shallow Trench Isolation (STI) in the CMOS process induces mechanical stresses on the transistor channel. These mechanical stresses alter the transistor channel mobility (μ) and voltage threshold (V_{th}) causing

deviations in the electrical performance of the transistors and subsequently in the circuit target specifications. Similarly the well proximity effects can cause changes in the doping profile of the transistors causing changes in the μ and V_{th} as well. These physical layout effects have high impact on sensitive analog designs such as current mirrors, differential pairs, amplifiers and others; causing circuit mismatches, DC current offsets and deviations from their original target specifications in the schematics [126].

Physical layout impact on analog design specifications: Moreover, at 90nm and above, neighboring devices in the layout are not of much effect. But at more advanced nodes as in 65nm, 45nm and beyond, geometrical and electrical parameters of the transistor are affected by the neighboring devices in the layout; which is totally unpredictable in the schematics phase or the FEP [127]. Another point to be considered is the common practice in analog designs to convert single schematic transistors into multi-finger devices in the layout. In this case each finger of the transistor will be affected with the layout proximity effects in a different way from the other fingers; this is depending on its location, although all fingers belong to the same single transistor in the schematics [15] (Figure 6.1).

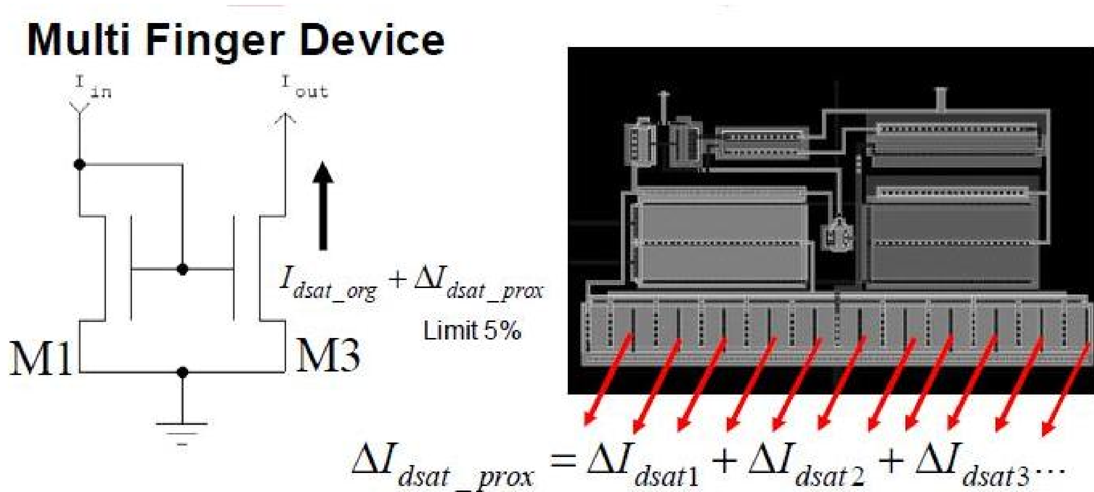


Figure 6.1: Layout proximity effects on multi-finger devices can not be predicted in schematics phase [15]

Electrical and physical constraints are getting more complex: Layout constrains are used as guidelines for the layout designer to lay correctly the transistors and to meet the design specifications. For example; schematics designer needs to define differential pair transistors. These two differential transistors to be laid in a matched way in the layout and to avoid any sources of mismatch between them; example: *M1 and M3 transistors are*

differential transistors to be matched in gate length within 5% tolerance. With the advances in the IC process technologies and the migration of analog designs from 130nm/90nm to 65nm/45nm and beyond, the proximity constrains have to be considered. For example; I_{dsat} *in M3 not to exceed 5% due to layout proximity effects*, this constrain can show up in the design cycle of a DC current bias cell (a common block in analog designs). In the traditional design flow, such new constrains are difficult to be handled either on the FEP side or on the BEP side.

There are some ways currently used in the design flow to account for layout effects but with drawbacks. For example, on the FEP side, CAD engineers can provide the IC schematics designers with Process Design Kit (PDK). This PDK gives estimates for the layout proximity effects in schematics. However this way is not accurate as proximity effects cant be accurately predicted unless the layout is fully finished. On the other side, layout designers do not use spice transistor simulations tools to perform simulations to predict electrical changes in performance due to the layout drawn. Also, it is not feasible to add a time consuming electrical simulator to the layout drawing phase as this is not a common practice and is fully disrupting the normal design flow. Thats why a different methodology is needed to handle such issues in the design cycle and this is proposed in the following CAD solution.

6.3 Proposed Flow Overview: Electrical Hot Spot Detection, Analysis and Correction Flow

The proposed CAD flow incorporates some key features that enable the proposed parametric DFM solution and differentiate it from the existing solutions. These features are highlighted as follows:

- **Electrical-aware DFM solution**

  Today, many of the approaches that are commonly referred to as physical DFM techniques only address catastrophic defects and systematic process variations. These techniques include spreading wires, doubling vias, identification of critical areas in the circuit that are especially susceptible to defects, and identification of proximity effects caused by the lithography process. However, physical DFM tools are purely "geometric", in that they work to preserve shape fidelity without any knowledge of the impact on the electrical characteristics of the shapes that are manufactured in silicon [112]. While these techniques have proven useful in reducing functional

failures and increasing overall yield by a few percentage points, they completely ignore the more important category of parametric failures. The proposed solution presented in this chapter specifically helps to address the parametric performance modeling problems encountered at smaller geometries. As this solution drives design requirements into physical layout design and moves layout awareness upstream into design, useful information about the design (on the physical and electrical level) is captured, analyzed, and simulated. Deviations in the electrical characteristics due to physical layout and process variations, are identified and highlighted on the design. These deviations are referred as electrical hotspots (e-hotspots).

- **Complete Parametric DFM Solution for Analog Designs: Detection, Analysis and Correction**

  Several prototype techniques exemplify how parametric DFM solutions can take into account design-specific information to improve design analyses and enhancements [112]. Examples include (1) iso-dense awareness of pitch-dependent through-focus CD variation, to reduce timing guardbands and improve timing robustness [128]; (2) Selfcompensating design techniques that minimize the inherent sensitivity of critical paths to various sources of process variation [129]. Other parametric DFM solutions are focused on analyzing and detecting the electrical variations, although some solutions include DFM-aware place-and-route remedies that propose e-hotspot fixes. However, all of these solutions mainly address digital designs [130]. In our work, we propose a variability-aware parametric DFM solution that detects and fixes e-hotspots for analog circuits.

- **Eliminate time-consuming full circuit SPICE simulations**

  Analyzing circuit behavior after extracting lithography and stress effects from the layout requires simulating the electrical circuit with the lithography and stress parameters added to the SPICE model. For analog circuits, the goal is to find transistors whose behavior are affected by the lithography process and silicon stress such that their DC current can impact the circuit specifications (such as gain, phase margin, frequency response). In other words, designers are looking for transistors whose electrical parameters such as current, V_{th}, and mobility experience variations before and after lithography and stress effects are applied and violates the designers' electrical constraints. This process of re-simulating the electrical circuit adds a new stage to the process flow, which consequently increases the total budget for the tapeout flow. To optimize the electrical simulation time and minimize the impact on the total design budget, the proposed parametric DFM solution can detect the electrical

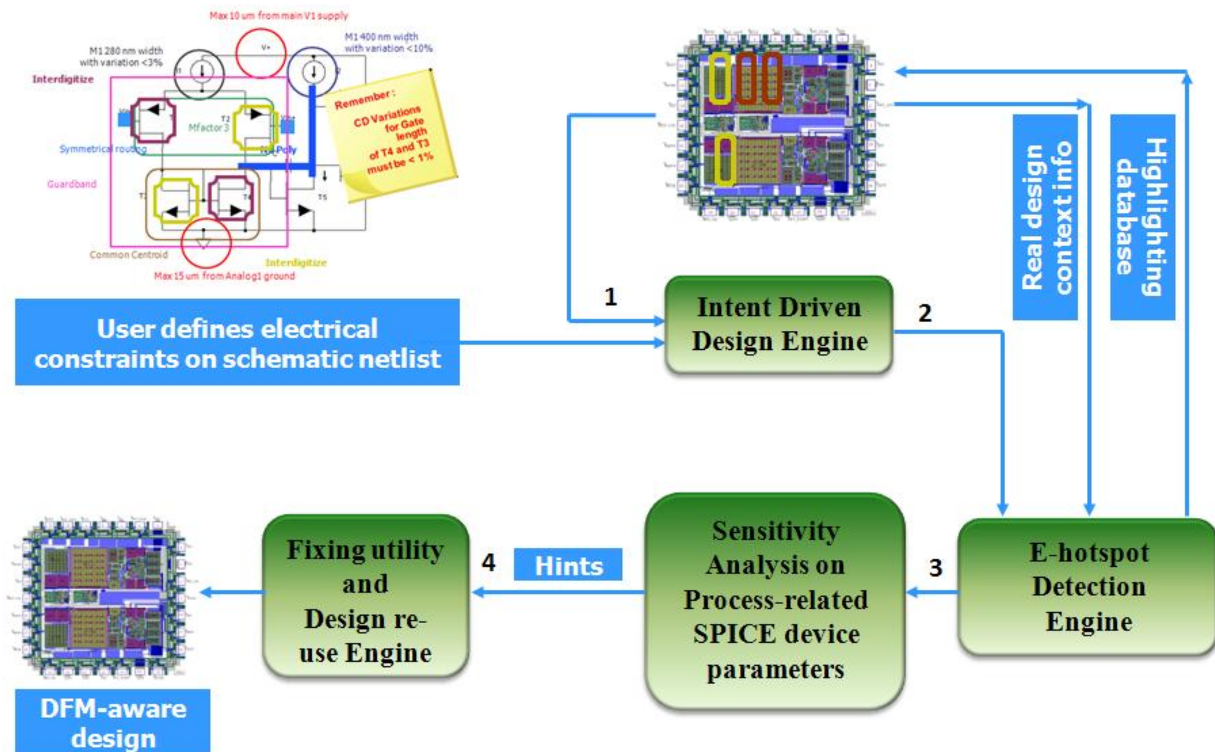


Figure 6.2: Proposed Flow Overview: Complete Electrical DFM Solution

behavior either through lookup tables or generates a minimized SPICE netlist that is used for the re-simulation.

- **Automated intent driven design solution**

  In the proposed flow, a fully automated engine is used to capture electrical variation constraints. These constraints are propagated from the schematic level into layout level to ensure actual process variation analysis based on real design context.

Figure 6.2 illustrates the proposed parametric DFM solution used for automated electrical hotspot detection and correction in analog circuit design [131]. The proposed flow is explained through the following steps:

1. **Design Intent Gathering:** Capture all user-defined electrical constraints and real design physical layout context information,

2. **E-hotspot Detection:** Detect the impact of process variations on the electrical behavior of the devices by identifying the devices with electrical behavior that deviates from the user-defined constraints,

3. **E-hotspot Analysis:** Perform sensitivity analysis on e-hotspot device parameters, and generate hints for fixing these e-hotspots,

4. **E-hotspot Correction**: Capture the fixing hints and re-design the circuit without impacting the original design specifications.

The first step in this methodology is gathering the different electrical constraints predefined by the designer and linking each of the constraints to the corresponding device on the physical layout. To do so, a fully automated engine is implemented. This engine not only captures the designer's intent in regard to electrical constraints, but also captures any other design intent information, such as layout crafting recommendations that will better manage the different electrical variation issues (for example: device symmetry and orientation, device width segment matching, net balancing, devices to be isolated by guard rings). This design intent information is propagated from the schematic level into the layout. This engine will be described further in section 6.3.1.

The second step performs lithography and stress analysis to update the SPICE device netlist with all the adjusted device parameters that were impacted by the process variations. Thirdly, certain devices in the updated device netlist are simulated to determine if there are any violations in the pre-defined electrical variations. If there are violations, this device is marked as an e-hotspot. This step is then followed by analyzing the e-hotspot device parameters to identify which parameters have been impacted the most by the lithography or stress variations. Fixing hints (in the form of geometrical modifications) are generated to guide the correction engine to the proper geometrical adjustments needed on the physical layout.

Finally, a correction engine applies adjustments to the e-hotspot devices. This correction engine interface with an in-house design reuse tool, ChameleonART^{TM} [132], which is used to re-design the layout with minimum geometrical changes without impacting the original design specification and while preserving the design rule checks (DRC).

In the following subsections, the intent-driven design engine, e-hotspot detection engine, sensitivity analysis algorithm, and correction engines are discussed in more depth.

6.3.1 Intent Driven Design Engine

The proposed intent-driven design engine (Figure 6.3), assumes that the front-end designer places design constraints in the form of annotations on the schematic netlist that are extracted from the electrical analysis stage. These annotations inform the physical layout engineers that the electrical behavior of certain devices and nets should be preserved after the physical layout implementation. Afterwards, this information is processed to generate text and marker layers on the layout to mark the annotated devices/nets. The type of annotation, together with the text and marker layer numbers, is used afterwards by the e-hotspot detection engine to generate a minimized SPICE netlist. The flow goes through the following steps:

1. Parsing the schematic annotations,
2. Linking the schematic database to the layout database,
3. Identifying the annotated devices and nets on the layout database.

The first step parses the annotations placed in the schematic netlist (such as electrical variation constraints or layout recommendations). This step also correlates each device/net with its corresponding design constraint check; this information may be reported in a separate output file containing each device or net and its associated annotation.

The second step links the parsed device or net to its corresponding mate in the layout. The main link between the schematic netlist and the layout is the Layout versus Schematic (LVS) rule file, as running LVS on a design generates the cross-referencing files that identify which device or net on the schematic corresponds to which device or net on the layout. It also assigns, and reports a corresponding coordinate for that device or net on the layout. For this part of the step to run correctly, the design must be LVS-clean, to ensure the generated cross-referencing files correctly map to the equivalent devices.

The third step uses the coordinates obtained from the cross-referencing files and places a text layer on each appropriate device. In this step, a software script is used to organize and apply the data obtained from the first two steps to generate the required text and marker layers. Because a net may be composed of many layers, one pair of coordinates is not enough to mark a complete net. For the net checks, the net coordinates from the LVS database are extracted to mark the net. When the device/net annotation calls for an electrical simulation, the associated text and marker layers are passed, along with a procedure that generates the required SPICE testcase.

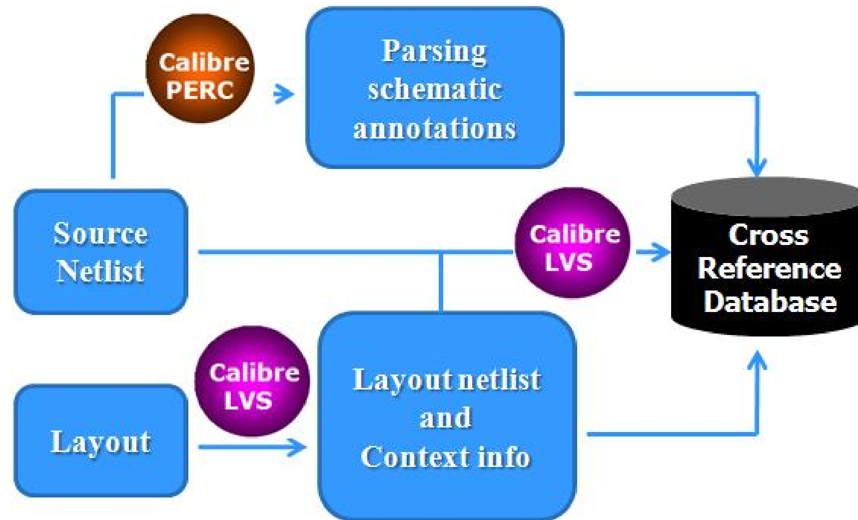


Figure 6.3: Intent Driven Design Engine

6.3.2 E-hotspot Detection Engine

The e-hotspot detection engine extracts the layout-dependent systematic manufacturing variations and study their impact on the parametric behavior of devices. The e-hotspot algorithm is described as follows:

1. **Extract real design context-aware stress parameters;** The Calibre LVS tool is used to extract the layout netlist that contains the physical design context information. The design-aware stress parameters (such as the SA and SB parameters) are extracted and used in place of the default stress parameters obtained from schematics. In other words, the transistor stress parameters are re-evaluated by taking into account the real design context. These proximity parameters are extracted and then updated in a new SPICE netlist.

2. **Extract the actual litho aware device dimensions;** Regarding lithography effects, a silicon simulation-based methods [133] [84] can be used to incorporate lithography variations to simulate the lithography contours and extract the equivalent L and W values from them. These equivalent L and W values are then updated in each device netlist. As an alternative approach, some foundries provide the option in their design kit to extract the equivalent L and W values directly from a pre-calculated look up table. For this solution, we used an industrial design kit containing look up

tables for the equivalent L and W values according to their context. Hence, the litho contour extraction flow through simulation is avoided.

3. **Calculate electrical variations;**

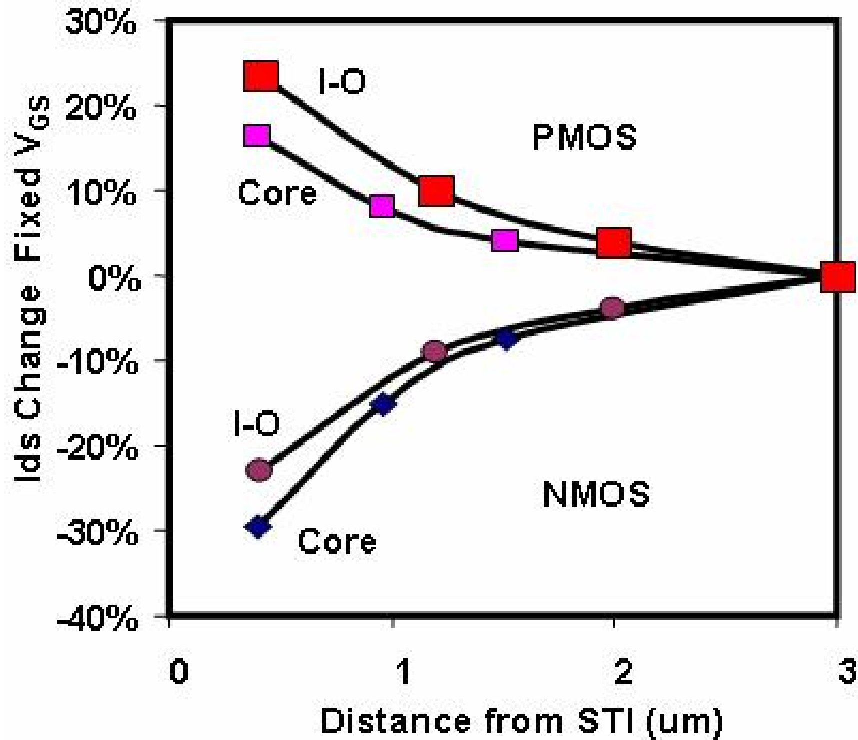


Figure 6.4: Ids vs. Distance from STI boundary Typical 0.13um process [16]

Lithography, STI and well proximity effects cause a shift in transistor dimensions, mobility and threshold voltage, which cause a shift in the I_{dsat} [126]. All these shifts can be detected using DC analysis for each device. Figure 6.4 is a plot showing an example of how I_{dsat} varies as a function of distance from the STI boundary when biased at low current where V_{gs} is within 200 mV of threshold. This has important ramifications in staple circuits of analog design like current mirrors and differential amplifiers. Figure 6.4 illustrates that I_{dsat} mismatches as large as 30% can be possible so in extreme cases. An example of a DC parametric failure for an amplifier circuit caused by stress effects, is illustrated in [16], proposing a solution to catch parametric failures by monitoring the DC bias currents in a post-layout extracted simulation.

Three different approaches are adopted in the proposed flow to calculate the shift in the I_{dsat}:

- Use built in functions exactly similar to the transistor model equations to calculate the variations in V_{th}, mobility, device dimensions and then pass these variations to look up table that consists of a list of a pre-simulated values of I_{dsat} as function of V_{th}, mobility, device dimensions. This flow avoids introducing more simulations, however in this case there is accuracy tradeoff depending on the lookup table database [15].
- Generate a minimized SPICE netlist updated with process-aware parameters for each of the annotated devices in the schematic design. In order to get accurate values, and once the process-aware netlist is generated for each device with electrical constraints in the design, a SPICE simulation is performed.
- For small sized analog circuits, there is an option to simulate the complete circuit. Afterwards variations in percentage for absolute and matched devices are calculated.

4. **Highlight violations;** Compare the original electrical behavior versus the results from the extracted device parameter. Whenever a device violates the user-defined electrical constraints, that device is highlighted as e-hotspot devices. Below is a sample of the output from the e-hotspot Detection Engine: List of the current variations in each device in the design followed by list of devices that violated the designer constraints.

```
##################################################################
##############         Device Reports           #################
##################################################################
Id variation in device M12: -7.90739272034%
Id variation in device M50: -5.56061578377%
Id variation in device M36: 3.60789458673%
Id variation in device M40: 2.35862956526%
##################################################################
##############         Device Violations        #################
##################################################################
Violation: M12 absolute Id variation = 7.90739272034 > 5%
Violation: M50 absolute Id variation = 5.56061578377 > 5%
##################################################################
```

This flow enables the layout engineers to identify areas in the layout considered to be e-hotspots and also checks that the designer's intent has been correctly passed and properly implemented on the physical layout design.

6.3.3 Sensitivity Analysis Algorithm

The goal of this step is to identify the exact source of the layout effects on integrated circuit performance.

Sensitivity analysis of the saturation current I_{dsat} can be expressed as [134]:

$$\frac{\delta I_{dsat}}{I_{dsat}} = \Sigma(\frac{\delta I_{dsat}}{\delta x_i}\frac{x_i}{I_{dsat}})\frac{\delta x_i}{x_i} \tag{6.1}$$

where x_i, denotes average of V_{th}, C_{ox}, L and any other parameters in interest, while δx_i denotes deviation for each parameters. The average and standard deviation of the independent SPICE parameters can be extracted from the layout aware SPICE netlist. The percentage of each component of I_{dsat} standard deviation is calculated. Based on this result, it can concluded if the I_{dsat} of specific device varies due to lithography effects (L,W) or stress effects (V_{th}, mobility) or even CMP effects (Tox).

As stress effects are one of the main causes of electrical variations in analog and mixed signal designs [15], we will further apply equation 6.1 on stress related parameters. Studies show that the width of the well proximity effects and shallow trench isolation (STI) effects induce mechanical stresses on the transistor channel, causing electrical and timing changes in the circuit performance [80] [5]. Hence, the electrical SPICE simulation of a schematic level netlist can perform differently when in a real design environment. Shallow Trench Isolation (STI) in the CMOS process induces mechanical stresses on the transistor channel. These mechanical stresses alter the transistor channel mobility μ and voltage threshold V_{th} causing deviations in the electrical performance of the transistors and subsequently in the circuit target specifications. Similarly the well proximity effects can cause changes in the doping profile of the transistors causing changes in the μ and V_{th} as well. These physical layout effects have high impact on sensitive analog designs such as current mirrors, differential pairs, amplifiers and others; causing circuit mismatches [135], DC current offsets and deviations from their original target specifications in the schematics.

Stress modeling equations are available in Berkley BSIM4 manuals [9]. The following is the V_{th} equation changes due to stress effects [15]:

$$VTH0 = VTH0_{original} + \frac{KTH0}{K_{stress_vth0}} * (Inv_sa + Inv_sb - Inv_sa_{ref} - Inv_sb_{ref}) \quad (6.2)$$

$$\Delta VTH = \frac{KTH0}{K_{stress_vth0}} * (Inv_sa + Inv_sb - Inv_sa_{ref} - Inv_sb_{ref}) \quad (6.3)$$

Assuming that mobility relative change is propotional to stress distribution. It can be described as function of SA, SB(LOD effect), L, W, and T dependence [9]:

$$\Delta\mu = \frac{KU0}{K_{stress_u0}} * (Inv_sa + Inv_sb) \quad (6.4)$$

Studying these equations closely; they are totally dependent on two different kinds of parameters:

1. Layout geometrical parameters; such as Inv_sa and Inv_sb. Where

$$Inv_sa = \frac{1}{S_A + 0.5 * L_{drawn}} \quad (6.5)$$

$$Inv_sb = \frac{1}{S_B + 0.5 * L_{drawn}} \quad (6.6)$$

2. Process model technology parameters; such as $KTH0$, $KU0$, K_{stress_u0} and K_{stress_vth0}.

The layout geometrical parameters will be extracted using the physical verification tool and the process technology parameters are available as constants in the spice model cards of the target IC technology.

As the saturation current is one of the main components in analog designs and it is highly dependent on μ, V_{th} and also saturation velocity. The following equations represent changes in I_{dsat} (using the simplified I_{dsat} equations) due to stress effects on V_{th} and mobility:

$$I_{dsat} = \frac{1}{2}\mu o C_{ox}\frac{W}{L} * (V_{gs} - V_{th})^2 \quad (6.7)$$

$$\delta I_{dsat} = \frac{\delta I_{dsat}}{\delta \mu_o} \delta \mu_o + \frac{\delta I_{dsat}}{\delta V_{th}} \delta V_{th} \tag{6.8}$$

where

$$\frac{\delta I_{dsat}}{\delta \mu_o} = \frac{1}{2} C_{ox} \frac{W}{L} * (V_{gs} - V_{th})^2 \tag{6.9}$$

$$\frac{\delta I_{dsat}}{\delta V_{th}} = -\mu_o C_{ox} \frac{W}{L} * (V_{gs} - V_{th}) \tag{6.10}$$

and δV_{th} is defined in equation 6.2 and $\delta \mu_o$ is defined in equation 6.4.

Schematic designers already have the operating point for all nodes of their circuit including transistor gate-source voltage (V_{gs}) and transistor drain-source voltage (V_{ds}). Schematic designers can pass transistor voltages and any other required electrical constrains similar to the layout constrains through the intent driven design engine. Applying these parameters to the target electrical equations will give the layout designer approximate values for electrical changes due to layout proximity effects. This is applied with no need to have a simulator on the layout side.

Using the suggested flow, and knowing the main contributor in I_{dsat} variation, a layout hint information to minimize the process effects, is provided. Speaking of stress parameters, the diffusion edges are increased in steps (Figure 6.5). Each time the diffusion edge is changed, multiple parameters in the transistor are impacted, such as Area Drain Capacitance (AD), Area Source Capacitance (AS), PD, PS, NRD, and NRS, as well as the stress and Nwell Proximity parameters (SA, SB, SCA, SCB and SCC). For each of these changes, the transistor model parameters in the SPICE netlist are updated, and simulation runs are performed to obtain the electrical variation change. In addition DRC is applied to make sure that the suggested movements do not violate any DRC rules. The diffusion edge values that correspond to the minimum change are identified as the fixing hint for the e-hotspot correction engine. Below is an example of the hint file generated:

```
####################################################################
//Cell Name:
PLL_TxVCO 1
//Layers Name
check.copy::M2_MWC_COPY = COPY M2_MSC
//First Hint
```

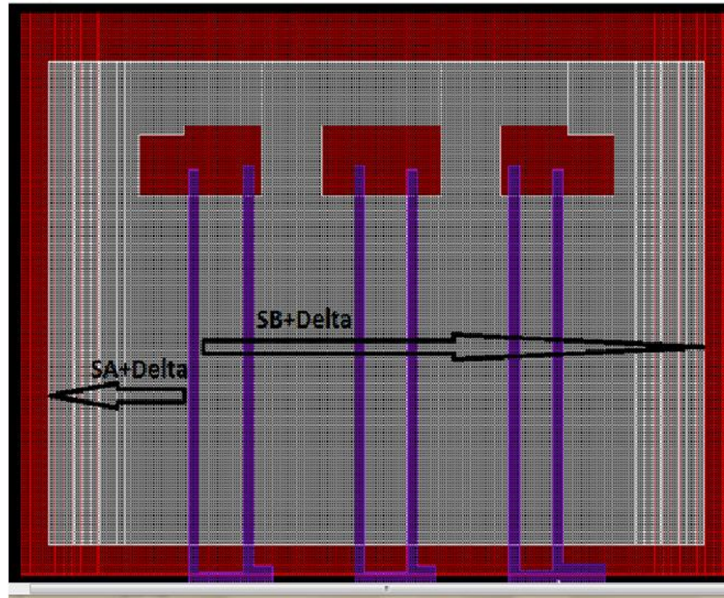


Figure 6.5: Incremental change in diffusion edges

```
p 1 4
LFD_FIX_HINTS  BEGIN
//Move Command
LFD_EDGE_MOVE_HINT BEGIN
LFD_LAYER 6 6080 8085 6080 5535
LEFT 450 LFD_EDGE_MOVE_HINT END LFD_FIX_HINTS
//New Coordinates
11460 -840   11760 -840   11760 2340   11460 2340
###################################################################
```

6.3.4 E-Hotspot Correction Engine

Analog design automation is a very important topic, especially with the rapid evolution of new technologies. Design automation helps resolve a lot of the design cycle issues that not only consume a lot of time and effort, but also get more complicated as the technology advances. Design reuse is a significant aspect of the analog design automation. ChameleonART [132], a non-optimization based design automation tool, exploits on design reuse to fulfill multiple functions, including:

- Netlist migration between a source and a target technology,
- Layout migration and compaction to a target technology,
- Processing tasks on the output layout to preserve a clean DRC profile.

The ChameleonART layout compaction engine moves all edges relative to each other, while respecting the target DRC constraints. This kind of compaction is useful for migration between different technologies, preserving the original layout intelligence inherent in the design, such as the topology, the symmetry and other design aspects. However, the design may incur a change in the layout shape and presence of void spaces due to the compaction. When used as a correction utility within the e-hotspot fixing flow, ChameleonART is limited to moving the faulty edges or implementing the user-defined fix hints, then adjusting other dependent edges. This fixing technique ensures minimum movement of most of the edges and polygons, while preserving DRC constraints and the original layout intelligence. The electrical variation flow provides ChameleonART with the layout changes that compensate for the effect of stress and lithography. These changes are defined in terms of minor device dimensions or stress dimension changes such as SA or SB in a multi-finger device. The ChameleonART tool implements these changes and adjusts any dependent edges, to preserve both a clean DRC profile and the design intelligence, while minimizing changes in all other edges and polygons. ChameleonART is characterized by its fast runtime and clean physical verification results.

The main layout correction algorithm (Figure 6.6) is described as follows:

1. The correction flow starts with the layout pre-processing stage. This pre-processing step helps identify a correct relative coverage of layers, lower number of edges, and other techniques that ensure the correct and optimum number of constraints.
2. Next, the tool generates the different types of constraints that guide the fixing process, such as target DRC constraints, device constraints, symmetry, and routing constraints. In addition, there are many directives that must be set initially for proper fixing, such as flat/hierarchical migration, use of basic/recommended DRC rules as the target DRC constraints, type of passive devices handling, and type of handle routing with different algorithms and user defined constraints.
3. Finally, correcting the e-hotspot by localized layout compaction, decreasing the number of edges, constraints, or layers. The post-processing stage generates a suitable version for the DRC/LVS verification phase.

The pseudo code of the layout correction algorithm is described in Figure 6.7

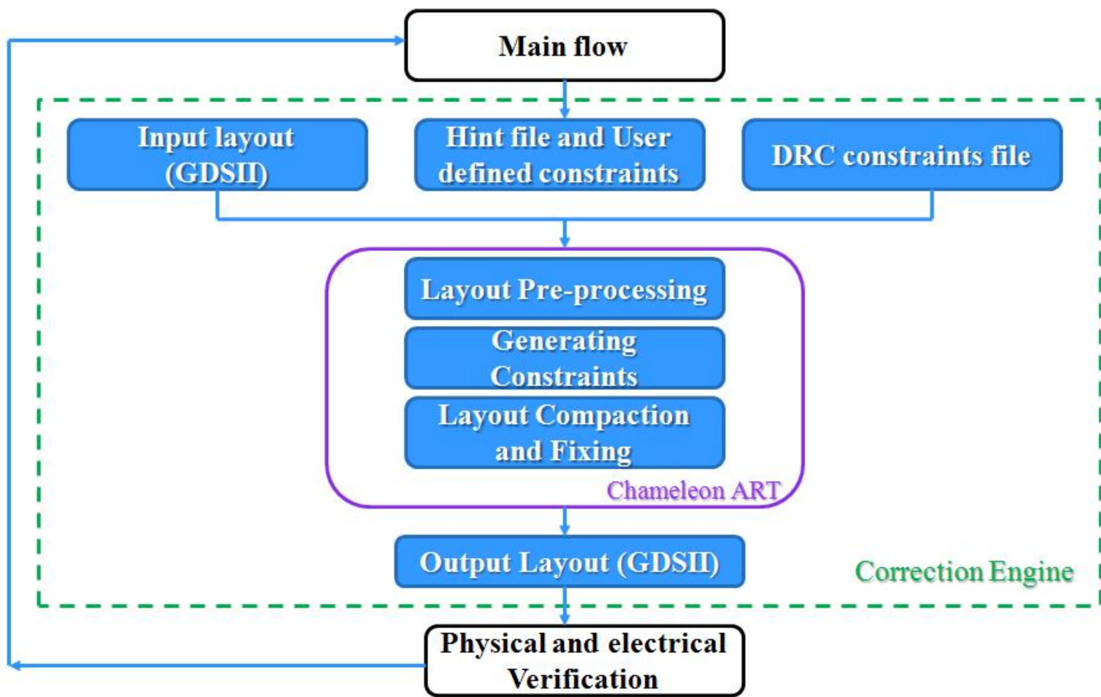


Figure 6.6: E-hotspot Correction Engine

Invoking reuse (ChameleonART tool) in hot spot mode: reuse –hotspot

1. **Exec_layer_mapping :** this command calls caliber DRC to:
   1. Perform any required mapping/adding/removing for the CAD layers of the layout, with ability to modify in datatypes, relative coverage of layers …etc (using SVRF derivations)
   2. Merges polygons of same layers. (using SVRF sizing commands)
   3. Decreases the number of contacts/vias through merging (no supported in hot spot)
2. **Load_gds: this command:**
   1. Reads the GDS format, dumps it in a text format, read all polygons, edges, vertices, SRef/Aref (arrays) ….etc.
   2. Arranges the edges and build CART database.
   3. Loads all database in the layout viewer of ChameleonART
3. **Calling _constraints_modules: it creates the database that will be used later as part of the migration engine.**
   I. Calling_DRC_constraint_generator:
      a) It uses caliber DRC to generate all the necessary DRC constraints guided by the DRC rule file to be used as part of the hot spot fixing. During the fix, the edges move to their target destinations guided by the DRC constraints or even push the nearby edge and any other dependent edges) preserving a clean DRC layout.
      b) The DRC result database is in form of error pairs.

   II. Calling Routing Handling generator:
      a) Loads the "tech file" which have data concerning the used metals, their widths, required scaling features … etc
      b) Use SVRF commands to generate constraints for the routing handling. It looks similar to the DRC constraints but handled different inside the code (as will be explained later)

   III. Calling text handling generator: (not currently supported in hot pot engine)
      a) Loads the "tech file" which have data concerning the used metals, the text/port layers attached to each metal.
      b) The preservation of text over its metal and relaive space between this text and the nearby edges are done during the hot-spot fixing phase through ChameleonART engine.

   IV. Calling cut(contact/via) handling generator: (not currently supported in hot pot engine)
      a) Loads the "tech file" which have data concerning the used metals, Cuts width/space/array sizes … etc

   V. Calling Device constraints module generator: (not currently supported in hot pot engine)
      a) It loads some sort of the netlist of the design
      b) It calls caliber LVS to verify the input layout and generate SVDB data where the IXF (instance cross reference) file will be used in a the device constraints generation

   VI. Calling Add UDC module (User Defined Constraints):
      a) Where the user through the GUI can add/remove (the removing feature is not implemented yet) DRC constraints customized for some shapes or to handling specific locations.
4. **Viewing constraints and Loading the generated constraints:** in order to debug the input constraints and evaluate it on many levels. This viewing utility has many features that help the user to select constraints per edge …. Etc
   1. For the DRC constraints:
      a) Filtered internally through specific algorithms to remove redundancies.
      b) The constraints are attached to each edge/polygon by its location and CAD layer number.
   2. For routing handling: loaded to the layout edges
   3. For cut constraints: (not supported in hot spot engine) It should generate internal constraints to be used during the hot spot fixing phase to guide the engine to preserve the number of cuts or fill the S/D by contacts …etc
   4. For Device constraints (not supported in the hot spot engine)
      a) It uses the generated IXF files in addition to the netlist to map each finger to it device knowing it finger width and length. It also maps the subcircuits.
   5. For UDC: the added constraints are attached to their edges to be evaluated and debugged.
   6. Generating other geometry dependent constraints that guide the shape preservations and the hierarchy preservation.
5. **Calling the migration/fixing algorithm:** When the migration/fixing phase starts, the engine solves the generated equations in different directions, resolve redundancies, use different algorithms to fulfill the fixing/migration function. Preserving all
6. **Output stage:** The output database is post processed by caliber DRC (using SVRF to do different functions):
   1. Cut handling (to divide the merged cuts) (not supported in the hotspot engine)
   2. The new layout is dumped in GDS format. To be used in the normal verification flow where DRC/LVS and PEX can be performed.

Figure 6.7: Pseudo Code for the E-hotspot Correction Engine

6.4 Experiments and Results

6.4.1 Silicon Wafer Full Speed Transmitter chip designed at 130nm Technology

Problem Description

An IP design house experienced a parametric yield failure in its Full Speed Transmitter, shown in Figure 6.8. Lab test results show that some of the chip running in full speed mode has a signal quality issue. The differential data outputs (DM and DP) are distorted under the 5 meters cable and sometimes on shorter cables (Figure 6.9). Investigations, simulations and lab experiments have been heavily conducted to discover the cause behind the distorted output signals in lab measurements.

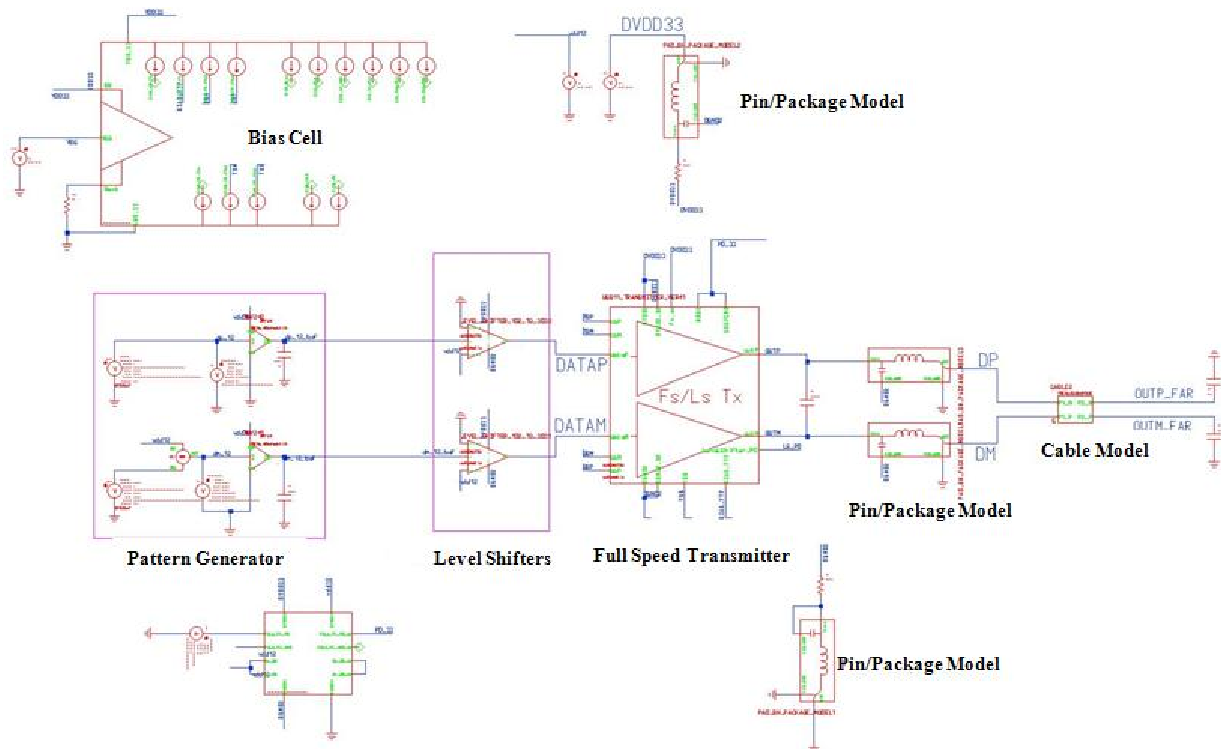


Figure 6.8: FullSpeed Transmitter testbench including supply/ground/DP/DM bonding and cable model. Cable end loaded with 50pF as for Full Speed specs.

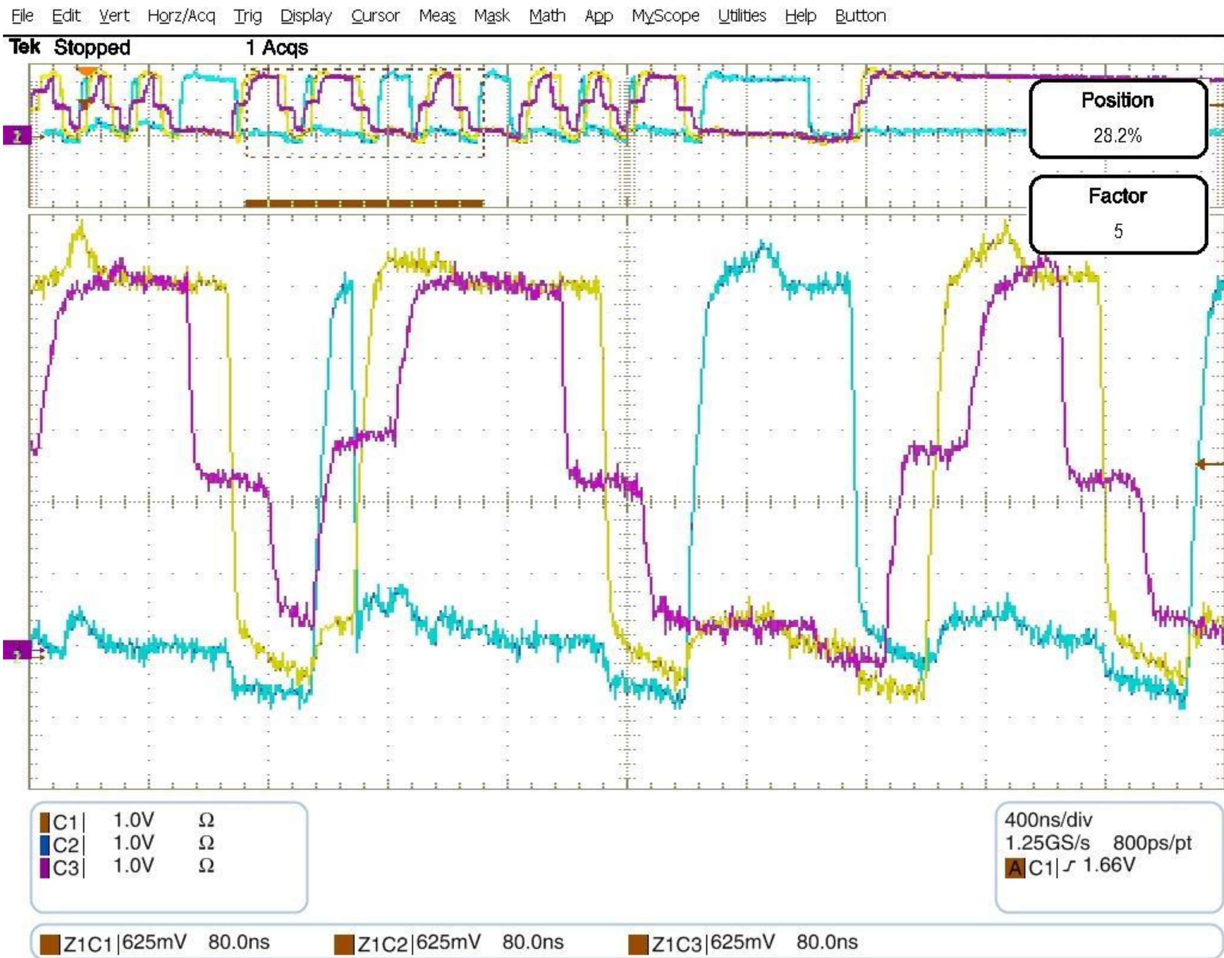


Figure 6.9: Lab Measurement: Full Speed Signals after 5m cable. DP(Yellow) DM (Blue) at end of cable. Also DP(Violet) before cable. DM is totally distorted by end of cable.

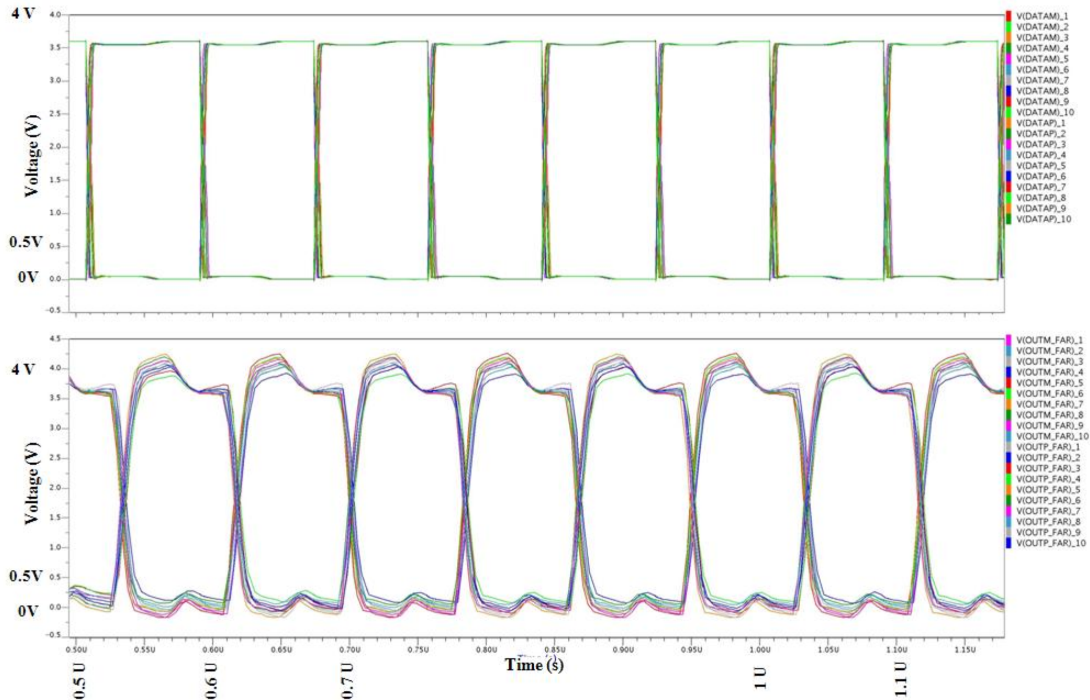


Figure 6.10: Simulation results without stress effects: a- Levelshifter DATAP/DATAM outputs b- Far end DP/DM

Investigations and conclusions from lab testing and simulations

The signal distortion issue appeared in the full speed transmitter signal outputs (DP and DM) was initiated at the level shifter output signals (DATAP and DATAM). The level shifter is located between digital core and the full speed transmitter. Each of the digital core and the full speed transmitter has different ground pins. In case of the presence of a very small delay between the DP and DM signals output from the digital core will cause disturbance to ground level, and the level shifter will amplify this delay and in the form of a duty cycle distortion that explains the lab measurements.

It is worth mentioning that the original circuit simulation results for the level shifter under different PVT (Process, Variation, Temperature) never show distortion when delay in the differential signal output from the digital core (Figure 6.10).

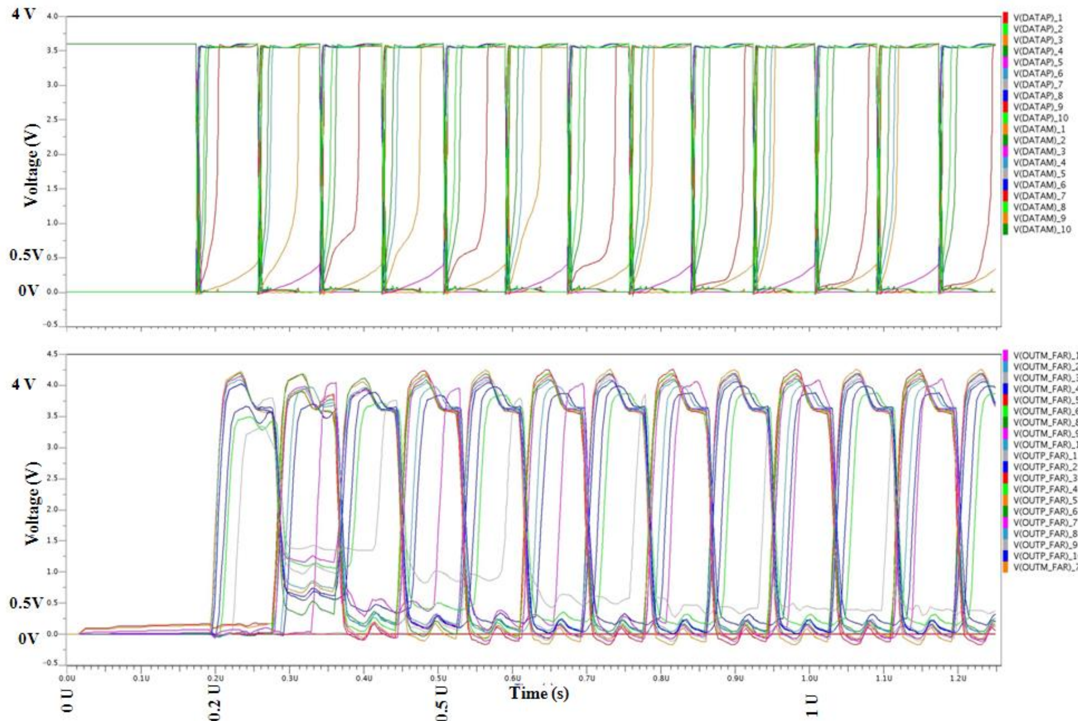


Figure 6.11: Simulation results with stress effects: a- Levelshifter DATAP/DATAM outputs b- Far end DP/DM

Running the proposed Electrical Hot Spot Detection on the level shifter

The simple description of a level shifter circuit functionality is raising the output voltage swing from a low voltage value (in this example V_{DDlow}=1.2v) to a higher voltage value (in this example V_{DDhigh}=3.3v). The basic circuit of a level shifter could be two inverters connected together however each inverter has different power and ground voltages, one use low V_{DD} and the other use high V_{DD}.

The level shifter circuit is tested using the proposed electrical hotspot detector. As a first step, the layout design context is extracted and the netlist with the actual lithography and stress parameters is accordingly updated. Electrical hot spot detection engine is applied to the different level shifter netlists (original netlist without the layout effects and the updated netlist with layout effects). From the lab experiments it seems that the level shifter has rise/fall time issues. Therefore two operating points are used to study the level shifter performance, one during the rise time and the other during the fall time.

The proposed flow is able to detect devices inside the level shifter that have parametric performance issues, as follows:

```
Case1: Rise Time
#################################################################
##############          Device Reports           ################
#################################################################
Id variation in device Mout_p: 75.183745307 %
#################################################################
##############          Device Violations        ################
#################################################################
Violation: Mout_p absolute Id variation = 75.183745307% > 5%
#################################################################

Case2: Fall Time
#################################################################
##############          Device Reports           ################
#################################################################
Id variation in device Mout_n: 0.63810392%
#################################################################
##############          Device Violations        ################
#################################################################

#################################################################
```

The highlighted e-hotspot device $Mout_p$ is the output stage device for the level shifter. The stress effects on this device impact the device's mobility, V_{th} and hence affect I_{dsat} and hence the rise time to load the following stage causing signal distortion, matching simulation results and lab measurements.

Full circuit simulations are repeated after extracting the actual layout stress effects, the simulations results (Figure 6.11) illustrated signal distortion from the level shifter output, concluding that the updated level shifter netlist (with layout stress effects) gave similar performance as the lab measurements due to the stress effects, which were not accounted during the design verification stage. This experiment illustrates the huge benefit of having this flow in the design verification stage, as an early stage detector for electrical hotspots and avoids wafer parametric yield failures. The physical layout was not shared by the IP design house, therefore the correction stage was not applied in this test case.

6.4.2 Voltage Control Oscillator designed at 45nm Technology

In another experiment, an industrial Voltage Control Oscillator (VCO) circuit, is used to detect and fix electrical hotspots. Figure 6.12 shows the VCO schematic design. The circuit designer simulated the impact of the process variation within the VCO transistors on the circuit specifications, mainly the VCO frequency response and its output voltage swing. A pre-defined values of DC current variations are defined on VCO devices as designer's constraint, to ensure that the VCO circuit meets the required specifications. The intent-driven design engine captures this information and all other electrical constraints, and links them to the schematic/layout database.

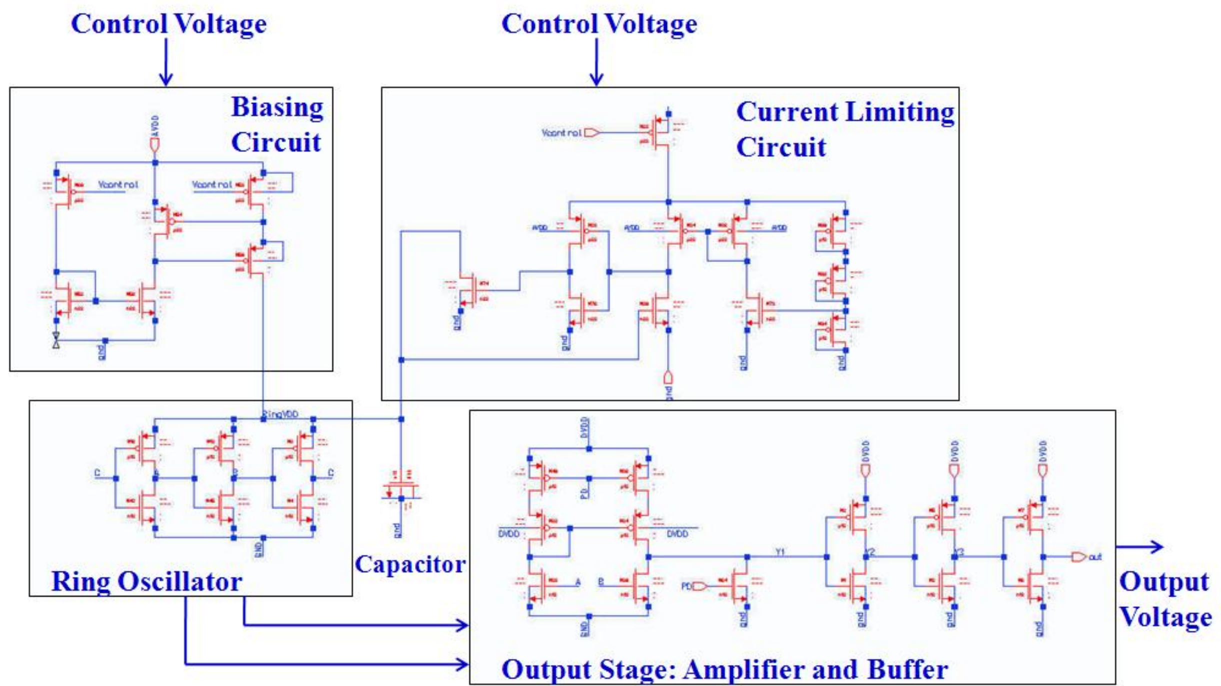


Figure 6.12: VCO schematic design

After extracting layout design context and updating the netlist with the actual lithography and stress parameters [136], the simulation results showed that one of the critical transistors that connects the biasing circuit with the ring oscillator circuit experienced a 5.5% change in the DC current due to process variations. The sensitivity analysis engine determined that the stress effect was the major contributor. Incremental changes in diffusion edges were applied on this e-hotspot device. The sensitivity analysis engine then

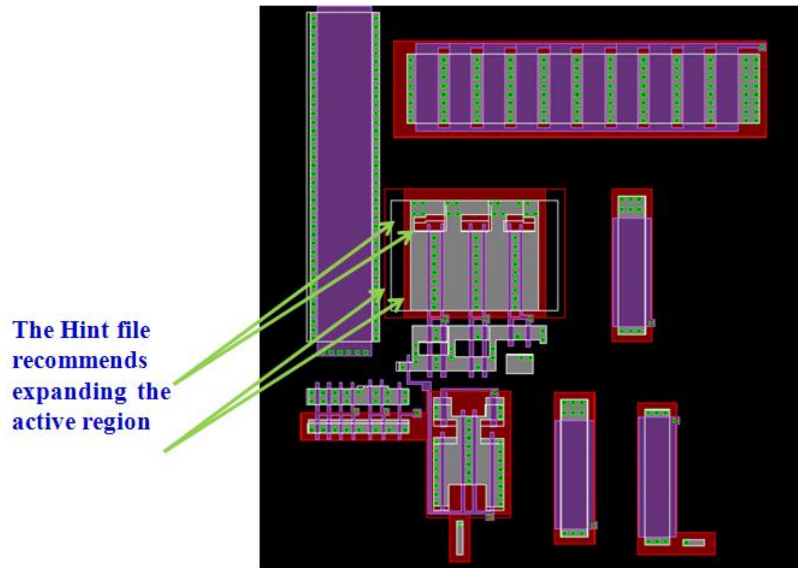


Figure 6.13: VCO layout after correction

recommended values for biasing the diffusion edges that ensured the DC current variations due to process variations would be less than the designer's pre-defined constraint, without violating the DRC rules. In this test case, the DC current variations after correction for the faulty device dropped to 0.9%. The final recommended values for biasing the diffusion edges were written in the proper syntax format used in the hint file.

This hint file is passed to the e-hotspot correction engine. As described earlier, this correction engine interfaces with ChameleonART, which in turn generates a modified VCO circuit, as shown in Figure 6.13. The corrected VCO layout was passed through Design Rule Check (DRC) and Layout versus Schematic (LVS) to verify a clean DRC and LVS layout. In addition, the corrected circuit is verified through full circuit SPICE simulation to ensure that the VCO specifications were met. Figures 6.14 and 6.15 show the VCO frequency response and output swing simulation results, respectively.

Interesting conclusions can be extracted from Figure 6.14. The physical layout design context highly impacted the design specifications, as simulation results showed that the frequency output at zero control voltage was 6.8GHz, while it shifted to 7.3GHz after layout. The proposed flow was able to decrease the design immunity to stress effects which consequently relax other PVT (process, voltage and temperature) simulations. This will allow the designer with more flexible design guard-bands which will improve the yield.

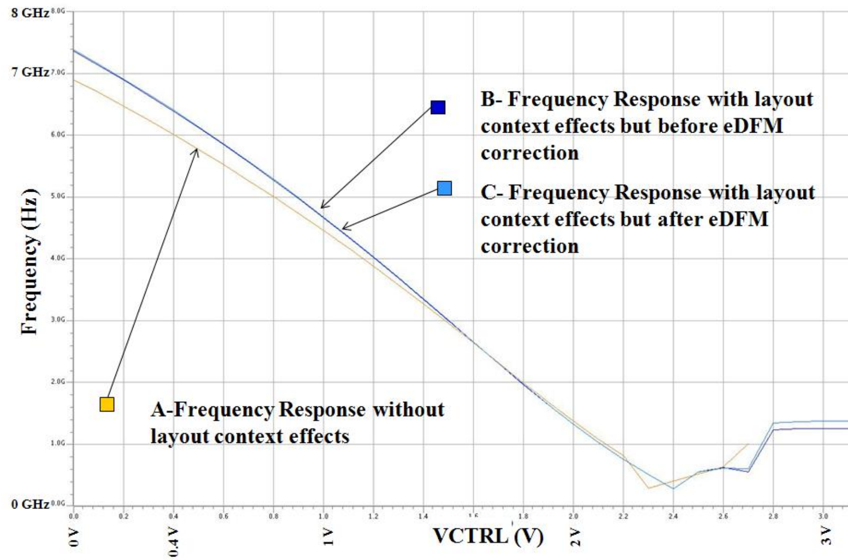


Figure 6.14: The VCO frequency response chart for different cases: a-No layout context effects b-Before e-DFM correction c-After e-DFM correction

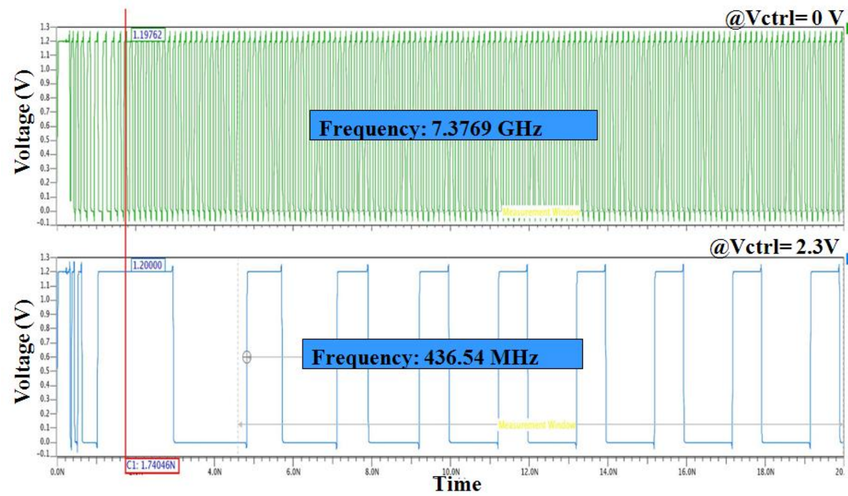


Figure 6.15: The VCO output swing chart after e-DFM correction

6.4.3 OPAMP designed at 65nm Technology

Applying the proposed flow on the OPAMP circuit, the layout designer is able to detect the change in the saturation current (I_{dsat}) within the OPAMP devices. The OPAMP is a small signal circuit, and from the schematic simulations the operating point for circuit can be extracted. The e-hotspot detection engine calculates the I_{dsat} values for the OPAMP devices with and without stress/lithography effects using lookup tables [15]. The e-hotspots devices which have variations in the saturation current I_{dsat} and violates the electrical constraints are highlighted on the layout shown in Figure 6.16. The sensitivity analysis engine indicates that the change in I_{dsat} values are mainly due to stress effects. The fixing hints recommend to bias up the active regions for the devices experiencing stress effects. The proper fix is applied to the OPAMP layout design.

To verify the design before fixing stress effects and after fixing stress effects, the full circuit netlists for these two cases are extracted and electrical simulations are applied to compare the OPAMP performance. Simulation results are shown in Figure 6.17. This simulation shows the phase margin of the OPAMP for the different cases, showing curve after fix close to designer's target curve.

Figure 6.16: Histogram for worst case changes in saturation current I_{dsat} and pointing out most affected transistors in the tested layout design. [15]

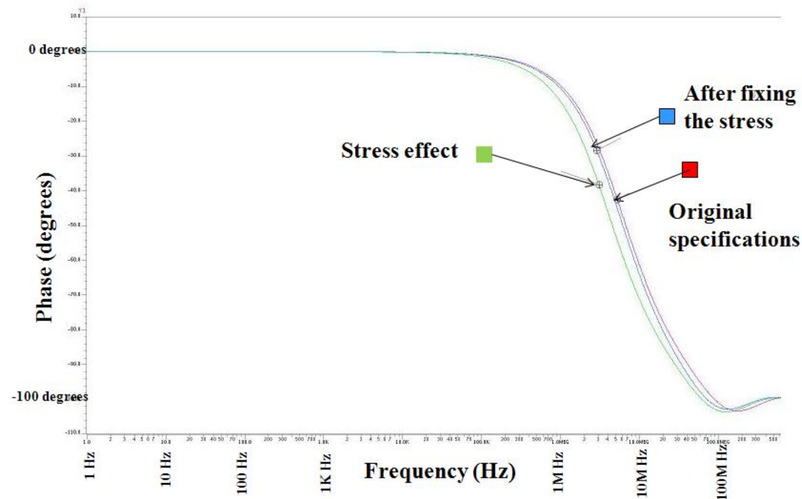


Figure 6.17: Simulation results for the Opamp analog circuit showing curve close to actual curve after fixing stress effects. [15]

6.5 Summary

The main contributions of the proposed CAD solution are listed as follows:

- The proposed CAD solution is a **complete** electrical DFM analyzer and fixing solution. This solution enables the analog designers to accurately detect, analyze, and minimize the electrical effects on design performance caused by systematic manufacturing variations which improves parametric yield on sub-65 nm designs.
- The proposed CAD solution is one of the first solutions that is dedicated to **automatically fix** parametric yield issues for analog circuits.
- This flow plugs directly into the designer's existing flows and **avoids** going through extensive post layout runtime simulations.

To validate this work, the solution automatically identified electrical hotspot devices on an industrial level shifter, VCO and OPAMP circuits. This followed by analyzing the root cause if these hotspots are lithography or stress related. The CAD solution applied the appropriate corrections to enable the design to comply with the electrical performance design constraints.

Chapter 7

Summary and Conclusions

7.1 Conclusion

As VLSI technology scales to 65nm and below, traditional communication between design and manufacturing becomes more and more inadequate. Gone are the days when designers simply pass the design tape out file to the foundry and expect very good manufacturing and parametric yield. This is largely due to the enormous challenges in the manufacturing stage as the feature size continues to shrink. Thus, the idea of Design for Manufacturing (DFM) is becoming very popular. Even though there is no universally accepted definition of DFM, the idea is to convey the manufacturing information into the design stage in a way that is understood by designers. Consequently, designers can act on the information to improve both the manufacturing and parametric yield . In this thesis, different CAD solutions are proposed to provide designers with high performance systematic variation analysis tools. In addition, these CAD solutions aid the designers with different design enhancement techniques.

At the process variation analysis level, in Chapter 3, we proposed a design context-aware and process-aware methodology to perform a quantitative study of the impact of systematic variations for different circuits' electrical and geometrical behavior. The existing simulation-based process variation analysis solutions face several challenges, mainly the high computational silicon simulation time. To address this issue, a novel solution is introduced in Chapter 4 to highly reduce the lithography simulation runtime with a minimum impact on accuracy. Another challenge for the existing simulation-based or even the pattern matching-based process variation analysis solution is the lack of the design and electrical awareness of the circuit. Therefore a high performance electrical driven hotspot

detection CAD solution for full chip design was presented in Chapter 5, using a novel device parameter matching technique. This solution delivered extraordinary fast and accurate results. Finally a parametric DFM hotspot detection and design driven analysis solution that is dedicated for analog circuits, was introduced in Chapter 6. This solution was verified with silicon wafer measurements for a level shifter circuit confirming the existence of parametric yield issues in the design.

At the design enhancement level, the proposed CAD solutions provide designers with various design enhancement and correction methods. For digital designers, a DFM-aware standard cell re-characterization flow was proposed in Chapter 4 to attain a more robust digital design. This solution was examined on 45nm FIR chip. For analog designers, an automated correction solution that is based on systematic variation sensitivity analysis was devised in Chapter 6. Several analog circuits were used for design enhancement and correction experiments.

Hopefully, the proposals in this thesis is to help designers successfully capture, understand, and fix manufacturing defects to improve both the manufacturing and parametric yield of a design.

7.2 Summary of Contributions

- Implement four CAD solutions to cover different aspects of the DFM:
  - Physical and parametric DFM
  - DFM analysis and design enhancement
- These technology independent solutions consider different design types:
  - Solutions for digital designs
  - Solutions for analog designs
  - Experiments on 130nm, 65nm, and 45nm technologies.
- New philosophies were adopted:
  - Ultra-fast electrical driven hotspot detection solution based on device parameters instead of their geometrical information.
  - Fix Before Design CAD solution: DFM-aware Standard Cell re-characterization

- Novel CAD engines which were deployed in the industry:
  - DFM-aware Smart Device Matching engine (patent pending).
  - Intent Driven Design engine (the engine is tested and deployed at On-Semi Corp. and implemented within industrial CAD tools: Calibre PERC LDL Mentor Graphics)
- Key features of the proposed CAD solutions:
  - Design intent driven solutions, electrical awareness
  - RET awareness, design context awareness, lithography and stress awareness
  - Fast runtime algorithms and avoid time-consuming simulations
  - Provide a complete DFM solution for Analog circuits: Detection-Analysis-Correction

7.3 Future Research Directions

Potential ideas could be added, as future work, to the design intent driven solutions. One challenge that can be addressed, is defining electrical constraints. Usually the designers find electrical constraints is somewhat cumbersome, as they must be able to translate the allowed variations of circuit specifications (for example: gain and bandwidth) to allowed variations in device parameters. One idea is to develop a tool based on *Symbolic Simulation* [137]. Symbolic simulation involves evaluating circuit behavior using special symbolic values to encode a range of circuit operating conditions. In one simulation run, a symbolic simulator can compute what would require many runs of a traditional simulator. Integrating such function or similar into the design intent driven system could add value to the parametric-aware DFM tools.

Appendix A

Appendix: Pending Patents from this Work

1. Working title of the invention: **Novel Technique for DFM-aware Device Matching**

   Short abstract of the invention: For technology nodes of 45nm and below; advanced device parameters extracted from the layout can have a significant effect on the on the results of the electrical simulations. Running full chip electrical simulation can be one way to accurately predict the behavior of the devices taking into account layout effects. However this approach is time consuming. In proposed method the devices that have their corresponding parameter values within a given tolerance are grouped together. The effect of the these parameters on the electrical simulations is the same for all the devices in the same group, thus grouping layout devices into sets of unique devices, in respect to these parameters, can greatly reduce the simulation time by simulating the unique devices only and mapping the electrical variation to the rest of the design devices.

2. Working title of the invention: **Electrical Driven Process Variations Checks**

   Short abstract of the invention: As IC design became a more complex process, the designer now not only takes care of the normal design and layout parameters as usual, but also needs to consider the process variation impact on his design to preserve the same chip functionality with no failure during fabrication. In the current process; schematic designers go through extensive simulations to cover all the possible variations on their design parameters and hence on the design functionality. At the

same time layout designers perform a time consuming simulations such as lithography simulations, chemical mechanical polishing simulations (CMP), critical feature analysis and then perform lithography/stress verification checks on the complete layout which is also time consuming that impacts the design turn-around-time. Our novel methodology provides a fully automated CAD flow that captures the designers process constraints from the schematic netlist; then define a prioritized list of the process sensitive devices and nets, that are then automatically linked and highlighted on the physical layout level; then followed by running verification checks using Calibre DFM, LFD suite tools avoiding the time consuming full chip simulations .

3. Working title of the invention: **A Parametric DFM Solution for Analog Circuits: Electrical Driven Hot Spot Detection, Analysis and Correction Flow**

   Short abstract of the invention: Recently, there has been an increased emphasis on parametric yield issues, referred to as electrical-DFM (e-DFM). In this invention, we present the only (up to our knowledge) complete electrical-driven design for manufacturing solution that detects, analyzes, and automatically fixes electrical hotspots (e-hotspots) caused by different process variations within the analog circuit design. Our proposed flow is examined on a 130nm, 65nm, and 45nm industrial analog circuits.

References

[1] A. Sreedhar, "Automatic techniques for modeling impact of sub-wavelength lithography on transistors and interconnects and strategies for testing lithography induced defects," Ph.D. dissertation, University of Massachusetts Amherst, 2008. xii, 1, 2

[2] P. Sharma, "Manufacturing-aware physical design techniques," Ph.D. dissertation, University of California, San Diego, 2007. xii, 9, 11, 13

[3] B. Lee, "Modeling of chemical mechanical polishing for shallow trench isolation," Ph.D. dissertation, Massachusetts Institute of Technology, 2002. xii, 15, 22, 31

[4] M. H. Abu-Rahma, "Design of variation-tolerant circuits for nanometer CMOS technology: Circuits and architecture co-design," Ph.D. dissertation, University of Waterloo, 2008. xii, 8, 15, 21

[5] P. G. Drennan, M. L. Kniffin, and D. R. Locascio, "Implications of proximity effects for analog design," *Custom Integrated Circuits Conference*, pp. 169 – 176, 2006. xii, 16, 17, 25, 113

[6] D. Abercrombie, P. Elakkumanan, and L. Liebmann, "Restrictive design rules and their impact on 22 nm design and physical verification," *Mentor Graphics Technical Library*, 2009. xii, 20

[7] N. Verghese and P. Hurat, "DFM reality in sub-nanometer IC design," *Design Automation Conference (ASP-DAC), 2007 16th Asia and South Pacific*, 2007. xii, 21, 22

[8] N. S. Nagaraj, T. Bonifield, A. Singh, U. N. Frank Cano, M. Kulkarni, P. T. Balsara, and C. D. Cantrell, "Benchmarks for interconnect parasitic resistance and capacitance," *ISQED*, 2003. xii, 24

[9] Y. Cheng and C. Hu, *MOSFET Modeling and BSIM User Guide*. Kluwer Academic Publishers. xii, xiv, 19, 23, 24, 89, 90, 113, 114

[10] C. Mack, "Modeling solvent effects in optical lithography," Ph.D. dissertation, University of Texas at Austin, 1998. xii, 32, 33

[11] *Calibre Workbench Manual*. Mentor Graphics Corp. xii, 42, 43

[12] T. Ma, "Improving design quality by managing process variability," *ISQED*, 2009. xii, xiv, 43, 86

[13] Tsien, "Context-specific leakage and delay analysis of a 65nm standard cell library for lithography-induced variability," *Proceedings of SPIE 6521, 65210F*, 2007. xiii, 44, 49, 50, 51, 61

[14] R. Martins, "On the design of very low power integrated circuits," Ph.D. dissertation, Swiss Academy of Engineering Science, 1998. xiii, 56

[15] H. M. Eissa and A. E. ElMously, "Physical aware design methodology for analog and mixed signal integrated circuits," *IEEE International Design and Test Workshop (IDT)*, 2009. xiv, xv, 104, 112, 113, 128, 129

[16] W. Evans and D. Burnell, "Deep submicron effects on data converter building blocks," *IEEE Custom Intergrated Circuits Conference (CICC)*, 2008. xiv, 111

[17] C. Robertson, "E-DFM: The future of parametric yield success," *Chip Design: Tools, Technologies and Methodologies*. 1, 102

[18] S. Nassif, "Design for variability in DSM technologies," *In Proceedings of First International Symposium on Quality of Electronic Design*, p. 451454, 2000. 1

[19] S. Borkar, "Designing reliable systems from unreliable components: The challenges of transistor variability and degradation," *IEEE Micro*, 2005. 1

[20] K. Harris, "The next generation of DFM tools: Die-level parametric variation monitoring," *ChipEstimate.com*, 2011. 1

[21] W. Poppe and L. Capodieci, "Platform for collaborative DFM," *Proc. SPIE Int. Soc. Opt. Eng 6156,124*, 2006. 2

[22] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuit and microarchitecture," *IEEE Micro*, pp. 30–39, 2006. 7

[23] B. Wong, A. Mittal, Y. Cao, and G. W. Starr, "Nano-CMOS circuit and physical design," *Wiley-Interscience*, 2004. 8, 9

[24] A. Chandrakasan, W. J. Bowhill, and F. Fox, "Design of high-performance microprocessor circuits," *Wiley-IEEE Press*, 2000. 8, 9

[25] A. Srivastava, D. Sylvester, and D. Blaauw, "Statistical analysis and optimization for VLSI: Timing and power (series on integrated circuits and systems)," *Springer*, 2005. 8, 9

[26] H. Masuda, S. Ohkawa, A. Kurokawa, and M. Aoki, "Challenge: variability characterization and modeling for 65- to 90-nm processes," *Proceedings of IEEE Custom Integrated Circuits conference*, pp. 593–599, 2005. 8

[27] S. Borkar, T. Karnik, S. Narendra, J. Tschanz, A. Keshavarzi, and V. De, "Parameter variations and impact on circuits and microarchitecture," *Proceedings of the 40th conference on Design automation*, pp. 338–342, 2003. 8

[28] K. Bowman, S. Duvall, and J. Meindl, "Impact of die-to-die and within-die parameter fluctuations on the maximum clock frequency distribution for gigascale integration," *IEEE Journal of Solid-State Circuits*, pp. 183–190, 2002. 8

[29] H. J. Levinson, *Mircolithography: Science and Technology.* SPIE The International Society of Optical Engineering, 2001. 9, 11, 19

[30] A. K.-K. Wong, *Resolution Enhancement Techniques in Optical Lithography.* SPIE The International Society of Optical Engineering, 2001. 11

[31] J. R. Sheats and B. W. Smith, *Principles of Lithography.* Marcell Dekker, 1998. 14

[32] B. Lee, D. S. Boning, D. L. Hetherington, and D. J. Stein, "Using smart dummy fill and selective reverse etchback for pattern density equalization," *Chemical-Mechanical Polish for VLSI/ULSI Multilevel Interconnection Conference*, pp. 255–258, 2000. 14

[33] B. Stine, D. Ouma, R. Divecha, D. S. Boning, J. Chung, D. Hertherington, C. R. Harwood, O. S. Nakagawa, and S.-Y. Oh, "Rapid characterization and modeling of pattern dependent variation in chemical-mechanical polishing," *in IEEE Trans. on Semiconductor Manufacturing*, pp. 129–140, 1998. 15

[34] T. Tugbawa, "Chip-scale modeling of pattern dependencies in copper chemical mechanical polishing processes," Ph.D. dissertation, Massachusetts Institute of Technology, 2002. 15, 22, 31

[35] A. Kahng, G. Robins, H. W. A. Singh, and A. Zelikovsky, "Filling and slotting: Analysis and Algorithms," *in Proc. IEEE International Symposium on Physical Design, 1998*, pp. 95–102, 1998. 15

[36] K. W. Su and et al., "A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characterization," *IEEE CICC*, pp. 245–248, 2003. 17, 23

[37] R. A. Bianchi, G. Bouche, and O. R. dit Buisson, "Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance," *IEDM Tech. Dig*, pp. 117–120, 2002. 17, 23, 38

[38] T. Hook and et al., "The dependence of channel length on channel width in narrow-channel CMOS devices for 0.25-0.13 um technologies," *IEEE Electron Device Letters*, pp. 85–87, 2000. 17

[39] G. Scott and et al., "NMOS drive current reduction caused by transistor layout and trench isolation induced stress," *IEEE IEDM*, pp. 827–830, 1999. 17

[40] T. R. Hook, J. Brown, and X. Tian, "Proximity effects and VLSI design," *ICICT*, pp. 167–170, 2005. 17, 18, 25

[41] M. Miyamoto and et al., "Impact of reducing STI-induced stress on layout dependence of MOSFET characteristics," *IEEE Trans. Elec. Dev.*, pp. 440–443, 2004. 17

[42] S. Wolf, "Silicon processing for the VLSI era," *Process Integration*, p. 389, 2004. 18

[43] T. B. Hook and et al., "Lateral ion implant straggle and mask proximity effect," *IEEE Trans. Elec. Dev.*, pp. 1946–1951, 2003. 18, 25

[44] S. Postnikove and S. Hector, "CD error budgets: Proposed simulation study methodology," *ITRS*, 2003. 19

[45] A. Wong, R. Ferguson, L. Liebmann, S. Mansfield, A. Molless, and M. Neisser, "Lithographic effects of mask critical dimension error," *SPIE Conference on Optical Microlithography*, pp. 106 – 116, 1998. 19

[46] T. Brunner, "Impact of lens aberrations on optical lithography," *IBM Journal of Research and Development*, 1997. 19

[47] A. Bourov, L. C. Litt, and L. Zavyalova, "Impact of flare on CD variation for 248nm and 193nm lithography systems," *SPIE Conference on Optical Microlithography*, p. 1388 393, 2001. 19

[48] C. Hedlund, H.-O. Blom, and S. Berg, "Microloading effect in reactive ion etching," *Journal of Vacuum Science and Technology*, pp. 1962–1965, 1994. 19

[49] S. Nassif, "Modeling and analysis of manufacturing variations," *IEEE Pmc. Of CICC*, pp. 223–228, 2001. 22

[50] R. Arunachaiam, K. Rajagopal, and L. T. Pileggi, "TACO: Timing analysis with coupling," *In Proc. Design Automation Conf.*, 2000. 23

[51] D. Wu, J. Hu, and R. Mahapatra, "Coupling aware timing optimization and antenna avoidance in layer assignment," *In Proc. Int. Symp. on Physical Design*, 2005. 23

[52] D. Wu, J. Hu, M. Zhao, and R. Mahapatra, "Timing driven track routing considering coupling capacitance," *In Proc. Asia and South Pacific Design Automation Conf.*, 2005. 23

[53] M. Cho and D. Z. Pan, "Wire density driven global routing for CMP variation and timing," *ICCAD*, 2006. 23

[54] H. S. L. et al., "An optimized densification of the filled oxide for quarter micron shallow trench isolation (STI)," *IEEE Symposium on VLSI Technology*, pp. 158–159, 1996. 25

[55] V. Moroz, L. Smith, X. Lin, D. Pramanik, and G. Rollins, "Stress-aware design methodology," *IEEE International Symposium on Quality Electronic Design*, pp. 807–812, 2006. 25

[56] L. W. L. et al., "High-performance circuit design for the RET enabled 65-nm technology node," *SPIE Conference on Design and Process Integration for Microelectronic Manufacturing*, pp. 20–29, 2004. 28

[57] C. Visweswariah, "Death, taxes and failing chips," *ACM/IEEE Design Automation Conference*, pp. 343 – 347, 2003. 28

[58] I. A. Wagner and I. Koren, "An interactive VLSI CAD tool for yield estimation," *IEEE Trans. on Semiconductor Manufacturing*, 1995. 29

[59] S. Nojima, S. Mimotogi, M. Itoh, O. Ikenaga, S. Hasebe, K. Hashimoto, S. Inoue, M. Goto, and I. Mori, "Flexible mask specifications," *SPIE BACUS Symposium on Photomask Technology*, pp. 187–196, 2002. 29

[60] A. Kahng, G. Robins, A. Singh, , and A. Zelikovsky, "Filling algorithms and analyses for layout density control," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 445–462, 1999. 31

[61] M. Fujinaga, N. Kotani, T. Kunikiyo, H. Oda, M. Shirahata, and Y. Akasaka, "Three-dimensional topography simulation model: Etching and lithography," *IEEE Transactions on Electron Devices*, 1990. 32, 34

[62] G. Nicosia, S. Rinaudo, and E. Sciacca, "An evolutionary algorithm-based approach to robust analog circuit design using constrained multiobjective optimization," *Knowledge-Based Syst*, pp. 175–183, 2008. 35

[63] M. Jack, D. Ertugrul, and O. Edward, "US20080005717," US Patent, 2008. 35

[64] S. Mansfield, L. Liebmann, A. Krasnoperova, and I. Grauer, "US20070261013," US Patent, 2007. 35

[65] T. Steven and J. Etienne, "US20077117468," US Patent, 2007. 35

[66] V. Ramnath, C. Ruggero, and R. Subramanian, "US20080013383," US Patent, 2008. 35

[67] L. Chih-Ming, L. Ru-Gun, S. I-Chang, K. Yao-Ching, and H. Cliff, "US20070266362," US Patent, 2007. 35

[68] H. Katsuhiko, "US20080003510," US Patent, 2008. 35

[69] L. Mu-Jing, "US20070300195," US Patent, 2007. 35

[70] D. Kevin, E. Robert, D. Mehul, J. Kirk, Y. Chimin, and T. Jason, "US20077284231," US Patent, 2007. 35

[71] F. Fiorante, G. Nicosia, S. Rinaudo, and G. Stracquadanio, "Automatic layout optimization of power mosfets using innovative distributed model technique," *IEEE International Symposium on Industrial Electronics ISIE*, 2008. 35

[72] S. Abdurrahman, "US20077266800," US Patent, 2007. 35

[73] W. Ting-Chi and L. Kuang-Yao, "US20077302662," US Patent, 2007. 35

[74] W. Hung-Sheng, T. Hui-Kuo, H. Chi-Yen, Y. Chi-Kuen, and M. DinGuow, "US20070277132," US Patent, 2007. 35

[75] V. Moroz, X. Xu, D. Pramanik, F. Nouri, and Z. Krivokapic, "Analyzing strained-silicon options for stress-engineering transistors," *Solid State Technology*, pp. 49–50, 52, July 2004. 38

[76] S.-M. Bae, H.-E. Kim, Y.-M. Ham, S.-C. Moon, , and S.-H. Choi, "Quantitative analysis of the proximity effect in optical lithographic process," *Proc. SPIE, Vol. 2196, 254*, 1994. 38, 42

[77] Cadence, "Increasing parametric yield," White Paper. 39

[78] B. Stine, D. Boning, J. Chung, D. Ciplickas, and J. Kibarian, "Simulating the impact of poly CD wafer-level and die-level variation on circuit performance," *International Workshop on Statistical Metrology*, pp. 24–27, 1997. 39

[79] R. Salem, H. AbdelGhany, A. Elmously, H. Eissa, and M. Anis, "Implementing a methodology for process variation awareness of design context and its impact on circuit analysis," *IEEE International Design and Test Workshop (IDT),*, November 2009. 42

[80] A. B. Kahng, P. Sharma, and R. O. Topaloglu, "Exploiting STI stress for performance," *Proc. 2007 IEEE/ACM International Conference on Computer-aided design*, pp. 83–90, 2007. 43, 113

[81] R. Fathy, H. AbdelGhany, A. R. ElMously, H. Eissa, and M. Anis, "Implementing a methodology for process variation awareness of design context and its impact on circuit analysis," *4th IEEE International Design and Test workshop IDT*, 2009. 44

[82] K. Cao, J. Hu, and S. Dobre, "Standard cell characterization considering lithography induced variations," *Qualcomm, Inc, San Diego, CA. Dept. of Electrical Engineering Texas A&M University College Station, TX presented at DAC*, 2006. 48, 61, 74

[83] M. Al-Imam, A. Torres, J.-M. Brunet, M. Fakhry, and R. Fathy, "Optimization of process window simulations for litho friendly design frame work," *BACUS*, 2006. 51

[84] W. J. Poppe, L. Capodieci, J. Wu, , and A. Neureuther, "From poly line to transistor: Building bsim models for non-rectangular transistors," *spie*, 2006. 51, 110

[85] S. X. Shi, P. Yu, and D. Z. Pan, "A unified non-rectangular device and circuit simulation model for timing and power." 51

[86] J. Rabaey and A. Chandrakasan, *Digital Integrated Circuits, A Design Prospective.* Second Edition. 52

[87] R. Fathy, A. ElMousley, H. Eissa, M. Dessouky, and M. Anis, "A DFM tool for analyzing lithography and stress effects on standard cells and critical path performance in 45nm digital designs," *5th IEEE International Design and Test workshop IDT*, 2010. 52

[88] R. Fathy, M. Al-Imam, H. Diab, M. Fakhry, J. Torres, B. Graupp, J. M.Brunet, and M. Bahnas, "Litho aware method for circuit timing/power analysis through process," *Proc. VOLUME=6521 Design for Manufacturability through Design-Process Integration, SPIE Advanced Lithography*, 2007. 55

[89] G. A. Gomba, "Collaborative innovation: IBMs immersion lithography strategy for 65 nm and 45 nm half-pitch nodes & beyond," *SPIE Conference on Optical Microlithography*, 2007. 60

[90] J.Wiley, "Future challenges in computational lithography (cover story)," *Solid State Technology*, p. 68, 2006. 60

[91] C. Albertalli and T. Kingsley, "Computational lithography (cover story)," *Semicond. Int.*, pp. 36–42, 2007. 60

[92] T. Kingsley, J. Sturtevant, S. McPherson, and M. Sexton, "Advances in compute hardware platforms for computational lithography," *SPIE*, 2007. 60

[93] S. Suh, S. Lee, K.-Y. Back, S. Lee, Y. Kim, S. Kim, and Y.-J. Chun, "Three-dimensional mask effect approximate modeling for sub-50-nm node device OPC," *SPIE*, 2007. 60

[94] J. Yang, L. Capodieci, and D. Sylvester, "Advanced timing analysis based on post-OPC extraction of critical dimensions," *DAC*, pp. 359–364, 2005. 60

[95] P. Gupta, A. Kahng, D. Sylvester, , and J. Yang, "Performance-driven OPC for mask cost reduction," *CM/IEEE International Symposium on Quality Electronic Design*, pp. 270–275, 2005. 62, 85

[96] R. Fathy and M. Anis, "An automated methodology for implementing a DFM-aware standard cell library under process window and context variations," *8th IEEE International NEWCAS Conference*, 2010. 62

[97] J. A. Torres and C. N. Berglund, "Integrated circuit DFM framework for deep subwavelength processes," *Proc. SPIE Int. Soc. Opt. Eng. 5756, 39*, 2005. 62, 73

[98] J. A. Torres and F. Pikus, "Unified process-aware system for circuit layout verification," *Proc. SPIE Int. Soc. Opt. Eng. Volume 6521*, p. 652108, 2007. 62

[99] J. A. Torres and N. Cobb, "Study towards model-based DRC verification," *SPIE Conference on Optical Microlithography*, 2005. 64, 70

[100] A. Krasnoperova, J. A. Culp, I. Graur, S. Mansfield, M. Al-Imam, and H. Maaty, "Process window OPC for reduced process variability and enhanced yield," *SPIE Conference on Optical Microlithography*, 2006. 70

[101] J. Word and K. Sakajiri, "OPC to improve lithographic process window," *SPIE Conference on Optical Microlithography*, 2006. 70

[102] J. A. Torres, T. Roessler, and Y. Granik, "Process window modeling using compact models," *SPIE Conference on Optical Microlithography*, 2004. 70

[103] S. Mansfield, G. Han, M. Al-Imam, and R. Fathy, "Through-process modeling in a DfM environment," *SPIE Conference on Optical Microlithography*, 2006. 70

[104] E. Chin, C. Levy, , and A. R. Neureuther, "Variability aware timing models at the standard cell level," *Design for Manufacturability through Design-Process Integration, SPIE Advanced Lithography*, 2010. 80

[105] H. Yao, S. Sinha, and et al., "Efficient process-hotspot detection using range pattern modeling," *Proceedings of the 2006 IEEE/ACM International Conference on Computer-Aided Design*, pp. 625–632, 2006. 83, 84

[106] P. Brooker, H. Koop, and H. Marquardt, "Using DFM hot-spot analysis for predicting resist patterns," *Solid State Technology*, pp. 47–49, 2005. 84

[107] J. K. et al., "Hotspot detection on post-opc layout using full chip simulation based berification tool: A case study with aerial image simulation," *SPIE*, 2003. 84

[108] E. R. et al., "Automated full-chip hotspot detection and removal flow for interconnect layers of cell-based designs," *SPIE*, 2007. 84

[109] M. Ma, M. Anderson, and et al., “A novel approach for hot-spot removal for sub-100nm manufacturing,” *SPIE*, 2006. 84

[110] S. Kyoh, T. Kotani, and et al., “Lithography oriented DfM for 65nm and beyond,” *SPIE*, 2006. 84

[111] F. Gennari, G. Robins, and A. Neureuther, “Validation of the aberration pattern-matching OPC strategy,” *SPIE*, pp. 444–453, 2003. 84

[112] A. Kahng and et.al, “DFM: Where’s the proof of value?” *DAC*, 2006. 85, 105, 106

[113] D. Ding, X. Wu, J. Ghosh, and D. Z. Pan, “Machine learning based lithographic hotspot detection with critical feature extraction and classffication,” *ICICDT Conf.*, 2009. 85, 100

[114] D. Reed, “Electrical-DFM: Focusing on what matters to chip designers,” *Chip Design: Tools, Technologies and Methodologies*. 85, 102

[115] J. Segura, A. Keshavarzi, J. Soden, and C. Hawkins, “Parametric failures in CMOS ICs: A defect-based analysis,” *Proceedings of the 2002 IEEE International Test Conference*, 2002. 85, 102

[116] S. Banerjee, P. Elakkumanan, L. Liebmann, J. C. M., and Orshansky, “Electrically driven optical proximity correction,” *SPIE*, 2008. 85

[117] B. Yongchan, S. Sundareswaran, and D. Pan, “Total sensitivity based DFM optimization of standard library cells,” *Proceedings of the 19th international symposium on Physical design*, pp. 113–120, 2010. 85

[118] S. Bhattacharya, N. Jangkrajamg, R. Hartono, and C.-J. R. Shi, “Hierarchical extraction and verification of symmetry constraints for analog layout automation,” *Asia and South Pacific Design Automation Conference*, 2004. 87

[119] R. Fathy, M. Al-Imam, A. Arafa, A. ElMousley, H. Eissa, and M. Anis, “High performance electrical driven hotspot detection CAD solution for full chip design using a novel device parameter matching technique,” 2011, unpublished. 87

[120] *Calibre Verification User Manual*. Mentor Graphics Corp. 91, 94, 95

[121] *Calibre Designrev Reference Manual*. Mentor Graphics Corp. 94, 95

[122] A. Arafa, H. Wagieh, R. Fathy, J. F. D. Morgan, M. Dessouky, and M. Anis, "Schematic-driven physical verification: Fully automated solution for analog IC design," *48th Design Automation Conference User Track*, 2011. 95

[123] D. Ding, A. J. Torres, F. G. Pikus, and D. Z. Pan, "High performance lithographic hotspot detection using hierarchically refined machine learning," *Design Automation Conference (ASP-DAC), 2011 16th Asia and South Pacific*, 2011. 100

[124] D. G. Drmanac, F. Liu, and L. Wang, "Predicting variability in nanoscale lithography processes," *Design Automation Conf.*, 2009. 100

[125] A. J. Torres, M. Hofmann, , and O. Otto, "Directional 2D functions as models for fast layout pattern transfer verification," *SPIE*, 2009. 100

[126] L. Lewyn, "Physical design and reliability issues in nanoscale analog cmos technologies," *NORCHIP*, pp. 1–10, 2009. 104, 111

[127] B. Wang, H. Nguyen, J. Mavoon, A. Horch, Y. Ma, T. Humes, and R. Paulsen, "Effect of layout orientation on the performance and reliability of high voltage N-LDMUS in standard submicron logic STI CMOS process," 2005. 104

[128] P. Gupta and F. Heng, "Toward a systematic-variation aware timing methodology," *DAC*, pp. 321–326, 2004. 106

[129] P. Gupta, B. Khang, and D. Sylvester, "Self-compensating design for focus variation," *DAC*, pp. 365–368, 2005. 106

[130] A. Kahng, "Key directions and a roadmap for electrical design for manufacturability," *Solid State Circuits Conference*, pp. 83 – 88, 2007. 106

[131] R. Fathy, A. Arafa, S. Hany, A. ElMousley, H. Eissa, D. Nairn, and M. Anis, "A parametric DFM CAD solution for analog circuits: Electrical driven hot spot detection, analysis and correction flow," *24th IEEE International SOC Conference*, 2011. 107

[132] S. Hammouda, H. Said, M. Dessouky, M. Tawfik, Q. Nguyen, W. Badawy, H. Abbas, and H. Shahein, "Chameleon ART: a non-optimization based analog design migration framework," *Proceedings of the 43rd annual Design Automation Conference*, 2006. 108, 116

[133] R. Salem, H. AbdelGhany, , and M. Anis, "An automated methodology for implementing a DFM-aware standard cell library under process window and context variations," *8th IEEE International NEWCAS Conference*, July 2010. 110

[134] A. Azuma, A. Oish, Y. Okayama, K. Kasai, and Y. Toyoshima, "Methodology of MOSFET characteristics flutuations description using BSIM3v3 SPICE model for statistical circuit siluations," *Int'l Workshop on Statistical Metrology Technical Papers*, 1998. 113

[135] P. B. Y. Tan, A. Kordesch, and O. Sidek, "Effect of shallow trench isolation induced stress on CMOS transistor mismatch," *IEEE International Conference ICSE*, 2004. 113

[136] S. Hany and M. Dessouky, "Operating point extraction techniques," 2011, unpublished. 125

[137] G. Gielen, H. Walscharts, and W. Sansen, "ISAAC: A symbolic simulator for analog integrated circuits," *IEEE Journal of Solid-State Circuits, VOL. 24, NO, 6*, 1989. 132