Design Techniques for Lithography-Friendly Nanometer CMOS Integrated Circuits

by

Ayman Hamouda

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

Waterloo, Ontario, Canada, 2015

Abstract

The Integrated Circuits industry has been a major driver of the outstanding changes and improvements in the modern day technology and life style that we are observing in our day to day life. The continuous scaling of CMOS technology has been one of the major challenges and success stories. However, as the CMOS technology advances deeply into the deep sub-micron technology nodes, the whole industry (both manufacturing and design) is starting to face new challenges. One major challenge is the control of the variation in device parameters. Lithography variations result from the industry incapability to come up with new light sources with a smaller wavelength than ArF source (193 nm wavelength).

In this research, we develop better understanding of the photo-lithography variations and their effect on how the design gets patterned. We investigate the state-of-the-art mask correction and design manipulation techniques. We are focusing in our study on the different Optical Proximity Correction (OPC) and design retargeting techniques to assess how we can improve both the functional and parametric yield. Our goal is to achieve a fast and accurate Model Based Re-Targeting (MBRT) technique that can achieve a better functional yield during manufacturing by establishing the techniques to produce more lithography-friendly targets. Moreover, it can be easily integrated into a fab's PDK (due to its relatively high speed) to feedback the exact final printing on wafer to the designers during the early design phase.

In this thesis, we focus on two main topics. First is the development of a fast technique that can predict the final mask shape with reasonable accuracy. This is our proposed Model-based Initial Bias (MIB) methodology, in which we develop the full methodology for creating compact models that can predict the perturbation needed to get to an OPC initial condition that is much closer to the final solution. This is very useful in general in the OPC domain, where it can save almost 50% of the OPC runtime. We also use MIB in our proposed Model-Based Retargeting(MBRT) flow to accurately compute lithography hot-spots location and severity. Second, we develop the fast model-based retargeting methodology that is capable of fixing lithography hot spots and improving the functional yield. Moreover, in this methodology we introduce to the first time the concept of distributed retargeting. In distributed MBRT, not only the design portion that is suffering from the hot-spot is moving to get it fixed but also the surrounding designs and design fragments also contribute to the hot-spot fix. Our proposed model-based retargeting methodology also includes the multiple-patterning awareness as well as the electrical-connectivity-awareness (via-awareness). We used Mentor Graphics Calibre Litho-API c-based programing to develop all of the methodologies we explain in this thesis and tested it on 20nm and 10nm nodes.

Acknowledgements

I would like to express my special appreciation and thanks to my supervisors Professor Dr. Mohab Anis and Professor Karim Karim for their tremendous support and understanding during the course of this research works. This has been a long journey with lots of challenges that would not have been possible without their understanding and support. I would like also to thank my PhD advisory committee Professor Chettypalayam Selvakumar, Prof. Siva Sivoththaman, Prof. Patricia Nieva, Professor Andrew Brzezinski and Professor Siddharth Garg for all their support.

I would like to express my gratitude to GLOBALFOUNDRIES for the financial and technical support that allowed this work to come out in its present quality after allowing me to work on the state-of-the-art technical problems and the wonderful exposure of the technical challenges and opportunities in the foundry world in deep sub-micron technologies. I'd like to thank Chidam Kallingal and Sarah McGowan for their continuous support of this work over the past few years inside GLOBALFOUNDRIES.

I am always very grateful to my wife, Doaa Soliman, for her continuous support and encouragement to have my PhD work done since day one. Also, my little angels Mariam and Youssef, I love you so much. And last but not least, to my late parents, I wish you were here beside me at this moment.

Dedication

To my late parents, I wish you were here.

To my lovely wife, without you I wouldn't have been standing here today.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

The IC industry has been successfully following an exponential trend in scaling down transistor sizes for more than 30 years [29], [30], [31]. Both the speed and density of the devices are doubling every two years, which offers a continuous improvement to the cost and mobility of the final products, where now it is possible to do full System On a Chip (SoC) for very small cost compared to its complexity. This continuous evolution opened (and still opening) the door to new markets, new concepts, and even to an evolution in our life style. A decade ago, no one would have ever expected our life as it is now. And today, no one can forecast how the world would be in ten years; Sky is the limit!

One of the main tools for this success is the advancement in the manufacturing technology especially in photo-lithography. Technologists have pushed (and still pushing) the manufacturing capabilities to the limits of the physics [32], [33]. Moreover, they have been using every scientific and technological trick to achieve what they want even when it sounds impossible. Manipulating the laws of diffraction with every smart technique made it possible for us to be developing the 20nm node using the same ArF exposure systems at a wavelength of 193 nm [3].

The aggressive scaling of CMOS technology and pushing the printing capabilities towards their physical limits to achieve deep sub-micron (sub-wavelength) designs doesn't come for free. Unfortunately, we have to deal with huge challenges in order to maintain this amazing scaling-down trend. Working at the edges of the diffraction limits makes it impossible to print the designed patterns well without going through the complicated and

computationally intensive process of mask manipulation. The effect of the surrounding designs clearly affects how a pattern is going to print on silicon. Moreover, even with perfect OPC the WYSIWYG (What You See Is What You Get) agreement between the designers and semiconductor companies is not valid anymore [34] as shown in figure 1.1 [1] and some rigorous but efficient techniques are needed to feedback such deviations to the designer for better control on performance of their design.

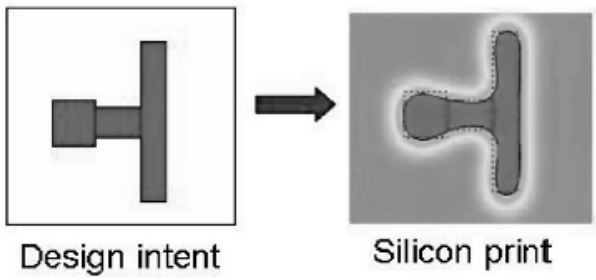


Figure 1.1: An illustration of the deviation of the Lithography to resolve the design intent without deformation. [1]

The industry is facing a huge challenge due to the process variability during manufacturing. The patterning process control is not scaling at the same pace as transistor's channel length. Such variations can be due to photo-lithography [35], Implantation, oxide thickness [4], Chemical Mechanical Polishing (CMP) [36], density related variations, as well as random die-to-die and wafer-to-wafer variations [37]. This variability issue is of great concern to the design community as it can affect the design performance and can severely affect the timing and power management of the system under design. In order to achieve and maintain the success of the digital design in the deep sub-micron technologies, it is essential to explore novel techniques to reduce and mitigate the process variations. Moreover, the diversity of the layout physical design style is becoming a huge challenge to the semiconductor industry. We are facing serious challenges, where the lithography process variations can easily cause a design catastrophic failure if not designed carefully. The design style and its compatibility to the technology used in semiconductor patterning have a direct impact on the yield. This is appearing in the increasing complexity in the physical design rules specified by the semiconductor manufacturers. Without the proper two-way feedback (the design-process co-optimization and the proper yield fixing tools) it is almost impossible to achieve the required final product quality.

1.2 Challenges and Motivation

The Patterning of semiconductor devices constitutes of many processing steps for every physically designed layer. Among these processes are the photo-lithography, etching [38], deposition and Chemical Mechanical Polishing (CMP) [36]. The accuracy and the variation control of all these processes are crucial for successful manufacturing and achieving good yield in deep sub-micron nodes. Characterization and modeling of each of these steps is very important during technology development to achieve the required CD control. In this thesis, we focus on the lithography-related patterning challenges. As we'll discuss in more details in the rest of this thesis, lithography control is a huge challenge in deep sub-wavelength technology nodes, where the final patterns CDs are almost as small as one tenth of the 193 nm wavelength used to transfer them to the wafer.

Lithography-related process variations and yield issues are growing and threatening the success of the semiconductor industry, where achieving good yield is becoming and extremely challenging task. Special attention is required to handle such variations and fix them. During the past decade there has been an outstanding progress in optical lithography, which would have never been possible without the evolution of Resolution Enhancement Techniques (RET) [21], [39] as a major player in the mask preparation technology, and all the computational lithography techniques like Optical Proximity Correction (OPC) [40], Lithography-Friendly Designing (LFD) [41], [42] and Design for Manufacturability (DFM) [43], [6], [1], [12], [44], [45], [46]. The resources and efforts needed for a technology node development increases dramatically from a node to the next. Novel new techniques for design optimization (from a lithography perspective) are required.

Many techniques are developed to improve the design immunity against process variations. Some techniques are applied during the mask tape-out flow in the Fab, while others are at the designer's side. For example, Process-Window OPC (PWOPC) and Rules-Based retargeting are both applied during the mask tape-out flow where the design layers (especially interconnects) are retargeted to be more Lithography-Friendly. As for designers, they always need to use guard-bands and statistical techniques to protect their designs from manufacturing process variability. In advanced nodes, it is necessary to use LFD and DFM tools to make sure that designs are immune to Lithography process variations. This helps in capturing the systematic catastrophic failures (opens, shorts, bad contact coverage, etc...) early enough during the design phase.

Heavy research and development have been focusing during the past several years on fixing the lithography process challenges. Some techniques include fixing the issues during the mask tape-out flow through developing more sophisticated OPC methodologies that

can handle such patterning defects more robustly. But these techniques are facing serious challenges in the 20nm and below. Other research topics focused on fixing the design as early as possible during the design phase by using different LFD methodologies, which also suffers from their extensive computation. A new methodology is needed to transform the physical design to be more lithography friendly efficiently and allow the proper feedback between the designer and the manufacturing. This new methodology needs to be capable of capturing and fixing lithography hot-spots very efficiently to handle the current weakness in the fab-designer design co-optimization needs.

It is the focus of this thesis to develop a generalized methodology for a computationally efficient model-based yield-improving retargeting during the tape-out flow. The speed requirement is essential to allow integrating it into Process Design Kits (PDKs) to allow the designers also from getting this useful feedback as very early during the design phase, which is a key element to achieve a faster yield ramp-up in advanced technology nodes.

1.3 Structure of the thesis

In chapter 2 we review the process variability challenges in semiconductor manufacturing and specifically for DUV photo-lithography. Then in Chapter 3, we review the state-of-the-art research and development in the field of computational lithography to improve the patterning quality and the overall yield using DFM techniques and reviewing the state-of-the-art OPC technology. This is followed by the summary of our very first work, where we managed to prove the feasibility of stand-alone OPC-independent model-based retargeting (AIR) in chapter 4. Then in chapter 5, we explain the evolution of the model-based retargeting into a full rigorous flow, where lithography hot-spots are captured more accurately and fixed using the LAYER methodology. In chapter 6, we introduce the concept of Model-based Initial Bias (MIB), which is a fast generic methodology that allows the prediction of an approximate post-OPC mask, which can be used to speed-up current OPC flows as well as its usefulness in improving the accuracy of model-based retargeting without having a big impact on the runtime. In Chapter 7, we explain our generalized methodology for distributed model-based retargeting, which is introduced to the first time to the best of our knowledge. In this methodology, we propose and develop the framework for distributing the lithography hot-spot problem overall adjacent designs to achieve an overall improvement in the patterning quality.In this chapter, we also summarize the potential and advantages of the work in this research in guiding the IC designers in improving the functional and parametric yield. Finally, we conclude the work and summarize the future work in chapter 8.

Chapter 2

Sources and Implications of Process Variations

The continuous scaling down of the transistor size towards the physics limit is an enormously challenging mission. Every technology node (1.5-2 years), billions of dollars are spent to achieve such difficult target. The timing and power budgets of the advanced technology nodes do not have enough margins to over-design. Also, trading off performance or area for the sake of maximizing yield might not be any better economically than having a higher performance design and accepting yield degradation.

One of the fundamental challenges is the variation control and mitigation. This problem has been associated with the semiconductor industry since its early days. It is always expected to have a degree of variability around the target specifications of the product and it needs to be accounted for in the product specifications in the form of a range of acceptable variation of each the product parameters. This variation is due to many origins among which come the tool tolerances, tool-to-tool variations tolerances, human operator error, environmental tolerances, randomness of the defects allowed, randomness in the doping profiles (RDF), variations in the photo-lithography exposure tools and resist planarity, as well as many other sources that will be discussed in details in this chapter.

The process variations affect the yield on many aspects depending on the origin of the variation and the strength of its impact. Variations could impact the functionality of the circuit if it manifests itself in a instantaneous yield limiting issue like a line short/open (due to lithography process variations or CMP variations), a missing via or extreme gate leakage (due to gate Line-End pull-back). It can also directly impact the circuit performance due to the deviation of the wafer parameters from the intended design parameters resulting in

a final product that doesn't meet the required specifications (large leakage or bad timing control). Finally, the process variation impact on yield can be delayed and appear as an aging issue (commonly known as reliability), where the process variations cause several weak spots on wafer that gradually deviate or fail after several hours of operation.

Understanding the sources of the process variations and studying their effect on the design performance is very important. It allows us to identify which process parameters are more influential on the product performance compared to less effective ones, as well as being able to identify the process parameters theoretical physical limits so that we can identify the floor of the process variation control. This better understanding also allows us to identify different techniques to mitigate such variations and to compare the improvement gain by each technique against its expected cost.

In this chapter, we are reviewing the sources of process variations in details, and how they manifest themselves and how they affect the functionality and reliability of the final product as well as their statistical nature.

2.1 The Impact of Variability on Yield and Performance

The ultimate goal of the IC industry (and any industry) is to maximize the profit. In our case maximizing the yield means that the number of fully functioning chip per wafer is maximized. By fully functioning (yielding) we mean that that a chip is both functional and meets the timing and power constraints required by the designer for their product. As we have discussed earlier, there are many sources of variability during the manufacturing process. Process variations degrade the yield on different aspects [47]. And it is a practice by the designers to allow a safety margin (guard bands) in their design so that their circuits are functioning despite all the variations it can be subject to during manufacturing.

There are multiple ways that the process variations can impact the circuit yield. It can degrade the functional yield, which is when the circuit is not electrically functioning such as in the case of opens and shorts in the circuit. This is observed as a result of some of the variability sources discussed in this chapter. Opens and shorts can happen due to random particles landing on a wire and they can be due to a lithography hot-spot (which are going to be discussed in more details in the next chapter).

Another possibility of yield loss manifests itself in the form of deviations in the transistor and circuit parameters from their nominal behavior. This deviation has to be considered

during the design or otherwise the circuit will not meet the required specifications. Unless modeled (in the case of systematic deviations), the deviations will be just widening the variance of the circuit parameters uncertainty. Accordingly, we define the term Parametric Yield as the percentage of the chips that are functioning and meeting the timing and power specifications required by the design.

The last important quality-degradation form that process variation affects is the reliability. The Reliability is defined as a measure of how long the product is going to function well without deviating away from the acceptable range of operation. For example, a reliable product would run for years without degrading in performance, while a bad one is that starts to heat up, to hang, to slow down or to totally stop functioning after a short period. Reliability is thus a quality-versus-time metric and it is usually coupled with time-dependent yield loss mechanisms such as electron migration, Time-Dependent Dielectric Breakdown (TDDB) and hot-carrier effects.

2.2 Statistical Nature of Process Variation

In this section we review the criteria used to classify the process variations. We are adopting the classification used in [1]. The process variation is classified based on its statistical nature as well as the extended range of the parameter. Accordingly, the main classification criteria are:

(I) Random vs. Systematic (II) Correlated vs. Uncorrelated

2.2.1 Random vs. Systematic.

Minimizing the variability of a process parameter to its minimum theoretical value is not always the most cost effective solution. A more practical approach is to minimize the variations to an economically acceptable range while allowing designs to be more tolerant to process variations.

Understanding the statistical nature of a process parameter is of great importance to identify how to control and mitigate the effect of its variation. If the process parameter is of random nature, then we need to identify its sources and to quantify its physical limits and to determine the cost of decreasing its variance.

Any normally distributed random process can be defined by its mean and standard deviation, which are enough to include its effect on the product performance during the

design phase using a statistical analysis technique such as Monte-Carlo simulation and analysis.

On the other hand, a systematic process parameter is deterministic and can be physically modeled through equations and simulated and fixed during the fabrication phase (or at least fed-back to the designer during the extraction phase). The only challenge is to efficiently model the physical effect using fast compact models, because an accurate (but slow) model is not very useful for production.

Examples of random process parameters are Random Doping Fluctuations (RDF), Line Edge Roughness (LER), and random particles defect. While examples of systematic variations are across wafer resist thickness variations during the spinning and coating, short, medium and long density-related variations (during CMP, photo-lithography and etching).

2.2.2 Correlated vs. Uncorrelated.

It is of great importance to determine the correlation behavior of any process parameter. The correlation behavior defines how a group of adjacent devices would jointly behave. Some parameters like the atomistic variability or Line Edge Roughness (LER) are totally uncorrelated, which means that for any two adjacent gates there is no way to know (even qualitatively) how the first behaves even if we know how the second behaves. Accordingly each device can be considered as an independent random variable when this process parameter is considered.

Spatially correlated variables are like dose and focus variations in optical lithography or like the temperature variation during the device operation. There is always a relation between adjacent gates due to the natural constraint on the maximum value for the gradient of the process parameter under study. The cross-correlation between random variables must be included during the statistical design phase so that we neither over-design nor under estimate the effects of any random process parameter on the product performance [48].

2.3 Sources of Process variations

Understanding the sources of process variability is as important as understanding their statistical nature. We need to separate and identify the underlying physical or technical sources in the fabrication-process in order to enable better process control and accordingly a better yield.

In this section, we review the main source of the yield challenges. Some of these challenges are due to the limitations of our systems as we approach the boarders of physics. Photo-lithography limitations present an example of such challenges in deep-sub-micron technologies. Other challenges are due to technical limitations because of our limited control over the fabrication process for example the random-particles defects presence during the fabrication. Figure 2.1 shows the sources of process-variability and their main categories.

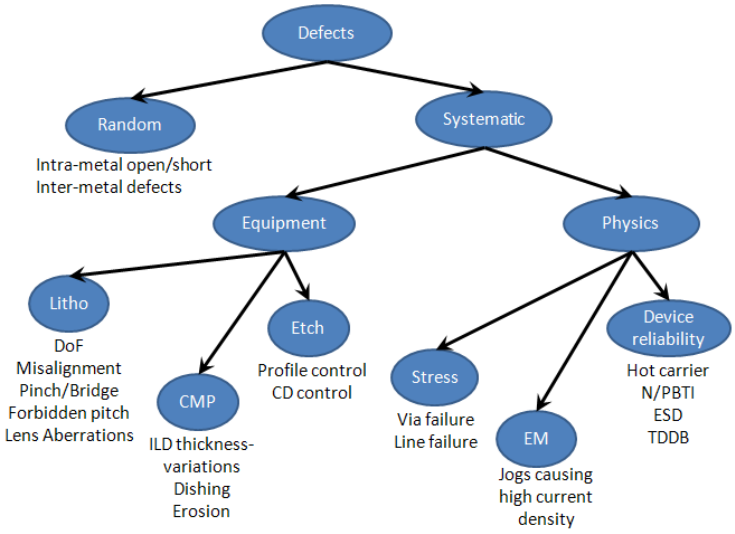


Figure 2.1: A map showing different sources of process variations and their classification.

2.3.1 Variability extension.

Wafers variability can be classified into different categories based on their extension like intra-wafer, wafer-to-wafer, lot-to-lot and fab-to-fab variability.

Intra-wafer (also known as die-to-die) variability is usually systematic in nature and it arises due to several reasons like the radial dependency of the photo-resist thickness during the resist-coating step due to the centripetal force. Wafer-to-wafer variability is also systematic and it is usually related to the order of the wafer within a fabrication lot due to process parameters drift like gases flow rate which can have an effect on the oxide thickness for example. Lot-to-lot variations are random in nature and it is caused by the drift in the fabrication tools, the very slight differences between the tools in the same fab

and also due to the human operator skill. Finally the fab-to-fab variability is obviously due to the impracticality of assuming that any two different fabs can run with exactly the same specifications even if they are running with the same process recipes, there will always be some differences that are purely random in nature.

Intra-die (also known as intra-field) variations on the other hand are a combination of systematic and random variations and can be attributed to process variability in many of the fabrication steps that are going to be discussed in the coming subsections (like photo-lithography, Chemical Mechanical Polishing, etc). The intra-die variations problem is becoming a significant source of variations that is affecting the circuits within-die [2] as shown in Figure 2.2

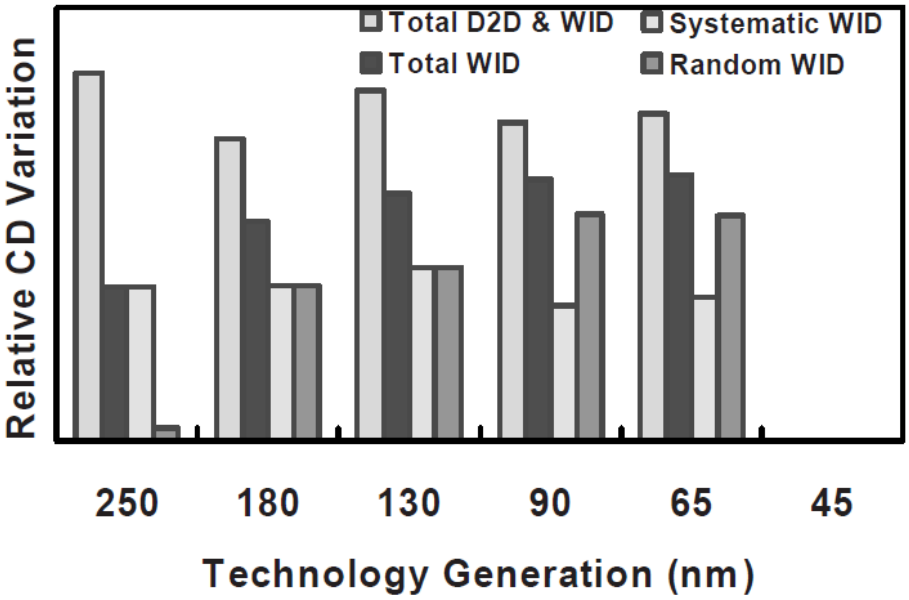
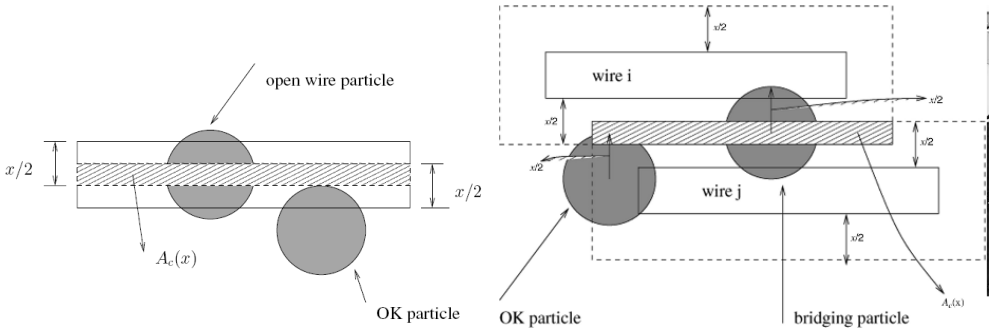


Figure 2.2: Total within-die (intra-die) contribution to CD variation is increasing each node and will be dominating in deep-subwavelength nodes. [2]

2.3.2 Random-particle defects

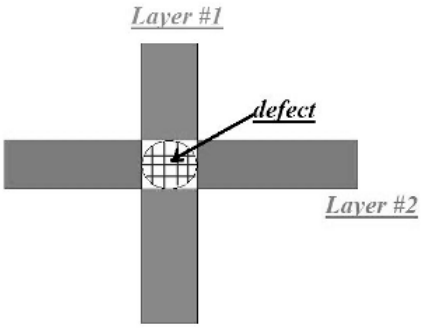
In the IC industry, it is crucial to have the wafer processing inside clean rooms. Clean rooms are fabrication rooms with very tight measures of cleanness and particles and contamination control. They are defined according to the control of the particles (defect-causing) density

per cubic meter and also according to the statistical distribution of the defects size. Ideally, we want the clean room to have absolutely no random particles flying around during the fabrication process, because whenever a particle lands on the wafer it could resolve on the photo-resist causing an open or a short or an inter-layer defect as shown in figures 2.3(a), 2.3(b) and 2.3(c) respectively. However, it is practically very hard to achieve (knowing that the ambient air outside in a typical urban environment contains 35,000,000 particles per cubic meter in the size range 0.5 microns and larger in diameter). The industry had to define a minimum accepted value for the random particles presence in the clean rooms. Accordingly, there is always a probability of such particles to land somewhere on the wafer during the fabrication of any level of the design.



(a) An open

(b) A short



(c) An interlayer defect

Figure 2.3: Examples of common problems caused by random particles defects. [1]

The statistical distribution of the random defects is known for each fab. However, the real randomness comes when we have absolutely no clue about where this particle is going to land and if it is going to cause a failure in the device (yield loss) or not. A relatively new

computational field arose to handle this specific issue; it is called Critical Area Analysis (CAA). The Critical Area Analysis is a DFM tool that measures design sensitivity to random particles and helps the designer to redesign the vulnerable designs.

2.3.3 Photo-lithography

Along with the start of the Sub-wavelength era it became clear that the industry is facing completely new challenges. For the first time since the CMOS industry started the minimum design features were getting smaller than the wavelength of light (193nm) used in the patterning process. This is where nature says its word and the laws of diffraction break the essential What You See Is What You Get (WYSIWYG) agreement between the designer and the fab. Moreover, due to the physical extension of the wavelength along several design features, the proximity effects of adjacent and surrounding designs can affect how the design is going to print. Even more, pushing the resolution to its limits (in order to print smaller and smaller nodes) makes the printability more and more vulnerable to photo-lithography process variations such as dose and focus variations.

In photo-lithography the design pattern is transferred from photomasks onto the actual wafer. Figure 2.4 [1] is an illustrative diagram of the step and scan projection lithography. The illumination source is a 193nm ArF Laser, a condenser lens on one side of the mask (which is called Kohler illumination [21]) which allows us to design the source (i.e. the wave-front of the illuminating source) and a projection lens between the mask and the wafer that is used to de-magnify the mask by a factor of 4.

As we continue to scale sub-wavelength, the lithography process cannot keep up with the scaling trend and many of the diffraction orders (mask information) are lost. The resolution limit is defined by the following equation

$$HP = k_1 \frac{\lambda}{NA} \tag{2.1}$$

where HP, λ and NA represent the half pitch (critical dimension), the wavelength, and the numerical aperture of optical system, respectively. The parameter K_1 depends on the process specifications and called the Rayleigh factor.

From equation 2.1, as k_1 gets smaller the lithography process can resolve a smaller half-pitch for the same wavelength and numerical aperture (NA), which is achieved by increasing the complexity of the process but also to the degradation of the pattern fidelity and process variation immunity. This translates to several forms of systematic distortions

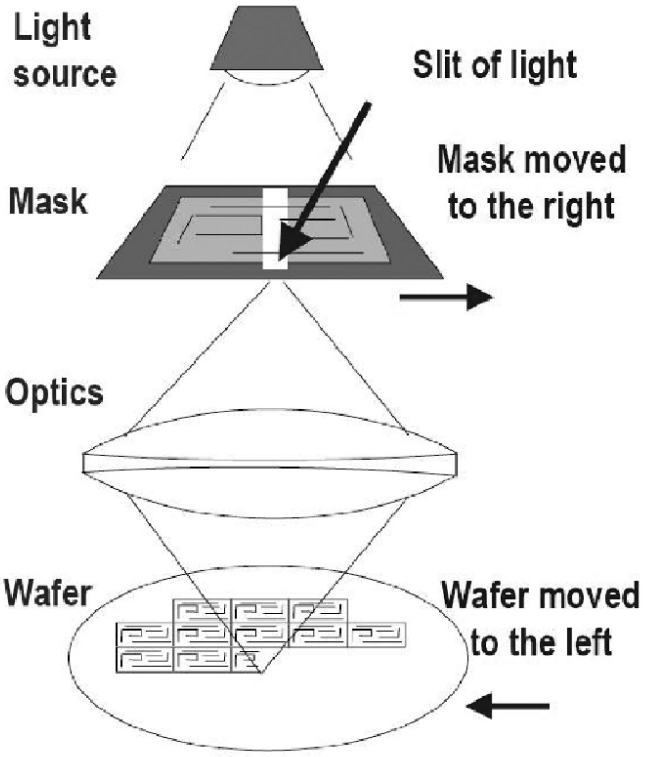


Figure 2.4: Step and scan lithography system illustration. [1]

(line-width variations, line-end shortening and corner rounding, etc). Figure 2.5 shows an example of how a pattern printability degrades due to approaching the diffraction limits [3].

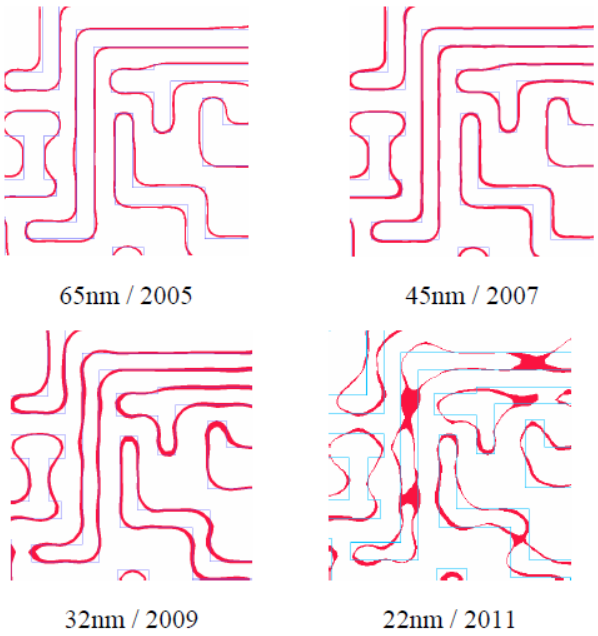


Figure 2.5: Contour bands illustrate how a first metal layout images under node-specific levels of process variability. Process variation design-immunity is shown to degrade from a node to the next until it completely fails at 22nm. [3]

The Depth of Focus (DoF) is defined as the focus variation beyond which the CD variation is beyond the specifications (usually 5% CD variation for advanced technology nodes) at constant Exposure Latitude (EL) value. The DoF is calculated as [21]

$$DoF = k_2 \frac{\lambda}{NA^2} \tag{2.2}$$

Which indicates the inverse proportionality of the DoF with the square of the numerical aperture. It also explains the sensitivity to focus variations in advanced nodes, where aggressive off-axis illumination combined with hyper-NA systems, in which the resolution is improved on the expense of the DoF.

Like any step during the fabrication of integrated circuits, it is impossible to achieve absolutely zero variability. However, it is always possible to minimize the lithographic

variations to the most economically possible and meanwhile start to find the best design practices to minimize the impact of process variations.

There are several sources of variations in photo-lithography. Among which we list

- The tool-to-tool (Photo-lithography exposure tools) difference is one important source of variations, where it is impossible to have all tools manufactured and operated exactly the same with extreme precision.
- The systematic photo-resist thickness variation that happens during the coating process [1]. This is a very small variation, at a high Numerical Aperture (NA) systems and a 193 nm wavelength, even a 10 nm will count.
- The flare [49], [50], [51] is a background noise that originates from the light scattering and depends on the design density. It affects the exposure light intensity and accordingly how much light is absorbed by the photo-resist.
- The mask errors can also affect the performance, where some deviations can happen during the mask manufacturing and could result to variability between the same designs paced at two different places in the chip. This is usually a sub-1nm but it has a reasonable effect especially in the sub-32nm nodes.
- Another very important source of variability is the multi-levels misalignment between adjacent layers in the design stack (for example via-wire misalignment). A misalignment margin has to be always included during design, but there is a random variation on how things are going to be misaligned from a wafer to another.

Figure 2.6 shows a Bossung plot showing the Critical Dimension (CD) variations at different dose values due to the photo-lithography process variation. This information is plotted in a more sophisticated manner to explain the concept of allowable process-window for a technology, where the objectives can be both the allowable percentage of through-PW CD variations as well as the acceptable resist loss and sidewall angle as shown in figure 2.7 [52]. The rectangle shown in the figure represents the maximum allowable Focus-Dose variations that still keep any CD variations and sidewall angle within the acceptable range.

In order to print smaller features (pitches) a new science was founded (borrowed from microscopy and developed further), which is the Resolution Enhancement Techniques (RET). RET is basically what enabled us to keep pushing our tools to the limits and decreasing the Rayleigh factor to reach its theoretical minimum. The Illumination is not standard illumination any more rather than pixels-based sources that are computed to fit

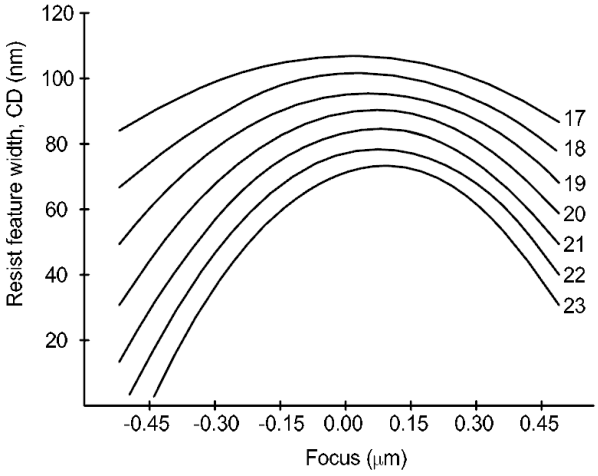


Figure 2.6: Bossung plot represents CD linewidth as a function of defocus at different dose values. [4]

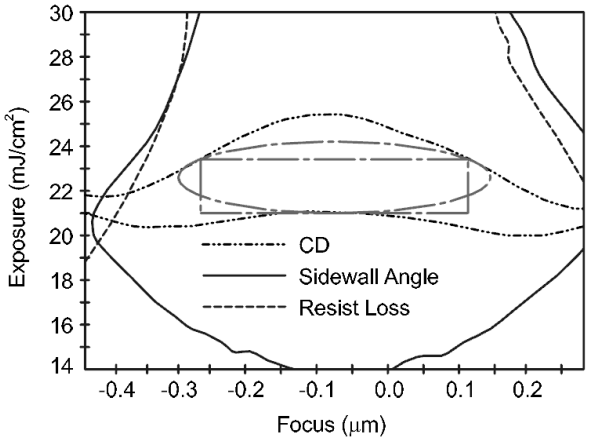


Figure 2.7: E-D process-window plots contours of constant CD, resist sidewall, and resist thickness values in the exposure-defocus space. Their overlap establishes the process-window which is given in the figure by the inscribed rectangle or ellipse. [4]

the technology requirements [53]. The masks are attenuated Phase Shift Masks (PSM) [54], [21], [55], the Sub-Resolution Assist Features (SRAFs) are a building component of any mask in deep sub-micron nodes [54], [56], [57], [58], [59], and [60], [61], and last but not least Optical Proximity Correction (OPC) [40], [62] and [63] is used to maximize the design fidelity by perturbing the mask to obtain the best printability. Figure 2.8 shows how the design fidelity can be improved by applying OPC.

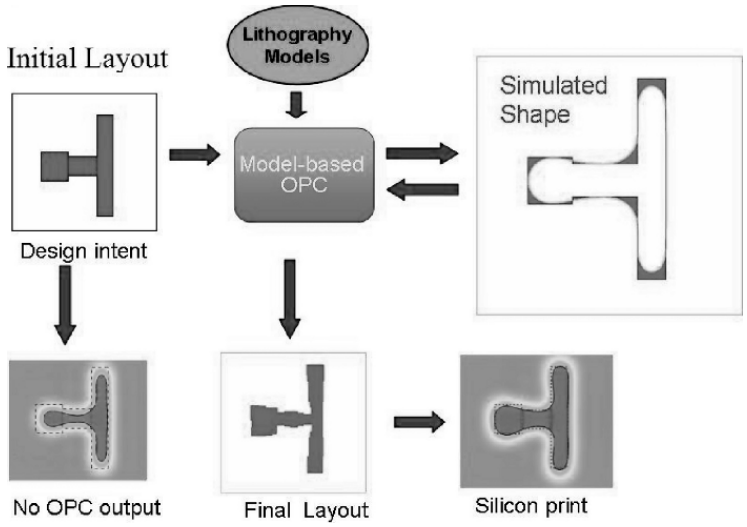


Figure 2.8: Design fidelity improvement due to OPC. [1]

Corner rounding is a phenomenon that results from the low-pass filter effect (or limitation) of the optical systems used in photo-lithography (filtering away higher diffraction orders). This loss of higher diffraction orders is equivalent to an information loss of sharp edges where the optical system cannot reconstruct the sharp edges back on the photo-resist and prints them rather corner-rounded. Figure 2.9 [21] is an example of the corner rounding of the active layer and how it can affect the gate width of the transistor. The corner rounding is not bad just because it affects the average gate width but also because it is more sensitive to dose and focus variations and accordingly it widens the variance of the gate width and not only shifts it.

Line-end shortening is another problem that can happen during manufacturing. Figures 2.10(a) and 2.10(b) show how line-end shortening can manifest itself due to either defocusing or dose variation. Line-ends patterns are very sensitive to proximity effects and process variations. This line-end shortening is always considered in the fab design rules so that the combined effect of line-end shortening, corner-rounding and any Poly-Active misalignment does not result in a high leakage of the device.

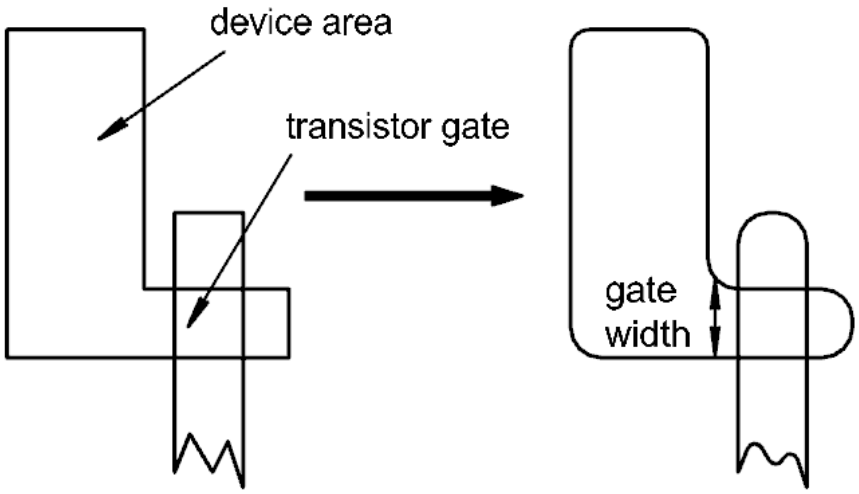
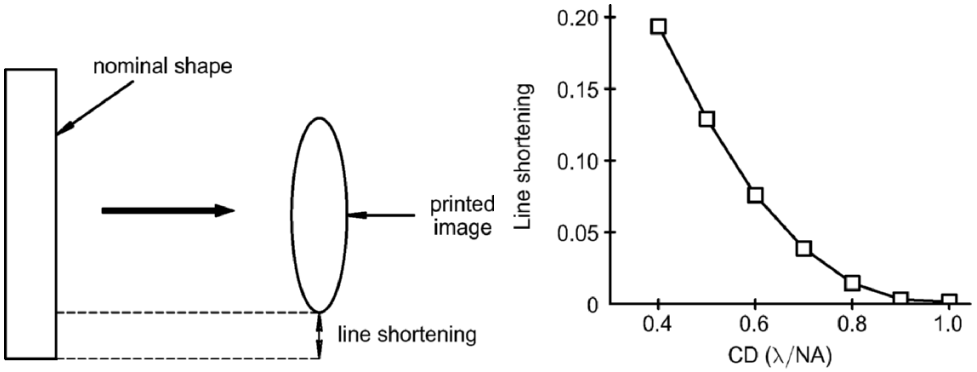


Figure 2.9: The effect of corner rounding and how it can impact the gate. [4]



(a) An illustration of Line-end Shortening. (b) Line-end Shortening variation with design CD

Figure 2.10: Line-end Shortening has a strong dependence on process variations and the design proximity. [4]

The pitch-dependent printing quality is an important phenomenon that needs careful attention. There is always a maximum defined deviation for the gate length across the chip that the manufacturers should never exceed and this is called Across Chip Linewidth Variation (ACLV). However, even when this is met, the variance of the line-width is going to be dependent on the pitch used, where some pitches are better than others. To solve this issue, the designers insert dummy poly to allow a homogeneous pattern and get the best performance by reshaping the diffraction orders to the optimum case [5] as shown in figure 2.11.

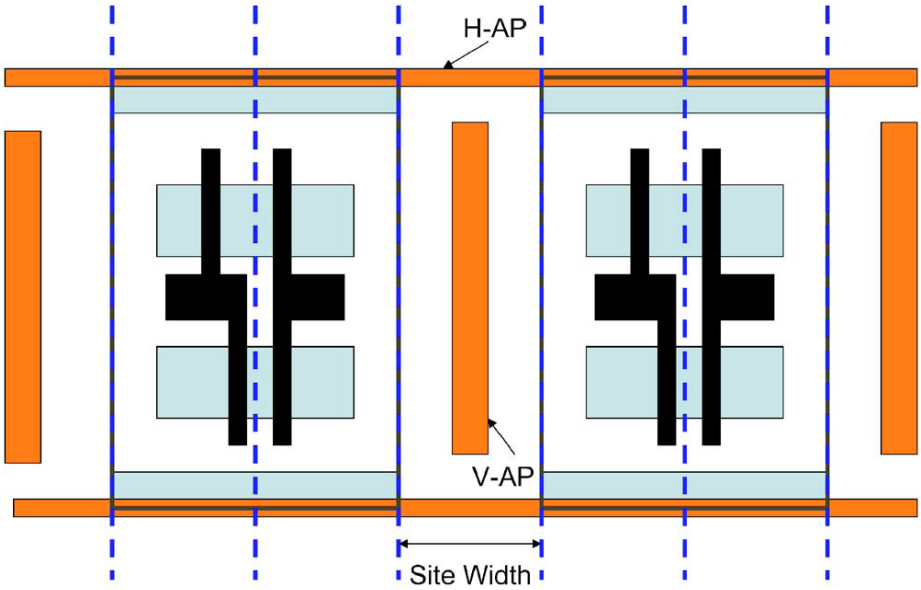


Figure 2.11: Inserting Auxiliary patterns to improve design immunity to process variations [5].

Although technically Line Edge Roughness (LER) is not a result of the optical lithography solely rather than being a byproduct of many other sources, we are mentioning it here because it is still strongly tied photo-lithography and due to its pronounced effect on the device performance. It is observed that line edge roughness (LER) contributes more significantly to the channel-length variability of nano-CMOS devices. A higher device OFF-current (I_{off}) variability, as shown in Figure 2.12. This is a totally undesirable feature (especially for low-power products like cell phones). The reasons for the increased LER in the deep sub-micron nodes processes include the random variation in the incoming photon count during exposure and the contrast of light intensity, as well as the absorption rate, chemical reactivity, and the molecular composition of resist [4]. Figure 2.13 shows

the randomness of the line edge through several steps of the via hole fabrication process.

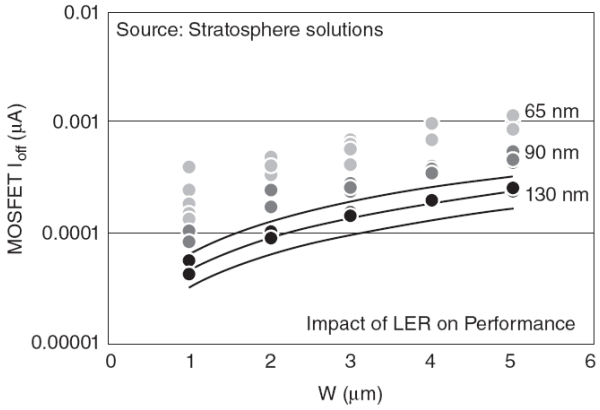


Figure 2.12: Ioff variability as a result of LER. [6].

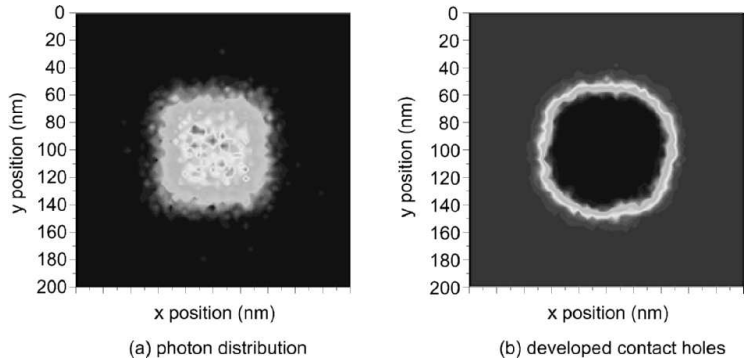


Figure 2.13: Simulation of the exposure and development of a via hole with extreme ultraviolet lithography. [7].

2.3.4 Etching

The Etching process is an essential step in the manufacturing of integrated circuits. The patterned photo-resist acts to transfer the same pattern to the wafer during the etching process, where the hard (etchant-stopper) resist prevents the underlying layer from being etched and allowing only the exposed areas to be etched. Since many nodes now, Reactive Ion Etching (RIE) has been the etching technology used in IC manufacturing. In RIE

the reactive ions are accelerated to bombard the wafer surface to chemically react with the material to be etched and evaporating due to the high energy used in the process as shown in figure 2.14 [38], [8], [64], and [65]. RIE proved to be an accurate and efficient technique that fits the needs of the IC manufacturing process from the process control and high etching aspect ratio and anisotropy.

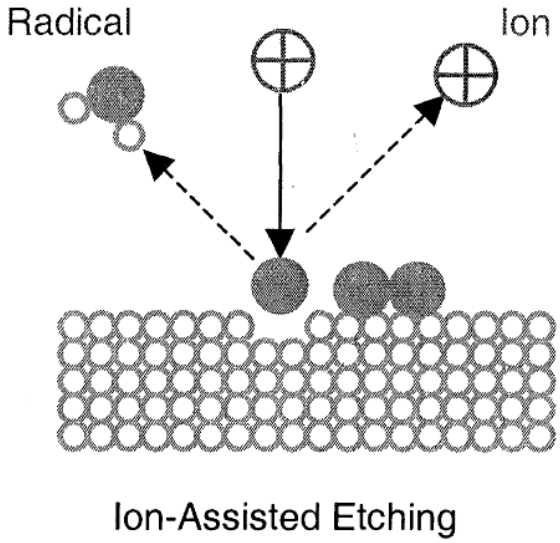


Figure 2.14: An illustrative diagram for RIE. [8]

However, the etching process (like any other) has its limitations. There are several factors that affect the quality of the etching and accordingly the accuracy of the printed feature. The impact of etching non-uniformity on the line-width accuracy is comparable to that of the photo-lithography. Etching non-uniformity manifests itself as variability of etching bias (which is the difference between the photo-resist and the etched poly-silicon or oxide) and accordingly affecting the CD Uniformity (CDU).

The most pronounced physical sources of etching bias variation across the chip are the aperture effect [38] and the micro- and macro-loading effects. The aperture effect is simply related to the size of the etching aperture (photo-resist opening size), where the time-varying etching rate depends on the aperture size as shown in figure 2.15 , where in small trenches the passivation layer is not uniform on the side walls and can reduce the etching rate inside small trenches. Both micro- and macro-loading are density-related effect, i.e. depends on the density of the design features per unit area. The etching rate depends on the density of the feature. The main difference between micro- and macro-loading effects is the range of the effect. The micro-loading is a short range effect (extends

from one to hundreds of microns), while the macro-loading can extend across the whole wafer.

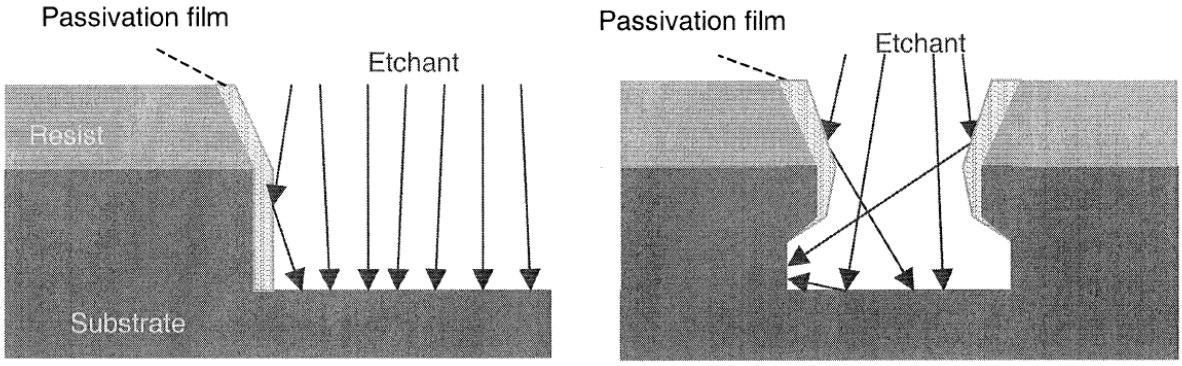


Figure 2.15: An illustration of the root cause of the aperture effect in etching. [8]

Aperture effect and micro-loading are design-dependent systematic variations that can be modeled and corrected for during the mask preparation flow. There has been some research and work done on producing compact etching models and a correction methodology for the etching process to minimize any systematic deviations from the original design [8]. However, macro-loading, which causes die-to-die systematic variation that cannot be mitigated during the mask manufacturing.

Other sources of variability in etching are related to the technicalities of the etching process and the resulting etched profile. Figure 2.16 [9] shows different etching profiles depending on the underlying technical effect such as the kinetic ion and neutral fluxes, the electron shading effects and the transport and depletion of the chemical etchant.

2.3.5 Chemical Mechanical Polishing (CMP)

CMP is a hybrid planarization process used in Integrated Circuits manufacturing for smoothing and flattening the wafer surfaces with the combination of chemical and mechanical forces during the back-end processing (metal wires). In interconnects, the main purpose of CMP is to completely remove the extra copper outside the trench area or extending higher than the trench height, leaving a flat surface that is nearly coplanar with the top of the surrounding dielectric and parallel to the wafer surface. Figure 2.17 shows an illustrative diagram of the CMP processing.

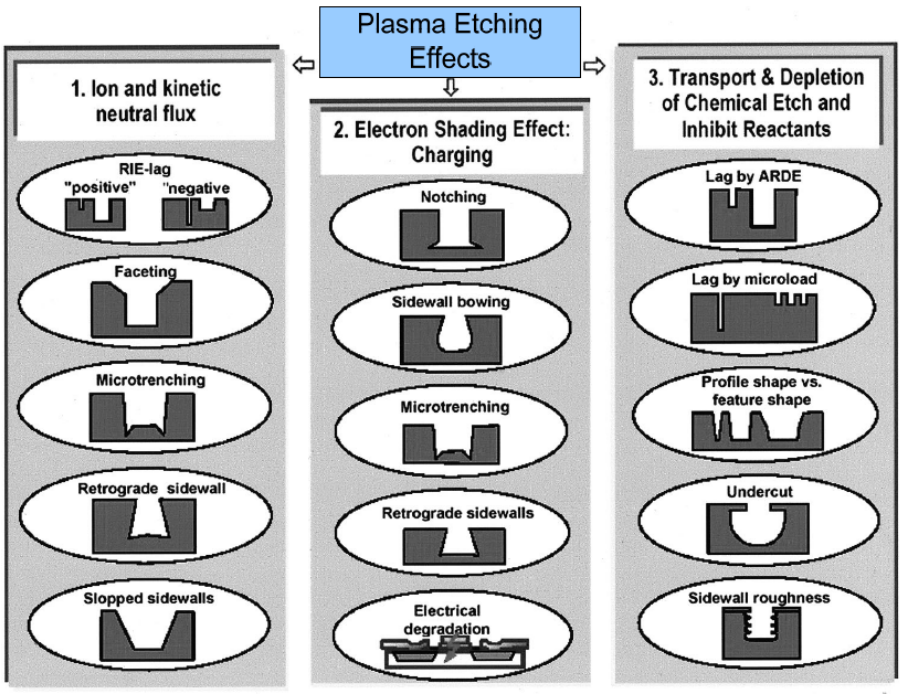


Figure 2.16: RIE non-uniformity issues. [9]

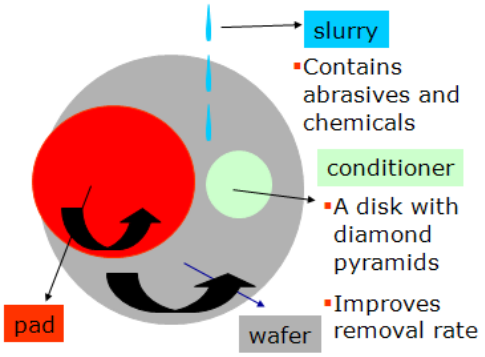


Figure 2.17: An illustration of CMP process.

Surface planarization is essential for the subsequent photo-lithography steps (i.e. vias and metal wires above). This is mainly due to the limited the Depth of Focus (DoF) [66] in photo-lithography (especially in advanced nodes), where the DoF in some cases (some design patterns) might not exceed a 100nm.

Unfortunately, CMP is not a variation-free process, where the metal density (i.e. the design) is a major factor in the polishing rate due to the difference of the mechanical strength between the copper and the oxide. Accordingly, CMP is also a significant source of systematic variations. The variations in CMP arise from two basic sources, which are the metal density and the process conditions (variations in down force, rotational speed, pad conditioning, and temperature, etc).

CMP is known to suffer from two important pattern dependent non-idealities, which are dishing and erosion as shown in figure 2.18 [10]. Dishing happens when there is an over-polishing within a feature with respect to the surrounding dielectric surface. While erosion happens when there is a loss in height of the surrounding dielectric compared to the wafer surface.

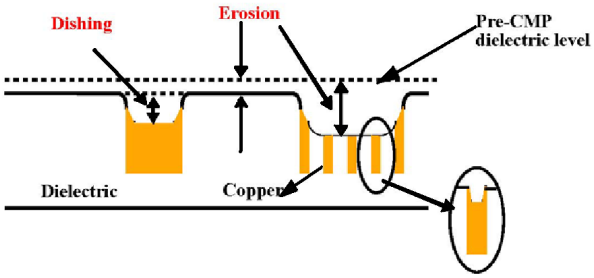


Figure 2.18: CMP Metal Dishing and Dielectric Erosion. [10]

The amount of dishing and erosion strongly depend on the designed pattern. Figure 2.19 [10] shows an example of the post-CMP surface with a metal wires density transition.

The systematic nature of the pattern-dependent CMP variations made it appealing for the fabs to do computational process correction and to insert metal fills and dielectric holes in the dielectric and metal wires respectively. These fills and holes have to be well designed to minimize the density variations across the chip. Moreover, it is essential to feedback the effect of the fills and holes on the coupling capacitance and the wire resistance to the designers to include them during the design phase.

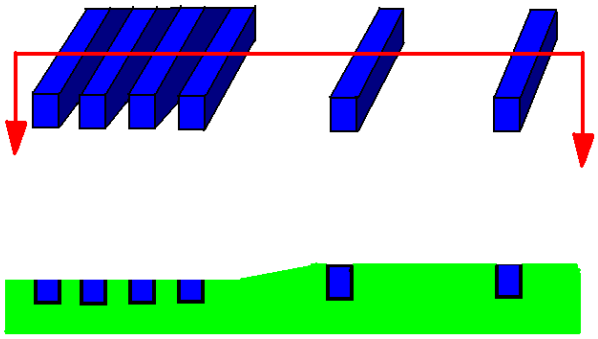


Figure 2.19: Density-dependent thickness variations during CMP.

2.3.6 Stress

One of the techniques for enhancing the mobility of electrons and holes in the FET is by intentionally creating mechanical stress in the channel. This is engineered by adjusting the mismatch in the lattice structure at the interface between the bulk silicon and the induced material - namely SiN, SiC, or SiGe. There are both compressive and tensile strain techniques depending on the type of the device we intend to enhance. The Mobility of PFET and NFET carriers has been reported to be enhanced by as much as 50% due to mechanical stress alone [1]. Among the techniques to introduce uni-axial stress is the Embedded SiGe (eSiGe) technology and the parasitic stress from Shallow Trench Isolation (STI). The strain-induced variability is also systematic as it depends on the shape and size of the transistor and its surroundings. The size of the active area, the distance between the gate and the STI have an apparent effect on amount of the stress and accordingly the carriers mobility.

2.3.7 Atomistic variability

The IC technology has reached an incredibly excellent precision in terms of the control of the fabrication parameters. However, as we go deeper in deep-sub-micron technologies we are facing a new type of challenge that we never faced before. This is the atomic-level precision, where the gate length is composed of a few tens of silicon atoms, and the oxide thickness is literally a few atoms in thickness. This means that we have reached a state where each atom counts!

Silicon dioxide has been used as a gate oxide insulating material for several decades. As transistors scale down, the thickness of the silicon dioxide insulator has steadily decreased to increase the gate capacitance and accordingly increase the drive current of the transistor and device performance. In deep sub-micron technologies the oxide thickness has reached a very small value (a few silicon dioxide molecules in size), which makes the control of the thickness variance very hard. The Si-SiO2 interface has the standard deviation on the order of $2A^o$ [4]. The thickness of oxide film of $10A^o$ corresponds to approximately five atomic layers of SiO_2, while the thickness variation is 1-2 atomic layers. This causes a random variance of the oxide thickness along the gate, which accordingly results in a random variance in the threshold voltage Vt. It also directly affects the current driving capabilities of the devices and consequently affects the ON current. It also can affect the gate leakage current, where the tunneling current varies exponentially with the barrier thickness.

High-K materials were introduced to fix the gate tunneling current issue [67], [68]. The gate tunneling current is mainly to the decrease of the insulating dielectric material width to very small values where the quantum mechanical tunneling starts to take effect through the potential barrier caused by the dielectric. High-K materials are introduced to increase the capacitance of the gate without decreasing the width. From the process control point, introducing the high-K material with a larger width keeps the limits away from the dangerous atomistic variability zone. However, there are several challenges that semiconductor manufacturers are working on such as the challenges of the deposition control of the high-K material compared to the tight control of the thermal growth of the silicon dioxide. Also it is heavily studied now whether the gate conducting material and gate fabrication order should change corresponding to the new high-K materials.

The other significant atomistic effect that appears during manufacturing is the Random Dopant Fluctuation (RDF) [69]. It is due to the extreme shrinking of the technology where the gate size becomes very small and accordingly the number of donor/acceptor atoms doped under the gate is very small to the level where the discrete-nature of the atoms starts to appear. This is very obvious in deep sub-micron technology when the number of dopants per channel is in the order of tens to a few hundred atoms. Figure2.20 [11] shows the Number of dopant atoms in the depletion layer of a MOSFET versus channel length, it is obvious that the variance of the dopant count in the channel increases as the channel length decreases [70], [71]. This dopant fluctuation has the impact of adding a random variation to the threshold voltage Vt . In newer technologies like in Fully Depleted SOI (FDSOI) [72], [33] and FINFETs [73], [74] it is investigated to use completely intrinsic channels to avoid RDF, where the doping has not much effect on these devices performance [75].

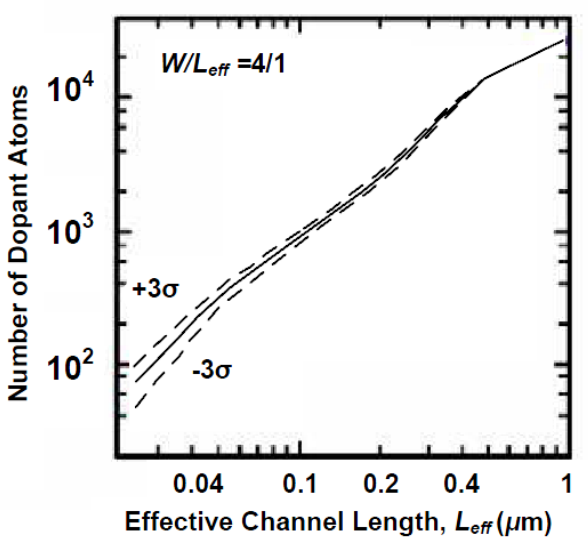


Figure 2.20: The doping concentration variance increases as the technology shrinks. [11]

2.4 Process Variations Control and Mitigation

The compensation and control of process variability is becoming a very important research (and industrial) topic. It is very important to imply our understanding of each individual source of variations and find the means to improve the process (from the manufacturing perspective) and to develop new methods to mitigate the process variations and avoid its impact on the device and circuit functionality and performance.

Process variation mitigation and compensation is considered on several levels. First, applying smart mitigation techniques on the circuit and architecture levels like using the Adaptive Body Bias and Statistical Static Timing Analysis respectively. Second, during manufacturing, where new compensation techniques are used to monitor the in-line process and compensate for any deviations like for example the application of dose-mapping [76] during optical lithography. Finally, the application of etch and optical proximity correction during the mask tape-out flow to mitigate diffraction effects during optical lithography (due to its importance, this topic will be covered in details in the next chapter) including the application of re-targeting for yield and Process-Window OPC.

2.5 Summary

In this chapter, we have reviewed the sources of variations in Deep Sub-micron CMOS technologies. We have shown how the process variability is getting more pronounced with the technology scaling. Process variations are deeply impacting both the functional and parametric yield. The variability problem cannot be overlooked or simplified to be only considered into the designer's guard band, but rather it needs to be thoroughly studied. Systematic deviations need to be modeled using fast and accurate models so that their effect can be practically implemented inside the designer's environment. The random variations impact on the wafer results needs to be quantified as a probability density function for each process parameter and to know the correlation distance (extension) of each process variation source. The photo-lithography is already a significant source of intra-die variations, where its effect extends between causing functional yield issues (like opens and shorts for non-Lithography friendly designs) to real parametric yield problems when large deviations from the design-intent can be observed.

Chapter 3

Design for Manufacturability and Optical lithography

3.1 Introduction

As the technology scaling continues into deep sub-micron technology nodes the Design for Manufacturability (DFM) becomes an essential practice. DFM covers a wide variety of design techniques and methodologies that focus on enabling better yield through applying the understanding of the manufacturing process strength and weakness to achieve a high-yield final design. DFM includes handling systematic and random process variations.

Optical Lithography is strongly tied to DFM, as it is a strong contributor to the process variations during manufacturing. The ultimate goal of optical lithography is to transfer a robust image of the IC design to the wafer. This is a complicated process which involves light source design [77], optical system design with a maximized Numerical Aperture (NA), imaging system performance maximization (minimizing aberrations) and last but not least the photo-resist stack design to maximize the light absorption while minimizing the defect printability issues. This spectacular process has been the main enabling tool that allowed the IC industry to keep up with the amazing scaling trend over the past three decades.

As the industry reaches the limits of lithography, the pattern transfer quality is degraded. At small dimensions, the loss of image quality in optical lithography erodes the design-to-wafer fidelity on silicon, and the manufacturers are not anymore able to hold to their side of the WYSIWYG agreement. The main reason behind the pattern transfer fidelity degradation is due to the continued use of the ArF (193 nm wavelength) sources

because of the technical difficulties of producing a shorter wavelength sources. To extend the lifetime of optical lithography, integrated circuit (IC) manufacturers have been seeking all the possible techniques to enhance the systems resolution, which include improving the illumination, and compensating for all pattern transfer non-idealities at the mask level.

3.2 Design for Manufacturing (DFM)

DFM emerged as a new methodology for yield, manufacturability and profit maximization. The way that DFM [78] functions is to allow early manufacturing process information to the designers to allow them to choose and alter their design practices to reach their best area-versus-yield and performance designs using the model shown in figure 3.1 [12] for fabless companies. This information can be in the form of a set of design rules that the designer is advised to follow to maximize their yield. This is a different set of rules in the Design Rules Checks (DRC) that the designers must obey. The DFM rules can take the form of recommending design pitches over others and recommending some design patterns over others. A simultaneous DFM technique is the Lithography-Friendly Design (LFD) tools [16], [79], [80], where the fabs supply the design community with process simulation tools that can highlight weak design practices (resulting in lithography hot-spots) and suggests correction mechanisms to them. Also, Lithography-Friendly Routing (LFR) [81], [82], [83] where the LFD tools are used to allow automatic re-routing if a lithography hot-spot is discovered in the routing levels. Accordingly, DFM goal is to maximize the ramping up of the yield learning curve as shown in figure 3.2 [13].

3.2.1 Recommended (Restrictive) Designs Rules (RDR)

Recommended Design Rules (RDR) is a recommended set of design rules that are agreed to improve the yield. In contrary to DRC these rules can be violated but the designer knows in advance that this comes on the expense of yield degradation. An example of RDR is the recommendation of a range of favorable design pitches which are known to have better control and suffer less from process variability (due to forbidden pitch for example [84]). Another example (which is now a standard practice in standard cell design) is the use of a dummy poly in between of standard cells to keep the cell spatial frequency homogeneity [5] as shown in figure 3.3, where a design that is more one-dimensional is more immune to process variations. Also, recommending the gate placement to be further than a certain value from two-dimensional features (corners for example) is another rule that helps in minimizing the Across Chip Line-width Variation(ACLV).

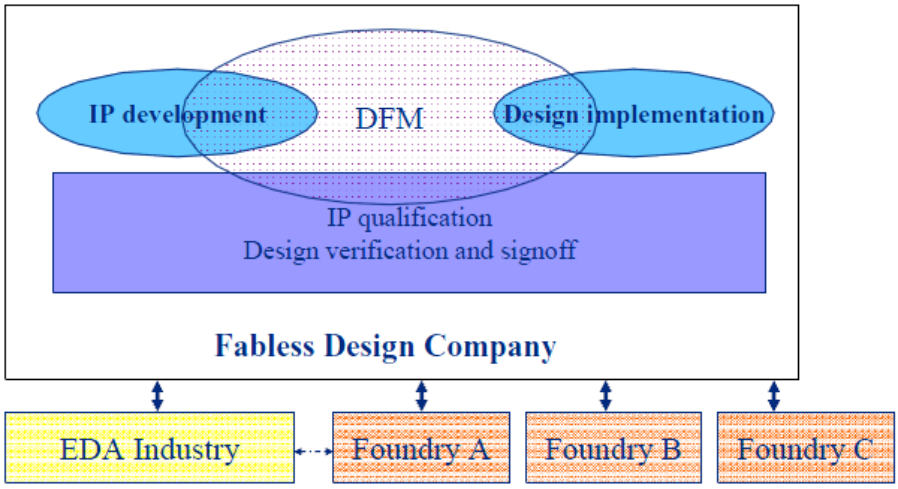


Figure 3.1: DFM Model for Fabless companies. [12]

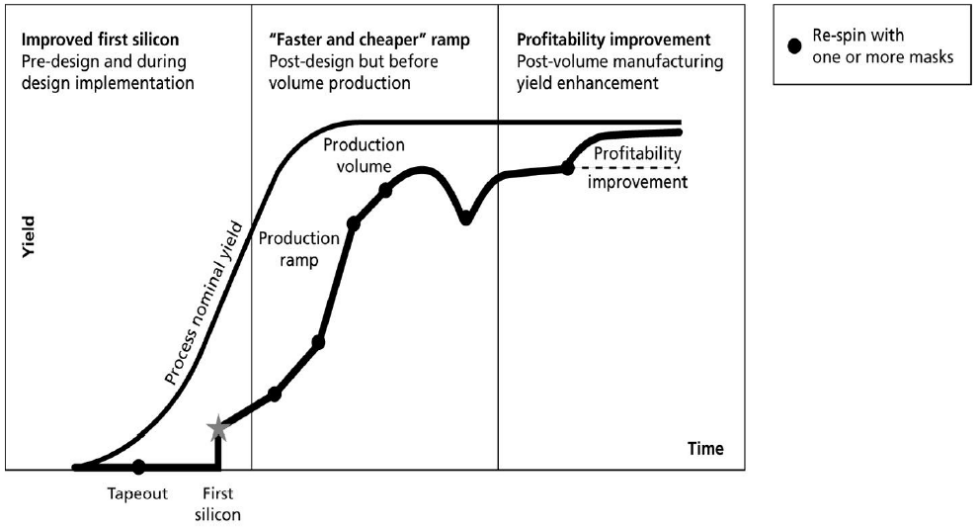


Figure 3.2: DFM methodologies play a key role in the yield improvement of the first silicon, in achieving a more effective ramp to volume production and in improving the profitability during volume production. [13]

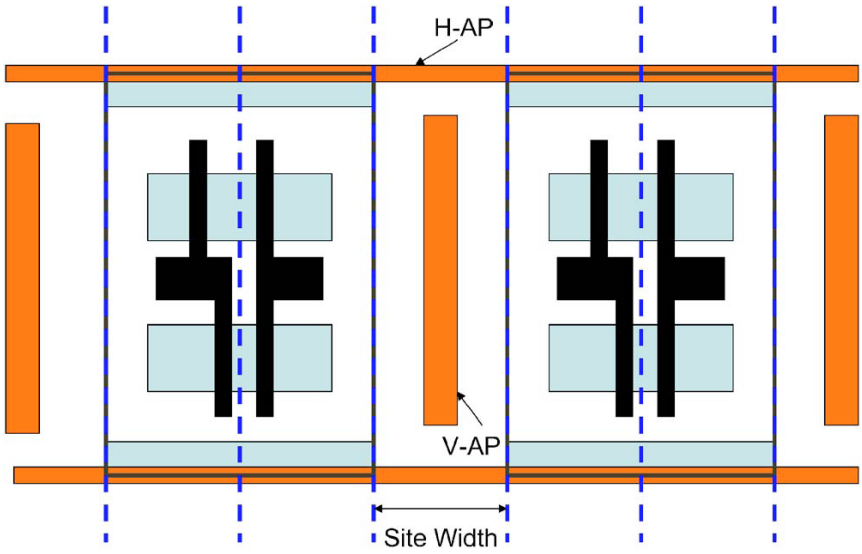
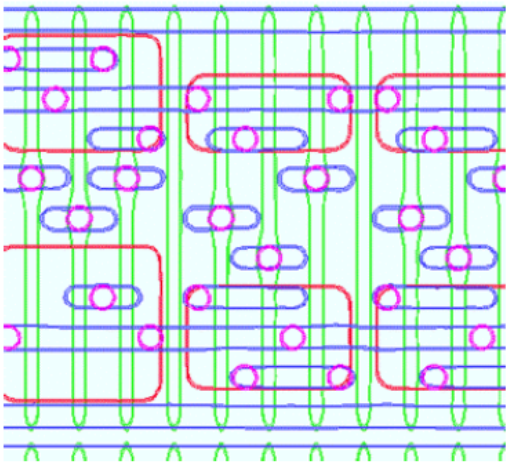


Figure 3.3: Auxillary pattern insertion for a more regular poly pitch. [14]

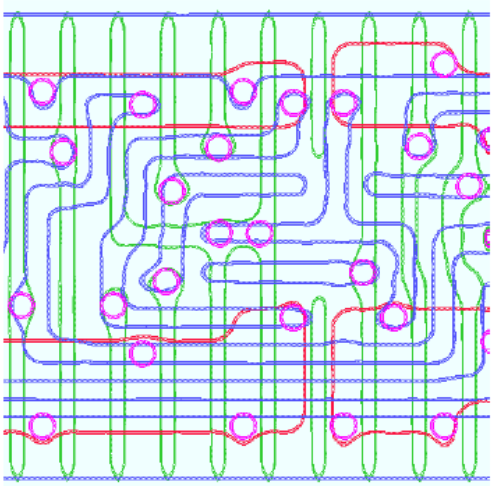
A more recent and mature form of RDR is the recommendation of new design practices that result in an extremely regular one-dimensional design fabric for all levels and adjusting the pitch for each level to its own recommended pitch as shown in figure 3.4[15]. It is proven that this design practice does not necessarily come on the expense of a design area penalty [85], [86]. The whole frame work for the design-lithography co-optimization and using regular design fabrics is a very promising technique for yield optimization.

3.2.2 Litho-Friendly Designs (LFD)

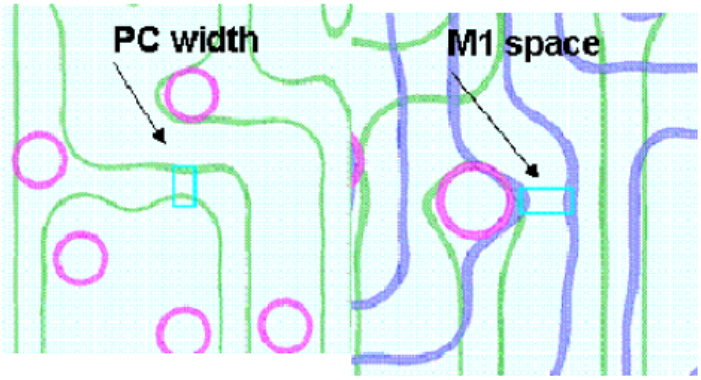
An alternative approach to improve the designs immunity to process variations is to provide the designers with the accurate means of identifying all lithography hot-spots during the early design phase and categorizing them according to their likelihood to fail during manufacturing. The earlier this information is available to the designer the faster the yield ramps up because the designer can start on fixing their designs by applying small perturbations or even changing the design strategy. Different techniques were proposed to enable LFD which varies between full chip OPC and through process simulations and verifications [35], [41], [42], [42], [87] (Model-Based LFD) to pure pattern recognition techniques [88]). The first is based on allowing the designer to get a glimpse on how their design is going to print by letting their design to pass through all the mask preparation steps and through accurate process simulations. Then all lithography hot-spots are given to the designer to



(a) Regular design (1D-like).



(b) Standard design



(c) 2D-designs are weaker against process variations

Figure 3.4: Regular design fabrics tend to exhibit better immunity against process variations. [15]

fix accompanied by a score-card representing how severe the hot-spots are.

Obviously, doing the full OPC and OPC verification on the full chip is extremely time and resources consuming and is totally not practical and as a result new methodologies were proposed to improve the efficiency of LFD flows. Among which is the speed improvement by running Model-Based LFD only on regions that are geometrically subject to lithography hot-spots (like small widths and small spaces). More recently pattern matching was proposed to run model-based LFD on even a more selective set of patterns that are close to known problematic designs as shown in figure 3.5. Using pure pattern-recognition approaches were also proposed, where all lithography simulations, verification and pattern learning is done at the fab side and the designer only gets an LFD kit that is based on the pattern trained Neural Network program [88].

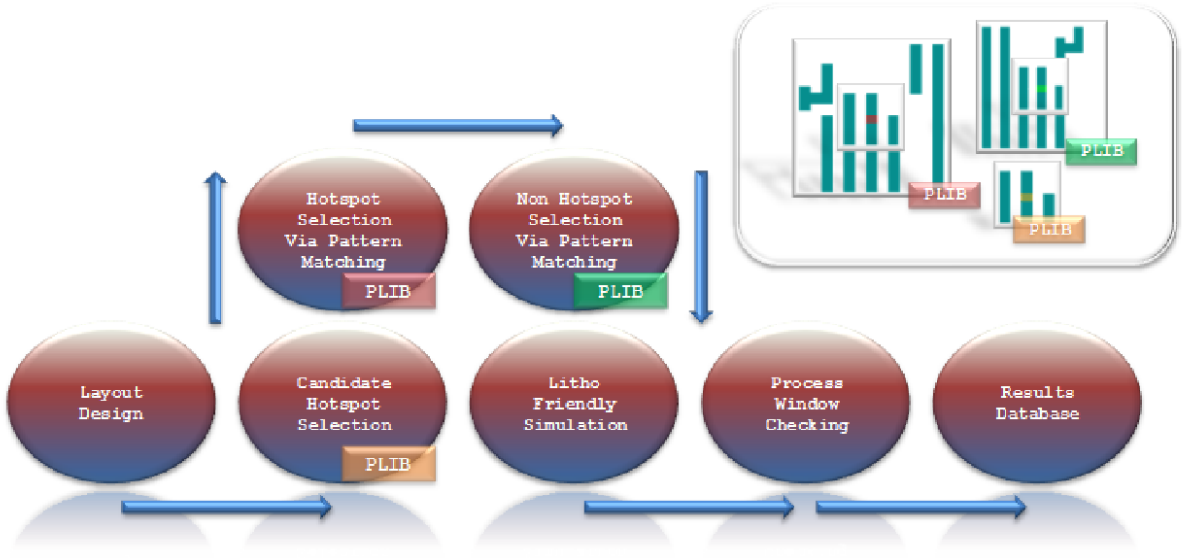


Figure 3.5: Using a Hybrid Pattern Recognition- MBLFD approach to improve the LFD flow speed. [16]

Meanwhile, an interesting concept is the to use LFD to give the designer a hint on how to fix the design or even to automatically fix the design is based on checking the design and capturing the Lithography hot-spots and applying a model-based hint (fix), then rerun the design to insure that the hot-spot is fixed [89]. This is a hot topic in the industry and is gaining attention due to its potential in improving the yield and restoring the design time to an acceptable level.

3.2.3 Parametric DFM

In contrary to functional yield issues, which are catastrophic failures (like opens and shorts) due to process variations, parametric yield extends beyond catastrophic failures to include circuit performance issues like meeting the timing and power constraints as well as product reliability specifications. Parametric DFM [90] focuses on developing methodologies that helps the designer to meet the product performance requirements. This includes process mitigation techniques and self compensation methodologies, as well as techniques that feed back the process variations impact on circuit performance and device parameters to the designers during early design stages. Accordingly, instead of doing full tape-out runs and doing the design centering based on silicon data, Parametric DFM tools and virtual fabs allow the designers to get the feedback required through simulations and save time and cost of the silicon re-spins as shown in figure 3.6 [13].

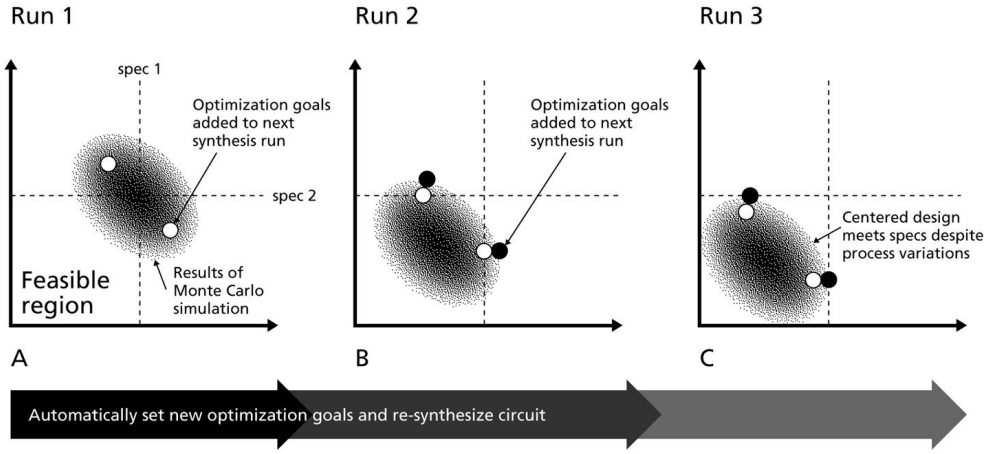


Figure 3.6: Yield-focused design tools automatically center a design by iteratively adding corners in successive runs until the design is re-centered in the optimum feasibility region. [13]

For earlier process nodes, designers have been able to successfully anticipate process variation by including safety margins in the timing requirements and layout rules. This approach helps to ensure early functioning silicon at the cost of reduced overall performance or larger die size. In deep sub-micron technologies, designing an early functioning silicon is complicated because the relatively larger process variations and the tighter design requirements would make over-designing a very lengthy and tedious process and can significantly lengthen the design cycle time. However, with more accurate DFM methodologies and proper analysis designers can reduce margins to the minimum. The proposal

of lithography-simulation-based parameter extraction was studied [91], [92], [17], [93], [94] [95], [34] to reduce the gap between the designer's expectation and the real circuit performance and to allow the designer to tweak and optimize the design during early product development stages. For device parameters extraction the simulated contour of the gate is segmented and the effective transistor properties are extracted as shown in figure 3.7.

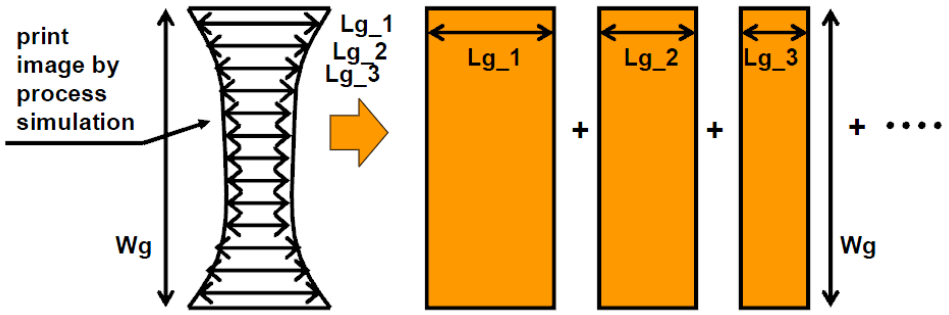


Figure 3.7: Lithography Simulation and Gate Segmentation to extract the realistic device parameters. [17]

Virtual fab parametric DFM tools also allow the designer to tweak their physical design to minimize the mismatch due to the process limitations between the design and on silicon device. Moreover, it is used to maximize the design immunity to process variations. Figure 3.8 tests the required distance between the poly contacting pad and the diffusion region. It is clear that due to the lithography process limitations, where there is a a finite curvature of the poly and the gate should start to overlap with the diffusion after the poly line has straightened so that the gate length is uniform everywhere. Of course this value is a preliminary value, that has to increase to include process variability and mask overlay issues.

The accurate parasitic extraction allows designers to account for systematic process variations and decouple them from random process variations. The interconnects parasitic extraction is important to include the circuit delay in the design considerations. Systematic deviations in interconnects can be due to many reasons like using Rules-Based ReTargeting (RBRT) in the Fab to improve the functional yield or due to the design density-dependent variation during CMP. The delay variation can result in serious discrepancies between timing simulations and silicon performance [13] and needless to say that accurately including these systematic variations in interconnect parameters significantly reduced the safety margins that the designer needs to consider in the circuit delay due to interconnects, leaving only the truly random variations in the design guard bands.

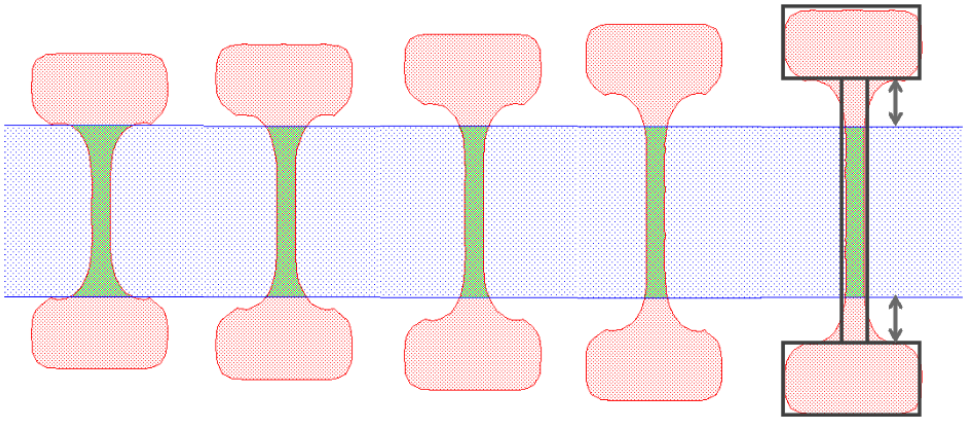


Figure 3.8: Using lithography simulations to improve the gate length control. [17]

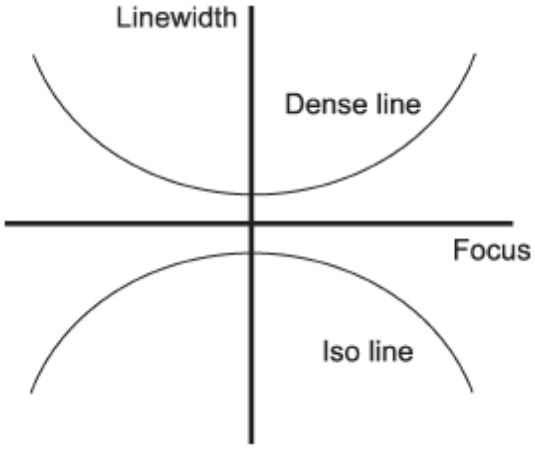


Figure 3.9: An idealized Bossung plot representing the systematic and opposing behavior of dense and iso patterns through defocus. [18]

Self compensating design techniques represent another prominent field that requires very good understanding of the design mitigation techniques as well as good process understanding. A good example of lithography self-compensating design is the work presented in [19] and [18], where it was proposed to use a phenomenon that the authors observed to automatically compensate for random through focus variations during photo-lithography. They observed that the CD variation behavior is different between dense and isolated features, where a dense feature tends to have a smiling Bossung plot in contrary to isolated CDs as shown in figure 3.9. The authors suggest to combine isolated and dense gates during the standard cell design such that any defocus-induced delay added to one gate is compensated by a speeding up in the other gate while keeping the design overall delay almost constant against focus variations as shown in figure 3.10. However, we'd like to note that the authors' assumption cannot be generalized for every level and every technology, it is strongly dependent on the RET designed for each layer.

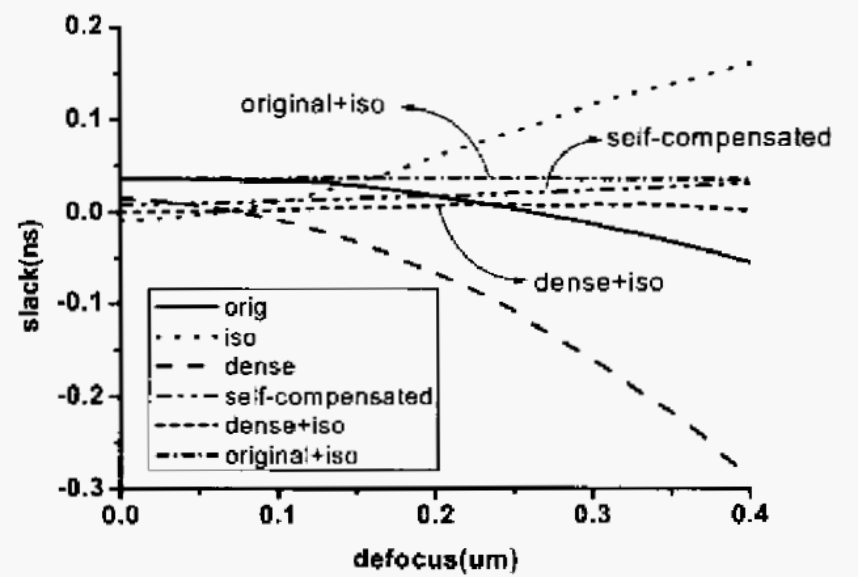


Figure 3.10: Slack(ns) vs. defocus for a testcase showing the effective compensation in self-compensating design options. [19]

3.3 Retargeting for Yield

3.3.1 Rules-Based ReTargeting (RBRT)

Retargeting has been used for years to prevent catastrophic failures during the Photolithography processing. It is also known as CATastrophic OPC (CATOPC) [21], where rules-based modifications are applied to the original design. The main goal of this retargeting step is to avoid any lithography-related weakness. Lithographic failures can manifest themselves in many forms such as the optics-related pinching and bridging or as the resist-related resist collapse issues. In other words, retargeting is applied during the mask tape-out step to transform non lithography-friendly patterns into more lithography-friendly ones.

RBRT is a form of table-driven correction where every design width-pitch combination gets the necessary bias needed to shift towards a more Lithography-safe design space as illustrated in figure 3.11. By definition, retargeting results in a systematic deviation between the on-wafer results and the original design. In relatively larger technology nodes (180nm 130nm nodes), such rules were simpler, and the relative deviation from the design was not noticeable, and could always fit within the designer's guard bands during design. However, as the design dimensions continued shrinking down into the deep-sub-wavelength nodes, this deviation cannot be overlooked anymore where its systematic deviation component needs to be fed-back to the designers instead of asking them to fit it within their guard bands.

RBRT is powerful for simple one-dimensional designs, however, things become much more complicated for two-dimensional designs, where the spatial frequency filtering effects makes it harder to recognize whether the design is Lithography safe or not. Moreover, in deep sub-micron technologies and due to the extreme shrinking of the design with respect to the wavelength, the proximity effect of the design extends even beyond the first neighboring patterns. Accordingly, a simple width-space measurement does not guarantee that the pattern can considered as a one-dimensional pattern even if the design polygon and the ones next to it are long enough. This is because if a 2D pattern is within proximity of the lines, it is going to affect the optical intensity distribution. This is demonstrated in figure 3.12, where technically for a 32nm node metal line with a length of 500nm can be considered as one-dimensional (especially when the middle space is shielded with two adjacent long lines), but once a two-dimensional design was added within its proximity, the intensity distribution starts to change noticeably. This complexity of describing the problem was one of the main reasons that Process-Window OPC (PWOPC) was necessary for deep sub-wavelength technology nodes.

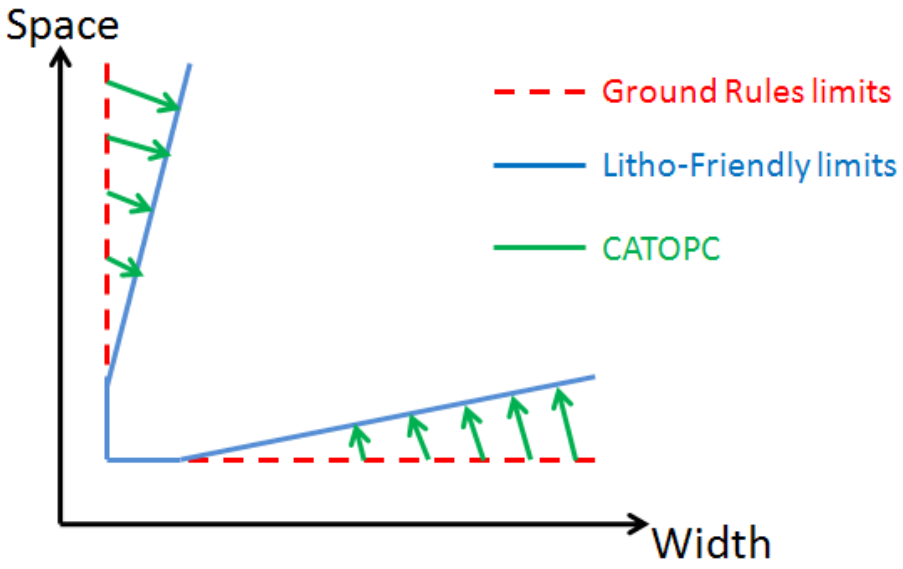
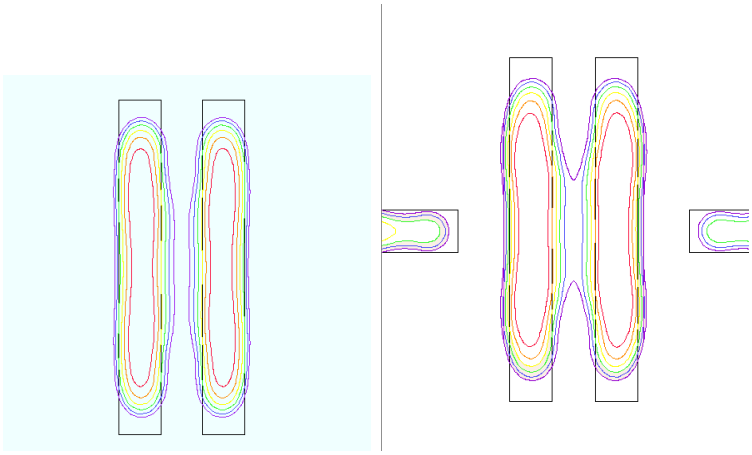


Figure 3.11: Rules-Based (Geometry-Based) Retargeting to improve the yield.



(a) Design A

(b) Design A + an added 2D feature

Figure 3.12: The Aerial Image Intensity distribution showing that in advanced nodes the proximity effects extend even beyond directly adjacent design patterns.

3.3.2 Model-Based Re-Targeting (MBRT)

In recent research, the idea of doing a Model-Based Re-targeting started to gain attention [20], [96]. It is evident that RBRT is not enough, and that designs are still suffering from PW issues (even with the application of PWOPC). The number of hot-spots increases every technology node and there is a significant increase in the sophistication of the mask correction flow to avoid hot-spots.

In [20], MBRT is introduced as a component of PWOPC, where the OPC target is modified when the OPC engine sees that the results don't fully satisfy the PW verification criteria. The flow chart is shown in figure 3.13. In concept, this technique is not very different from the PWOPC flow except that they started to identify that MBRT is the core difference between nominal OPC and PWOPC. In another approach [96], MBRT is also proposed as an OPC component, where the authors try to explore the Normalized Image Log Slope (NILS) as the metric for identifying the direction and the amount of the retargeting.

MBRT research is still pre-mature. There are several points that need to be addressed. First, the proposed techniques, so far, function only as new PWOPC algorithms rather than a stand-alone MBRT algorithm. Second, it is of great importance to identify and study the model parameters upon which the PW variations have the biggest impact and use them in the stand-alone MBRT. Third, for practical application of MBRT in the industry the speed and the computation efficiency of the MBRT algorithm need to be maximized to allow its insertion in both the mask-correction flow during manufacturing and in the parameter extraction flow. And last but not least, it is of great importance that the algorithm can evolve into a full methodology for handling multi-backend levels simultaneously to ensure that one level retargeting (Metal wire retargeting for example) is consistent with the interconnections above and below (i.e. a metal wire retargeting doesn't expose a via above or below it and resulting in a multi-level yield issue).

3.4 Optical Proximity Correction (OPC)

The technology used for compensating the pattern transfer non-idealities is called Optical Proximity Correction (OPC). The goal of the OPC is to enhance the final accuracy by making adjustments (perturbations) to the mask. This is accomplished by compensating mask geometry for known effects due to the optics or resist transfer functions. It can be explained that OPC is correcting for the the transfer functions of the optics and the photo-resist to achieve a final patterning that matches the design intent. This technology alone

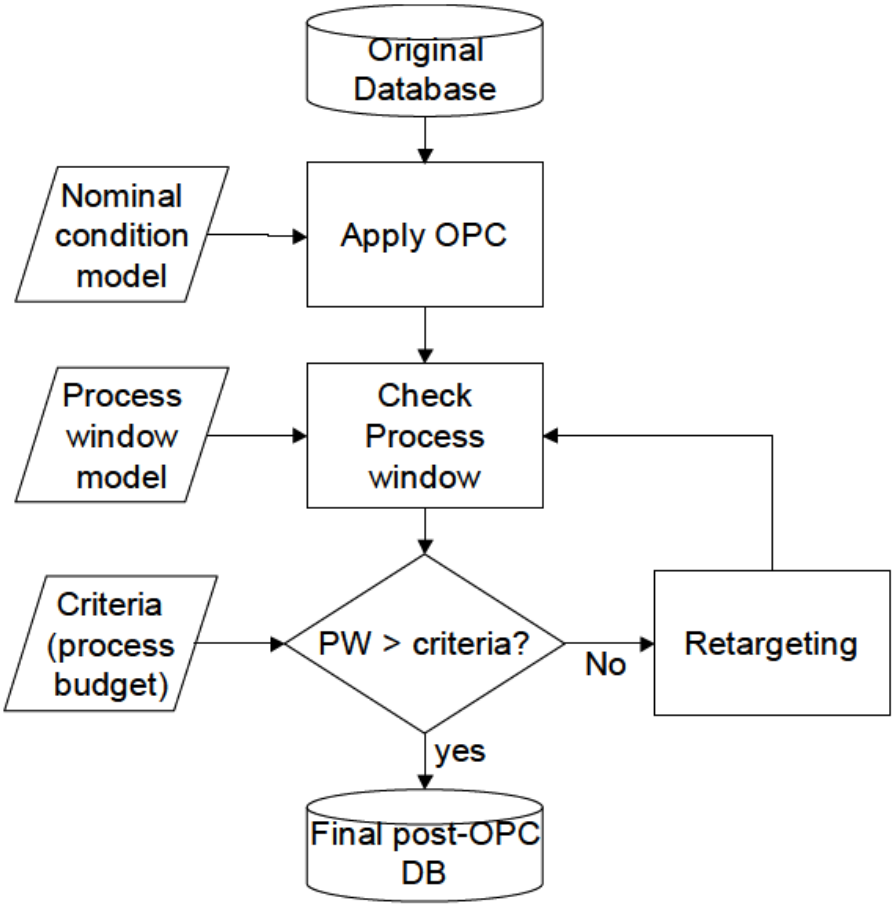


Figure 3.13: The flow chart of during-OPC MBRT. [20]

was very successful down to the 90nm node (although it involved both the Rules-Based OPC and Model-Based OPC generations).

As the technology scaling further proceeds while continuing to use the same 193nm wavelength, more advanced Resolution Enhancement Techniques (RET) were needed to keep the success going. Advancing from the 90nm node, to the 65nm node and then to the 45nm node many wonderful scientific and engineering techniques were developed to increase the resolution like the introduction of Off-Axis Illumination (OAI) [97], [98], Attenuated Phase Shift Masking (Att. PSM) [21], [54], , Sub-Resolution Assist Features (SRAF) [56], [57], [58] and Immersion Lithography [99], [100], [101], [102]. The photolithography technical difficulties of subwavelength nodes and the RET evolution is shown in figure 3.14, where the resolution challenges, the pattern information loss and the process sensitivity increase as the Rayleigh Factor (k_1) decreases.

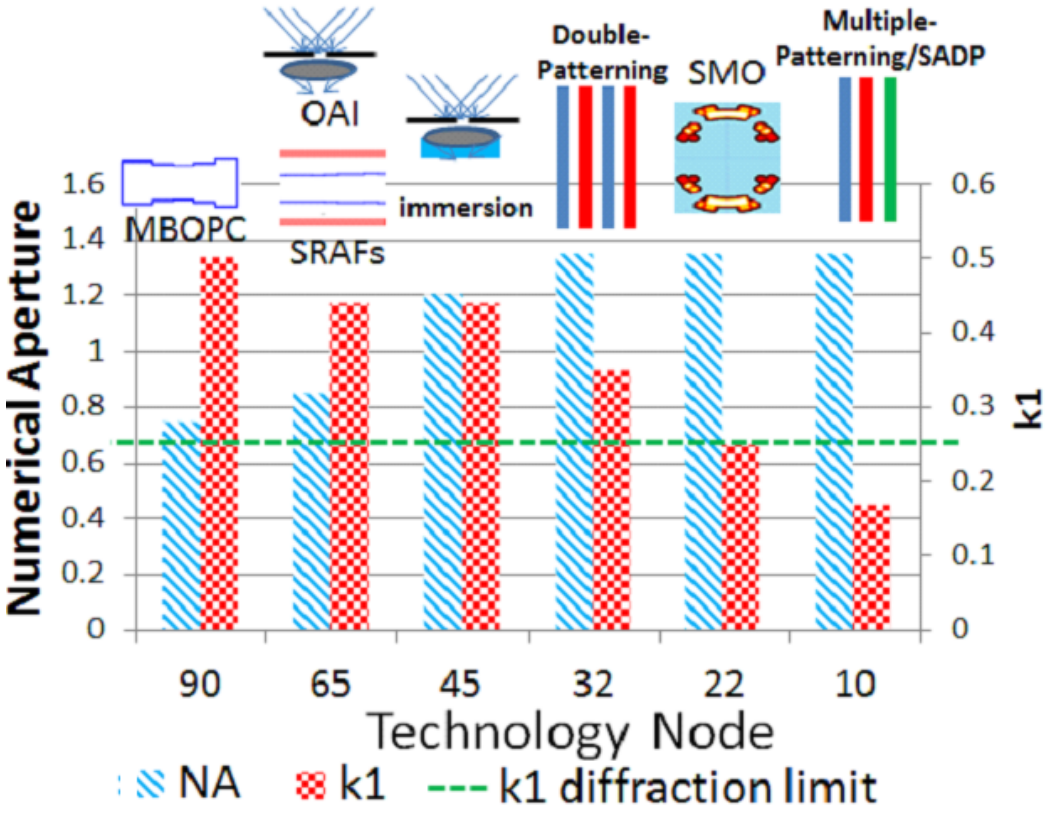


Figure 3.14: An illustration of the RET evolution in sub-wavelength nodes, demonstrating both k_1, NA and the mainstream RET development.

Starting from the 32nm and 22nm nodes, even more challenges were discovered and had to be overcome. First, double Patterning was inevitable for some layers [82], [103], [104], [105], [106], [107] . Second, OPC alone (i.e. OPC with pattern transfer fidelity as the only objective) is not sufficient anymore to achieve an acceptable yield. Process-Window OPC (PWOPC) was introduced to fix any through-PW hot-spots during the OPC operation. PWOPC didn't come for free, but rather on the expense of a noticeable addition to the computational power required during the mask tape-out. The third challenge is that even with PWOPC, lithography hot-spot-free chips are not totally insured. A new set of tools were introduced to the design community to help them achieve Lithography-Friendly Designs (LFD) to allow them to start working on any lithography hot-spots during the design phase.

Pushing of the lithography tools to their physical limits (where the k_1 factor is close to the 0.25 theoretical limit value), the lithography exposure system starts to lose important pattern information. This loss of information (which is directly related to the loss of higher order diffraction orders of the pattern) results in a systematic deformation in the pattern that is transferred to the wafer. There are many forms of this loss of pattern information could manifest itself into. For example, it appears in the form of corner-rounding, line-end shortening, through-pitch Critical Dimension (CD) variations and last but not least comes the non-linearity in the Mask Error Enhancement Factor (MEEF) [108], [109], where the same change in the mask dimensions transfers differently on wafer depending on the neighboring patterns (i.e. depends on the short-range proximity). These four types of non-idealities are illustrated in figure 3.15 [21].

Accordingly, OPC is a technique to iteratively perturb the mask feature to counter the effects of the non-idealities in the optical system. This means that each shape in the design is manipulated depending on its proximity so that its final printing on the wafer is as close as possible to the designer's intention. Figure 3.16 shows how an isolated line is printing for both the original shape and its OPCed version. It is clear that the OPC is capable of allowing the final wafer pattern to be more similar to the original design intent.

The IC industry started using Model-Based OPC (MBOPC) since the 130nm node, where earlier the much simpler Rules-Based OPC has been used to correct for simple lithography non-idealities. All MBOPC algorithms share the need to have a process model that is capable of accurately predicting how the features on the mask are going to print on wafer. This model should take all of the optical system specifications and limitations, mask manufacturing non-idealities and the photo-resist specifications and composition. In this section we are reviewing the OPC Models and the different OPC algorithms.

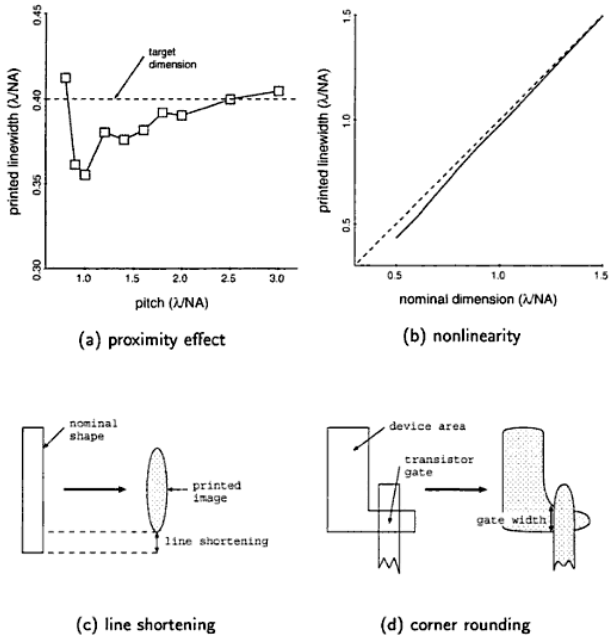


Figure 3.15: Non-idealities in optical lithography. [21]

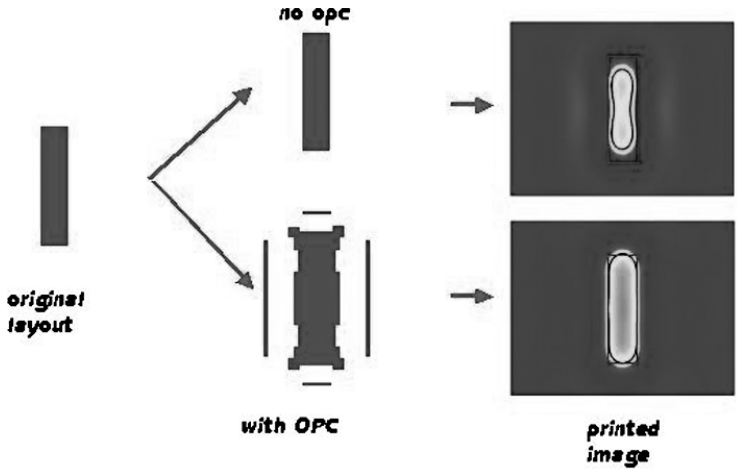


Figure 3.16: An OPC example showing how an isolated pattern is printing with and without OPC. [1]

3.4.1 OPC Process Models

OPC process models are required to model all the primary and secondary optical effects. This includes modeling the imaging system [110], [111], the optical system non-idealities like the lens aberrations and the appodization loss [112], the low-pass spatial-filtering effects, the optical system response to different wave polarizations which is of great importance for hyper NA systems in immersion lithography [113], [114].

It is also essential that the OPC process model considers the photo-resist response to the light in terms of how the standing wave is formed inside the resist stack and also in terms of how the light absorption process happens inside the resist. Moreover, it is essential to include the post-exposure baking and the resist development process. In addition, the model also has to be capable of including any mask manufacturing non-idealities (like the corner-rounding on the mask and the chrome Side-Wall Angle (SWA)).

For practical reasons, the OPC models need to be computationally efficient. Fourier Optics [115] [116] is one of the strongest candidates used in the optical simulation of the lithography imaging system, where the internal algorithms are using Fast Fourier Transform (FFT) to model the spatial frequency decomposition of the mask and to simulate the propagation of all the resulting plane waves. Also, because these propagating waves are neither coherent neither incoherent, the Hopkins equations for partial coherence[111], [117] are used to include the partial coherence into the simulations.

The accuracy limit of the model is predefined for each technology node depending on the criticality of the level of interest (for example, the accuracy of a poly level is much tighter compared to the accuracy required for a higher level metal). The model building is a fitting process, where many wafer measurements are taken for a wide range of design space and fitted into the physical or the empirical models to reach a final set of parameters values that accurately predicts the pattern transfer process. Among the challenges that face the model building is the measurements accuracy, where this is usually handled by gathering several measurements per measurement location and averaging them and excluding any measurement statistical outliers.

3.4.2 OPC Algorithms and their objectives

3.4.2.1 Iterative Nominal Model Based OPC

At early stages when the feature size was larger or comparable to the wavelength the OPC was a very simple operation. It simply consisted of a set of rules-based corrective

operations (Rules-Based OPC). This includes correcting the through-pitch CDs using a simple two-dimensional width-space table. Moreover, the corner rounding and line-end shortening were fixed by adding hammerheads to the line-ends, which also was simply explained in terms of a few simple rules. Figure 3.17, shows a few examples of RBOPC in earlier technology nodes.

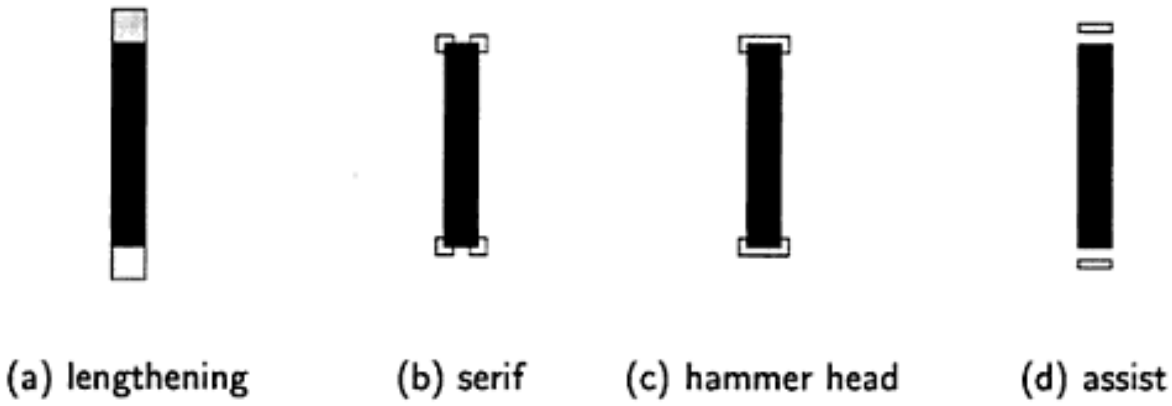
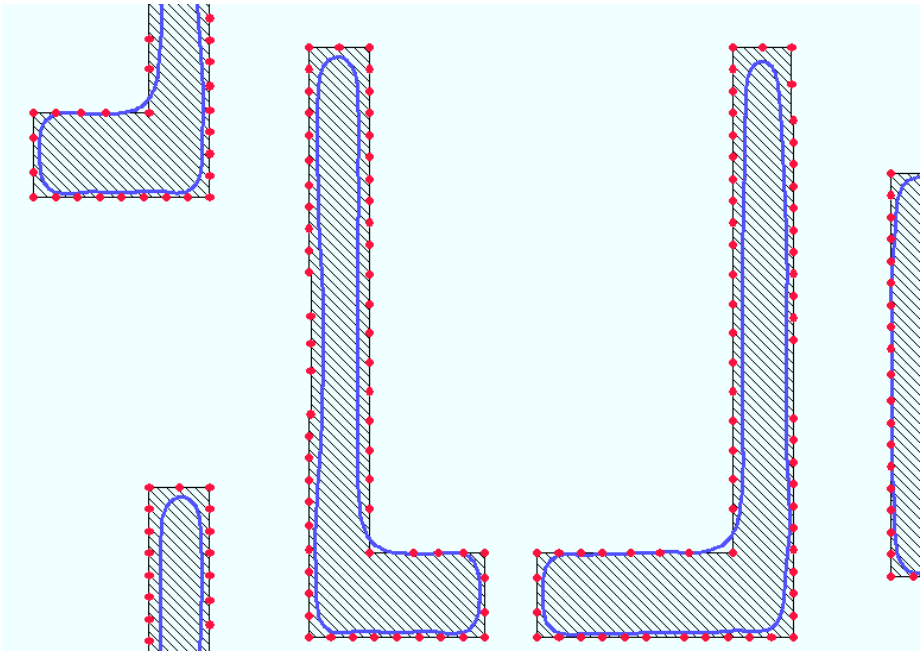


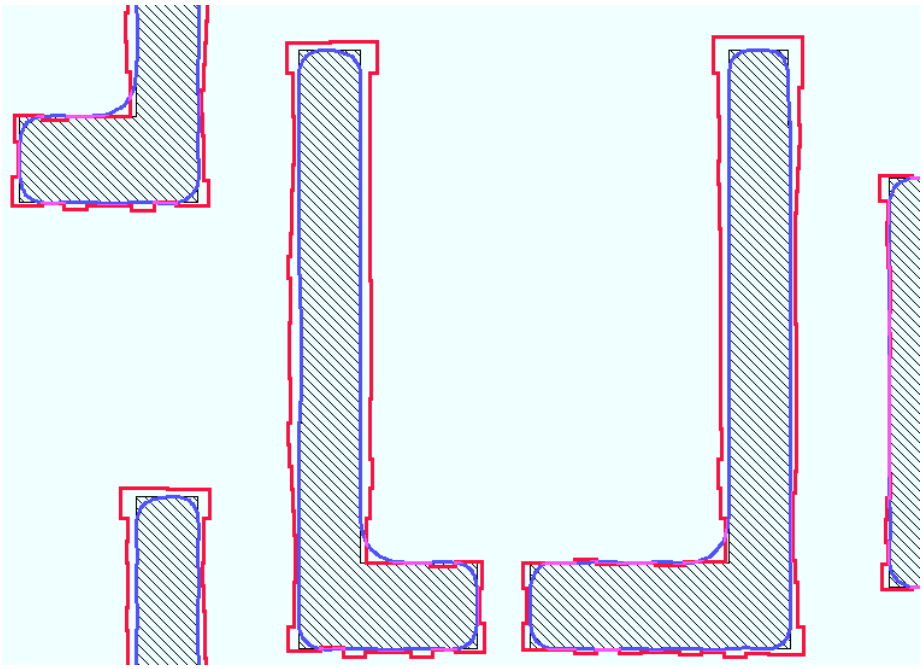
Figure 3.17: An example of Rules-Based correction for Line-Ends. [21]

When the technology started to reach the sub-wavelength nodes, Rules-Based OPC was not sufficient anymore and more sophisticated techniques were needed. This is when the Model-Based OPC concept was introduced as a feasible solution. Nominal condition MBOPC [118] is the first model-based OPC technique. It is developed to achieve design fidelity for sub-wavelength technology nodes. This technique has been used on the critical levels since the 130nm node and later on widely used in the 65nm node for both front-end and back-end levels.

In this technique each design edge is divided into small fragments, and each fragment is allowed to move independently in the inward and outward directions as shown in figures 3.18(a) and 3.18(b) . Then the correction engine iteratively tries to forward solve this inverse problem. During every OPC iteration the engine simulates the optical and resist responses to the mask in hand. Then the difference between the designed pattern and the on-wafer printed pattern (called the Edge Placement Error or EPE) is computed for each fragment and then moves each fragment according to how much deviation it is suffering. This flow is summarized in details in the flow chart in figure 3.19.



(a) Pre-OPC fragmentation and lithography simulation



(b) Post-OPC mask shape and Lithography-simulation

Figure 3.18: An example of Nominal Model-Based OPC.

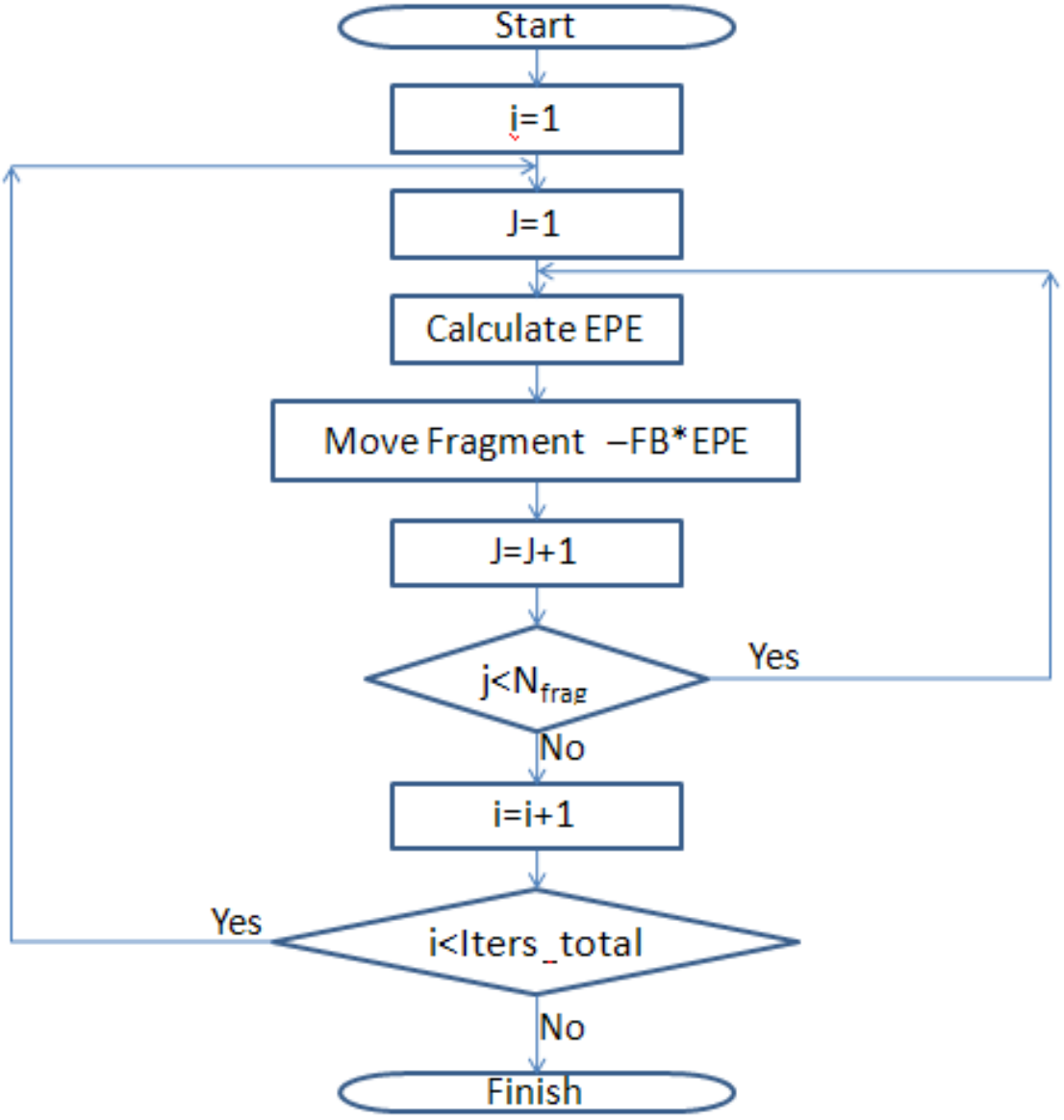


Figure 3.19: Nominal MBOPC Flow.

There are several factors that determine the success of the OPC correction recipe built inside the fab. The most important factor is the final EPE and whether the recipe successfully maximizes the design-to-transferred-pattern fidelity. In the nominal MBOPC the only objective function used is the nominal process conditions EPE, which means that the variance in the photo-lithography process and its effect on the wafer-printed results is not considered. Another success (or preferable) factor is the speed of the correction algorithm (which is translated into time and computational resources). Also the robustness of the OPC correction recipe for all the patterns in the allowable design space is very desirable.

3.4.2.2 Process-Window OPC (PWOPC)

As the industry started to move forward towards the 45nm node, the photo-lithography resolution improved with the successful introduction of immersion lithography [99]. The resolution limit is better (smaller) in immersion lithography due to the increased numerical aperture of the optical system, where we have shown in the previous chapter that the resolution limit is inversely proportional with the system's numerical aperture. This was a widely celebrated success within the manufacturers' community. This was not an easy job, many technical obstacles had to be faced, like the technicalities of introducing the high refractive index fluid without introducing any defects and without causing any resist lift-off during processing [102]. An additional challenge to the computational lithography world is that the OPC models need to robustly consider the polarization. This is mainly because the oblique incidence angles on the mask are larger than before and as a result both the TE and the TM polarizations start to behave differently, which has to be considered during OPC in mask preparation.

However, even with such improvement in the resolution, we still faced a new challenge that required careful attention. Through-process-window lithography hot-spots generation rate increased in several front-end levels. These hot-spots are resulting from a few serious reasons. First, with the increase of the Numerical aperture the optical system's Depth of Focus (DoF) decreases [21] which means that the statistical variance of the printed patterns increases for the same defocusing control in the exposure system. Second, with the design shrinking to almost a quarter of a wavelength, the proximity effect of neighboring designs is extended even to farther designs. The printability and through-PW performance (i.e. through the possible range of variation in Focus and Dose) is not only affected by the adjacent space, but to how the nearby shapes look like too. This means that it is possible that the same design could pass or fail within a lithography process condition depending only on how its neighboring pattern looks like. This is a new added level of complexity to both the designers and manufacturers. Figures 3.20(a) and 3.20(b) show an example of

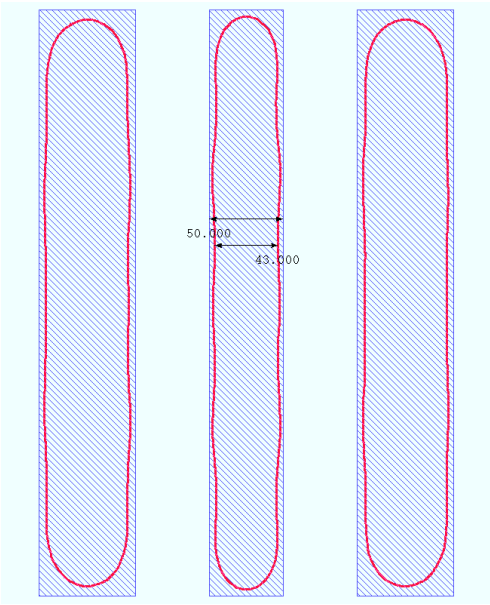
how the through-PW innermost contour of the two metal lines in different designs could behave differently through PW just because their neighbor shapes changed even if the spacing was the same.

Several PWOPC techniques were proposed to consider PW conditions during OPC and automatically fix them during the mask tape-out flow [119], [120], [121], [122]. In PWOPC, the mask is simulated against different through PW variations, which includes dose, focus and mask error variations. And if at any iteration the OPC simulations show a probable lithography hot-spot then a retargeting is internally applied to prevent the hot-spot. This retargeting takes the form of widening the design widths or spaces that are subject to through-PW pinching or bridging respectively. Figure 3.21 shows the same design (B) after applying PWOPC, it is obvious that PWOPC automatically fixed the soft pinching hot-spot by widening the nominal printing.

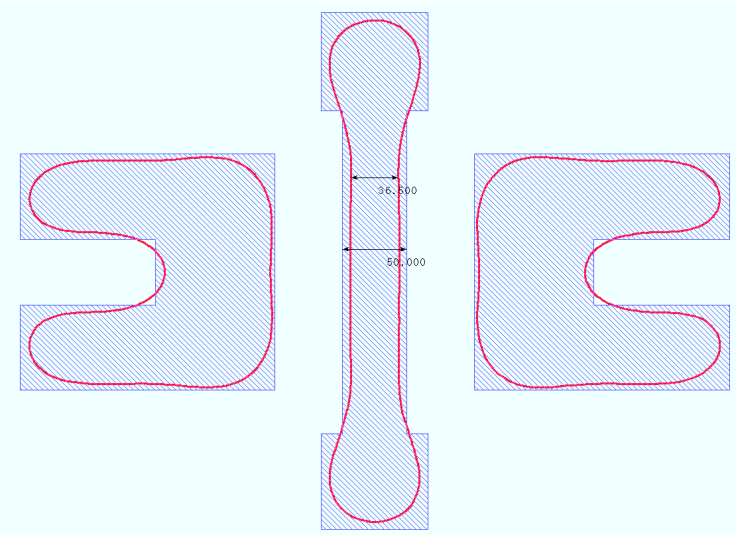
PWOPC is a building block in all technology nodes mask-tape out flow since the 45 nm node. Many of the lithography hot-spots can be fixed. However, there are a few challenges that are facing PWOPC. First, there is an increased demand for the computational power required by PWOPC; this is mainly due to the increased number of optical simulations. Second, the quality of the final solution depends on how good the correction recipes are prepared, where all sort of recipe parameters like the fragmentation sizes, the feedback and the number of iterations directly affect the quality of the wafer results. Finally, the PWOPC retargeting results in a systematic deviation in the device parameters (gate length and width, wires resistance and capacitances, etc). For this reason, PWOPC is not usually applied to device-forming levels and instead nominal OPC is used simultaneously with a set of Restrictive Design Rules (RDR) which is a part of the DFM flow explained later in this chapter. However, PWOPC is still being used on Metal levels, and accordingly the deviation in the resistance and capacitances can still be observed on interconnect levels.

3.4.2.3 Inverse Lithography OPC

The Inverse Lithography Technology (ILT) has been developing since several years [22], [123], [23], [124]. In this powerful computational technique, the OPC problem is backward solved instead of forward solving the OPC problem in an iterative fashion. This means that the correction algorithm is based on finding the inverse of the collective transfer function of the lithographic system (i.e. the optical system and the resist response) and using it to find the mask that would result in the desired wafer printing. This technique has an edge over the standard (iterative forward solving) OPC techniques, which is the ability to compute the best mask as a whole including the best SRAF sizes and positions. Figure 3.22 shows an example of the generated mask for a contact level [22]. Generating the



(a) Design A



(b) Design B

Figure 3.20: The through-PW innermost resist contours comparison for to designs with the same width-space.

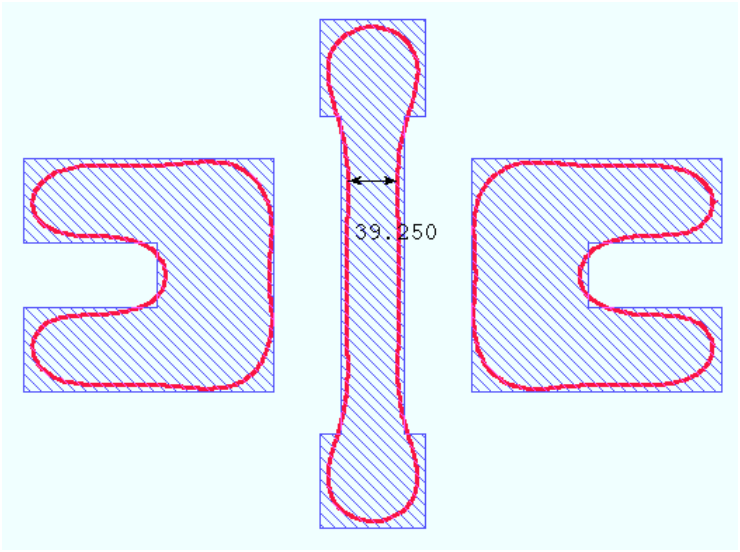


Figure 3.21: Design B from figure 3.20(b) after applying PWOPC.

Model-Based SRAFs (MBSRAF) results in a noticeable PW improvement as shown in figures 3.23(a) and 3.23(b) showing the average of the printing variability and its standard variations respectively.

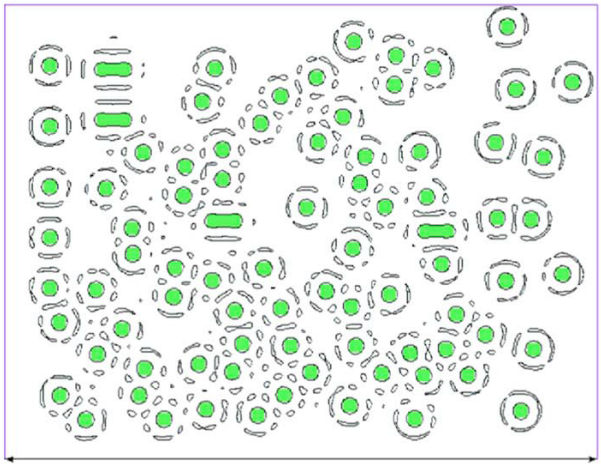
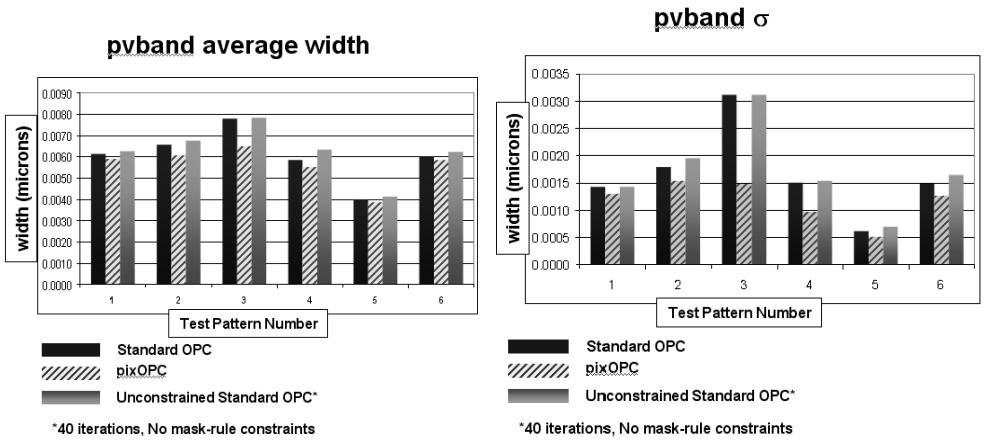


Figure 3.22: An example of ILT of a contact level. [22]

The formulation of ILT is not a straightforward task. This is mainly due to the in-



(a) Average PVBand width (b) PVBand width standard deviation

Figure 3.23: Through-Process-Window performance improvement in ILT. [23]

formation loss in the optical system to high spatial frequencies on the mask. In other words, the transferred optical image and resist latent image is not ideal (i.e. binary like on the intensity axis) as shown in figure 3.24. Accordingly, it is hard to define the best achievable aerial or latent image spatial-distribution that is going to be used as the input to the inverse problem solving. This increases the computation power required to achieve the best solution, where the inverse problem is iteratively solved trying to maximize the image fidelity, where image fidelity combines the terms like the Normalized Image Log Slope (NILS) and the image contrast and any terms that measures how close the image is to the binary-like spatial intensity distribution.

ILT has a great potential and it is capable of improving the lithography-related yield issues (especially due to the introduction of MBSRAFs). However, there are a few technical challenges that are preventing it from being the front runner in the industry. An important factor is that the created mask (best mask) is in a free-form shape, which doesn't follow the manufacturability constraints defined by the mask house. In order to create a manufacturable mask, the edges need to be either horizontal, vertical or at 45 degrees. Accordingly, a post ILT processing step is required to transform the free-form mask into the final mask and ensure that not much of the image quality is not lost. Figure 3.25 shows how a final mask looks like compared to the free-form output from ILT [24]. This post processing step is complicated because if it is not done right, the final image quality could degrade and results even in a degraded yield. The second challenge that faces ILT is the much higher computational power it requires compared to the conventional OPC

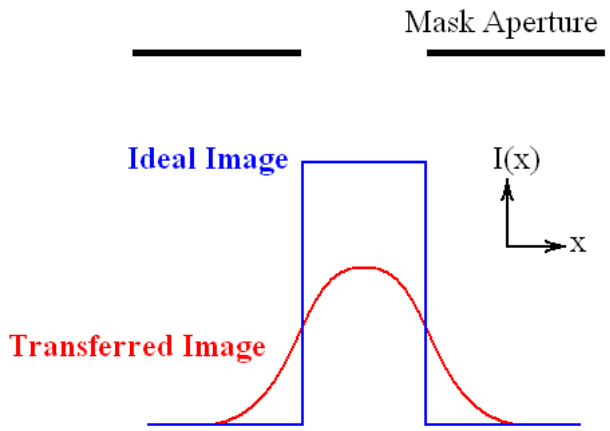


Figure 3.24: An ideal (lossless with an infinite NA) vs. the practical transmitted image.

techniques, where it could sometimes reach an order of magnitude.

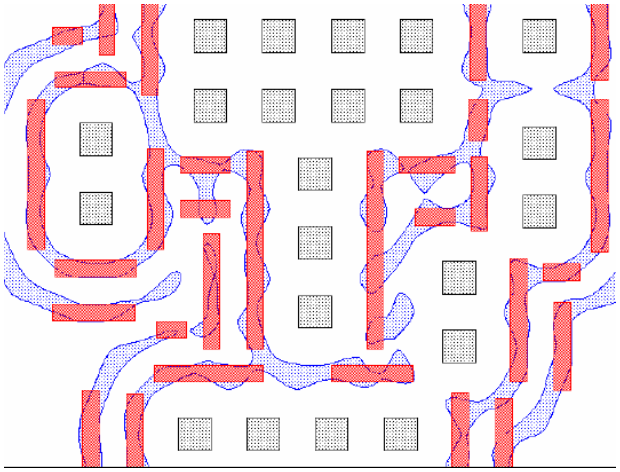
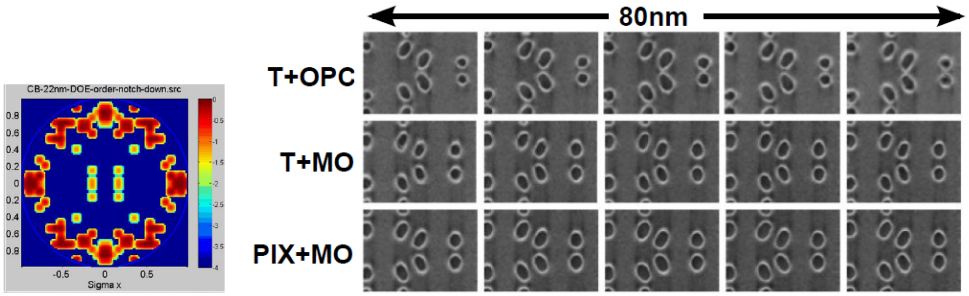


Figure 3.25: Free-Form SRAFs in ILT (blue) vs. manufacturable SRAFs (red) [24].

3.4.2.4 Source Mask Optimization (SMO)

Source Mask Optimization [53], [125], [126], [25], [127] is a more recent technique that simultaneously optimizes the mask and the light source illumination distribution. The Source optimization is the process of engineering the wavefront of the light wave as it falls

on the mask to maximize the resolution and the through-PW performance of a selected group of important design patterns like the SRAMs, important design pitches, etc. In SMO, the source is optimized in conjunction with the mask optimization. Figure 3.26 [25] shows an example of how the optimized source and mask look like for an SRAM contact level and how the wafer printed image looks like. It is observed that for very dense like structures, the combined Mask-SRAF shapes can be non-intuitive and can never be achieved using conventional OPC algorithms as shown in figure 3.27 [26]. SMO is also a computationally expensive technique, and recent work[128], [129] was investigating the possibility the using efficient distributed computing techniques to allow the utilization of more resources efficiently to save the computation time, which is in some cases more important when the TAT is the important metric.



(a) Optimized source (b) SEM for Traditional OPC, Mask Optimization+Traditional RET, and Full SMO

Figure 3.26: Optimized Source and on-wafer improved printing quality. [25]

3.4.2.5 Miscellaneous OPC techniques

Many other miscellaneous OPC techniques were proposed to address some issues or propose new metrics for the OPC problem. In one approach, Dual-Metric OPC [27] is proposing to use multiple objectives for the OPC. The design fragments would move to solve their own EPE, while simultaneously the SRAFs are fragmented and allowed to move around with the objective of improving the PW performance on the main mask features. This approach is capable of optimizing the SRAFs placement to improve the process-window, while having the merits of the iterative forward-correction OPC family. Figures 3.28(a) and 3.28(b) shows how the initially inserted SRAFs are mapped to the main features and corrected to improve the PW performance respectively. This technique is very useful in cases where the rules-based SRAFs can be inserted in close range to the optimum SRAFs so that the

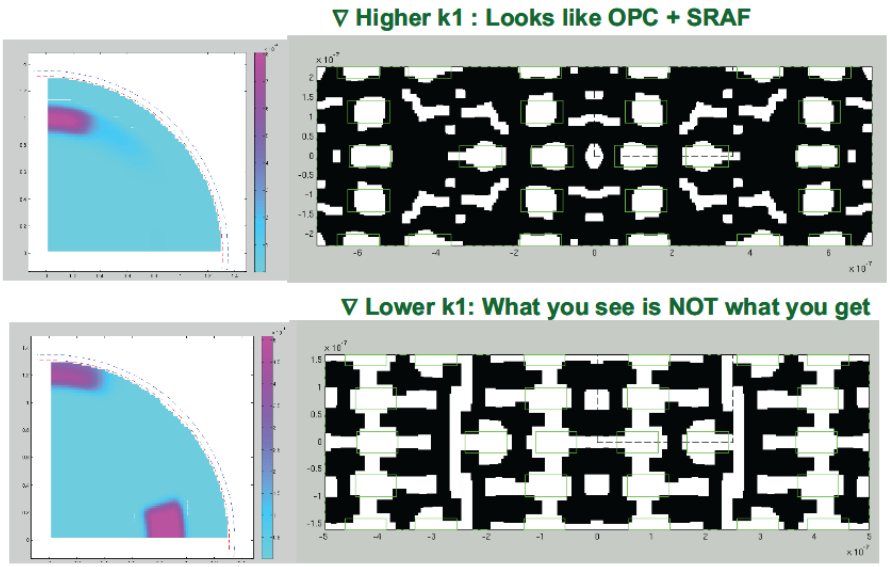


Figure 3.27: SMO Mask does not necessarily look close to the original target (green) at low k_1 values (i.e. at extremely aggressive RET). [26]

engine can manipulate the initial SRAFs to reach their optimum sizes and positions. This technique is useful down to the 20nm node and it is very fast in comparison to the ILT solution.

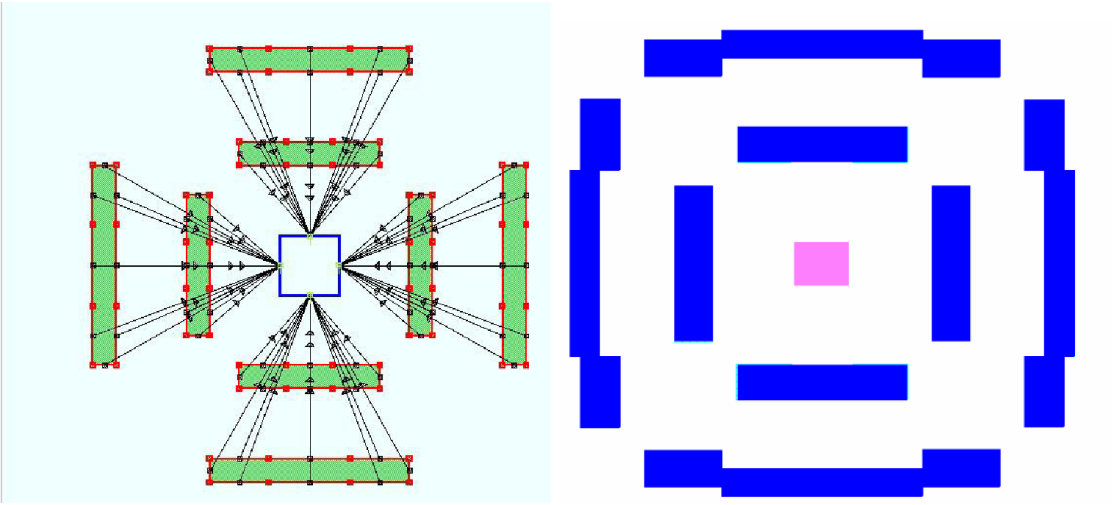
Another creative OPC technique was suggested to totally change the OPC metrics from geometric to electric [28], [130], [131]. In this work, the authors highlight that the ultimate goal of the pattern transfer to wafer is to meet the electrical requirement by the designers rather than the geometrical objectives. As shown in figure3.29, they segment the simulated gate region and calculate the ON current for this non-uniform gate and accordingly they extrapolate the gate-length difference from the designer's intent

The position varying current value is a function of the gate length at a specific point and can be considered as

$$I(x) = f[L(x)] \tag{3.1}$$

And due to segmentation, the total gate current can be calculated as the sum of the currents in all segments

$$I_{shape} = \sum I(x)W(x) \tag{3.2}$$



(a) Initial mapping of seed SRAFs (b) Post-Correction SRAFs

Figure 3.28: The Basic Concept of DMOPC [27].

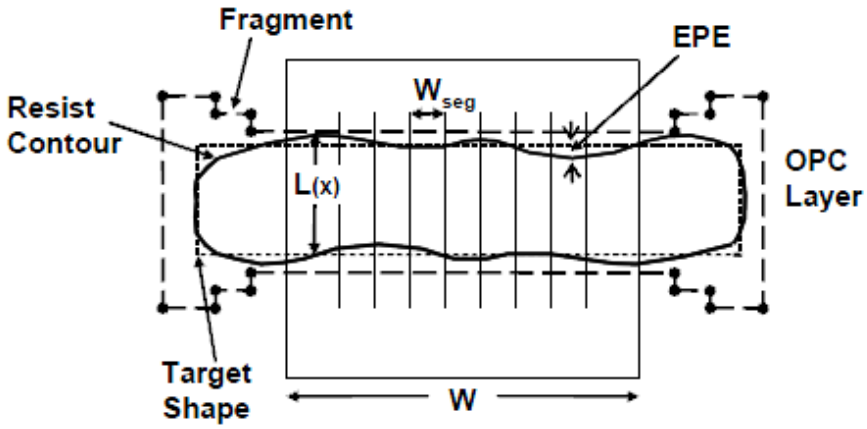


Figure 3.29: Nominal Contour segmentation into multiple transistor equivalent. [28]

Where, I(x) is the extrapolated circuit current from SPICE simulation of a gate with width Wseg and w(x) is a weighting parameter for each transistor slice to account for the gate edge effect. Finally, we can compute the effective gate length to be

$$L_{eff} = f^{-1}(I_{shape}/W) \tag{3.3}$$

The computed Leff is then used to decide whether the timing and power of the transistor is within the allowed specifications and use this as the objective function to be minimized in this OPC algorithm. The flow chart of this OPC methodology is compared to standard OPC in figure 3.30.

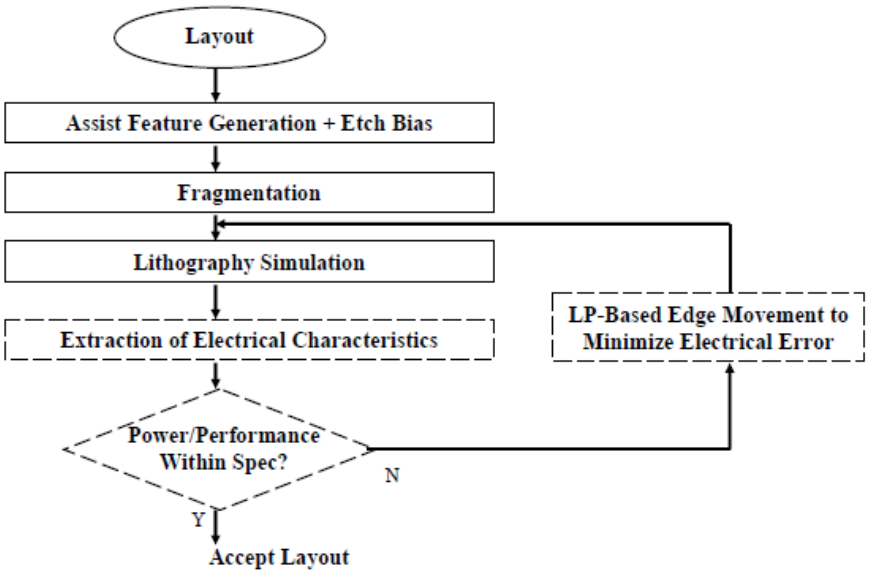


Figure 3.30: The flow chart of Electricall-Driven OPC flow. [28]

Other techniques were proposed to take the produced mask shot count into considerations [132], [133], [134], [135], where the shot count of a mask directly affects the writing time and the cost of the mask. These techniques are pre-OPC fragmentation alignment, post-OPC fragments smoothing, and post-OPC jogs alignment. All these techniques apply the understanding of how the mask fracturing and writing tools work to minimize the number of shots needed to manufacture the mask. It was reported that a shot count reduction of approximately 20% can be achieved without affecting the wafer results quality.

3.5 Summary

In this chapter we reviewed the state of the art technologies that are applied to achieve design fidelity, which includes reviewing several OPC techniques and comparing their capabilities and their objective functions. We also reviewed the design retargeting-for-yield and DFM techniques, which can be classified into two main categories that target Lithography-related yield improvement. The first is the in-fab retargeting-for-yield, where the design is perturbed to prevent lithography hot-spots and is usually classified into rules-based and model-based approaches. The second technique is the recommended design rules and LFD-based retargeting done during the design phase, where the designer is provided by the means (lithography simulation tools or just yield improving rules and information) to identify the lithography hot-spots and fix them. Parametric DFM and virtual fab methodology are also of very important to achieve the parametric yield of the product while reducing the development time and cost.

Chapter 4

Aerial Image Retargeting (AIR): A proof of feasibility

4.1 Introduction

In the previous chapter we studied the state of the art techniques for maximizing the design fidelity by correcting for all optical and process proximity effects during the mask creation to minimize the differences between the design intent and what is going to be finally printed on silicon. We also studied the state of the art concepts and tools for maximizing both the functional and the parametric yield. Functional yield is improved by modeling the design printing behavior and identifying the weak design spots and fixing them through the proper retargeting. This retargeting is applied anywhere during the mask tape-out flow inside the fabs. Retargeting can be applied before the OPC step in a rules-based (geometrical-based) approach, where the weak spots geometries are highlighted and redesigned to improve the design immunity to process variations. Retargeting is also implicitly applied during PWOPC, where the lithography hot-spots are identified during OPC using the process models and fixed instantly during tape-out. Other Model-Based Re-Targeting (MBRT) techniques were also proposed [20], [96], but as they are applied during OPC, we prefer to categorize them still as PWOPC with new correction metrics as they are completely inseparable from OPC.

Rules-Based ReTargeting (RBRT) is facing serious challenges starting from the 32nm node due to the extended optical proximity effect, which makes the design vulnerable to the nearby design shapes and the design style. It is becoming very hard to have a simple set of rules that can fix all possible hot-spots and the in-fab retargeting flow is growing to

be much more complex with every technology node. It is very obvious that model-based techniques are needed. Meanwhile, Model-Based techniques that were proposed so far are very time and resources consuming and always result in at least doubling the time needed for each mask layer tape-out. Moreover, because they are computationally exhaustive techniques it is totally impractical to integrate them in the parameter extraction flow to feedback the systematic deviation from the design intent to the designers.

Due to their importance, in our research we address the automatic model-based retargeting techniques. First, we identify the potential weaknesses in the current parametric and functional DFM methodologies and that a more accurate feedback is required from the manufacturing to the designer. Also, the need for having a computationally efficient model-based retargeting (fast automatic design-tuning for yield), where we identify the process parameters required for applying the retargeting. In this chapter, we present our early research work that proved the feasibility of fast MBRT that can improve the patterning quality.

4.2 Aerial-Image Retargeting (AIR)

In the previous section we concluded that an efficient (accurate and fast) model-Based OPC target correction is needed to achieve better lithography-Friendly targets. This technique needs to be accurate and computationally efficient so that it can be effectively used in the OPC production environment. Also, if this approach can catch most of the retargeting needed to have a PW-friendly OPC target, then this would enable two important benefits. First, on the fabs side, reducing the computation needed for better yield. This is mainly because the post-AIR OPC target is more lithography-friendly and accordingly simpler OPC recipes are needed. Second, it is possible to model the PWOPC systematic deviation between the design intent and the final on-wafer characteristics.

Aerial Image Retargeting (AIR) identifies the weak PW designs (hot-spots) and classifies them into width sensitive and space-sensitive. The basic concept is based on evaluating the optical signature of the design (which also includes the proximity effects). The optical signature (also known as the aerial image signature, because the resist effects are not included in it) is computed for the OPC target (i.e. its a pre-OPC simulation). Then the design is divided into small fragments (similar to what happens in regular OPC), but their movement (retargeting value) is extracted from a look-up table. This look-up table is developed by the DFM and OPC Engineers after training the recipe on a wide and yield-challenging design space. This movement is a change in the OPC target and should

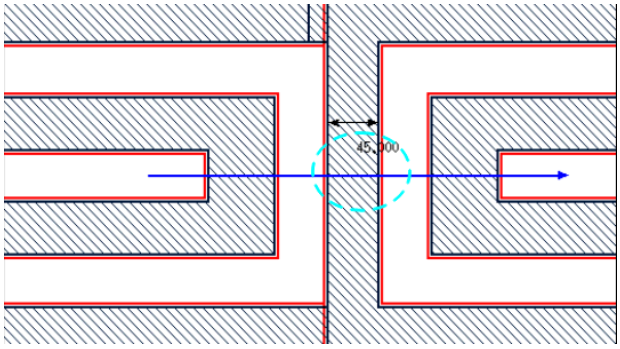
not be confused with the OPC where the target fragments move to minimize their Edge Placement Error (EPE) between the printing image and the design intent.

Similar to RBRT, where the resizing value is determined by the width-space combination, the AIR resizing value is coded in a table based on the Aerial Image signature using parameters like I_{max}, I_{min} (Maximum and Minimum optical intensity along the simulation site respectively), Aerial Image Slope, as well as the optical field curvature at the simulation site. This adds several degrees of freedom and is more capable of identifying Lithography hot-spots based on the Optics which is the root cause of Lithography limitations. Accordingly, even if the design is 2D it can be fixed independent of all different proximity effects and any surrounding designs.

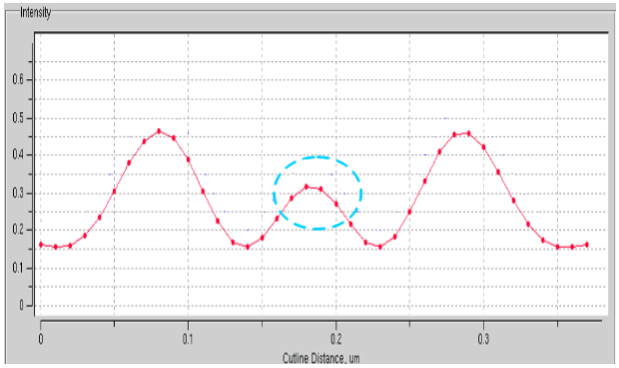
Figures 4.1 and 4.2 show two different hot-spots (with the same width-space combination), but in the same time they exhibit radically different post-OPC yield issues. It is obvious that both designs exhibit two very different aerial image signatures, which we claim to be the origin of the lithography yield issues. We propose using the aerial image signature in capturing and classifying lithography hot-spots because they can generically identify the lithography-related yield issues better than the geometrical description techniques. Moreover, using aerial image signatures in capturing hot-spots is much faster than relying on the full model-based lithography hot-spots capturing techniques (i.e. doing full OPC+lithography simulation).

To test the assumption above, we ran OPC on a chip that is designed to contain all the lithography-challenging designs. Then we mapped the PW worst widths (pinching) and spaces (bridging) to the aerial image maps. Figure 4.3 shows the AI map for the width-sensitive designs, where each contour represents hot-spots having the same measurement through the process-window, the smaller the number gets (on the legend) the worse it is from a lithography (yield) perspective. So for example in the figure the point (0.3,0.14) represents the worst design signature and results into severe pinching, then as the designs signatures starts to get away from this point, the failure severity starts to get better. Accordingly, we can conclude that for such map the designs with aerial image signature closest to (0.3,0.14) need a large positive bias (retargeting) to counter the effect of the bad initial design. One important observation from figure 4.3 is that the contours follow a monotonic behavior, which indicates that the AI signature can solely identify the PW weak areas. Moreover, using the aerial image contours curvature and the aerial image slope would add extra degrees of freedom in describing the PW more accurately and allows better separation between the both width and space weak designs signatures.

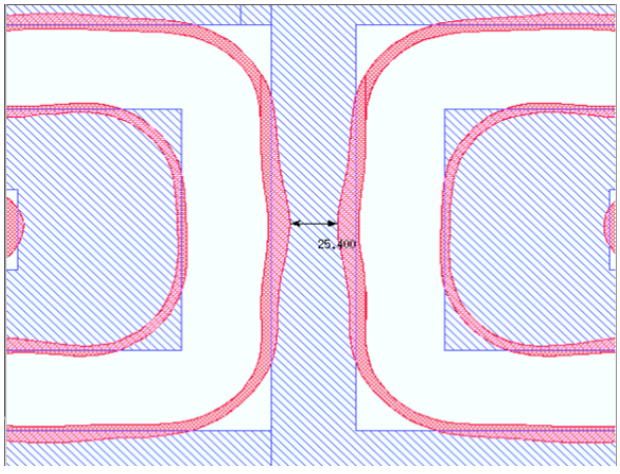
It is of particular interest to see if the aerial image contrast could alone serve as a single variable for identifying weak PW designs.



(a) Hot-Spot position and the critical direction

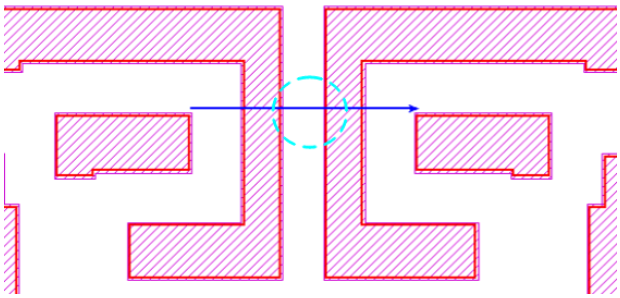


(b) Aerial Image distribution along the critical direction

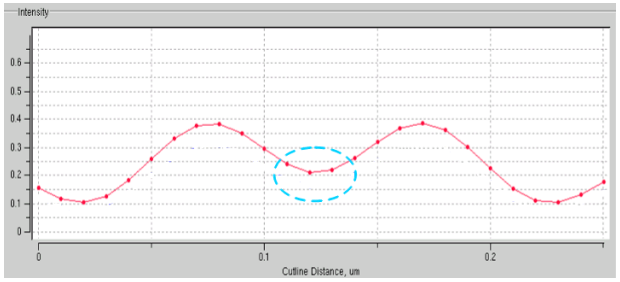


(c) Post-OPC through PW verification of the design showing bad soft pinching

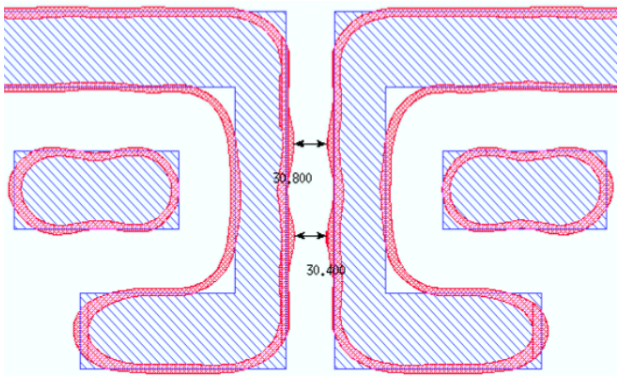
Figure 4.1: Analyzing hot-spot (1).



(a) Hot-Spot position and the critical direction



(b) Aerial Image distribution along the critical direction



(c) Post-OPC through PW verification of the design showing bad soft pinching

Figure 4.2: Analyzing hot-spot (2).

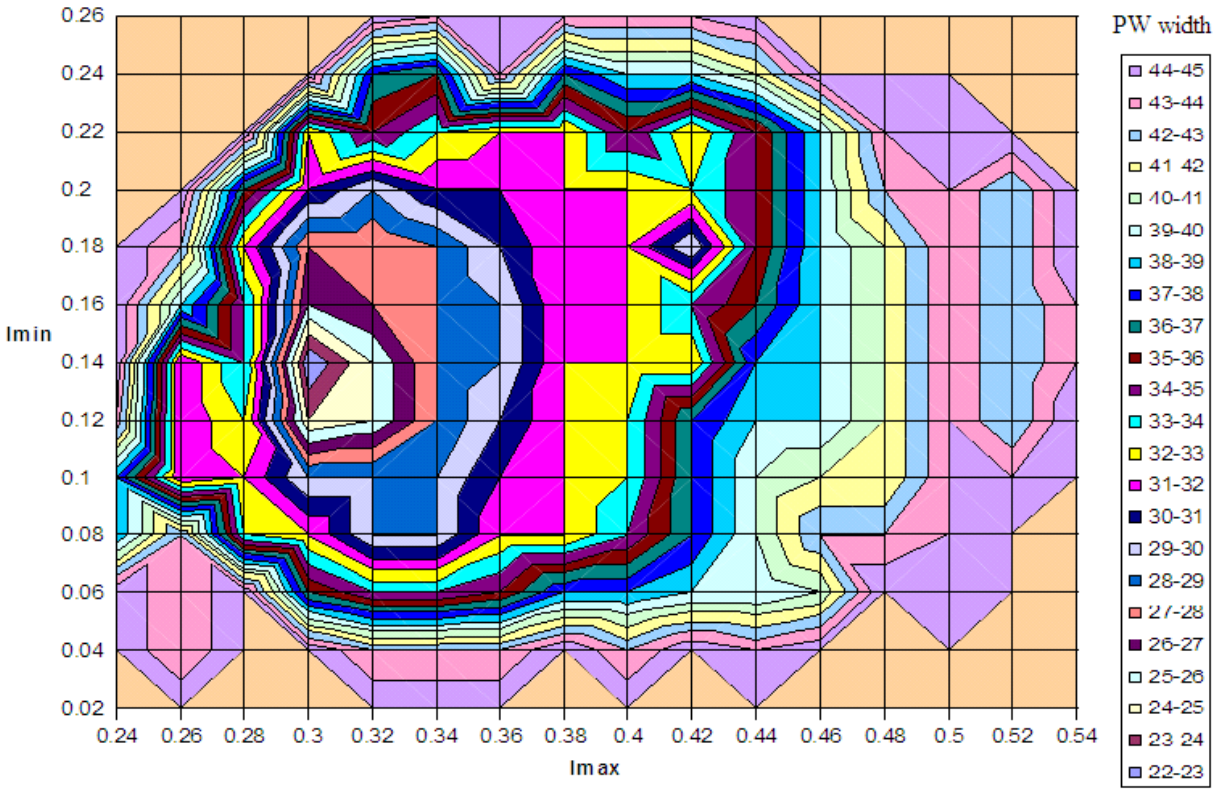
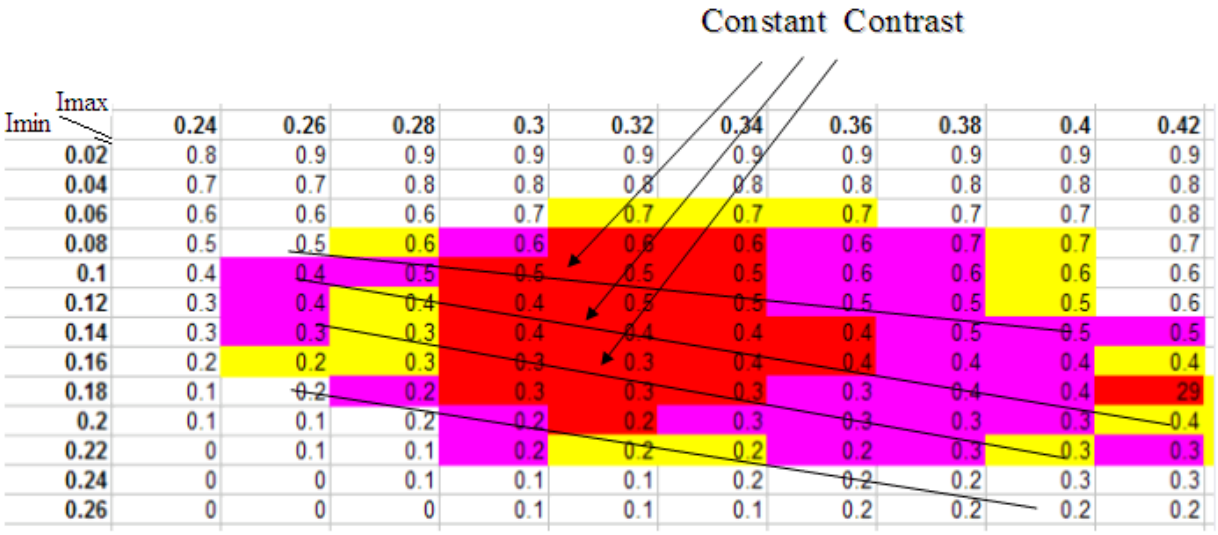


Figure 4.3: The PW width errors mapped on the I_{max}-I_{min}.

$$C = \frac{(I_{max} - I_{min})}{(I_{max} + I_{min})} \tag{4.1}$$

In figure 4.4, we are plotting the contours of constant contrast overlaid on the I_{max}-I_{min} map shown earlier, while the color coding represents the severity of the hot-spot. It is very evident that the contrast cannot describe the weak PW areas by itself.



Imin \ Imax	0.24	0.26	0.28	0.3	0.32	0.34	0.36	0.38	0.4	0.42
0.02	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
0.04	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
0.06	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.8
0.08	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7
0.1	0.4	0.4	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6
0.12	0.3	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.6
0.14	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5
0.16	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.4
0.18	0.1	0.2	0.2	0.3	0.3	0.3	0.3	0.4	0.4	29
0.2	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.4
0.22	0	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.3	0.3
0.24	0	0	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3
0.26	0	0	0	0.1	0.1	0.1	0.2	0.2	0.2	0.2

Figure 4.4: Mapping the contours of constant contrast on the I_{max}-I_{min} map (while the severity level is color coded). It is shown that for the same contrast value passes through different regions of the map that goes between very critical and non critical.

AIR is thus capable of doing a Lithography-aware retargeting of the design that is very similar to that done during PWOPC (as PWOPC sacrifices the design fidelity (EPE) if the design is suffering from poor PW performance). However, AIR is capable of doing this retargeting 1) As a pre-OPC step, i.e. not linked to OPC, 2) Very computationally efficient (consumes less resources than that needed by a single nominal OPC iteration), 3) fits better in the parasitic-extraction flow and the modeling of the systematic PWOPC deviation from the original design as presented in figure 7.28.

Figures 4.5(b) and 4.5(b) show the standard mask tape-out flow and the proposed flow after the insertion of AIR as a MBRT module for lithography-related yield improvement. Note that with AIR, it is not necessary to use full PWOPC anymore and nominal-OPC is sufficient because the OPC target is more lithography friendly after MBRT. The AIR

correction flow is summarized in the flow chart in figure 4.6 , where the design is simulated using the process optical model for all fragments, then looping over all fragments and comparing the aerial image signature to the AIR lookup table that defines the bias (re-targeting amount) needed for each AI signature, after than all fragments are allowed to move according to the pre-set values in the lookup table, followed by a cleanup step that ensures that no lithography target is being pushed into a different non-litho friendly regime.

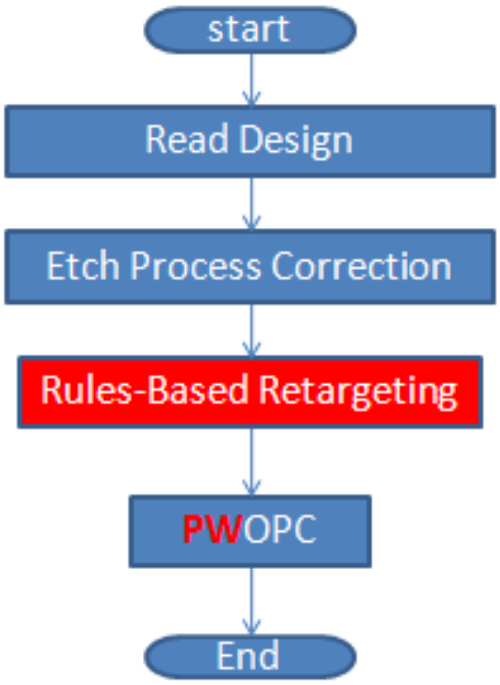
Figure 4.7 shows the methodology used for creating the AIR table, First a library of lithography challenging designs is used to evaluate the AI-PW performance correlation, then the generated table is used by the OPC engineer to set the proper fragments biasing based on their AI signature and then running the AIR mask tape-out flow and do PW OPC verification iteratively until the best AIR table testing is achieved. The last part of the flow is not easy to automate because of the importance of the human judgment, which makes building the AIR table a tedious job during the development but once the final table is achieved the impact of AIR is outstanding in terms of quality and runtime.

4.3 AIR Testing Results

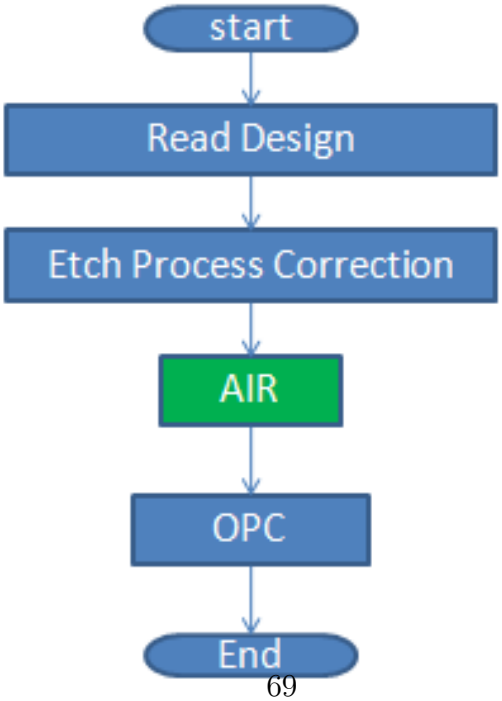
In our testing of AIR, we bench-marked a 28nm metal flow with AIR against full PWOPC. A generic AIR recipe (table) was generated based on the known problematic designs and interpolation was used to guess the proper biases for missing data. The AIR recipe was also optimized for run-time optimization, where the simulations were only performed on edges that have a potential of bridging or pinching (i.e. relatively small width or space), as well as optimizing the simulation parameters for speed.

4.3.0.6 Basic Testing Flow

In this section we explain the flow we used to do the Aerial Image analysis and how we correlate the aerial image map to the corresponding through-process-window performance. The flow is explained in figure 4.8, where the design goes through the default OPC flow but with two additional analysis steps. The first step is applied to the OPC target and an Aerial image signature for each fragment is stored as shown illustratively in figure 4.9. Then this signature is mapped to the worst condition printing of the same fragment. Finally, the worst printing value for each aerial image signature is selected and added to the final AIR map as shown in figure 4.10.



(a) The standard mask tape-out flow



(b) The mask tape-out flow after adding AIR

Figure 4.5: Mask tapeout Flow after AIR insertion.

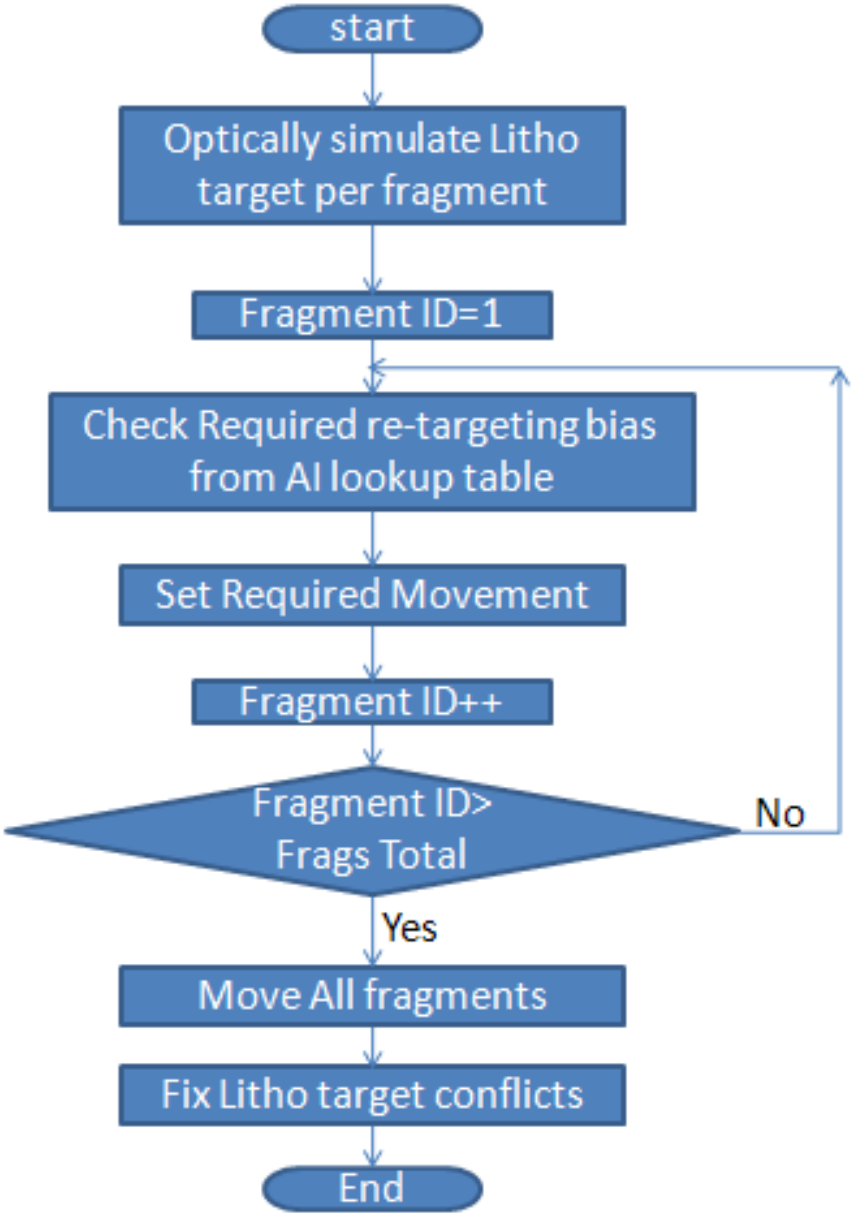


Figure 4.6: AIR correction flow.

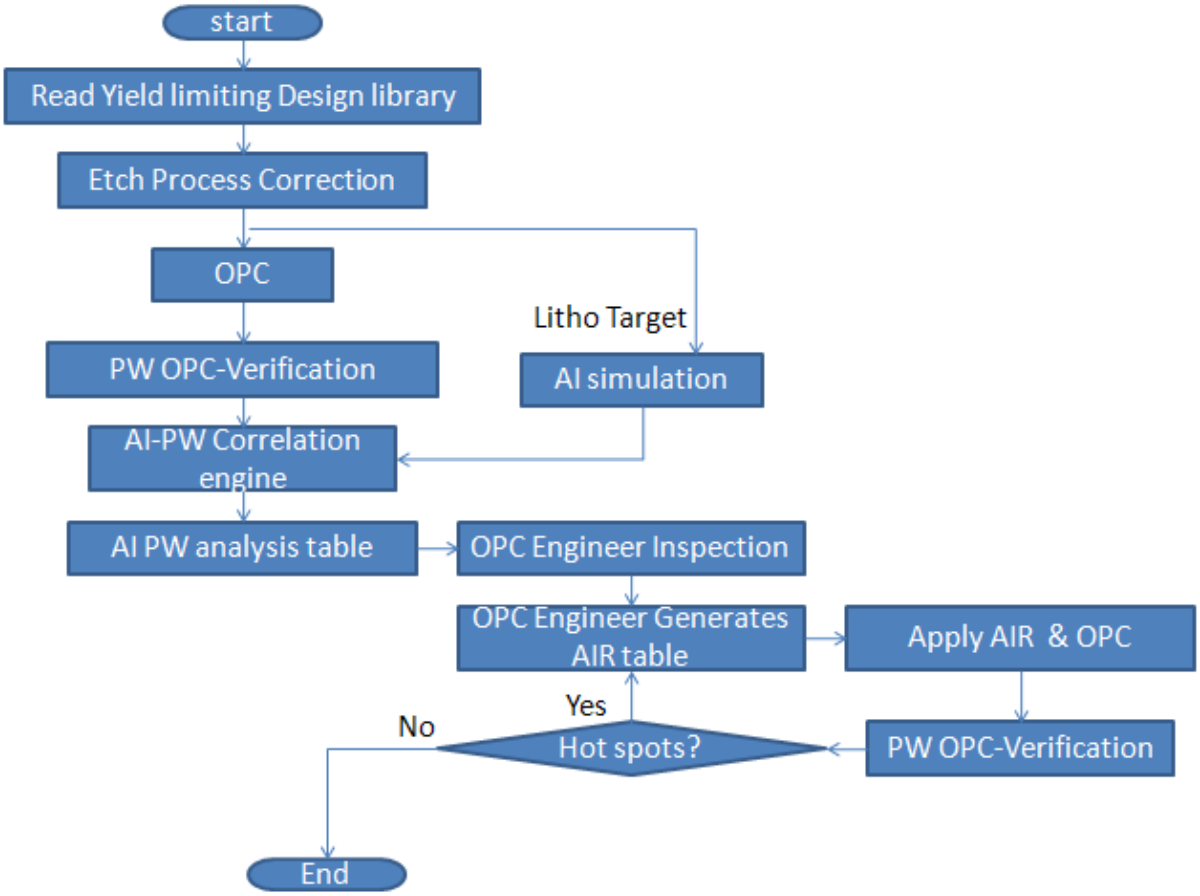


Figure 4.7: AIR lookup table preparation flow.

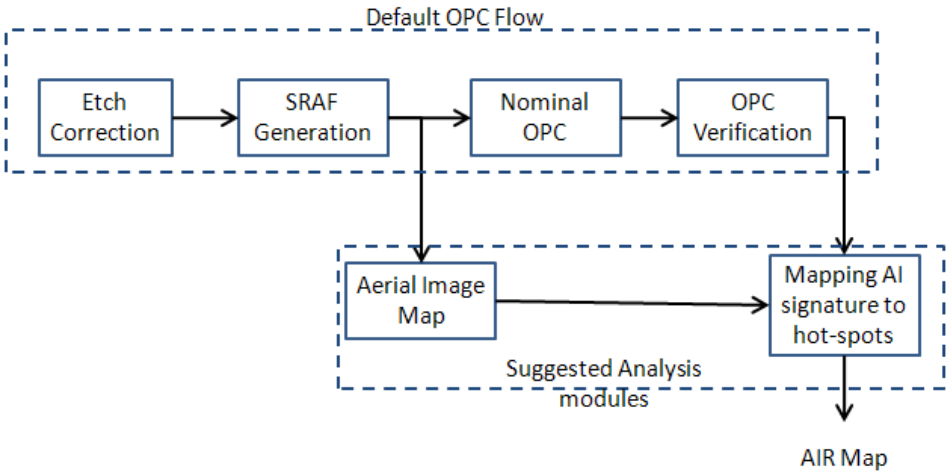


Figure 4.8: The Flow used in generating the Aerial Image Maps.

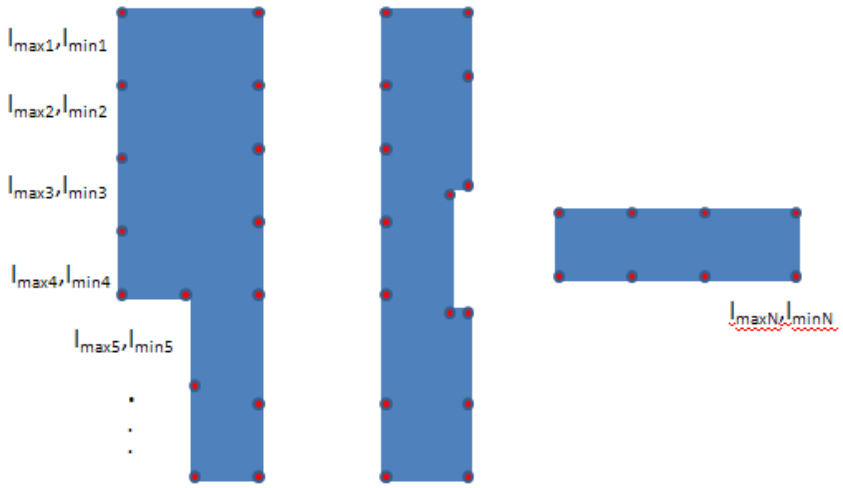


Figure 4.9: Calculating the I_{max}, I_{min}, etc... for each fragment.

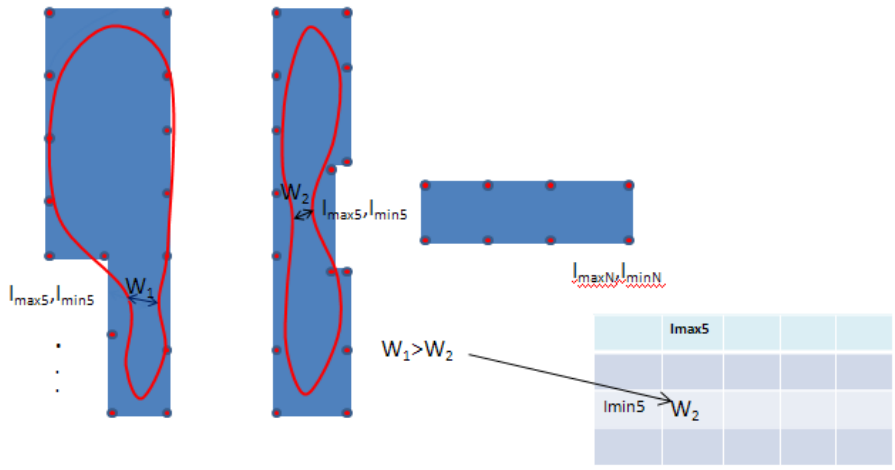


Figure 4.10: Mapping the worst PW measurement to the AI map.

4.3.0.7 Computation-Power Optimization

Aerial Image Retargeting is studied with the purpose of improving the functional yield by resizing and reshaping the intended design OPC target into a more lithography friendly one. However, the computational efficiency (speed) of the Aerial Image Retargeting is very important, where this methodology needs to consume less than a fraction of the time-resources required by PWOPC in order to be added to the mask tape-out flow without any noticeable runtime penalty. Moreover, with the OPC target transformation towards the more lithography-friendly regime, it is even possible to reduce the number of OPC iterations as the convergence criteria is met in a smaller number of iterations.

To compute the aerial image signatures we used a Sparse OPC commercial tool (Mentor Graphics Calibre) to calculate the aerial image signature values for each fragment (specifically the I_{max} , I_{min}, slope and the intensity distribution curvature). The basic concept of sparse OPC tools is built on creating an intensity measurement site for each fragment, where each site consists of a number of calculation points. The aerial intensity value is calculated at each point and then the whole intensity distribution among the site is computed through proper interpolation as shown in figure 4.11. Choosing the correct number of calculation points and the site length is very important in optimizing the runtime needed to compute the intensity distribution. The number of calculation points is directly proportional to the CPU time consumed in the option. Also, the calculation points separation is very important in setting the accuracy of the calculations, where setting it to the proper sampling rate makes the interpolation of the continuous intensity distribution accurate

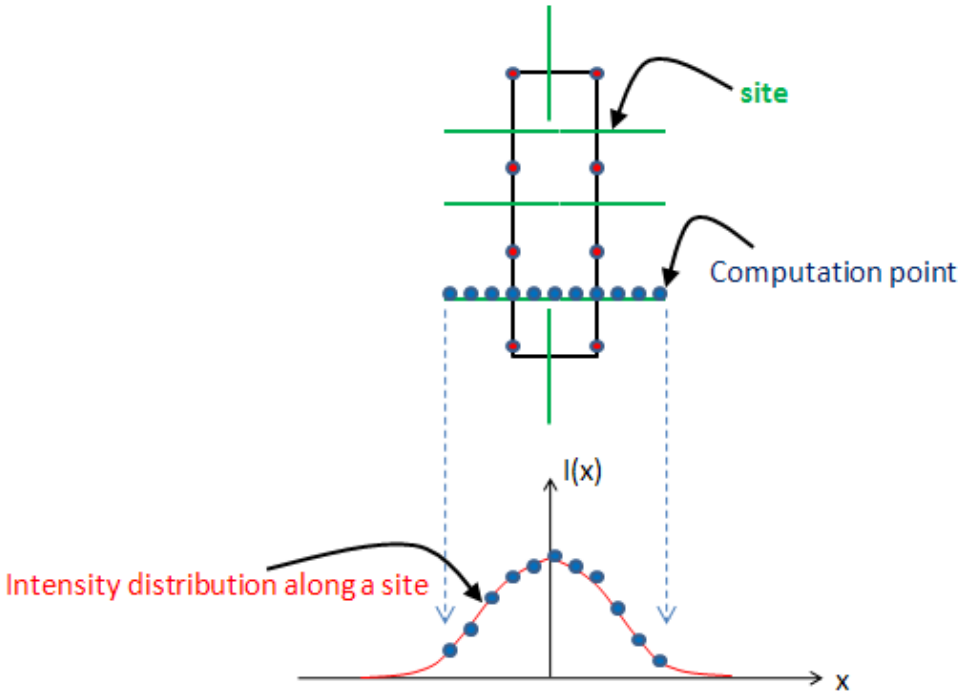


Figure 4.11: The basic concept of Sparse Aerial Image intensity computation.

enough without any oversampling (extra unnecessary computation).

4.3.0.8 AI Analysis

Performing the aerial image signature analysis requires the data gathering of the process-window width and space in the multi-dimensional array of I_{max}-I_{min}-slope-curvature. We limit our attention in this section to the I_{max}-I_{min}curvature. The tables 4.3.0.8 through4.3.0.8 show different PW width and space I_{max}-I_{min} maps for different curvature ranges.

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	35	35	35	37	NA	NA	NA	NA	NA	NA	NA
0.11	NA	NA	36	37	37	36	NA	NA	NA	NA	NA	NA
0.13	NA	35	34	35	33	35	NA	NA	NA	NA	NA	NA
0.15	NA	NA	NA	NA	37	35	36	36	36	37	NA	NA
0.17	NA	NA	NA	NA	NA	31	34	35	36	35	NA	NA
0.19	NA	NA	NA	NA	28	36	35	37	37	NA	NA	NA
0.21	NA	NA	NA	NA	NA	NA	37	37	NA	NA	NA	NA

Table 4.1: Intensity-PW width map for curvature between -6 and -2

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.13	NA	NA	NA	NA	NA	NA	33	33	NA	NA	NA	NA
0.15	NA	NA	NA	NA	34	34	34	36	36	36	NA	NA
0.17	NA	NA	NA	NA	37	37	37	35	36	36	37	NA
0.19	NA	NA	NA	NA	37	35	34	34	36	37	35	36
0.21	NA	NA	NA	NA	37	35	35	31	31	36	35	37

Table 4.2: Intensity-PW space map for curvature between -6 and -2

It is obvious that in the selected technology, there are two distinct signatures for the width and space lithography failures. The good thing is that there is no width-space competition for the same signature, and accordingly we can clearly identify a seperable signature for each of the width and space failures. In the tables above, we colored the tier 1 critical failures in red and then the risky but less critical failures in orange.

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	35	35	34	35	34	37	NA	NA	NA	NA	NA
0.11	NA	NA	30	36	34	35	34	35	37	NA	NA	NA
0.13	NA	NA	35	32	32	34	35	34	36	36	NA	NA
0.15	NA	NA	NA	30	30	33	34	35	35	35	NA	NA
0.17	NA	NA	26	29	28	32	34	35	35	36	NA	NA
0.19	NA	NA	NA	30	32	32	35	36	37	NA	NA	NA
0.21	NA	NA	NA	NA	32	32	34	34	NA	NA	NA	NA

Table 4.3: Intensity-PW width map for curvature between -2 and 2

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.13	NA	NA	NA	NA	NA	NA	37	36	NA	NA	NA	NA
0.15	NA	NA	NA	NA	35	36	36	37	37	38	NA	NA
0.17	NA	NA	NA	NA	37	37	37	36	37	37	39	NA
0.19	NA	NA	NA	NA	37	36	35	35	36	37	36	36
0.21	NA	NA	NA	NA	37	36	35	35	35	36	36	37

Table 4.4: Intensity-PW space map for curvature between -2 and 2

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	NA	NA	33	34	33	34	NA	NA	NA	NA	NA
0.11	NA	NA	NA	31	34	34	34	35	36	NA	NA	NA
0.13	NA	NA	NA	32	32	34	34	32	32	36	NA	NA
0.15	NA	NA	NA	32	32	33	33	31	33	NA	NA	NA
0.17	NA	NA	NA	34	30	32	35	33	36	NA	NA	NA
0.19	NA	NA	NA	31	33	32	34	34	NA	NA	NA	NA
0.21	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.5: Intensity-PW width map for curvature between 2 and 6

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.13	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
0.15	NA	NA	NA	NA	36	37	36	37	37	37	NA	NA
0.17	NA	NA	NA	NA	37	37	37	36	37	37	39	NA
0.19	NA	NA	NA	NA	37	37	35	35	35	37	36	36
0.21	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.6: Intensity-PW space map for curvature between 2 and 6

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	NA	NA	NA	NA	36	33	NA	NA	NA	NA	NA	NA
0.11	NA	NA	NA	NA	NA	35	NA	NA	NA	NA	NA	NA
0.13	NA	NA	NA	NA	33	34	36	36	36	NA	NA	NA
0.15	NA	NA	28	25	33	NA	35	35	35	36	NA	NA
0.17	NA	NA	NA	NA	NA	31	31	33	34	34	NA	NA
0.19	NA	NA	NA	NA	NA	NA	36	35	NA	NA	NA	NA
0.21	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Table 4.7: Intensity-PW width map for curvature between 6 and 10

4.3.0.9 Wafer-Simulation Results

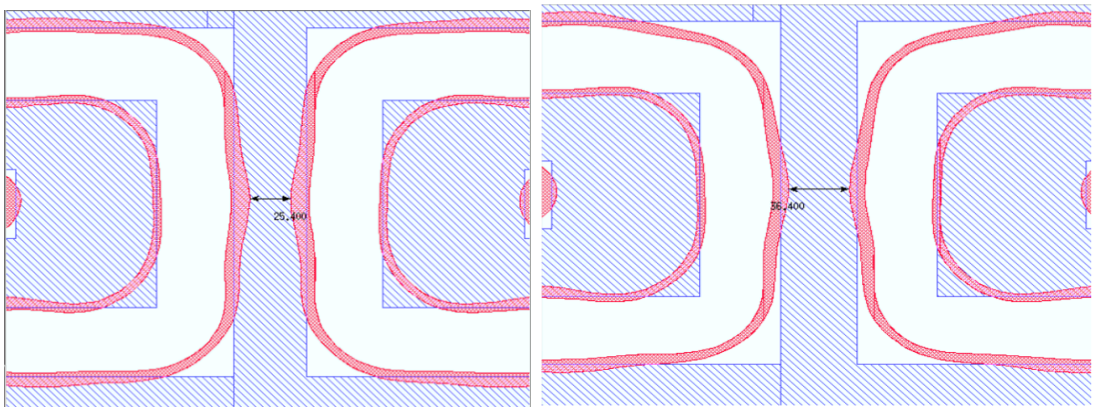
Based on the analysis of the AI maps and their relation to process-window lithography failures, we were able to compile the AIR correction table. In this table we decide the retargeting value needed for each fragment of the design to transform it into a lithography friendly design. Table 4.3.0.9 shows the AIR table where the values listed in it are the correctional movement per edge. During the table building we gave extra care to interpolate for missing biases and smoothing the table to have a reasonable gradient because the process is not expected to vary radically between two adjacent ranges in the aerial image map. We also were able to simplify the AIR table to make it dependent only on the I_{max}-I_{min} (although technically we can still see a present dependence on the intensity curvature too.

Then we applied the AIR module on a full chip to test and quantify the improvement of

I_{min} \ I_{max}	0.25	0.27	0.29	0.31	0.33	0.35	0.37	0.39	0.41	0.43	0.45	0.47
0.09	0	0	1	0	0	0	0	0	0	0	0	0
0.11	0	0	1	1	0	0	0	0	0	0	0	0
0.13	0	0	1	1	1	1	1	1	1	0	0	0
0.15	0	1	2	3	1.5	1	1	1	1	1	0	0
0.17	1	3	3	2	2	1	1	1	0	0	0	0
0.19	0	1	1.5	1.5	1	1	0	0	0	0	0	0
0.21	0	0	1	1	1	1	0	-1	-1	0	0	0
0.23	0	0	0	0	0	0	0	-1	-1	0	0	0

Table 4.8: AIR table

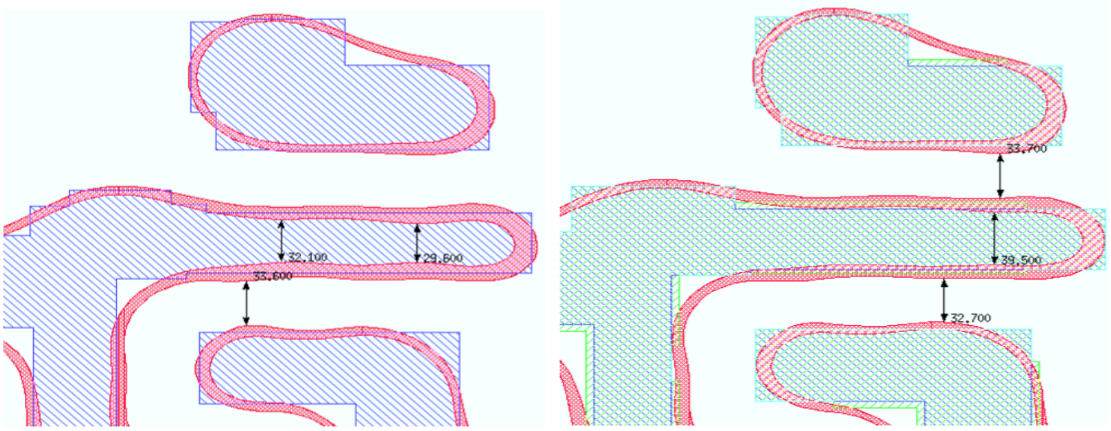
the yield (the reduction of the number of the hot-spots and shifting the worst case width and space towards larger values). Also to demonstrate the real power of AIR and its capability for improving the OPC convergence (as a byproduct next to the yield improvement), we are comparing the full 10-iteations PWOPC recipe to a 6 nominal OPC iterations applied on an AIR target. A sample of the process-window improvement and the retargeting effect on the contour is shown in figures 4.12, through 4.17. It is obvious that AIR automatically transformed the designs into lithography-friendly designs and OPC is capable of converging to the target more quickly.



(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

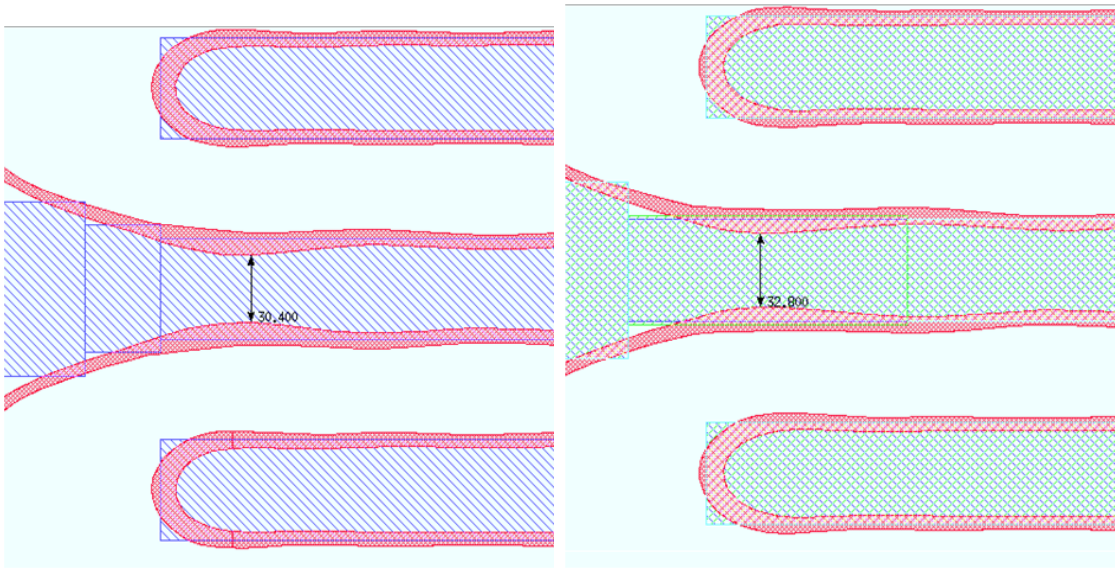
Figure 4.12: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (1).

Figures 4.18 and 4.19 show the distribution of the PW errors on the test chip, where the



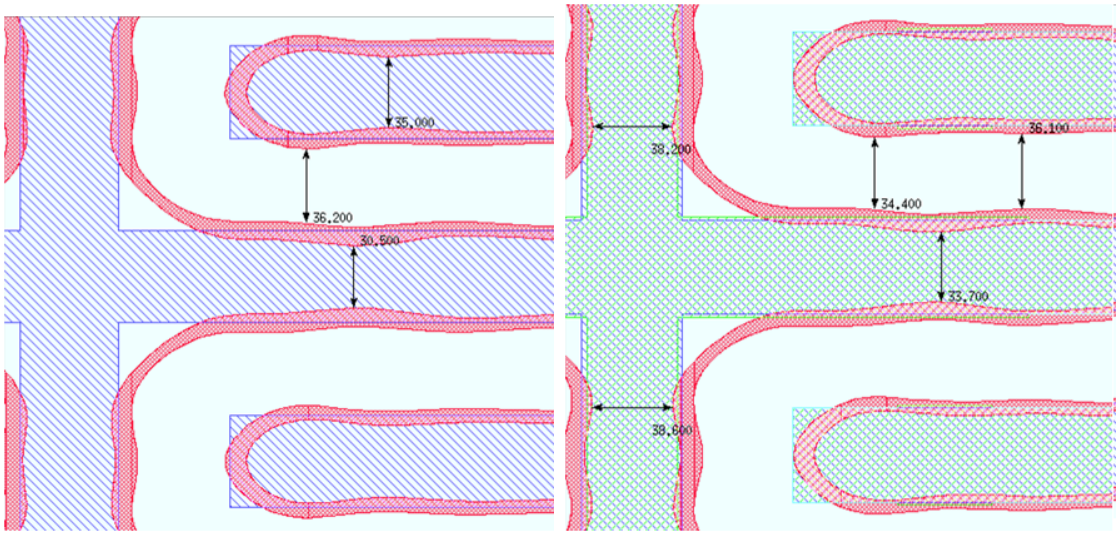
(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

Figure 4.13: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (2).



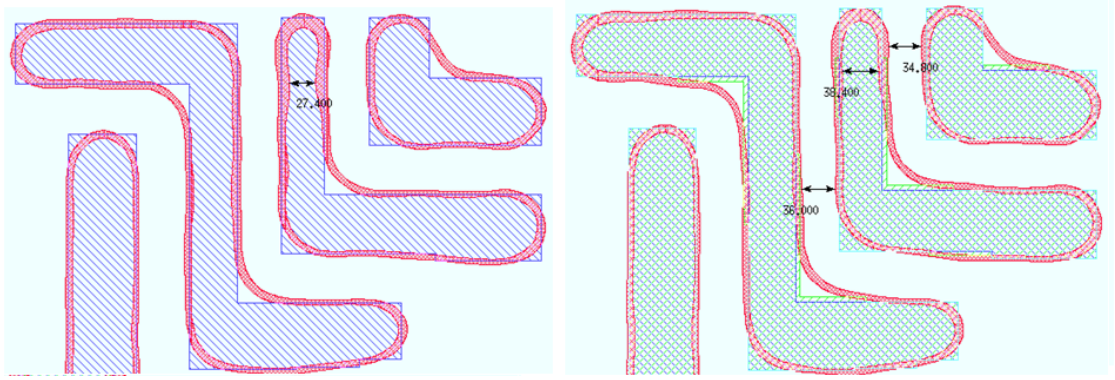
(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

Figure 4.14: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (3).



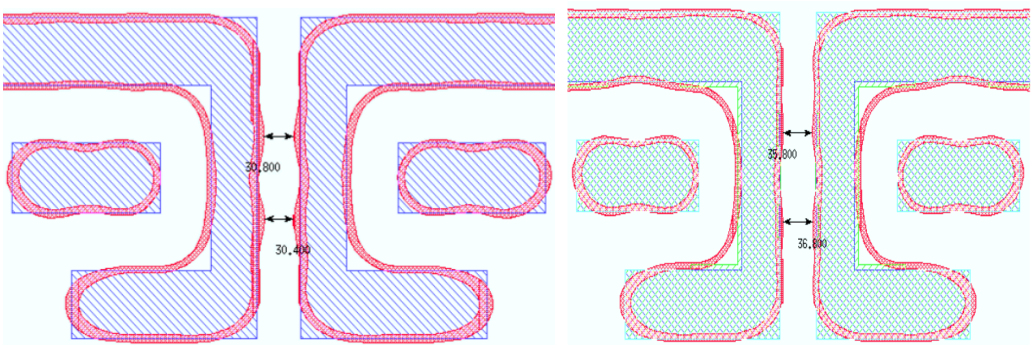
(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

Figure 4.15: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (4).



(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

Figure 4.16: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (5).



(a) PWOPC (10 iterations). (b) AIR_OPC (6 iterations).

Figure 4.17: Comparing PWOPC recipe vs. OPC+AIR for hot-spot (6).

x-axis is the PW error measurement (the smaller the value the worse the design becomes from a lithography perspective) and the Y axis is the PW error count in the chiplet. It is evident from the distribution that the hot-spots distribution has improved a lot for width-related hot-spots and many of the critical errors were translated to the safe range (i.e. transformed into lithography-friendly OPC targets), while for space distribution, it is kept almost the same (which was almost all in the safe region) and moreover fixed a couple of critical soft bridging issues.

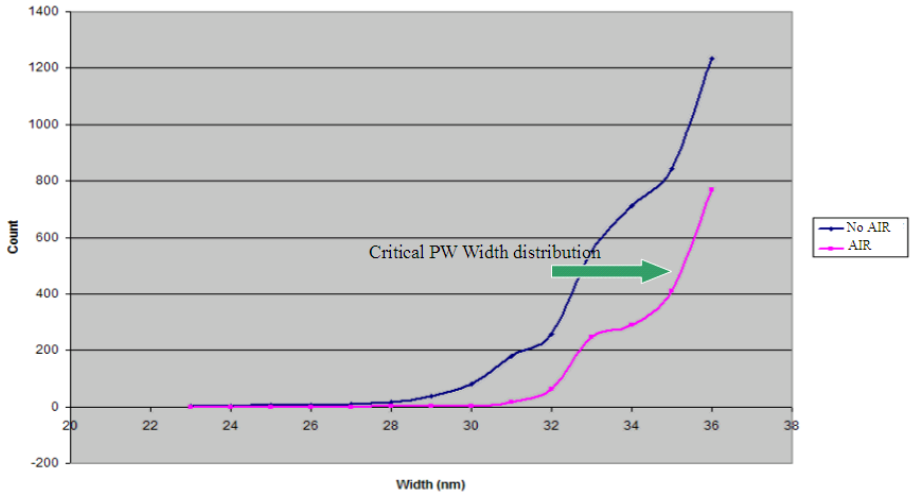


Figure 4.18: Hot-spots Width distribution.

Figure 4.20 shows the overall run-time of the mask tape-out flow with and without

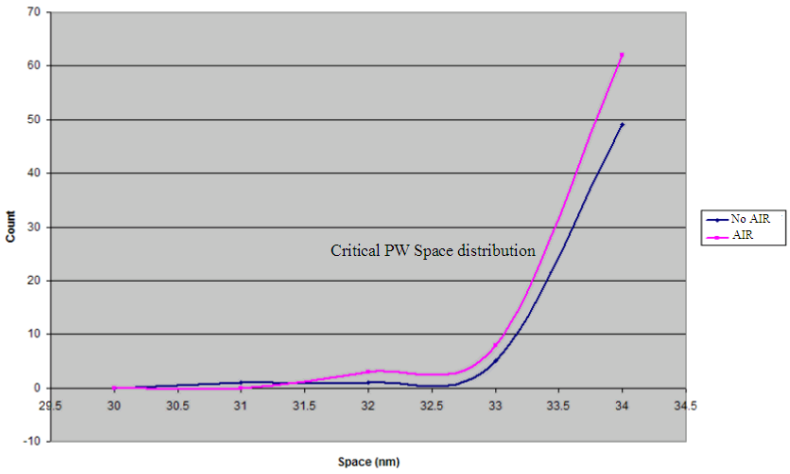


Figure 4.19: Hot-spots Space distribution.

AIR, the introduction of AIR itself didn't consume more than 2% of the total run-time, while due to the clear benefit it gives on OPC, we can see up to 55% overall run-time reduction. The more appealing feature is presented in figure 4.21, where it is obvious that AIR was capable of reducing the hot-spots count from 300 instance to only 5 in the test chip, which represents a significant yield improvement.

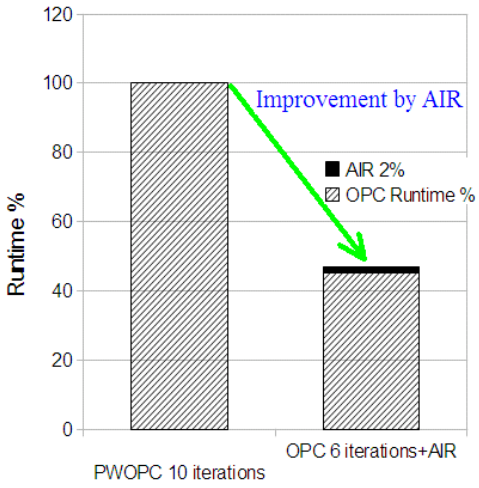


Figure 4.20: Comparing the Runtime of PWOPC and AIR+OPC.

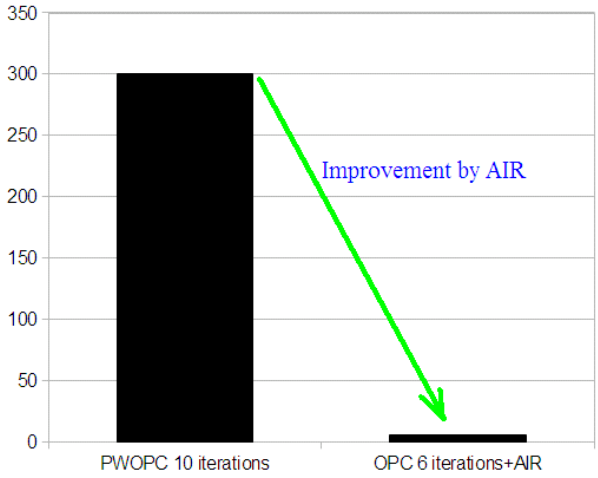


Figure 4.21: Comparing Hot-spots count between PWOPC and AIR+OPC.

4.4 Summary

In the chapter, we presented our early work to prove the feasibility of MBRT and demonstrate its benefits in the mask-tape-out flow as well as its potential role in improving the CD extraction accuracy when included in PDK. We used the aerial image signatures simulated on the target design to identify lithography hot-spots.Then using a lookup table, it is possible to set the re-targeting value based on the aerial image signature to improve the final target patterning quality (i.e. improving the lithography friendliness). The main drawback of AIR is how tedious building the AIR lookup tables. Moreover, if any simple update in the process (a CD change or a dose change, requires rebuilding the tables all over again. In the coming chapters we'll be discussing these drawbacks and developing more robust and applicable MBRT techniques.

Chapter 5

Litho And Yield-Enhancing Retargeting (LAYER): A full Model-Based Retargeting Methodology

5.1 Introduction

In the previous chapter, we demonstrated the feasibility of a fast OPC-independent MBRT. These results were encouraging as they showed the great potential for OPC-independent MBRT, which is capable of capturing the lithography hot-spots using simple simulations and fixing them. This approach is very fast compared to standard flows. In standard verification flows, hot-spots are captured after performing ful OPC and through PW simulations. All these advantages are very nice and encouraging. However, we highlighted that there are still some challenges that need to be addressed in order to present the AIR as a robust solution that can be presented to the science and technology world. The lookup table approach used in the previous chapter is not a practical approach, special if it is based on human judgment to choose the proper bias parameters and update them if a new hot-spot is discovered.

In this chapter, we present a full methodology for MBRT that addresses the above mentioned challenges. We first explain the validity of using the OPC target as the simulation layer and its limitations. Then we discuss the details of our LAYER methodology the

compact LAYER models that we propose to model the hot-spots severity and the necessary retargeting to achieve a final litho-friendly design.

5.2 LAYER: Simulation Layer Construction

The simulation performed in LAYER is using a simplified sized clone of the OPC target and not the actual mask layer. This is a first order approximation that allows decoupling the model-based retargeting from the actual OPC recipe, where this is based on the fact that the mask shape is nothing more than a model-based perturbation to the OPC target to achieve the best design fidelity (CD uniformity) as shown in figure (5.1).

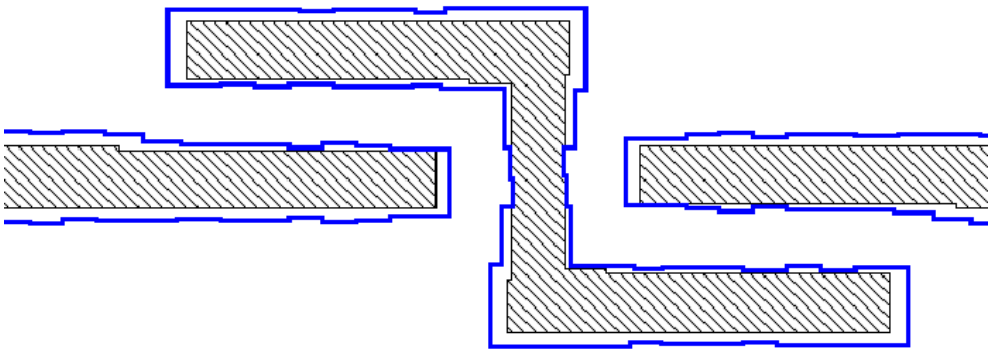


Figure 5.1: An actual example comparing the final OPC mask to the OPC target, showing that the final mask shape is nothing but a perturbed joggy clone of the target shape.

To demonstrate our hypothesis, we plot the distribution of the target-to-mask delta in figure (5.2). It shows that from a geometrical point of view, the mask shape is a model-based perturbation of the target shape, which can be simplified into a global sizing of the target shape with the mean of the distribution.

A comparison between the contours simulated using the final mask versus the sized OPC target is shown respectively in figure (5.3). It is evident that (although there is a shift in the contours CD) the general imaging weakness signature can still be recognized. Using this approximation is of great benefit because it allows us to apply this model-based retargeting methodology with a minimal runtime impact.

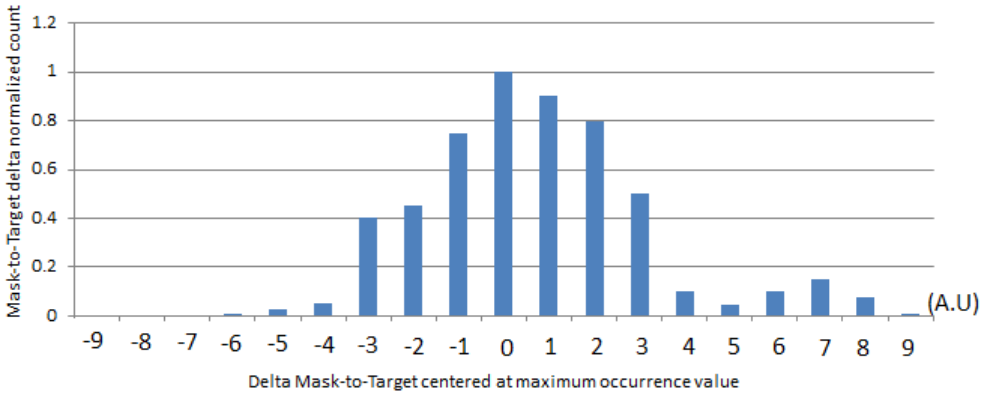


Figure 5.2: An example of the normalized distribution of the mask-to-target delta distribution.

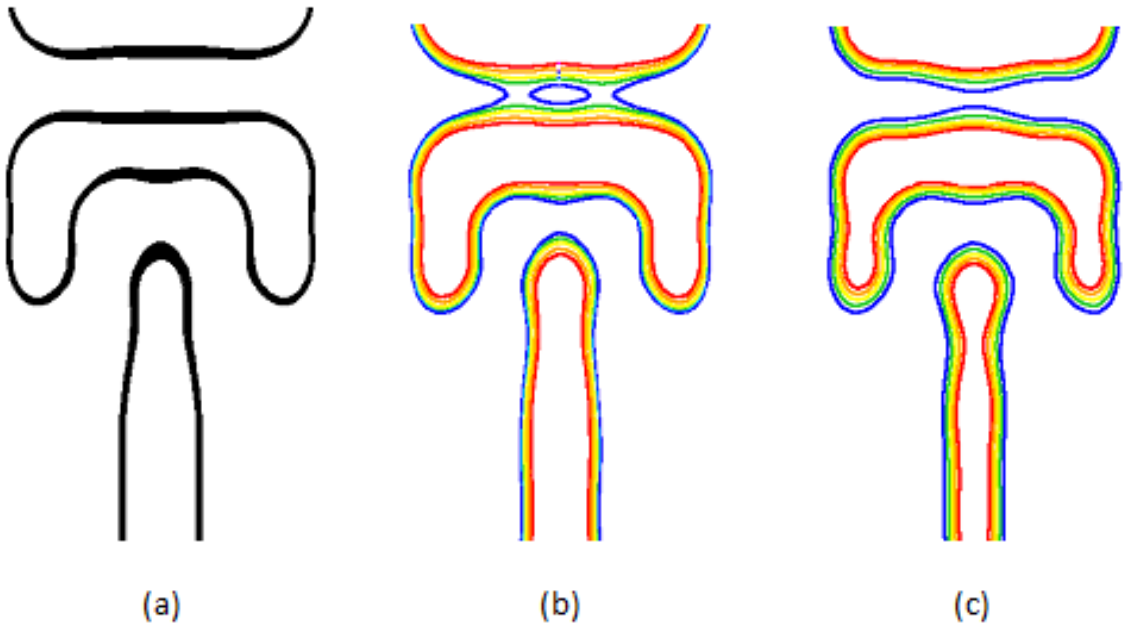


Figure 5.3: Comparing the (a) post-OPC PVBands, optical simulation contours for both (b) post-OPC mask and (c) the pre-OPC target, they could have different magnitudes but they still share a reasonably similar behavior.

5.3 LAYER: Optical Simulation Conditions

Weak designs (from a patterning perspective) can be identified using a number of imaging parameters, among which are the poor imaging contrast, very high Mask Error Enhancement Factor(MEEF) and small Depth of Focus (DoF). We base our work on using specific optical imaging signals to identify patterning hot spots. We use the optical simulation contours to capture such weak spots. As shown in figure (5.4), optical simulation of the perturbed target design can capture areas where the image contrast is weak due to the overall imaged pattern of the design and its proximity.

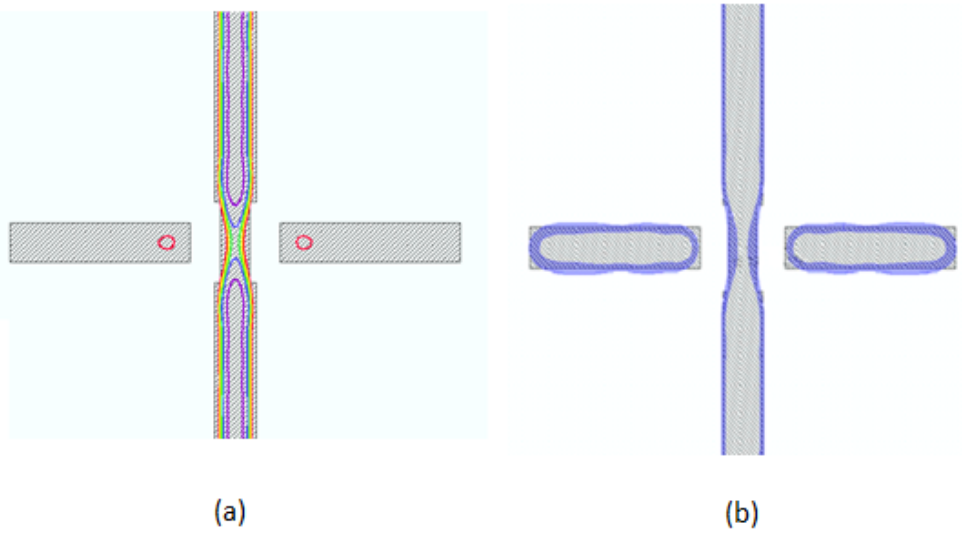


Figure 5.4: Real simulation example showing the validity of using the optical simulation of the OPC target to capture potential lithography hot-spots. (a) Optical simulation contours at different threshold values. (b) Post-OPC PW verification PVBands, showing that the PW weakness is due to design aggressiveness and that it correlates to the contours in (b).

The required simulations for this technique can be as simple as a single optical simulation at nominal focus and using two threshold values to create pinching and bringing contours that will be used in the retargeting recipe development. For a more comprehensive analysis, a more sophisticated simulation strategy can be used to capture all potential hot spots as shown in the Bossung plots in figure (5.5). This more sophisticated strategy doubles the simulation time, but results in a much better capability of capturing lithography hot spots as it combines all of the Mask Error Enhancement Factor (MEEF), Normalized Image Log Slope (NILS), imaging contrast and Depth of Focus (DoF) into the analysis.

The first simulation condition is focusing on capturing design areas that are tending to bridge due to an extra positive sizing of the mask in addition to a decreased illumination dose. On the other hand, the pinching weakness situation can be captured using an off-focus simulation at a mask size down and an overexposed dose.

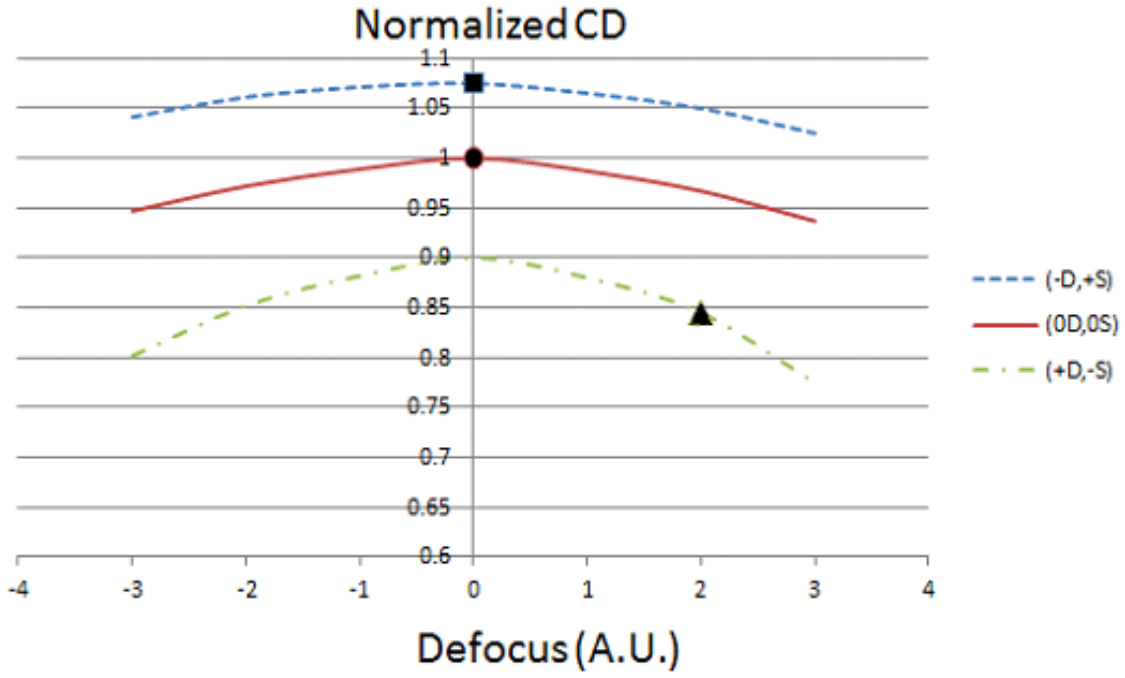


Figure 5.5: An illustration of the process variations Bossung plots and the choice of the simulation conditions to capture lithography hot-spots.

5.4 LAYER: Hot-Spots Modeling

In this section, we explain the proposed flow for lithography hot spots modeling and correlating the post-OPC verification defects to the pre-OPC optical simulations. This is one of the fundamental building blocks of our methodology that will enable the accurate detection of Lithography hot spots without performing any OPC, which allows the decoupling between the hot spot detection from the OPC step as well as saving the very large time required to run OPC.

Figure (5.6) shows the LAYER hot spots modeling flow, where a hot-spots database (HSDB) is used to provide the designs to be used in the HS model calibration. This HSDB constitutes of a large sample of known hot spots and tight challenging designs as well systematically altered versions of these designs to provide a wider coverage of the design space. To quantify the proper retargeting amounts for different hots-pots fragments, the designs would go through an iterative series of manual retargeting, nominal OPC, and through-PW OPC Verification until the proper retargeting is finally achieved. Then to correlate the retargeting values to the simulated conditions, the litho-target is perturbed to approximate the OPC final mask shape as discussed in section (B) and simulated using both the bridging and the pinching conditions simultaneously to extract their optical simulation parameters. This is followed by the HS calibration step, where the required retargeting (fragments displacement) is fitted as a polynomial function of the optical simulation parameters as defined in equation (6.1).

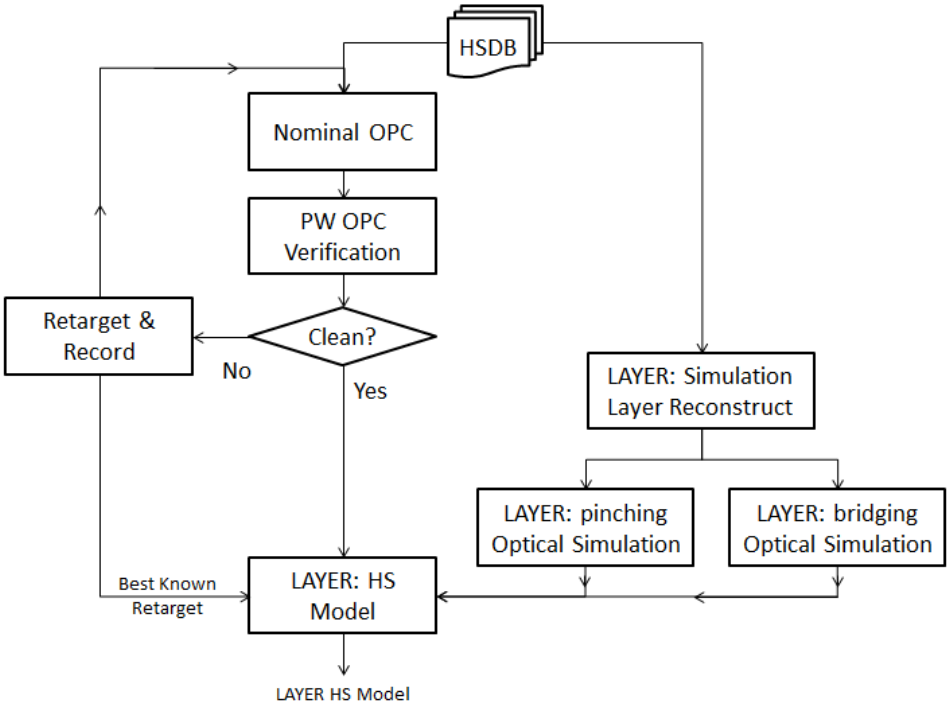


Figure 5.6: Flow chart of the LAYER model calibration.

$$B_j = \sum_{i=1}^{N} K_i (OSP)_{i,j} \tag{5.1}$$

where (B_j) is the retargeting value for the fragment j. (K_i) is the i^{th} coefficient of the retargeting model, and $OSP_{i,j}$ is the i^{th} Optical Simulation Parameter representation of the j^{th} fragment, this can come in all forms of the first or second order combination of all the optical simulation parameters like the maximum intensity (I_{max}), minimum intensity (I_{min}), Slope (S) and Curvature (C).

The full retargeting matrix for all fragments can be described as

$$\begin{pmatrix} B_1 \\ \vdots \\ B_M \end{pmatrix} = \begin{pmatrix} (OSP)_{1,1} & \cdots & (OSP)_{1,N} \\ \vdots & \ddots & \vdots \\ (OSP)_{M,1} & \cdots & (OSP)_{M,N} \end{pmatrix} \begin{pmatrix} K_1 \\ \vdots \\ K_N \end{pmatrix} \tag{5.2}$$

which is a matrix representation of an overdetermined system that can be solved using Singular Value Decomposition (SVD) for this Least Squares problem. This LAYER retargeting model is valid only within a specific range of the of the optical simulation parameters where it is observed to correlate to potential hot-spots, otherwise the retargeting value is zero (i.e. the fragment patterning is robust through PW and doesn't need any retargeting or PWOPC).

This compact model is very fast and adds virtually no additional computation overhead to the simulation and auto-retargeting time. Moreover, it has a very promising application, where not only it captures litho hot-spots and design weaknesses (without doing any extensive OPC or through-PW resist contour simulations), but in addition it is capable of determining the proper retargeting amount that each design fragment needs. The outcome of this calibration stage are two models, one that predicts the pinching hot spots retargeting values and another that predicts the bridging hot spots values. Moreover, it is worth explaining that the reason that nominal OPC is only used in this calibration flow rather than PWOPC is that it is our intention to be able to eliminate the need of PWOPC from the flow and predict any necessary retargeting independent from it.

To verify the validity of the LAYER HS models, we propose the verification flow shown in figure (5.7), where a different set of hot spots database designs (VHSDB) are used in this step than the one used in calibration. The purpose of this step is to verify the LAYER HS models stability and predictive coverage for a wide design space range.

In this flow, first, the VHSDB designs are retargeted through the auto-retargeting LAYER step (will be explained in more details in the next section), then the output is

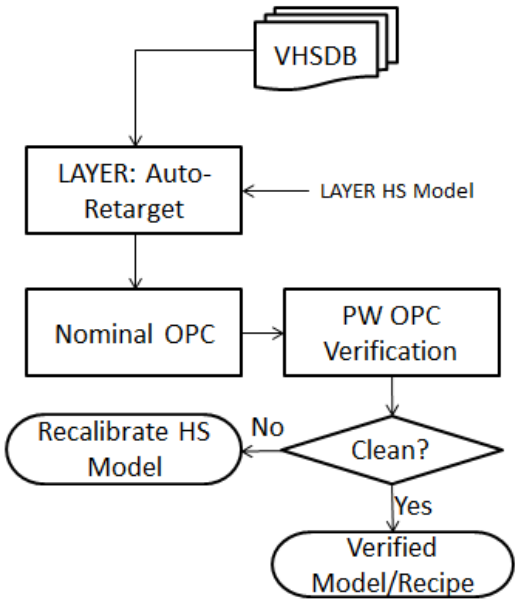


Figure 5.7: Flow chart of the LAYER model verification.

used as the new target for a nominal OPC recipe. This is followed by a through-PW resist contour verification step to confirm that the LAYER HS models were capable of fixing all hot spots through the pre-OPC LAYER step. If the final printing quality does not meet the requirements, then that's an indication of a poor model which requires to further analyze the failing designs to identify the course of action. The most common failure modes are 1) Missing a whole family of problematic designs from the calibration HSDB, where a wide range of the allowed design space was not covered at all, or 2) in a similar sense, some designs were there but had a lower weight during the calibration and the model fitted them marginally. The solution for these situations is to recalibrate the LAYER HS models but this time after improving the design-space coverage/weights in the calibration HSDB.

5.5 LAYER: Auto-Retargeting Methodology and Implementation

The implementation of LAYER is based on the c-based Litho API tool provided by Mentor Graphics Calibre. This is required to create all the required functionalities that is needed

for LAYER modeling and correction. The LAYER auto-retargeting flow is straight forward once the proper LAYER HS models are created as shown in figure (5.8).

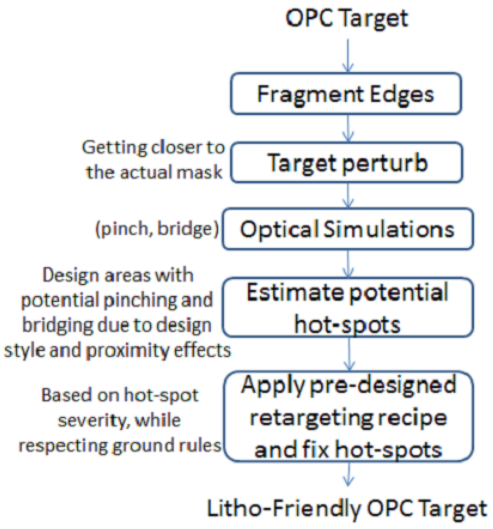


Figure 5.8: The flow diagram of the design Auto-Correction using LAYER.

In this flow, first, the input design is reconstructed to approximate the OPC final mask (this has to match the LAYER HS model calibration simulation geometries reconstruction step). Then this approximate mask is fragmented very similarly to how its edges are going to be fragmented during OPC. This is followed by the simulation using both the pinching and bridging conditions explained earlier (this also has to match the conditions used in the LAYER HS models calibration). After that, the program loops over all fragments and evaluates whether the fragment is considered as a lithography hot-spot or not. If fragments are considered to suffer from a patterning weakness, then the program will compute the required retargeting for each fragment using both the pinching and bridging LAYER HS retargeting models. Finally, the fragments are allowed to move to their final retargeting value, but after enforcing a geometrical minimum width and space as a sanity check that ensures that nothing wrong happens during the retargeting. These minimum width and space can be slightly less than the minimum allowed technology width and space values in order to allow the retargeting and hot-spots fixing in congested design areas, where the designs suffering from patterning weakness are prioritized over healthier designs. In concept, this flow can be an iterative process until the best final target is achieved; however, our current testing show that a single retargeting iteration is good enough to meet the required quality.

5.6 LAYER: Testing Results

In our testing, we focus on 20nm metal levels, where we benchmark LAYER against a PWOPC solution using industrial patterning models. Figures (5.9)−(5.11) show different examples of the actual hot-spots and how their corresponding pinching contours look like before and after retargeting as well as the final post-OPC PW verification contours. The comparison demonstrates the improvement of the through-PW patterning. In this comparison, we are comparing the through PW 3σ PVBands for the LAYER solution (followed by nominal only OPC) vs. a PWOPC solution. These results prove the ability of the pre-OPC optical simulations in capturing potential PW hot spots and fix them using LAYER methodology even before doing any actual OPC.

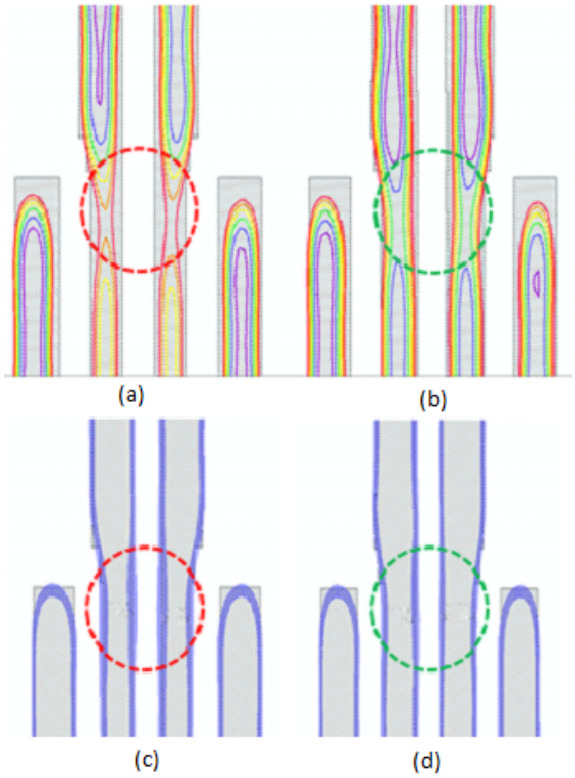


Figure 5.9: An example showing the improvements of the patterning using LAYER. (a) The optical simulation contours of the Target shapes. (b) The optical simulation contours post LAYER. (c) Post-PWOPC PW Verification contours without LAYER. (d) Post-Nominal OPC PW Verification contours with LAYER

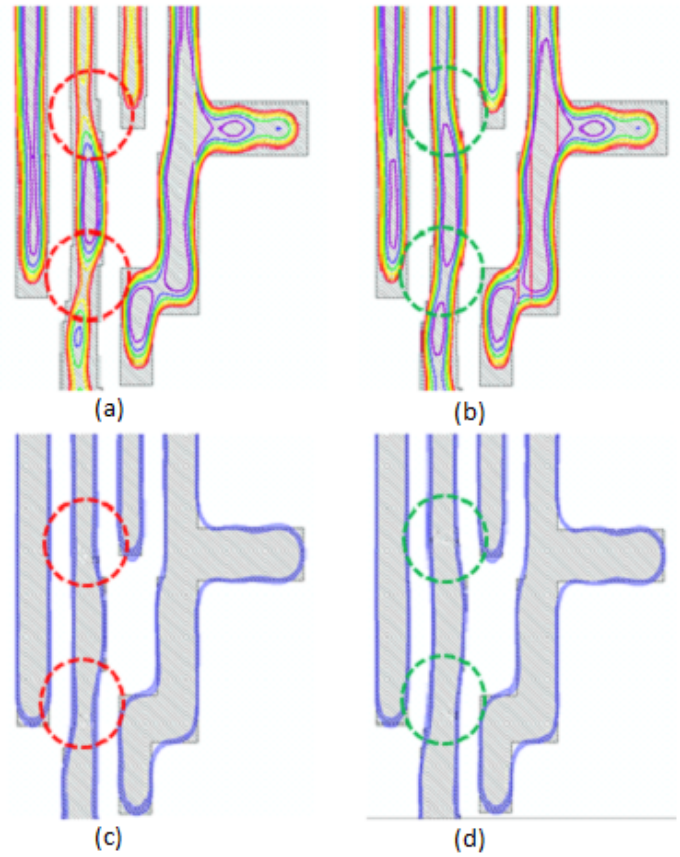


Figure 5.10: An example showing the improvements of the patterning using LAYER. (a) The optical simulation contours of the Target shapes. (b) The optical simulation contours post LAYER. (c) Post-PWOPC PW Verification contours without LAYER. (d) Post-Nominal OPC PW Verification contours with LAYER

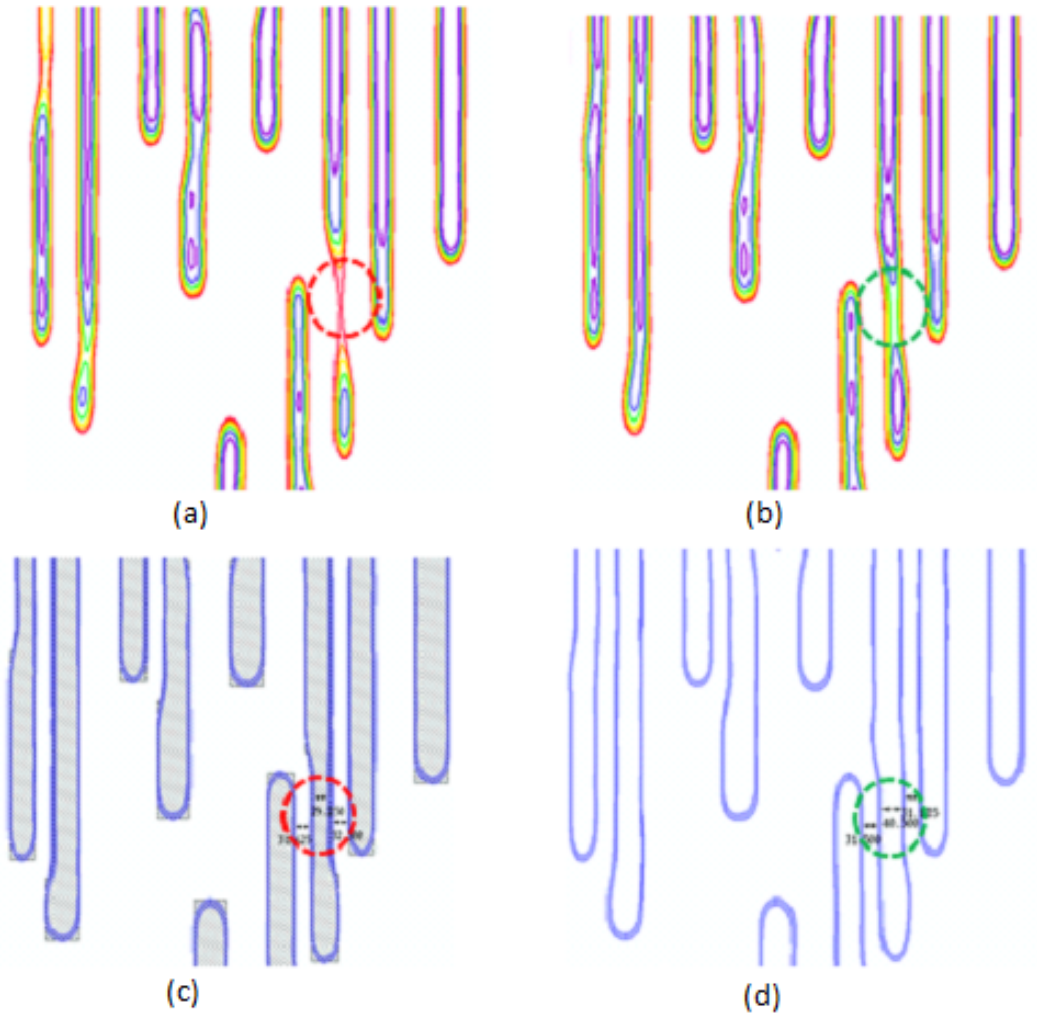


Figure 5.11: An example showing the improvements of the patterning using LAYER. (a) The optical simulation contours of the Target shapes. (b) The optical simulation contours post LAYER. (c) Post-PWOPC PW Verification contours without LAYER. (d) Post-Nominal OPC PW Verification contours with LAYER

The pinch/bridge hot-spots distribution of large testing testcase is shown in figure(5.12).

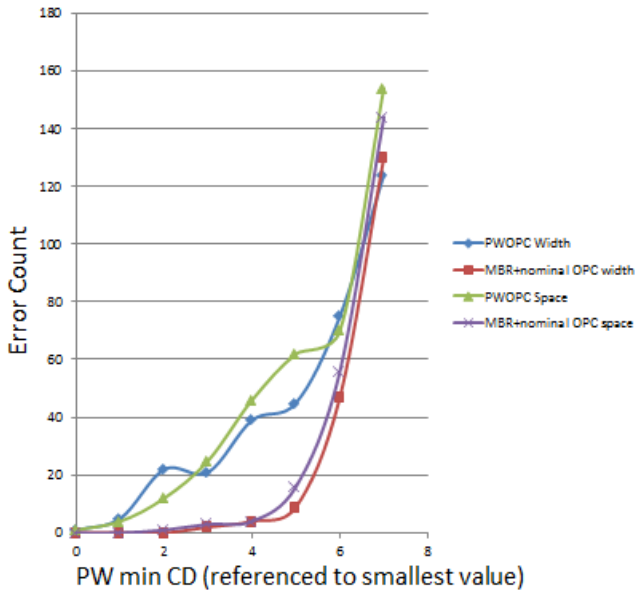


Figure 5.12: The PW width and space hot-spots distribution comparison between the reference PWOPC solution and the proposed LAYER solution. X-axis is the PW CD delta to the minimum CD of the PWOPC hot-spots.

Moreover, for an optimized LAYER recipe, the need for PWOPC is eliminated (or at least highly reduced), which has two main benefits. First it would save a noticeable amount of the computation CPU-runtime needed to tape-out a mask as shown in figure (5.13), where although LAYER consumes almost 5% of the original PWOPC runtime, but the runtime reduction of the migration of the computationally intensive PWOPC into nominal OPC still saves an overall runtime of 41%. Second, the decoupling of the LAYER from the OPC recipe offers a computationally efficient flow to feed back more accurate final design dimensions to the designers in their parasitics-extraction without any need to go through the full OPC flow.

5.7 Conclusions

In this chapter, we presented LAYER as a novel methodology to quickly identify potential lithography hot spots using their optical simulation imaging parameters. The simulation is

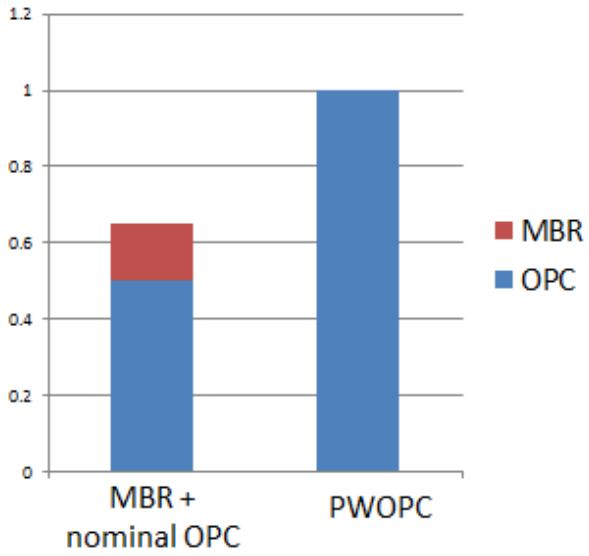


Figure 5.13: Normalized resource-runtime product comparison between the reference PWOPC recipe and the LAYER+nominal OPC solution.

performed on a perturbed clone of the OPC target and did not use the actual final mask. This full flow solves the problems of AIR explained in the last chapter, and still keeps the speed advantage, which allows LAYER to be efficiently integrated in both the mask tape-out flow as well as LFD flows.

We explained the overall flow as well as the details of the creation of the LAYER compact models. These models are calibrated to capture the hot-spots locations and severity based on the purturbed target layer. We developed the algorithms explained in this chapter using the Mentor Graphics c-based Litho-API tool. The full flow was tested on a 20nm metal level and showed strong potential in capturing hot spots and correcting them even without going through any OPC. This technique is very fast compared to any other known approach and consumes no more than a 5% of a typical OPC runtime. All these advantages give this proposed methodology the potential to be used also in both the Fab's mask tape-out flow as well as their LFD kits.

This new MBRT flow shows many advantages on 20nm designs. However, we need to investigate the requirements and challenges for more advanced technology nodes like the 10nm node. As we'll explain in more details in the coming chapters, there are two main challenges that appear more severely in 10nm. First, The simulation approximation that we used in LAYER, to approximate the mask is not very useful anymore due to

the larger mask bias in the 10nm node as well as the tighter ground rules. Second, we are observing some situations, where some aggressive designs are not solvable using the standard approaches anymore. This is again due to the tighter ground rules and the more complex multiple patterning requirements.

Chapter 6

Model-based Initial Bias (MIB): Towards a Single-Iteration OPC

6.1 introduction

In the previous chapter, we demonstrated the MBRT solution using the LAYER methodology. LAYER used a simplified version of the target as an approximation for the final mask and to capture potential hot-spots that appear mainly due to the overall constructive and destructive interference of the design imaging. This was acceptable during our 20nm metal level testing.

However, this approximation is not very valid in more advanced nodes. The tighter design ground rules result in two effects that weaken this approximation.

1. The smaller CDs, specially with multiple patterning, result in a larger mask-to-target bias. Which means that the mask distribution is wider.
2. Tighter ground rules result in closer interaction of adjacent features and accordingly, the same design OPC has a bigger variance than before due to stronger neighboring design interaction.

It is still within our scope to decouple the MBRT from the actual OPC recipe mainly due to the speed benefit. As running the full OPC before MBRT makes the solution loses its primary advantages of speed (and decoupling from the OPC recipe details, which has its benefits from an LFD point of view as we'll explain in a coming chapter).

In this chapter we develop a new methodology for a fast model-based technique to approximate the final OPC mask based on the optical simulation of the initial target. This technique has two main benefits. First, in the field of OPC itself, this model-based initial bias approach is capable of improving the initial condition of the iterative OPC process and resulting into an improvement in the OPC convergence into the final mask. Second, with a good accuracy (good OPC initial bias) this technique's output is a much better starting point for MBRT as the simulation layer more accurately represents the final mask.

In the implementation of MIB we used the c-based Litho API tool provided by Mentor Graphics Calibre to do all the necessary coding to create the required functionalities on both the modeling and the correction aspects. We test it on 10nm metal and via layers as two of the most design-diverse critical layers in 10nm technology.

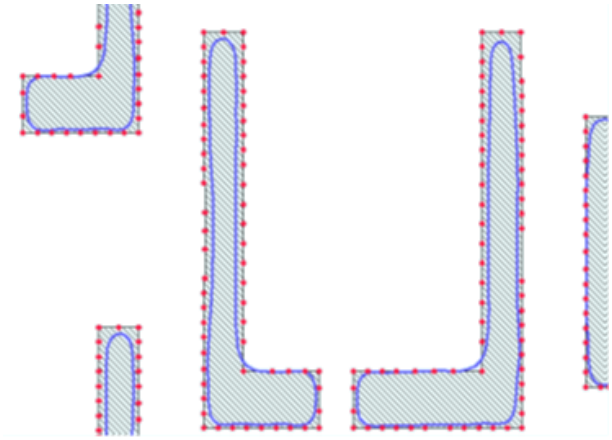
6.2 Background

In OPC, the design edges are fragmented into small fragments and each fragment is allowed to move independently to improve its own design fidelity objective function (dominated by Edge Placement Error (EPE)). The OPC fragmentation and how the no-OPC patterning contours are shown in (6.1(a)), where it is clear that the final printing is not meeting the necessary pattern fidelity. In figure, (6.1(b)) we show how the final mask shape is perturbed to achieve the design fidelity dictated by the industry.

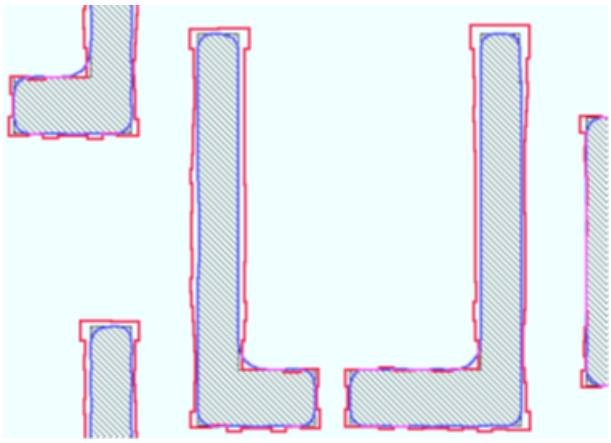
Figure (6.2) shows a simplified flow chart to the standard OPC flow, where the OPC correction engine iteratively tries to forward solve this problem in order to find the optimum mask shape that compensate for all the extreme diffraction that the design information is suffering from during the photo-lithography process. During each OPC iteration, the OPC engine simulates the optical and resist responses to the mask in hand. Then the difference between the designed pattern target and the simulated on-wafer printed pattern (EPE) is computed for each fragment, followed by movement of fragments by the OPC engine according to the deviation from its own target.

6.3 MIB: Model-based Initial Bias

Model-Based OPC is the process of iteratively perturbing the OPC target shape to compensate for the optical diffraction and photo-resist chemistry in order to compute the mask shape that can achieve the best final patterning on silicon. Figure (6.3) shows an example



(a)



(b)

Figure 6.1: A demonstration of OPC basic concepts, (a) Pre-OPC fragmentation and Litho simulation for a pre-OPC mask, where the dots represent the fragmentation points, while the contour predicts the wafer printing if no OPC is used. (b) Post-OPC mask-shape (red) and Litho-simulation (blue).

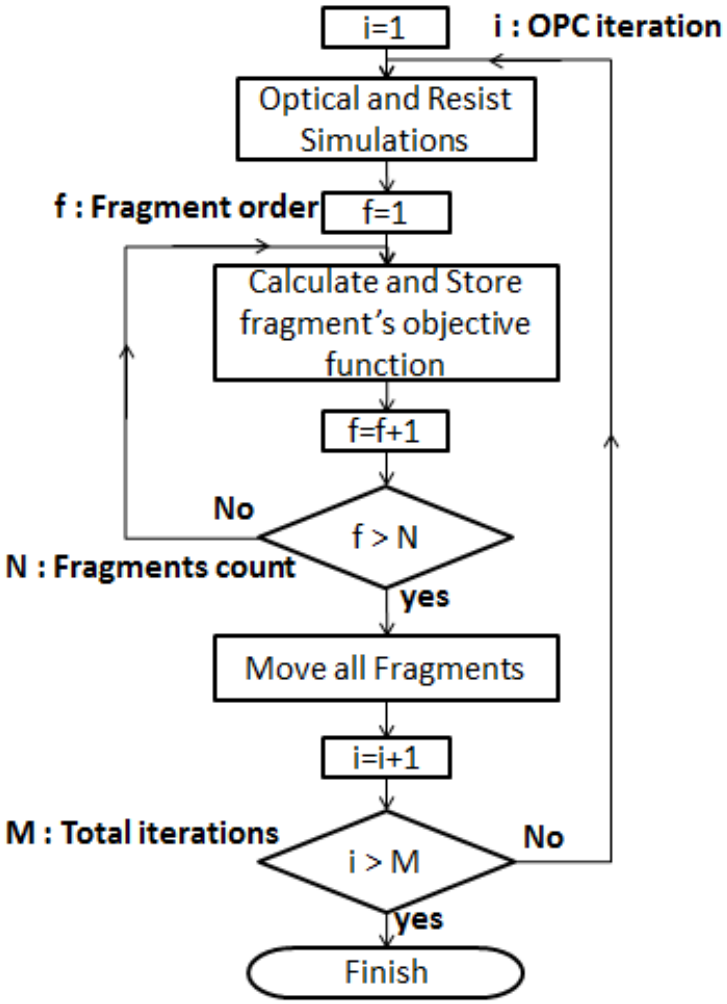


Figure 6.2: A simplified flow chart explaining the how traditional OPC works.

of the evolution of the mask shape from the initial target into the final mask shape and how the printing improves along with the iterations' progress. We can notice that in the initial iterations, the OPC engine is trying just to achieve reasonable printing before spending the last few iterations in doing the actual CD control job.

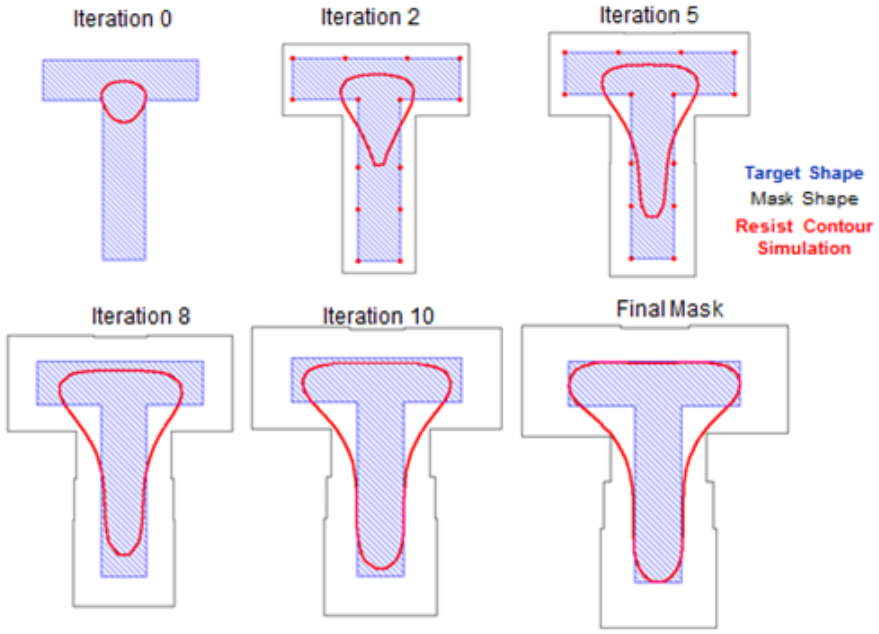


Figure 6.3: An example of the evolution of the mask shape.

It is our goal to speed up the OPC by improving the convergence process. This is very valuable on many aspects, among which is to save the cost and Turn-Around-Time of the mask tape-out. We build our fast OPC methodology on improving the OPC initial condition and making it as close as possible to the final mask shape. In other words, we try to eliminate the need for the first several iterations of OPC and saving all that unnecessary computation. To achieve that, we are proposing the Model-based Initial Bias (MIB) methodology. This methodology is explained in detail in the coming subsections, where in section (A) we introduce the basic concept. This is followed by the detailed description of our proposed MIB compact model for achieving the concept of the Model-based Initial Bias in section (B). Then finally, we explain the whole OPC correction flow after the incorporation of MIB in section (C). In the implementation of MIB we used the c-based Litho API tool provided by Mentor Graphics Calibre to do all the necessary coding to create the required functionalities on both the modeling and the correction aspects.

6.3.1 MIB: The concept

The basic concept behind MIB is to model the fragments' displacement between the OPC target to the final OPC as shown in figure (6.4). Then using this model to aid the OPC in starting with an improved initial state as shown in the illustration in figure (6.5). It is of great importance that this compact model satisfies two main criteria in order to successfully achieve the proposed task. First, the MIB model must be capable of capturing the OPC displacements accurately, where MIB models need to extend its coverage across all the supported design space. Second, MIB model must to be very fast compared to the regular OPC iterations, where the only way for MIB to be efficiently inserted in the OPC correction flow is to show an actual benefit on the overall OPC runtime.

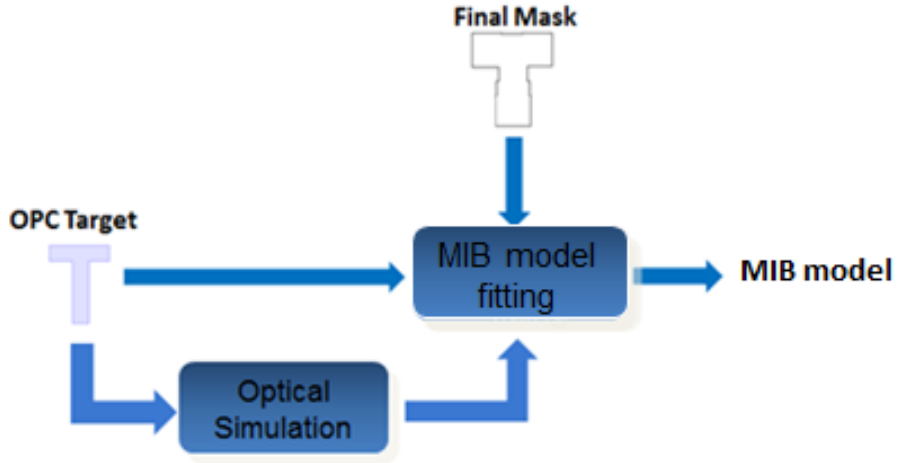


Figure 6.4: An illustrative diagram of MIB model calibration.

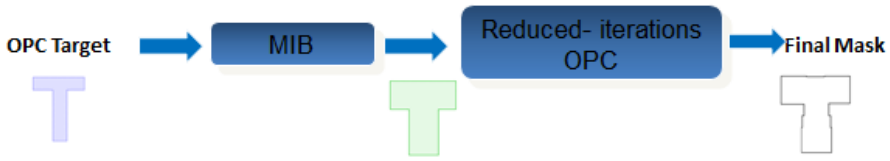


Figure 6.5: An illustrative diagram of using MIB to improve the OPC initial condition.

6.3.2 MIB: Modeling

6.3.2.1 General MIB Modeling

In our work, we propose using the optical simulation parameters as the main parameters of our MIB models. This is due to two main reasons. First, because the phenomenon that OPC is trying to compensate is mainly the optical diffraction in addition to the resist response, which is highly dependent on the imaging signal inside the resist too. Second, the optical signal is still capable of implicitly capturing variations of the proximity of the design as well as the density variations. Optical simulation parameters are not only capturing design variations, but also they capture the Sub-Resolution Assist Features (SRAFs) surrounding the design. The optical simulation parameters of interest are explained in figure (6.6), where I_{max} is the maximum optical intensity value along the simulation site placed at the center of the OPC fragment. Similarly, I_{min} is the minimum optical intensity value along the simulation site. S is the slope of the intensity distribution along the site and finally, C is the contour curvature at the simulation site. The MIB model is described in equation (6.1).

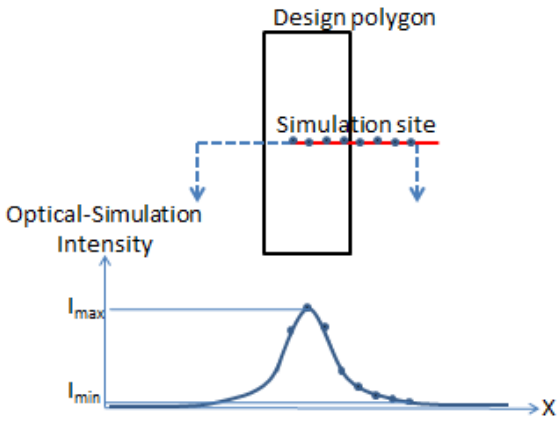


Figure 6.6: An illustration of the main intensity parameters along the simulation sites.

$$B_j = \sum_{i=1}^{N} K_i (AI)_{i,j} \tag{6.1}$$

where B_j is the model-based initial bias for the fragment j. K_i is the i^{th} coefficient of the bias model, and $(AI)_{i,j}$ is the i^{th} Aerial Image representation of thej^{th} fragment, this can come in all forms of the first or second order combination of all the major aerial image parameters I_{max}, I_{min}, Slope and Curvature). In addition to the linear AI terms, our analysis and testing results (as will be shown in the results section) show that the presence of inverse proportionality terms of the AI parameters improves the model predictive capabilities. The following inverse proportionality terms are added to the model parameters.

$$Inverse\ Imax\ Term(IX) = \frac{1}{I_{max} - I_{max0}} \tag{6.2}$$

$$Inverse\ Imin\ Term(IN) = \frac{1}{I_{min} - I_{min0}} \tag{6.3}$$

where I_{max0} is the I_{max} reference value. This value is optimized with a restriction from getting larger than a certain value that could risk making the correction value unstable. Similarly, I_{min0} is the I_{min} reference value designed also to capture the increased OPC-to-target delta as the Imin value increases.

The full solution matrix for all fragments can be described as

$$\begin{pmatrix} B_1 \\ \vdots \\ B_M \end{pmatrix} = \begin{pmatrix} (AI)_{1,1} & \cdots & (AI)_{1,N} \\ \vdots & \ddots & \vdots \\ (AI)_{M,1} & \cdots & (AI)_{M,N} \end{pmatrix} \begin{pmatrix} K_1 \\ \vdots \\ K_N \end{pmatrix} \tag{6.4}$$

which is a matrix representation of an overdetermined system that can be solved using Singular Value Decomposition (SVD) for this Least Squares problem as shown in the flow diagram shown in figure(6.7).

6.3.2.2 MEEF-Aware MIB Modeling

Improving the OPC convergence is the ultimate goal of MIB. This should not be overlooked during the MIB model calibration. Accordingly, it is important to understand the mechanisms that can affect the OPC convergence. One of the most important mechanisms is the Mask Error Enhancement Factor (MEEF), which is the quantitative correlation

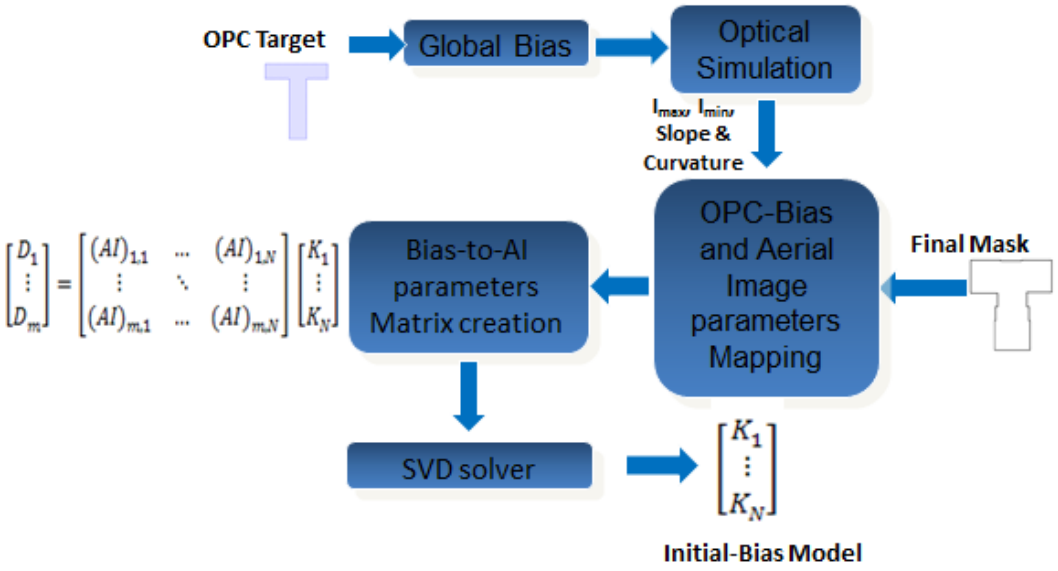


Figure 6.7: An illustration of MIB model calibration flow.

between the mask perturbation and the corresponding on wafer (simulated contour) perturbation. MEEF is explained in equation (6.5) as

$$MEEF = \frac{\partial w}{\partial m} \tag{6.5}$$

where w is the wafer printing CD and m is the mask CD. A variation in the mask CD affects the printing CD depending on the fragment's MEEF value.

A high MEEF fragment has a large impact effect on the printing on wafer, where a small movement results in a large printing variation. These fragments are the ones who suffer the most during OPC, where they are subject to an oscillatory behavior. Accordingly the MIB model needs to be very accurate for those. Moreover, self-MEEF is not the only MEEF description. In its generalized form, MEEF can be represented in its matrix form to include the effect of a fragment on the printing of nearby fragments too [**?**].Equation (6.6) describes the generalized printing variation as a function of the fragment's mask movement as well as its surrounding fragments' mask variations.

$$EPE_i = \sum_{j=1}^{N} \frac{\partial w_i}{\partial m_j} . \Delta m_j \tag{6.6}$$

where the EPE (Edge Placement Error) of the i^{th} fragment is represented as a function of the mask errors of the adjacent fragments and the cross-MEEF terms.

It is necessary to have the MIB model error minimized also for fragments with high cross-MEEF values. This is to ensure that the initial condition provided to OPC is of the best accuracy for fragments that are sensitive to model error variations.

Finally, there is another set of important structures that could arise in some designs, which is the relatively low MEEF fragments. These are the fragments that react very slowly to the mask's movement and require lots of OPC iterations (and mask displacement) to reach their final proper solution. These fragments also need to be prioritized during MIB modeling.

The fragments that are easiest to converge during OPC are the ones that have a self-MEEF that is close to the inverse value of the feedback value used during OPC. Where theoretically, if the cross-MEEF is zero, a fragment with a MEEF equal to 1/feedback would converge in a single iteration. Accordingly, in the MEEF-aware MIB modeling we use equation (6.7) to provide the weights during the MIB model calibration.

$$w_i = f(GMEEF_i) = w_o + W(GMEEF_i - \frac{1}{FB_o})^2 \tag{6.7}$$

where w_i is the weight of the i_{th} fragment during the model calibration. w_o is the minimum weight value. W is an additional weighting factor that we use to control the weight variation as a function of the MEEF. Additionally, we define $GMEEF_i$ as the combined printing sensitivity to the fragment and all its adjacent fragments and is presented by equation (6.8), which is an approximation as if all the surrounding fragments moved the same movement and then the EPE in equation (6.6) is normalized to this movement. And FB_o is the average feedback value used in OPC.

$$GMEEF_i = \sum_{j=1}^{N} \frac{\partial w_i}{\partial m_j} \tag{6.8}$$

The final MEEF-aware MIB model calibration flow is explained in figure (6.8) and the pseudo-code 1, where first the OPC target (T) is read in. This OPC target is a set of structures that are designed to cover the full allowed design space. Then regular OPC is executed for this target to compute the final mask shapes (M). The Nominal lithography simulation (NS) is then calculated to be used as a reference during the GMEEF calculation. This is followed by post-OPC mask perturbation (P) using a global sizing (in the order of

1nm) and then another nominal lithography simulation (PS) is performed on the perturbed mask. This is then followed by the looping over all the design fragments to collect the OPC displacement, optical simulation information and GMEEF to use them in the creation of the matrices that will be used in the fitting of the MIB model. In our work, we use Singular Value Decomposition (SVD) to solve this Least Squares problem. This whole flow is repeated several times to optimize for the best inverse proportionality terms I_{max0} and I_{min0}, which cannot be easily optimized altogether with the MIB model coefficients using SVD. The model selection is based on the fitting quality comparison across different I_{max0} and I_{min0}.

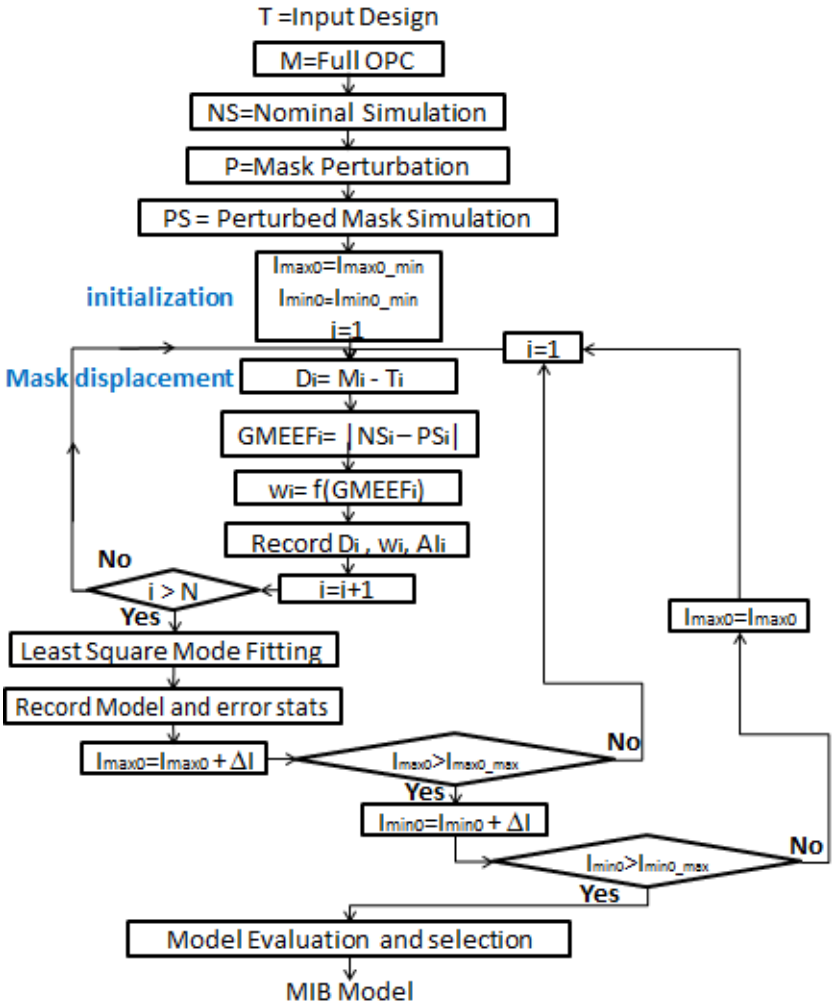


Figure 6.8: A detailed Flow Chart for the MEEF-Aware MIB Model Calibration.

Algorithm 1 MEEF-Aware MIB model Calibration

```
1: procedure MIBCalibrate(LithoTarget,FinalMask)
2:     NS ← Nominal Contour Simulation
3:     SizedMask ← Size FinalMask by 1nm
4:     PS ← Contour Simulation on SizedMask
5:     for each Fragment Frag ∈ LithoTarget do
6:         Disp(Frag) ← FinalMask(Frag)-LithoTarget(Frag)
7:         GMEEF(Frag) ← PS(Frag)-NS(Frag)
8:         weight(Frag) ← f(GMEEF(Frag))
9:         for Initial Sizing S in 0,1,....10 do
10:            AI(Frag,S) ← Imaging parameters of S-sized LithoTarget
11:    for Initial Sizing S in 0,1,....10 do
12:        MIB(S) ← Least Square Fit (AI(S),weight,Disp)
13:        errorrms(S) ← Fitting error (AI(S),weight,Disp)
14:        if errorrms(S) < errorrms_min then
15:            MIBFinal ← MIB(S)
16:            errorrms_min ← errorrms(S)
```

6.3.3 MIB: Correction

The correction flow using MIB is relatively straight forward once the well-predicting MIB models are calibrated. MIB is inserted at the very beginning of the OPC, where it performs an optical simulation to extract the MIB model parameters, then looping over the target fragments and assigning the proper initial bias that each needs as shown in figure (6.9) and the pseudo-code 2. Then an iterations-reduced OPC follows to do the final tweaking to achieve the final mask shapes. A c-based Litho API code is developed to provide this additional add-on model-based bias functionality.

Algorithm 2 MIB-Based OPC Flow

```
1: procedure MIBCorrect(LithoTarget)
2:     AI ← Optical Simulation
3:     for each Fragment Frag ∈ LithoTarget do
4:         MIBDisp(Frag) ← MIBModel(AI(Frag))
5:     OPCInitialMask ← Move LithiTarget Fragments by MIBDisp
6:     FinalMask ← OPC(LithoTarget,OPCInitialMask)
```

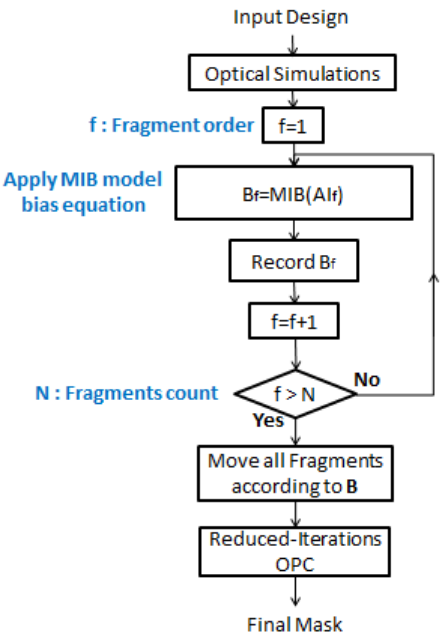


Figure 6.9: A detailed Flow Chart for the integration of MIB inside the OPC flow.

6.4 Testing Results

In this paper, we focus our testing on the 10nm technology. We chose two critical and very different (from a geometrical point of view) layers. First we applied MIB to metal levels, where metals are usually the most complicated from an OPC point of view due to the very wide variety of designs to support in addition to the complicated two-dimensional design variations that they are allowed to take. Moreover, metals one-dimensional shapes (lines) are still of great importance due to their large recurrence in designs. The other layer-set that we used in our testing are the vias, which are usually small square or rectangular shapes but their final mask CD is very dependent on their proximity and the adjacent via shapes. In our testing, we'll focus on the models fitting quality as well as the final OPC convergence improvement for both layers.

6.4.1 Metal Layer Testing

In our testing for the metal level, we used industrial 10nm OPC models to quantify the benefits when MIB is applied in an actual OPC benchmarking. First, we generated a large

set of 1D and 2D structures that cover a wide spectrum of the allowed design space for the metal layer. Then we created two MIB models, the first has all structures equally weighted, while the second is a MEEF-aware MIB model. The fitting rms value for the models is 3.294 nm and 3.865 nm respectively. These relatively small values indicate that our MIB model is accurate enough to get the majority of the OPC fragments a few nanometers close to their final position.

We benchmarked the generated MIB models against regular OPC on a large random logic block design, which consists of a very large collection of standard cells placed in all sort of random placements to generate a very comprehensive practical test case for 10nm designs. In our analysis, we rely on two very important metrics when we compare the proposed MIB to the reference OPC. First, is the rms value of the delta-to-final OPC mask evolution during the OPC iterations. Second, is the EPE rms evolution during the OPC iterations. These two metrics are somehow equivalent as they both test the conversion speed of MIB compared to the reference OPC, but studying them separately allows us to demonstrate the key improvements offered by MIB.

Figure (6.10), shows the how the delta between each fragment's position and its final position (combined rms value for the whole random logic block) improves during OPC. We plot the reference OPC, which obviously converges at 10 iterations. MIB is shown to improve the initial condition of the OPC, which accordingly reduces the necessary OPC iterations. There is around 30%-35% improvement in the number of iterations due to MIB. However, the more interesting results are for the MEEF-Aware MIB, where although the initial rms value seems worse than the regular MIB solution, but it actually is converging faster, and the fragments that are not at their very best starting position are having a medium range MEEF value, which allows them to converge faster, while in contrast to the regular MIB and the reference OPC its high MEEF fragments are already closer to their final position. And accordingly, the OPC is capable of converging less than 50% of the number of iterations of the reference OPC.

As for the EPE convergence, we show the EPE rms progression through OPC in figure (6.11). The results are in agreement with what we explained earlier that the MEEF-Aware MIB solution has a better convergence. Moreover, the because of their better prediction of high-MEEF fragments the overall starting EPE rms value for the MEEF-Aware is actually better than the regular MIB model and orders of magnitude better than the reference OPC EPE rms. This is because even for a small mask error a high MEEF fragment results in a large EPE.

In figure (6.12), we demonstrate a sample of the MEEF-Aware MIB output compared to the final mask shape, where it is obvious how MIB is capable of predicting the final

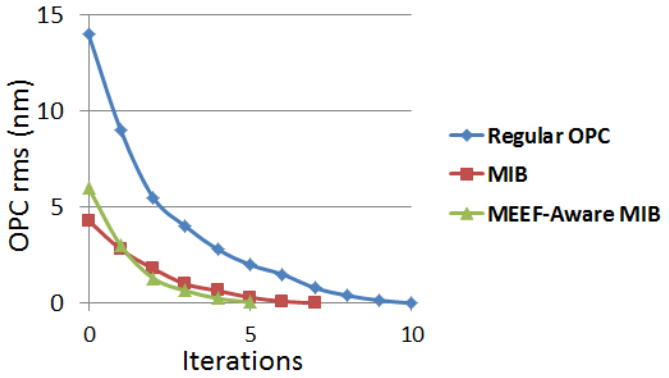


Figure 6.10: OPC mask conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of metal layer.

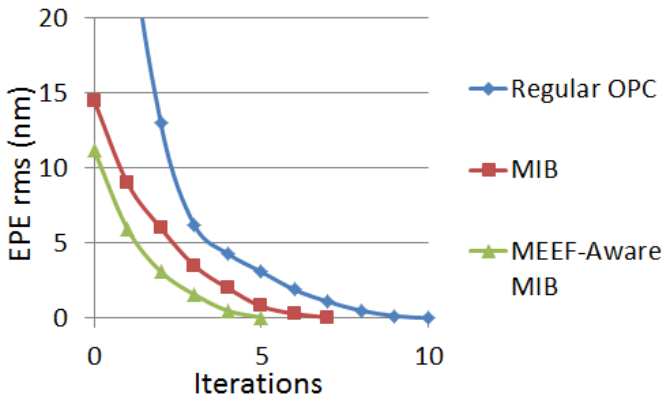


Figure 6.11: EPE conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of metal layer.

mask shape just from its target and how each fragment did get its own model-based bias that would make it very close to the final mask. Also, in this same figure, we show how the contour simulation of the MIB output closely matches the final target.

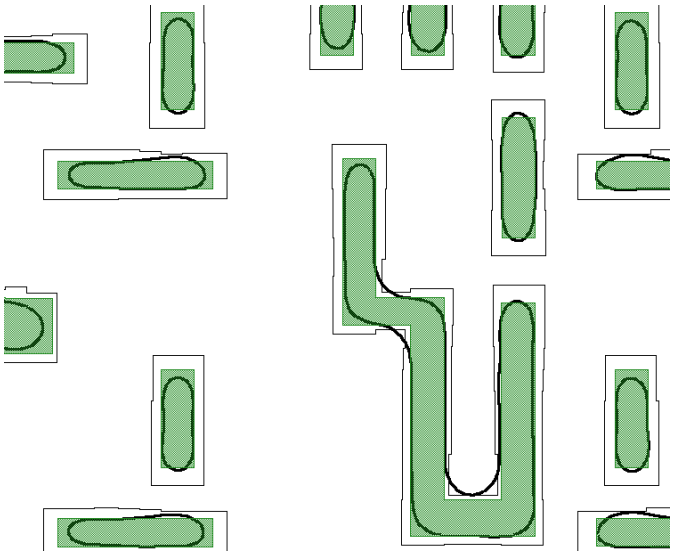


Figure 6.12: EPE conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of metal layer.

6.4.2 Via Layer Testing

In our testing for the Via level, we also used industrial 10nm OPC models to assess the improvement to the OPC convergence. Due to the outstanding difference in design style between the Metal layer design and Via layers, we created a totally different calibration and verification test pattern that match the via layer design style. In addition, we included some design patterns from critical designs such as the SRAMs and random logic. The generated MEEF-Aware MIB models had a fitting rms value of 2.734nm, which is an indication of the success of our proposed model parameters to predict most of the OPC movement information. Similar to the metal testing, we compare the convergence behavior of OPC with and without MIB. Figure (6.13), shows the OPC delta from final position and its evolution during OPC. The OPC convergence is still following a similar trend to what we showed earlier in the metal testing, where approximately 50% of the iterations can be spared when we use MIB.

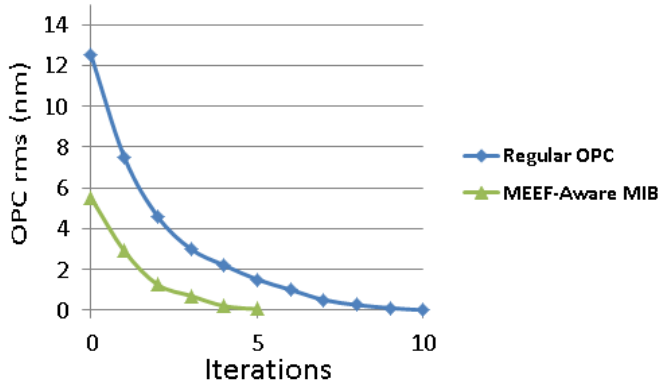


Figure 6.13: OPC mask conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of via layer.

The EPE convergence results are shown in figure (6.14), where we show how MEEF-Aware MIB is assisting OPC to reach a minimum EPE more quickly.

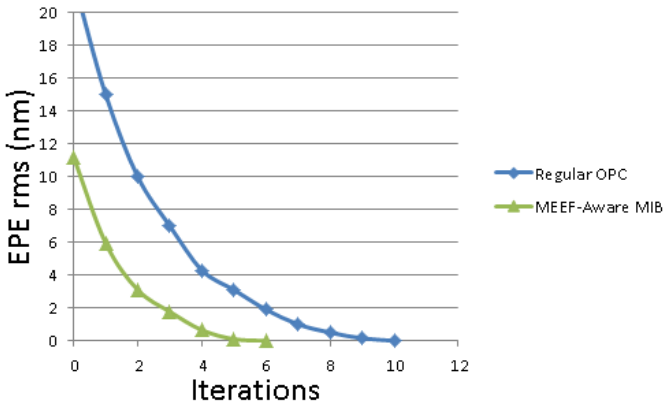


Figure 6.14: EPE conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of via layer.

Finally, In figure (6.15), we show an example of the MEEF-Aware MIB output and its output, where we demonstrate the capability of MIB to predict the final mask shape just from its target for via levels and how each fragment did get its own model-based bias that would make it very close to the final mask. The contour simulation of the MIB output show how close it is from the final target.

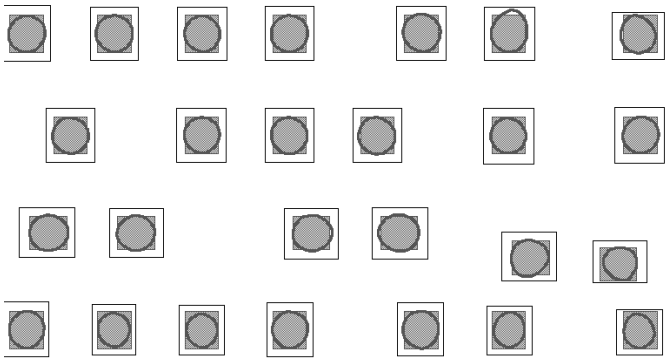


Figure 6.15: EPE conversion comparison showing how MIB and MEEF-Aware MIB are improving the OPC convergence of via layer.

6.5 Conclusions

In this chapter, we presented a novel methodology to model the OPC correction displacements using optical simulation-based compact bias models. These Model-based Initial Bias models are capable of accurately predicting the mask perturbation from the resist target shape. Using these models as a pre-step to OPC noticeably improves the OPC initial condition and accordingly improves the convergence time. We tested our methodology on 10nm layers and our testing proves that at least a 50% of the OPC runtime can be saved if MIB is used. We recognize this as a proof of potential of the concept and will focus on exploring more possible improvements techniques to our proposed MIB models to achieve more runtime reduction. This is an important achievement in the MBRT area, where we can now achieve a more accurate assessment of the lithography hot-spots based after using MIB to compute a more accurate mask quickly.

Chapter 7

Distributed Model-Based Retargeting

7.1 Introduction

In the previous chapters, we proposed (LAYER) as a methodology to achieve robust MBR to fix lithography hot spots. LAYER is a pre-OPC operation that captures and fixes lithography hot-spots even before going through the very computationally expensive OPC and PW lithography simulations. It is based on a single simulation of the OPC target (as an approximate version of the final mask), where the potential lithography hot-spots can be identified by capturing poor imaging that results from the destructive interference occurring whenever the design is not friendly with the illumination source. This was successfully demonstrated in 20nm node metal layer.

In the 10nm technology node, we are facing even more challenges to the patterning quality in DUV, where the need for more robust computational patterning techniques are needed to improve the final yield of the technology. It is not sufficient anymore to rely on simplified PWOPC or MBR techniques to automatically fix lithography hot spots. More sophisticated methodologies are required to achieve that. Another challenge that the technology is facing in 10nm is that the required accuracy to capture a lithography hot-spot has increased. To solve this we used the Model-based Initial Bias (MIB) methodology that uses a single optical simulation to estimate final mask shapes with a good accuracy.

In this chapter, we introduce the concept of Distributed-LAYER (D-LAYER) and the details of the algorithm and its building blocks and how it is integrated with MIB in the full

solution. Finally, we present our testing results of this new methodology on a 10nm metal layer using industrial OPC and lithography simulation models, where we demonstrate the added value of this algorithm and its potential to improve the patterning quality in optical lithography.

7.2 D-LAYER: Distributed Litho- And Yield-Enhancing Retargeting

Distributed LAYER is a more generalized methodology to improve the patterning by developing more robust model-based retargeting and accordingly improve the final yield. We still focus on increasing the quality of the retargeting solution while maintaining the speed advantage that MBR has over regular PWOPC and LFD solutions [136].

Figure 7.1 shows a simplified block diagram of the proposed D-LAYER flow, where first we use MIB to compute a more accurate estimation of the final OPC mask. Then this generated approximate mask is used in the optical simulation in LAYER hot spot analysis, where we use a pre-calibrated model to capture hot-spots severity using a single optical simulation. Then the hot-spot information from LAYER is passed to our proposed D-LAYER to achieve a generalized distributed MBR. All these steps are implemented using the c-based Litho API tool provided by Mentor Graphics Calibre. We used this tool to create all the newly proposed functionalities and methodologies.

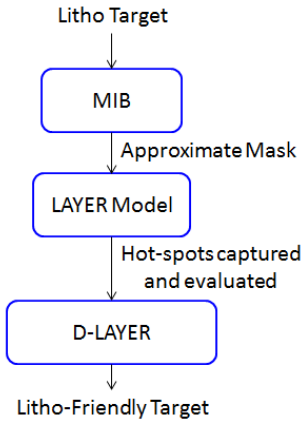


Figure 7.1: An illustration of the overall MBR flow after integrating MIB and LAYER models.

7.2.1 Problem Definition

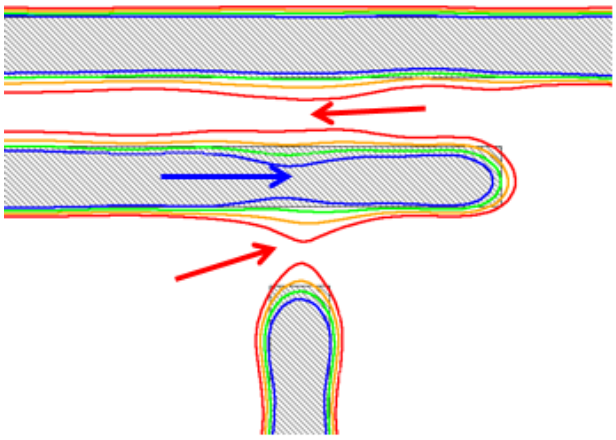
In this section, we first explain how we map the hot spot problem in a physical domain to enable the distributed retargeting and illustrating how this could improve the overall patterning quality. In the same section we formulate the equations needed to model this system. Then we discuss the necessity for the restoring forces to prevent the retargeting from deviating too much from the design intention and prevent any poor connectivity problems to the metal level. Finally, we discuss the whole D-LAYER flow that is proposed and tested in this paper.

Figure 7.2(a) shows an actual example from a 10nm metal level, where LAYER hot spot analysis detects a weak design. In this design, three adjacent weak points (with different severity magnitudes) are detected based on their imaging quality. Such a pattern is very sensitive to process variation and the slightest change in the mask to fix one problem forces the printability issue into another location as seen in figures 7.2(b), 7.3(a) and 7.3(b). Proper retargeting is necessary to solve this problem. Regular MBR and PWOPC solutions don't offer the means to fix this issue due to two main challenges. The first is that fragments during PWOPC and MBR are selfish, where in most solvers there is no way to share information about their objective functions without expensive calculation methods [137], [138]. The second challenge is that complex retargeting, where designs far from the hot-spots move around to contribute to a neighboring hot-spot is not established. Based on these limitations it is clear that any retargeting that doesn't include the collective sensation of the surrounding hot-spots is going to be severely limited [139].

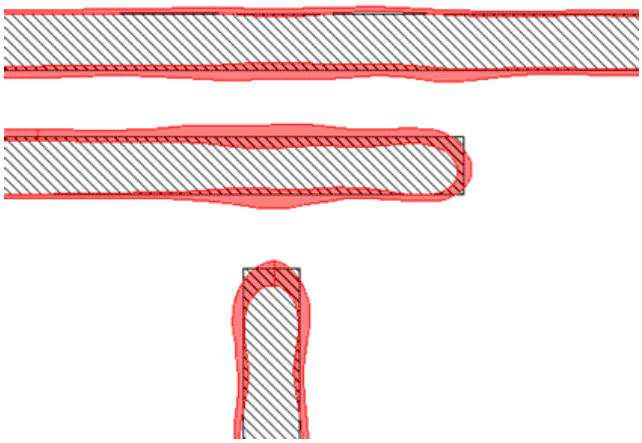
7.2.2 D-LAYER: Retargeting Forces

A more appropriate solution is to do simultaneous retargeting, where the designs would start to shift around to give more space to allow the proper patterning of the design and fixing the design hot-spots. In this paper, we present a novel distributed retargeting methodology to address these challenges and provide a generalized MBR methodology that is capable of doing more efficient retargeting in order to solve complex hot-spots situations in optical lithography.

We first represent the hot-spot as outward flux that pushes the design away from the hot spots and generating what we model as a retargeting force per unit fragment length. The magnitude of the Hot Spot Flux (HSF) is proportional to the severity of the hot-spot calculated using the LAYER model as shown in figure 7.4.

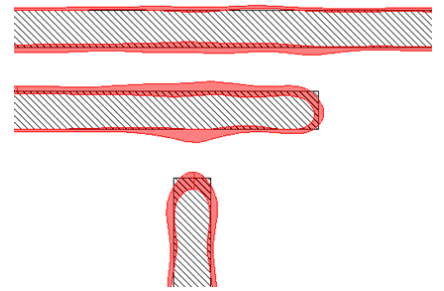


(a) Multiple potential hot-spots captured by the LAYER model indicating a weak design.

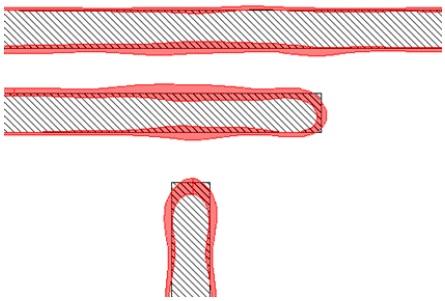


(b) PVBands post PWOPC showing a general agreement of the hot-spots.

Figure 7.2: Potential hot spots captured by the LAYER model and verified using PW simulations after PWOPC.



(a) variation 1.



(b) variation 2.

Figure 7.3: PW simulations for PWOPC recipe variations and it is clear that no solution can ultimately fix all hot spots, but rather making one of them worse on the expense of the others.

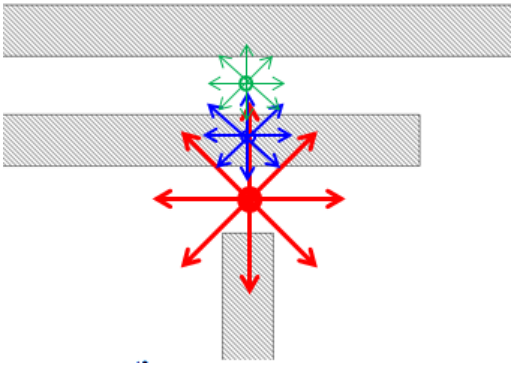


Figure 7.4: Retargeting force flux representation of lithography hot-spots.

This constant flux generates a constant force per unit length that is inversely proportional to the radial distance between the hot spot and the location on the fragment as per the inverse-square law in a two-dimensional plane (design plan) as represented in equation 7.3.

$$\nabla . \overrightarrow{P} = HSF \tag{7.1}$$

where P is the retargeting force per unit length affecting the surrounding fragments. and HSF is the hot-spot flux as calculated by the LAYER model. Then applying a two-dimensional divergence theorem we can show the inverse proportionality to the radial distance from the design hot-spot when we approximate it as a point source.

$$\oint_L \overrightarrow{P} . \overrightarrow{dL} = HSF \tag{7.2}$$

And accordingly, the retargeting pressure P can be calculated as follows

$$\overrightarrow{P} = \frac{HSF}{2\pi r}\hat{r} \tag{7.3}$$

Then the retargeting force affecting any fragment due to neighboring hot-spots can be calculated by integrating the force per unit length over the fragments length as shown in figure 7.5. Where the fragments are represented by match-sticks that are allowed to move only in one direction perpendicular to the fragments with no rotation allowed. The Final retargeting force R can be calculated as shown in equation 7.4.

$$\overrightarrow{R} = \int_{y_1}^{y_2} \overrightarrow{P} . \overrightarrow{dn}\hat{n} = \int_{y_1}^{y_2} \frac{HSF}{2\pi r} cos(\theta) dy \hat{n} \tag{7.4}$$

where dn is the fragment length differential directed normal to the fragment, r and θ are the distance and angle between the hot-spot center and the differential fragment location. Solving this equation we get the final retargeting force per fragment as

$$\overrightarrow{R} = \frac{HSF}{2\pi}[arctan(\frac{y_2}{x_o}) - arctan(\frac{y_1}{x_o})]\hat{n} \tag{7.5}$$

This is an intuitive result arising from the fact that the net force acting on a fragment is equal to the amount of the retargeting flux passing through it.

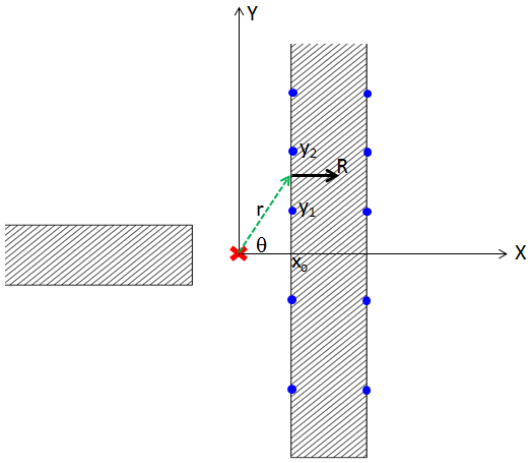


Figure 7.5: An illustrative diagram explaining the retargeting force R originating from the hot-spot center and forcing the fragment shift away from it.

Calculating the Retargeting force per fragment for all fragments surrounding the hot spot is the first step in achieving the distributed retargeting effect that we are looking for. As a result the final retargeting is what we call distributed retargeting as shown in figure 7.6, where all the neighboring fragments to the hot spots contribute to solving the problem. This results in a better overall quality as shown in figure 7.7 compared to the single fragment retargeting (as highlighted in [139]).

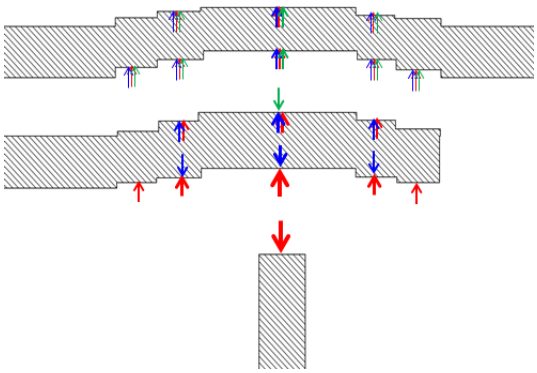
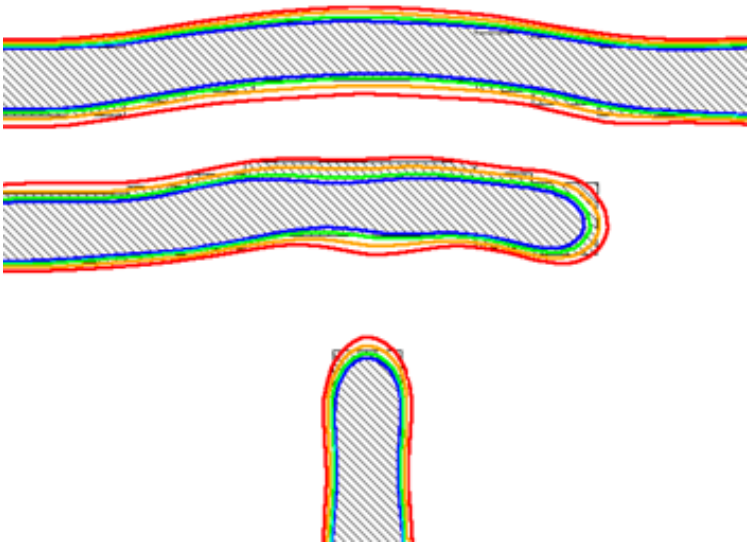
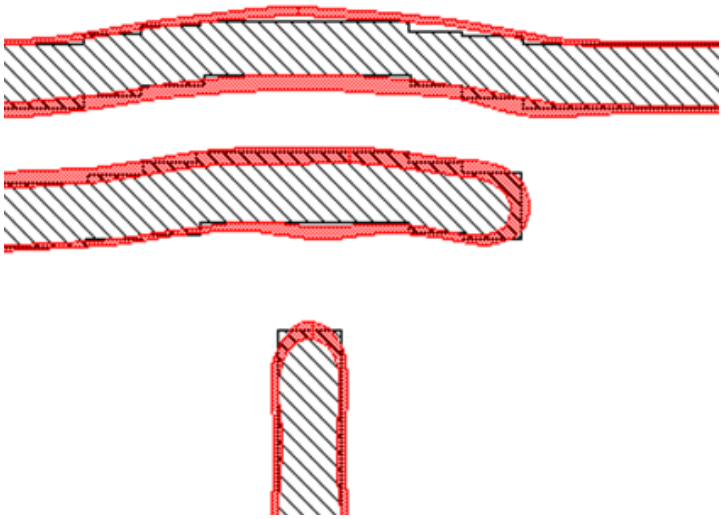


Figure 7.6: Distributed retargeting output from D-LAYER methodology.



(a) LAYER model output post-D-LAYER showing an improvement in the overall patterning quality after the distributed retargeting of the hot-spot.



(b) Process Window simulations (PVBands) after OPC of the D-LAYER retargeting.

Figure 7.7: D-LAYER successfully applying the distributed retargeting and fixing the lithography hot-spots.

7.2.3 D-LAYER: Restraining Forces

In the retargeting operations it is also very important to address the limitations that the retargeting needs to respect. Without these limitations retargeting could be doing more harm than good. In this subsection, we'll be discussing the main restraining forces we are considering in our methodology and how we model them in our proposed distributed MBR methodology.

7.2.3.1 Retargetability

By retargetability we refer to the tolerance of the fragments to move around under retargeting forces. For this we assign a retargetability coefficient, where each fragment has a retargetability coefficient for inward movement and another for outward movement. These restraining forces can be observed as if the fragments are held by springs to their original location and if a retargeting force is applied over them, then the springs will try to resist this movement according to their patterning strength. The retargetability coefficient is calculated from the LAYER model, where a fragment that has very good width patterning quality can be allowed to move inside more easily than those who already suffer from weak width patterning and vice versa for the outer movement correlation to the space patterning quality. Accordingly, the final movement d_i of the i^{th} fragment is equal to

$$d_i = \begin{cases} \frac{R_{total}}{k_{ini}} & R_{total} < 0 \\ \frac{R_{total}}{k_{outi}} & R_{total} > 0 \end{cases} \tag{7.6}$$

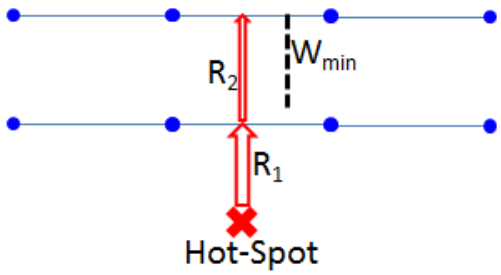
where k_{ini} and k_{outi} are the inward and outward retargetability coefficients respectively, and R_{total} is defined as the vectorial sum of the retargeting forces from all N neighboring hot-spots

$$\overrightarrow{R}_{total} = \sum_{j=1}^{N} \overrightarrow{R}_j \tag{7.7}$$

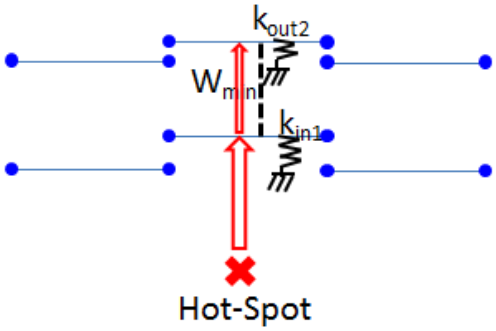
7.2.3.2 Minimum CD constraints

Even using the retargetability restoring forces, it is still very important to assign a minimum CD (width and space) that the fragments have to respect, as violating such minimum

CDs (even if the originally calculated patterning strength seems good) could cause the patterning quality to degrade so quickly. This is forced upon the fragments movement and acts as the means to transfer retargeting forces (if too big) through opposite fragments, where the opposite fragment will move to allow more space for the original fragment retargeting. At the same time, it would contribute to resisting the aggressive retargeting forces by forming a parallel spring system as shown in figure 7.8.



(a) Pre-retargeting.



(b) Post-retargeting, where the minimum CD is respected in addition to triggering a parallel spring configuration to resist the retargeting forces.

Figure 7.8: Applying minimum CD constrains together with retargetability as restraining forces to prevent from over-retargeting.

Figure 7.8 illustrates how the minimum width is enforced, where the retargeting forces R_1 and R_2 are applied on fragments 1 and 2 respectively, if the difference between both forces (combined with their respective retargetability coefficients) results in a tendency to have a width that is smaller than W_{min}, then the width is maintained at W_{min}, while the second fragment starts get extra retargeting while contributing to resisting the over-

retargeting pressure on the first fragment and the two fragments move together reaching their steady state when their combined resistance equals that of the the retargeting forces.

7.2.3.3 Multi-Layer: connectivity considerations

One of the complexities of the metal level retargeting is how to ensure the multi-layer connectivity. The awareness of where the contacts/vias above and below to the metal layer receiving the retargeting is very important otherwise new connectivity problems (circuit opens) could result as shown in figure 7.9. Accordingly, we adapted a via-aware retargeting methodology that is capable of handling such situations.

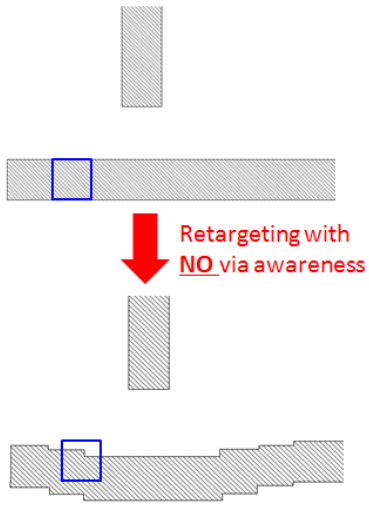


Figure 7.9: Retargeting needs to be Multi-layer aware otherwise poor connectivity situations could arise due to via layer exposure.

In this via-awareness algorithm, we first compute the safe range for via shift that the via can move around without loosing connectivity to the "upper/lower" metal layer. This is done before any MBR operations. It is just done to have a pre-judgment on which directions vias are allowed to shift towards and by how much. For example, in figure 7.10 the via is allowed to move up but not down due to the extension of the upper metal to the upper direction.

On the other had, via1 in figure 7.11 is not allowed to shift around because of a minimum via-to-via space constraint (to via2). We are keeping the level of complexity of the via awareness to the first degree in which we consider only the via and its closest via interaction

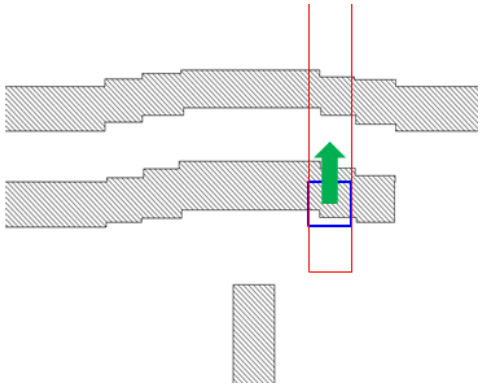


Figure 7.10: Metal-above (red) extension affects how a via (blue) would affect the retargeting.

and we don't go further to consider if via2 can shift away to give more space for via1 to move.

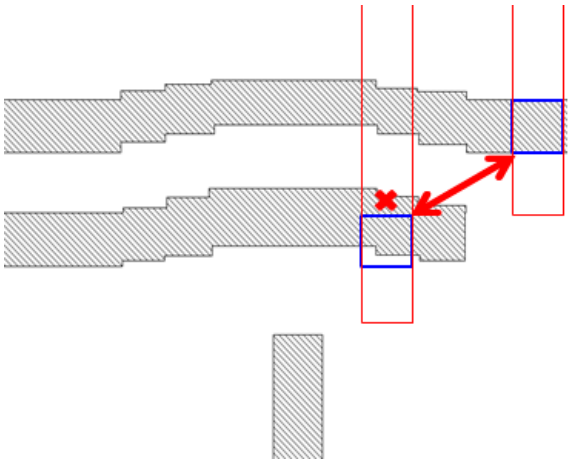


Figure 7.11: Via-to-via spacing (red arrow) affects how a via would affect the retargeting retargeting.

To feedback this information to the metal D-LAYER algorithm, we pass a reference layer that represents how much the via is allowed to move and in which directions. Then this information is passed to the via-surrounding metal fragments to increase their retargetability constant (the spring equivalent holding the fragments from moving around), such that fragments that are near a via (and are likely to expose it) will get an increased resis-

tance against retargeting as demonstrated in figure 7.12. The output from this via-aware D-LAYER is presented in figure 7.13, where the via coverage is still maintained, where the retargeting forces had more resistance at the metal fragments that are near to the via. In other cases where the via is allowed to shift freely to follow the metal layer retargeting, the via would be shifted after D-LAYER to be placed at the new intersection with the upper/lower meta as shown in figure 7.14.

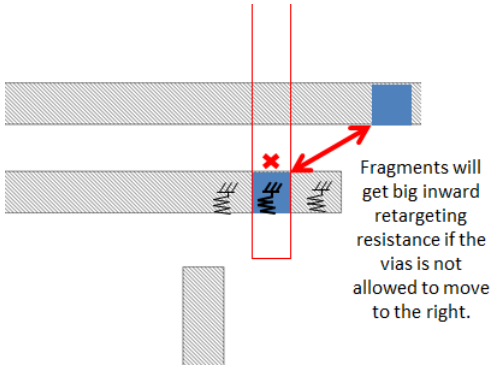


Figure 7.12: An illustration of how the via retargetability is linked to the metal fragments retargetability by increasing their spring constant.

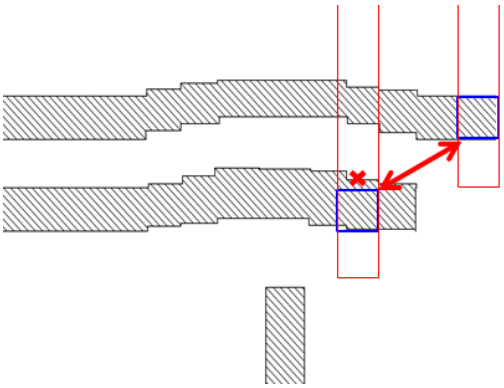


Figure 7.13: D-LAYER output after adding the via awareness, the retargeting is minimized near where the via is positioned to ensure its coverage post D-LAYER.

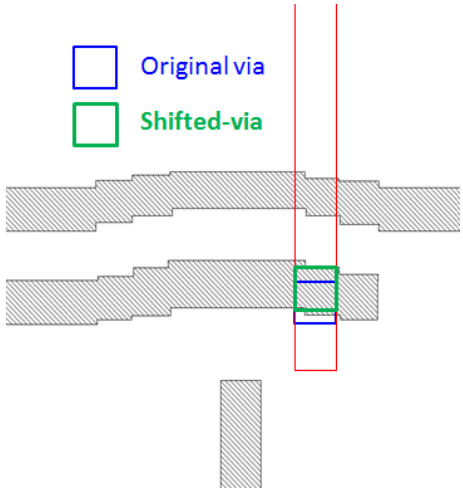


Figure 7.14: Retargeting needs to be Multi-layer aware otherwise poor connectivity situations could arise due to via layer exposure.

7.2.3.4 Multiple-Patterning considerations

Another important consideration also is the handling of the spacing between different exposures of the multiple patterning metal. Without this awareness different color bridging is a possibility as shown in figure 7.15 where the post-etch contours are showing an inter-color bridging risk. Although these are two different exposures, but a minimum different-color spacing has to be respected otherwise short circuit situations could arise when the spacing becomes too small (specially in poor overlay conditions).

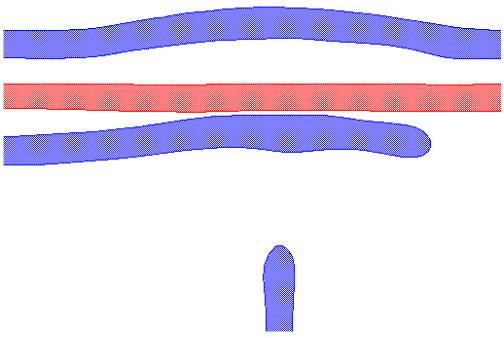


Figure 7.15: Post-etch simulations show how model-based metal retargeting can result into short circuit risks in multiple patterning if the different color distances are not respected.

To protect against this different-color patterns bridging we add the different color awareness by setting a minimum distance to the different color shapes that should never be violated. This is very similar to the minimum CD constraint solution that implemented for same color CDs, except that this time it's between different color designs as shown in figure 7.16. This is a preliminary solution to prevent the inter-color bridging. This minimum different color spacing rule results in an implicit retargeting of the different color as shown in figure 7.17. In the future, we will be working on the simultaneous retargeting of all multiple patterning layers to better handle the inter-color interactions and enable more sophisticated retargeting methodologies.

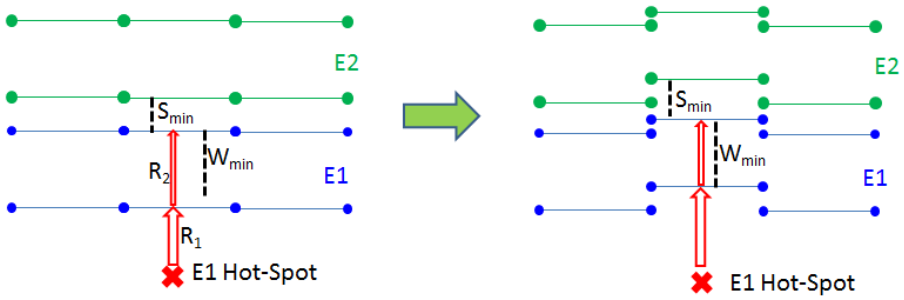


Figure 7.16: Multiple patterning-aware D-LAYER by enforcing minimum spacing between different color designs and allowing the different color design to move around to give more space to the hot-spot fix.

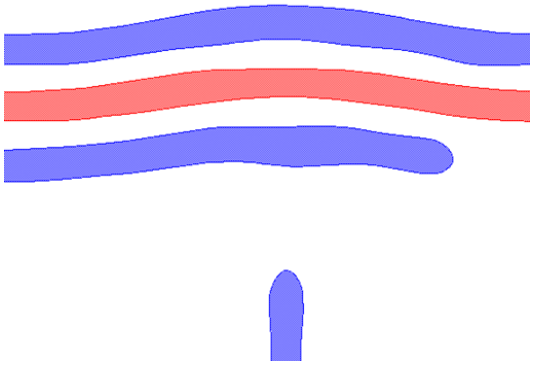


Figure 7.17: Multiple patterning-aware D-LAYER can apply implicit retargeting to different color designs to avoid inter-color bridging.

7.2.4 D-LAYER: Overall Algorithm

In this sub-section, we summarize the overall flow to implement the distributed model-based retargeting methodology as shown in figure 7.18. Where the metal layer first gets its normal etch correction and SRAF insertion to reach the final litho-target intended on wafer. Then To add the via awareness, an additional step is used to compute the via flexibility to retargeting (shifting around based on the D-LAYER needs to fix the metal hot-spots). Then we use MIB [140] to perturb the target layer into an approximate mask, which results in a more robust hot-spots capturing and analysis using the LAYER model simulations. This is followed by the application of the full distributed retargeting using D-LAYER.

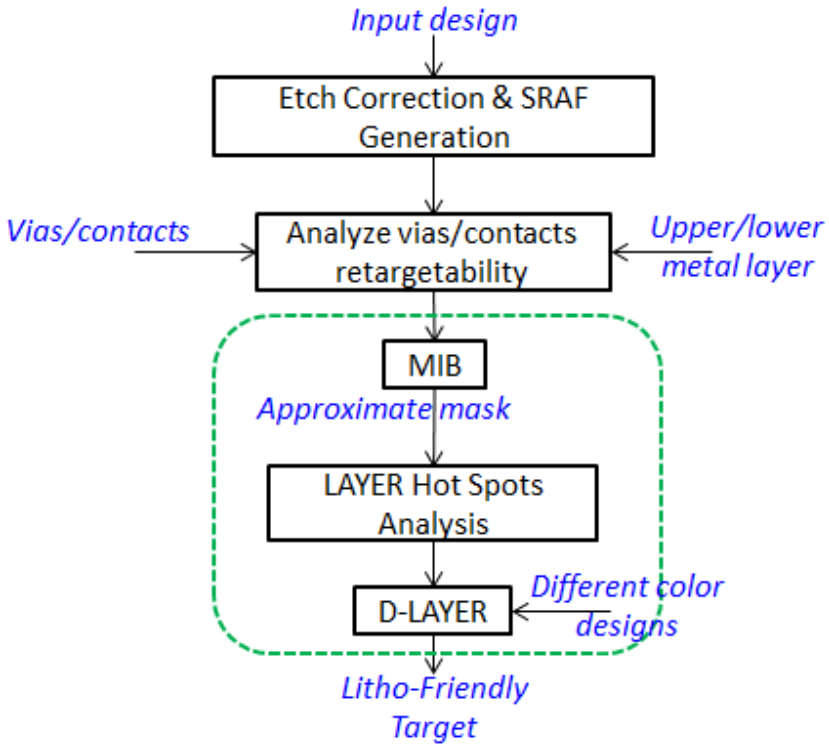


Figure 7.18: The general flow for using D-LAYER in the mask tape-out flow or in LFD.

The flow chart shown in figure 7.19 and the pseudo code in 3 summarize the whole D-LAYER algorithm and our implementation. First, the inputs (both the layer to retarget and the adjacent different color designs) are passed to the engine to get fragmented to

allow their retargeting. Also, the via retargetability information is passed along to be used during the retargeting.

Algorithm 3 D-LAYER algorithm

```
1: procedure DLAYER(LithoTarget,DiffColorTargets,Vias)
2:     ApproxMask ← MIB(LithoTarget)
3:     HotSpots(LithoTarget) ← LAYERSimulate(ApproxMask)
4:     for each Fragment Frag ∈ LithoTarget do
5:         Calculate k_in,k_out
6:         RetargetingForce(Frag) = 0
7:         for each HotSpot HS ∈ HotSpots(LithoTarget) do
8:             RetargetingForce(Frag) ← RetargetingForce(Frag) + R(HS,r)
9:         if RetargetingForce(Frag) > 0 then
10:            Displacement(Frag) ← RetargetingForce(Frag)/k_out
11:        else
12:            Displacement(Frag) ← RetargetingForce(Frag)/k_in
13:    Move all Fragments(LithoTarget) respecting W_min,S_min
```

After that, the LAYER model simulations are done on the MIB-output (the approximate mask) to capture the hot spots, recording their center locations and using the patterning strength signals to calculate the directional retargetability coefficients k_{in} and k_{out}. These retargetability coefficients also include the effect of nearby vias and their retargetability. This is then followed by nested loops of looping over all fragments and looping over all nearby hot-spots to each fragment to calculate the effective retargeting forces over all the design fragments and to calculate the required movement per fragment based on the applied retargeting force and each fragment retargetability constant.

Finally, the fragments are moved to follow the retargeting displacement per fragment, however, this movement is constrained by the minimum same-color and different-color CDs. Then another loop over all the fragments is done to find fragments that are constrained and update their k values (as they will be now operating in a parallel spring configuration) to resist back the retargeting forces and calculating the new retargeting displacements. As for the constraining fragments, the retargeting forces on them is updated to include the transferred force from the opposite fragment that they are constraining. After that, the design fragments are moved and then followed by another adjustment on the different color and vias retargeting to follow the distributed retargeting that was applied.

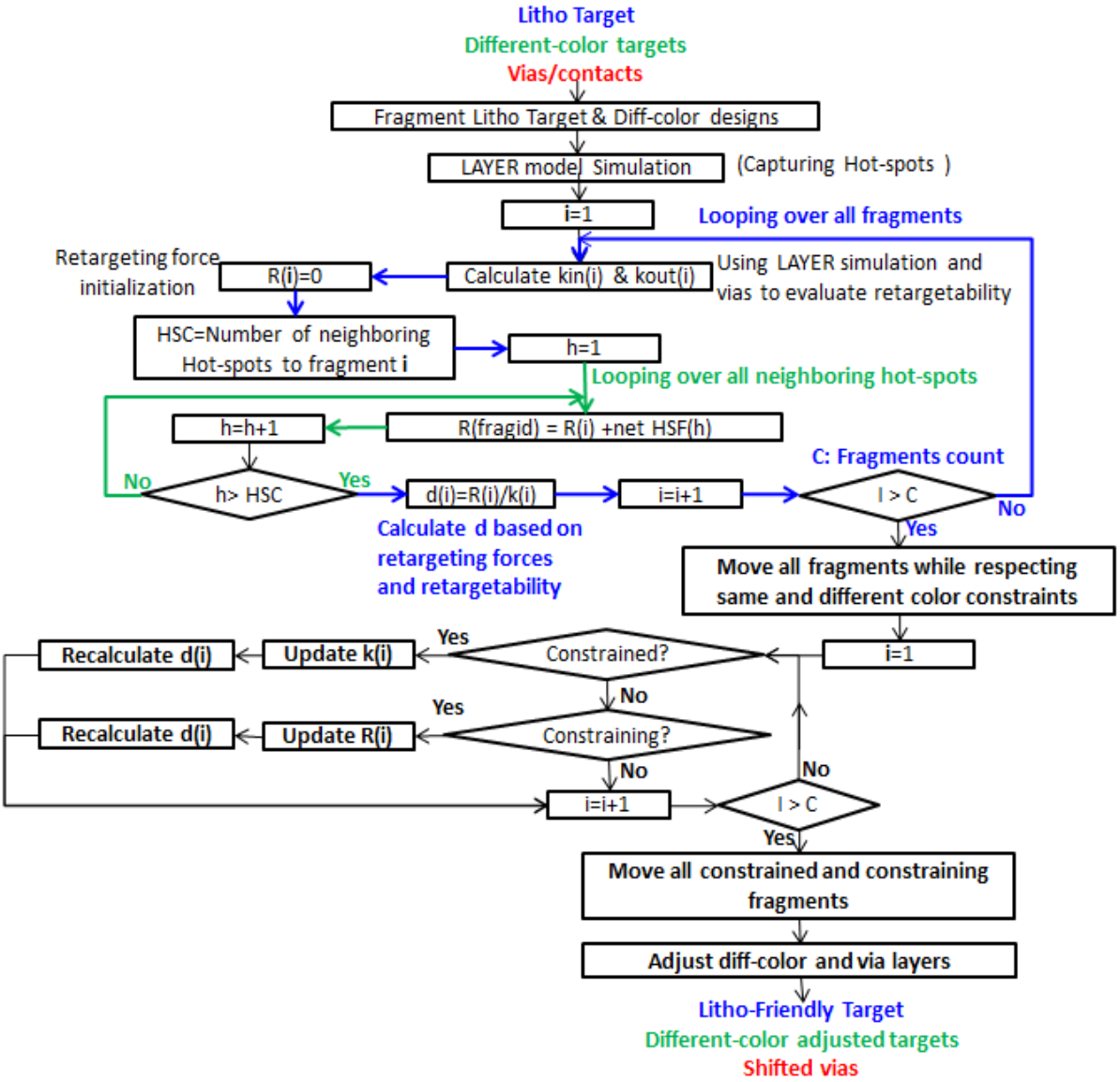


Figure 7.19: The flow chart explaining the details of the D-LAYER implementation.

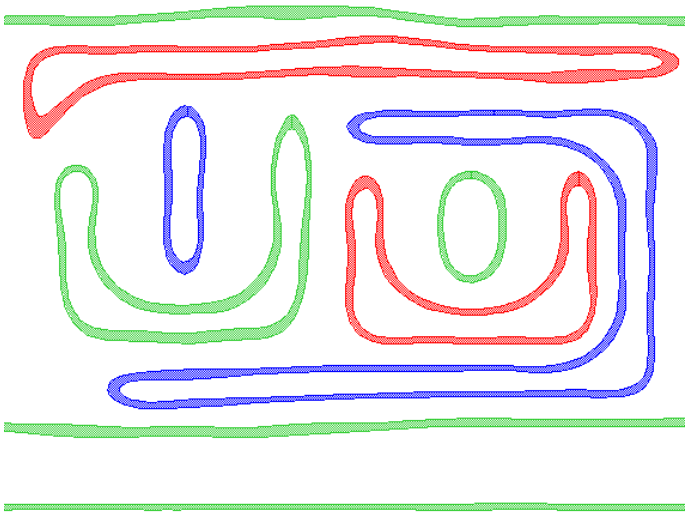
7.3 Testing Results

We implemented the D-LAYER methodology by using the c-based Litho API tool provided by Mentor Graphics Calibre, where all the functionalities mentioned in this paper were coded to provide distributed model-based retargeting. We test it on a 10nm metal layer and compare it to a regular PWOPC solution. We performed our testing on a 0.5 mm x0.5 mm sized design that constitutes a wide variety of designs and a significant contribution of random logic blocks. In the comparison, we focus on both the same-color and different-color patterning quality as well as the speed. The OPC recipe used with D-LAYER is a nominal OPC recipe, in which the goal of OPC is to converge on target with a minimum EPE. This is mainly because the D-LAYER output is more patterning friendly and a much less aggressive OPC solution can be used to do the mask correction. This proves the validity of our methodology, where with a minimal amount of computation, a litho-friendly target can be generated instead of wasting too much computation iteratively during PWOPC.

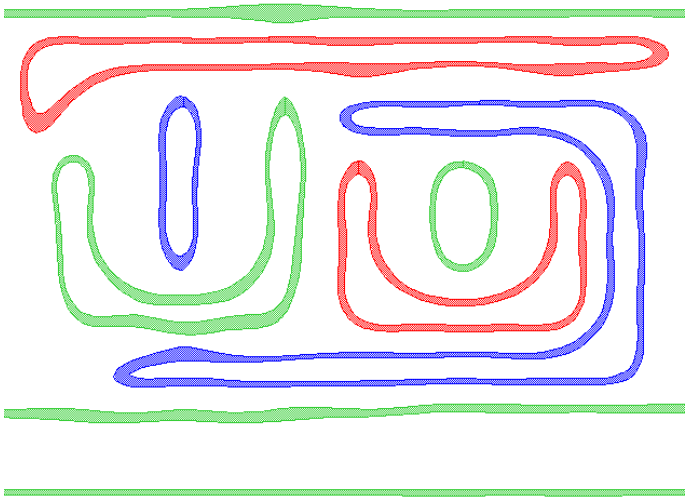
Figures 7.20 and 7.21 show snapshots of the post-etch simulations different colors for this LELELE multiple patterning metal, where the different color-aware D-LAYER shows the capability of improving both the same-color patterning without creating any different-color issues.

A more quantitative comparison is shown in figure 7.22, where the through-PW width and space distribution is plotted for both solutions. D-LAYER is showing a better overall distribution specially at the tail, where the yield-limiting failures are located. In this plot it is clear that the D-LAYER algorithm managed to eliminate the worst tens instances of issues and fixing them (by shifting their through PW CDs towards the larger CDs). Achieving this simultaneously for both width and space in the 10nm technology node is not a trivial thing due to the complexity of the patterns interaction as discussed earlier. This improvement has a direct impact on the yield, where each of these hot-spots contributed to the yield (or reliability) degradation of the final product.

The computation power assessment is also a fundamental metric that needs to be analyzed. Figure 7.23 shows the normalized CPU-runtime product comparing both the D-LAYER-based solution vs the reference PWOPC solution. It is clear that although there is an overhead for the D-LAYER solution, it saves that overhead and more by allowing the OPC to switch to nominal OPC, which saves an overall runtime of almost 12%.

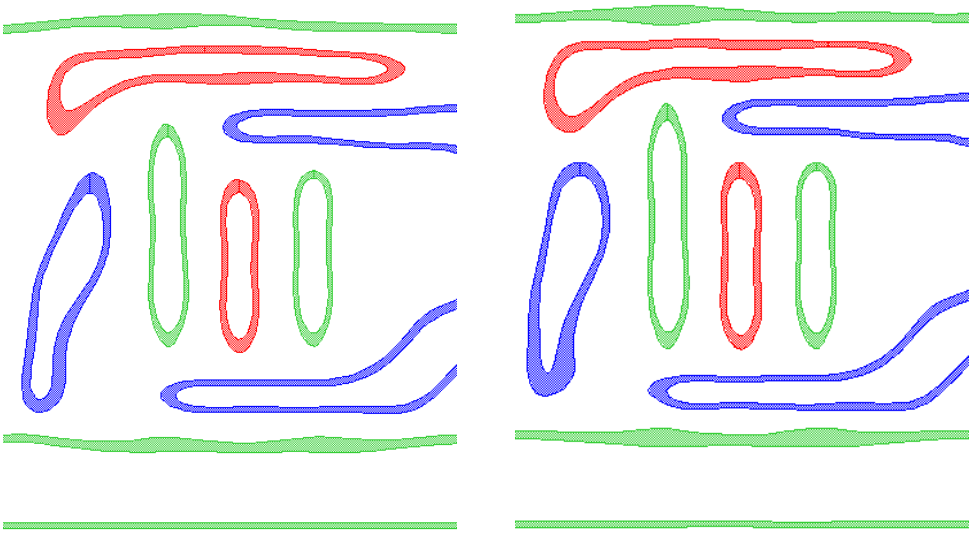


(a) D-LAYER + Nominal OPC.



(b) PWOPC.

Figure 7.20: Comparing the PVBands of the D-LAYER solution to the reference PWOPC solution as expected on wafer for the multiple patterning solution.



(a) D-LAYER + Nominal OPC. (b) PWOPC.

Figure 7.21: Comparing the post-etch PVBands of the D-LAYER solution to the reference PWOPC solution, D-LAYER shows better overall results.

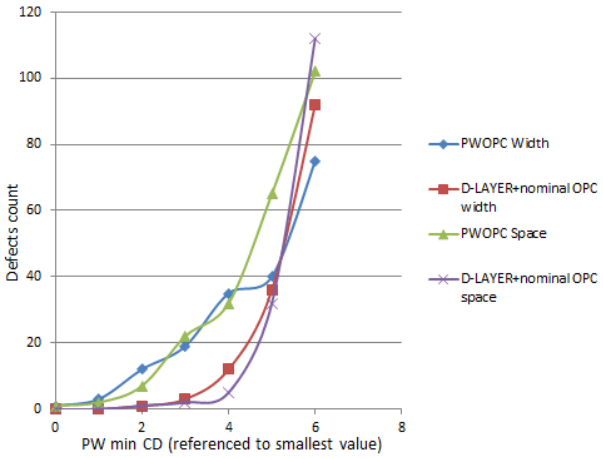


Figure 7.22: Through PW width and space CD distribution, comparing the tail distribution for both D-LAYER and the reference PWOPC.

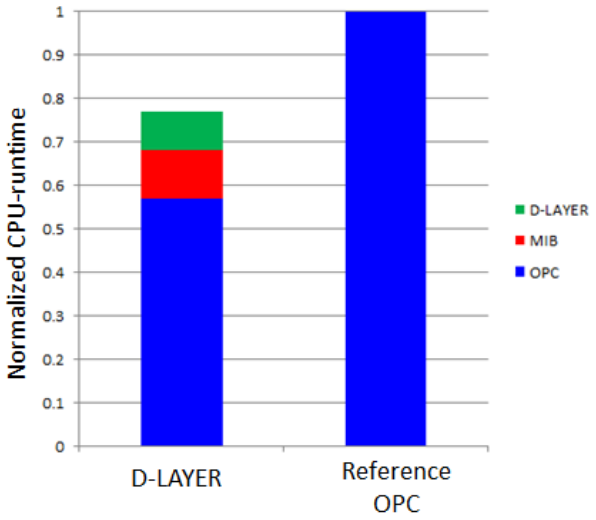


Figure 7.23: Runtime comparison between D-LAYER flow and the reference PWOPC.

7.4 D-LAYER: An LFD perspective

One of the main motivation of this work is to address and intercept the expected weak spots in DFM that are expected to start causing problems in the advanced technology nodes. Some of these issues are directly related to the lithography-related functional yield where PWOPC limitations are starting to affect both the quality and the computational resources which is expected to cause serious yield issues. In this thesis, we developed a computationally efficient MBRT that can be used to improve the yield by automatically fixing the designs without the need to go into the full flow of MBOPC and through PW verification (as proposed by regular LFD solutions), which makes it by far more computationally efficient.

Current LFD flows are very computationally expensive. In standard LFD flows, the design has to go through the full OPC then through PW simulations in order to capture the non-lithography friendly designs as shown in figure (7.24). This is a very computationally expensive flow. However, with the introduction of D-LAYER as a very efficient methodology for capturing and fixing lithography hot-spots, the LFD flow can be orders of magnitude faster as well as providing the final design fix to the designers as shown figure (7.25). This fix can be accepted by the designer or even used as a guideline for further enhancement and achieving a more robust design.

It is also in our goals use this full simultaneous multi-layer distributed retargeting into

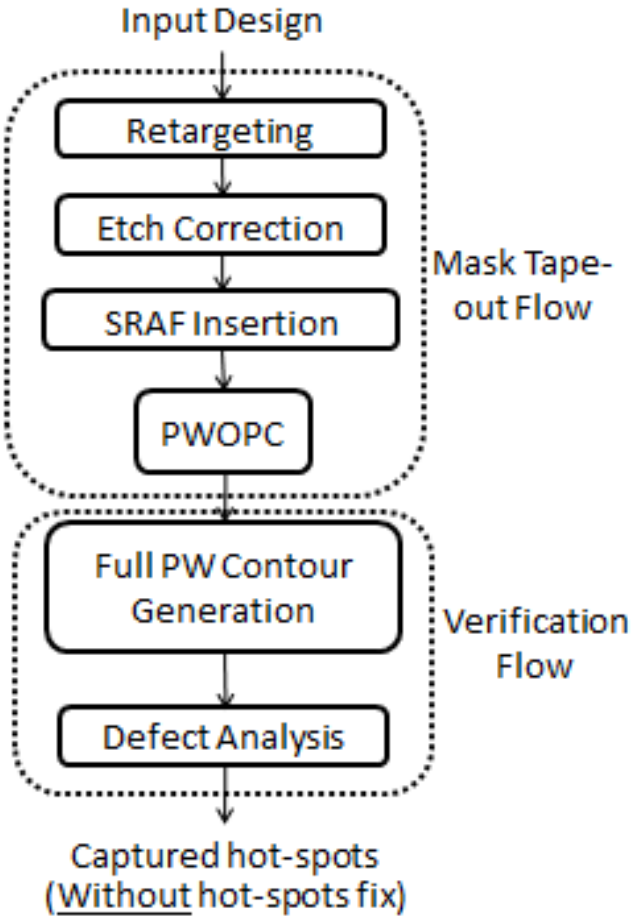


Figure 7.24: Flow diagram of a standard LFD Flow.

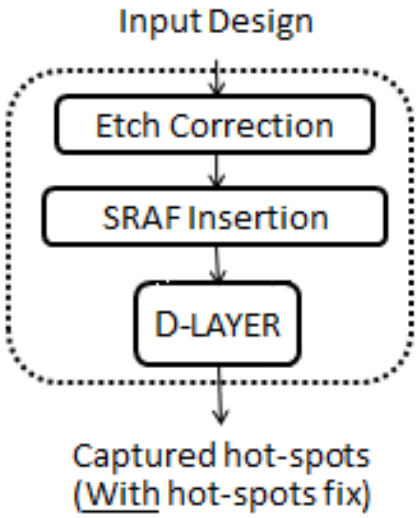


Figure 7.25: Flow diagram of a much more computationally efficient and useful D-LAYER-based LFD flow, where not only it is orders of magnitude faster, but also providing the design auto-retargeting that fixes the hot-spots.

the auto-correction LFD domain as shown in figure 7.26. This flow can be integrated to do connectivity-aware simultaneous retargeting of the design levels to maximize the patterning quality (and yield) simultaneously with improving the electrical connectivity.

This auto-correction LFD flow has three main advantages. First, from a speed point of view, it is much faster as in contrast to standard LFD flows ours is based on D-LAYER which doesn't require full OPC and PW simulations just to identify lithography hot-spots. Figure 7.27 shows a comparison of the normalized computation required by standard LFD compared to the D-LAYER-based LFD per design layer. D-LAYER does all this much faster and is totally decoupled from OPC. Second, it serves to guide the designers on the changes they need to do to achieve better yield, or they can simply accept the redesign done by D-LAYER based auto-correction LFD. Finally, D-LAYER-based LFD can be integrated into a fab's PDK to provide the exact final CDs to the designers to use in their final electrical verification early enough during the design phase. This is never possible with the current flows, where final CD extraction from the LFD contours is too expensive to do on a full-chip level.

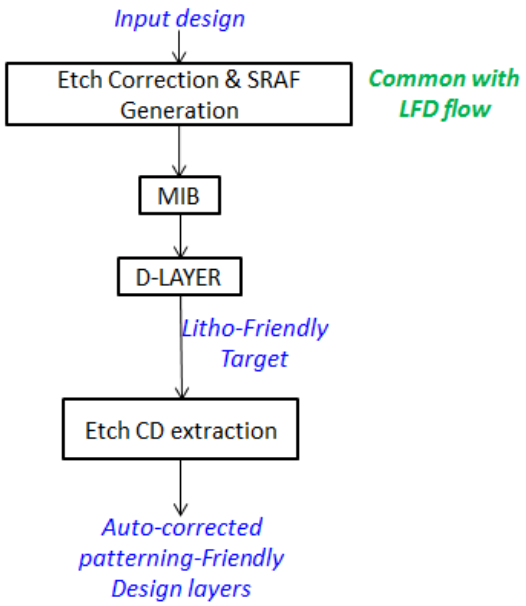


Figure 7.26: A Proposal of D-LAYER-based LFD flow.

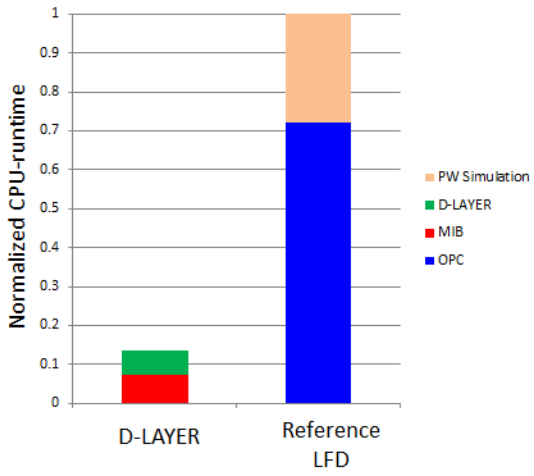


Figure 7.27: Runtime comparison between D-LAYER-based LFD flow and a reference LFD, where the standard LFD is spending too much runtime in OPC and PW simulations.

7.5 D-LAYER: Parasitic-Extraction Applications

In the design process, it is desirable to include all systematic deviations accurately using fast models. This is very important for the yield ramp-up and the final product availability and maturity. The current approach for including the design deviations is shown in figure 7.28, where all geometric-based yield enhancing re-targeting is included before the deviation in the electrical parameters is fed-back to the simulation. The problem with this approach is that 1) it does not include all the design-deviations that happens during the mask tape-out. For example it assumes perfect etch-process correction, which is a good approximation, but yet not perfect. 2) Also it does not include the PWOPC retargeting which can reach a value of few nanometers. The problem with PWOPC retargeting is that it cannot be included without running full OPC and full verification, which is very time-consuming. Figure 7.29 shows a more accurate approach to include all the design deviations in the mask-tape-out flow, but unfortunately it is very computationally intensive. It is of great benefit to develop a technique that allows the reproduction of the Lithography friendly retargeting while being much faster than PWOPC so that this design deviation can be included in the design process.

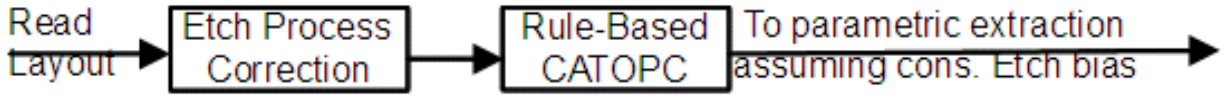


Figure 7.28: A schematic diagram explaining the current extraction flow.

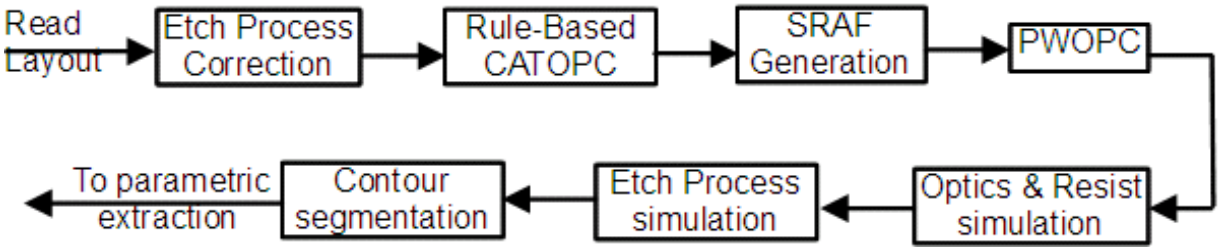


Figure 7.29: A schematic diagram of an ideal extraction flow.

D-LAYER can play a role in improving the accuracy of the parasitic-extraction. This is mainly because using D-LAYER allows the extraction tools to use the most accurate final wafer CDs compared to the current flow, which neglects the implicit secondary retargeting introduced by PWOPC. Figure (7.30) illustrates the simplification to the accurate final-design extraction and how this can be used in predicting the actual final prediction on

wafer. Current commercial etch correction and simulation tools is much faster when done on design edges rather than fine contours, which means that the final CD computation is very fast and compatible with the MBRT flow.

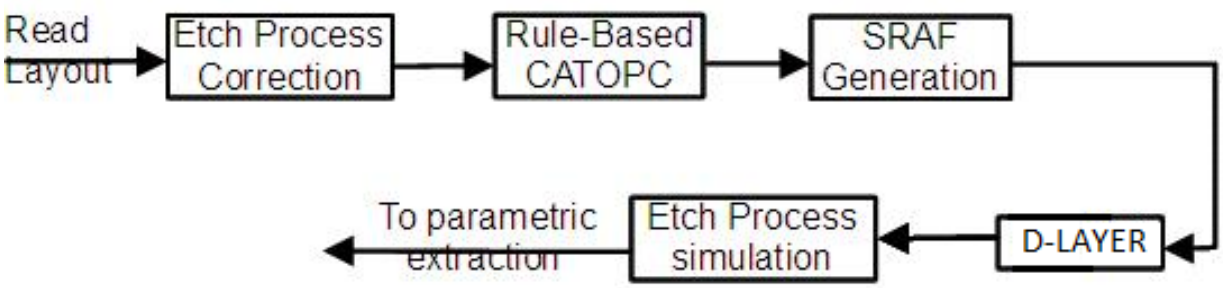


Figure 7.30: A schematic diagram explaining the proposed fast and accurate D-LAYER-Based extraction flow.

Figure (7.31) shows a simplified illustration diagram of the design flow in sub-micron technologies. The rules-based retargeting output is fed into the parasitic extraction tools to be used in a more accurate R and C estimation and accordingly more accurate timing analysis. However, with the increasing deviation due to PWOPC, the actual R and C calculation starts to deviate from the ideal calculations both on the standard cell level and the routing levels. Using LAYER to robustly fix lithography hot-spots eliminates the need for PWOPC and provides a very computationally efficient way to provide a more accurate CDs to be used in the parasitic extraction, where the model-based retargeting step (D-LAYER) would substitute the less accurate retargeting step.

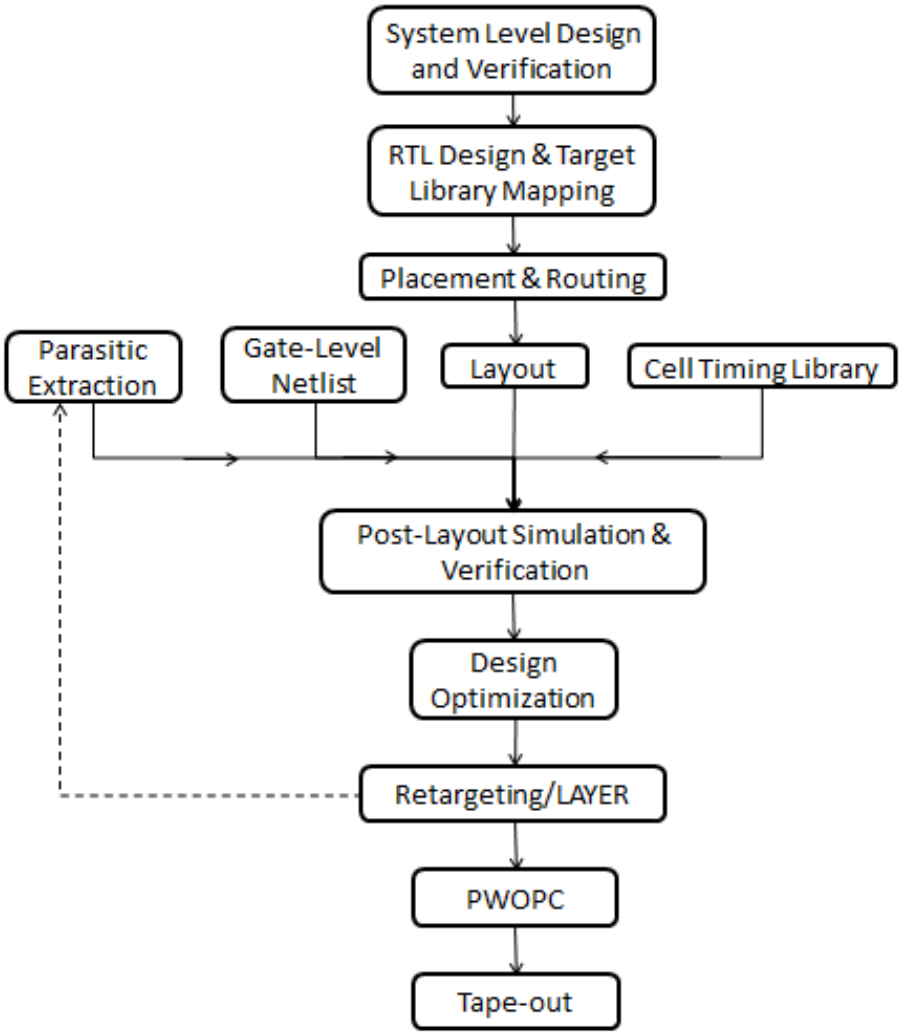


Figure 7.31: A Simplified Flow diagram of the design flow in sub-micron technologies showing where D-LAYER can be introduced to improve the electrical parameters accuracy.

7.6 Conclusion

In this chapter, we present a novel methodology for distributed model-based retargeting. We also demonstrate the advantages of distributed model-based retargeting and how it is capable of solving very complex through-PW lithography patterning problems by distributing the problem over the neighboring designs. We developed the basis of the transformation of the retargeting problem into the physical domain by introducing the concept of retargeting flux and retargeting forces. This transformation enables the distributed retargeting and makes it achievable. We also cover multi-layer-aware retargeting. This is done by feeding the adjacent layers ability to retarget to the D-LAYER solver in the form of retargetability constants, which are used by the solver to ensure the electrical connectivity of the final design. We implemented this methodology and tested it on a 10nm metal layer, where we demonstrated a better overall patterning solution for both same-color and different color hot-spots. We also demonstrate the runtime advantage, as using D-LAYER has a potential to eliminate the need for PWOPC, which would save a lot of the computation time and resources spent in it. In our testing, we demonstrated a 12% overall improvement in runtime while achieving better patterning quality using this new concept.

We also discussed the main advantages of the MBRT when integrated in the LFD and parasitic extraction flows. D-LAYER shows a very good agreement with the main requirements for integrating it efficiently into next generation LFD tools. This MBRT-based LFD is capable of improving the functional yield by helping the designer fix the weak designs (or fix it for them). It is also capable of improving the final parametric yield by providing more accurate final CDs to the designers for their parasitic extraction tools. And finally, all this is done much more quickly than the current sate-of-the-art tools, which is very important to integrating it into the technology PDKs.

Chapter 8

Conclusion and Future Work

As we proceed into more advanced technology nodes using the DUV lithography, the patterning challenges increase to an extent that seriously affects the yield, the technology maturity and the technology time-to-market. Many development and evolutions happened to the computational lithography field in the past two decades to allow the patterning of the deep sub-wavelength designs. Also, many new fields evolved to achieve this among which are RET, Multiple-Patterning, PWOPC and LFD (as a part of the wider field of DFM). In this chapter, we summarize the research work in this thesis and the future research directions.

8.1 Summary of Contribution

In this thesis, we introduced the new concept of Model-Based ReTargeting (MBRT) and specified its goals and advantages. The main goal of the MBRT is to improve the yield in advanced technology nodes by improving the design patterning quality for a given technology. In chapter 4, we presented the concept and proved its feasibility. Then in chapter 5, we developed Litho And Yield Enhancing Retargeting (LAYER) as a more general framework to do MBRT. In LAYER, we developed the framework and the algorithms to calibrate a compact model that uses the optical simulations to capture lithography hot spots and their severity. Moreover, LAYER model also computes the necessary displacement for the design fragments to move and fix the lithography hot-spots they are suffering from.

To improve the MBRT accuracy, we also developed a new methodology to very quickly compute an approximate version of the final mask. Mode-based Initial Bias (MIB) is

a technique to mode the design fragments displacement from the original design to the final mask position. This model is then used to perturb the OPC target into a better approximation of the final mask. This has two main fields of application. First, in the field of OPC, where it saves many of the OPC iterations by allowing the OPC to start from a better initial condition, which has its benefit in the mask-tape-out flow. Second, using the approximate mask computed by MIB gives the LAYER better accuracy in predicting the lithography hot-spots.

Also, in chapter 7, we present a generalized MBRT methodology for achieving a better yield by introducing the concept of distributed retargeting. In distributed retargeting, the fragments suffering from lithography hot-spot move all together with their neighboring fragments to fix the hot-spot in a form of distributed retargeting. To achieve this, we proposed a new way to look at the retargeting problem, where we mapped it into the physical domain and introduced the concept of the retargeting flux and retargeting forces. Moreover, to achieve solution robustness, we introduced the concept of retargetability, where the fragments would follow or resist the retargeting forces based on their patterning strength and how their retargeting is going to affect the inter-layer electrical connectivity (mainly via-awareness). The formulation of this methodology and the retargeting and resistance forces was first introduced and derived as a part of this research. This methodology was developed and tested on 10nm metal layers using industrial models and showed an improvement in the overall patterning.

8.2 Limitations and Recommendations

The design trends/recommendations in advanced technology nodes is to use regular (gridded) design style for the device-forming layers, while using additional cut/trim layers to create the final shape. This is mainly to minimize the device performance variability. Due to this design-style, the lithography-patterning complexity is shifted from Front-End-Of-Line (FEOL) to Back-End-Of-Line (BEOL). In this research, we focus mainly on Back-End-Of-Line (BEOL) layers, which are mainly the interconnect levels. Interconnect levels patterning improvement is very important from a yield point of view.

Practically, the techniques and methodologies applied in this thesis are useful only for non-uniform deign styles (which usually suffer in lithography patterning). If BEOL layers design style starts to follow a uniform style, or switch to side-wall deposition patterning techniques like Self-Aligned Double Patterning (SADP) and Self-Aligned Quadruple-Patterning SAQP), then the added value of this research would mainly be on the MIB part and speeding up the OPC. The mode-based retargeting portion of this work would

be useful only (if needed) in the cut/trim layers rather than the patterning of the actual layers.

Extreme UltraViolet (EUV) lithography is showing stronger signs for intercepting the 7 nm or 5 nm nodes. EUV has a much better resolution capabilities compared to DUV (13 nm wavelength versus a 193 nm wavelength). However, as the EUV is coming so late, it itself is facing some challenges of its own, where by the 5 nm technology EUV is challenged to the extent that might require double patterning EUV to meet the technology requirements. Accordingly, the techniques developed in this thesis can be extended to address the challenges in EUV patterning. However, additional considerations might be required to capture and fix EUV-hot-spots, where including the shadowing and the flare effects is important to consider.

8.3 Future Work

The formulation of the distributed model-based retargeting in this work has the ability to extend into fully simultaneous multi-layer distributed retargeting. In this simultaneous retargeting, the full retargeting solution of all the physical design layers can be done simultaneously. In this flow, the hot spots of each layer can be recorded first and then using the retargeting and resistance forces of all design layers are handled together, while using the design connectivity restraining forces to control the extend of the retargeting of each layer, while simultaneously respecting the electrical connectivity of the design. This solution offers the maximum freedom and potentially the best overall retargeting solution as all layers are being retargeted while simultaneously aware of the hot-spots and the constraints on the adjacent design layers.

However, this fully simultaneous model-based retargeting is not trivial. This is due to the complexity of handling all the design and connectivity requirements of all design layers without causing any new issues. This approach still has the potential to improve the overall patterning quality as an evolution of what was proposed in this thesis (where we mainly anchored the contact/via layers to allow the decoupling of different design layers)

APPENDICES

Appendix A

Publications from this work

1. A. Y. Hamouda, J. Word, M. Anis, and K. S. Karim, "Aerial Image Retargeting (AIR): Achieving Litho-Friendly Designs" vol. 7974, p. 797411, SPIE, 2011.

2. A. Y. Hamouda, M. Anis, and K. S. Karim, "Air (Aerial Image Retargeting): A Novel Technique For in-Fab Automatic Model-Based Retargeting-for-Yield" in DATE, pp. 1603-1608, 2012.

3. Ayman Hamouda ; Mohab Anis ; Karim S. Karim, "Using segmented models for initial mask perturbation and OPC speedup" . Proc. SPIE 8880, Photomask Technology 2013.

4. Ayman Hamouda ; Mohab Anis, "Methods for Modifying and Integrated Circuit Layout Design". U.S. Patent 8,997,027, March 31, 2015.

5. A. Hamouda, M. Anis, and K. S. Karim, LAYER (Lithography And Yield-Enhancing Retargeting) : A Novel Automatic Model-Based Retargeting and Auto-Correction LFD Flow", IEEE TCAD Accepted with Modifications, 2015.

6. A. Hamouda, M. Anis, and K. S. Karim, "Mib (Model-based Initial Bias) : Towards a Single-Iteration Optical Proximity Correction", IEEE TCAD Submitted for publication, 2015.

7. A. Hamouda, M. Anis, and K. S. Karim, "A Novel Generalized Methodology for Distributed Model-Based Retargeting for Advanced Integrated Circuits Technology Nodes", IEEE TCAD Submitted for publication, 2015.

Appendix B

Glossary

λ	=	Wavelength
ArF	=	Argon Fluoride Lasers
BEOL	=	Back end of Line
CAA	=	Critical Area Analysis
CD	=	Critical Dimension
CDU	=	CD Uniformity
CMP	=	Chemical Mechanical Polishing
DFM	=	Design For Manufacturing
DoF	=	Depth of Focus
DP	=	Double Patterning
DUV	=	Deep Ultra Violet
EL	=	Exposure Latitude
EPE	=	Edge Placement Error
EUV	=	Extreme Ultra Violet
ILT	=	Inverse Lithography
K_1	=	Empirical Lithography resolution factor (Rayleigh factor)
LAYER	=	Litho And Yield Enhancing Retargeting
LER	=	Line Edge Roughness
LFD	=	Litho-Friendly Design
LELE	=	Litho-Etch-Litho-Etch
LELELE	=	Litho-Etch-Litho-Etch-Litho-Etch
MBRT	=	Model-Based ReTargeting
MBSRAF	=	Model-Based SRAF
MEEF	=	Mask Error Enhancement Factor

MIB	=	Model-based Initial Bias
MP	=	Multiple Patterning
NA	=	Numerical Aperture
OAI	=	Off-Axis Illumination
OPC	=	Optical Proximity Correction
PWOPC	=	Process Window OPC
RBRT	=	Rules-Based ReTargeting
RDF	=	Random Dopant Fluctuations
RDR	=	Restricted Design Rules
RET	=	Resolution Enhancement Techniques
RIE	=	Reactive Ion Etching
SMO	=	Source Mask Optimization
SRAF	=	Sub-Resolution Assist Features

References

[1] K. J. Chiang, C. C., Design for Manufacturability and Yield for Nano-Scale CMOS. Dordrecht, The Netherlands: Springer, 2006.

[2] S. Borkar, T. Karnik, and V. De, "Design and reliability challenges in nanometer technologies," in Proceedings of the 41st annual Design Automation Conference, DAC '04, (New York, NY, USA), pp. 75–75, ACM, 2004.

[3] L. Liebmann, "Dfm, the teenage years," vol. 6925, p. 692502, SPIE, 2008.

[4] N. S. B. D. Orshansky, Michael, Design for Manufacturability and Statistical Design. Dordrecht, The Netherlands: Springer, 2008.

[5] A. B. Kahng, S. Muddu, and C.-H. Park, "Auxiliary pattern-based optical proximity correction for better printability, timing, and leakage control," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 7, no. 1, p. 013002, 2008.

[6] B. P. Wong, A. Mittal, G. W. Starr, F. Zach, V. Moroz, and A. Kahng, Nano-CMOS Design for Manufacturability: Robust Circuit and Physical Design for Sub-65nm Technology Nodes. New York, NY, USA: Wiley-Interscience, 2008.

[7] J. L. Cobb, F. A. Houle, and G. M. Gallatin, "Estimated impact of shot noise in extreme-ultraviolet lithography," vol. 5037, pp. 397–405, SPIE, 2003.

[8] Y. Granik, "Correction for etch proximity: new models and applications," vol. 4346, pp. 98–112, SPIE, 2001.

[9] I. W. Rangelow, "Critical tasks in high aspect ratio silicon dry etching for microelectromechanical systems," vol. 21, pp. 1550–1562, AVS, 2003.

[10] K. S. Andrew B. Kahng and R. O. Topaloglu, "Recent topics in cmp-related ic design for manufacturing," Proc. Advanced Metallization Conference, 2008.

[11] D. Frank, R. Dennard, E. Nowak, P. Solomon, Y. Taur, and H.-S. P. Wong, "Device scaling limits of si mosfets and their application dependencies," Proceedings of the IEEE, vol. 89, pp. 259 –288, mar 2001.

[12] J. Hogan, C. Progler, A. Chatila, B. Bruggeman, M. Heins, R. Pack, and V. Boksha, "Design process optimization, virtual prototyping of manufacturing, and foundry-portable dfm (invited paper)," vol. 5756, pp. 21–38, SPIE, 2005.

[13] M. Levitt, "Using yield-focused design methodologies to speed time-to-market (invited paper)," vol. 5756, pp. 13–20, SPIE, 2005.

[14] A. B. Kahng and C.-H. Park, "Auxiliary pattern for cell-based opc," vol. 6349, p. 63494S, SPIE, 2006.

[15] L. Liebmann, L. Pileggi, J. Hibbeler, V. Rovner, T. Jhaveri, and G. Northrop, "Simplify to survive: prescriptive layouts ensure profitable scaling to 32nm and beyond," vol. 7275, p. 72750A, SPIE, 2009.

[16] M. C. Simmons, J.-H. Kang, Y. Kim, J. I. Park, S. weon Paek, and K. sup Kim, "A state-of-the-art hotspot recognition system for full chip verification with lithographic simulation," vol. 7974, p. 79740M, SPIE, 2011.

[17] K. Koike, K. Nakayama, K. Ogawa, and H. Ohnuma, "Optimization of layout design and opc by using estimation of transistor properties," vol. 6283, p. 62830P, SPIE, 2006.

[18] P. Gupta, A. Kahng, Y. Kim, and D. Sylvester, "Self-compensating design for reduction of timing and leakage sensitivity to systematic pattern-dependent variation," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 26, pp. 1614 –1624, sept. 2007.

[19] P. Gupta, A. Kahng, Y. Kim, and D. Sylvester, "Self-compensating design for focus variation," in Design Automation Conference, 2005. Proceedings. 42nd, pp. 365 – 368, june 2005.

[20] E. Yang, C. H. Li, X. H. Kang, and E. Guo, "Model-based retarget for 45nm node and beyond," vol. 7274, p. 727428, SPIE, 2009.

[21] A. K.-K. Wong, "Resolution enhancement techniques in optical lithography," in SPIE Press, 2001, SPIE Press, 2001.

[22] Y. Granik, "Fast pixel-based mask optimization for inverse lithography," Journal of Microlithography, Microfabrication, and Microsystems, vol. 5, no. 4, p. 043002, 2006.

[23] J. Word, Y. Granik, M. Medvedeva, S. Rodin, L. Capodieci, Y. Deng, J. Kye, C. Tabery, K. Yoshimoto, Y. Zou, H. Diab, M. Gheith, M. Habib, and C. Zhu, "Inverse vs. traditional opc for the 22nm node," vol. 7274, p. 72743A, SPIE, 2009.

[24] B.-S. Kim, Y.-H. Kim, S.-H. Lee, S.-I. Kim, S.-R. Ha, J. Kim, and A. Tritchkov, "Pixel-based sraf implementation for 32nm lithography process," vol. 7122, p. 71220T, SPIE, 2008.

[25] D. Melville, A. E. Rosenbluth, K. Tian, K. Lai, S. Bagheri, J. Tirapu-Azpiroz, J. Meiring, S. Halle, G. McIntyre, T. Faure, D. Corliss, A. Krasnoperova, L. Zhuang, P. Strenski, A. Waechter, L. Ladanyi, F. Barahona, D. Scarpazza, J. Lee, T. Inoue, M. Sakamoto, H. Muta, A. Wagner, G. Burr, Y. Kim, E. Gallagher, M. Hibbs, A. Tritchkov, Y. Granik, M. Fakhry, K. Adam, G. Berger, M. Lam, A. Dave, and N. Cobb, "Demonstrating the benefits of source-mask optimization and enabling technologies through experiment and simulations," vol. 7640, p. 764006, SPIE, 2010.

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

[27] A. Yehia Hamouda, "Dual metric opc for complex sraf implementation," in Design and Test Workshop, 2007. IDT 2007. 2nd International, pp. 253 –256, dec. 2007.

[28] S. Banerjee, P. Elakkumanan, L. W. Liebmann, J. A. Culp, and M. Orshansky, "Electrically driven optical proximity correction," vol. 6925, p. 69251W, SPIE, 2008.

[29] G. Moore, "Progress in digital integrated electronics," in Electron Devices Meeting, 1975 International, vol. 21, pp. 11 – 13, 1975.

[30] M. Bohr, "A 30 year retrospective on dennard's mosfet scaling paper," Solid-State Circuits Newsletter, IEEE, vol. 12, pp. 11 –13, winter 2007.

[31] G. Shahidi, "Device scaling for 15 nm node and beyond," in Device Research Conference, 2009. DRC 2009, pp. 247 –250, june 2009.

[32] G. Shahidi, "Design-technology interaction for post-32 nm node cmos technologies," in VLSI Technology (VLSIT), 2010 Symposium on, pp. 143 –144, june 2010.

[33] G. Shahidi, "Fdsoi for low power cmos (invited)," in SOI Conference, 2009 IEEE International, pp. 1 –2, oct. 2009.

[34] R. C. Pack, V. Axelrad, A. Shibkov, V. V. Boksha, J. A. Huckabay, R. Salik, W. Staud, R. Wang, and W. D. Grobman, "Physical and timing verification of subwavelength-scale designs: I. lithography impact on mosfets," vol. 5042, pp. 51–62, SPIE, 2003.

[35] S. Mansfield, G. Han, and L. Liebmann, "Through-process modeling for design-for-manufacturability applications," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031007, 2007.

[36] C. S. Xu, J. Liu, and Y. Xia, "Quantitative study of chemical mechanical planarization process affected by bare silicon wafer front surface topography," vol. 17, pp. 2210–2215, AVS, 1999.

[37] M. Orshansky, L. Milor, and C. Hu, "Characterization of spatial intrafield gate cd variability, its impact on circuit performance, and spatial mask-level correction," Semiconductor Manufacturing, IEEE Transactions on, vol. 17, pp. 2 – 11, feb. 2004.

[38] Y. Granik, "Dry etch proximity modeling in mask fabrication," vol. 5130, pp. 86–91, SPIE, 2003.

[39] F. M. Schellenberg, "Resolution enhancement technology: the past, the present, and extensions for the future," vol. 5377, pp. 1–20, SPIE, 2004.

[40] N. Cobb and Y. Granik, "New concepts in opc," vol. 5377, pp. 680–690, SPIE, 2004.

[41] J. A. Torres, "Layout verification in the era of process uncertainty: target process variability bands vs actual process variability bands," vol. 6925, p. 692509, SPIE, 2008.

[42] K. Hashimoto, S. Usui, S. Nojima, T. Kotani, E. Yamanaka, S. Tanaka, and S. Inoue, "Hotspot management and its applications in ultralow k1 lithography," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 8, no. 3, p. 033007, 2009.

[43] J. A. Torres and C. N. Berglund, "Integrated circuit dfm framework for deep sub-wavelength processes," vol. 5756, pp. 39–50, SPIE, 2005.

[44] K. Keutzer and M. Orshansky, "From blind certainty to informed uncertainty," in Proceedings of the 8th ACM/IEEE international workshop on Timing issues in the specification and synthesis of digital systems, TAU '02, (New York, NY, USA), pp. 37–41, ACM, 2002.

[45] J. A. Carballo and S. Nassif, "Dfm requirements and solution roadmaps: the multi-layer approach," vol. 6156, p. 61560Y, SPIE, 2006.

[46] Y. Cheng, "Design for manufacturing (dfm) in nano-cmos era," in Nanoelectronics Conference, 2008. INEC 2008. 2nd IEEE International, pp. 524 –529, march 2008.

[47] D. Sylvester, K. Agarwal, and S. Shah, "Invited paper: Variability in nanometer cmos: Impact, analysis, and minimization," Integr. VLSI J., vol. 41, pp. 319–339, May 2008.

[48] J. A. Carballo and S. Nassif, "Manufacturing-aware design methodologies for mixed-signal communication circuits," vol. 5379, pp. 119–127, SPIE, 2004.

[49] J. Word, J. Belledent, Y. Trouiller, Y. Granik, O. Toublan, and W. Maurer, "Full chip model based correction of flare-induced linewidth variation," vol. 5754, pp. 1209–1219, SPIE, 2004.

[50] T. Yao, M. Osawa, T. Minami, N. Yamamoto, H. Aoyama, G. Okuda, T. Sawano, I. Kamatsuki, F. Sugimoto, H. Futatsuya, K. Kobayashi, K. Ogino, H. Hoshino, Y. Machida, H. Arimoto, and S. Asai, "Local flare effects and correction in arf lithography," in VLSI Technology, 2003. Digest of Technical Papers. 2003 Symposium on, pp. 43 – 44, june 2003.

[51] C. A. Mack, "Measuring and modeling flare in optical lithography," vol. 5040, pp. 151–161, SPIE, 2003.

[52] C. A. Mack, Field Guide to Optical Lithography. Bellingham, WA: SPIE Press, 2006.

[53] A. E. Rosenbluth, S. Bukofsky, C. Fonseca, M. Hibbs, K. Lai, A. F. Molless, R. N. Singh, and A. K. K. Wong, "Optimum mask and source patterns to print a given shape," Journal of Microlithography, Microfabrication, and Microsystems, vol. 1, no. 1, pp. 13–30, 2002.

[54] B. W. Smith, "Spatial filtering effects of the attenuated psm and assist bar opc," vol. 4404, pp. 412–419, SPIE, 2001.

[55] Z. M. Ma and C. A. Mack, "Impact of illumination coherence and polarization on the imaging of attenuated phase-shift masks," vol. 4346, pp. 1522–1532, SPIE, 2001.

[56] T. Fühner, A. Erdmann, and S. Seifert, "Direct optimization approach for lithographic process conditions," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031006, 2007.

[57] B. Yenikaya, O. Alexandrov, S. Chen, A. Liu, A. Mokhberi, and A. Sezginer, "Model-based assist features," vol. 7379, p. 73792Z, SPIE, 2009.

[58] L. A. Joesten, M. T. Reilly, J. DeSisto, and C. Jehoul, "Effect of scattering bar assist features in 193-nm lithography," vol. 4691, pp. 861–870, SPIE, 2002.

[59] N. V. Lafferty, G. Vandenberghe, B. W. Smith, M. Lassiter, and P. M. Martin, "Gray assist bar opc," vol. 5377, pp. 381–392, SPIE, 2004.

[60] J. A. Torres, Y. Granik, and F. Schellenberg, "Contrast analysis and optimization for resolution enhancement technique," Journal of Microlithography, Microfabrication, and Microsystems, vol. 2, no. 2, pp. 119–128, 2003.

[61] P. Gupta, A. B. Kahng, D. Sylvester, and J. Yang, "Performance-driven optical proximity correction for mask cost reduction," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031005, 2007.

[62] Y. Granik, N. Cobb, and D. Medvedev, "Extreme mask corrections: technology and benefits," vol. 6924, p. 69243W, SPIE, 2008.

[63] J. Word, D. Dudau, and N. Cobb, "Automatic opc mask shape repair," vol. 6521, p. 65211G, SPIE, 2007.

[64] D. F. Beale and L. S. M. III, "Etch analysis of patterns to comprehend manufacturability," vol. 6283, p. 62832T, SPIE, 2006.

[65] J. Xue, J. Huang, A. Kazarian, and B. Falch, "A new etch-aware after development inspection (adi) technique for opc modeling," vol. 7640, p. 76402Q, SPIE, 2010.

[66] Y. Zhang, L. Wagner, and P. Golbutsov, "Importance of wafer flatness for cmp and lithography," vol. 3050, pp. 266–269, SPIE, 1997.

[67] G. Ribes, J. Mitard, M. Denais, S. Bruyere, F. Monsieur, C. Parthasarathy, E. Vincent, and G. Ghibaudo, "Review on high-k dielectrics reliability issues," Device and Materials Reliability, IEEE Transactions on, vol. 5, pp. 5 – 19, march 2005.

[68] M. Bohr, R. Chau, T. Ghani, and K. Mistry, "The high-k solution," Spectrum, IEEE, vol. 44, pp. 29 –35, oct. 2007.

[69] K. Agarwal and S. Nassif, "Characterizing process variation in nanometer cmos," in Design Automation Conference, 2007. DAC '07. 44th ACM/IEEE, pp. 396 –399, june 2007.

[70] W. Hansch, K. Anil, P. Bieringer, C. Fink, F. Kaesen, I. Eisele, M. Tanaka, and M. Miura-Mattausch, "Channel engineering for the reduction of random-dopant placement-induced threshold voltage fluctuations in vertical sub-100nm mosfets," in Solid-State Device Research Conference, 1999. Proceeding of the 29th European, vol. 1, pp. 408 –411, sept. 1999.

[71] T. Mizuno, J. Okumtura, and A. Toriumi, "Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in mosfet's," Electron Devices, IEEE Transactions on, vol. 41, pp. 2216 –2221, nov 1994.

[72] K. Itoh, N. Sugii, D. Hisamoto, and R. Tsuchiya, "Fd-soi mosfets for the low-voltage nanoscale cmos era," in SOI Conference, 2009 IEEE International, pp. 1 –4, oct. 2009.

[73] S. Inaba, K. Okano, T. Izumida, A. Kaneko, H. Kawasaki, A. Yagishita, T. Kanemura, T. Ishida, N. Aoki, K. Ishimaru, K. Suguro, K. Eguchi, Y. Tsunashima, Y. Toyoshima, and H. Ishiuchi, "Finfet: the prospective multi-gate device for future soc applications," in Solid-State Device Research Conference, 2006. ESSDERC 2006. Proceeding of the 36th European, pp. 49 –52, sept. 2006.

[74] T.-J. King, "Finfets for nanoscale cmos digital integrated circuits," in Computer-Aided Design, 2005. ICCAD-2005. IEEE/ACM International Conference on, pp. 207 – 210, nov. 2005.

[75] J.-P. Colinge, FinFETs and Other Multi-Gate Transistors. Dordrecht, The Netherlands: Springer, 2008.

[76] T. E. Zavecz, "Bossung curves: an old technique with a new twist for sub-90-nm nodes," vol. 6152, p. 61522S, SPIE, 2006.

[77] Y. Granik, "Source optimization for image fidelity and throughput," Journal of Microlithography, Microfabrication, and Microsystems, vol. 3, no. 4, pp. 509–522, 2004.

[78] S. Banerjee, P. Elakkumanan, D. Chidambarrao, J. Culp, and M. Orshansky, "Analysis of systematic variation and impact on circuit performance," vol. 6925, p. 69250K, SPIE, 2008.

[79] J.-Y. Wuu, F. G. Pikus, A. Torres, and M. Marek-Sadowska, "Detecting context sensitive hot spots in standard cell libraries," vol. 7275, p. 727515, SPIE, 2009.

[80] S. Mansfield, I. Graur, G. Han, J. Meiring, L. Liebmann, and D. Chidambarrao, "Lithography simulation in dfm: achievable accuracy versus requirements," vol. 6521, p. 652106, SPIE, 2007.

[81] R. März, K. Peter, and K. Engelhardt, "Rerouting and guided-repair strategies to resolve lithography hotspots," vol. 7974, p. 79740R, SPIE, 2011.

[82] M. Mirsaeedi, J. A. Torres, and M. Anis, "Self-aligned double-patterning (sadp) friendly detailed routing," vol. 7974, p. 79740O, SPIE, 2011.

[83] D. Pan, M. Cho, K. Yuan, and Y. Ban, "Lithography friendly routing: From construct-by-correction to correct-by-construction," in Solid-State and Integrated-Circuit Technology, 2008. ICSICT 2008. 9th International Conference on, pp. 2232 –2235, oct. 2008.

[84] X. Shi, S. Hsu, J. F. Chen, C. M. Hsu, R. J. Socha, and M. V. Dusa, "Understanding the forbidden pitch and assist feature placement," vol. 4562, pp. 968–979, SPIE, 2002.

[85] T. Jhaveri, V. Rovner, L. Liebmann, L. Pileggi, A. Strojwas, and J. Hibbeler, "Co-optimization of circuits, layout and lithography for predictive technology scaling beyond gratings," Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on, vol. 29, pp. 509 –527, april 2010.

[86] T. Jhaveri, V. Rovner, L. Pileggi, A. J. Strojwas, D. Motiani, V. Kheterpal, K. Y. Tong, T. Hersan, and D. Pandini, "Maximization of layout printability/manufacturability by extreme layout regularity," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031011, 2007.

[87] J. Ho, Y. Wang, J. Wu, Y.-C. Hou, and K. Wu, "Lithography-simulation-based design for manufacturability rule development: an integrated circuit design house's approach," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031008, 2007.

[88] N. Nagase, K. Suzuki, K. Takahashi, M. Minemura, S. Yamauchi, and T. Okada, "Study of hot spot detection using neural network judgment," vol. 6607, p. 66071B, SPIE, 2007.

[89] Y. Yang, M. Chew, T. Endo, and M. Simmons, "Model-based hints for litho-hotspots repair," vol. 7488, p. 74883O, SPIE, 2009.

[90] J. Gookassian, B. Pack, M. Heins, J. Garcia, H. Divecha, B. Gordon, D. Frazier, D. White, G. Lachinyan, B. Dillon, C. Suzor, A. Adamov, K.-Y. Min, S. Bakarian, R. Marutyan, and V. Boksha, "Building an infrastructure for parametric yield analysis: concept and implementation of a dfm platform (invited paper)," vol. 5756, pp. 208–218, SPIE, 2005.

[91] J. Yang, L. Capodieci, and D. Sylvester, "Advanced timing analysis based on post-opc extraction of critical dimensions," in Proceedings of the 42nd annual Design Automation Conference, DAC '05, (New York, NY, USA), pp. 359–364, ACM, 2005.

[92] W. J. Poppe, L. Capodieci, J. Wu, and A. Neureuther, "From poly line to transistor: building bsim models for non-rectangular transistors," vol. 6156, p. 61560P, SPIE, 2006.

[93] S. Shi, P. Yu, and D. Pan, "A unified non-rectangular device and circuit simulation model for timing and power," in Computer-Aided Design, 2006. ICCAD '06. IEEE/ACM International Conference on, pp. 423 –428, nov. 2006.

[94] P. Wright and M. Fan, "A dfm methodology to evaluate the impact of lithography conditions on the speed of critical paths in a vlsi circuit," in Proceedings of the 7th International Symposium on Quality Electronic Design, ISQED '06, (Washington, DC, USA), pp. 813–817, IEEE Computer Society, 2006.

[95] K. Agarwal, R. Rao, D. Sylvester, and R. Brown, "Parametric yield analysis and optimization in leakage dominated technologies," Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 15, pp. 613 –623, june 2007.

[96] K. B. Agarwal and S. Banerjee, "Integrated model-based retargeting and optical proximity correction," vol. 7974, p. 79740F, SPIE, 2011.

[97] J. G. Tsacoyeanes, "Off-axis illumination for improving depth of focus for isolated features," vol. 3679, pp. 750–761, SPIE, 1999.

[98] B. J. Lin, "Off-axis illumination–working principles and comparison with alternating phase-shifting masks," vol. 1927, pp. 89–100, SPIE, 1993.

[99] S. Owa and H. Nagasaka, "Immersion lithography: its history, current status and future prospects," vol. 7140, p. 714015, SPIE, 2008.

[100] B. J. Lin, "Immersion lithography and its impact on semiconductor manufacturing," Journal of Microlithography, Microfabrication, and Microsystems, vol. 3, no. 3, pp. 377–395, 2004.

[101] K. Adam and W. Maurer, "Polarization effects in immersion lithography," Journal of Microlithography, Microfabrication, and Microsystems, vol. 4, no. 3, p. 031106, 2005.

[102] B. J. Lin, "Immersion lithography," 2010.

[103] R. O. Topaloglu, "Is manufacturability with double patterning a burden on designer? analyses of device and circuit aspects," vol. 7974, p. 79740A, SPIE, 2011.

[104] Y. Wei and R. L. Brainard, "Double exposure and double patterning," 2009.

[105] J. Finders, M. Dusa, B. Vleeming, B. Hepp, M. Maenhoudt, S. Cheng, and T. Vandeweyer, "Double patterning lithography for 32 nm: critical dimensions uniformity and overlay control considerations," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 8, no. 1, p. 011002, 2009.

[106] R. Rodrigues and S. Kundu, "Model based double patterning lithography (dpl) and simulated annealing (sa)," in Quality Electronic Design (ISQED), 2011 12th International Symposium on, pp. 1 –8, march 2011.

[107] M. Mirsaeedi, J. Torres, and M. Anis, "Self-aligned double patterning (sadp) layout decomposition," in Quality Electronic Design (ISQED), 2011 12th International Symposium on, pp. 1–7, March 2011.

[108] Y. Granik and N. B. Cobb, "Meef as a matrix," vol. 4562, pp. 980–991, SPIE, 2002.

[109] Y. Granik, "Generalized mask error enhancement factor theory," Journal of Microlithography, Microfabrication, and Microsystems, vol. 4, no. 2, p. 023001, 2005.

[110] T. M. Tawfik, A. H. Morshed, and D. Khalil, "Modeling mask scattered field at oblique incidence," vol. 7274, p. 727433, SPIE, 2009.

[111] A. Erdmann, G. Citarella, P. Evanschitzky, H. Schermer, V. Philipsen, and P. D. Bisschop, "Validity of the hopkins approximation in simulations of hyper-na (na ¿ 1) line-space structures for an attenuated psm mask," vol. 6154, p. 61540G, SPIE, 2006.

[112] Q. C. Zhang, K. Lucas, P. van Adrichem, J. Tyminski, and J. S. Gordon, "Novel apodization and pellicle optical models for accurate optical proximity correction modeling at 45 and 32 nm," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031003, 2007.

[113] K. Lai, A. E. Rosenbluth, G. Han, J. Tirapu-Azpiroz, J. Meiring, A. Goehnermeier, B. Kneer, M. Totzeck, L. de Winter, W. de Boeij, and M. van de Kerkhof, "Modeling polarization for hyper-na lithography tools and masks," vol. 6520, p. 65200D, SPIE, 2007.

[114] A. Erdmann and P. Evanschitzky, "Rigorous electromagnetic field mask modeling and related lithographic effects in the low k[sub 1] and ultrahigh numerical aperture regime," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031002, 2007.

[115] J. W. Goodman, "Introduction to fourier optics," in McGraw-Hill Book Company, 19681.

[116] C. A. Mack, "Fourier optics," 2006.

[117] K. Adam and M. C. Lam, "Hybrid hopkins-abbe method for modeling oblique angle mask effects in opc," vol. 6924, p. 69241E, SPIE, 2008.

[118] N. Cobb, "Fast optical and process proximity correction algorithms for integrated circuit manufacturing," PhD Thesis, University of California, Berkeley, 1998.

[119] S. Kobayashi, S. Kyoh, T. Kotani, S. Tanaka, and S. Inoue, "Automated hotspot fixing system applied to the metal layers of 65-nm logic devices," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031010, 2007.

[120] P. Yu, S. X. Shi, and D. Z. Pan, "True process variation aware optical proximity correction with variational lithography modeling and model calibration," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031004, 2007.

[121] A. Balasinski, J. Cetin, and L. Karklin, "Layout techniques and rules to reduce process-related variability," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 6, no. 3, p. 031009, 2007.

[122] J. T. Azpiroz, A. Krasnoperova, S. Siddiqui, K. Settlemyer, I. Graur, I. Stobert, and J. M. Oberschmidt, "Improving yield through the application of process window opc," vol. 7274, p. 727411, SPIE, 2009.

[123] M. L. Kempsell, E. Hendrickx, A. Tritchkov, K. Sakajiri, K. Yasui, S. Yoshitake, Y. Granik, G. Vandenberghe, and B. W. Smith, "Inverse lithography for 45-nm-node contact holes at 1.35 numerical aperture," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 8, no. 4, p. 043001, 2009.

[124] Y. Granik, "On the uniqueness of optical images and solutions of inverse lithographical problems," Journal of Micro/Nanolithography, MEMS and MOEMS, vol. 8, no. 3, p. 031405, 2009.

[125] K. Tian, M. Fakyry, A. Dave, A. Tritchkov, J. Tirapu-Azpiroz, A. E. Rosenbluth, D. Melville, M. Sakamoto, T. Inoue, S. Mansfield, A. Wei, Y. Kim, B. Durgan, K. Adam, G. Berger, G. Bhatara, J. Meiring, H. Haffner, and B.-S. Kim, "Applicability of global source mask optimization to 22/20nm node and beyond," vol. 7973, p. 79730C, SPIE, 2011.

[126] K. Tian, A. Krasnoperova, D. Melville, A. E. Rosenbluth, D. Gil, J. Tirapu-Azpiroz, K. Lai, S. Bagheri, C. chen Chen, and B. Morgenfeld, "Benefits and trade-offs of global source optimization in optical lithography," vol. 7274, p. 72740C, SPIE, 2009.

[127] A. E. Rosenbluth, D. O. Melville, K. Tian, S. Bagheri, J. Tirapu-Azpiroz, K. Lai, A. Waechter, T. Inoue, L. Ladanyi, F. Barahona, K. Scheinberg, M. Sakamoto, H. Muta, E. Gallagher, T. Faure, M. Hibbs, A. Tritchkov, and Y. Granik, "Intensive optimization of masks and sources for 22nm lithography," vol. 7274, p. 727409, SPIE, 2009.

[128] M.-C. Tsai, S. Hsu, L. Chen, Y.-W. Lu, J. Li, F. Chen, H. Chen, J. Tao, B.-D. Chen, H. Feng, W. Wong, W. Yuan, X. Li, Z. Li, L. Li, R. Dover, H. yu Liu, and J. Koonmen, "Full-chip source and mask optimization," vol. 7973, p. 79730A, SPIE, 2011.

[129] M. Fakhry, Y. Granik, K. Adam, and K. Lai, "Total source mask optimization: high-capacity, resist modeling, and production-ready mask solution," vol. 8166, p. 81663M, SPIE, 2011.

[130] S. Banerjee, K. B. Agarwal, J. A. Culp, P. Elakkumanan, L. W. Liebmann, and M. Orshansky, "Compensating non-optical effects using electrically driven optical proximity correction," vol. 7275, p. 72750E, SPIE, 2009.

[131] Q. C. Zhang and P. van Adrichem, "Determining opc target specifications electrically instead of geometrically," vol. 6730, p. 67303V, SPIE, 2007.

[132] A. Y. Hamouda, "Mask-friendly opc for a reduced mask cost and writing time," vol. 6520, p. 65203Y, SPIE, 2007.

[133] J. Word, K. Mizuuchi, S. Fu, W. Brown, and E. Sahouria, "Mask shot count reduction strategies in the opc flow," vol. 7028, p. 70283F, SPIE, 2008.

[134] H. Zhang, Y. Du, M. Wong, and K.-Y. Chao, "Mask cost reduction with circuit performance consideration for self-aligned double patterning," in Design Automation Conference (ASP-DAC), 2011 16th Asia and South Pacific, pp. 787 –792, jan. 2011.

[135] P. Gupta, A. Kahng, D. Sylvester, and J. Yangt, "Performance-driven opc for mask cost reduction," in Quality of Electronic Design, 2005. ISQED 2005. Sixth International Symposium on, pp. 270 – 275, march 2005.

[136] A. Hamouda, M. Anis, and K. S. Karim, "Layer (lithography and yield-enhancing retargeting) : A novel automatic model-based retargeting and auto-correction lfd flow," vol. Submitted for publication, 2015.

[137] J. Lei, L. Hong, G. Lippincott, and J. Word, "Model-based opc using the meef matrix ii," 2014.

[138] Y. Ping, S. McGowan, Y. Gong, Y. M. Foong, J. Liu, J. Qiu, V. Shu, B. Yan, J. Ye, P. Li, H. Zhou, T. Pandey, J. Liang, C. Aquino, S. Baron, and S. Kapasi, "Process window enhancement using advanced ret techniques for 20nm contact layer," 2014.

[139] J. Wang, P. Verma, A. Wei, and W. Wilkinson, "Conducting opc retargeting as guided by principles of classical dynamics," 2014.

[140] A. Hamouda, M. Anis, and K. S. Karim, "Mib (model-based initial bias) : Towards a single-iteration optical proximity correction," vol. Submitted for publication, 2015.