



HAL
open science

Diminution of the lithographic process variability for advanced technology nodes

Anna Szucs

► **To cite this version:**

Anna Szucs. Diminution of the lithographic process variability for advanced technology nodes. Micro and nanotechnologies/Microelectronics. Université Grenoble Alpes, 2015. English. NNT : 2015GREAT129 . tel-01278511

HAL Id: tel-01278511

<https://theses.hal.science/tel-01278511>

Submitted on 24 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ GRENOBLE ALPES

THESIS

To obtain the degree of

DOCTOR FROM UNIVERSITE DE GRENOBLE

Specialty: **Nano technology and Nano electronics**

Ministerial Order: 6 August 2006

Presented by

Anna SZ UCS

Directed by **Cécile GOURGON** and co supervised by **Jonathan PLANCHOT** and **Maxime Besacier**

Prepared at **Laboratoire des Technologies de la Microélectronique (LTM)**,
and at **STMicroelectronics**, at **École Doctorale d'Electronique,**
Electrotechnique (EEATS)

Diminution of the lithographic process variability for advanced technology nodes

Thesis defended publically « 10 December 2015»,

In front of the jury composed of:

Mr. Peter DeBisshop

Engineer - IMEC (Reviewer)

Mr. Régis Orobthouk

Professor assistant - INL (Reviewer)

Mr. Gérard Granet

Professor - Institut Pascal (Examiner)

Mr. Ahmad Bsiesy

Professor - LTM (President)

Mr. Jonathan PLANCHOT

Engineer - STMicroelectronics (co-supervisor)

Mme. Cécile GOURGON

Directrice de Recherches - LTM (supervisor)

Mr. Maxime BESACIER

professor assistant - LTM (co-supervisor- invited)

Mr. Jo Finders

Engineer- ASML (invited)



UNIVERSITÉ GRENOBLE ALPES

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Nano électronique et Nano Technologies** Arrêté ministériel: 7 août 2006

Présentée par

Anna SZUCS

Thèse dirigée par **Cécile GOURGON** and Co-Encadrée par **Jonathan PLANCHOT** et **Maxime Besacier**

Préparée au sein du **Laboratoire des Technologies de la Microélectronique (LTM)** et de **l'entreprise STMicroelectronics** à l'École Doctorale **d'Electronique, Electrotechnique (EEATS)**

Diminution de la variabilité du procédé lithographique pour les nœuds technologiques avancés

Thèse soutenue publiquement le « **10 Décembre 2015** »
devant le jury composé de :

Mr. Peter DeBisshop

Ingénieur - IMEC (Rapporteur)

Mr. Régis Orobtcouk

Maître de Conférences - INL (Rapporteur)

Mr. Gérard Granet

Professeur - Institut Pascal (Examineur)

Mr. Ahmad Bsiesy

Professeur - LTM (Président)

Mr. Jonathan PLANCHOT

Ingénieur - STMicroelectronics (Encadrant)

Mme. Cécile GOURGON

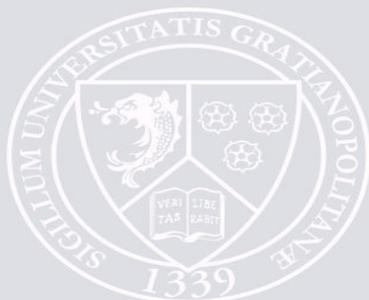
Directrice de Recherches - LTM (Directeur de thèse)

Mr. Maxime BESACIER

Maître de Conférences - LTM (encadrant - invité)

Mr. Jo Finders

Ingénieur - ASML (invité)



Köszönöm Atyám!

Acknowledgment

I'm very grateful for all, who helped me to accomplish this work.

First of all, I would like to express my gratitude to my supervisors: Cécile Gourgon, Maxime Besacier and Jonathan Planchot for their guidance on this project. Their kind help, dedication and knowledge have helped to convert this thesis work from some initial ideas to a completed thesis.

I would like to express my gratitude to all the RET (OPC) team at STMicroelectronics for their kindness and helpful discussions. They have contributed knowledge and friendship into my life. Beside, special thanks for this team for all the delicious cakes during the coffee breaks – except for the little specialty from Thailand.

I would like to thank to all the members from ASML Brion, ASML, LTM, Leti and CIME Nanotechnology who helped to complete this thesis work.

I am especially grateful to my family for their support and kind words. They have given me endless love and support from the beginning to end, being there all the time I needed. Also, I would like to express special thanks to Wafa and Lama for their devoted friendship.

Abstract

The diminution of the lithographic process variability for advanced technology nodes

The currently used 193 nm optical lithography reaches its limits from resolution point of view. It is despite of the fact that various techniques have been developed to push this limit as much as possible. Indeed new generation lithography exists such as the EUV, but are not yet reliable to be applied in mass production. Thus in orders to maintain a robust lithographic process for these shrunk nodes, 28 nm and beyond, the optical lithography needs to be further explored. It is possible through alternatives techniques: e.g. the RETs (Resolution Enhancement Techniques), such as OPC (Optical Proximity Correction) and the double patterning. In addition to the resolution limits, advanced technology nodes are dealing with increasing complexity of design and steadily increasing process variability requiring more and more compromises.

In the light of this increasing complexity, this dissertation work is addressed to mitigate the lithographic process variability by the implementation of a correction (mitigation) flow explored mainly through the capability of computational lithography. Within this frame, our main objective is to participate to the challenge of assuring a good imaging quality for the process window limiting patterns with an acceptable gain in uDoF (usable Depth of Focus).

In order to accomplish this task, we proposed and validated a flow that might be later implemented in the production. The proposed flow consists on simulation based detection methodology of the most critical patterns that are impacted by effects coming from the mask topography and the resist profile. Furthermore it consists of the mitigation and the compensation of these effects, once the critical patterns are detected. The obtained results on the completed flow are encouraging: a validated method that detects the critical patterns and then mitigates the lithographic process variability been developed successfully.

Keywords: Microelectronics, 193 nm Optical Lithography, RET (Resolution Enhancement Techniques), OPC (Optical Proximity Correction), mask and resist 3D

Résumé

A l'heure actuelle, la lithographie optique 193 nm arrive à ses limites de capacité en termes de résolution des motifs dans la fenêtre du procédé souhaitée pour les nœuds avancés. Des lithographies de nouvelle génération (NGL) sont à l'étude, comme la lithographie EUV (EUV). La complexité de mise en production de ces nouvelles lithographies entraîne le fait que la lithographie 193 nm continue à être exploitée pour les nœuds 28 nm et au-delà. Afin de suivre la miniaturisation le rôle des techniques alternatives comme les RET (en anglais Resolution Enhancement Techniques) tels que l'OPC (Optical Proximity Correction) est devenu primordial et essentiel. Néanmoins, la complexité croissante de design et de la variabilité du procédé lithographique font qu'il est nécessaire de faire des compromis.

Dans ce contexte de complexité croissante du procédé de fabrication, l'objectif de la thèse est de mettre en place une méthode de boucles de correction des facteurs de variabilité. Cela signifie une diminution de la variabilité des motifs complexes pour assurer une résolution suffisante dans la fenêtre de procédé. Ces motifs complexes sont très importants, car c'est eux qui peuvent diminuer la profondeur du champ commune (uDoF).

Afin d'accomplir cette tâche, nous avons proposé et validé un enchaînement qui pourra être plus tard implémenté en production. L'enchaînement en question consiste en une méthodologie de détection basée sur la simulation des motifs les plus critiques étant impactés par les effets issus de la topographie du masque et du profil de la résine. En outre cette méthodologie consiste en une diminution et la compensation de ces effets, une fois que ces motifs les plus critiques sont détectés. Le résultat de l'enchaînement complété sont encourageants : une méthode qui détecte et diminue les variabilités du processus lithographique pour des nœuds de technologie de 28nm a été validée. En plus elle pourrait être adaptée pour les nœuds au-delà de 28 nm.

Mots clé: Microélectronique, Lithographie optique 193 nm, RET (Resolution Enhancement Techniques), modèle OPC (Optical Proximity Correction), masque 3D, résine 3D

Publications

A. Szucs et al. « Best focus shift mitigation for extending the depth of focus», In: Proceedings of SPIE 8326, Optical Microlithography (2013).

A. Szucs et al. « Advanced OPC Mask-3D and Resist-3D modeling », In: Proceedings of SPIE 8326, Optical Microlithography (2014).

Co-author:

B. Le-Gratiet et al. « Patterning critical dimension control for advanced logic nodes », In: Journal of Micro/Nanolithography, MEMS, and MOEMS Volume 14, Issue 2 (2015).

Table of contents

Acknowledgment.....	4
Abstract	5
Résumé	6
Publications	7
List of acronyms.....	11
General Introduction	13
1 Chapter	16
1.1 From simple transistor to IC – CMOS technology	17
1.2 Fundamental concept of photolithography	26
1.2.1 The principle of patterning (photolithography process)	26
1.2.2 The Photolithography system elements.....	27
1.3 The Lithographic Performance Indicators	41
1.4 Resolution Enhancement Technology (RET).....	44
1.5 Conclusion	46
2 Chapter	47
2.1 Introduction.....	48
2.2 OPC model	51
2.2.1 Optical model	51
2.2.2 Resist model	53
2.3 OPC model calibration.....	57
2.3.1 The basic calibration patterns	58
2.3.2 Process window – a reliable metrology setup.....	59
2.3.3 Modeling requirements (the accuracy rate)	60
2.4 Current limitations to “everything is linked” solution.....	61
2.5 Conclusion	63
3 Chapter	64
3.1 Introduction.....	65
3.2 Experimental (modeling and simulation) constraints	67
3.3 Advanced 3D modeling: the mask 3D model (M3D)	67

3.3.1	The motivation for a M3D model - the phase offset related the best focus deviation	67
3.3.2	Demonstration of the best-focus shift problematic through simulation	69
3.3.3	Results and discussion on the mask 3D model.....	70
3.4	Advanced 3D modeling - The Resist 3D model.....	72
3.4.1	Possible approaches of a resist 3D model construction.....	73
3.4.2	What is behind the applied resist profile aware 3D modeling?	75
3.5	Conclusion	88
4	Chapter	89
4.1	Detection methodology.....	90
4.2	Detection approach among the mask topography induced effects.....	90
4.2.1	The problematic among the mask topography induced effects	90
4.2.2	The proposed solution: mask topography conducted detector.....	90
4.2.3	Results and discussion	92
4.3	Detection approach among the resist profile induced effects.....	94
4.3.1	The problematic among the resist profile induced effects	94
4.3.2	The proposed solution: resist profile conducted detector.....	97
4.3.3	Results and discussion among the resist profile related detectors.....	98
4.4	Conclusion	104
5	Chapter	106
	Introduction.....	Erreur ! Signet non défini.
5.1	Mitigation of the mask topography related effects at scanner level	107
5.1.1	The best focus shift mitigation by wavefront optimization	107
5.1.2	Experimental setup for wavefront optimization	109
5.1.3	Results and discussion	112
5.1.4	Partial Conclusion on mask topography related mitigation at scanner level by wavefront optimization.....	118
5.2	Mitigation through the resist profile related effects	119
5.2.1	The source optimization (SO)	120
5.2.2	Data setup for the simulation.....	122
5.2.3	Results and discussion	123
5.2.4	Validation of the optimized source	124
5.2.5	Partial conclusion along the resist profile related mitigation	133

5.3	Conclusion	133
	General conclusion	134
	Annex 1.....	137
6	Résumé en français	142
6.1	Définitions et concepts.....	143
6.2	Plan expérimental.....	146
6.3	Les résultats et les faits marquants	146
6.3.1	La Détection par application des modèles avancés : masque 3D / résine 3D.....	146
6.3.2	La Diminution	161
6.3.3	Conclusion	174
	Bibliography.....	175

List of acronyms

AFM	Atomic Force Microscope
ARC	Anti-Reflective Coating
ArF	Argon-Fluor
BARC	Bottom Anti Reflective Coating
CD	Critical Dimension
CMOS	Complementary Metal Oxide Semiconductor
CMP	Chemical Mechanical Planarization
DoF	Depth of Focus
DRC	Design Rule Check
DUV	Deep Ultra Violet
E-beam	Electron Beam
EL	Energy Latitude
EUV	Extreme Ultra Violet
FEM	Focus Exposure Matrix
FLEX	Focus Latitude enhancement Exposure
IC	Integrated Circuit
ITRS	International Technology Roadmap for Semiconductors
LE	Line End
LMC	Lithography Manufacturability Check
M3D	Mask 3D
MEEF	Mask Enhancement Error Factor
MOS	Metal Oxide Semiconductor
MRC	Mask Rule Check
NA	Numerical Aperture
NGL	Next Generation Lithography
OAI	Off Axis Illumination
OPC	Optical Proximity effects Correction
OPD	Optical Path Difference
OPE	Optical Proximity Effect
p	Pitch
PAG	Photo Acid Generator
PB	Post Bake
PEB	Post Exposure Bake
PSM	Phase Shifting Mask
PW	Process Window
PWQ	Process Window Qualification
R	Resolution
R3D	Resist 3D

RET	Reticle Enhancement Technology
RMS	Root Mean Square
SB	Scatter Bar
SEM	Scanning Electron Microscopy
SMO	Source Mask Optimization
SRAF	Sub Resolution Assist Feature
TARC	Top Anti Reflective Coating
uDoF	usable Depth of Focus
UV	Ultra Violet
λ	Wavelength
σ	Partial Coherence

General Introduction

It has been now for at least a half century, that the microelectronic field¹ is searching for pushing its boundary with the objective to engineer more and more useful electronic devices on a given surface. The “why” sometimes is hard to understand, but the semiconductor industry is mainly driven by this fact. The currently applied Integrated Circuits’ (IC) major role is to implement as much electronic function as possible with various complexity levels on the same chip. Then, these IC are used in a large scope of applications from entertainment industry to military applications. Indeed to satisfy these various needs, related to semiconductor industry, any perspective progresses towards the improvement of such electronic systems.

The major conducting parameters in research, with the objective to achieve higher performance ICs, are the speed, the integration density, the reduced consumption of power and, of course, the reduced fabrication costs. Linked to the cost reduction, a pronounced effort needs to be put on the capability and robustness of the process.

Focusing on the complex processing stages of IC’s mass-production, without boasting, we can say that its centerpiece is the lithography. It is coming from the fact that the IC’s speed and the component’s integration density depend on the minimum feature size or critical dimension (CD) that can be resulted by the lithography. Indeed, the higher integration density on a given surface enables the circuit a higher complexity level. It is coupled with higher velocity for similar fabrication costs. Beyond this aspect, the lithography accounts around 30 % of the total manufacturing costs of a chip (Mack C. , 2015). This percentage is unceasingly increasing and is one of the most critical point for the process.

As current state in mass production, the mostly applied technology is the photolithography that enables an efficient compromise in between robustness and high productivity capability. It uses deep ultraviolet monochromatic (DUV) light to rapidly pattern all the design of the transistor. It is realized by transferring the object, the photomask image, into the photo sensible material (photoresist) on a substrate through a projection system. Its exposure wavelength and numerical aperture (NA) are directly related to the minimum feature size – thus the resolution limit. The decrease of the exposure wavelength or the augmentation of the NA would enable to minimize the smallest dimension that can be patterned. Currently the most reliable photolithography tool in industry for advanced nodes (28nm) in term of capability, robustness is the projection system with 193 nm of exposure wavelength and 1.35 NA. Several techniques were tested in the beginning of this century to decrease the wavelength: as example 157 nm aiming the 65 nm node technology or 126 nm aiming the 45 nm node technology. Nevertheless there was no success due to the uncontrollable illumination source of the tool and due to difficulties in the projection optics (lens). Thus the 193 nm projection systems seemed to be the best choice that required further enhancement, as it was initially intended to use for 90 and 65 nm nodes technologies and not for “28 nm and beyond”.

¹ The microelectronic field encompass: the study and the fabrication of electronic components at the micro and nano scale, connected together on the same substrate to form Integrated Circuits (IC).

Indeed, there is a physical limit, but still the resolution limit is pushed as far as possible. It means that the lithographers are faced to this challenge to figure out new imaging strategies (and tool enhancement). The ensemble of the various original ideas and the creativity, usually referred as Resolution Enhancement Techniques (RETs), temporarily offers solution to push the limiting factors of the projection printing. Thus it is still possible to follow the miniaturization tendency by the support of RETs, including e.g. Optical Proximity Correction (OPC) to compensate the process induced effects by modifying the sizes and the shapes of patterns at mask level.

The processing solutions for 28 nm and beyond nodes would not be possible in fact without these knobs. In addition to the shrunk dimension, these advanced nodes have to deal with complex features at design level, such as metal layers. The resolution within a sufficient process window is in fact a challenge. The process window means the range of tolerated variation of the position of the focal plane and the dose for which the resolution is within a specified limit. It is usually referred by the usable Depth of Focus (uDoF) – the higher uDoF is, the more stable the process is. So, the aim is to expand as much as possible the uDoF, since for 28 nm and beyond nodes the uDoF is now indeed reduced.

The thesis work then is put in this increasingly complex production environment with the objective to establish a process variability correction methodology that would enable for these complex patterns an expanded process window. In order to achieve this set goal, it is first necessary to detect and select the most critical features by an efficient methodology supported by computational lithography. Then the process variability needs to be mitigated through these detected patterns. The thesis work layout is as follow:

The **first chapter** describes in details the context. A brief history and substances on the semiconductor industry are presented for a deeper understanding of important aspects of integrated circuits (ICs) fabrication and its challenge faced to engineering. The emphasis is then put on the photolithography by the description of a classical exposure system. Its boundaries are listed as well as a possible way of amelioration for each of them - outlined of the importance of the computational lithography (thus OPC modeling).

The **second chapter** deals with the State-of-Art focusing on the interrelation of OPC modeling and production environment. For deeper understanding of the challenge hide behind, a description is provided on the modeling portion of the scanner. Then from the current limitations we arrive to an “everything is linked”. It means that we anticipate in the realization of a dynamic link between OPC and process environment by the setup of our above mentioned correction methodology.

The **third chapter** is dedicated to ground of the detection part by the enhancement of the available OPC models’ upfront prediction. The necessity is coming from the fact that the conventional models became obsolete for shrunk nodes. In the light of this fact, the intention is put on the introduction of the construction background of 3D aware modeling (mask 3D and resist 3D) enabling an efficient support for a higher level detection of critical features.

The **forth chapter** is dedicated subsequently to the detection and selection of the most critical patterns based on the calibrated 3D aware OPC models. It includes discussions from the basic

consideration to the proof of concept. The detection and selection is conducted among the mask topography related and the resist profile related effects. Our objective is to validate the set detection methodologies as efficient and reliable support for production on design layout.

Finally, the **fifth chapter** provides the mitigation solution of the process variation on the previously detected critical patterns in continuity of the forth chapter. Subsequent to the forth chapter, two approaches are demonstrated, one for compensation of the mask topography effects and one for the resist profile effects.

1 Chapter

Introduction to the photolithography

The space developing and shrinking micro components of a circuit board are followed by a steadily increasing complexity of process. Behind this engineering feat, the machinery needs to respond to this high complexity task. Current IC (Integrated Circuit) process technology owns fabrication processes consisting of hundreds process steps applied on a raw semiconductor material wafer. The majority of these steps can be aggregated into the following ensembles:

- Photo mask pattern transfer by lithography and etching
- Layer forming by deposition, oxidation or metallization
- Modification of the layer by ion implantation and diffusion.

The lithography is one of the essential processes in the fabrication plant that can be called also the heart of the micro component technology. Its main task is to transfer the designed photomask pattern image on the silicon wafer surface. This patterning method had and has an efficient adaptive capability to support different technologies from 10 μm technology nodes from early stage of IC fabrication (1960) to today's 28 nm and beyond. Therefore the extension of lithography limits is steadily on-going, which is driven by transistor dimension shrinking.

So this requirement provided the bedrock of this thesis work that is placed into the lithography world. Thus this chapter is dedicated first to a brief historical insight in the semiconductor industry from the invention of the first transistor to current challenges (through the elementary MOS transistor). More precisely, at this point, we will investigate in the basic concept of lithography, more specifically of the photolithography. This part will include a description of a classical exposure system emphasizing on its characteristics, performance indicators, but also limitation factors and possible improvement.

1.1 From simple transistor to IC – CMOS technology

As foreword of this thesis work, the main entrances to the microelectronic world is the concept of the first bipolar-junction transistor on a germanium monocrystalline by the team of W. B. Shockley dated at 1949. A transistor is a semiconductor component or device that can be used to switch or to amplify electronic signals and power.

The concept of the first transistor is a result of an initialized research project on semiconductor technology to replace vacuum tubes for telephone applications. This investigation was started because of the weak reliability and pronounced power consumption of vacuum tubes used for phones. Their ability could not follow the increasing phone call capacity. The patent of the junction transistor is the result of W. B. Shockley's theoretical activity on the point-contact transistor, an early transistor solution by Bardeen and Brattain. Besides the deeper understanding of his theoretical work, the junction transistor showed more potential to be a good candidate for high volume production.

Five years later from its first introduction, as the initial stage of the electrical micro component fabrication, the transistor commercialization has been already started in the same Bell Laboratories. For many years, transistors were fabricated as individual components and the connection between these electronic components (diodes, capacitors, etc.) was realized on a board to obtain an electronic circuit (circuit board). The need of a smaller, faster circuit board came with the fact that there was a pronounced issue due to time delays for electrical signals to propagate on such a long distance in these wildish boards. Thus after many developments J. Kilby from Texas Instruments came up with the solution to this problematic of a multitude of components. So the first integrated circuit (IC) was built, placing several transistors, in additional some diodes and resistors, by the same process on the same semiconductor surface.

In 1968 the team of R. Noyce put down the technology enabling the IC realization of the MOS (Metal Oxide Semiconductor) transistors. Different MOS transistors, with different conduction – N and P type – enabled the construction of logic functions. The technology making use of both P and N channel devices within the same substrate material, so called CMOS (*Complementary Metal Oxide Semiconductors*) technology, is then the milestone of industrial IC development. Until present-day solutions, this *CMOS* processing technology is the most currently applied in the integrated circuit industry. This technology method's flexibility allows an efficient adaptation to a large scale of existing and innovative applications. It is still a key solution for the current global trend realizing electronic components with a better tradeoff between energy consumption and performance (Zeggaoui, 2011) (Michel J.-C. , 2014) (CHM, 2015) (Wikipedia, 2015).

The following sections are dedicated to the description of the basic mechanism of MOS transistor.

1.1.1 The basic elements of an Integrated Circuit – MOS transistor and its operation

The basic element of an electronic gadget or an IC is the MOS (Metal Oxide Semiconductor) transistor. It is composed of three electrode terminals: Gate, Source and Drain. The illustrated example, a MOS transistor, on Figure 1, is "processed" on bulk silicon substrate and consists on the following principal elements:

- Substrate : silicon substrate doped N (N type transistor) or P (P type transistor);
- Gate : commanding electrode of the transistor that can be poly-silicon or a stack of metal gate for advanced technologies; noted by number 3 on Figure 1
- Source/Drain : corresponding to a highly doped charge tank with an opposite doping type to the substrate's - indicated by number 5 and 6 respectively Source and Drain on Figure 1;
- Insulating dielectric layer: usually SiO_2 , or for advanced application, it can be a high K dielectric in addition on SiO_2 - indicated as number 4. Gate oxide on Figure 1;
- Shallow Trench Isolation (STI): applied to electrically isolate transistors in between each other - represented by number 2 on Figure 1.

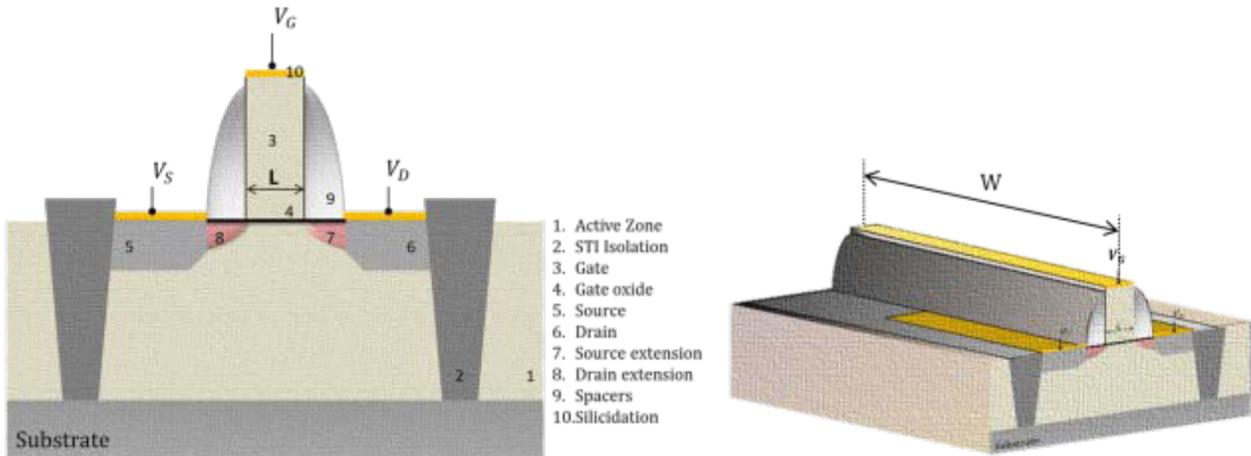


Figure 1 Simplified cross section and 3D view of a MOS transistor with its basic elements

A key point to note is that a technology node in fabrication is denoted (most usually) by the dimension of the Gate as marked “L” on Figure 1. Mostly the shorter size of the Gate is called Gate Length “L” and the longer is called Gate Width “W”

This MOS device is voltage controlled. By means that it is analogous to an electrical switch and enables the passage of the electric current according to the applied voltage on the gate. The operational principle is based on the ‘field effect’, on the electrostatic modulation of the density of mobile charges in the semiconductor. It can be occurred either by a potential variation between the Gate and the Source (V_{GS}) that creates inversion charges at the surface of the semiconductor by a transversal field effect through the Gate Oxide. Or either a potential difference among the Source and the Drain (V_{DS}) that enables the circulation of the minority carriers between them that generates the Drain Current (I_D) (Rahhal, 2014). Figure 2 represents the named Field effect on a MOS stack. Usually it is called MOSFET referring to Metal Oxide Semiconductor Field Effect Transistor.

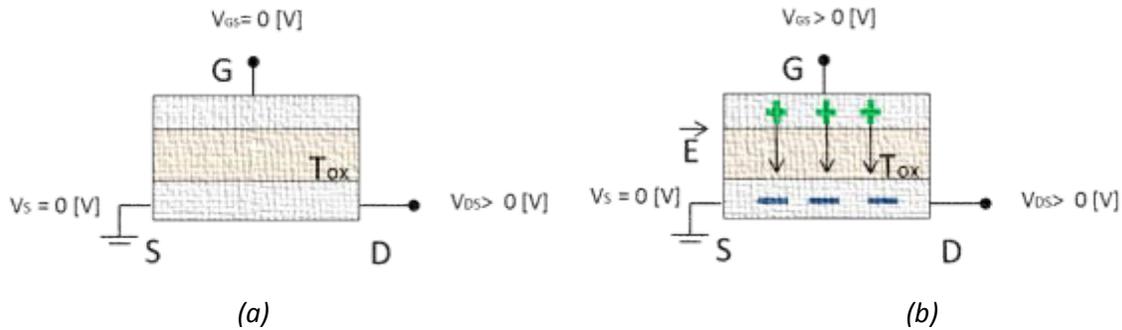


Figure 2 Representation of the Field effect for a MOS transistor (a) $V_{GS}=0 [V]$ so no channel is induced therefor $I_{DS}=0 [A]$ (b) $V_{GS} > 0 [V]$ channel is filled with minority carriers thus $I_{DS} \neq 0 [A]$

Now, if the applied V_{GS} is lower than the threshold voltage (V_{Th}), the transistor remains blocked or non-conducting. This means that there is no circulating current between the Source and the Drain. Still if V_{GS} is equivalent or higher than V_{Th} , the involved minority carriers are vertically modulated to the semiconductor's surface thus a channel region is obtained under the Gate's insulator, the T_{ox} . If the minor carrier's modulation is realized vertically by applied V_{DS} , an I_{DS} current will pass from the Drain to the Source (see Figure 3).

The simplified representation of the MOS transistor's different Operating Regimes in function of the applied V_{DS} is represented on Figure 3. Three regions are differentiated (from left to right on Figure 3):

- The linear Operation Region, where the V_{DS} is negligible, thus I_{DS} shows linear tendency in function of the increasing V_{DS} ;
- The Saturation mode at point of pinch-off, where the applied V_{DS} is equal to the $V_{DSaturation}$;
- The Saturation mode , where the applied V_{DS} is higher than the $V_{DSaturation}$ (the V_{DS} is a pronounce value), thus the device is controlled by the V_{DS} and the I_{DS} is constant

Note: for further details, the reader is kindly invited to check the following references (Arfaoui, 2012) (Rahhal, 2014).

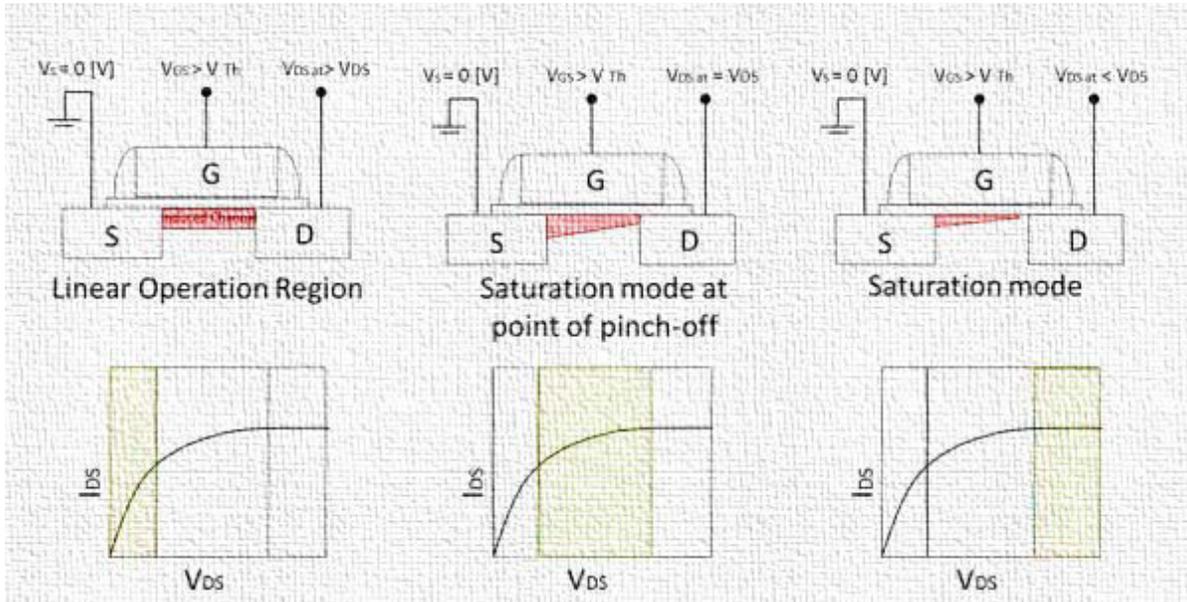


Figure 3 the MOS transistor's different Operating Regimes in function of the applied V_{DS} .

The proper functioning of a complex electronic circuit, based on these elementary adobes, has well defined requirements that need an optimal operation of the transistor components. As example to obtain a transistor with low consumption, the output or Drain current (I_{ON} or I_{DS}) needs to be the highest possible with a low level of the leakage current (I_{OFF} or I_{GS}). Note: the I_{OFF} leakage current is when the channel is blocked and the I_{ON} Drain current is appearing when the transistor is in a passing mode of operation, when the channel is induced:

- The leakage current, I_{OFF} , is obtained while $V_{GS} = 0 [V]$ and $V_{DS} = V_{DD}[V]$ (NMOS) and $V_{DS} = -V_{DD}[V]$ (PMOS)
- The Drain current, I_{ON} , is obtained while $V_{GS} = V_{DS} = V_{DD}$ (NMOS) ; $V_{GS} = V_{DS} = V_{DD}$ (PMOS)

I_{OFF} and I_{ON} can be calculated as follow:

$$I_{ON} = \frac{w}{L} \mu C_{ox} \frac{(V_{GS} - V_{Th})^2}{2}, \text{ where } C_{ox} = \frac{\epsilon_{ox}}{T_{ox}}$$

Equation 1

W: Channel width [μm]
L: Channel length [μm]
 μ : minor carrier's mobility
 T_{ox} : Dielectric thickness [μm]
 ϵ_{ox} : Permittivity of dielectric
 C_{ox} : Dielectric capacity
 V_G : Gate voltage [V]
 V_{Th} : Threshold voltage [V]

These current parameters, I_{ON} and I_{OFF} , then are usually applied as a good performance indicator of such a device. I_{OFF} indicates the leakage level of the transistor, thus the static consumption. Besides, I_{ON} reveals in the operational velocity and speed. Thus a high I_{ON}/I_{OFF} ratio is better for a device. Overall, an ideal transistor needs to have an I_{OFF} as low as possible, an I_{ON} as pronounced as possible with a more abrupt possible transition between them.

The objective of the mainstream production is always making compromises, so a real end product cannot meet these “ideal” criteria. Depending on the final functional application, the priority of the operational velocity and power can be less determinative. Thus the following groups can be distinguished related to the operation specifications:

- HP = High Performance transistors : high speed, but with relatively high leakage current ; high I_{OFF} and I_{ON} (e.g. microprocessors)
- LP = Low Power transistor : less power consumption , but higher saturation current ; low I_{OFF} and I_{ON} (cell phones, media players)
- GP = General purpose transistors: these types of transistors are a compromise in between high speed and less consumption.

1.1.2 The CMOS Integrated circuit

A complex CMOS Integrated Circuit includes a large scale of NMOS and PMOS transistor couples built as blocks on a wafer. The purpose is the realization of logical operations (based on OR, XOR, AND, NAND... circuits connection solutions (Babaud, 2010)). From process point of view, an IC block can be divided into the two main parts, which are the following (figure 4):

- FEOL (Front End Of Line) – situated closer to the wafer surface and consisting of all the main layers defining the transistor, also called active part
- BEOL (Back End Of Line) – including all the metal interconnection layers that enable to supply or to control the active part of the circuit

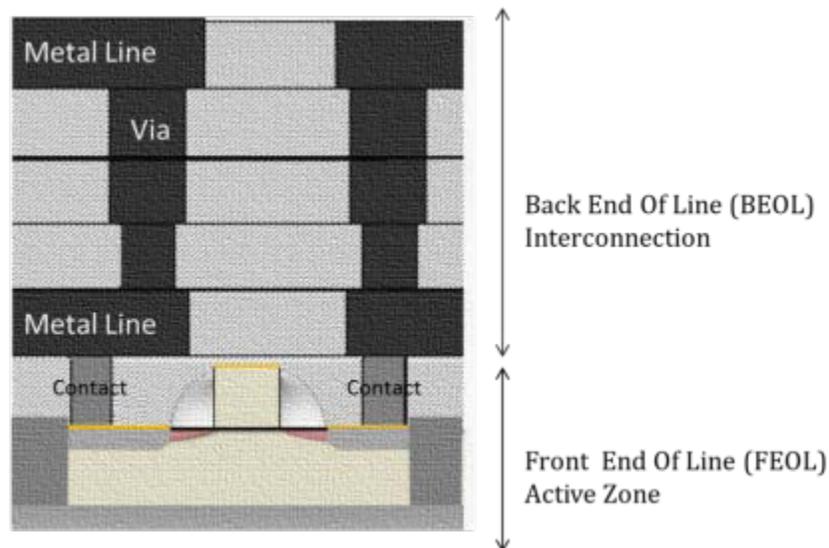


Figure 4 Schema of an IC block illustrating all the superposed layers (including contacts, metal lines, and vias)

1.1.3 From data sheet to IC

Due to these complex requirements for the fabrication of the transistors, the process in integrity needs to be well composed from the very first step following the predefined need. The first stage of this sequence is the definition of the electric schema so called “design” followed by the data sheet or also called specifications. The design will meet with the final application need on a chip, the rectangular area containing the complete IC design. It will also define the transistor’s exact positioning within. Once the electrical pre-simulations are accomplished the complete design is transliterated on the substrate wafer. It is enabled by a given number of insulator, metal and semiconductor layers. Each of these layers is represented on a design level by corresponding geometrical polygons. In order to maintain the operational requirements, design rules are applied focusing on these polygons. Basically it means a series of parameters that characterize certain geometric and connectivity restrictions ensuring the good functionality of the designed device (See Figure 5). Besides it has to ensure an adequate margin of the process variability so that each component works well.

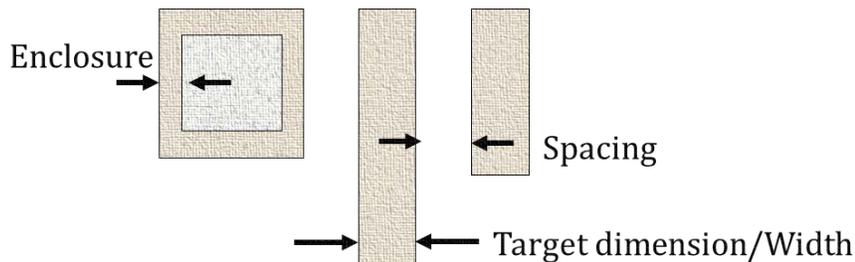


Figure 5 Simplified illustration of design rule

Note that each technology node has its fixed design rules, specific to each layer. It is important to accomplish these rules as they are the bedrock of the transistor’s proper functioning.

1.1.4 The miniaturization and its impact

From the early development stage on transistors until present-day innovation, the main barrier imposed to semiconductor engineers is the dimension shrinking, the miniaturization. Smaller is better in the sense that the major goal is to obtain higher transistor density on raw silicon wafer maintaining the advantageous properties (or even better performance regarding the speed and the consumption as example) in addition to the application and functional implication. Figure 6 illustrates the problematic of shrinking dimensions: it shows how a shrink of the dimensions by a α factor implies the characteristics of a transistor (Colidre, 2014).

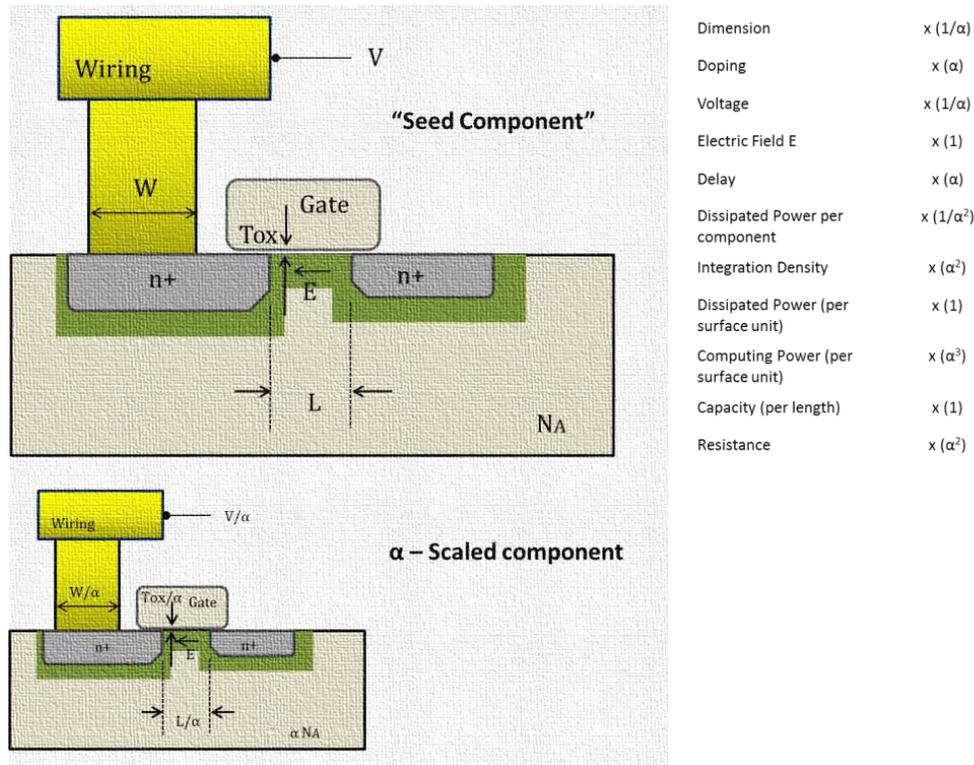


Figure 6 Illustration of the dimension scaling “smaller is the better law” (Colidre, 2014) indicating the parameter changing in case of a shrink by α factor

The maintenance of the proper functioning at this high complexity level requires a well composed integration and information manipulation. It includes a large scale of electronic components; therefore the compilation has to be without any degradation. This needs to settle on highly reliable transistors (as example). It has to be mentioned that by this apace miniaturization of MOS transistors, new physical barriers and new parasitic effects appeared. These effects might degrade the component’s functioning such as the short channel effects, the depletion of the grid, parasitic resistances or other gate control loss related phenomena in the channel (Colidre, 2014).

1.1.5 The process fabrication sequence

In order to follow the above mentioned, steadily decreasing dimension tendency, all the machinery behind the scene has to ensure a well-controlled, robust process. The current IC process technology consists on numerous process steps applied on a wafer. The worldwide applied semiconductor material for the substrate is the silicon monocrystalline. Indeed new solutions are in progress. As industrial consideration and knowing-how, the size of the wafer is increasing in parallel with the chip size shrinking. It means that the wafer dimension became six times larger than the first generation product from 1970’s. Present-day factories capability reaches the 300 nm diameter, but the 450 mm is already under thinking. It is important from the processed chip number point of view, thus from the productivity effectiveness.

A fully processed chip requires several hundreds of process steps synchronized by the process integration – optimization of process step sequence combination in an economic way. The Figure 7 illustrates the simplified schema arriving to an operational transistor.

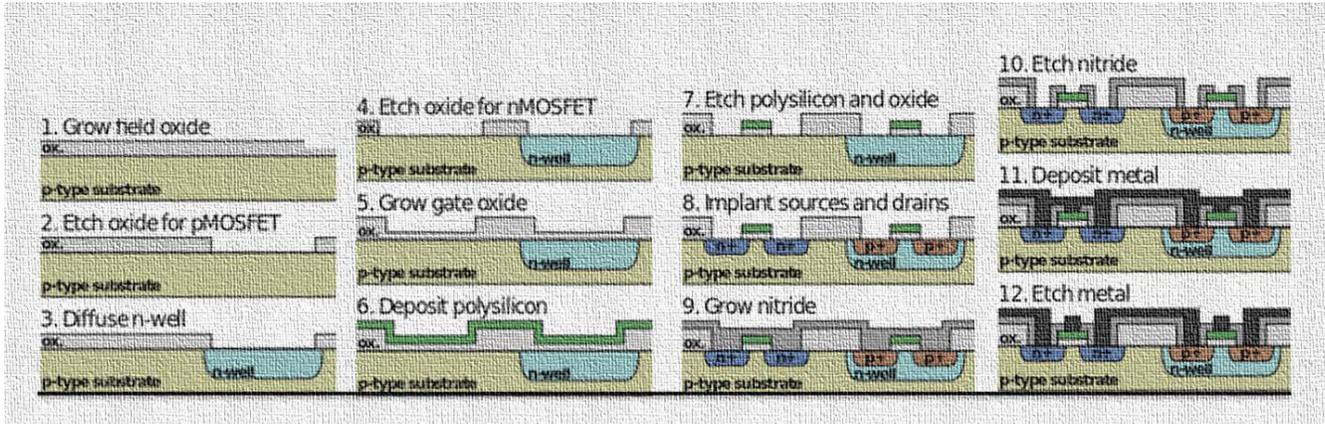


Figure 7 Schema representing the main process step

Starting from a single crystal silicon wafer it is pattern using repetitive process steps such as lithographic, deposition, diffusion and etching steps as far as a functional IC is built. A set of about thirty separate wafer processing cycles, forming modules, are accomplished. Obviously the number of these cycles depends on the technology complexity. The figure 8 shows a simplified IC fabrication cycles flow chart including the following sequences/stages:

- Layer forming consisting of a film/layer deposition on top of the material:
 - Deposition of a material on top of another material;
 - Epitaxial growing of the crystal of a material on another material;
 - Metallization ;
- Photo mask pattern transfer:
 - Photolithography : it consists on the designed photomask pattern image transfer on the photoresist coated wafer;
 - Etching : it consists on etching out the undesired resist and consists of transferring the designed pattern image (negative) into the substrate;
- Modification of the layer
 - Ion implantation/Doping: it consists on the chemical introduction of an element into the material;
- Chemical Mechanical Polishing (CMP): it consists on flattening the surface of the material;
- Cleaning consisting on the removal of residual material;
- Electrical test/Optical inspection: each process steps are assisted by metrology for verification purpose as follow:
 - In-line characterization by Scanning Electron Microscope (SEM)
 - Defectivity analysis
 - Electrical tests

The fabrication flow is represented by a simplified process sequence flow chart on figure 8.

It indicates the entering wafer (silicon substrate) in the fab that will go through the shown sequence's stages until it arrives after inspection.

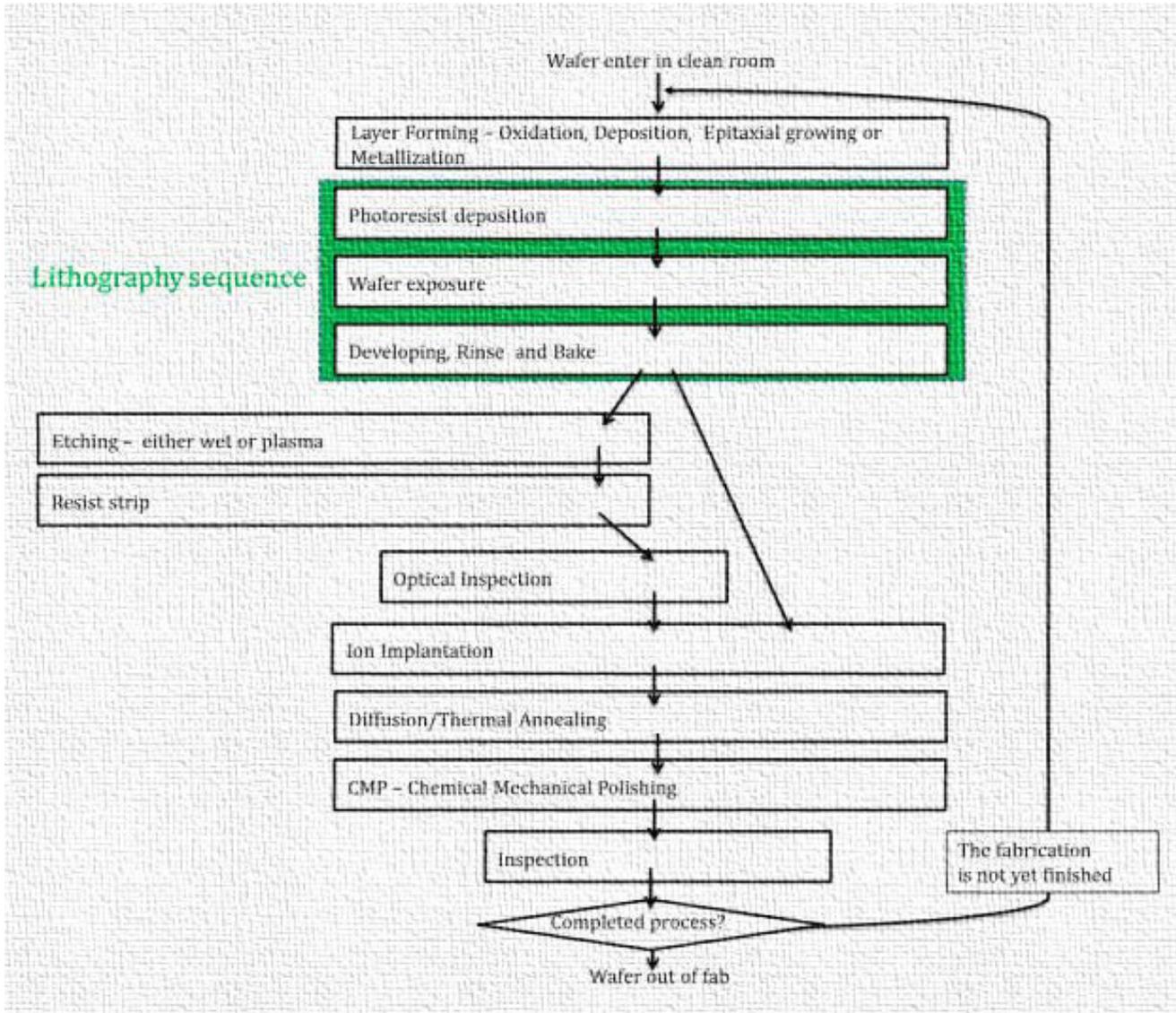


Figure 8 Simplified process sequence flow chart of IC fabrication

Within this flow chart, the lithography can be considered as the heart of the process. It is due to the fact that the lithography step defines the dimension of the fabricated patterns. The thesis work is situated within this critical domain of photolithography. Note that a variety of possible patterning technologies exist, but the best compromise in term of throughput and favorable defectivity in semiconductor manufacturing is still the photolithography. Other patterning solutions are not discussed here.

1.2 Fundamental concept of photolithography

The photolithography is defined by the creation of patterns within a layer of photoresist that covers the silicon wafer. The patterns are then transferred into the substrate by etching or ion implantation.

Thereby the precision and the control of this process step is a major task following the shrinking dimensions. The capability scale of the printing can be measured from a few microns to submicron's (current state around 32 nm) by applying a wavelength (λ) of 436 nm, 365 nm, 248 nm or 193 nm. The current critical dimension (CD) that can be realized by DUV (Deep Ultra Violet) projection with 193nm wavelength is 32 nm that follows the lithography requirements from the International Technology Roadmap for Semiconductors (ITRS). Note that basically the ITRS is designated to ensure an economic advancement in the performance of integrated circuits and the product that apply them (ITRS, 2014). Within this section the basic conception of the photolithography itself and its machinery will be described.

1.2.1 The principle of patterning (photolithography process)

The photolithography process, or so called the patterning, operational principle is based on the resist aptness to store an image of the pattern to be replicated/printed. The basic operation behind such an apparatus are shown on Figure 9 indicating the system elements.

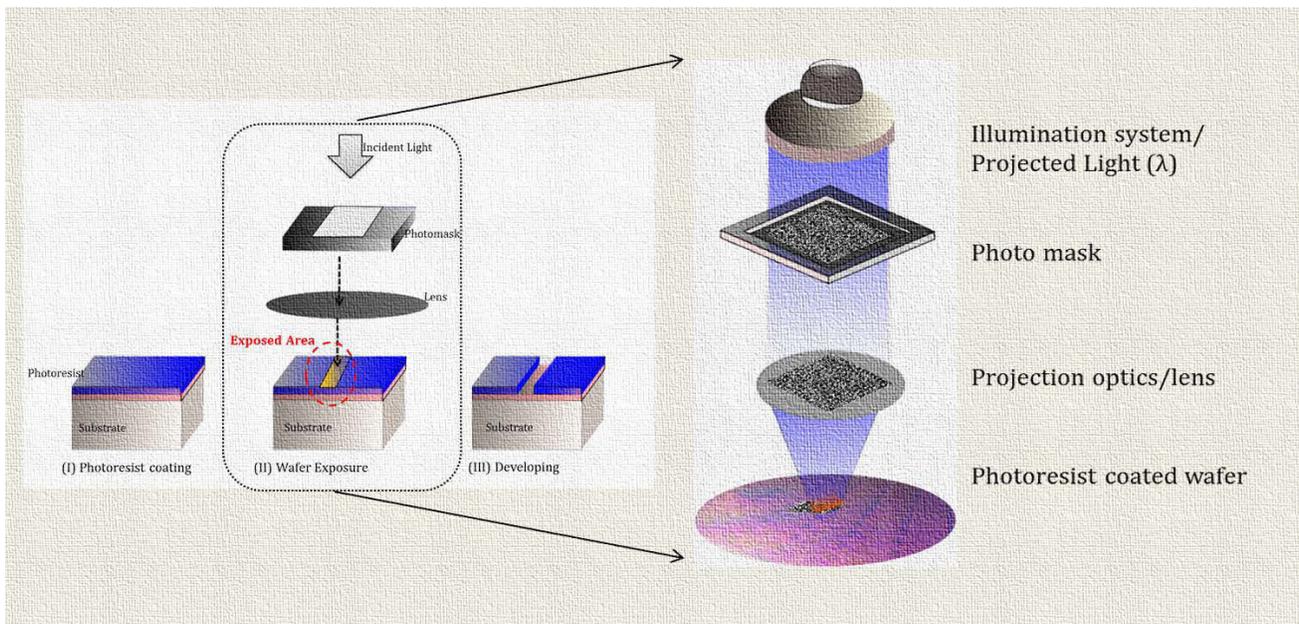


Figure 9 Principle of optical lithography - Note: figure not in scale

As initial point, the illumination system is represented, which main function is to modify the resist dissolution rate. Then the light will reach the Photomask that is composed of a large set of patterns. The light is basically transmitted only through the transparent region. As the mask is partially transparent, it creates an indirect optical source. Its objective is to modify locally the

energy into the photoresist. Therefore the layout pattern is then replicated onto the photoresist coated wafer. The optical system, composed by a set of lens, enables the image transfer by reducing the mask image. The resist coated wafer is placed in the image plane in order to be exposed by the incoming light from the optical system. In general the intensity of the light impinging on top of the resist is named the *aerial image*. Thus the pattern now is stored as latent images² within until the wafer is put into a developer solution. During the development, the exposed areas, depending on the resist polarity, remain or dissolve, completing the lithography process. Finally the opening areas in the resist are formed and are ready for subsequent processing: etching or implantation.

1.2.2 The Photolithography system elements

Under the term of photolithography, a projection printing tool is referred to a scanner system or a step-and-repeat (reduction stepper) system. The *scanning projection printing systems*, the earlier machinery, use reflective optics (mirrors) to transmit the light from the photomask onto the photoresist coated wafer, while the wafers are moved in the same moment by the slit. This scanner's magnification is $M=1$. The *step-and-repeat* (so called stepper / figure 10) systems expose the coated wafer with a rectangular area, an image field, at a time. The magnification can vary typically from $M=1$ to $M=1/5$. Reflective systems/lenses are employed within these complex machineries.

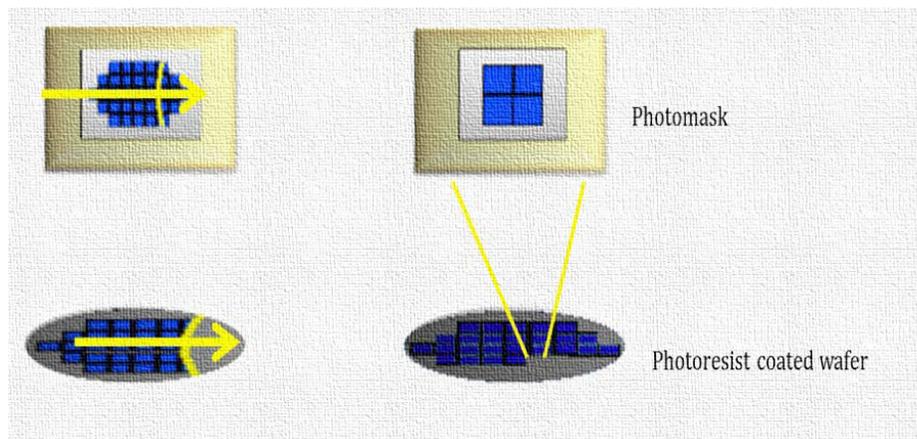


Figure 10 Illustration of (left) a scanner and of (right) a step-and-repeat concept

Hereafter hybrid approaches appeared such the step-and-scan that became dominant. It uses a fraction of a standard stepper field then it scans the field in one direction. These tools are designed to meet a high volume production with a high resolution capability, thus their application is frequent in fab. Note that in the framework of this thesis work, a step-and-scan tool was put into focus (Twinscan®)

² The latent image is the reproduction of the aerial image in the resist as a spatial variation of the chemical compounds

1.2.2.1 The subsystems of a photolithographic tool

The sub-systems of a projection printing system (figure 11) comprise:

- An illumination system;
- A photomask;
- A projection optics/lens.

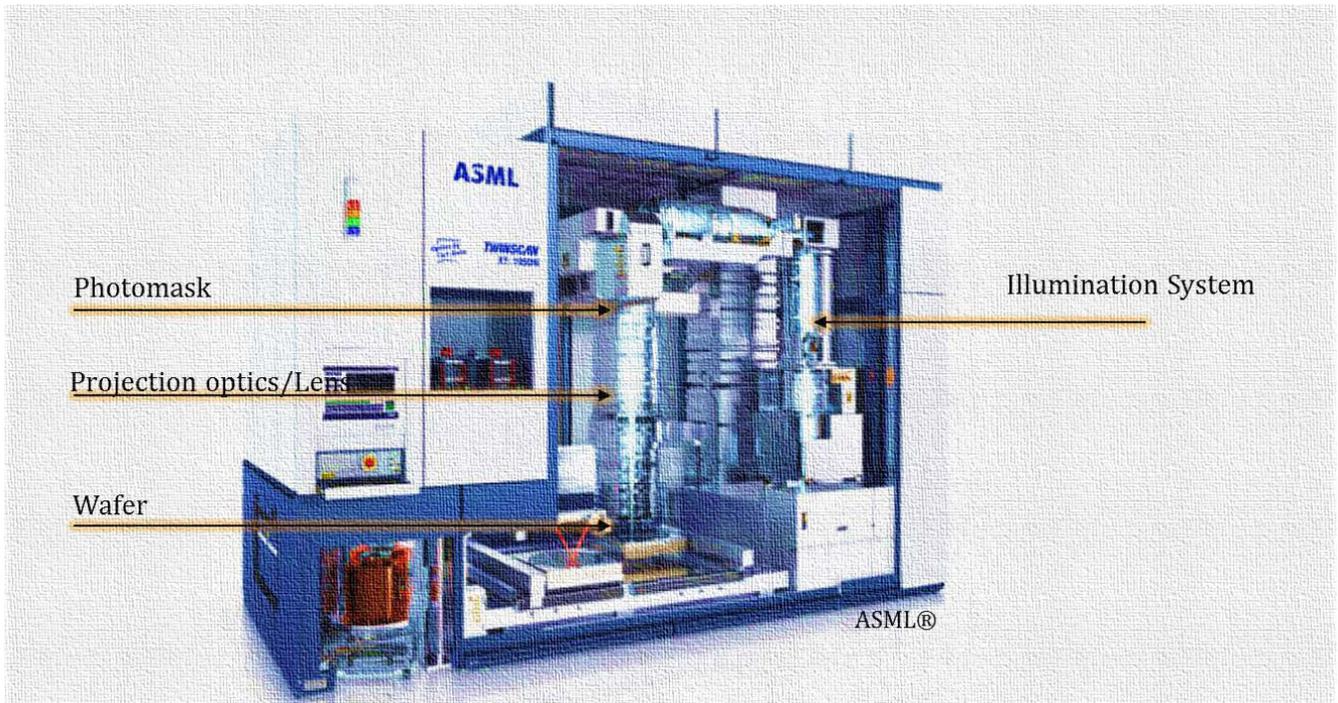


Figure 11 a projection printing system indicating the location of each sub-system that will be described later in this chapter 1

Figure 12 shows a simplified schema of the imaging theory behind such a tool. The image formation is referred to Abbé imaging theory (1872): the plane waves from a light source are diffracted by the reticle (object). Then they are recombined by the projection lens. The lens focuses the incident light from infinity to the Focal point. The diffraction pattern functions as a light source and it propagates to the image plane where it is formed (Cambridge, 2014).

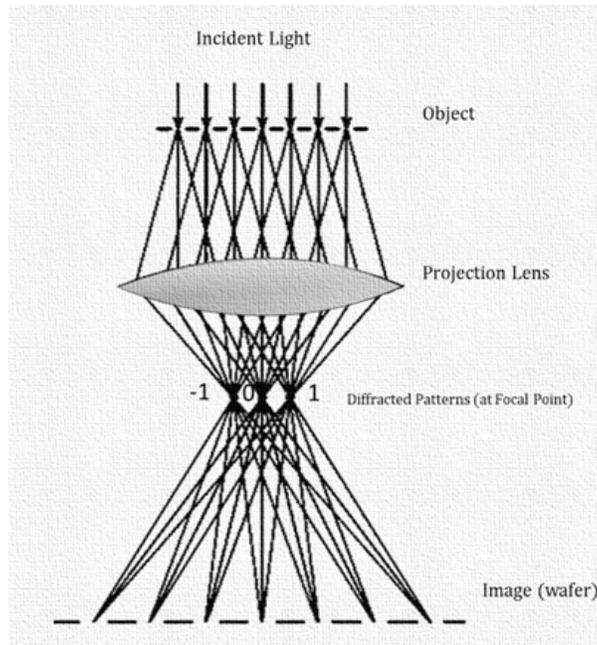


Figure 12: Representation of the image formation based on the Abbé theory

1.2.2.1.1 The illumination system

The first sub system of a lithographic tool, from incident light point view is the illumination system (see Figure 11). It is configured to supply a projection beam of radiation that consists of the combination of light source and condenser lens.

1.2.2.1.1.1 The light source

For the *light source* in lithography equipment, monochromatic³ light is used. The coupled wavelengths (λ) are ranging within the ultraviolet (UV), deep-ultraviolet (DUV) and extreme ultraviolet domain (EUV). The decrease of the exposure wavelength is needed to follow the shrinking device dimensions demands. Therefore the supply of the projection beam of radiation with a mitigated λ needs a light source that can maintain the robustness for this change (EUV supply systems are not at this point yet). As current state, in mainstream manufacturing and development, the more stable embodiment is the excimer laser light source (Figure 13). The excimer lasers are pulsed discharge gas lasers providing light output in the UV region of the spectrum. The term of “laser” is an acronym for: Light Amplification by Stimulated Emission of Radiation (Stanley Wolf, 2004). While the “excimer” is referred to the excited dimer⁴, molecules that exist only in this excited state.

³ Monochromaticity: \sim refers to the single frequency / wavelength property of the radiation. A monochromatic laser radiation has its origin in the stimulated emission process that supports the laser to emit light.

⁴ Dimer : a molecule with two identical or similar parts

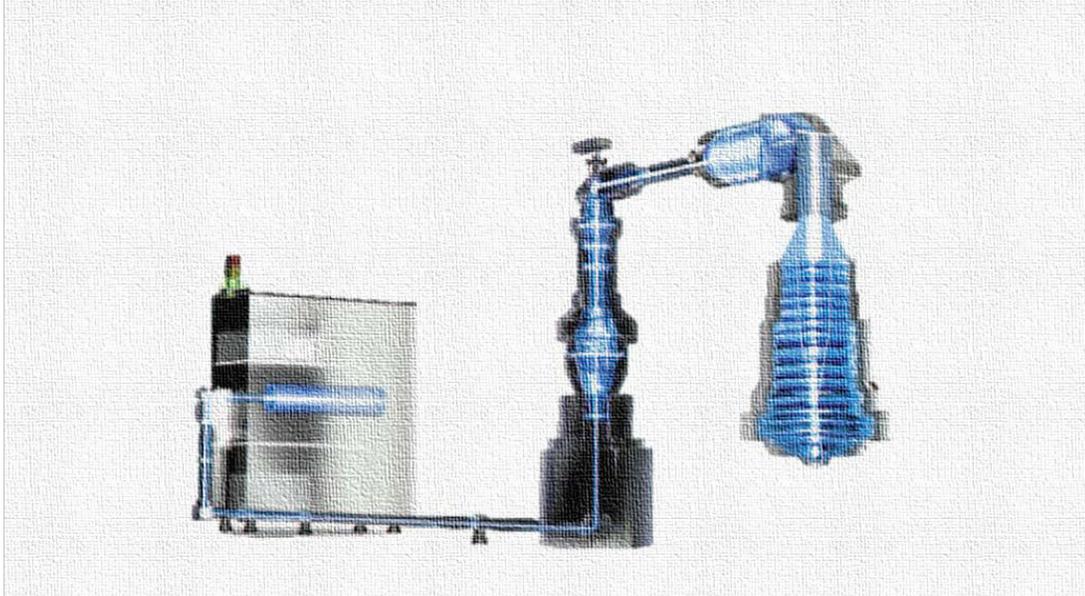


Figure 13 Light source module schematic view showing the excimer laser (the “invisible” part of a scanner built under the tool in the basement)

The excimer laser applies a combination of a noble gas and a reactive gas. Under specific conditions of high pressure and electrical stimulation an excimer is created, that exists only in this excited state. The excited state is a short period of time and is not stable. The forming excimer molecule corresponds to optical amplification needed to initiate a laser action (Maini, 2013). Note that in most current applications, the Argon Fluoride (ArF) excimer laser is employed to emit $\lambda=193\text{nm}$ DUV light.

In order to enable and to ensure an adequate lighting uniformity on the photomask, the most common applied illumination systems is configured with Köhler illumination using condenser (see Figure 14).

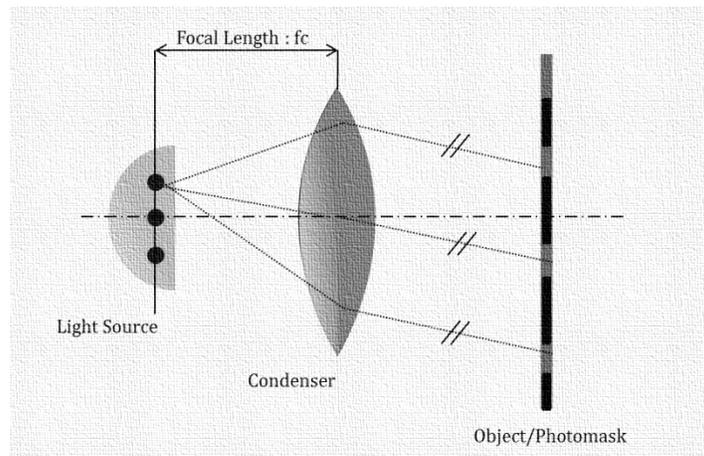


Figure 14 : Köhler illumination system providing a plane wave illumination of the photomask

The implied method means that all the subsystems are positioned in a way that the light source coincides with the focal plane of the converging lens, the condenser⁵. Thus this method has an important property: the rays from each source point illuminate the photomask as a parallel beam. Therefore the reticle benefits of the same amount of lighting or illumination energy. Further, this arrangement enables that irregularities in the source brightness distribution will not result in any disadvantageous irregularities in the intensity of the object (mask) illumination.

1.2.2.1.1.2 The coherence or the fill factor of the illumination

In addition to this adequate across-field dose uniformity, the directional uniformity has to be insured so that the same patterns are replicated identically. The illumination is characterized by the coherence⁶, or fill factor has two important aspects. First, due to the assumption of the light source monochromaticity, in the time domain the light is considered strongly coherent. In spite of that, in the spatial domain, the coherence becomes variable (measure of the physical extent of the light source, its radius) having a pronounced influence capability on imaging performance. The spatial coherence in projection apparatus is defined as the light phase relationship as it propagates from the illuminator system towards the photomask. For tight numerical aperture (NA) in the illuminator system, no angular distribution to the illumination, there is no distribution of the light phase. Thus this can be considered as spatially coherent or a point source. In spite of that increased numerical aperture in the illuminator lead to increased angular distribution of the illumination towards the photomask. Then the illumination is called partial coherent. The aspect of “partial” is referring to occurrence of a certain amount of spatial coherence. In the limit of an infinite (or maximum NA), where all the possible illumination angles are present at mask level, the imaging is spatial incoherent. Note that the NA reveals in the ability of the lens to collect light.

As shown on Figure 15, the level of the spatial coherence is quantified by the ratio of the illuminator system NA, the condenser NA_C to the projection lens NA_P . It is called partial coherence factor or sigma (σ) and is calculated as following (Eq2):

$$\sigma = \frac{n \sin\alpha}{n \sin\theta} = \frac{NA_C}{NA_P}$$

Equation 2

The magnitude of the partial coherence varies between 0 and 1 as follow:

- $\sigma = 0$: spatial coherent illumination (or point source) ;
- $0 < \sigma < 1$: partially coherent illumination ;
- $\sigma = 1$: incoherent illumination.

⁵ Condenser lens: converging lens in the illumination system that consists of a large set of lenses, filters, mirrors and additional optical elements.

⁶ Coherence : the phase relationship of light at two different points in space at any instant in time (Mack C. A., 2006); a fixed relationship between the electric field values at different locations or at different times

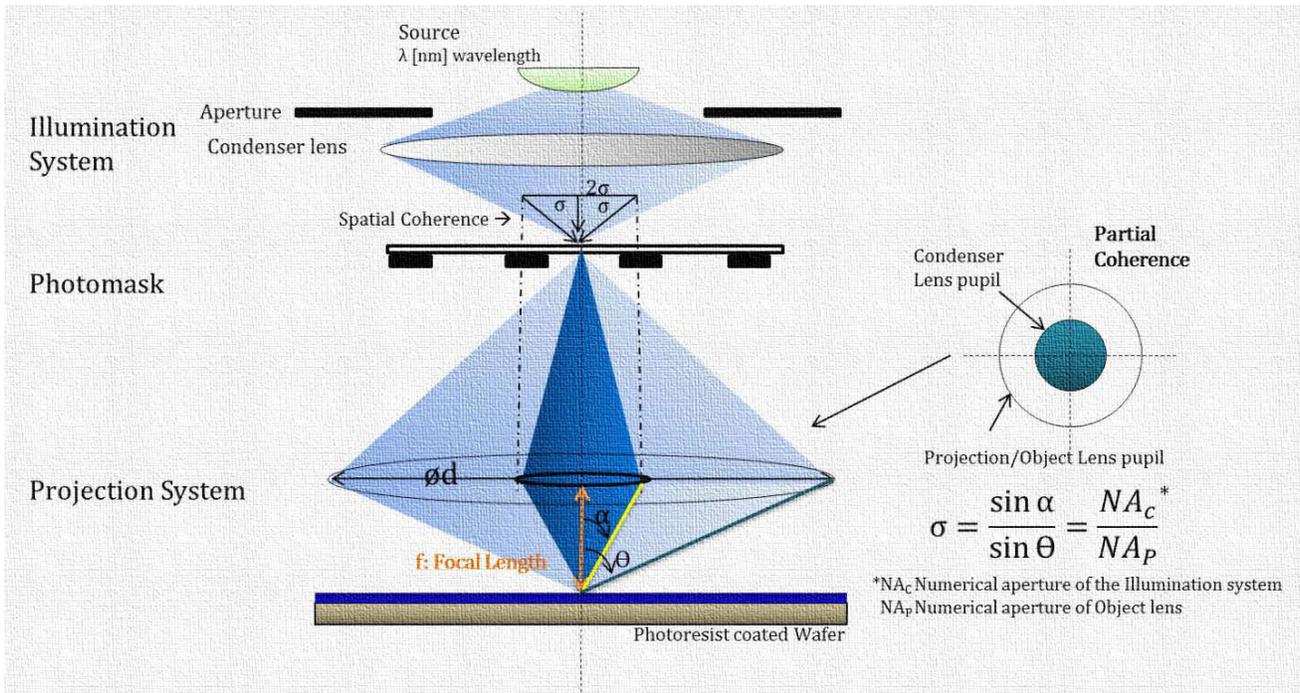


Figure 15 : Projection imaging apparatus simplified schema on optical elements (the Illumination system, the Photomask and the Projection System) with Köhler illumination indicating the image of the condenser in the object lens

Further improvement on the illumination system, in respect to the coherence, can be obtained by the application of an aperture between the light source and the condenser lens. It will function as a wavefront filter. The shapes of the additional aperture are typically circular, annular or quadripolar. A circular element reveals then a low pass filter. Meanwhile an annular or a quadripolar element's main function is to block the vertical or zero-order waves (without carried information). In current optical lithography towards sub-wavelength resolution such modified configurations with off-axis illumination are applied. More detailed description will be given in the context of resolution enhancement techniques within chapter 1.

1.2.2.1.2 The Photomask

The next subsystem of a lithographic apparatus is the photomask on which the patterns are defined. It has an important role, because these patterns are the structures that are replicated. Thus the final product quality can be considered as direct dependent on the photomask. The mask quality is mainly based on the mask set components, dimensions, flatness, mask shop precision and the defectivity. Besides, as all elements of the lithography tool, it should correspond to the demand for high-resolution.

As shown on Figure 16, the simplest mask type is the binary (or COG = Chrome on Glass). It consists on a substrate made of an "exposure-wavelength transparent" material. The patterns are etched into the chrome layer whose thickness is a few nanometers that behaves as an absorber at this wavelength. The term of binary comes from the fact that it has a transparent and an absorber region. The central limit of this type of mask is related to the opaque pattern. It means that pattern

with critical dimension lower than the exposure wavelength produces incorrect patterns at wafer level due to interferences and diffraction. Binary masks are usually applicable up to 0.18 μm technology nodes. Advanced binary mask are Opaque MoSi on Glass (OMOG) that result in better CD performance and in higher resolution. These materials are applied in this thesis work.

Beside the standard binary intensity masks, there are large scales of other types of photomask concepts that enhance the optical resolution for shrunk dimension technology nodes. These types include the alternating phase-shift, attenuated phase-shift (PSM) as well as assorted hybrid types. These mask technologies are applied for line width lower than the exposure wavelength. As current state, the mostly implemented concept in addition to the OMOG is the attenuated PSM. Figure 16 represents basics of these mask concepts by illustrating a top and cross-sectional view of each with the corresponding intensity (indicating the resulting Sum Intensity by yellow color).

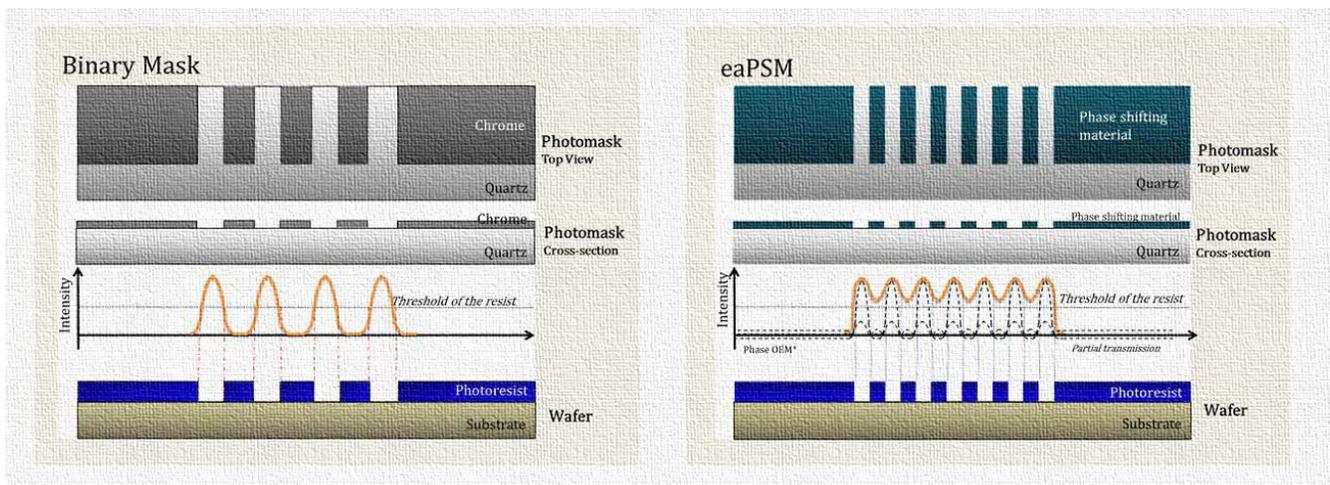


Figure 16: Representation of the Binary vs eaPSM photomask [Toppan] illustrating the comparison of the obtained light intensity at wafer level.

The Attenuated PSM consists on a chrome layer and a semitransparent layer that shifts the phase angle of the light compared to light passing through the opening area of the photomask. The passing light's speed is altering, through the material shifting, its phase angle. Thus this kind of material is so called phase shifting material or semitransparent material. Therefore the phase shifting mask is changing the transmitted light, the intensity distribution and as well as it varies the transmitted light's phase at some region of the mask (Figure 16, Figure 17).

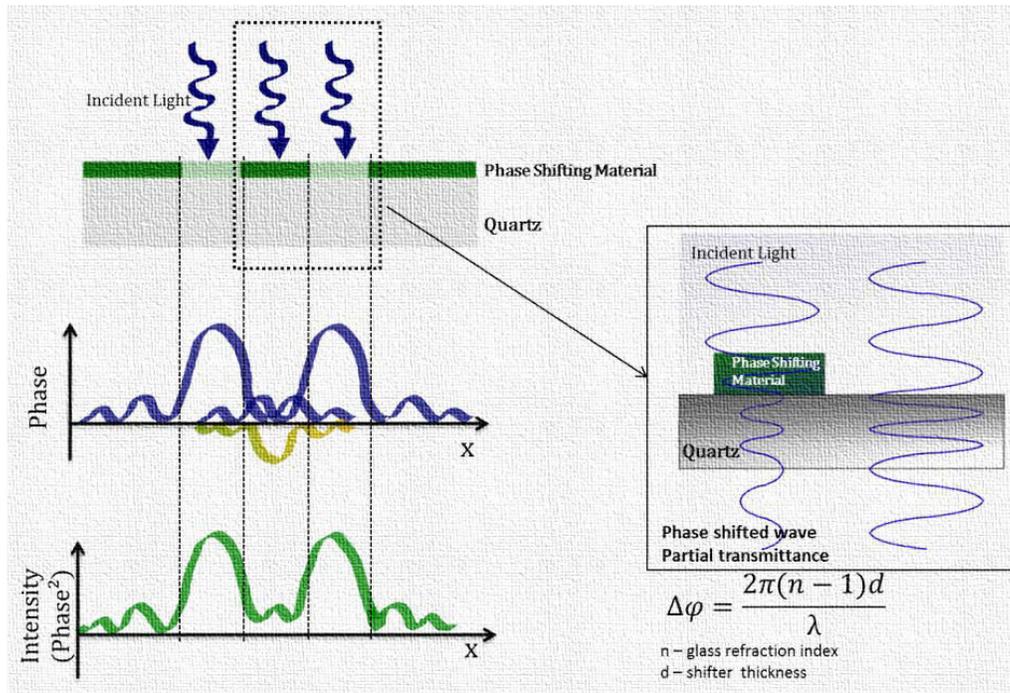


Figure 17: Simplified (not in scale!) schema of the Phase shift through a phase shifting material (Laube, 2014)

The Phase Shifting material thickness is determined to obtain the phase shift of 180 degree. For advanced nodes, the Electromagnetic Field effect needs to be taken into account (Laube, 2014). It requires the mitigation of the thickness that is challenging for PSM masks, as it impacts the phase shifting. The solution for that problematic is either the application of the OMOG (as it is in the case in this thesis work) or the application of an eaPSM using an optimized absorber.

1.2.2.1.3 The projection optics

The projection optics is composed in a way that the lens system enables to collect the incoming light. It means a part of the diffracted orders at mask level are participating to image formation at wafer level. Then it enables to project it further on the wafer (image plane). Thus the reconstitution of the pattern image (from the mask) is realized by the phenomenon of interference of the collected waves in the photoresist. The aerial image designates the light intensity formed at wafer level in the air, or in the immersion liquid in case of immersion photolithography. In general a projection optic system is characterized by the value of its magnification and the NA. As earlier mentioned, the magnification (M) is referring to the reduction factor in between the mask pattern dimension and the obtained pattern dimension on wafer. The numerical aperture defines the maximum angle that makes a light beam diffracted and then captured by the projection optic.

Further improvement is possible to increase the magnitude of the numerical aperture, as it is limited by the lens geometry (diameter, curvature). One of a widely applied key to improve is an immersion liquid inserted in between the exit of the projection optic and the wafer stage (see Figure 18). The inserted liquid enables an increase of the NA: whereas in dry optics $NA = \sin(\alpha)$, in

immersed optic $NA = n \sin(\alpha)$, where n is the refractive index of the liquid. Thus in case of water as immersion liquid it results in an increase factor 1.44.

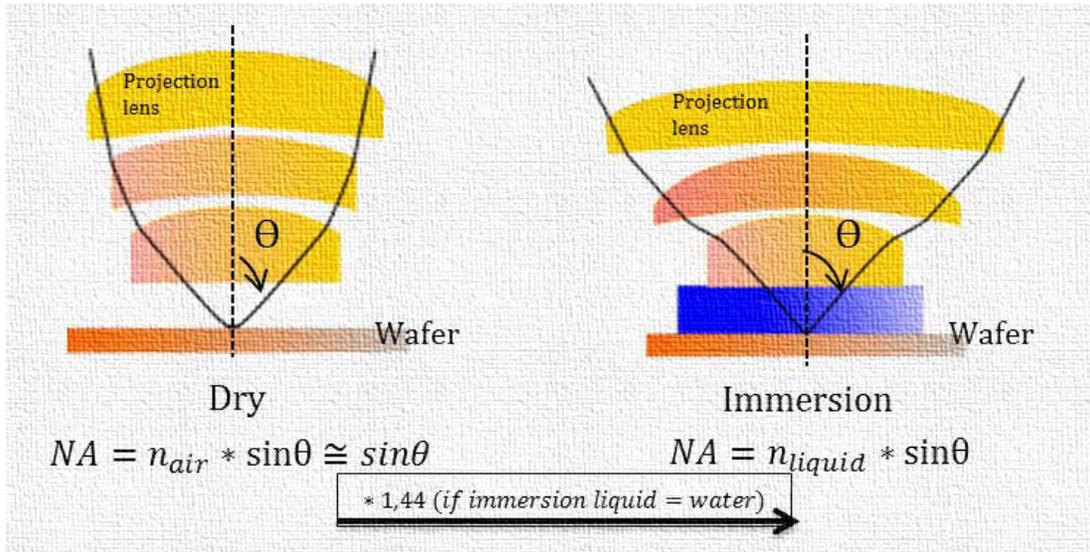


Figure 18 : Simplified schema on dry optics vs. immersion optics

Note that the other aspect of the way of improvement is related to the optical aberration⁷ and the lens quality that will be more detailed in Chapter 5 (FlexWave).

1.2.2.2 The formation of the aerial image

The incident light passing through the mask is diffracted at different directions and angles. The interactions of the diffracted orders captured by the projection pupils form the image of the mask pattern that is projected onto the photoresist. It is so called aerial image. The goal is to create a high-fidelity image compared to the mask pattern.

The diffraction provides the Fourier transform of the mask. The imaging limitation is coming from the fact that the projection lenses have finite size. Even if they are constructed to behave so as they could result in inverse Fourier transform of the diffraction pattern creating the “ideal” image of the mask, this finite size of the NA limits spatially the incoming diffracted orders. Thus only a certain amount of diffracted orders are participating to image formation (see Figure 19).

⁷ Aberration : ~ is any divergence from the ideal « diffraction limited » imaging performance of a given lens (Mack C. A., 2006)

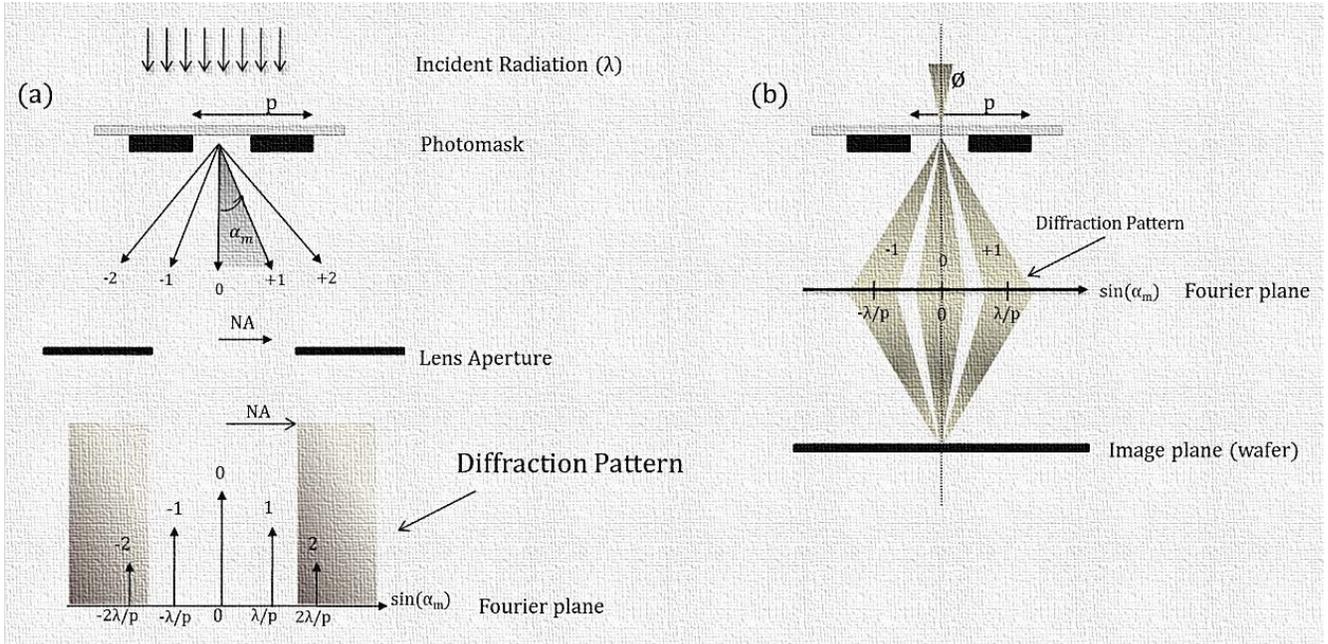


Figure 19 : Simplified schema illustrating the diffracted phenomena on a (periodic) line vs. pitch configured structure: (a) diffracted orders distribution under coherent illumination by indicating the filtering behavior of the exit pupil (b) diffracted orders distribution under partially coherent illumination

Due to the diffraction limited behavior of an optical system, there is loss in term of mask pattern information related to the captured (or non-captured) diffracted orders, as well as there is a relatively high light intensity loss compared to an ideal imaging case. Despite of the degradation, the capture of first diffracted orders is sufficient enough to form an image.

In order to calculate the diffracted light direction, a simple example is taken (see Figure 19 (a)) with a monochromatic coherent radiation that arrives on a periodic line vs. pitch structured mask. The direction $\sin(\alpha_m)$ is defined by the Bragg's law (Landis, 2011) as follow:

$$\sin(\alpha_m) = m * \frac{\lambda}{p} ; m = 0, \pm 1, \pm 2, \dots$$

Equation 3

Where α_m is the diffraction angle, m is an integer indicating the diffraction orders, λ is the wavelength of the exposure and p defines the pitch or period of the structure on mask (see Figure 19). As pattern dimensions and so pattern pitches mitigate, it would impose an increased angle of diffracted orders. It also means that fewer diffracted orders are captured by the imaging system, the optical system. Thus it leads to provide an image with less frequency details, resulting in degradation on the resolution capability. The limitation of the image quality is then defined by a minimum pitch associated with the "feasibility" of at least one interaction in between two diffracted orders.

Let consider a partially coherent illumination with a conventional source (σ) as example. Each scattered, incoming plane wave is referred as a diffracted order. The diffracted orders distribution

behavior for each point is identical to the simple previously described coherent case and calculated as follow:

$$\sin\phi = \sigma * NA$$

Equation 4

Where ϕ is the light beam angle, σ is the radius of the diffracted order at pupil level and NA is the numerical aperture. Identical to the previous case, minimum the two first diffracted orders are required for imaging (image formation) at the image plane, and at wafer level by the phenomena of interference. Based on source behaving properties (Zeggaoui, 2011), an interference occurs between two beams if they are coupled from the same initial point. For each captured diffracted order in the projection lens, each captured point interacts with its counterpart from the identical point source. Therefore the interactions can occur between two diffracted orders (-1/0 or 1/0), as well as between three (-1/0/+1) or more orders in function of the pitch size. Figure 20 shows a simple example for the case of the two first diffracted orders interference for a periodic line vs. pitch structure (p) with a conventional partially coherent source (σ). Each point at the surface of the first order colored with light blue that is captured by the pupil is assisting in the image formation due to the interference. Although the center part belonging to the zero order filled with white color does not participate to imaging.

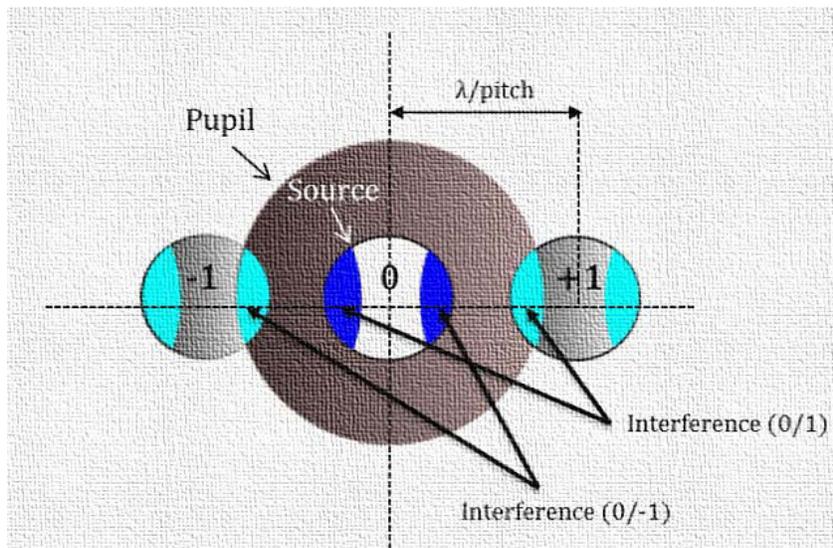


Figure 20 : Captured diffraction orders by the pupil

1.2.2.2.1 The aerial image qualification metric

Note that in function of the pitch p, the wavelength λ ; the partial coherence σ , and the numerical aperture NA, the different interactions are predictable that assists the imaging. The aerial image quality is improved by the number of interactions (more is better), which is quantified by the

image contrast⁸ and expressed as the intensity variation over the total intensity of the intensity distribution as follow:

$$Contrast = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Equation 5

Where I_{max} is the maximum intensity of the aerial image, while I_{min} is the minimum. The more the number of constructive interferences between the diffracted orders, the higher the gap between I_{max} and I_{min} resulting in a better aerial image contrast (see Figure 21).

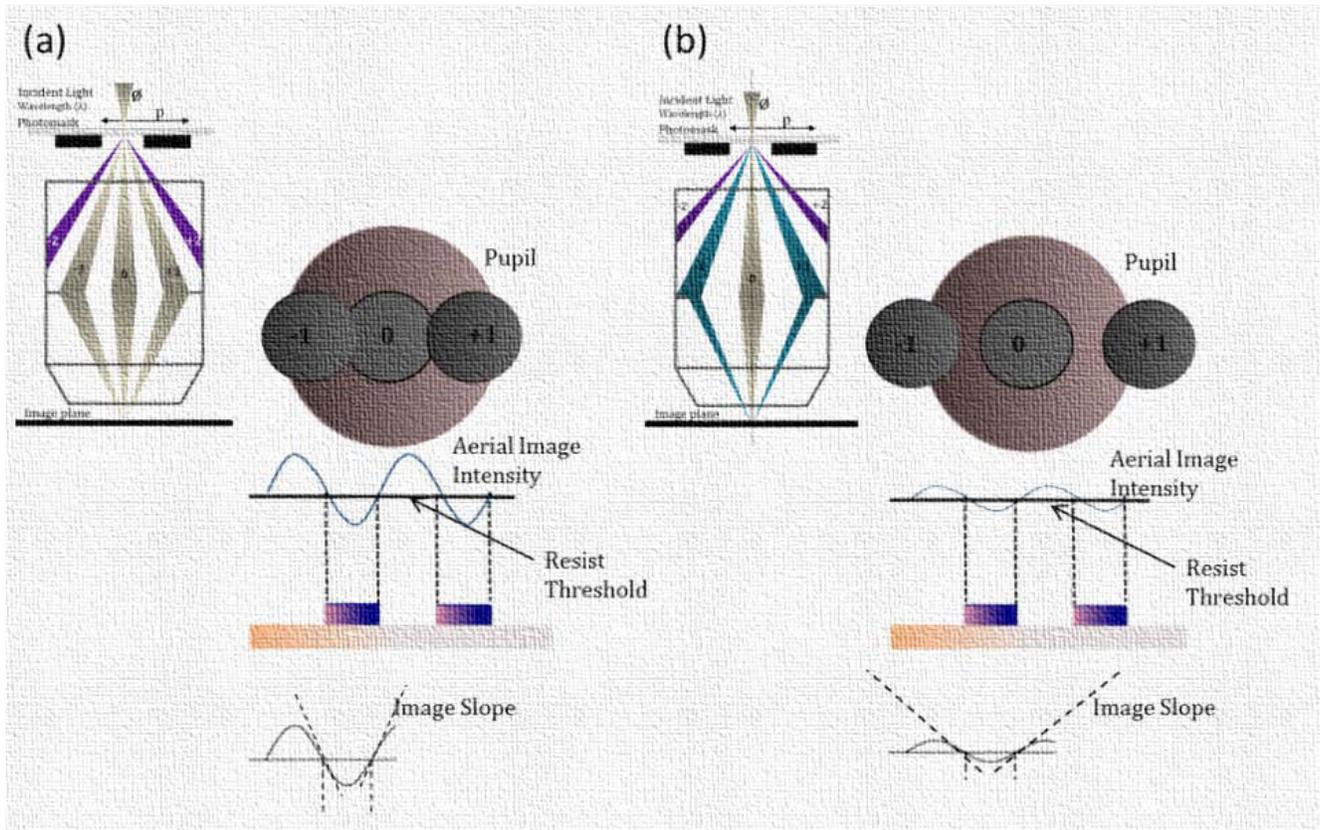


Figure 21 : Illustration of aerial image characterization emphasizing the fact that increased number of interferences corresponding to (a) is related to a better aerial image contrast, thus a higher image slope than for the case (b)

As Equation 5 is only applied for small identical line vs. space structures, another approach can be applied for non-periodic cases. It is referring to the normalized slope of the intensity curve as a function of position dI/dx . This metric is so called Image Log Slope (ILS):

$$ILS = \frac{1}{I} \frac{dI}{dx} = \frac{d \ln(I)}{dx}$$

⁸ The image contrast is also known as fringe visibility of two plane wave that interfering

Equation 6

The ILS is measured at the nominal line edge. The higher the ILS value is, the better the imaging is. Further normalization of the ILS is the NILS (normalized ILS):

$$NILS = w \frac{d \ln(I)}{dx}$$

Equation 7

w is referring to the nominal line width. The NILS is often used metric for aerial image characterization, which is considered sometimes as more informative than the ILS – since it might carry more information on optical parameters that can impact the image quality.

1.2.2.3 The photoresist

Once the aerial image is created by the lithographic apparatus it is transferred further into the photoresist and it propagates into the wafer stack. The photoresist should have a well-defined and sharp energy value at which the resist begins to change, so called the threshold, for a good imaging (see Figure 22).

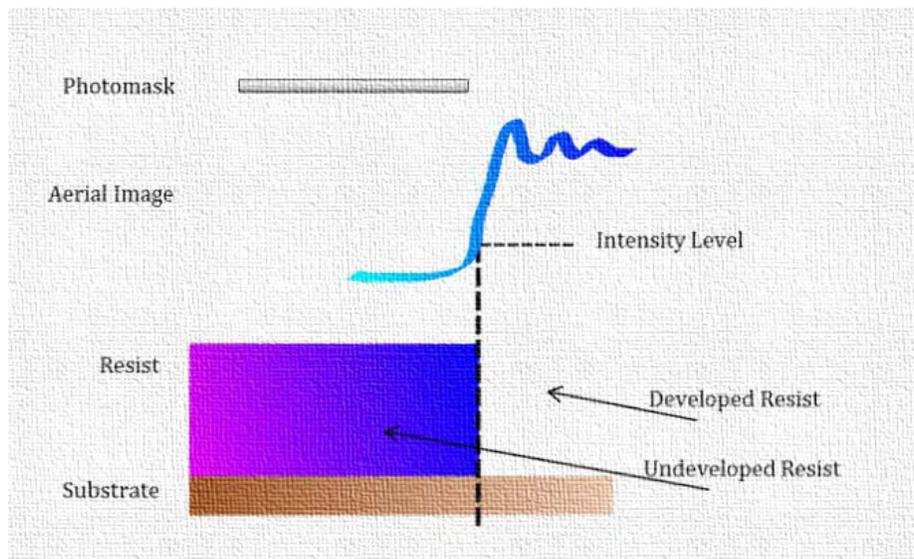


Figure 22 : Simplified schema of aerial image transfer to the photoresist by indicating where the resist will be developed and undeveloped following the threshold

The photoresists, for DUV spectra light emitting source, are usually radiation⁹ sensitive organic polymers. The majority of photoresist applied currently in fab for DUV are multi-component formulations. Chemical amplification resists (CAR) (Landis, 2011) are required for wavelength applied in DUV spectra (or shorter λ) in order to improve the sensitivity to the exposure energy, so as to enable a better absorption.

⁹ The radiation of the exposure tool

This specific chemistry material requires two differentiated chemical reactions in order to change the solubility behavior of the photoresist:

- First : Aerial image to be turned into a latent image by the exposure (without any significant change in solubility)
- Second: diffusion and amplification reactions during the post-exposure bake (PEB) catalyzed by the exposure reaction products (soluble photoresist as result), which is the case for a positive tone resist.

It is realized in a way that this material is in liquid form and is coated by spin-casting technique on the substrate (wafer). Subsequently the coated wafer is driven from the solvent coating by the substrate baking. Thus this step yields in a durable polymer film on the substrate and is referred as the coating process.

In order to ensure the transfer of this image into the underlying substrate, providing high pattern fidelity, the resist performance needs to fulfill several requirements such: uniform resist thickness on substrate and weak (weakest as possible) line width roughness (LWR).

1.2.2.3.1 The Antireflective Coating (ARC)

An important photoresist related aspect needs special attention: the reflectivity. A certain amount of incident light reaches the bottom part of the resist, thus the stack underneath. As the resist is partially reflective, this amount of light is reflected. Therefore interferences occur – constructive or destructive – between the incoming and reflected light. As consequence a periodic intensity distribution is introduced along the resist thickness that alters wavy the photoresist edges (see Figure 23). Thus it provides undulated resist profile after development that can have an important role on contrast¹⁰ degradation so as on pattern dimension.

¹⁰ Resist contrast: \sim corresponds to the remaining photoresist thickness after the development as a function of the exposure energy. The contrast is typically applied for resist characterization purpose.

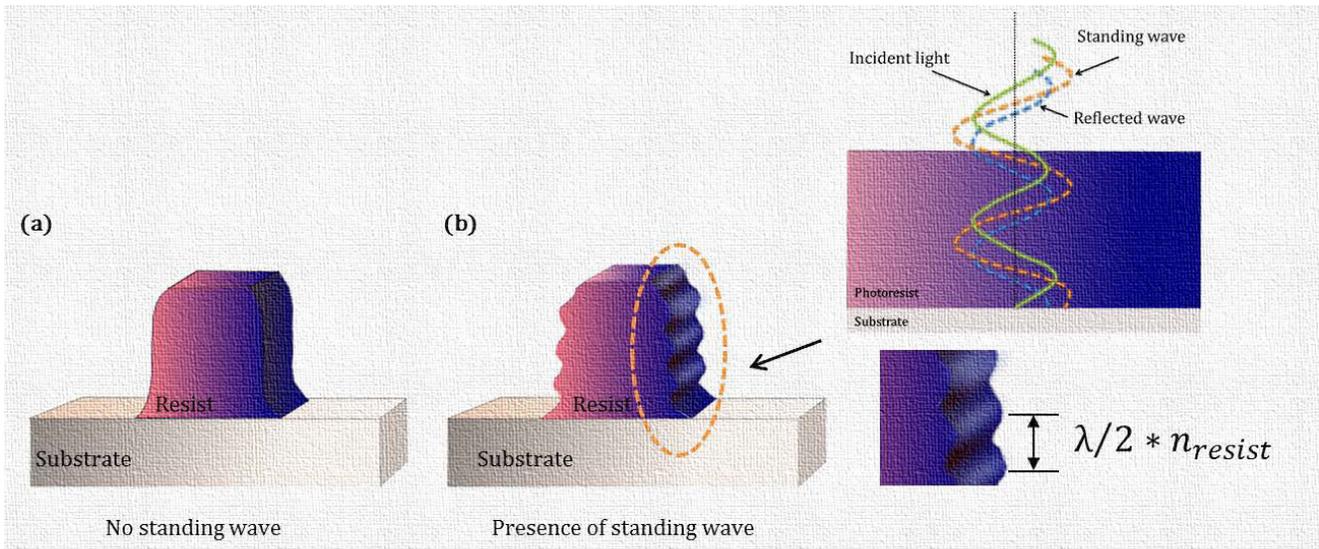


Figure 23: Illustration of a resist line after development that shows (a) no presence of standing wave (b) profile with standing wave

Several methods exist to minimize this swing curve profile effect such as the optimization of Post Exposure Bake (PEB) that induces the diffusion of the exposure generated compounds resulting in a smoothed out profile (due to the photoresist thermal reflow). The other option is the applications of Antireflective Layer Coating (ARC) to reduce the reflection, which can be introduced either on the top of the resist (TARC¹¹) or either underneath (BARC¹²). Usually this coated layer consists of a tenth to several hundreds of nanometer thickness of organic or non-organic material. Current technology applications rely on both solutions in fab.

1.3 The Lithographic Performance Indicators

The performance of photolithography, its constraints and the feasible strategies for further improvement are affected mainly by the following indicators of merit.

1.3.1 The Resolution

The limit of conventional projection optical lithography is defined by the resolution capability with an appropriate process control. The resolution of an optical system might be translated by the notion of the minimum printable feature size:

- If designed dimension on mask > minimum feature size : two separate points will remain separated after the aerial image diffraction into the photoresist;

¹¹ Top Antireflective Coating : introduction of a thin film on top of the resist to reduce reflection from air-photoresist interface and so to reduce undulated curves (caused by thin-film interference effects)

¹² Bottom Antireflective Coating: introduction of an antireflective layer underneath the photoresist to reduce reflection from the substrate (Mack C. A., 2006)

- If designed dimension on mask < minimum feature size: two separate points will be merged after the aerial image diffraction into the photoresist.

Basically it is bounded by the smallest image, which is projected onto the photoresist coated wafer. The other limiter factor is the resolving capability of the photoresist. From machinery/projection image viewpoint, the resolution is (more often used) described by the well-known Rayleigh's criteria:

$$R = k_1 * \frac{\lambda}{NA}$$

Equation 8

Where R , λ , NA represent respectively the resolution, the exposure wavelength and the numerical aperture of the optical system. The parameter k_1 is a process specific parameter determined by lithographic process (illumination conditions, photoresist, development...). The k_1 value varies between 0.2 and 0.6 in lithography – the practical lower value is approximately 0.25. Note that the parameter NA reveals in the ability of the lens to gather light. It is a critical parameter in respect of the resolution as it has impact on the primary system property.

Based on Rayleigh's criteria, several solutions exist to make improvements, by means the mitigation of the R value. These solutions are related to mitigate k_1 factor, to decrease the exposure wavelength λ or to increase the NA :

- The k_1 factor is a complex element of several variables in the photolithography. Its mitigation can be realized by the application of enhanced photoresist (with better solvents) and the application of resolution enhancement techniques such the off-axis illumination (OAI), phase shifting mask (PSM) or optical proximity correction (OPC)(more in detail in next chapter).
- The mitigation of the exposure wavelength λ needed the less investigation, thus it was realized first by passing from first 435 nm to 248 nm, then now to 193nm¹³. For different reasons, the usable wavelength for industry is currently fixed at 193nm.
- The amelioration of the NA needs a high complexity investigation for the projection lens, where the dimension of the lens diameter is increasing.

1.3.2 The concept of the Process Window (PW)

In addition to make advance in the mitigation of the R value, the process needs to be ensured, providing a proper functioning. Therefor the stability of certain parameters should be well controlled such the ARC thickness, the photoresist thickness, substrate topology, focal plane or scanner defocus. Overall an adequate control of the defocus fluctuation and the dose can enable a quite stable process functioning. The defocus is referring to the distance measured along the optical axis between the wafer position and the position of best focus. The dose is referring to the

¹³ Next generation lithography solutions are challenged to go further with the decrease of the wavelength, which is currently only a concept, not yet reliable – such as 13.5nm (EUV lithography).

exposure energy, the amount of energy per unit area that is the light intensity multiplied by the exposure time.

For this analytical purpose, the Bossung curves¹⁴ are the most usually applied. The Bossung curves plot the line width as a function of exposure dose and defocus for a given feature (see Figure 24.) The shape of each curve, belonging to different exposure-dose, contains information on the process window and the possible process issues. The changing CD through defocus defines the feature sensibility to process variation. Usually the dense lines are “smiling” while the isolated lines are “frown” through focus. The process window is determined among the Bossung curve shape at isofocal dose. If these curves tend to have less curvature or are almost flat, it means that the feature is highly sensible to the focal plane and called Isofocal Dose. The dose variation sensitivity reveals in spaced curves.

Further information is carried in parallel by these curves as the depth of focus (DOF) or energy Latitude (EL). The Depth of focus corresponds to the maximum range of focus around the image plane, in which the printed image is kept within a variety of specifications. These specification factors can be as follow: side wall angle, resist top loss, line width... Thus the depth of focus is a tolerance criterion of the focus, where the tolerance is a fixed value upon the printed CD. In general it is ±10% of the nominal CD. Another important parameter is the usable DOF which is the DOF available for a set of different patterns.

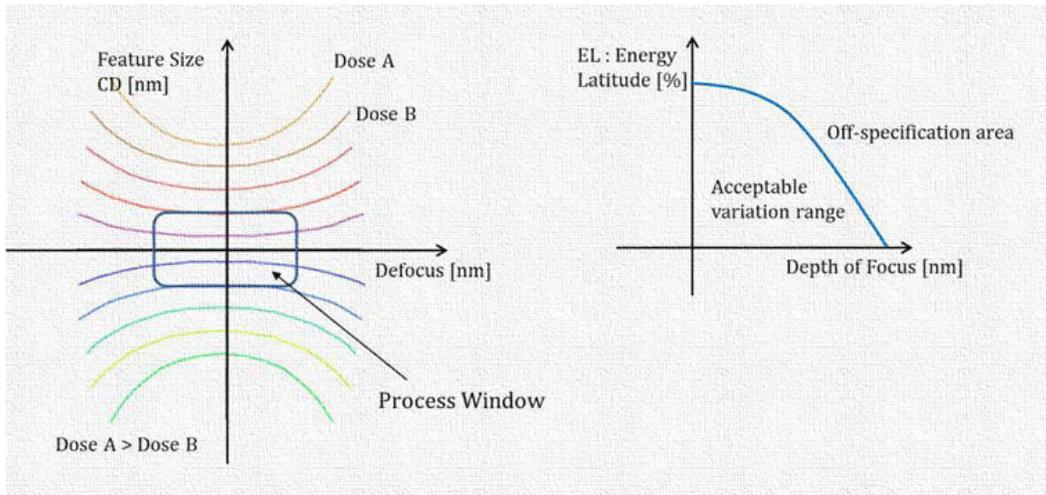


Figure 24 : Bossung curve illustration indicating the Process Window with the Energy Latitude Curve reflecting on the acceptable variation area (whereat the CD is within the tolerance value)

In practice it is based on Bossung analysis. The simplified way to express is as follow:

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

Equation 9

¹⁴ Bossung curves : named after John Bossung referring to the focus exposure matrix

Where k_2 is a process specific factor (likewise k_1), λ is the wavelength and NA is the pupil numerical aperture. Figure 25 illustrates how the change of NA impacts the depth of focus, which supports the resolution as well.

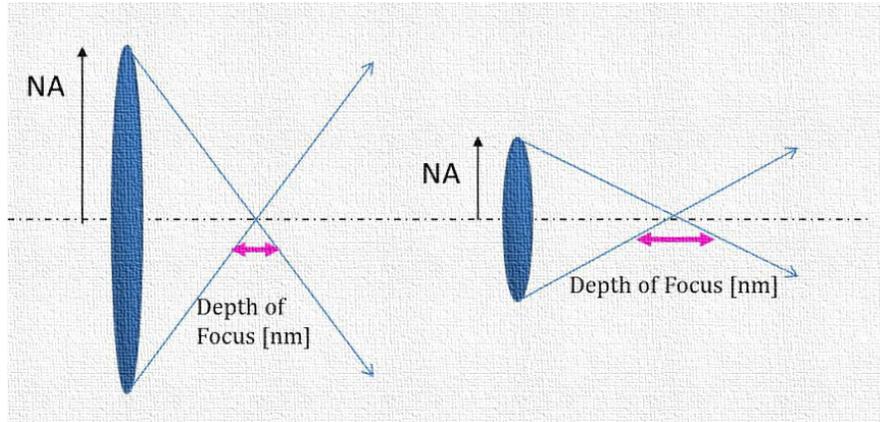


Figure 25 : Depth of focus variation for different NA

The other significant information from the Bossung analysis is related to the dose and so called Energy Latitude. The energy latitude is a tolerance criterion of the dose. The proper functioning of the process requires a competent control of the dose in order to support a possible issue not only coming from the lithographic aperture, but also from the substrate topology variation, from non-uniform wafer exposure, from photomask issue or emitting light issue. Thus the energy latitude corresponds to the exposure dose fluctuation that results in printed feature CD within the specifications (tolerances) such the case for the depth of focus. In practice it is also defined during the Bossung curve analysis.

Note that there is another parameter, the Mask Error Enhancement Factor (MEEF), which is widely used as a lithographic parameter for CD control. The MEEF represents the sensitivity of a mask level variation. It means that it reveals the impact of a small CD variation at mask level that would impact the printed CD at wafer level. The MEEF is calculated as follow:

$$MEEF = M \frac{\partial CD_{wafer}}{\partial CD_{mask}}$$

Equation 10

M is the magnification factor of the photomask. A value of 1 means a linear imaging of mask pattern to wafer.

1.4 Resolution Enhancement Technology (RET)

In addition to mitigate the wavelength or to improve the NA, a large scale of methods have been (or are under development) developed under the wings of Resolution Enhancement techniques. All that development is motivated aiming a possible amelioration of the resolution for the steadily shrinking features and the increase of the process latitude. Note that a well composed theoretical

discussion is provided on the resolution enhancement technique by Wong [Wong, 1992]. Some of the key techniques related to this study are mentioned here, that are also applied by industry in respect of a large variety of RETs. Several examples were already introduced beforehand such the eaPSM, the application of an immersion liquid between the projection system and the wafer and briefly the off axis illumination.

Thus in order to continue the introduction of the RETs, the emphasis is on how to mitigate the k_1 imaging dependent factor. As earlier mentioned it can be realized by the application of advanced types of photoresist that have a more solvent behavior enabling an improved imaging quality despite the diffraction limits. Another approach to mitigate the k_1 factor is to modify or optimize the illumination adapted by a set of calibration structures. The illumination that is passed through an evaluation among the downscaling: from a conventional source to through off axis to freeform source or to a pixelated source (see Figure 26). The source evaluation is realized in the shadow of the coherency, the partial coherency that is an essential point for lithography performance improvements. The application of a large partial coherency factor (σ_{MAX}) might be an efficient solution to print dense features. Nevertheless a large factor would decrease the process latitude for isolated pattern or for semi-isolated pattern. In addition the contrast might degrade for dense patterns. Thus key point of the application of a non-conventional source, an off axis (OAI) as example, is to eliminate the zero diffracted order surfaces that do not participate in the imaging while it is increasing the contrast of the formed image. Thus it is possible to realize twice as small (mask level) pitches than in a conventional case. At current state the parametric source and the freeform source are promising the more improvement. The parametric source consists of discrete segments described by an σ , an NA, intensity and a position. The freeform might allow a higher level of freedom in term of intensity and position of the radiation. These types of sources are optimized for future advanced nodes including denser and denser features assisted by the Source Optimization or the Source Mask Optimization (ASML®). Note that essentially the selected source for a given technology process is based on the geometry to be printed and the lithography performance indicators. Some design level includes critical pattern orientations thus it is an additional limitation or criteria.

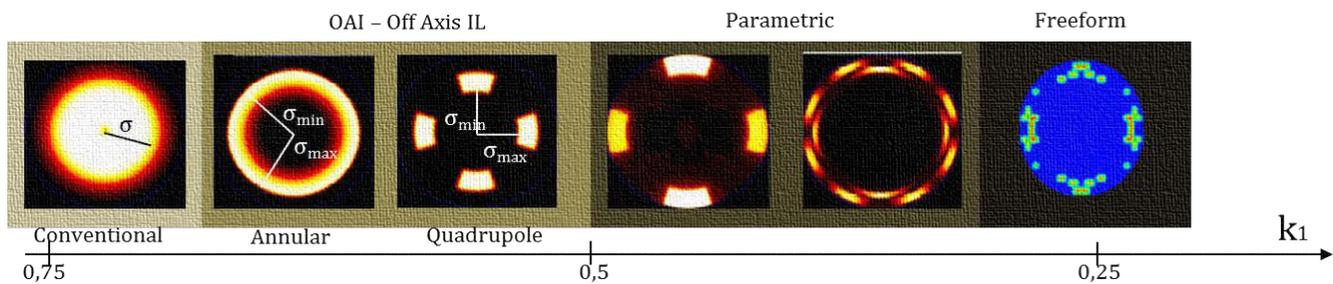


Figure 26 Illumination types under use in optical lithography

Other possible RETs is the Optical Proximity Correction (OPC), a method that allows to selectively modify the pattern shape and the size on the mask in order to more precisely obtain the coveted pattern at wafer level. It became crucial since the diffraction limits are approached. Thus OPC fine tunes the patterns to compensate the image fidelity loss due to increased proximity effects. Proximity effects are referring to features with identical CD printing though differently because of

environment variation (see Figure 27), diffraction (line end shortening), and sharp corners filtering by the pupil (corner rounding).

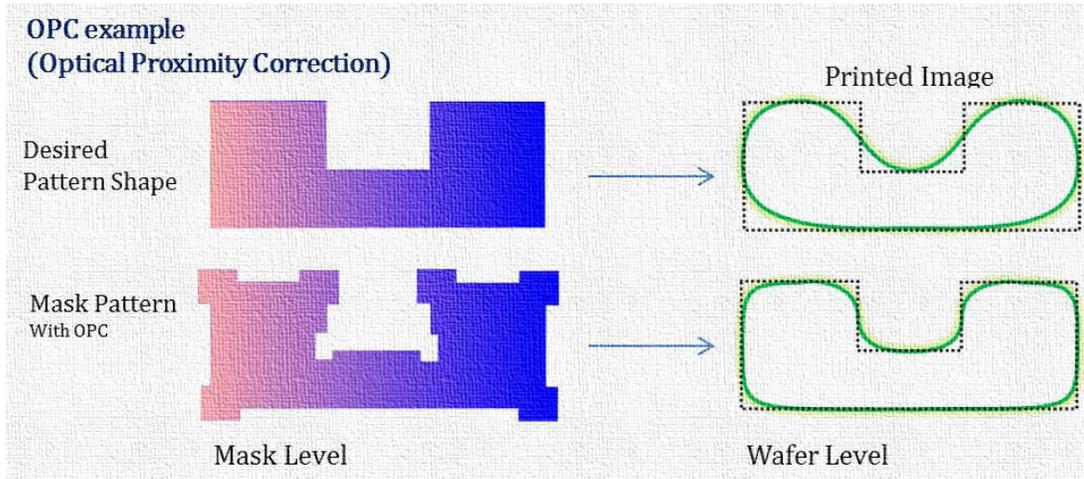


Figure 27 : Example that illustrates an example of OPC

Further improvement is the application of sub-resolution assist features (SRAF), still within the frame of OPC that will subsidize a dense design environment. The SRAFs would manipulate the proximity effects by modifying the diffracted patterns. Overall all this RETs enables to maintain the high adaptive capability of immersion lithography for downscaled technology nodes (22 nm and beyond) that is a quite notable accomplishment.

1.5 Conclusion

This first chapter formed a review from the history of the invention, the bedrock of the microelectronic industry, to the heart element of the fabrication process, the optical lithography. Through a brief introduction to the photolithography, the basic elements with their complexity and “beauty” were shown as well as the possible way of amelioration of each. As seen at current stage of the advanced nodes, the photolithography reaches its boundary that implies the increasing demand for RETs such as the phase-shift mask, the off-axis illumination and optimized sources...

Thus this chapter provided a global view of the environment in which this thesis work is placed. Within the next section, the context and the objective will be presented through the OPC modeling and its limitations.

2 Chapter

Interrelation between OPC model and the process environment

State of art

The immersion DUV (projection lithography) is the current technology of patterning in production for critical layers and, as seen in previous chapter, it arrives at the resolution capability limits. The RETs (Resolution Enhancement Techniques) assist to extend these limits enabling a high-fidelity mask-to-silicon transfer. They include Optical proximity corrections (OPC). Proximity effects reveal in the variation in line width or shape of a structure as a function of the proximity to neighboring features. Optical Proximity Effects refer those that occur beneath optical lithography. A conventional OPC model corresponds to a single technology process including information on the scanner, the photoresist and the exposure conditions. These OPC models cover a rich theory of optics, optimization and imaging work in order to respond to the high degree of complexity related to the apace scale down. This section provides a description on the modeling portion of the scanner. In parallel it poses the current challenge of conventional models providing the bedrock of the thesis.

2.1 Introduction

As seen in Chapter 1, the Resolution Enhancement Techniques (RETs) enfold all the knobs that enable a further assured mask-to-silicon transfer. Within these knobs, the Optical Proximity Correction (OPC) is one of the key optical lithography enabler to improve a reliable resolution and pattern transfer fidelity for IC industry. Thus OPC is the most current applied technique to handle design shape distortions for sub wavelength lithography. While immersion lithography tools are now reaching their limits of optical resolution, OPC is more and more constrained so need tighter attention.

In order to have a deeper insight into OPC, its different correction types are defined. The simplest and oldest correction concept is the rule based OPC. It is based on geometrical rules. This rule based concept consists of adding or removing parts from the photomask artificially. Its intention is to predict and compensate the resolution loss during the image formation. Thus this kind of pre-compensation of the mask pattern is a manner to optimize the aerial image that would take the expected distortion due to diffraction into consideration. As seen in Chapter 1, as consequence of the diffraction effects, the precise dimension and the shape of the structure to be patterned at wafer level¹⁵ depend on the proximity to another feature. Therefore, following the shrunk trend of pattern dimension that is now comparable or even less than the exposure wavelength (λ), the imaging of individual features is not realized independently. They will interact with adjacent patterns.

So the solution to improve the printed pattern fidelity sequent to rule based concept is as follow: the rules are fixed for simple patterns and their values are fixed empirically. Figure 28 illustrates several examples on the rule based modifications. It includes the following ones:

- Additional material in the corners, so called Serif and additional parts to line ends, so called Hammer Head : to improve corner and line end fidelity. Serifs can be convex or concave.
- Removal of the material: to improve corner fidelity.
- Additional unresolved patterns: The unresolved structures (sub resolved assist features, SRFAR, and scattering bars, SBAR) are applied near isolated pattern to mitigate the variation through focus. Thus isolated patterns profile variation is comparable and they locally reproduce dense lines. Nevertheless their dimensions are infinitesimal, thus they won't be resolved in the photoresist.

¹⁵ Pattern that will be printed on the wafer

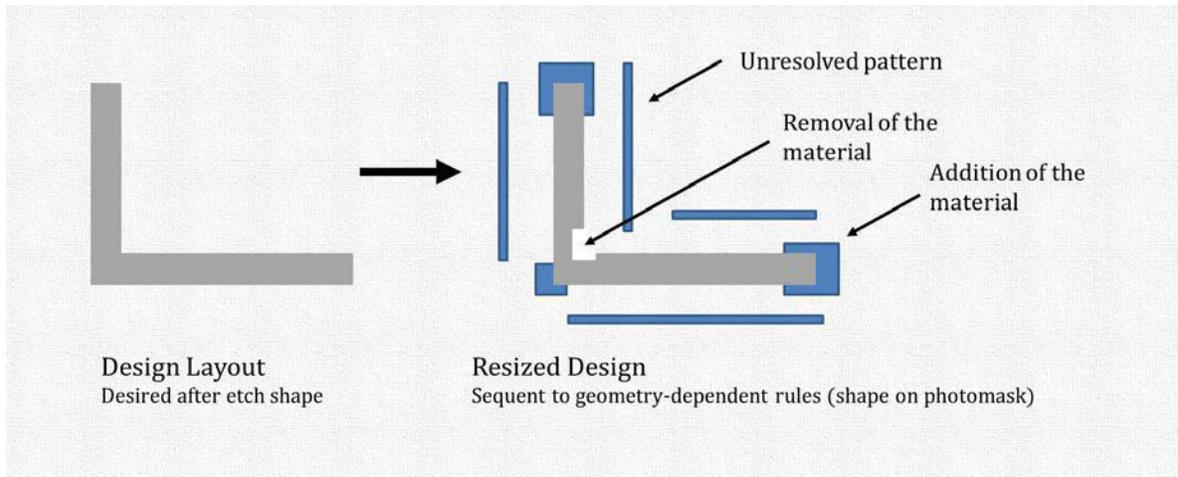


Figure 28 Geometrical modification of a simple L shape pattern to compensate the proximity effects (rule based OPC)

As another correction type, a more refined concept can be named: the model based OPC. A conventional model based OPC is supported by models simulating the lithographic process. These models are anticipating the possible geometric distortions during the printing within the photoresist in order to correct these distortion areas at photomask level. The complex OPC treatment process is done iteratively, where at each iteration stages the photomask is adjusted for example by edge displacement. And finally the resist contours are simulated. So as result, the adjusted photomask design differs from the desired design.

In practice, the two above described types of correction are combined for higher efficiency. An example of OPC treatment is illustrated by Figure 29, which is applied at STMicroelectronics. The draw, simplified schema considers a simple initial designs described by the specified target post-etch dimensions. The first step is to resize the design by rule based treatment. So it is sequent to geometry dependent rules in function of the desired target dimensions within the photoresist. Further improvements can be done at this stage, still rule based treatment, by the application of the SRAF. In rule based OPC treatment, the main task of these features is to subsidize a dense design environment. It is due to the fact that they will manipulate the appearing proximity effects by modifying the diffracted pattern. Once the SRAFs are inserted, the design edges are fragmented in order to prepare the mask design for the model based treatment. Each fragment is separately moveable. Thenceforth an iterative process is realized within the frame of model based OPC. It is composed by contour simulation within the photoresist, which is enabled by empirical models. In addition, an adjustment is done at design edge level. It is based on minimizing the delta CD between the obtained and desired “printed” shape with the goal of a better matching to target dimension. The result is that will be etched into the photomask. The global objective with the entire treatment is to provide the less difference in between the simulated results and the desired draw contours at wafer¹⁶ level.

¹⁶ In this context wafer level means during the modeling stage

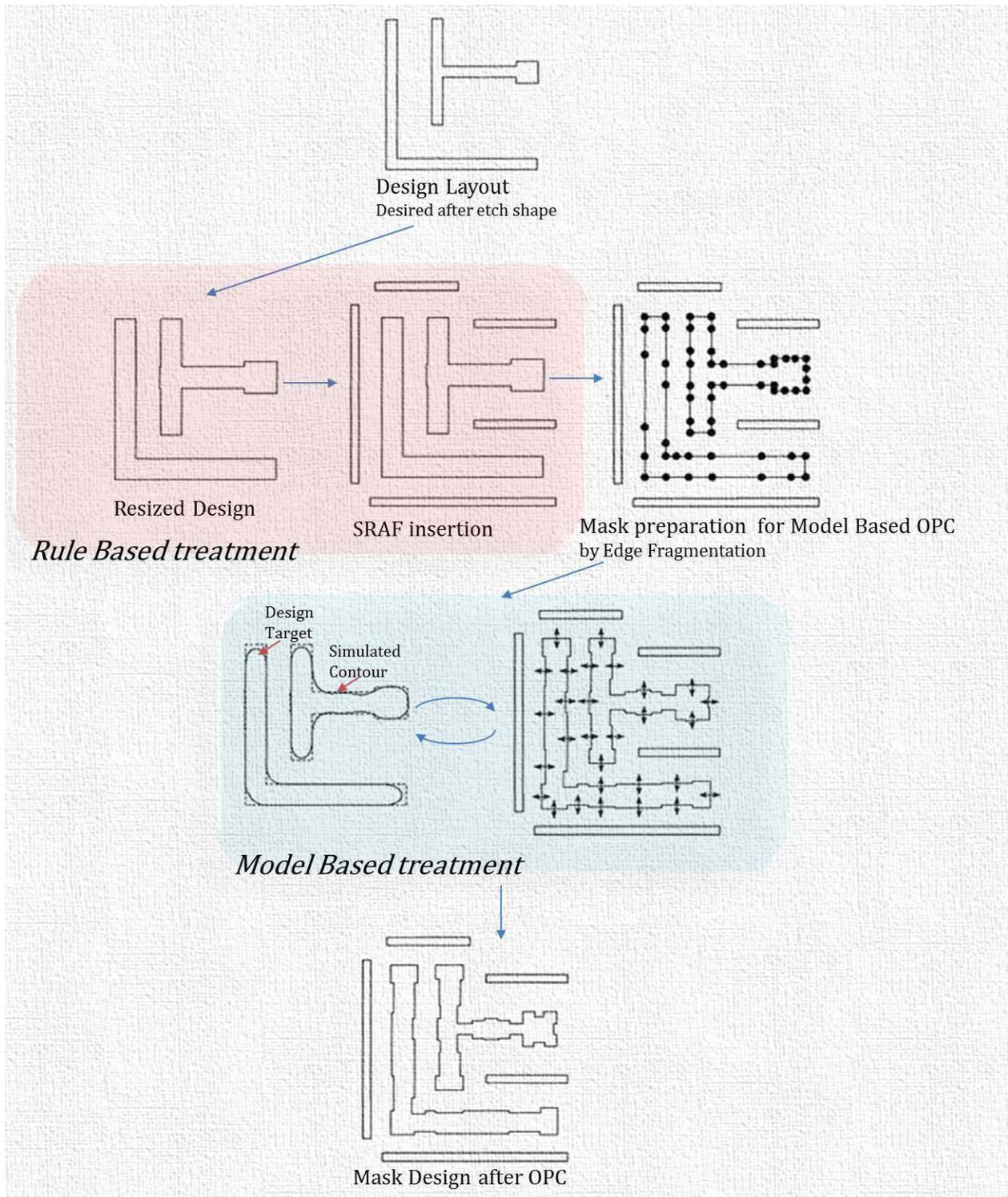


Figure 29 simplified OPC treatment schema indicating the different stages from the design layout to the final mask design (Saied, 2011)

Once the OPC treatment is completed the set models used for model based OPC are further applied in Lithography Manufacturing Check (LMC). The LMC is full-chip verification. It comprehends the simulation and prediction of wafer CD across the field. Its major purpose is highlight critical patterns in the field before the photomask order is sent to the mask shop. In addition it enables to pull out design regions that are sensitive to process condition variation. LMC basically comprehends of a comparison of the simulated full-chip resist contour against the design target. Note that an OPC model (thus an LMC model) is constructed for one particular process, for one particular optical condition (scanner) and applies an ideal photomask.

As every element of the industry, the above described design treatment targets also to be less time consuming. Indeed one of the main advantages of the modeling is its ability that makes possible to realize experiments in a virtual environment in a faster way. The basics on OPC modeling will be introduced in the following sections.

2.2 OPC model

A reliable model is a core element for OPC treatments (model based OPC part). Therefore a pronounced effort is put on the creation and the calibration of the OPC models. In current lithographic applications, OPC models can be decomposed into two main simulator parts:

- Aerial image simulator: A purely optical simulation of the light intensity arriving on the photoresist, so called optical model. It covers the beginning of the optical chain including the light diffraction, propagation from the radiation source to the photoresist (through the photomask and the projection optics);
- Photoresist simulator: A resist model describing the exposure light vs. photoresist interaction related phenomena coupled with the diffusion effects. It computes the resist contour based on the aerial image.

In the next sub sessions the above mentioned optical model and resist model will be described. Note that current OPC applies almost rigorously predictive model-based optical models and hybrid approach is used in resist modeling since a physical model is still CPU time consuming.

2.2.1 Optical model

By the optical modeling part, the primary goal is to describe the aerial image, then to determine the light intensity at wafer level. A conventional optical model can be considered as a transfer function that enables the light intensity calculation (thus the aerial image) within the photoresist based on an input data from the photomask design. It is applied to describe the physical behavior of the exposure stage in lithography and it represents the lithography imaging sequence. Standard modeling method for aerial image formation of the photomask is based on Fourier optics. It is coming from the fact that the light intensity is obtained from the interference of the diffracted orders at mask level. The objective lenses are performing as Fourier Transform resulting in diffraction patterns of the object in the focal plane. Then an inverse Fourier transform creates the

interference of the diffracted radiation back to (real space image) image plane thus resulting in the desired image (see Figure 30).

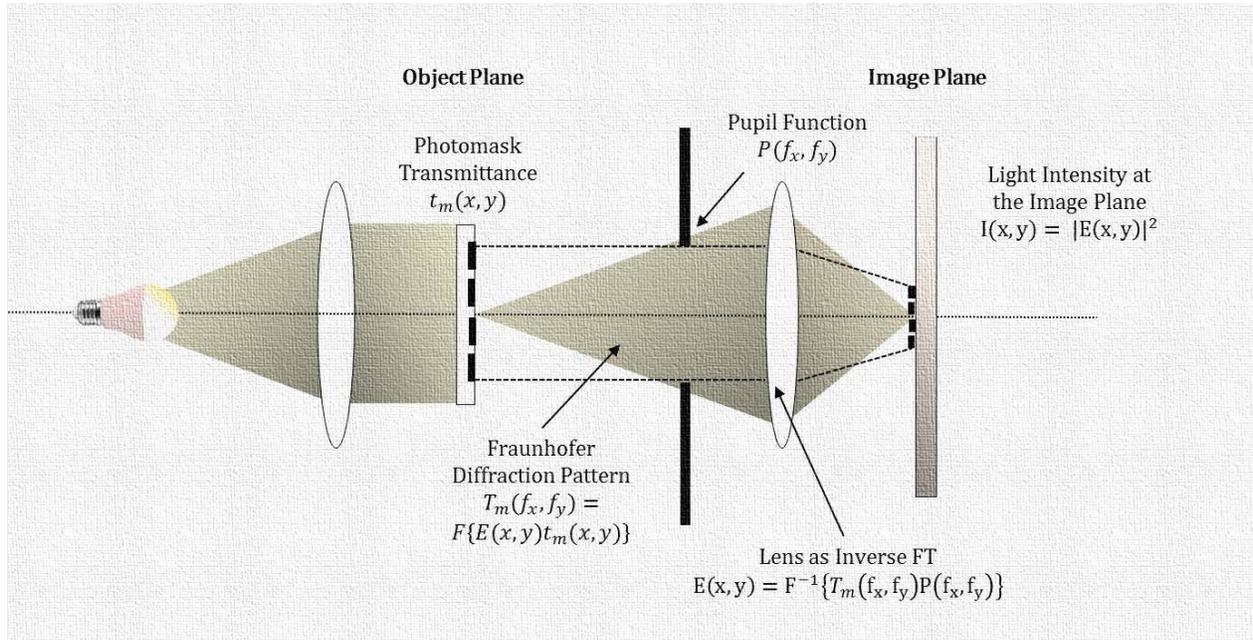


Figure 30 Schema of passing light indicating the basic considerations for lithographic modeling (for more details please check Annex 1)

The Fourier optic considers that the diffracted orders propagate in respect to the angle that corresponds to the spatial frequency of the photomask pattern. Each diffracted order is defined by the corresponding angle. Through the exposure apparatus the waves propagate in accordance to the angle θ_i . Although it should be mentioned that the captured orders by the pupil are limited by the numerical aperture NA; it means the light emerging from the mask at $\sin\theta_i > NA$ will not participate in image forming within the photoresist. Therefore only a part of the incident light will contribute to the final image at wafer level. Thus the replica of the photomask image is constructed then by the resulting, interfering light.

The basics of the lithographic modeling for such a complex exposure apparatus can be simplified and interpreted from mathematical point of view (see Figure 30). In a nutshell the mathematical description behind the optical model is linked to the Abbé's method. In Abbé's imaging, the source is discretized. The final aerial image is a sum of the calculated aerial images for each grid point. As much as this method is reliable, as the source is an integration of the entire set of source points, it remains computationally intensive and less applied for advanced nodes. Therefore, it is further simplified and approximated by the Hopkins formula. H. H. Hopkins basically reformulated the Abbé's method to a partial coherent image calculation (Mack C. , 2007) (Rodrigue, 2010) (Saied, 2011).

The Hopkins formula is extended usually by an intermediate term, the Transfer Cross Coefficient (TCC). The application of the TCC has the advantage that it remains independent of the mask function. It is due to the fact that it is a function of the spatial frequencies depending only on the

illumination source and on the pupil (Mack C. , 2007) (Michel J.-C. , 2014). Therefore it enables the possibility of computing and storing data in a pre-phase resulting in a time efficient computing. Hence it is widely used in commercial software's embodiments for OPC applications since OPC will modify the pattern geometry for better pattern replica at wafer level.

At this point the sequential step to the up-to-now obtained aerial image resulting in a light intensity map at wafer level is the contour formation modeling that will be described by the next session.

2.2.2 Resist model

The emphasis of this part is put on the modified chemical composition of the photoresist during the exposure. In accordance to the literature different modeling methods exist for the different stages of the contour formation. It is due to the fact that the light spreading within the photoresist and the resulting photoactive compound reaction dispose of a high level of complexity. It requires a large scale of parameters that need to be set and well composed. Thus these models are varying depending on the applied approaches beginning with the N.Cobb's constant threshold approach arriving to models referring to as semi-empirical, lumped, compact or behavioral... (Cobb, 1998) (Saied, 2011) (Zuniga, 2014) (Liu, 2012). Although all of them are placed in the framework of mathematical characterizations providing a transfer function in between the system inputs and the measurement outputs of interest.

Note that all the above mentioned physical models in lithographic software applications are rooted from Dill et al mathematical framework for positive tone photoresist (Dill, Neureuther, Tuttle, & Walker, 1975) (see Figure 31). Dill's method is taking into account both the exposure and the development stage effects. These two stages are coupled since the light propagation will cause the photoactive compound reaction to occur. It results in a change in the resist material composition. The developed model by Dill is using the light intensity at different resist heights (or depths) and its impact on chemical reaction over time. He perceived that positive tone resist becomes more transparent among the exposure, while the light intensity exponentially decreases with the resist thickness. Finally the resist exposure model among Dill is a model with three parameters: A, B and C. The named parameters define the light absorption within the resist and the reaction efficiency:

- "A" corresponds to the absorption coefficient of the resist, which becomes transparent during the exposure [μm^{-1}];
- "B" corresponds to the absorption coefficient of the unbleached resist, which is not changing during the exposure [μm^{-1}];
- "C" corresponds to the exposure rate constant describing the chemical reaction efficiency [cm^2/m];

The computation of the profile image in the resist film or, in other word, the distribution of the photo-generated acid during the exposure step is based on the aerial image. This forms the so called latent image resulting of the exposure process simulation on the concentration distribution of the photoactive compound in the resist noted as (M). The dissolution rate of the resist during

the development is related to the instantaneous concentration of the acid at the surface of the developer. Finally, the calculation of resist profile is acquired from the obtained latent image (Saied, 2011).

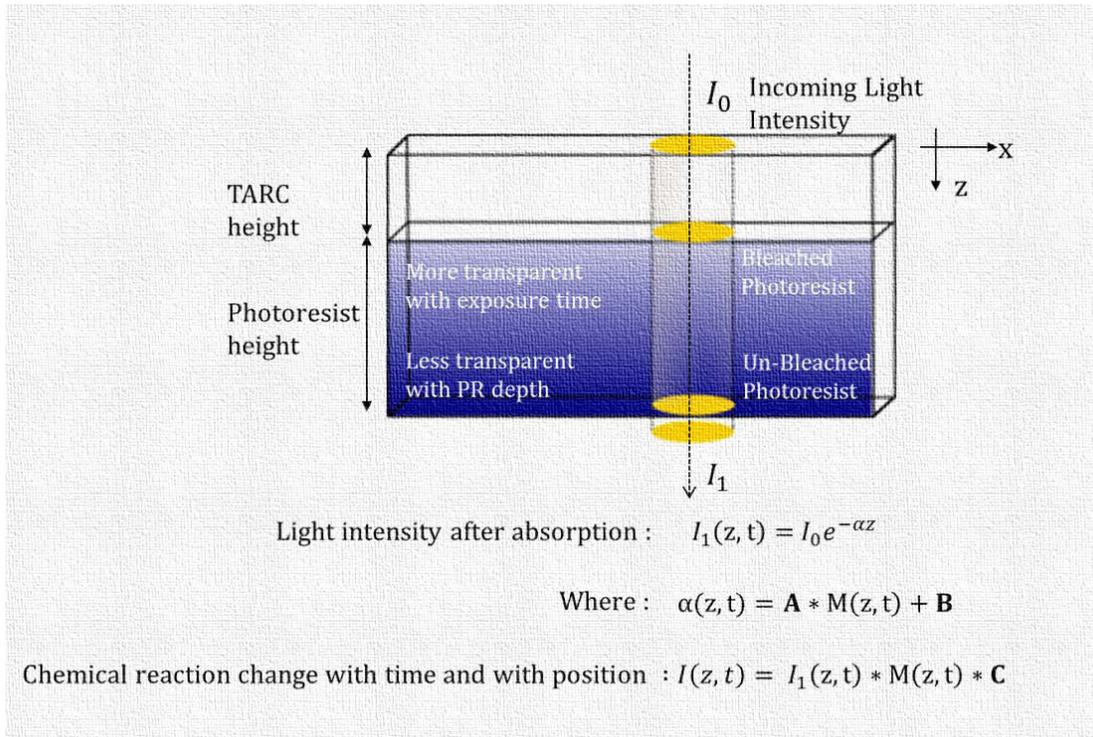


Figure 31 Illustration of intensity expression (Dill's method) taking a TARC coated photoresist case – here A, B and C corresponds to Dill's parameters (Dill, Neureuther, Tuttle, & Walker, 1975)

By the appearance of the most commonly applied resist type, which is the Chemically Amplified Resist (CAR)¹⁷, an additional stage, the Post Exposure Bake (PEB)¹⁸, is inset just before the development. Although it is not as simple since the PEB step is accompanied by a de-blocking reaction that results in solubility change of the resist.

“For chemically amplified resists, things are a bit more complicated. Diffusion during post-exposure bake is accompanied by a de-blocking reaction that changes the solubility of the resist. Thus, it is not the final, post-diffusion distribution of exposure reaction products that controls development but rather the integral over time of these exposure products. Things become even more complicated when acid loss is accounted for. In particular, the presence of base quencher (that also may diffuse) leads to both complexity and advantage in tailoring the final latent image shape. However, for the simplified case of no acid loss an analytical solution is possible.” (C Mack)

¹⁷ Chemically Amplified Resist (CAR): ~ is a type of photoresist that is the most commonly applied for DUV lithography processes that upon the PEB will multiply the chemical reactions by the support of chemical catalysis (Mack C. A., 2006)

¹⁸ Post Exposure Bake (PEB): ~ is the heating process of the wafer at post exposure stage. It stimulates the diffusion of the photoactive compound of the resist and it reduces the possible standing waves through the resist height. For CAR, it causes a reaction changing the solubility of the resist.

The photo-generated acid diffusion at PEB can be modeled by applying Fick's second law to the initial image (Saied, 2011). The photo-generated acid concentration in addition to the protecting compounds is calculated iteratively. So the latent image is obtained by calculating the concentration of the non-protected polymer parts. The dissolution rate is calculated based on the covering resist. Several ways of dissolution rate calculations were introduced since, although among the proposed solutions/papers C. Mack's model is a good approach that is a well-tried one for now on (Mack C. , Advanced topics in lithography modeling, 1986). Figure 32 illustrates Mack's model by indicating the PEB stage.

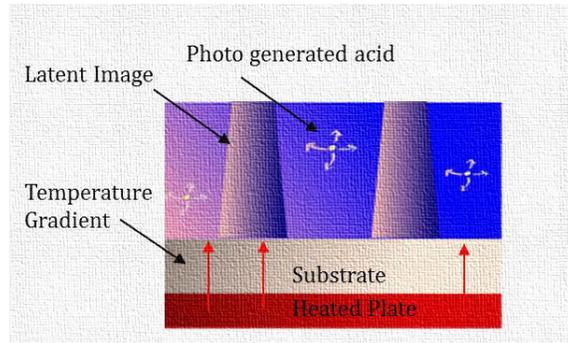


Figure 32: Simplified schema of PEB process step of C. Mack

This coupled model (the optical & the resist) is considered working correctly, when it fulfills several criteria that make it adequate for OPC applications. One of the key points is the time effectiveness since such a model would require a high level synchronization of a large number of parameters. This synchronization is time consuming. It means that higher the number of parameters is to be set, higher the CPU¹⁹ time is for the simulation. Thus OPC tools users make efforts to apply the simplest models. Here within the framework of this thesis only one modeling approach is gripped, the Brion Technology's modeling approach (the extended convolution).

2.2.2.1 The extended convolution type of modeling

As an understanding of the resist behavior, the resist model is expanded with the diffusion and density by extended convolution. The basic idea of this method is to enable the computation on a simulation grid the latent image surface within the high contrast recording medium. This determines an effective resist image that will match the wafer measurements at a given threshold value. This resist image is denoted as $R(x, y)$ and it is a linear combination of a set of term images (the aerial images and the mask images), Equation 11:

$$R(x, y) = \sum_i coefficient_i \cdot term_{image}$$

Equation 11

¹⁹ CPU = Central Processing Unit

Each image term is determined from the aerial image and from the mask image. The different terms correspond to the aerial image terms, to the photomask images terms, to the truncated aerial image terms, to the aerial image threshold and to theta the CD offset.

- *The aerial image terms includes:* A: The aerial image intensity; AG₁: The Gaussian representation of the acid diffusion for the reduced size patterns; AG₂: The Gaussian representation of the acid diffusion for the increased size patterns and the local slope of the light intensity at this (x, y) point.
- *The photomask image terms includes:* MG₁: The Gaussian representation of the base diffusion for the reduced size patterns; MG₂: The Gaussian representation of the base diffusion for the increased size patterns; M_{average}: the local mean of M, the photomask image within a radius around the point (x, y).
- *The truncated aerial image terms includes:* A_P: The Gaussian representation of the truncated aerial image on the acid diffusion; B_P: The Gaussian representation of the truncated aerial image on the base diffusion; A_M: The Gaussian representation of the truncated photomask image on the acid diffusion; B_N: The Gaussian representation of the truncated photomask image on the base diffusion.

These different terms correspond to different physical and chemical effects of the lithographic process. They are calculated from the aerial image of the photomask image. The contour within the photoresist can be written as follow²⁰, see Equation 12:

$$R(x, y) = c_A A + c_{M_{average}} M_{average} + c_{A_p} A_p \otimes G_{A_p} + c_{B_p} B_p \otimes G_{B_p} + c_{A_m} A_m \otimes G_{A_m} + c_{B_n} B_n \otimes G_{B_n} \\ + c_{AG_1} AG_1 \otimes G_{AG_1} + c_{AG_2} AG_2 \otimes G_{AG_2} + c_{MG_1} M \otimes G_{MG_1} + c_{MG_2} M \otimes G_{MG_2} + c_{slope} Slope \\ + \theta - threshold$$

Equation 12

The aerial image intensity leaving the optical system (so at wafer level) is noted by A. The model parameters are: AG₁; AG₂; A_P; B_P; A_M; B_N; MG₁; MG₂. They reveal in the Gaussian width and they have all a weighting coefficient to be optimized and to be fitted. These coefficients are: c_{AG1}; c_{AG2}; c_{Ap}; c_{Bp}; c_{Am}; c_{Bn}; c_{MG1} and c_{MG2}. The \otimes operator represents the convolution. The G_{Ap}; G_{Bp}; G_{Am} ... and G_{AG1} are Gaussian filters that models diffusion effects or 3D resist development. The threshold corresponds to the resist image threshold. Note that the Equation 12 is referring to the resist image equation what is provided by Brion Technology.

Overall the extended convolution model is an extension of Gaussian distribution with the ability to model diffusion phenomena with other optical or chemical phenomena such as the effects on short, medium and long distances (by convolution of the mask image). The diffusion is directly taken into account right after the light intensity computation by its integration into the TCC matrix (Liu, A full-chip 3D computational lithography framework, 2012). Note this method would not much increase the CPU time.

²⁰ Same for negative and positive tone photoresist

As summary for this part, the basics of applied optical model and resist model are shown. Figure 33 illustrates a summarized schema on the OPC modeling for a given design layout. It indicates the two sub models, the optical and the resist model. It shows the necessary inputs process parameters such as the type of the source, the photomask characteristics, the NA (cf. chapter 1), the magnification of the scanner and information on the wafer stack. As the OPC model needs to be calibrated it would need experimental data on the pattern of interest (mostly applied metrology is the CD SEM). The calibration is described more in details in the next section.

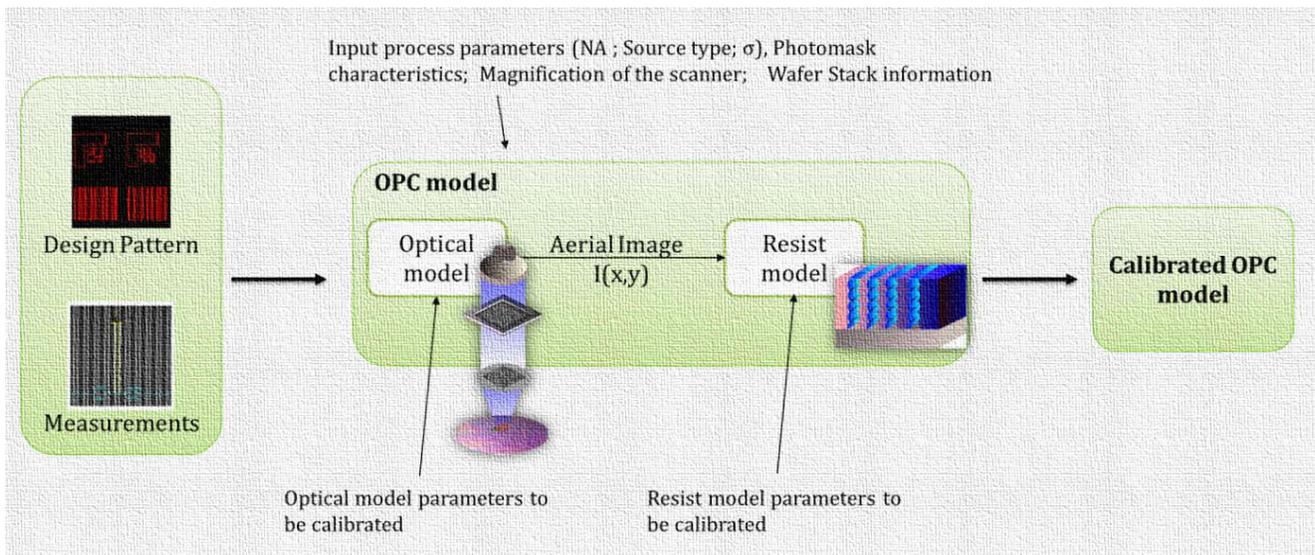


Figure 33 Schematic view on OPC modeling illustrating the input and output data

2.3 OPC model calibration

The mask fabrication process related modeling enables the description of patterning process in integrality that includes the photomask, the scanner specific optics, the photoresist and the etching stage. Although current modeling methods does not apply the etch model due its “immature” nature, but the re-adjustment after the etching table (post-etch bias).

The optical and the resist related parts are characterized separately and they are delineated by an ensemble of parameters. A part of these parameters are users known input parameter values, the other part is obtained empirically during the calibration step. During the calibration process of the OPC model a photomask layout including design specific patterns or a test photomask layout is measured by printing the target process design. The CD measurements at wafer level are done on photoresist (post-lithography) by CD-SEM²¹ apparatus metrology tool. Finally the OPC model is tuned iteratively by varying the certain parameter values that are quiet difficult (or impossible) to be measured. Hence their values are estimated by comparing the model prediction vs. the experimental wafer data.

²¹ SEM = Scanning Electron Microscopy

These photomask, optical, photoresist and etch process calibration related parameters can be distinguished into three sub categories:

- Software options associated parameters, which are alternating the different approximations used in models (e.g. photomask representation type, resist model form)
- Physical phenomena associated parameters, which are built in the model as mathematical proxies (they are related for example to complex PEB and development)
- Optical chain, lithographic apparatus associated parameters, which can be directly measured such as: the photomask stack characteristics as the resist thickness and refractive index; the incident light wavelength; NA, lens aberrations, source shape parameters as partial coherence, mask line width etc.

In parallel to parameter variation, the OPC model is tuned by reducing to minimum the errors in between the model predicted CD and the measured CD data based on RMS error. So the simulation prediction capability is required to be ensured as high as possible. It means that it should be able to predict the most proximately the image and the photoresist characteristics of interest based on measurement data.

Note that various modeling platforms exist with different modeling flows- depending on the suppliers. Here a general idea of the calibration flow is described.

2.3.1 The basic calibration patterns

In order to build up a reliable model that has the capability to make the most precise possible²² mask pattern replica within the process window, experimental measurements need to be set up. For these purpose test patterns, so called calibration patterns are drawn within the photomask and then printed. In general the calibration features consists of two populations of these patterns. One is representative of the transfer function of the optical device, thus it is the mainstay of the calibration model. The other is the device layout specific including all the geometries that are assessed to a given layer. In practice the calibration features consist mainly in 1D structures (see Figure 34) providing enough information to characterize a process (see Chapter 1 for 1D and 2D pattern definition). It is due to the fact that a 1D structure is less complicated to be measured. The CD of these 1D patterns is measured perpendicular to the long line, as the only important dimension.

Once a process model is finished and fine-tuned, OPC is able to be performed on any geometry. Geometry that meets the design rule check (DRM - see Chapter 1) regulation. The undermentioned patterns are examples for the basic test structures that are:

- Dose sensitive structures such as dense line stripe structures (line vs. space);
- Focus sensitive structures such as isolated trenches;
- Both focus and dose sensitive structures such as line stripe structures with varied pitch (as if the pitch is slightly below the pitch of the second diffracted feature) or line between blocks or line ends

²² Means that the simulated results meet well with experimental data (on wafer)

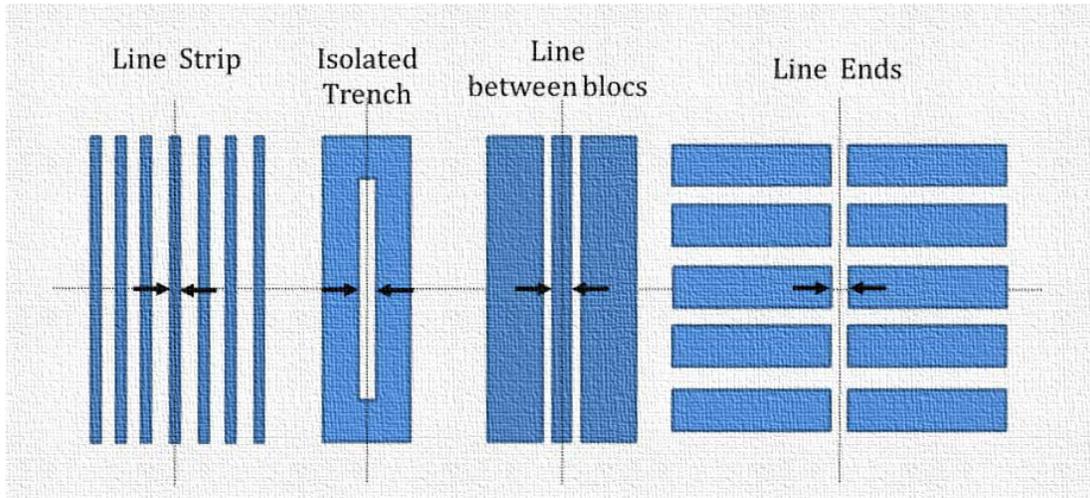


Figure 34 Simple calibration structure examples that are used in practice

2.3.2 Process window – a reliable metrology setup

The model calibration through the process window requires the extraction of the measurements through the corresponding process conditions. For this purpose the observed wafer's exposure is based on the Focus-Exposure-Matrix (FEM, also called FEM wafer) so that each die²³ is printed at different focus and different exposure energy conditions. The model calibration and construction is preferred to have the more data possible; but at least three different dose positions and three different focus positions (including the nominal condition: nominal dose with nominal focus) are considered. Thus the selected fields can cover more or less the whole Process window with the nominal and four additional conditions. Figure 35 illustrates how an FEM wafer looks like indicating the process window where the pattern resolution is within the tolerance²⁴ (green area) and also how the minimum calibration fields can be chosen. Note that the described method is an example, a generic method for calibration.

²³ Die : a naked chip before photoresist coating (plural: dice); in practice the coated chips are called die also

²⁴ Usually the tolerance is given as : $CD \pm 3\%$

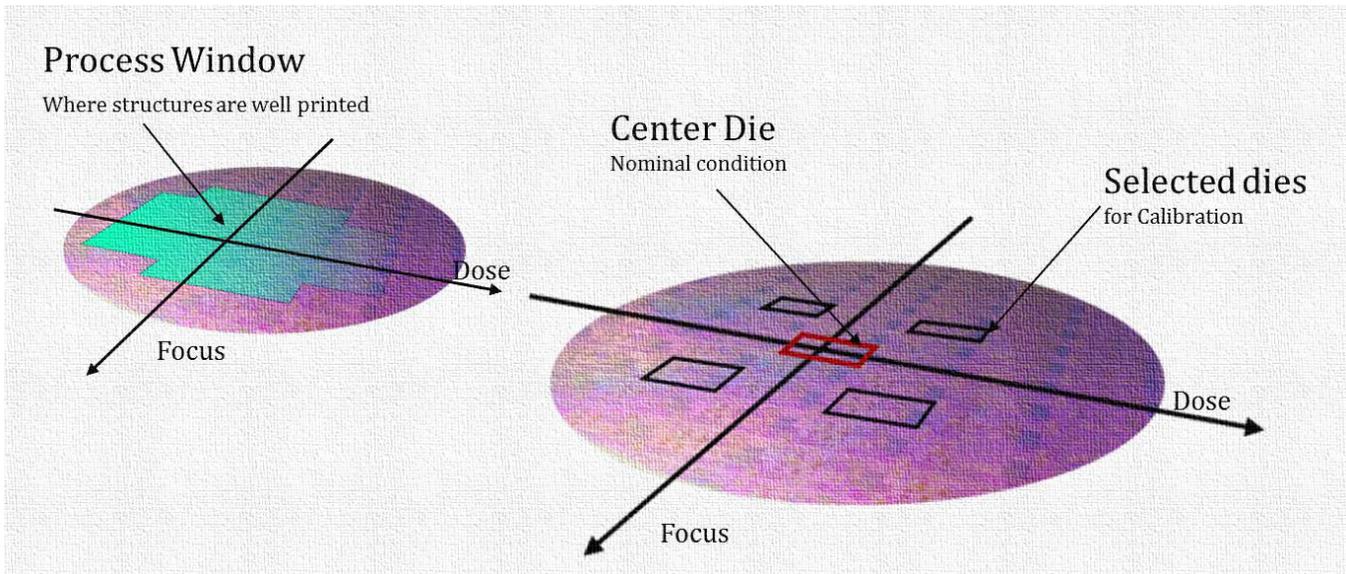


Figure 35 Example for a FEM wafer and the calibration die selection

By selecting a small number of fields and structures including the maximum depth of field and the exposure dose variation the analysis covers the different behavior of interest in the process window.

2.3.3 Modeling requirements (the accuracy rate)

As mentioned the accuracy rate of the calibrated model is measured by the prediction capability of the calibrated model and the corresponding wafer data. The model accuracy is influenced by several factors. Within these factor, the more pronounced is its ability to “reproduce” the patterning trends through pitch, through pattern shape and through target size for a process condition.

The applied metric is the errRMS (Root Mean Square) value that is the best metrics to evaluate the model ability vs. experimental data including the measurement’s noises. Note that the errRMS is the most currently applied metric, although it might be recommended to check e.g. on the Bossung curve’s predictions.

The errRMS is coupled with the test structures ensemble and it is weighted enabling to assign higher importance level to certain critical patterns. The errRMS is expressed as Equation 13:

$$errRMS = \sqrt{\frac{\sum w_i (CD_{simulated} - CD_{measured})^2}{\sum w_i}}$$

Equation 13

Where $CD_{simulated}$ is the model CD for each measured location; $CD_{measured}$ is the CD data for related measured location and w is the user-specified weighting factor value for each i^{th} point. The

weighting factor's value varies between 1 and 0 regarding the image acceptability and quality. Thus it would be 1 if good quality and 0 if failed.

2.4 Current limitations to “everything is linked” solution

In addition to these accuracy requirements as the OPC models are put into the increasingly complex process optimization it should now be interconnected with other elements of the lithography process loop. Figure 36 illustrates how conventional OPC models (OPC and LMC models) are connected to the process elements. As first stage the OPC treatment is done then the corrected mask design is sent to the Photomask shop. It is important to note that a standard OPC model is created at the beginning of the process and it will remain constant all along the process in the production environment. Therefore a conventional OPC model does not consider the different root causes of process variation.

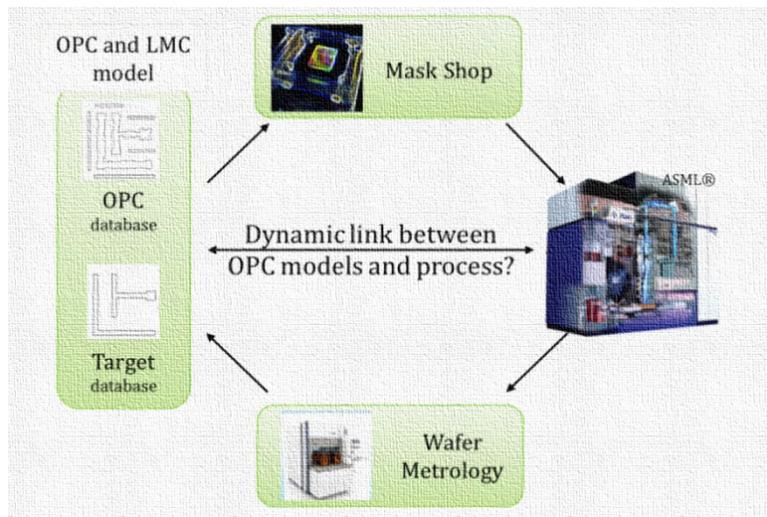


Figure 36 Lithography process loop indicating the OPC position within and the missing dynamic link between the OPC models and the process.

Following the continuous scaling of semiconductor components, the predictive modeling in lithography is put into spotlight. It is a result of the patterning process control requirement in order to remain competent for advanced nodes. At this stage, the process control is needed to provide a deeper insight on the root cause of process variability since at this low- k_1 domain certain sources of variability can be no longer neglected, which means every nanometer counts. Hence it will involve a larger domain of the all-encompassing process. Thus the computational lithography²⁵ support is one of the key enablers guide²⁶ for advanced technology node generations from comprehension-prediction-mitigation point of view.

²⁵Computational lithography is the set of mathematical and algorithmic approaches designed to improve the resolution attainable through photolithography.

²⁶Guide referring to “As an industry, we use the development of lithography simulation as a tool to test advance our understanding of lithography” (Mack C., 30 Years of Lithography Simulation, 2004)

Currently, all the computational lithography efforts, such as OPC, are entirely dependent on the lithography modeling ability and the metrology processes. Today full chip simulations apply approximate image models and empirical resist models that enable a reliable OPC technology. But still, there is a requirement for an even faster simulation including deeper physics (physically based models) providing higher level accuracy over a wider range of process conditions.

Hence, process simulation works would now tend to use approaches that integrate the computational lithography abilities, the wafer patterning and the process control (Figure 37). Therefore “everything is linked”, where everything covers the following assets:

- The computational aspect that has the important role to reduce the process window shrinkage (coming from the shrunk patterns with high level complexity that result in tight process window);
- The advanced lithography tool that permits the high-volume production; in addition it provides process fingerprints, which are then the process control, in parallel they can be implanted into the OPC models with the ultimate goal to compensate or mitigate them by the following;
- The innovative wafer metrology methods ensure that the lithographic apparatus remain well-controlled and centered on the desired process window.

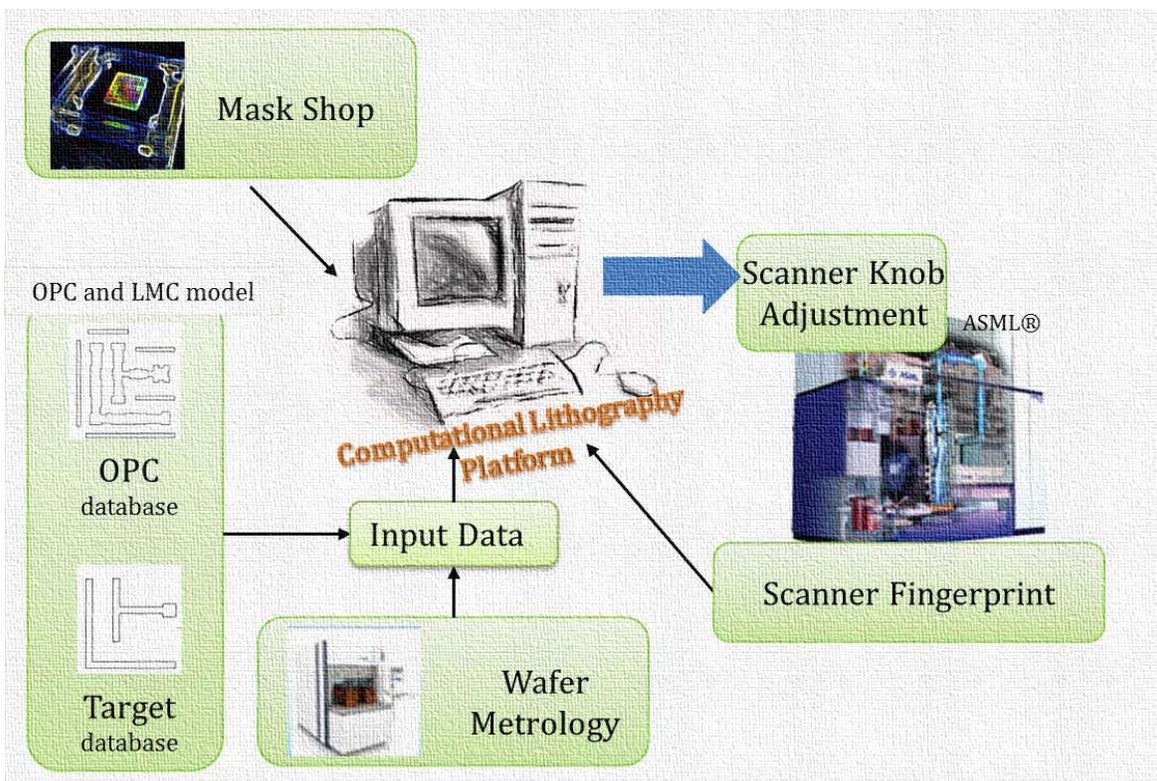


Figure 37 Targeted lithography process loop indicating the interconnection in between the OPC models and the process (covering the tool and the related metrology)

In the literature the following term is proposed to cover the “everything is linked” expression: the holistic lithography (Mulken, 2014) (Le-Gratiet, 2015). This term is referring more precisely to an approach within the IC design. The photomask, the lithography and metrology process are matched with each other in order to enable an optimal process for shrunk nodes in fab. In addition the scanner and the other process behaviors are optimized in a better way following this approach. Note that for less critical nodes the various process stages were optimized independently of one another, but now for 28 nm and beyond it’s not enough. Although at this point a question can be posed, whether the modification targeted to obtain the interconnection of the above mentioned stages would result in any impact on the process/device or not?

As consequence to the challenges faced to current 193nm immersion lithography solutions for 28 nm and beyond technology nodes, within the framework of this thesis the emphasis is put on the anticipation in the dynamic link realization. It will be based on the prediction, the quantification and the diminution of the process variability impact relying on computational simulation capabilities: “It is fair to say that current optical lithography capabilities would be impossible without the use of computational lithography tools, and these techniques will only become more essential in the future” says Dr. Mack.

2.5 Conclusion

This chapter formed a review on the OPC model and described how a conventional OPC model is built on with the basic ideas behind. Through a brief overview the two distinguished model parts: the optical and the resist part are shown. The limitation of standard OPC model was discussed that enabled an opening on the thesis problematic: the current OPC does not consider the different variability roots that reveal in the importance of the interrelation of OPC model vs. process. The realization of the interrelation will be then based on the prediction, the quantification and the diminution of the process variability impact supported by the computational lithographic abilities. Thus, it provides the bedrock of the solution posed by this thesis work, as suggested in the title “the diminution of the lithographic process variability for advanced nodes”.

3 Chapter

Enhancement of available OPC models' upfront process variation prediction ability

In the continuity of the previous chapter, the emphasis is still put on the OPC models and their upfront process variation prediction capability on which the upcoming chapters rely on. These variations are mainly related to 3D effects that are referring to mask topography, resist profile and wafer topography induced effects. Indeed a conventional, recently presented OPC model would not take into account these effects. As the 3D effects are not anymore negligible for shrunk dimension nodes, which is the case for 28 nm and beyond, these models become obsolete. So by this chapter, the main intention is to introduce the construction background of accurate 3D modeling including the mask 3D model and the resist 3D model. These 3D aware models provide then efficient support for high-level detection of the most critical patterns.

3.1 Introduction

An accurate simulation of the aerial image at virtual wafer level is an essential point for prosperous OPC algorithms. In addition to that key point, as mentioned in previous chapter, OPC modeling is required to provide a high level velocity in term of simulation time to accomplish a faster mask preparation. Indeed there is always a tradeoff in between the speed and the accuracy on this type of operation.

For current advanced nodes (28 nm and beyond), the emphasis is on the 3D lithography simulation enabling the modeling of 3D effects in all lithographic processes. These effects concern the mask topography, the resist profile and the topography at wafer level. The framework of this thesis is expanded up to the resist profile including the mask topography, but not the wafer topography. For further detail on wafer topography related work, the reader is kindly invited to check on the following reference (Michel j.-C. , 2014).

3D effects over the mask topography: currently the most widely applied method for photomask modeling is the thin-mask approximation for its acceptable accuracy and speed. This is called the Kirchhoff approximation, which considers the mask thickness to be infinitely thin (Saied, 2011). At the region, where the photomask pattern dimensions are comparable to or less than the wavelength λ , the thin mask model would become inadequate (Saied, 2011) (Liu, 2012) (Finders, 2014) (Le-Gratiet, 2015). It is due to the fact that the mask topography effects lead to image contrast loss, pattern-dependent best focus shift, and dependence of diffraction efficiency on illumination. So accordingly, the 3-dimensional nature of the mask needs to be taken into account via a mask 3D model. It is therefore thought to be indispensable to build an accurate model.

3D effects over the resist profile: for instance resist with significant top loss may be etched away during the etching step. Furthermore an ill-controlled resist sidewall angle (SWA) can lead to undesired etch bias and cause the after etch feature to deviate from the design target, see Figure 38. Figure 39 represents a real example on that problematic. Two post-etch & post CMP X-SEM images are shown. The left image belongs to the case, where the desired target design is altered, what is obtained after etch and CMP: CD = 22 nm. The right image represents the results by applying an OPC on the mask design, taking the profile issue into account: CD = 27 nm, within tolerance.

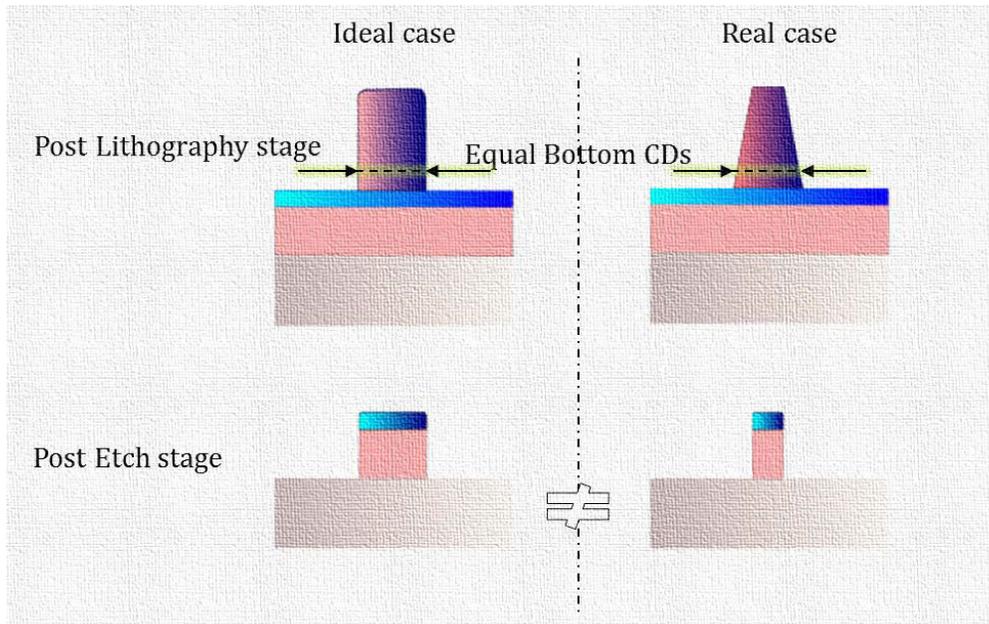


Figure 38 Illustration to highlight the importance of an accurate resist 3D model enabling the prediction of resist profile behavior that could results in deviated after etch feature from the design target

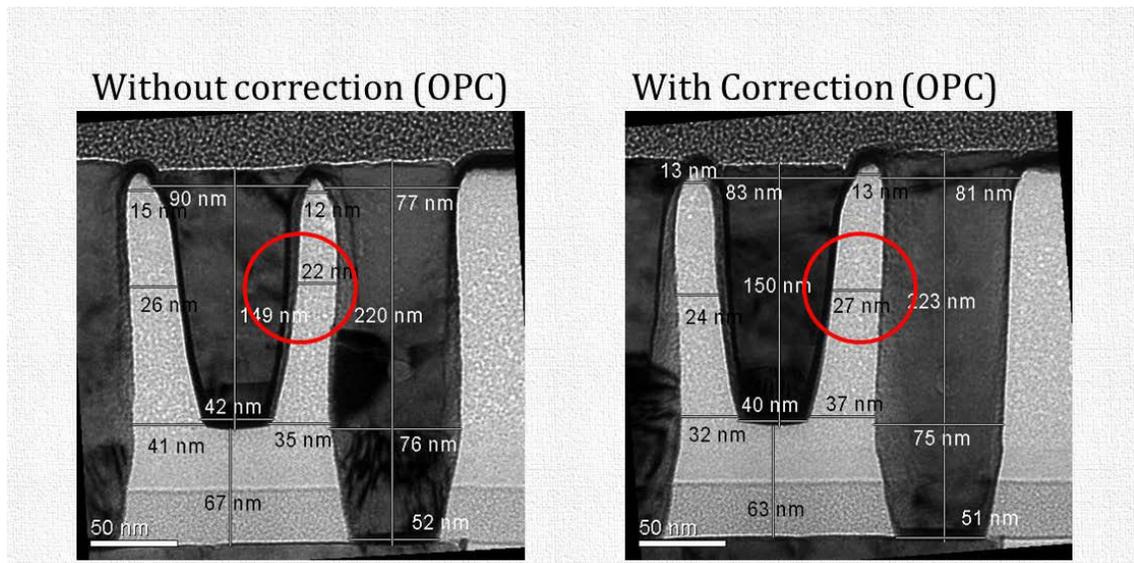


Figure 39 Example to highlight the importance of the resist profile; Post Etch (and Post CMP) cross section SEM; the light gray color represents the resist protected area

In consequence to catch these critical patterns, computational lithographic applications need to take the resist profile information into consideration. Therefore the main objective of this chapter is the construction of a reliable mask 3D model, and resist 3D model in order to achieve a more accurate predictability of the most critical patterns and the process variations.

3.2 Experimental (modeling and simulation) constraints

Except for it is expressly mentioned, the results are carried applying an ASML NXT: 1950i scanner (4X reduction immersion exposure lithography exposure system with numerical aperture NA=1.35 and exposure wavelength $\lambda=193$ nm). The model calibration and the simulations are acquired with Tachyon computational lithography platform and panoramic lithography simulation software. The thesis work is focusing on sub 32 nm node metal layers. The corresponding resist stack consist of 40 nm of BARC²⁷, 100 nm of CAR²⁸ (chemically amplified resist) resist, 30 nm of top coat, coated on a process representative test wafer. The photomask is a MoSi type mask.

3.3 Advanced 3D modeling: the mask 3D model (M3D)

In accordance to the literature and the lithography experiences in production, the primary purpose of the mask 3D model application is coming from decreased usable depth of focus (for DoF see chapter 1.3.2). It is due to the fact that the individual design structure's depth of focus mitigates with the shrinking trend of the pattern dimension. Therefore the usable DoF and the overlapping process window of the entire design become even smaller. Indeed there are numerous factors that would contribute to this decreased common DoF²⁹ (so called usable DoF, uDoF), but one of the most significant factors is the pattern dependent best focus shift. This best-focus shift deviation among different mask patterns is mainly controlled by the delta phase³⁰ between the 0th and 1st diffracted orders and pitch variation. The delta phases are varying among design patterns resulting then in a global best focus shift variation among different features.

3.3.1 The motivation for a M3D model - the phase offset related the best focus deviation

In order to demonstrate the problematic of phase deviation and so of the best-focus shift variation, let assume an example with a two-beam interference condition (Finders & Hollink, 2011) (Liu, 2011) (Wang, 2013). For that example a 1D dense line versus space is taken with a pitch p . The source is considered to be a symmetric dipole source and the planewave polarization a TE (transverse electric) parallel to the lines. The poles of the source are considered to be small enough, thus they can be taken as point sources being incoherent relative to each.

In this case the image produced by each of the represented point sources³¹ will be the result of two-beam interference in the photoresist. The best-focus shift can be derived from the total intensity (final image) calculation. Thus starting with one point source, produced total electric

²⁷ Bottom Antireflective Coating (BARC): \sim is used in order to reduce the reflection from the substrate. The antireflective coating is under the photoresist.

²⁸ Chemically Amplified Resist (CAR): \sim is a type of photoresist that is the most commonly applied for DUV lithography processes that upon the PEB will multiply the chemical reactions by the support of chemical catalysis (Mack C. A., 2006)

²⁹ For advanced technology nodes a usable DoF (uDoF) of 80 nm is considered to be acceptable in production

³⁰The phase variation is led by several contributors such as the mask topography, the imperfection of the objective lens (inducing wavefront aberrations) and the wafer stack. The total phase difference is the sum of these individual phase variations.

³¹ As we considered small enough the poles of the source so they act *as point sources*

field under the condition of the two-beam interference is determined for the two beams as follow in Equation 14 :

$$E_{total} = E_0 e^{ik_0 x + i\beta_0 \Delta} + E_1 e^{ik_1 x + i\beta_1 \Delta}$$

Equation 14

Where E_0 and E_1 are the complex magnitude for the 0th and 1st diffraction orders and Δ is the delta focus. β is a function of the optical parameters of the imaging system and equal to (see Equation 15):

$$\beta = \beta_0 - \beta_1$$

Equation 15

Here k_0 , β_0 , k_1 and β_1 are given as:

$$k_0 = \frac{2\pi}{\lambda} NA\sigma \quad \text{and} \quad \beta_0 = \sqrt{\left(n \frac{2\pi}{\lambda}\right)^2 - k_0}$$

Equation 16

$$k_1 = k_0 - \frac{2\pi}{p} \quad \text{and} \quad \beta_1 = \sqrt{\left(n \frac{2\pi}{\lambda}\right)^2 - k_1}$$

Equation 17

Where σ is referring to the point source location; λ is the wavelength of the emitting light; NA is the numerical aperture of the lens; p is the pitch of the pattern on the photomask and n is the refractive index of the medium the entire system remain in.

Then the final intensity can be written under the following expression (Manakli, 2004) (Liu, 2007):

$$I = |E_0|^2 + |E_1|^2 + 2|E_0||E_1|\cos\left(\frac{2\pi}{p} + \beta\Delta_{BF} - \varphi\right)$$

Equation 18

Where φ reveals in the phase difference of the 0th and 1st diffraction orders and $\beta = \beta_0 - \beta_1$.

The other point source's final intensity calculation is deduced likewise. Finally the total intensity (partial coherent now) of the dipole source is given as follow after assuming the symmetry behavior of the source and the photomask. It is equal to the weighted sum of the two final images of the two point sources. It is derived as:

$$I = |E_0|^2 + |E_1|^2 + 2|E_0||E_1|\cos(\beta\Delta_{BF} - \varphi)\cos\left(\frac{2\pi x}{p}\right)$$

Equation 19

From Equation 19 the image contrast (cf. Chapter 1) is derived as follow that will result in the equation for the delta best-focus:

$$\frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{2|E_0||E_1|}{|E_0|^2 + |E_1|^2}\cos(\beta\Delta_{BF} - \varphi)$$

Equation 20

The maximal image contrast is at best focus, so as consequence the best-focus deviation is now determined as:

$$\Delta_{BF} \equiv \frac{\varphi}{\beta}$$

Equation 21

Equation 21 reveals in the fact that the best-focus offset is directly proportional to delta phase occurring between the 0th and 1st diffracted orders. In addition it is inversely proportional to β (see Equation 17) so to the pitch.

Therefor there is present requirement on the delta phase prediction capability for advanced model that would predict then the best-focus offset accurately. It is important since the best-focus offset among different device patterns is one of the process window limiting factors. It reveals in reduced overlapping process window. Thus the main intention is to enable an increased capture of the pattern dependent best focus variation introduced by mask topography with mask 3D model.

3.3.2 Demonstration of the best-focus shift problematic through simulation

With regard to the pattern dependent best focus shift problematic, the main expectation of the M3D model application is to provide a more efficient best focus shift predication capability than the currently used flat model (an M2D) would do.

In order to demonstrate our expectation, an experimental case by simulation is set. The objective is to determine the best focus shift for varied pitch on dense lines, whereat the line CD and the trench CD is equal. The value of the pitch is varied from 90 nm up to 170 nm. Figure 40 shows the comparison between M3D and M2D models. The results of the M3D are provided with a 3rd order polynomial fit, since the 2D approach only takes the best-focus shift caused by the resist and the underlying stack into account – so not the mask topography related. Therefore it remains constant in opposite to the M3D (Wong A. , 1992) (Erdmann, 2005) . As consequence of this experiment showing that the best-focus vs. pitch is not constant, the pattern dependent best-focus shift cannot be predicted accurately by a conventional, currently used flat mask model.

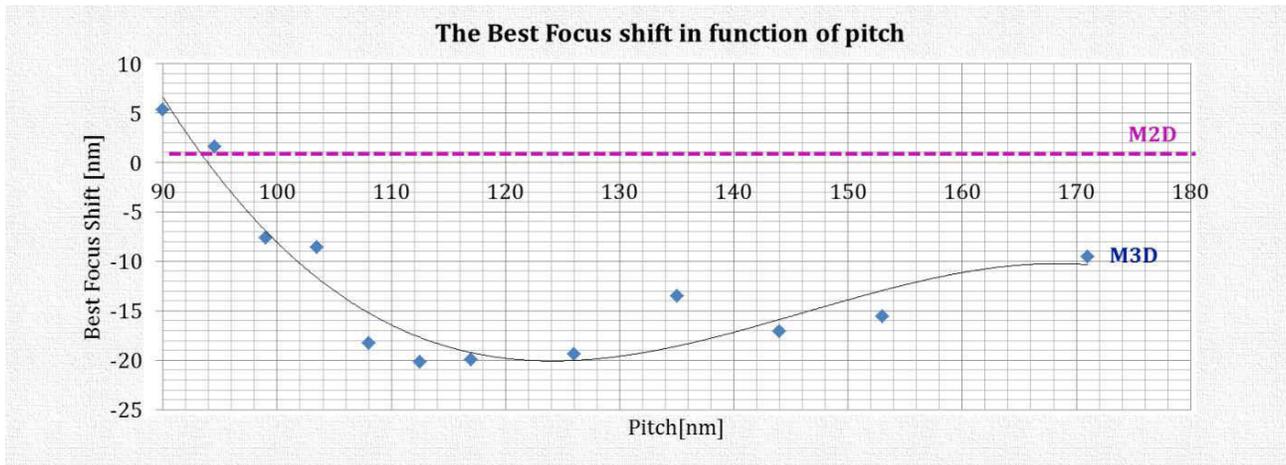


Figure 40 Best-focus through pitch (Annular illumination, 1.35NA) representing the pattern dependent best-focus shift with M2D and with M3D

Overall, as the best focus shift is one of the major contributors to the decreased uDoF, it is clear that a best focus offset aware model (an M3D) will provide a better result on uDoF prediction accuracy.

Indeed advanced nodes have reduced uDoF. For these nodes the magnitude of the phase error, thus the best focus variation is highly increased with the highly decreased pitch (Erdmann, Topography effects and wave aberration in advanced PSM-technology, 2001) (Liu, 2011). Therefore the awareness of the mentioned phase variation becomes important - the mask topography induced phase error needs to be taken into account via a well-established M3D model.

3.3.3 Results and discussion on the mask 3D model

In order to meet the previously mentioned requirements, an M3D model is calibrated. The general calibration of the applied mask 3D model (Brion Technologies) is shown on Figure 41. The model is setup with the design layout, the illumination and the topography details, similar as to a conventional flat model. In our case, the setup is proper to section 3.2. Furthermore in the general calibration flow, a mask 3D library engine is introduced. It is photomask specific (mask material, n and k values, geometries, film thickness, sidewall slope...) and it is independent on the mask layout. The mask 3D library contains the results of a rigorous treatment on the electromagnetic field (EMF) scattering effects. So it can be seen as a mask 3D engine with the task of processing any input photomask into a near-field mask 3D image. The provided 3D image is then pointed to the optical engine and finally the resist engine. In order to arrive at the calibrated model, the data fitting is completed with the wafer resist CD data (Brion Technologies, 2008) (Liu, 2012).

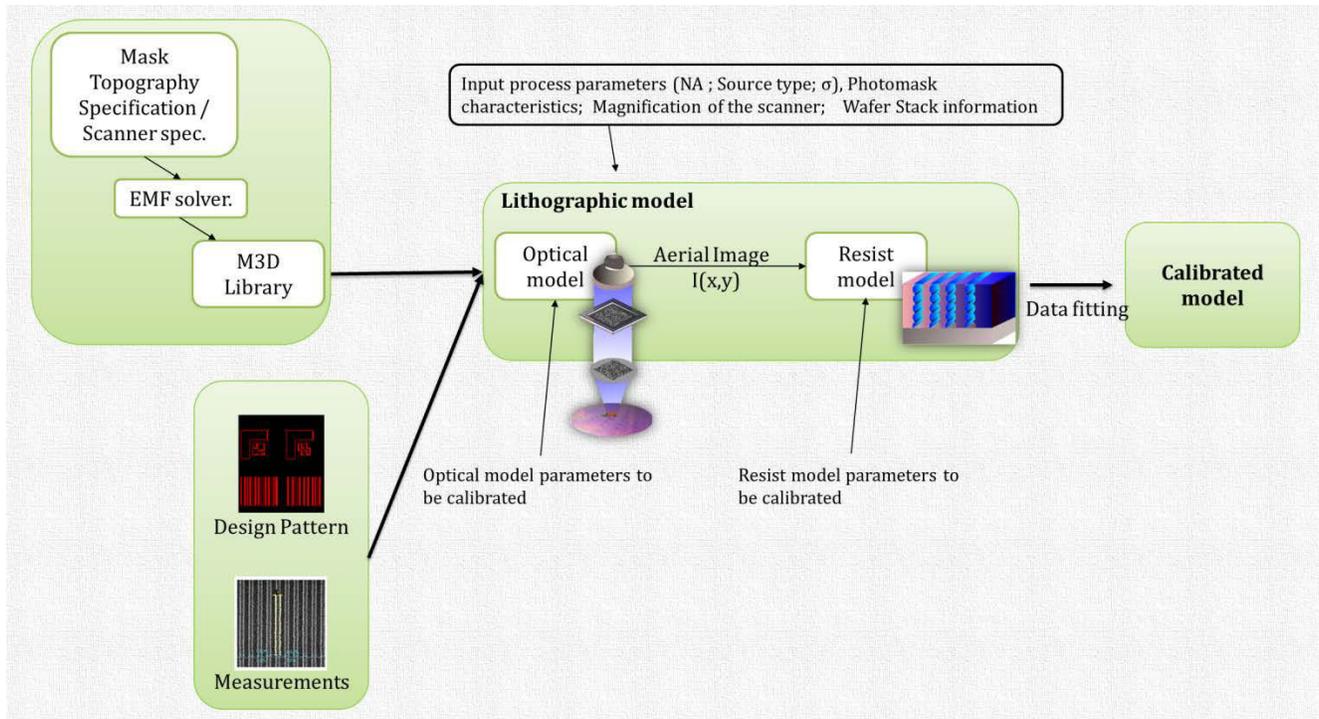


Figure 41 The mask 3D modeling flow

Our model is calibrated with an ensemble of 1380 CD-SEM measurements including: gauges through dose; gauges through best conditions (nominal dose and nominal focus); gauges for process window evaluation and gauges for contour evaluation. The total data set consists of 164 different gauges.

The calibrated model's results are shown on Figure 42 in reflect of the best-focus shift problematic, but more focusing on the DoF. The energy latitude [%] plotted as a function of the DoF [nm] is one of the most currently used representation form of the usable DoF. The energy latitude corresponds to a range of exposure energy (in % variation of the nominal one), which keeps the CDs within specified limits. On this figure two graphs are presented. The left one corresponds to the calibrated M3D results on a selected set of critical patterns indicating the usable or common DoF, which is approximately 78 nm. The right graph represents results on the Flat model versus the calibrated M3D showing the corresponding usable DoF. According to the presented results on Figure 42, the M3D model provides a smaller simulated usable DoF, due to the pattern dependent best focus shift. The delta between the two predictions is around 20 nm, which is not negligible in term of prediction accuracy. It is also important to note smaller depth of focus provided by the 3D model is closer to the real wafer data. In addition it is closer to 60 nm, which is the post etch measured usable DoF for this layer.

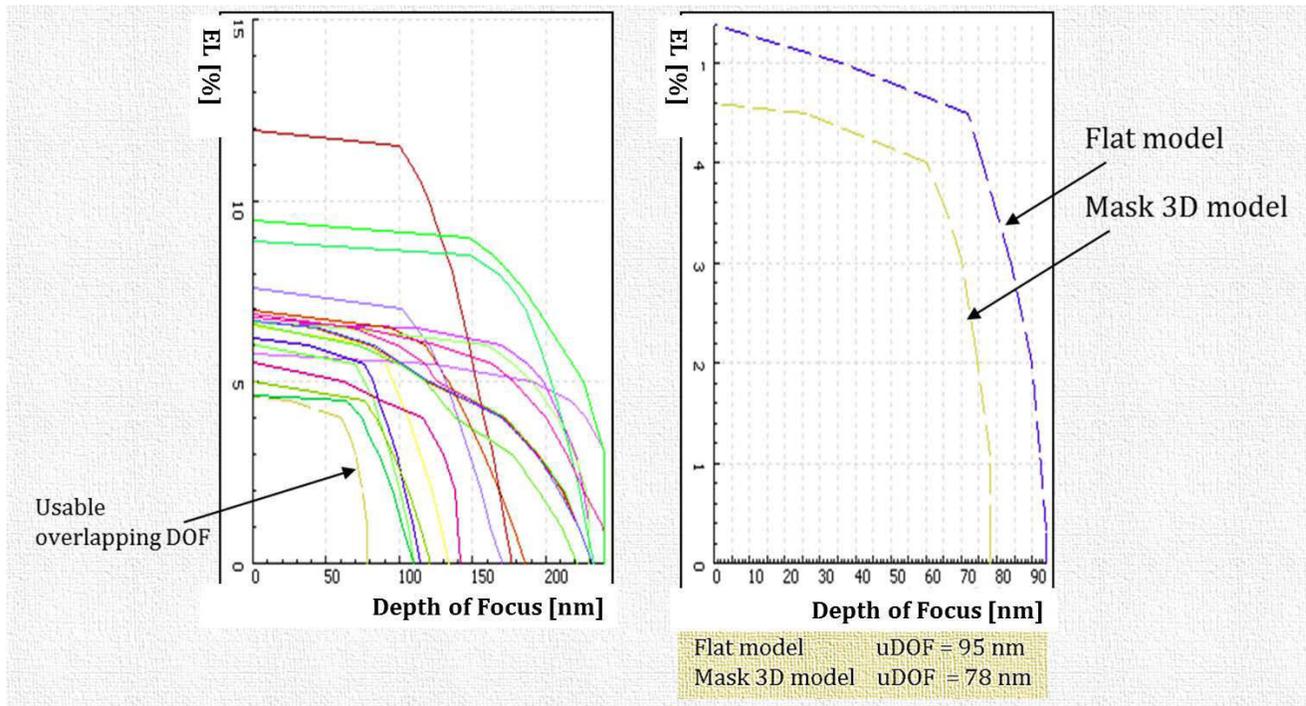


Figure 42 The Energy Latitude [%] as a function of the depth of focus [nm] on a set of selected critical patterns showing the improvement by applying a mask 3D aware model: (left) results only by mask 3D model indicating the overlapping/usable Depth of Focus; (right) results on flat and the mask 3D model indicating the usable DoF for each

As conclusion on the calibrated mask 3D model, it enables a proper detection of the best focus shift thus proper usable DoF prediction. The created M3D library serves afterwards as input data (CD-SEM) for the sequent resist 3D calibration. And the mask topography induced best focus shift will be further unwrapped in Chapter 5 describing the compensation of mask induced effects by wavefront optimization.

3.4 Advanced 3D modeling - The Resist 3D model

The sub-sequent step is the setup and the calibration of the resist 3D model. Its main purpose is to enable an efficient prediction of the resist profiles at lithography level for all layout patterns at varying focus and dose conditions. The objective is the best possible wafer exposure methods and OPC correction strategies that can be determined to avoid profile-induced hotspots³².

Conventional OPC (and LMC) models are usually applying compact 2D resist models. The input to these 2D models is a 2D optical image that is sampled in a single plane. This single plane corresponds to bottom resist, or near bottom resist height in general. Then in order to imitate the resist process, various operations are used such as the application of the threshold or Gaussian

³² A hotspot is a critical pattern for which the risk to fail at post-lithography (or post-etch) process stage is extremely high.

smoothing (cf. Chapter 2 for resist model). The result of the different operations is a 2D resist contour obtained in the initial single plane (at bottom of the resist). So this 2D model would not enable to provide information on the resist contour at different heights (Moulis, 2014). Thus in opposite to a resist 3D model, the 2D one is not capable to estimate the side-wall angles or the resist top loss. Therefore the effort is put on the resist 3D model in the following sections.

3.4.1 Possible approaches of a resist 3D model construction

A resist 3D model construction (or the resist profile acquisition) can be approached from different directions respectively to the specifications of a given process and layer. They are listed below with their comparison. Note that the challenge posed at this level is related to the acquisition of an accurate resist profile on enough representative patterns.

- Approach 1 : The contour determination based on Aerial Image Location change

The first approach is one of the most simple: the contour determination based on the variation of the aerial image location (see Figure 43). The aerial image location, or so called also imaging depth, is defined as the distance between interfaces of the first and second layer of film stack (here air and photoresist) and the imaging plane, where the image is taken for the simulation of the photoresist image without taking the reflection into account. Note that the aerial location positive sign would refer to a location inside the resist.

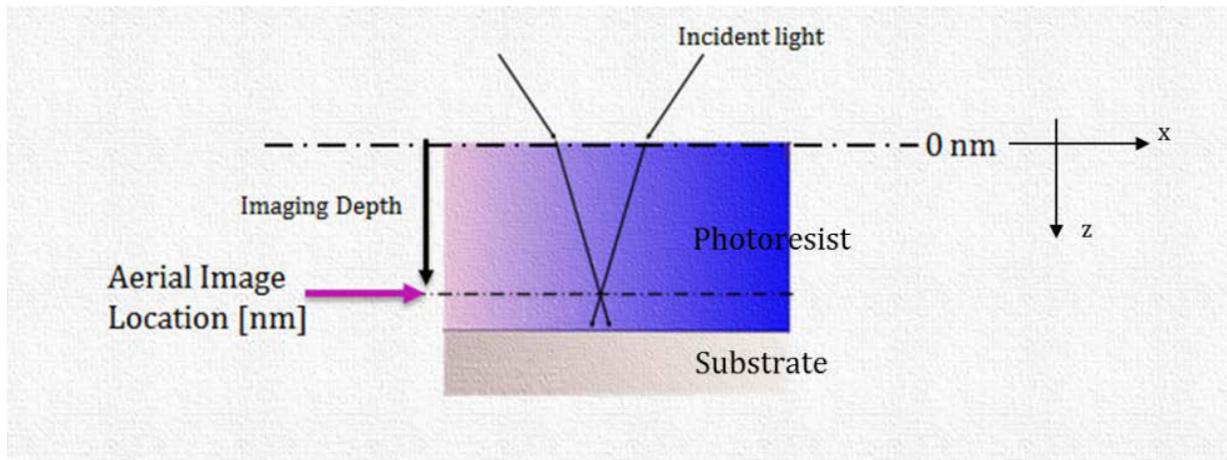


Figure 43 Simplified cross sectional view of the photoresist coated wafer illustrating the convention of the Aerial Image location [nm] – the top of the resist is considered to be zero nm

This method is capable only to provide estimation on the photoresist profile trend in a fast way. It is carried out by taking a resist 2D model (that can be a M3D for example as it will be shown later) and changing the aerial image location. The resist parameters are considered to be constant.

- Approach 2 : The Resist Profile aware modeling

Finally it is possible to calibrate and so construct a resist model based on the application of resist profile measurement data. This resist profile aware modeling is described later in this chapter, but in nutshell, it is built on a baseline model (mask 3D model) and extended with experimental 3D profile data supported by joint calibration.

- *Comparison : the “Contour determination based on Aerial Image Location change” versus the “Resist Profile aware modeling”*

For validation purpose of our choice for the resist model construction, a comparison is realized on the two above mentioned calibration methods. The choice is made by time effectiveness and available data point of view. A part of this comparison is based on the foreshadowed results as the resist profile aware model construction and calibration work is described in the following section.

The first candidate for the “Contour determination based on Aerial Image Location change” is the recently calibrated M3D model through different aerial image location. For this comparison it is named model 1. The second candidate belonging to “the Resist Profile aware modeling” is the upcoming described calibrated resist 3D model based (see later in 3.4) on profile data. This candidate is denoted model 2.

The comparison is realized by an approach shown on Figure 44. It is reflecting on the resist profile by taking the half CD for a selected cutline, a trench of a line vs. space feature. For this line vs. pitch configuration, the pitch is equal to 90 nm. This pattern is usually applied for process monitoring reasons. In order to demonstrate the results, the half CD is plotted through the resist height from zero corresponding to resist bottom. The plotted half profile CD data as a function of the resist height reveals in a quiet pronounced difference between the model 1 (denoted by red color) and model 2 (denoted by blue color). The resist profiles show similar tendency. At the middle part of the resist, around 35-55 nm of height, the model 1 has no useable results. This lack of results for this middle part of the resist height can be explained by a strong stack reflectivity. Overall, it can be concluded that at this level the resist effects are really pronounced. Therefor the requirement to take into account these effects by an accurate resist 3D model is evident.

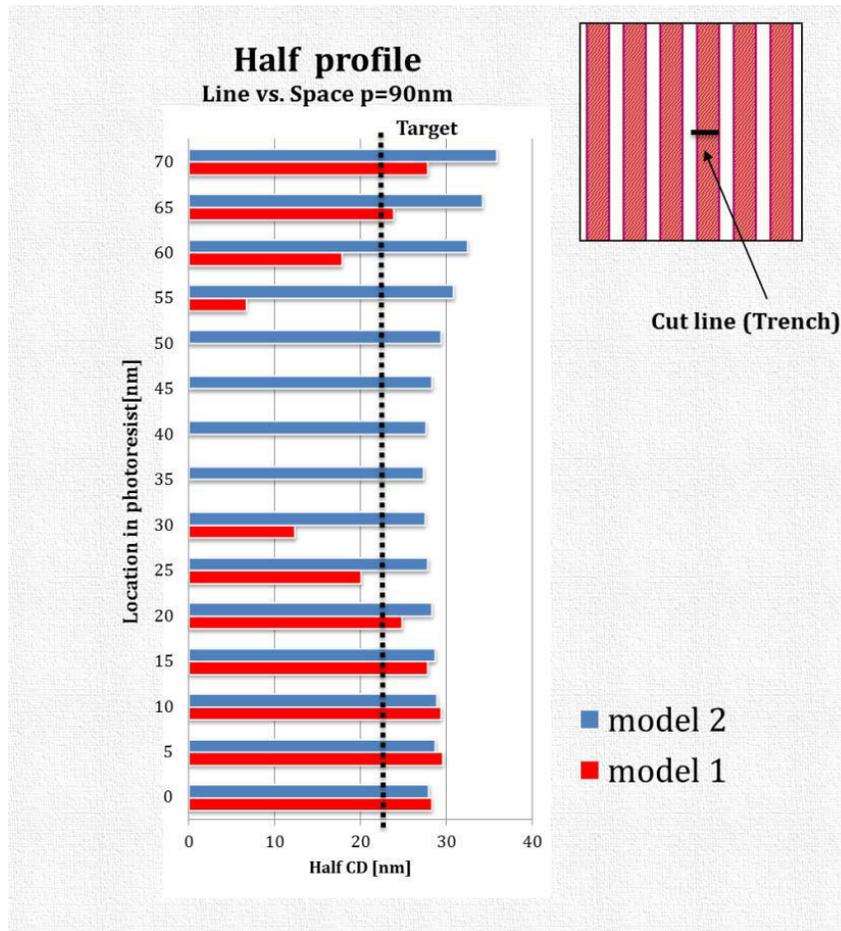


Figure 44 Comparison between the mask 3D model for different aerial image location and the resist 3D model on pitch 90 nm pattern (trench) showing the half CD [nm] through resist height [nm]

3.4.2 What is behind the applied resist profile aware 3D modeling?

Indeed the ultimate goal is to obtain a robust model with a high level prediction capability. Although as in every area of the computational lithography, one of the most important barriers is imposed by the time. It is due to the cost of full chip applications (weeks in CPU time). Therefore the formulation of a full chip 3D resist model is needed to be derived and simplified from a physical 3D resist model (Liu, 2012) (Moulis, 2014).

3.4.2.1 The physical resist model for a chemically amplified resist (CAR)

The basic equations for a 3D resist model are expressed as indicated in this section for chemically amplified resist (CAR) as it is the case for the applied experimental case. A generic resist model is described by the exposure process step, the post exposure bake stage and a final development (cf.

Chapter 1). The main objective of the resist modeling part is to arrive to describe the resist profile contour.

The generic resist model step:

- 1) **Exposure** – the exposure process is modeled by the Equation 22 and Equation 23, where PAG indicates the photo-acid-generator; “I” is the intensity of the aerial image in the resist, C is the Dill’s parameter (see Chapter 2) that reveals in the photo speed of the photoresist and H, the acid.

$$\frac{\partial[PAG]}{\partial t} = -C \cdot I \cdot [PAG]$$

Equation 22

$$[H] = 1 - [PAG]$$

Equation 23

Note that square brackets [...] mean the normalized concentration of the selected chemical component.

- 2) **Post Exposure bake – PEB** – the sub sequential step to the exposure is the PEB whereat the resist is baked in order to de-protect the inhibitors enabling the photoresist to be soluble. The PEB is a quiet complex process stage, since it includes the diffusion and the reaction between the acid (H), the base (Q) and the inhibitor (M).

$$\frac{\partial[M]}{\partial t} = -k_a \cdot [H] \cdot [M]$$

Equation 24

Where k_a is the de-protection rate of the inhibitor coming from the reaction of the inhibitor (M) and the acid (H).

$$\frac{\partial[H]}{\partial t} = -k_Q \cdot [Q] \cdot [H] + D_H \cdot \nabla^2 \cdot [H]$$

Equation 25

Where k_Q is the neutralization rate among the base (Q) and the acid (H) and D_H is the diffusion rate of the acid.

$$\frac{\partial[Q]}{\partial t} = -k_Q \cdot [Q] \cdot [H] + D_Q \cdot \nabla^2 \cdot [Q]$$

Equation 26

Where k_Q is the neutralization rate among the base (Q) and the acid (H) and D_Q is the diffusion rate of the base.

- 3) **Development** – the ending process step is the development whereat the development rate, denoted by R , can be written as a function of M , the inhibitor :

$$R = f([M])$$

Equation 27

For most of the cases and most of the applications, the rate function is based on an empirical model where the related parameters are provided by data fitting to either lithographic data or either experimental data. One of the applied models is the Mach model (Moulis, 2014):

$$R([M]) = R_{max} \frac{(a + 1)([M])^n}{a + ([M])^n} + R_{min}$$

Equation 28

Where:

- $a = \frac{n+1}{n-1} (1 - R_{threshold})^n$;
- R_{max} : development rate of the de-protected resist ($R_{max} > R_{min}$);
- R_{min} : development rate of the protected resist;
- n : developer sensitivity

Furthermore the development rate, R , can be understood as a function of x, y, z – 3D spatial coordinates. Thus $R(x, y, z)$ reveals in the development rate at the point (x, y, z) that can be calculated by the local concentration of M , as $[M](x, y, z)$. During the development, the area of interest is first the surface of the photoresist than can be expressed by the support of the Eikonal equation³³ from geometrical optics (Liu, 2012) (Khan, 1997). It relates the arrival time $T(x, y, z)$ to the development rate that results in the following equation:

$$|\nabla T(x, y, z)| = \frac{1}{R(x, y, z)}$$

Equation 29

At this point, the photoresist profile (contour) can be defined by the development time as:

$$T(x, y, z) = t_{development}$$

Equation 30

- 4) **Surface effects** – for advanced nodes the surface effects such as the surface contamination at the top and bottom of the resist have to be taken into account since they have impact on resist top loss or resist bottom footing. The surface effects can be modeled empirically by the support of a z -dependent scaling moreover the initial acid distribution during exposure.

³³ The Eikonal equation is referring to a non-linear partial differential equation usually applied in wave propagation problems (Pérez, 1996)

3.4.2.2 The applied Resist 3D model from Brion Technologies

The resist 3D model is a derived form of the physical resist 3D model that provides a faster computation by its simplifications and assumption (Liu, 2012). In order to determine the resist contour, the model starts to calculate the initial acid concentration after the exposure by linearizing Equation 22. It results in:

$$[H] = 1 - e^{-Ct_{exposure}I} \approx Ct_{exposure}I$$

Equation 31

The $t_{exposure}$ refers to the exposure time. Then the resist 3D model determines the normalized concentration of the de-protected area in the photoresist, the concentration of the soluble, [S], written as follow:

$$[S] = 1 - [M] = 1 - e^{-k_a \int_0^{t_{PEB}} [H] dt} \approx k_a \int_0^{t_{PEB}} [H] dt$$

Equation 32

Here by simplification purpose, the development is considered to be limited at the first order of e^x function at zero. Further this integration is calculated as a weighted sum of the acid at any instant of the PEB:

$$[S] = k_a \sum_{i=0}^N [H](t_i) \Delta_t$$

Equation 33

Where $t_0=0$; $t_N=t_{PEB}$ the PEB process time and $\Delta_t=t_{i+1}-t_i$. This is the base of the resist 3D model formulation. Furthermore focusing on the PEB process equations, the concentration of the acid is determined by a linear combination of Equation 25 under three different hypothesis or assumptions.

1. $[Q]=0$

Under this condition there is no base, thus no reaction decreasing [H], the acid concentration. Thus it reduces the problem to a simple diffusion problem of the acid, so Equation 25 can be reformulated as:

$$\frac{\partial [H]}{\partial t} = D_H \nabla^2 [H]$$

Equation 34

In order to solve this diffusion problem of the acid – considering that there is zero loss through the resist surface (top and bottom) – a Gaussian convolution is applied to the initial acid that will result in [H] at a given time, t_i :

$$[H](t_i) = [H](t_0) \otimes G(\sigma_i)$$

Equation 35

Where $\sigma_i = \sqrt{2D_H t_i}$ is the sigma of the Gaussian convolution kernel – or so called diffusion length and \otimes is the convolution operator.

2. $k_Q = \infty$ and $D_H = D_Q = 0$

Under this condition the base and the acid interact immediately, meaning an instant neutralization in between them without any diffusion. It means no acid and no base diffusion. Thus the acid concentration, [H], stays constant among time and can be expressed as follow:

$$[H] = \max([H] - [Q], 0)$$

Equation 36

The determination of the base concentration, [Q], is obtained similar to Equation 36.

3. $k_Q = \infty$

Under the third condition, the base and the acid interact immediately, similar to previous case, but now there is diffusion between the acid and the base. Here the acid concentration, [H] is approached by several steps. First part is the neutralization of the acid and the base and determined as in Equation 36. The second part is the diffusion that takes no chemical reaction into account. The acid and the base diffusion are considered as diffusing independently. The formulation is as follow:

$$[H](t_{i+1}) = \max([H](t_i) - [Q], 0) \otimes G(\sigma_i^H)$$

Equation 37

$$[Q](t_{i+1}) = \max([Q](t_i) - [H], 0) \otimes G(\sigma_i^Q)$$

Equation 38

Here $\sigma_i^x = \sqrt{2D_x(t_{i+1} - t_i)}$ is the sigma of the Gaussian convolution kernel and x refers to H or Q. The \otimes is the convolution operator.

Up to this point, as final step of the resist process development, the model has finished calculations on the acid (and eventually base) concentration under the three different conditions. The normalized concentration in the non-de-protected area of the resist, [M], is calculated by a linear combination. As next step, instead of determining the development rate and applying a

threshold on a development time (cf. Equation 29), the resist 3D model would consider the development is being implicitly modeled by a calibration among experimental data. In addition to that it would apply a threshold, $[M]_{\text{threshold}}$, on $[M]$.

As sequent ingredient, the resist 3D model takes into account the development effects among the resist height. It is done by applying a z-dependent (height dependent) threshold in the resist: $([M]_{\text{threshold}}(z))$ – corresponding to a vertical diffusion, together with the acid/base contamination through resist top and bottom interfaces. Figure 45 represents a simplified schema on the vertical diffusion and the contamination within the resist.

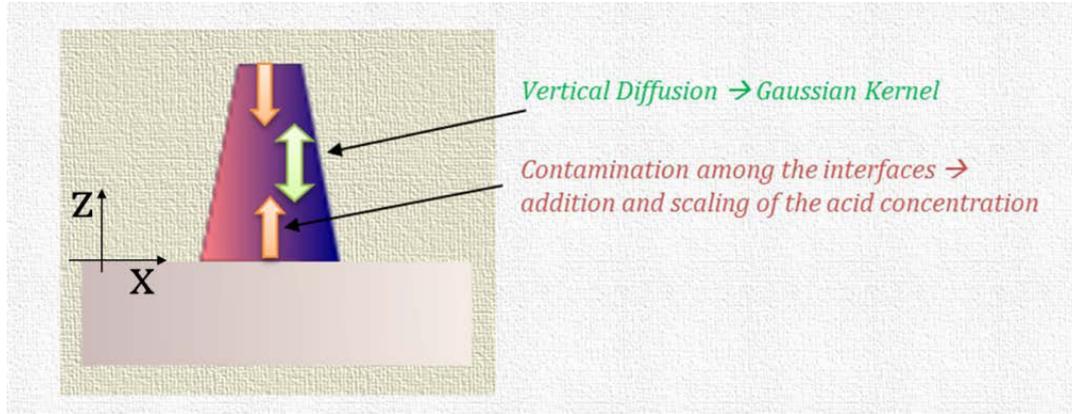


Figure 45 simplified schema on effects within the resist indicating the vertical diffusion and the contamination among the interfaces (at bottom and at top of the resist)

The contamination can modify the acid/base quantity. It is modeled empirically by a z-dependent addition or scaling on the acid concentration through a pre-defined distance from the top to the bottom of the resist height. It is as follow on the acid concentration:

$$[H]_{\text{contamination}} = \begin{cases} Sc_{\text{top}}[H](t_0) + A_{\text{top}}, & \text{if } z > z_{\text{top}} \\ [H](t_0), & \text{if } z_{\text{bottom}} < z < z_{\text{top}} \\ Sc_{\text{bottom}}[H](t_0) + A_{\text{bottom}}, & \text{if } z < z_{\text{bottom}} \end{cases}$$

Equation 39

Note: z_{bottom} corresponds to the bottom resist, while z_{top} to the top.

The Sc_x and H_x are the scaling and addition coefficient of the acid. It is similar for the base. It is important to note that the resist 3D model ignores the acid-base neutralization (no considered vertical acid-base reaction) effects among vertical direction. It considers only one diffusion problem along z. The acid/base vertical diffusion is derived from the calculation of the vertical diffused optical image, I, in the resist of the linear model (cf. Equation 31). In addition to all these,

the way the resist 3D model describes the image diffusion in the resist would also enable to handle the standing waves (Liu, 2012) (Moulis, 2014).

As final step of the resist 3D model, the parameter calibration of variable threshold is provided directly by the computer software with no need of user intervention. The parameters are set for the resist profile determining effects to solve the surface interaction and vertical diffusion. These terms are the following (see Table 1):

Parameter	Description
zSigma	referring to the vertical diffusion with a z-dependent Gaussian filter [μm]
A_top	referring to the acid addition with a z-dependent Gaussian distribution on the top of the resist [-]
Sc_top	referring to acid scaling with a z-dependent exponential function on the top of the resist [-]
z_top	referring to the effective range of A_top and Sc_top [μm]
A_bottom	referring to the acid addition with a z-dependent Gaussian distribution on the bottom of the resist [-]
Sc_bottom	referring to acid scaling with a z-dependent exponential function on the bottom of the resist [-]
z_bottom	referring to the effective range of A_bottom and Sc_bottom [μm]

Table 1 Resist 3D model parameters that are to be calibrated with description of each

The above indicated seven terms need to be set and calibrated by the user by accomplishing the main goal which is the “best” prediction capability of the model against to experimental data. It is based on the RMS³⁴. After this brief review on “what is behind the applied resist 3D model”, the next section will introduce the calibration work starting with the data preparation.

3.4.2.3 Experimental plan and the applied metrology

In a pre-phase of the model application for defect predictions and defect fixing, the model needs to be ensured to predict the image to most precise possible together with the photoresist characteristics of interest. Therefore the calibration of a resist 3D model requires profile data through the resist height in addition to CD measurements. These 3D data can be acquired from scatterometry, AFM3D (3D atomic force microscopy), cross section SEM. A brief summary on finally applied metrology methods is given here below.

The AFM3D: the working principle of the AFM3D is based on its vibrating tip that palpates the edges of the unit of inspection in order to obtain information on the profile in a non-destructive way. This metrology method is limited by the size of the trench to be measured, where the barrier is posed by tip specifications, dimensions (Landis, 2011).

The cross section SEM: it is basically working with the same principles as a CD SEM that is used to measure critical dimension by bombarding the region of inspection with electrons detecting the

³⁴ Root Mean Square (cf. Chapter 2)

backscattering of them. The only big difference is that the observed sample is movable for cross section view. Note that this metrology method remains destructive once the sample is inspected (Zeiss, 2014).

By exploiting the described metrology methods in the framework of this study, we had the following available measurements data for both calibration and validation purpose.

- 1380 CD-SEM measurements for dense, isolated, critical CD and large features to sample the desired design patterns all through dose and defocus conditions ;
- AFM 3D measurements and cross-section SEM to validate profile shapes;
- Post-etch CD-SEM images for validation purpose on critical layouts through dose and defocus to validate defect insertion points.

Figure 46 summarizes the applied metrology methods indicating the tool properties: its strength and weakness that play an important role in tool choice. At the bottom of the figure the available data are shown for each metrology. It is clear that it is important to combine metrology techniques for a relevant and reliable resist profile information at the nanometer level,. It is not negligible for the calibration of an accurate resist 3D model.

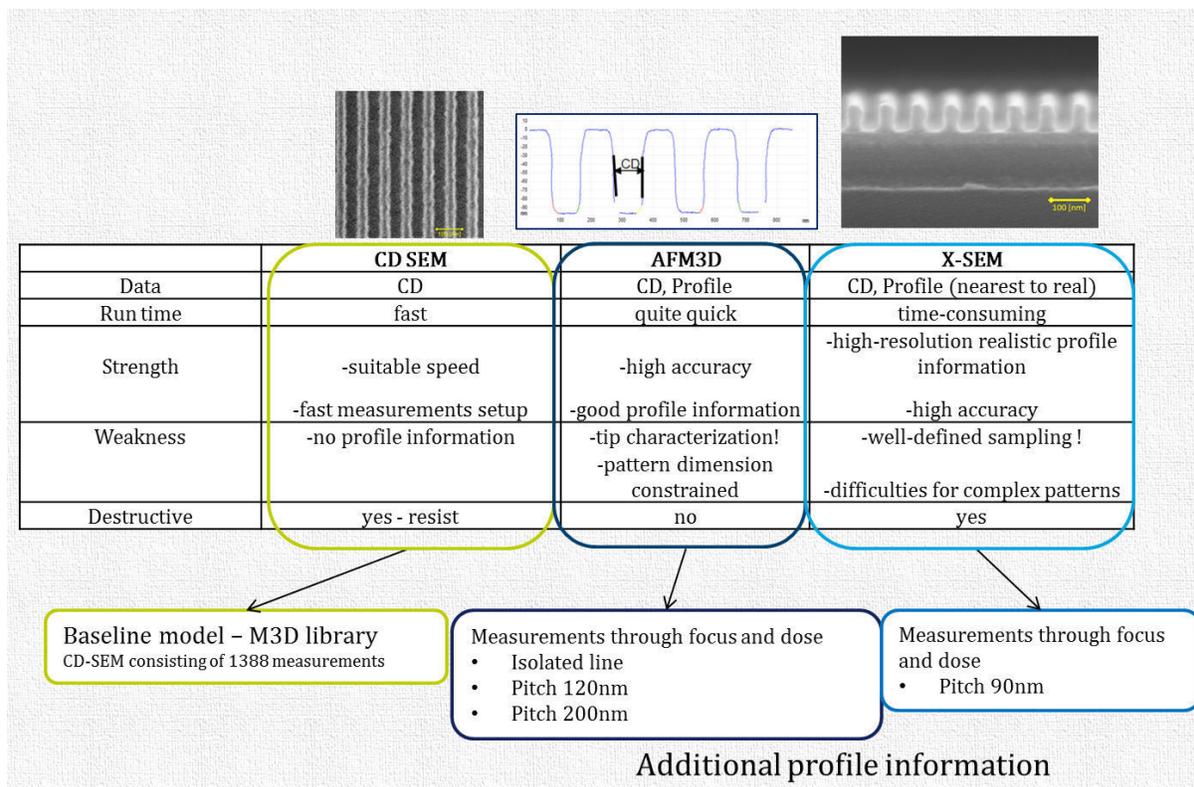


Figure 46 Applied measurements with indicated metrology tool properties and collected data for model calibration

3.4.2.3.1 The Cross Section-SEM data

For the X-SEM measurements, the emphasis is put on pitch 90 nm, since this structure is the mostly used as process control/indicator structure as this pitch is the barrier of the resolution capability for this node. Note that the well-defined sampling is a limiting factor for this technique, which requires tight attention. Thus this method becomes impossible for 2D structures coming from the fact that the cleave process is not able to cut with that high precision.

Samples are cleaved from a Focus Exposure Matrix (FEM) post-lithographic wafer. The set of samples includes cross-sections at overall three different focus and six different dose conditions using a Zeiss Ultra Plus scanning electron System (see Table 2). As it can be seen, the data exploitation is quiet a hard task as the resist lines are not that identical. Therefor different approaches are applied for the manual data treatment. Howbeit behind the X-SEM images there is a lot of solved challenges, they serve only for the validation part of the model construction due to data treatment difficulty. Although it is important to mention that these images are the closest to reality what is on the wafer.

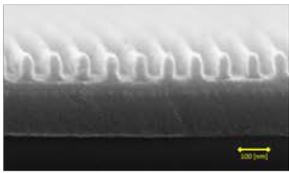
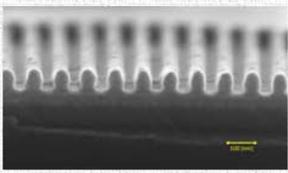
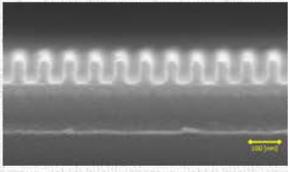
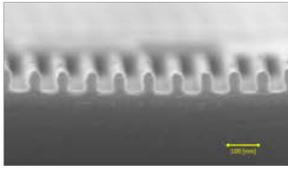
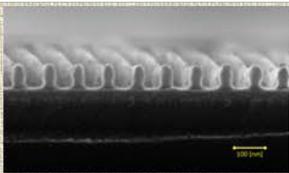
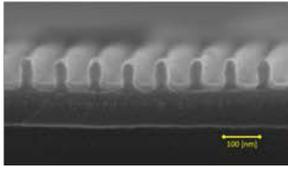
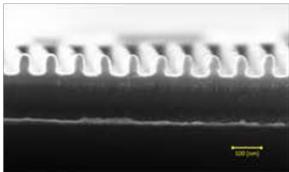
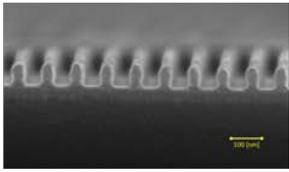
Dose\Focus	-20 nm Defocus	Nominal Focus	20 nm Defocus
0.95 · Dose			
0.975 · Dose			
Nominal Dose			
1.025 · Dose			
1.05 · Dose			
1.075 · Dose			

Table 2 Post-lithographic Cross - Section SEM samples on pitch 90 nm using a Zeiss Ultra Plus scanning electron System; missing pictures are due to sampling difficulties

3.4.2.3.2 The AFM3D data

The second applied metrology is the 3D AFM. The AFM measurement data set consists on isolated trench, pitch 120 nm and pitch 200 nm patterns through three focus conditions and three dose conditions (trenches are measured). Unfortunately pitch 90 nm, on which we have X-SEM data, is not available on AFM3D. It is due to tool limitation coming from tip dimension ($\varnothing 20$ nm).

In order to convert the obtained profile information into CD information, the top-middle-bottom CD are measured systematically at 80% (top), 50% (middle) and 20% (bottom) of the resist height as indicated on Figure 47.

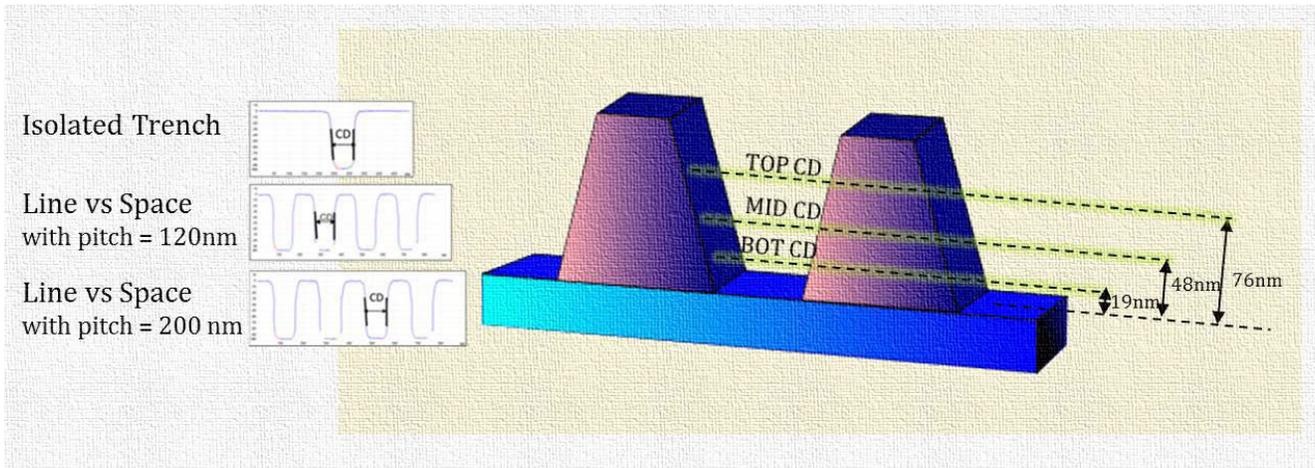


Figure 47 AFM 3D measured patterns – isolated trench, pitch 120 nm and pitch 200 nm and the illustration of CD acquirement

These measured AFM profiles provide clear and low-noise signals from which CD values (at bottom, mid and top of the resist) can be extracted fairly easily versus other profile metrology data. So according to these measurement findings the AFM3D data are used predominantly as source of input for the resist 3D model-calibration. And the X-SEM data remains complementary profile information.

3.4.2.3.3 The calibration and its validation

So, as described, the resist 3D model calibration is an extension of conventional 2D OPC model calibration platform. The calibration flow relies on the previously generated mask 3D Library serving as bottom CD data, which will provide the baseline model. Figure 48 shows a simplified schematic view on the data preparation flow as it is established by indicating the attuned different measurement data: CD-SEM; X-SEM and AFM 3D. Together with defocus and dose, resist height is used as a process window space dimension. The final model optimization is done using linear solver to minimize CD RMS.

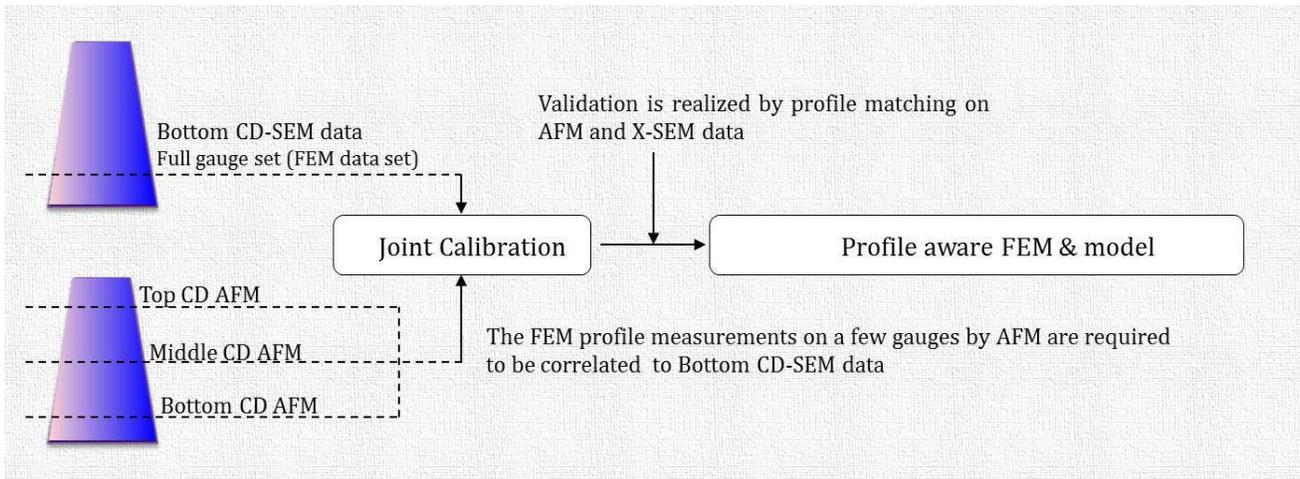


Figure 48 Overview on the calibration flow for the resist 3D model

3.4.2.4 The metrology based model validation

As described in the data preparation step, the bottom AFM CD is biased toward the baseline model CD. During the entire calibration the attention is emphasized on the profile fit. The Figure 49 shows the simulated profile for pitch 200 nm overlaid with AFM CD and with cross-section SEM. Both profile matching reveals that the simulated profile is in good agreement with real wafer data.

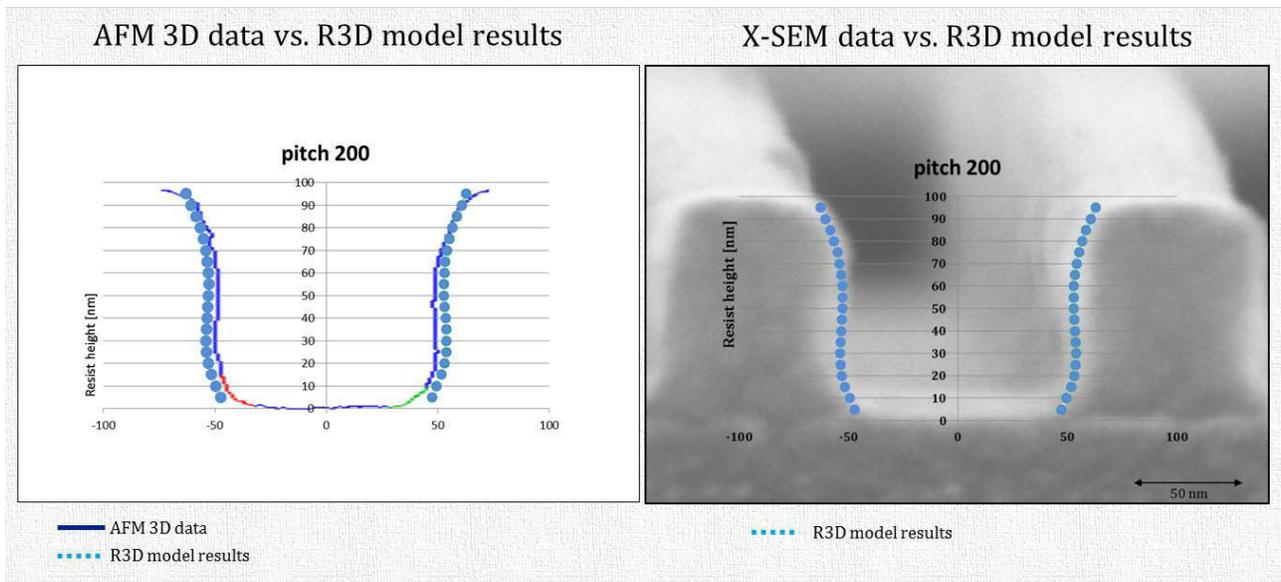


Figure 49 Resist profile for pitch 200nm illustrating the match between the Wafer AFM profile solid line vs. resist 3D model filled dots(left) and the Cross section SEM vs. resist 3D model filled dots (right); Note : the red and the green curves are part as they could not be measured with AFM tip

It is important to highlight that with the calibrated resist 3D model, the simulated resist profile of pitch 90 nm overlays very well with the corresponding X-SEM image (Figure 50). In point of fact

even if the X-SEM profile data shows no regular shapes on that pattern the matching can be considered good. In addition, the model predicts accurately the observed resist loss on wafer – which is 30 nm of magnitude. Therefore it can be concluded that the predicted resist profile has excellent matching with the X-SEM measurement. Thus the validation on these simple patterns makes the opening for further validation on structures that are not included in the calibration data set.

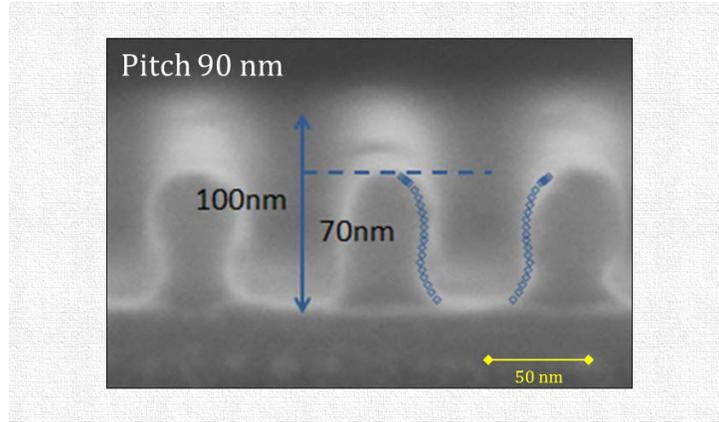


Figure 50 Results on pitch 90 nm by overlaying the resist 3D model profile on the corresponding X-SEM image

3.4.2.5 Validation of detection capability of the calibrated resist 3D model on real, critical structures

As mentioned the calibrated resist 3D model is now verified and validated on the available resist profile data, the next step is to check it on real critical patterns, or so called hotspots. For this reason the resist 3D model is applied in order to check on weak points that can be highlighted in term of strong resist slope effects. For instance several hotspots as results of a Process Window Qualification (PWQ)³⁵ analysis were subjects for this validation test. Furthermore they are identified also as most critical patterns by the usage of appropriate CD detector settings on a small clip verification run with resist 3D aware LMC (see Chapter 2 for LMC) at the top of the resist. Note that a standard model without 3D effect awareness would not detect these patterns as critical structures.

As a test case on the calibrated resist 3D model a critical pattern showing strong resist top loss is taken. The simulated contours are observed for the bottom (20% of total height); for the middle (50% of total height) and for the top (80% of total height) of the resist height. Figure 51 shows the simulated contours versus post-lithography and post-etch. At post-lithography level the top loss is not visible, but the additional post-etch (and post CMP) defectivity CD-SEM data/image highlights it. Thus Figure 51 results in the fact that the resist 3D model enables a good matching between simulation and defectivity (CD-SEM).

³⁵ The PWQ flow is a well-established wafer inspection method coupled with design based binning algorithm in order to determine the lithographic critical structures' process window within a lithographic dose and focus matrices. It is applied for OPC quality check, for new mask quality check and for lithography process check. For metal layers it is done "After Etch Inspection" (AEI) and the "Chemical Mechanical Planarization" (CMP) step.

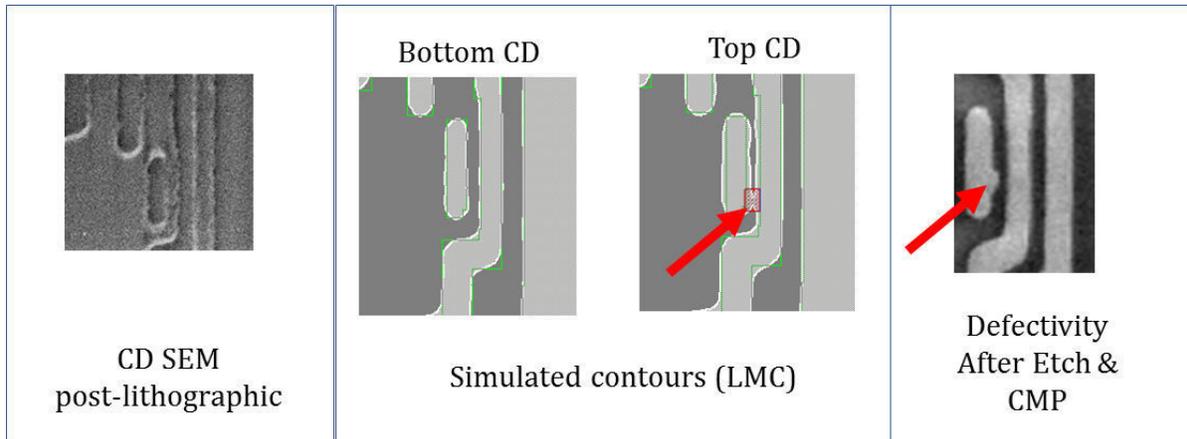


Figure 51 Result on a critical pattern showing the resist 3D model capability that would enables to highlight critical areas that would not be caught without the 3D notion. It provides a good matching between simulation and defectivity

3.5 Conclusion

The requirement is growing for accurate 3D modeling for today's nodes. As seen, the conventional 2D modeling is become obsolete for advanced technology nodes, for which the mask and resist (and wafer) 3D induced effects are not negligible anymore. Therefor the constructions of reliable and efficient 3D models are needed. In the light of this problematic, this chapter is providing the construction of a mask 3D model and a resist 3D model from their background until the verification step. The calibrated models are shown to be efficient and accurate enough, which is validated on experimental data for 28 nm technology node metal layer. These models are now capable to provide an accurate support for the detection of the most critical features that will be presented in the next section of this dissertation work.

4 Chapter

Detection of the most critical structures assisted by the resist-profile aware methodology

In the previous chapters, the context of the thesis work and the important notions for this work has been presented. It mainly focused on OPC models' upfront process variation prediction ability and the construction of advanced 3D models. The sub sequent step now is to select the most process sensible patterns, supported by the simulation.

The layer of interest within this thesis is metal layer. Howbeit, it mainly includes 1D features, the failing ones are complex 2D structures in which the task's challenge lies. These patterns have a large deviation in CD uniformity derived from mask and resist 3D effects.

In order to set the direction for the detection methodology's groundwork, the emphases is on the mask topography and resist profile related approaches. Then, in this chapter the problematics of detection for both mask topography and resist profile induced effects are presented. Solutions, with different approaches are proposed and finally discussions on results are developed. Although note that in the framework of this study, the focus is more on the resist related effects. The detectors are based on simulation data and mainly used LMC, a model based verification tools already presented in chapter 2. Simulation data are validated by SEM measurement, especially for detection approach among resist profile induced effects. The studied patterns are critical features.

4.1 Detection methodology

In order to setup a detection methodology, “behind effects” need to be taken into account. These effects can be related to mask or resist 3D induced phenomena as seen previously (cf. Chapter 3). They cannot be neglected anymore at this shrunk node level. Hence a good understanding on these phenomena is a useful support. It would lead to a reliable detection methodology. Note that the detection methodology can implicate simulation and non-simulation based detectors³⁶.

Under the simulation based, the LMC tool is used. Models used in LMC are 2D models and currently used in production environment. For the work of this thesis 3D models are required, and then developed with associated detectors. Here only the simulation based one is discussed, because this work relies on the capabilities of computational lithography. In the following sections, two possible detection approaches will be described providing reliable detection methodologies. One is related to the mask 3D effects, while the other is related to the resist 3D.

4.2 Detection approach among the mask topography induced effects

4.2.1 The problematic among the mask topography induced effects

As first approach, the mask related one is introduced. The idea, to setup a “smart” detection method, is referring to the pronounced best focus offset among different patterns for that layer. The behind effects are related to the finite thickness of the photomask that leads to phase errors in the diffracted orders appearing in the best focus offset. In the literature, the best focus variation is indicated in the range of 40 to 60 nm (Finders, 2014)(see Chapter 3). Besides the mask topography, the resist contributes also a small portion in the best focus offset through patterns. It is attributed to different refractive index. Howbeit, it remains negligible compared to the mask topography driven offset.

4.2.2 The proposed solution: mask topography conducted detector

Under the mask effects, the detector’s objective is to catch all the patterns showing large best focus shift and then analyze them. For this reason the main idea is to setup a simulation based detector that catches patterns with high best focus shift. The simplest way is to focus on the CD through focus response: the determination of the Bossung curvature slope (see Chapter 1 for Bossung curve definition).

The basic of the slope determination is illustrated on Figure 52, whereat possible curve trends are indicated for line CD and for trench CD examples. For each example, the ideal Bossung curve is shown at left hand side. These ideal curves are considered to be well centered, with a Best Focus³⁷ Shift within 10 nm of tolerance. Then, the three other curves for each case represent the possible non-ideal cases: the shifted (positive and negative) and non-parabolic, shape Bossung curves. For these curves the absolute value of the BF shift is higher than the tolerated 10 nm. The slope of the

³⁶ Under the non-simulation based detector the following examples can be listed: The Mask Rule Check (MRC) or The Overlapping check

³⁷ Maximum or minimum of the Bossung curve

Bossung curve is characterized by s_1 and s_2 . These s_1 and s_2 calculation areas have the same width (Δf_0) and are adjusted depending on set of features analysis. In the case of ideal bossung curve s_1 and s_2 are centered on the best focus of the model. Non ideal Bossung curves are shifted from this best focus (M3D effect). In order to check the undesired or reluctant results, the sign of $s_1 \times s_2$ are inspected. It means that e.g. for line CD, we consider they have parabolic shapes. So if s_1 is a negative value, s_2 should be a positive one. If it is not the case, the Bossung curve has a non-regular shape. Indeed for specific cases, we can check the curve shape by a cut-line³⁸ based Process Window Analysis (PWA).

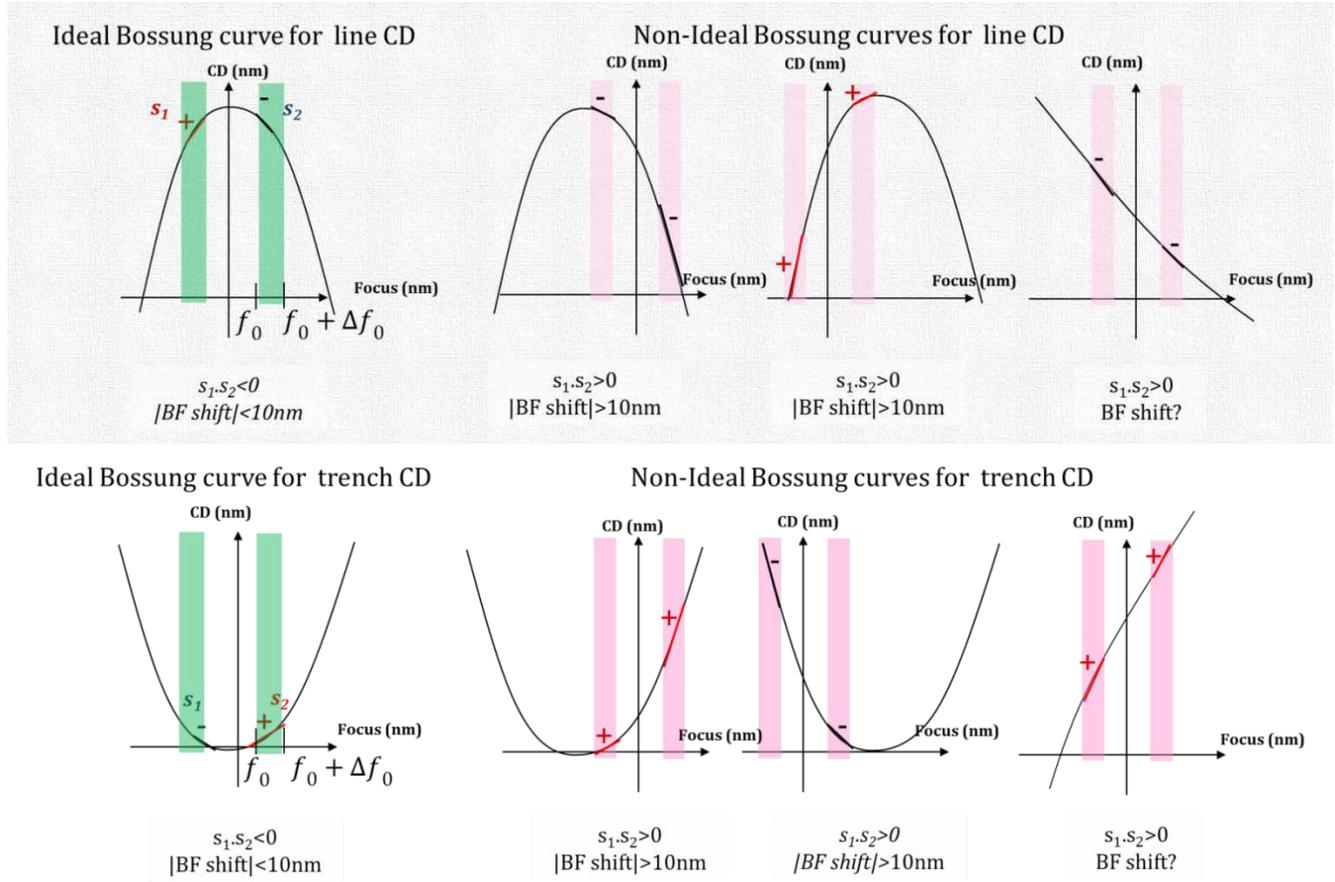


Figure 52 Schematic representation on the Bossung curve analysis focusing on its slope by showing different possible curves for line cut-line (CD) and for trench cut-line (CD). The ideal Bossung curve is indicated (left) for each

In order to realize the above mentioned idea the best focus shift detection, we used the capability of an LMC³⁹ job to select the most potential best focus shifted patterns on the design layer. It is possible by applying the existing Contour-to-Contour comparison LMC job: a model-based contour-to-contour check focusing on the CD variation in nm for different process conditions (see figure 53).

³⁸ A cut line refers to the position that we use for the simulation, the modeling or the Process Window analysis...

³⁹ An LMC, as described already, is a model-based verification tool that can be used on a reduced surface as well as on full chip across the full process window (Brion, 2015). For more details see Chapter 3.

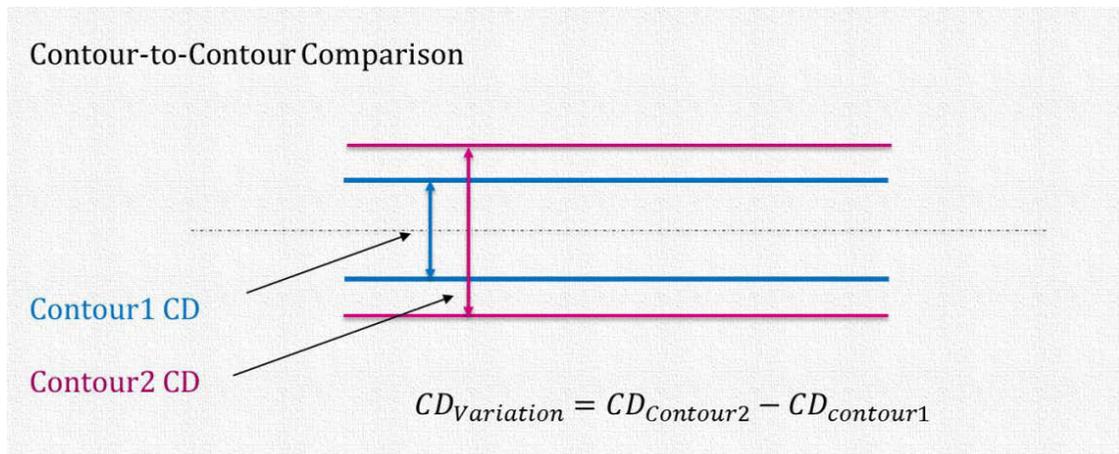


Figure 53 Representation of the Contour-to-Contour Comparison

In our case, as we are searching to analyze the Bossung slopes, s_1 and s_2 . It is determined in a symmetrical way based on the delta CD at different focus position: $\pm f_0$ and $\pm (f_0 + \Delta f_0)$. Therefore we check the following conditions:

- For slope, s_1 : the delta CD in between the conditions : $-f_0$ and $-(f_0 + \Delta f_0)$;
- For slope, s_2 : the delta CD in between the conditions: f_0 and $(f_0 + \Delta f_0)$.

Once all is setup and the detector is launched on a set of patterns, it is necessary to check the detected patterns by a process window analysis (see Chapter 3). This analysis enables to plot all the Bossung curves (of the highlighted patterns, gauges) for shape verification. In addition it is resulting in the overlapping process window that is very useful support to check whether these newly found patterns are really process window limiters or not.

4.2.3 Results and discussion

In order to test the above set LMC based detector, it has been applied to a set of patterns. It includes a thousand of different features representatives of the layer of interest. After a zoom on the detected structures, patterns on board of the best focus shift margin are taken and are further analyzed. The results are represented on figure 54, by the Best Focus shift [nm] in function of structure names. Two examples from our findings are shown by their snapshots indicating the cut lines – the best focus shift positions.

According to the results, a delta best focus of 30 nm is present for this set of patterns. Remark: the magnitude of the found deviation by this best focus shift detector is coherent with the 27 nm of delta best focus introduced in Chapter 3 by Figure 42 on the pattern dependent best-focus shift.

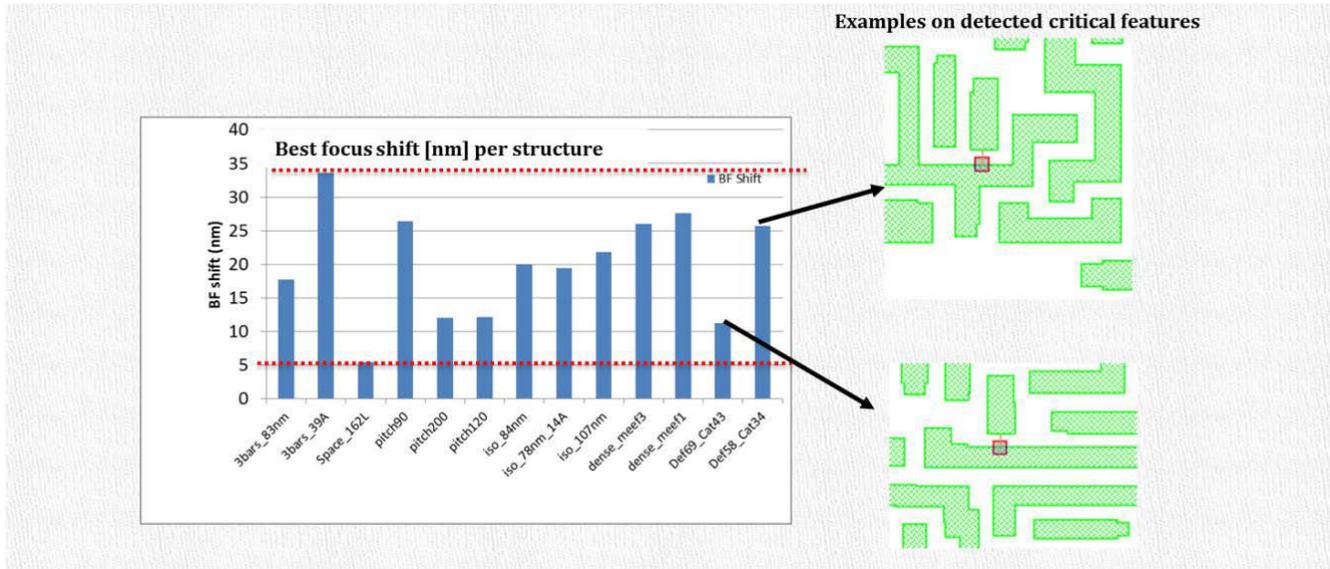


Figure 54 best focus shift detector results on a test case showing the best focus variation [nm] per structure per features

At this point we should go back to Chapter 3, to notion of the defectivity analysis using Process Window Qualification⁴⁰ (PWQ). It is because of the fact that we are searching, by the support of this detector, to catch all patterns that would limit our lithography process. It means that our findings should match with the results from the defectivity. Our results are promising, since our detector provided matching value. Although it should be mentioned, that there is a limitation. It is due to the fact that, even if the uDoF of the detector and the uDoF of the PWQ fit well, the limiter hotspots are not identic. Thus, further investigation is required to smooth our detection method, which is added to the list of thesis perspectives.

Thus, if we consider all the results into account, our detector is validated. But beside the reliable capability of the detector, it is not well-marked that patterns with stronger best focus shift are the process window limiters. It is based on a complementary overlapping process window analysis on these patterns. The limiters are more features with small, or zero process window (further discussion on this topic is given at the profile related part).

Overall, it can be concluded that this detector works well on this set of structures. But in point of fact, this best focus shift approach could have been further developed on the most critical site in the exposure field for this layer. It refers for example to a development how to handle patterns with tilted or strongly flat CD response through focus, which is one of the weaknesses at this point of the discussed detector.

Despite of the fact that the layer of interest has a large set of patterns owning pronounced best focus shift responses, the test process window analysis revealed more in the resist profile

⁴⁰ The PWQ flow is a well-established wafer inspection method coupled with design based binning algorithm in order to determine lithographic critical structures PW within lithography dose and focus matrices. For metal layers it is done "After Etch Inspection" (AEI) and the "Chemical Mechanical Planarization" (CMP) step.

importance. It came from the fact that the process window limiting factors represented more patterns with tight process window related to resist profile compoment and not best focus shift. Figure 55 shows an example through a PWQ found hotspot reflecting on the fact, that profile has a high important role.

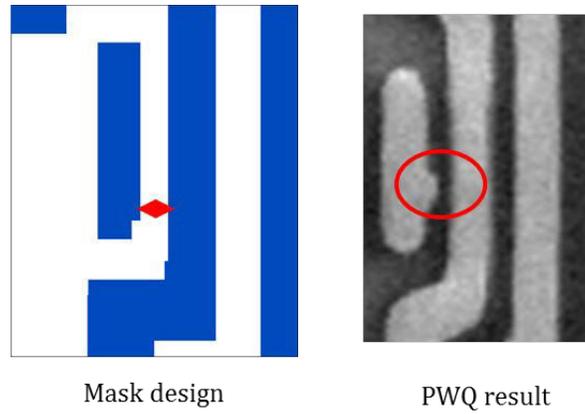


Figure 55 a PWQ (Process Window Qualification) hotspot SEM image indicating the mask design by left crop. Note that on the mask design the white color represents the photoresist coated areas.

It means that the resist 3D induced effects are more present than the ones coming from the mask 3D effects leading to best focus shift⁴¹. Therefore our main point of interest is linked to the resist profile that is presented in the next section.

4.3 Detection approach among the resist profile induced effects

4.3.1 The problematic among the resist profile induced effects

In the literature, as well as in practice, the resist profile is more and more concerned for advanced nodes. As already mentioned it is due to the fact that the profile for the shrunk nodes is not any more negligible – means that it cannot be considered simply as straight. Along this work, a profile change trend is first identified on experimental scatterometry data for this layer on a process control pattern, a line vs. space (pitch 90 nm) structure as shown on figure 56. Here the calculated profile sensitivity to dose ($\Delta CD/\Delta E$) is plotted versus dose. The examined exposure energy field covers a variation from -10% to 5% from the nominal dose. It is clear that the resist profile is not constant through dose. Besides that, the top, middle and bottom part of the profile shows different evolutions through dose variation.

⁴¹ Best Focus shift among different patterns has been reported in numerous publications. The global message: the best focus shift can be affected by many factors, but the mask 3D or mask topography is the leading cause.

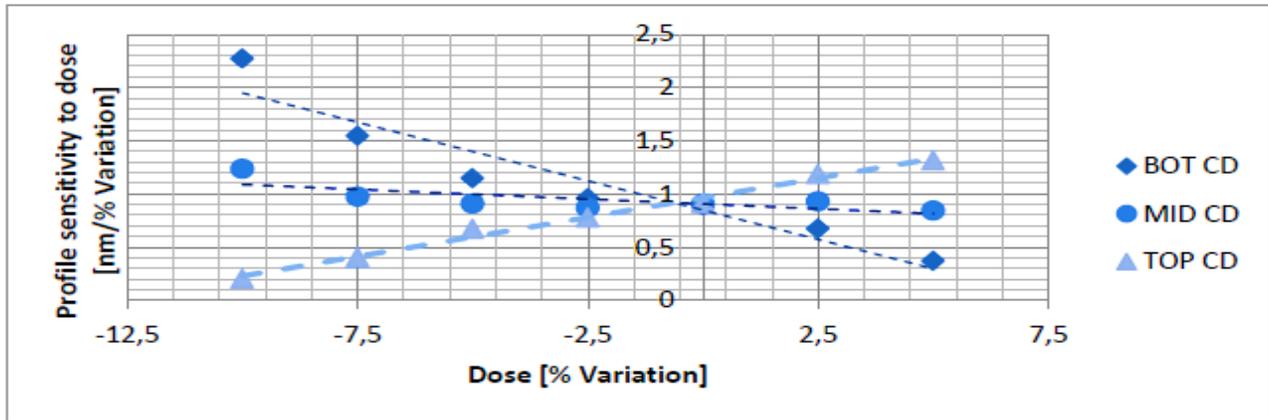


Figure 56 calculated profile sensitivity to dose. Scatterometry data for top-middle-bottom Line CD for line vs. space pattern among dose at nominal focus (Finders, 2014)

The sensitivity of the profile ($\Delta CD/\Delta E$) is further examined. It can be said that it is an interesting indicator for critical features having small CD. It means the higher is the sensitivity, the less robust the process is. As example, when the profile shows non-straight shape compartments it indicates process limitations coming either from the mask blank, either the stack reflectivity or the illumination choice etc.

Further through the scatterometry data come-off, Figure 57 illustrates along a schematic sketch the difference between the ideal and real resist shape through dose. The trend of the changing CD to dose and the resist profile sensitivity is represented by simplified curves indicating variation for the top-middle-bottom. Note that these curves are here only to provide a summarized representation on the profile compartment.

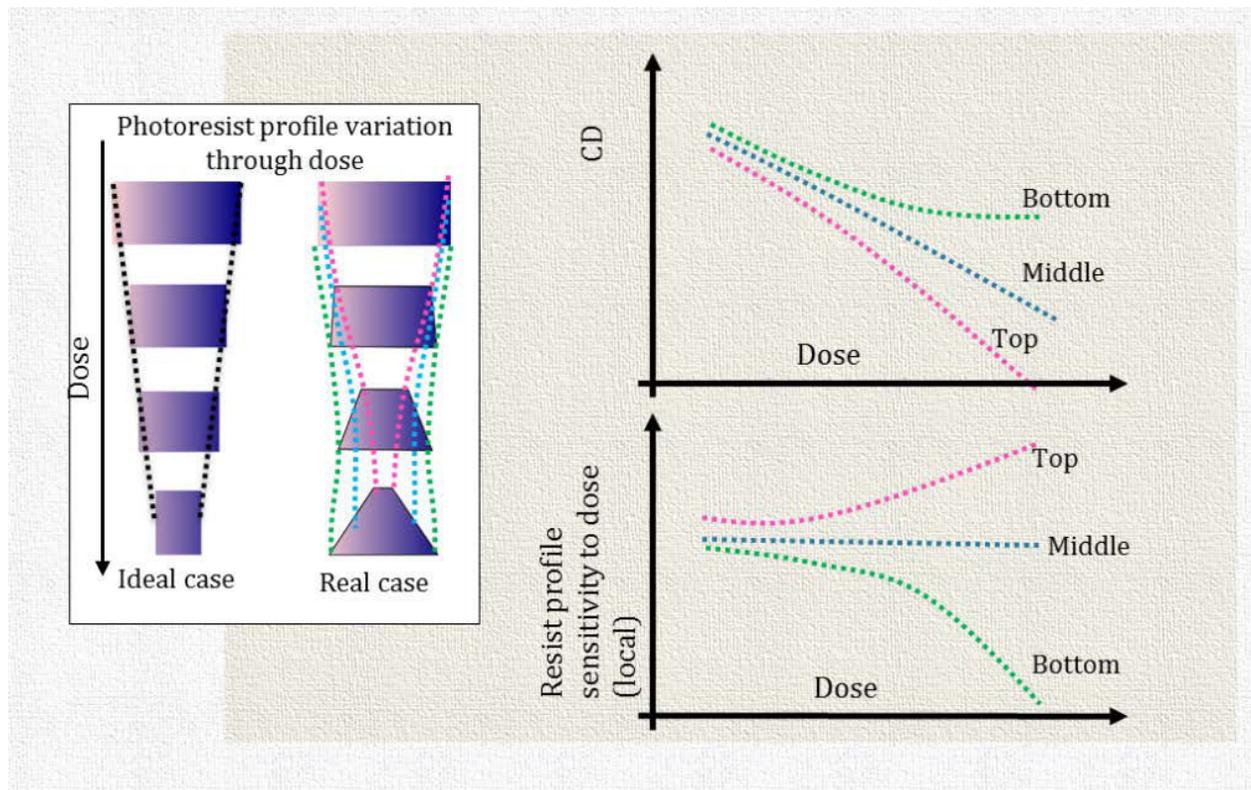


Figure 57 Illustration of the resist profile changing through dose means: different compartments of top-middle-bottom of the resist through dose

For verification purpose on this changing profile concept, the calibrated R3D model is a good support. It is applied on two candidates. One is the line vs. space pattern (pitch = 90 nm), whose importance reveals in the fact that this pitch is on the boundary of resolution of the layer of interest. And this pitch 90 nm is the smallest dimension authorized by the design rule for this layer. The second candidate for the preliminary simulation case is a pattern that represents one of the most critical structures from the PWQ defectivity analysis, also called hotspots.

The results are focusing on the resist contour only at top and bottom height (see Figure 58). The calculated profile sensitivity to dose is plotted versus dose (nominal $\pm 10\%$ dose deviation). The selected hotspot shows twice more pronounced sensitivity than the 1D pitch 90 nm. It means that even at nominal dose there is a critical issue for this pattern. Indeed it is coming from the fact the process is not that robust and it has high top loss effect. So the preliminary simulation results proved the profile tendency.

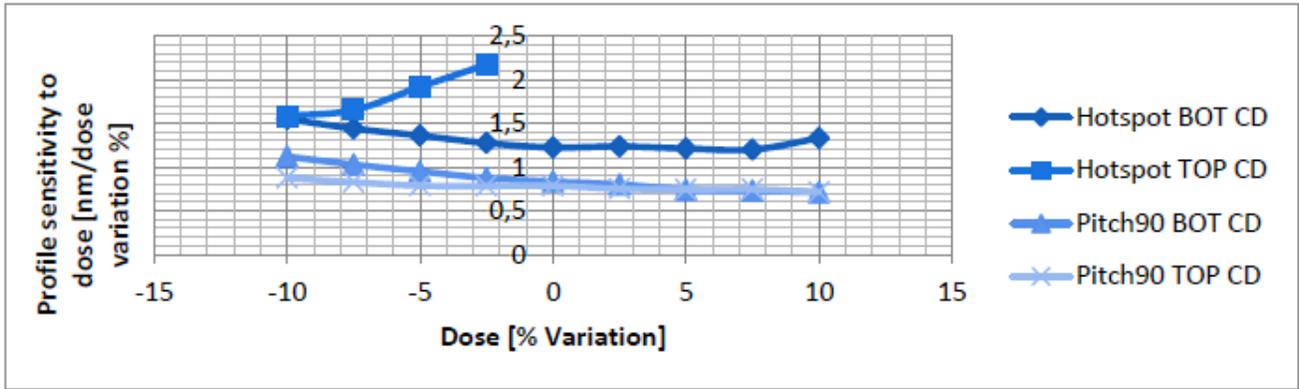


Figure 58 calculated profile sensitivity to dose. Simulated data of top and bottom resist among dose (nominal $\pm 10\%$ variation) at nominal focus

Based on these results through dose, the accuracy of the prediction can be improved with the knowledge of the profile sensitivity $\Delta CD/\Delta E$ to dose. This profile related item can lead to improvement on process control providing an analysis metrics: the resist profile sensitivity to dose. It seems a good support to enhance the most critical pattern detection that will be unwrapped in the following section by the introduction of our methodology for critical pattern detection (on full chip). So in order to evaluate these learnings that are reported on wafer data, the resist profile is put into reflection providing a good support for the detection.

4.3.2 The proposed solution: resist profile conducted detector

By looking deeper the resist 3D induced problematic, two primary approaches enable to obtain detection solution:

- Approach A: application of a Contour-to-Contour (C2C) comparison and inspection. It means a delta CD comparison, for which *the higher delta CD [nm] corresponds to the most critical pattern.*
- Approach B: application of a bridge detector, as in standard LMC (cf. Chapter 2), where *the smallest bridging value [nm] refers to the worst hotspot.*

Approach A: the Contour-to-Contour detector

The basic idea of this methodology among approach A consists in assuming a Tachyon Lithography Manufacturability Check. Since standard LMC disposes only of the Bottom CD as information, it is combined with calibrated resist 3D model. The detector main purpose is a contour-to-contour inspection at different height of the resist. It is realized by comparing the contours through the height. Here all the patterns are detected for which the CD difference is higher than a certain tolerance. The tolerance values are corresponding to those in production. The CD difference reveals first in the difference in between the Bottom and Top CD contour values that would provide patterns with pronounced Top Loss. Of course the other type can be tried as well, where the line collapses are detected: $CD_{\text{Bottom}} < CD_{\text{Middle}}$.

Approach B: the resist profile aware bridge detector

The approach B is staying with the basic profile related concept as well. It turns back to a standard bridge detector that highlights all patterns with a bridge failure, patterns for which the bridge CD value is out of tolerance. Figure 59 represents how the bridging detector works, whereat the contour CD fails. Such a failing hotspot example is shown on figure 60.

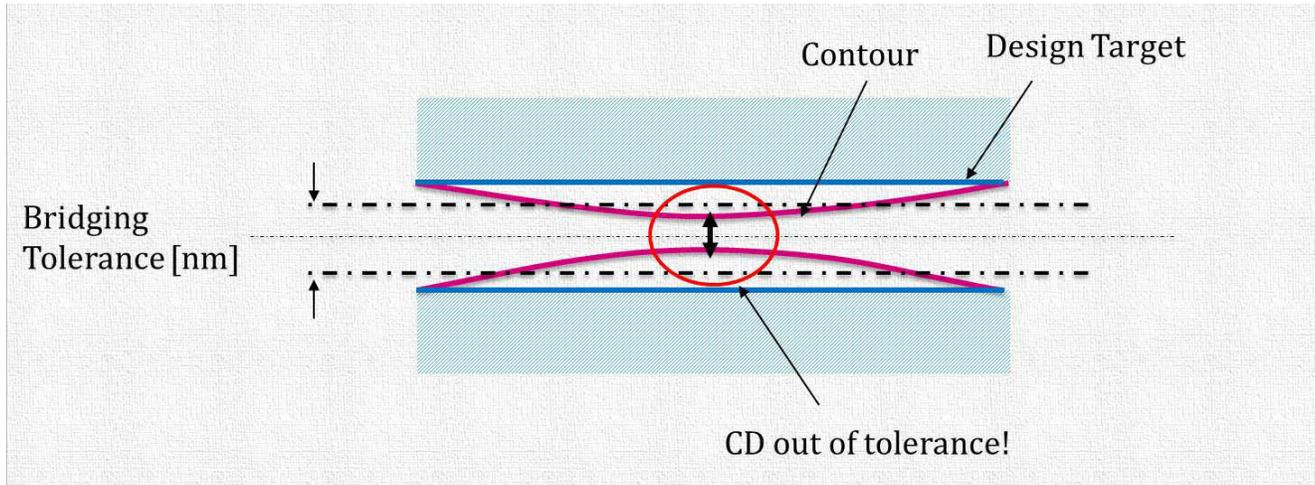


Figure 59 Illustration for bridging representing the design target, the contour and the range of the tolerance



Figure 60 SEM Post-Lithography example on a bridge hotspot (line end pattern)

Note that the only pronounced change against standard bridge detection solutions is to apply the calibrated resist 3D model. The profile aware bridge detector is set for the top of the resist. Thus the top of the resist height with no more resist, meaning a top loss, is caught then easily. So the major concern is on the bridge length: the less bridge length turn out the higher the criticality is.

4.3.3 Results and discussion among the resist profile related detectors

The basic considerations are followed by a preparatory test and then an investigation on full chip. At this step both approaches' detector are well-tuned, by means all the tolerance values are set.

For approach A, preliminary results are realized on a set of patterns including design representative pattern set: process sensitive patterns such as the pitch 90nm (line vs. space dense pattern) and critical patterns from existing studies on this layer of interest. The simplest way to represent results on these preliminary experiments is to realize a process window analysis. It is done in two steps: first a C2C type LMC simulation job that detects all the risky, significant patterns for the full chip. Then as a second stage, the highlighted patterns are further analyzed among their process windows by a second run on the selected gauges, or so called also cut lines.

Results are shown on figure 61 and 62 . Both figures represent the overlapping process window with the multiple exposure latitude for the detected patterns. Figure 61 shows results from simulations at bottom resist, at 20 % of the resist height. In this case, as it can be observed the simulated overlapping process window and the uDoF (around 70 nm) are acceptable for this layer. The overlapping PW is defined as the surface of the overlapping Bossung curves. While the determination of the uDoF is possible: it can be based on the EL-DOF curve or it can be read from the overlapping PW as the major axis of the ellipse.

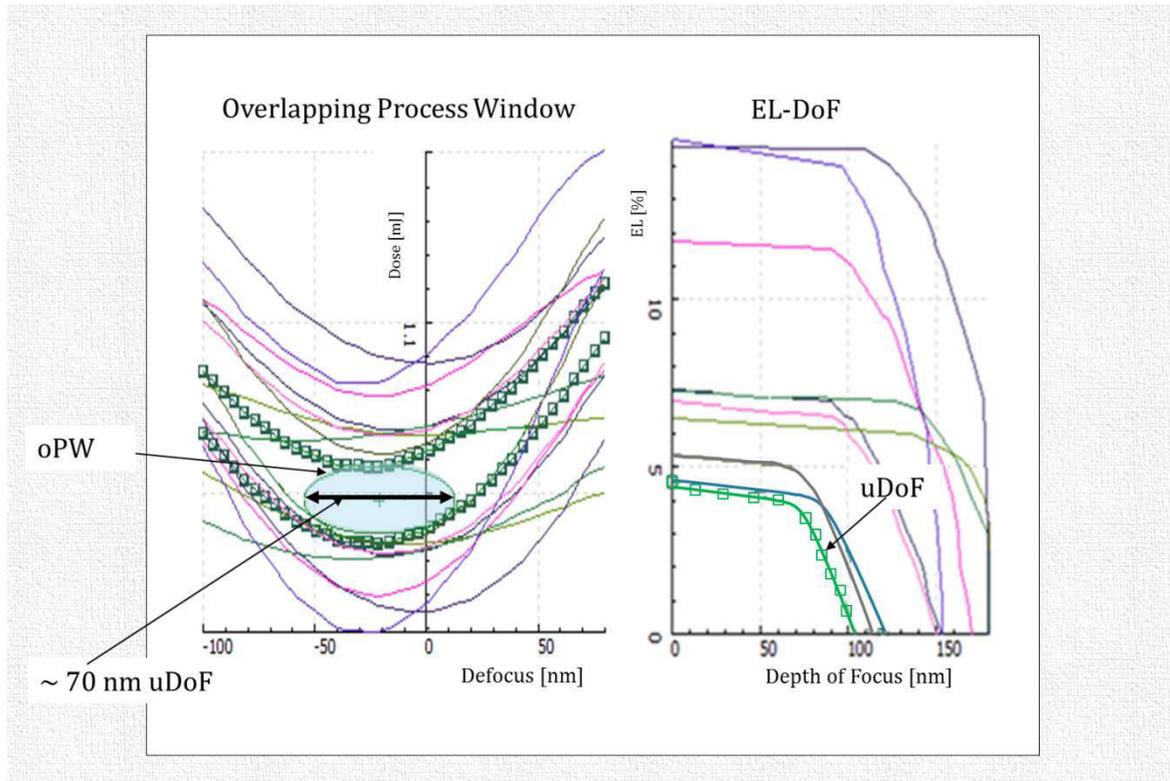


Figure 61 Process Window Analysis results at **bottom resist height at 20 %**: left: the overlapping PW right : the uDoF

Figure 62 illustrates the results for the top of the resist at 80% of the resist height. It provides no overlapping process window and zero usable depth of focus. For this particular preliminary case, the two most critical patterns are shown on the snapshots (Figure 62) indicating the top contour. For certain patterns a pronounced top loss can be seen from the highly reduced process window

conducting to failure. A snapshot is added in order to verify the contour. It proves that there is already a failure at nominal condition. As subsequent step in the next session a validation on wafer is shown. Remark returning on 4.2 section, the mask related section: both process window analyses prove the fact that the limiters are patterns owning reduced process window (coming from the resist profile issue).

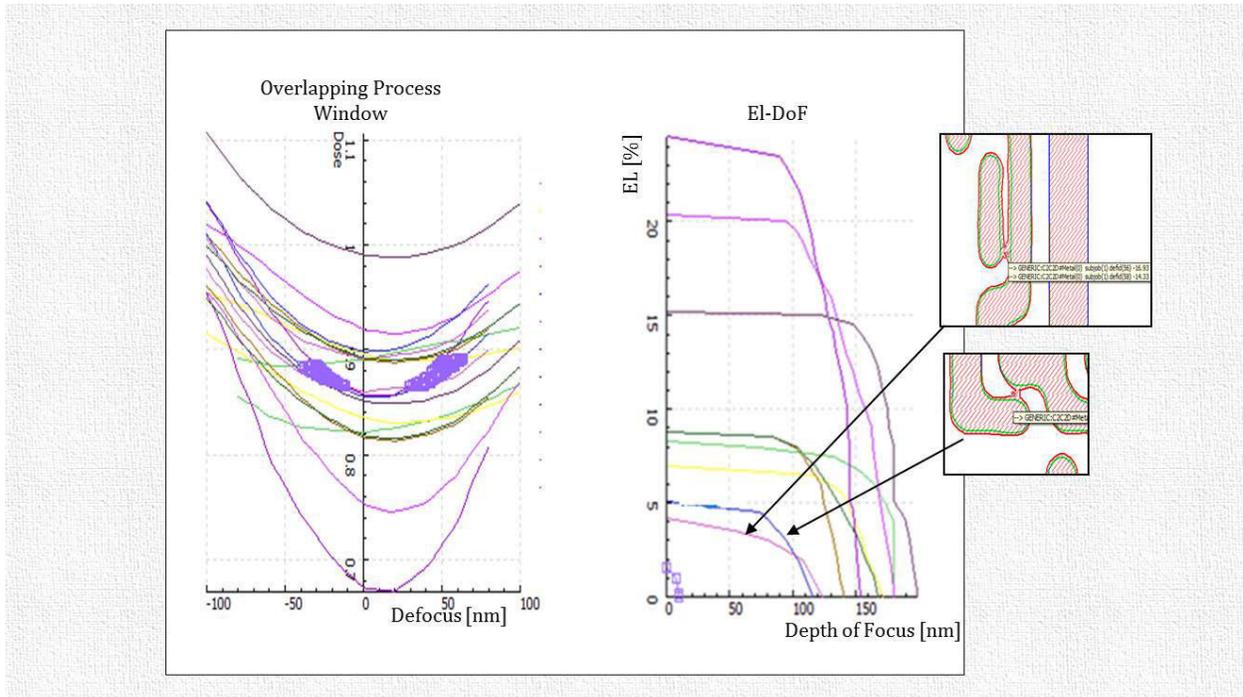


Figure 60 Process Window Analysis results at **top resist height at 80 %**: (left) the overlapping PW; (right) the uDoF indicating the two process limiters (on snapshots: red color = top contour; green = bottom contour)

To validate the approach based on simulated data, a wafer inspection is realized. Hence we are searching to inspect patterns with a presence of a pronounced top loss, a post-etch SEM observation is more justified. It is due to the fact that for a post-lithographic SEM observation it is unassured that it will provide information on this top loss notion. In opposite a post-etch image reveals on this information, less resist height will cause an etching away, so a failure. Figure 63 shows results on wafer data. For results representation only one of the process window limiter structure is taken, which is also considered to be a highly dose sensitive feature. Note that two wafers are inspected. One is an intra-wafer, means each die had the same nominal exposure condition, where the nominal dose is equal to 19.1 mJ and focus -105 nm. The other wafer is a FEM (Focus Exposure Matrix, see Chapter 1) wafer, exposed with a focus nominal ± 20 nm and nominal dose ± 0.3 mJ.

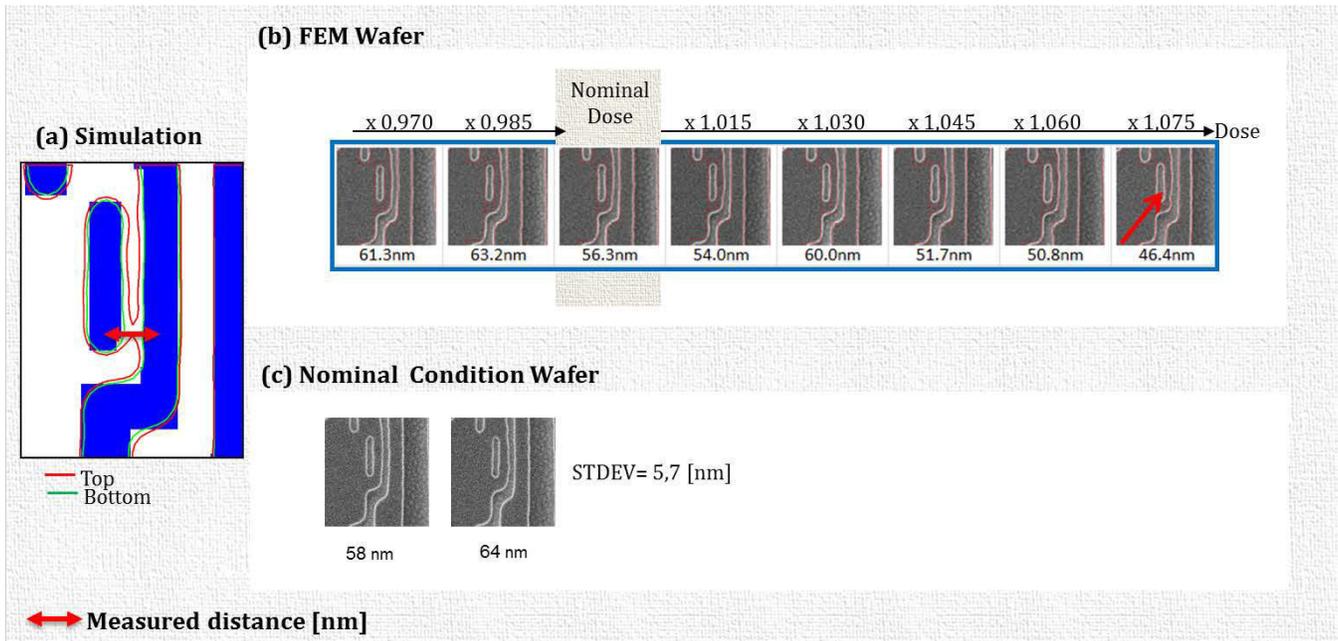


Figure 63 Validation on a detected pattern (a) the simulated contours (b) Post Etch CD SEM wafer data at nominal focus for varying dose (c) Intra wafer Post Etch CD SEM wafer data at nominal condition. Note: STDEV for all the measured data, but only 2 snapshots are presented

The simulated results show acceptable behaviors bottom (no bridging) whereas hard bridge on top. This bridging behavior is also observed on FEM at best focus for high dose post etch.. Although on experimental data the failure occurs nearer to the board. The nominal condition wafer provided data conducted to a conclusion that this pattern or more precisely the detected process window limiters need tight attention. It is from the reason that a strong CD uniformity deviation with a standard deviation of 5.7 nm is observed, while for pitch 90 nm (the control pattern for this layer) it is equal to 2.6 nm. It means that the top loss (resist profile) related failure is a kind of random phenomenon leading to random defects. It can be considered design systematic.

Our results and findings, more precisely the hotspot shown on figure 63, are also approved by defectivity analysis using Process Window Qualification (PWQ) flow. As consequence of these experimental findings the detected features are assessed indeed critical.

As further result of a first launch of the approach A, as it is, on the entire chip, a new hotspot gallery is found. It is considered to be strongly critical as they are showing hard failures line collapses. Figure 64 shows some examples from this gallery. On each snapshot, the critical spots are indicated by the red flash. These critical places tend to be hard failures that are related to top loss. Howbeit it is interesting that by a closer view on this gallery, these design configurations approach closely the hotspots that are found by the test case. At least it is the case for the middle one, which is almost the same one presented by figure 63.

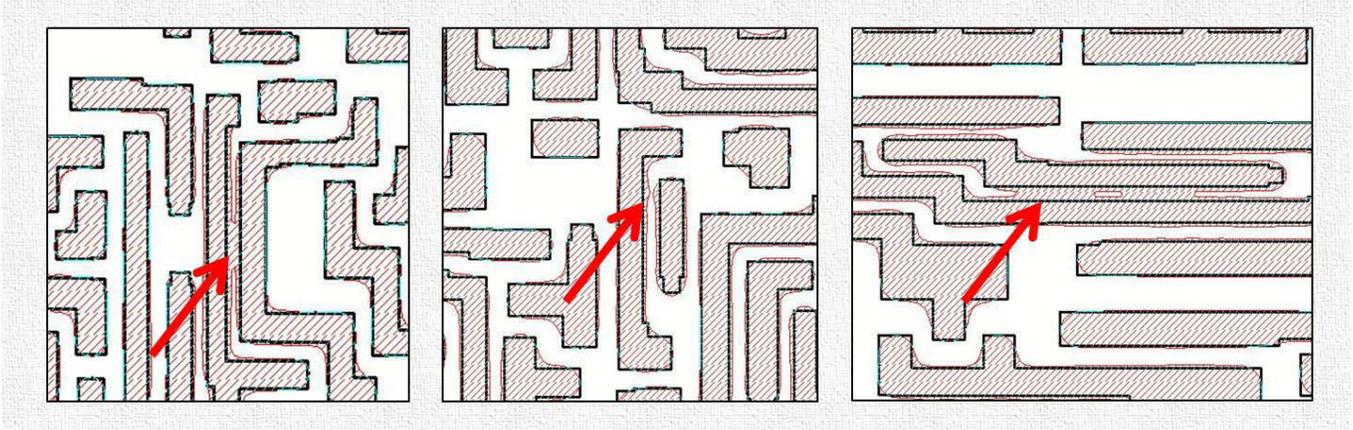


Figure 64 Simulation examples from the identified hotspot gallery that was not detected by standard bottom LMC applying 2D models

Subsequent to simulation work, wafer inspection is accomplished among the new hotspot gallery. According to previous validation concept, the same post-etch wafers are taken as these kind of defects would not be caught by post-lithography observation. Figure 65 shows results on FEM wafer data. For results representation only a few examples are taken, which are the more representative. They are also considered to be highly dose sensitive features.

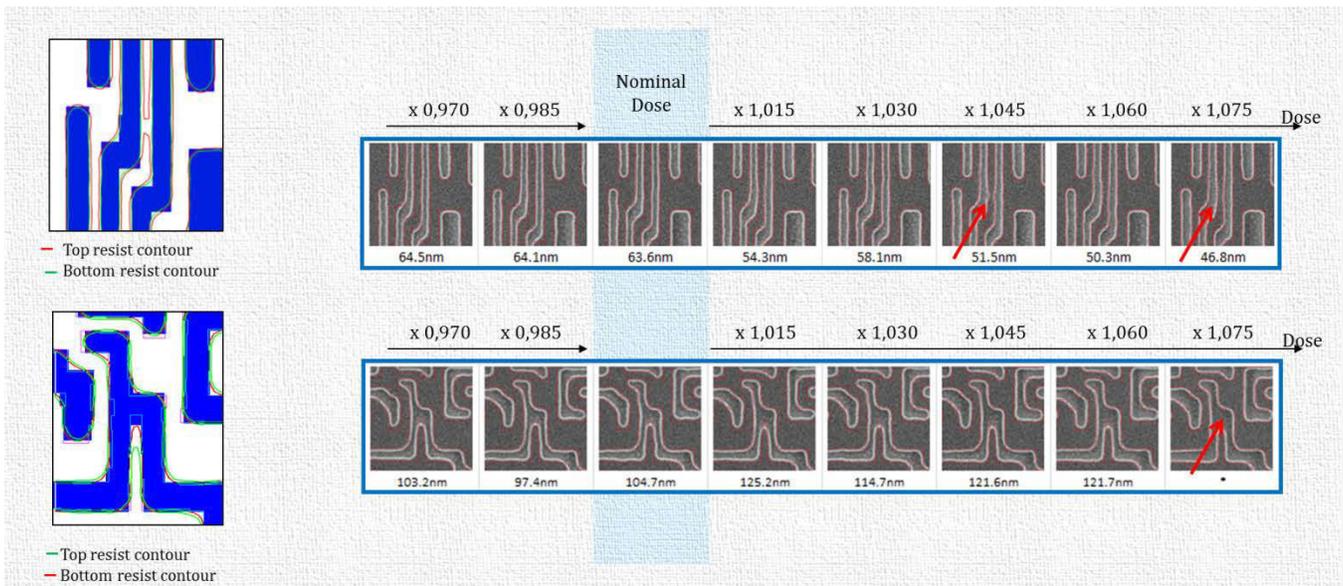


Figure 65 Validation on full chip application on detected patterns: the simulated contours vs. post-etch CD SEM wafer data at nominal focus for varying dose. Data fits well with top loss detector findings.

The experimental validation confirmed the criticality of the detected new hotspot gallery by the fine-tuned detector. The simulated contours fits well with the post-etch CD SEM data. The previously stated conclusion can be made: the resist profile related top loss failure is a kind of random phenomenon that drives to random defects. As consequence of these CD SEM

observations, the detected features are assessed indeed critical. In addition, the detected defects match well with the PWQ findings.

The defect showed in figure 63c is a real critical post etch defect than can lead failures. The problematic of detection is to know if we are capable to catch it by simulation and to rank as a most critical defect.

Although, unlikely to these promising results, approach A highlights more than one hundred thousand defects as critical patterns. This amount of data is far too much that can be handled. Figure 66 represents this problematic by the defect number count from a full chip run. The detected patterns are split into three different groups along the delta CD (CDBottom vs. CDTop) range. Note that a range high means a strong criticality. It reveals in the fact that the majority of the defects can be found between 8 nm and 9 nm of delta CD (CDBottom vs. CDTop). But still the amount of data is out of scope that can be handled in this case. Important remark is that, all of the PWQ defects, such as the post etch defect detailed on figure 63c, can be found in the defects of 8 nm to 10 nm groups counting more than 140 thousands defects, so providing the pronounced challenges.

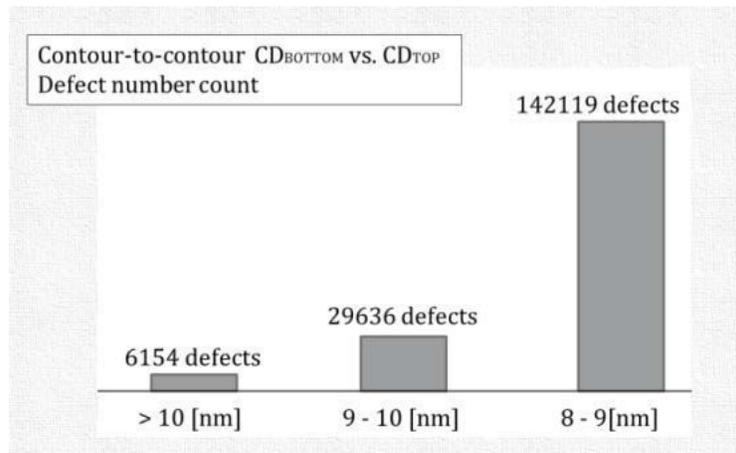


Figure 66 Contour-to-contour detector defect number count on full chip launch

Hence, the other approach B is tried to complete our methodology for full chip application. Once it is run, bridge detector defect revealed slightly more efficient compared to the contour-to-contour method. Note that a low value means a strong criticality. Inasmuch as it resulted in less highlighted defects (see figure 67), but still covering all the PWQ defects. In addition the most critical patterns are matching for both detection methods.

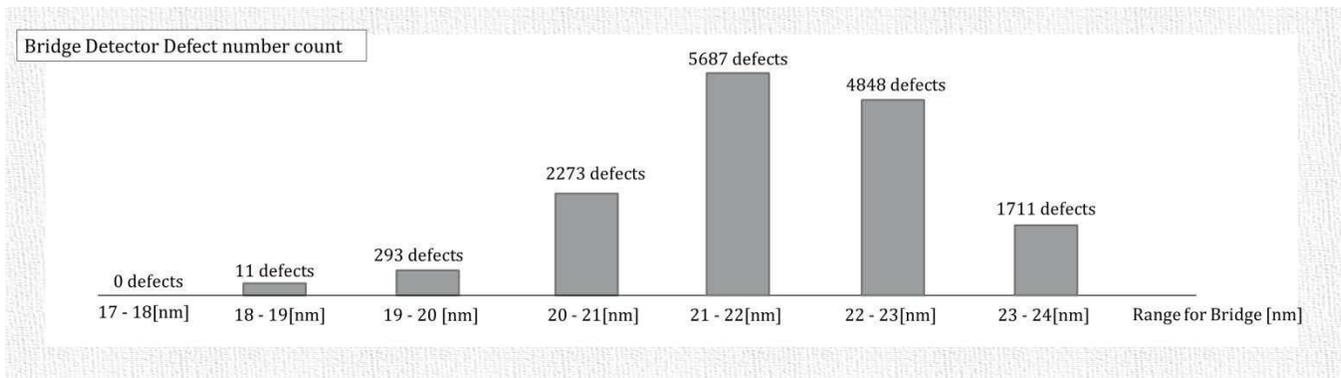


Figure 67 Bridge detector defect number count on full chip run

To conclude, the detection methodology concepts are proved for full chip application with a reliable performance and capability. Although, the approach B has more efficient performance from defects count point of view. Thus it drives the user to apply more the second detector as in production one of the key parameter is time. Note that as current state of the detector status : each approaches were developed and tried, but indeed as further steps, they requires to be fine-tuned with additional emphasize on each.

4.4 Conclusion

This chapter major objective was to enable an efficient detection methodology of the most critical, process limiter features. We approached the solution by two passes, one along the mask topography induced effects and the other along the resist profile induced effects:

- The detection method through the mask topography effects revealed in the best focus shift detection focusing on the Bossung curve analysis. Our results showed matching uDoF versus results provided by the defectivity, thus it is promising. Although, we should mention that the found critical patterns are not identic, which means that further fine tuning might be needed on this detector.
- The detection method through the resist profile effects revealed in two proposed approaches: a contour-to-contour comparison and a bridge detection supported by our calibrated resist 3D model. Both based on the use of the capability of LMC. According to presented results, both approaches enable a reliable performance and capability on detection of the most critical patterns affected by the resist profile induced effects. Results were compared to defectivity results showing good matching. Further improvement might be applied in order to enable less defects handling time, as for it is a fables of both approaches.

Furthermore, based on the fact that the observed PW limiters are mainly critical patterns affected by resist profile related effects, the resist related effects can be considered to have more impact on the reduced overlapping PW. Thus it can be considered to have more impact on the decreased

uDoF. As consequence for these nodes, the resist profile related effects might be considered as higher priority task.

Overall, the presented solutions provided promising results that enable a stable and efficient basic for the next step the mitigation of the process variability within the next chapter.

5 Chapter

Solutions for process variability mitigation

By the construction of a detection methodology, the most critical patterns can be selected as seen previously in Chapter 4. In the continuity, the sub-sequent step is to find variability mitigation solutions to reduce the set of the detected hotspots. It means the mitigation of the root cause or its compensation.

Following the reasoning flow from previous chapters, the process variability mitigation will be approached from mask and from resist side. From mask side, we are going to discuss how to compensate the mask topography related effects supported by wavefront optimization resulting in decreasing the best focus offset. And from resist side, we are going to focus on the mitigation of resist profile related effects supported by source optimization. By the source optimization, we have the intention to enhance the contrast through the resist height resulting in a straighter profile for sub-sequent etching process.

In both cases our main objective is to enhance the overlapping PW, thus increase the uDoF.

In order to complete the logical pathway of the thesis work, the process variability mitigation will be exploited still for immersion lithography systems with 193 nm wavelength and 1.35 NA. It is possible by various approaches coming from different effects as discussed in earlier chapters: from mask topography related effects, from resist profile related effects and wafer topography related effects. In the following section the mask and the resist related mitigation will be introduced and described by obtained results. For further details on wafer topography please check the following reference (Michel, 2014).

5.1 Mitigation of the mask topography related effects at scanner level

The mask topography related effects in projection systems are studied and known by almost all the engineers implied in lithography processes. It is driven by the reduced usable DoF for the advanced nodes. As suggested in previous chapters, these effects have gained significant attention with shrunk technology nodes as imaging challenges arise relative to mask thickness. It is due to the fact, that pattern dimensions become comparable to mask thickness. These mask 3D effects lead then to the imaging effects because of the non-negligible thickness of the mask. These mask 3D effects would depend, logically, on the CD, the thickness of the absorber, the mask type (material) and the manufacturing profile of the chrome for the different patterns. The mask absorber thickness would be responsible for the pattern dependent focus shift and higher order spherical aberration (Staals, 2011) (Finders & Hollink, 2011). Depending on the mask (or pattern) layout and properties, these mask 3D effects can be compensated by wavefront optimization.

5.1.1 The best focus shift mitigation by wavefront optimization

As seen before in Chapter 4, the topological effects are responsible for the difference in phase of the diffracted orders due to asymmetry between them. This fact implies the best focus deviation for the variety of structures. Further investigation of the delta phase assignment on line-versus-pitch structures has already shown that the phase difference of the diffracted orders has a corresponding phenomenon to the one induced by wavefront⁴² aberrations of the projection optics (Wong, 1992). It means that the mask topography related effects result in a wavefront distortion acting as lens aberration. The most commonly used term in conventional lithography is wavefront aberration referring to the wavefront of the pupil plane. It is determined by the Optical Path length Difference (OPD) with polar coordinates $W(\rho, \theta)$. The OPD refers to the distance of the wavefront at pupil plane from the reference sphere, which is the “ideal wavefront for a perfect lens”, see figure 68. In this figure the left part is an ideal wavefront with spherical wavefronts from object point and on the right part after an aberrated lens wavefronts are deviated from a perfect sphere but always perpendicular to all light ray direction.

⁴² Wavefront are defined as the planes of the incident lights that have the same phase and that are perpendicular to the beam direction.

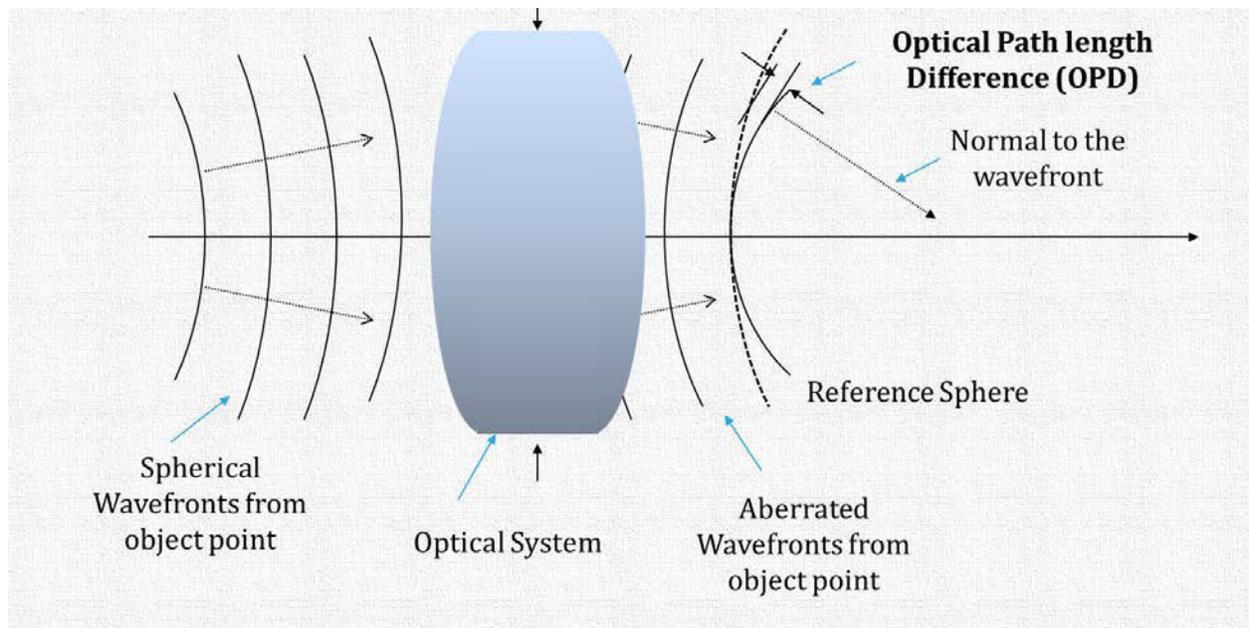


Figure 68 schematic representation of the Optical Path length Difference (OPD) between a reference spherical wavefront and an aberrated wavefront

Furthermore, the lens total wavefront aberration can be described by a series of Zernike polynomials, since the common representation of the OPD applies Zernike polynomials (Staals, 2011) (Alleaume, 2014).

A series of Zernike polynomials combines a radial (ρ) and an angular (θ) function with respect to the coordinates of the pupil and can be expressed as follow, see Equation 40:

$$W(\rho, \theta) = \sum Z_n * f(\rho, \theta)_n$$

Equation 40

Where $f(\rho, \theta)_n$ are referring to the Zernike polynomials and $n=1,2,3 \dots$ that are independent from each other. They are orthogonal units. Note that usually a finite number of independent Zernike polynomials is used for lens characterization. Any other aberration captured furthermore remains residual. So Equation 40 is as follow with a slight modification:

$$W(\rho, \theta) = \sum Z_n * f(\rho, \theta)_n + Residuals$$

Equation 41

The objective here is then to predict the actual (aberrated) wavefront, which has to be optimized in order to compensate the mask 3D effects, namely the pronounced best focus offset. This wavefront optimization is possible at scanner level, supported by computational lithography. Under the term of delta best focus shift mitigation, we mean the compensation of the mask

induced aberrations by projection wavefront optimization, based on the capability of FlexWave (courtesy of ASML, see later for description within this section) and based on its supporting software, the Tachyon simulations software (from Brion-ASML).

In order to accomplish the wavefront optimization, we are going to follow a conventional flow proposed by Brion-ASML. The optimization is realized in a reactive way. It means that as a starting point, a set of patterns need to be chosen with large topography induced effects, with pronounced best focus shift. Therefore, the previously set best focus shift detector (cf. Chapter 4) is now an excellent support to finalize the choice of patterns with the largest best focus offset.

Then, the diffraction of these structures is calculated and a phase offset is applied through projection wavefront control. It is possible by the generation and the application of a wavefront correction map by using Tachyon simulations software. This wavefront correction map denotes the required change on the reference, aberrated wavefront in order to obtain a requested one that would provide an enhanced imaging versus the actual status. The wavefront correction map refers to the Zernike polynomials: $z_1...z_n$. Inside the scanner the wavefront optimization is feasible by applying a projection lens modulator, FlexWave, in which the generated wavefront correction map can be inserted as an option.

5.1.2 Experimental setup for wavefront optimization

The experimental setup is almost identical as laid in Chapter 2: the simulation experiments are carried out on a 28 nm Metal layer using an ASML NXT: 1950i scanner (4X reduction immersion lithography exposure system with a wavelength $\lambda=193$ nm and numerical aperture NA=1.35). The used photomask is a binary mask with an optimized thin absorber called on this study thin blank.

Experimental setup at scanner level: The FlexWave lens manipulator for wavefront optimization

At scanner level, the FlexWave lens manipulator capabilities are explored for the wavefront optimization. Figure 69 shows the placement of this optional element in the pupil plane⁴³ as an additional (extra) part of the projection optics.

⁴³ The NA plane is referred as the pupil plane in practice in optical lithography

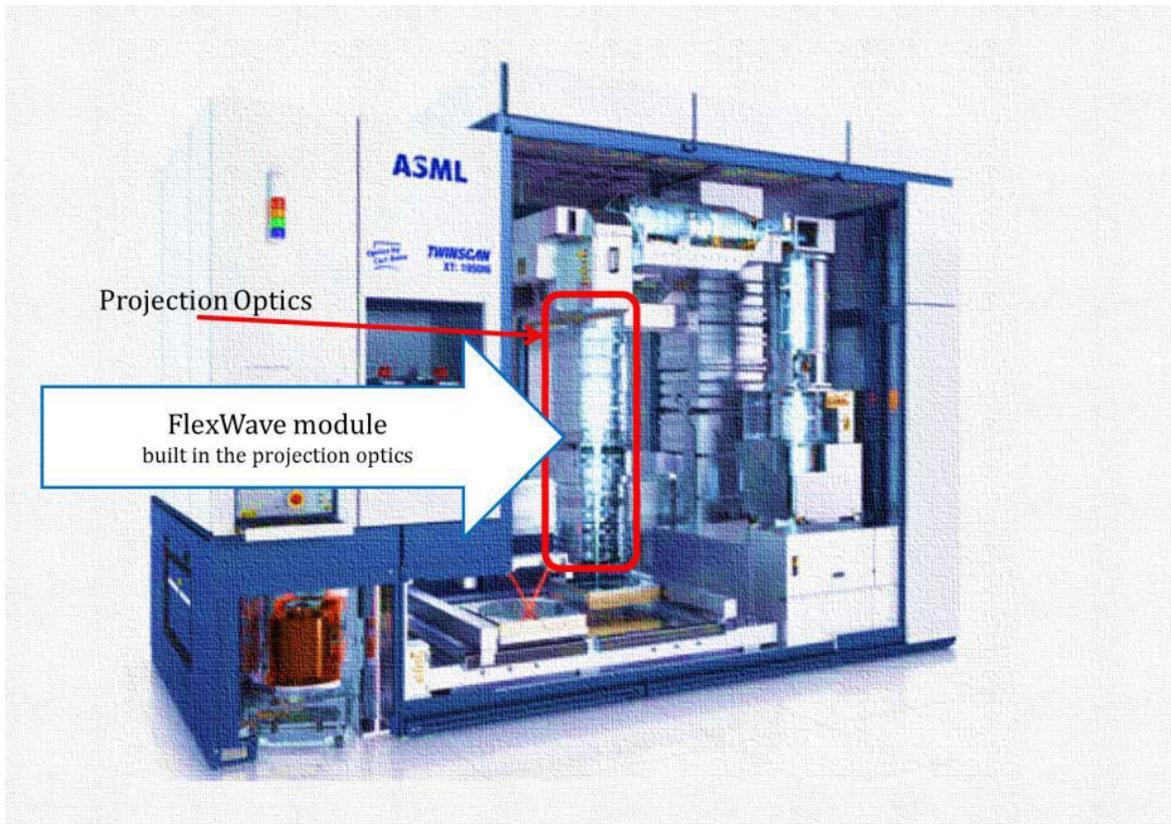


Figure 69 The FlexWave module in the scanner within the projection optics – representation of its physical position inside the tool

This manipulator's main goal is the optimal use of the projection optics. In nutshell about the operational principles: this advanced actuator enables the adjustment of the wavefront to compensate for process induced effects, like M3D effects, which can be responsible for focus and pattern shifts. This module is composed of two optical elements placed near the projection lens pupil plane, figure 70. They consist of individually heated segments translating into a change in optical path length as a function of the position. It is due to the fact that, when the individual segments are heated, the refractive index of the quartz is changing. By this change the optical path changes too. It allows the adjustment of higher order aberration terms in a complex way and tuning the phase of the wavefront. This phase wavefront change in the projection pupil can thus mitigate M3D effects. Note that besides the capability of mask topography effects reduction, this module can also minimize the overlay error performance of the scanner and better control lens heating that can occur over time (Brion, Brion, 2015) (Staals, 2011). Note that the exploitation of the lens heating and overlay related options are not covered by this thesis. The needed and optimization of input wavefront correction maps are generated in our study from simulation tools.

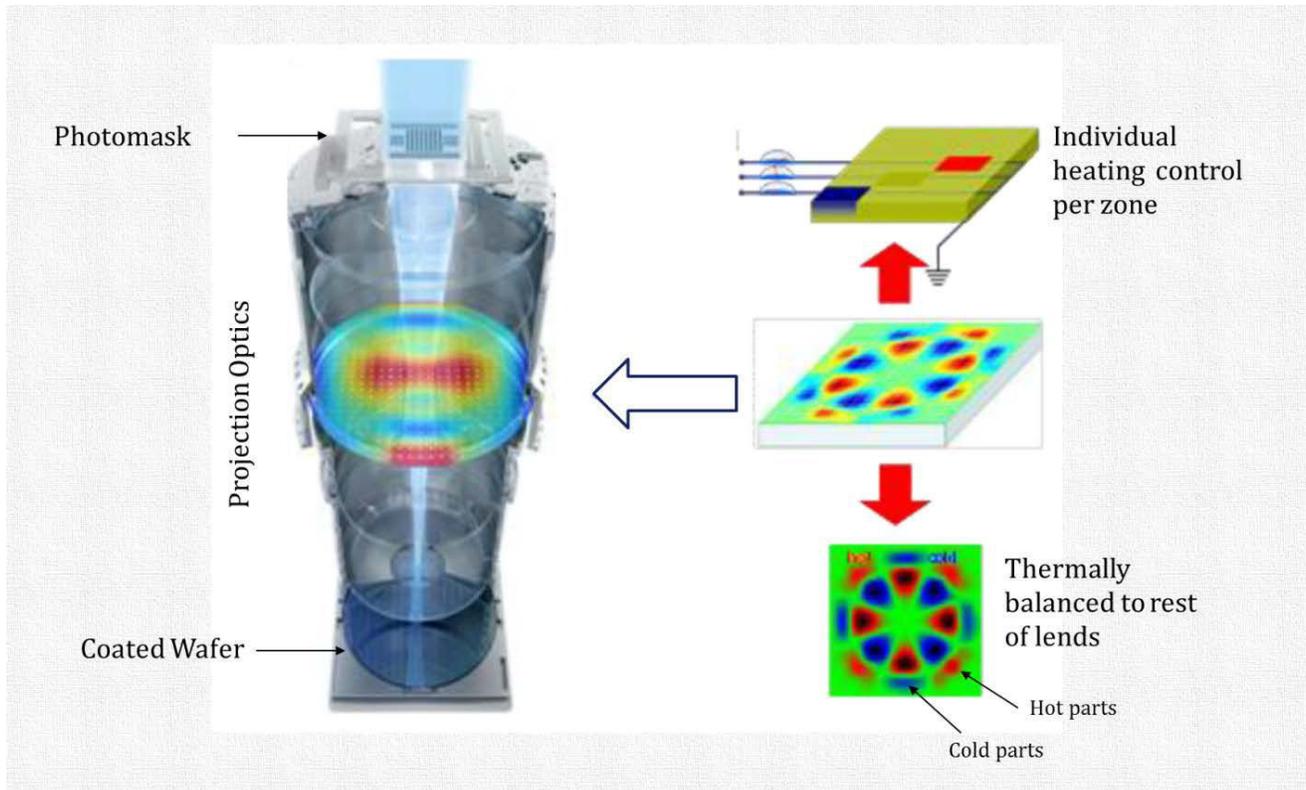


Figure 70 the principle of the FlexWave optical element manipulator available in ASML NXT: 1950i scanner; placed near the projection lens pupil plane; depicted here by a plane-parallel optical plate consisting in individual heating segments (Staals, 2011) (Szucs, 2013)

The FlexWave user interface and the supporting software are based on 64 Zernike polynomials. As described before, the Tachyon SMO-FW software enables to optimize the actual hardware capability and predict the achieved wavefront. As the FlexWave module is a pupil lens element, it would correct the Zernike offsets.

Experimental setup at computational lithography level for the optimization of the wavefront

The Tachyon computational lithography platform is used for optimization in this study. Important to note that even if this software offers a SMO (Source Mask Optimization) option, we explored only the FW (FlexWave) part of it.

The simulation experiments are split into two parts regarding the illumination modes (both XY polarized): Annular (σ outer = 0.9; σ inner = 0.7) and C-Quad (σ outer = 0.8; σ inner = 0.6). Table 3 and table 4 show the structures used for wavefront correction map generation. The set includes 1D and 2D features which are named calibration features. These patterns from 28 nm Metal layer (minimum pitch is 90nm) are selected as the most focus sensitive features from a preliminary focus shift study (cf. Chapter 4).

1D Structures	
Test Structure Name	Structure Notation
Dense	1
3 Lines_ Pitch A	2
3 Lines_ Pitch B	3
2 Lines	4
4 Lines	5

Table 3 Information on selected 1D calibration structures for the simulations

2D structures	
Test Structure Name	Structure Notation
Line End	6
2D _ A	7
2D _ B	8

Table 4 Information on selected 2D calibration structures for the simulations

Note that we have chosen 1D and 2D features from two purposes. One is to create the wavefront correction map on selected patterns separately on 1D and on 2D. And the other purpose is to setup a subsequent post optimization monitoring on these patterns. As logical consideration, we are going to inspect both 1D and 2D configurations, to check whether simpler⁴⁴ 1D pattern would be efficient enough for the wavefront map optimization and the ensuing monitoring.

5.1.3 Results and discussion

As first component of the wavefront optimization, simulations are accomplished. They are conducted on the determination and mitigation of best focus shift (coming from mask 3D effects) providing an enhanced uDoF. For validation purpose, the feasibility of best focus shift decrease by wavefront optimization correction maps are introduced into a computational lithographic rigorous simulator, Panoramic Technology®. This simulator resolves rigorously Maxwell equations for 3D patterns and is used as a reference optical result. Finally the obtained results on best focus shift and uDoF are compared to wafers exposed using FlexWave. For the comparison with experimental data, wafers are measured by scanning electron microscopy (SEM). Both simulation and measured data analysis are consisted mainly of process window (PW) and Bossung curves analysis. Note that the best focus is determined from the Bossung curves. In case

⁴⁴ Simpler in a way : more simple from inspection point of view – metrology consideration

of simulation, it is based on the output of the simulation software. And in case of SEM measurements, it is obtained by the support of in-house measurement data analysis software.

5.1.3.1 Simulations

Data reported here are including results from optimizations on 1D and 2D calibration features separately with both illumination modes. Four separate wavefront optimizations are done among the nature of the patterns (1D or 2D) and among the illumination modes for which four different wavefront correction maps are generated. Note that description, discussion and choice of illumination will be done on part 5.2

- 1st simulation : C-Quad source on 1D calibration patterns ;
- 2nd simulation : C-Quad source on 2D calibration patterns ;
- 3rd simulation : Annular source on 1D calibration patterns ;
- 4th simulation: Annular source on 2D calibration patterns.

The summarized results from the experiments using Tachyon SMO-FW simulations are shown in table 5 for the delta best focus and table 6 for the uDoF.

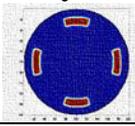
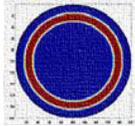
Illumination mode	Calibration structure type for correction map generation	Δ BF [nm] without wavefront optimization	Δ BF [nm] with wavefront optimization
C-Quad 	1D	71	26
	2D	12	1
Annular 	1D	20.9	3.3
	2D	13.1	2.6

Table 5 Best focus dispersions, from Tachyon SMO-FW simulation, before and after wavefront tuning on the four calibrations sets individually

Illumination mode	Calibration structure type for correction map generation	uDOF [nm] without wavefront optimization	uDOF [nm] with wavefront optimization	Percentage [%] of uDoF augmentation
C Quad	1D	90	130	44.5
	2D	122	134	9.9
Annular	1D	120	153	27.5
	2D	60	72	20

Table 6 the uDoF from Tachyon SMO-FW simulation on the four calibrations sets individually

By the simulation results, we obtained promising results for wavefront optimization. It appears, from these results on simulation data that the application of a wavefront correction decreases the

dispersion of the best focus offset on the observed four case studies. This mitigated, so re-centered, best focus shift revealed in further positive improvement: the gain in uDoF, which is a major goal with this optimization exercise. As indicated on Table 6, this gain includes an improvement up to 44.5 % of uDoF augmentation. The results in percentage augmentation showed in table 6 is more important for C-Quad illumination than annular one. It can explain by the fact that the using mask is optimized for annular source, then the optimization step has more impact on C-Quad configuration since the starting point is not optimized.

These results drive further study in to the use of the FlexWave module to explore its effects, which are validated mainly by SEM measurements.

5.1.3.2 Validation by rigorous simulation

The purpose of this validation is the feasibility of focus shift decrease by wavefront tuning using the capabilities of the Panoramic Technology® simulation package. This simulation package corresponds to a computational lithographic platform that resolved Maxwell equation. The FlexWave correction maps or more specifically the Zernike coefficients, the output data from Tachyon are used. These FlexWave correction maps consist in information of the wavefront specified with 64 Zernike coefficients which are introduced as input. The impact of the utilization of the wavefront map is estimated assuming the same structures and illumination conditions as applied in the Tachyon SMO-FW M3D simulations. The result shows a same trend as in the previous section for the Tachyon SMO-FW simulation. They are reported below in table 7.

Illumination mode	Calibration structure type for correction map generation	Δ BF [nm] without wavefront optimization	Δ BF [nm] with wavefront optimization
C-Quad	1D	55	35
	2D	NA*	NA*
Annular	1D	13.71	5.20
	2D	20	5

Table 7 Delta best focus dispersion from Panoramic® simulation on the four set of calibration structures; *NA means that it was not possible to investigate due to no usable Bossung curves

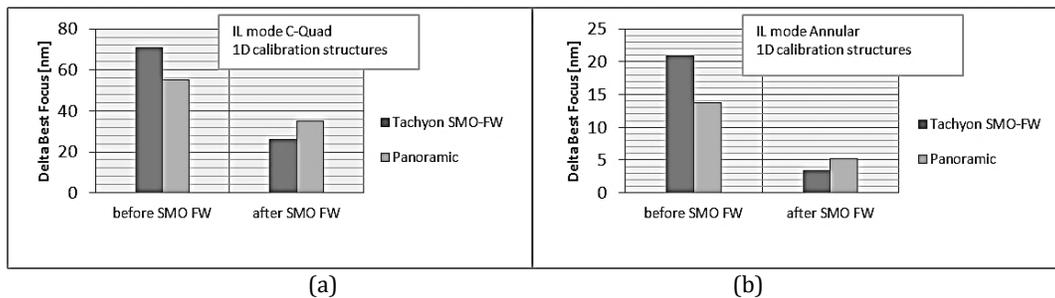


Figure 71 Delta Best Focus comparison (a) for C-Quad illumination mode (b) for annular illumination mode; illustrating the difference before and after wavefront optimization by Tachyon SMO-FW M3D and Panoramic®; simulation results based on 1D calibration structures

Figure 71 (a) with setup for C-Quad illumination mode and (b) with setup for Annular illumination mode illustrates clearly the comparable tendency between the Tachyon SMO-FW and Panoramic® software simulations by indicating the status without and with wavefront optimization for the delta best focus.

5.1.3.3 Validation by SEM measurement analysis

Wafers were exposed using FlexWave on an ASML NXT: 1950i scanner through dose and focus (FEM) to confirm the simulation results and to check the impact on silicon. Here we used the generated wavefront correction maps from the Tachyon SMO-FW M3D simulations as input data into the scanner. Two additional wafers were exposed without any correction maps serving as references. Measurements are realized by scanning electron microscopy (SEM).

In this evaluation on processed wafers, the SEM measurement data are split into:

- Evaluation of the simulation on calibration structures;
- Evaluation on non-calibration structures;
- Additional evaluation on hotspots from production.

Evaluation of the simulation on calibration structures

As a starting point, to combine simulated and measured results, the calibration structures are analyzed by full-map SEM measurements. The table below shows the study on calibration features with the difference in best focus dispersion before and after optimization in table 8.

Illumination mode	Calibration structure type for correction map generation	Δ BF [nm] without wavefront optimization	Δ BF [nm] with wavefront optimization
C-Quad	1D	28.9	10.1
	2D	16.6	3.5
Annular	1D	27.7	6.7
	2D	14.3	11.1

Table 8 Best focus dispersion from SEM analysis results on proceeded wafers on each calibration features sets using the four simulated wavefront corrections maps

It shows an impact in best focus shift for the calibration features: a noticeable decrease can be observed. The best focus of each calibration structure demonstrates improvements on alignment, which also have uDoF impact as follows (see table 9):

Illumination mode	Calibration structure type for correction map generation	uDOF [nm] without wavefront optimization	uDOF [nm] with wavefront optimization	Percentage [%] of uDoF augmentation
C Quad	1D	60	110	83.3
	2D	60	90	50
Annular	1D	80	120	50
	2D	80	110	37.5

Table 9 uDoF from SEM analysis on the four calibrations sets individually indicating the augmentation in percentage for each of the cases

These results show tight correlation with Tachyon SMO-FW simulation results. Although should be mentioned in term of uDoF simulation that results are more optimistic than reality is on wafer certainly due to additional process variability not simulated on scanner resist or mask. It can come from the fact that the best focus calculated based on the Bossung curves would not provide an exact value if a Bossung tilt is present. Idem for plate Bossung curves.

Evaluation on non-calibration structures

The objective of these measurements on non-calibration patterns is the inspection of the impact on additional set of features (beyond the calibration structures) that were not used to optimize the wavefront. The inspected impact refers to a risk of degradation. Therefore the major goal is to explore the validity of the optimization. Beside the validation of the optimization, the objective of this analysis is to check whether the calibration could be done with 1D structures only.

This analysis includes a set of 21 features. A similar trend can be observed as by simulation: the best focus of each individual structure is re-centered reducing the best focus shift deviation (see Table 10) in other word the range of best focus for all features is reduced with wavefront optimization

Illumination mode	Calibration structure type for correction map generation	Δ BF [nm] without wavefront optimization	Δ BF [nm] with wavefront optimization
C-Quad	1D	47.1	26
	2D	47.1	30.2
Annular	1D	41.6	12.7
	2D	41.6	31.5

Table 10 Best Focus dispersion results from SEM analysis results on proceeded wafers on non-calibration features using the four simulated wavefront corrections maps

Table clearly shows the impact of the application of wavefront tuning using FlexWave, which is reducing the best focus offset. Along the best focus mitigation, a pronounced augmentation can be observed under the change in usable DoF (table 11):

Illumination mode	Calibration structure type for correction map generation	uDOF [nm] without wavefront optimization	uDOF [nm] with wavefront optimization	Percentage [%] of uDoF augmentation
C Quad	1D	80	110	37.5
	2D	80	110	37.5
Annular	1D	60	120	100
	2D	60	110	83.3

Table 11 The uDoF from SEM analysis on non-calibration features indicating the augmentation in percentage for each of the cases

With the goal of investigation on calibration structures selection, the validation on additional features is completed by checking the possible penalties derived from the application of wavefront

tuning. The SEM measurements analysis shows no pronounced degradation that can be observed. Furthermore the application of the correction map, generated on 1D calibration structures, shows more significant reduction in delta best focus and uDOF augmentation. So means that for further improvements on calibration structure selection: the choice is on 1D features regarding measurement complexity aspect.

SEM measurements analysis on hotspots

Finally, in order to explore the impact of wavefront tuning by FlexWave, a set of hotspots structures is chosen from production full-chip verification using Tachyon Lithographic Manufacturing Check (LMC) simulations (cf. Chapter 2). These experiments consist of bridging and necking defects with overall about ten critical patterns for each case. Focus shift analysis is addressed first to assess whether applying the correction map using FlexWave re-centers best focus or not. In parallel the impact on usable DoF improvement is investigated. Note that the study of these critical defects is not obvious. The extraction of the results is quite complex, since the reliable determination of the best focus and uDoF is not possible automatically. A commonly applied method is visual analysis, thus results are handled this way.

These results corroborate the computational wavefront optimization in the way that the overlapping process window is improved: the position of the best focuses is re-centered (Figure 72) for checked hotspots thus uDoF (dotted line) is extended as show Table 12 and Table 13 .

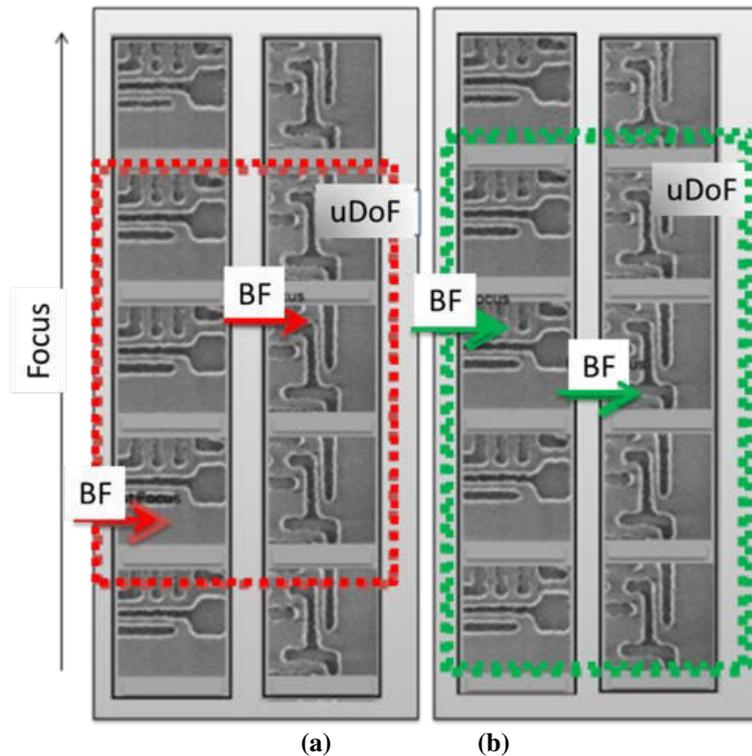


Figure 72 C-Quad illumination mode: Example of the wavefront optimization impact on metal 28 nm hotspots bridging at best dose illustrating the re-centered best focus position and evaluated uDoF (dotted line) for taken hotspots (a) without wavefront optimization (b) with wavefront optimization

Illumination mode	Calibration structure type for correction map generation	Δ BF [nm] without wavefront optimization	Δ BF [nm] with wavefront optimization
C-Quad	1D	30	10
	2D	30	10
Annular	1D	60	30
	2D	60	20

Table 12 Best Focus dispersion results from SEM analysis results on proceeded wafers on critical patterns using the four simulated wavefront corrections maps

Illumination mode	Calibration structure type for correction map generation	uDOF [nm] without wavefront optimization	uDOF [nm] with wavefront optimization	Percentage [%] d' uDOF augmentation
C Quad	1D	60	90	50
	2D	60	90	50
Annular	1D	60	80	33.3
	2D	60	90	50

Table 13 The uDoF from SEM analysis on non-calibration features indicating the augmentation in percentage for each of the cases

By these results, again, the tendency is the same – although we need to be conscious of the fact, that measuring complex 2D patterns is not that simple, in addition the determination of the uDoF is usually realized in a manual way.

Overall, according to these data, the hotspot evaluation showed close correlation between SEM measurements data and simulation data from Tachyon SMO-FW and the Panoramic Technology® simulation package.

5.1.4 Partial Conclusion on mask topography related mitigation at scanner level by wavefront optimization

For compensation purpose of mask 3D effects, wavefront correction is investigated by the support of FlexWave lens manipulator. According to the performed wavefront optimization using computational lithography and SEM measurements analysis there is indeed a pronounced effect from the mask topography. The wavefront manipulator enables an imaging enhancement by minimizing mask topographic effects.

Based on comparative analysis, on the layer of interest, of computational lithography and SEM measurements data, consistent results were found. By presented experiments it was shown that by computational lithography a pronounced increase in uDoF can be achieved by applying wavefront optimization on two illuminations tested. The same tendency is shown on experimental wafer.

Beyond the exploitation of uDoF enhancement, investigation was performed in order to assist calibration features selection. Here we found that 1D features are the best candidates to serve for wavefront correction map generation. So we can conclude that the task to mitigate, or more precisely to compensate, the mask topography related effects is accomplished. In the next section we are going to then discuss the mitigation of the resist profile related effects.

5.2 Mitigation through the resist profile related effects

As seen in previous chapters, a good control of the resist profile now becomes more and more crucial for shrunk nodes. It is due to the fact that the aspect ratio between the CD and the resist thickness is more and more pronounced and of course the complexity of 2D features. As example, figure 73 shows a critical hotspot – same as in previous chapter – whereat the top loss effect might induce a failure after etch.

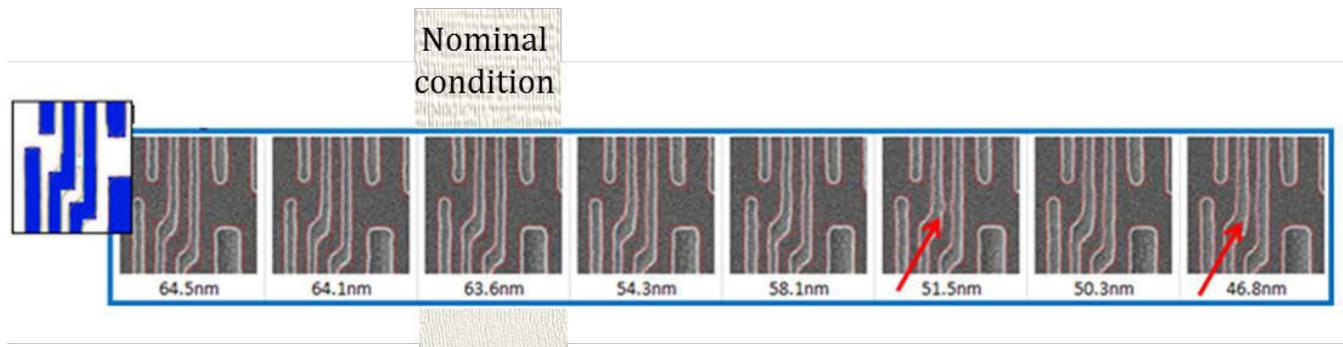


Figure 73 top loss effect's example on a critical hotspot at nominal focus through exposure dose indicating the nominal condition (whereat the dose step = 0.3 mJ/cm^2)

Indeed the named resist profile is first related to chemistry facts, so it is tuned usually through the baking and the development stages of the lithography process. But it can be also optimized by the optical intensity distribution through the thickness of the resist. Therefore such an example for this approach, a resist profile aware OPC can be listed. By means, during the correction, the final mask design is customized by top CD enabling a better top CD control, and of course maintaining the required bottom CD resist contour as well (Seongbo, 2011). Furthermore, since the mask layout and the illumination source plays important role on the intensity distribution of the pattern image, the resist profile can be further optimized by mask and/or source optimization (Alleaume, 2014) (Brion, Brion, 2015) (Wu, 2011).

Within the framework of this section on the resist profile related mitigation, a resist profile aware source optimization (SO) is presented. The previously (cf. Chapter 3) calibrated resist 3D model provides the basis for this approach. The optimization's cost function contains the image properties calculation from the bottom and from the top resist image planes as well. As consequence the final expectation is to obtain a good compromise in between the bottom CD's overlapping process window and the top CD's overlapping one. All that by assuring a better, straighter resist profile for the sub-sequent etching. Thus the objective is to reduce the "bumped" shape resist profile by the achievement of a straighter one and to provide an enhanced uDoF.

Indeed an additional item at this point, a beneficial optimization without any image quality degradation on design layout.

5.2.1 The source optimization (SO)

The source optimization (SO)⁴⁵, has become a key point for advanced nodes in lithography only since a few times ago. Indeed, the application of simple sources is adequate enough for less complicated design layers, not the case for the advanced ones. The simple sources show excellent results in term of uDoF, thus imaging for 1D feature (Alleaume, 2014). Also, they are reliable for mass production. Some simple source examples are shown on Figure 74: an annular source, a quasar source, a C-Quad source and a dipole source. They are represented by the incident light intensity I , where $I_{\max}=1$ (red color on figure). They are characterized by their polar coordinates sigma (σ): sigma inner and sigma outer; and the opening angles.

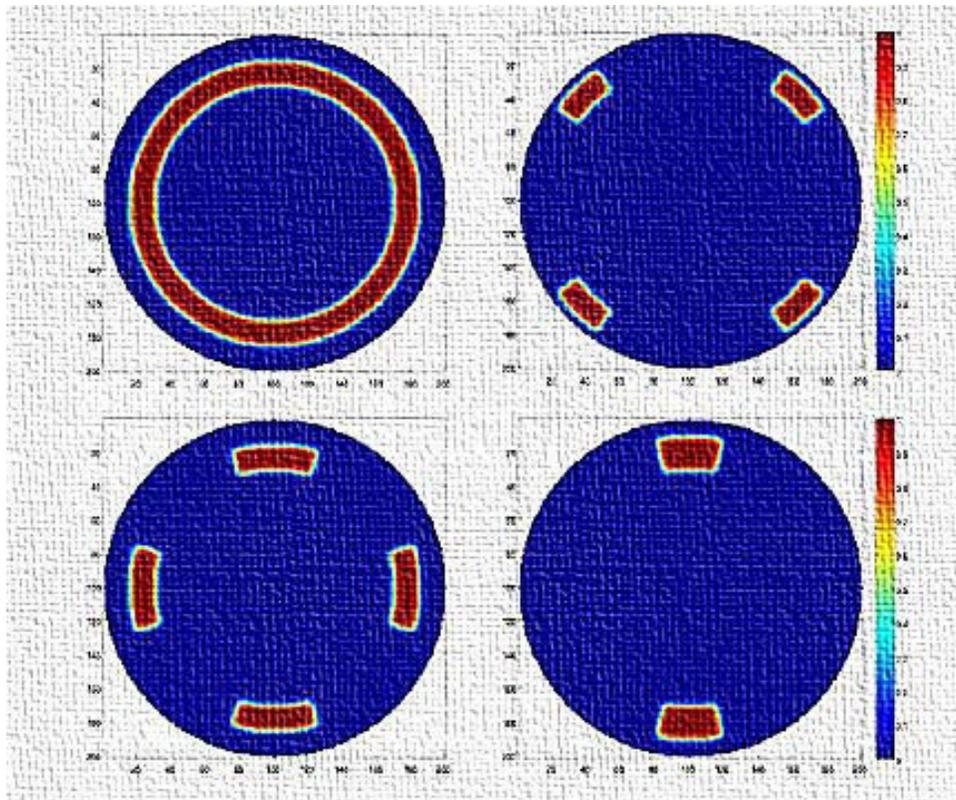


Figure 74 Examples for simple sources – the mostly applied ones; (Upper left to right) annular source, Quasar source, C-Quad source, Dipole source

By the support of the SO, we are searching to optimize these types of simple sources for a larger overlapping PW and, of course, an increased uDoF. Several types of optimized source exist, such as

⁴⁵ In general the Source Mask Optimization (SMO) term is used, if only the source optimization option is applied, we used to say SO only.

the parametric source of the Freeform⁴⁶. Each has its own strength and weakness, although the freeform source has been shown resulting in better results in term of uDoF (Alleaume, 2014) (Wu, 2011). Thus a Freeform SO is applied here supported with the calibrated resist 3D model.

Here below, on Figure 75, the applied simplified flow of the resist profile aware source optimization is introduced; whereat the input source is our annular source as set in Chapter 2 by the experimental setup description. It is possible, by feasibility reasons in fab, to constraint the output source. In our case this constraint is as follow: optimization within a frame of $\sigma_{inner}=0.7$ and $\sigma_{outer}=0.9$ based on our input source.

Then, a set of location where the aerial image is calculated (cut lines) on selected pattern clips are set. These patterns are belonging to the design representative patterns. The identified aerial image planes are the following: one at 20 % of the resist height, and one at 80 % of the resist height, corresponding respectively to the bottom and the top of the pattern. Source optimization is performed to minimize edge placement errors (EPE), It refers to the delta value (error) in between the simulated contour and the target contour in nm. This minimization is done at different Process Windows conditions (dose, focus, mask variability) and including also different Aerial image location (20% and 80% of resist height). (Brion, Brion, 2015) (Chong, 2014).

In this optimization case, all the computational work is realized by the support of The Tachyon computational lithography platform (Tachyon SO & SMO software for RET, courtesy of ASML Brion). For confidential reasons, details are not discussed.

Once the optimization is accomplished, a source validation needs to be placed. For this performance investigation, different methods are developed. If the source validation shows that the optimized source's performance is adequate and efficient enough on critical patterns other than the ones used for the optimization, then it can be tried and validated on real wafer in fab as the final goal is the application in production. Note that in our case due to time constraint this flow is completed only at simulation level.

⁴⁶ Freeform source: ~ is an illumination type for which there is unlimited freedom in intensity and position of the light in the pupil (Alleaume, 2014).

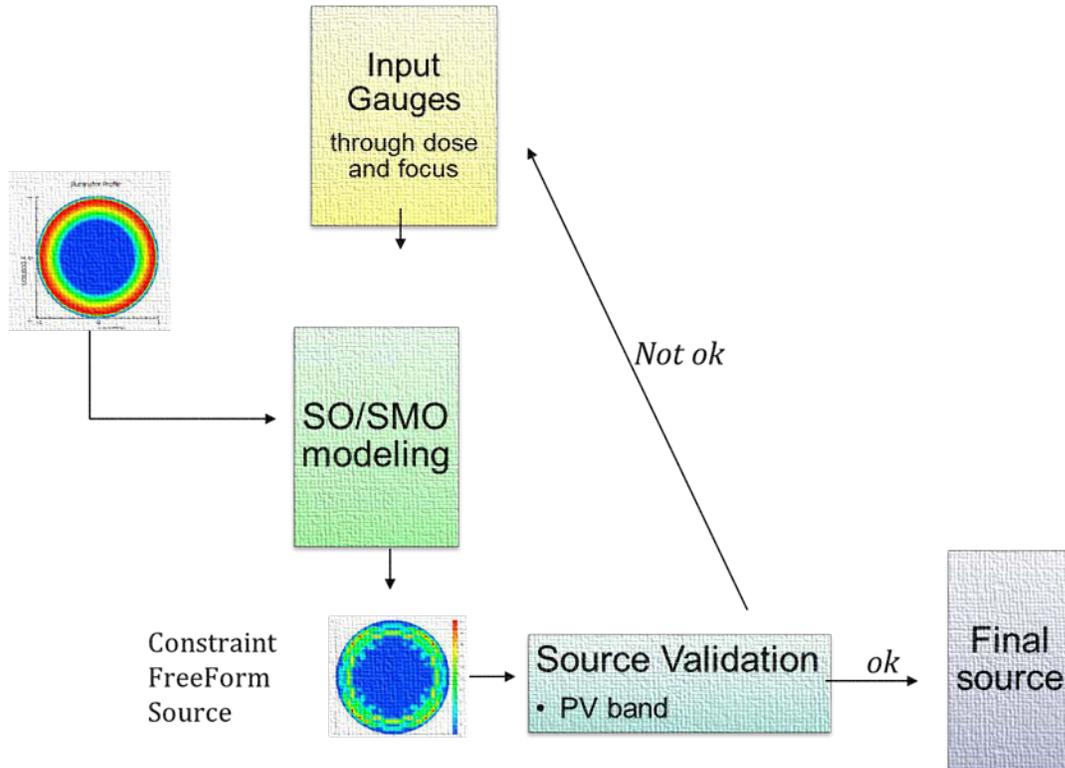


Figure 75 the flow of the source optimization (simplified one)

5.2.2 Data setup for the simulation

We investigate the use of source optimization (SO) through resist height in order to optimize also resist profile supported by Tachyon SMO (SO only). The baseline model is the calibrated resist 3D OPC model. The reference baseline source is the up-to-now applied annular source (0.7/0.9 sigma). As optimization parameter on the source, a constraint is set: the final source must be in the boundary of 0.7/0.9 sigma.

Furthermore, for the optimized source generation, a set of about thirty patterns is chosen for this case study. Indeed before the optimization, we have to carefully define, which patterns we are going to use for the complex optimization as it can change everything. Herein the choice is realized following some criteria. Each of the selected patterns is mask design representative. The belonging layout crop of each should be at least a 100 nm X 100 nm square optimization area for a repeating, periodic area. It is recommended to have overlapping PW limiter critical patterns within the set of these features that would widen the optimization performance. In figure 76, simple periodic 1D feature are selected, such as line versus space covering all the palette of appearing pitches on reticle including the critical ones – e.g. pitch 90 nm (as less pitch is considered as forbidden pitch).

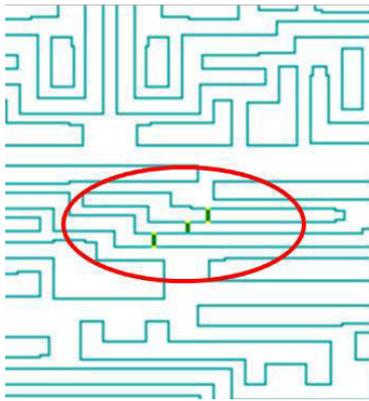
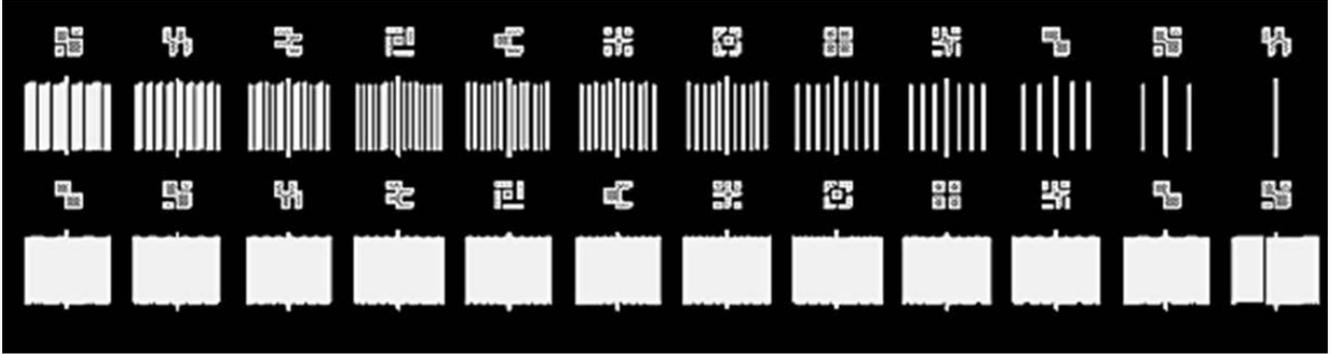


Figure 76 1D and 2D examples on selected patterns for the source optimization

Remark: at this point, an interesting question can be posed: whether the use of a lot of patterns in the optimization has an impact on the final optimal source or not? Will it always converge to the same or a similar solution? It might be an important point as by adding numerous patterns in the input gauges, on the optimization mask layout, the increase of CPU time will be very important.

5.2.3 Results and discussion

As first results on the initial source versus the optimized source a simple verification check is realized, see figure 77. The verification check consists of a check on the half CD through the resist height for a dense pattern (pitch 90 nm) and a hotspot. The CD variation through the resist height (max - min value) for the initial source is equal to 17 nm for the dense pattern and 56.7 nm for the hotspot. The application of the optimized source provides less variation: 15.7 nm for the dense patterns and 53.7 nm for the hotspot. Thus there is an improvement for both features (3nm for the hotspots). Indeed there is a pronounced and significant variation for the hotspot due to top loss. The next step in the following section is to validate this optimized source.

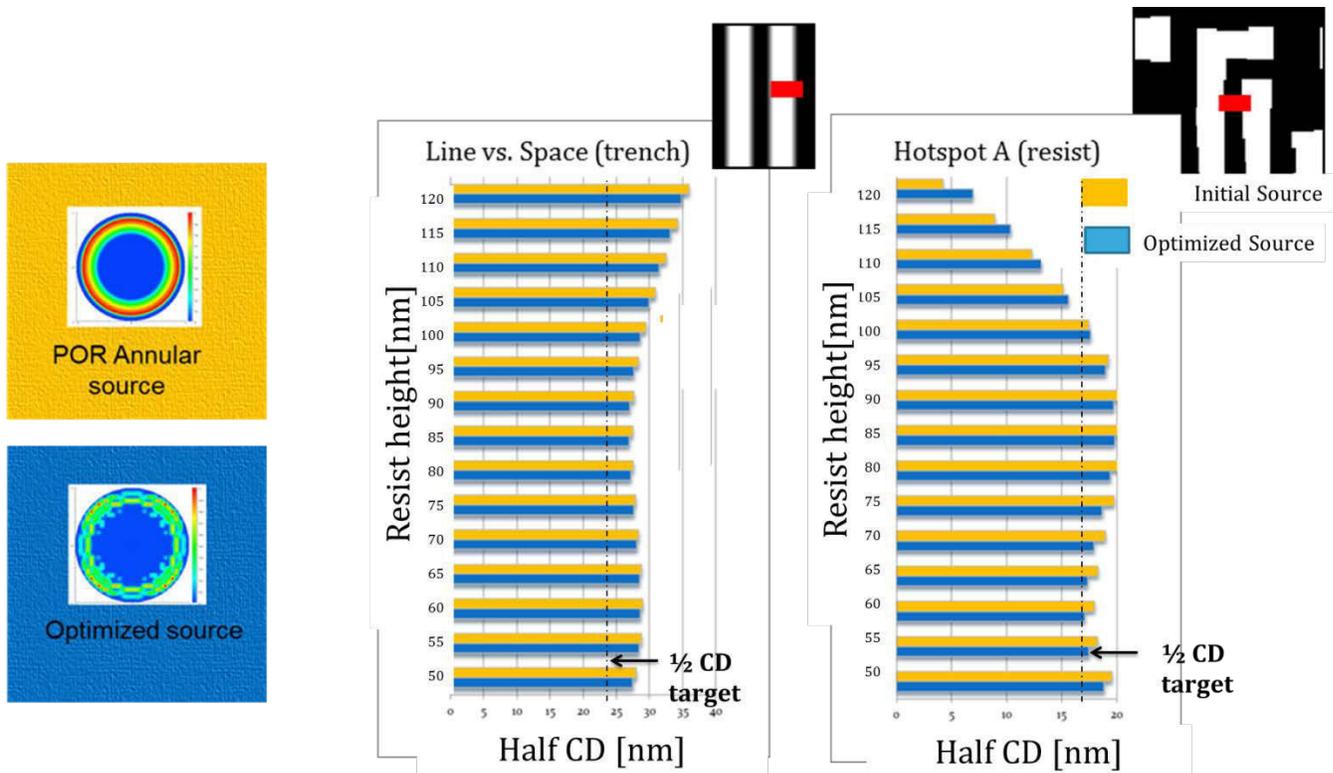


Figure 77 input source vs. optimized source results by half CD [nm] through resist height representation on two examples. Left graph: line vs. space with pitch = 90 nm showing the half CD variation through the resist height on trench. Right graph: a selected hotspot showing the half CD variation through the resist height on resist line

5.2.4 Validation of the optimized source

The sub-sequent step of the source optimization is to investigate the impact whether the application of the optimized source will result in any degradation or whether it results in Process Variability Band degradation. For that purpose three different approaches were developed along the resist profile.

5.2.4.1 Method 1: Normalized CD (means contour CD divided by the target CD) for various cut lines by the support of PV Band Analysis

The Method 1 is supported by the Process Variability (PV) band performance investigation. A PV band is the representation of simulated contours obtained by the variations in the process parameters such as dose, focus or mask bias. It is used as a metrics to quantify design sensitivity to process. Thus it is focusing on changes in the lithography process parameters that may results in non-desired impact (e.g. bridging failure).

The PV band investigation is supported by computational lithography software's, whereat the user gives all the process conditions including different doses, focuses and mask biases as input. Then from these data, the simulation results in the extreme contours. It means the maximum and the minimum displacement after exposure (= printed image). The CD difference in between will provide a PV band value showing how much a design is sensitive to process variations.

The PV band width is very useful as it can be considered adequate as a quantitative measure for judging the mask design manufacturability. Accordingly it is often used for optimization validation purpose (Zou, 2010) (Alleaume, 2014). So as we applied for our validation of the optimized source versus the initial one. Figure 78 represents our metric, which is based on the normalized CD (contour CD per post-lithography target CD) for various cut lines.

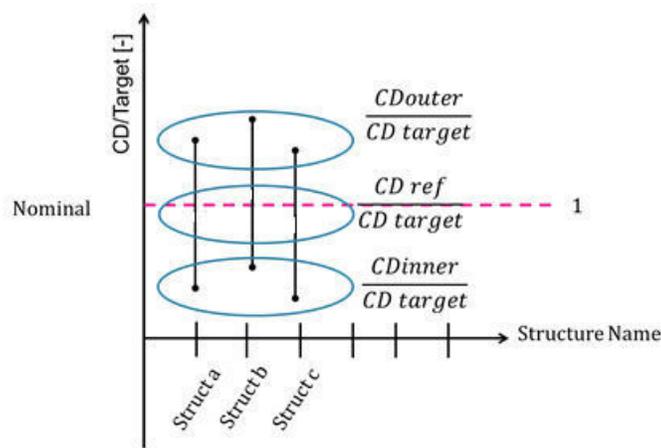


Figure 78 the used metric for the PV band analysis

We define as success criteria the following items: the less PV band width means the more performant the process is. And the closest the normalized CD for the reference condition is to the nominal 1 value, the better the quality of OPC is. Indeed in our case the first item is the most important from results point of view. Note even if it is a method with good capability to judge the process, it remains cut-line based, means no full chip capability. In addition, it stays limited for the resist heights (Top & Bottom).

Our results for Method 1 are shown on figure 79 and figure 80 for 1D pattern. Figure 82 and figure 83 are representing for 2D patterns. On both of these figures the normalized CD (means contour CD divided by the target CD) are represented for the various selected cut lines. A part of them are chosen as simple 1D layer representative features consisting line vs. space configuration patterns. The 1D ensemble consists of 25 cut lines. . The 2D ones are examples from the most critical structures from our findings from Chapter 4, see figure 81.

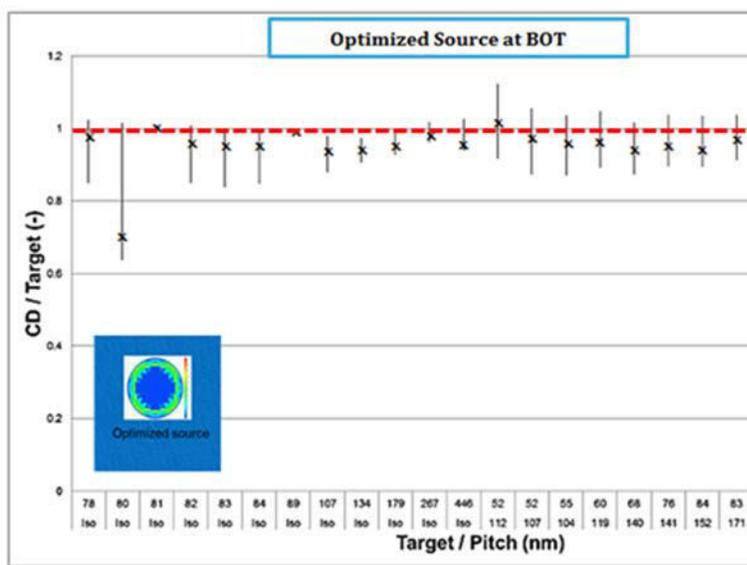
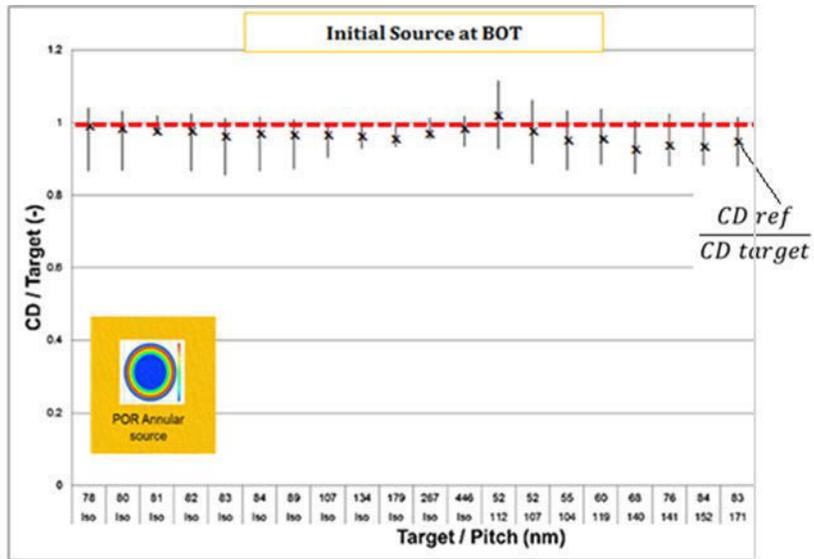


Figure 79 PV band comparison results: Contour CD / the target CD for different 1D patterns : upper part representing the results for POR Annular source, while the other for the optimized source for Bottom resist height

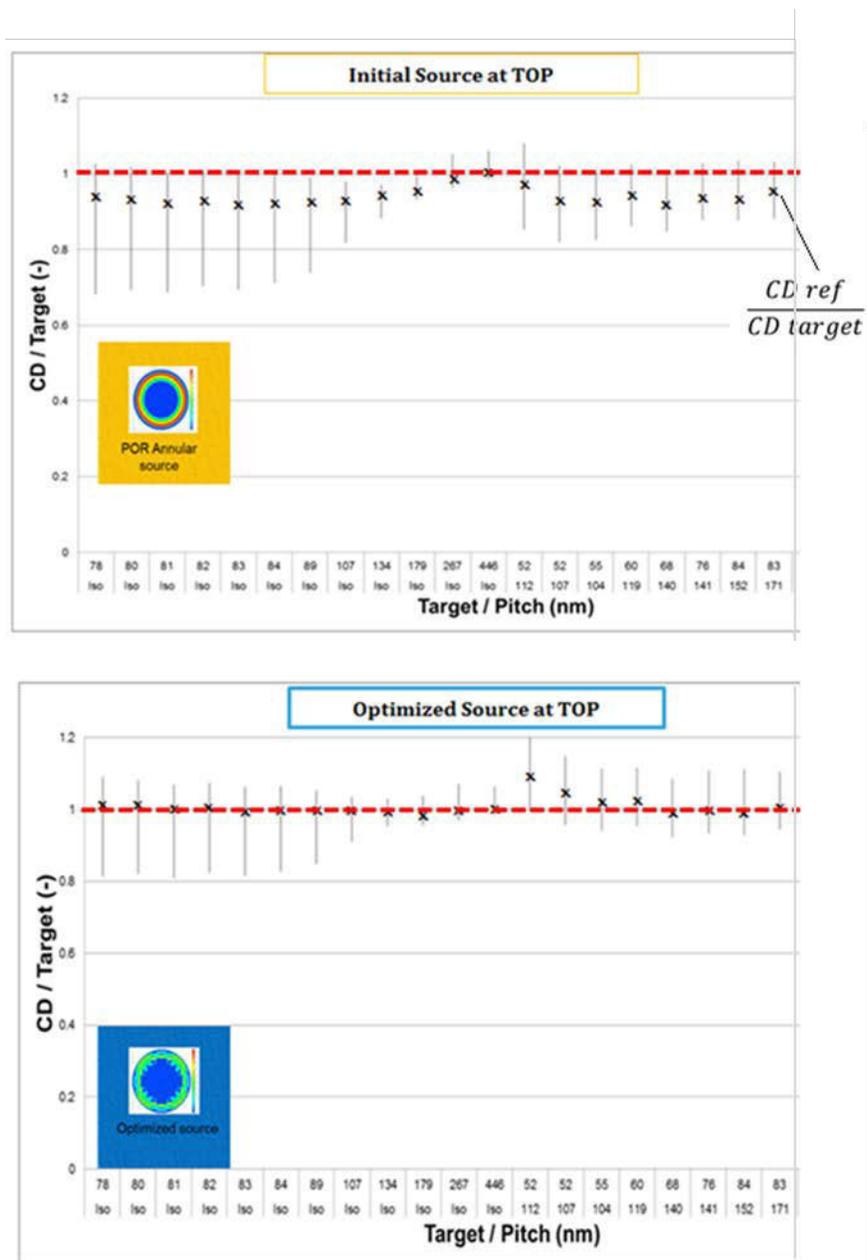


Figure 80 PV band comparison results: Contour CD / the target CD for different 1D patterns: upper part representing the results for POR Annular source, while the other for the optimized source for Top resist height

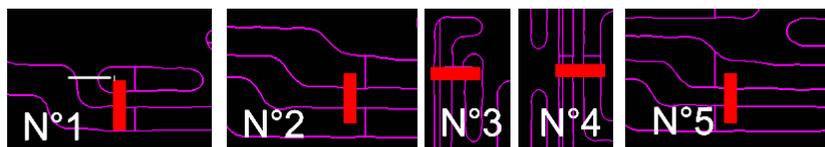


Figure 81 the used Cut Lines for 2D patterns

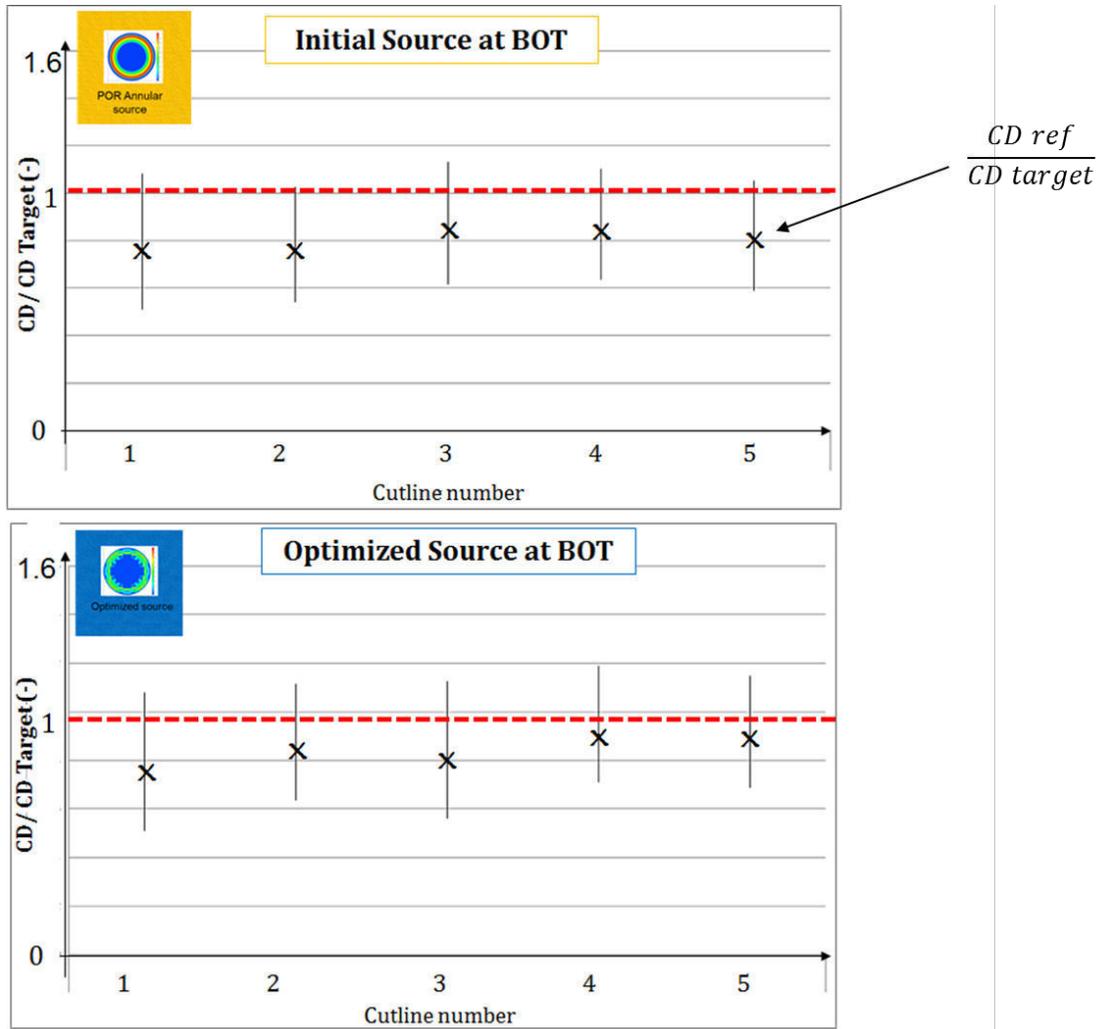


Figure 82 PV band comparison results: Contour CD / the target CD for different 2D patterns: upper part representing the results for POR Annular source, while the other for the optimized source for Bottom resist height

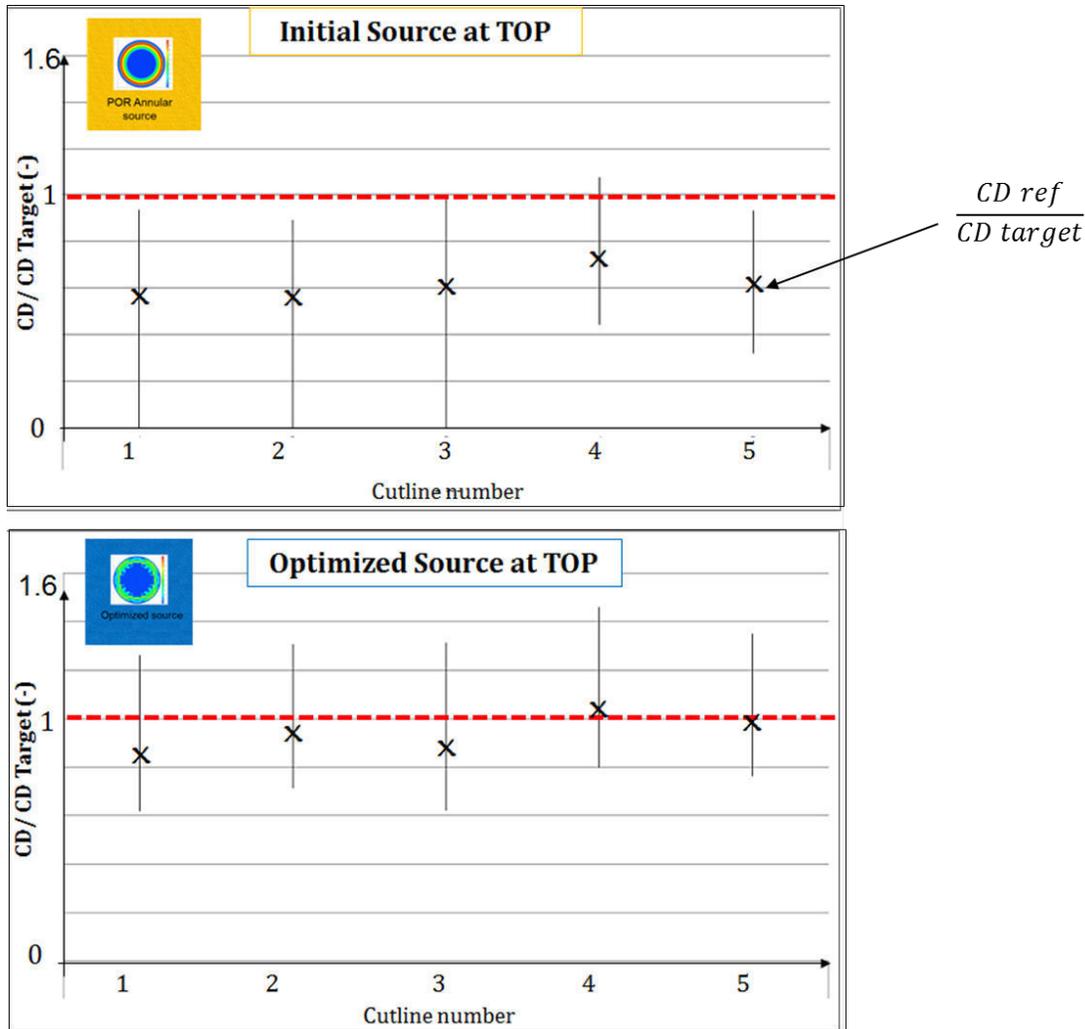


Figure 83 PV band comparison results: Contour CD / the target CD for different 2D patterns: upper part representing the results for POR Annular source, while the other for the optimized source for Bottom resist height

The PV band analysis for 1D feature provides no performance degradation using optimized source. It is valid as well for the bottom and for the top of the resist height. The before and after optimization results show quiet a similar tendency. Although by applying the optimized source a tight improvement can be seen as the obtained PV bands, the normalized CD_{REF} are closer to the target (red line). While focusing on the 2D patterns a PV band improvement can be observed by applying the optimized source. Meanwhile it results in a closer target for the top by the optimized source. Therefore the optimized source showed to be better according to these results.

5.2.4.2 Method 2: Application of an LMC at different resist height

The Method 2 refers to the application of a conventional LMC (see chapter 2 for LMC). The additional item on standard LMC is the notion of the resist profile by the calibrated resist 3D model. As setup, several resist heights can be set as resist height of interest whereat the defects

are checked. So these set heights are predefined resist heights. The choice is done at locations, where user expects having issues: such as at the top of the resist, where the possible top loss occurs. Therefore this method 2 remains resist profile limited. It is due to the fact that it will not provide a global scan along the resist height – only at selected heights. But even if it is resist profile limited, it is still a fast method in terms of CPU time that can be used for full-chip application.

In order to try this method’s capability for the source comparison and validation, the set resist profile aware LMC, is tried on a small layout. The objective, so the success criterion is to demonstrate that the optimized source will enable less found defects versus the initial source. For this purpose different kinds of defects were checked, similar to production setup. They include bridging and necking faults. Here for simplicity reason, they are noted by letters from A to G (80% are bridging defects). This check is done at the middle and at the top of the resist height considering that, it might reflect more bridging and necking defects.

The results are shown on figure 84. It reveals in the fact that the application of the optimized source provides less defects occurrence principally for bridging detector (ex. on figure 60): a decrease of 74 % of the number of detected defects against the initial source for the middle resist height case. It provides a decrease of 17 % for the top resist height. Thus it can be concluded, that this method works well and the optimized source is proved to be more efficient than the initial one. Note that bottom data are not presented as for these defects top and middle are more pronounced.

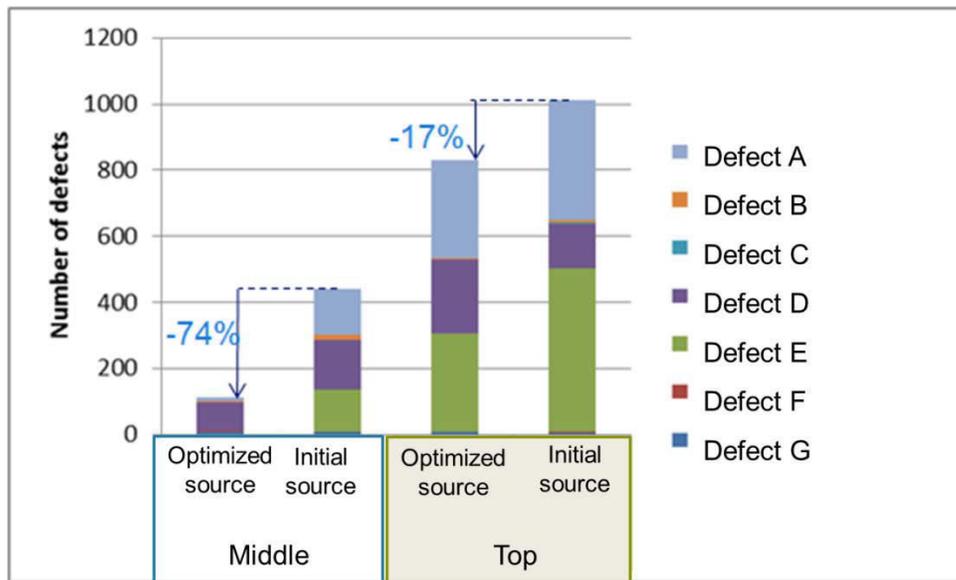


Figure 84 Resist profile aware LMC at the middle and at the top of the resist height indicating the decreased defect number in percentage by the optimized source for the middle an top resist height

5.2.4.3 Method 3: Full chip scan detecting hotpots with the higher delta CD through resist

The Method 3 refers to a full chip scan. It checks the CD through the resist height and takes the maximum and minimum CD, see figure 85. Note that illustrated example on the “bumped” shape resist profile is volunteer, since we met this issue what the optimization is targeted to handle. This range in CD ($CD_{max} - CD_{min}$) reveals then in profile information: it means the smaller the CD range is the straighter the resist profile – thus the better the process performance is.

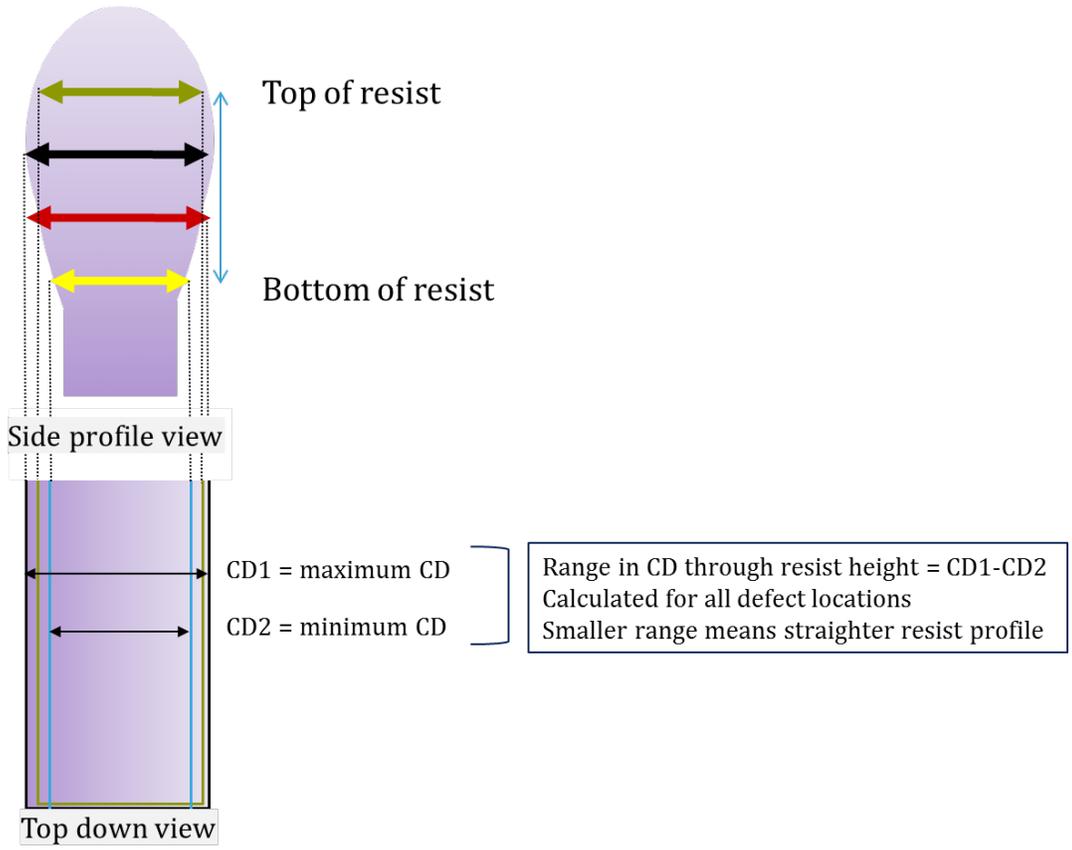


Figure 85 the Full chip scan detecting hotspots with the higher delta CD through resist height

Thus method 3 works as a detector. The operational principle is similar to the PV band analysis; the only difference is that, it focuses on the resist profile, the CD range through the resist height. The objective is to detect all the patterns that have high delta CD through the resist height. The success criterion of the source validation is to validate the optimized source enabling less detected defects versus the initial source.

This method is tried on the optimization layout, which was applied for the source optimization. In addition we launched it on full chip. The obtained results are shown on figure 86 and 87. The emphasis is on the defect counts: the defect counts in function of the range of CD through the resist height. As first look on these figures, the application of the optimized source enables to obtain a smaller range of the CDs through the resist height. By means the bumped shape of the resist profile becomes straighter.

Figure 86 , belonging to a test case on the optimization layout clip, results in an calculated average range in CDs along the resist height in 7 nm for the annular source and 4.9 nm for the optimized source. It means much less variation along the resist height with the optimized source.

Figure 87 shows results from a real full chip run. It results in a calculated average range in CDs along the resist height in 7.1 nm for the annular source and 5.5 nm for the optimized source. Thus it means that the optimized source could meet the requirement of reducing the bumped shape against the initial source.

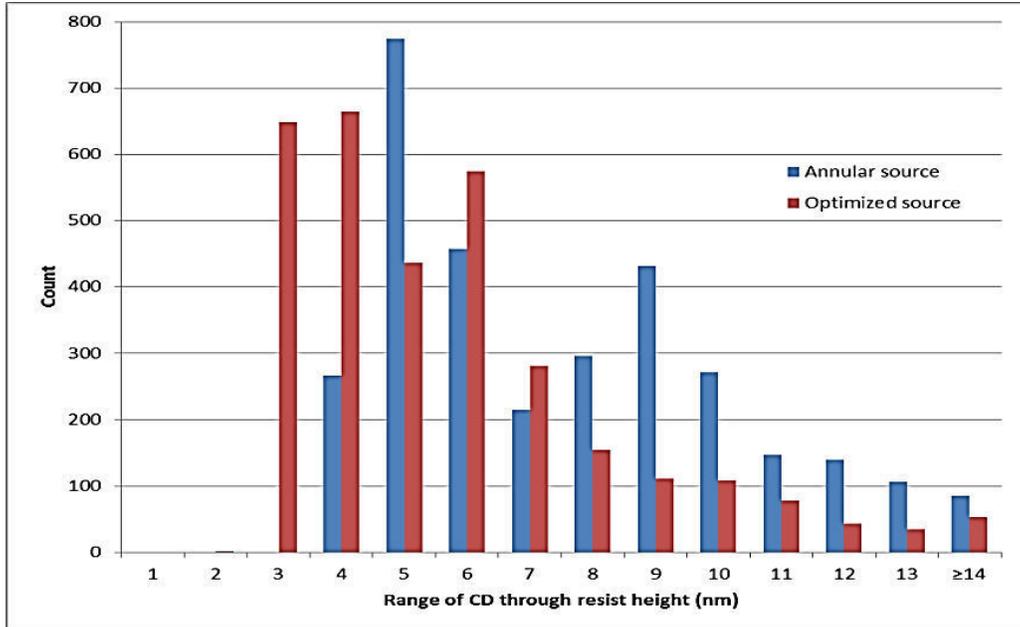


Figure 86 Analysis results on patterns on the optimization layout; annular source (blue color) versus optimized source (red color)

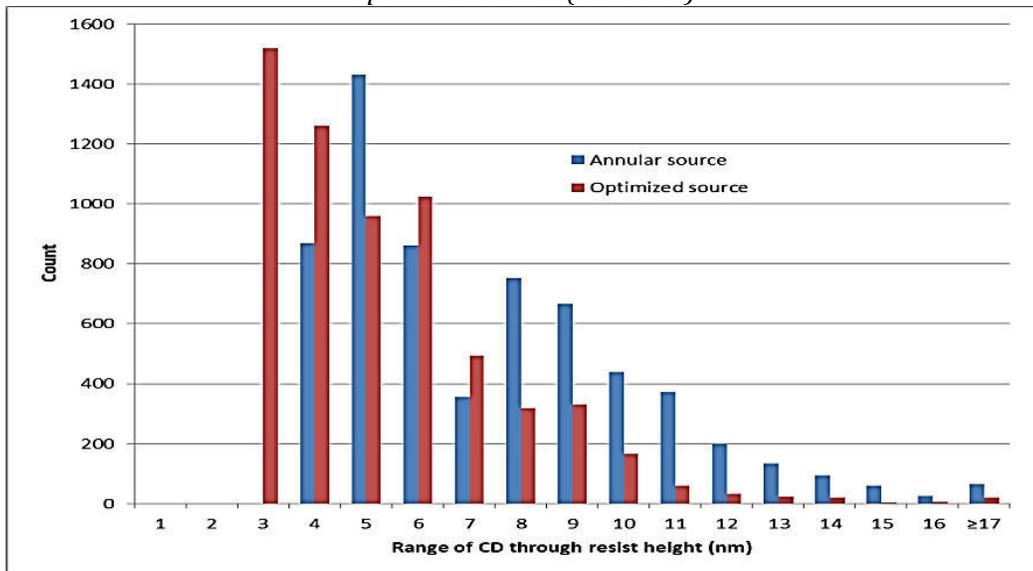


Figure 87 Analysis results on product layout annular source (blue color) versus optimized source (red color)

5.2.5 Partial conclusion along the resist profile related mitigation

The resist profile aware source optimization concept is demonstrated with success by simulation results showing that the optimized source performance is superior to the initial one. So it is promising. Although further efforts need to be put in this optimization in order to establish a more stable flow. It is due to the fact that the setup of the cost function in the very beginning of the optimization requires a fine tuning that might need higher level of experiments on this topic.

5.3 Conclusion

By this chapter, our main objective was to provide performant solution for process variability mitigation. Following the logical pathway of the previous chapter, the diminution (or compensation) of the process variability were discussed through the mask 3D and through the resist 3D induced effects.

The mitigation, or more precisely the compensation, among the mask induced effects refers to the investigation at scanner level with the support of the computational lithography. For the compensation of these mask related effects, we applied a wavefront optimization. For this purpose, the optimization is realized by the support of the FlexWave lens manipulator at scanner level. The performed optimization work refers to simulation and experimental investigation. Our results enabled an imaging enhancement by minimizing mask topographic effects provided a pronounced increase in uDoF.

The mitigation among the resist induced effects refers to the optimization through the optical intensity distribution through the resist thickness. For this optimization, a source optimization was applied to decrease such resist effects. Our optimization results, based on simulation work, the optimized source enhanced the imaging quality providing straighter resist profile. In order to validate the obtained source, various methods were discussed. They were developed and tried showing promising results. According to these results from the source validation, the optimized source's performance showed to be superior to the original source.

Overall, our solutions for process variability mitigation showed to be efficient enough to be applied in fab. Of course further fine tuning might be required for a smoother operation of the introduced methodology combined with more technical experience on e.g. the source optimization part. As consequence the construction of our global flow is accomplished by this last step. Note it might be completed with a quantification stage that would monitor the process once the presented flow is applied.

General conclusion

The currently used optical lithography reaches its limits from resolution point of view. It is despite of the fact that various techniques have been developed to push this limit as much as possible. Indeed new generation lithography exists such as the EUV, but it is still not reliable to be applied in mass production. Thus in order to maintain a robust process for advanced nodes 28 nm and beyond, alternatives techniques are required. That is emphasizing on the “everything is linked”: process (tool, stack, and mask), computational lithography & metrology. For this purpose, for our investigation, the metal layer of the 28 nm node is a prominent candidate containing numerous complex features to be well patterned. Thus it provides pronounced challenges to be solved and to be studied.

In the light of the above mentioned items, this dissertation work was addressed to mitigate the lithographic process variability. Within this frame, our main objective was to participate to the challenge of assuring a good imaging quality for these shrunk dimension layers with an acceptable uDoF. In order to accomplish this task, we proposed and validated a flow that might be later implemented in the production. The proposed flow consists of the detection of the most critical patterns that are impacted by effects coming from the mask topography and the resist profile (and wafer topography, which was not explored here). Furthermore it consists of the mitigation and compensation of these effects, once the critical patterns are captured.

In order to construct our process variation flow, as a starting point, we took tight attention on current modeling prediction capability. It is due of the fact that it relies on the capability of computational lithography, more precisely on OPC modeling. So it is because of the fact that the conventional OPC models became obsolete for shrunk nodes technologies. Today, our comprehension should be widened further on the related mask & resist (& wafer) 3D effects occurring during the exposure. And this comprehension needs to be implemented in lithographic modeling work.

In the continuity in this dissertation, the emphasis was put then on the construction of a mask 3D and resist 3D models with enhanced upfront process variation prediction capability (versus standard OPC models). The calibrations involved the ensemble of work from the proof of concept until the validation on wafer. Our mask and resist 3D model showed an efficient capability for the layer of interest providing an efficient support for further work for the detection and for the mitigation part.

As sub-sequent stage, using the calibrated 3D models, a simulation based detection methodology was set. We approached this task among the mask related effects through the pattern dependent best focus offset and the resist related effects through resist profile:

- *The detection method among the mask induced effects* includes a delta best focus shift analysis through the Bossung curves. The behind idea is to setup an efficient method to catch patterns showing high best focus shift. For this reason the proposed method refers to the Bossung curves’ slopes investigation. This method is validated through PW analysis and its findings correlate well with 60 nm of uDoF (usable Depth of Focus) from the PWQ

(Process Window Qualification). In addition the tilted or flat Bossung curves are a challenge for this method since it is based on a slope analysis. Therefore this method needs further fine-tuning and investigations before its implementation in fab.

- For *the detection method among the resist profile induced effects*, two approaches are proposed. One refers to the Contour-to-Contour (C2C) application that focuses on delta CD comparison through the resist height. While the other refers to an LMC supported bridge detector through the resist height. Both approaches were developed and tried with success on full chip. They showed reliable performance and capability of detection matching well with PWQ findings; proofed results by wafer inspection. Further improvement might be applied in order to handle the detected defects counts enabling less “defects handling” time.

Our results showed, in addition, that the resist profile related effects have more impact on the reduced overlapping PW, thus the decreased uDoF – the observed PW limiters are mainly critical patterns showing pronounced resist profile related effects. As consequence for these nodes, the resist profile related effects might be considered as higher priority task.

As following step, based on the findings from by the set detectors, the emphasis was put on the mitigation solutions of the lithographic process variability completing our proposed flow. Anew, we approached this task among the mask related effects through the pattern dependent best focus offset and the resist related effects through the profile.

- *The mitigation, or more precisely the compensation, among the mask induced effects* reveals in an investigation at scanner level with the support of the computational lithography. For the compensation of the mask topological effects, we applied a wavefront optimization on the basis of the fact: the mask topological effects are the responsible for the delta phase assignment having corresponding phenomenon to the one induced by wavefront aberrations of the projection optics (Wong A. , 1992). As consequence, the mask topography related effects result in a wavefront distortion acting as lens aberration. Thus the compensation of these effects is possible by wavefront correction. For this purpose, we used the support of the FlexWave lens manipulator at scanner level. The performed wavefront optimization refers to simulation and experimental investigation. According to our results, an imaging enhancement by minimizing mask topographic effects was provided by a pronounced increase in uDoF.
- *The mitigation among the resist induced effects* reveals in an optimization along the optical intensity distribution through the resist thickness. For this optimization, a source optimization was applied to decrease such resist effects because the source has a crucial rule on the intensity distribution on the pattern image. Thus by its optimization, a straighter resist profile can be assured after exposure having less issues for the subsequent etching stages with an enhanced uDoF. According to our source optimization results, based on simulation work, the optimized source enhanced the imaging quality providing straighter resist profile and an increased uDoF. For source validation purpose, various methods were developed and tried with success. By the support of these source validation methods, the optimized source’s performance showed to be superior to the

original source. Therefore our results are promising, although further effort might be needed to establish a more robust optimization flow applicable in fab.

Accordingly, the last part permitted to complete and confirm our proposed flow by this dissertation work. It participates well to the diminution of the lithographic process variability along critical patterns of the 28 nm technology node metal layer. The presented results are promising providing key points for further improvements to be realized.

The implementation of the presented ideas in production will require more in-depth investigation and fine-tuning. Although as an already applicable achievement, the R3D model construction methodology is already under implementation for 28 nm technology nodes. Furthermore, it can be adapted for 14 nm technology nodes as well.

Annex 1

The Abbé's method

As example to introduce Abbé's method, let consider a photomask having regions where the light pass through made of glass otherwise opaque covered by chrome. So the photomask is described by two following discrete scalar functions: the mask transmittance $t_m(x, y)$ and the phase $\varphi(x, y)$ derived from the electrical field described at the x-y mask plane as follow:

$$E_{mask}(x, y) = \sqrt{t_m(x, y)}e^{j\varphi(x, y)}$$

Equation 1

Where t_m and φ are expressed as follow – here t_0 and φ_0 reveal in the clear area of the photomask where light pass through or in other word where feature is present - :

$$t_m(x, y) = \begin{cases} t_0, & \text{if } (x, y) \in \text{transparent region of photomask} \\ 0 & \text{otherwise} \end{cases}$$

Equation 2

$$\varphi(x, y) = \begin{cases} \varphi_0, & \text{if } (x, y) \in \text{transparent region area of photomask} \\ 0, & \text{otherwise} \end{cases}$$

Equation 3

Once the emitted light is diffracted, the electric field of the diffracted pattern $T_m(f_x, f_y)$ is referred to the Fraunhofer diffraction⁴⁷ integral based on the electric field incident to the photomask and the mask transmittance:

$$T_m(f_x, f_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} E(x, y)t_m(x, y)e^{-2\pi j(f_x x + f_y y)} dx dy$$

Equation 42

Where f_x, f_y reveals to the spatial frequencies of the diffraction pattern as scaled coordinates in the x'-y' diffraction plane, the entrance to the objective lens (Mack C. A., 2006) :

$$f_x = \frac{nx'}{z\lambda}$$

Equation 4

$$f_y = \frac{ny'}{z\lambda}$$

⁴⁷ Fraunhofer Diffraction also called far-field diffraction: ~ region is defined where $z \gg \frac{\pi w^2}{\lambda}$, where "z" is the distance of the aperture from the screen, "w" slit width and "λ" incident light wavelength. Optical Lithography is placed in this region.

Equation 543

Where z is the distance of the photomask to the x' - y' diffraction plane, λ is the applied light wavelength, n is the refractive index of the medium the complete system remain in.

Furthermore the Equation 45, the diffracted pattern or electric field distribution as it arrives to the objective lens; it can be understood as the Fourier Transform of the mask pattern transmittance:

$$T_m(f_x, f_y) = F \{E(x, y)t_m(x, y)\}$$

Equation 6

Then the light intensity is calculated as square of the electric field magnitude:

$$I(f_x, f_y) = |T_m(f_x, f_y)|^2 = |F\{E(x, y)t_m(x, y)\}|^2$$

Equation 7

Once the light emitting the photomask arrives at the pupil level only a portion in function of the NA is collected by the objective lens in the x' - y' plane (Figure 1). The pupil is characterized by P , the objective lens pupil function that describes the transmittance of the objective lens from the inlet to the exit pupil:

$$P(f_x, f_y) = \begin{cases} 1, & \sqrt{f_x^2 + f_y^2} < NA/\lambda \\ 0, & \sqrt{f_x^2 + f_y^2} > NA/\lambda \end{cases}$$

Equation 8

The light passing out the objective lens can be defined by the product of the pupil function and the diffraction pattern. The related electric field at this level is resulting in the inverse Fourier transform of the transmitted diffracted pattern as follow:

$$E(x, y) = F^{-1}\{T_m(f_x, f_y)P(f_x, f_y)\}$$

Equation 9

Thereafter the intensity distribution (the aerial image), $I(x, y)$ at the imaging plane is now a simple square of the magnitude of the expressed electric field by Equation 51 :

$$I(x, y) = |E(x, y)|^2$$

Equation 10

In summary so far : a Fourier transform occurs in front of the objective lens, while an inverse Fourier transform takes place at objective lens level, see figure below.

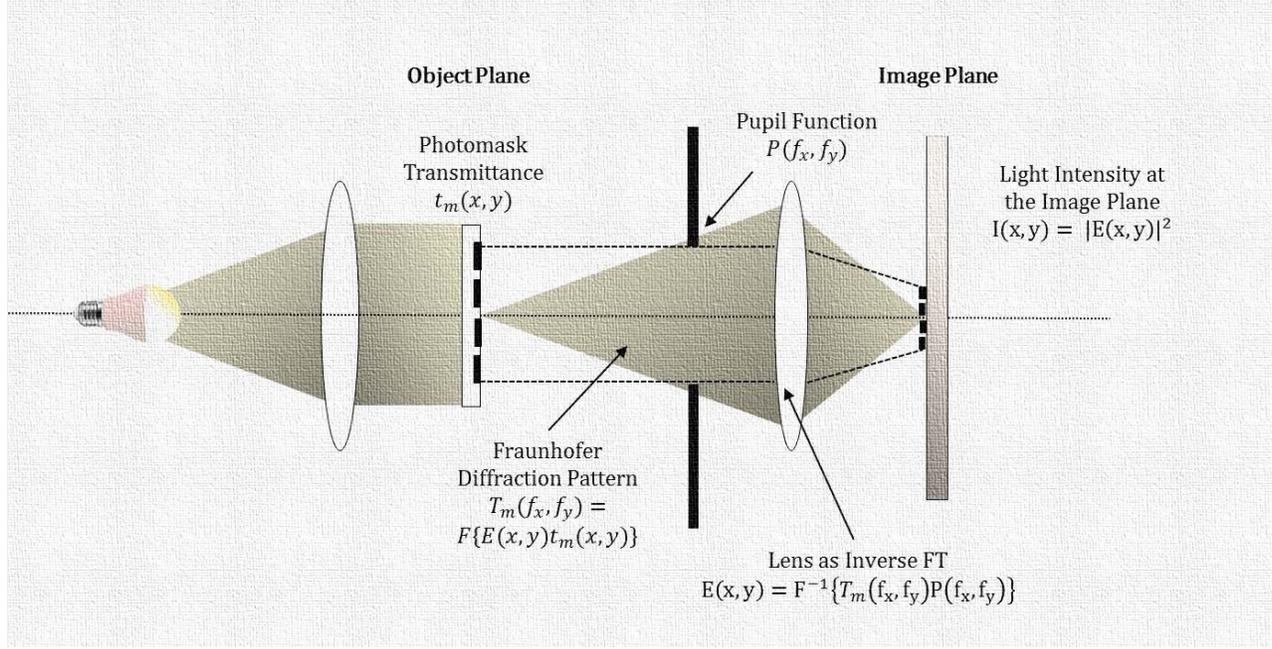


Figure 1 Schema of passing light indicating the basic considerations for lithographic modeling

Note that the above defined $I(x, y)$ under this mathematical expression (Equation 52) is only valid for point source and is much more complex for a partial coherent light where the diffraction pattern is spread in space. As consequence the electric field sustain a spatial shift of f'_x and f'_y . Thus at wafer level the electric field is modified as, if magnification is equal to one:

$$E(x, y, f'_x, f'_y) = F^{-1}\{T_m(f_x - f'_x, f_y - f'_y)P(f_x, f_y)\}$$

Equation 11

Then the light intensity at the image plane will be the sum of the individual intensities from each point source, known as the Abbé's method (Mack C. , 2008)for partially coherent image calculation. Let the source transmittance be $S(f'_x, f'_y)$ then the total image intensity can be determine as follow:

$$I(x, y) = \frac{\iint I(x, y, f'_x, f'_y)S(f'_x, f'_y)df'_x df'_y}{\iint S(f'_x, f'_y)df'_x df'_y}$$

Equation 12

The source transmittance, $S(f'_x, f'_y)$ in the pupil plane and can be further expand in function of the coherence property, σ . Applying the coherence factor as follow $\sigma_x = f_x \lambda / NA$ and $\sigma_y = f_y \lambda / NA$, the normalized source shape is then (Rodrigue, 2010) :

$$S'(\sigma'_x, \sigma'_y) = \frac{S(\sigma'_x, \sigma'_y)}{\iint S(\sigma'_x, \sigma'_y)d\sigma'_x d\sigma'_y}$$

Equation 13

Therefore the final form of the total light intensity is written as:

$$I(x, y) = \iint I(x, y, \sigma'_x, \sigma'_y) S(\sigma'_x, \sigma'_y) d\sigma'_x d\sigma'_y$$

Equation 14

So in summary, Abbé's imaging formulation considers the source as composed by a number of source points. Furthermore the total light intensity, the aerial image, is calculated as the superposition of the entire light intensity distribution corresponding to these source points. However for the fact that the intensity calculations for every source points would require a Fast Fourier Transform, this method remains limited in term of simulation time.

The Hopkins formula

In order to further simplify and approximate the Abbé's method or model, the Hopkins formula is applied. It is also called Hopkins diffraction model that will be shown by a simplified, one dimensional example. Here the electrical field of light at the image plane is written as:

$$E(x, f'_x) = \int P(f_x + f'_x) T_m(f_x) e^{-2\pi j(f_x + f'_x)x} df_x$$

Equation 15

The light intensity is obtained then :

$$\begin{aligned} I(x, f'_x) &= E(x, f'_x) E^*(x, f'_x) = \\ &= \left(\int P(f_x + f'_x) T_m(f_x) e^{-2\pi j(f_x + f'_x)x} df_x \right) \times \left(\int P^*(f''_x + f'_x) T_m^*(f''_x) e^{-2\pi j(f''_x + f'_x)x} df''_x \right) \\ &= \iint P(f_x + f'_x) P^*(f''_x + f'_x) T_m(f_x) T_m^*(f''_x) e^{-2\pi j(f''_x - f_x)x} df_x df''_x \end{aligned}$$

Equation 16

The total light intensity is obtained similar as for Abbé's method by an integral over the source. Although for Hopkins the integration over the source is done before the inverse Fourier integrals. An intermediate term is usually applied, so called the Transfer Cross Coefficient (TCC)

$$TCC(f_x, f''_x) = \int P(f_x + f'_x) P^*(f''_x + f'_x) S''(f'_x) df'_x$$

Equation 17

$I_{total}(x)$, the output image intensity at position x is calculated as follow in the image plane:

$$I_{total}(x) = \iint TCC(f_x, f''_x) T_m(f_x) T_m^*(f''_x) e^{-2\pi j(f''_x - f_x)x} df_x df''_x$$

Equation 18

The application of this intermediate term of the TCC has the advantage that it is independent of the mask function due to the fact that it is a function of the spatial frequencies that depends only on the illumination source and on the pupil. Thus it enables the possibility of computing and storing data in a pre-phase resulting in a more efficient computing than Abbé's method. Hence it is widely used in commercial software's embodiments for OPC applications since OPC will modify the pattern geometry for better pattern replica at wafer level.

6 Résumé en français

Mots clés : lithographie, RET (*Resolution Enhancement Techniques*), modèle OPC (*Optical Proximity Correction*), masque 3D, résine 3D

Dans le domaine de la microélectronique, la constante problématique de la miniaturisation des circuits électroniques se traduit par une diminution des dimensions des structures afin d'augmenter la densité et les performances des circuits. Si ces dimensions critiques s'éloignent de la valeur pour laquelle le rendement du circuit est maximal, le circuit ne respecte plus les spécifications définies dans le cahier des charges. Ceci peut être le résultat d'une dégradation de la chaîne de production. Par conséquent il est indispensable de développer des équipements de fabrication et de métrologie fiables et robustes, nécessitant un savoir-faire de plus en plus complexe.

Les procédés de fabrication d'un circuit intégré actuel requièrent une centaine d'étapes à réaliser sur la plaque de Silicium. Il est possible de diviser en 3 groupes distincts l'ensemble de ces étapes :

- La définition des motifs à partir d'un réticule par lithographie puis gravure ;
- La formation des couches par dépôt, oxydation ou métallisation ;
- La modification d'une couche (d'un matériau) par implantation ionique et par diffusion.

Parmi ces étapes essentielles, c'est la lithographie qui limite principalement la miniaturisation du fait que c'est elle qui permet de définir les structures nécessaires au circuit. A l'heure actuelle, c'est la lithographie optique, ou photolithographie, qui est la plus utilisée dans l'industrie.

Le but de développements lithographiques est de diminuer la dimension critique correspondant à la plus petite dimension réalisable par le procédé. Il est possible en passant par l'optimisation de la résine dans laquelle l'image est formée et par l'optimisation de l'équipement. Des lithographies de nouvelle génération (NGL) sont à l'étude, et notamment la lithographie Extrême UV (EUV), mais les retards accumulés dans son développement font que la technique actuelle, la lithographie 193 nm, continue à être exploitée pour les nœuds 28 nm et au-delà.

Le cœur du développement de la photolithographie est l'optimisation de son équipement, qui est très complexe et qui arrive à ces limites de capacité en termes de résolution des motifs dans la fenêtre du procédé souhaitée. L'une des techniques alternatives, afin de suivre la miniaturisation malgré ces limitations, est d'utiliser les techniques d'améliorations de la résolution : RET (en anglais *Resolution Enhancement Technique*) tels que la correction des effets de proximité : OPC⁴⁸ (*Optical Proximity Correction*) au niveau du dessin sur le masque.

L'objectif de la thèse

Dans ce contexte de complexité croissante du procédé de fabrication, le contrôle dimensionnel (ou variabilité) toléré pour un nœud technologique 28 nm (et au-delà) de taille de grille est inférieur à 1nm

⁴⁸ OPC : ~ est une technique qui modifie sélectivement la forme et la taille des motifs sur le masque afin d'obtenir plus précisément les motifs envisagés au niveau de la plaque.

(3 σ). La quantification, la modélisation, le contrôle puis la correction des facteurs de variabilité sont donc devenus des missions essentielles de l'équipe procédé lithographique et de l'équipe RET de la société STMicroelectronics.

Dans ce cadre-là, l'objectif de la thèse est de mettre en place une méthode de correction des facteurs de variabilité afin de diminuer la variabilité des motifs complexes pour assurer une résolution suffisante dans la fenêtre de procédé. Ces motifs complexes sont très importants, car ce sont eux qui peuvent diminuer la profondeur du champ commune (uDoF). Cette profondeur de champ est la plage de focus commune à toutes les structures du dessin, pour laquelle les CD des motifs obtenus dans la résine restent conformes aux spécifications. Le but global est toujours d'augmenter ce paramètre uDoF.

Afin d'atteindre l'objectif, il convient dans un premier temps de détecter et sélectionner les motifs les plus complexes par simulation, en construisant une méthodologie efficace. Dans un deuxième temps le but est de diminuer la variabilité en se basant sur ces motifs. Les travaux effectués, se focalisant sur les résultats et les faits marquants, sont résumés dans ce document.

6.1 Définitions et concepts

Dans le but d'une meilleure compréhension des résultats de cette thèse, il est important de définir au préalable les concepts les plus importants.

- *La profondeur de champ* (Depth of Focus, DoF, en anglais), (figure 88) : pour un système optique, il existe une distance de part et d'autre du plan focal pour laquelle le CD reste dans une gamme de tolérance compatible avec les spécifications. Ce paramètre est défini par l'équation 1, pour une longueur d'onde λ donnée. La DoF est définie pour une structure donnée. Dans le cas d'un layout avec des structures multiples, la profondeur de champ commune à tous les motifs sera appelée uDoF (usable Depth of focus).

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

Equation 44

Où k_2 est un facteur de procédé, λ est la longueur d'onde d'exposition (193 nm) et NA est l'ouverture numérique de la lentille.

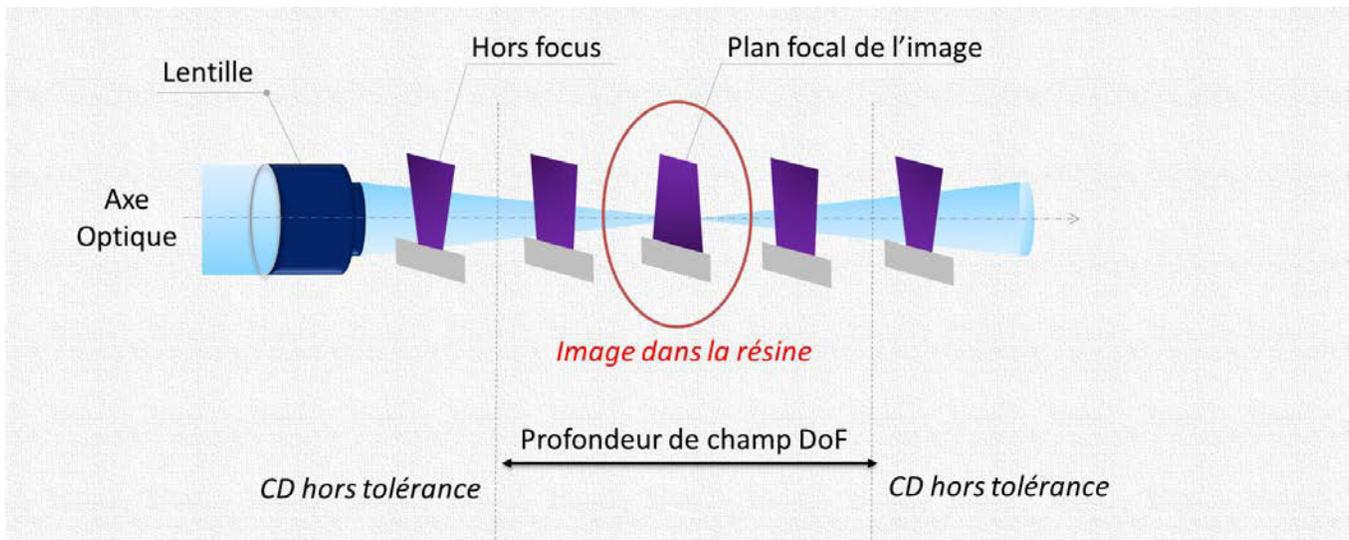


Figure 88 Illustration de la profondeur de champ (DoF) en termes de dimension critique (CD) dans la résine. Le CD dépend de la position d'observation sur l'axe optique.

- La latitude en énergie (EL) : de la même manière que la profondeur de champ DoF, le concept de latitude en énergie correspond à une plage de dose d'exposition pour laquelle le CD reste dans les spécifications. Cette zone de tolérance est définie à un focus constant. Donc la latitude en énergie est définie pour une plage de focus donnée.
- La courbe EL-DoF : afin d'analyser la variation en dose et en focus permise, la courbe EL-DoF est tracée, voir figure 89, où tous les points sous la courbe correspondent aux plages de variation valides qui permettent de garder le CD dans les spécifications.

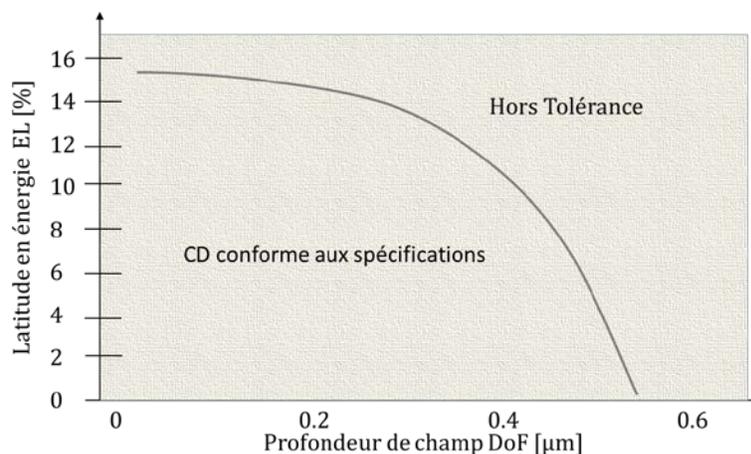


Figure 89 Courbe EL-DoF ; Exemple de la gamme en énergie et profondeur de champ qui permettent de garder le CD d'un motif donné conforme aux spécifications

- La fenêtre de procédé (PW pour process window) représente les plages de variation tolérées pour la position du plan focal et la dose d'exposition, pour lesquelles les motifs s'impriment dans la limite des spécifications concernant le CD. Cette fenêtre de procédé est

sensée être la plus grande possible afin de minimiser la sensibilité du procédé à certaines instabilités liées à l'équipement, à la topographie présente sur les plaques, etc.

- Les courbes de Bossung, voir figure 90 : elles représentent la variation de CD d'un motif donné en fonction de la dose et du focus. De telles courbes permettent également de déterminer la sensibilité des motifs aux variations du procédé et de connaître la fenêtre du procédé (PW) en observant leur allure. Lorsque la courbe est plutôt incurvée, le motif sera fortement sensible à la position du plan focal. Si les courbes sont très espacées, alors le motif sera plus sensible aux variations de dose.

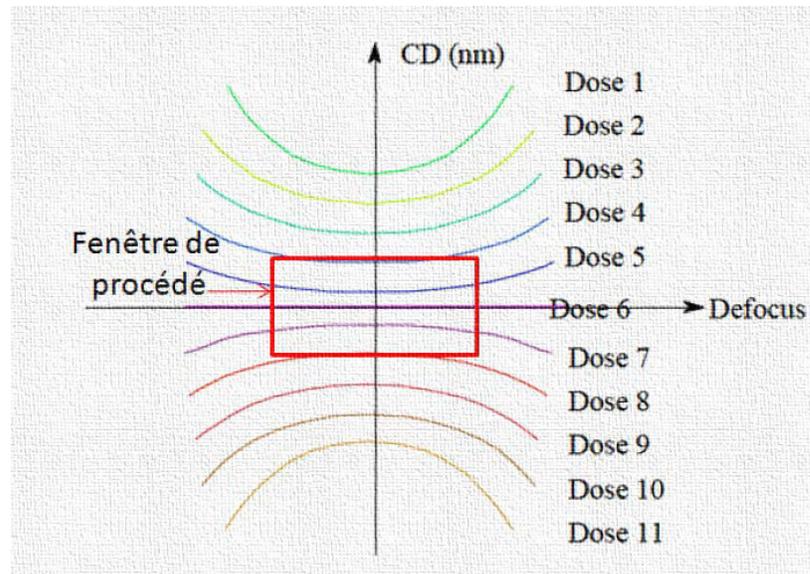


Figure 90 Illustration schématique des courbes de Bossung

- Le procédé appelé « FEM » (Focus Exposure Matrix en anglais) : Exposition d'une puce identique sur l'ensemble de la surface de la plaque de Silicium mais pour différentes conditions de dose et de focus selon la position sur la plaque. Cela permet un bon centrage du procédé ; une mesure de la fenêtre de procédé (PW).
- La qualification de fenêtre de procédé, le PWQ (Process Window Qualification en anglais): Cette analyse correspond à une méthode d'inspection des plaques, à partir de contrôles CD-SEM, couplée à un algorithme permettant de détecter les défauts critiques. Le PWQ est utilisé pour qualifier l'OPC, comme ces résultats sont utilisés comme référence pour les analyses de PW (simulé). En outre, il est appliqué pour qualifier le nouveau masque et pour qualifier le procédé lithographique. Pour un niveau spécifique, le PWQ est réalisé sur les plaques FEM après l'étape de gravure et la planarisation mécano-chimique (Le CMP).
- Le «Lithography Manufacturability Check, LMC » : c'est un outil de vérification utilisé systématiquement. Il est basé sur des simulations. Cet outil permet de repérer les motifs critiques dans un champ et d'extraire les régions du design sensibles aux variations de

procédés. Le LMC est fait après OPC, donc avant la commande du masque. Il peut être appliqué soit sur une partie du layout, soit sur le layout entier (Brion, 2015).

6.2 Plan expérimental

Les résultats sont obtenus sur un équipement de type scanner à immersion ASML NXT : 1950i (Caractéristiques : réduction 4X avec une ouverture numérique $NA=1.35$ et longueur d'onde $\lambda = 193$ nm). La plateforme de Tachyon et le logiciel Panoramic ont été appliqués pour toute la partie de modélisation et de simulation. L'étude est focalisée sur un nœud sub-32 nm d'un niveau métal. L'empilement correspondant sur les plaques testées représentatives du procédé est composé des couches suivantes : 40 nm de couche antireflective BARC (Bottom Antireflective Coating), 100 nm de résine à amplification chimique (CAR), 30 nm de matériau déposé sur la résine appelé top coat.

6.3 Les résultats et les faits marquants

6.3.1 La Détection par application des modèles avancés : masque 3D / résine 3D

Pour répondre au premier objectif, c'est-à-dire construire une méthodologie efficace de détection des motifs les plus sensibles aux variations de procédé, il faut être capable de prédire leur exposition. Pour sélectionner ces motifs par simulation il faut donc un modèle prédictif pour une détection efficace (pour construire des méthodologies de détection efficaces). L'effort dans un premier temps est mis sur la capacité prédictive du modèle OPC par la calibration des modèles masque et résine 3D avec le fournisseur ASML/Brion. Remarque : pour plus de détails sur les modèles « wafer » 3D, qui n'est pas détaillé ici, le lecteur est invité à lire la référence suivante (Michel, 2014).

6.3.1.1 Les modèles de détection

6.3.1.1.1 Le modèle masque 3D

On appelle communément un modèle de masque 3D, un modèle qui considère les épaisseurs et les profils des motifs présents sur le masque. Conformément à la littérature sur le modèle lithographique de masque 3D et aux expériences dans l'industrie, ce type de modèle 3D de masque a été appliqué car la diminution des dimensions des motifs des nœuds technologiques avancés se traduit généralement par une réduction de la profondeur de champ uDoF (section 6.1) et de la fenêtre de procédé (oPW). Plusieurs facteurs influencent ces paramètres, mais celui qui devient le plus en plus impactant est le fait que la déviation du meilleur focus est fortement dépendante du design de layout (« pattern dependent best focus shift » en anglais) (Finders, 2014). Ceci vient des effets de la topographie du masque qui conduisent à la déviation de phase et qui induit une mauvaise définition de l'image sur plaque.

Cette déviation du meilleur focus en fonction du design correspond à un effet de masque 3D. Elle ne peut donc pas être prédite par un modèle 2D conventionnel (type approche de Kirchhoff). Néanmoins, à l'heure actuelle, la méthode la plus appliquée pour la modélisation du masque est l'approche de Kirchhoff qui consiste à considérer que le masque est infiniment fin et que son épaisseur est donc négligeable. (Saied, 2011). Néanmoins pour le domaine où les dimensions des

motifs au niveau du masque deviennent comparables ou plus petites que la longueur d'onde λ , ce modèle Kirchhoff est limité (Saied, 2011) (Liu, 2012) (Finders, 2014).

La figure 91 représente la problématique importante liée au meilleur focus en prenant un modèle masque 3D comparé à un modèle masque 2D. L'objectif de cette simulation est de déterminer la déviation de meilleur focus venant de la topographie du masque sur un réseau de lignes dont la période varie de 90 nm à 170 nm (p). Pour cette étude le CD des lignes est égal à l'espace entre lignes. Les résultats de modèle 3D sont montrés avec une courbe d'interpolation polynomiale de 3^{ième} ordre. Comme on le voit sur la figure, l'approche 2D reste constante (ligne violette) – ceci est lié au fait que les effets provenant de la topographie du masque ne sont pas pris en compte (Wong, 1992) (Erdmann, 2005). Par conséquent, pour un procédé ayant une uDoF inférieur à 100 nm cela devient donc non négligeable : l'utilisation d'un modèle masque 3D est nécessaire.

Remarque : la déviation du meilleur focus peut être considérée comme le facteur principal de la diminution de la profondeur de champ uDoF.

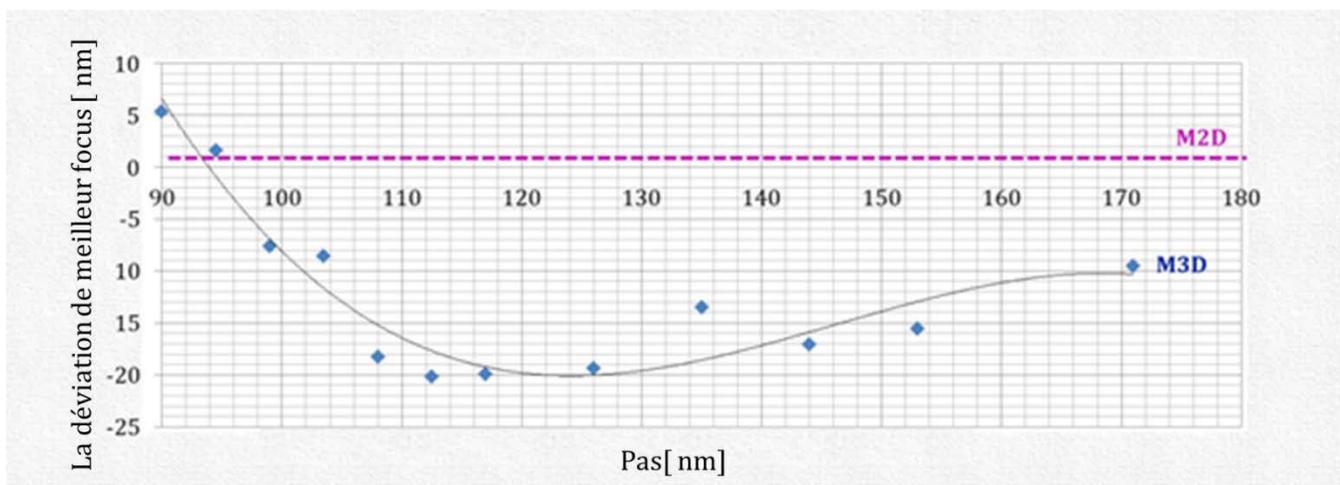


Figure 91 La déviation de meilleur focus (Illumination annulaire) en fonction du pas du réseau, obtenue avec un modèle masque 2D (ligne violette) et avec un modèle masque 3D (points bleus avec une interpolation d'ordre 3)

Le caractère 3D du masque doit donc être pris en compte pour construire un modèle précis du masque. Le déroulement de la calibration est montré sur la figure 92 – ce déroulement correspond à celui du modèle de Brion Technologies, fournisseur de solutions OPC. Le modèle est établi à partir du layout, de la source d'illumination et de la topographie du masque. Une librairie de masques 3D est introduite. Elle dépend des caractéristiques du masque (valeurs n et k et épaisseur des matériaux ...), mais est indépendante du layout. Cette librairie contient des résultats rigoureux de simulation du champ électromagnétique (EMF). Le moteur masque 3D peut donc traiter n'importe quel design et simuler l'image obtenue à la sortie du réticule. Cette image est ensuite envoyée vers le moteur optique puis le moteur résine. Finalement pour compléter la calibration, les données sont comparées avec des données de mesure sur plaque (Brion Technologies, 2008) (Liu, 2012).

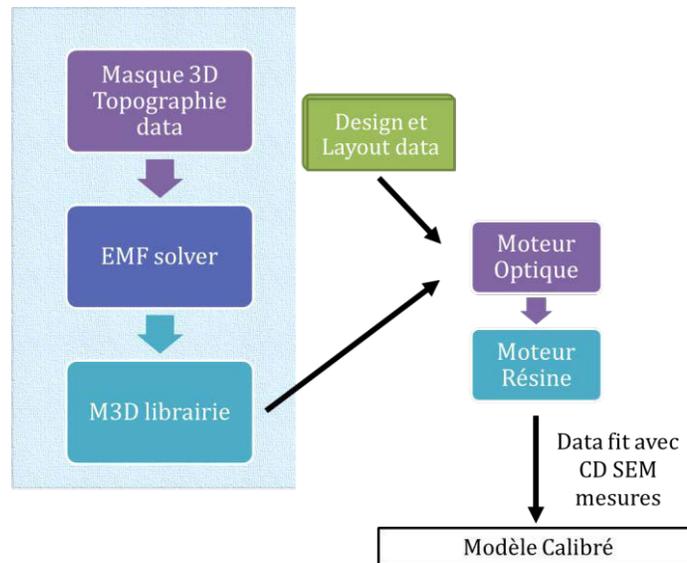


Figure 92 Déroulement global de calibration. La partie encadrée en bleu clair correspond aux étapes du moteur masque 3D.

Notre modèle masque 3D est calibré à partir d'un ensemble de données issues de 1380 mesures CD-SEM comprenant les motifs exposés en fonction de la dose, les motifs obtenus pour les conditions nominales, les motifs spécifiques à l'évaluation de fenêtrage du procédé et les motifs pour évaluer le contour obtenu dans la résine. Au total, cet ensemble contient 164 structures différentes présentes sur le layout qui ont été exposées à différentes conditions de dose et focus (FEM : Focus Exposure Matrix).

Les résultats de calibration sont montrés sur la figure 93 pour un ensemble de motifs complexes. La latitude en énergie est présentée en fonction de la profondeur de champ. Ces courbes EL-DoF sont couramment utilisées pour ce genre d'analyses. Les courbes de gauche permettent d'identifier la profondeur de champ commune à tous les motifs uDoF. Sur la figure de droite, cette courbe correspondant à la uDoF et issue du modèle 3D est comparée à celle obtenue avec le modèle 2D. Le paramètre uDoF est vraiment réduit avec une valeur de 78 nm, alors que le modèle 2D conduit à une uDoF de 95 nm. Cette différence de 20 nm n'est pas négligeable. Ces valeurs issues de la simulation ont été comparées à des mesures expérimentales, obtenues par analyse de PWQ. Les mesures de défektivité ainsi réalisées ont conduit à une valeur de 60 nm, qui est plus proche de celle obtenue avec le modèle 3D, ce qui permet de justifier l'utilisation de ce modèle. A noter que la valeur fournie par le PWQ est une valeur basée sur les inspections après gravure. Donc si on considère qu'il y a une perte après gravure qui s'ajoute aux effets provenant de la résine, la valeur obtenue par le modèle masque 3D nous semble pertinente.

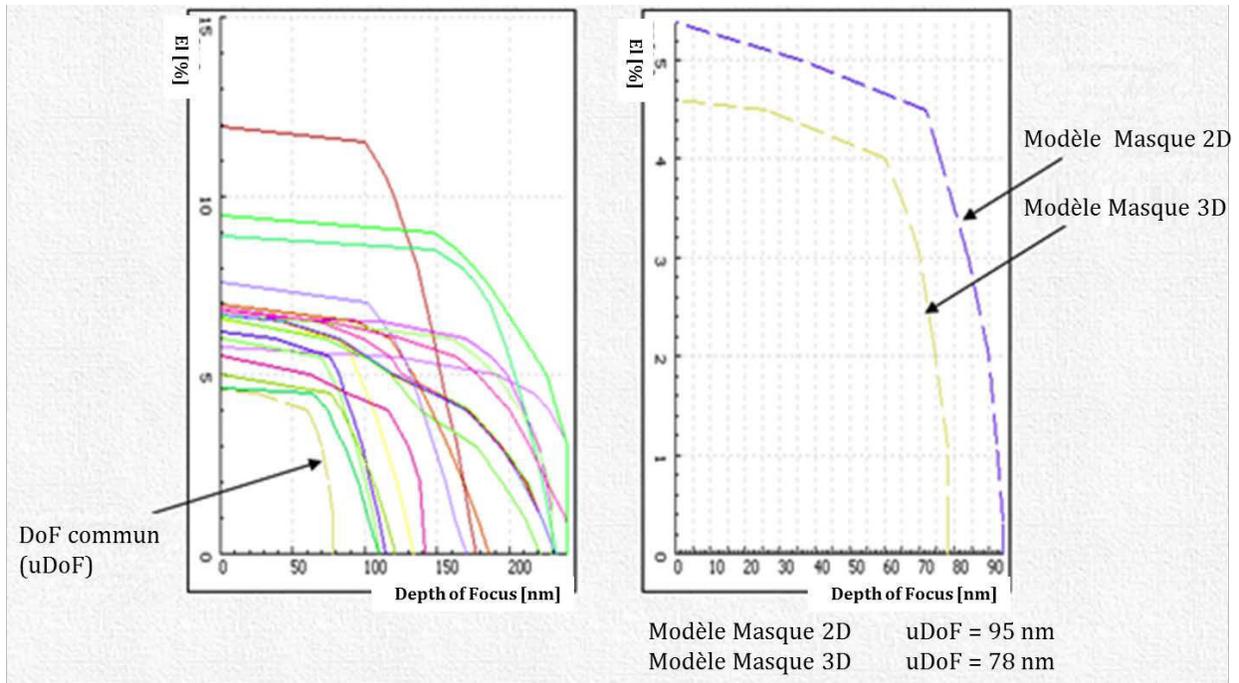


Figure 93 Courbe EL [%] – DoF [nm] pour un ensemble de motifs sélectionnés (motifs complexes 2D)

A partir des travaux effectués sur le modèle masque 3D, le modèle résine 3D est alors indispensable pour améliorer la capacité de la prédiction de notre modèle de détection.

6.3.1.1.2 Le modèle résine 3D

Actuellement, le grand challenge face aux solutions dans l'industrie pour les nœuds 28 nm et au-delà est le fait que le profil du motif de résine après la lithographie a un fort impact sur la forme du motif que l'on obtiendra après l'étape de gravure. Ceci est notamment lié à la perte significative de matière au sommet de la résine au cours du procédé, appelée « top loss ». Ce phénomène correspond aussi bien à une diminution globale de la hauteur de résine qu'à la formation de profil arrondi au sommet. Cette perte d'épaisseur peut avoir des conséquences très importantes lors des procédés gravures, certaines zones n'étant plus suffisamment protégées. Par ailleurs, un manque de contrôle du profil de résine et notamment de la pente sur les flancs (Side Wall Angle, SWA, en anglais) cause des effets indésirables sur les motifs gravés, ce qui se traduit par une déviation du CD (voir figure 94).

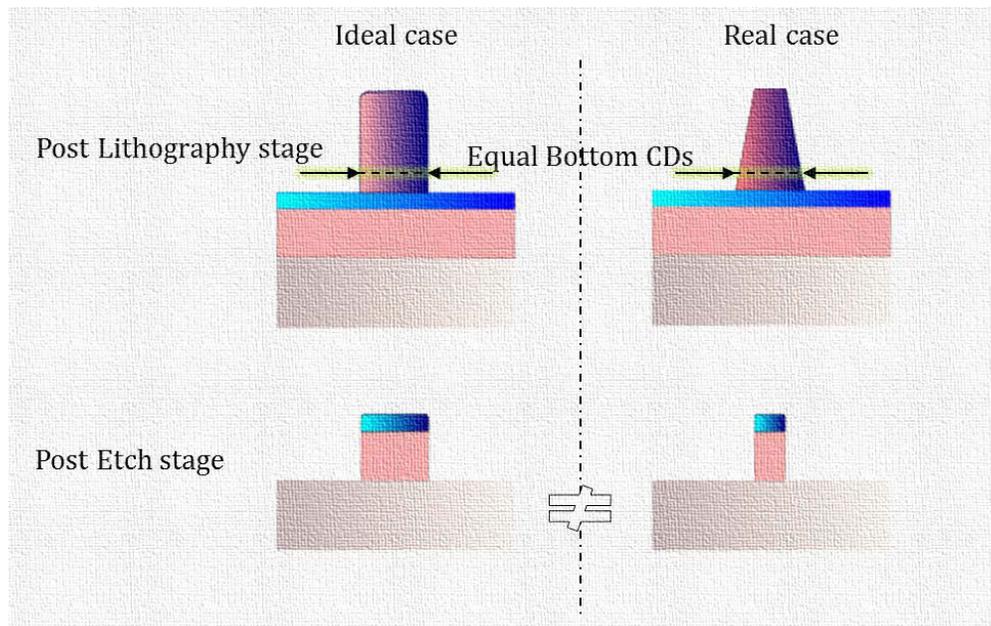


Figure 94 Illustration de l'importance d'avoir un modèle résine 3D précis et conséquences possibles sur les résultats obtenus après gravure

Par conséquent, il est nécessaire d'implémenter une information sur le profil de résine dans le modèle afin de pouvoir prédire de façon fiable les motifs critiques qui conduisent à une variation du procédé. L'objectif est donc la construction et la calibration d'un modèle résine 3D pour arriver à une capacité plus fiable de prédiction de ces motifs critiques.

Le modèle résine 3D utilisé est également celui de Brion Technologies. Ce modèle résine est utilisé pour prédire une modification du profil de la résine, ainsi qu'une perte d'épaisseur⁴⁹ qui apparaissent lors du développement. La construction de ce modèle est composée de deux étapes : une liée à l'optique et une à la résine. La partie optique –comme pour le modèle masque 3D – modélise le masque et le scanner et permet la prédiction de l'image aérienne du masque. La partie résine modélise l'interaction entre la résine photosensible et la lumière, les réactions chimiques dans la résine, le recuit thermique et le développement. Cependant les paramètres au niveau de la résine sont nombreux et le modèle physique est couplé avec un modèle empirique pour le développement (Liu, 2012) (Moulis, 2014) afin d'optimiser le temps de calcul.

Ce modèle résine 3D est calibré selon le processus de la figure 95. Les métrologies appliquées sont les suivantes : les mesures CD SEM pour le modèle masque 3D à partir des 1380 données;; les mesures AFM3D⁵⁰ pour plusieurs conditions (dose et focus) sur deux motifs denses avec un pas = 120 nm et 200 nm et un motif d'une ligne isolée ; les mesures X-SEM pour plusieurs conditions (dose et focus) sur un motif dense de pas = 90 nm.

⁴⁹ Ces effets causent très possiblement des défauts lors des gravures subséquentes

⁵⁰ AFM 3D : Atomic Force Microscope

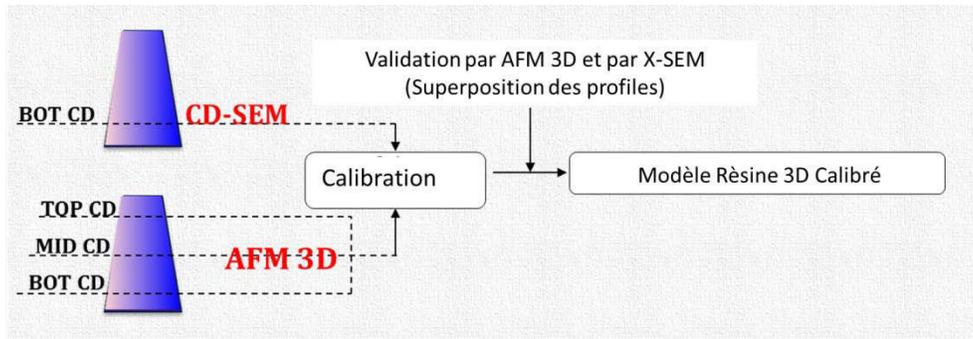


Figure 95 Le déroulement simplifié de la calibration du modèle résine 3D

Lors de la calibration de ce modèle résine, notre modèle masque 3D fournit le modèle de « Baseline ». Les mesures CD AFM3D sont utilisées pour compléter la calibration jointe. Afin d'optimiser au mieux le profil, les résultats issus du modèle sont comparés avec les données de l'AFM 3D sur les structures qui font partie de l'ensemble des motifs de calibration. La figure 96 (image de gauche) montre un exemple avec la superposition de contour du profil sur le motif de pas = 200 nm. Cette superposition est faite sur la mesure AFM3D et sur une image de X-SEM. Le contour de profil de résine simulé reproduit bien les données expérimentales. La comparaison entre le profil simulé et l'image X-SEM a été faite également sur le pas 90 nm. On voit ici également un bon accord entre le profil simulé et l'image X-SEM. Le modèle est capable de prédire la perte d'épaisseur au sommet de 30 nm.

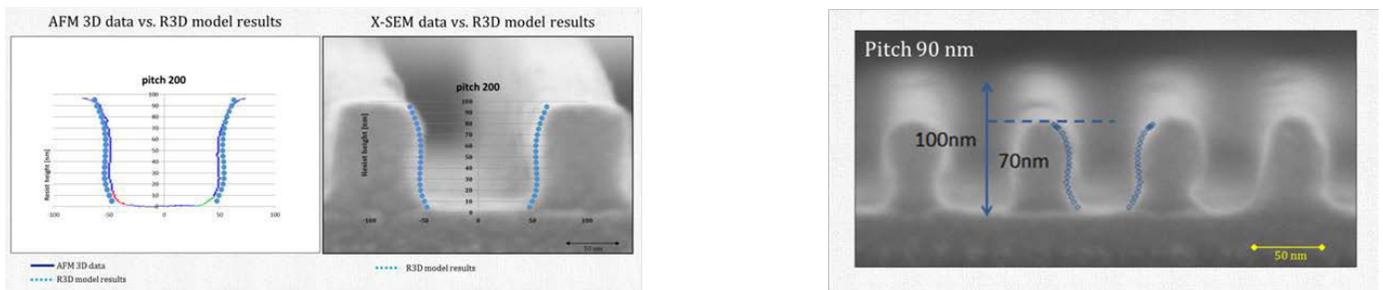


Figure 96 : comparaison des profils de la résine obtenus par simulation et métrologie : (à gauche) sur structure de pas = 200 nm; (à droite) sur structure de pas = 90 nm. A noter : chaque fois la ligne pointillée bleu représente le profil simulé

Cependant, en allant plus loin avec la validation, nous testons notre modèle résine 3D à l'aide de l'outil LMC sur une partie du design. Il est intéressant de noter qu'un des motifs critiques trouvé par cet essai est équivalent à un motif détecté par le PWQ. La figure 97 nous montre ce motif avec, successivement de gauche à droite : l'image CD SEM obtenue après l'étape lithographique à l'endroit correspondant à la zone de défaut potentiel, les contours simulés en bas de la résine à 20 % de la hauteur (Bottom) – au milieu à 50 % de la hauteur (Middle) – en haut à 80 % de la hauteur (Top), et l'image de CD SEM obtenue après gravure et CMP issue de l'analyse de PWQ. On voit que le défaut apparaissant après gravure n'est pas détectable sur l'image CD SEM obtenue après lithographie. Cette figure montre que les modèles standards qui prédisent simplement le CD Bottom ne permettent pas d'anticiper l'apparition de ce type de défaut.

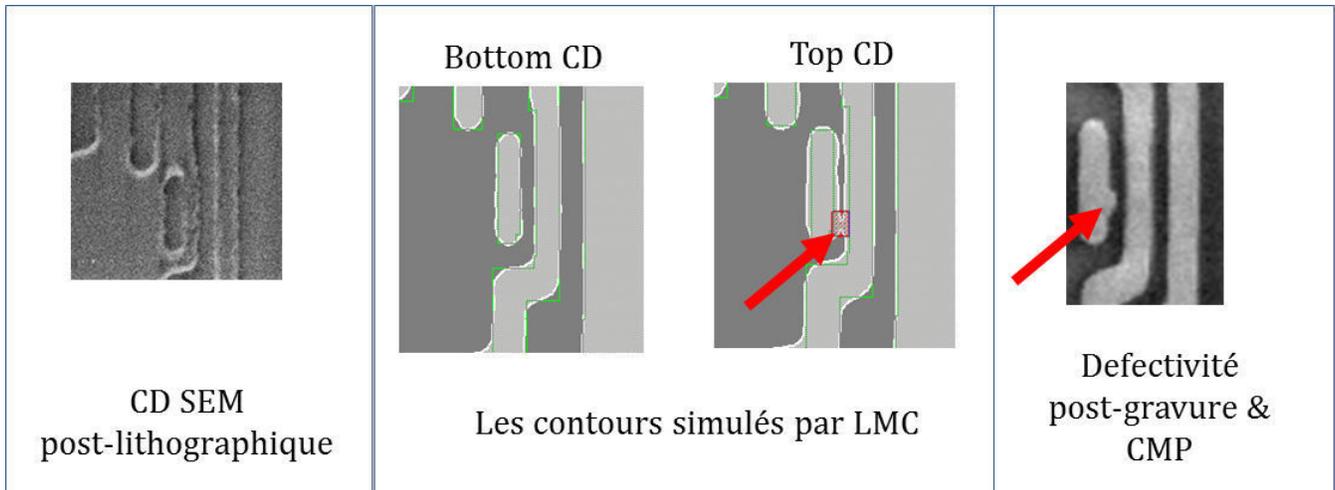


Figure 97 : exemple de motif critique montrant la performance du modèle résine 3D. Les images montrent (gauche à droite) : CD SEM après lithographie ; les contours simulés par en bas et en haut de la résine ; CD SEM après gravure & CMP fourni par PWQ.

A noter que sur les images de simulation, les parties en gris foncé correspondent à des motifs et en gris clair à des trous.

Ces résultats montrent que l'on peut conclure que le modèle résine 3D est capable de fournir un bon accord entre simulation et réalité sur plaque.

6.3.1.2 La méthodologie de détection

Une fois les modèles définis avec une bonne capacité de prédiction, l'étape suivante est de construire une méthodologie intelligente de détection. Comme précédemment, cette problématique est traitée en étudiant dans un premier temps les effets liés à la topographie de masque et dans un deuxième temps par les effets liés au profil de la résine.

6.3.1.2.1 La détection conduite par la topographie de masque

L'objectif est de détecter toutes les structures montrant une déviation de meilleur focus prononcée et de les analyser. L'idée principale est de construire un détecteur basé sur des simulations qui soit capable de repérer les motifs sujets à une grande différence de meilleur focus. La façon la plus simple pour le réaliser est de regarder et analyser la réponse du CD en fonction du focus, c'est à dire la pente de la courbe de Bossung, à l'aide de LMC (section 6.1) (Brion, Brion, 2015). A noter que ceci est limité par la taille du design ayant une grande variété des motifs.

L'analyse des courbes de Bossung est illustrée sur la figure 98, pour deux types de motifs : CD ligne et CD espace. Pour chacun, la courbe de Bossung idéale est indiquée (à gauche de la figure). La courbe de Bossung est bien centrée avec une déviation de meilleur focus inférieur à 10 nm. Tandis qu'à droite, nous montrons des courbes non-idéales possibles pour ces exemples : courbes avec forte déviation et courbes avec une allure non-régulière.

La pente est caractérisée par les paramètres s_1 et s_2 , qui sont déterminées de manière symétrique sur une plage de focus (f_0) et ($f_0 + \Delta f_0$). Afin de vérifier les cas indésirables, nous inspectons le produit de s_1 et s_2 . Cela signifie, si nous prenons un cas de CD ligne où nous considérons parabolique la forme de Bossung : le produit de s_1 et s_2 doit être négatif. Si ce n'est pas le cas, la courbe est soit déviée, soit possède une forme non-régulière. Remarque : nous pouvons toujours vérifier l'allure de Bossung par une analyse de fenêtre de procédé (PWA).

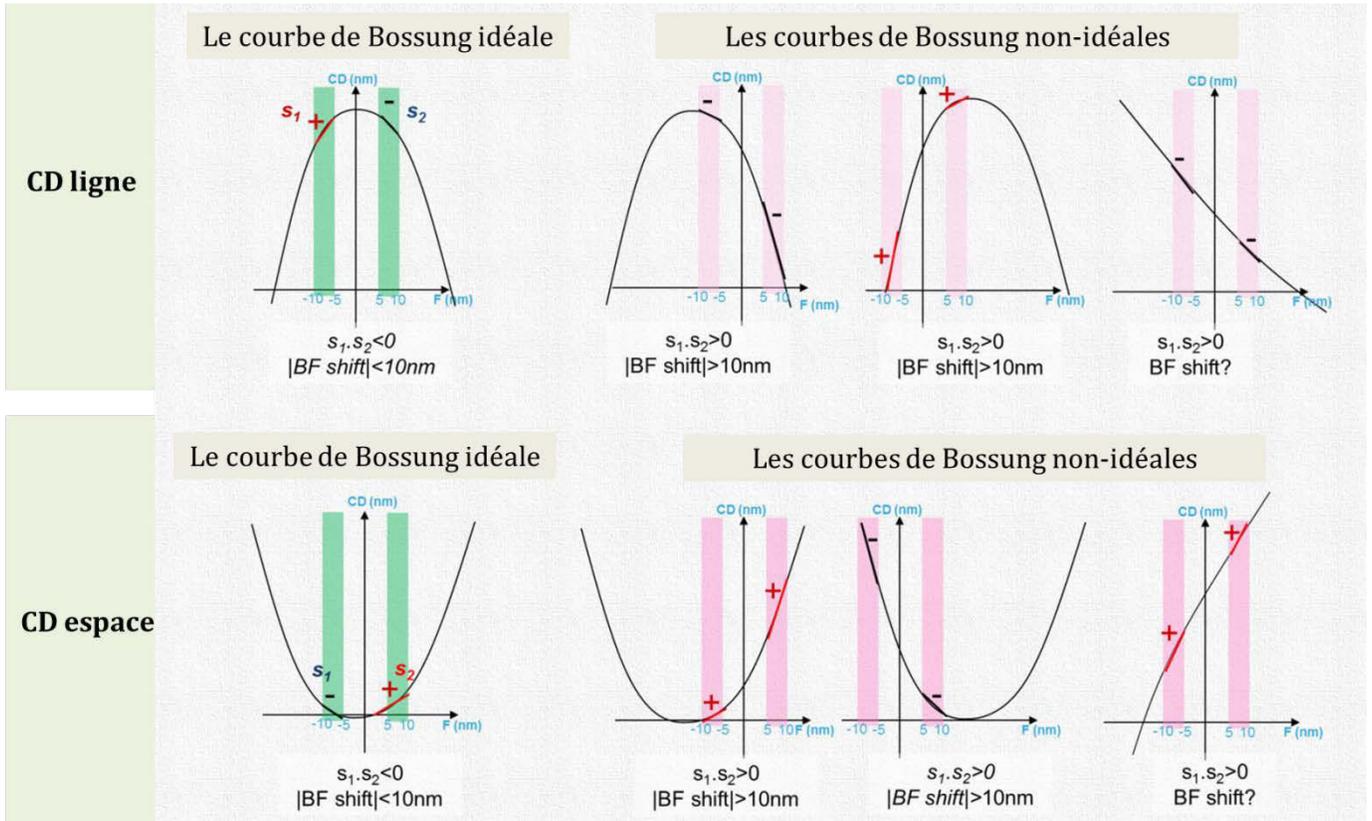


Figure 98: représentation schématique de l'analyse de la courbe de Bossung montrant les différentes courbes de Bossung possible sur deux exemples : CD ligne et CD espace.

Afin d'implémenter cette idée pour sélectionner les motifs potentiel ayant une forte déviation de meilleur focus dans une méthode de détection applicable, nous utilisons la capacité de l'outil LMC (section 6.1). Nous utilisons une comparaison contour-à-contour à l'aide de LMC (C2C LMC), vu que s_1 et s_2 sont déterminés de manière symétrique basés sur le delta CD pour des conditions de focus différents : $\pm (f_0)$ et $\pm (f_0 + \Delta f_0)$. A noter que la valeur de meilleur focus est également fournie.

Une fois ce détecteur bien défini, une analyse de fenêtre de procédé (PWA) est indispensable sur les structures détectées comme étant les plus critiques du point de vue de la déviation du meilleur focus. A l'aide de cette analyse nous arrivons également à vérifier si ces motifs limitent vraiment la fenêtre de procédé commune ou pas.

Afin de tester cette méthodologie, nous l'appliquons sur un layout contenant des milliers de motifs différents, représentatifs du niveau métal considéré. Nous trouvons une déviation de meilleur focus, parmi les motifs, de l'ordre de grandeur de 30 – 35 nm, ce qui est cohérent avec la valeur d'environ 27 nm présentée précédemment sur la figure 92. La figure 99 représente le delta de meilleur focus en fonction du nom des motifs, en particulier pour des motifs dont la variation de meilleur focus est proche des bords de la gamme considérée 5 – 35 nm.

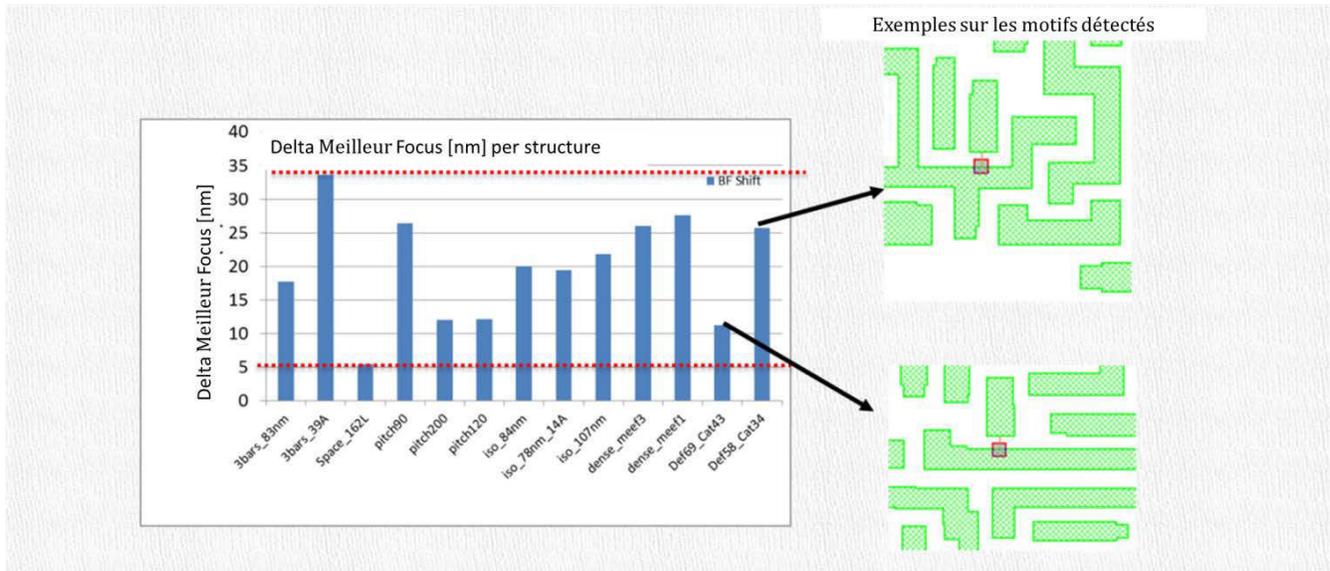


Figure 99 Représentation du décalage du meilleur focus en fonction des structures

Outre l'analyse sur la déviation de meilleur focus, nous faisons également une analyse de la fenêtre de procédé. On trouve que la profondeur de champ DoF la plus petite est de 60 nm, qui est une valeur très faible. Cette faible valeur de 60 nm est également la profondeur de champ commune (uDoF). Cela signifie que la déviation de meilleur focus a un impact négligeable (ou zéro) sur la variation de procédé. C'est le DoF individuel qui possède un impact important.

Nous avons également des mesures de défektivité provenant de PWQ dont les résultats sont cohérents avec cette valeur. Malgré la cohérence des deux analyses, les structures critiques détectées par notre détecteur ne sont pas équivalentes à celles de l'analyse PWQ. Par conséquent notre modèle de détection doit encore être affiné, ce qui constituera une des perspectives de ce travail.

Nous pouvons dire que ce détecteur, lié aux effets de la topographie du masque, fonctionne bien sur cet ensemble de motifs. Mais il y a quelques points faibles encore à affiner, comme développer notre méthode sur la mise au point de détecteur. Ce point se traduit par la problématique des motifs possédant une courbe de Bossung fortement inclinée ou plate. Donc logiquement notre méthode a sa faiblesse pour déterminer les valeurs précises de meilleur focus pour ces cas. Il faut également noter que, malgré le fait que ce niveau possède un large ensemble de motifs ayant un

meilleur focus décalé, nous avons trouvé par une analyse de fenêtre de procédé que les motifs limitant une PW commune sont les motifs présentant un problème de profil. Ceci montre que le profil de la résine devient plus marquant que celui du masque : comme nous l'avons montré dans notre analyse précédant où la déviation de meilleur focus n'apparaît plus limitant. La figure 100 montre un exemple de ces motifs qui est détecté par le PWQ.

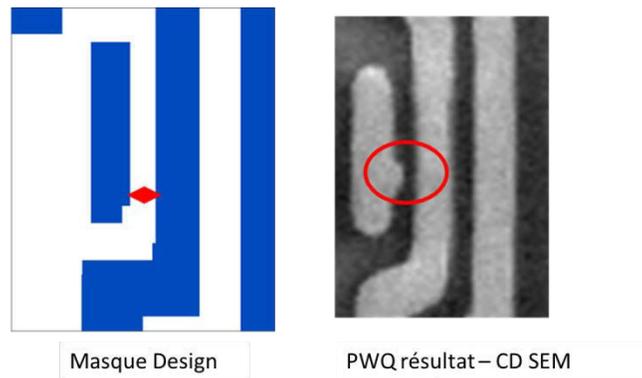


Figure 100 Un exemple lié à la problématique de profil de la résine dans le cas d'un défaut typique. (A gauche) le design du niveau (à droite) le CD SEM obtenu par PWQ, donc après gravure et CMP

Les effets liés au profil de la résine ayant une plus grande influence que ceux de la topographie du masque, nous allons maintenant considérer le profil de résine.

6.3.1.2.2 Détection conduite par les effets liés au profil de la résine

Nous avons utilisé deux approches pour la méthodologie de détection. Toutes les deux sont basées sur la capacité de l'outil de LMC en appliquant le modèle résine 3D. Les approches nommées sont les suivantes.

- *L'approche A* : comparaison et inspection contour-à-contour (C2C). Cela veut dire une comparaison de la différence de CD à différentes hauteurs de la résine, les différences les plus grandes correspondant aux motifs les plus critiques. A l'aide de l'outil LMC, le détecteur va repérer tous les motifs pour lesquels le delta CD est supérieur à une certaine tolérance, que nous avons jugé la plus pertinente et cohérente avec la production. Nous pouvons utiliser cette approche pour deux types de défauts, notamment le top loss et l'effondrement des lignes⁵¹. Dans le cas du « top loss », nous ne considérons que les structures pour lesquelles le CD top est inférieur au CD bottom.
- *L'approche B* : détection de pont⁵², le CD pont identifié correspondant au plus petit espace détecté entre les lignes rouges. Les valeurs de CD pont les plus faibles correspondent aux motifs les plus critiques. Cette analyse se fait de la même manière qu'une inspection LMC tout en utilisant le modèle calibré de résine R3D. Tandis qu'un LMC standard n'est

⁵¹ Line collapse est le terme utilisé en anglais

⁵² Pont : le terme utilisé en anglais est : Bridge, où il y a un pont entre deux lignes de résine

applicable qu'en bas de la résine, en l'étendant avec notre modèle R3D sa capacité couvre une analyse à plusieurs hauteurs dans la résine. Ce détecteur repère les motifs pour lesquels la valeur de CD de « pont » est hors tolérance (figure 101). Un tel exemple de défaut est montré sur la figure 102. La limitation de cette approche est que l'analyse est faite à une hauteur donnée.

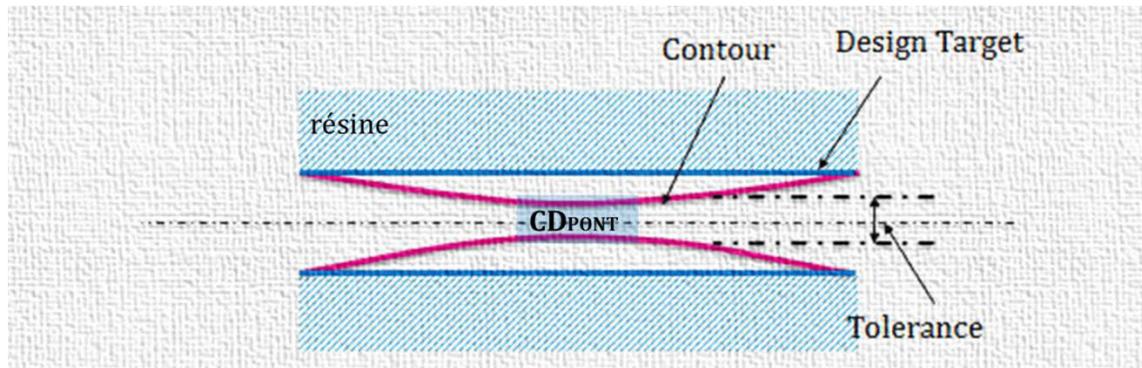


Figure 101 Illustration de « pont » en montrant le design souhaité et la plage du contour hors tolérance (la surface bleue correspond ici à la résine)



Figure 102 Un exemple sur un défaut de « pont », CD SEM après l'étape de la lithographie

Dans les deux approches, nous faisons un essai préliminaire sur une partie de la puce entière afin de disposer de détecteurs aux tolérances bien définies. Etant donné que le travail de validation est volumineux, nous ne montrons que les résultats de l'approche A. Nous avons vu que l'approche B fournit les mêmes défauts les plus critiques donc donne une analyse de PW équivalente. Nous montrons une comparaison des deux méthodes, A et B, en se focalisant sur le compte de défauts détectés.

Les résultats de l'approche A sont analysés avec les mesures de fenêtre de procédé PW. Le détecteur « contour-à-contour » détecte par principe les motifs les plus critiques, et nous déterminons la fenêtre de procédé commune à tous ces motifs à partir de leurs courbes de Bossung. Nous déterminons également la DoF commune pour ces motifs à partir de leur courbe d'EL-DoF. La figure 103 et la figure 104 représentent les résultats respectivement à 20 % (CD BOT) et 80 % (CD TOP) de la hauteur de la résine. A partir de ces courbes, nous avons extrait la

PW commune en prenant la surface commune maximale des Bossung de chaque motif. Tandis que l'extraction de uDoF est possible à partir des courbes EL-DoF à une énergie latitude choisie ou bien à partir de PW commune (l'axe principal d'ellipse).

Nos résultats montrent que tandis que la fenêtre de procédé commune pour le CD BOT reste assez large avec une uDoF de 70 nm, le TOP CD ne fait pas apparaître de fenêtre de procédé commune à tous les motifs critiques, donc pas de uDoF.

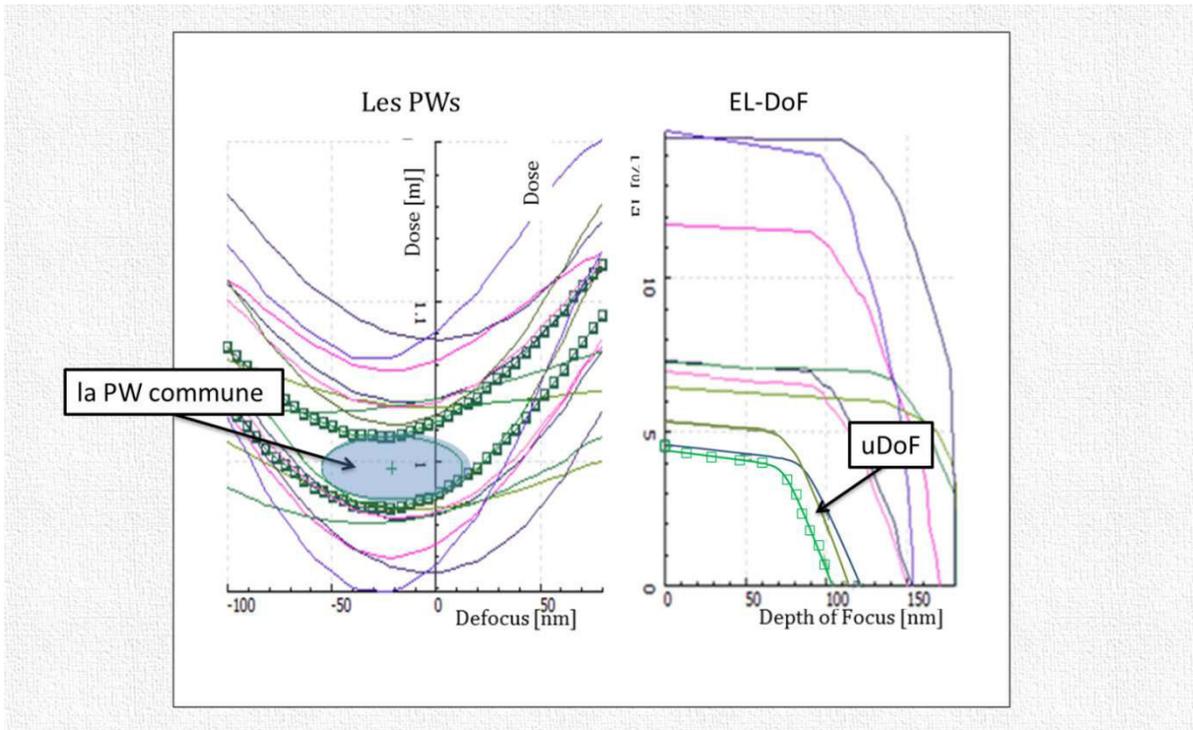


Figure 103 Les résultats d'analyse à 20% de la hauteur dans la résine : (à gauche) la PW commune (à droite) la DoF commune

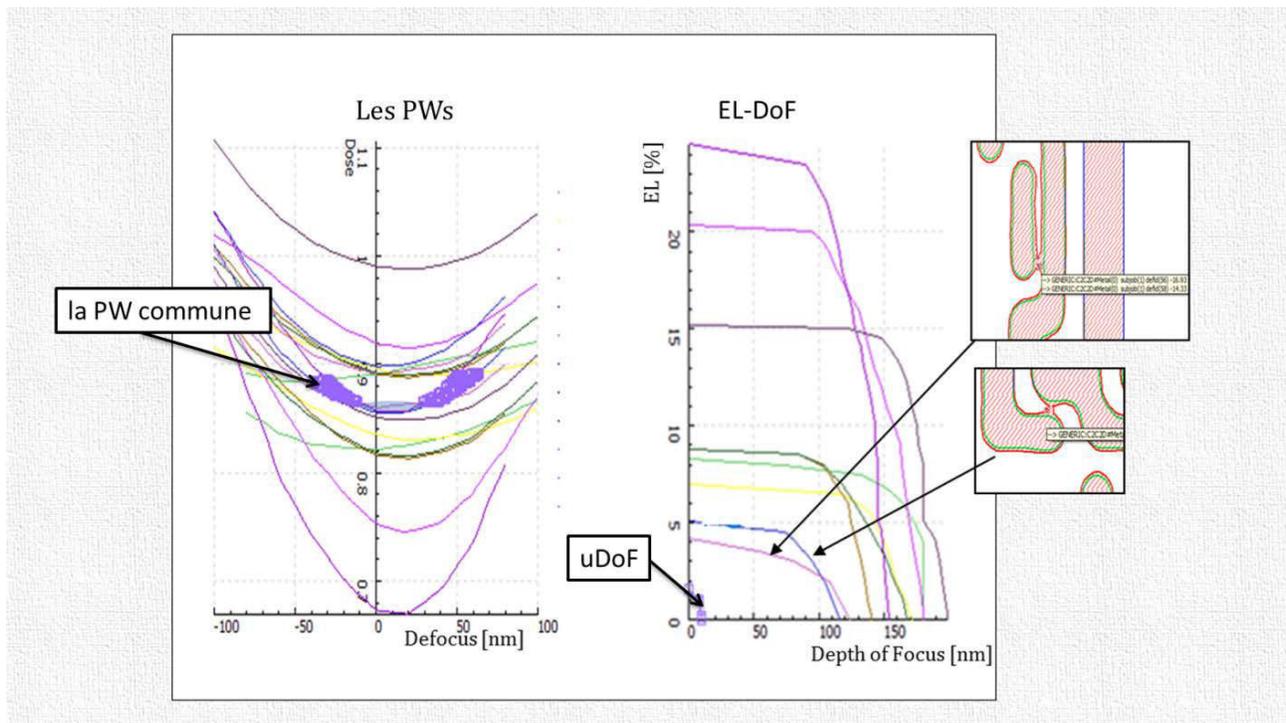


Figure 10' Les résultats d'analyse à 80% de la hauteur dans la résine : (à gauche) la PW commune (à droite) la DoF commune en montrant les deux motifs limitant la fenêtre de procédé

Pour cette étude préliminaire les deux motifs limitant la fenêtre de procédé sont indiqués sur la figure 104. Cela indique bien que l'effet de « Top Loss » est très présent et qu'il est plus limitant que l'effet provenant de la topographie du masque. En outre, ces défauts sont déjà présents dans les conditions nominales selon les résultats simulés. L'étape suivante est donc la validation de toute cette expérience.

Pour la validation on réalise des mesures CD SEM sur silicium après gravure, puisque certains défauts, dus à une trop grande perte de matière « top loss » pendant l'étape lithographique, ne sont visibles qu'après gravure. Deux plaques sont exposées : une à conditions nominales et l'autre en faisant varier la dose et le focus (plaque FEM). La figure 105 montre les résultats obtenus pour le motif critique déjà présenté sur la figure 104, dans les conditions de focus nominal et de dose variable. Le contour simulé montre bien une cohérence avec les mesures SEM après gravure, même si le défaut apparaît dans des conditions de dose proches du bord de la fenêtre de procédé. Sur cette figure nous montrons également des résultats de mesure d'une plaque aux conditions nominales pour laquelle nous observons une déviation de CD augmentée (chiffre mentionné en dessous de chaque image) : STDEV⁵³ est égale à 5.7 nm. A noter que pour le motif de contrôle de ce niveau de métal, cette valeur de STDEV est égale à 2.1 nm. Par conséquent nous pouvons dire que les motifs détectés comme limitant la fenêtre de procédé PW nécessitent une attention plus approfondie. De même en analysant la déviation de CD, les défauts de Top Loss (donc effet profil de la résine) sont une sorte de phénomène aléatoire conduisant à des défauts éventuels plus ou

⁵³ STDEV = l'écart-type : est une façon de mesurer la déviation (dispersion) d'un ensemble des données/ des valeurs parmi la valeur moyenne.

moins systématiques. A noter que ce motif est aussi détecté par l'analyse PWQ, ce qui signifie qu'il s'agit bien d'un défaut critique.

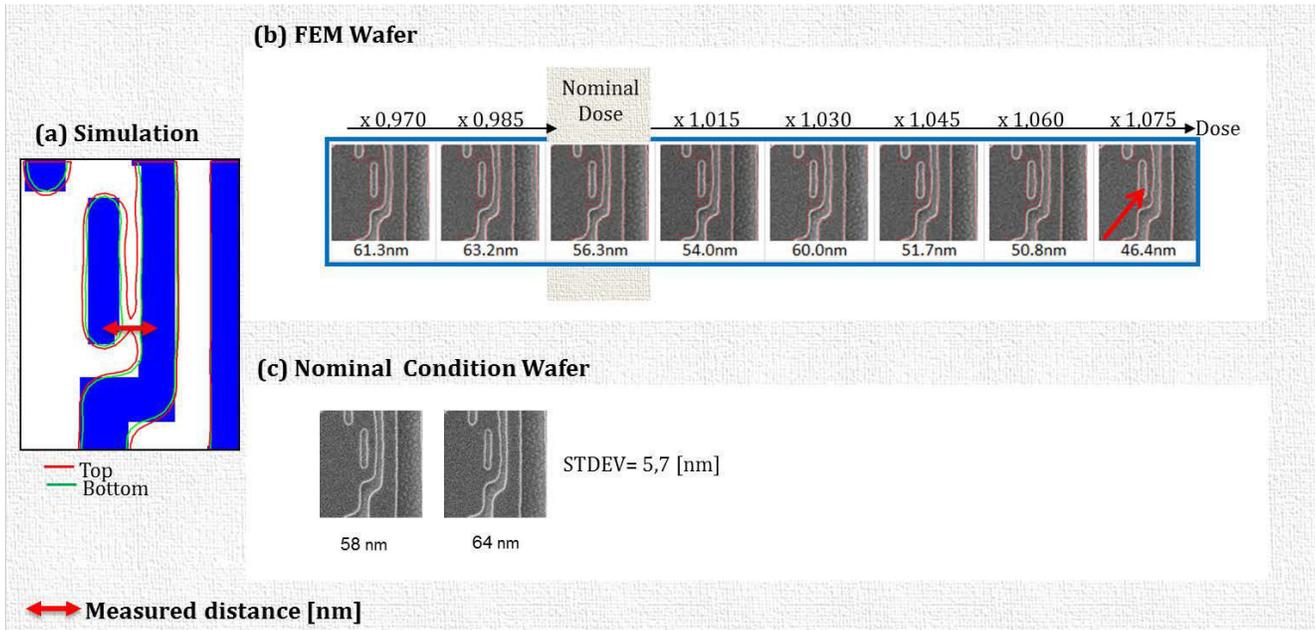


Figure 105 Résultats de validation sur plaque (a) les contours simulés par notre modèle résine 3D (b) les CD SEM obtenus après gravure dans les conditions de focus nominal et de dose variable (c) Deux images CD SEM provenant de la plaque exposée entièrement aux conditions nominales

Dans une deuxième étape, l'approche A de contour-à-contour est appliquée sur la puce entière. Nous identifions un nouvel ensemble de motifs critiques. La figure 105 montre ces motifs en indiquant les endroits critiques sur les résultats simulés. D'ailleurs il est intéressant de noter que ces trois exemples ont un design proche de celui du motif détecté par notre détecteur et par PWQ de la figure 105.

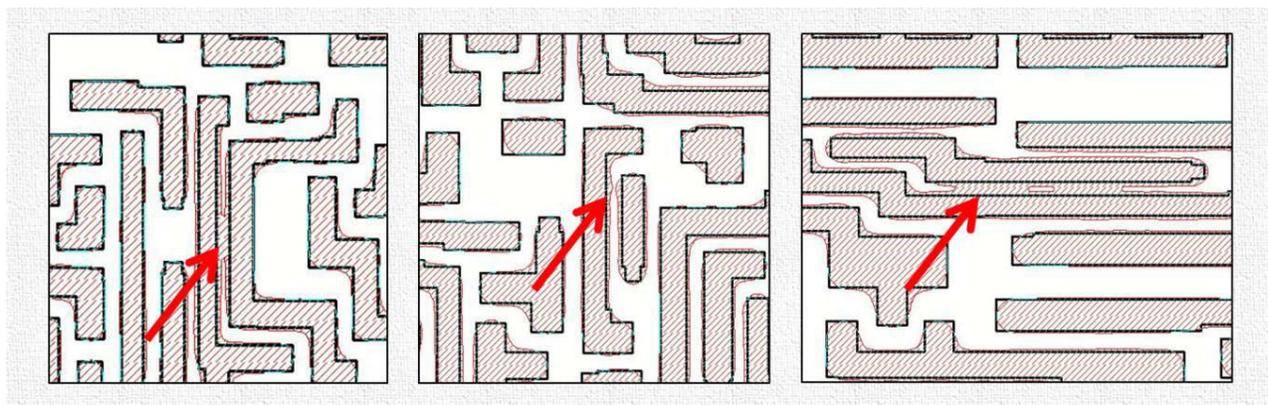


Figure 106 Le nouvel ensemble des motifs critiques par application de notre détecteur indiquant les endroits d'échec

A la suite des résultats simulés, une validation sur plaque est réalisée. Pour cette validation, les mêmes plaques sont utilisées. La figure 107 montre les résultats sur deux exemples de motif de type effondrement de ligne (« line-collapse ») et bout de ligne (« line-end »).

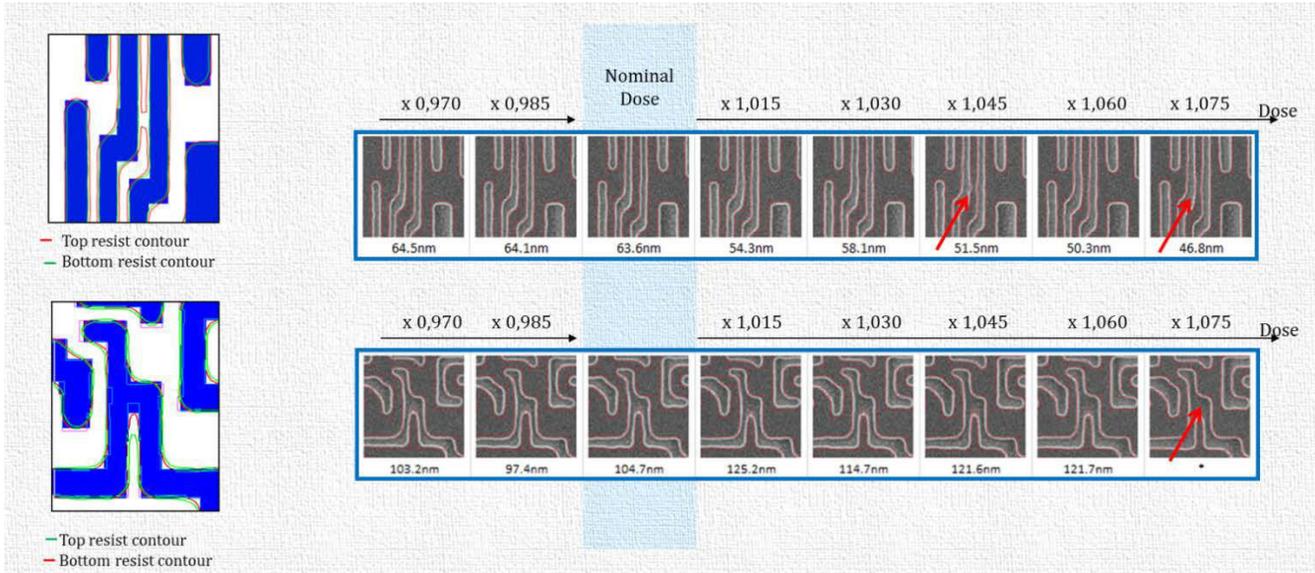


Figure 107 Deux exemples des motifs du nouvel ensemble, qui sont les plus représentatifs de l'effet lié au profil de la résine. CD SEM de plaque FEM après gravure en indiquant les endroits critiques (flèches rouges)

Ces motifs représentent et valident bien la problématique liée au profil de la résine. De la même manière les contours simulés sont cohérents avec les observations CD SEM après gravure. Par conséquent les motifs détectés sont en effet critiques, ce qui montre le fort potentiel de ce détecteur. En outre les résultats sont en harmonie avec ceux de PWQ. Remarque : l'approche B nous a conduit à détecter les mêmes motifs comme étant les plus critiques.

Malgré tous ces résultats prometteurs, il faut cependant tenir compte d'un point faible de cette méthode : l'ensemble des motifs détectés contient approximativement une centaine de milliers de défauts sur la puce entière. Cette quantité de données devient difficile à traiter – surtout de savoir lesquels sont les plus critiques. Par ailleurs les motifs issus de l'analyse PWQ se trouvent au milieu de ce large ensemble.

Nous pouvons conclure que les concepts de ces deux méthodes de détection sont validés pour une application sur puce entière (full chip) avec une performance fiable. L'approche B est plus efficace en termes de nombres de défauts, ce qui peut conduire l'utilisateur à la choisir en premier au détriment de l'approche A.

6.3.2 La Diminution

Des outils de détection de défauts ayant été optimisés, nous allons nous intéresser à la solution à appliquer pour limiter l'impact par ces défauts et ainsi diminuer la variabilité du procédé. Nous allons, comme dans le cas de la détection, analyser les effets de topographie du masque et les effets de profil de résine.

6.3.2.1 Les effets liés au masque 3D

L'étude des effets provenant du masque n'est pas nouvelle, ni dans la littérature, ni au niveau industriel. A l'heure actuelle ces effets sont déjà bien étudiés, ceci vient du fait que les CD dimensions des motifs sont devenus comparables avec l'épaisseur du masque pour ces niveaux. Ces effets masque 3D dépendent du CD des motifs, de l'épaisseur de l'absorbeur du masque, du type de masque etc...

Dans la littérature, nous trouvons, que l'épaisseur de l'absorbeur est la responsable de la présence de déviations prononcées du meilleur focus parmi les différentes structures et des aberrations sphériques d'ordre important (Staals, 2011) (Finders & Hollink, 2011) (Finders, 2014). Il existe donc déjà une grande variété de solutions pour diminuer et simuler ces effets.

Au cours de cette thèse, nous avons étudié la diminution de ces effets liée au meilleur focus par l'optimisation de front d'onde à l'aide de l'outil FlexWave. L'outil de FlexWave est un manipulateur des lentilles implémenté dans l'optique de projection. Cela est dû au fait que les effets de topologie du masque causent une différence de phase des ordres diffractés en raison de l'asymétrie entre eux. Cela implique, comme nous avons vu en section 6.3.1.1.1, une déviation du meilleur focus parmi les motifs. Dans la littérature, il est démontré que cette différence de phase est un phénomène correspondant à celui induit par les aberrations de front d'onde dans l'optique de projection. En outre la différence de phase liée à la déviation du meilleur focus, donc ces effets topologiques, peut être compensée par l'optimisation du front d'onde.

6.3.2.1.1 Résultats de l'optimisation du front d'onde

Pour accomplir notre objectif de diminuer la déviation du meilleur focus afin d'augmenter la DoF commune, nous avons réalisé dans un premier temps des simulations à l'aide de la plate-forme Tachyon SMO-FW (Source Mask Optimization avec FlexWave). Par ces simulations nous obtenons une augmentation importante de la DoF commune. Afin de valider ces résultats prometteurs, nous réalisons des simulations rigoureuses (logiciel de simulation Technology® Panoramic) en y implémentant la carte de correction du front d'onde (« wavefront correction map » en anglais) fourni par notre simulation Tachyon FW. Les cartes de correction de front d'onde de FlexWave (ou plus précisément les coefficients de Zernike) se réfèrent aux données de sortie fournies par Tachyon SMO-FW. Ces cartes de correction de FlexWave consistent en des informations du front d'onde spécifiées avec 64 coefficients de Zernike qui sont introduits comme données d'entrée de la plateforme.

En parallèle, des mesures CD SEM sont faites pour vérifier les résultats au niveau du système de projection.

6.3.2.1.1.1 L'optimisation à l'aide de Tachyon SMO-FW

Les résultats simulés à l'aide de la plate-forme Tachyon FW sont prometteurs. Nous avons essayé deux sources lumineuses différentes, une source Annulaire⁵⁴ et une source C-Quad⁵⁵. A noter que nous avons réalisé quatre simulations différentes pour lesquelles nous avons généré quatre cartes de correction de front d'onde différentes :

- 1^{ière} simulation : source C-Quad sur les motifs de calibration 1D
- 2^{ième} simulation : source C-Quad sur les motifs de calibration 2D
- 3^{ième} simulation : source Annulaire sur les motifs de calibration 1D
- 4^{ième} simulation : source Annulaire sur les motifs de calibration 2D

Le Tableau 14 résume ces résultats sur la déviation de meilleur focus avant et après l'optimisation avec l'amélioration d'uDoF (DoF commune) en pourcentage pour chaque ensemble de simulation. Nous observons un recentrage du meilleur focus. Ce recentrage nous permet d'améliorer la DoF commune. L'optimisation réduit la déviation de meilleur focus, donc augmente le gain pour la DoF commune.

Mode d'Illumination	Type de motifs de calibration pour le "wavefront map"	Δ BF [nm] Avant SMO FW	Δ BF [nm] Après SMO FW	Pourcentage [%] d'uDoF
C-Quad	1D	71	26	44.5
	2D	12	1	9.9
Annulaire	1D	20.9	3.3	27.5
	2D	13.1	2.6	20

Tableau 14 Résultats simulés à l'aide de Tachyon SMO-FW sur les quatre ensembles de simulation en indiquant la déviation de meilleur focus avant et après l'optimisation et en indiquant le pourcentage d'amélioration de la DoF commune

6.3.2.1.1.2 Validation par simulation rigoureuse

A la suite des simulations pour valider la faisabilité de la diminution du delta meilleur focus par l'optimisation, nous utilisons les capacités du logiciel de Technology® Panoramic pour les calculs lithographiques rigoureux. Les mêmes ensembles de simulations sont utilisés avec les mêmes conditions qu'avec Tachyon SMO-FW. Nous implémentons dans Panoramic les cartes correction de front d'onde générées récemment comme données d'entrées. Les résultats obtenus avant et après l'optimisation du front d'onde sont indiqués dans le Tableau 15 . Le terme « avant et après »

⁵⁴ Source Annulaire : ~ est un type de source lumineuse hors-axes pour laquelle la forme est un anneau. (Mack C. A., 2006)

⁵⁵ Source C-Quad : ~ est un type de source lumineuse hors-axes pour laquelle la forme est symétrique selon axe x et y. Il a quatre pôles. (Alleaume, 2014)

signifie dans ce cas, que la carte de correction a été appliquée ou pas. Ces valeurs de déviation de meilleur focus montrent une tendance cohérente avec les résultats montrés dans le Tableau 14, dans la mesure où les ΔBF (nm) les plus faibles et les plus élevés sont obtenus dans les même cas. En conséquence les résultats de Panoramic ont validé nos résultats de Tachyon SMO-FW.

Mode d'illumination	Type de motifs de calibration pour le "wavefront map »	ΔBF [nm] sans l'optimisation	ΔBF [nm] avec l'optimisation
C-Quad	1D	55	35
	2D	NA*	NA*
Annulaire	1D	13.71	5.20
	2D	20	5

Tableau 15 Résultats de simulations rigoureuses à l'aide du logiciel de Technology® Panoramic indiquant la dispersion de meilleur focus parmi les motifs de calibration. « NA » signifie qu'il y a un manque de données.

6.3.2.1.1.3 Validation par CD SEM

Pour la validation sur silicium par CD SEM, les plaques sont exposées avec des doses et des focus différents (plaque FEM). Un système de projection avec l'option de FlexWave est appliqué. Les cartes de correction de front d'onde sont directement introduites au niveau du scanner. Par cette évaluation sur plaque nous réalisons :

- (1) Une évaluation sur les motifs de calibration pour valider la fiabilité de notre simulation
- (2) Une évaluation sur les motifs simples pour valider le fait que l'optimisation du front d'onde n'implique aucun dégât sur les autres structures
- (3) Une évaluation sur quelques motifs critiques pour vérifier l'impact de l'optimisation du front d'onde sur les structures critiques

A noter que les courbes de Bossung sont issues d'analyses de CD SEM et donc de mesures de fenêtre de procédé.

(1) Evaluation sur les motifs de calibration

Cette évaluation nous a servi de point de départ pour la validation. Le tableau 16 montre les résultats de la déviation du meilleur focus avant et après l'optimisation en indiquant le pourcentage d'amélioration de la DoF commune, pour des conditions différentes de source et motif. Par ces résultats on voit que la déviation du meilleur focus diminue bien grâce à l'optimisation. On constate une diminution de la déviation du meilleur focus, donc un recentrage qui implique les changements suivants en termes d'uDoF. Ces résultats montrent une bonne corrélation avec ceux de Tachyon SMO-FW, notre simulation initiale.

Mode d'Illumination	Type de motifs de calibration pour le "wavefront map	Δ BF [nm] sans FW	Δ BF [nm] avec FW	Pourcentage [%] d'uDoF
C-Quad	1D	28.9	10.1	83.3
	2D	16.6	3.5	50
Annulaire	1D	27.7	6.7	50
	2D	14.3	11.1	37.5

Tableau 16 résultats de CD SEM sur les motifs de calibration indiquant la déviation du meilleur focus sans et avec l'optimisation par FlexWave ; en indiquant le pourcentage d'amélioration de la DoF commune

(2) Evaluation sur les motifs simples, non utilisés pour la calibration

L'objectif de cette campagne de mesures CD SEM est d'analyser l'impact de l'optimisation sur les motifs supplémentaires qui n'ont pas servi pour optimiser le front d'onde. Par conséquent, l'objectif principal est d'explorer la validité de notre optimisation. En parallèle nous cherchons à déterminer si une optimisation sur les motifs 1D est suffisante, afin d'éviter celle sur les motifs 2D, vu que la criticité ou complexité des motifs 2D est plus prononcée.

Cette analyse est faite à partir d'un ensemble de 20 motifs 1D et 2D. Nous pouvons voir la même tendance que celle obtenue avec les résultats simulés, voir tableau 17. Cela veut dire que les meilleurs focus sont recentrés en diminuant la déviation totale. L'optimisation du front d'onde en utilisant FlexWave permet de réduire la déviation du meilleur focus. Cette diminution se traduit par un recentrage et donc par une augmentation de la DoF commune

Mode d'Illumination	Type de motifs de calibration pour le "wavefront map	Δ BF [nm] sans FW	Δ BF [nm] avec FW	Pourcentage [%] d' uDoF
C-Quad	1D	47.1	26	37.5
	2D	47.1	30.2	37.5
Annulaire	1D	41.6	12.7	100
	2D	41.6	31.5	83.3

Tableau 17 CD SEM résultats sur les motifs non-calibration en indiquant le delta meilleur focus sans et avec l'optimisation par FlexWave; en indiquant le pourcentage d'amélioration de la DoF commune

Notre objectif de valider l'impact positif de l'optimisation sur les motifs supplémentaires est atteint. Nous pouvons également dire que l'optimisation sur les motifs 1D est suffisante, ce qui permet d'éviter un travail plus complexe sur les motifs 2D.

(3) Une évaluation sur quelques motifs critiques

L'objectif de cette campagne de mesures CD SEM est identique à celle d'avant : analyser l'impact de l'optimisation sur les motifs n'ayant pas servi à la calibration, dans ce cas les motifs critiques détectés par LMC. Il faut noter que l'évaluation sur ces motifs n'est pas si simple puisque ces motifs complexes sont difficiles à traiter. Cela veut dire que l'extraction de la déviation du meilleur focus parmi les motifs et de la DoF commune est réalisée manuellement.

La figure 108 montre un exemple représentatif de notre campagne de mesures, où on voit bien le recentrage des meilleurs focus sur les deux structures choisies – indiqué par les flèches (les rouges et les vertes signifiant respectivement avec et sans carte de correction de FlexWave). La DoF commune pour toutes les mesures faites est comme suit, voir Tableau 18.

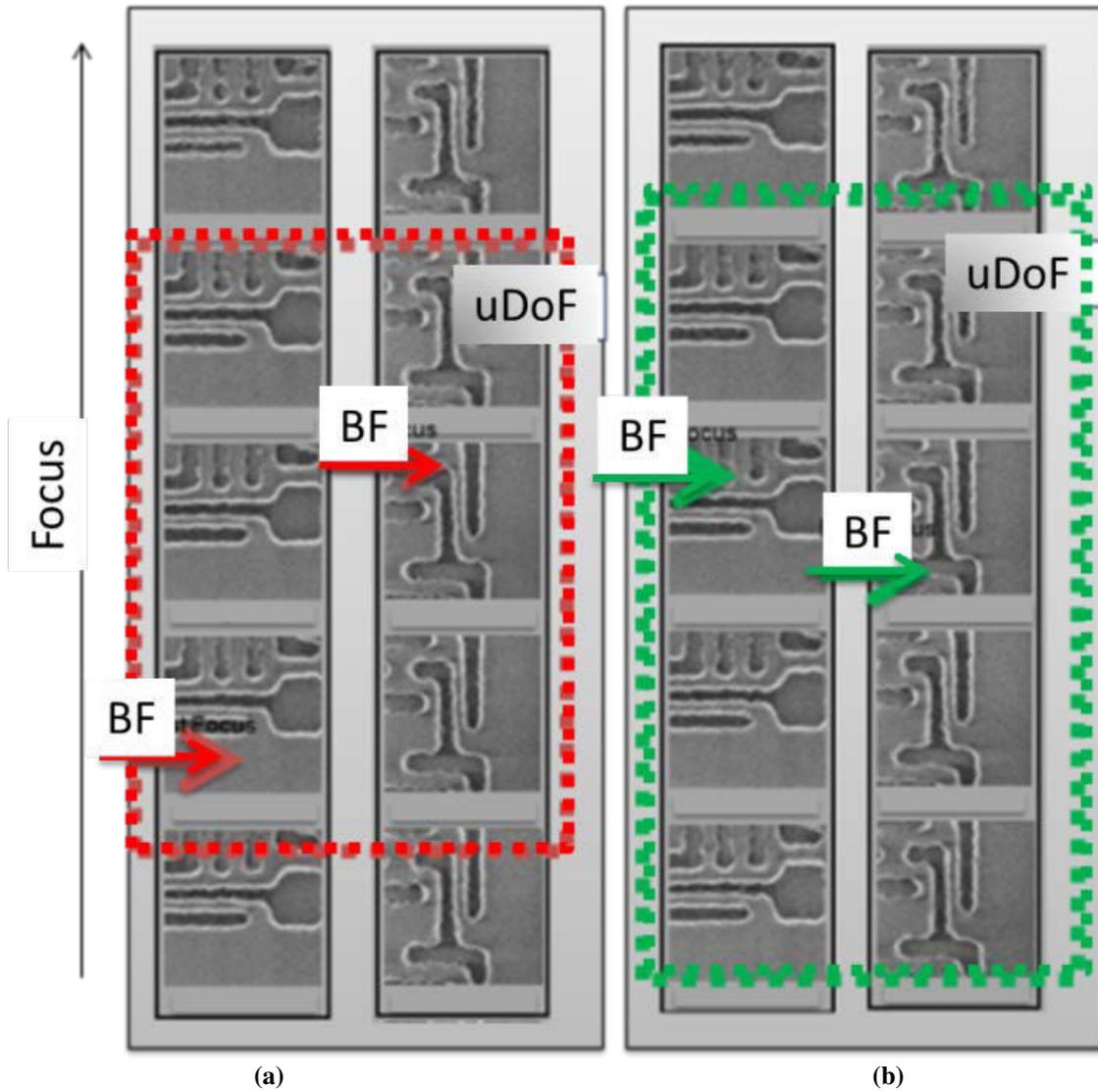


Figure 108: exemple de l'optimisation du front d'onde (mode d'illumination C-Quad) sur deux motifs critiques en indiquant la plage de DoF commune (uDoF) avant (a) et après (b) FW. Les flèches indiquent la position de meilleur focus.

Mode illumination	Type de motifs de calibration pour le "wavefront map"	ΔBF [nm] sans FW	ΔBF [nm] avec FW	Pourcentage [%] d'uDoF
C-Quad	1D	30	10	50
	2D	30	10	50
Annulaire	1D	60	30	33.3
	2D	60	20	50

Tableau 18 CD SEM résultats sur les motifs non-calibration en indiquant le delta meilleur focus sans et avec l'optimisation par FlexWave; en indiquant le pourcentage d'amélioration de la DoF commune

A partir de ces résultats, nous pouvons conclure que l'évaluation sur les motifs critiques est bien corrélée avec les données simulées par Tachyon SMO-FW. Ces résultats sont satisfaisants en montrant que par l'optimisation du front d'onde on arrive à compenser des effets topographiques. A noter qu'un point à améliorer dans l'optimisation est de développer une méthode pour la sélection des motifs de calibration les plus pertinents pour le nœud et le niveau en question.

A la suite de cette partie nous explorons les solutions pour diminuer les effets liés au profil de la résine.

6.3.2.1 Les effets liés à la résine 3D

L'objectif est de proposer des améliorations possibles pour réduire les problèmes de profil de résine non vertical. Par exemple, si le modèle prédit des flancs présentant une oscillation, on pourra essayer d'optimiser l'empilement. Si le contraste prédit par le modèle n'est pas constant sur toute l'épaisseur de la résine, il sera possible de l'améliorer par optimisation de la source (SO) (en utilisant le modèle R3D pour l'optimisation). Le nœud technologique considéré dans cette thèse fait que les problèmes de profil de résine sont à considérer systématiquement avec une optimisation de la source.

6.3.2.1.1 L'optimisation de la source

Il faut savoir que l'optimisation de la source est généralement destinée à identifier une source optimale pour un ensemble de motifs (Alleaume, 2014). Cela signifie qu'une source optimisée supportera une gamme de CD. Si ces motifs ne sont plus des motifs 1D, la détermination (le calcul) de la meilleure configuration de la source par optimisation devient très complexe. Après optimisation de la source, il est nécessaire de vérifier que cette optimisation n'a pas un impact négatif sur la DoF commune. Une validation de la source optimisée est obligatoire, car elle peut impliquer des gros changements sur les motifs imprimés dans la résine. Cette validation peut se caractériser par une évaluation de PV⁵⁶; une évaluation sur la fenêtre de procédé (PW) en se

⁵⁶ La construction de PV band (Processus Variation en anglais) est une technique pour exprimer les variations des paramètres du procédé tels que la dose, la concentration, et le biais de masque. La largeur de PV band, elle-même, est adéquate comme une mesure quantitative pour juger la faisabilité de conception du masque. Ainsi, c'est souvent utilisé comme un bon support de l'optimisation (Zou, 2010) (Alleaume, 2014).

focalisant sur la profondeur de champ commune uDoF ; une évaluation sur le contraste dans la résine etc.

Dans cette optique, la figure 109 montre le déroulement appliqué dans notre cas. A noter que toute l'optimisation est basée sur les capacités de Tachyon SMO/SO⁵⁷. La source d'entrée est la source annulaire utilisée pour ce niveau de métal. Le modèle OPC de l'entrée est notre modèle résine 3D calibré. Un ensemble d'une trentaine de motifs de calibration est considéré. Il contient des motifs 1D, des motifs 2D, y compris celui de l'analyse PW. Pour des raisons de confidentialité de plus amples détails ne sont pas abordés ici.

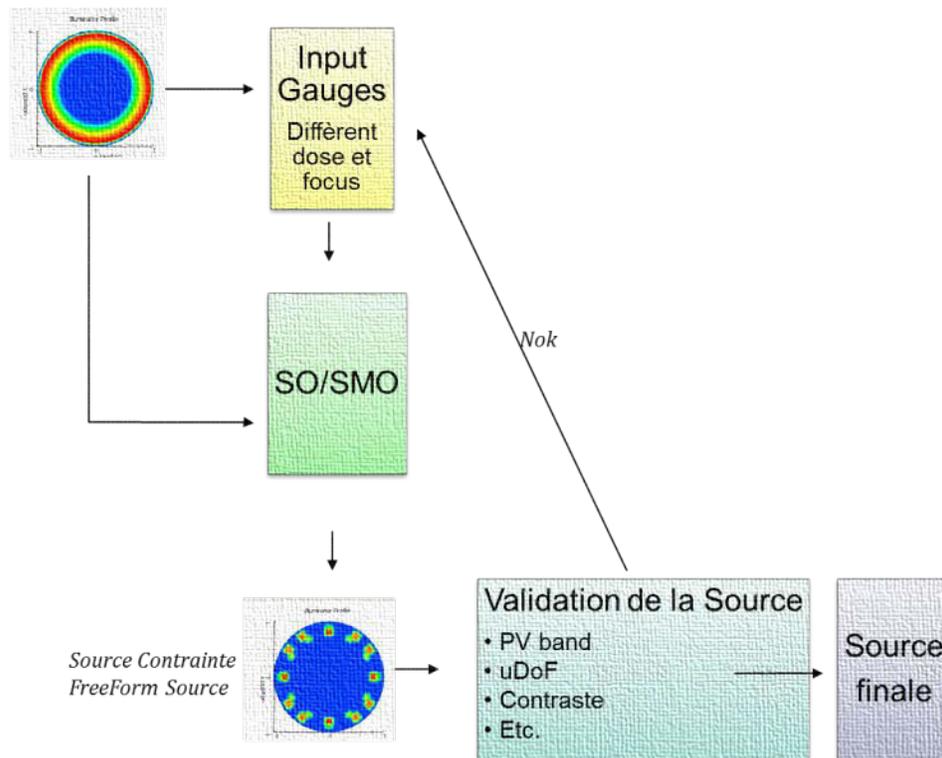


Figure 109 Le déroulement de l'optimisation de la source

Une validation de la source optimisée est faite, voir figure 110. Cette figure a également pour but de comparer la source initiale et la source optimisée. Nous avons considéré deux structures : des lignes denses avec un pas = 90 nm⁵⁸ (le trait rouge représente l'espace entre lignes) et un motif critique de PWQ (le trait rouge représente la ligne). Nous étudions le demi-CD à travers la hauteur dans la résine, et nous trouvons que la source optimisée, en effet, fournit un meilleur profil de la résine – plus proche de la dimension attendue :

⁵⁷ SMO = Source Mask Optimization ; SO = Source Optimization (only)

⁵⁸ La structure qui est utilisée pour raison de control en salle blanche pour ce métal layer, tant que ce pas = 90 nm et le pas le plus petit qui peut être imprimé par lithographie ; au-delà c'est les pas interdit par règle de dessin.

- Lignes denses :
 - o CD var = 17, 03 nm avec la source initiale ;
 - o CD var = 15,73 nm avec la source optimisée.
- Motif critique :
 - o CD var = 56, 69 nm avec la source initiale ;
 - o CD var = 53, 73 nm avec la source optimisée.

Il y a donc une amélioration de l'ordre de 2 nm pour les lignes denses. L'influence est plus importante pour le motif critique (droite), en particulier au niveau de la mise en évidence de la diminution du CD au sommet (top loss). Cette source optimisée doit encore être validée.

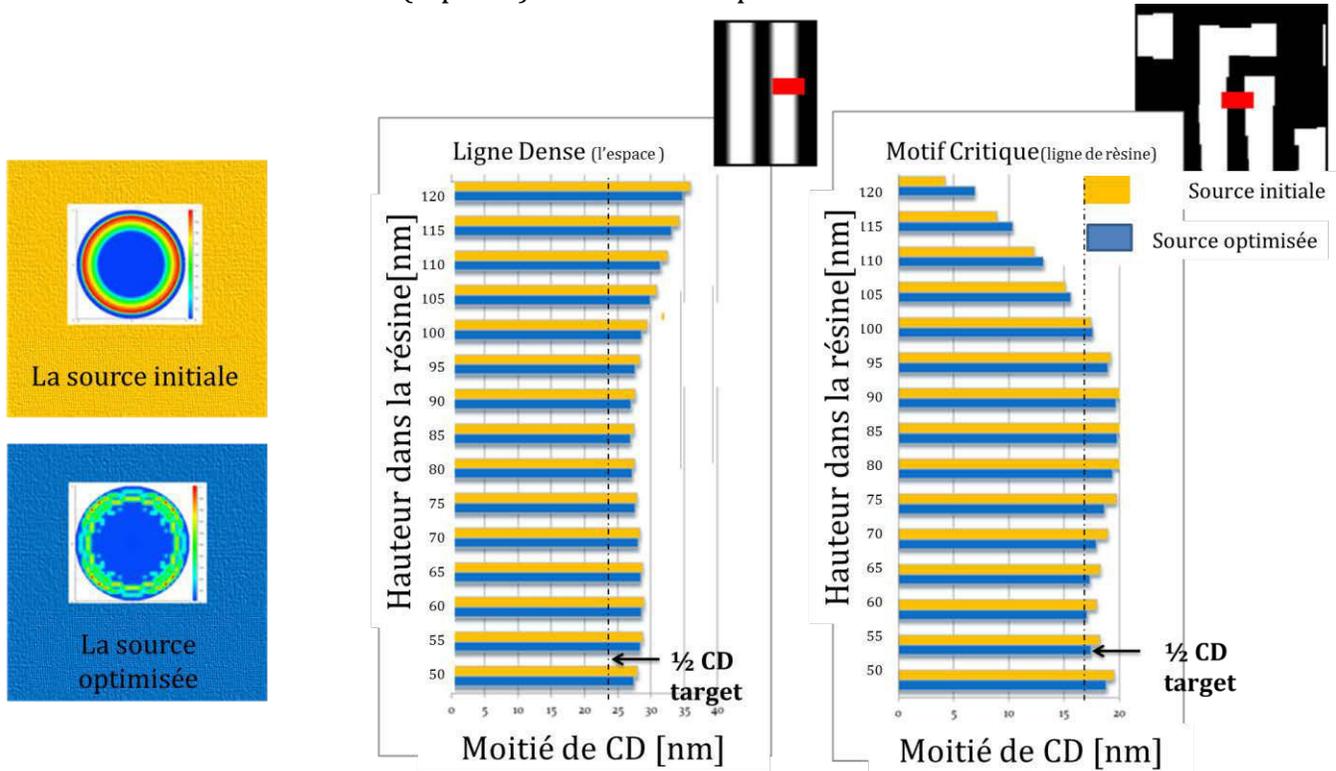


Figure 110 La source initiale vs. La source optimisée. Exemple sur une ligne de réseau dense et un motif critique (couleur jaune : source initiale et couleur bleu : source optimisée)

La validation de notre source optimisée intègre une étude de son impact sur la défektivité et sur le PV band. Afin de mieux répondre à cette question, nous analysons cette problématique par plusieurs approches :

- 1^{ière} méthode : cette méthode est basée sur le CD normalisé, ce qui veut dire le contour CD divisé par le CD nominal, sur différentes lignes de coupe (« cut line » en anglais) d'une analyse PV band. L'avantage principal de cette méthode est qu'elle fournit des informations sur la qualité de l'OPC, qui représente l'écart par rapport au CD envisagé). Egalement, elle fournit des informations sur le PV band, qui représente la variabilité de procédé (dose, focus, perturbation du masque). Par contre, étant donné qu'elle est basée sur les lignes de

coupe, c'est n'est pas vraiment applicable pour la puce entière. Elle est de plus limitée par la hauteur choisie dans la résine.

- 2^{ième} méthode : cette méthode est basée sur l'utilisation d'un LMC conventionnel à différentes hauteurs dans la résine pour les deux sources (initiale et optimisée) en même temps. La limitation principale est la hauteur de la résine (une hauteur choisie à la fois), mais l'avantage est qu'on peut l'utiliser sur une puce entière.
- 3^{ième} méthode : cette méthode est basée sur un scan de la puce entière qui détecte les motifs critiques ayant une variation importante de CD sur la hauteur de la résine, voir figure 111. Cela veut dire qu'un motif ayant une grande variation de CD va avoir un profil très perturbé alors qu'une petite variation de CD signifie que le profil est plus vertical, ce qui est le cas souhaité. L'avantage principal de cette méthode est qu'elle est applicable sur puce entière à travers la hauteur de la résine.

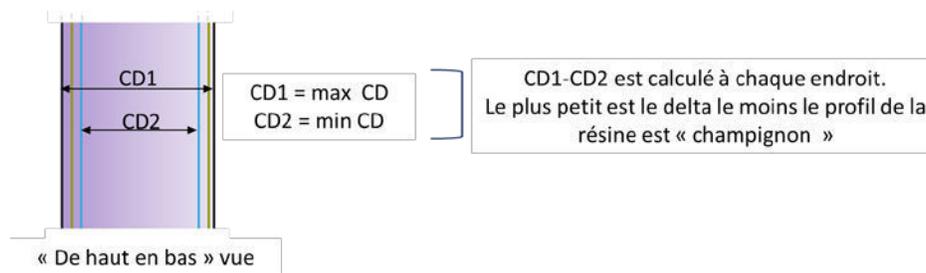


Figure 111 Illustration de la 3^{ième} méthode qui détecte les motifs ayant une variation de CD prononcée à travers la hauteur de la résine

(1) Méthode basée sur le CD normalisé à l'aide de l'analyse PV band

Les résultats de la 1^{ière} méthode sont montrés sur les Figures 112 et 113 à l'aide d'une analyse de PV band sur les motifs 1D et 2D. Sur ces figures, le CD normalisé, signifiant le CD contour divisé par le CD nominal est représenté en fonction du réseau (CD nominal= target et pas = pitch). Sur l'ensemble des motifs 1D, il n'y a pas de dégradation du CD, ni à 20 %, ni à 80% de la hauteur de la résine. Ceci est vrai aussi bien avec la source initiale qu'avec la source optimisée, même si on constate une très légère amélioration dans ce deuxième cas. Une analyse similaire est présentée sur la Figure dans le cas des motifs 2D. Cette fois, l'amélioration liée à la source optimisée est plus visible, les CD normalisés étant plus centrés sur le CD nominal indiqué par le trait rouge. Par conséquent, nous pouvons dire que la source optimisée à partir des simulations ne dégrade pas les motifs et permet une amélioration pour certains. Par contre, comme déjà mentionné, l'inconvénient de cette méthode est qu'elle nécessite une analyse à chaque hauteur dans la résine.

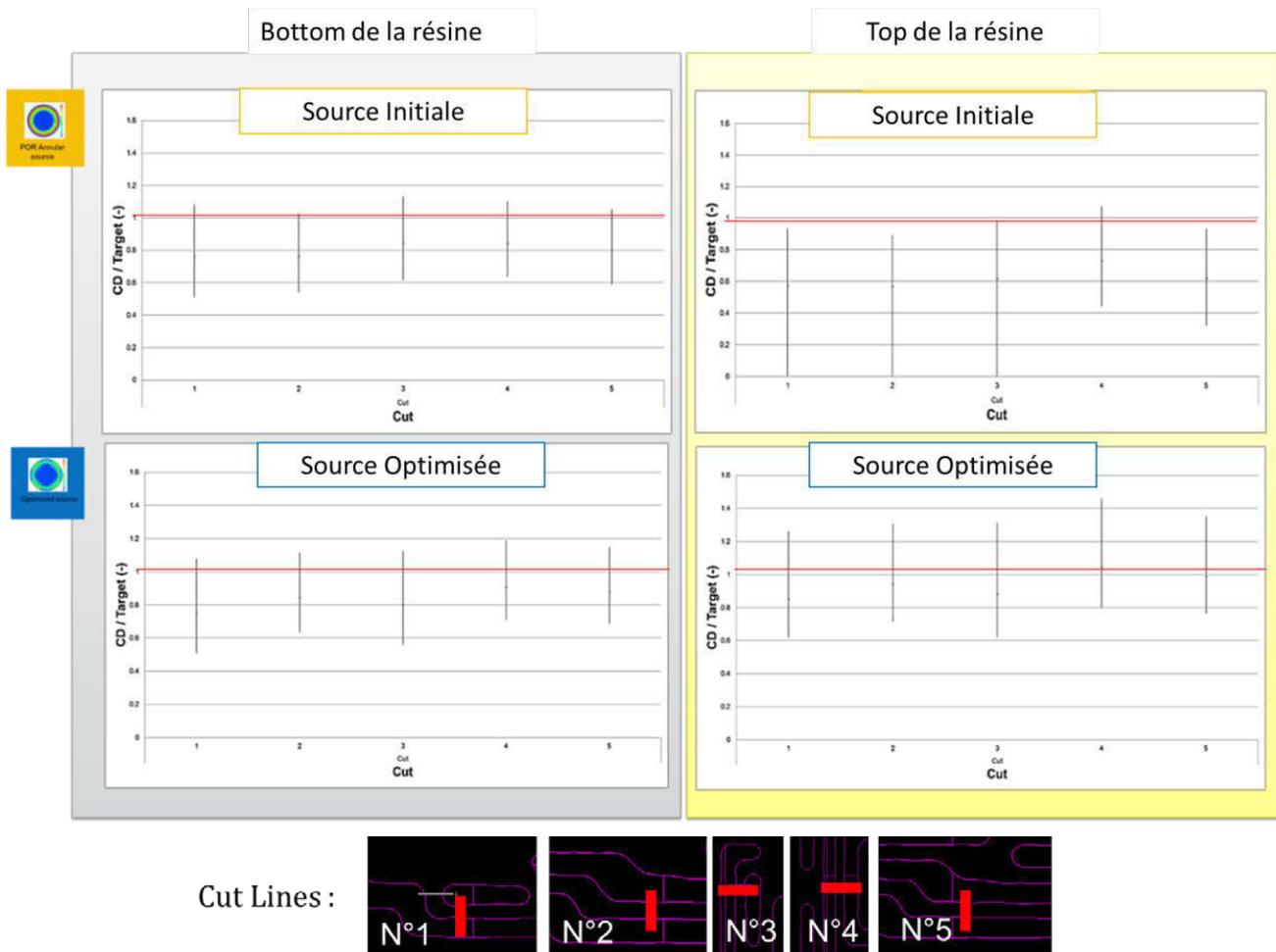


Figure 113 Le CD normalisé ($CD \text{ contour} / CD \text{ target}$) sur les motifs 2D. Les différentes lignes de coupe sont indiquées.

(2) Méthode basée sur l'analyse LMC à travers la hauteur de la résine

Les résultats de la 2^{ème} méthode sont montrés sur la figure 114. Ce procédé est similaire à une analyse LMC standard, sauf que par cette méthode nous rajoutons la notion 3D de la résine. Nous l'essayons sur un layout réduit. Les résultats montrent le nombre de défauts détectés à 20% et 80% de la hauteur de résine, en fonction de la source, initiale ou optimisée. La source optimisée conduit à une réduction du nombre de défauts de 74 % au milieu de la résine, et 17 % pour la partie haute de la résine. Quoique cette méthode ait l'avantage de pouvoir être utilisée sur la puce entière, la limitation est toujours le fait de ne faire l'analyse qu'à une hauteur donnée pour chaque mesure LMC. A noter que sur la figure 114 les notations de A à G correspondent aux différents types de défauts.

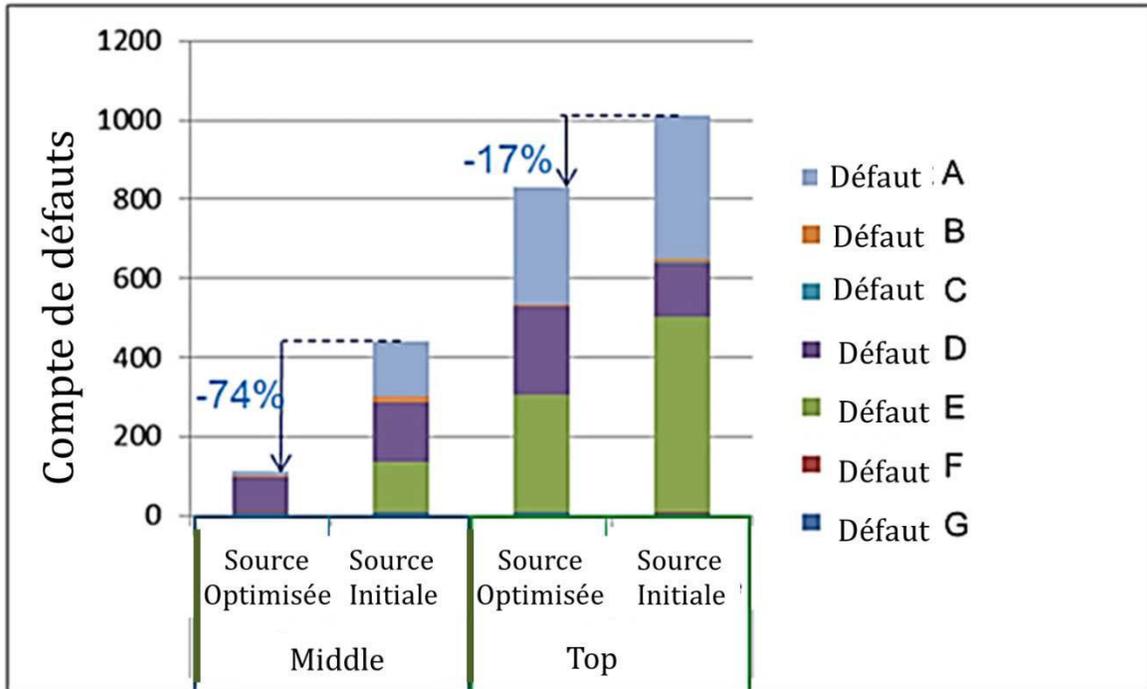


Figure 114 Les résultats sur l'utilisation de l'analyse LMC à travers la hauteur de la résine indiquant le compte de défauts pour chaque source utilisée.

(3) Méthode basée sur un scan de la puce entière détectant les motifs critiques avec une variation de CD prononcée à travers la hauteur de la résine

Les résultats pour la 3^{ème} méthode sont montrés sur les figures 115 et 116. Ces figures représentent le nombre de défauts en fonction de la variation de CD à travers la hauteur de la résine. Nous attendons que la source optimisée fournisse moins de défauts ayant une grande variation de CD signifiant un profil très perturbé.

Parmi ces résultats représentés sur ces 2 figures, l'application de la source optimisée nous permet d'obtenir une gamme moins prononcée de delta CD. Cela signifie que la source optimisée nous permet d'avoir un profil plus droit – qui est le cas idéal – en évitant le profil de type « champignon ». Nous montrons deux cas parmi ces figures, sur le layout de l'optimisation de la source puis sur la puce de production entière (de ce niveau). La figure 115 montre que l'utilisation de la source optimisée permet de réduire la gamme moyenne de variation de CD de 7 nm (source initiale annulaire) à 4.9 nm. Ce résultat est prometteur. La figure 116 présente le même type d'analyse sur puce entière, en se focalisant sur les défauts configuration necking, qui constituent la majorité des défauts réels obtenus. Nous constatons une baisse de variation de CD de 7.1 nm (source initiale) à 5.5 nm (source optimisée) sur le moyen. Ainsi, cela veut dire que la source optimisée peut mieux répondre à la problématique liée au profil de la résine.

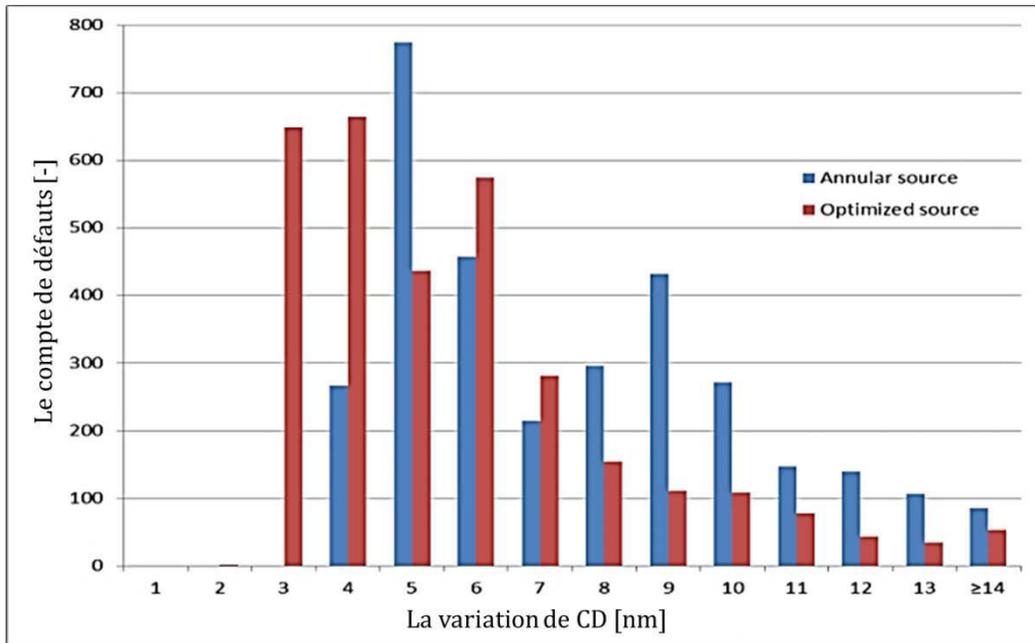


Figure 115 Les résultats sur le compte de défauts en fonction de la variation de CD sur le layout utilisé pour l'optimisation de la source

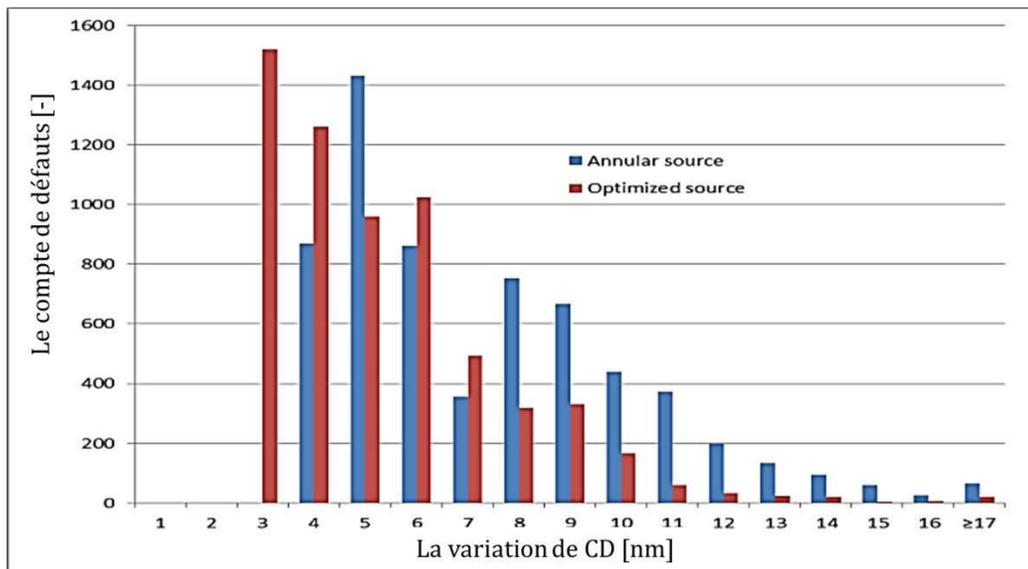


Figure 116 Les résultats sur le compte de défauts en fonction de la variation de CD sur la puce entière

Par conséquent, tous ces résultats préliminaires ont été satisfaisants en utilisant une source qui a élargi la fenêtre du procédé pour un ensemble de structures de calibration. Les résultats sont donc prometteurs. Cependant, un travail supplémentaire doit être mis en place pour compléter cette étude afin d'obtenir un procédé complet plus robuste et fiable. En outre un réglage sur les structures de l'optimisation peut être mis en place également.

6.3.3 Conclusion

L'objectif de ce travail de thèse est de proposer une solution pour diminuer la variabilité de procédé en lithographie afin d'assurer une bonne qualité d'impression pour les dimensions critiques ainsi qu'une profondeur de champ commune (uDOF) adéquate. Une démarche est donc proposée sur la base de la détection des motifs les plus critiques étant impactés par les effets inhérents à la topographie du masque ainsi que le profil de la résine. On évoque ici la notion d'effets 3D. Cette méthodologie consiste en la diminution et la compensation de ces effets, une fois que ces motifs les plus critiques sont détectés. Afin de mettre en place cette démarche sur la variation de procédé, une attention particulière est portée sur la capacité des outils de modélisation prédictive comme les modèles OPC. A ce jour les modèles OPC conventionnels s'avèrent obsolètes pour les applications sur les technologies des nœuds avancés. Il est donc nécessaire d'étendre la compréhension des systèmes en incluant les effets 3D liés aux masques et à la résine apparaissant lors de l'étape d'exposition. Ceci devant par la suite être implémenté dans les outils de modélisation en lithographie.

L'accent a donc été mis, lors de ce travail, sur la construction d'un modèle prenant en compte les effets masque et résine 3D et possédant une capacité de prédiction sur les variations de procédé (par opposition aux modèles OPC standards). Le modèle développé dans ce travail montre une réelle capacité à apporter une amélioration pour les travaux à venir sur les aspects de détection et de diminution des défauts.

D'autre part, une méthodologie de simulation fondée sur la détection a été mise en place. Celle-ci utilise les modèles 3D calibrés. Cette tâche est abordée en considérant les effets liés au masque par le biais de motifs sensibles au décalage du meilleur focus ainsi que les effets liés à la résine par le biais de l'évolution du profil observé. Les résultats montrent que les effets liés au profil de résine doivent être considérés comme les plus influents.

A partir de ces différentes constatations, une nouvelle approche a été proposée afin de réduire les variabilités liées au procédé de lithographie. Sur les effets liés au masque, partant du principe que les effets de topographie du masque induisent des aberrations de front d'onde, une correction du front d'onde est appliquée à partir de l'outil FlexWave. Sur les effets liés au profil résine, une optimisation de la source est appliquée considérant que la source possède un rôle primordial dans la distribution de l'intensité pour l'impression du motif. Dans les deux cas, les résultats obtenus sont prometteurs et participent directement à la diminution des variabilités du procédé de lithographie. Des travaux futurs permettront la mise en place d'un outil encore plus robuste pour être utilisé en production. Néanmoins, une méthodologie de construction du modèle 3D résine est en cours d'implémentation pour le nœud technologique 28nm et pourrait être adapté pour le 14nm.

Bibliography

- Alleaume, C. (2014). *Etude de la modification de la source dans l'utilisation de la méthode de cooptimisation source masque en lithographie optique, mise en oeuvre et application*. Saint Etienne: Université de Saint Etienne.
- Arfaoui, W. (2012). *Etude de la fiabilité par porteurs chaud effets AC/DC*. Crolles: ST Microelectronics.
- Babaud, L. (2010). *Développement et optimisation d'un procédé de gravure grille polysilicium pour les noeuds technologiques 45 et 32 nm*. Grenoble: L'Institut polytechnique de Grenoble.
- Brion Technologies, I. (2008). *Tachyon M3D*. Consulté le 03 16, 2015, sur Brion Technologies: <http://www.brion.com/m3d.pdf>
- Brion, T. (2015). *Brion*. Consulté le 03 05, 2015, sur Brion Technologies Inc.: <http://www.brion.com/lmc.pdf>
- Cambridge, U. o. (2014). *Dissemination of IT for the Promotion of Materials Science*. Récupéré sur <http://www.doitpoms.ac.uk/>.
- CHM. (2015). *Computer History Museum*. Consulté le 07 22, 2015, sur Computer History Museum: <http://www.computerhistory.org/>
- Chong, N.-Y. (2014). "Smart" source, mask, and target co-optimization to improve design related lithographically weak spots. *Proceeding of SPIE 9053, Design-Process-Technology Co-optimization for Manufacturability 9053*.
- Cobb, N. (1998). *Fast Optical and Process Proximity Correction Algorithms*. Berkeley.
- Colidre, B. d. (2014). *Colidre*. Récupéré sur Comité d'information et de Liaison de cadres Dirigeants Retraités de France Telecom: <http://www.colidre-ft.asso.fr/html/weisbuch060704.html>
- Dill, F., Neureuther, A., Tuttle, J., & Walker, E. (1975). Modeling Projection Printing of Positive Photoresist. *IEEE Transactions on Electron Devices*, 456-464.
- Erdmann, A. (2001). Topography effects and wave aberration in advanced PSM-technology. *Proceeding of SPIE 4346, Optical Microlithography*.
- Erdmann, A. (2005). Mask and Wafer Topography Effects in Immersion Lithography. *Proceeding of SPIE 5754, Optical Microlithography*.
- Finders, J. (2014). The impact of Mask 3D and resist 3D in optical lithography. *Proceeding of SPIE 9052, Optical Microlithography*.
- Finders, J., & Hollink, T. (2011). Mask 3D effects: impact on imaging and placement. *Proceeding of SPIE 7985, 27th European Mask and Lithography Conference*.
- ITRS. (2014). <http://www.itrs.net>. Consulté le 2014, sur <http://www.itrs.net>
- Khan, M. (1997). A Semi-Empirical Resist Dissolution Model for Sub-micron. *MSM*.
- Landis, S. (2011). *Lithography*. Wiley-ISTE.
- Laube, P. (2014, 6 8). *Halbleitertechnologie von A bis Z*. Consulté le 2014, sur <http://www.halbleiter.org/>
- Le-Gratiet, B. (2015). Patterning CD control for advanced logic nodes. *Journal of Micro/Nanolithography, MEMS, and MOEMS Volume 14, Issue 2*.
- Liu, P. (2007). Fast and accurate 3D mask model for full-chip OPC and verification. *Proceedings of SPIE 6520, Optical Microlithography*.

- Liu, P. (2011). Accurate prediction of 3D mask topography induced best focus variation in full chip photolithography applications. *Proceedings of SPIE 8166, Photomask Technology* .
- Liu, P. (2012). A full-chip 3D computational lithography framework. *Proceeding of SPIE 8326, Optical Microlithography*.
- Mack, C. (1986). Advanced topics in lithography modeling. *Proceedings SPIE 0631, Advances in Resist Technology and Processing*.
- Mack, C. (2004). 30 Years of Lithography Simulation. *Proc. SPIE 5754, Optical Microlithography*.
- Mack, C. (2007). *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. Chichester: Wiley.
- Mack, C. (2008). *Fundamental Principles of Optical Lithography: The Science of Microfabrication*. John Wiley & Sons, 2008.
- Mack, C. (2015). *Lithography*. Consulté le 06 26, 2015, sur <http://www.lithoguru.com/scientist/lithobasics.html>
- Mack, C. A. (2006). *Fiel Guide to Optical Lithography*. SPIE The International Society for Optical Engineering.
- Maini, A. K. (2013). *Lasers and Optoelectronics : Fundamentals, Devices and Applications*. John Wiley & Sons.
- Manakli, S. (2004). *Amélioration de la profondeur de champ de la lithographie CMOS sub 0,1 µm par des méthodes de double exposition*. Grenoble: Insitut National Polytechnique de Grenoble.
- Michel, J.-C. (2014). *Caractérisation et modélisation des effets d'empilement des couches minces sous la résine photosensible pendant le procédé de photolithographie optique*. Saint Etienne: Université de Saint Etienne.
- Moulis, S. (2014). *Lithographie par division de pas de réseau pour les ciurcuits logiques avancés*. L'Institut National des Sciences Appliquées de Lyon.
- Mulken, J. (2014). Holistic optimization architecture enabling sub-14-nm projection lithography. *Journal of Micro/Nanolithography, MEMS, and MOEMS Volume 13, Issue 1*.
- Pérez, J. (1996). *Optique. Fondements et applications*. Paris: Masson.
- Planchot, J. (2005). *Etude de la scattérométrie pour une application en au suivi en temps réel du procédé de gravure*. Grenoble: INSA.
- Rahhal, L. (2014). *Analyse et modélisation des phénomènes de mismatch des transistors MOSFET avancées*. Grenoble: Université de Grenoble.
- Rodrigue, R. (2010). *Optical Lithography Simulation using Wavelet*. University of Massachusetts.
- Sacek, V. (2006, July 14). Récupéré sur Telescope Optics: <http://www.telescope-optics.net/>
- Saied, M. (2011). Etude de la stabilité et de la précision des modèles utilisés dans la correction des effet de proximité optique en photolithographie. Grenoble.
- Seongbo, S. (2011). Physical Simulation for Verification and OPC on Full Chip level. *Proceedings of SPIE 7973, Optical Microlithography*.
- Staals, F. (2011). Advanced Wavefront Engineering for Improved Imaging and Overlay Application on a 1.35 NA Immersion Soucre. *Proceedings of SPIE 7973, Optical Microlithography*.
- Stanley Wolf, P. (2004). <http://www.latticepress.com/cymer.html>.
- Szucs, A. (2013). Best focus shift mitigation for extending the depth of focus. *Proceeding of SPIE 8683, Optical Microlithography* .
- Wang, C. (2013). Benchmarking study of 3D mask modeling for 2X and 1X nodes. *Proceeding of SPIE 8683, Optical Microlithography*.
- Wikipedia. (2015, 07 21). *Invention of the integrated circuit*. Consulté le 07 22, 2015, sur Wikipedia: https://en.wikipedia.org/wiki/Invention_of_the_integrated_circuit

- Wong, A. (1992). Edge Effects in Phase Shifting Mask for 0.25 μ m Lithography. *Proceedings of SPIE 2197, Optical/Laser Microlithography*.
- Wu, C.-W. (2011). Freeform source optimization for improving litho-performance of warm spots. *Proceedings of SPIE 8166, Photomask Technology*.
- Zeggaoui, N. (2011). *La lithographie par double impression pour les noeuds technologiques avancés*. Grenoble: Université de Grenoble.
- Zeiss. (2014, 06). *Carl Zeiss Microscopy GmbH*. Consulté le 03 2015, sur Carl Zeiss Microscopy GmbH: http://www.zeiss.com/microscopy/en_de/products/scanning-electron-microscopes.html
- Zuniga, C. (2014). Resist toploss and profile modeling for optical proximity correction applications. *Proceedings of SPIE 9052, Optical Microlithography*.