

In-Situ, Portable Monitoring Methods for Photolithography Characterization

Jing Xue



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2009-88

<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-88.html>

May 29, 2009

Copyright 2009, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

In-Situ, Portable Monitoring Methods for Photolithography

Characterization

by

Jing Xue

B.S. (Huazhong University of Science & Technology) 1997

M.S. (Beijing University) 2000

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Costas J. Spanos, Chair

Professor Andrew R. Neureuther

Professor David Attwood

Professor Arun Majumdar

Spring 2009

Abstract

In-Situ, Portable Monitoring Methods for Photolithography Characterization

by

Jing Xue

Doctor of Philosophy in Electrical Engineering

University of California, Berkeley

Professor Costas J. Spanos, Chair

The effective characterization of the Photolithographic performance needs inline information of the fully-assembled tools during production, and this requires *in-situ* monitors to track tool and process information over time. To avoid the costly integration of complex sensors into the processing equipment, *in-situ*, yet *portable* monitoring methods are desirable. In this thesis, we pursue the concept of *in-situ*, *portable* monitoring methods, and more specifically, we introduce two such novel monitors, a Probe Pattern Grating Defocus Monitor (PPG_{DM}) and an on-wafer Integrated Aerial Image Sensor (IAIS). These monitors can be deployed non-invasively at the reticle and the wafer stages in order to characterize lithographic performance.

The first objective of this thesis is to develop an *in-situ*, *portable* monitor to characterize the defocus effect in photolithography by utilizing 90° phase shifted periodic grating in conjunction with scatterometry metrology. The monitor achieves its high sensitivity to defocus by placing transparent lines spaced at the strong focus spillover distance from the centerline of a 90° phase-shifted probe that functions as an interferometer detector. Theoretical simulations show an aerial image defocus sensitivity of 0.83CF/RU at $\sigma=0.1$.

This defocus error is translated into the probe line photoresist trench depth after exposure and development. The relationship between defocus and trench depth is largely linear. Following that, printing experiments and scatterometry calibration confirm this linearity, and demonstrate a high sensitivity of around 1nm defocus/nm trench depth. An excellent correlation between the measured defocus and the scanner programmed defocus is achieved in the experiment. These experimental results show that the PPG monitor is likely the most sensitive defocus monitor currently available.

A small defocus sensitivity difference between horizontal and vertical lines is due to the coma aberration, astigmatism aberration and the illumination angular non-uniformity. Other effects that may impact the experimental results are also discussed: the linewidth effect, the High NA effect, the mask edge effect, the partial coherence effect, the resist effect, the Optical Digital Profilometer (ODP) sensitivity effect, and the measurement noise. The PPG monitor can facilitate run-time process monitoring and support fine tool adjustments and can be conceivably integrated into production reticles and wafers. Depending on the monitoring strategy, the additional cost involved in adding 90° phase-shifted features on the mask may be well compensated by the reduced number of focus-exposure wafers and the overall increase of printing quality and process throughput.

The second objective of this thesis is to explore an *in-situ, portable* monitor to measure the aerial image in photolithography directly and in real time. The Integrated Aerial Image Sensor (IAIS) concept we explored utilizes an electronic aerial image detector integrated into the surface of a tetherless wafer. The detector includes an aperture mask that is integrated on top of a CCD array. The aperture mask features a series of “moving” aperture groups which sample the aerial image to form a low spatial frequency

interference pattern. This low spatial frequency image can be captured by the CCD array with an effective spatial resolution of only a few nanometers. Simulations show that an aperture width of $35\text{nm}\pm 15\text{nm}$ and thickness of $80\text{nm}\pm 20\text{nm}$ in amorphous silicon is nearly optimal for the 65nm node, and that fabrication precision is relatively not critical.

Viable process techniques to fabricate such a sensor are explored, the main steps being the patterning of the IAIS aperture mask and the assembly of photo-detector chips onto wafer carrier templates. Because of limited time and budget, we have not been able to experimentally demonstrate the integration of the aperture mask onto the real CCD array, but made considerable progress towards patterning the mask on the dummy chips and the hybrid assembly of the CCD array within a carrier wafer. Adhesive bonding and capillary force assisted self-assembly were explored and experimental results point to limitations in the silicon bulk etching non-uniformity leading to height differences and some tilt of the finished wafer-CCD chip assembly. Capillary force is used to improve the self-alignment accuracy. Experimental results of the capillary force assisted assembly show negligible tilt and height difference down to about $1\mu\text{m}$.

A proof-of-concept prototype was designed and tested in the i-line GCA6200 stepper in the Berkeley Microlab, but securing and aligning this tethered assembly onto the wafer chuck proved to be a big challenge. It is clear that the proper vehicle for testing the IAIS concept will require a completely wireless, self-contained platform, which has the form factor of a regular silicon wafer.

Functioning as a recording tool, IAIS is able to detect aerial image features, such as intensity, contrast, slope, and feature size, as verified by first-principle simulations. Because of the structure of the design, IAIS is also a natural polarizer. A set of rotated

IAIS gratings with 90° , 45° and 135° could be integrated into the same aperture mask to investigate the linear polarization components and ratios with a single exposure. Finally, combining the purpose-designed reticle, such as PPG, with the on-wafer IAIS and a suitable aperture mask, will bring forth a new family of *in-situ, portable* sensors that will dramatically improve the state of the art in lithography characterization.

Professor Costas J. Spanos
Dissertation Committee Chair

To memorize my father

Table of Contents

Acknowledgements	iv
Chapter 1 Introduction.....	1
1.1 Thesis Contributions	4
1.2 Thesis Organization.....	5
Chapter 2 Photolithography and Metrology Techniques.....	7
2.1 Photolithography Imaging System and Imaging Theory	7
2.1.1 Photolithography Imaging System	9
2.1.2 Photolithography Imaging Theory	23
2.2 Metrology Techniques.....	30
2.2.1 Scatterometry Based Metrology	31
2.2.2 On-Wafer Sensors.....	35
Chapter 3 Probe Pattern Grating Defocus Monitor Based on Scatterometry Metrology.....	42
3.1 Review of Defocus Effects on Lithographic Images	43
3.2 Review of Defocus Monitor Techniques	46
3.3 Design of Probe Pattern Grating (PPG) Defocus Monitor	52
3.3.1 Concept of Probe Pattern Grating (PPG) Defocus Monitor	52
3.3.2 Design of the Probe Pattern Grating (PPG) Defocus Monitor	56
3.3.2.1 Optimizing the Reference Line Width of the PPG	57
3.3.2.2 Optimizing the Pitch of the PPG	65
3.3.3 Mathematical Justification of the PPG Defocus Monitor	71
3.3.4 Aerial Image Simulation of the PPG Defocus Monitor	82
3.3.5 Resist Image Simulation of the PPG Defocus Monitor	87
3.4 Printing Experiments and Scatterometry Calibration of PPG Defocus Monitor	98
3.4.1 Printing Experiments of PPG Defocus Monitor	98
3.4.2 Scatterometry Calibration of PPG Defocus Monitor	101
3.4.3 Analysis of the Experiment and Simulation Results	116
3.4.3.1 Defocus Sensitivity	116
3.4.3.2 Dose Sensitivity	118
3.4.3.3 Line Width Effects	120

3.4.3.4 Partial Coherence Effects	122
3.4.3.5 High NA and Mask Edge Effects	123
3.4.3.6 Resist Effects	124
3.4.3.7 ODP Effects	125
3.5 Conclusions	130
Chapter 4 Integrated Aerial Image Sensor	135
4.1 Review of Aerial Image Monitor Techniques	136
4.2 Design of the Integrated Aerial Image Sensor (IAIS)	142
4.2.1 IAIS System and Interference Sampling Concept.....	142
4.2.2 IAIS Aperture Mask	146
4.2.2.1 IAIS Aperture Mask Material	147
4.2.2.2 IAIS Aperture Mask Geometry	152
4.2.2.3 Other Features of IAIS Aperture Mask	165
4.2.2.4 Summary of IAIS Aperture Mask Design	165
4.2.3 IAIS Photo-Detector Considerations.....	166
4.3 Applications of the Integrated Aerial Image Sensor (IAIS).....	168
4.3.1 IAIS as an Aerial Image Metric Detector.....	168
4.3.2 IAIS as a linear Polarization Monitor	180
4.4 Fabrication of the Integrated Aerial Image Sensor (IAIS).....	187
4.4.1 Aperture Mask Material Deposition	189
4.4.2 Aperture Mask Patterning	191
4.4.3 IAIS Assembly	198
4.4.4 IAIS Prototype Testing in the Berkeley Micro-fabrication Laboratory	209
4.4.5 IAIS Sources of Error.....	212
4.5 Conclusions	213
Chapter 5 Conclusions and Future Work	220
5.1 Conclusions	220
5.2 Future Work	223

Acknowledgements

First of all, I would like to thank my advisor, Professor Costas J. Spanos, for his guidance and support during my years at Berkeley. I thank him to open the door of the semiconductor manufacturing to me. His remarkable knowledge and broad vision in the semiconductor metrology and the statistical process control have been inspiring my interests and potentials on the research career development. His excellent guidance and invaluable suggestions help me in every step of my study at Berkeley. His gracious understanding and open-minded tolerance allow me to make a slow but considerable progress during these years. I specially thank him to “push” me but company me to finish this dissertation. He is the source of advices, knowledge, and encouragement. I would like to express my deepest gratitude to him again.

Professor Andrew R. Neureuther opened another door of the lithography modeling to me. I learned lithography, from the basic concept to the advanced modeling from him. His unique and intelligent thinking method impressed me and inspired me deeply. Without his insight, persistent support and fully understanding, the first part of this dissertation work is impossible. Although tons of knowledge and time are required to equip me to become a good lithographer, the precious leaning process from him will benefit my current and future career for sure.

I thank Professor David Attwood, Professor Arun Majumdar for serving as the committee in my Qualification Exam and dissertation. Their encouragements always give me confidence to proceed.

I thank Kurt Moen and Chaohao Wang to help me on the IAIS project. In addition to

their contributions to the project, they let me think and know how to perform a good team work. I wish they have great achievements at graduate schools.

I am grateful to Dr. Xiaofan Meng, Matthew Wasilik, Joe Donnelly, Marilyn Kushner, Bill Flounders, and other staffs in the Berkeley Microlab for helping me in the equipment training and fabrication process.

Thanks so much to Changrui Yin for his support in the ODP station, BCAM server and computers, and all related issues. Thanks to BCAM fellows for their friendship and supports: Yu Ben, Jason Cain, Runzi Chang, Paul Friedberg, Justin Ghan, Ning Ma, Kedar Patel, Kun Qian, Qianying Tang, Dekong Zeng, and Qiaolin (Charlie) Zhang. I am especially grateful to Qiaolin (Charlie) Zhang for introducing me to Synopsys Inc., where my career has a brand new start.

I would like to thank Ruth Gjerde, Charlotte Jones for their assistance with administrative and grant issues. I appreciate Ruth's great help during my Qualification Exam and dissertation periods.

Thanks Kwame Eason, Jeffrey Schefske, and Phillip Jones from Spansion Inc. for supporting lithography and scatterometry calibration.

Most importantly, thanks for the unconditional love and support from my dear mom and sisters. This dissertation is dedicated to them and to memorize my father.

Finally, I would like to thank my husband, Zeyu Li, for his love and support through all the years, and being with me to spend all the up and down at Berkeley. Without his support, this dissertation would not be possible. I would like to cheer this moment with him and wish we have a better and happier life in the future!

Chapter 1

Introduction

Printing is one of the most indispensable techniques contributing to our culture, commerce, law, religion, and other social functions. Printing streamlines the process of communication and enables the scientific revolutions. Printing originated in China. The history of printing can be traced back to the later Han Dynasty (25-220 CE). The Chinese processed the three necessary ingredients of printing: Paper, the recording medium; ink, the writing material; and template, the model to make replicas. The art of lithography was invented by Alois Senefelder in 1798 in Austria by printing ink on a series of metal or stone carvings. The Greek word *litho* means stone, and *graphy* denotes writing. Lithography is the precursor of the modern printing technologies [1-1] [1-2].

Photolithography is a conceptual extension of lithography. Rather than ink, it uses light as the writing material. E-beam lithography is another example which uses electron beam as the writing material. The template of the greased and moistened areas on the limestone becomes the clear (non-chrome) and dark (chrome) region in the reticle or mask in photolithography. The typical recording material in photolithography is a photoresist that receives the light energy and generates the photochemical reactions to form the printing patterns. Nowadays photolithography is one of the most critical processes in semiconductor manufacturing. It patterns nearly all integrated circuits fabricated today. Lithography (Lithography means photolithography in the following contents) has been the biggest contributor to drive the semiconductor industry to validate the prediction of

Moore's law: "the complexity of the integrated circuit measured by the number of components (transistor, capacitors, diodes, resistors) per chip, would double approximately every two years" [1-3]. The advancement of the optical lithography resolution relies upon the improved exposure tool precision and enhanced imaging strategies.

Exposure tools evolved from contact, proximity, full wafer scanner to step-and-repeat and step-and-scan system. With the enhanced printing capability, the wafer stepper prices have been increased by around 17% per year. A step and scan system now costs close to \$20M, and their prices are projected to increase in the future [1-4]. As the feature size reduces to the nanometer range, slight imperfections and variations in certain part of the tool can have disastrous effect to the image quality and hence the device performance. Advanced monitoring techniques to characterize the lithography tool, and therefore improving the equipment precision and the process yield has become critical components to IC production. Most of the conventional monitoring techniques either off-line measure the element of the tool before tool assembly or integrate the detector into the individual equipment to monitor the specific tool. However, an *in-situ, portable* monitor is needed to characterize the equipment status in line over the real time and effectively compare multiple tool performances using a shared sensor. The available stages for the end user to access and apply the monitors to the wafer stepper without impacting the original hardware are the reticle stage and the wafer stage. Reticle engineering can be used to create purpose-designed mask to detect and qualify a specific tool effect, such as defocus and illumination. In combination with a purpose-designed mask, special wafers equipped with sensors can be deployed on the wafer stage to directly detect the aerial image effect.

In this thesis, we investigate in detail these two kinds of monitors at the reticle stage and the wafer stage.

Maintaining an optimal focus level of the exposure tool is critically important and particularly challenging in the high NA lithography. The available Depth of Focus (DOF) for 65 to 32nm technology nodes is only around 200nm to 100nm with 3% exposure latitude [1-5]. Another trend is that as k_1 decreases, it makes the process quite sensitive to focus and dose variations, lens aberrations, and other process variations [1-6]. The need for an advanced monitoring method to address subtle defocus variations when printing small geometries presents a new challenge to metrology. A highly sensitive Probe-Pattern Grating (PPG) Defocus monitor is introduced in this thesis to intently address this challenge [1-7].

Optical and resist effects interact and impact the end printed features in a complex way. In order to characterize the exposure tool performance and isolate the sources of process variation, a direct, resist-less measurement of the aerial image is required. However, because of the nature of wafer stepper design, it is difficult to measure the aerial image directly and in real-time without modifications of the exposure equipment. Dedicated aerial image sensors, even when available, cannot measure the aerial image truly *in-situ*. We address this problem by exploring an on-wafer Integrated Aerial Image Sensor (IAIS) [1-8].

1.1 Thesis Contributions

This thesis proposes *in-situ, portable* monitoring methods through lithography stepper's reticle stage and wafer stages for lithography characterization. The first major contribution is the development of a Probe-Pattern Grating (PPG) Defocus Monitor, including pattern design, theoretical verification and experimental verification. The concept of reciprocity interferometry is utilized to create a highly defocus-sensitive one-dimensional periodic grating. The aerial image defocus sensitivity of around 0.83CF/RU at $\sigma=0.1$ is achieved through *SPLAT* simulation and theoretical analysis. A high sensitivity around 1nm defocus/nm trench depth is confirmed experimentally with good repeatability. PPG experimental result matches to the theoretical analysis and the simulation results with the similar trend of probe trench depth versus defocus. This monitor achieves a higher sensitivity than the known defocus monitors based on the phase shift mask and the binary mask designs. In addition, PPG has a strong linear response to defocus and hence significantly reduces the cost of the numerical computing required in the traditional scatterometry based process monitor techniques.

The second major contribution is introducing a novel on-wafer sensor to lithography process. The Integrated Aerial Image Sensor (IAIS) can retrieve the aerial image in nanometer range by sufficient sampling through an on-wafer aperture mask, and then the spatially amplified aerial image can be captured by a near field photodetector array with pixel size in the micrometer range. The method to determine the optimum parameter design window at a given technology node is provided. The sensor application, such as aerial image metrics calibration, and linear polarization monitoring, is discussed and verified by simulations. Viable process techniques to fabricate such a sensor are also

explored. Being aware of the full challenges in the assembly, patterning, and testing, we propose and test various promising process flows. Although more collaboration with the MEMS industry is needed to overcome some fabrication challenges, this work provides the needed insights to improve the sensor fabrication and operation, and extend the on-wafer sensor concept to a new process area.

1.2 Thesis Organization

The remainder of this thesis is organized as follows: The background of photolithography imaging system and imaging theory, and relevant metrology techniques is reviewed in Chapter 2. In Chapter 3, we extensively discuss the design, modeling, simulation, printing experiment and scatterometry calibration of the Probe-Pattern Grating (PPG) Defocus Monitor. The Integrated Aerial Image Sensor (IAIS) are presented in Chapter 4. The design, manufacturing, and the applications of IAIS are investigated in this chapter. Finally, chapter 5 summarizes the dissertation and provides suggested future directions of relevant inquiry.

References

- [1-1] http://en.wikipedia.org/wiki/History_of_printing
- [1-2] A. K. Wong, "Resolution Enhancement Techniques", *SPIE*, 2001
- [1-3] G. E. Moore, "Lithography and the Future of Moore's Law", *Proc. of SPIE*, vol. 2440, pp.2, 1995
- [1-4] H. J. Levinson, "Principles of Lithography", *SPIE*, 2005
- [1-5] B. L. Lin, "Immersion lithography and its impact on semiconductor manufacturing", *Journal of Microlithography, Microfabrication and Microsystem*, vol.3, pp.337, 2004
- [1-6] T. A. Brunner and C. P. Ausschnitt, "Process Monitoring Gratings", *Proc. of SPIE*, vol. 6518, 651803-1, 2007
- [1-7] J. Xue, C. J. Spanos, and A. R. Neureuther, "Probe-Pattern Gating Focus Monitor Through Scatterometry Calibration", *Proc. of SPIE*, vol 6922, 692225, 2008
- [1-8] J. Xue, and C. J. Spanos, "Design of an Integrated Aerial Image Sensor", *Proc. of SPIE*, vol. 5752, pp392, 2005

Chapter 2

Photolithography and Metrology Techniques

Photolithography is the major process to print the integrated circuit patterns on wafer. Metrology is fundamental to semiconductor manufacturing control. An accurate measurement is particularly significant for lithographers, where the extremely small features of state of the art printing processes have pushed measurement capabilities to their limits and beyond [2-1].

In this chapter, the basics of photolithography and metrology techniques are reviewed, which provides the background of this thesis work. In section 2.1.1, the lithography system and the related issues are discussed, in which wavefront aberration at the projection lens pupil and the phase shift mask are emphasized. The imaging theory is shown in section 2.1.2. Two modeling methods, Hopkins' method and Abbe's method are reviewed. The optical diagrams are provided to understand the methods. The metrology techniques are discussed in section 2.2, with special attention to scatterometry and the on wafer sensors in section 2.2.1 and 2.2.2 respectively.

2.1 Photolithography Imaging System and Imaging Theory

The wafer stepper was first introduced commercially in the late 1970s by the GCA Corporation of North Andover, Massachusetts [2-1]. DSW4800 is the first GCA system, which could achieve resolution of $1.25\mu\text{m}$, overlay of $\pm 0.7\mu\text{m}$ and depth of focus of $1.5\mu\text{m}$. Its list price was about \$300,000. The name of DSW refers to Direct Step on the

Wafer. Following GCA, several companies began to produce wafer steppers commercially, and today Nikon, Canon, and ASML are the major stepper suppliers. Modern steppers provide much greater imaging capability and overlay than the original GCA steppers, but the operating principles are essentially the same. Nowadays, GCA 6200 is still well maintained in the UC Berkeley Micro-fabrication laboratory to support students' MOS/MEMS manufacturing experiments.

Wafer stepper undergoes the evaluation from step-and-repeat to step-and-scan. The reticle stage scanning synchronized motion with the wafer stage is the key in the step-and-scan system. Step-and-scan system achieves improvement in overlay and better linewidth control without requiring significant improvement in lens technology, provides compensation for reticle errors because magnification is adjusted independently in the directions parallel and perpendicular to the scan, and is capable to print large field sizes cost effectively. Synchronizing the stage positions at high speed scanning and controlling the mechanical vibrations are significant challenges in the modern step-and-scan system.

The advancement in lithography has been facilitated by the predictive theoretic models. Lithography simulations enable lithographers to understand the relationship of the lithographic parameters, predict process yields and determine the process parameters before the real fabrications, and help the new lithographic process methods and monitor techniques to be developed and applied. The first generally available program for calculating optical image and the resist profile was SAMPLE [2-2], developed by researchers at UC Berkeley in 1980. A notable implement of Hopkins' method was implemented by O' Toole *et al.*[2-3], as part of SAMPLE. Later Toh make the addition of the program to calculate lens aberrations and also upgraded the name of the simulator to

SPLAT [2-4]. As feature size smaller than light wavelength, the finite thickness and shapes of the apertures have effects on the resulting image. Another program, TEMPEST, was developed at UC Berkeley for modeling three-dimensional mask effects and the polarization effect [2-5]. Later the lithographic simulation programs can be performed over the internet at the website at UC Berkeley [2-6].

PROLITH [2-7], another modeling program was developed at the United States department of Defense. Later generations of PROLITH became the product of a commercial product of FINLE Technologies (now a division of KLA-Tencor). Several EDA companies sell lithography modeling program, such as Solid-C from Sigma-C [2-8], Calibre from Mentor Graphics [2-9], Proteus from Synopsys [2-10].

2.1.1 Lithography Imaging System

The schematic diagram of a projection lithography system is shown in Figure 2-1 [2-11]. The system comprises of a narrow bandwidth light source radiating at a wavelength of λ , a condenser lens, a mask with the circuit patterns to be fabricated, the projection optics consisting of multiple lenses, and the wafer chuck.

The light source and the condenser lens constitute the illumination system, where the source is placed at the focal plane of the condenser. Most DUV systems today use excimer lasers as the light source, in which ArF laser is used for 193nm lithography. The source can be regarded as a collection of incoherent point sources [2-12]. Each point source independently radiates a spherical wave to be captured by the condenser lens. The condenser lens redirects the light and forms the image of the source at the entrance pupil plane of the projection lens. The illumination system establishes a specified level of

coherence, controls the exposure dose, and provides the uniform light across the reticle. The degree of coherence, expressed in σ , describes the correlation of the electric field interactions between two nearby mask locations at the image plane. A small σ represents high coherence and therefore the electric fields of two far apart points will interact with each other in phase sufficiently. Large σ indicates that the interaction distance is small. The degree of mutual coherence, represented by μ , depends on the spatial extent of the illumination source. The mutual coherence function [2-12] is calculated by the inverse Fourier transform of the source distribution.

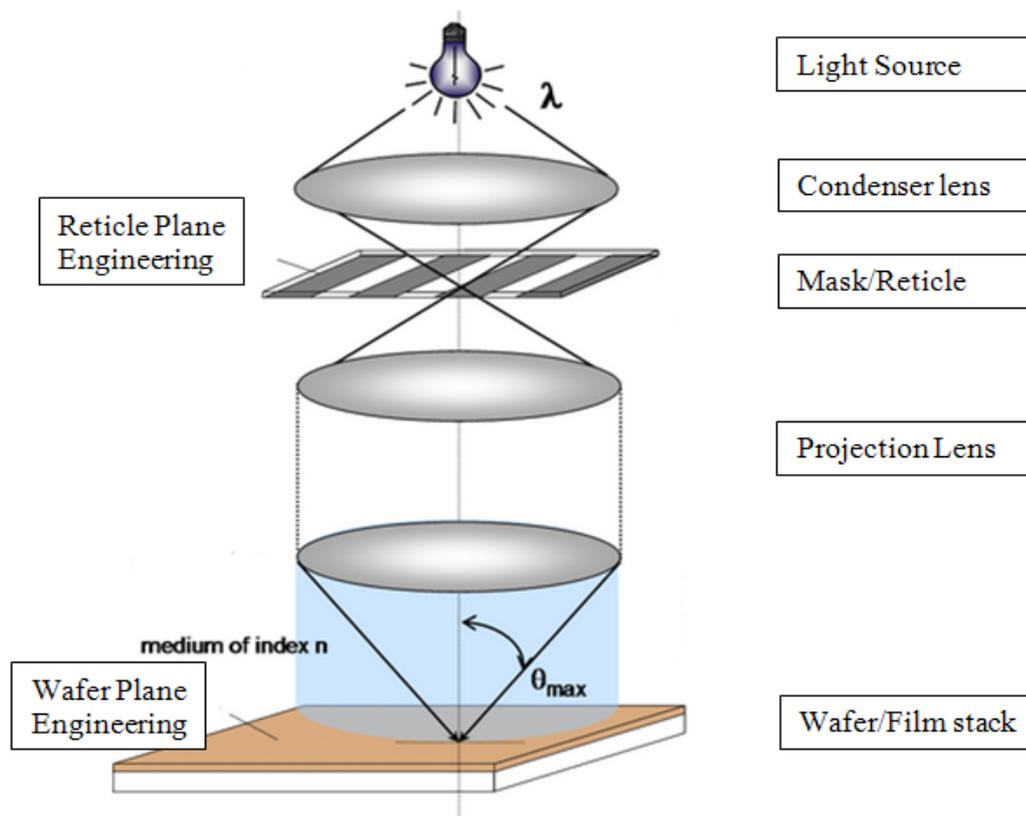


Figure 2-1 Schematic diagram of a projection lithography system. Refer to [2-11].

Illumination uniformity is achieved by a number of approaches, such as fly's eyes [2-13], in which an array of lens-lets is arranged to average the amount of light over the field. Generally, we consider each point source illuminates the entire mask uniformly, which ensures the across field dose uniformity. For a step and scan system, with a uniform illumination achieved over the entire slit, a good dose control requires the same number of pulses for every exposure. To avoid the so-called quantization problem [2-1], the illumination is generally non-uniform in the scan direction, a trapezoidal shape being typical.

The mask is located at the exit pupil of the condenser optics and the focal plane of the projection lens. In order to distinguish the mask being made with multiple chips and the mask containing the pattern of only one chip, the single chip mask is called a reticle. The mask diffracts the illumination fields into scattered orders and relays them to the projection lens. The diffraction angles depend on the features on the mask. For each source point, the diffraction orders from each mask location are correlated to each other. The resulting electric field at the image plane is the vector superposition of the interference electric field of the diffraction orders. The image intensity is therefore the absolute sum of the intensities over all illumination source points.

The mask patterns are transferred to the wafer by the repeated imaging of the reticles. The quality of the wafer patterns, such as the linewidth, defects, and overlay, is strongly affected by the quality of the mask parameters. The pellicle, a thin transparent polymer film is generally attached to the patterned mask to avoid printable defects during the extend mask usage [2-1]. Photo-masks are fabricated with techniques similar to those used in wafer processing, where electron-beam lithography or optical mask writers are

generally used to pattern chromium (opaque) film on the glass or quartz substrate. Typical chromium etches involve mixtures of Cl, O₂ and other additional gas. Since Chromium is a difficult film for dry etch, some alternative opaque materials, such as MoSi, has been considered. Nevertheless, chromium continues to be the dominant choice of absorber material, MoSi has found considerable applications in the attenuated phase shift masks.

To improve image capabilities, the masks have evolved from simple binary mask to the complex phase shift mask, where 180°, 90° and other phase differences are created on the mask by etching into the quartz. The relationship between the etch thickness and the phase shift amount measured in λ is expressed as:

$$d = \frac{\text{Phaseshift}(\lambda)}{(n_1 - n_2)} \quad \text{Eq. (2-1)}$$

Where n_1 is the refractive index of the glass, and n_2 is the index of the surrounding medium. For a 180° degree phase shift in air, the etch thickness is: $d = \frac{\lambda}{2(n_1 - 1)}$. Since

glass $n_1 \approx 1.5$ at 193nm, the etch thickness is about a λ . And the etch thickness for the 90° phase shift is around $\lambda/2$. The 180° phase difference is used in the alternating phase shift mask (ALT-PSM) and attenuated phase shift mask (ATT-PSM). In this thesis, we designed a 90° phase shift grating mask as the defocus monitor which will be discussed in Chapter 3. One of the concerns of the phase shift mask is the intensity imbalance effect between phase shifted and non-shifted regions, which was first revealed by A.Wong [2-14]. Wong observed that near material interfaces electromagnetic fields tend to concentrate in high index materials. This phenomenon was verified by Rigorous finite difference time domain (FDTD) simulation, where Maxwell's equation and the boundary

conditions associated with the thick mask topography are solved to predict the electromagnetic interactions. Since the edge effect [2-15], which may extend 10% of a wavelength from the edge is related to the topography feature height which does not scale with feature size, electromagnetic vector effects in phase shift mask have a larger effect on the pattern image as feature sizes become smaller, where a distortion of the fields transmitted through the mask and therefore impacting the photo-resist image is expected.

The demagnified images of the mask patterns are projected onto the wafer surface by the projection lens, mounted a few millimeters above the wafer surface. Most step-and-scan systems demagnify the image from the reticle by 4x. The incident field at the pupil is the Fourier transform of the mask geometry. The projection lens then creates the images on the wafer by inverse Fourier transforming the diffraction orders collected by the lens. Some of the diffracted light which carries the high frequency information from the mask may not be captured by the pupil due to the limited pupil size. So the pupil of the lens serves as a low-pass filter. The nature of the diffraction limited lens causes the image distortions, such as corner rounding, line-end shortening, and etc. Optical proximity correction strategies are generally used to tune the mask layout geometry to counteract such image distortions.

Lenses for 193nm lithography typically use fused silica and a small quantity of CaF_2 . Since fused silica is slightly absorbing at 193nm, lenses heat up and expand during exposure, causing defocus [2-16], changes in magnification, and other aberrations. Focus is the parameter affected most by lens heating. Changes to other aberrations are typically small. Flare occurs when the surface of the lens elements is rough, which is particularly serious for CaF_2 because the hardness of CaF_2 depends upon crystallographic orientation

[2-17]. Flare can also arise by scattering from surface defects, or contaminations from the deposited material of photochemical reaction, or can be caused by damaged or nonoptimized antireflection coating. Flare results to background light. For long range flare, the background light is throughout the imaged area and thereby reduces the image contrast.

Actual optics differs from the ideal diffraction-limited concept. These differences are typically known as aberrations. The potential sources of aberrations are imperfect design, distorted lens and mirror surfaces, lens material inhomogeneity or imperfection, imperfect lens assembly, lens heating, and etc. Since light rays travel across different parts of the lens surface, various imperfect curvature of the lens surface and variations of the glass optical constants cause optical path difference (OPD) of rays, especially when a large image field size is combined with a high numerical aperture (NA). Aberrations that vary across the image field contribute to across-field linewidth variations [2-18]. To meet modern lithography requirements, the index of refraction for optical-materials must be uniform to less than one part per million, and lens surfaces may deviate from design value by no more than a few nanometers (rms) [2-1].

The concept of wavefront aberration is illustrated in Figure 2-2. Because of lens aberrations, the actual wavefront may deviate from a perfect spherical wave.

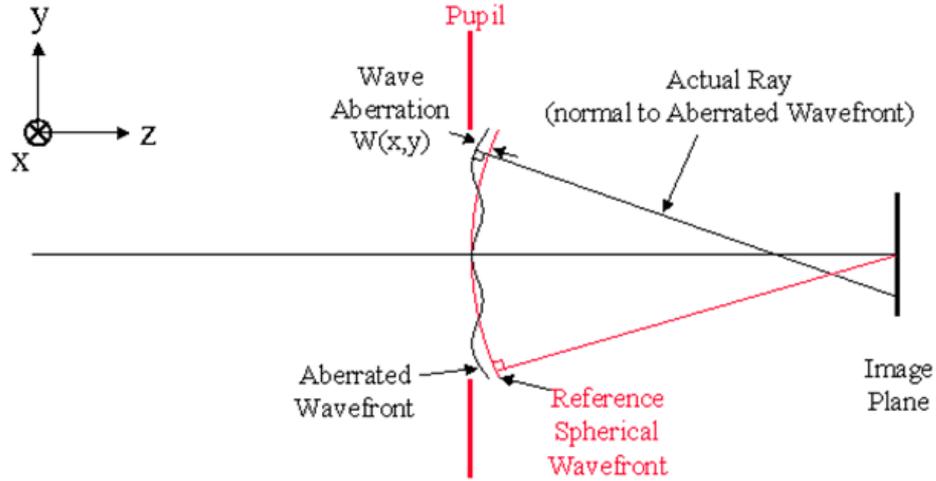


Figure 2-2 Illustration of wavefront aberrations [2-19]

The deviations are expressed as the phase error: $\Phi = \frac{2\pi}{\lambda} OPD$, at points of the exit pupil. Since every different ray passes through a different part of the aperture, the OPD could be defined as a surface across the aperture whose shape is much like the aberration plate. Cylindrical coordinates (ρ, θ) can conveniently to define the location of the ray within the aperture. The shape of the $OPD(\rho, \theta)$ surface specifies the aberration at a particular pupil point and is normally represented as a sum of Zernike Polynomials.

$$OPD(\rho, \theta) = \sum_j a_j \lambda Z_j(\rho, \theta) \quad \text{Eq. (2-2)}$$

Where Z_j is the j th orthonormal Zernike polynomials. The a_j presents the root-mean-square deviation of the OPD surface contributed by the j th Zernike term in units of wavelength λ . To fully characterize a lens system, measuring the Zernike coefficients at many points across the pupil is needed [2-20]. The orthonormal Zernike polynomials are expressed as [2-21]:

$$\begin{aligned} Z_{even,j}(\rho, \theta) &= R_n^m(\rho) \cos m\theta \\ Z_{odd,j}(\rho, \theta) &= R_n^m(\rho) \sin m\theta \end{aligned} \quad \text{Eq. (2-3)}$$

The even j corresponds to a symmetric polynomial given by $\cos m\theta$, while odd j is the antisymmetric polynomial varying as $\sin m\theta$. The first 36 Zernike terms are shown in reference [2-21]. We list the first 11 terms in table 2-1. The defocus aberration is of special interest and will be discussed at length in Chapter 3.

Lens aberrations cause lithographic image degradations and variations, such as image position shift, image contrast decrease, image asymmetry, process window reduction, appearance of undesirable image artifacts, etc. Two examples of spherical and coma aberrations are discussed now. Spherical is a typical even aberration in addition to defocus, and it can be seen as a focus shift depending on the pupil radius. Indeed spherical Z_{11} in table 2-1 can be rewritten including the defocus term Z_4 :

$$Z_{11} = \frac{3}{2}(2\rho^2 - 1)Z_4 - \frac{1}{2} \quad \text{Eq. (2-4)}$$

Equation (2-4) shows that spherical aberration leads the best focus plane to be dependent on the pattern pitch and the illumination. The angle at which a ray is diffracted by a grating is a function of the pitch and the illumination, and therefore the radius point where the light reaches the pupil is dependent on the pitch and the illumination. Since the focus shifts with radius, the best focus plane varies with the pattern density and the method of illumination. Another aspect of spherical aberration is that it affects the symmetry of the image above and below the focal plane. Eq. (2-5) indicates that the focus degrades more at the negative direction than the positive direction, so the process window might be cut off at the negative defocus.

Table 2-1 The first 11 Zernike Polynomial terms

Z_j	Equation of Z_j	Name
Z_1	1	Piston
Z_2	$\rho \cos(\theta)$	x tilt
Z_3	$\rho \sin(\theta)$	y tilt
Z_4	$-1 + 2\rho^2$	Defocus
Z_5	$\rho^2 \cos(2\theta)$	Astigmatism H/V
Z_6	$\rho^2 \sin(2\theta)$	Astigmatism $\pm 45^\circ$
Z_7	$(3\rho^3 - 2\rho) \cos(\theta)$	Coma x
Z_8	$(3\rho^3 - 2\rho) \sin(\theta)$	Coma y
Z_9	$\rho^3 \cos(3\theta)$	Trefoil x
Z_{10}	$\rho^3 \sin(3\theta)$	Trefoil y
Z_{11}	$6\rho^4 - 6\rho^2 + 1$	Primary Spherical

Coma is a typical odd aberration, which causes image contributions from different pupil radii to shift relative to each other. The consequences of coma are that symmetric patterns may print asymmetrically, and the pattern center position may shift. The three bar image with and without coma aberration in [2-20] is a good example that shows the asymmetric effects. The image position shift depends on the mask pattern and the illumination, in contrast to tilt (Z_2, Z_3) aberration where the image shift is independent of the details of the pattern. The coma in table 2-1 can be rewritten as:

$$(Z_7, Z_8) = (3\rho^2 - 2) \cdot (Z_2, Z_3) \quad \text{Eq. (2-5)}$$

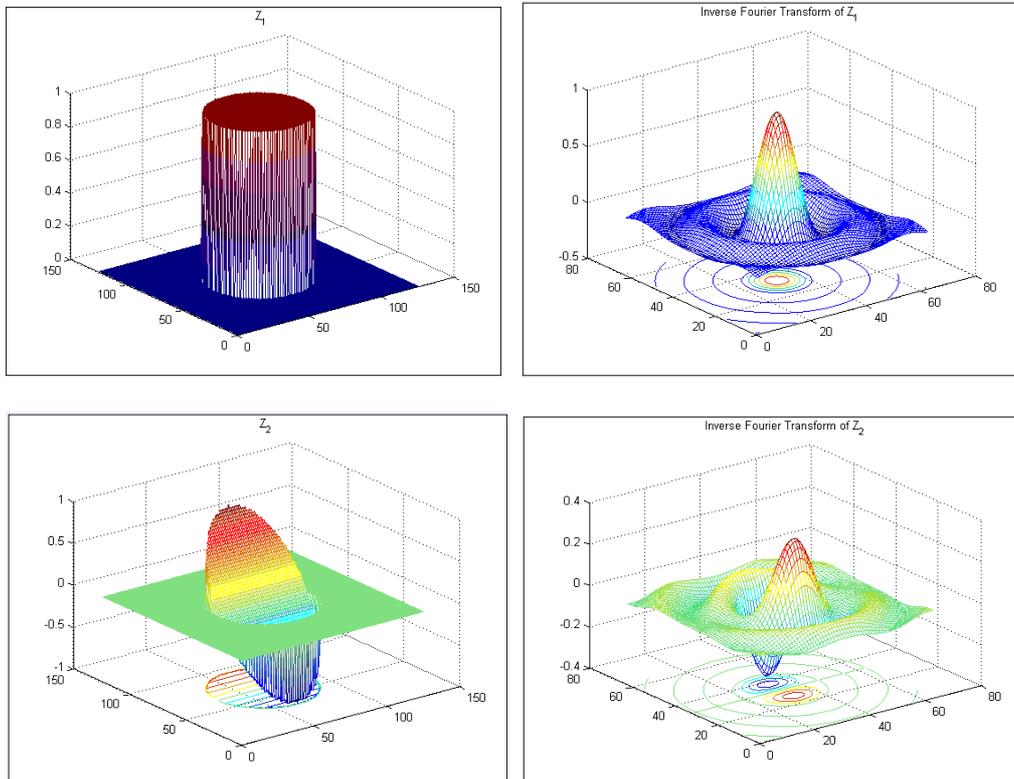
So the shift is a function of the pupil radius. Similarly, the coma effect depends on the mask pattern as well as the illumination.

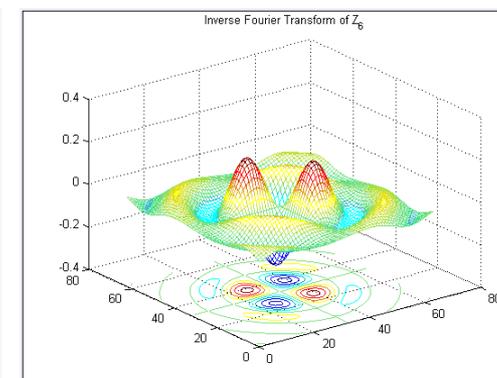
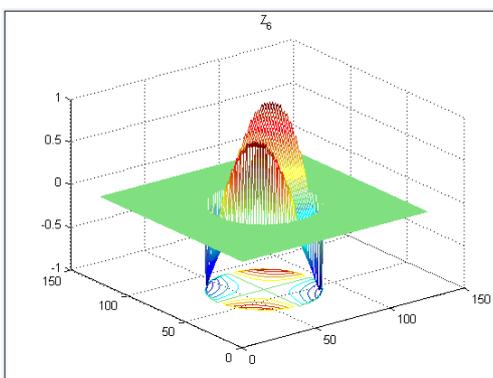
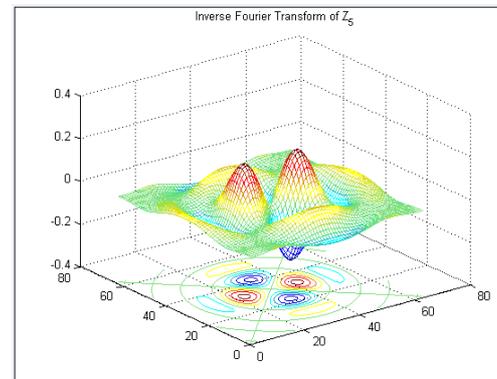
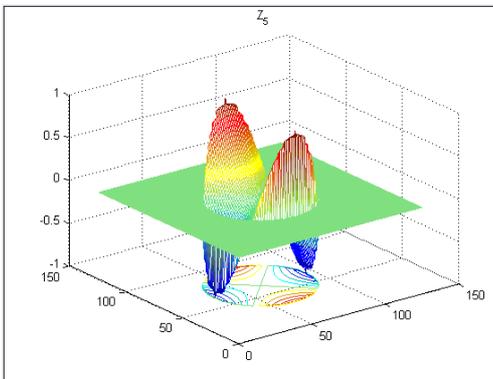
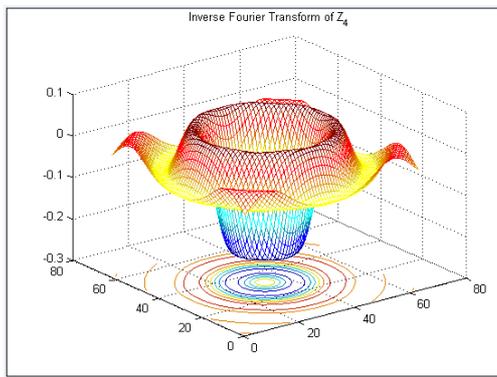
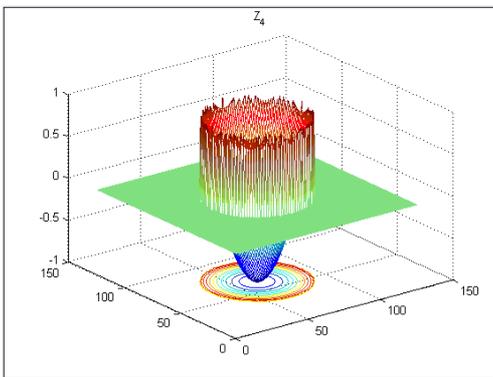
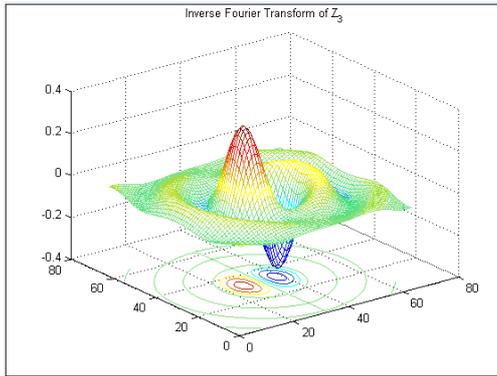
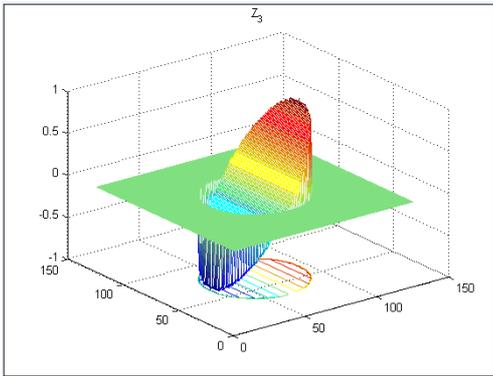
Another type of aberration arises when the light source has a range of wavelengths. For the refractive lens, the refractive indices of glass material vary with wavelength causing imaging to change at different wavelengths. The inability of the lens to focus light over a range of the wavelengths due to the variations of the refractive indices of the lens material is called chromatic aberration. The chromatic aberration has been described in the literature [2-23] [2-24] [2-25].

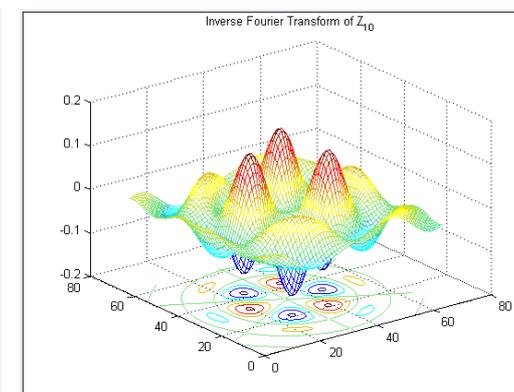
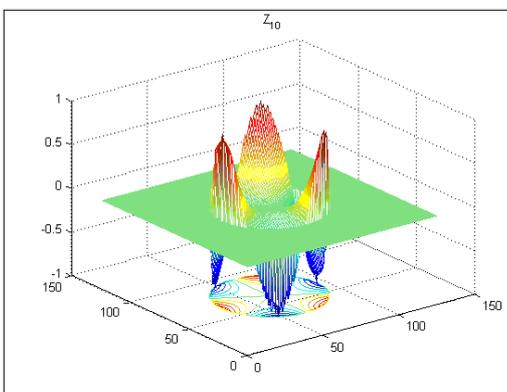
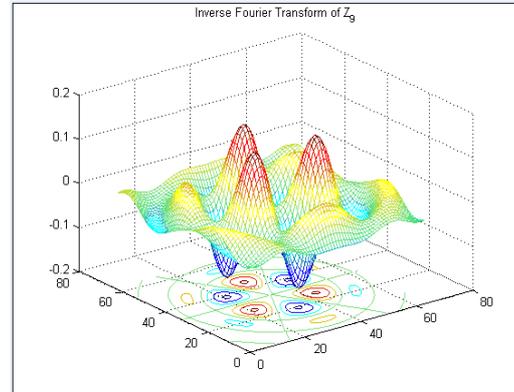
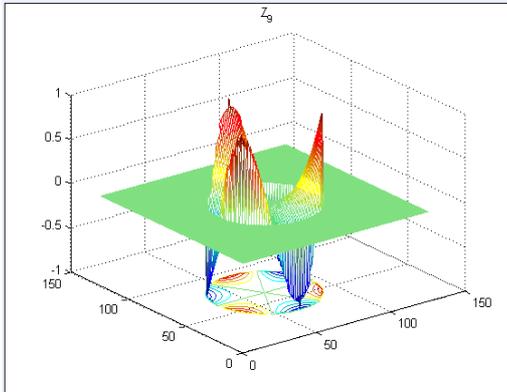
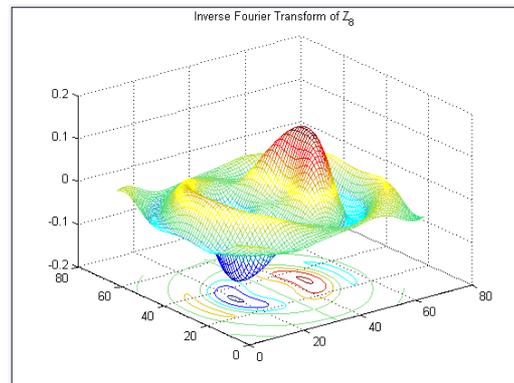
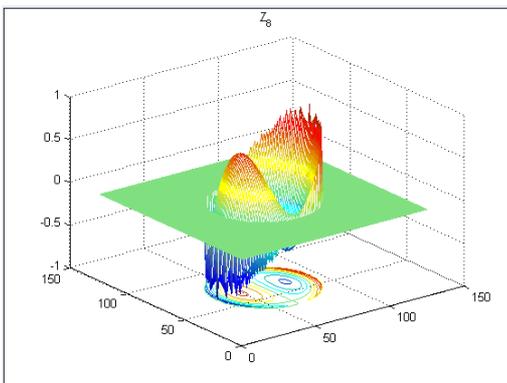
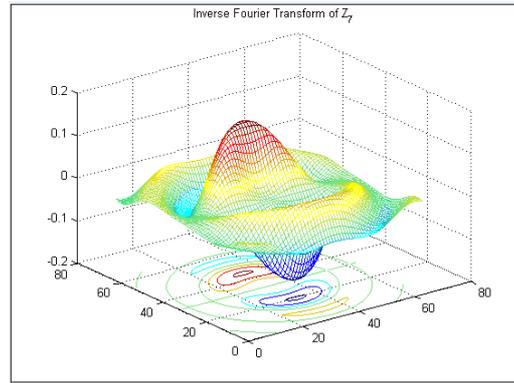
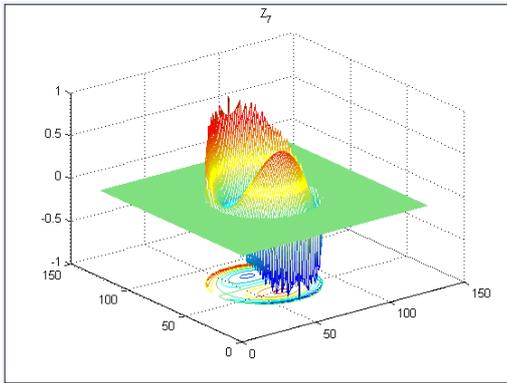
An aberration measurement technique based on observing the printability of side-lobes using knife-edge transition from a clear area to a 6% attenuated and $\lambda/2$ phase-shifted area is demonstrated in reference [2-26][2-27][2-28]. A. R. Neureuther and et al. [2-29] and G. C. Robins [2-45] described that each aberration has a unique effect upon the side-lobe spillover around a point source. For example the coma aberration takes energy from one side of the pattern and throws it to the other side. In this fashion, an interferometric aberration monitor was developed by inversely mapping the point spread function field with Zernike aberrations to locate positions and phase of target components that coherently add electric fields at the center of the probe position. This inverse mapping technique constitutes the basis of the Probe Pattern Grating (PPG) Defocus Monitor in this thesis. Figure 2-3 depicts the pairs of the first eleven Zernike aberrations and their inverse Fourier transform functions. The concept, modeling, simulation and experiments of the PPG defocus monitor are extensively discussed in chapter 3.

The quantity of $n\sin\theta$, as shown in Figure 2-1, is called the numerical aperture (NA) of the lens, where n is the index of refraction of the medium surrounding the lens. The

imaging resolution could be improved by increasing NA. For dry lithography, where the medium is air and $n=1$, decreasing the focal length or increasing the diameter of the lens enlarges NA. However in practice, the focal length is limited by the necessity of moving wafers underneath it without hitting it and the needs to protect the lens from photo-resist outgassing during exposure. Increasing the lens size and using more glass elements make the lens construction much more difficult and expensive. Today the lens cost represents one-third or more of the total exposure system cost.







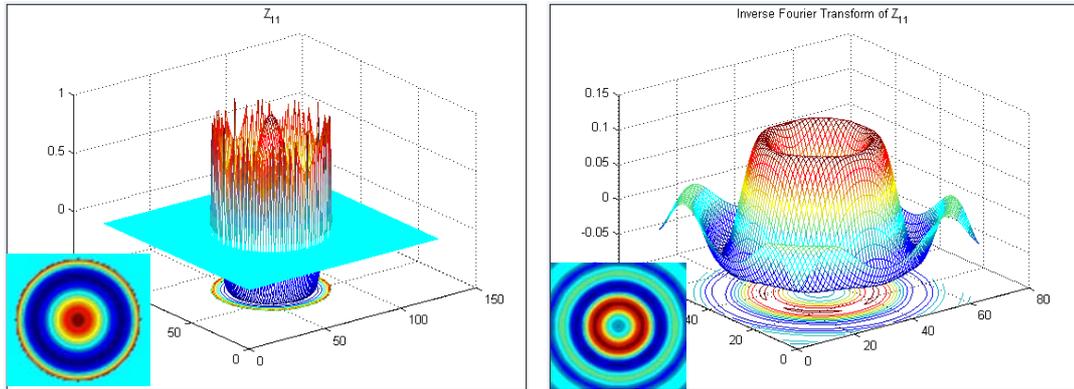


Figure 2-3 Illustration of the first eleven Zernike aberration functions and their inverse Fourier transforms. The two contour plots of Z_{11} are shown for clarification.

The ability to maintain extreme precision of wafer stage position during moving is important during the step and expose sequence. The position of the wafer stage is measured by a modern variant of the Michelson interferometer with great accuracy. The stage control is accomplished by precision motors, bearings, and feedback loops. In order to achieve high stepper throughput, the stages must be capable of moving at high velocity while maintaining good position control. Vibrations present a special challenge during this process. Scanning a reticle stage at four times (assuming 4x lens reduction) velocity of a wafer stage moving in the opposite direction, and keeping the position of the two stages synchronized to an error of less than 10nm, requires the highest degree of active vibration control available. Image blur occurs if the reticle-stage scanning is not synchronized with the wafer-stage scanning [2-30].

2.1.2 Lithography Imaging Theory

Three kinds of image terminologies are generally classified in lithographic community: aerial image, latent image, and resist image. Aerial image is the optical image formed at the wafer plane without considering the resist process effect, latent image is the resist image generated after the exposure and post-exposure bake steps, but before the resist development, and resist image is the final pattern formed in the resist after development. Corresponding to the image definitions, lithography modeling involves three key steps: modeling the aerial image, predicting the photochemical reaction during the resist exposure process and the chemical distributions in the post exposure bake process, and finally calculating the final resist profile after resist development.

A practical theory describing the effects of resist exposure was developed by Dill et al. [2-31] in 1975, and then a series of resist models have been proposed [2-32] [2-33], in which modeling the post-exposure bake in chemical amplified resist is an important and active area of research. An efficient solution to model the photo-acid behavior during the quenching and diffusion process is still one of the main challenges in lithography modeling today.

A fully rigorous model to calculate the optical images involves solving the Maxwell's equations throughout the photolithography system in Figure 2-1. Because of the immensity of the computations, a practical solution uses certain approximations. The degree of approximations impacts the achieved accuracy. Hopkins' Method is one of the major approximate lithographic image modeling methods, which is extensively used in the lithography communities [2-34]. Many simulation programs, including SPLAT [2-

35], PROLITH [2-7], SOLID-C [2-8] and *etc.*, use Hopkins' method to calculate the optical image.

Figure 2-4 shows the imaging process with the on-axis point source. The analysis closely follows reference [2-36] but it is consistent with most of the advanced textbooks in optics. The mask spectrum $O(f, g)$ is defined as the Fourier transform of the mask pattern $O(\xi, \eta)$. The pupil function $P(f, g)$ represents the transmission function in the pupil plane, which is one for the transmitted frequency, and is zero for the blocked frequency.

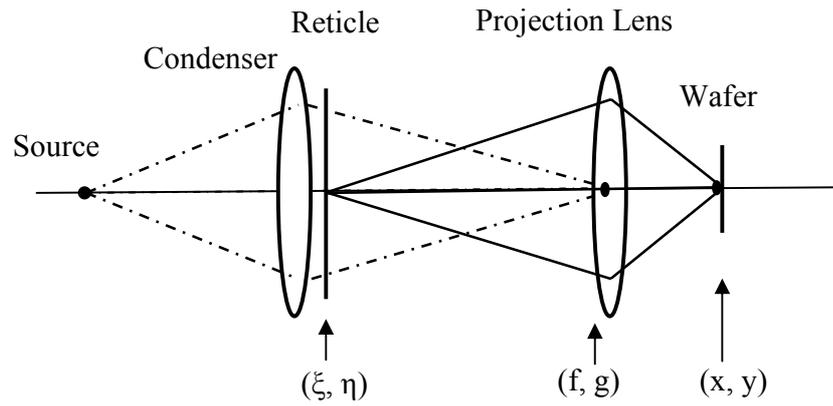


Figure 2-4 Illustration of the imaging process with on-axis point source illumination

Then the field distribution at the wafer image plane is the Fourier transform of the product of the pupil function and the mask spectrum:

$$E(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(f, g) O(f, g) e^{-i2\pi[fx+gy]} df dg \quad \text{Eq. (2-6)}$$

The image intensity is:

$$I(x, y) = |E(x, y)|^2 = \left| \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(f, g) O(f, g) e^{-i2\pi[fx+gy]} df dg \right|^2 \quad \text{Eq. (2-7)}$$

For a circularly symmetric pupil typical in photolithography, the pupil function can be represented by:

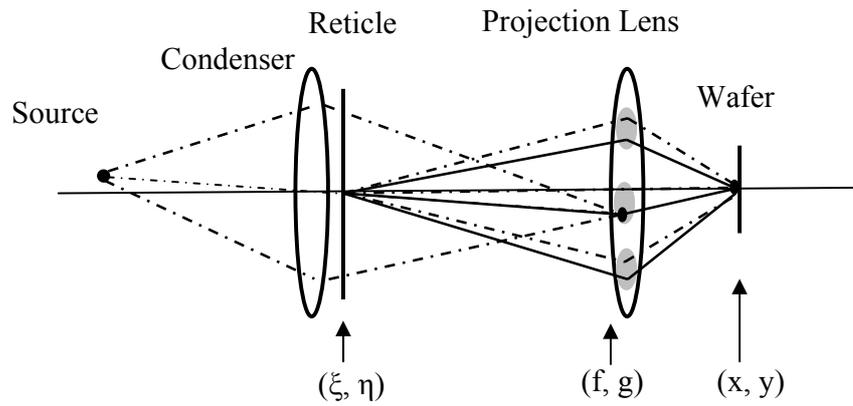
$$P(f, g) = \begin{cases} 1 & \text{if } \sqrt{f^2 + g^2} \leq \frac{NA}{\lambda} \\ 0 & \text{otherwise} \end{cases} \quad \text{Eq. (2-8)}$$

Figure 2-5(a) illustrates an off-axis point source illumination. If we associate the source point with a coordinate (f_s, g_s) corresponding to the location of its image in the pupil, for an off-axis point source, $(f_s, g_s) \neq (0,0)$. An approximation is made that the optical system is shift-invariant such that the spectrum of the reticle due to the source point s is described by a shifted spectrum $O(f - f_s, g - g_s)$. So the image arising from source point s can be described by:

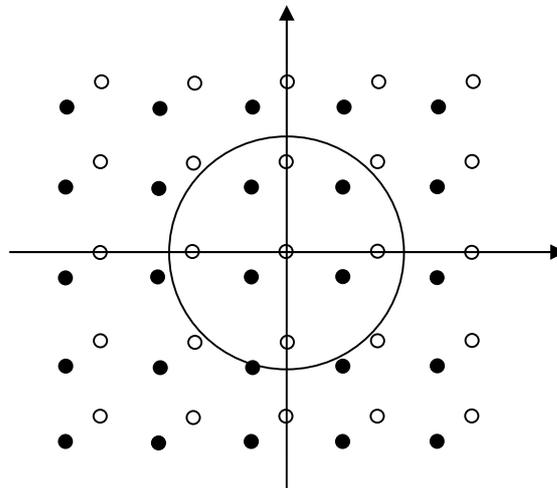
$$I(x, y) = |E(x, y)|^2 = \left| \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(f, g) O(f - f_s, g - g_s) e^{-i2\pi[fx+gy]} df dg \right|^2 \quad \text{Eq. (2-9)}$$

where f_s and g_s represent the angle of the incident ray. The pupil map of the mask spectrum resulting from an on-axis ray and an off-axis ray is shown in Figure 2-5 (b). Some of the diffracted orders captured before in the on-axis illumination become lost in the off-axis illumination, but some of the orders are introduced into the pupil.

The concept of effective source is helpful in visualizing the partial coherent imaging [2-34]. As shown in Figure 2-6, the effective source is the spectrum of the light source in the pupil without including mask diffraction. For a circular illumination, the effective source fills a circle with radius σ in the pupil plane. Assuming equal intensity of the light rays, the effective source can be described by:



(a) Schematic of the imaging process of the off-axis point source illumination



(b) Pupil map with two beam illuminations. The empty circles represent on-axis rays and the filled circles represent off-axis rays.

Figure 2-5 Illustration of the imaging process with off-axis illumination

$$J(f, g) = \frac{1}{\pi\sigma^2} \text{circ}\left(\frac{\sqrt{f^2 + g^2}}{\sigma}\right) = \begin{cases} \frac{1}{\pi\sigma^2} & \text{if } \sqrt{f^2 + g^2} \leq \sigma \\ 0 & \text{otherwise} \end{cases} \quad \text{Eq. (2-10)}$$

where σ is the partial coherence factor. The image with interference pairs of waves traveling with angles (f', g') and (f'', g'') can be formulated in terms of an effective source:

$$I(x, y) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} J(f, g) P(f + f', g + g') P^*(f + f'', g + g'') O(f', g') O(f'', g'') e^{-i2\pi[(f' - f'')x + (g - g'')y]} df dg df' dg' df'' dg''$$

Eq. (2-11)

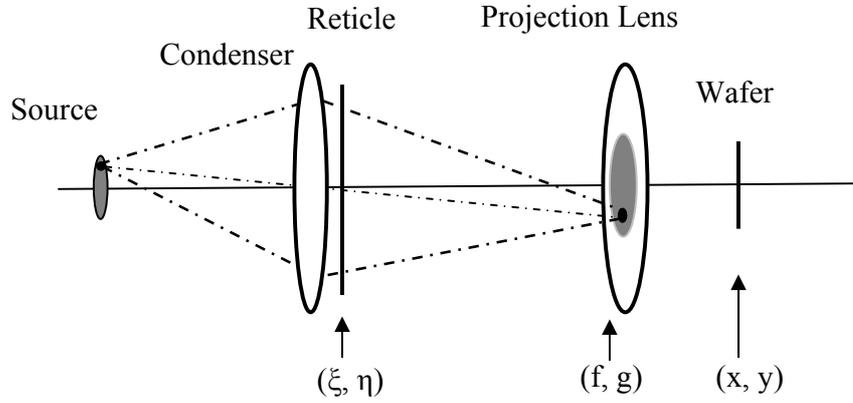


Figure 2-6 Illustration of the effective source in the pupil

The interaction between each pair on the mask is weighted by a factor known as the Transmission Cross-Coefficient (TCC) [2-37]:

$$TCC(f', g'; f'', g'') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} J(f, g) P(f + f', g + g') P^*(f + f'', g + g'') df dg \quad \text{Eq. (2-12)}$$

So Eq. (2-10) can be re-expressed by:

$$I(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} TCC(f', g'; f'', g'') O(f', g') O^*(f'', g'') e^{-i2\pi[(f' - f'')x + (g - g'')y]} df' dg' df'' dg''$$

Eq. (2-13)

TCC is proportional to the overlapping area among the effective source and displaced pupil centered at $(-f', -g')$ and $(-f'', -g'')$, as shown in Figure 2-7. The integration of the shaded overlap area among the three circles is TCC. The significant advantage of the TCC approach is that all illumination and pupil parameters are completely included in TCCs, so that they can be pre-calculated and stored. Then a number of different mask objects can be quickly simulated by Eq. (2-13).

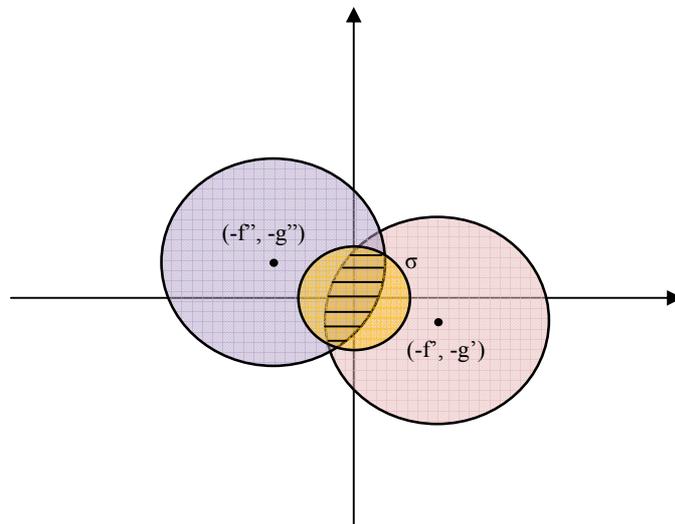


Figure 2-7 Illustration of Transmission Cross-Coefficient (TCC). The yellow circle represents the effective source, the other two circles indicates the shifted pupil pairs centered at $(-f', -g')$ and $(-f'', -g'')$ respectively.

Abbe's method is another important image modeling method [2-38]. The illumination source is assumed to be spatially incoherent, so the image intensity produced from each source point is calculated individually. The intensity contributions from all points that comprise the source are added with a weighting factor proportional to the source intensity distribution. The image intensity is expressed as [2-12]:

$$I(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I_s(\alpha, \beta) \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} K(x, y; \xi_1, \eta_1) K^*(x, y; \xi_2, \eta_2) F(\xi_1, \eta_1; \alpha, \beta) F^*(\xi_2, \eta_2; \alpha, \beta) O(\xi_1, \eta_1) O^*(\xi_2, \eta_2) d\xi_1 d\eta_1 d\xi_2 d\eta_2 d\alpha d\beta$$

Eq. (2-14)

Where (α, β) are the coordinates at the source plane. $F(\xi, \eta, \alpha, \beta)$ and $K(x, y, \xi, \eta)$ are the amplitude spread functions (impulse response) of the illumination and the imaging system respectively. $O(\xi, \eta)$ is the mask amplitude transmittance. Abbe's formulation is conceptually simple, but it does not lend itself to computationally efficient implementation. However, Abbe's method is much easier than Hopkins' method to incorporate a thick mask model, in which at least one rigorous simulation is needed to determine mask scattering. The thick mask model is necessary when the mask edge effect extends to a significant fraction of the feature size. In this case the transitions of the actual fields at the edges, polarization dependent effects due to different boundary conditions, and the lateral cross-mixing of the field components passing through the mask cannot be neglected [2-39].

The simulation program SPLAT, TEMPEST, and Panoramic [2-40] are utilized for this thesis work. Most of the PPG simulations use the Hopkins modeling method, and some of the IAIS on-wafer topography simulations employ Abbe's method. Both thick and thin mask model are used throughout this work, as appropriate.

2.2 Metrology Techniques

Various metrology schemes are required for developing lithographic processes. As feature size continues to decrease, an effective metrology is one of the main challenges in

semiconductor industry. Scanning Electron Microscope (SEMs) and Automatic Force Microscope (AFMs) can deliver direct topographic images of very small features and are hence the common tools to measure line-width and pattern profile. However only a good top-down view of linewidth is available to SEM, and it is difficult to monitor pattern sidewall profile in detail without additional patterning and operation strategies. Sample charging is a significant problem when using SEM for measuring photo-resist patterns. Cross-sectional SEM is capable of measuring feature profiles but still needs help with other tools for the CD measurement. CD-SEM is commonly used in the productions and is effective in measuring CD, but they are unable to extract the thickness values or detailed topographic information, and hence the extracted image information is incomplete. AFM is purely mechanical and can be applied on various materials. However because of probe tip size and shape limitations, AFMs do not provide reliable measurement of tight pitch structures. AFM is also a very slow measurement process. Profilometer is another general pattern profile monitor which used primarily for measuring transitions between various thin films, and shares the similar measurement mechanisms and limitations of AFM. Both SEM and AFM are very expensive, and even destructive, and thus not suitable for run to run monitoring. In this thesis, SEM and Profilometry are the measurement tools used in the IAIS fabrication work frequently.

Electrical measurement has the significant advantage of speed. It can collect considerable data in a small amount of time once the wafer has been patterned and etched. Electrical measurement can also provide accurate information of final CD linewidth, but does not provide pattern profile details. Further, electrical measurement targets

etched conducting material, such as polysilicon or metal, so it cannot be used as an *in-situ* monitor.

Optical metrology is preferred in manufacturing environments because of its low cost, high speed, small footprint, high accuracy, and nondestructive properties. Scatterometry is one of the optical metrologies, which is based on the reconstruction of the diffraction grating profile from its optical diffraction response [2-41]. In this thesis, the PPG defocus monitor developed in chapter 3 is based on scatterometry metrology.

2.2.1 Scatterometry Based Metrology

Two main scatterometry configurations have been extensively investigated. H. Giovannini, and *et al.* use angle resolved scatterometry, where the beam wavelength is fixed but the incident angle is varied [2-42]. A big concern of the single wavelength variable angle approach is that it needs special purpose hardware. Also, the required mechanical motion makes it slow, and it does not seem to support the practical extraction of detailed topographic information from the grating structures. In contrast to variable angle scatterometry, a Specular Spectroscopic Scatterometry scheme was proposed by researchers at UC Berkeley to measure intensity and phase of the 0th order diffraction at a fixed incident angle and multiple wavelengths [2-43] [2-41]. Coupled with a very efficient RCWA (Rigorous Coupled-Wave Analysis) implementation, this method is adequate for detailed reconstruction of the periodic grating profile. Specular Spectroscopic Scatterometry can make direct use of conventional spectroscopic ellipsometer and spectroscopic reflectometer systems.

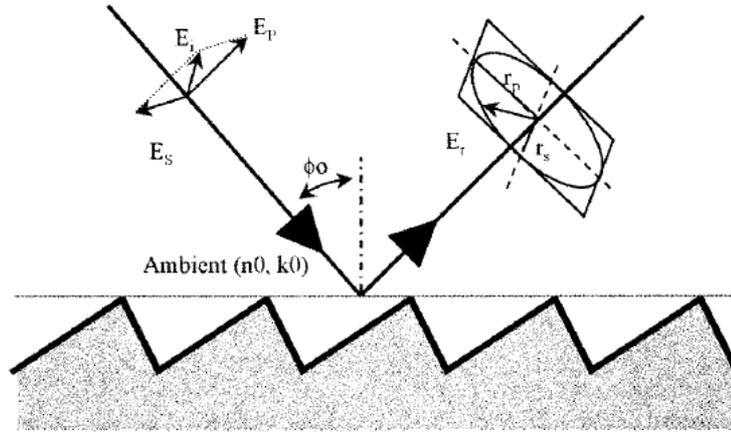


Figure 2-8 Specular Spectroscopic Scatterometry for 1D gratings [2-41]

The principle for Specular Spectroscopic Scatterometry is shown in Figure 2-8 [2-41]. The ratio of the 0th order complex TE and TM reflectivity: $\rho = r_{p,0} / r_{s,0} = (\tan \Psi)e^{i\Delta}$ is measured, where $r_{p,0}$ is the 0th-order TM reflectance coefficient and $r_{s,0}$ is the 0th-order TE reflectance coefficient. $\tan \Psi$ is the amplitude of ρ . The real part of the phase term $\cos \Delta$ is used for analysis. The ratio of TE and TM provides additional sensitivity and robustness over just measuring either TE or TM. There is a variety of ways to configure the hardware of a spectroscopic ellipsometer, but most result in the measurement of the ellipsometric parameters which are based on the two independent variables Ψ and Δ [2-44]:

$$\begin{aligned}
 N &= \cos(2\Psi) \\
 S &= \sin(2\Psi) \sin \Delta \\
 C &= \sin(2\Psi) \cos \Delta
 \end{aligned}
 \tag{2-15}$$

The parameters N, C and S are not independent since $N^2 + S^2 + C^2 = 1$. The challenge for the scatterometry simulation software is to handle various dependent variables by different optical tools. Spectroscopic Ellipsometry measures parameter of Ψ and Δ , then

the simulation software generally maps the measured values to the fully bounded parameters N , S , C . Timbre Technologies' Optical Digital Profilometry (ODP) [2-46] is a scatterometry based metrology solution software, and is the main simulation platform of the thesis work described in Chapter 3. Timbre ODP uses the three parameters of N , S , C as the measurement metrics.

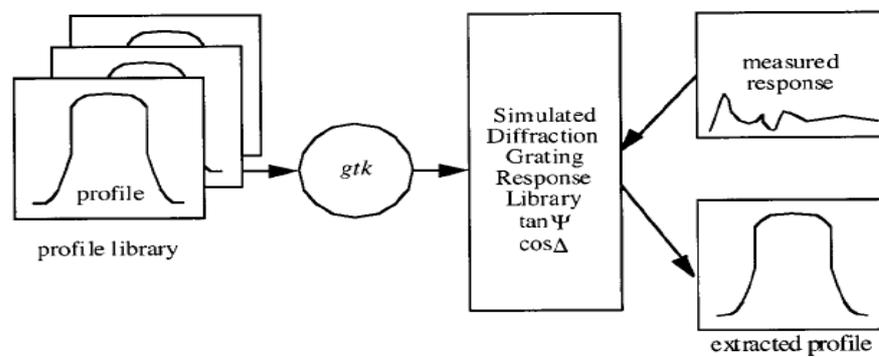


Figure 2-9 A library-based methodology for CD profile extraction [2-41]

The objective of the simulation program is to find a profile whose simulated diffraction response matches the measured response. Because of the numerical complexity involved in directly solving for the simulated profile in real-time, this problem is usually addressed by pre-calculating a large library of profiles and their respective scatterometric response. Figure 2-9 shows the extraction flow of Berkeley Specular Spectroscopic Scatterometry [2-41], where *gtk* (grating tool-kit) is a high performance implementation of RCWA. Achieving high sensitivity over a range of different profiles is one of the main objectives of *gtk*, and that requires the efficient generation of large library that can be searched in real time during production metrology. The uniqueness of the solution may be an issue of

this approach, since it is possible that two different profiles may lead to similar diffraction response. This problem will be further addressed in the context of our application in Chapter 3.

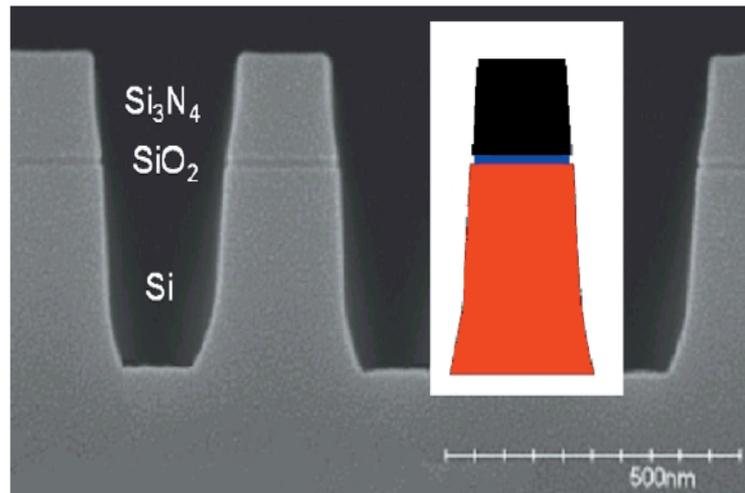


Figure 2-10 ODP measured shallow trench isolation structure profile after etch and clean as compared to cross-sectional SEM of the same grating profile [2-47]

The scatterometry based metrology provides full pattern profile information, which significantly facilitates the PPG defocus monitor experimental work in this thesis. It is also fast, repeatable, non-destructive and offers good accuracy and precision, suitable for *in-situ*¹ deployment. It also has the benefits of fast, repeatable and non-destructive measurements with good accuracy and precision, suitable for *in-situ* metrology. Figure 2-

¹ As of today, most commercial applications of spectroscopic scatterometry are based on off-line, dedicated ellipsometers or reflectometers. In addition, small-footprint ellipsometers and reflectometers have been developed by Tokyo Electron, Nanometrics and others, and integrated into various tools, such as wafer tracks and multi-chamber etchers, in an *in-line* fashion. *In-situ* deployment is also conceptually possible, and has been demonstrated, even though, as of this writing, it is not commercially available.

10 shows an example of ODP measured shallow trench isolation structure after etch as compared to cross-sectional SEM [2-39]. ODP is used to measure silicon trench depth, stack trench depth, trench CD at different locations, trench profile, rounding, barrier oxide notching, and sidewall slope, all through a single inline measurement of several seconds per site. To fully characterize these parameters using traditional methods, several metrology tools are needed, including a profiler for trench depth, a thin film tool for film thickness, a CD-SEM for CD, a cross-section SEM for profile and a side wall angle.

2.2.2 On-Wafer Sensors

The autonomous, wireless On-Wafer Sensor concept was first proposed and sensor prototypes were developed by researchers at UC Berkeley [2-48] [2-49], and then a startup company OnWafer Technologies [2-50] was funded focusing on novel sensing methods for the semiconductor industry. This on-wafer sensor is a paradigm shift where a complete metrology system migrates from dedicated tools to the surface of a wafer that can be handled with regular material handling robotics in a production fab.

The major advantages of on-wafer sensor were discussed in reference [2-49]. First of all, it improves process characterization and design. Traditional process and equipment development procedures involve processing so-called dummy wafers, measuring the finished wafer parameters, adjusting the machine parameters, and then repeating the process. Since this is the post-process measurement, only the composite effect of multiple steps can be measured. Each iteration can take a number of days, and the entire process can take several months. By using a real-time, direct measurement scheme, far fewer

iterations could be performed because much more information is gathered during each iteration.

Second, to calibrate the process, traditionally, dummy wafers are put into the production flow every so often, and measurements from these processed wafers are used to monitor and adjust the process. It has been estimated that 15-18% of the production capacity is wasted on process characterization. By placing the autonomous sensor into the process chamber for the real-time monitoring, process parameters can be optimized automatically in a single step, and far fewer test wafers would be needed, thereby improving the useful throughput of the production line.

Third, in a typical fabrication facility machine, downtime accounts for over 10% of the total equipment time. Traditional diagnosis of equipment problems involves venting and disassembling the process chamber to place wired sensors on the wafer-chuck, which can take several hours. By instead using an *in situ, portable* sensor that can be loaded into the tool just like a regular wafer, possible sources of the problem can be eliminated or confirmed much more rapidly.

Finally, some measurements are difficult or destructive to measure using traditional techniques. Two examples of this type are etched sidewall profile and lithography aerial image. Because of the nature of the on-wafer sensors, they conceptually offer the opportunity to measure some of these quantities *in situ*. This would greatly increase the ability to evaluate the process and develop the novel monitoring techniques.

The first commercial on wafer sensors are temperature-sensor wafers for plasma-etch equipment and for relatively low temperature, high precision (<150°C) bakeplates, such as the ones used for post-exposure bake (PEB). Today's commercial on-wafer sensors

have been much improved since the early laboratory prototypes tested in the Berkeley Microfab.

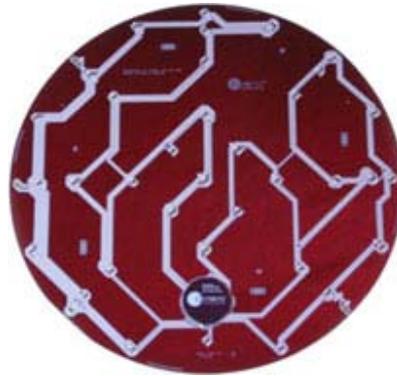


Figure 2-11 Wireless temperature sensor wafer from OnWafer Technologies [2-50]

Figure 2-11 shows one commercial wireless temperature sensor wafer from OnWafer technologies [2-50]. The sensor wafer has an array of temperature sensors mounted on the wafer surface to spatially resolve PEB temperature and on-board memory and communication module to store and transfer temperature measurement data. The wireless sensor wafer supplies temperature measurement with $<0.05^{\circ}\text{C}$ absolute precision, and $<0.05^{\circ}\text{C}$ repeatability [2-51]. And additional plasma-oriented sensors have been marketed for measuring the uniformity of the RF field.

In this thesis, we utilize the on-wafer sensor concept, and explore the feasibility of constructing an on-wafer aerial image sensor for lithography characterization, as described in Chapter 4.

References

- [2-1] H. J. Levinson, Principles of Lithography, *SPIE*, 2005
- [2-2] W. Oldman, S. nandgaonkar, A. Neureuther, and M. O'Toole, "A general simulator for VLSI lithography and etching process: part I- application to projection lithography", *IEEE Trans. on Electr. Dev.*, ED-26, pp. 717, 1980
- [2-3] M. M. O'Toole, "Simulation of Optically Formed Image Profiles in Positive Photoresist", Ph. D. Dissertation, UC Berkeley, 1979
- [2-4] K. Toh, "Two-Dimensional Images with Effects of Lens Aberrations in Optical Lithography", *Proc of SPIE*, vol. 4066, pp.376, 2000
- [2-5] A. Wong and A. neureuther, "Examination of polarization and edge effects in photolithography masks using three-dimensional rigorous simulation", *Proc of SPIE* 2197, pp521, 1994
- [2-6] A. R. Neureuther, "Lithography simulation using the internet", *Microolithography world*, pp.18, 1999
- [2-7] C. A. Mack, "PROLITH: a comprehensive optical lithography model", *Proc. of SPIE*, vol 538, pp 207, 1985
- [2-8] www.sigma-c.de
- [2-9] www.mentorgraphics.com
- [2-10] www.synopsys.com
- [2-11] <http://www.almaden.ibm.com/st/chemistry/lithography/immersion/immersion.png>
- [2-12] J. W. Goodman, "Statistical Optics", *John Wiley & Sons, Inc.*, 1985
- [2-13] R. Voelkel, and et al., "Microlens array imaging system for photolithography", *Optical Engineering* 35, pp.3323, 1996
- [2-14] C. Pierrat, A. Wong, And S. Vaidya, "Phase-shifting mask topography effects on lithographic image quality", *Electron Devices Meeting*, Dec., 1992
- [2-15] M. A. Miller, A. R. Neureuther, and et al., "Characterization and Monitoring of photomask edge effects", *Proc. of SPIE*, vol.6730, 67301U, 2007
- [2-16] T. A. Brunner, S. Cheng, A. E. Norton, "A stepper image monitor for precise setup and characterization", *Proc. of SPIE*, vol 922, pp.377, 1988

- [2-17] D. Goloni, M. DeMarco, W. Kordonski, and J. Bruning, "MRF polishes calcium fluoride to high quality", *Laser Focus World*, pp. S5, 2001
- [2-18] H. Sewell and T. Friedman, "High resolution imagery: the matching of optical and Resist system in the mid-UV", *Proc. of SPIE*, pp328, 1988
- [2-19] online discussions of aberrations
- [2-20] T. A. Brunner, "Impact of lens aberrations on optical lithography", *IBM J. RES. DEVELOP.*, Vol 41, No. 1/2, 1997
- [2-21] V. N. Mahajan, "Zernike Circle Polynomials and Optical Aberrations of Systems with Circular Pupils", Eng. & Lab. Note, in *Opt. & Phot. News* 5, S-21-S-24, 1994
- [2-22] J. R. Sheats, B. W. Smith, "Microlithography: Science and technology", CRC Press, 1998
- [2-23] T. Brunner, D. Corliss, "Laser bandwidth and other sources of focus blur in lithography", *J. Microlith, Microfab, Microsyst*, vol. 5(4), 043003, 2006
- [2-24] P. D. Bisschop, I. L. F. Trinchouk, "Impact of finite laser bandwidth on the critical dimension of L/S structures", *J. Micro/Nanolith, MEMS MOEMS*, vol 7(3), 033001, 2008
- [2-25] C. P. Ausschnitt, T. A. Brunner, "Distinguishing dose, focus and blur for lithography characterization and control", *Proc. of SPIE*, Vol. 6520, 65200M-1, 2007
- [2-26] K. Hayano and *et al.*, Proceedings IEICE Technical Report, SDM98-163, 1998
- [2-27] H. Fukuda and *et al.*, "Determination of high-order lens aberrations using phase/amplitude linear algebra", *J. Vac. Sci. Technol B*. 17(6), pp3318, 1999
- [2-28] A. Imai and *et al.*, "Lens aberration measurement technique using attenuated phase-shifting mask", *Proc. of SPIE* Vol. 4000, 2000
- [2-29] A. R. Neureuther, S. Hotta, K. Adam, "Modeling defect-feature interactions in the presence of aberrations", *Proc. of SPIE*, Vol. 4186, pp405, 2001
- [2-30] G. de Zwart and *et al.*, "performance of a step-and-scan system for DUV lithography", *Proc. of SPIE*, vol 3051, pp817, 1997
- [2-31] F. H. Dill and *et al.*, "Characterization of positive photoresist", *IEEE Trans. on Electr. Dev.* ED-22(7), pp. 445, 1975

- [2-32] H. Fukuda, and *et al.*, "Impact of Acid/Quencher Behavior on Lithography Performance", *Proc. of SPIE*, Vol. 4346, pp. 319, 2001
- [2-33] K. Kattori, and *et al.*, "The Accuracy of Simulation Based on The Acid-Quencher Mutual Diffusion Model in KrF Process", *Proc. of SPIE*, Vol. 4691, pp. 1243, 2002
- [2-34] H. Hopkins, "On the diffraction theory of optical images", *Proc. Roy. Soc. A.* vol.217, pp408, 1953
- [2-35] K. Toh, "Two-Dimensional Images with Effects of Lens Aberrations in Optical Lithography", M. S. Thesis, UC Berkeley, 1988
- [2-36] A. K. Wong, "Resolution Enhancement Techniques", *SPIE*, 2001
- [2-37] M. Born and E. Wolf, "Principle of Optics", *Pergamon Press*, section 10.4, pp505, 6th ed., 1980
- [2-38] E. Abbe, "Beitrage zur Theories des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv f. Mikroskopische Anat*, vol.9, pp. 413 -468, 1873
- [2-39] K. Adam, "Domain Decomposition Methods for the Electromagnetic Simulation of scattering from three-dimensional structures with Applications in Lithography", Ph. D. Dissertation, UC Berkeley, 2001
- [2-40] <http://www.panoramictech.com/>
- [2-41] X. Niu, N. Jakatdar, J. Bao, and C. J. Spanos, "Specular Spectroscopic Scatterometry", *IEEE trans. on Semiconductor Manufacturing*, Vol. 14, No. 2, pp.97, 2001
- [2-42] H. Giovannimi, and et al., "Angle-resolved polarimetric phase-measurement for the characterization of gratings", *Optics letters*, vol. 21, No 20., pp.361, 1997
- [2-43] X. Niu, "An Integrated System of Optical metrology for Deep Sub-Micron Lithography", Ph. D. Dissertation, UC berkeley, 1999
- [2-44] K. Lensing and et al., "A Comprehensive comparison of Spectral Scatterometry Hardware", *Proc. of SPIE*, Vol.5752, pp.337, 2005
- [2-45] G. Robins, "Interferometric Pattern and Probe Based Aberration Monitors", Ph. D. Dissertation, UC Berkeley, 2005
- [2-46] <http://www.tel.com/>
- [2-47] K. Barry and et al., "Optical Digital Profilometry: A Breakthrough Technology for

Non-Destructive Inline Profile and CD Metrology", Timbre Technologies Inc, Semiconductor Fabtech, 15th edition

- [2-48] M. Freed, "in situ Etch Rate Sensor Arrays for Plasma Etch Processes", UC Berkeley, May 1999
- [2-49] M. Freed, M. Kruger, C. J. Spanos, and K. Poolla, " Autonomous On-Wafer Sensors for Process Modeling, Diagnosis, and Control", *IEEE trans. on Semiconductor Manufacturing*, Vol. 14, No, 3, pp.255, 2001
- [2-50] www.onwafer.com
- [2-51] Q. Zhang, "IC manufacturing Performance Enhancement through Advanced Process/Equipment Modeling, Control and Metrology", Ph. D. Dissertation, UC Berkeley, 2006

Chapter 3

Probe Pattern Grating Defocus Monitor Based on Scatterometry Metrology

As CMOS device sizes continue to decrease, the numerical aperture of lithography scanner increases, and therefore the depth of focus (DOF) continues to shrink. Since the pattern transfer is very sensitive to focus variations, the focus must be maintained to about 0.5 of a Rayleigh Unit (RU) of defocus, and monitored to about 0.1 RU. Therefore an accurate and quick focus metrology method is increasingly important to focus control. In this chapter, we introduce a new highly sensitive Scatterometry based Probe-Pattern Grating (PPG) Defocus Monitor. The monitor translates the focus error into the probe line trench depth, which can be measured by commercial spectroscopic scatterometry techniques. The transparent lines spaced at the strong focus spillover distance from the centerline of a 90 degree phase-shifted probe line are placed in the monitor to achieve its high sensitivity. In section 3.1 and section 3.2, we overview the defocus effects on lithographic images and the defocus monitor techniques. In section 3.3, we describe the design of the defocus monitor. The concept of the parameter sensitive monitor is introduced in section 3.3.1, the design of the Probe Pattern Grating (PPD) defocus monitor is shown in section 3.3.2, the aerial image simulation is shown in section 3.3.3, and the resist simulation is shown in section 3.3.4. In section 3.4, we describe the printing assessment and scatterometry calibration results of the PPD defocus monitor. The conclusions are presented in section 3.5.

3.1 Review of Defocus Effects on Lithographic Images

The effect of defocus in lithography is to cause an error in curvature of the actual wavefront relative to the desired wavefront. As shown in Figure 3-1, the solid curve is the un-aberrated wavefront, and the dashed curve is the defocus wavefront. The distance between the desired wavefront and the defocused wavefront is zero at the center of the exit pupil, and increases as approaching to the edge of the pupil.

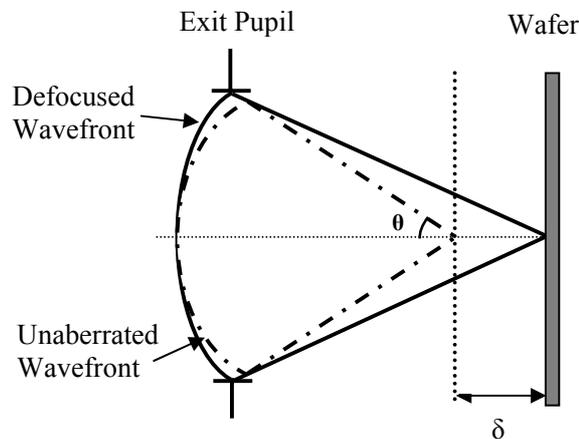


Figure 3-1 Illustration of defocus effect in lithography

The optical path difference (OPD) is a function of the defocus distance δ and the position within the pupil. Describing the position within the exit pupil by an angle θ , the OPD is given by:

$$OPD = \delta(1 - \cos\theta) \quad \text{Eq. (3-1)}$$

In the case that the zero and the first diffraction orders have 90° phase difference, the two orders would not interfere with each other and no pattern would be formed. So in order to generate the image, the tolerable OPD is $\lambda/4$, as expressed in Equation (3-2).

$$OPD_{\max} = k_2 \frac{\lambda}{4} \quad \text{Eq. (3-2)}$$

Where $k_2 < 1$. Depth of Focus (DOF) is defined as the range of the defocus distance that can be tolerated.

$$DOF = 2\delta_{\max} = 2 \cdot k_2 \frac{\lambda}{4(1 - \cos \theta)} = k_2 \frac{\lambda}{2(1 - \cos \theta)} \quad \text{Eq. (3-3)}$$

Equation (3-3) is applied to all small line and space patterns. Lord Rayleigh, who developed the now-famous Rayleigh criteria for resolution and depth of focus [3-1], essentially looked at the resolution limit: the smallest pitch that can be printed would put the first diffraction order at the largest angle that could pass through the lens, defined by the numerical aperture, NA, and assumed that lens NA is relatively small. Then Equation (3-3) can be specified as Rayleigh DOF:

$$DOF = k_2 \frac{\lambda}{NA^2} \quad \text{Eq. (3-4)}$$

The restrictions of the conventional Equation (3-4) are that the relatively low numerical apertures imaging a binary mask pattern of lines and spaces at the resolution limit, while Equation (3-3) is an exact ODP expression without the low NA restriction. Equation (3-3) can be further modified to account for immersion lithography of high NA Rayleigh criteria:

$$OPD = n_{\text{fluid}} \delta(1 - \cos \theta)$$

$$DOF = k_2 \frac{\lambda}{2n_{\text{fluid}}(1 - \cos \theta)} \quad \text{Eq. (3-5)}$$

For a diffraction limited image system, the impulse response of the point spread function is the Fraunhofer diffraction pattern of the exit pupil [3-2]. When wavefront errors exist, the pupil function can be expressed as: $P(\xi, \eta) = P(\xi, \eta) \exp[jkW(\xi, \eta)]$,

where $P(\xi, \eta)$ is the un-aberrated pupil function, $W(\xi, \eta)$ is the effective phase error, and $k=2\pi/\lambda$. So the electric field of the point spread function of the aberrated wavefront is:

$$h(\mu, \nu) = \iint P(\xi, \eta) \exp[jkW(\xi, \eta)] \exp\left\{-j\frac{2\pi}{\lambda z}[\mu\xi + \nu\eta]\right\} d\xi d\eta \quad \text{Eq. (3-6)}$$

where (μ, ν) are the image plane coordinates, (ξ, η) are the pupil plane coordinates, and z is the distance between pupil and image planes. For the small amount of aberrations, the aberration phase term can be expressed as a Taylor series of: $1 + jkW(\xi, \eta)$, thus the electric field at the image plane is a superposition of the ideal field and an aberration induced perturbation field:

$$h(\mu, \nu) = \iint P(\xi, \eta) \exp\left\{-j\frac{2\pi}{\lambda z}[\mu\xi + \nu\eta]\right\} d\xi d\eta + jk \iint P(\xi, \eta) W(\xi, \eta) \exp\left\{-j\frac{2\pi}{\lambda z}[\mu\xi + \nu\eta]\right\} d\xi d\eta \quad \text{Eq. (3-7)}$$

By including the Zernike polynomial of aberrations in the second term of Equation (3-7), $\sum a_j w_j(\xi, \eta)$ [3-3], where a_j and $w_j(\xi, \eta)$ are the aberration coefficient and function of j th term, the point spread function field becomes:

$$\begin{aligned} h(\mu, \nu) &= h_0 + \sum a_j h_{a,j} \\ &= F[P(\xi, \eta)] + \sum jk F[P(\xi, \eta)W(\xi, \eta)] \cdot a_j \end{aligned} \quad \text{Eq. (3-8)}$$

Where F represents Fourier transform. The first term in Equation (3-8) is the original point spread function field and the second term is the aberration induced electric field perturbation. By calculating the perturbation field with even and odd aberrations separately, we obtain that even aberrations produce imaginary field, and odd aberrations create real field. This conclusion indicates that a pattern of phase 90° or 270° will be co-phase and sensitive to the defocus electric field perturbations.

T. A. Brunner and *etal.*, proposed the concept of “focus blur”, which encompasses the effect of laser bandwidth chromatic aberrations, vertical stage vibrations, and stage tilts [3-3]. A modified Lorentzian form is used to fit the laser bandwidth related focus blur distribution, and a Gaussian form is to model the axial vibration. The overall focus blur is the convolution of the three components.

Defocus leads to a quadratic OPD surface which tends to degrade the image contrast, edge slope, patterns fidelity, and resolution [3-4]. Defocus variation always combined effects of many problems: wafer nonflatness and topography, lens heating, chuck nonflatness, mask tilt, across slit/scan focus tilt, mask-to-wafer scan synchronization, source wavelength aberration, and process variations in the resist coating, post-exposure bake, and development process. The variation in exposure dose is always accounted for with defocus variations. The process window, which describes the defocus and dose space with acceptable lithography quality, is the standard way to measure the robustness of the process to variations.

3.2 Review of Defocus Monitor Techniques

A number of defocus monitor techniques have been introduced using advanced imaging techniques in the past. Phase Shift Focus Monitor (PSFM) [3-5], developed by T. A. Brunner et.al, converts the focus error into overlay shifts in the printed pattern. The work on Phase shift Masks (PSM) has revealed that non-180 shifters cause image asymmetries through focus. Simulations of a narrow line with no shifter on the left and 90° shifter on the right, as shown in Figure 3-1, indicates that with positive defocus, the image shifts toward the 90° shifter, and with negative defocus, the image shifts away from the shifter.

The image shifting is roughly proportional to focus in the interesting region near best focus. This effect can be utilized in various focus monitor patterns, including line-in-line structure and box-in-box structure [3-6], in which the latter structure is designed to be consistent with the registration measurement used on the lithography tools.

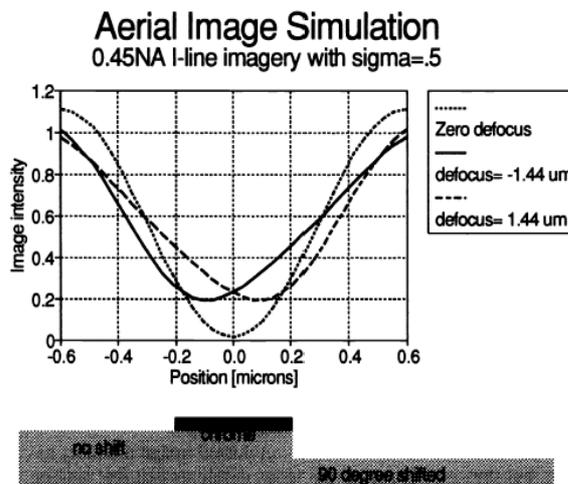
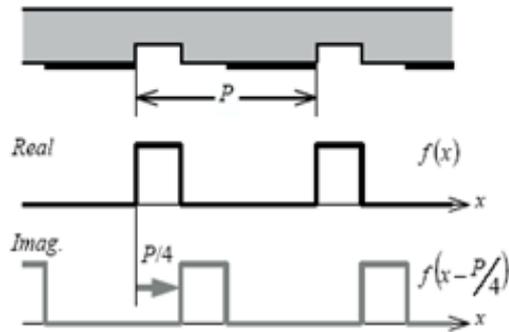
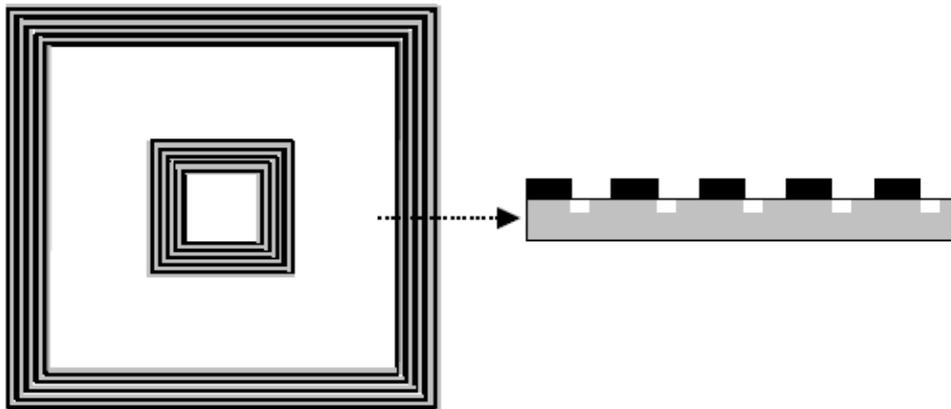


Figure 3-2 Aerial Image simulation of PSFM for 3 different focus positions [3-5]

The repeated Phase Shift Grating (r-PSG) [3-7], or Phase shift Grating Focus Monitor (PGFM) [3-8], is based on the same idea of converting focus errors into alignment errors and measuring them with an overlay measurement tool. However, the isolated lines with large 90° patches used for PSFM are replaced with phase grating, which generate the equivalent of a two-beam interferometer. The r-PSG (PGFM) grating patterns comprising opaque line, naked line and 90° phase-shift groove with their width ratio equivalent to 2/1/1, makes either of the first order rays of diffraction disappear entirely. The other first-order ray interferes with the zeroth-order ray to form interference fringes with the ability of moving proportional to a defocus.



(a) Basic structure of r-PSG. Zero diffraction is realized at the diffraction order of -1, because the imaginary vector is in the direction opposite the real vector [3-7].



(b) Schematic of the PGFM. Black lines represent Cr and gray lines are 90° phase trenches etch in the glass [3-8].

Figure 3-3 The Basic structures of repeated Phase Shift Grating focus monitor (r-PSG) and Phase Grating Focus Monitor (PGFM).

The r-PSG and PGFM structures are shown in Figure 3-3. Since the interferometer insides the lens is very sensitive to focus errors, the r-PSG (PGFM) has more than twice as sensitive to defocus as PSFM, it thus better suited to the measurement of tighter focus distribution.

Since focus effects always dependent on dose, the focus-exposure matrix is mostly used to judge a process and achieve a good model fit of the test patterns. Some early efforts [3-9] to fit such data had involved the polynomials in focus (F) and exposure (E) with the expression: $CD = \sum_{i=0}^M \sum_{j=0}^N a_{ij} E^i F^j$. C. A. Mark and J. D. Byers improved the model by

offering the physical meaning of the fitting coefficients [3-10]. They assumed a thin and infinite contrast photo-resist will be removed when the dose exceeds the threshold dose E_{th} , and verified that a CD variation with one over exposure dose is more physically accurate than assuming that CD is proportional to powers of dose. This approximation leads to a model form: $CD = \sum_{i=0}^M \sum_{j=0}^N a_{ij} (1 - \frac{E_s}{E})^i F^j$. A. P. Ausschnitt and A. Y. Cheng [3-

11] analyzed the focus offset S_t between the top and bottom CDs, and assumed high order polynomial terms ($N > 2$, $M > 4$, and $N+M > 4$), focus asymmetry terms (a_{i1}) and the dose-defocus coupling term (a_{12}) are negligible, then the model can be simplified into two

$$\text{equations of bottom and top CDs: } CD_{bot} \approx CD_{0b} + a_{10b} \left[\left(1 - \frac{E_0}{E}\right) + \frac{a_{20}}{a_{10}} \left(1 - \frac{E_0}{E}\right)^2 \right] + a_{02b} F^2 ,$$

$$\text{and } CD_{top} \approx CD_{0t} + a_{10t} \left[\left(1 - \frac{E_0}{E}\right) + \frac{a_{20}}{a_{10}} \left(1 - \frac{E_0}{E}\right)^2 \right] + a_{02t} (F^2 - 2S_t F + S_t^2) .$$

In the case that the model simplified assumptions are valid, e.g. the pattern is an isolated pattern, the above two equations can be inverted to solve the unique dose and focus. K. Lensing and *etal.* [3-12] pointed out that the above equations can be only applicable to patterns that exhibit parabolic response to dose, but a model structure that accommodates a variety of lithography system, would require higher polynomial orders to adequately fit the sample data. They found that the (4, 1) model is most satisfactory to the process, and then they

have two 10-parameter polynomial expansions for the top and bottom CDs. This model can be inversely solved to estimate the focus and dose by numerical optimization techniques. The inversion of 10-parameter model returns eight complex solutions. The solutions closest to the command dose and focus should be considered the correct value.

With the knowledge that isolated patterns are generally more sensitive to defocus variations than dense patterns, and the images without high diffraction orders have poor exposure latitude and therefore high sensitivity to dose changes, T. A. Brunner and C. P. Ausschnitt [3-13] developed a Process Monitor Grating (PMG) by using assist features to zero out higher diffraction orders and keep the 0 and ± 1 orders in the image formation, which is shown in Figure 3-4. At the same time, they included the focus blur into the CD expressions, such that the resulting response of the profile is the convolution of the focus blur distribution and the general profile expression described in [3-14]. To solve the three parameters of dose, focus and blur, three scatterometry measures, the top and bottom CD of PMG patterns, and the bottom CDs of a dense near-isofocal pattern, are required. An 11 parameter analytic model is reversely solved to obtain three unknowns. The PMG is an improved binary mask design, which is more sensitive to process variations than the general binary patterns. However, the PMG sensitivity is limited by its intrinsic pattern properties whose electric field is orthogonal to defocus variations, and the isolate pattern design reduces the sensitivity of the scatterometry system.

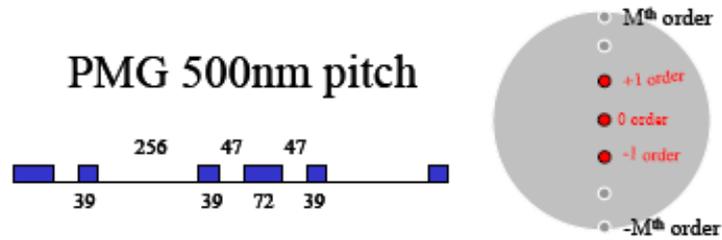


Figure 3-4 Layout of Process Monitor Grating PMG 500.

A class of interferometric pattern and probe-based aberration monitors were developed by G. C. Robin [3-16] and A. R. Neureuther, S. Hotta, and K. Adam [3-15]. The monitors are based on the principle of coherent electromagnetic spillover among transmitting phase shifted regions. The mathematical theory shows that the optimum patterns for measuring aberrations are the two-dimensional inverse Fourier transform of the original aberration functions expressed as Zernike polynomials. Therefore each monitoring target consists of a sub-resolution probe circle surrounded by a number of patterns rings. The ring spills electric field to the probe in an amount linearly proportional to the given aberrations. The monitor has been characterize in AIMS and photo-resist exposure, and shows great promise in lithography and vision science applications. However, the observations of the two-ring defocus target show discrepancies between the experimental behavior and the simulations. The theoretical sensitivity and linearity through focus were not observed during the experiment due to the weakness of the probes. An efficient measurement method to calibrate and quantify the monitor resist behavior is needed.

In this chapter, we develop a Probe Pattern Grating (PPG) defocus monitor by extending the interferometric aberration monitor concept into one dimensional grating design, and using the advanced metrology technique of scatterometry measurement. The PPG has

higher sensitivity to defocus than most of the phase shift mask patterns and the binary monitors, and at the same time has sufficiently high sensitivity to the scatterometry system. A simple linear model is build to describe the defocus response, which significantly reduces the cost of the numerical computing in the scatterometry based process monitor techniques.

3.3 Design of Probe Pattern Grating (PPG) Defocus Monitor

3.3.1 Concept of Probe Pattern Grating (PPG) Defocus Monitor

For a given aberration, a two dimensional pattern sensitive to that aberration is generated by inverse-mapping the aberrated pupil to give the location and phase of the target pattern at the mask plane. The mask pattern then creates an aberration compensated wavefront, which will interact appropriately with the specific wavefront aberration. G. C. Robin described this two dimensional pattern design in [3-16]. Figure 3-5 emphasizes the inverse mapping concept for the parameter sensitive mask design.

The example of the mask design optimized for defocus sensitivity is given in Figure 3-6. The side lobe positions around the center in Figure (c) provide the positions of the patterns on the mask. The first side-lobe is around $0.57 \lambda/NA$ distance to the center. While the inverse mapping of the defocus Optical Path Difference (OPD) over the unit pupil circle gives the completed two dimensional mask pattern, the projection of the inverse Fourier transform onto the horizontal axis will generate the grating pattern which samples the pupil along horizontal axis at the most sensitive position to defocus. Figure

3-7 shows the 1D defocus optical path difference (ODP) at the pupil plane with the amplitude normalized to one. The spherical aberration, another typical even aberration, is shown in the graph for a comparison.

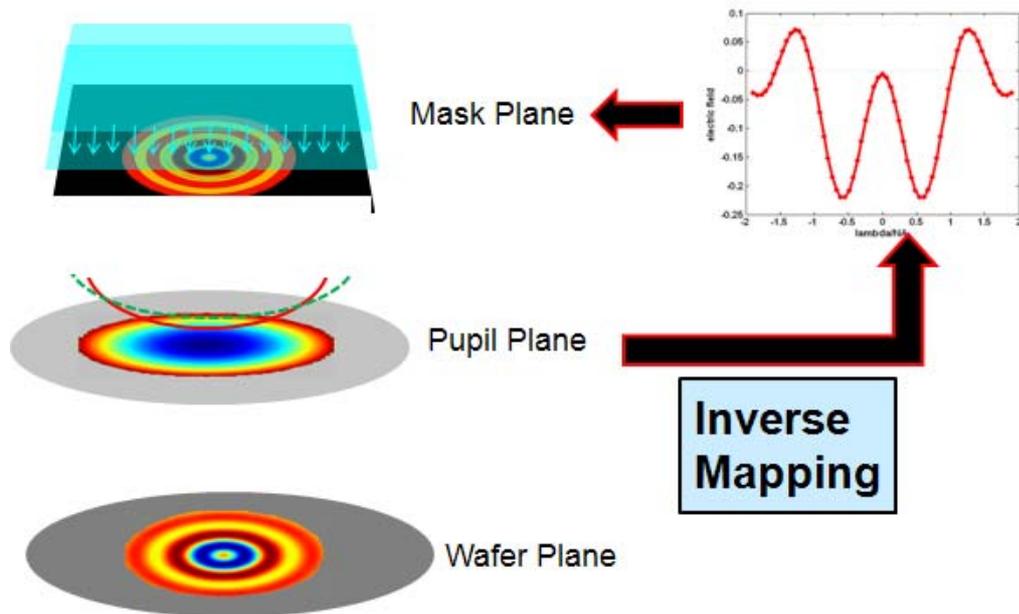


Figure 3-5 Illustration of inverse mapping process of parameter sensitive mask design, referring to [3-16].

For the on-axis illumination, at which the center of the pupil has no aberration effects, the most sensitive point to defocus is at the edge of the pupil, around 1NA position, while the spherical aberration has no effects at the pupil edge. The most sensitive point to spherical is at 0.7NA. This pupil map illustrates that if we sample the signal at the pupil edge, the highest sensitivity to the defocus aberration and a good orthogonality to the spherical aberration can be achieved. Based on this observation, we design the pitch of the grating patterns that will be suitable as a defocus monitor at the mask plane.

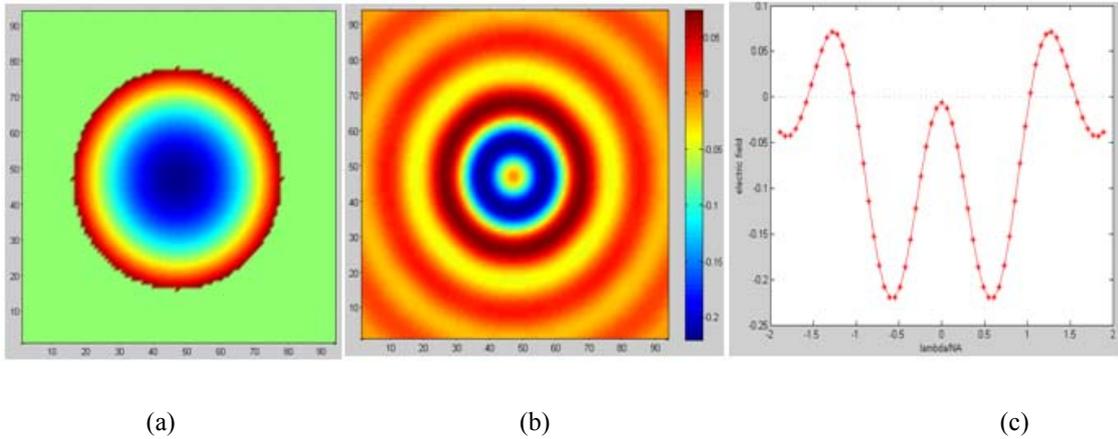


Figure 3-6 Illustration of the defocus mask design through the inverse mapping process: (a) Optical path difference of defocus at the pupil plane; (b) Inverse Fourier transform of the defocus pupil to give the location and phase of the mask design; (c) Cut-plane along x direction of (b). The first side lobe is at $0.57 \lambda/\text{NA}$ to the center.

Figure 3-8 shows the projection of the inverse Fourier transform of the defocus OPD function onto horizontal axis. The first side lobe of the curve is at $\pm 0.43\lambda/\text{NA}$, which indicates that the pitch of the grating is $0.9\lambda/\text{NA}$. We also include the projection of the inverse Fourier transforms of the spherical and piston aberrations in Figure 3-7. The first side-lobe of spherical is at $\pm 0.80\lambda/\text{NA}$, so the pitch of the spherical sensitive grating is around $1.6\lambda/\text{NA}$. The piston curve has 90° phase difference than the defocus electric field spillover. It is note that the electric field spill-over properties keep the same as that of the two dimensional design. The center of the inverse Fourier transform is zero. However, the optimized design parameters for the one dimensional grating is changed as described in Figure 3-8.

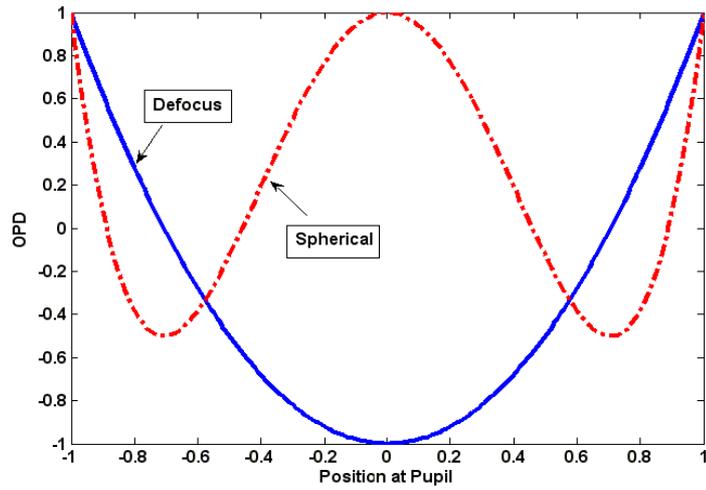


Figure 3-7 1D defocus and spherical optical path difference (ODP) at the pupil plane.

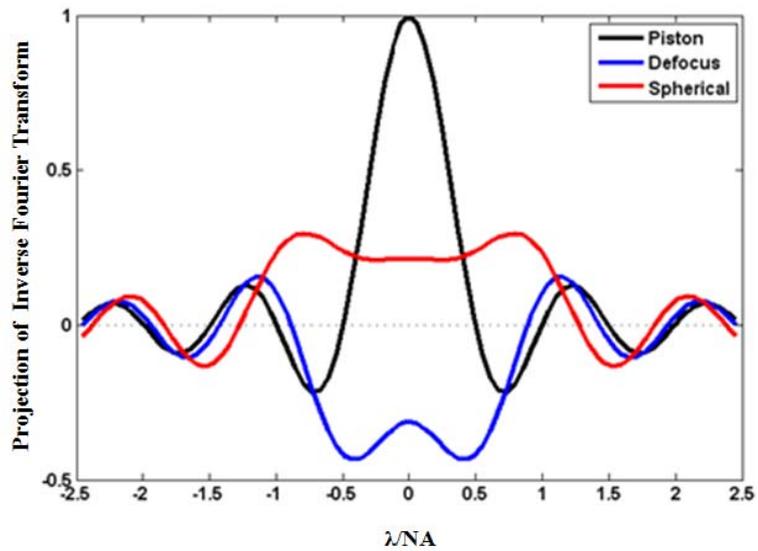


Figure 3-8 Projection of inverse Fourier transform of the defocus ODP function onto horizontal axis. The intensity is normalized so that the Piston peak intensity is one.

On the other hand, in order to obtain at least two order wave interference, the pitch of the grating needs to satisfy:

$$P > \frac{\lambda}{NA(1 + \sigma)} \quad \text{Eq. (3-9)}$$

For the fully coherent illumination, the minimum pitch of the grating is $1\lambda/NA$. As partial coherence increases to 0.1, the minimum pitch becomes $0.91\lambda/NA$. When the partial coherence equals to 0.34, the minimum pitch is $0.746\lambda/NA$. In order to realize the defocus sensitive grating design, first, the grating pitch must be larger than the minimum pitch from Equation (3-9); Then, the pitch should be close to the distance of the defocus electric field spillover side-lobe; Third, the grating pitch will be tuned by considering the partial coherence effect, which will influence the defocus monitor sensitivity since part of the first order ray may be out of the pupil filter.

3.3.2 Design of the Probe Pattern Grating (PPG) Defocus Monitor

The PPG defocus monitor is a periodic grating, composite of a 90° phase shift probe line and a non-phase shift pattern line. Figure 3-9 shows the design of the PPG defocus monitor. Since the defocus aberration induces the 90° electric field spillover, and the probe line is a 90° phase shift grating, the aerial image electric field and intensity at the probe line will be modulated by the defocus aberration. Consequently, the energy inside the resist and the resist profile at the probe line will change according to the defocus variation.

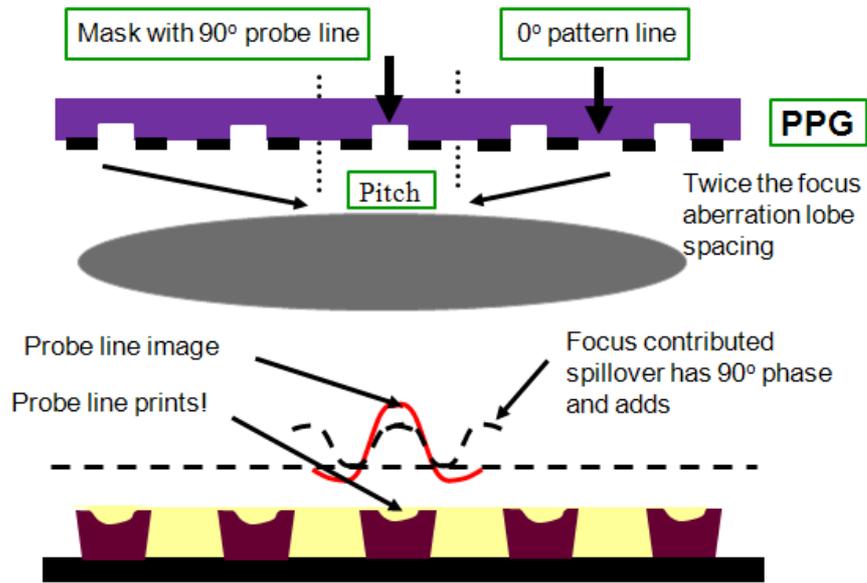


Figure 3-9 Illustration the design of the PPG defocus monitor.

3.3.2.1 Optimizing the Reference Line Width of the PPG

The width of the probe line is designed to make the center intensity of the probe line around 25% CF (Clear Field) without aberration. For the small grating opening of lateral width: $w \leq 0.4 \lambda/NA$, the peak intensity [3-17] is approximated by:

$$I = \sqrt{8.5 \times \left(\frac{w}{\lambda/NA}\right)^4} \quad \text{Eq. (3-10)}$$

The intensity I is normalized to the clear field intensity. Therefore, if the reference intensity is 25%CF, the line width is around $0.29\lambda/NA$. The isolated probe line simulation is done in SPLAT with monopole on-axis illumination of $\sigma=0.1$, and $\lambda/NA = 0.5/0.5$. Figure 3-10 indicates that when the probe line width equals to $0.26\lambda/NA$, the aerial image intensity is around 25%CF.

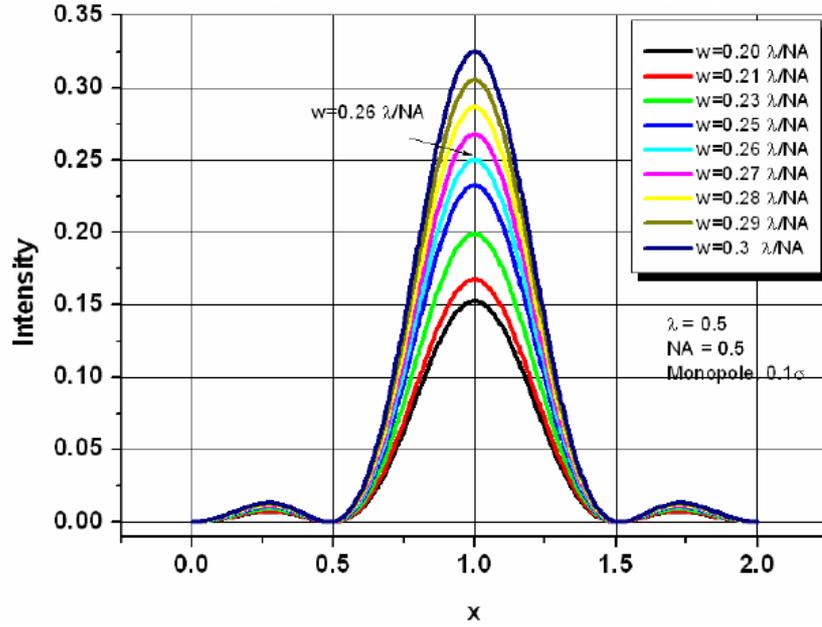


Figure 3-10 Aerial image simulation of the isolated probe line.

For the periodic grating, the optical proximity effect is considered. The partial coherence of the illumination will impact the proximity effect of multiple openings on the mask. The degree of coherence μ_{12} describes the correlation of the electric-field disturbances at the two points on the mask. The mutual coherence is therefore a function of the distance and phase between the two points. The value of the mutual coherence is given by the inverse Fourier transform of the wave angle and magnitude that make up the illumination. For the circular or “top hat” illumination, the mutual coherence is given by [3-18] as:

$$\mu_{12} = \frac{2 \times J_1(2\pi \cdot r \cdot \sigma)}{(2\pi \cdot r \cdot \sigma)} \quad \text{Eq. (3-11)}$$

Where J_1 is the first order Bessel function of the first kind, r is the distance between two points on the mask normalized by λ/NA , and σ is the partial coherence of the

illumination. Equation (3-11) is plotted in Figure 3-11 for four partial coherence coefficients, in which $\sigma=0.1$ represents the approximately fully coherent illumination, $\sigma=0.34$ is the real PPG experimental partial coherence in the ASML 193nm scanner, $\sigma=1$ is an example of the fully incoherent illumination, and $\sigma=0.5$ is the intermediate partial coherence value.

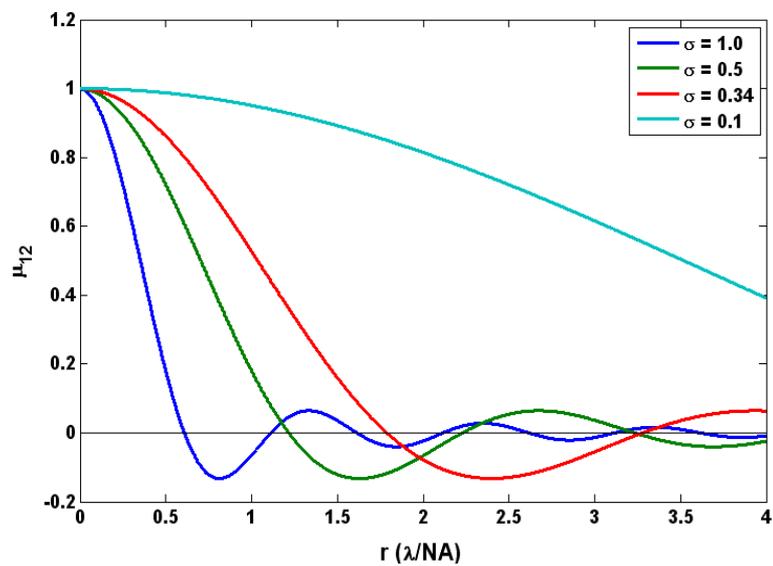
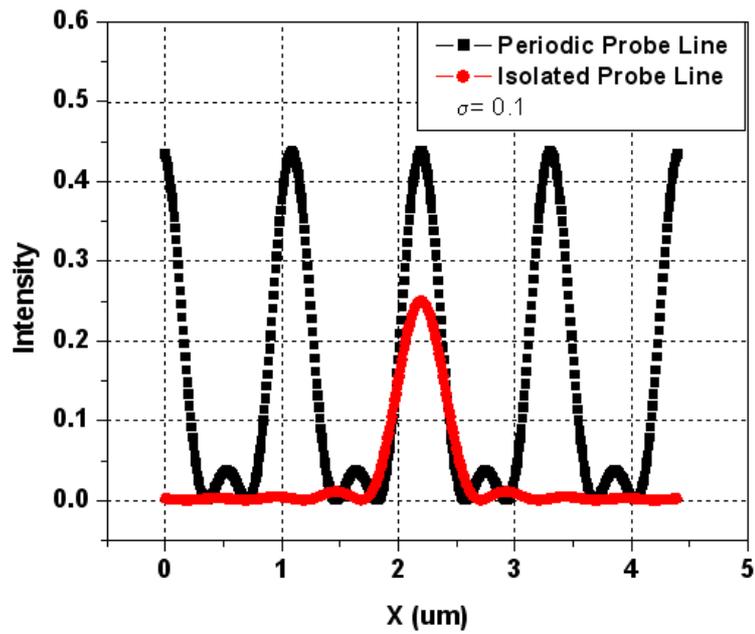


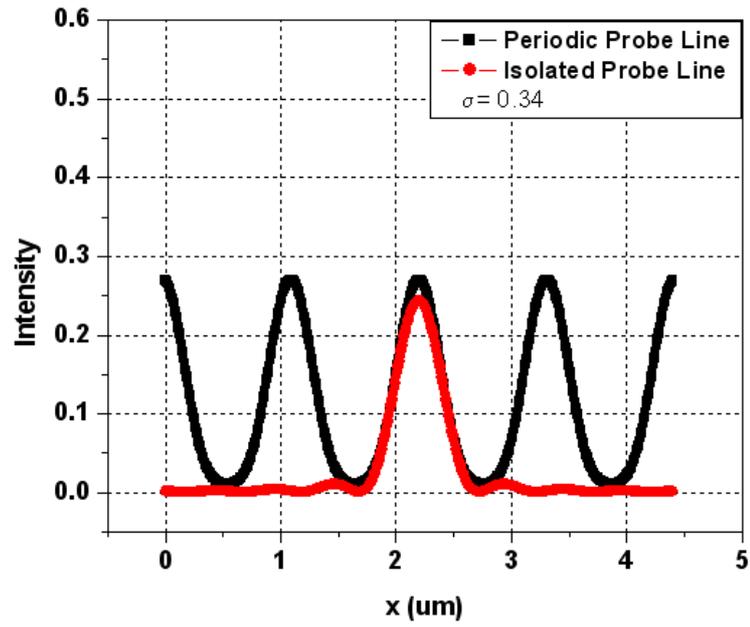
Figure 3-11 Mutual Coherence Function of circular illumination for four partial coherence coefficients

For any partial coherence coefficient, the degree of coherence has the same shape except for the scaling of the x-axis by $1/\sigma$. The plots fall to zero when $r = 0.61 \frac{\lambda}{\sigma NA}$. So the periodic grating has the stronger proximity effect at the smaller partial coherence. For example, when the pitch of the grating equals to $1.11\lambda/NA$, the mutual coherence of the four partial coherence of 1.0, 0.5, 0.34, and 0.1, is about 0, 0.08, 0.44, and 0.94 respectively. The grating proximity effects can be seen in Figure 3-12. At the coherence

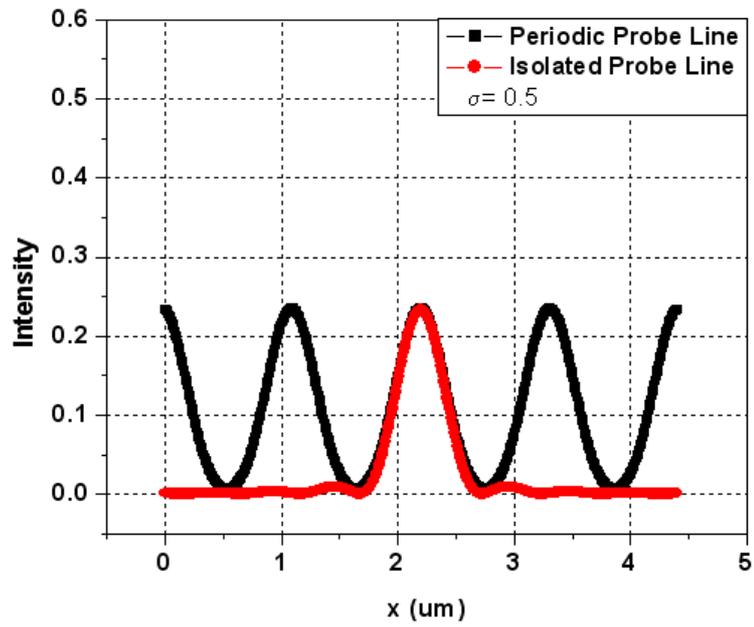
illumination of $\sigma=0.1$, as shown in Figure 3-12 (a), the periodic probe grating peak intensity is 1.75 times larger than the isolated line due to the electric field spill over from the neighbored lines. When σ increases to 0.34, the periodic line intensity is about 1.11 times larger than the isolated line. However, when σ increases to 0.5, the intensity difference between periodic and isolated lines becomes insignificant. Based on the partial coherence effect on the proximity interactions, we can design the PPG probe line width at different illumination conditions. Figure 3-13 lists the periodic probe line width at six partial coherence conditions.



(a)



(b)



(c)

Figure 3-12 Aerial image simulation of periodic and isolated probe lines at (a) $\sigma=0.1$ (b) $\sigma=0.34$ and (c) $\sigma=0.5$

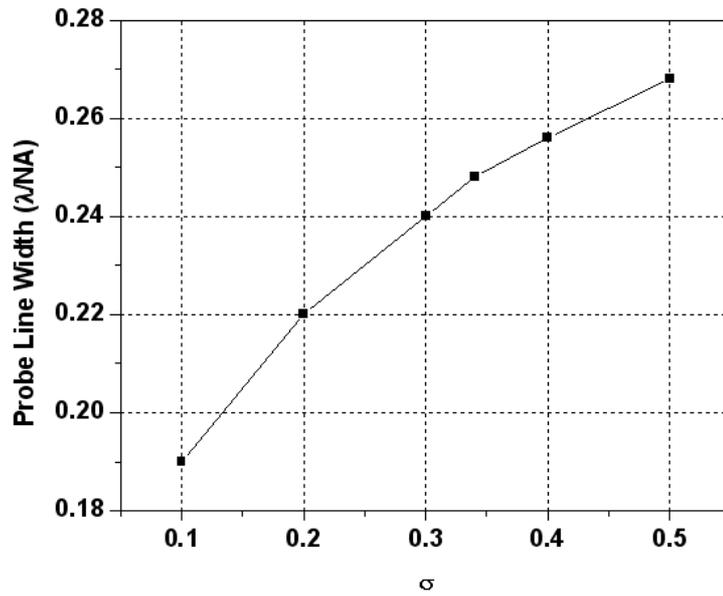


Figure 3-13 PPG probe line width at around 0.25CF intensity at six partial coherence conditions with probe grating pitch of $1.11 \lambda/NA$

On the other hand, we expect the pattern line to print well into the resist at best focus condition. The reference intensity of the pattern line is about 80%CF. The simulation of the isolated pattern line is performed in SPLAT, as shown in Figure 3-14. The simulation condition is the same as that of the probe line simulation. Figure 3-14 indicates that at the 80%CF intensity, the pattern line width is around $0.52 \lambda/NA$. Similarly, the proximity effect of the periodic pattern line is simulated and the line width at partial coherence from 0.1 to 0.5 is shown in Figure 3-15.

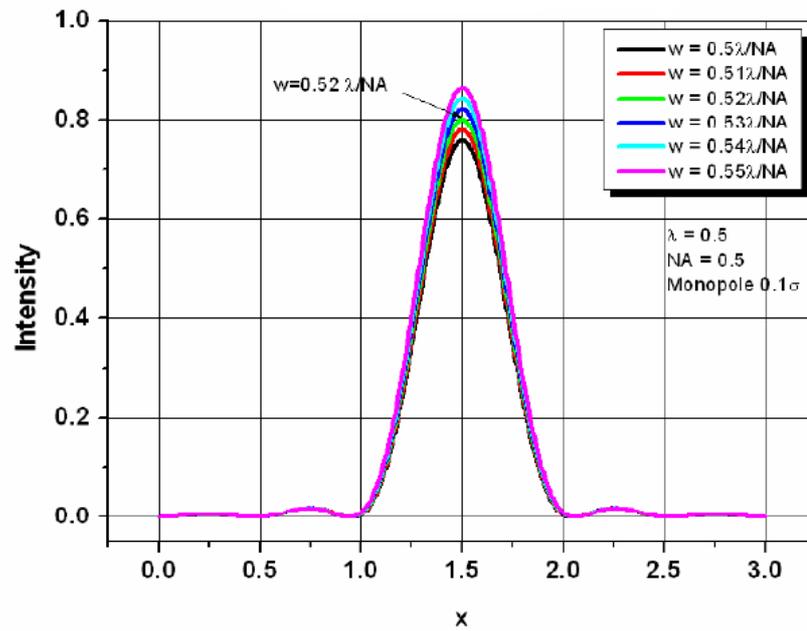


Figure 3-14 Aerial image simulation of isolated pattern line.

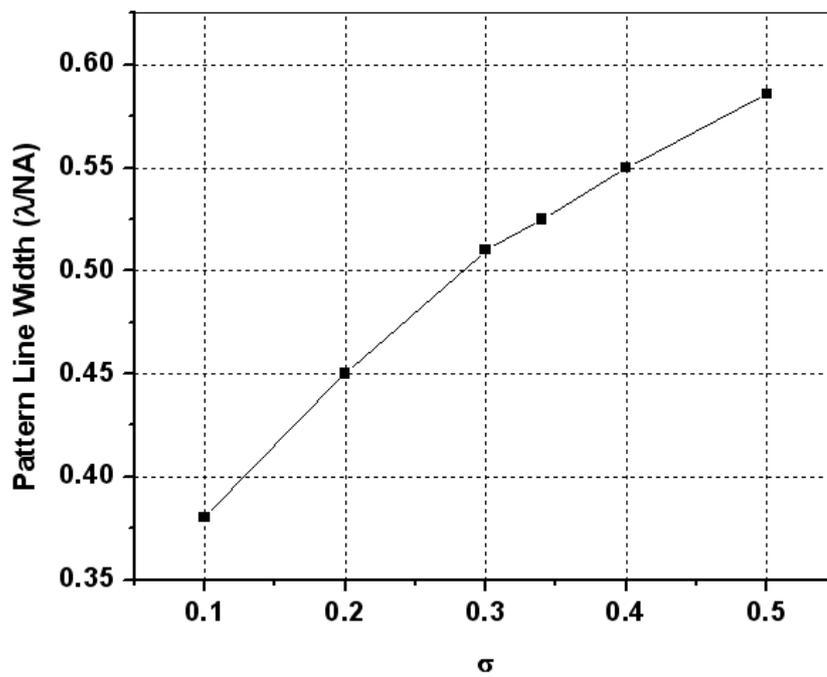


Figure 3-15 Pattern line width at around 0.8CF intensity at six partial coherence conditions with pattern grating pitch of $1.11 \lambda/NA$

Next, the proximity effect between the probe line and the pattern line of the PPG pattern is calculated. The PPG is composed of the 90° probe line and the 0° pattern line with pitch $1.11\lambda/NA$, as shown in Figure 3-9. The peak intensity difference before and after combining the probe grating and the pattern grating is computed. From table 3-1, the pattern line intensity gain due to electric field spillover is less than that of the probe line. As partial coherence increases, the proximity effect becomes less significant to the pattern line, because mutual coherence and hence the electric field interaction decrease.

Table 3-1 Proximity effects (aerial image intensity increases as a portion of clear field intensity) between probe and pattern lines as a function of partial coherence

ΔI \ σ	0.1	0.2	0.3	0.34	0.4	0.5
Probe	0.04721	0.02798	0.02369	0.02476	0.0289	0.04108
Pattern	0.02401	0.01726	0.01213	0.0107	0.00871	0.00652

The maximum electric field spill-over to the pattern line is $\sqrt{0.024} = 0.155$. The intensity increase compared to 0.8CF reference is about 3%, which can be ignored in the design. However, the maximum proximity field to the probe line is $\sqrt{0.04721} = 0.217$ with intensity increase about 19% at $\sigma = 0.1$. The electric field spill-over is $\sqrt{0.02476} = 0.157$ at $\sigma = 0.34$ and its intensity increase is about 10% of 0.25CF. As σ increases continually, since the chrome dimension between the pattern and probe line is only $0.3\lambda/NA$, insufficient resolution causes the contrast to decrease and the intensity at the probe position to increase accordingly. The electric field interaction effect to the probe line hence must be considered in the design. We can either adjust the probe line width as

shown in Table 3-2, to match the reference intensity level, or include the proximity effect in the PPG resist simulation and experimental analysis.

Table 3-2 PPG Probe line width and pattern line width at the reference intensity levels. Probe line reference intensity is 0.25CF. Pattern line reference intensity is 0.8 CF. The line width unit is λ/NA

Width \ σ	0.1	0.2	0.3	0.34	0.4	0.5
Probe line	0.17	0.206	0.228	0.234	0.24	0.243
Pattern line	0.38	0.45	0.51	0.525	0.55	0.586

In the above analysis, we assume that the pitch is a constant of $1.11\lambda/NA$, however the proximity effect becomes more significant as the pitch decreases. Conversely, the electric field interactions between the gratings are reduced as pitch increases. In summary, the reference grating width is a function of the reference intensity level and the partial coherence of the illumination conditions. At the same time, the chrome size must meet the mask making ability and the resolution limitation. The phase shift mask edge effect [3-19], will also affect the effective line width of the opening of air. In order to approximately compensate for the edge effect, the line width of the reference design is increased by about 10nm at each edge.

3.3.2.2 Optimizing the Pitch of the PPG

The PPG pitch determines the sensitivity of the defocus monitor. As discussed in section 3.3.1, the PPG pitch must follow the image generation rule at first, then the pitch should be close to the defocus electric field spillover side-lobe, and it should be tuned by the

partial coherence of the system. As shown in Figure 3-8, the first side lobe of the defocus electric field spill-over is at $0.45 \lambda/NA$, therefore the detector pitch is around $0.9\lambda/NA$. Considering the circular illumination of the two partial coherence conditions, $\sigma = 0.1$ and $\sigma = 0.34$, from equation 3-1, the minimum pitch for $\sigma = 0.1$ is $0.91\lambda/NA$ and the minimum pitch for $\sigma = 0.34$ is $0.746 \lambda/NA$.

We define the sensitivity as the intensity variation at the probe line position divided by the defocus variation of one Rayleigh Unit. The sensitivity plot for various grating pitches at $\sigma = 0.1$, $\lambda = 193\text{nm}$ and $NA = 0.85$, is drawn in Figure 3-16. Since at the minimum pitch of $0.91\lambda/NA$, the available exposure energy is almost zero, we use the minimum pitch of $0.95\lambda/NA$ instead. The probe line is $0.17\lambda/NA + 20\text{nm}$, about 59nm , and the pattern line is $0.38\lambda/NA$, about 86nm respectively. When the pitch is smaller than or equal to $1\lambda/NA$, as shown by pupil map (a) and (b) in Figure 3-16, more than half or half of the first order ray is out of the pupil, the sensitivity is weakened due to loss of the exposure energy.

As pitch increases to $1.1\lambda/NA$, as shown in pupil map (c), the entire first order comes into the pupil and is right at the edge of the pupil, the sensitivity reaches its peak. As the first order ray away from the pupil edge, defocus sensitivity decreases accordingly. Pupil map (e) indicates one of these conditions. When the pitch becomes larger than $1.8 \lambda/NA$, the first order ray is at about $0.55 NA/\lambda$ in the pupil. As seen in Figure 3-8, the electric field spill-over due to defocus changes its phase at this position, and therefore the sensitivity phase in Figure 3-16 is reversed. The second peak of the sensitivity is around $0.4 NA/\lambda$, which is consistent with the position at the projection, as attached in plot (f). As the pitch becomes larger than $1.8\lambda/NA$, the second order ray appears at the edge of

pupil, however the effect is insignificant. The electric field spill-over from the two dimensional inverse Fourier transform of the defocus ODP is shown in the plot (d) at pitch $1.14\lambda/\text{NA}$. The sensitivity is 9% lower than the maximum sensitivity at $1.1\lambda/\text{NA}$. In summary, the most sensitive defocus grating is obtained when the pitch equals to $1.1\lambda/\text{NA}$, and the pitch of an effective design should be smaller than $1.4\lambda/\text{NA}$ to obtain the sufficient sensitivity.

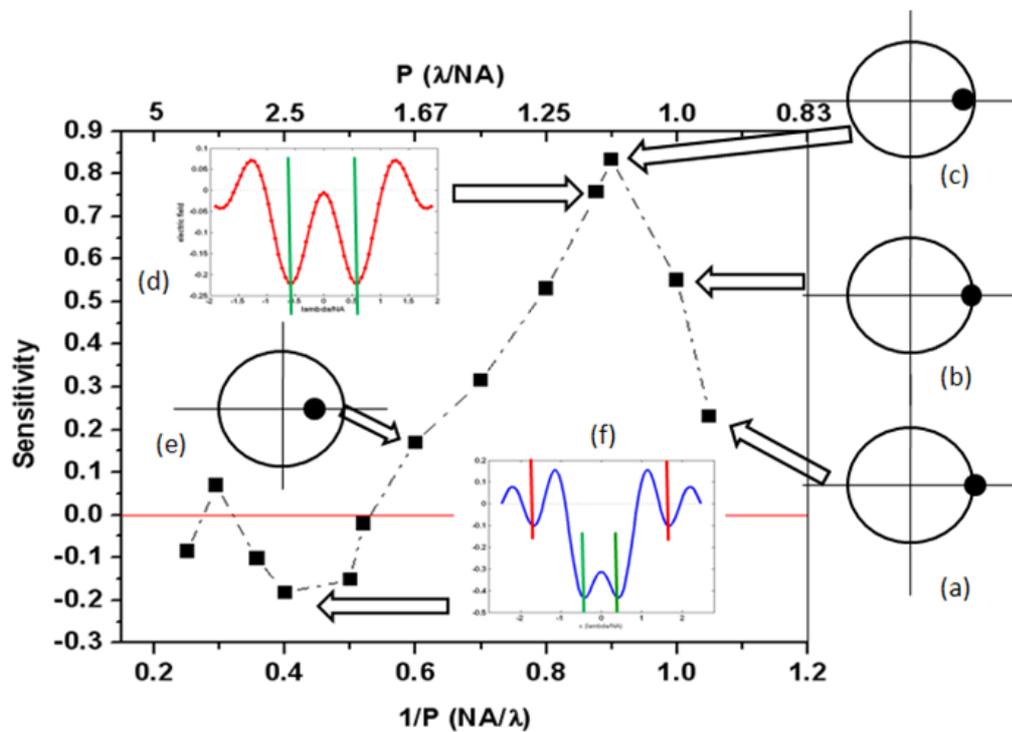


Figure 3-16 PPG defocus sensitivity as a function of the grating pitch. $\sigma=0.1$, $\lambda/\text{NA} = 0.193/0.85$, probe line width 59 nm, and pattern line width 86nm. The pupil maps at different conditions and the related projection of the Fourier transform plots are attached to the sensitivity curve as labeled as (a) – (f)

In order to explore the sensitivity of the commercial exposure tool that is used in the experimental part of this work and compare it with the ideal design, we plot the sensitivity curve at a partial coherence of 0.34 and $\lambda/NA = 0.193/0.85$. The same geometry of the grating is used in Figure 3-17 as that of Figure 3-18. The peak sensitivity reduced to about half of that in $\sigma=0.1$, while the optimum pitch is same at around $1.11\lambda/NA$, at which about 65% of the first order ray is inside the pupil. The sensitivity curve becomes less sharp than the fully coherent condition. The maximum 2D sensitivity at pitch of $1.14\lambda/NA$ is pretty close to the maximum 1D sensitivity at pitch of $1.11\lambda/NA$. The overall trend is similar to the approximately fully coherent case.

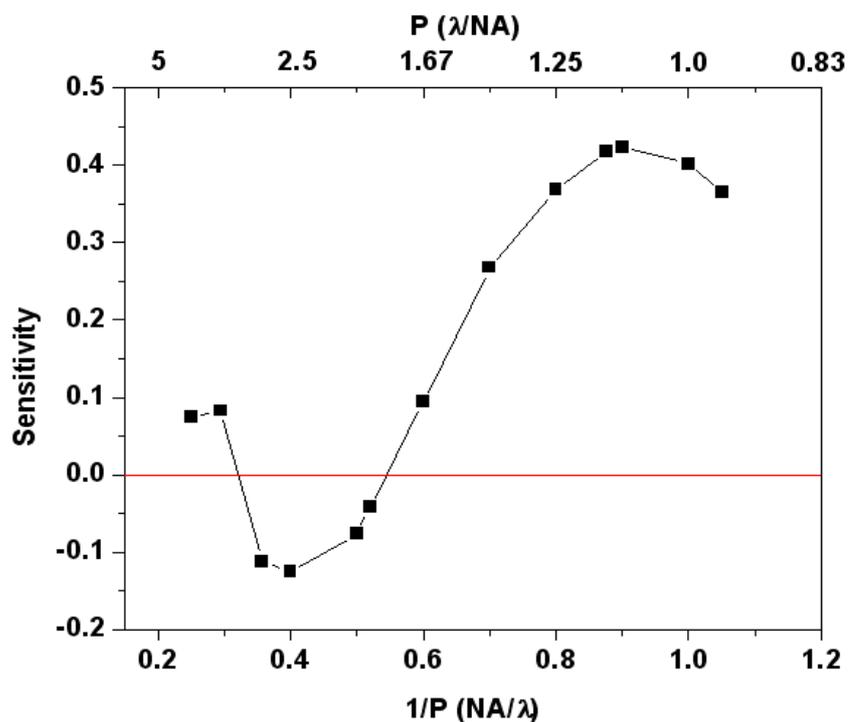


Figure 3-17 PPG defocus sensitivity as a function of the grating pitch. $\sigma=0.34$, $\lambda/NA = 0.193/0.85$, probe line width 59 nm, and pattern line width 86nm.

Further the PPG defocus sensitivity with the probe line width variation is investigated. In Figure 3-8, we observed that the electric field at the central position is not zero, which indicates that the probe line also contributes to the defocus electric field spill-over to itself, and therefore the probe line width will influence the monitor sensitivity. Table 3-3 lists the sensitivity variation with the probe line width. The base probe line width is $0.17\lambda/\text{NA}$, about 39nm, and the pattern line width is $0.38\lambda/\text{NA}$, around 86nm. The probe line width variation step is 10nm. As line width increases, the sensitivity increases accordingly. A similar conclusion applies to the pattern line, which is the major component of the defocus electric field spill over. In section 3.3.2.1, we study the proximity effects with the probe and pattern line width and obtain the ideal design of the line width to keep the reference intensity level at the non-aberration condition. However, Enlarging the probe and pattern line width until the chrome size reaching the mask manufacturing extreme around 40nm, is an available strategy to attain the high sensitivity, which is the major target of the PPG design. For example, if the pitch of the grating is 252nm, the total chrome size of the two sides is 80nm, the probe line and the pattern line width can be 172nm in total, which improves the design margin. As the probe line width increases, the reference intensity level will be larger than 25%CF, as discussed in section 3.3.2.1, the photo-resist may print at this reference zero defocus condition. However, the printed photo resist profile provides additional advantages to the scatterometry calibration. It will have a measurable linear relationship and better calibration precision at the best focus and the small defocus ranges.

Table 3-3 PPG defocus sensitivity variation as probe line width changed with a step of 10nm. The base probe line width is 39nm, the pattern line width is 86nm, and the pitch is 252nm

Δ Width (nm)	0	10	20	30	40
Sensitivity (CF/RU)	0.6957	0.7704	0.8335	0.8858	0.9283

In the real experiment mask layout, the PPG pitch is $1.14\lambda/\text{NA}$ around 260nm, the probe line width is 80nm and the pattern line width is 100nm. The PPG pitch is designed to fit the 2D defocus electric field spill over condition. As shown in Figure 3-17, at $\sigma=0.34$, the sensitivity compromise at pitch of $1.14\lambda/\text{NA}$ is insignificant compared to the best pitch design of $1.11\lambda/\text{NA}$. We performed SPLAT simulation of the probe line aerial image intensity at variable defocus in Figure 3-18. In the defocus range of -0.7 and 0.4 RU, a linear approximation can be approached. The slope is around -0.585 CF/RU. Compared to Figure 3-17 of the same pitch design, the sensitivity improves due to the enlarged probe and pattern line width.

In summary, we discussed the design of the reference line width in section 3.3.2.1, while taking into account the proximity effects within the various features of the grating. The optimum design with the ideal reference intensity level is listed in Table 3-2. The probe and pattern line width can be enlarged further appropriately to increase the PPG sensitivity without compromising the scatterometry calibration at the reference level of zero defocus. Second, we investigated the optimum pitch design to realize the high sensitivity as shown in section 3.3.2.2. The optimum pitch is $1.11 \lambda/\text{NA}$ at both fully coherent circular illumination of $\sigma = 0.1$ and practical illumination of $\sigma = 0.34$. From the SPLAT simulation, the aerial image sensitivity of the real experimental PPG design is

about 0.585 CF/RU at $\sigma=0.34$ and $\lambda/NA = 0.193/0.85$. This design yields a linear defocus range that extends from -0.7RU to 0.4 RU.

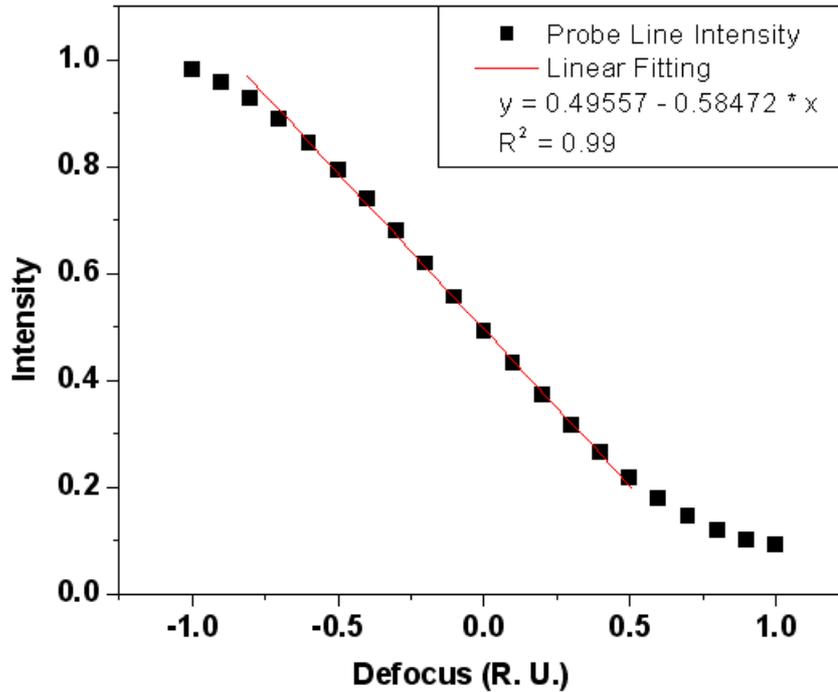


Figure 3-18 PPG aerial image intensity of probe line vs. defocus. PPG pitch is 260nm, probe line width is 80nm and pattern line width is 100nm. The simulation is performed in SPLAT. Positive defocus is defined as the positive distance below the best focus plane.

3.3.3 Mathematical Justification of the PPG Defocus Monitor

Figure 3-19 shows the schematic graph of the PPG defocus monitor. The center is the 90° phase shift probe line with line width W_{pb} . The pattern line is non-phase shifted with line width W_{pa} . The monitor is a periodic grating with pitch P . The pitch design realizes three beam (0 and ± 1 rays) interference.

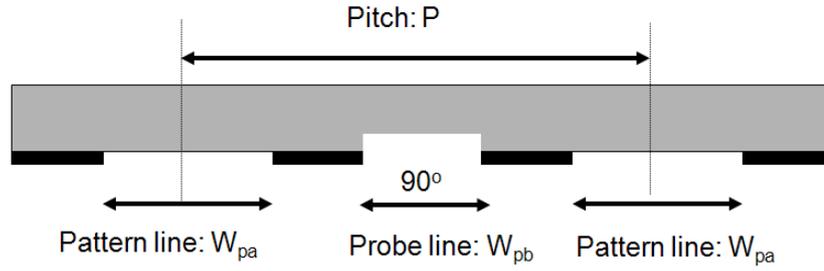


Figure 3-19 Schematic graph of PPG defocus monitor. Probe line width is W_{pb} , pattern line width is W_{pa} , and pitch is P .

The PPG transmission can be mathematically expressed by:

$$t(x) = \text{rect}\left(\frac{x}{W_{pb}}\right) \exp(j\frac{\pi}{2}) \otimes \text{comb}\left(\frac{x}{P}\right) + \text{rect}\left(\frac{x}{W_{pa}}\right) \otimes \text{comb}\left(\frac{x - P/2}{P}\right) \quad \text{Eq. (3-12)}$$

Where $\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n)$. The first part represents the pattern line grating, and the second part represents the probe line grating. The x axis is in the object coordinate plane of the lithography system. The coordinates of the exposure system are shown in Figure 3-20. Since the aspect ratio (length to width) of the grating is larger than 2000, we ignore the y axis in the transmission equation.

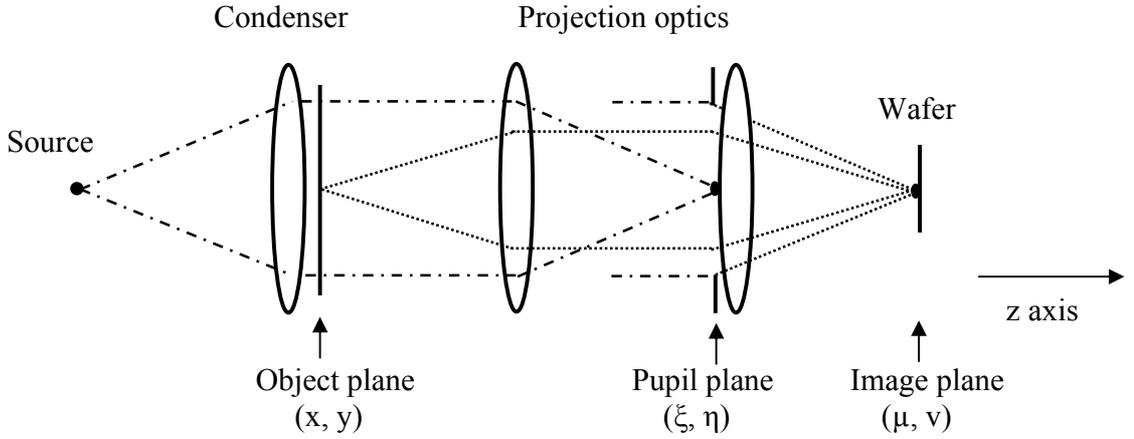


Figure 3-20. The coordinates of the object plane, pupil plane, and the image plane of the projection exposure system. The schematic approximation refers to [3-4].

The Fourier transform of the PPG transmission is the frequency domain representation of the PPG at the pupil plane :

$$\begin{aligned}
 T(\xi) &= F\{t(x)\} = j(W_{pb} \cdot P) \times \text{sinc}(f_x \cdot W_{pb}) \times \text{comb}(f_x \cdot P) \\
 &\quad + (W_{pa} \cdot P) \times \text{sinc}(f_x \cdot W_{pa}) \times \text{comb}(f_x \cdot P) \times \exp(-j2\pi \cdot f_x \cdot \frac{P}{2}) \\
 &= j(W_{pb} \cdot P) \times \text{sinc}(n \cdot \frac{W_{pb}}{P}) \times \text{comb}(n) \\
 &\quad + (W_{pa} \cdot P) \times \text{sinc}(n \cdot \frac{W_{pa}}{P}) \times \text{comb}(n) \times \exp(-j \cdot n\pi) \\
 &= j(W_{pb} \cdot P) \times \text{sinc}(n \cdot \frac{W_{pb}}{P}) \times \text{comb}(n) \\
 &\quad + (W_{pa} \cdot P) \times \text{sinc}(n \cdot \frac{W_{pa}}{P}) \times \text{comb}(n) \times \cos(n\pi)
 \end{aligned} \tag{3-13}$$

where ξ is the axis in the pupil coordinate plane, $f_x = n/P$, and n is the number of the diffraction orders. The respective amplitudes of 0th and ± 1 st diffraction beams are:

$$\begin{cases} T(\xi) = W_{pa} \cdot P + jW_{pb} \cdot P, & n = 0 \\ T(\xi) = -W_{pa} \cdot P \cdot \sin c\left(\frac{W_{pa}}{P}\right) + jW_{pb} \cdot P \cdot \sin c\left(\frac{W_{pb}}{P}\right), & n = +1 \\ T(\xi) = -W_{pa} \cdot P \cdot \sin c\left(\frac{W_{pa}}{P}\right) + jW_{pb} \cdot P \cdot \sin c\left(\frac{W_{pb}}{P}\right), & n = -1 \end{cases} \quad \text{Eq. (3-14)}$$

The three-beam components are shown in the complex coordinates of Fourier space in Figure 3-21. Both the zero and the first order beams have the real and imaginary components, whose magnitudes are determined by the grating geometry parameters: W_{pa} , W_{pb} , and P .

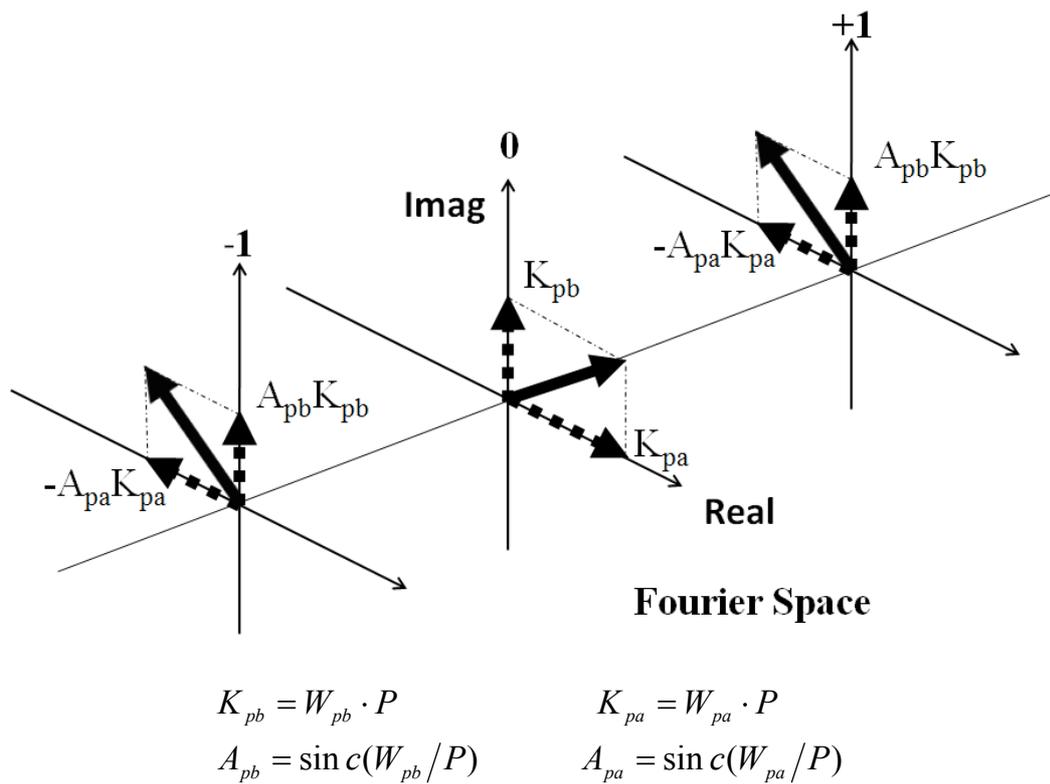


Figure 3-21 PPG three-diffraction-order components in the Fourier space.

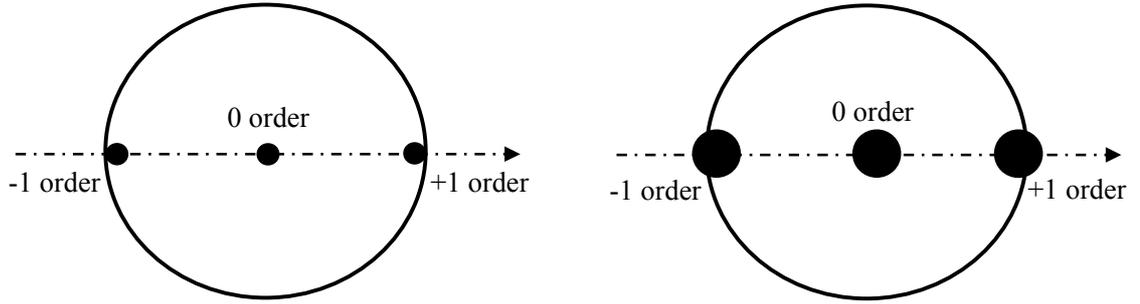


Figure 3-22 Pupil map of three order beam interference: (a) top-hat illumination with $\sigma=0.1$. All three beams are fully inside the pupil. The grating pitch is $1.11\lambda/\text{NA}$ (b) top-hat illumination with $\sigma=0.34$. Around 2/3 of the first orders are inside the pupil. The grating pitch is $1.14\lambda/\text{NA}$.

The pupil map of the three beam interference is shown in Figure 3-22. Figure (a) is the three diffraction orders in the pupil with the on-axis approximately fully coherent illumination of $\sigma=0.1$. The three beams are fully inside the pupil. Figure (b) shows the partial coherence top-hat illumination of $\sigma=0.34$. Around two thirds of the ± 1 orders are inside the pupil. The propagation vector is defined as:

$$k_n = k_{\mu_n} \mu + k_{\nu_n} \nu + k_{z_n} z = k_0 \sin(\theta_{\mu_n}) \mu + k_0 \sin(\theta_{\nu_n}) \nu + k_0 \cos(\theta_{z_n}) z \quad \text{Eq. (3-15)}$$

Consider the rays travel in μ - z plane, and $k_{\nu}=0$, then,

$$k_n = k_{\mu_n} \mu + k_{z_n} z = k_0 \sin(\theta_{\mu_n}) \mu + k_0 \cos(\theta_{z_n}) z \quad \text{Eq. (3-16)}$$

When the PPG is exposed to an approximately coherent light with wavelength λ , the electric field at the image plane is expressed by:

$$\begin{aligned}
E(\mu, z) &= E_0 \exp[-i(k_{\mu_0} \mu + k_{z_0} z + \Phi_0)] + E_1 \exp[-i(k_{\mu_1} \mu + k_{z_1} z + \Phi_1)] + E_{-1} \exp[-i(k_{\mu_{-1}} \mu + k_{z_{-1}} z + \Phi_{-1})] \\
&= E_0 \exp\left[-i \frac{2\pi}{\lambda} (\sin\theta_0 \mu + \cos\theta_0 z + \lambda\varphi_0)\right] + E_1 \exp\left[-i \frac{2\pi}{\lambda} (\sin\theta_1 \mu + \cos\theta_1 z + \lambda\varphi_1)\right] \\
&\quad + E_{-1} \exp\left[-i \frac{2\pi}{\lambda} (\sin\theta_{-1} \mu + \cos\theta_{-1} z + \lambda\varphi_{-1})\right] \\
&= E_0 \exp\left[i \frac{2\pi}{\lambda} (z + \lambda\varphi_0)\right] + E_1 \exp\left[i \frac{2\pi}{\lambda} \left(\frac{\lambda}{P} \mu + \sqrt{1 - \left(\frac{\lambda}{P}\right)^2} z + \lambda\varphi_1\right)\right] \\
&\quad + E_{-1} \exp\left[i \frac{2\pi}{\lambda} \left(-\frac{\lambda}{P} \mu + \sqrt{1 - \left(-\frac{\lambda}{P}\right)^2} z + \lambda\varphi_{-1}\right)\right]
\end{aligned}$$

Eq. (3-17)

Where the magnitudes of E_0 and E_1 are indicated in Eq. (3-14) and Figure 3-21,

$$E_0 = I_{ill} \cdot \sqrt{K_{pa}^2 + K_{pb}^2}, \quad \text{and} \quad E_1 = E_{-1} = I_{ill} \cdot \sqrt{(A_{pa} K_{pa})^2 + (A_{pb} K_{pb})^2}.$$

I_{ill} is the illumination intensity, determined by the exposure dose. The θ_0 is zero for the 0th order beam. The additional phase terms caused by the phase shift mask configurations are:

$$\begin{aligned}
\Phi_0 &= 2\pi\varphi_0 = \arctan\left(\frac{K_{pb}}{K_{pa}}\right) = \arctan\left(\frac{W_{pb}}{W_{pa}}\right) \\
\Phi_1 &= 2\pi\varphi_1 = \arctan\left(\frac{-A_{pb} K_{pb}}{A_{pa} K_{pa}}\right) = \arctan\left(-\frac{W_{pb} \sin c(W_{pb}/P)}{W_{pa} \sin c(W_{pa}/P)}\right) \\
\Phi_{-1} &= 2\pi\varphi_{-1} = \Phi_1
\end{aligned} \tag{3-18}$$

Considering the wave-front phase errors $W_{1,0}$ and $W_{-1,0}$ between 0 and ± 1 orders caused by lens defocus aberration, we introduce :

$$\begin{aligned}
W_{1,0} &= \frac{2\pi}{\lambda} (w_{1,0}) = \frac{2\pi}{\lambda} (w_1 - w_0) \\
W_{-1,0} &= \frac{2\pi}{\lambda} (w_{-1,0}) = \frac{2\pi}{\lambda} (w_{-1} - w_0)
\end{aligned} \tag{3-19}$$

For the on-axis illumination, we have:

$$\begin{aligned}
w_0 &= 0 \\
w_1 &= w_{-1}
\end{aligned} \tag{3-20}$$

Therefore,

$$W_{1,0} = \frac{2\pi}{\lambda} w_1$$

$$W_{-1,0} = \frac{2\pi}{\lambda} w_{-1} = W_{1,0}$$
Eq. (3-21)

The phase errors in Equation (3-21) are added into Equation (3-17):

$$E(\mu, z) = E_0 \exp\left[-i \frac{2\pi}{\lambda} (z + \lambda \varphi_0)\right] + E_1 \exp\left[-i \frac{2\pi}{\lambda} \left(\frac{\lambda}{P} \mu + \sqrt{1 - \left(\frac{\lambda}{P}\right)^2} z + \lambda \varphi_1 + w_1\right)\right]$$

$$+ E_{-1} \exp\left[-i \frac{2\pi}{\lambda} \left(-\frac{\lambda}{P} \mu + \sqrt{1 - \left(\frac{\lambda}{P}\right)^2} z + \lambda \varphi_{-1} + w_{-1}\right)\right]$$
Eq. (3-22)

$$= E_0 \exp\left[-i \frac{2\pi}{\lambda} (z + \lambda \varphi_0)\right] + 2E_1 \cos\left(\frac{2\pi}{P} \cdot \mu\right) \exp\left[-i \frac{2\pi}{\lambda} \left(\sqrt{1 - \left(\frac{\lambda}{P}\right)^2} z + \lambda \varphi_1 + w_1\right)\right]$$

Therefore the intensity at the image plane:

$$I = E \cdot E^* = E_0^2 + 4E_1^2 \cos^2\left(\frac{2\pi}{P} \mu\right)$$

$$+ 4E_0 E_1 \cos\left(\frac{2\pi}{P} \mu\right) \cos\left[2\pi \left(\frac{\sqrt{1 - \left(\frac{\lambda}{P}\right)^2} - 1}{\lambda} z + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda}\right)\right]$$
Eq. (3-23)

When consider the intensity at the center probe line, $\mu = 0$, then,

$$I = E_0^2 + 4E_1^2 + 4E_0 E_1 \cos\left[2\pi \left(\frac{\sqrt{1 - \left(\frac{\lambda}{P}\right)^2} - 1}{\lambda} z + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda}\right)\right]$$
Eq. (3-24)

We note that the definition of the Raleigh Unit (RU) defocus:

$$\delta = \pm \frac{\lambda}{4(1 - \cos \theta)} = \pm \frac{\lambda}{4(1 - \sqrt{1 - \sin^2 \theta})} = \pm \frac{\lambda}{4(1 - \sqrt{1 - (\frac{\lambda}{P})^2})} = \pm RU \quad \text{Eq. (3-25)}$$

So Eq. (3-22) can be rewritten as:

$$I = E_0^2 + 4E_1^2 + 4E_0E_1 \cos \left[2\pi \left(\frac{z}{-4RU} + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right) \right] \quad \text{Eq. (3-26)}$$

As normalizing z to RU and express it as a defocus variable: $f = -\frac{z}{RU}$, we get:

$$I = E_0^2 + 4E_1^2 + 4E_0E_1 \cos \left[2\pi \left(\frac{1}{4}f + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right) \right] \quad \text{Eq. (3-27)}$$

The intensity profile along z axis has bright lines of fringes when the cosine term becomes unity:

$$\frac{1}{4}f + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} = 0 \quad \text{Eq. (3-28)}$$

The above equation represents that the focus plane moves linearly with defocus aberration. We can further express the intensity to a constant intensity I_0 and a defocus derivative intensity I_{df} , where,

$$I_0 = E_0^2 + 4E_1^2$$

$$I_{df} = 4E_0E_1 \cos \left[2\pi \left(\frac{1}{4}f + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right) \right] \quad \text{Eq. (3-29)}$$

The first term f in the cosine function represents the system stage defocus, the second term $(\varphi_1 - \varphi_0)$ represents the phase error due to mask design, and the third $\frac{w_1}{\lambda}$ is the lens aberration. Consider I_{df} at small stage defocus $f \cong 0$, then I_{df} can be expressed by:

$$\begin{aligned}
I_{df} &= 4E_0E_1 \cos \left[2\pi \left(\frac{1}{4}f + (\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right) \right] \\
&\approx 4E_0E_1 (k_0 + k_1f + k_2f^2) \\
&= K_0 + K_1f + K_2f^2
\end{aligned} \tag{3-30}$$

Where,

$$\begin{aligned}
k_0 &= \cos \left\{ 2\pi \left[(\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right] \right\} \\
k_1 &= \frac{\pi}{2} \sin \left\{ -2\pi \left[(\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right] \right\} \\
k_2 &= -\frac{\pi^2}{8} \cos \left\{ 2\pi \left[(\varphi_1 - \varphi_0) + \frac{w_1}{\lambda} \right] \right\} \\
K_0 &= 4E_0E_1k_0, \quad K_1 = 4E_0E_1k_1, \quad K_2 = 4E_0E_1k_2
\end{aligned} \tag{3-31}$$

The cosine function in I_{df} is shown in Figure 3-23. We ignore w_1 in the plot, where

$$I_{\cos} = 4E_0E_1 \cos \left[2\pi \left(\frac{1}{4}f + (\varphi_1 - \varphi_0) \right) \right], \text{ and } I_{\text{linear}} = K_0 + K_1f.$$

In the plot, at the defocus range around $(-0.6RU, 0.3RU)$, the linear term is adequate to approach the cosine function. Therefore, we can use the coefficient K_1 to estimate the sensitivity of the defocus monitor.

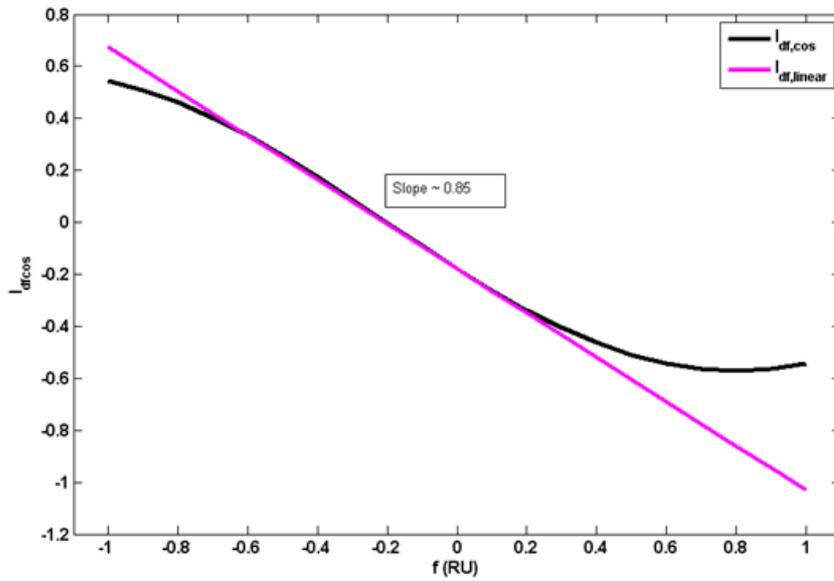


Figure 3-23 Indication of linear approximation of cosine function in I_{df} .

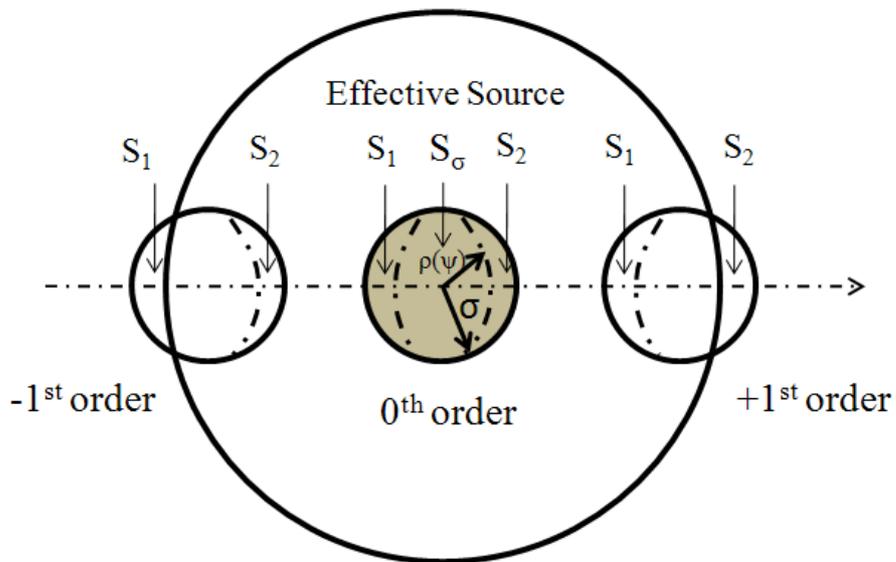


Figure 3-24 Illustration the pupil map of the partial coherent illumination. S_1 and S_2 indicate the area, in which the rays are out of pupil. S_σ is the area of the effective source. $\rho(\psi)$ represents a polar vector in the source. The shaded area in the center is the effective source, which has the same shape as the 0th order.

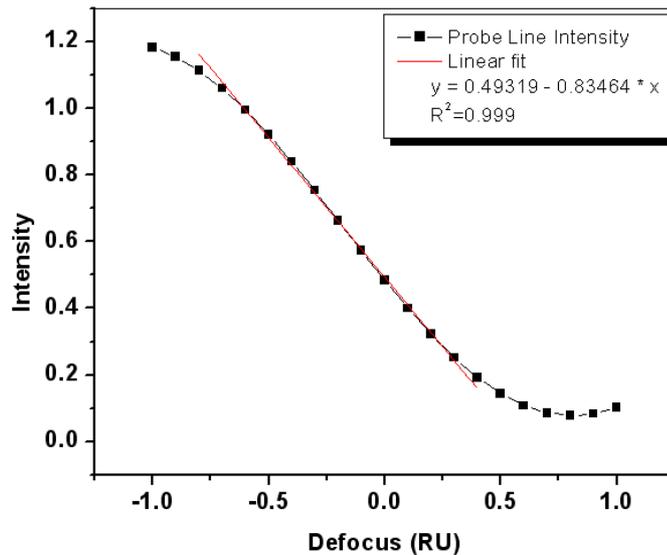
For the circular illumination with partial coherence σ , the pupil map shown in Figure 3-22 (b) indicates that part of the shifted source positions for the diffracted orders falls outside of the lens pupil. The loss of the signal in these lost orders is a major source of loss of modulation and hence sensitivity to focus. Figure 3-24 shows this loss of diffracted orders in detail. Here S_σ is the area of each order with partial coherence σ . S_l is the area of the -1 diffracted order lost and S_2 is the area of diffraction order +1 that are lost. These loss areas can be moved to the actual source in the center that is integrated over to find the intensity. Here S_l and S_2 represent two wave imaging and $S_\sigma - (S_l + S_2)$ is the remaining 3-wave imaging area. Thus the integration over the central source must be broken up into three separate integrals. The correct expressions for each plane wave in the electric field for an angle of incidence θ are given in the top line of Eq. (3-17). These can be specialized to three waves with the off-axis effects on the lateral and vertical directions as well as two waves to determine the intensity contributed by each source point. For illustration in Equation 3-32 the very approximate case is shown where both the z direction behavior and the $w+1$ and $w-1$ values are those for the on-axis source ray. Note that for the two-wave imaging areas the modulation by the focus term is roughly only half as large.

$$I = \begin{cases} \int_{S_\sigma - S_l - S_2} E_0^2 + 4E_1^2 + 4E_0E_1 \cos[2\pi(-\frac{1}{4}f + (\varphi_1(\rho(\psi)) - \varphi_0(\rho(\psi))) + \frac{w_1(\rho(\psi))}{\lambda}]d\rho(\psi) \\ \int_{S_l} E_0^2 + E_1^2 + 2E_0E_1 \cos[2\pi(-\frac{1}{4}f + (\varphi_1(\rho(\psi)) - \varphi_0(\rho(\psi))) + \frac{w_1(\rho(\psi))}{\lambda}]d\rho(\psi) \\ \int_{S_2} E_0^2 + E_{-1}^2 + 2E_0E_{-1} \cos[2\pi(-\frac{1}{4}f + (\varphi_{-1}(\rho(\psi)) - \varphi_0(\rho(\psi))) + \frac{w_{-1}(\rho)}{\lambda}]d\rho(\psi) \end{cases}$$

Eq. (3-32)

3.3.4 Aerial Image Simulation of the PPG Defocus Monitor

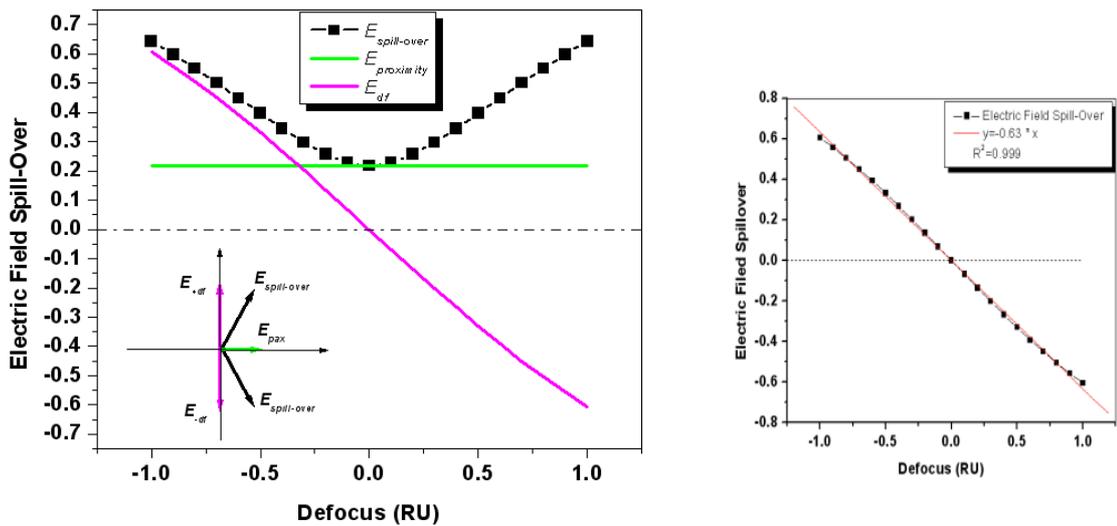
The PPG defocus monitor mask design is shown in Figure 3-19. We use the grating pitch $1.11\lambda/NA = 252nm$, the pattern line width $0.38\lambda/NA = 86nm$, and the probe line width $0.17\lambda/NA + 20nm = 59nm$. The aerial image simulation is performed in SPLAT. Figure 3-25 (a) shows the peak intensity of the probe line at (-1RU, 1RU) defocus range. The positive defocus is defined as the positive distance below the best focus plane. In the small defocus range from $-0.6RU$ to $+0.3RU$, the intensity is a linear function of defocus with slope 0.835.



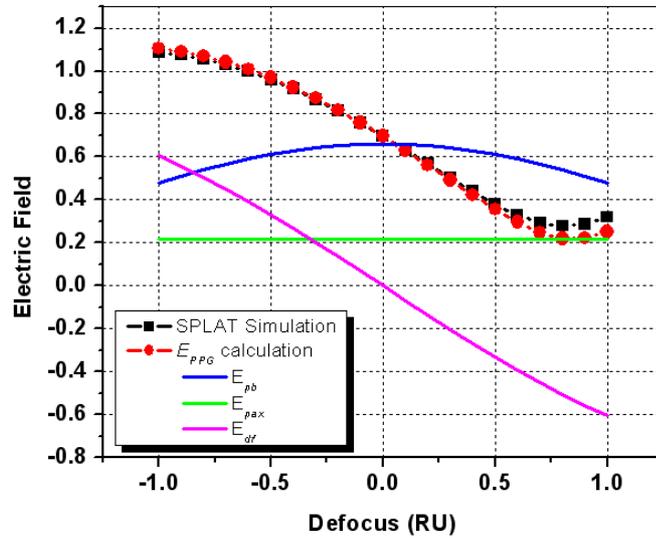
(a) Aerial image simulation of PPG probe line peak intensity vs. defocus. A linear approximation is obtained in the $-0.6RU$ to $0.3RU$ defocus range. The defocus monitor sensitivity is 0.835.

$$\begin{array}{r}
 \text{Imag} \\
 E_{PPG} \\
 E_{pb} \\
 E_{+df} \\
 E_{-df} \quad E_{pax} \\
 \text{Real}
 \end{array}$$

(b) Illustration of PPG probe line electric field components, assuming coherent illumination.



(c) Electric field spill over from pattern line to probe line due to proximity effect and defocus perturbation. The curve of the electric field spillover due to defocus E_{df} is linearly fitted in the right plot with $y = -0.63 * x$.



(d) Comparison of the SPLAT simulation and the electric field phasor model calculation of the PPG probe line intensity vs. defocus.

Figure 3-25 Aerial image simulation and electric field phasor model of PPG probe line peak intensity vs. defocus.

The PPG probe line electric field E_{PPG} is composed of 3 parts: the probe line field E_{pb} , the proximity effect field due to the pattern line E_{pax} , and the electric field spill-over from the pattern line to the probe line due to defocus perturbation E_{df} . Since the probe line is 90° phase shifted, while the pattern line is non-phase shifted, it follows that the pattern line proximity spill over is purely real, and the defocus perturbation field is purely imaginary, and we can draw the probe line electric field components in Figure 3-25 (b). The PPG probe line electric field E_{total} is a vector addition of E_{pb} , E_{pax} , and E_{df} .

We simulated the pure probe line to estimate the E_{pb} as a function of defocus in SPLAT. For the pure pattern line simulation, a signal varied with defocus is at the probe line position, which is expected to be the proximity effect field and defocus perturbation on the electric field spill over. Assuming that the proximity effect field E_{pax} is constant over

small defocus changes, the electric field spill over due to defocus can be extracted from the total electric field spill-over. The proximity effect field and the estimated defocus perturbation are shown in Figure 3-25 (c). The total electric field spill-over $E_{spill-over}$ is the vector addition of E_{df} and E_{pax} .

Then the PPG probe line electric field E_{PPG} can be calculated by E_{pb} , E_{pax} , and E_{df} as shown in Figure 3-25(d), which depicts both the calculation and the SPLAT simulations results. At the small defocus range from $-0.6RU$ to $0.3RU$, a good match is obtained between the SPLAT simulation and the calculation. At larger defocus values, the assumption of the constant proximity effect is no longer valid, and this causes the difference between the simulation and the calculations. At small defocus, the \vec{E}_{PPG} can be expressed as: $\vec{E}_{PPG} = \vec{E}_{pb} + \vec{E}_{df} + \vec{E}_{pax}$, and the amplitude can be written as:

$E_{PPG} = \sqrt{(E_{pb} + E_{df})^2 + E_{pax}^2}$. Therefore the PPG intensity is:

$$\begin{aligned}
 I_{PPG} &= (E_{pb} + E_{df})^2 + E_{pax}^2 \\
 &= E_{pb}^2 + E_{pax}^2 + 2E_{pb}E_{df} + E_{df}^2 \\
 &= E_{pb}^2 + E_{pax}^2 - 1.26E_{pb}f + 0.397f^2
 \end{aligned}
 \tag{Eq. (3-33)}$$

where we use: $E_{df} = -0.63 \cdot f$ Eq. (3-34)

Eq. (3-34) is obtained from the Figure 3-25 (c), and f is the defocus expressed in RU. In Eq. (3-33), the linear term is about 3 to 21 times bigger than the quadratic term in the defocus range from $-0.6RU$ to $0.3RU$. At smaller defocus around the best focus, the linear term becomes more dominant over the quadratic part, and therefore a linear approximation is valid.

Further we consider the practical illumination partial coherence conditions in 193nm exposure tools. The aerial image simulation is shown in Figure 3-26. The sensitivity is worse as the partial coherence increases. The slope decreases from 0.83 at $\sigma=0.1$ to 0.42 at $\sigma=0.34$, and 0.29 at $\sigma=0.5$. The sensitivity loss is due to mutual coherence effect and part of the shifted source positions for the diffracted orders being outside of the lens pupil, as discussed in section 3.3.3. So it might be better to make the grating period larger to move the shifted orders further away from the pupil edges. But the larger period reduces the focus spillover value discussed earlier. This some improvement for sigma of 0.34 might be possible. A large sigma of 0.5 has very little 3 wave and should be about half of the sensitivity even without considering the non-telecentric z-variation and dependence of the value of w_{+l} and w_{-l} on position.

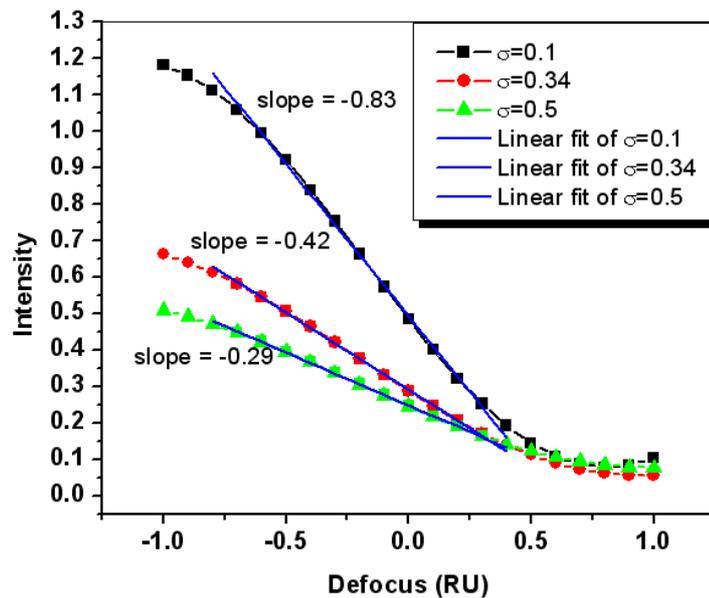


Figure 3-26. Aerial image simulation of PPG at practical partial coherence illumination conditions of $\sigma=0.1$, $\sigma=0.34$ and $\sigma=0.5$.

As we discuss in the section 3.3.2.2, the mask used in the experiment has a pitch of 260nm, a probe line width of 80nm and a pattern line width of 100nm. The aerial image simulation is shown in Figure 3-18. As the probe line and the pattern line size become wider while keeping the pitch optimal, we obtain a better sensitivity of 0.586 at $\sigma=0.34$.

3.3.5 Resist Image Simulation of the PPG Defocus Monitor

The resist image simulation is performed in the Panoramic EM-Suite [3-20]. To match the aerial image simulation in SPLAT, the positive defocus is redefined as the positive distance below the best focus plane. The thick mask topography effect is ignored in this section and we use the “thin” mask model in the simulations. The mask geometry parameters are pitch 260nm, probe line width 80nm and pattern line width 100nm. The illumination is top hat illumination with partial coherence of 0.34, and NA of 0.85.

A maximum of 30% aerial image intensity difference between the Panoramic EM-suit and the SPLAT defocus simulations has been noticed. Since EM-suit uses a vector calculation of the aerial image by taking into account the high-NA effects, we expect the main source of the difference comes from the vector imaging effects of a NA of 0.85. We then repeat the sensitivity study in Panoramic. Although a discrepancy of the sensitivity values is observed, similar trends and conclusions of the optimized mask design range at both $\sigma=0.1$ and $\sigma=0.34$ are obtained. Hence, we confirm the validity and the consistency of using Panoramic as a PPG simulation tool.

Shipley resist at 193nm exposure is used in the simulation and the experiments. The resist thickness is 198nm, and the BARC thickness is 30nm. The film stack and the resist image depth coordinates are shown in Figure 3-27. The z-coordinate has the reversed

direction of that in Figure 3-20. The aerial images formed inside the resist before development is named the latent image. The latent image is different than the aerial image is mainly due to the bending of the waves towards the normal as they enter the resist and the standing wave created by reflection from the bottom surface of the resist [3-21].

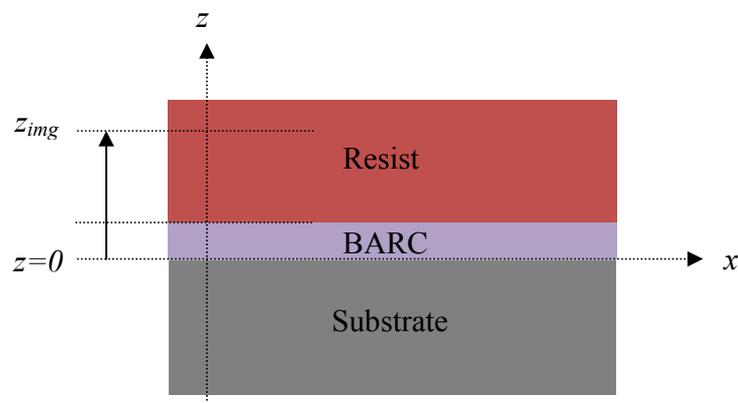


Figure 3-27 PPG resist film stack and coordinates.

The focus plane z_{img} inside the resist is generally determined by finding a plane with the best expected image performance [3-20]. Figure 3-28 shows the PPG latent image profile at various z_{img} . As discussed in the section 3.3.2.1, the preferred PPG performance at the focus condition is that the Pattern line is well printed, but the probe line is right at the threshold or printed weakly. Figure 3-28 indicates that the image plane around 80% of the resist thickness from bottom to top is the best image plane. Since the acid and the base pieces of the CAR resist may diffuse in the x, y, and z direction, and the wavelength aberration may cause the image blur along the z direction, the predication of z_{img} from the

ideal optical condition may not be consistent with the real process. Practically the z_{img} value is obtained by matching the simulation and the experimental results.

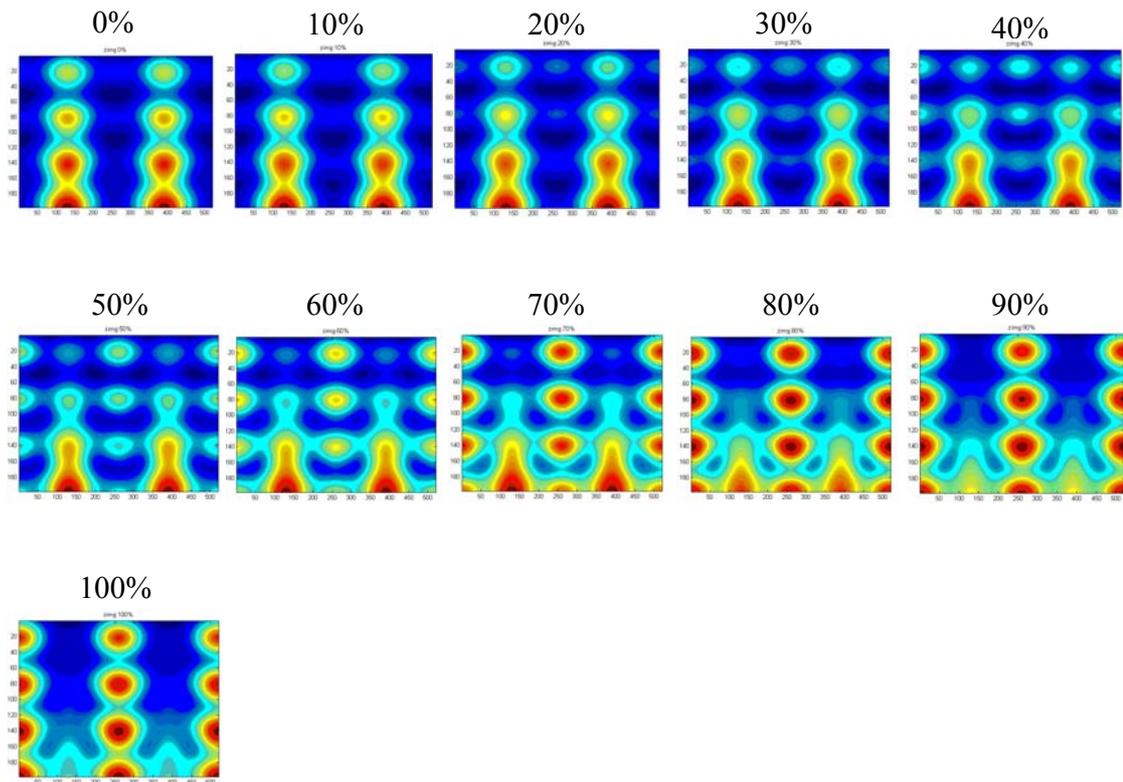
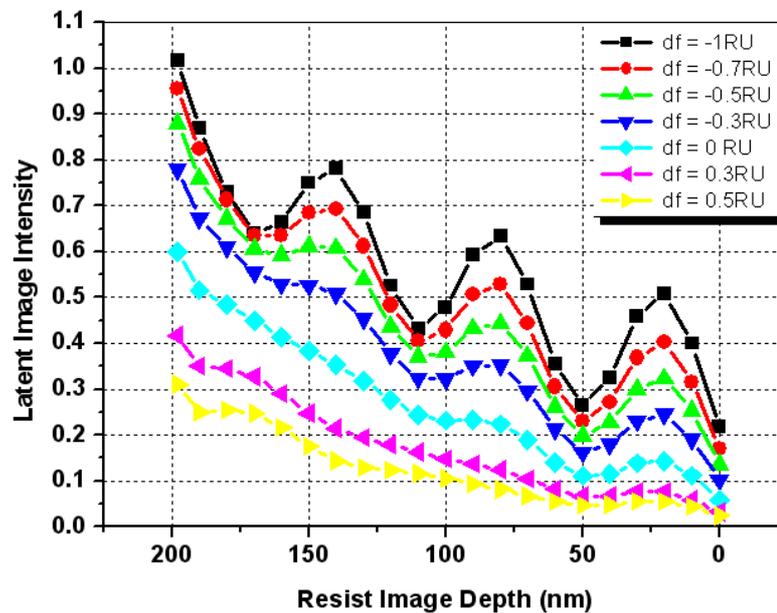


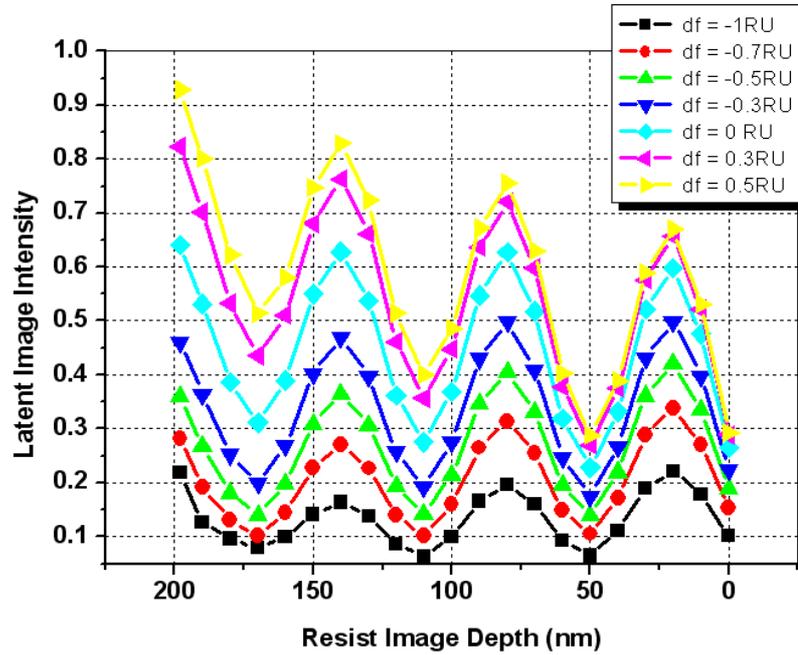
Figure 3-28 Latent image profile at the best focus condition with various image planes inside the resist. The labels on the plot are the percentage of the resist thickness. $Z_{img} = \text{percentage}\% * t + t_{BARC}$, where t is the resist thickness and t_{BARC} is the BARC thickness.

The latent image at various image depths vs. defocus is shown in Figure 3-29. The resist focus plane z_{img} is set 70% of the resist thickness. Figure (a) is the probe line latent image intensity, which has the similar trend as that of the aerial image. The intensity increases as the defocus towards the negative direction, where the resist trench is expected to be deeper after development. The standing wave effect has been seen at the large negative

defocus with a peak intensity period of around 60nm. The intensity attenuated along the resist depth. If we consider a constant threshold of 0.25, an increasing trench depth is expected as defocus towards negative direction. In Figure 3-29 (b), the pattern line has a reversed trend than the probe line. As the defocus increases toward the negative, the intensity becomes smaller. A strong standing wave effect has been in the pattern line. The pattern line intensity decreasing trend along resist depth with defocus is not so significant. We expect that the pattern line resist trench depth to have a sharp change instead of a linear transition with defocus. If we use the threshold of 0.25 in the figure (b), the printed trench depth may jump from 150nm to 20nm when defocus drops from 0.3RU to -0.3RU. The strong standing wave effect will cause a perturbation of the latent image.



(a) Probe line latent image profile vs. defocus



(b) Pattern line latent image profile vs. defocus

Figure 3-29 The Probe line and the pattern line latent image profiles vs. defocus. Image depth = 0 is at the bottom of the resist.

The strong standing wave effect will cause a perturbation of the latent image. The standing wave period [3-22] is estimated by:

$$P = \frac{\lambda}{2n \cos \theta} \quad \text{Eq. (3-35)}$$

where P is the period, n is the refractive index of the resist, and θ is the incident angle of the light. For the small incident angle and the resist of $n=1.7$, the standing wave period is around 57nm. The simple calculation range matches with the simulation results around 60nm in Figure 3-29.

We use the diffusion-reaction and develop equations in EM-suit to calculate the resist image. The develop rate, fudge factor in the development and the kinetic and the

temperature parameters in the post exposure bake employ the suggested values of the general Shipley 193nm resist from Panoramic software. The resist diffusion length is around 0.05 μm . The development time is 60sec. We perform the focus exposure matrix (FEM) on the PPG pattern. The focus factors are from -1.5RU to +1.5RU with 0.05RU step, and the dose parameters are from 18mJ/cm² to 28mJ/cm² with 1mJ/cm² step. The image depth is fixed at 70% of the resist thickness. In order to compensate the variations in the resist process, a variable image depth is considered in the experimental analysis. The resist image contours are shown in Figure 3-30. The fields along horizontal direction from left to right are the defocus from -1RU to 1RU with a step of 0.2RU. The fields along vertical direction from top to bottom are the dose from 18mJ/cm² to 28mJ/cm² with the 2mJ/cm² step. Each plot shows two periods of PPG. The probe line is at the center of one pitch and the pattern line is at the border of a pitch. The probe line trench depth increases as defocus towards the negative direction. The pattern line becomes significant at another defocus side. As dose increases, the exposure energy increases, and therefore the printed trench depth becomes deeper.

The probe trench depth is defined as the distance from top of the original resist film of 198nm to the remaining resist surface at the probe line center position. In Figure 3-30, we observed that the total resist thickness may be attenuated to be less than 198nm, but we use the same definition to calculate the probe trench depth as a difference between the original resist film thickness to the current probe line center thickness. The extracted probe line trench depth is shown in Figure 3-31(a). An approximately linear relationship of the trench depth with defocus is obtained at each dose condition. The linear fitting is

shown in Figure 3-31(b). The curves with $2\text{mJ}/\text{cm}^2$ dose step are shown in the figure for a clear view.

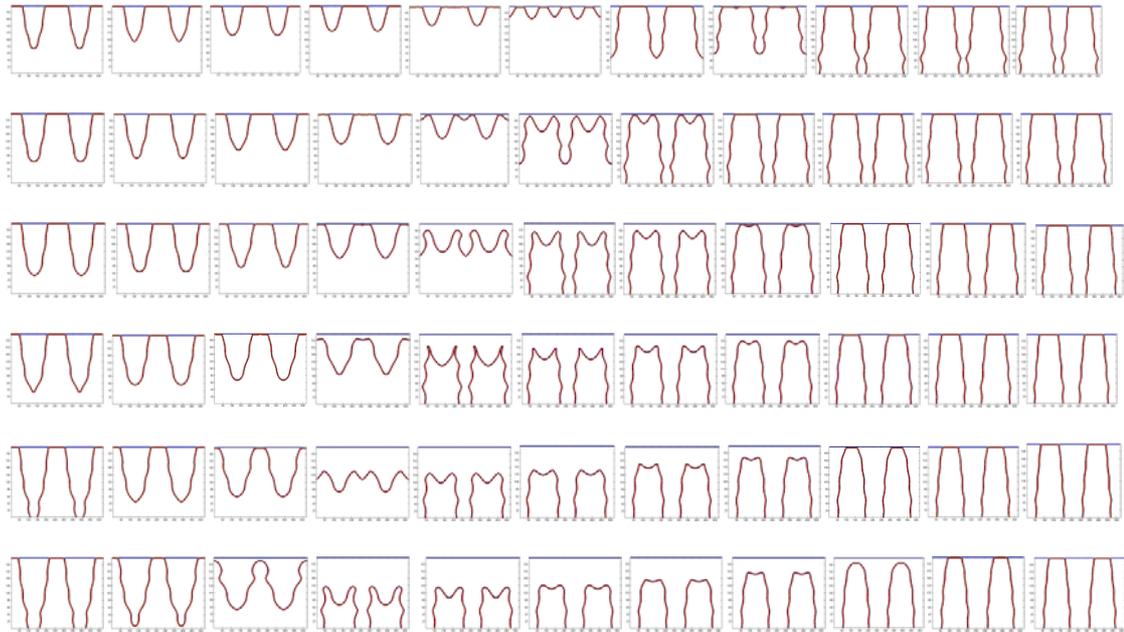
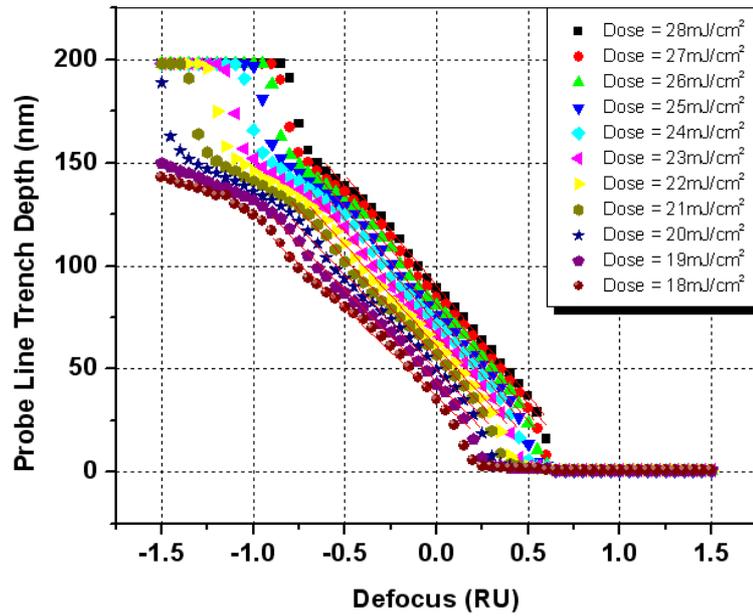
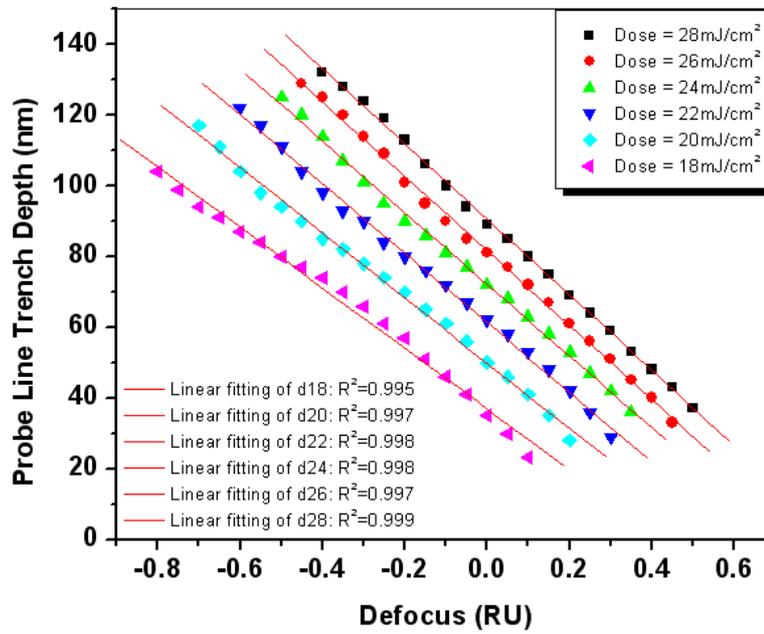


Figure 3-30 Illustration of PPG Focus Exposure Matrix (FEM). In the completed FEM, the fields along x direction from left to right are the defocus with a step of 0.05RU and nominal defocus is zero. The fields along the y direction from top to bottom are the dose with a step $1\text{mJ}/\text{cm}^2$ with a nominal dose $22\text{mJ}/\text{cm}^2$. The defocus from -1RU to 1RU with 0.2RU step, and dose from $18\text{mJ}/\text{cm}^2$ to $28\text{mJ}/\text{cm}^2$ with $2\text{mJ}/\text{cm}^2$ step are shown in this figure.



(a) The probe line trench depth with defocus from -1.5 RU to 1.5RU, and with dose from 18mJ/cm² to 28mJ/cm²



(c) The linear approximation of the trench depth vs. defocus with a 2mJ/cm² dose step

Figure 3-31 Probe line trench depth vs. defocus vs. dose.

The slopes of the fitting curves are weakly related with dose with the mean of -98nm/RU and the standard deviation of 6.255. We define the intercept as the probe trench depth at the zero defocus condition. Therefore the intercept represents the probe line depth with pure dose variations. The linear and the quadratic approximations of the depth vs. dose are shown in Figure 3-32. The quadratic approximation has a higher goodness-of-fit, while the linear approach is sufficient in Figure (a) with R^2 of 0.9957. The linear relationship of the intercept and the dose can be described by:

$$I(D_{ose}) = 57.03 + 5.329 \times D_{ose} \quad \text{Eq. (3-36)}$$

We assume that the defocus offset differences between fields are insignificant. Therefore, the probe line trench depth in the linear region of Figure 3-32 can be represented by:

$$T_{ij} = S_i \cdot F_{ocus,j} + I_i(D_{ose,i}) \quad \text{Eq. (3-37)}$$

where T_i is the probe line trench depth, F_{ocus} is the focus, D_{ose} is the dose, S_i is the slope, I is the interception as a function of dose, i is the index of dose, and j is the index of defocus. As considering the focus and dose variations, Eq. (3-37) can be further expressed by:

$$\begin{aligned} T_{ij} &= S_i \cdot (F_j + \Delta F_j) + I_i(D_i + \Delta D_i) \\ &= S_i \cdot (F_j + \Delta F_j) + A + B(D_i + \Delta D_i) \end{aligned} \quad \text{Eq. (3-38)}$$

where F_j is the programmed defocus setting, ΔF_j is the defocus variation, D_i is the programmed dose setting, and ΔD_i is the dose variation. A is the fitting parameter 57.03, and B equals to 5.329.

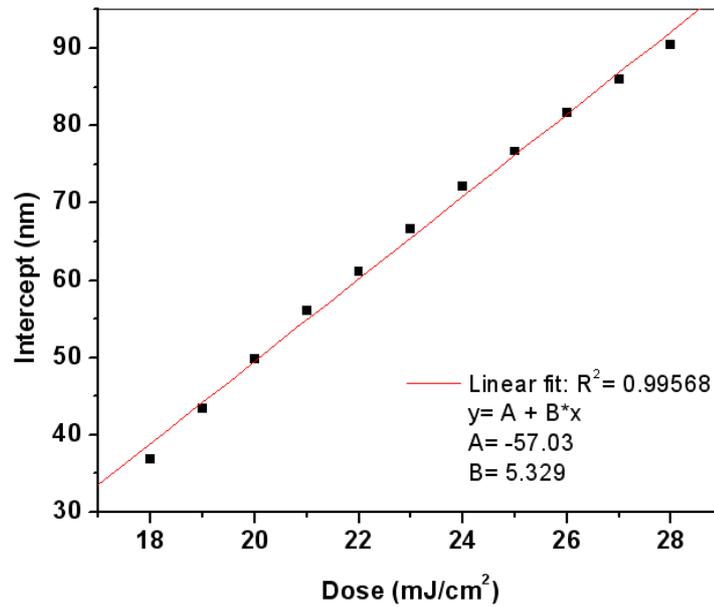


Figure (a) the linear approximation of the interception as a function of dose

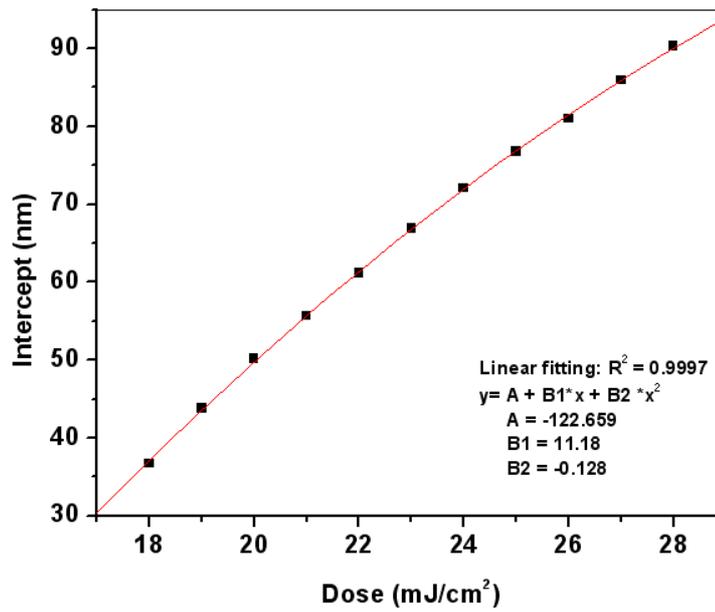


Figure (b) the quadratic approximation of the interception as a function of dose

Figure 3-32 Analysis of the simulated probe line trench depth vs. defocus

In the simulation, we have 11 dose parameters from 18 to 28mJ/cm². Since the maximum linear defocus range is from -0.8RU to 0.5 RU, we have defocus parameters. So there are 38 variables, and 297 measurements. The Eq. (3-38) can be re-expressed as Eq. (3-39).

$$\begin{bmatrix} T_{0,0} \\ T_{0,1} \\ T_{0,2} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ T_{11,26} \end{bmatrix} = \begin{bmatrix} S_0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & B & 0 & \cdot & \cdot \\ 0 & S_0 & 0 & \cdot & \cdot & \cdot & \cdot & B & 0 & \cdot & \cdot \\ 0 & 0 & S_0 & \cdot & \cdot & \cdot & \cdot & B & 0 & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ \cdot & & & \cdot & \cdot & & & & & & \\ 0 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & S_{10} & 0 & 0 & \cdot & B \end{bmatrix} \cdot \begin{bmatrix} \Delta F_0 \\ \Delta F_1 \\ \Delta F_2 \\ \cdot \\ \cdot \\ \cdot \\ \Delta D_0 \\ \cdot \\ \cdot \\ \Delta D_{10} \end{bmatrix} + \begin{bmatrix} S_0 \cdot F_0 + B \cdot D_0 + A \\ S_0 \cdot F_1 + B \cdot D_0 + A \\ S_0 \cdot F_2 + B \cdot D_0 + A \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ S_{10} F_{26} + B \cdot D_{10} + A \end{bmatrix}$$

Eq. (3-39)

The left side of Equation (3-39) is the measured probe line trench depth. The first term of the right side of the equation is: $S_i \cdot \Delta F_j + B \cdot \Delta D_i$, and the second term is the constants of the known parameters: $S_i \cdot F_j + B \cdot D_i + A$. If we express the above matrix by the symbols as: $T = x \cdot \beta + \gamma$, then T is 297x1 matrix, x is 297x38 matrix, β is 38x1 matrix, and γ is 297x1 matrix. It is note that at certain doses, the linear defocus ranges are smaller than the maximum range. The measurements corresponding to these rows should be removed from the properties. Through solving the linear equation of Eq. (3-39), we can get $\beta = (x^T x)^{-1} x^T T$, Hence the process defocus offset ΔF_j and the dose offset ΔD_i are obtained.

Figure 3-32(b) shows the quadratic approximation of the dose properties. A more accurate fitting is achieved. We represent Eq. (3-36) and Eq. (3-38) as:

$$I(D_{ose}) = A + B_1 \times D_{ose} + B_2 \times D_{ose}^2 \quad \text{Eq. (3-40)}$$

$$\begin{aligned} T_{ij} &= S_i \cdot (F_j + \Delta F_j) + I_i(D_i + \Delta D_i) \\ &= S_i \cdot (F_j + \Delta F_j) + A + B_1(D_i + \Delta D_i) + B_2(D_i + \Delta D_i)^2 \end{aligned} \quad \text{Eq. (3-41)}$$

We define: $f(\Delta F_j, \Delta D_i) = S_i \cdot (F_j + \Delta F_j) + A + B_1(D_i + \Delta D_i) + B_2(D_i + \Delta D_i)^2$

The solution of ΔF_j and ΔD_i can be found by solving the minimization problem:

$$\{\{\Delta F_j\}, \{\Delta D_i\}\} = \arg \min \sum_{i,j} (T_{i,j} - f(\Delta F_j, \Delta D_i))^2 \quad \text{Eq. (3-42)}$$

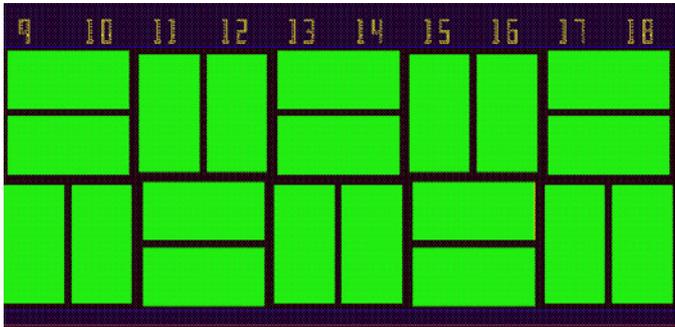
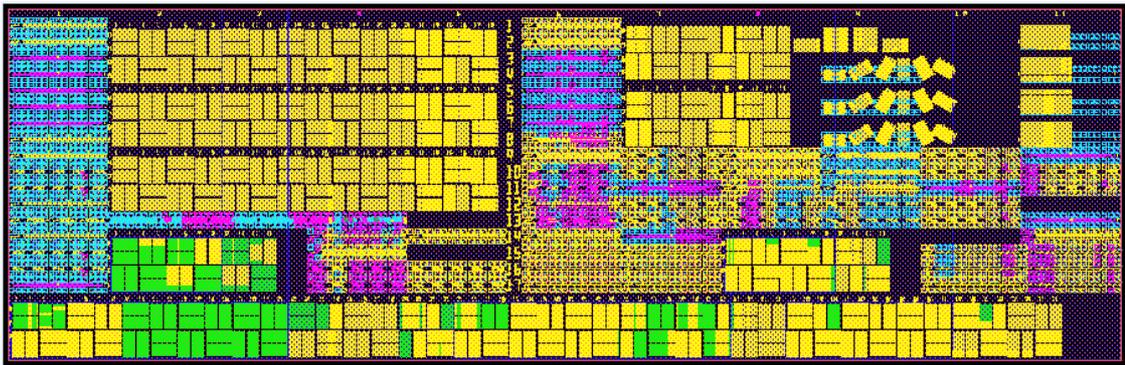
Taking the first-order derivatives and using Matlab optimization (non-linear least square) toolbox, we can easily solve the optimization problem. The details will be shown in the section 3.4.2.

3.4 Printing Experiments and Scatterometry Calibration of PPG Defocus Monitor

3.4.1 Printing Experiments of PPG Defocus Monitor

The periodic gratings of the PPG mask are designed to investigate the first side lobe effect of the defocus spillover. The one-pitch pattern schematic is shown in Figure 3-19 with the pitch of 260nm, the probe line width of 80nm and the pattern line width of 100nm. The mask layout is shown in Figure 3-33. The PPG patterns are at the bottom of the field. There are 76 columns and 4 rows in one field, where the different pitch patterns are drawn to investigate the second and the third side lobe of the defocus spillover,

however the measurements are focused on the primary pitch patterns. Both the vertical and the horizontal gratings are included on the mask. The grating area in one block is $250\ \mu\text{m} \times 500\ \mu\text{m}$. The same pattern repeats three times in the left, middle and the right side of the die, as shown in Figure 3-34.



Probe Line Pattern Line

Figure 3-33 Illustration of the PPG mask layout. The PPG patterns are at the bottom of the field. There are 76×4 grating groups in one field. The top figure is to show the whole mask layout. The bottom left figure is an example of the PPG grating blocks, and the bottom right figure is the enlarged PPG gratings in two periods.

The printing experiments were performed at Spansion Inc. A 300mm wafer was exposed by a 193nm ASML dry scanner with NA of 0.85 and σ of 0.34. The top-hat illumination is used. The resist thickness is 193nm, and the BARC thickness is 30nm. The focus exposure matrix was prepared for scatterometry measurement. The fields along y-direction are programmed with 10nm focus steps centered on nominal focus. From top to bottom, the defocus is from -150nm to 150nm. The fields along x-direction are programmed with 1mJ/cm² dose step, centered on 18mJ/cm². The dose is from 12mJ/cm² to 24mJ/cm². Figure 3-35 shows the printing layout. The area of the green color indicates the fields measured by the scatterometry ellipsometry techniques.

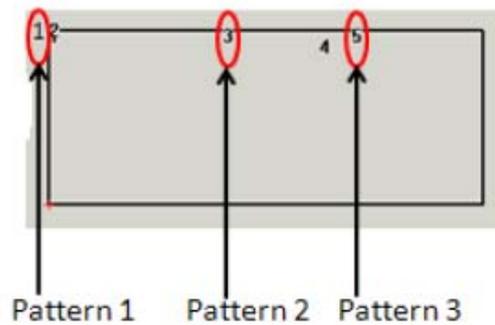
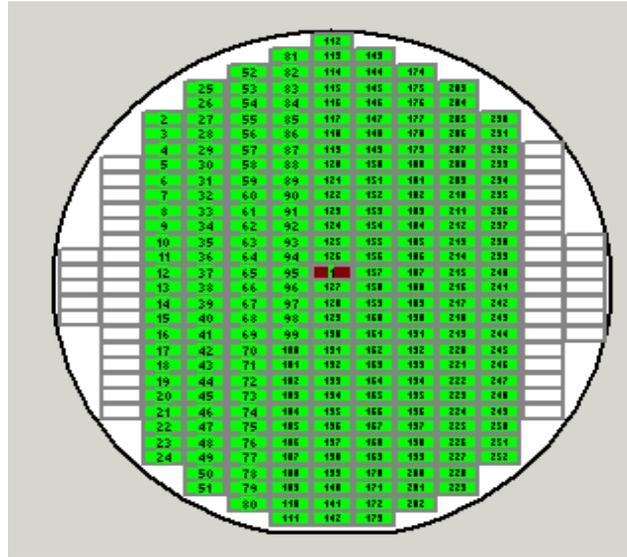


Figure 3-34 Illustration of the pattern positions of the repeated measurements in one field.



Nominal dose: 18mJ/cm²; Step: 1mJ/cm²

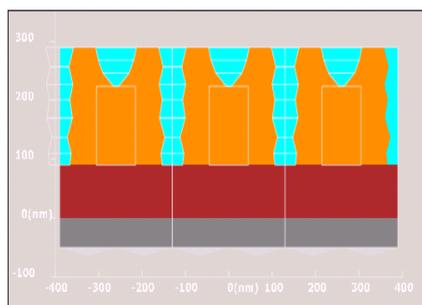
Figure 3-35 Printing layout of the focus exposure matrix. Defocus is from -150nm to 150nm with 10nm step. The nominal defocus is 0. Dose is from 12mJ/cm² to 24mJ/cm² with 1mJ/cm² step. The nominal dose is 18mJ/cm². In the scatterometry measurement, the fields with the green color are measured.

3.4.2 Scatterometry Calibration of the PPG Defocus Monitor

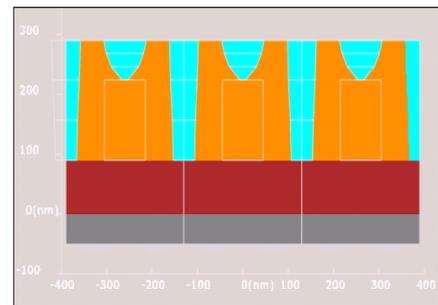
The measurement of the printed pattern is performed in the Spansion Inc. The spectroscopic ellipsometry is used to collect the spectra of the PPG resist cross section profile. The Timbre Optical Digital Profilometry (ODP) is the software platform to do the simulation and analysis of the scatterometry measurement results.

The typical probe-pattern line resist profile is shown in Figure 3-36(a). The standing wave effect impacts the pattern line shape. In order to reduce the complexity of the ODP model, we assume that the standing-wave induced profile is not critical to the ODP

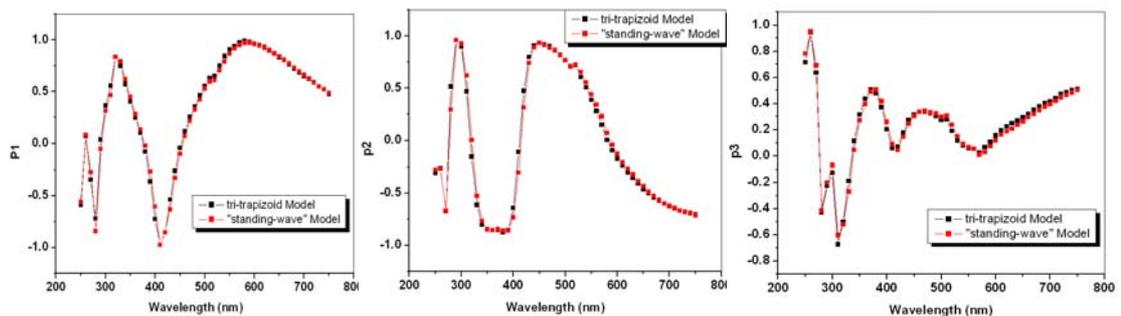
modeling, and we can use a tri-trapezoid model to approximate the real pattern line, which is shown in Figure 3-36(b). To verify this assumption, a model of real standing wave structure and a simplified tri-trapezoid model profile are generated in ODP to test the sensitivity of the spectrum. The three associated parameters p_1 , p_2 , and p_3 , which correspond to the ellipsometry parameters of N, S, and C, of each model are compared in Figure 3-36(c). We confirmed that there is no significant difference between the two models, and we therefore can use the simple trapezoid model in the ODP simulation.



(a) “standing-wave” model



(b) tri-trapezoid model



(d) Ellipsometry spectrum parameters p_1 , p_2 and p_3 for “standing-wave” model and tri-trapezoid model

Figure 3-36 Illustration of the “standing-wave” model and simplified “tri-trapezoid” model in the ODP calibration

The multiple shape indices of the probe-pattern line are shown in Figure 3-37. We use seven main parameters and three dependent parameters to define the PPG resist behavior: the probe trench depth (T_HT), probe top CD (T_TCD), probe bottom CD (T_BCD), pattern line depth (P_HT), pattern top CD (P_TCD), pattern bottom CD (P_BCD), resist height (TP_HT), probe mid CD (T_MCD), pattern mid-top CD (P_MCD), and pattern mid-bottom CD (P_MBCD). In the ODP simulations, the model is first evaluated according to the sensitivity and correlation of the parameters. Then the model will be modified and reevaluated until an optimal model is validated. The optimized model is fed into ODP-NOW regression to determine the resist profile solutions by matching the simulated spectra and the experiment measured spectra.

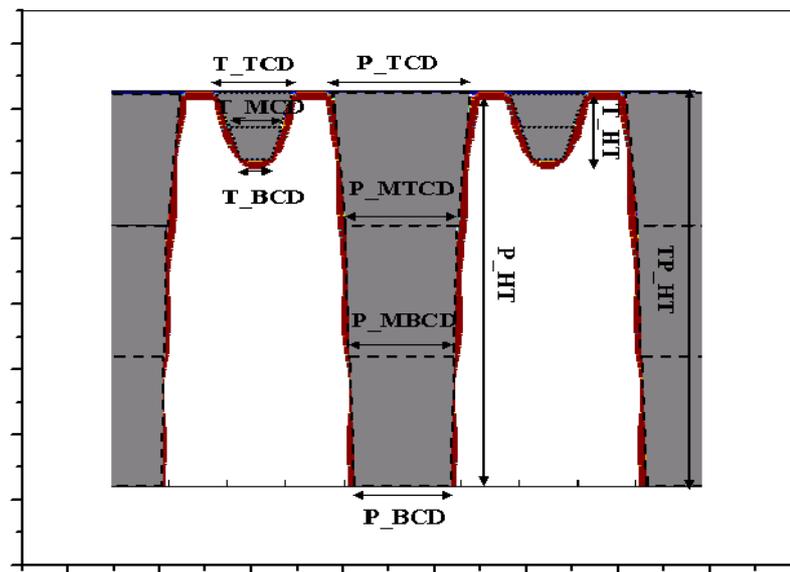


Figure 3-37 Multiple parameter characterization of PPG resist profile. T represents probe line, and P represents pattern line in the graph.

The measurement and the calibration results of the vertical gratings are shown in Figure 3-38. The measurement on the same pattern repeated three time in one field. The average values of the three measurements are used to build the model. The Bossung curve based on the five dose conditions from 18 to 22mJ/cm² is plotted in Figure 3-38 (a), while we ignore the doses from 14 to 17mJ/cm² which turned out to be too low to print the reliable patterns. Figure 3-38 (b) shows the linear approximation in the defocus range of -1 to +0.6 RU. To avoid over-fitting to the noise and effectively remove the outliers, the RANSAC fitting method is used. The fitting R² is above 0.98. The average slope is 116.9nm/RU, and the standard deviation is 1.57nm/RU. We show the slopes of the trench depth vs. defocus in Figure 3-39 (a). The slope is approximately independent of dose. The sensitivity of the vertical PPG lines is therefore calibrated to be around 0.85 nm defocus / per nm of trench depth. This sensitivity is higher than the Phase-Grating Focus Monitor (PGFM) sensitivity of around 3nm defocus /nm overlay [3-8].

Figure 3-39(b) shows the intercept vs. dose. The second order or higher order polynomial fitting is used with R² above 0.989. With the measured probe trench depth, the linear function of the slope and the defocus, we can solve the equation (3-37) to obtain the predicted defocus. The results are shown in Figure 3-40. An excellent correlation between the predicted defocus and the scanner programmed defocus is achieved. The larger mean offset at the edge is not a concern cause PPG is highly sensitive and better suited to the measurement of tigher defocuse distributions. The maximum standard deviation of the defocus spread is 4.2nm, which indictates in the measured FEM matrix, the maximum cross field defocuse variation is 4.2nm.

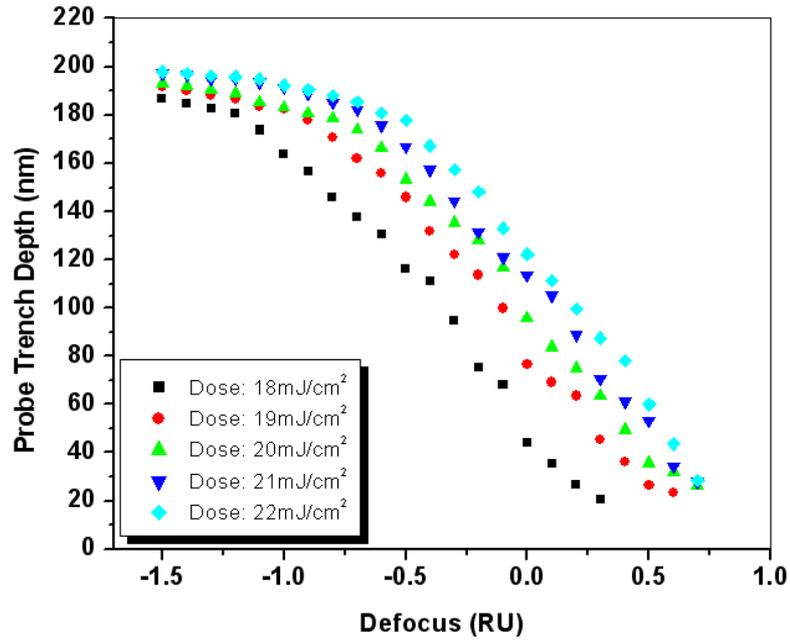


Figure (a) Vertical probe line trench depth vs. defocus with dose from 18mJ/cm² to 22mJ/cm²

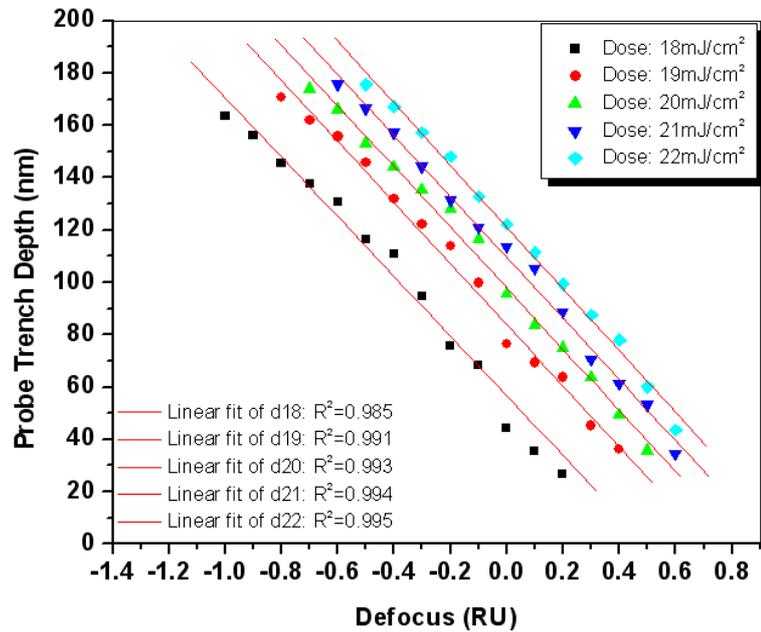


Figure (b) Linear approximation of vertical probe line trench depth vs. defocus with dose from 18mJ/cm² to 22mJ/cm²

Figure 3-38 Vertical probe line trench depth measurement and calibration results

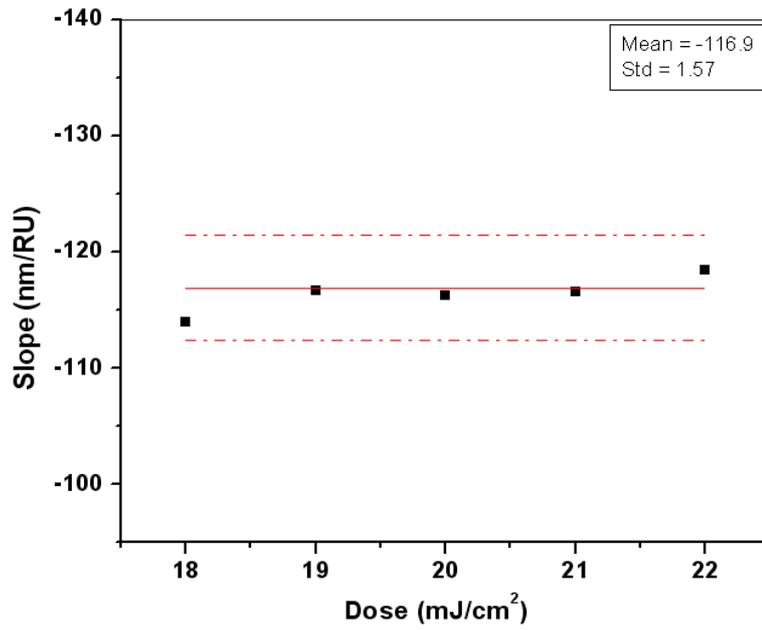


Figure (a) Slope of vertical probe line trench depth vs.defocus

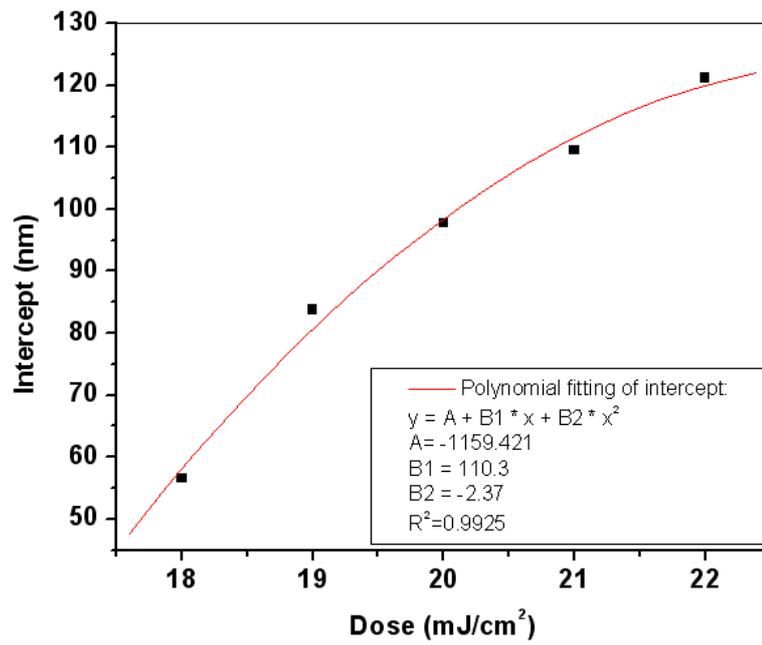


Figure (b) Intercept of vertical probe line trench depth vs.defocus as a function of dose

Figure 3-39 Analysis of the claibration results of vertical probe line trench depth vs. defocus

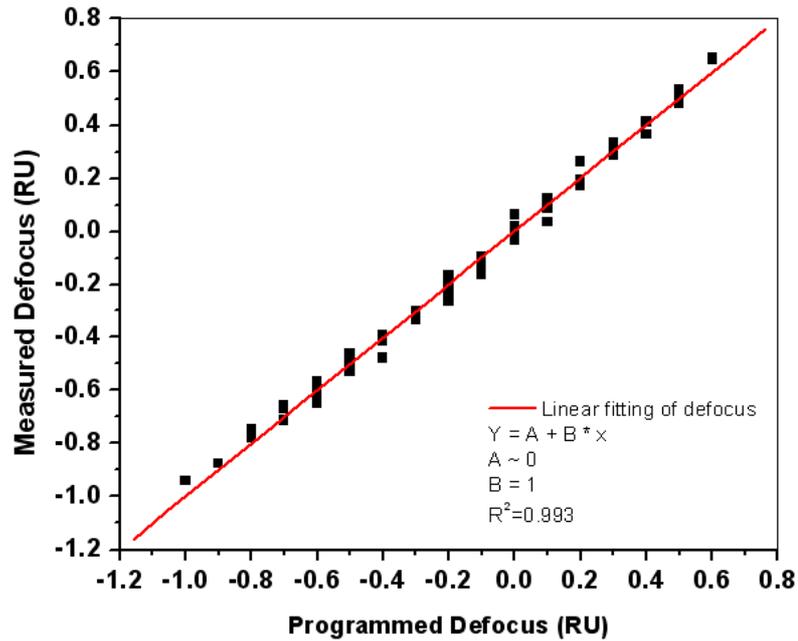


Figure 3-40 Vertical line measured defocus vs. scanner programmed defocus

According to the semi-empirical model in equation (3-38), the defocus and dose variations as offsets respect to the programmed defocus and dose can be calculated. In Equation (3-41), we represent the measured probe trench depth as T_{ij} , the scanner programmed defocus as F_j , and the programmed dose as D_i . The fitting parameter A, B_1 , and B_2 are determined from the fittings of the intercept vs. dose. The unknown parameters are the defocus variation ΔF_j , and the dose perturbation ΔD_i . The Eq. (3-41) is rewritten as follows:

$$\begin{aligned}
 T_{ij} &= S_i \cdot (F_j + \Delta F_j) + I_i (D_i + \Delta D_i) \\
 &= S_i \cdot (F_j + \Delta F_j) + A + B_1 (D_i + \Delta D_i) + B_2 (D_i + \Delta D_i)^2 + B_3 (D_i + \Delta D_i)^3 \dots
 \end{aligned}$$

Eq. (3-43)

The fitting parameters are determined by the polynomial fitting of intercept vs. dose. More than two orders can be used to improve the fitting accuracy, but with the awareness

of the balance between a good fit and the over-fit to the noise. Eq. (3-35) can be solved by Eq. (3-42), which is rewritten below:

$$\{\{\Delta F_j\}, \{\Delta D_i\}\} = \arg \min \sum_{i,j} (T_{i,j} - f(\Delta F_j, \Delta D_i))^2 \quad \text{Eq. (3-42)}$$

Where, $f(\Delta F_j, \Delta D_i) = S_i \cdot (F_j + \Delta F_j) + A + B_1(D_i + \Delta D_i) + B_2(D_i + \Delta D_i)^2 + B_3(D_i + \Delta D_i)^3 \dots$

In the empirical model, five dose settings are used with i equals to 5. The focus range is from -1.0 RU to 0.6RU in the linear approximation, therefore j equals to 17. There are 78 measurements of the probe trench depth in the calculations. By taking the first-order derivatives and using the nonlinear least square toolbox in Matlab, the ΔF_j and ΔD_i are solved in Figure 3-41. Figure (a) is the defocus variations. The single variation value is extended to two-side error bar in the plot. The mean defocus offset to programmed defocus is around 0, and the standard deviation is 2.3nm. The maximum defocus offset is 4.49nm, around 7.48% of 0.6RU. The defocus residual is defined as the measurement defocus minus the modeled defocus. The mean of the residual is 0.7nm, and the standard deviation is 0.73nm. The dose variation is shown in Figure (b). The maximum dose offset is 0.244mJ/cm², about 1.16% of dose 21mJ/cm². The mean dose offset is -0.05mJ/cm², and the standard deviation is 0.16mJ/cm².

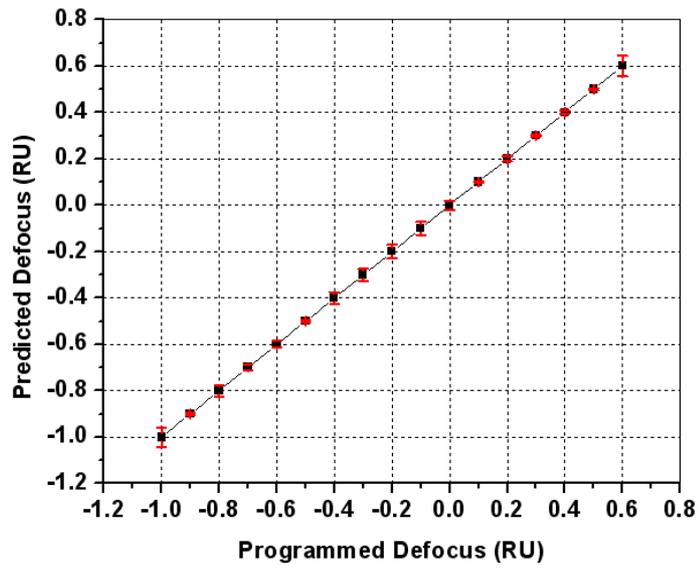


Figure (a) The calculated defocus offsets on programmed defocus. The single variation value is extended to two-side error bar in the plot.

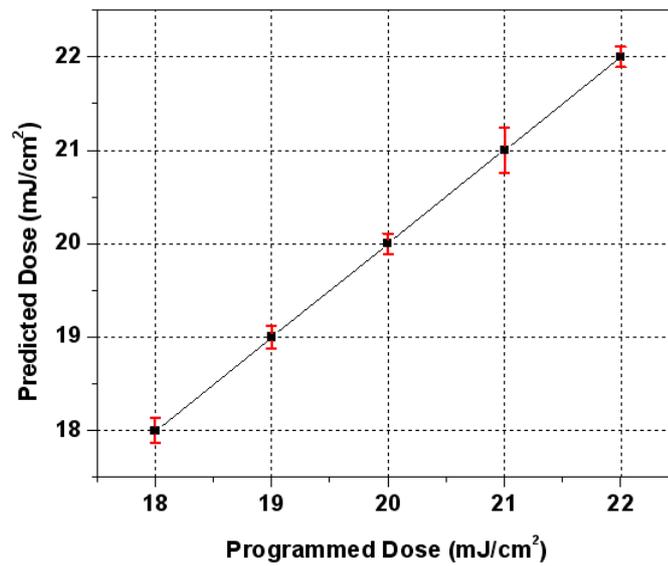


Figure (b) The calculated dose offsets on programmed dose. The single variation value is extended to two-side error bar in the plot.

Figure 3-41 The predicted defocus and dose offsets for vertical gratings

The Bossung curve of the horizontal line is shown in Figure 3-42 (a) with five dose conditions from 18 to 22mJ/cm². Figure 3-42 (b) shows the linear approximation in the defocus range from -0.7RU to +0.6 RU. The fitting R² is above 0.987 for all the doses except 21mJ/cm² with R² of 0.969. The average slope is 153.49nm/RU, and the standard deviation is 9.354. The slopes and the intercepts of the five doses are shown in Figure 3-43. The slopes are approximately independent of dose. Both the mean and the standard deviation are larger than those of the vertical lines. The former indicates a larger sensitivity to defocus, and the latter shows a bigger measurement noise or the internal discrete distributions of the sensitivities. The sensitivity of the horizontal line is about 0.65nm defocus/per nm of trench depth. Similar to the vertical line at dose of 18mJ/cm², there is a sharp drop of trench depth at -0.1RU, and the intercept at this dose is therefore lower than the trend expected from the large defocus. The second order polynomial fitting is used to the intercept vs. dose. The fitting R² is 0.996. By only considering the depth beyond -0.1RU at the negative direction, the intercept of the dose of 18mJ/cm² will rise about 12% of the intercept in Figure 3-43 (b).

As we know the information of the measured probe trench depth, the linear function of the slope and the defocus, and the intercepts, the predicted defocus, are obtained by solving the Equation (3-37). The results are shown in Figure 3-44. The correlation between the predicted defocus and the scanner programmed defocus are close to one without significant shift. The maximum standard deviation of the defocus spread is 6.0nm, which indicates in the measured FEM matrix, the maximum cross field defocuse variation in horizontal direction is 6.0nm.

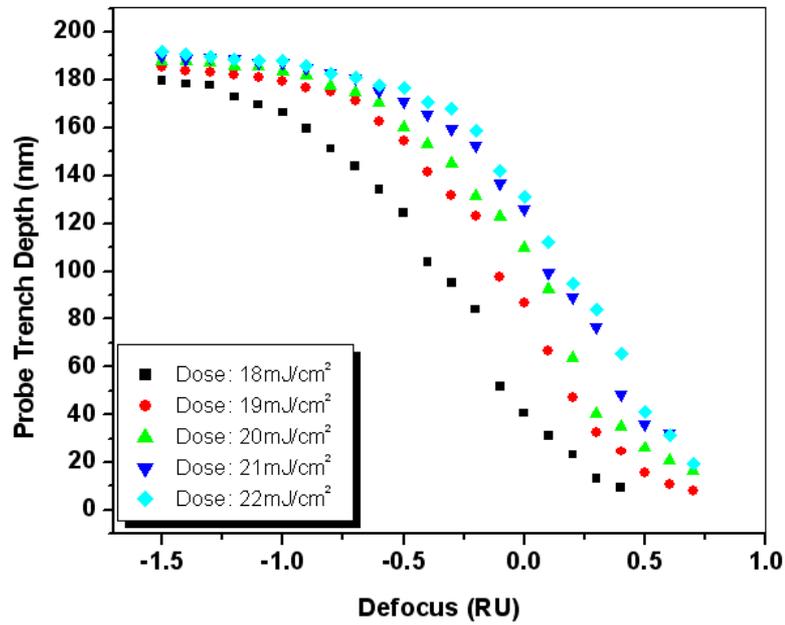


Figure (a) Horizontal probe line trench depth vs. defocus with dose from 18mJ/cm² to 22mJ/cm²

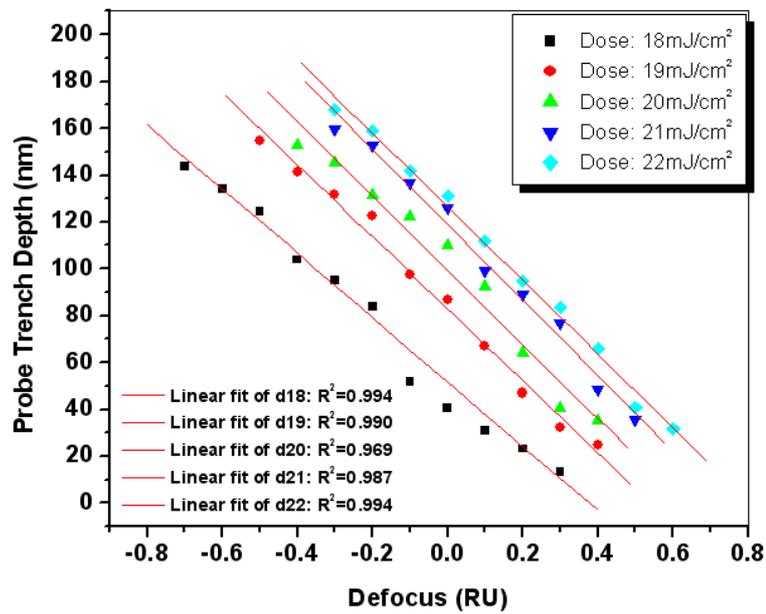


Figure (b) Linear approximation of horizontal probe line trench depth vs. defocus with dose from 18mJ/cm² to 22mJ/cm²

Figure 3-42 Horizontal probe line trench depth measurement and calibration results

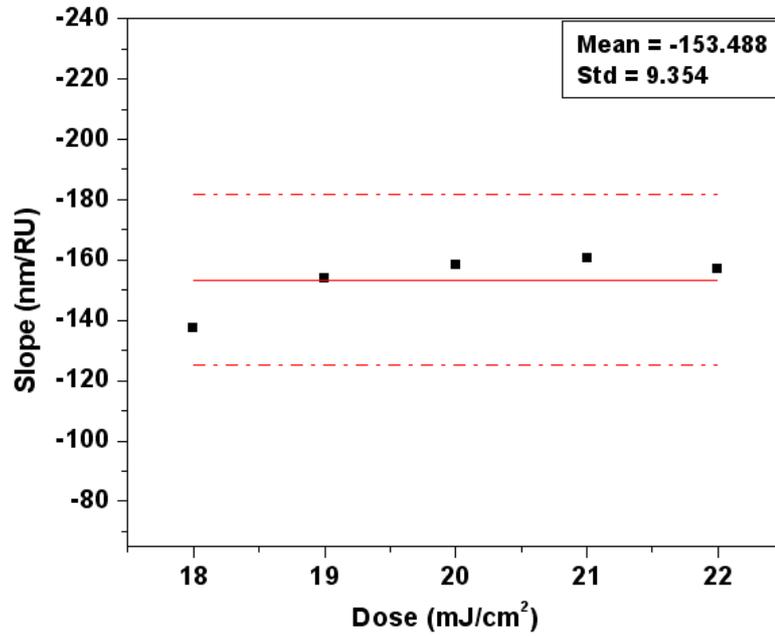


Figure (a) Slope of horizontal probe line trench depth vs.defocus

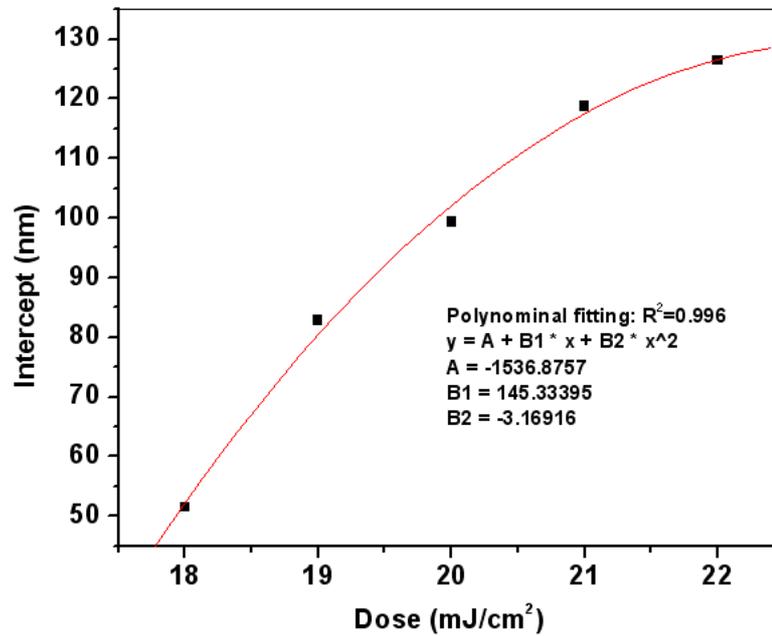


Figure (b) Intercept of horizontal probe line trench depth vs.defocus as a function of dose

Figure 3-43 Analysis of the claibration results of horizontal probe line trench depth vs. defocus

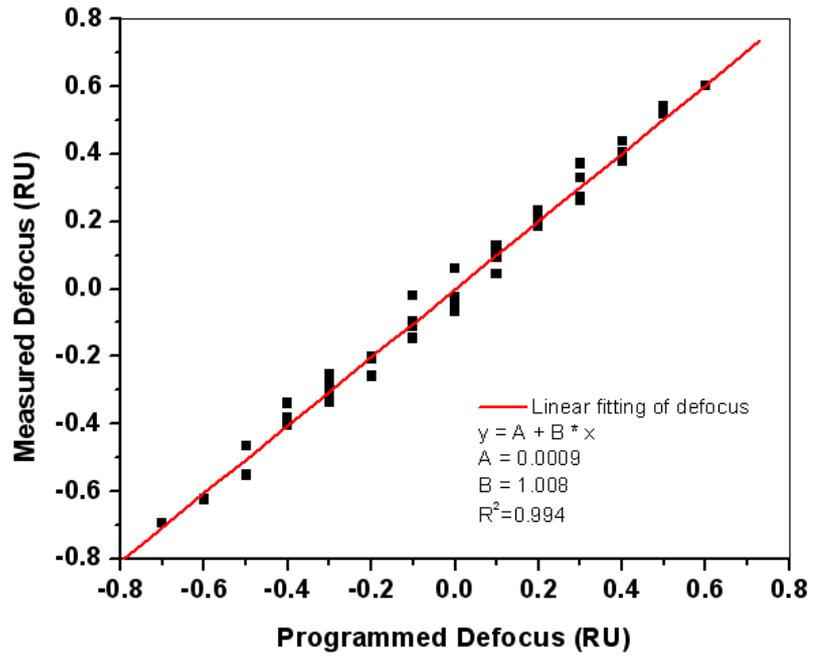


Figure 3-44 Horizontal line measured defocus vs. scanner programmed defocus

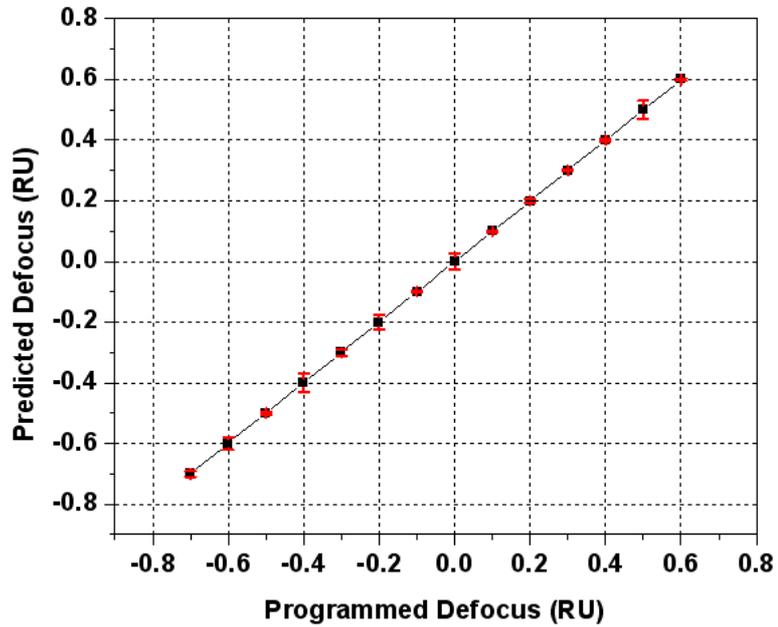


Figure (a) The calculated defocus offsets on programmed defocus. The single variation value is extended to two-side error bar in the plot.

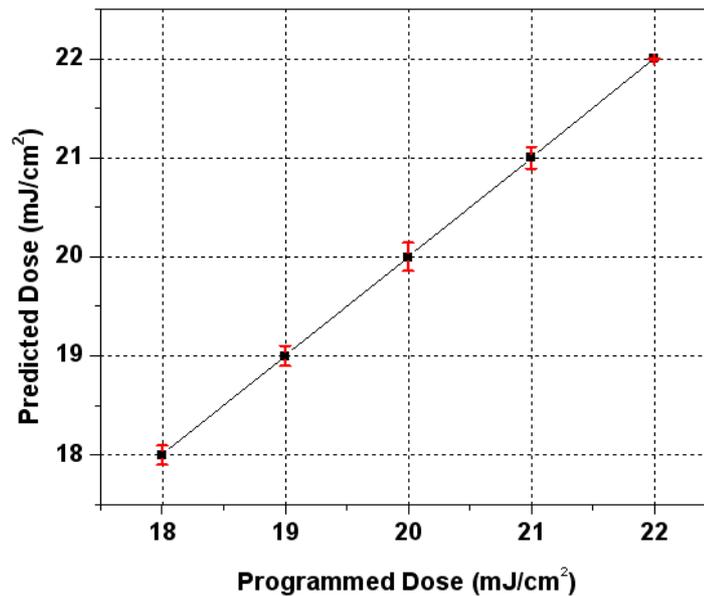


Figure (b) The calculated dose offsets on programmed dose. The single variation value is extended to two-side error bar in the plot.

Figure 3-45 The predicted defocus and dose offsets for horizontal gratings

The defocus and dose variations are calculated by solving Equation (3-43) and (3-42). The focus range is from -1.0RU to 0.6RU in the linear approximation. There are 49 measurements involved in building the model. Figure 3-45 (a) is the defocus variations. The single variation value is extended to two-side error bar in the plot. The maximum defocus offset to the programmed defocus is 3nm, around 6% of 0.5RU. The mean offset is 0.17nm, and the standard deviation is 1.69nm. The defocus residual of the measured defocus and the model defocus has the mean of -0.16nm and standard deviation of 0.28nm. Figure 3-45 (b) is the dose variations. The maximum dose variation is 0.143mJ/cm², about 0.7% of 20mJ/cm². The mean dose offset is around 0, and the standard deviation is 0.11mJ/cm².

Finally the defocus offset differences along the horizontal dies and the vertical dies are calculated. The defocus offset difference as a function of the position across the x-dimension of the fields is shown in Figure 3-46. The results indicate that the average defocus difference along the horizontal direction across the five fields is around 0.46nm. The defocus offset between the vertical and the horizontal measurements as a function of the position across the y-dimension of the fields is shown in Figure 3-47. The results indicate that the average defocus difference along the vertical die direction cross 14 dies is around 2.2nm

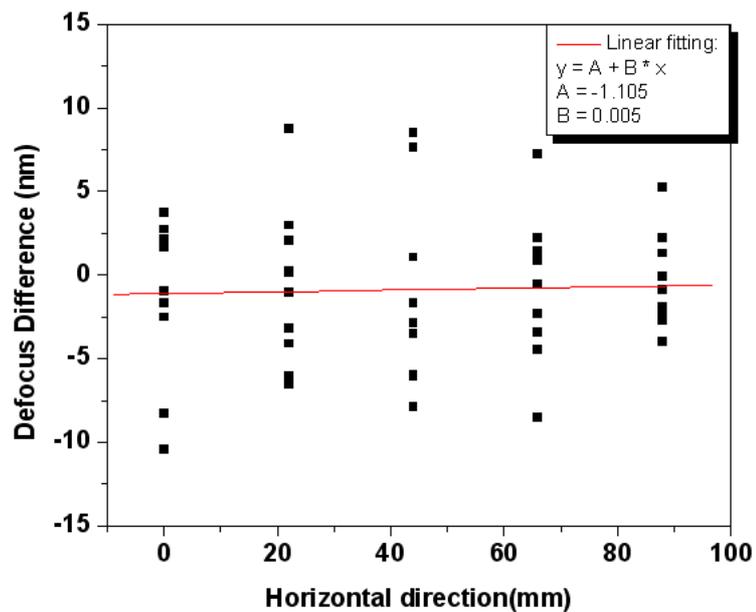


Figure 3-46 Defocus difference between vertical line and horizontal line along horizontal directions of wafer fields. The average difference is 0.46nm.

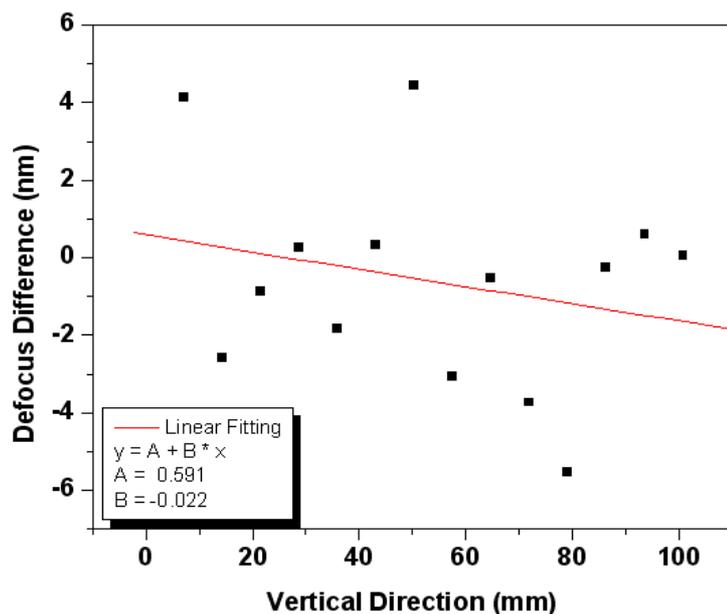


Figure 3-47 Defocus difference between vertical line and horizontal line along vertical directions of wafer fields. The average difference is 2.2nm.

3.4.3 Analysis of Experiment and Simulation Results

3.4.3.1 Defocus Sensitivity

The sensitivity value of the PPG detector for both the vertical and the horizontal gratings is below 1nm defocus / per nm of trench depth. In other words, one nanometer trench depth measurement reflects smaller than one nanometer defocus variation. This sensitivity is higher than the known defocus monitors based on phase shift mask techniques, e.g., Phase-Grating Focus Monitor (PGFM) sensitivity of around 3nm defocus /nm overlay, and the other defocus monitors based on the binary mask design.

The measurement defocus sensitivity of the horizontal line is around 1.3 times larger than that of the vertical line. We suspect that the coma and astigmatism aberrations, and illumination angular non-uniformity across the field cause the defocus difference in the

horizontal and the vertical directions. Limited by the layout dimensions and the measurement data, we only have three measurements for the same pattern in one field, which are shown in Figure 3-34. Site one is at the left edge of the field, site two is close to the middle of the field, site three is close to the right of the field.

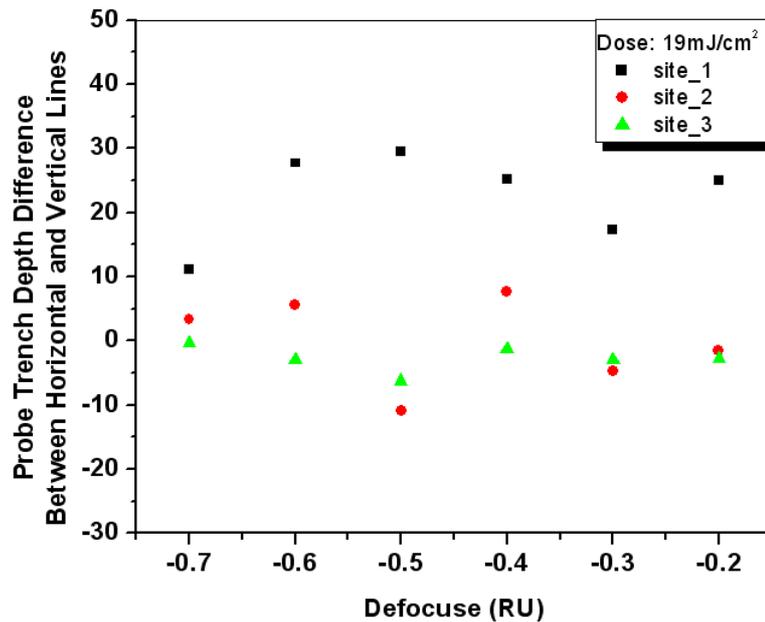


Figure 3-48 Measurement difference of probe trench depth between the horizontal line and the vertical line for dose 19mJ/cm².

The measurement difference between the vertical and the horizontal gratings for the three sites at dose 19mJ/cm² are shown in Figure 3-48. The measurement difference in site one is much bigger than those in site two and site three, but there is no clear trend from site two to site three. Similar phenomena are seen in other dose conditions. Since the three measurements are averaged to get the slopes of the vertical and the horizontal

gratings, the sensitivity difference between the vertical and the horizontal gratings is expected to mainly come from the site one measurements. So the coma and astigmatism aberrations, and illumination angular non-uniformity across the field are expected to be the main reason.

3.4.3.2 Dose Sensitivity

The intercept of the probe trench depth vs. defocus represents the trench performance at zero defocus with dose. The intercept of both the horizontal line and the vertical line are drawn in Figure 3-49. The average intercept difference is 2nm, with no clear trend between the vertical and the horizontal lines. Therefore the dose performance is expected to be independent of the grating directions.

In Figure 3-39(b) and Figure 3-43 (b), a second order or higher order polynomial are used to fit the intercept vs. dose. We found that, in Figure 3-38(a), there is a big drop at zero defocus of dose $18\text{mJ}/\text{cm}^2$, possibly due to scanner mechanical defocus deviations, the noise during the ellipsometer measurement, or the standing wave effects to make the trench open sharply. At smaller dose, the resist depth response to the defocus at the transition region from negative to positive may not be consistent, and the measurement fidelity thereby degrades. By not considering the defocus from zero to positive direction of dose $18\text{mJ}/\text{cm}^2$, we get a higher intercept of 69.77 nm, and the slope of 95.75nm/RU. The new intercept value is marked in Figure 3-50 with open squares. A good linear approximation between the intercept and the dose is approached.

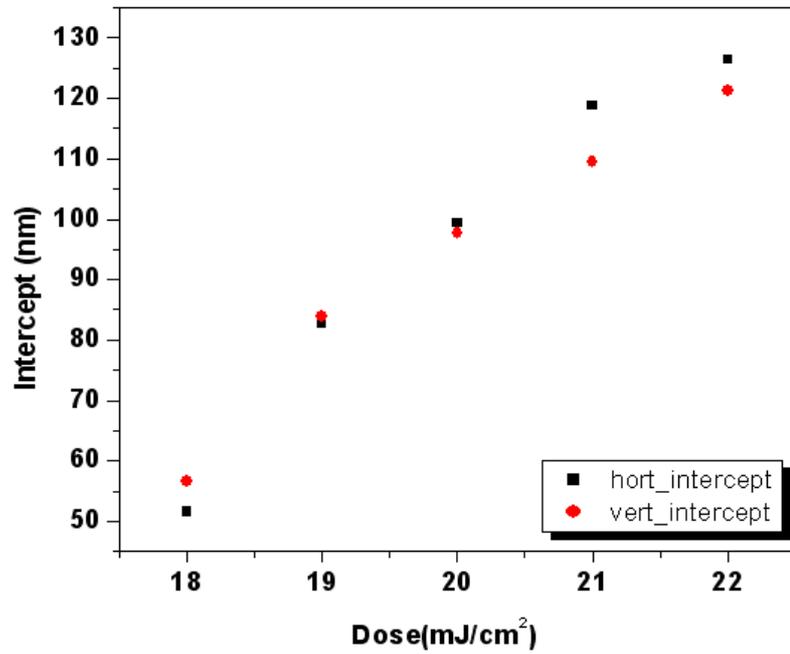


Figure 3-49 Intercept vs. dose for both horizontal line and vertical line.

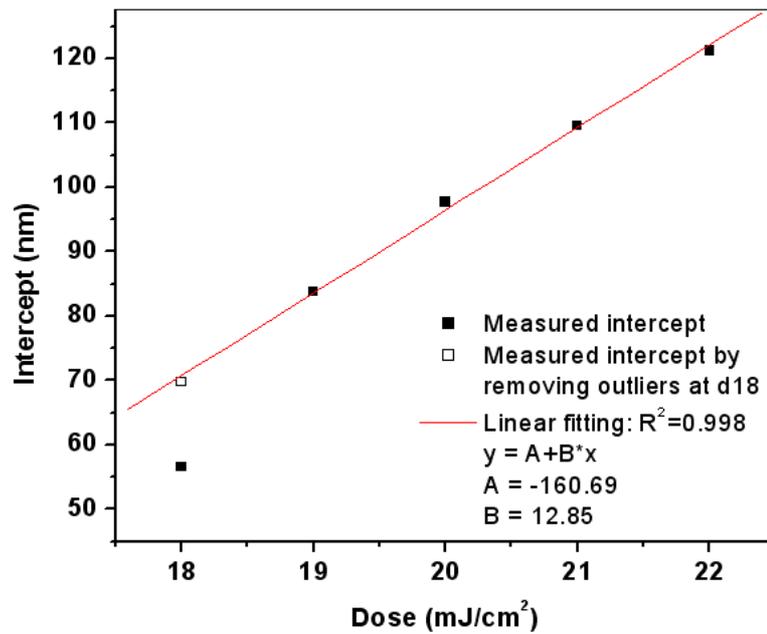


Figure 3-50 Analysis of the intercept at dose 18mJ/cm²

3.4.3.3 Line Width Effects

In table 3-3, we show that the PPG sensitivity increases as the probe line width. The probe line electric field is simulated in SPLAT with varying line width from 38nm to 78nm. No pattern line is involved. As shown in Figure 3-51, the probe line electric field increases with line width. The shape of electric field vs. defocus with variant line width is similar.

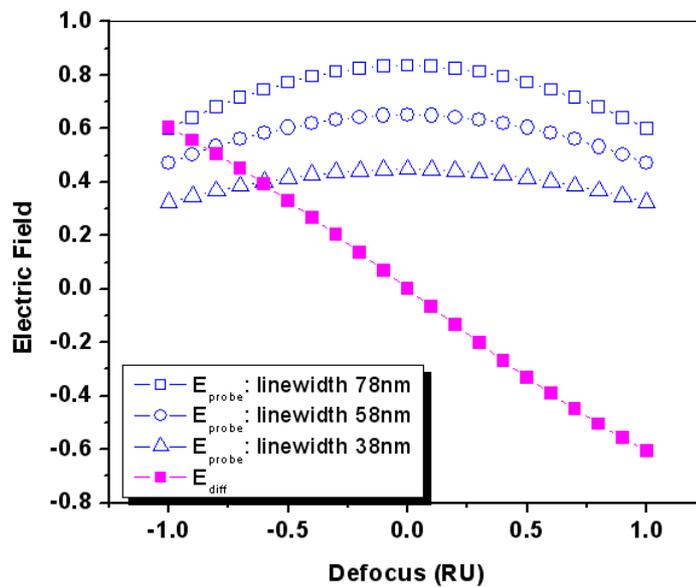


Figure (a) Probe line electric field vs. defocus at variable linewidth. For the E_{probe} , only the probe line is included in the simulation. The E_{diff} is the defocus spillover from the pattern line to the probe line with a fixed probe linewidth of 86nm. The E_{diff} is extracted from Figure 3-25.

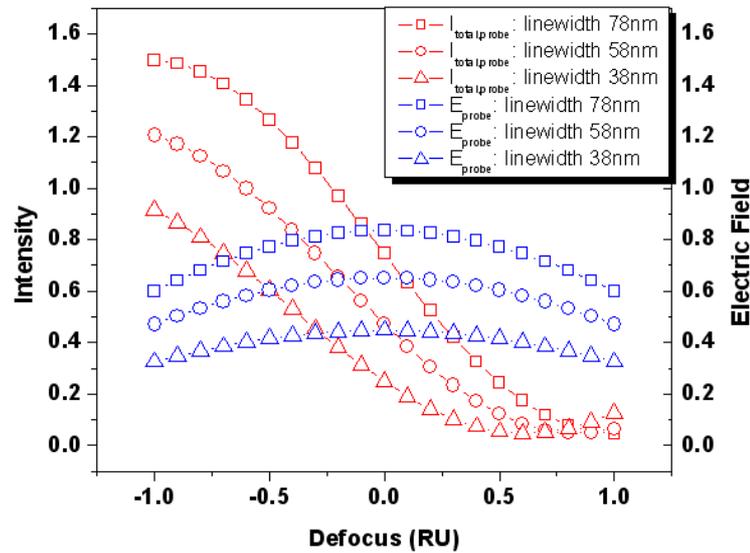


Figure (b) The red symbols mark the total intensity of the probe line with the variable linewidth but the fixed pitch and pattern linewidth. The blue curves are the electric field only the probe line.

Figure 3-51 Analysis of probe linewidth effects on the PPG sensitivity

With the fixed probe line width and the pitch, the defocus spill over from the pattern line to the probe line is approximately constant. The defocus spill-over is extracted from Figure 3-25, which has the same phase as the probe line electric field. So a higher electric field is expected at the imaginary axis for the larger probe line width. The total intensity of the PPG at the center of the probe line is shown in 3-51(b). Since the intensity at the 90° phase dominant over the 0° proximity effect, the total intensity and the PPG sensitivity increase with the probe line width. Table 3-3 shows that at $\sigma=0.1$, the sensitivity increases/decreases around $0.05(CF/RU)/10nm$. At partial coherence of $\sigma=0.34$, the sensitivity is expected to increase/decrease about $0.025(CF/RU)/10nm$, which is 6% of the peak sensitivity. Therefore the PPG has a good tolerance to the mask feature size variations.

3.4.3.4 Partial Coherence Effects

In section 3.3.2.1, we discussed the partial coherence effect on the design of the PPG probe line and pattern line width. The periodic grating has the stronger proximity effect at the smaller partial coherence. For the pitch of the grating equals to $1.11\lambda/NA$, the mutual coherence increases from 0 to around 0.94, as the partial coherence reduces from 1.0 to 0.1. The mutual coherence is about 0.44 at the experimental illumination of $\sigma=0.34$. The smaller partial coherence causes the probe line base intensity at the nominal condition to increase at the same probe line width.

In Figure 3-16 and Figure 3-17, we show that the PPG sensitivity decreases as the partial coherence. The PPG electric field as the vector addition of $\vec{E}_{PPG} = \vec{E}_{pb} + \vec{E}_{df} + \vec{E}_{pax}$ and the magnitude of the intensity as $I_{PPG} = (E_{pb} + E_{df})^2 + E_{pax}^2$ are expressed in Eq. (3-33) and Eq. (3-34). By taking into account of the partial coherence effects, the intensity can be re-expressed as:

$$I_{PPG} = E_{pb}^2 + \mu_{pb,pa} E_{pb} E_{df} + E_{df}^2 + E_{pax}^2 \quad \text{Eq. (3-44)}$$

Where $\mu_{pb,pa}$ is the mutual coherence function of the probe line and the pattern line. E_{df} can be written as a function of defocus f , e.g. $E_{df} = -0.63 \cdot f$, as shown in Figure 3-25, so the sensitivity of I_{PPG} to defocus f is mostly determined by: $-0.63\mu_{pb,pa}E_{pb}$. As partial coherence increases, the mutual coherence factor decreases from one, and the PPG sensitivity to defocus is therefore attenuated. It is note that the E_{pb} may also be impacted by the partial coherence value since the three wave interference may change to two wave interference as described in Figure 3-24.

3.4.3.5 High NA and Mask Edge Effects

The exposure system in the experiment has an NA of 0.85, so the maximum angle of the rays reaching the image plane is 58.2° . Referring to Figure 3-52, the TM mode in the electric field has the $E_{TM,xy}$ and $E_{TM,z}$ components, the former constitutes 52.7% of E_{TM} , and the latter has 85% of E_{TM} . The $E_{TM,xy}$ contributes the contrast in the image plane. The energy in the $E_{TM,z}$ provides the background intensity, which degrades the image contrast. Considering the high NA effect, the electric field spill over from the pattern line to the probe line will be attenuated, which will impact the PPG sensitivity.

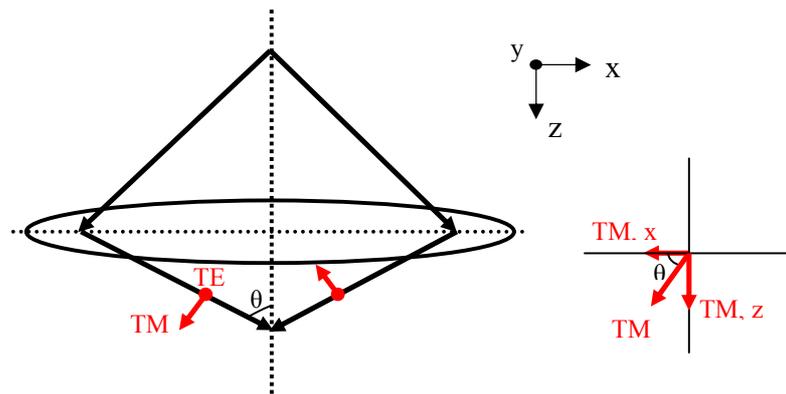


Figure 3-52 Illustration of the high NA conditions. TM mode decreases the image contrast.

The phase shift mask edge effects on the order of 10% of a wavelength was first proposed by A. Wong [3-23]. The electromagnetic vector effects associated with the etched sidewall cause less intensity of the opening than the non-phase shift ones of the same size. For 180° line, A. Wong indicated that the bias necessary to compensate the edge effects is $0.021\lambda/NA$ per edge. In the PPG mask design, we extend 10nm at each edge of the 90° probe opening to address this problem.

3.4.3.6 Resist Effects

The experiment shows consistent trend but different sensitivity values than those predicted by the simulations. One of the major reasons of the difference comes from uncertainties in the resist simulations. The Shipley 193nm resist is used in both the simulations and the experiments. The refractive index of the BARC and the resist are recommended by the Panoramic EN-suite and the Timbre ODP platform for the general 193nm Shipley resist. However, the resist parameters in the experiment are possibly different with those in the simulations. We are not able to measure the refractive index in the experiments. Since the refractive index impacts the transmission, absorption, and reflection of light inside the resist, at the interface of resist/ambit, resist/BARC, and BARC/substrate, the unknown parameters of the experimental resist will cause the mismatch of the experimental results and the simulation results.

The focus plane z_{img} inside the resist is determined by finding a plane with the best expected image performance, as shown in Figure 3-28. In the simulation we use 70% of the resist thickness as the best image plane. However, because of the acid and the base diffusion in the x, y, and z direction, and the image blur along z-direction caused by the wavelength aberration, the determination of z_{img} from the ideal optical condition may not be consistent with the real process.

The diffusion-reaction and develop equations in the Panoramic EM-suite are used to calculate the resist image. The parameters of the resist diffusion length, the develop rate, fudge factor, and the development time in the development, the kinetic and the temperature parameters in the post exposure bake employ the suggested values of the

Shipley 193nm resist. These parameters are not optimized for the resist process in the experiments.

Another uncertainty in the resist simulation is the exposure dose. Either the clear field dose and the dose scaling factors in the simulation is not the same as that in the experiments, which causes the difference of the detailed values of the defocus sensitivity at the individual dose and the overall dose performance. However, the performance trends of the PPG probe line trench depth to defocus and dose are consistent in both the simulation and the experiment.

3.4.3.7 ODP Effects

Timbre Technologies' Optical Digital Profilometry (ODP) is the simulation platform of the Specular Spectroscopic Scatterometry to determine the PPG probe line trench depth. In the common implementation scheme, ODP captures the diffracted broadband light from a diffraction grating and matches the spectra to a set of previously generated profile-spectra pairs in a library file. In order to create the library, Timbre ODP – NOW software regressively determine the solution for the manufacturing process and the parameters of the device profile. In our experiments, due to the limited requirement of one wafer application, the metrology results are the results of the ODP-NOW regressions. Although a different categorical approach from the large-scale library matching techniques, the regression results are a reliable substitute [3-24].

The complexity of the model increases as more floating variables are contained in the model. In Figure 3-37, seven main parameters are used to describe the PPG resist profile.

Since a high correlation, about 90%, between the top CD and mid CDs, we define the mid CDs as the dependent variables to the top CDs, and fix the thickness ratio between top CDs to mid CDs. Similarly, the BARC thickness is set as a fixed value in the simplified model. The refractive indices of the resist and the BARC employ the default value of the ODP software. As discussed in section 3.4.2, the double and triple trapezoid models are used to approximate the real standing wave effects. Both the simplified model and the uncertainty of the material are expected to impact the goodness of fitting in the regressions and the unique solutions of the device structures.

The PPG profile in the ODP simulation is constructed by matching the simulated spectrum of the device model and the experimental spectrum. A good matching between the spectra is critical to build the device. In order to achieve the unique solution, a high sensitivity of the simulated spectrum over the different profiles is required. One of the main problems to achieve the uniqueness is that two or multiple similar device profiles may lead to the similar diffraction responses. In the PPG device, the probe line and the pattern line have the close line width and the same pitch, although the defocus responses of the two lines have a reversed trend, as shown in Figure 3-53, at the certain range, the probe line and the pattern line may be confounded in the ODP simulations. The gray area in Figure 3-53 indicates the range in which both the probe line trench depth and the pattern line trench depth are close to and below the crossover depth of the two lines. This is a low-sensitive region to the probe line and the pattern line in the PPG simulations. The resist simulation settings in Figure 3-53 are different to those in Figure 3-39, but the PPG responses to defocus keep the same trend. We also find that the non-uniqueness

phenomenon occurs when the probe trench depth is smaller but close to 100nm, which is in the near-crossover region.

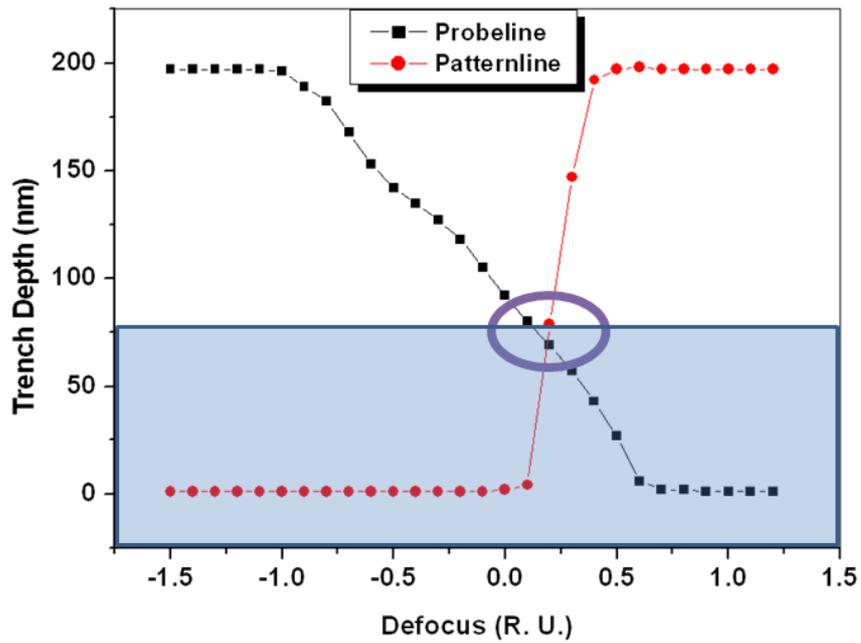
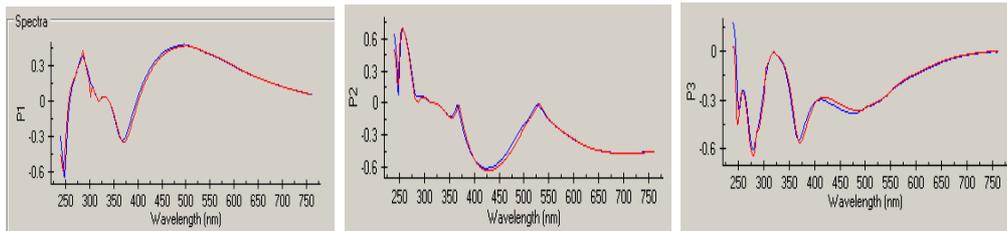


Figure 3-53 Illustration of the PPG probe line and pattern line behavior through defocus.

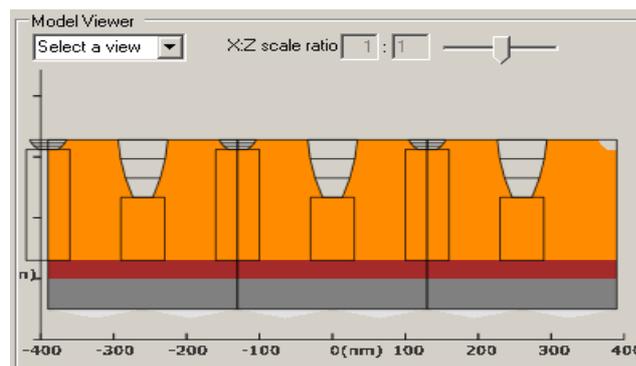
The ODP simulation spectra and the constructed device profile of the PPG vertical grating at dose $19\text{mJ}/\text{cm}^2$, and defocus -0.1 RU , are shown in Figure 3-54. Figure (a) is the first simulation spectra of the device model compared to the experimental spectra collected by the spectroscopic ellipsometer. The goodness-of-fit of the spectra matching is 0.9815. Figure (b) is the simulated device profile. The probe trench depth is 94.4nm , and the pattern trench depth is 15.9nm .

For the same pattern, we perform the second simulation, as shown in Figure (c) and Figure (d), with decreasing the initial probe depth value of 20nm in the model regression. The goodness-of-fit of the spectra is 0.9815. The reconstructed PPG probe trench depth is

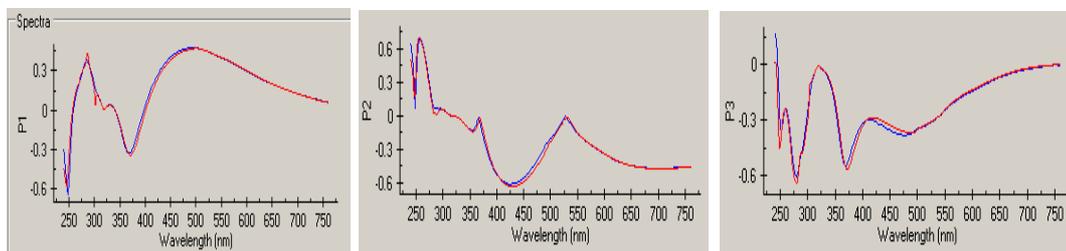
15.6nm, and the pattern trench depth is 92.4nm. The probe line and the pattern line placements are reversed in the second simulations with the similar fitting results as the first simulation.



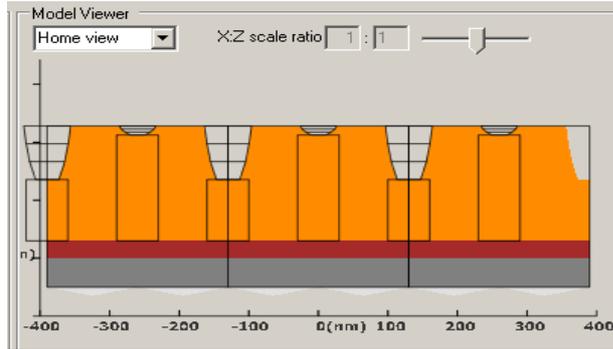
(a)ODP simulated spectrum vs. experiment measured spectrum in the first simulation



(b)ODP simulated device profile in the first simulation



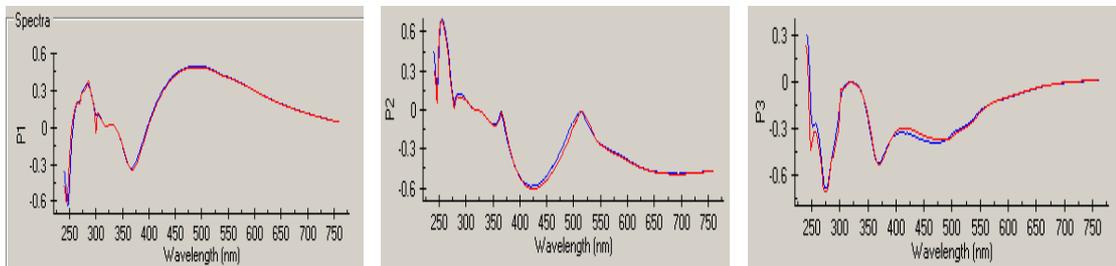
(c)ODP simulated spectrum vs. experiment measured spectrum in the second simulation. The pattern is the same as that in Figure (a) and Figure (b)



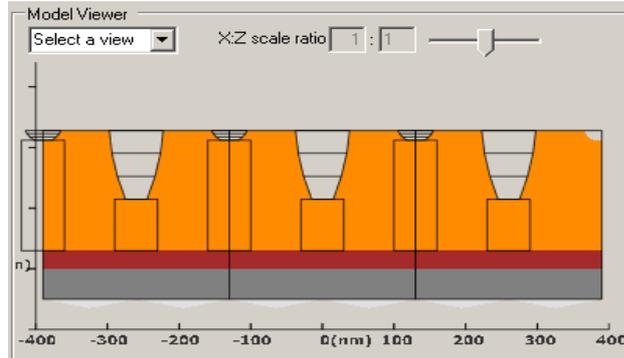
(d) ODP simulated device profile in the second simulation. The pattern is the same as that in Figure (a) and Figure (b)

Figure 3-54 Analysis the sensitvity of the ODP simulation to the PPG pattern.

When the trench depth is bigger than 100nm, the simulation results are quite repeatable with an unqiue solution. The deeper trench simulations also provide the clues to determine the trend of the low sensitive region from multiple solutions. Figure 3-55 shows the simulation spectra and the device structure of the pattern at dose 19mJ/cm², and defocus -0.2 RU. The goodness-of-fit is 0.9824. The probe trench depth is 113.3nm. The simulation results at defocus of -0.2RU provide an evidence of an appropriate choice from the two deviec profiles at defocus of -0.1RU. The first simulation in Figure 3-54 has the consistent performace with the neighbored defocus conditions.



(a)ODP simulated spectrum vs. experiment measured spectrum



(b) Constructing the PPG profile by back-forward calculation of the ODP spectrum

Figure 3-55 Analysis of the sensitivity of the ODP simulation

3.5 Conclusions

In this chapter, we present a new highly sensitive Probe-Pattern Grating (PPG) Defocus Monitor. The monitor is a periodic grating with a ninety degree phase-shifted probe line at the center of the pitch, and a zero degree pattern line at an optimized position of the strong defocus spillover to the probe line. The aerial image electric field and the energy transmitted to the resist will be modulated by the defocus aberration. The defocus variations are therefore translated into the probe line trench depth variations. A good linear approximation of the probe trench depth versus defocus is obtained in the Bossung plot.

Specular Spectroscopic Scatterometry is an advanced process metrology techniques, which can addressing the subtle variations at small geometries. Timbre ODP is the scatterometry software platform used in the experiment calibration. The ODP construct the PPG device profile by matching the simulated spectra from the device model and the experiment measured spectra from the spectroscopic hardware.

We perform the PPG printing experiments by the ASML 193nm dry lithography scanner at Spansion Inc. The diffraction spectra are measured by spectroscopic ellipsometer at Spansion. The calibration software package is the Timbre ODP platform. A linear relationship between the probe trench depth and the defocus is attained in the range of -1RU to +0.6RU. By reversely solving the linear models, the PPG predicted defocus is obtained. An excellent correlation between the measured defocus and the scanner programmed defocus is achieved. For the vertical grating, the mean defocus offset is 0, and the maximum standard deviation of the defocus spread is 4.2nm. For the horizontal grating, the mean offset is around 0, the maximum standard deviation is 6nm. The residuals of the measured defocus and the modeled defocus have the mean of 0.7nm, and the standard deviation of 0.73nm for the vertical grating, and the mean of -0.15nm, and standard deviation of 0.28nm for the horizontal lines.

The PPG sensitivity value is smaller than 1nm defocus / per nm of trench depth, which indicates a higher sensitivity than the known defocus monitors based on phase shift mask techniques of around 3nm defocus /nm overlay, and the other types of defocus monitor with the binary mask design. The defocus sensitivity of the horizontal line is around 1.3 times larger than that of the vertical line. The main reason of the difference is from the site one measurements. The coma aberration, astigmatism aberration, and the illumination angular non-uniformity are expected to be the main reasons.

We use the intercept of the Bossung plot to describe the PPG dose performance. A quadratic function of intercept vs. dose is approached. The dose response is consistent between the vertical and the horizontal gratings. With removing the outliers, a good linear approximation can be accessed. The mean dose offset is close to 0, the standard

deviation is $0.16\text{mJ}/\text{cm}^2$, and the maximum dose offset for the vertical line is $0.244\text{mJ}/\text{cm}^2$. For the horizontal line, the mean dose offset is around 0, and the standard deviation is $0.11\text{mJ}/\text{cm}^2$, and the maximum dose variation is $0.143\text{mJ}/\text{cm}^2$.

Besides the measurement noise, seven effects are analyzed to impact the simulation and the experiment results. The High NA effect, mask edge effect, and the resist effects are the main reasons to cause the difference between the simulation and the experiment results. The astigmatism and coma aberrations within the field, the linewidth effect, the partial coherent effect, and the ODP sensitivity effect are expected to impact the experiment results. As a side effect of the high sensitivity to defocus variations, the measurement deviations and noises will affect the measurement precisions. The devices at site one of the die edge and at the smaller dose of $18\text{mJ}/\text{cm}^2$ majorly suffer this effect.

PPG defocus pattern is an interferometric monitor and thus has a strong linear response to the defocus aberration with a very high sensitivity. An important implication of the high sensitivity is that the direction and the amount of drift in focus or of a tool parameter can be identified from a single exposure after the linear process window model is determined. Therefore the PPG monitor has the potential to facilitate run-time process monitor and adjustments directly from production wafers.

References

- [3-1] B. J. Lin, "The k_3 coefficient in nonparaxial $1/NA$ scaling equation for resolution, depth of focus, and immersion lithography", *Journal of Microlithography, Microfabrication, and Microsystems*, Vol. 1, No. 1, pp7, 2002
- [3-2] J. W. Goodman, "Introduction to Fourier Optics", 2nd edition, *McGraw-Hill*, 1996
- [3-3] T. Brunner, D. Corliss, "Laser bandwidth and other sources of focus blur in lithography", *J. Microlith, Microfab, Microsyst*, vol. 5(4), 043003, 2006
- [3-4] T. A. Brunner, "Impact of lens aberrations on optical lithography", *IBM J. RES. DEVELOP.*, Vol 41, No. 1/2, 1997
- [3-5] T. A. Brunner, et al., "Quantitative stepper metrology using the focus monitor test mask", *Proc. of SPIE*, Vol. 2197, pp.541, 1994
- [3-6] G. M. Pugh, "Detailed study of a phase-shift focus monitor", *Proc. of SPIE*, Vol. 2440, pp. 690, 1995
- [3-7] H. Nomura, "New phase shift gratings for measuring aberrations", *Proc. of SPIE*, Vol. 4346, pp. 25, 2001
- [3-8] B. La Fontain, et al., "study of the influence of substrate topography on the focusing performance of advanced lithography scanners", *Proc. of SPIE*, Vol. 5040, pp.570, 2003
- [3-9] C. A. Mack, D. A. Legband, S. Jug, "Data analysis for photolithography", *Micro- and Nano- Engineering 98, Proc.*, pp. 65, 1998
- [3-10] C. A. Mack, and J. D. Byers "Improved model for focus-exposure data analysis", *Proc. of SPIE*, Vol. 5038, pp.396, 2003
- [3-11] C. P. Ausschnitt, and S. Y. Cheng, "Modeling fro profile-based process window metrology", *Proc. of SPIE*, Vol. 5378, pp.38, 2004
- [3-12] K. Lensing, et al., "Lithography process control using scatterometry metrology and semi-physical modeling", *Proc. of SPIE*, Vol. 6518, 651804, 2007
- [3-13] T. A. Brunner and C. P. Ausschnitt, "Process Monitoring Gratings", *Proc. of SPIE*, vol. 6518, 651803-1, 2007
- [3-14] C. Ausschnitt, and T. A. Brunner, "Distinguishing dose, focus, and blur for Lithography characterization and control", *Proc. of SPIE*, Vol. 6520, 65200M, 2007

- [3-15] A. R. Neureuther, S. Hotta, and K. Adam, "Modeling defect-feature interactions in the presence of aberrations", *Proc. of SPIE*, Vol. 4186, pp.405, 2001
- [3-16] G. Robins, "Interferometric Pattern and Probe Based Aberration Monitors", Ph. D. Dissertation, UC Berkeley, 2005
- [3-17] K. Toh, A. R. Neureuther, "Charaterization of Voting Suppression of Optical Defects Through Simulation", *Proc. of SPIE*, 1998
- [3-18] A. R. Neureuther, EE243 Lecture Notes, UC Berkeley, 2003
- [3-19] M. A. Miller, A. R. Neureuther, and et al., "Characterization and Monitoring of photomask edge effects", *Proc. of SPIE*, vol.6730, 67301U, 2007
- [3-20] <http://www.panoramictech.com/>
- [3-21] T. V. Pistor, "Electromagnetic Simualtion and Modeling with Applications in Lithography", Ph. D. Dissertation, UC Berkeley, 2001
- [3-22] N. Cheung, EE143 Lecture Notes, UC Berkeley, 2002
- [3-23] C. Pierrat, A. Wong, And S. Vaidya, "Phase-shifting mask topography effects on lithographic image qualify", Electron Devices Meeting, Dec., 1992
- [3-24] K. Lensing, et al., "A comprehensive comparison of spectral scatterometry hardware", *Proc. of SPIE*, Vol. 5752, pp.337, 2005

Chapter 4

Integrated Aerial Image Sensor

Lithographic performance is largely dependent on the quality of the projected aerial image. A sensor which directly measures the aerial image incident on the wafer, could provide real time evaluation of the lithography system without the time consuming steps involved in the resist process and avoid the bias imposed by the variation of resist parameters and related SEM measurements. This metrology technique would speed up the characterization and optimization of the exposure system, and be important and useful to distinguish the source of the lithography process variations at separated stages.

In this chapter, we address the design and manufacturing of a novel Integrated Aerial Image Sensor (IAIS) system suitable for integrating all the elements within the surface of an autonomous test wafer. The IAIS is mainly composed of an interference grating aperture mask and an array of CCD detectors. The aperture mask interferes with the incident aerial image to form a low spatial frequency interference pattern, which could be captured by the large pixels typically found in commercially available photo-detector arrays. Through sufficient sampling, the aerial image pattern can be reconstructed with a spatial resolution of only a few nanometers. Such a design would be useful for the analysis of sub 65-nm periodic features.

Section 4.1 overviews the aerial image monitor techniques. The concept and the design of the IAIS system are presented in section 4.2. In section 4.3, we discuss the potential

applications of the sensor. The manufacturing details and the implementation challenges are described in section 4.4. The conclusions are presented in section 4.5.

4.1 Review of Aerial Image Monitor Techniques

Several researchers and companies have developed different methods for evaluating the aerial image by indirect and direct measurements during the last three decades. An early method to measure the aerial image of a lithography exposure tool was developed by Brunner *et al.* [4-1] in 1985, known as the Stepper Image Monitor (SIM). The key technique is the fabrication of narrow fluorescent structures, accomplished by patterning sub-micrometer photo-resist structures doped with a laser dye on the special testing wafer, so that blue photons in the aerial image cause fluorescence at some longer wavelength like green. The fluorescence is monitored when the narrow lines are swept through the lithographic image profile. The fluorescent structures are fabricated with bi-level resist processing utilizing a PMMA bottom layer which has been heavily doped with Coumarin 6 dye. The basic structure is shown in Figure 4-1.

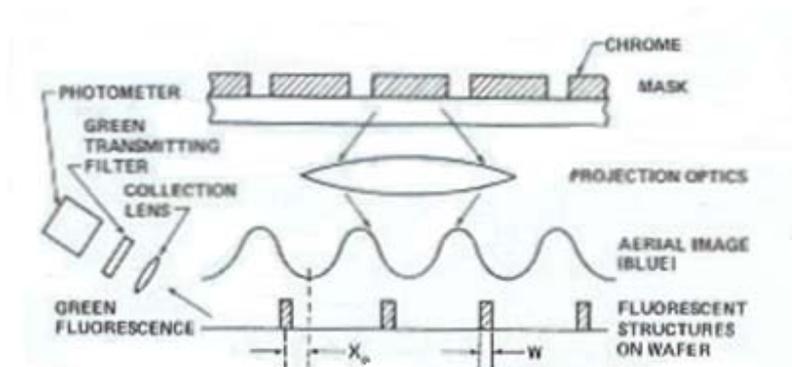


Figure 4-1 Illustration of the schematic of SIM [4-1]

While usable results were obtained with the fluorescence approach, the signal decay due to photo-chemical degradation of the dye limits this technique. The soft photo resist is not well suited for a permanent in-situ artifact wafer. To solve the problem associated with the use of photo-resist, Brunner used narrow poly lines on oxide wafers in the new version [4-2] [4-3]. The blue light scattered from the poly line is measured. Further improvements involved a brief etch of the poly line to increase roughness and improve scattering efficiency.

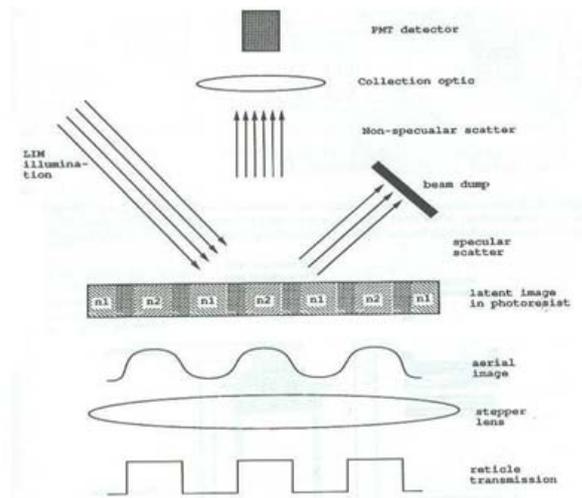


Figure 4-2 Illustration of the principle of LIM [4-4]

Latent Image Metrology (LIM) [4-4] is another interesting technique, which collects a “snapshot” in a layer of undeveloped photo resist. Since the analysis is done in the latent image thus avoiding the develop step, it eliminates the effects of resist develop chemistry, and it allows the observation of small variations in the printed pattern. Fig 4-2 illustrates the LIM technique. When a grating projected by a stepper onto a photo-resist-covered wafer, the resist changes its index of refraction only in the illuminated region due to the

photo-chemical reaction occurring there. The sharpness of the transition region between the light and dark regions is an indication of the contrast of the aerial image. After post-exposure bake, the resulting photo-resist pattern is probed with the AT&T lithographic Analysis Workstation (LAWS). It has been shown that the peak signal in the first-order of the non-specular scatter occurs when the resist pattern is exposed at best stepper focus. While LIM has many practical applications, it is not suitable for DUV resists, since the change of optical parameters is minimal and very hard to detect even after the PEB step.

The method introduced by Brunner was improved by Pfau and *etal.* [4-5]. The image monitor was implemented on the Berkeley micro-stepper. The main components of the sensor are a photodiode and a chromium coated quartz wafer with a 10x10 array of 200nm pinholes. The original DUV-capable aerial image monitor was further extended to EUV lithography systems by Charles Fields and *etal.* [4-6]. Indium Antimonide (InSb) was chosen as the photon absorber material as a compromise between its high extinction coefficient at EUV wavelength and ease of deposition. The cross-section of aperture and holding-and-scanning equipment at Sandia National Lab are illustrated in Figure 4-3. The EUV aerial image monitor was used to quantify the EUV system vibration and its impact on imaging performance.

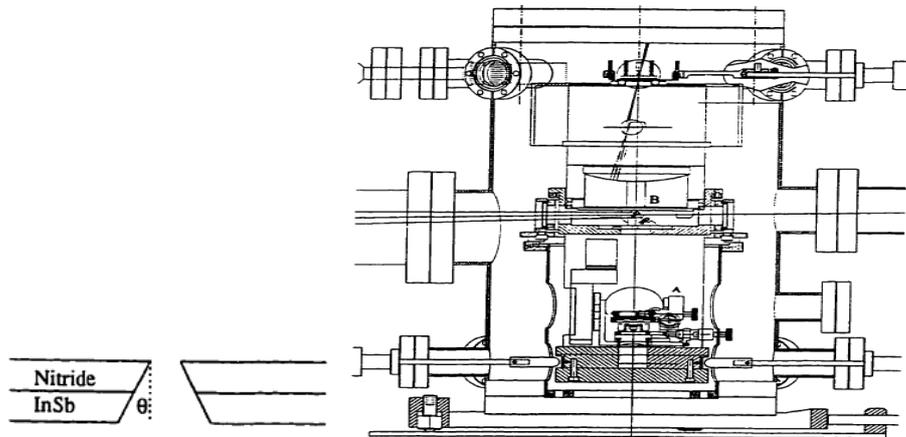


Fig 4-3 Illustration of cross-section of the aperture and the equipment of EUV aerial image monitor [4-6]

An aerial image sensor for self-calibration of wafer steppers was integrated into steppers and scanners by Nikon Corporation [4-7]. Figure 4-4 shows the schematic of the sensor. Light passing through a reticle of an aperture structure, either slit or knife-edge, is collected by the first collector lens, then passes through the optical fiber bundle, and reaches to the second collector lens before reaching a photo multiplier tube (PMT). The PMT converts the light intensity to an electrical signal, which is further processed. The experiments showed good results for best focus calibration.

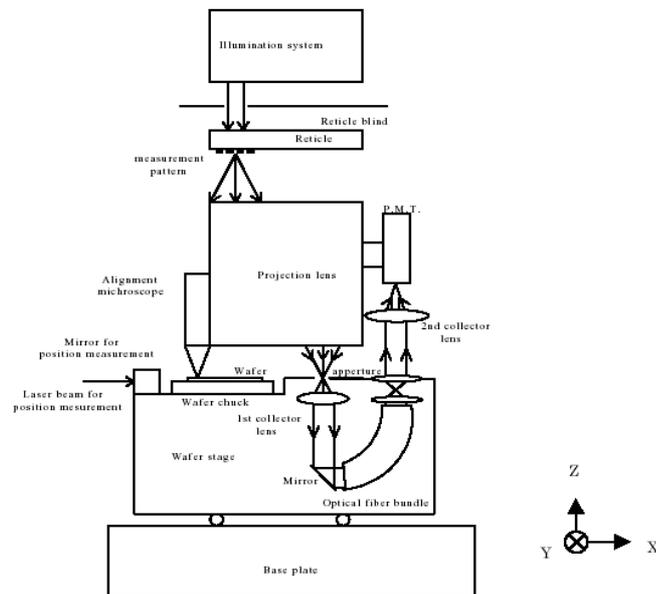


Figure 4-4 Illustration of the aerial image sensor structure in Nikon Corporation [4-7]

It is also worth mentioning the IBM Aerial Image System (AIMS), commercially known as Microlithography Simulation Microscope (MSM) [4-8]. The MSM is a UV microscope based imaging system that could measure masks under conditions which *mimic* (rather than measure) the stepper of interest. Unlike stepper optics, the mask inspection optics enlarges the reticle image onto a CCD camera, see Figure 4-5. Many of the defects on masks detected by the traditional inspection methods are not transferred to the wafer, and therefore do not affect the CD uniformity of the wafer. AIMS based on aerial image can alert the operator to defects that actually print, allowing them to ignore those that may be present on the mask but have no impact on the final result. AIMS can be used to verify that the mask is free of basic phase and other errors, to predict resist feature size for a given exposure or threshold, and to predict the magnitude of resist

perturbations due to mask defects. The AIMS/MSM has become a standard tool in mask shops.

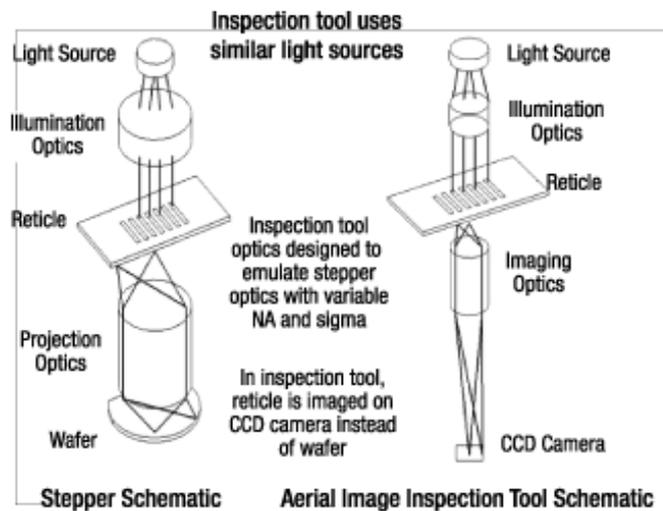


Fig 4-5 Illustration of the basic concept and structure of AIMS (MSM) [4-8]

While many ingenious techniques have been suggested and used today, all the methods involve the integration of the aerial image sensor to the exposure tool or require additional complicated measurement instruments. The application of the sensor system is thus strictly dedicated to the projection equipment. Problems associated with equipment modification or build-up will strongly influence the measurement stability and reliability. In addition, the money and time spent to modify equipments by installing these types of dedicated sensors, are not insignificant. To improve the performance of the aerial image measurement and to extend the practicability of the sensor applications, we propose a novel aerial image sensor that integrates most elements of the system on the testing wafer without the costly and complex sensor-on-equipment integration process. This sensor is

meant as a lithography processing monitor, in order to facilitate advanced process control and diagnosis.

4.2 Design of the Integrated Aerial Image Sensor (IAIS)

4.2.1 IAIS System and Interference Sampling Concept

The main components of IAIS, as shown in Figure 4-6, are the on-wafer interference grating, which we refer to as the *aperture mask*, and an array of photo-detectors. The aperture mask is an opaque contact mask within near field distance from the photo detectors. With a series of spatial phase shift aperture groups, the grating modulates and gates the radiation of the aerial image to the photo detectors underneath. This on-wafer aperture mask interferes with a periodic aerial image, ultimately forming a low spatial frequency interference pattern. This pattern is sufficiently large to be captured by the array of μm -scale photo-detectors. Through adequate sampling, periodic aerial image features in the order of a few nanometers can be extracted from the interference patterns. The aperture mask and the respective photo detector array are designed to sample an entire period of the aerial image, assuming a stationary wafer stage. In order to increase light transmission, multiple apertures are opened at the same spatial phase position in one aperture group over one or several detector pixels.

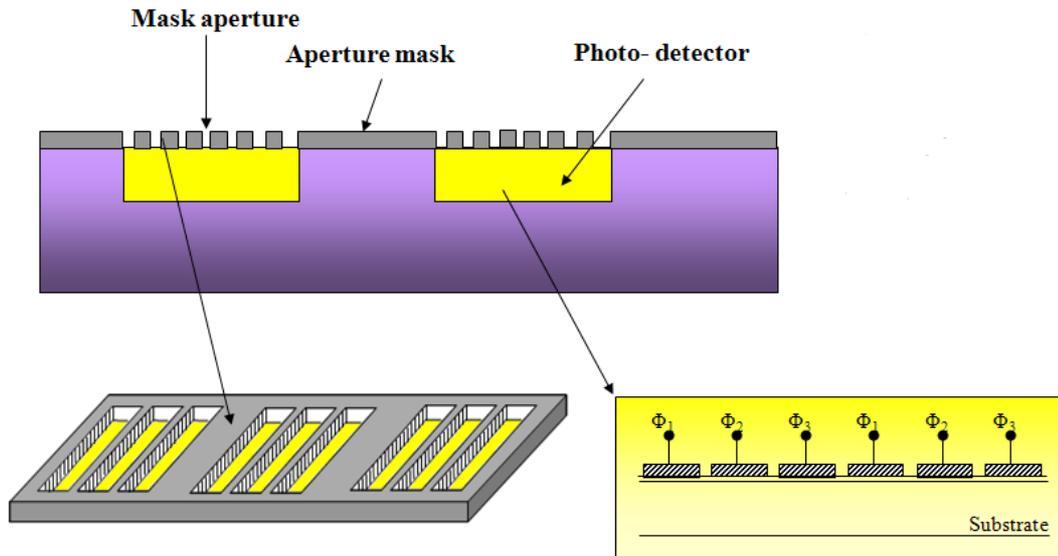
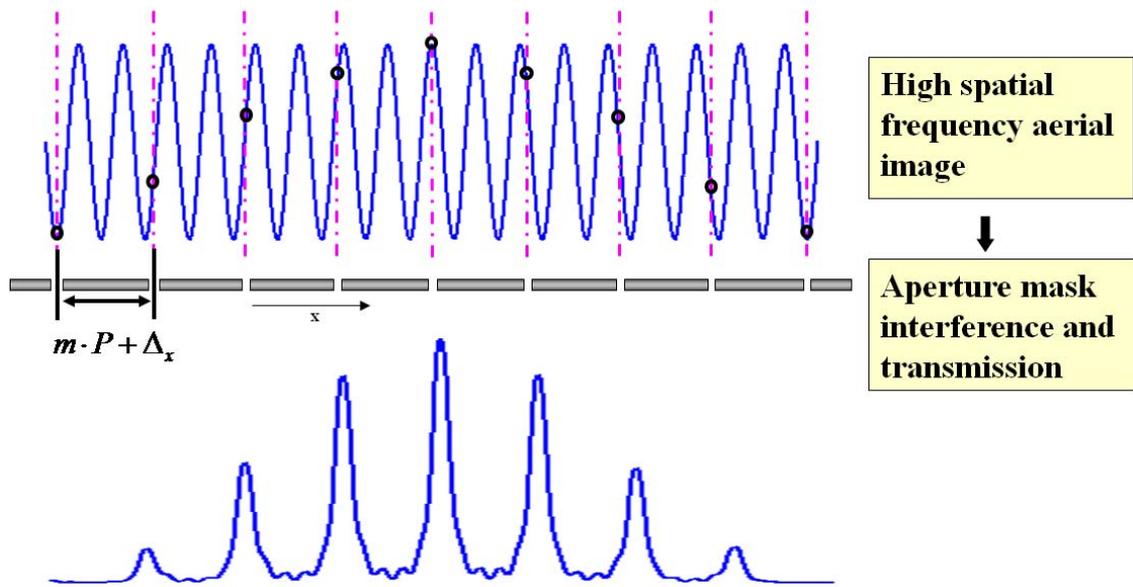
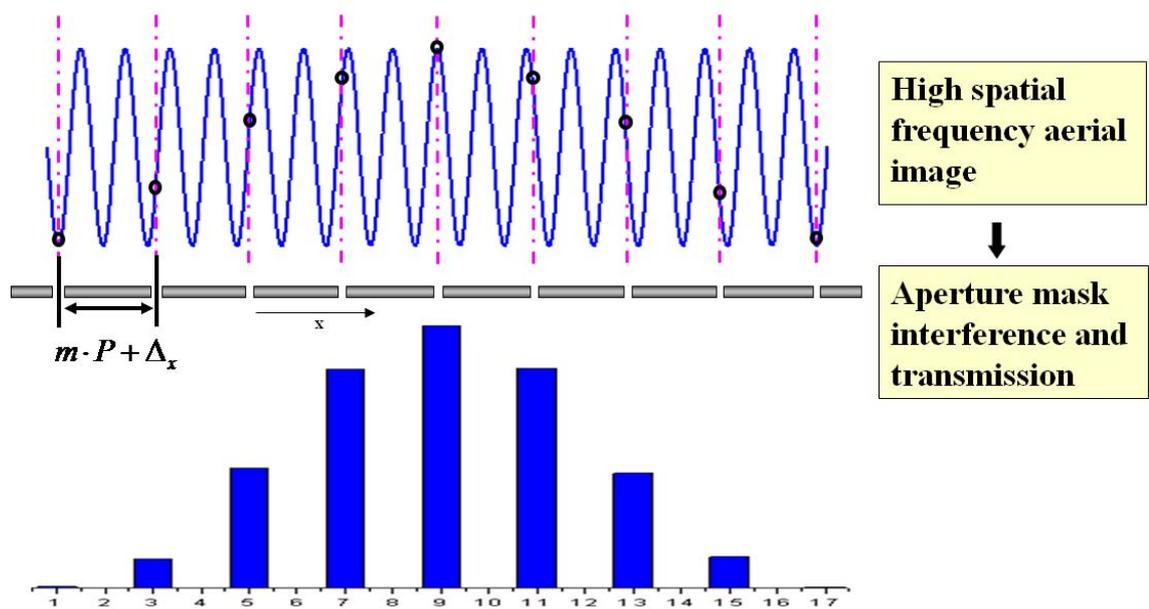


Figure 4-6 Illustration of the main components of IAIS. The aperture mask is an interference grating pattern.

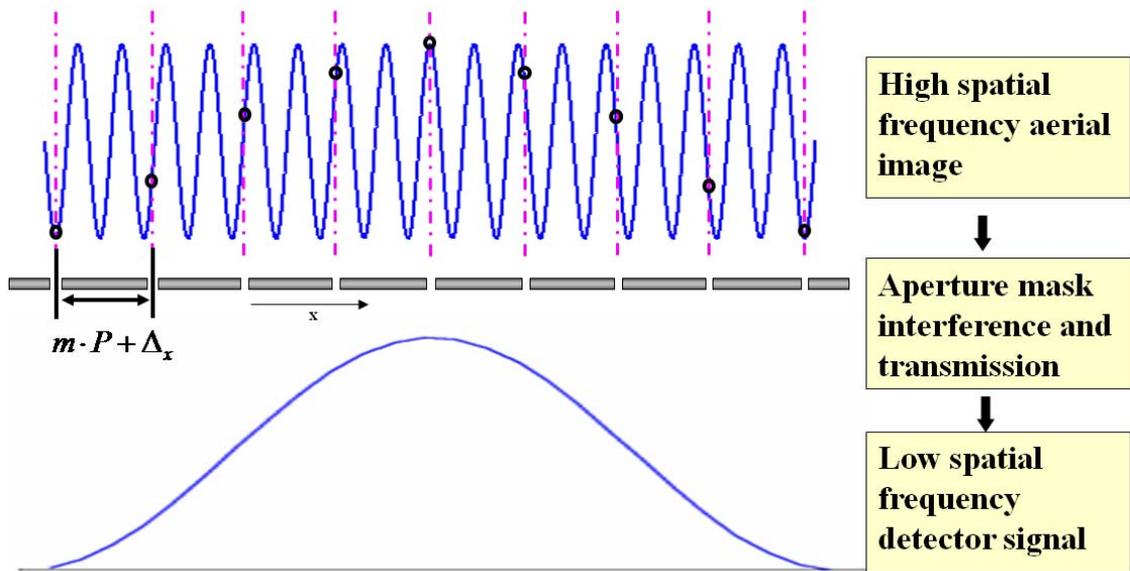
The IAIS interference pattern and the signal acquisition process are shown in Figure 4-7. Both the aperture mask and the aerial image are the periodic patterns with a spatial phase difference Δ_x . Thus each aperture in the mask will sample the signal at different positions in one period of the incident aerial image. The sampling point shift Δ_x sequentially. Through sufficient sampling, the aerial image signal could be reconstituted from the sampled version. The image at the top of Figure (a) is the incident aerial image. The dash-dot line indicates the center of apertures in the grating mask. The circle is the center of the image part to be interfered with the gratings. The bottom image is the sampled signal from the aperture grating. The near field simulation, around 20nm below the mask plane, was performed to show the interference image passing through the aperture mask. The highest intensity corresponds to the peak intensity region of the aerial image, and the smallest intensity is sampled at the tail of the aerial image.



(a) Aperture mask spatial phase shift grating interferes and samples the incident periodic aerial image to form a low spatial frequency interference pattern.



(b) The interference pattern is received by the photo-detector underneath the aperture mask. The signal at each aperture group is integrated over the effective size of each photo-detector pixel.



(c) The high spatial frequency aerial image is transferred into the low spatial frequency detector signal, which can be monitored and further processed.

Figure 4-7 Illustration of IAIS interference pattern and signal acquisition process.

The total power to the photo-detector is the integral of the power density over the detector pixel area. The optical intensity is directly proportional to the detector power density. The sampled image intensity is integrated over the effective area of the photo-detector pixels of each aperture group, as shown in Figure (b). Then the signals of all the sampling groups are combined to form the detector image. The nanometer scale aerial image is therefore translated into a micrometer scale detector image, which can be detected by available commercial grade photo-detectors.

The IAIS is designed as an on-wafer sensor, which is placed at the wafer chunk of the exposure tool. The process flow is shown in Figure 4-8. We define that the image output of the lithography exposure tool is Aerial Image (AI), the near field output of the aperture mask is Detector Image (DI), and the final output of photo-detector is photocurrent.

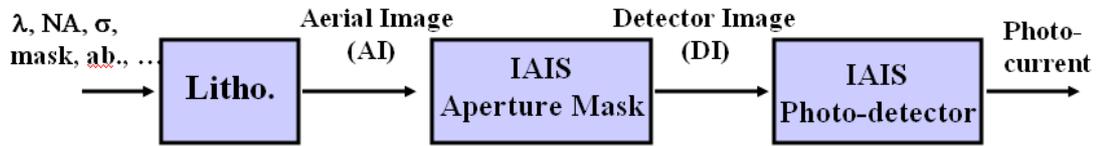


Figure 4-8 Illustration of IAIS system and process flow.

A practical IAIS implementation has two requirements: First, the output of the aperture mask, DI, must meet the needed performance specifications. For example, the DI contrast must be larger than the value limited by the measurement accuracy requirement. Second, the intensity of the DI must be significantly larger than the noise-limited detection threshold of the photo-detectors.

4.2.2 IAIS Aperture Mask

The image retrieved by the detectors depends on the optical and the geometrical properties of the aperture mask. The mask is mainly composed of spatial phase shift gratings. Figure 4-9 shows the aperture mask basic structure.

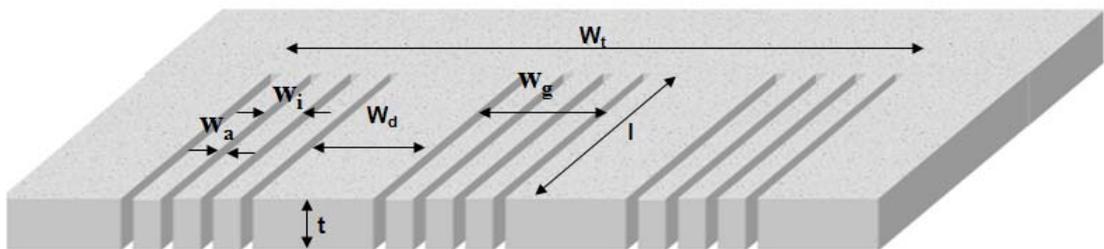


Figure 4-9 Illustration of the basic structure of the IAIS aperture mask

The gratings can be geometrically represented by:

$$T_{m_{-\xi}} = \{[\text{rect}(\frac{\xi}{W_a}) \otimes \text{comb}_n(\frac{\xi}{P})] \cdot \text{rect}(\frac{\xi}{W_g})\} \otimes \text{comb}_N(\frac{\xi}{W})$$

$$W_g = n \cdot P; \quad W_d = m \cdot P + \Delta_{\xi}, \quad W = W_g + W_d \quad \text{Eq. (4-1)}$$

Where, $\text{comb}(\xi) = \sum_{-\infty}^{\infty} \delta(\xi - n)$, $\text{rect}(\xi)$ is the rectangle function, ξ indicates the aperture mask

top plane, P is the period of the grating within one sampling group, W_g is the width of one group, W_d is the width between groups, and W_t is the total horizontal size. The aperture group spatial phase shift step is represented by Δ_{ξ} . The number of apertures within one aperture group is n , and the total number of the aperture groups is N . The number of image periods between aperture group m defines a “dead” space between aperture groups, which can be used to align the aperture groups to the photo detector pixels. Later in this chapter we will discuss the practical sizing of these parameters given the performance requirements of the overall IAIS system.

4.2.2.1 IAIS Aperture Mask Material

A material suitable for the aperture mask should have high absorption, low deposition temperature, must be easy to pattern in the nm-scale, and have long term physical and chemical stability. A high signal to noise ratio requires lower background transmissions and therefore higher absorption coefficient. If α is the absorption coefficients of the material, and T represents the transmission of the blank material film, the transmission is inversely proportional to the absorption coefficient [4-9]: $T = e^{-\rho ad}$, where ρ is the material density, and d is the film thickness. The extinction coefficient, or the imaginary

part of the refractive index, indicates the material absorption property, while the real part indicates the transmission property.

Standard IC fabrication materials with high extinction coefficients at 193nm include chrome, crystalline silicon, amorphous silicon, aluminum, etc. The refractive indices of four materials at 195nm are listed in table 4-1 [4-10].

Table 4-1 Refractive indices of five materials at 195nm [4-10]

Material	Chrome	c - Si	α - Si	Al
n	0.861	0.9096	0.9955	0.1149
k	1.668	2.7970	2.1278	2.234

The deposition temperature for crystalline silicon is up to more than 1000°C, and the temperature for poly-silicon is usually larger than 580°C, while the temperature for α -Si is lower than 580 °C, so amorphous silicon is a better choice than crystalline silicon in the sensor design. Chrome has smaller values of both the real and imaginary indices than others. Aluminum has a much lower real index than the imaginary index, which could lead to surface plasmons [4-11] and is sensitive to a specific polarization. From the aspects of absorption, deposition temperature, patterning ability and hardness, amorphous silicon and aluminum seem to be good candidate materials for IAIS applications.

As discussed in Chap 3, high NA polarization causes image contrast degradation. IAIS has the potential to act as a polarization monitor. We perform the simulation in Panoramic to investigate the polarization ratio as a function of material refractive index. Two apertures with aperture width of 35nm, thickness of 80nm, and pitch of 130nm are

placed in the simulation domain. We use two-beam, +1 order and -1 order, 193nm illumination to generate 130nm periodic aerial image pattern. The near field values are taken at 10nm below the aperture mask. Silicon dioxide is placed below the aperture mask, as it would be in the practical IAIS stack. The topography at z-x cut-plane in the simulation is shown in Figure 4-10.

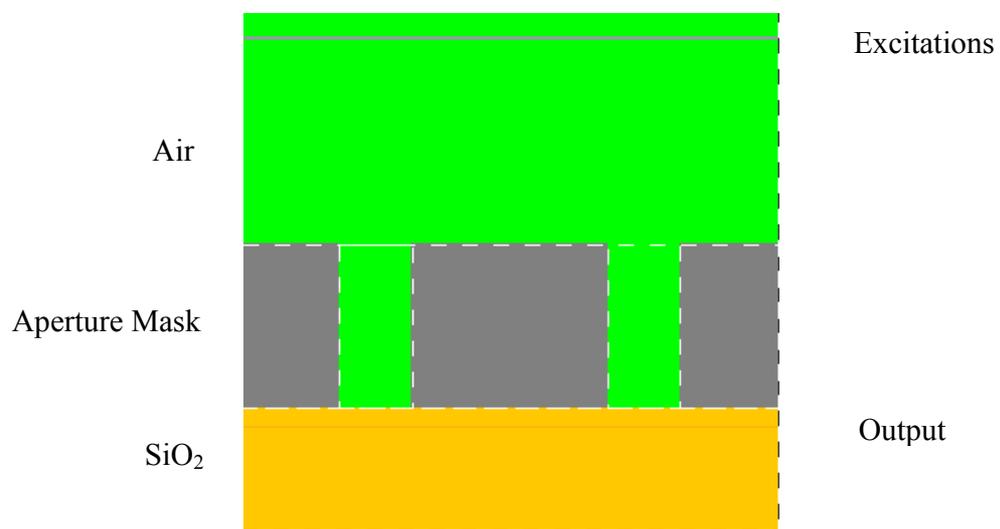


Figure 4-10 Illustration of simulation topography in z-x cut-plane.

The ratio between the maximum value of E^2 for the electric field orthogonal (TM) to the gratings and that of the electric field parallel (TE) to the gratings is plotted as a function of real refractive index in Figure 4-11. The real index is varied from 0.1 to 3.0, and the extinction coefficient is fixed at 2.1. The orthogonal polarization transmission is larger than that of the parallel polarization. The TE polarization will excite currents along the gratings that reradiate the energy, while TM polarization is unable to generate these currents and will be less affected by the gratings [4-12]. As the real index increases,

parallel polarization is attenuated through the apertures, causing the polarization ratio to increase significantly.

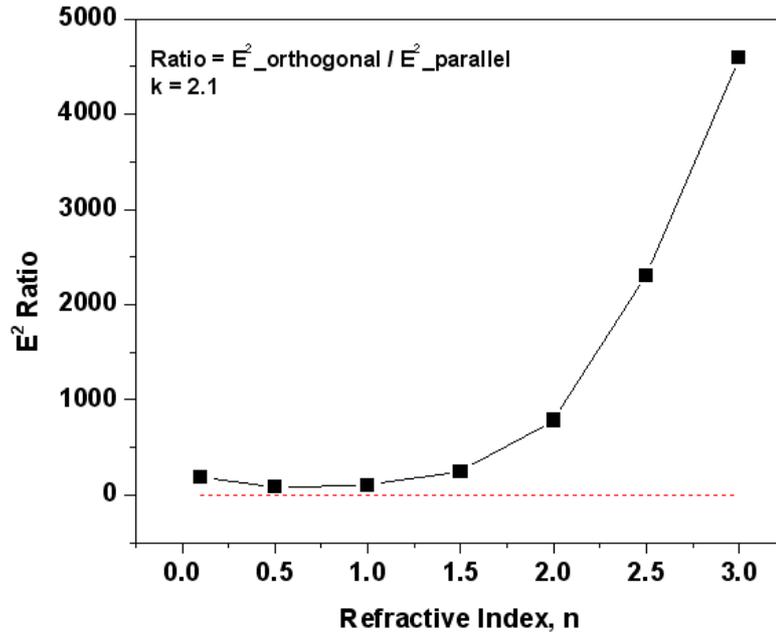


Figure 4-11 Polarization ratio of the maximum E^2 of the electric field orthogonal to the gratings to the maximum E^2 of the electric field parallel to the gratings is plotted as a function of real refractive index. The imaginary index is fixed as 2.1.

The polarization ratio as a function of the imaginary indices with a fixed real part is plotted in Figure 4-12. The TM transmission is improved, while the TE polarization is attenuated as k increases, therefore the polarization ratio becomes larger as the extinction coefficient grows. In summary, larger refractive index in both real and imaginary parts will improve the polarization ratio. Detecting this ratio is important, since as we discussed in Chapter 3, the high NA lithography causes the TM polarization of the aerial image out of phase to degrade the image contrast, while TE polarization is not affected.

To solve this problem, polarized illumination is used recently to increase TE polarization and to limit TM light, and therefore improve image contrast and exposure latitude. In order to detect TE polarization, the grating orientation can be tuned 90° to make the electric field orthogonal to gratings consistent with the aerial image TE polarization.

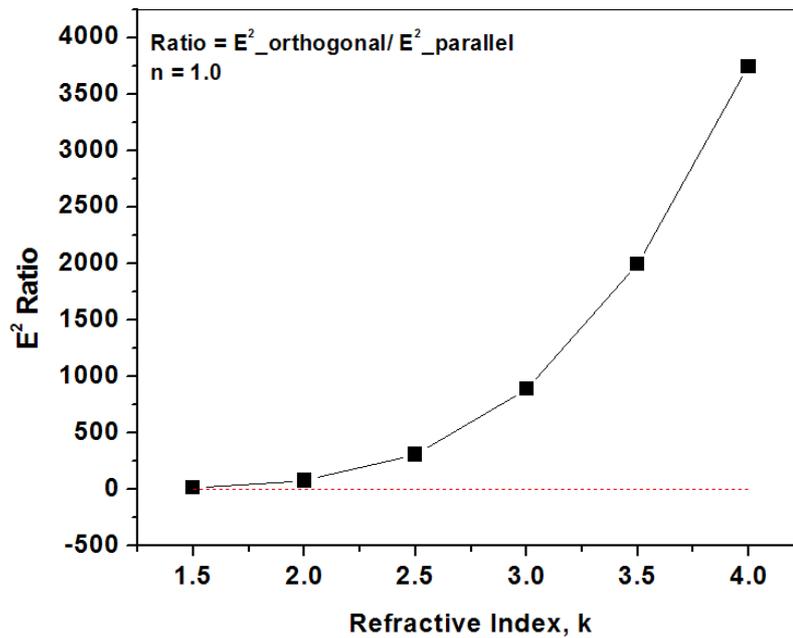


Figure 4-12 Polarization ratio of the maximum E^2 of the electric field orthogonal to the gratings to the maximum E^2 of the electric field parallel to the gratings is plotted with varied imaginary refractive index. The real index is set to 1.0.

Therefore, the high extinction coefficient of the refractive index will benefit the signal to noise ratio, the polarization ratio, and the transmission throughput. A large real index will improve the polarization ratio, while the effect on the transmission throughput is not significant. The optimal choice of mask material will depend on the intended design and the specific applications. By taking into account the detector gain and a desire to detect

polarization, a material with high imaginary index is a good candidate. A high real index is suitable to the polarization detector, while the background transmission noise has to be considered simultaneously. In addition to amorphous silicon and aluminum, tungsten with a refractive index of around $1.337 + i3.081$ at 195nm is also a good candidate. For the practical device, the aperture mask is fabricated on silicon dioxide with temperature below 200°C. From the overall aspects of the device performance, fabrication ability and equipment availability in the Berkeley Microlab, amorphous silicon is selected as the aperture mask material in the IAIS fabrication.

4.2.2.2 IAIS Aperture Mask Geometry

The aerial image sensor converts the aerial image into digital signal through discrete sampling by the aperture gratings. The grating period P , width W_a , and thickness t are three important parameters. The grating period is designed to be equal to the aerial image period. The aperture width determines the sampling pulse. As the pulse becomes narrower, it approaches a delta function. If we ignore the sub-wavelength effect, the nonzero values of the image will be the value of the sampled signal. The practical sampling width cannot be infinitesimally small, which is limited by the transmission throughput and our patterning capability. Therefore the aperture width is determined by the trade-off of more precise sampling and the higher transmissions. Considering the polarization ratio, reference [4-12] indicates that a gap width greater than 1/8 wavelength and smaller than 1/3 wavelength ensures enough throughput and polarization percentage. The light transmission through the apertures and the transmission ratio between the background and the gaps determine the aperture mask thickness.

We define the detector image as the near field transmission through the apertures before being received by the photo-detector. The detector image contrast can be expressed by:

$$C_{DIC} = \frac{I_{o,max} - I_{o,min}}{I_{o,max} + I_{o,min}} \quad \text{Eq. (4-2)}$$

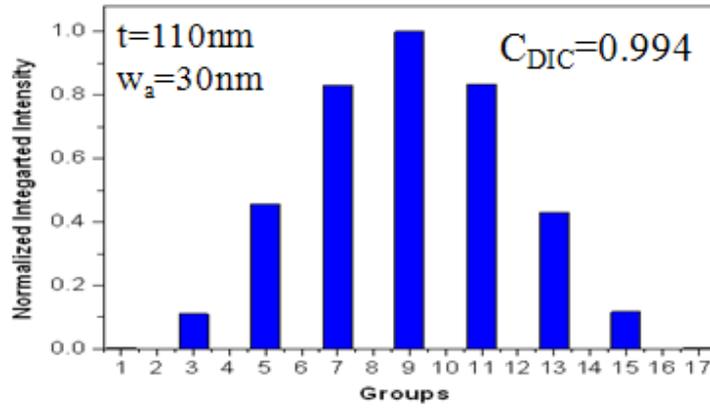
where I_o is the integrated detector image intensity over one aperture group. $I_{o,min}$ is the minimum detector image intensity, sampled at the point of the minimum aerial image intensity, and $I_{o,max}$ is the maximum detector image intensity, sampled at the position of maximum aerial image intensity. Detector image contrast provides one of the criteria to evaluate the detector image quality.

Figure 4-13 Detector image intensity of nine sampling apertures. The intensity is normalized to the maximum intensity when the aperture thickness is 10nm. The horizontal scale is μm . The aperture width is kept constant at 30nm.

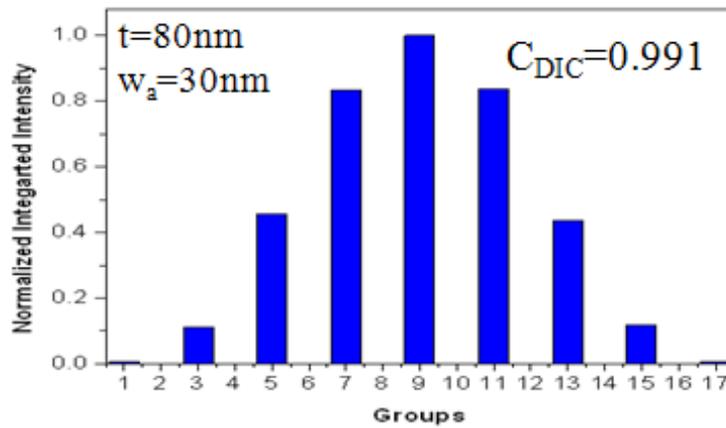
We perform the detector image simulation in Panoramic. The illumination settings are the same as those in Figure 4-10. The aerial image is a periodic pattern with pitch 130nm. The output is 20nm below the interface of aperture mask and SiO₂. Nine apertures are placed at the mask plane with spatial phase shift of 16.25nm. The aperture width is fixed at 30nm, and the thickness varies from 10nm to 110nm. The detector images of the nine groups are shown in Figure 4-13.

The intensity includes both TE and TM transmissions, which is normalized to the maximum intensity. As aperture thickness decreases from 110nm to 10nm, the minimum transmission $I_{o,min}$ and the background transmission increase significantly, while the detector image contrast drops accordingly. The histograms of the integrated intensity of each aperture are plotted in Figure 4-14. The detector image contrast decreases 72.4% as film thickness reduces to 10nm. To minimize cross-interference between aperture groups, we require an opaque zone between aperture groups that has the equivalent width of at least two photo-detector pixels. The background transmissions between apertures are tuned to match this background intensity.

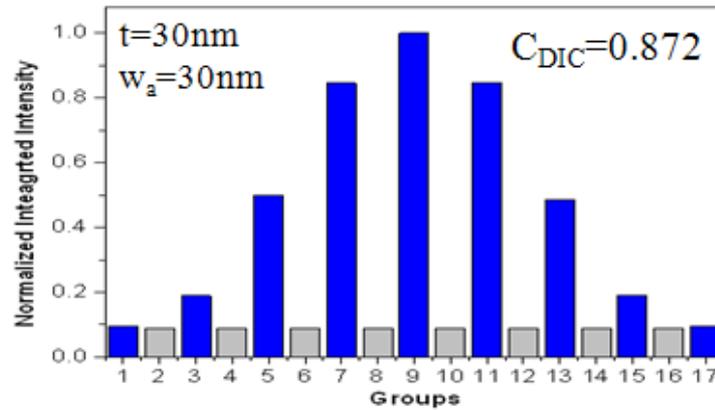
The aperture width effect is also investigated through simulations. The aperture thickness is fixed at 80nm. Figure 4-15 lists three examples of widths set at 30nm, 50nm, and 80nm respectively. The gray box at the right top of the graph illustrates the width and thickness of the aperture mask in the certain design. The detector image contrast decreases with the aperture width, while the transmission increases. An optimal design must balance transmission and image contrast. Figure 4-16 plots the maximum and minimum integrated intensity as functions of thickness and width. Both transmissions increase with aperture width and attenuate with aperture thickness.



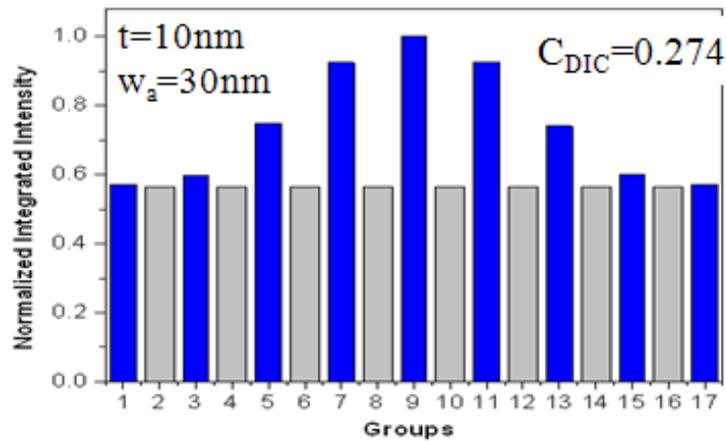
(a) Aperture thickness is 110nm, and aperture width is fixed at 30nm. The detector image contrast is 0.994.



(b) Aperture thickness is 80nm, and aperture width is fixed at 30nm. The detector image contrast is 0.991.

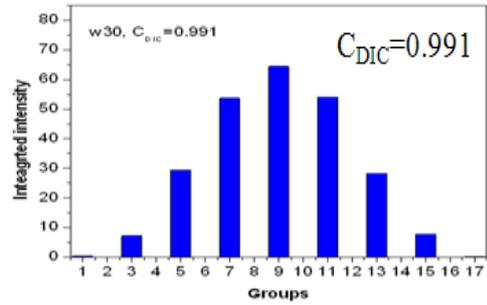
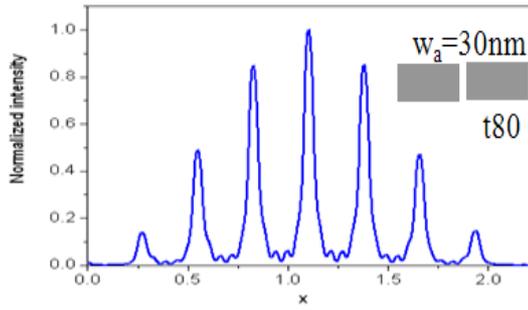


(c) Aperture thickness is 30nm, and aperture width is fixed at 30nm. The detector image contrast is 0.872.

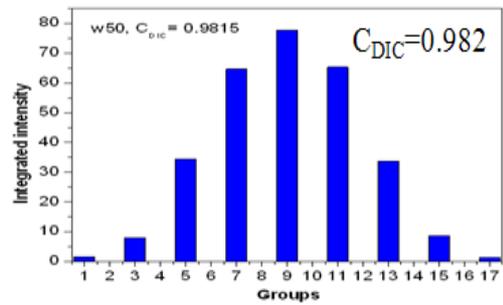
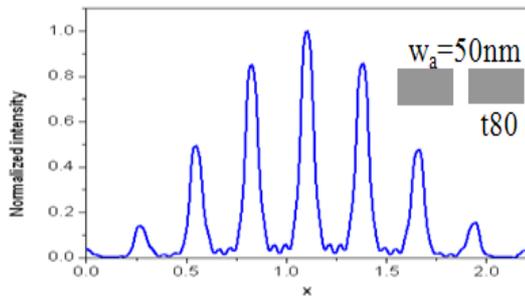


(d) Aperture thickness is 10nm, and aperture width is fixed at 30nm. The detector image contrast is 0.274.

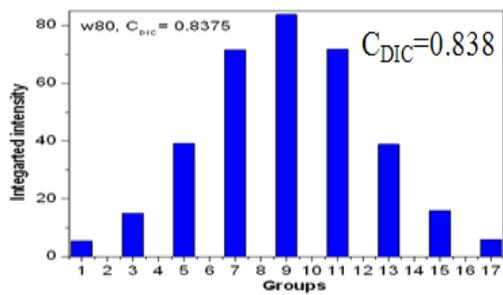
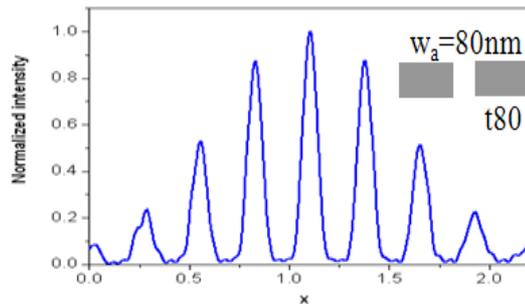
Figure 4-14 Illustration of the integrated detector image transmissions and contrasts as functions of aperture mask thickness. The aperture width is fixed at 30nm. Gray bars indicate background transmission through the aperture mask over the space separating the aperture groups. The width of that space is equivalent to two photo-detector pixels.



(a) Aperture width 30nm. Aperture thickness is fixed at 80nm



(b) Aperture width 50nm. Aperture thickness is fixed at 80nm



(c) Aperture width 80nm. Aperture thickness is fixed at 80nm

Figure 4-15 Illustration of the integrated detector image transmissions and contrasts as functions of aperture mask thickness. The aperture width is fixed at 30nm. The left side is the FDTD simulation results of the detector image, and the right side is the integrated detector image at each aperture.

Figure 4-16 The maximum and the minimum integrated intensities as functions of thickness and width.

Figure 4-17 takes the detector image contrast C_{DIC} into account. Since both the maximum and the minimum intensities increase simultaneously as thickness becomes smaller, the detector image contrast C_{DIC} attenuates at the smaller thickness. If aperture width increases to the size of the pitch, the maximum intensity is equal to the minimum intensity, and then the C_{DIC} becomes zero. On the other hand, if aperture width is extremely small, no significant light is transmitted, and the contrast therefore drops towards zero. Between the two extremes, the contrast is determined by the competition between the maximum and the minimum intensities, calculated by Eq. (4-2). The red line in Figure 4-17 indicates the 98% contrast value, which is the contrast of the good line-space dense pattern, and is used as one of our standards to judge a good detector image.

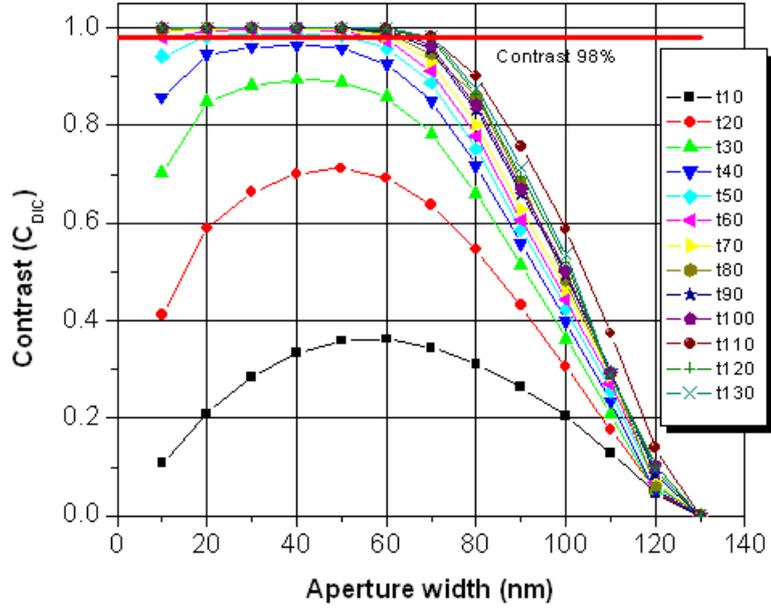


Figure 4-17 The detector image contrast C_{DIC} as functions of thickness and width.

In summary, transmission throughput and image contrast are two important properties to determine the detector gain and the image quality. By choosing the transmission level, the design window can be determined to achieve the required contrast value. Figure 4-18 plots the contrast C_{DIC} as a function of transmission.

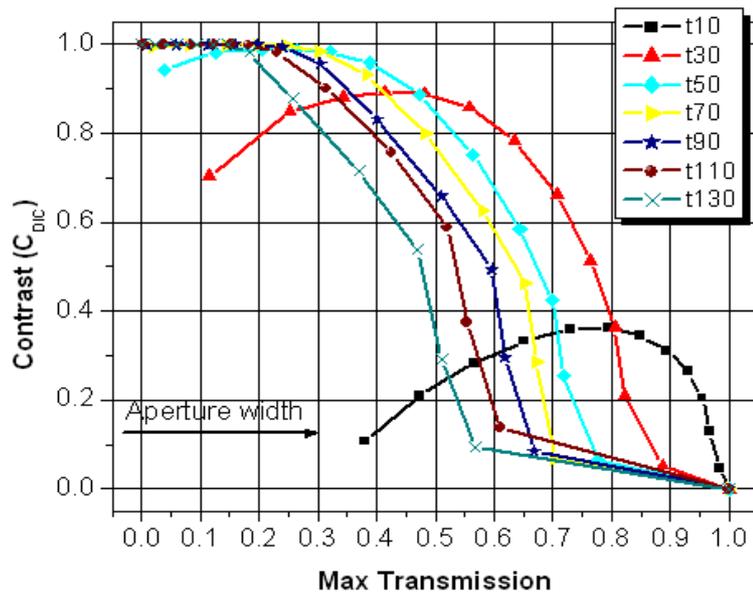
We define the signal to noise ratio (SNR) as a function of C_{DIC} :

$$SNR(C_{DIC}) = \frac{C_{DIC}}{C_{DIC,noise}} \quad \text{Eq. (4-3)}$$

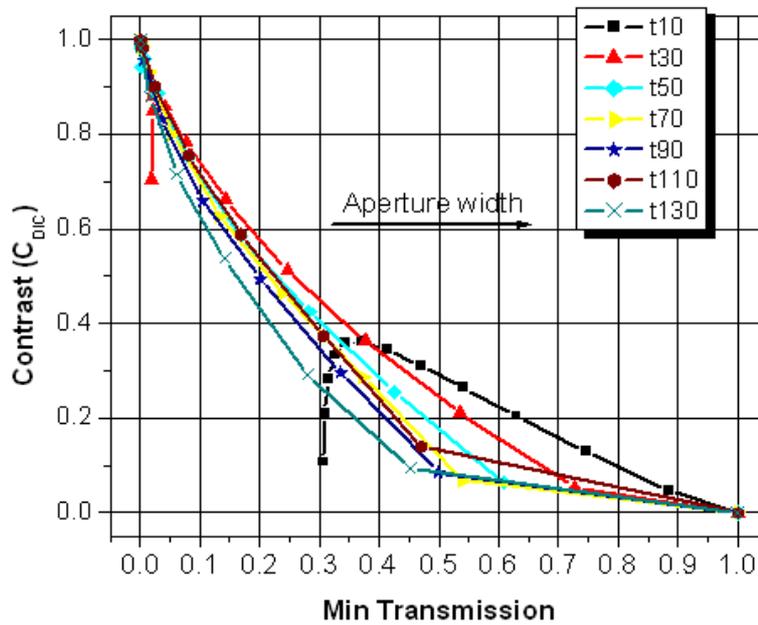
Where,

$$C_{DIC,noise} = \sqrt{\text{var}(C_{DIC})} = \sqrt{\sigma_{I_{o,max}}^2 \left(\frac{\partial C_{DIC}}{\partial I_{o,max}}\right)^2 + \sigma_{I_{o,min}}^2 \left(\frac{\partial C_{DIC}}{\partial I_{o,min}}\right)^2} \quad \text{Eq. (4-4)}$$

Assuming that the photo density follows a Poisson distribution, the photon noise has a square root relationship with the transmitted photon flux corresponding to $I_{o,max}$ and $I_{o,min}$.

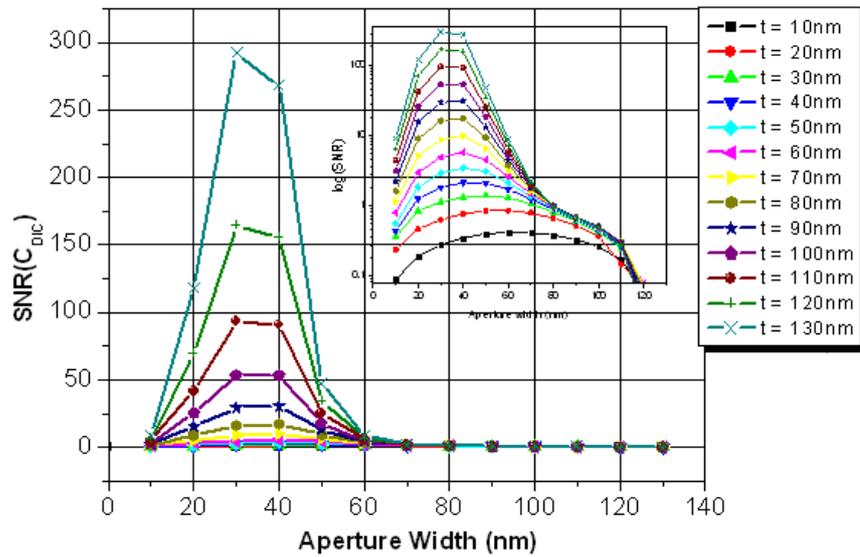


(a) Detector image contrast (C_{DIC}) vs. maximum transmissions

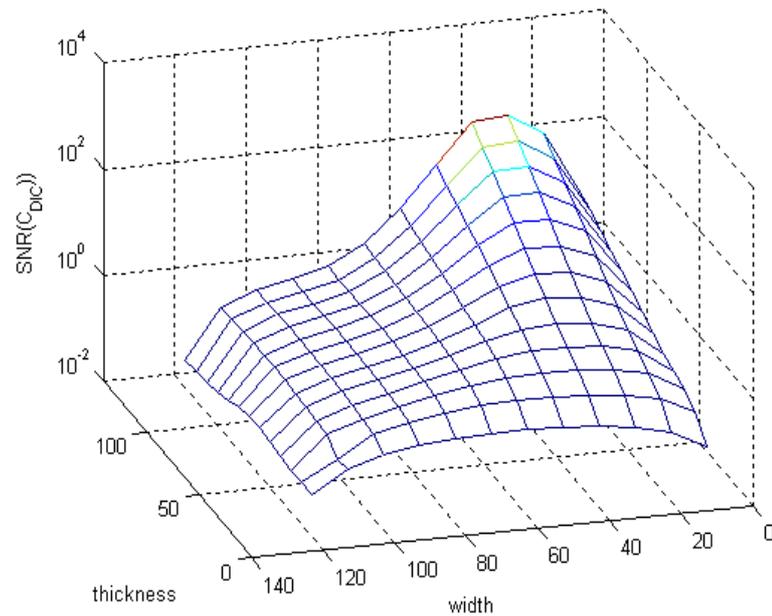


(b) Detector image contrast (C_{DIC}) vs. minimum transmissions

Figure 4-18 Illustration of the design window of detector image contrast (C_{DIC}) vs. transmissions.



(a) Signal to noise ratio of detector image contrast ($SNR(C_{DIC})$) vs. aperture width and thickness.



(b) Three dimensional plot to indicate the design window based on $SNR(C_{DIC})$

Figure 4-19 Illustration of the aperture size design window by signal to noise ratio of the detector image contrast.

The SNR (C_{DIC}) is plotted in Figure 4-19 (a). When aperture thickness is larger than 40nm, the peak SNR occurs when aperture width is around 20 to 50nm. The design window is clearly indicated in the three dimensional plot in Figure 4-19 (b).

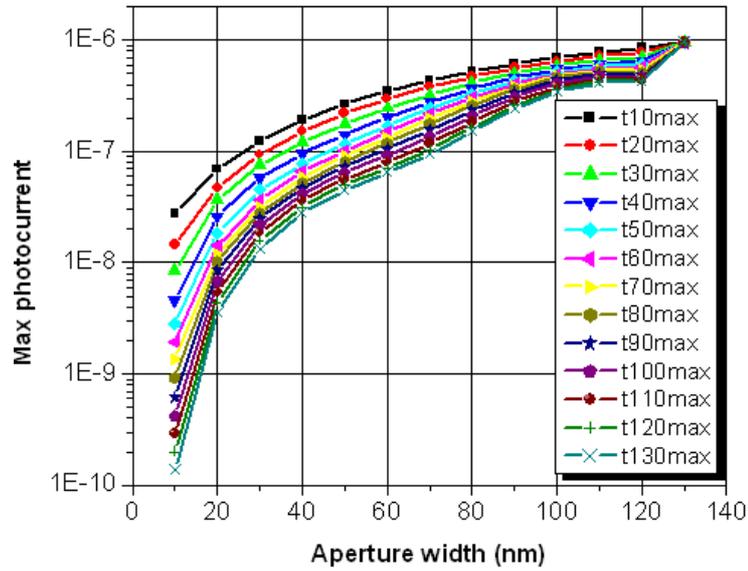
The third stage of IAIS flow in Figure 4-8 is photocurrent of photo-detectors, which is the final output of IAIS. The photocurrent [4-13] [4-14] is calculated by:

$$I_{photocurrent} (A / m^2) = \zeta \cdot \frac{q}{\hbar\omega} (A / W) \cdot P_{opt} (W / m^2) \quad \text{Eq. (4-5)}$$

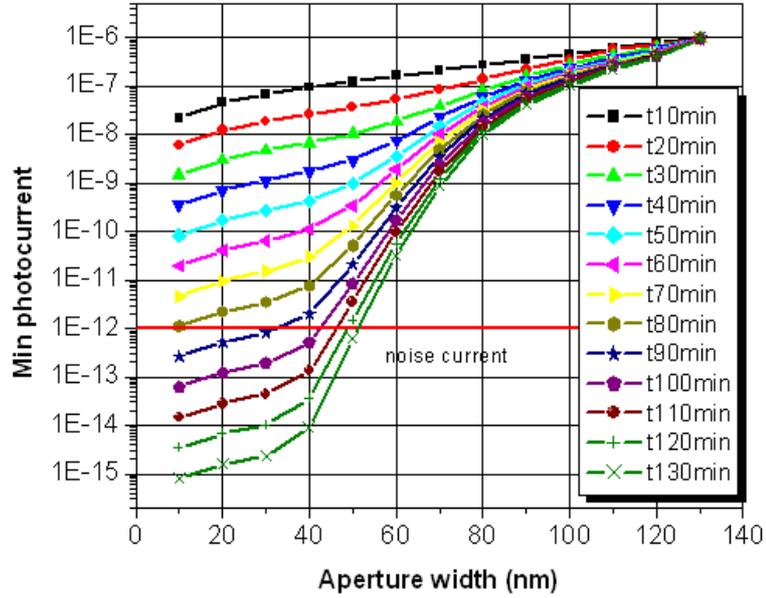
Where ζ is the quantum efficiency (QE) defined by: $\zeta = \frac{\#electrons}{\#photos}$, and P_{opt} is the

power density: $P_{opt} = \frac{E \cdot E^*}{2\eta}$. η is the characteristic impedance of the medium. The

photocurrent is calculated and plotted in Figure 4-20, assuming $QE=40\%$. The trend between photocurrent vs. aperture width and thickness is similar to that of the detector image. The photo-current contrast is considered in Figure 4-21, which indicates the measurement signal contrast. In this work we adopt a requirement that the photocurrent contrast is at least 98%. The optimized output signal contrast is at aperture thickness larger than 50nm, and aperture width between 20 to 60nm. The three dimensional plot is used to indicate the aperture size design window based on the photocurrent contrast. From Figure 3-19 and Figure 4-21, we conclude that the aperture width of $35nm \pm 15nm$ and the thickness of $80nm \pm 20nm$ define the optimal design window of the aperture mask.

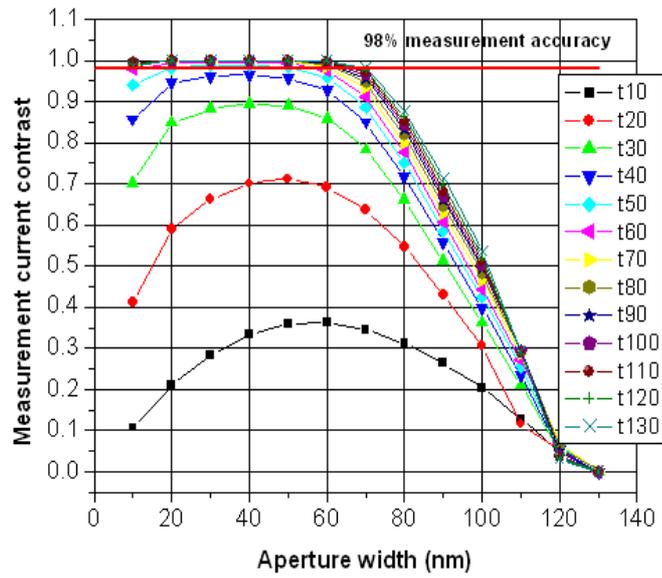


(a) The maximum photocurrent, shown in A/cm², as a function of aperture width and aperture thickness

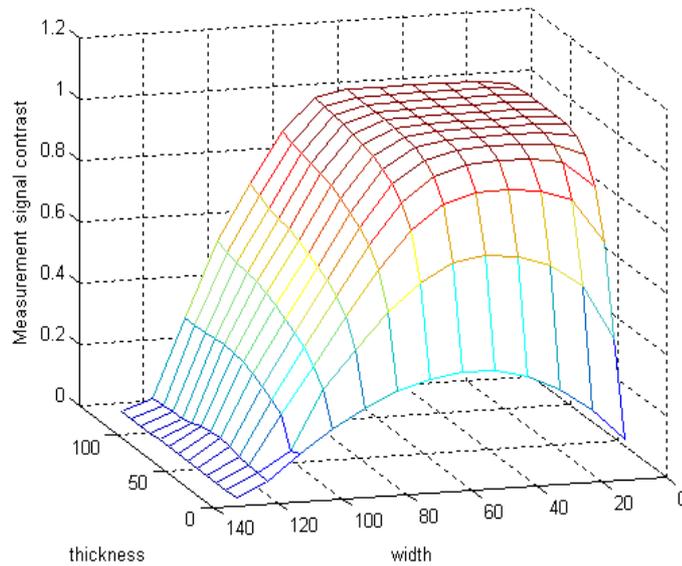


(b) The minimum photocurrent as a function of aperture width and aperture thickness. The red line indicates the photo-detector noise level of 10pA/cm².

Figure 4-20 Photo-current vs. aperture width and thickness



(a) The measurement current contrast as functions of aperture width and aperture thickness. The red line indicates the 98% contrast standard.



(b) Three dimensional plot of the design window based on photo-current contrast.

Figure 4-21 Illustration of the aperture size design window based on the estimated measurement signal contrast.

4.2.2.3 Other Features of IAIS Aperture Mask

The Nyquist-Shannon sampling theorem [4-15] states that in order for a band-limited signal to be fully reconstructed, the sampling frequency must be at least twice of the highest frequency of the input. The critical frequency is called Nyquist frequency. No information is lost if the signal is sampled at or above the Nyquist frequency, and no additional information is gained when sampling is faster than this rate. The Nyquist criterion dictates the aperture mask sampling frequency. For the 65nm nodes with aerial image of 130nm pitch, the maximum sampling interval is 65nm. In order to ensure high measurement fidelity, a smaller sampling step is necessary. A step of 10nm is feasible in the IAIS design.

For the sampling pitch of 130nm, sampling step of 10nm, 13 aperture groups are needed. In order to improve the transmission throughput, multiple apertures are placed in one sampling group. For the intent design of four detector pixels ($\sim 24\mu\text{m}$) per aperture group, 184 apertures are used in one group, which extends $23.92\mu\text{m}$ in the horizontal direction.

The dark field size of around two pixels is designed, with a size approximately equal to ninety-two pitches. The vertical grating size is around $100\mu\text{m}$, which takes the effective photo-detector area and the transmission into account.

4.2.2.4 Summary of IAIS Aperture Mask Design

The design of the IAIS aperture mask is summarized in Figure 4-22, where W_a is the aperture width, t is the aperture thickness, l is the aperture length, n is the number of apertures in one group, m is the equivalent number of apertures in the dark field, N is the

total number of sampling groups, Δ_ξ is the sampling step, W_i is the sampling pitch, W_g is the dimension of one aperture group, W_d is the dark field size, and W_t is the aperture mask effective dimension in the horizontal orientation. This design is intent to 65nm periodic patterns, but the same methodology can be extended to other technology-node applications.

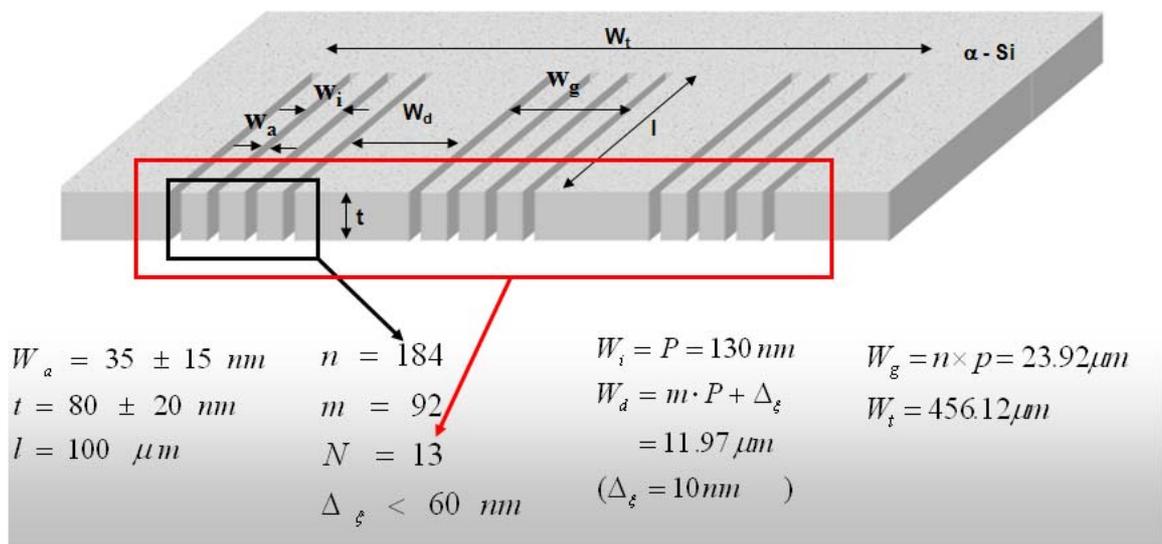


Figure 4-22 Summary of IAIS aperture mask design.

4.2.3 IAIS Photo-Detector Considerations

The criteria for IAIS photo-detectors are: high quantum efficiency at DUV wavelength, low noise at room temperature, the number of pixel larger than 250, the chip size smaller than 10mm, and the CCD material is compatible with aperture mask process.

The Kodak KAF0402ME CCD [4-16] is a suitable candidate for the IAIS. The sensor is built with a two-phase CCD technology employing a transparent gate. This technology

simplifies the support circuits that drive the sensor, reduces the dark current without compromising charge capacity, and significantly increase optical response compared to traditional front illuminated full frame sensors. The ME configuration adds a micro lens to the surface of the CCD sensor. These lenses focus the majority of the light through the transparent gate, further increasing the optical response.

The CCD is shown in Figure 4-23. The total number of pixels is 768(H) x 512(V). The pixel size is $9\mu\text{m} \times 9\mu\text{m}$. The dark current is $<10\text{pA}/\text{cm}^2$ @ 25°C , and the quantum efficiency is around 40% at DUV. The imager size is 6.91mm (H) x 4.6mm (V), and the die size is 8.4mm (H) x 5.5mm (V). There are 24 package pins. The packaged CCD is de-bonded by phenol to get the naked CCD chip for on-wafer integration. The CCD chip is wire-bonded and then tested and verified that the functionality is good after de-packaging.

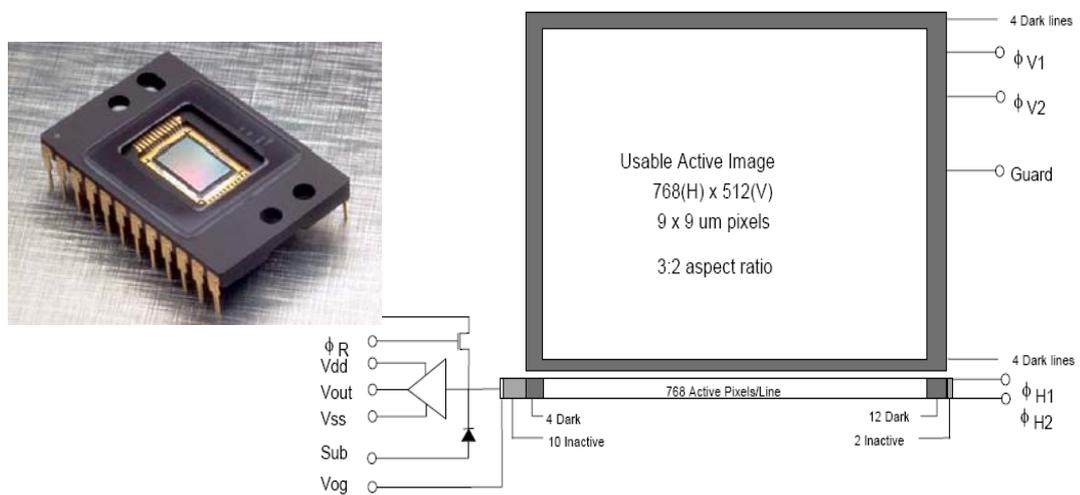


Figure 4-23 Kodak KAF-0402ME photo-detector functional block diagram. The upper left picture is the packaged CCD chip.

4.3 Applications of the Integrated Aerial Image Sensor (IAIS)

4.3.1 IAIS as an Aerial Image Metrics Detector

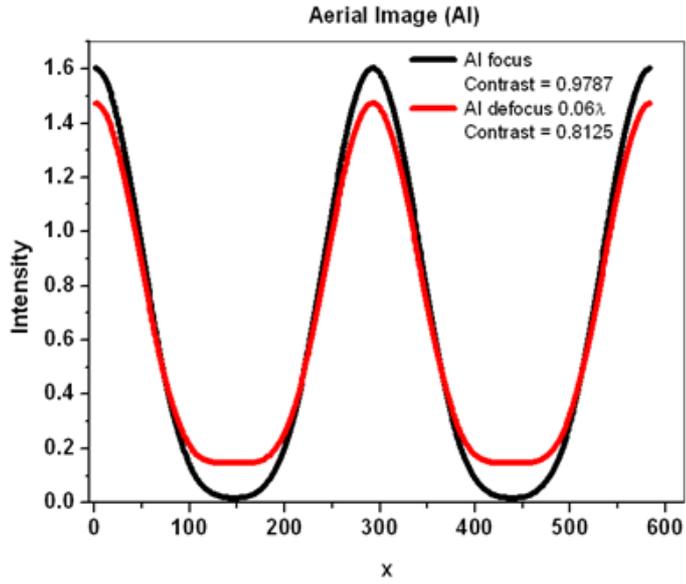
Aerial image intensity, contrast, slope, and feature size are the most important metrics to determine the aerial image qualities. IAIS has the capability to detect these aerial image features. The simulation is performed in Panoramic EM-suite with topography imaging. The two topographies are built: mask topography, representing the top mask at the reticle plane, and wafer topography, representing the on-wafer aperture mask at the wafer plane. Five steps are involved in the simulations:

1. Discretization of the illumination source
2. Simulation of the top mask and projection optics
3. Extraction of the scattering orders at the wafer plane
4. Simulation of the IAIS aperture mask topography
5. Combination of the partial images into integrated detector image.

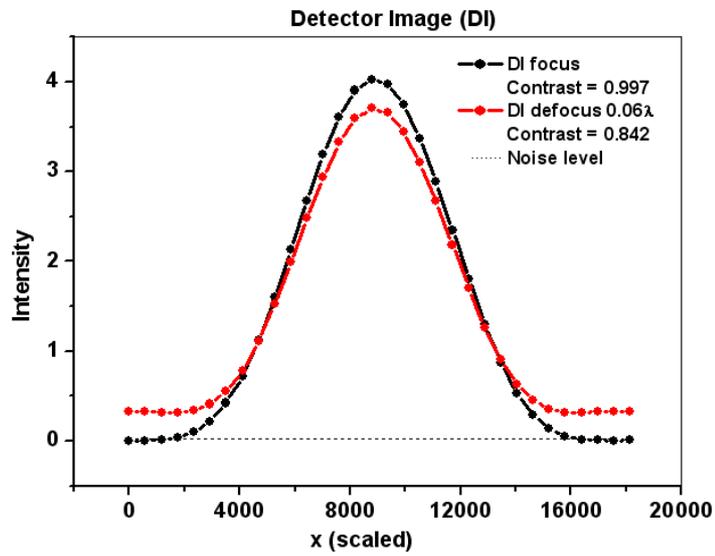
We generate the aerial image of 292nm 1:1 line/space periodic patterns, with wavelength of 193nm, NA of 0.75, and σ of 0.1. The top mask uses scalar mask model. The intensities at best focus and defocus of 0.06λ are shown in Figure 4-24 (a). The aerial image contrast at defocus of 0.06λ decreases to 83% of that at best focus condition. We place the IAIS aperture mask of width 50nm, thickness 90nm, and sampling step 10nm on top of SiO₂ at the wafer plane. Through wafer topography reimaging, we extract the near field detector image at 10nm below the aperture mask, which is the assumed effective position of the photo-detector. Figure 4-24 (b) shows the detector image sampling results. The contrast at defocus of 0.06λ is 84% of that at the best focus condition, which matches with the aerial image result. The noise level on the

graph is the equivalent intensity calculated from a dark current of $10\text{pA}/\text{cm}^2$. The minimum intensity of the best focus is compromised by the photo noise, while the defocus minimum intensity is larger than that level. The maximum DI intensity at defocus drops to 8% of the intensity at the best focus, which has the same result as that of the aerial image. Both the DI shapes at best focus and 0.06λ defocus are similar to those of the AI.

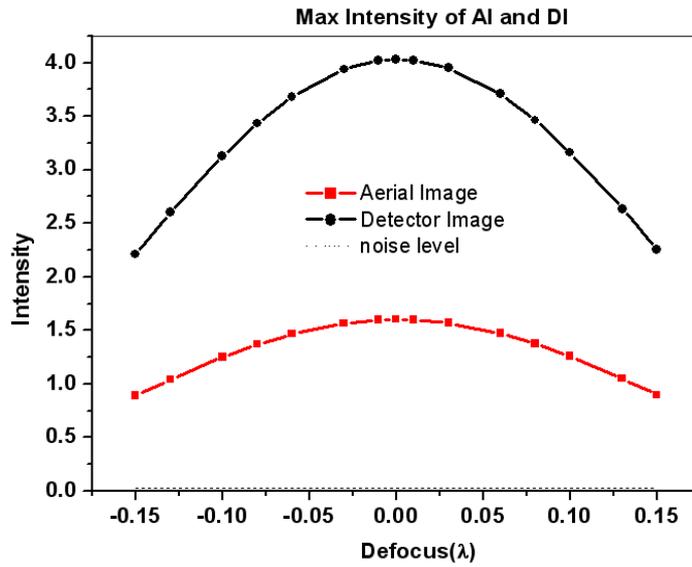
Defocus aberration causes aerial image intensity variations. Figures 4-24 (c) and (d) show the maximum and minimum intensities as a function of defocus. The detector image can capture the aerial image variation trends with defocus, while the max and min intensity curves of DI are sharper than those of AI. These results show that the detector image is more sensitive to defocus variations than the aerial image, and therefore IAIS provides an efficient way to measure defocus variations using the binary periodic patterns.



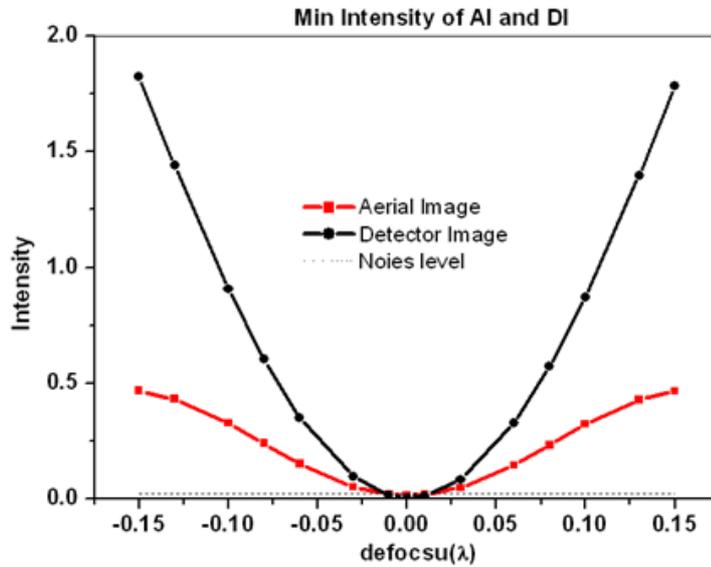
(a) Aerial image of 292nm 1:1 line/space periodic patterns, with wavelength of 193nm, NA of 0.75, and σ of 0.1. The two curves are intensities at best focus and defocus of 0.06λ respectively.



(b) IAIS detector image of monitored aerial image in (a).

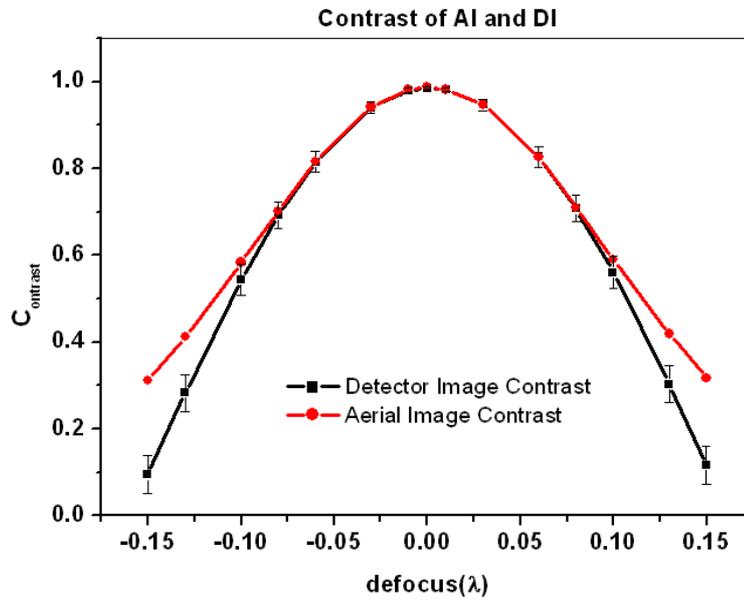


(c) Maximum intensities of AI and DI as a function of defocus

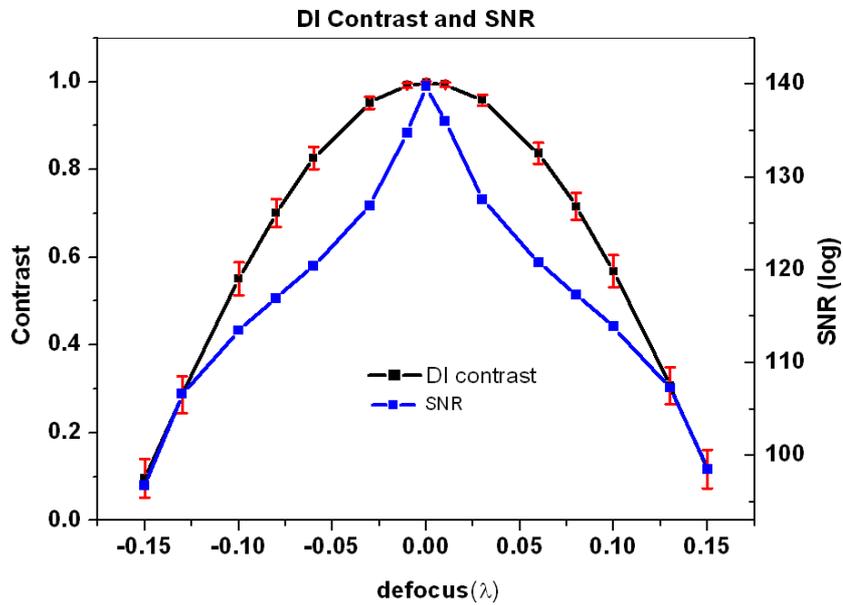


(d) Minimum intensities of AI and DI as a function of defocus

Figure 4-24 Illustration of IAIS capabilities to detect aerial image intensity and intensity variations.

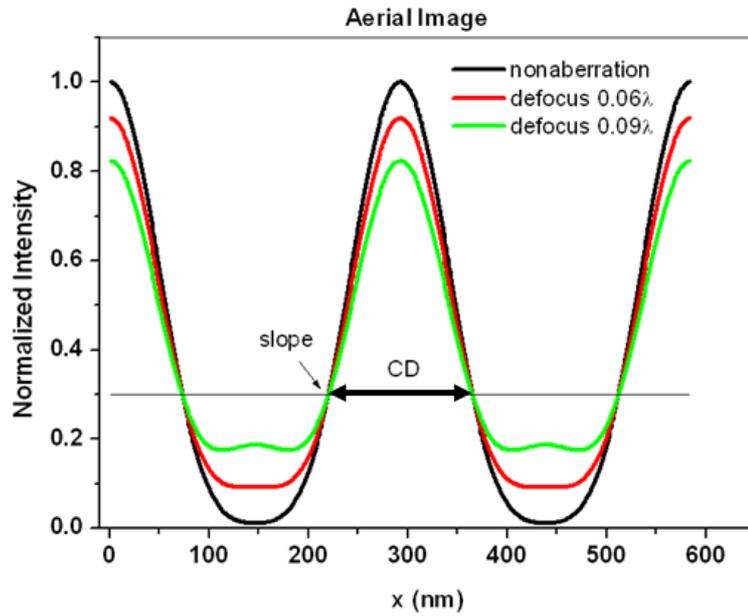


(a) Aerial image and detector image contrast vs. defocus. The Noise bar is magnified 1000x for clarity.

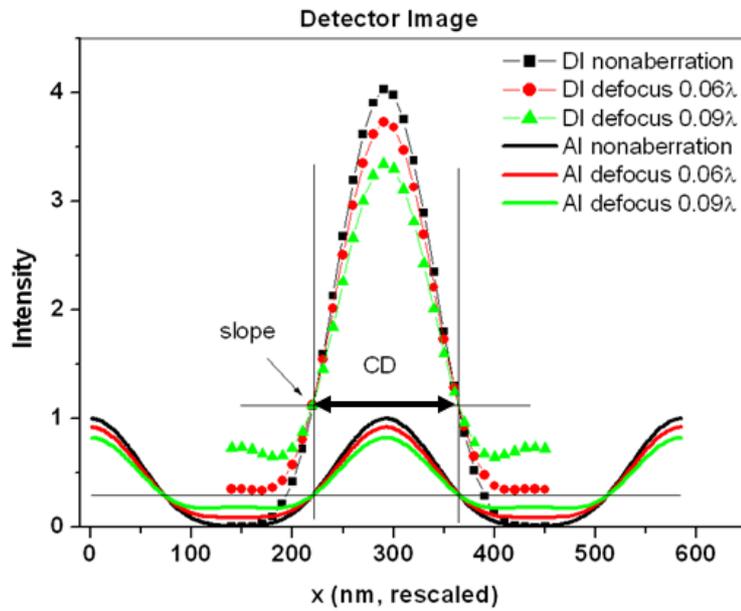


(b) Detector image contrast and signal to noise ratio (SNR) vs. defocus variations.

Figure 4-25 Illustration of IAIS capabilities to detect aerial image contrast and contrast variations.



(a) Aerial image slopes at best focus, and defocus of 0.06λ and 0.09λ respectively. A constant threshold is set at intensity of 0.3.



(b) Detector image slopes of monitored aerial image in (a)

Figure 4-26 Illustration of IAIS capabilities to detect aerial image slope and slope variations.

The aerial image and detector image contrasts from the above simulations are shown in Figure 4-25. The detector image contrast is more sensitive to defocus variations than the aerial image contrast. We calculate the contrast noise based on the standard deviations of the contrast in Eq. (4-4). The contrast noise increases with defocus. The noise bar is magnified 1000x for clarity in the figure. The contrast difference between best focus and 0.01λ defocus is 50 times larger than the contrast noise, which indicates the IAIS' ability to detect defocus aberration is in the order of 0.01λ . The signal to noise ratio (SNR) is calculated in Figure 4-25 (b) by Eq. (4-3). The maximum signal to noise ratio is achieved around the best focus plane. At small defocus range, e.g. around 0.03λ defocus, SNR is a more sensitive parameter than contrast to monitor defocus variation.

Aerial image slope is another important aerial image property. A larger slope indicates a better image quality with sharper sidewall angle, better Line Edge Roughness (LER) characteristics and more stable to process variations, such as defocus and dose, and resist contrast variations. The slope in one dimension is defined as:

$$S_{slope} = \frac{\partial I}{\partial x} \quad \text{Eq. (4-6)}$$

Since exposure can increase and therefore improve the slope, a slope normalized to the intensity is desired. The normalized image log slope (NILS) [4-17] is defined as:

$$NILS = L \frac{\partial(\ln I)}{\partial x} = \frac{L}{I} \left(\frac{\partial I}{\partial x} \right) \quad \text{Eq. (4-7)}$$

In order to scale to small features, the slope should also be scaled to feature size. We use the CD normalized slope as:

$$S_{slope} = \left(\frac{\partial I}{\partial x} \cdot \frac{\lambda}{NA} \right) / \left(\frac{\lambda}{NA} \right) = S_{slope,CD} / \left(\frac{\lambda}{NA} \right) \quad \text{Eq. (4-8)}$$

Table 4-2 The aerial image slope and CD at best focus and defocus of 0.06λ and 0.09λ

	best focus	defocus 0.06λ	defocus 0.09λ
$S_{slope,CD} (I/(\lambda/NA))$	2.76	2.32	1.79
NILS	5.19	4.35	3.35
CD (nm)	145	145	143

The aerial image slopes at best focus, and defocus of 0.06λ and 0.09λ are shown in Figure 4-26 (a), using a constant threshold of intensity of 0.3. The top mask and the on-wafer aperture mask are the same as those in Figure 4-24 and 4-25. The CD normalized slope of $S_{slope,CD}$ with unit $I/(\lambda/NA)$, and NILS are listed in table 4-2. We extracted the CD values at the intersection of the threshold and the image edges, which are shown in table 4-2. The CD extraction resolution is one nanometer. The $S_{slope,CD}$ decreases 34% at 0.09λ defocus, while CD changes 1.3%. The slope is very sensitive to defocus variations, even as CD does not vary significantly at the small defocus.

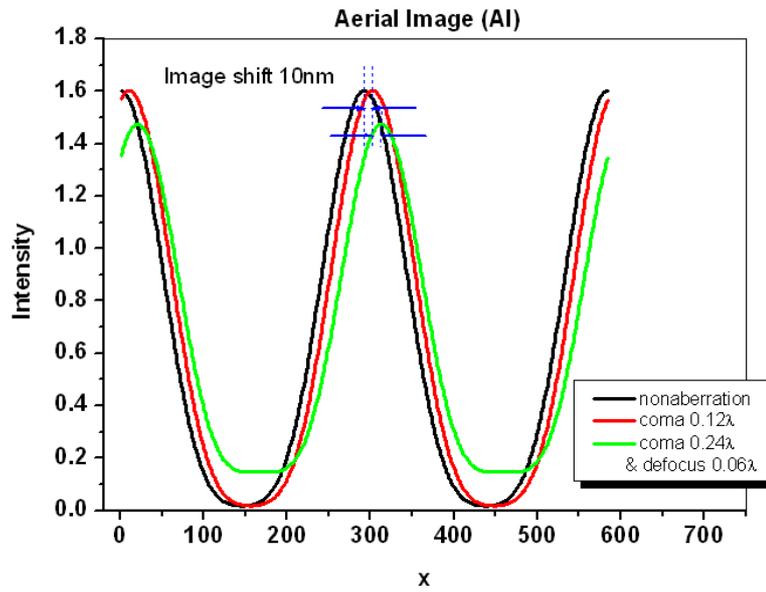
The corresponding detector image is shown in Figure 4-26 (b). The x coordinate is scaled to the same dimension as the aerial image, so that we can consistently extract the slope and CD. Since the detector image (DI) integrated intensity is much larger than that of the aerial image (AI), the slope is enhanced accordingly, while the CD values are the same. In the case that the DI and the AI have different intensity scales, *NILS* is a good criterion to compare slope results. The slope and CD values are listed in Table 4-3. At the best focus condition, the DI *NILS* increases slightly than that of AI. For defocus at 0.06λ , DI *NILS* decreases to 84% of the best focus value, and AI *NILS* has almost the same scale as DI, which reduces to 83.8% of the best focus value. Similarly, at 0.09λ defocus, both

DI and AI *NILS* have the same ratio, around 64% of the best focus value. Therefore, although IAIS detector image intensity increases compared to the aerial image, the intensity normalized slope is preserved.

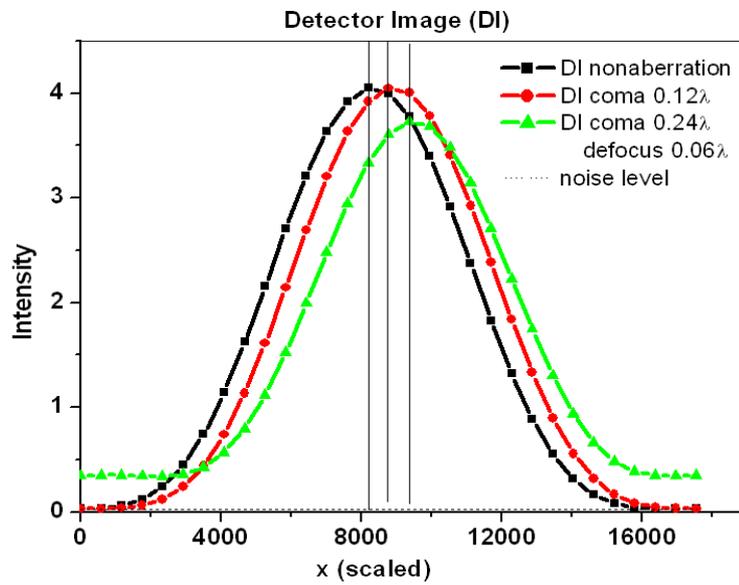
Table 4-3 The detector image slope and CD at best focus and defocus of 0.06λ and 0.09λ

	best focus	defocus 0.06λ	defocus 0.09λ
$S_{\text{slope,CD}} (I/(\lambda/NA))$	11.27	9.49	7.36
NILS	5.57	4.69	3.59
CD (nm)	145	145	143

Odd aberrations cause image shift. Aerial images at zero-aberration, coma of 0.12λ , and coma of 0.24λ and defocus of 0.06λ are shown in Figure 4-27 (a). The image shifts 10nm per coma of 0.12λ without contrast degradations. As defocus combines with coma, both image shift and contrast distortion occur simultaneously. Figure 4-27 (b) shows the detector image at three conditions. The sampling step is 10nm, and the distance between sampling groups is two pitches. The peak intensity shifts one sampling step, which indicates the detector image shift is 10nm. IAIS has the capability to determine the image shift, while the precision of determining the shift depends on the sampling frequency and the detector noise. As long as the intensity change between the sampling groups is larger than the detector noise, the smaller sampling step will benefit the precisions, and advanced fitting techniques, such as spline interpolation, can further improve the monitoring precision.



(a) Aerial image at non-aberration, coma of 0.12λ and zero defocus, and coma of 0.24λ and defocus of 0.06λ .

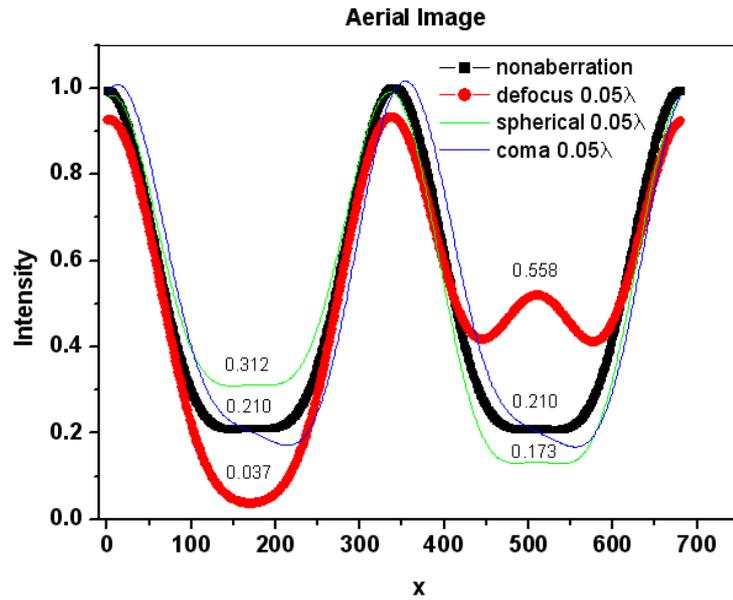


(b) Detector image shift of the monitored aerial image in (a)

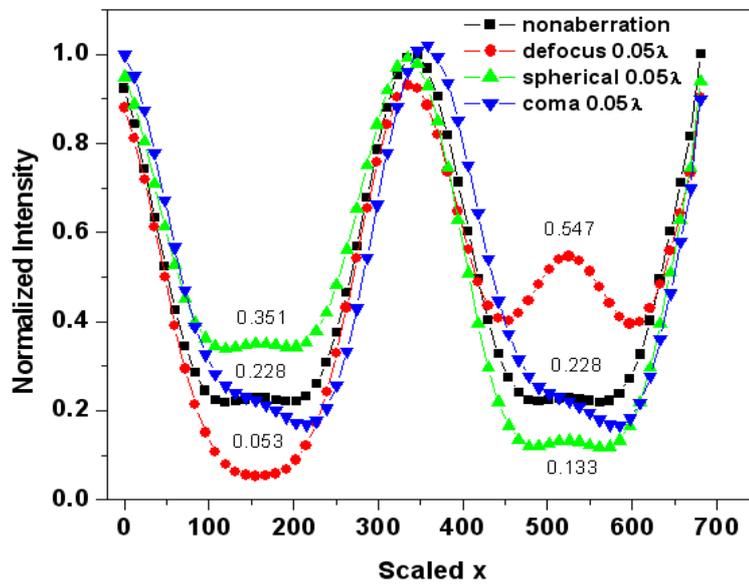
Figure 4-27 Illustration of IAIS capabilities to detect aerial image shift

Phase shift masks are used as advanced process monitors to detect process variations, as we discussed in Chapter 3. IAIS can be used as an on-wafer aerial image monitor of the top phase shift masks. Figure 4-28 (a) is the aerial image of the top mask pattern with 0° phase of $0.6\lambda/\text{NA}$ line width at the center, and $+90^\circ$ (left) and -90° (right) phase of $0.3\lambda/\text{NA}$ lines alternating to form a periodic patterns with pitch of $2.16\lambda/\text{NA}$. The 0.05λ defocus, spherical, and coma aberrations are considered, which cause the side-lobe intensity variations at the phase shift lines. At $+90^\circ$ side lobe position, defocus intensity decreases to 18% to the ideal condition, and spherical intensity increases to 148%. Coma causes lateral shift and asymmetric intensity around the center of the side-lobe, but the average side-lobe intensity has no significant change. At -90° side lobe position, defocus peak intensity increase to 266%, with a spherical decrease to 82% of the ideal condition.

Figure 4-28 (b) shows the detector image simulations with normalized intensities. The left and right side-lobe intensities are asymmetric at the edge at the ideal condition because the sampling window is not perfectly symmetric around the center peak intensity. The DI intensity variation values in the presence of modest aberration are close to but not exactly same as those of AI, which may be due to the averaging effect of the sampling pulse and the thick phase shift mask effect, however the detector images closely follows the aerial image trend at the side lobes and the main features at each aberration, which efficiently monitors the aberration effects.



(a) Aerial image of top phase shift mask patterns with non-aberration and 0.05λ of defocus, spherical and coma aberrations.



(b) Detector image of the monitored aerial image in (a).

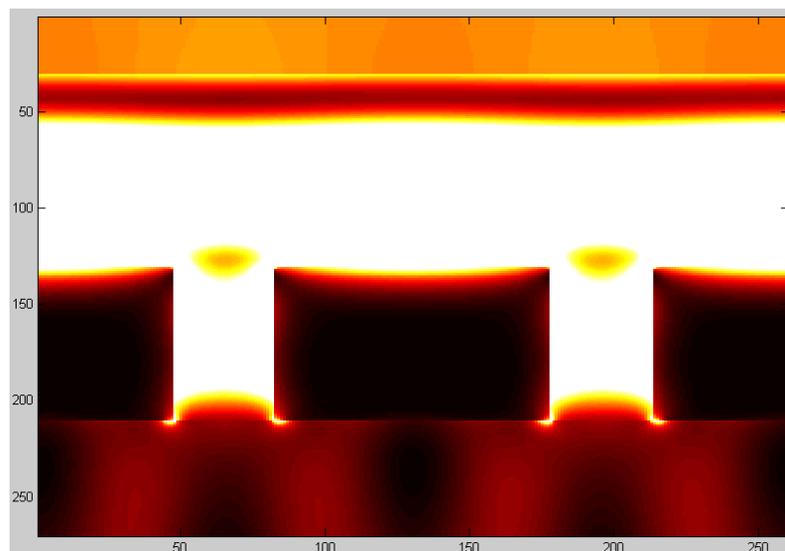
Figure 4-28 Illustration of IAIS capabilities to detect phase shift mask aerial image patterns.

In summary, IAIS can detect aerial image intensity, contrast, slope variations, and image shift. Even aberrations, such as defocus, cause image intensity and contrast degradations. The detector image can capture the aerial image intensity variation trends with defocus, and provides an efficient and sensitive way to measure defocus variations on binary periodic patterns. The maximum signal to noise ratio of contrast is achieved at the best focus plane, and the contrast difference between best focus and 0.01λ defocus is 50 times larger than the contrast noise. These indicate the IAIS' ability to detect defocus aberration, as limited by photo-detector noise is in the order of 0.01λ . The detector image slope is larger than that of the aerial image, while the intensity normalized slope (*NILS*) and slope variations follow the same trend and range of those of the aerial image. The DI CD feature size keeps the same value of that of the aerial image with the proper dimensional scale. Coma and other odd aberrations cause an image shift. IAIS can detect the image shift, and the precision is determined by the sampling frequency and the detector noise. Advanced fitting technologies, such as spline and cubic interpolations, can further improve the detector precision. Phase shift masks, such as PPG defocus monitor, are used as the advanced process monitors. IAIS can be used as the aerial image monitor of these process monitors, which removes the additional variations caused by the photo-resist effects. Simulation results of 90° phase shift masks indicate that the detector images closely follow the aerial image trends and could efficiently monitor aberration effects.

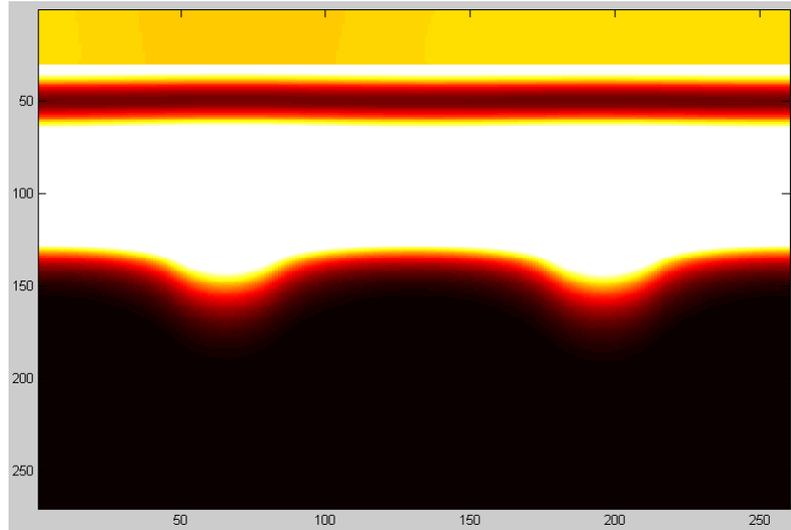
4.3.2 IAIS as a linear Polarization Monitor

IAIS has the potential to function as a linear polarization monitor. Figure 4-29 shows the near-field contour of the orthogonal polarization (TM) and parallel polarization (TE)

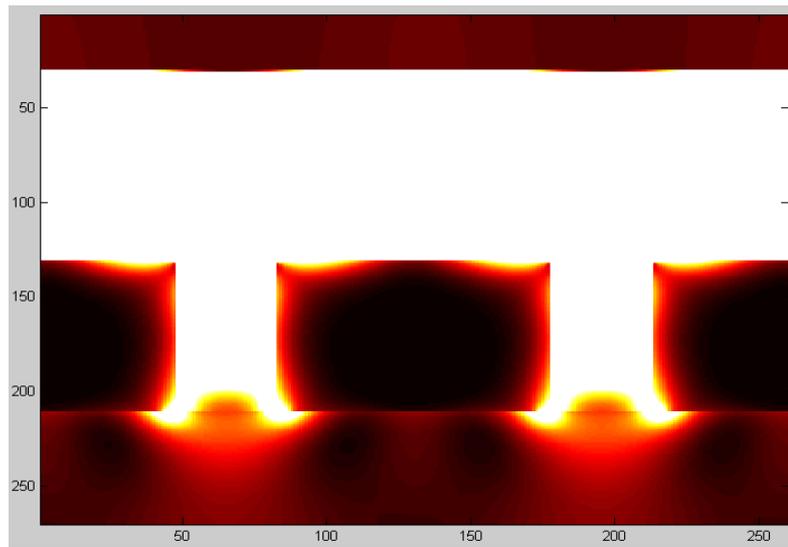
through the aperture mask. The simulation is performed in the Panoramic EM-suite with normal illumination in Figures (a) and (b). Two apertures are placed in the simulation domain. The aperture width is 35nm, the thickness is 80nm, and the pitch is 130nm. The TM polarization transmits through the apertures, but the TE polarization is mostly blocked, which is consistent with the polarization analysis in section 4.4.2.1. The two beams of ± 1 order illumination are combined to form the aerial image of pitch 130nm in Figures (c) and (d) at an off-axis angle of 47.93° . The on-wafer topography is the same as the above. For the TM polarization in Figure (c), most of the light transmits through the apertures, however the aerial image contrast is very low, and hence no image pattern is generated above the aperture mask. The TE polarization is shown in Figure (d). The aerial image is formed with a good contrast, while no detectable light is transmitted through the open portion of the aperture mask. The aperture mask and the film stack are shown in the figure for clarity.



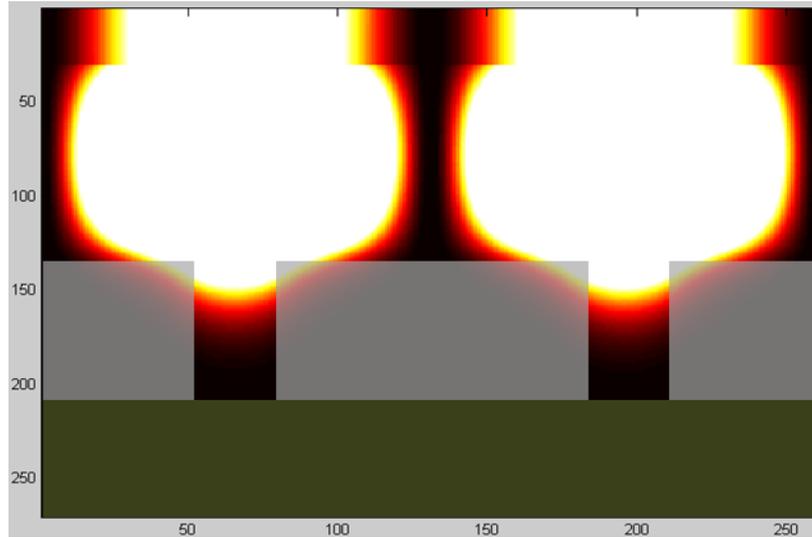
(a) Near field contour of orthogonal polarization (TM) intensity with normal illumination



(b) Near field contour of parallel polarization (TE) intensity with normal illumination



(c) Near field contour of orthogonal polarization (TM) intensity with off-axis illumination at an angle of 47.93°



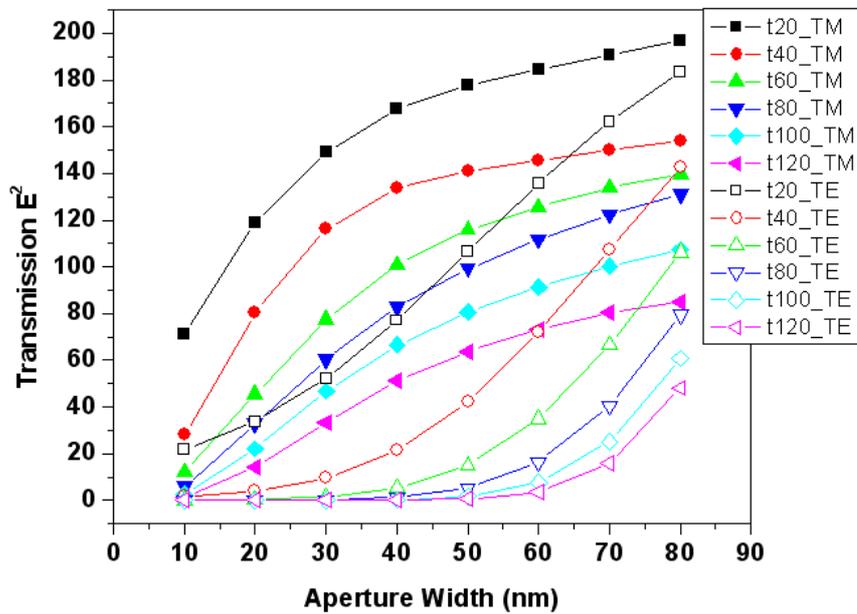
(d) Near field contour of parallel polarization (TE) intensity with off-axis illumination at an angle of 47.93°

Figure 4-29 Near field contours of orthogonal polarization and parallel polarization intensities.

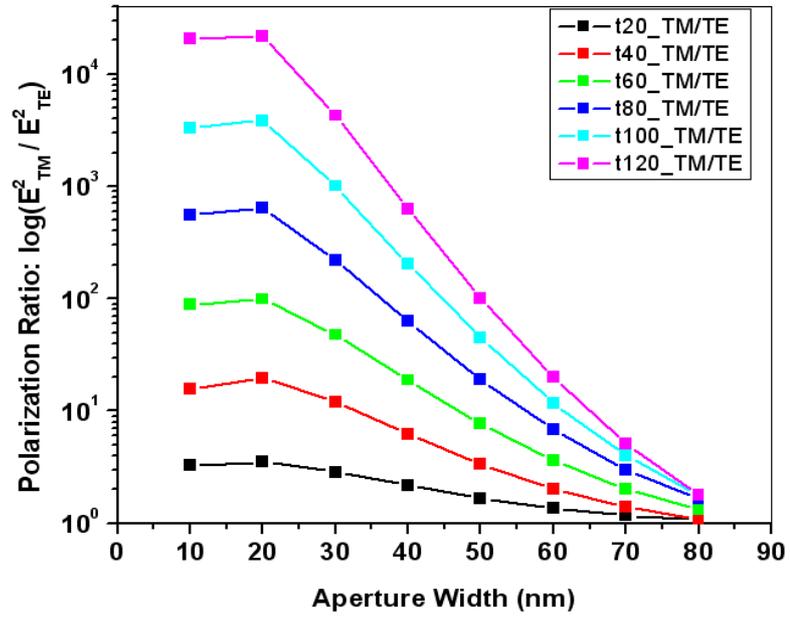
The polarization ratio is defined as the maximum integrated E^2 of the electric field orthogonal to the gratings (TM) divided by the maximum integrated E^2 of the electric field parallel to the gratings (TE). We simulate the impact of aperture geometry on polarization ratio and transmission. The aperture width is varied from 20 to 80nm, and the thickness is varied from 10 to 130nm with a step of 20nm. Figure 4-30 (a) shows the maximum E^2 of TM and TE polarizations vs the aperture width and thickness. Both the TM and TE transmission increases with aperture width, while it decreases as aperture gets thicker. The polarization ratio is shown in Figure 4-30 (b). The peak ratio is at aperture width around 20nm, and the ratio becomes worse as width increases. Thicker aperture masks improve the polarization ratio. So a smaller width and larger thickness will improve the polarization ratio, however a larger width and smaller thickness increase the transmissions. A polarization monitor has to balance the design to meet the

requirements in both polarization ratio and transmission throughput. A width W_a around 30nm with ± 10 nm range, and thickness t around 80nm with ± 15 nm range will ensure a polarization ratio larger than 20, which means the TE polarization is smaller than 5% of TM polarization, and at the same time allow enough light to go through to the detector.

As shown in Figures 4-29 (c) and (d), and discussed in section 4.2.2.1, TE polarization contributes to image contrast mostly in high NA imaging, whereas TM polarization degrades image contrast. The analysis of the two beam interference pattern is shown in Figure 4-31, which is adapted from Figure 3-52. Consider that the interference beams transmit in x, z plane, and the wafer image plane is at $z=0$. The TE polarization is along the y direction. The TM polarization has both x and z components, so the electric fields at x, z , and y directions [4-18] are:



(a) Maximum TM and TE transmission E^2 vs. aperture width and thickness



(b) Polarization ratio of TM over TE vs. aperture width and thickness

Figure 4-30 Illustration of the polarization ratio as a function of aperture width and thickness

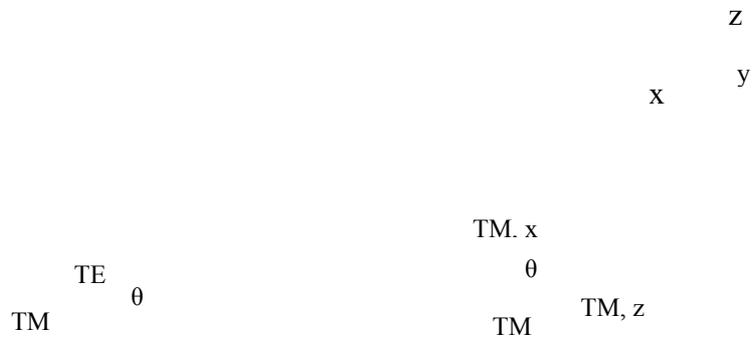


Figure 4-31 High NA imaging of two beam illuminations. The $NA (\sin\theta)$ is 0.74.

$$\begin{aligned}
E_x &= E_{TM} \cos \theta \cdot e^{-ik_x x} \vec{x} + E_{TM} \cos \theta \cdot e^{ik_x x} \vec{x} = 2E_{TM} \cos \theta \cos(k_x x) \vec{x} \\
E_z &= E_{TM} \sin \theta \cdot e^{-ik_x x} \vec{z} + E_{TM} \sin \theta \cdot e^{ik_x x} \vec{z} = 2E_{TM} \sin \theta \sin(k_x x) \vec{z} \\
E_y &= E_{TE} e^{-ik_x x} \vec{y} + E_{TE} e^{ik_x x} \vec{y} = 2E_{TE} \cos(k_x x) \vec{y}
\end{aligned} \tag{4-9}$$

Therefore the image intensity of the TM polarization is:

$$\begin{aligned}
I_{TM} &= E_x \cdot E_x^* + E_z \cdot E_z^* = 4E_{TM}^2 [\cos^2 \theta \cos^2(k_x x) + \sin^2 \theta \sin^2(k_x x)] \\
&= 4E_{TM}^2 [\cos^2(k_x x)(1 - 2\sin^2 \theta) + \sin^2 \theta] \\
&= 4E_{TM}^2 [\cos^2(k_x x)(1 - 2NA^2) + NA^2]
\end{aligned} \tag{4-10}$$

Hence the TM image contrast is:

$$C_{TM} = \frac{I_{TM,\max} - I_{TM,\min}}{I_{TM,\max} + I_{TM,\min}} = |1 - 2NA^2| \tag{4-11}$$

If $NA = \frac{\sqrt{2}}{2}$, $\theta = 45^\circ$, $C_{TM} = 0$. In Figure 4-31, $\theta = 47.93^\circ$, and $NA = \sin \theta = 0.74$, so the

TM image contrast is only around 0.09. For TE polarization, we have:

$$\begin{aligned}
I_{TE} &= E_y \cdot E_y^* = 4E_{TE}^2 \cos^2(k_x x) \\
C_{TE} &= \frac{I_{TE,\max} - I_{TE,\min}}{I_{TE,\max} + I_{TE,\min}} = 1
\end{aligned} \tag{4-12}$$

Therefore high NA imaging does not impact the TE contrast. However, as the incident beam angle approaches 45° , the TM contrast becomes zero. In order to detect TE polarization, the grating orientation can be tuned 90° to make the electric field orthogonal to the gratings. The rotated gratings with 45° and 135° can monitor 135° and 45° linear polarizations respectively. The horizontal, vertical, and the diagonal gratings could be

designed on the aperture mask simultaneously to investigate the linear polarization components and ratios with a single exposure.

4.4 Fabrication of the Integrated Aerial Image Sensor (IAIS)

IAIS fabrication consists of two main flows: the assembly photo-detector chips onto wafer carrier templates, and the patterning of the IAIS aperture mask onto the photo-detector chips. The fabrication flows are shown in Figure 4-32, where label (1) indicates the wafer carrier assembly, and label (2) indicates the aperture mask patterning.

The main challenge of the assembly is to place the photo-detector chips co-planar to the wafer carrier surface for lithography process and its precise integration (well within near-field) with an aperture mask. This chip assembly problem has no mature, high yield technique solutions. Patterning sub-wavelength grating, e.g. 30nm openings, presents another challenge to MOS manufacturing. Advanced techniques to pattern small features, such as immersion lithography, e-beam lithography, EUV lithography, and etc, are employed in recent years. In the Berkeley Micro-fabrication Laboratory, we use e-beam system Jeo1107 EBL to pattern aperture mask.

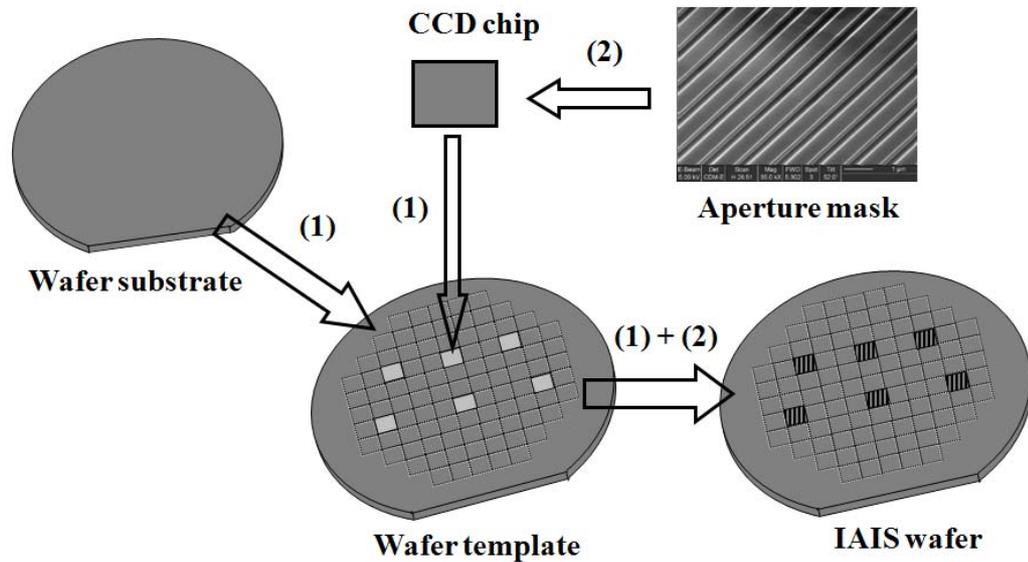


Figure 4-32 Illustration of IAIS fabrication flows

Since the e-beam system is designed to pattern small chip, and the assembly technique is in development, we developed the two processes, assembly chip to wafer and patterning mask to chip separately and in parallel, with given process limitations imposed by the presence of the photo-detector. The aperture mask was designed for proof-of-concept testing in the GCA-6200 G-line stepper. The third challenge is the insertion of the IAIS prototype onto the stepper wafer chuck, while meeting requirements of wafer topography flatness, lithography depth of focus, and overall compatibility with the existing automated wafer handling system.

While full assembly and testing of an IAIS prototype proved to be beyond our limited resources, in this section we discuss what we learned about the key steps involved in such an experiment.

4.4.1 Aperture Mask Material Deposition

Amorphous silicon is deposited by sputtering technique. In the Berkeley Micro-fabrication Laboratory we used Novellus m2i sputtering system. Physical Vapor Deposition (PVD) process is used to sputter amorphous silicon to the target chip. The temperature is below 100°C, and the deposition rate is well controlled and suitable for this application. To avoid damaging the CCD chip, the collimator is employed between target and chip. In order to optimize the process conditions, we use 2² factorial experiment, as shown in Figure 4-33, to investigate the power and temperature effects on the refractive index of the deposited material. The power of 10% and 30% (out of 3kW max), and temperature of 20°C and 75 °C were used. Each condition was repeated two times.

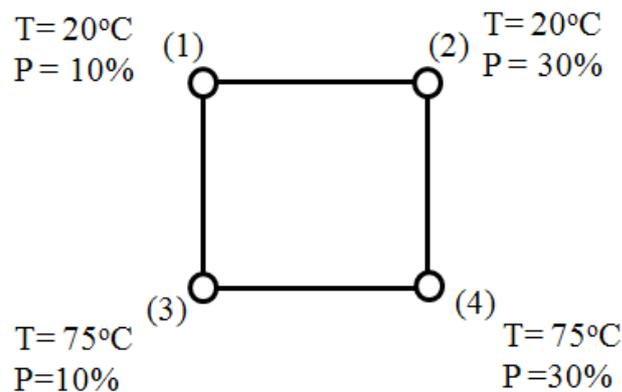


Figure 4-33 Illustration of 2² experimental design of the sputtering process. The temperature is 20°C and 75°C, and the power is 10% and 30% of 3kW.

The refractive index and the film thickness were measured by the Sopra ellipsometer, as shown in Figure 4-34. The ellipsometer wavelength is ranging from 220nm to 500nm. We show RTI [4-10] refractive index measurement results in the graphs as a reference.

The measured imaginary index shifted up to the maximum value of 1.0 to the reference value across wavelengths, while the real index is close to the reference at small wavelengths. The index becomes smaller at high temperature and low power, but the curve shapes remain similar. The measurement repeatability at high power is better than that at the low power. The enlarged imaginary part provides advantages of high signal to noise ratio and polarization ratio. We finally select temperature of 20°C and power of 30% (out of 3kW) as the deposition conditions.

The film thickness uniformity is measured by Ellipsometer on the 4 inch dummy wafer depositions. The mean of the film thickness is 150nm, and the across wafer variation is about 10nm. The wafer to wafer variation of the mean thickness is 3nm.

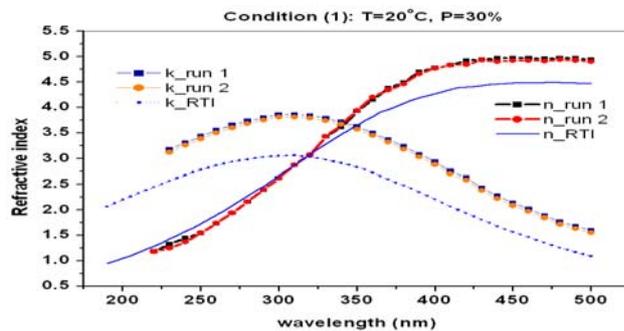


Figure (a) refractive index at temperature 20°C, and power 30% of 3kW.

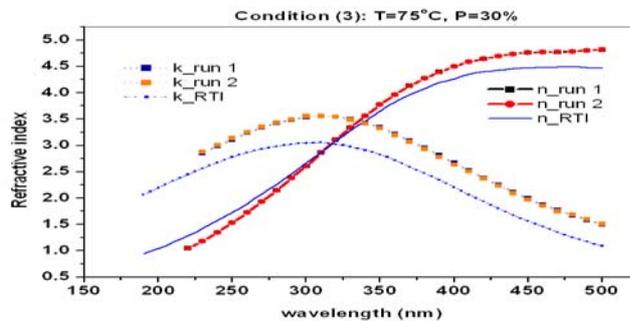


Figure (b) Refractive index at temperature 75°C, and power 30% of 3kW.

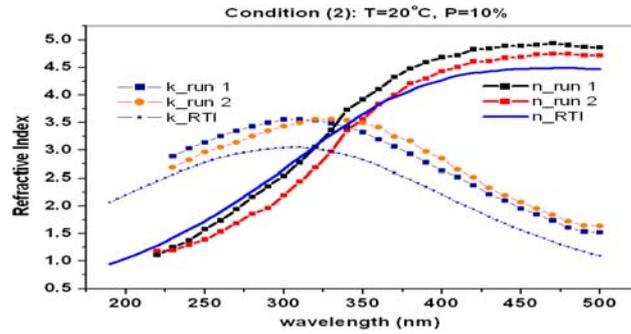


Figure (c) Refractive index at temperature 20°C, and power 10% of 3kW.

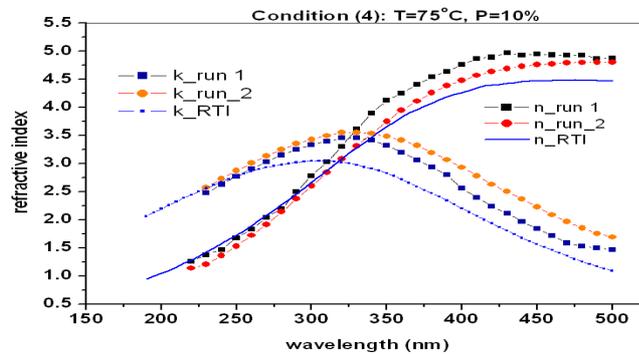


Figure (d) Refractive index at temperature 75°C, and power 10% of 3kW.

Figure 4-34 Amorphous silicon refractive index measurement results at 2² experiment conditions of the sputtering deposition.

4.4.2 Aperture Mask Patterning

We targeted the GCA 6200 wafer stepper in the Berkeley Micro-lab as the exposure tool to do IAIS proof-of-concept experiments. The resolvable feature size is around 1.1 μ m in the production environment and can be pushed around half of these limits in the research facility. Accordingly, we designed the aerial image pattern with a pitch of 1.2 μ m, 1:1 line/space pattern. By using the optimization method in section 4.2.2.2, we target the

aperture mask thickness of around 100nm to 300nm, and the aperture width of 100nm to 300nm.

PMMA is the positive resist used in E-beam lithography. It is capable of very high resolution of around 20nm, but it suffers from poor sensitivity and etching resistance. The aperture mask patterning process flow is as follows:

- HMDS coating and spin PMMA 950
- Prebake: 75°C, 1 hour, hotplate
- E-beam lithography
- Development: 1:3 MIBK to IPA
- Hard bake: 75°C, 3 minutes
- Plasma etching

Exposure dose and line spacing pitch are two important parameters that impact the printed CD in E-beam lithography. We vary the grating duty cycle of 125nm line-width by 1:1 and 1:9.6, and 250nm line-width by 1:1 and 1:4.8. The exposure dose is from $20\mu\text{C}/\text{cm}^2$ to $400\mu\text{C}/\text{cm}^2$, with a step of $20\mu\text{C}/\text{cm}^2$. The electron energy is fixed at 40kV.

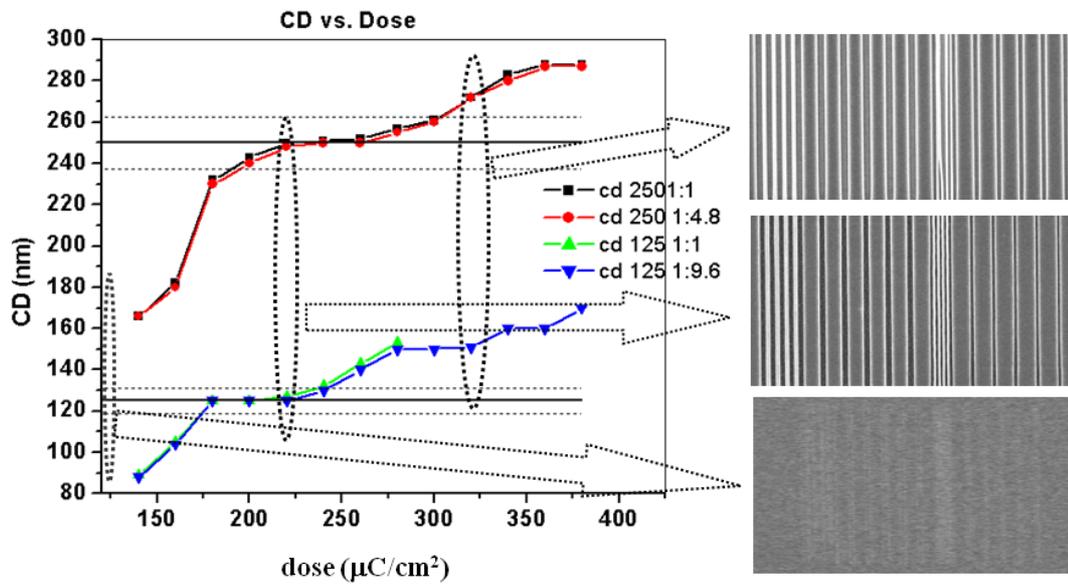


Figure (a) The plot on the left shows printed CD vs. dose vs. drawn CD and pitch. The right SEM picture shows the printed image at the small, medium and large dose conditions.

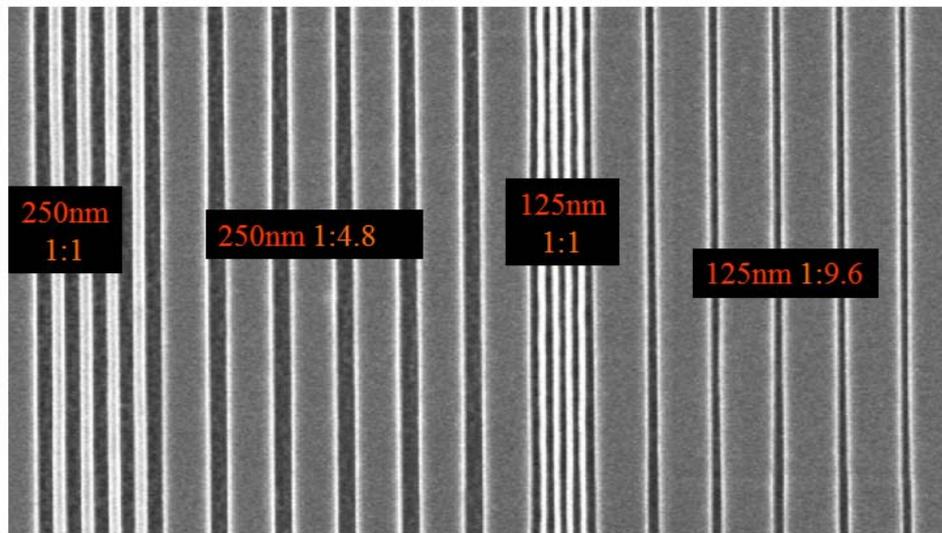
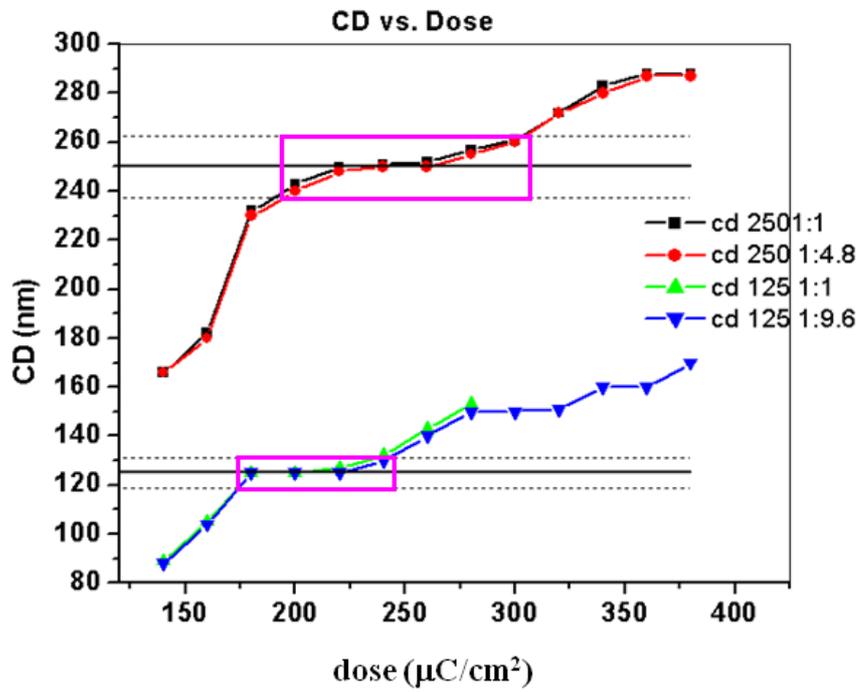


Figure (b) The top plot shows the optimal exposure dose window of the two draw CD patterns.

The bottom is the SEM picture of the printed patterns in the dose window.

Figure 4-35 E-Beam lithography printed CD vs. dose vs. drawn CD and pitch

Figure 4-35(a) shows the printed CD as a function of dose, drawn CD and pitch. The right SEM pictures are the printed image at three dose conditions. At a small dose of around $80\mu\text{C}/\text{cm}^2$, the pattern starts to print but not clearly. The second picture is the printed pattern at the medium dose around $220\mu\text{C}/\text{cm}^2$, where the pattern is well defined. At large doses bigger than $320\mu\text{C}/\text{cm}^2$, over-exposure occurs. The optimal dose window is shown in the purple box in Figure 4-23(b) top plot. For 125nm line-width, the optimal dose range is from $170\mu\text{C}/\text{cm}^2$ to $250\mu\text{C}/\text{cm}^2$, while $200\mu\text{C}/\text{cm}^2$ to $320\mu\text{C}/\text{cm}^2$ is the optimal range for 250nm line-width. The bottom plot in Figure 4-35(b) shows the SEM picture of the printed patterns at dose of $250\mu\text{C}/\text{cm}^2$.

One of the limitations of the Jeol107 EBL system is the small field of $100\mu\text{m} \times 100\mu\text{m}$ at the resolution of interest. A large writing field has to be divided into small areas that must be stitched together. Since the EBL system has no laser control stage, the limited stage moving accuracy causes a field to field stitching error of around $2\mu\text{m}$. Using alignment marks is another option, however with 100nm alignment marks, the alignment error is around $1\mu\text{m}$. By taking the above concerns into account, we folded the aperture mask grating array into a matrix. The grating pitch is $1.2\mu\text{m}$, and the length is $20\mu\text{m}$. There are 24 aperture groups with 10 apertures in one group. The sampling step is 50nm . The distance between aperture groups is: $7.2\mu\text{m} + 0.05\mu\text{m} = 7.25\mu\text{m}$. The aperture mask grating matrix is shown in Figure 4-36. The spatial shift measurement results are shown in the figure around the main matrix. For the horizontal array, the designed distance between apertures is $7.25\mu\text{m}$, and the alignment difference of the first aperture between the vertical arrays is: $5 \times 50\text{nm} + 50\text{nm} = 300\text{nm}$. From the SEM measurement results, the aperture group spatial shift is $50\text{nm} \pm 5\text{nm}$. The small variation of the sampling steps

will cause the sampled signal to be slightly asymmetric around the peak intensity; however the sampling fidelity will not be affected significantly. Some defects were evident on the mask, which may cause signal perturbations at the contaminated aperture groups.

The Lam5 Poly-Si TCP (Transformer Coupled Plasma) Etcher is used to etch aperture mask. Cl_2 is the main process gas in the etch recipe, and HBr is used to improve the selectivity over oxide. The E-Beam pictures of the etched patterns are shown in Figure 4-37. Figure (a) is the pattern image from top to bottom view. The etched pattern CD uniformity is good over the fields. We found that there is a relatively wide transition region at the edge. The image is tilted at the right side to see the sidewall profile, where a trench at the grating upper edge is noticeable. In order to see the sidewall section, a cut hole inside the grating was made by E-Beam cut, as shown in Figure 4-37 (b). This confirmed the over-etching at the aperture edge. The upper side of the aperture edge was etched away, which caused the aperture side-wall to have two parts with different aperture widths, the upper being bigger than the lower. The source of the over etch is the poor etch resistance of PMMA. An additional hard mask could be added to strengthen the etch resistance.

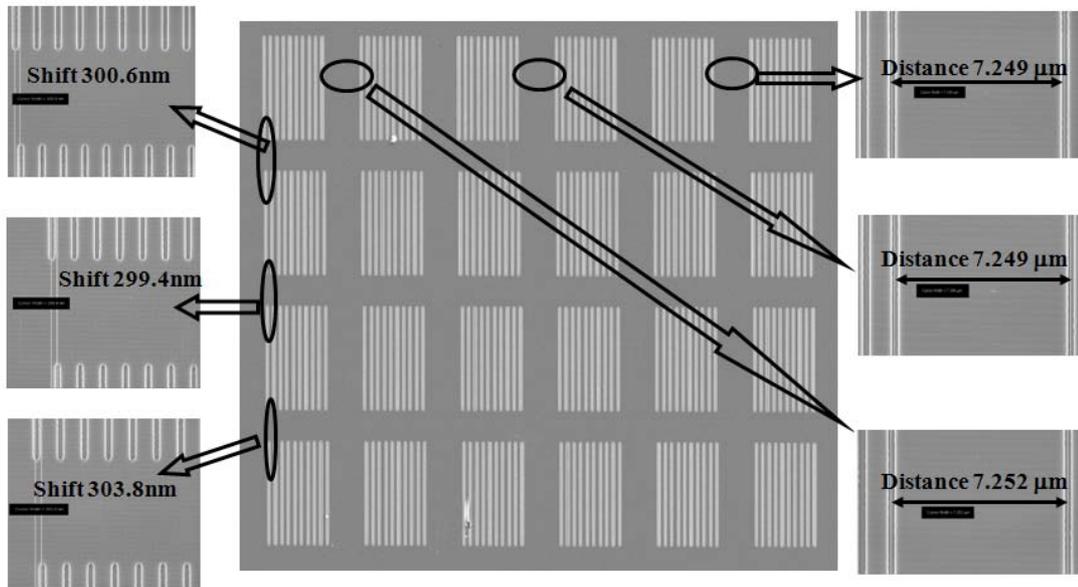
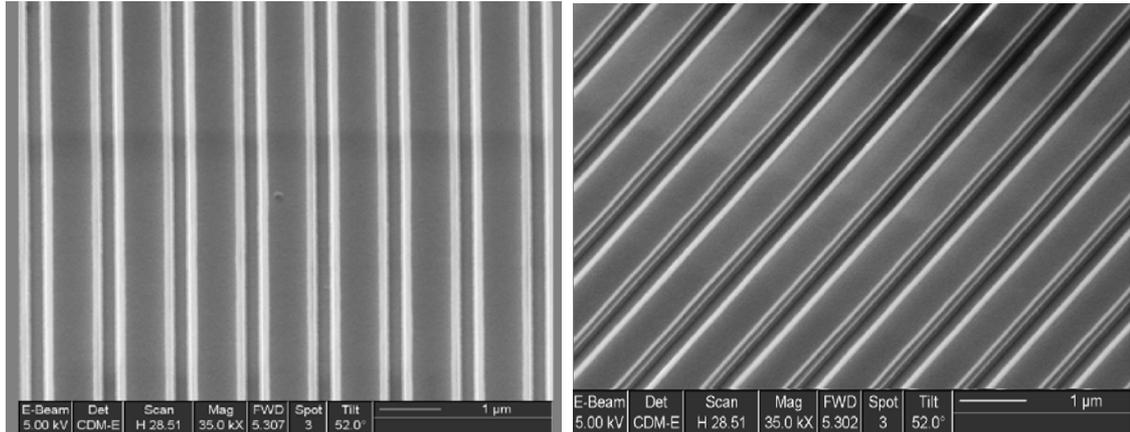
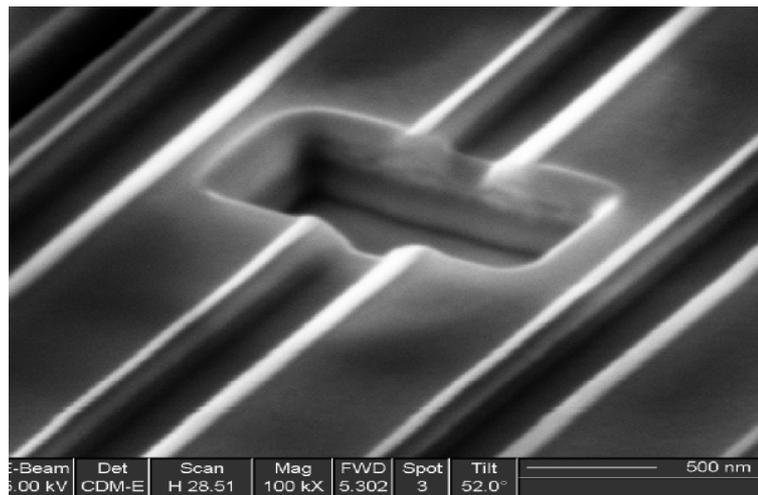


Figure 4-36 Aperture mask grating folding matrix by E-Beam lithography

Since the two process flows, aperture mask patterning and chip assembly, were executed in parallel and the work of integration of the IAIS circuit board into the stepper proved to be very time consuming, we have not been able to perform the hard mask process. However, our patterning work provides the information of the feasibility of patterning IAIS aperture mask by the current fabrication facilities, and provides the basis to improve the process robustness, which is the main objective of this fabrication effort.



(a) E-Beam picture of the etched aperture mask patterns. The image on the left is from top to bottom view, and the image on the right is tilted to show the sidewall profile.



(b) E-Beam cut hole in the gratings to show the side wall cross section.

Figure 4-37 E-Beam pictures of the etched aperture mask patterns

4.4.3 IAIS Assembly

Assembly of a chip co-planar to the wafer carrier surface is one of the challenging topics in MEMS processing. Two techniques have been carried out in this project: adhesive bonding and capillary force assisted self-assembly. In this section, we explore

the possible techniques, investigate the advantages and disadvantages of each method, and provide useful information for the future integration work.

Adhesive bonding has the advantage of low temperature, high bonding strength, and compatibility with the standard clean-room processing. The main disadvantage is the limited long-term temperature stability of the polymer material. The negative epoxy-type polymer SU-8 is used as the intermediate bonding material in the IAIS assembly. Diced silicon chips are used to demonstrate the bonding process. In our experiment, nine square holes slightly larger than the

CCD chip were opened on the wafer carrier by silicon bulk etching around 130minutes. The depth of the hole is 330 μm , and the thickness of the CCD chip is 323 μm . The bonding process is shown in Figure 4-38.

The wafer carrier has a dehydration treatment on a hot plate at 200°C for 2 hours. Then a 10 μm layer of SU-8 was spun on the wafer carrier, and a 2 μm layer of SU-8 was spun on the small dummy chips. After 2 minutes soft-bake at 75°C, the wafer carrier and the small chips are bonded together. The pairs are then put into the vacuum environment of the Suss MA6 aligner with vacuum contact exposure of 85mJ/cm² UV. A pressure of 5-6 barr is applied. Post bake is at 65°C for 3 minutes. Then two minutes development with SU-8 developer is used to remove the SU-8 at the un-exposed area. Finally, the assembled pairs are placed in a vacuum for 7 minutes to further remove the potential bubbles.

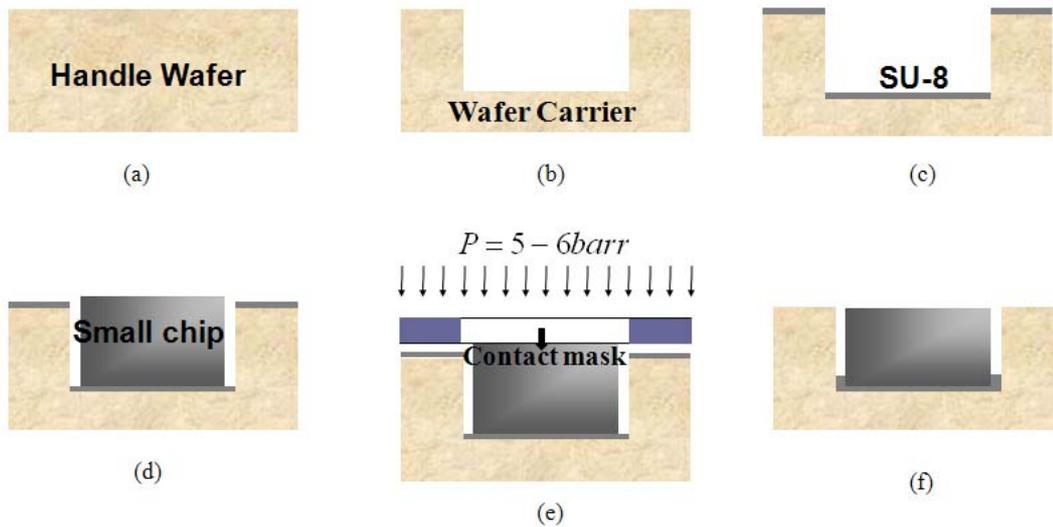


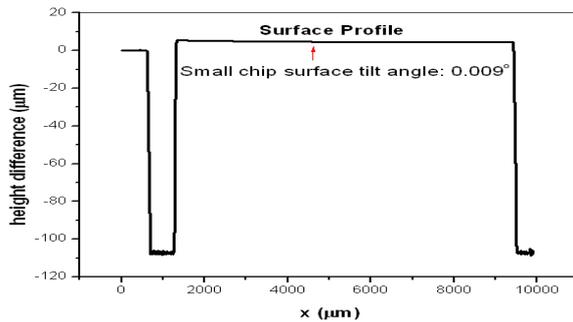
Figure 4-38 Illustration of the adhesive polymer bonding process

The profilometer measurement results of the two sampled positions are shown in Figure 4-39. The measurement is along the longer side of the chip at the center line. The first profile is from the layout receptor at location d as shown in Figure 4-39. In this experiment we have achieved very good co-planarity, as the surface of the small chip is tilted only about 0.009 degrees from the surface of the wafer holder. However the small chip is higher than the wafer carrier by about $5.47\mu\text{m}$. Since it is likely that non-uniform pressure was applied across the wafer, it was apparently insufficient at the receptor location d to make the chip coplanar to the carrier. The gap between the chip and the receptor hole at the left side is larger than that of the right side, due to the absence of in-plane alignment control during the bonding process. The second profile is from layout location i , as shown in Figure 4-39, where the Si bulk etching has the largest etch rate and therefore the receptor hole is the deepest at $345\mu\text{m}$, which is $15\mu\text{m}$ deeper than the target. Even without the contact mask pressure, the CCD dummy chip is easily embedded into

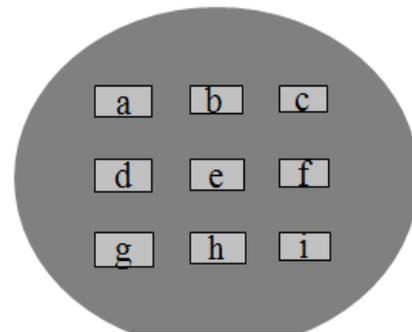
the receptor hole. The surface of the CCD dummy chip is lower than that of the wafer carrier with the maximum height difference of $8.79\mu\text{m}$, while the surface tilt with reference to the wafer surface is still coplanar with only about 0.002° of tilt.

One of the challenges in achieving the correct height of the embedded chip is the limited silicon bulk etch rate uniformity. The etch uniformity is $\pm 5\%$, where the minimum receptor thickness is $318\mu\text{m}$ at position e , and the maximum thickness is $348\mu\text{m}$ at position i . The second problem is the small chip in-pane tilt in the receptor trench. In order to solve the etch rate non-uniformity and the difficulty of the application of the standard bonding tool's pressure to the small chip-large carrier pairs, we propose a new process flow in Figure 4-40. On that figure, steps (a) through (e) illustrate the wafer receptor fabrication process. The basic idea is to pattern small holes around the large receptor holes, so that the polymer can be drained away when the CCD chips bond with the carrier. The small drain is etched from the front side of the wafer. The mask layout is shown in Figure 4-40(c). Then the wafer is etched through from the backside with the size of the big receptor hole and bonded with another wafer substrate, as shown in Figure 4-40 (e). The width W_2 is slightly larger than the chip size, while the W_1 is much smaller than W_2 and t_1 is smaller than two thirds of the wafer thickness. The detailed value of W_2 and t_1 needs more experimental optimization to take the bonding process and the fluid flow properties into account. To alleviate the need for a precise careful alignment of the big receptor hole and the small drain hole, the front-side process can be simplified to a larger square, whose width is equal to $W_1 + W_2$ and thickness equal to t_1 . Based on this design, the bottom surface non-uniformity from the bulk Si etch can be avoided, and the thickness of the wafer receptor will be well controlled. Then SU-8 is placed into the

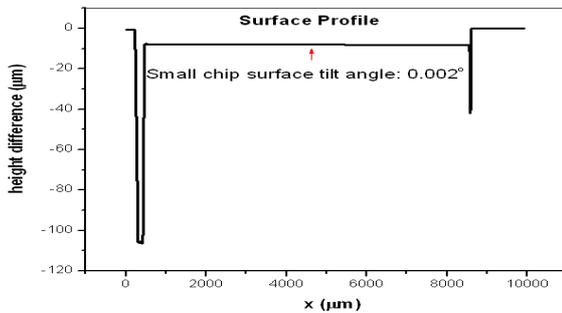
wafer carrier by the EFD 1500 dispenser, the small chip is picked and placed into the wafer receptor. At the same time pressure is added to the chip surface. The assembly process can be implemented by the die-bonding tool automatically, which is under construction in Berkeley Micro-fabrication Laboratory.



(a) Measured surface profile at wafer receptor d.



Wafer receptor layout



(b) Measured surface profile at wafer receptor i.

Figure 4-39 Profilometer measured surface profile of the assembly wafer. The wafer receptor layout is shown in the right side of the plots.

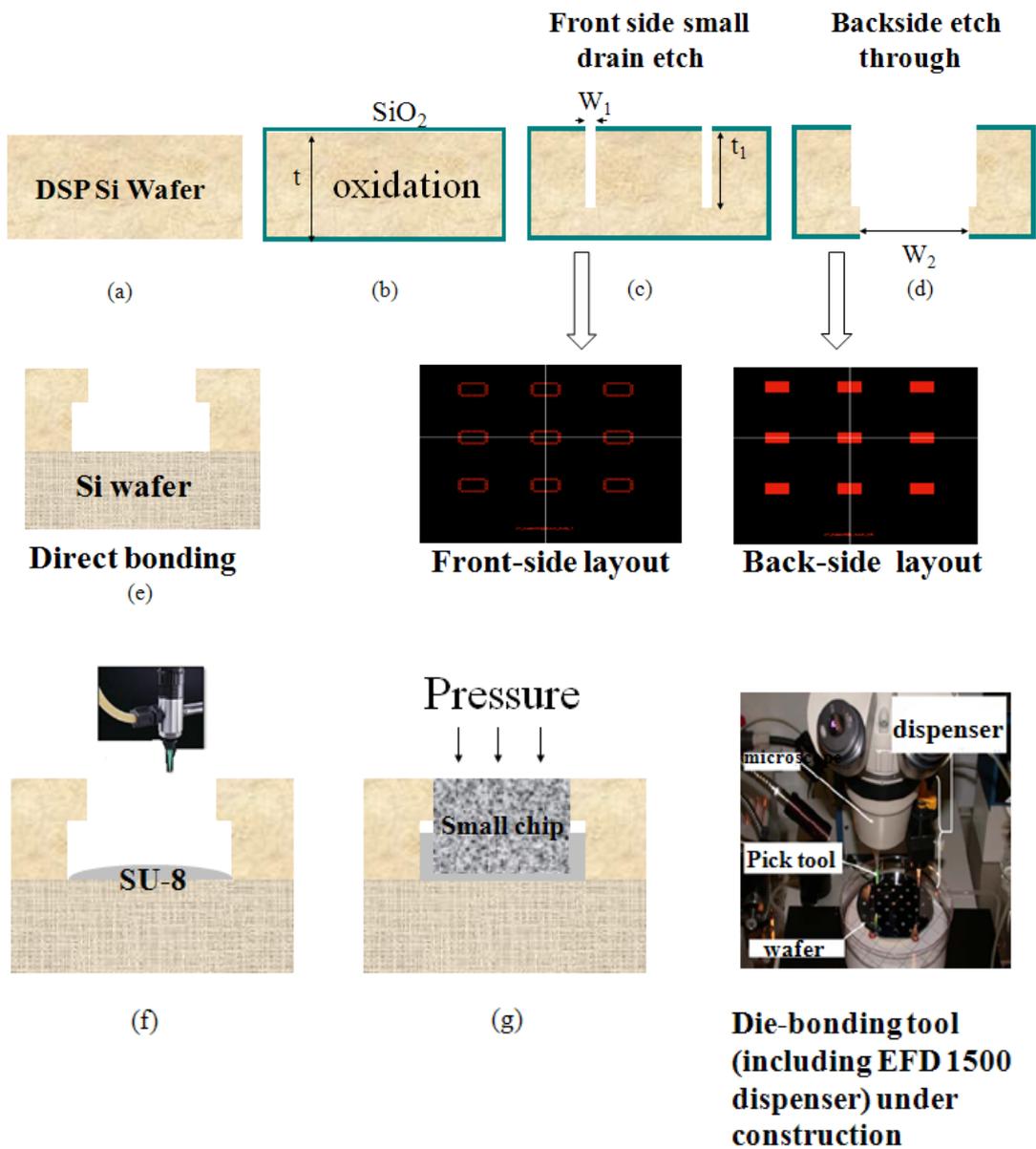


Figure 4-40 Proposed sequence of the new adhesive bonding process

To prevent the CCD chip front side from being damaged in the bonding process, we modify the above process from (e) to (g), as shown in Figure 4-41. Before the etched wafer carrier and the Si substrate are bonded together, a layer of Omnicoat is deposited

onto the wafer substrate, and the CCD chip is picked and placed upside down into the wafer carrier. Then the SU-8 is dispensed to the receptor hole and the wafer surface, and the pressure from the die-bonding tool is applied to the back side of the CCD chip. Another wafer substrate is bonded quickly from the front side of the assembled wafer and chips. The Omnicoat is removed by the MCC developer, and the assembled wafer and chips are separated with the lower wafer substrate. Finally a hard bake is performed to remove bubbles and solidify the bonding.

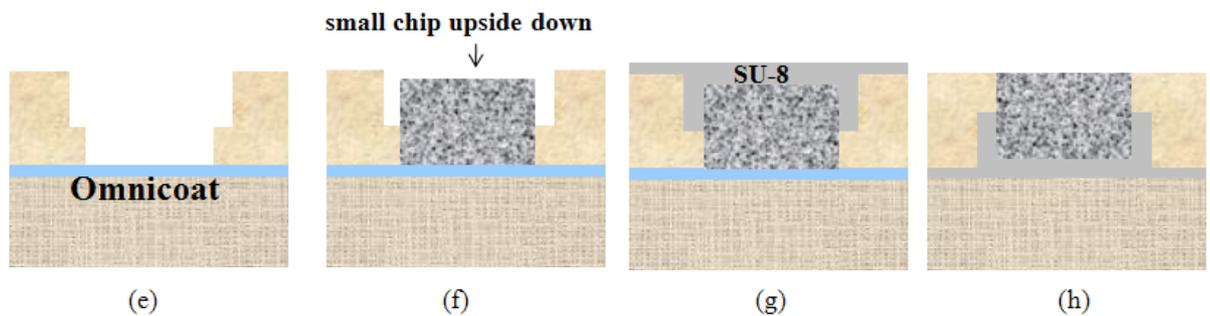


Figure 4-41 Modified process of the adhesive bonding in Figure 4-40 (e) – (g)

A potential problem of the adhesive bonding is that the in-plane alignment of the small chips to the slightly oversized holes during the bonding cannot be well controlled. Therefore a lateral alignment force in the direction of the plane is necessary to assist the assembly. Capillary force is one of the candidates that can deliver sub-micron self-alignment accuracy. We used and slightly modified the process flow proposed by E. P. Quévy, and *et.al* [4-19], which is shown in Figure 4-42. The silicon wafer carrier is prepared with 1 μ m LPCVD low stress nitride film. The front-side nitride membrane is patterned to match the CCD contact pads, as shown in Figure 4-43. Nine square holes

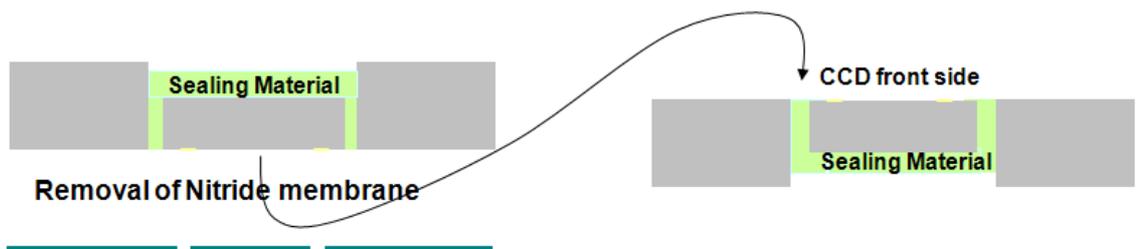
slightly larger than the CCD chip were opened from the backside of the wafer using 25% TMAH.



(a) Wafer carrier and CCD dummy chip alignment pads



(b) Self-assembly of CCD chip with wafer carrier



(c) Sealing CCD chip with wafer carrier

Figure 4-42 Illustration of the capillary force assisted assembly process

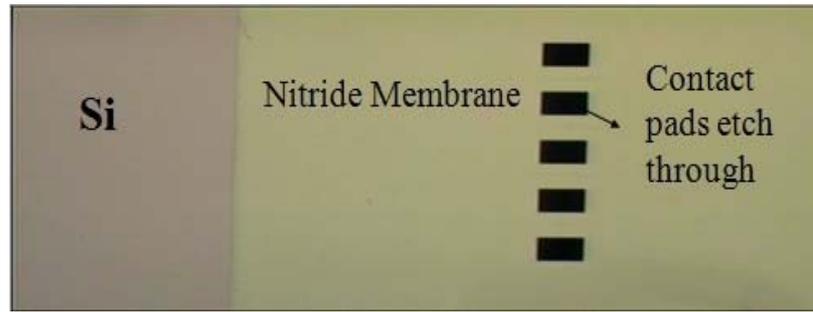


Figure 4-43 Picture of the nitride membrane with contact pads etched through at the backside of the wafer carrier.

The single layer mono-molecular film, Fluorinated Organosilane (FOTS: Trichloro(1*H*,1*H*,2*H*,2*H*-perfluorooctyl) silane—CF₃(CF₂)₅(CF₂)₂SiCl₃), is selectively coated on the aluminum oxide of the contact pads on the chips, and makes the surface hydrophobic. The contact angle measurement of the dummy CCD chip is shown in Figure 4-44. The Al alignment pads become highly hydrophobic with a contact angle of 119.1°, and the passivation layer of SiO₂ remains hydrophilic (or lightly hydrophobic) with an angle of 64°. Next, a droplet of DI water is deposited into the Si carrier hole, and then the CCD dummy chips are manually placed into the holes from the backside of the wafer. As the water progressively evaporates, the capillary and gravitational forces make the small chips adhere to the nitride membrane, and at the same time, the fluidic interaction provides the lateral alignment due to the hydrophobic alignment pads and the surface tension around the membrane holes. Finally, the assembled chips are sealed into the carrier holes with the temporary material, high temperature low viscosity epoxy glue.

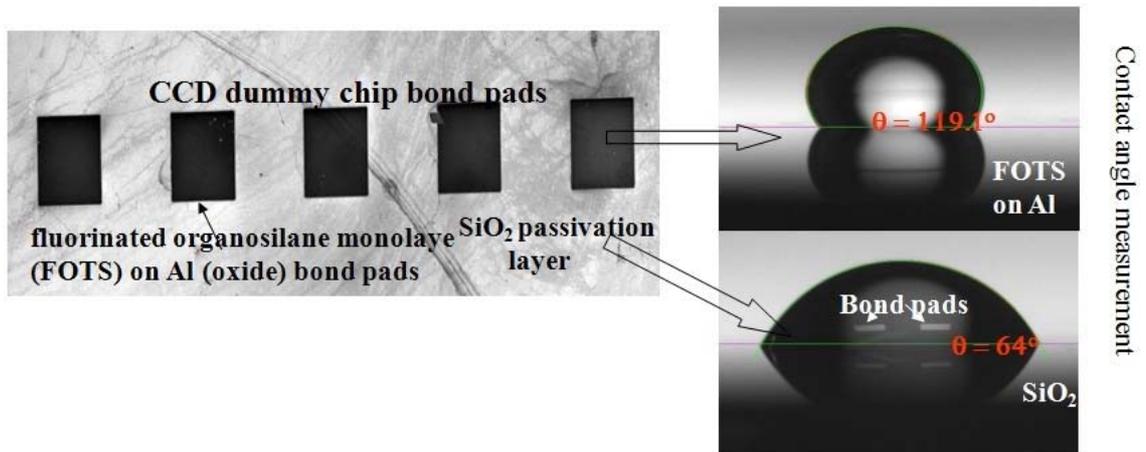


Figure 4-44 Contact angle measurements of the alignment pads and the passivation layer. Left is the SEM picture of CCD dummy chip pads. Right is the contact angle measurement results.

Profilometer measured surface profiles are shown in Figure 4-45. The measurement is conducted along the two sides (x and y directions) of the dummy chip. The chip size along the x direction is 8.4mm, and the size along the y direction is 5.5mm. The cut line at the center position of each side is shown in the attached 2D plots. The surface of the CCD dummy chip is higher than that of the wafer carrier with the maximum height difference of 1.17 μ m, and it is tilted about 0.00034 $^\circ$ to the surface of the wafer holder along the x-direction. The reason that the edges of the chip appear to be elevated, is that the epoxy glue flows into the boundary of the chip. Moreover, the epoxy glue makes the sealing process difficult to operate and control. Sedimentation of the SiO₂ beads followed by depositing Ge is reported as a permanent sealing method [4-19]. However, the higher temperature (350 $^\circ$ C) Ge deposition is too high for this process and the multiple-cycle sedimentation might further limit its yield.

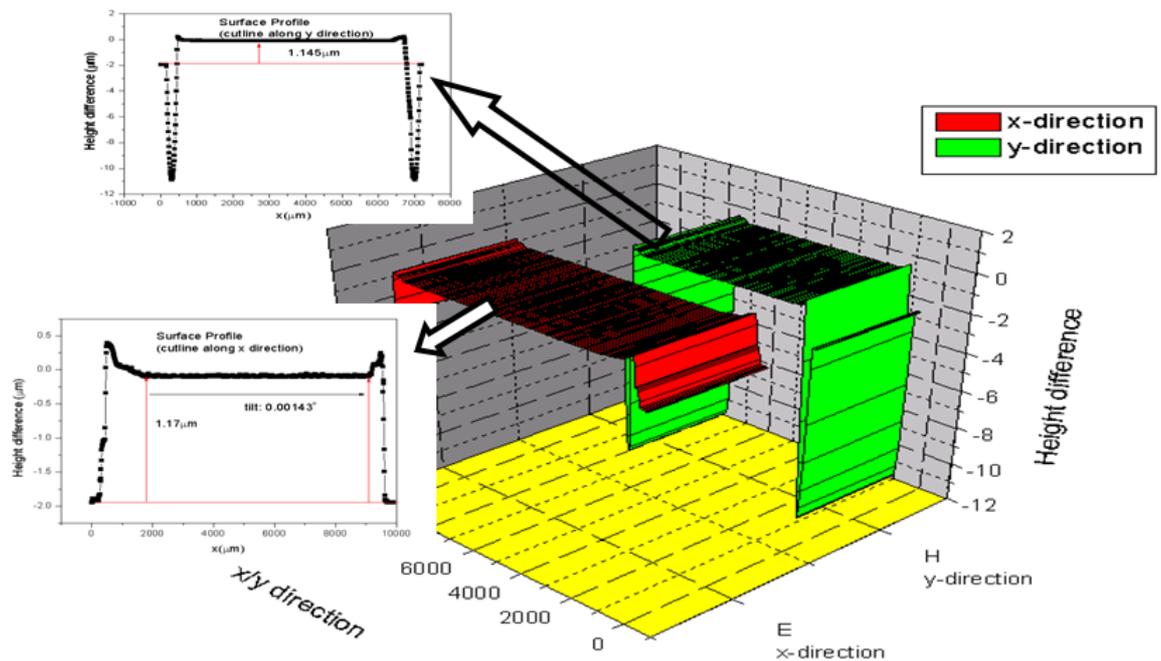


Figure 4-45 Profilmeter measurement results of surface profile of assembly pairs

During this assembly, the easily damaged nitride membrane stands alone to support the assembly chips. In order to avoid breaking the membrane and integrate the aperture mask to the CCD chips densely, we modify the process as follows: Amorphous Si, which is the material of the aperture mask, is deposited on the nitride film before the assembly of the small chips. Then, both the amorphous Si and the nitride membrane are patterned to match the alignment pads on the CCD chip. After etching the carrier holes from the backside of the wafer, the nitride film is etched away in SF₆ plasma to expose the surface of amorphous Si. Next, the CCD chips will be embedded and sealed into the carrier holes. Finally, amorphous Si is patterned as the aperture mask patterns.

In this section, we discussed two assembly techniques: adhesive bonding and capillary force assisted self assembly. The processes are performed in Berkeley Micro-lab. Both

process sequences have advantages and disadvantages. Adhesive bonding provides a feasible assembly method, but wafer receptor etch uniformity and assembly pressure uniformity may limit our ability to achieve consistent CDD chip height. A new process flow is proposed to solve the above problems. Capillary force applies a lateral in-plane alignment force to the CCD chips during assembly. A better assembly surface profile is achieved in this process. We proposed a new process flow to improve the fabrication yield, while the sealing material is still open for improvement.

4.4.4 IAIS Prototype Testing in the Berkeley Micro-fabrication Laboratory

In this experiment we designed an IAIS proof of concept vehicle that involved a wired PC board carrying the CCD chip, mounted on a 4-in carrier wafer. These tests were carried out by manually loading this prototype onto the stage of a G-line stepper in the Berkeley Micro-fabrication Laboratory. The sequence is described next.

The 4 inch dummy wafer was patterned with a large alignment grid using the GCA 6200 wafer stepper. The field size is slightly larger than the CCD chip size. The dummy printed chip is a model to align the CCD chip in a field. The CCD chip is then fixed on the thin wafer in one field by adhesive bonding. We designed the IAIS circuit board as shown in Figure 4-46 for the Kodak KAF0402 CCD. The board connects the CCD chip at one side, and communicates to the computer from the other side. The CCD contact pads are connected with the circuit board connection pads by using the 7400B Wire-bonders in Berkeley Micro-lab. The entire assembly is manually loaded on the stepper stage, and the wire leads must be very carefully handled.

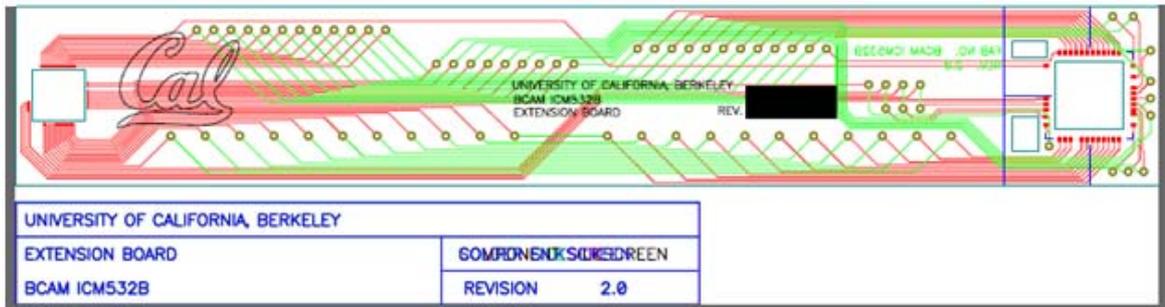


Figure 4-46 IAIS circuit connection board layout

The IAIS testing prototype is shown in Figure 4-47. We access the inside of GCAWS2 chamber through the side doors to manually load and position the assembled wafer at the chuck. The stage vacuum switch is turned off, then we place the alignment bar for proper y positioning. The wafer slides in the y direction to align with the bar and slides in the x direction to be centered on the chuck. The vacuum switch is turned on after the alignment. The next exposure procedures follow the manual.

One of the most challenging problems in this sequence is securing the wafer on the chuck. Since GCAWS2 vacuum on the wafer chuck is not sufficiently strong to support the assembled wafer, and the length of the circuit board is 15cm, which also connects to another cable switch to the computer, any slight vibration of the cable during the testing will move the wafer away from the alignment. Although we can use other materials, such as tapes, to fix the wafer on the chuck, this must be done after the alignment and right before the exposure, where the sensitive projection lens is in close proximity to the wafer and the CCD chip. The solution at the time was to increase the stepper wafer vacuum strength, align the wafer, circuit board, cable and cable switch at the same level, and keep the wafer movement to minimum. We made the holders to level the circuit board and

cables with the wafer at a good alignment. However, in the end the wafer chuck vacuum was proved to be insufficient for this operation. Right before the exposure after we press the EXPO button, the wafer chuck has another automatic alignment to the projection lens, which cannot be controlled manually. During this alignment, the wafer, the circuit board and the cable have to move with the chuck, and this proved impossible to control. This technique problem causes a big challenge to the IAIS testing in the GCA6200 stepper.

In retrospect, the proper vehicle for testing this concept should have been completely wireless, self-contained platform that has substantially the size of a real wafer. While this is technically possible, and it has been used for both temperature testing on bake plates, plasma chambers and scanner stages, it was not possible for us to implement with our CCD sensor with our limited time and resources.

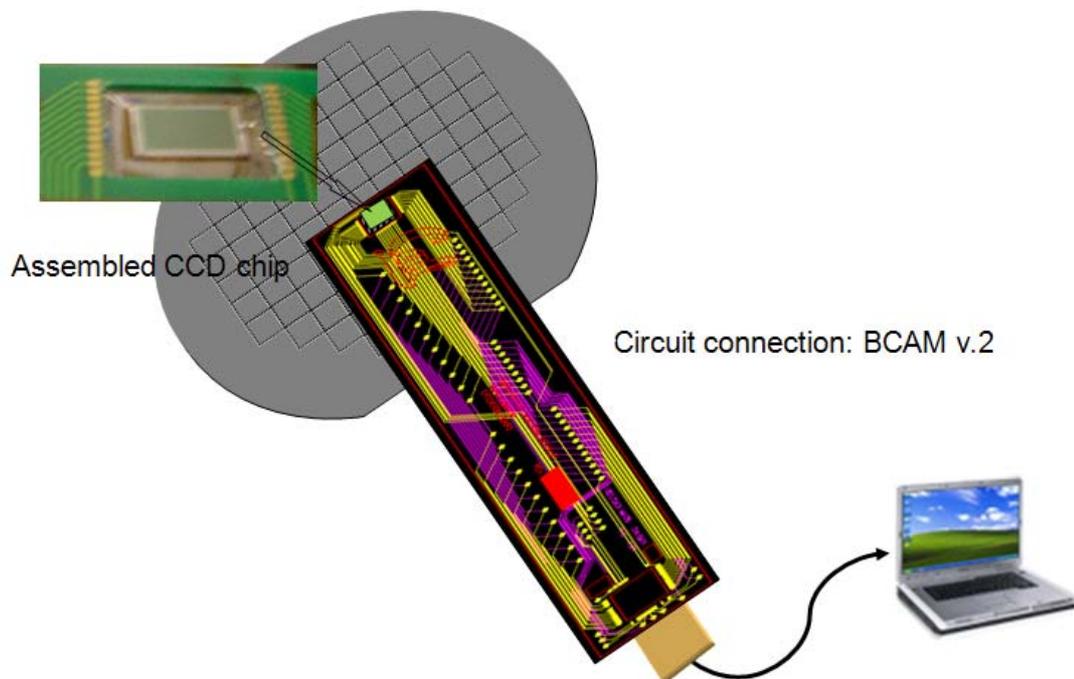


Figure 4-47 Illustration of the IAIS testing prototype.

4.4.5 IAIS Source of Error

The major sources of error of IAIS fabrications are summarized in table 4-4.

Table 4-4 IAIS Source of Error

Source of error	Status	Estimation	Process	Future solution
Aperture width	measured	250±10nm	E-beam & Plasma Etch	DUV lithography with large field size
Aperture thickness	measured	150±5nm	Sputtering	Deposition rate well controlled
Aperture sidewall	not measured	None	Etch	Etching recipe
Sampling step	measured	50 ±5nm	E-beam & Etch	Litho & etch process and data fitting technique needed
Aperture mask material	measured	Refractive index measured	Sputtering	Adjust material for certain applications
IAIS assembly (CCD tilt)	measured	F: 0.0005°	Assembly	New process proposed & robust solution needed
CCD noise	provided	10pA/cm ²	Signal measurement	Appropriate aperture mask design
Communication noise	Not measured	None	Signal measurement	Reduce CCD, circuit board, cable, and image processing program communication noise
Defocus	not measured	None	Signal measurement	Improve assembly system surface height and flatness

4.5 Conclusions

In this chapter, we discuss the design, manufacturing and applications of an integrated aerial image sensor. Since a practical photo-detector pixel size is still on the order of a few micrometers, while the area image under testing is generally in the order of a few nanometers, we designed a series of “moving” sub-wavelength sampling groups, which we incorporated into an aperture mask that interferes with the aerial image, and ultimately forms a low spatial frequency interference pattern. This low spatial frequency reconstructed image can be captured by the photo-detector arrays, and achieves a spatial resolution of a few nanometers. The aperture mask is placed on the CCD detectors and integrated on the wafer surface.

Targeting the 65nm technology node (but capable to scale to more advanced nodes), the optimum aperture mask geometry parameters, such as aperture width and thickness, are designed based on the detector image contrast, signal to noise ratio and photocurrent contrast. Aperture width of $35nm \pm 15nm$ and thickness of $80nm \pm 20nm$ is nearly optimal for the 65nm node, and the simulations show that fabrication precision is not critical. A good choice of the mask material depends on the specific applications. A high extinction coefficient of the refractive index benefits the signal to noise ratio, the polarization ratio, and the transmission throughput. A large real index can improve the polarization ratio. Based on the device performance, fabrication ability and equipment availability in Berkeley Microlab, amorphous silicon is selected as the aperture mask material in the IAIS fabrication.

Defocus aberration causes aerial image intensity and contrast degradations. The IAIS detector image can capture the intensity variation trends with defocus, and provide an

efficient way to measure defocus variations on the binary periodic patterns. The maximum signal to noise ratio of the contrast achieved at the best focus plane, and the contrast difference between the best focus and 0.01λ defocus being 50 times larger than the contrast noise indicate that the IAIS' ability to detect defocus aberration is within 0.01λ . The intensity normalized slope (*NILS*) and slope variations of the detector image follow the same trend and range of those of the aerial image, and the CD feature size keeps the same value of the aerial image with the proper dimensional scale. IAIS can also detect the image shift, whose precision is determined by the sampling frequency and the detector noise. Phase shift masks are used as the advanced process monitors. IAIS as the on-wafer mask can monitor the aerial performance of the phase shift mask. Simulation results of the 90° phase shift mask indicate that the detector images closely follow the aerial image trends and could efficiently monitor the aberration effects. IAIS may also be the basis of an on-wafer linear polarization monitor. A width around 30nm with ± 10 nm range, and thickness around 80nm with ± 15 nm range will ensure the ratio of TM polarization to TE polarization to be larger than 20, and at the same time keep the enough transmission through-put. By considering high NA lithography, the grating orientation can be tuned 90° to make the electric field orthogonal to gratings consistent with the aerial image TE polarization.

IAIS fabrications consist of two main steps, the patterning of the IAIS aperture mask and the assembly of photo-detector chips onto wafer carrier templates. E-beam lithography and plasma etch are the main processes to pattern the aperture mask. Because of the stitching problem of EBL system, the grating array is folded into a matrix. The grating pitch is $1.2\mu\text{m}$, and length is $20\mu\text{m}$. There are 24 aperture groups with 10

apertures in one group. The measured sampling step is $50\text{nm} \pm 5\text{nm}$. A good CD uniformity is achieved after plasma etching. However, over-etch at the upper side wall is noticed due to the poor etch resistance of PMMA. Because of the overhead parallel process work and the limited time and budget, we have not been able to perform the hard mask process, but the patterning work provides the useful information to improve the process robustness and verifies the feasibility of the patterning work at the available fabrication facilities.

Adhesive bonding and capillary force assisted self-assembly are explored as the assembly methods. The profilometer measurement results of the surface profile indicate that the silicon bulk etching non-uniformity causes the assembly chip to chip surface non-planar. To solve this problem, a new process flow is proposed. The second problem of the assembly is the small chip in-pane tilt in the receptor trench. Capillary force is used to improve the self-alignment accuracy. The capillary and gravitational forces make the CCD dummy chips adhere to the nitride membrane, and the fluidic interaction provides the lateral alignment due to the hydrophobic alignment pads and the surface tension around the membrane holes. Profilometer measurement results show that the surface of the CCD dummy chip is higher than that of the wafer carrier with the maximum height difference of $1.17\mu\text{m}$, and the chip tilts about 0.00034° to the surface of the wafer holder along the x-direction. The epoxy glue flowing into the boundary of the chip causes the edges of the chip to be elevated slightly. In order to avoid breaking the membrane and improve the fabrication yield, we propose a new process flow. The permanent sealing material remains challenge in the process.

A Micro-lab testing prototype is described. A circuit board is designed to connect the CCD chip, and the cable to the computer. The testing system is placed in GCA 6200 wafer stepper chuck. A good alignment of the wafer, the circuit board, and the cable is achieved. Since the wafer chuck vacuum is too weak to support the assembly wafer, the wafer may slide away from the wafer chuck alignment center right before the exposure. This technique problem causes a big challenge to the IAIS testing in the GCA6200 stepper. More customer supports from the stepper vendor is expected to solve this problem.

Viable process techniques are explored in the IAIS fabrication work. Being aware of the full challenges in the assembly, patterning, and testing, we propose the technique solutions to the process flow, investigate the advantages and disadvantages of each process, and discuss the difficulties and feasibilities of the fabrication work. We hope that this work provides the needed insights to improve the sensor fabrications and improves its operations in the future.

The Integrated Aerial Image Sensor (IAIS) proposes a direct and efficient way to measure the aerial image in the wafer surface. This sensor would speed up the diagnostics of the exposure system, and avoid the bias imposed by the resist process. Measuring the lens aberrations of lithography exposure tool is one of the examples. The lens aberrations, such as defocus are traditionally tested off-line through measuring the exposed and developed photo-resist profile using scanning electron microscope (SEM), or optical microscopy. However, the resist processing has strong impact on the resist image characteristics which makes it difficult to extract the exposure tool information. The SEM based measurement is sensitive to the variations of different SEM equipments and the

performance of different operators. By direct measurement of the aerial image, the sensor is capable of several tasks, such as aerial image feature calibrations and lens aberration extractions. The aerial image information is important and useful to distinguish the source of process variations. By combining with the variable design of the top mask, IAIS can play a role of an advanced process monitor. After conquering the fabrication and technique challenges, IAIS is expected to provide a more stable metrology platform with higher throughput and lower cost.

Finally, it is conceptually possible to merge the features of the PPG defocus monitor described in Chapter 3, with the on-wafer CCD metrology described in this Chapter. This application would potentially provide a very sensitive *in-situ* focus calibration method.

References

- [4-1] T. A. Brunner and R. R. Allen, "In Situ measurement of an image during lithography exposure", *IEEE Electron Device Letter*, Vol. EDL-6, NO. 7, 1985
- [4-2] T. A. Brunner, and R. R. Allen, "In Situ resolution and overlay measurement on a stepper", *Proc. of SPIE*, Vol. 565, pp6, 1985
- [4-3] T. A. Brunner and A. M. Stuber, "Characterization and setup techniques for a 5X stepper", *Proc. of SPIE*, Vol. 633, pp106, 1986
- [4-4] T. E. Adams, "Application of Latent Image Metrology in Microlithography", *Proc. of SPIE*, Vol. 1464, pp294, 1991
- [4-5] A. K. Pfau, R. Hsu, W. G. Oldham, "A two-dimensional high-resolution stepper Image monitor", UC. Berkeley, 1990
- [4-6] C. H. Fields, "Direct Aerial Image monitoring for extreme ultraviolet lithography systems", Ph. D. Dissertation, UC Berkeley, Spring 1997
- [4-7] T. Hagiwara et. al., "Self calibration of wafer scanner using aerial image sensor", *Proc. of SPIE*, Vol. 4691, pp871, 2002
- [4-8] S. Hemar et. al., "CD uniformity control using aerial image – based mask inspection", *Proc. of SPIE*, Vol. 5038, pp929, 2003
- [4-9] David Attwood, "Soft X-Rays and Extreme Ultraviolet Radiation", *Cambridge University Press*, 2007
- [4-10] <http://www.rit.edu/kgcoe/microsystems/lithography/thinfilms/thinfilms/>
- [4-11] Dan Ceperley, "Surface Plasmon Portion of Qualifying Examination Talk ", UC Berkeley, 2006
- [4-12] M. Lam and A. R. Neureuther, "Polarization masks: concept and initial assessment", *Proc. of SPIE* , Vol. 4691, pp437, 2002
- [4-13] S. L. Chuang, *Physics of Optoelectronic Devices*, 1995
- [4-14] J. W. Goodman, "Introduction to Fourier Optics", 2nd edition, *McGraw-Hill*, 1996
- [4-15] Nyquist-Shannon sampling theorem
- [4-16] <http://www.kodak.com/>

- [4-17] A. R. Neureuther, EE243 Lecture Notes, UC Berkeley, 2003
- [4-18] G. R. McIntyre, "Characterizing Polarized illumination in High Numerical Aperture Optical Lithography with Phase Shifting Masks", Ph. D. Dissertation, UC Berkeley, 2006
- [4-19] E. P. Quévy, et.al., "Reconstituted wafer technology for heterogeneous integration", *IEEE MEMS 2006*, pp.302, 2006

Chapter 5

Conclusions and Future Work

5.1 Conclusions

In this work we explore *in-situ*, portable monitoring methods that can be deployed at the reticle and wafer stages for lithography characterization. Two novel monitors, Probe Pattern Grating (PPG) Defocus Monitor and Integrated Aerial Image Sensor (IAIS), are introduced in this thesis. Both monitors can be conceptually implemented for *in-situ* monitoring. The PPG responds the defocus stimulus and transmits the resulting image to the photoresist whose pattern can be measured by scatterometry. The IAIS samples the aerial image and records its frequency-shifted results *in-situ* and in real time into a photodetector array.

Chapter 3 presents the pattern design, theoretical development, printing experiment and scatterometry measurement of the PPG defocus monitor. The PPG defocus sensitivity of around $0.83CF/RU$ at $\sigma=0.1$ is confirmed in both mathematical modeling results and the *SPLAT* simulation results. A good linear relationship between probe line intensity and the defocus aberration is shown from $-0.6RU$ to $0.3RU$ defocus range.

The defocus variations are translated to the probe line resist trench depth changes. A good linear approximation of the probe trench depth *versus* defocus is obtained in the simulation Bossung plot. This linear relationship is again verified in the printing experiments. An excellent correlation between the measured defocus and the scanner

programmed defocus is achieved. For the vertical grating, the mean defocus offset is 0nm, and the maximum standard deviation is 4.2nm. For the horizontal grating, the mean offset is around 0nm, and the maximum standard deviation is 6nm. A semi-empirical model is built based on the measured probe trench depth. The maximum mean of the residual is 0.7nm, and the maximum standard deviation of the residual is 0.73nm. The coma aberration, the astigmatism aberration, and the illumination angular non-uniformity are expected to be the main reason to cause the difference between the vertical and the horizontal lines. The sensitivity of around 1nm defocus / per nm of trench depth is higher than the existed defocus monitors based on phase shift mask techniques of around 3nm defocus /nm overlay, and other types of defocus monitors based on binary mask designs. The PPG monitor is likely the most sensitive defocus monitor currently available.

Chapter 4 presents the design and the possible ways to fabricate the IAIS. A sampling aperture mask transfers the high spatial frequency aerial image into a low spatial frequency detector pattern, which can be captured by the photo-detector arrays, and achieves a spatial resolution of a few nanometers. A parameter design targeting the 65nm technology node, but scalable to more advanced nodes is described. An optimal parameter window of aperture width of $35nm \pm 15nm$ and thickness of $80nm \pm 20nm$ is achieved, and the simulations show that fabrication precision is not critical.

The fabrication process of patterning the IAIS aperture mask and the assembly of photo-detector chips onto wafer carrier are explored. A considerable process is achieved by patterning the aperture mask “matrix” onto the dummy chips, however, the over etch at the grating upper sidewall indicates that a hard mask is needed for the future patterning process. Adhesive bonding and capillary force assisted self-assembly were explored. Two

new process flows are proposed to address the silicon bulk etching non-uniformity and the small chip in-pane tilt problems. Profilometer measurements of the capillary force assisted assembly show that the surface of the CCD dummy chip is higher than that of the wafer carrier with the maximum height difference of $1.17\mu\text{m}$, and the chip tilts about 0.00034° to the surface of the wafer holder. Epoxy glue flowing into the edge of the chip is the main source of the surface tilt. The permanent sealing material remains a challenge in the process sequence.

A Berkeley Micro-lab prototype has been deployed on the GCA 6200 wafer stepper chuck. Good alignment of the wafer, the circuit board, and the cable has been achieved. Unfortunately, a weak wafer chuck vacuum has proved to be too difficult to overcome without equipment modifications. A completely wireless, self-contained platform that has substantially the size of a real wafer would have been the appropriate solution to this challenge.

IAIS' capability to detect aerial image features, such as intensity, contrast, slope, and feature size is shown by simulation. Because of the structure of the design, IAIS is a natural polarizer. The grating orientation can be tuned 90° to detect TE polarization at high NA imaging. By combining the special design features of the top mask, IAIS can play the role of an advanced process monitor. It is specifically possible to merge the features of the PPG defocus monitor as described in Chapter 3, with the on-wafer CCD metrology in Chapter 4. This application would potentially provide a very sensitive *in-situ* focus calibration system.

5.2 Future Work

I. PPG Future Applications

First of all, a pair of gratings with opposite tone of phase can be designed on the PPG mask, then the intersection of the two linear plots will provide an accurate reference to the best focus position, as shown in Figure 5-1.

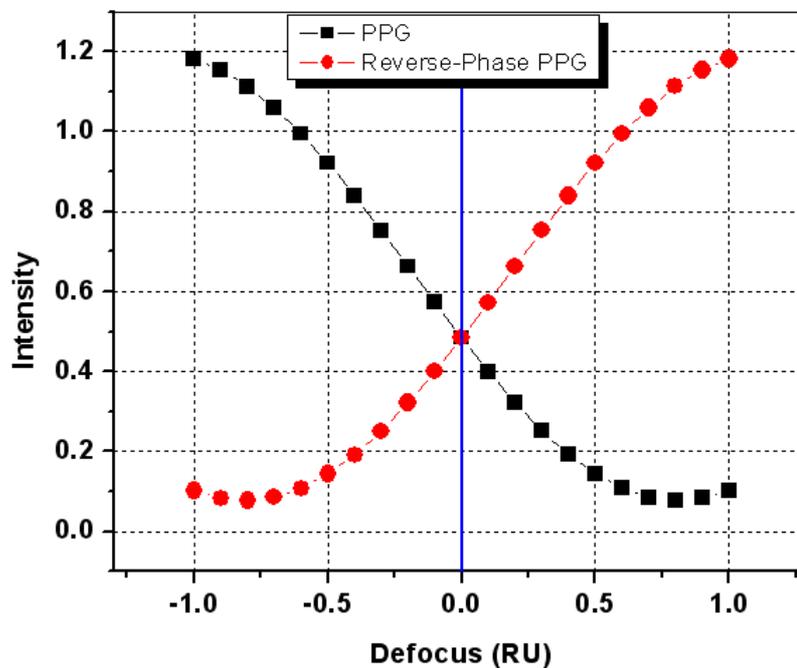


Figure 5-1 PPG vs. Reverse-Phase PPG. Aerial image simulation by SPLAT. The intersection of the reverse-phase gratings provides an accurate reference of the best focus condition.

Second, it is possible in tools with flexible illuminators (as used for source mask co-optimization) to include PPG on the product mask to monitor production chips. Advanced off-axis illumination schemes, such as Dipole, Quadrupole, and Annular illumination, are usually used in a production environment. For one dimensional grating,

quadrupole illumination has similar effects to dipole. The mutual coherence function of dipole [5-1] is expressed as:

$$|\mu(\hat{r}; \Delta_x, \Delta_y)| = \cos(2\pi\Delta_x \hat{v}) \cdot 2 \frac{J_1(2\pi\sigma\hat{r})}{2\pi\sigma\hat{r}} \quad \text{Eq. (5-1)}$$

where the distance of the pole center from origin is v and the radius of the circular poles is σ . J_1 is the first order Bessel function of the first kind. We use $\Delta_x = \hat{r}$, and $\Delta_y = 0$ and assume $v=0.7$. The mutual coherence function is drawn in Figure 5-4, where μ_{12} equals to μ in Eq. (5-1).

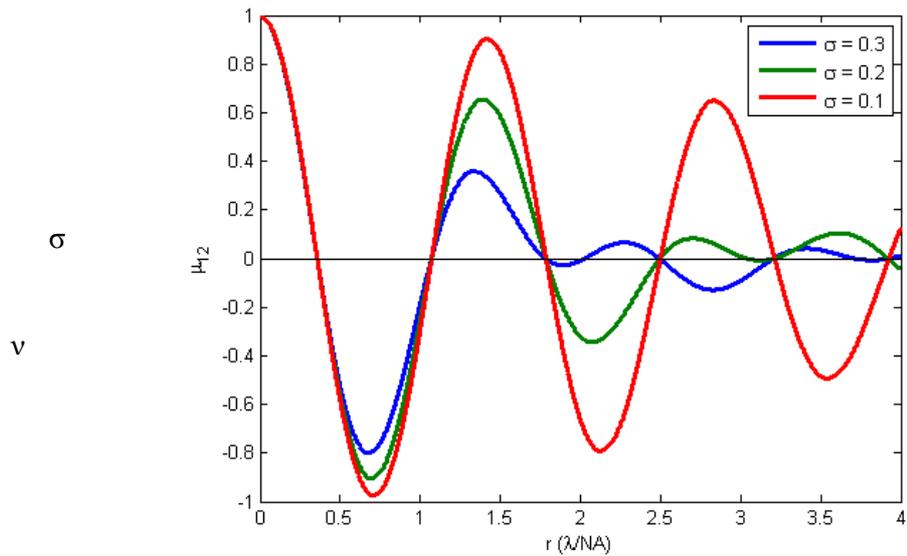


Figure 5-2 Mutual coherence function of dipole illumination. The left graph shows the Dipole scheme, and the right plot shows the mutual coherence function at three pole radius. μ_{12} is the mutual coherence function.

When r equals $0.55\lambda/NA$, the mutual coherence is around 0.6, and if v is increased to 0.85, the μ_{12} at $r = 0.55\lambda/NA$ is enhanced to about 0.85. The sufficiently large mutual

coherence ensures that the PPG designed in this thesis can be deployed in the production environment that employs Dipole illumination without significantly compromising the defocus sensitivity. The probe line phase may be reversed to -90° (or 270°) to be consistent with the negative mutual coherence phase. It may also be possible to synchronize the location of the grating lines and probes with the lateral variation of the mutual coherence function to enhance the sensitivity and create a practical image monitor for focus under double patterning production conditions.

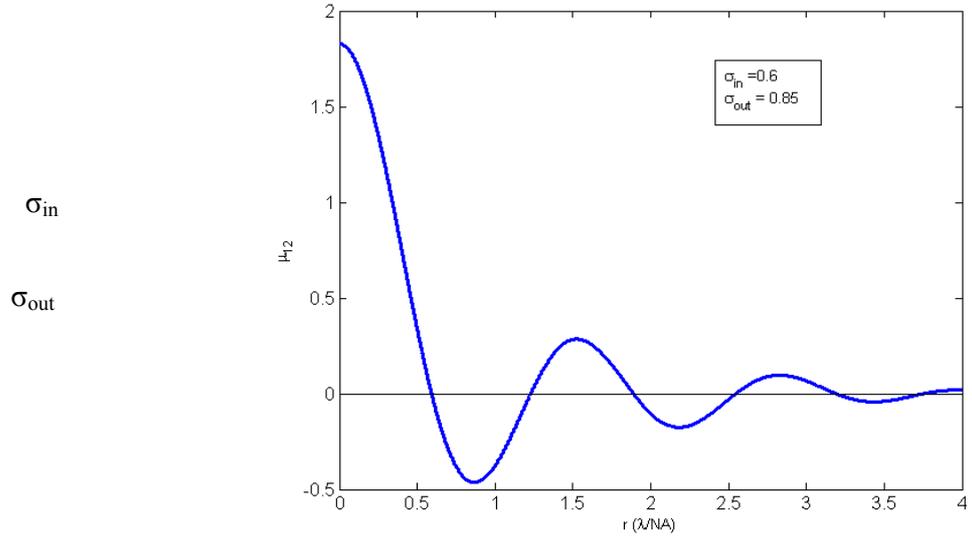


Figure 5-3 Mutual coherence function of Annular illumination. The left graph shows the Annular scheme, and the right plot shows the mutual coherence function at $\sigma_{in}=0.6$ and $\sigma_{out}=0.85$.

The mutual coherence function of Annular illumination [5-1] can be expressed as:

$$|\mu(\hat{r})| = \frac{1}{1 - \left(\frac{\sigma_{in}}{\sigma_{out}}\right)^2} \left[2 \frac{J_1(2\pi\sigma_{out}\hat{r})}{2\pi\sigma_{out}\hat{r}} - \left(\frac{\sigma_{in}}{\sigma_{out}}\right)^2 2 \frac{J_1(2\pi\sigma_{in}\hat{r})}{2\pi\sigma_{in}\hat{r}} \right] \quad \text{Eq. (5-2)}$$

The function with $\sigma_{in} = 0.6$ and $\sigma_{out} = 0.85$ is shown in Figure 5-3. At $r=0.55\lambda/NA$, μ_{12} is close to 0.2, which indicates that the annular illumination may not be an appropriate illumination condition for the PPG application.

Third, in a truly flexible future tool, an off-axis monopole maybe used to “program” the phase of the grating lines, thus eliminating the need for the extra cost of a phase shift mask. The binary mask strategy in Figure 5-4 is an example of this method [5-2]. The distance between the central probe line and the side pattern line is around $0.85\lambda/NA$. Thus, the probe line is at the minimum position of the first side lobe of the pattern line electric field spread function. At on-axis illumination, the probe line field and the proximity field spill-over from the pattern lines are at 180° phase difference, the central intensity is therefore at the minimum. As the monopole shifts away form the center, the phase difference between the probe line and the pattern line changes. The phase difference can be expressed as:

$$\Delta\varphi = \frac{2\pi}{\lambda} \frac{P}{2} \sin\theta = \frac{2\pi}{\lambda} \cdot (0.85 \frac{\lambda}{NA})(\sigma NA) = 1.7\pi\sigma \quad \text{Eq. (5-3)}$$

where σ indicates the lateral shift of the pole. So as $\sigma=0.6$, $\Delta\varphi \approx \pi$. Then the two surrounding pattern lines are in phase, and at the center probe line the original 180° phase difference between the probe line and the pattern lines is eliminated. So the electric field of the probe line and the proximity field of the patterns line is summed together. At this point, the intensity at the center line reaches a maximum. The SPLAT simulation result in Figure 5-5 verifies this conclusion. The mutual coherence effect may impact the results. The monopole radius can be increased relatively to eliminate the neighboring pattern proximity effect. The radius of 0.4 is used in this example.

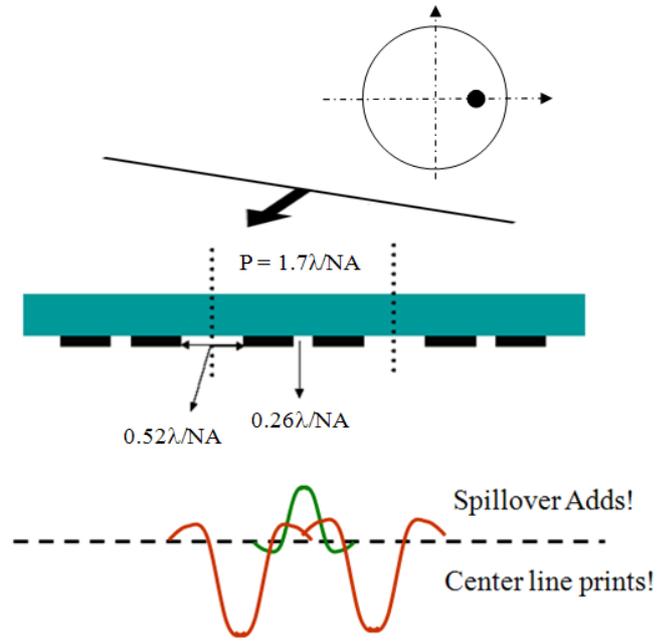


Figure 5-4 Illustration of the concept of an off-axis illumination monitor

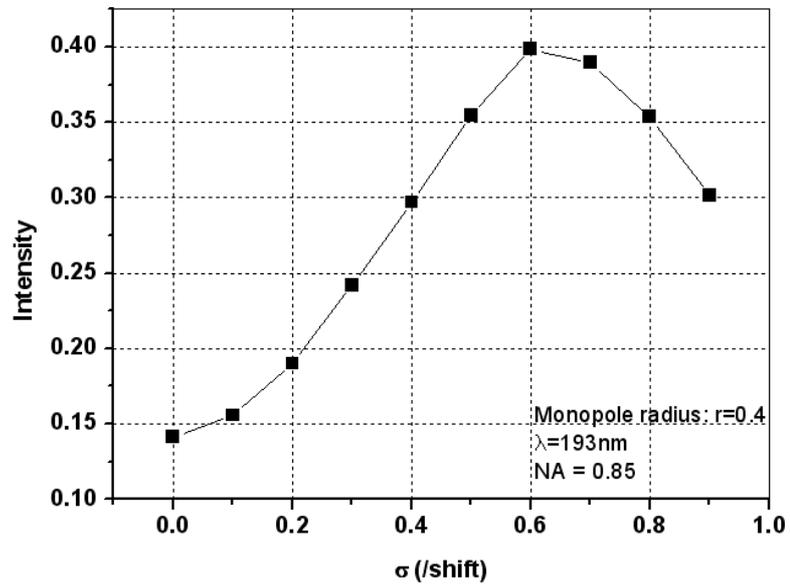


Figure 5-5 Illustration of probe line intensity vs. monopole shift, expressed by σ , by SPLAT simulation. The pole radius is 0.4.

II. IAIS Fabrication and Applications:

A proper vehicle for testing IAIS is a completely wireless, self-contained platform. Such a platform is technically possible, and has been used for temperature testing on bake plates, and plasma chambers initially at OnWafer Technologies, later by SensArray and now by KLA-Tencor. An enhanced bonding method has been recently developed and it appears promising as a potential platform for the IAIS. A robust solution of the heterogeneous assembly of a CCD detector into a base wafer will greatly speed up and ensure the successful IAIS deployment for routine, production use. In addition to its function as a recording tool, IAIS can be optimized as a special purpose monitoring system. A polarization monitor is the first candidate because of the nature of IAIS. A set of rotated IAIS gratings with 90° , 45° and 135° could be integrated into the same aperture mask to investigate the linear polarization components and ratios with a single exposure. Further, combining the purpose designed reticle with the on wafer IAIS will bring forth a broad application of the on-wafer sensors for lithography characterization.

References

- [5-1] A. K. Wong, "Resolution Enhancement Techniques", *SPIE*, 2001
- [5-2] A. R. Neureuther, "ODP Illumination Test Pattern", *FLCC Project Presentation*, UC Berkeley, April, 2007