

Copyright © 2002, 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.

**CAUSAL ANALYSIS OF
SYSTEMATIC SPATIAL VARIATION
IN OPTICAL LITHOGRAPHY**

by

Haolin Zhang

Memorandum No. UCB/ERL M02/27

12 September 2002

**CAUSAL ANALYSIS OF
SYSTEMATIC SPATIAL VARIATION
IN OPTICAL LITHOGRAPHY**

by

Haolin Zhang

Memorandum No. UCB/ERL M02/27

12 September 2002

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Causal Analysis of Systematic Spatial Variation in Optical Lithography

by

Haolin Zhang

B.S. (Beijing University, Beijing, China) 1987

M.S. (Beijing University, Beijing, China) 1990

A dissertation submitted in partial satisfaction of the
requirement 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 Ching-Shui Cheng

Fall, 2002

The dissertation of Haolin Zhang is approved:

Chair

Date

Date

Date

University of California, Berkeley

Fall, 2002

Causal Analysis of Systematic Spatial Variation in Optical Lithography

Copyright © 2002

By

Haolin Zhang

Abstract

Causal Analysis of Systematic Spatial Variation in Optical Lithography

by

Haolin Zhang

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Costas J. Spanos, Chair

The central theme of this thesis is the systematic analysis of the sources of variation in optical lithography. This work pursues data collection and data analysis techniques that are used to identify and model the various sources of variability in pattern transfer.

The first contribution of the thesis is that an automatic focus/exposure control method has been proposed based on digitized SEM scan using statistical feature extraction and neural network classification. Digitized CD-SEM traces are used to study die-to-die spatial variation and wafer-to-wafer or lot-to-lot temporal variation, induced by focus/exposure fluctuations for a 0.35 μm i-line process. Seventy-four percent perfect classification or ninety-six percent $\pm 5\text{mJ}$ match has been achieved for automatically detecting exposure, with 1σ prediction accuracy approximately equal to 1/18 of exposure window; forty perfect classifications or eighty-three percent $\pm 0.1\mu\text{m}$ match has been achieved for automatic detecting defocus settings, with 1σ prediction accuracy approximately equal to 20% of depth of focus window. This automatic focus/exposure control method can be extended to further extracting pattern profile information. Through intelligent data analysis techniques,

more process information can be obtained besides routine CD measurements, without resorting to complicated modeling of the interaction of the electron beam with samples. The work is of practical interest and research for implementation is being conducted in industry.

The second contribution of this thesis is a new technique to measure full-field lens aberrations using printed linewidth patterns. The basic concept of this method is that the dependency of the linewidth on individual aberration terms can be approximated by a Taylor series expansion about aberration free imaging. This approximation is valid, since the aberration of a state of the art lithography system is typically small. The expansion can be conducted under different process conditions. Therefore the aberrations can be deduced by means of numerical analysis. The experiment of this part of work utilizes electrical test patterns for off-line analysis, combined with reticle measurements to decompose the linewidth variability to the contributions of the optics, the reticle, the resist, and other random sources for a 0.22 μm process using 248 nm DUV lithography. This data set captures the spatial distribution of CD variation across die as well as across wafer. It is assumed that the deterministic within field variation is a major variation component, which consumes a large portion of the error budget. Experimental results indicate that the across wafer variation is at the order of 2.7nm; while across field variation is about 5.9nm for isolated vertical feature for this standard 0.22 μm process. Lens aberration residuals are estimated to be one of the most important sources of the within field systematical spatial variations. Generally all Zernikes have different sensitivity to linewidth under different process settings, such as defocus. Certain aberration terms, such as Z_3 , Z_8 and Z_{11} , have been found having significant impact on linewidth under non-zero nominal defocus settings. This aberration measurement method assumes an additive model ignoring interaction between different terms. Aberrations have been extracted at 456 points through exposure field using this method. Even though the results obtained from this method are currently not satisfactory, we believe that this method has promise once more

sophisticated, full-profile measurements become widely available and a broader set of feature types and orientations are measured under more varied process conditions. A detailed error analysis and future improvement are presented.

The third contribution of the thesis is a compact formulation of the mask error factor. Mask error magnification, as another important source of systematic spatial variation, has been studied theoretically. The rigorous formulation of mask error factor has been derived based on closed form aerial image calculation for coherent illumination condition. The results agree very well with first principle simulation and experiment.

Professor Costas J. Spanos
Committee Chairman

Table of Contents

List of Figures.....	iv
List of Tables.....	vii
Acknowledgements.....	viii
Chapter 1	
Introduction.....	1
1.1 Motivation.....	1
1.2 Thesis Goals.....	3
1.3 Challenge, Approach and Results.....	4
1.4 Summary.....	6
1.5 Thesis Organization.....	6
Chapter 2	
Background.....	11
2.1 Sources of Variation in Optical Lithography.....	11
2.2 Lens Aberration Background.....	16
2.2.1 A Brief Review of Lens Aberration Theory.....	16
2.2.2 Previous Work on Measuring Lens Aberrations.....	21
2.3 The Mask Error Factor.....	23
Chapter 3	
CD-SEM Based Pattern Recognition for Focus/Exposure Control.....	29
3.1 Introduction.....	29
3.2 Experiment and Results.....	36
3.3 Conclusion.....	39

Chapter 4	
Spatial Variation Characterization in DUV Lithography Using Electrical Metrology.....	42
4.1 Introduction.....	42
4.2 Experiment.....	45
4.2.1 Electrical Linewidth Measurement (ELM) Mask Designs ELM-1 and ELM-2...45	
4.3 Requirements of The Electrical Testing Instrumentation.....	55
4.4 Across-Field and Across-Wafer Spatial CD Distribution.....	57
Chapter 5	
Full Field Lens Aberration Extraction Using Printed Patterns.....	64
5.1 Introduction.....	64
5.2 A Simple Method to Extract Field-Dependent Lens Aberration.....	67
5.3 Sensitivity of Linewidth Variation to Individual Zernike Terms.....	69
5.4 Error Analysis of the Extraction Method.....	84
5.5 Simulation Based Verification of the Extraction Method.....	89
5.6 Improvement Using Scatterometry Metrology.....	92
5.7 Resist Parameter Optimization.....	94
5.7.1 Optimization Using Downhill Simplex Method.....	95
5.8 Discussion on Issues in Solving the Linear System.....	100
5.9 Summary.....	104
Chapter 6	
Compact Formulation of Mask Error Factor for Critical Dimension Control..	109
6.1 Introduction.....	109
6.2 A Compact Theoretical Analysis of Mask Error Factor.....	111
6.3 Experiment and Comparison with Theory and Simulation.....	117
6.4 Conclusion.....	117

Chapter 7	
Conclusions and Future Work.....	120
7.1 Thesis Summary.....	120
7.2 Future Work.....	122
Appendix A	
List of Symbols.....	124
Appendix B-1	
VB Source Code for UV110 Parameter Extraction Using Downhill Simplex Method.....	129
Appendix B-2	
VB Source Code for Extracting Petzval, Primary 90° Astigmatism and Primary Spherical Aberration.....	136

List of Figures

1-1	Chip size increases to accommodate more functionality and higher performance..	3
2-1	Source of variation decomposition in semiconductor manufacturing.....	12
2-2	CD variation components at different scales.....	14
2-3	Illustration of aberrated imaging system.....	16
2-4	Point spread function of 0.2λ coma.....	19
2-5	Point spread function of 0.2λ spherical.....	19
2-6	Point spread function of 0.2λ defocus.....	20
2-7	Point spread function of 0.2λ 90 degree astigmatism.....	20
3-1	Concept of CD-SEM based focus/exposure control in lithography.....	30
3-2	CD-SEM scan for five-line dense pattern.....	31
3-3	CD-SEM scan for five-line dense pattern.....	32
3-4	Neural network configuration (30 input nodes, one hidden layer with 50 nodes and 26 output nodes).....	35
4-1	Cross-sectional SEM image showing footing features.....	43
4-2	Comparison of CD-SEM measurement versus electrical test.....	44
4-3	Floor plan of basic electrical test module.....	46
4-4	Layout of linewidth test structure for dense pattern.....	47
4-5	Layout of basic linewidth test module.....	48
4-6	Sub-block consisting of two basic module frames.....	48
4-7	CD linearity module.....	49
4-8	CD-MEF module layout.....	51

4-9	CD-MEF module layout.....	52
4-10	Cross-sectional SEM lines.....	53
4-11	Floor plan of basic module of ELM-2 mask.....	54
4-12	E-test setup.....	55
4-13	Updating E-test using high resolution source/monitor units.....	56
4-14	Chip image.....	56
4-15	Wafer CD and mask CD measurement for isolated vertical features.....	59
4-16	Bossung curves of resist pattern measured from focus/exposure matrix wafer.....	61
5-1	Reduction of across-field linewidth variability by compensating position-specific lens aberrations.....	66
5-2	Simulated CD sensitivity to Z_1 - Z_6 for isolated and dense lines.....	70
5-3	Simulated CD sensitivity to Z_7 - Z_{12} for isolated and dense lines.....	71
5-4	Simulated CD sensitivity to Z_{13} - Z_{18} for isolated and dense lines.....	72
5-5	Simulated CD sensitivity to Z_{19} - Z_{24} for isolated and dense lines.....	73
5-6	Simulated CD sensitivity to Z_{25} - Z_{30} for isolated and dense lines.....	74
5-7	Simulated CD sensitivity to Z_{31} - Z_{36} for isolated and dense lines.....	75
5-8	Simulated CD sensitivity of isolated line to Z_8 spherical under different defocus	76
5-9	Relation of aberration orientation and pattern orientation (an example of 90° astigmatism).....	80
5-10	Field map of extracted Z_4 primary 90° astigmatism in λ	82
5-11	Field map of extracted $Z_{11}((4\rho^2-3)\rho^2\cos 2\theta)$ in λ	82
5-12	Comparison of aberrated wavefronts across exit pupil for two different field positions.....	83
5-13	CD field map measured using electrical test for chip corresponding to focus – $0.1\mu\text{m}$ and exposure 21mJ on FEM wafer.....	85
5-14	Etching bias and noise measured from uniform wafer.....	86
5-15	Simulation based verification of Zernike extraction method.....	91

5-16	Extracting aberration by measuring profile of resist pattern.....	93
5-17	Resist parameter optimization.....	94
5-18	Geometrical illustration of downhill simplex algorithm.....	97
5-19	Comparison of simulated CD using optimized resist parameters with poly CD under various focus and exposure conditions.....	100
6-1	Closed form aerial image calculation.....	113
6-2	Illustration mask error factor assuming fixed resist threshold.....	114
6-3	Calculated MEF as function of feature size compared with simulation and other work.....	116

List of Tables

2-1	Primary aberration terms.....	18
3-1	Summary of experiment and results for NN-based focus/exposure automatic control.....	36
3-2	Comparison of predicted setting and actual settings, (a) exposure dose, (b) defocus distance.....	38
4-1	Design variables for ELM1 mask.....	46
5-1	CD variation range in nm for individual Zernikes changing within $\pm 0.05\lambda$	77
5-2	Zernike terms of high sensitivity with CD variation range larger than 10nm within $\pm 0.05\lambda$	78
5-3	Downhill Simplex Algorithm.....	99

Acknowledgments

I would first like to express my deepest gratitude to my advisor Professor Costas J. Spanos for his support, guidance and encouragement during my graduate years in Berkeley. I thank him for his consistent involvement, suggestion, enlightenment and help in every detail of my research. Without his guidance, this work would not be possible. I thank him for his gracious understanding and support on many aspects of life beyond research. I especially thank him for his time in reading and correcting the thesis. I would also like to express my gratitude to Professor Andy Neureuther for his helpful insight, invaluable suggestion and comments on my research. I thank him for his detailed guidance for improving the thesis work and his precious expertise in lithography field as well as his professional attitude towards research. Without his suggestion and enlightenment I benefit from his long years lithography research, this work would not be like what it is. I thank Professor Ching-Shui Cheng for his guidance and for spending time in reading and correcting my thesis. I thank Professor Tsu-Jae King for serving as the committee in my Qualification Exam.

I acknowledge Fariborz Nadi and Olga Kobezera for their help in measuring CD-SEM wafers in Cypress Semiconductor. I am grateful to Joe Bendik and Matt Hankinson of National Semiconductor for their support in NIST project. I thank them for their generous help. I thank John Love of TestChip for his help in mask design. I am grateful to Chris Spence of AMD for his help in printing wafers and Richard Lozes of Etec for his help in measuring mask.

I would like to thank Tim Duncan for his assistance in upgrading electrical test setup. I thank my fellow student Jason Cain for his help in experiment. I appreciate his help in technical discussion as well as his kindly offering of reading my thesis.

I thank members of Berkeley Computer-Aided Manufacturing (BCAM) group for their help and friendship. I thank all of you, Junwei Bao, Jason Cain, Runzi Chang, Roawen Chen, Mason Freed, Michiel Krugger, John Musacchio, Anna Ison, Nickhil Jakatdar, Jae-wook Lee, Mian Li, Jeff Lin, Greg Luurtsema, Xinhui Niu, Sundeep Rangan, Nikhil Vaidya, Jiangxin Wang, and Dongwu Zhao. You make all these years of experience in BCAM unforgettable. I thank Thanh-Xuan Nguyen and T K Chen for their support. I would like to thank Mosong Cheng for helpful discussion in many lithography issues.

I would like to thank my parents, Jianjun Zhang and Xi-An Xie, for their unconditional love and support. I thank my sister Weihua Zhang and her husband Long Wang, my sister Weiping Zhang and her husband Zhongfeng Liu for their encouragement.

Finally and most importantly, I declare my deepest debt of gratitude to my wife Yun for her genuine understanding and encouragement. I thank her for enduring all these “hard years” of our life back in China and here in America. My gratitude is beyond any words. Without her love and companion, this thesis would not be possible. I look forward to spending more time with her and our incoming baby. “Thanks”, from my heart! This degree is shared with her.

Chapter 1

Introduction

1.1 Motivation

The rapid growth of the IC industry, which has been seen in past decades and will continue in the 21st century, is driven by demands for improved applications, including data processing and communication with higher speed at a constant or lower cost [1.1]. In 1964 Gordon Moore observed that the market demand for functionality (bits, transistors) doubles every 1.5 to 2 years [1.2][1.3]. As technology nodes have been passed one by one in last 30 years, “Moore’s Law” has been validated as a consistent macro trend of successful leading-edge semiconductor products and companies.

However, as feature size decreases to 90nm nm and beyond for today’s IC technology, the semiconductor industry is facing increasingly difficult challenges. One of the challenges is the overwhelming cost of research and development for each new technology generation. For example, the total cost of a 300 mm wafer manufacturing facility is more than \$3 billion and the development of 193nm lithography is about \$1 billion [1.4].

Therefore, in addition to implementation of next generation technologies, it is highly desirable for the industry to push the existing technology to its full potential. Various resolution enhancement techniques (RET), such as optical proximity correction, phase shifting masks and off-axis illumination, have been commonly implemented to print

smaller features without migration to shorter wavelengths [1.5][1.6][1.7].

The RET approach, however, leads to dramatically smaller process windows [1.8][1.9][1.10] that make process control an increasingly critical issue in maintaining high yield. As the International Technology Roadmap for Semiconductors (ITRS) indicated, the largest tolerable critical dimension (CD) variation is roughly one tenth of the minimum printable linewidth. For example, the error budget for 180 nm technology is only 18 nm. Considering that a wafer will typically go through approximately 400 processing steps before completion, we must identify the critical steps that consume major portions of the error budget for the purpose of efficient process control. Therefore, as the first step to effectively tighten the CD spread, one needs to understand and accurately quantify the structure of the CD error, whether systematic or random, spatial or temporal.

Temporal variability, such as wafer-to-wafer or lot-to-lot, is mainly caused by material variation and process drifts, and it can be reduced by run-to-run or in-situ process control. The manufacturing performance of semiconductor fabrication plants varies significantly depending on whether an appropriate run-to-run or in-situ process control approach is implemented, although there are great similarities in production equipment, manufacturing processes and product produced [1.11]. A successful example [1.12] of run-to-run and in-situ process control shows that 1 nm reduction in CD variation is equivalent to 1 MHz chip speed improvement. This may be translated to an increase of approximately \$7.50 in the selling price of each chip.

In spite of the importance of reducing temporal random process variation, this work proves that it is systematic spatial within-chip variation that is the major component of overall CD variation. The first reason is that chip sizes are becoming larger and larger. For example, today's DRAM chip area is as large as 438 mm² and will continue to

increase by 12% a year, as shown in figure 1-1, in order to accommodate 59% more components per year, required by “Moore’s Law”.

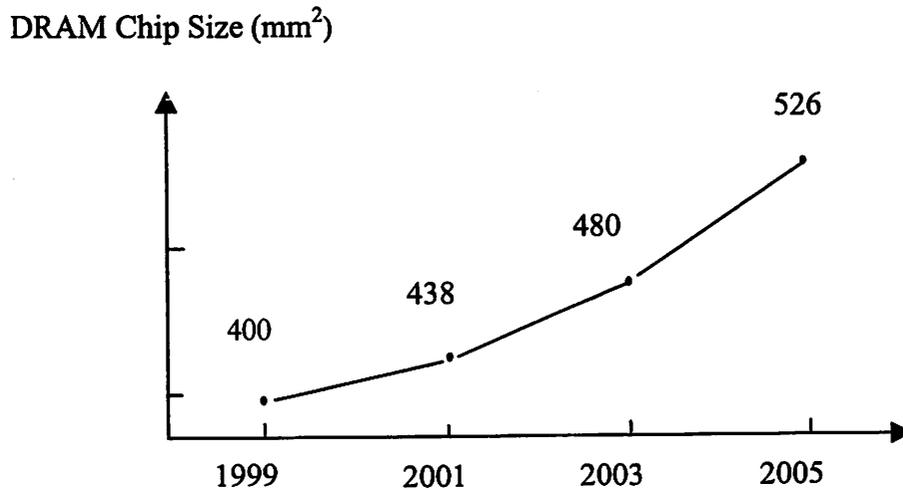


Figure 1-1 Chip size increases to accommodate more functionality and higher performance

Secondly, previous work by Yu *et al* [1.13] shows that CD variation is mostly attributed to the lithography step, rather than other steps. Yu found that there is a significant spatial periodic pattern of CD distribution across wafer for I-line lithography. This suggests that spatial variability is mainly systematic instead of random. Thus process optimization or layout engineering could compensate for this variation, once we are able to decompose the variability into different sources and accurately model them.

1.2 Thesis Goals

The goal of this thesis is to find and classify the physical causes of the major sources of variability in photolithography, by collecting and analyzing a large amount of experimental data. As the first step to effectively tighten the CD spread, one needs to

understand and accurately quantify the structure of the CD error. Through statistical analysis of experimental data, CD variation is decomposed into different components, contrasting random error to systematic error, and temporal variation to spatial variation. Spatial variation is further decomposed into different scales, such as across-field or across-wafer. Once systematic spatial variation is identified and modeled, its underlying physical causes are analyzed. The purpose is to find the important sources of major systematic variation in DUV lithography patterning. In this way we can identify efficient process control schemes and other techniques, such as mask-based error correction, in order to control and minimize the variability.

1.3 Challenge, Approach and Results

Various sources will contribute to final CD variations. As will be explained in the following chapters, different categories of error component are caused by different error sources. For example, illumination uniformity and flare may cause systematic CD variation. But these are assumed to be generally small when comparing with lens and reticle error. Their contributions to CD error are assumed to be insignificant and are lumped into the random component. Field-to-field exposure dose change and leveling error cause die-to-die variation and are generally random. Non-uniformity in film deposition, resist coating, development and etching lead to across wafer variation. All these factors, combined with metrology errors, make the CD variation analysis a big challenge. Therefore it is necessary to repeatedly measure linewidth with high-density spatial coverage to extract systematic CD variation within the field in order to relate it to physical causes and accurately model it. CD-SEM is not adequate because it is relatively slow when collecting the large amounts of data which is necessary for this study. Meanwhile, the reliability of CD-SEM is also in doubt because the edge roughness of the pattern degrades the repeatability of the measurement. On the other hand, electrical

linewidth measurement (ELM) has the advantage of high speed, is independent of pattern profile and can offer sub-nanometer precision. ELM also has some limitations, namely that it includes the confounding variability of the plasma etch process, and that the electrical profile is not the same as the physical profile.

From the study presented here, it is found that a major component of CD variation appears within the exposure field. Across wafer variation is at the order of 2.7nm; while across field variation is about 5.9 nm for a standard 0.22 μ m process. Experimental results indicate that the systematic across-field linewidth variation (AFLV) mainly comes from lens aberrations, such as astigmatism, spherical, petzval, etc [1.17]. Reticle error is an additional important source leading to systematic CD variation.

A new method has been proposed to quantitatively extract lens aberrations, which is identified to be one of the most important sources of spatial variation. Magnification of reticle error, as a second important variation source, has been formulated in a straightforward manner. The thesis aims to lay groundwork for systematically compensating error sources in order to increase manufacturing yield and product performance. Specifically, the newly proposed full-field lens aberration measurement technique is potentially useful for optical proximity correction (OPC) technology. If local lens aberration information can be incorporated into this widely used resolution enhancement treatment, it may improve OPC performance. This method of lens aberration extraction is easy to implement without requiring additional processing or metrology hardware. It is also helpful to evaluate the quality of lithography equipment. Analytical formulation of mask error factor may also be helpful in finding the interaction of local lens aberration with OPC patterns. It will also provide an insight into the mechanism of mask error magnification and its relation with process settings.

This study is crucial in determining the levels of control needed and justified for 248 and 193nm technologies. The principles discovered might also be of use in 157nm and EUV technologies as well. It could also lead to off-line techniques for reticle/stepper matching and mask engineering in order to reduce the variability in DUV lithography patterning. Position-specific nonlinearity in pattern transfer for different features can also be obtained. This information can then be used to systematically pre-modify mask design in order to compensate the CD variation due to lens aberration signature.

1.4 Summary

In summary, the objectives of causal analysis of lithography variability are three-fold: major sources of variation need to be found; the underlying physical causes of these major sources should be accurately characterized and modeled; the knowledge in modeling the physical causes can be used in effectively reducing lithography variations.

1.5 Thesis Organization

The thesis is organized as follows. Chapter 1 gives the introduction, motivation and overview of the thesis project. In chapter 2, the background of lithography variation is explained. As the most critical step in semiconductor manufacturing, lithography presents a lot of challenges in efficient process control. Different variation factors are categorized and discussed. The basics of lens aberration theory are reviewed in this chapter, followed by a brief discussion of mask error factor.

Focus and exposure dose are the basic controllable process settings in photolithography. Fluctuations in these two process parameters are the important source in die-to-die spatial variation as well as lot-to-lot temporal variations. The calibration of focus and

exposure dose setting is a common requirement in lithography. A CD-SEM based focus/exposure control method is introduced in Chapter 3. This method utilizes statistical data reduction to extract characteristic feature in digitized CD-SEM trace. Then feed-forward neural network is used to classify the traces based on the focus-exposure conditions of the samples.

Chapter 4 introduces the full-scale experiment for characterizing spatial linewidth variation in lithography. The chapter starts with an experiment overview and consideration of the metrology issue. The mask design for electrical linewidth measurement (ELM) is explained. Wafer processing and electrical measurement setup are then introduced. The measurement results have shown two indications. First, major variation comes from within exposure field instead of across wafer. Second, the variation is mostly systematic rather than random in nature. This observation suggests that lens aberrations and mask errors are two important variation sources.

Chapter 5 introduces a technique to extract full-field lens aberration using printed patterns in chapter 4. The basic concept of this method is that the dependency of linewidth on individual aberration terms can be approximated by linearization around zero. This approximation is valid, since the aberration of a state of the art lithography system is typically small. The linearization can be conducted under different process conditions. Therefore the aberrations can be deduced by means of numerical analysis. The accuracy of this method depends on the quality of the resist model. In this chapter, several numerical algorithms (Simplex, Powell and Simulated Annealing) are used to extract resist parameters. The performance and other attributes of these algorithms are studied and compared. The reliability of this aberration extraction method is also dependent on the choice of the combination of aberration terms, as well as the process conditions. Some important aberration terms may cause image distortion or shifting, but have relatively small impact on linewidth. This is one of the limitations of this method.

At the end of this chapter, the criteria for selecting aberration terms and the impact on the accuracy of the results are discussed.

Mask errors are one of the most important sources of systematic spatial lithography variation. Chapter 6 introduces a method of formulating the mask error factor in a straightforward way. First, the aerial image of a one-dimensional pattern is derived under coherent illumination. Based on the closed form aerial image calculation, the analytical expression of mask error factor is then derived. The results are compared to first principle simulation and experiments. The formulation of the mask error factor is helpful in theoretically understanding its physical cause and its relation with process setting, so that the variation caused by mask errors can be controlled by selecting the optimum process window. This knowledge is also useful in understanding the interaction of mask errors and local lens aberrations.

Chapter 7 concludes the thesis with summary, discussion and future work. A promising improvement of extracting lens aberration is to use the information of profile of printed pattern instead of linewidth only. This approach is described in this chapter as future work. Since one-dimensional profile measurement is available [1.14-1.16], the remaining issue of this method is still precise resist modeling. Compared to the method introduced in chapter 5, this improved method may be able to extract more aberration terms with higher accuracy.

References for Chapter 1

- [1.1] "International Technology Roadmap for Semiconductors", 2001 Edition.
- [1.2] Moore, G.E., "Intel-Memories and the Microprocessor" *Daedalus* 125, No.2, 1964.
- [1.3] Gordon E. Moore. "Lithography and the Future of Moore's Law". *Microlithography, Proceedings of the SPIE*, vol. 2440. pp.2-17. 1995.
- [1.4] Huff, H.R., Goodall, R.K., Nilson, R.H., and Griffiths, S.K., "Thermal Processing Issues for 300 mm Silicon wafers: Challenges and Opportunities," *Symposium Proceedings of ULSI*, 1997.
- [1.5] M. D. Levenson, et al, "Improving Resolution in Photolithography with a Phase-shifting Mask", *IEEE, Trans. Electron Devices*, Vol. ED-29, No.12, pp.1828-1836, 1982.
- [1.6] T. Terasawa. "Sub-wavelength Optical Lithography". *Proceedings of the SPIE - The International Society for Optical Engineering*, vol.4181, (Challenges in Process Integration and Device Technology, Santa Clara, CA, USA, 18-19 Sept. 2000.) *SPIE-Int. Soc. Opt. Eng.* 2000. p.8-16.
- [1.7] T. Terasawa. "Subwavelength Lithography (PSM, OPC)". *Proceedings ASP-DAC 2000. Asia and South Pacific Design Automation Conference 2000 with EDA TechnoFair 2000. IEEE*, 2000. p.295-300.
- [1.8] M.V. Plat, et al, "Impact of Optical Enhancement Techniques on the Mask Error Enhancement Function", *Optical Microlithography XIII*, 4000, SPIE, 2000.
- [1.9] Akio Misaka and Shinji Odanaka, "Optical Proximity Correction Considering Process Latitude", *Optical Microlithography XII*, SPIE vol. 3679, 1999.
- [1.10] R. Schenker. "Effects of Phase Shift Masks on Across Field Linewidth Control". *Microlithography, Proceedings of the SPIE*. Vol. 3679. pp. 18-26. 1999.
- [1.11] R. Leachman and D. Hodges, "Benchmarking Semiconductor Manufacturing," *IEEE Trans. On Semiconductor Manufacturing*, vol. 9, no.2, pp. 158-167, May 1996.
- [1.12] J. Sturtevant, et al, "CD Control Challenges for Sub-0.25 μm Patterning", *SEMATECH DUV Lithography Workshop*, Austin, TX, Oct., 1996.

- [1.13] C. Yu and C. J. Spanos, "Patterning Tool Characterization by Causal Variability Decomposition", IEEE Trans. on Semiconductor Manufacturing, vol. 9, no.4, pp527-535, Nov. 1996.
- [1.14] X. Niu, N. Jakatdar, and C. J. Spanos. "Deep Ultraviolet Lithography Simulator Tuning by Resist Profile Matching". International Symposium on Microelectronic Manufacturing Technologies". pp. 245-252, 1999.
- [1.15] X. Niu, N. Jakatdar, J. Bao, and C. J. Spanos. "Specular Spectroscopic Scatterometry in DUV Lithography". Microlithography, Proceedings of the SPIE, pp. 159-168, 1999
- [1.16] Xinhui Nu. "An Integrated System of Optical Metrology for Deep Sub-Micron Lithography". PhD thesis, University of California, Berkeley, 1999.
- [1.17] H. Liu, C. Yu and B. Gleason, "Contributions of stepper lenses to systematic CD errors within exposure fields" SPIE Vol. 2440, pp868-877

Chapter 2

Background

The objective of this chapter is to review the background of lithography variation. As the most critical step in semiconductor manufacturing, lithography presents a lot of challenges in efficient process control. Different variation factors are categorized and discussed. Since lens aberration and mask error magnification are two important variation sources in lithography, the background of these two issues are introduced in this chapter. The basics of lens aberration theory are reviewed in this chapter followed by an introduction of previous works on measuring lens aberrations. Finally a brief discussion of mask error factor is given.

2.1 Sources of Variation in Optical Lithography

Lithography is perhaps the most important processing step in determining IC performance. Because of economic factors as well as advances in optics and resist materials, optical lithography in the near future will remain more practical than other candidates, such as electron beam, extreme ultraviolet (EUV), and X-ray lithography [2.1]. Several resolution enhancement techniques, including phase shift masks, off-axis illumination, pupil filtering, and optical proximity correction, have extended the resolution limits of optical lithography [2.18-2.28].

There are many factors that contribute to the final variation of the printed pattern. First of all, there may be material imperfections in the resist and in the wafer itself. These factors include non-uniformity in resist coating and wafer substrate reflectivity and topography. Secondly, variations also come from errors due to equipment and instrument-related effects, including the reticle and the exposure tool. The quality of the mask is critical since any deviation of mask CD from the designed value will be magnified by a factor called mask error factor (MEF) for features with a k_1 factor less than 0.8. Lens aberrations and focus/dose fluctuations of lithography tools have a strong influence on CD uniformity. Thirdly, process control in PEB, development and etching will also affect the final CD variability. The relevant process parameters include PEB temperature, time, plasma density, gas flow rates etc. Finally, disturbances in the manufacturing environment such as amines and humidity are also a source of CD variation in photolithography.

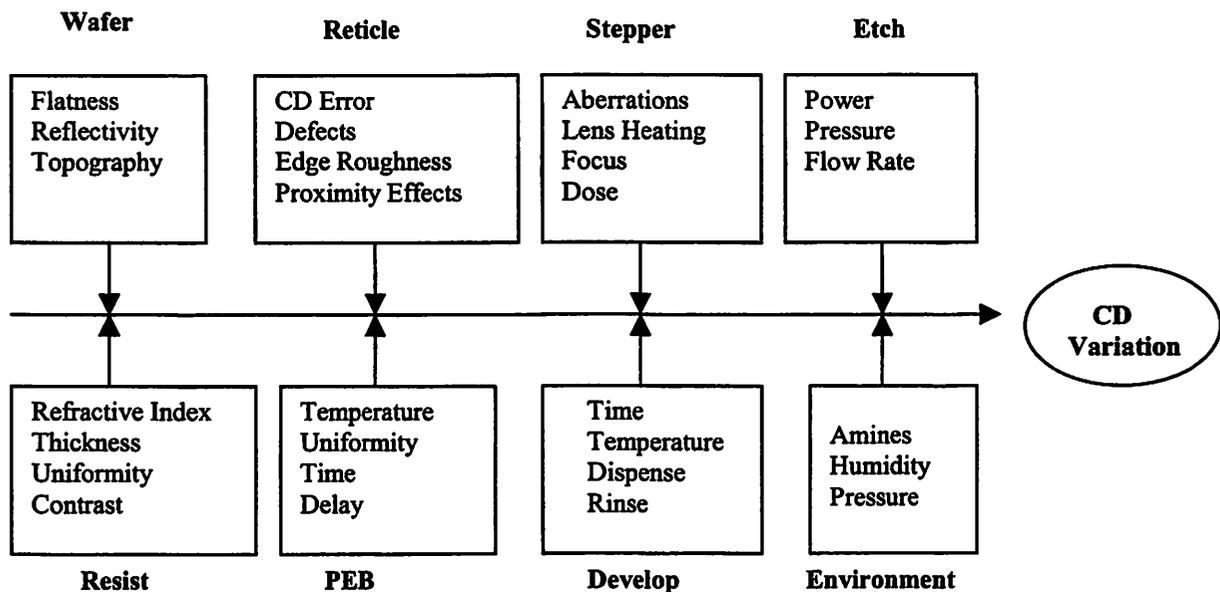


Figure 2-1 Source of variation decomposition in semiconductor manufacturing

An effective CD control approach requires a “divide and conquer” strategy, which means it is necessary to decompose the overall CD variation into individual physical causes and find the most important error sources. Then relevant control or compensation techniques can be implemented.

Variation occurs across time and across space. Temporal variation, including wafer-to-wafer and lot-to-lot, is mainly due to equipment drifting and long-term material variability. It is reduced by real time, in situ process monitoring, run-to-run or statistical process control. Spatial variation manifests itself in different scales, such as intra-field, die-to-die, and across wafer. Most spatial variation is systematic in nature. For example, mask errors and lens aberrations cause a relatively static pattern of CD distribution across the field. As will be demonstrated in chapter 4, systematic spatial variation is the most important error component for 248 nm lithography and it is the focus of this work. Die-to-die spatial variation is mainly attributed to random fluctuations of exposure dose and defocus distance during the step-and-scan process. Chapter 3 will introduce an efficient focus/exposure control method using pattern recognition techniques based on SEM traces. Across wafer variation is generally caused by equipment non-uniformity and various physical effects such as thermal gradients in film deposition, resist coating, development and etching. Pattern dependent variation occurs mainly within the field. Among this category are optical diffraction, process proximity effect, and micro-loading in development as well as in etching.

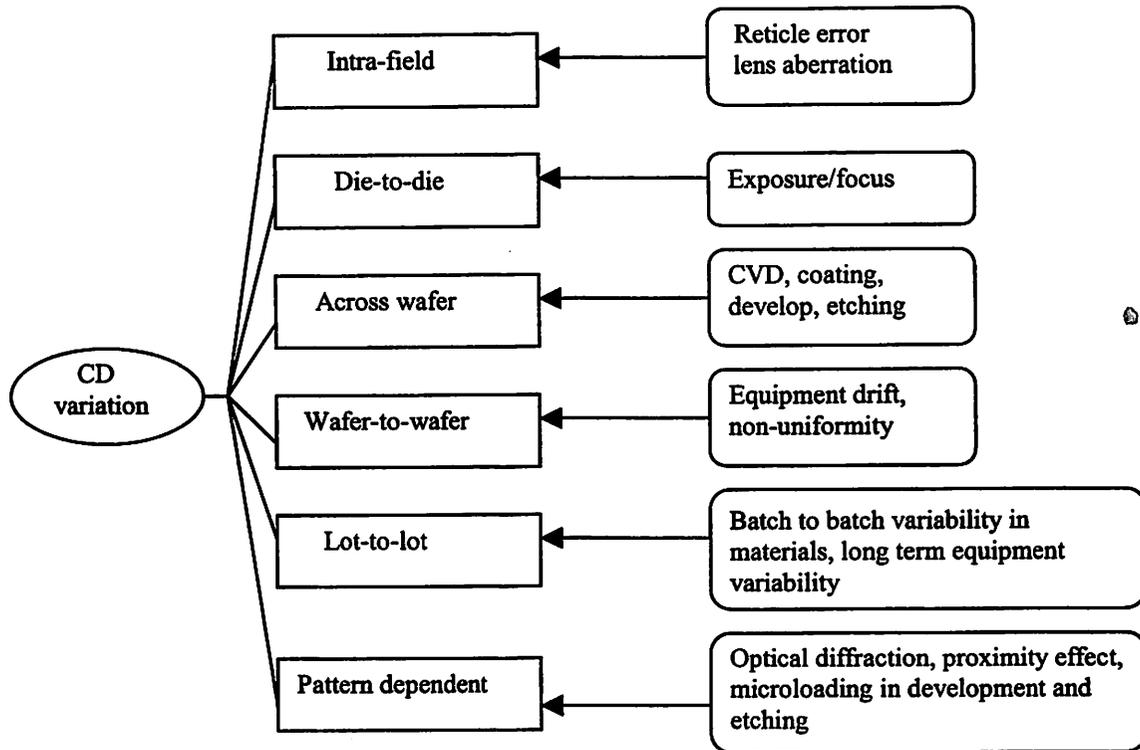


Figure 2-2 CD variation components at different scales

This work attempts to identify the major spatial systematic variation in optical lithography. Lens aberrations and mask errors are of special interest in this analysis.

Other error sources may also cause systematic variation, but they are not the subjects of this study. In I-line lithography for patterning 0.35 μm features, it is found that the lack of uniformity of local partial coherence is possibly a critical source leading to systematic across field linewidth variation (AFLV). Borodovsky [2.43] published a research to study the effect of this local illumination variation on exposure tool performance. He observed a certain type of I-line stepper has significantly larger AFLV than another type of I-line stepper. The only difference between these two machine types is the nominal partial coherence, while all other settings are the same, including numerical aperture,

best focus and exposure, as well as photoresist. He also found from simulation that partial coherence would change the characteristics of linewidth vs. defocus dramatically. These observations suggested that illumination may influence AFLV, because he observed that other systematic causes, such as reticle error and aberration residual failed to explain the larger AFLV for this particular type of stepper. Since proximity effects are different under different partial coherence, he proposed a curve fitting method to measure local effective partial coherence. The linewidths at various pitches are measured at different field positions. Then the experimental proximity effect curve is compared with simulated curves under varying partial coherence, and the partial coherence value at this position is obtained. It is found that the equipment with poorer AFLV does have large variation of partial coherence across the field. It is also found that vertical and horizontal features have different partial coherence even at the same field location. It is explained by the author that this is mainly due to the lack of rotational symmetry of the extended light source. In other words, the asymmetry of the light source will cause the difference in partial coherence for features with different orientations. However, since aberrations may compound the proximity effect curve under different partial coherence, the difference in extracted partial coherence for vertical and horizontal is possibly due to astigmatism to some extent.

AFLV has been a concern in terms of process control for quite a long time. This issue becomes more of interest since it is known that systematic CD error is a significant fraction of the total linewidth error. Liu [2.44] et al reported a study of CD distribution as function of field coordinates. It is found both reticles and steppers contribute to the systematic AFLV for three different I-line exposure tools. After removing reticle errors, the systematic linewidth variation pattern suggests a correlation to aberration residuals. Isolated lines exhibited more AFLV than dense lines.

2.2 Lens Aberration Background

Today's¹ lithography systems are capable of resolving $0.07\ \mu\text{m}$ features using a wavelength of $193\ \text{nm}$. Because of advances in lens design, inspection and manufacturing and progress in mounting technique and precision engineering, the optical image is improved so much that it is often said that lithography systems are "diffraction-limited," which means image quality is not affected by lens aberrations. However, this will never be true due to unavoidable defects in optical design and manufacturing. On the contrary, aberrations have become increasingly important in limiting the performance of modern lithography systems [2.33][2.34][2.35].

2.2.1 A Brief Review of Lens Aberration Theory [2.42]

Consider an optical system with an off-axis point source P_o as shown in the figure 2-3. Two rays from object P_o on the mask plane to the image point at the wafer plane are shown. One of these two rays is called the chief ray (CR), which is defined as the ray passing through the center of the exit pupil. The second ray is an arbitrary ray (AR) from P_o to the image plane.

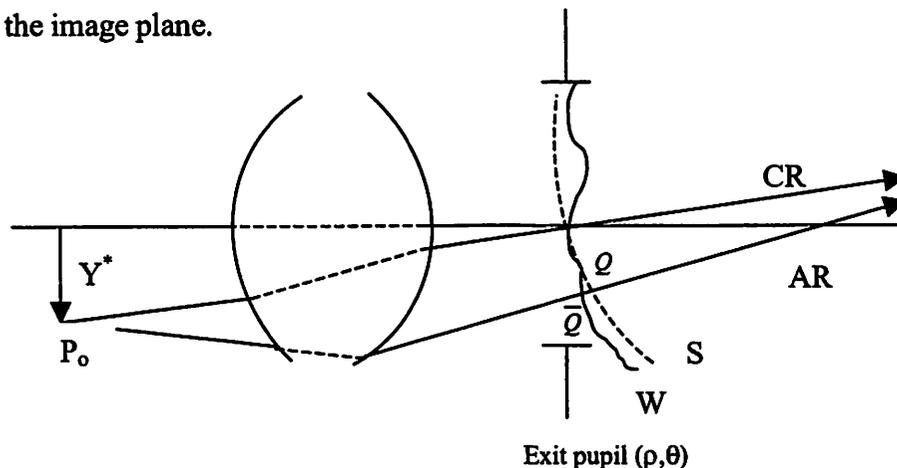


Figure 2-3 Illustration of aberrated imaging system

¹ August 2002.

For an aberration-free system, the optical path of CR and all other rays are equal. In reality, lens aberration occurs when there is difference in the optical path among different rays. S represents the wavefront of a Gaussian reference sphere² corresponding to CR. The deformation of the wavefront in the region of the exit pupil is represented as W . The aberration function Φ is defined as the optical path difference (OPD) of AR and CR. If \bar{Q} and Q are the points in which AR intersects with W and S , respectively, the aberration function is measured as $\bar{Q}Q$ times the local refraction index. Since each different AR passes through a different part of W , the aberration function is a function of the exit pupil coordinates (ρ, θ) . Moreover, different image field positions will also correspond to a different aberrated wavefront. Thus, aberration is also a function of field position. If the optical system is rotationally symmetric around the main axis, the aberration function is only dependent on the image height Y^* .

Due to the property of orthonormality³, and because of the convenience of balancing aberrations with lower order terms to minimize the net aberration variance, the aberration function is usually expanded as a Zernike circle polynomial power series.

$$\Phi(Y^*, \rho, \theta) = \sum_l \sum_n \sum_m a_{lmn} Y^{*2l+m} R_n^m(\rho) \cos m\theta \quad (2-1)$$

Here l , n and m are nonnegative integers, where $n \geq m$, and $n-m$ is even. In this equation, a_{lmn} is the Zernike coefficient and R is the radial Zernike polynomial.

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! [(n+m)/2 - s]! [(n-m)/2 - s]!} \rho^{n-2s} \quad (2-2)$$

The orthonormality of the Zernike polynomials is expressed as

² The sphere surface with the center at Gaussian image in geometric optics.

³ Zernike circle polynomials are orthonormal over a unit circular pupil as defined in eqn. 2-3.

$$\int_0^1 \int_0^{2\pi} R_n^m(\rho) \cos m\theta R_n^{m'}(\rho) \cos m'\theta \rho d\rho d\theta / \int_0^1 \int_0^{2\pi} \rho d\rho d\theta = \delta_{nn'} \delta_{mm'} \quad (2-3)$$

where δ is the Kronecker delta function.

The order of an aberration term is equal to the sum of the powers of Y^* and ρ . The fourth-order aberrations are called the primary aberrations and they are represented in table 2-1, in which $\sin(m\theta)$ terms are also included, considering general conditions without rotational symmetry.

Table 2-1 Primary aberration terms

Symbol	Name	l n m	Term
Z1 Z2	Distortion	1 1 1	$2\rho\sin\theta$ $2\rho\cos\theta$
Z3	Petzval	1 2 0	$v3(2\rho^2-1)$
Z4 Z5	Astigmatism	0 2 2	$v6\rho^2\sin 2\theta$ $v6\rho^2\cos 2\theta$
Z6 Z7	Coma	0 3 1	$v8(3\rho^3-2\rho)\sin\theta$ $v8(3\rho^3-2\rho)\cos\theta$
Z11	Spherical	0 4 0	$v5(6\rho^4-6\rho^2+1)$

Figure 2-4 to figure 2-7 illustrate the point spread functions with 0.2λ individual primary aberrations, simulated using Prolith [2.36][2.37], with X Y coordinates corresponding to image plane and Z to image intensity. Prolith is a lithography simulation software

package. It simulates aerial image formation based on scalar and vector diffraction theory. It also simulates basic lithographic steps of resist exposure, PEB diffusion and development. Similar simulation results can be obtained using Splat [2.38-2.41].

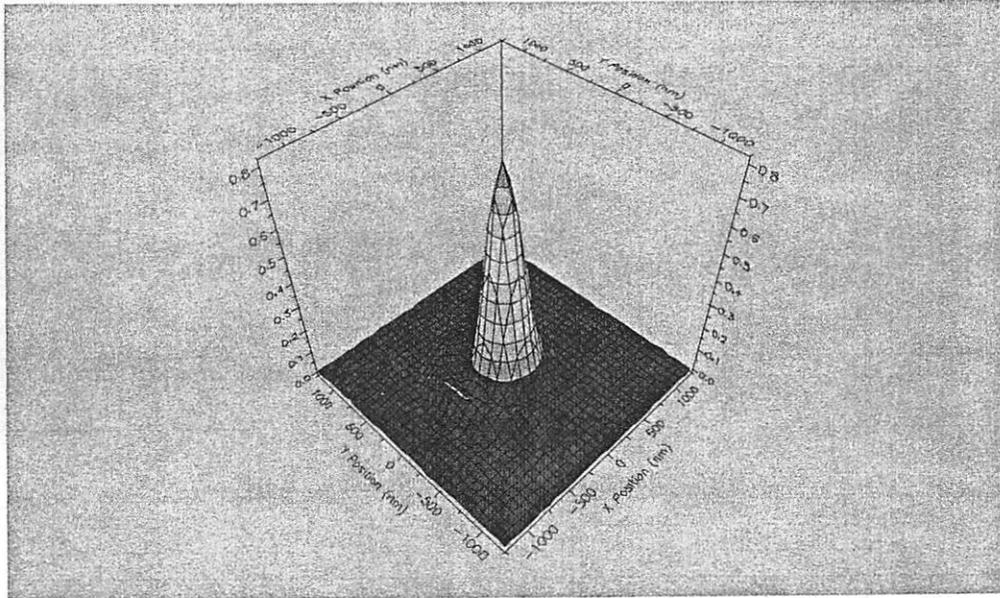


Figure 2-4 Point spread function of 0.2λ coma

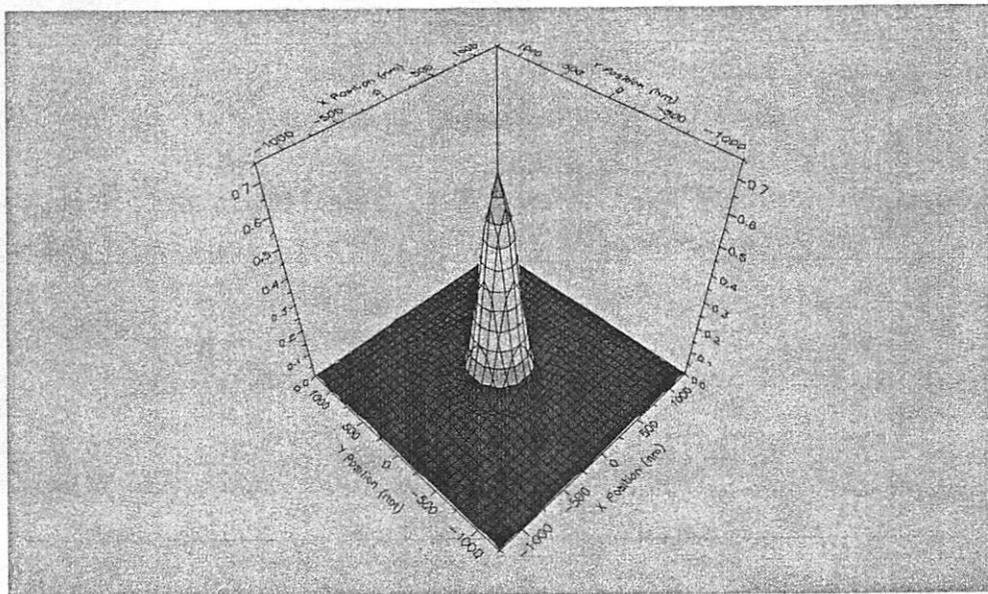


Figure 2-5 Point spread function of 0.2λ spherical

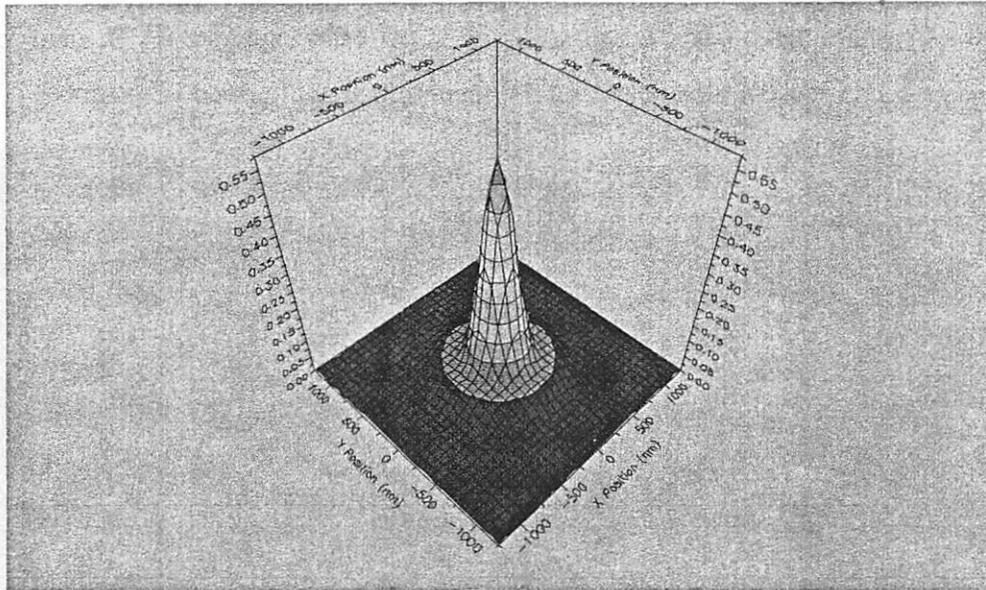


Figure 2-6 Point spread function of 0.2λ defocus

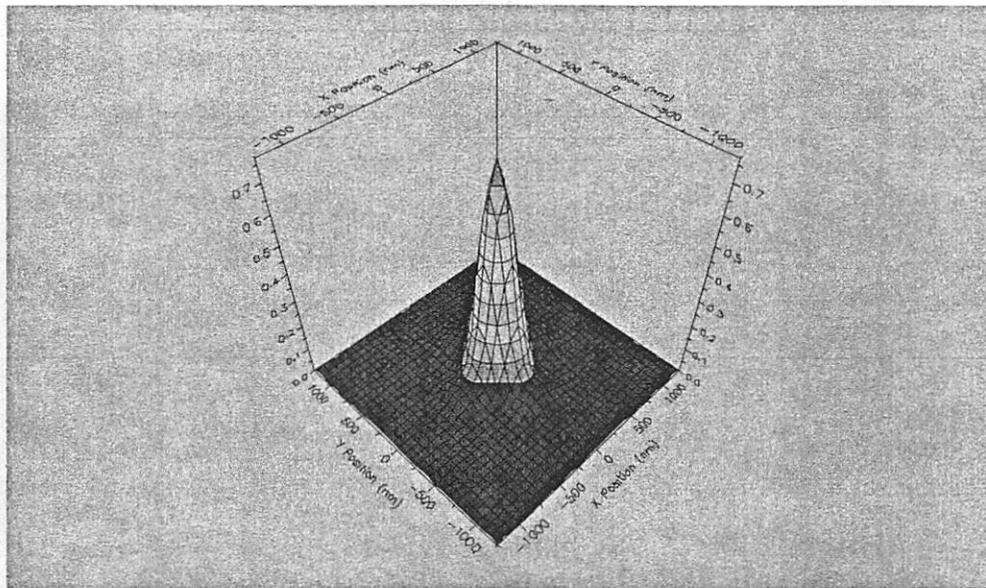


Figure 2-7 Point spread function of 0.2λ 90 degree astigmatism

A commonly used index of quantitatively characterizing the lens quality is the Strehl ratio, which is defined as the ratio of the peak image irradiance with aberrations to that of the image irradiance without aberration. It is expressed as,

$$Sr = \pi^{-2} \left| \int_0^1 \int_0^{2\pi} \exp[i\Phi(\rho, \theta)] \rho d\rho d\theta \right|^2 \quad (2-4)$$

When the amount of aberration is small, the Strehl ratio can be approximated [2.2] as

$$Sr \approx \exp(-RMS^2) \quad (2-5)$$

$$\text{where } RMS = \sqrt{\langle [\Phi(\rho, \theta) - \langle \Phi(\rho, \theta) \rangle]^2 \rangle} \quad (2-6)$$

It can be seen that the Strehl ratio is maximized when RMS is at a minimum. This is achieved by mixing the primary aberration with lower order terms to reduce the net RMS in lens design. State-of-the-art optical lithography lens elements have an average Strehl ratio of above 0.95 across entire image field [2.3].

2.2.2 Previous Work on Measuring Lens Aberrations

As discussed earlier, lens aberrations are becoming an increasingly important concern in extending the limits of optical lithography. Although generally the lens maker can measure the wavefront of the lens with a through-the-lens-interferometer (TTLI), this is not available once the lenses are mounted to the system. Also, the lens aberration of the completed system is usually not the same as when it is last measured by TTLI. So the aberration signature of the whole system is generally unknown to the end user.

If aberration cannot be measured directly, it may be possible to infer their magnitude from the printing behavior of special test patterns. To this end, a number of resist pattern based techniques have been investigated.

Neureuther [2.10-2.14] proposed an effective method of measuring lens aberrations by designing phase shifting patterns that have enhanced sensitivity to individual aberration factors. Sensitivity levels of 0.01λ and good orthogonality with respect to different individual Zernike terms have been shown. Based on an elegant theory utilizing the orthogonality property of Zernike terms, Neureuther et al designed various probe and target to detect individual aberration orders. The shape of target is the inverse Fourier transform of Zernikes. 180° phase shift pattern is used to measure odd term. Since the inverse Fourier transform of most Zernikes is rotational symmetric, it is a challenge to fabricate small circle pattern instead of Manhattan pattern on mask. The levels of individual aberration are determined by measuring the intensity of the various targets using optical detector or inspecting the printed resist images.

Kirk reported a technique for measuring astigmatism [2.4][2.5]. Astigmatism is determined by measuring the focal positions of printed photoresist lines oriented at four different angles, 0° , 45° , 90° , and 135° . The author claims that a precision of standard deviation of 7nm of astigmatism is obtained using this technique. Dirksen [2.6] [2.7] presented a method for evaluating lens aberrations by examining the SEM picture of a $1/2\lambda$ circular phase object. The deformation of a printed resist ring is related to the amount of aberrations. Besides the measurement of focal offsets of lines with various orientations similar to Kirk's work, Yeung [2.8] measured the positional shifts of a set of printed photoresist gratings with different periods and angles under almost coherent illumination. Nomura *et al* [2.9][2.17] modified the approach by three-beam interference, that is, imaging by only three diffracted beams. Litel [2.15][2.16]

demonstrated a method by which the aberrated wavefront is reconstructed by measuring local pinhole displacement with limited sampling aperture.

Although these techniques have shown promising ways to measure the lens aberrations, some ambiguity still exists in interpreting the relationship between resist image and individual Zernike terms. Either approximation assumptions or inherent confounding phenomena may affect the accuracy of measured Zernike coefficients. Moreover, some techniques are only applicable to a few low order aberration terms, which limits their utility when higher orders need to be taken into account. The new aberration measurement technique proposed in this work is theoretically valid for higher order aberrations. This issue will be discussed in later chapters.

2.3 The Mask Error Factor

The second most important source of within-field systematic spatial variation is mask errors. It is a well-known phenomenon that lithography systems tend to enlarge mask errors when transferring the mask pattern to the wafer. This effect becomes significant when the k_1 factor is less than 0.8. The mask error factor is defined as the ratio of wafer CD deviation over de-magnified mask CD deviation.

$$MEF = \frac{\Delta CD_{wafer}}{\Delta CD_{mask}} \quad (2-7)$$

The larger than unity MEF is caused by degradation of the image integrity when the photomask gradually loses control of the image shape as feature size decreases. Nonlinearity in both the imaging system and resist response exacerbates CD error on the mask. This makes CD control a more difficult challenge.

Many theoretical and experimental works study the effects of the mask error factor as well as its relation to processing conditions [2.37-2.39]. Although this knowledge is helpful in understanding the implication of the mask error factor on process control, a rigorous formulation of the mask error factor are still needed. It is very useful to quantitatively determine the dependency of the mask error factor on process conditions in order to effectively control it. Furthermore, such a formulation is potentially useful in improving the efficiency and accuracy of optical proximity correction rule generation. Finally, since lens aberrations and mask errors exist simultaneously, the actual linewidth variation is the result of the interaction of these two effects. Therefore, theoretical understanding of the mask error factor is necessary to determine the MEF in presence of local lens aberrations. This issue will be addressed in detail in a later chapter of the thesis. Our interest not only lies in rigorous formulating the mask error factor, but also aims to study the importance of this important process phenomenon in term of causing spatial process variation and limiting the process latitude.

References for Chapter 2

- [2.1] J. H. Bruning, "Optical lithography: Thirty years and three orders of magnitude". *Microlithography, Proceedings of the SPIE*, vol. 3049. pp.14-27. 1997.
- [2.2] V. Mahajan, "Strehl ratio for primary aberrations in terms of their aberration variance" *Journal of the Optical Society of America, A* 73, pp860-861, 1983.
- [2.3] D. Flagello, Jos de Klerk, Guy Davies, Rich Rogoff, "Towards a comprehensive control of full-field image quality in optical photolithography", *Optical Microlithography*, vol.3051, SPIE, pp672-685.
- [2.4] J. P. Kirk, "Astigmatism and field curvature from pin-bars", *Optical/Laser Microlithography IV*, SPIE 1463, pp282-291, 1991.
- [2.5] J. P. Kirk, "Measurement of astigmatism in microlithography lenses", *Optical Microlithography XI*, SPIE 3334, pp 848-854, 1998
- [2.6] P. Dirksen, C. Juffermans, R. Pellens, M. Maenhoudt, and P. De Bisschop, "Novel aberration monitor for optical lithography", *Optical Microlithography XII*, SPIE 3679, pp77-86, 1999.
- [2.7] P. Dirksen, C. Juffermans, A. Engelen, P.De Bisschop, and H. Muellerke, "Impact of high order aberrations on the performance of the aberration monitor," *Optical Microlithography XIII SPIE 4000*, pp9-17. 2000.
- [2.8] Michael S. Yeung, "Measurement of wave-front aberrations in high-resolution optical lithographic systems from printed photoresist patterns", *IEEE trans. On semiconductor manufacturing*, vol. 13, no. 1, pp24-33, February 2000.
- [2.9] H. Nomura and T. Sato, "Techniques for measuring aberrations in lenses used in photolithography with printed patterns", *Applied optics*, vol. 38, no. 13, pp2800-07, May 1999.
- [2.10] A. Neureuther, S. Hotta and K. Adam, "Modeling defect-feature interactions in the presence of aberrations".
- [2.11] G. Robins, K. Adam and A. Neureuther, "Measuring optical image aberrations with pattern and probe based targets". *J. Vac. Sci. Tech. B* 20(1), Jan/Feb 2002, pp338-343.

- [2.12] A. Neureuther, P. Flanner III, and S. Shen, J. Vac. Sci. Technol. B.5, 308 (1988).
- [2.13] V. Mastromacro, K. H. Toh, and A. Neureuther, J. Vac. Sci. Technol. B 6, 224 (1988).
- [2.14] A. Neureuther, Proc. SPIE 1496, 80 (1990).
- [2.15] N. Farrar, A. Smith, D. Busath, D. Taitano, "In-situ measurement of lens aberrations", Optical Microlithography XIII SPIE 4000, pp18-29. 2000.
- [2.16] www.litel.net
- [2.17] H. Nomura, K. Tawarayama, and T. Kohno, "Higher order aberration measurement with printed patterns under extremely reduced σ illumination", Optical Microlithography XII SPIE 3679, pp358-367. 1999.
- [2.18] M.D. Prouty and A. R. Neureuther. "Optical imaging with phase shift masks". Optical Microlithography III, Proceedings of SPIE, vol. 470, pp.228-232, 1984.
- [2.19] M. Levinson, et al. "Improving resolution in photolithography with a phase shifting mask". IEEE, ED-29, p.1828, 1982.
- [2.20] B. Lin. "The attenuated phase-shifting mask". Solid State Technology, vol.35, pp.43-47, Jan.1992.
- [2.21] K. Derbyshire. "Issues in advanced lithography". Solid State Technology, p133, May, 1997.
- [2.22] H. Fukuda, A. Imai, T. Terasawa, and S. Okazaki, "New approach to resolution limit and advanced image formation techniques in optical lithography", IEEE, Tran. On Electron Devices, Vol. ED-38, No.1, pp67-75, 1991.
- [2.23] P.Ackmann, et al. "Phase shifting and optical proximity corrections to improve CD control", Microlithography, Proceedings of the SPIE, vol. 3051. pp.146-153. 1997.
- [2.24] Y. Trouiller, et al. "0.18 μ m optical lithography performances using an alternating DUV phase shift mask". Microlithography, Proceedings of the SPIE, vol.3334, pp. 25-35, 1998.

- [2.25] H. Fukuda, et al. "New approach to resolution limit and advanced image formation techniques in optical lithography". IEEE trans. Electron device, vol. Ed-38, No.1 pp.67-75, 1991.
- [2.26] O. Otto, "Automated optical proximity correction: A rule-based approach". Microlithography, Proceedings of the SPIE, vol.2197, pp.278-285, 1994.
- [2.27] C. Park, "Automatic optical proximity correction using aerial image simulation and Newton optimization scheme". Symposium on Semiconductor and Integrated Circuits Technology, Vol.50, pp. 44-52, 1996.
- [2.28] N. Cobb and A. Zakhor, "Experimental results on optical proximity correction with variable threshold resist model". Microlithography, Proceedings of the SPIE, vol. 3051, pp.458-468, 1997.
- [2.29] J. Chen, et al. "Optical proximity correction for intermediate-pitch features using sub-resolution scattering bars". J. Vacuum Science and Technology, vol. 15, pp.2426-2433, Nov. 1997.
- [2.30] A. Wong, R. Ferguson, L. Liebmann, S. Mansfield, A. Molless, and M. Neisser. "Lithographic effects of mask critical dimension error". Microlithography, Proceedings of the SPIE, vol.3334, pp.106-116, 1998.
- [2.31] P.Yan and J. Langston. "Mask CD control requirement at 0.18 μm design rules for 193 nm lithography". Microlithography, Proceedings of the SPIE, vol. 3051, pp. 164-169, 1997.
- [2.32] J. Van Schoot, J. Finders, K. van Ingen Schnau, M. Klaassen, and C. Bujik. "The mask error factor: Causes and implications for process latitude". Microlithography, Proceedings of the SPIE, vol. 3679, pp.250-260, 1999.
- [2.33] A. M. Davis, A. E. Bair, B. D. Lantz, J. R. Johnson, and C. R. Spinner III. "The effect of lens aberrations as a function of illumination condition on full field process windows". Microlithography, Proceedings of the SPIE, vol. 3334, pp.839-847, 1998.
- [2.34] T.A. Brunner. "Impact of lens aberrations on optical lithography", Proceedings of Olin Microlithography Seminar, Interface 1996, pp-1-28 (1996).
- [2.35] M. V. Dusa. "Effects of lens aberrations on critical dimension control in optical lithography". Microlithography, Proceedings of the SPIE, vol. 3804, 1997.

- [2.36] C. A. Mack. "PROLITH: A comprehensive optical lithography model". Proceedings of SPIE, vol. 538, pp.207-220, 1985.
- [2.37] C. A. Mack. "Inside PROLITH, A Comprehensive Guide to Optical Lithography Simulation", FINLE Technology, 1997.
- [2.38] F. H. Dill, A.R. Neureuther, J.A. Tuttle and E. J. Walker. "Modeling projection printing of positive photoresists". IEEE Trans. Electron Devices, ED-22, no. 7, pp.456-464, 1975.
- [2.39] W.G. Oldham, S.N. Nadgaonkar, A. R. Neureuther, and M. M. O'Tole, "A general simulator for VLSI lithography and etching process: Part I – Applications to projection lithography", IEEE Trans. Electron Devices, vol. ED-26, no. 4, pp.717-722, 1979.
- [2.40] A. R. Neureuther and F. H. Dill. "Photoresist modeling and device fabrication applications". Optical and Acoustical Micro-Electronics, Polytechnic Press (New York: 1974).
- [2.41] C. Hsu, R. Tang, J. Cheng, P. Chien, V. Wen, and A. R. Neureuther. "Lithography analysis using virtual access (LAVA) web resource". SPIE Vol. 3334, pp. 197-201, 1998.
- [2.42] M. Born and E. Wolf. "Principles of Optics" 7th (expanded) edition. Cambridge University Press, 1999.
- [2.43] Y. Bordovsky, "Impact of local partial coherence variations on exposure tool performance", SPIE Vol. 2440, pp750-770.
- [2.44] H. Liu, C. Yu and B. Gleason, "Contributions of stepper lenses to systematic CD errors within exposure fields" SPIE Vol. 2440, pp868-877

Chapter 3

CD-SEM Based Pattern Recognition for Focus/Exposure Control

Fluctuations in focus and exposure dose are important sources of die-to-die spatial variation. In this chapter, a method is introduced to utilize digitized CD-SEM traces collected from small test patterns to infer these two critical process parameters, focus distance and exposure dose. The basic premise is that these signals contain latent information about not just the Critical Dimension, but also about the shape of the profile. Extracting this latent information is accomplished in two steps: First, principal component analysis (PCA) is applied to extract the characteristic features of the SEM image. Then a feedforward neural network trained by back-propagation is used to classify the traces based on the processing conditions of the samples. The ultimate goal is to further exploit the information content of CD-SEM scans for lithography control and diagnosis.

3.1 Introduction

Two key controllable parameters in lithography are the focus distance and the exposure dose. As the printable feature size has been pushed to the resolution limit, the process window (particularly the exposure/focus latitude) shrinks dramatically. New resolution enhancement techniques, such as off-axis illumination and optical proximity correction, phase shift mask are very sensitive to the setting of defocus distance and exposure dose.

Therefore, setting these two parameters precisely is of great importance in reducing CD variation across the exposure field, within wafer and lot-to-lot [3.1][3.3]. Traditionally, the optimal settings are selected by operators who examine the patterns projected under different settings across a wafer. This is time consuming and error-prone, and thus not suitable for today's IC manufacturing needs. In addition, the stepper, light source and chemicals tend to age, and the optimal settings drift with time. It is therefore desirable to find an automatic way to control these two important parameters [3.2][3.4].

Previous work has automatically recognized the focus/exposure related problems for 1 μ m I-line lithography using digitized images from an optical microscope [3.6]. This work uses a CD-SEM to examine sub-micron patterns. The full wave form of the CD-SEM scan is used to infer corresponding process conditions.

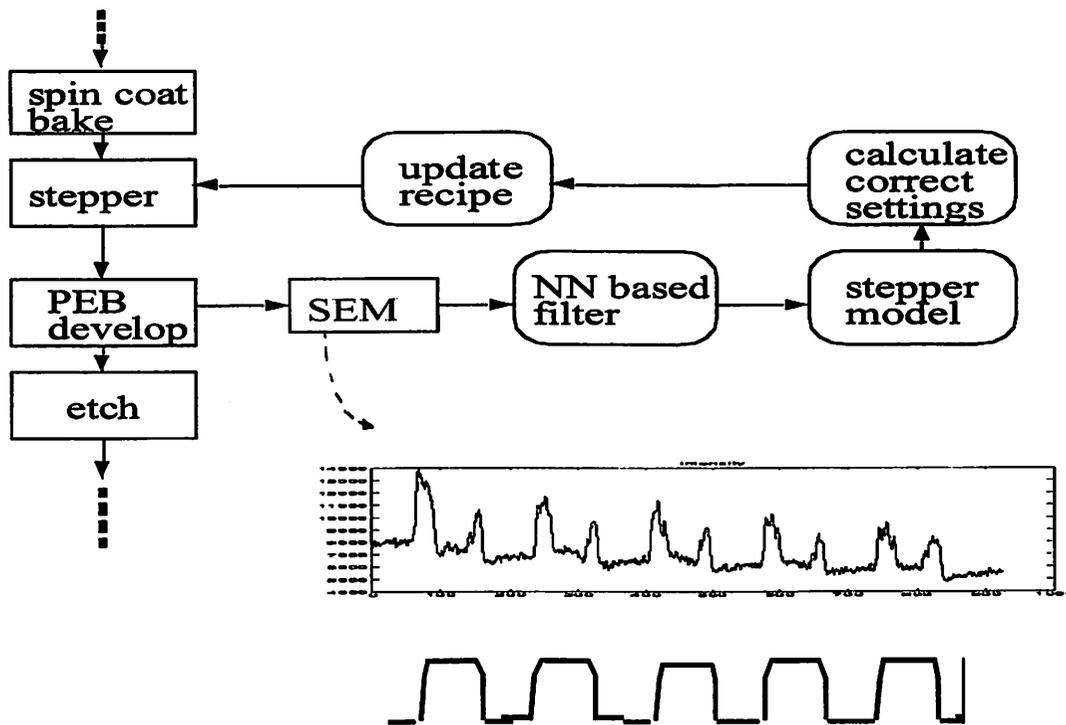


Figure 3-1 Concept of CD-SEM based focus/exposure control in lithography

CD-SEM is a commonly used metrology tool in production with nanometer order precision. Compared to cross section SEM, it has the advantage of relatively high speed and it is not destructive. Figure 3-2 and figure 3-3 show examples of CD-SEM traces for two different line-space patterns with different CD profiles. The direction of the scan is perpendicular to the lines. The CD-SEM image intensity is not directly related to the height of the feature under testing. The CD sidewall region corresponds to the peak in the SEM scan. The valley between two peaks is related to the top of CD pattern. So a pair of peaks represents a line pattern.

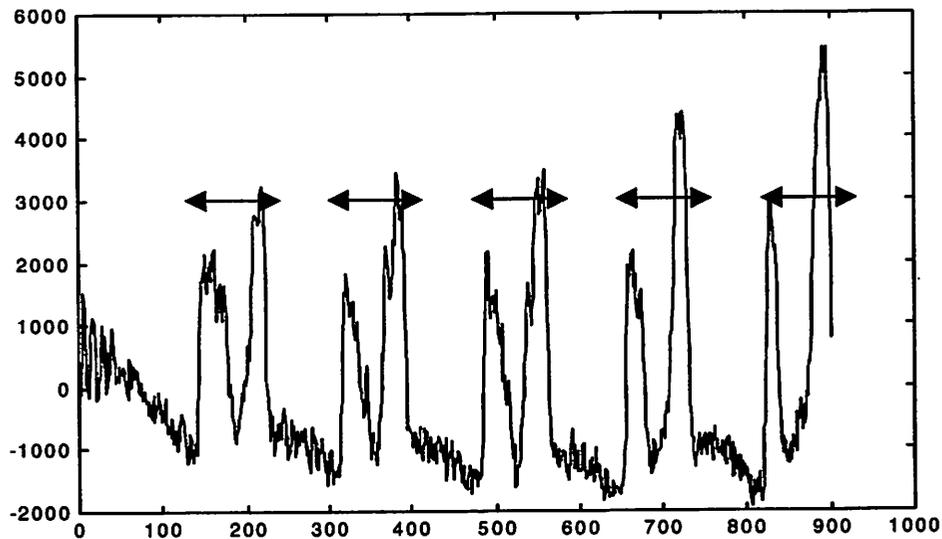


Figure 3-2 CD-SEM scan for five-line dense pattern

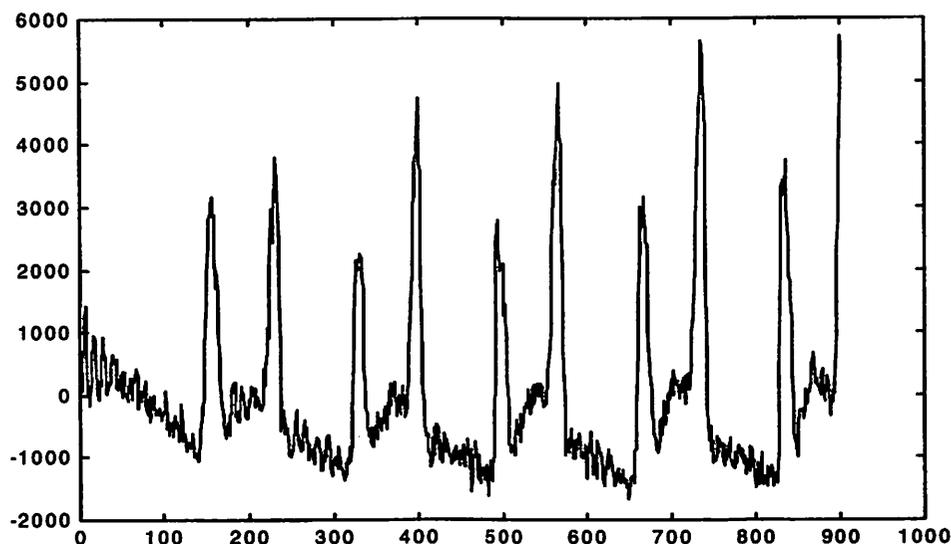


Figure 3-3 CD-SEM scan for five-line dense pattern

The algorithms for determining the linewidth based on CD-SEM scan are of heuristic nature. Also, the methods for determining the actual edge position can vary. However, for a given material property of the sample, such as surface roughness, absorption and reflectivity, as well as a given chamber condition, such as configuration and pressure, the CD-SEM trace is supposed to correspond to actual pattern profile. Therefore, the information about the pattern profile (width, height, sidewall) is expressed in some way through the CD-SEM trace. However, the detailed relation between actual profile and CD-SEM can only be obtained by accurate modeling of the interaction between the electron beam and the material. Such a model will need to correctly characterize sample properties and chamber conditions, which are formidably difficult to capture.

So, our approach will be to extract this information by “pattern matching”. This approach does not guarantee a unique or even an existing solution, so it has to be demonstrated experimentally. The basic tool for our approach is a neural network that

will be used to model the complex relationship between the CDSEM trace and the photolithography settings.

When it comes to creating NN models that utilize the CD-SEM trace as an input in order to infer the process settings, some simplifications are in order. Typically, there are about one thousand pixels for a CD-SEM scan of several microns. Obviously, it is impossible to treat all these pixels directly as inputs. Through dimensional reduction, we can obtain a reduced data set that is easier for subsequent handling. The signal noise and measurement error included in the original data may also be reduced through data preprocessing.

Principal component analysis (PCA) is a commonly used technique for data reduction, and has been shown to facilitate many types of data analysis in process engineering. PCA uses linear functions to model the relationship between the original data and a reduced set of “latent variables” [3.7][3.8][3.9]. The aim of PCA is to extract normalized orthogonal vectors U_i , for $i=1,2, \dots,M$, in the input space that account for as much of the variance in the data as possible. The N -dimensional input data is transformed to a lower M -dimensional space ($M < N$) without losing essential intrinsic information. PCA is a linear technique in the sense that it uses linear functions to model relationship between the original data and the latent variables.

The i th principal component of the original N -dimensional data vector \tilde{X} is

$$P_i = U_i'[\tilde{X} - \tilde{X}'] \quad (3-1)$$

where \tilde{X}' is the mean vector of \tilde{X} and U_i' is the i th normalized eigenvector of the covariance matrix associated with the i th largest eigenvalue e_i .

$$\tilde{L} = \tilde{U}'\tilde{D}\tilde{U} \quad (3-2)$$

where \tilde{L} is the covariance matrix and \tilde{D} is the diagonal matrix with e_i as diagonal entries.

Once a data set has been reduced to a manageable size, a neural net [3.10-3.14] can recognize and classify the different traces with very good accuracy. The network inputs are the first 30 principal components calculated from the original digitized SEM data. The output nodes represent various focus/exposure combinations that were used to produce the samples being tested.

We used the first 30 principal components, calculated from original CD-SEM trace. 95% variance is captured. Figure 3-4 depicts the configuration of the neural network.

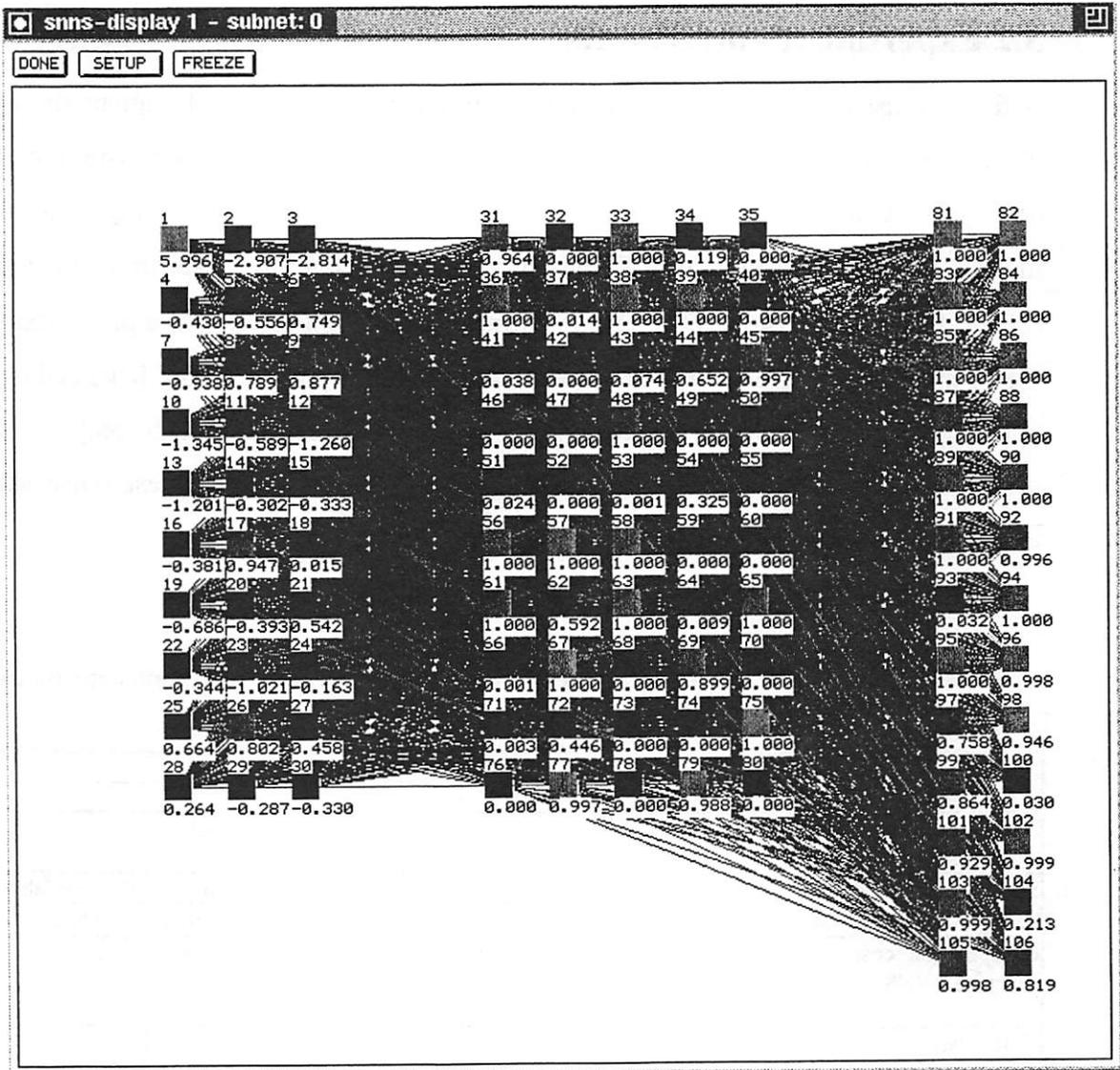


Figure 3-4 Neural network configuration (30 input nodes, one hidden layer with 50 nodes and 26 output nodes)

3.2 Experiment and Results

A five line/space pattern with nominal 0.35 μm line width is exposed sequentially under different focus/exposure settings across a wafer using an I-line stepper. The numerical aperture is set as 0.6 and partial coherence is 0.6. The exposure dose changes from 255 mJ to 315 mJ in 13 steps of 5mJ. The focus varies from $-0.9 \mu\text{m}$ to $0.3 \mu\text{m}$ in 13 steps of $0.1 \mu\text{m}$. In total, there are 169 different settings. Five six-inch wafers are processed with the 13 x 13 exposure/focus matrix on each of them. After post-exposure bake and development, scanning electron microscope traces are taken from each pattern. Each measurement is actually the average of 32 scans in order to filter test noise and to improve measurement repeatability.

Table 3-1 Summary of experiment and results for NN-based focus/exposure automatic control

Number of wafers processed	5
Number of samples tested by CD-SEM	845
Samples used for training	676 samples from the 1st, 2nd, 4th and 5th wafer
Samples used for testing	169 samples from the 3rd wafer
Computation cost for training using neural network software SNNS	30 minutes on Sparc2
Computation cost for testing	\ll 1 second
Classification accuracy for exposure settings	perfect classification: 74.0%; within +/-5mJ: 96.4%
Classification accuracy for focus settings	perfect classification:40.1%; within +/-0.1 μm : 82.8%

In order to test the network's capability of recognizing patterns made under the same focus/exposure conditions, we use training data from four wafers and testing data from

the remaining wafer. Table 3-2 displays the classification results in comparison with the actual settings. The entry in the table is the prediction frequency. Since focus latitude for the dense and isolated 0.35 μm features is approximately 0.6 μm and 0.8 μm respectively [3.15], the 0.3 μm control precision using this method is approximately 50% of the process window. 0.3 μm range has 82.8% (approximately 2.7σ) prediction accuracy. Therefore 1σ prediction is about 0.11 μm . The percentage of 1σ prediction is thus 20% ($0.11 \mu\text{m}/0.6\mu\text{m} = 20\%$) of depth of focus window. The dose window for dense pattern for 365 nm i-line process is approximately 14% of the dose to clear [3.16]. Since the dose to clear data is not available, we assume the maximum dose (315 mJ) is the upper limit of the process window and minimum dose (255 mJ) is the lower limit of the process window for the i-line process under study. Three 5 mJ bins contain 96.4% (approximately 4.2σ). Thus 1σ is $15\text{mJ}/4.2 = 3.6 \text{ mJ}$ and this is 1.3% of the average dose or 1/18 of the dose window..

Through examining SEM traces, the neural net can recognize the corresponding focus/exposure settings. Similarly, other information about the resist profile, such as sidewall angle and critical dimension variation, can be provided to the network during the learning phase. Therefore, this method could lead to an automated metrology to evaluate the pattern quality in lithography. The pattern with the sharpest sidewall and least CD variation will be chosen during instrument calibration. In addition, a deteriorated pattern can be automatically classified during processing, and the appropriate lithography settings could be automatically deduced.

Table 3-2 Comparison of predicted setting and actual settings, (a) exposure dose, (b) defocus distance

Predicted Exposure Setting (in 5 mJ/cm ² steps)	13	1										6	8	
	12											6	7	2
	11										2	6		1
	10									3	10			
	9								2	10	1			1
	8								9					1
	7							10	2			1		
	6						11	2						
	5				1	11	2							
	4			1	10	2								
	3			10	2									
	2		10	2										
	1	12	3											
	1	2	3	4	5	6	7	8	9	10	11	12	13	

Actual Exposure Setting (in 5 mJ/cm² steps)

(a)

Predicted Defocus Distance (in 0.1 μm steps)	13										1	1	6
	12										2	9	7
	11									2	2	4	1
	10						1			1	2	4	
	9				1		1	2	3	5	5	2	1
	8						3	2	6	3	1		1
	7					2	3	4	3	2	2		
	6		1	4	2	3	5	3			1		
	5	1			2	6		1	11				
	4			5	6	1		1					
	3		7	3	1	1							
	2	2	5	1	1								
	1	10											
	1	2	3	4	5	6	7	8	9	10	11	12	13

Actual Defocus Distance (in 0.1 μm steps)

(b)

3.3 Conclusion

CD-SEM traces can provide good reference for process control and diagnosis. The method we are pursuing is very easy to implement in a manufacturing process by combining this method with the routine SEM inspection. A good pattern recognition neural network for focus/exposure control has been established in this project. We successfully applied PCA and neural network in focus/exposure control of lithography and very high recognition accuracy has been achieved. One σ of dose prediction accuracy is approximately 1/18 of the dose window and one σ of defocus prediction accuracy is approximately 1/5 of the depth of focus window.

This work is meant to demonstrate the concept and it is complete. Future analysis is possible, including obtaining better data sets using deep ultraviolet lithography. We also plan to create more classification categories including specific develop problems. In some situations, a CD-SEM fails to measure CD due to serious deterioration in the pattern. It will be helpful to automatically classify special patterns and diagnose such cases in the future. It will also be interesting to apply the proposed method in the presence of more detailed profile data such as what can be collected either with Atomic Force Microscopy, or with full-profile Scatterometry.

References for Chapter 3

- [3.1] L. Pain, Y. Trouiller, a. Barberet, O. Guirimand, G. Fanget, N. Martin, Y. Quere, M. E. Nier, E.Lajoinie, D. Louis, M. Heitzmann, P. Scheiblin, A. Toffoli, "193 nm scanner characterization by SEM and electrical CD measurements". Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 3998, pp.96-107, 2000.
- [3.2] P. K. Govil, J. Tsacoyeanes, R. Eron, D. Walters, "Contributors to focal plane nonuniformity and their impact on linewidth control in a DUV and scan system", Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 3334, pp. 92-103, 1997.
- [3.3] Y. Kim, G. Yeo, H. Shin, H. Kim, H. hang, and U. Chung, "CD control for quarter micron logic device gates by using iso-pitch bias", Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 3334, pp.629-639, 1997.
- [3.4] J. Sturtevant et al, "CD control challenges for sub-0.25 μm patterning", IEEE Lithography Symposium, 1996.
- [3.5] L. Liebmann, et al, "Understanding across chip line width variation: the first step toward optical proximity correction", Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 3051, pp.124-132, 1997.
- [3.6] P. Tsai and C.J. Spanos, "A neural-net based, in-line focus/exposure monitor", Advanced semiconductor manufacturing conference and workshop, 1994.
- [3.7] J. Karhunen etc, "Generalization of principal component analysis, optimization problems, and neural networks", Neural networks, pp549, vol.8 No.4 1995.
- [3.8] S. Tan, "Reducing data dimensionality through optimizing neural network inputs", AIChE, pp1471, Juen 1995.
- [3.9] M. Kramer, "Nonlinear principal component analysis using autoassociative neural networks". AIChE. pp233, Feb. 1991, Vol.37.
- [3.10] A. Cichoki and R. Unbehauen. "Neural networks for optimization and signal processing", John Wiley and Sons, Chichester, 1993.
- [3.11] C. M. Bishop, "Neural networks for pattern recognition", Oxford University Press, 1995.

[3.12] J. J. Hopfield and D. W. Tank, "Neural computation of decision in optimization problems", *Biological Cybernetics*, vol. 52, pp.141-152, 1985.

[3.13] L. B. Almeida and C. Wellekens, "Neural networks", vol. 412 of *lecture indexes in computer science*, Springer Verlag, February 1990.

[3.14] SNNS: a neural network simulator software.

[3.15] S. Sethi, M. Barrick, J. Massey, C. Froelich, M. Weilemann, F. Garza, "Lithograph strategy for patterning 0.35 μ m devices". *SPIE*. 2440, pp. 619-632, 1995.

[3.16] J. Sturtevant, B. Ho, K. Lucas, J. Peterson, C. Mack, E. Charrier, B. Peterson, N. Koshiba, G. Barnes, "Considerations for the use of application-specific photoresists", *SPIE*, vol.3678, 1999.

Chapter 4

Spatial Variation Characterization in DUV Lithography Using Electrical Metrology

4.1 Introduction

In order to characterize spatial variation in lithography, it is necessary to measure linewidth distribution across the field as well as across the wafer. Based on the measurement data, the variation can then be decomposed into different error categories, including random and systematic components. Once the deterministic spatial pattern has been identified, one can analyze the physical causes underlying such systematic variation patterns and accurately model them. Therefore, it is necessary to repeatedly collect this spatial distribution data with relatively high spatial density coverage as well as high accuracy.

We chose to base the metrology for this study on electrical testing. The reason is based on the following considerations. To extract systematic CD variation within the field or across the wafer, extensive measurement is needed. Therefore, sampling speed is of utmost concern in selecting the metrology method. CD-SEM is not adequate because it is relatively slow when collecting large amounts of data. Meanwhile, the CD-SEM precision is also in doubt because of, among other things, the edge roughness of the pattern that degrades the repeatability of measurement. Although CD-SEM typically has a precision of ~4-5 nm, this is still relatively large when compared to electrical test [4.7-4.10]. Electrical metrology has demonstrated the capability of resolving linewidths well

below 100 nm [4.1][4.2][4.3]. Moreover, CD-SEM is highly pattern profile dependent [4.12]. It is a common phenomenon that CD-SEM fails to give meaningful readings when the pattern degrades to some extent, such as when T-topping or footing features are present. The following figure of cross-sectional SEM shows a typical example of footing for dense lines, which may negatively impact top-down CD-SEM data.

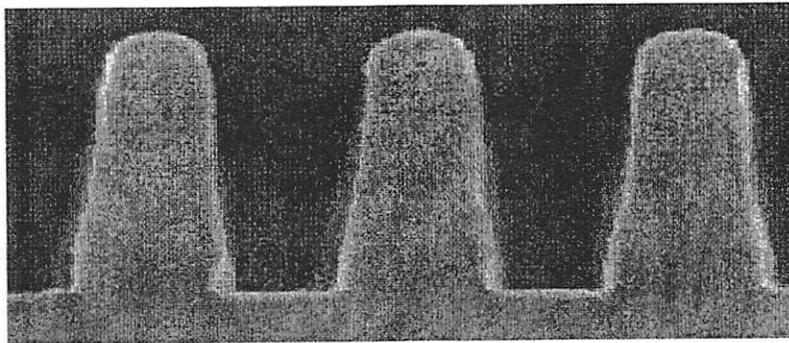


Figure 4-1 Cross-sectional SEM image showing footing features

On the other hand, electrical measurement has the advantage of high speed when an automated probe station is used. Further, it is independent of the pattern profile and has sub-nanometer precision if the measurement conditions are optimized to take into account the doping level, the forced current and current source resolution, and joule heating [4.4][4.6][4.11].

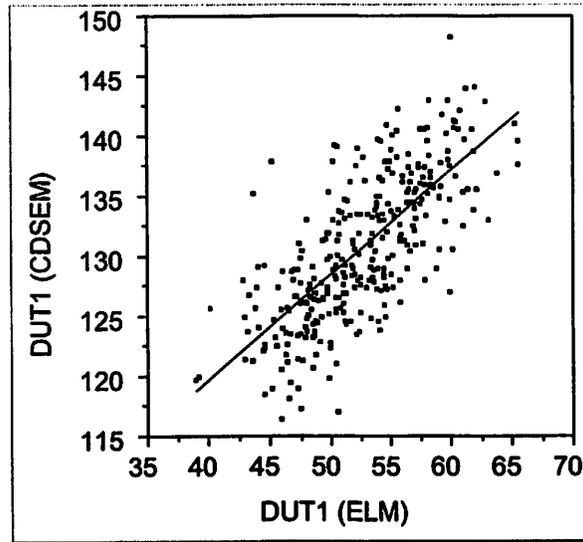


Figure 4-2 Comparison of CD-SEM measurement versus electrical test

Figure 4-2 shows the comparison of the measurement results for the same test sample using the two metrologies [4.16]. It is noted that CD-SEM has a very large measurement noise compared to electrical test. 3σ of CD-SEM repeatability is ± 14.2 nm. The relation between CD-SEM and electrical test can be modeled as,

$$\text{CD-SEM} = 84.87 + 0.87 * \text{ELM} \quad (4-1)$$

It is clear that a significant bias, as large as 85nm, exists between measured linewidth from the two metrologies.

A single layer mask is designed so that electrical test structures are patterned on doped polysilicon. The CD is then measured using electrical test. Generally, short-loop experiments are necessary to characterize micro-loading when pattern density varies significantly. This step can be skipped if the mask is designed in such a way that the pattern density for any particular feature is uniform throughout the patterning field. In addition to measuring the printed patterns on the wafer, the mask must also be

characterized. Mask characterization can be conducted either by SEM or by optical means.

With wafer CD map and mask CD data at hand, full field lens aberration information can be deduced using the algorithm proposed in section 5.3. By incorporating existing OPC, one can devise a set of systematic mask based error correction rules. These rules may be scanner/stepper specific and field position dependent. The final goal of this work is an integrated solution of mask engineering, consisting of lens system characterization, OPC with lens aberrations and PSM with lens aberrations.

4.2 Experiment

4.2.1 Electrical Linewidth Measurement (ELM) Mask Designs ELM-1 And ELM-2

An electrical linewidth measurement (ELM) mask has been designed so that the spatial coverage of identical features in the field is maximized (19×24 points in stepper field). Figure 4-3 shows the basic module with one Van der Pauw structure for measuring sheet resistance and eight Kelvin structures for measuring linewidth, along with mask linearity test patterns, an MEF module, and cross-section SEM lines.

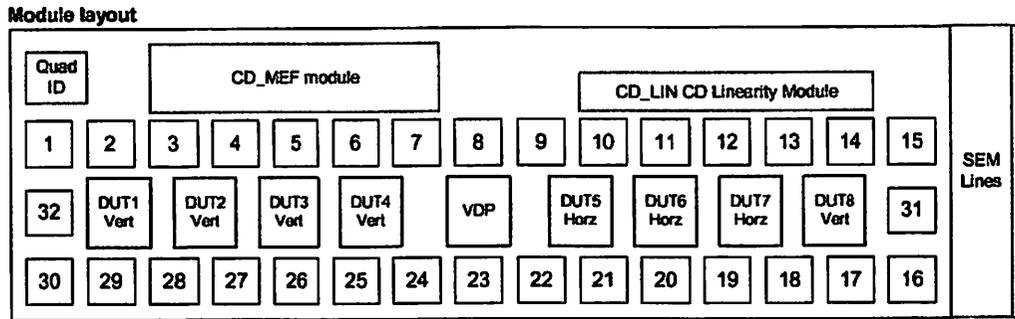


Figure 4-3 Floor plan of basic electrical test module

Eight Kelvin linewidth test structures (also referred to as devices under test, or DUTs) are laid out on either side of the center Van der Pauw (VDP) sheet resistance test structure. Linewidth is calculated from the VDP and DUT measurements. Four different combinations of orientation (horizontal and vertical) and line spacing (isolated and dense) are repeated twice in each module. The OPC module is identical to the non-OPC module, with the exception of the scattering bars that are added next to each main linewidth measurement feature. Therefore, there are eight different feature cases. The design variables are summarized as follows.

Table 4-1 Design variables for ELM1 mask

<u>Space</u>	<u>Orientation</u>	<u>OPC</u>
Isolated	Vertical	Yes
Dense	Horizontal	No

The pad frame module contains 32 active pads. Four of them are assigned to the Van der Pauw structure and the remaining pads are used by the eight Kelvin structures that also share the common potential node). The mask is designed for a single polysilicon layer. The contact pads are all made of polysilicon so as to simplify the processing as well as the analysis, since a single lithography step is needed for patterning.

The nominal printed linewidth of the first ELM mask (designated as ELM-1) is set at 220 nm. Figure 4-4 illustrates the layout of individual linewidth measurement structures.

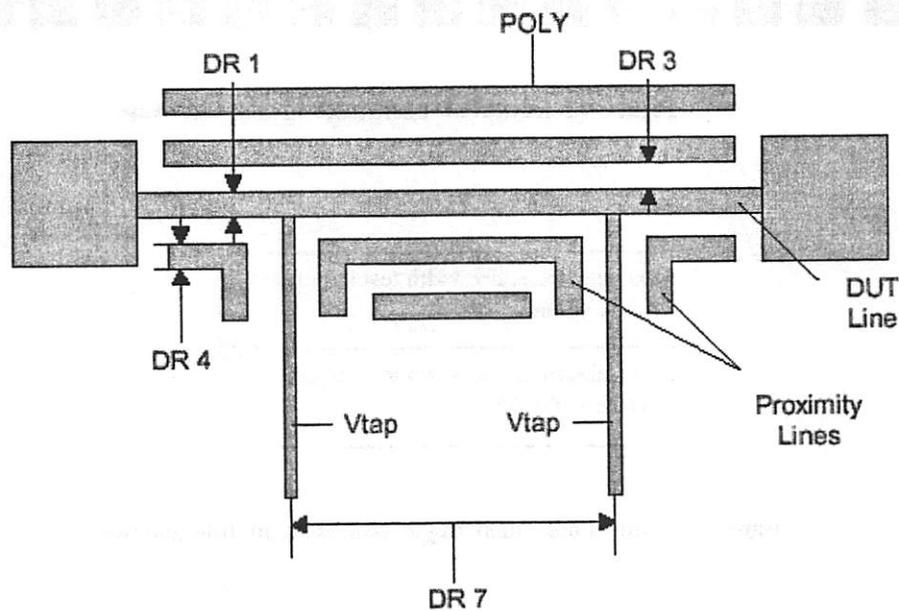


Figure 4-4 Layout of linewidth test structure for dense pattern

The two pads on either side are used to force current and the Vtap nodes are dedicated to voltage-sensing. Isolated features are defined so that the spacing between main feature and proximity lines are relatively far apart (DR3=1.2 μ m). The actual layout differs from the figure in terms of positioning of pads and proximity lines. Figure 4-4 shows the

actual layout with CD-MEF module and CD-linearity module hidden. The orientation and the relative size of the features with respect to the pads can also be seen from this figure.

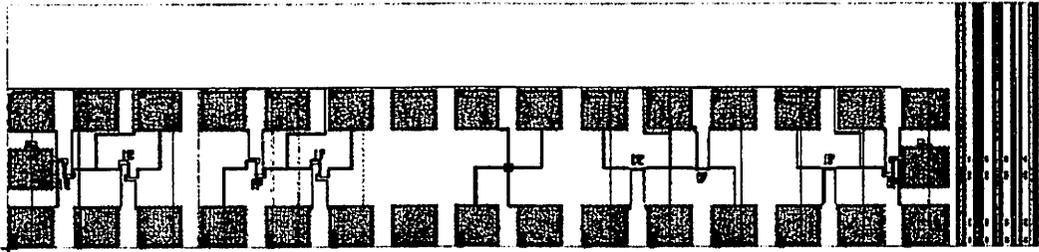


Figure 4-5 Layout of basic linewidth test module

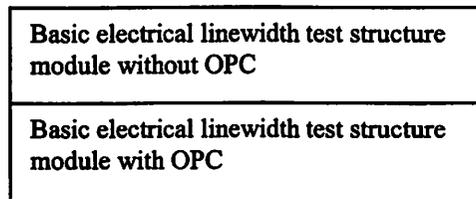


Figure 4-6 Sub-block consisting of two basic module frames

Figure 4-6 shows a sub-block that consists of two test modules of electrical linewidth, with and without OPC, respectively. This sub-block is repeated 228 times on a 19×12 grid across the field. Since each feature is duplicated twice within the basic module, the spatial density of every feature case is 19×24 . The de-magnified (4 \times) mask area is $35.04 \times 29.07 \text{ mm}^2$. Because this is larger than the typical stepper field size, some portions of the mask may not be able to be exposed. The mask is designed in this way so that the maximum field size can be printed to capture the spatial linewidth variability.

In addition to the electrical test pattern, two CD-SEM measurement patterns are also designed into the basic module. The CD-linearity module is designed to calibrate the mask-making linearity. It can also be used to measure mask error factor (MEF). The experimental results will be explained in chapter 5. Both the mask pattern and resist pattern can only be measured by CD-SEM. A wide range of CD values from 25 nm to 1.25 μm is laid out within the module. This will provide measured MEF vs. linewidth for very large CD range and it will be used to compare with theoretical MEF formulation, which will be introduced in chapter 6.

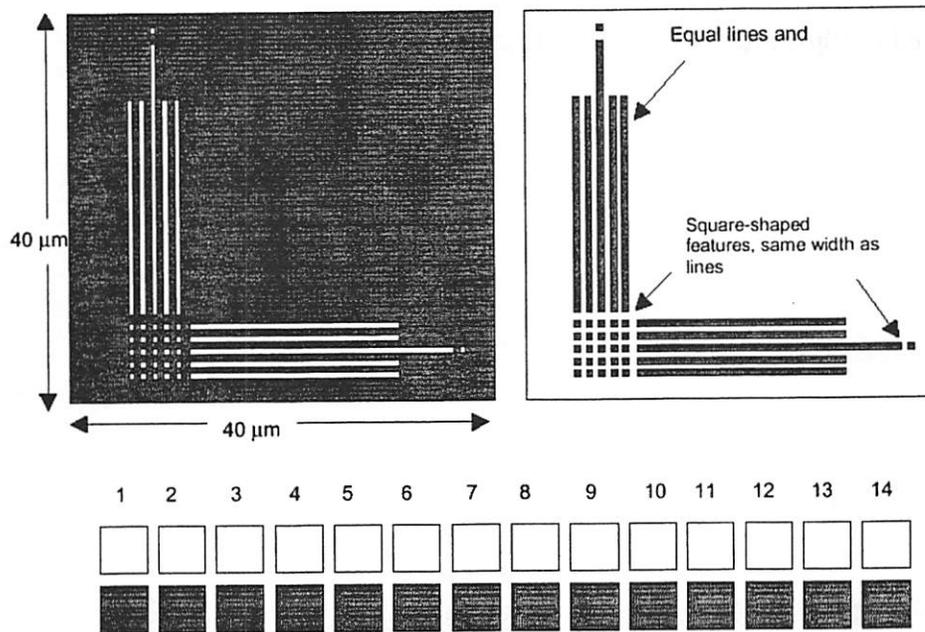


Figure 4-7 CD linearity module

Fourteen different CD values have been used in both clear and dark field patterns. Isolated and dense lines, as well as square-shaped features, are taken into account. Each pattern occupies only $40 \times 40 \mu\text{m}^2$ area and they are densely packed into two rows above the electrical test patterns.

The CD-MEF module is designed to physically measure how much mask correction is needed to obtain the printed CD that is closest to the target value. Two types of correction have been designed. One type is designed to correct CD for both main features and scattering bars, while the other type only corrects the center main features. Five different levels of correction are used, $0, \pm 1.5\epsilon, \pm 3\epsilon$, where ϵ is the unit. Both vertical and horizontal directions are taken into consideration. Nine different spacings from isolated ($>10\mu\text{m}$) to $0.32 \mu\text{m}$ are designed.

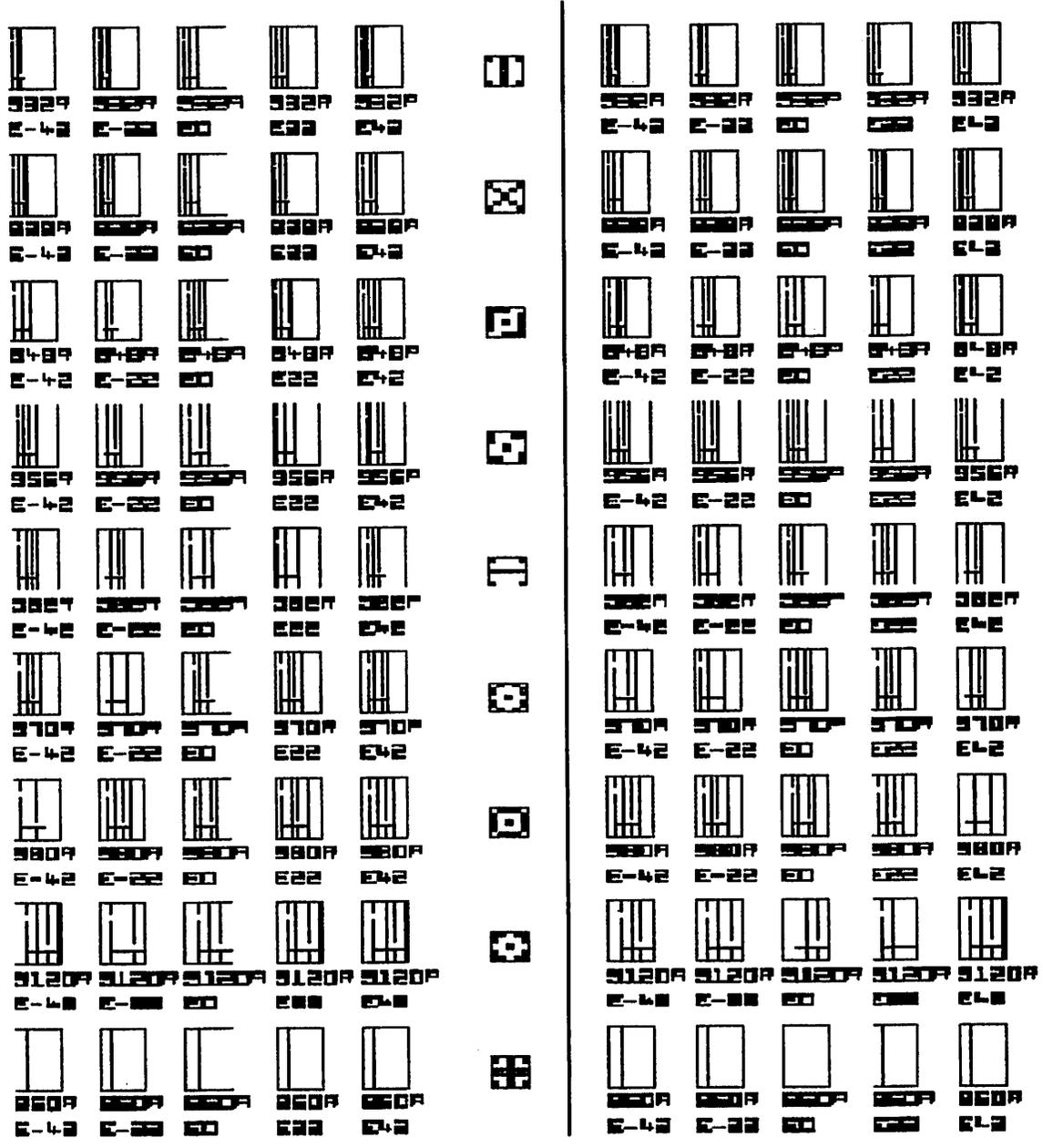


Figure 4-9 CD-MEF module layout

Cross-section SEM lines are useful in calibrating electrical CD metrology. The SEM lines are placed at the right side of each basic e-test module. They are continuous across the entire chip, so that a cut at any point will yield usable cross-section data. Both OPC and non-OPC lines are included, and different feature spaces are considered.

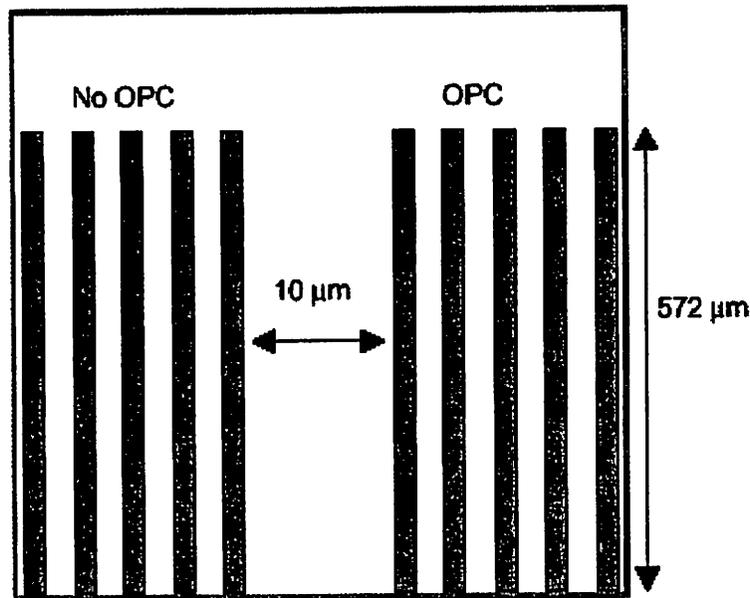


Figure 4-10 Cross-sectional SEM lines

A second ELM mask (ELM-2) is designed to be similar to the ELM-1 mask, except in this case the CD is decreased to 180 nm (the designed final CD after etching is 130 nm) and scatterometry grating patterns are added to the non-OPC modules. The purpose of these eight grating patterns will be explained in section 5.6 of chapter 5 and chapter 7. They can be used as test structures for future improvement of an aberration extraction method which will be introduced in chapter 5.

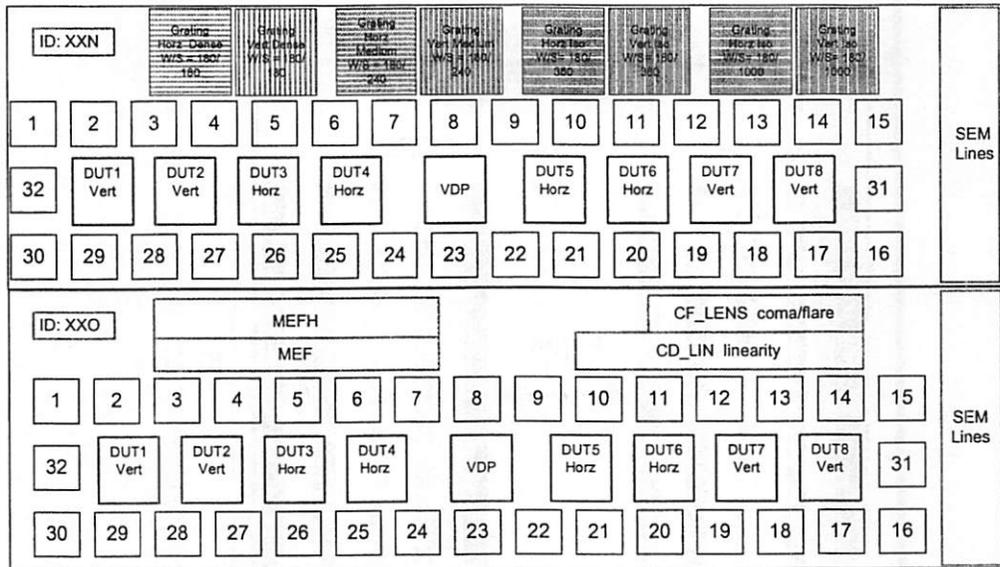


Figure 4-11 Floor plan of basic module of ELM-2 mask

The pad frame module is the same as the ELM-1 mask, so the probe card and measurement setup is compatible with those for the ELM-1 mask. The addition of the scattering feature is meant to facilitate scatterometry measurements. More specifically, it has been shown that a broadband ellipsometer is capable of measuring the detailed pattern profile with high speed and accuracy [4.13]. Certainly, the pattern profile will yield more information than the CD value alone. On the other hand, it is meaningful to compare and calibrate these two metrologies in a single design.

4.3 Requirements of the Electrical Testing Instrumentation

The electrical parameter test system consists of a Sun workstation, an Electroglass 2001x [4.14], HP 4085A switching matrix, HP 4084A switching matrix controller and HP 4142 source monitor. The system is controlled by SUNBASE [4.15], a program that interprets the test program input files, defines measurement routines, and outputs measurement results. Input files determine the X-Y coordinates of all test structures at both the die scale and the wafer scale. The newly written SUNBASE software is also capable of allowing the user to define different voltage and current source/monitor units to be connected to test structures so that the best measurement conditions can be selected.

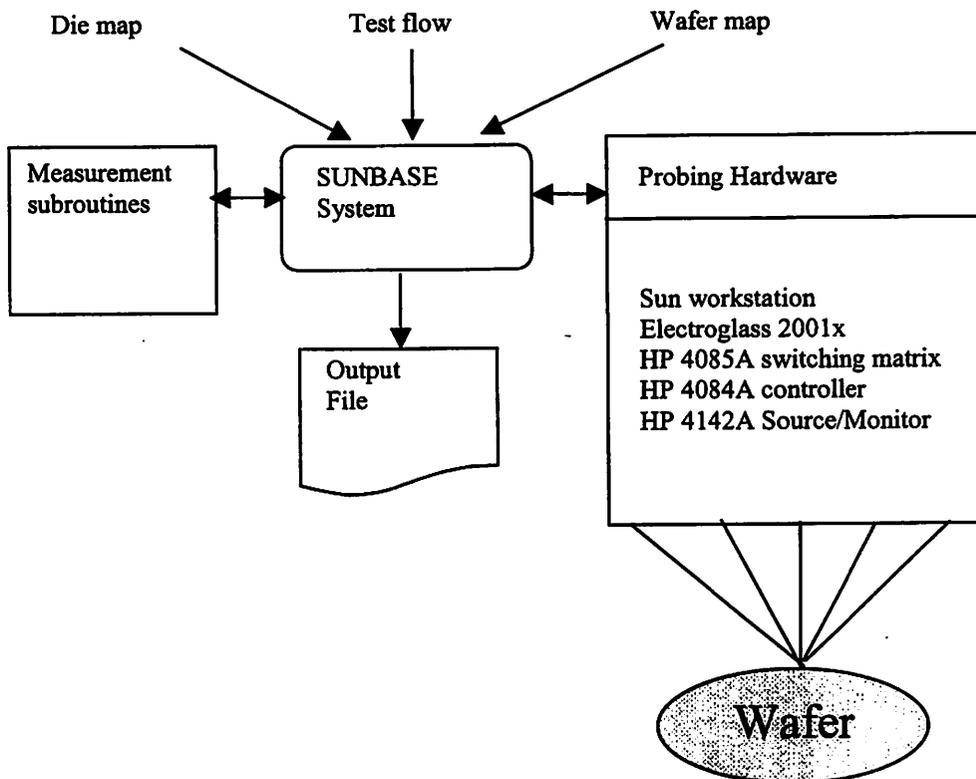


Figure 4-12 E-test setup

The original source monitor unit model is the HP4141, whose resolution is too coarse to measure sub-micron features. Low sheet resistance of the VDP required a large current

to be used to obtain a significant voltage difference between the two voltage measurement units under the old system ($\sim 0.5\text{mV}$ resolution), which caused significant heating problems. The measurement consistency was not satisfying in the sense that measured sheet resistance varies significantly with forced current and delay time.

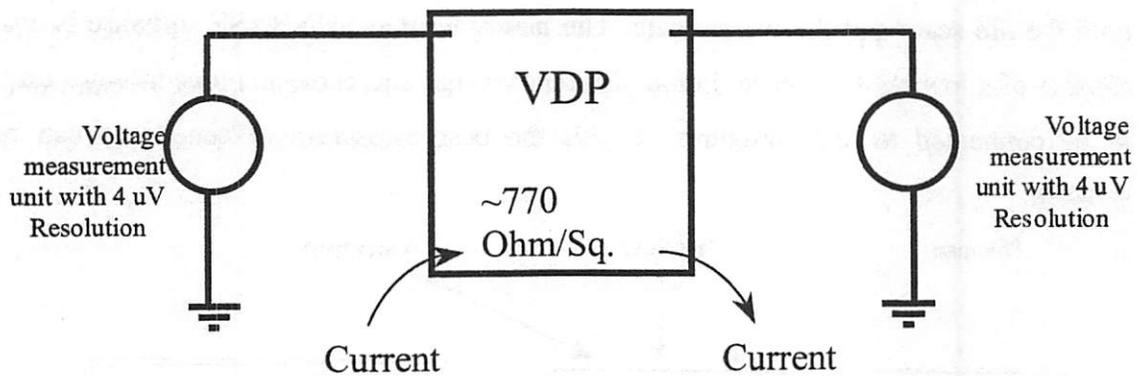


Figure 4-13 Updating E-test using high resolution source/monitor units

The system has been upgraded with much higher resolution ($4\mu\text{V}$) Source/Monitor Units to allow uses of much smaller currents, reducing the heating problem. Figure 4-14 is the

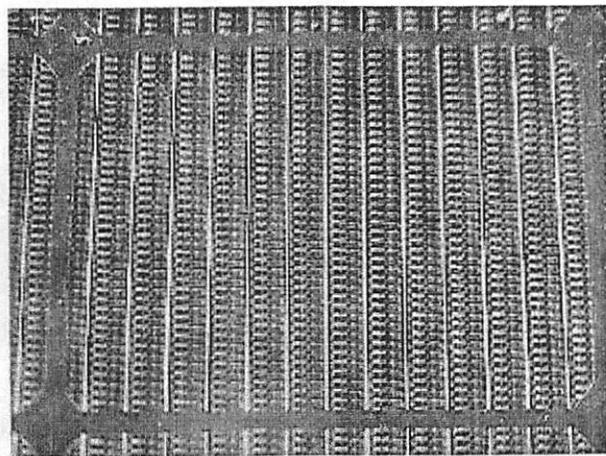


Figure 4-14 Chip image

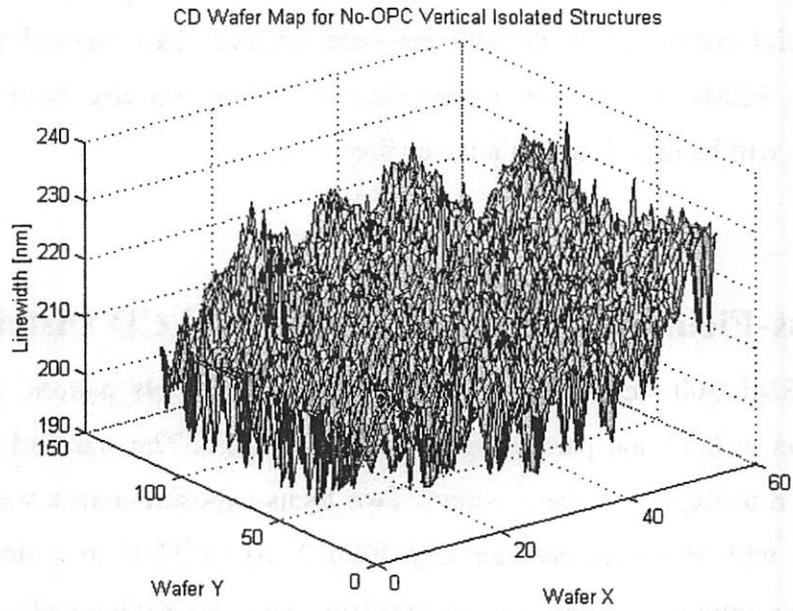
chip image of 24mm x 30mm size. Various requirements of pitch, orientation, and optical proximity correction (OPC) features have been taken into account. Multiple wafers have been processed using standard mass production conditions, as well as under different partial coherence, focus and exposure settings. The nominal feature size is 220nm. (The ELM-2 mask with feature size of 180nm has also been designed and fabricated, as will be introduced in later section)

4.4 Across-Field and Across-Wafer Spatial CD Distribution

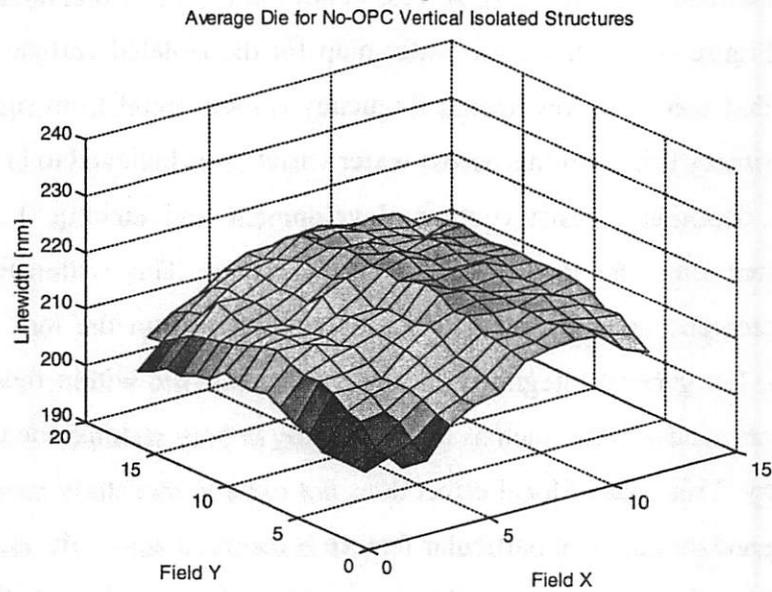
A 248nm ASML/300 stepper is used to print wafer with poly pattern. The numerical aperture is set at 0.57 and partial coherence is set at 0.6. The standard dose and best focus are used to expose uniform wafers. Two focus-exposure-matrix wafers have also been printed with exposure dose ranging from 21mJ to 27mJ in a step of 1mJ and defocus from 0.0 μ m to -0.4 μ m in a step of 0.1 μ m. The nominal linewidth is 220nm with k1 factor 0.51. Photoresist is UV110 with thickness of 5400A.

The uniform wafers, with the same process condition for all 31 die, have been measured electrically. Figure 4-15a shows the wafer map for the isolated vertical line features. It can be seen that there is a low spatial frequency smooth trend from right to left across wafer. The primary source of this across wafer variation is believed to be non-uniformity in thin film deposition, resist coating, development and etching. From the plot, a periodic pattern from die to die can also be observed. This within-field variation is generally systematic in nature, which results primarily from the lens aberrations and reticle errors. For general integrated circuits, the systematic within-field variation may also come from local effects, such as micro-loading in poly etching due to different local pattern density. This type of local effect does not exist in this study because the ELM-1 mask is designed so that each particular feature is identical across the chip. Figures 4-15 b, c and d show the average field CD map, wafer residual and mask CD respectively,

plotted in the same scale and range. It is evident that most CD variability appears within the field.

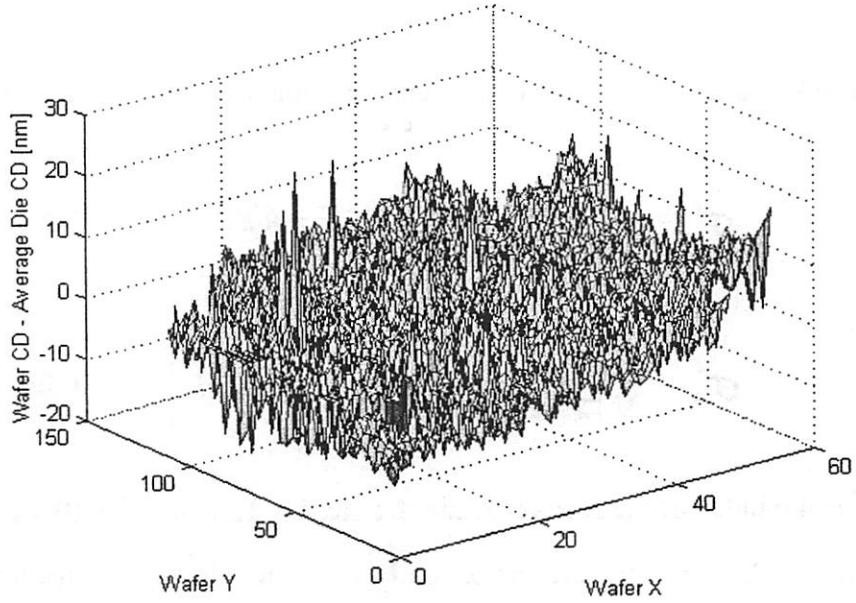


a) Wafer map



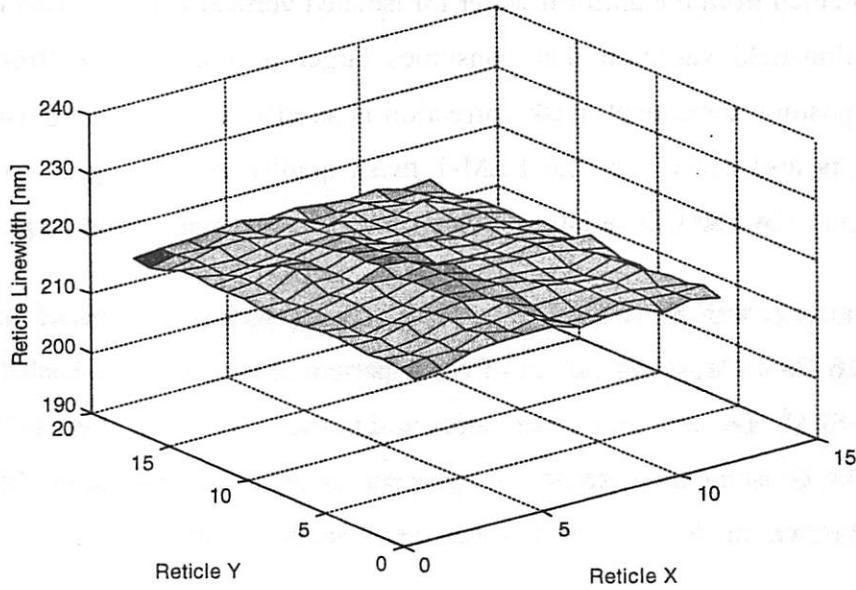
b) Die average

CD Wafer Map with Average Die Subtracted for No-OPC Vertical Isolated Structures



c) Across wafer residual

Reticle Measurements for No-OPC Vertical Isolated Structures



d) Reticle map

Figure 4-15 Wafer CD and mask CD measurement for isolated vertical features

Across wafer variation s_{aw}^2 and across field variation s_{af}^2 are defined as follows.

$$\sigma_{aw}^2 = \frac{1}{n^f} \sum_{i^f} \left[\frac{1}{N-1} \sum_{k=1}^N (CD_{i^f}^k - \overline{CD}_{i^f})^2 \right] \quad (4-2a)$$

$$\sigma_{af}^2 = \frac{1}{N} \sum_{k=1}^N \left[\frac{1}{n^f-1} \sum_{i^f} (CD_{i^f}^k - \overline{CD}^k)^2 \right] \quad (4-2b)$$

where i^f is the index of the feature within the die, $i^f = 1, 2, \dots, n^f$. $n^f = 19 \times 24 = 456$ for this design. $k = 1, 2, \dots, N$ is the die index. \overline{CD}_{i^f} is mean value of all dies for the feature of the same position within the die. \overline{CD}^k is the mean value of all features for the k^{th} die. From the measurement data excluding apparent outliers, $s_{aw} = 2.70$ nm and $s_{af} = 5.89$ nm are obtained from the uniform wafer for isolated vertical features. This implies that it is the within-field variation that consumes larger portion of CD error budget and therefore position dependent mask correction is an effective way to tighten overall CD spread. It is also noted that the ELM-1 mask quality is fairly good in terms of its contribution to overall CD variation given mask error factor is not too large.

Besides uniform wafers, two focus-exposure-matrix wafers are printed and measured. Figure 4-16 shows Bossung curves of resist pattern at the center of each die, measured using CD-SEM. The exposure dose varies within the range from 21mJ to 27mJ in steps of 1mJ. The Bossung characteristic is apparent for most exposure dose. Once exposure setting is known, the defocus can be estimated from these curves.

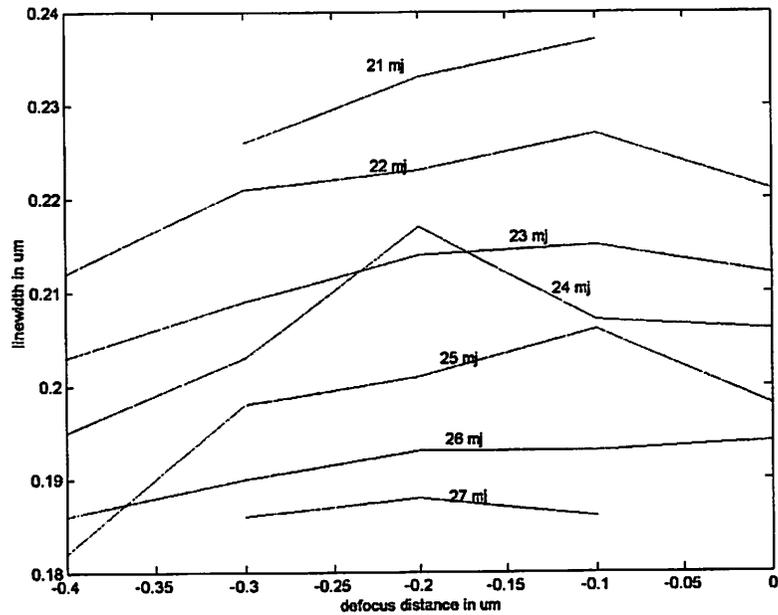


Figure 4-16 Bossung curves of resist pattern measured from focus/exposure matrix wafer

In summary, the following conclusions are drawn from the experiment. First, the major component of CD variation is from within the field. Secondly, most of this variability is systematic and deterministic in nature, so it may be assumed that lens aberrations and mask error magnification are among the most important sources of spatial systematic variation.

References for Chapter 4

- [4.1] A. Grenville, B. Coombs, J. Hutchinson, K. Kuhn, D. Miller, and P. Troccoli, "Electrical critical dimension metrology for 100-nm linewidths and below", Proceedings of SPIE, Optical Microlithography XIII, ed. C. Proglor, vol. 4000, pp.452-459, 2000.
- [4.2] T. Marschner, I. Pollentier, C. Baerts, I. Boltz, K. Ronse, J. Finders, H. K. Gangala, and L. Capodieci, "Qualification of electrical linewidth measurements as a metrology tool for 0.18 μ m and below", Proceedings of Olin Interface Conference 1998, pp.263-pp270.
- [4.3] P. Troccoli, L. Mantalas, R. Allen, and L. Linholm, "Extending electrical measurements to the 0.15 μ m regime", Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 1464, pp.90-103, 1991.
- [4.4] P.M. Hall, "Resistance calculations for thin film patterns", Thin Solid Films, Vol.1, pp.227-295, 1967-1968.
- [4.5] E. E. Chain and M. Griswold, "In-line electrical probe for CD metrology", Metrology, Inspection, and process control in Microlithography, Proceedings of the SPIE, vol. 2876, pp.135-146, 1996.
- [4.6] M. W. Cresswell, N. M. P. Guillaume, W. E. Lee, R. A. Allen, W. F. Guthrie, R. N. Ghostagore, Z. E. Osborne, N. Sullivan, and L. W. Linholm, "Extraction of sheet resistance from four-terminal sheet resistors replicated in monocrystalline films with nonplanar geometries", IEEE trans. on semiconductor manufacturing, vol. 12, no. 2, May, 1999.
- [4.7] R. I. Scarce, "Metrology: what next?" Solid State Technology, pp.43-47, March, 1994
- [4.8] M. W. Cresswell, J. Bonevich, T. J. Headley, R. A. Allen, L. A. Giannuzzi, S. C. Everist, R. N. Ghostagore, and P. J. Shea, "Comparison of electrical CD measurement and cross-section lattice-plane counts of sub-micrometer feature replicated in (100) silicon-on-insulator material", Metrology, Inspection, and process control in Microlithography XIV, Proceedings of the SPIE, vol. 3998, pp.74-83, 2000.
- [4.9] S. K. Jones, R. L. Van Asselt, J. C. Russ, B. W. Dudley, G. Johnson, B. Cohen, M. Schwartz, P. Besser, and P. Herman, "Comparison of SEM, confocal light microscope and electrical resistance measurements of microelectronic devices", Journal of computer assisted Microscopy, vol.2, pp.211-221, 1990.

[4.10] J.S. Villarrubia, A.E. Vladar, J. R. Lowney, M. T. Postek, R. A. Allen, M. W. Cresswell, and R. W. Ghostagore, "Linewidth measurement intercomparison on a BESOI sample", Metrology, Inspection, and process control in Microlithography XIV, Proceedings of the SPIE, vol. 3998, pp.84-95, 2000.

[4.11] S. Smith, I.A. B. Lindsay, A. J. Walton, M. W. Cresswell, L. W. Linholm, R. A. Allen, M. Fallon, A. M. Gundlach, "Analysis of current flow in mono-crystalline electrical linewidth structures", Proceedings of IEEE 1999 International Conference on Microelectronic Test Structures, Vol. 12, March, 1999.

[4.12] R. A. Allen, P. M. Troccoli, J. C. Owen III, I. E. Potzick, and L. W. Linholm, "Comparison of measured line-widths of sub-micrometer lines using optical, electrical, and SEM metrologies", Proceedings of SPIE, Integrated Circuit Metrology, Inspection, and Process Control VII, vol. 1926, pp. 34-43, 1993.

[4.13] X. Niu, N. Jakatdar, J. Bao and C. J. Spanos, "Specular spectroscopic scatterometry in DUV lithography", SRC Techcon'98.

[4.14] Electroglas manual.

[4.15] David Rodriguez, "Electrical testing of a CMOS baseline process", Master thesis. University of California, Berkeley. Memorandum No. UCB/ERL M94/63..

[4.16] Jason Cain, private communication.

Chapter 5

Full Field Lens Aberration Extraction Using Printed Patterns

Lens aberrations are becoming an increasing concern in limiting the performance of today's optical lithography systems [5.1-5.6]. In chapter 2, we explained that processing variability is attributed to several sources of variation. Among them lens aberrations are believed to be one of the major sources that cause spatial linewidth variation. As explained in section 4.2, experiments have confirmed that major variation comes from within field and lens aberration is possibly an important error source. Consequently, this chapter is devoted to lens aberration issues. In chapter 4, we described the experiment, including mask design; wafer processing, electrical test setup, measurement results, data analysis and experiment conclusion. After an introduction and a project outline in this chapter, a novel full-field lens aberration measurement technique is introduced in section 5.2. The methodology, optimization, its limitation and error analysis are explained. Finally, issues in numerical analysis which limit the performance of this method are discussed.

5.1 Introduction

As has been discussed in chapter 2, it is necessary to decompose the variability of printed patterns into relevant sources, and to accurately characterize or model them, in order to efficiently reduce the overall variation. Lens aberrations are believed to be one of most

critical error sources leading to within-field systematic spatial variations. This has been assumed in this study. Therefore, it is highly desirable to find an efficient way to extract lens aberrations. This can either be used to evaluate the quality of exposure tools or to provide the necessary information to related process control techniques to effectively reduce the error caused by lens aberrations. First of all, in production, there are usually multiple pieces of exposure equipment running simultaneously for the same product. It is very important to assign the piece of equipment with the best quality to pattern the most critical layer, such as polysilicon gate. Secondly, for any aggressive design to fully utilize the capacity of either 248 nm or 193 nm tools, resolution enhancement techniques (RET), such as double exposure, phase shift mask (PSM) and optical proximity correction (OPC), have been routinely used in the manufacturing of high performance products. Notice that all these RETs, especially PSM, are extremely sensitive to lens aberrations. The design rules have been pushed so far that lens aberrations cannot be ignored any more and they are becoming an increasingly critical factor in limiting the yield. The economic concerns regarding the cost of designing a specific mask for each stepper or scanner will eventually be balanced by the fact that the benefits from improved performance and yield will exceed the cost of this investment. This claim is based on the assumption that lens system design and quality control will not improve dramatically in near term.

The ultimate goal of this work is to identify efficient process control schemes and other techniques, such as mask-based error correction, in order to minimize the variability due to lens errors. More explicitly, once we can extract full-field lens aberration data, the position-specific mask error factor of different features can be obtained. This information can then be used to systematically pre-modify a mask design in order to compensate for the CD variation due to lens aberration signature. This sequence is illustrated in the following figure.

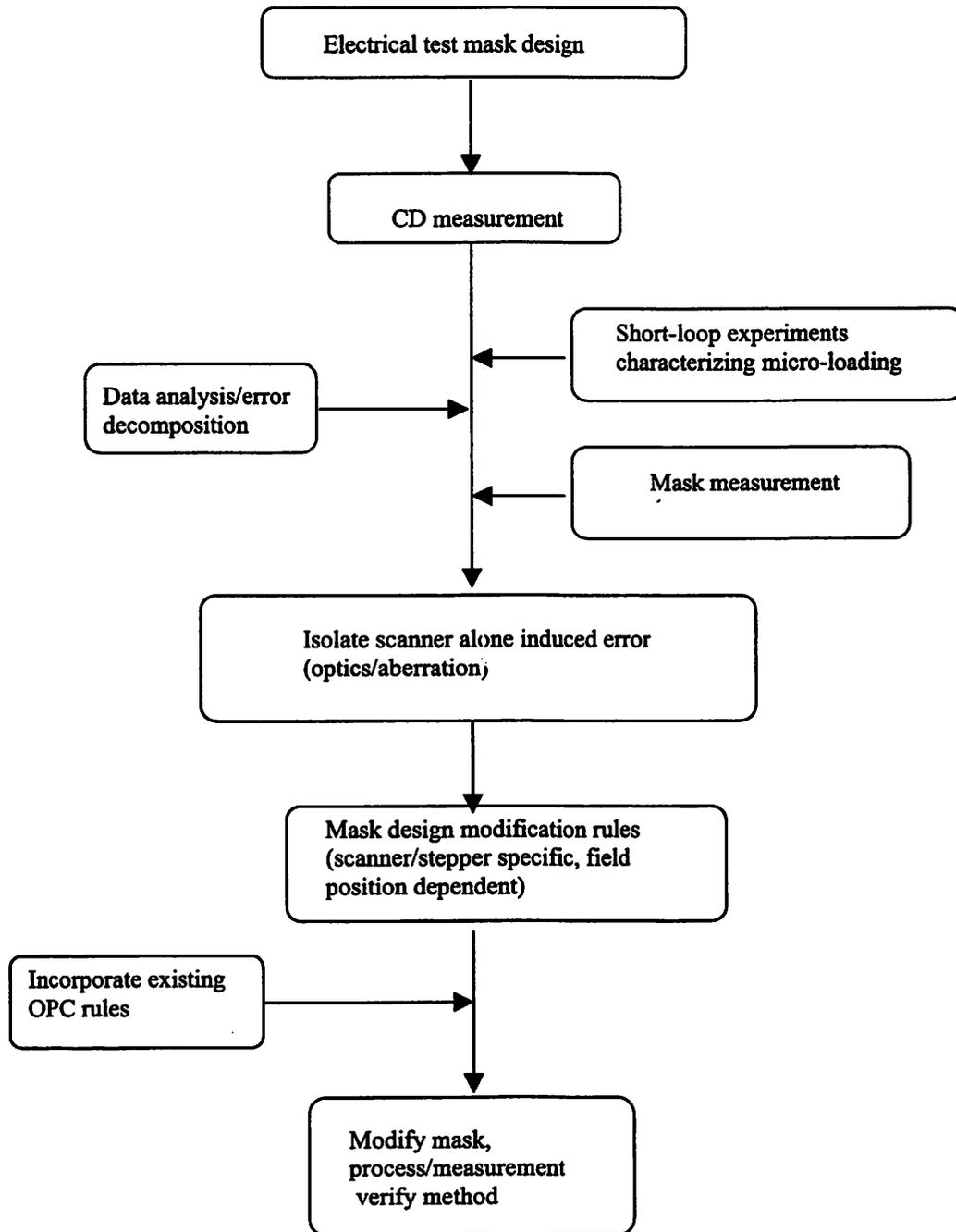


Figure 5-1 Reduction of across-field linewidth variability by compensating position-specific lens aberrations

5.2 A Simple Method to Extract Field-Dependent Lens Aberration

With chip size becoming larger and larger, the isoplanatic approximation⁴ is no longer valid, which means the aberration function of optical lithography systems is field position dependent. Although generally lens makers can measure the wavefront of the lens by through-the-lens interferometry (TTLI), the aberration signature of the whole system is usually unknown to the end user. However, this information is critical for the analysis of AFLV (across field linewidth variation). As introduced in section 2.2.2, different methods have been studied to measure residual lens aberration through analyzing the printing behavior.

In this section, we introduce a simple method to extract the field-dependent aberration using high-density electrical measurement. This method is possibly applicable to Zernike terms that are sensitive to linewidth variation.

Lens aberrations of modern optical lithography systems are generally small. Therefore, we can model CD as the function of aberration as follows

$$CD(x, y, F, E) = CD_0(x, y, F, E) + \sum_{i=1}^{\infty} Z_i(x, y) \times \frac{\partial CD}{\partial Z_i}(F, E) \quad (5-1)$$

Where CD_0 is the target linewidth, (x, y) is the field coordinates, F represents the defocus distance, while E is the exposure dose, and Z_i is the i^{th} Zernike coefficient.

The infinite Zernike polynomial series are often limited to 36 terms, since higher order aberrations have negligible influence on the image. In this study we only consider the first 31 terms, determined by the limited focus-exposure set of experimental samples.

For a given field position (x_o, y_o) , we can rewrite (5-1) as

$$\left\{ \begin{array}{l} CD(FE^{(1)}) = CD_0(FE^{(1)}) + Z_1 \times \frac{\partial CD}{\partial Z_1}(FE^{(1)}) + Z_2 \times \frac{\partial CD}{\partial Z_2}(FE^{(1)}) + \dots + Z_{31} \times \frac{\partial CD}{\partial Z_{31}}(FE^{(1)}) \\ CD(FE^{(2)}) = CD_0(FE^{(2)}) + Z_1 \times \frac{\partial CD}{\partial Z_1}(FE^{(2)}) + Z_2 \times \frac{\partial CD}{\partial Z_2}(FE^{(2)}) + \dots + Z_{31} \times \frac{\partial CD}{\partial Z_{31}}(FE^{(2)}) \\ \vdots \\ CD(FE^{(31)}) = CD_0(FE^{(31)}) + Z_1 \times \frac{\partial CD}{\partial Z_1}(FE^{(31)}) + Z_2 \times \frac{\partial CD}{\partial Z_2}(FE^{(31)}) + \dots + Z_{31} \times \frac{\partial CD}{\partial Z_{31}}(FE^{(31)}) \end{array} \right. \quad (5-2)$$

where $CD(FE^{(i)})$ is the linewidth resulting from the i^{th} focus-exposure setting, which can be measured from a focus-exposure-matrix (FEM) wafer. $CD_0(FE^{(i)})$ is the CD without the aberration, simulated under corresponding focus/exposure conditions, using reticle CD, which can be also measured by an SEM or an optical tool. The sensitivity factor $\frac{\partial CD}{\partial Z_i}$ can be conveniently obtained by calibrated simulation under various focus-

exposure settings. The performance of lithography simulation is related to the accuracy of the resist model to some extent. The calibration of resist parameters is crucial for this analysis. This issue will be addressed in the later part of the thesis. Two FEM wafers have been processed and are being electrically measured. At the same time, the reticle CD has also been extensively measured. Thus, in equation (5-2), we have 31 relations and 31 unknowns. The Zernike coefficients Z_1 to Z_{31} can be uniquely solved, provided that the resulting linear system has the required number of degrees of freedom.

By repeating the above procedure through all the positions in the field, we can derive the aberration distribution across the field.

⁴ The approximation holds only when the object is small so that it falls in isoplanatic region. Isoplanatic region (or space-invariant region) is defined as a region, in which point source object changes only in location, not in functional form when point source moves in this region.

5.3 Sensitivity of Linewidth Variation to Individual Zernike Terms

A set of typical Zernike coefficient ranges has been used to simulate the CD sensitivity to individual aberrations. Since some Zernike aberrations (such as coma) predominately result in an image shift, which mostly causes overlay problems, they do not affect linewidth variation. Also features with particular orientations may not be influenced by a given aberration. They cannot be determined by the method discussed above. So it is necessary to find what Zernike terms can or cannot be extracted by this method through simulating their sensitivity. Otherwise, as will be discussed in later sections, the insensitive terms will cause singularity problem in linear equations and lead to unstable results. We can effectively set all insensitive Zernikes to zero, and thus solve the linear system for only the remaining coefficients.

The following results are based on simulation. Aerial image CD instead of resist CD is used to remove the effect of photoresist and uncertainty in resist model parameters. The numerical aperture is 0.57 and the wavelength is 248nm. The partial coherence is set at 0.3 to enhance the response of aberration. Mask patterns are line/space features, with 200nm linewidth and different pitches of 600nm and 2000nm, for isolated and nested features respectively. Aerial image threshold is adjusted so that under defocus of $-0.2 \mu\text{m}$ the aerial image CD is equal to the mask CD of 200 nm for isolated line. For nested lines, the image threshold is set separately to facilitate the analysis. As an example, figures 5-2 to 5-7 show how CD varies with individual Zernikes Z_1 to Z_{36} for isolated lines (a) and (b) dense lines.

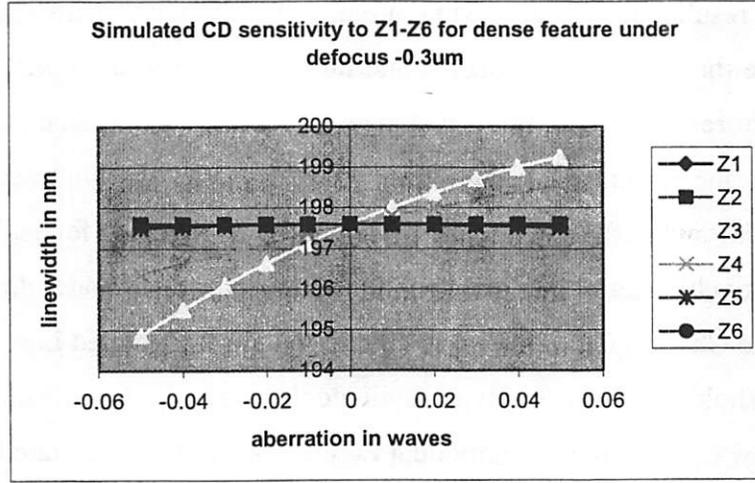
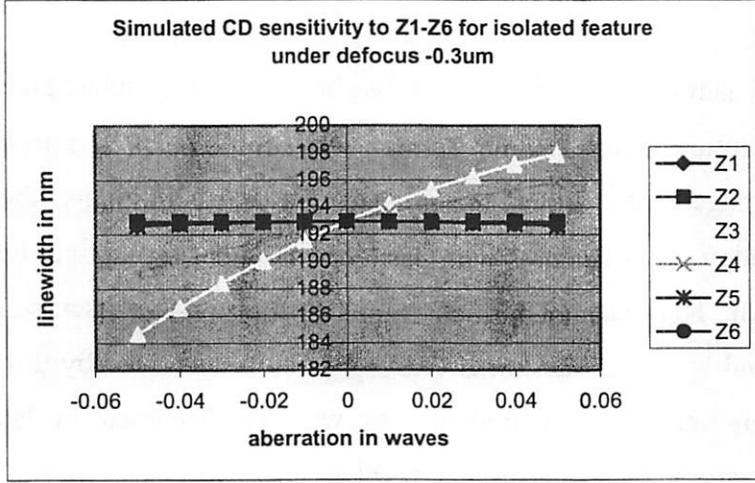


Figure 5-2 Simulated CD sensitivity to Z₁-Z₆ for isolated and dense lines

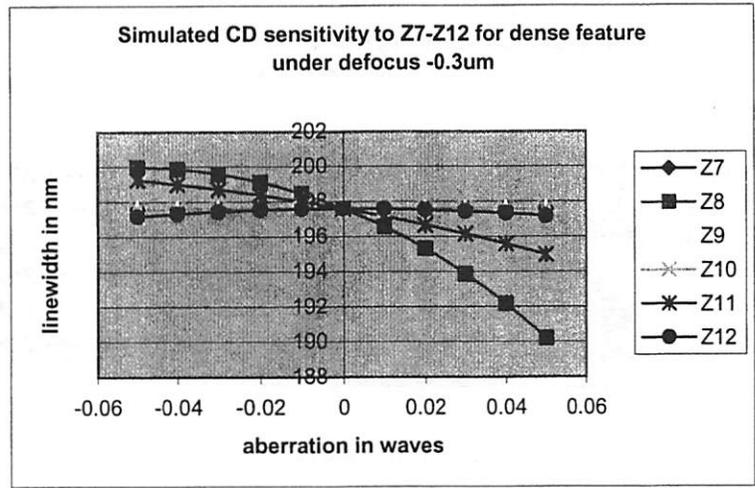
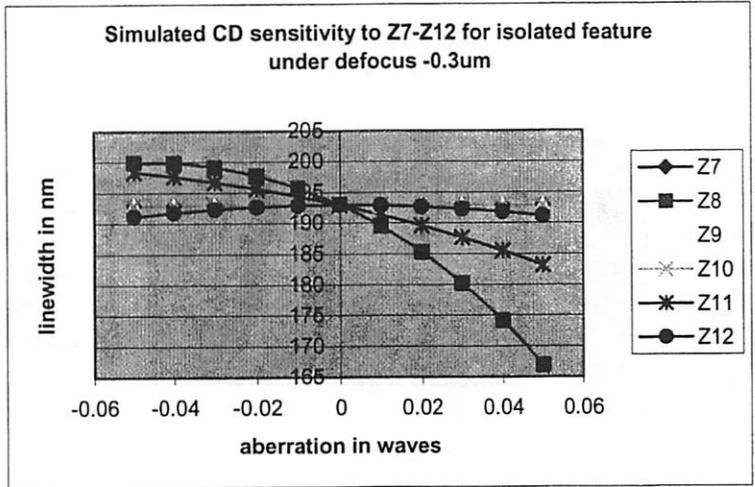


Figure 5-3 Simulated CD sensitivity to Z₇-Z₁₂ for isolated and dense lines

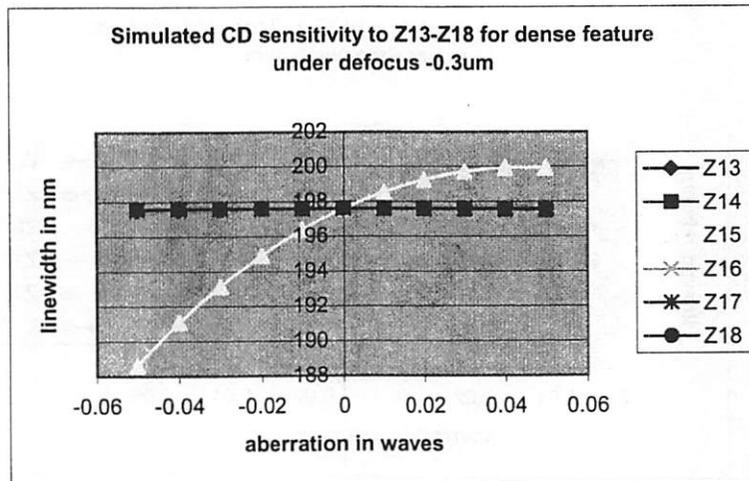
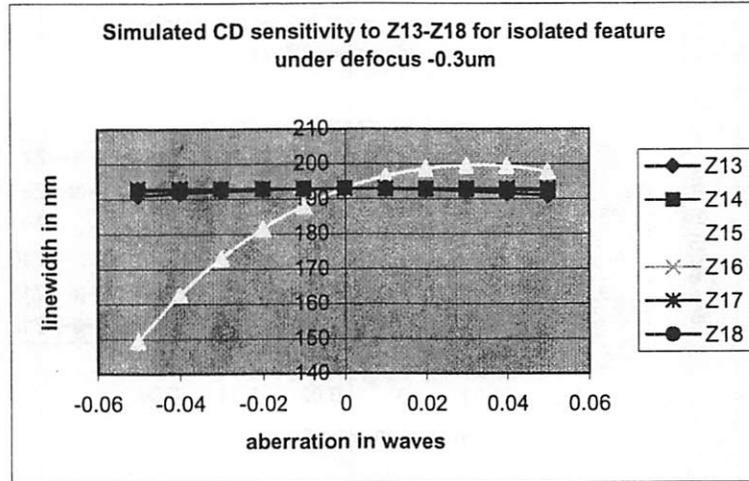


Figure 5-4 Simulated CD sensitivity to Z₁₃-Z₁₈ for isolated and dense lines

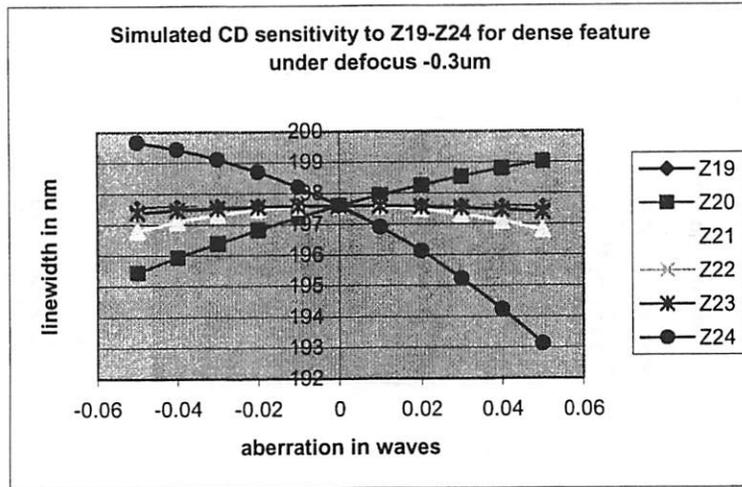
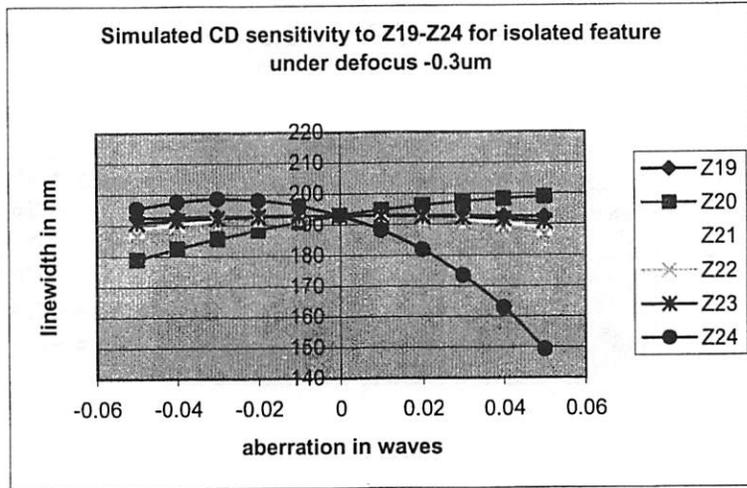


Figure 5-5 Simulated CD sensitivity to Z₁₉-Z₂₄ for isolated and dense lines

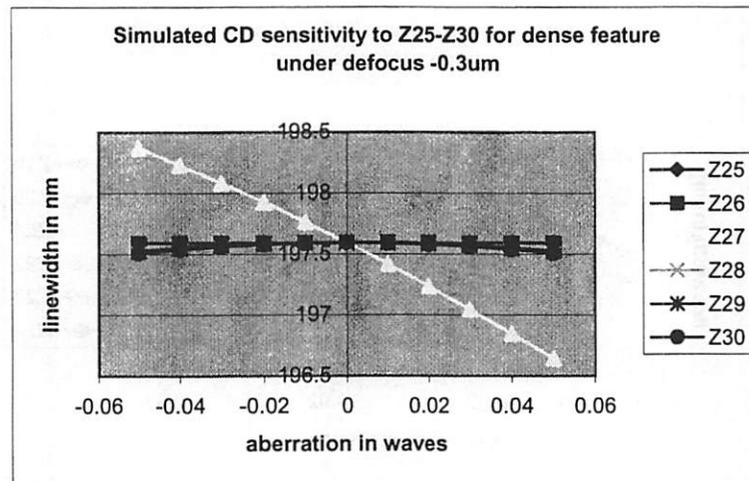
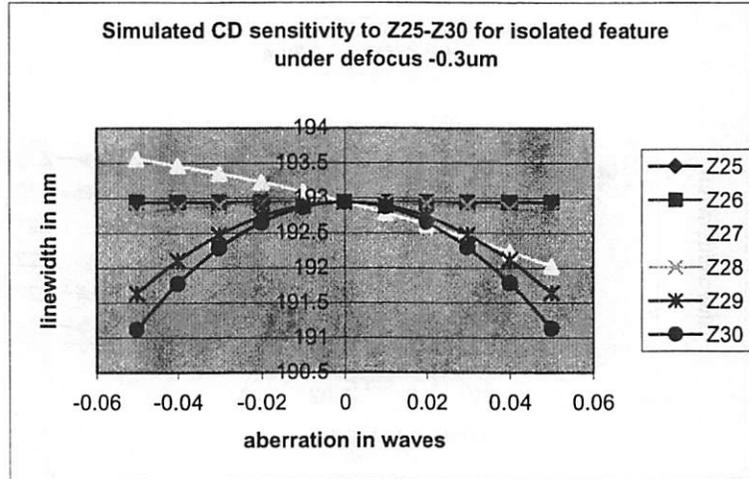


Figure 5-6 Simulated CD sensitivity to Z_{25} - Z_{30} for isolated and dense lines

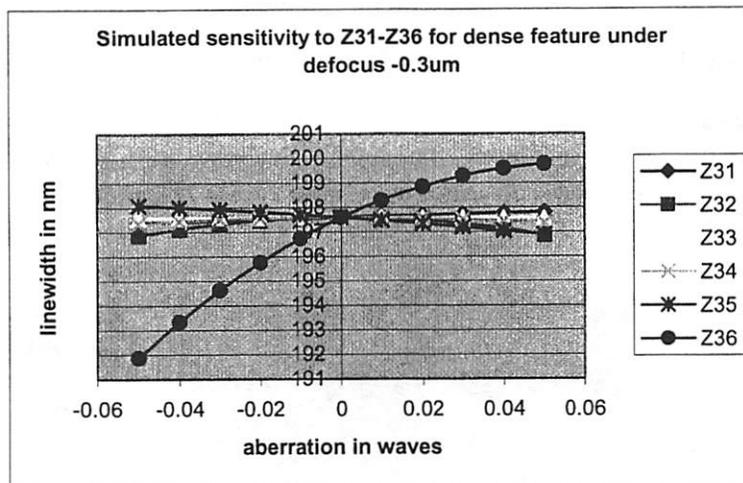
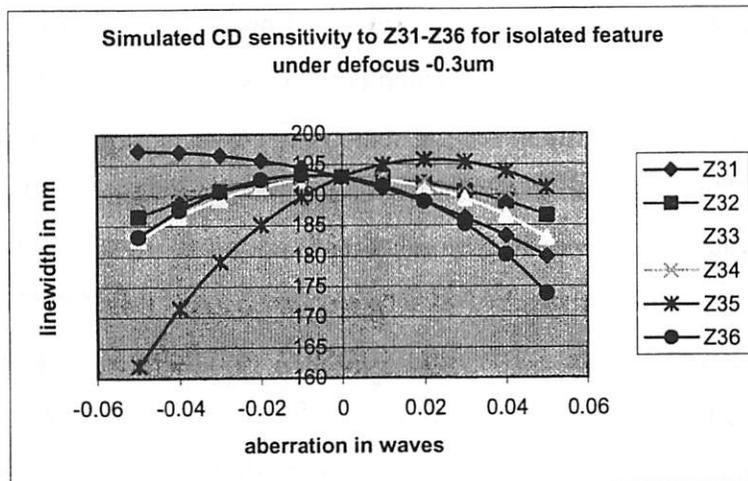


Figure 5-7 Simulated CD sensitivity to Z₃₁-Z₃₆ for isolated and dense lines

It can be seen that some aberration terms have a strong influence on the CD change and it is almost linear over the -0.05λ to 0.05λ range for isolated features as well as dense features under non-zero defocus settings. So the linear assumption in equation (5-1) is valid when lens aberrations are small, which is the case for most 248 nm systems. If exact zero aberration is used in simulation, the sensitivity curves will have a dip around zero for most Zernikes. This discontinuity at zero is a simulation artifact. Prolith uses different algorithms in simulation depending on whether aberrations are specified. Applying a very small value of aberration instead of exact zero can bypass this problem. The aberration range in figure 5-2 to figure 5-7 is from -0.0501λ to 0.0499λ in steps of 0.01λ . For all aberration terms, the slope of sensitivity curves for dense features are generally smaller than isolated features. This suggests that dense lines are less sensitive to lens aberration in terms of linewidth variation.

Figure 5-8 shows the simulated CD sensitivity of isolated lines to Z_8 spherical under different defocus settings. Apparently, all the aberration sensitivity factors are different for each different defocus and exposure dose. A careful calibration of image intensity threshold and resist parameters is important for this analysis.

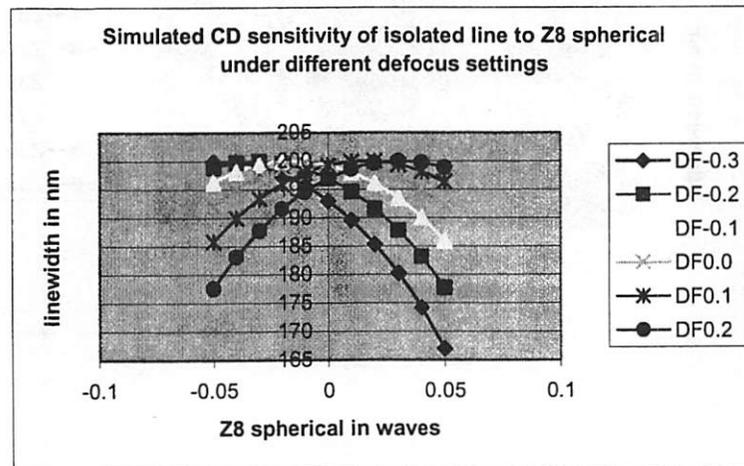


Figure 5-8 Simulated CD sensitivity of isolated line to Z_8 spherical under different defocus

It has been found that some aberration components have a very small impact on linewidth under a large range of focus-exposure settings. Those components that have a strong influence are almost linear over the practical lens aberration range. From the simulation, odd terms and even terms do not present a dramatic difference in CD influence.

Table 5-1 lists the CD variation range when Zernike changes within $\pm 0.05\lambda$. The entries of the table are the linewidth variation range in nm.

Table 5-1 CD variation range in nm for individual Zernikes changing within $\pm 0.05\lambda$

	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅	Z ₆	Z ₇	Z ₈	Z ₉	Z ₁₀	Z ₁₁	Z ₁₂
isolated	0.01	0.00	13.24	6.55	0.30	0.22	0.06	32.97	0.04	0.06	15.21	1.86
dense	0.01	0.00	4.39	2.18	0.08	0.01	0.01	9.78	0.01	0.01	4.27	0.44

	Z ₁₃	Z ₁₄	Z ₁₅	Z ₁₆	Z ₁₇	Z ₁₈	Z ₁₉	Z ₂₀	Z ₂₁	Z ₂₂	Z ₂₃	Z ₂₄
isolated	2.07	0.65	50.36	0.41	0.01	0.40	0.60	20.05	4.31	6.39	2.18	49.24
dense	0.05	0.08	11.32	0.47	0.01	0.02	0.06	3.59	0.80	0.08	0.21	6.54

	Z ₂₅	Z ₂₆	Z ₂₇	Z ₂₈	Z ₂₉	Z ₃₀	Z ₃₁	Z ₃₂	Z ₃₃	Z ₃₄	Z ₃₅	Z ₃₆
isolated	0.00	0.00	1.53	0.15	1.31	1.82	17.40	6.41	10.12	3.92	33.76	19.46
dense	0.00	0.00	1.73	0.09	0.07	0.09	0.45	0.74	0.03	0.25	1.20	7.91

Among all first 36 Zernikes studied, the following terms have large (greater than 10nm) influence on linewidth variations.

Table 5-2 Zernike terms of high sensitivity with CD variation range larger than 10 nm within $\pm 0.05\lambda$

Z_3	$2\rho^2-1$
Z_8	$6\rho^4-6\rho^2+1$
Z_{11}	$(4\rho^2-3)\rho^2\cos 2\theta$
Z_{15}	$20\rho^6-30\rho^4+12\rho^2-1$
Z_{20}	$(15\rho^4-20\rho^2+6)\rho^2\cos 2\theta$
Z_{24}	$70\rho^8-140\rho^6+90\rho^4-20\rho^2+1$
Z_{31}	$(56\rho^6-105\rho^4+60\rho^2-10)\rho^2\cos 2\theta$
Z_{33}	$(126\rho^8-280\rho^6+210\rho^4-60\rho^2+5)\rho\cos\theta$
Z_{35}	$252\rho^{10}-630\rho^8+560\rho^6-210\rho^4+30\rho^2-1$
Z_{36}	$924\rho^{12}-2772\rho^{10}+3150\rho^8-1680\rho^6+420\rho^4-42\rho^2+1$

These highly sensitive Zernike terms must be carefully balanced and kept small in lens design. For those terms with CD variation range less than 5nm in table 5-2, the extraction method explained above may not be able to capture them, considering the metrology error and resist model uncertainty. 5nm variation range for isolated line would be an approximate criterion to choose Zernike terms that can be extracted by this method.

The coefficients selected by this criterion can be extracted using the method explained above, while the rest of the coefficients, since they do not influence CD, are simply regarded as zero. This simplification is not important since our main purpose is to

characterize the influence of lens aberration on AFLV. With a reduced number of unknowns, equation (5-2) becomes an over-determined system and least square fitting is used to calculate the Zernikes. Furthermore, equation (5-2) can be expanded at the initially extracted Zernikes instead of zero. Thus, non-linear optimization can be employed for the best fit.

Some aberration terms are strongly rotationally asymmetric, such as Z_4 90° astigmatism, Z_7 y-coma. Their impact on printed feature is dependent on the orientation of the features.

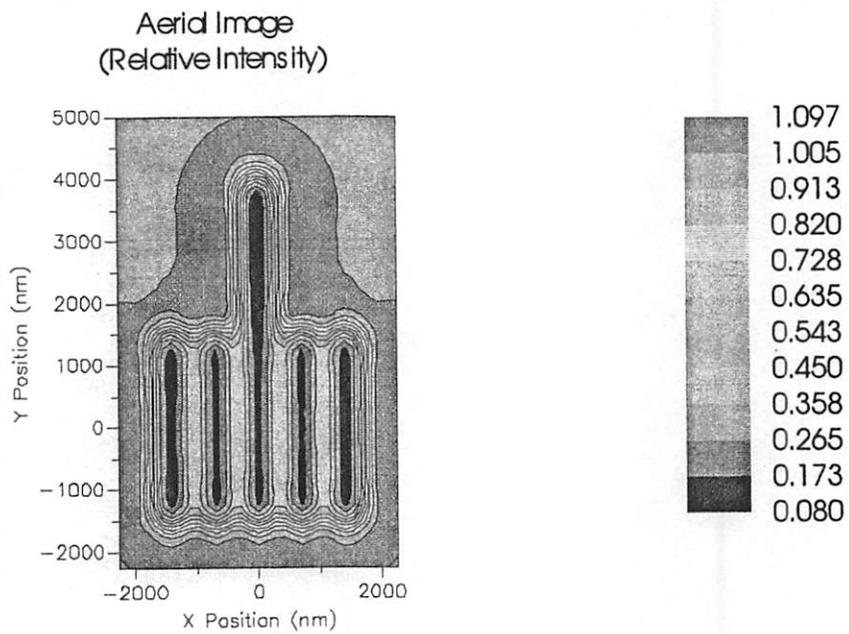
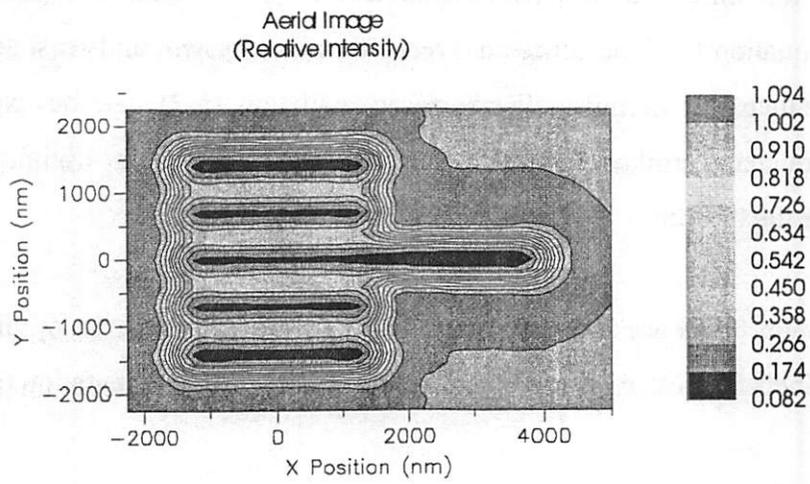


Figure 5-9 Relation of aberration orientation and pattern orientation (an example of 90° astigmatism)

Figure 5-9 is the simulated aerial image of horizontal and vertical 5-bar structures under 0.05λ 90° astigmatism. The saddle shape of the intensity distribution is apparent for

horizontal lines; vertical lines do not show this shape. 45° astigmatism (Z_5) shows equal impact on vertical line and horizontal line since the orientation of Z_5 is diagonal.

As an example, the extracted Zernike coefficients are illustrated in figure 5-10 to figure 5-11, and are shown as the field map for 19×24 positions. These values are the amplitudes of the corresponding function of the Zernike terms. Figure 5-12 shows aberrated wavefront across the exit pupil for two points at field center and field edge respectively using extracted Petzval, astigmatism, spherical aberrations and Z_{11} . An optimized choice of the combination of Zernike terms of interest should be considered for accurate extraction.

We note that the field map of extracted Zernikes is quite noisy. Data at many field points exceeds the reasonable level of state-of-art exposure tools, which is typically within $\pm 0.02\lambda$. The obtained results do not capture the reality of aberration level of the equipment under study.

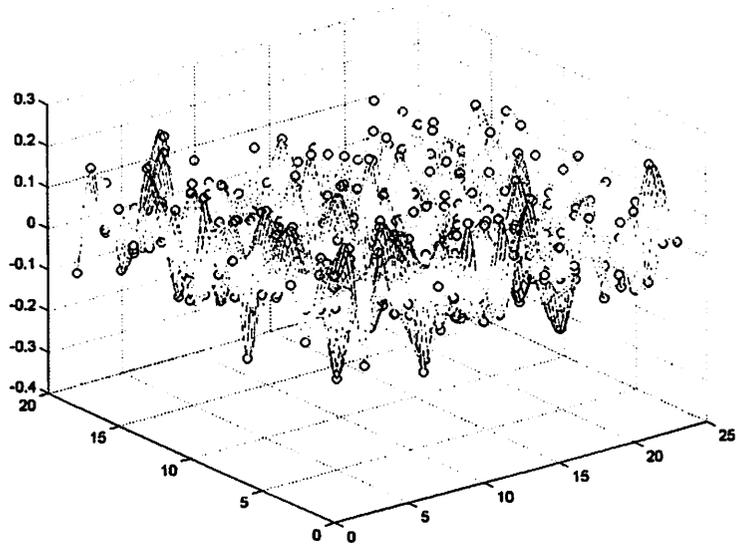


Figure 5-10 Field map of extracted Z_4 primary 90° astigmatism in λ

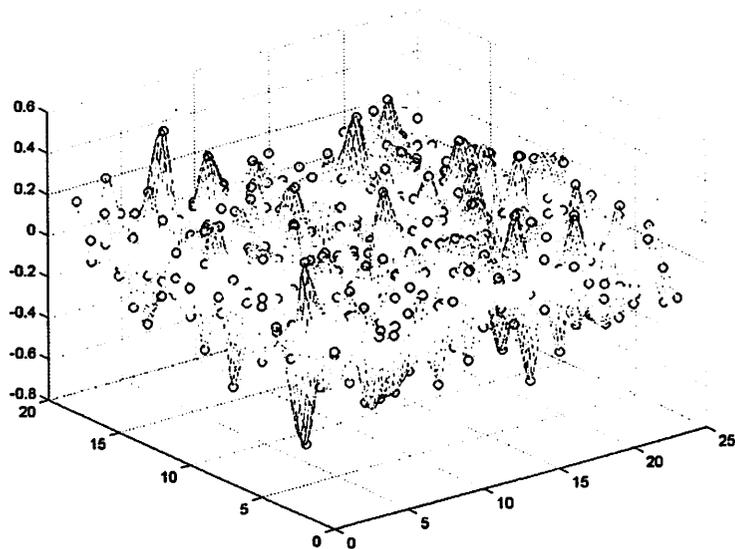


Figure 5-11 Field map of extracted Z_{11} $((4\rho^2-3)\rho^2\cos 2\theta)$ in λ

OPD

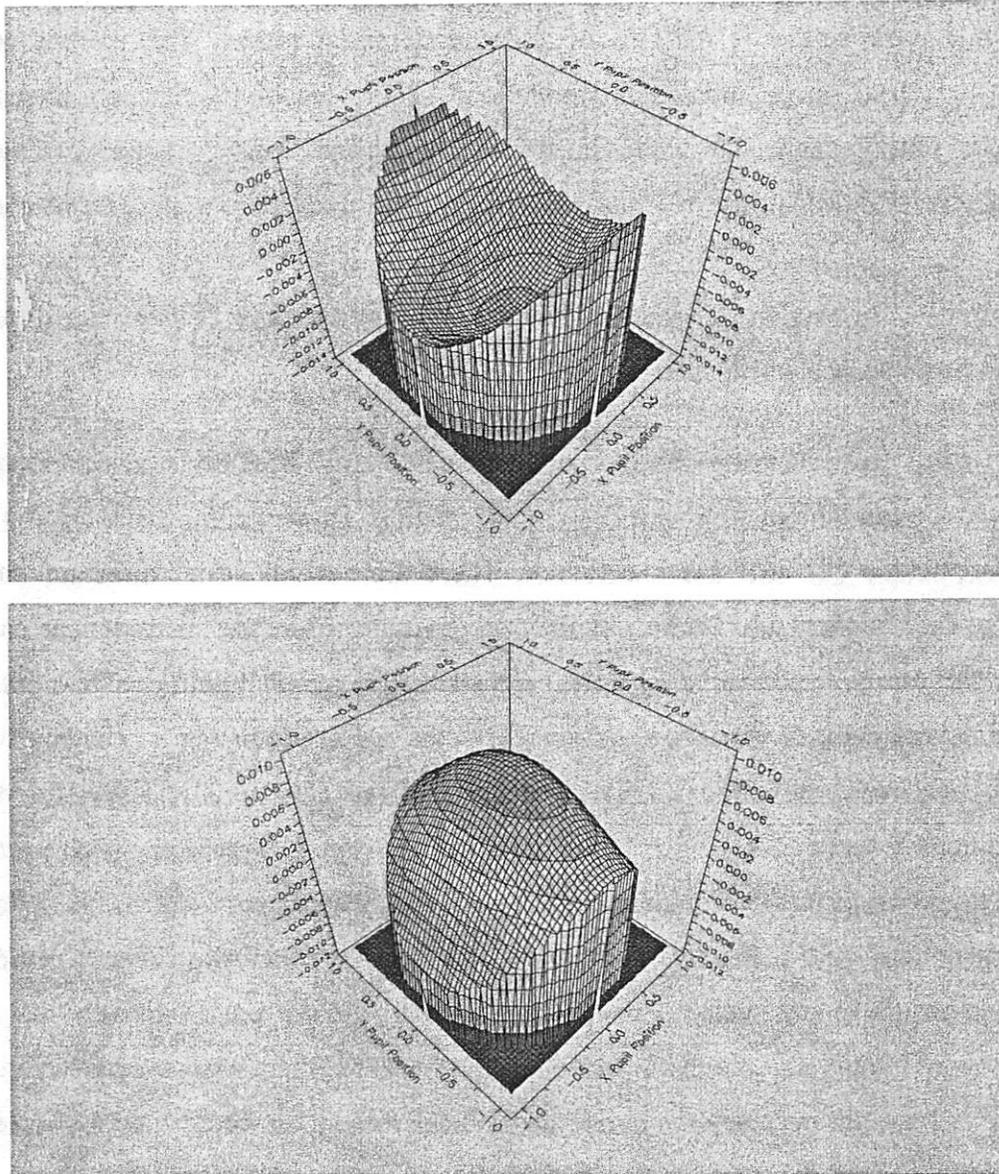


Figure 5-12 Comparison of aberrated wavefronts across exit pupil for two different field positions

5.4 Error Analysis of the Extraction Method

There are four factors contributing to the error in extracted Zernikes using this method. The first error source is the lack of an accurate resist model. This method relies on extensive aerial image and resists simulation. All the sensitivity factors under different process settings and CD value in diffraction-limited condition are obtained from simulation. The reliability of resist model is the key to ensure the correctness of these data. The resist model provided by resist vendor is not accurate enough to fit the experiments. In section 5.7, multi-dimensional optimization technique to extract resist model parameters is introduced.

The second error factor is the measurement deviations. The mask measurement using optical tool can be assumed with fairly good quality. However the electrical test used to measure wafer CD still has uncertainty. There are several issues affecting the e-test results. One is the Joule heating of the test structure when the measurement current is high. This has been addressed by using high-resolution source/monitor unit. Another issue is the contact between the probe and the pad of the test structure. The mechanical portion of the e-test setup is rather old. The contact is controlled manually. The e-test structure uses single layer of poly with thickness of about 200 nm. It is possible to scratch through with the probe or in other cases the pad is not touched properly. Figure 5-13 shows the CD field map of FEM wafer with focus $-0.1\ \mu\text{m}$ and exposure dose $21\ \text{mJ}/\text{cm}^2$ on focus/exposure matrix wafer. This plot has already filtered out 18 outliers with reading larger than 230nm and one point less than 210 nm. Since there is no average step in FEM wafer to filter out measurement noise, as has been done in uniform wafer shown in figure 4-15b, the measurement error will propagate into the noise of extracted Zernikes.

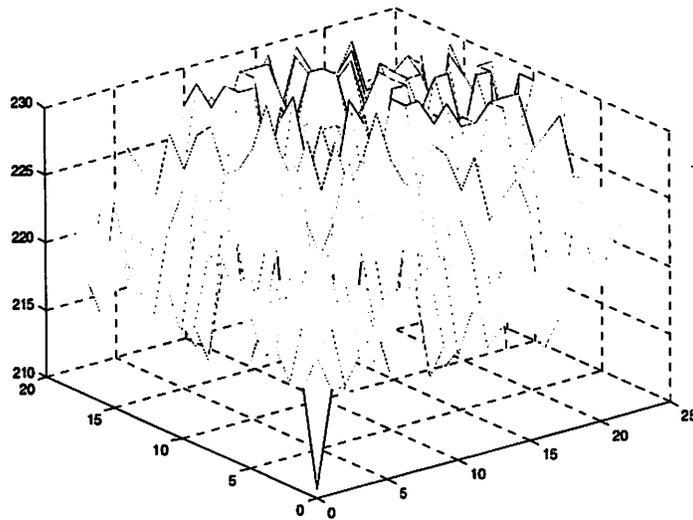


Figure 5-13 CD field map measured using electrical test for chip corresponding to focus -0.1 um and exposure 21 mJ on FEM wafer

The third source that causes the error in extracted aberrations is the etching bias. The resist pattern is more directly related to optics. However, since e-test is only applicable to poly pattern, a precise characterization of the relation of resist CD and poly CD is required. Generally, a simple “bias” cannot accurately model the relation between resist CD and poly CD. It is the shape of resist pattern that determines the final poly CD. Figure 5-14 illustrates the relation of poly CD (FI) and resist CD, measured from uniform wafer at different locations. R^2 value is only 0.49 with the average bias of 10nm and standard deviation of bias (1σ) 3nm.

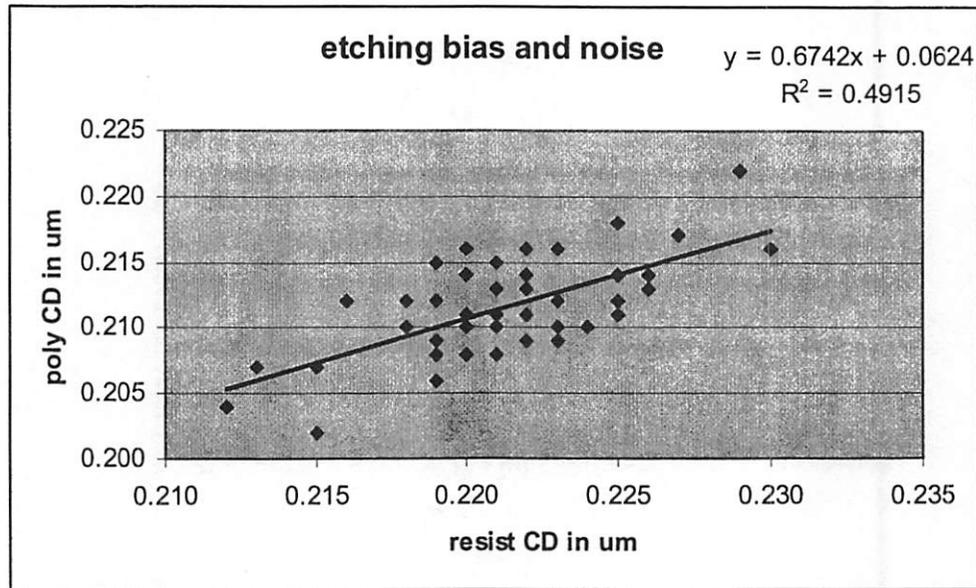


Figure 5-14 Etching bias and noise measured from uniform wafer

The last error source is the algorithm itself. It is a convenient choice to use least square fitting to numerically solve Zernike terms. If the system is ill-conditioned or even rank-deficient, least square fitting will give unstable results. However other methods, such as QR decomposition or singular value decomposition, may be of use, even if they come at considerable computational cost. This issue is addressed in section 5.8. A best combination of extractable Zernikes is the key to ensure the condition number of the linear system to be sufficiently small. Some even term Zernikes, such as defocus and spherical, have parabolic instead of linear dependency to linewidth when nominal defocus is small, so the linearization assumption will cause instability of the extracted results. In addition to linear terms, an expansion with high order terms would increase the performance of this method. Furthermore, the underlying physical compounding among different Zernikes is not taken into account. The additive model as expressed in equation (5-1) may have large error if interaction between Zernikes is not negligible. Another possible improvement is to take into account the orientation dependency of certain

Zernikes. It is a better approach to design patterns with different orientations and pitches to enhance the response of these Zernike terms.

The following analysis assumes that the third factor (etch bias) and the fourth factor (algorithm error) are not very significant. Then the only sources of extraction error are metrology uncertainty and simulation deviation.

The newly upgraded autoprobe station at the Berkeley Microfabrication Laboratory has linewidth measurement precision of about 0.3nm and the simulation error is less than 8nm. From equation (5-2), without iterated optimization, calculated Zernikes are expressed as

$$\tilde{Z} = (\tilde{G}'\tilde{G})^{-1}\tilde{G}'(\tilde{\alpha} - \tilde{\beta}) \quad (5-3)$$

where $\tilde{\alpha} = WaferCD|_{m \times 1}$ is the measured CD for m different focus/exposure settings;

$\tilde{\beta} = CD_o(\tilde{Z} = \tilde{0})|_{m \times 1}$ is the simulated CD from a perfect lens.

$$\tilde{G} = \begin{bmatrix} \frac{\partial CD}{\partial Z_1}|_{FE^{(1)}} & \cdots & \frac{\partial CD}{\partial Z_n}|_{FE^{(1)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial CD}{\partial Z_1}|_{FE^{(m)}} & \cdots & \frac{\partial CD}{\partial Z_n}|_{FE^{(m)}} \end{bmatrix} \quad (5-4)$$

is the sensitivity factor matrix of individual aberration terms for different focus/exposure settings.

Thus, i^{th} Zernike coefficient is

$$Z_i = \tilde{A}_i(\tilde{\alpha} - \tilde{\beta}) \quad (5-5)$$

where $A = (\tilde{G}'\tilde{G})^{-1}\tilde{G}$.

The error associated with the extracted coefficients is

$$\delta Z_i = \delta A_i(\tilde{\alpha} - \tilde{\beta}) + A_i(\delta\tilde{\alpha} - \delta\tilde{\beta}) \quad (5-6)$$

The expected value of the squared error is thus

$$E(\delta Z_i^2) = (\tilde{\alpha} - \tilde{\beta})' E(\delta\tilde{A}_i' \delta\tilde{A}_i) (\tilde{\alpha} - \tilde{\beta}) + A_i E[(\delta\tilde{\alpha} - \delta\tilde{\beta})(\delta\tilde{\alpha} - \delta\tilde{\beta})'] A_i' \approx (0.031\lambda)^2 \quad (5-7)$$

where

$$E[(\delta\tilde{\alpha} - \delta\tilde{\beta})(\delta\tilde{\alpha} - \delta\tilde{\beta})'] = \begin{bmatrix} \text{var}(\alpha_1 - \beta_1) & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & \cdots & \text{var}(\alpha_m - \beta_m) \end{bmatrix} = 65I \quad (5-8)$$

and $E(\delta\tilde{A}_i' \delta\tilde{A}_i) \approx 1.18 \times 10^{-5} I$ from a Monte Carlo run.

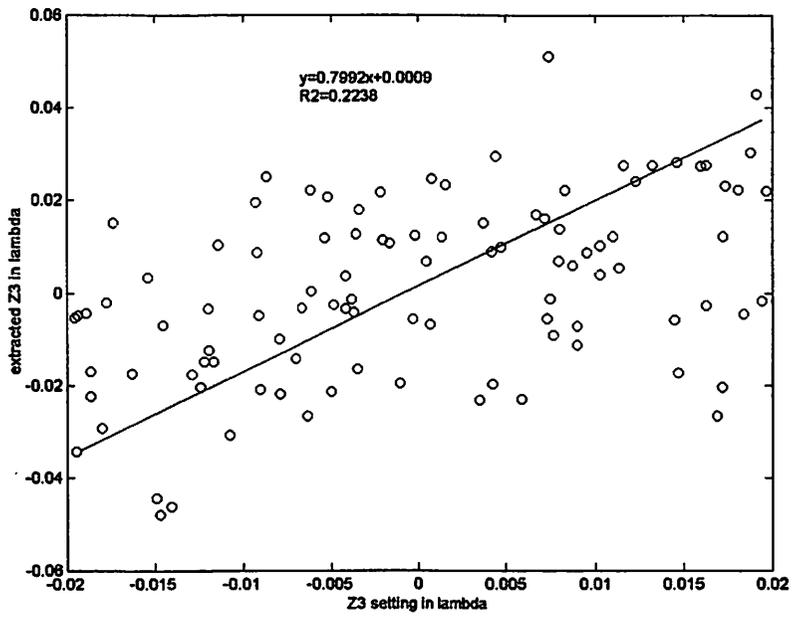
Therefore, the average error of the extracted Zernike coefficients is about 0.031 wavelengths, and that, unfortunately, is almost as large as the values that we are trying to extract. If the etching bias and algorithm error are taken into account, the estimated error will be larger than this number. Still, this method might hold considerable promise if one can use high-end scatterometry for measuring resist profiles, and if much more accurate simulation models may be obtained. We examine this hypothesis next by performing a simulation-based study.

5.5 Simulation Based Verification of the Extraction Method

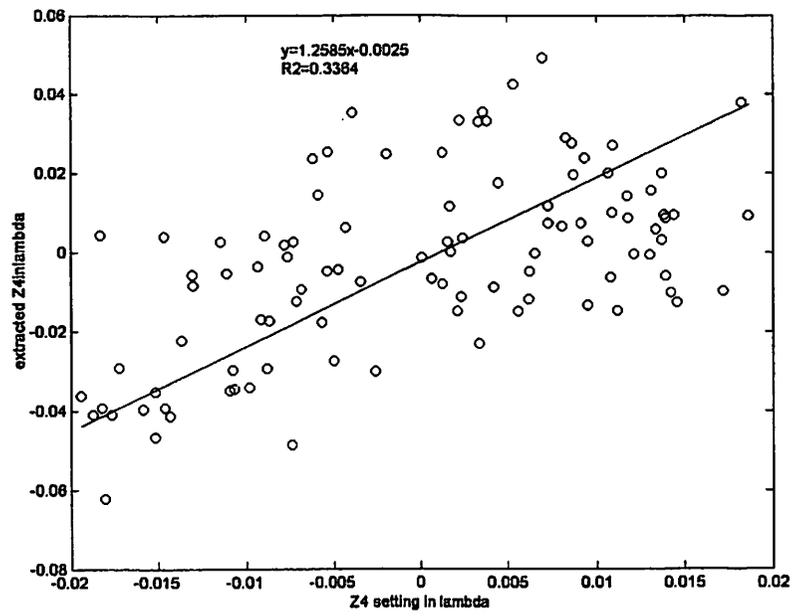
As discussed in the previous section, there are four major sources contributing to the error in extracted Zernikes. It is meaningful to evaluate the algorithm itself by ignoring the effect from resist modeling and etching noise. In this section, a simulation-based verification is conducted by generating artificial data sets and by extracting the aberrations.

We still focus on the first 4 sensitive aberration terms Z_3 , Z_4 , Z_8 and Z_{11} . An artificial data set with 100 data points is randomly generated. At each data point, the aberration level for these four Zernike terms is within $\pm 0.02\lambda$. The aerial image CD is then simulated at all these 100 data points and at four different defocus conditions and partial coherences. The mask pattern is a one-dimensional isolated line with linewidth 200 nm and pitch 2000nm. Ideally it is helpful to include features with different orientations. However Prolith is not able to specify the orientation of one-dimensional line/space patterns. The constraints of 2D simulation are that the speed is slow and that there is no function to automatically read out simulated output. The aerial image threshold is set at 0.37 so that at condition of defocus $-0.2\mu\text{m}$ and zero aberration, the aerial image CD is equal to target 200nm. Then a normal distributed noise with standard deviation of 3 nm is added to these 400 calculated CD points to simulate metrology error. Based on this generated CD points, the extracting method is used to calculate the four aberration terms at the 100 points.

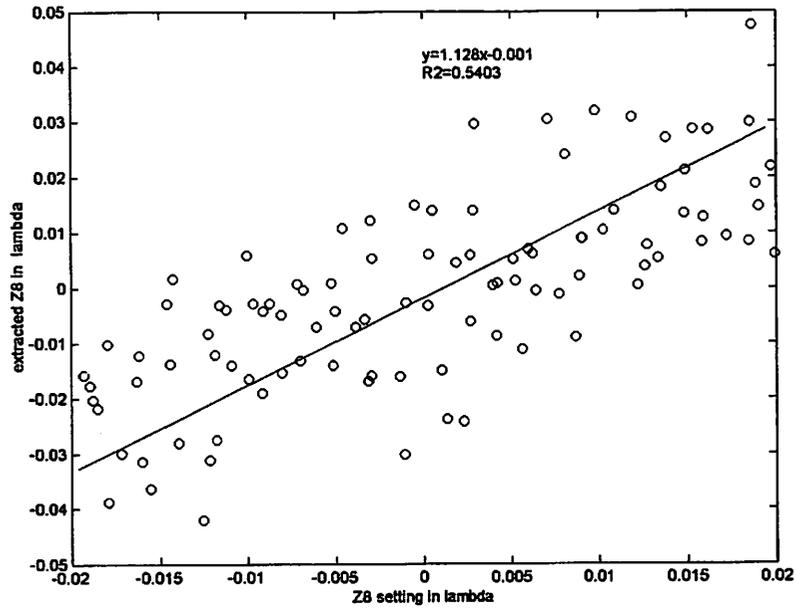
Figure 5-15 shows the comparison of extracted aberration versus the value previously set.



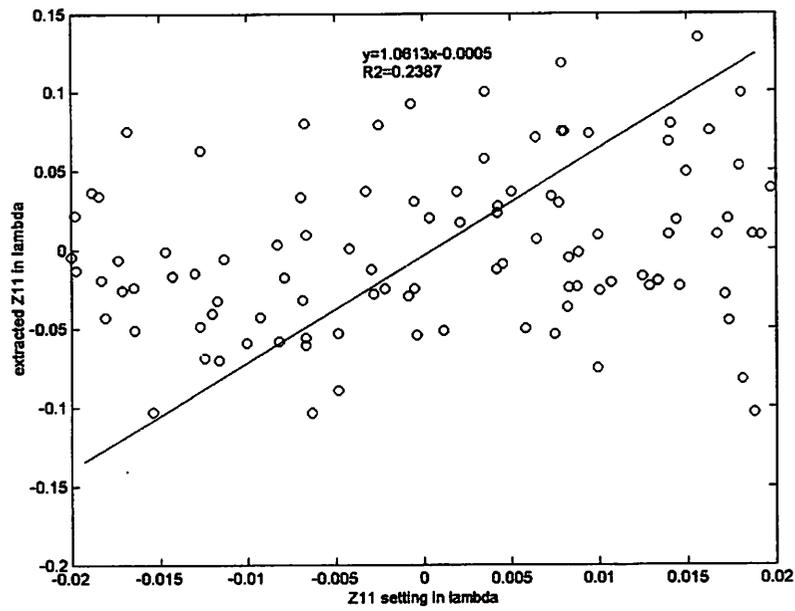
(a)



(b)



(c)



(d)

Figure 5-15 Simulation based verification of Zernike extraction method

The simulation results show that the extracted Zernikes agree with the pre-generated value well but still have relatively large error. One sigma 3nm metrology error is estimated from standard deviation value of poly CD bias versus resist CD (figure 5-14). If resist CD, instead of poly CD, is measured directly using scatterometry and if we assume the metrology error of scatterometry is less than 3 nm, this method should give a better performance. However, a data set with reduced metrology error of 1-sigma of 0.5nm does not improve the extraction accuracy significantly. This suggests that other factors may play important roles in limiting the performance of this method. These factors include choice of combination of Zernike terms and high order approximation as well as accuracy of simulation software. Furthermore, this analysis is conducted by assuming a linear relationship between the Zernikes and the CD, and without iterating in order to capture the non-linearities. If we use iteration around the Taylor expansion terms at the intermediately generated value and apply more advanced techniques, such as singular value decomposition, to solve the linear systems, we can anticipate better extraction results.

5.6 Improvement Using Scatterometry Metrology

Scatterometry is the proven technology to measure profile of thin film of various materials. Similar to equation (5-1), we can model the profile of test pattern as the function of Zernikes as follows,

$$\left\{ \begin{array}{l} X_1(FE^{(1)}) = X_1^0(FE^{(1)}) + Z_1 \times \frac{\partial X_1}{\partial Z_1}(FE^{(1)}) + Z_2 \times \frac{\partial X_1}{\partial Z_2}(FE^{(1)}) + \dots + Z_{31} \times \frac{\partial X_1}{\partial Z_{31}}(FE^{(1)}) \\ X_2(FE^{(2)}) = X_2^0(FE^{(2)}) + Z_1 \times \frac{\partial X_2}{\partial Z_1}(FE^{(2)}) + Z_2 \times \frac{\partial X_2}{\partial Z_2}(FE^{(2)}) + \dots + Z_{31} \times \frac{\partial X_2}{\partial Z_{31}}(FE^{(2)}) \\ \vdots \\ Y(FE^{(31)}) = Y^0(FE^{(31)}) + Z_1 \times \frac{\partial Y}{\partial Z_1}(FE^{(31)}) + Z_2 \times \frac{\partial Y}{\partial Z_2}(FE^{(31)}) + \dots + Z_{31} \times \frac{\partial Y}{\partial Z_{31}}(FE^{(31)}) \end{array} \right. (5-9)$$

where $X_1, X_2 \dots Y$ are the profile parameters shown in figure 5-16.

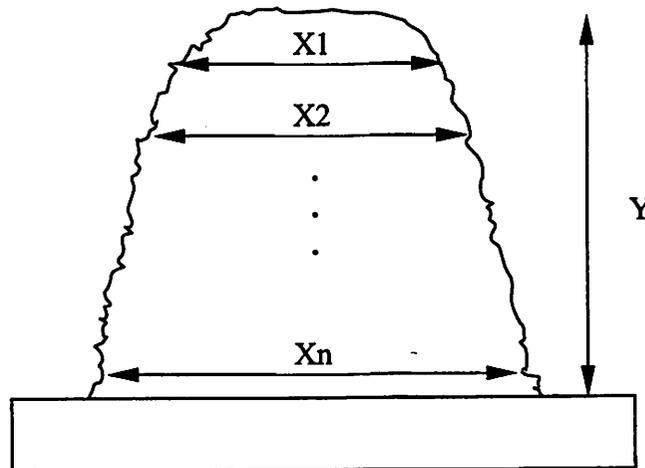


Figure 5-16 Extracting aberration by measuring profile of resist pattern

Since a full profile gives more information about the optics compared to only linewidth, more aberration terms could be extracted. Using scatterometry will improve the metrology uncertainty, the second error source in section 5.4. It will also get rid of the problem of the etching bias, the third error source, since we can directly measure photoresist profile. However, the availability of an accurate resist model is still a critical issue that must be resolved before this method can be fully tested.

5.7 Resist Parameter Optimization

As described in section 5.4, the accuracy of the lens aberration extraction technique is dependent on the accuracy of the resist model parameters. Due to the complexity of mechanisms in DUV resist and in extracting parameters for their models, the models provided by photoresist vendors are almost always far off from the real values. The two sets of resist parameters available from the resist manufacturer used for either dense or semi-dense patterns usually cannot be directly applied to an arbitrary resist pattern. Therefore, it is highly desirable to obtain the resist model for the specific pattern used in the experiment with reasonable accuracy, in order to assure the correctness of extracted lens aberration coefficients.

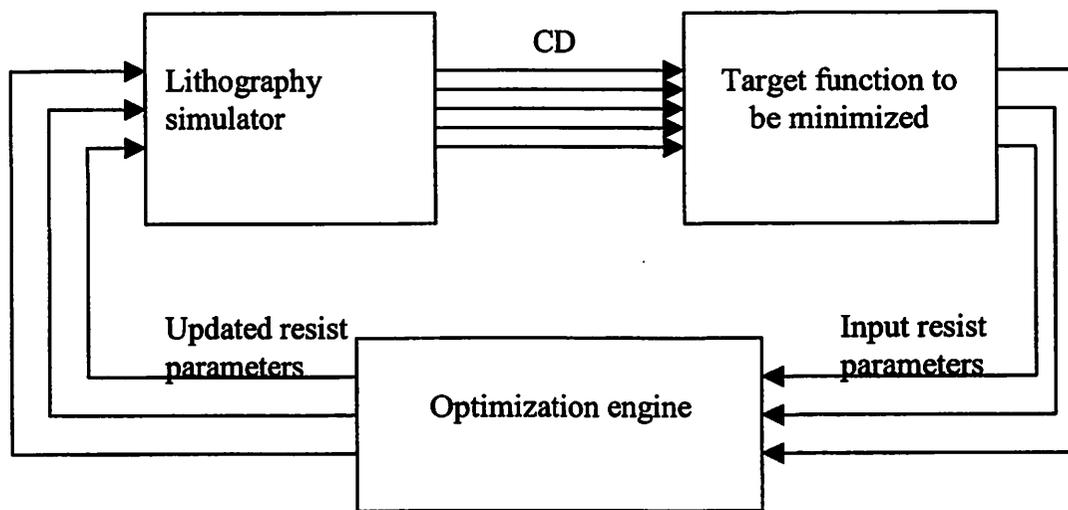


Figure 5-17 Resist parameter optimization

Figure 5-17 depicts the optimization loop. The electrically measured CD data for the central site of each die on the focus-exposure matrix wafer is used for optimizing the UV110 resist parameters. The objective function is defined as the sum of the squared errors of poly CD versus simulated CD for the current resist parameters.

$$SSE(a, b, c, \dots) = \sum_{F, E} \left[\text{measuredCD}(F, E) - \text{simulatedCD}(F, E) \Big|_{(a, b, c, \dots)} \right]^2 \quad (5-10)$$

Here, a, b, c, \dots are the parameters need to be optimized and F, E are focus and exposure dose, respectively.

Alternatively, photoresist CD measured by SEM may be used for the extraction. However, several factors need to be taken into account in order to choose resist CD data for optimization. First, photoresist CD is seemingly more directly related to the optics than poly CD due to the influence of etching bias during polysilicon patterning. On the other hand, however, since the data used in extracting aberration coefficients is from electrically measured poly CD and this data is used in simulations to calculate the sensitivity factor of CD on each individual Zernike term, it is more appropriate to use poly CD instead of resist CD. In this way the effect of etching bias is factored in to extracted resist parameters and will be more or less cancelled out in extracted lens aberration data. Secondly, the repeatability of SEM measurements is not as good as that of electrical measurements, as mentioned in the previous chapter. This will negatively affect the accuracy of the final results.

The following section introduces simplex method, a multi-dimensional optimization algorithm, which form the basis of the above optimization engine.

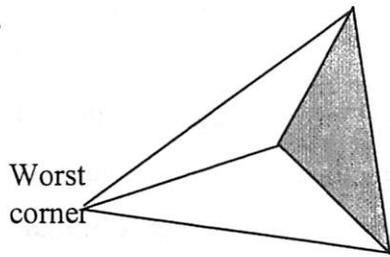
5.7.1 Optimization Using Downhill Simplex Method

Optimization techniques play an important role in various fields of science and engineering by maximizing or minimizing an object function subject to constraint conditions. It is the art of allocating limited resources for the best possible effect, or finding the very best condition or setting, which yields the most satisfying performance or

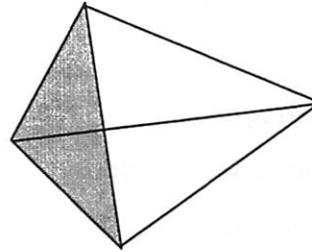
results. Multidimensional optimization is the technique of finding the minimum or maximum of a function with more than one independent variable [5.22]. For example, several dozen parameter values define the resist characteristics. The simplex method is one of the first practical, large-scale optimization techniques. It is “a systematic procedure for generating and testing candidate vertex solution” [5.23]. Its advantage is that the method can be utilized in large-scale, multi-dimensional systems without computing the derivative of the object function, which may save a lot of execution time. The method selects the best choice at every iteration step without using information from previous or future iterations. This is also a very attractive property in terms of memory space when the problem involves large amounts of data. The disadvantage is that the simplex method usually requires evaluating the object function many times and may increase computational cost for some cases as compared to other multidimensional optimization methods.

The algorithm begins at an initial guess of the solution set. At each iteration step, the simplex method chooses the point that will have the largest possible update towards the best solution. This point replaces one of the worst corners, thus moving the interim solution closer to its final solution.

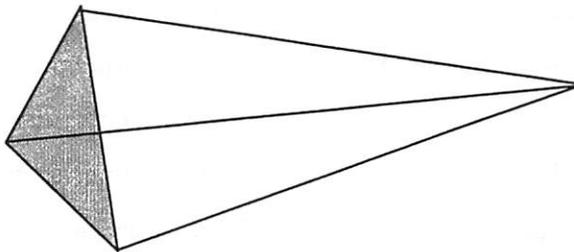
If the optimization problem has N independent variables (N dimensional problem), the simplex is defined as a geometrical object consisting of $N+1$ points connected to each other.



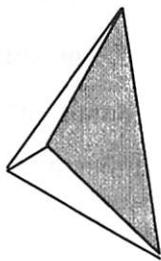
(a) Initial state of simplex



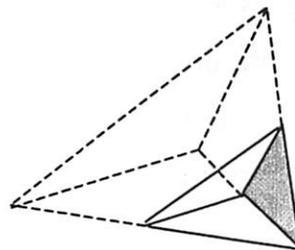
(b) reflection



(c) reflection and expansion



(d) contraction



(e) multiple contraction

Figure 5-18 Geometrical illustration of downhill simplex algorithm
(after ref [22])

To visualize the structure of a simplex and how it evolves as it iterates towards convergence, let us assume a three-dimensional case, as illustrated in the figure 5-18.

At the starting step, the simplex method needs as input an initial guess of the solution. This of course will affect the final result and convergence time. If the initial guess is close to a local minimum, the method is likely to trap into it. Therefore, a good initial guess input is critical for the performance of the algorithm. Instead of just a single starting point, the simplex method requires input of $N+1$ points. This is $N+1$ sets of N dimension data, an $(N+1) \times N$ matrix. Then the method takes a series of steps to transform the shape and position of the simplex until certain termination criteria have been met, at which point the final solution has been located. The most common step is called reflection, in which the object function is computed at all $N+1$ simplex corners and the worst point is replaced by its image point with respect to the surface consists of the remaining N points, as shown in figure 5-18(b). When this is done, the method further expands or contracts the simplex in either direction at a different scale. If a situation occurs in which reflection and contraction about different surfaces does not lead to a large function change, the algorithm pulls the simplex towards its best point along all dimensions.

Four ending criteria of the iteration should be used selectively. The first criterion is that the update to the object function is smaller than a predefined value. The optimization loop can also be stopped when the change in any input dimension is small enough, when the volume enclosed by simplex is small, or when a maximum number of cycles has been reached.

Table 5-3 Downhill Simplex Algorithm

Basic Downhill Simplex Method Algorithm [5.24]

Begin

Initialize Simplex

Repeat

Compute $f_{obj}(\text{simplex})$

$rtol = \text{fractional range from highest point to lowest point}$

$nfunk ++$ /* iteration cycle */

reflect simplex from the high point

extrapolate by factor a through the face of simplex from the high point

if ($f_{obj}(\text{new point}) < f_{obj}(\text{old point})$) then

extrapolate by factor b

else

extraction

if ($f_{obj}(\text{contracted point}) > f_{obj}(\text{original})$) then

all dimensional contraction

endif

endif

Until ($rtol < ftol$) || ($nfunk > NMAX$)

End

Figure 5-19 depicts the comparison of simulated CD using optimized resist A B C parameters with poly CD under various focus and exposure conditions. The optimization

assumes a reasonably good initial guess to have relatively fast convergence, since aerial image simulation is computationally expensive. The results agree fairly well with measured poly CD for the entire focus-exposure range. However, it should be noted that the simplex method does not perform satisfactorily when the initial guess is too far away from the optimal value.

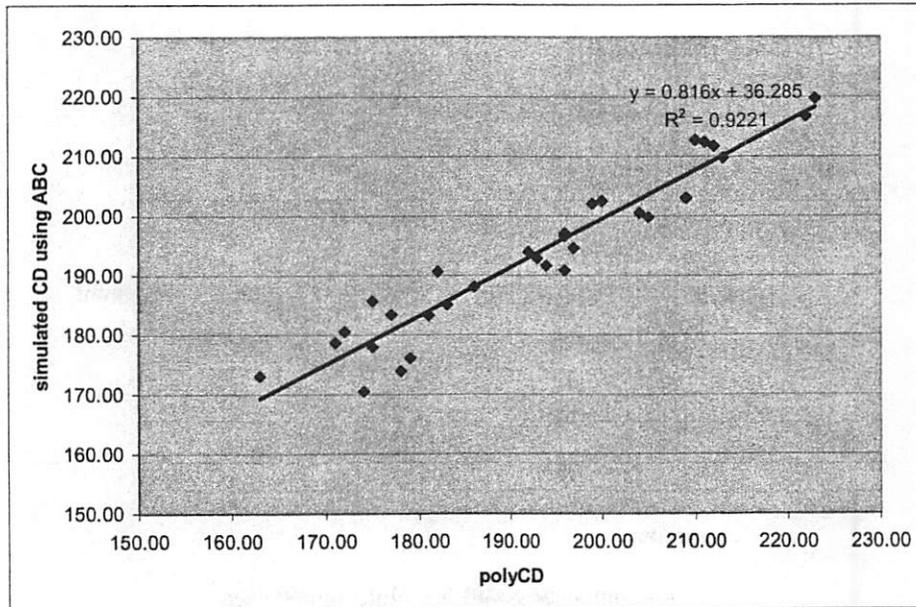


Figure 5-19 Comparison of simulated CD using optimized resist parameters with poly CD under various focus and exposure conditions

5.8 Discussion on Issues in Solving the Linear System

As introduced in previous sections, the basic premise of this method is to linearize the linewidth as the function of individual aberration terms under different exposure conditions, and to solve the resulting linear system. Since some aberration terms have a very small impact on linewidth of printed pattern, they should not be included into the set of Zernike terms to be extracted. Otherwise, the condition number of the matrix of

sensitivity factors will be too large. This will lead to a large numerical error in calculated Zernike coefficients, because small errors (perturbations) in the measurement data will cause a dramatic change in the results. Another problem is that the vectors of sensitivity factors for two different Zernike terms under various exposure settings may be linearly dependent or nearly dependent on each other. If this is the case, the matrix is rank deficient or close to rank deficient. There may be no unique solution even for an over-determined system with an exact rank less than the number of unknowns. However, it is ideally desirable to extract as many Zernike terms as possible using printed patterns. Some aberration coefficients may have a strong effect on image distortion or position shift, but a relatively small influence on linewidth only. Here the question arises of studying the best possible combination of Zernike terms and the best choice of experiment settings (the focus-exposure matrix selection in this case), so that a balance can be achieved among inherent algorithm error, measurement error, resist modeling error and largest set of extractable lens aberration coefficients.

The following discussion assumes that the above best selection of Zernike terms and experiment conditions are determined. We only focus on the issues in solving the linear equation system.

Given m-by-n sensitivity factor matrix

$$\tilde{G} = \begin{bmatrix} \left. \frac{\partial CD}{\partial Z_1} \right|_{FE^{(1)}} & \cdots & \left. \frac{\partial CD}{\partial Z_n} \right|_{FE^{(1)}} \\ \vdots & \ddots & \vdots \\ \left. \frac{\partial CD}{\partial Z_1} \right|_{FE^{(m)}} & \cdots & \left. \frac{\partial CD}{\partial Z_n} \right|_{FE^{(m)}} \end{bmatrix} \quad (5-11)$$

and m -by- 1 vector $\tilde{b} = \tilde{\alpha} - \tilde{\beta} = \text{WaferCD}|_{m \times 1} - \text{CD}_o(\tilde{Z} = \tilde{0})|_{m \times 1}$, we want to find the best \tilde{Z} by minimizing the objective function $\|\tilde{G}\tilde{Z} - \tilde{b}\|_2$. If $m > n$, this means that we have more equations than unknowns. In that case, the linear system is over-determined.

There are three explicit solutions for this linear least squares problem. The simplest one is the normal equation, which is used in the previous section for extraction. It is the fastest but least accurate. This solution is expressed as

$$\tilde{Z} = (\tilde{G}'\tilde{G})^{-1}\tilde{G}'(\tilde{\alpha} - \tilde{\beta}) \quad (5-12)$$

This method is adequate when we are only interested in a small number of Zernike coefficients with large sensitivity factors. Since the condition number of \tilde{G} is small in this case, the normal equation is a convenient choice with the advantage of the least computational cost, which is especially attractive when the required aerial image calculation is numerically intensive.

The second method is QR decomposition, which may require up to twice the cost of the normal equation method but with improved accuracy. As with the normal equation method, QR decomposition requires that \tilde{G} is a full rank matrix, $\text{rank}(\tilde{G}) = n$. Then \tilde{G} is decomposed as $\tilde{Q}\tilde{R}$, where \tilde{Q} is m -by- n orthogonal matrix ($\tilde{Q}'\tilde{Q} = I_n$) and \tilde{R} is n -by- n upper triangular matrix. The solution of this linear least square problem is then expressed as,

$$\begin{aligned}
\tilde{Z} &= (\tilde{G}' \tilde{G})^{-1} \tilde{G}' (\tilde{\alpha} - \tilde{\beta}) \\
&= (\tilde{R}' \tilde{Q}' \tilde{Q} \tilde{R})^{-1} \tilde{R}' \tilde{Q}' (\tilde{\alpha} - \tilde{\beta}) \\
&= (\tilde{R}' \tilde{R})^{-1} \tilde{R}' \tilde{Q}' (\tilde{\alpha} - \tilde{\beta}) \\
&= \tilde{R}^{-1} (\tilde{R}')^{-1} \tilde{R}' \tilde{Q}' (\tilde{\alpha} - \tilde{\beta}) \\
&= \tilde{R}^{-1} \tilde{Q}' (\tilde{\alpha} - \tilde{\beta})
\end{aligned} \tag{5-13}$$

The last method is singular value decomposition (SVD). It is more commonly used for general cases in which \tilde{G} is possibly ill-conditioned or even rank deficient. SVD requires several times more computational cost but achieves best accuracy when compared to the first two methods.

SVD is defined as $\tilde{G} = \tilde{U} \tilde{\Sigma} \tilde{V}'$, where \tilde{U} is m-by-n matrix satisfying $\tilde{U}' \tilde{U} = I$, \tilde{V} is n-by-n matrix satisfying $\tilde{V}' \tilde{V} = I$, and $\tilde{\Sigma}$ is diagonal matrix $\text{diag}(s_1, s_2, \dots, s_n)$ with $s_1 > s_2 > \dots > s_n > 0$ are called singular values. If the over-determined system has full rank n, and the QR decomposition and singular value decomposition of \tilde{G} are

$\tilde{G} = \tilde{Q} \tilde{R} = \tilde{U} \tilde{\Sigma} \tilde{V}'$ respectively, then the so-called Moore-Penrose pseudo-inverse matrix of \tilde{G} is expressed as $\tilde{G}^+ = (\tilde{G}' \tilde{G})^{-1} \tilde{G}' = \tilde{R}^{-1} \tilde{Q}' = \tilde{V} \tilde{\Sigma}^{-1} \tilde{U}'$. The solution of this over-determined full rank linear least squares problem is simply $\tilde{Z} = \tilde{G}^+ (\tilde{\alpha} - \tilde{\beta})$.

If \tilde{G} is possibly rank deficient, the solution is still the same, $\tilde{Z} = \tilde{G}^+ (\tilde{\alpha} - \tilde{\beta})$, but

$\tilde{G}^+ = \tilde{V} \tilde{\Sigma}^+ \tilde{U}'$, where $\tilde{\Sigma}^+ = \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\tilde{G} = [\tilde{U}_1, \tilde{U}_2] \begin{bmatrix} \tilde{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} [\tilde{V}_1, \tilde{V}_2]'$ is exactly singular

and $\tilde{\Sigma}_1$ is r-by-r nonsingular matrix ($r < n$).

As mentioned above, if Zernike terms and exposure settings are selected so that the sensitivity factor matrix \tilde{G} is well-conditioned, the normal equation method is a natural choice given its high speed, but is potentially unstable. If \tilde{G} is not well-conditioned but is far from rank deficient, QR decomposition may be used with better accuracy. If \tilde{G} is close to rank deficient, singular value decomposition should be chosen, although it will result in a higher computational cost. In summary, the tradeoff between reliability and speed lies not only in choosing the method of solving linear least squares equations, but also in choosing best combination of unknown terms and experimental conditions.

5.9 Summary

Lens aberration is believed to be one of the most important underlying physical causes leading to systematic spatial variation and thus limiting the lithography performance. This chapter introduces a simple method to measure full-field Zernike aberrations using conventional printed linewidth pattern. The essence of this technique is the surface representation of the linewidth and aberration levels. The coefficients of sensitivity factors in linear expansion are obtained from simulation. The linear expansion can be conducted at different process settings to obtain more equations to solve multiple unknowns. A careful selection of the combination of Zernike terms and process conditions is the key to avoid degenerate case or ill-condition situations. This method is tested by experiment and simulation. The results from experiment are not satisfactory due to the fact that they are dominated by various noise sources, including metrology error, resist modeling, etching biasing, algorithm error, and other error factors. The simulation results show a best correlation of 0.5 when 3 nm noise is applied.

This method might be improved if use scatterometry to measure the full profile of the resist pattern instead of linewidth only. The partial coherence should be set to a small

value to enhance the sensitivity of aberrations. Different pattern orientations should be considered to obtain information of certain aberration terms that are orientation specific. In addition to first-order additive linear expansion in modeling, higher order or interaction terms could be included in order to consider high order effects. Finally, advanced optimization techniques, such as singular value decomposition, are potentially helpful to improve the stability of extracted results.

In summary, this technique has the advantages of high speed and low cost but has limitations in its reliance on a precise resist model and insensitivity to certain aberration terms. Various issues which affect the performance of this method and future improvement are discussed. Even though this method holds considerable promise, its application is limited by the present day lithography error and simulation accuracy.

References for Chapter 5

- [5.1] T. A. Brunner, "Impact of lens aberrations on optical lithography," IBM Journal of Research Development, vol. 41, No.1/2 pp. 57-67, March 1997.
- [5.2] H. Y. Liu, C. Yu and R. Gleason, "Contributions of stepper lenses to systematic CD errors within exposure fields," Optical Microlithography, Proceedings of the SPIE, vol. 2440, pp. 868-877, 1995.
- [5.3] S. Nakao, A. Nakae, K. Tsujita, and W. Wakamiya, "Impact of spherical aberrations on printing characteristics of irregularly aligned patterns of alternating phase shift mask," Japan Journal of Applied Physics, vol. 38, pp. 1919-1926, Apr. 1999.
- [5.4] A. Kroyan, M. D. Levenson and F. K. Tittel, "Coping with the impact of lens aberrations in the context of wavefront engineering," Optical Microlithography, Proceedings of the SPIE, vol. 3334, pp. 832-838, 1998.
- [5.5] J. P. Kirk, "Astigmatism and field curvature from pin-bars," Optical/Laser Microlithography IV, vol. 1463, pp. 282-291, 1991.
- [5.6] P. Dirksen, C. Juffermans, A. Engelen, P. De. Bisschop and H. Meullerke, "Impact of high order aberrations on the performance of the aberration monitor," Optical Microlithography, Proceedings of the SPIE, vol. 4000, 2000.
- [5.7] J. P. Kirk, "Review of photoresist based lens evaluation methods," Optical Microlithography, Proceedings of the SPIE, vol. 4000, 2000.
- [5.8] T. Sato and H. Nomura, "Coma aberration measurement by relative shift of displacement with pattern dependence," Japan Journal of Applied Physics, vol. 37, Part 1, No. 6A, pp. 3553-3557, June 1998.
- [5.9] A. Imai, K. Hayano, H. Fukuda, N. Asai, and N. Hasegawa, "Lens aberration measurement technique using attenuated phase-shifting mask," Optical Microlithography, Proceedings of the SPIE, vol. 4000, 2000.
- [5.10] P. Dirksen, C. Juffermans, R. Pellens, M. Maenhoudt and P. De. Bisschop, "Novel aberration monitor for optical lithography," Optical Microlithography, Proceedings of the SPIE, vol. 3679, pp. 77-86, 1999.

- [5.11] H. Nomura and T. Sato, "Techniques for measuring aberrations in lenses used in photolithography with printed patterns," *Applied Optics*, vol. 38, No. 13, pp. 2800-2807, May 1999.
- [5.12] H. Fukuda, K. Hayano and S. Shirai, "Determination of high-order lens aberrations using phase/amplitude linear algebra," *Journal of Vacuum Science and Technology*, vol. B, 17(6), pp. 3318-3321, Nov/Dec 1999.
- [5.13] T. A. Brunner, A. L. Martin, R. M. Marino, C. P. Ausschnitt, T. H. Newman and M. S. Hibbs, "Quantitative stepper metrology using the focus monitor test mask," *Optical Microlithography*, Proceedings of the SPIE, vol.2197, pp. 541-549, 1994.
- [5.14] D. G. Flagello and B. Geh, "Lithographic lens testing: Analysis of measured aerial images, interferometric data and photoresist measurements," *Proceedings of SPIE's Ninth Annual International Symposium on Microlithography*, vol. 2726, 1996, pp. 788-798.
- [5.15] J. P. Kirk, "Measurement of astigmatism in Microlithography lenses," *Optical Microlithography*, Proceedings of the SPIE, vol. 3334, pp. 848-854, 1998.
- [5.16] G. Robins, K. Adam and A. R. Neureuther, "Measuring optical image aberrations with pattern and probe based targets," *Journal of Vacuum Science and Technology*, vol. B 20(1), pp. 338-343, Jan/Feb 2002.
- [5.17] A. R. Neureuther, "Modeling phase shifting masks," *Optical Microlithography*, Proceedings of the SPIE, vol.1496, pp. 80 – 88, 1990.
- [5.18] V. Mastromarco, K.H. Toh and A.R. Neureuther, "Printability of defects in optical lithography: polarity and critical location effects," *Journal of Vacuum Science and Technology*, vol. B, pp. 224-229, Jan/Feb 1988.
- [5.19] K. H. Toh and A. R. Neureuther, "Identifying and monitoring effects of lens aberrations in projection printing," *Optical Microlithography*, Proceedings of the SPIE, vol. 772, pp. 202-209, 1987.
- [5.20] A. R. Neureuther, P. Flanner III and S. Shen, "Coherence of defect interactions with features on optical imaging," *Journal of Vacuum Science and Technology*, vol. B, pp. 308-312, Jan/Feb. 1987.
- [5.21] G. C. Robins and A. R. Neureuther, "Illumination, mask and tool effects on pattern and probe-based aberration monitors," *Optical Microlithography XV*, Proceedings of the SPIE, vol. 4691, 2002

[5.22] R. J. Larsen and M. L. Marx, "An introduction to mathematical statistics and its applications," Prentice Hall, Inc., 1981.

[5.23] Gill, Philip E.. Numerical linear algebra and optimization, Philip E. Gill, Walter Murray, Margaret H. Wright. Redwood City, Calif. : Addison-Wesley Pub.

[5.24] Numerical recipes in C: The art of scientific computing.

Chapter 6

Compact Formulation of Mask Error Factor for Critical Dimension Control

This chapter introduces a rigorous theoretical formulation of the mask error factor. The results are compared with experiments as well as simulations and a good match has been found. This work provides an insight into the cause of the MEF and it is helpful to find its dependency on process settings in order to effectively control critical dimension variation in photolithography.

6.1 Introduction

A significant component of systematic within-field critical dimension (CD) variation is contributed by mask errors. As the k_1 factor becomes smaller, the photomask gradually loses control of the image [6.1][6.2][6.3]. When that happens, the sensitivity of the image transfer mechanism is such that any deviations of the mask pattern are magnified by a factor known as the Mask Error Factor (MEF) before they reach the wafer plane. The mask error factor is defined as the ratio of wafer CD deviation over de-magnified mask CD deviation.

$$MEF = \frac{\Delta CD_{wafer}}{\Delta CD_{mask}}$$

The MEF larger than unity is caused by degradation of image integrity when feature size decreases. Nonlinearity in both the imaging system and resist response exacerbates CD

error on the mask. This effect increases CD variation and makes CD control a more difficult challenge.

Several theoretical and experimental works have been reported to study the effects of mask error factor as well as its relation with processing conditions [6.4-6.7]. Although this knowledge is helpful in understanding its implications in process control, the theoretical physical cause, coupled to a rigorous formulation of the mask error factor are still lacking. Such a formulation in particular will be very useful in quantifying the dependence of the mask error factor on process conditions, and in ultimately controlling it. Furthermore, they are potentially useful in improving the efficiency and accuracy of optical proximity correction rule generation. Finally, since lens aberrations and mask errors exist simultaneously, the actual linewidth variation is the result of the interaction of these two causes. Therefore, theoretical understanding of mask error factor is necessary to find the MEF in presence of local lens aberrations.

In summary, a theoretical understanding of the cause of MEF is helpful to control across field linewidth variation (AFLV). Our analysis gives a coherent theoretical basis and a straightforward formulation of the MEF. In this way, the optimum process window can be chosen. This formulation will also provide an understanding of the dependency of MEF on local lens aberrations.

The chapter is organized as follows. Section 6.1 gives the introduction and background. In section 6.2, the theoretical formulation of MEF is introduced. The experiment and results are explained in section 6.3. Finally, section 6.4 concludes this subject with a summary, remarks and suggestion for future work.

6.2 A Compact Theoretical Analysis of Mask Error Factor

This analysis assumes that the illumination is coherent and we focus on isolated features. We start by deriving an analytical expression of the aerial image of isolated features under coherent illumination. The Fraunhofer diffraction pattern of an isolated space with width a is expressed as a *sinc* function.

$$E_1(f_x) = A_o \frac{\sin(\pi a f_x)}{\pi a f_x} \quad (6-1)$$

where $A_o = I_o a$, f_x denotes the spatial frequency and I_o is illumination intensity. The pupil function due to limited NA is,

$$E_2(f_x) = \begin{cases} 1, & \text{when } |f_x| \leq \frac{NA}{\lambda} \\ 0, & \text{else} \end{cases} \quad (6-2)$$

The electrical field of the aerial image is expressed as the convolution of the inverse Fourier transform of (6-1) and (6-2).

$$E(x) = \int_{-\infty}^{\infty} A_1(\tau) A_2(x - \tau) d\tau \quad (6-3)$$

where $A_1 = \mathcal{F}^{-1}[E_1(f_x)] = A_o/a$; $A_2 = \mathcal{F}^{-1}[E_2(f_x)] = \frac{\sin T_2 x}{\pi x}$. Substituting into (6-3) and calculating the integration, we can obtain the analytical expression of the aerial image.

$$\begin{aligned}
E(x) &= \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{A_o}{a} \frac{\sin T_2(x-\tau)}{\pi(x-\tau)} d\tau \\
&= -\frac{A_o}{\pi a} \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{\sin T_2(x-\tau)}{T_2(x-\tau)} d[T_2(x-\tau)] \\
&= \frac{I_o}{\pi} \int_{T_2(x-\frac{a}{2})}^{T_2(x+\frac{a}{2})} \frac{\sin y}{y} dy
\end{aligned} \tag{6-4}$$

$$\text{Since} \quad \int \frac{\sin y}{y} dy = \frac{1}{2j} \int \left(\frac{e^{jy}}{y} - \frac{e^{-jy}}{y} \right) dy \tag{6-5}$$

$$\text{while} \quad \int \frac{e^{cx}}{x} dx = \ln x + cx + \frac{(cx)^2}{2 \cdot 2!} + \frac{(cx)^3}{3 \cdot 3!} + \dots \tag{6-6}$$

substitute (6-6) into (6-5), we have

$$\begin{aligned}
\int \frac{\sin y}{y} dy &= \frac{1}{2j} (\ln y + jy - \frac{y^2}{2 \cdot 2!} - j \frac{y^3}{3 \cdot 3!} + \frac{y^4}{4 \cdot 4!} + \dots) - \frac{1}{2j} (\ln y - jy - \frac{y^2}{2 \cdot 2!} + j \frac{y^3}{3 \cdot 3!} + \frac{y^4}{4 \cdot 4!} + \dots) \\
&= y - \frac{y^3}{3 \cdot 3!} + \frac{y^5}{5 \cdot 5!} - \frac{y^7}{7 \cdot 7!} + \frac{y^9}{9 \cdot 9!} - \dots \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} y^{2n-1}}{(2n-1)(2n-1)!}
\end{aligned} \tag{6-7}$$

substitute (6-7) into (6-4), we obtain the final aerial image as follows,

$$E(x) = \frac{1}{\pi} \sqrt{I_o} \sum_{n=1}^{\infty} \frac{(-1)^{n+1} \{ [T_2(x+a/2)]^{2n-1} - [T_2(x-a/2)]^{2n-1} \}}{(2n-1)(2n-1)!} \tag{6-8}$$

where x is the coordinate on the image plane; T_2 is NA/λ , and a is mask CD. As an example, Figure 6-1 illustrates the calculated aerial image ($I=E^2(x)$) using the above equation with the first 100 terms. NA is 0.5 and illumination intensity I_0 is unity.

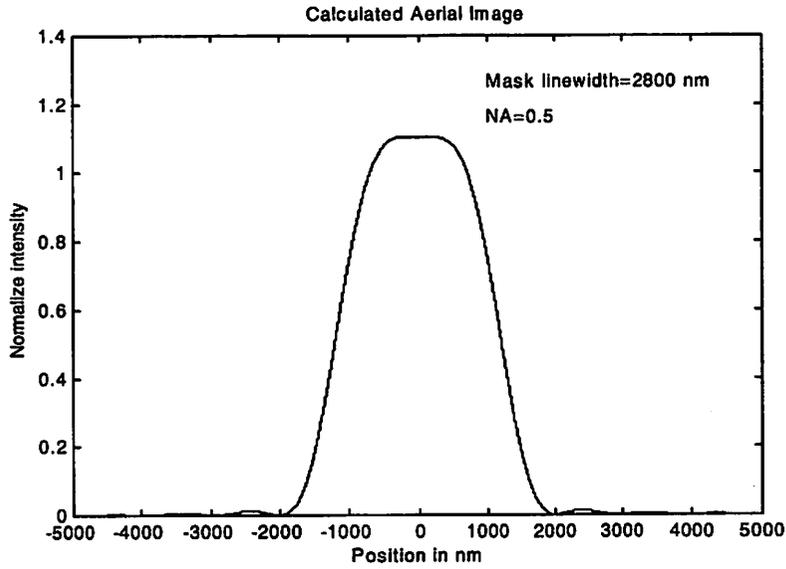


Figure 6-1 Closed form aerial image calculation

Mask error factor is defined as the ratio of wafer CD error over mask CD error, then

$$MEF = \frac{\Delta CD_{wafer}}{\Delta CD_{mask}} = \frac{\Delta x}{\Delta a} \quad (6-9)$$

where x and a correspond to coordinates of the wafer plane and mask plane respectively.

We then can express the above as

$$\frac{\Delta x}{\Delta a} = \frac{\partial I / \partial I}{\partial a / \partial x} \quad (6-10)$$

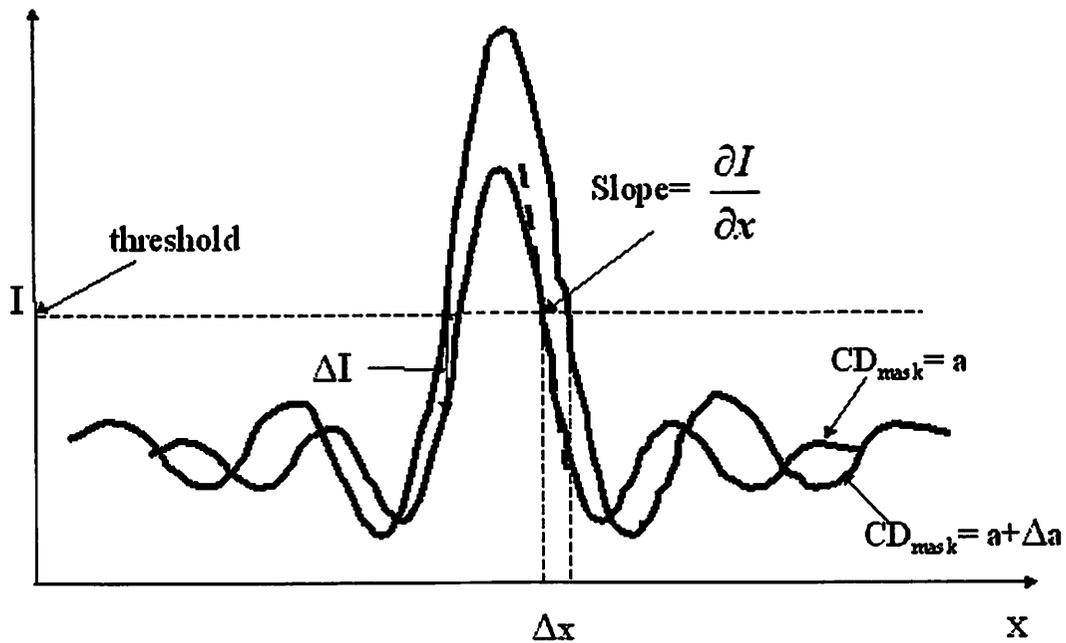


Figure 6-2 Illustration mask error factor assuming fixed resist threshold

Figure 6-2 shows the physical meaning of the mask error factor assuming a fixed resist threshold. The two curves shown are aerial images corresponding to the isolated space pattern on mask with and without mask error respectively. When resist threshold is constant, the deviation of CD on wafer depends on the image slope and the deviation of image intensity at this threshold.

$$\frac{\partial I}{\partial a} = \frac{\partial E^2}{\partial a} = 2E \frac{\partial E}{\partial a} ; \frac{\partial I}{\partial x} = \frac{\partial E^2}{\partial x} = 2E \frac{\partial E}{\partial x} \quad (6-11)$$

$$\text{Thus } MEF = \frac{\partial E}{\partial a} / \frac{\partial E}{\partial x}$$

Using the aerial image expression in (6-3) and (6-4), we can calculate the derivative as,

$$\frac{\partial E}{\partial a} = \frac{T_2 \sqrt{I_o}}{2\pi} [\text{sinc}(u) + \text{sinc}(v)] \quad (6-12a)$$

$$\frac{\partial E}{\partial x} = \frac{T_2 \sqrt{I_o}}{2\pi} [\text{sinc}(u) - \text{sinc}(v)] \quad (6-12b)$$

$$\text{where} \quad u = T_2 \left(x + \frac{a}{2}\right); \quad v = T_2 \left(x - \frac{a}{2}\right)$$

and x corresponds to resist threshold.

Then we have

$$MEF = \frac{1 \text{ sinc}(u) + \text{sinc}(v)}{2 \text{ sinc}(u) - \text{sinc}(v)} \quad (6-13)$$

Figure 6-3 shows the calculated MEF using the above equation.

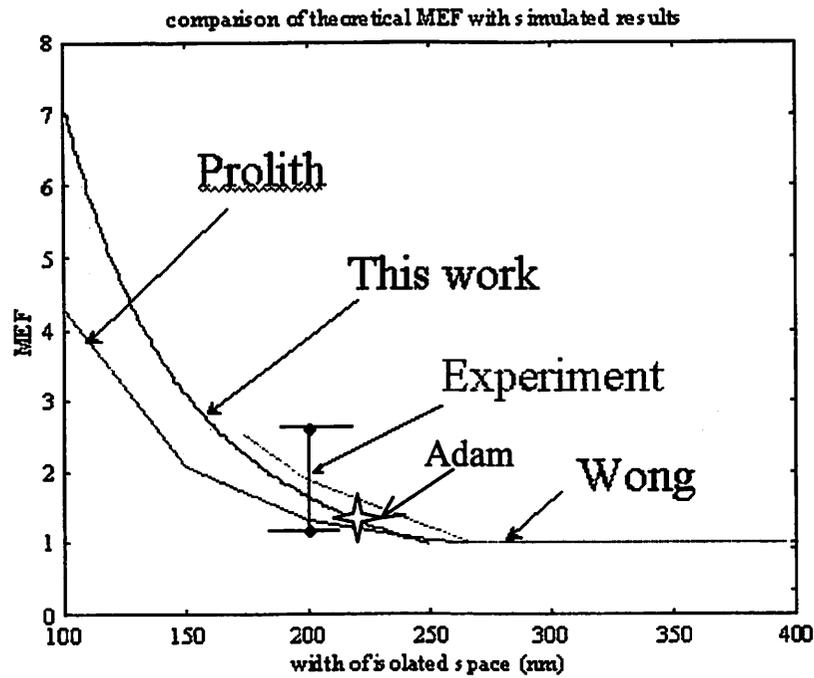


Figure 6-3 Calculated MEF as function of feature size compared with simulation and other work

When linewidth a is large, the MEF is approximately equal to unity, and it increases as line width decreases. Thus mask quality is more important for small features. It can be seen that the MEF also depends on the resist threshold. Therefore the MEF can be controlled by adjusting resist threshold and exposure dose. On the other hand, this will be a limiting factor for process latitude.

6.3 Experiment and Comparison with Theory and Simulation

The experimental value of the MEF is obtained with the help of direct mask measurements. The nominal CD is 220 nm. The theoretical MEF calculated using equation (6-9) is shown in the figure 6-3. Also plotted are the simulated MEF using Prolith, simulation results by Wong [6.5] and experimental results by Kostas Adam [6.9]. The closed form MEF matches simulation and Adam's measurements fairly well in most regions. Adam's results indicate that resist tends to improve the MEF and OPC feature (scattering bar) also help to reduce MEF of isolated lines by 10%. There is a relatively large range of experimental MEF. The discrepancy between theory and experiment is most likely due to resist effects, partial coherence and lens aberrations, which are not taken into account in theoretical formulation. It suggests that in addition to lens aberrations and other effects, the mask error factor is another important source causing within-field linewidth variations.

6.4 Conclusion

A theoretical formulation of the mask error factor is introduced. An analytical formulation of the MEF is useful to understand its physical cause in order to effectively control the linewidth variations. It is also potentially helpful to improve the existing resolution enhancement techniques (OPC). A fairly good match has been found between the theory and simulations. The experiment shows that lens quality is another important source of variation besides the mask error. Future work includes comparing theoretical results and simulation results with experiments, with and without the consideration of lens aberrations, in order to determine the relative importance of the roles of both the MEF and lens aberrations in linewidth variations. It would also be meaningful to intentionally design a relatively large range of errors on a mask to evaluate the theory.

Future work will also quantify the relation between MEF and process settings, such as exposure dose.

References for Chapter 6

- [6.1] W. Maurer, K. Sato, D. Samuels and T. Fischer, "Pattern transfer at $k_1=0.5$: Get 0.25 μm lithography ready for manufacturing," *Optical Microlithography, Proceedings of the SPIE*, vol. 2726, pp. 113-124, 1996.
- [6.2] P. Yan, R. Hainsey, J. Farnsworth and J. Neff, "Sub-micron low- k_1 imaging characteristics using a DUV printing tool and binary masks," *Optical Microlithography, Proceedings of the SPIE*, vol. 2440, pp. 270-276, 1995.
- [6.3] B. J. Lin, "Methods to print optical images at low- k_1 factors," *Optical Microlithography, Proceedings of the SPIE*, vol. 1264, pp. 2-13, 1990.
- [6.4] P. Yan and J. Langston, "Mask CD control requirement at 0.18 μm design rules for 193nm lithography," *Optical Microlithography, Proceedings of the SPIE*, vol. 3051, pp. 164-169, 1997.
- [6.5] A. Wong, R. Ferguson, L. Liebmann, S. Mansfield, A. Molless and M. Neisser, "Lithographic effects of mask critical dimension error," *Optical Microlithography, Proceedings of the SPIE*, vol. 3334, pp. 106-116, 1998.
- [6.6] J. Wiley and J. Reynolds, "Device yield and reliability by specification of mask defects," *Solid State Technology*, vol. 36, pp. 65-66, 70, 72, 74, 77, July 1993.
- [6.7] J. van Schoot, J. Finders, K. van Ingen Schenau, M. Klaassen and C. Bujik, "The mask error factor: Causes and implications for process latitude," *Optical Microlithography, Proceedings of the SPIE*, vol. 3679, pp. 250-260, 1999.
- [6.8] H. Zhang, J. Cain and C. J. Spanos, "Compact formulation of mask error factor for critical dimension control in optical lithography", *Metrology, Inspection, and process control in Microlithography XVI, Proceedings of the SPIE*, vol. 4689, 2002.
- [6.9] K. Adam, R. Socha, M. Dusa, A. Neureuther. "Effect of reticle CD uniformity on wafer CD uniformity in the presence of scattering bar optical proximity correction". *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 3546, (18th Annual BACUS Symposium on Photomask Technology and Management, Redwood City, CA, USA, 16-18 Sept. 1998.) *SPIE-Int. Soc. Opt. Eng.* 1998. p.642-50. 7 references

Chapter 7

Conclusions and Future Work

7.1 Thesis Summary

Semiconductor manufacturing has entered the era of extremely fine feature size and exceedingly complex integrated systems. Today's state-of-art semiconductor fabs are capable of printing 0.09 μm pattern in mass production. Devices with different gate oxide thickness, threshold voltages and voltage supplies are now fabricated on the same silicon to meet the various requirement of digital core, IO part, memory module, and analog/RF blocks, which together enable the implementation of system-on-chip solutions. As many as 8 layers of copper metal and low-k dielectric material ease the constraints of interconnect delay, and enable very complicated systems with functionality unimaginable just a few years ago. Another achievement of today's semiconductor technology is the revival of bipolar technology on SiGe material with its superior transition time, which enables many high-speed applications including wireless. It is a difficult task to integrate tens of millions of transistors, resistors, capacitors and inductors on very different material onto the silicon.. Therefore, advanced process/equipment control is the key factor enabling the semiconductor foundry to maintain high production yield in order to succeed in this highly competitive environment, given the fact of similarity in all basic process modules among different fabs. A silicon wafer of 0.13 μm process typically goes through 20 lithography steps ignoring the backend steps such as bond pad and packaging before completion. It is the lithography that determines the feature size; therefore process control in lithography

plays the most important role in the entire semiconductor manufacturing process. The k_1 factor is pushed to be very small in today's DUV lithography in printing $0.09\mu\text{m}$ and beyond using 193nm system. Furthermore, a variety of resolution enhancement techniques (RET), such as OPC, PSM and off-axis illumination, have been routinely put into application. They have dramatically extended the DUV lithography lifetime while they have significantly shrunk the process latitude, since all these RET techniques are highly sensitive to process and equipment conditions. Therefore process control in lithography is becoming a more and more critical task.

This thesis conducted a systematic characterization of process variability for the purpose of efficient process control in DUV lithography. Through extensive linewidth measurement, spatial distribution of CD variation has been captured across the die as well as across the wafer. It has been found that within field variation is a major variation component, consuming a large portion of error budget. This kind of variation is mainly deterministic instead of random in nature. Two important sources of within field variation are found to be reticle error and lens aberrations. This thesis has presented a new approach to extract the full-field lens aberration signature using printed patterns. The accuracy of this methodology is analyzed and a precise resist model is found to be the key for the reliability of results obtained using this approach. Mask error magnification, as another important source of systematic spatial variation, has been studied theoretically. The rigorous formulation of the mask error factor has been derived based on closed form aerial image calculation under coherent illumination conditions. Focus/exposure fluctuation is an important error source for die-to-die spatial variation. In chapter 3, an automatic focus/exposure control method has been proposed based on digitized SEM scan using statistical feature extraction and neural network classification.

7.2 Future Work

As a widely used metrology tool, CD-SEM is generally used to measure linewidth. The method presented in chapter 3 of CD-SEM based focus-exposure control can be extended by further mining of pattern profile information. This can be achieved by building a one-to-one correspondence between CD-SEM trace and CD-AFM profile for any given material. Through intelligent data analysis techniques, more process information will be obtained besides routine CD measurement, without requiring complicated modeling of the interaction of electron beam with samples.

Scatterometry has been proven to be a cost-effective and high-speed in-situ metrology in deep sub-micron CD and profile analysis. It has shown the ability of measuring diffraction gratings using a library search method with a great success. This method can be utilized to improve the performance of the lens aberration extraction method proposed in chapter 5. Instead of using only the linewidth, the full profile of printed pattern can be used to extract Zernike coefficients. The resist profile will contain more information compared to CD, since certain aberration terms not sensitive to CD variation will have impact on the pattern shape. Again, this will also depend on accurate resist modeling and good data analysis techniques in creating an over-determined system with adequate noise immunity. There is no fundamental limitation in extending this lens aberration measurement technique into future technology nodes, given the availability of the metrology for measuring and simulating small features.

The theoretical formulation of the mask error factor helps to understand the underlying physical cause of this common phenomenon and its negative effect on spatial CD uniformity. However, a detailed analysis is needed to take into account the influence of partial coherence. Since it is impractical to formulate MEF in the presence of partial coherence, a statistical modeling approach can be adopted. Mask error magnification and

lens aberration exist at the same time in affecting the linewidth variation in lithography. It is worthy to study the interaction of MEF with local lens aberrations. This knowledge in turn is helpful to improve the performance of existing OPC technique by including the impact of lens aberrations.

Appendix A

List of Symbols

Symbol	Description
ρ	radius coordinate within the plane of exit pupil
θ	angular coordinate within the plane of exit pupil
CR	chief ray
AR	arbitrary ray
S	Gaussian reference sphere
W	aberrated wavefront across exit pupil
P_o	point source
Y^*	height of point source
Φ	aberration function
R	radial Zernike polynomial
a_{lmn}	Zernike coefficients expressed in circular power series
l, n, m	non-negative integer in Zernike polynomial
δ	Kronecker delta function
$Z_1, Z_2 \dots Z_n$	Zernike coefficients in conventional notation
Sr	Strehl ratio

<i>MEF</i>	mask error factor
<i>U</i>	normalized orthogonal vector in calculating principal component
<i>M</i>	number of the dimension of the space spanned by principal components
<i>N</i>	dimension of original input data
\tilde{X}	matrix of input data consisting of N_d dimensional vectors
\tilde{X}'	vector of mean value of matrix \tilde{X}
U'_i	i th normalized eigenvector of the covariance matrix associated with the i th largest eigenvalue
\tilde{D}	diagonal matrix with the eigenvalues as elements
e_i	i^{th} eigenvalue
<i>P</i>	principal component
σ_{aw}^2	across wafer CD variance
σ_{af}^2	across field CD variance
i^f	index of the feature within a die
k	index of the die within a wafer
n^f	total number of each individual feature within a die
N	total number of die within a wafer
\overline{CD}_{i^f}	average value of all dies for the feature of the same position within the die

\overline{CD}^k	average value of all features for the k^{th} die
x, y	wafer plane coordinates
F	defocus distance from the center of focus
E	exposure dose
$CD(FE^{(i)})$	measured linewidth resulting from the i^{th} focus-exposure setting
$CD_o(FE^{(i)})$	simulated linewidth without lens aberration under the i^{th} focus-exposure setting
$\frac{\partial CD}{\partial Z_i}(FE^{(i)})$	linewidth sensitivity factor with respect to the i^{th} Zernike coefficient under the i^{th} focus-exposure setting
$(\cdot)'$	matrix transpose
$(\cdot)^{-1}$	matrix inverse
$\tilde{\alpha}$	vector of measured wafer CD; the number of vector elements is equal to the number of dies on the FEM wafer
$\tilde{\beta}$	vector of simulated CD under zero aberrations
\tilde{G}	matrix of CD sensitivity factors
\tilde{A}	derived matrix from matrix \tilde{G}
$E(\cdot)$	expected value
$VAR(\cdot)$	variance
$\delta(\cdot)$	error

I	unit matrix
a, b, c	resist parameters
SSE	sum square error
$\ \cdot \ _2$	least square fitting residual
\tilde{Q}	orthogonal matrix in QR decomposition
\tilde{R}	upper triangular matrix in QR decomposition
\tilde{U}	left orthogonal matrix in singular value decomposition
\tilde{V}	right orthogonal matrix in singular value decomposition
$diag(\cdot)$	diagonal matrix with diagonal elements contained in vector
$\tilde{\Sigma}$	diagonal matrix in singular value decomposition
s_1, s_2, \dots, s_n	singular values
\tilde{G}^+	Moore-Penrose pseudo-inverse matrix
Kl	The ratio of minimum printable feature size divided by λ/NA
E_1	diffraction pattern on exit pupil
f_x	spatial frequency
E_2	pupil function
NA	numerical aperture
λ	Wavelength
I_o	illumination intensity

$E(x)$	electrical field of aerial image
ΔCD_{wafer}	wafer CD error
ΔCD_{mask}	mask CD error
I	aerial image intensity
T_2	NA/λ

Appendix B-1

VB source code for UV110 parameter extraction using Downhill Simplex method

```
*****
'This program extracts the resist parameters (UV110)
'from measured focus-exposure-matrix data
'using Simplex method, a mutidimensional optimization algorithm
*****
'global prolith objects to be used in this VB module
Dim ProlithApp As Prolith.Application
Dim ProDoc As Prolith.Document
Dim ProSimEng As Prolith.SimEngine
Dim ProlithWasRunning As Boolean

Const CDm = 216.3   'Mask CD
Dim F As Double    'Focus
Dim E As Double    'Exposure dose

'matlab object
Dim MatLab As Object
Dim Result As String
Dim MImag() As Double

Dim FECDvb(30, 2) As Double 'FEM data
Dim WaferCD(30) As Double   'CD column of above FEM data

'Excel objects
Dim xlApp As Excel.Application
Dim xlbook As Excel.Workbook
Dim xlsheet As Excel.Worksheet

*****
'Function_Calculate SSE
'The function to be minimized
'It is defined as sum square error of measured CD
'against simulated CD under current resist parameters
```

```

*****
Public Function SSE(A As Double, B As Double, C As Double) As
Double

Dim FE As Integer
SSE = 0

ProSimEng.SetInput(29117, 0) = A      '0.03
ProSimEng.SetInput(29118, 0) = B      '0.18
ProSimEng.SetInput(29119, 0) = C      '0.11

For FE = 0 To 30 Step 1
  F = FECDvb(FE, 0)
  E = FECDvb(FE, 1)
  WaferCD(FE) = FECDvb(FE, 2)

  ProSimEng.SetInput(29101, 0) = E 'exposure dose
  ProSimEng.SetInput(29102, 0) = F 'focus

  ProSimEng.SingleRun
  SimulatedCD = ProSimEng.GetOutput(8, -1) 'resist CD

  SSE = SSE + (SimulatedCD - WaferCD(FE)) ^ 2

Next FE
  xlsheet.Cells(jjj, 8) = jjj

Close 1

End Function

*****
'Function_amo try extrapolate by a factor fac thru
'the face of simplex across the high point. if the
'new point is better, replace the high point w/ the new one
*****
Public Function amo try(p() As Double, y() As Double, psum() As
Double, ndim As Integer, ihi

As Integer, fac As Double) As Double

Dim j As Integer
Dim fac1 As Double
Dim fac2 As Double
Dim ytry As Double

```

```

Dim ptry(3) As Double

fac1 = (1# - fac) / ndim
fac2 = fac1 - fac

For j = 1 To ndim Step 1
    ptry(j) = psum(j) * fac1 - p(ihi, j) * fac2
Next j

ytry = SSE(ptry(1), ptry(2), ptry(3))

If ytry < y(ihi) Then
    y(ihi) = ytry
    For j = 1 To ndim Step 1
        psum(j) = psum(j) + ptry(j) - p(ihi, j)
        p(ihi, j) = ptry(j)
    Next j
End If

amotry = ytry

Close 1

End Function

'*****
'main program
'*****
Private Sub Form_Load()
'Variables for simplex algorithm
Dim i As Integer
Dim ihi As Integer      'index of high point in simplex
Dim ilo As Integer      'low point index
Dim inhi As Integer     'next high point
Dim j As Integer
Dim mpts As Integer     '= ndim + 1

Dim nfunk As Integer    'iteration number
Dim y(4) As Double      'SSE on simplex points

Dim p(4, 3) As Double   'simplex
Dim rtol As Double      'fractional convergence
Dim sum As Double
Dim swap As Double
Dim ysave As Double
Dim ytry As Double

```

```

Dim psum(3) As Double

Const TINY = 0.0000000001
Const ftol = 0.1          'fractional convergence tolerance
Const NMAX = 20          'maximum iteration number

'connect to prolith
On Error Resume Next
Set ProlithApp = GetObject(, "Prolith.Application")
If Err.Number <> 0 Then
ProlithWasRunning = False
'if not running, launch prolith
Set ProlithApp = CreateObject("Prolith.Application")
ProlithApp.Visible = True
Err.Clear
Else
ProlithWasRunning = True
End If

Set prolithDoc = ProlithApp.ActiveDocument
Set ProSimEng = prolithDoc.SimulationEngine

'set input parameters
ProSimEng.SetInput(29100, 0) = 540      'resist thickness
ProSimEng.SetInput(29105, 0) = 0.57    'NA
ProSimEng.SetInput(29106, 0) = 0.6     'partial coherence
ProSimEng.SetInput(29139, 0) = 248     'wavelength
ProSimEng.SetInput(29116, 0) = 0.258   'Aerial image threshold

'launch Excel
Set xlApp = CreateObject("excel.application")
Set xlbook = xlApp.Workbooks.Add
Set xlsheet = xlbook.Worksheets(1)
xlApp.Visible = True

xlsheet.Cells(1, 1) = "A"
xlsheet.Cells(1, 2) = "B"
xlsheet.Cells(1, 3) = "C"

xlsheet.Cells(1, 10) = "rtol"

'read in FEM data
Set MatLab = CreateObject("MatLab.Application")
Result = MatLab.Execute("load C:\haolin1\feSEMcd.txt")
'file exist?
Call MatLab.GetFullMatrix("feSEMcd", "base", FECDvb, MImag)

```

'FECDvb is waferCD file/array, whose format is 3 columns: F E CD

'-----
mpts = 4
nfunk = 0

'set initial simplex points

p(1, 1) = 0.032 'A1
p(1, 2) = 0.182 'B1
p(1, 3) = 0.112 'C1
p(2, 1) = 0.031 'A2
p(2, 2) = 0.181 'B2
p(2, 3) = 0.111 'C2
p(3, 1) = 0.03 'A3
p(3, 2) = 0.18 'B3
p(3, 3) = 0.11 'C3
p(4, 1) = 0.029 'A4
p(4, 2) = 0.179 'B4
p(4, 3) = 0.109 'C4

'get SSE on initial simplex points

For i = 1 To 4 Step 1
 y(i) = SSE(p(i, 1), p(i, 2), p(i, 3))
Next i

'get PSUM

For j = 1 To 3 Step 1
 sum = 0
 For i = 1 To 4 Step 1
 sum = sum + p(i, j)
 Next i
 psum(j) = sum
Next j
iii = 1

Do While 1

 ilo = 1
 'determine highest(worst), next-highest, and lowest(best) points
 in the simplex
 If y(1) > y(2) Then
 ihi = 1
 inhi = 2
 Else

```

        ihi = 2
        inhi = 1
    End If

    For i = 1 To mpts Step 1
        If y(i) <= y(ilo) Then
            ilo = i
        End If
        If y(i) > y(ihi) Then
            inhi = ihi
            ihi = i
        ElseIf y(i) > y(inhi) And i <> ihi Then
            inhi = i
        End If
    Next i

    rtol = 2 * Abs(y(ihi) - y(ilo)) / (Abs(y(ihi))+Abs(y(ilo))+ TINY)

    xlsheet.Cells(iii + 1, 10) = rtol

    'compute the fractional range from highest to lowest and return
    if satisfactory
    If rtol < ftol Then
        swap = y(1)
        y(1) = y(ilo)
        y(ilo) = swap
        For i = 1 To 3 Step 1
            swap = p(1, i)
            p(1, i) = p(ilo, i)
            p(ilo, i) = swap
        Next i
        Exit Do
    End If

    'exit if maximum number of iteration reached
    If nfunc >= NMAX Then
        Exit Do
    End If

    iii = iii + 1
    For i = 0 To 3 Step 1
        xlsheet.Cells(iii + i, 1) = p(i, 0)
        xlsheet.Cells(iii + i, 2) = p(i, 1)
        xlsheet.Cells(iii + i, 3) = p(i, 2)
    Next i

```

```

    nfunk = nfunk + 2
    'begin a new iteration. First extrapolate by a factor -1 thru
the face
    'of the simplex across from the high point. reflect the simplex
from the high point
    ytry = amotry(p(), y(), psum(), 3, ihi, -1#)
    If ytry <= y(ilo) Then
        ytry = amotry(p(), y(), psum(), 3, ihi, 2#)
    ElseIf ytry >= y(inhi) Then
        ysave = y(ihi)
        ytry = amotry(p(), y(), psum(), 3, ihi, 0.5)
        If ytry >= ysave Then
            For i = 1 To mpts Step 1
                If i <> ilo Then
                    For j = 1 To 3 Step 1
                        p(i, j) = 0.5 * (p(i, j) + p(ilo, j))
                        psum(j) = p(i, j)
                    Next j
                    y(i) = SSE(psum(1), psum(2), psum(3))
                End If
            Next i
            nfunk = nfunk + 3
            'get PSUM
            For j = 1 To 3 Step 1
                sum = 0
                For i = 1 To 4 Step 1
                    sum = sum + p(i, j)
                Next i
                psum(j) = sum
            Next j
        End If
    Else
        nfunk = nfunk - 1
    End If
Loop
'Print out the results
'xlsheet.Cells(1, 1) = "A"
'xlsheet.Cells(1, 2) = "B"
'xlsheet.Cells(1, 3) = "C"

For i = 1 To mpts Step 1
    xlsheet.Cells(iii + 1, 1) = p(i, 0)
    xlsheet.Cells(iii + 1, 2) = p(i, 1)
    xlsheet.Cells(iii + 1, 3) = p(i, 2)
Next i
End Sub

```

Appendix B-2

VB source code for extracting Petzval, primary 90° astigmatism and primary spherical aberration

```
'global constant
Const Zinumber = 8

'global prolith objects to be used in this VB module
Dim ProlithApp As Prolith.Application
Dim ProDoc As Prolith.Document
Dim ProSimEng As Prolith.SimEngine
Dim ProlithWasRunning As Boolean
'matlab object
Dim MatLab As Object

'Excel objects
Dim xlApp As Excel.Application
Dim xlbook As Excel.Workbook
Dim xlsheet As Excel.Worksheet
'the main part of SUB1 is function ZiCD; the bottom form_load SUB
'is just for testing this function. the excel_
'objects are just for testing too, delete them when link

Public Function ZiCD(Zinterim() As Double, CDm As Double, stp As
Double, F As Double, E As Double, i As Integer) As Double()

'Zinumber is how many Zi to extract (=8)
Dim x(1, 1) As Double
Dim newfile As String
Dim ZernikeVarString As String
Dim zlocal(Zinumber - 1) As Double          '7=zinumber -1

ProSimEng.SetInput(29101, 0) = E          'exposure dose
ProSimEng.SetInput(29102, 0) = F          'focus

Open "C:\haolin1\ZrnFileHead.txt" For Input As 1
zrnfileheader = StrConv(InputB(LOF(1), 1), vbUnicode)
Set fso = CreateObject("Scripting.filesystemobject")

For indx2 = 0 To Zinumber - 1
```

```

zlocal(indx2) = Zinterim(indx2)
Next indx2

For indx1 = 0 To 1 Step 1
    zlocal(i) = zlocal(i) + indx1 * stp
    newfile = zrnfileheader

    For j = 0 To Zinumber - 1 Step 1
        ZernikeVarString = zlocal(j)
        newfile = newfile + ZernikeVarString + " "
        'xlsheet.cells(j + 1, 1) = Z(j)
    Next j

    'add rest 36-Zinumber Zero's at tail
    For j = Zinumber To 35 Step 1
        newfile = newfile + "0 "
    Next j

    newfile = newfile + Chr(13) + Chr(10)
    Set file1 = fso.createtextfile("C:\haolin1\ZrnFile.ZRN",
True)

    file1.write (newfile)
    file1.Close
    ProSimEng.AddFileData "C:\haolin1\ZrnFile.ZRN", 0
    ProSimEng.SingleRun
    x(indx1, 0) = zlocal(i)
    x(indx1, 1) = ProSimEng.GetOutput(8, -1) 'resist CD

    'xlsheet.cells(Indx1 + 1, 2) = x(Indx1, 1)
    Set file2 = fso.getfile("C:\haolin1\ZrnFile.ZRN")
    file2.Delete
Next indx1

ZiCD = x()

Close 1
Exit Function

End Function

Public Function SimplyCD(CDm As Double, F As Double, E As Double,
Z() As Double) As Double

Dim newfile As String
Dim ZernikeVarString As String

```

```

ProSimEng.SetInput(29101, 0) = E      'exposure dose
ProSimEng.SetInput(29102, 0) = F      'focus

Open "C:\haolin1\ZrnFileHead.txt" For Input As 1      'file
exists?
zrnfileheader = StrConv(InputB(LOF(1), 1), vbUnicode)
Set fso = CreateObject("Scripting.filesystemobject")

    newfile = zrnfileheader
    For j = 0 To Zinumber - 1 Step 1
        ZernikeVarString = Z(j)
        newfile = newfile + ZernikeVarString + " "
        'xlsheet.cells(j + 1, 1) = Z(j)
    Next j
    'add rest 36-Zinumber Zero's at tail
    For j = Zinumber To 35 Step 1
        newfile = newfile + "0 "
    Next j

    newfile = newfile + Chr(13) + Chr(10)
    Set file1 = fso.createtextfile("C:\haolin1\ZrnFile.ZRN",
True)

    file1.write (newfile)
    file1.Close
    ProSimEng.AddFileData "C:\haolin1\ZrnFile.ZRN", 0
    ProSimEng.SingleRun
    SimplyCD = ProSimEng.GetOutput(8, -1) 'arreal image CD
    'xlsheet.cells(Indx + 1, 2) = x(Indx, 1)
    Set file2 = fso.getfile("C:\haolin1\ZrnFile.ZRN")
    file2.Delete

Close 1
End Function

Private Sub Form_Load()

Dim F As Double
Dim E As Double
Dim i As Integer
Dim WaferCD(30) As Double
Dim CDo(30) As Double
Dim Zinterim(Zinumber - 1) As Double
Dim ZiCDallZiallFE() As Double
Dim tempZiCDarray() As Double

```

```

Dim FECDvb(30, 2) As Double
Dim MImag() As Double

Const CDm = 200 'not used in whole program. need to know how to
specify mask CD in Prolith

Const stp = 0.0005 'delta Zi to calculate derivative

'ReDim Zinterim(Zinumber - 1)

''connect to prolith
On Error Resume Next
Set ProlithApp = GetObject(, "Prolith.Application")
If Err.Number <> 0 Then
ProlithWasRunning = False
''if not running, launch prolith
Set ProlithApp = CreateObject("Prolith.Application")
ProlithApp.Visible = True
Err.Clear
Else
ProlithWasRunning = True
End If

Set prolithDoc = ProlithApp.ActiveDocument
Set ProSimEng = prolithDoc.SimulationEngine

'set input parameters
ProSimEng.SetInput(29100, 0) = 540      'resist thickness

ProSimEng.SetInput(29117, 0) = 0.7115   'resist para A
ProSimEng.SetInput(29118, 0) = 0.22     'resist para B
ProSimEng.SetInput(29119, 0) = 0.0194   'resist para C

'ProSimEng.SetInput(29101, 0) = 24      'exposure dose
'ProSimEng.SetInput(29102, 0) = -0.2    'focus
ProSimEng.SetInput(29105, 0) = 0.57     'NA
ProSimEng.SetInput(29106, 0) = 0.6      'partial coherence
ProSimEng.SetInput(29139, 0) = 248      'wavelength
'ProSimEng.SetInput(29116, 0) = 0.258   'Aerial image threshold

'launch Excel
Set xlApp = CreateObject("excel.application")
Set xlbook = xlApp.Workbooks.Add
Set xlsheet = xlbook.Worksheets(1)
xlApp.Visible = True

```

```

'xlsheet.Cells(1, 1) = "Zernike"
'xlsheet.Cells(1, 2) = "CD"
'Dim zz(35) As Double
'Dim zcd As Double
'For j = 0 To 35 Step 1
'zz(j) = 0
'Next j

'tempstr = ""
'zcd = SimplyCD(200, -0.2, 24, zz())
'  xlsheet.Cells(1, 1) = zcd

'Text1.Text = tempstr
*****
'Open  "C:\Program Files\Finle\Prolith\haolin\FECDvb.txt"  For
Input As 1
'zrnfileheader = StrConv(InputB(LOF(1), 1), vbUnicode)

For i = 0 To Zinumber - 1 Step 1 'reset Zinterim
Zinterim(i) = 0
Next i

Set MatLab = CreateObject("MatLab.Application")
Result = MatLab.Execute("load C:\haolin1\fecd.txt")      'file
exist?
Call MatLab.GetFullMatrix("fecd", "base", FECDvb, MImag)

'For ii = 1 To 31 Step 1          'test data transfer
from matlab to vb
'xlsheet.Cells(ii, 1) = FECDvb(ii - 1, 0)  'and load file
'xlsheet.Cells(ii, 2) = FECDvb(ii - 1, 1)
'xlsheet.Cells(ii, 3) = FECDvb(ii - 1, 2)
'Next ii

ReDim ZiCDallZiallFE(31 * Zinumber * 2 - 1, 1)

'FECDvb is waferCD file/array, whose format is 3 columns: F E CD
For FE = 0 To 30 Step 1
F = FECDvb(FE, 0)
E = FECDvb(FE, 1)
WaferCD(FE) = FECDvb(FE, 2)

CDo(FE) = SimplyCD(CDm, F, E, Zinterim()) 'transfer to matlab!

```

```

'xlsheet.Cells(1, 1) = CDo(FE)

'ReDim Preserve ZiCDallZiallFE((FE + 1) * n * 31 - 1, 1)

ReDim tempZiCDarray(1, 1)

For i = 0 To Zinumber - 1 Step 1
    tempZiCDarray = ZiCD(Zinterim(), CDm, stp, F, E, i)

    indx = FE * Zinumber * 2 + i * 2

    ' For ii = 1 To 31 Step 1
    'xlsheet.Cells(ii, 1) = tempZiCDarray(ii - 1, 0)
    'xlsheet.Cells(ii, 2) = tempZiCDarray(ii - 1, 1)
    ' Next ii

    For j = 0 To 1 Step 1
        ZiCDallZiallFE(indx + j, 0) = tempZiCDarray(j, 0)
        ZiCDallZiallFE(indx + j, 1) = tempZiCDarray(j, 1)
    Next j

Next i
Next FE

    'For ii = 1 To 2 Step 1
    'xlsheet.Cells(ii, 1) = tempZiCDarray(ii - 1, 0)
    ' xlsheet.Cells(ii, 2) = tempZiCDarray(ii - 1, 1)
    'Next ii

'xxx = 1
For ii = 1 To 496 Step 1
    xlsheet.Cells(ii, 1) = ZiCDallZiallFE(ii - 1, 0)
    xlsheet.Cells(ii, 2) = ZiCDallZiallFE(ii - 1, 1)
Next ii

    Call MatLab.PutFullMatrix("all", "base", ZiCDallZiallFE,
MImag)
    Result = MatLab.Execute("save C:\haolin1\allZiCD.txt all -
ascii")

    Call MatLab.PutFullMatrix("CDo", "base", CDo, MImag)
    Result = MatLab.Execute("save C:\haolin1\CDo.txt CDo -ascii")

End Sub

```