

Copyright © 2003, 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.

**AN OPTICAL METROLOGY SYSTEM  
FOR LITHOGRAPHY PROCESS  
MONITORING AND CONTROL**

by

Junwei Bao

Memorandum No. UCB/ERL M03/2

1 January 2003

**AN OPTICAL METROLOGY SYSTEM  
FOR LITHOGRAPHY PROCESS  
MONITORING AND CONTROL**

by

Junwei Bao

Memorandum No. UCB/ERL M03/2

1 January 2003

**ELECTRONICS RESEARCH LABORATORY**

College of Engineering  
University of California, Berkeley  
94720

# An Optical Metrology System for Lithography Process Monitoring and Control

by

Junwei Bao

B.S. (Peking University, Beijing, China) 1996

M.S. (University of California at Berkeley) 2000

A dissertation submitted in partial satisfaction of the requirements for  
the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Costas J. Spanos, Chair

Professor Andrew R. Neureuther

Professor Kameshwar Poolla

Spring, 2003

The dissertation of Junwei Bao is approved:

---

Chair Date

---

Date

---

Date

University of California, Berkeley

Spring, 2003

# **An Optical Metrology System for Lithography Process Monitoring and Control**

**Copyright © 2003**

**by**

**Junwei Bao**

# Abstract

An Optical Metrology System for Lithography Process Monitoring and Control

By

Junwei Bao

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Costas J. Spanos, Chair

As the semiconductor industry approaches the first sub-100 nm technology node on the SIA roadmap, process equipment and materials are stretched toward their limits, thus making the process very sensitive to even small perturbations of process conditions. At the same time, building and maintaining a semiconductor production fab has become increasingly expensive, and lithography alone can account for as much as one third of the overall chip manufacturing cost. Therefore, it is very important to monitor and control the lithography process in order to enhance the production yield. This thesis presents a framework for lithography process monitoring and control using the full line-edge profile information obtained from scatterometry, and discusses in detail the performance issues of scatterometry as they pertain to lithography control.

In this thesis we review various optical configurations of optical metrology tools that have been used for scatterometry, and analyze some related theoretical problems. The results show that signal matching with weighting based on signal noise or in the domain close to the raw detector signal can yield a greatly improved CD and thickness precision. These results provide

ample evidence that full-profile scatterometry is easily scalable for use on the 70 nm technology node on the International Technology Roadmap for Semiconductors.

Some experimental results on characterizing a scatterometry system used in a production environment are also presented. The accuracy and precision of the system are demonstrated on various applications. The scatterometry profiles and line widths of shallow trench isolation and polysilicon gate structures are sufficiently close to those measured with cross-section SEM and CD-SEM. Most importantly, however, the long-term precision of scatterometry is several times better than that of the CD-SEM.

It is still a big challenge to integrate scatterometry into the advanced process control (APC) framework. Research and development on using the wafer level and full-profile information of scatterometry is still in the early stages. A framework on lithography process monitoring and control using scatterometry is proposed first in this thesis. The simulation results show significant benefits in using CD and sidewall angle information obtained from scatterometry as controlled variables. Namely, the resulting CD and sidewall angle variations are more than 30% smaller than those with the traditional CD-based control method. The requirements and challenges in applying scatterometry in a real production environment are also discussed.

---

Professor Costas J. Spanos

Committee Chairman



# Table of Contents

<b>Chapter 1 Introduction.....</b>	<b>1</b>
1.1 Background and motivation.....	1
1.1.1 Background and motivation for inline process control.....	2
1.1.2 Scatterometry and its application to inline process control.....	4
1.2 Thesis organization.....	5
Reference.....	7
<b>Chapter 2 Background.....</b>	<b>8</b>
2.1 Introduction.....	8
2.2 Semiconductor process strategies.....	8
2.3 Optical imaging based microscopy.....	9
2.4 Overview of scatterometry technology.....	11
2.4.1 Sensitivity-based metrology.....	11
2.4.2 Review of simulation methods.....	12
2.4.3 Profile extraction methods.....	14
2.5 Summary.....	19
Reference.....	19
<b>Chapter 3 Theoretical Analysis of Optical Tools Used for Scatterometry.....</b>	<b>22</b>
3.1 Introduction.....	22
3.2 Optical metrology tools for scatterometry.....	22
3.2.1 Single-wavelength variable-angle reflectometer.....	22
3.2.2 Normal-incidence reflectometer.....	24
3.2.3 Spectroscopic ellipsometer.....	26
3.2.4 Other optical configurations.....	32
3.3 Study of signal matching domain based on noise analysis.....	33
3.3.1 Criteria for judging the goodness of a matching method.....	33
3.3.2 Noise analysis of SOPRA tool.....	33
3.3.3 Simulation of matching domain effect for SOPRA tool.....	36
3.4 The effect of azimuth angle on scatterometry.....	48
3.5 Summary.....	56

Reference.....	57
<b>Chapter 4 Scatterometry Experiment Data Analysis.....</b>	<b>59</b>
4.1 Introduction.....	59
4.2 Review of experimental work on scatterometry.....	59
4.2.1 Research on scatterometry in the early stage.....	59
4.2.2 Comparison between spectroscopic and 2- $\theta$ scatterometry.....	62
4.3 Specular spectroscopic scatterometry tool characterization.....	66
4.3.1 Introduction of applications.....	66
4.3.2 Library-based vs. regression method for profile extraction.....	68
4.3.3 Accuracy of the scatterometry system.....	71
4.3.4 Precision of the scatterometry system.....	74
4.3.5 Therma-Wave SE tool light spot size characterization.....	76
4.4 Scatterometry application on metal process.....	80
4.4.1 Scatterometry work on metal applications.....	80
4.4.2 Potential application to CMP end-point detection.....	86
4.5 Summary.....	87
Reference.....	88
<b>Chapter 5 Lithography Process Monitoring and Control Using Scatterometry.....</b>	<b>91</b>
5.1 Introduction.....	91
5.2 Review of previous work on lithography process monitoring and control.....	91
5.2.1 Background.....	91
5.2.2 CD-based lithography process control method.....	93
5.2.3 Recent development on lithography process control using scatterometry.....	94
5.3 Focus-exposure monitor experiment.....	95
5.3.1 Background.....	95
5.3.2 FEM wafer processing.....	96
5.3.3 FEM characterization using scatterometry.....	97
5.4 A simulation framework for lithography process control.....	103
5.4.1 A lithography process control framework using scatterometry.....	103
5.4.2 Lithography simulation and model parameter extraction.....	105
5.4.3 A simulation framework for lithography process control.....	107
5.4.4 Simulation results.....	110

5.4.5 Discussion.....	114
5.5 Conclusion.....	117
Reference.....	117
<b>Chapter 6 Requirements of Scatterometry for Lithography Process Control.....</b>	<b>120</b>
6.1 Introduction.....	120
6.2 Key elements of run-to-run process control framework.....	120
6.3 Requirements of scatterometry for lithography process control.....	122
6.3.1 Precision and accuracy.....	122
6.3.2 Sensitivity to fine profile features.....	125
6.3.3 Inline compatibility.....	125
6.3.4 Requirement of the semiconductor manufacture environment.....	126
6.4 Summary.....	126
Reference.....	127
<b>Chapter 7 Conclusions and Future Work.....</b>	<b>129</b>
7.1 Thesis summary.....	129
7.2 Future work.....	130
Reference.....	131

# List of Figures

Figure 1-1	Lithography process trends in history and prediction.....	1
Figure 1-2	Effect of metrology delay on control effectiveness.....	3
Figure 2-1	Illustration of the principle of the microscope.....	10
Figure 2-2	Illustration of grating profile slicing for RCWA simulation.....	13
Figure 2-3	Illustration of the regression method for profile extraction.....	15
Figure 2-4	Illustration of library-based method for profile extraction.....	17
Figure 3-1	Single wavelength, variable angle reflectometer.....	23
Figure 3-2	Illustration of broadband reflectometer configuration.....	24
Figure 3-3	Polarized broadband reflectometer.....	26
Figure 3-4	Illustration of rotating-polarizer ellipsometer setup.....	28
Figure 3-5	PSCA ellipsometer setup.....	30
Figure 3-6	Simulation flowchart for matching domain effect on profile extraction...	38
Figure 3-7	Grating structure used for matching domain study.....	40
Figure 3-8	Extracted profile parameters using signal matching at different domain..	41
Figure 3-9	Extracted profile parameter at different signal matching domain with wavelength weighting.....	45
Figure 3-10	Extracted profile parameter precision with different line widths.....	47
Figure 3-11	Incident light with same angle of incidence $\theta$ and symmetric azimuth angles.....	48
Figure 3-12	Ellipsometry signals measured at different azimuth angles.....	49
Figure 3-13	Ellipsometry signals at 650 nm for different azimuth angles.....	50
Figure 3-14	Top-down view of Figure 3-10.....	51
Figure 3-15	GOF vs. azimuth angle.....	56

Figure 4-1	Stack structure used for early scatterometry study.....	60
Figure 4-2	Structure used to illustrate scatterometry sensitivity.....	62
Figure 4-3	Example signal of a 2- $\theta$ scatterometry system.....	63
Figure 4-4	Signal sensitivity of the 2- $\theta$ scatterometry system.....	64
Figure 4-5	Sample signal of a spectroscopic scatterometry system.....	65
Figure 4-6	Signal sensitivity of the spectroscopic scatterometry system.....	65
Figure 4-7	Standard deviation of repeatability measurement signals on the ellipsometry system.....	66
Figure 4-8	Polysilicon gate DI structure.....	67
Figure 4-9	Polysilicon FI structure.....	68
Figure 4-10	Shallow Trench Isolation FI structure.....	69
Figure 4-11	Measured versus simulated signals using regression and library-based Methods.....	70
Figure 4-12	Measured STI FI structure profile using scatterometry.....	71
Figure 4-13	Cross-section SEM image of STI FI structure.....	72
Figure 4-14	Comparison of measured bottom CD using scatterometry and KLA- Tencor CD-SEM for all the dies on a production wafer.....	73
Figure 4-15	Comparison of long-term repeatability results using scatterometry and CD-SEM.....	74
Figure 4-16	Long-term repeatability measurement signals.....	75
Figure 4-17	Testing grating for light spot size characterization.....	77
Figure 4-18	Bottom CD and GOF for horizontal line scan.....	79
Figure 4-19	Bottom CD and GOF for vertical line scan.....	79
Figure 4-20	Example of metal interconnect profile variation.....	80

Figure 4-21	Illustration of metal DI structure.....	80
Figure 4-22	Extracted FEM resist profiles plotted as their positions on the wafer....	81
Figure 4-23	Comparison of CD-SEM and scatterometry bottom CD.....	82
Figure 4-24	Illustration of the metal FI structure.....	83
Figure 4-25	Extracted FEM metal profiles plotted as their positions on the wafer....	84
Figure 4-26	Comparison of scatterometry and CD SEM top CD for metal features..	85
Figure 4-27	Deviations between CD-SEM and scatterometry bottom CD versus sidewall angle.....	86
Figure 4-28	One-dimensional structure for CMP end-point detection.....	87
Figure 5-1	Lithography process flow and sources of process variation.....	92
Figure 5-2	Manual focus-dose calibration procedure with FEM.....	95
Figure 5-3	Resist grating structure used for FEM characterization.....	96
Figure 5-4	FEM resist profile across the full wafer map.....	96
Figure 5-5	Tilt-angle SEM images of selected sites of the FEM wafer.....	98
Figure 5-6	Profile with standing wave and the approximated smoothed profile.....	99
Figure 5-7	Rectangle slices for RCWA simulation.....	99
Figure 5-8	Signal responses for the profile with standing waves and the approximated smoothed profile.....	100
Figure 5-9	Measured signals from neighboring dice on the FEM wafer.....	100
Figure 5-10	Bossung curves for the FEM wafer using scatterometry middle CD measurement.....	102
Figure 5-11	Bossung curves generated using CD-SEM measurement.....	103
Figure 5-12	The control framework using scatterometry.....	104
Figure 5-13	Example of simulator tuning using profile matching.....	107
Figure 5-14	Simulation of the control framework using scatterometry.....	108

Figure 5-15	Parameter disturbance of the simulated lithography process.....	111
Figure 5-16	Extracted control-parameter values vs. their corresponding input values.....	112
Figure 5-17	Comparison of resulting bottom CD and sidewall angle with and without control.....	113
Figure 5-18	Signal sensitivity to focus drift for profiles with 160 nm CD and 200 nm CD.....	116
Figure 6-1	Structure used for accuracy study.....	123
Figure 6-2	Parameter offsets for accuracy study.....	124

# List of Tables

Table 3-1	Summary of profile extraction without wavelength weighting .....	42
Table 3-2	Summary of extraction with wavelength weighting .....	44
Table 3-3	Comparison of standard deviations of extracted profile parameters without and with weighting .....	44
Table 4-1	Comparison of scatterometry and other metrology tools.....	61
Table 4-2	Comparison of scatterometry and other metrology tools.....	69
Table 4-3	Library parameters for polysilicon FI structure.....	70
Table 4-4	Correlation of measured spectrum to site 8 for horizontal scan.....	78
Table 4-5	Correlation of measured spectrum to site 8 for vertical scan.....	78
Table 5-1	Process parameters for the FEM wafer.....	97
Table 5-2	Standard Deviation of the disturbed parameter noise.....	110
Table 5-3	Comparison of standard deviations with and without control.....	113
Table 5-4	Scatterometry signal sensitivity to focus drift.....	115
Table 6-1	Nominal and $1\sigma$ for the sample structure generation.....	123
Table 6-2	Extracted profile parameter offset and $1\sigma$ .....	124

## Acknowledgements

I would like to express my deepest gratitude to my research advisor, Professor Costas J. Spanos, for his guidance and supervision, and for the opportunities he has created for me during my graduate research and career. His vision of the semiconductor manufacturing industry was instrumental in helping me to determine the direction of my research. I thank him for his consistent guidance, suggestion, enlightenment and help in every detail of my research.

I would also like to thank Professor Andrew R. Neureuther and Professor Kameshwar Poolla for their insightful advices and inputs to this work and for reviewing my dissertation. Their experiences in the field of lithography and process control provide invaluable to me. I thank Professor Vivek Subramanian and Professor David Brillinger for serving as committee members in my qualifying exam.

Special thanks to my mentors and buddies at Timbre Technologies, Xinhui Niu, Nickhil Jakatdar, and Alan Nolet. Their friendship, advice and encouragement make my experience at Timbre unforgettable in my life. I acknowledge many valuable discussions and collaborations with my colleagues in the Applied Research Group, Vi Vuong, Srinivas Doddi, Emmanuel Drege, and Wen Jin. I also thank Lawrence Lane, Johan Saleh, Kelly Barry, Youngho Chung, Wenge Yang, Craig Michaelis, and all other Timbre employees, for their help and friendship.

Thanks for the past and present BCAMers, Jason Cain, Runzi Chang, Rowen Chen, Mareike Claassen, Weng Loong Foong, Mason Freed, Paul Friedberg, Anna Ison, Nickhil Jakatdar, Michiel Kruger, Jae-Wook Lee, John Musacchio, Xinhui Niu, Jiangxin Wang, Jing Xue, Haolin

Zhang, Dongwu Zhao, for their help and support. Many thanks to the staff members of the Berkeley Microelectronics Laboratory, Katalin Voros, Bob Hamilton, Sia Parsa, Kim Chan, Xiaofan Meng, for their support on some of the experimental work for the dissertation. I also acknowledge the staff of the EECS department, Ruth Gjerde, Dianne Chang, Charlotte Jones, T.K. Chan, Tim Duncan, for their help and support.

I thank my former industry colleagues, Bhanwar Singh, Joe Bendik, Matt Hankinson, Carmen Morales, Cristina Cheung, for sharing their knowledge and experience with me during my internship at AMD and National Semiconductor Corporation.

I thank my friends at Berkeley, Peiqi Xuan, Weidong Luo, Yuwei Li, Yashesh Shroff, Yu Cao, Qiang Lu, Chunlong Guo, Wei Mao, Yunjian Jiang, Kanyu Cao, and all other friends. They made my life at Berkeley colorful and memorable.

I would like to thank my Mom and sister for their love and support during my graduate years. Finally and most importantly, I would like to thank my wife, Haijing, for her companion and love during my thesis writing. Together we have overcome many difficulties. I thank her for her love, encouragement, and understanding. No matter what happened, she was always behind me. I am looking forward to having a wonderful life with her.

This work was funded by the State of California SMART program under research contract SM97-01, and by the following participating companies: Advanced Energy, ASML, Atmel Corp., Advanced Micro Devices, Applied Materials, Asyst Technologies Inc., BOC Edwards, Cymer, Etec Systems Inc., Intel Corporation, KLA-TENCOR, Lam Research Corp., Nanometrics, Inc, Nikon Research Corp., Novellus Systems Inc., Silicon Valley Group, Texas Instruments Inc., and Tokyo Electron America.

# Chapter 1 Introduction

## 1.1 Background and motivation

The demand for faster and larger scale integrated circuits (IC) has pushed the continuing down-scaling of the transistors printed on the silicon wafer. Although many people have doubted the continued validity of Moore's Law, the semiconductor industry has actually accelerated the rate of circuit feature shrinkage, and will be able to keep a similar pace in the next few years.

As IC production equipment and materials are stretched towards their limits, semiconductor processes become more expensive and very sensitive to disturbances. Photolithography is seen by many as the key driver in feature shrinkage. Therefore, it is very important to monitor and control the lithography process to enhance the production yield. This thesis presents a

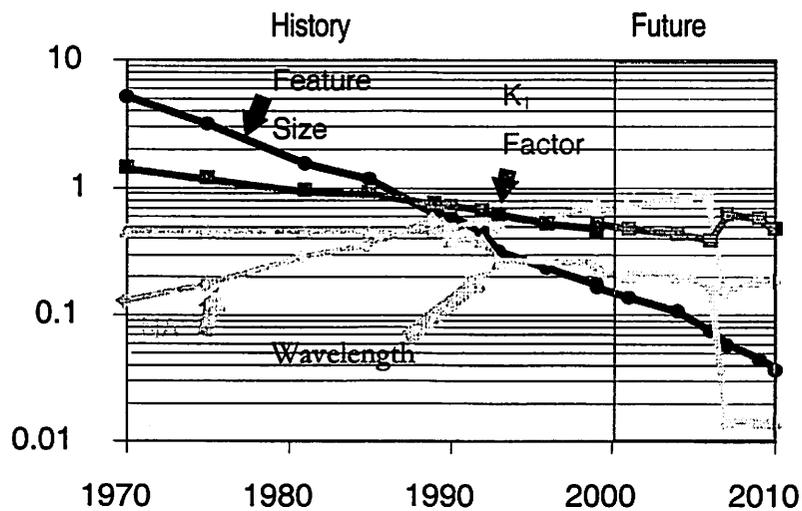


Figure 1-1 Lithography process trends in history and prediction. Source: W.G. Oldham, SPIE Microlithography 2000.

simulation framework for lithography process monitoring and control using the full-profile information obtained from scatterometry, and discusses in detail the performance and progress of scatterometry technology.

### 1.1.1 Background and motivation for inline process control

Optical lithography is the key process step that enables the continuing feature size reduction. As shown in Figure 1-1, the smallest resist line width printed on production wafers has been decreasing exponentially, and it will be below 100 nm during 2003. For optical lithography, thanks to continuing efforts to reduce aberrations, the resolution limit is mainly due to diffraction, and it can be represented as

$$w = k_1 \frac{\lambda}{NA} \quad (1.1)$$

where  $\lambda$  is the wavelength,  $NA$  is the numerical aperture of the optical system, and  $k_1$  is a scaling factor. As we can also see from Figure 1-1, the stepper light source wavelength of 193 nm will be used for a few more years, and we are approaching the 1.0 limit of non-emersion optics for  $NA$ . The  $k_1$  factor will continue to decrease in order to enable the line width reduction.

There are many resolution enhancement techniques that can help reduce the  $k_1$  factor, such as phase shifting masks (PSM), optical proximity corrections (OPC), off-axis illumination, and resist property improvement, etc. [1.1]. On the other hand, many of these techniques and  $NA$  scaling extend optical patterning process toward physical limits, and make the process

extremely sensitive to small perturbations of process parameters. For example, the depth of focus of an optical system can be represented as

$$d.o.f. \propto \frac{\lambda}{NA^2} \quad (1.2)$$

As  $NA$  increases, the depth of focus reduces, making the process more sensitive to focus drift.

There have been two general strategies of process control. The first one focuses on keeping the process stable. Once a process is developed, all the process parameters are kept constant during production, and statistical process control (SPC) can be used to monitor the process output to check if it is within the statistical limit. As the process becomes increasingly sensitive to the disturbance of operating parameters, the process window becomes smaller and smaller, necessitating active process control. In the second strategy, process is monitored *in situ* or inline with integrated metrology sensors, and feedback and/or feed-forward control is used to

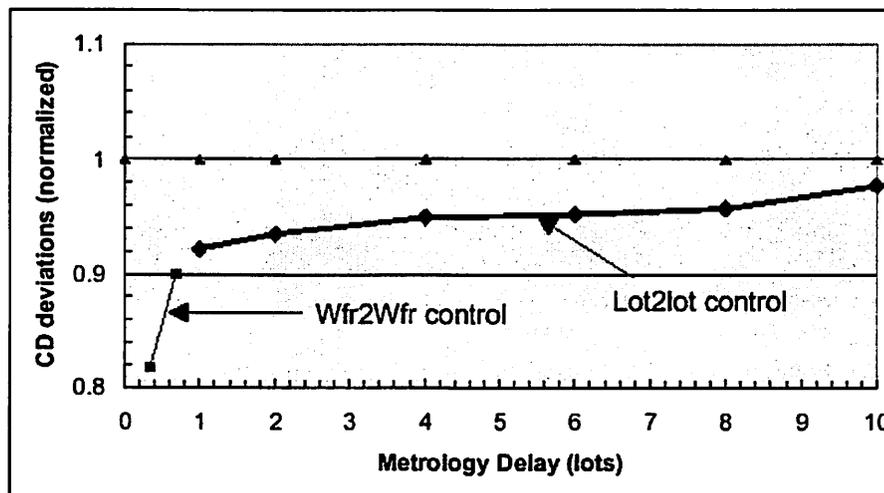


Figure 1-2 Effect of metrology delay on control effectiveness. Source: M. Hankinson, AEC/APC 2000

compensate the drift detected by the sensors.

When process drift occurs, it is preferable to detect it and react to it as early as possible to prevent the extra cost. Figure 1-2 shows the effect of metrology delay on control effectiveness presented by Matt Hankinson on AEC/APC Symposium XII [1.2]. The simulation shows that the maximum benefit in lot-to-lot control is 8% in critical dimension (CD) deviation reduction, and 18% in wafer-to-wafer control. The study was done on a 200 mm wafer process, and the cost savings could be more significant in the incoming 300 mm wafer process technology.

### **1.1.2 Scatterometry and its application to inline process control**

The developed resist line profile is the ultimate output of the lithography step. The traditional critical dimension scanning electron microscope (CD-SEM) is used as the metrology tool to monitor the lithography process. However, because the wafer needs to be measured in a vacuum environment, CD-SEM metrology is a batch process. Furthermore, since only a top-down view of line width information is available, it is difficult to monitor the process in detail on a wafer-by-wafer basis.

Scatterometry is an ideal candidate for inline process monitoring. The optical instrument can be made small enough to fit in the space of the bake module on a wafer track, enabling a true wafer-by-wafer metrology scheme. Furthermore, as we will see in later chapters, the quality (full profile versus top-town view) and quantity (accuracy, precision, and throughput) of the collected data underscore its clear advantage for inline application.

Advanced process control (APC) for lithography has been widely used recently in state-of-the-art semiconductor production lines [1.3]. However, most of the control frameworks are based on the measurement data obtained using CD-SEM in the lot-to-lot level. Since scatterometry

can measure the features right after the resist is developed, and full-profile information is available, the control framework needs to be adapted and improved to make better use of the scatterometry data.

## 1.2 Thesis organization

This thesis presents the latest development on the scatterometry technology and discusses its integration into the process control framework.

In chapter 2, we will first introduce the status of advanced process control in the semiconductor industry, and then review the history of optical metrology used for semiconductor processing. Traditional optical metrology, such as image-based optical microscopy, forms directly from the object a virtual image, and uses a geometric scale to measure the image and get the size of the object. The resolution, however, is limited by the diffraction of the optical imaging process. Sensitivity-based optical metrology, such as scatterometry, can obtain higher resolution by matching the simulated signal with the measured one based on signal sensitivity to small feature changes. We will review the popular methods for optical scattering simulation, as well as various profile extraction techniques.

Many of the optical metrology tools used for scatterometry were originally designed and optimized for thin film metrology. It is essential to analyze their performance when applied to patterned feature measurement. In Chapter 3, we use the SOPRA broadband ellipsometer as an example for analyzing the system noise performance. Many of the ellipsometer systems designed for thin film metrology do not have strict specification on azimuth angle alignment, while this becomes a key factor when doing patterned feature measurement. We will therefore analyze the effect of azimuth angle to the measured profiles.

Chapter 4 presents some experimental results on scatterometry. After reviewing some pioneering scatterometry work done on the  $2\theta$  and other hardware configurations, we show some results from characterizing a scatterometry system using a Thermo-Wave's Optical-Probe 5240 ellipsometer, including accuracy, precision, light spot size management, as well as the effect of "library-based" versus "regression" methods for profile extraction. Finally, we demonstrate the potential application of scatterometry on metal layer processes, and foresee its possible application to chemical-mechanical planarization (CMP).

Chapter 5 illustrates how to fit scatterometry into the process control framework. We first demonstrate some experimental results on focus and exposure monitors using scatterometry. After some discussion on lithography process parameter extraction, we present a simulation framework for lithography process monitoring and control using scatterometry. The simulated results show that the bottom CD and sidewall angle variations are reduced when doing process control using scatterometry's full-profile information, compared to the traditional open-loop or CD-based process control methods.

In chapter 6, we will discuss the scatterometry requirements when implementing it in process control. We will first introduce some basic concepts of process control, and analyze the four elements of the run-to-run control framework. Metrology is one of the key elements of the control framework. We will discuss the current status of the scatterometry technology in meeting the requirement of the control framework, and what is needed to apply scatterometry in the production environment.

Chapter 7 provides concluding remarks and outlines the challenges and future work in scatterometry as an integral part of process control.

## Reference

- [1.1] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors," <http://public.itrs.net/Files/2002Update/Home.pdf>, 2002.
- [1.2] M. Hankinson, "Intelligent Control of the Semiconductor Patterning Process," AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [1.3] AEC/APC Symposium XIV, Salt Lake City, UT, September 7-12, 2002.

# Chapter 2 Background

## 2.1 Introduction

Advanced process control (APC) becomes increasingly important in state-of-the-art semiconductor production fabs. Much progress has been made in applying various control algorithms to semiconductor processes, especially to the lithography module. In this chapter we introduce the status of advanced process control in the semiconductor industry. We also give a brief introduction of the traditional optical methods, and then overview scatterometry-based metrology.

Metrology is one of the essential elements in the process control framework. Ever since the discovery of the magnification capability of convex lenses, people have tried to build various optical instruments to measure small objects. Traditional optical metrology, such as image-based optical microscopy, forms a virtual image of the object, and uses a geometrical scale to measure the image to get the size of the object. One can also measure small objects indirectly by detecting the optical response to the variation of feature sizes. As long as the mapping from the structure to the optical signal is one-to-one, one can inverse this relation empirically or through simulation to get the actual structure shape.

## 2.2 Semiconductor process control strategies

There are four essential functions for APC as defined by SEMATECH in 1994 [2.1]: fault detection, fault classification or diagnosis, fault prognosis, and process control. Fault detection identifies that an operating condition is different from a normal one; fault classification classifies what caused the difference; fault prognosis predicts whether the fault will happen in

the future processes; and process control determines values for the manipulated variables to achieve desired performance of the control variable. In broader definition [2.1], APC monitors the process and determines the condition and methods to adjust the process so that it remains on target. It also monitors and optimizes the adjustment to keep the process stable and robust, and generates an alarm if a violation occurs.

The lithography process is a multiple-input-multiple-output (MIMO) process. The input variables are process parameters, such as exposure dose, focus, PEB temperature and time, etc., and the output variables are resist profile parameters. There are two general control strategies for MIMO processes. In the past decade, there has been much progress in applying single-input-single-output (SISO) process control to the lithography process [2.2, 2.3]. Usually resist profile CD and exposure dose are selected as controlled and manipulated variables, respectively, because CD is the most critical parameter, and it is most affected by exposure dose. However, as the requirement to resist profile by later pattern transfer steps becomes increasingly strict, it is more important to monitor and control the full resist profile, which is affected by multiple process parameters. Multiple-input-multiple-output (MIMO) control demands new requirements for the key elements of process control, as will be discussed in Chapter 5 and Chapter 6.

## **2.3 Optical imaging based microscopy**

The apparent size of an object is determined by the size of its image on the retina, which depends on the angle the object subtends for an unaided eye. Optical imaging based microscopes use a combination of objectives and eyepieces to enlarge this angle by forming a

virtual image. With the help of scale and crosshair on the eyepieces, one can measure the size of an object up to the resolution of the virtual image.

As shown in Figure 2-1 [2.4], the image of an object P in front of the objective, P', is formed between the objective and the eyepiece. The Eyepiece then forms an enlarged virtual image P'' at the favorable distance (about 25 cm) to the eye. The subtended angle of the image P'' is much larger than that of the object P if P is placed at the same location.

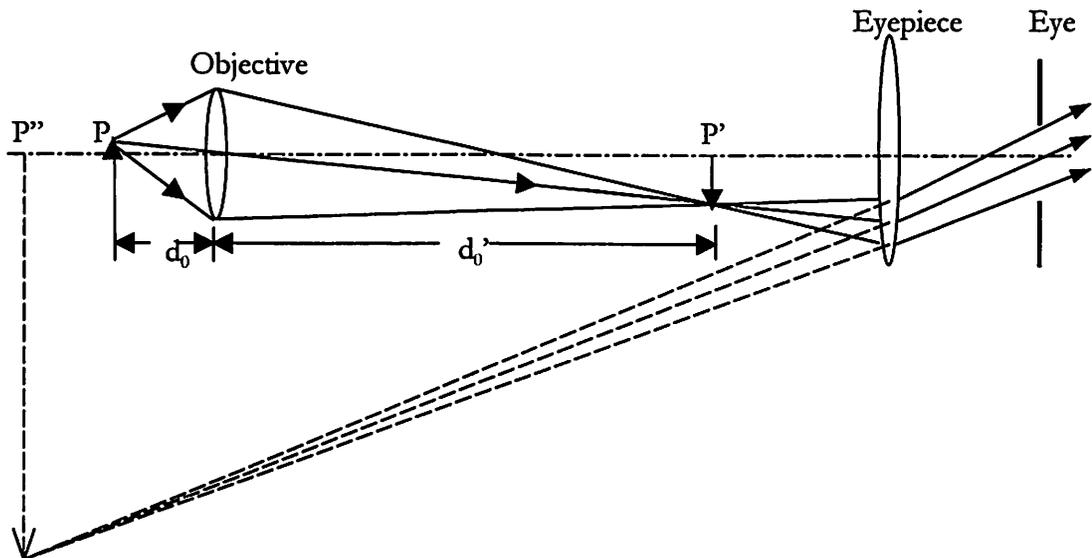


Figure 2-1 Illustration of the principle of the microscope

However, due to the diffraction at the apertures in the optical system, the detail of the object that is smaller in size than approximately half of the light wavelength becomes “fuzzy.” The resolving power of the microscope is determined by the resolution of the images limited by diffraction, which can be represented by [2.4]

$$R \sim C \cdot \frac{\lambda_0}{NA} \quad (2.1)$$

where  $\lambda_0$  is the wavelength in vacuum, NA is the numerical aperture of the objective, and C is a factor depending on the coherence of illumination, where  $C = 0.61$  for self-luminous objects, and  $C = 0.82$  for coherent illumination if Rayleigh criterion is used [2.4].

## 2.4 Overview of scatterometry technology

### 2.4.1 Sensitivity-based metrology

The imaging-based microscopy can be considered as a “direct” method, where the size of an object is measured directly on its image. Another way of measuring feature size is sensitivity-based metrology. This method usually involves detecting the signal response of a sample structure from an incident light beam. As long as the variation of the feature size causes detectable change on the response signal, and this mapping relation is unique within a certain metrology range, the feature size can be “extracted” from the measured signal. Sensitivity-based optical metrology has been widely used in many applications, such as ellipsometry and reflectometry for thin film characterization, where the film thickness that can be resolved (usually in sub-nanometers) is much smaller than the wavelength of light.

Many people have employed empirical methods to monitor and characterize patterned features using optical metrology tools [2.5]. During the initial phase, wafers are measured on both the optical tool and other referencing metrology tools such as critical dimension scanning electronic microscope (CD-SEM) or atomic force microscope (AFM). The diffraction signals and the corresponding reference results are used to build a model based on neural network or other statistical methods. During the measuring phase, the optical signal is fed into the model to extract the profile information. With this method, one does not need to simulate the optical response from the sample structure, which makes it applicable to more general types of sample

structures, such as the actual device region in DRAM cells. However, the model built normally only covers the ranges of the data from pilot runs, and a prohibitively large number of pilot runs are usually needed to build a model with reasonable accuracy. Therefore, the empirical method is mainly used for fault detection and classification. Accurate simulation of the optical response is needed to further improve the profile extraction result.

## 2.4.2 Review of simulation methods

A one-dimensional grating is one of the key testing structures used for optical metrology. Although a typical circuit usually consists of lines in different densities and orientations, they can be represented well enough using testing grating structures of different line-space ratios and orientations. Furthermore, the measurement region size (about 50  $\mu\text{m}$  to 100  $\mu\text{m}$ ) is usually much larger than the feature size (typically less than 1  $\mu\text{m}$ ), so repetitive structures in the measurement light spot would strengthen “the signal content of the structure” (i.e., make the effect of the structure more eminent) in the response signal. In many cases, with additional knowledge of periodical condition of the structure, one can more efficiently construct and solve the optical response problem. Therefore, in this thesis, we mainly focus on simulation and metrology on one-dimensional grating structures.

Various techniques have been developed to solve the problem of electromagnetic wave diffraction from periodical structures. The scalar wave modeling method only offers an analytical solution for simple binary gratings [2.4]. The Rayleigh method decomposes the field above and below the grating region using plane waves in the direction of diffraction orders, and assumes that the field in the grating region can also be decomposed using the same basis of waves [2.6]. But this assumption has been proved to be invalid for sinusoidal gratings with a

high aspect ratio [2.7]. With the advances of computer technology, many numerical techniques have been developed to rigorously solve the grating diffraction problem. Most of these methods start from Maxwell equations, formulate either integral or differential equations to take into account the periodic change of refractive index, and solve these equations

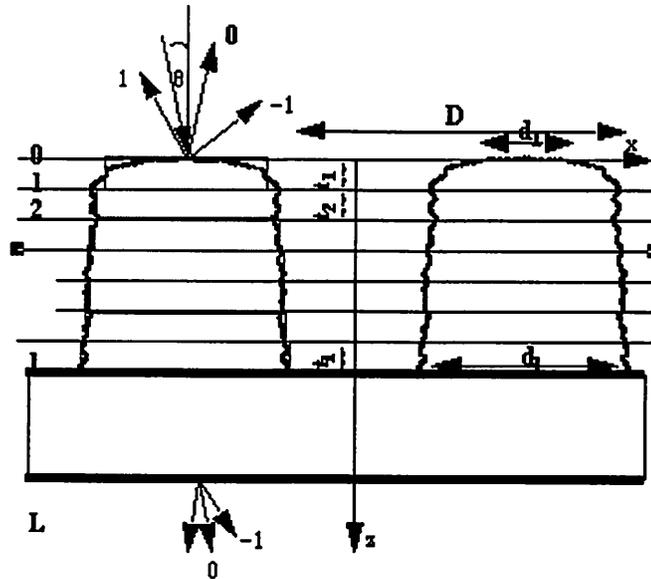


Figure 2-2 Illustration of grating profile slicing for RCWA simulation

numerically. Some methods are limited to gratings of certain sizes [2.8], while others do not have such limitations. The integral method was used by Neureuther, Zaki [2.9], and Maystre [2.10], and the differential method was proposed by Neviere, Cadilhac, and Petit [2.11]. Finite-difference time-domain (FDTD) algorithm was introduced by Yee [2.12]. However, the integral method is hard to formulate and implement; the differential method is iterative, and sometimes it takes a very long time to converge.

Moharam and Gaylord proposed the Rigorous Coupled Wave Analysis (RCWA) method in 1980 [2.13]. It is a non-iterative method that is adequate for non-planar periodic dielectric surfaces. As shown in Figure 2-2, the grating is first sliced into multiple layers and each layer is

approximated by a rectangular slab. The periodic permittivity distribution in each slab of the grating region is first expanded using the Fourier series. Also, from the “Floquet condition” [2.14], the field in the grating region can be expressed as Fourier expansion. Substituting them into the Maxwell equations, with some expansion of the product of the infinite series, comparing coefficients of corresponding terms, and using boundary conditions between neighboring slab layers, we can derive a set of first-order coupled wave equations in matrix form. Its homogeneous solution can be obtained by solving for the eigenvalue and eigenvector of the matrix. On the other hand, similar to the Rayleigh method, the electromagnetic field above and below the grating region can be represented using linear superposition of plane waves in the direction of various orders of grating diffraction. This can serve as the boundary conditions to determine the specific solution of the diffraction efficiency. This method can be very accurate if enough diffraction orders are preserved and enough rectangular slices are used to approximate the grating structure profile. With some later developments in the formulation of modeling TM wave diffraction to improve the order convergence performance [2.15, 2.16], RCWA has shown advantages in solving one-dimensional grating diffraction problems over other techniques in terms of computation speed and accuracy, and it has been widely used in various application areas [2.17, 2.18, 2.19].

### **2.4.3 Profile extraction methods**

An optical response simulator only solves the “forward” problem, i.e., given the dimensions of a certain structure, one can simulate its optical response, but not *vice versa*. What an optical metrology tool directly measures is the optical response. Thus the “reverse” problem needs to be solved to extract the feature size. This type of problem also exists in many other areas, such as ellipsometry and reflectometry for thin film characterization.

The regression method [2.20] is one of the most popular methods for the “reverse” problem. As illustrated in Figure 2-3, a set of initial parameter values are input to the object response simulator, and then the simulated response is compared with the measured response signal. If they are not close enough (as determined by an appropriately chosen convergence criterion), the result is fed into the optimizer to generate a new set of trial object parameter values for the next iteration, until the simulated signal matches the measured signal well enough. This is a typical flow of an optimization process, where the goal is to minimize the difference between the measured and the simulated signals. This difference is often called “cost function.”

Many algorithms have been developed for this type of minimization process. It is important to know that the optical response from the multi-dimensional parameters of the object is often highly non-linear, so usually there are many local minima for the cost function across the

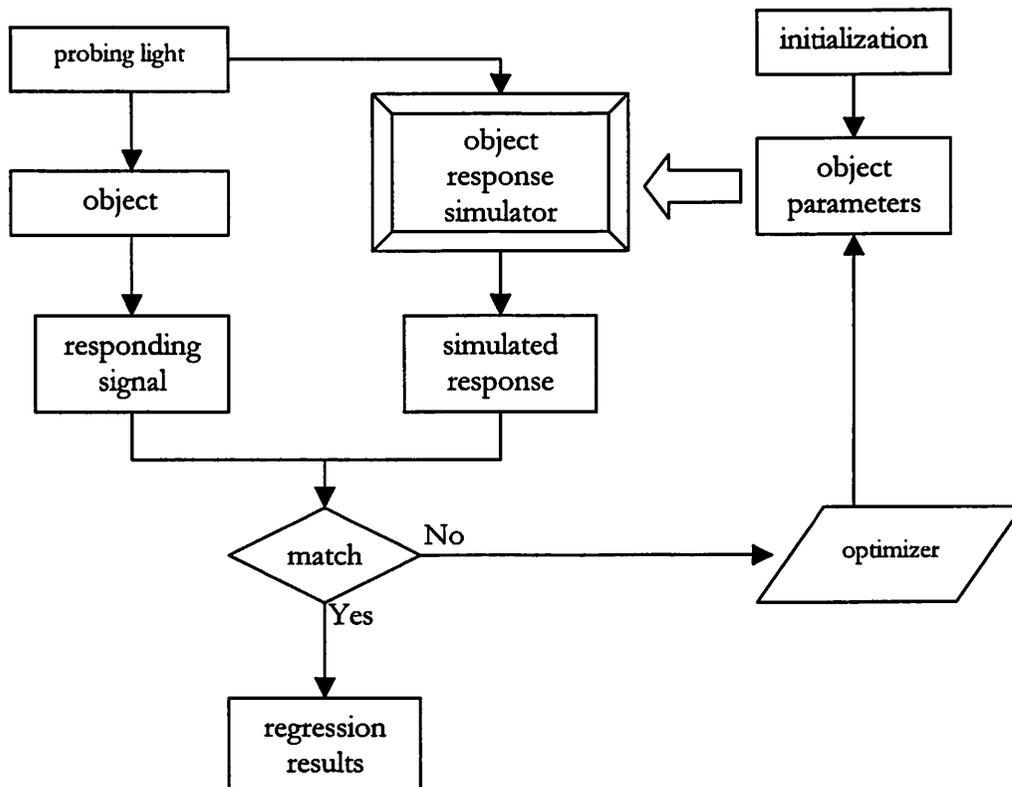


Figure 2-3 Illustration of the regression method for profile extraction

parameter domain. Local optimization techniques, such as the Gauss-Newton method [2.21] and the Levenberg-Marquardt (L-M) method [2.22], may end up converging at one of the local minima. Some global optimization methods, such as simulated annealing [2.23], can do a good job finding the global minimum, but they suffer from the problem of slow convergence rate comparing to the local optimization methods. A heuristic combination of global and local optimization techniques can usually achieve better performance.

Depending on the characteristics of the cost function and the starting point, it can take from a few to many hundreds of iterations to converge in the regression process. Within each iteration, one or more optical response simulations are needed, which may take from a few seconds to several minutes on the fastest single-CPU computer. Therefore, the whole regression process usually takes tens of seconds or longer. Using advanced computer techniques, such as distributed computing or parallel computing, we can significantly reduce the regression time [2.24], but this requires much more investment on computer hardware and software.

Another method for profile extraction is the library-based method [2.17, 2.18], as illustrated in Figure 2-4. During the off-line process, one can first input film stack information as well as the expected profile features for each section such as rectangle, trapezoid, hour-glass shape, T-topping, footing, etc., to define the structure. Profiles are then parameterized to describe these geometric shapes. For example, a rectangle can be defined by its width and height, and a trapezoid can be defined by top width, bottom width, and height, etc. The ranges of these parameters are determined based on the process window, and their resolutions are determined based on the trade-off between the library size and the required accuracies of these parameters. The combination of these multi-dimensional parameter values is used to generate the profile

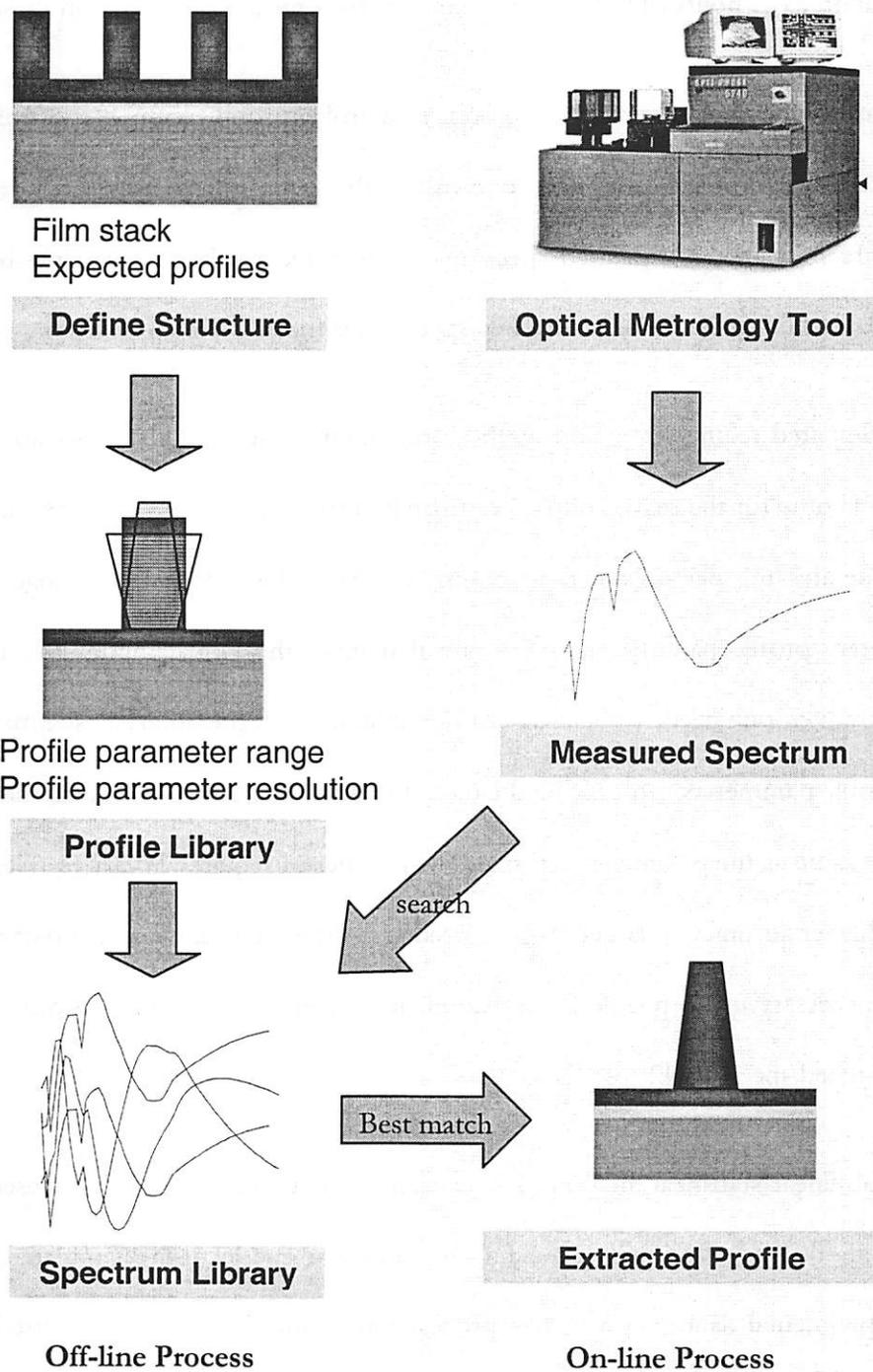


Figure 2-4 Illustration of library-based method for profile extraction

library, and the optical response from each of these profiles is simulated to populate the spectrum library. In order to cover the typical process variation with accuracy of about 1 nm

for 5 to 6 profile parameters, a few hundred thousand profiles are needed, and it usually takes hundreds of hours of CPU time to simulate the optical response from these profiles.

During the production process, optical metrology tools, such as spectroscopic ellipsometers and reflectometers, are used to measure the optical responses from testing structures on wafers. Then the measured spectrum is compared to those in the pre-built library, and the profile that gives the best- matched spectrum is the extracted profile.

Compared to the regression method, the library-based method does not require a very long CPU time for the search of the best-matched spectrum during production, and the profile that generates the best-matched spectrum is usually close (within the range of one step of the library profile parameter) to the one that gives the global minimum of the cost function. However, one needs to be aware of the limitations of the library-based method. The extracted profile parameters can only be the discrete grid values, and the measurement accuracy is about the same as the parameter step size. The number of required library profiles can easily blow up if higher accuracy or larger range is needed for the multi-dimensional parameters. Therefore, it is necessary to do profile interpolation or refinement around the profile that gives the best-matched spectrum [2.25].

Non-linear statistical methods, such as neural network, are also widely used [2.26,2.27]. These methods typically involve building a non-linear model using simulated spectra or sample experimental data over a certain process range, and feeding the measured spectrum into the model to extract the profile during the on-line phase. Usually this type of method suffers from poor accuracy due to the limitation of the model, and they are often used for the purpose of fault detection or fault classification.

## 2.5 Summary

In this chapter, we introduced the traditional optical metrology methods, where one measures the optical image of the object formed by the metrology instrument. Sensitivity-based metrology can overcome the difficulty of image formation in the traditional optical metrology, and can greatly extend the resolution beyond the diffraction limit. After a brief overview of the simulation methods, various profile extraction techniques, such as the regression method, library-based method, and non-linear statistical method, were discussed. In the next chapter we will do some theoretical analysis of the optical tools used for scatterometry.

### Reference

- [2.1] T. Edgar, S.J. Qin, W.J. Campbell, "Run-to-Run Control and Fault Detection," Short course, AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [2.2] E. Robertson, S. George, J.Holt, "Advanced CD Control Architecture and Implementation in Motorola," AEC/APC Symposium IX, Lake Tahoe, CA, September 20-24, 1997.
- [2.3] J. Stuber, F. Pagette, S. Tang, "Device Dependent Run-to-Run Control of Transistor Critical Dimension by Manipulating Photolithography Exposure Settings," AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [2.4] M. Born, E. Wolf, "Principle of Optics," Cambridge University Press, 7<sup>th</sup> edition, 1999.
- [2.5] N. Benesch, A. Hettwer, C. Schneider, L. Pfitzner, H. Ryssel, "Phi-Scatterometry for Integrated Linewidth Control in DRAM Manufacturing," Proceedings of IEEE International Symposium on Semiconductor Manufacturing, 129-132, 2001.
- [2.6] Lord Rayleigh, "On the Dynamical Theory of Gratings," Proc. Of the Royal Society A, vol. 79, 399-416, 1907.

- [2.7] R. Petit and M. Cadilhac, "Sur la diffraction d'une onde plane par un réseau infini-ment conducteur," C. R. Acad. Sci., B, vol. 262, no. 7, 468-471, 1966.
- [2.8] K. Knop, "Rigorous Diffraction Theory for Transmission Phase Gratings with Deep Rectangular Grooves," Journal of the Optical Society of America, vol.68, (no.9), 1206-1210, Sept. 1978.
- [2.9] A. R. Neureuther and K. Zaki, "Numerical Methods for the Analysis of Scattering from Nonplanar Periodic Structures," Intn'l URSI Symposium on Electromagnetic Waves, Stresa, Italy, 282-285, 1969.
- [2.10] D. Maystre, "A New General Integral Theory for Dielectric Coated Gratings," J. Opt. Soc. Am., vol. 68, no. 4, 490-495, April 1978.
- [2.11] R. Petit (editor), "Electromagnetic Theory of Gratings," Berlin: Springer-Verlag, 1980.
- [2.12] K.S. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media," IEEE Trans. Ant. Prop. vol. 14, pp. 302-307 (1966).
- [2.13] M. G. Moharam and T. K. Gaylord, "Rigorous Coupled-Wave Analysis of Planar-Grating Diffraction," J. Opt. Soc. Am., vol. 71, 811-818, July 1981.
- [2.14] Leon Brillouin, "Wave Propagation in Periodic Structures", Dover Publications, Inc., 1951.
- [2.15] P. Lalanne and G. M. Morris, "Highly Improved Convergence of the Coupled-Wave Method for TM Polarization," J. Opt. Soc. Am. A, 779-784, 1996.
- [2.16] L. Li and C. Haggans, "Convergence of the Coupled-Wave Method for Metallic Lamellar Diffraction Gratings," J. Opt. Soc. Am. A, 1184-1189, June 1993.
- [2.17] C.J. Raymond, M.R. Murnane, S.L. Prins, S.S. Naqvi, J.R. McNeil, "Multi-parameter Process Metrology Using Scatterometry," SPIE Vol. 2638, Optical Characterization Techniques for High-Performance Microelectronic Device Manufacturing II, 84-93, 1995.
- [2.18] X. Niu, N. Jakatdar, J. Bao, C.J. Spanos, S. Yedur, "Specular Spectroscopic Scatterometry in DUV Lithography," SPIE vol. 3677, Metrology, Inspection and Process Control for Microlithography XIII, 159-168, 1999.

- [2.19] J. Bischoff, J.W. Baumgart, H. Truckenbrodt, J.J. Bauer, "Photoresist Metrology based on Light Scattering," SPIE Vol. 2725, Metrology, Inspection and Process Control for Microlithography X, 678-689, 1996.
- [2.20] D.M. Bates, D.G. Watts, "Nonlinear Regression Analysis and Its Applications," New York: Wiley, 1988.
- [2.21] G.C. Kossakes, N. Kyriakopoulos, "An Improvement on the Newton-Gauss Algorithm," IEEE Transactions on Circuits and Systems, vol.CAS-24, (no.5), 271-273, May 1977.
- [2.22] D. Marquardt, "An Algorithm for Least Squares Estimation of Nonlinear Parameters," SIAM J. Appl. Math. Vol. 11, 431-441, 1963.
- [2.23] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, "Optimization by Simulated Annealing," Science 220, 4598 (May), 671-680, 1983.
- [2.24] J. Opsal, H. Chu, Y. Wen, Y.C. Chang, G. Li, "Fundamental Solutions for Real-time Optical CD Metrology," SPIE vol. 4689, Metrology, Inspection and Process Control for Microlithography XIII, 163-176, 2002.
- [2.25] K.C. Johnson, F.E. Stanke, "Database Interpolation Method for Optical Measurement of Diffraction Microstructures," United States Patent Application, No. 20020038196, US Patent and Trademark Office, March 28, 2002.
- [2.26] S. Robert, A. Mure-Ravaud, D. Lacour, "Characterization of Optical Diffraction Gratings by Use of a Neural Method," Journal of the Optical Society of America A (Optics, Image Science and Vision), vol.19, (no.1), Opt. Soc. America, 24-32, Jan. 2002.
- [2.27] J. Saarinen, I. Kallioniemi, A. Niinisto, A.T. Friberg, "Surface Roughness Measurement with Optical Scatterometry [and Neural Network Model]," Proceedings of the SPIE - The International Society for Optical Engineering, vol.3897, (Advanced Photonic Sensors and Applications, Singapore, 30 Nov.-3 Dec. 1999.) SPIE-Int. Soc. Opt. Eng, 570-577, 1999.

# Chapter 3 Theoretical Analysis of Optical Tools Used for Scatterometry

## 3.1 Introduction

In Chapter 2 we reviewed various profile extraction methods for scatterometry. In this chapter, we first introduce popular optical metrology tools that can be used for scatterometry. Many of these tools were originally developed for thin film metrology. However, many new issues arise when using these tools for scatterometry. We use the SOPRA broadband ellipsometer as an example to analyze the effect of signal matching domain on extracted profile noise. When doing simulation of signal response from grating structures, we always assume that the incidence plane is perpendicular to the grating lines. In a real ellipsometer system, this condition may not hold due to the difficulty of wafer stage rotation control. We will analyze the effect of the azimuth angle on the measured ellipsometer signals.

## 3.2 Optical metrology tools for scatterometry

### 3.2.1 Single-wavelength variable-angle reflectometer

Single-wavelength variable-angle reflectometer is the first optical configuration used on commercial scatterometry systems [3.1]. Intuitively, one can imagine that in order to “see” the profile of a one-dimensional grating, the system should shine a beam of light perpendicular to the direction of the grating lines, and analyze the reflectance from the grating at multiple angles of incidence. Figure 3-1 illustrates a typical single-wavelength variable-angle reflectometer system. A laser light source directs single-wavelength light on the sample structure after

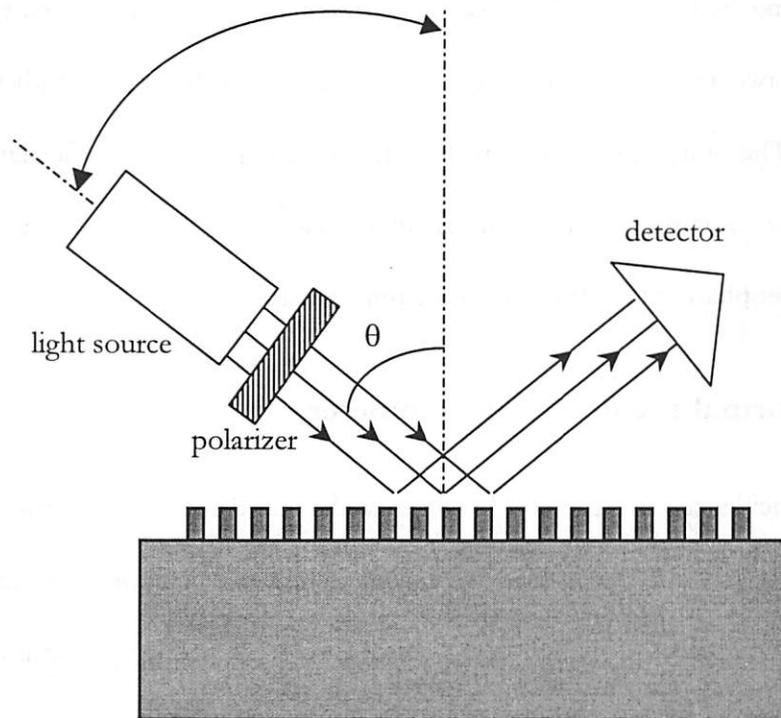


Figure 3-1 Single wavelength, variable angle reflectometer

passing a polarizer. Depending on the grating pitch and the wavelength of the light, there can be multiple orders of diffraction from the grating, (which are not drawn in the figure,) but only the 0<sup>th</sup> order of diffracted light is collected by the detector. The angle of incidence  $\theta$  is varied, and the detector angle varies accordingly as well. Therefore, this configuration is also called 2- $\theta$  scatterometry.

Since a laser source can be used for 2- $\theta$  scatterometry, its optical setup is relatively easy compared to configurations with broadband light sources, and the signal to noise ratio can be quite high. Typically He-Ne laser or semiconductor lasers at a wavelength of about 600 nm are used as light sources. The key drawback of the 2- $\theta$  scatterometry configuration is also due to its single-wavelength light source. For most of the grating structures, the sensitive wavelength range can vary from UV to IR depending on the structure, and very often it does not cover the wavelength of the 2- $\theta$  scatterometry light source. Furthermore, it is unable to distinguish

neighboring film stacks with similar refractive indices at the measurement wavelength, while rarely do two materials have similar indices across the full wavelength range for broadband systems. The other disadvantage is that only the intensity of the reflectance signal is obtained, which may contain less “profile information” than those systems such as the ellipsometer that can also get phase information from the reflectance.

### 3.2.2 Normal-incidence reflectometer

Normal-incidence reflectometry has been widely used for thin film thickness measurements. Recently, it has also been used by a few commercial vendors for scatterometry [3.2]. As illustrated in Figure 3-2, light coming out of the broadband lamp is split into two beams. One

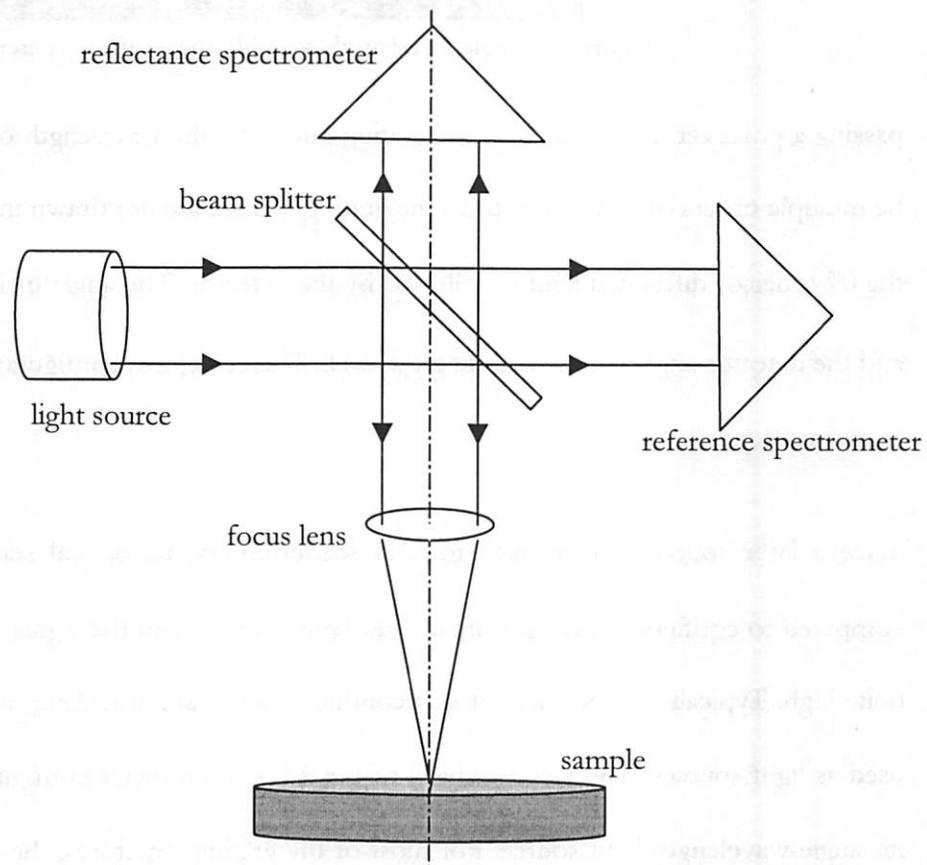


Figure 3-2 Illustration of broadband reflectometer configuration

serves as a reference, the other is focused by the lens on the sample, and the reflected light is collected by the lens and directed to the reflectance spectrometer. By comparing the intensity at these two detectors, the system can calculate the reflectivity at various wavelengths. Because of the simplicity of its optics, the reflectometer can be made very compact with high throughput, thus is ideal for integrated metrology. However, only the combined reflectance of the TE (transverse electric, meaning that the E-field vector is parallel to the grating lines) and TM (transverse magnetic, when the E-field is perpendicular to the grating lines) lights are detected, so a lot of information is not preserved in the combined signal. Thus it is difficult to resolve detailed profile features with a reflectometer because of the low signal sensitivity.

A commonly used method for decoupling the TE and TM reflectance is to add a polarizer in the light path. Since there are different optical responses for these two polarization states, one can measure both the intensity and phase responses by adjusting the polarizer angle. An illustration of the optical system design used by Nanometrics is shown in Figure 3-3 [3.3].

If the angle between the polarization direction of the polarizer and the grating lines is  $\varphi$ , then the overall reflectivity can be written as

$$R(\varphi) = r_s^2 \cos^4 \varphi + r_p^2 \sin^4 \varphi + 2r_s r_p \cos \Delta \cos^2 \varphi \sin^2 \varphi \quad (3.1)$$

where  $\tilde{r}_s = r_s e^{i\delta_1}$  and  $\tilde{r}_p = r_p e^{i\delta_2}$  are the complex reflectivity for TE and TM waves, respectively, and  $\Delta = \delta_1 - \delta_2$ .

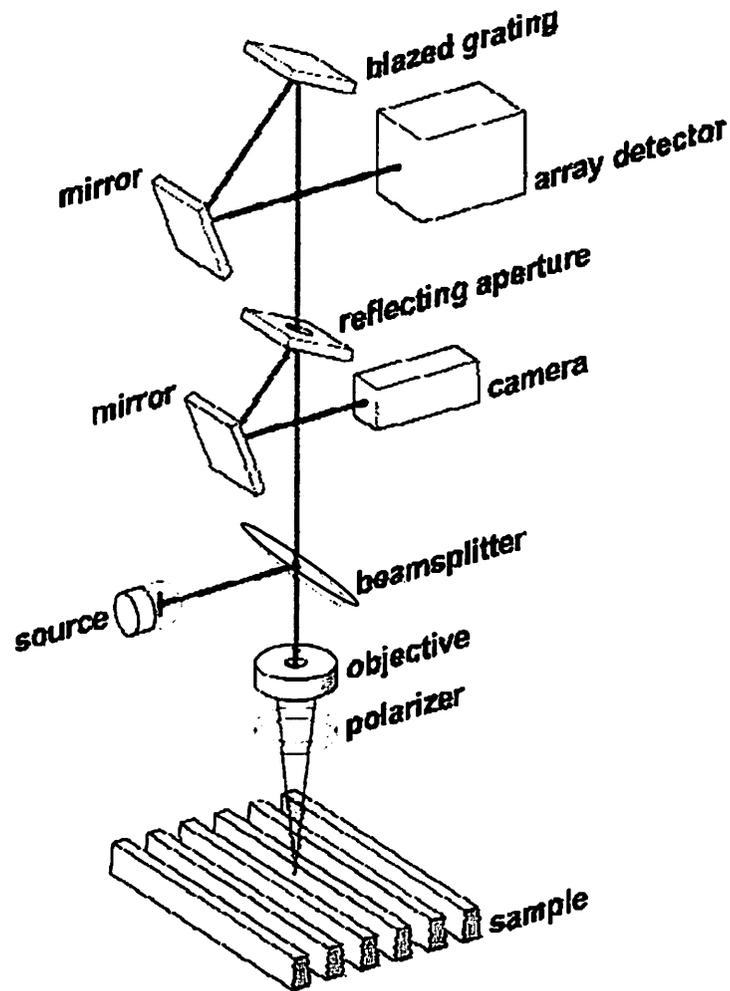


Figure 3-3 Polarized broadband reflectometer. Source: J. Holden, AEC/APC symposium 2001

By measuring  $R(\varphi)$  at multiple  $\varphi$  angles, one can calculate  $r_s$ ,  $r_p$ , and  $\cos \Delta$ , which can be used to extract the profile information.

### 3.2.3 Spectroscopic ellipsometer

Spectroscopic ellipsometry is widely used for thin film characterization and as of the last couple of years for scatterometry as well. When linearly polarized light shines on the sample, the reflected light becomes elliptically polarized due to the different magnitude and phase responses of TM and TE light. We can represent the ratio of the responses as

$$\frac{\tilde{r}_p}{\tilde{r}_s} = \tan \Psi \cdot e^{i\Delta} \quad (3.2)$$

where  $\tilde{r}_p$  and  $\tilde{r}_s$  are the complex reflectance of TM and TE light, respectively. Typical commercial ellipsometer systems obtain the  $\Psi$  and  $\Delta$  signals indirectly through transformations of measured intensity signals, and, unlike a reflectometer, there is no need for an absolute reference. Since both the intensity response  $\Psi$  and the phase response  $\Delta$  are measured, more information about the sample can be extracted from these signals.

The rotation-polarizer configuration has been used by many state-of-the-art commercial ellipsometer systems, such as SOPRA's GESP5 and KLA-Tencor's F5. The basic optical configuration is illustrated in Figure 3-4. The optical path consists of the broadband light source, two rotatable polarizing filters known as the polarizer and the analyzer, the sample, and the spectrometer. During the measurement, the analyzer stays at a certain position where the angle between the polarization direction and the incidence plane is  $A$ , and the polarizer rotates continuously to create time-variant signal at the spectrometer. When the angle between the polarizer polarization direction and the incidence plane is  $P$ , the resulting electrical fields at the end of the optical path can be represented using Jones matrices as [3.4]

$$\begin{pmatrix} \tilde{E}_p \\ \tilde{E}_s \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos A & \sin A \\ -\sin A & \cos A \end{pmatrix} \begin{pmatrix} \tilde{r}_p & 0 \\ 0 & \tilde{r}_s \end{pmatrix} \begin{pmatrix} \cos P & -\sin P \\ \sin P & \cos P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{E}_0 \\ \tilde{E}_0 \end{pmatrix}, \quad (3.3)$$

where the matrices on the right hand side of the equation represent the effects of analyzer, coordinate rotation for analyzer, sample, coordinate rotation for polarizer, polarizer, and light source, respectively, from left to right. Then the light intensity can be simplified as

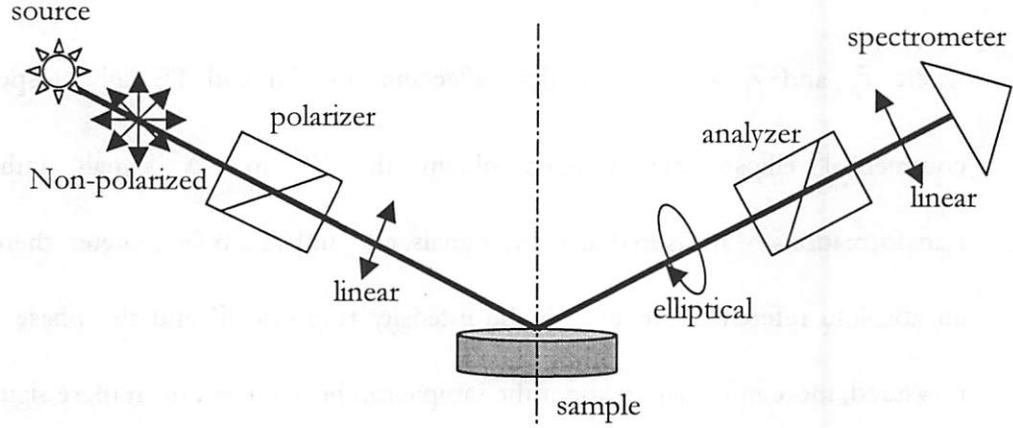


Figure 3-4 Illustration of rotating-polarizer ellipsometer setup

$$I(P) = |\tilde{E}_p|^2 + |\tilde{E}_s|^2 = I_0(1 + \alpha \cos 2P + \beta \sin 2P), \quad (3.4)$$

where

$$\alpha = \frac{\tan^2 \Psi - \tan^2 A}{\tan^2 \Psi + \tan^2 A}, \quad (3.5)$$

$$\beta = \frac{2 \cos \Delta \tan \Psi \tan A}{\tan^2 \Psi + \tan^2 A}, \quad (3.6)$$

$$I_0 = \frac{1}{2} |\tilde{E}_0 \tilde{r}_s|^2 \cos^2 A (\tan^2 \Psi + \tan^2 A), \quad (3.7)$$

and  $\tan \Psi$  is the same as that defined in equation (3.2).

The time-variant signal is sampled for a few periods while the polarizer is rotating, then Hadamard transform is used to extract the coefficients  $\alpha$  and  $\beta$  for each wavelength. The signal is integrated every quarter of the half-turn of the polarizer:

$$S_1 = \int_0^{\pi/4} I(P) dP = \frac{I_0}{2} \left( \frac{\pi}{2} + \alpha + \beta \right), \quad (3.8)$$

$$S_2 = \int_{\pi/4}^{\pi/2} I(P) dP = \frac{I_0}{2} \left( \frac{\pi}{2} - \alpha + \beta \right), \quad (3.9)$$

$$S_3 = \int_{\pi/2}^{3\pi/4} I(P) dP = \frac{I_0}{2} \left( \frac{\pi}{2} - \alpha - \beta \right), \quad (3.10)$$

$$S_4 = \int_{3\pi/4}^{\pi} I(P) dP = \frac{I_0}{2} \left( \frac{\pi}{2} + \alpha - \beta \right). \quad (3.11)$$

Then  $\alpha$  and  $\beta$  can be expressed as

$$\alpha = \frac{1}{2I_0} (S_1 - S_2 - S_3 + S_4), \quad (3.12)$$

$$\beta = \frac{1}{2I_0} (S_1 + S_2 - S_3 - S_4), \quad (3.13)$$

and  $\tan \Psi$  and  $\cos \Delta$  can be derived from equations (3.5) and (3.6):

$$\tan \Psi = \tan A \sqrt{\frac{1+\alpha}{1-\alpha}}, \quad (3.14)$$

$$\cos \Delta = \frac{\beta}{\sqrt{1-\alpha^2}}. \quad (3.15)$$

The advantage of the rotating polarizer technique is that it is optically and mechanically simple.

Only polarizers and focusing lenses (focusing reflecting mirrors in production configuration)

are used in the light path, and these optical elements are relatively easy to make and characterize. Furthermore, since the analyzer angle is fixed during the measurement, the spectrometer does not need to be insensitive to the polarization of the incidence light. But as we can see in the equations, only  $\cos \Delta$  can be measured instead of  $\Delta$ , so it is difficult to determine the sign of  $\Delta$  using this technique. With further noise analysis, we can also find out that the  $\cos \Delta$  signal can be very noisy when its absolute value is close to 1.

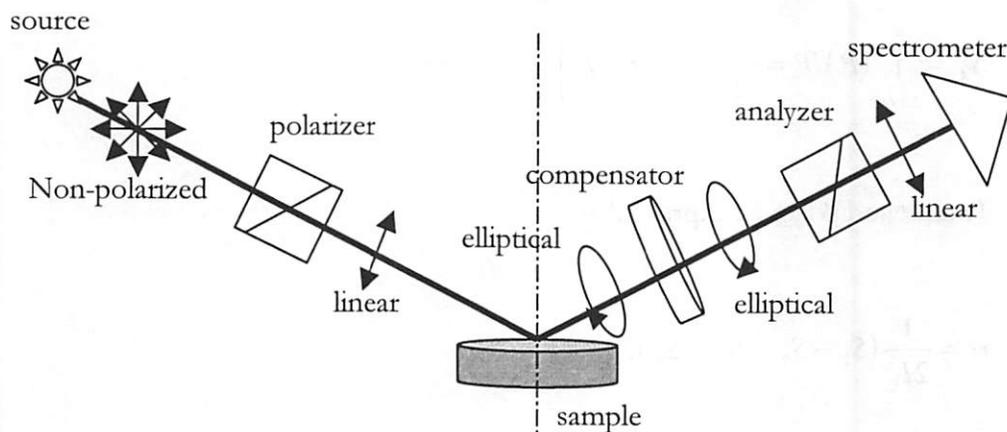


Figure 3-5 PSCA ellipsometer setup

The other popular ellipsometer configuration involves a rotating compensator. Compared to the rotating polarizer configuration, there is a wave plate (compensator) inserted between the polarizer and analyzer in the optical path, as shown in Figure 3-5, so this type of system is also called PSCA (polarizer-sample-compensator-analyzer) ellipsometer. During the measurement process, both the polarizer and the analyzer are fixed, while the compensator rotates to create a time-variant signal. When the fast wave direction of the compensator is at angle  $C$  from the incidence plane, using a similar method to that for the rotating polarizer configuration, the electrical field at the spectrometer can be represented using Jones matrices as

$$\begin{pmatrix} \tilde{E}_p \\ \tilde{E}_s \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos(A-C) & \sin(A-C) \\ -\sin(A-C) & \cos(A-C) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix} \begin{pmatrix} \cos C & \sin C \\ -\sin C & \cos C \end{pmatrix} \quad (3.16)$$

$$\times \begin{pmatrix} \tilde{r}_p & 0 \\ 0 & \tilde{r}_s \end{pmatrix} \begin{pmatrix} \cos P & -\sin P \\ \sin P & \cos P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} E_0 \\ E_0 \end{pmatrix}$$

where the matrices on the right hand side of the equation represent the effects of analyzer, coordinate rotation for analyzer, compensator phase shift, coordinate rotation for compensator, sample, coordinate rotation for polarizer, polarizer, and light source, respectively, from left to right, and  $C = \omega t$  if the compensator rotates at angular speed of  $\omega$ .

Then the time-variant intensity of the light at the spectrometer can be written in the form of

$$I(t) = a_0 + a_2 \cos 2\omega t + b_2 \sin 2\omega t + a_4 \cos 4\omega t + b_4 \sin 4\omega t, \quad (3.17)$$

where parameters  $a_i$  and  $b_i$  ( $i = 0, 2, 4$ ) are Fourier coefficients of compensator rotating frequencies. If analyzer angle  $A$  is set to  $A = 0$ , then these coefficients can be simplified as

$$a_0 = \frac{1}{2} E_0^2 |\tilde{r}_s|^2 \left[ \sin^2 \frac{\delta}{2} \sin^2 P + \tan^2 \Psi \left( 1 + \cos^2 \frac{\delta}{2} \right) \cos^2 P \right], \quad (3.18)$$

$$a_2 = 0, \quad (3.19)$$

$$b_2 = -\frac{1}{2} E_0^2 |\tilde{r}_s|^2 \tan \Psi \sin 2P \sin \delta \sin \Delta, \quad (3.20)$$

$$a_4 = \frac{1}{2} E_0^2 |\tilde{r}_s|^2 \sin^2 \frac{\delta}{2} [\tan^2 \Psi \cos^2 P - \sin^2 P], \quad (3.21)$$

$$b_4 = \frac{1}{2} E_0^2 |\tilde{r}_s|^2 \tan \Psi \sin 2P \sin^2 \frac{\delta}{2} \cos \Delta, \quad (3.22)$$

where  $\Delta$  is the same as that defined in equation (3.2). The Fourier coefficients can be extracted from the time-variant intensity signal, and then  $\Psi$  and  $\Delta$ , if needed, can be calculated using equations (3.18) ~ (3.22).

As we can see from these equations, both  $\sin \Delta$  and  $\cos \Delta$  appear in the equations for Fourier coefficients, so the sign of  $\Delta$  can be determined during the calculation.

Due to their popularity in the scatterometry applications, we will focus on rotating polarizer and PSCA ellipsometer configurations for the analysis in the remainder part of this thesis.

### 3.2.4 Other optical configurations

In addition to those already introduced in this section, many other configurations of optical sensors can also be used for scatterometry, such as broadband polarized reflectometer [3.5], variable azimuth angle reflectometer [3.6], multiple-wavelength multiple-AOI ellipsometer [3.7], etc. As discussed in Chapter 2, as long as the optical response is sensitive to the feature variation on the sample, and this response relation is unique within the metrology range, theoretically the sensor can be used for scatterometry. When determining which sensor configuration is the best for scatterometry, one needs to balance various factors, such as optical sensitivity of signal to measured features, system noise characteristic, ease of optical response simulation, physical size, etc., depending on the requirement of the application. Very often, one also needs to consider cost, popularity and robustness of the configuration for a production environment. In the next section, we will focus on noise analysis using the rotating polarizer setup.

## 3.3 Study of signal matching domain based on noise analysis

### 3.3.1 Criteria for judging the goodness of a matching method

As discussed in Chapter 2, when extracting profiles from the measured signals using the regression method or the library-based method, the goal is to make the simulated signal match the measured one as closely as possible. Usually the sum-of-squares-error of the two signals is used to judge how closely they match each other. However, this matching metric implies that the measured signals have independent identical distribution of noise across the wavelengths, and the typical signals obtained from the optical tool may not have this attribute.

As we can see in section 3.2, the ellipsometry signals  $\tan\Psi$  and  $\cos\Delta$  are measured indirectly through other physical variables. For example, a rotation polarizer ellipsometer system measures the four S signals, and they can be transformed to  $\alpha$  and  $\beta$  domain or  $\tan\Psi$  and  $\cos\Delta$  domain. In the next subsection, we will analyze the noise distribution of a SOPRA system at different domains of the optical signals.

### 3.3.2 Noise analysis of Sopra tool

SOPRA ellipsometer configuration utilizes a rotating polarizer. The raw signal directly seen by the detector is the time-variant intensity, and it is integrated for every quarter of the half-turn of the polarizer, as shown in formulae (3.8-3.11). So we can treat the signals  $S_1$ ,  $S_2$ ,  $S_3$  and  $S_4$  as the direct output of the detector, and their noise can be shown to be identically and independently distributed. The source of noise can be from light source intensity fluctuation, thermal and other noise of the detector, and random error due to the polarizer rotation angle control.

From equations (3.8-3.11), we have

$$I_0 = \frac{1}{\pi}(S_1 + S_2 + S_3 + S_4), \quad (3.22)$$

$$\alpha = \frac{1}{2I_0}(S_1 - S_2 - S_3 + S_4) = \frac{\pi}{2} \frac{S_1 - S_2 - S_3 + S_4}{S_1 + S_2 + S_3 + S_4}, \quad (3.23)$$

$$\beta = \frac{1}{2I_0}(S_1 + S_2 - S_3 - S_4) = \frac{\pi}{2} \frac{S_1 + S_2 - S_3 - S_4}{S_1 + S_2 + S_3 + S_4}. \quad (3.24)$$

Then

$$\frac{\Delta\alpha}{\pi} = \frac{S_2 + S_3}{I_0^2}(\Delta S_1 + \Delta S_4) - \frac{S_1 + S_4}{I_0^2}(\Delta S_2 + \Delta S_3), \quad (3.25)$$

$$\frac{\Delta\beta}{\pi} = \frac{S_3 + S_4}{I_0^2}(\Delta S_1 + \Delta S_2) - \frac{S_1 + S_2}{I_0^2}(\Delta S_3 + \Delta S_4). \quad (3.26)$$

Assuming that the variances of the four intensity integrals are

$$\sigma_{S_1}^2 = \sigma_{S_2}^2 = \sigma_{S_3}^2 = \sigma_{S_4}^2 \equiv \sigma^2, \quad (3.27)$$

then the variances of  $\alpha$  and  $\beta$  can be represented as

$$\sigma_\alpha^2 = \pi^2 \frac{(S_2 + S_3)^2 + (S_1 + S_4)^2}{I_0^4} \sigma^2 = \frac{(\pi^2 + 4\alpha^2)\pi^2}{2I_0^2} \sigma^2, \quad (3.28)$$

$$\sigma_\beta^2 = \pi^2 \frac{(S_1 + S_2)^2 + (S_3 + S_4)^2}{I_0^4} \sigma^2 = \frac{(\pi^2 + 4\beta^2)\pi^2}{2I_0^2} \sigma^2. \quad (3.29)$$

In equation (3.27), we assume the variance of the four S signals are the same, because the major component of the detector noise is proportional to the integration time, and the S signals have the same integration time.

From equations (3.28) and (3.29), we can see that the variances of  $\alpha$  and  $\beta$  are not identical any more. If one wants to do signal matching in the  $(\alpha, \beta)$  space, proper weighting of signals at different wavelengths is needed to reduce the noise effect.

Since  $\tan \Psi$  and  $\cos \Delta$  are “standard” signals for ellipsometry equipment, much work has been done on signal matching in the  $\tan \Psi$  and  $\cos \Delta$  domain or  $\log \tan \Psi$  and  $\cos \Delta$  domain. Using a similar method, from equations (3.14) and (3.15), we have the variances of these signals as follows:

$$\sigma_{\tan \Psi}^2 = \frac{\tan^2 A}{(1-\alpha)^3(1+\alpha)} \sigma_{\alpha}^2 = \frac{\tan^2 A}{(1-\alpha)^3(1+\alpha)} \frac{(\pi^2 + 4\alpha^2)\pi^2}{2I_0^2} \sigma^2, \quad (3.30)$$

$$\sigma_{\log \tan \Psi}^2 = \frac{1}{\tan^2 \Psi} \sigma_{\tan \Psi}^2 = \frac{1}{(1-\alpha^2)^2} \frac{(\pi^2 + 4\alpha^2)\pi^2}{2I_0^2} \sigma^2, \quad (3.31)$$

$$\sigma_{\cos \Delta}^2 = \frac{1}{1-\alpha^2} \sigma_{\beta}^2 + \frac{\alpha^2 \beta^2}{(1-\alpha^2)^3} \sigma_{\alpha}^2 = \left[ \frac{1}{1-\alpha^2} \frac{(\pi^2 + 4\beta^2)}{2I_0^2} + \frac{\alpha^2 \beta^2}{(1-\alpha^2)^3} \frac{(\pi^2 + 4\alpha^2)}{2I_0^2} \right] \pi^2 \sigma^2. \quad (3.32)$$

Many SOPRA ellipsometers have two measurement modes: scanning mode and CCD mode. The scanning mode uses diffraction grating and a photomultiplier tube to detect optical response wavelength-by-wavelength, and it is possible to optimize the analyzer angle for each wavelength to maximize the signal-to-noise ratio [3.4] with the tracking method, i.e., set the analyzer angle  $A$  close to the  $\Psi$  value. In practice, this is achieved by doing a test

measurement first using the analyzer angle of the previous wavelength, then measure the actual signal after adjusting the analyzer angle according to the test measurement. If the tracking method is used, then  $\alpha \cong 0$ , and  $\cos\Delta \cong \beta$ , and equations (3.30-3.32) can be simplified to the following form:

$$\sigma_{\tan \Psi}^2 \cong \frac{\pi^4}{2I_0^2} \tan^2 \Psi \sigma^2, \quad (3.33)$$

$$\sigma_{\log \tan \Psi}^2 \cong \frac{\pi^4}{2I_0^2} \sigma^2, \quad (3.34)$$

$$\sigma_{\cos \Delta}^2 \cong \frac{(\pi^2 + 4\cos^2 \Delta)\pi^2}{2I_0^2} \sigma^2. \quad (3.35)$$

As we can see from the above equations,  $\log \tan \Psi$  shows better<sup>1</sup> noise behavior than  $\tan \Psi$ , especially when sharp peaks exist in the  $\tan \Psi$  signal. This is the reason why in the early works of scatterometry, better results were achieved when doing signal matching in the  $\log \tan \Psi$  and  $\cos \Delta$  domain for the rotating polarizer ellipsometer data.

### 3.3.3 Simulation of matching domain effect for SOPRA tool

Typically it takes a few minutes to measure one site with the scanning mode, so only CCD mode is feasible for production environments. For CCD mode, usually the analyzer angle is fixed at an angle close to  $45^\circ$ , and CCD or photodiode array is used as a detector, whose noise is much higher than the photomultiplier tube. In this case, it is difficult to analytically

---

<sup>1</sup> Here “better” means that the noise scaling factor at different wavelengths does not vary drastically.

approximate the noise distribution of  $\tan \Psi$ ,  $\log \tan \Psi$  and  $\cos \Delta$ . Therefore, we resort to Monte Carlo simulation to study the noise effect at various matching domains.

The flowchart for the simulation process is illustrated in Figure 3-6. A grating optical response simulator, gtk (Grating Tool Kit), [3.8] is used to simulate  $\tan \Psi$  and  $\cos \Delta$  signals of the “actual” profile. With the addition of information about the analyzer angle (it is set to  $45^\circ$  for the simulation), they are converted to the S signals  $S_1$ ,  $S_2$ ,  $S_3$  and  $S_4$  using formulae (3.5-3.11). We assume the major source of noise is from the integral of the time-variant signal at the detector, so we add randomly generated Gaussian noise to the S signals. In [3.9] the noise sources of a SOPRA system in spectrometer scanning mode was analyzed, and the conclusion was that the spectrometer noise contributes a similar amount of noise as other sources, such as analyzer angle error. In a production environment, the tool can only run in CCD mode, where the spectrometer noise is much higher. Thus our assumption above is valid. The signals in the matching domain are then calculated with information of analyzer angle  $A$  using formulae (3.12-3.15). These signals can be treated as simulated repeatability measurement data in the matching domain.

The regression method is used to extract profiles from simulated perturbed signals, because only a few hundred profiles need to be extracted, and there is no uncertainty as in those induced from the discrete parameter values with library-based method. To study the effect of noise on extracted profiles, we can start from a profile that is close enough to the “actual” profile, so that only the local optimizer is needed. In this study, we use a simplified version of Gauss-Newton method [3.10], linear inversion method [3.11] as the local optimizer, and the procedure is explained below.

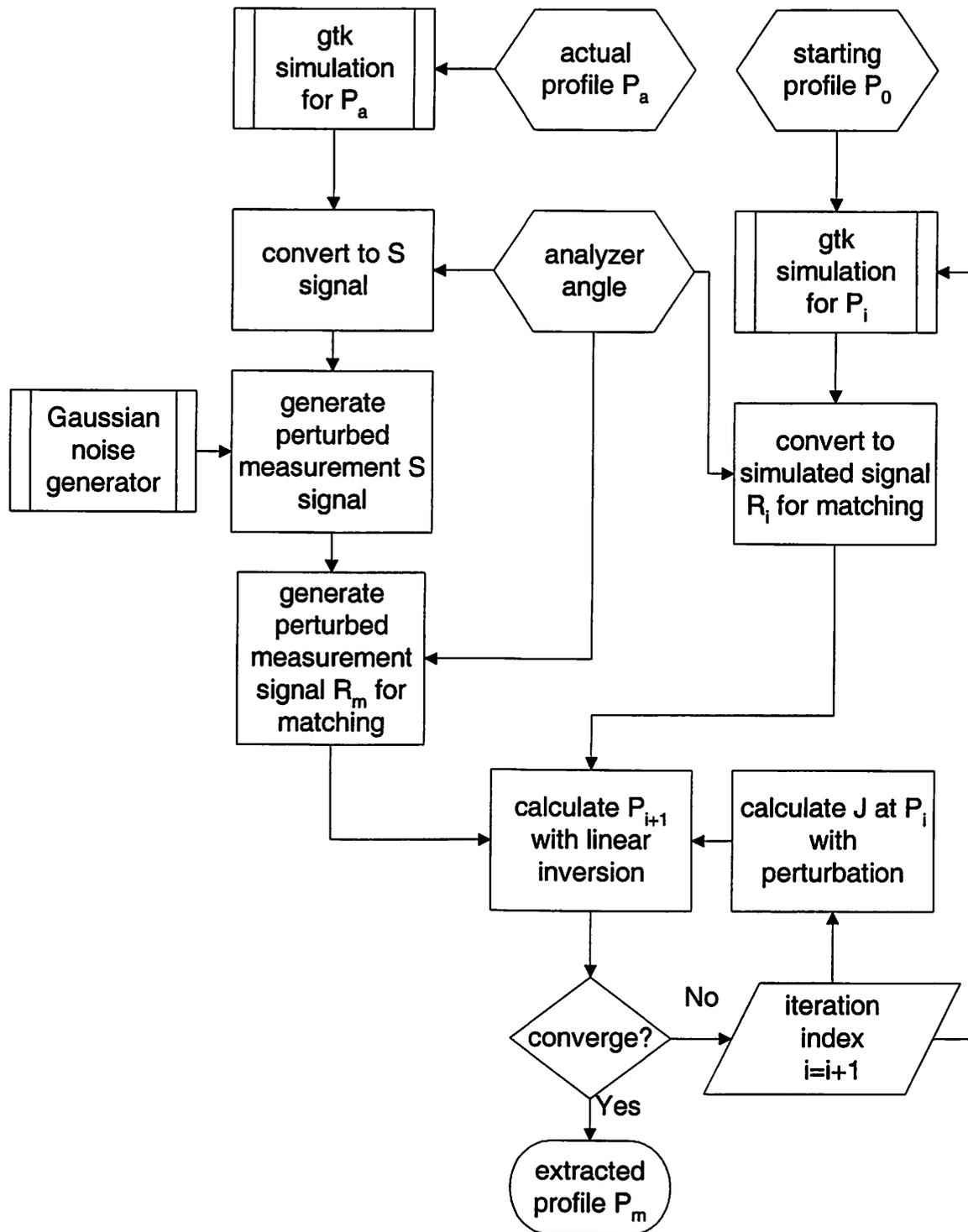


Figure 3-6 Simulation flowchart for matching domain effect on profile extraction

During the profile extraction process, the signal  $R_i$  in the matching domain is converted from the simulated  $\tan \Psi$  and  $\cos \Delta$  responses for the profile  $P_i$  provided by the optimizer, then it

is compared with the “measured” signal  $R_m$  to determine the profile for the next iteration using linear inversion:

$$P_{i+1} = P_i + (J^T J)^{-1} J^T (R_m - R_i), \quad (3.36)$$

where  $P_i$  and  $P_{i+1}$  are vectors of  $n$  profile parameters,  $R_m$  and  $R_i$  are response signal vectors of  $k$  elements in the matching domain, and  $J$  is the Jacobean matrix defined as

$$J = \frac{\partial R}{\partial P} = \begin{pmatrix} \frac{\partial R_1}{\partial P_1} & \frac{\partial R_1}{\partial P_2} & \dots & \frac{\partial R_1}{\partial P_n} \\ \frac{\partial R_2}{\partial P_1} & \frac{\partial R_2}{\partial P_2} & \dots & \frac{\partial R_2}{\partial P_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial R_k}{\partial P_1} & \frac{\partial R_k}{\partial P_2} & \dots & \frac{\partial R_k}{\partial P_n} \end{pmatrix}. \quad (3.37)$$

In practice,  $J$  is calculated using small perturbations of each of the profile parameters from  $P_i$ , i.e., the  $l^{\text{th}}$  column of the  $J$  matrix is calculated as follows:

$$J_l = \begin{pmatrix} \frac{\partial R_1}{\partial P_l} \\ \frac{\partial R_2}{\partial P_l} \\ \vdots \\ \frac{\partial R_k}{\partial P_l} \end{pmatrix} = \frac{\delta R}{\delta P_l}, \quad (3.38)$$

where  $\delta P_l = 0.1 \text{ nm}$ , and  $\delta R$  is the signal change induced by  $\delta P_l$ .

At the end of each iteration, the profile change is checked to see if it meets the convergence criterion. If not, the index is adjusted and we start the next iteration; otherwise, we stop and report the extracted profile. Here we use the following criterion:

$$\|P_{i+1} - P_i\|_{\infty} < 0.1 \text{ nm.} \quad (3.39)$$

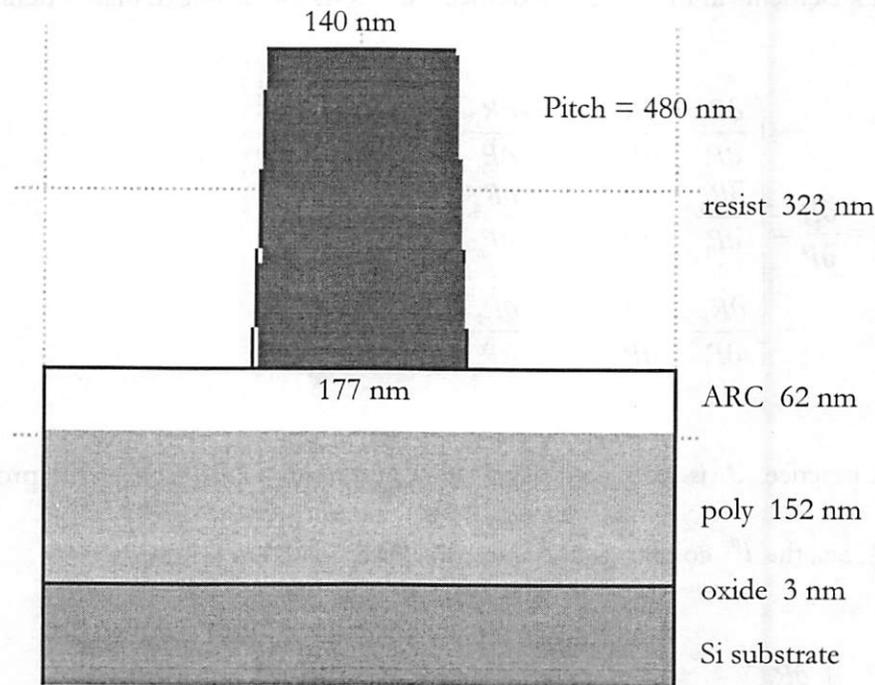


Figure 3-7 Grating structure used for matching domain study

The grating structure used in this study is shown in Figure 3-7. It is a typical poly gate DI (develop inspect) structure, with patterned resist on ARC (anti-reflective coating) on poly silicon on gate oxide on silicon substrate. The grating pitch is 480 nm. The resist profile is modeled with a trapezoid, and its parameters are 140 nm, 177 nm, and 323 nm for top CD, bottom CD, and height, respectively. The film layer thicknesses under the resist are 62 nm, 152 nm, and 3 nm for ARC, poly, and gate oxide, respectively. Since the optical response is not very sensitive to the gate oxide layer thickness variation, and its thickness can be well

controlled in normal processes, we fixed the oxide layer thickness during the profile extraction.

Therefore, the “actual” profile vector is

$$P_a = (140, 177, 323, 62, 152)^T \text{ nm.} \quad (3.40)$$

When doing simulation, we use the typical SOPRA system wavelength range of 240 nm to 760

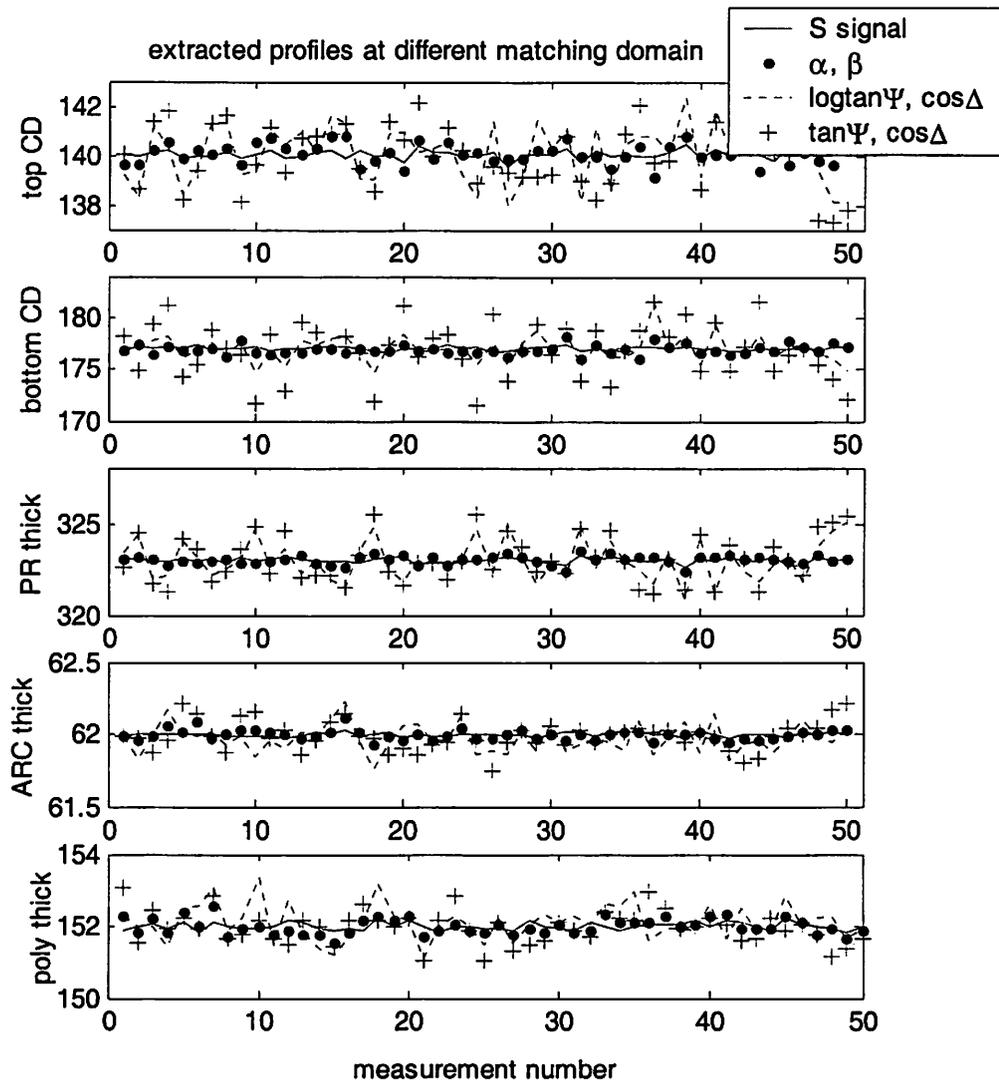


Figure 3-8 Extracted profile parameters using signal matching at different domain

nm with 10 nm step, and angle of incidence of 65°.

We generate 50 normal distributed noise vectors whose mean is 0 and  $3\sigma$  is 1% of the overall S signal magnitude. Then these noise vectors are added to the simulated S signal using profile  $P_a$  to obtain the “measured” signal. The simulated signal noise in the  $(\tan\Psi, \cos\Delta)$  domain has a similar magnitude to that measured on a rotating polarizer ellipsometer for production, which justifies the noise parameter selection above.

The starting profile vector  $P_0$  for regression is a few percent away from  $P_a$ :

$$P_0 = (150, 170, 320, 61, 150)^T \text{ nm.} \quad (3.41)$$

The same set of noise vectors are added to the S signal, then converted to the four matching domains: S signal,  $\alpha$  and  $\beta$ ,  $\log \tan \Psi$  and  $\cos \Delta$ , and  $\tan \Psi$  and  $\cos \Delta$ . Each of the five profile parameters extracted from the noisy signal at these four matching domains for the 50 measurements are plotted in Figure 3-8, and their means and standard deviations are listed in Table 3-1.

Table 3-1 Summary of profile extraction without wavelength weighting

		Top CD	Bottom CD	Height	ARC Th	Poly Th
mean (nm)	S signal	140.05	177.03	322.98	62.00	152.01
	$\alpha, \beta$	140.07	176.92	323.01	62.00	152.01
	$\log \tan \Psi, \cos \Delta$	140.17	177.02	323.00	61.97	152.12
	$\tan \Psi, \cos \Delta$	139.97	176.89	323.06	61.99	151.97
$\sigma$ (nm)	S signal	0.18	0.17	0.13	0.01	0.10
	$\alpha, \beta$	0.41	0.47	0.24	0.04	0.22
	$\log \tan \Psi, \cos \Delta$	1.22	1.27	1.07	0.10	0.50
	$\tan \Psi, \cos \Delta$	1.25	2.67	1.28	0.10	0.46

Although same amount of noise is added to the raw signal, there is a significant difference for the standard deviations of the profiles obtained from regression at different domains. Matching signals at S signal domain gives best results, and standard deviations of profiles from signal matching at  $\alpha$  and  $\beta$  domain are about two times larger. The results from doing regression at  $\log \tan \Psi$  and  $\cos \Delta$ , and  $\tan \Psi$  and  $\cos \Delta$  domains are much worse. The major reason is that when converting from the S signal to  $\alpha$  and  $\beta$ ,  $\log \tan \Psi$  or  $\tan \Psi$  and  $\cos \Delta$ , the noise distribution is deformed by the transformation. There are different magnifications of the original noise for different wavelengths, and those wavelengths with larger noise have a stronger effect on the profile variation.

One way to deal with the unequal noise distribution problem is through wavelength weighting, i.e., the noisier the signal is, the less weight it has when doing signal matching. The weighting matrix  $W$  can be defined from the inverse of the signal variances:

$$W = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_n^2} \end{pmatrix}, \quad (3.42)$$

where  $\sigma_i^2$  ( $i=1, 2, \dots, n$ ) are the variances of the signals, as expressed in equations (3.28) ~ (3.32). The linear inversion equation (3.36) becomes

$$P_{i+1} = P_i + (J^T W J)^{-1} J^T W (R_m - R_i). \quad (3.43)$$

Using the same set of noise signals above, we extract profiles using regression at  $\alpha$  and  $\beta$  domain,  $\log \tan \Psi$  and  $\cos \Delta$  domain, and  $\tan \Psi$  and  $\cos \Delta$  domain, as plotted in Figure 3-9. Their means and standard deviations are listed in Table 3-2.

Table 3-2 Summary of extraction with wavelength weighting

		Top CD	Bottom CD	Height	ARC Th	Poly Th
mean (nm)	S signal	140.05	177.03	322.98	62.00	152.01
	$\alpha, \beta$	140.08	176.92	323.01	62.00	152.00
	$\log \tan \Psi, \cos \Delta$	140.12	176.84	323.01	62.00	152.00
	$\tan \Psi, \cos \Delta$	140.17	176.83	322.99	62.00	152.01
$\sigma$ (nm)	S signal	0.18	0.17	0.13	0.01	0.10
	$\alpha, \beta$	0.40	0.47	0.23	0.03	0.21
	$\log \tan \Psi, \cos \Delta$	0.42	0.52	0.27	0.04	0.27
	$\tan \Psi, \cos \Delta$	0.41	0.52	0.27	0.04	0.27

Table 3-3 Comparison of standard deviations of extracted profile parameters without and with weighting

		Top CD	Bottom CD	Height	ARC Th	Poly Th
S signal		0.18	0.17	0.13	0.01	0.10
$\sigma$ (nm) without weighting	$\alpha, \beta$	0.41	0.47	0.24	0.04	0.22
	$\log \tan \Psi, \cos \Delta$	1.22	1.27	1.07	0.10	0.50
	$\tan \Psi, \cos \Delta$	1.25	2.67	1.28	0.10	0.46
$\sigma$ (nm) with weighting	$\alpha, \beta$	0.40	0.47	0.23	0.03	0.21
	$\log \tan \Psi, \cos \Delta$	0.42	0.52	0.27	0.04	0.27
	$\tan \Psi, \cos \Delta$	0.41	0.52	0.27	0.04	0.27

First we notice from the data that there is an outlier (data point 12) when doing signal matching in  $\log \tan \Psi$  and  $\cos \Delta$  domain. This indicates that the regression process for this data point is trapped in a local minimum, because only the local optimizer is used for the regression, and the noise for this data point happens to make the regression path converge to a

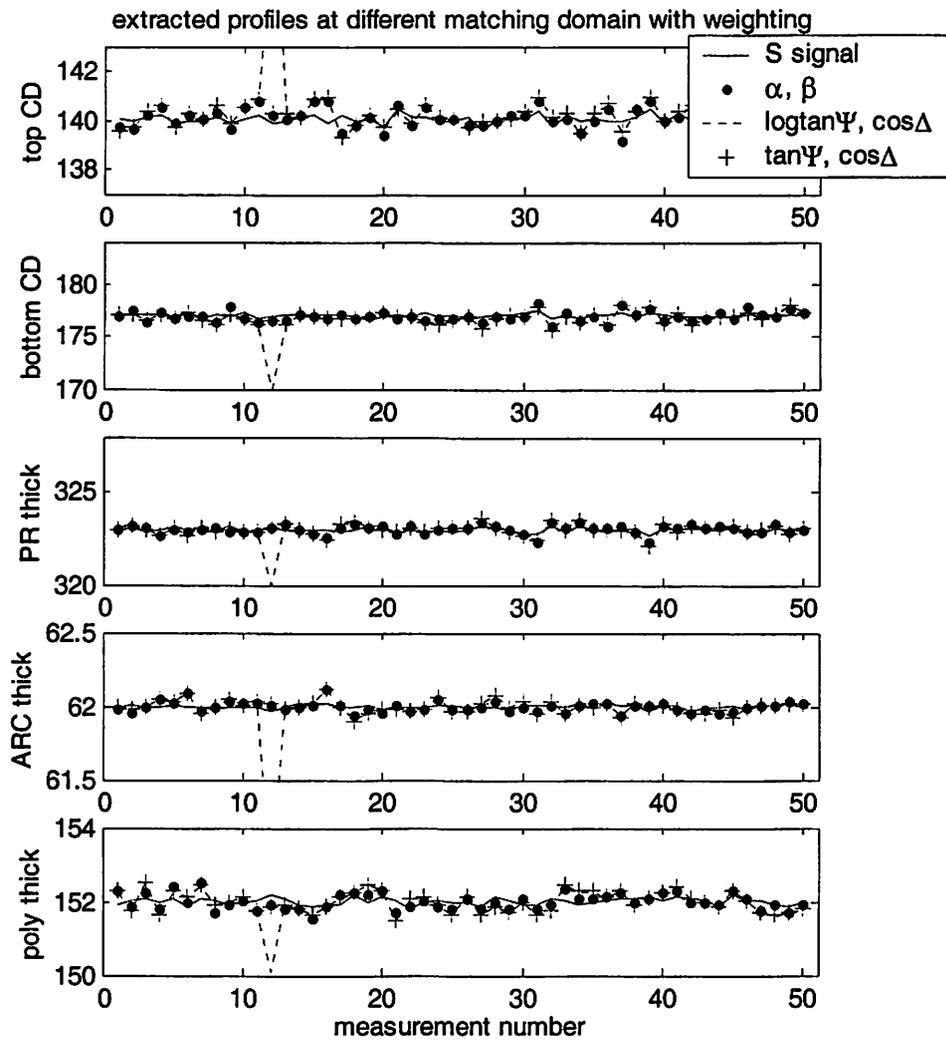


Figure 3-9 Extracted profile parameter at different signal matching domain with wavelength weighting

local minimum far away from the “actual” profile. If we use the global optimization technique before doing local optimization, or use the library-based profile extraction method, this data point will not be an outlier any more. When computing the statistics in Table 3-2, we exclude this outlier point.

With the inverse-noise weighting method, the variances for signal matching in  $\log \tan \Psi$  and  $\cos \Delta$  domain and  $\tan \Psi$  and  $\cos \Delta$  domain are significantly improved, as we listed in Table

3-3. Their variances become similar to that of signal matching in  $\alpha$  and  $\beta$  domain, which is about two times larger than matching in S signal domain.

The simulation above uses the grating structure with 160 nm nominal CD and 480 nm pitch. It is valuable to check how the extracted profile parameter variation changes for smaller features. We select two structures with nominal CDs of 100 nm and 70 nm, and scale other parameters accordingly. The profile parameter vectors are  $P_1 = (86, 112, 253, 46, 102)^T$  nm and  $P_2 = (60, 78, 202, 46, 102)^T$  nm with pitches 300 nm and 210 nm, respectively. We run the simulation at different signal matching domains for these two structures, and compare the bottom CD standard deviations to the 160 nm CD structure. The result is plotted in Figure 3-10.

As was observed for the 160 nm nominal CD case, signal matching with weighting achieves similar bottom CD variance for the  $\alpha$  and  $\beta$  domain,  $\log \tan \Psi$  and  $\cos \Delta$  domain, and  $\tan \Psi$  and  $\cos \Delta$  domain. An interesting finding is that as feature size scales down, the measured bottom CD variance also scales down. Since we also scale pitch with the same ratio as line width, it appears that the ratio of the measurement noise to the feature size is constant as feature size shrinks. This means that the scatterometry measurement precision is scalable to the 70 nm technology node on the roadmap.

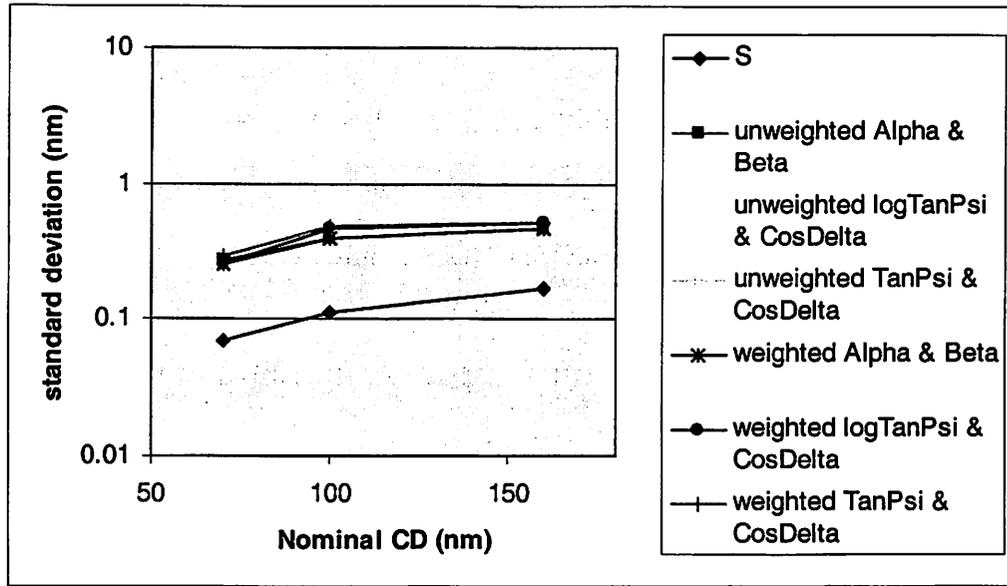


Figure 3-10 Extracted profile parameter precision with different line widths

With the data and analysis above, we can conclude that S signal domain is the best for signal matching in terms of minimizing the detector noise effect on the extracted profiles, and signal matching in the  $\alpha$  and  $\beta$  domain also gives acceptable results. However, converting simulated signals to the S signal domain requires additional optical hardware parameters, such as the analyzer angle. When these parameters are not available, we can only do signal matching in  $\log \tan \Psi$  or  $\tan \Psi$  and  $\cos \Delta$  domains, which significantly magnifies the signal noise effect. Doing signal-matching with weighting according to the noise level at different wavelengths can reduce the noise effect to a similar level as that in the  $\alpha$  and  $\beta$  domain.

In the analysis above, we assume that the major sources of system noise are the variations of the light source, the detector, and the polarizer angle control, which can be attributed to the S signal noise. In reality, there are also many other noise sources, such as analyzer, incidence and azimuth angles, spectrometer wavelength calibration, etc. But their effects are typically much smaller than that of the S signals.

As explained in Chapter 2, scatterometry is a sensitivity-based metrology, and the variation of a certain profile parameter can only be detected if the signal is sensitive to that parameter. Therefore, besides considering the effect of signal noise, one should also take into account signal sensitivity to parameter variations when determining the weights at different wavelengths. This topic can be studied in depth in future work.

### 3.4 The effect of azimuth angle on scatterometry

In all the analysis above, we assume the incident plane is perpendicular to the grating lines. (We define the azimuth angle as  $0^\circ$  in this case.) For non-zero azimuth angle incidence, we can also apply the same scatterometry method for profile extraction, but instead of the 1D-TE and 1D-TM model, the 1D-conical RCWA model [3.12] should be used for grating response simulation, which is about two times slower. Furthermore, due to coupling between TE and TM waves with non-zero azimuth angle incidence, the optical instrument modeling would be more complicated, because the non-diagonal elements of the Jones matrix for the sample are

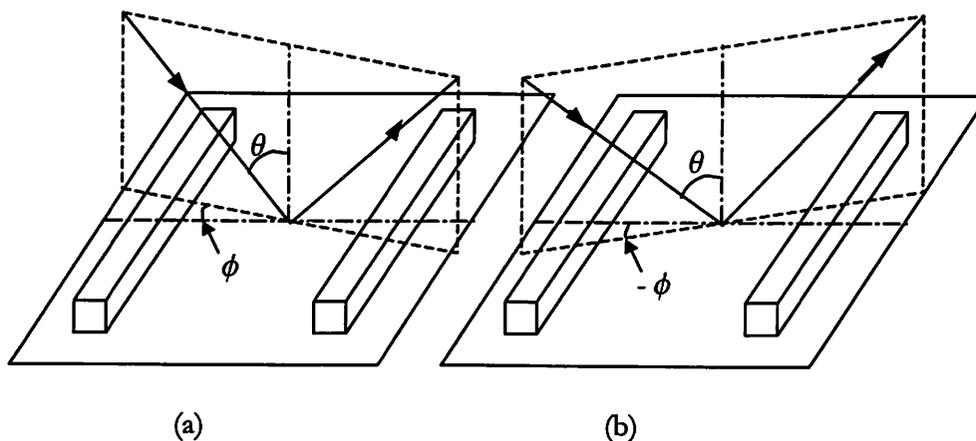


Figure 3-11 Incident light with same angle of incidence  $\theta$  and symmetric azimuth angles: (a)  $\phi$ ; (b)  $-\phi$ .

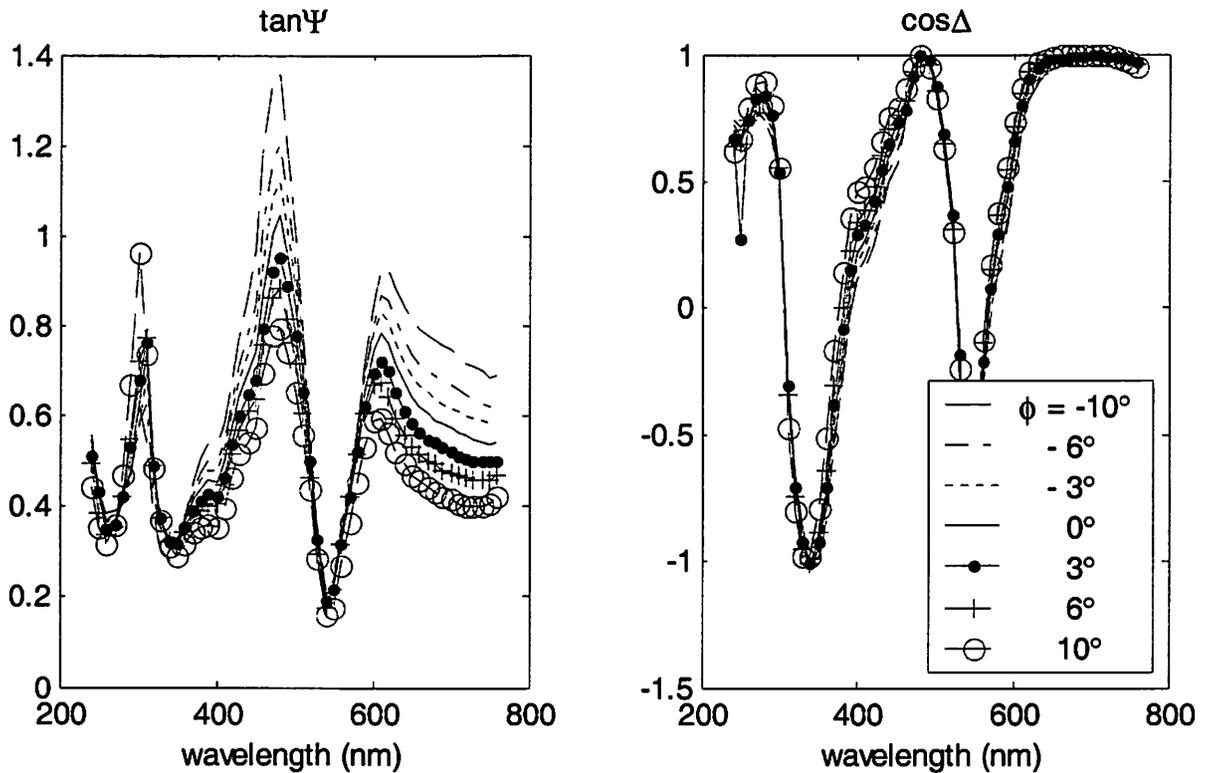


Figure 3-12 Ellipsometry signals measured at different azimuth angles

not zero anymore, as we can see later in this section.

Currently, most of the optical metrology tools used for scatterometry in academia and industry were originally designed for thin film characterization. Although there is an azimuth angle alignment function in these tools, the accuracy cannot meet the requirement of scatterometry, as the azimuth angle does not critically affect the optical response of the thin film stacks. Therefore, it is important to better understand the effect of azimuth angle on scatterometry.

One may think of developing a method using the symmetry of optical settings to monitor and calibrate the azimuth angle of optical tools. As illustrated in Figure 3-11 (a) and (b), with the same angle of incidence  $\theta$ , but different azimuth angles  $\phi$  and  $-\phi$ , the light response  $R$  from the 1D grating, which could be  $\tan \Psi$ ,  $\cos \Delta$ , or reflectivity, is the same. Since it is symmetric

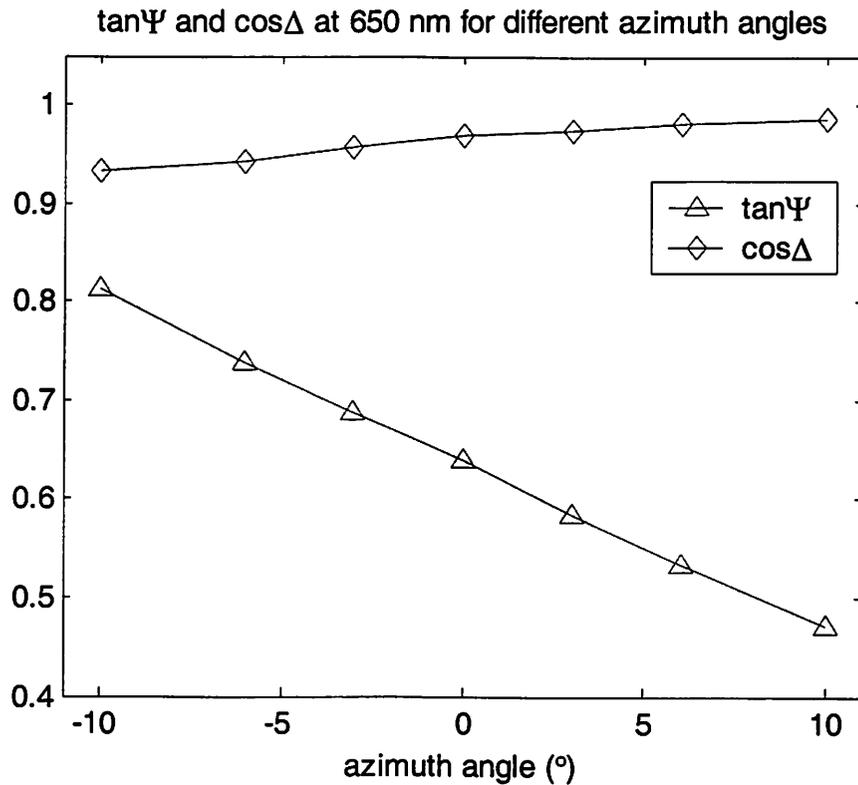


Figure 3-13 Ellipsometry signals at 650 nm for different azimuth angles

to the zero-azimuth-angle incidence, one can measure the response  $R$  at a few azimuth angles close to zero, then determine the zero azimuth angle location from the center of the symmetric response curve with respect to the azimuth angle.

However, measured results on a sample at different azimuth angles on the SOPRA ellipsometer do not show symmetry in the response curve. As illustrated in Figure 3-12, when changing the azimuth angle from  $-10^\circ$  to  $10^\circ$ , the  $\tan\Psi$  and  $\cos\Delta$  signals change monotonically. The phenomenon is more clearly shown in Figure 3-13, where the  $\tan\Psi$  and  $\cos\Delta$  signals at 650 nm are plotted versus the azimuth angle. Signals at other wavelengths have similar behavior. This is contradictory to our analysis that the response signal is symmetric to the zero-azimuth-angle incidence.

Further analysis shows that this is due to coupling between TE and TM waves. Figure 3-14 uses the top-down view of Figure 3-11 to show the coupling from TM to TE wave. In Figure 3-14 (a), the azimuth angle is  $\phi$ , the E-field in the incident TM wave is  $\tilde{E}_p^i$ , and the TM and the TE components of the reflected waves are  $\tilde{E}_p^r$  and  $\tilde{E}_{ps}^r$ , respectively. Then the reflectivity of the TM wave is

$$\tilde{r}_p = \frac{\tilde{E}_p^r}{\tilde{E}_p^i}, \quad (3.44)$$

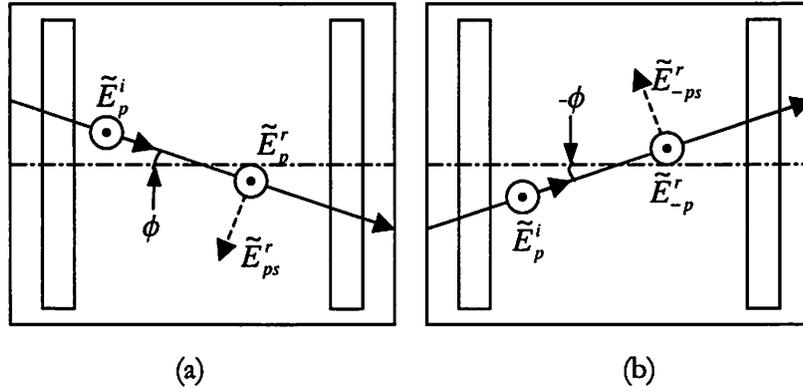


Figure 3-14 Top-down view of Figure 3-10.

and the coupling coefficient from TM to TE wave is defined as

$$\tilde{r}_{ps} = \frac{\tilde{E}_{ps}^r}{\tilde{E}_p^i}. \quad (3.45)$$

When the azimuth angle is  $-\phi$ , as shown in Figure 3-14(b), if similar coordinate system is used, from the symmetry of the optical system, we have the following relations:

$$\tilde{E}_{-p}^r = \tilde{E}_p^r, \quad (3.46)$$

$$\tilde{E}_{-ps}^r = -\tilde{E}_{ps}^r. \quad (3.47)$$

So,

$$\tilde{r}_{-p} = \frac{\tilde{E}_{-p}^r}{\tilde{E}_p^i} = \tilde{r}_p, \quad (3.48)$$

$$\tilde{r}_{-ps} = \frac{\tilde{E}_{-ps}^r}{\tilde{E}_p^i} = -\tilde{r}_{ps}. \quad (3.49)$$

Similarly, we have the relations for the TE wave reflectivity and the coupling coefficient from TE to TM wave:

$$\tilde{r}_{-s} = \frac{\tilde{E}_{-s}^r}{\tilde{E}_s^i} = \tilde{r}_s, \quad (3.50)$$

$$\tilde{r}_{-sp} = \frac{\tilde{E}_{-sp}^r}{\tilde{E}_s^i} = -\tilde{r}_{sp}. \quad (3.51)$$

When azimuth angle is  $\phi$ , with non-zero coupling terms as the non-diagonal elements of the Jones for the sample, equation (3.3) becomes

$$\begin{pmatrix} \tilde{E}_p \\ \tilde{E}_s \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \cos A & \sin A \\ -\sin A & \cos A \end{pmatrix} \begin{pmatrix} \tilde{r}_p & \tilde{r}_{sp} \\ \tilde{r}_{ps} & \tilde{r}_s \end{pmatrix} \begin{pmatrix} \cos P & -\sin P \\ \sin P & \cos P \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{E}_0 \\ \tilde{E}_0 \end{pmatrix}. \quad (3.52)$$

Similar to the definition of  $\tan \Psi$  and  $\cos \Delta$ , we can define

$$\frac{\tilde{r}_{ps}}{\tilde{r}_s} = \tan \Psi_{ps} \cdot e^{i\Delta_{ps}}, \quad (3.53)$$

$$\frac{\tilde{r}_{sp}}{\tilde{r}_s} = \tan \Psi_{sp} \cdot e^{i\Delta_{sp}}. \quad (3.54)$$

For small azimuth angles, e.g.,  $|\phi| \leq 10^\circ$ , the coupling terms are usually very small compared to the reflectivity coefficients. Then we have

$$\tan \Psi_{ps} = \left| \frac{\tilde{r}_{ps}}{\tilde{r}_s} \right| \ll 1, \quad (3.55)$$

$$\tan \Psi_{sp} = \left| \frac{\tilde{r}_{sp}}{\tilde{r}_s} \right| \ll 1. \quad (3.56)$$

Ignoring the second order terms of  $\tan \Psi_{ps}$  and  $\tan \Psi_{sp}$ , we can derive the formula for the intensity at the detector using similar procedure as that in section 3.2.3:

$$I(P) = \left| \tilde{E}_p \right|^2 + \left| \tilde{E}_s \right|^2 = I_0 (1 + \alpha \cos 2P + \beta \sin 2P), \quad (3.57)$$

where

$$\alpha = \frac{\tan^2 \Psi - \tan^2 A}{\tan^2 \Psi + \tan^2 A} \cdot \frac{1 + \delta_2}{1 + \delta_1}, \quad (3.58)$$

$$\beta = \frac{2 \cos \Delta \tan \Psi \tan A}{\tan^2 \Psi + \tan^2 A} \cdot \frac{1 + \delta_3}{1 + \delta_1}, \quad (3.59)$$

$$I_0 = \frac{1}{2} \left| \tilde{E}_0 \tilde{r}_s \right|^2 \cos^2 A (\tan^2 \Psi + \tan^2 A) (1 + \delta_1). \quad (3.60)$$

The equations above are very similar to those without coupling coefficients, except the extra terms with  $\delta_i$  ( $i = 1, 2, 3$ ), which can be expressed as

$$\delta_1 = 2 \left( \tan \Psi_{ps} \tan \Psi \cos(\Delta - \Delta_{ps}) + \tan \Psi_{sp} \cos \Delta_{sp} \right) \frac{\tan A}{\tan^2 \Psi + \tan^2 A}, \quad (3.61)$$

$$\delta_2 = 2 \left( \tan \Psi_{ps} \tan \Psi \cos(\Delta - \Delta_{ps}) - \tan \Psi_{sp} \cos \Delta_{sp} \right) \frac{\tan A}{\tan^2 \Psi - \tan^2 A}, \quad (3.62)$$

$$\delta_3 = \frac{\tan \Psi_{ps} \cos \Delta_{ps} \tan A}{\tan \Psi \cos \Delta} + \frac{\tan \Psi_{sp} \cos(\Delta - \Delta_{sp})}{\tan A \cos \Delta}. \quad (3.63)$$

When the azimuth angle changes to  $-\phi$ , substituting equations (3.48) ~ (3.51) into equation (3.52), we have the time-variant intensity in the similar form to that in equation (3.57):

$$I(P) = \left| \tilde{E}_p \right|^2 + \left| \tilde{E}_s \right|^2 = I_{0-} (1 + \alpha_- \cos 2P + \beta_- \sin 2P), \quad (3.64)$$

and that the expressions of  $\alpha_-$ ,  $\beta_-$  and  $I_{0-}$  are:

$$\alpha_- = \frac{\tan^2 \Psi - \tan^2 A}{\tan^2 \Psi + \tan^2 A} \cdot \frac{1 - \delta_2}{1 - \delta_1}, \quad (3.65)$$

$$\beta_- = \frac{2 \cos \Delta \tan \Psi \tan A}{\tan^2 \Psi + \tan^2 A} \cdot \frac{1 - \delta_3}{1 - \delta_1}, \quad (3.66)$$

$$I_{0-} = \frac{1}{2} \left| \tilde{E}_0 \tilde{r}_s \right|^2 \cos^2 A (\tan^2 \Psi + \tan^2 A) (1 - \delta_1). \quad (3.67)$$

where  $\delta_1$ ,  $\delta_2$ , and  $\delta_3$  are the same as those defined in equations (3.61) ~ (3.63). It is obvious from the equations above that as azimuth angle switches from  $-\phi$  to 0 to  $\phi$ , both  $\alpha$  and  $\beta$  vary monotonously. Therefore, it is easy to prove from equations (3.14) and (3.15) that the measured signals  $\tan \Psi$  and  $\cos \Delta$  also change monotonically with the azimuth angle. Since there is no symmetry around the zero-azimuth angle location for the response curves, it is impossible to calibrate the azimuth angle in this way.

For rotation compensator configuration, we can apply the same method to analyze the coupling effect, and the similar monotonous response of the  $\tan \Psi$  and  $\cos \Delta$  signal with respect to the azimuth angle can be derived. When measuring samples with this type of ellipsometer, we got similar signals as those shown Figure 3-12 and Figure 3-13.

When using either a library-based or a regression method to do profile extraction with the signals at various azimuth angles, if we only do simulation with the zero azimuth angle, we would expect the highest fit from simulation to measurement occurs with the signals obtained at the zero-azimuth-angle location. This is verified by experiment. In Figure 3-14, the goodness of fit (GOF) of the best-matched signals to the measured ones at various azimuth angles are plotted, where GOF is defined as

$$GOF = 1 - \frac{\sum_i [S_s(\lambda_i) - S_m(\lambda_i)]^2}{\sum_i [S_m(\lambda_i) - \bar{S}_m]^2}, \quad (3-68)$$

where  $S_s(\lambda_i)$  and  $S_m(\lambda_i)$  are the simulated and measured signals at wavelength  $\lambda_i$ , and  $\bar{S}_m$  is the average of the measured signals for all wavelengths. The highest GOF points in the figure occur at the location between  $0^\circ$  and  $1^\circ$ , where the actual zero-azimuth angle should be.

However, the GOF versus azimuth angle curve is relatively smooth, and the GOF curve might not be symmetric, so it is difficult to determine the zero-azimuth angle accurately.

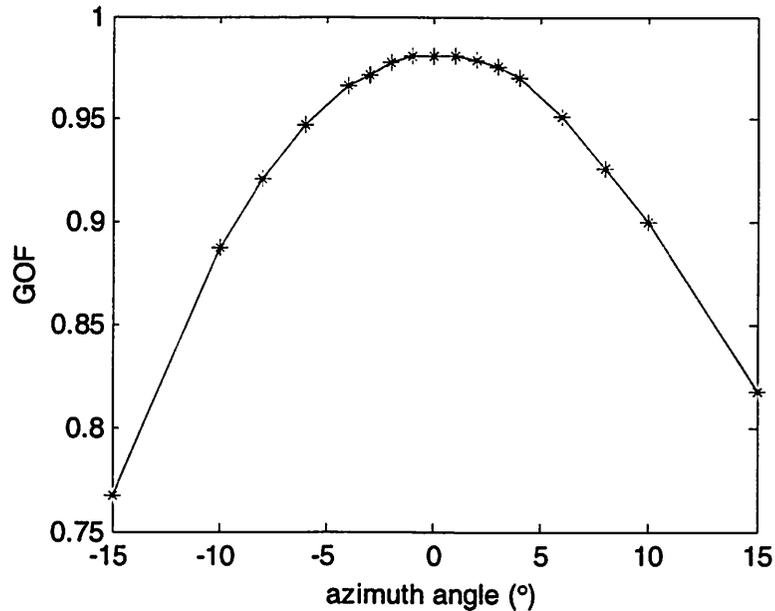


Figure 3-15 GOF vs. azimuth angle

Since cross-polarization occurs when the incident light beam is not perpendicular to the direction of the grating lines, another method of azimuth angle alignment is to illuminate TM light on the grating, and monitor the intensity of the TE component of the reflected light. When the incident light beam is perpendicular to the grating, the TE reflection is zero. This method involves only a small change in the ellipsometer control software, and the accuracy depends on the ratio of the magnitude of the cross-polarization to the detector background noise.

### 3.5 Summary

In this chapter, we first introduced various optical devices that can be used for scatterometry application. Then we analyzed the noise effect of the signals to extracted profiles and

simulated the different effects when doing signal matching at different signal domains. Since optical response from a grating structure is sensitive to the azimuth angle of the incident light beam, azimuth angle calibration brings a new requirement to the optical tools for scatterometry. We explained the monotonous behavior of measured signal change versus azimuth, and discussed possible azimuth angle calibration methods.

In the next chapter, we will present some experimental results of scatterometry measured on sub-quarter- $\mu\text{m}$  features.

## Reference

- [3.1] C.J. Raymond, M.R. Murnane, S.L. Prins, S.S. Naqvi, J.R. McNeil, "Multi-parameter Process Metrology Using Scatterometry," SPIE Vol. 2638, Optical Characterization Techniques for High-Performance Microelectronic Device Manufacturing II, 84-93, 1995.
- [3.2] E. Maiken, K. Johnson, D. Likhachev, A. Norton, A. Sezginer, F. Stanke, D. Yonenaga, "Integrated Critical Dimension Metrology for Lithography Process Control," AEC/APC Symposium XIII, Banff, Canada, October 6-11, 2001.
- [3.3] J.M. Holden, V. Zhuang, W. McGahan, "Optical CD for Metrology and Control of Line and Space Critical Dimension," AEC/APC Symposium XIII, Banff, Canada, October 6-11, 2001.
- [3.4] SOPRA Theory Document, <http://www.sopra-sa.com/index2.htm>, SOPRA Corporation.
- [3.5] B.S. Stutzman, H.T. Huang, F.L. Terry, "Two-Channel Spectroscopic Reflectometry for in situ Monitoring of Blanket and Patterned Structures During Reactive Ion Etching," Journal of Vacuum Science & Technology B, vol. 18, (no. 6), 2785-2793, Nov. 2000.

- [3.6] N. Benesch, A. Hettwer, C. Schneider, L. Pfitzner, H. Ryssel, "Phi-scatterometry for Integrated Linewidth Control in DRAM Manufacturing," Proceedings of IEEE International Symposium on Semiconductor Manufacturing, 129-132, 2001.
- [3.7] [http://www.rudolphtech.com/products/spectralaser\\_200.html](http://www.rudolphtech.com/products/spectralaser_200.html), SpectraLaser200, Rudolph Technologies, Inc.
- [3.8] X. Niu, "An Integrated System of Optical Metrology for Deep Sub-micron Lithography," Ph.D. dissertation, UC Berkeley, Spring, 1999.
- [3.9] W. Foong, "Characterizing the Sensitivity of Scatterometry for sub-100 nm Technologies, MS report, UC Berkeley, Spring 2001.
- [3.10] G.C. Kossakes, N. Kyriakopoulos, "An Improvement on the Newton-Gauss Algorithm," IEEE Transactions on Circuits and Systems, vol.CAS-24, (no.5), 271-273, May 1977.
- [3.11] E.M. Drege, D.M. Byrne, "Lithographic Process Monitoring Using Diffraction Measurements," Proceedings of the SPIE, vol.3998, (Metrology, Inspection, and Process Control for Microlithography XIV) 147-57, 28 Feb.-2 March 2000.
- [3.12] M.G. Moharam and T.K. Gaylord, "Three-dimensional Vector Coupled-wave Analysis of Planar-grating Diffraction," J. Opt. Soc. Am. A, vol. 73, 1105-1112, Sept., 1983.

# Chapter 4 Scatterometry Experiment Data Analysis

## 4.1 Introduction

After analyzing a few theoretical problems for scatterometry in Chapter 3, we will demonstrate some experimental results in this chapter. As explained in the second chapter, scatterometry is one of the key elements of the advanced process control framework. Since scatterometry is still a new technology, before integrating it into the control framework, it is important to characterize a scatterometry system, and understand the system specifications and limitations through experiments. In this chapter we first review some pioneering scatterometry work done on the  $2\theta$  and other hardware configurations, and then compare its sensitivity with specular spectroscopic scatterometry. Next, we show some results from characterizing a scatterometry system using Thermo-Wave's Optical-Probe 5240 ellipsometer, including accuracy, precision, light spot size measurement, as well as the effect of the library-based versus regression method for profile extraction. Finally, we demonstrate the potential application of scatterometry on metal layer processes, and foresee its possible application to chemical-mechanical planarization (CMP).

## 4.2 Review of experimental work on scatterometry

### 4.2.1 Research on scatterometry in the early stage

As reviewed in Chapter 2, much work has been done on line width and profile characterization with optical techniques. Computer simulation plays a major role in understanding and

modeling the process of optical response from small features (in the same scale as the illuminating wavelength). It was not until the early 1990s that computers became fast enough to do rigorous vector electromagnetic simulations for the optical scattering process within a practically short time. One early experimental result shows that the measured spectra obtained

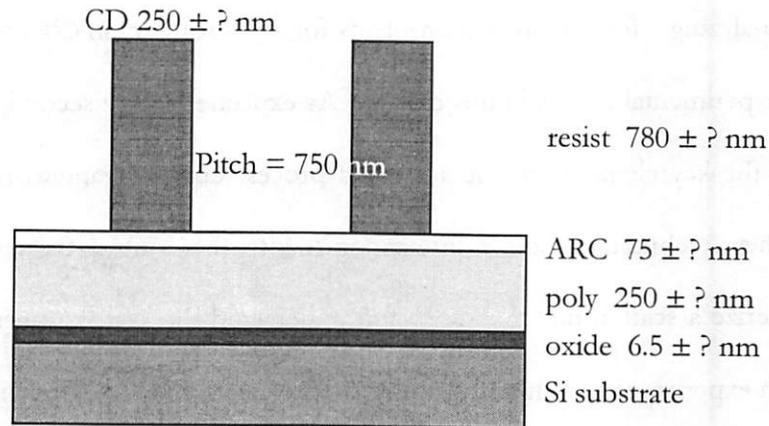


Figure 4-1 Stack structure used for early scatterometry study (figure not to scale)

from scattering at features about 1.2  $\mu\text{m}$  wide and 476 nm thick are sensitive to the line width and thickness variations of a few nanometers [4.1]. The experimental result is also in good agreement with simulation using TEMPEST [4.2], a rigorous electromagnetic scattering simulator. This and other similar observations helped motivate people to develop a systematic method for optical metrology of profiles.

The application of the rigorous coupled wave analysis technique in modeling optical response from grating structures has greatly improved the simulation speed, which makes it possible to do line width and profile extraction from the measured spectrum [4.3]. Raymond, etc. showed the capability of simultaneous measurement of resist line width and film thickness using 2- $\theta$  scatterometry [4.4]. As shown in Figure 4-1, the cross section of the photoresist line is

modeled as a rectangle, with the thicknesses of ARC, and polysilicon films as variables as well. The scatterometry line width measurement is compared with CD-SEM, and the film thickness measurement is compared with ellipsometry. The results are listed in Table 4-1 below.

Table 4-1 Comparison of scatterometry and other metrology tools

Sample	CD (nm)		Resist height (nm)		ARC thickness (nm)		Poly thickness (nm)	
	SEM	Scat.	Ellip.	Scat.	Ellip.	Scat.	Ellip.	Scat.
Wafer 2	239	200	782	700	68.7	80	247	230
Wafer 7	263	210	782	710	68.0	65	292	260
Wafer 12	254	220	781	710	67.8	75	248	230
Wafer 17	253	200	781	710	67.5	65	290	260
Wafer 22	241	210	780	720	66.7	75	267	250
Avg. bias	42.0		71.2		6.4		23.4	
1 $\sigma$ of bias	10.4		7.8		3.7		7.6	
Scat. 1 $\sigma$	0.25		1.29		0.0		0.72	
Comment	SEM – CD-SEM; Ellip. - Ellipsometry; Scat. - Scatterometry; 1 $\sigma$ - Std. deviation							

The results above demonstrate that there is sufficient information in diffraction signals to determine both line width and film thickness. The scatterometry data show good agreement with data from other referencing metrologies (CD-SEM for line width, and ellipsometry for film thickness). The repeatability (1 $\sigma$ ) results were also shown to be sum-nanometer for the measured parameters.

The same group of researchers also demonstrated scatterometry to be a promising technique in many other applications, such as etched polysilicon gate structure [4.5], chrome mask features [4.6], etched metal features [4.7], etc. In the meantime, many other researchers have obtained comparable results with 2- $\theta$  scatterometry on various applications as well. [4.8] [4.9]. All these progresses have induced more and more interests in scatterometry in both academia

and industry [4.10], which results in the rapid development and application of scatterometry in the semiconductor industry.

#### 4.2.2 Comparison between spectroscopic and 2- $\theta$ scatterometry

Most of the early scatterometry work was performed on the 2- $\theta$  configuration, as it is based on the intuitive thinking of “observing” objects at various angles. The majority of these published experimental results use simple rectangles to model the cross section of grating profiles. In addition to the limited processing power available during that era, the major reason for choosing a simple profile modeling method is the low sensitivity of the signals for the detailed profile structures.

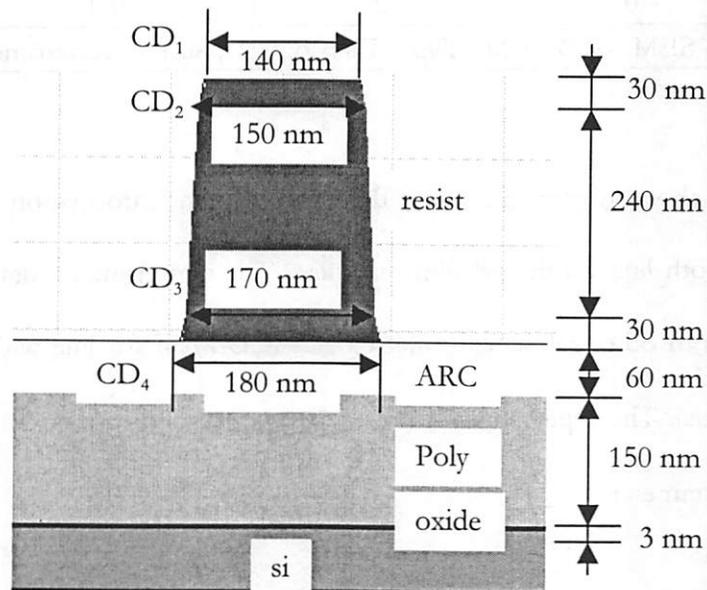


Figure 4-2 Structure used to illustrate scatterometry sensitivity

We use the structure in Figure 4-2 to illustrate this point. It is a typical post-develop photo resist for a polysilicon gate structure, consisting of a 300 nm high resist on 60 nm thick ARC on 150 nm thick poly on 3 nm thick oxide on silicon substrate. We model the profile using three stacked trapezoids, where the two inflection points occur at 10% and at 90% from the bottom of the resist line. As illustrated in the figure, for the nominal profile, the widths at the bottom, 10%, 90%, and the top ( $CD_1$ ,  $CD_2$ ,  $CD_3$ , and  $CD_4$ ) are 180 nm, 170 nm, 150 nm, and 140 nm, respectively. The TE and TM reflectance from the nominal profile for the 2- $\theta$  configuration are simulated and plotted in Figure 4-3, assuming a He-Ne laser is used for the light source, where wavelength is 633 nm, and the range of incidence angle is from  $0^\circ$  to  $50^\circ$ .

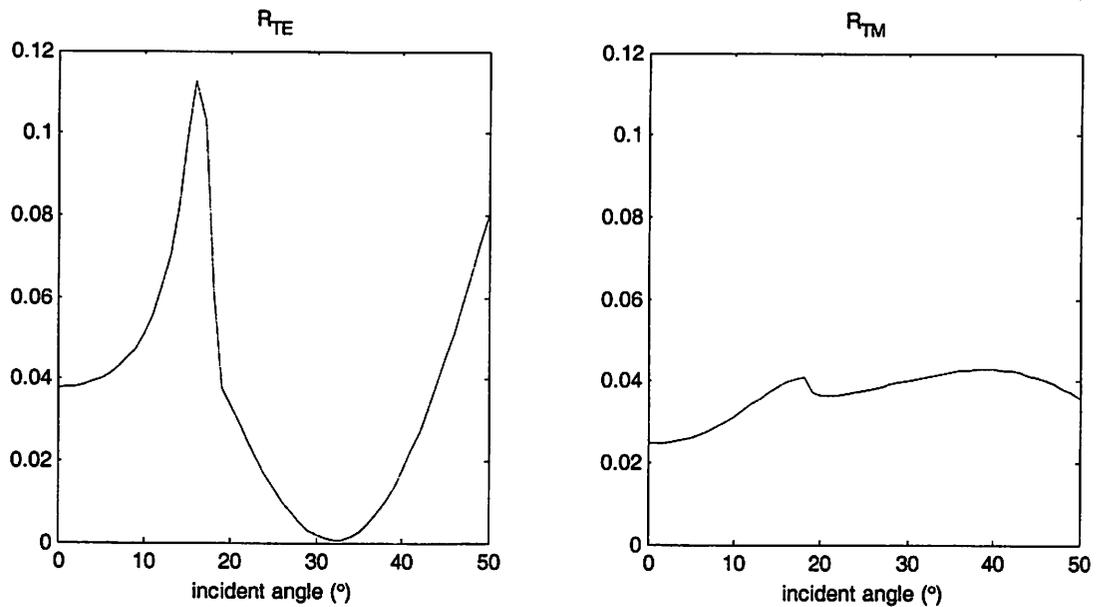


Figure 4-3 Example signal of a 2- $\theta$  scatterometry system

To study the sensitivity of the 2- $\theta$  scatterometry signal to detailed profile features, we vary the four CDs by 5 nm each time, and simulate the difference of reflectance for TE and TM waves, as shown in Figure 4-4. We can see that the signal variations due to the variation of detailed profile features ( $CD_1$  and  $CD_4$ ) are less than 0.1% for the TE wave, and they are even smaller for the TM wave. Although we could not obtain the noise behavior of the actual 2- $\theta$  scatterometry system, we see the signal variation due to the detailed features is comparable to the noise level (standard deviation of the repeatability measurement signals) of other reflectometer systems, thus we conclude that it would be difficult to detect the detailed profile change of less than 5 nm from the signal.

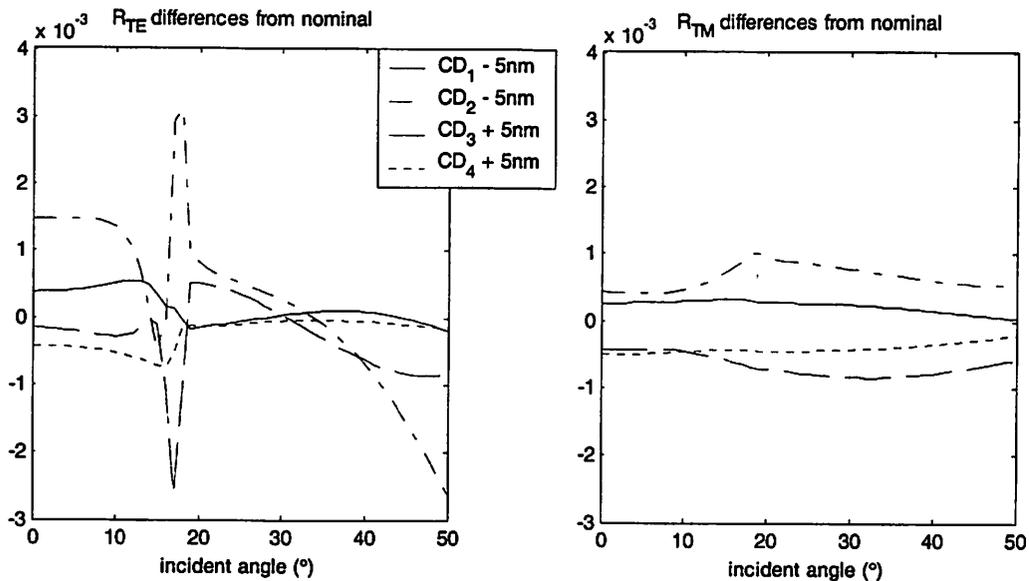


Figure 4-4 Signal sensitivity of the 2- $\theta$  scatterometry system

We could do a similar study on a spectroscopic ellipsometer system. The ellipsometer signals  $\log(\tan\Psi)$  and  $\cos\Delta$  simulated from the same nominal profile in Figure 4-2 are shown in Figure 4-5, and the signal variations due to width changes at each of the four CD measurement

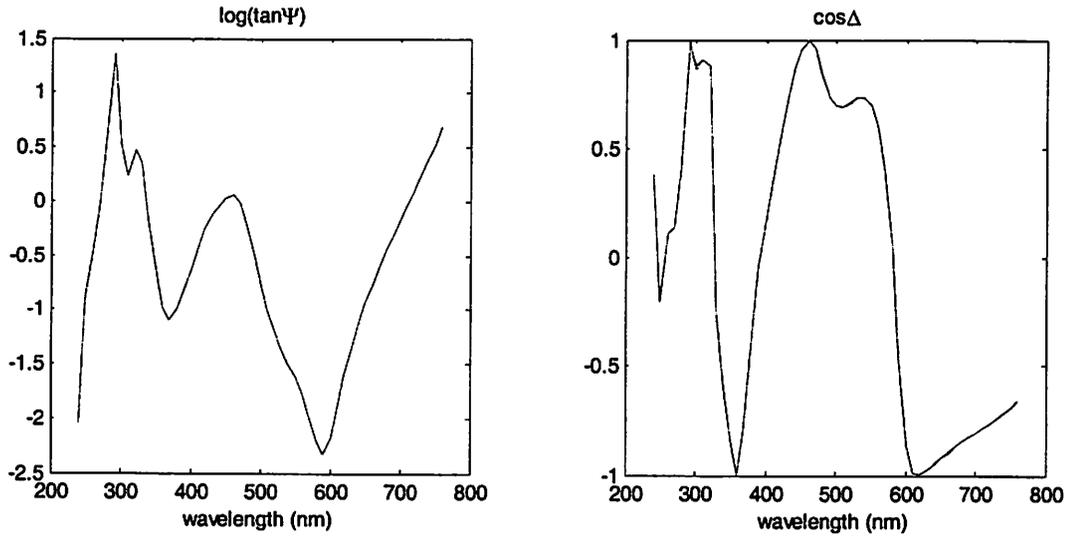


Figure 4-5 Sample signal of a spectroscopic scatterometry system

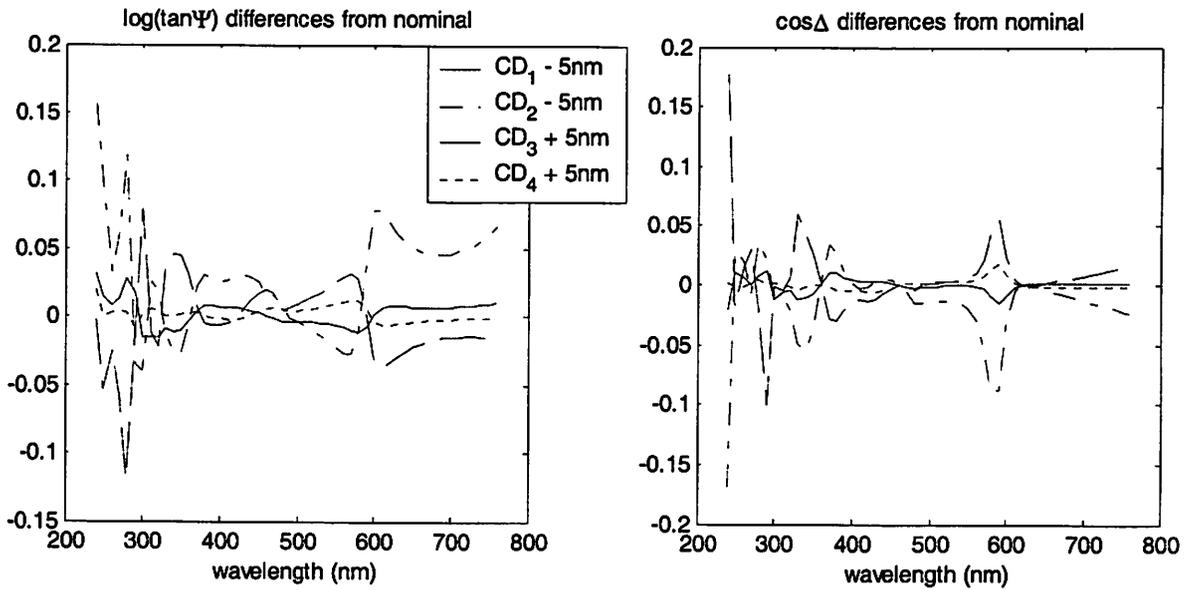


Figure 4-6 Signal sensitivity of the spectroscopic scatterometry system

heights are plotted in Figure 4-6. We can see that the signal variations due to changes of  $CD_1$  and  $CD_4$  are larger than the noise levels at many wavelengths, which means that these changes can be detected by scatterometry using spectroscopic ellipsometer signals.

Since the introduction of specular spectroscopic scatterometry [4.11], various optical configurations have been applied for spectroscopic scatterometry by both academic researchers and commercial developers, including polarized reflectometry [4.12], unpolarized normal incidence reflectometry [4.13], polarized normal incidence reflectometry [4.14], and ellipsometry [4.15] [4.16] [4.17]. Among these works, scatterometry has been demonstrated to be feasible for a wide spectrum of applications, such as etched polysilicon gate, shallow trench

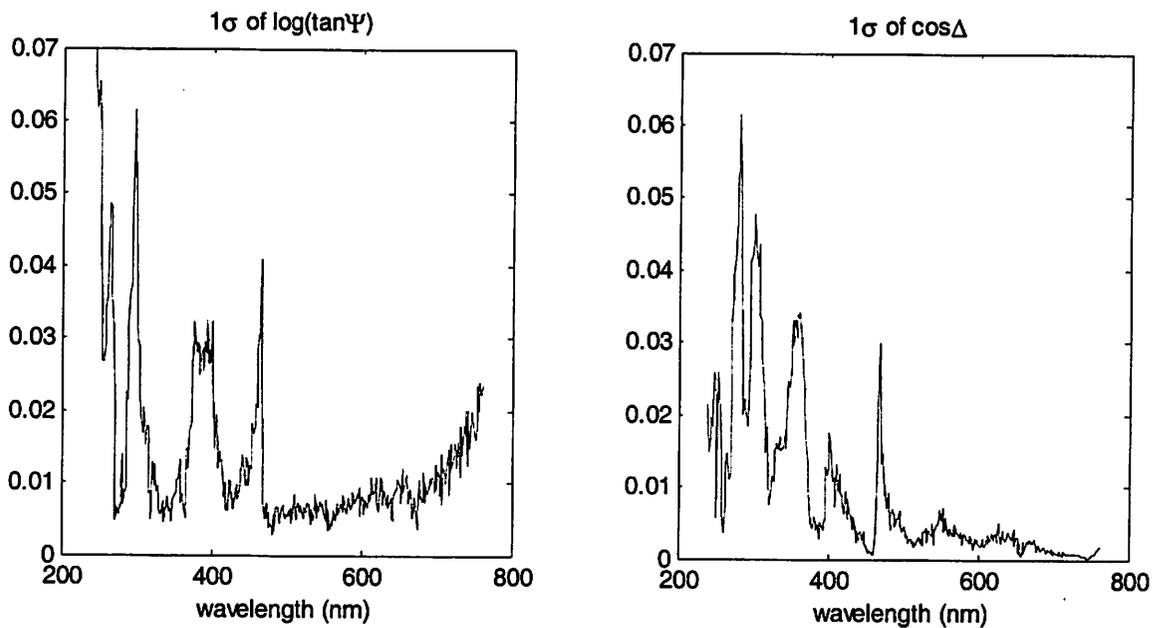


Figure 4-7 Standard deviation of repeatability measurement signals on the ellipsometry system

isolation, and dual damascene trench for final inspection (FI), as well as the resist features for develop inspection (DI) for these applications.

## 4.3 Specular spectroscopic scatterometry tool characterization

### 4.3.1 Introduction of applications

In this section, we present the experimental results of characterization of a specular spectroscopic scatterometry system. We use the spectroscopic ellipsometer of a Therna-Wave

[4.18] Opti-Probe 5240 system, a popular thin film metrology tool in semiconductor fabs, to collect signals for scatterometry analysis. Three applications in a state-of-the-art semiconductor production line are used for accuracy, precision and measurement light spot size characterization.

The first application is a polysilicon gate develop inspection (DI) structure. As shown in Figure 4-8, it consists of developed resist lines on top of a thin film stack of SiON (as ARC), polysilicon, and gate oxide.

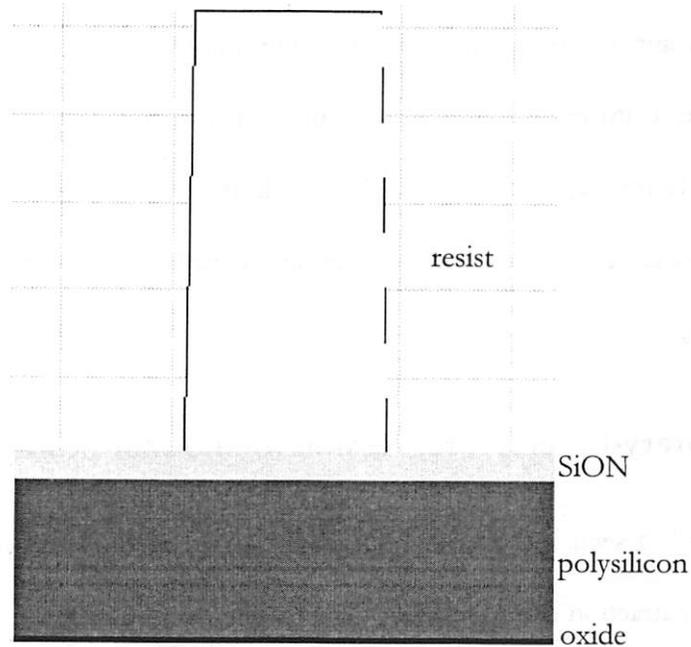


Figure 4-8 Polysilicon gate DI structure

After the wafer is etched, it is called polysilicon gate final inspection (FI), and the structure becomes patterned SiON and polysilicon on top of unpatterned gate oxide on silicon substrate, as illustrated in Figure 4-9. For the above two applications, both the bottom width and the shape of the profiles are important because they directly link to the device performance.

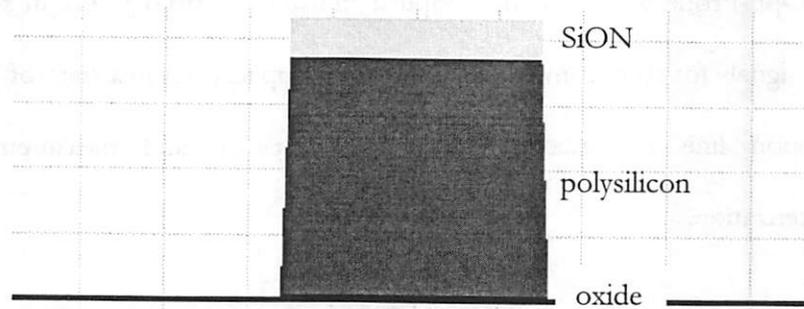


Figure 4-9 Polysilicon FI structure

The third application is shallow trench isolation (STI) final inspection (FI). The structure consists of a patterned stack of  $\text{Si}_3\text{N}_4$  on oxide on silicon, as shown in Figure 4-10. There is a small reentrant feature (about 10 nm to 20 nm) at the oxide layer. This is due to the fact that oxide etches isotropically during this particular plasma etch process. For this structure, the key parameter is the depth of the trench as well as the width and sidewall angle of the silicon trench because the effect of the inter-device isolation has strong correlation with these parameters.

#### 4.3.2 Library-based vs. regression method for profile extraction

As reviewed in section 2.3.3, both the library-based method and regression method are used for profile extraction from measured signals. The detailed procedure for doing regression with the linear inversion method is described in section 3.3.3. We use the polysilicon gate FI structure (Figure 4-9) to illustrate the effect of these two methods on profile extraction.

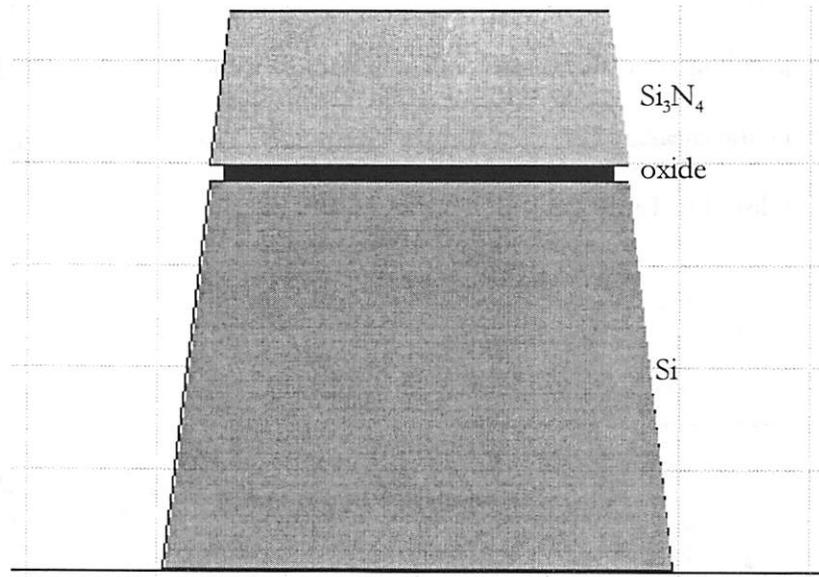


Figure 4-10 Shallow Trench Isolation FI structure

The polysilicon section is modeled as a trapezoid, and the SiON layer on top is modeled as a rectangle, where its width is set equal to the top width of the polysilicon trapezoid. So there are five independent profile parameters: the polysilicon top and bottom CDs, and the thicknesses of SiON, polysilicon, and unpatterned oxide. Their nominal values are 150 nm, 160 nm, 31 nm, 180 nm, and 2 nm, respectively.

Table 4-2 Comparison of extracted profiles using regression and library-based methods

Parameter Method	Polysilicon top CD	Polysilicon bottom CD	SiON thickness	Polysilicon thickness	Oxide thickness	GOF
Regression (nm)	147.8	158.3	30.8	178.2	3.0	0.9935
Library (nm)	148.0	158.0	31.0	178.0	3.0	0.9932

We first run regression starting from the nominal profile to match a measured signal. The result parameter values are listed in Table 4-2. Based on the regression result and process specifications, we make a library to cover the possible process range. The profile parameter

ranges and resolution are listed in Table 4-3. There are 143,360 profiles and their corresponding simulated signals in the library. We then find the signal in the library that best matches the measured signals, and the parameter values of the profile that generate this signal are also listed in Table 4-2.

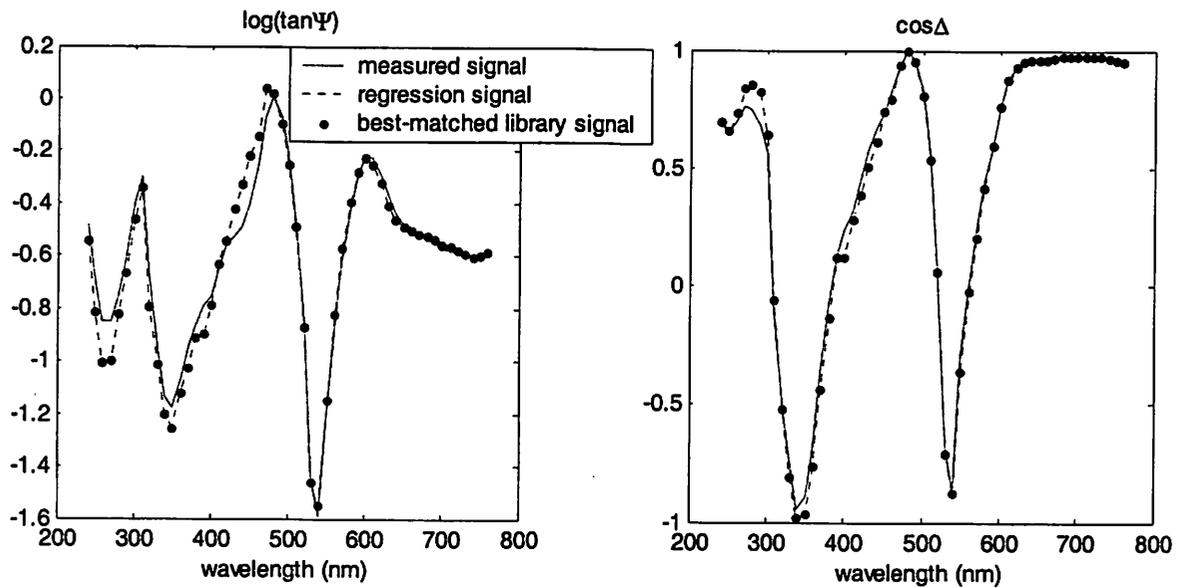


Figure 4-11 Measured versus simulated signals using regression and library-based methods

Table 4-3 Library parameters for polysilicon FI structure

Parameter	Polysilicon top CD	Polysilicon bottom CD	SiON thickness	Polysilicon thickness	Oxide thickness
Minimum (nm)	145.0	148.9	30.0	170.0	1.5
Maximum (nm)	160.0	162.2	31.5	182.0	3.3
Resolution (nm)	1.0	0.7	0.5	0.8	0.3

As we can see in these tables, the extracted profiles using regression and library-based methods are very close to each other, and the difference of the profile parameter values using these two methods is less than half of the library parameter step sizes. If we consider the profile

extracted using the regression method as the best result one can obtain using the current profile model, then the uncertainty induced by the library-based method due to discrete profiles is about half of the library parameter step sizes (resolution). In this example, the signal difference from these two profiles is very small, as illustrated in Figure 4-11.

### 4.3.3 Accuracy of the scatterometry system

The cross section scanning electron microscope (XSEM) has long been used in the semiconductor industry as a tool to obtain detailed profile information. To check the accuracy of the scatterometry system, we use the STI FI structure. We first measure the wafer using the

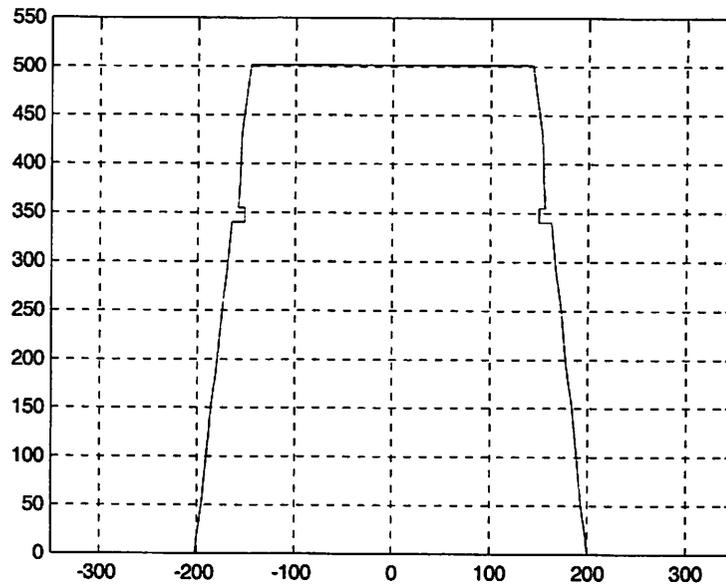


Figure 4-12 Measured STI FI structure profile using scatterometry

ellipsometer, and then run the measured signals through a pre-built library to extract the profile, as shown in Figure 4-12. Next the wafer is cross-sectioned at the grating section. To improve the contrast under the XSEM, a layer of gold of about 20 nm is coated on the structure. The resulting XSEM picture is shown in Figure 4-13. Compared with Figure 4-12,

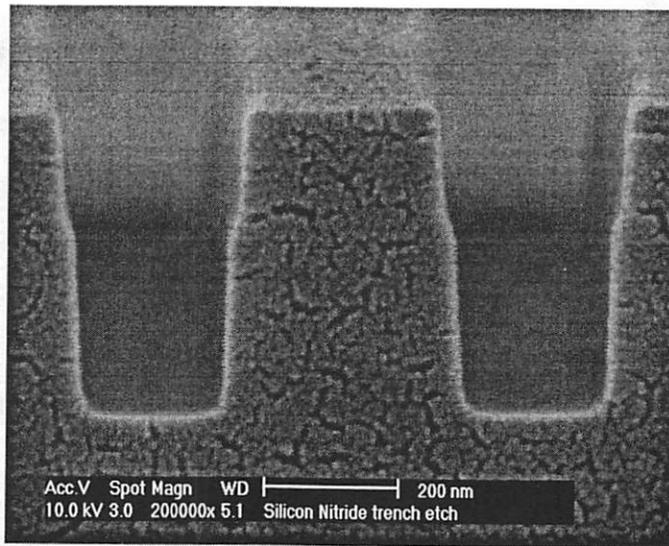


Figure 4-13 Cross-section SEM image of STI FI structure

we can see that the XSEM picture has a similar shape to the scatterometry profile, except that there is no reentrant feature at the oxide layer in the profile. We believe the actual profile does have the small reentrant feature, as has been shown in some pictures obtained using tunneling electron microscope (TEM), a metrology tool that has higher resolution than SEM. In the picture in Figure 4-13, the reentrant feature is covered by the gold layer.

From the scale in the XSEM picture, we can estimate that the resolution of the image is about 10 nm. In order to study the scatterometry accuracy quantitatively, we need referencing data with higher resolution. Critical dimension scanning electron microscope (CD-SEM) is the workhorse in semiconductor fabs for process monitoring and control. An electron gun scans a small beam (a few nm in diameter) of electrons on the sample, and the detector collects back-scattering or second order scattering electrons during the scan. The scanning trace is then analyzed with various edge-detection algorithms to obtain the width at a certain height of the profile.

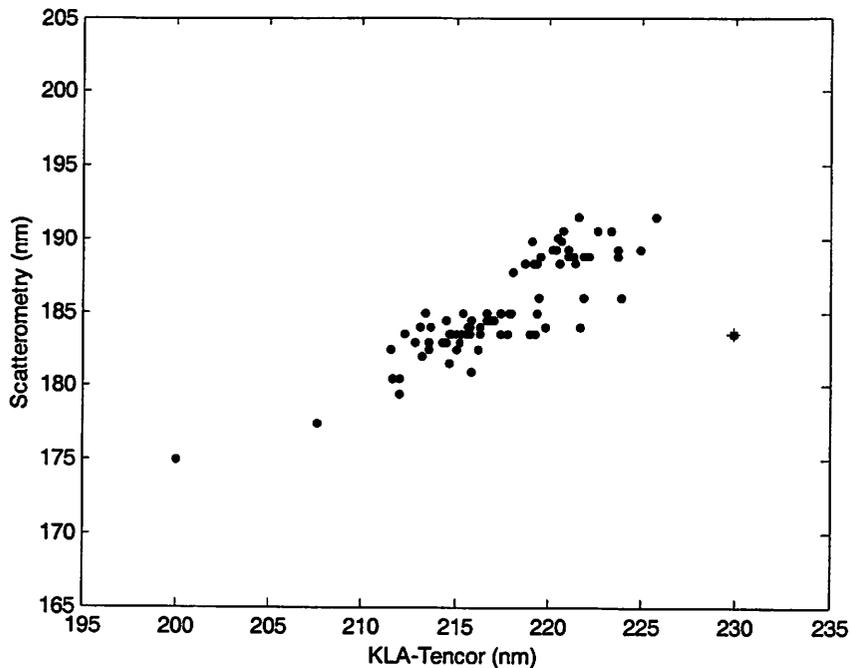


Figure 4-14 Comparison of measured bottom CD using scatterometry and KLA-Tencor CD-SEM for all the dice on a production wafer

Usually for resist structures, the electron beam of the CD-SEM can cause charging effect, thus there is higher uncertainty in the measurement results. We use a wafer with a polysilicon gate FI structure for comparing scatterometry with CD-SEM. The bottom CDs of the testing structures on all the dice of the wafer are plotted in Figure 4-14, where the X-axis is the data from a KLA-Tencor CD-SEM, and the Y-axis is that from scatterometry using the library based method. There is obviously an outlier in the data, as marked with the ‘\*’ label. Without considering this data point, the correlation between CD-SEM and scatterometry is 0.87.

There are various factors that affect the correlation. For scatterometry, there are modeling errors and uncertainty due to the discrete library profiles. For CD-SEM, the CD measured depends on the effect of profile sidewall shapes on the edge detection algorithms. For example, the data point with a CD of about 200 nm measured by CD-SEM may have larger

uncertainty since the scatterometry measurement shows that this profile has a smaller sidewall angle. Furthermore, scatterometry measures the average profiles across the measurement light spot, while CD-SEM measures specific lines.

There is an obvious offset of about 30 nm between the CD-SEM and scatterometry data. This is mainly due to calibration of the CD-SEM edge-detection algorithm, which has also been reported by many other researchers. Depending on the calibration procedure, the offset can be different for tools in different fabs. As long as the offset is consistent for a certain tool, it is not very critical for process monitoring and control.

#### 4.3.4 Precision of the scatterometry system

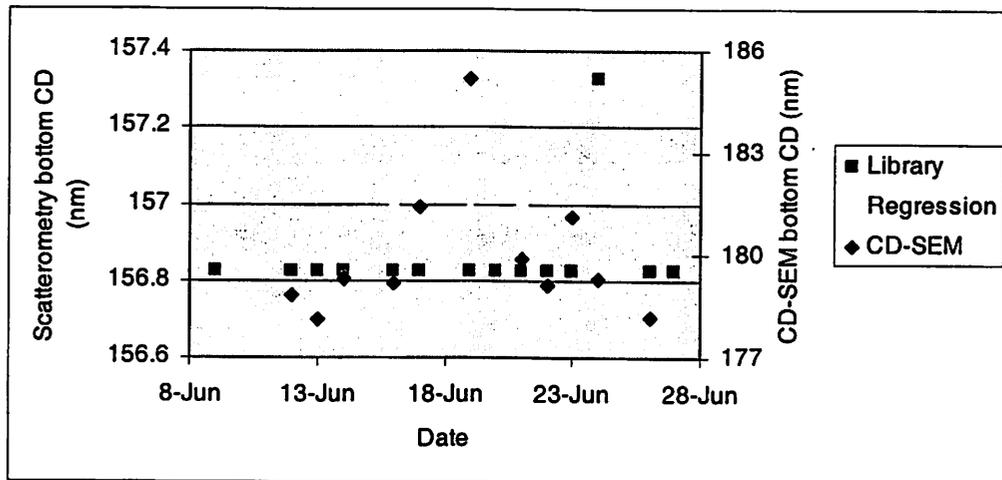


Figure 4-15 Comparison of long-term repeatability results using scatterometry and CD-SEM.  
(Note: scatterometry and CD-SEM data have different scales)

We use the Polysilicon gate FI structure to characterize the long-term repeatability of the scatterometry system. From June 9<sup>th</sup> to June 27<sup>th</sup>, we measured the same location on the same test wafer with the scatterometry system, except weekends and a few days for hardware

maintenance. In the same period, we also measured the same sample with an OPAL CD-SEM for most of the days.

Both regression and library-based methods are used for scatterometry profile extraction. The measured bottom CDs are plotted in Figure 4-15. Because the variation of the measurement spectra is very small, as shown in Figure 4-16, the profiles given by the library-based method are all the same, except the one measured on June 24<sup>th</sup>, which jumps by one step in the library parameter space. The bottom CD  $3\sigma$  from this data is 0.40 nm. However, in order to take into account the effect of discrete profiles, about half of the bottom CD step size needs to be added to the number above, resulting in the actual bottom CD  $3\sigma$  of 0.75 nm. The CDs extracted with the regression method show small variations for these measurements, and they are all distributed between the two discrete CD values of the library-based method. The

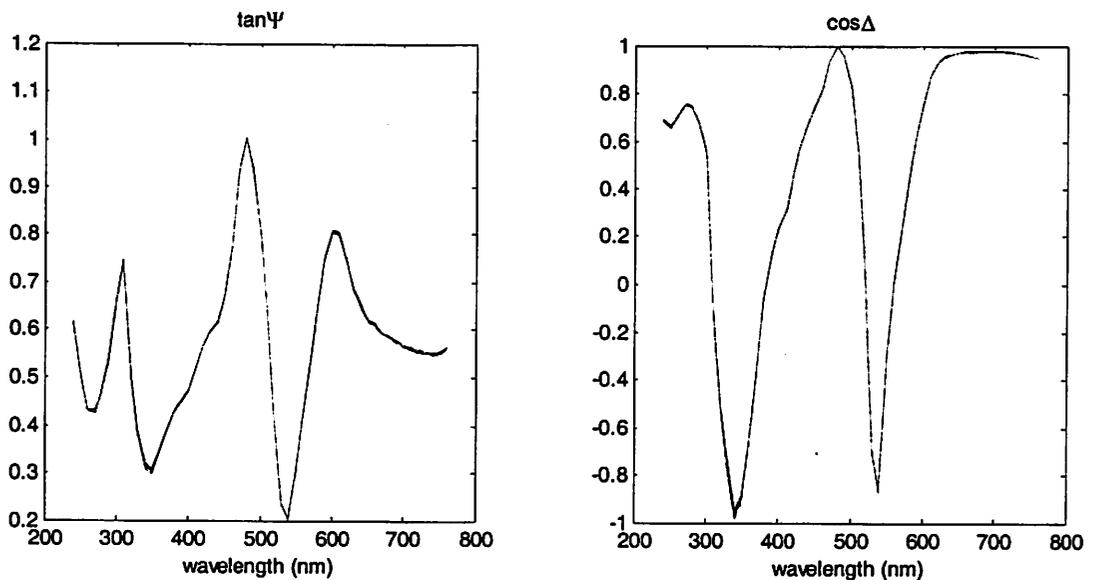


Figure 4-16 Long-term repeatability measurement signals

resulting bottom CD  $3\sigma$  is 0.25 nm.

The measured line widths with CD-SEM during this period are also shown in Figure 4-15. We can see a much larger range of spread than scatterometry. The overall  $3\sigma$  is 6.07 nm. There was periodic tool maintenance during this period, which might have caused the big deviation of the CD measurement on June 19<sup>th</sup>. Without this data point, the measurement  $3\sigma$  is 3.28 nm, which is still significantly higher than that of scatterometry.

#### **4.3.5 Thermawave SE tool light spot size characterization**

One key requirement for scatterometry using RCWA as a simulation algorithm is the periodic sample structure. While the grating structure is available in many memory circuit layouts, it is difficult to find periodic lines with a large area in logic circuits. For today's state-of-the-art broadband optical metrology tools, the test sample area required is about 30  $\mu\text{m}$  for a normal incidence reflectometer, and about 50  $\mu\text{m}$  to 150  $\mu\text{m}$  for an elliptically incidence reflectometer or ellipsometer. Usually people can design this kind of testing structure on a scribe line. But with the progress of the semiconductor process development, the width of the scribe line will soon be shrunk to less than 50  $\mu\text{m}$ . Furthermore, there are many other test modules competing for the precious scribe line area. Therefore, it is important to determine the minimum testing sample area.

The test sample area specification takes into account both light spot size and stage location error due to pattern recognition and mechanical movement. The pattern recognition error is typically less than 10  $\mu\text{m}$ . There are two factors that determine the light spot size. One is the size of the ideal image of the light source; the other is the light spread due to diffraction. Since

the RCWA simulation typically assumes a parallel incident light beam for fast simulation speed, the numerical aperture (NA) of the system cannot be too large. This sets the minimum light

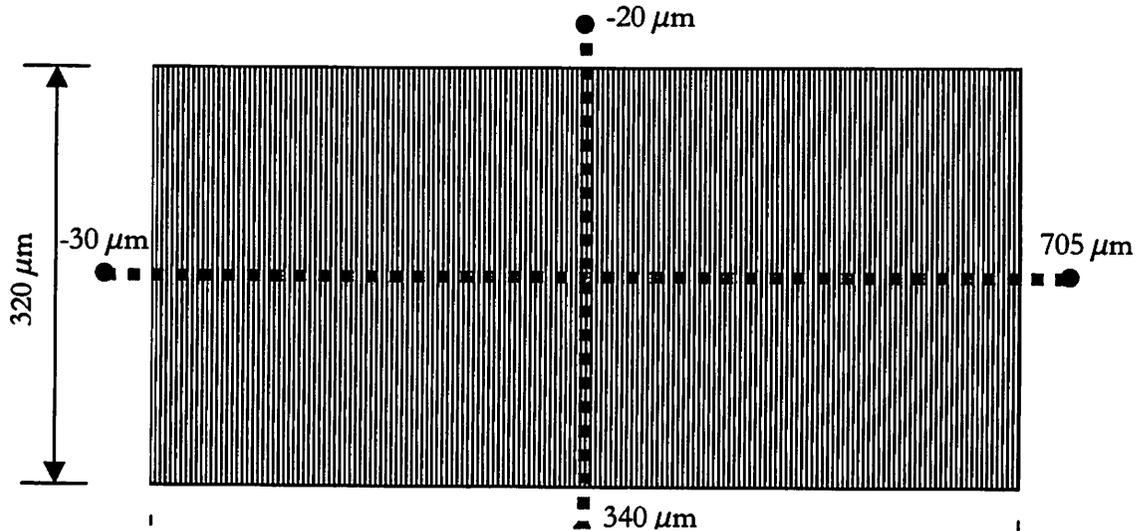


Figure 4-17 Testing grating for light spot size characterization

spread due to diffraction. While one can theoretically calculate each of the factors that contribute to the final light spot, or measure light spot size with thin film testing areas, it is also important to characterize the effect of light spot to the measurement result for grating structures.

A large area of grating is used for this study, where the structure is STI FI stacks. As shown in Figure 4-17, the size of the grating area is 675  $\mu\text{m}$  by 320  $\mu\text{m}$ , and the grating lines go in the vertical direction. By doing measurements in horizontal and vertical scans, and analyzing the spectra, we can characterize the light spot size.

We use the line scan function of the OptiProbe ellipsometer to measure 16 sites along the horizontal direction, and 15 sites along the vertical direction. The distances from the left edge of the grating to the centers of the first and the last measurement sites of the horizontal scan, as well as those from the top edge to the sites of the vertical scan, are shown in Figure 4-17.

The distances from the left and the top edges for all the sites of these scans are listed in the first rows of Table 4-4 and Table 4-5, respectively.

Since the grating area is more than three times larger than the light spot size in specification, it is guaranteed to hold the light spot completely for the measurements at the center of the grating. One way to tell at which measurements part of the light lands outside of the grating region is to correlate all the measured spectra to the one measured at the grating center, which is site 8 for both horizontal and vertical scans. The resultant correlation coefficients are listed in Table 4-4 and Table 4-5. It is obvious from the tables that if we take 0.99 at the threshold value, significant parts of the light spots for the first two measurements from the starting and the ending scanning points fall outside of the grating region, and a small amount of light for the third measurement goes outside. This confirms the specification from the tool vendor that the light spot is less than 80  $\mu\text{m}$  by 120  $\mu\text{m}$ .

Table 4-4 Correlation of measured spectrum to site 8 for horizontal scan

Location ( $\mu\text{m}$ )	-30	0	30	60	90	120	207	294
Correlation to center	-.3617	-.3548	0.9873	0.9930	0.9947	0.9942	0.9983	1.0000
Location ( $\mu\text{m}$ )	381	468	555	585	615	645	675	705
Correlation to center	0.9989	0.9992	0.9979	0.9990	0.9986	0.9921	-.0840	-.3618

Table 4-5 Correlation of measured spectrum to site 8 for vertical scan

Location ( $\mu\text{m}$ )	-20	0	20	40	60	80	120	160
Correlation to center	-.2529	-.3272	0.9377	0.9991	0.9990	0.9996	0.9998	1.0000
Location ( $\mu\text{m}$ )	200	240	260	280	300	320	340	
Correlation to center	0.9997	0.9995	0.9992	0.9995	0.9980	0.4403	-.2748	

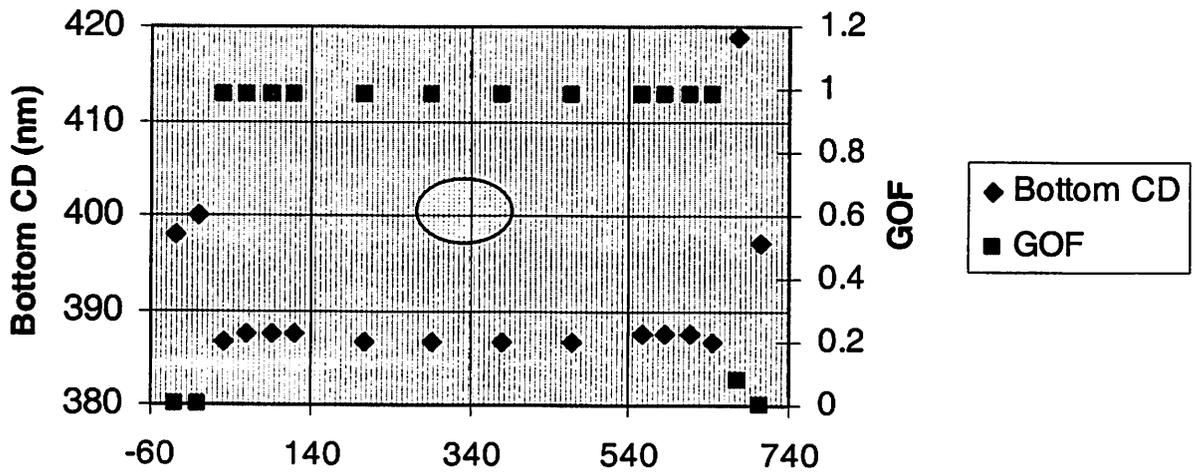


Figure 4-18 Bottom CD and GOF for horizontal line scan

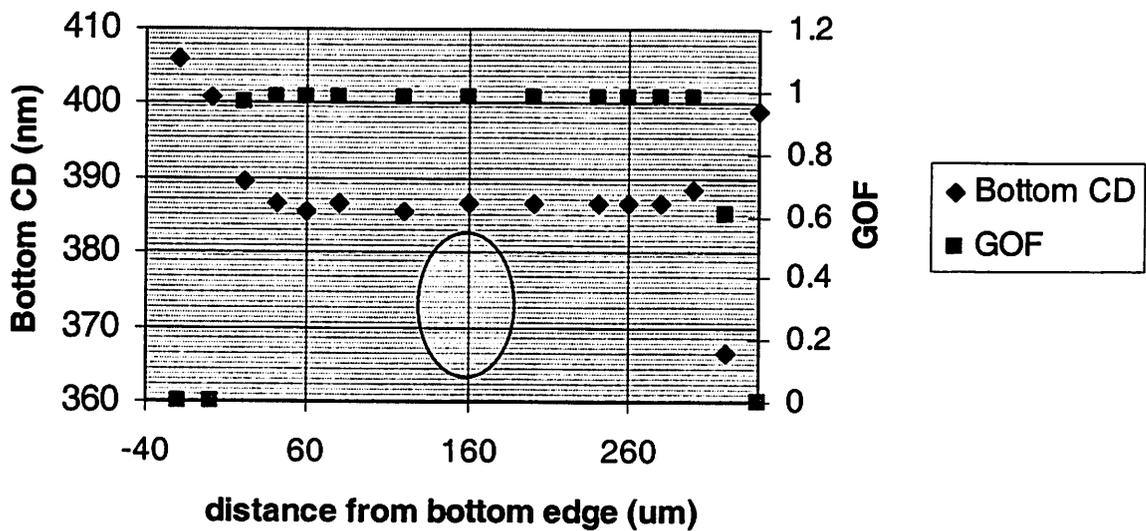


Figure 4-19 Bottom CD and GOF for vertical line scan

The output of the scatterometry measurement is grating line profiles. Therefore, a more direct way of characterizing light spot size is to see when the output profiles change for the line-scan measurements. We run the measured spectra to a pre-built library, and the output bottom CDs

and GOF are plotted in Figure 4-18 and Figure 4-19 for horizontal and vertical scans, respectively. The orientations of the gratings with respect to the scanning direction (x axis), as well as the approximate light spot, are also illustrated. From these figures, we can reach a similar conclusion to that obtained using spectrum correlation.

## 4.4 Scatterometry application on metal process

### 4.4.1 Scatterometry work on metal applications

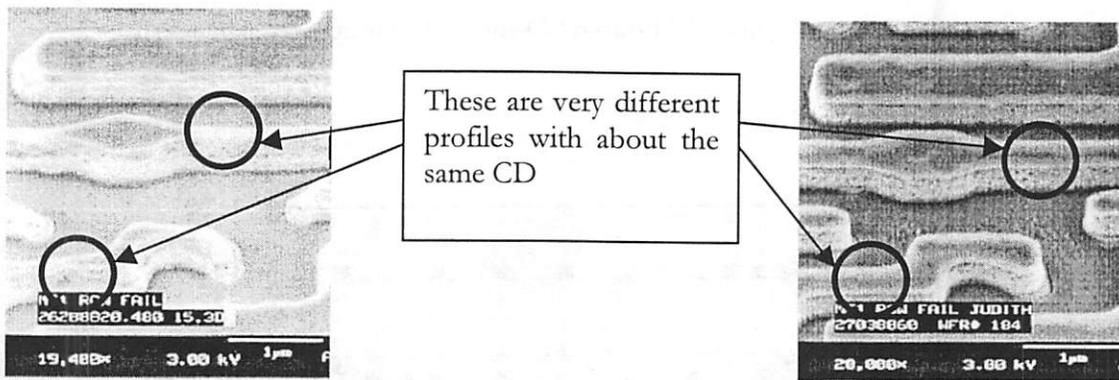


Figure 4-20 Example of metal interconnect profile variation

Besides the applications introduced in the last section, another key application for scatterometry is metal interconnect layers. With continuing technology scaling into the deep

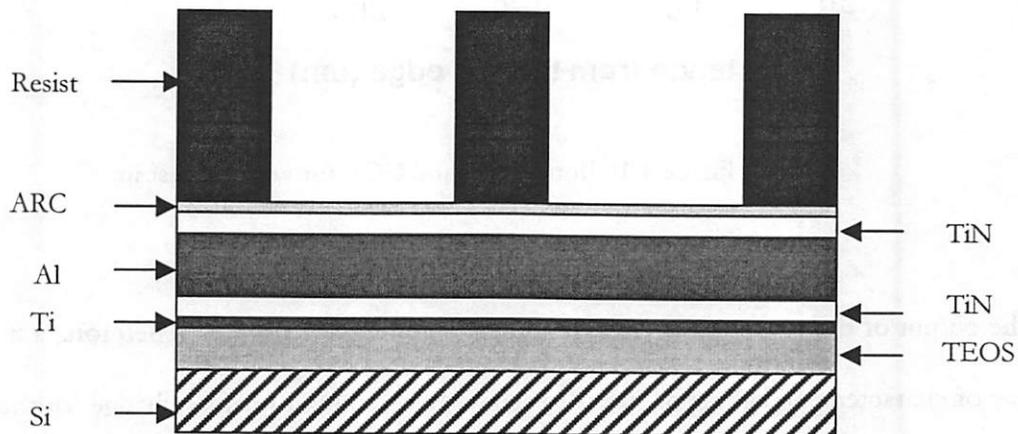


Figure 4-21 Illustration of metal DI structure

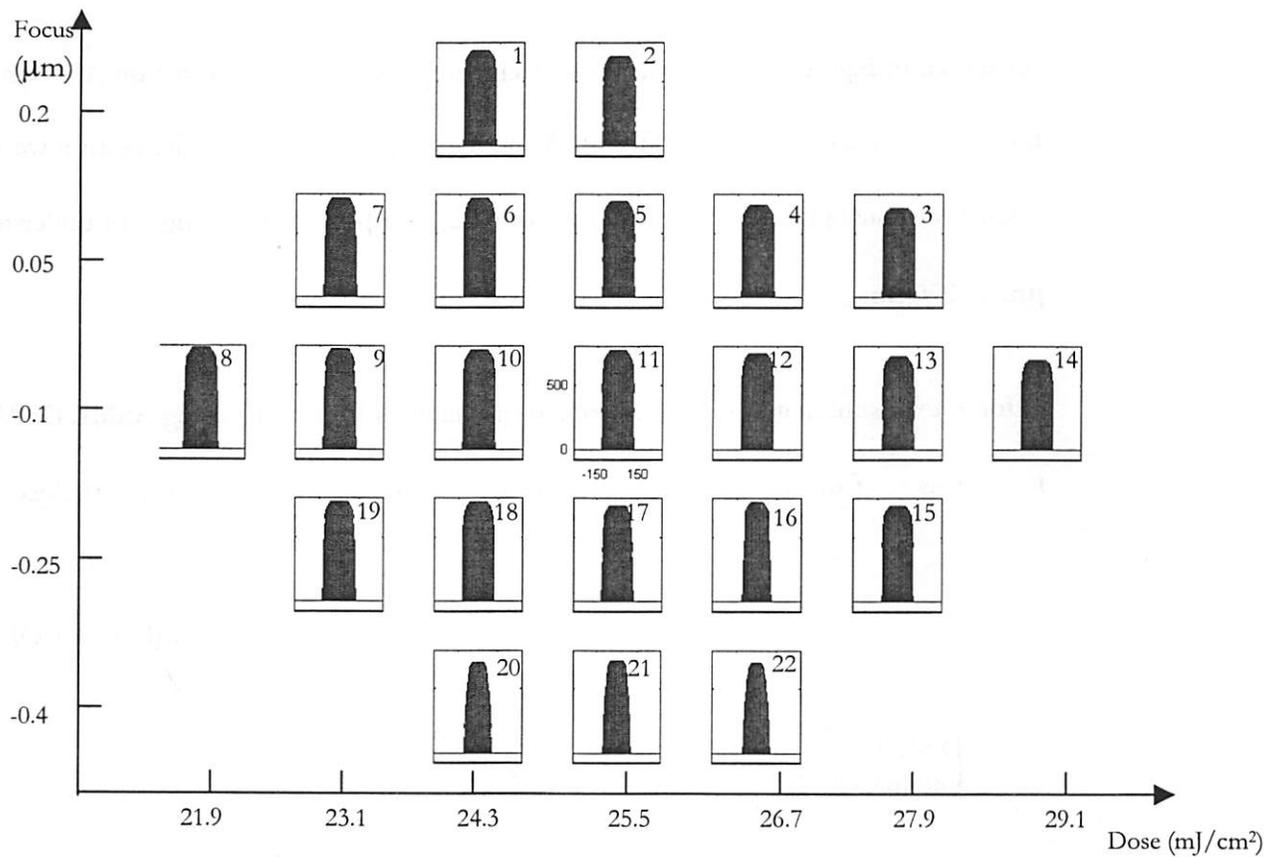


Figure 4-22 Extracted FEM resist profiles plotted as their positions on the wafer. All figures are in the same scale. The numbers on the upper right corners are site numbers.

sub-micron regime, interconnect constitutes an increasing portion of the overall circuit delay [4.19]. Not only line width, but also height and sidewall profile measurements are required to accurately characterize RC delay. However, with today's standard process monitoring tool, CD-SEM, the structures with different profiles in Figure 4-20 will have about the same CD, while their RC characteristics are quite different. Instead, scatterometry can be used here for better monitoring the metal process. In this section, we demonstrate scatterometry measurement results for both metal DI and FI structures.

As shown in Figure 4-21, the metal DI stack comprises developed resist on ARC on thin film metal layers, which consist of TiN on Al on TiN on Ti on TEOS. The feature we study is a one-dimensional grating, with line/space of 0.22/0.44  $\mu\text{m}$ . The grating area under test is 250  $\mu\text{m}$  by 250  $\mu\text{m}$ .

A focus exposure matrix (FEM) is printed on Shipley's UV6 resist using ASML DUV stepper. Focus was set from  $-0.4 \mu\text{m}$  to  $0.2 \mu\text{m}$  with a step of  $0.15 \mu\text{m}$ , and exposure doses were set

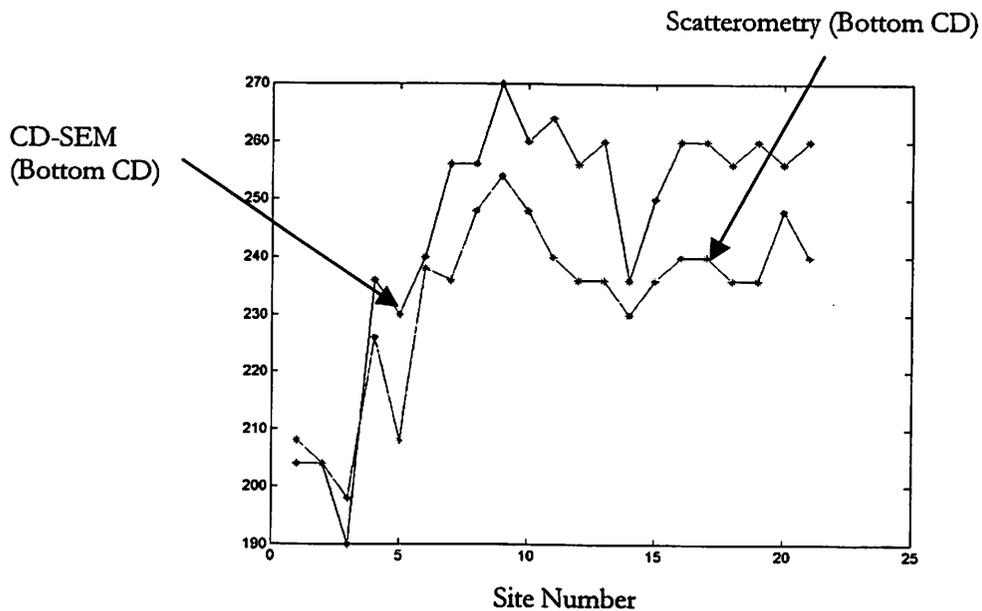


Figure 4-23 Comparison of CD-SEM and scatterometry bottom CD

from  $21.9 \text{ mJ}/\text{cm}^2$  to  $29.1 \text{ mJ}/\text{cm}^2$  with a step of  $1.2 \text{ mJ}/\text{cm}^2$ . After exposure and PEB, the resist is developed. A KLA-Tencor 1280 spectroscopic ellipsometer is used to measure the response at an incidence angle of  $70^\circ$ . Twenty-two dies were measured, and the extracted resist profiles are shown in Figure 4-22 in the same layout as they are on the wafer. The profile distribution across the wafer looks quite reasonable for an FEM.

We also did CD-SEM measurements on these gratings, and compared the bottom CD measurement produced from scatterometry and CD-SEM, as plotted in Figure 4-23. The two data sets show good correlation. The deviation is mostly due to the constant offset for the CD-SEM, as explained in section 4.3.3. Some of the deviations between the two metrologies are not consistent with others. This is partly due to the charging effect of the CD-SEM on resist features. Upon further examination, we found that those deviations correspond to profiles that have small sidewall angles or some footing.

The FEM was then patterned onto the metal layer using plasma etching, and the resist and ARC were removed. The resulting stack of patterned metal features (TiN on Al on TiN on Ti) on TEOS thin film on silicon wafer is shown in Figure 4-24.

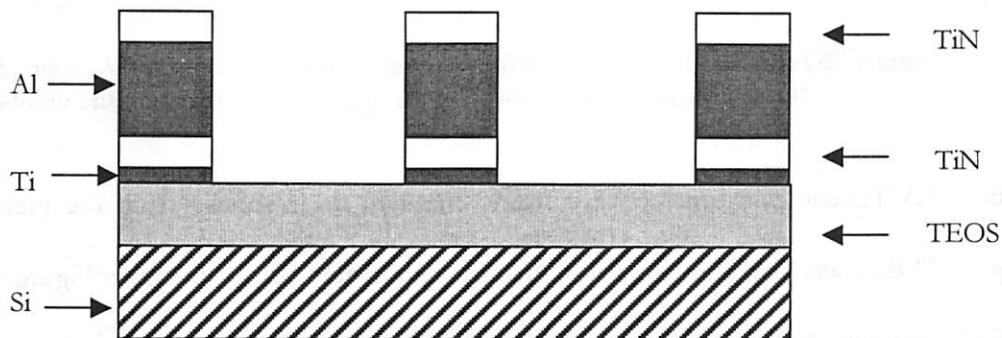


Figure 4-24 Illustration of the metal FI structure

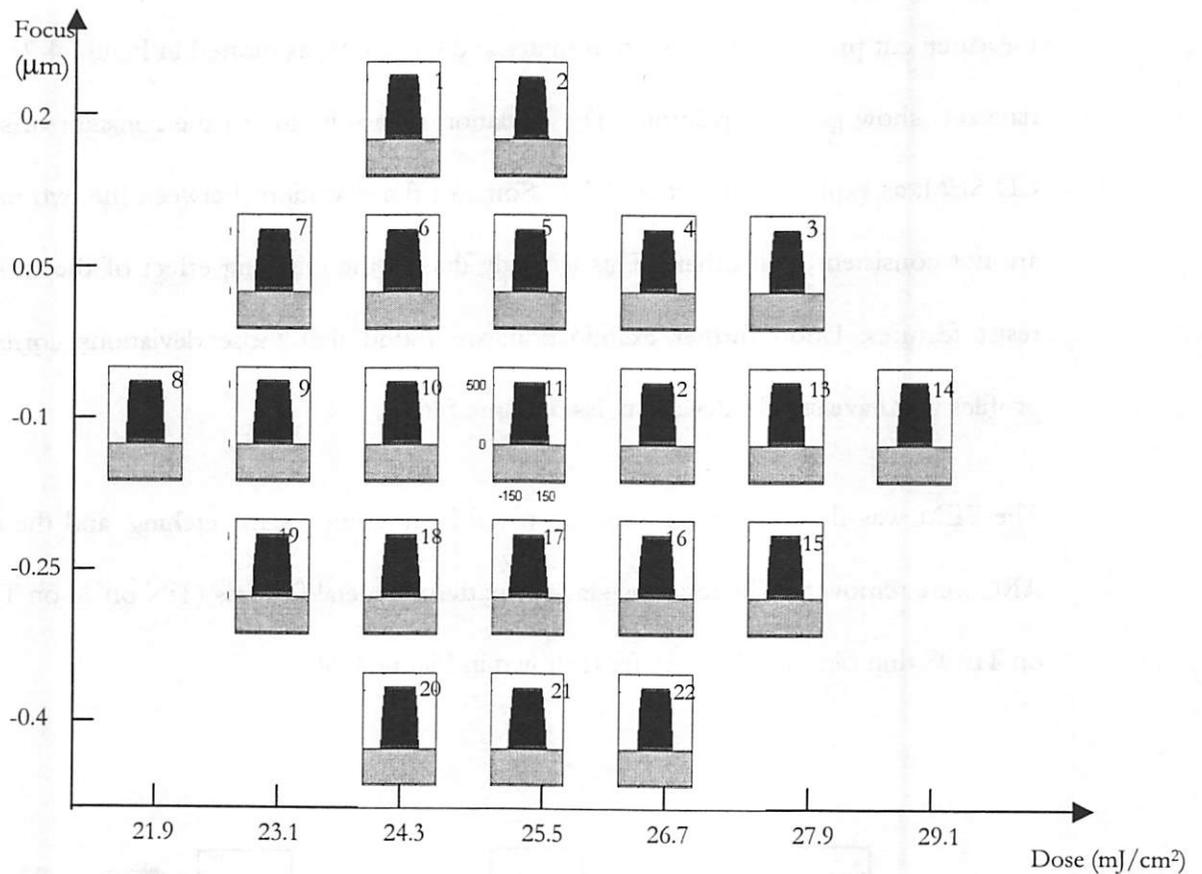


Figure 4-25 Extracted FEM metal profiles plotted as their positions on the wafer. All figures are in the same scale. The numbers on the upper right corners are site numbers.

The KLA-Tencor ellipsometer was used to measure the response from the metal gratings. Again 22 dice were measured and the profiles were extracted, as shown in Figure 4-25. CD-SEM measurements were also made, and the comparison of the top CD from both methods is plotted in Figure 4-26. We can see that the deviation between CD-SEM and profilometry is more consistent than the result for resist features in Figure 4-23, because the sidewalls for the metal profiles are fairly straight compared to those for the resist profiles. We also note that the SEM charging effect that can distort photoresist measurements is much less significant when measuring metal features. This helps the CD-SEM get more consistent results.

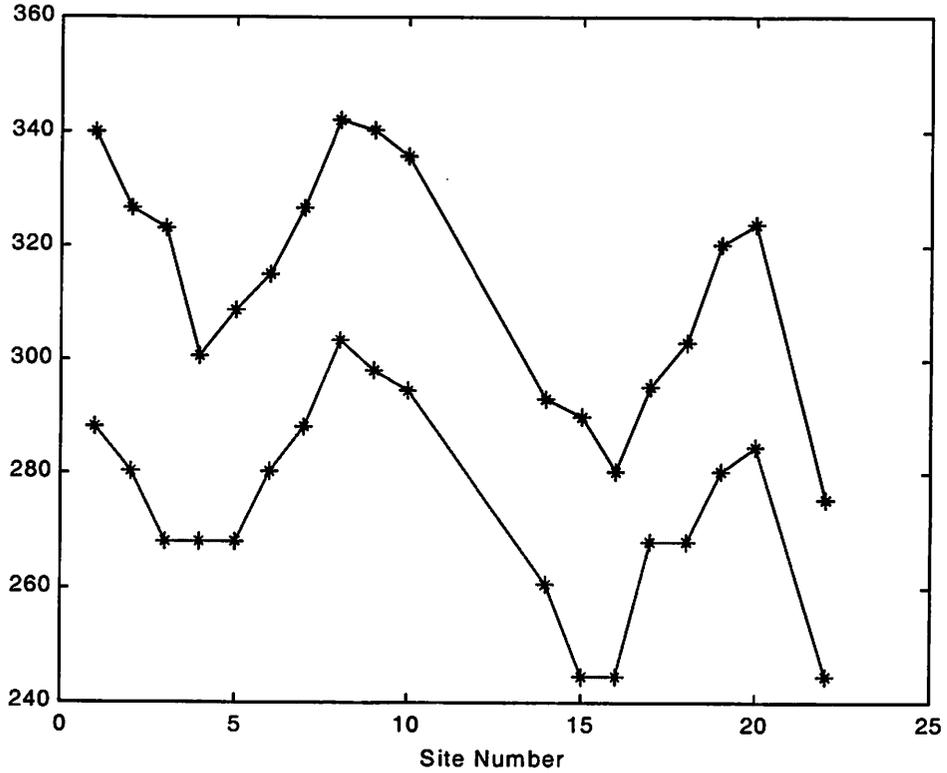


Figure 4-26 Comparison of scatterometry and CD SEM top CD for metal features. CD SEM has no valid measurements at the missing data points.

To study the effect of sidewall angles on CD-SEM measurements, we take our scatterometry as the reference, and plot the deviation of CD-SEM measurement versus sidewall angle, as shown in Figure 4-27. We can see that for smaller sidewall angles, the distribution of deviations between the two metrologies is larger. This might be because the CD-SEM algorithm cannot easily determine top/bottom boundaries when the profiles are not very steep.

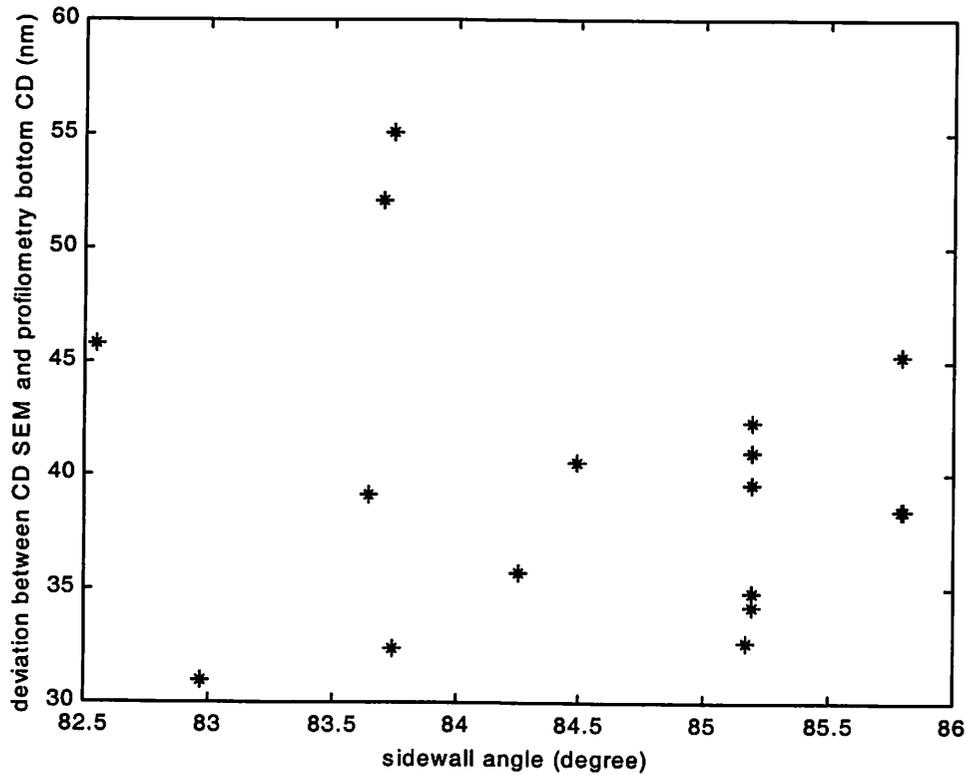


Figure 4-27 Deviations between CD-SEM and scatterometry bottom CD versus sidewall angle.

#### 4.4.2 Potential application to CMP end-point detection

Damascene processes bring a new, difficult challenge to CD measurement and end point detection. Several methods have been proposed based on CD-SEM [4.20]. One is to measure CD on the dielectric lines, but this suffers from sample charging effect. Another approach is to measure the dielectric trench structure after the barrier layer is deposited. Although this would reduce the charging problem, sidewall shapes could not be characterized, and the electron beam may damage the dielectric layer and make the process more complicated.

Scatterometry is a promising candidate for measuring trench profiles with or without metal filled in. It can also be used for end point detection during the CMP process, as illustrated in Figure 4-28. As the metal layer is being polished, the response from the structure can be

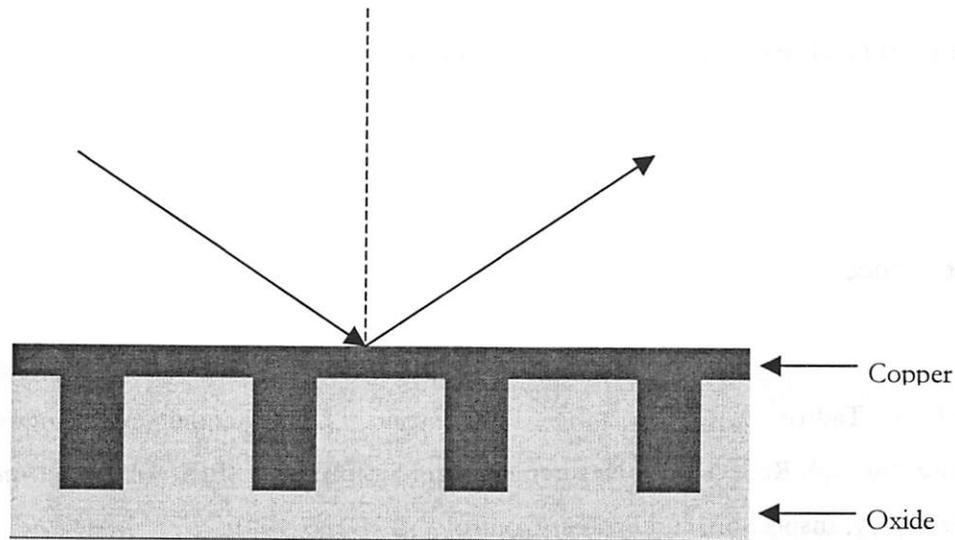


Figure 4-28 One-dimensional structure for CMP end-point detection

measured using ellipsometry or reflectometry. At first, the signal is much like that from metal film, because most of the light going through has been absorbed. When the metal layer gets thinner, it becomes more transparent, and the scattered light from the structure below can be detected. If the metal film thickness is included as a variable in the library, the remaining film thickness can be measured and used for end point detection.

## 4.5 Summary

In this chapter, we first reviewed some early experimental work on various hardware configurations, and then presented the experimental results for a scatterometry system characterization using Therma-Wave's Opti-Probe 5240 ellipsometer system. Application of scatterometry to metal DI and FI structure using KLA-Tencor's 1280 ellipsometer, as well as its potential application to CMP end-point detection, was also demonstrated. The experimental data presented in this chapter show that the accuracy, precision, and throughput of scatterometry is comparable or better than CD-SEM. Together with its full-profile

measurement capability and inline compatibility, scatterometry can fit very well in the next generation advanced process control framework.

## Reference

- [4.1] K. Tadros, A.R. Neureuther, R. Guerrieri, "Understanding Metrology of Polysilicon Gates through Reflectance Measurements and Simulation," SPIE vol. 1464, Integrated Circuit Metrology, Inspection, and Process Control V, 177-186, 1991.
- [4.2] J. Gamelin, R. Guerrieri, A.R. Neureuther, "Exploration of Scattering from Topography with Massively Parallel Computers," J. vac. Sci. Technol. B., 7(6): 1984-1990, Nov/Dec, 1989.
- [4.3] R.H. Krukar, S.M. Gaspar, S.R. Wilson, D.R. Hush, S.S. Naqvi, J.R. McNeil, "Wafer Examination and Critical Dimension Estimation Using Scattered Light," SPIE vol. 1661, Machine Vision Applications in Character Recognition and Industrial Inspection, 323-332, 1992.
- [4.4] C.J. Raymond, M.R. Murnane, S.L. Prins, S.S. Naqvi, J.R. McNeil, "Multi-parameter Process Metrology Using Scatterometry," SPIE Vol. 2638, Optical Characterization Techniques for High-Performance Microelectronic Device Manufacturing II, 84-93, 1995.
- [4.5] C.J. Raymond, S.S. Naqvi, J.R. McNeil, "Scatterometry for CD Measurements for Etched Structures," SPIE Vol. 2725, Metrology, Inspection and Process Control for Microlithography X, 720-728, 1996.
- [4.6] S.M.G. Wilson, S.S. Naqvi, J.R. McNeil, H.M. Marchman, B. Johs, R.H. French, F.D. Kalk, "Metrology of Etched Quartz and Chrome Embedded Phase Shift Gratings Using Scatterometry," SPIE vol. 2439, Integrated Circuit Metrology, Inspection, and Process Control IX, 479-494, 1995.
- [4.7] C.J. Raymond, S.W. Farrer, S. Sucher, "Scatterometry for the Measurement of Metal Features," SPIE vol. 3998, Metrology, Inspection and Process Control for Microlithography XIV, 135-146, 2000.

- [4.8] J. Bischoff, J.W. Baumgart, H. Truckenbrodt, J.J. Bauer, "Photoresist Metrology based on Light Scattering," SPIE Vol. 2725, Metrology, Inspection and Process Control for Microlithography X, 678-689, 1996.
- [4.9] C.C. Baum, R.A. Soper, S. Farrer, J.L. Shohet, "Scatterometry for Post-etch Polysilicon Gate Metrology," SPIE vol. 3677, Metrology, Inspection and Process Control for Microlithography XIII, 148-158, 1999.
- [4.10] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors," <http://public.itrs.net/Files/2002Update/Home.pdf>, 2002.
- [4.11] X. Niu, N. Jakatdar, J. Bao, C.J. Spanos, S. Yedur, "Specular Spectroscopic Scatterometry in DUV Lithography," SPIE vol. 3677, Metrology, Inspection and Process Control for Microlithography XIII, 159-168, 1999.
- [4.12] B.S. Stutzman, H.T. Huang, F.L. Terry, "Two-Channel Spectroscopic Reflectometry for in situ Monitoring of Blanket and Patterned Structures during Reactive Ion Etching," Journal of Vacuum Science & Technology B, vol. 18, (no. 6), 2785-2793, Nov. 2000.
- [4.13] E. Maiken, K. Johnson, D. Likhachev, A. Norton, A. Sezginer, F. Stanke, D. Yonenaga, "Integrated Critical Dimension Metrology for Lithography Process Control," AEC/APC Symposium XIII, Banff, Canada, October 6-11, 2001.
- [4.14] J.M. Holden, V. Zhuang, W. McGahan, "Optical CD for Metrology and Control of Line and Space Critical Dimension," AEC/APC Symposium XIII, Banff, Canada, October 6-11, 2001.
- [4.15] K.R. Lensing, R.J. Markle, B. Stirton, M.A. Laughery, "Shallow Trench Isolation Scatterometry Metrology in a High Volume Fab," Proceedings of IEEE International Symposium on Semiconductor Manufacturing, 195-198, 2001
- [4.16] V.A. Ukraintsev, M. Kulkarni, C. Baum, K. Kirmse, M. Guevremont, S. Lakkapragada, K. Bhatia, P. Herrera, U. Whitney, "Spectral Scatterometry for 2D Trench Metrology of Low-K Dual-damascene Interconnect," SPIE vol. 4689, Metrology, Inspection and Process Control for Microlithography XIII, 189-195, 2002.

[4.17] J. Opsal, H. Chu, Y. Wen, Y.C. Chang, G. Li, "Fundamental Solutions for Real-time Optical CD Metrology," SPIE vol. 4689, Metrology, Inspection and Process Control for Microlithography XIII, 163-176, 2002.

[4.18] Therma-Wave, Inc., 1250 Reliance Way, Fremont, California 94539.

[4.19] A. K. Stamper, M. B. Fuselier, X. Tian, "Advanced Wiring RC Delay Issues for Sub-0.25-micron Generation CMOS," Proceedings of the IEEE International Interconnect Technology Conference, 1998, 62-64.

[4.20] A. C. Diebold, R. K. Goodall, "Interconnect Metrology Roadmap: Status and Future," Proceedings of the IEEE International Interconnect Technology Conference, 1999, 77-79.

# Chapter 5 Lithography Process Monitoring and Control Using Scatterometry

## 5.1 Introduction

In Chapter 4 we presented some experimental scatterometry results. The accuracy, precision, throughput, and amount of profile information provided by scatterometry show a clear advantage over CD-SEM, which is the current standard metrology tool for process monitoring and control. However, there has not been much work done on applying the full-profile information obtained from scatterometry to process control. Since both the line width and sidewall angle can be measured with scatterometry, many of the current APC methods need to be adapted to make use of the additional sidewall angle information. In this chapter we first review the work on lithography process control by other researchers, and then demonstrate some experimental results on focus and exposure monitor using scatterometry. After some discussion on lithography process parameter extraction, we present a simulation framework for lithography process monitoring and control using scatterometry.

## 5.2 Review of previous work on lithography process monitoring and control

### 5.2.1 Background

As the semiconductor industry is approaching the sub-0.1  $\mu\text{m}$  technology node on the SIA roadmap [5.1], critical dimension (CD) control becomes more and more important. One major source of CD variation is the parameter drift in the lithography process. Figure 5-1 illustrates a

typical DUV lithography process flow. After deposition of under-layer thin film materials or processes in previous steps, a thin layer of photoresist is spun on the wafer and softbaked, and the wafer is sent to the stepper for exposure. During the exposure process, photoacid generators in the photoresist produce acid when reacting with DUV photons. Then, the wafer is brought back to the wafer track for post-exposure-bake (PEB), in which the acid diffuses and acts as a catalyst to induce deprotection. In the develop step, the deprotected molecules are resolved by the developer, and removed after the rinse. Finally the wafer is etched with an optional resist trim step in plasma chambers to transfer the pattern to under layer materials.

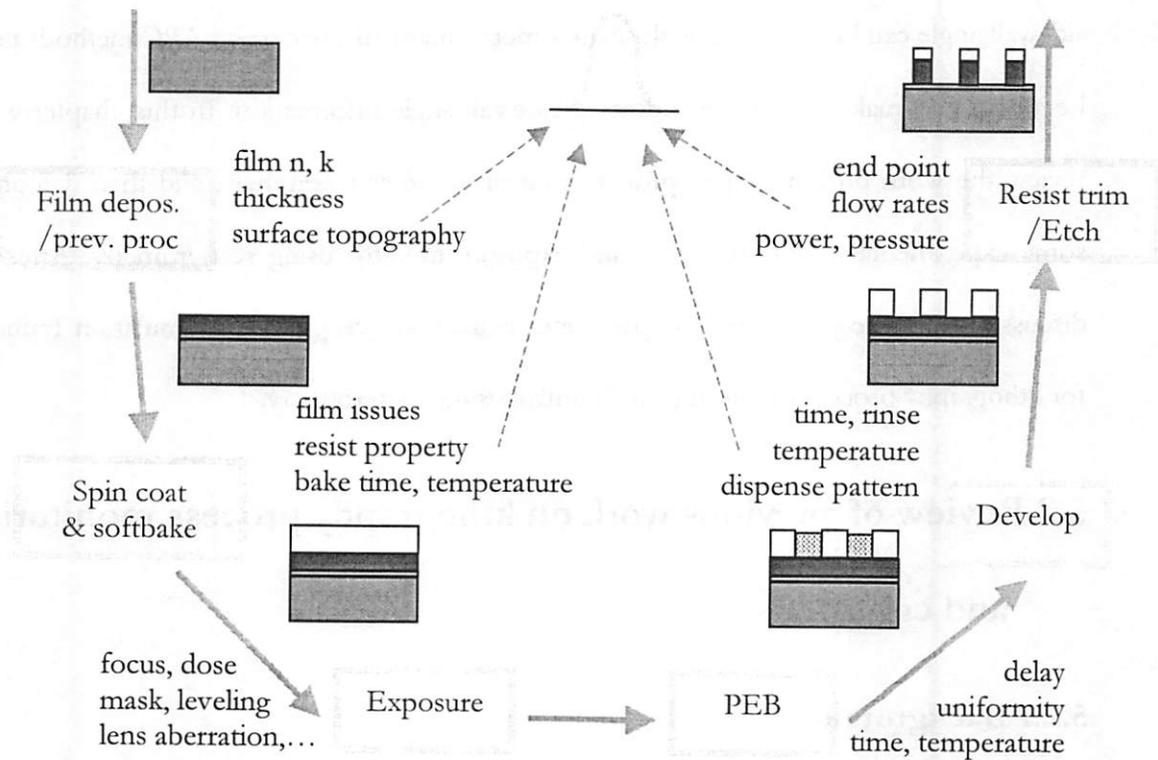


Figure 5-1 Lithography process flow and sources of process variation

As listed in Figure 5-1, there are many process and material parameters that may drift in the lithography flow, for example, film refractive index and thickness for under layer materials in the thin film deposition step, resist property and bake time and temperature in the spin-coating and softbake step, focus, dose, mask and optical parameters in the exposure step, bake time and temperature in the PEB step, developer rinse pattern and temperature in the develop step, and plasma etching control parameters in the resist trim step. The drifts of all these parameters contribute to the variation of the resist profiles that will affect the etched line width after pattern transfer. Therefore, it is important to monitor and compensate these drifts through advanced process and equipment control.

### **5.2.2 CD-based lithography process control method**

Various methods have been proposed to monitor and compensate for process drift in order to reduce CD variation. One of the early works presented at AEC/APC workshop IX showed that by integrating the metrology tool and process equipment into a CD run-to-run control framework, the Cpk can be increased by 67%, resulting in a potential \$1 million per week added revenue [5.2]. Stuber and coworkers also demonstrated similar improvement over an open-loop process with process control through partitioning of model parameters [5.3].

All the works mentioned above involve building an empirical relation between the CD and exposure dose, and adjusting the dose to compensate the CD drift, while CD-SEM is used as the metrology tool. However, the sub-0.1 $\mu$ m patterning process has greatly reduced the depth of focus, and the printed structures are very sensitive to the stepper focus drift. Due to the requirement of later process steps, a tight control of the sidewall angle is also needed, so it is not desirable to compensate the focus shift by adjusting the exposure.

### 5.2.3 Recent development on lithography process control using scatterometry

With the recent progress of scatterometry, many people have applied it as a tool for routine process monitoring. Allgair and coworkers [5.4] demonstrated that spectroscopic scatterometry is able to characterize a focus and exposure matrix (FEM), and showed good correlation to a top-down CD-SEM and cross-section SEM. The scatterometry measurements of resist height and sidewall angle illustrate clearly how they are affected by focus and exposure dose changes, and can be used to monitor stepper condition. Furthermore, scatterometry can also be used as a fault-detection tool to detect whether the process is out of the control window.

Lafferty et al. compared the performance of  $2\theta$  scatterometry to CD-SEM for process control [5.5]. The gauge capability study on a metal-level product step showed that scatterometry has a better precision to tolerance ratio than CD-SEM. When comparing lithographic cell monitor trends, a shift in the process halfway through the trending chart resulted in a 10 nm increase in CD-SEM measured CD. If the same sample is measured with a scatterometer, the results showed that it was actually due to a  $1^\circ$  increase in the sidewall angle since full profile information can be obtained. From this example, it is clearly shown that it is important to detect the correct source of process drifting (focus in this example) using scatterometry's full profile information.

It has also been demonstrated that scatterometry can be used to monitor *in-situ* latent image formation during post exposure baking (PEB) of a chemically amplified resist [5.18]. During PEB, the refractive index of the exposed resist becomes different from that of the unexposed

part, which can be detected by a scatterometric sensor. Using data from the first 45 seconds of a nominal 60 seconds PEB process for 0.35  $\mu\text{m}$  resist lines, developed linewidth predictions were achieved with standard prediction error of 5.08 nm for bare Si wafers and 6.89 nm for poly-Si/oxide/Si wafers.

## 5.3 Focus-exposure monitor experiment

### 5.3.1 Background

In the lithography process, focus and exposure dose are the two most convenient parameters to adjust to affect resist line width and profile. On the other hand, these two parameters are also subject to drifting. On a state-of-the-art DUV stepper, typically focus offset can drift a few tenths of a micron from one eight-hour work shift to another, and exposure dose may also drift a few percent due to laser light source instability. Usually in production fabs, a focus-exposure matrix (FEM) is printed at the beginning of each work shift in order to calibrate the focus offset and optimal dose. A matrix of dies with incremental focus settings in one direction and incremental dose settings in the other direction is exposed. Then the wafer is inspected and the resist line width distribution for different focus-dose settings is plotted on Bossung curves. The best focus and dose settings are selected at the point with the largest depth of focus and closest to the target CD. This manual calibration process is illustrated in

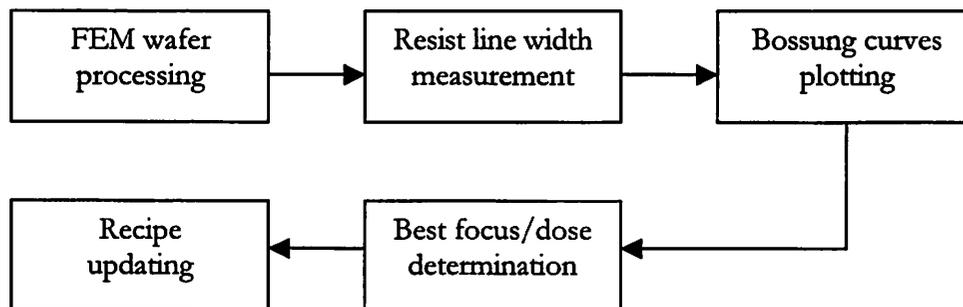


Figure 5-2 Manual focus-dose calibration procedure with FEM

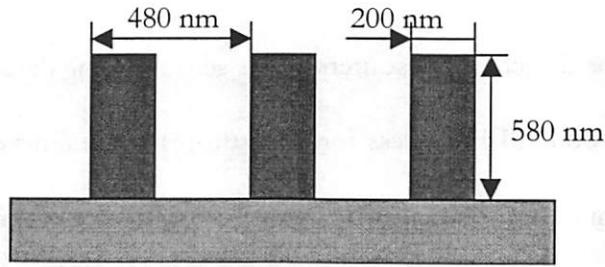


Figure 5-3 Resist grating structure used for FEM characterization

Figure 5-2.

Usually CD-SEM is used as the inspection tool for this purpose. If there is a test one-dimensional grating structure on the wafer, it is better to use scatterometry for FEM characterization, since scatterometry can provide more accurate and precise profile information.

### 5.3.2 FEM wafer processing

In this section, we demonstrate some experimental results on determining the optimal focus

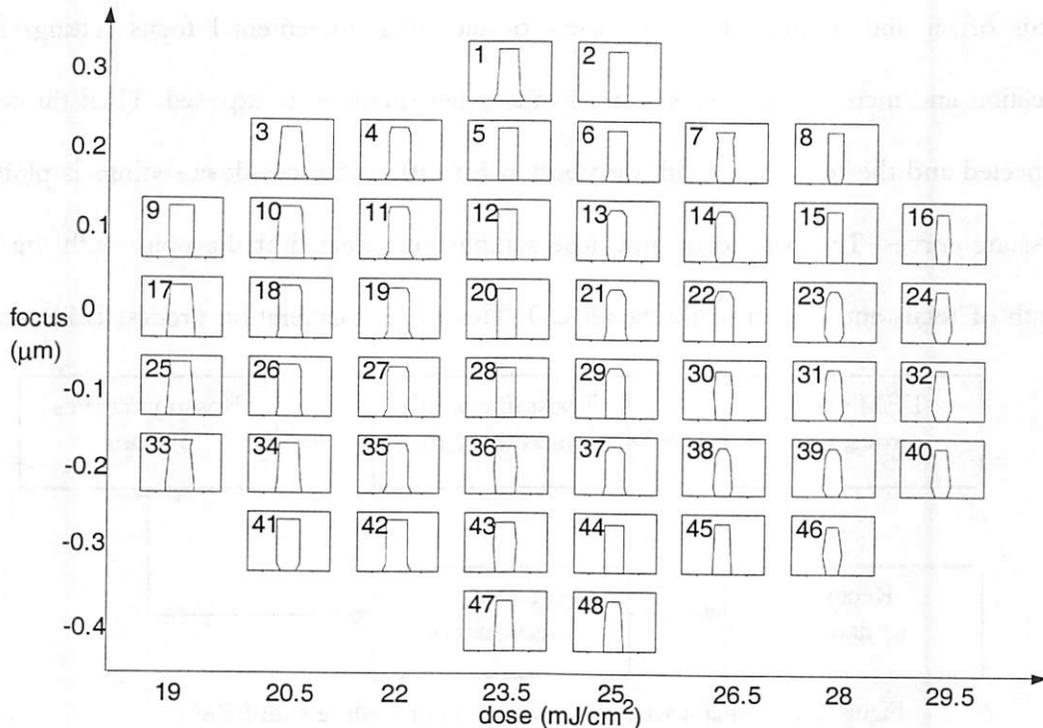


Figure 5-4 FEM resist profile across the full wafer map

and dose settings by inspecting an FEM wafer using scatterometry. A simple structure of a resist grating on a silicon wafer is used in this demonstration. As shown in Figure 5-3, the resist thickness is 580 nm, the pitch is 480 nm, and the target line width is 200 nm.

The wafer is processed on a wafer track and a DUV (248 nm) stepper, where an annular light source is used for exposure. The partial coherence of the inner and outer apertures is 0.25 and 0.75, respectively. The center focus and dose of the FEM are set at the nominal process recipe values, 0.0  $\mu\text{m}$  and 23.5  $\text{mJ}/\text{cm}^2$ , and their steps are 0.1  $\mu\text{m}$  and 1.5  $\text{mJ}/\text{cm}^2$ , respectively. All the other process parameters are also set at the nominal condition, and their values are listed in Table 5-1.

Table 5-1 Process parameters for the FEM wafer

Wafer track parameters		Stepper parameters	
Softbake time (sec)	90	FEM center dose ( $\text{mJ}/\text{cm}^2$ )	23.5
Softbake temperature ( $^{\circ}\text{C}$ )	100	FEM dose step ( $\text{mJ}/\text{cm}^2$ )	1.5
PEB time (sec)	90	FEM focus center ( $\mu\text{m}$ )	0.0
PEB temperature ( $^{\circ}\text{C}$ )	110	FEM focus step ( $\mu\text{m}$ )	0.1
Develop time (sec)	60	Stepper numerical aperture	0.54
Developer temperature ( $^{\circ}\text{C}$ )	20	Stepper partial coherence (inner)	0.25
Wafer rinse time (sec)	20	Stepper partial coherence (outer)	0.75

### 5.3.3 FEM characterization using scatterometry

After the wafer was processed, we measured the spectra reflected from the testing grating area using a Thermo-Wave [5.6] OptiProbe ellipsometer, and extracted the resist line profiles with the regression method described in Chapter 3, as plotted in Figure 5-4. In general, the resist line widths show a clear trend as the exposure dose increases from left to right, but the profile

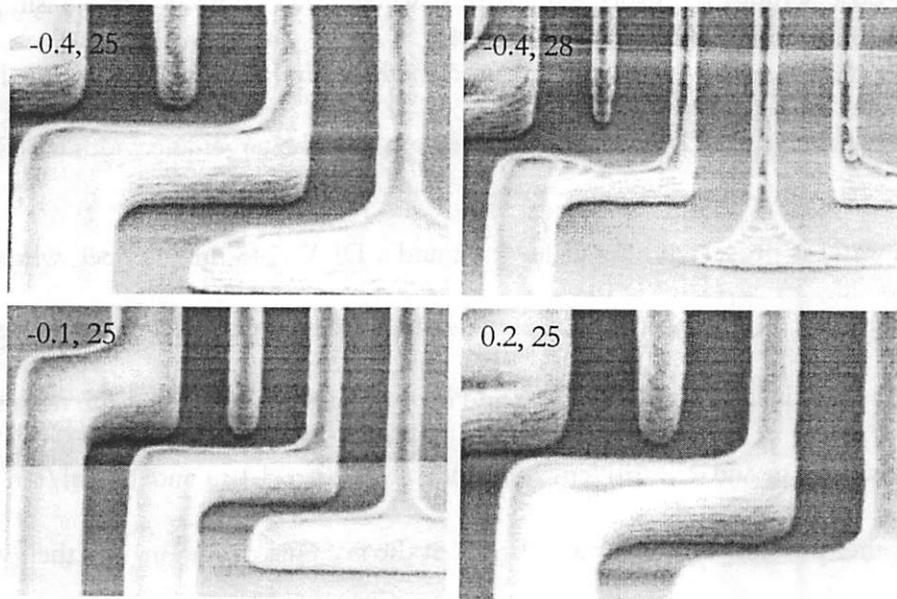


Figure 5-5 Tilt-angle SEM images of selected sites of the FEM wafer change has no obvious trend with respect to focus change from top to bottom of the wafer, especially for the high exposure dosage sites.

We then inspected the wafer under a tilt-angle SEM, and the pictures of a few sites are shown in Figure 5-5, with their focus and exposure settings labeled on the upper left corners. The pictures were taken at the edge of the grating region where there are also some turning and “T” shape lines for ease of inspection. For the three sites exposed with  $25 \text{ mJ/cm}^2$  with different focus settings, the profiles have fairly straight sidewalls, with more top rounding as focus changes from negative to positive. This is consistent with the profiles extracted with scatterometry. We also inspected a site with a high exposure dose ( $28 \text{ mJ/cm}^2$ ) and  $-0.4 \text{ }\mu\text{m}$  focus. Although this site was not measured with scatterometry, it follows the trend of profile change with respect to focus and dose in Figure 5-4.

Many of the tilt-angle SEM pictures show clear standing waves on the sidewalls of the resist lines. This is because no anti-reflective coating (ARC) was used between the resist and silicon

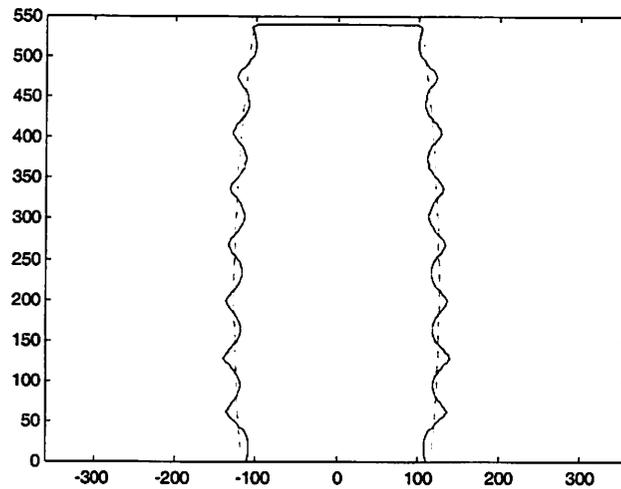


Figure 5-6 Profile with standing wave (solid line) and the approximated smoothed profile (dashed line)

wafer, and the process was not optimized for film stacks without ARC. When modeling profiles for scatterometry profile extraction, we did not take into account the effect of standing waves, because the reflected signals are not sensitive to them, as shown in the following simulation results.

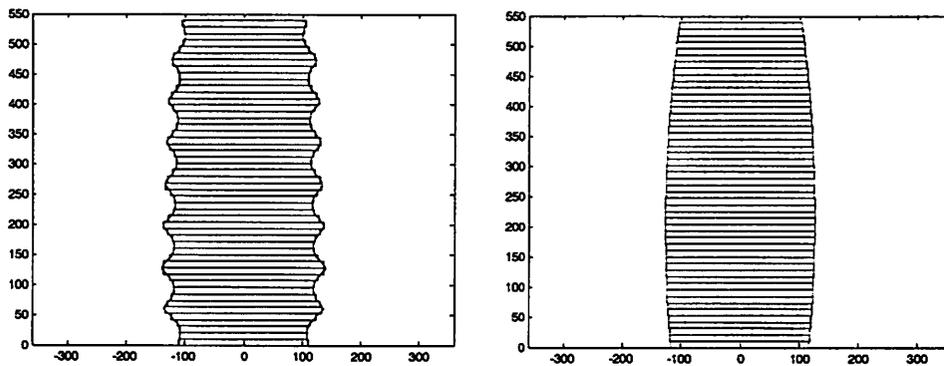


Figure 5-7 Rectangle slices for RCWA simulation. The left figure is for the profile with standing waves, and the right figure is for the approximated smoothed profile.

We first generate a resist line profile with strong standing waves using PROLITH [5.7], a lithography simulator, as shown by the solid line in Figure 5-6. The standing wave amplitude is about 15 to 20 nm on the resist line sidewall. In order to accurately simulate the signal

response from this profile using RCWA, we slice the profile into 50 layers of rectangles, which can represent the profile well enough, as shown in the left graph in Figure 5-7. The simulated  $\tan\Psi$  and  $\cos\Delta$  signals are plotted in solid lines in Figure 5-8. We then approximate the waved profile sidewall with a second order polynomial curve, as illustrated with the dashed line in Figure 5-6, and also slice the approximated profile into 50 layers for RCWA simulation, as shown in the right graph in Figure 5-7. The simulated  $\tan\Psi$  and  $\cos\Delta$  signals are plotted with dashed lines in Figure 5-8.

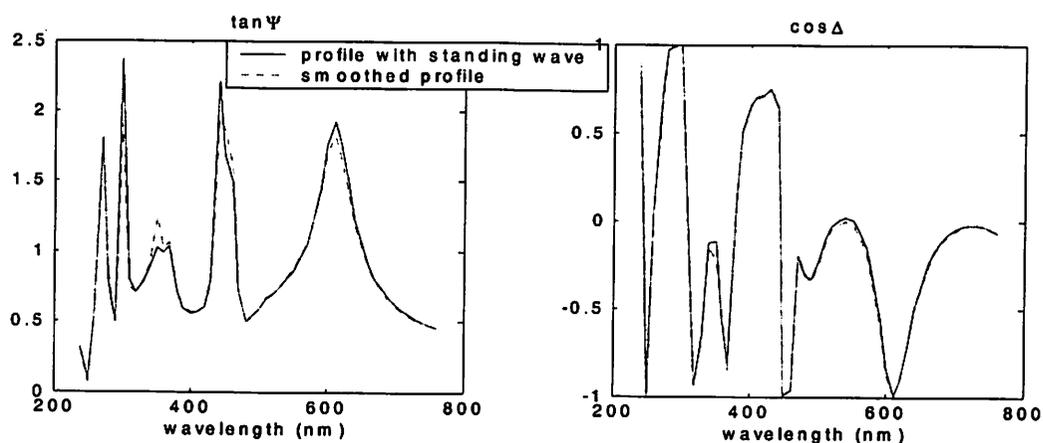


Figure 5-8 Signal responses for the profile with standing waves and the approximated smoothed profile.

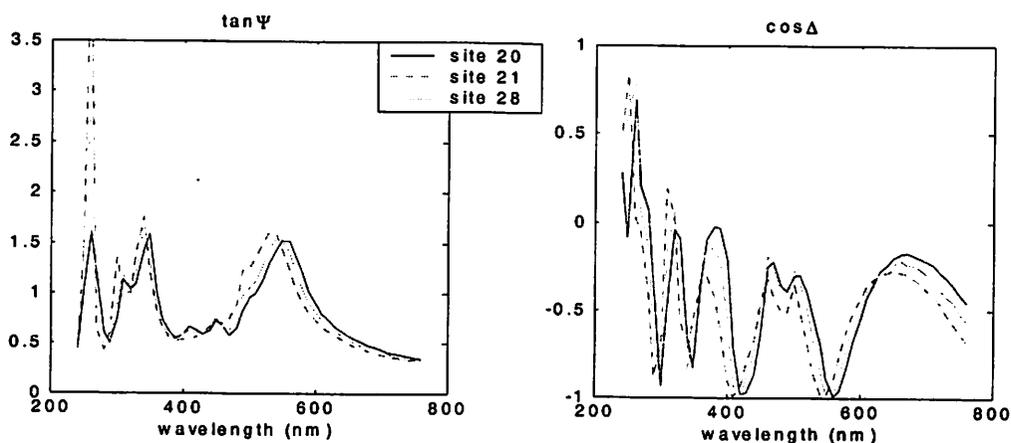


Figure 5-9 Measured signals from neighboring dice on the FEM wafer

Compared to the measured signals from neighboring dice (dice 20, 21, and 28), as shown in Figure 5-9, it is obvious that standing waves have a small effect on the magnitude of some of the peaks in the scattered signal, while line width and sidewall angle changes due to focus and exposure affect both the location and magnitude of the peaks. Therefore, the effect of standing waves can be negligible. If it is necessary to model the standing wave effect, the average material method [5.19] can be used to improve the simulation accuracy.

It is typical to use Bossung curves to determine the optimal focus and exposure settings. The Bossung curve generated with a middle-height CD measurement (obtained via scatterometry) is plotted in Figure 5-10. The target CD of this process is 200 nm, so the resist lines printed with the dose of 22 mJ/cm<sup>2</sup> is closest to this target. Also considering that this dose setting has the largest depth of focus, and the center of the curve of resist line width response to focus is approximately at 0.0 μm, we can conclude that 0.0 μm and 22 mJ/cm<sup>2</sup> are the optimal focus and exposure settings.

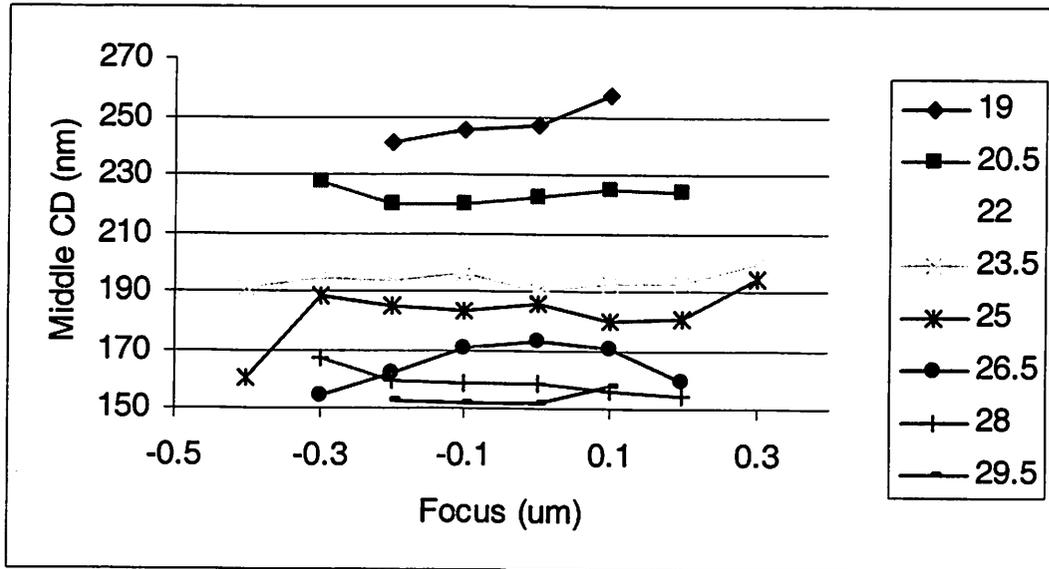


Figure 5-10 Bossung curves for the FEM wafer using scatterometry middle CD measurement

We also measured the full wafer (with a few dies missing due to pattern recognition failure) with CD-SEM, and we plot the Bossung curve in Figure 5-11. The curves have a similar trend to those plotted with the scatterometry measurements, but with approximately 25 nm offset. The cause of the offset has been discussed in Chapter 4. If this offset is properly calibrated, and the measurement for the site of 0.0  $\mu\text{m}$  focus and 22  $\text{mJ}/\text{cm}^2$  dose is considered an outlier, it is possible to determine similar optimal focus and exposure settings as scatterometry with the same method described in the last paragraph.

Since the grating structure used in this is fairly large, no pattern recognition is needed. To measure all the 48 sites on the FEM wafer, it takes about 7 minutes on a Therna-Wave OptiProbe 5240 ellipsometer. If pattern recognition is needed for smaller testing regions, the throughput can be slower, but it can be improved with later ellipsometer models. On an OPAL CD-SEM, it takes about 9 minutes to measure 48 dice. Considering its high throughput, lack of charging effect on resist wafers, and scalability down to a sub-100 nm technology node [5.8], scatterometry shows a clear advantage over CD-SEM on FEM characterization.

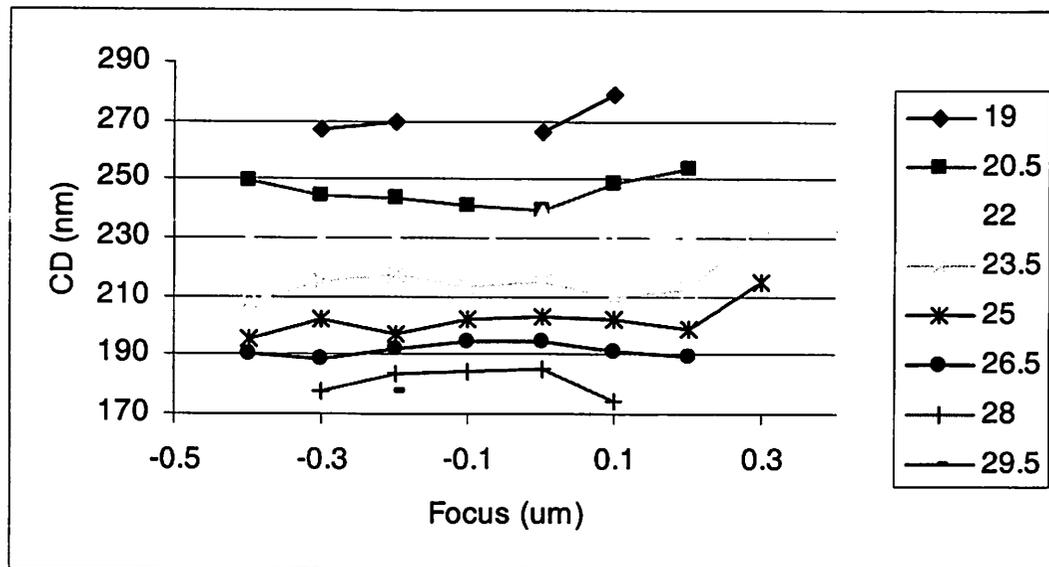


Figure 5-11 Bossung curves generated using CD-SEM measurement

## 5.4 A simulation framework for lithography process control

### 5.4.1 A lithography process control framework using scatterometry

An important advantage of scatterometry over CD-SEM is its capability to obtain full-profile information, because the optical response signal is sensitive not only to line width, but also to

other profile parameters, such as sidewall angle, top rounding, etc. The measured profile can then be used to monitor both exposure and focus drifts and to make adjustments accordingly.

A feedback control framework using scatterometry and a calibrated lithography simulator is illustrated in Figure 5-12. Wafers are first patterned using control-parameter settings in the process recipe, and the grating<sup>2</sup> is measured using an ellipsometer. The profile is then extracted from a pre-built scatterometry library using the ellipsometer signal, and fed into a control-parameter extractor to get the equivalent control-parameter input for this wafer. Compared to the original control-parameter input, the equivalent control-parameter drift is compensated

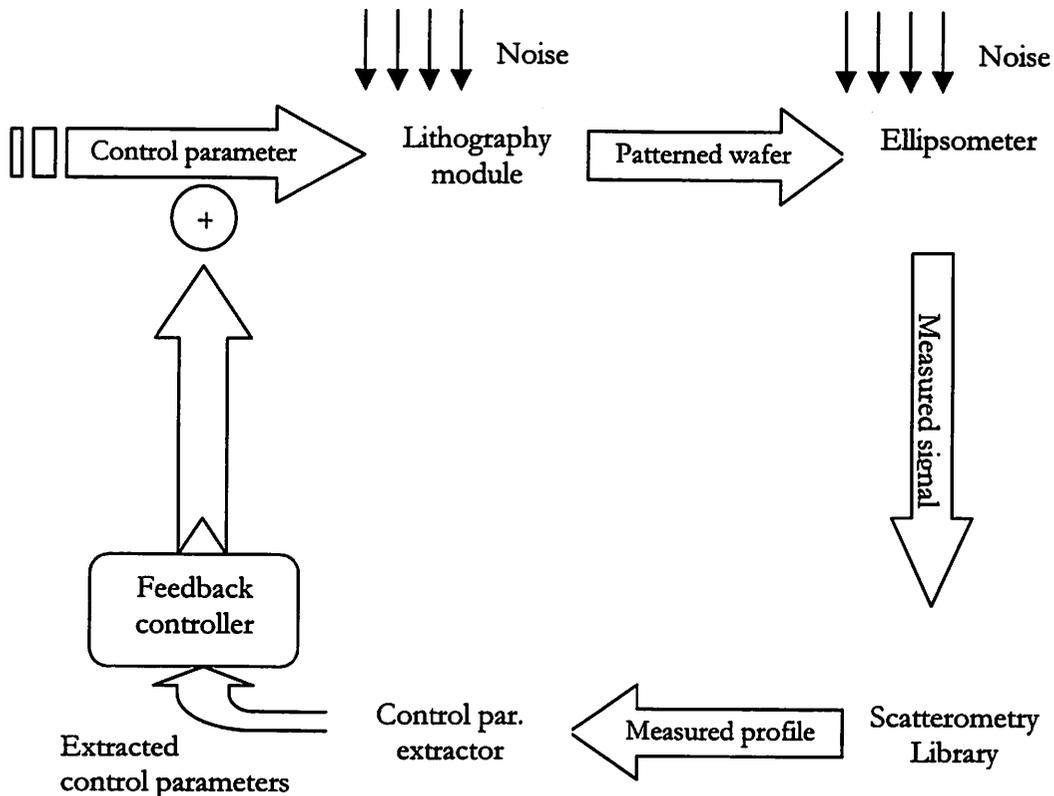


Figure 5-12 The control framework using scatterometry

<sup>2</sup> This grating can be either the actual device structure on some memory chips, or that printed on scribe line for the purpose of process monitoring.

using a feedback controller, and the next wafer is patterned using the adjusted parameters.

The control-parameter extractor in the above framework is a profile database used to get the equivalent control-parameter input from the measured profile. We first calibrate a lithography simulator using the procedure described in the next section, and then simulate resist profiles for a large combination of control-parameter settings with the calibrated lithography simulator. During the lithography process, after the resist line profile is measured with scatterometry, it is fed into the pre-simulated profile database. The control-parameter set that generates the profile in the database that best matches the measured profile is the extracted setting.

Ideally we want to include all the input parameters for the lithography module in the control-parameter set. However, considering the fact that focus and exposure dose are the most convenient knobs to adjust in the process, and most other parameters, such as resist and develop parameters, are confounded with them in the process window, only the combination of focus and dose are included in the control-parameter extractor. The drifts of other parameters can be considered as equivalent drifts of focus and dose.

#### **5.4.2 Lithography simulation and model parameter extraction**

One of the key components in the control framework presented in the last section is the control-parameter extractor, which is generated using a calibrated lithography simulator. We will review the development of lithography modeling and simulator calibration in this section.

Since the early work on photoresist modeling [5.9, 5.10], great progress has been made on lithography process simulation. There are three important elements required in order to simulate a process [5.1]: models, simulators, and calibration/validation. Models capture the behavior of a physical phenomenon by means of mathematical formulae; simulators

implement models with computer codes; and finally the simulated results need to be compared to relevant experimental data to determine numeric parameter values and to demonstrate suitability for purpose.

There are many academic (SAMPLE, SPLAT, TEMPEST, STORM [5.11]) and commercial (PROLITH [5.7], SOLID-C [5.12], etc.) technology computer-aided design (TCAD) tools that simulate the entire lithography process and use efficient numerical models. The value of a lithography simulator can be embodied in three aspects [5.13]. In the early stage of the technology development, simulation can help us understand the basic physical process, such as optics, resist reaction, and provide insight and guide innovation. During process development, simulation can help balance trade-offs of various process parameters, improve yield ramping rates and reduce development cost. After the process is developed and applied to production, a well-calibrated simulator can be used for process monitoring and control.

Parameter extraction and simulator calibration are essential simulation steps. There have been many works on lithography parameter extraction and simulator tuning [5.14-5.17]. Usually the simulator tuning process involves wafer processing and measurement after design of the experiment and matching experimental data to simulator output through the optimization process. The experimental data can be measurement on unpatterned wafers, such as film thickness and Fourier transform infrared (FTIR) responses, or on patterned wafers, such as line widths and profiles. After calibration, the simulator can simulate the fab process very well within the process window and can be combined with analysis and optimization tools for process development and control.

An example of lithography simulator tuning was presented in [5.16]. The parameters are categorized into three groups. The first group consists of parameters that can be obtained through direct, well-recognized methods. For example, thin film thickness and refractive index can be measured very accurately using ellipsometry. The second category consists of parameters that can be extracted by running an unpatterned wafer experiment. For example, the develop rate parameters can be measured using a develop rate monitor (DRM) experiment. Finally, a FEM wafer is processed, and the simulated resist profiles using PROLITH are matched with the measured ones by tuning the parameters in the second and third categories. The measured and fitted profiles for different focus and mask layouts are plotted in Figure 5-13. The average mismatch between the measured and fitted profiles is about 5 nm.

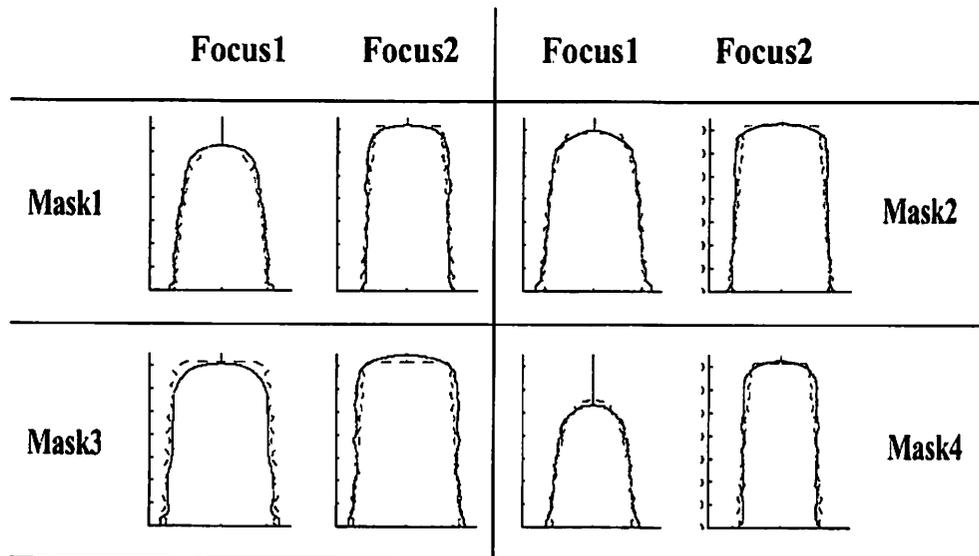


Figure 5-13 Example of simulator tuning using profile matching

### 5.4.3 A simulation framework for lithography process control

Before implementing the control framework, it is very important to do a simulation to test the effect of using the full-profile metrology to control the process drift. Since we do not do experiments on the lithography module, we use a PROLITH process simulator (we call it

“plant simulator”) to simulate the patterned profiles from the input control-parameter settings, and use a grating response simulator to simulate the spectrum measured by the ellipsometer. The “plant simulator” here is used to substitute the lithography module for the simulation framework, while the calibrated simulator in the control-parameter extractor explained in section 5.4.1 is used to match the lithography process. Considering errors in calibrating the simulator for control-parameter extraction, as presented in section 5.4.2, we set some resist parameters in the “plant simulator” different from those in the calibrated simulator, so that the output profiles from these two simulators have a mismatch of about 3 to 5 nm, which is close to the typical simulator tuning mismatch error.

Various disturbance models have been developed to simulate the process parameter drift. In this work, we use the first-order integrated moving average model, which is commonly used in

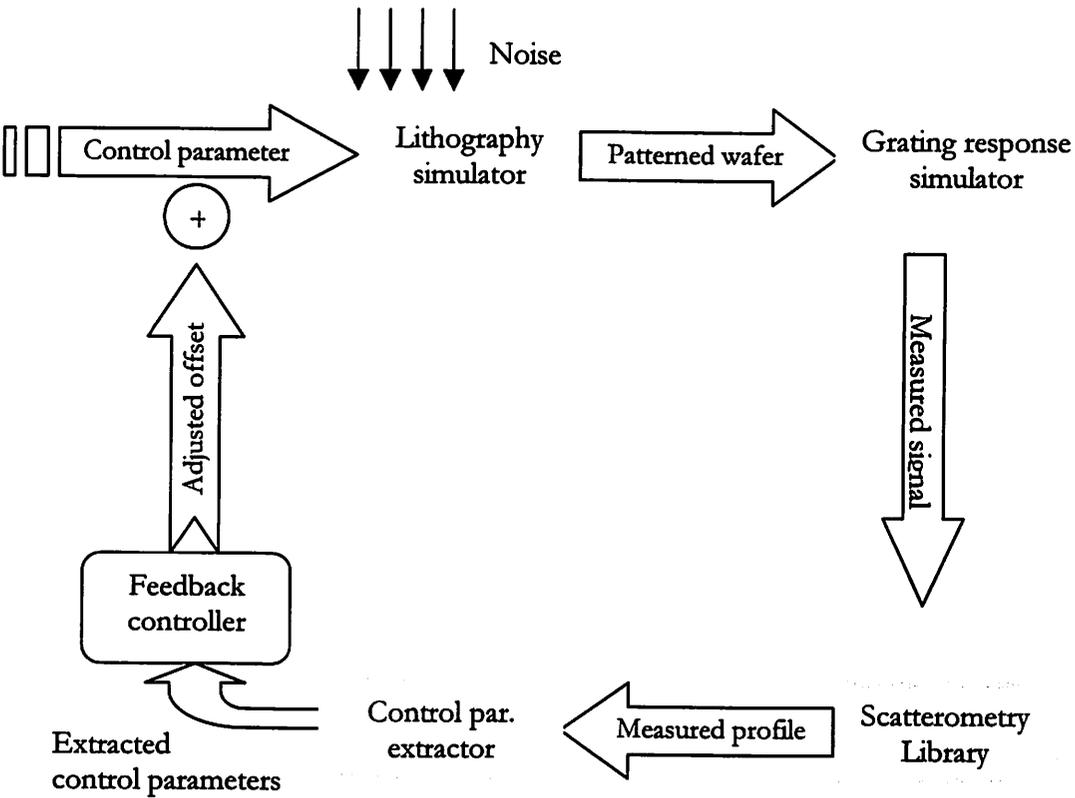


Figure 5-14 Simulation of the control framework using scatterometry

the statistics process control (SPC) environment. The disturbance  $d_j(k)$  for parameter  $j$  at step  $k$  can be modeled as:

$$d_j(k) - d_j(k-1) = a_j(k) - \theta_j a_j(k-1) \quad (5-1)$$

where  $a(\cdot)$  is an independent, uncorrelated stationary white noise sequence with mean zero and standard deviation  $\sigma_a$ , and  $\theta$  is an integrated moving average model parameter.

An exponentially weighted controller is used to adjust the  $j^{\text{th}}$  control-parameter setting  $p_j$  on the equipment at step  $k+1$  according to the extracted disturbance  $\tilde{d}_j$ :

$$p_j(k+1) = p_{j0} - (\lambda \tilde{d}_j(k) + (1-\lambda) \tilde{d}_j(k-1)) \quad (5-2)$$

where  $p_{j0}$  is the target recipe setting, and  $\lambda$  is the controller parameter. In equation (5-2),

$$\begin{aligned} \tilde{d}_j(k) &= d_j(k) + b_j(k) \\ \tilde{d}_j(k-1) &= d_j(k-1) + b_j(k-1) \end{aligned} \quad (5-3)$$

where  $b_j(k)$  and  $b_j(k-1)$  are the parameter extraction errors for the  $j^{\text{th}}$  control parameter in step  $k$  step  $k-1$ , respectively, and we assume they both have standard deviation  $\sigma_b$ .

Considering the disturbance  $d_j(k+1)$  to the  $j^{\text{th}}$  control parameter in step  $k+1$ , its real value would be

$$\hat{p}_j(k+1) = p_{j0} - (\lambda \tilde{d}_j(k) + (1-\lambda) \tilde{d}_j(k-1)) + d_j(k+1) \quad (5-4)$$

We can choose the controller parameter  $\lambda$  to minimize the variance of  $\hat{p}_j(k+1)$ . Substituting equations (5-1) through (5-3) in to equation (5-4), after some derivation, we have that the standard deviation of  $\hat{p}_j(k+1)$  is minimized when

$$\lambda = \frac{(1 + \theta^2 - \theta)\sigma_a^2 + \sigma_b^2}{(1 + \theta^2)\sigma_a^2 + 2\sigma_b^2} \quad (5-5)$$

#### 5.4.4 Simulation results

The structure used to monitor the process is 505 nm thick resist patterned on 61 nm thick ARC layer on silicon substrate, with pitch of 480 nm and 1:2 line-space ratio. There are about 500,000 profiles in the library. The control-parameter extractor was also generated offline, consisting of focus and exposure ranges of  $-0.3 \mu\text{m}$  to  $0.1 \mu\text{m}$  and  $22.0 \text{ mJ}/\text{cm}^2$  to  $27.0 \text{ mJ}/\text{cm}^2$ , respectively.

Table 5-2 Standard Deviation of the disturbed parameter noise

Focus ( $\mu\text{m}$ )	0.04
Exposure ( $\text{mJ}/\text{cm}^2$ )	0.5
PEB temperature ( $^\circ\text{C}$ )	0.2
Resist thickness (nm)	1.5
ARC thickness (nm)	0.5

We choose focus and exposure dose as the control parameters, and post-exposure bake temperature, resist and ARC thickness as disturbance noise parameters. The deviations of these parameters are listed in Table 5-2. The  $\theta$  parameter in the disturbance model (1) is 0.3 for all the parameters. Considering the fact that the control-parameter extraction error variance

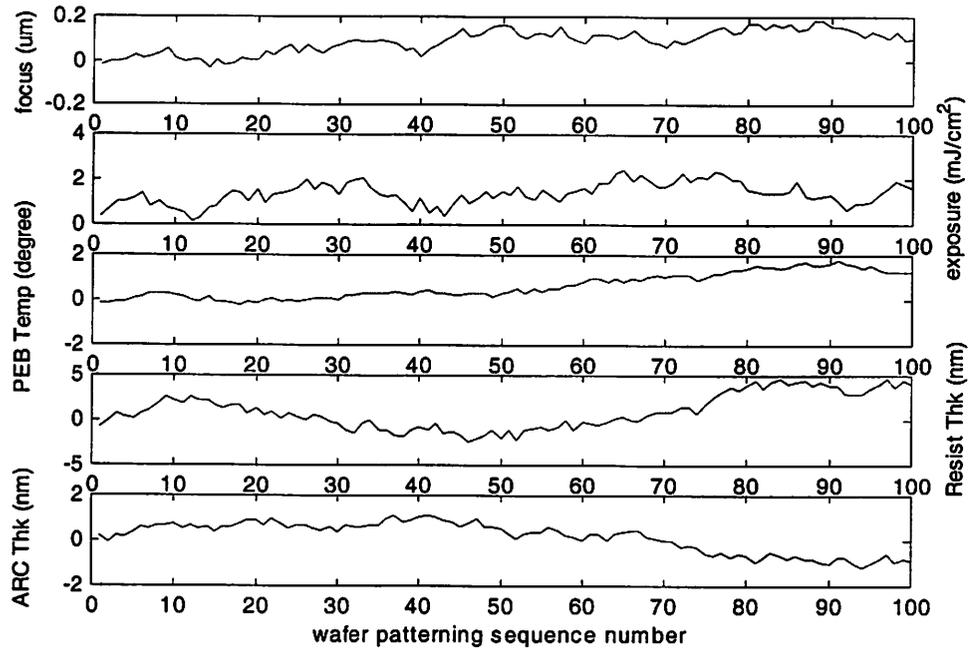


Figure 5-15 Parameter disturbance of the simulated lithography process

$\sigma_b^2$  should be less than that of the corresponding control-parameter disturbance  $\sigma_a^2$ , from equation (5-5), we can set the value of the controller parameter  $\lambda = 0.7$ .

After Prolith generates the profile, it is sliced into 50 layers and approximated using rectangles to simulate the response. This response (“measured” signal) is compared to those in the scatterometry library, and the profile that generates the best-fitted signal is the measured profile. The measured profile is then fed into the control-parameter extractor to find the best matched profile, which will give the equivalent control-parameter setting for this wafer.

The patterning process of 100 wafers is simulated using the control framework described in the last section. The disturbance of focus, dose, PEB temperature, resist thickness and ARC thickness are plotted in Figure 5-15. The extracted focus and exposure from the control-parameter extractor versus their corresponding input settings are plotted in Figure 5-16. The

difference between the extracted and the input settings is the noise in the monitoring loop, and the standard deviation is  $0.025\ \mu\text{m}$  for focus and  $0.23\ \text{mJ}/\text{cm}^2$  for exposure.

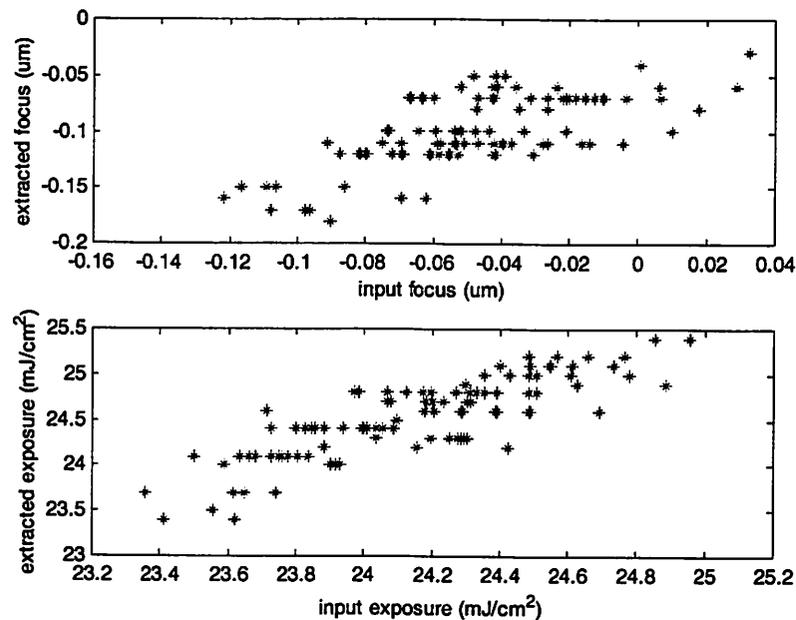


Figure 5-16 Extracted control-parameter values vs. their corresponding input values

Using the disturbance shown in Figure 5-15 and the control framework, the resulting bottom CD and sidewall angle have a much smaller variation comparing to the open-loop process, as shown in Figure 5-17.

As mentioned in the introduction section, it is also possible to do process control using CD SEM as the metrology tool by monitoring CD drift. In a production environment, usually an exposure meander (i.e., patterning scores of dice with a small increment of a dose that covers the process window for exposure with a fixed focus setting) is patterned with zero focus offset, and measured under CD-SEM. Then an empirical model relating CD to exposure can be developed from the CD-SEM measurement. In this study, we use the “plant simulator” to obtain the CD data for different dose settings, and fit to a linear model:

$$CD_{\text{bottom}} (\text{nm}) = 354.8 - 7.642 \times \text{dose} (\text{mJ} / \text{cm}^2) \quad (5-6)$$

Using the same disturbance model and controller parameter, the resulting bottom CD and sidewall angle are also plotted in Figure 5-17. We can see that although this method shows some improvement over the open-loop process, since it cannot monitor the focus drift, the output has a larger distribution than the scatterometry monitored method. The standard deviations of output variables using these three methods are listed in Table 5-3.

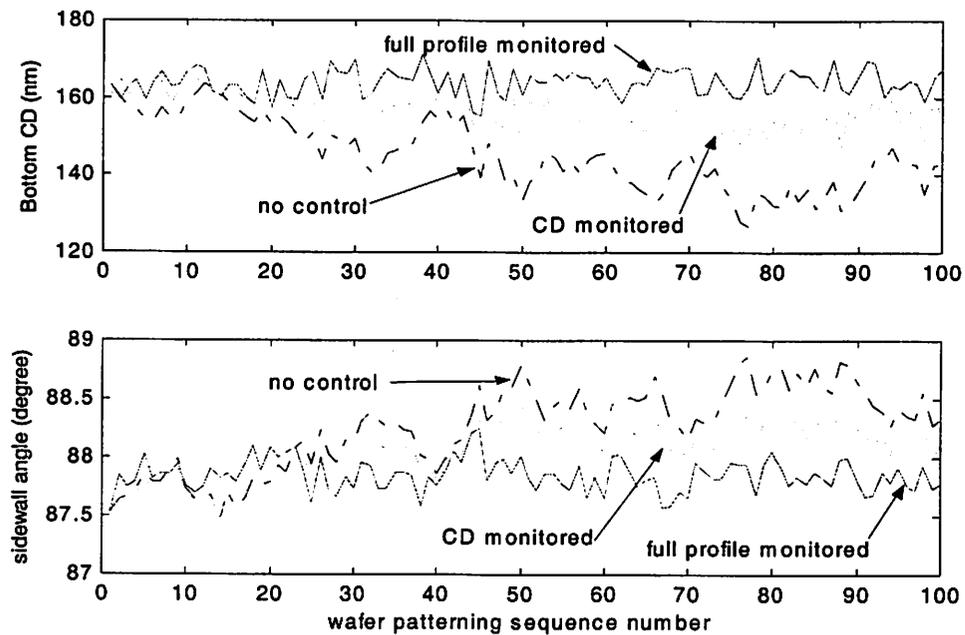


Figure 5-17 Comparison of resulting bottom CD and sidewall angle with and without control

Table 5-3 Comparison of standard deviations with and without control

	Bottom CD (nm)	Sidewall angle (°)
No control	8.9	0.35
CD only	4.5	0.23
Full profile	3.4	0.14

### 5.4.5 Discussion

There are many assumptions in the simulation study presented in the last section. In the simulation framework, we vary five parameters (focus, dose, PEB temperature, resist thickness, and ARC thickness) in the “plant simulator” to simulate process drift, and only focus and dose are adjusted to compensate the disturbance. In a real production environment, all the process parameters may drift. However, since different focus and dose combinations can usually generate most of the possible profiles if all the other parameters drift within the process window, it is sufficient to compensate the parameter drifts with focus and dose.

Another assumption in the simulator framework is that the developed resist profile information is available before the next wafer is exposed. On a modern wafer track system, wafers are queued in the stepper reception module, and it takes about 3 to 5 minutes after exposure to obtain the profile information using an integrated scatterometry system. Considering the throughput of 60 wafers per hour on a DUV stepper, it is equivalent to have about 3 to 5 wafers delay for metrology. Therefore, the benefit of process control may be less than that shown in the simulation.

In the simulation study, we use the first-order integrated moving average model to simulate the parameter drift. The model parameter  $\theta$  is 0.3, and the controller parameter  $\lambda$  is set at 0.7. In reality, it is difficult to obtain the accurate value of  $\theta$  and the ratio of the control-parameter extraction variance over model parameter drifting variance  $\frac{\sigma_b^2}{\sigma_a^2}$ . However, usually both  $\frac{\sigma_b^2}{\sigma_a^2}$  and  $\theta$  should be between 0 and 1, thus the optimal  $\lambda$  value given by equation (5-5) can only be between 0.5 and 1. Further derivation shows that the maximum parameter variance can only

be increased by 50% if non-optimal  $\lambda$  value is used due to inaccurate information of  $\frac{\sigma_b^2}{\sigma_a^2}$  and  $\theta$ .

One of the key advantages of the control framework using scatterometry is its capability to detect focus drift through profile change. In the study above, we use grating of 160 nm nominal CD and 480 nm pitch as the testing structure. Different structures have different sensitivity to focus change, so it is possible to optimize the feature design to make it very sensitive to focus drift. Since the profiles are measured using scatterometry, it is preferable to study the sensitivity of the scatterometry signals to focus drift.

We select four different line widths, 120 nm, 160 nm, 200 nm, and 240 nm, with the same pitch (480 nm) for this study. For each structure, we simulate the profiles with zero and  $-0.02 \mu\text{m}$  focus offset and  $22 \text{ mJ}/\text{cm}^2$  using the “plant simulator,” and then simulate the optical responses from these profiles. We use the sum-of-squares of the signal difference between the two focus settings as the metric for sensitivity:

$$SSE = \sum_{\lambda} [\log(\tan \Psi_m(\lambda)) - \log(\tan \Psi_s(\lambda))]^2 + [\cos \Delta_m(\lambda) - \cos \Delta_s(\lambda)]^2 \quad (5-6)$$

The simulation results for the four structures are listed in Table 5-4, and the signals for the 160 nm and 200 nm nominal line width gratings are plotted in Figure 5-18.

Table 5-4 Scatterometry signal sensitivity to focus drift

CD (nm)	120	160	200	240
SSE	8.16	1.10	0.10	0.02

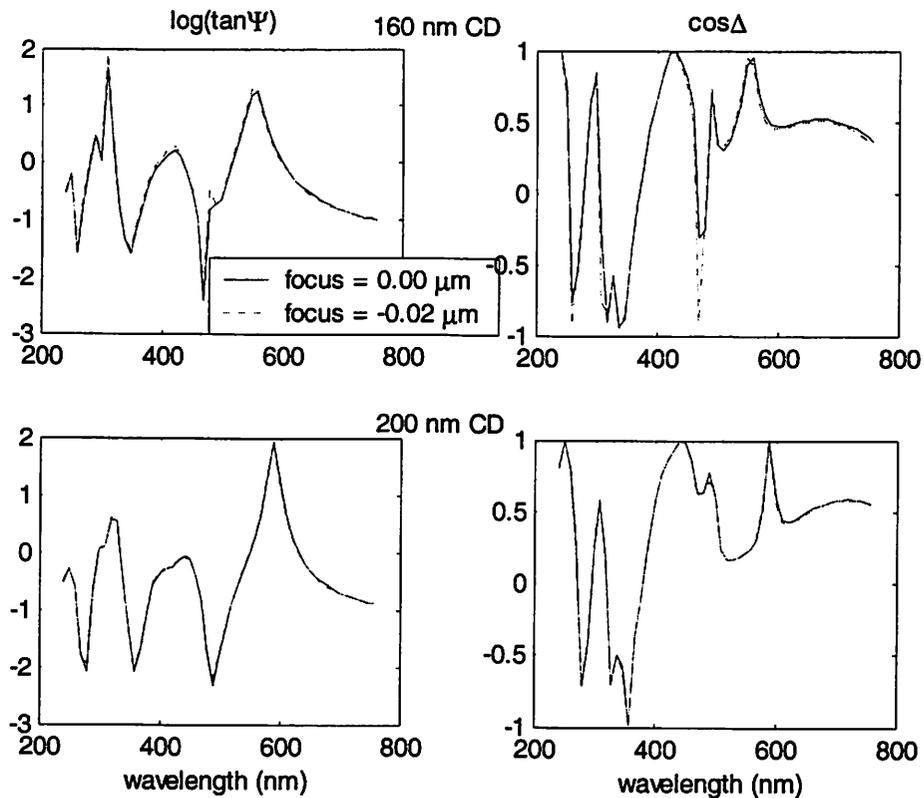


Figure 5-18 Signal sensitivity to focus drift for profiles with 160 nm CD and 200 nm CD

There is significant sensitivity difference for different structures. The smaller the nominal line width, the more sensitive it is to focus drift. On the other hand, the grating line width cannot be too small so that it can still be printed when other parameters drift within the process window. For the “plant simulator” we used for this study, the process window for grating structure of 120 nm line width is very small, so the 160 nm grating is the best of these four structures based on the trade-off between the process window and the sensitivity.

More advanced mask making technologies, such as optical proximity correction (OPC) and phase shifting mask (PSM), can be used together with line width and pitch optimization to further improve the sensitivity to focus and other parameters.

## 5.5 Conclusion

Many theoretical and experimental studies have demonstrated that scatterometry has clear advantage over current metrology tools for process monitoring and control due to its high accuracy, precision, scalability, and in-line compatibility. However, it is still a brand new area for applying the full-profile information obtained from scatterometry to process control. In this chapter, we first demonstrated some experimental work on FEM characterization using scatterometry, then presented a simulation framework on monitoring the process with scatterometry, and compensating parameter drift using a calibrated lithography simulator. Since both focus and dose drift can be monitored and corrected in the subsequent step, this method shows a smaller CD and sidewall angle variation than the traditional CD-SEM based control method.

## Reference

- [5.1] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors," <http://public.itrs.net/Files/2002Update/Home.pdf>, 2002.
- [5.2] E. Robertson, S. George, J.Holt, "Advanced CD Control Architecture and Implementation in Motorola," AEC/APC Symposium IX, Lake Tahoe, CA, September 20-24, 1997.
- [5.3] J. Stuber, F. Pagette, S. Tang, "Device Dependent Run-to-Run Control of Transistor Critical Dimension by Manipulating Photolithography Exposure Settings," AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [5.4] J. Allgair, D. Benoit, M. Drew, R. Hershey, L.C. Litt, P. Herrera, U. Whitney, M. Guevremont, A.Levy, S. Lakkapragada, "Implementation of Spectroscopic Critical Dimension

- (SCD\_TM) for Gate CD Control and Stepper Characterization,” SPIE vol. 4344, Metrology, Inspection and Process Control for Microlithography XII, 462-471, 2001.
- [5.5] N. Lafferty, C. Gould, M. Littau, C.J. Raymond, “Gauge Control for sub 170 nm DRAM Product Features,” SPIE vol. 4344, Metrology, Inspection and Process Control for Microlithography XII, 454-461, 2001.
- [5.6] Therma-Wave, Inc., 1250 Reliance Way, Fremont, California 94539.
- [5.7] KLA-Tencor Texas, Finle Division, 8834 N. Capital of Texas HWY, Suite 301, Austin, TX 78759.
- [5.8] W.L. Foong, “Characterizing the Sensitivity of Scatterometry for sub-100 nm Technologies,” MS Thesis, University of California, Berkeley, Spring 2001.
- [5.9] A.R. Neureuther, F.H. Dill, “Photoresist Modeling and Device Fabrication Applications,” Proceedings of the Symposium on Optical and Acoustical Micro-Electronics, (Proceedings of the Symposium on Optical and Acoustical Micro-Electronics, New York, NY, USA, 16-18 April 1974) 233-249, 1975.
- [5.10] F.H. Dill, A.R. Neureuther, J.A. Tuttle, E.J. Walker, “Modeling Projection Printing of Positive Photoresists,” IEEE Transactions on Electron Devices, vol.ED-22, (no.7), 456-641, July 1975.
- [5.11] <http://cuervo.eecs.berkeley.edu/Volcano/>
- [5.12] SIGMA-C Software, 901 Campisi Way #248, Campbell, CA 95008. <http://www.sigma-c.de>
- [5.13] A.R. Neureuther, T. Pistor, E. Croffie, K. Adam, B. Wu, M. Cheng, T. Horie, M. Brukman, M. Williamson, Y. Deng, L. Yuan, “Lithography Simulation is Becoming Essential,” UCB SAMPLE Group 2000, [http://www.eecs.berkeley.edu/~neureuth/SG\\_WEB\\_001.htm](http://www.eecs.berkeley.edu/~neureuth/SG_WEB_001.htm)
- [5.14] S. Thornton, C. Mack, “Lithography Model Tuning: Matching Simulation to Experiment,” SPIE Vol. 2726, 223-234, 1996
- [5.15] E. Croffie, “Simulation Tools for Optical Resist Models STORM,” PhD Thesis, University of California, Berkeley, 2001.

- [5.16] N. Jakatdar, "Deep Sub-Micron Photolithography Control through In-Line Metrology," PhD Thesis, University of California, Berkeley, Spring 2000.
- [5.17] J. Bao, "Deep Submicron Photoresist Modeling and Parameter Extraction," MS thesis, University of California, Berkeley, Spring 2000.
- [5.18] S.L. Prins, J.R. McNeil, S.S.H. Naqvi, J.W. Hosch, "Scatterometric Sensor for PEB Process Control," SPIE vol. 2725, Metrology, Inspection and Process Control for Microlithography X, 710-719, 1996.
- [5.19] Y. Deng, A.R. Neureuther, "EUV Phase-Shifting Masks and Aberration Monitors," SPIE vol. 4688, Emerging Lithographic Technologies VI, 495-502, 2002.

# Chapter 6 Requirements Of Scatterometry for Lithography Process Control

## 6.1 Introduction

In the previous chapters, we have demonstrated that scatterometry is a promising candidate for lithography process monitoring and control. However, many requirements need to be met before scatterometry is implemented as a key element for the control framework in a real production environment. The accuracy, precision, and throughput of scatterometry need to match or be better than other metrology tools, while it should also be inline compatible and be able to detect small features in the measured profile. Furthermore, it needs to be adaptive to many practical issues encountered in the semiconductor manufacture environment. In this chapter we first introduce the key elements of a run-to-run process control framework, and then discuss the requirements of scatterometry for the control framework. As an “indirect” metrology, it was unclear about the effect of systematic errors of the optical tools on the accuracy of system output. We will illustrate this effect through simulation in section 6.3.

## 6.2 Key elements of run-to-run process control framework

Run-to-run control is one of the most popular control methods for the semiconductor process [6.1-6.4]. During a process using run-to-run control, key process variables are kept constant at set points, and controlled variables are measured after each run. A model relating controlled variables to manipulated variables and a control algorithm are used to determine the necessary adjustment of manipulated variables to keep the controlled variables on target, and the adjustment is implemented in the next run.

There are four key elements of a run-to-run control framework: process model, metrology system, process model inversion method, and control algorithm. The process model describes the relation between controlled variables and manipulated and other input variables. For the SISO control strategy, usually a simple linear or polynomial model is good enough for lithography, and the model parameters can be extracted through a pre-designed experiment [6.5]. For the MIMO control strategy, it is essential to resort to lithography models derived from basic physics or chemistry processes, and solved numerically with computer simulation [6.6-6.8].

The metrology system measures controlled variables to determine the process drift. CD-SEM has been used as the standard tool for process control, but it can only monitor the wafer after the full lot is done in a real production environment, thus the unit for each run of run-to-run control is "lot." As illustrated in Figure 1-2, the earlier the metrology system can measure the wafer, the more effective the control is. Therefore, it is necessary to have in-line metrology sensors installed on the wafer track, so that the wafer can be measured right after the resist is developed. This can turn the unit of each run to "wafer." In order to implement MIMO control, it is necessary to measure the full profiles of the resist lines. Currently scatterometry is the only candidate that can satisfy both of these requirements.

After the controlled variables are measured by the metrology system, they are used to extract the equivalent drifts of the manipulated variables. This requires solving the problem of process model inversion. Either the regression [6.9] or library method [6.10] can be used for the MIMO process, but much work is needed to further improve the performance of this process.

The control algorithm determines how to adjust the manipulated variables for the next run to keep the controlled variables on target. One important assumption is that the process disturbance is auto-correlated. Without this assumption, there is no benefit to adjusting the process input based on the output of the last run. Fortunately, in the real world, autocorrelation does exist in the lithography process. Depending on the process drifting model, various controllers can be implemented to make the process stable and robust [6.2].

## **6.3 Requirement of scatterometry for lithography process**

### **control**

#### **6.3.1 Precision and accuracy**

The projected CD control range for the lithography process will soon be below 5 nm. This requires precision of less than 1 nm to meet the relaxed 20% measurement precision of the process tolerance metric [6.11]. In state-of-the-art ellipsometer systems used in production, the optical systems are optimized and have very low noise level. It has been demonstrated in many publications that the precision can be well below 1 nm. However, for reflectometer systems designed for integrated metrology, due to the limitation of the equipment size, signal noise is typically higher. Considering also that the reflectometer signal is usually less sensitive to critical dimensions, some development work is needed to meet the 1 nm precision specification.

Accuracy is another important metrology system specification for process control. For many metrology systems, there is relatively easy mapping between the measured feature and the signal. For example, the distance between the two scattering peaks in the CD-SEM signal directly relates to the measured CD. Therefore, the measurement inaccuracy appears as an offset that is close to constant, and can be characterized during tool calibration.

Scatterometry is an indirect metrology; it determines feature size based on signal sensitivity in a small range. When there is inaccuracy in the measured signal, it is impossible to tell whether the signal inaccuracy can be transformed to a fixed offset for different profiles.

Table 6-1 Nominal and  $1\sigma$  for the sample structure generation

	Top CD	Bottom CD	Resist thick	ARC thick	Poly thick
Nominal (nm)	65	75	200	45	100
$1\sigma$ (nm)	2.2	2.5	2.0	0.5	1.5

We use the structure in Figure 6-1 to simulate the effect of signal inaccuracy to measurement.

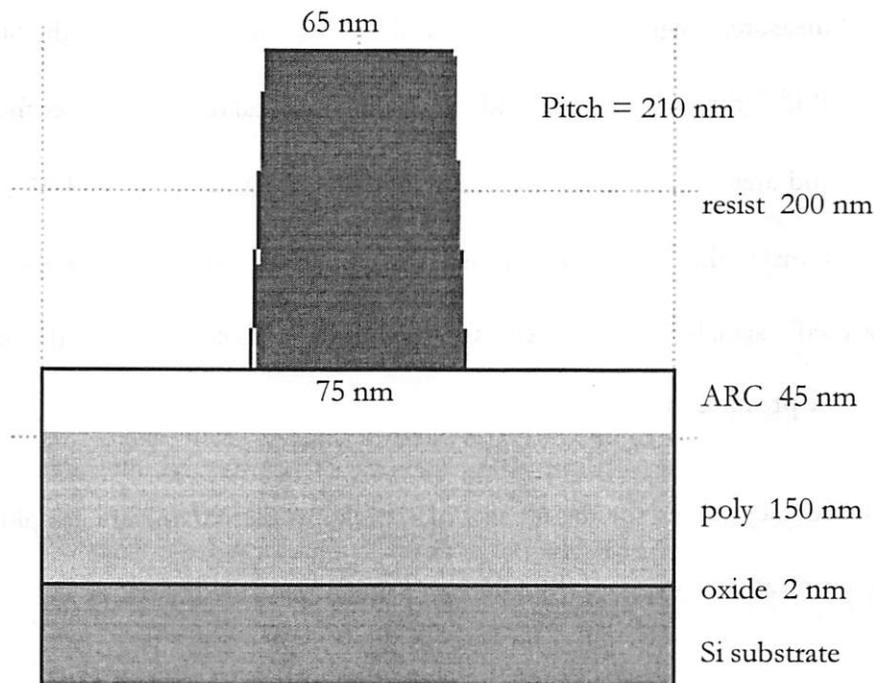


Figure 6-1 Structure used for accuracy study

The nominal structure parameter values, as well as the range of the process variations for these parameters, are listed in Table 6-1. We randomly generate 60 sample profiles based on the process specifications, and calculate the scattered signal response from a SOPRA ellipsometer

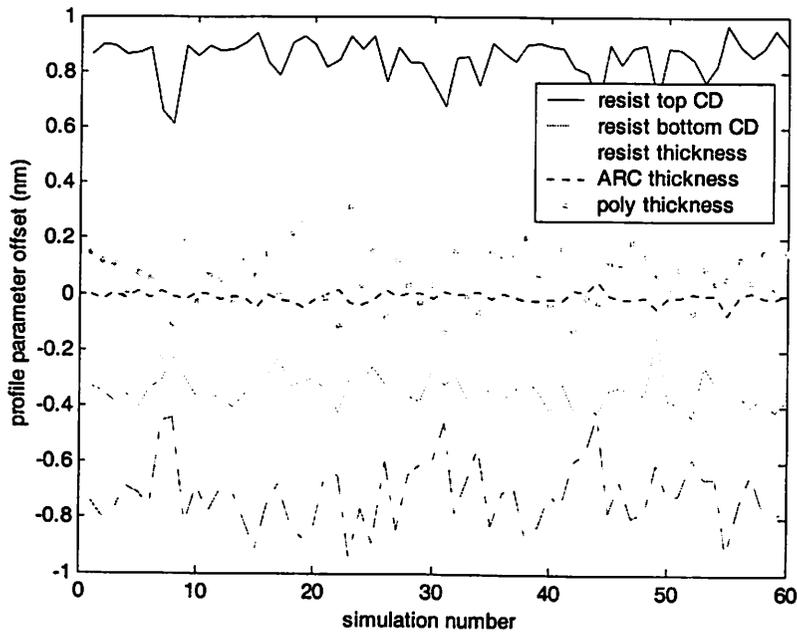


Figure 6-2 Parameter offsets for accuracy study

as the “measured” signal. We assume that there is some inaccuracy in the SOPRA ellipsometer system:  $0.1^\circ$  for incidence angle, and  $0.2^\circ$  for analyzer calibration. We set the incidence angle to be  $65^\circ$  and analyzer angle  $45^\circ$ , so the actual angles will be  $65.1^\circ$  and  $45.2^\circ$ , respectively, and this will make the “measured” signal inaccurate. We then extract the profiles using the “measured” signals with the regression method, and compare to the original profiles to determine profile extraction inaccuracy.

The inaccuracy values for the 60 sets of sample profile parameters are plotted in Figure 6-2, and their mean and standard deviation are listed in Table 6-2.

Table 6-2 Extracted profile parameter offset and  $1\sigma$

	Top CD	Bottom CD	Resist thick	ARC thick	Poly thick
Mean (nm)	0.85	-0.72	-0.34	-0.01	0.07
$1\sigma$ (nm)	0.07	0.12	0.05	0.02	0.11

The results above show that there are offsets of extracted profile parameters due to signal inaccuracy. The variations of the offsets for different profiles are comparable to the 0.1 nm convergence criterion. This means that the offset is approximately constant within the small process range. Therefore, like other traditional metrologies such as CD-SEM, this offset can be characterized during the calibration process, and can be corrected with scatterometry used in advanced process control. The effect of signal inaccuracy due to other factors can also be simulated similarly.

### **6.3.2 Sensitivity to fine profile features**

One of the expectations of scatterometry for process control is its capability to resolve very fine process features, such as top rounding, footing, etc., because these fine features can reflect the drift of some process parameters. As demonstrated in Chapter 4, the scattered signal is sensitive to these small feature changes. However, the signal change is approaching the equipment noise level, and very often the signal sensitivity to fine features is confounded with the sensitivity to other features, such as CDs and under-layer thin films. More research and development is needed to further improve the equipment performance and profile extraction algorithm to resolve the fine features.

### **6.3.3 Inline compatibility**

In order to maximize the benefit of process control in variation reduction, scatterometry sensors must be installed in the process module, such as wafer track. The allowed size for an integrated metrology tool is much smaller than a stand-alone tool. Usually the tool performance, such as wavelength range, signal to noise ratio, or optical configuration, is compromised to meet the requirement. Furthermore, an integrated metrology tool must be

able to follow the throughput of the process tool, which is usually faster than that requested for a stand-alone tool.

#### **6.3.4 Requirement of the semiconductor manufacture environment**

There are also some other requirements demanded by the semiconductor manufacture environment. Because scatterometry usually requires periodical structures, it is important to understand the correlation between the grating test structures and the circuit lines. The grating structures need to be optimized so that they can approximate the circuit line structure very well.

For practical reason, the refractive indices of the materials are fixed during scattering simulation for library generation or regression. However, it is inevitable that the material indices might drift from lot to lot, or even within wafer (for metal films). The scatterometry technology needs to be adaptive to variations of refractive index and other unaccounted factors.

There are also some other issues that will occur in a semiconductor manufacture environment, such as tool-to-tool matching, line edge roughness, profile asymmetry, etc. All these issues need to be solved before implementing scatterometry to production.

### **6.4 Summary**

In this chapter, we first introduced some basic concepts in process control. The tighter and tighter process range brings many new challenges to all of the four elements of process control. On the metrology aspect, scatterometry is a promising candidate to meet the new

requirement, but more research and development is needed to be able to implement it to the real production environment.

## Reference

- [6.1] T. Edgar, S.J. Qin, W.J. Campbell, "Run-to-run Control and Fault Detection," Short course, AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [6.2] J. Musacchio, "Run to Run Control in Semiconductor Manufacturing," MS report, UC Berkeley, Fall, 1998.
- [6.3] E. Sachs, A. Hu, A. Ingolfsson, "Run by Run Process Control: Combining SPC and Feedback Control," IEEE Trans. Semiconductor Manufact., vol. 8, no. 1, 26 -43, Feb. 1995.
- [6.4] D. Bonning, W. Moyne, T. Smith, "Run by Run Control of Chemical-Mechanical Polishing," 1995, IEEE CPMT Int'l Electronics Manufacturing Technology Symposium, 1995.
- [6.5] J. Stuber, F. Pagette, S. Tang, "Device Dependent Run-to-Run Control of Transistor Critical Dimension by Manipulating Photolithography Exposure Settings," AEC/APC Symposium XII, Lake Tahoe, CA, September 23-28, 2000.
- [6.6] Berkeley TCAD group, <http://cuervo.eecs.berkeley.edu/Volcano/>
- [6.7] KLA-Tencor Texas, Finle Division, 8834 N. Capital of Texas HWY, Suite 301, Austin, TX 78759
- [6.8] SIGMA-C Software, 901 Campisi Way #248, Campbell, CA 95008. <http://www.sigma-c.de>
- [6.9] N. Jakatdar, "Deep Sub-micron Photolithography Control through Inline Metrology," Ph.D. dissertation, UC Berkeley, Spring 2000.
- [6.10] J. Bao, C. Spanos, "A Simulation Framework for Lithography Process Monitoring and Control Using Scatterometry," AEC/APC Symposium XIII, Banff, Canada, October 6-11, 2001.

[6.11] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors," <http://public.itrs.net/Files/2002Update/Home.pdf>, 2002.

# Chapter 7 Conclusions and Future Work

## 7.1 Thesis summary

Since its debut in the mid 1990s, scatterometry has proven to be a promising candidate as an *in-situ*, full-profile metrology tool [7.2, 7.3]. We introduce various optical configurations of metrology tools that have been used for scatterometry, and analyze a few theoretical problems that relate to applying these tools to scatterometry in chapter 3. In chapter 4, we show some experimental results on characterizing a scatterometry system used in a production environment. The accuracy and precision of the scatterometry system are demonstrated on various applications such as polysilicon gate DI, polysilicon gate FI, shallow trench isolation FI, metal layer DI and FI. The results show the clear advantage of scatterometry over CD-SEM for production monitoring and control.

It is still a big challenge to integrate scatterometry into the advanced process control (APC) framework. Research and development on using the wafer level and full-profile information of scatterometry is still in the early stages. In chapter 5, we present a simulation framework on lithography process monitoring and control using scatterometry. The simulation results show that using CD and sidewall angle information obtained from scatterometry as controlled variables, the CD and sidewall angle variations are much smaller than those with the traditional CD-based control method. The requirement and challenges in applying scatterometry in a real production environment are discussed in Chapter 6.

## 7.2 Future work

Scatterometry is a typical inversion problem solving process. It is possible to simulate the forward problem accurately, but there is no direct solution to the inverse problem, i.e., given grating structure parameters, it is easy to simulate its optical response, but one cannot solve the grating structure parameters directly from its optical response. The regression method and library-based method have been used for scatterometry, but a lot of research can be done on combining these methods with other techniques such as neural networks to further improve the speed, accuracy and robustness.

Currently only one-dimensional gratings with symmetric profiles have been frequently demonstrated in scatterometry. However, there are many other circuit structures that need to be measured during the semiconductor process, such as two-dimensional contact hole arrays, overlay, profiles with line edge roughness, etc. Modeling and studying the sensitivity of the optical response from these structures can be an interesting research topic.

Most of the optical tools used for scatterometry today were originally designed for thin film metrology. For example, the incident angle of  $65^\circ$  to  $70^\circ$  for an ellipsometer is optimized based on the Brewster angle of the thin film stack on the silicon wafer. At this angle range the system is the most sensitive to the film thickness change. When the system is used for scatterometry on patterned structures, many of the instrument design parameters can be optimized to improve the system performance.

Another interesting area of future research is integrating scatterometry to the advanced process control framework for both lithography and other process steps, such as plasma etching, CMP, etc. The request for utilizing the wafer-level full-profile measurement data obtained from

scatterometry for process control requires more research on the improvement of the current APC framework. Since scatterometry still requires gratings as testing structures, it is also important to study how to optimize the design of the testing gratings so that they can be used to monitor the profiles of the actual circuit lines, or to detect the process condition drift.

## Reference

- [7.1] Semiconductor Industry Association, "International Technology Roadmap for Semiconductors," <http://public.itrs.net/Files/2002Update/Home.pdf>, 2002.
- [7.2] X. Niu, N. Jakatdar, J. Bao, C. J. Spanos, and S. Yedur, "Specular Spectroscopic Scatterometry in DUV Lithography," 159-168, SPIE vol. 3677, 1999.
- [7.3] C. Raymond, M. Murnane, S. Prins, S. S. H. Naqvi, et; al, "Multiparameter Grating Metrology Using Optical Scatterometry," J. Vac. Sci. Tech. B, vol. 15, no.2, 361-368, 1997.