

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.

**FIRST PRINCIPLE-BASED STATE
ESTIMATOR FOR PHOTOLITHOGRAPHY
CONTROL USING FULL PROFILE METROLOGY**

by

Paul D. Friedberg

Memorandum No. UCB/ERL M03/16

23 May 2003

**FIRST PRINCIPLE-BASED STATE
ESTIMATOR FOR PHOTOLITHOGRAPHY
CONTROL USING FULL PROFILE METROLOGY**

by

Paul D. Friedberg

Memorandum No. UCB/ERL M03/16

23 May 2003

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

First Principle-Based State Estimator for Photolithography

Control using Full Profile Metrology

M.S. Report

Paul D. Friedberg

Department of Electrical Engineering and Computer Sciences

University of California, Berkeley

May 23, 2003

**First Principle-Based State Estimator for Photolithography Control using
Full Profile Metrology**


by Paul D. Friedberg

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of Master of Science, Plan II.

Approval for the Report and Comprehensive Examination:

Committee:

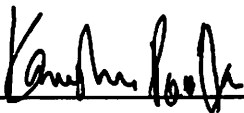
 C.J. SPANOS

Professor Costas J. Spanos

Research Advisor

5/23/2003

Date



Professor Kameshwar Poola

Second Reader

5/23/2002

Date

Abstract

In DUV photolithography, mask patterns and processes are increasing in complexity, while IC critical dimensions continue to shrink at a rapid pace. As a result, the proportional variability of the process will increase to unacceptable levels unless a means of more advanced process control is introduced. Previously standard offline pilot-lot experiments now prove to be too costly and difficult. One attractive and potentially highly viable alternative is simulation-based advanced process control [1, 2].

The proposed control framework exploits scatterometry, which provides in-situ, full-profile metrology [3]. The major obstacle to implementing scatterometry in a process control setting is profile inversion—deriving estimated input conditions from the measured profile. In this work, a first-principle-based process simulator (PROLITH [4]) is used to simulate the lithography process and create a library of profile-to-input-conditions pairs. These profiles are then used to generate simulated diffraction responses, resulting in a library of diffraction-response-to-input-conditions pairs. Finally, the empirically measured diffraction response will be matched to a simulated diffraction response in this library, whose accompanying set of input conditions should estimate the actual input conditions well.

Preliminary, simulation-only results suggest that the framework has the potential to be successful, particularly if approximate values of the input conditions are provided during the matching step. However, it is expected that the success of the framework in reality will hinge on how well PROLITH models the actual lithography process, the levels of measurement noise, and the method of constructing the simulated library. The current focus of effort in this research is determining how to build a library with balanced sensitivity across all input parameters, as well as diagnosing the empirical performance of the framework.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Thesis Organization	3
2	Overview of Lithography	4
2.1	Exposure System	4
2.1.1	Resolution Enhancement Techniques	9
2.2	Rest of the Lithography System	10
2.2.1	Photoresist	10
2.2.2	Bake Steps	11
2.2.3	Development	12
2.3	Metrology for Lithography	12
2.3.1	CD-SEM	13
2.3.2	Scatterometry	14
2.3.3	Metrology Tools for Scatterometry	17
3	Review of Process Control for Lithography	21
3.1	Statistical Process Control	21
3.1.1	The Control Chart	22
3.1.2	Shortcomings of SPC	24
3.2	Advanced Process Control	24
3.2.1	Run-to-Run Control	25
3.2.2	Predictive Process Control	28
3.3	Process Control Summary	29
4	Development of the Process Control Framework	30
4.1	Lithography Process Control Using Scatterometry	30
4.1.1	Potential Profile Inversion Solutions	31
4.2	Considerations for Library Development	34
4.2.1	Library-Building Inefficiencies in Standard Scatterometry	35
4.2.2	Proposed Library Building Methods	35
4.3	Potential Pitfalls	41
4.4	PROLITH	41
5	Experiment and Results	47
5.1	Experimental Setup	47
5.1.1	Evaluation Method	51
5.2	Results of the Small (2,420-entry) Seed Library	51
5.2.1	Estimation Results with No Additional Information Supplied	52

5.2.2	Estimation Results for 2,420-entry Library with Approximate PEB Temperature Supplied	53
5.2.3	Estimation Results for 2,420-entry Library with Approximate PEB Time Supplied	56
5.2.4	Estimation Results for 2,420-entry Library with Approximate PEB Temperature and Time Supplied	56
5.2.5	Summary of Evaluation for the Seed Library	56
5.3	Results of the Iteratively-Built Library	57
5.4	Results of the Alternate Large (27,783-entry) Library	59
5.5	Analysis and Conclusions	60
6	Conclusions	65
7	Acknowledgements	66
A	Sources of Circuit Performance Variability	67
A.1	Circuit Variability Study	67
B	Software Framework	70
B.1	Matlab—Core of the Software	70
B.2	PROLITH Simulation Engine	71
B.3	Profile Digitization	74
B.4	Scatterometry Library Simulation	75

List of Figures

1.1	ITRS roadmap requirements for gate CD control.	2
2.1	An illustration of the lithography process flow.	5
2.2	Illumination wavelength and $\frac{1}{2}$ -pitch vs. lithography node.	7
2.3	State-of-the-art ASML lithography system.	8
2.4	Example of optical proximity correction.	9
2.5	Resolution enhancement technologies.	10
2.6	CD-SEM shortcomings.	13
2.7	Geometry of the ellipsometric measurement for SSS.	15
2.8	Flow diagram for library-based SSS.	16
2.9	Variable-angle, single-wavelength reflectometer.	18
2.10	Ellipsometer illustration.	19
2.11	Normal-incidence reflectometry.	20
3.1	Control chart of photoresist thickness.	23
3.2	Block diagram for R2R feedback control model.	25
3.3	Generic illustration of PID control.	27
4.1	Illustration of the lithography process control flow.	31
4.2	Software flow for simulation of PROLITH-based scatterometry library.	33
4.3	PROLITH-based lithography process control flow.	34
4.4	SSD calculation illustration.	37
4.5	Illustration of the iterative library building method.	39
4.6	Illustration of the iterative library building method, second section.	40
4.7	Snapshot of the PROLITH user interface.	43
4.8	Sample aerial image result from PROLITH.	44
4.9	Sample image-intensity-in-resist result from PROLITH.	44
4.10	Sample result of latent image following PEB step from PROLITH.	45
4.11	Sample resist profile result from PROLITH.	45
5.1	Initial library sensitivities vs. processing parameter values.	49
5.2	Sensitivity balance evolution of iteratively-built library.	50
5.3	Parameter estimation for 2,420-entry library.	52
5.4	PEB temperature estimation error vs. exposure estimation error.	53
5.5	Parameter estimation for 2,420-entry library, with approximate PEB temperature provided.	54
5.6	Estimation error correlations.	55
5.7	Bossung curves.	55
5.8	Parameter estimation for 2,420-entry library, with approximate PEB temperature and time provided.	57

5.9	Exposure dose estimation for 2,420-entry library, with approximate PEB temperature and time and estimated focus value provided.	58
5.10	Parameter estimation for 27,783-entry library.	61
5.11	Parameter estimation for 27,783-entry library, with approximate PEB temperature and time provided.	62
5.12	Exposure dose estimation for 27,783-entry library, with approximate PEB temperature and time and estimated focus value provided.	63
A.1	Circuit propagation delay breakdown for successive technology nodes.	68
A.2	ANOVA breakdown of delay variation for successive technology nodes.	68
B.1	Flow-diagram of the software framework.	71

List of Tables

5.1	Experimental lithography process simulation setup: fixed nominal parameter values.	47
5.2	Experimental range and step size for variable processing parameters.	48
5.3	Initial library sensitivity for each processing parameter.	48
5.4	Experimental range and step size for variable processing parameters in 27,783-entry library.	50
5.5	Comparison of results from seed and iteratively-built libraries, with no additional measurement information supplied.	59
5.6	Comparison of results from seed and iteratively-built libraries, with approximate PEB temperature supplied.	59
5.7	Comparison of results from seed and iteratively-built libraries, with approximate PEB temperature and time supplied.	59
5.8	Comparison of best estimation results from seed libraries of different size.	60

Chapter 1

Introduction

1.1 Motivation

Since the introduction of the integrated circuit, the economic and technological driving force in the production of semiconductor-based microelectronics has been photolithography. Economically, the ability to batch-produce numerous copies of a circuit on a single silicon wafer has enabled the industry's exponential growth. Technologically, lithography has set the pace for performance improvement of integrated circuits by enabling the fabrication of increasingly smaller devices in successive technology generations at an exponential pace according to Moore's law [5]. Although Moore's law has repeatedly been proven true despite persistent doubt and worry on the part of those both inside and outside the industry, the present undercurrent of uncertainty about its future applicability as a predictive model is stronger than ever. Concerns about managing power (driven to higher levels by the increasing leakage of transistors) and increasing interconnect delays threaten to curtail the performance improvement promised by each new device generation, while increasing costs of production raise the question of profitability. From experience we know that difficulties such as these can be solved by novel design methodologies—but now there exists the potential for a more fundamental problem, one involving the role of lithography in manufacturing.

The semiconductor industry has commonly used the gate length (or gate critical dimension—CD) of the MOS transistor as a metric by which to judge its technological progress. In turn, minimum gate length is mainly determined by the resolution of lithography. In addition to the challenge of improving lithographic resolution to enable the reduction of gate length, the tolerances in gate length

must scale down in tandem with device size reduction in order to maintain acceptable reproducibility, presenting a related but separate challenge. Unfortunately, it has been shown that as device size shrinks, the variability of many critical device parameters, including gate length (as well as threshold voltage, oxide thickness, and others), increases in proportion to the nominal value of those parameters. The resulting variability in the quality of manufactured transistors (combined with the variability of intrinsic circuit operational parameters such as the supply voltage and temperature) has profound implications for circuit design and performance. Specifically, these variations engender a distribution of performance for a given circuit, as measured by maximum clocking frequency beyond which the circuit will fail. As variations increase, the performance distribution spreads, leading to an unacceptable reduction in chip yield, as fewer manufactured circuits will meet increasingly aggressive specifications. Particularly worrying for the field of lithography, it has been shown that variation in gate length is the most influential parameter in determining variability of circuit performance (see Appendix A). Critical dimension control is therefore a vital requirement, and one that is becoming increasingly difficult. In fact, according to the ITRS 2002 update, the semiconductor community has already fallen behind in the arena of CD control, as shown in Figure 1.1.

<i>Year of Production</i>	<i>2001</i>	<i>2002</i>	<i>2003</i>	<i>2004</i>	<i>2005</i>	<i>2006</i>	<i>2007</i>
<i>MPU gate in resist (nm)</i>	90	70	65	53	45	40	35
<i>MPU gate length after etch (nm)</i>	65	53	45	37	32	28	25
<i>Gate CD control (3 sigma) (nm)</i>	5.3	4.3	3.7	3	2.6	2.4	2




-  -- manufacturable solutions are known and are being optimized
-  -- manufacturable solutions are known
-  -- manufacturable solutions are NOT known

Figure 1.1: ITRS roadmap requirements for gate CD control. Note that beginning in 2003, there is not a known manufacturable solution for achieving the required control. (In fact, industry has settled for less stringent CD control of ~5-6 nm; therefore, CD control is one of the rare areas where industry lags rather than leads the roadmap.)

There are two distinct directions from which to attack this problem, and in order for the semiconductor industry's growth to be maintained, both must be taken into consideration. First, circuit designers must develop a methodology for creating robust circuits from unreliable components. This approach will not be addressed in this thesis. Second, the root cause of the problem, namely the variability of device characteristics, must be reduced during manufacturing. This thesis presents work in the effort to reduce variability in gate length through lithography process control. Specifically, a first principle-based state estimator for the lithography module will be proposed, designed

to be the principal component in an advanced process control (APC) framework.

1.2 Thesis Organization

The remainder of this thesis will be organized as follows. An overview of the lithographic pattern transfer process and associated metrology will be presented in Chapter 2. Previous work in advanced process control for lithography will be reviewed in Chapter 3. Next, a novel method for deriving process conditions from a measured profile (a critical component to an advanced process control framework) will be presented in Chapter 4. The results of simulations of this state estimation software tool will be presented in Chapter 5. Finally, Chapter 6 will present a summary of the work and give conclusions as well as immediate directions for future work.

Chapter 2

Overview of Lithography

Photolithography is a method for transferring a desired pattern into a thin film. Pattern transfer is achieved by using light exposure to change the chemical properties of a photosensitive material, called photoresist, which has been deposited on top of the device layer (thin film) to be patterned. By changing the chemical properties of the photoresist, it becomes either more resistant (in the case of positive resist) or less resistant (in the case of negative resist) to a particular solvent called a developer solution. With careful design, the exposed sections of a positive photoresist will have a pattern identical to that desired in the underlying device layer (or a negative image of the desired pattern in the case of negative resist), so that after the photoresist layer is placed in the developer, the pattern intended to be transferred into the device layer will remain in the form of photoresist. Finally, the underlying regions which are not protected by photoresist are etched away; when the remaining photoresist is stripped away, the layer has been patterned as desired.

2.1 Exposure System

The most critical, difficult, and expensive step in the general lithography flow (shown in Figure 2.1) is the exposure step. The exposure system consists of three main components. First, a source of light at a particular wavelength is needed. Second, a reticle or mask is used as a master copy of the desired pattern, through which the light is passed to transfer that pattern onto the photoresist-coated wafer. In early lithography systems, the mask was placed directly onto the wafer and illuminated by a lamp to expose the photoresist, in a process known as contact printing. However, this process was not suitable for batch production since the mask suffered contamination as a result of direct contact with the wafer. Thereafter, the mask was lifted just above the wafer during exposure, in a

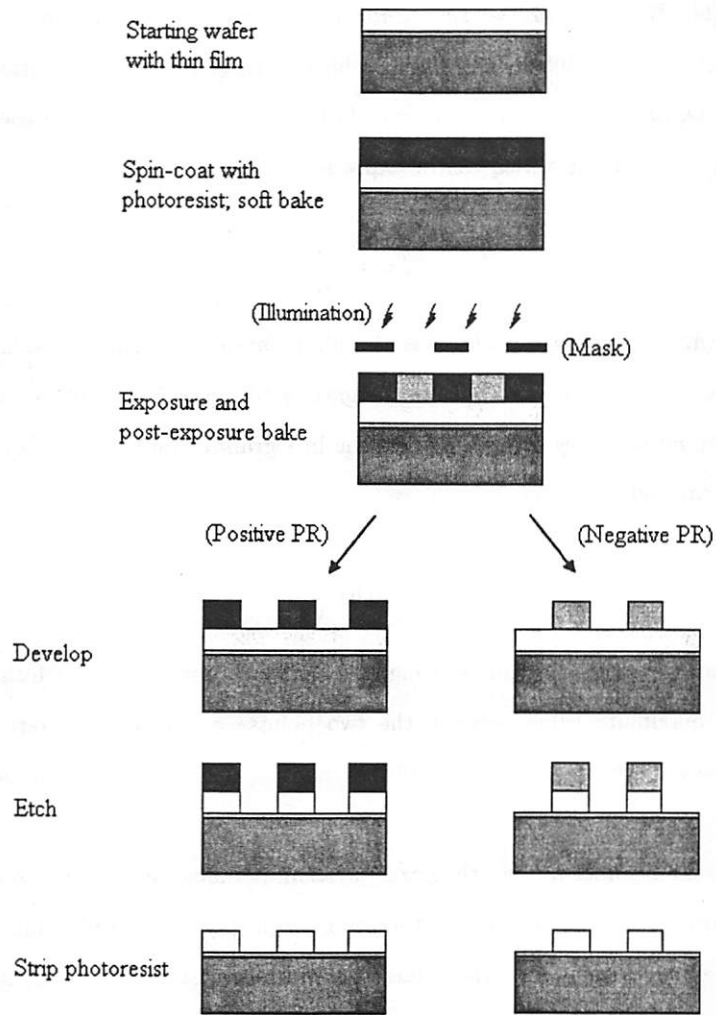


Figure 2.1: Illustration of the main steps in the lithography process flow.

method known as proximity printing. As the feature size decreased, however, it became impossible to accurately fabricate the mask, since proximity and contact printing transfer the mask feature size to the wafer in a one-to-one fashion. At this point, the industry shifted to the current standard of projection printing, wherein the mask pattern is demagnified as it is transferred to the wafer. To accomplish this image reduction, the third component of the exposure system, a lens consisting of several optical elements (called projection optics), is required to project a reduced image of the mask onto the target wafer.

Projection printing is the modern industry standard, and promises to be the standard for at least

a few more years [6]. However, the relative simplicity of contact/proximity imaging is sacrificed. Projection optics give rise to diffraction effects, which in turn limit the resolution, or minimum printable feature size, of the lithography system. Indeed, the principal metric used to describe a lithography system is resolution, which can be expressed by the equation:

$$w_{min} = k_1 \frac{\lambda}{NA}, \quad (2.1)$$

where w_{min} is the minimum feature size, λ is the illumination wavelength, NA is the numerical aperture of the final lens element, and k_1 is a lithography “difficulty” parameter, which effectively captures information about every other aspect of the lithography system. This equation is similar to the well-known Rayleigh criterion:

$$w_{min} = 0.61 \frac{\lambda}{NA}, \quad (2.2)$$

where w_{min} corresponds to the minimum spacing between two points of light such that the intensity falls to 80% of its maximum value between the two points—a subjective criterion of minimum discernability. As we will see, however, in photolithography k_1 may be pushed well below Rayleigh’s value of 0.61.

In order to improve the resolution of a lithography system, perhaps the most obvious solution is to reduce the illumination wavelength. Indeed, the industry has shifted the illumination wavelength from the mercury e-line (577 nm) to the g-line (436 nm) and i-line (365 nm), and on to deep ultraviolet (DUV) wavelengths 248 nm (KrF excimer laser illumination), 193 nm (ArF excimer laser), and finally 157 nm (F₂ laser). The lithography system of the future will most likely use extreme ultraviolet (EUV) light, at a wavelength of ~ 13 nm. However, the improvement in lithographic resolution has greatly out-stripped the reduction of illumination wavelength. In fact, at the 248-nm illumination wavelength, lithographers began printing feature sizes that were smaller than the illuminating wavelength, as shown in Figure 2.2.

This improvement has been possible due to the other two parameters in Equation 2.1, namely NA and k_1 . The numerical aperture is given by

$$NA = n \sin(\alpha), \quad (2.3)$$

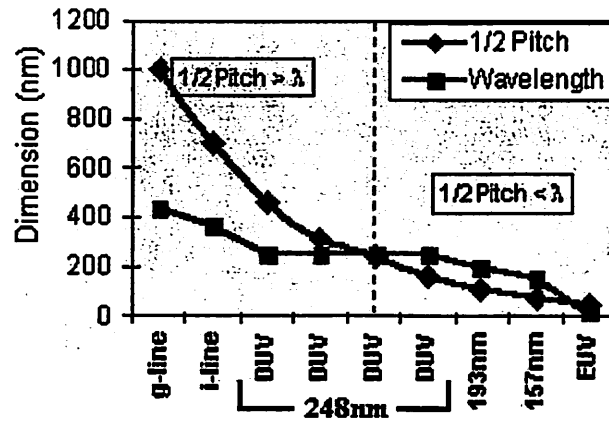


Figure 2.2: Illumination wavelength and $\frac{1}{2}$ -pitch vs. lithography node. From the Intel Lithography Roadmap [7].

where α is one-half the angle of acceptance of the objective lens (the final lens element, which projects the mask image onto the wafer) and n is the refractive index of the medium separating the objective and the wafer. With the advent of projection printing, it became impossible to expose the entire wafer at one time. Now, only small partitions of the wafer, called fields, are exposed at one time; each partition is exposed in step-wise fashion, giving rise to the common name “stepper” for a projection lithography system. The step-wise exposure of single, small portions of the wafer allows systems to be built with very high NAs , since the acceptance angle is greatly increased. Modern systems commonly have NAs over 0.7, compared to an NA around 0.25 in the 1970s, allowing for a three-fold improvement in resolution at a given wavelength. In order to achieve such high NA , modern lithography systems contain massive lenses (see Figure 2.3).

As lens size increases, imperfections in the lens—or aberrations—become more difficult to avoid during lens manufacturing. To combat this difficulty, modern lithography systems limit the portion of the lens that is actually used during imaging. This is accomplished by using a slit centered over the lens that allows only a portion of the mask pattern to be imaged onto the wafer at one time. To print the entire field, the mask is then translated in one direction while the wafer is translated in the opposite direction in a scanning fashion.

However, there are limits to the potential continued increase in NA . First, when imaging in air (as all current lithography systems do), NA cannot exceed $n_{air} \cdot \sin(\alpha) \approx 1$ since the index of refraction for air is approximately 1 and the $\sin(\alpha)$ term cannot exceed 1. Research is underway to investigate the possibility of increasing NA to ~ 1.25 through the use of a medium of higher refractive index

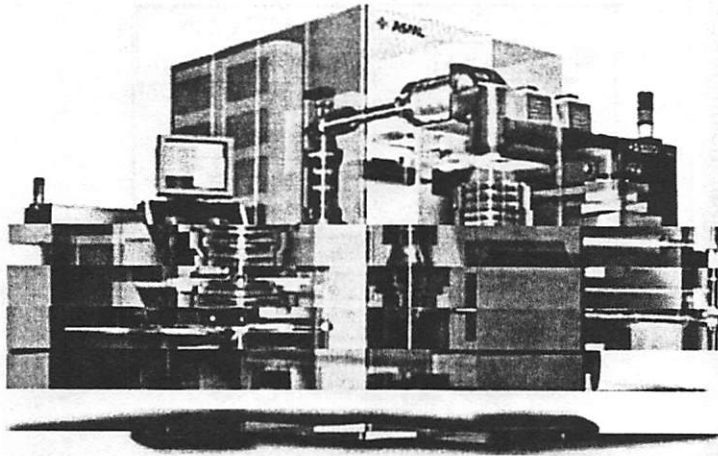


Figure 2.3: The ASML PAS 5500/1100B 193 nm Step and Scan system. With NA of 0.75, this tool is capable of sub-100 nm resolution. From [8].

(nominally a perfluoropolyether liquid) between the lens and wafer [9]. However, liquid immersion is unlikely to be implemented in the near future, with significant obstacles to overcome including turbulent flow issues with the liquid medium and contamination of the wafer. In addition, the depth of focus—defined to be the range over which the wafer can be moved along the optical axis such that the image stays in focus—is inversely proportional to the square of NA :

$$DOF = k_2 \frac{\lambda}{NA^2}, \quad (2.4)$$

If NA is increased too drastically, lithography will fail for even modest levels of topography on the wafer, since some features will fall outside of the usable focal range. Therefore, increasing NA has both physical limitations and additional processing constraints—fortunately, NA increase is not expected to be the only vehicle for further improvement of lithographic resolution. In fact, the majority of the achievement in resolution improvement has been accomplished by decreasing the lithography “difficulty” parameter, k_1 . For modern lithography systems, k_1 is pushed well below Rayleigh’s 0.61 through the use of high-contrast photoresist and a host of resolution enhancement techniques.

2.1.1 Resolution Enhancement Techniques

As features grow smaller in proportion to the wavelength, diffraction effects become more noticeable. Features with sharp corners on the mask are printed with rounded corners. Extremely narrow features will not print if they are isolated (since they do not benefit from diffracted intensity-spillover from neighboring lines as they would if the lines were densely packed). To combat these effects, mask engineers employ optical proximity corrections (OPC). Sharp corners can be achieved by adding “serif” or “hammerhead” structures (see Figure 2.4) to the features on the mask, and the iso-dense bias can be similarly relieved by adding scattering bars (sub-printable features) near isolated features. When the mask pattern is printed, these additional features serve to create a sharper pattern, as shown in Fig. 2.4.

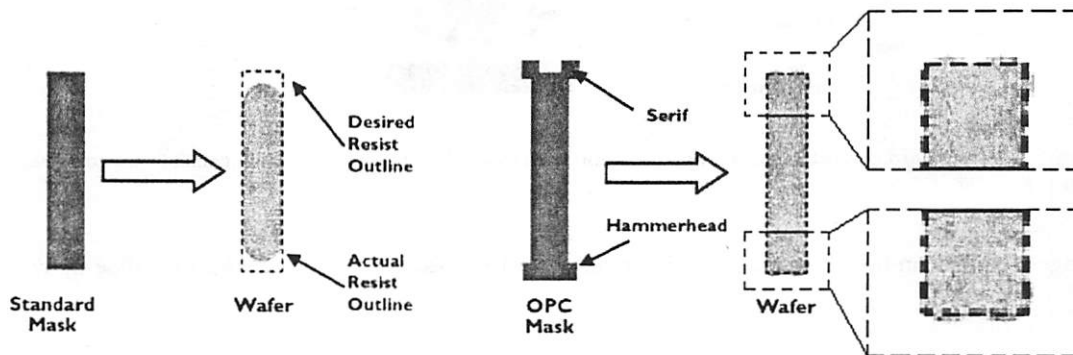


Figure 2.4: Standard optical proximity corrections. On the left, no OPC adjustments have been made. On the right, the serif and hammerhead features enhance the printing of the desired feature. From [10].

Other resolution enhancement techniques include phase-shifting masks, off-axis illumination, and multiple exposure. Virtually every component of the exposure system has been explored for sources of potential reduction in k_1 (see Figure 2.5).

Not surprisingly, there is a price to pay for employing these techniques to enable the printing of sub-wavelength features: increased variability in the lithography process. As additional RET are employed, the lithography process gains additional sensitivities. In addition, existing sensitivities may become stronger, as evidenced by the mask error factor (MEF), which captures the degree to which mask errors are magnified during the transfer to photoresist. As more sophisticated RET are incorporated, the MEF grows. Therefore, for a typical drift in a given lithography process parameter (such as exposure dose), the resulting resist profile will display a corresponding distortion in shape that will be magnified by the addition of more complex components. Put simply, more can go

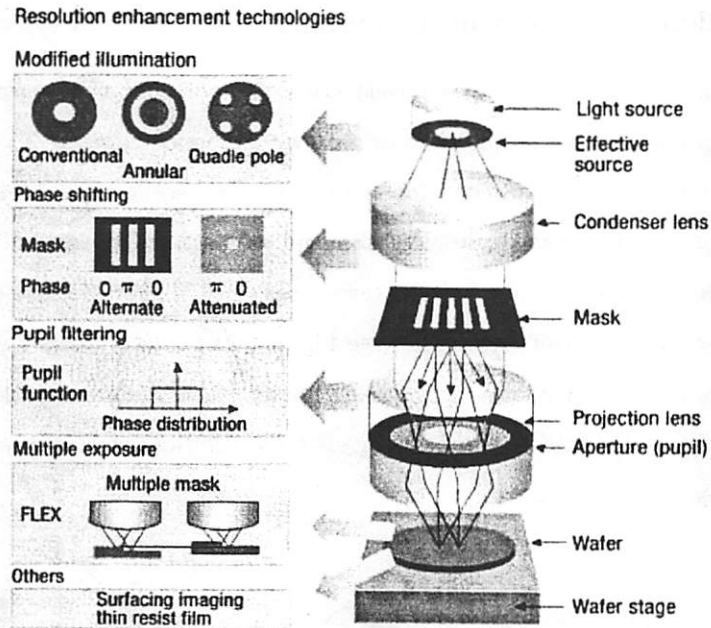


Figure 2.5: A standard projection lithography system, with examples of resolution enhancement technologies. From [11].

wrong—and therein lies the motivation for an advanced process control framework for lithography.

2.2 Rest of the Lithography System

To gain insight into the various aspects of the lithography module which may trigger this amplified variation in printed features, let us look briefly at each of the other steps in the lithography process in turn.

2.2.1 Photoresist

Photoresists have three main components: a resin (base material), a photoactive compound (PAC), and a solvent that controls the structural properties of the substance, particularly viscosity. In positive resists, the PAC behaves as an inhibitor, nominally preventing the resist from dissolving in developer solution. Exposure to light triggers a chemical reaction which changes the PAC into a sensitizer—a compound which enhances the dissolution rate in developer by breaking the resin structure material down. Resist contrast is determined by the sensitivity of the resist to changes in exposure dose. The narrower the range of exposure energies between zero resist removal and total

resist removal, the higher the resist contrast. In turn, the higher the resist contrast, the sharper the line edge in the printed feature. Resist resolution—the minimum feature size which may be printed in the resist—also depends on resist contrast, since the contrast of the aerial image (i.e., the light before it interacts with the resist) must exceed the contrast of the resist in order for the pattern to print [12].

Photoresist is applied to the wafer with a spin-coating process, which yields an exceptionally uniform film across the wafer. However, from wafer to wafer and lot to lot, variability in spin speed, resist viscosity, and adhesive properties between the resist and substrate (usually improved through the use of a specially adhesive material such as hexamethyldisilazane (HMDS) or an anti-reflective coating (ARC)) lead to film thickness variation. This variation will in turn lead to variation in the printed CD—the same amount of exposed intensity will yield wider features for a thicker resist layer.

2.2.2 Bake Steps

Below about 250 nm, traditional resists begin to strongly absorb incident light, leading to nonuniform exposures through the depth of the film. At the DUV (248 nm) lithography technology node, this absorptive property rendered traditional resists virtually obsolete, and new chemically amplified resists (CAR) were developed. Chemical amplification is achieved by an agent contained in the resist which greatly increases the potency of the photochemical process. The photoacid generator (PAG) of a CAR is designed so that an incident photon will create a cascade of chemical reactions that catalyze the scission of the resin [13]. The resist sensitivity is dramatically increased, but relies heavily on thermal activation, which is carried out in a post-exposure bake (PEB) step. The resulting latent image (prior to development) in the resist depends strongly on the thermal dose—both PEB temperature and duration are key factors. Therefore, PEB bake stations must be designed to place the wafer perfectly level on the heated surface so that the same thermal dose is received at all points on the wafer; also, the nominal temperature must be carefully maintained, particularly difficult since the bake plate encounters a cold, room-temperature wafer at the beginning of each bake cycle. At a nominal temperature of roughly 100°C, deviations of a single degree will lead to ~3 nm deviations in resulting feature width.

There are two other bake steps in a standard lithography flow: the post-application bake (PAB) and an optional hard bake. The PAB step serves to drive off some of the solvent in the film following spin-coat, serving to create a firmer film prior to exposure. The hard bake step follows development,

and is intended to drive off any remaining solvent in the remaining resist to strengthen the masking features prior to the etch process. Since neither step has a strong effect on the resulting pattern, they generally do not receive much attention from a process control standpoint.

2.2.3 Development

Development transforms the latent resist image (following the PEB step) into the actual relief image which serves as a mask for subsequent subtractive etch steps. As described previously, the development of the photoresist is a selective dissolution of exposed resist. Although the resist feature width is strongly dependent on the development process at the beginning of the develop step, this dependence quickly reduces to well below the sensitivity level of exposure, focus, resist thickness, and PEB characteristics. Therefore, lithography modules are designed such that the nominal development time is much longer than the time period over which the feature is sensitive to this step, ensuring that slight variations will not contribute much variability in the process. Thus, the development step is a secondary concern as a source of variability, and is commonly ignored in the process control setting.

2.3 Metrology for Lithography

In addition to an understanding of the sources of variability within the lithography process flow, we require a metrology which can accurately measure the resulting variability in printed resist features. In the interests of devising a process control framework, the ideal metrology must meet several criteria. First, the metrology should convey as much information as possible. Although critical dimension is the final measurement of interest, the entire resist profile contains more information about the conditions of its “parent” lithography process. Second, the metrology must be non-invasive; obviously, implementation of process control in a production environment cannot come at the price of a fixed overhead in yield loss for measurement purposes—the devices which are measured must still be saleable. Third, the metrology must be rapid, and should not present a significant financial burden in an already expensive process (state-of-the-art lithography systems cost upwards of ~15 million). This condition implies the requirement of in-situ measurement, by equipment with minimal footprint. In this section, two metrology options will be discussed. CD-SEM, which nominally provides only CD measurement, but which may be adapted to provide profile

data, and scatterometry, which provides full-profile measurement.

2.3.1 CD-SEM

CD-SEM (short for critical dimension scanning electron microscopy), uses a standard scanning electron microscope to take top-down images of resist features and extract linewidth (CD) measurements from those images. This metrology is relatively fast and easily automated, and has been the industry standard throughout the 1990's. However, CD-SEM has several drawbacks. One primary shortcoming of CD-SEM is the requirement that this measurement be made under vacuum, precluding in-situ measurement. The other primary argument against CD-SEM is that it provides only two-dimensional information. Since only the top surface of the resist line is imaged by the CD-SEM, there may exist a bias between the measurement and the actual electrical linewidth of interest. For example, in the case of measurement of MOS transistor gate length, the critical parameter is the length of the gate at the interface between the gate and the substrate (ignoring the gate oxide for purposes of simplicity). If the resist line has perfectly vertical sidewalls or sub-90° tapered sidewalls, then the CD-SEM should theoretically return the correct measurement of interest. However, if the profile has a re-entrant (inverted) trapezoidal shape—that is, a sidewall angle of greater than 90°—then the measured CD will exceed the actual electrical CD, giving an inaccurate measurement of the true gate length (see Fig. 2.6).

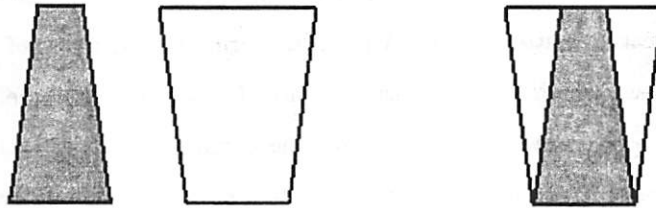


Figure 2.6: These two profiles have the same bottom CD, but the inverted profile (sidewall angle $> 90^\circ$) will have a larger measured CD when measured with CD-SEM, since this is a top-down measurement. In the case of the MOS transistor gate, this bottom CD is the critical parameter.

Recent research suggests that some profile information can be derived from a top-down view of the resist structure, based on the intensity of the image at several points along its transverse cross-section [14]. However, this technique is still immature, and does not present any obvious advantages over scatterometry. In addition, CD-SEM accuracy is suffering increasingly from issues of focus in the

top-down image. As linewidths shrink, the fixed value of uncertainty due to image focus limitations (arising because the location of line edge in a fuzzy image—due to variable charging effects in the CD-SEM—is uncertain) gives rise to a larger relative variability in the final measurement. Also, since the CD-SEM profile extraction method relies mostly on information obtained from the near-edge portions of the image, we might expect that accuracy of profile extraction to be limited as well.

Applied Materials has recently introduced the NanoSEM-3D [15], which allows the CD-SEM beam to be tilted up to an angle of 15°. By looking at a sample from several angles, stereographic vision is achieved and the profile shape may be roughly extracted. However, in order to provide quick measurement, the system assumes a trapezoidal feature shape—if the true, accurate profile is desired, offline modeling of the particular process in question is required. Thus, for the time being it appears that NanoSEM-3D is not ideal for production metrology, although it does provide a useful visual tool for research purposes.

2.3.2 Scatterometry

Profile information is highly desirable, and cannot be quickly and reliably measured using a CD-SEM. In the 90's, profile measurement was typically done by cross-section SEM, or by AFM (atomic force microscopy). However, both methods are destructive, and neither can be carried out in-line, let alone with speed. Therefore, scatterometry has emerged as the leading source of profile metrology. The principle of scatterometry is basic. A periodic grating test-structure of regularly spaced lines is created (separate from the device resist pattern). Light shines onto the grating and the resulting diffracted light is measured (Fig. 2.7). From this derived optical diffraction response, the grating profile is reconstructed. Thus, although the actual device parameters are not measured, the measurement of several test gratings of varying line-space ratios provides a sufficient snapshot of the typical resist feature and therefore detailed information about the state of the lithography process. A grating is used for a number of reasons. In the most practical sense, the measurement spot size (roughly 50 μm by 100 μm) greatly exceeds the minimum feature size (less than 1 μm); therefore, several features should be sampled in a single measurement. In addition, the cumulative effect of repeated structures improves the "signal content". Similarly, the use of the grating allows the extraction of an average profile shape, which precludes the possibility of the measurement of a line which happens to be an outlier in CD. Finally, the 1-D periodic structure simplifies the forward

calculation of the diffraction response, making the method more efficient. Since scatterometry does not involve charging of the sample as in the case of CD-SEM, or sacrifice of the sample as in the case of cross-section SEM, this metrology is said to be non-invasive. Scatterometry is also extremely efficient in terms of speed, holding a decisive advantage over the alternative metrology options.

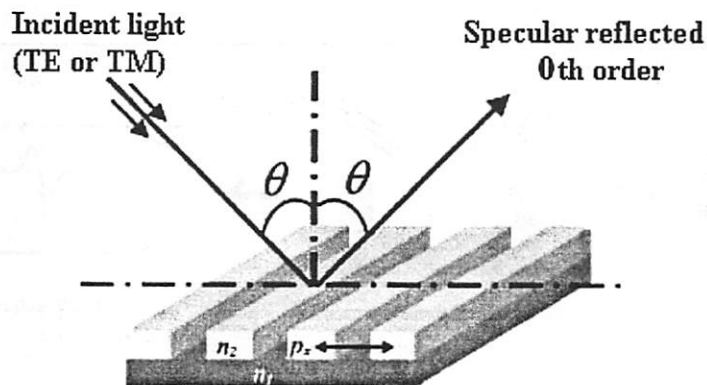


Figure 2.7: Geometry of the ellipsometric measurement for specular spectroscopic scatterometry. From [16].

There are several flavors of optical measurement for scatterometry, each of which employs a different combination of setup conditions. In variable-angle scatterometry, multiple incident angles are measured, but for only a single wavelength of incident light. Specular spectroscopic scatterometry (SSS), of primary interest in this work, measures only the zeroth-order diffraction response at a single fixed angle of incidence, but for multiple wavelengths. SSS is found to be a superior implementation for several reasons which will be discussed in section 2.3.3.

Similarly, there are different methods of profile reconstruction from the optically measured diffraction response. Although the forward calculation of the diffraction response from a candidate profile is tractable, the inverse—deriving the profile shape from the diffraction response—is not analytically possible due to the extreme nonlinearity of the relationship between profile and spectral response. In linearized inversion, small deviations from a known profile can be detected by changes in the diffraction response at particular wavelengths, and the resulting adjusted profile can be determined [16]. However, this method becomes less accurate as the profile shape “drift” grows. Another approach is a library-based approach [3]. Here, a RCWA simulation tool is used to calculate the diffraction responses resulting from gratings composed of copies of various profile, creating a library of diffraction response/profile pairs. (For more details on the forward simulation of the library, refer to Appendix B.) In order to extract a measured profile, the measured diffraction response is compared

to each of the diffraction responses contained in the simulated library. By finding the best-matching diffraction response (determined by basic least-squares comparison), the corresponding matching profile is selected from the library (Fig 2.8).

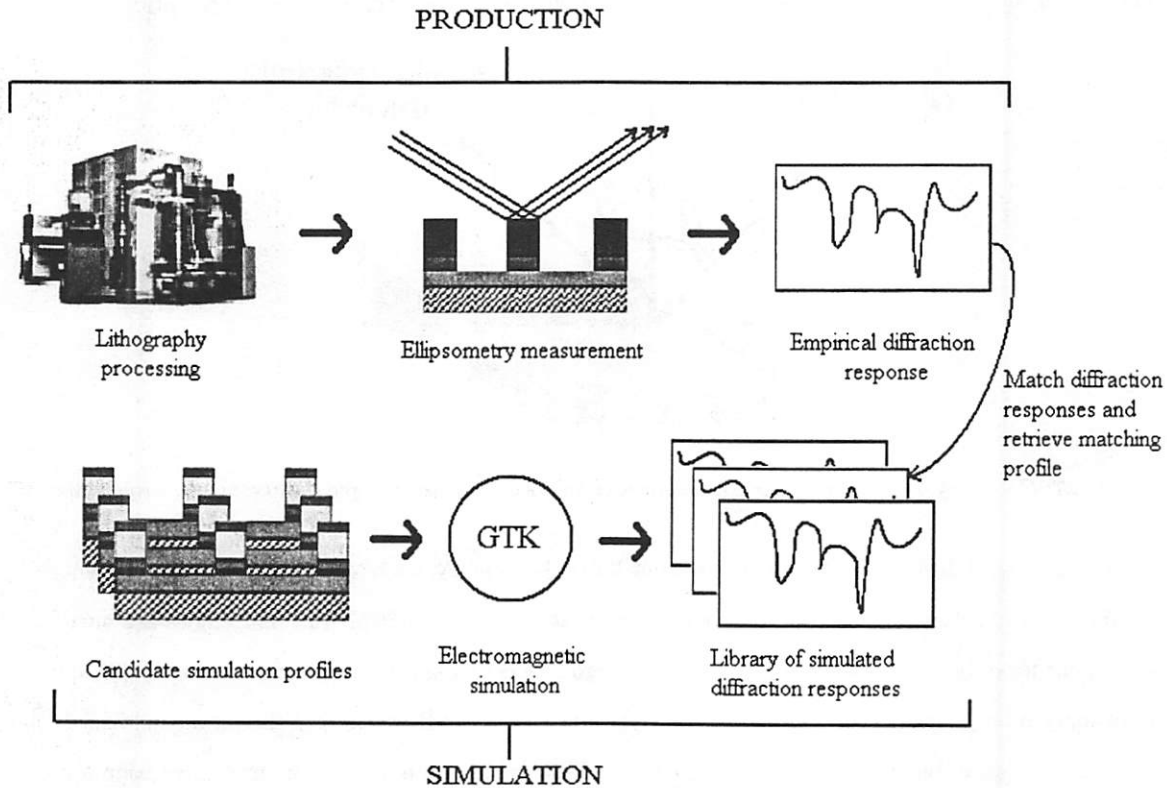


Figure 2.8: Flow diagram for SSS with library-based profile reconstruction. After [17].

In order for this library-based technique to provide accurate measurement, the library must be sufficiently large such that it always contains at least one profile that matches the actual profile to within some level of uncertainty. In addition, the library should also be efficient, designed to contain only as many profiles as are needed, thereby minimizing simulation time and guaranteeing a unique solution. In order to fulfill these requirements, typical scatterometry libraries contain ~300,000 - 500,000 profiles, and take roughly one day to create via simulation. However, scatterometry is currently most often used in a research setting, when a process may not be “well-behaved”; therefore, the library must be large enough which cover somewhat “unlikely” profiles. In an advanced process control setting, only a limited-size process window would need to be captured by the library—therefore, we might expect that libraries used for this purpose could be reduced in size. This subject

is investigated in this work.

Not surprisingly, scatterometry has proven to yield higher precision measurements than the CD-SEM because it avoids the focal dependence and impact of image variability. For example, when a ~ 180 nm shallow-trench isolation structure was repeatedly measured with CD-SEM over a course of roughly three weeks, the resulting distribution of measurements had a 3σ -variation of ~ 6.07 nm [18]. Scatterometry measurements taken over the same period yielded significantly more precise results; using the linear inversion method, the 3σ -variation was just 0.25 nm, and using the library matching method, the 3σ -variation was 0.75 nm (larger than the result for the linear inversion method due to discretization of the profile space). (All of the linear inversion results fell between the two discrete profiles reported using the library matching method, hence the slightly better precision for linear inversion.) Furthermore, it has been shown that the sensitivity of the diffraction response is highest at lower end of the current ellsoetric wavelength range (DUV spectral range), imdicating that sensitivity will be maintained as feature size (and hence metrology wavelengths) decreases. Recent evidence suggests that using the current inspection wavelength range of ~ 240 -780 nm, scatterometry will easily survive to at least the 65-nm technology node [19]. Beyond that, predictions are mixed—however, if the wavelength range can be extended deeper into the UV region, scatterometry should be applicable to any hypothetical device size.

2.3.3 Metrology Tools for Scatterometry

We have already mentioned two configurations for scatterometry; the single-wavelength, variable-angle configuration, and the multiple-wavelength, single-angle configuration. In both cases, the metrology instruments used to take the empirical measurement existed before the advent of scatterometry. For single-wavelength, variable-angle scatterometry, a reflectometer is the tool of choice. The light source is scanned through a specified range of angles as the detector is scanned through the reflection of that angular range, measuring the 0^{th} diffraction order. This configuration is called the $2\text{-}\theta$ configuration, since the variable angle of incidence (θ) is “counted” once each for the incident and detector arms of the apparatus.

$2\text{-}\theta$ scatterometry allows for a relatively easy configuration of metrology equipment and high signal to noise ratio, since a laser source is typically used (instead of a broadband light source). However, this configuration has several drawbacks. First, different grating structures have different wavelength ranges (falling anywhere from IR to UV) of high sensitivity—these information-rich ranges are the

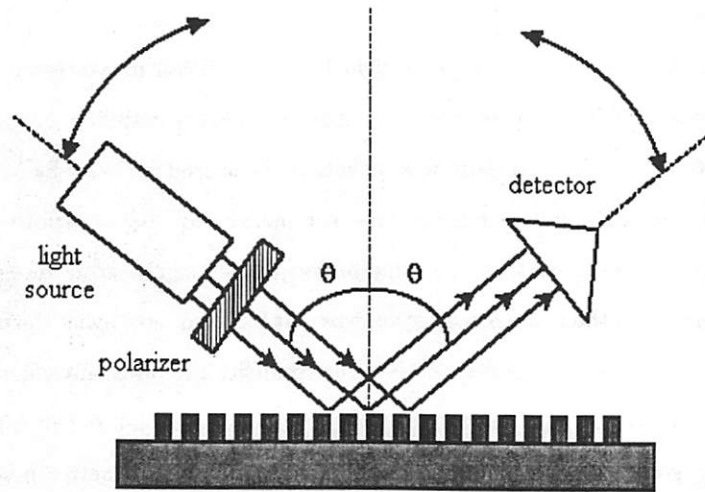


Figure 2.9: Variable-angle, single-wavelength (2θ) reflectometer. After [18].

most important to capture, but frequently the measurement wavelength for 2θ scatterometry will not fall within the sensitive range of the given grating. Second, the 2θ method cannot differentiate between two materials with similar refractive indices at the measurement wavelength. Finally, only the intensity, and not the phase, of the diffraction signal is measured [18].

Spectral spectroscopic scatterometry addresses each of these problems. Since this is a single-angle, multiple-wavelength method, the sensitive wavelengths of every grating will be measured. Because two different materials very rarely have similar refractive indices over a wide range of wavelengths, materials with similar optical properties are more easily differentiated. Finally, the SSS configuration measures both intensity and phase information, using a standard ellipsometer, as shown in Figure 2.10. When TE and TM linearly polarized light is directed onto a grating, elliptically polarized light is reflected due to the different magnitude and phase responses of the two polarizations. The ratio of the resulting responses can be expressed as

$$\frac{\tilde{r}_p}{\tilde{r}_s} = \tan\Psi * e^{i\Delta}, \quad (2.5)$$

where Ψ and Δ are the intensity and phase components, respectively, of the optical response, and \tilde{r}_p and \tilde{r}_s are the complex reflectance of TM and TE light, respectively [18]. Standard ellipsometer systems report $Tan\Psi$ and $Cos\Delta$ signals through transformations of the measured intensity—thus, the phase information is captured as well.

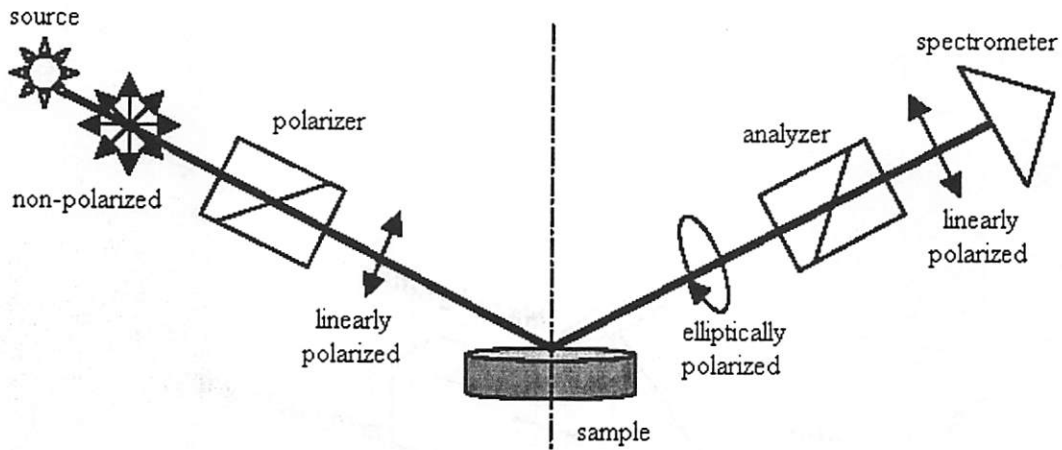


Figure 2.10: Variable-wavelength, single-angle ellipsometry. From [18].

In order to minimize the footprint of the metrology tool for scatterometry (and to facilitate its integration with existing hardware), normal-incidence reflectometry has been implemented for scatterometry. Used commonly for thin-film measurements, the normal-incidence reflectometer can essentially implement SSS with the single-angle set to 0° (with respect to the normal). Additional optical components must be added to the optical path to decouple TE and TM light, but \tilde{r}_s , \tilde{r}_p , and $\cos\Delta$ can all be measured to extract the grating profile. A diagram of the normal-incidence scatterometry system from Nanometrics is shown in Figure 2.11.

Several other optical configurations for scatterometry are available, implementing the many permutations of multiple- or single-wavelength, multiple- or single-angle, and variety of polarization. For the remainder of this thesis, we will consider the SSS method described above as the nominal metrology configuration.

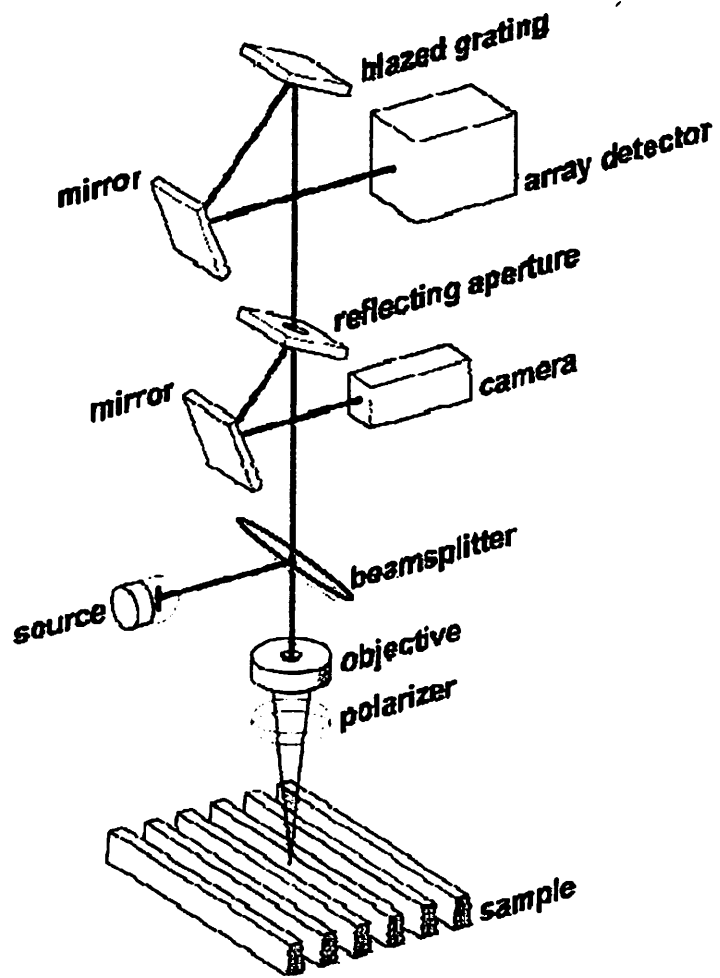


Figure 2.11: Schematics of Nanometrics normal-incidence reflectometry for scatterometry. From [20].

Chapter 3

Review of Process Control for Lithography

Motivated by the economic impact of efficiency and reproducibility, semiconductor manufacturing has grown increasingly automated. In turn, this automation has allowed for the implementation of long-understood concepts of scientific management of production [21]. In lithography, automated inspection techniques (such as CD-SEM and now scatterometry) have allowed engineers to efficiently collect data about the lithography process which can then be used to alert the operator when a lithography tool has malfunctioned, through statistical process control techniques. Over the last decade, a desire for even better automated process control has engendered real-time, adaptive, advanced process control. This chapter reviews these main areas of process control.

3.1 Statistical Process Control

In order to summarize statistical process control (SPC) we must first have an understanding of a few basic concepts. A process is in “statistical control” when it displays only random variations that are considered routine. A process that is out of statistical control displays systematic variation, defined as any variation that cannot be dismissed as random. When a process is diagnosed as being out of statistical control, then there exists an assignable cause that is the source of the systematic variability; once this source is eliminated, the process will again be in statistical control. The role of SPC in a production environment is to diagnose when a process is or is not operating under statistical control (i.e., SPC must differentiate between systematic and statistical variation). The potential improvement in efficiency possible through SPC is impressive and well-documented. When a manufacturing process can be quickly diagnosed as being out of statistical control, the cause of the

problem may be addressed and corrected before too many faulty parts are manufactured. For the lithography module in particular, fast detection of a tool malfunction reduces the volume of rework (removal of incorrectly-patterned photoresist and subsequent repetition of the lithography process) and therefore improves throughput. There are several SPC methods for detecting an out-of-control process. To get a flavor of typical SPC methods, we will briefly review the most basic concept—the control chart.

3.1.1 The Control Chart

The control chart is used to compare measured data to a known distribution of data from an in-control process. Most simply, the measured data can be compared to upper and lower limits derived from the “healthy” distribution, referred to as the control limits. These limits are usually set to be equal to $(\mu \pm 3\sigma)$, where μ and σ are the mean and standard deviation, respectively, of the distribution. That is, when the measured value of the parameter to be controlled exceeds $\mu+3\sigma$ or is less than $\mu-3\sigma$, then an alarm will be raised, indicating that the process has been determined to be out of statistical control. Of course, there will be points which fall outside of these control limits even if the process is in statistical control. We can expect a false alarm, or type I error, periodically based on the relationship

$$ARL = \frac{1}{\alpha}, \quad (3.1)$$

where ARL is the average run length—or number of runs between false alarms—and α is the probability of a type I error (assuming each measurement is independently and identically normally distributed—a common assumption). For 3σ control limits, it can be shown that $\alpha = 0.0027$, or that the ARL is about 370 measurements.

Similarly, there may be instances in which the process is in fact out of statistical control, but that the subsequent measurement falls between the control limits. Missing the detection of an out-of-control process is called a type II error. Correspondingly, we may define an out-of-control average run length ARL_{out} which gives the expected number of runs before an alarm when the process is out of statistical control:

$$ARL_{out} = \frac{1}{1 - \beta}. \quad (3.2)$$

Here β is the probability of a type II error—it is related to the magnitude and type of shift in the process. Naturally, we'd like to maximize ARL and minimize the ARL_{out} —that is, we'd like to minimize the occurrence of false alarms when the process is in control and also minimize the average delay before an alarm is raised when the process is out of control. These variables are strongly impacted by the frequency of data sampling, which must in turn be budgeted according to the delay in production due to data collection and the cost incurred by misprocessing allowed when out-of-control processes are not quickly detected.

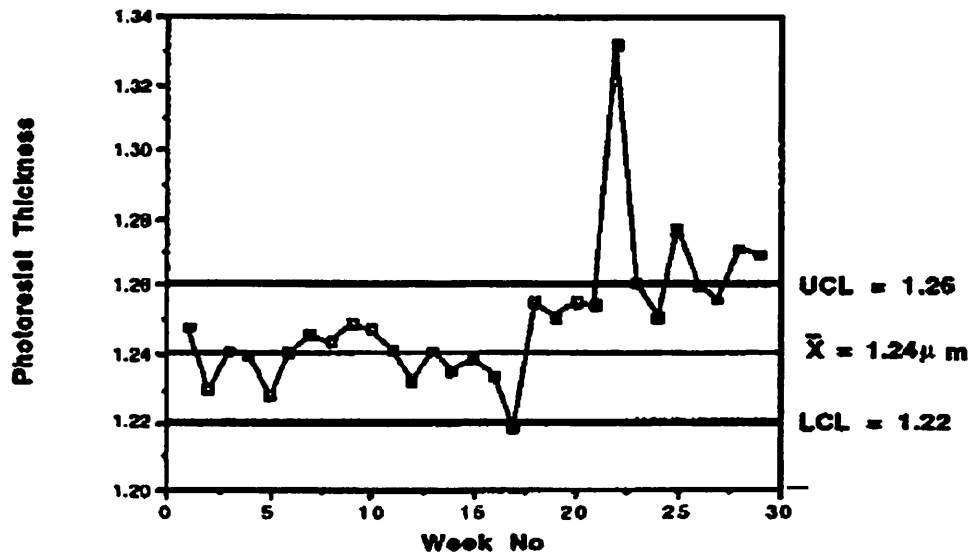


Figure 3.1: Control chart of photoresist thickness in the Berkeley Microfabrication Laboratory. From [21].

The detection of an an out-of-control process is not always as intuitive as watching for measurements that fall outside of the control limits. For example, it is not unusual for a process to oscillate within the range specified by the control limits. If only control-limit violations were taken into account, an alarm would never be raised in this situation, despite the fact that the oscillation could be attributed to an assignable cause which could in turn be corrected. Therefore, a system for analyzing patterns on a control chart was devised and standardized in the Western Electric rules. Essentially, these rules describe several criteria, each of which indicates that the process in question is out of statistical control. Examples of alarms include not only measurements that fall outside of the control limits, but also consecutively increasing or decreasing runs of measurements, sets of consecutive measurements that fall outside of narrower, 2σ “warning limits”, oscillating patterns in measurements, and other

unusual or nonrandom patterns [22]. In addition, several alternative types of control charts have been devised, each possessing heightened sensitivity to certain types of production processes.

3.1.2 Shortcomings of SPC

SPC techniques have become highly automated in the semiconductor industry, and are fully integrated with production. Unfortunately, SPC techniques do not provide ideal process control. Although statistically out-of-control processes may be identified, the SPC framework requires equipment shut-down, diagnosis of the source of the problem, and manual correction of that problem. In short, the SPC framework lacks a means of adaptive control. Additionally, in a high-mix, high-production fabrication facility, several products are run on a given lithography tool, with several different reticles and processing conditions. Therefore measurements on particular product lines are frequently appropriate only for their corresponding, specific control charts. Long delays between successive measurement collections weakens the sensitivity of SPC techniques, and the usage of different operating conditions increases the difficulty in diagnosing the problem when an alarm occurs. In short, SPC is a invaluable tool for process control but is too coarse to incorporate real-time corrective action to minor deviations in process state. To address this shortcoming, we must look to advance process control.

3.2 Advanced Process Control

Advanced process control (APC) described a control framework in which the process is adjusted adaptively on a real-time basis. In order to implement advanced control for a given process, three general tasks must be accomplished. First, we require a metrology system which can give an estimate of the state of the process to be controlled. (For lithography, we will argue that full-profile measurement is necessary to provide sufficient data for process-state estimation; in turn, deriving the state of the process from the profile measurement will be the focus of the remainder of this thesis.) Next, we need a method by which we can compare the estimated state of the process to the desired state of the process. Finally, we must adjust the process dynamically to reduce the disagreement between the measured state of the system and the desired state of the system. These latter two components have been thoroughly studied, and should not present a limiting difficulty in the implementation of a lithography-specific process control framework. Several methods by which a process may be

adaptively controlled via APC techniques are described in the following sections.

3.2.1 Run-to-Run Control

Run-to-run (R2R) control is a control method in which process parameters are tuned via feedback (FB), feed forward (FF), or FB/FF combination models between successive runs (iterations) of a given process. Adjustments in process parameter values are calculated using process response measurements applied to some process control algorithm. A rudimentary feedback control loop block diagram is shown in Figure 3.2.

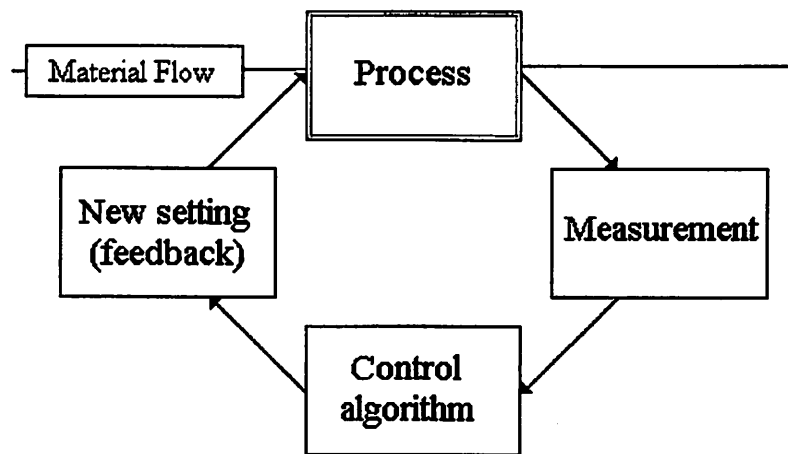


Figure 3.2: Block diagram of R2R feedback control model. After [1].

As an example of a simple approach for a R2R control model for lithography, suppose that the measured parameter (CD) has been trending normally (i.e., within statistical control) until several consecutive runs indicate an increasing measured CD. This information will be passed on to the control algorithm, which in turn will adjust a controlling process parameter (feedback). In this instance, the exposure dose—the most common process parameter used to tune lithography CD—would be increased, thereby reducing the measured CD in subsequent runs, preventing the trend from continuing. In a feed forward situation, properties of the incoming wafer will be used to adjust the upcoming lithography process. For example, the thickness of the incoming film stack might suddenly shift; a measurement of the new film stack thicknesses might be passed to the control algorithm, which would in turn adjust process parameters to compensate for the shift in the incoming material.

The Control Algorithm—PID Control

Both types of R2R control require a sophisticated control algorithm. One basic feedback control algorithm is PID control. (Similar algorithms exist for feed forward control and for FB/FF combinations, but PID feedback control will give the reader sufficient sense of the level of complexity of a standard control algorithm.) The first element of PID control is proportional control. The equations describing proportional-only control are simple:

$$\text{error} = e = \text{measurement} - \text{setpoint} \quad (3.3)$$

$$\text{gain} = g = \frac{\text{Output change}}{\text{Input change}} \quad (3.4)$$

$$\text{adjustment} = g * e + \text{bias} \quad (3.5)$$

The *setpoint* is the desired value of the quantity being measured (i.e., CD), the *gain* is the ratio of change in the output to the change in the input, and the *bias* term is correction to the control loop engaged whenever an offset is observed. If only proportional control is used, the bias must be tuned manually by an operator. To automate this correction, the second element of PID control—integral control—is implemented. In PI control, the bias is captured by the integral of the error over time; the corresponding equation is

$$\text{adjustment} = g * (e + K_i \int (e)dt), \quad (3.6)$$

where K_i is a constant factor that scales the integral element by the time it takes for the integral element to repeat the action of the proportional element, and has units of $1/\text{time}$ (i.e., repeats/minute). By using PI control, cumulative drift in the control loop is automatically accounted for. However, the adjustment result displays undesirable lag compared to the error in many instances, and strong oscillation in the signal can result. To compensate for the lag introduced by the integral control term, derivative—or pre-act—control is used as the final component of PID control. With derivative control, the corrective adjustment is proportional to the rate of change of the error—thus, action is taken to inhibit rapid changes in the process (such as overshoot). Under full PID control, the adjustment equation becomes

$$\text{adjustment} = g * (e + K_i \int (e)dt + K_d \frac{de}{dt}), \quad (3.7)$$

where K_d is a derivative-term proportionality constant—the amount of time that the derivative action advances the output. In order to achieve effective R2R control under the PID framework, the values of the constants g , K_i , and K_d have to be tuned to balance the effects of the proportional, integral, and derivative action, which depend on the nature (i.e., frequency and magnitude) of typical error in the process.

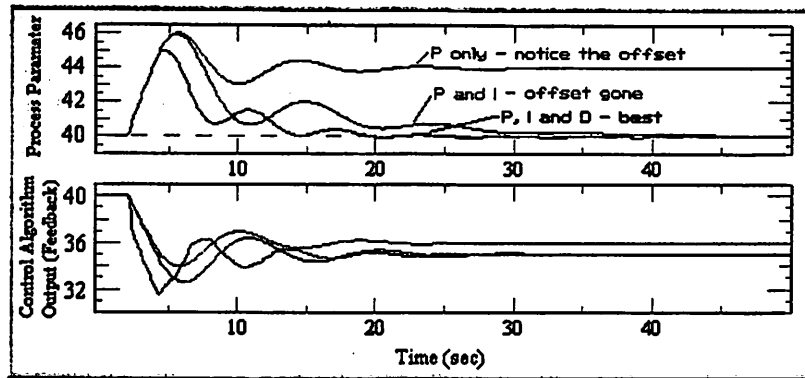


Figure 3.3: Illustration of PID control, compared to PI and P-only control. Courtesy of ExperTune Loop Simulator, from [23].

Shortcomings of Modern R2R Control

R2R control has been demonstrated to be a useful control tool in several processes. In fact, R2R control has been successful in the lithography module, as first implemented by Motorola in 1998 [24]. By dynamically tuning the lithography process, manufacturing delays are reduced as compared with SPC, where tools must be taken off line when a problem is detected. Although R2R control has proved to be an indispensable technique for enabling a heightened level of process control, it is not perfect. In the case of modern lithography, R2R control suffers from the fact that literally dozens of process parameters have significant impact on the resulting resist CD, because R2R methods do not attempt to monitor all of these parameters. Usually, the response variable is simply the resist linewidth, and the controlling process parameter is exposure dose (as in the case of the Motorola effort) and occasionally also focus. However, the exposure energy frequently is not the cause of CD variation—instead, fluctuations in the PEB step or variations in the incoming film stack may be the source of variability. In these instances, tuning the exposure and/or focus to compensate for variation in other process parameters will ultimately fail, particularly as process windows shrink in order to meet demand for higher performance electronics while maintaining yield. In order to accomplish

efficient and thorough process control, more parameters must be monitored, which requires more sophisticated metrology.

In addition, basic R2R control algorithms suffer from variability due to the bias introduced by the unique attributes of exposure tools, reticles, and various product flows in the modern high-mix fab environment. Since each exposure tool suffers from unique aberrations, each reticle possesses unique linewidth errors, and each product has a different incoming material stack, there are an intimidating number of permutations of total process bias that must be accounted for in order to avoid a steep adjustment in the first few runs of an R2R framework.

3.2.2 Predictive Process Control

Predictive process control (PPC) refers to an advanced process control framework for lithography which is enhanced by a methodology of collecting not only measurement results, but also the bias attributes of the various exposure tools, reticles, and incoming processes in a “signatures database”. With PPC, the signature database is used to predict the most-likely-correct initial process setup for a given product on a particular tool with a specific reticle before the lot processing begins [25]. If implemented effectively, PPC promises to reduce the large corrections (caused by the factors listed above) at the beginning of each processing run that are typical for a standard R2R control framework.

Many anticipate that PPC will become the control method of choice in high-mix ASIC fabs [2]. But in some instances, impressive process control has been achieved in ASIC fabs using much simpler techniques. For example, LSI has recently presented results on an extremely simplified control framework for lithography overlay control in an ASIC fab [26]. Rather than calling on separate models for each tool-reticle-product-layer combination, a more general process model was applied to all combinations, with substantial results of reduced rework and improved process capability. However, the crux of this overlay control method was the observation that overlay errors add linearly, and are directly proportional to alignment input parameter values—this simple relationship allows for quick and accurate extraction of the overlay state. An equally simplified framework would fail in the case of CD control since the profile does not have a linear relationship with the inputs to the lithography process.

3.3 Process Control Summary

Process control has been implemented effectively in semiconductor manufacturing. Control methodology has evolved from the manual-adjustment model of SPC to the automated-adjustment framework of APC. APC models have proven to improve process capability and in turn increase yield. However, current sources of metrology do not capture the state of the system; instead, current systems rely on the measurement of a single an output variable, such as CD. While CD drift may be compensated by tuning one or two of the most significant controlling parameters (such as exposure and focus settings), this approach will eventually fail if the actual drifting parameter is not considered within the control algorithm. In order to achieve the level of CD control that will be required in future technologies, a more sophisticated control framework, using a more thorough means of metrology and process state estimation, must be devised. The transition from CD-SEM metrology to scatterometry will improve APC methods since the profile measurement yields much more information than the top-down linewidth measurement; however, procedures to extract the precise process state from the profile (a procedure called “profile inversion”) are still immature. The remainder of this thesis presents a potential state-estimation method and results from simulations.

Chapter 4

Development of the Process Control Framework

As discussed in Chapter 2, scatterometry provides a non-invasive, highly accurate, and fast means of full-profile metrology. As such, it is a highly suitable metrology for process control in lithography. Although scatterometry requires the use of an “off-circuit” test grating, tight wafer-to-wafer and/or die-to-die control over the characteristics of the grating—accomplished through feedback to the lithography module—will extend a similar level of control of the resist features which define device patterns on the rest of the chip. Within-die variation will require other metrology and alternate control techniques, since scatterometry does not directly measure the actual device pattern.

4.1 Lithography Process Control Using Scatterometry

A general scatterometry-based lithography process control flow is shown in Figure 4.1. After each wafer is processed in the lithography module, standard scatterometry is carried out. The diffraction response of a test grating pattern on each wafer (or perhaps on several die of each wafer) is measured using an ellipsometer; by matching the measured diffraction response to an entry in a prepared library of simulated diffraction responses, the resist profile is extracted. Then, based on the measured profile, the actual input conditions of the lithography module (resist thickness, bake parameters, exposure dose, focus setting, etc.) are determined. Finally, a control algorithm is employed to compare the actual input parameter values to the ideal input parameter values and thus determine appropriate adjustments to the lithography equipment settings. In this fashion, any offsets or drifts in the lithography module will be reconciled on a wafer to wafer basis.

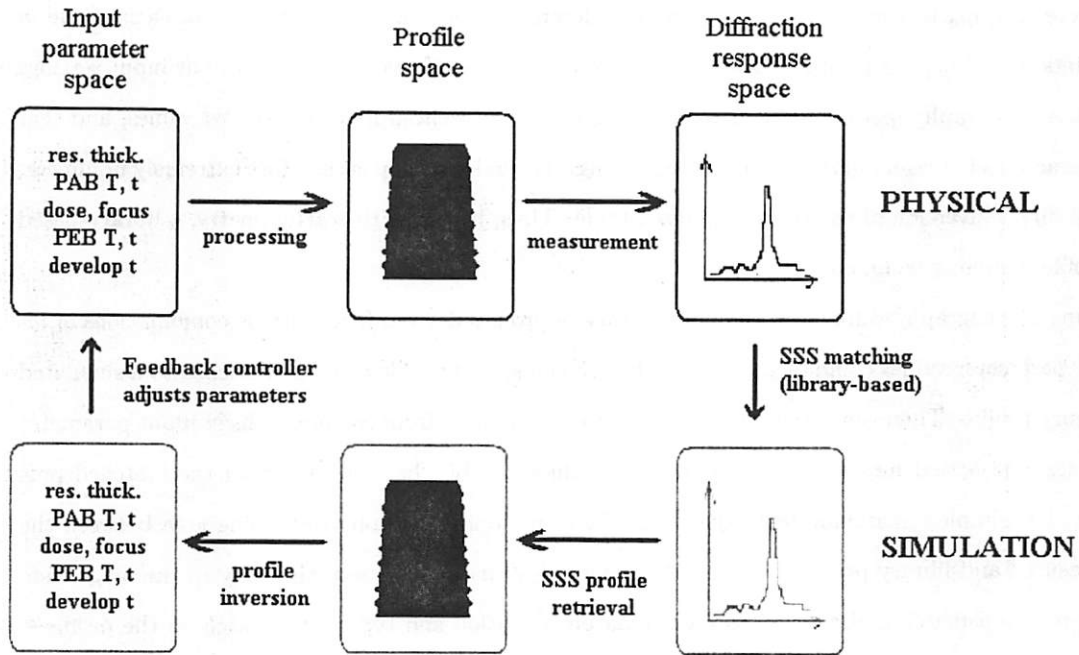


Figure 4.1: An illustration of the lithography process control flow.

Notice that this general scatterometry-based process control flow relies on two separate steps to extract the actual input conditions from the measured diffraction response. Although obtaining the profile from the diffraction response is a well-understood practice, obtaining the input conditions from the profile—a step called “profile inversion”—is less straight-forward.

4.1.1 Potential Profile Inversion Solutions

Based on the way the process control flow is illustrated in Figure 4.1, one might expect that profile inversion will entail direct analysis of the profile. For example, one possibility might be to establish a correspondence between fluctuations in particular process parameters (such as resist thickness, exposure dose, focus setting, bake parameters, and develop parameters) and resulting fluctuations in resist profile features (such as footing, top rounding, sidewall angle, and CD). If we describe the input process parameters as a vector X_{inputs} of dimension $n * 1$ where n is the number of inputs to be considered, then the resulting profile can be determined by some function

$$X_{profile} = f(X_{inputs}) \quad (4.1)$$

where $X_{profile}$ is a $m * 2$ dimensional matrix describing the x and y coordinates of each of the m points describing the profile. However, it is known that the effects due to the various input settings of the lithography module are variable over the range of potential input parameter values, and that interactions between input parameters are strong. This relationship is therefore extremely nonlinear, and direct inversion of the relation is intractable. Thus, just as with scatterometry, a library-based profile matching technique might be pursued.

Using a lithography simulation engine, a library of profiles derived from various combinations of input parameter values could be generated; then the measured profile could be matched to a simulated library profile. Then, since the library profiles were simulated from combinations of input parameter values, the actual measured input parameter values can be “backed out” from the matched profile. The simplest matching technique would be to minimize the non-overlapping area between the measured and library profile. However, this method of matching is sensitive only to the magnitude of profile mismatch and not to the specific nature (location and type of mismatch on the profile—sidewall angle, width, height, T-topping, footing, etc.) of the disagreement. With sufficient density of the library, this problem might be avoidable, but if we move a level higher in abstraction and take advantage of the diffraction response, inversion may be simplified.

It is evident that the diffraction response displays strong sensitivity to different profile features, as shown by the proven accuracy of scatterometric profile measurement. Therefore, if we extend the library-based profile matching technique to the diffraction response domain, the sensitivity of the diffraction response can be exploited to better distinguish among profiles with similar non-overlapping areas. (For two candidate matches to a measured profile with similar profile disagreement, the one whose corresponding simulated diffraction response better matches that of the measured profile will have more closely corresponding input parameter values.) Continuing the mathematical formulation of the profile inversion problem, we can in turn express the resulting diffraction response of a grating composed of copies of a given profile as a function of the points which comprise that profile:

$$X_{diff} = g(X_{profile}) \quad (4.2)$$

where X_{diff} is a $p*2$ dimensional matrix containing the $\text{Tan}(\Psi)$ and $\text{Cos}(\Delta)$ values of the diffraction response for each of the p wavelengths used—in this research, 53 wavelengths ranging from 240 nm

to 760 nm in steps of 10 nm.

The convolution of these two functions gives a function relating the diffraction response to the values of the input parameters:

$$X_{diff} = g(f(X_{inputs})) = h(X_{inputs}) \quad (4.3)$$

This relation, also, is extremely nonlinear and cannot be solved analytically. However, the library-based method can be used simply by extending the profile library to a diffraction response library using the established simulation techniques that are already used for scatterometry library construction.

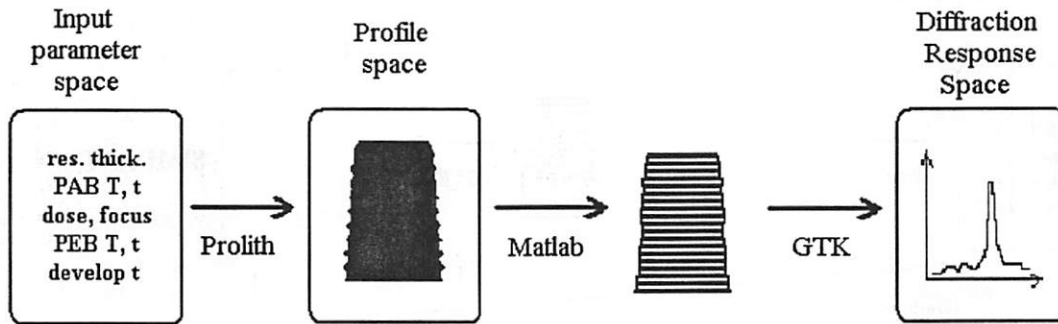


Figure 4.2: Software flow for simulation of PROLITH-based scatterometry library.

As shown in Figure 4.2, a first-principle based lithography simulation engine (nominally PROLITH, which is used in this work) is used to generate profiles from simulated input conditions. These profiles are then digitized (using Matlab) into stacks of discrete rectangular segments, as is required for RCWA simulation. Finally, the digitized profiles are submitted as input profiles to the standard scatterometry library construction framework (here, the grating tool kit—GTK [27]—is used), which executes electromagnetic simulation of the profiles to generate a collection of diffraction responses. By linking the “input parameter space” to the “profile space” and in turn to the “diffraction response space” during forward simulation, the matched diffraction response can be tracked back not only to the corresponding resist profile but also to the corresponding set of input conditions which “created” that profile. Assuming that the lithography simulator can be tuned to match the production tool and environment—a topic discussed in Section 4.4—and that the diffraction response matching method yields enough sensitivity to input conditions, this method could potentially allow for determination

of the actual process conditions of the lithography module in a modified scatterometry-based process control framework, as shown in Figure 4.3.

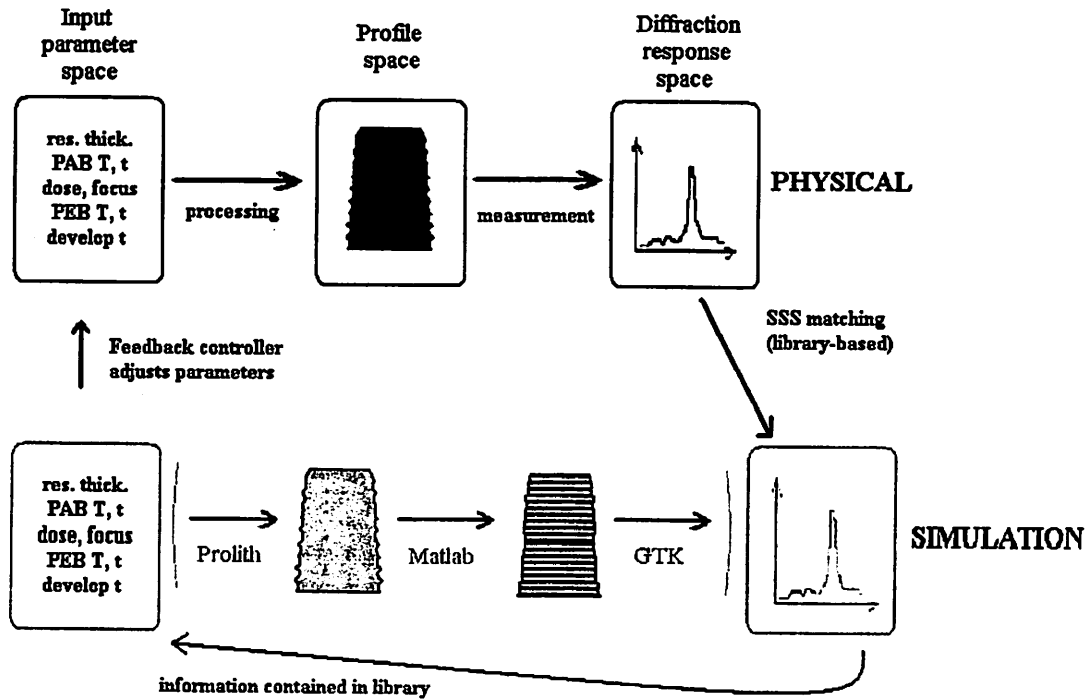


Figure 4.3: Illustration of the modified, PROLITH-based lithography process control flow.

4.2 Considerations for Library Development

Just as in the case of a standard library-based scatterometry framework, the library for this process control framework should strike a balance between sufficiency and efficiency. That is, the library must contain a sufficient number of entries such that within a given process window, a match to any measured diffraction response is guaranteed to within a specified error; but it must also be small enough to guarantee a unique solution—additional profiles would needlessly waste additional simulation time. Based on these rules, it has been observed that for a typical scatterometry application, a library should contain ~300,000 - 500,000 profiles. However, scatterometry is generally used in a research and development setting, where extremely wide profile coverage is desired. Therefore, the ~500,000 number may be much larger than is required for a well-behaved process that is a candidate for process control, where the required number of profiles is dictated only by the process window.

In short, standard scatterometry library-building techniques are inefficient for the process control scenario.

4.2.1 Library-Building Inefficiencies in Standard Scatterometry

Standard scatterometry libraries are generated by starting in the profile space, where candidate library profiles are systematically chosen for simulation based strictly on the goal of achieving thorough coverage of all potential profile geometries. For example, a library will find the nominal (most desired) profile shape, and then vary geometric parameters around the nominal values over a specified range. A desired profile might have vertical sidewalls—the library might then contain profiles with sidewall angles ranging from 85° to 95° in steps of 0.5° . If the nominal CD is 130 nm, then the range from 125 nm to 135 nm might be covered in steps of 1 nm, and so on. The upshot of this method is that traditional scatterometry libraries may have poor sensitivity to certain ranges of input parameter values and excessive sensitivity to others since there is no correlation between the geometric profiles simulated and the input conditions which could actually produce them. For example, if only CD and sidewall angle are accounted for in the library construction (for the sake of illustration) then the library might have good sensitivity to exposure dose and focus settings, but poor sensitivity to bake properties and resist thickness, which might more strongly affect other geometric parameters. The candidate profiles are also chosen without consideration of the diffraction response sensitivity—if the diffraction response changes dramatically for small fluctuations in one range of a geometric parameter but minimally for another range, this gives rise to sensitivity imbalance, creating further inefficiency. For standard scatterometry usage, this is not important, since the only figure of interest is the profile, and the library is designed to capture every potential profile. However, in a process control setting, it may prove useful to utilize the sensitivity of the diffraction response space in balanced fashion. Thus, the method for building the library can be improved for the purpose of creating a library with the highest potential for success in the parameter extraction framework introduced in section 4.1.

4.2.2 Proposed Library Building Methods

In order to address the first of these potential shortcomings, all libraries studied in this work are “initiated” in the input parameter space. That is, entries are introduced to the library’s collection

by first choosing combinations of the input parameter values. As shown in Fig. 4.2, the profile and diffraction response components of the entry are then calculated based on these input parameter values, ensuring that the library contains only physically-realizable entries. To address the second potential shortcoming, the procedure could also be modified to include a dynamic awareness of the "state" of the diffraction response space; that is, the candidate combinations of input parameter values could be chosen such that the sensitivity of the diffraction response space is balanced, since this is where the matching will take place. To achieve this second potential improvement, an iterative library building method is proposed and evaluated.

An initial "seed" library is chosen such that the entries in the seed library will cover the entire range of the desired process window sparsely but evenly in the input parameter space. These entries (combinations of input parameter values) are simulated using PROLITH, Matlab, and GTK as described above to generate corresponding profiles and diffraction responses. Next, the sensitivity balance of the diffraction response space must be evaluated. Library entries in proximity to each other in the input parameter space (i.e., two points neighbor each other if they differ only in one input parameter value, and if the difference is of minimum step size) are considered in discrete groups such that the groups completely cover the diffraction response space without overlap, and contain 2^n entries where n is the number of input parameters being considered. For example, let us suppose that $n = 2$, and that the input parameters are x and y . Then each group will consist of a "corner" point, a second point with the same x -coordinate but with a y -coordinate that is the minimum positive distance away from the corner point in the y -direction, a third point with the same y -coordinate but with an x -coordinate that is the minimum positive distance away from the corner point in the x -direction, and a fourth point with both an x -coordinate and a y -coordinate that are the minimum positive distances away from the corner point in the x -direction and y -direction, respectively. So, in two dimensions, each group describes a square in the input parameter space; if three parameters are considered, each group will contain the 8 corner points of a cube in the input parameter space. In any number of dimensions, by beginning in the appropriate corner of the seed library and continuing systematically to all points which do not form non-adjacent edges of the library, the entire "area" ("volume" if three input parameters are considered, etc.) of the library will be covered by the union of the groups, with zero intersection.

Next, for each of these groups, a measure of the group "area" will be calculated. There is no guarantee that the groups will have convex shape, and in higher dimensions when more input

parameters considered, direct “area”-like calculations are difficult. Therefore, the group “area” is calculated by determining sum of squared differences (SSD) between each pair of entries in a group, where the SSD is the sum of squared differences at each sampled wavelength in the two diffraction response signals A and B :

$$SSD = \sum_{\lambda=240nm}^{760nm} [(Tan\Psi)_A - (Tan\Psi)_B]^2 + [(Cos\Delta)_A - (Cos\Delta)_B]^2 \quad (4.4)$$

$$area = \sum_{i=1}^n \sum_{j=i+1}^n SSD_{ij} \quad (4.5)$$

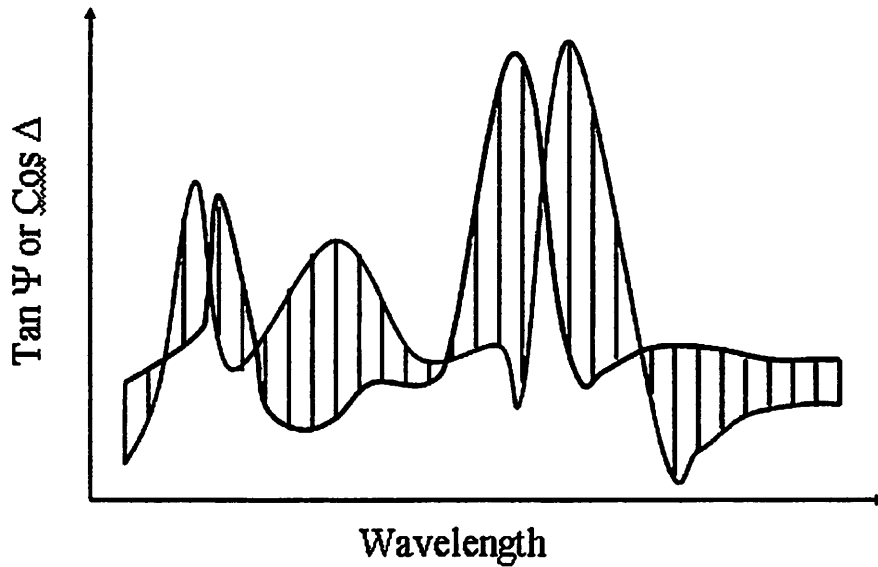


Figure 4.4: Sum of squared differences calculation illustration. The sum of squared lengths of the line segments comprises the SSD metric.

It is hoped that the SSD metric will identify the groups with the largest inherent diffraction response sensitivity, even if that metric does not capture the true area (or volume, in three dimensions, etc.)—its efficacy will be evaluated in the following chapter. After calculating the areas of all groups, the groups which have the “largest” area can be picked out and targeted for further population in the next iteration. The rule for selecting which groups will receive additional points can be tuned according to desired library size and sensitivity balance in the diffraction response space. In this work, the groups whose area exceeded the $mean + 2\sigma$ value were commonly chosen at each iteration.

In this fashion, the input parameter values are chosen to maximize the balance of sensitivity in the diffraction response space. As an example, Figures 4.5 and 4.6 illustrate the iterative method for a nonlinear example of two parametric functions,

$$F(a, b) = a^{\frac{1}{2}} * b \quad (4.6)$$

$$G(a, b) = a^8 * b^{\frac{1}{8}} \quad (4.7)$$

The initial seed library is shown, both for the “ (a, b) -space”, which symbolizes the input parameter space, and for the “ $[F(a, b), G(a, b)]$ -space”, which symbolizes the diffraction response space. In each of 5 successive iteration, points are added to the (a, b) -space to maximize uniformity of the $[F(a, b), G(a, b)]$ -space.

Although the library will have an increasingly balanced overall sensitivity with each successive iteration, the possible imbalance in sensitivity between two given input parameters may not be improved. For example, if the original seed library is chosen such that there is greater sensitivity in one parameter than another parameter (which results in the entries being close together along one “axis” in the diffraction response space but far apart along another axis), the ratio of the imbalance will remain the same despite the iterative method. This imbalance is preserved because the method does not account for discrepancies between sensitivities to individual parameters, but rather to groups of parameters. That is, at each iteration, a group selected to be populated with more entries will have equal density increase in all parameters. While it is possible to accommodate further analysis in the iterative method to correct for this problem, the complexity presents quite a challenge, and in this work, the initial seed library was chosen with care to achieve approximate balance between average initial sensitivity in processing parameters.

In this work, three libraries will be built and evaluated in the state estimation framework. First, the “seed” library (of 2,420 entries) will be evaluated by itself. Second, a much larger library of 26,991 entries will be built using the iterative method. Third, a similarly larger library of 27,783 entries will be built, but rather than using the iterative method, this library will be constructed to evenly cover the input parameter space; in effect, this library is just like the “seed” library except with increased density in every processing parameter.

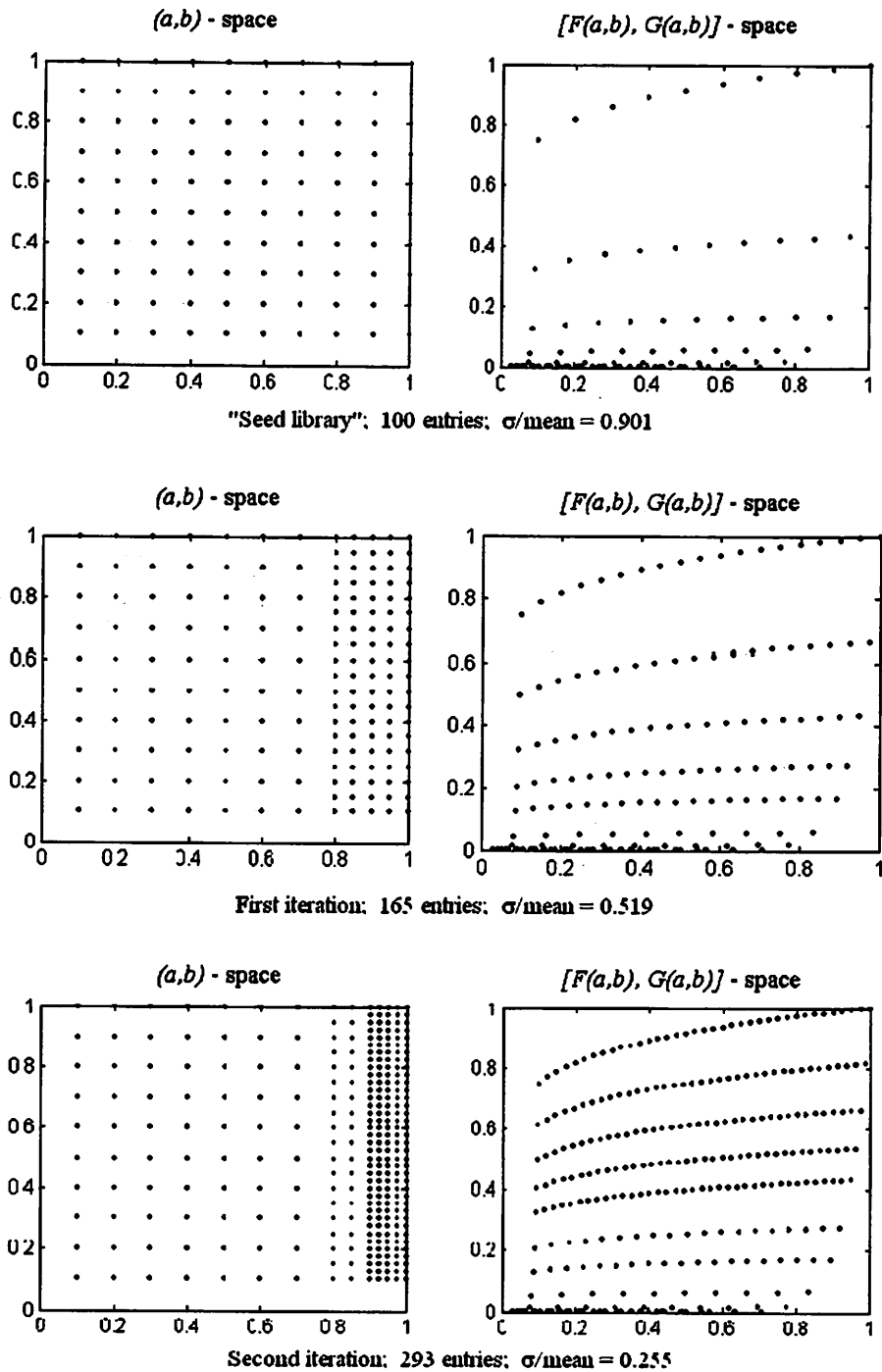


Figure 4.5: Illustration of the iterative library building method using a simple nonlinear function. The seed library and first two iterations are pictured here. The number of entries as well as σ/mean of group area are listed under each set of plots.

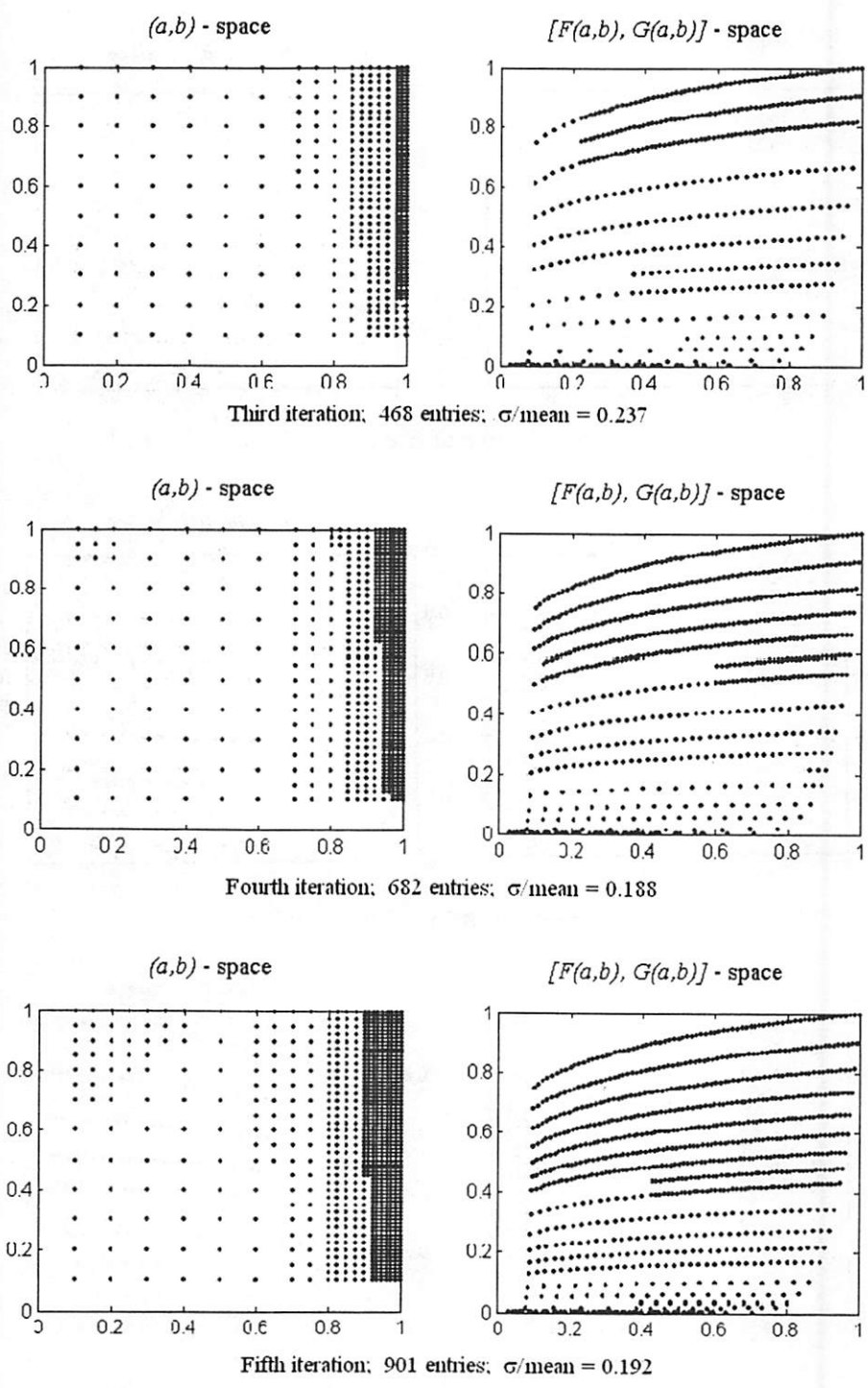


Figure 4.6: Illustration of the iterative library building method using a simple nonlinear function. Iterations 3 through 5 are shown. The number of entries as well as σ/mean of group area are listed under each set of plots. For comparison, 961 entries distributed evenly in the (a, b) -space lead to a σ/mean of 0.912 in the $[F(a, b), G(a, b)]$ -space.

4.3 Potential Pitfalls

The proposed library-building technique will yield a library whose entries are all physically realizable, and (in the case of the iteratively-built library) whose diffraction response space is well-balanced. However, there may still be problems with profile inversion by this method. These potential problems are all introduced because neither profile uniqueness nor diffraction response uniqueness is guaranteed. That is, there is a possibility that two different sets of input parameter values might create virtually identical resulting resist profiles. Similarly, it is possible that two different profiles could yield the same (or very similar) diffraction response. In the latter case, scatterometry's success has shown that although there is no proof of diffraction response uniqueness, this issue is not problematic in practice—there have been no observations of multiple disparate profiles having the same diffraction response. However, it is certain that multiple combinations of input parameter values will lead to virtually the same profile. For example, exposure dose and PEB temperature both most strongly influence the resist feature width. For a given profile, an exposure dose increase can likely be offset by an appropriate PEB temperature decrease to yield the exact same profile. Although this relationship is not perfect—for instance, higher PEB temperature should lead to a smoother resist profile since the acid diffusivity will be increased—it will, particularly in combination with adjustments in other input parameter values, give rise to nonuniqueness in the profile space.

Therefore, it is likely that the proposed profile inversion method will suffer from some inaccuracy as it stands. However, by supplying the profile inversion framework with “extra” information, we can improve the accuracy. Extending the example from above, we might supply the framework with an approximate value of the PEB temperature (supposing we can measure the temperature periodically to calibrate the expected temperature), narrowing the scope of the search within the library, which in turn might lead to a more accurate extraction of the exposure dose. Therefore, the analysis of the proposed profile inversion method will necessarily include such an investigation of how much information must be supplied to the framework to ensure a given level of accuracy in predicting the unknown parameter values.

4.4 PROLITH

In order to generate a library with a basis in the input process parameter domain, we require a first-principle based lithography simulator that generates simulated profiles from simulated input

conditions. Moreover, this lithography simulator should be tuned to the actual lithography process in order for the library to be applicable for measurement of actual input conditions. PROLITH is an excellent candidate for this purpose, because it is arguably the semiconductor industry's most-used lithography simulator, and has therefore received significant research attention. A brief overview of this software is required in order to discuss PROLITH's adaptability to matching the actual performance of a given lithography process.

The Positive Resist Optical Lithography model (PROLITH) was initially developed in 1985 in a Department of Defense project [4]. The model was used to predict resist profiles for submicron projection, proximity, and contact printing, using first-principle-based models for optical projection systems, the exposure reaction, and the development process. Since its introduction, PROLITH has been commercialized, first under FINLE Technologies, and now KLA-Tencor. Over the course of PROLITH's evolution, the user has been provided with increasing control over the model parameters, and functionality and simulation speed has increased. Currently, PROLITH is widely used as an offline and/or diagnostic tool, mainly for process development. The stand-alone user interface allows for adjustment of all the major processing parameters. The substrate stack is specified by a set of model files, one for each material layer. In particular, the resist is captured by a model containing information about the resist tone (positive or negative), development parameters and development model to be used (Mack, enhanced, or notch), bake response properties (acid diffusivity, amplification in the case of a CA resist), and Cauchy coefficients for calculation of the index of refraction over a range of wavelengths. Next, the PAB parameters—temperature and time—are specified, and the user creates a mask image, such as a contact hole or a resist line, of specified size and spacing. Then the exposure system and parameters are defined: illumination source, wavelength, partial coherence, and numerical aperture, as well as exposure dose and focus offset, are entered by the user. (The exposure tab is pictured in Figure 4.7.) Finally, the PEB parameters (time and temperature) and develop time are specified.

Once all of the models, parameters, and input conditions are specified, the user may retrieve the several simulation results. The aerial image (shown in Figure 4.8), the image intensity in the resist (Fig. 4.9), the latent resist image following the PEB step (Fig. 4.10), and the final post-development resist profile (Fig. 4.11) can all be extracted. PROLITH can also be used in stand-alone mode to find optimal processing conditions. For instance, focus-exposure matrix simulation sets can be run to determine optimal exposure conditions; also, dose-to-clear calculations can be carried out quickly

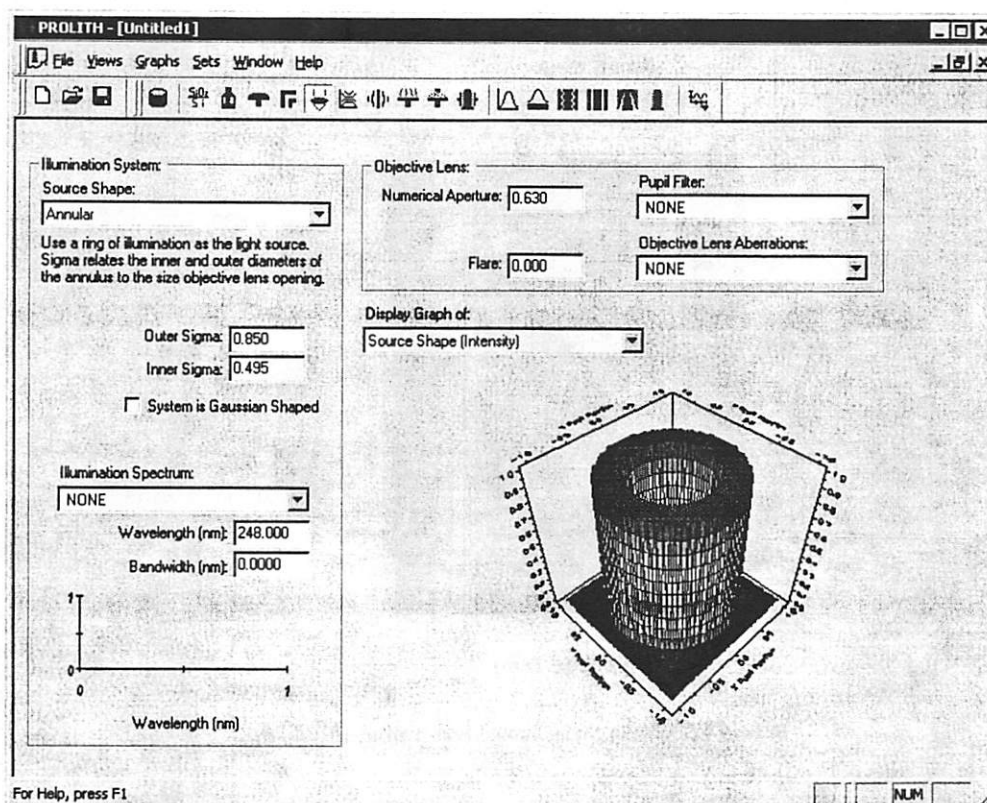


Figure 4.7: A snapshot of the exposure system tab of the PROLITH v6.1.2 user interface.

to evaluate resist sensitivity.

Finally, and most convenient for this work, PROLITH can be accessed remotely to execute very large batches of simulations. Using a Microsoft Excel/Visual Basic-based software framework, the user can specify all of the parameters available in stand-alone mode and can extract specified results in text-file format. In the interests of this work, for example, several thousand profiles are simulated at a time, and for each profile, a data file containing the (x,y)-coordinates describing that profile are saved to a file. Later, each file can be accessed in order for the profile to be digitized and passed to the GTK software. (For more specific details about the developed software package, refer to Appendix B). When accessed in remote mode, PROLITH not only offers enhanced data collection, but also increased speed—simulation of a single resist profile takes roughly 1 second on a 1.7 GHz processor. If a PROLITH batch-mode job were split up among several CPUs as is commonly done in industry for simulation of scatterometry libraries, the effective simulation time can be reduced easily by an arbitrary factor.

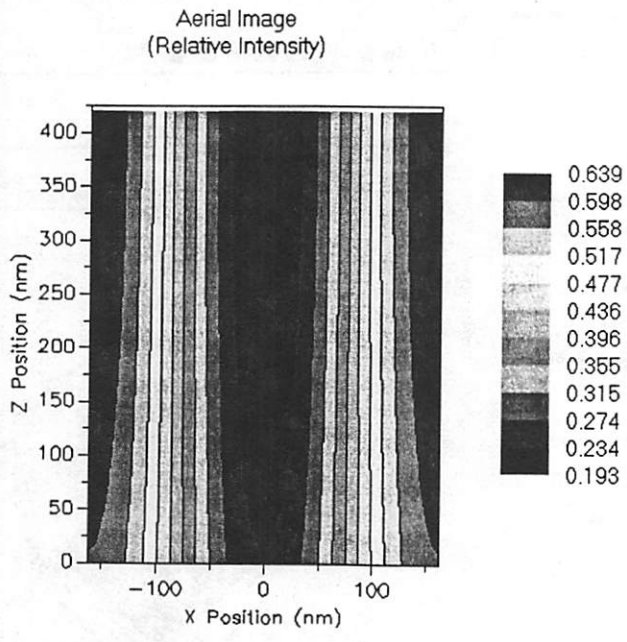


Figure 4.8: Sample aerial image result from PROLITH.

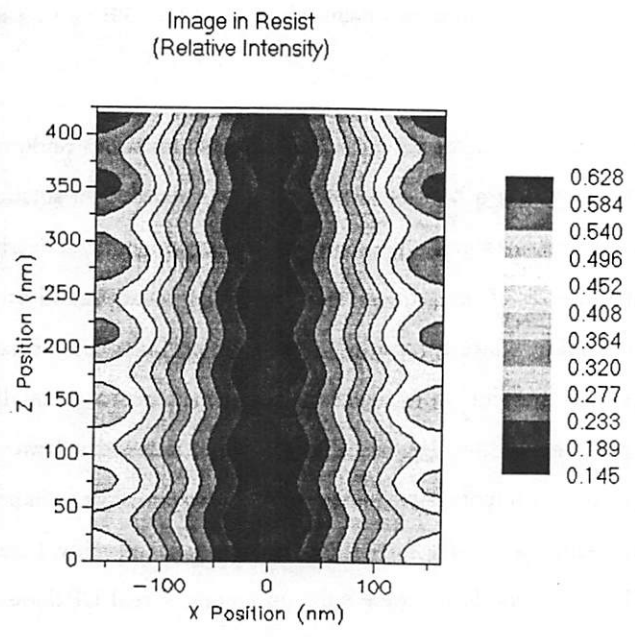


Figure 4.9: Sample image-intensity-in-resist result from PROLITH.

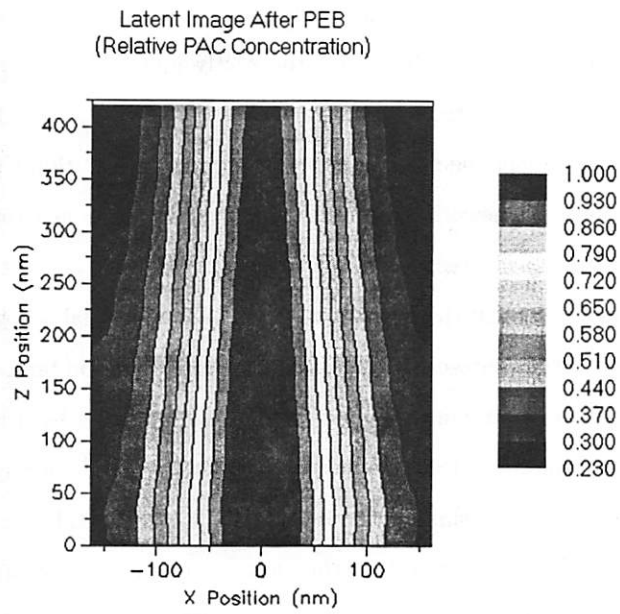


Figure 4.10: Sample result of latent image following PEB step from PROLITH.

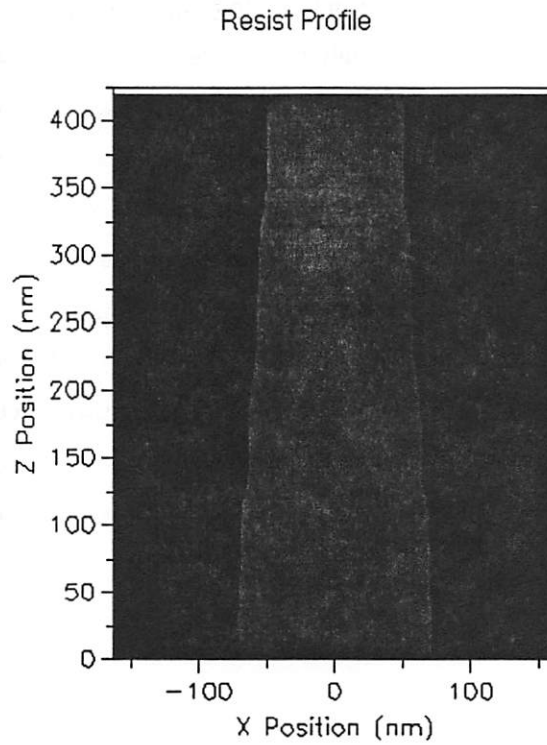


Figure 4.11: Sample resist profile result from PROLITH.

In the context of its usage in the proposed control framework, PROLITH does have a handful of weaknesses. The difficulty is in tuning PROLITH to exactly match the lithography process it is modeling. For process development, the slight differences between the actual lithography process and PROLITH can be swept under the rug, because an approximate solution is sufficient; however, in order for the process control framework to be successful, an exact (or an as-close-to-exact) solution is ideal. Prior research has demonstrated how PROLITH's flexibility can be used to achieve process-specific tuning of the simulator so that the software closely mimics an actual lithography module. For example, measurement of resist dose-to-clear as a function of PEB time and temperature can be used to calibrate the thermal model of amplification for a CA resist [28]. And by using contrast curves, swing curves, and focus-exposure matrices, simulation parameters can be automatically modified to minimize the difference between the simulated data and the experimental results [29, 30]. While the results of any of these efforts was imperfect, the authors are quick to explain that most of the shortcomings were due to PROLITH resources left untapped for simplicity. However, PROLITH does completely lack a model which can capture the PAB bake step. Although the user may enter ridiculous values for the PAB bake temperature and time (e.g., 500°C temperature), the resist profile will not change. However, the PAB step certainly has an effect on the process in reality (albeit a small one). Similarly, in PROLITH the photoresist is assumed to never age, nor vary from batch to batch. This is not a weak approximation in a high-volume fab where the resist never sits on a shelf for more than a few days, but could present a long-term drift in cases of a rarely used resist. These factors which are not captured by PROLITH must simply be treated as contributions to random noise, which cannot be controlled. Fortunately, these effects have been shown to affect the profile shape very slightly, and future versions of PROLITH may very well incorporate additional specifications if they are required. Therefore, we may safely assume that PROLITH can likely be sufficiently well tuned to match a lithography process; for the purposes of this work, all experiments are simulated, implicitly assuming that the lithography module can be accurately matched by the PROLITH simulation engine.

Chapter 5

Experiment and Results

This chapter contains the experimental setup and results of simulations carried out based on the framework proposed in Chapter 4.

5.1 Experimental Setup

As mentioned in Chapter 4, three libraries were created for experimental evaluation by simulation. All three libraries are based on the same nominal lithography process to pattern a standard polysilicon gate stack (consisting of SiO₂, poly-Si, bottom anti-reflective coating, and photoresist), which is summarized in Table 5.1:

Process Parameter	Fixed Nominal Value
Mask Linewidth	130nm
Mask Pitch	325nm (1:1.5 line:space)
Illumination Wavelength	248nm
Source Shape	Annular
Inner Sigma	0.495
Outer Sigma	0.85
<i>NA</i>	0.63
Resist Type	Shibley UV210
Resist Thickness	420 nm
BARC Type	Shibley AR3
BARC Thickness	60 nm
poly-Si Thickness	200 nm
SiO ₂ Thickness	3 nm
PAB Temperature	130°C
PAB Time	60 s
Developer	LDD-26W
Develop Time	45 s

Table 5.1: Experimental lithography process simulation setup: fixed nominal parameter values.

Four input parameter values were made variable—exposure dose, focus setting, PEB temperature, and PEB time. The range and step size of these values for the initial “seed” library are given in Table 5.2. These parameters are the parameters of interest (that is, the parameters to be estimated) in the state-estimation framework: The range of each parameter represents a ~ 15 nm range in

Process Parameter	Minimum Value	Maximum Value	Step Size
Exposure Dose	28 mJ/cm ²	30 mJ/cm ²	0.2 mJ/cm ²
Focus Offset	-0.5 μ m	0.5 μ m	0.1 μ m
PEB Temperature	128°C	132°C	1°C
PEB Time	57 s	63 s	2 s

Table 5.2: Range and step size in 2,420-entry “seed” library for variable processing parameters.

critical dimension, holding all other parameters constant; this corresponds to a total range of 40 nm between extreme corners of the library, with all four parameters variable.

Using the sensitivity measurement technique described in Chapter 4, we can investigate the nature of the initial seed library sensitivity balance. Using the central values of the variable process parameters as the “nominal” values (i.e., exposure dose = 29 mJ/cm², focus offset = 0 μ m, PEB temperature = 130°C, and PEB time = 61 s), we can vary each parameter across its full range while holding the other parameters at nominal values. Then, by measuring the SSD between consecutive steps in that parameter, we can extract a sensitivity cross-section through the center of the library for each processing parameter. These cross-sections are shown in Figure 5.1.

Clearly, the diffraction response sensitivity is variable within each processing parameter, and is also higher for focus and PEB characteristics compared to exposure dose. The statistics of the initial sensitivities are listed in Table 5.3: Although this sensitivity cross-section only illustrates the

Process Parameter	Mean Sensitivity	Standard Deviation of Sensitivity
Exposure Dose	0.807	0.328
Focus Offset	2.603	3.313
PEB Temperature	3.017	1.076
PEB Time	2.513	0.903

Table 5.3: Initial library sensitivity for each processing parameter.

sensitivity variability around one area of the library, it confirms the motivation for the iterative library-building procedure; for areas in the library where sensitivity is initially relatively high, ad-

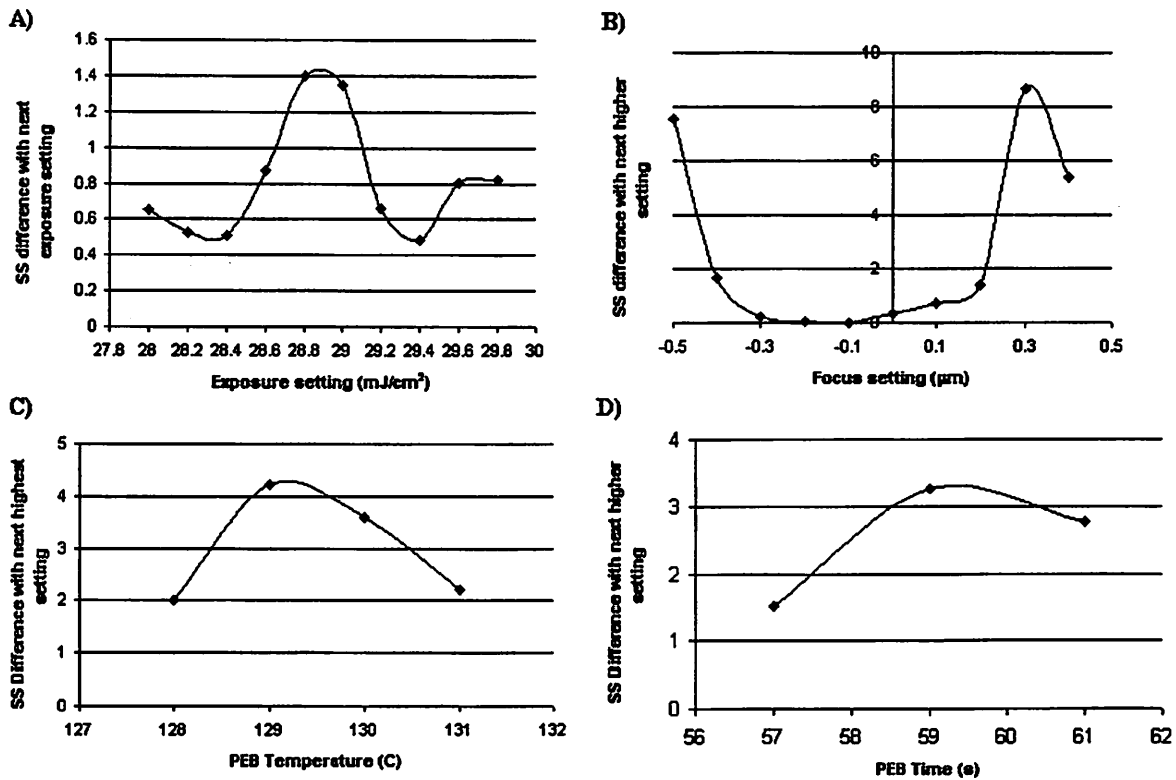


Figure 5.1: Initial library sensitivities vs. processing parameter values. A) exposure dose; B) focus setting; C) PEB temperature; D) PEB time.

ditional entries will be inserted. As a result, the overall diffraction response space sensitivity should become more balanced.

Using the 2,420-entry seed library, the library-building process was run for 9 iterations. The resulting improvement in sensitivity balance is captured in Figure 5.2. Here, the normalized mean group sensitivity (as defined in Section 4.2.2) and normalized standard deviation of all measured group sensitivities are charted against the iteration number. The ratio of these two, a normalized measure of sensitivity variability in the entire diffraction response space, clearly decreases. By iteration 9, the marginal improvement in sensitivity balance has decreased compared to the initial few iterations, as the worst sensitivity outliers have been reconciled.

To serve as a basis of comparison to this sensitivity-balanced library, a similarly large (27,783 entries) library was constructed using the same systematic input parameter space coverage as was used for the “seed” library, but with double the density in values of each processing parameter. The entries for this library are summarized in Table 5.4. The $\sigma/mean$ of group sensitivities measured in the

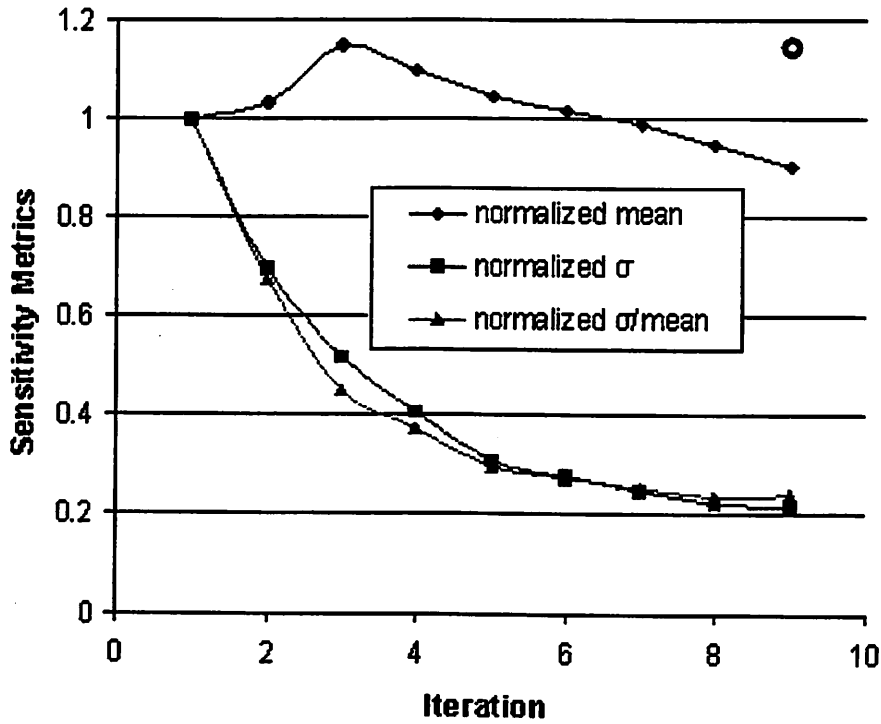


Figure 5.2: Sensitivity balance evolution of iteratively-built library. The large circle point at the ninth iteration represents the σ/mean of the sensitivities measured in the alternate large (27,783-entry) library.

alternate large (27,783-entry) library is shown as a comparison to the iterative result at iteration 9 in Figure 5.2. The sensitivity is more variable in this 27,783-entry library, as is expected since this library was built to evenly cover the input parameter space but without special awareness of the sensitivity balance of the diffraction response space.

Process Parameter	Minimum Value	Maximum Value	Step Size
Exposure Dose	28 mJ/cm^2	30 mJ/cm^2	0.1 mJ/cm^2
Focus Offset	-0.5 μm	0.5 μm	0.05 μm
PEB Temperature	128°C	132°C	0.5°C
PEB Time	57 s	63 s	1 s

Table 5.4: Range and step size in 27,783-entry library for variable processing parameters.

5.1.1 Evaluation Method

To evaluate the three libraries, 100 test measurements were generated. First, 100 combinations of values of the four processing parameters were drawn randomly from uniform distributions covering the full range of each processing parameter. Then, these 100 input parameter combinations were submitted to the PROLITH-based simulation framework, generating mock ellipsometer diffraction response measurements. No additional noise was introduced during the simulation of the 100 test measurements—some noise is incorporated due to the discretization of the profile, but in reality, noise will be introduced also by the ellipsometer measurement [31]. Here, though, the ellipsometer measurement noise is assumed to be less critical than the noise introduced by the profile discretization. Each of these 100 test point measurements is matched to the most similar library entry using the SSD method illustrated in Figure 4.4. This matching is repeated for varying conditions of additional supplied information; for example, if the PEB temperature is measured (e.g., by an inline temperature sensor) to within $\sim 1^\circ\text{C}$, then the search can be narrowed to consider only those entries whose PEB temperature falls within 1°C of the measured temperature. For each matching scenario, the resulting estimated parameter values (estimated exposure dose, estimated focus setting, and so on) are plotted against the actual test point parameter values. The success of the estimation will be measured by the R^2 value of a linear fit to the plotted data.

$$R^2 = 1 - SS_{error}/SS_{total}, \quad (5.1)$$

where SS_{error} is the error sum of squares (calculated by summing the squares of the differences between each actual value and the corresponding estimated value) and SS_{total} is the total sum of squares (calculated by summing the squares of the differences between each actual value and their average). If the estimation is accurate, the estimated parameter values will match the actual parameter values and the R^2 value of the fit will be 1. R^2 has a lower bound of 0, which represents the situation where the estimated and actual values are completely uncorrelated.

5.2 Results of the Small (2,420-entry) Seed Library

This section presents the evaluation of the state-estimation framework using only the 2,420-entry seed library. Each subsection contains the results under different matching scenarios.

5.2.1 Estimation Results with No Additional Information Supplied

As shown in Figure 5.3, when no additional approximate measurement information is supplied, the state estimation framework is extremely successful in estimating the focus ($R^2 = 0.9717$) but is unsuccessful in estimating the exposure dose value ($R^2 = 0.2254$) and PEB parameter values ($R^2 = 0.1150$ and 0.0060 , respectively). Examining the nature of the error in estimation of the PEB characteristics and the exposure dose leads to the conclusion that the estimation error in one is compensated by an opposite error in the other, as is illustrated in Figure 5.4. Clearly, if we provide an approximate PEB temperature measurement, the estimation should improve.

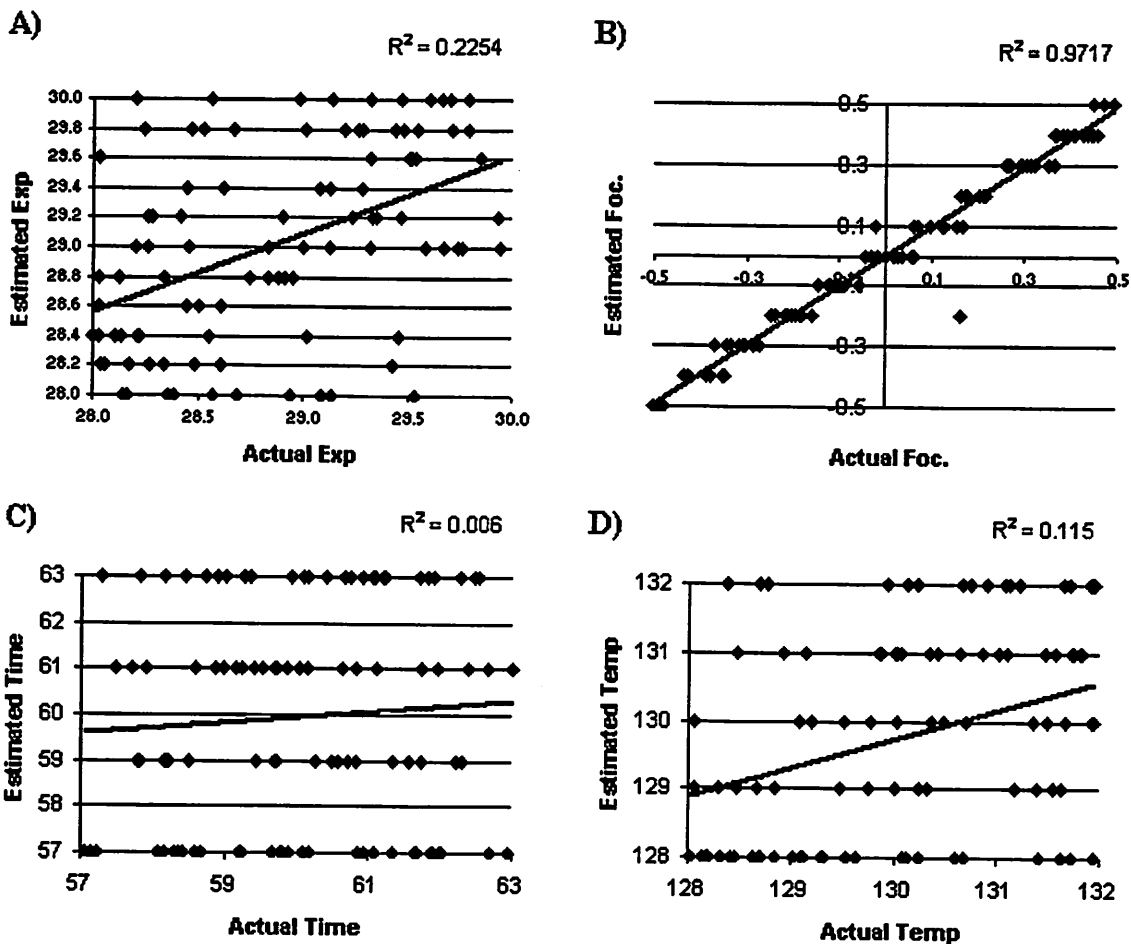


Figure 5.3: Parameter value estimation for 2,420-entry library. A) Exposure dose estimation; B) Focus estimation; C) PEB time estimation; D) PEB temperature estimation.

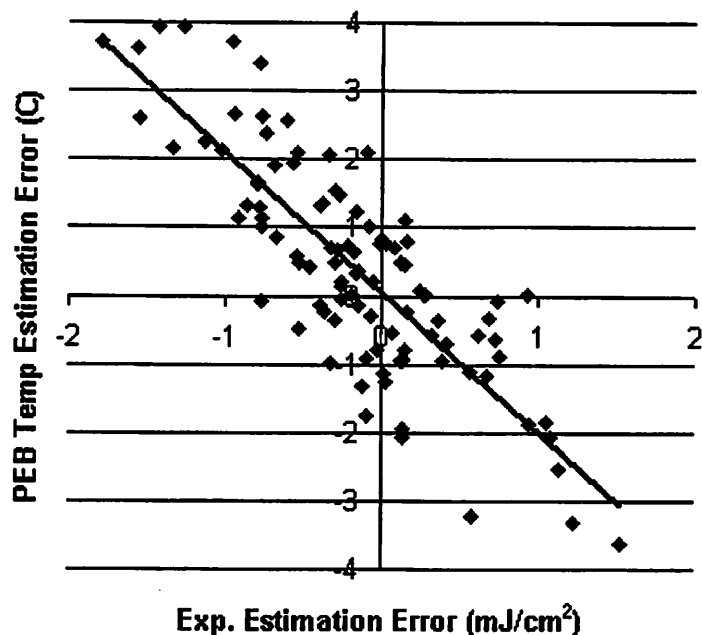


Figure 5.4: PEB temperature estimation error vs. exposure estimation error.

5.2.2 Estimation Results for 2,420-entry Library with Approximate PEB Temperature Supplied

By limiting the search scope to library entries whose PEB temperature value differs from the PEB temperature value of the test point by 1°C or less (supplying a mock approximate PEB temperature measurement), the exposure dose estimation is improved somewhat ($R^2 = 0.6854$) and the PEB temperature estimation becomes excellent ($R^2 = 0.9585$). The PEB time estimation is still poor ($R^2 = 0.0037$)—but it is now evident that the PEB time estimation error is correlated with the exposure dose estimation error now that the PEB temperature has been “pinned” to a value that is correct to within 1°C. This implies that the remaining exposure dose estimation error is now being compensated by the PEB time estimation error. Therefore, if we provide the approximate PEB time in addition to the PEB temperature (as is presented in Section 5.2.4), we should be able to improve the exposure dose estimation once more.

In addition, the focus estimation worsens significantly ($R^2 = 0.6161$); however, it appears that the test points whose focus setting is estimated incorrectly frequently have the right magnitude focus offset estimation, but wrong sign. This is not completely surprising, since there are two focus

solutions for a given printed CD, as can be noticed from standard Bossung curves relating CD to exposure dose and focus. Apparently, as the other parameters are estimated more successfully, the alternate focus “solution” for a given resist CD is occasionally selected as the best match, since the corresponding slight adjustment to the profile (since the two alternate “solutions” do not give rise to identical profiles) gives rise to a more closely matching diffraction response.

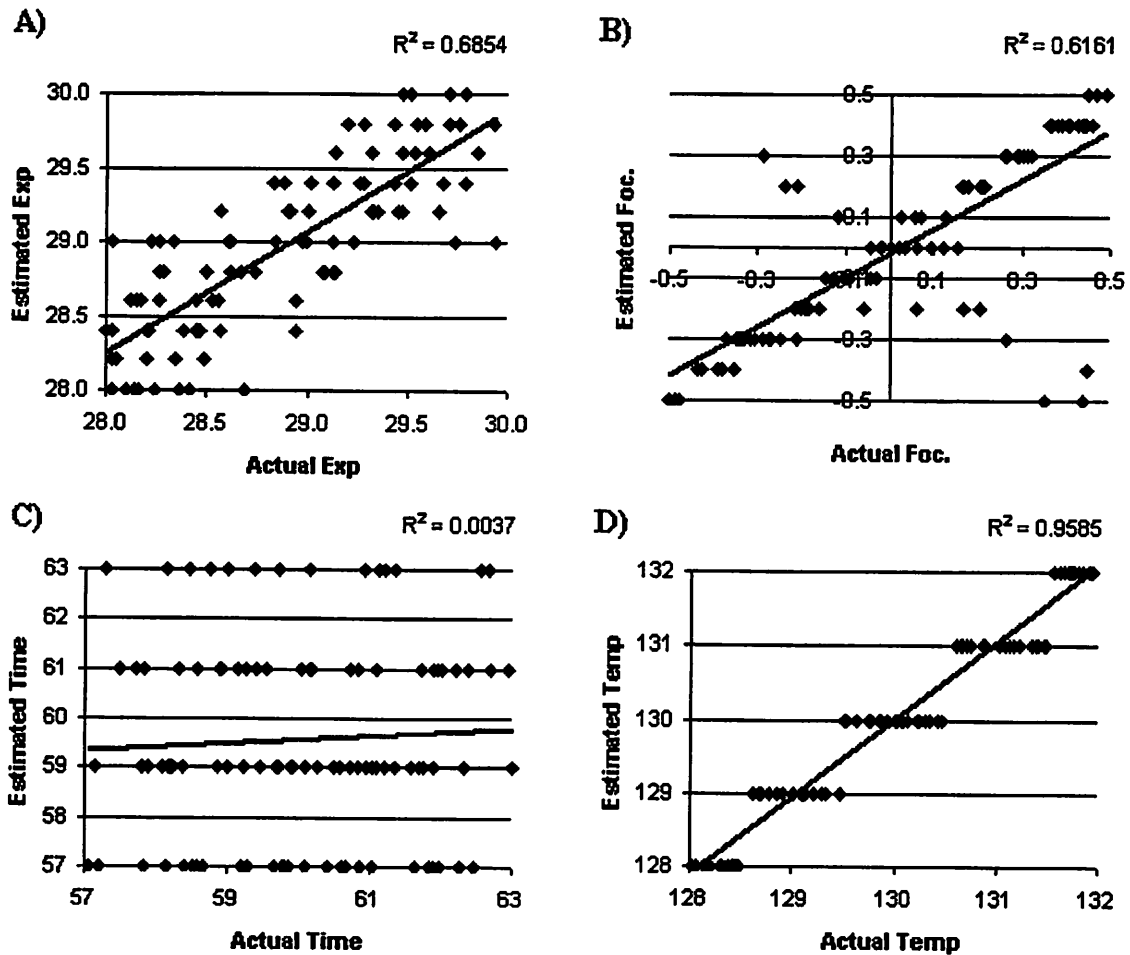


Figure 5.5: Parameter value estimation for 2,420-entry library with approximate PEB temperature supplied to narrow search scope. A) Exposure dose estimation; B) Focus estimation; C) PEB time estimation; D) PEB temperature estimation.

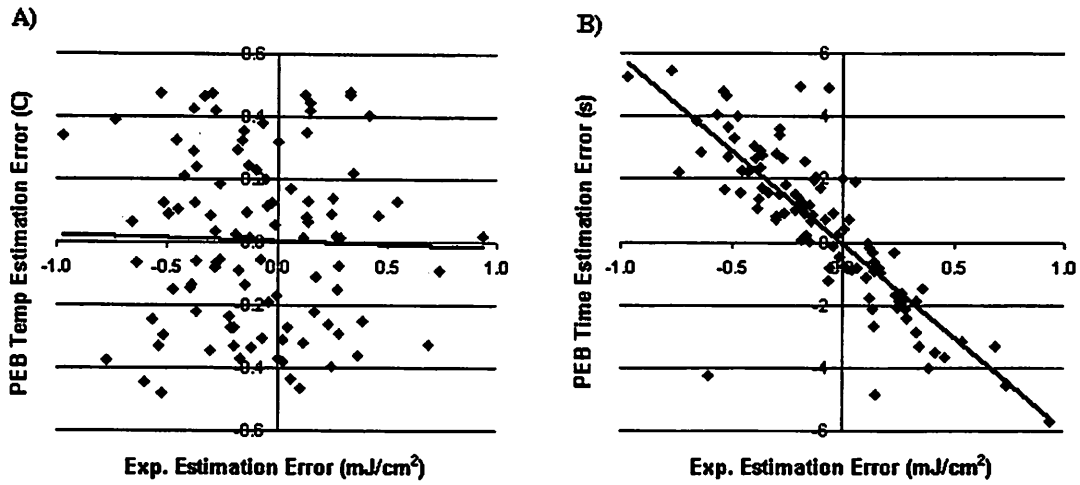


Figure 5.6: Estimation error correlations for 2,420-entry library with approximate PEB temperature supplied: A) PEB temperature estimation error vs. exposure estimation error; B) PEB time estimation error vs. exposure estimation error.

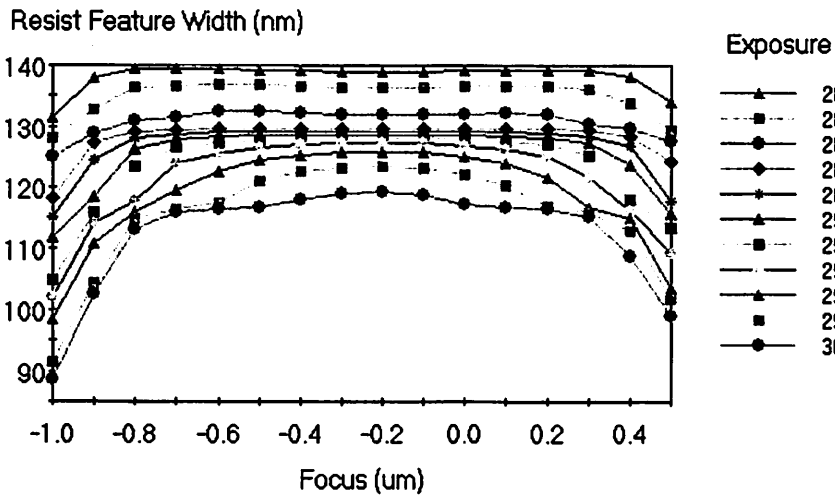


Figure 5.7: Bossung curves for this lithography process. For a given exposure dose setting, two focus settings yield the same CD.

5.2.3 Estimation Results for 2,420-entry Library with Approximate PEB Time Supplied

Instead of providing the approximate PEB temperature, external sensors might provide an approximate measurement of the duration of the PEB step. (The exact PEB time is not supplied because actual PEB trajectories can have an “effective” duration at fixed temperature, and this effective time is not exactly known.) By supplying the approximate PEB time (to within two seconds of the actual time) the exposure dose estimation is identical to that achieved with approximated PEB temperature supplied ($R^2 = 0.6854$) and the PEB time estimation reaches its maximum achievable value ($R^2 = 0.8637$) with the given library density—that is, every test point finds the closest available PEB time estimation value. The PEB temperature estimation is now still poor ($R^2 = 0.1573$), and the correlation between PEB temperature error and exposure dose estimation error remains. Finally, the focus estimation suffers considerably ($R^2 = 0.3539$). However, it is again observed that the noticeable focus estimation errors are only errors in sign of focus offset—the magnitude is usually correctly estimated, due to the dual-solution nature of the Bossung curve.

5.2.4 Estimation Results for 2,420-entry Library with Approximate PEB Temperature and Time Supplied

By supplying the approximate PEB time (to within 2 seconds) in addition to the approximate PEB temperature, the search scope is further limited to library entries whose PEB time value differs from the PEB time value of the test point by 2 seconds or less. As expected, the exposure dose estimation improves dramatically ($R^2 = 0.9129$), and both the PEB temperature and time are well estimated ($R^2 = 0.9585$ and 0.8637 , respectively). The focus estimation becomes quite poor in terms of the R^2 value ($R^2 = 0.3573$), but again the magnitude appears to be estimated quite well, with the sign error due to the Bossung curve dual-solution phenomenon. If the estimated focus is also supplied (from the matching performed with no additional information provided), the exposure estimation R^2 improves further to 0.9215 (Fig. 5.9).

5.2.5 Summary of Evaluation for the Seed Library

Under varying conditions of additional supplied information, all four processing parameters can be adequately estimated. Focus setting is best estimated with no additional information supplied; expo-

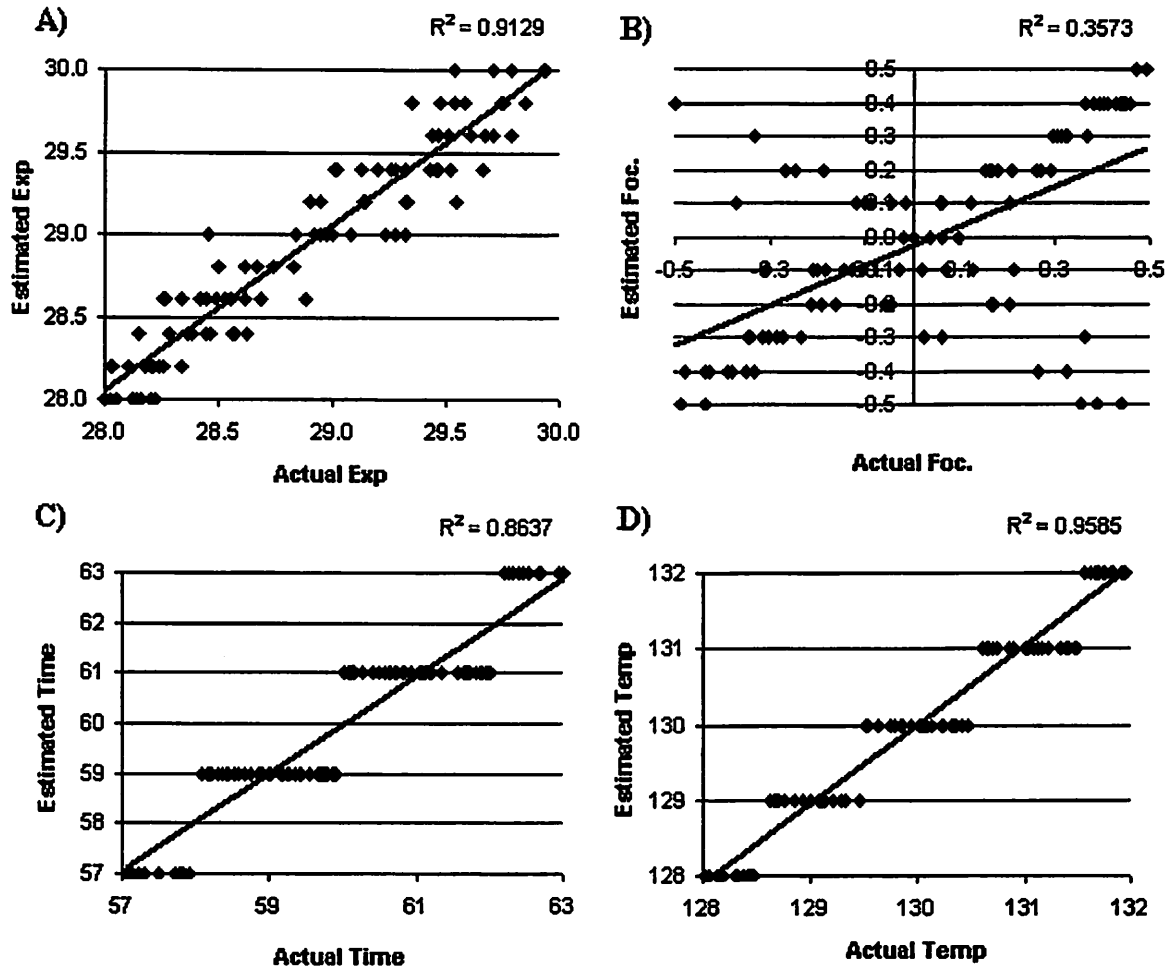


Figure 5.8: Parameter value estimation for 2,420-entry library with approximate PEB temperature and time supplied to narrow search scope. A) Exposure dose estimation; B) Focus estimation; C) PEB time estimation; D) PEB temperature estimation.

sure dose as well as PEB parameters are best estimated when the estimated focus and approximate PEB parameter values are supplied to fix the correct focus estimation and limit the compensating relationship between the exposure step and post-exposure bake step on resulting linewidth.

5.3 Results of the Iteratively-Built Library

Upon evaluation, it is clear that the iteratively-built library offers minimal improvement in estimation success over the seed library. As listed in Tables 5.5, 5.6, and 5.7, the R^2 values for all parameters

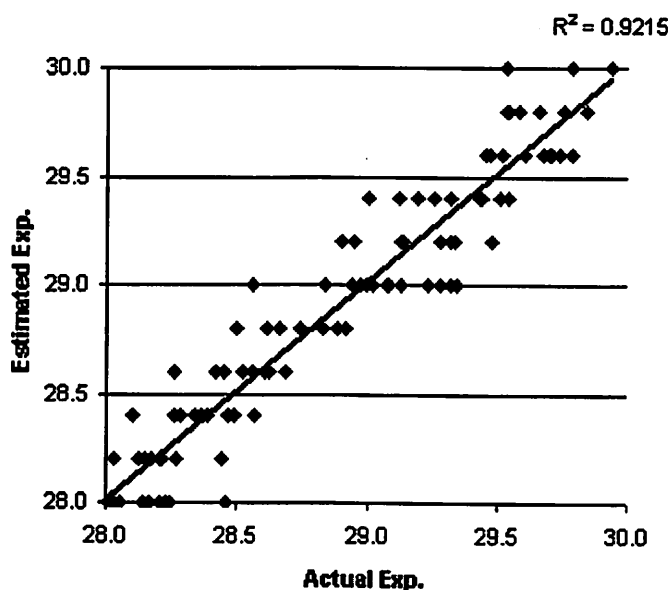


Figure 5.9: Exposure dose estimation for 2,420-entry library, with approximate PEB temperature and time and estimated focus value provided.

being estimated under each of the various conditions of supplied approximate PEB measurement values have virtually identical values. Inspection of the iterative library reveals that all of the additional $26,991 - 2,420 = 24,571$ entries were placed into a small segment of the library at the high ends of the focus and exposure dose parameter ranges, where the sensitivity is initially extremely high. Thus, the majority of the library sees very little if any increase in number of entries—and therefore most of the test points “fall” in regions of the new library that have virtually the same density as the corresponding regions of the original 2,420-entry library. As a result, most of the matches for the iteratively-built library are identical to those of the seed library; only the very few number of test points which happen to fall in the high-exposure dose, high-focus setting range stand a chance of being matched more accurately. Therefore, although this library construction technique optimizes the use of the diffraction response space’s inherently variable sensitivity, it is not useful for increasing the accuracy of parameter value estimation.

Process Parameter	Seed Library R^2	Iteratively-Built Library R^2
Exposure Dose	0.2254	0.2325
Focus Offset	0.9717	0.9739
PEB Temperature	0.1150	0.0730
PEB Time	0.0060	0.0164

Table 5.5: Comparison of results from seed and iteratively-built libraries, with no additional measurement information supplied.

Process Parameter	Seed Library R^2	Iteratively-Built Library R^2
Exposure Dose	0.6854	0.6793
Focus Offset	0.6161	0.7575
PEB Temperature	0.9585	0.9566
PEB Time	0.0037	0.0068

Table 5.6: Comparison of results from seed and iteratively-built libraries, with approximate PEB temperature supplied.

5.4 Results of the Alternate Large (27,783-entry) Library

In the case of the 27,783-entry library, the increased density is spread evenly over the input parameter space. Therefore, every test point falls in a region with higher density compared to the 2,420-entry library. As a result, we might expect the estimation accuracy to improve. Indeed, using this library, all estimations show significant improvement. As shown in Fig. 5.10, the R^2 values of fits for both the exposure and focus estimation increase for the 27,783-entry library compared to the 2,420-entry library—from 0.2254 to 0.5195 for exposure and from 0.9717 to 0.9955 for focus. The PEB temperature and time estimation are still poor. The same compensating relationships exist between exposure and PEB characteristics; thus, when both the approximate PEB temperature and time are provided, much higher R^2 values are achieved for exposure (0.9646), PEB temperature (0.9868), and PEB time (0.9733) estimations. The same Bossung-curve phenomenon is seen in the focus estimation, but to a lesser degree than in the smaller library (the R^2 for the focus estimation fit decreases to

Process Parameter	Seed Library R^2	Iteratively-Built Library R^2
Exposure Dose	0.9129	0.9196
Focus Offset	0.3573	0.3429
PEB Temperature	0.9585	0.9586
PEB Time	0.8637	0.8545

Table 5.7: Comparison of results from seed and iteratively-built libraries, with approximate PEB temperature and time supplied.

0.6958). Finally, if we supply the estimated focus setting (from the match performed with no additional information) in addition to the approximate PEB characteristics, the exposure estimate reaches a maximum R^2 value of 0.9838 (Fig. 5.12). Therefore, we again find that the focus estimation is most successful when no additional information is supplied; exposure and PEB characteristics are better estimated when approximate PEB characteristic measurements and estimated focus are supplied to the framework.

Process Parameter	R^2 using 2,420-entry Library	R^2 using 27,783-entry Library
Exposure Dose	0.9215	0.9838
Focus Offset	0.9717	0.9955
PEB Temperature	0.9585	0.9868
PEB Time	0.8637	0.9733

Table 5.8: Comparison of best estimation results from the 2,420-entry and 27,783-entry libraries. For focus estimation, no additional information is provided; for exposure and PEB estimations, the focus estimate and approximate PEB measurements are provided.

5.5 Analysis and Conclusions

Successful estimation of each input parameter is possible using this state-estimation method. Successful focus estimation is best achieved with no additional information supplied to the framework; however, under this condition, estimation of exposure dose and PEB characteristics is poor due to the fact that these parameters have a very strongly correlated, compensating relationship. In fact, it is likely that this relationship is exploited in favor of correct focus setting estimation. Consider a given test point; there exists a “best” matching library entry for which the error between each library parameter value and the corresponding actual parameter value is minimized. However, this solution frequently will not be found, because the minimal error in each parameter will contribute to a potentially large SSD in diffraction responses. Instead, the estimated values of the parameters with compensating relationships will likely drift from the “best” values to alternate values, where the impact of each parameter’s drift on the profile shape will *nearly* cancel but will still create some slight overall adjustment to the profile, thereby tuning the resulting SSD lower. Since the exposure dose and PEB characteristics have the strongest compensatory relationship, their estimation will therefore suffer, while the focus estimation will be accurate. When the approximate PEB characteristics are supplied, estimation of the exposure dose and the PEB characteristics improves, because

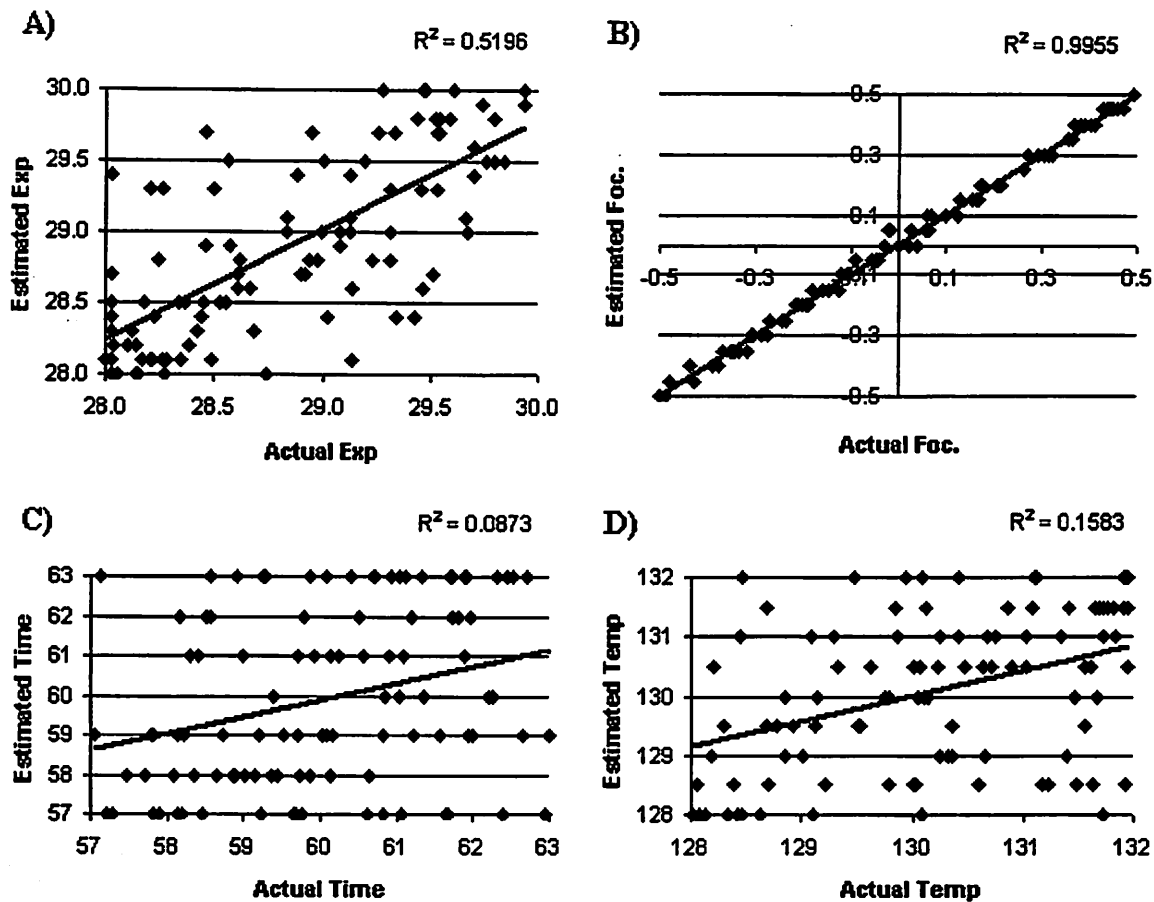


Figure 5.10: Parameter value estimation for 27,783-entry library. A) Exposure dose estimation; B) Focus estimation; C) PEB time estimation; D) PEB temperature estimation.

the compensating relationship is limited to a narrowed possible range of drift. That is, since the PEB values cannot be estimated outside of a certain range, the range of exposure estimation is necessarily limited to a narrower range as well. The upshot of limiting the tuning afforded by the compensation of exposure dose by PEB characteristics is that the focus estimation becomes less accurate. The nature of the focus inaccuracy also is not surprising—in cases where the focus is incorrectly estimated, the estimated value usually has nearly correct magnitude but incorrect sign, reflecting the well-known relationship of similar CD for equal magnitude and opposite sign focus settings shown in Bossung curves. Thus, to best estimate all of these parameters, two separate matches should be performed; the first should estimate the focus, with no additional supplied information, and the second should provide this estimated focus, as well as the approximate PEB characteristics, to best

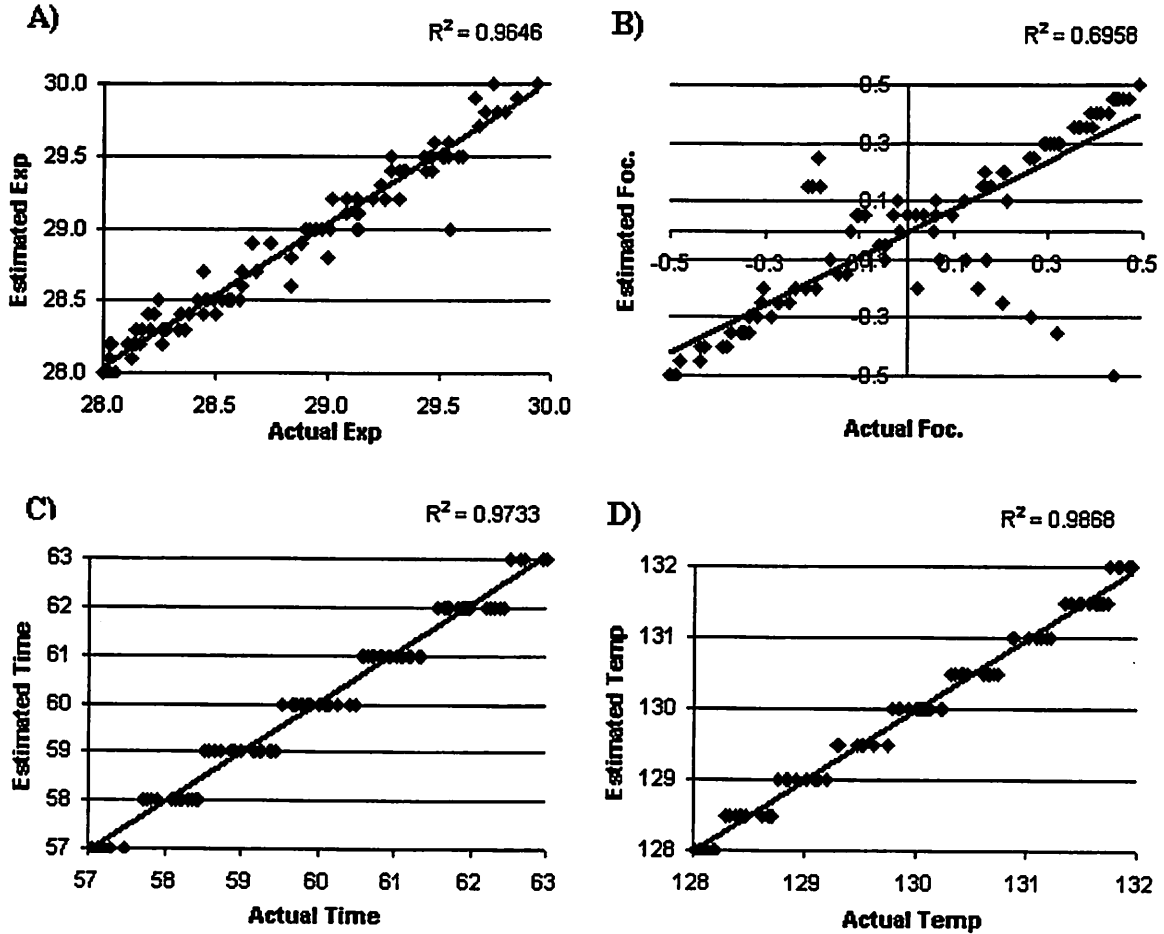


Figure 5.11: Parameter value estimation for 27,783-entry library with approximate PEB temperature and time supplied to narrow search scope. A) Exposure dose estimation; B) Focus estimation; C) PEB time estimation; D) PEB temperature estimation.

estimate the exposure dose and PEB characteristics.

This state estimation technique promises to extend relatively smoothly to future generations of photolithography. Mainly, the success of scatterometry-based state estimation relies on the accuracy of optical measurement. As mentioned in Chapter 2, projections for the adequacy of current scatterometry techniques are mixed, but standard scatterometry is expected to provide sufficient accuracy at least through the 65 nm node. Also, with the introduction of 193 nm illumination, new CA resists have been developed which display some critical differences with 248 nm resists. In particular, the sensitivity to PEB temperature increases—it has been observed that a 1°C variation in temperature of the resist during the PEB step can result in a 10 nm variation in resulting line

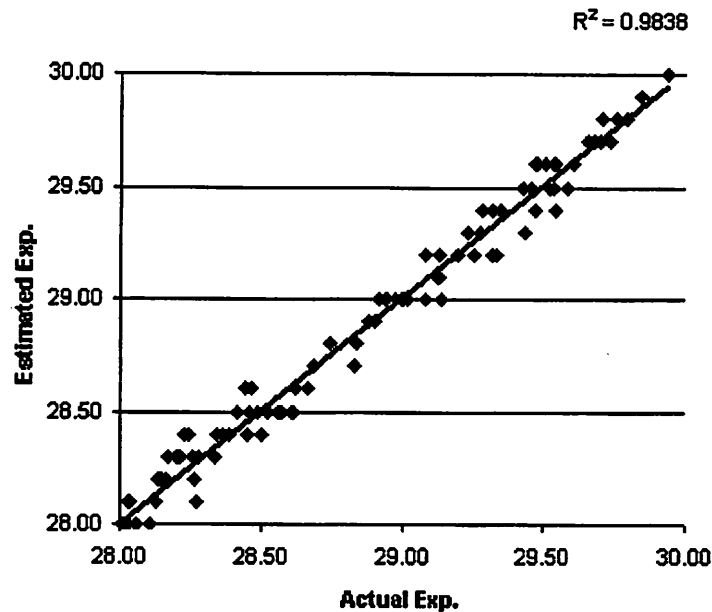


Figure 5.12: Exposure dose estimation for 27,783-entry library, with approximate PEB temperature and time and estimated focus value provided.

width [32]. Thus, an equal uncertainty in the provided PEB temperature measurement (as required for accurate estimation of exposure dose as well as PEB characteristics) will cover a wider range in CD for state estimation at the 193 nm node than at the 248 nm node. Although current temperature sensors are capable of 0.1°C accuracy—which corresponds to $\sim 1\text{nm}$ accuracy in terms of effect on resulting CD for worst-case scenario at 193 nm—this accuracy may need to be improved to meet future, more stringent CD control requirements if this state estimation framework is to be used.

In terms of cost of implementation, the size and nature of the library is the key factor. Good estimation accuracy can be achieved for an extremely small library (here, 2,420 entries). Improved estimation can be achieved if the library density is increased evenly in the input parameter space. (If the density increase is aimed toward improving the sensitivity balance in the diffraction response space, the estimation accuracy may not be improved—as was the case in this work—because the added entries may cluster in relatively narrow ranges of the library, offering little potential improvement in estimation of most parameter value combinations.) However, the accuracy improvement over the 2,420-entry library afforded by the 27,783-entry library was marginal, despite the ten-fold increase in library size—this estimation accuracy improvement comes at the considerable expense of increased off-line simulation time required for the library. Using an industrial-quality scatterometry

simulation engine as the standard, the time required to simulate a single diffraction response (the most time-consuming component of the library building flow) would be roughly 0.2 second. Thus, a library of 2,420 entries could be created in roughly 10 minutes, and a library of 27,783 entries could be created in less than 2 hours. The result that fairly accurate estimation can be achieved for such small libraries (and short simulation times) is an auspicious sign. Compared to standard scatterometry libraries of ~500,000 profiles that take ~1 day to create, the libraries successfully used in this work are quite small, confirming the hypothesis that a control-based use of scatterometry would not call for a typical library of such large size. In practice, specific (i.e., 3σ) requirements of CD control will dictate the required accuracy in parameter value estimates, which in turn will dictate the lower bound on the library size. A process engineer might then decide which library size (necessarily exceeding this lower bound) to use based on the tradeoff between estimation accuracy and required simulation time.

Chapter 6

Conclusions

A first principle-based lithography process state estimation framework has been proposed and implemented in simulation. The framework, as evaluated in simulation, demonstrates accurate estimation of exposure dose, focus setting, and PEB temperature and time. It was found that the library should be evenly distributed throughout the input parameter space to gain the best estimation, and that the accuracy of estimation improves as the size of the library increases. In this work, libraries of sizes one to two orders of magnitude smaller than traditional scatterometry libraries were found to be successful. This suggests that the library construction framework used here eliminates unnecessary profiles that are considered in typical scatterometry applications, since a process must be well-behaved in the first place to be a control candidate, and therefore requires a smaller range of profiles to be covered in the library. Future work on this project will include an empirical implementation of the state estimation framework. Topics of interest will be the quality of matching between the PROLITH simulator and lithography module, effects of measurement-induced noise, and use of embedded temperature sensors to gather real-time PEB measurements to supply in the matching step.

Chapter 7

Acknowledgements

I would like to first thank my advisor, Professor Costas Spanos, for his support and guidance over the past two years. Additional thanks go to Professor Kameshwar Poolla for comments and suggestions as second reader of this report.

All scatterometry-based simulation in this work was carried out using the grating tool kit developed by Junwei Bao, Nickhil Jakatdar, and Xinhui Niu. Many thanks Junwei for invaluable assistance with GTK and other scatterometry-related issues. I'd like to specially thank Jason Cain for offering sage advice in regard to GTK, Matlab, LaTeX, and other miscellaneous software, as well as for his general wealth of knowledge of lithography. In addition, I thank the other members of the BCAM group for their friendship and support: Runzi Chang, Michiel Kruger, Jiangxin Wang, Jing Xue, Charlie Zhang, Haolin Zhang, Dongwu Zhao.

Thanks also to friends from the UCB Microfabrication Laboratory, who provided Shipley material documentation as well as training on lithography equipment: Sia Parsa, Kim Chan, Lazlo Voros.

This work was supported by Advanced Energy, Advanced Micro Devices, Applied Materials, ASML, Atmel, Cymer, EBARA, EVG, Intel, KLA-TENCOR, Mycrolis, Nikon Research, Novellus Systems, Numerical Technologies, Tokyo Electron America, and the UC Discovery Grant.

Appendix A

Sources of Circuit Performance Variability

A.1 Circuit Variability Study

As a class project for EE244 (Introduction to Computer Aided Design) the effects of device parameter variability on circuit performance was investigated. Projected device parameter values and levels of variation were taken from the 1999 and 2001 ITRS roadmaps and other literature for several device generations. A canonical circuit was created with one segment (a simple inverter chain) designed to model a local datapath and another segment (a sized inverter chain driving a large capacitive load) to represent the a global interconnect and the buffer used to drive its large capacitance. This circuit skeleton was submitted to Monte Carlo analysis, wherein for each trial, device parameters were drawn from normal distributions described by the corresponding nominal value and level of variation. Then each fully-specified circuit was simulated in HSPICE, and the two components (local, global) of the resulting propagation delay were measured. For each of several device generations, 1000 such trials were carried out. As expected, the component of delay due to interconnect was increasing and dominated the total propagation delay, as shown in Fig. A.1. However, a simple ANOVA analysis of the data yielded a measure of timing variation broken down into the component of device parameter variation; this variability of total delay was dominated by gate length variation (Fig. A.2).

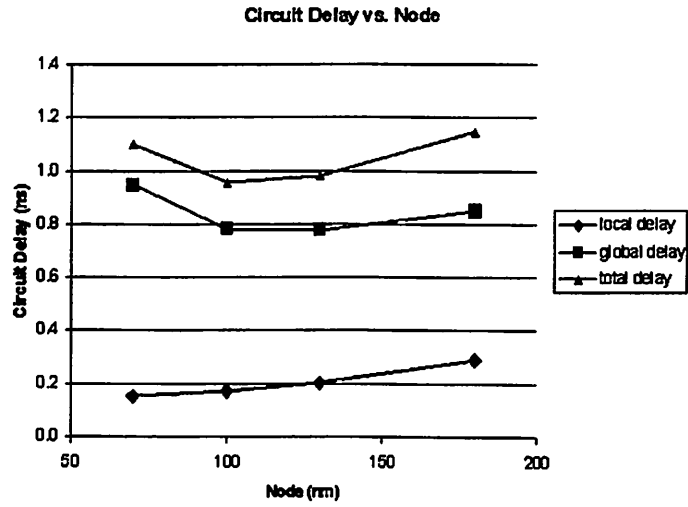


Figure A.1: Circuit propagation delay breakdown for successive technology nodes.

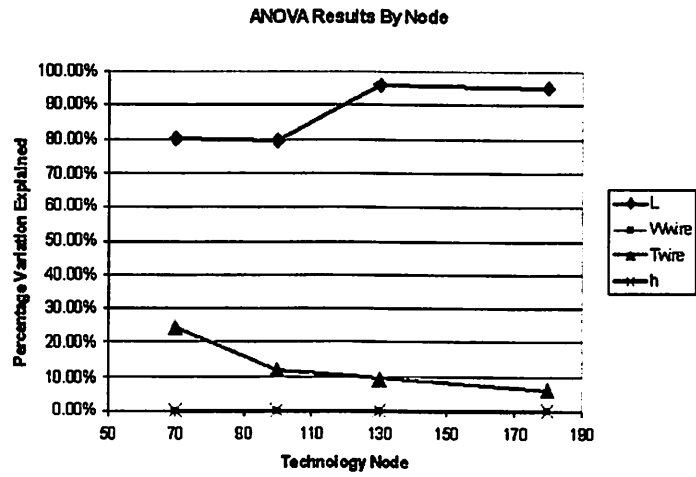


Figure A.2: ANOVA breakdown of delay variation for various device parameters at successive technology nodes.

Thus, it is found that as device size shrinks, the effective load presented by the interconnect and other fanout is increasing in proportion to the transistor drive strength. Therefore, variability in global propagation delay, which in turn most strongly depends on variability in the gate length of the transistor, is becoming an increasingly large problem. Clearly, in order to maintain sufficient circuit yield while achieving the performance improvement that the industry has come to promise its customers, tight CD control will be a necessity in semiconductor manufacturing.

Appendix B

Software Framework

The simulation framework described in Chapter 4 is implemented using a core Matlab script that calls three subroutine segments using three respective pieces of software: Matlab itself, Prolith, and the grating tool kit (GTK). In this appendix, each segment of the software will be described in the order prescribed by an iteration of the iterative library construction method. The library construction method used for the non-iteratively built libraries is essentially captured by a single iteration.

B.1 Matlab—Core of the Software

The software framework is governed by a Matlab script. This script creates the library seed, which is comprised of a specified number of combinations of input parameter values. These combinations are generated using basic nested for-loops, each of which delineates a range and step size for a given parameter. The core script then enters its iterative mode. For each iteration, the script calls external subroutines to simulate the resulting profiles, then digitize those profiles, and simulate the corresponding diffraction responses. Finally, the core script examines the diffraction response space to evaluate sensitivity imbalance, and chooses to add entries (further combinations of input parameter values) to the library such that the imbalance is reduced. This loop repeats for a specified number of iterations.

In non-iterative mode, the routine simply lists all of the entries to be constructed in the first step, and calls external subroutines to simulate the resulting profiles, create digital representations, and

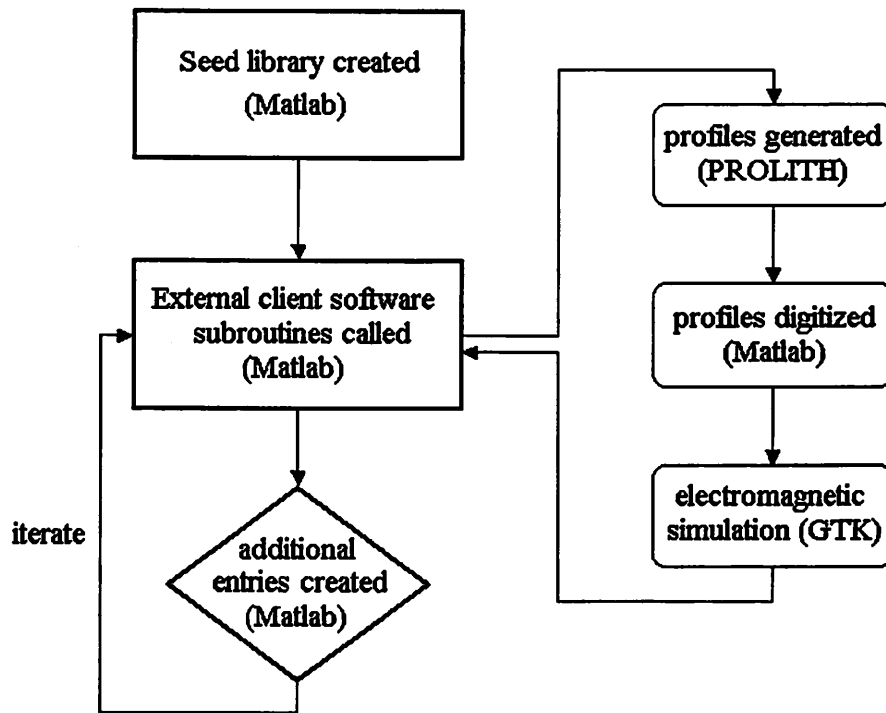


Figure B.1: Flow-diagram of the software framework.

the corresponding diffraction responses. No examination of the diffraction response sensitivity is required.

B.2 PROLITH Simulation Engine

Since PROLITH cannot be used to efficiently extract large batches of profiles in stand-alone mode, it is called remotely through a Microsoft Excel/Visual Basic-governed shell. This shell loads a file saved in the core Matlab script which contains a list of input parameter combinations for which to generate profiles. The shell then connects to PROLITH and passes these combinations to the PROLITH engine one at a time, saving each output profile to disk. In addition to specifying the input parameter values, this shell also describes the setup of the entire lithography module: the mask pattern, substrate film stack, illumination source characteristics, resist type, and development model to be used are all contained. The heart of this Excel/VB code is shown below:

```
Public Sub ProfileGenerator()
```

```

'Connecting to/Starting PROLITH
    ConnectProlith

'Declare constants
    Dim res_thickness As Single
    Dim PEB_time As Single
    Dim PEB_temp As Single
    Dim PAB_time As Single
    Dim PAB_temp As Single
    Dim dev_time As Single
    Dim exp As Single
    Dim foc As Single
    Dim th_2 As Single
    Dim th_3 As Single
    Dim th_4 As Single

' Set up grating mask pattern
    SetMaskBinary Mask_Binary_Line
    SetInputParameter Input_Mask_Binary_Width_Constant_Pitch, 130
    SetInputParameter Input_Mask_Binary_Pitch, 325
    SetInputParameter Input_Mask_Binary_Bias, 0

' Set up the illumination source
    SetSourceShape Stepper_Source_Annular
    SetInputParameter Input_Image_Annular_Inner, 0.495
    SetInputParameter Input_Image_Annular_Outer, 0.85
    SetInputParameter Input_Image_NA, 0.63
    SetInputParameter Input_Image_WaveLength, 248

' Define the photoresist and substrate stack materials

```

```

AddInputFile "E:\Paul\pfriedbe\proolith_stuff\UV210.res"
AddMaterialFile "E:\Paul\pfriedbe\proolith_stuff\AR3Shipley.mat", 2
AddMaterialFile "E:\Paul\pfriedbe\proolith_stuff\a-Polysilicon.mat", 3
AddMaterialFile "E:\Paul\pfriedbe\proolith_stuff\Si_Dioxide.mat", 4

' Set the constant input values
  res_thickness = 420
  th_2 = 60
  th_3 = 200
  th_4 = 3
  PAB_time = 60
  PAB_temp = 130
  dev_time = 45

' Loop which executes simulations
  i = 1
  Do

    i = i + 1

  ' Load each set of input parameter values
    exp = Range("e" + Format(i)).Value2
    foc = Range("f" + Format(i)).Value2
    PEBtime = Range("g" + Format(i)).Value2
    PEBtemp = Range("h" + Format(i)).Value2
    SetInputParameter Material_Layer_2_Thickness, th_2
    SetInputParameter Material_Layer_3_Thickness, th_3
    SetInputParameter Material_Layer_4_Thickness, th_4
    SetInputParameter Input_Prebake_Time, PAB_time
    SetInputParameter Input_Prebake_Temperature, PAB_temp
    SetInputParameter Input_PEB_Time, PEB_time
    SetInputParameter Input_PEB_Temperature, PEB_temp

```

```

SetInputParameter Input_Development_Time, dev_time
SetInputParameter Input_Resist_Thickness, res_thickness
SetInputParameter Input_Focus, foc
SetInputParameter Input_Exposure, exp

'Run the simulation and save the profile data to disk
RunSingleSim
OutputResistProfileData "E:\Paul\130nm\profiles\th" + Format(thickness)
                        + "_PABtime" + Format(PAB_time) + ... + ".txt"

Loop Until (Range("e" + Format(i)).Value2 = 0)

'4. Disconnect from PROLITH
DisconnectProlith
End Sub

```

B.3 Profile Digitization

In the next step, the profiles are digitized using a brief Matlab script. The script scans through the (x,y)-data describing each profile one time to select the subset of that data which describes a single sidewall (the top of the resist profile is ignored once a flat surface is detected). The selected data is scanned through again to determine the number of points retained, n . Then, based on the number of rectangles m desired for the digitized profile, the script averages the x-coordinates of the first $\frac{n}{m}$ points to determine the first rectangle half-width, and records the m^{th} y-coordinate value of the last of those selected points to specify the rectangle height (subtracting out the height recorded for the previous rectangles created each time). This procedure is repeated along the entire sidewall, creating the n rectangles which will best fill the given profile, thus yielding a digitized representation of the profile.

This step introduces some level of noise into the simulation framework, since avergaing over several consecutive points along a profile's sidewall may "miss" some profile features. Increasing the number of rectangles n will mitigate the introduction of this inaccuracy, but it will also increase the time

required for simulation of the electromagnetic response. Thus, there is a tradeoff between speed and accuracy at this level of the software framework. It has been noticed that ~ 20 rectangles is sufficient to accurately represent a profile with moderate standing wave features in its sidewalls.

B.4 Scatterometry Library Simulation

The grating tool-kit (GTK), developed by Niu et al [3], is the principal software component for construction of a diffraction response library for scatterometry. GTK employs rigorous coupled-wave analysis (RCWA), a state-variable method for determining a numerical representation of the diffraction response. The RCWA method basically creates expressions for the electric fields above the grating, within the grating, and in the substrate; then through iterative calculation, the boundary conditions of the three fields are matched, and the corresponding diffraction response can be calculated. GTK implements both the one-dimensional TE (transverse electric polarization) and TM (transverse magnetic polarization) diffraction cases. The accuracy of the simulation depends on the number of diffraction orders retained in the electromagnetic simulation for each type of polarization, with more retained orders giving rise to more accurate computation. However, the computation time increases exponentially in the number of orders retained, so a balance between accuracy and simulation time is required here as well as with the profile digitization step. Generally, and in this work, roughly 10 retained orders yields sufficient accuracy without presenting unacceptable simulation times. In any case, the flexibility of the software (in the tradeoff between accuracy and simulation time) is an exceptional feature and requires additional exploration in the venue of applying scatterometry to process control, where accuracy requirements may change over time.

Bibliography

- [1] C. Gould, "Advanced Process Control: Basic Functionality Requirements for Lithography," in *12th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference, Proc. IEEE*, pp. 49–53, 2001.
- [2] A. Zeidler, K. J. Veenstra, and T. E. Zavecz, "Advanced statistical process control: Controlling sub-0.18 μ m Lithography and other processes," in *Metrology, Inspection, and Process Control for Microlithography XV*, N. T. Sullivan, ed., *Proc. SPIE 4344*, pp. 312–322, 2001.
- [3] X. Niu, N. Jakatdar, J. Bao, and C. J. Spanos, "Specular Spectroscopic Scatterometry," *IEEE Trans. Semiconduct. Manufact.* **14**, pp. 97–111, May 2001.
- [4] C. A. Mack, "PROLITH: a comprehensive optical lithography model," in *Optical Microlithography IV*, H. L. Stover, ed., *Proc. SPIE 538*, pp. 207–220, 1985.
- [5] G. E. Moore, "Lithography and the Future of Moore's Law," in *Optical/Laser Microlithography VIII*, T. A. Brunner, ed., *Proc. SPIE 2440*, pp. 2–17, 1995.
- [6] "International Technology Roadmap for Semiconductors: 2002 Update." <http://public.itrs.net/Files/2002Update/Home.pdf>.
- [7] "Intel Lithography Roadmap." http://www.intel.com/technology/itj/2002/volume06issue02/art06_lithogra/phyroadmap/p03_roadmap.htm.
- [8] "ASML." <http://www.asml.com>.
- [9] M. Switkes and M. Rothschild, "Resolution Enhancement of 157-nm Lithography by Liquid Immersion," in *Optical Microlithography XV*, A. Yen, ed., *Proc. SPIE 4691*, pp. 459–465, 2002.
- [10] "ASML MaskTools Division." <http://www.masktools.com/content/mttechno.htm>.
- [11] T. Ito and S. Okazaki, "Pushing the Limits of Lithography," *Nature* **406**, pp. 1027–1031, 2000.
- [12] S. A. Campbell, *The Science and Engineering of Microelectronic Fabrication*, Oxford University Press, New York, NY, 2001.
- [13] M. Madou, *Fundamentals of Microfabrication*, CRC Press, Boca Raton, FL, 1997.
- [14] C. Shishido, Y. Takagi, M. Tanaka, O. Komure, H. Morokuma, K. Sasada, and L. (Hitachi, "Characterizing cross-sectional profile variations by using multiple parameters extracted from top-down SEM images," in *Metrology, Inspection, and Process Control for Microlithography XVI*, D. J. Herr, ed., *Proc. SPIE 4689*, pp. 453–460, 2002.
- [15] B. Bunday, M. Bishop, and J. Swyers, "Quantitative profile-shape measurement study on a CD-SEM with application to etch-bias control and several different CMOS features," in Herr [33].

- [16] E. M. Drege, J. A. Reed, and D. M. Byrne, "Linearized inversion of scatterometric data to obtain surface profile information," *Optical Engineering* **41**, pp. 225–236, January 2002.
- [17] N. H. Jakatdar, "Deep sub-micron photolithography control through in-line metrology," Master's thesis, University of California, Berkeley, 2000.
- [18] J. Bao, *An Optical Metrology System for Lithography Process Monitoring and Control*. PhD thesis, University of California, Berkeley, 2003.
- [19] E. Barouch, S. Knodle, and K. Matsunaga, "Scatterometry as a practical in-situ metrology technology," in Herr [33].
- [20] "Nanometrics—Basic Principles of CD Measurement and OCD Technology." <http://www.nanometrics.com>.
- [21] C. J. Spanos, "Statistical Process Control in Semiconductor Manufacturing," *Proceedings of the IEEE* **80**, pp. 819–830, June 1992.
- [22] D. Montgomery, *Methods and Philosophy of Statistical Process Control, 2nd Edition*, John Wiley & Sons, New York, 1991.
- [23] "ExperTune: PID Tuning and Performance Monitoring Software." <http://www.expertune.com/tutor.html>.
- [24] J. Sturtevant, M. Weilemann, K. Green, J. Dwyer, E. Robertson, and R. H. M. S. Products), "Implementation of a Closed-loop CD and Overlay Controller for sub-0.25 μm Patterning," in *Metrology, Inspection, and Process Control for Microlithography XII*, B. Singh, ed., *Proc. SPIE* **3332**, pp. 461–470, 1998.
- [25] T. E. Zavecz and R. Blanquies, "Predictive process control for sub-0.2 μm lithography," in *Metrology, Inspection, and Process Control for Microlithography XIV*, N. T. Sullivan, ed., *Proc. SPIE* **3998**, pp. 312–322, 2000.
- [26] S. A. Middlebrooks, "Optimal model-predictive control of overlay lithography implemented in an ASIC fab," in *Advanced Process Control and Automation I*, M. Hankinson, ed., *Proc. SPIE* **5044**, 2003.
- [27] X. Niu, *An Integrated System of Optical Metrology for Deep Sub-Micron Lithography*. PhD thesis, University of California, Berkeley, 1999.
- [28] C. Mack, M. Ercken, and M. Moelants, "Matching Simulation and Experiment for Chemically Amplified Resists," in *Optical Microlithography XII*, L. V. den Hove, ed., *Proc. SPIE* **3679**, pp. 183–192, 1999.
- [29] S. Jug, R. Huang, J. Byers, and C. Mack, "Automatic Calibration of Lithography Simulation Parameters," in *Lithography for Semiconductor Manufacturing II*, C. Mack and T. Stevenson, eds., *Proc. SPIE* **4404**, pp. 380–393, 2001.
- [30] M. Mason, R. Soper, R. Terry, and C. Mack, "Process-specific tuning of lithography simulation tools," in *Optical Microlithography X*, G. E. Fuller, ed., *Proc. SPIE* **3051**, pp. 491–498, 1997.
- [31] W. L. Foong, "Characterizing the sensitivity of scatterometry for sub-100nm technologies," Master's thesis, University of California, Berkeley, 2001.
- [32] Y.-M. Lee, M.-G. Sung, E.-M. Lee, Y.-S. Suh, H.-J. Bak, and H.-K. Oh, "Temperature Rising Effect of 193nm Chemically Amplified Resist during Post Exposure Bake," in *Advances in Resist Technology and Processing XVII*, F. Houlihan, ed., *Proc. SPIE* **3999**, pp. 1000–1008, 2000.

[33] D. J. Herr, ed., *Metrology, Inspection, and Process Control for Microlithography XVII, Proc. SPIE 5038*, 2003.