INTEGRATED CIRCUIT PROCESS DESIGN FOR MANUFACTURABILITY USING STATISTICAL METROLOGY

by

Crid Yu

Memorandum No. UCB/ERL M96/47

6 August 1996

INTEGRATED CIRCUIT PROCESS DESIGN FOR MANUFACTURABILITY USING STATISTICAL METROLOGY

by

Crid Yu

Memorandum No. UCB/ERL M96/47

6 August 1996

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

Integrated Circuit Process Design for Manufacturability Using Statistical Metrology

by

Crid Yu

Doctor of Philosophy in Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Costas J. Spanos, Chair

The design of a VLSI process can be described as a procedure to center and optimize process settings. Using a suitable design of experiments, such as Response Surface Methodology (RSM), the space relating output process parameters (i.e. film thickness, critical dimensions, etch rate) to input process settings (equipment types and settings) is mapped out. The process designer then explores this mapped space to determine the vector of settings which will produce a set of parameters that meets the process specifications and is manufacturable. In this context, the manufacturability of a process is related to the variability of an output process parameter. Higher manufacturability requires lower variability. To incorporate process variability into the design, the process manufacturability as well as the output process parameters (nominal values) must be determined over the range of process settings.

Statistical Metrology, which combines conventional metrology with statistical filtering, has been presented as a methodology to extract and calculate error budget items versus process controls (equipment and recipe) over a range of circuit design variables (critical dimension, proximity). As an example, we have developed a comprehensive error framework and error characterization methodology and applied it on the decomposition of the error budget of photolithography. We formulated an error structure which describes both deterministic and random critical dimension variability as the sum of variability components contributed by physical sources (process, equipment, recipe). This information was then used to impact process Design for Manufacturability (DFM) in: 1) processing equipment selection and matching, 2) recipe optimization for manufacturability, and 3) accurate and efficient calculation of process margins. DFM was implemented by managing the process error budget, which is a tabulation of the error introduced by each process step and the way in which errors accumulate over a sequence of steps.

Chair Date

Acknowledgments

I am very grateful to Professor Costas Spanos for taking responsibility for supervising my thesis research when I went to him with a half-baked idea, for overseeing my personal and professional development, and for being almost always available. Especially I am eternally grateful to him for the time he spent on this thesis.

I am indebted to Professor Nathan Cheung, who invested his mentorship and attention to me as an undergraduate and then graduate, whose style of active debate I take with me out of Berkeley.

Many thanks to Professor David Brillinger and Professor Avideh Zakhor for their patience and participation in this thesis.

I thank my industrial mentors, Dirk Bartelink and Niel Berglund for their interest in my research and career, and for the many doors they have opened for me.

I am grateful to the staff at Hewlett Packard. Thanks to Hua-Yu Liu, Bob Gleason, and Nader Shamma for being mentors and colleagues, and for providing resources conducive to academic research.

Many thanks to the staff at Advanced Micro Devices- Bill Arnold, Anna Minevielle, Kenji Morisaki, and Anita Fumar, for their willingness to collaborate and openness to share information.

Thanks to past and present BCAMers, for their friendship and support, and for providing an environment which will most likely make all others I encounter pale in comparason.

This work was supported by a Doctoral Fellowship for Integrated Manufacturing provided by the Department of Energy, the Semiconductor Research Corporation, California Micro, and Advanced Micro Devices.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

While aggressive CD control specifications are projected for future device generations, process or equipment variability is not necessarily falling at the same pace. This is a result of both pushing equipment technologies to the limit, especially in the area of optical lithography, and the decrease in nominal critical dimensions. In particular, systematic equipment variability in some segments of the IC process chain has become a significant fraction of the error budget. These trends call for careful management of the process error budget.

Ideally, there would be three types of information contained in an error budget: (1) a classification of process error in useful, quantified components, (2) the correlation structure between components, and (3) the relationship between process variability components and process design variables. Such a complete understanding of the error budget can provide many opportunities to engineer process variability. Variability bottlenecks can be identified so as to direct the focus of engineering efforts. Process error can be manipulated by changing settings and/or redesigning equipment. Trade-offs between process performance and yield become possible. In short, process variability becomes designable. However, not all items of interest in a process error budget can be determined by conventional means.

Statistical Metrology [1] is a methodology which combines measurements from existing metrology with subsequent data filtering to extract items in the error budget. An ideal experimental or test structure design would provide data free of error or confounding variations, though this is almost never the case in the real world. However, simple error reduction techniques by statistical manipulation (such as averaging or regression) of raw experimental data has been common practice. Statistical Metrology uses statistical proce-

dures to extend the capability of experimental techniques in order to decompose raw measurements into variability components. To demonstrate this by example, we have applied this methodology towards extracting equipment and measurement errors from electrical and automated SEM CD measurements.

1.1 Background

The growing visibility of systematic error, that is variability with some structure in either space or time, suggests that there may be some lower limit to the amount of variability which can be achieved. Spatial variability of this sort is usually an intrinsic feature of the equipment and as such is a concern in process design.

Process parameter variation is often deterministic and can thus be assigned a cause. For example, the variation of film thickness across a wafer may appear as a continuous function (such as radial variation) of the coordinates of the wafer, and may be caused by non-uniformity introduced by the equipment. These types of systematic variation are routinely observed in processing, and have been used by operators as qualitative visual indicators of process well-being. This suggests that the causal nature of these variations is physical and therefore can be characterized and modeled. A causal classification of process error may also provide an opportunity to control that error.

Variation generally classified as "random" may simply be deterministic variation that is uncharacterized. Moreover, the truly random process variation associated with today's well-characterized process equipment is very well bounded and subject to fundamental quantification. This gives us the (as yet discarded) opportunity to assign causes to variation and to systematically remove and/or model these causes.

In addition to spatial and causal qualities of process error, a correlation structure between different variabilities can provide another lever to control CD's. For example, consider two or more systems in sequence which have relatively poorer control on the wafer edge than in the center. The resulting parameter distribution on the wafer will be different for different individual spatial distributions of error, even if the individual contributions of each equipment remain the same.

To formulate an error budget, it is also necessary to understand and accurately quantify the structure of process error, whether systematic or random, spatial or temporal. Error with causal and spatial structures will require statistical descriptions different from the uncorrelated gaussian distribution assumed in most conventional error budget calculations. Work done by other authors have characterized spatial features of process error and used them in process diagnosis [2]. Obviously, the spatial structure in process error must be taken into account in order to formulate an accurate error budget. Furthermore, the spatial distribution of process parameters has implications for both process and circuit design. Processes producing opposite trends over the wafer may be combined to cancel one another. Similarly, process variation within a die may be of concern to circuit designers.

From the error budget, yield bottlenecks can be identified and engineering efforts can be focused to the particular process steps introducing the most process variation. A design space for process variability can be mapped versus controllables, such as the equipment and the recipe used. The space is analogous to the response surface used in determining the mean or target performance of a process. Both the performance and manufacturability responses will be used to determine an optimal operating point (Figure 1.1). Given quantified responses for performance and manufacturability, informed trade-offs between the two may be possible.

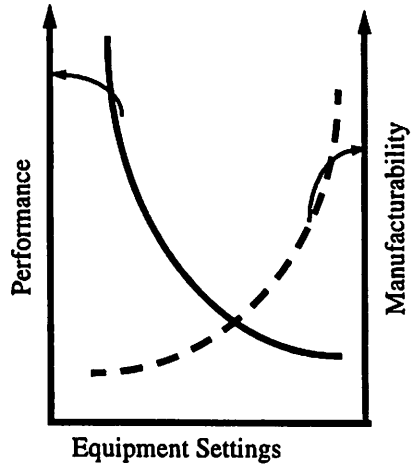


FIGURE 1.1 Concurrent performance and manufacturability design curves.

1.2 The Structure of Patterning Variability

Determinism, spatiality, and causality are three attributes that describe the variability one might observe in VLSI fabrication. These constitute a complete process error framework shown in Figure 1.2. Spatial classifications, such as within die and between die, are frequently used to factor process error. However, the exact spatial patterns of variation have not been extensively characterized and cannot be fully known with a limited amount of data samples. Moreover, spatial classification alone does not provide the means to control process error, which is of ultimate engineering use.

The next category, causality, classifies process variability by its contributing sources. Just as in the case of spatial hierarchies, there are many levels of causality. Assignment of variability by equipment and process recipe is the most common form of causal classification. Environmental factors, which are more difficult to identify and control, are also variability sources. It is unlikely that all sources of process error will ever be identified.

Finally, process variability can be deterministic or random. As an example, consider an etcher (cause) that introduces a radial (deterministic) etch rate variation within the wafer. The same etcher could also introduce random variability over the same wafer. In this conceptual framework, the error budget is primarily a causal decomposition of process error.

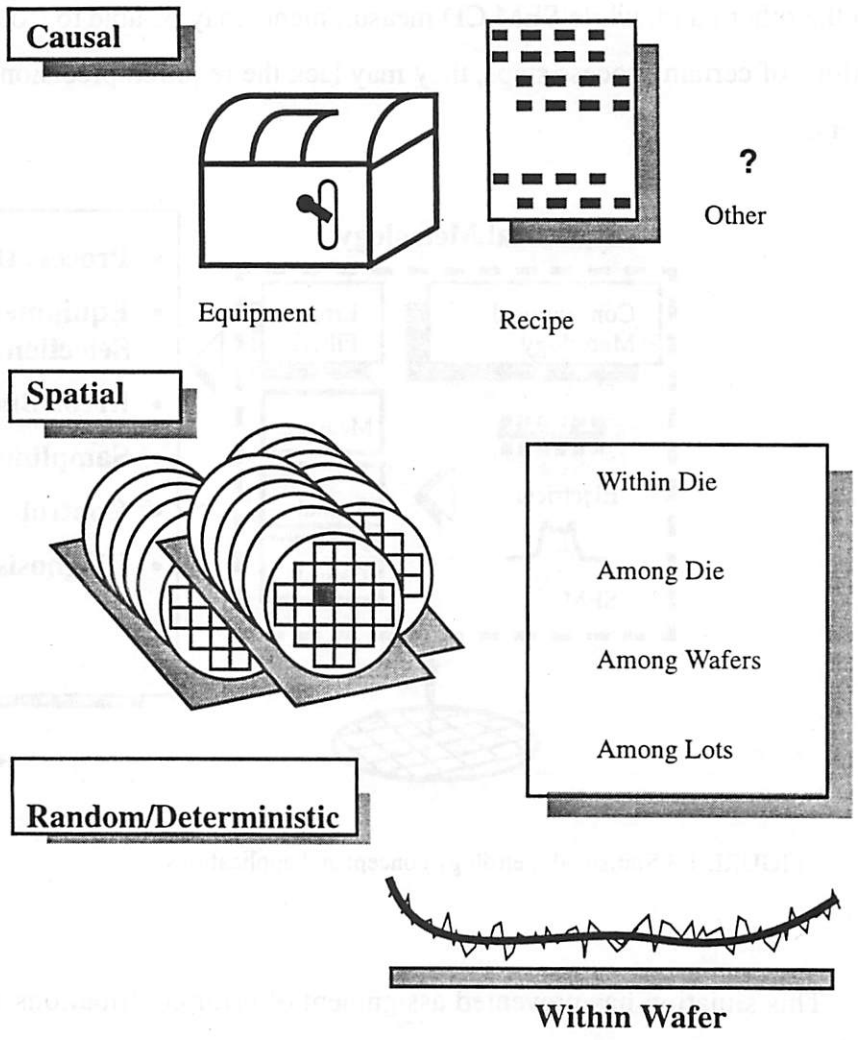


FIGURE 1.2 Process variability categories.

1.3 Statistical Metrology- Critical Dimension (CD)

One of the basic requirements of an IC process is the ability to pattern features of specified sizes in a repeatable manner. Critical dimension (CD) control is a popular metric for gauging fab performance because it indicates whether equipment is functioning within control and also anticipates the final circuit yield. CD error is the result of aggregate errors in the sequence of film deposition, resist patterning, and film etching. The error contribution of each process step cannot always be determined by conventional metrology. Electrical CD measurements of patterned conductors provide fast and precise counts of

variability. Further, because the measurements can only be taken at the end of the patterning sequence they include confounded error contributions from all steps in the sequence. On the other hand, while SEM CD measurements may be able to isolate variability contributions of certain process steps, they may lack the required precision due to measurement errors.

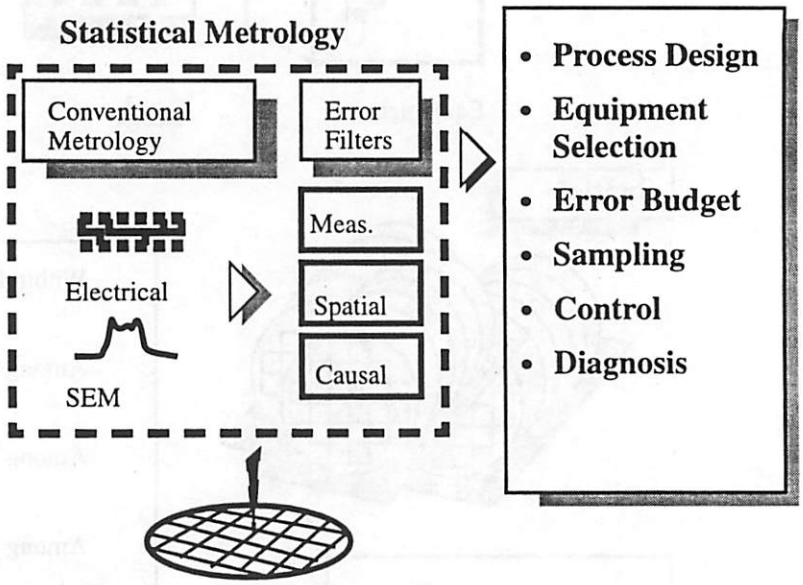


FIGURE 1.3 Statistical metrology concept and applications.

This situation has prevented assignment of error contributions amongst various process segments, most notably between lithography and etch, and between different choices of process equipment. Ultimately, it is the electrical performance that must be delivered to the customer, so a manufacturability metric should be based on electrical performance variation. Following Statistical Metrology concepts, a methodology has been developed in this thesis to extract individual equipment variability from electrical CD measurements. Similarly, a methodology has also been developed to characterize and to remove SEM measurement artifacts in variability sampling.

In this thesis, a novel metrology system is demonstrated by examples in electrical or SEM CD decomposition (Figure 1.3). The measurement of process parameters is by nature a sampling of errors: measurement error and components of process error. While

the raw data is rich in information about process error, this information is usually inscrutable because the errors are confounded. Error filtering and signal extraction has been associated with metrology, even in simple forms such as averaging to reduce random error. Statistical Metrology is proposed in this thesis as a methodology which extends error filtering in metrology to include process error components as well as experimental error. Measurements from conventional metrology, such as SEM or electrical CD's are decomposed into error components with the use of a series of filters. Metrology is "enhanced" because the error structure can be discerned.

1.4 Statistical Metrology- Methodology

A methodology has been developed to extract the error budget of a process sequence by accurate and systematic characterization of the variation, both deterministic and random, associated with each process parameter. Shown in Figure 1.4, this methodology consists of four phases: (1) Experimental design and data collection, (2) Physically based statistical filtering, (3) Error budget decomposition and, (4) Verification.

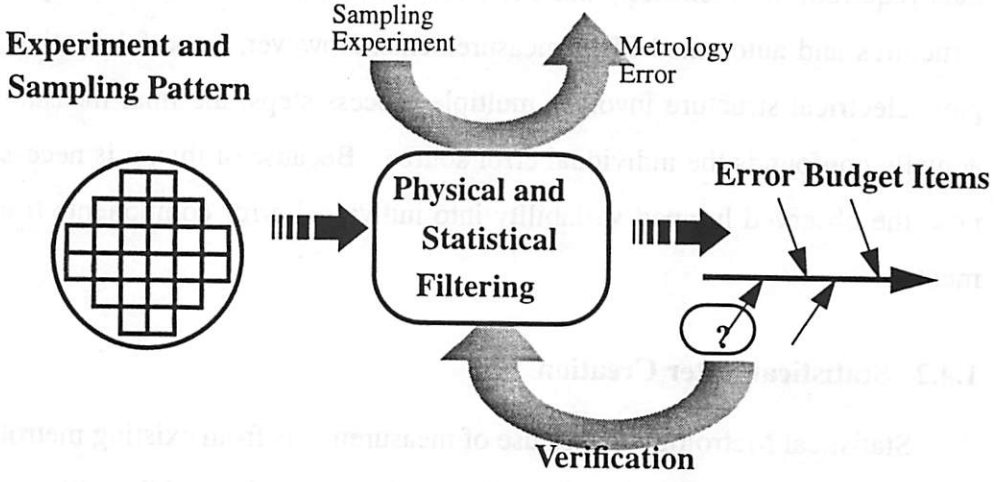


FIGURE 1.4 Schematic of decomposition methodology.

1.4.1 Experimental Design

Statistically significant quantities of data can be collected efficiently with short loop processing experiments in the real time environment of a wafer fab. An experiment was designed to collect data over the deterministic, spatial, and causal error dimensions of

interest by sampling CD variability at a fixed operating point in the process using linewidth measurements [3][4]. This is in contrast to conventional process design experiments which seek to determine the optimal process setting, or roughly the mean value of the process response.

Spatial variability is sampled by repeating CD patterns over the reticle in a grid and repeating the reticle patterns over the wafer. The dimension and proximity conditions of the CD patterns were varied in each location on the reticle as sample systematic variability due to mask CD variations and proximity effects. A variety of linewidth topologies can be considered. The linewidth structures are electrically measurable and use the Kelvin 4 point measurement design, with the sheet resistance determined by adjacent Van der Pauw structures. The raw data from the experiment consists of periodic repetitions of the reticle over the wafer.

The experiment is conducted over a short-loop process sequence as to sample a minimum number of process error components. Because of the large amounts of high precision data required, the metrology has been constrained to that of electrically measurable test structures and automated SEM measurements. However, since fabricating even the simplest electrical structure involves multiple process steps, the final measurable variability actually confounds the individual error sources. Because of this, it is necessary to decompose the observed lumped variability into individual error components from each equipment.

1.4.2 Statistical Filter Creation

Statistical Metrology makes use of measurements from existing metrology combined with subsequent data filtering algorithms to infer causal variability. Ideally, experimental or test structure design would provide data free of error or confounding variations. However, this is almost never the case. Meanwhile, simple statistical manipulation (such as averaging or regression) of raw data has always been used in practice to reduce measurement error. Statistical procedures are used to extend the capability of experimental techniques in order to decompose raw measurements into equipment contributions.

The data collected exhibit particular spatial structure due to the operation of the processing equipment (for example, within the exposure field vs. over the wafer). Furthermore, these spatial patterns can be linked to physical causes. We propose a two step data decomposition process to exploit these data features. A statistical signal filter first decomposes the raw data into signal subcomponents based on spatial characteristics. Because different processing equipment affect different spatial domains, the physical sources of variability for a particular signal subcomponent is only a subset of the total variability components of the experiment. Then, a second physical decomposition involving additional experimentation is carried out for particular signal subcomponents to determine error contributions from specific equipment. These filters may be implemented through separate experiments and analysis.

1.4.3 Error Budget Decomposition

Error decomposition is performed using a set of test structures sensitized to particular process disturbances *in conjunction* with subsequent statistical and physically based filtering techniques. Raw data is collected from an experiment which samples variability over a particular spatial extent and over a particular process sequence. Through the use of data filters described in Section 1.4.2, the deterministic components of the variation can be identified and decomposed into components associated with individual process equipment. This decomposition process may be aided by multiple structures with sensitivity optimizations for each tool. Thus, the system consisting of conventional CD measurements and data filtering is actually a metrology system with enhanced ability to resolve confounded causes over various spatial extents.

1.4.4 Filter Verification

The system of enhanced metrology yields outputs which are essentially spatial and causal error components which can be used as responses in process design. The effectiveness of the filters need to be verified before such a metrology system is to be put to use. An additional set of experiments should be carried out to corroborate the estimates from the filters. In some cases independent, alternate metrology would suffice, such as the use of SEM measurements to confirm the decomposition of electrical measurements. In other

instances deliberate known perturbations to equipment variability can be introduced to verify the causal filter.

The combined effects and interactions of a complete IC process are enormously complex. Smaller, more tractable portions of this problem will be approached first in characterizing process modules. Photolithography is used extensively in IC manufacturing and is a relatively well-characterized process. Our work will start first with applying the proposed methodology to the lithography process module and demonstrating the feasibility of this methodology with a 0.35μm poly gate patterning sequence.

1.5 Impact of Statistical Metrology on DFM

The Statistical Metrology system yields extracted error components which can be used to impact process DFM in several ways. The effects of the stepper and the reticle on CD error has been determined. Methodology has being developed to extract the variability contribution from plasma etch in the metal patterning sequence. Characterization and removal of measurement errors from automated SEM tools have allowed the estimation of develop and etch contributions to CD variability.

1.5.1 IC Manufacturability as it Relates to Processing Equipment

Increasing complexity in IC fabrication has put tremendous demands on equipment manufacturers to meet next-generation process specifications. This is true, for example, of stepper manufacturers. However, technologies exist that extend the usefulness of lithography tools half or even a full generation, i.e. higher contrast resist, phase-shift masks, off-axis illumination schemes. In process design, the option exist between using existing technology with enhancements or going to more advanced but less familiar technology. Inevitably, the manufacturability of the process becomes the deciding factor. Statistical Metrology can be used to extract an estimate of the isolated equipment error contribution as a basis for machine comparison to determine equipment and technology choices.

In the case of poly CD's, a statistical filter extracted the periodic signal component over the wafer. Subsequently, additional experimentation and physical filtering yielded

separate reticle and stepper components of CD variability. This allows equipment benchmarking studies independent of process variability from other parts of the processing sequence. The sensitivity of CD error to reticle error makes it possible to provide realistic specifications to reticle vendors. Manufacturability bottlenecks can be identified in the error budget to optimize design priorities.

1.5.2 DFM in Process Design

Statistical metrology can be used to generate a manufacturability space in which to design a process or a circuit. Equipment and process variability is determined as a function of operating controls so that manufacturability is designed concurrently with process performance. In this way, a process is designed to be manufacturable as well as functional. This can help speed up time to market as well as increase yield learning. With a quantified manufacturability metric, informed trade-offs can be made between process performance and manufacturability.

1.5.3 Flexible Design Rules

A measure of manufacturability can be assigned to a circuit layout and comparisons made between alternative circuit designs. Layout design rules become flexible because aggressive design can be justified by taking quantifiable penalties in manufacturability. A simple example would be to use narrower metal lines to improve performance at the expense of linewidth control.

The IC industry has become quite homogenous. That is, constrained by the availability of technology and their manufacturability, process technology of a given generation is essentially the same across the industry. Furthermore, equipment is in general available to all parties. This has made it more difficult to gain a competitive advantage from the fabrication process alone. The focus of competition has shifted to manufacturability issues, such as yield ramping rate and time to market. Statistical metrology can lend competitive advantage in these critical areas.

1.6 Thesis Organization

Statistical Metrology as described in Section 1.4 will be demonstrated through examples in pattern transfer. Chapter 2 describes the reticle and measurement setup used to sample electrical CD's. Various versions of this reticle will be documented. Chapter 3 is an example of a complete iteration of the methodology implemented on the stepper/reticle decomposition of a poly CD patterning sequence. The use of automated SEM measurements to estimate develop and etch components of poly patterning is described in Chapter 4. In Chapter 5, the contribution of metal etch to metal CD variation will be characterized. Applications of Statistical Metrology to process DFM and equipment manufacturability analysis will be discussed in Chapter 6. In Chapter 7, concluding remarks will be made and future work will be discussed.

References for Chapter 1

[1] Dirk Bartelink, "Statistical Metrology- at the Root of Manufacturing Control", *Journal of Vacuum Science and Technology* B, vol. 12, no.4, July/August 1994.

[2] John K. Kibarian and Andrzej J. Strojwas, "Using Spatial Information to Analyze Correlations Between Test Structure Data", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 4, No. 3, Aug. 1991, pp. 219-225.

[3] C. Yu, *et al.* "Use of Short Loop Electrical Measurements For Yield Improvement", *IEEE Transactions on Semiconductor Manufacturing,* vol. 8, pp. 150-159, May 1995.

[4] H-Y Liu, *et al.* "Contributions of Stepper Lenses to Systematic CD Errors Within Exposure Fields", Proceedings of SPIE's 1995 International Symposium on Microlithography, Feb. 19-24, Santa Clara, CA.

Chapter 2

CD Sampling Experiment and Short-loop Process Flow

In this chapter, the metrology issues (Section 2.1) associated with variability sampling as well as the experimental designs (Section 2.2) used in this thesis will be described. The process parameter monitored in this thesis is critical dimension (CD) or linewidth. Most of the experiments in this thesis have been performed on the polysilicon gate patterning sequence. Because of increasing concern over the control of interconnect dimensions, metal CD variation will also be treated in Chapter 5. Electrical CD measurements will be used to characterize polysilicon gate patterning, specifically the reticle and stepper error contributions (Chapter 3) as well as the metal CD variations. Automated SEM measurements will be used to estimate the contributions of the etch and develop steps in the gate patterning sequence. In all experiments, the process sequences used were part of the .35μm generation technology.

Ultimately, the goal of this work is to isolate equipment contributions to CD variability. However, to form measurable patterns in a thin film, a number of process steps must be involved, and their contributions to process error are confounded in the final measurements of CD. While both the metrology and the test structures used can be sensitized to decouple equipment variability, such as separate sheet resistance and linewidth structures, this is not always possible. A feature of the methodology proposed in this thesis is the use of physically based statistical filters to decouple sources of variability present in raw data from conventional measurements. This topic will be addressed in subsequent chapters. The experiments described in this chapter have been designed to sample the variability of interest and to facilitate error decomposition. In general, short-loop sequences in state of the art processes were used.

2.1 Metrology for Variability Sampling

Metrology used to sample process variability must fulfill two criteria: 1) Cost effectiveness- because of the high volume of data required for statistical characterization of a complex space of spatial and causal attributes, economies in both time and resources are crucial. As an example, the number of CD test structures placed in the experiment described in (2.1.1) total to 51,840. At one measurement per second the test time per wafer would be ~14 hours. 2) High data precision- the magnitude of the variability interest is in some cases on the order of nanometers. Currently, only electrical measurements satisfy these criteria. Recently, automated SEM's with a sampling rate of ~1 Hz have also become attractive metrology choice. However, the data may suffer from significant measurement artifacts which confound the actual spatial variability. This will be discussed in Chapter 4.

Electrical measurements used in this thesis were made at Hewlett Packard Co.'s ULSI Laboratories, using the HP ICMS [5] automatic electrical test platform. The testing methodology, including mask design and metrology system, follows the description given by C. Weber [6]. The hierarchy of structures, with increasing specificity, is as follows: wafer, die, module device. Dice refer to an instance of a periodic pattern on the wafer, and could be the whole or part of the stepper field. The module refers to the set of structures that can be tested with one touch-down of the probe card. "Devices" are measured structures in the module. The test program hierarchy is structured in parallel to the levels of physical abstraction. A library of all the structures and locations is referenced. At the device level standard measurement algorithms (such as sheet resistance or linewidth) are applied.

The hardware consists of measurement, analysis, and storage devices connected on an HP-IB bus. An automated probe station with movable stage and pattern recognition capability measures the structures automatically. A switching matrix multiplexes the measurement meters and power sources between the various probe pins. Test execution is driven by a UNIX workstation.

For the polysilicon structures, the test current for the Van der Pauw and the Kelvin structures is set at 0.01mA. This gives ~10mV of voltage difference across the CD struc-

ture. For the metal structures, the current is set at 1mA to compensate for the much lower sheet resistivity (60mΩ vs. 12 Ω for poly) to yield a voltage drop of similar magnitude.

SEM CD measurements used in Chapter 4 were taken at the Advanced Process Development Laboratory of Advanced Micro Devices. SEM experimental setup and error analysis will be discussed in Chapter 4.

2.2 Experimental Design For CD Analysis

2.2.1 CD Test Structures and Reticle Design

A test mask designed at Hewlett Packard was used to generate a range of CD structures in various layout configurations over the space of interest [7]. The mask design is summarized here for reference.

Five topologies consisting of isolated lines, double lines, triple lines (measuring the center line), triple lines (measuring the side line), and a nested line in 5 lines were created (Figure 2.1). Lines were placed in both x and y orientations over 5 nominal dimensions ranging from 0.25 to 0.45 μm in .05μm increments. All nested lines were designed at 1 μm pitch to reflect the process requirements. The range of CD variations was designed to span the linewidth choices of interest to the process, nominally targeted at .35μm. In process design the pitch is usually fixed for a given generation of technology, while the actual dimensions on the wafer are manipulated by adjusting the mask dimensions as well as the processing conditions, such as exposure dose and etch bias. This mask design was also adapted at Advanced Micro Devices to characterize lithographic performance. In this version the pitch (spatial frequency) of the lines were varied. Additional structures were also

included to characterize aberrations in optical lithography. In addition, this design has

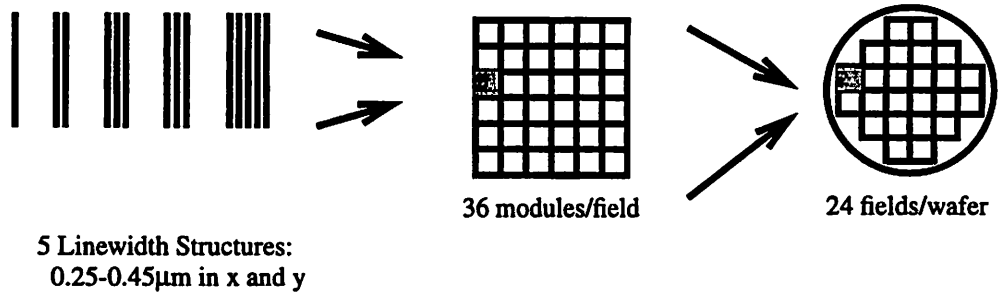


FIGURE 2.1 Schematic of linewidth topologies and placement in the reticle and wafer.

been adapted and scaled for the 2μm CMOS process at the UC Berkeley Microfabrication Laboratory. This is shown in Figure 2.2 as an example. For metal CD decomposition described in Chapter 5, another version of this test mask was created where the nominal

CD dimension is .5 μm and the patterns are placed in a 5x5 array within the stepper field.

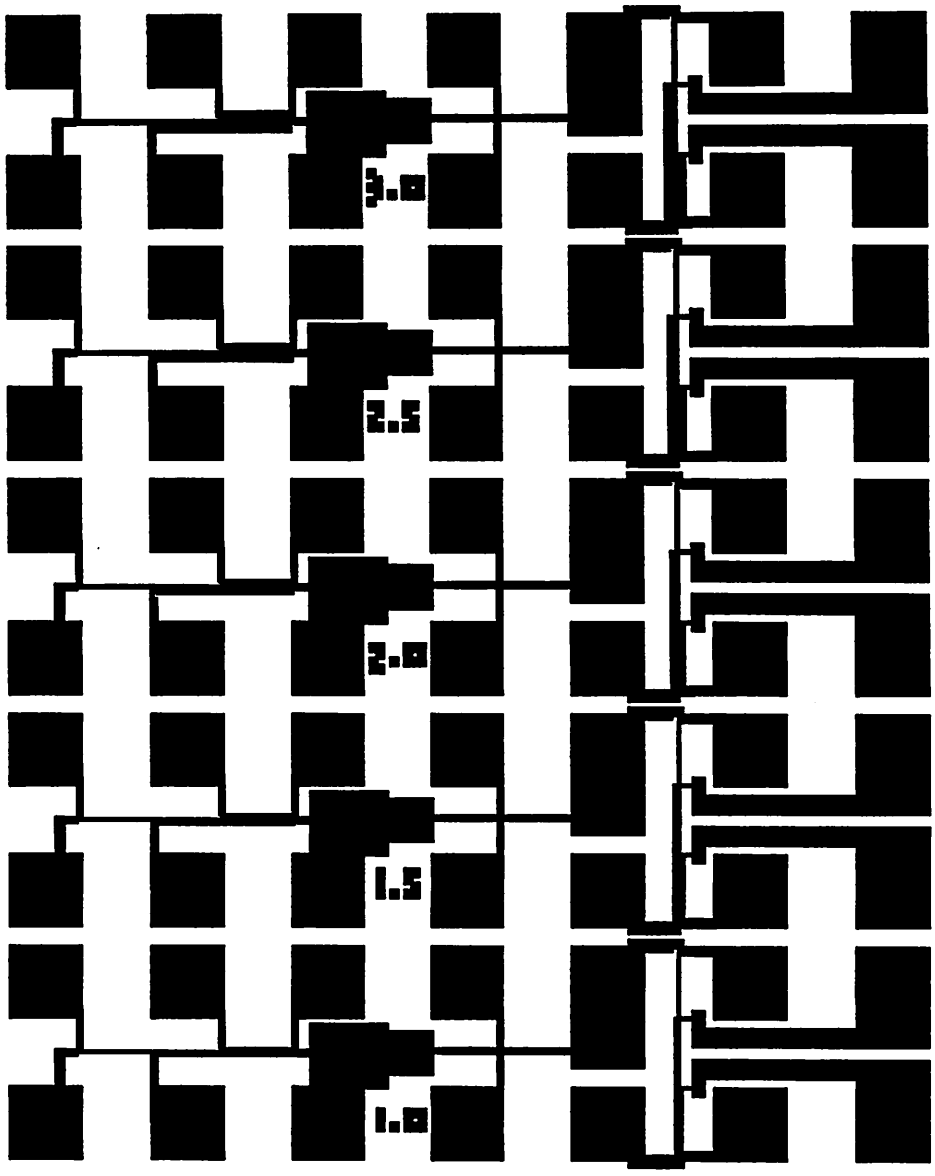


FIGURE 2.2 Test structure design for electrical linewidth measurements over 5 nominal dimensions. Design shown is adapted and scaled for a 2μm CMOS process. Original design was fabricated on 0.35μm CMOS process.

Structures of all topologies with one nominal dimension were placed in a "track" of 2x12 probe pads. These structures could be visited by two touch-downs of the 2x12 probe card. All linewidth structures use variations of the 4-point Kelvin structure, with two sets of separate current and voltage taps to eliminate contact resistance. To conserve pad use, different structures may share common current taps and one of two voltage taps. Van der

Pauw structures were used to obtain the sheet resistance value used in the linewidth calculations. Sheet resistance values in proximity to the CD structures were used to calculate linewidth in order to eliminate the effects of sheet resistance non-uniformities. Two sheet resistance values were used in each exposure field for all linewidth measurements in the field. Sheet resistance variability was characterized after sampling to detect large gradients within one field. In general the variation over the field was less than one percent.

These structures were placed in a group and were replicated 36 times within the stepper field in a 6 x 6 array and the resulting 20mm by 20mm die was repeated 24 times over a 6 inch wafer (Figure 2.1). Depending on the stepper, 22x22mm fields were also used. These test structures span a continuum over design parameters in gate length and layout configuration, as well as over physical space- over the field and over the wafer.

2.2.2 Measurement Error Characterization

Measurement error from electrical metrology was characterized. Both polysilicon and thin TiW film patterns were characterized as appropriate for the gate CD (Chapter 3-5) and metal interconnect CD experiments (Chapter 6). Measurements (600 on TiW and 864 on poly) were repeated on each wafer in order to estimate the measurement error. Both sheet resistances and linewidths CD, nominally .35μm for poly and .50μm for TiW, were measured and calculated. The results are shown in Table 2.1. The poly and TiW results are

Table 2.1 Electrical CD Measurement Error Summary

	Sheet Resistance Std. Err. (Ω/□)	Sheet Resistance % Error	CD Std. Error (nm)	CD % Error
Poly	1.3	0.765%	1.5	0.68
TiW	.066	0.465%	1.72	0.45

comparable, and representative of the repeatability of electrical measurements used throughout this work. This is an advantage derived from the inherent averaging in current flow over a length of material and the repeatability of the electrical instrumentation. High precision, the fact that electrical metrology does not interact with the sample to produce measurement proximity or memory effects (such as SEM), as well as superior throughput

and sampling rates make electrical measurements suitable for variability analysis.

In Chapter 4, automated SEM CD measurements will be used to estimate the etch and develop contributions to CD variability. SEM measurement error will be characterized and compared to electrical results.

2.3 Pattern Transfer Sequence Process Flow

Gate length control in an IC process is crucial because channel length variations strongly impact circuit performance. Furthermore, poly CD's are the smallest features of an IC process and usually challenge existing lithographic and etch technology. CD measurements were obtained from electrically testable structures that involve a minimum of processing steps so that large amounts of data can be collected with relatively few sources of variation. The sequence of short loop gate patterning processing steps includes poly CVD, resist coat, exposure, development, and finally gate definition by plasma etch (Figure 2.3). This short-loop process represents the minimum steps required for patterning, and this simplicity facilitates decomposition. Because pattern transfer constitutes the basic module of IC fabrication, understanding derived from short-loop experiments is easily leveraged to other sections of the process sequence. In this work, the patterning sequences involved used I-line lithography technology. This is used in the .35μm generation of process technology without optical or process enhancements, such as annular illumination or antireflective coatings.

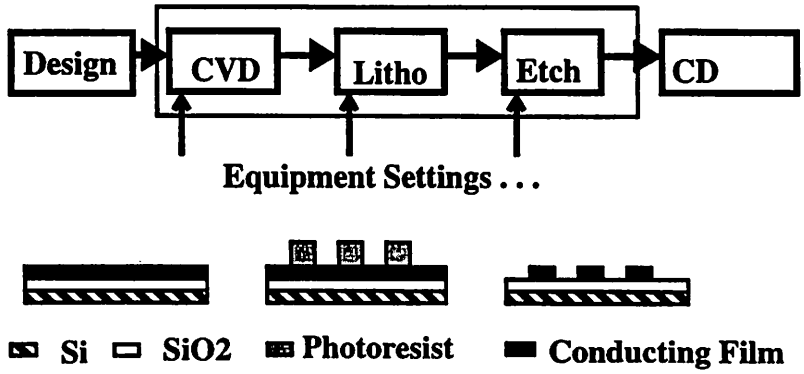


FIGURE 2.3 Schematic of pattern transfer process sequence.

For demonstration purposes, the effects of wafer topography or parameters in profile

control, such as sidewall angle, were not considered. Gate profile control was verified independently by SEM measurements. This experiment takes advantage of the relatively high density of structures that can be placed on a reticle to explore systematic variability as a function of process settings. However, only a small subset of the total process space will produce a process which is acceptable, i.e. meets the specified nominal performance targets. Process variability is only meaningful in this subspace of process conditions. In the experiment, independent metrology and experiments have been used to determine the process operating points. For example, prior to this experiment, SEM measurements were used to determine appropriate stepper exposure conditions to pattern lines with the target dimensions.

2.4 Thin Film Variability

Measurements were taken to assess the uniformity of the deposited films in the poly patterning process. Film thickness measurements over the wafer were taken using standard reflectometry methods. Resist films were deposited over bare silicon wafers to facilitate measurement. Polysilicon film thicknesses on 4 wafers selected from a lot of 24 were measured by reflectometry. The results are summarized on Table 2.2. The variation over the wafer in each case was less than 15Å, 3-sigma. All metrics are very consistent over the sampled wafers. Moreover, the systematic thickness trend over the wafer is consistent across all the wafers.

Table 2.2 Polysilicon Film Uniformity Summary

Wafer #	Mean (Å)	Stnd. Dev. (Å)	% Stnd. Dev.	Max-Min(Å)
1	2470.5	4.594	0.186	15.5
2	2470.9	4.686	0.190	15.4
3	2470.9	4.832	0.196	16.3
4	2469.6	4.253	0.172	15.1

Similarly, resist thickness measurements were made over a number of wafers as well

as over a range of thicknesses. The uniformity ranges between over ~10 to 25 Å, 3-sigma.

Table 2.3 Photoresist Film Uniformity Summary

Wafer #	Mean (Å)	Stnd. Dev. (Å)	% Stnd. Dev.	Max-Min(Å)
1	7034.6	7.816	0.111	34.8
2	7105.3	2.589	0.036	11.0
3	7019.6	7.733	0.110	24.1

A simple sensitivity analysis can be performed to estimate the impact of resist variability on CD variation. The effect of resist thickness uniformity on CD is estimated from dose-to-clear measurements. In short, dose-to-clear (DTC) is the amount of exposure necessary to completely develop an area of resist and is proportional to the dose required to pattern a line, or dose-to-size (DTS). From empirical DTC curves [8], the DTS varies 3.89% for 100Å of thickness variation around the nominal resist thickness used in these experiments. Using 24Å as an upper-bound of within-wafer thickness variation from Table 2.2, the DTS varies correspondingly by approximately 0.93%. In other words, the equivalent dose modulation from resist thickness variations over the wafer is <1%. From this analysis, resist and poly thickness variation typically obtained in this process have negligible effect on the final CD variation and do not have to be considered in the decomposition.

2.5 Sampling Experiment and Process Flow Summary

Large quantities of data are easily collected using electrical or SEM measurements. However, resulting CD variation includes contributions from confounded sources. In the past, this has been a shortcoming of short-loop electrical measurements. However, taking advantage of the quantity and placement of the measurement structures and applying statistical and signal-processing techniques, it is possible to extract the errors introduced by a particular process tool. In this way the observed error can be decomposed into individual sources and attributed to specific modules, steps, or equipment. This procedure is demonstrated in the following chapters.

References for Chapter 2

[5] *IC-MS User's Manual*, Hewlett-Packard Co., Semiconductor Systems Center.

[6] C. Weber, "A Standardized Method for CMOS Unit Process Development", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 5, No. 2, May 1992, pp. 94-100.

[7] H-Y Liu, *et al.* "Contributions of Stepper Lenses to Systematic CD Errors Within Exposure Fields", Proceedings of SPIE's 1995 International Symposium on Microlithography, Feb. 19-24, Santa Clara, CA.

[8] Nader Shamma, Hewlett Packard Co. internal communications.

Chapter 3

Stepper/Reticle Decomposition of Poly CD Variation

The variability in the raw electrical CD measurements is the aggregate process variability from each step in the patterning sequence. However, ultimately it is the categorization of variability by cause that is of engineering importance. The itemization of the systematic spatial variability is significant because only a subset of the process steps contribute variability to each spatial component. While the raw electrical measurements from this experiment contain confounded variability from all parts of the process, the number of variability sources are reduced for any particular spatial component. Effectively, isolating a spatial component will reduce the number of variability causes that have to be decomposed. This chapter will focus on the decomposition of error contributions from the stepper and reticle from the collected raw data. We proceed by first extracting the periodic component $p(x,y)$ through the use of a spatial filter. Then, the contributions of the stepper and reticle will be determined by a separate physical filter. Our approach exploits the spatial structure of the measured data so that the causal variability is isolated through the use of a statistical and a physical filter used in sequence.

3.1 Polysilicon Electrical CD- Raw Data

In this chapter causal variability decomposition will be demonstrated by assigning variability contributions to the stepper and reticle pair from sampled poly CD's. The reticle used was described in Chapter 2. Spatial variability was sampled by repeating CD patterns over the reticle in a 6X6 grid and repeating the reticle patterns over the wafer. The dimension and proximity conditions of the CD patterns were varied in each location on the reticle as sample systematic variability due to reticle CD variations and proximity effects. A variety of linewidth topologies were considered including isolated, dual, triple, and 5 line arrays in both x and y orientations spanning a range of nominal linewidths from .25 to

.45 μm. The raw data from the experiment consists of periodic repetitions of the reticle over the wafer and can be seen for the case of .35μm isolated lines in Figure 3.1.

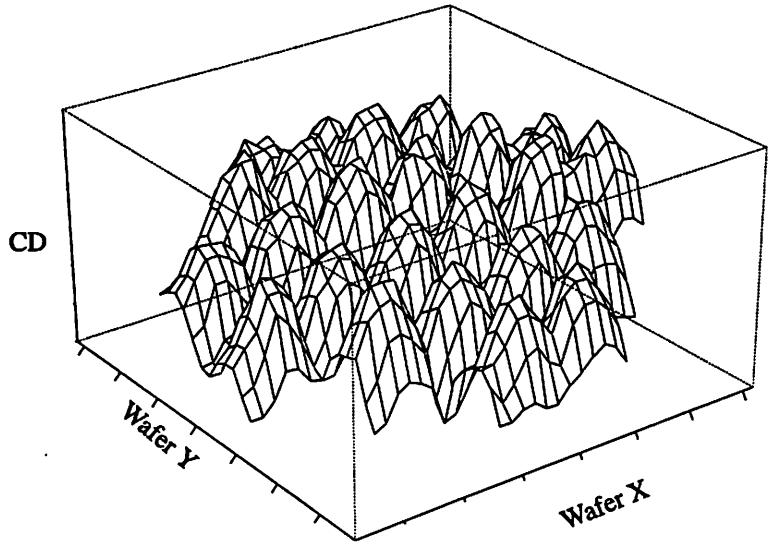


FIGURE 3.1 Raw poly CD measurements over the wafer for nominally .35 μm isolated lines.

Several observations can be made about the CD response over the wafer. First, most of the variability appears to be systematic. This would not have been determined without the periodic full field sampling. Second, the systematic variability appears to have several components. The variations within each stepper field appear as "dome"-shaped response surfaces which correspond exactly to the periodicity of the stepper fields. This variability component is repeatable with each exposure and appears to be the dominant source of variability within the wafer. Other, less obvious, sources of variability can also be observed. For example, there appears to be a slowly varying trend over the entire wafer. Systematic variability is easily associated with equipment error. These associations will be used to structure the decomposition strategy.

3.2 Poly CD Variability Components

The possible types of spatial variation are determined a priori to construct a list of effects which contribute to the CD response over the wafer (Figure 3.2). This formulation is based on physical understanding of the process mechanics, and is one of many ways to decompose the CD variation. As will be discussed in a later section, this decomposition is

advantageous because it facilitates subsequent decomposition of CD variance into equipment causes. Next we define the following variability components:

1) The stepper operates by exposing the pattern on the reticle repeatedly over the wafer. In these experiments the same reticle is used for the entire wafer. Automatic leveling of the exposure field is performed at each exposure to compensate for wafer or film flatness. Given identical exposure conditions and perfect spatial uniformity from the rest of the process steps, the stepper/reticle pair will generate perfectly periodic and identical CD responses in each exposure, corresponding to the periodicity of the stepping pattern. This is denoted by $p(x,y)$, where the variables x and y denote spatial coordinates over the wafer. The periodicity of p stipulates that

$$p(x, y) = p(x + ns_x, y + ms_y)$$

where s_x and s_y correspond to the field dimensions and n and m the indices of periodicity.

2) Exposure to exposure variation is modeled by the signal $f(x,y)$. This represents perturbations to the intra-field CD responses described in p, and are assumed to be additive. These perturbations may be exposure dose errors, which change the CD response over the entire field by a constant value. Similarly, focus errors can be modeled by a mean shift over the field. Another source of variation in f is leveling error, in which case the field is tilted linearly, translating into linearly varying defocus across the field. The signal f has the characteristic that it is systematic within the field and random from field to field.

3) Slowly varying trends over the wafer, $w(x,y)$, are introduced by batch or single-wafer processes, such as etch, develop, and resist or film coat (CVD). No functional form will be assumed.

4) A stationary error series over the wafer, $\varepsilon(x,y) \sim N(0,s^2)$. This represents random (white) noise from the process or measurement. Randomness was established in the repeatability measurements from Chapter 2.

The spatial dependence of $p(x, y)$, $w(x, y)$, and $f(x, y)$ is stated explicitly by the functional dependence on spatial coordinates x and y. We assume an additive nature of the error components so that,

$$cd(x,y) = p(x, y) + f(x, y) + w(x, y) + \varepsilon \tag{3.1}$$

where *cd* is the electrically measured CD value after processing. The random component of the error, or "white noise" is normally distributed but has no spatial dependence. Also, although p, f, and w all vary systematically over the wafer, p and f affect one stepper field per exposure. Thus, p consists of the CD response of the stepper and reticle occurring identically within each exposure field. The error components are illustrated in Fig. 3.3 in one dimension only for simplicity of representation.

The causal associations of the signals are also shown in Figure 3.2. The periodic variation p is caused by the stepper and reticle pair. The signal f is caused by variations in the stepper exposures. The signal w is caused by the etch, develop, and film depositions. From Chapter 2, the film thickness variations for resist and polysilicon have negligible effect on CD. Thus the dominant sources of w in poly gate patterning come from the resist develop and plasma etch steps. A subset of causes contribute to each isolated signal component. The decomposition strategy applies signal decomposition to estimate one of the signals. If applicable, subsequent causal decomposition of the isolated signal into the subset of equipment causes is performed.

The additive model stated in (3.1) assumes no interactions between the signal components. Similarly the independence of variation between stepper/reticle (f) and etch/develop (w) is assumed. An obvious scenario of interacting effects would be the presence of variations in the CD profile due to stepper field imperfections. The variability in resist profiles would interact with the etch process. Similarly, the additive model would not hold if the etch responses for different resist CD values were not uniform. In short, these assumptions

hold for a small range of values around the process operating point. Therefore, this methodology is appropriately applied at legitimate process operating points and when the process response is well controlled over the spatial extent of interest.

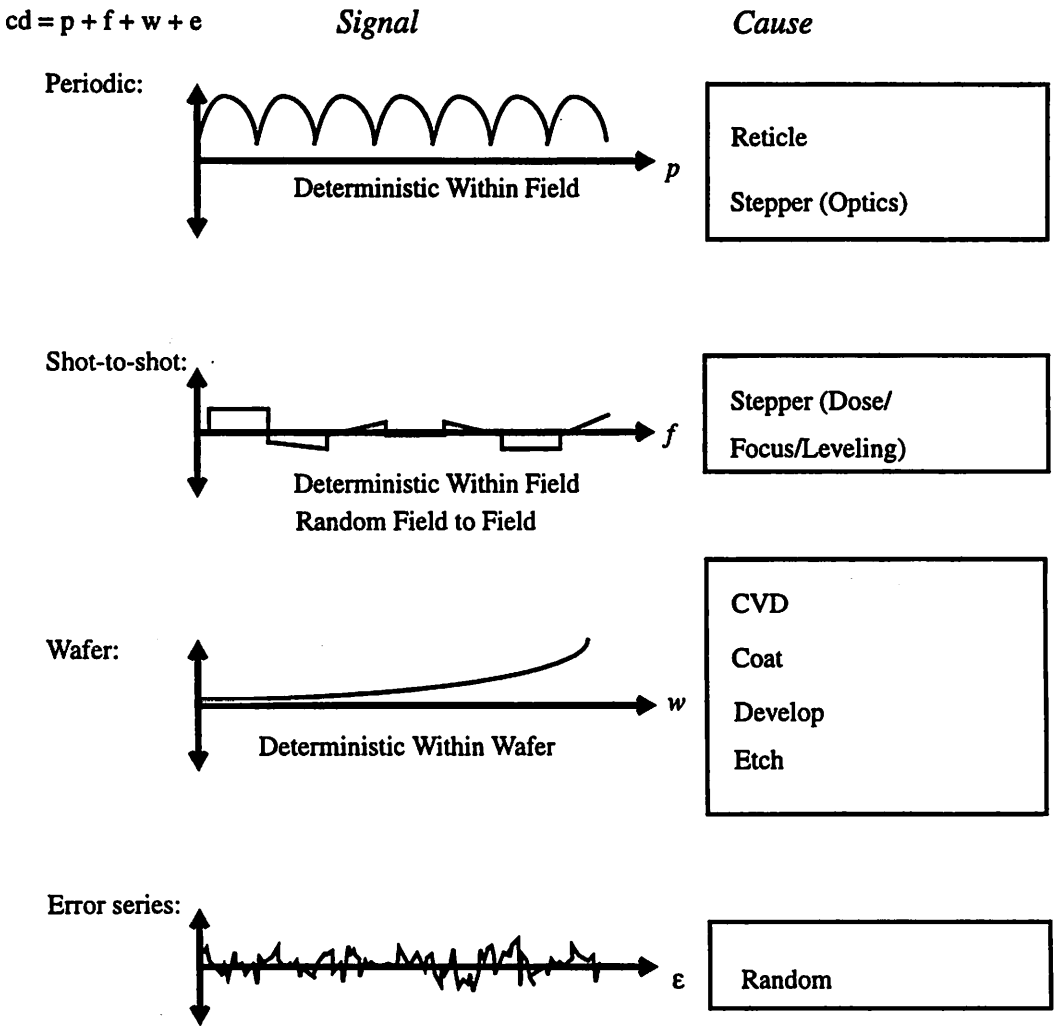


FIGURE 3.2 One dimensional illustration of variability signal components.

3.3 Signal Decomposition

We will now focus on the decomposition of systematic error contributions from the stepper and reticle from the collected raw data. A step-and-repeat exposure system oper-

ates by printing mask images periodically over a wafer, modeled by p. Because the same mask is used in the same optical system for all exposures, CD variability over each exposure field should be identical. In reality, this periodic variability is perturbed by errors in focus and dose, modeled by f, and wafer effects from the rest of processing, modeled by w. The decomposition scheme takes advantage of this error structure. We proceed by first extracting the periodic component $p(x,y)$ through the use of a spatial filter. Then, the contributions of the stepper and reticle will be determined by a separate physical filter.

Several methods are available to isolate the periodic variability $p(x, y)$ from the rest of the error components. Analysis using discrete Fourier transforms in frequency and two-dimensional time series models in space will be presented. In the case of very distinct periodic signals that are the dominant variability component, a simple ANOVA model can be sufficient.

3.3.1 Frequency Domain Analysis for Periodic Component Identification

Taking advantage of its periodic nature of over the wafer, $p(x, y)$ can be approximately decoupled from other sources of variability so that an isolated systematic stepper error can be assigned. $cd(x, y)$ can be represented as a two-dimensional image. In our experiment, this signal is sampled digitally in the x and y directions with period T_x and T_y respectively to yield a 2D discrete signal $cd(n_x, n_y)$ where

$$cd(n_x,n_y) = cd_c(x,y)\Big|_{x = n_xT_x, y = n_yT_y}, n_x = 1, 2, 3, \ldots; \quad n_y = 1, 2, 3, \ldots \tag{3.2}$$

The discrete Fourier transform (DFT) of $cd(n_x,n_y)$ is $CD(k_x,k_y)$. DFTs can be quickly computed by fast Fourier transform (FFT) algorithms [8]. We will follow the convention of using lower case variables to denote digitally sampled space domain signals, and variables in capital to denote the corresponding signal is the spatial frequency domain.

A data decomposition scheme has been implemented to extract the isolated within-field stepper variation from the conglomerated wafer data. This is done by filtering the wafer CD response signal in the spatial frequency domain and then reconstructing the filtered portions via the inverse FFT into $p(n_x,n_y)$ and the remainder components lumped in

$w+f+e$. In this decomposition scheme, first the raw data is transformed to the spatial frequency domain as $CD(k_x, k_y)$ by FFT. A perfectly periodic signal in space only contains energy at the N/n_p multiples of the sampling frequency

$$\begin{aligned} k_x' &= i_x \frac{N_x}{n_{px}}, i_x = \left(-\frac{N_{px}}{2}\right) \ldots -1,0,1,\ldots \left(-\frac{N_{px}}{2}\right) \\ k_y' &= i_y \frac{N_y}{n_{py}}, i_y = \left(-\frac{N_{py}}{2}\right) \ldots -1,0,1,\ldots \left(-\frac{N_{py}}{2}\right) \end{aligned} \tag{3.3}$$

where N_{px} and N_{py} are the numbers of sample points within each repetition of the periodic signal and n_{px} and n_{py} are the numbers of periods of $p(n_x, n_y)$ in either direction. $p(n_x, n_y)$ is periodic so that in spatial frequency domain

$$P(k_x, k_y) = P(k_x', k_y') \ , \tag{3.4}$$

where the primed variables indicate the frequency multiples. A filter can be implemented to select only these harmonics:

$$R(k_x, k_y) = \begin{cases} 1, k_x = k_x', k_y = k_y' \\ 0, \text{otherwise} \end{cases} \tag{3.5}$$

Applying this filter on $CD(k_x, k_y)$ yields:

$$CD'(k_x, k_y) = P(k_x, k_y) + F(k_x', k_y') + W(k_x', k_y') + \varepsilon \tag{3.6}$$

which is a purely periodic signal with spatial periodicity determined by the size of an exposure field. The remaining frequency content of CD is M' and 0-valued at k_x' and k_y'. The first term on the right side of (1.6) is the stepper error contribution and is the desired output of this decomposition. The next two terms are lumped contributions from W and F, which degrade the periodic signal P. However, the values of the degradation terms can be estimated because their frequency spectrum is continuous. Thus, $W+F$ can be estimated from M' by interpolation and can be removed from CD'. In this way, we obtain an estimate p from the inverse FFT of the estimated P. The data decomposition algorithm is summarized in Fig. 3.3.

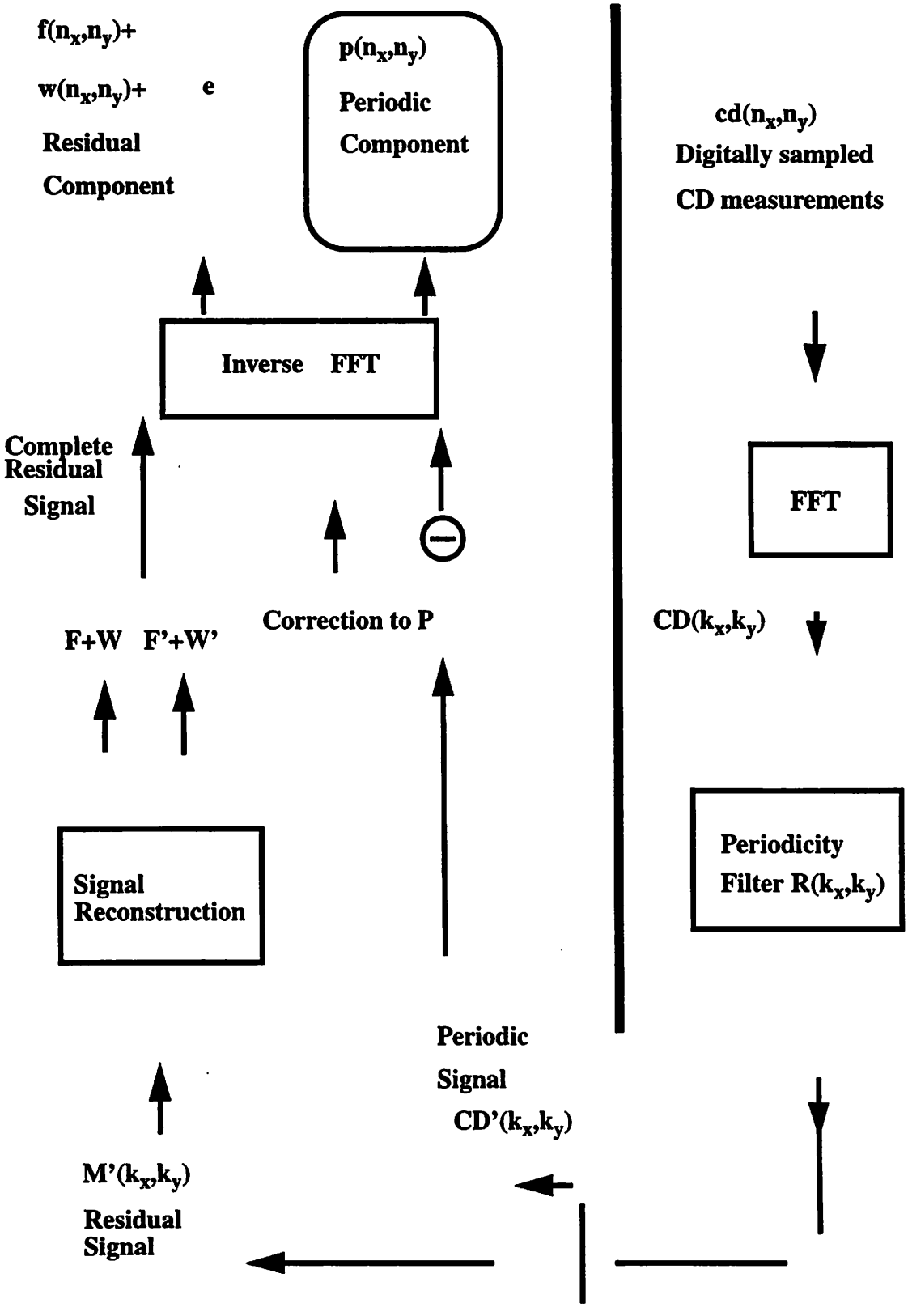


FIGURE 3.3 Data decomposition architecture for the extraction of *p(x,y)*.

Figure 3.4 shows one example of the data extraction applied to wafer CD measurements obtained from the poly gate CD experiment described in Chapter 2. Figure 3.4a shows the three-dimensional mesh plot of raw data $cd(n_x,n_y)$. After filtering, the components $f+w+\varepsilon$ is shown in Fig. 3.4b. The stepper variability p has been isolated and is shown for one period in Fig. 3.4c. Note that without appropriate grouping of the data by fields and subsequent filtering, the entire collection of data points would generate a histogram that appears random. However, given some physical insight into the causes of the variation, it has been possible to isolate some components of the observed variability and identify them as systematic, i.e., a deterministic effect having a specified origin. This presents the opportunity to understand, model, and control variations which until now we have thought of as random.

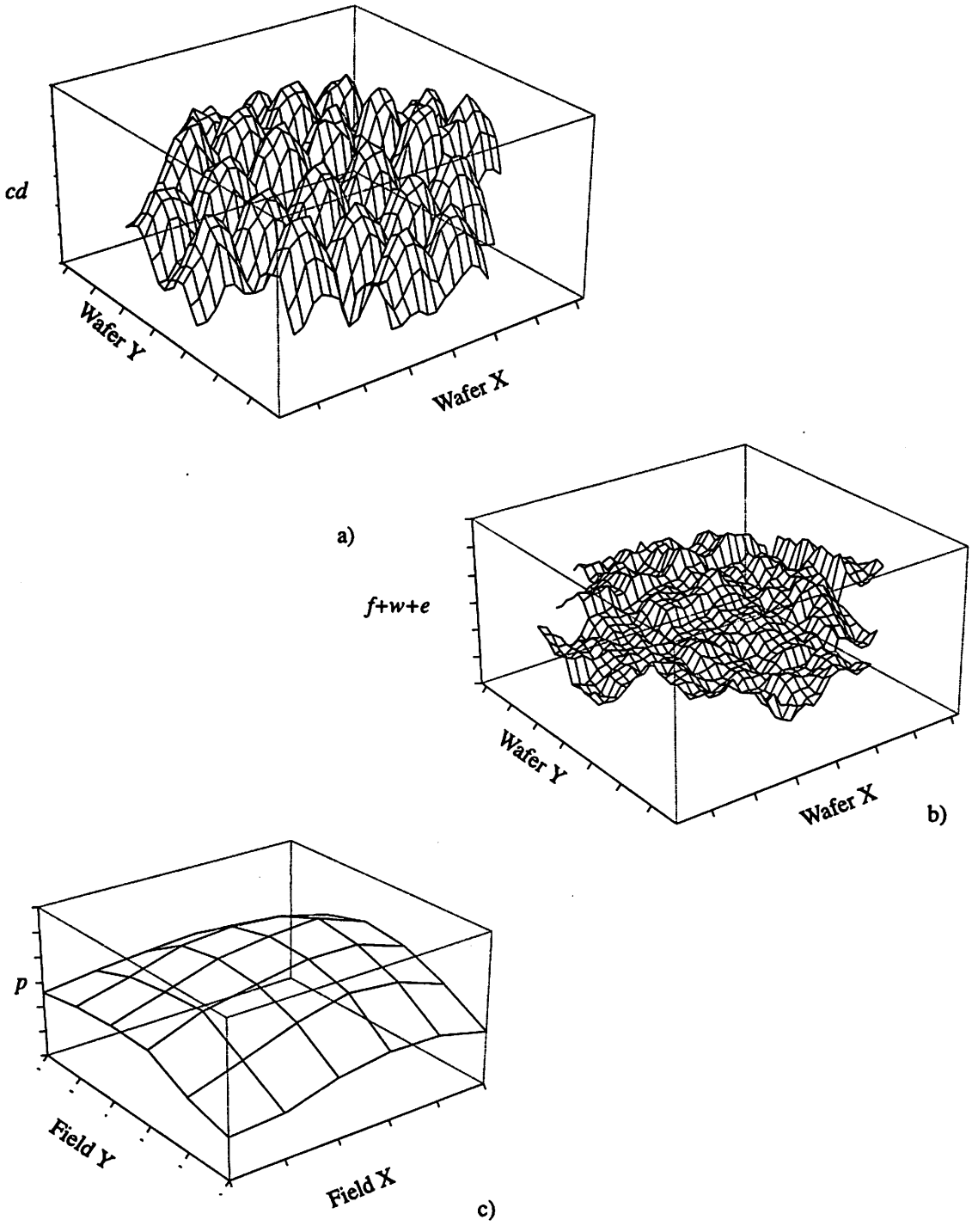


FIGURE 3.4 Three-dimensional mesh plots of: a) Raw CD data over a wafer. b) Residual wafer CD variation $f+w+\varepsilon$. c) Extracted CD variation $p(n_x, n_y)$ within an exposure field. Vertical axis are on identical scales.

3.3.2 ANOVA Model for Periodic Component Identification

Because p is perfectly periodic by definition, it can be estimated by estimating one of its periods. In a simple case, p can also be estimated from an ANOVA model,

$$cd(x', y') = p(x', y') + e(x', y'), \tag{3.7}$$

where $e(x', y')$ is has mean zero and x', y' are coordinates within the field. This condition is satisfied when both the field-to-field and within wafer variabilities behave like random error and can be found in the case where both w and f are negligible within a stepper field. This is generally true for fields printed in the center of the wafers in a well controlled process. In this case, p can be estimated simply by

$$\hat{p}(x', y') = \frac{1}{n}\sum_{i=1}^{n} cd_i(x', y') \quad , \tag{3.8}$$

where i indexes different exposure fields over a wafer, or over several wafers. The extracted $\hat{p}(x', y')$ is the isolated spatial signature of the reticle/stepper combination.

3.3.3 Verification of Signal Decomposition

Independent metrology was used to verify the results of the signal decomposition. In this case SEM metrology was used to measure resist CD's over several fields in the middle of a wafer processed in the same sequence. The assumption is that the periodic component extracted would contain negligible variability from the etch or develop processes. Therefore the extracted results from electrical measurements (which included the etch step) and the SEM measurements of resist CD's (which do not include the etch step) would yield the same results. The results are shown for one decomposition example in Figure 3.5. The correlation coefficient between the two sets of measurements is 0.64, and the spatial features of the two sets of CD contours are also slightly different. The contour in the verification measurements is slightly rotated from the extracted contours. This could be explained by the fact that a period of several months had elapsed between the electrical and SEM measurements and a change had occurred in the process or equipment. Additional verification is provided in Section 6.1.2.3, where it is shown that the intrafield spatial characteristics reflect changes in the equipment state.

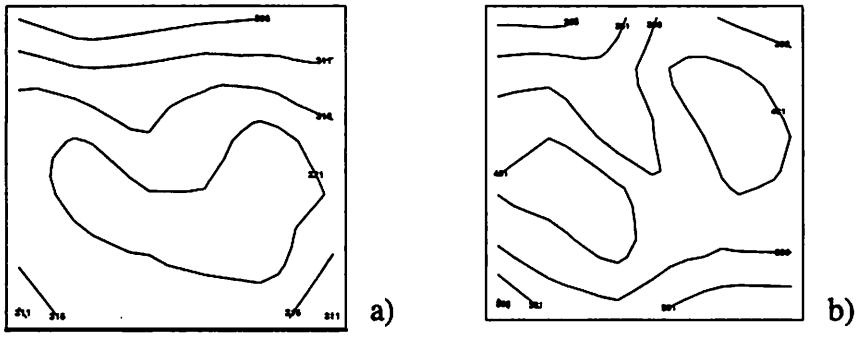


FIGURE 3.5 Countours of periodic component p within stepper field: a) extracted from electrical measurements b) obtained from independent metrology. Contour intervals are the same for both plots.

3.4 Further Causal Decomposition of the Periodic Component

In this section, the analysis will focus on the stepper/reticle pair where a second physically based variability filter is implemented to separate the reticle and stepper error contributions. Experimental verification of the physical filter will be provided. The results of decomposition and its impact on error budget analysis will also be presented. All analysis will be performed across the spatial coordinates x' and y' within one stepper field.

3.4.1 Causal Decomposition Method

The variation in the periodic spatial pattern p has two sources, the stepper/reticle pair. The reduced number of variability contributors to the extracted periodic signal $\hat{p}$ facilitates separation of the variability contributions of the stepper and reticle. The remaining causal decomposition proceeds by considering the printed CD signal p as a response of the reticle CD through the stepper. Spatial variability is accounted for when we consider each of the 36 reticle CD's and the corresponding 36 printed CD's on the wafer. Spatial correlation between the reticle and printed CD's may exist. However, no interaction is considered among the 36 CD's.

In subsequent discussions, we adopt the printed stepper field as the frame of reference. In this frame, mask dimensions are quoted at 1X to account for the stepper magnification. For each point x', y' in the reference frame of the stepper field, the mask CD can be related to the printed CD by an empirical relationship:

$$p(x', y') = A(x', y') + B(x', y') \cdot cd_r(x', y') + C(x', y') \cdot cd_r^2(x', y') \quad (3.9)$$

over a continuous range of CD's of interest, where cd_r refers to the reticle CD dimensions. This second order approximation is justified for a range of reticle CD variation about a nominal value that and can be verified empirically by examining the printed vs. the reticle values on the test reticle. Physically, the coefficients *A*, *B*, and *C* are purely empirical parameters that describe the lumped optical effects of the projection system.

The actual reticle dimensions can be measured directly with great accuracy because the chrome patterns are magnified by a factor of 4 or 5 and placed on a transparent plate. A range of CD patterns from .25 to .45 μm at .05 μm intervals has been placed on the test reticle so *A*, *B*, and *C* can be estimated using linear regression techniques. The extracted value of the within field *p* is used so as to not include variability from other portions of the process. Measurements from the 36 positions over the field are used to generate a set of transfer functions, one at each field position. Collectively, these describe the full field mask CD/printed CD transfer function of the stepper system. Because the nature of the variability is systematic within the stepper field, the parameters in the transfer functions are smooth over the stepper field. In this way the variability is spatially correlated. However, this spatial stepper CD transfer function does not take into account the interaction of optical effects of the features over the spatial extent, as in the case of optical proximity effects. Thus, the variability in this formulation is spatially correlated but not spatially interacting.

The formulation shown in equation (3.9) can be used to decouple stepper and reticle errors. Note that each of the variables in (3.9) are spatially varying. The reticle dimensions cd_r will contain some error in actual cases, and will vary over the field coordinates x', y'. In the absence of reticle error, or in the case when a single value cd_r^t (the superscript *t* indicating the target value) has been measured over the reticle, the quantity $p^t(x',y')$, or *p* evaluated at a single value over x' and y', will describe the printed CD's without reticle error. We rearrange the transfer function so that,

$$p(x', y') = p^t(x', y') + \Delta p c d_r(x', y') \quad \text{and} \tag{3.10}$$

$$\Delta p[cd_r(x', y')] = p[cd_r(x', y')] - p^t(x', y') \quad . \tag{3.11}$$

The variability can be decomposed if we take the sample variance of (3.10):

$$var(p) = var(p^t) + var(\Delta p) + 2 \cdot cov(p^t, \Delta p) \quad . \tag{3.12}$$

In the case where the reticle error is small compared to the change in the printed CD, (3.11) can be locally linearized:

$$\Delta p(x', y') = M \cdot \Delta cd_r(x', y') \tag{3.13}$$

where $M = \frac{d}{dcd_r} p(x', y')|_{cd_r^t}$ is the slope of the transfer function evaluated at the nominal mask CD value. In this form, it is easy to see that the slope factor M magnifies reticle error. In diffraction limited optics the value of M is greater or equal to 1, thus reticle error is always enhanced. Because M could be non-uniform over the stepper field, the spatial correlation between M and Δcd_r could also impact the net spatial distribution of the $M\Delta cd_r$ product.

The first of factored variance terms on the right side of (3.12) is the stepper-alone within-field variance. The second term is the variance from the impact of reticle error on printed CD error. According to our formulation, reticle error is defined as CD's over the reticle differing from the target value p^t. Then, Δp is not a pure reticle nor stepper error because it represents the combined effect of propagating reticle errors through the stepper system, or the stepper/reticle interaction term. Last of the terms in (3.12) is the covariance between the stepper alone error and the reticle induced errors. This accounts for the spatial correlation between the two error terms where the spatial distribution of individual errors could enhance or cancel each other as to increase or reduce the net error.

3.4.2 Implementation of Causal Decomposition

Causal decomposition of $p(x', y')$ is implemented following the procedure described in the previous section. First the CD transfer function of the stepper is mapped

versus field coordinates for .30 to .45 μm nominal mask CD's. Actual measured reticle values will be used to reduce fitting error. Figure 3.6 shows a subset of the transfer curves over positions in the field. For clarity, only 3 field positions are shown. Symbols mark data points derived from spatial filtering, which are essentially an averaged value for the fields sampled. Therefore, we assume that there is very little error associated with them. A second order polynomial is fitted to each of the transfer curves in order to obtain continuous values around the nominal CD, the results of which are also plotted. A collection of such fitted transfer functions will constitute an empirical form of $p(x', y')$ as a function of reticle dimensions described in (4) and will allow subsequent causal decomposition.

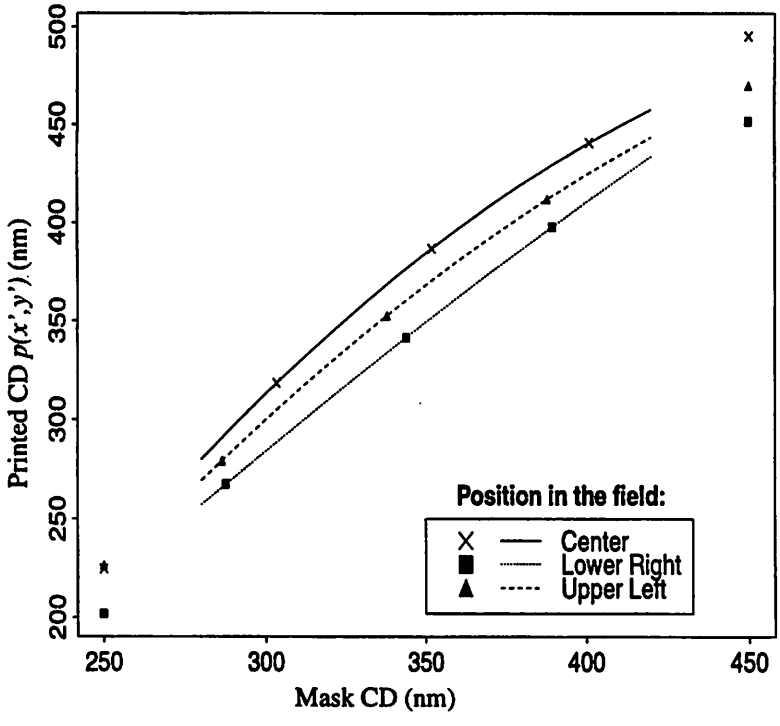


FIGURE 3.6 Mask vs. printed CD transfer function. Symbols mark measured data and lines show fitted values.

Before proceeding with the decomposition, we examine the spatial behavior of the transfer function parameter estimates. Figure 3.7 shows the spatial variation of the parameters $A(x',y')$, $B(x',y')$, and $C(x',y')$ from (4) corresponding to plots a), b), and c). The spatial variation is radially symmetric, as one would expect for an optical system. As a result of both spatial and causal decomposition, the *variation* of these parameters describes the

variability within the stepper field only, and is most likely due to optical effects. However, because these parameters are empirical in nature, they are a lumped description of the aberrations and nonuniformities in the stepper system, and thus have no physical significance as such. Also implicit in the decomposition is that the process conditions are fixed. While the decomposition has isolated within field stepper variability from other process variability, the CD transfer function will still depend on stepper defocus, as well as the resist and develop process operating points, and will be demonstrated in a later section.

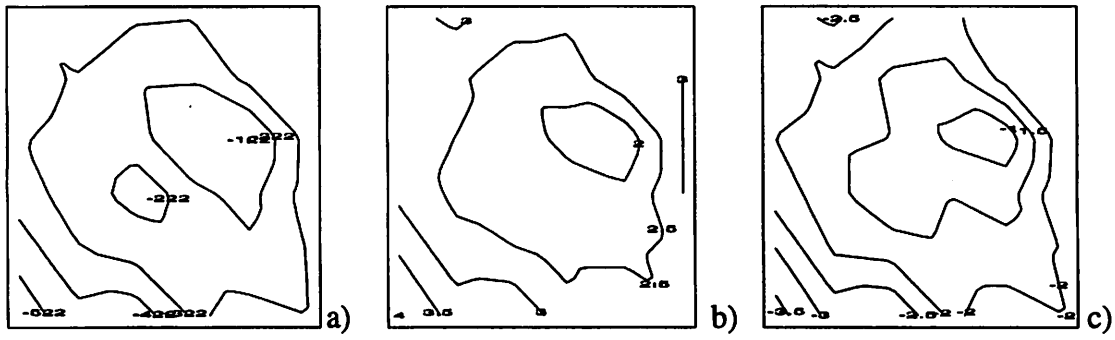


FIGURE 3.7 Contours over the stepper field of fitted parameters a) A, b) B, and c) C of the CD transfer function.

With the determination of the empirical transfer function, poly CD variation can be

factored by cause according to (3.12) and (3.13). Figure 3.8 shows the contours of the ret-

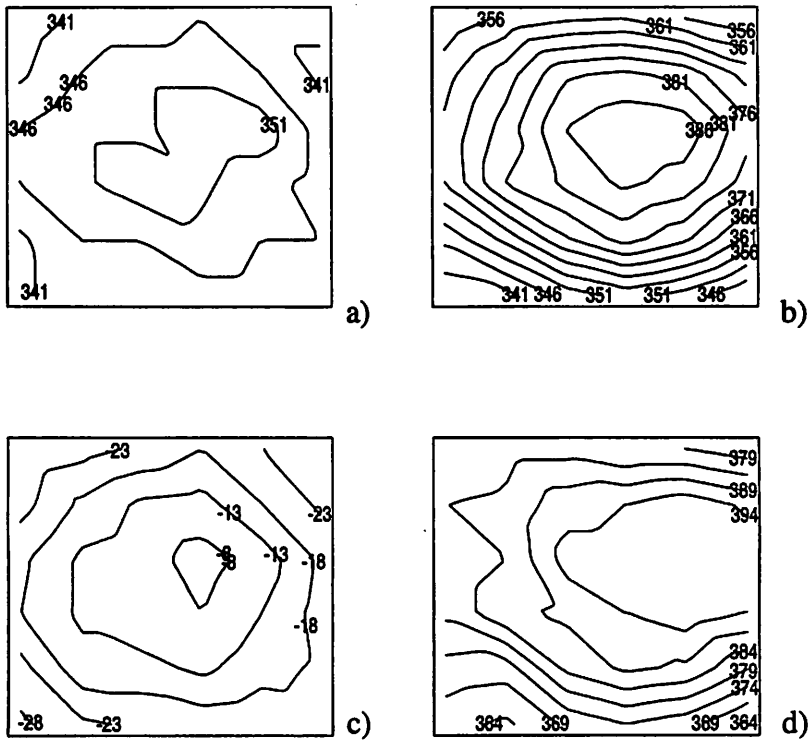


FIGURE 3.8 Contour plots of CD's over the stepper field after causal decomposition: a) Measured reticle dimensions at 1X. b) Stepper contribution. c) Reticle induced error. d) Resulting stepper alone error. (Interval between contours is constant.)

icle CD cd_r, the printed CD p, the reticle error correction Δp, and the stepper alone term p^t evaluated at the target dimension of 350 nm, all are spatial quantities. The variance contributions are tabulated below:

Table 3.1 Variance Summary of Causal Decomposition of $p(x,y)$

Symbol	Variance (nm^2)	Percentage
p^t	107.5	48
Δp	36.1	16
2*Cov	79.4	36
Total	223.0	100

From (1.11) and (3.13), the variability introduced through the reticle/stepper system has two parts: the stepper alone component p^t and the reticle induced error $\Delta p \cong M \Delta cd_r$. Of the latter M is also a stepper dependent term. In fact, it is obvious that

M is the first derivative of the CD transfer function (3.9) with respect to cd_r. Ideally, M=1 over all field coordinates, or perfect transfer of the reticle CD pattern. For reference, the reticle CD variance is only 17.9 nm^2, which is smaller than the reticle induced error contribution term Δp. This can be explained by examining (3.13), which shows that the reticle error is in fact magnified by M, the local slope of the transfer function, shown in Figure 3.9. In this example, the spatial variation of M tended to be similar to cd_r. This and the fact

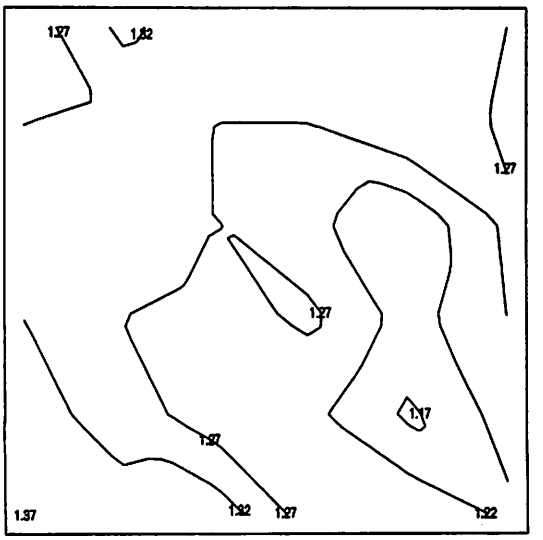


FIGURE 3.9 Contour of the reticle error gain factor M over the field.

that M is greater than one for these process conditions leads to the net effect that reticle error on printed CD was enhanced. Similarly, there is a covariance term in the table above due to correlation between the reticle induced and the stepper-alone errors. This is due to the fact that the spatial distributions of p^t and Δp are similar and additive with the effect of increasing net CD variability. From the table above the covariance contributes 36% of the total within-field variance. While p^t depends on the stepper optics and may not be easily manipulated, it is possible to manipulate the spatial distribution of Δp by manipulating reticle CD error. In this way, it would be possible to tune the correlation between the variance components so as to reduce or even eliminate net CD variation.

Following the previous section, the stepper alone component of CD variability can be extracted. Figures 3.10 and 3.11 show the field CD contours for the stepper-alone variability as a function of the target reticle dimensions from 310 to 390 nm for both isolated hor-

izontal lines as well horizontal lines in an array. As expected, a trend of decreasing variability can be observed with increasing reticle dimensions. Similarly, M can be calculated from the fitted transfer function and its contours plotted as a function of reticle pattern dimensions (Figure 3.12 and 3.13). Again, a trend of decreasing variability and decreasing mean value of M is seen versus reticle dimensions.

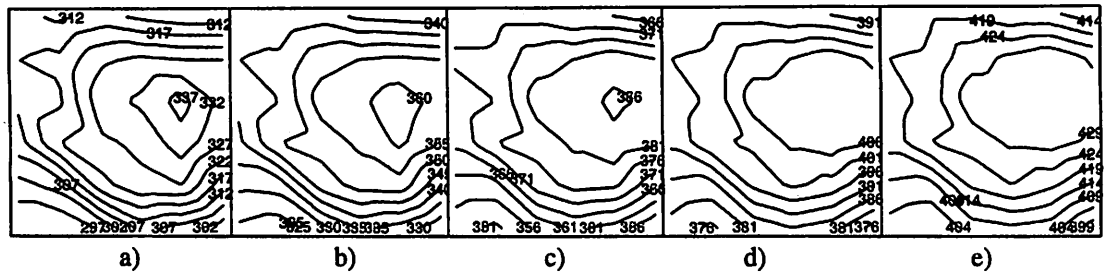


FIGURE 3.10 CD contour plots of the stepper alone contribution to CD error for horizontal isolated lines. Each of plots a)-e) evaluated at nominal reticle CD's from 310nm to 390nm by increments of 20nm.

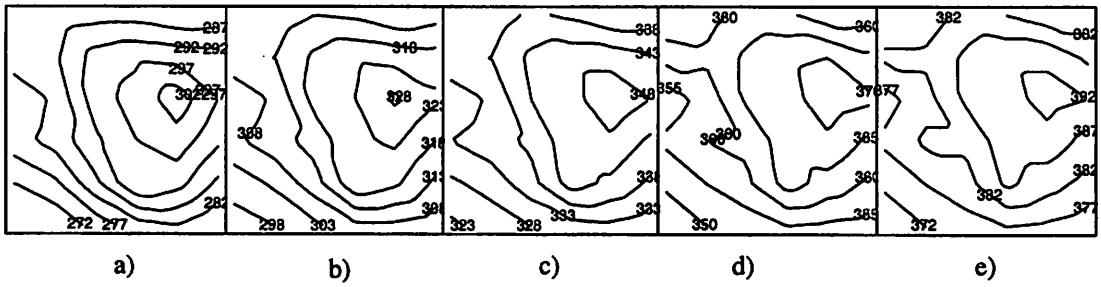


FIGURE 3.11 CD contour plots of the stepper alone contribution to CD error for horizontal lines in an array. Each of plots a)-e) evaluated at nominal reticle CD's from 310nm to 390nm by increments of 20nm.

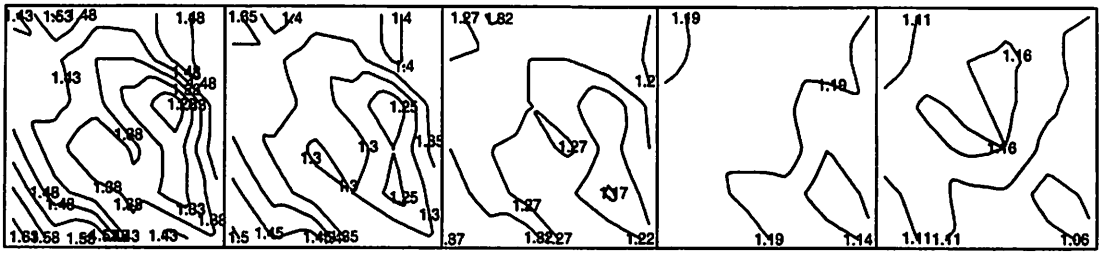


FIGURE 3.12 Contour plots of the reticle error enhancement factor M for horizontal isolated lines. Each of plots a)-e) evaluated at nominal reticle CD's from 310nm to 390nm by increments of 20nm.

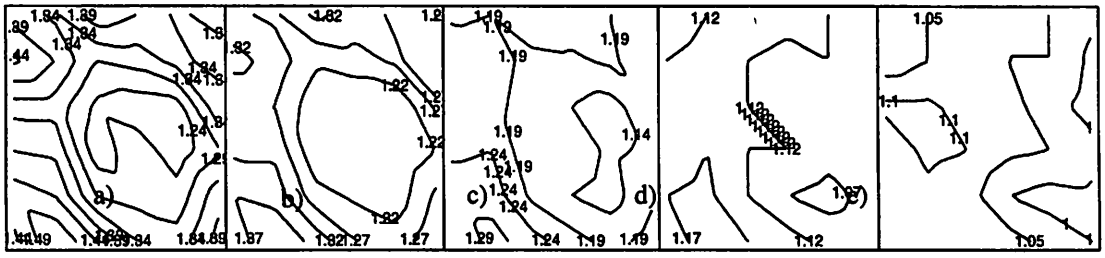


FIGURE 3.13 Contour plots of the reticle error enhancement factor M for horizontal lines in an array. Each of plots a)-e) evaluated at nominal reticle CD's from 310nm to 390nm by increments of 20nm.

3.4.3 Verification of Causal Decomposition of Periodic Component

Causal verification was carried out in a separate experiment. The stepper error component p^t accounts for the variability introduced by the stepper alone. (What the stepper would print given no reticle CD variability.) A number of reticles were fabricated. From these a "uniform" reticle with very little CD variability was chosen and printed. The resulting comparison is shown in Figure 3.14. As expected, the as-printed variability from the "uniform" reticle agrees well with the extracted stepper variability from the original reticle (r^2=0.92, n=36).

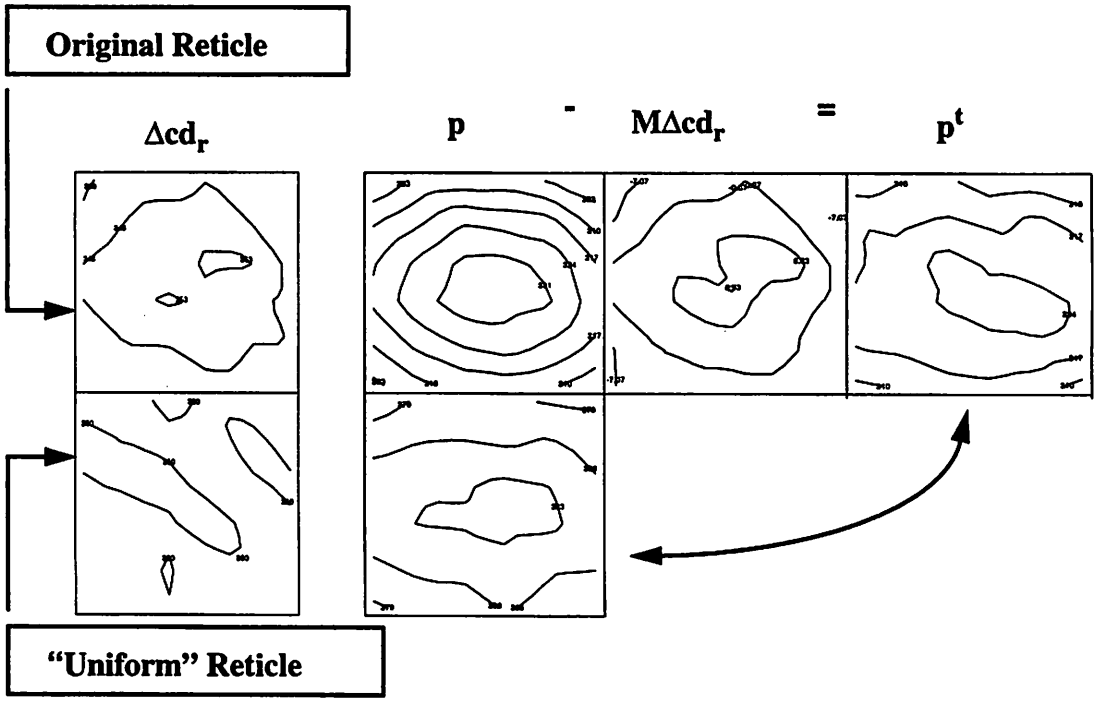


FIGURE 3.14 Causal verification by comparing decomposition results using two reticles. All contours are within the exposure field and equal intervals. The columns are: reticle error (1X), intrafield error, reticle induced error, and extracted stepper error. The contours in the uniform reticle represent the same value.

3.5 Stepper/Reticle Decomposition Summary

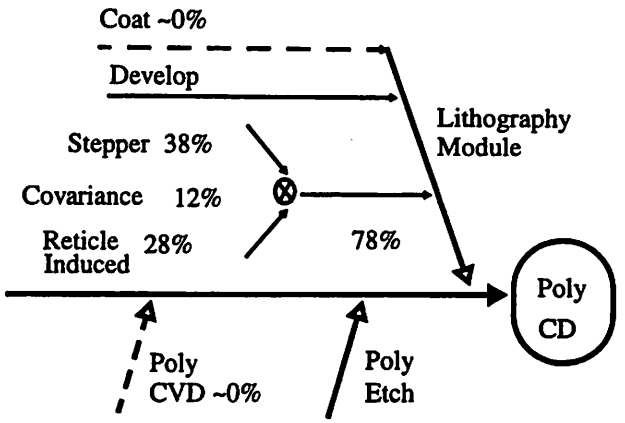


FIGURE 3.15 Error budget diagram for .35μm horizontal isolated lines as a fraction of total wafer CD variance.

We have developed a methodology which decomposes and categorizes systematic spatial CD variability into individual equipment contributions and is advantageous for several reasons: 1) It identifies bottlenecks in manufacturability and offers an opportunity to manipulate various variability components through the choice of equipment hardware and equipment settings. 2) The isolated equipment variability components can be used as a benchmarking metric between equipment choices. 3) The correlation between spatial variability can be manipulated to reduce net variability.

Systematic, or roughly time-invariant process error in the form of poly CD variability is captured by sampling spatial variability. The large amounts of data required for this analysis was provided by electrical measurements. However, raw electrical measurements contain confounded sources of error from the fabrication sequence. Decomposition of the net CD error was achieved by a series of statistical and physical filters. The statistical spatial filter isolated the systematic variability within the exposure field while the physical filter captured the reticle CD/printed CD transfer function of the stepper system. From this empirically determined transfer function, CD variability was decomposed into stepper-alone and reticle-induced components. Because these variability components are spatial in nature, their correlation also affects the net CD variability. In the next chapter, we will address metrology variability introduced by automated SEM measurements.

Chapter 4

Poly Etch/Develop Decomposition Using SEM

Chapter 3 demonstrated Statistical Metrology on one component of the poly CD pattern transfer sequence. Using this methodology we were able to assign poly CD variability to individual equipment contributions from the stepper and reticle. Raw data from electrical metrology was decomposed using a series of filters developed through experimental and statistical methods. In this chapter, the methodology is extended to the determination of CD error contributions from the remaining components in the patterning sequence: resist develop and plasma etch. Automated top-down SEM CD measurements can be used to provide estimates of the develop and etch CD variability. However, metrology error, specifically in the form of sequencing artifacts, needs to be characterized and removed. An updated error budget diagram is shown in Figure 4.1, where the dashed entries have already been determined.

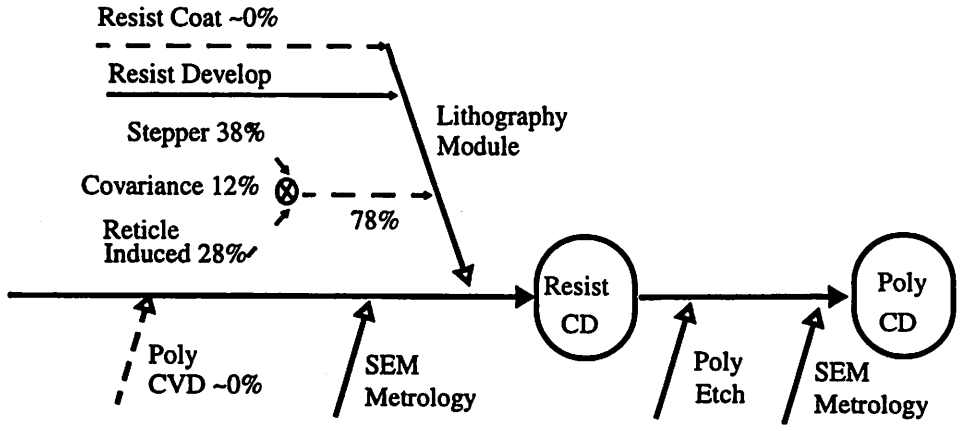


FIGURE 4.1 Error budget diagram for .35μm poly gate patterning.

4.1 SEM CD Metrology

4.1.1 Sequencing Trend and Autocorrelation

Electrical measurements confound the etch and develop contributions to CD variability because the measurements are taken after the entire pattern transfer sequence. SEM CD measurements, on the other hand, can sample variability at intermediate steps in the pattern transfer sequence: after resist develop and after etch. This and the increased efficiency of automated CD SEM's make SEM metrology a viable means of sampling CD variability. However, experimental evidence suggests that SEM CD precision is degraded by the sampling sequence as well as structure proximity.

Figure 4.2 shows sequential measurement trends from an automated CD SEM. The observed errors take two forms. First, systematic trend or drift vs. sequence in the measured values can be observed. The bottom plot in Figure 4.2a shows the top-down SEM measurement values of CD's over a wafer vs. sampling sequence. By comparison, the top plot shows the measurement from the same locations on the wafer made with alternate metrology. A clear trend can be observed for the SEM sequence. The magnitude of the errors introduced by the SEM are comparable to the cross-wafer variability observed (~10nm). Moreover, these errors are confounded with the true spatial variation on the wafer.

Another form of sequencing error is autocorrelation in the measurement error. Analysis of the measurement error series, after removal of the measurement trend and the estimated values of the CD, shows significant autocorrelation. Finally, these two sequencing errors, the trend and the autocorrelation, must be accounted for in order to estimate CD

variability over the wafer.

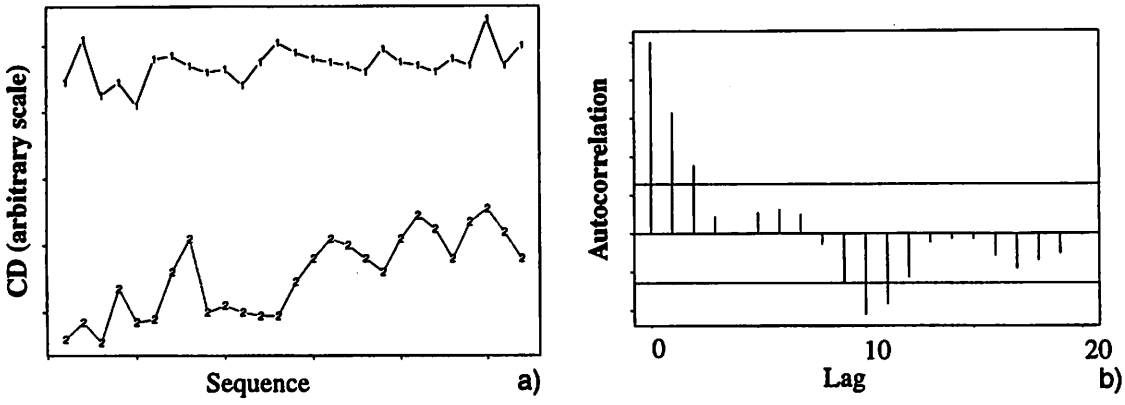


FIGURE 4.2 Metrology sampling artifacts in forms of: a) systematic trend observed in the bottom graph compared to the reference data in the top plot and b) autocorrelation structure in the measurement error series.

4.1.2 Methodology for SEM Measurement Error Analysis

Following the signal decomposition strategy outlined in Section 3.2, the across-wafer variability signal w is extracted. This can be done simply by fitting a low order polynomial surface to the spatial variability. From film thickness measurements and sensitivity analysis in Chapter 2, it was determined that the effects of resist and poly film thickness variations on CD variation were negligible. Of the contributors to the error, only the develop and etch components remain unquantified. From Figure 4.1, it is obvious that the wafer component w obtained from resist CD measurements after develop includes primarily the variability from develop, while the variability of the difference between the after-etch and after-develop CD's can be assigned to the etch step. Automated, top-down SEM metrology can provide high quantities of samples nondestructively, and can be used to measure non-conductive patterns such as photoresist. Thus, the develop and etch variability can be simply determined by taking SEM samples of CD's after develop and after etch.

As shown in Figure 4.1, SEM metrology can be used to sample at intermediate points in a process sequence to avoid confounding etch and develop errors. However, it may introduce error from the sampling sequence which could confound the spatial variability being measured. Evidence suggests that SEM CD measurements can be tainted by sequencing artifacts such as trend and autocorrelation. A methodology is developed in this

chapter to characterize the CD variability introduced by automated SEM measurements and remove this variability through statistical filtering. This technique is applied to estimate variability from resist develop and plasma etch.

4.2 Poly CD Experiment for SEM Sampling

A .35 μm poly gate patterning process using I-line lithography was used to demonstrate the methodology. Following Chapter 2, a test mask of electrically measurable structures was used to sample CD over the wafer. Experiments were performed to sample the within-wafer variability of patterned resist as well as patterned poly CD structures using automated SEM measurements. Another series of sampling experiments were carried out to determine the measurement trend.

4.2.1 Test Reticle and Processing Sequence

The reticle and process sequence follow the description given in Chapter 2. A test mask was designed to sample systematic CD variability at a fixed operating point in the poly gate patterning sequence. Spatial variability was sampled by repeating CD patterns over a reticle in a 6X6 grid spanning a 22x22mm stepper field. This reticle pattern was repeated 21 times over a 6" wafer. The dimension and proximity conditions of the CD patterns were varied in each location on the reticle to sample systematic variability due to mask CD variations and proximity effects. Linewidth topologies of both isolated and dense arrays were considered with drawn .35μm lines at 1μm pitch. The linewidth structures are electrically measurable using the Kelvin 4-point measurement design. The sheet resistance was determined by adjacent Van der Pauw structures to compensate for errors in film resistivity.

Short loop processing on flat wafers minimized the number of confounding processing effects. Poly deposition and oxide growth on bare Si wafers yielded a stack of 2000Å undoped poly / 1000Å SiO_2. Polysilicon was doped after patterning for electrical testing. Resist was coated, patterned with the reticle, and developed in an I-line lithography process. Three sets of 4 wafers were exposed at nominal dose +/- 10mJ. Film patterning was accomplished by plasma etching.

4.2.2 Automated SEM Metrology Setup and Operation

Automated SEM metrology was provided by Advanced Micro Devices in an OPAL 7830i®. The information provided in this section can be found in the tool operation manual [9]. In this metrology system, secondary electrons are detected by a collector mounted in the beam column over the sample. This is in contrast to previous generations of tools where a pair of detectors is placed to the sides of the sample at an angle. The side detector setup required signal averaging from both detectors to account for the asymmetrical collection of electrons from the sides of specimen. It was also ineffective for the detection of electrons emitted around obstructing structures, such as those electrons from the bottom of the space between closely spaced line pairs.

Autofocus is performed in two parts. The focus levels at the measurement sites are determined by optical microscope before SEM measurements. Within-die target selection is done by pattern recognition. With positional information over the wafer, die to die selection is accomplished. CD values are extracted from the intensity scan of a structure. Three CD values are given for each measurement location. Reference points are determined by points of maximum slope along the intensity trace. This is the estimate for the CD value at the middle, or half height of the structure. The CD value at the bottom of the structure is estimated by linear extrapolation of the point of maximum slope with the wafer surface. The CD value on the top of the structure is estimated by the distance between intensity peaks. The accelerating voltages used were 0.6kV for resist structures and 1.2kV for polysilicon.

4.2.3 Sampling Schemes and Decomposition Strategy

The trend and correlation errors introduced by the metrology can be confounded with the spatial variability being measured. A separate set of experiments and analysis was conducted to characterize the measurement errors. Section 4.3.1 will describe how statistical analysis and modeling coupled with experiments in the CD sampling sequence are used to differentiate the process induced variability signatures from the measurement errors. This analysis is implemented in the form of a data filter. Raw CD measurements are filtered to yield accurate estimates of CD variability. By comparing the measurements after resist

patterning, the develop step contribution to CD variability is assessed using the filtered data.

To sample the intra-wafer variability between exposure fields, one measurement in the center of each field was taken yielding a total of 21 measurements over the wafer. The sampling algorithm on the CD SEM ordered the measurements so that adjacent points were measured in sequence. This had the effect of confounding measurement errors with the true spatial CD variability on the wafer. Another experiment was carried out to characterize the measurement error. Two sequences of measurements were taken on the same structures over the wafer in different order to break up the spatial structure and to isolate the sequencing artifacts. Two rows of thirty measurements each span the center of a wafer. The sixty points were visited in two different sequences, as shown in Figure 4.3. This was done for all the wafers, on resist as well as poly CD's. Because of the pattern recognition and targeting algorithm of the measurement tool, it was not possible to program a randomized sampling pattern.

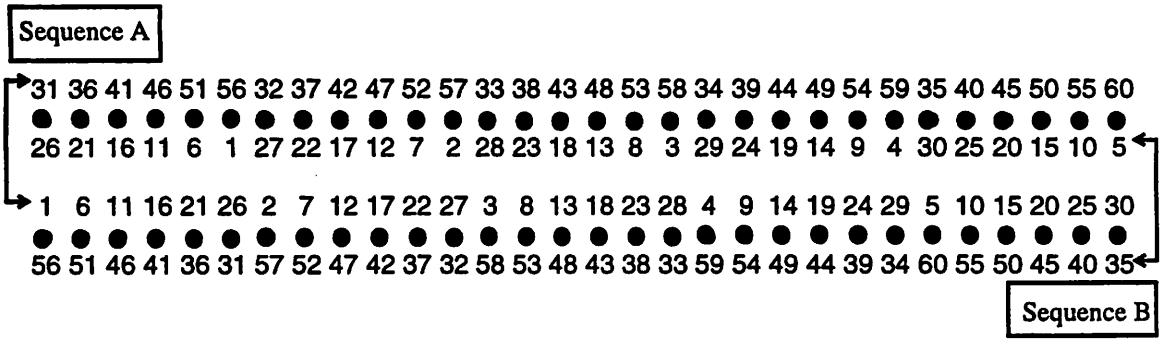


FIGURE 4.3 Sequence order versus location. Sixty positions (symbol "●") are sampled along two rows on the wafer with 30 positions each. The 60 positions are visited by two sequences in different order.

4.3 Analysis of SEM CD Variability Decomposition

Following the methodology described in Chapter 3, decomposition of CD variability is accomplished through both spatial and causal filtering. The general variability model over the wafer is restated here:

$$cd(x, y) = p(x, y) + f(x, y) + w(x, y) + \varepsilon \ , \tag{4.1}$$

where the variables x and y are spatial coordinates over the wafer. The signal $p(x,y)$ is the within exposure field signature from the stepper/reticle pair and is perfectly periodic field to field. The exposure to exposure variability $f(x,y)$ from dose, focus, or leveling error is also systematic within the field. Variability over the wafer $w(x,y)$ is continuous and slowly varying. Random error, ε, is normally distributed and has no spatial structure. These components are assumed to be additive.

4.3.1 Analysis of Intra-wafer Poly CD Variation

The separation of the variability into systematic components within the exposure field and between exposure fields is effectively a separation of the stepper/reticle variation (within field) and the variation from the rest of the process as slowly varying variability over the wafer (intra-wafer). The sampling pattern described in Section 4.2.3 achieves this separation by fixing the sampling point within the exposure field and removing, to first order, the contributions from p and f. The variability of interest, w, includes variability from film deposition, resist coat resist develop, and etch. From independent measurements, the intra-wafer variation in film thickness for both the resist and the poly films are <1% (Chapter 2). Their effects on CD variability are negligible. The components of w are then from etch and develop. Again assuming additivity and implicit spatial dependence of the variables, w, in this specific case w_{poly} can be further decomposed as:

$$w_{poly} = w_{develop} + w_{etch} \ . \tag{4.2}$$

Using automated SEM, CD measurements can be taken over the wafer after resist patterning and after etch. The sampling plan described above eliminated the intra-field contributions p and f. The signal w_{poly} is the wafer component of final poly variation and can be estimated from the final poly measurements, $w_{poly}^{measured}$, which also includes measurement error, the superscript "measured" indicating measured data. Because there is no significant contribution from the resist or poly thickness variations, $w_{develop}$ can be estimated from the resist CD measurements $w_{resist}^{measured}$. Summarized in the following equations:

$$w_{poly}^{measured} = w_{poly} + \delta_{poly} + \xi_{poly} \text{ and}$$

$$w_{resist}^{measured} = w_{develop} + \delta_{resist} + \xi_{resist} . \quad (4.3)$$

Both $w_{poly}^{measured}$ and $w_{resist}^{measured}$ are measured quantities, and contain measurement errors in the forms of δ, or a systematic trend with measurement sequence, and ξ, a random error with zero mean. The trend δ can be characterized and subtracted, as is done in the following section. The parameters w_{poly} and $w_{develop}$ can then be estimated by regression over the spatial variables x and y over the wafer. Finally, the etch variability w_{etch} can be obtained by subtracting $w_{develop}$ from w_{poly}. Therefore, in principle, the quantities $w_{develop}$ and w_{etch} can be measured separately after develop and after etch. However, SEM metrology itself needs to be characterized in terms of δ before this analysis can be accomplished. For reference, the variables used in this chapter are summarized in the following table:

Table 4.1 Variable Definitions in Section 4.3.1

Variable	Definition
w_{poly}	**parameter for intra-wafer poly CD variability**
$w_{develop}$	**parameter for develop contribution to intra-wafer CD variability**
w_{etch}	**parameter for poly etch contribution to intra-wafer CD variability**
$w_{resist}^{measured}$	**measured values for intra-wafer resist CD variability**
$w_{poly}^{measured}$	**measured values for intra-wafer poly CD variability**
$\hat{w}_{develop}^{measured}$	**estimated values for develop contribution to intra-wafer CD variability**
$\hat{w}_{poly}^{measured}$	**estimated values for intra-wafer poly CD variability**
$\hat{w}_{etch}^{measured}$	**estimated values for poly etch contribution to intra-wafer CD variability**
δ_{resist}	**measurement trend error for resist CD's**
ξ_{resist}	**measurement random error for resist CD's**
δ_{poly}	**measurement trend error resist for CD's**
ξ_{poly}	**measurement trend error resist for CD's**

4.3.2 Analysis of SEM CD Measurement Errors

In order to characterize the measurement errors δ and ξ, separate sequencing experiments, described in the previous section, are conducted on both the poly and resist levels. The following general model is assumed for the measured values in this experiment:

$$cd_i^{measured} = cd_i^{actual} + \delta_i + \xi_i \tag{4.4}$$

Here $cd^{measured}$ are the measured values, cd^{actual} the actual values on the wafer, δ the mean offset due to measurement drift, and ξ the measurement error series, a random variable with zero mean. The subscript i denotes the index of the sequence of the measurements. The quantity $w_{develop}$ is systematic over x and y while the measurement offset and noise series may have correlation structure over the sampling sequence i. This formulation is general, and is applied in the following sections to both poly and resist CD's. Note that in this experiment, we attempt to estimate δ and ξ, while the actual values cd^{actual} are not of interest.

A randomized sampling sequence over the wafer would separate the measurement error from the actual CD values. However, the sampling algorithm used by the tool in this experiment arranges the sequence to measure neighboring points. This may confound the spatial structure that we wish to observe with the measurement error. In the following sections we develop an experiment and methodology to estimate the mean drift values δ_i as a function of sampling sequence so that it can be subtracted from the raw data.

The effects of measurement sequence are determined by a separate sampling experiment. First, the resist CD measurement errors are characterized. On the resist two sets of measurements were taken on 60 sites on the wafer. For this experiment the measurement sites were different from the sites sampled to determine the wafer variability. The two sets of measurements traverse the identical sites in different order so that we obtain vectors of the measured values

$$\boldsymbol{cd}_A^{measured} = \boldsymbol{cd}_A^{actual} + \boldsymbol{d} + \boldsymbol{z} \quad \text{and} \tag{4.5}$$

$$\boldsymbol{cd}_B^{measured} = \boldsymbol{cd}_B^{actual} + \boldsymbol{d} + \boldsymbol{v} \ . \tag{4.6}$$

Variables in bold denote 60x1 vectors with the *i*th element corresponding to the *i*th measurement in that sequence. The vectors $\boldsymbol{cd}_A^{measured}$ and $\boldsymbol{cd}_B^{measured}$ denote vectors of measured values in each of the two sampling sequences *A* and *B*. Their entries are the measured CD values corresponding to the measured sites in the sequence. The vectors $\boldsymbol{cd}_A^{actual}$ and $\boldsymbol{cd}_B^{actual}$ are the vectors of the actual values on the wafers. The entries of $\boldsymbol{cd}_A^{actual}$ and $\boldsymbol{cd}_B^{actual}$ are permuted with respect to each other to reflect to the different sampling sequence. The permutation can be represented by a 60x60 matrix operator $\boldsymbol{R}$ such that $\boldsymbol{Rcd}_B^{actual} = \boldsymbol{cd}_A^{actual}$. Entries of $\boldsymbol{d}$ are the sequence trend values δ_i. The vector $\boldsymbol{d}$ is the systematic measurement sequence trend, and is therefore the same sequence for both $\boldsymbol{cd}_A^{measured}$ and $\boldsymbol{cd}_B^{measured}$. The vectors z and v are different realizations of the error series ξ.

The measurement sequence trend $\boldsymbol{d}$ is of interest and is estimated as follows. The operator *R* is applied to $\boldsymbol{cd}_B^{measured}$ so that:

$$\boldsymbol{Rcd}_B^{measured} = \boldsymbol{Rcd}_B^{actual} + \boldsymbol{Rd} + \boldsymbol{Rv} \quad \text{and}$$

$$\boldsymbol{Rcd}_B^{measured} = \boldsymbol{cd}_A^{actual} + \boldsymbol{Rd} + \boldsymbol{Rv} \tag{4.7}$$

The measurement trend $\boldsymbol{d}$ and error series v are randomized by $\boldsymbol{R}$ so that their respective correlations to sequence are lost. $\boldsymbol{Rd}$ and $\boldsymbol{Rv}$ are not correlated to each other and are lumped into one random vector

$$\boldsymbol{u} = \boldsymbol{Rd} + \boldsymbol{Rv}$$

where $\boldsymbol{u}$ has a sample average equal to the sample average of $\boldsymbol{d}$ and (4.7) becomes

$$\boldsymbol{Rcd}_B^{measured} = \boldsymbol{cd}_A^{actual} + \boldsymbol{u} \ . \tag{4.8}$$

Then, subtracting (4.8) and (4.5) results in

$$cd_A^{measured} - Rcd_B^{measured} = d - u \equiv e_A \quad (4.9)$$

Similarly, by defining a reverse operator Q such that $cd_B^{actual} = Qcd_A^{actual}$, we obtain

$$cd_B^{measured} - Qcd_A^{measured} = d - u' \equiv e_B$$

where u', the sum of a randomized d and a random error series with zero mean, and is analogous in this derivation to u in (4.7). The vectors e_A and e_B contain the trend d with the addition of a random vector u or u'. The trend d in vectors e_A and e_B can then be estimated by linear regression techniques. For reference, the variables used in this section are summarized in the following table:

Table 4.2 Variable Definitions in Section 4.3.2

Variable	Definition
$cd_i^{measured}$	sequence of measured CD values
cd_i^{actual}	sequence of actual CD values
δ_i	sequence of measurement trend error
ξ_i	sequence of measurement random error
$cd_A^{measured}$	60x1 vector of measured CD values, sequence A
$cd_B^{measured}$	60x1 vector of measured CD values, sequence B
cd_A^{actual}	60x1 vector of actual CD values, sequence A
cd_B^{actual}	60x1 vector of actual CD values, sequence B
d	60x1 vector of measurement trend error values δ_i
z	60x1 vector of a realization of ξ
v	60x1 vector of a realization of ξ
R	60x60 matrix of a matrix to permute sequence B
Q	60x60 matrix of a matrix to permute sequence A
u	60x1 vector of a lumped random variable in sequence A
u'	60x1 vector of a lumped random variable in sequence B
e_A	60x1 vector containing trend series d from sequence A
e_B	60x1 vector containing trend series d from sequence B

4.4 Poly CD Develop Error Through Resist Measurements

4.4.1 Resist CD Measurement Sequence Error

The resist CD measurement errors are characterized to yield estimates for the develop contribution $w_{develop}$. First, the effects of measurement sequence are determined by a separate sampling experiment. On the resist two sets of measurements were taken on 60 sites on the wafer. For this experiment the measurement sites were different from the sites sampled to determine the wafer variability. The sequencing experiment was performed on each of the three wafers in Figure 4.3 and the result examined to detect the

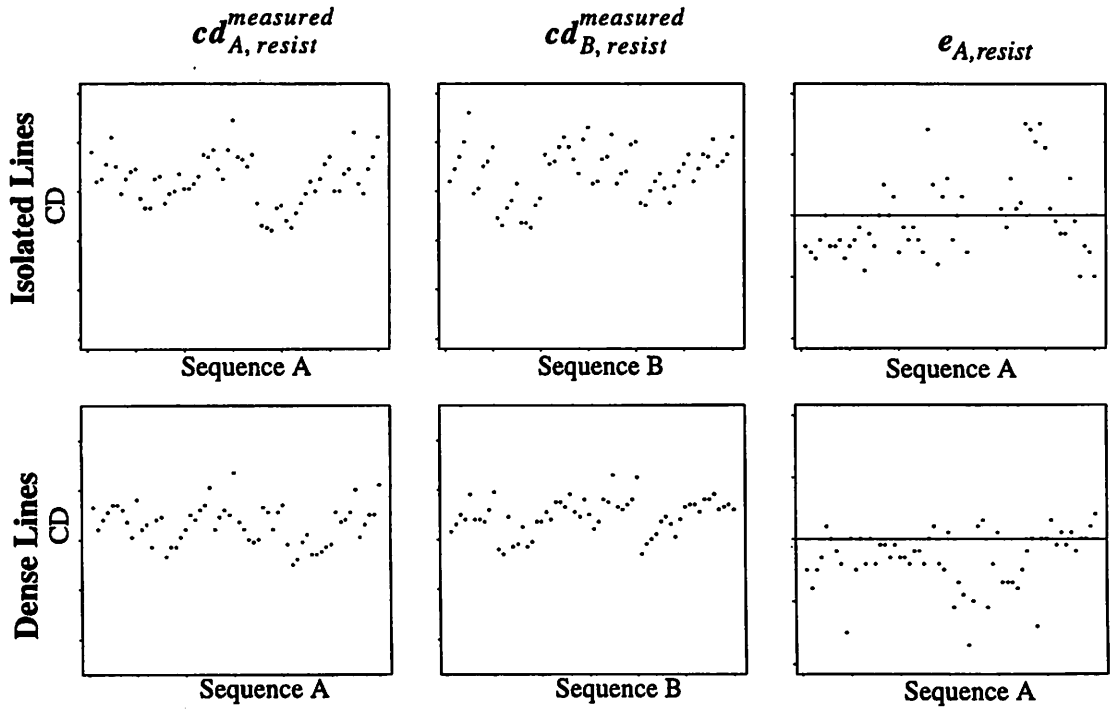


FIGURE 4.4 Example of the raw data from the sampling experiment on one wafer. The scale of the vertical axis is the same for all plots. Sixty points each of isolated and dense features are visited in two sampling sequences. The resulting error sequence is an estimate of the measurement error trend.

presence of the measurement trend ***d*** (Figure 4.4). There the addition subscript "resist" refers to the analysis carried out on resist CD's. The slightly negative mean of the points in the three graphs indicates a mean shift between the two measurement sequences. This will not affect the calculated variability over the wafer. Some structure in the data remains with respect to sequence. However, these structures are found in small segments over the entire

sequence and are probably due to incomplete randomization of $cd^{measured}_{B,\,resist}$ by R. Because of their relatively short ranges of effect, these localized structures are not likely to impact the trend we are trying to estimate if the trend is limited to low orders.

The systematic measurement drift d_{resist} is of most interest with respect to the estima-

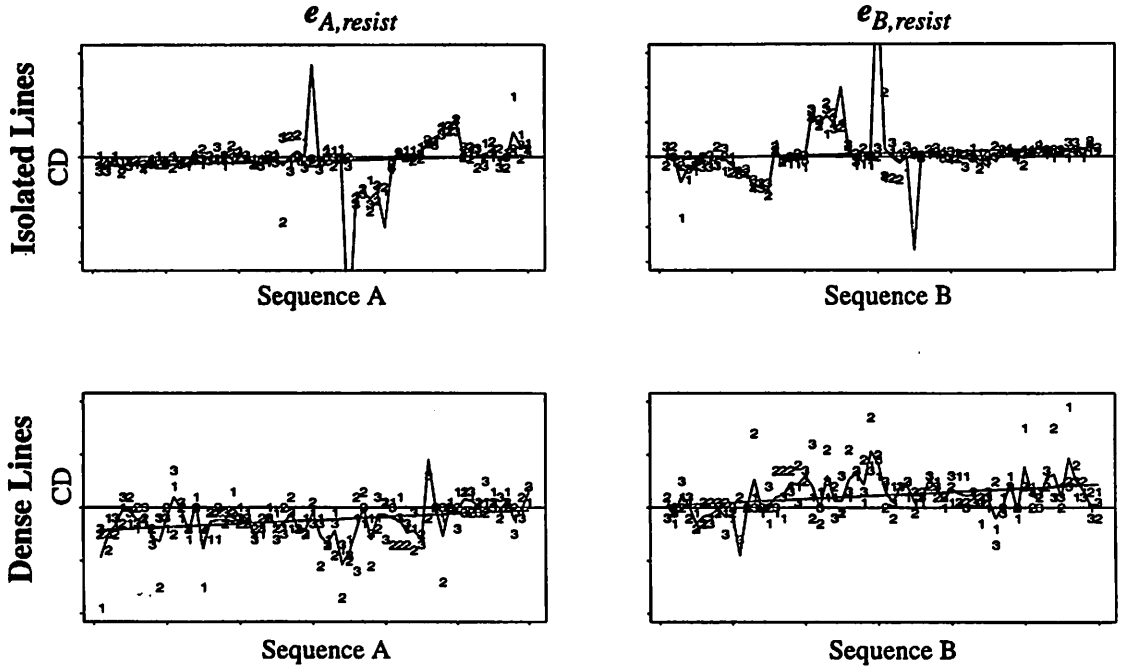


FIGURE 4.5 Measurement error versus measurement sequence. Scatter points (numbers) represent raw data from three wafers in one split. Solid line represents the average of three wafers. The estimated linear trend from linear regression is shown.

tion of $w_{develop}$. The assumption is that the error characteristics depend on the factors in the experimental splits: dense/isolated and exposure dose. Three wafers were measured for the high and medium dose splits, and the resulting error trends were averaged. The trends were obtained by fitting a simple linear model using least squares regression. As an example, the high dose wafers are shown in Figure 4.5. Measurements from the low dose wafers are quite noisy. This might be due to poor linewidth resolution as a result of under-exposing the wafers. Summary statistics for the high and medium splits are shown in Table 4.3.

Table 4.3 Summary Statistics of Trend Regression for SEM Resist CD Measurements

Exposure Dose	Feature Type	Sequence	Slope	% of Error	Significant at 5%
High	Isolated	A	-0.04	0.7	
High	Isolated	B	-0.054	0.4	
High	Array	A	0.064	11.53	✔
High	Array	B	0.0667	12.51	✔
Medium	Isolated	A	0.099	0.95	
Medium	Isolated	B	0.1023	1.024	
Medium	Array	A	0.02	1.3	
Medium	Array	B	0.018	1.0	

In general, the trends are very weak for even the most prominent conditions of arrays and high exposure doses. A slope of .1 amounts to 6nm of measurement drift over 60 measurements. The magnitude of the measurement drift is quite small by metrology standards, but is significant compared to the magnitude of the error we are trying to detect on the wafer. The measurement trend for each experimental condition was the same regardless of measurement sequence. This is consistent with our model which decouples the wafer values from the trend values. The "% of Error" column indicates the amount of variance explained in the measurement error series by the linear trend. The random portion of the measurement error dominates in all cases. The significance criterion is the t-value of the estimated slope with a threshold of 5%. The linear measurement trend has to be accounted for in the estimates of the wafer trends.

4.4.2 Resist CD Wafer Component Estimation

The double sequence experiment yielded estimated of the measurement trend δ_{resist} for measured resist CD variation w_{resist}. Following (4.3), the develop contribution, $w_{develop}$, can be estimated from $w_{resist}^{measured}$ by removing the measurement noise through subtraction of δ_{resist} and filtering of ξ_{resist} by regression.

Three sets of three wafers were exposed at three doses. CD measurements were taken in the center of each of the 21 fields on the wafer, on both isolated and dense lines. Note that here the intra-wafer variability $w_{resist}^{measured}$ was sampled in a separate experiment from the double sequence experiment. The measurement algorithm produced three values for each structure, representing CD values at the top, center, and bottom of the pattern. Some summary statistics are provided in Table 4.4.

Table 4.4 Aggregate Sample Variance (nm^2) of Resist CD Measurements for Three Wafers

Exposure Dose	Wafer #	Isolated (Top)	Isolated (Center)	Isolated (Bottom)	Dense (Top)	Dense (Center)	Dense (Bottom)
High	1	32.06	24.95	26.26	21.25	23.21	23.21
High	2	20.83	27.73	32.25	9.13	14.09	13.73
High	3	55.46	47.36	55.05	27.85	21.16	21.16
Medium	1	18.51	15.91	15.99	14.45	13.85	257.35
Medium	2	32.83	32.66	32.66	23.23	13.05	31.06
Medium	3	17.41	21.83	26.13	7.35	8.91	178.90
Low	1	30.96	29.76	31.45	27.83	109.05	1369.26
Low	2	19.95	25.50	39.29	24.10	73.59	423.53
Low	3	16.45	23.96	34.10	8.80	78.13	478.53

For the high dose split, sample variances are consistent for top, center, or bottom measurements. In the cases of medium or low dose, the bottom measurements of arrays show increasingly higher variation. This is consistent with the fact that with decreasing dose, more resist remains between the lines, causing non-vertical slopes on the line structures. However, there is also the accompanying effect of weaker electron signals from the bottom of the structures and obstruction from neighboring lines. Therefore, the center and bottom measurements from arrays in splits 2 and 3 are not reliable. For this work the top measurement is used. In general the variability over the wafer is quite small (~5nm).

The intra-wafer signal $w_{develop}$ is estimated by empirical removal of the measurement error. The systematic error is removed by simply subtracting the estimated trend from the raw data. This was performed only for the cases where the linear trend model was signifi-

cant. The random measurement error is removed by taking advantage of the spatial correlation of $w_{develop}$. Because of the relatively small number of measurements over the wafer (21), the systematic component of the wafer variability is best extracted by fitting a slowly varying surface to the measurements. A low order polynomial form is used:

$$w_{develop} = a_0 + a_1 x^2 + a_2 y^2 + a_3 x + a_4 y + a_5 xy \quad \text{and} \tag{4.10}$$

$$w_{resist}^{measured} = w_{develop} + \delta_{resist} + \xi_{resist}$$

where ξ_{resist} is the random experimental noise ~$N(0,\sigma^2)$.

Figure 4.6 shows the resulting contours from the filtering for the three wafers in the medium dose split and figure 4.7 shows results for the high dose split. In general, the model adequately describes the spatial features on the wafers. Moreover, the spatial features between the array and dense CD's agree for the same wafer. Also, the spatial contours between wafers suggest that the systematic pattern has rotational freedom. Both of these facts are consistent with the mechanisms of the develop process. The bar graph in Figure 4.6 shows the relative amounts of variation between the wafers as well as between the signal $\hat{w}_{develop}$ and the measurement errors. Table 4.5 shows the summary statistics of

the filtering for these splits wafers. The "$\hat{w}$ % meas" column refers to the percentage of

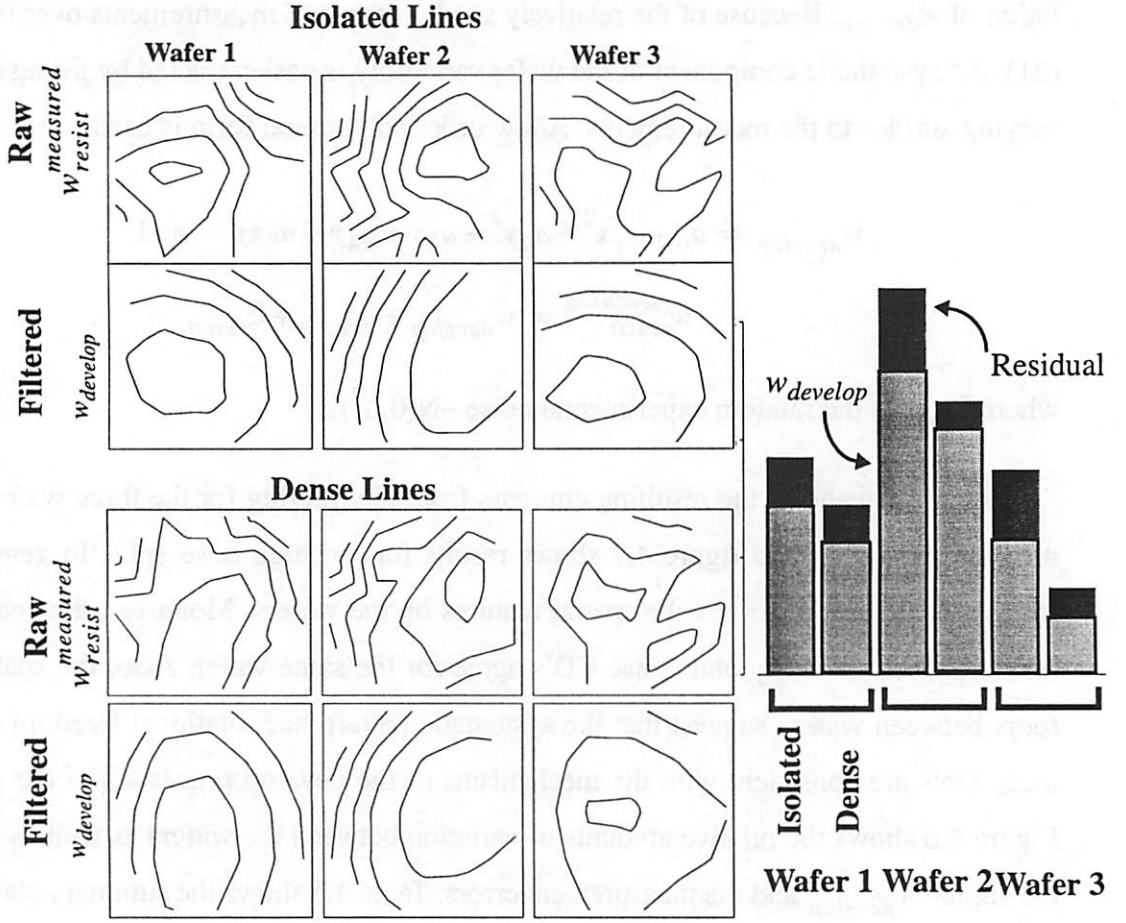


FIGURE 4.6 Raw and filtered intra-wafer CD contours for 3 wafers in the medium dose split. Interval between contours are the same for all plots. Bar plot shows the relative amounts of variance between the wafers, and between the systematic intra-wafer signal $w_{develop}$ and regression residual portion (Error).

variance $\hat{w}_{develop}$ relative to the variance in the raw data. This is the R^2 value of the regression in the absence of significant measurement trends. "P-Value" is based on the F-statistic for model significance, which we set at 5%. The models are significant, and the systematic trend over the wafer accounts for 10-20% of the measured variation.

Table 4.5 Variation Summary for Medium Dose Wafers Resist CD's

Wafer #	Isolated $\hat{w}$ %meas	Isolated P-Value	Dense $\hat{w}$ %meas	Dense P-Value
1	0.88	0	0.8	0.001
2	0.81	0.001	0.9	0
3	0.76	0.003	0.64	0.029

Similarly, results are shown for the high dose split. In this case, the measurement trend calculated from Sect 4.3.1 has been removed from the raw data before the spatial regression. Figure 4.7 shows the results of the filtering. Again, the systematic patterns of

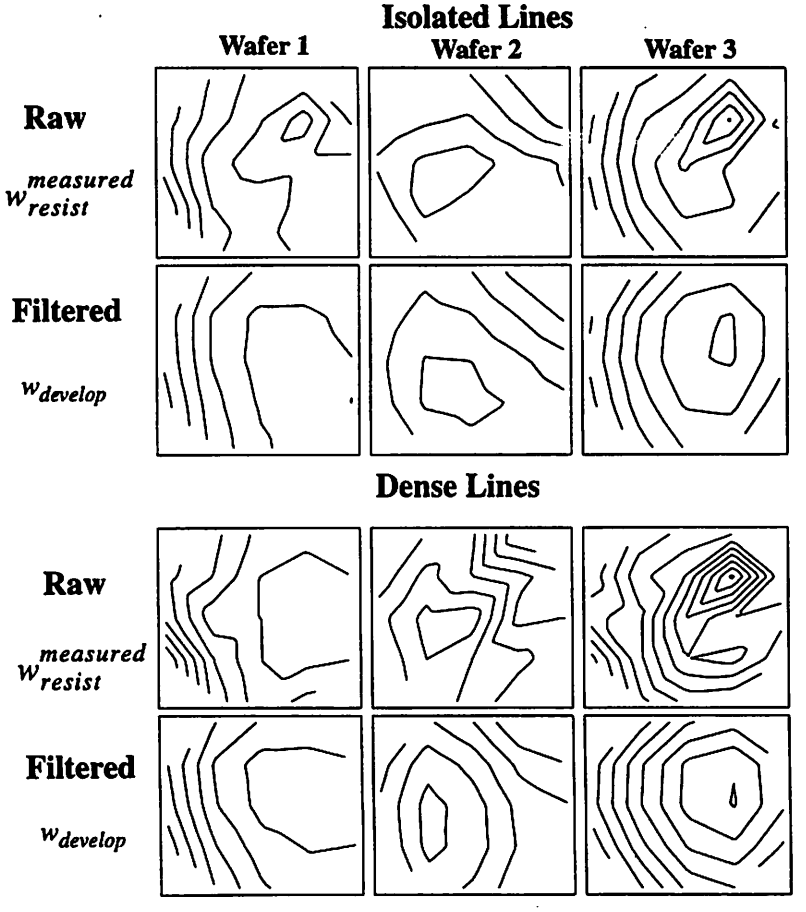


FIGURE 4.7 Raw ($w_{resist}^{measured}$) and filtered ($w_{develop}$) intra-wafer CD contours for 3 wafers in the medium dose split. Interval between contours are the same for all plots. Bar plot shows the relative amounts of variance between the wafers and between error vs. systematic portions.

$\hat{w}_{develop}$ show no rotational registration. They are also consistent for both isolated and dense features on the same wafer. The variance summary is shown in Table 4.6. The "δ % total" column refers to the percent variance adjustment to the raw sample variance after removing the measurement trend whereas "$\hat{w}$ % meas" refers to the percent variance in the measured data explained by fitting the surface $\hat{w}_{develop}$. Positive quantities for the trend correction indicate that variation increased after removing measurement trends. The sign of this contribution depends on the spatial correlation of the trend to $w_{develop}$.

Table 4.6 Variation Summary for High Dose Wafers Resist CD's

Wafer #	Isolated $\hat{w}$ %meas	Isolated P-Value	Dense $\hat{w}$ %meas	Dense δ % meas	Dense P-Value
1	0.89	$<10^{-5}$	0.9	3.3	$<10^{-5}$
2	0.85	$<10^{-5}$	0.88	12.8	$<10^{-5}$
3	0.79	0.001	0.87	0.3	$<10^{-5}$

4.5 Analysis of Poly CD Etch Error

Following the analysis performed in Section 4.3, the etch error can be estimated. Measurements of after-etch poly CD's yielded (4.3):

$$w_{poly}^{measured} = w_{poly} + \delta_{poly} + \xi_{poly} .$$

The poly intra-wafer component w_{poly} can be estimated by removing the measurement errors δ_{poly} and fitting a low order polynomial such that

$$w_{poly} = a_0 + a_1 x^2 + a_2 y^2 + a_3 x + a_4 y + a_5 xy \quad .$$

The etch component is then estimated by

$$\hat{w}_{etch} = \hat{w}_{poly} - \hat{w}_{develop} \quad .$$

Two wafers from each of the three dose splits were plasma etched. Top-down SEM measurements were taken on the polysilicon CD's in a series of experiments that parallel the resist CD analysis. The etch contribution to CD variability was found by subtracting

the develop contribution from the final CD variation.

4.5.1 Poly CD Measurement Sequence Error

Using the sequencing experiment in Section 4.3.1, two wafers were measured for each dose splits, and the resulting error trends averaged. Only isolated lines yielded satisfactory measurements and are analyzed in this section. The trends were obtained by fitting a simple linear model using least squares regression. The results are shown in Figure 4.8, where the subscript "poly" refers to the sequencing experiment carried out on the poly layer. Summary statistics for the splits are shown in Table 4.7. The "% of Error" column

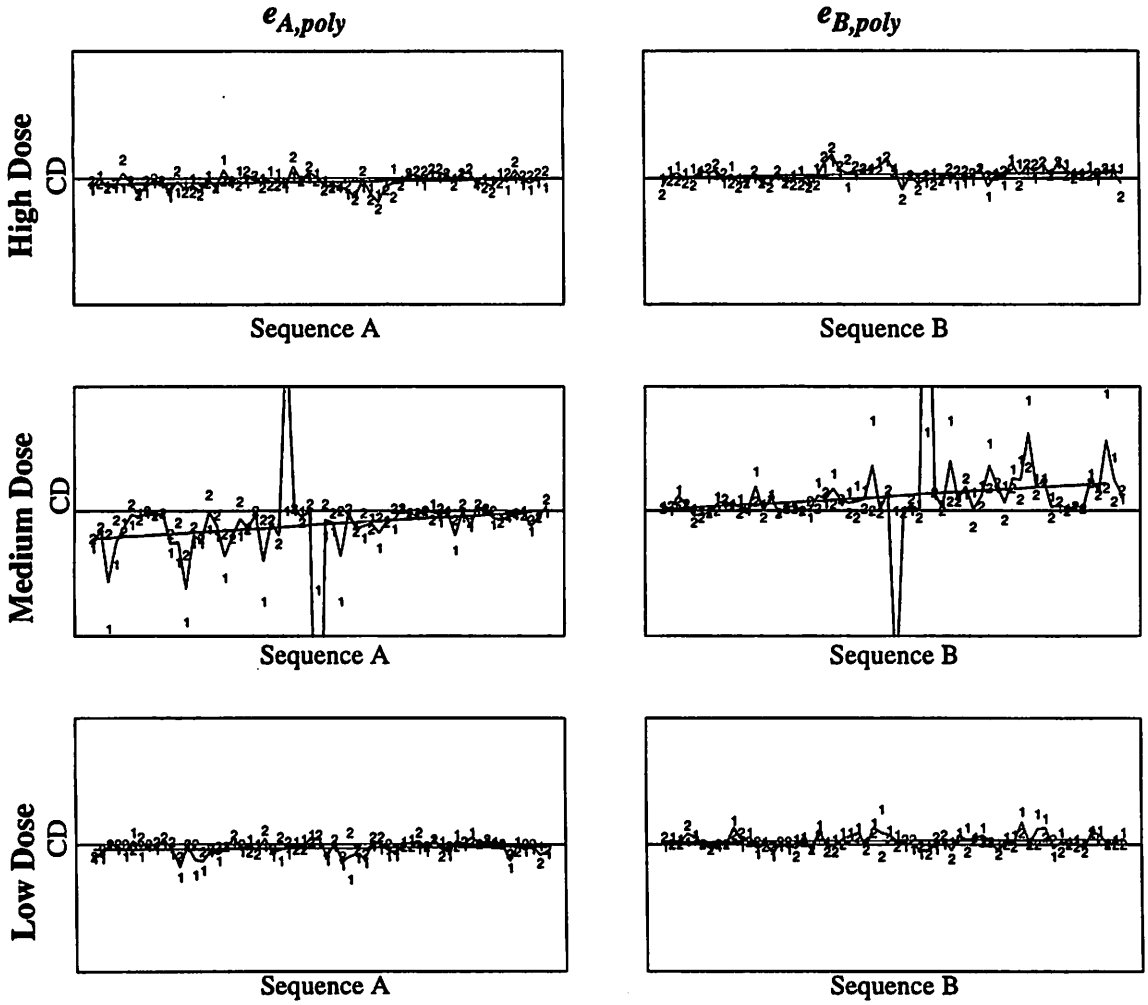


FIGURE 4.8 Measurement error versus measurement sequence. Scatter (numbers) points represent raw data from two wafers in each split. Solid line represents the average values from the two wafers. The estimated linear trend from linear regression is shown.

indicates the amount of variance explained in the measurement error series by the linear

trend. The random portion of the measurement error dominates in all cases. The significance criterion is the t-value of the estimated slope with a threshold of 5%. As in the case of resist CD measurements, the trends are very slight in all but the medium dose split. Again, the measurement trends for each experimental condition were the same regardless of measurement sequence. In these measurements, the split with medium dose exposures showed significant measurement trend. This trend is also much greater in magnitude than the one observed in the resist CD measurements. On average there is 1nm drift for every 5 measurements. The trend is positive in slope, as in the resist case, indicating CD gain with measurement time. The wide scatter in part of the measurement error sequence corresponds to the fact that the measurements were over a portion of the wafer. The linear measurement trend is accounted for in the next section where the wafer variability is estimated.

Table 4.7 Summary Statistics of Trend Regression for Poly CD Measurements

Exposure Dose	Feature Type	Sequence	Slope	% of Error	Significant at 5%
High	Isolated	A	0.0365	4.80	
High	Isolated	B	0.0350	4.41	
Medium	Isolated	A	0.1918	21.02	✔
Medium	Isolated	B	0.1938	21.46	✔
Low	Isolated	A	0.0203	1.58	
Low	Isolated	B	0.0196	1.47	

4.5.2 Poly CD Wafer Component Filtering

As in the case of resist structures, CD measurements were taken in a separate experiment in the center of each of the 21 fields on the wafer, on both isolated and dense lines. Three values for each structure, representing CD values at the top, center, and bottom of the line were produced. Summary statistics are provided in Table 4.8. As in the resist case, in the high dose split, sample variances are consistent for top, center, or bottom measurements. In the cases of medium or low dose, the dense measurements of arrays show unrea-

sonably high variation. This could be due to either measurement errors, such as electron obstruction by neighboring structures, or by process errors, such as partially cleared poly film. In this section only the top measurement from isolated structures are used. There seems to be significantly more variability in the raw data from the resist measurements. This is at least in part due to the introduction of variation by the etch process.

Table 4.8 Sample Variance (nm^2) of Polysilicon CD Measurements

Exposure Dose	Wafer #	Isolated (Top)	Isolated (Center)	Isolated (Bottom)	Dense (Top)	Dense (Center)	Dense (Bottom)
High	1	32.06	24.95	26.26	21.25	23.21	23.21
High	2	20.83	27.73	32.25	9.13	14.09	13.73
Medium	1	18.51	15.91	15.99	14.45	13.85	257.35
Medium	2	32.83	32.66	32.66	23.23	13.05	31.06
Low	1	30.96	29.76	31.45	27.83	109.05	1369.26
Low	2	19.95	25.50	39.29	24.10	73.59	423.53

As in the case of post-develop measurements, the poly intra-wafer signal w is estimated by removal of the measurement error: subtraction of the estimated trend from the raw data for the cases where the linear trend model was significant and removal of random error by extracting a continuous CD surface over the wafer from the measurement data. Figure 4.9 shows the resulting contours after filtering. In general the model adequately describes the spatial features on the wafers by visual inspection as well as by the amount

of variance explained ("$\hat{w}$ % meas" column in Table 4.9).

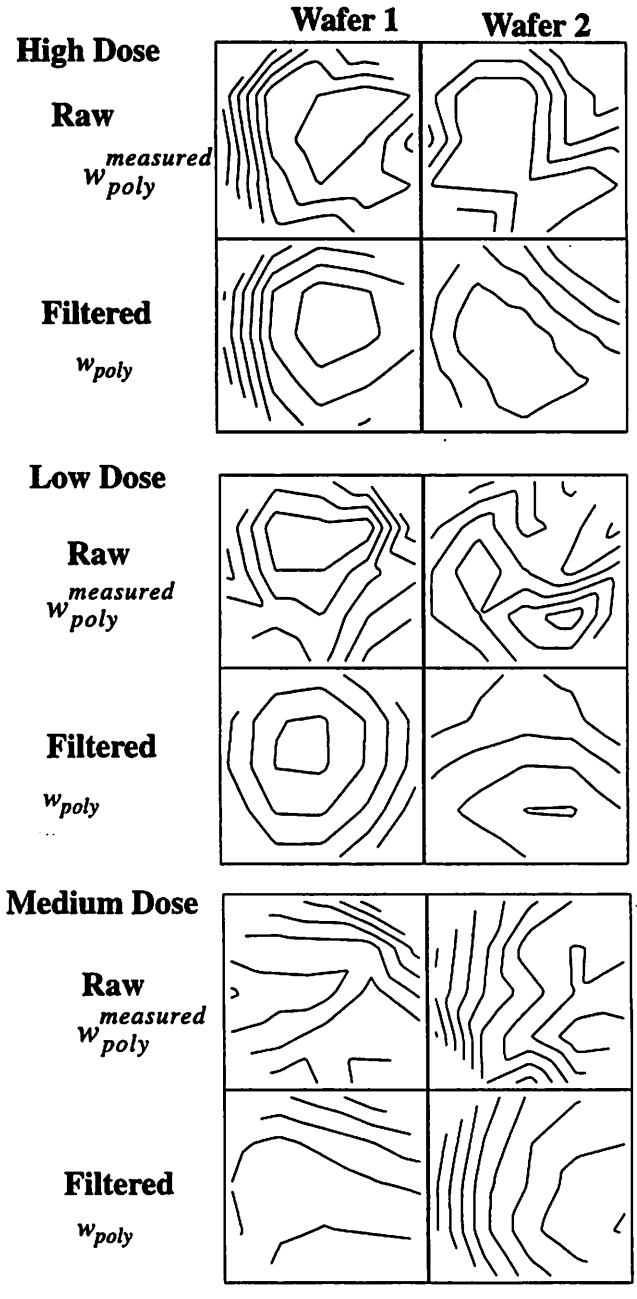


FIGURE 4.9 Raw ($w_{poly}^{measured}$) and filtered (w_{poly}) intra-wafer CD contours for 2 wafers in each split. Interval between contours are the same for all plots.

Table 4.9 Variance Summary for Polysilicon (Isolated CD) Regression

Exposure Dose	Wafer #	$\hat{w}$ %meas	P-Value	δ % meas
High	1	70	0.017	
High	2	89	$<10^{-5}$	
Medium	1	72	0.019	36.8
Medium	2	90	$<10^{-5}$	7.1
Low	1	71	0.008	
Low	2	57	0.078	

Table 4.9 shows the relative amounts of variation between the wafers as well as between the signal w_{poly} and the measurement errors. "$\hat{w}$ % meas" refers to the percent variance in the measured data explained by fitting the surface $\hat{w}_{poly}$. "δ % meas" refers to the percent variance adjustment to the raw sample variance after removing the measurement trend. The effects of trend variation on net variation depend on the spatial correlation of the two quantities. In all cases, the models are significant from the p-values of the f statistic. Using these results, etch decomposition can be carried out. From (4.2)

$$\hat{w}_{etch} = \hat{w}_{poly} - \hat{w}_{develop} \tag{4.11}$$

and can be estimated by subtracting the resist measurements from the poly, and then extracting the systematic portion of the result. The measurement sequencing drift has to be accounted for when appropriate. Figure 4.10 shows the contours of the decomposition sequence for two wafers. In both cases, the difference between the two measurements is quite noisy, as expected. The extracted etch variation is very small in magnitude (~6nm

over the wafer).

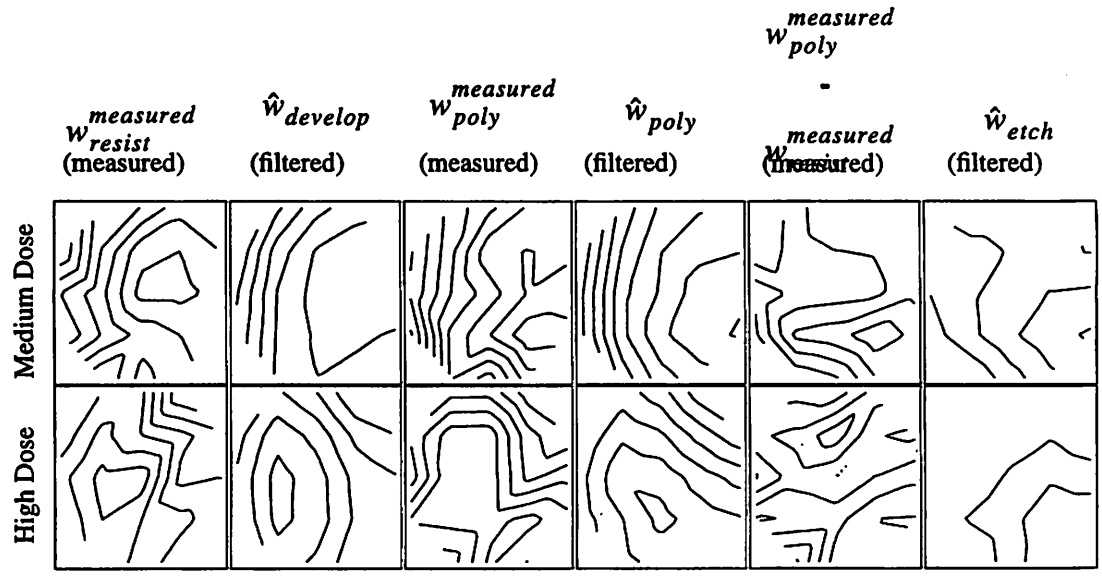


FIGURE 4.10 Etch and develop decomposition sequence for 2 wafers: medium dose and high dose. Contour intervals are the same in all cases.

Table 4.10 Variance Summary of Intra-wafer SEM Poly CD Decomposition

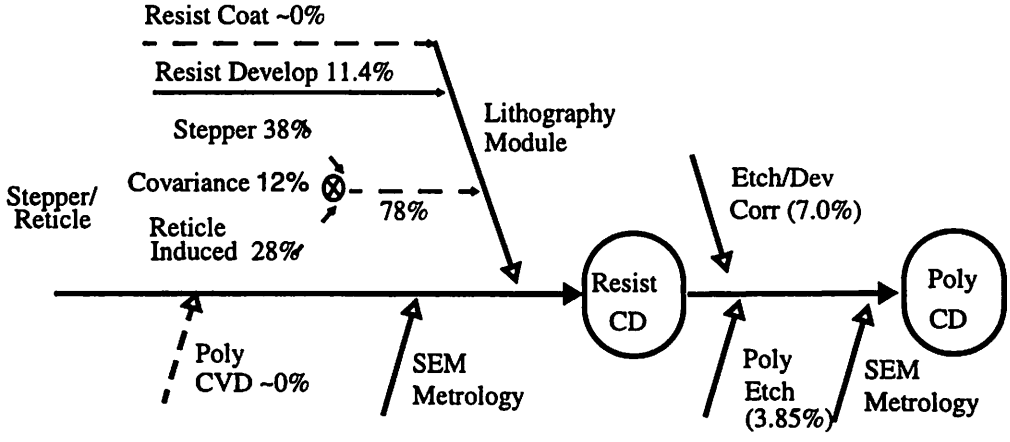
Exposure Dose	Wafer #	$\hat{w}_{poly}$ (%)	$\hat{w}_{develop}$ (%)	$\hat{w}_{etch}$ (%)	Cov (%)	$\hat{w}_{etch}$ (% meas)	$\hat{w}_{etch}$ p-value
High	1	100	65.6	17.9	16.5	42.6	.292
High	2	100	81.0	34.1	-10.9	43.5	.321
Medium	1	100	81.6	48.0	-40.7	45.2	.336
Medium	2	100	52.1	17.5	32.1	60.4	.069
Low	1	100	132	18.6	-48.7	25.9	.747
Low	2	100	79.4	56.3	-35.8	42.8	.289

Decomposition results are summarized for all six wafers in Table 4.10. Variability components are expressed in percentages normalized to the final poly CD variation (from filtered data). Taking the variance of (4.2) yields

$$var(\hat{w}) = var(\hat{w}_{develop}) + var(\hat{w}_{etch}) + 2cov(\hat{w}_{develop}, \hat{w}_{etch}) \cdot \quad (4.12)$$

From this the poly variability w can be decomposed into three components: develop, etch, and the spatial covariance between develop and etch. Variation due to spatial correlation between the etch and develop components is listed in the "Cov (%)" column and is defined as $2cov(\hat{w}_{etch}, \hat{w}_{develop})$. Because the develop and etch contributions are determined using two separate measurements, additional measurement error is injected into the analysis. Therefore, the variability components do not always sum to 100%, which is the measured final poly CD variation and includes error from one set of measurements. The last two columns summarize the wafer trend regression for the etch component. "$\hat{w}$ (% meas)" refers to the percentage of systematic variability extracted relative to the raw data. In general the systematic component is ~40% of the total, the rest can be attributed measurement noise. We find that the estimated model parameters are significant only for one of the wafers.

4.6 Develop/Poly Etch Decomposition Summary



(3.3%) SEM metrology error for poly etch
(13.2%) SEM metrology error for resist CD

FIGURE 4.11 Error budget diagram for .35μm poly gate patterning. Percentages of variance normalized to the final poly CD variability. Experimental error listed separately.

Automated SEM measurements have been used to estimate the develop contribution to CD variation within the wafer. Separate sequencing experiments and statistical analysis

were performed to characterize SEM measurement error and determine sampling artifacts. Filtering of the raw data yielded estimates of the systematic error attributed to the develop process. After plasma etch, the experiments were repeated to determine the poly measurement errors and the poly CD variability. The etch variability was then extracted from the difference between the poly and resist measurements. The etch component is in general quite small (~20% of total poly CD variability) and in most cases dominated by experimental noise.

References for Chapter 4

[9] "OPAL 7830i Automated SEM-CD Wafer Metrology System", Applications Reference Manual.

Chapter 5

Metal CD Variability

5.1 Metal Etch Variability

In this chapter, the contribution of etch variability to metal CD patterning is determined. By extending the methodology for variability decomposition using Statistical Metrology, metal CD variability was measured using short loop electrical measurements. Subsequent statistical and causal decomposition of the raw data isolated the variability contribution from plasma etch. In Chapters 3 and 4 we developed a methodology which decomposes variability in electrically measured poly CD's into individual equipment contributions using a series of statistical and causal filters. In this chapter this methodology is extended to determine CD error contributions from metal etch.

5.1.1 Metal Etch Decomposition Methodology

A .50 μm metal interconnect patterning process using I-line lithography is used to demonstrate the methodology. Following the experiment described in Section 2.2.1, a test mask of electrically measurable structures was designed to print CD samples over the stepper field and the wafer. Then the pattern was transferred through a short-loop process sequence as to minimize the number of confounding sources of variation. A spatial filter was implemented to separate the within-wafer variability from the variability within the stepper field. The intra-wafer variability excludes error introduced by the stepper and reticle pair, which consists of variations within the stepper field.

Separate experiments were conducted to isolate the error introduced by metal etch. The short-loop process sequence was repeated on a thin metal film using the exact lithographic process as the original metal stack with respect to the resist, exposure, and develop process (Figure 5.1). Metal etch for both the thin and thick films was carried endpoint with

the same equipment and chemistries. The much shorter etch times used for the thin film wafers minimize the effect of etching on the CD structures. The variation of the difference in CD variation between both thick and thin film CD's is used to quantify metal etch variability.

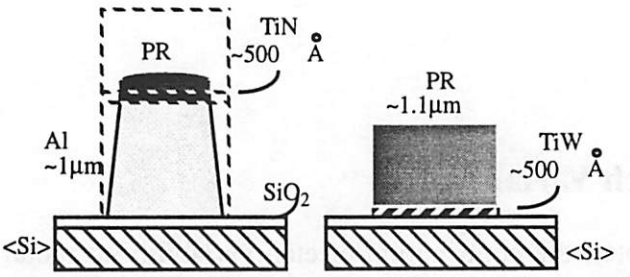


FIGURE 5.1 Thin versus thick metal film stacks.

5.1.2 Metal CD Sampling Experiment

A test mask was designed to sample systematic CD variability at a fixed operating point in the metal interconnect patterning sequence using electrical linewidth measurements. Spatial variability was sampled by repeating CD patterns over a reticle in a 5X5 grid spanning a 20x20mm stepper field. This reticle pattern was repeated 24 times over a 6" wafer.

The dimension and proximity conditions of the CD patterns were varied in each location on the reticle to sample systematic variability due to mask CD variations and proximity effects. Both isolated and dense arrays of linewidth topologies were considered with nominally .5μm lines at 1μm pitch. The linewidth structures are electrically measurable using the Kelvin 4-point measurement design. The sheet resistance is determined by adjacent Van der Pauw structures to compensate for errors in film resistivity. The raw data from the experiment consists of periodic repetitions of the reticle over the wafer. Prior to the experiment, SEM measurements were used to determine appropriate stepper exposure conditions to pattern lines with the target dimensions.

No wafer topography was included in order to minimize the number of confounding processing effects. The thick metal stack consists of ~1μm thick Al film with small per-

centages of Cu topped by ~50nm of TiN on top of 100nm isolation oxide. The thin metal film in causal decomposition uses thin TiW (~50nm) in place of the thick metal stack. Both types of wafers were patterned in the identical process sequence consisting of film deposition, resist coat, pattern exposure, resist develop, and metal etch (to endpoint). Resulting CD structures were probed using an automated electrical measurement system.

5.2 Metal CD Results and Analysis

5.2.1 Metal CD Raw Data

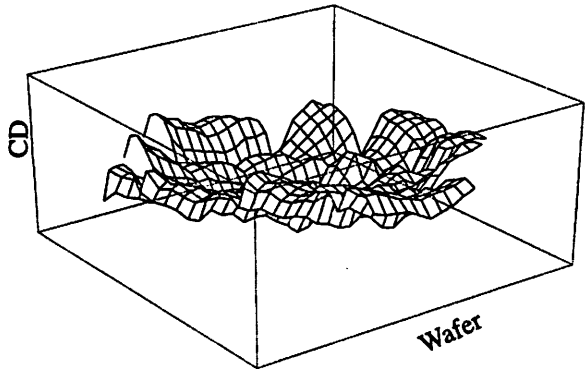


FIGURE 5.2 Elevation map of nominally .5μm isolated metal CD's over the wafer for a thick metal stack.

Figure 5.2 shows the elevation map for the electrically measured values of .5μm isolated CD's over the wafer for the thick metal film stack. In all CD plots, the vertical scales are the same. Roughly radially symmetric spatial variability over the wafer can be observed. Also spatial variation, less in magnitude, can be observed within exposure fields.

Figure 5.3 shows the CD variability for the same structure on the patterned thin film. The variability over the wafer is much less than for the thick film case. Also, periodic within-field variability structure is evident. This could be a result of the effect of reduced

etching over the wafer from the shorter etch time.

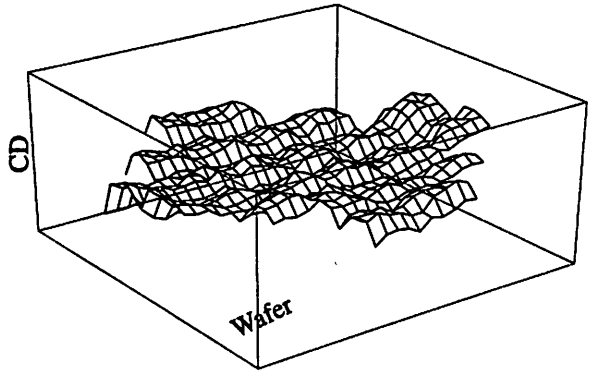


FIGURE 5.3 Elevation map of nominally .5μm isolated metal CD's over the wafer for a thin metal stack.

5.2.2 Statistical and Physical Filtering of Metal Etch Variability

Restating the error formulation in Chapter 3, the variability over the wafer can be expressed as:

$$cd_{metal}(x, y) = p_{metal}(x, y) + w_{metal}(x, y) + f_{metal}(x, y) + \varepsilon, \qquad (5.1)$$

where $p_{metal}(x,y)$ is the within exposure field signature from the stepper/reticle pair and is perfectly periodic field to field. The exposure to exposure variability $f_{metal}(x,y)$ from dose, focus, or leveling error is also systematic within the field. Variability over the wafer $w_{metal}(x,y)$ is continuous and slowly varying. Random error, ε, is normally distributed and has no spatial structure. These components are assumed to be additive. The explicit subscripts in the variables refer to the metal CD's used in this chapter.

The etch variability of interest is incorporated in w_{metal} along with variability from film deposition, resist coat and resist develop. From independent measurements, the variation in film thickness for both the resist and the metal films are <1%. Their effects on CD variability are negligible. The components of w_{metal} are then primarily from etch and develop. Again assuming additivity:

$$w_{metal} = w_{Mdevelop} + w_{Metch}. \tag{5.2}$$

Here the prefix "M" in the subscript distinguish these variables from their counterparts in poly decomposition (Chapter 4). Combining equations (5.11) and (5.2), the terms p_{metal} and $w_{Mdevelop}$ will be approximately the same for both thin and thick film wafers because of identical process conditions.

$$w_{Thick} = w_{Mdevelop} + w_{MetchThick} \quad \text{and}$$

$$w_{Thin} = w_{Mdevelop} + w_{MetchThin} \quad . \tag{5.3}$$

Subtracting the CD values between the thick and thin film wafers yields:

$$\Delta w = w_{MetchThick} - w_{MetchThin} + \varepsilon' \tag{5.4}$$

where ε' combines the differences of f_{metal} and ε between the two wafers. Because f_{metal} and ε are not systematic over the entire wafer and the values between the wafers are not correlated ε' is assumed to be random over the wafer.

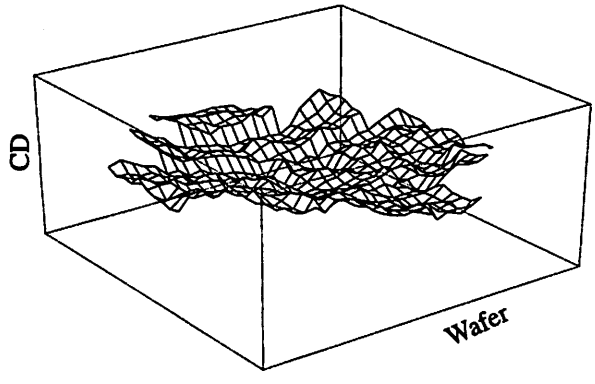


FIGURE 5.4 Elevation map of the difference in nominally .5μm isolated metal CD's (Δcd_{metal}) over the wafer between the thick and thin metal stacks.

Figure 5.4 shows the elevation plot for Δw. While the variability over the wafer remains, the spatial structure within the fields is mostly eliminated. The differences in the etch effects can easily be obtained by extracting the systematic component of Δw. This can be done most easily by fitting a low order polynomial surface over the wafer through linear regression:

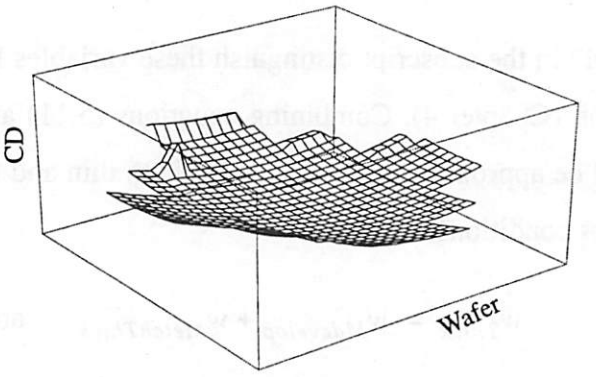


FIGURE 5.5 Fitted polynomial surface extracting the systematic variability over the wafer.

$$\Delta w = a_0 + a_1 x^2 + a_2 y^2 + a_3 x + a_4 y + a_5 xy + \delta \tag{5.5}$$

where δ is random error. The resulting surface from the regression is shown in Figure 5.5 and has an R^2 value of 0.72.

With the additional assumption that the effect of etch on a very thin film is small compared to that of a thick film, the fitted surface can be used to estimate the metal etch contribution to CD variability so that

$$\Delta \hat{w} = \hat{w}_{MetchThick} - \hat{w}_{MetchThin} \cong \hat{w}_{MetchThick}. \tag{5.6}$$

Also, (5.3) becomes

$$w_{Thin} \cong w_{Mdevelop} \quad . \tag{5.7}$$

Summarizing the variability components, (5.3) becomes

$$w_{Thick} \cong w_{Thin} + w_{MetchThick} \quad . \tag{5.8}$$

The variance is given by

$$var(w_{Thick}) \cong var(w_{Thin}) + var(w_{MetchThick}) + 2cov(w_{Thin}, w_{MetchThick}). \tag{5.9}$$

The etch component $w_{MetchThick}$ is estimated by regression on Δw, and the quantities w_{Thin} and w_{Thick} are measured.

The decomposition results are summarized in the following table

Table 5.1 Intra-wafer Metal CD Decomposition Summary

Variability Component	% Variance of w_{Thick}
Thin film wafer (w_{Thin})	**30.4**
Estimated etch portion $\hat{w}_{MetchThick}$	**32.4**
$2cov(w_{Thin}, w_{MetchThick})$	**30.1**

The correlation term refers to the variability contributed by the spatial correlation between the etch contribution and the contribution from the rest of the process (estimated by variability from the thin film wafer).

5.3 Metal CD Etch Decomposition Summary

In this chapter, Statistical Metrology has been extended to extract etch variability from a metal CD patterning process sequence. An experiment using a thin metal film was used to approximate the effects of the process without the etch. Subsequent subtraction of measured CD's and extraction of the systematic component yielded the estimate for the etch contribution.

Chapter 6

Applications of Statistical Metrology

In the previous chapters, we have developed a methodology which decomposes and categorizes systematic spatial CD variability into individual equipment contributions. Two applications of the methodology in Chap. 3 and 4 yielded the variability assignments in the poly patterning sequence. This type of variability assignment is advantageous for several reasons: 1) It identifies bottlenecks in manufacturability and offers an opportunity to manipulate various variability components through the choice of equipment hardware and process settings. 2) The isolated equipment variability components can be used as a benchmark metric between equipment choices. 3) The spatial correlation between variability components can be manipulated to reduce net variability. In this Chapter, several results of Statistical Metrology will be applied towards improvements in process design.

In an IC manufacturing sequence, all of the process tools introduce some variation into the product. In this way, each tool used along the process sequence attenuates product yield in some way. However, systematic variability can be manipulated through judicious design- by changing the process operating point or by modifying the hardware. With appropriate filtering and decomposition of the measured data, it is possible to determine the response surface of the variability of particular process steps over a range of process conditions. Process variability can be also be manipulated through the choice of process equipment. Specifically the reticle/stepper and the develop variability will be examined.

Lastly, we show an example in this chapter in which manufacturability is incorporated as an additional criterion in process design. In this way, process manufacturability can be optimized by minimizing the systematic variability and the sensitivity to random variability. Process variability is also a function of the layout of the circuit being manufactured. We show that a continuum of process variability versus circuit design parameters

can be established using Statistical Metrology to facilitate circuit design for manufacturability.

6.1 Reticle/Stepper Variability Applications

Spatial features in process variability can be used to gain insight into the process and equipment. In Chapter 3 and 4, causal decomposition provided a methodology to estimate CD error budget items. Since the decomposition procedures, consisting of a test mask and filtering algorithms, are generic, they can be applied to various process segments and across fabrication sites. A causal error breakdown also reveals ways to manipulate process error. The results in Chapter 3 already suggest that the correlation between error components can reduce net variability. Similarly, altering equipment and equipment settings can affect CD variability. In Section 3.4.2, causal decomposition of intrafield CD variability yielded separate error contributions from the stepper and the reticle. The results of this decomposition can be used to formulate a numerical metric to gauge equipment manufacturability.

6.1.1 Intrafield Systematic Variability

The detailed spatial information obtained from extensive variability sampling (Chapter 2) can provide additional levers to manipulate process variability. Moreover, to incorporate spatial information may place additional constraints on process design. The effects of structure topology, in the form of CD proximity designs, were also measured in the experimental. These measurements, also sampled over the field and the wafer, may provide useful insights into the equipment or process mechanics. The following are a few examples in which spatial information has supplemented process design or equipment operation.

Column Tilt

The intrafield variation p, obtained as a result of signal decomposition (Section 3.4.2), shows roughly radial variability which is not perfectly centered within the field. In subsequent equipment maintenance, an adjustment was made on the optical column of the

stepper. The signal p obtained after the adjustments shows centered contours. (Figure

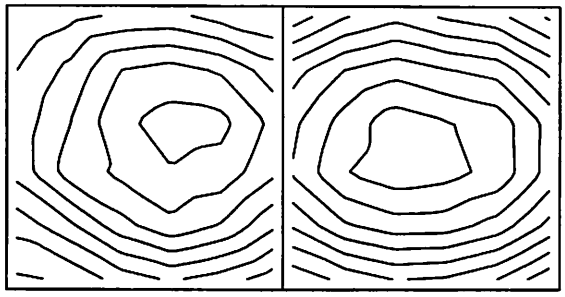


FIGURE 6.1 Intrafield CD contours of $p(x,y)$ a) before and b) after column tilt alignment. CD contours drawn at same intervals.

6.1b) This example is additional validation of the signal decomposition, in that the signal responded predictably to equipment manipulation.

Isolated-Dense Bias

Besides spatial variation, other systematic errors can be observed IC processing. One type of error which has been of concern is proximity effect error, or error between CD's with different neighboring structures. A common metric of this effect is the difference between isolated and dense features. In actuality, this metric is a measure of both optical proximity effects and etch microloading effects. Usually this quantity is characterized by one scalar quantity, without considering effects of spatial variability. In Figure 4.2, both the within field and over the wafer variability of isolated minus dense CD's in the horizon-

tal orientation is shown. The intrafield variability is, by arguments similar to those for

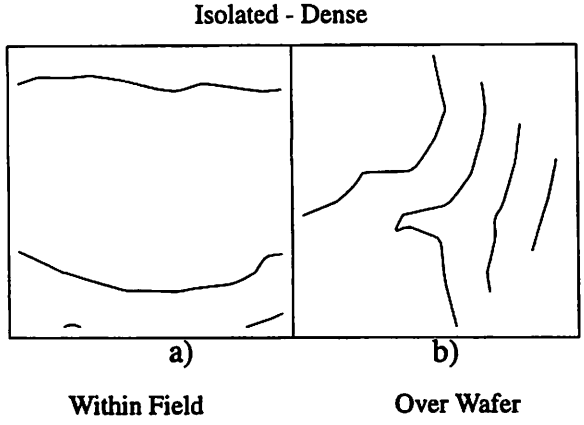


FIGURE 6.2 As printed CD contours a) within field and b) over wafer of isolated minus dense horizontal CD's. CD contours NOT drawn at same intervals.

causal decomposition, due to stepper optics. The spatial trends observed tend to follow that of individual CD variability shown in Figure 4.1. This is to be expected, since both CD and variability in iso-dense differences are manifestations of the same optical aberrations. By parallel arguments, over the wafer variability of iso-dense differences are due to plasma etching. This is shown in Figure 4.2b, in which a trend of variation low in the right to high in left side of the wafer is observed. This trend is repeatable for other wafers measured.

Orientation Dependence

Similarly, trends in orientation dependence are telling of optical aberrations. In Figure 4.3, the intrafield contours for horizontal minus vertical (x-y) lines are shown. The resulting pattern is in the shape of a "saddle", or 0 along the diagonals, positive in the top and bottom quadrants, and negative in the left and right quadrants. This can result as a difference of two elliptical surfaces, one with the major axis along x and one along y, and is

indicative of astigmatism in the optics.

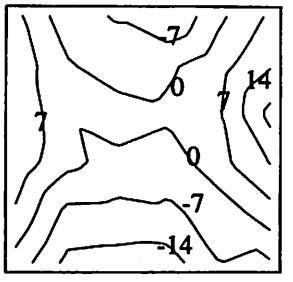


FIGURE 6.3 CD contours within field of horizontal minus vertical printed CD's in (nm).

6.1.2 Stepper Benchmarking

The stepper/reticle decomposition methodology developed in Chap. 3.4 identified the stepper contribution to CD variability apart from the process and reticle variability. This methodology can be applied to benchmark steppers. Three advanced i-line steppers, each from a different manufacturer, are characterized. The tools have comparable functionality and are designed for manufacturing .35μm technology. In this generation of i-line exposure tools, the CD requirements are very close to the resolution limit of the equipment. Tool variability is expected to be very visible.

Following the electrical experimental procedure described in Chapter 2, short loop patterning was performed on flat polysilicon wafers made in the same batch. Resist coat for all wafers was also performed on the same process line. Stepper exposure and develop were carried out at the three tool demonstration sites. Wafers were collected, etched and measured at the original process line. CD's were sampled in a 6x6 array in 22x22mm

square fields. Decomposition results are shown in Figs. 6.4 and 6.5.

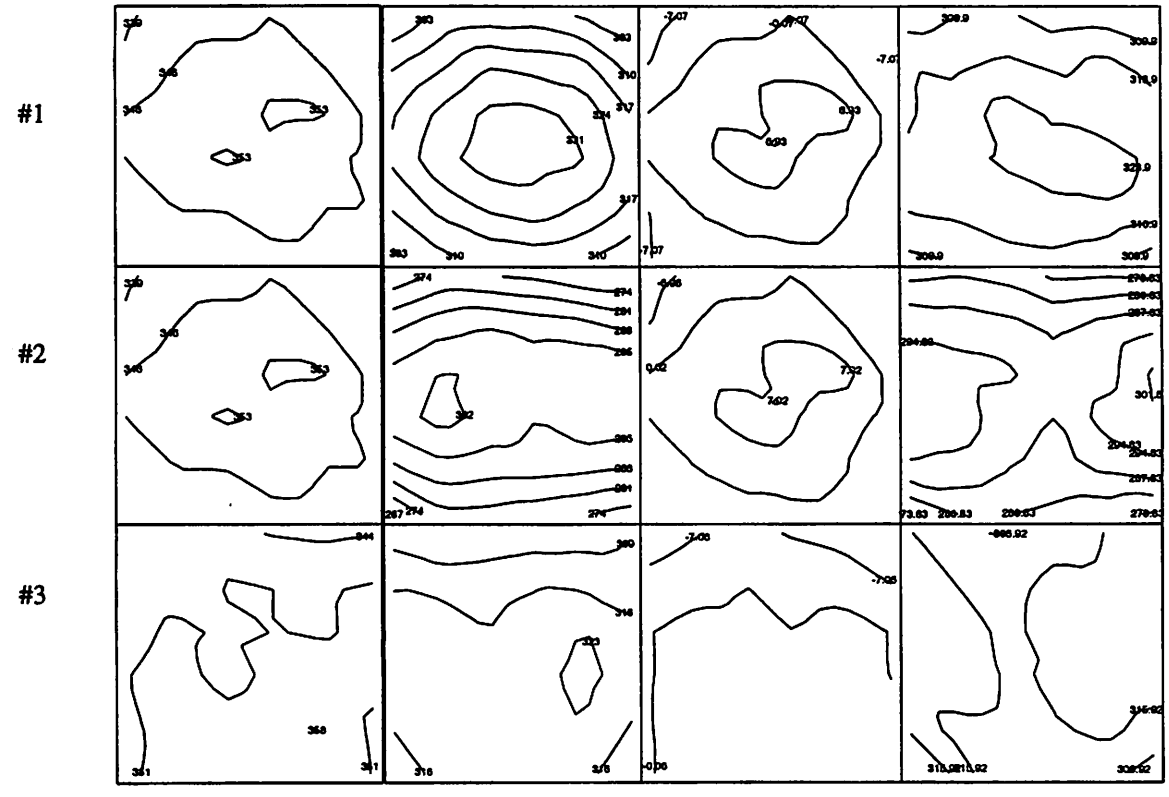


FIGURE 6.4 Stepper/reticle decomposition results for three stepper models. Contours shown are isolated pattern .35μm CD variation within the field. The intervals between contours are equal in all plots.

In Figure 6.4, stepper/reticle decomposition results are shown for the .35μm isolated features. The mask errors, Δcd_r, are quite small. Note that in stepper #3, a different reticle was used; it's CD is distributed differently over the reticle than the others. Because the reticle induced error is subtracted, use of different reticles with different variability does not compromise the benchmark metrics. The periodic intrafield component p has been extracted from the electrically measured raw data. By characterizing the stepper CD transfer function (following Chapter 3.4.1), the effect of reticle error on p has been calculated as $M\Delta cd_r$. In all cases, the reticle error is enhanced. Again, this is because $|M|>1$ as well as in steppers #1 and #2 the mutually spatial enhancement of $M(x,y)$ and $cd_r(x,y)$. The rightmost column shows the CD contours of p^t, the stepper contribution. The different spatial features indicate different optical aberrations present in the tool designs, with stepper #3 showing the least CD variability. In Figure 6.5, the decomposition has been performed for

the dense features. Qualitatively results are similar to the isolated case.

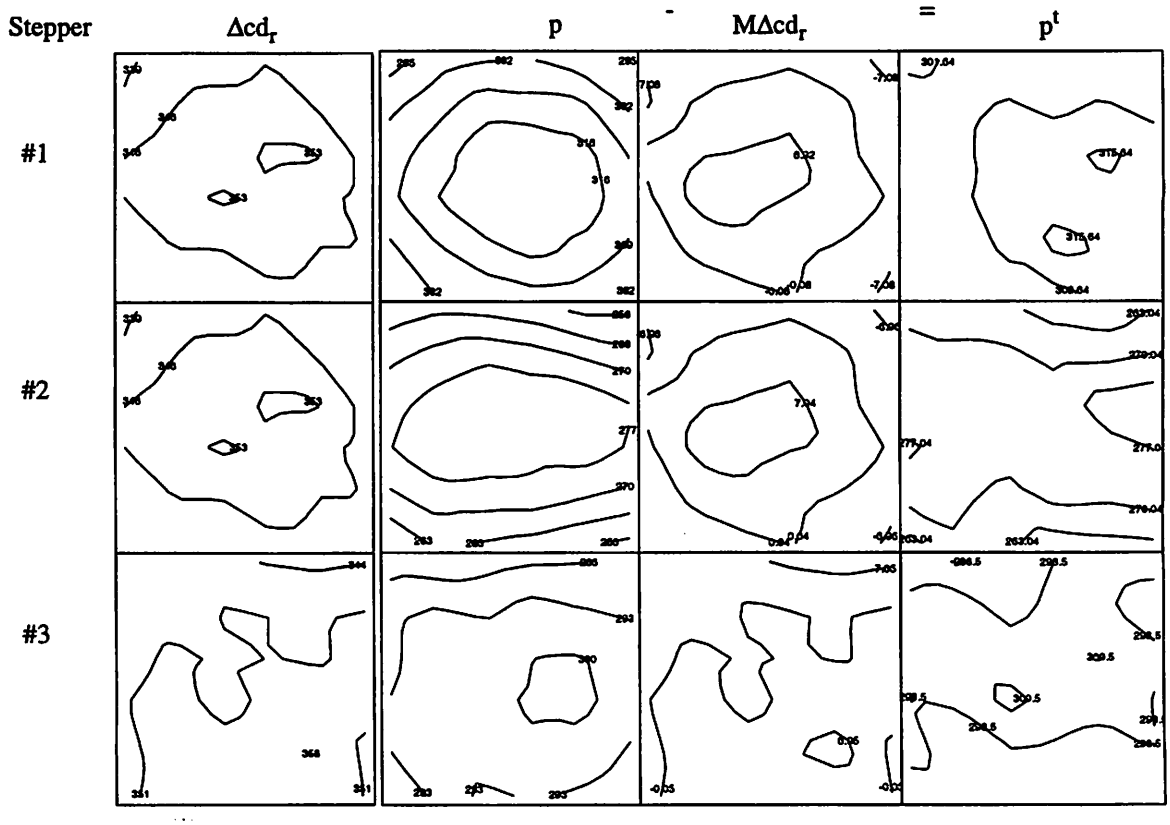


FIGURE 6.5 Stepper/reticle decomposition results for three stepper models. Contours shown are dense pattern .35μm CD variation within the field. The interval between contours are equal in all plots.

The benchmark study is summarized on Table 6.1, where variability components are expressed in terms of variance (nm^2). In general the reticle contributions to variability are comparable, even accounting for the enhancement factor *M*. While the total extracted variability of *p* from steppers #1 and #2 were comparable, the estimated stepper contributions were different by a factor of close to two. The spatial correlation between the stepper and reticle-induced components account for a significant amount of variability in steppers #1 and #2. However, in #3, the spatial distributions of the components were such that there was very little, even somewhat negative correlation between them.

Since the net CD variability depends on the exact spatial distribution of the reticle CD's, it is difficult to formulate convenient manufacturability metrics which incorporate all the spatial information of stepper variability. In the following sections the variation in

Table 6.1 Error Components as Result of Stepper Benchmarking (nm^2)

Structure	Component	Symbol	Stepper #1	Stepper #2	Stepper #3
Isolated	**Stepper**	p^t	77.4	41.1	15.5
Isolated	**Reticle-Induced**	$M\Delta cd_r$	20.9	26.2	28.4
Isolated	**Covariance**	**2*Cov**	39.5	46.8	-6.7
Isolated	**Total**		137.8	114.1	37.2
Array	**Stepper**	p^t	21.0	37.3	16.3
Array	**Reticle-Induced**	$M\Delta cd_r$	27.4	27.3	21.9
Array	**Covariance**	**2*Cov**	18.6	23.4	-8.1
Array	**Total**		66.9	87.9	30.0

the stepper alone error as well as the value and variability of the reticle error gain factor M will be presented as the manufacturability metrics of interest.

Following Chapter 3, p^t, the stepper alone component of CD variability, can be extracted. Furthermore, it is a function of the reticle dimensions. Figures 3.10 and 3.11 show the field CD contours for the stepper-alone variability as a function of the target reticle dimensions from 310 to 390 nm for both isolated horizontal lines as well horizontal lines in an array. A trend of decreasing variability can be observed with increasing reticle dimensions, as expected. Similarly, M can be calculated from the fitted transfer function and it's contours plotted as a function of reticle pattern dimensions (Figure 3.12 and 3.13). Again, a trend of decreasing variability and decreasing mean value of M is seen versus reticle dimensions.

These results can be summarized at the expense of the spatial information contained. As an example, Figure 6.6 compares the manufacturability of the stepper used in the analysis of Chapter 3 with that from a more advanced model made by the same manufacturer. The stepper only variation, var(p^t), and the scatter of the reticle error enhancement factor M is shown for both isolated lines and arrays of nominally .35μm horizontal CD's. For either structure, the stepper-only variability is smaller for stepper #2 over the range of reticle CD's. Furthermore, M for stepper #2 is both smaller in mean value and more uniform over the field over the range of reticle CD's. This shows that stepper #2 is less sensitive to

reticle errors. Using the same technique, comparisons of manufacturability can be made between manufacturers, different models, and even different steppers of the same model.

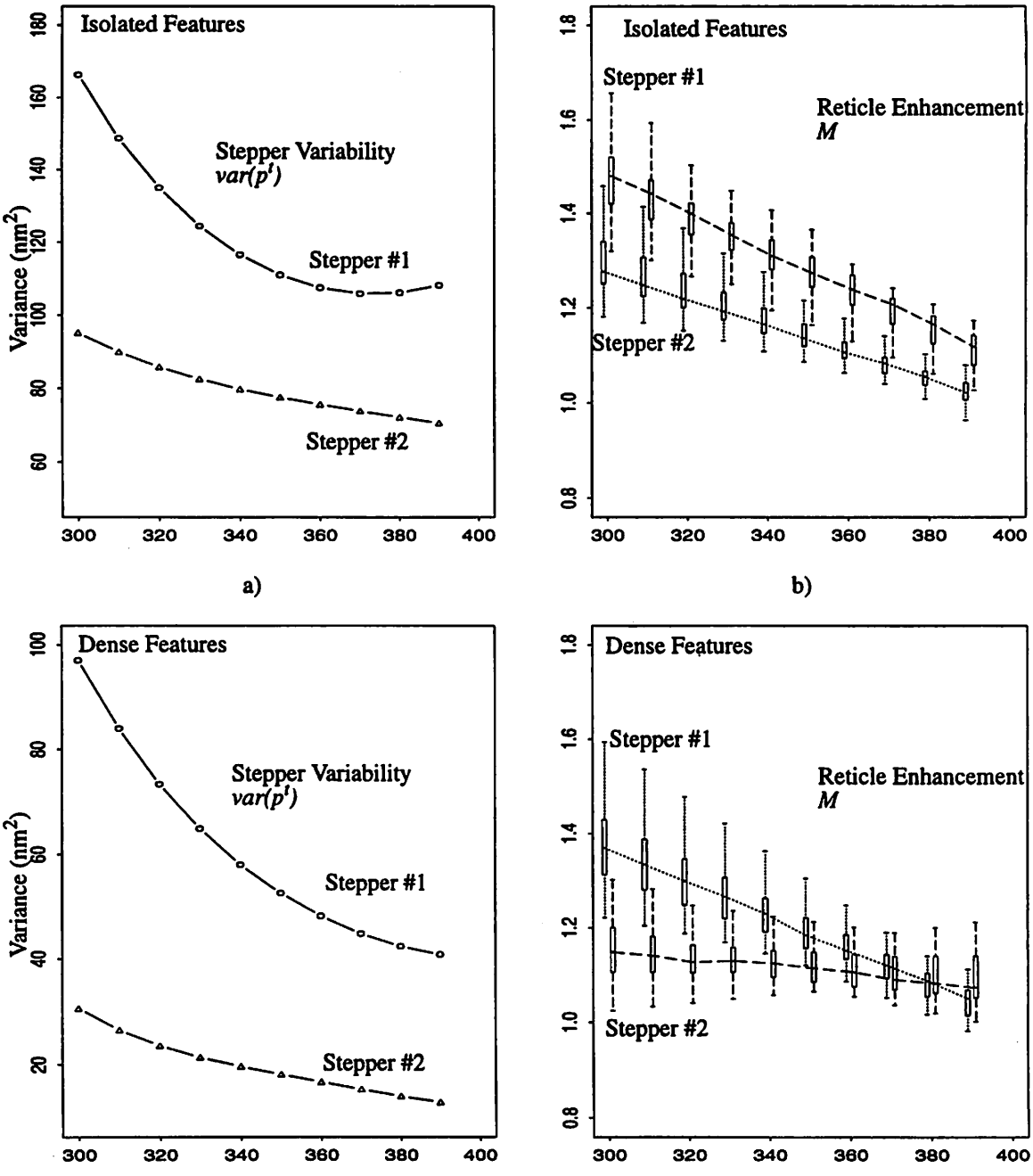


FIGURE 6.6 Variance of the stepper only variation $var(p^t)$ (figures a) and c)) and scatterplot of the reticle error enhancement factor M (figures b) and d) versus nominal reticle CD's evaluated from 310nm to 390nm by increments of 20nm. Isolated (a) and b)) and lines in an array (c) and d)) for steppers #1 and #2 are compared. Separations in the X axis between scatter points for #1 and #2 for ease of display only.

The variability decomposition discussed so far is an analysis of systematic variation

about a fixed process point. Therefore, decomposition results would depend on the process operating point. In fact, this has already been observed in the case of stepper comparisons, if we consider the equipment choice among the list of process settings. Similarly, variability decomposition, specifically the causal decomposition via transfer function characterization, can be affected by process conditions. As examples, the effects of resist process choice and stepper defocus will be demonstrated. Figure 6.7 shows the comparative effects on stepper-alone and reticle-induced variability components for CD's patterned using three different resist processes. Both variability components show significant effects from

changes in the resist process. Similarly, Figure 6.8 shows the effect of changing defocus

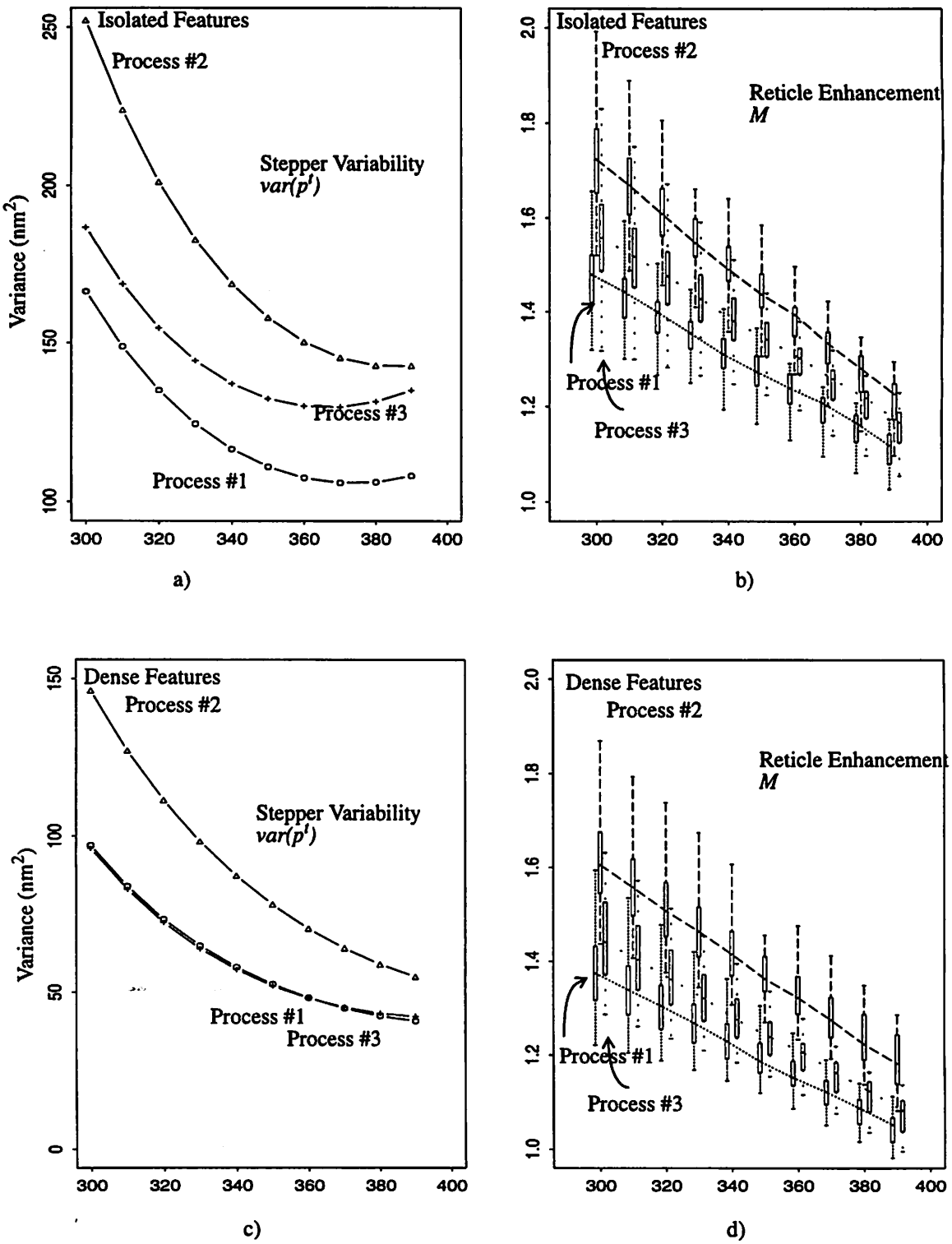


FIGURE 6.7 Variance of the stepper only variation $var(p^t)$ (figures a) and c)) and scatter of the reticle error enhancement factor M (figures b) and d) versus nominal reticle CD's evaluated from 310nm to 390nm by increments of 20nm. Isolated (a) and b)) and lines in an array (c) and d)) for resist processes #1, #2, and #3 are compared. Separations in the X axis between scatter points for ease of display only

values on the variability components. This demonstrates the importance of fixing the pro-

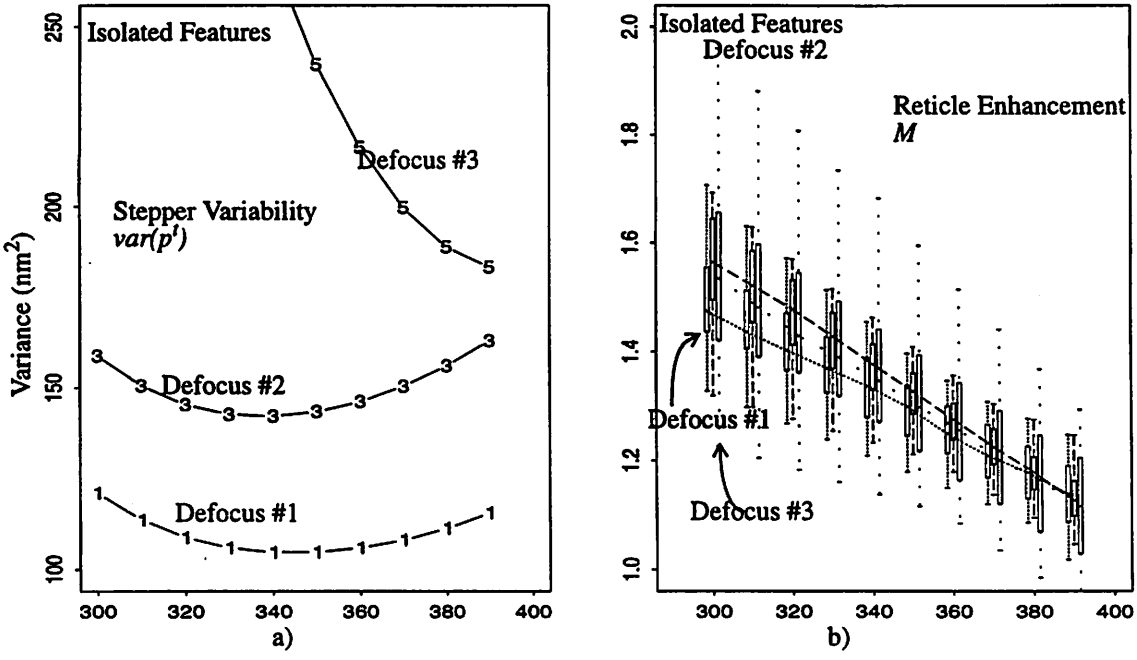


FIGURE 6.8 Variance of the stepper only variation $var(p^t)$ (figure a)) and scatter of the reticle error enhancement factor M (figure b)) versus nominal reticle CD's evaluated from 310nm to 390nm by increments of 20nm. Isolated lines for three values of defocus. Separations in the X axis between scatter points for ease of display only.

cess operating point while applying causal decomposition. While the variability in the process has been separated, the effects from changes in the process operating point is not compensated. For example, changes in the resist process and focus conditions would both affect stepper CD transfer linearity M. This would alter the results of error decomposition. Also, different process conditions would have different sensitivities to the optical aberrations which cause intrafield variability. These are all interactions not accounted for in the simple additive model of process variation. Therefore the additive model is only valid at a fixed process operating point and for small perturbations.

6.2 Applications Towards Resist Develop

CD variability over the wafer, w, is easily decoupled from the rest of the variability by sampling at a fixed location in printed fields over the wafer. Moreover, in Chapter 4, it was shown that plasma etching contributes relatively little variability to w. This allows the use of final measurements, made electrically or by SEM, of poly CD's as estimates of w_{de}-

$_{velop}$. Simple experiments were performed to observe the effects of both changes in develop recipe and equipment. In Figure 6.9, $\hat{w}_{develop}$ is shown as a response to two develop recipes. Effects in both the magnitude and the spatial features of $\hat{w}_{develop}$ can be observed. In manufacturing situations, new equipment undergo a period of characteriza-

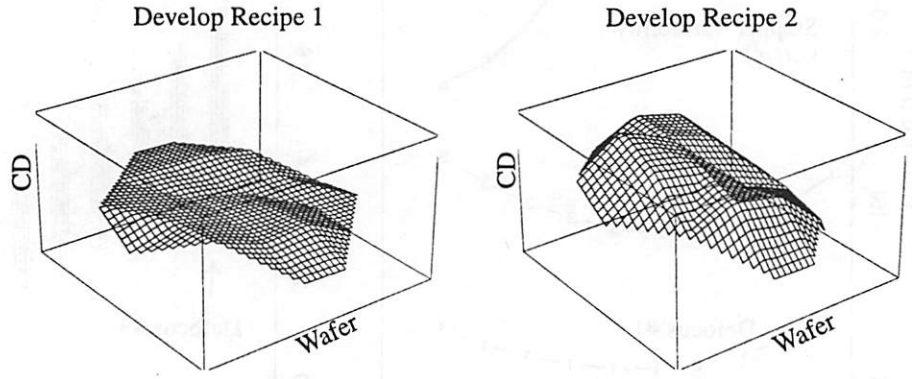


FIGURE 6.9 Response of CD variability over the wafer (w) for changes in develop recipe.

tion and qualification, sometimes with the intention of tuning tool performance to match existing equipment. Figure 6.10 shows the performance of a new develop track (B) compared to an existing track (A). These examples show how variability can be used to both debug process equipment problems as well as provide an additional criterion in process design.

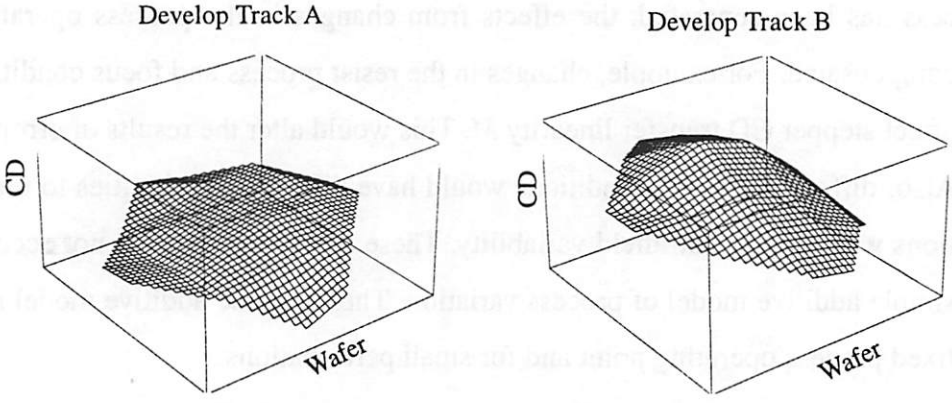


FIGURE 6.10 Response of CD variability over the wafer (w) for changes in develop equipment.

6.3 Process Design for Manufacturability

The electrical measurements obtained from the experiments described in Chapter 2 span a range over gate length and gate configurations. The process designer usually gener-

ates a response surface for process centering. However, simultaneous optimization of process center and process margin is necessary since we have shown that process variability has a significant systematic component and is a function of the process operating point. Figure 6.10(a) shows the response surface of poly CD as a function of mask bias and exposure dose. This is an example of a response surface characterizing process performance, which is typically the mean value of the profile of the parameter in question. The process calls for 0.35 μm poly gate dimensions, which can be achieved by different combinations of these dose and bias. The target CD intersects the response surface as a horizontal contour, so that all operating points along this contour will, in principle, produce the target gate length. The poly gate length variability due to the stepper may be calculated along this same contour for all points on a wafer using the data decomposition filter in Sect. 3.4. Figure 10(b) shows the variance for isolated lines and for lines in the center and at the sides of a group, each including both over the wafer and within the field variations. This figure is one form of process margin response surface which may be used in process design to identify an optimum region, as shown, in the operating space which simultaneously constraints on both the center point and spread of a process.

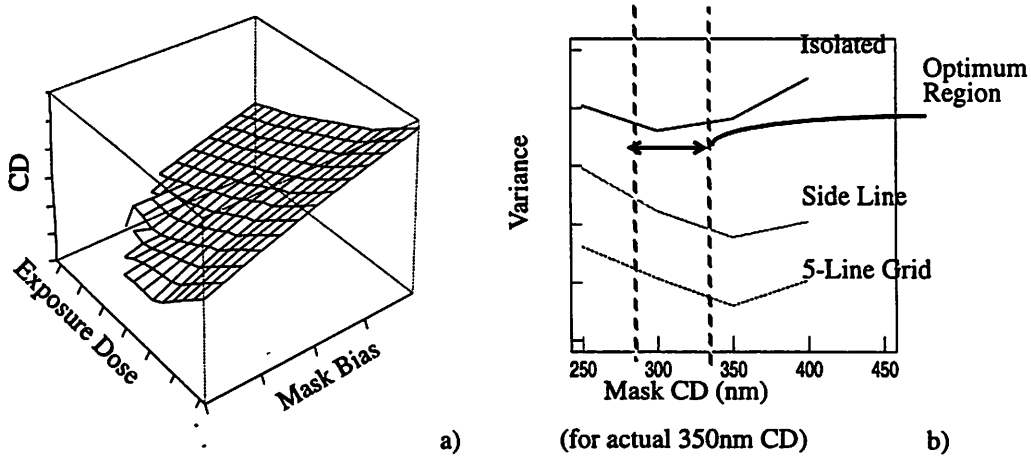


FIGURE 6.11 a) Average response surface for linewidth versus exposure dose and mask bias. Vertical scale obscured to protect proprietary data. b) Total wafer variance for 3 types of linewidths plotted versus mask bias. Dose was adjusted to meet target CD (350 nm) at each bias.

6.4 Statistical Metrology Applications Summary

With decomposition methodology established by Statistical Metrology, several variability features can be extracted from IC process sequences. Systematic spatial variability

reveals process errors, as well as provide additional criteria in process design. Spatial correlation also provides an additional lever in process manufacturability. Causal decomposition allows the estimation of equipment contributions to variability. This facilitates equipment benchmarking. Moreover, metrics produced through Statistical Metrology can be calculated over a continuum of parameters in order to generate a design surface.

Chapter 7

Conclusions

7.1 Thesis Summary

The management of variability in IC manufacturing is a complex undertaking. Whereas the issue of managing process variability over time is relegated to the activity of process control, in this thesis we investigated the structure of systematic process variability at a fixed time. The results of this investigation yielded a methodology to characterize the spatial and causal components of process variability. Spatial variability is important for several reasons. First, spatial information provides additional criteria for process and circuit design. In process design, the process latitude would be different if one considers full-field response versus a single-point response. In circuit design, systematic variability affects the performance of circuit elements in a predictable and reproducible manner. This offers an opportunity for designers to optimize the circuit layout in anticipation of known variability scenarios. Second, spatial variability contains causal information. Based on physical understanding of the process and equipment, the spatial structure can be formulated and variability components estimated. The signal decomposition performed in Chapter 3.3 is such an example. The estimated spatial signatures provide insights into equipment performance, or can be further decomposed into individual equipment components. Furthermore, given predictable systematic trends and the ability to control the variability components, it is possible to manipulate spatial correlation to reduce total variability.

Understanding of the causal structure of variability offers further opportunities to reduce process variation. Error budget items can be used as outputs in designed experiments, yielding response surfaces as a function of process control variables as well as process equipment. By changing the process settings, individual variability items can be

manipulated in order to optimize the overall error budget. Systematic, time-invariant process error constitutes a lower limit to process variation, with time-varying components contributing over this lower limit. Then, it becomes crucial to characterize systematic process variability and to design the process to establish an acceptable lower limit.

To actually obtain the causal and spatial error structures, several challenges were presented to available metrology techniques. Highly precise metrology was needed to quantify the very small differences between measurements. Furthermore, the technique had to be cost-effective for the large amounts of data required. Very significant was also the issue of the incorporation of process variation into the sampled variability. From these, a general view of measurements data was formulated. Variability sampling unavoidably includes error from process variation components as well as the metrology. Causal and spatial error decomposition, then, required the separation of variability from each process component of interest as well as from the measurements themselves. To this end, we developed a methodology which used statistical and experimental algorithms in the form of filters to decomposed raw measurement data into equipment components.

Statistical Metrology has been demonstrated through a series of case studies. The basic pattern transfer sequence, which served as the experimental basis in this thesis, involved process modules frequently used in IC fabrication- film deposition, lithography, and plasma etch. The polysilicon gate patterning sequence was an effective tool to demonstrate the decomposition of lithographic variability (Chapter 3), the results of which directly impacted gate-length control. For the .35μm generation of IC technology, stepper capability presented a serious concern because of limits in I-line resolution. This phenomenon exasperated the periodic intrafield signatures which were observed in CD measurements and facilitated variability decomposition through signal analysis. Subsequently, the intrafield component obtained from signal decomposition was decomposed into separate contributions from the stepper and reticle.

In Chapter 4, the etch and develop contributions to poly CD variability were characterized. The use of top-down SEM measurements provided a non-intrusive and cost-effective technique to sample CD variability directly after develop and after etch. However, we

showed that measurement errors in the form of systematic linear drifts of CD's was present. A separate experiment to characterize and model the measurement trend was devised so that the errors can be subtracted from the raw data. In Chapter 5, this methodology was also applied to estimate the contribution of metal etch to metal CD variation.

In both electrical and SEM CD's, process variability between the process steps was separated. However, the dependence of the variability responses on the process operating point was not accounted for. For example, we showed the use of different resist or exposure conditions yielded different stepper or reticle contributions. Also, the estimated SEM trend depended on the exposure conditions of the resist CD's. These phenomena were symptomatic of the empiricism of the decomposition.

Once the filters and experiments had been developed, the resulting variability components could be used to impact process design. In Chapter 6, several examples were shown.The spatial information provided insights into equipment behavior whereas the estimates of equipment variability provided a means to benchmark equipment. The dependence of process variability on controllable inputs was explored in the case of CD targeting, and provided an example of simultaneous determination of process performance and manufacturability.

7.2 Future Work

The decomposition of the critical dimension measurement was useful because pattern transfer constitutes a basic, and therefore highly leverageable sequence in IC fabrication. The methodology developed in this work can be extended to include other portions of the process as well as increased process complexity and interactions. Other features worthy of investigation include contact holes and spacers. These features have traditionally been difficult to measure because they are non-conductive and because, besides their dimensions, their profiles are important. Bulk parameters, such as those defined by doping, are critical to the understanding of device performance. Also, dielectric material stacks, such as oxides, measured along the vertical direction are of interest.

The exercise of characterizing systematic behavior in process variability was moti-

vated by the possibility of manipulating process variation. Causal decomposition revealed the structure of equipment variability contributions to overall CD variation, and allowed the manipulation of CD variability through recipe and equipment changes. Spatial correlation between variability contributions can be exploited to reduce net variability. Another category of systematic variability is the CD differences between structures with different neighboring topology, such as isolated versus dense lines. An example of this type of variability, specifically to lithography, is optical proximity correction. Similarly, it would be possible to compensate systematic biases introduced by other parts of the process, such as microloading effects in plasma etch. These relative spatial effects would interact with the systematic spatial trends discussed in this thesis, and vary over the field, the wafer, etc. The behavior of systematic variability over time adds yet another axis of complexity to this problem.