PROPAGATION EFFECTS OF PARTIALLY COHERENT LIGHT IN OPTICAL LITHOGRAPHY AND INSPECTION

by

Robert John Socha

Memorandum No. UCB/ERL M97/55

6 August 1997

PROPAGATION EFFECTS OF PARTIALLY COHERENT LIGHT IN OPTICAL LITHOGRAPHY AND INSPECTION

by

Robert John Socha

Memorandum No. UCB/ERL M97/55

6 August 1997

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

Propagation Effects of Partially Coherent Light in Optical Lithography and Inspection

by

Robert John Socha

Doctor of Philosophy in Engineering - Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Andrew R. Neureuther, Chair

This thesis describes a formulation of a computationally efficient method for analyzing topography scattering with vector polarized partially coherent spatial illumination; presents the implementation of this method into TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition); and shows systematic studies and rules of thumb for the effects of partial coherence in mask imaging, wafer patterning, and wafer inspection by using TEMPEST-PCD.

The effects of partial coherent light scattering from topography on a mask or wafer during printing and inspection have been a long-standing unresolved concern in optical projection printing. This thesis attacks this problem through generalizing the optimal decomposition to vector electromagnetic scattering. The decomposition method was initially developed by Gamo for scalar imaging and applied to optical proximity correction by Patti and Cobb and later used in pupil filter design by von Bünau. The generalization here decomposes the coherency matrix developed by Mandel and Wolf. Since this vector decomposition is a transform technique that involves the diagonalization of a matrix, this decomposition is both optimal in an energy compaction sense (since the least number of excitations is needed), and has orthogonal eigenvectors such that the partially coherent case is the direct sum of coherent cases.

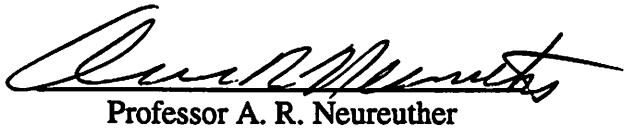
The inclusion of partial coherence effects in scattering from topography typically requires simulation time that is at least two orders of magnitude greater than a single coherent plane wave. However, with the decomposition technique, up to one order of magnitude can be saved in simulation time. Abbe's technique, which represents the illumination cone as a series of obliquely incident plane waves, typically requires 256 simulations for inspection problems, and 45 simulations for mask imaging and wafer patterning problems. The

decomposition technique, however, typically requires 30 simulations for the same inspection problem, and 8 simulations for the same mask imaging problem. The decomposition technique is recommended for all simulation problems except for truly periodic mask structures, where aliasing reduces the effectiveness of the decomposition method.

Simulation with experimental verification is used to evaluate effects and develop rules of thumb for the effect of the partial coherence in optical lithography and inspection. Using the mutual coherence, a perturbational model is developed that predicts the interaction and impact of a phase defect on the aerial image of a phase mask by considering the previously known behavior of isolated defects and line features. Where large topographies are encountered in phase shift masks, it is demonstrated that the assumption that the diffraction efficiencies must be independent of the illumination angle used in Hopkins' formulation is valid only when the aspect ratio (depth-to-width) of the mask feature is less than 0.2 for a stepper with σNA less than 0.5.

When simulating the patterning of an image inside a resist on a wafer with device features, interference effects at low σ (0.3) due to lateral scattering from the topography are more pronounced. While these interference effects are pronounced in the latent image it is shown that when Fickian diffusion occurs after exposure in a subsequent post exposure bake (PEB) process, the net effect can be similar to exposing with less coherent light. A consequence is that for most purposes the electromagnetic analysis can be simplified to a simple normally incident plane wave followed by a PEB. Lateral scattering effects in resist also play an important role in printing a dark line from a pre-patterned edge of a phase shift resist layer directly on top of the resist. Lateral refraction of energy into the phase shift edge increases the linewidth beyond the anticipated resolution of $\frac{\lambda}{4n}$ by severalfold.

Later, a sound methodology for improving wafer inspection will be demonstrated. The reciprocity theorem is used to create a filter which compensates for both the topography and the thin-film thickness. Such a filter guides light down the hole and increases the amount of light reaching the bottom of the hole by threefold.



Professor A. R. Neureuther
Committee Chairman

Table of Contents

List of Figures

List of Tables

Acknowledgments

Many people have assisted me in my education, and I would like to thank all of those who have made this thesis possible. First of all, I would like to thank my research advisor, Professor Andrew Neureuther. He has not only been my mentor but also has become a good friend during my graduate studies here at Berkeley. His guidance and encouragement is greatly appreciated. I am grateful to Professor Jeff Bokor and Professor William Oldham who have given me valuable insight into various aspects of optical lithography. I would like to thank Professor Bokor and Professor Raymond Chiao for serving on my dissertation committee, and to thank Professor Avideh Zakhor for acting as chairperson on my qualifying examination and Professor Ronald Gronsky for serving on my qualifying examination committee. Finally, I would also like to acknowledge Professor Andreas Cangellaris from the University of Arizona for helping me to prepare for my qualifying examination and for assisting me with the development of a point source excitation in TEMPEST.

The friendly research atmosphere at Berkeley has helped to promote my learning of optical lithography. Besides the great collaboration among the students in the lithography group, they have also become my good friends. I would especially like to thank Richard Schenker for his friendship. In addition, fellow graduate students, Konstantinos Adam, Seema Butala, Charlie Fields, Anna Ison, Derek Lee, Barry Linder, Bernice Lum, Tom Pistor, Edita Tejnil, Alfred Wong, Andrew Zenk, and Marco Zuniga, have assisted me through their friendship and support.

Finally, I would like to express my special thanks to my family for their love and support in whatever I endeavor to do. My father, Gerald, and my mother, Josephine, have always supported me and have been a great inspiration to me. My brothers, Paul and Thomas, have also been my closest friends and biggest supporters. I would also like to thank my best friend, Craig Linne; although he is not truly my brother, he has become one through the years. I really do not believe that I could accomplish many of my goals without

these five people. I hope to give back what you all have given to me through out the years. I love all of you very much.

I would also like to acknowledge the financial support provided over my graduate studies. This thesis work was funded by a National Science Foundation Graduate Research Fellowship, by a National Semiconductor and Semiconductor Research Corporation (SRC) Graduate Fellowship, and by the SRC and SEMATECH contract number MC-500. I would also like to thank the National Science Foundation for awarding me a Summer Institute in Japan Fellowship which allowed me to study in Japan for a summer.

1 Introduction

Computer aided design (CAD) tools have come to play an important role in the design of integrated circuits (ICs). CAD tools currently available analyze circuit design performance, device performance, and semiconductor processing. Circuit design tools such as SPICE[80] were among the first tools to be developed and to be readily used in the IC community. As the dimensions of semiconductor devices continue to shrink, CAD tools that analyze device performance and semiconductor processing are becoming increasingly important because these tools can find problems in design prior to running costly experiments. These CAD tools that simulate device performance include BSIM[110], MINIMOS[107], and PISCES[100], and tools that simulate semiconductor processing include DEPICT[124], PROLITH[69], SAMPLE[94], SOLID[50], and SUPREM[63]. As the cost of computation decreases while the cost of experimentation increases, these CAD tools are essential in the IC community as a cost effective method to produce rapid feed back and analysis of designs.

During the optical lithography step of the semiconductor process, a mask is transferred into photoresist covering the wafer surface through the use of imaging optics. Since the light is projected onto two surfaces, the mask and the wafer, it is important to understand how light scatters from topographical features on these two surfaces. For example, nonplanar topographies in photomasks cause lateral scattering which may degrade the image collected by the imaging optics. Further, topography on the wafer may cause light to scatter into undesired areas causing reflective notching and standing waves in the photoresist. As dimensions continue to shrink in semiconductor manufacture, the aspect ratio of these topographies are becoming more severe, and understanding scattering mechanisms are becoming increasingly important. Therefore, simulation of optical lithography

can give valuable physical insight and can predict problems resulting from scattering mechanisms prior to observing these problems at the manufacturing stage.

The geometric sizes in these electromagnetic problems are on the order of one wavelength. Electromagnetic simulation of these wavelength features is computationally intensive, and neither geometric optics nor Rayleigh's method is sufficient to accurately analyze scattering effects. Furthermore, frequency domain analysis of these structures is not feasible because the features together with the surrounding features are on the order of several wavelengths which require many spatial frequencies in multiple dimensions to represent the scattered light. This would involve tracking many spatial frequencies in an iterative approach or would involve the inversion of a large matrix in an integral approach. Consequently, time domain solutions have been formulated for electromagnetic simulation for features that are on the order of a wavelength. These time domain solutions are either iterative, such as finite difference or finite element, or are integral solutions.

A notable simulation which solves Maxwells' equations using a time domain finite difference algorithm is TEMPEST[152,153]. This program, developed at UCB, is based on the algorithm proposed by Yee[155]. The initial version of TEMPEST was built by Guerrieri[46] and Gamelin[36] in two-dimensions. This version was expanded to three dimensions by Wong, and was improved computationally (less memory requirements with increased speed) by Pistor[101]. Electromagnetic wave propagation and scattering are simulated by solving iteratively the discretized Maxwell's equations until the electromagnetic field inside the simulation domain reaches steady-state.

Another physical aspect which must be included in lithography simulation is the coherence of the optical systems. Previous time domain methods to include partial coherence assume that the light is fully coherent in the space domain (spatial coherence) and in the time domain (temporal coherence). When light is fully coherent in space, fields add, however, when light is fully incoherent, intensities add. Although the light used in optical lithography often has full temporal coherence, the light is rarely fully spatial coherent. It is important to consider the simulation of this spatial coherence in optical lithography and to understand the effects of spatial coherence on image formation and on scattering from

topography. Previous simulation techniques have taken the spatial coherence into account by brute force by repeated simulations. However, this brute force method is computationally time consuming and a more optimal approach is needed.

This thesis will present: a systematic study of the effect of partial coherence in mask imaging and in wafer patterning; the first formulation of a computationally efficient method for analyzing topography scattering with vector polarized, partially coherent spatial illumination; and the implementation of this method into TEMPEST-PCD[115] (TEMPEST with Partial Coherence Decomposition). The effect of partial coherence in mask imaging is studied by understanding the influence of the partial coherence on the printability of defects on the mask. Its effect in wafer patterning is also researched by examining reflective notching from the wafer substrate. Finally, the effect of the partial coherence on wafer inspection is analyzed as well.

This thesis begins by providing a historical account of the different techniques proposed to simulate electromagnetic problems in photolithography in Chapter 2. These techniques include both scalar imaging simulation and vector electromagnetic simulation. Two techniques to simulate scalar imaging have been used in optical lithography; these are Abbe's formulation[1] and Hopkin's formulation[54]. Abbe's formulation samples an incoherent source into a number of incoherent excitations. Since one simulation must be run for each source sample, it is computationally expensive. For example, if this technique was used to find the influence of the partial coherence on contact hole inspection in Chapter 8, about 650 simulations are needed to accurately model the partial coherence. However, Hopkin's formulation is computationally efficient for calculating the aerial images, but it is based on the assumption that diffraction from the structure is independent of the polarization and angle.

Fundamental to many of these methods reviewed in Chapter 2 is the theory of imaging scalar partial coherent light. This theory is described in Chapter 3 by reviewing many sources in order to present a concise description of scalar imaging. This description is used in future chapters to explain the influence of the partial coherence on optical lithography and inspection and to serve as a foundation for extending this scalar theory to include the

vector polarization of the partial coherent light. In Chapter 3, the theory of imaging in partially coherent light is described by first defining the partial coherence and by making several assumptions applicable to optical lithography and inspection. These assumptions assume that the light is quasi-monochromatic, i.e., the temporal coherence is high. Consequently, only efficient methods of simulating the spatial coherence are considered in this thesis. Chapter 3 continues by describing the theory of scalar imaging for this partially coherent light and by describing the use of this theory in optical lithography and inspection. A method for decomposing this partially coherent light into a summation of orthogonal coherent fields is presented. The energy compaction, accuracy, and simulation implementation of this decomposition technique are discussed.

Using the theoretical foundation built in Chapter 3, Chapter 4 presents an efficient method for calculating the effect of the scalar partial coherence on defect printing in an attenuating phase shift mask. It is based on a perturbational model that is capable of rapid evaluation of the impact of the illumination, focus, defect size, defect location, defect type, and feature type on defect printing. From the model, guidelines for defect printability for attenuated phase shift mask technologies are presented. These guidelines found through the model are then compared to experimental data from printed wafer. Through the simulations and experiments, the role of the partial coherence on defect printing is emphasized in order to gain physical insight into suppressing the defect impact.

In Chapter 5, the vector nature of light neglected in the scalar imaging theory is investigated. The effect of the vector nature of light is examined as it scatters off a topographical structure that has physical dimensions on the order of one wavelength. The scattering is evaluated with TEMPEST simulation. In these simulations, a gate is patterned over an active area well. The nonplanar topography of the well causes light to scatter laterally off the sidewalls leading to reflective notching, i.e., a local region where resist has been exposed and etched away due to lateral reflection of light. As mentioned previously, these TEMPEST simulations assume that the light is monochromatic and is fully coherent in space. As well as understanding the reflective notching through TEMPEST simulation, Chapter 5 presents experimental SEM's which show reflective notching in the gate and which validate the predictive capability of the TEMPEST simulations.

In Chapter 6, the formulation for modelling the partial coherence is extended to include the vector polarization of the light. In this extension, the coherence matrix is introduced, and the elements of this coherence matrix are calculated. As shown in Chapter 3 for the case of the scalar mutual intensity matrix, this coherence matrix is decomposed in Chapter 6 into a summation of orthogonal coherent electric or magnetic fields. Since these electric or magnetic fields represent coherent excitations, TEMPEST can then be used to simulate the scattering of these individual excitations from topography. This extension for including the decomposition of the vector polarized partial coherence is included in the existing TEMPEST program, and this new program is called TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition). The energy compaction of TEMPEST-PCD is compared to Abbe's formulation. The accuracy of TEMPEST-PCD is checked in Chapter 6 by comparing aerial images of an attenuated phase shift mask calculated with TEMPEST-PCD to aerial images calculated with Abbe's formulation.

Chapter 7 explains effects that vector polarized partial coherence have on mask imaging and reflective notching. For mask imaging, aerial images of an attenuated phase shift mask are calculated with TEMPEST-PCD and compared to aerial images calculated with Abbe's formulation and with scalar image theory. To determine the effect of the partial coherence on reflective notching, TEMPEST is used to pattern a gate into photoresist over an active area well. This well structure is similar to the topographical structure used in Chapter 5. Since the photoresist used in this example is chemically amplified, a post exposure bake (PEB) is needed to diffuse the acid concentration [160]. Therefore, the effect of the PEB versus the partial coherence is also studied in Chapter 7.

The difficult problem of simulating optical inspection of contact holes with partially coherent light is investigated in Chapter 8. TEMPEST-PCD is used to evaluate the effectiveness of a pupil filter to improve contact hole inspection. The chapter begins by using a ray tracing method to find the induced spherical aberration to detect a defect at the bottom of a thin-film stack. Building on this ray tracing work, TEMPEST-PCD is used to simulate the inspection of a contact hole. In order to optimize the inspection, a pupil filter is devised by using the reciprocity theorem of electromagnetics [4]. A point source is placed at the bottom of the contact hole, and the radiated fields are measured at the top of

the simulation domain. By the reciprocity theorem, these fields represent the excitations needed to image a point at the bottom of the hole. These fields can then be projected back through the imaging system to any plane in order to create an optimal pupil filter. TEMPEST-PCD is then used to determine the effectiveness of this pupil filter. Since this filter is only optimum for coherent illumination, TEMPEST-PCD is used to explore how partial coherence reduces the influence of process variation to produce a filter that is still advantageous to inspection.

Chapter 9 concludes the thesis. In the conclusion, the complexity of implementing the partial coherence into the TEMPEST simulation program is discussed. Since simulation of the partial coherence requires more computational time than simulation of coherent light, guidelines are presented which outline when simulation of the partial coherence is necessary. Furthermore, if the partial coherence is simulated, the increased computational cost is discussed as well.

2 Simulation Techniques for Modelling the Partial Coherence in Optical Lithography and Inspection

2.1. Introduction

In this chapter, previous studies concerning the effects of the partial coherence in optical lithography and inspection are described with emphasis on simulation techniques. Simulation techniques for scalar imaging have been developed over the past century. In 1873, Abbe formulated one of the first techniques to describe the imaging of a microscope. Later, in 1951, Hopkins developed an alternative approach to Abbe's formulation. Hopkins' approach is computationally more efficient, however it assumes that diffraction is independent of the illumination angle. Recently, another alternative approach that Gamo formulated in 1964 is gaining much attention in the lithography community as an optimal methodology for including the partial coherence in optical proximity correction and in pupil filter design. With the advent of faster computers with larger memories beginning in the 70s, techniques that directly simulate the propagation and scattering of the electromagnetic field are now feasible. These techniques can be classified as either frequency-domain or time-domain methods and are briefly highlighted in this chapter by literature review. This chapter concludes by presenting previous work concerned with important effects observed in lithography. These include a description of the effect of the partial coherence on linewidth variation, on defect printing, on reflective notching, and on pupil filtering. This work will serve as a foundation for developing a theory that will help explain phenomena observed in these problems.

2.2. Scalar Imaging Simulation Techniques

In scalar imaging, the coupling between the electric and magnetic fields is ignored at the physical boundaries of the object under study. When the coupling is ignored, all the vector components of the electric and magnetic field can be imaged separately through scalar Kirchhoff diffraction theory [120]. The coupling between the electric and magnetic field is negligible when the structure under study is large compared to the wavelength of the light. This is equivalent to the requirement that the diffraction angles caused by the structure are small.

Although optical lithography and inspection are almost always concerned with studying structures that are on the order of a wavelength, scalar theory is often used as a first order model to simulate optical imaging. An important consideration in optical lithographic imaging and inspection is that the illumination source is partially coherent in the spatial domain. Historically, two techniques have been formulated to model scalar imaging which include partial coherence. These techniques are Abbe's formulation and the more widely used Hopkins' formulation. These techniques and their historical background are discussed in the following sections.

2.2.1. Abbe's Formulation

Abbe's formulation [1] was originally proposed as a method to model the scalar imaging performed by a microscope when the illumination source is incoherent. Since the source is incoherent, the total intensity in the image plane can be found by sampling the incoherent source with a number of point sources. In a microscope with Köhler illumination [12], each of these point source samples produce a fully coherent plane wave that is obliquely incident on the object. These oblique waves, when projected into a simulation domain, produce an excitation with a sine wave modulation. Since the source is incoherent, each excitation is orthogonal to the other excitations. The intensity at the image plane is then found by using the Kirchhoff diffraction integral to calculate the image intensity from each of these incident plane waves. The total intensity is then the summation of the intensities produced by each of the plane wave excitations. Thus, the coherence is modelled by

summing (integrating) over the plane waves after their individual intensities have been evaluated.

Since the optical system used in lithography and inspection is essentially a microscope, this formulation can be used in lithography to simulate scalar imaging. However, since the source is integrated over last in Abbe's Formulation, it is computationally intensive and is rarely used to simulate scalar imaging. The more computationally efficient Hopkins' Formulation is the method of choice for scalar imaging with partially coherent light.

2.2.2. Hopkins' Formulation

In Hopkins' formulation[54] the coherence is modelled by integrating over the source first before integrating over the diffraction orders. That is, the order of integration is switched. The order of integration can be switched if the plane wave excitations produced by Abbe's formulation create diffraction orders having magnitudes that are independent of the oblique angle of incidence. Since scalar imaging theory always implies that the coupling between the electric and magnetic fields is negligible, the magnitude of the orders are always assumed to be independent of the incident excitation angle. By integrating over the source first, a transmission cross coefficient is created that takes the partial coherence of the system into account. By creating these transmission cross coefficients which depend only on the optical system and are independent of the mask, Hopkins' formulation is capable of even further improving the computational efficiency.

Following the publication by Hopkins in 1953, his formulation was widely used in optics through the late 50s and 60s to find the aerial images of objects illuminated by partially coherent light with some of the earlier work being highlighted here. Since elements of the optical system are circular, Hopkins' formulation has been used extensively in the literature to find the image of circular objects. One of the earliest applications of using Hopkins' formulation was by Weinstein who calculated the image of disks with diameters on the same order of magnitude as the Airy disk [140]. Later, De and Som also used the formulation to calculate the image of circular apertures[28], of circular phase objects[29], and of annular objects[118] when illuminated with partially coherent light. Imaging a circular hole with Hopkins' formulation was investigated by Charman[23]. Canals-Frau and

Rousseau described the images of an opaque disk[19]. As well as using Hopkins' formulation to image circular objects, many people, Hopkins[52], Steel[121], Canals-Frau and Rousseau[19], Considine[27], Thompson[128], and many others, have used the formulation to find the diffraction images of non-rotationally symmetric objects such as an edge, a slit, a line, and a three bar target. In addition to these studies, the formulation, along with a mathematical framework for the partial coherence theory, has been presented in books, with Born and Wolf[9], Beran and Parrent[7], Marathay[73], Goodman[41], and Mandel and Wolf[72] being consulted for this thesis work.

Hopkins' formulation has been widely used to simulate imaging in the optical lithography community. Hopkins' formulation has been used to find the image degradation due to partial coherent illumination through focus. O'Toole[95] implemented Hopkins' formulation to find the image of a square aperture in photoresist. Kintner[61] also used a method based on Hopkins' formulation to describe partially coherent imagery through focus. Subsequently, Subramanian[122] later modified this work by calculating the transmission cross coefficient. Based on this work, Toh implemented Hopkins' formulation into a lithography simulation tool called SPLAT[126,64]. Many other aerial imaging tools based on Hopkins' formulation, such as PROLITH[69] and iPHOTO[104], have also been developed.

2.2.3. Decomposition Formulation

As described above, Abbe's formulation takes the partial coherence of a source into account by expanding the partial coherence of the source into a number of fully coherent, orthogonal, sine wave excitations. Since the source is incoherent, each sine wave excitation is incoherent to the other excitations. A sine wave excitation results when taking the Fourier transform of a sampled point on the source which implies that Abbe's technique is a Fourier transform technique.

In addition to Abbe's formulation, alternative expansions of the partial coherence produced by the source are also possible. In these alternative expansions, the partial coherence in the object plane is first calculated as a function of the source coherence. The partial coherence is described by the mutual intensity. The mutual intensity at the object plane can

be found with the van Cittert-Zernike theorem [73], which states that the mutual intensity of an incoherent source is the Fourier transform of the source. This implies that light gathers coherence as it propagates. This can be understood by considering a stellar example. The light produced by stars is incoherent. The light of our sun observed on earth is still incoherent because the light has traveled over a short distance relative to the sun's size. However, when Alpha Centauri, which is also an incoherent source, is viewed from earth, the light from Alpha Centauri is measured to be fully coherent because the light has propagated over a great distance. By using the fact that the light from an incoherent source gathers coherence as it propagates, an alternative transform technique is sought that takes advantage of this gathered coherence.

Several transformations have been studied in communication theory including the Fourier transform, cosine transform[66], wavelet transform[135], and Karhunen-Loeve transform[41]. The latter transform, Karhunen-Loeve, has been proven to have the optimal energy compaction [67], i.e., the least number of mutually uncorrelated signals are needed to represent a partially correlated signal. It is important to note that the Fourier transform (Abbe's formulation) is not optimal in this sense. The Karhunen-Loeve transform diagonalizes the correlation matrix into a set of eigenvalues and eigenfunctions. Each of these eigenfunctions represent fully coherent excitations yet are incoherent to the other eigenfunction excitations. This diagonalization technique is called the decomposition formulation throughout this thesis.

The Karhunen-Loeve transformation technique has been applied previously in optics to remove the correlation in the mutual intensity. Gamo[38] showed that the mutual intensity can be decomposed into a set of orthogonal eigenvalues and eigenfunctions. With this technique, the mutual intensity is represented by an incoherent sum of coherent fields. This technique, as proposed by Gamo, was later rediscovered by Wolf[150] and applied by Saleh *et al.*[106] to find the aerial image of an edge and of a star pattern. In optical lithography, this decomposition technique has also been successfully applied to optical proximity correction [97,24] and to depth of focus enhancement [16].

2.3. Vector Electromagnetic Simulation Methods

In this section, methods of simulating the propagation of vector electromagnetic fields are discussed with those used in lithography being highlighted. Most of the methods simulate light that is monochromatic and fully coherent in space while methods to simulate monochromatic, partially coherent, vector electromagnetic fields are relatively few. In Section 2.3.1, simulation of fully coherent fields is first discussed followed by the simulation of partially coherent fields in Section 2.3.2.

2.3.1. Coherent Excitations

Many sources that describe the simulation of monochromatic, fully coherent electromagnetic fields exist in the literature. Those used in photolithography simulation are highlighted in this section. In electromagnetic problems of interest in photolithography, typical feature sizes are on the order of one wavelength. In this regime, the coupling between the electric and magnetic fields cannot be neglected; consequently, the scalar techniques discussed previously are inaccurate. The problem of addressing this coupling between the fields and of considering the vector components of these fields has been addressed by various techniques. These techniques can be classified as either frequency-domain or time-domain methods.

2.3.1.1. Frequency Domain Methods

Several frequency domain methods have been proposed to simulate the scattering of light from topographic features used in lithography. These methods include Rayleigh's method, the waveguide method, and differential and integral methods.

In Rayleigh's method, the fields are expressed as a linear superposition of propagating and evanescent waves[105]. Later, Petit and Cadihac found that Rayleigh's method was only valid when the product of the wave number times the grating depth is less than a unitless quantity 0.448 [99]. This implies that the grating depth must be less than 0.07 wavelengths or 26nm for features illuminated with light having a wavelength of 365nm. This is a severe restriction in the simulation of lithography because structures studied almost always have depths that vary by 0.07 wavelengths. Despite this restriction, Gall-

atin[35] as well as Bobroff and Rosenbluth[8] used Rayleigh's method to simulate imaging of alignment marks under photoresist.

In the waveguide method, Maxwell's equations are solved by a separation of variables method. In this method, the electric fields are equivalent to a truncated infinite series that satisfy a set of equations found by separating the variables. Burckhardt used this waveguide method to find the fields diffracting from a sinusoidal dielectric grating [18], and later, Kasper used the method for non-sinusoidal lossy dielectric gratings [59]. Nyyssonen and Kirk extended this work to examine scattering from alignment marks [85]. Lucas also used the technique to study the image formed by a three dimensional phase shift mask [68]. The one disadvantage of the waveguide method is that it is computationally expensive. For three dimensional problems, Yeung states that the method has a storage cost on the order of n^4 and a cpu cost of n^6 where n is the number of grid points in each dimension [159].

In frequency domain differential methods, the electromagnetic field and the wave number squared are expanded into a Fourier series. When this expansion is substituted into the Helmholtz equation, a set of coupled differential equations results. Neviere used this method to study resonance in holographic film couplers [83]. However, this method produces incorrect results for highly conducting materials. Neureuther and Zaki proposed an integral method which alleviates this problem in differential methods when highly conducting materials are present [81]. In this method, a set of integral equations are solved where the kernels of these integrals are the periodic Green's function and its derivative. Several other integral methods have also been proposed. Most notably, a finite element method to solve the integral equations has been used by Matsuzawa to find a photoresist image above a conducting substrate [75]. This approach of Matsuzawa was improved by Urbach and Bernard to include more general domains and to include partial coherence [132]. Using a set of Legendre polynomial basis functions, Barouch simulated the reflective notching from a three-dimensional substrate using a finite element method [6]. Direct solution of any of these integral equations can be solved through the method of moments. However, solution of these integral frequency domain equations is impractical in three

dimensions because the solution involves inverting a large matrix which is computationally expensive. In Barouch's choice of basis functions, however, the computational cost is reduced because the matrix may be condensed. Yeung further reduced the computational cost by using a fast multipole method to find an iterative solution of these integral equations [158,159]. In Yeung's method the storage cost is reduced to n^2 and the cpu cost reduce to $n^{3.33} \times N$ where n is the number of grid points in one dimension and N is the number of iterations [159].

2.3.1.2. Time Domain Methods

In the previous section, frequency methods were described to solve Maxwell's equations in which the excited field is monochromatic and fully coherent. These techniques are rigorous and accurate. However, since lithography problems have feature dimensions that are on an order of a wavelength, many spatial frequencies are needed to represent the scattering from these features. Since many spatial frequencies are required, frequency methods prior to Yeung's work required inverting a large matrix. Time domain methods, however, require no matrix inversion since Maxwell's differential equations are solved by a time marching iterative approach. These differential, time domain, iterative methods require many additions and multiplications. In these time domain methods, the storage cost is on the order of n^3 and the cpu cost is on the order of $n^3 \times N$ where n is the number of grid points in one dimension and N is the number of iterations [159].

The finite difference time domain (FDTD) method is one time domain method that has been used to simulate lithography. In a FDTD method, Maxwell's equations are solved at discrete points on a cubic staggered grid where the excitation field is monochromatic and fully coherent, i.e., electric and magnetic fields add. These equations are then iterated in time through a leap frog technique proposed by Yee [155]. In order to study problems of interest in photolithography, Guerrieri[46] formulated and Gamelin[36] implemented a two-dimension FDTD program called TEMPEST[153] on a connection machine. TEMPEST was extended to three dimensions by Wong[152], and the speed and memory efficiency was improved by Pistor[101]. In addition to this finite difference method, Wojcik applied a finite element method to solve Maxwell's equations in the time domain [149]. A

finite element method has the advantage that a conformal grid can be placed on a surface in contrast to a finite difference grid which represents a surface as a staggered stair case. In the past, because of the computational demands required by time domain methods, super computers were required to solve problems of interest in lithography. With the advent of fast computers having large memories, time domain methods because of their ease of implementation are becoming more feasible.

2.3.2. Partially Coherent Excitations

The previous section described methods for simulating fields that are monochromatic and fully coherent in the spatial domain. Although the light used in a photolithographic system or inspection system is often monochromatic, the light is rarely fully coherent. Therefore, this section describes previous methods that simulate the partial coherence in optical lithography and optical microscopy.

In optical lithography, the source is incoherent in the spatial domain and produces Köhler illumination. Through Abbe's formulation as described in Section 2.2.1, the source is sampled into a discrete number of point sources where each point source is imaged as an obliquely incident plane wave. Since the source is incoherent, the total intensity is then the summation of the intensities produced by each of these obliquely incident plane wave. In the simulation of optical lithography, Abbe's formulation has been implemented by Urbach and Bernard[132] and by Wojcik[149]. In the approach of Urbach and Bernard, the source is sampled into a number of point sources. These point sources are imaged by the optical system as sine wave excitations. The scattered electromagnetic fields produced by these sine wave excitations are calculated through a frequency domain, finite element method. Like Urbach and Bernard, Wojcik used Abbe's formulation and calculated the scattered electromagnetics fields through a time domain, finite element method. Abbe's formulation, however, suffers from the disadvantage that it is computationally intensive because one simulation must be run for each sampled source point.

In Section 2.2.2, Hopkins' formulation for scalar light which efficiently models the partial coherence can be applied to simulate vector, partially coherent fields under one certain restriction. This restriction is that the magnitude of the diffraction orders produced by

a structure must not be a function of the angle of oblique incidence. Under this restriction, the diffraction orders need only be calculated with a single normally incident plane wave. The partial coherence of the system is then modeled by using Hopkins' formulation. Using Hopkins', the diffracted orders from the single, incident plane wave are collected and weighted by the transmission cross coefficients of the optical system. This weighting by the transmission coefficients produces an aerial image resulting from partially coherent illumination. This method has been implemented into TEMPEST, and in this thesis it is referred to as TEMPEST-HN (TEMPEST with Hopkins' imaging for a Normally incident plane wave). Wojcik[149] has also successfully applied this method to evaluate the aerial images due to a phase shift mask. However, this method only applies to aerial image calculation and does not apply to imaging into the photoresist over some wafer topography. Furthermore, for some masks, the assumption that the diffraction orders are independent of excitation angle is not valid.

2.4. Effect of Partial Coherence in Optical Lithography and Inspection

In this section, the effect of partial coherence on printing and inspecting small features on wafers is examined by presenting previous work in the literature. In optical lithography, the performance of a optical system is characterized by measuring the linewidth, as known as the critical dimension (CD). In the first section, previous work describing the effect of the partial coherence on the CD is presented. Later sections highlight previous work in defect printing, in reflective notching, and in pupil filter design.

2.4.1. Effect of the Partial Coherence on Critical Dimension

The performance of steppers is significantly affected by a change in numerical aperture and in partial coherence (σ) when operating at the resolution limit. For example, when printing a grating with a line-space width on the order of a wavelength, the CD remains constant over a larger focus window when the partial coherence increases, i.e., the depth of focus increases as σ increases. This is often called the proximity effect in the lithography community, and has lead to a trend of using higher σ. This proximity effect is shown through experiment and through simulation by Mack[71], Canestrari[21], Partlo[96], Yamanaka[154], André[2], and many others. In contrast to this work, King[60]

observed an increased depth of focus for structures which are far from the resolution limit. Due to this trade-off between resolution and depth of focus, σ must be adjusted to best maintain the CD uniformity over the wafer.

As well as affecting the CD, the uniformity of the partially coherent source can impact the uniformity of the CD across the exposure field. Canestrari[21] found that the uniformity of the source varies by $\pm 4.5\%$. This can adversely affect the printing of dense features and can reduce the effectiveness of optical proximity correction. Borodovsky[14] observed that variations in the local partial coherence across the exposure field might be responsible for the excessive linewidth variation and for poor uniformity across the field. Progler[103] also observed that, in addition to optical aberrations, variation in partial coherence, variation in intensity, variation in NA, and stray light cause CD variation across the field.

2.4.2. Defect Printing

Experiment, aerial image simulation and algebraic modelling are useful in characterizing the effect of defects. The impact of defect printing on features for standard chrome masks has been studied through experiment [143,144,20,146] and through modelling and simulation [82,74,56]. However, characterizing the printing of defects is now more complicated due to the advent of phase shift masks (PSM's)[65]. Problems in manufacturing PSM's may cause phase shift defects which are more difficult to detect and to repair [136,145,138,161,87]. Detection of these phase defects depends on parameters associated with the tool, with the mask, and with the defect. The problem is compounded by interactions between the defect and the feature. These interactions depend on defocus, on defect location, and on coherence. Simulation and experimental studies on PSM's have shown the complexity of PSM's by demonstrating the tendency of phase shift defects to print when out of focus [136,111,114]. The printing of phase defects is further complicated by using modified illumination [108,113] and by printing through the thin-film layers lying on the wafer [113].

2.4.3. Reflective Notching

As the wavelength used in lithography decreases, the reflectivity of silicon increases. This increase in reflectivity causes an increase in the standing wave amplitude on a flat silicon wafer. Numerical solutions for calculating the standing wave intensity in photoresist have been studied [77,62,57,141,32]. An analytical expression for the standing wave intensity in photoresist has been calculated by Mack[70] based on the work of Berning[5]. This standing wave effect is important because the coupling of light into the photoresist causes the CD of the feature to vary through focus where this variation is a function of resist thickness and of illumination partial coherence [131].

The topography of the wafer substrate can cause light to scatter light into unexposed areas which may adversely effect the CD. The effect is particularly noticeable where positive tone photoresist lines must cross concave upward features. This reflected light may cause reflective notching where concentrated light causes linewidth narrowing as observed by Widmann[142]. In order to reduce this narrowing, Petersen[98] has proposed a design methodology to minimize the linewidth variation in photoresist patterns over polysilicon topography by systematically varying the resist thickness, the bake temperature, and the bake time. Reflective notching in gate lines have also been corrected by using anti-reflective coatings [84,86,31] and by increasing the absorption of the photoresist by adding dye [55,127]. As well as observing the notching experimentally, simulation has proven to be a valuable tool for analyzing various corrections prior to performing time consuming experiments. [148, 75, 132, 37, 123, 129, 112, 116].

2.4.4. Pupil Filter Design

In the printing and inspection of wafers, spherical aberration is induced by photoresist when printing with a high numerical aperture stepper [33]. Yeung[157] showed that the effect of arbitrary thin-film layers on a wafer can be modelled by ray tracing through the thin-film layers to find the optical path difference (OPD). This OPD induces aberrations which can be modelled by modifying the pupil function of the imaging system. Similar to this work, Progler[102] then proposed using a merit function that is capable of determining which aberrations may actually benefit printing. This merit function is dependant on the mask feature and on the various thin-film layers covering the wafer.

As well as studying the effect of thin-film layers on optical printing, designing pupil filters that improve depth of focus have been studied. In these studies, Tsujiuchi[130] first described a cosine filter to improve depth of focus. This filter was later re-invented by Ojeda-Castañeda[92] and was implemented by Fukuda[34]. The effect of filtering with a Fresnel zone plane [58], a partitioned pupil [133], and various analytic pupil functions was also studied by several authors[79,91,109]. By using McCutchen's theorem [76], Ojeda-Castañeda designed filters for increased depth of focus [88,89]. Alternatively, von Bünau[15-17] designed a filter for increased depth of focus by using the decomposition technique described in Section 2.2.3. Pupil filters which reduce the effect of aberrations other than defocus have been studied as well [90,93,130].

2.5. Summary

Electromagnetic simulation of photolithography is difficult because many factors are involved in forming an image. These factors include parameters of the optical system, of the mask, and of the wafer. Since many parameters must be considered, simulation of photolithography problems can be computationally expensive. Therefore, efficient algorithms are needed to solve these expensive problems, and guidelines are needed to understand which parameters are important. The following chapters present an efficient algorithm for the simulation of partially coherent light and present guidelines when this simulation is necessary. The algorithm involves an extension to TEMPEST to model the vector scattering of electromagnetic fields when these fields are partially coherent. This extension is based on generalizing the decomposition of the scalar mutual intensity proposed by Gamo to include the polarization of the partially coherent electromagnetic field. After developing this algorithm, it is applied to the simulation of problems encountered in optical lithography and inspection. These simulations are accompanied by experimental verification. From these simulations and experiments, physical insight into the effects of partial coherence is presented. From this insight, guidelines for understanding the impact of the partial coherence are developed for problems encountered in optical lithography and optical inspection.

3 Theory of Scalar Partial Coherent Light

3.1. Introduction

This chapter presents the theory of scalar imaging which creates a foundation for extensions of this theory in subsequent chapters of this thesis. The theory of scalar imaging with partially coherent light is widely discussed in the literature [1,9,41,42,49,72,73]. Therefore, the purpose of this chapter is not intended to justify the validity of scalar imaging, but to bring together many sources from the literature to form a theoretical foundation which will be presented in this chapter and expanded in future chapters. The expansion of this theoretical foundation will then allow the development of models capable of analyzing problems in optical lithography and inspection that were previously difficult to model.

Before considering these new problems, the theoretical foundation is first built by describing the temporal and spatial coherence in Section 3.2. Scalar imaging from the object plane to the image plane is then described in Section 3.3 for three states of spatial coherence. These three states are coherent light, partially coherent light, and incoherent light. The pupil function of the imaging system is presented in Section 3.4. This pupil function represents the transfer function of the optical system in the frequency domain and describes the imaging from the object plane to the image plane. This scalar theory is used to model imaging in the optical lithography system that is described in Section 3.5. Since the light in a lithography and inspection system is partially coherent, Section 3.6 uses the theory presented in Section 3.3 to describe two methods for imaging partially coherent light. These two methods are Abbe's Formulation and Hopkins' Formulation. In Section 3.7, the decomposition of the mutual intensity calculated by Hopkins' Formulation which was developed by Gamo[38] is presented using the notation of von Bünau. The

accuracy of this decomposition is analyzed in Section 3.8 by comparing the mutual intensity calculated with the decomposition technique to the mutual intensity calculated analytically. Since the source is radially symmetric, the decomposition technique is performed in radial coordinates which may cause problems when the simulation domain is rectangular. These problems are discussed in Section 3.9.

3.2. Temporal and Spatial Coherence

The coherence of light is divided into two classifications, temporal and spatial. The ability of light emitted from the same point to interfere with a delayed wavefront of itself is described by means of temporal coherence. The ability of light along the same wavefront to interfere is described by means of spatial coherence.

The degree of temporal coherence is quantified by the coherence time, τ_c. The coherence time is the temporal interval over which the phase of a light wave can be reasonably predicted at a given point in space. The coherence time is related to the bandwidth of the light source, $\Delta\nu$, and is proportional to $1/\Delta\nu$. Another parameter used to measure the temporal coherence is the coherence length, l_c, which is given by $c\tau_c$ where c is the speed of light. When $\Delta\nu$ is zero, the coherence time is infinite, and the light is said to be monochromatic, even though truly monochromatic light is never attainable. However, it is often the case in optical lithography that the maximum path length difference in the passage of light from the source to the mask or to the wafer is much smaller than the coherence length. Under this condition the light is quasi-monochromatic and the field $\mathbf{E}(\mathbf{r},t)$ is given by $\mathbf{E}(\mathbf{r})e^{i\omega t}$. Throughout this dissertation, quasi-monochromatic light is assumed. This is not a limitation since light used in optical lithography is often nearly monochromatic. Furthermore, if the assumption that light is quasi-monochromatic is not valid for a particular lithography system, the temporal coherence can be modelled by sampling the frequency spectrum at discrete frequencies with a sampling interval in which the light is quasi-monochromatic over this interval.

The degree of spatial coherence is quantified by the mutual coherence, $J(\mathbf{x}_1,\mathbf{x}_2,\tau)$. Mutual coherence measures the correlation between two points separated in space and is

related to the spatial interval over which the phase of the light wave can be reasonably predicted at a given point in time. The mutual coherence at two points depends on the propagating mechanisms of the light and is also related to the physical size of the source. For a field that is statistically wide sense stationary, i.e., the mean is independent of time and the correlation function is a function of the time difference only, the mutual coherence is a time averaged intensity between the field at two points, x_1 and x_2, over a period of τ as shown in Equation 3-1. These two points, x_1 and x_2, can be located at any reference plane, such as at the source or at the object plane. Statistically, Equation 3-1 implies that the mutual intensity is the autocorrelation between two electric fields. Equation 3-1 can be simplified under the quasi-monochromatic approximation. The optical spectral intensity, $J(\nu)$, of quasi-monochromatic light is approximately a delta function; therefore, the mutual coherence is the product of the mutual intensity, $J(x_1,x_2)$, and $e^{-i2\pi\nu\tau}$ as shown in Equation 3-2.

Equation 3-1. $$J(x_1, x_2, \tau) \equiv \langle E(x_1, t) E^*(x_2, t+\tau) \rangle$$

Equation 3-2. $$J(x_1, x_2, \nu) = J(x_1, x_2)\,\delta(\nu) \Rightarrow J(x_1, x_2, \tau) = J(x_1, x_2)\, e^{-i2\pi\nu\tau}$$

The mutual intensity can be normalized by the intensity at the two points, x_1 and x_2, and is called the complex degree of mutual coherence, $\mu(x_1,x_2)$. The complex degree of the mutual coherence is defined in Equation 3-3 and its value ranges between 0 and 1.

Equation 3-3. $$\mu(x_1, x_2) = \frac{J(x_1, x_2)}{\sqrt{J(x_1, x_1)}\sqrt{J(x_2, x_2)}}$$

If the light at two points is completely correlated, μ=1, the light is spatially coherent and the radiated fields add. If the light is completely uncorrelated, μ=0, the light is spatially incoherent and the intensities at these two points add. Mathematically, incoherence implies that the electric fields between two points are orthogonal. When μ is between 0 and 1, the light is partially coherent. Because the quasi-monochromatic assumption is used throughout this thesis, the word, coherence, will always refer to the spatial coherence rather than the temporal coherence. Therefore, incoherent light refers to spatial incoherent light, coherent light to spatial coherent light, and partial coherent light to spatial partial coherent light.

3.3. Imaging in the Spatial Domain

3.3.1. Coherent Light

The image of an object illuminated by coherent light serves as the foundation for imaging scalar partially coherent light and as the foundation for second order coherence theory. The image of a coherent object as it propagates through an imaging object is illustrated in Figure 3-1. The object diffracts the light which is collected by an optical system

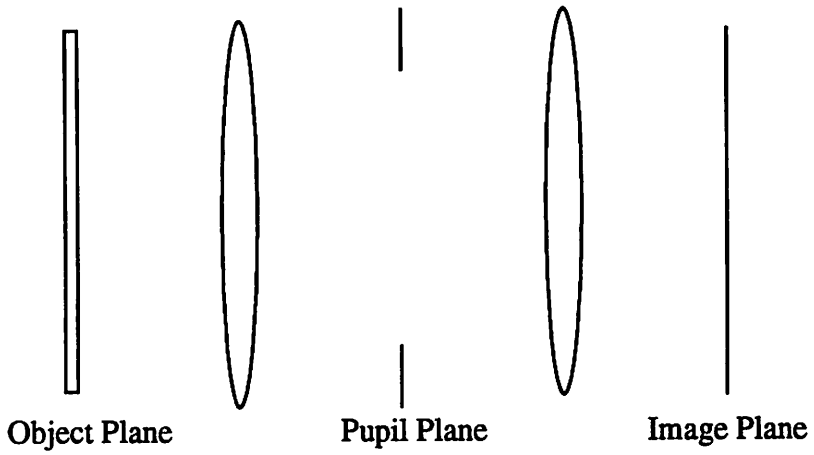


Figure 3-1. Coherent optical imaging system

The object plane diffracts the light which is collected by optical system and imaged at the image plane. Optical system here is represented as two lens and a pupil.

and is projected onto the image plane. The field of the object, a_o, is first formed at the pupil plane, a_p, through Fresnel-Kirchhoff diffraction theory [43]. Under the Fresnel approximation, the quadratic phase factors are eliminated because of the lens law, and the image at the pupil plane is given by Equation 3-4. In Equation 3-4, the paraxial approximation has been made as the light is assumed to diffract into small angles. The paraxial approximation is not a limitation because high angular diffraction due to the lens can be modelled by modifying the pupil function, $P(\tilde{x}_p, \tilde{y}_p)$, to include an obliquity factor. The field at the pupil plane, a_p, is then formed on the image plane, a_i, again through Fresnel-Kirchhoff diffrac-

tion theory as shown in Equation 3-5. In Equations 3-4 and 3-5, $\tilde{x}$ and $\tilde{y}$ represent physical

Equation 3-4. $$a_p(\tilde{x}_p, \tilde{y}_p) = \frac{1}{i\lambda f}\int_{-\infty}^{\infty}\int a_o(\tilde{x}_o, \tilde{y}_o)\, exp\left[i\frac{2\pi}{\lambda f}(\tilde{x}_o\tilde{x}_p + \tilde{y}_o\tilde{y}_p)\right]d\tilde{x}_o d\tilde{y}_o$$

Equation 3-5. $$a_i(\tilde{x}_i, \tilde{y}_i) = \frac{1}{i\lambda f}\int_{-\infty}^{\infty}\int a_p(\tilde{x}_p, \tilde{y}_p)\, P(\tilde{x}_p, \tilde{y}_p)\, exp\left[i\frac{2\pi}{\lambda f}(\tilde{x}_p\tilde{x}_i + \tilde{y}_p\tilde{y}_i)\right]d\tilde{x}_p d\tilde{y}_p$$

dimensions and have length units and f is the focal length of the imaging system. These equations are simplified by transforming the physical dimension variables, $\tilde{x}$ and $\tilde{y}$, into normalized variables, x and y as shown in Equations 3-6, 3-7, and 3-8. The normalized

Equation 3-6. $$\mathbf{x}_o = (x_o, y_o) \equiv \frac{NA}{\lambda}(\tilde{x}_o, \tilde{y}_o)$$

Equation 3-7. $$\mathbf{x}_i = (x_i, y_i) \equiv \frac{NA}{\lambda}(\tilde{x}_i, \tilde{y}_i)$$

Equation 3-8. $$\mathbf{x}_p = (x_p, y_p) \equiv \frac{1}{a}(\tilde{x}_p, \tilde{y}_p)$$

dimensions at the object plane and image plane in Equations 3-6 and 3-7 are normalized by the wavelength, λ, and the numerical aperture, NA, which is equivalent to a/f under the paraxial assumption. The pupil dimensions in Equation 3-8 are normalized by the physical size of the lens, a. These normalized units are used throughout this dissertation unless otherwise stated. After substituting in the normalized variables, Equations 3-4 and 3-5 simplify to Equations 3-9 and 3-10, respectively. Finally, the field at the image plane is found

Equation 3-9. $$a_p(\mathbf{x}_p) = \frac{\lambda f}{ia^2}\int_{-\infty}^{\infty} a_o(\mathbf{x}_o)\, e^{i2\pi\mathbf{x}_o\bullet\mathbf{x}_p}d\mathbf{x}_o$$

Equation 3-10. $$a_i(\mathbf{x}_i) = \frac{a^2}{i\lambda f}\int_{-\infty}^{\infty} a_p(\mathbf{x}_p)\, P(\mathbf{x}_p)\, e^{i2\pi\mathbf{x}_p\bullet\mathbf{x}_i}d\mathbf{x}_p$$

by substituting Equation 3-9 into Equation 3-10, and shown in Equation 3-11. From

Equation 3-11, the field at the image plane is the convolution of the field at the object plane with the image transfer function, $K(\mathbf{x})$ which is given by Equation 3-12.

Equation 3-11.
$$a_i(\mathbf{x}_i) = \int_{-\infty}^{\infty} a_o(\mathbf{x}_o) K(\mathbf{x}_i - \mathbf{x}_o)\, d\mathbf{x}_o$$

Equation 3-12.
$$K(\mathbf{x}) = \int_{-\infty}^{\infty} P(\mathbf{x}_p)\, e^{-2\pi i \mathbf{x}_p \bullet \mathbf{x}} d\mathbf{x}_p$$

3.3.2. Partial Coherent Light

The mutual intensity of an object illuminated by partial coherent light as shown in Figure 3-1 can be found by generalizing the theory of imaging coherent light in Section 3.3.1. The mutual intensity in the image plane, $J_i(\mathbf{x}_{i1},\mathbf{x}_{i2})$, is defined by Equation 3-13 and is found by substituting the fields in the image plane, $a_i(\mathbf{x}_{i1})$ and $a_i(\mathbf{x}_{i2})$, given by Equation 3-11 into this definition. After substituting, the time average is then taken resulting in Equation 3-14 which gives the mutual intensity in the image plane as a function of the image transfer function, K, and as a function of mutual intensity in the object plane, $J_o(\mathbf{x}_{o1},\mathbf{x}_{o2})$. In deriving Equation 3-14, the definition of $J_o(\mathbf{x}_{o1},\mathbf{x}_{o2})$ was used, i.e., $J_o(\mathbf{x}_{o1}, \mathbf{x}_{o2}) \equiv \langle a_o(\mathbf{x}_{o1})\, a_o^*(\mathbf{x}_{o2}) \rangle$.

Equation 3-13.
$$J_i(\mathbf{x}_{i1}, \mathbf{x}_{i2}) \equiv \langle a_i(\mathbf{x}_{i1})\, a_i^*(\mathbf{x}_{i2}) \rangle$$

Equation 3-14.
$$J_i(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \int_{-\infty}^{\infty}\int J_o(\mathbf{x}_{o1}, \mathbf{x}_{o2})\, K(\mathbf{x}_{i1} - \mathbf{x}_{01})\, K^*(\mathbf{x}_{i2} - \mathbf{x}_{o2})\, d\mathbf{x}_{o1} d\mathbf{x}_{o2}$$

3.3.3. Incoherent Light

The mutual intensity of an object illuminated by incoherent light is a limiting case of imaging with partial coherent light. The mutual intensity in the image plane is found by using the equations of Section 3.3.2. The fields at two separate points, $\mathbf{x}_{01}$ and $\mathbf{x}_{02}$, in an incoherent source placed in the object plane are, by definition, completely uncorrelated. This implies that the mutual intensity is zero when $\mathbf{x}_{o1} \neq \mathbf{x}_{o2}$ and non-zero when

$\mathbf{x}_{o1} = \mathbf{x}_{o2}$ because the field must be correlated with itself. Mathematically this implies that the mutual intensity at the object plane is given by Equation 3-15.

Equation 3-15. $$J_o(\mathbf{x}_{o1}, \mathbf{x}_{o2}) = I(\mathbf{x}_{o1})\,\delta(\mathbf{x}_{o1} - \mathbf{x}_{o2})$$

The mutual intensity at the image plane is found by substituting the mutual intensity at the object plane, J_0, (Equation 3-15) into Equation 3-14. After replacing the image transfer function, K, in Equation 3-14 with Equation 3-12, the mutual intensity at image plane for an incoherent object is given by:

Equation 3-16. $$J_i(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \int_{-\infty}^{\infty}\int d\mathbf{x}_{p1} d\mathbf{x}_{p2} P(\mathbf{x}_{p1}) P^*(\mathbf{x}_{p2}) e^{-i2\pi(\mathbf{x}_{p1}\bullet\mathbf{x}_{i1} - \mathbf{x}_{p2}\bullet\mathbf{x}_{i2})}$$

$$\times \int_{-\infty}^{\infty} d\mathbf{x}_o I(\mathbf{x}_o) e^{i2\pi[(\mathbf{x}_{p1} - \mathbf{x}_{p2})\bullet\mathbf{x}_o]}$$

The single integral in Equation 3-16 is the Fourier transform of the source. Under the assumption that the source is uniform and the source is large, i.e., it completely fills the pupil, the single integral can be replaced by $I_0\delta(\mathbf{x}_{p1}\text{-}\mathbf{x}_{p2})$. By replacing the single integral, a Hopkins' effective source is said to be located in the exit pupil plane [54]. This Hopkins' effective source is incoherent; consequently, Equation 3-16 reduces to Equation 3-17.

Equation 3-17. $$J_i(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = I_o \int_{-\infty}^{\infty} d\mathbf{x}_p |P(\mathbf{x}_p)|^2 e^{-i2\pi(\mathbf{x}_{i1} - \mathbf{x}_{i2})\bullet\mathbf{x}_p}$$

Equation 3-17 implies that the spatial coherence of a incoherent source is equivalent to the spatial Fourier transform of the optical intensity distribution. This result is known as the van Cittert-Zernike theorem [73]. The van Cittert-Zernike theorem implies that the light gathers spatial coherence as it propagates. This fact will be used in future sections and chapters where it will be shown that a partially coherent field can be decomposed into a minimum number of coherent fields.

Equation 3-17 implies that the mutual intensity of an incoherent source with a circular pupil aperture is an Airy function, which is given in Equation 3-18 below.

Equation 3-18. $$J_i(x_1, y_1; x_2, y_2) = \int_{\sqrt{x_p + y_p} < 1} dx_p dx_p e^{-i2\pi(\Delta x_i x_p + \Delta y_p y_p)} = 2\frac{J_1(2\pi r_{12})}{2\pi r_{12}}$$

where $\Delta x_i = x_{i1} - x_{i2} = r_{12}\cos\theta_{12}$, $\Delta y_i = y_{i1} - y_{i2} = r_{12}\sin\theta_{12}$, $x_i = \rho_i\cos\phi$, $y_i = \rho_i\sin\phi$, and $r_{12} = \sqrt{\Delta x_i^2 + \Delta y_i^2} = \sqrt{\rho_{i1}^2 + \rho_{i2}^2 - 2\rho\rho\cos(\phi_{i1} - \phi_{i2})}$. It should be noted that the mutual intensity is a function of only three variables, ρ_1, ρ_2, and $\Delta\phi$. Since the mutual intensity is a function of $\Delta\phi$ rather than ϕ_1 and ϕ_2, the optical system is rotationally symmetric. The rotational symmetry of the mutual intensity implies that the mutual coherence has some order or correlation.

3.4. The Pupil Function

The image transfer function, *K*, determines the ability of the imaging system to resolve a feature. As shown by Equation 3-12, the image transfer function is the Fourier transform of the pupil function, *P*. The pupil function describes the imaging characteristics by ray tracing the optical path difference (OPD) through each point of the optical system. These characteristics include aberrations, high numerical aperture effects, defocus, pupil filtering, and the effect of imaging into a thin stack. The pupil function is most often modified to include non-idealities such as a focus error (defocus) [53,45] and aberrations. The aberrations of a system are characterized by the aberration function [10], $\Phi(\tilde{x}_p, \tilde{y}_p)$, which is given in wavelengths of error. The pupil function for an aberrated optic with defocus is shown Equation 3-19 below where circ is one when the argument is less then zero otherwise circ is zero. The effect of high numerical aperture on imaging can be included by mod-

Equation 3-19.
$$P(\mathbf{x}_p) = P(x_p, y_p) = exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p)\right] \times exp\left[-i2\pi z\frac{1}{NA^2}\sqrt{1 - NA^2(x_p^2 + y_p^2)}\right] circ\left(\sqrt{x_p^2 + y_p^2}\right)$$

ifying the pupil function with an obliquity factor [156,26]. For an optical system with of a large numerical aperture, aberrations, and defocus, the pupil function is given by

Equation 3-20. The pupil function can be generalized further to include the effect of a thin-

Equation 3-20.

$$P(x_p, y_p) = \left[\frac{1-NA^2(x_p^2+y_p^2)/M^2}{1-NA^2(x_p^2+y_p^2)}\right]^{1/4} exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p)\right] \times exp\left[-i2\pi z\frac{1}{NA^2}\sqrt{1-NA^2(x_p^2+y_p^2)}\right] circ\left(\sqrt{x_p^2+y_p^2}\right)$$

film stack by using ray tracing inside the stack to find the OPD[157]. Equation 3-21 shows the effect of the thin-film stack on the pupil function where P_{ik} is a matrix describing the OPD at a xy plane at a level z in the stack and where M_{kj} is a matrix describing the polarization rotation between the entrance pupil and the exit pupil of the optic.

Equation 3-21.

$$P(x_p, y_p) = \left[\frac{1-NA^2(x_p^2+y_p^2)/M^2}{1-NA^2(x_p^2+y_p^2)}\right]^{1/4} P_{ik}(x_p, y_p; z)\, M_{kj}(x_p, y_p) \times exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p)\right] exp\left[-i2\pi z\frac{1}{NA^2}\sqrt{1-NA^2(x_p^2+y_p^2)}\right] circ\left(\sqrt{x_p^2+y_p^2}\right)$$

3.5. Optical Lithography and Inspection Systems

In optical lithography, a mask that is either transmissive or reflective is located in the object plane. This mask is imaged into photoresist covering a wafer located in the image plane. Similarly in an optical inspection system, the light diffracted or reflected from a mask or a wafer is located in the object plane and is imaged onto a CCD camera located in the image plane.

A model for the lithography system or for the inspection system is shown in Figure 3-2. This is the same model used in microscopy. In this model, the condenser optic projects light from an incoherent source on to a mask (in lithography). The mask diffracts the light into a series of diffracted orders as shown in Figure 3-2. Typically, the projection optic is capable of collecting and imaging only the lowest diffracted orders on to a wafer.

3.6. Optical System Simulation

Simulation of the optical system requires calculating the mutual coherence as it propagates through the mask, through the lens, and onto the wafer. Topographical struc-

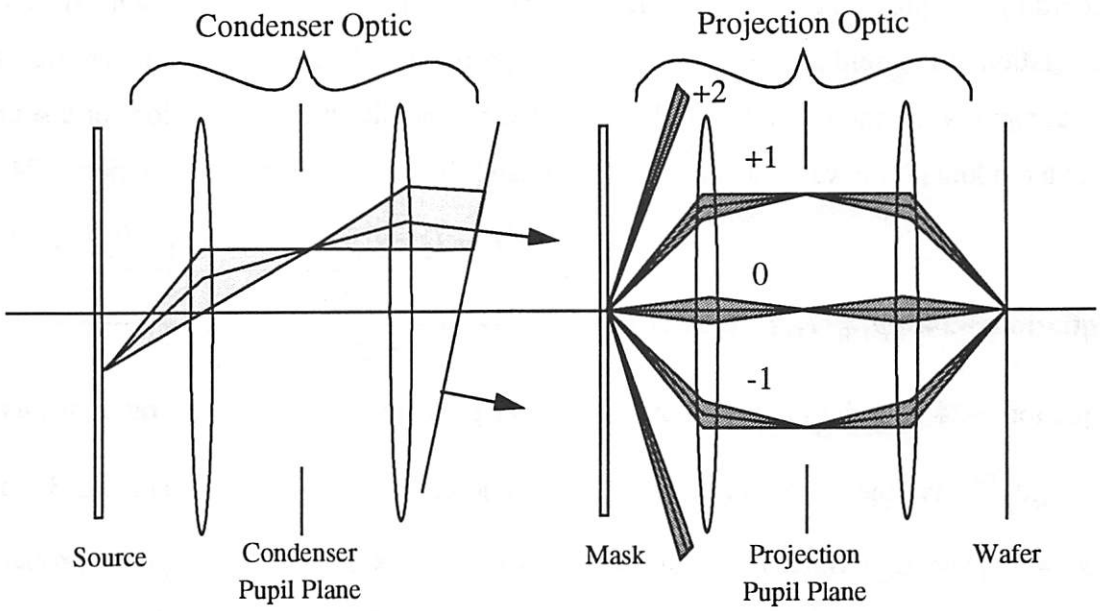


Figure 3-2. Optical Lithography and Inspection System Optic

tures on the mask and on the wafer scatter the mutual coherence. This scattering could be simulated by solving the set of coupled wave equations [13] shown in Equations 3-22 and 3-23 below. Equations 3-22 and 3-23 describe how partially coherent light interacts with matter through the coupled set of equations for the mutual intensity, *J*. These equations could be solved through an integral or differential technique described in Chapter 2. However, since the mutual coherence is a function of seven variables, x_1, y_1, z_1, x_2, y_2, z_2, and τ, the simulation would be memory and time intensive. In addition, this approach does not take advantage of the fact that the field gathers coherence as it propagates. This gathering of coherence implies that there is redundancy between the two coupled equations, since Equation 3-22 would be equivalent to Equation 3-23 for a fully coherent field.

Equation 3-22. $$\nabla_1^2 J(x_1, y_1, z_1; x_2, y_2, z_2; \tau) = \frac{1}{c^2}\frac{\partial^2}{\partial \tau^2} J(x_1, y_1, z_1; x_2, y_2, z_2; \tau)$$

Equation 3-23. $$\nabla_2^2 J(x_1, y_1, z_1; x_2, y_2, z_2; \tau) = \frac{1}{c^2}\frac{\partial^2}{\partial \tau^2} J(x_1, y_1, z_1; x_2, y_2, z_2; \tau)$$

Another possibility for calculating the partial coherence as it propagates and scatters the mask and wafer involves simulating the statistics of the light. The statistics are simulated by modulating the excited field with a random transmittance and phase that varies

according to a probability density function. Since the mutual intensity represents the auto-correlation of the field, it is possible to define a probability density function that represents the transmittance and the phase of the light. The probability density function for a source with a random phase was calculated by Goodman [40] and is repeated in Equation 3-24. In

Equation 3-24. $$p_{I,\theta}(I_1, I_2, \theta_1, \theta_2) = \frac{exp\left[-\frac{I_1 + I_2 - 2\sqrt{I_1 I_2}|\mu|\cos(\theta_1 + \theta_2 - \psi)}{2\sigma^2(1-|\mu|^2)}\right]}{16\pi^2\sigma^4(1-|\mu|^2)}$$

Equation 3-24 I_i is the random transmittance at point i=1 or 2, θ_i is the random phase, $\mu = |\mu|e^{i\psi}$ is the complex degree of coherence given by Equation 3-3, and $2\sigma^2 = \langle I_1 \rangle = \langle I_2 \rangle$. Since μ is a function of two points, (x_1, y_1) and (x_2, y_2), the probability density is a function of (x_1, y_1) and of (x_2, y_2) as well. This implies that the random transmittance and phase depends on the coupling between two points in the excitation plane. Therefore, the excitation at one point depends on the amount of coupling between all the other points in the excitation plane. Consequently, in order to describe the propagation of the mutual coherence through statistical modeling, two coupled wave equations, Equation 3-22 and Equation 3-23, must also be solved. Since solving these coupled equations is memory and time intensive, statistical modeling is also memory and time intensive.

The simulation time and memory required in propagating the mutual coherence can be reduced by decoupling wave equations 3-22 and 3-23. These two coupled partial differential equations can be decoupled by transforming the mutual intensity, *J*, into an orthogonal basis. In an orthogonal basis, the basis vectors are correlated or coherent with itself but are uncorrelated or incoherent with the other vectors. As well as orthogonality between basis vectors, it is desirable to represent the mutual intensity with least number of basis vectors in order to reduce the simulation time. The transform technique that decomposes the mutual intensity into the least number of basis vectors is said to have the optimal energy compaction. In summary, a transform technique that has the following two properties is sought:

- a sum of excitations where each excitation is coherent with itself yet incoherent or orthogonal with other excitations.
- optimal energy compaction.

3.6.1. Abbe's Formulation

In the lithography and inspection system, the source is incoherent and is imaged onto the mask by Köhler illumination [12]. In Köhler illumination rays from each discrete point on the source emerge from the condenser as a plane wave. Each of these plane waves are diffracted by the mask, collected by the projection optic, and imaged on to the wafer. An incoherent source implies that these plane wave excitations form an orthogonal basis of coherent fields. Since the excitations are orthogonal, the total intensity on the wafer is the sum of the intensities produced by each of these individual plane waves. This method is known as Abbe's formulation [1] which is shown schematically in Figure 3-3. In Figure 3-3, each of the plane wave excitations are simulated individually. The total intensity, I_{TOT}, is found by summing the intensities produced by these individual simulations. This implies that in Abbe's formulation the source is integrated over last. Since the excitations form an orthogonal basis of coherent fields, Abbe's formulation has the first desired property stated in Section 3.6.

Since Abbe's formulation has the first desired property of orthogonal coherent excitations, Abbe's formulation is next tested to determine if it has the second desired property, that of optimal energy compaction. In Abbe's formulation, the Hopkins' effective source is sampled at discrete points into a set of delta functions. The coherence of each of these individual delta functions is found by using the van Cittert-Zernike theorem. Using this theorem, the Fourier transform is taken for each delta function resulting in a plane wave excitation. Since these excitations are found by taking a Fourier transform, Abbe's formulation is a Fourier transform technique for representing the partial coherence of the source. Since each of these excitations require one simulation, it is desirable to represent the source with the least number of samples that accurately represents the mutual intensity.

The number of samples needed in the Abbe Formulation to represent the partial coherence of the source can be found through the Bragg condition. For a periodic square simulation domain with length, w, the Bragg condition states that the diffracted orders occur at discrete angles as stated in Equation 3-25. The maximum order produced by the illumination source is related to the numerical aperture of the illumination lens and is given

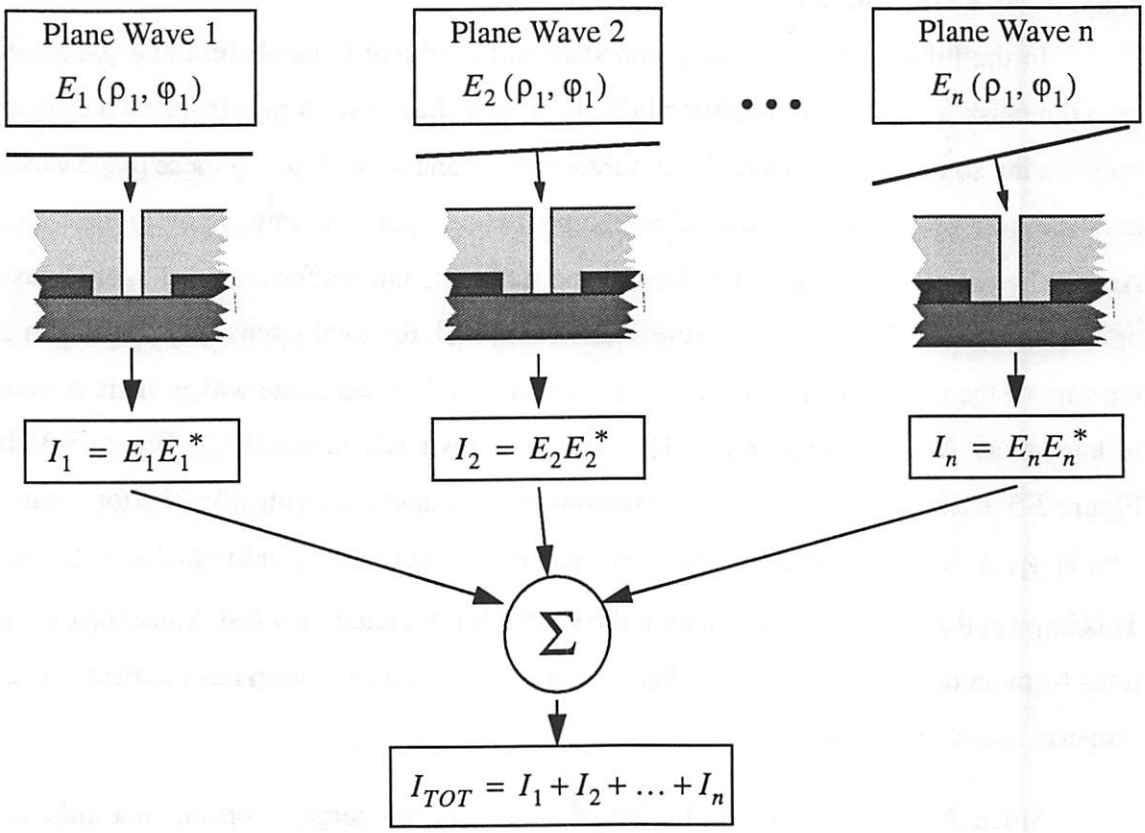


Figure 3-3. Pictorial representation of Abbe's formulation.

by $\sin\theta_{max} = \sigma NA$. This is substituted into Equation 3-25; however, since the lens collects both the maximum positive m order and the maximum negative m order, a factor of 2 must also be substituted into Equation 3-25. After making both these substitutions, Equation 3-25 becomes Equation 3-26. The number of samples needed in a one dimensional simulation domain is calculated from Equation 3-26 by solving for m and is shown in Equation 3-27. The number of samples needed in two dimensions is given in Equation 3-28, and is found by squaring Equation 3-27 and multiplying by $\frac{\pi}{4}$ because the source is circular rather than square. In Equations 3-27 and 3-28, an over sampling term N_s has been included. This over sampling term is chosen by slowly increasing N_s until the best representation of the mutual intensity is found.

Equation 3-25. $$m\lambda = w\sin\theta_m$$

Equation 3-26.

$$m\lambda = 2\sigma NAw$$

Equation 3-27.

$$n_{1D} = N_s \frac{2\sigma NAw}{\lambda}$$

Equation 3-28.

$$n_{2D} = \frac{\pi}{4}\left(N_s \frac{2\sigma NAw}{\lambda}\right)^2$$

The ability of Abbe's formulation to represent the mutual intensity is demonstrated in Figure 3-4. Figure 3-4 is a plot of the mutual intensity of a condenser optic with σ of 0.5,

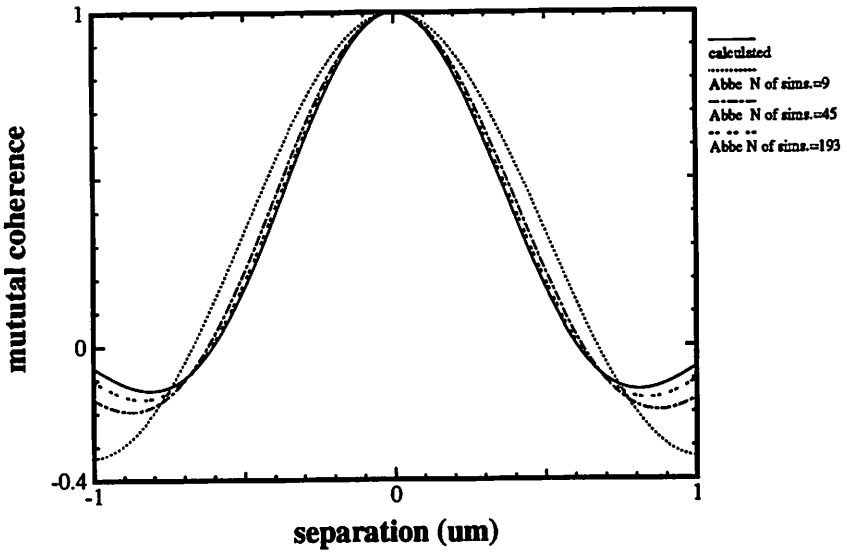


Figure 3-4. Comparison of the Mutual Intensity for Various Over Sampling Values, N_S

An incoherent circular source with NA=0.5, σ=0.6 and l=248nm was sampled and mutual intensity was imaged to the mask plane. The solid line corresponds to the analytical solution and the dashed lines to a source sampled a number of times in cartesian coordinates.

NA of 0.5 and λ of 248nm. The solid line in Figure 3-4 corresponds to the analytical mutual intensity which is an Airy function as given by Equation 3-18. The analytical mutual intensity is compared to the mutual intensity of a sampled illumination source in which N_S=1, 2, and 4. When N_S is 1, 2, and 4, the source is sampled 9, 45, and 193 times, respectively. Comparing the analytical mutual intensity of Figure 3-4 with sampled source mutual intensities (dotted lines), the source is better represented as N_S increases. When N_S is 4, the sampled mutual intensity approximates the main lobe of the analytic solution to 1% accuracy, and the first side lobe to 5% accuracy. This accuracy, however, requires many samples, 193

for N_s=4; consequently, many simulations are needed to accurately model the partial coherence.

Although Abbe's formulation has the desirable property that the simulation excitations form an orthogonal basis of coherent fields, it requires many simulations, i.e., it has poor energy compaction. It is important to note that this poor energy compaction property of Abbe's formulation can be understood physically. As a result of the van Cittert-Zernike theorem, the fields gather spatial coherence as a result of propagating. This gathering of coherence is not utilized in Abbe's formulation because the source is integrated over after the intensities produced by each of its rays has been computed. In addition, it is also known in signal and image processing that the Fourier transform technique requires many basis vectors [66]. In order to represent the mutual intensity with the least number of basis vectors, it is necessary to find an alternative method to transform the mutual intensity into the least number of orthogonal coherent fields.

3.6.2. Hopkins' Formulation

Alternatively to Abbe's formulation, the mutual intensity at the wafer can be found by calculating the mutual intensity incident onto the mask by integrating over the source first rather than after finding the intensities produced by each of the rays in the source as in Abbe's formulation. This technique is known as Hopkins' formulation [12,54]. By assuming that the incoherent source is uniform and large, the condenser optic can be replaced by a Hopkins' effective source located a focal length in front of the mask as shown in Figure 3-5. The radius of the effective source is equivalent to the numerical aperture of the condenser optic, NA_c. In optical lithography as in microscopy, it is convenient to define a partial coherence factor, σ. The partial coherent factor is the ratio of the radius of the illumination cone to that of the acceptance cone of the lens that is $\sigma=NA_c/NA$, where NA is the numerical aperture of the projection optic. Since the effective source is incoherent, the mutual intensity incident onto the object is given by the van Cittert-Zernike theorem. Through this theorem, the Fourier transform of this circular effective source with radius σNA is taken resulting in an Airy function representation for the mutual intensity as shown in Equation 3-30. The mutual intensity at the wafer plane is found through Equation 3-14

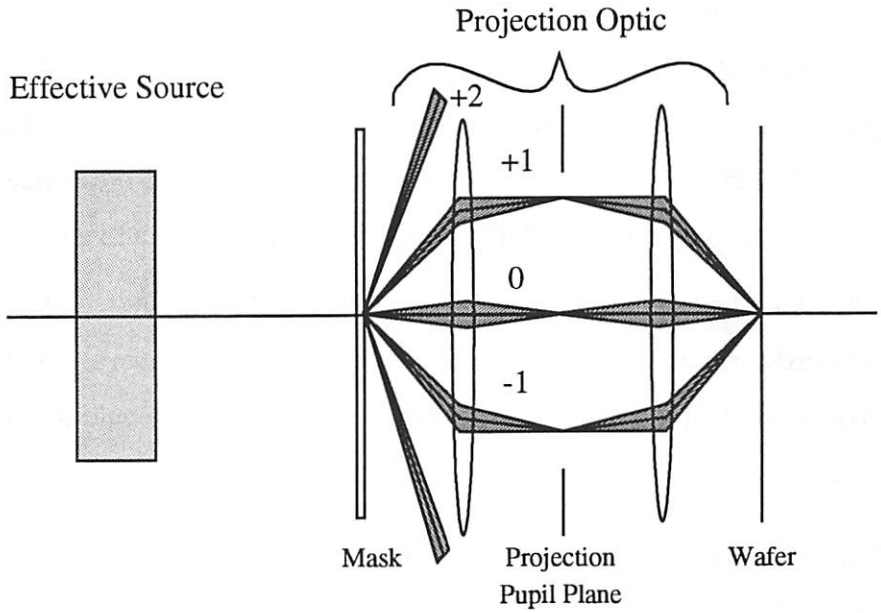


Figure 3-5. Lithography Optic with a Hopkins' Effective Source

The illumination optic can be replaced by a Hopkins' effective source with radius σNA.

of Section 3.3.2 with a slight modification in which the mutual intensity at the mask is modulated by the transmission of the mask as shown in Equation 3-29. The intensity in this

Equation 3-29.

$$J_i(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \int_{-\infty}^{\infty}\int J_o(\mathbf{x}_{o1}, \mathbf{x}_{o2})\, t_o(\mathbf{x}_{o1})\, t_o^*(\mathbf{x}_{o2})\, K(\mathbf{x}_{i1} - \mathbf{x}_{01})\, K^*(\mathbf{x}_{i2} - \mathbf{x}_{o2})\, d\mathbf{x}_{o1} d\mathbf{x}_{o2}$$

plane is given when $\mathbf{x}_{i1}$ is equal to $\mathbf{x}_{i2}$, i.e., $I(\mathbf{x}_i) = J_i(\mathbf{x}_i, \mathbf{x}_i)$. For a condenser system with a circular pupil, J_o is an Airy function as given by Equation 3-30 where $|\mathbf{x}_1 - \mathbf{x}_2|$ is given by Equation 3-31 in Cartesian coordinates and by Equation 3-32 in polar coordinates.

Equation 3-30.

$$J_o(\mathbf{x}_{o1}, \mathbf{x}_{o2}) = 2\frac{J_1(|\mathbf{x}_{o1} - \mathbf{x}_{o2}|)}{|\mathbf{x}_{o1} - \mathbf{x}_{o2}|}$$

Equation 3-31.

$$|\mathbf{x}_{o1} - \mathbf{x}_{o2}| = \frac{2\pi}{\lambda} NA\sigma \sqrt{(\tilde{x}_{o1} - \tilde{x}_{o2})^2 + (\tilde{y}_{o1} - \tilde{y}_{o2})^2}$$

Equation 3-32.

$$|\mathbf{x}_{o1} - \mathbf{x}_{o2}| = \frac{2\pi}{\lambda} NA\sigma \sqrt{\tilde{\rho}_{o1}^2 + \tilde{\rho}_{o2}^2 - 2\tilde{\rho}_{o1}\tilde{\rho}_{o2}\cos(\phi_1 - \phi_2)}$$

This formulation is an efficient method for calculating the mutual intensity in the image plane, $J_i(\mathbf{x}_{i1},\mathbf{x}_{i2})$. Due to this efficiency, Hopkins' formulation is widely used in the lithography community to calculate the scalar image of a mask produced by a lithography system. Although Hopkins' formulation is widely used to calculate the scalar image, the formulation calculates only the mutual intensity, $J_i(\mathbf{x}_{i1},\mathbf{x}_{i2})$ which is a function of two coupled spatial points, $\mathbf{x}_{i1}$ and $\mathbf{x}_{i2}$. Since a coherent excitation is a function of only one spatial variable, Hopkins' formulation does not produce coherent excitations, and simulating the propagation of the mutual intensity as it interacts with topography would be computationally expensive.

Although Hopkins' formulation does not produce coherent excitations, the formulation does take advantage of the fact that the light has gathered coherence through its propagation from the incoherent source to the object plane or to the image plane. This gathering of coherence implies that the light is more correlated and that there is some redundancy of information. Since there is redundancy, it is possible to find a more efficient representation for this light. This representation is described in the next section, Section 3.7.

3.7. Decomposition of the Scalar Mutual Intensity

A more efficient representation for the mutual intensity is developed by using techniques from communication theory. By noting that the mutual intensity is the correlation matrix of the light at a particular imaging plane, the Karhunen-Loeve transform[41] from communication theory can be used to decompose this correlation matrix. The Karhunen-Loeve transform[41] has been proven to have the optimal energy compaction, i.e., the least number of mutually uncorrelated signals are needed to represent a partially correlated signal. The Fourier transform used in Abbe's formulation is not optimal in this sense. The Karhunen-Loeve transform decomposes the correlation matrix into a set of eigenvalues and eigenfunctions. These eigenfunctions represent coherent excitations which are by definition orthogonal or incoherent with the other excitations.

The Karhunen-Loeve transformation technique has been applied previously to optics to remove the correlation in the mutual intensity. Gamo[38] showed that the mutual intensity can be decomposed into a set of orthogonal eigenvalues and eigenfunctions. With

this technique, the mutual intensity is represented by an incoherent sum of coherent electric fields. This decomposition technique has also been successfully applied to optical proximity correction[97,24], to depth of focus enhancement[15], and to topography scattering with partially coherent illumination[115].

The mathematical detail of using the Karhunen-Loeve transform to decompose the mutual intensity was developed by von Bünau[17] and is repeated below. The Karhunen-Loeve transform diagonalizes the correlation matrix (in this case the mutual intensity, J, at either the object plane or the image plane) into a set a of eigenvalues and eigenfunctions. This is represented mathematically by Equation 3-33.

Equation 3-33.
$$\int_{-\infty}^{\infty} J(\mathbf{x}_{i1}, \mathbf{x}_{i2})\, \Phi_k(\mathbf{x}_{i2})\, d\mathbf{x}_{i2} = \lambda_k \Phi_k(\mathbf{x}_{i1})$$

with the diagonalization properties of Equations 3-34 and 3-35.

Equation 3-34.
$$\int_{-\infty}^{\infty} \Phi_k(\mathbf{x})\, \Phi_l^*(\mathbf{x})\, d\mathbf{x} = \delta_{kl}$$

Equation 3-35.
$$J(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \sum_{k=1}^{\infty} \lambda_k \Phi_k(\mathbf{x}_{i1})\, \Phi_k^*(\mathbf{x}_{i2})$$

The diagonalization of Equation 3-33 can be simplified by recognizing that the mutual intensity, J, depends on ϕ_1-ϕ_2 as given by Equations 3-30 and 3-32. This implies that the mutual intensity is rotationally symmetric and the diagonalization can then be written as:

Equation 3-36.
$$\int_0^{\infty} \int_0^{2\pi} J(\rho_1, \rho_2, \phi_1 - \phi_2)\, \Phi_k(\rho_2, \phi_2)\, d\phi_2 \rho_2 d\rho_2 = \lambda_k \Phi_k(\rho_1, \phi_1)$$

Since the mutual intensity or the kernel of the transform depends upon the different angle $\delta\phi=\phi_1-\phi_2$, ϕ_1 can be replaced with $\phi_1+\delta\phi$ and ϕ_2 with $\phi_2+\delta\phi$ as shown in Equation 3-37.

Equation 3-37.
$$\int_0^{\infty} \int_0^{2\pi} J(\rho_1, \rho_2, \phi_1 - \phi_2)\, \Phi_k(\rho_2, \phi_2 + \delta\phi)\, d\phi_2 \rho_2 d\rho_2 = \lambda_k \Phi_k(\rho_1, \phi_1 + \delta\phi)$$

By subtracting Equation 3-37 from Equation 3-36, dividing by $\delta\phi$ and taking the limit

$\delta\phi \rightarrow 0$, Equation 3-38 results. Equation 3-38 implies that the angular derivatives are also eigenfunctions.

Equation 3-38. $$\int_0^{\infty}\int_0^{2\pi} J(\rho_1, \rho_2, \phi_1 - \phi_2)\frac{\partial}{\partial\phi_2}\Phi_k(\rho_2, \phi_2)\, d\phi_2 \rho_2 d\rho_2 = \lambda_k \frac{\partial}{\partial\phi_1}\Phi_k(\rho_1, \phi_1)$$

The eigenfunctions given by Equation 3-38 are chosen such that the operator $\frac{\partial}{\partial\phi}$ is diagonal to the basis formed by the eigenfunctions given by Equation 3-36. This implies that:

Equation 3-39. $$\Phi_k(\rho, \phi) \propto \frac{\partial}{\partial\phi}\Phi_k(\rho, \phi)$$

and requiring that Φ_k is continuous in ϕ.

Equation 3-40. $$\Phi_k(\rho, \phi + 2\pi) = \Phi_k(\rho, \phi)$$

Therefore, Equations 3-39 and 3-40 imply that the eigenfunctions must be products of the form given in Equation 3-41 where the index n specifies the radial order of the eigenfunc-

Equation 3-41. $$\Phi_{nm}(\rho, \phi) = \Phi_{nm}(\rho)\, e^{im\phi} \qquad m = 0, \pm 1, \pm 2, \ldots$$

tions and the index m specifies the angular order. After substitution of Equation 3-41 into Equation 3-36 and diving through by $e^{im\phi_1}$, Equation 3-36 becomes:

Equation 3-42. $$\int_0^{\infty} \bar{J}(\rho_1, \rho_2)\, \Phi_{nm}(\rho_2)\, \rho_2 d\rho_2 = \lambda_{nm}\Phi_{nm}(\rho_1)$$

where $\bar{J}(\rho_1, \rho_2)$ is given by Equation 3-43. Equation 3-42 implies that $\Phi_{nm}(\rho)$ are eigen-

Equation 3-43. $$\bar{J}(\rho_1, \rho_2) = \int_0^{2\pi} J(\rho_1, \rho_2, \Delta\phi)\, e^{im\phi} d\Delta\phi$$

functions of $\bar{J}(\rho_1, \rho_2)$.

Since the diagonalization of the correlation matrix into its eigenfunctions implies that the basis is orthogonal, the eigenfunctions, $\Phi_{nm}(\rho,\phi)$, are incoherent with the other eigenfunctions yet each eigenfunction is coherent with itself. These eigenfunctions essentially represent the electric field contributions to the mutual intensity. Since each eigen-

function field is coherent with itself, the field can be propagated to another plane by using the scalar coherent theory developed in Section 3.3.1. After propagation, the intensity at this other plane is found by summing the squared contributions from the eigenfunctions weighted by the eigenvalues as shown in Equation 3-44.

Equation 3-44.

$$I(\mathbf{x}) = \sum_{n=1}^{\infty} \sum_{m=0}^{\infty} \lambda_{nm} |\Phi_{nm}(\mathbf{x})|^2$$

This approach is the same technique proposed by Gamo in 1964 as an alternative method to Hopkins' formulation. In this approach, Gamo found the mutual intensity at the object plane by using the van Cittert-Zernike theorem, and he decomposed this mutual intensity into a set of eigenfunctions. Gamo then used these eigenfunctions to calculate the intensity of a grating in an image plane. In this calculation, each of these eigenfunctions was propagated through the grating to the image plane by using the scalar diffraction theory of Section 3.3.1. Through Equation 3-44, the intensity in the image plane is calculated by squaring and summing over the diffracted eigenfunctions as weighted by the eigenvalues. Like Hopkins' formulation, the total intensity from this decomposition technique models the effect of using partially coherent illumination in scalar imaging. This decomposition technique proposed by Gamo was later re-invented by Wolf[150] and used by Saleh[106] to calculate the image of a grating.

3.7.1. Decomposition of the Mutual Intensity Block Matrix

In Section 3.7, the mutual intensity, which depends on four spatial variables, (ρ_1,ϕ_1) and (ρ_2,ϕ_2), was decomposed into a set of eigenfunctions by first removing the angular, ϕ, dependence. It is also possible to find the eigenfunctions without removing the angular dependence. This is accomplished by transforming the four dimensional mutual intensity into a two dimensional mutual intensity by using block matrices[25].

Since the source is circular, the decomposition is performed in polar coordinates where the location of point $\mathbf{x}_1$ and of point $\mathbf{x}_2$ is represented by the coordinates, (ρ_1,ϕ_1) and (ρ_2,ϕ_2), respectively. Representing these polar coordinates, ρ_1, ϕ_1, ρ_2, and ϕ_2, with i, j, k, l, respectively, the block notation of the four dimensional mutual intensity, $J(i,j,k,l)$, can be

written in two dimensions as shown in Equation 3-45. This block matrix is then diagonal-

Equation 3-45.

$$J(i,j,k,l) = \begin{bmatrix} J(1,1,1,1) & J(1,1,2,1) & \cdots & J(1,1,N,1) & J(1,1,1,2) & J(1,1,2,2) & \cdots & J(1,1,N,N) \\ J(2,1,1,1) & & & & & & & \\ \vdots & & & & & & & \\ J(N,1,1,1) & & & & & & & \\ J(1,2,1,1) & & & & & & & \\ J(2,2,1,1) & & & & & & & \\ \vdots & & & & & & & \\ J(N,2,1,1) & & & & & & & \\ \vdots & & & & & & & \\ J(N,N,1,1) & & \cdots & & & & & J(N,N,N,N) \end{bmatrix}$$

ized. The eigenvector, $\Phi_n(\mathbf{x}_1)$, of the diagonalization is represented by a column vector shown in Equation 3-46, and the eigenvector, $\Phi_n(\mathbf{x}_2)$, is represented by a row vector shown in Equation 3-47. These eigenvectors like those in Section 3.7 represent the electric field

Equation 3-46.

$$\Phi_n(\mathbf{x}_1) = \begin{bmatrix} \Phi_n(1,1) \\ \Phi_n(2,1) \\ \vdots \\ \Phi_n(N,1) \\ \Phi_n(1,2) \\ \Phi_n(2,2) \\ \vdots \\ \Phi_n(N,2) \\ \vdots \\ \Phi_n(N,N) \end{bmatrix}$$

Equation 3-47.

$$\Phi_n(\mathbf{x}_2) = \left[\Phi_n(1,1)\ \Phi_n(2,1)\ \cdots\ \Phi_n(N,1)\ \Phi_n(1,2)\ \Phi_n(2,2)\ \cdots\ \Phi_n(N,N)\right]$$

excitations where each excitation is coherent with itself yet incoherent or orthogonal with other excitations. The eigenvector and eigenvalue solution when the block matrix is diagonalized is equivalent to the solution of the previous section where the rotational ϕ dependence is removed. However, when the ϕ dependence is removed, a matrix, $\bar{J}(\rho_1, \rho_2)$, that

is two-dimensional must be diagonalized which is computationally faster to diagonalize than a four dimensional block matrix.

3.8. Energy Compaction and Accuracy of the Decomposition

In this section, the decomposition method proposed in Section 3.7 is evaluated for its ability to accurately represent the mutual intensity with the least number of coherent excitations. The mutual intensity, J, at the mask plane is given by the van Cittert-Zernike theorem as shown in Equation 3-30 of Section 3.6.2. Using Equations 3-41 and 3-42, the eigenvalues and eigenfunctions of this mutual intensity were found for two optical systems, optical system A with λ=365nm, NA=0.9, and σ=1 and optical system B with λ=248nm, NA=0.5, σ=0.5.

The eigenvalues and eigenfunctions of optical systems A and B are shown in Figures 3-6a, 3-6b, 3-6c, and 3-6d, respectively. The eigenvalues in Figures 3-6a and 3-6b are plotted in descending order. In both optical systems, the mutual intensity matrix is hermitian which implies that the eigenvalues are real. Furthermore, since the eigenvalues are non-negative, the mutual intensity matrix is non-negative. Physically, a non-negative mutual intensity matrix implies that the intensity is non-negative as well. The first five non-degenerate eigenvalues multiplied by the square root of its corresponding eigenvalue are plotted in Figures 3-6c and 3-6d. Figures 3-6c and 3-6d show that the lower order eigenvectors represent the center of the domain and the higher order eigenvectors represent the edges of the domain. A closed form solution for these eigenfunctions has been found by Tejnil[125] in which the eigenfunctions are linear combinations of the Lommel functions[9].

Since the eigenfunctions form an orthogonal basis set, each eigenvalue and eigenfunction pair requires one simulation. Since the eigenvalues in Figures 3-6a and 3-6b decay with a 1/n relationship, the summation that represents the mutual intensity in Equation 3-35 can be truncated once an error criteria is met. The error criteria consists of calculating the mutual intensity with the decomposition method by summing the eigenfunctions through Equation 3-35 and comparing this mutual intensity to the analytical mutual intensity at the center and at the edge of the simulation domain. If the error between the analyt-

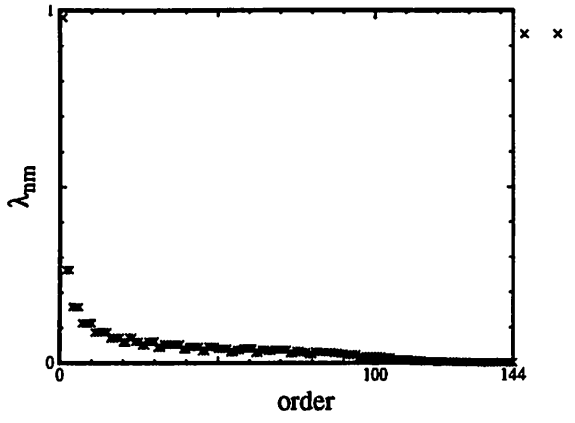


Figure 3-6a. Eigenvalues of an optical system with λ=365nm, NA=0.9, and σ=1.0.

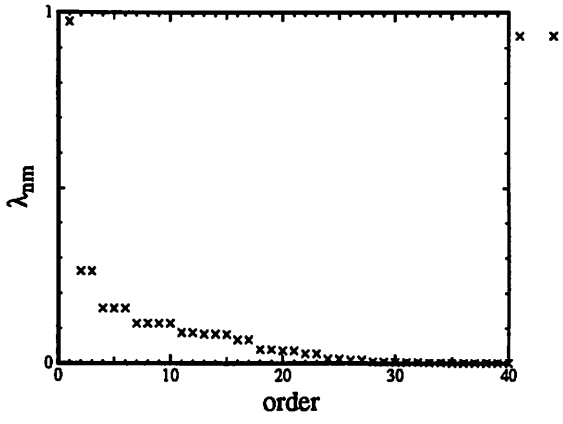


Figure 3-6b. Eigenvalues of an optical system with λ=248nm, NA=0.5, and σ=0.5.

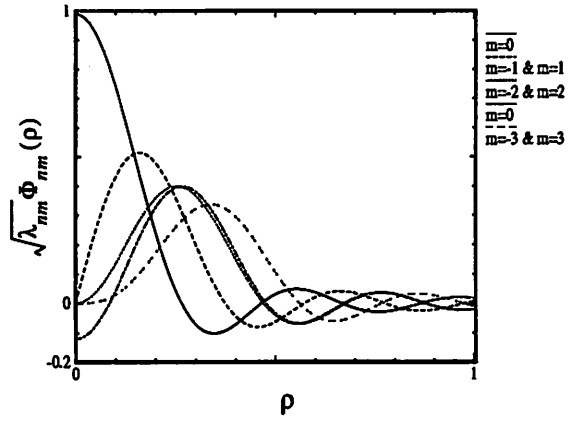


Figure 3-6c. First five non-degenerate eigenvectors of an optical system with λ=365nm, NA=0.9, σ=1.0.

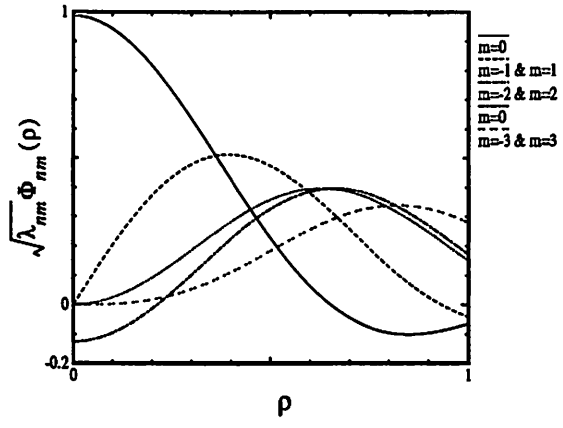


Figure 3-6d. First five non-degenerate eigenvectors of an optical system with λ=248nm, NA=0.5, σ=0.5.

Figure 3-6. Plots of the eigenvalues and eigenfunctions calculated with the decomposition formulation.

ical solution at the edge of domain is less than a certain percentage, the summation is truncated. The number of eigenfunctions or simulations needed before the truncation is directly proportional to the σNA of the system, i.e., as σNA increases, the number of simulations increases. This is demonstrated in Figure 3-7 where the mutual coherence of the eigensolution between points x_1 and x_2 for optical systems A and B is compared as a function of x_1 at x_2=0.0μm, 0.5μm, and 1.0μm over a 2x2μm simulation domain. In Figure 3-7 the mutual coherence is truncated after 82 and after 19 eigenfunctions in optical systems A and B, respectively. Figure 3-7 shows that the mutual coherence from the decomposition

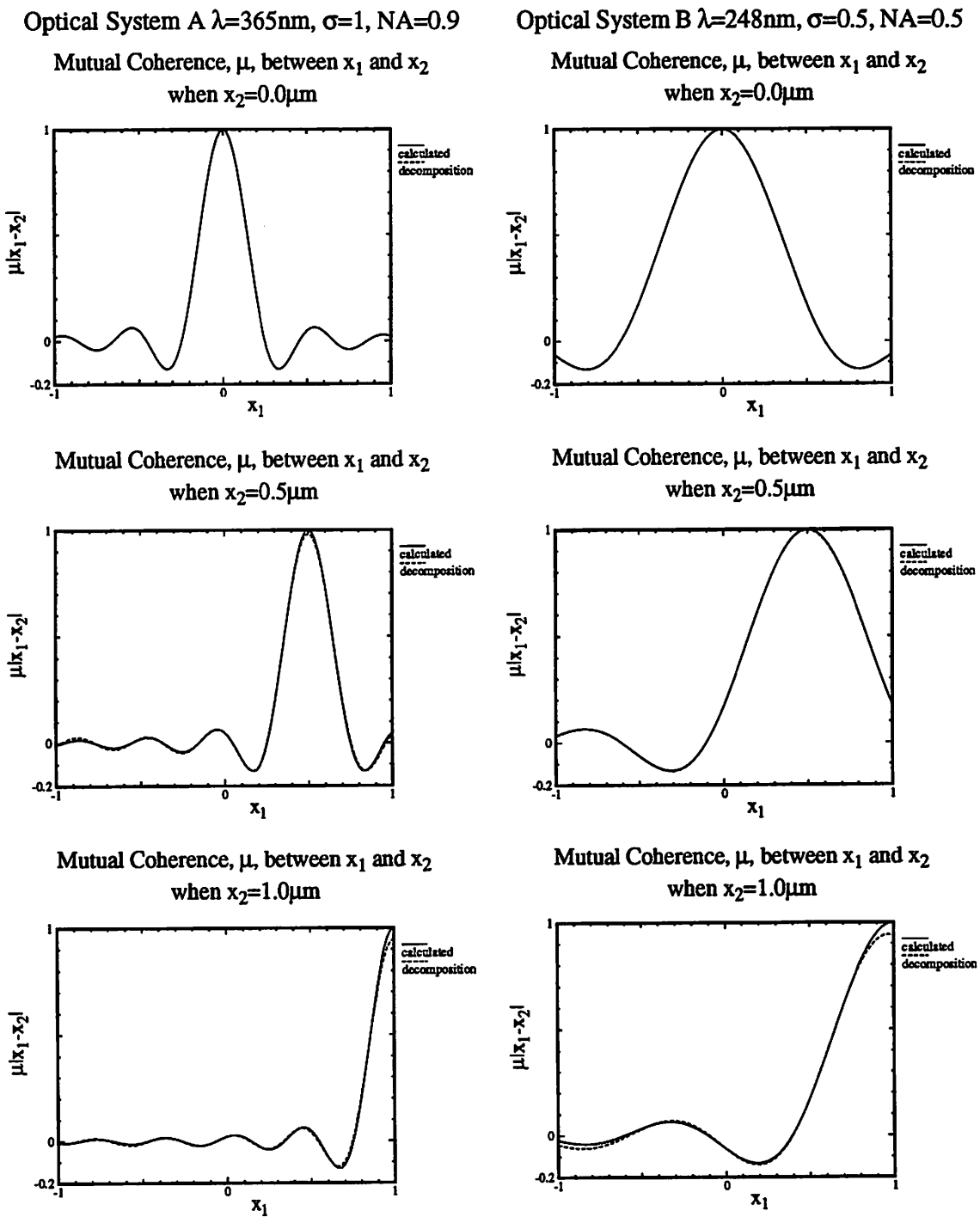


Figure 3-7. Comparison of the mutual coherence calculated with the decomposition formulation and with analytical result.

The mutual coherence was calculated for optical systems A (left) and B (right) at the center of the domain (top), at three quarters of the domain (middle) and at the edge of the domain (bottom). Figure 3-7 shows that throughout the domain, the decomposition formulation is in close agreement with the analytical mutual coherence.

is in excellent agreement with the analytical mutual coherence at the center (x_2=0.0μm) and half-way to the edge of the simulation domain (x_2=0.5μm). The peak of the main lobe does show about 4% discrepancy between the analytical solution and the decomposition representation at the edge of the domain (x_2=1.0μm).

The energy compaction of the decomposition representation is compared to Abbe's formulation in Figure 3-8a and in Figure 3-8b for optical systems A and B, respectively. In these figures, the analytical mutual coherence is compared to the mutual coherence generated with the decomposition technique and with Abbe's formulation at the center of the simulation domain (x_2=0.0μm) as a function of x_1. The mutual coherence generated by Abbe's formulation is shown when the illumination source is sampled with N_s=1, 2, and 4. In both Figure 3-8a and Figure 3-8b, the plot of the mutual coherence from the decom-

Optical System A λ=365nm, σ=1, NA=0.9

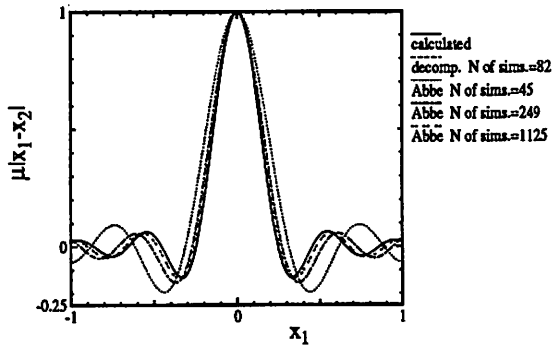


Figure 3-8a. Comparison of the mutual coherence for optical system A (λ=365nm NA=0.9 σ=1.0).

Optical System B λ=248nm, σ=0.5, NA=0.5

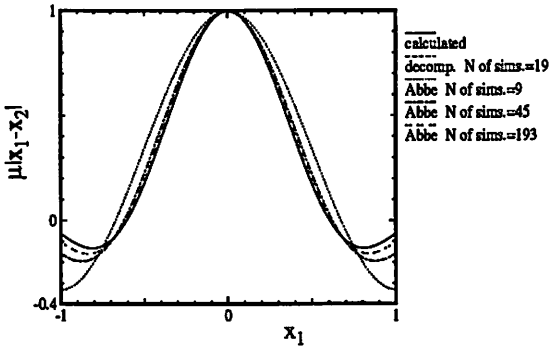


Figure 3-8b. Comparison of the mutual coherence for optical system B (λ=248nm NA=0.5 σ=0.5).

Figure 3-8. Comparison of the mutual coherence.

The mutual coherence calculated analytically with Equation 3-30 to the mutual coherence calculated with the decomposition formulation and with Abbe's formulation when N_s is 1, 2, and 4. The comparison is shown for optical systems A (left) and B (right).

position technique is the same as the analytical mutual coherence; consequently, there is only one solid line representing both the analytical result and the decomposition result. In optical system A the decomposition technique is capable of accurately representing the mutual coherence with 82 simulations while Abbe's formulation would require over 1125

simulations to represent the mutual coherence with slightly less accuracy. Also, in optical system B, the decomposition requires 19 simulations while Abbe's needs over 193 simulations to produce the mutual coherence.

In the previous paragraphs, the accuracy of the decomposition method and Abbe's method were determined by qualitatively comparing the mutual coherence calculated with both techniques to the analytical mutual coherence. The accuracy can also be quantified by calculating the root mean square error. The root mean square error is defined by Equation 3-48 where μ_a is the analytical mutual coherence and where μ_s is the mutual coherence from the decomposition technique or from Abbe's formulation. In Table 3-1, the root mean square error between the analytical mutual coherence and the decomposition mutual coherence is calculated and compared to the root mean square calculated for Abbe's method. From Table 3-1, the decomposition representation is most accurate in the center of the simulation domain, and the accuracy with the decomposition representation at the edge of the domain increases as the number of eigenfunctions increases. The decomposition is accurate in the center of the domain because as shown in Figure 3-6 the lower order eigenfunctions are largest in the center and the higher order eigenfunctions are largest at the center of the domain. Consequently, as the number of eigenfunctions increases the edge of the domain is better represented. Also from Table 3-1, the root mean square error at the center and at the edge of the simulation domain is less when using the decomposition technique than when using Abbe's formulation. The plots in Figure 3-7 and the error in Table 3-1 indicate that fewer simulations are needed with the decomposition technique than with Abbe's formulation. This implies that the decomposition technique has better energy compaction than Abbe's formulation.

Equation 3-48.
$$RMSE = \left(\sqrt{\sum_i \sum_j \left[\mu_s(i,j) - \mu_a(i,j)\right]^2}\right) / (N_i N_j)$$

Table 3-1. Root mean square error between the analytical mutual coherence versus those calculated with the decomposition formulation and versus those calculated with Abbe's formulation.

Table 3-1 shows comparison results for optical system A (top) and B (bottom) at the center of the domain, at three quarters of the domain, and at the edge of the domain. The coherency elements are calculated with Abbe's formulation when N_s is 1, 2, and 4.

Optical System A (λ=365nm NA=0.9 σ=1.0)			
Mutual Coherence Representation	Center of Domain (0.0μm)	Three quarters of Domain (0.5μm)	Edge of Domain (1.0μm)
Decomposition technique with 82 samples	0.01%	0.08%	0.12%
Abbe's formulation with 45 samples	1.19%	1.19%	1.19%
Abbe's formulation with 249 samples	0.45%	0.45%	0.45%
Abbe's formulation with 1125 samples	0.18%	0.18%	0.18%
Optical System B (λ=248nm NA=0.5 σ=0.5)			
Decomposition technique with 19 samples	0.01%	0.02%	0.15%
Abbe's formulation with 9 samples	1.14%	1.14%	1.14%
Abbe's formulation with 45 samples	0.41%	0.41%	0.41%
Abbe's formulation with 193 samples	0.18%	0.18%	0.18%

In conclusion, the decomposition technique has both of the desired properties stated in Section 3.6: it is capable of representing the mutual intensity with a sum of excitations, where each excitation is coherent with itself yet incoherent or orthogonal with other excitations; and the decomposition technique has superior energy compaction as compared to Abbe's formulation.

The eigenfunctions and eigenvalues in Table 3-1 where calculated by discretizing the spatial variables, (ρ_1,ϕ_1) and (ρ_2,ϕ_2), and by calculating the eigenfunctions and eigenvalues of the block matrix as discussed in Section 3.7.1. In Table 3-1, all four spatial variables, ρ_1, ρ_2, ϕ_1, and ϕ_2, were sampled with 50 points. The radial variables, ρ, extended from 0 to 1μm and the angular variables, ϕ, from 0 to 2π. The number of sample points were then increased to 100 points. The eigenvalues with 50 sample points and 100 sample points were then compared. In both optical systems, A and B, the eigenvalues with 50 points differed by less than 1% than the eigenvalues with 100 points. This indicates that the spatial variables need only be sampled with 50 points and the eigenfunctions can be interpolated to 100 points if needed. Sampling with the least number of variables is desired

because the computational time to decompose the matrix is proportional to the number of sample points. This independence of the eigenvalues on the number of samples was first stated by von Bünau and applied in this section to test its validity [17].

3.9. Circular Simulation Domain versus Rectangular Simulation Domain

In the previous sections, since the illumination source is circular, the mutual intensity is analyzed in polar coordinates rather than in cartesian coordinates. Consequently when the mutual intensity is decomposed into its eigensolution, the field excitations are in polar coordinates. Since simulation programs such as TEMPEST and SPLAT use rectangular periodic simulation domains, the use of polar coordinates can pose some aliasing problems. Since many of the eigensolutions have odd rotational symmetry, the rectangular periodic domain introduces high frequencies due to the discontinuity at the edge of the simulation domain. This is demonstrated in Figure 3-9. In Figure 3-9 the odd signal is forced

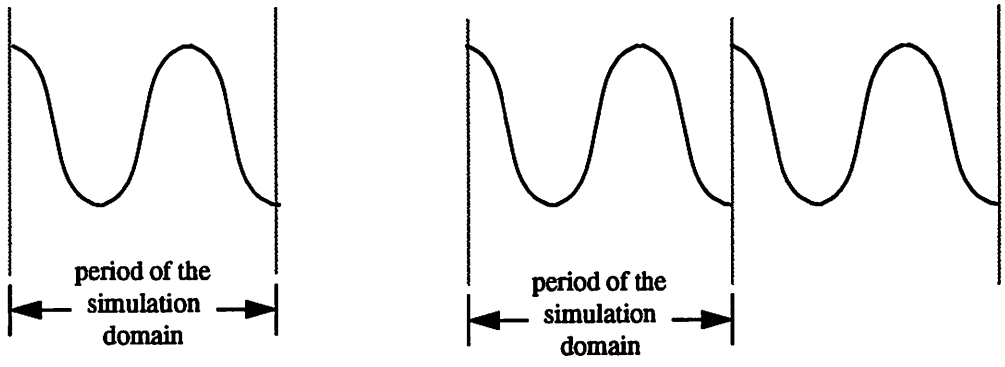


Figure 3-9. Schematic showing the aliasing resulting from simulating a sine wave excitation in a periodic domain.

to be periodic resulting in a discontinuity at the boundary which introduces nonphysical high frequency excitation components into the simulation. In order to completely cover a periodic rectangular simulation domain with a circular excitation, the circular excitation must extend into the neighboring simulation domain as is shown in Figure 3-10. In

Figure 3-10, the circular excitations overlap in the shaded regions resulting in aliasing when the decomposition method is used.

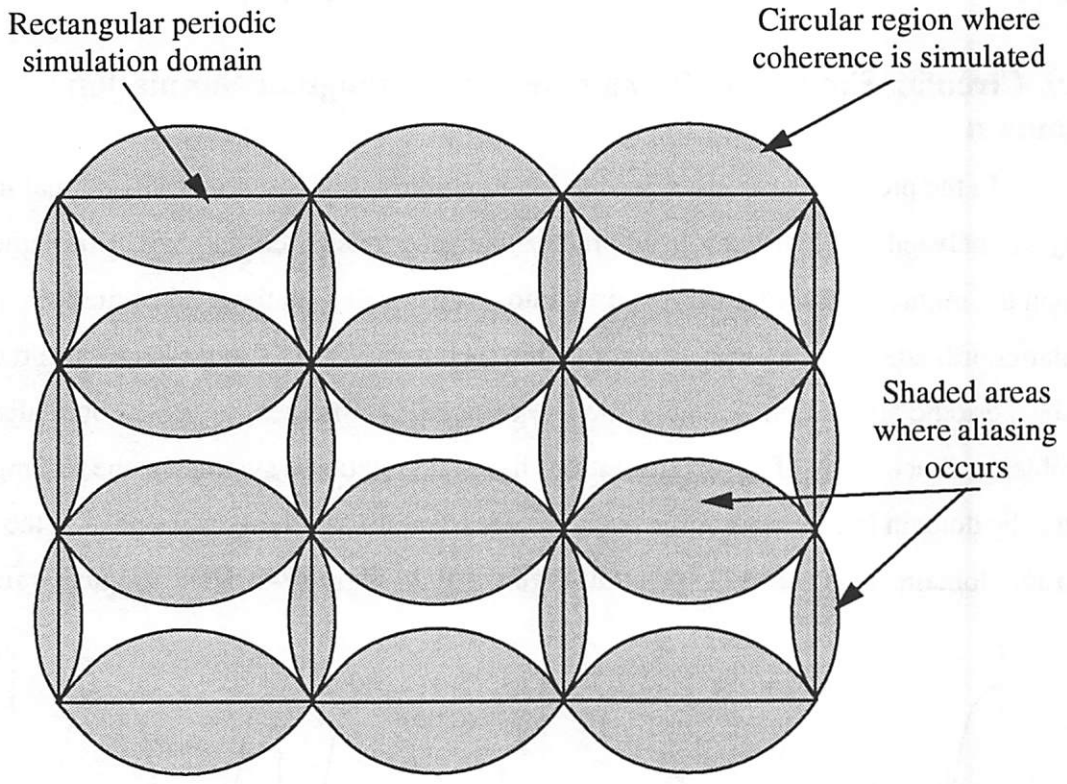


Figure 3-10. Schematic showing aliasing resulting from using an excitation calculated in radial coordinates when the simulation domain is rectangular.

Although the polar coordinate treatment of the decomposition may introduce aliasing, lithography simulation most often is concerned about simulation of isolated features. For example, patterning a gate over an active area well or inspecting a contact hole are isolated features. Of course in these cases, it is desirable to have a simulation tool that simulates scattering from truly isolated features. However, currently TEMPEST and other 3D scattering simulators assume a periodic simulation domain. Another alternative does exist when using a periodic simulation tool. Since the intensity inside the active area well or the intensity scattering from the contact hole are of more concern than the intensity far from the well or hole, it is possible to simulate the coherence using the decomposition technique by using a circular excitation that surrounds the feature of concern. This is demonstrated

in Figure 3-11 where the rectangular periodic simulation domain completely encloses the circular source region. The field excitations outside the circle are allowed to decay to zero, resulting in no aliasing at the boundary. Since the simulation domain must be increased in this method, it may seem disadvantageous because a larger simulation domain requires more simulation time and memory. However, since an isolated boundary condition does not currently exist, the simulation domain must also be increased in order to model an isolated feature in a periodic simulation domain. When modeling this isolated feature, an isolation buffer is placed around the feature. Since the field is not important inside this buffer, the coherence is only simulated in a circle that contains the isolated feature and not the buffer.

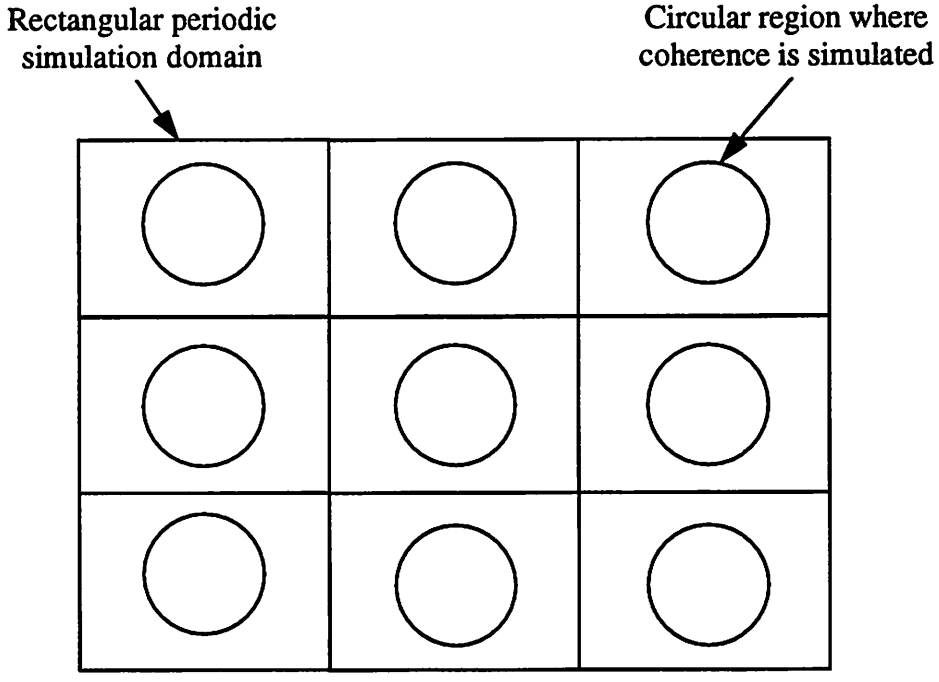


Figure 3-11. Schematic showing an excitation calculated in radial coordinates in which the feature simulated is contained within an excitation circle and fields outside this circle are not important.

Although the decomposition technique offers improved energy compaction and fewer simulations for an isolated feature, in cases where a truly periodic feature is to be simulated, such as diffraction grating, the decomposition technique may still be used. For a periodic domain, each of the field excitations produced by the decomposition technique

are transformed to k space through the Fourier transform. The k space representation is then sampled with period $2\pi/w_x$ and $2\pi/w_y$ where w_x and w_y is the period of the simulation domain in the x and y directions, respectively. The effectiveness of using the decomposition technique over a periodic domain is measured by comparing the mutual coherence with the decomposition technique to Abbe's formulation and to the analytical solution. These are shown in Figure 3-12 for optical system A (λ=365nm, NA=0.9, σ=1.0) and B

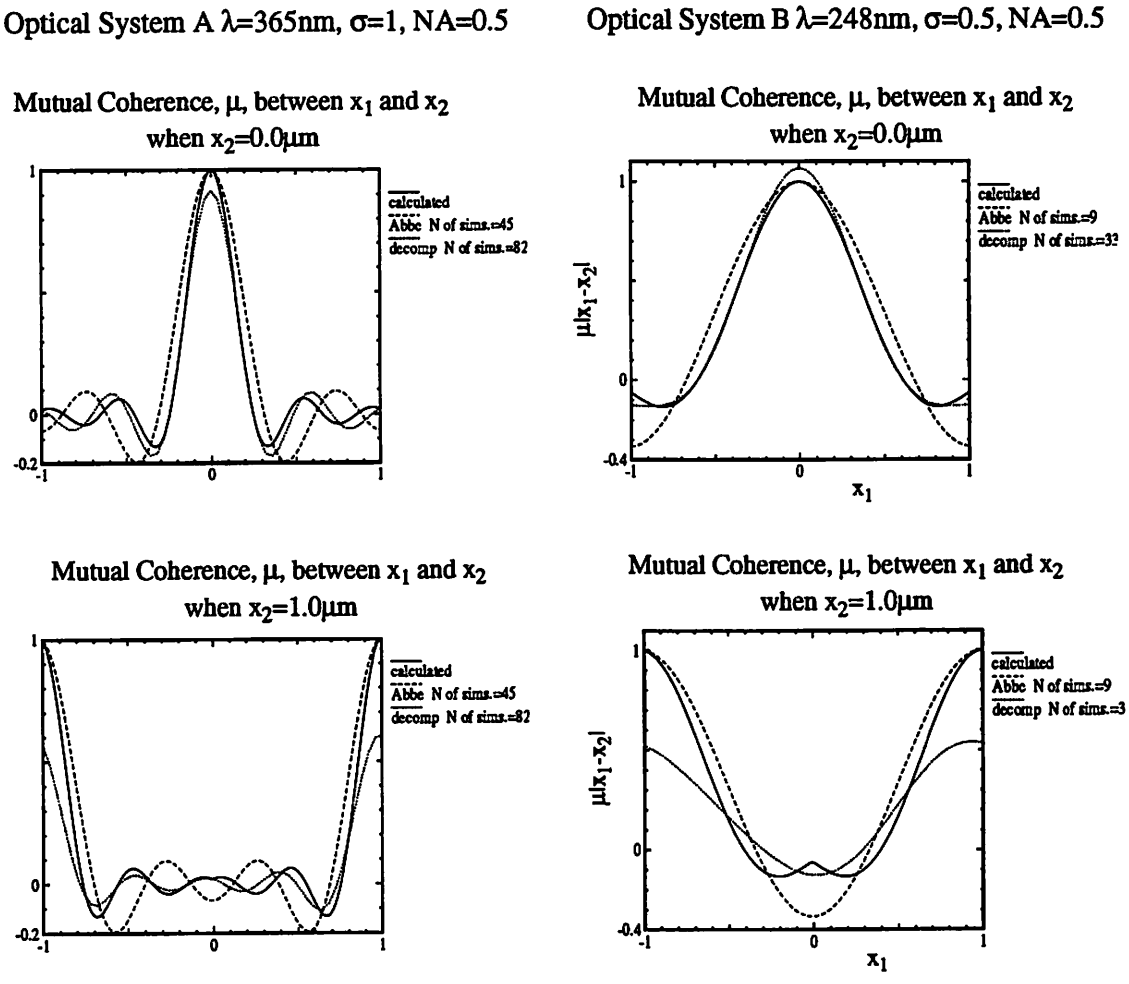


Figure 3-12. Comparison of the mutual coherence when the coherence calculated with the decomposition formulation is forced to be periodic.

The comparison of the mutual coherence between points x_1 and x_2 is shown for optical systems A (left) and B (right) at the center (top) and at the edge (bottom) of the simulation domain.

(λ=248nm, NA=0.5, σ=0.5) for locations at the edge and at the center of the domain. The

mutual coherence using the decomposition technique in Figure 3-12 was generated with 19 and 82 excitations for optical systems, A and B, respectively and is compared to the mutual coherence using Abbe's formulation with 45 and 9 excitations for systems, A and B. The root mean square error from using both these techniques is given in Table 3-2. As seen in Table 3-2 and in Figure 3-12, the decomposition technique more accurately models the analytical mutual coherence in the center of the domain as compared to Abbe's method. However, at the edge of the simulation domain the mutual coherence calculated with the decomposition method significantly differs from the analytical mutual coherence, while the mutual coherence calculated with Abbe's method is in better agreement. Furthermore, Abbe's method has better correlation with the analytical mutual coherence throughout the simulation domain in comparison to the decomposition method. Therefore, Abbe's technique is best suited for simulating truly periodic structures such as a line-space array on a phase shift mask, and the decomposition technique is best suited for simulating isolated features such as a contact hole in a mask or scattering from wafer topography.

Table 3-2. Root mean square error between the analytical mutual coherence versus those calculated with Abbe's formulation and versus those calculated with the decomposition formulation over a periodic rectangular domain.

Table 3-2 shows comparison results for optical system A (top) and B (bottom) at the center of the domain and at the edge of the domain. The coherency elements are calculated with Abbe's formulation when N_s is 1

Mutual Coherence Representation	Center of Domain ($0.0\mu m$)	Edge of Domain ($1.0\mu m$)
Optical System A (λ=365nm NA=0.9 σ=1.0)		
Sampled decomposition technique with 82 samples	0.56%	1.75%
Abbe's formulation with 45 samples	1.19%	1.19%
Optical System B (λ=248nm NA=0.5 σ=0.5)		
Sampled decomposition technique with 19 samples	0.44%	3.10%
Abbe's formulation with 9 samples	1.14%	1.14%

3.10. The Summary

In this chapter, the imaging of scalar partially coherent light was presented in order to form a theoretical foundation for future chapters. Two techniques, Abbe's formulation and a scalar decomposition technique, were presented in order to simulate imaging with partially coherent illumination. Both these techniques modeled the partial coherence

through a summation of coherent excitations where each excitation is incoherent with the other excitations. Although both techniques involve a summation of coherent excitations, Abbe's technique requires more excitations than the decomposition technique in order to represent the partial coherence of the system. For example, an inspection system having a wavelength of 365nm and a NA of 0.9 and σ of 1 requires 1125 excitations with Abbe's formulation while the decomposition technique requires 82 excitations to model the mutual intensity over a $2 \times 2\mu m$ simulation domain. The decomposition technique is accurate over a circular window. This was shown to cause some inaccuracies in periodic structures because aliasing occurs when a radially symmetric source is applied to a rectangular simulation domain. However, for isolated structures, the circular window used by the decomposition technique does not lead to inaccuracies because the mutual intensity need only be decomposed in a region than encompasses the isolated structure of interest.

The scalar theory and the scalar decomposition technique in this chapter will be used in future chapters. In Chapter 4, the scalar theory is used to develop a perturbation model to describe defect printing in phase shift masks. In Chapter 6, the scalar decomposition is expanded to include polarization of the vector electromagnetic field.

4 Effect of Scalar Partial Coherence on Defect Printing

4.1. Introduction

In this chapter, the scalar imaging theory described in Chapter 3 is used to characterize defect printing. Understanding defect printing is important as the ability to make defect free masks is a key requirement in optical projection printing. The term "defect free" usually implies "printable defect free," and this chapter characterizes acceptable limits on defect size, phase, location, etc. The characterization depends on many parameters such as tool, mask, and defect parameters. Understanding the tendency of defects to print is compounded by interactions between the defect and feature due to defocus, defect location, and coherence interactions. Because printing depends on all these interactions, an algebraic model which can rapidly assess the impact of a defect on a feature is developed in this chapter.

Using the scalar theory developed in Chapter 3, this chapter presents a modeling based approach in Section 4.2 to provide physical insight into the mechanisms that effect defect printing. In this approach a perturbational model is developed which characterizes defect interaction with features. In Section 4.4, the perturbational model is used to calculate the impact of a defect on a line, on an array, and on a contact hole. Using the perturbational model, this impact is evaluated by examining the aerial image, the image inside a thin-film stack, and the image due to modified illumination. In Section 4.5, the physical attributes of the optical system used in the perturbational model are characterized and those attributes that have the largest impact on defect printing are discussed. In Section 4.6, the trends in defect printability found by using the perturbational model are verified by experiment.

4.2. Perturbational Model

In this section the perturbation model developed for isolated phase shift defects of Ref. 111 is extended to include the interaction between a defect and a feature. In this extension, the mask is modeled by breaking the problem into five contributions. These five contributions add to give the same electric field transmitted through the mask and are shown in Figure 4-1. The first contribution is modeled as a plane wave transmitted through the background of the mask with phasor representation $\sqrt{T_B}e^{i\phi_B}$. Since this is a uniform field and the effect of the lens system is normalized to a clear field image field of $1.0\angle 0°$, the electric field at the wafer is just $\sqrt{T_B}e^{i\phi_B}$. The feature is constructed by subtracting the geometry of the feature in an opaque background with the geometry having transmittance and phase of the feature which is represented with the phasor, $\sqrt{T_B}e^{i\phi_B}E_F$ and then adding. These phasor notations for the feature contain an additional term, E_F, which represents the imaged electric field of the feature in an opaque background. Next, the imaged electric field of the defect is taken into account. This defect electric field is modeled similarly to the feature in that it is represented by subtracting a pinhole in an opaque background with the pinhole having the same transmittance and phase of the background, and then by adding a pinhole with phase and transmittance of the defect in an opaque background. These fields are represented with phasors, $\sqrt{T_B}e^{i\phi_B}E_P$ and $\sqrt{T_D}e^{i\phi_D}E_P$, respectively where the E_P term represents the imaged electric field of a pinhole. These pinholes act as point sources, and the image of this point source is calculated in the isolated defect perturbational model with the Lommel functions[9].

The total intensity of the feature and defect is calculated by taking the magnitude squared of the sum of the five electric field contributions as shown in Equation 4-1. The total intensity contains cross terms between the feature and the defect weighted by the mutual coherence. After expanding Equation 4-1, the terms are grouped into the intensity

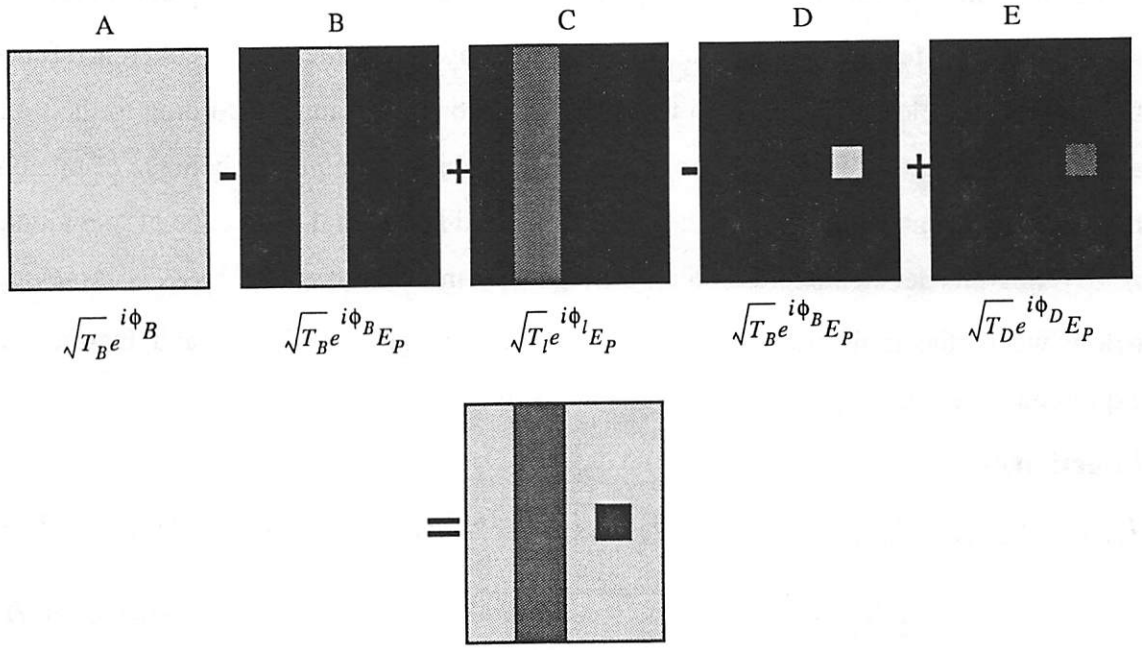


Figure 4-1. Pictorial of the extended perturbation model.

This example shows a defect near an attenuating phase line, where the line and defect are decomposed into electric fields.

Equation 4-1.
$$I_{TOT} = (A+B+C+D+E)\,(A+B+C+D+E)^*$$

of the feature and defect with additional cross terms to include the interaction between the defect and the feature in Equation 4-2.

Equation 4-2.
$$I_{TOT} = I_{feature} + I_{defect} - AA^* + 2Re\,(BD^*)$$
$$+ (2Re\,(BE^*) + 2Re\,(CD^*) + 2Re\,(CE^*))$$

The intensity of the feature is calculated through SPLAT and the intensity of the defect is calculated with the perturbational model for an isolated defect of Ref. 111. The cross terms between the feature and the defect are now expressed in terms of the electric fields, E_F and E_P, of a feature and a pinhole in an opaque background. The influence of E_P on E_F may be expressed through the mutual coherence theory developed in Chapter 3 and is defined below in Equation 4-3 where μ_{12} and γ is the magnitude and phase of the mutual

Equation 4-3.
$$E_F E_P^* \equiv \mu_{12} e^{i\gamma} \sqrt{I_F}\sqrt{I_P}$$

coherence function (MCF) between a pinhole and a feature, and where I_F and I_P are the intensities of the feature and pinhole, respectively. Physically, the mutual coherence function weighs the electric fields from the defect and from the feature according to the time average cancellations; therefore, the electric fields add for a mutual coherence of magnitude one (coherent illumination) while intensities add for mutual coherence of magnitude zero. Using this detailed expression for each pattern, the total intensity, I_{TOT}, is expressed below where the transmittance and phase of the background, defect, and feature are expressed by T_B, ϕ_B, T_D, ϕ_D, T_F, and ϕ_F, respectively.

Equation 4-4.

$$I_{TOT} = I_{feature} + I_{defect} - T_B + 2\mu_{12}T_B\sqrt{I_F I_P}\cos\gamma - 2\mu_{12}\sqrt{T_B T_D}\sqrt{I_F I_P}\cos(\phi_B - \phi_D + \gamma)$$
$$-2\mu_{12}\sqrt{T_F T_B}\sqrt{I_F I_P}\cos(\phi_F - \phi_D + \gamma) + 2\mu_{12}\sqrt{T_F T_D}\sqrt{I_F I_P}\cos(\phi_F - \phi_D + \gamma)$$

This extension of the perturbational model to include the impact of a defect on a feature reduces the complexity of the problem as the transmittance and phase parameters of the mask have been removed through Equation 4-4. Therefore, the mutual coherence is only a function of the mask geometry, illumination, and imaging system. Once this mutual coherence function is determined, the total image intensity may be evaluated for any background transmittance and phase and for any defect and feature transmittance and phase. The mutual coherence in Equation 4-4 above is the coherence measured after the mask is imaged onto the wafer. Consequently, this coherence differs from the coherence measured at the mask. The coherence on the wafer includes propagation effects through the imaging system and propagation effects of multiple reflections of rays in a thin film substrate. This total intensity is evaluated by calculating two separate images, the intensity of the feature and the intensity of the isolated defect, plus the contribution of the mutual coherence between the feature and defect. The image of the feature can be simulated and stored in advance as this image is the desired intensity result. The image of the isolated defect can be calculated with the isolated defect perturbational model. Finally, the mutual coherence must be calculated which will be shown as a function of parameters that may be simulated and stored in advance, as described in the next section.

4.3. Calculation of the Mutual Coherence Function

In this section, the mutual coherence used in the perturbational model of Section 4.2 is calculated. The mutual coherence between a defect and a feature weighs how much the defect electric field interacts with the feature electric field. This mutual coherence function is a function of the type of illumination, such as conventional, quadrupole, and annular, and is also a function of the condenser optics and of the thin-film substrate. While these interactions are rigorously modeled in thin-film SPLAT for any particular defect and pattern, both physical insight into the nature of this effect and rapid assessment are possible using an image perturbational approach based on modelling the mutual coherence.

In order to calculate the mutual coherence function between a feature and a defect, the electric fields from the pinhole and from the feature in an opaque background must be calculated. As discussed in the previous section the mutual coherence is represented by the electric field of the feature and defect normalized by the square root of the intensity of the feature and defect. In Section 3.3.1, it was shown that the electric field at the image plane for a feature is related to the electric field at the object plane by Equation 3-11. Equation 3-11 is repeated below as Equation 4-5. Similarly the imaged electric field of a pinhole is given in Equation 4-6 where t_o is the transmittance of the object, K is the transfer function

Equation 4-5.

$$a_i(x_{Fi}, y_{Fi}) = \int_{-\infty}^{\infty}\int K(x_{Fi}, y_{Fi}; x_{Fo}, y_{Fo})\, t_o(x_{Fo}, y_{Fo})\, a_o(x_{Fo}, y_{Fo})\, dx_{Fo} dy_{Fo}$$

Equation 4-6.

$$a_i(x_{Pi}, y_{Pi}) = K(x_{Pi}, y_{Pi}; x_{Po}, y_{Po})\, t_o(x_{Po}, y_{Po})\, a_o(x_{Po}, y_{Po})$$

of the imaging optics, and a_i and a_o is the electric field at the image and object plane, respectively. In Equation 4-5, (x_{Fo},y_{Fo}) and (x_{Fi},y_{Fi}) represent the spatial coordinates of the feature in the object plane (mask) and in the image plane (wafer), respectively. In Equation 4-6, (x_{Po},y_{Po}) and (x_{Pi},y_{Pi}) represent the coordinates of the pinhole at the mask and wafer, respectively. From these equations the cross term $E_F E_P^*$ is found by Equations 4-5 and 4-6 above and is shown in Equation 4-7 after rearranging the terms of integration.

The first term after the integral in Equation 4-7 is the mutual intensity incident on the

Equation 4-7.

$$a_{Fi}(x_{Fi}, y_{Fi})\, a_{Pi}{}^{*}(x_{Pi}, y_{Pi}) = \int_{-\infty}^{\infty}\int a_{Fo}(x_{Fo}, y_{Fo})\, a^{*}{}_{Po}(x_{Po}, y_{Po})$$

$$\times h(x_{Fi}, y_{Fi}; x_{Fo}, y_{Fo})\, h^{*}(x_{Pi}, y_{Pi}; x_{Po}, y_{Po})\, t_o(x_{Fo}, y_{Fo})\, t_o{}^{*}(x_{Po}, y_{Po})\, dx_{Fo} dy_{Fo}$$

object under the quasi-monochromatic assumption. Assuming that an incoherent source illuminates the object, this mutual intensity is given by the Van Cittert-Zernike theorem discussed in Section 3.3.3. From this theorem, the mutual intensity at the object plane is the Fourier transform of the incoherent source as given by Equation 3-17 and is repeated in Equation 4-8 below. For a circular source the mutual intensity is an Airy function as shown in Equation 4-9 below, where d is the separation between two points on the object plane. As shown in Section 3.3.1 through Equation 3-12, the transfer function of the imag-

Equation 4-8.

$$a_{Fo}(x_{Fo}, y_{Fo})\, a^{*}{}_{Po}(x_{Po}, y_{Po}) = \int_{-\infty}^{\infty}\int I_S(\alpha, \beta)\, exp\{j\frac{2\pi}{\lambda}[\alpha(x_{Po} - x_{Fo}) + \beta(y_{Po} - y_{Po})]\}\, d\alpha d\beta$$

Equation 4-9.

$$a_{Fo}(x_{Fo}, y_{Fo})\, a^{*}{}_{Po}(x_{Po}, y_{Po}) = 2\frac{J_1[2\pi\sigma d/(\lambda/NA)]}{2\pi\sigma d/(\lambda/NA)}$$

ing optics, K, is given by the Fourier transform of the pupil function, and Equation 3-12 is repeated below as Equation 4-10 where P is the pupil function. The pupil function including high NA, thin film effects, aberrations and defocus is given in Equation 3-21 of Section 3.4 and is repeated below as Equation 4-11.

Equation 4-10.

$$K(x_i, y_i; x_o, y_o; z) = \int_{-\infty}^{\infty}\int P(x_p, y_p; z)\, exp\{-i2\pi[x_p(x_o - Mx_i) + x_p(y_o - My_i)]\}\, dx_p dx_i$$

Equation 4-11. $$P(x_p, y_p) = \left[\frac{1 - NA^2(x_p^2 + y_p^2)/M^2}{1 - NA^2(x_p^2 + y_p^2)}\right]^{1/4} P_{ik}(x_p, y_p; z)\, M_{kj}(x_p, y_p)$$

$$\times exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p) \right] exp\left[-i2\pi z\frac{1}{NA^2}\sqrt{1-NA^2(x_p^2+y_p^2)} \right] circ\left(\sqrt{x_p^2+y_p^2}\right)$$

As outlined above, to calculate the impact of a defect on a feature, the mutual coherence function (MCF) between the feature and the defect must first be evaluated. To calculate the MCF, the geometry of the mask is evaluated in an opaque background which corresponds to a pinhole near a space as shown in Figure 4-2. In calculating the MCF the mutual coherence at the object plane which is only a function of the illumination source is calculated and stored. The image transform function, K, is next calculated and stored since it is a function of the condenser optics and the thin-film substrate. The transfer function and consequently the MCF is calculated along an imaging plane, which can either form an aerial image or form an image at a depth z in a thin-film stack.

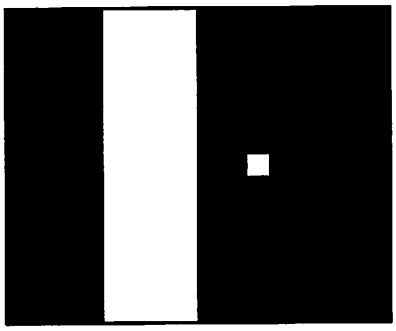


Figure 4-2. Mask used for calculating the MCF which corresponds to the geometry of Figure 4-10.

4.4. Comparison of the Perturbation Model with SPLAT Image Simulation

In this section, the accuracy of the perturbational model is validated through comparison to SPLAT simulation in Section 4.4.1. In Section 4.4.2, the perturbational model is used to systematically evaluate the impact that various size and phase defects have on the printing of attenuating phase shift lines, arrays, and contact holes. Section 4.4.3 extends the results of Sections 4.4.1 and 4.4.2 for aerial images to analyze defect impact inside a thin-film stack and for various modified illumination sources.

4.4.1. Defect Impact on the Aerial Image

The accuracy of the perturbational model is compared to SPLAT simulation for an isolated line and for an isolated space in an attenuating PSM. Using the perturbational model, the impact of a defect on a feature may be evaluated. The impact of a defect on an isolated line and space for attenuating phase shift masks will be discussed; however, the theory may also be used to evaluate the impact of a defect on a contact and an array. First Equations 4-3 and 4-7 are used to calculate the mutual coherence, μ_{12}, between a 0.35μm line and a 0.25 λ/NA square pinhole. Next, the images of an isolated attenuating phase shift line and of an isolated space in an opaque background are simulated with SPLAT. The images of an isolated defect and of a pinhole in an opaque background are calculated with the perturbational model through Ref. 111. Finally, Equation 4-4 may be used to calculate the total intensity for the defect interacting with the feature.

The perturbational model is compared to SPLAT for the image of a quartz bump defect on the edge of an isolated line as shown in Figure 4-3. In this example, the desired

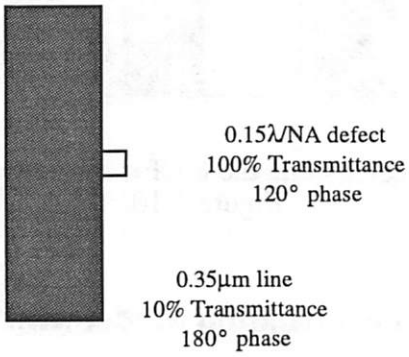


Figure 4-3. Geometry of an example showing a transparent phase defect on edge of a attenuating phase shift line.

feature is an isolated line that has 10% transmittance and 180° phase. The perturbational model is then used to calculate the impact of a square 0.25 λ/NA 150° 100% transmitting quartz bump that is on the edge of the line. Using SPLAT and the perturbational model, Equation 4-4 was used to calculate the total image due to the defect and the isolated line. Figure 4-4 shows the images calculated by the perturbational model and by SPLAT for an

attenuating phase shift mask at 0.0μm defocus. The perturbational model and SPLAT show close agreement along the image slope of the line with less agreement in the minimum. This discrepancy is caused by approximations used in calculating the image of the isolated defect. In the isolated defect perturbational model, the image of the defect is approximated to be the image of a defect created by fully coherent light. As previously shown, this approximation predicts more ringing in the image side lobes. This ringing causes the discrepancy between the image generated from SPLAT and by the perturbational model. The accuracy of the model can be improved by first breaking the defect into smaller defects, and then calculating the image as created by the sum of the interactions of the smaller defects. However, the line width is the parameter of interest, and the perturbational model and SPLAT are in agreement at the 0.3 intensity contour which is assumed to be the developed resist line width.

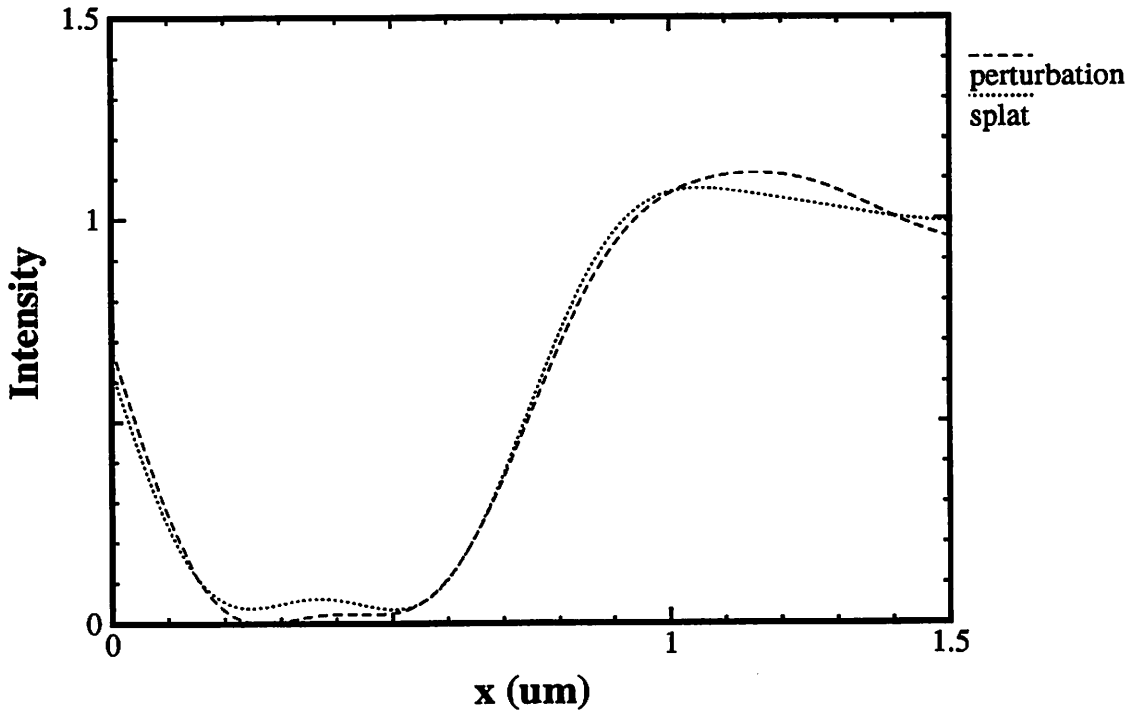


Figure 4-4. Comparison of aerial image calculated with the extended perturbational model and with SPLAT for the geometry shown in Figure 4-3.

The geometry is that of a 120° quartz bump on edge of an attenuating phase shifted line at 0.0μm defocus and σ=0.5.

Figure 4-6 compares images calculated with the perturbational model to images calculated with SPLAT at 0.5μm defocus. In this example, a defect is on edge of an isolated 0.35 μm space in a 10% transmitting, 180° phase background as shown in Figure 4-5. The

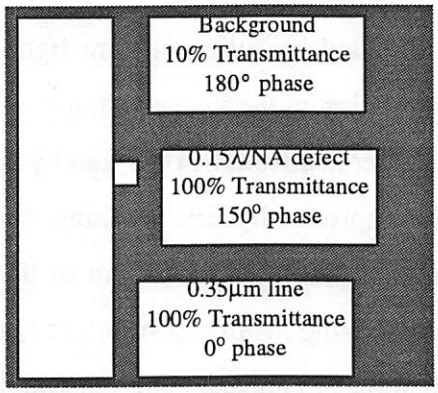


Figure 4-5. Geometry of an example showing a pinhole defect on edge of a space in an attenuating phase shift background.

pinhole is a square 0.25 λ/NA 100% transmitting defect with a 150° phase as compared to the 0° phase space. Since the geometry of Figure 4-5 is the same as the geometry of Figure 4-3, the same mutual coherence used to generate Figure 4-4 is used to Figure 4-6. This demonstrates the versatility of the perturbational model, as Equation 4-4 shows that the impact of a defect on feature is only a function of the geometry and imaging system; not the mask parameters. Figure 4-6 shows that the perturbational model and SPLAT are in close agreement. The isolated space is more accurate compared to the isolated line because the 10% transmittance of the background when printing an isolated space attenuates the ringing from the isolated defect, reducing the defect impact on the feature.

The extended perturbational model can also be used in the inspection process to generate contours which cause 10% line width variation at 0.3 intensity. Intensities are normalized to unity for a clear field mask. These contours were generated in Figure 4-7 using the same geometry in Figure 4-3, i.e., a quartz bump defect on edge of a 0.35μm 10% transmitting 180° phase isolated line. Figure 4-7 shows 10% line width variation contours as a function of defect phase and size for variable defocus. Defects of phase and size that lie to the right of a contour are printable while defects to the left are unprintable. Figure 4-7 also

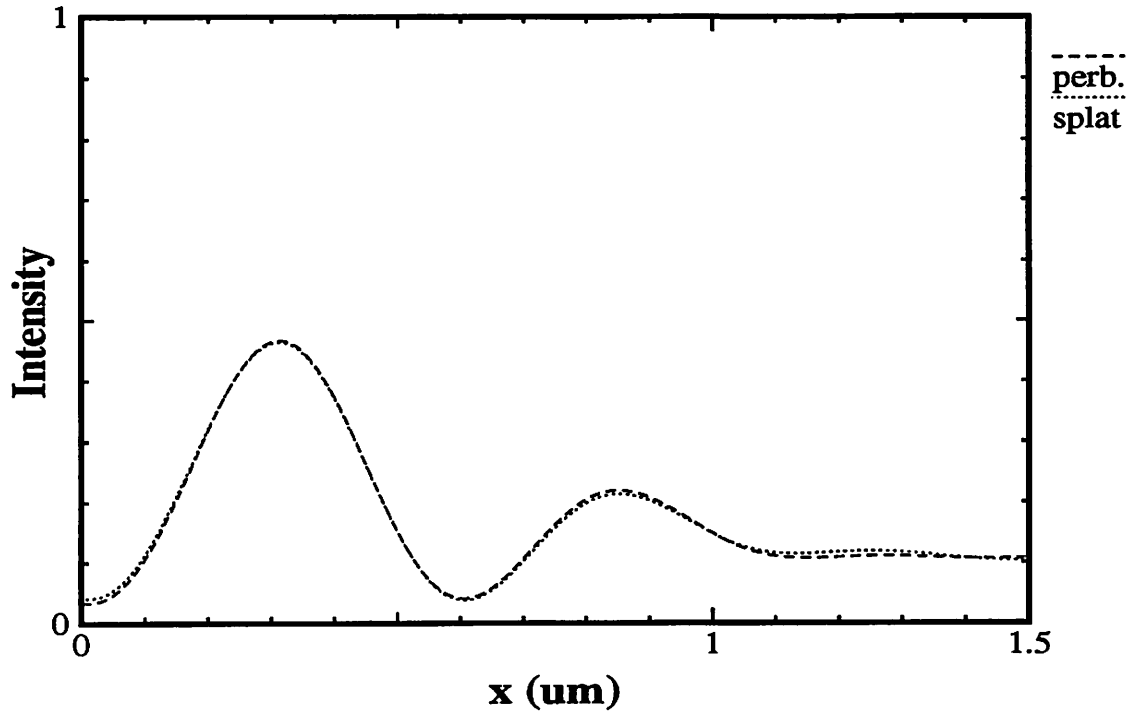


Figure 4-6. Comparison of aerial image calculated with the extended perturbational model and with SPLAT for the geometry shown in Figure 4-5.

The geometry is that of a 150° pinhole defect on edge of an attenuating phase shifted space in a 180° 10% attenuated phase shifted background at 0.5μm defocus and σ=0.5.

shows the tendency of a defect to cause more line width variation when out of focus. For example, a 120° phase quartz bump defect as small as 0.12 λ/NA will cause 10% line width variation at 0.75μm defocus while a larger 150° phase 0.12 λ/NA defect is needed to cause 10% line width variation at 0.0μm defocus. As an accuracy check, experimental results are plotted at 0, 0.25, 0.5, and 0.75μm defocus for phases of 150°, 120°, and 90°. The perturbational model shows more discrepancy from experiments when out of focus, and due to small errors in the isolated defect perturbational model. However, the extended perturbational model yields worst case results as contours generated by extended perturbational model always lie to the left of the experimental results.

4.4.2. Defect Impact on Lines, Arrays, and Contacts in Attenuating PSM

In this section, the perturbation model will be used to analyze several phase shifting mask (PSM) technologies. These PSM technologies are thin chrome attenuating PSM and

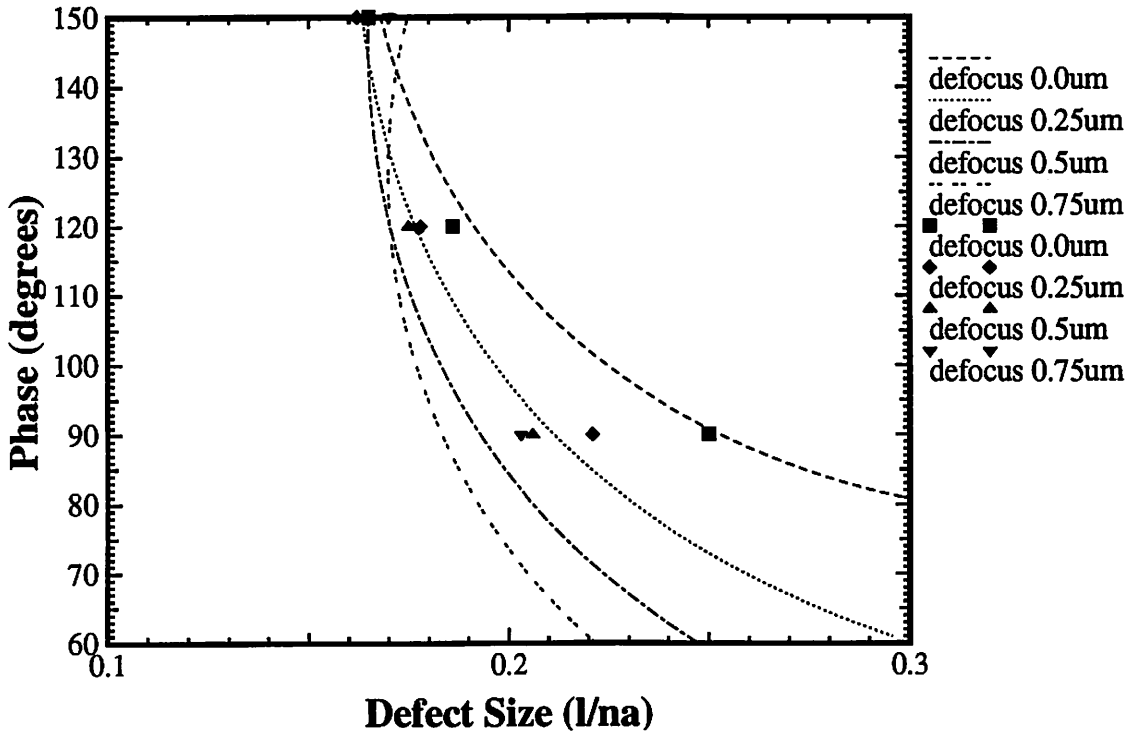


Figure 4-7. Using the extended perturbational model to generate printable defect guidelines.

Solid lines are from the model with discrete points from the experimental results.

embedded shifter attenuating PSM. In thin chrome attenuating PSM, as the name implies, a thin layer of 9-30° of chrome is deposited on a quartz blank or quartz with a spin on glass (SOG) layer blank. The light passing through this thin layer of chrome is not fully attenuated and the transmittance of light through the mask ranges between 6-10%, depending on the thickness of chrome. The chrome is then patterned and etched. Following the chrome etch, the quartz or SOG backing on the mask is typically etched 150° in order to achieve a total phase change of 180° for light passing through the feature and through the background. In the embedded shifter attenuating PSM, the attenuation and phase transformation takes place in a single material and no quartz or SOG etch is needed. In this approach the embedded shifter is deposited to a thickness of 180° and to the desired attenuation. The mask is then patterned and the full 180° of the attenuating material is etched where features are to be defined. This type of technology is advantageous since only one material needs to be etched as opposed to two materials in thin chrome attenuated PSM. Later, it

will also be shown that the embedded shifter attenuating PSM is more defect tolerant than the thin chrome attenuating PSM.

In these two types of attenuating phase shift technologies, different types of defects result from problems in patterning and etching the mask. To facilitate discussion of these defect types, we need to develop appropriate terminology. In thin chrome attenuating PSM, a common defect occurs when, due to patterning or adhesion problems, the thin chrome attenuating material flakes off. When this type of defect occurs inside the feature, it is called a pinhole; when this type of defect occurs outside the feature it leaves a pillar of quartz or glass behind and is called a quartz bump. These quartz bump defects may also be partially etched during the quartz etch leaving behind a variable phase quartz bump defect. Also, pinhole defects in thin chrome attenuating PSM may be fully etched for a total of 180° or may be partially etched for a variable phase defect. In both the pinhole and quartz bump cases, the defect will be 100% transmitting in thin chrome attenuating PSM. Another defect may result when the attenuating material is left on top of the quartz bump, and this defect will be called a pinspot defect which, in the case of thin chrome attenuating PSM, will be a 10% transmitting 180° phase defect. Similar defects occur in embedded attenuating PSM. Like the thin chrome attenuating PSM, when a defect occurs inside of a feature it will be called a pinhole defect; however, unlike the thin chrome PSM, a defect that occurs outside of a feature will always be called a pinspot. Both the pinhole and pinspot defects may be fully or partially etched which causes a variable defect phase and transmittance. Note that a quartz bump with the attenuating material in thin chrome attenuating PSM and a bump of material in embedded attenuating PSM are both called a pinspot defect. This is because the perturbational model and the SPLAT simulations are scalar approaches, and both defects electrically appear to be the same.

Examples of defect interactions with features from simulation and from the perturbational model provide insight into the complexity of the impact of a defect on a feature. Defect interactions are discussed for three features: defects near isolated attenuated phase shift lines, defects near arrays, and defects near contacts. These three cases are summarized in tables which show worst case scenarios. The following simulations are studied for an i-

line system (365nm) with a 0.5 σ and an NA of 0.5. Results for other systems (such as NA=0.6) can be determined from the dimensions in λ/NA.

The impact of a defect was simulated for isolate lines and spaces. These defects had variable type, focus, defect size, defect phase, and separation between the defect and the feature. In the case of an isolated line, the line is 10% transmitting with a 180° phase and is 0.35μm (wafer dimension) wide, while an isolated space was 0.5μm wide in a 10% transmitting 180° phase background (a 0.15μm bias was applied to the space in order to allow sufficient light to pass). From these simulations, the placement of a quartz bump or a pinspot defect near an attenuating phase shift line is critical in causing line width variation. Quartz bump and pinspot defects have greatest impact when separated from the line; however, pinhole defects have the greatest impact when on edge of a space. Also, defects of phase other than 180° produce more line width variation when out of focus (Figure 4-8). For example, a 150° quartz bump defect produces the largest line width variation when at a defocus of about 0.25μm, and a 90° quartz bump has greatest impact on line width at close to 0.75μm defocus. This defocus effect on line width variation agrees with the isolated quartz bump defect in which the maxima in line width variation occurs at approximately the same focus. Figure 4-8 also shows that the impact of a quartz bump defect on line width variation is reduced if the phase of the defect is decreased which may possibly be achieved through an etching repair.

For defects near arrays, several observations about the impact of quartz bump and pinhole defects in attenuating phase shift arrays can be made from the simulations for a 10% transmitting, 180° phase, 0.35μm wide line/space pattern. Figure 4-9 shows that the line width variation is greater when the defect is separated from the array edge (the convention used throughout this paper is that separation is measured from the defect edge to the line edge). From Figure 4-9 the line width variation is proportional to the defect area and independent of aspect ratio for small area and relatively symmetric aspect ratios. For example, the $0.04 \times 0.08\mu m$ defect and the $0.02 \times 0.16\mu m$ defect produce about the same amount of line width variation, and as the area increases, the $0.04 \times 0.16\mu m$ defect and $0.08 \times 0.08\mu m$ defect produce similar line width variations (within 10%). However, if the

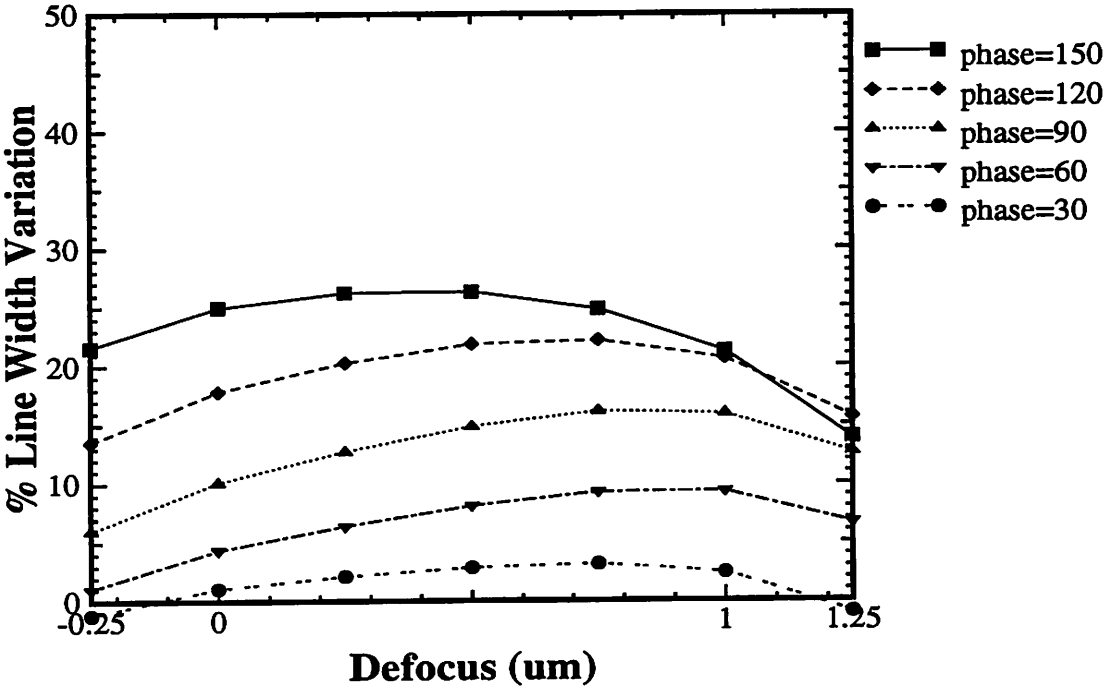


Figure 4-8. Percent linewidth variation versus defocus for a quartz bump defect near an isolated attenuating phase shift line.

The defect size is $0.25 \times 0.25\lambda/NA$ and the separation distance is 0.1μm.

area is increased further or if the aspect ratio is made more unsymmetrical, shape dependent line width variation differences occur. This is due to the fact that the defect can no longer be modeled as a pinhole and the mutual coherence function becomes dependent upon shape. Trends for defect impact on an array agree with trends for the isolated line. These trends are: quartz bump and pinspot defects cause more line width variation than pinhole defects, quartz bump and pinspot defects produce more line width variation when separated, and pinhole defects produce the most line width variation when on edge of the line.

For defects inside and outside a contact, several observations can be made about the effect of quartz bump defects, pinspot defects, and pinhole defects in attenuating PSM for a 0.5×0.5μm contact in an 10% transmitting 180° background (a 0.15μm bias was applied to the space in order to allow sufficient light to pass). A quartz bump defect in thin chrome PSM has a greater effect on contact width variation than the fully etched pinhole defect in embedded PSM. The quartz bump defect in thin chrome PSM causes more con-

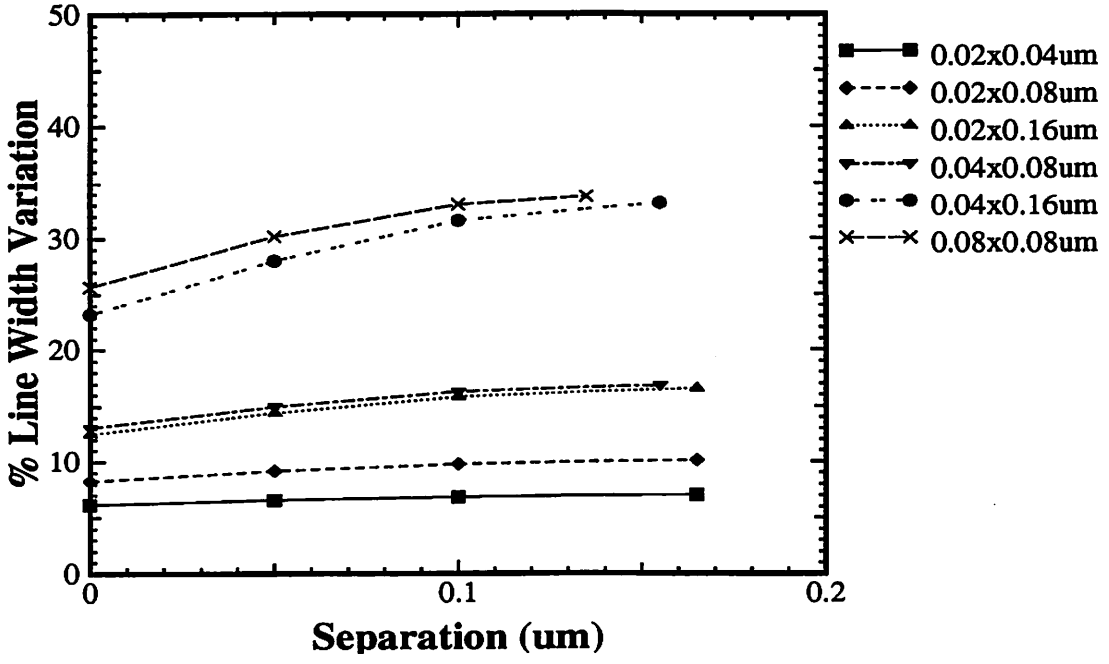


Figure 4-9. Percent linewidth variation for a quartz bump defect in an open region of an array.

The defect is missing the attenuating layer producing a 150° phase error when the defocus is 0.5μm.

tact width variation than a pinspot defect in embedded PSM. Also, defects in the middle of a contact have greater impact on the contact width variation, with a defect on edge causing the second most variation, and a defect in the corner producing the least.

From these simulations, we now attempt to systematically tabulate the minimum wafer dimension square defect. In the isolated defect case the smallest defect to cause an intensity minimum of 0.3 is recorded, and in a defect near a feature the defect which causes a 10% feature variation at an intensity of 0.3 is tabulated. In all the guidelines, the minimum size defect was assumed to be in the worst defect location such as the middle of an array, the middle of a contact, etc. The tables also apply through a 1.5μm (1 Rayleigh Unit) defocus range in which the defect size was determined at the worst focus condition. The composite set of guidelines for the printability of defects based on defect type, size, and phase and are shown in Tables 4-1, 4-2, 4-3, and 4-4.

The smallest defects capable of producing a 10% feature variation in thin chrome

attenuating PSM are highlighted in Tables 4-1 and 4-2. These tables show the guidelines for defects in which a pinhole is missing in an attenuating thin chrome layer of 10°, 20°, or 30° phase and for defects in which a quartz bump is missing the attenuating material. As stated previously, the isolated line has a width of 0.35μm while an isolated space and contact are oversized with a 0.50μm width. From Tables 4-1 and 4-2 the smallest defect to

Table 4-1. Pinhole absorber defects in thin chrome attenuating PSM

Missing Absorber Defects	Phase Error $\Delta\phi$	Minimum Printable Defect Size		
		$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
Isolated pinhole	10°	0.25	0.18	0.15
	20°	0.25	0.18	0.15
	30°	0.25	0.18	0.15
Pinhole near an isolated space	10°	0.35	0.26	0.21
	20°	0.35	0.26	0.21
	30°	0.35	0.26	0.21
Pinhole in an array	30°	0.18	0.13	0.11

Table 4-2. Quartz bump defects in thin chrome attenuating PSM

Quartz Bump Defects	Phase Error $\Delta\phi$	Minimum Printable Defect Size		
		$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
Isolated quartz bump	150°	0.30	0.22	0.18
	120°	0.33	0.24	0.20
	90°	0.40	0.29	0.24
	60°	0.49	0.36	0.30
Quartz bump near a line	150°	0.16	0.12	0.10
	120°	0.17	0.12	0.10
	90°	0.19	0.14	0.12
	60°	0.26	0.19	0.16

Table 4-2. Quartz bump defects in thin chrome attenuating PSM

Quartz Bump Defects	Phase Error $\Delta\phi$	Minimum Printable Defect Size		
		$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
	30°	0.43	0.31	0.26
Quartz bump in an array	150°	0.07	0.05	0.04
Quartz bump in a contact	150°	0.13	0.10	0.08

cause 10% line width variation is the quartz bump. A 150° quartz bump defect as small as 0.07 λ/NA in the middle of an array causes 10% feature variation while a 0.13 λ/NA 150° quartz bump in the middle of a contact produces 10% variation. The third worst defect and the worst pinhole defect in thin chrome attenuating PSM is a 0.18 λ/NA pinhole with a 30-degree phase error in the thin chrome material.

Tables 4-3 and 4-4 list the smallest size defects in embedded attenuation PSM that are capable of causing a 10% variation in the feature. Table 4-3 shows pinhole defects for

Table 4-3. Missing absorber in embedded attenuating PSM

Missing Absorber 180° phase 10% Att. Error	Minimum Printable Defect Size		
	$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
Isolated pinhole	0.35	0.26	0.21
Pinhole near an isolated space	0.25	0.18	0.15
Pinhole in an array	0.13	0.10	0.08
Pinhole near a contact	0.21	0.15	0.13

Table 4-4. Extra absorber in embedded attenuating PSM

Extra Absorber 180° phase 10% Att. Error	Minimum Printable Defect Size		
	$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
Isolated pinspot	0.34	0.25	0.21
Pinspot near an isolated line	0.15	0.11	0.09

Table 4-4. Extra absorber in embedded attenuating PSM

Extra Absorber 180° phase 10% Att. Error	Minimum Printable Defect Size		
	$\frac{\lambda}{NA}$	NA=0.5 (μm)	NA=0.5 (μm)
Pinspot in an array	0.10	0.07	0.06
Pinspot in a contact	0.14	0.10	0.09

missing absorber defects which produce a phase 10% transmittance error. Table 4-4 lists the pinspot defects or extra absorber defects that have a 180° phase 10% transmittance error. It should be noted that partially etched defects of variable phase and attenuation are also possible. However, a 180-degree phase 10% transmitting defect will always cause the most feature variation in embedded attenuating PSM and are the only defects listed in Tables 4-3 and 4-4.

Several defects from the table should be highlighted. Once again a pinspot defect in embedded attenuating PSM causes the most feature variation, and a pinspot as small as 0.10 λ/NA in the middle of an array is capable of producing a feature width variation of 10%. The second smallest pinspot defect to cause 10% line width variation is a 0.14 λ/NA pinspot defect in the middle of a contact. The pinhole defect that produces the most line width variation in embedded attenuating PSM is a 0.13 λ/NA pinhole in the middle of an array. In summary, as in the case of a quartz bump in an array, a pinspot defect as small as 0.10 λ/NA may be difficult to detect during mask inspection.

Comparing the results of all four tables, several observations about defect printability should be highlighted. First, the minimum pinspot defect size is smaller than the minimum pinhole defect size. This is consistent with the previous case study results, where it was shown that a pinspot defect impacts a feature more than pinhole defects. Second, a 150° quartz bump defect causes more feature variation than embedded 10% transmitting 180° phase pinspots; however, the minimum size quartz bump defect depends on the phase error, as defect size is inversely proportional to defect phase error. Third, a 150° quartz bump defect as small as 0.07 λ/NA will cause a 10% line width variation in the array pat-

tern. This implies that if an i-line 5X stepper with NA=0.6 is used, a mask defect of 0.2μm needs to be detected prior to printing. Finally, an important observation is that the minimum size defects in embedded attenuating PSM capable of producing 10% feature variation are larger than the minimum size defect in thin chrome attenuating PSM. This implies that embedded attenuating PSM's are more defect tolerant than thin chrome attenuating PSM's.

4.4.3. Effect of Modified Illumination and Thin-Film Substrates on Defect Printing

In this section the effect of modified illumination and thin-film substrates on defect printing is not only evaluated using the perturbational model formulated in Section 4.2, but also is compared to SPLAT simulation. The mutual coherence function (MCF), μ_{12}, is calculated with Equations 4-3 and 4-7. The effect of modified illumination on the MCF is found by first using Equation 4-8 to calculate the MCF at the object plane and substituting this into Equation 4-7 which gives the MCF in the image plane. The effect of thin-film substrates on the MCF is found by using the modified pupil function of Equation 4-11. In Equation 4-11, the thin-film substrate is taken into account through $P_{ik}(x_p,y_p;z)$ which is the OPD of ray tracing the light to a plane, z, inside the thin-film stack.

After calculating the MCF, Equation 4-4 is then used to calculate the impact of a defect on a feature. The mask defect and feature considered in this section is that of a 0.15λ/NA defect with phase of 150° and transmittance of 100% separated from a 0.35μm line

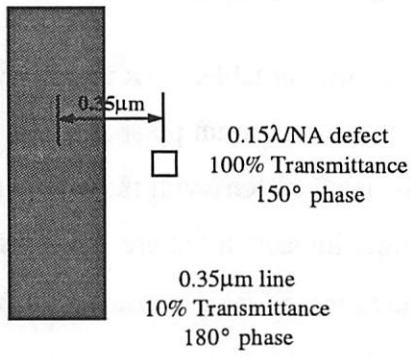


Figure 4-10. Mask used in the perturbational model example.

that is 10% transmitting with a 180° phase. This geometry is shown in Figure 4-10 where the separation from the center of the line to the center of the defect is 0.35μm. This mask pattern was illuminated at i-line with three different sources and collected with a condenser having an NA of 0.5. These three illumination sources are the conventional "top-hat," annular, and quadrupole illumination; however, any illumination source of arbitrary shape may be evaluated. Figure 4-11 shows the pupil of these three illumination sources normalized by the NA of the condenser, i.e, the conventional illumination source has a σ of 0.5. The effect of the substrate on defect printing was determined for three substrates using three different illumination systems. These three substrates are: aerial image (no substrate), resist over a polysilicon gate stack, and resist over a TiN ARC (Anti-Reflective Coating), poly gate stack. The dimensions and material parameters of the later two stacks are shown in Figure 4-11. Unless otherwise stated for all three illumination sources, a 0.5μm defocus was used, and the impact of the defect on a feature was found in the middle of the resist (0.5μm) for the two gate stacks.

A comparison of the MCF at the wafer for the three different illumination sources shows that the magnitude of the MCF for quadrupole illumination is less than the magnitude of the MCF for annular or conventional illumination. Using Equations 4-3 and 4-7 the magnitude and phase of the MCF was calculated for the geometry of Figure 4-10 for the three illumination sources and the three substrates. For the three illumination sources, the MCF was found with no substrate at a defocus of 0.5μm as shown in Figure 4-12. In this figure, the line is centered at 1.0μm and the defect is centered at 1.35μm, the MCF is calculated in the image plane at 0.5μm into the resist. The ripples in the magnitude of the MCF in Figure 4-12 occur when the image of the space or pinhole are near or at a null in the intensity image. If the image transfer function becomes a constant due to an infinitely large numerical aperture or at extreme defocus, these ripples disappear and the magnitude of the MCF is a constant. Therefore, the ripples are due to the low pass filtering action of the imaging system. The discrete points in Figure 4-12 indicate the magnitude of the MCF (μ_{12}) at the line edge (1.175μm) where the point marked with a square, a diamond and a triangle indicate conventional, annular, and quadrupole illumination, respectively. These discrete points occur at a MCF magnitude of 0.80, 0.73 and 0.57 for conventional, annular,

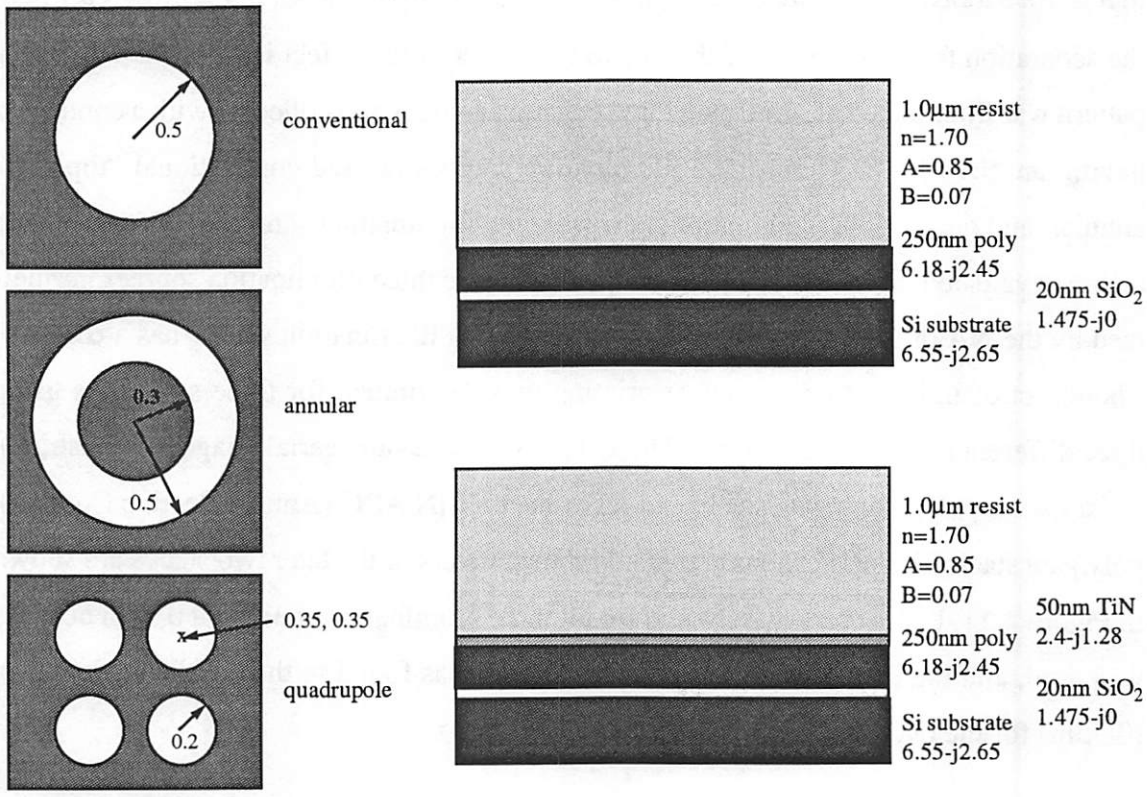


Figure 4-11. Illumination sources and thin-film stacks used in calculating the MCF in the perturbational model.

and quadrupole illumination, respectively. It should also be noted from Figure 4-12 that the phase of the MCF (γ) is almost independent of illumination as the phase deviates by 1° to 3°.

The MCF was calculated for the geometry of Figure 4-10, for the thin-film substrates of Figure 4-11, and for the three illumination sources. The magnitude and phase of the MCF for the poly gate stack with and without TiN ARC are shown in Figure 4-13. The MCF's in Figure 4-13 was calculated in the middle of the resist at 0.5μm defocus, and the line and defect are located in the same position of Figure 4-10. Once again, Figure 4-13 shows not only that the magnitude of the MCF is smaller in quadrupole illumination but also that conventional illumination has the greatest MCF magnitude. The spike in the magnitude of the MCF at 1.45μm occurs at a null in the image intensity of a space. When this MCF is substituted into the total perturbational model intensity Equation 4-4, the spike is

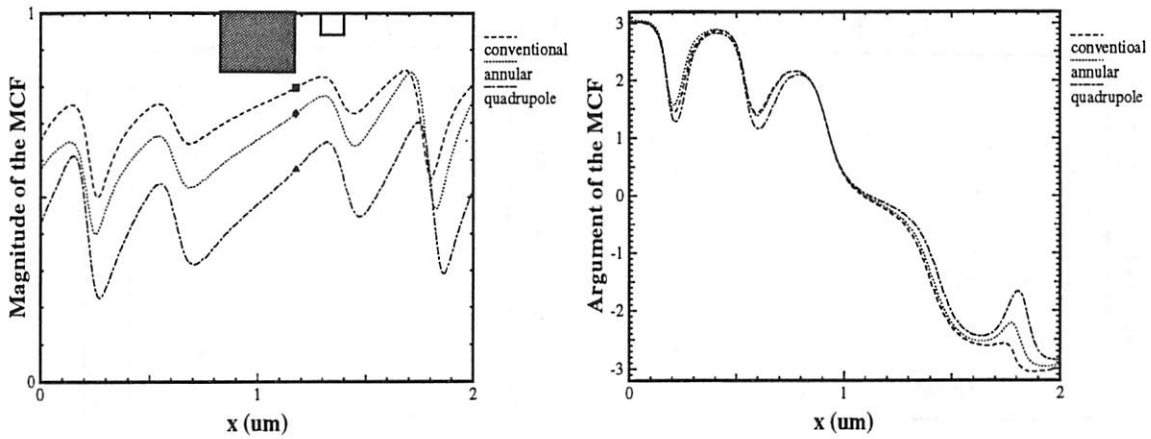


Figure 4-12. Magnitude (left) and argument (right) in radians of the MCF for the thin-film stack without a substrate at 0.5μm defocus.

cancelled by the null in the image intensity of the space; therefore, this spike is a result of the numerical calculation and does result from a physical phenomenon. By comparing the three discrete points in Figures 4-12 and 4-13 which indicate the magnitude of the MCF at the line edge, the magnitude of the MCF for the three types of substrates are all within one percent for a given illumination; consequently, the MCF is a weak function of the substrate. The phase of the MCF shown in Figures 4-12 and 4-13 also show that the phase of the MCF is a weak function of the thin-film substrate and illumination as the phase is only a function of the mask geometry and the condenser optics.

Both the thin-film stacks of Figure 4-13 and the MCF were calculated for various positions within the resist at variable defocus, resist thickness, and defect-to-line separation. Varying the defocus causes both the magnitude and phase of the MCF to become smoother as the MCF approaches a constant DC value with a 0° phase. The resist thickness was increased in quarter-wave increments to insure that the MCF was not being calculated at an electric field null. By varying the defect-to-line separation, the maximums and minimums in the magnitude of the MCF where found to be a function of the mask geometry. In all the calculations, the MCF was found to be a weak function of the thin-film substrate.

Intensity images calculated with the perturbational model are in excellent agreement with images from thin-film SPLAT. Using the perturbational model and the MCF in the previous figures, the image of Figure 4-10 was calculated for various substrates and

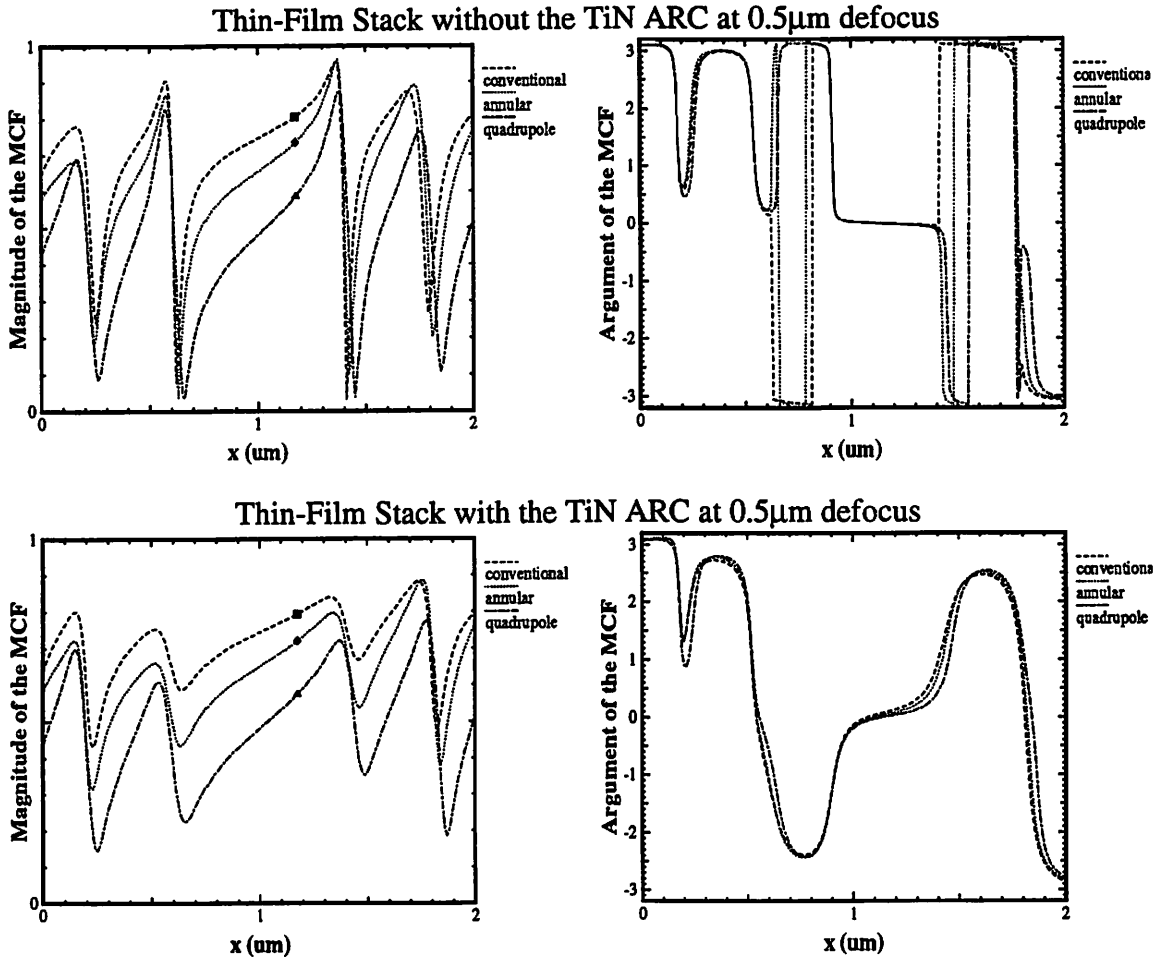


Figure 4-13. Magnitude and phase of the MCF for the thin-film stacks in Figure 4-11.

illuminations and compared to thin-film SPLAT. The perturbational model is compared to thin-film SPLAT in Figure 4-14 for the two different thin-film stack structures of Figure 4-11: a thin-film stack without the TiN ARC at 0.5μm defocus, and a thin-film stack with the TiN ARC at 0.5μm defocus. In Figure 4-14 the image of the isolated line was calculated with SPLAT while the image of the defect was estimated using an algebraic expression for the intensity from the defect perturbational model of Ref. 111. The impact of this defect on the image of the line was then calculated with the perturbational model of Equation 4-4. Figure 4-14 demonstrates that the perturbational model is in close agreement with thin-film SPLAT and can be used for spread sheet type estimates of defect printability. The slight error in the images was due to the grid size discretizing the number of nearby pixels in Equation 4-7.

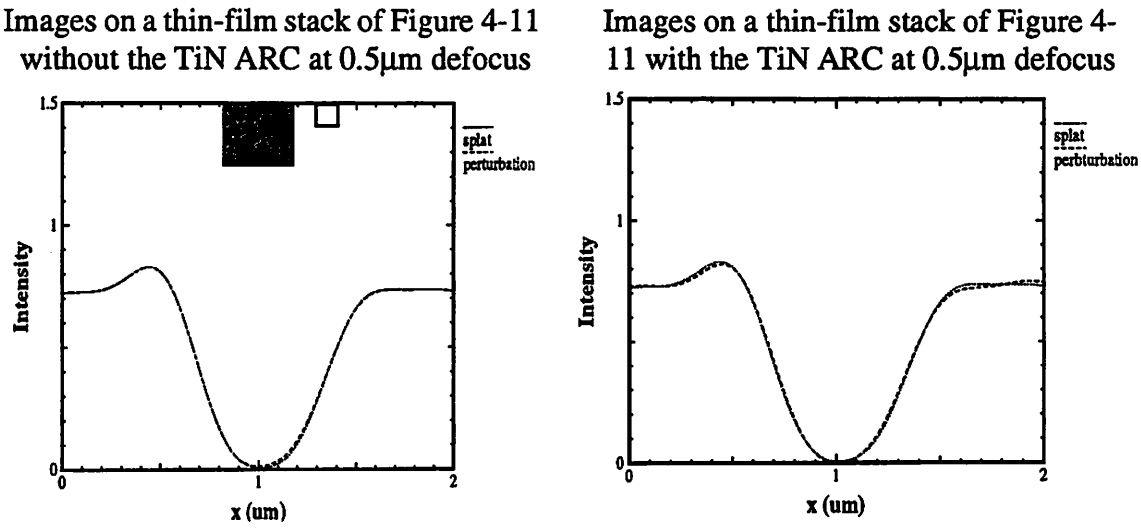


Figure 4-14. Comparison of the images from thin-film SPLAT and from the perturbational model.

4.5. Importance of the Aerial Image Edge Slope on Defect Printing

In Section 4.4.3, it was shown that the magnitude of the MCF is least for quadrupole illumination, followed by annular, with conventional illumination having the largest MCF magnitude. The magnitude of MCF was shown to be a weak function of the thin-film substrate. This would seem to indicate that since the MCF is least for quadrupole illumination, quadrupole illumination is more defect tolerant, i.e., under quadrupole illumination defects are less likely to print. Although the MCF is less for some illumination schemes, the impact the defect has on the feature also depends on the intensity of the feature and the pinhole defect. This is seen from Equation 4-4 of the perturbational model. In Equation 4-4, the impact the defect has on the feature is represented in the cross terms of Equation 4-3, $E_F E_P^* \equiv \mu_{12} e^{i\gamma} \sqrt{I_F} \sqrt{I_P}$. These cross terms are a function not only of the MCF, μ_{12}, and the intensity of the feature but also the pinhole defect, I_F and I_P. Therefore, in order to evaluate the impact that a defect has on a feature, the role of the MCF, as well as I_F and I_P must be considered.

The role of the defect is small because the defect size is small, the intensity of the defect is independent of the type of illumination. The role of the feature, however, is more significant because the intensity of the feature is a strong function of the type of illumina-

tion. The significance of the feature intensity on defect printing in binary masks was shown by Neureuther to be a function of the line edge slope of the image intensity [82]. Neureuther showed that the linewidth variation is inversely proportional to the line edge slope of the feature intensity, i.e., an image of a feature with large slope is more defect tolerant than a feature producing a smaller image intensity slope. Since the perturbational model shows that the mutual intensity, μ_{12}, is a function of the mask geometry only, this observation also applies to phase shift masks.

The role of the image slope on defect printing is best understood by example. The aerial images produced by a binary mask with a 0.5μm line-space (1.0μm pitch) array and with a 0.35μm (0.7μm pitch) array are shown in left and right plots of Figure 4-15, respec-

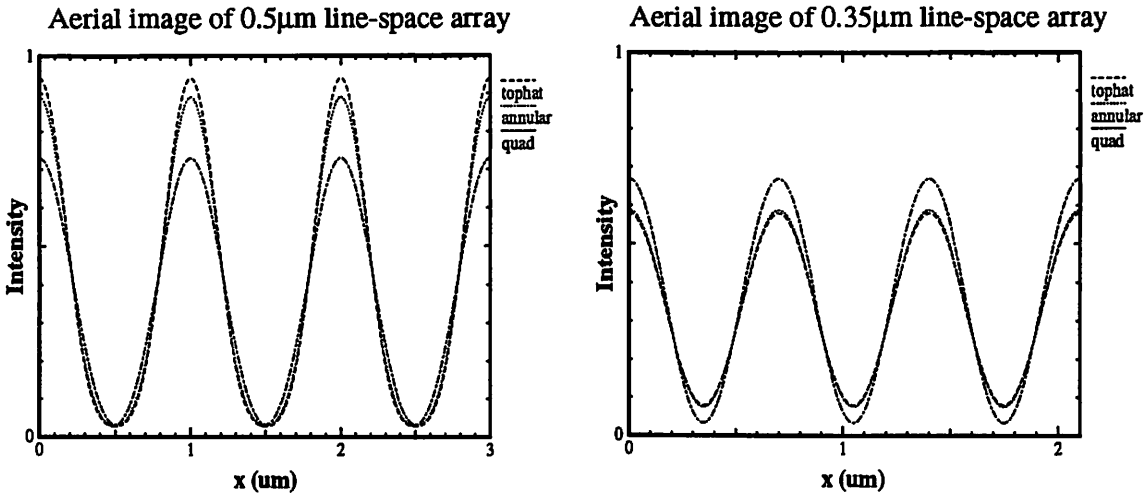


Figure 4-15. Aerial images of a chrome line-space array.

The image on the left is for a 0.5μm line-space array while the image on the right is for a 0.35μm line-space array. Both aerial images are for an i-line stepper with NA of 0.6 and σ of 0.5.

tively. The images are those produced by an i-line (λ=365nm), NA=0.6, and σ=0.5 stepper with conventional (top hat), annular, and quadrupole illumination sources. Figure 4-15 indicates that the image edge slope is greatest with conventional illumination in the 0.5μm line-space array while quadrupole illumination produces the greatest slope in the 0.35μm line-space array. Since the linewidth variation is inversely proportional to the image slope, the larger image slope for quadrupole illumination for dense line-space arrays implies that

quadrupole illumination is more defect tolerant when the line-space pattern is dense. This is demonstrated in Figure 4-16 where the percent critical dimension (CD) loss is plotted as a function of line-space array 0.15×0.15μm chrome defect is in the middle of a chrome array. Figure 4-16 also indicates that when a defect is in the middle of the array, quadrupole

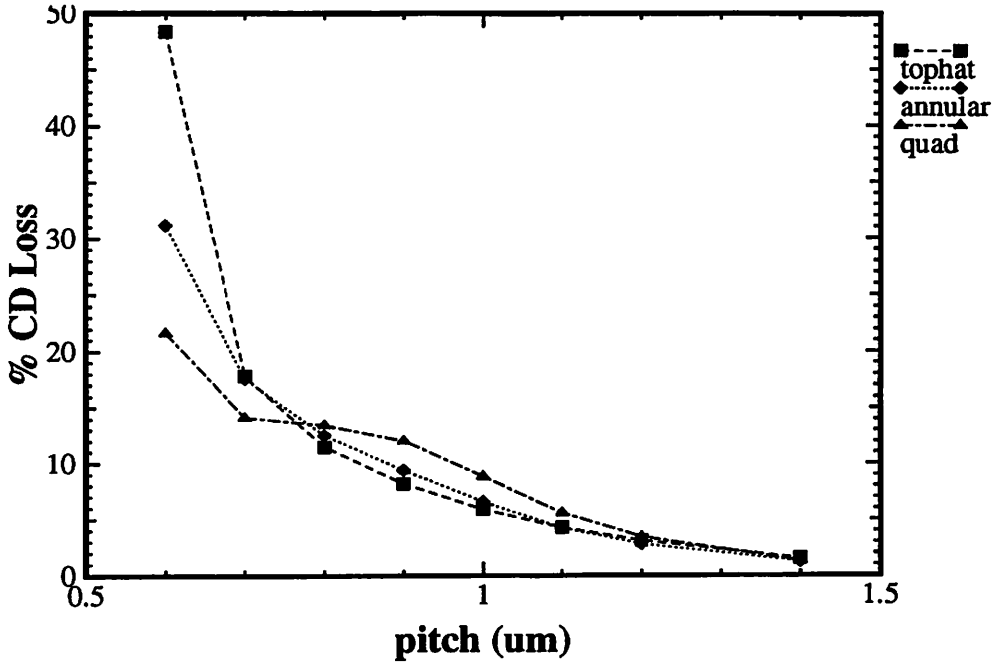


Figure 4-16. Effect of the aerial image line edge slope on the printing of a chrome defect in the middle of a chrome array.

The chrome defect is 0.15×0.15μm and is in the middle of a chrome array. The stepper is an i-line stepper with NA=0.6 and σ=0.5. Figure 4-16 shows that the line edge slope (or contrast) near a pitch of 0.8μm inverts the effect of the illumination on defect printing.

illumination produces more CD variation when the pitch is larger than 0.75μm. This result agrees with the work of Shaw[108] who found experimentally that a 1.0μm pitch array illuminated with a quadrupole source produces the most CD variation, and conventional illumination has the least CD variation. However, when the pitch is less than 0.7μm (dense array), the simulation results in Figure 4-16 show that the CD variation is less with quadrupole illumination. This reduction in CD loss is due to the increased line image edge slope when a dense array is illuminated with a quadrupole source as seen in Figure 4-15. In conclusion, although the MCF may be less for one type of illumination, the impact a defect has on a feature depends on both the line edge slope of the image and on the MCF.

4.6. Experimental Verification

In this section, the perturbational model is compared to experiments conducted under the guidance of Rajeev Singh of SEMATECH, and the results were originally published in Ref. 114. In this experimental work, a thin chrome attenuating mask (6% transmission at a wavelength of 365nm) was manufactured with a test pattern having programmed defects of different sizes. The mask was quartz etched to provide the 180° phase shift between the light transmitted through the transparent area and through the attenuating film area. Experiments were performed on a mask with three different programmed defect test structures in a 1.75μm (mask dimension, throughout Section 6 mask dimensions are reported) line/space array. The first test structure (C1EG175) is a variable size pinspot defect protrusion with 6% transmittance and 180° phase on edge of the array line. The second test structure (P1EG175) is a variable size, 100% transmitting, 170° quartz bump protrusion on edge of the array line. The last test structure (M1EG175) exposed is the fully etched, 100% transmitting, 180° phase error pinhole defect on edge of the array line. The mask was then exposed on a bare silicon wafer with a 1.07μm thick high contrast positive resist. An i-line stepper with a NA=0.6, 5X magnification, and σ=0.6 was used, and the best exposure and best focus was found to be 400mJ/cm^2 and -0.3μm, respectively. Critical dimension (CD) measurements were then made on the exposed wafer for three test structures by using an AMRAY 1830 SEM while a Hitachi 6000 SEM was used to measure the mask line, space, and defect dimensions.

The fabrication of programmed defects in attenuating PSM's is complicated by the etch of the mask quartz backing. During this etch step the anisotropic etch will cause under-sizing of the defects. This under-sizing occurs during both the chrome and quartz etch. The etching process also causes rounding of the defect corners. This corner rounding makes measuring the defect area difficult, and some estimation of the defect area is needed. Throughout this paper, when measured defect dimensions are reported, the maximum amount the defect protrudes or intrudes on the feature is reported first, and the width of the defect at 50% of the maximum protrusion or intrusion is reported last. The rounding of

defect corners is clearly seen in Figure 4-17a for a SEM taken of a 170° quartz bump protrusion on edge of a 1.75μm wide array line.

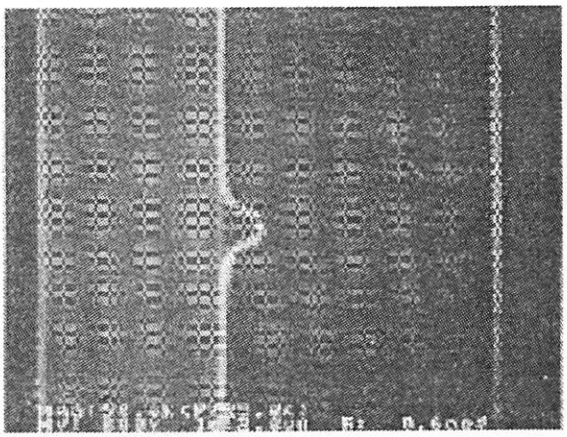


Figure 4-17a. Mask image of quartz bump defect on edge of an 1.75μm APS line.

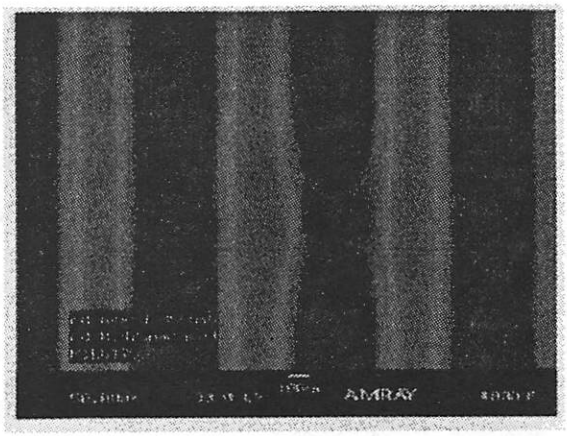


Figure 4-17b. Effect of a quartz bump defect on an array (P1EG175) after exposure.

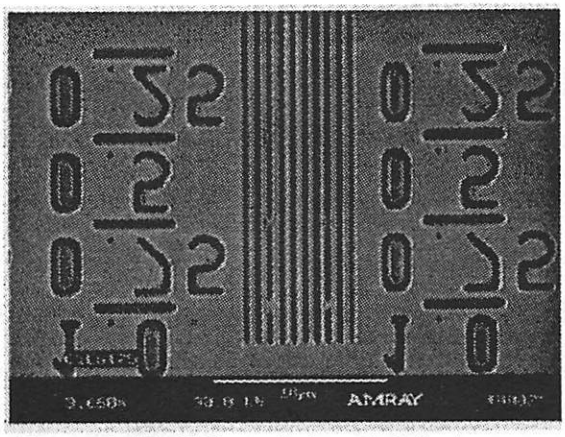


Figure 4-17c. Effect of pinspot defects on an array (C1EG175). Note the bridging.

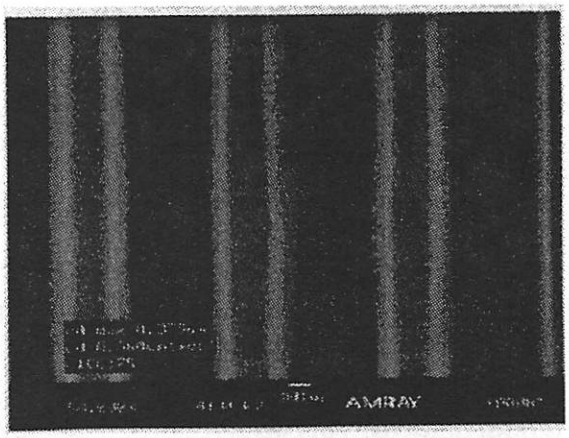


Figure 4-17d. Magnified image of a pinspot impacting an array (C1EG175).

Figure 4-17. SEM's showing a defect on a mask and showing the impact of a phase shift defect on patterning an attenuating phase shift array.

SEM's of experimental results were taken for a quartz bump, a pinspot, and a pinhole defect that are approximately 0.5μm square with the results shown in Figures 4-17b, 4-17c, and 4-17d. In 4-17b, the developed pattern is shown for the quartz bump protrusion (P1EG175), and this figure shows that the defect is just large enough to cause the array pattern to bridge. Experimental results for protruding pinspot defects on edge of an array line (C1EG175) are shown in Figures 4-17c and d. In Figure 4-17c, the developed resist pattern

is shown for pinspot defects of several sizes with defects to the left being rectangular defects and with defects to the right being square defects. This figure agrees with the simulation result that line width variation is proportional to area as rectangular pinspots impact the array more than square pinspots. A magnified image of Figure 4-17c for the impact 0.5μm square pinspot defect is shown in Figure 4-17d. From Figures 4-17b and d, the experimental results agree with the previous simulation results of Section 4.4.2 where Section 4.4.2 showed that a quartz bump defect impacts the feature more then a pinspot defect.

The measured CD of an array line with a quartz bump on edge (P1EG175) is shown in Figure 4-18. Experimental results are shown for smaller defects because larger defects cause significant bridging. The measured line width for a line without a defect is shown as a control with ±10% CD variation tolerances shown as dotted lines. From the figure the small 0.28×0.27μm quartz bump defect remained within the 10% tolerance; however, the larger 0.66×0.73μm and 0.33×1.70μm quartz bumps cause the most variation. The 0.66×0.73μm quartz bump produced the most line width variation when the focus was at its best; which in this case is -0.3μm. This result is consistent with the previous simulation results because the 170° quartz bump which is close to a 180° bump should cause the most line width variation when near best focus. The 0.66×0.73μm quartz bump, however, produced more line width variation when at -0.9μm and +0.3μm defocus then at -0.6μm and 0.0μm defocus. This disagrees with the previous simulation result that line width variation for a 170° quartz bump should decrease when going out of focus. This result may be due to statistical uncertainty in the line width measurement in which more sample measurements would produce more reliable results. The experimental results are compared to the extended perturbational model for the 0.66×0.73μm quartz bump. The line width of the simulated array without a defect is also shown for comparison. The simulation without a defect predicts a line width that is approximately 35nm greater than the experimental result. This increased line width predicted by simulation may be due to vector electromagnetic scattering and to resist amplification. In the scalar perturbational model, neither vector electromagnetic scattering effects nor resist amplification is taken into account. The perturbational model for the 0.66×0.73μm defect predicts a line width that is approxi-

mately 50nm greater at 0.0, -0.3, and -0.6μm. This increase in line width is approximately the same increase predicted by the simulation for an array without a defect.

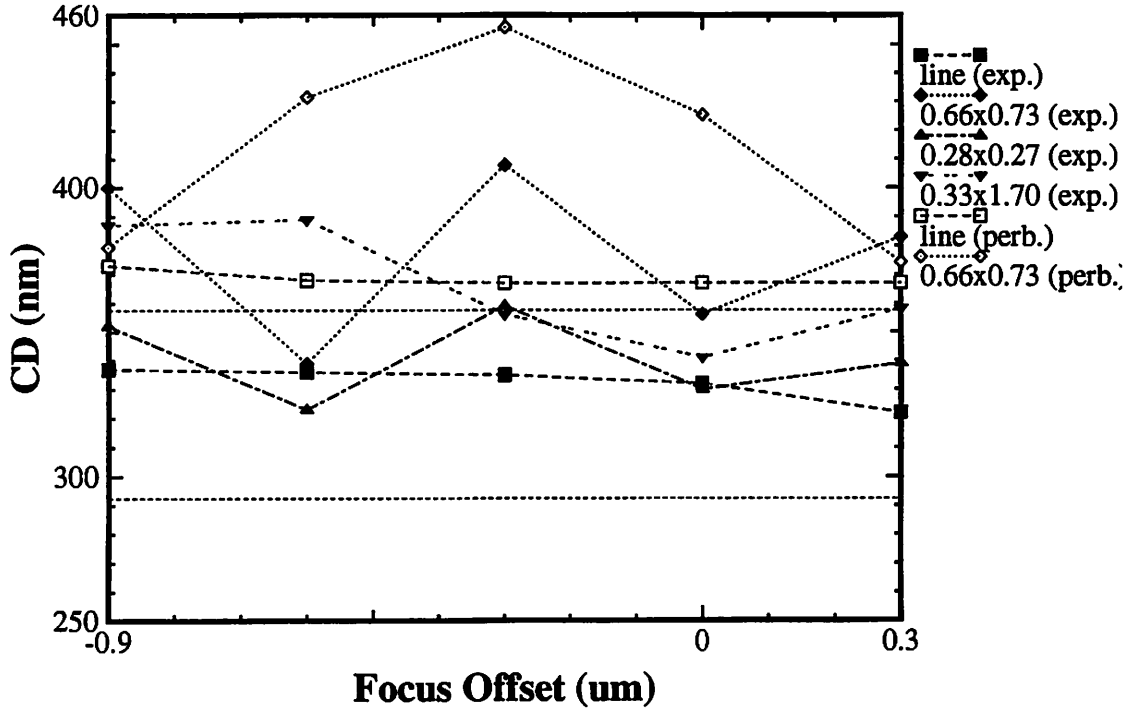


Figure 4-18. Measured CD for a quartz bump defect impacting an array line (P1EG175). The experimental results (exp.) are compared to the perturbational model (perb.).

Figure 4-19 shows the measured CD for an array line with a pinspot defect on edge (C1EG175). From the figure the smaller pinspot defects of 0.42×0.57μm and of 0.23×2.01μm cause line width variation that are within the 10% CD tolerance. The larger 0.48×2.02μm pinspot, however, causes significant line width variation. From the simulations in the previous sections, the 0.48×2.02μm pinspot should have the greatest impact while focussed. The experimental results for the 0.48×2.02μm pinspot, however, does not show the trend that the line width should steadily decrease when out of focus. This discrepancy is also seen in Figure 4-18 for the 0.66×0.73μm quartz bump, and more experimental samples may improve the experimental accuracy. Also from comparing Figures 4-18 and 4-19, the quartz bump defect causes more line width variation. For example, the smaller 0.66×0.73μm quartz bump (0.48μm^2 area) begins to cause 10% variation while a pinspot of 0.48×2.02μm (0.97μm^2 area) is needed to cause 10% variation. The perturbational

model was used to generate results for an array without a defect and for an array with a 0.42×0.57μm pinspot defect on edge of the line. As in the quartz bump results, the perturbational model for the array without a defect and for the 0.42×0.57μm pinspot predict a slightly greater line width than observed in experiment, and this increase may be due to resist effects or measurement uncertainty. The perturbational model and experiment, however, agree in predicting the trends that the 0.42×0.57μm pinspot have on line width variation.

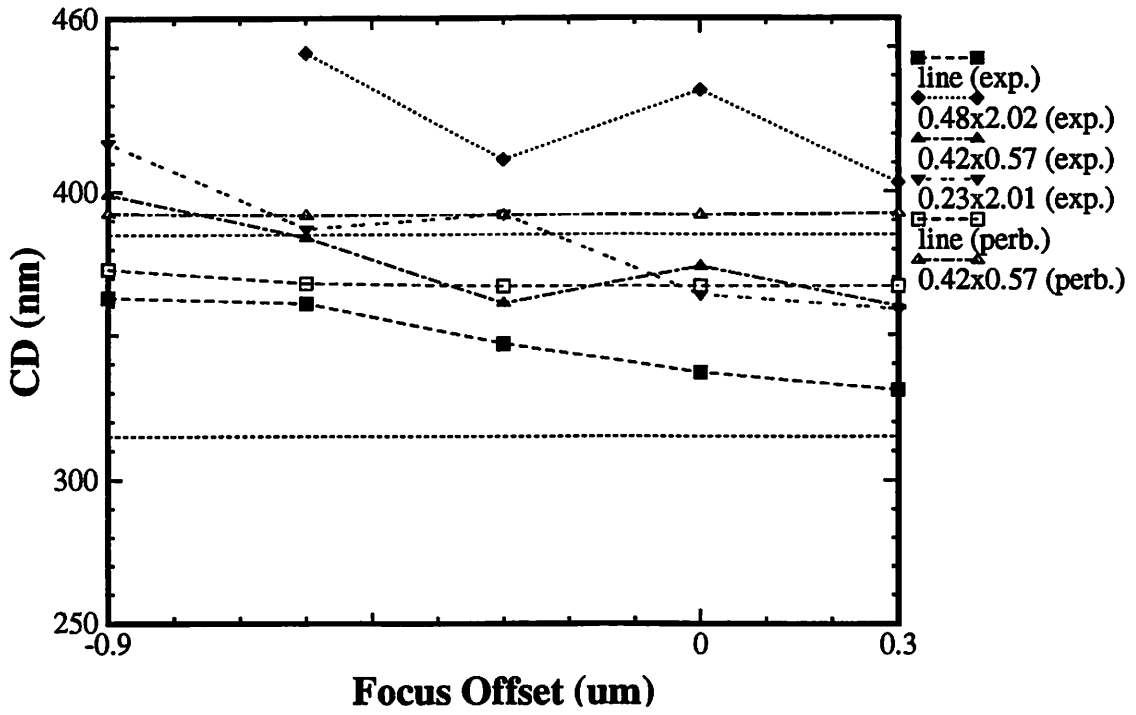


Figure 4-19. Measured CD for a pinspot defect impacting an array line (C1EG175).

The experimental results (exp.) are compared to the perturbational model (perb.).

In Figure 4-20, the CD variation caused by a fully etched pinhole defect on an array is shown. Figure 4-20 shows that attenuating PSM arrays are more tolerant to pinhole defects as all the defect sizes cause less than 10% CD variation. Once again comparing Figure 4-20 with Figure 4-18, the 0.62×0.70μm pinhole causes less variation than the 0.66×0.73μm quartz bump. This result agrees with previous simulation results in which it was seen that the quartz bump defect causes more feature variation than the pinhole defect. The effect of a 0.62×0.70μm pinhole defect on an array was simulated with the perturba-

tional model and is shown in Figure 4-20 for comparison. The simulated line width is approximately 40nm greater than the experimental results for an array without a defect and for an array with a 0.62×0.70μm pinhole. These simulation results, however, show that the 0.62×0.70μm pinhole produces little line width variation which is consistent with the experimental trends.

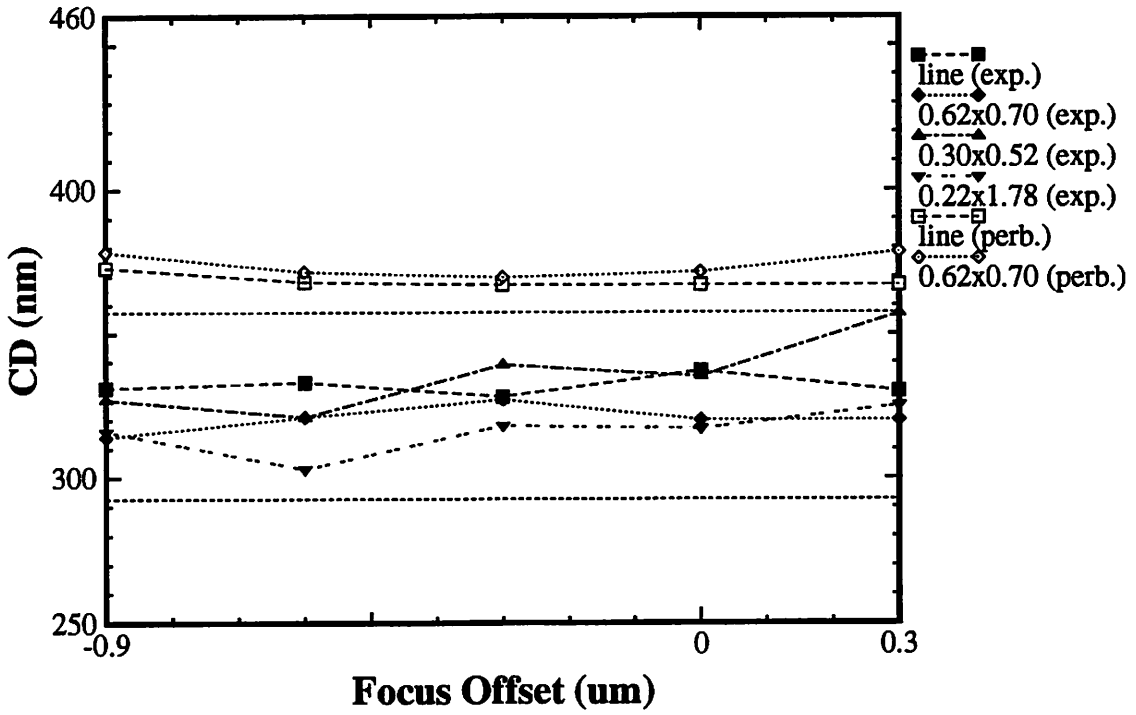


Figure 4-20. Measured CD for a fully etched pinhole defect impacting an array line (M1EG175).

The experimental results (exp.) are compared to the perturbational model (perb.).

4.7. Conclusions

In this chapter, both direct simulation and the perturbational model have been used to determine guidelines for defect printability. These guidelines include focus as well as illumination parameters, defect size, defect location, defect type and feature type. Because of the many factors involved in defect printability, the perturbational model can give physical insight and allow the effects of as many as 9 mask and 3 stepper variables to be rapidly assessed. The perturbational model in Section 4.2 calculates algebraically the effect of the defect on a feature image by considering the electric field passing through the defect to be

a perturbation on the field from the feature. The perturbational model separates the influence of the 9 mask parameters and the influence of the 3 stepper variables. These mask parameters include feature transmittance and phase, defect size, transmittance, and phase, and the background transmittance and phase. The effect of the stepper parameters on defect printing influence the mutual coherence function (MCF).

The MCF is a weighting of the electric field from the defect and of the electric field from the feature and was calculated in Section 4.3. The effect of modified illumination, of high numerical aperture, and of thin-film interference on defect printing is included in the MCF. Because the MCF is able to consider many stepper parameters, the MCF gives physical insight into defect printing as it determines how much the defect image will perturb the feature image. Using the algebraic model for linewidth perturbation and the MCF, the impact of a defect on a feature can be rapidly assessed through parameters which are independent of the mask and may be calculated and stored prior to mask inspection. These parameters include the image transform function, the mutual coherence function at the object plane, and the desired image of the feature. Therefore, only the MCF at the image plane and the image of the defect need to be calculated during inspection.

Using the MCF in the perturbational model, the type of illumination can reduce the impact a defect has on a feature. In annular and quadrupole illumination, the center part of the source, which is coherent, is removed. The quadrupole illumination source examined in this chapter blocks illumination on the horizontal and vertical axes; consequently, when a defect is oriented along an axis where the source is blocked, the MCF between a defect and a feature is reduced. This reduction of coherence in quadrupole and annular sources causes the MCF to decrease. Although the MCF is less for quadrupole and annular illumination than for conventional illumination, Section 4.5 showed that defect printing depends strongly on the line edge slope of the image intensity. Depending on this line edge slope, some illumination sources are more defect tolerant. For example, it was shown in Section 4.5 that quadrupole illumination is more defect tolerant for dense arrays because of its improved line edge slope.

In Section 4.4, the perturbational model was used to determine the trends in defect impact on attenuating phase-shift lines, arrays, and contact holes. These trends, determined by the perturbational model in Section 4.4, were compared to trends found through experiment in Section 4.6. In Section 4.4, the perturbational model was shown to provide more than adequate accuracy in determining design rules for defect inspection and in determining the trends observed in experiments. The experiments as well as perturbational model results show that the quartz bump defect has more impact on line width variation than the pinspot and pinhole defects. Due to mask fabrication tolerances and noise in line width measurements through focus, the experimental results were difficult to interpret for quartz bump defects. The experiment results, however, did not exceed the line width offset between the experimental array results and the results predicted by the perturbational model. In the case of chrome pinspot and pinhole defects the perturbational model was able to predict the impact of these defects on an array.

The scalar image theory of Chapter 3 used to develop the perturbational model in this chapter will be used again in Chapter 6 to develop a technique for simulating vector polarized partial coherent light. Before considering the partial coherence, Chapter 5 presents TEMPEST which simulates scattering from topography of vector polarized light that is fully coherent. In Chapter 5, TEMPEST results are presented for simulating the gate patterning process over an active area well.

5 Simulation of Vector Coherent Light Scattering from Topography

5.1. Introduction

As the wavelength used in lithography becomes shorter, the reflectivity of silicon increases. This increase in reflectivity causes the topography to scatter additional light into unexposed areas. The effect is particularly noticeable where positive-tone resist lines must cross concave upward features. This reflected light may cause what is known as reflective notching, where concentrated light causes linewidth narrowing [142]. This reflected notching problem has been observed experimentally and has been corrected by using anti-reflective coatings [84,86,31] and by increasing the absorption of the photoresist [55,127]. Dyed photoresist is the least costly processing solution and shows some reduction of reflective notching; however, it has been shown that dyed photoresist reduces resolution as well. A TiN anti-reflective coating (ARC) has been shown to eliminate notching, but is the most costly processing solution since another layer is deposited and may cause heavy metal contamination. A related phenomena also mitigated by dyed resist and ARC is the variation in energy coupling characterized by the swing curve. While topography can produce resist thickness variations which affect energy coupling and linewidth, the reflective notching problem, which involves lateral reflection effects as well as vertical coupling effects, can be far more devastating. In this chapter, the three-dimensional simulation with TEMPEST is used to gain insight into the effectiveness of dyed photoresist and into the effectiveness of anti-reflective coatings in suppressing reflective notching from a three-dimensional substrate feature.

This chapter begins by briefly summarizing TEMPEST, which was described earlier in Chapter 2. TEMPEST is used throughout this dissertation to simulate scattering

from topography. In collaboration with Myron Cagan and Zoran Krivokapic of Advanced Micro Devices (AMD), we demonstrated the accuracy of TEMPEST in predicting reflective notching from an active area well. The TEMPEST results presented here were compared to those obtained by Cagan at AMD. These results were presented at *Optical/Laser Microlithography VIII* Conference in 1995 and published in its Proceedings [112]. The active area well structure studied in this collaboration is described in Section 5.3 and is formed by a LOCOS process [151]. In the LOCOS process, a field oxide step is formed through thermal oxidation. After the oxidation, a conformal polysilicon layer is then deposited. In Section 5.4, experimental results are presented that show reflective notching caused by light reflecting off this step covered in polysilicon. Experimental results of using a TiN anti-reflective coating and of adding dye to the resist are also investigated in Section 5.4. In Section 5.5, experimental results are compared to three-dimensional TEMPEST simulation results.

5.2. Finite Difference Time Domain Electromagnetic Simulation: TEMPEST

The simulation of reflective notching from wafer topography has been examined through various methods. These methods can be divided into frequency domain-simulation techniques and time domain-simulation techniques. By using a frequency domain-simulation technique, Matsuzawa[75] was one of the first to simulate the image inside photoresist above a conducting substrate. Matsuzawa's approach employs a finite element method to solve the frequency domain integral equations. Urbach and Bernard[132] extended this frequency domain-simulation approach to find notching from non-conducting substrates. Yeung[158] also used a frequency domain approach to simulate reflective notching by using a multipole method to solve the frequency domain integral equations. In time domain-simulation techniques, a finite difference time domain (FDTD) method is used to simulate reflective notching. In a FDTD method, Maxwell's equations are solved at discrete points on a cubic staggered grid, where the excitation field is monochromatic and fully coherent, i.e., electric and magnetic fields add. In order to study problems of interest in photolithography, Guerrieri[46] formulated and Gamelin[36] implemented a two-dimension FDTD program called TEMPEST[153] on a connection machine. Tadros[123]

used this two-dimensional version of TEMPEST to study reflective notching, and verified these simulations with experiment. TEMPEST was extended to three dimensions by Wong[152], and the speed and memory efficiency was improved by Pistor[101]. In addition to this finite difference method, Wojcik[149] applied a finite-element method to solve Maxwell's equations in the time domain.

TEMPEST-3D was used here to study reflective notching from an active area well. In these TEMPEST-3D simulations, the active area was modeled with six planes rather than by the true experimentally observed topography in order to gain some simple physical insight. This topography structure is then excited by a plane wave normal to the topography. The amplitude of the plane wave excitation is modulated by the intensity distribution calculated by SPLAT. In this chapter, the TEMPEST-3D simulations were based on the assumption that the light of this SPLAT intensity-modulated excitation is fully coherent. In the next chapter, Chapter 6, a method is developed that allows the partial coherence of vector polarized light to be simulated. This method is then used in Chapter 7 to explain the effect of partial coherence in reflective notching. These TEMPEST-3D simulations were run on a single-program multiple-data supercomputer called the Connection Machine 5 (CM-5). The CM-5 used in the simulations of this chapter is located at the National Center for Supercomputing Applications (NCSA). This CM-5 has 512 processors and 4G bytes of total memory and is capable of operating at speeds of 20G flops.

In these TEMPEST simulations on the CM-5, the electric field in the resist material is first calculated for each photoresist bleaching cycle. In a photoresist bleaching cycle, the intensity of the resist is first calculated and then the absorption constant is updated with Dill's A, B, C model. After updating the absorption constant, the fields in the next bleaching cycle are simulated. A total of five bleaching cycles was simulated on the CM-5, where each of these bleaching cycle simulations requires approximately 10 minutes of cpu time on 128 of 512 processors and needs approximately 2G bytes of memory, which is half the total memory of the CM-5 machine. After completing all five bleaching cycles, the photoactive compound (PAC) concentration inside the resist is calculated from the electric field with Dill's A,B,C model. Once the PAC concentration is found, other simulation tools such as SAMPLE can perform a post-exposure bake (PEB) and develop the resist.

5.3. Topography of the Gate Patterning Process

The active area-topography structure examined here was formed through a LOCOS process. First a pad oxide was grown, and then a layer of silicon nitride was deposited and patterned with an active area mask. The field oxide was then grown using LOCOS to a thickness of 0.27μm. The lateral encroachment was approximately 0.3μm leading to a "bird's beak" whose upper surface is approximately a plane at an angle of 23°. Growing this field oxide for the structure under study resulted in an approximately rectangular active area of 1.15μm × 5μm. After growing the oxide, 0.25μm of polysilicon was deposited over the structure. One set of wafers was patterned with the as-deposited polysilicon and a second set received an additional 50nm of deposited TiN ARC layer. To pattern the gate, an experimental positive i-line resist was spun on to achieve a thickness of 1.0μm over the field region. In addition to examining the case of leaving the resist undyed, two concentrations of dye were added to the photoresist and are called "low dye" and "medium dye" resist in the following sections. A schematic of the active area topography is shown in Figure 5-1, where the topography has been approximated with planar surfaces. In this figure, a top view and a cross-section show the resulting topography, material parameters, and dimension. As shown in the top view, the two planar structures will form a 45° intersection in the corner of the active area well. The bird's beak at the sides and at the ends of the active area forms a 23° angle with respect to the substrate. The gate is then patterned in the middle of the active area well in order to achieve a 0.35μm line in a 5X i-line stepper with NA of 0.5 and a conventional illumination source with σ of 0.6. The resist was then developed and the resulting photoresist gate lines were examined with a SEM to determine the extent of reflective notching.

From optical ray tracing, some intuitive results can be gained by considering how a plane wave will reflect off the structure depicted in Figure 5-1. Using optical ray tracing, reflections from the corners, ends, and sides are shown in Figure 5-2. A plane wave of light incident on the ends and sides of the poly step will reflect at an angle of 46° with respect to the incident light; therefore, light will expose the middle of the active area at a height of 0.56μm, which is given by (active area width)/2tan(reflection angle) (=1.15/2tan46°). The

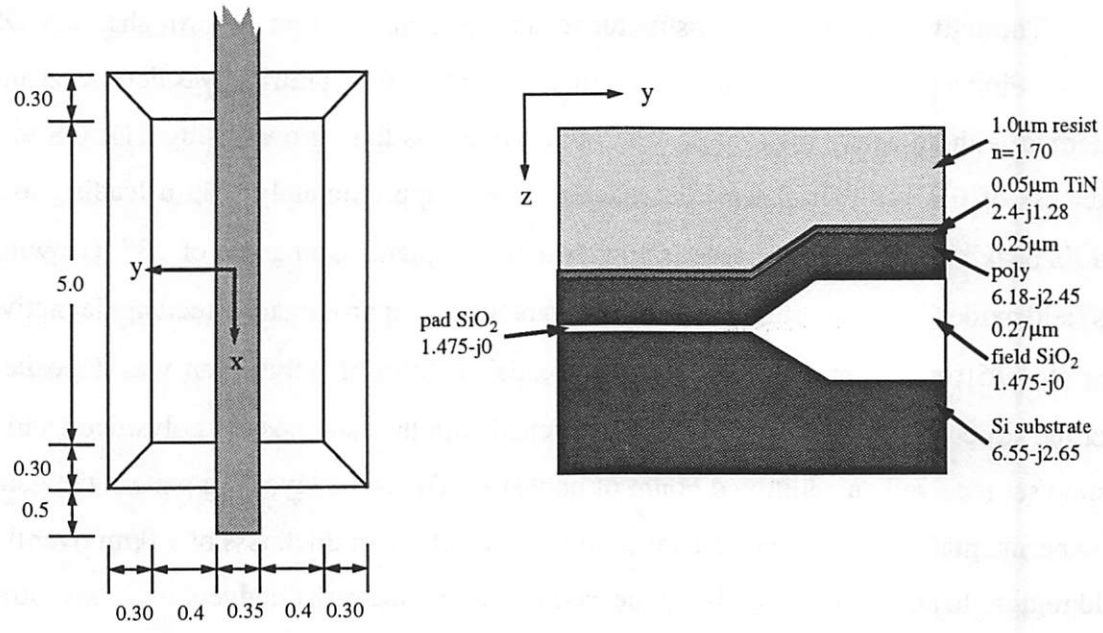


Figure 5-1. Model of the recessed active area with six planes.

The top view is shown on the right and a cross-sectional view of the topography on the right. All dimensions are in micrometers.

reflection of light from the ends and sides interfere, and the superposition of these reflected fields acts as a corner reflector. This light is reflected at an angle which is the bisector of the angle forming the corner that is 45°. Therefore, for the coordinate system defined in Figure 5-1, light is reflected at an angle of 45° in the xy plane and at an angle of 46° in the xz and yz planes. The summation of these reflections causes the greatest amount of light to be reflected into the middle of the active area at a height of 0.56μm in the active area well and at a distance of 0.575μm from the end of the active area (x=1.925μm). This is the location where the maximum notching should occur. The impact of these reflections can be lessened by reducing the reflectivity of the substrate or by increasing the absorption of the resist.

5.4. Reflective Notching - Experimental Results

In experiments done by Cagan of AMD, polysilicon gates were patterned over wafers with an active area well, as described in Section 5.3. SEM's were taken of these

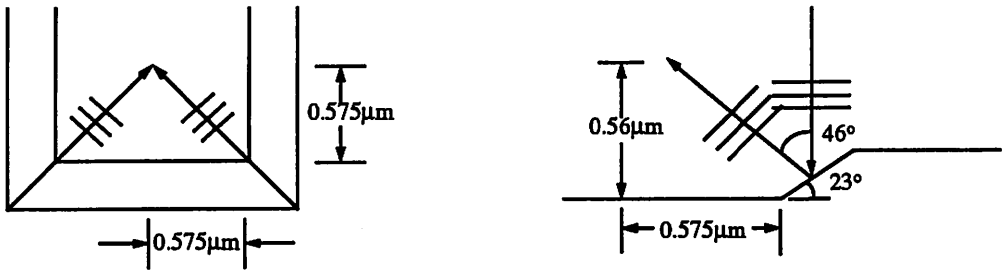


Figure 5-2. First-order model for calculating the location most susceptible to reflective notching through ray tracing.

This location is 0.575µm away from the well edge at a height of 0.56µm above the bottom of the well.

wafers with as-deposited polysilicon (on left) and with TiN (on right) for no dye, low dye, and medium dye added to the resist (see Figures 5-3, 5-4, and 5-5, respectively). In the SEM's of wafers with a TiN layer, the actual outline of the recessed active area can be clearly seen as a rounded bowl and differs from the modeled planar structure of Figure 5-1. In Figures 5-3, 5-4, and 5-5 the bright areas of the SEM's are from the sloping resist profiles; therefore, the width of the gate at the top of the resist is the dark area, and the width at the bottom of the gate is at the edge of the bright area. From these measured linewidths, the amount of slope in the resist line edge can be quantified in a slope parameter. This slope parameter is defined as line width at the top of the gate divided by the line width at the bottom of the gate. These widths can be measured from the SEM's.

In Figures 5-3, 5-4, and 5-5, the SEM's without a TiN layer show a very slight necking in the linewidth at the location predicted to be the most susceptible to notching, i.e., at an angle normal to the bowl formed by the LOCOS process. The SEM's also show that the deposited polysilicon layer appears slightly more grainy than the TiN layer, and this graininess may contribute to the necking of the resist line. In comparing the six SEM's, adding a TiN ARC layer clearly reduces the reflective notching and improves the resist line slope. In the three SEM's on wafers with as-deposited poly, there is considerable degradation of the resist line near the bird's beak. Adding dye to the resist shows some reduction of the reflective notching problem; however, the improvement from adding dye is considerably less apparent than the improvement in adding a TiN ARC.

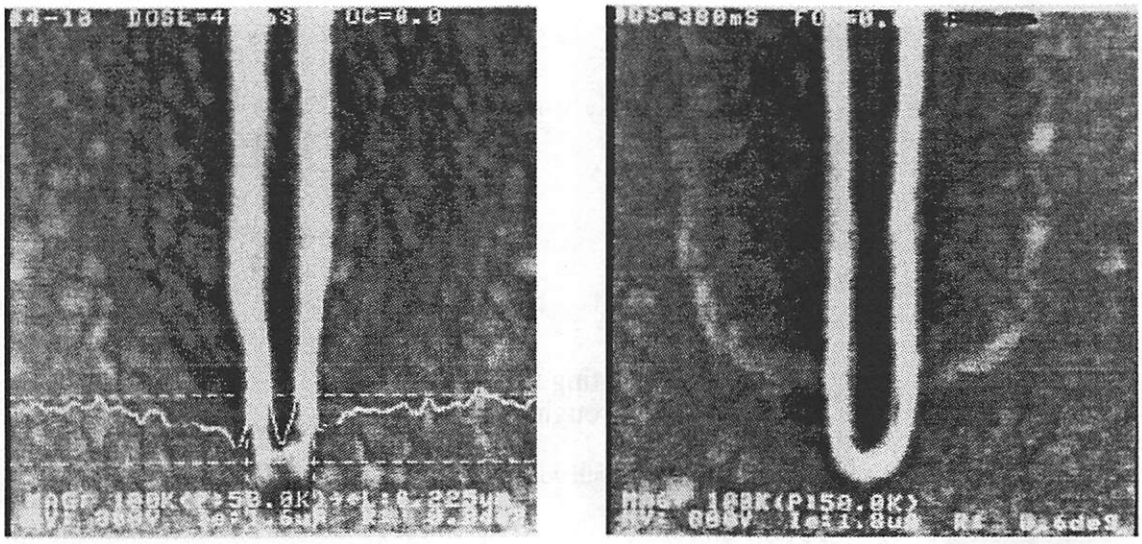


Figure 5-3. SEM pictures of photoresist gate lines without dye.

The SEM on the left is for a wafer without a TiN ARC and that on the right is for a wafer with a TiN ARC. In both pictures, no dye has been added to the photoresist. The TiN ARC removes the necking problem at the bird's beak.

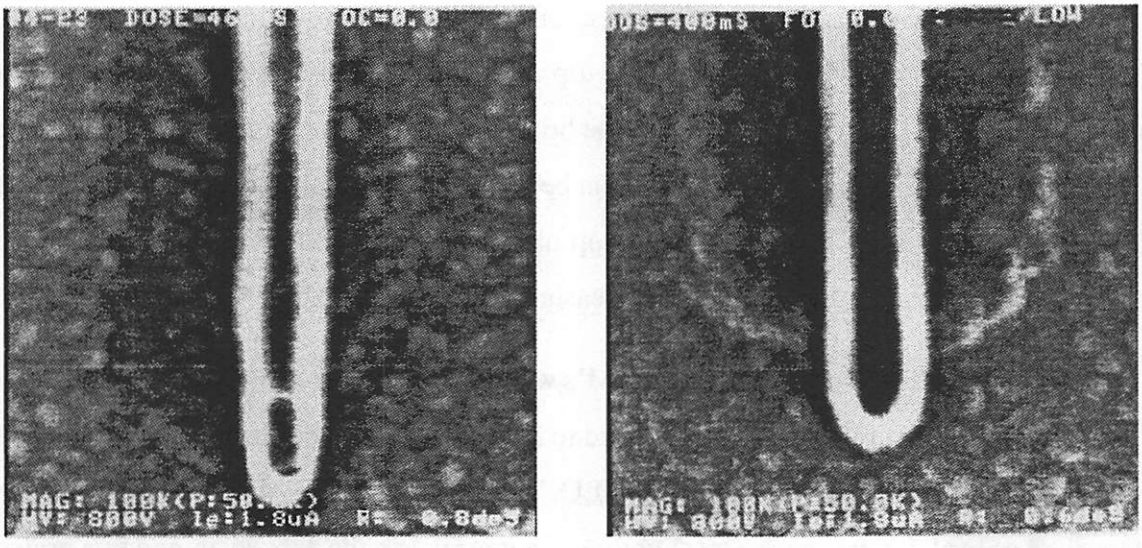


Figure 5-4. SEM pictures of photoresist gate lines with low dye.

The SEM on the left is for a wafer without a TiN ARC and that on the right is for a wafer with a TiN ARC. In both pictures, low dye has been added to the photoresist. The TiN ARC removes the necking problem at the bird's beak.

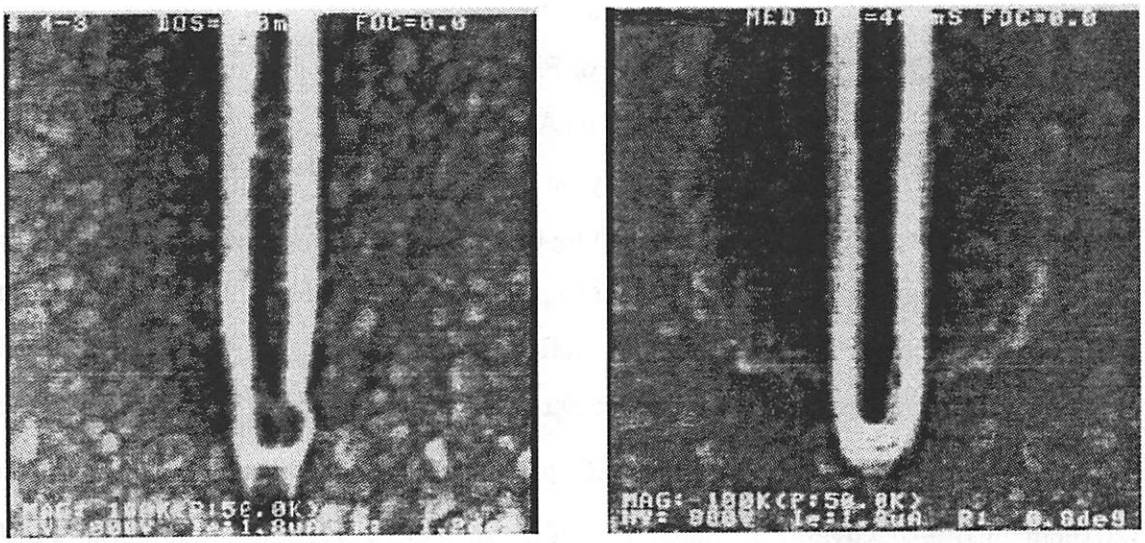


Figure 5-5. SEM pictures of the photoresist gate lines with medium dye.

The SEM on the left is for a wafer without a TiN ARC and that on the right is for a wafer with a TiN ARC. In both pictures, medium dye has been added to the photoresist. The TiN ARC removes the necking problem at the bird's beak.

In addition to reducing the risk of reflective notching, adding a TiN layer improves the resist line edge slope. In wafers without the TiN layer, the slope parameter was measured to be 0.30, 0.22, and 0.24 for the no dye, low dye and medium dye conditions, respectively; however, in wafers with TiN, the edge slope is improved and was measured as 0.55, 0.50, and 0.45 for the same respective conditions. In comparing these slope parameters, adding dye to the photoresist causes degradation of the resist line slope when a TiN ARC is present, and adding dye does not seem to improve the resist slope when a TiN ARC is omitted.

Although the TiN ARC helps prevent the reflective notching problem, the value of TiN as an effective ARC must be weighed against problems associated with its implementation. Using an TiN ARC increases the processing complexity as its use requires the additional steps of deposition over poly, TiN and poly etch, and TiN strip after poly etch. Not only is there an increase in processing complexity, but failure to completely remove the TiN may result in furnace contamination during subsequent poly dopant diffusion steps.

5.5. Reflective Notching: Simulation Results

Using TEMPEST, the topography of Figure 5-1 with the gate mask was simulated with as-deposited polysilicon and with a TiN ARC with no dye, low dye, and medium dye. To determine the effect of the topography on the resist line, PAC concentration profiles were examined at several cross sections along the yz plane at various positions along the x axis for the topography in Figure 5-1. These positions include the middle of the resist at x=0.0μm, the location most susceptible to reflective notching at x=1.925μm as predicted by ray tracing, and outside the well at x=3.0μm. As previously mentioned, the PAC was calculated with Dill's model with A,B,C parameters of $0.844\mu m^{-1}$, $0.068\mu m^{-1}$, and $0.016cm^2/mJ$, respectively, for undyed photoresist. After generating the PAC in TEMPEST, the PAC profiles underwent a post-exposure bake (PEB) in SAMPLE 2D with a diffusion length of 0.04μm. The profiles, however, were left undeveloped due to a lack of a good development model.

In Figure 5-6, PAC profiles are first shown for the recessed active area for wafers with as-deposited polysilicon and with undyed photoresist. In Figure 5-6, a planar substrate without a TiN layer (upper-left corner) is the desired profile. Working down from the planar substrate in the upper-left corner of Figure 5-6, PAC results are shown in the middle of the active area well (at x=0.0μm) and at the location most susceptible to notching (x=1.925μm). Continuing from the bottom left to the top right, cross-section cuts are shown in the well near the bird's beak (at x=2.4μm), along the sloping bird's beak (at x=2.7μm), and at the top of the well in the field-oxide region (at x=3.0μm). Examining the five PAC profiles when the topography is present, a wave is scattered off the field oxide step at an angle of approximately 46°, which consequently causes a sharply angled resist line edge. This degradation in the resist line-edge angle is in agreement with the ray tracing explanation given above in that a 46° angle is observed in the PAC profile. Also, as can be seen by comparing profiles, while moving from the center of the well (0.0μm) to the top of the field oxide (3.0μm), the PAC concentration increases and the line-edge slope improves. This indicates that the light scattered off the birds beak is notching the resist

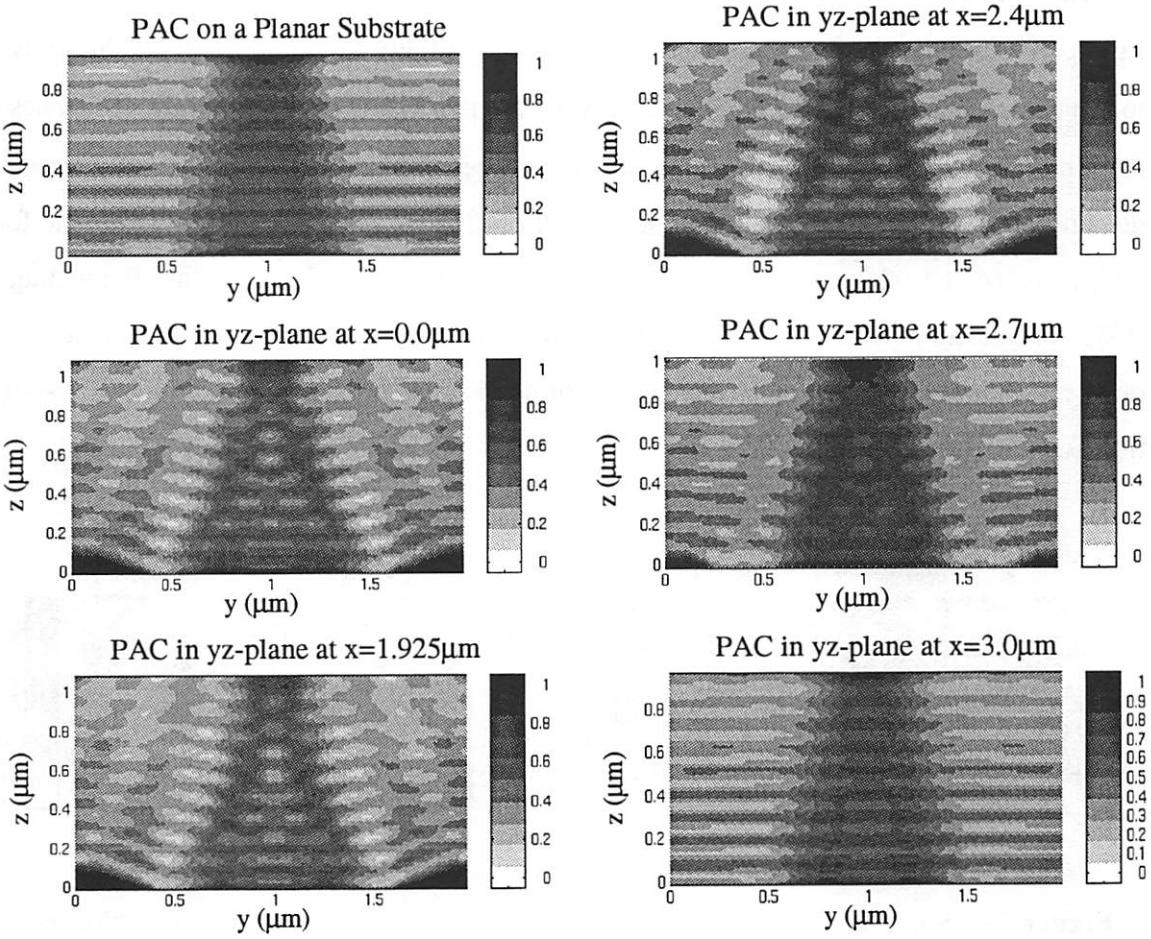


Figure 5-6. Simulated PAC concentrations after PEB on wafers with the as-deposited polysilicon and with undyed resist.

The simulations show profiles with poor edge slope and with reflective notching. The profile at x=0.0μm occurs in the middle of the active area in Figure 5-1 with the following profiles moving closer to the end of the active area. The profile with the least average PAC occurs at 1.925μm and is the location most susceptible to reflective notching. The profile at x=3.0μm is outside the active area on top of the field oxide.

PAC profile. As predicted by ray tracing, the cross section with the least average PAC and consequently the greatest notching seems to occur at 1.925μm.

Through simulation, the effect of adding dye to the photoresist was shown to be ineffective at reducing reflective notching and at improving the resist line edge slope. Figure 5-7 shows the PAC profiles when leaving the wafers with as-deposited polysilicon and when a low and medium concentration of dye is added to the resist. Adding dye to resist increases the B parameter of Dill's model to 0.19 and 0.335 in the low- and medium-

dye cases, respectively. The PAC profiles along the yz plane at x=1.925μm are shown in Figure 5-7 for low dye (left) and medium dye (right). Figure 5-7 indicates that adding dye to the resist does not change the line-edge slope, which is still approximately 46°. These PAC profiles agree with experiment, where it has been shown that adding dye to the resist does not change the edge slope. Comparing the profiles of Figure 5-7 with the profile for undyed resist at 1.925μm in Figure 5-6, adding dye to the resist does increase the average PAC concentration in the resist line. This increase may reduce the notching in the developed resist line; however, the PAC profiles for dyed resist are still poor, and a better solution is needed.

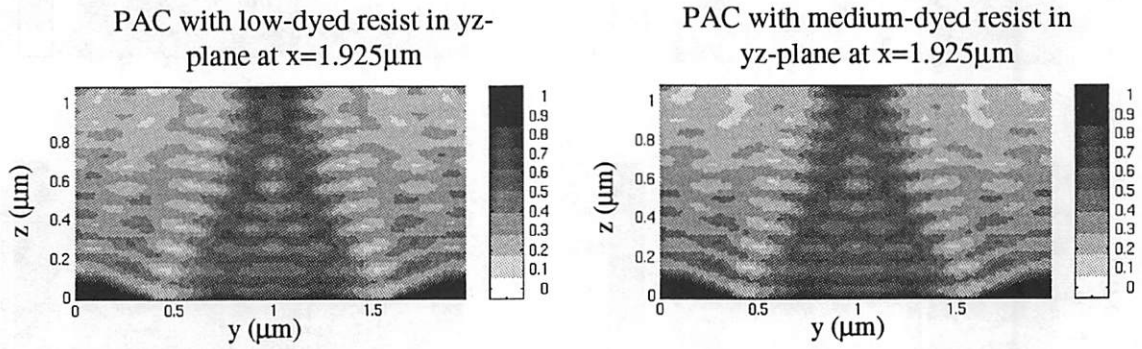


Figure 5-7. Simulated PAC concentration after PEB on wafers with the as-deposited polysilicon and with low-dye (left) and medium-dye (right) resist.

The profiles are located at the position most susceptible to notching. The dye is ineffective at reducing the reflective notching and at improving the resist line-edge slope.

Simulation results of adding a TiN ARC layer to the topography show that adding TiN improves the line edge-slope angle and eliminates the reflective notching problem. Figure 5-8 shows the case of undyed resist when a TiN ARC layer is added. Four PAC profiles along the yz plane are shown in Figure 5-8, beginning with a planar substrate in the upper-left corner, the PAC profile at x=0.0μm in the lower left, at x=1.925μm in the upper right, and at x=3.0μm in the lower right. Comparing the four profiles of Figure 5-8 with their corresponding profiles in Figure 5-6 for the as-deposited polysilicon, the line-edge angle improves and the average PAC in the resist line increases when a TiN layer is present. It should also be noted that the profiles with TiN in the well are in closer agreement with the desired planar substrate profile than the profiles without TiN are to the desired

planar substrate profile. These PAC profiles agree with experiment in showing that the deposition of a TiN ARC eliminates notching and improves the resist line-edge slope.

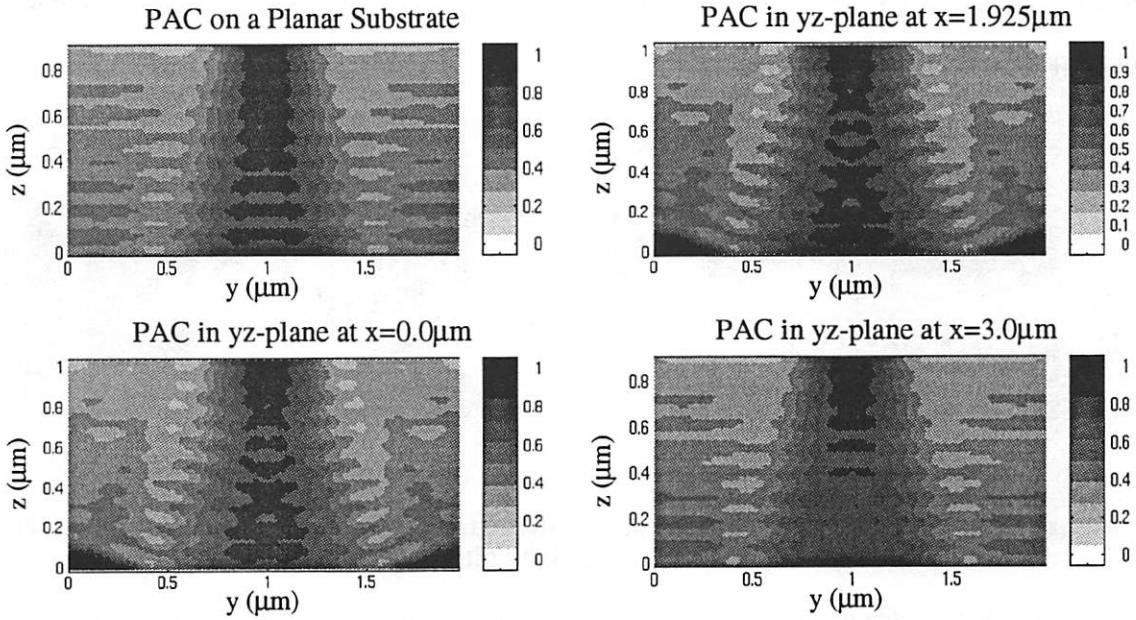


Figure 5-8. Simulated PAC concentrations after PEB on wafers with a TiN ARC layer on undyed photoresist.

The TiN layer improves the line-edge slope and eliminates the reflective notching problem.

Increasing the oxide slope of the bird's beak from 23° to 32° degrades the sharpness of the line-edge angle and increases the susceptibility of the resist line to reflective notching. As predicted by ray tracing in the 32° case, the location into which the most light will scatter into still occurs at x=1.925μm in the yz plane though at a different height in the resist, 0.28μm above the bottom of the well. In Figure 5-9, the PAC profiles for undyed resist are shown when the TiN is left undeposited (left) and when the TiN ARC is deposited (right). Examining and comparing Figure 5-9 to previous figures for an oxide slope of 23°, the incident light is now scattered off the oxide step at a greater angle, approximately 64°. This increased scattering angle now causes more light to be reflected toward the unexposed area and consequently causes a decrease in the steepness of the resist line-edge slope. The PAC profile simulated without a TiN ARC has less average PAC compound than the profile with the TiN, which again indicates that the TiN ARC layer decreases the risk of reflec-

tive notching, even at this steep angle. In comparing the 32° PAC profiles with the 23° profiles, the 32° profiles seem to have less average PAC concentration in the unexposed line area than the 23° profiles. This demonstrates that increasing the oxide slope increases the risk of reflective notching.

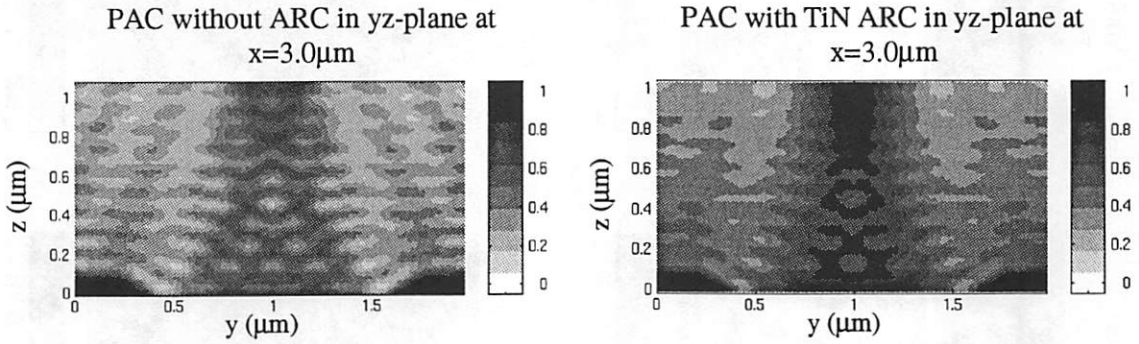


Figure 5-9. Simulated PAC concentrations after PEB on wafers with a 32° slope at the bird's beak in undyed photoresist.

The PAC profile on the left is for the as-deposited polysilicon wafer; the profile on the right is for the wafer with a TiN ARC. Even at this steep slope, the TiN layer improves the line-edge slope and eliminates the reflective notching problem.

5.6. Conclusions

Comparing the experimental results to simulation has demonstrated the effectiveness of TEMPEST in accurately predicting problems of reflective notching from a three-dimensional wafer topography. From the PAC profiles, TEMPEST simulations of the location most susceptible to notching and of the edge slope agree with experiment. Both simulation and experiment have shown that the use of dyed photoresist neither reduces the reflective notching nor improves the resist edge slope. However, both simulation and experiment have also demonstrated the effectiveness of using a TiN ARC layer to prevent reflective notching in three-dimensional topography structures. As materials become more reflective at shorter wavelengths and as feature dimensions continue to shrink, the ability to test a topography prior to costly experiments will become increasingly important. TEMPEST simulations can help locate problems and test solutions that eliminate scattering from the wafer topography. These include problems which are considerable, such as,

reflective notching from severe disk head topography; arbitrary 3D-shaped topography; and non-planar resist.

In this chapter, the TEMPEST simulations were based on the assumption that the light is fully coherent. In Chapter 6, a method is developed that allows the partial coherence of vector-polarized light to be simulated. This method is then used in Chapter 7 to explain the effect of partial coherence in imaging a phase-shift mask and in patterning a gate line over an active area well.

6 Theory of Vector Polarized Partial Coherent Light

6.1. Introduction

Since the electric and magnetic fields are vectors, the theory developed in Chapter 3 for the decomposition of scalar partial-coherent light is extended in this chapter to include the vector nature of the electric and magnetic fields. This extension is similar to the decomposition technique presented in Chapter 3. In Chapter 3, a matrix that represents the mutual intensity of the system was decomposed into a set of orthogonal eigenvalues and eigenfunctions. This chapter presents a method in which a matrix that represents the partial coherence of vector polarized light is decomposed into a set of orthogonal eigenvalues and eigenfunctions as well. Since this decomposition is based on diagonalizing a matrix, this technique is called the vector decomposition. Since transforms involving matrix diagonalization have the optimal energy compaction, this vector decomposition, like the scalar decomposition, is also optimal in an energy compaction sense. Furthermore, matrix diagonalization techniques by definition produce eigenfunctions that are coherent with themselves yet incoherent with the other eigenfunctions. These eigenfunctions represent the field excitations. Since each point in these field excitations is coherent with all the other points in the excitation, TEMPEST can be used to simulate the scattering of these excitations from a topographical structure.

Section 6.2 describes field polarization and the calculation of the image of a point source with transverse electric (TE) field polarization located in the object plane. Using the calculation of this field in the object plane, a coherence matrix is defined in Section 6.3, and the elements of this matrix are calculated in Section 6.4. The coherency matrix, for TE polarized light, describes the state of mutual coherence between the vector components of

the electric field for a given spatial location. In Section 6.5, this coherency matrix is then diagonalized into a set of orthogonal eigenvalues and eigenfunctions. These eigenfunctions represent the electric field excitations for TE polarized light. The accuracy of this vector decomposition is then compared to Abbe's formulation in Section 6.6. In Section 6.7, the vector decomposition and Abbe's formulation are used to find the aerial image of a contact hole in an attenuated phase-shift mask.

6.2. Polarization

The scalar theory of Chapter 3 is expanded here to include the effect of the field polarization. Any electromagnetic field can be represented as the sum of a transverse electric (TE) field and of a transverse magnetic (TM) field. The TE field is defined here to have no E_z vector component while the TM field is defined to have no H_z component. The TE fields and TM fields are defined in Equations 6-1 and 6-2, respectively. The direction of propagation, $\bar{k}$, is defined from the electric and magnetic fields as shown in Equation 6-3.

Equation 6-1. TE Fields

$$\bar{E}_{TE} = -\hat{x}\sin\phi + \hat{y}\cos\phi$$

$$\bar{H}_{TE} = -\hat{x}\cos\theta\cos\phi - \hat{y}\cos\theta\sin\phi + \hat{z}\cos\theta$$

Equation 6-2. TM Fields

$$\bar{H}_{TM} = x\sin\phi - y\cos\phi$$

$$\bar{E}_{TM} = -\hat{x}\cos\theta\cos\phi - \hat{y}\cos\theta\sin\phi + \hat{z}\cos\theta$$

Equation 6-3. Direction of Propagation

$$\bar{k} \propto \bar{S} = \bar{E} \times \bar{H} = \hat{x}\sin\theta\cos\phi + \hat{y}\sin\theta\sin\phi + \hat{z}\cos\theta$$

In Abbe's formulation, a point in the effective source is imaged as a plane wave onto the object. By definition, this effective source is located in the exit pupil of an illumination optic, where it is assumed to be fully incoherent. An effective source with TE polarization can be represented by a collection of point sources where each point in this effective source has TE polarization and can be represented mathematically by Equation 6-4. In Equation 6-4, x_p and y_p are the coordinates of the effective source located in the exit pupil and x_s and y_s are a sampled source point where ϕ is $\operatorname{atan}\frac{y_s}{x_s}$. The image from each of these

polarized points can be calculated by using the scalar diffraction theory on each vector component. This image under the Fresnel approximation was given by Equation 3-12 and is repeated as Equation 6-5, where x_0 and y_0 are the coordinates at the object plane. As dis-

Equation 6-4.
$$\bar{E}_s(\mathbf{x}_s) = (-\hat{x}\sin\phi + \hat{y}\cos\phi)\,\delta(x_p - x_s)\,\delta(y_p - y_s)$$

Equation 6-5.
$$\bar{E}(\mathbf{x}_o) = e^{\frac{i2\pi z}{\lambda}} \int_{-\infty}^{\infty} P(\mathbf{x}_p)\,\bar{E}_p(\mathbf{x}_p)\, e^{\frac{i2\pi}{\lambda}\mathbf{x}_p \bullet \mathbf{x}_o} d\mathbf{x}_p$$
$$= (-\hat{x}\sin\phi + \hat{y}\cos\phi)\,P(x_s, y_s)\, e^{\frac{i2\pi}{\lambda}(x_s x_o + y_s y_o + z_s z_o)}$$

cussed in Section 3.6.2, Equation 6-5 assumes that aberrated and large NA systems can be taken into account by using a modified pupil, *P*, as given in Equation 3-20 and repeated in Equation 6-6. The pupil, *P*, can be simplified in accord with the following assumption that

Equation 6-6.
$$P(x_p, y_p) = \left[\frac{1-NA^2(x_p^2+y_p^2)/M^2}{1-NA^2(x_p^2+y_p^2)}\right]^{1/4} exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p)\right]$$
$$\times exp\left[-i2\pi z\frac{1}{NA^2}\sqrt{1-NA^2(x_p^2+y_p^2)}\right] circ\left(\sqrt{x_p^2+y_p^2}\right)$$

magnification of the illumination optic is one (M=1), which implies that the high NA term (the first term in Equation 6-6) is 1. With this assumption, the magnitude of the pupil filter term, *P*, in Equation 6-5, equals one. With this simplification and after replacing cartesian coordinates of the source with spherical coordinates (θ and ϕ), Equation 6-5 becomes Equation 6-7.

Equation 6-7.
$$E(\mathbf{x}_o) = (-\hat{x}\sin\phi + \hat{y}\cos\phi)\, e^{i\frac{2\pi}{\lambda}(x\sin\theta\cos\phi + y\sin\theta\sin\phi + z\cos\theta)}$$

Equation 6-7 is the electric field incident on to the object plane that is generated by a point in the TE polarized source. As in Abbe's formulation, the source can be sampled at a number of points. Each of these points generates a field incident onto the object with polarization and phase shown in Equation 6-7. Since the source is spatially incoherent, the mutual intensity is non-zero only when $\mathbf{x}_{01}$ is not equal to $\mathbf{x}_{02}$. Consequently, the intensity incident onto the object is given by adding the intensities produced by each sample source point, i.e., the source is integrated over last.

6.3. The Coherence Matrix

Instead of integrating over the source last, the source can be integrated over first to produce a mutual intensity at the object plane. This mutual intensity can then be decomposed as was done in Chapter 3 for scalar light to produce a mathematical representation of the mutual intensity with better energy compaction. This can be done by first generalizing the scalar decomposition technique developed in Chapter 3 to include polarization of the electromagnetic field. Consider a TE field propagating (E_z=0), the state of partial coherence for the vector electric field that s given by the coherence matrix. The coherency matrix is defined in Equation 6-8 where Equation 6-8 uses the notation of Mandel and Wolf of Ref. [72]. With this notation, $\mathbf{J}(\mathbf{x}_{o1}, \mathbf{x}_{o2}, 0)$ is the coherence matrix for quasimonochromatic light and J_{xx}, J_{yy}, J_{xy}, and J_{yx} are the mutual intensity elements which represent the coupling between the vector components.

Equation 6-8.

$$\mathbf{J}(\mathbf{x}_{o1}, \mathbf{x}_{o2}, 0) = \begin{bmatrix} J_{xx}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) & J_{xy}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) \\ J_{yx}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) & J_{yy}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) \end{bmatrix} = \begin{bmatrix} \langle E_x(\mathbf{x}_{o1}) E_x^*(\mathbf{x}_{o2}) \rangle & \langle E_x(\mathbf{x}_{o1}) E_y^*(\mathbf{x}_{o2}) \rangle \\ \langle E_y(\mathbf{x}_{o1}) E_x^*(\mathbf{x}_{o2}) \rangle & \langle E_y(\mathbf{x}_{o1}) E_y^*(\mathbf{x}_{o2}) \rangle \end{bmatrix}$$

The coherence matrix is a measure of the correlation between the different vector components of the light, which in this case is the correlation between the electric field components for TE polarization. The diagonal elements are the mutual intensity for the x and y vector components and when $\mathbf{x}_{o1} = \mathbf{x}_{o2}$, these elements represent the intensity in the x or y components. The off-diagonal elements represent the cross-correlation between the two components. From the definition of the off-diagonal elements in Equation 6-8, J_{xy} and J_{yx} are complex conjugates of each other, i.e., $J_{yx} = J_{xy}^*$. This implies that the coherence matrix is hermitian and that the eigenvalues are real [134]. If $J_{xy} = J_{yx} = J_{yy} = 0$, the field is linear polarized in the x direction; if $J_{xy} = J_{yx} = J_{xx} = 0$, the field is linear polarized in y direction; and, if $J_{xy} = J_{yx} = 0$, the field is unpolarized. The elements of the coherence matrix are also related the Stoke's parameters. The Stoke's parameters are linear combinations of the elements of the coherence matrix as described in Ref. 72.

6.4. Calculation of the Mutual Intensity Elements of the Coherence Matrix, J_{xx}, J_{yy}, and J_{xy}

Using the definition of the coherence matrix of Equation 6-8 and the scalar theory developed in Chapter 3, the mutual intensity elements of the coherence matrix, J_{xx}, J_{yy}, J_{xy}, and J_{yx} can be calculated by integrating over the source first. The electric field with TE polarization is $\bar{E}_{TE} = -\hat{x}\sin\psi + \hat{y}\cos\psi$. For a TE polarized spatially-incoherent source, by using the definition of J_{xx} in Equation 6-8, the J_{xx} mutual intensity at the source is given by Equation 6-9. Similarly, by calculating J_{yy}, J_{xy}, and J_{yx}, the coherence matrix for this TE polarized-incoherent source results are generated as shown in Equation 6-10.

Equation 6-9.

$$J_{xx}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) = \langle E_x(\mathbf{x}_{o1}) E_x^*(\mathbf{x}_{o2}) \rangle = I_o \sin\psi_1 \sin\psi_2 \delta(\mathbf{x}_{o1} - \mathbf{x}_{o2})$$

$$= I_o \sin^2\psi \delta(\mathbf{x}_{o1} - \mathbf{x}_{o2})$$

Equation 6-10.

$$\mathbf{J}(\mathbf{x}_{o1}, \mathbf{x}_{o2}) = \begin{bmatrix} \sin^2\psi & -\cos\psi\sin\psi \\ -\cos\psi\sin\psi & \cos^2\psi \end{bmatrix} I_o \delta(\mathbf{x}_{o1} - \mathbf{x}_{o2})$$

Using scalar wave theory and Equation 6-5, the mutual intensity elements of the coherence matrix in the image plane, $\mathbf{J}(\mathbf{x}_{i1}, \mathbf{x}_{i2})$, can be calculated from the elements in the object plane, $\mathbf{J}(\mathbf{x}_{o1}, \mathbf{x}_{o2})$, where the object plane is the source in this case. The scalar wave theory can be used to evaluate all the elements of the coherence matrix provided that the matrix elements are found in a dielectric medium (in this case, air), which has the properties of being linear, isotropic, homogeneous, and non-dispersive [42]. For example, the J_{xx} component in the image plane is shown in Equation 6-11 and simplifies to Equation 6-12 for a circular source. In Equation 6-12, r_{12} and θ_{12} are the same as those used in

Equation 6-11.

$$J_{xx}(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \int dx_p I_o \sin^2\psi |P(\mathbf{x}_p)|^2 e^{-i2\pi(\mathbf{x}_{i1} - \mathbf{x}_{i2}) \bullet \mathbf{x}_p}$$

Equation 6-12.

$$J_{xx}(\mathbf{x}_{i1}, \mathbf{x}_{i2}) = \int_0^1 \int_0^{2\pi} w dw d\psi \sin^2\psi e^{-i2\pi r_{12} w \cos(\psi - \theta_{12})}$$

Equation 3-18. After substituting the complex representation for $\sin^2\psi$ into Equation 6-

12, Equation 6-12 becomes Equation 6-13. The integral of ψ can be evaluated by using the definition of the Bessel function shown in Equation 6-14. After using Equation 6-14,

Equation 6-13. $$J_{xx}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = \int_0^1\int_0^{2\pi} w\,dw\,d\psi\,(\frac{1}{2}-\frac{1}{4}e^{j2\phi}-\frac{1}{4}e^{-j2\phi})\,e^{-i2\pi r_{12}w\cos(\psi-\theta_{12})}$$

Equation 6-14. $$2\pi i^m e^{im\theta_{12}}J_m(u) = \int_0^{2\pi} e^{im\psi}e^{iu\cos(\psi-\theta_{12})}\,d\psi$$

Equation 6-13 becomes Equation 6-15. Since the integrals of $uJ_0(u)$ and $uJ_2(u)$ have closed-form solutions, Equation 6-15 can be simplified to Equation 6-16, which is J_{xx} in **Equation 6-15.**

$$J_{xx}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = \frac{1}{2}(2\pi)\int_0^1 J_0(2\pi w r_{12})\,w\,dw + \frac{1}{4}(2\pi)\,i^2\cos2\theta_{12}\int_0^1 J_2(2\pi w r_{12})\,w\,dw$$

Equation 6-16.

$$J_{xx}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = \frac{J_1(2\pi r_{12})}{2\pi r_{12}} - \frac{1}{2\pi^2 r_{12}^2}\cos2\theta_{12}\,[1-J_0(2\pi r_{12})-\pi r_{12}J_1(2\pi r_{12})]$$

the image plane. Using the procedure outlined above of using the complex representation for the trigonometric functions, the other elements of the coherency matrix, J_{yy}, J_{xy}, and J_{yx}, can be calculated, resulting in Equation 6-17 and Equation 6-18.

Equation 6-17.

$$J_{yy}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = \frac{J_1(2\pi r_{12})}{2\pi r_{12}} + \frac{1}{2\pi^2 r_{12}^2}\cos2\theta_{12}\,[1-J_0(2\pi r_{12})-\pi r_{12}J_1(2\pi r_{12})]$$

Equation 6-18.

$$J_{xy}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = J_{yx}(\mathbf{x}_{i1},\mathbf{x}_{i2}) = -\frac{1}{2\pi^2 r_{12}^2}\sin2\theta_{12}\,[1-J_0(2\pi r_{12})-\pi r_{12}J_1(2\pi r_{12})]$$

The mutual intensity elements of the coherency matrix given by Equations 6-16, 6-17, and 6-18 represent the correlation between the electric field components of the TE polarized light. The mutual coherence for optical system A, λ=365nm, NA=0.9, σ=1, between two points, $\mathbf{x}_1$ and $\mathbf{x}_2$, is plotted in Figure 6-1 versus $\mathbf{x}_1$, where $\mathbf{x}_2$ is held constant

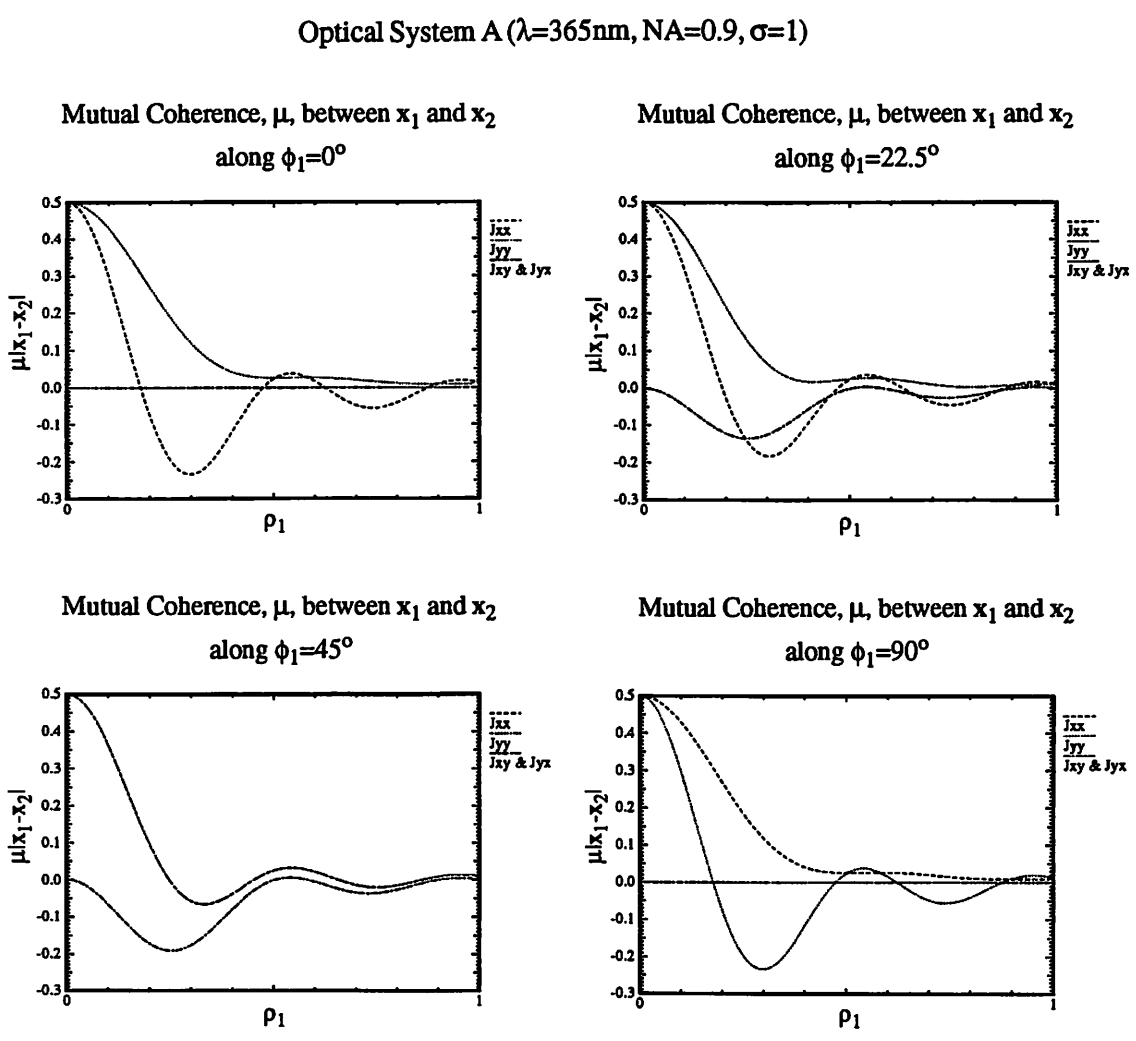


Figure 6-1. Plots J_{xx}, J_{yy}, and J_{xy} between $\mathbf{x}_1$ and $\mathbf{x}_2$ for four values of ϕ_1.

along the optical axis, i.e. $\mathbf{x}_2=(\rho_2=0; \phi_2=0^o)$, and where the radial component, ρ_1, of $\mathbf{x}_1$ is varied continuously from 0 to 1 and the azimuthal component, ϕ_1, is varied in discrete steps of 0^o, 22.5^o, 45^o, and 90^o, i.e. $\mathbf{x}_1=(\rho_1=0$ to 1; $\phi_1=0^o$, $\phi_1=22.5^o$, $\phi_1=45^o$, $\phi_1=90^o)$. From Figure 6-1, the mutual coherence of the J_{xx} element when $\phi_1=0^o$ is the same as the mutual coherence of the J_{yy} element when $\phi_1=90^o$. This occurs because the TE polarization, $\overline{E}_{TE} = \hat{x}\cos\psi - \hat{y}\sin\psi$, is symmetric with respect to $\psi=45^o$. Furthermore, J_{xx}, when $\phi_1=45^o$, equals J_{yy}, when $\phi_1=45^o$, where both J_{xx} and J_{yy} are Airy functions. The fact that

J_{xy} and J_{yx} are non-zero at angles other than multiples of 90^o implies that the matrix is polarized, in this case with TE polarization.

6.5. Diagonalization of Coherence Matrix

As described in Section 3.7.1 in which block matrices were used to decompose the mutual intensity, the eigenvalues and eigenfunctions of the coherency matrix given by Equation 6-8 are found by transforming $J(\mathbf{x}_1,\mathbf{x}_2)$ into a block matrix. Since the source is circular, the decomposition is performed in polar coordinates, where the location of point $\mathbf{x}_1$ and of point $\mathbf{x}_2$ is represented by the coordinates, (ρ_1,ϕ_1) and (ρ_2,ϕ_2), respectively. The block notation of the four-dimensional coherence matrix, $\mathbf{J}(i,j,k,l)$, in two dimensions is shown in Equation 6-19 , where i, j, k, l represent the variation in ρ_1, ϕ_1, ρ_2, and ϕ_2, respec-

Equation 6-19.

$$\mathbf{J}(i,j,k,l) = \begin{bmatrix} J_{xx}(1,1,1,1) & J_{xx}(1,1,2,1) & \cdots & J_{xx}(1,1,N,1) & \cdots & J_{xx}(1,1,N,N) & J_{xy}(1,1,1,1) & J_{xy}(1,1,2,1) & \cdots & J_{xy}(1,1,N,1) & \cdots & J_{xy}(1,1,N,N) \\ J_{xx}(2,1,1,1) & & & & & \vdots & J_{xy}(2,1,1,1) & & & & & \vdots \\ \vdots & & & & & & \vdots & & & & & \\ J_{xx}(N,1,1,1) & & & & & & J_{xy}(N,1,1,1) & & & & & \\ \vdots & & & & & & \vdots & & & & & \\ J_{xx}(N,N,1,1) & & \cdots & & & J_{xx}(N,N,N,N) & J_{xy}(N,N,1,1) & & \cdots & & & J_{xy}(N,N,N,N) \\ J_{yx}(1,1,1,1) & J_{yx}(1,1,2,1) & \cdots & J_{yx}(1,1,N,1) & \cdots & J_{yx}(1,1,N,N) & J_{yy}(1,1,1,1) & J_{yy}(1,1,2,1) & \cdots & J_{yy}(1,1,N,1) & \cdots & J_{yy}(1,1,N,N) \\ J_{yx}(2,1,1,1) & & & & & \vdots & J_{yy}(2,1,1,1) & & & & & \vdots \\ \vdots & & & & & & \vdots & & & & & \\ J_{yx}(N,1,1,1) & & & & & & J_{yy}(N,1,1,1) & & & & & \\ \vdots & & & & & & \vdots & & & & & \\ J_{yx}(N,N,1,1) & & \cdots & & & J_{yx}(N,N,N,N) & J_{yy}(N,N,1,1) & & \cdots & & & J_{yy}(N,N,N,N) \end{bmatrix}$$

tively. This is the same notation used in Equation 3-45 for the scalar mutual intensity, $J(i,j,k,l)$. The size of the block coherency matrix is $2N_\rho N_\phi$ x $2N_\rho N_\phi$, where N_ρ is the number of samples in ρ and N_ϕ is the number of ϕ samples. For example, if ρ and ϕ are sampled 25 times, the size-block coherency matrix is 1250×1250. Although this matrix is large, the matrix can be diagonalized using SuperLU developed by Demmel[30] in seconds. As shown in Equation 6-20, the eigenvector, $\Phi_n(\mathbf{x}_1)$, of the diagonalization is repre-

sented by a column vector and the eigenvector, $\Phi_n(x_2)$, is represented by a row vector .

Equation 6-20.

$$\Phi_n(x_2) = \left[\Phi_{xn}(1,1)\ \Phi_{xn}(2,1)\ \ldots\ \Phi_{xn}(N,1)\ \ldots\ \Phi_{xn}(N,N)\ \Phi_{yn}(1,1)\ \Phi_{yn}(2,1)\ \ldots\ \Phi_{yn}(N,1)\ \ldots\ \Phi_{yn}(N,N)\right]$$

These eigenvectors represent the electric field excitations and include polarization where the electric field in x is given by Φ_{xn} and the field in y is given by Φ_{yn}. Each excitation is coherent with itself yet incoherent or orthogonal with the other excitations.

The eigenvalues of the decomposition of the coherency matrix are plotted in Figure 6-2 for optical system A (λ=365nm,NA=0.9, σ=1) and optical system B (λ=248nm, NA=0.5, σ=0.5). Similar to the eigenvalues of the scalar mutual intensity, the eigenvalues of the coherency matrix are real and non-negative. This implies that the coherence matrix is Hermitian, i.e. the coherence depends only on the separation between two points. Furthermore, since the eigenvalues are non-negative, the intensity is non-negative. As seen in

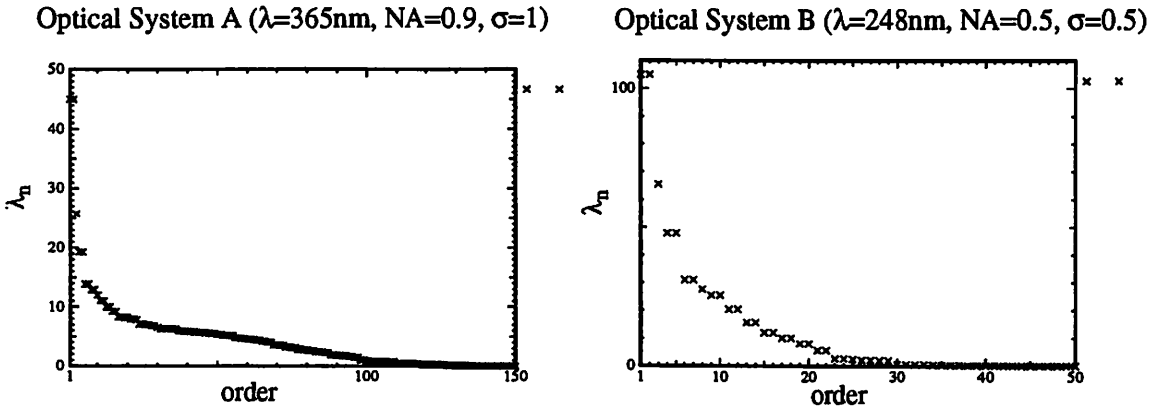


Figure 6-2. Plots of the eigenvalues for two different optical system.

The system on the left, optical system A, is an inspection system where λ=365nm, NA=0.9, and σ=1. The system on the right, optical system B, is a lithography system where λ=248nm, NA=0.5, and σ=0.5.

Figure 6-2, some of the eigenvalues are also degenerate. For example, Figure 6-2 indicates that the first and second orders in both optical systems are degenerate. This degeneracy of the first and second order is due to the polarization of the field and due to the symmetry of the source. The third eigenfunction is non-generate, while the fourth and fifth are once again degenerate. Similar to the eigenvalues of the scalar mutual intensity, the eigenvalues

descend with an approximately 1/n relationship, where n is the order. This indicates that the number of field excitations needed to represent the coherence can be truncated.

The eigenfunctions of the first three orders are plotted in Figure 6-3 and Figure 6-4 for optical system A and optical system B, respectively. These eigenfunctions represent the E_x and E_y field excitations for TE polarized light or represent the H_x and H_y field excitations for TM polarized light. Comparing the eigenfunctions of systems A and B, the eigenfunctions have the same functional representation in which the positions of the minima, the maxima, and zero crossings differ between the optical systems. The degenerate first and second orders are similar in that the first order E_x field is the transpose of the second-order E_y field, and due to symmetry, the first-order E_y is the transpose of the second order E_x. In fact, for all degenerate eigenfunctions, NA=0.9 and E_y are transposes of each other. In comparing Figure 6-3 to Figure 6-4, the eigenfunctions are similar in that they have the same functional relationship. The only difference between the eigenfunctions of Figure 6-3 and Figure 6-4 is the position of the maxima and minima. This is because the maxima and minima occur at different locations, where this location depends on $\frac{\lambda}{\sigma NA}$. For example, the first-order E_x eigenfunction of Figure 6-3 is similar to the first-order E_x eigenfunction of Figure 6-4, except that it is compressed.

These eigenfunctions form an orthogonal basis set in which each eigenfunction is incoherent with the other eigenfunctions in the set. The elements of the coherency matrix, J_{xx}, J_{yy}, and J_{xy}, are given by summing these eigenfunctions weighted by the eigenvalues. The elements, J_{xx}, J_{yy}, and J_{xy} are given in Equations 6-21, 6-22, and 6-23, respectively.

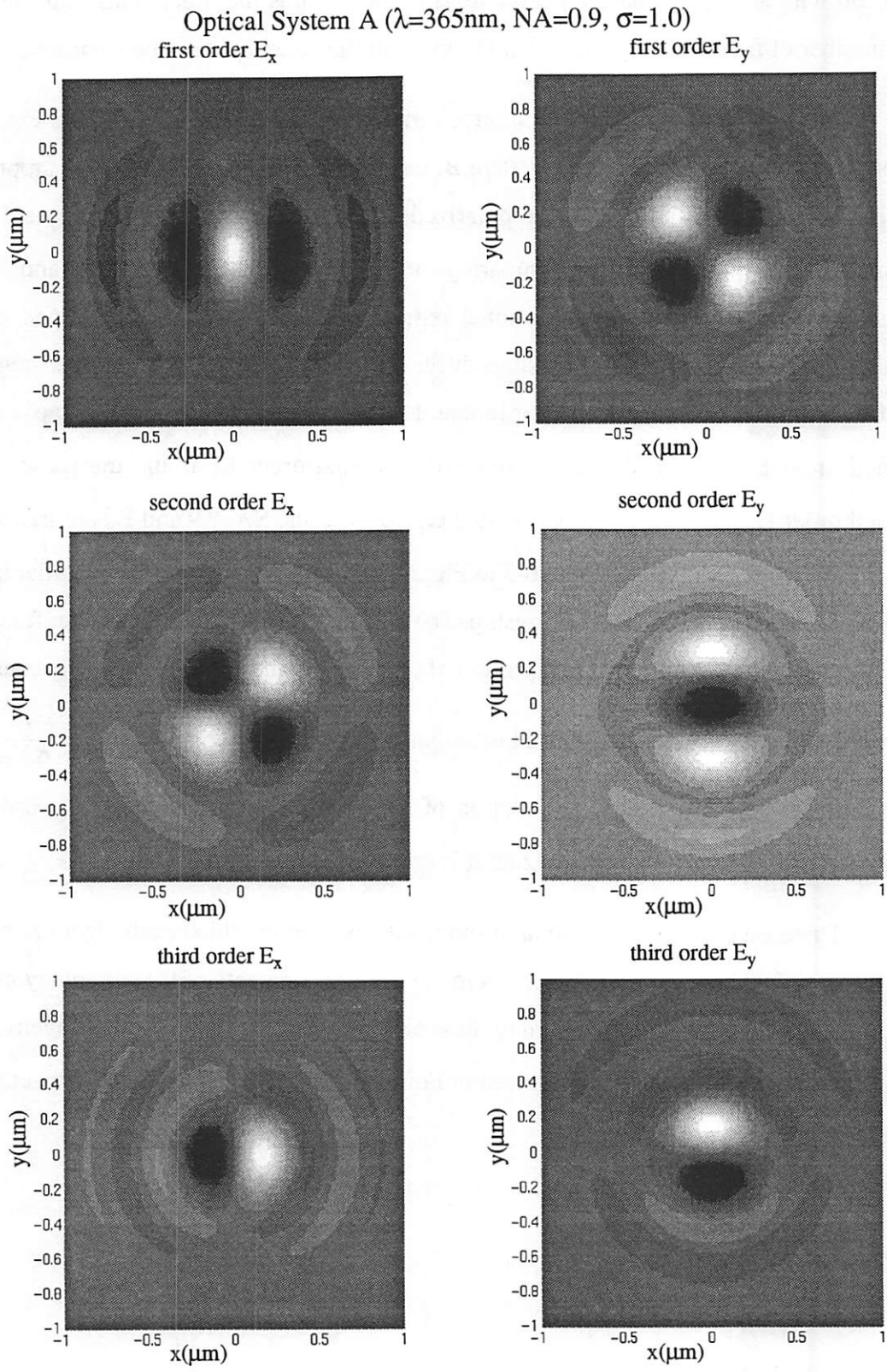


Figure 6-3. First three eigenfunctions of optical system A.

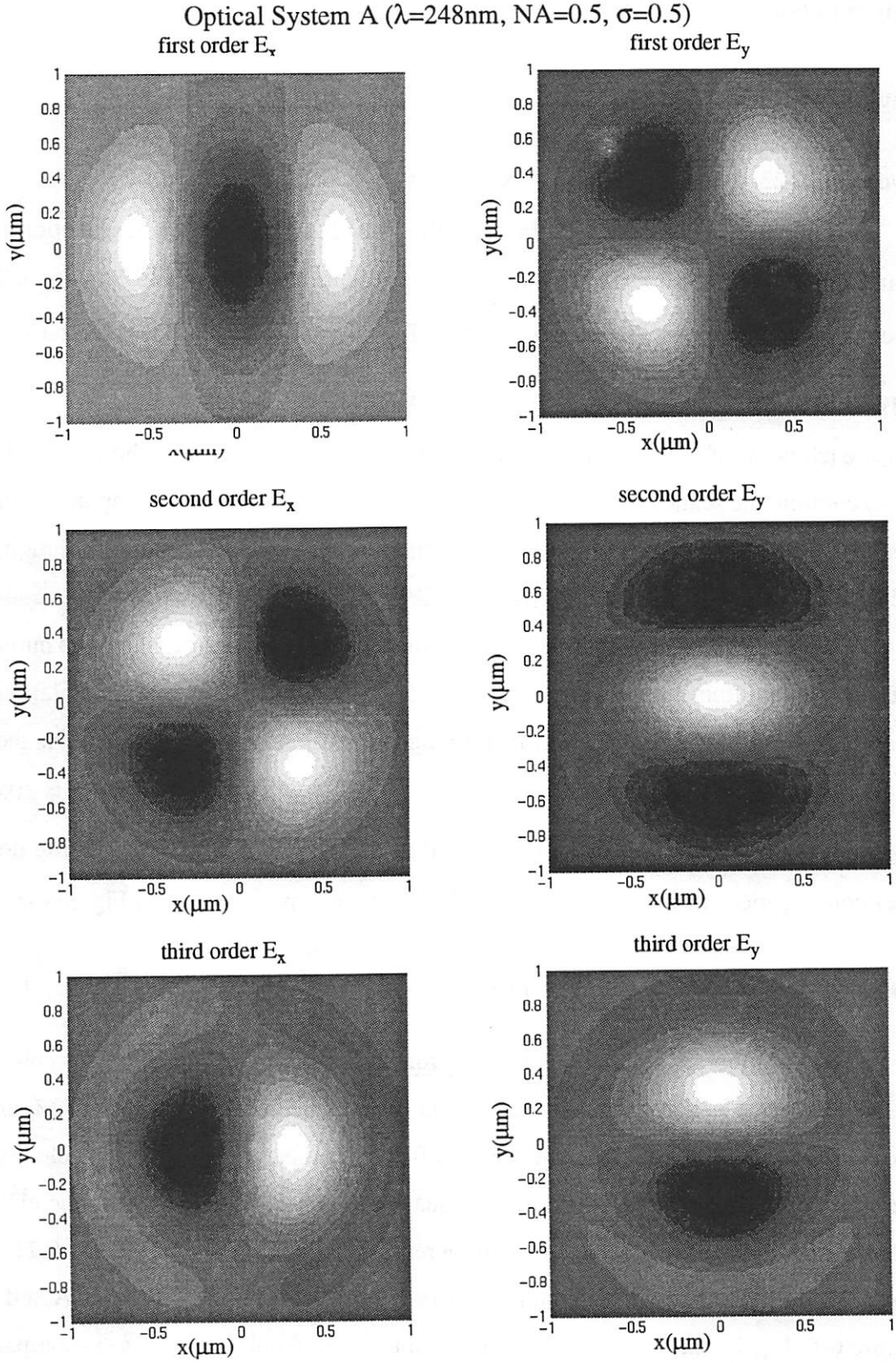


Figure 6-4. First three eigenfunctions of optical system B.

Equation 6-21. $$J_{xx}(\mathbf{x}_1, \mathbf{x}_2) = \sum_i \lambda_i E_{xi}(\mathbf{x}_1) E_{xi}^*(\mathbf{x}_2)$$

Equation 6-22. $$J_{yy}(\mathbf{x}_1, \mathbf{x}_2) = \sum_i \lambda_i E_{yi}(\mathbf{x}_1) E_{yi}^*(\mathbf{x}_2)$$

Equation 6-23. $$J_{xy}(\mathbf{x}_1, \mathbf{x}_2) = \sum_i \lambda_i E_{xi}(\mathbf{x}_1) E_{yi}^*(\mathbf{x}_2)$$

Since the eigenfunctions form an orthogonal basis set, each eigenfunction represents one simulation, where the eigenfunction represents the electric field excitation. These electric field excitations, $\bar{E}_i(\mathbf{x})$, are given by Equation 6-24 for the TE polarization case.

Equation 6-24. $$\bar{E}_i(\mathbf{x}) = \sqrt{\lambda_i}\,[\hat{x}E_{xi}(\rho, \phi) + \hat{y}E_{yi}(\rho, \phi)]\, e^{-jkz}$$

Since each point of these electric field excitations is fully coherent to the other points in the excitation, the scattering from these excitations incident upon a topographical structure can be simulated by TEMPEST. The implementation of this decomposition technique into TEMPEST is called TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition). The number of TEMPEST-PCD simulations needed is proportional to the numerical aperture of the illumination system, σNA, i.e., as σNA increases, more simulations are needed. Since the eigenfunctions form a orthogonal basis set, the total intensity is the sum of the intensities produced from the individual simulations. This total intensity is given by Equation 6-25, where $\bar{E}_j^a(\mathbf{x})$ is electric field in the TEMPEST-PCD simulation domain after convergence. This decomposition technique is shown pictorially in Figure 6-5.

Equation 6-25. $$I_{TOT} = \sum_{j=1}^{N} \bar{E}_j^a(\mathbf{x}) \bullet \bar{E}_j^a(\mathbf{x}) = \sum_{j=1}^{N} \left(\left|\bar{E}_{xj}^a(\mathbf{x})\right|^2 + \left|\bar{E}_{yj}^a(\mathbf{x})\right|^2 \right)$$

The number of simulations or eigenfunctions needed to represent the coherency matrix can be truncated once an error criteria is met since the eigenvalues of Figure 6-2 decay. The amount of this error can be quantified by comparing the analytical elements of the coherency matrix, as calculated with Equations 6-16, 6-17, and 6-18, to the elements calculated with the truncated decomposition representation given in Equations 6-21, 6-22, and 6-23. This comparison is demonstrated in Figure 6-6 for optical systems A and B. In Figure 6-6, J_{xx}, J_{yy}, and J_{xy} of the decomposition between points $\mathbf{x}_1$ and $\mathbf{x}_2$ is compared to the analytical solution as a function of $\mathbf{x}_1$, where $\mathbf{x}_2$ is located at the center of the domain

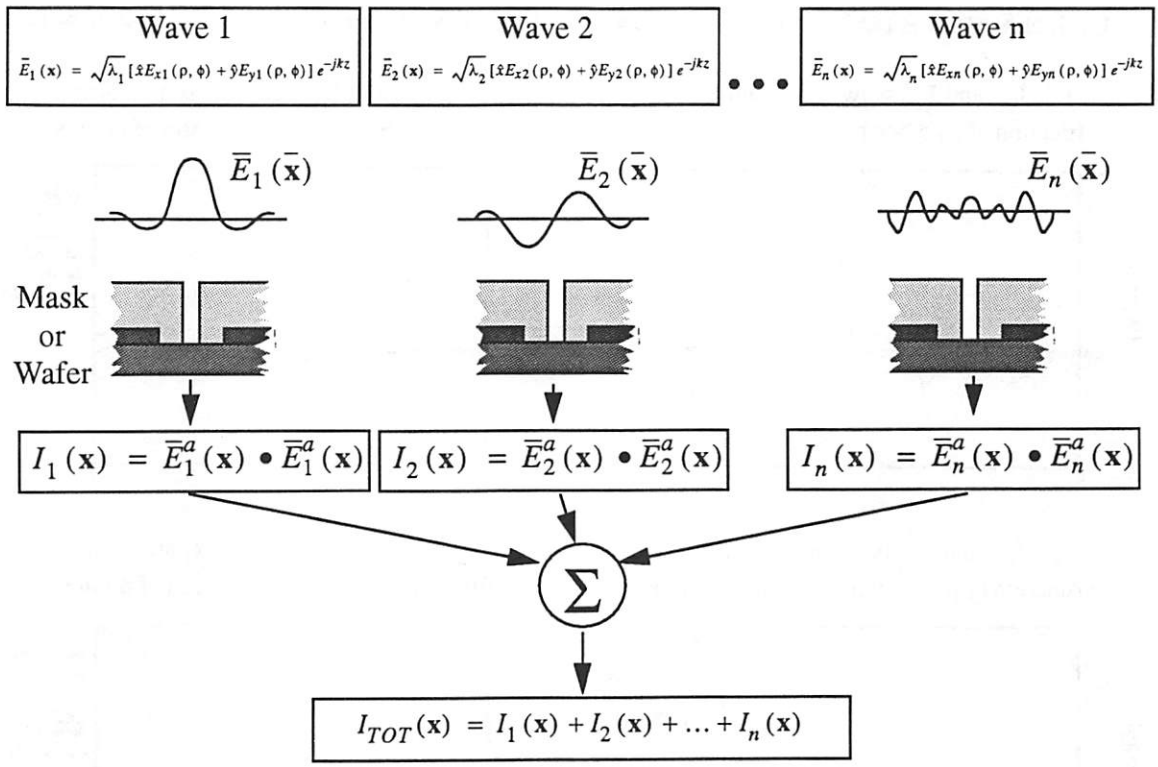


Figure 6-5. Pictorial representation of the decomposition formulation.

(ρ_2=0μm) and at the edge of the domain (ρ_2=1.0μm). The elements of the coherency matrix given by the decomposition technique are plotted when the decomposition sum is truncated after 93 eigenfunctions for optical system A and after 26 eigenfunctions for optical system B. The elements of the coherency matrix calculated after truncating the decomposition summation are in excellent agreement with the analytical solution at the center of the domain (ρ_2=0μm). The peak of the main lobe shows about 5% error between the analytical coherence and the decomposition coherence at the edge of the domain (ρ_2=1.0μm).

6.6. Comparison of the Decomposition Formulation versus Abbe's Formulation

In this section, the energy compaction and accuracy of the decomposition representation are compared versus Abbe's formulation for optical systems A and B. In Figure 6-7a and 6-7b, each element of the coherency matrix, J_{xx}, J_{yy}, and J_{xy}, is calculated with the decomposition formulation and with Abbe's formulation, and are then compared to the elements calculated analytically with Equations 6-16, 6-17, and 6-18. The elements are plot-

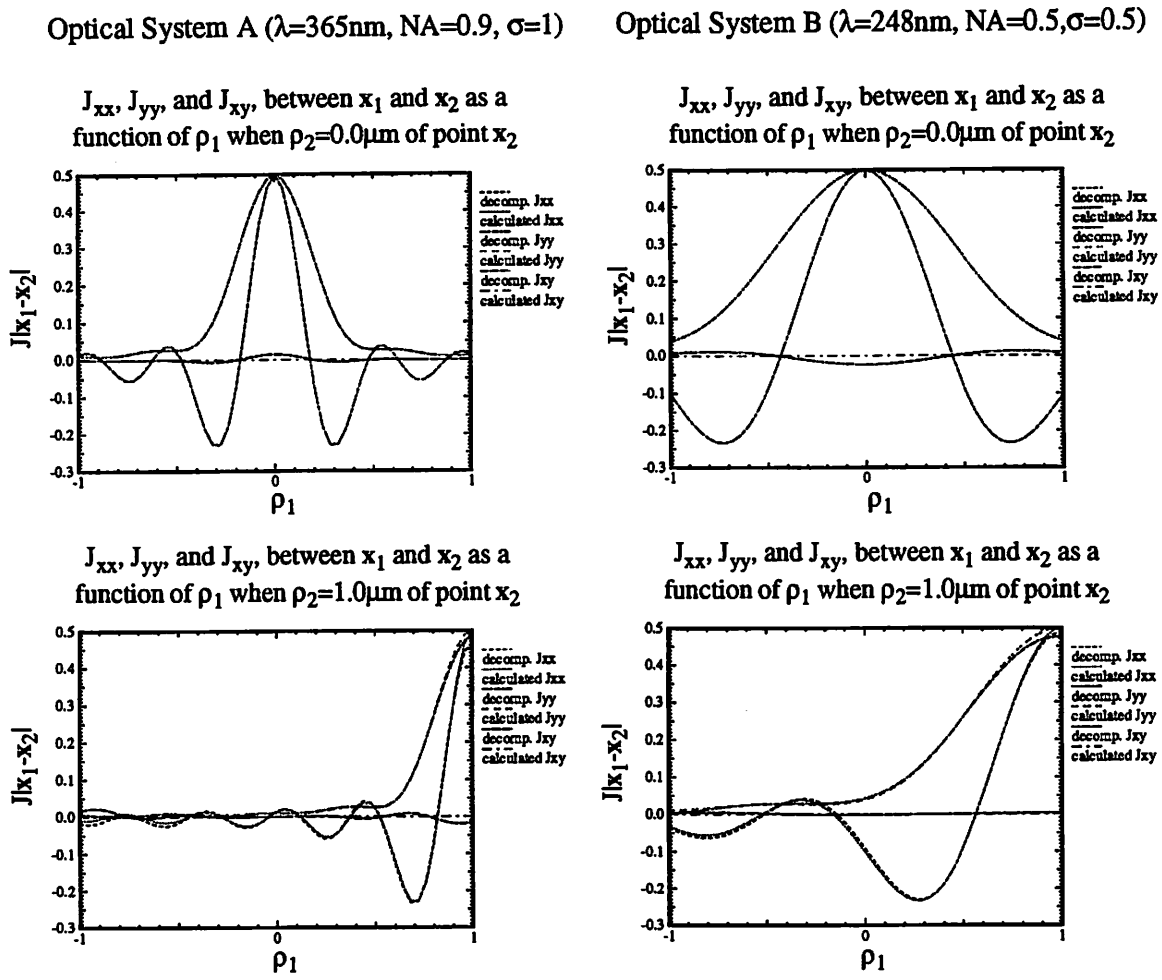


Figure 6-6. Comparison of J_{xx}, J_{yy}, and J_{xy} calculated with the decomposition formulation and calculated analytically with Equations 6-16, 6-17, and 6-18.

J_{xx}, J_{yy}, and J_{xy} between points x_1 and x_2 is compared for optical system A and B and is compared when ρ_2=0.0μm and when ρ_2=1.0μm.

ted as a function of ρ_1, where ϕ_1 of x_1 is fixed at 0^o and where ρ_2 and ϕ_2 of x_2 are both 0. For optical system A, the decomposition representation with 93 eigensolutions is in exact agreement with the analytical solution for all elements, J_{xx}, J_{xy}, and J_{yy}. For optical system B, the decomposition representation with 26 eigensolutions is again in exact agreement with the analytical solution. However, it should be noted that both optical system A and B required a few more eigensolutions than the decomposition on the scalar mutual intensity as the decomposition of the scalar mutual intensity required 82 and 19 eigensolutions for optical systems A and B, respectively. This increase in the number of eigensolutions is

most likely due to polarization causing more degenerate eigensolutions. In both optical systems A and B, the analytical solution is compared to Abbe's formulation when the illumination source is sampled with N_s=1, 2, and 4. As in the case of the scalar light of Chapter 3, the illumination source of optical system A was sampled 45, 249, and 1125 times when N_s=1, 2, and 4, respectively, and optical system B was sampled 9, 45, and 193 times. In both optical systems, the error difference between the analytical solution and Abbe's formulation decreases as the number of source samples increases. Abbe's formulation also shows more error in the side lobes than in the main lobe. The root mean square error for both optical systems is tabulated and shown in Table 6-1. Table 6-1 shows the error

Table 6-1. Root mean square error between the analytical mutual coherency elements, J_{xx}, J_{yy}, and J_{xy}, versus those calculated with the decomposition formulation and versus those calculated with Abbe's formulation.

Table 6-1 shows comparison results for optical system A (top) and B (bottom) at the center of the domain and at the edge of the domain. The coherency elements are calculated with Abbe's formulation when N_s is 1, 2, and 4.

Optical System A (λ=365nm, NA=0.9, σ=1.0)						
Mutual Coherence Representation	J_{xx} Center of the Domain (0μm)	J_{xx} Edge of the Domain (1μm)	J_{yy} Center of the Domain (0μm)	J_{yy} Edge of the Domain (1μm)	J_{xy} Center of the Domain (0μm)	J_{xy} Edge of the Domain (1μm)
Decomposition technique with 93 samples	0.01%	0.11%	0.00%	0.07%	0.07%	0.12%
Abbe's formulation with 45 samples	1.03%	1.03%	0.24%	0.24%	0.10%	0.10%
Abbe's formulation with 249 samples	0.40%	0.40%	0.11%	0.11%	0.10%	0.10%
Abbe's formulation with 1125 samples	0.15%	0.05%	0.05%	0.05%	0.10%	0.10%
Optical System B (λ=248nm, NA=0.5 , σ=0.5)						
Decomposition technique with 19 samples	0.00%	0.09%	0.00%	0.11%	0.02%	0.04%
Abbe's formulation with 9 samples	1.14%	1.14%	0.82%	0.82%	0.11%	0.11%
Abbe's formulation with 45 samples	0.45%	0.45%	0.39%	0.39%	0.11%	0.11%
Abbe's formulation with 193 samples	0.28%	0.28%	0.26%	0.26%	0.11%	0.11%

between the analytical solution for the coherency matrix elements and between both Abbe's formulation and the decomposition representation. The error is tabulated at the center of the domain (ρ_2=0, ϕ_2=0°) and at the edge of the domain (ρ_2=1.0μm, ϕ2=0°). From Table 6-1, the root mean square error at the center is less than at the edge in the decomposition method. Although the error of the decomposition method increases at the edge, the decomposition error is less than the error introduced with Abbe's formulation.

Optical System A (λ=365nm, NA=0.9, σ=1)

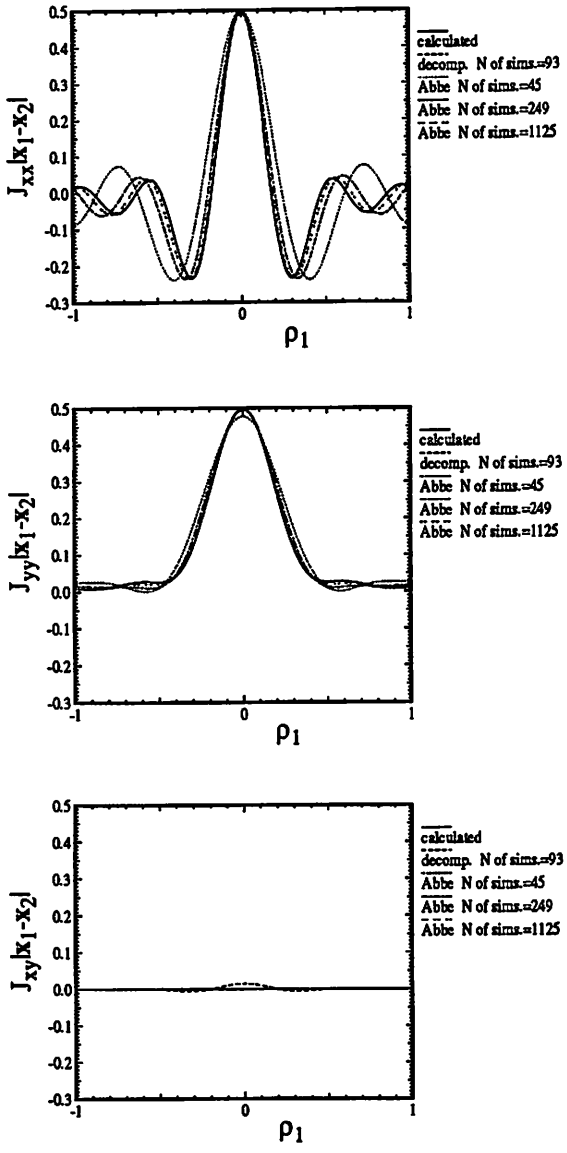


Figure 6-7a. Comparison of the elements of the coherency matrix, J_{xx}, J_{yy}, and J_{xy}, for optical system A (λ=365nm, NA=0.9, σ=1.0).

Optical System B (λ=248nm, NA=0.5, σ=0.5)

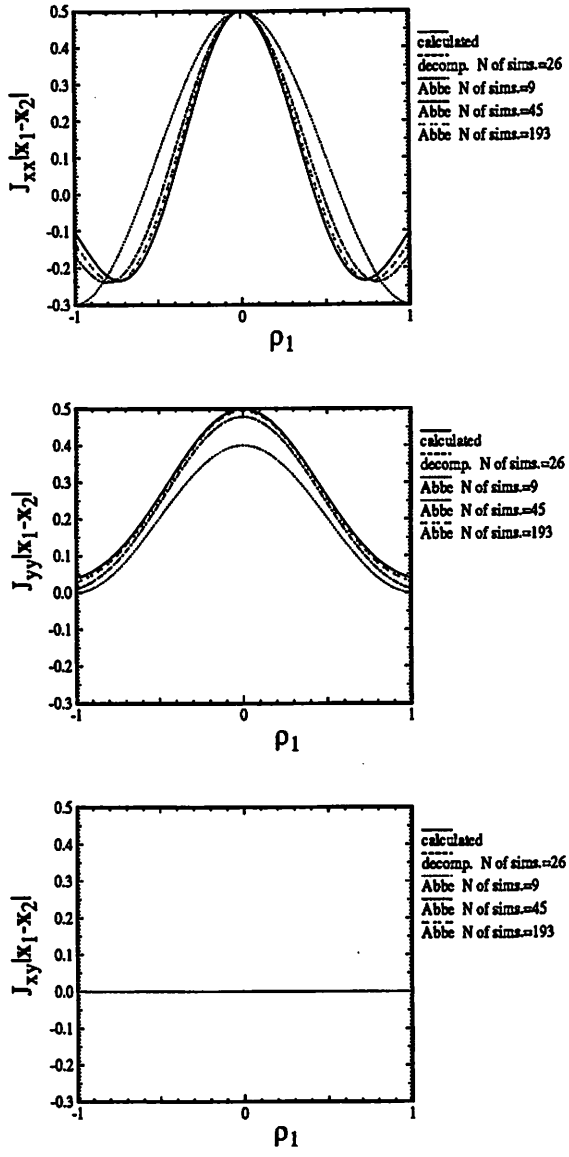


Figure 6-7b. Comparison of the elements of the coherency matrix, J_{xx}, J_{yy}, and J_{xy}, for optical system B (λ=248nm, NA=0.5, σ=0.5).

Figure 6-7. Comparison of the elements of the coherency matrix.

The elements calculated analytically with Equations 6-16, 6-17, and 6-18 to those calculated with the decomposition formulation and with Abbe's formulation when N_s is 1, 2, and 4. The comparison is shown for optical systems A (left) and B (right).

The number of simulations required for a 2% deviation of the partial coherence over a $6\lambda \times 6\lambda$ mathematical plane (a 3D-simulation domain) as a function of σ*NA is plotted in Figure 6-8 for a circular symmetric source. The number of simulations for the decomposition technique, as plotted in Figure 6-8, was found by truncating the eigenfunction summation. With this summation, the mutual coherence was calculated for two instances: the mutual coherence between a point in the center of the plane to all points across the entire $6\lambda \times 6\lambda$ plane, and the mutual coherence between a point at the edge of the plane (6λ) to the entire plane. When the mutual coherence of the eigenfunction summation in these two instances differed by less than 2% across the entire plane from the true mutual coherence, the eigenfunction summation was truncated. The number of eigenfunctions needed to represent this truncated summation is equivalent to the number of simulations. The number of simulations needed with Abbe's formulation was given by Equation 3-28 and is repeated in Equation 6-26 for a circular symmetric source, where N_S is the over-sam-

Equation 6-26.

$$n_{2D} = \frac{\pi}{4}\left(N_s \frac{2\sigma NAw}{\lambda}\right)^2$$

pling period. In this case, N_S of 4 was used because, as shown in Chapter of 3, N_S of 4 produced the most accurate mutual coherence. As observed in Figure 6-8, a comparison of the number of simulations needed with the decomposition technique to the number needed with Abbe's formulation indicates that the decomposition technique requires 5 to 10 times fewer simulations than Abbe's formulation. For example, for a stepper with NA=0.6 and σ=0.5, TEMPEST-PCD requires 22 simulations while Abbe's formulation would require 136 simulations for a 6.2 reduction in the number of simulations. For an inspection system with NA=0.9 and σ=1, Abbe's formulation would require 1350 simulations while TEMPEST-PCD requires 125 simulations, for a 10.8 $\times$ reduction.

Much of this computational reduction is due to the Karhunen-Loeve transform, which has the optimal energy compaction property. The Karhunen-Loeve transform represents the mutual intensity with fewer signal excitations as compared to the Fourier transform or Abbe's formulation. The optimal decomposition is in radial coordinates since the source is radially symmetric. This, as in the case of the scalar decomposition of Chapter 3, causes inaccuracies due to aliasing of applying a radially symmetric source to a rectangular

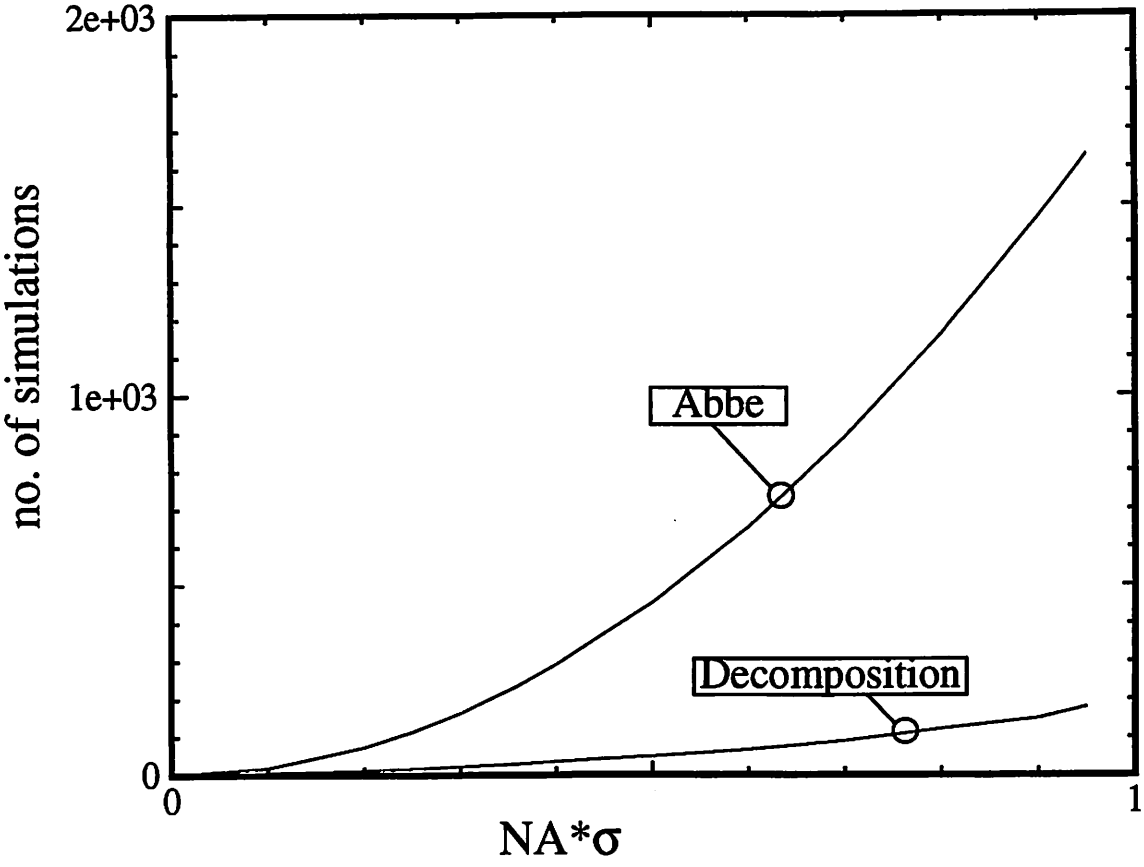


Figure 6-8. Number of simulations needed with the decomposition formulation and with Abbe's formulation.

The number of simulations was calculated in order to maintain 2% error across the simulation domain. The number of simulations with Abbe's formulation was found when N_s=4.

TEMPEST simulation domain. However, as demonstrated in Chapter 3, physics helps because the area of cross-talk from partially-coherent light scattering off the topography is often much smaller than the entire simulation domain. For example, in a reflective notching problem light scattering from the active area well is of interest and the mutual intensity need only be decomposed in a region that encompasses the active area well. Therefore, the partially coherent light should be decomposed in an area where the light scattering is significant.

6.7. Simulation Example: A 3D Hole in an Attenuated Phase-Shift Mask

In order to demonstrate the accuracy of the decomposition technique in three dimensions, the aerial image intensity produced by a contact hole in an attenuated phase-

shift mask was simulated using TEMPEST-PCD and using TEMPEST with Abbe's formulation. In this comparison example, the contact hole is square, $0.3 \times 0.3\mu m$. This hole is slightly over-etched in order to maintain a peak contrast that will cause a $0.25 \times 0.25\mu m$ hole after resist development. The hole is etched in an embedded attenuated phase-shift mask. The embedded attenuated phase-shift material is generated by the technique described by Mikami et al. [78]. In this technique, a 23nm chrome layer is first deposited on a quartz blank. This chrome layer causes most of the attenuation and provides some phase shift since its complex refractive index is 1.36–j1.91. After deposition of the chrome, a 55nm thick chromium fluoride layer is deposited onto the chrome. This chromium fluoride layer provides most of the 180^o phase shift. With this technique a phase-shift mask with 180^o phase shift and with 6.4% transmittance is generated. The $0.3 \times 0.3\mu m$ hole is then etched into the mask. A schematic of this mask is shown in Figure 6-9, which gives a cross-sectional view and a bottom-up view of the mask. The fields propagating through this mask are then simulated with TEMPEST using a 248nm stepper with NA of 0.7 and σ of 0.6. The diffracted orders are then calculated by taking the Fourier transform of the fields at a plane which is parallel to the mask and is at the bottom of the simulation domain. These orders are collected by the projection optic and imaged with SPLAT. The numerical aperture is chosen to be large in order to insure that off-axis illumination of sufficiently high angle is incident onto the mask. The mask was also chosen to have a magnification of 1. In this 1X stepper, the aspect ratio of the hole is 3.8 to 1 (300:78). The aspect ratio was chosen to be small because both Wong[152] and Wojcik[149] have found that Hopkin's formulation is not applicable in phase-shift masks with small aspect ratios.

The aerial images calculated with four different methods are compared in Figure 6-10. In Figure 6-10, the aerial image is plotted along a cut-line passing through the middle of the hole. These images are calculated with the decomposition method presented in this chapter (TEMPEST-PCD) and are then compared to the aerial image calculated with Abbe's method, TEMPEST-HN (TEMPEST with Hopkin's imaging for a Normally incident plane wave, as discussed in Chapter 2), and SPLAT (scalar imaging). The aerial image calculated with SPLAT predicts a larger intensity in the middle. This is most likely due to neglecting the boundary conditions assumed in scalar imaging. The boundary conditions

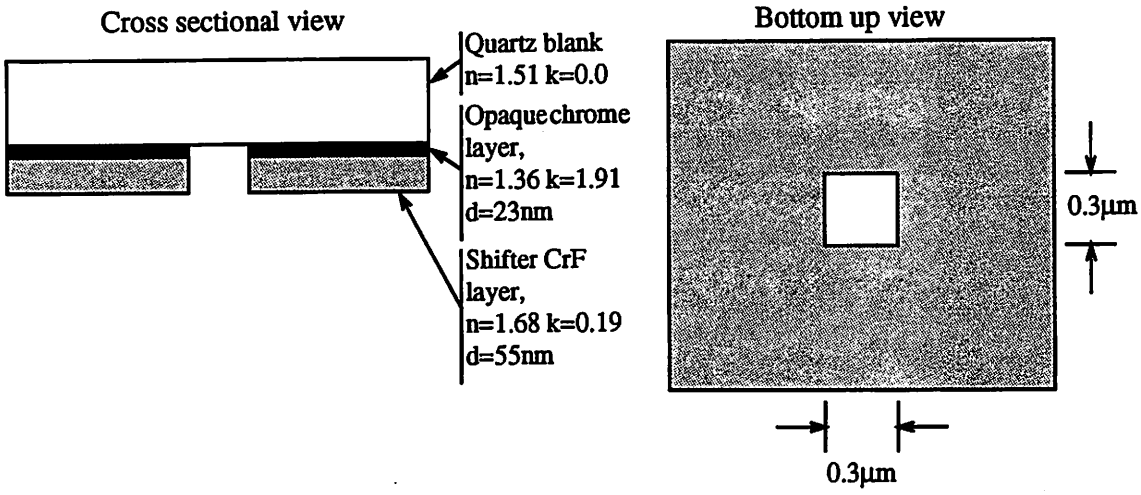


Figure 6-9. Attenuated phase-shift mask used to compare simulation results calculated with the decomposition formulation and with Abbe's formulation.

for vector polarized light create currents along the hole. These currents effectively reduce the amount of light passing through the hole. Consequently, the peak intensity calculated with SPLAT is greater than the intensity calculated with TEMPEST, in which the vector fields are solved. The aerial image calculated with TEMPEST and using Hopkin's formulation (TEMPEST-HN), using Abbe's formulation, and using the decomposition method are in close agreement with each other. The TEMPEST-HN image is in close agreement because, for this NA and σ, the diffraction orders are independent of the illumination angle and Hopkin's formulation remains valid. It is important to note that TEMPEST-HN requires only one simulation. Further, though the TEMPEST-PCD aerial image is in excellent agreement with Abbe's formulation, it requires fewer simulations than does Abbe's formulation. The aerial images using the decomposition method required 12 simulations while Abbe's formulation required 37 simulations.

6.8. Conclusions

A vector decomposition representation that includes polarization and partial coherence of the source was formulated by generalizing the scalar decomposition technique presented in Chapter 3. Each of these field excitations generated by this vector decomposition is coherent with itself yet incoherent with the other excitations. Since each excitation is coherent, TEMPEST can be used to analyze the scattering of this field from a topographi-

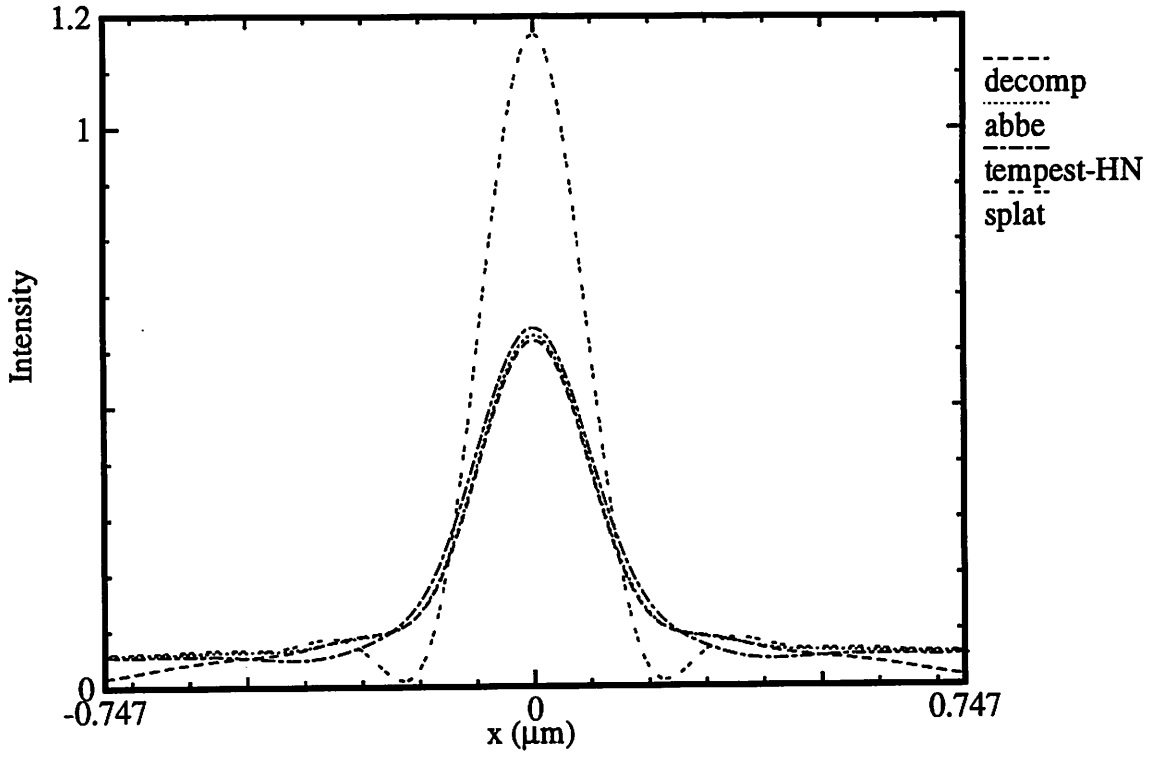


Figure 6-10. Comparison of aerial images for a contact hole in attenuating phase-shift mask of Figure 6-9.

The aerial image intensity was calculated with four techniques: SPLAT, TEMPEST-HN, TEMPEST with Abbe's formulation, and TEMPEST with the decomposition method. The hole is printed with a stepper having NA of 0.7, σ of 0.6, and λ of 248nm.

cal structure. Furthermore, since this vector decomposition is a transform technique that involves diagonalizing a matrix, this decomposition is optimal in an energy compaction sense, i.e., the least number of excitations are needed to represent the coherency and the polarization of the system.

The accuracy of this vector decomposition was analyzed by comparing the mutual intensity elements, J_{xx}, J_{xy}, and J_{yy}, of the coherence matrix generated with the decomposition to elements calculated analytically. The vector decomposition technique is accurate in the center of the simulation and suffers some loss of accuracy at the edges of the simulation. However, in non-periodic simulation domains the decomposition technique is more accurate than Abbe's formulation and requires fewer simulations. For example, in an inspection system with NA of 0.9 and σ of 1, the decomposition technique would require

93 simulations for a 0.01% error in the center of the domain while Abbe's formulation would require over 1125 simulations to produce the same error.

Using the vector decomposition, field excitations were generated for TE polarization. The propagation of these excitations incident upon an attenuating phase-shift contact hole was simulated with TEMPEST. Through these simulations, aerial images were calculated with the vector decomposition and with Abbe's formulation. The aerial image from the vector decomposition is nearly equivalent to the aerial image from Abbe's formulation. Since the vector decomposition is accurate in the center of the domain, the aerial image with the vector decomposition is most accurate in the center of the domain, and the aerial image with the decomposition begins to differ from the image using Abbe's formulation at the edge of the domain.

The vector decomposition technique is used in Chapters 7 and 8. Chapter 7 uses the vector decomposition technique to simulate two-dimensional structures for finding the aerial image from a phase-shift mask and of patterning a gate line into a two-dimensional active area well. Chapter 8 uses the vector decomposition technique to determine the effectiveness of using a pupil filter to analyze a three-dimensional structure, a contact hole etched in silicon dioxide.

7 Effect of Vector Polarized Partial Coherence on Phase Shift Methods and Reflective Notching

7.1. Introduction

In this chapter, the decomposition technique discussed in Chapter 6 and implemented through TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition) is used to simulate scattering from two-dimensional topographical structures. The goal is to assess the role and importance of partial coherence in lateral scattering at the wafer level in optical projection printing. The simulations studied in this chapter include imaging an attenuated phase-shift mask, patterning a line created by a two-layer phase-shifting resist edge, and patterning a gate line over an active-area trench. In Section 7.2, the accuracy of the TEMPEST-PCD simulations is verified by comparing aerial images of an attenuated phase-shift mask calculated with TEMPEST-PCD to aerial images calculated with TEMPEST and Abbe's formulation. The TEMPEST-PCD results are also compared to aerial images calculated with SPLAT and TEMPEST-HN (TEMPEST with Hopkin's imaging for a normally incident plane wave discussed in Chapter 2). Section 7.3 examines the feasibility of using a 180^0 phase-shifting two-layer resist edge in producing a sub-wavelength resolution line through TEMPEST-PCD simulation. In Section 7.4, TEMPEST-PCD is used to simulate the patterning of a gate line over an active-area trench. This active-area trench topography is similar to the topography reviewed in Chapter 5. However, the simulations in Section 7.4 are conducted with incident light that is partially coherent. Since a post-exposure bake (PEB) is normally performed on such simulations, the effect of PEB and the effect of the partial coherence (σ) on the final image profile will be discussed in Section 7.5.

7.2. Phase Shift Mask

In the first example, aerial images are calculated by collecting the light diffracting from a 1X two-dimensional attenuating phase-shift mask. This section was originally published in Ref. 115. At the time of publication, only a two-dimensional version of TEMPEST-PCD was available. Consequently, all the simulations in this section were done with a 1D cylindrical lens. Despite this, some valuable insight is to be gained on the validity of using TEMPEST-HN. It is important to recall from Chapter 2 that TEMPEST-HN makes the assumption that the diffraction orders are independent of the incidence angle and only one simulation is needed. Consequently, the goal of this section is to determine when the TEMPEST-HN simulation, as opposed to the TEMPEST-PCD, is valid.

The mask structure simulated with TEMPEST-HN and TEMPEST-PCD is shown in Figure 7-1a and consists of two 0.25μm attenuating phase-shift lines separated by an etched 0.45μm space. The dimensions were chosen to produce an image of nearly equal 0.35μm lines and spaces in a 1X i-line stepper with an NA of 0.6. In this phase-shifting mask technology approach, a 6% attenuation is designed into the mask by etching a 0.108μm thick attenuating material with n=2.115 and k=-j0.756 followed by a 0.130μm etch of the glass in order to achieve a 180^0 phase-shift.

Simulation results are shown in Figures 7-1b and 7-1c for a σ of 0.5 and of 0.7, respectively. The aerial images are calculated by four simulation methods: SPLAT (scalar imaging), TEMPEST-HN, TEMPEST using Abbe's formulation, and TEMPEST-PCD. The aerial images from a scalar mask using a simple vertical ray mask model (1D SPLAT simulation) shows considerable intensity in the middle of the feature. In 1D SPLAT, the transmission cross coefficient is calculated for a cylindrical lens. The aerial image is then given by integrating overlapping step functions of the transmission cross coefficient with the transmission function of the mask rather than overlapping circles in 2D SPLAT. These SPLAT images are compared to images from Hopkin's formulation with TEMPEST-HN. Using TEMPEST-HN, only one simulation is performed. This simulation calculates the diffracted orders due to a normally incident plane wave. These diffracted orders are collected by 1D SPLAT, which forms the aerial image intensity. The results with TEMPEST-

HN show reduced intensity of the feature as compared to SPLAT. The reduced intensity is expected due to the mask edges shorting out the parallel electric fields. It is important to note that TEMPEST-HN requires only one simulation. For both the σ=0.5 and σ=0.7 cases, the TEMPEST-PCD aerial image is in excellent agreement with Abbe's formulation, which yields the most accurate solution provided, however, that many obliquely incident plane waves are needed to accurately represent the illumination. Note, though, that a minor asymmetry is apparent in the images, which is most likely due to the numerical implementation of Maxwell's equations on a discrete grid on TEMPEST. This asymmetry can be eliminated by decreasing the grid spacing.

Due to the energy compaction property of the decomposition, TEMPEST-PCD is capable of analyzing the angular scattering dependence of the image with fewer simulations than the Abbe formulation. For σ=0.5, the aerial image using TEMPEST-PCD required 7 simulations while the image with Abbe's formulation required 21 simulations. Likewise in Figure 7-1c for σ=0.7, TEMPEST-PCD required 11 simulations while Abbe's formulation required 27 simulations.

Insight regarding the dominant sources of error in SPLAT can be gained through examination of the aerial images in Figure 7-1. These images, calculated with TEMPEST, predict that the peak intensity decreases compared to the images calculated with SPLAT. This decrease is due to a breakdown in the scalar assumption used in SPLAT. The scalar assumption neglects the currents created by the vector-like nature of the light, which creates currents in the mask layers. These currents decrease the amount of light passing through the opening. Consequently, the peak intensity calculated with TEMPEST is less than that calculated with SPLAT. The aerial images calculated with TEMPEST-PCD differ slightly from the images calculated with TEMPEST-HN. This is because Hopkin's formulation assumes that the diffraction orders are independent of the oblique angle of incidence, and for this mask structure this assumption is not valid. However, the mask studied in this section is used on a system with 1X magnification. On 1X masks the aspect ratio of opening width versus opening depth is small, and for the mask of Figure 7-1 this aspect ratio is 0.45μm to 0.238μm or 1.9:1. For 5X mask, however, this aspect ratio is much larger, in the order of 9.5:1. Wojcik found that when the aspect ratio is below 0.2 (depth to width) for a

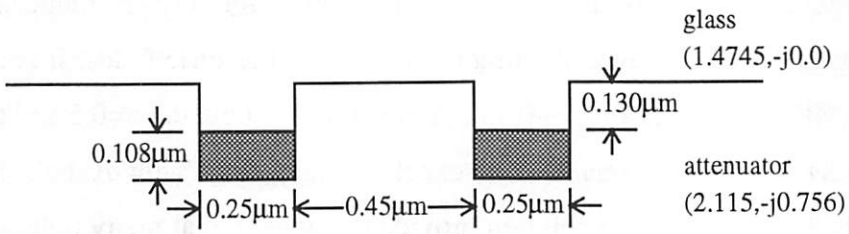


Figure 7-1a. Attenuating phase-shift mask that produces an image of nearly equal 0.35μm lines and spaces.

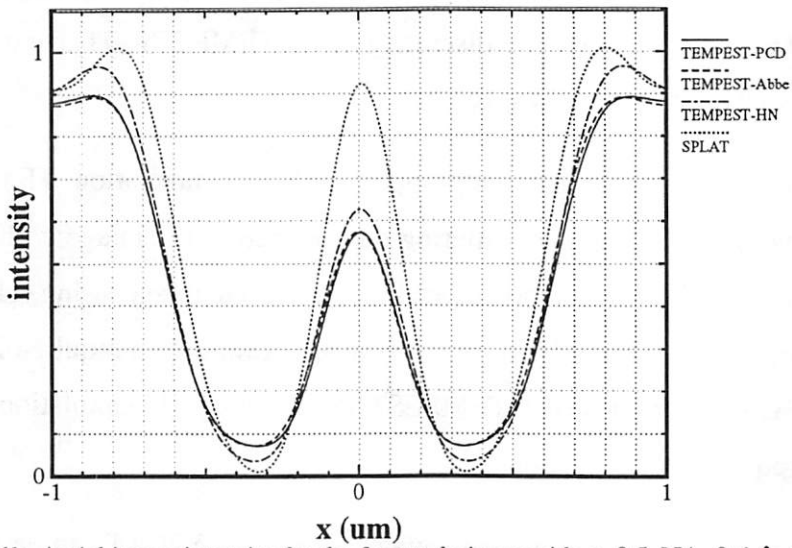


Figure 7-1b. Aerial image intensity for the four techniques with σ=0.5, NA=0.6, λ=365nm.

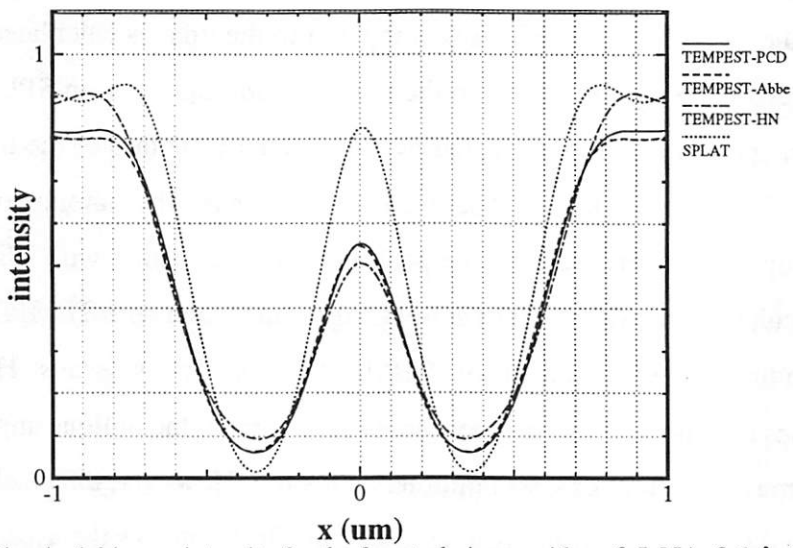


Figure 7-1c. Aerial image intensity for the four techniques with σ=0.5, NA=0.6, λ=365nm.

Figure 7-1. Comparison of aerial images produced by TEMPEST-PCD, TEMPEST-Abbe, TEMPEST-HN, and SPLAT

phase-shift mask, the diffraction orders are independent of incident angle for a mask imaged with an i-line stepper having NA of 0.6 and σ of 0.6 [149]. This implies that for 5X and 4X systems with typical NA's (<0.6) and σ's (0.3 to 0.7), imaging a PSM with the paraxial approximation used in Hopkin's formulation is valid. As a rule of thumb, when the system magnification is 4X and 5X, the aspect ratio is most likely less than 0.2; TEMPEST-HN, which requires one simulation, is accurate; and it is only necessary to use TEMPEST-PCD, which requires more than one simulation, when the system magnification is 1X and 2X and the aspect ratio is greater than 0.2.

7.3. Phase Shifting, Two Layer Resist Process

In the second example, TEMPEST-PCD is used to simulate patterning of a line by using a two-layer resist process. Mike Watts of Hewlett Packard suggested that this two-layer resist process would be effective at creating sub-wavelength resolution resist lines. This section is a simulation study of results he observed through experiments [139]. When the phase-shift layer is on the mask, light diffracting from the mask is collected by the collection optic. This collection optic low pass filters the light and eliminates the higher spatial frequencies. However, when the phase-shift layer is on the wafer, the light diffracted from the edge does not pass through the lens, which low pass filters the light. Consequently, higher spatial frequencies expose the resist in the phase-shifting, two-layer resist process. These higher spatial frequencies should improve the resolution when the shifter is on the wafer. Although the higher spatial frequencies expose the resist in the two-layer resist process as compared to the phase-shift mask, the two-layer resist may cause asymmetry in the image inside the resist or may cause alignment problems. Therefore, this section attempts to determine the feasibility of using a phase-shifting, two-layer resist process through TEMPEST-PCD simulation.

In this two-layer resist process, the top layer of resist is sensitive to one wavelength of light (such as 365nm) and the bottom layer is sensitive to a different wavelength such as 248nm. The top layer of resist is spun on, such that it provides a 180^0 phase-shift. This top layer of resist is exposed, in this case to 365nm light, and developed to form a 180^0 phase-shift edge. The resulting topography from this process is shown in Figure 7-2. This topog-

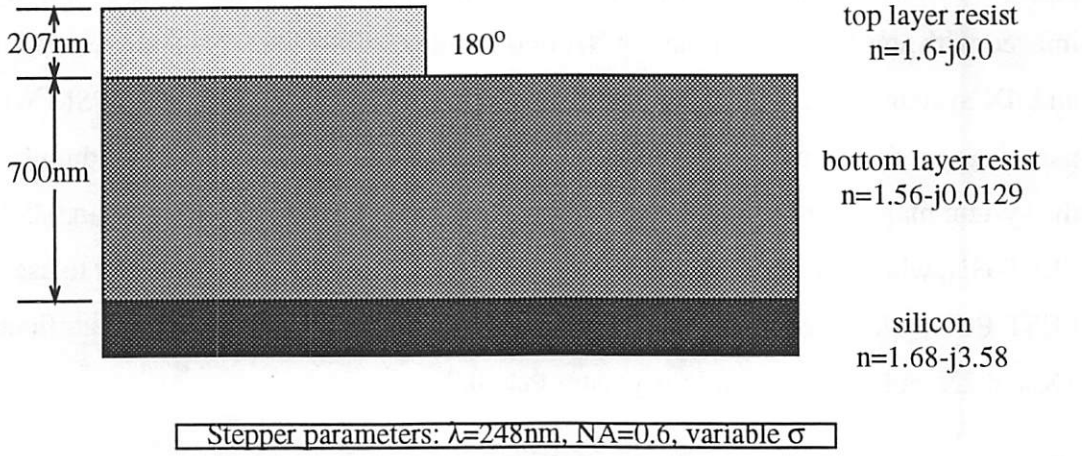


Figure 7-2. The topography of the two-layer 180° phase-shifting resist process.

raphy is then exposed at the second wavelength (248nm) and the 180° resist edge creates an exposed area below the line because of the 180° phase-shift. This unexposed area creates a photoresist line after development. The ability of this process to successfully create a line with sub-wavelength resolution that does suffer from process variation is examined with TEMPEST-PCD.

The topography in Figure 7-2 was simulated with TEMPEST-PCD for variable σ of 0.3, 0.5, and 0.7 at a wavelength of 248nm and an NA of 0.6. The intensity profiles for these three σ cases are plotted in Figure 7-3 prior to the post-exposure bake (PEB). In all of the intensity profiles, the 180° layer of resist is outlined in white. This 180° layer produces an unexposed area under the edge. This unexposed area will remain after development resulting in a photoresist line. The profiles in Figure 7-3, however, are not symmetrical, inasmuch as the contrast under the top layer resist (180° shifter area) is better than the contrast in the area not covered by the top layer (0° shifter area). These non-symmetric profiles are probably due to diffraction from the top layer resist and due to an index mismatch between the top layer resist (n=1.6) and the bottom layer resist (n=1.56). These non-symmetric intensity profiles may lead to non-symmetric developed line as the slope of the developed line in 0° shifter area will probably be poorer than the line slope in the 180° shifter area. In examining the effect of σ on the intensity, the profiles also indicate that the

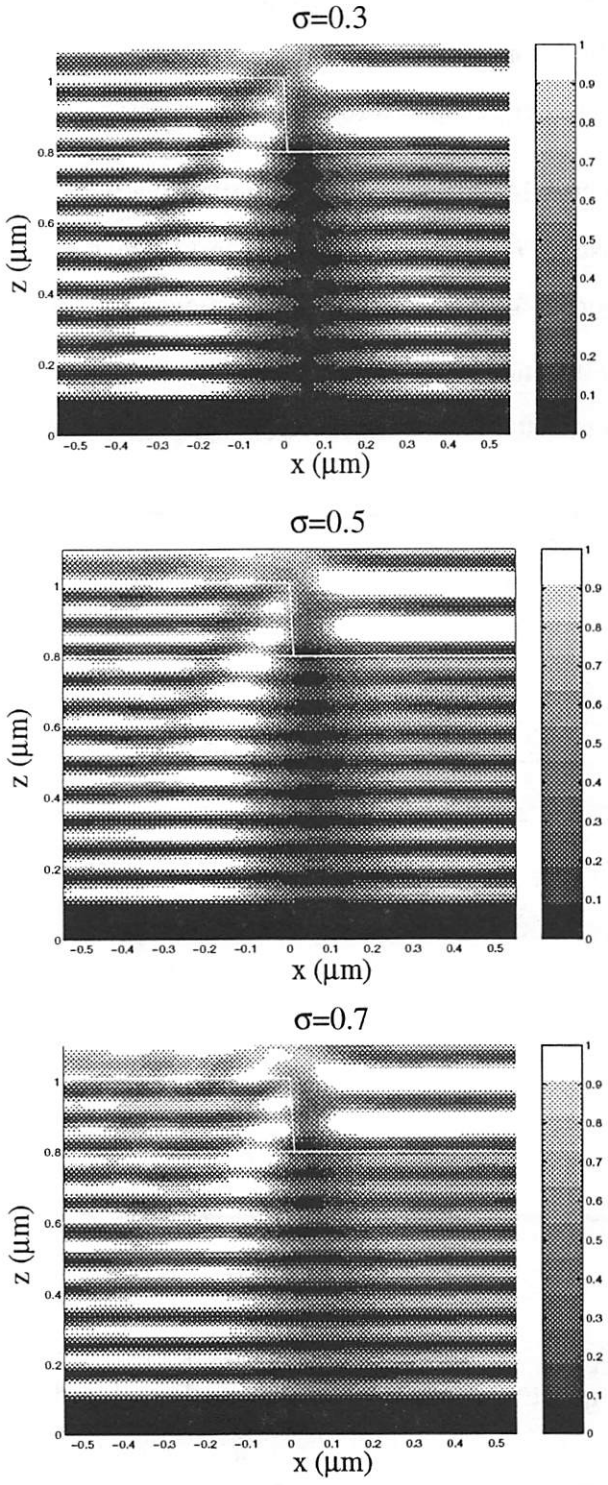


Figure 7-3. Intensity profiles before the PEB due to top layer phase-shifting edge for σ of 0.3, 0.5, and 0.7.

intensity contrast between this unexposed area and the exposed area is greatest at lower σ (σ=0.3) throughout the 700nm of resist. This improved contrast at σ of 0.3 produces a developed line with better edge slope throughout the resist.

To better measure the intensity contrast better, the intensity a quarter wavelength above the silicon substrate (z=0.14μm) is plotted in Figure 7-4 as a function of x at σ of 0.3, 0.5, and 0.7. Figure 7-4 then compares these images for the two-layer resist process with the image of a 180° phase-shift on the mask produced with an NA of 0.6 and σ of 0.5 stepper. Figure 7-4 demonstrates the improved image contrast and improved image inten-

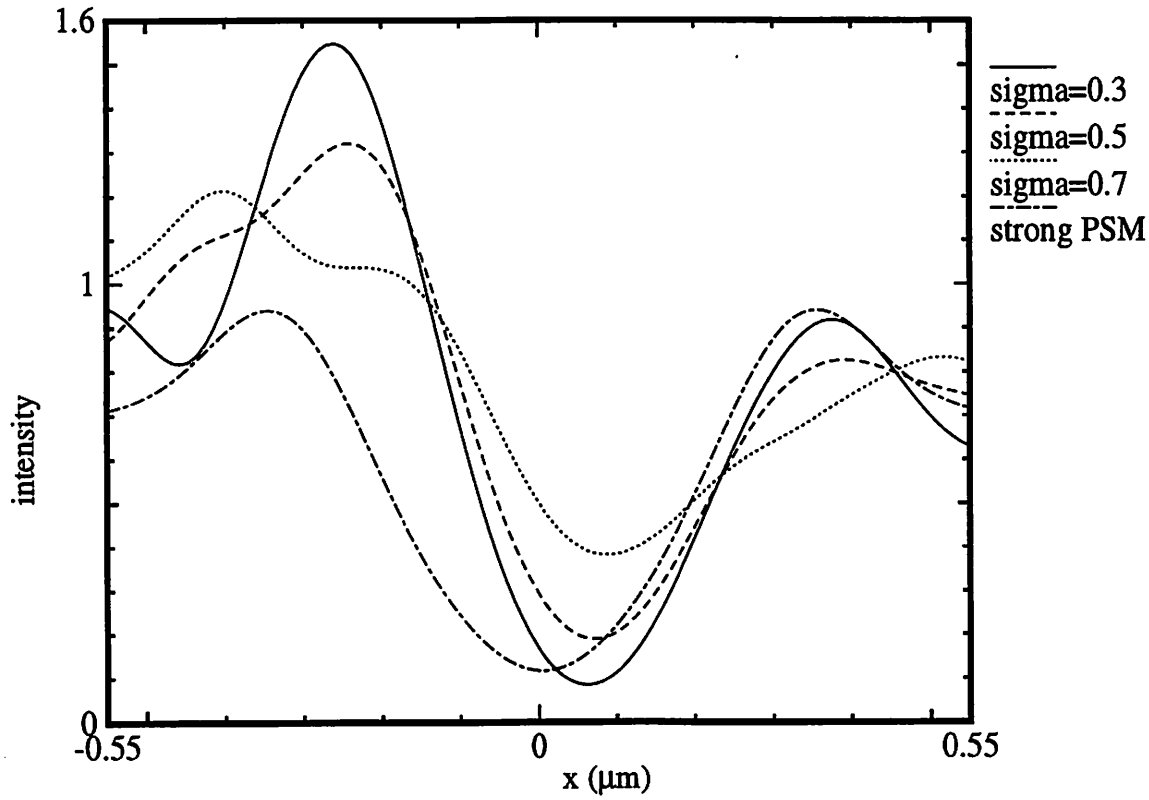


Figure 7-4. Intensity plot of the profiles in Figure 7-3 along the bottom of the resist a quarter wavelength above the silicon interface.

When σ is 0.3, the contrast is largest, which indicates that σ of 0.3 is probably the optimal σ. For comparison sake, the image from a strong PSM printed at NA of 0.6 and σ of 0.5 is plotted as well. Note that the resolution is better with the two-layer resist process than with the strong PSM.

sity slope at lower σ. In addition to indicating the reduction in line slope, Figure 7-4 also

shows that the intensity minimum does not occur at 0μm, which is where the edge located. When σ is 0.3, the intensity minimum is shifted 62nm away from the edge, and this distance increases as σ increases. In comparing the images produced with the phase-shifting two-layer resist process to the image produced with a strong phase-shift mask, the image minimum with the phase-shift mask is aligned with the phase edge on the mask. However, the linewidth at the 0.3 intensity threshold produced with phase-shift mask is 256nm while the linewidth with two-layer process is 206nm. This indicates that the resolution with two-layer process is better than the resolution with the strong phase-shift mask. Although the two-layer process produces a line with better resolution, the resolution is still far from the theoretical resolution limit. The theoretical limit occurs when the NA is 1, which would produce a wave that propagates with a direction parallel to the substrate. This would cause a wave propagating in both directions to interfere with a peak-to-valley interference distance of $\frac{\lambda}{4n}$, which is 40nm for this process. In the two-layer process, however, the peak-to-valley distance is approximately 300nm, which is far from the theoretical limit.

Although the two-layer process is capable of producing a line after development, process variation may limit the effectiveness of this two-layer resist process. To understand the role of the process variation, the thickness of the resist was varied, the slope of the 180^0 edge was changed, and top layer resist was made attenuating. The intensity profiles from these variations are plotted in Figure 7-5 for σ=0.3 only, since this σ leads to the best intensity contrast. In Figure 7-5, the intensity of profile is shown at the top left when the top resist layer is 180^0 (207nm thick resist), the edge of this top layer is 90^0, and the layer is non-attenuating (n=1.6-j0.0). The effect of matching the real and imaginary parts of the top layer resist index to the bottom-layer resist index (index is 1.56-j0.0129 in both top- and bottom-layer of the resist) is shown in the top right of Figure 7-5, effect of attenuating top layer resist (n=1.6-j0.013) is shown in the middle right, the effect of reducing the thickness of the resist 10% (thickness of 187nm which is 163^0) in the middle left, effect of increasing the thickness 10% (227nm or 198^0) in the lower left, and the effect of an 80^0 edge slope rather than a 90^0 edge slope in the lower right. These process variations give an indication

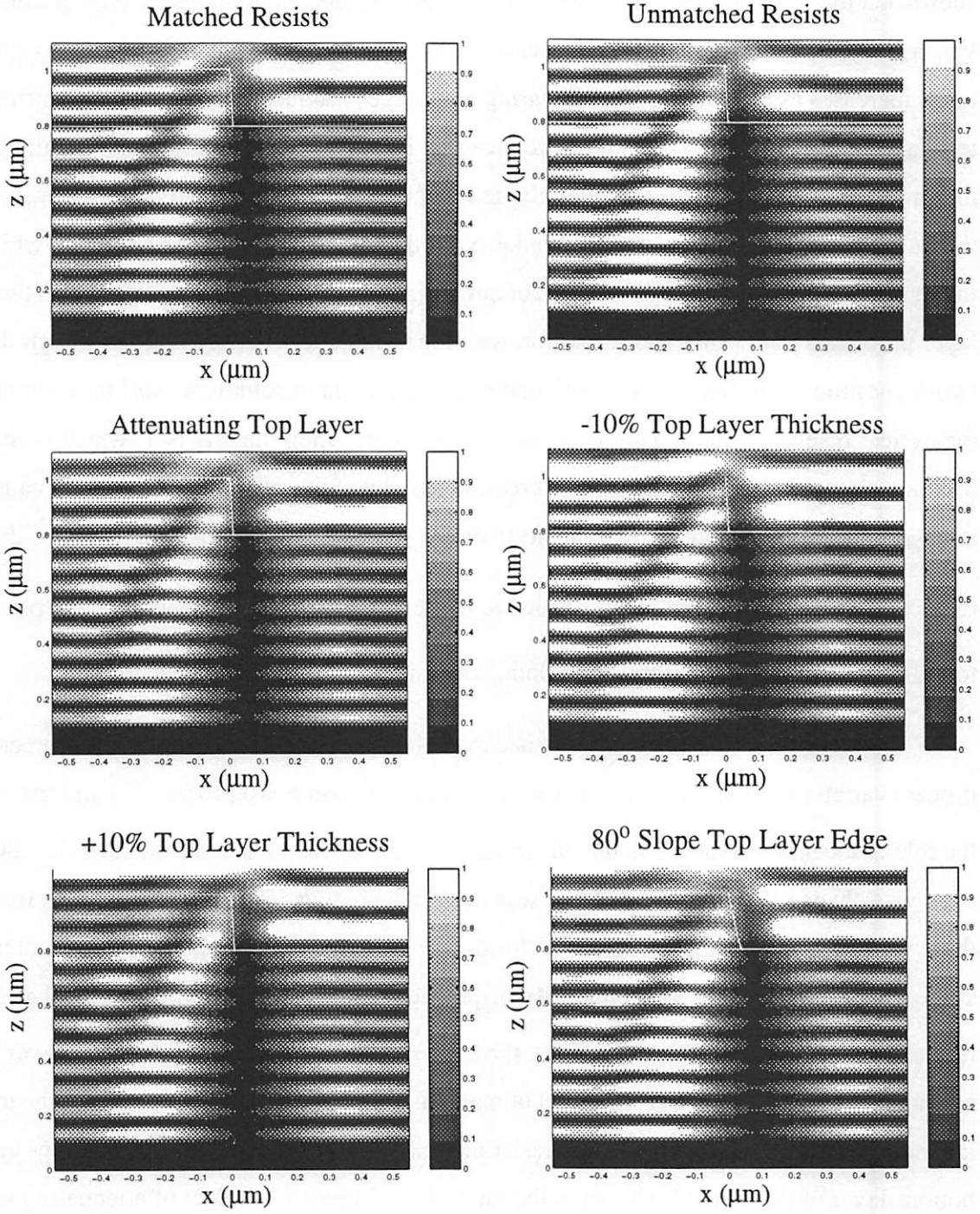


Figure 7-5. Intensity profiles for various process variations when σ is 0.3.

These variations include using an attenuating top layer (middle left), a 10% decrease in top layer thickness (middle right), a 10% increase in the top layer thickness (bottom left), and a 80° sloped top layer edge.

of how tightly the index of the resist, the resist thickness, and edge slope must be controlled in order to successfully use this two-layer phase-shifting resist process. In all of the profiles the top-layer resist edge is still capable of producing an unexposed area under the edge of the resist, which will lead to a line after the PEB and development. In comparing the top left profile (matched resists) to the top right profile (unmatched resists), some insight into the nature of the phase-shift due to the resist can be gained. Both profiles indicate that light is refracted into the top layer of the resist due to the phase edge created by the index change between the air and the top layer of resist. This refraction is due to a critical angle effect. Light incident at an angle of 90^0 with respect to the phase edge is refracted into the higher index material at an angle equivalent to the critical angle. Consequently, this critical angle effect increases the intensity under the top layer resist layer. This increase causes asymmetry in the aerial image and alignment problems with this two-layer process.

Although the profiles in Figure 7-5 look similar, there are slight differences. These differences are more apparent when the intensity in the resist is plotted a quarter wavelength above silicon interface as shown in Figure 7-6. In Figure 7-6, the intensity profiles for all the process variations are plotted as a function of x. These profiles show that all the intensity profiles are asymmetric. This asymmetry may produce an asymmetric photoresist line after development. In addition to this asymmetry, the profiles also indicate that the location of the minimum intensity changes as the process is varied. This change in minimum intensity location will shift the location of the printed line and may lead to alignment problems. In addition to these alignment problems, the line width was measured at 0.3 intensity threshold. The line width for the ideal process is 206nm and changes as the process is varied. The effect of these process variations on the location of the intensity minimum and on the line width at 0.3 intensity are displayed in Table 7-1.

In Table 7-1, the alignment error is the separation on the minimum intensity from the phase-shift edge, and the line width percent change error (last column) is the percent deviation from the line width with respect to the matched resist process. Table 7-1 indicates that the alignment error without process variations is 62nm away from the phase-shift edge. This 62nm offset is probably due to a shadowing effect created by the diffraction

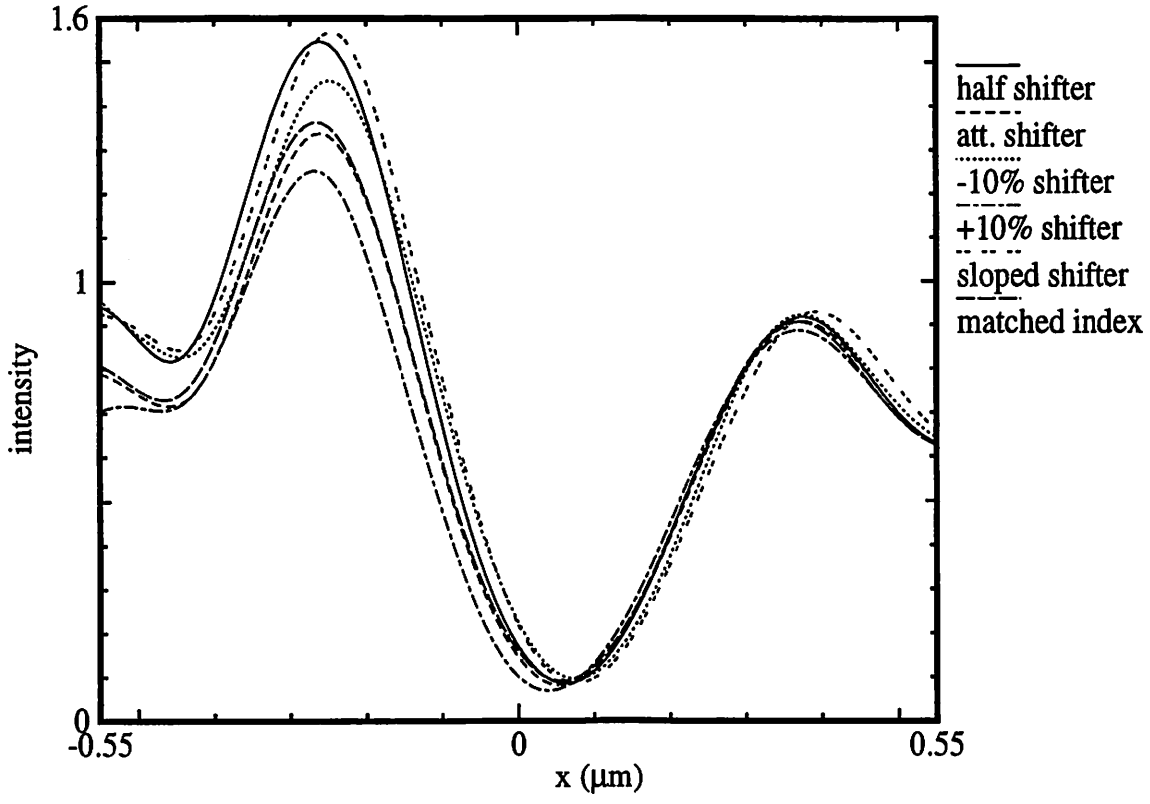


Figure 7-6. Intensity plot of the profiles in Figure 7-5 for the various process variations along the bottom of the resist a quarter wavelength above the silicon interface.

from the phase-shifting edge. This error is correctable by applying a 62nm bias to the edge. Again, from Table 7-1, we observe that a change in the resist thickness produces the most alignment error have a 10% increase in thickness causes the alignment to deviate by 26nm as compared to the alignment error without variation. This large deviation in alignment indicates that the thickness of the resist must be tightly controlled. When the resist edge slope is changed from 90^0 to 80^0, the alignment deviates by 11nm as compared to the alignment without variation. Although only one angle change was tested, this seems to suggest

Table 7-1. Alignment and line width errors in the two-layer phase-shifting resist process

Process Variation	Alignment error (nm)	Line width (nm)	Line width (% change)
Matched Resists	58	211	0.0%
Unmatched Resists	62	206	-2.4%
Attenuating Top Layer Resist	57	216	+2.4%
-10% Top Layer Thickness	75	200	-5.2%

Table 7-1. Alignment and line width errors in the two-layer phase-shifting resist process

Process Variation	Alignment error (nm)	Line width (nm)	Line width (% change)
+10% Top Layer Thickness	36	225	+6.6%
80° Slope Top Layer Edge	73	205	-2.8%
Strong Edge on PSM	0	256	+21.3%

that the thickness must be more tightly controlled than the angle. The line width at the 0.3 intensity threshold is also listed in Table 7-1. We see that change in the linewidth is also most pronounced when the thickness is varied, as a 10% increase in thickness causes 9.2% increase in line width.

Table 7-1 indicates that process variations in the top layer resist thickness of a few nanometers in the spun on resist causes several nanometer changes in the line width and in the alignment. Consequently, the top layer resist must be tightly controlled in order to limit changes in the line width and in alignment. Since this section was intended to examine only the feasibility of a two-layer phase-shifting resist process, the simulations were performed for a flat substrate and topography on the wafer was neglected. Based on the results for the flat substrate, it was shown that a few nanometer deviation in the resist thickness causes alignment and line width control problems. Since topography on the wafer may produce changes in the thickness of the spun on resist, the wafer topography plays a significant role in limiting the performance of this two-layer process, topography should not be neglected in future studies.

In conclusion, Table 7-1 and Figures 7-5 and 7-6 show that using a 180° phase-shifting two-layer process is capable of producing a sub-wavelength line. The resolution of this line with the two-layer process was shown to be on the order of a wavelength, while the resolution with a strong PSM is on the order of two-thirds of a wavelength. The simulations also indicate that the two-layer process is capable of producing an intensity minimum that is less than 0.1 when σ is less than 0.3. This intensity minimum is necessary in order to produce a well-defined photoresist line after development. The phase-shift edge of the top layer resist causes light to refract into the bottom of the resist into a region directly below the top layer. This refraction occurs even when the index of the top resist layer is matched to the index of the bottom resist layer. This refraction is due to the critical

angle effect of light incident at a 90^0 incident angle with respect to the phase edge. This critical angle effect causes asymmetry in the intensity profile which may cause asymmetry in the developed resist line. The refraction at the edge also causes the 25nm shift in alignment of the line with the top layer resist edge. Finally, ±10% thickness variation in the top layer resist causes as much as ±5% linewidth variation and 25nm shift in alignment, while the slope of the phase edge was shown to produce little change in the linewidth and alignment.

7.4. Reflective Notching Due to Patterning over a 2D Active Area Trench

In the third example, TEMPEST-PCD is used to simulate patterning of a gate over a two-dimensional active-area trench. This active-area trench topography is similar to the topography simulated in Chapter 5. However, in this chapter, the partial coherence of the field has been included in the simulations through TEMPEST-PCD whereas the simulations of Chapter 5 assumed that the excitation field is fully coherent. In addition to looking at the influence of σ, the results of simulations examined the effect of the post-exposure bake (PEB) on the resist image when the acid diffusion due to the PEB is Fickian. This section was done in collaboration with Chris Progler of Texas Instruments (TI). The experimental results generated at TI were than compared to the simulation results with TEMPEST-PCD. This work was presented at the *Optical Microlithography X SPIE* Conference and published in the proceedings [116].

7.4.1. Topography of the Gate Patterning Process

The active-area topography structure examined in this paper is that of a LOCOS process. First a pad oxide was grown, and then a layer of silicon nitride was deposited and patterned with an active-area mask. The field oxide was then grown using LOCOS to a thickness of 2500Å. The lateral encroachment lead to a "bird's beak" forming a trench with a side wall step angle of 25^0 with rounded corners. Active area trenches were grown with varying moat widths of 0.6, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 7.0, and 10.0μm. After growing the oxide, 250nm of polysilicon was deposited over the structure. Top down SEM's show that the polysilicon has significant granularity. In order to examine the trench topography, a wafer was cleaved and an SEM was taken of the resulting topography, which

is shown in Figure 7-7. Although the SEM has a bottom anti-reflective coating (BARC), a BARC was not deposited on the bare polysilicon in the wafers patterned, simulated, and studied here. To pattern the gate, a chemically-amplified positive DUV resist was spun on to achieve a thickness of 700nm over the field region. A gate was then patterned over the active-area trench in order to achieve a 0.25μm line on a DUV 248nm stepper having an NA of 0.5 and a σ of 0.3 and 0.6. The resist was then developed and the resulting photoresist lines were examined with top down SEM's to determine the extent of reflective notching.

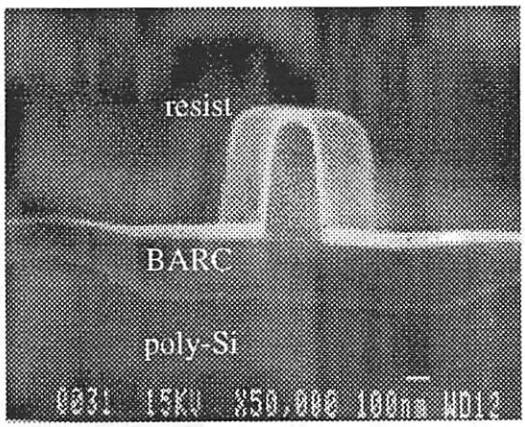


Figure 7-7a. SEM cross section of the wafer topography. Although this SEM has a BARC, a BARC was not used in any of the experimental or simulation results contained herein.

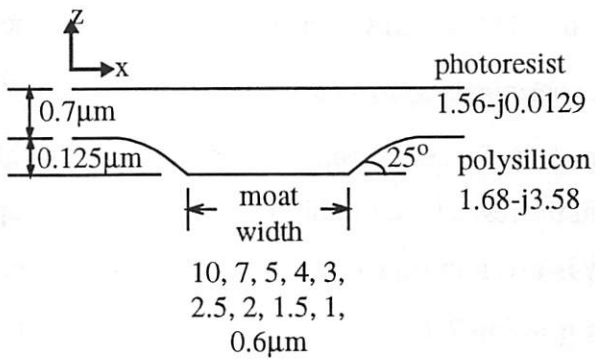


Figure 7-7b. Simulation topography representation of the wafer used in the 2D TEMPEST-PCD simulations.

Figure 7-7. The topography used in the experimental and simulation studies to determine the effect of σ on reflective notching.

The 2D active-area trench from the SEM in Figure 7-7a was modeled by the structure shown in Figure 7-7b for simulation. The trench consists of a planar structure having moat widths of 0.6, 1, 1.5, 2, 2.5, 3, 4, 5, 7, and 10μm. The polysilicon deposited over the LOCOS trench formed a 25° step with rounded corners. The spun on resist was modeled as a planar surface that fully covers the trench with a resist thickness of 700nm. The image formation in these 2D trenches was then simulated using TEMPEST-PCD, as described in Section 7.4.2.

7.4.2. Simulation Technique

The aerial image of a 0.25μm (wafer dimension) line formed by a DUV 248nm stepper with NA of 0.5 and σ of 0.3 and 0.6 was calculated by SPLAT. From this incident aerial image the intensity inside the photoresist of the topography structure shown in Figure 7-7 was simulated by TEMPEST-PCD. These TEMPEST-PCD simulations were run on a Sun Ultra 2 workstation. Each 2D simulation, including the partial coherence required only 5MB of RAM and ran in 5 minutes. The topography structure scatters the incident image from exposed areas into unexposed areas, causing reflective notching. In each TEMPEST-PCD simulation, the intensity inside the photoresist was calculated by using Dill's A,B,C model. The A,B,C parameters of the positive photoresist used in the experiments were measured to be A=-0.48μm^{-1}, B=1.134μm^{-1}, and C=0.0006cm^2/mJ with n=1.56. Since C is small and the dose used in the experiments was 12mJ/cm^2, the bleaching of the resist is negligible and A and B can be replaced by complex refractive index γ, where γ is given in Equation 7-1. From the intensity, I(x,z), the exposure state, M, which is lin-

Equation 7-1. $$\gamma = n - jk = n - j\lambda (A+B)/4\pi = 1.56 - j0.0129$$

Equation 7-2. $$M(x,z) = exp[-DCI(x,z)]$$

early proportional to photoactive compound concentration (PAC), can be calculated by Dill's model as shown in Equation 7-2, where D is the dose and I(x,z) is the intensity at a point (x,z) in the resist. In some chemically-amplified DUV resists, a post-exposure bake (PEB) diffuses the PAC through a non-linear concentration-dependent diffusion [160]. Since we are concerned with the optical effects of the spatial coherence, the PEB diffusion is modelled by a Fickian diffusion process rather than the concentration-dependent diffusion. The exposure state after PEB for a Fickian diffusion process, N, is given by the exposure state, M, convolved with a Gaussian function with diffusion length, $a = \lambda/4n$. Since D multiplied by C is much less than one, the PAC can be expended into a Taylor expansion [39] as shown in Equation 7-3, where * is the convolution operator. The PAC after PEB is

Equation 7-3.
$$N(x,z) = M(x,z) * exp\left[-\frac{1}{2}\sqrt{(x^2+z^2)/a^2}\right]$$
$$\equiv [1-DCI(x,z)] * exp\left[-\frac{1}{2}\sqrt{(x^2+z^2)/a^2}\right]$$

proportional N and since N is proportional to the intensity, I(x,z), the PAC after PEB is proportional to the intensity, I. The intensity profiles in Section 7.4.3 are shown before and after the PEB convolution. The profiles after the PEB were left undeveloped to better understand the effect of the optical model.

7.4.3. Simulation Results

In Figure 7-8, the intensity profiles are shown before the PEB over the recessed active-area trench with varying moat widths of 0.6, 1.0, 1.5, and 2.0μm and with σ of 0.3 and of 0.6. The profiles are arranged with σ of 0.3 on the left-hand column and with σ of 0.6 on the right-hand column, the moat width increasing from the top of the page to the bottom of the page. In the following paragraphs, the CD profile before the PEB is first considered first, then the CD profile after the PEB is examined in order to determine the effect of a PEB on CD profile.

The topography caused light to scatter from exposed areas into unexposed areas. For example, the 25^{o} poly step in Figure 7-7 causes light to scatter specularly off the topography at an angle of approximately 50^{o}. This scattered light interferes to form a bright intensity area along a line bisecting the poly step angle. When the light is normally incident (0^{o}), the electric field reflectivity at the resist and polysilicon interface is 0.74, while at 25^{o} incidence, the reflectivity is 0.76. When the moat width is 2.0μm, the light scattering off the poly step reflects off the top of the photoresist. This light reflecting off the top of the photoresist is scattered into the unexposed gate area, thus increasing the total intensity in the gate area. This increase in total intensity causes a narrower CD. In comparing the simulation results for the 2.0μm moat width at various σ, we find that the σ has little impact on the CD profile, the profiles being almost identical at σ of 0.3 and of 0.6. When the moat widths are 1.5μm and 1.0μm, the light reflects off the poly step and interferes near the top of the resist. As the moat width increases, the light interferes further down from the top of the resist in the unexposed gate line area. In comparing theses results versus σ at a moat width of 1.5μm, the intensity in the σ of the 0.3 case is approximately 0.7 at a position of x=0μm and z=0.7μm. However, when σ is 0.6, the intensity is approximately 0.4 at the same location. This indicates that the interfering intensity is larger in the unexposed area

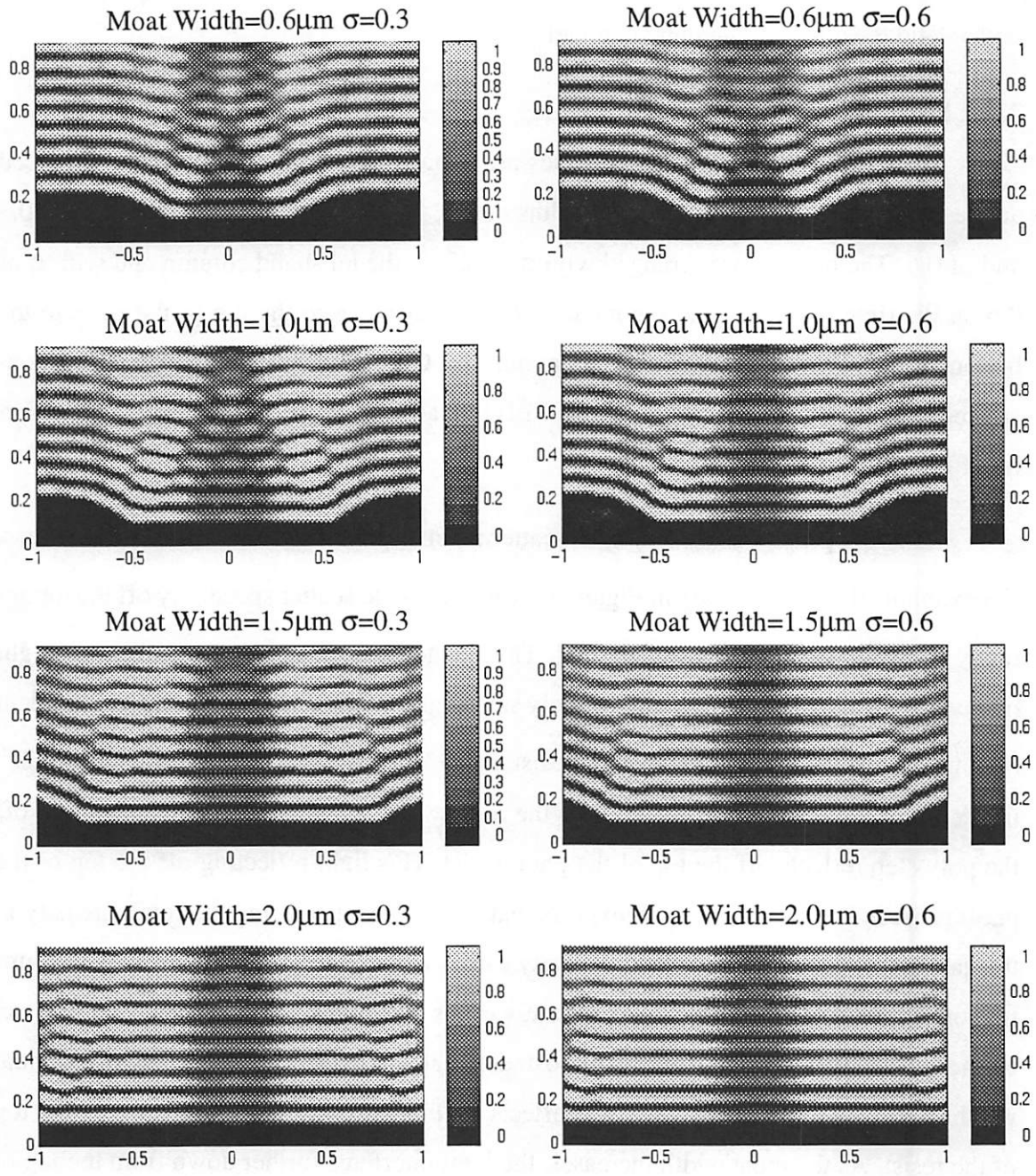


Figure 7-8. Two-dimensional TEMPEST-PCD simulated intensity profiles before the PEB when patterning a gate line over an active-area trench with topography of Figure 7-7.

The simulations show an increased intensity in the middle of the unexposed line near the top of the line at σ of 0.3 (left) as compared to a σ of 0.6 (right).

when σ is 0.3 than when σ is 0.6. This effect is more readily seen in the profiles when the moat width is 1.0μm and 0.6μm. In both of these cases, the light reflects of the step into the unexposed gate area. The intensity is higher in this area when σ is 0.3 than when σ is 0.6 in both 1.0μm and 0.6μm moat widths. This would seem to indicate that higher σ have less reflecting notching problems than lower σ. However, since a post-exposure bake is always performed on chemically-amplified 248nm resists, the effect of the PEB must be taken into account prior to reaching a conclusion on the role of σ on reflective notching and on CD variation.

In order to understand the effect of the post-exposure bake, a PEB was performed on the intensity profiles of Figure 7-8 by using Equation 7-3. These intensity profiles after a PEB are shown in Figures 7-9 and 7-10. In Figures 7-9 and 7-10, intensity profiles are given for wafers with varying moat widths of 0.6, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, and 5.0μm and with σ of 0.3 and 0.6. In order to parallel the profiles before the PEB in Figure 7-8, the profiles in Figures 7-9 and 7-10 are arranged with σ of 0.3 on the left and σ of 0.6 on the right and with increasing moat width from the top to the bottom. The effect of the topography and σ is measured by using the CD length at the bottom of the trench and by using the CD quality (steep-profile edge slope) as performance metrics.

The effect of the PEB and of the topography is first considered by examining the effect of the moat width on the CD quality. When the moat width is 5.0μm, the steps from the moat are separated by such a large distance that light reflecting from the steps no longer impacts the CD. As the moat width decreases to 2.0μm, the simulations show that light diffracting from the step forms a beam of radiation along the line bisecting the poly step angle. This beam reflects off the top of the photoresist into the unexposed area causing an increase of the total intensity, as shown in the cartoon of Figure 7-11a. This increase in the total intensity in the unexposed area leads to a narrower CD. At a moat width of 2.5μm, the step separation is larger, and the light reflecting off the top of the resist is farther from the unexposed area. This reflected light still increases the total intensity in the unexposed area, but to a lesser extent than the 2.0μm moat width. Consequently, the CD is slightly wider at 2.0μm than at 2.5μm. As the moat width increases from 2.5μm, the CD continues to increase until a steady CD is achieved when the moat width approaches 5.0μm. In the cases

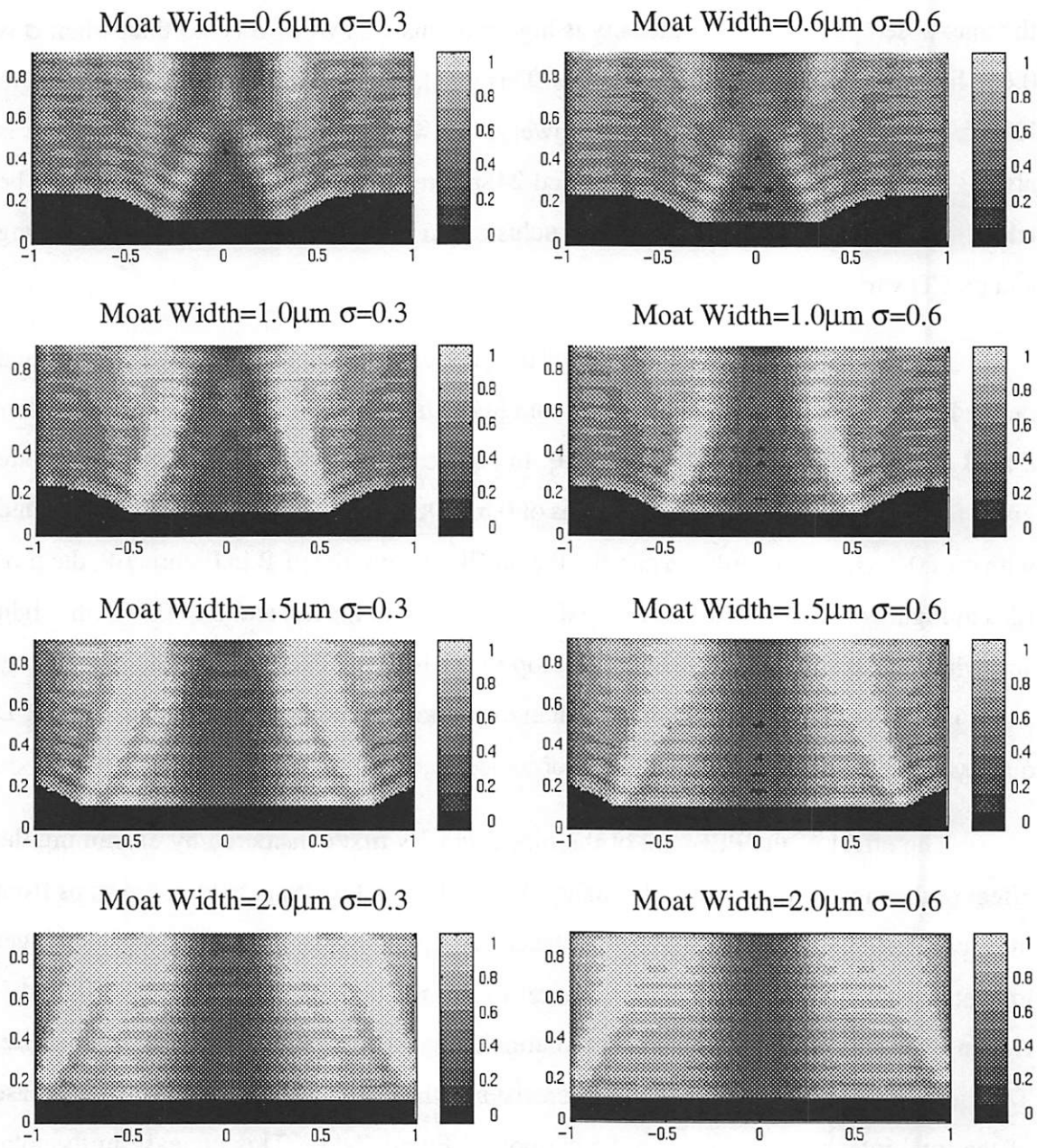


Figure 7-9. Two-dimensional TEMPEST-PCD simulated intensity profiles after the PEB when patterning a gate line over an active-area trench with the topography of Figure 7-7.

The top of the gate is knocked off when the moat is 0.6μm wide. The CD at the bottom of the line increases on the 1.0 and 1.5μm wide moats due to diffraction from the step. The CD decreases on the 2.0μm wide moats due to reflection from the top of the resist. The CD is slightly wider at σ of 0.3 than at σ of 0.6 due to the larger coherence area at σ of 0.3. All the dimensions above are in micrometers.

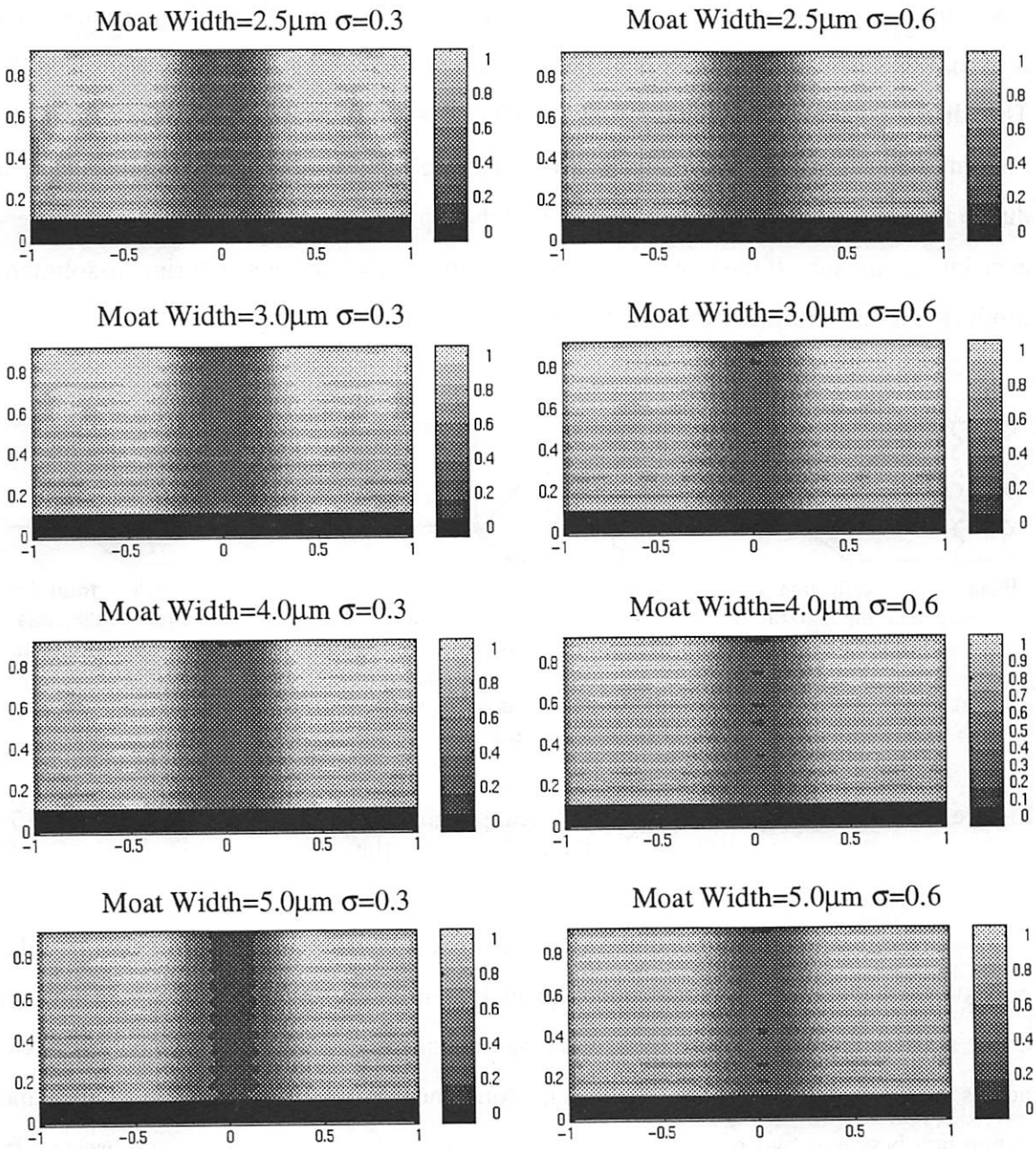


Figure 7-10. Continuation of Figure 7-9

The CD continues to increase as the moat width increases because of the wider separation between light reflecting off the step into the unexposed area.

with moat width of 1.0μm and 1.5μm, the light on either side of the dark gate area interacts with a step topography which is on the order of 2 to 3 wavelengths. The light incident onto this step is diffracted into a beam in which the diffraction area of the beam is given by the aperture radiation equation. This equation states that light diffracts into a beam with an area that is inversely proportional to the aperture size. Since the step topography is on the order of 2 to 3 wavelengths, the light is diffracted into a large area, as shown in Figure 7-11b. This diffraction causes degradation of the CD edge slope and increases the CD at the bottom of the trench. For the simulations with a 0.6μm moat width, the bright intensity area due to light reflecting off the step interferes at the top of the unexposed line. This interference knocks the top off the line and is subject to rapid linewidth change during dissolution, producing a line with wide CD at the bottom, as seen in Figure 7-11c.

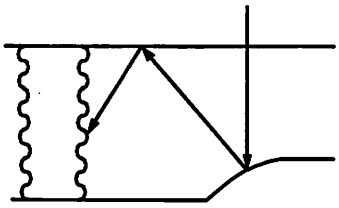


Figure 7-11a. Reflection from the step and the top of the resist reduce the CD when the moat width is 2.0μm and 2.5μm.

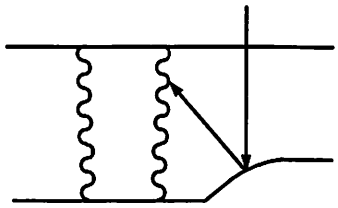


Figure 7-11b. Reflection from the step reduces the edge slope but leaves the CD at the wafer surface unaffected when the moat width is 1.0μm and 1.5μm.

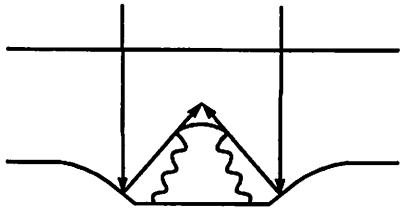


Figure 7-11c. Reflection from the step exposes the top of the resist line when the moat width is 0.6μm.

Figure 7-11. Schematics showing the physical phenomena observed in Figures 7-9 and 7-10 for reflective notching into the gate.

In considering the effect of σ on CD, the light from the exposed area scatters off the poly step into the unexposed gate area. The total intensity in the unexposed gate is the sum of the scattered intensities weighted by the partial coherence. The coherence between two points on the wafer is given by the mutual coherence. For a circular source, the mutual coherence between two points, $\mathbf{r}_1$ and $\mathbf{r}_2$, is given by the Airy function, as observed in

Equation 7-4. The first null of the Airy function occurs at $0.61\lambda/\sigma NA$ which corresponds

Equation 7-4.

$$\mu(\mathbf{r}_1, \mathbf{r}_2) = 2\frac{J_1\left(\frac{2\pi}{\lambda}\sigma NA|\mathbf{r}_1 - \mathbf{r}_2|\right)}{\frac{2\pi}{\lambda}\sigma NA|\mathbf{r}_1 - \mathbf{r}_2|}$$

to 1.0μm and 0.5μm for σ of 0.3 and 0.6, respectively. The location of this first null is critical in determining the notching in the line from the topography. At low σ, the first null occurs at a larger separation than at higher σ. This implies that the light is coherent over a larger area and the light interacts over a larger distance at low σ. However, images formed at low σ have a larger intensity slope, which leads to better edge slope in the photoresist. Therefore, at low σ the light reflecting off the topography step has a greater tendency to cause notching, but results in a sharper resist edge profile. This trade-off is seen in the simulations when the moat widths are 1.0 and 1.5 μm. From Figure 7-8, before the PEB, there is a slight increase in intensity and notching at z=0.7μm when σ is 0.3 as compared to a σ of 0.6 because the light is coherent over a larger area at low σ. Although there is a slight increase in intensity in the unexposed area for the narrower moat widths at σ of 0.3 as compared to σ of 0.6, the PEB diffusion averages the increase over a large area, as seen in Figure 7-9. Because of the PEB diffusion, there is little impact of σ on notching at these narrow moat widths. On the wider moats (2.0μm and greater), the simulations show that the CD is influenced by multiple reflections off the step and off the top of the photoresist. At such wide moat distances, the CD is impacted by interference in the unexposed area from light reflecting off the topography step. The intensity of this light is added as weighted by the mutual coherence. Since the moat steps are separated by a distance that is much larger than the first zero in the mutual coherence, the mutual coherence is small, and the intensities add at a σ of both 0.3 and 0.6. This implies that, similar to the narrow moat width cases, there is little effect of σ on notching at large moat widths as well.

7.4.4. Experimental Results

In Figures 7-13, 7-14, and 7-15, top down SEM's of developed photoresist gate lines are shown inside the active-area trench for wafers with varying moat widths of 0.6, 1.0, 2.0, 2.5, 3.0, 4.0, and 5.0 and with a σ of 0.3 and of 0.6. The profiles in Figures 7-13, 7-14, and 7-15 are arranged to parallel the simulation results with σ of 0.3 on the left-hand

column and σ of 0.6 on the right-hand column and with the moat width increasing from the top of the page to the bottom of the page. The CD at the bottom of the trench was measured and recorded above the SEM's in Figures 7-13, 7-14, and 7-15. Since the exposure dose was not changed from 12mJ/cm^2 in exposing the lines with σ of 0.3 and 0.6, the lines with σ of 0.6 were over exposed as compared to the lines exposed with σ of 0.3. Consequently, the CD at σ of 0.6 is slightly smaller than the CD with σ of 0.3.

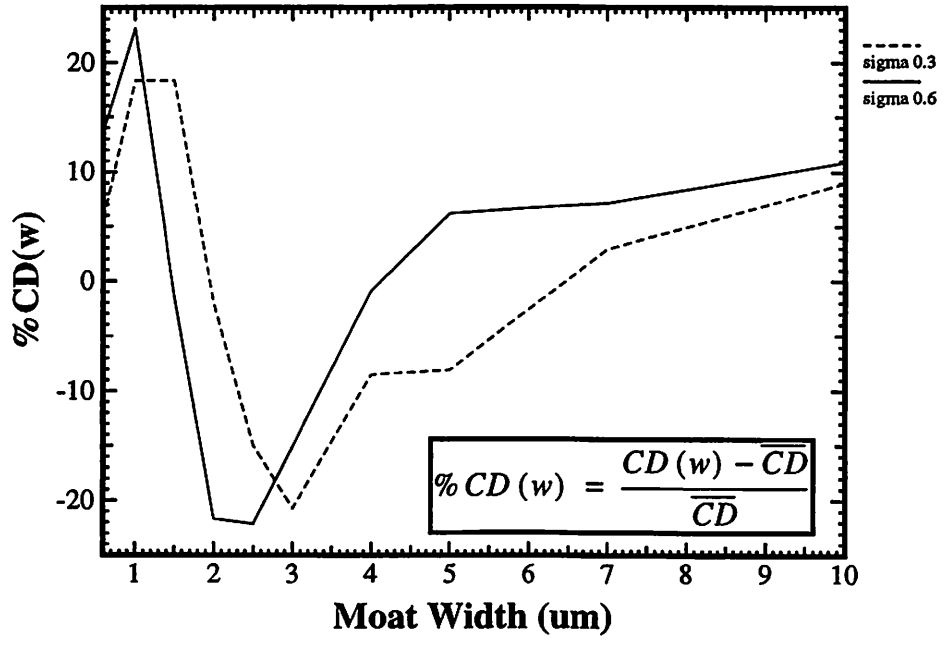


Figure 7-12. Experimental measurement of the percent CD change as a function of moat width for σ of 0.3 and for σ of 0.6.

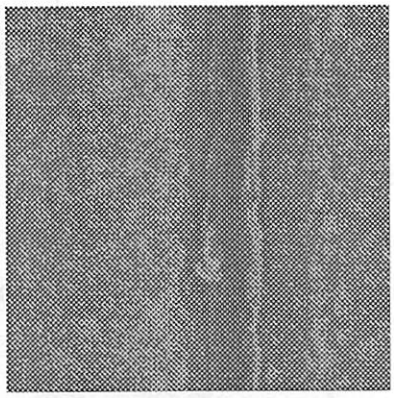
The functional relationship is similar for both σ's; however, the percent CD change curve with σ of 0.3 is shifted to larger moat widths where the coherence area is larger.

The percent CD change from the mean CD at σ of 0.3 and at 0.6 as a function of moat width was determined from the experiments and is plotted in Figure 7-12. This plot shows that percent CD change has a damped sinusoidal behavior as a function of moat width. The reasons for this damped sinusoidal CD behavior can be understood from the observations made from the simulations. As shown for the simulation of a moat with width 2.5μm and greater, the CD slowly increases because the moat step is separated far from the unexposed area where it has less effect on the percent CD change. In the case of 1.5μm-2.5μm wide moats, the percent CD change is negative and slowly decreases. By definition,

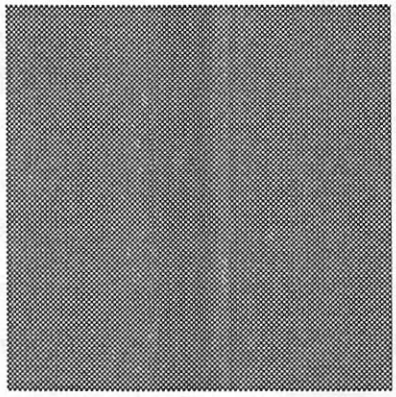
the negative percent CD change occurs when the CD is less than the average CD. The simulations indicate that smaller CD occurs in this moat width regime because light reflects off the step and off the top of the resist into the unexposed area, thereby increasing the intensity. This increased intensity in the unexposed area results in smaller CD after development. In 1.0 and 1.5μm wide trenches, the light is diffracted into a large beam area, as shown in the simulations. This leads to a larger than average CD (positive percent CD change). The percent CD decreases as the moat width increases because the light is diffracted into a smaller area at larger moat widths as given by the aperture radiation area equation. As shown for the simulation with a 0.6μm wide moat, the CD slightly increases and the top of the line is knocked off from light scattering off the step, as seen in the simulation of Figures 7-9 and 7-10 and in the SEM of Figures 7-13, 7-14, and 7-15.

In Figure 7-12, the percent CD curves have the same functional damped sinusoidal relationship at σ of 0.3 and at σ of 0.6, with a slight shift in curve at σ of 0.3. This shift is due to the larger coherence area at smaller σ being able to affect the line CD at larger moat widths. Since the curves at σ of 0.3 and at 0.6 are similar in shape, the curves indicate that the topography has a similar impact on the percent CD change regardless of the coherence. The lines in the SEM's of Figures 7-13, 7-14, and 7-15, however, show more CD variation along the line when σ is 0.3 than when σ is 0.6. As seen in the SEM's of Figures 7-13, 7-14, and 7-15, the polysilicon deposited onto the LOCOS trench shows a large amount of granularity as seen in the SEM's of Figures 7-13, 7-14, and 7-15. These polysilicon grains scatter light from the exposed area into the unexposed gate area, causing CD variation and slight notching along the line. The total intensity scattering into the unexposed gate area is given by intensity contributions of light scattering from the various grains as weighted by the mutual coherence. As discussed in Section 7.4.3, the location of the first null is critical in determining the notching from the topography. In this case the topography is caused by the polysilicon grains. When σ is 0.6, light scattering off polysilicon grains separated by 0.5μm or less add in the unexposed area and impact the CD. Likewise, when σ is 0.3, grains separated by 1.0μm or less impact the CD. The amount of CD variation depends on the intensity in the area bounded by the first null. Since the intensity is smaller in the unexposed area at a separation of 0.5μm than at 1.0μm, the gates patterned with σ of 0.6 show

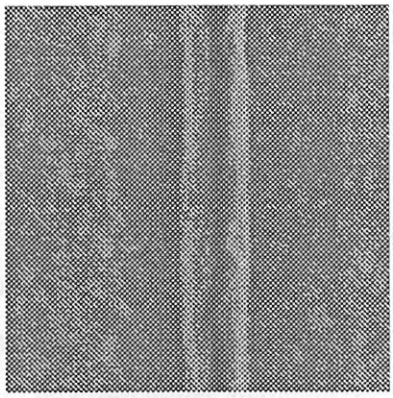
Moat Width=0.6μm σ=0.3 CD=260nm



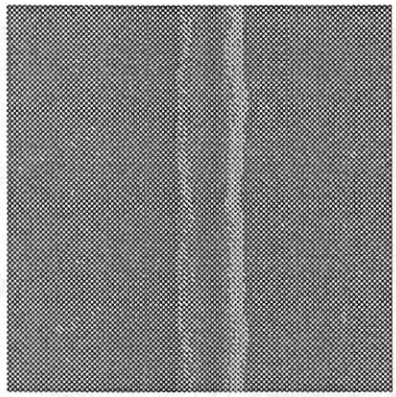
Moat Width=0.6μm σ=0.6 CD=241nm



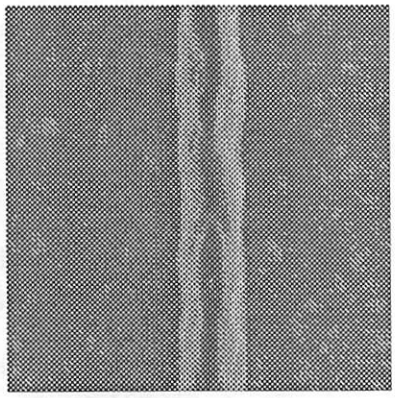
Moat Width=1.0μm σ=0.3 CD=291nm



Moat Width=1.0μm σ=0.6 CD=261nm



Moat Width=2.0μm σ=0.3 CD=241nm



Moat Width=2.0μm σ=0.6 CD=166nm

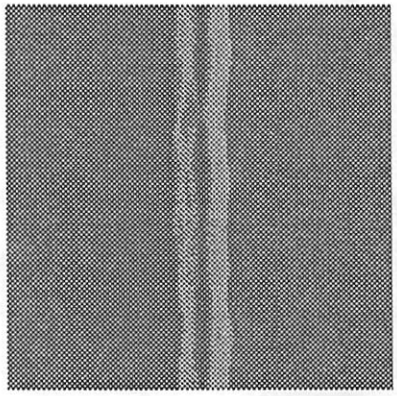
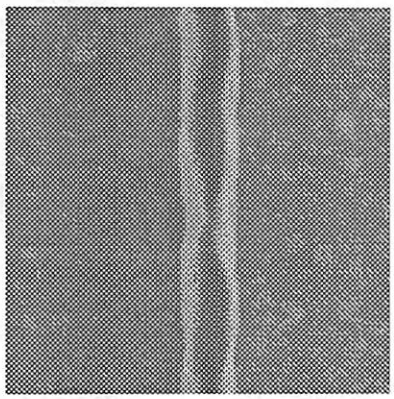


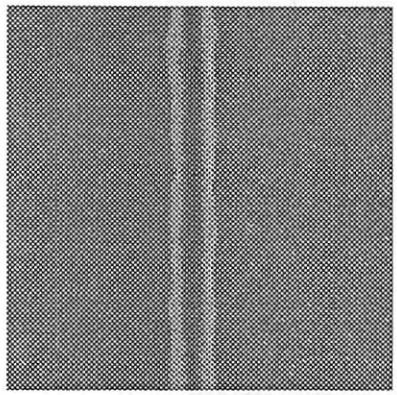
Figure 7-13. Top down SEM's showing the photoresist gate as a function of moat width (0.6, 1.0, and 2.0 μm) and σ.

The CD is recorded as measured at the wafer surface. The CD trends agree with the explanations learned from simulation. The CD on SEM's patterned when σ is 0.3 show more variation when σ is 0.6 due to an increase in the coherent interaction between the light and the polysilicon grains at lower σ.

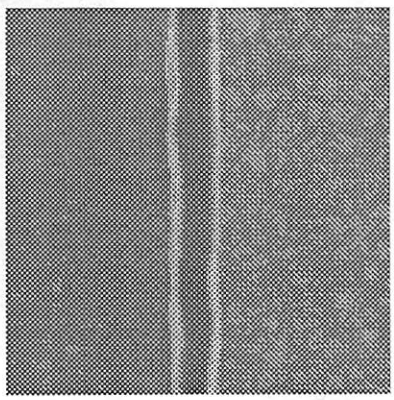
Moat Width=2.5μm σ=0.3 CD=209nm



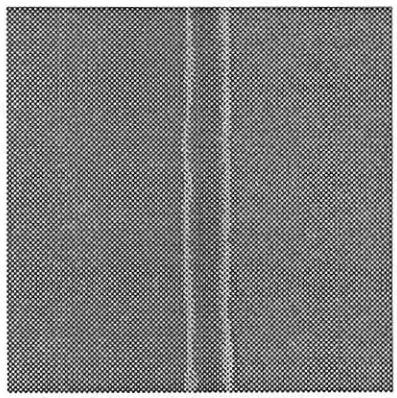
Moat Width=2.5μm σ=0.6 CD=165nm



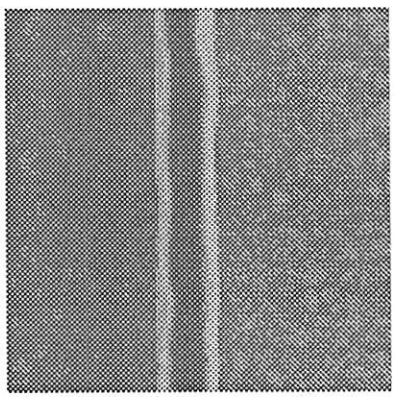
Moat Width=3.0μm σ=0.3 CD=195nm



Moat Width=3.0μm σ=0.6 CD=180nm



Moat Width=4.0μm σ=0.3 CD=225nm



Moat Width=4.0μm σ=0.6 CD=210nm

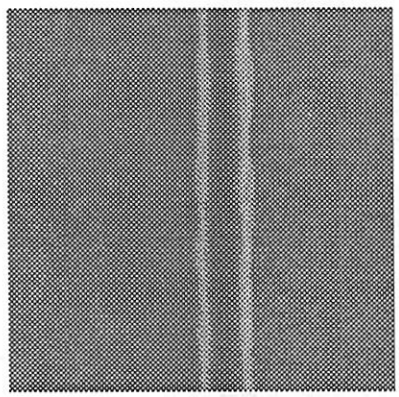


Figure 7-14. Continuation of Figure 7-13.

Top-down SEM's showing the photoresist gate as a function of moat width (2.5, 3.0, and 4.0 μm) and σ. The CD is recorded as measured at the wafer surface.

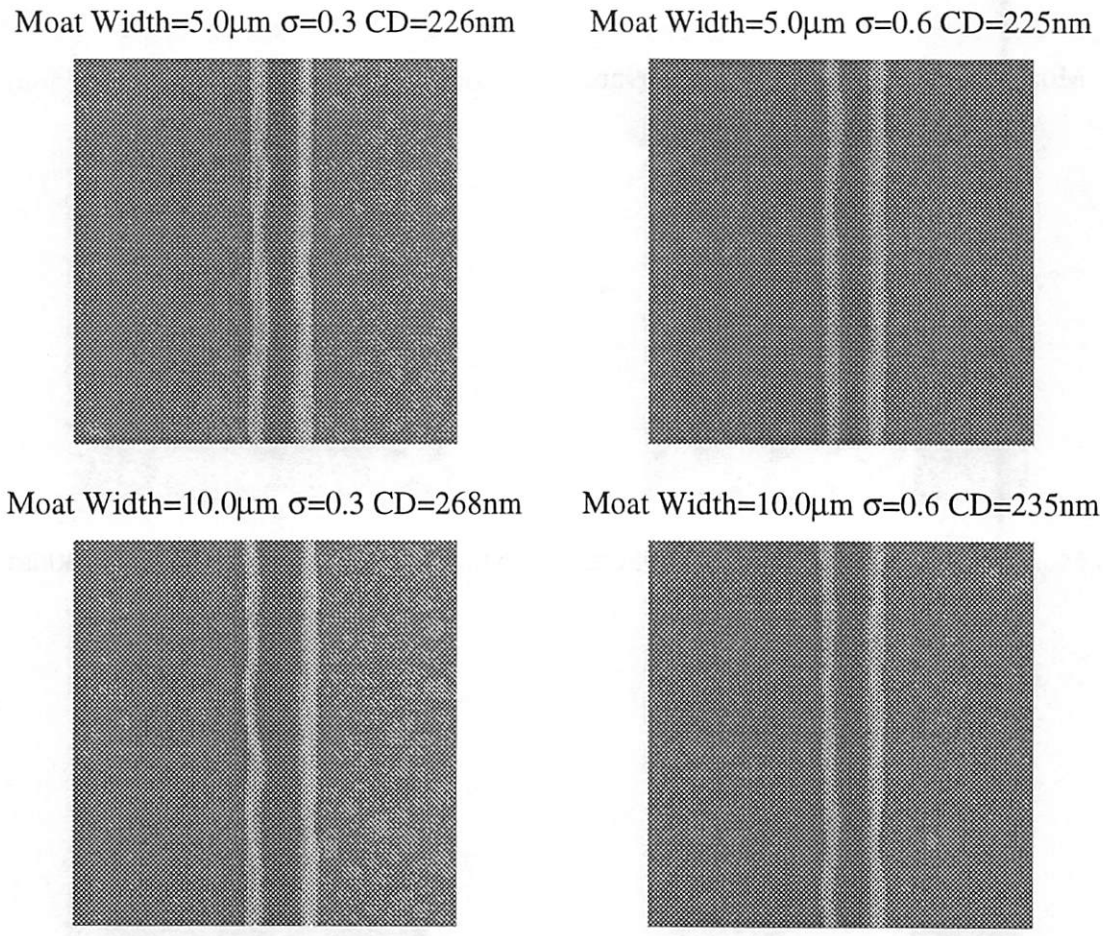


Figure 7-15. Continuation of Figure 7-14.

Top down SEM's showing the photoresist gate as a function of moat width (5.0 and 10.0 μm) and σ. The CD is recorded as measured at the wafer surface.

less CD variation and roughness than gates patterned with σ of 0.3, which is consistent with the experimental results.

7.5. The Influence of the Post-Exposure Bake on Reflective Notching

Although the simulation results prior to a post-exposure bake show that the intensity increases in the unexposed area as σ decreases, the simulations and experimental results after the PEB indicate that σ has little effect on the CD profile. The lack of an effect after the PEB is due to the fact that the PEB diffusion averages the intensity increase over a large area, as seen in the simulations of Figures 7-9 and 7-10. Therefore, because of the PEB diffusion, there is little impact of σ on notching at these narrow moat widths.

To demonstrate the effect of the post-exposure bake on imaging inside silicon, Young's double slit interference experiment [11] was simulated. In Young's experiment, the image of two slits in a opaque mask are illuminated with an incoherent source. This image is projected onto a CCD camera, as shown in Figure 10.4 of Ref. 11. In this simulation of Young's experiment, however, the image is recorded in 1μm of photoresist above a silicon substrate instead of using a CCD camera. In the 2D simulation, the slits are separated by 0.5μm and are illuminated with a 248nm wavelength source with a variable partial coherence factor (σ) of 0.0, 0.3, and 0.6. The light is then collected by a projection optic a NA of 0.5. The optical system used in this simulation is assumed to act like the optic used in a Köhler illumination system; consequently, the two slits are imaged as plane waves incident onto the photoresist. Each of the waves have an angle of incidence of 25^0, which is shown in the schematic of Figure 7-16. The intensity image of these two plane waves in

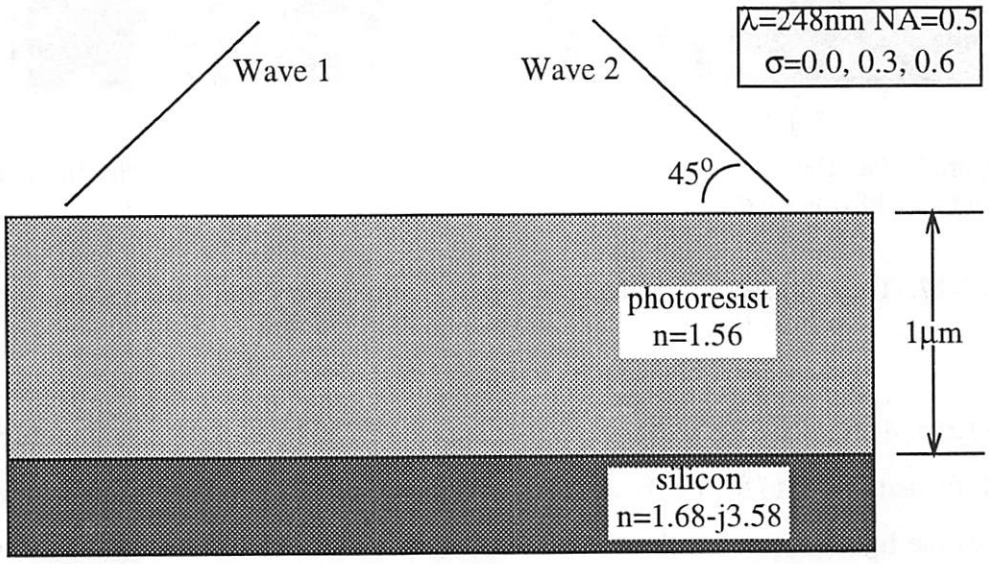


Figure 7-16. Schematic of the two planes produced by Young's experiment imaged in 1μm of photoresist.

photoresist is shown in Figure 7-17 for σ=0.0 (fully coherent light). In Figure 7-17, the fields in the photoresist should be found through the integral technique proposed by Sommerfeld[119] because the silicon is lossy. However, since this example is intended to show the effect of the PEB rather than the fields above a lossy surface, the silicon was assumed

to be a perfect conductor and the intensity in the photoresist was found through image theory [3]. The intensity inside the photoresist prior to the PEB is shown in Figure 7-17a, and the intensity after the PEB, in Figure 7-17b. The intensity after the PEB was calculated with Equation 7-3, in which a 0.04μm diffusion length, *a*, was assumed. Figure 7-17 indicates that there is a large standing wave ratio prior to the PEB; after the PEB, however, the standing waves nearly disappear. In Figure 7-17, the intensity was taken at a constant z=0.5μm cut-line and plotted in Figure 7-18. In Figure 7-17, the intensity is plotted prior

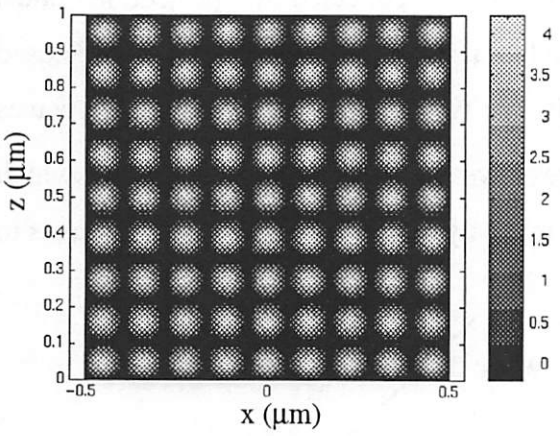


Figure 7-17a. The intensity in the resist before the PEB when σ=0.0

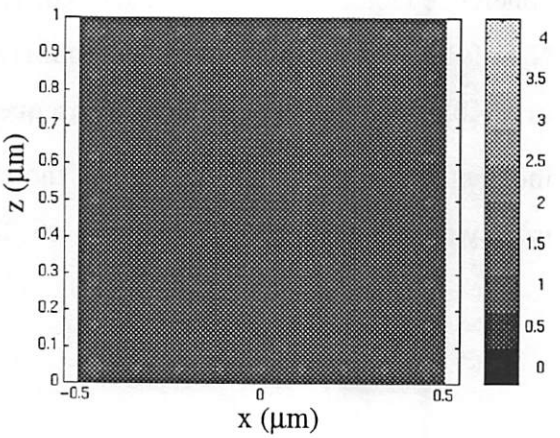


Figure 7-17b. The intensity in the resist after the PEB when σ=0.0

Figure 7-17. The intensity inside the photoresist for coherent illumination (σ=0.0) before a PEB (Figure 7-17a) and after a PEB (Figure 7-17b).

to the PEB and after the PEB when σ is 0.0, 0.3, and 0.6. In Figure 7-18, the standing wave ratio of intensity prior to the PEB decreases as σ increases. This is due to the fact that as σ increases the light is more incoherent, and in the extreme of incoherent illumination, the intensity of the two fields adds, which results in an image without standing waves. The intensity in Figure 7-18 for all the σ cases after the PEB is within ±10% of 1, the clear field intensity. This shows that the post-exposure bake effectively eliminates the standing waves, and the intensity inside the resist is that of an image resulting from incoherent illumination. Therefore, since the image after the PEB with σ=0.0 is approximately the image with both 0.3 and 0.6, as shown in Figure 7-18, coherent illumination (σ=0.0) can be assumed when simulating the image in a chemically-amplified resist.

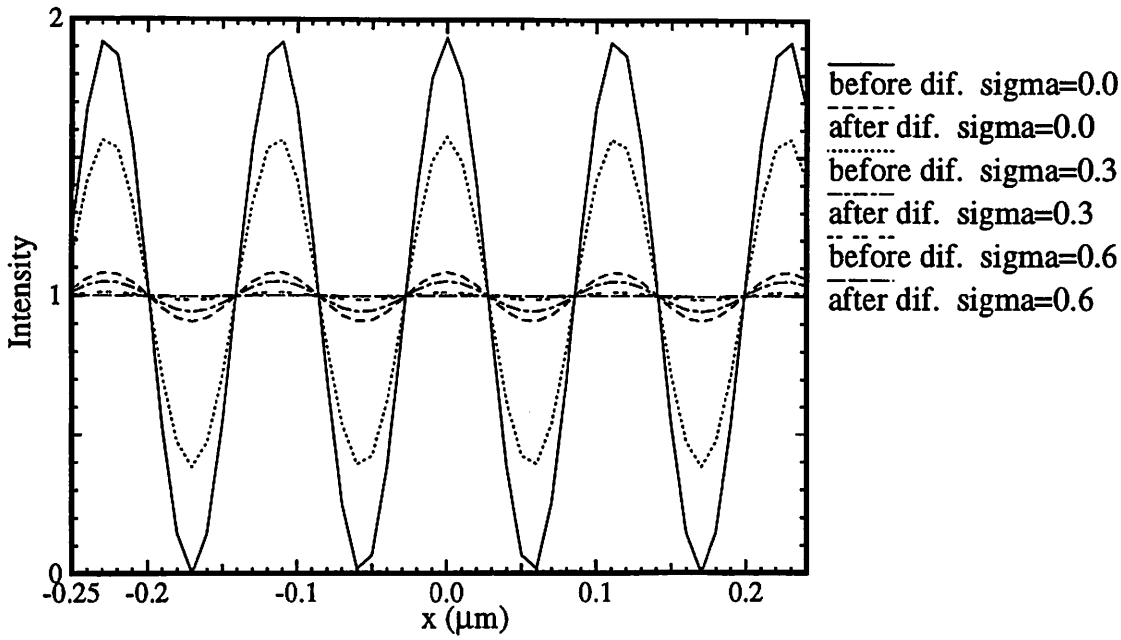


Figure 7-18. The intensity inside photoresist along a cut-line taken in the middle of the resist.

The intensity is plotted and before and after at post-exposure bake for σ of 0.0, 0.3, and 0.6. For all values of σ after the PEB, the intensity is within 10% of 1 which indicates that the PEB removes the standing waves in the resist.

7.6. Conclusions

The goal of this chapter was to define the role of the partial coherence in projection printing through the simulation of the imaging of an attenuated phase-shift mask, the patterning of a line created by a two-layer phase-shifting resist edge, and the patterning of a gate line over an active-area trench

In Section 7.2, a two-dimensional version of TEMPEST-PCD was used to study the effect of partial coherence on imaging a phase-shift mask. The aerial images calculated with TEMPEST predict that the peak intensity decreases from the images calculated with SPLAT. This decrease is due to a breakdown in the scalar assumption used in SPLAT which neglects the currents induced by the vector nature of the light. Although the scalar assumption is invalid for the PSM studied in this section, it is still possible to use Hopkins' formulation through TEMPEST-HN. When the aspect ratio is less than 0.2 (depth to width)

for a feature on a mask, the diffraction orders are independent of incident angle. This implies that for 5X and 4X systems with typical NA's (<0.6) and σ's (0.3 to 0.7), imaging a PSM with the paraxial approximation used in Hopkin's formulation is valid. As a rule of thumb, when the system magnification is 4X and 5X, TEMPEST-HN, which requires one simulation, is accurate; on the other hand, one need only use TEMPEST-PCD, which requires more than one simulation, when the system magnification is 1X and 2X.

In Section 7.3, the feasibility of using a two-layer resist process with a 180^0 phase-shifting edge to print a sub-wavelength line was studied by using TEMPEST-PCD simulations. These simulations showed that it is possible to print a sub-wavelength 200nm line with this process, but this resolution is far greater than the theoretical limit of $\frac{\lambda}{4n}$, which is 40nm for this process. Thickness changes in the resist cause alignment and line width problems. Although this resolution is far from the theoretical limit, the resolution improves with the two-layer process as compared to the resolution with a strong PSM when using a 0.3 threshold model for the resolution. The linewidth with the two-layer process is approximately 21% less than the linewidth with a strong PSM. The phase-shift edge of the top-layer resist causes light to refract into the bottom of the resist into a region directly below the top layer. This refraction is due to a critical angle effect and occurs even when the index of the top-resist layer is matched to the index of the bottom-resist layer. This critical angle effect causes asymmetry in the intensity profile which may cause asymmetry in the developed resist line. The refraction at the edge also causes the 25nm shift in alignment of the line with the top-layer resist edge. Finally, $\pm 10\%$ thickness variation in the top-layer resist cause as much as $\pm 5\%$ linewidth variation and 25nm shift in alignment, while the slope of the phase edge was shown to produce little change in the linewidth and alignment. Consequently, the thickness of the top layer resist must be tightly controlled in order to limit changes in the line width and in alignment. Since topography on the wafer may produce changes in the thickness of the spun on resist, the wafer topography plays a significant role in limiting the performance of this two-layer process and the topography should not be neglected in future studies.

As shown in Section 7.4, TEMPEST-PCD simulations also accurately modelled the patterning of a gate over an active-area trench. The image inside the resist before the PEB predicted that the intensity contrast increases for hot spots in the unexposed area as σ decreases. When the PEB diffusion is Fickian, however, the PEB diffusion low pass filters the initial PAC concentration and eliminates these hot spots and standing waves in the resist. The PEB diffusion eliminates these interference effects when the diffusion length is near the theoretical resolution limit of the resist, $\frac{\lambda}{4n}$. For example, in the profiles of Section 7.4 a diffusion length of 40nm was used, which is equivalent to the theoretical resolution limit. Consequently, reflective notching in the resist is eliminated. When the diffusion length of a Fickian process is near the theoretical resolution limit, the profiles with partially coherent illumination after the PEB are similar to the profiles that would be generated with incoherent illumination. This implies that simulation of the partial coherence is not necessary, and coherent illumination can be assumed when the diffusion length is near the theoretical resolution limit.

This chapter presented TEMPEST-PCD simulation results for a lithography system. These results are expanded in the next chapter, Chapter 8, where TEMPEST-PCD results are presented for a wafer inspection system.

8 The Effect of Partial Coherence on Wafer and Mask Inspection

8.1. Introduction

In an optical inspection system, a condenser optic with a large numerical aperture images light from an incoherent source onto a wafer or onto a mask. The large numerical aperture produces light that is obliquely incident at large angles. This obliquely incident light reflects off the topography of the wafer or of the mask and is collected by the same condenser optic, i.e., σ=1. Since the van Cittert-Zernike theorem, discussed in Chapter 3 implies that obliquely incident light is partially coherent, the illumination incident on the wafer or on the mask is partially coherent as well. The partial coherence of the light causes it to scatter in unexpected ways off the wafer or off the mask. For example, when inspecting for a defect at the bottom of a thin-film stack, the changes in refractive index from the stack induce aberrations which degrade the optical system resolution. In addition to these induced thin-film aberrations, the topography of the inspected structure can cause the light to refract out of areas of interest leaving these areas unresolved.

In printing into a thin-film stack, Flagello[33], Yeung[157], and Progler[102] have shown that spherical aberration is induced by the photoresist when printing with a high numerical aperture stepper. Yeung showed that the effect of arbitrary thin-film layers on imaging can be modelled by ray tracing through the thin-film layers to find the optical path difference (OPD). This OPD induces aberrations which can be modelled by modifying the pupil function of the imaging system. Similar to this work, Progler then proposed using a merit function that is capable of determining which aberrations may actually benefit printing.

Since some aberrations may actually help in printing certain features, it is possible to deliberately introduce beneficial aberrations by using a pupil filter. One type of pupil filter that has received considerable study is a filter that improves printing through focus. In these studies, Tsujiuchi[130] first described a cosine filter to improve depth of focus. This filter was later re-invented by Ojeda-Castañeda[92] and was implemented by Fukuda[34]. By using McCutchen's theorem [76], Ojeda-Castañeda designed filters for increased depth of focus [88,89]. Alternatively, von Bünau[15-17] designed a filter for increased depth of focus by using the decomposition technique described in Section 2.2.3.

In this chapter, the aberrations introduced by printing and by inspecting through a thin-film stack are first discussed in Section 8.2. These aberrations degrade both the resolution of the printing system and of the inspection system. The inspection of a thin-film stack is generalized to include topographical structures in Section 8.3. In Section 8.3, the inspection of a contact hole is studied through simulation. A methodology for improving this contact hole inspection by introducing a filter into the optical system is presented in Section 8.4. The effect of the partial coherence on the filter performance is studied in Section 8.4.4 and the effect of topography changes induced by process variation in Section 8.4.5.

8.2. Aberrations Induced by Inspecting Through a Thin-Film Stack

In optical lithography, a mask must be imaged into a thin-film stack whereas in optical inspection an object under the thin-film stack must be found. Due to refractive index changes in the thin-film stack, rays from different annuli of the optical system are refracted, different amounts leading to spherical aberration. The amount of spherical aberration can be found by ray tracing through the stack and finding the optical path difference introduced by the stack. This is demonstrated in Figure 8-1 for an air-to-glass interface. Figure 8-1 models a glass layer over a silicon substrate in which the reflection from the substrate is neglected as the glass is assumed to be infinite in the z direction. The optical path difference induced by the glass can be understood through the use of an analogy. In this analogy, a fish that is swimming a distance, w, below the surface would actually be perceived by an observer outside the water to be swimming a distance, d, below the sur-

face. The difference in path length that the light must travel between the true location of the fish and the perceived location multiplied by the index of refraction of the water is the optical path difference. This optical path difference depends on the index change from the air to the water, on the distance from the interface to the fish, and on the angle of incidence onto the interface which is related to the numerical aperture of the optical system.

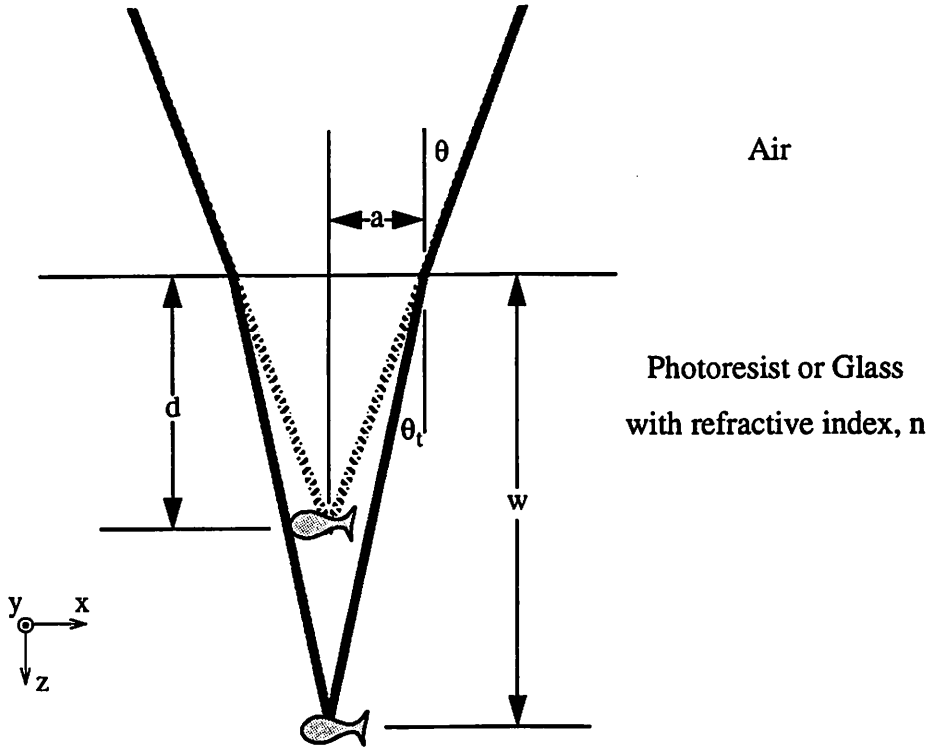


Figure 8-1. Schematic showing the optical path difference when printing or inspecting into a dielectric.

This optical path difference introduces aberrations into the optical system. The optical system can take these aberrations into account by modifying the pupil function to include an optical path difference term, Φ, as discussed in Section 3.4. In Section 3.4, the pupil function is given by Equation 3-19, which is repeated in Equation 8-1 where the defocus term has been included in Φ. The optical path difference term, Φ, is calculated by

Equation 8-1. $$P(\mathbf{x}_p) = P(x_p, y_p) = exp\left[i\frac{2\pi}{\lambda}\Phi(\tilde{x}_p, \tilde{y}_p) \right] circ\left(\sqrt{x_p^2 + y_p^2}\right)$$

ray tracing into the thin-film stack as shown in Equation 8-2, where x_2 is the actual path

length and x_1 is the induced path length. Since $a = d\tan\theta$, Equation 8-2 simplifies to Equation 8-3. In addition to the optical path difference induced by the stack, $W_s(\theta)$, the

Equation 8-2. $$W_s(\theta) = n(x_2 - x_1) = \frac{na}{\sin\theta_t} - \frac{na}{\sin\theta} = \frac{n^2 a}{\sin\theta} - \frac{na}{\sin\theta}$$

Equation 8-3. $$W_s(\theta) = -n(n-1)d\sec\theta$$

defocus error, Δ, induces some aberration, and the amount of this aberration is given as $W_d(\theta)$ in Equation 8-4. The total aberration, $W_t(\theta)$, is the sum of $W_s(\theta)$ and of $W_d(\theta)$ as

Equation 8-4. $$W_d(\theta) = \frac{n\Delta}{2}\tan^2\theta$$

shown in Equation 8-5. The angular coordinates, θ, used in Equation 8-5 can also be converted to radial pupil coordinates, ρ, shown in Equation 8-6 by using the fact that $\tan\theta = \sqrt{\frac{NA^2}{1-NA^2}}\rho$.

Equation 8-5. $$W_t(\theta) = \frac{n\Delta}{2}\tan^2\theta - n(n-1)d\sec\theta$$

Equation 8-6. $$W_t(\rho) = \frac{\Delta}{2}\left(\frac{NA^2}{1-NA^2}\right)\rho^2 - n(n-1)d\sqrt{1+\left(\frac{NA^2}{1-NA^2}\right)\rho^2}$$

By using Equation 8-6, the Strehl ratio[10], which is a measure of the optical system aberrations, can be calculated as a function of NA. This ratio is defined in Equation 8-7 and has a value between 0 and 1, where 1 corresponds to a perfect unaberrated system. The defocus, Δ, offers a degree of freedom to choose the best focal plane such that the Strehl ratio due to $W_t(\theta)$ is minimized.

Equation 8-7. $$U(r,\psi) = \frac{\left|\int_0^{2\pi} d\phi \int_0^{\sin^{-1}NA} d\theta\tan\theta\exp\left[jkW(\theta,\phi)\right]\right|^2}{\left|\int_0^{2\pi} d\phi \int_0^{\sin^{-1}NA} d\theta\tan\theta\right|^2}$$

Here, after finding this optimal defocus, the Strehl ratio as a function of numerical aperture is plotted (see Figure 8-2) for an air-to-glass interface. In Figure 8-2, the ratio is

calculated at various depths, d=0.2, 0.4, 0.6, 0.8, and 1.0μm, into the glass, where the glass has a refractive index of 1.47453 and is illuminated with light having a wavelength of 365nm. The Strehl ratio in Figure 8-2 does not begin to deviate from 1 until the NA

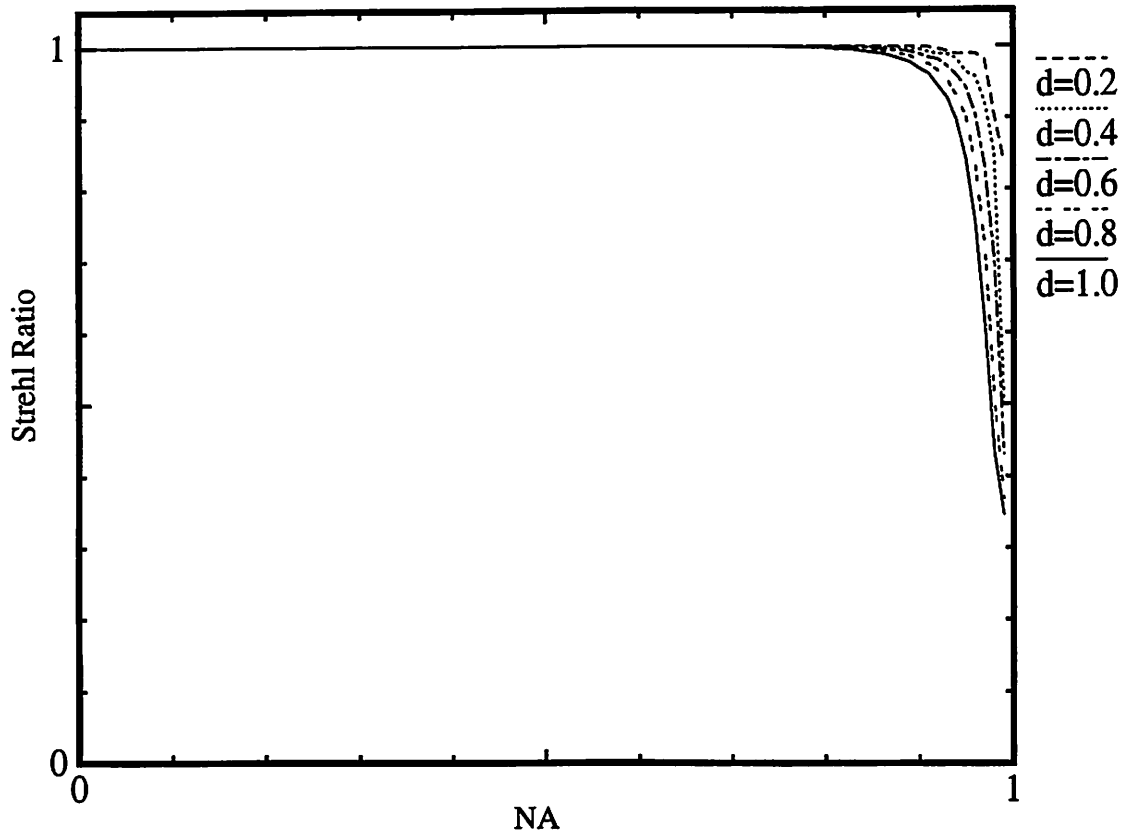


Figure 8-2. The Strehl ratio as function of NA when printing or inspecting through glass.

The ratio is plotted for an air-to-glass interface at various depths, d, into the glass of 0.2, 0.4, 0.6, 0.8, and 1.0μm. The glass has a refractive index of 1.47453 and a 365nm wavelength optic is used.

approaches 0.8. Figure 8-2 indicates that the Strehl ratio decreases as NA increases and indicates that the ratio decreases as the depth, *d*, into the glass increases. Only when d is 1.0μm and NA is 0.9 does the Strehl ratio drop below 0.96. Therefore, the Strehl ratio predicts that an inspection system with an NA of 0.9 is capable of imaging through about 1.0μm of glass before the aberration induced by glass begins to degrade the image quality.

Since the Strehl ratio is a general figure of merit that indicates the resolution of the inspection system, it does not necessarily predict the overall resolution of the system. Con-

sequently, a better indication of the resolution is given by the image transfer function of the system, K. The image transfer function is equivalent to the field produced by a point source located in the object plane. The image transfer function was given by Equation 3-12, which is repeated in Equation 8-8 in angular coordinates, θ. Using Equation 8-8, the

Equation 8-8.

$$K(r,\psi) = \int_0^{2\pi} d\phi \int_0^{\sin^{-1}NA} d\theta \tan\theta \exp\left[jkW(\theta,\phi)\right] exp\left[-jkr\sin\theta\cos\left(\phi-\psi\right)\right]$$

magnitude and phase of K is plotted as a function of r (in μm) in Figure 8-3 for a 365nm wavelength illumination optic having a NA of 0.9 at a depth, *d*, into the glass of 0.2, 0.4, 0.6, 0.8, and 1.0μm. In Figure 8-3, the magnitude of the transfer function is relatively unaffected by imaging deeper into the glass. However, the phase of the transfer function begins to show significant deviation from an ideal square wave when *d* is 0.8μm and is most pronounced when *d* is 1.0μm. These slight changes in the phase of *K* are not negligible since the phase has a large impact on image quality.

In addition to using the image transfer function to measure optical system performance, the total optical path difference, $W_t(\rho)$, can also be represented as a sum of Zernike polynomials [10]. Since the interface in Figure 8-1 is assumed to be infinite in the *x* and *y* directions and since the optical system is rotationally symmetric, the aberrations induced by the interface are rotationally symmetric as well. This implies that the aberrations are spherical aberrations only. These spherical aberrations caused by imaging into the glass are shown in Table 8-1. In Table 8-1, the piston and defocus aberration has been omitted since these are correctable by moving the plane of focus. The first four non-correctable spherical aberrations are displayed along with the total rms aberration. In an optical lithography system, the total rms aberration should be less than 0.02λ in order to maintain high resolution [147]. Since the total rms aberration at a depth of $\geq 1.0\mu m$ into the glass is greater

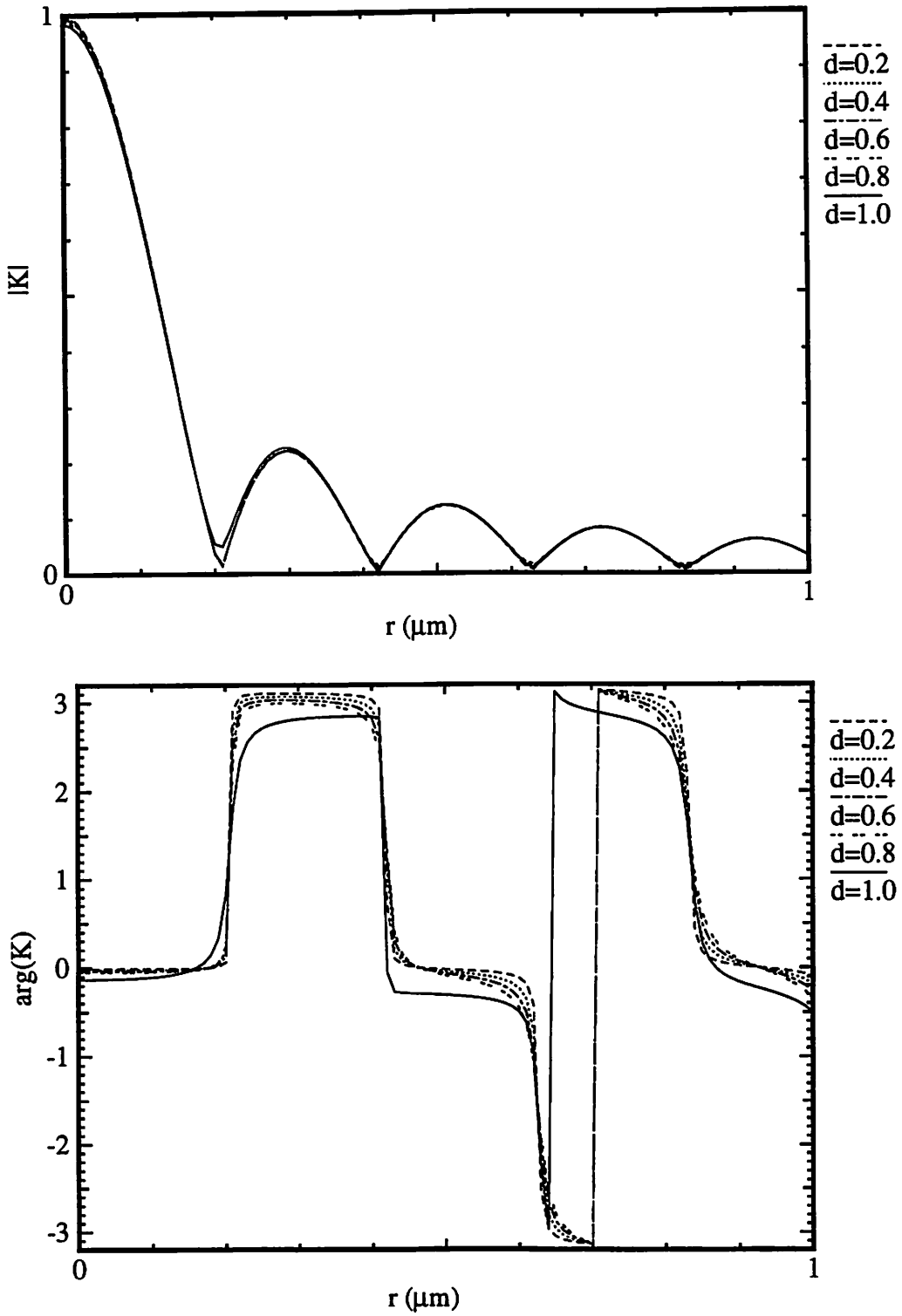


Figure 8-3. The magnitude and phase of the image transfer function, K.

The image transfer function is plotted for 365nm wavelength illumination optic having a NA of 0.9 at a depth, d, into the glass of 0.2, 0.4, 0.6, 0.8, and 1.0μm.

than 0.02λ, the inspection system may be incapable of producing high enough resolution to properly image a defect in the glass.

Table 8-1. First four non-correctable spherical Zernike polynomials.

The polynomials are calculated as a function of depth, d, and the total rms aberration is calculated.

depth, d, in μm	spherical aberrations a_n^m				RMS
	a_2^0	a_3^0	a_4^0	a_6^0	
0.2	0.0029λ	-0.0012λ	0.0008λ	-0.0004λ	0.0032λ
0.4	0.0057λ	-0.0024λ	0.0016λ	-0.0007λ	0.0065λ
0.6	0.0086λ	-0.0036λ	0.0024λ	-0.0011λ	0.0097λ
0.8	0.0115λ	-0.0048λ	0.0032λ	-0.0015λ	0.0129λ
1.0	0.0165λ	-0.0073λ	0.0045λ	-0.0025λ	0.0188λ

8.3. Contact Hole Inspection

The inspection of a thin-film stack discussed in the previous section is generalized in this section to include the topographical structure of a contact hole. The goal of the contact hole inspection system is to locate defects at the bottom of the hole which may form from underetching. If a defect is found, the source of the error can be corrected prior to processing more wafers, which amounts to a savings of time and money.

The contact hole studied in this chapter is formed by plasma etching a hole into a layer of dielectric isolation such SiO_2 that has been deposited over a silicon substrate. In this study, the hole is perfectly cylindrical having a diameter of 0.25μm and stops at the silicon and glass interface. The resulting contact hole topography is shown pictorially in Figure 8-4. The inspection system illuminates the hole with 365nm wavelength light that has been imaged with a 0.9 NA condenser optic. This incident light is scattered in all directions and the reflected light is collected by the same optic used to illuminate the hole, i.e., the optical system is a reflective system and σ is 1. This collected light is projected onto the image plane. If a defect is present at the bottom of the hole, a well designed inspection system should be able to distinguish between the aerial image with a defect and the aerial

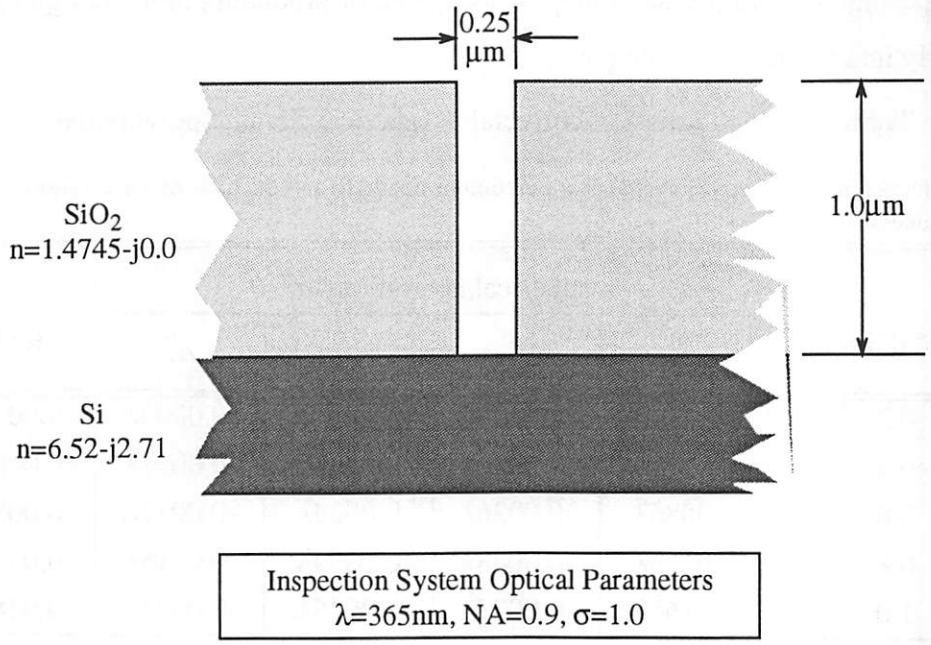


Figure 8-4. Schematic of the contact hole geometry to be inspected.

The hole is etched in 1.0μm of glass over a silicon substrate and is inspected at 365nm with the optic having a NA of 0.9 and σ of 1.0.

image without a defect. In order for this inspection system to distinguish these changes in the aerial image, the optical system must first deliver light to the bottom of the hole.

Inspection of a contact hole is difficult because of vertical propagation in the hole and in the oxide. However, the topography of the etched hole causes normally incident light to refract out of the hole. This is shown in Figure 8-5. The normally incident plane wave used in Figure 8-5 is generated by imaging a single plane wave (σ=0) onto the wafer topography. This normally incident plane wave is refracted out of the hole, producing an intensity that is nearly zero throughout the hole. This refraction is due to a critical angle effect. Since the light is incident at an angle of 90^{o} with respect to the sidewall, the light is refracted into the glass with an angle equivalent to the critical angle. Because of this refraction, virtually no light reaches the bottom of the hole. Since little if any light reaches the bottom, inspection of a defect at the bottom of the hole is impossible.

Since coherent illumination does not assist imaging the bottom of the hole because of critical-angle refraction, applying obliquely incident illumination may overcome this

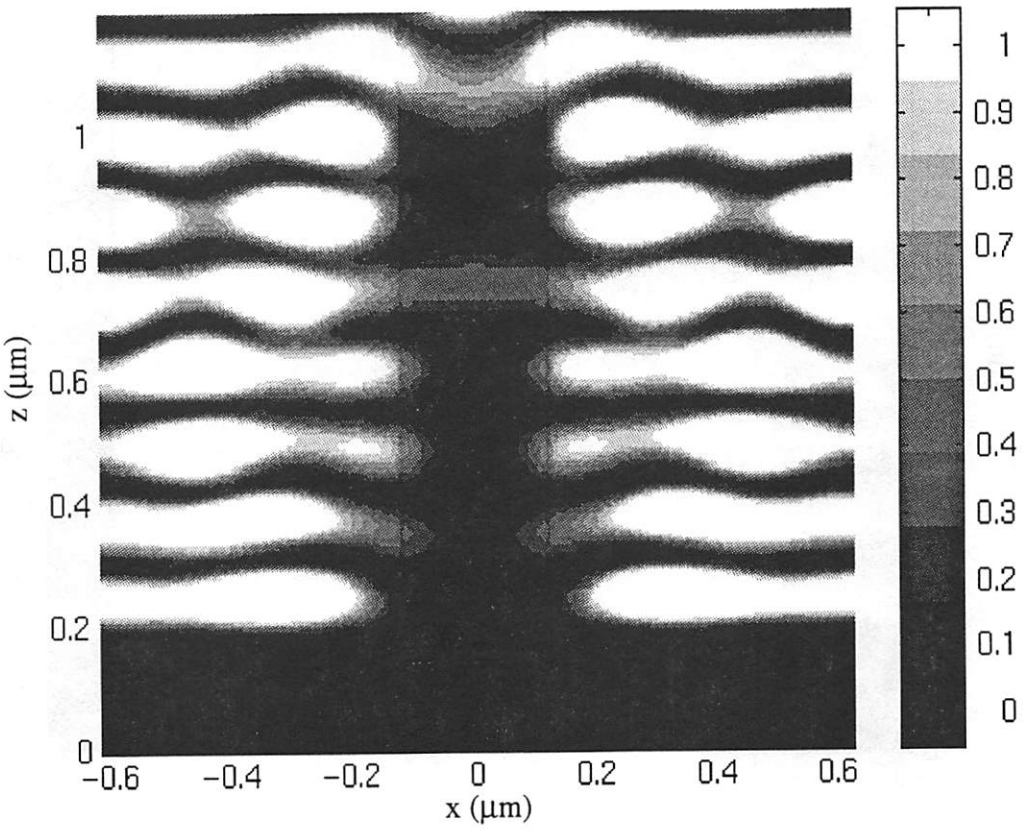


Figure 8-5. Intensity inside the contact hole when illuminated with a single normally incident plane wave.

Note that the hole topography refracts light out of the hole and little light reaches the bottom of the hole.

and enable imaging the bottom of the hole. Obliquely incident illumination is produced in the optical system by increasing the size of the illumination source (i.e., increasing σ). Here, the size of the illumination source was increased to from σ of 0 to σ of 1. The scattering of this obliquely incident light produced by this larger source was simulated with TEMPEST by using Abbe's formulation. The intensity profile from the simulation is plotted in Figure 8-6. Comparing Figure 8-5 and Figure 8-6, we observe that after appropriate normalization, the intensity at the bottom of the hole increases when σ=1 (as in Figure 8-6) as opposed to when σ=0 (as in Figure 8-5). This increased intensity, however, is due to a standing wave interaction with the substrate.

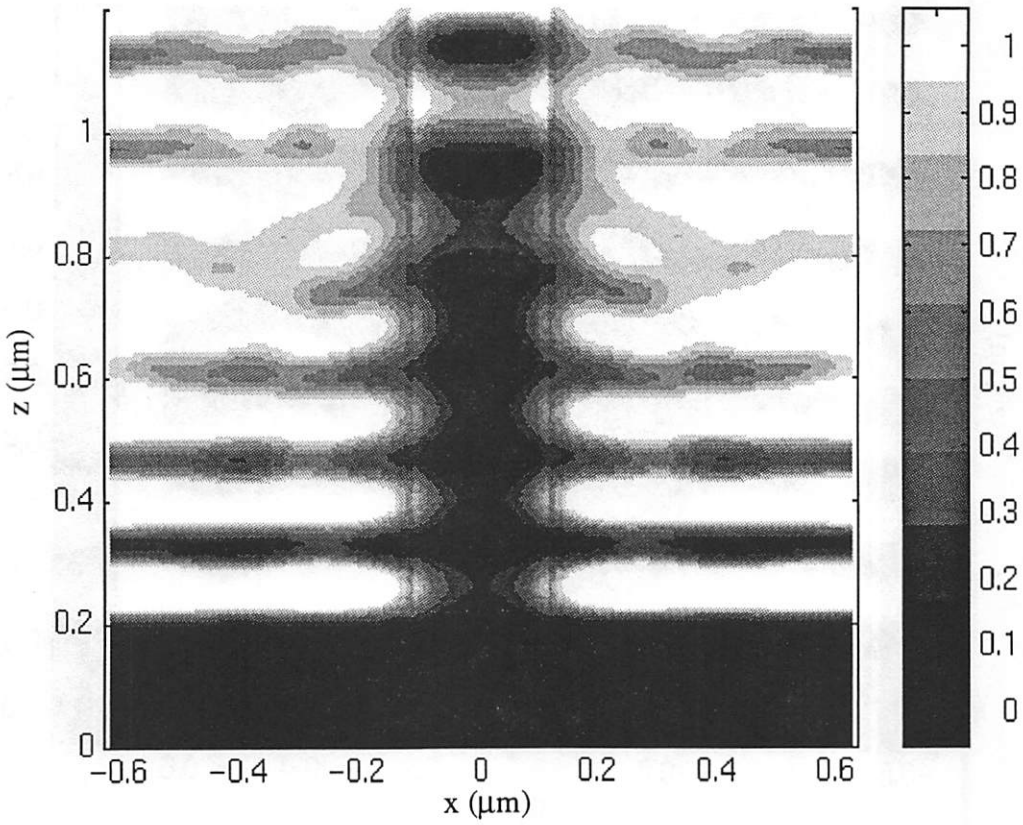


Figure 8-6. Intensity inside the contact hole when illuminated with a partially coherent source with σ of 1.0 and NA of 0.9.

8.4. Filter Design to Optimize Contact Hole Inspection

In the previous section, it was shown that using obliquely incident light from a larger source results in light at the bottom of the hole. In this approach, these waves are individually incoherent and are incapable of delivering substantial light to the bottom. Another approach consisting of a collection of coherent excitations may result in more light at the bottom. These coherent excitations are found by using a synthesis approach based on the reciprocity theorem.

8.4.1. Point Source Excitation and the Reciprocity Theorem

A point source in free space creates a spherical wavefront, and the converse is true as well, that is, a point image is generated by time reversing the propagation of a spherical wavefront. This fact is known as the reciprocity theorem. This theorem can be used in contact hole inspection to synthesize an excitation field that most efficiently guides light to the

bottom of the hole. This optimal excitation is found by placing a point source at the bottom of the hole and measuring the fields leaving the hole in a plane that is parallel and above the SiO_2-air interface, as shown in Figure 8-7. Extending the reciprocity theorem, we hypothesized that the excitation field that most efficiently guides light to the bottom of the hole is created by time reversing the field leaving the hole.

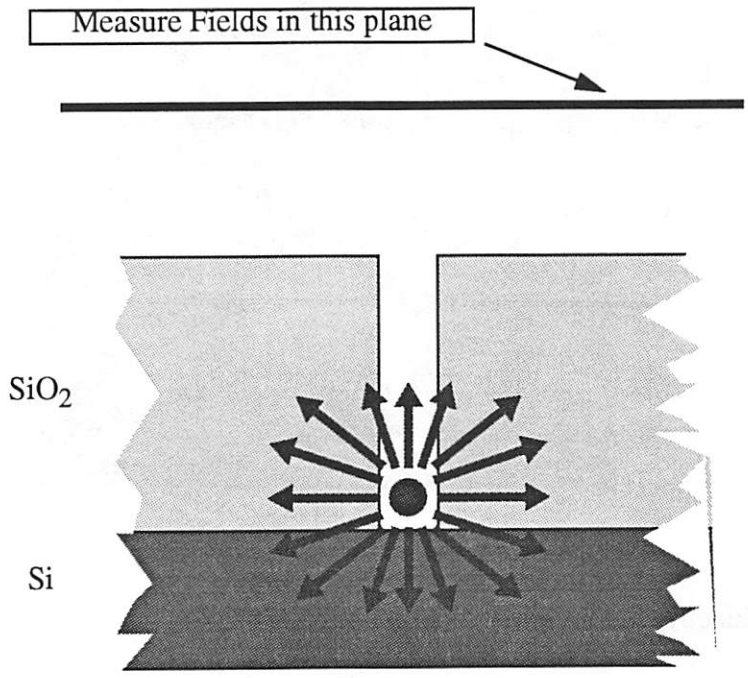


Figure 8-7. Schematic demonstrating the reciprocity principle.

The radiated fields are measured in the plane shown by the thick black line. From the reciprocity theorem, these fields are taken to represent the optimal excitation needed to produce a large intensity at the bottom of the hole.

Using TEMPEST, a point source excitation was placed at the bottom of the hole and allowed to radiated. Figure 8-8 depicts the radiated intensity for a point source having an electric field polarized in the y direction in the xz plane. The radiated E_y field leaving the hole was measured in an xy plane one wavelength above the SiO_2-air interface; the magnitude and phase of this field is plotted in Figure 8-9. This field has been low passed filtered such that the field contains only the spatial frequencies that an NA of 0.9 illumination optic is capable of producing.

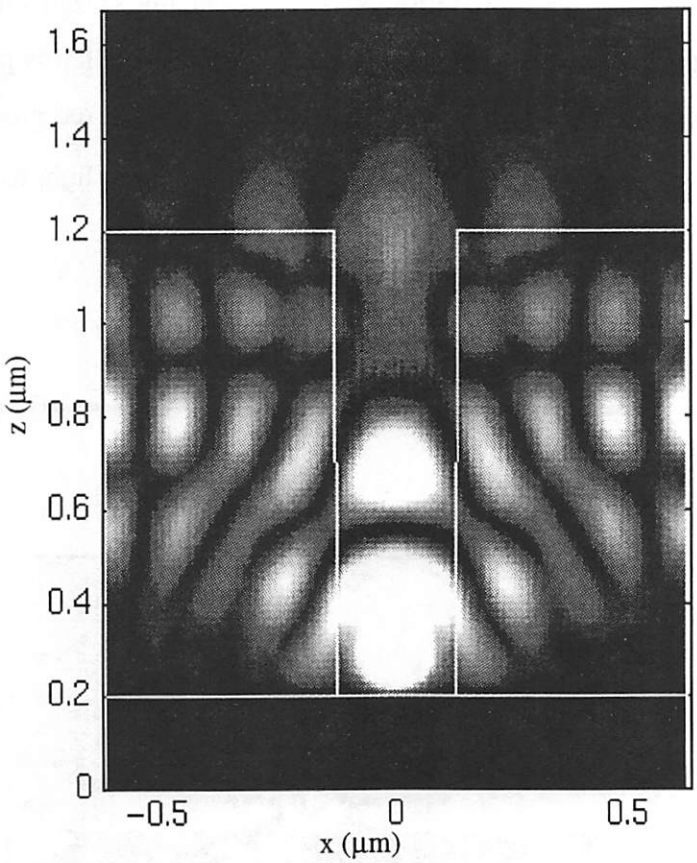


Figure 8-8. Electric field in the y direction radiated from a point source placed at the bottom of the hole.

8.4.2. Magnitude and Phase of the Filter

The radiated fields exiting the hole in Figure 8-9 can be time reversed and allowed to propagate back into the hole. According to the reciprocity theorem, these fields should produce a field that efficiently guides light to the bottom of the hole. These fields can also be propagated to a reference plane in the optical system by using Green's theorem. By propagating these fields to any plane, a filter that has both magnitude and phase can be designed. However, the magnitude of the filter is generated by using an absorbing material in the construction of the filter. This absorbing material would heat during the operation of the inspection system, thus causing the properties of the filter to change. For this reason, the filter is forced to be phase only. Since the field can be propagated to any plane, a filter that is situated in either the exit pupil plane or in the object plane of the optic can be constructed. When a phase-only filter is placed in the exit pupil plane, the light passes through

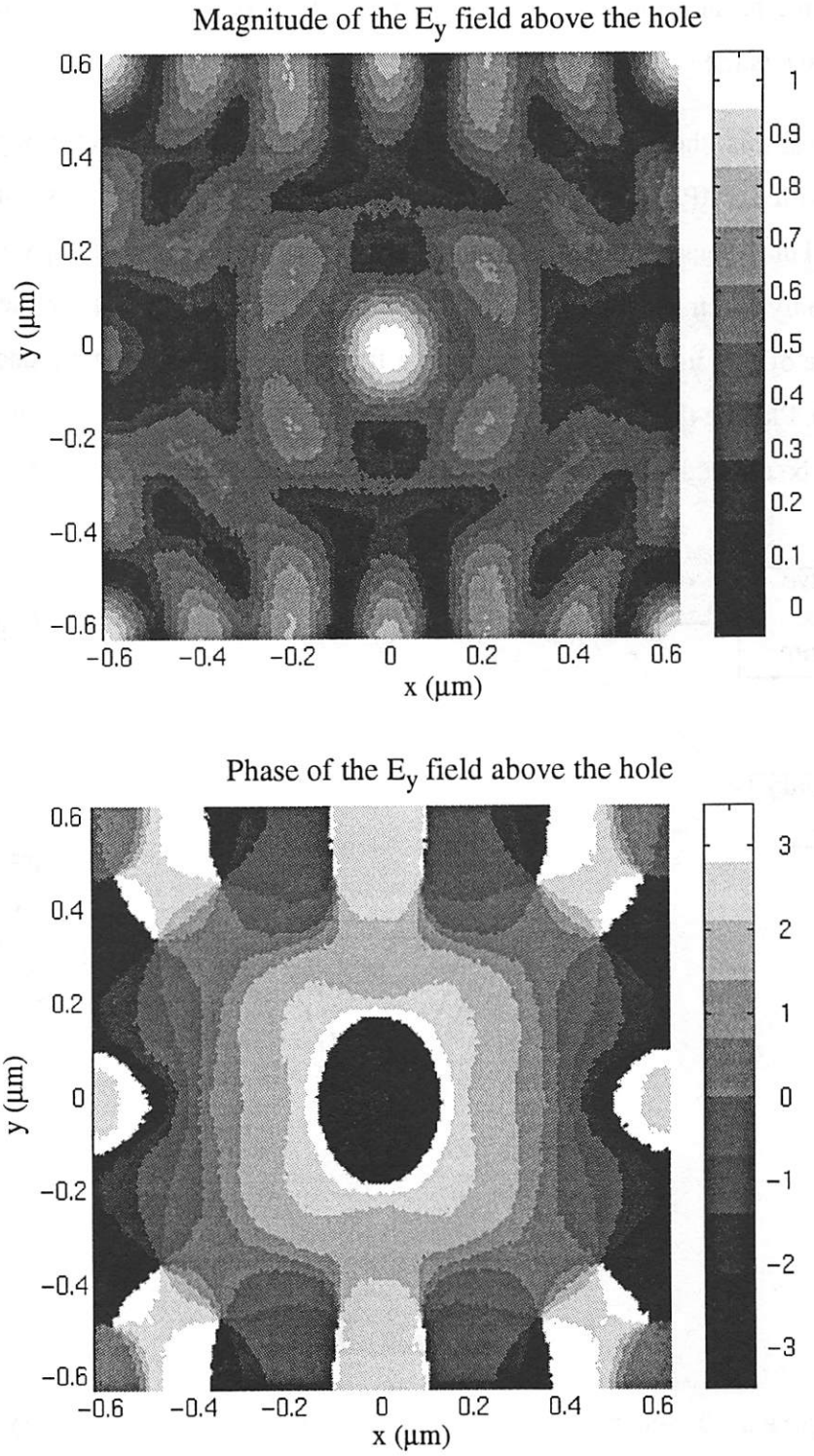


Figure 8-9. Magnitude and phase of the E_y field measured in the plane above the hole shown in Figure 8-7.

the filter twice because the condenser optic in a reflective system projects light onto the hole and collects the scattered light.

To test this, the effectiveness of placing a phase filter in the exit pupil plane was simulated with TEMPEST and was found to be less effective than a phase filter placed in the object plane. Consequently, in the simulation results presented in this chapter, results are shown only when the filter is placed in the object plane rather than in the pupil plane. A schematic of this inspection system with a filter placed in the object plane is shown in Figure 8-10. Placing the filter in the object plane, however, does have one disadvantage in that it must be aligned to the wafer.

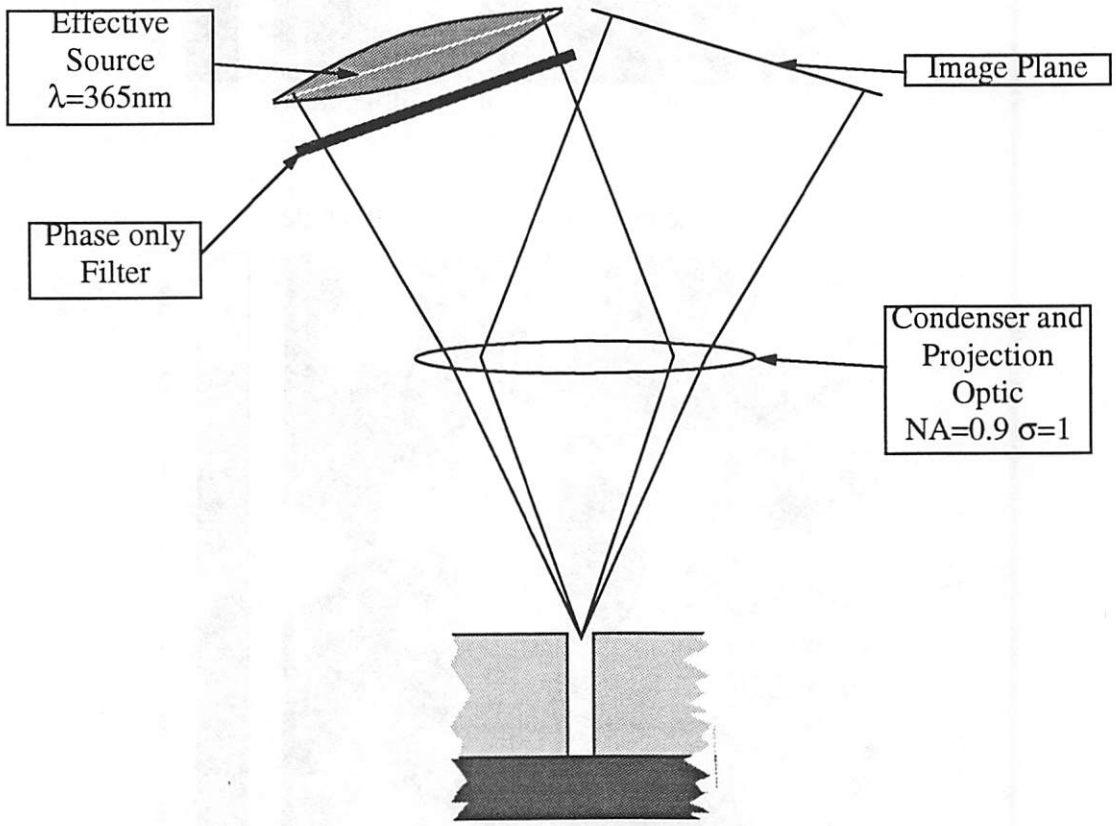


Figure 8-10. Schematic showing the reflective optical inspection system.

The 365nm wavelength source is modulated by the phase filter which is imaged onto the hole by an 0.9 NA condenser optic. The reflecting from the hole is collected by the same condenser optic and is projected onto the image plane.

Using this phase-only filter located in the object plane, the incident fields were calculated by imaging the filter to an excitation plane situated in an xy plane above the hole topography. These incident fields are plotted in Figure 8-11. In Figure 8-11, the fields have approximately the same phase relationship as the phase-only filter prior to its projection through the illumination optic of Figure 8-9. The phase transitions in Figure 8-11, however, are not as sharp as the transitions in the phase-only filter shown in Figure 8-9. Furthermore, spatial resolution in the magnitude of that observed in Figure 8-11 is reduced slightly as compared to the spatial resolution of the radiated field leaving the hole as in Figure 8-9. These variations in magnitude and phase of the incident field from the radiated field degrade the performance of the filter.

8.4.3. Coherent Fields Inside the Contact Hole Topography

In this section, the light incident upon the contact hole topography is first assumed to be fully coherent. The fully coherent light is generated by replacing the effective source found in Figure 8-10 with a point source (σ=0). With this point source in conjunction with the phase filter, an excitation field that is fully coherent is generated, as shown in Figure 8-11. The scattering of this coherent excitation from the hole topography was then simulated using TEMPEST. We find that the intensity in the contact hole due to this coherent excitation is shown in Figure 8-12 in the xz plane. The intensity at the bottom of the hole using the filter is significantly larger than the intensity without a filter under fully coherent illumination (σ=0) as was shown in Figure 8-5. The intensity at the bottom of the hole with the filter and when σ=0 is approximately equivalent to the intensity in Figure 8-6, which was produced by a source with σ=1. This indicates that the filter uses the illumination from a point source as efficiently as the illumination from a large source. In addition to improving the intensity at the bottom of the hole, the filter conditions the light such that it is guided down the hole. This guided light creates a leaky propagating mode in the hole.

This hypothesis was tested by depositing a 0.1μm thick layer of absorbing aluminum over the oxide, and then the hole was etched. TEMPEST simulation results when illuminating this structure with filtered coherent light (σ=0) are plotted Figure 8-13. The aluminum prevents the refraction through the oxide due to off-axis illumination; consequently, only light that is guided propagates in the hole as a leaky mode. Due to this guid-

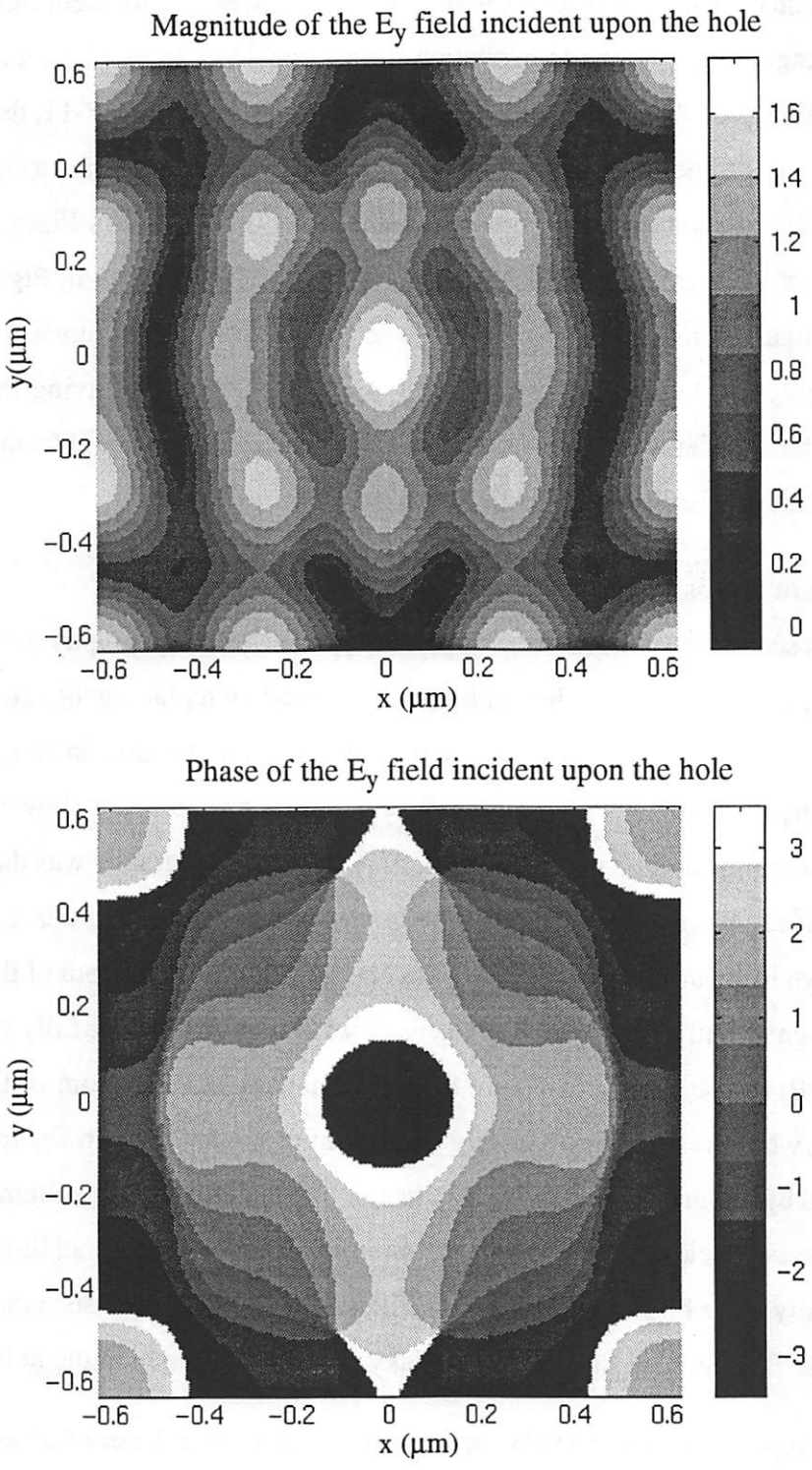


Figure 8-11. Magnitude and phase of the E_y. field incident upon the contact hole a wavelength above the hole.

The field is produced by modulating the source illumination with a phase-only filter and projecting it upon the wafer with the condenser optic.

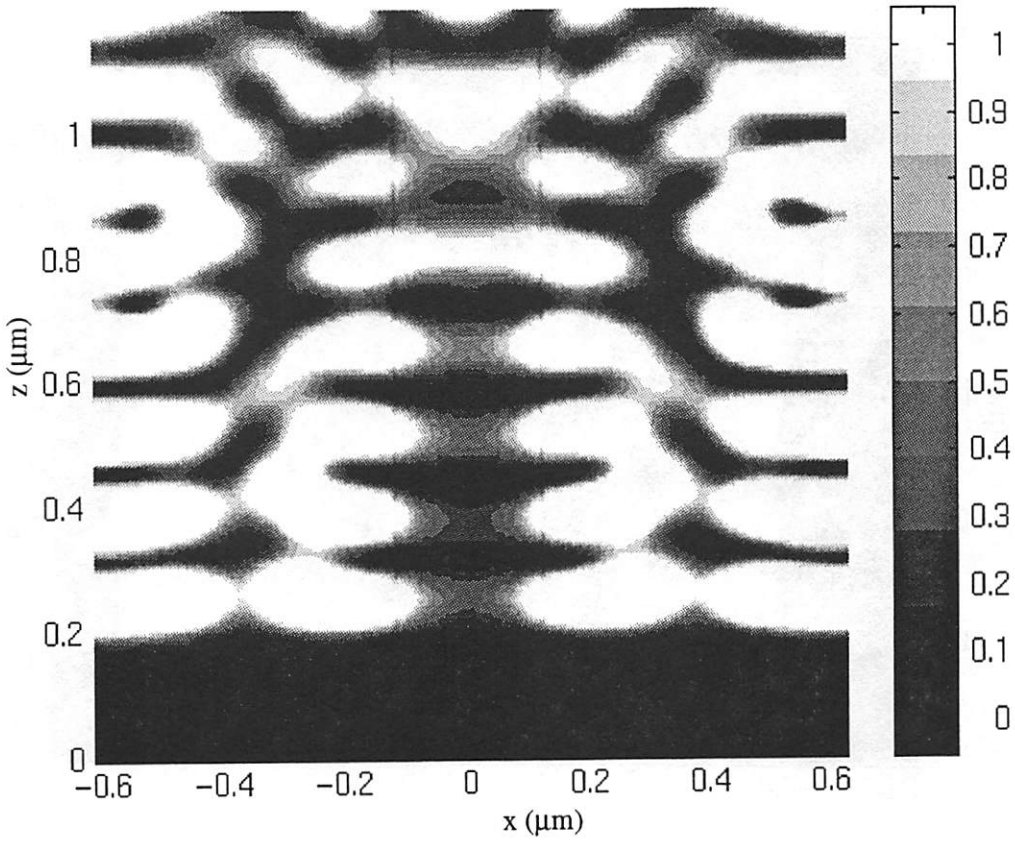


Figure 8-12. Intensity inside the contact hole when using the phase filter and a coherent point source excitation (σ=0).

Note that the intensity at the bottom of the hole is approximately equal to the intensity shown in Figure 8-6 when using a partially-coherent unfiltered excitation.

ing, the light can now propagate back out of the hole once it has interacted with defects at the bottom of the hole.

Although the filter conditions light such that it is guided down the contact hole, the intensity at the bottom of the hole, as in Figure 8-12, is less than the intensity at the bottom generated by the point source in Figure 8-8. This indicates that the filter is incapable of producing the same large field at the bottom of the hole when a point source is placed at the bottom. This occurs because the filter is forced to be phase only. Consequently, the phase and magnitude of the excitation generated by the filter seen in Figure 8-11 differs from the phase and magnitude of the radiated field in Figure 8-9. These differences cause the reduction in the intensity at the bottom of hole.

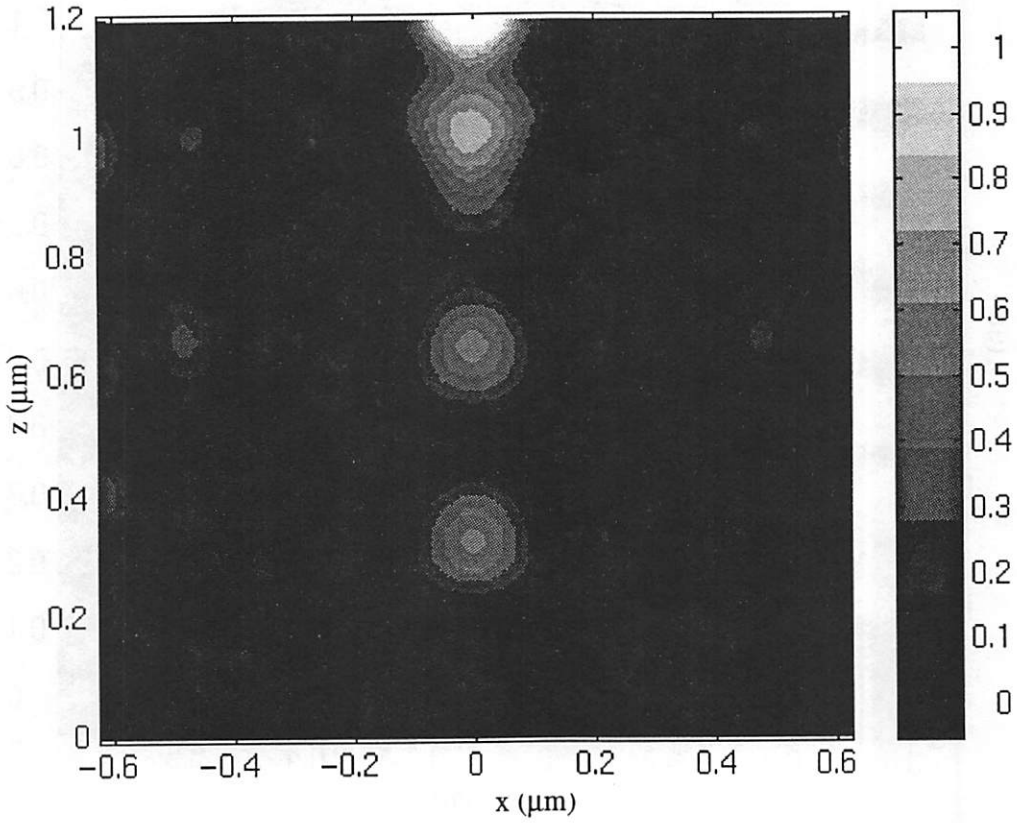


Figure 8-13. Effectiveness of the filter when an attenuating layer is placed on top of the oxide.

This intensity profile is generated with a point source (σ=0). The attenuated layer blocks off-axis illumination through the oxide; consequently, the filter launches a leaky mode into the hole.

8.4.4. Effect of the Partial Coherence on Filter Performance

In the previous section, the filter was shown to be effective at guiding coherent light down the hole. In this section the effect of the partial coherence on the performance of the filter is examined. Since the contact hole simulation domain is assumed to be periodic, the partial coherence of the source was included by using Abbe's formulation rather than the decomposition technique discussed in Chapter 6. In Abbe's formulation, the source is sampled such that the periodicity of the simulation domain is maintained. The periodicity is maintained by forcing the excitation field to be continuous over the simulation domain, which is accomplished by forcing the angular excitation angles, θ and ϕ, to be discrete. These angles are discrete when m and n in Equations 8-9 and 8-10 are of integer value

where P_x and P_y are the period of the simulation domain in the x and y directions, respectively. The field incident on the contact hole is then found by modulating this continuous

Equation 8-9.
$$m = \frac{P_x}{\lambda}\sin\theta\cos\phi \qquad m = 0, 1, 2, \ldots$$

Equation 8-10.
$$n = \frac{P_y}{\lambda}\sin\theta\sin\phi \qquad n = 0, 1, 2, \ldots$$

excitation field with the filter. This modulated field is projected onto the contact hole by collecting the light diffracting from the filter with the condenser optic. The scattering of this projected field from the contact hole topography is calculated by performing one TEMPEST simulation for each sampled point of the source. In this contact hole simulation, the source is sampled 29 times, which requires 29 TEMPEST simulations. The total intensity due to the partial coherent illumination is then calculated by summing the individual intensities produced by each of these 29 simulations.

This intensity is plotted in Figure 8-14, in which the partial coherence of the excitation has been included. In Figure 8-14, the partially coherent light has been modulated by the filter, and the scattering of this light from the contact hole was calculated through TEMPEST simulation. The intensity at the bottom of the hole in Figure 8-14 when σ=1 is approximately half the intensity at the bottom in Figure 8-12 when σ=0. This indicates that the partial coherence actually degrades the performance of the filter. This degradation occurs because the filter was designed to perform optimally for a single normally incident plane wave. However, since the partial coherence is modelled by modulating the filter with an obliquely incident plane wave in Abbe's formulation, the modulation with an obliquely incident plane wave changes the phase of the light propagating through the filter. Due to this phase change, the phase of the filter is no longer optimal when it is modulated with obliquely incident illumination; consequently, the performance of the filter is no longer optimal.

Although oblique illumination degrades the performance of the filter, it is possible to modify the illumination to maximize the amount of light reaching the bottom of the hole. The illumination source that maximizes the amount of light at the bottom of the hole is shown in Figure 8-15. In Figure 8-15, the source is sampled in five locations. In this

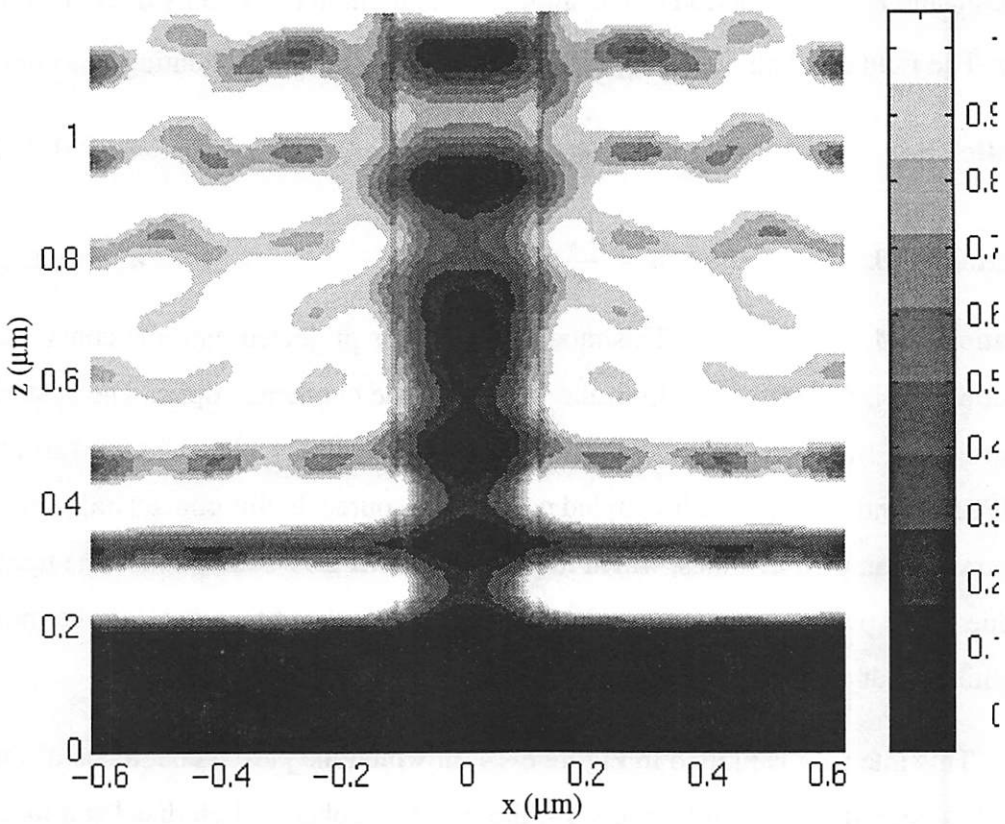
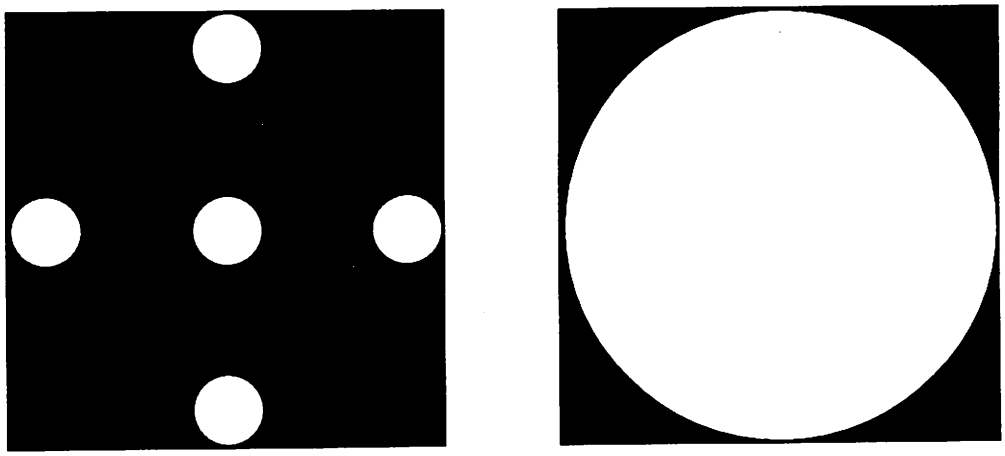


Figure 8-14. Intensity in the contact hole when using the filter and a conventional illumination source having σ of 1.0

source, each hole in the illumination pupil is 0.3 units in diameter relative to the size of the numerical aperture. These holes are located at (0, 0), (0, 0.85), (0, –0.85), (0.85, 0), and (–0.85, 0) in relative units of the pupil plane where the pupil is 1 unit in diameter. Using this source, the intensity in the hole was calculated with TEMPEST simulation and is plotted in Figure 8-16. Figure 8-16 shows that more light reaches the bottom of the hole when the source is sampled five times as compared to the intensity observed in Figure 8-12, when the source is sampled once, or when compared with the results depicted in Figure 8-14, where the source is sampled 29 times.

The effectiveness of using both the modified illumination and the filter situated in the object plane can be inferred by comparing results with this combination in Figure 8-16, to results of using conventional illumination without a filter, as those observed in Figure 8-6. Figure 8-16 shows that the intensity reaching the bottom when using both modified illu-



Modified illumination
Each circular source is 0.3 units in diameter and is located at (0,0), (0,0.85), (0,-0.85), (0.85,0), and (-0.85,0)

Convention illumination
The source is sampled 29 times

Figure 8-15. Pupils of the modified illumination source and the convention illumination source.

mination and the filter is approximately 3 times the intensity at the bottom in Figure 8-6 when neither modified illumination nor the filter is used. To better demonstrate this, the intensity as a function of z is plotted in Figure 8-17 when x=0.0μm. Figure 8-17 indicates that the intensity at the bottom of the hole (z=0.2μm) up to the middle of the hole (z=0.7μm) is greater when using the filter with modified illumination than when using conventional unfiltered illumination.

Using this modified illumination, the ability of the filter to image defects at the bottom of the hole was tested by comparing the aerial image of a hole without a defect to the aerial image of a hole with a defect. The defect in this test is a 0.05μm thick layer of glass at the bottom of the hole, i.e., the hole was underetched 0.05μm. The aerial image with and without the defect is plotted along the x axis and along the y axis in Figure 8-18. These aerial images are found by imaging the reflected light from the hole with the 0.9 NA condenser optic. The effect of the partial coherence on these aerial images is modelled by using Abbe's formulation, in which the total intensity is the sum of intensities produced by each of the five sources in the modified illumination. In these aerial images, the peak inten-

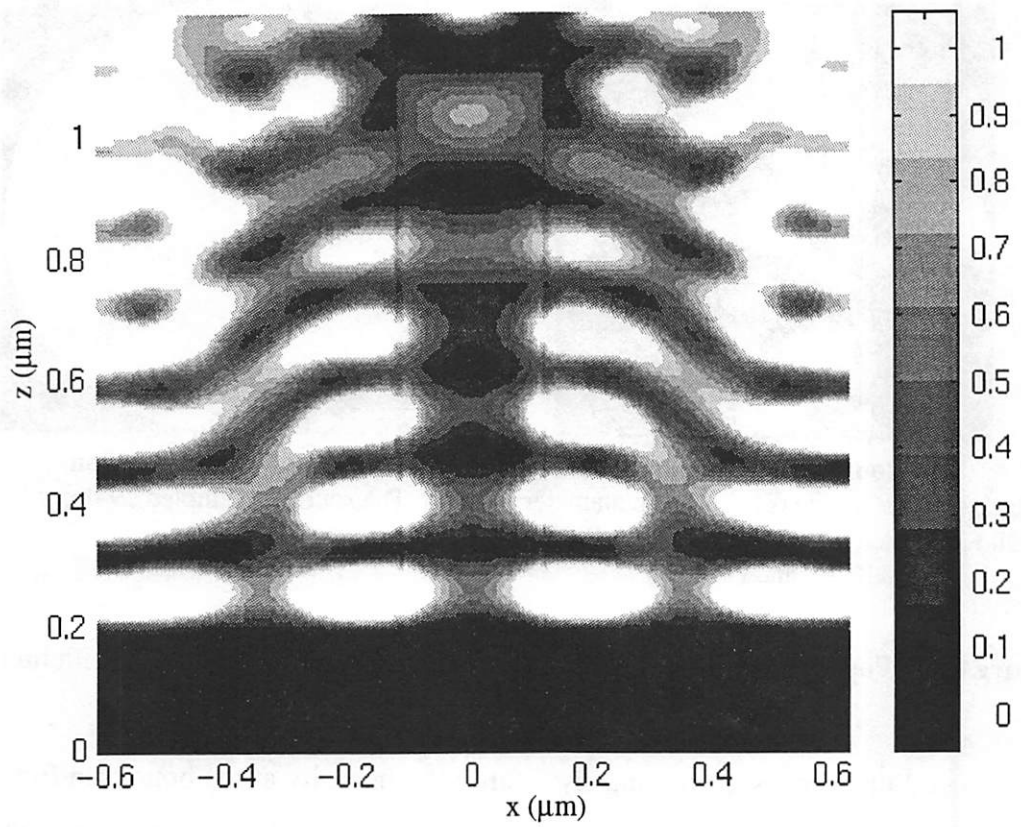


Figure 8-16. Intensity in the contact hole when using the filter and the modified illumination source of Figure 8-15.

sity at the center of the hole (x=0, y=0μm) is reduced when the hole is underetched, which indicates that the filter is capable of imaging defects at the bottom of the hole.

8.4.5. Effect of Process Variation on Filter Performance

It was shown in Section 8.4.4 that the filter is capable of imaging defects at the bottom of the hole. Due to variations in the process, the actual topography may deviate from the modeled topography of Figure 8-4. These variations may degrade the performance of the filter. In this section, the performance of the filter is tested when the oxide thickness is reduced from 1.0μm to 0.9μm and when the hole is overetched from 0.25μm to 0.3μm. Reflected aerial images due to these two variations and from the hole without any variations again are plotted along the x axis and y axis. In comparing the images in which the hole diameter changes, the intensity reflecting off the larger diameter hole is reduced in comparison to the intensity reflecting off the hole with 0.25μm diameter. This

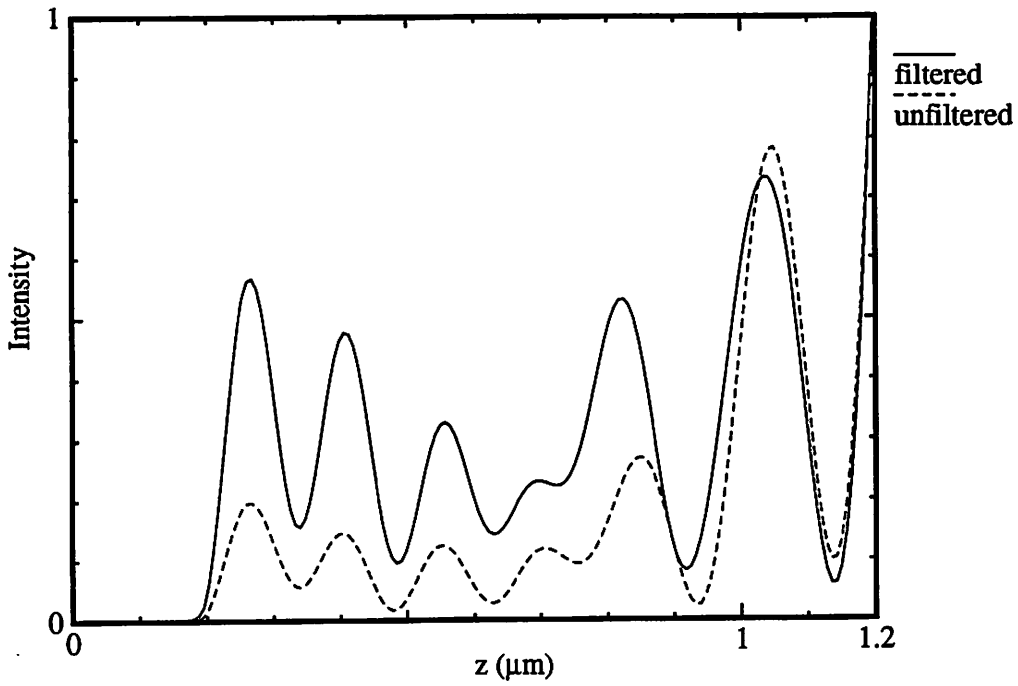


Figure 8-17. Intensity in the center of the contact hole for unfiltered and filtered illumination.

The intensity is plotted as a function of z when x=0.0μm. The intensity with the unfiltered illumination was given in Figure 8-6 when using conventional illumination with σ of 1.0. The intensity with the filtered illumination was given in Figure 8-16 when using the modified illumination of Figure 8-15.

is due to light being refracted out of the hole. A larger hole would propagate a leaky mode with a larger guide wavelength. Since the filter is incapable of creating this leaky mode, the light no longer propagates down the hole as efficiently as in the case of the 0.25μm diameter hole, and instead of propagating down the hole, the light is refracted out of the hole. In comparing images in which the oxide thickness changes, the aerial image with a thinner oxide differs significantly from the image produced from a 1.0μm thick oxide. This difference is most likely due to a focusing problem since the interference in the hole and in the oxide is occurring at a different plane. Both these variations, larger diameter and thinner oxide, cause large deviations in aerial image in comparison to the image produced by the 0.25μm diameter hole in 1.0μm thick oxide. These variations may make it difficult for the filter to determine when a defect is actually present or when a process variation is causing the difference.

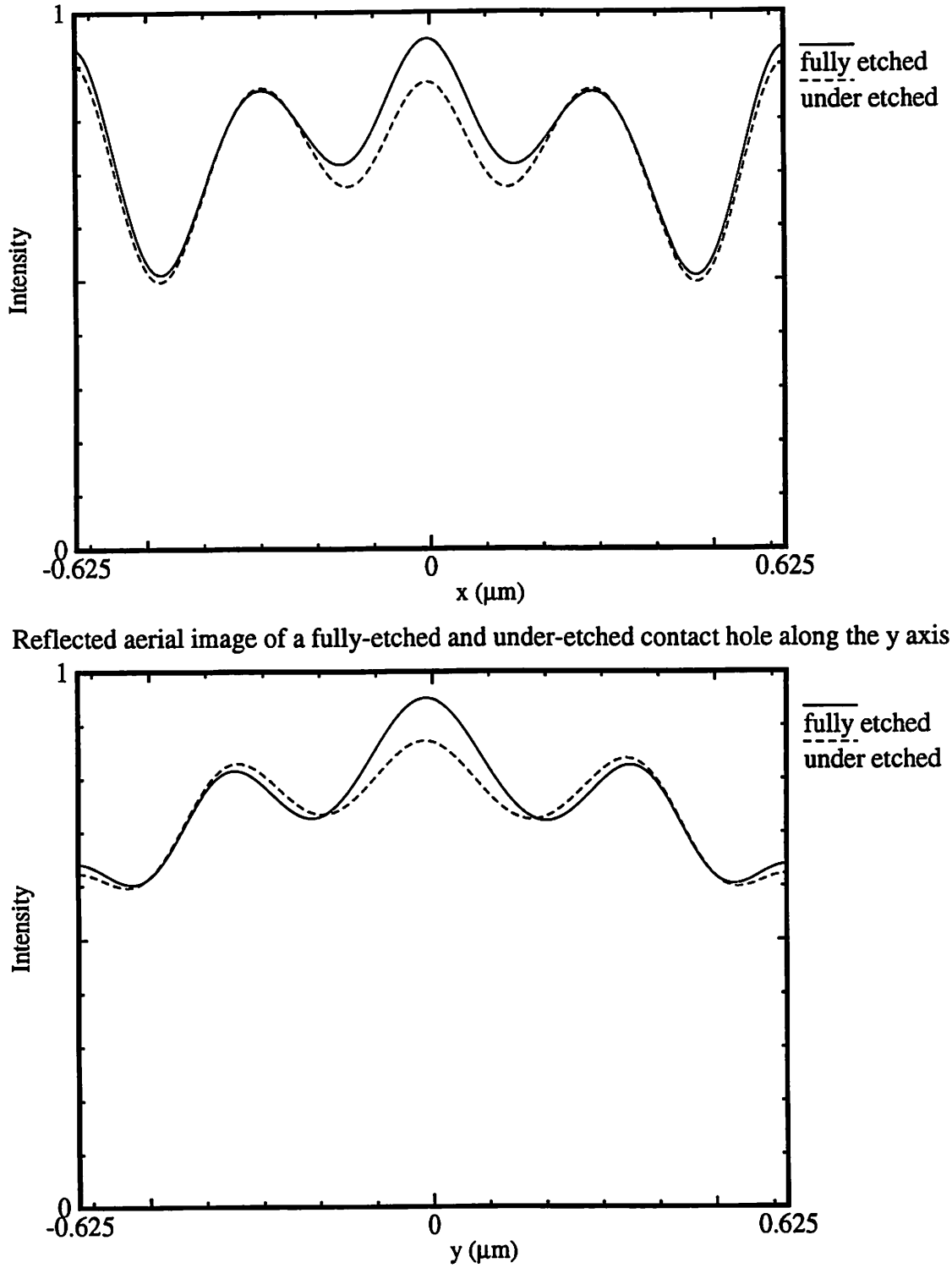


Figure 8-18. Reflected aerial images of fully-etched and under-etched contact holes.

In both the images the modified illumination source of Figure 8-15 is used.

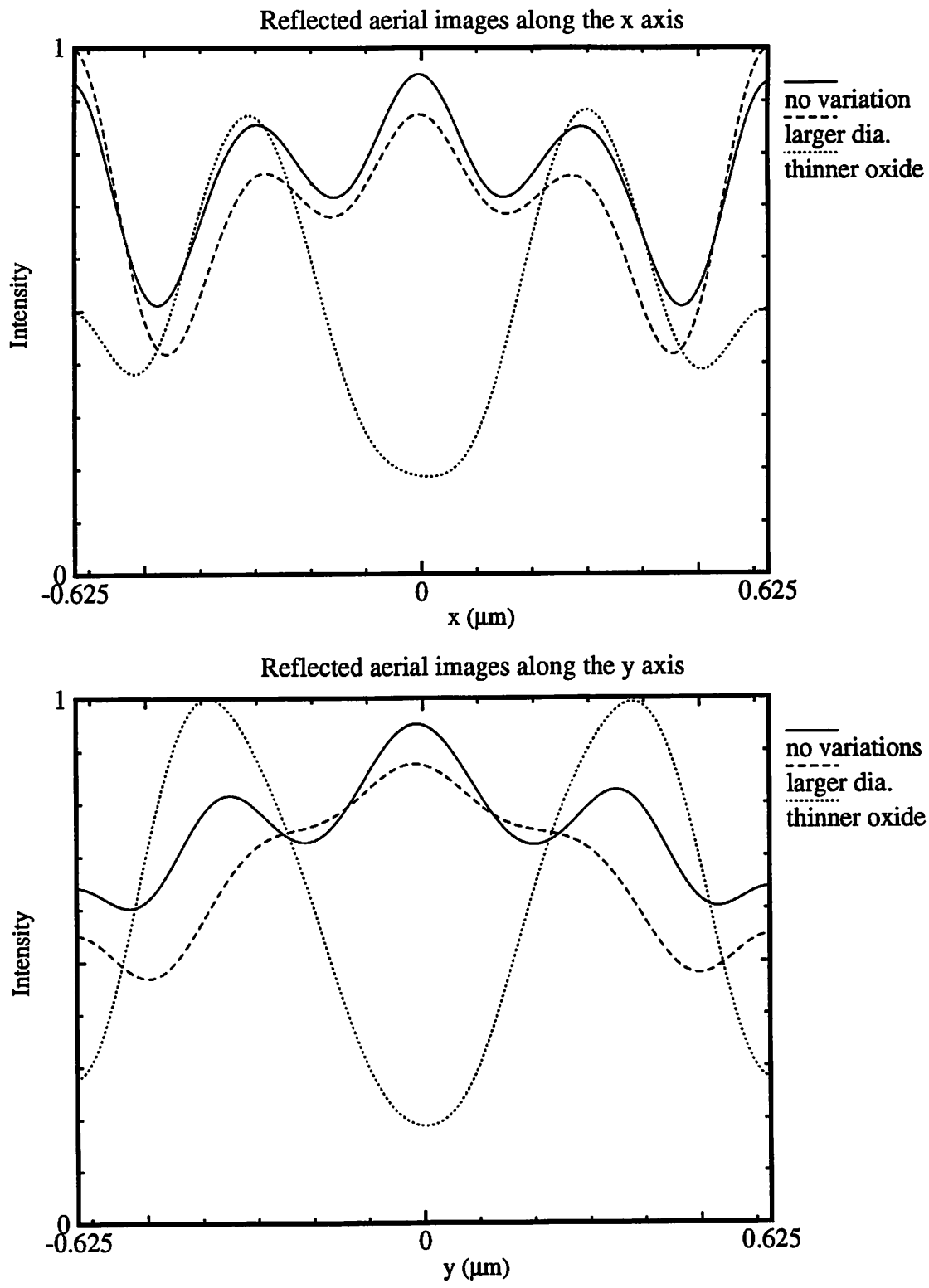


Figure 8-19. Reflected aerial images of the 0.25μm dia. hole in 1.0μm oxide as compared to the image of a 0.3μm dia. hole and to the image of a hole in 0.9μm oxide.

8.5. Conclusions

Thin-film layers induce spherical aberration that becomes significant for high NA inspection systems. Through a ray tracing model, these spherical aberrations were calculated for a glass-and-air interface. The aberrations are quantified by calculating the Strehl ratio, the point spread function, and the Zernike polynomials. These calculations indicate that 0.8μm of glass can be inspected through before the induced spherical aberrations degrade the image quality. However, since reflected inspection systems are used, the light must travel through the glass twice. Consequently, a defect below 0.4μm of glass may be difficult to distinguish because of a loss resolution in the reflected image of the defect.

The success of the filter in imaging the bottom of a contact hole shows that simulation tools are capable of optimizing contact hole inspection. By using this phase filter the amount of light reaching the bottom of the hole increases by three times. Furthermore, this filter launches a leaky mode that propagates down the hole. This propagation allows the light to interact with defects at the bottom of the hole and to propagate up the hole. When using this filter, the reflected aerial image differs by as much as 10% when the contact hole is under-etched by 0.05μm. Although this initial success shows that the filter is capable of imaging the bottom of the hole, the introduction of process variations such as a larger hole diameter or a thinner oxide causes the filter to fail. However, this initial study did not produce an optimal filter since the partial coherence and defocus was ignored in the design. In the future, the effect of the partial coherence will be added to the point source excitation. By fully considering the partial coherence, it may be possible to design a filter that is not susceptible to process variation.

9 Conclusions

As dimensions continue to shrink in semiconductor manufacture, understanding how light scatters from these smaller features becomes increasingly important. Simulation of these scattering effects in optical lithography can provide valuable process information and can predict problems due to these scattering mechanisms prior to their observation in manufacturing. Moreover, simulation is capable of identifying the scattering mechanisms which adversely effect lithographic performance. By isolating these mechanisms, simulation can be used to find approaches that correct problems due to scattering and to determine the effectiveness of novel technologies to overcome them. Furthermore, as the cost of performing experiments continues to increase and as the cost of computer time and memory continues to decrease, simulation is expected to grow as a cost effective means of creating and analyzing new processing techniques.

This thesis has extended the domain of electromagnetic simulation to include the propagation effects of partial coherent light in optical lithography and inspection. It presented the first formulation of a computationally efficient method for analyzing topography scattering with vector polarized, partially coherent spatial illumination. The implementation of this method into TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition) was also described and a systematic study of the effect of partial coherence in mask imaging and in wafer patterning by using TEMPEST-PCD was also given.

With regard to a computationally efficient method that includes polarization and partial coherence of the source, a vector formulation was given in Chapter 6. It is based on generalizing the scalar decomposition technique developed by Gamo which was presented in Chapter 3. In this method, the coherency matrix is expanded into a set of orthogonal eigenfunctions. Since the eigenfunctions are orthogonal, each of these eigenfunction exci-

tations are considered independently. This decomposition is optimal in an energy compaction sense, i.e., the least number of excitations are need to represent the coherency and the polarization of the system. This energy compaction reduces the simulation time by 7X for typical lithography problems and reduces the time by 10X for inspection problems. For example, in an inspection system with NA of 0.9 and σ of 1, the decomposition technique would require only 93 simulations for a 0.01% error in the center of the domain while Abbe's formulation would require over 1125 simulations to produce the same error. The vector decomposition is accurate in the center of the simulation and suffers some loss of accuracy at the edges of the simulation. The decomposition technique is more accurate in non-periodic simulation domains; Abbe's formulation is more accurate, however, in small periodic simulation domains when aliasing occurs.

This optimal decomposition technique was implemented by creating a new version of the program called TEMPEST. The new version is called TEMPEST-PCD (TEMPEST with Partial Coherence Decomposition). TEMPEST-PCD integrates SPLAT and TEMPEST into one software package. The software package decomposes the vector polarized partial coherent illumination and simulates the propagation of the partial coherent illumination. The software is capable of predicting how the propagation of partial coherent light influences scattering from the topography in optical lithography and inspection. In optical lithography or inspection, the aerial image due to the scattering of the partially coherent vector field from the topography on the wafer can be calculated by the software, and the image inside the photoresist can be evaluated as well.

The formulation and implementation of methods to include the partial coherence were made in support of a series of investigations of advanced innovations in optical lithography. Chapter 4 examined the effect of the partial coherence in mask imaging. By considering the mutual coherence between two points, $\mathbf{x}_1$ and $\mathbf{x}_2$, it was possible to develop a perturbational model which factors the problem into known contributions for rapid algebraic evaluation. The impact of a defect on the aerial image of a feature was found to be directly proportional to mutual coherence and inversely proportional to line edge slope of the image intensity. Depending on this line edge slope, some illumination sources are more

defect tolerant. For example, it was shown in Section 4.5 that quadrupole illumination is most defect tolerant for dense arrays because of its improved line edge slope.

A problematic issue in rapid evaluation of images in optical projection printing is the validity of the approximation that a mask's diffraction efficiencies are independent of the angle of illumination. This approximation is the foundation of Hopkins' method, which allows the order of integration to be changed such that the transmission cross-coefficients are a property of the optical system and are independent of the mask. To understand the effect of vector-polarized partially coherent light, simulations were run for a 2D phase shifting line pattern (see Section 7.2) and for a 3D phase-shifting contact hole (see Section 6.7). These simulations calculated the aerial image through four methods: scalar SPLAT simulation, TEMPEST-HN simulation, TEMPEST-PCD, and TEMPEST with Abbe's formulation. For both of these phase shifting features, the aerial images with TEMPEST-PCD were in very close agreement with the images calculated with TEMPEST-Abbe. In all the cases simulated, TEMPEST-PCD required fewer simulations than TEMPEST-Abbe. The aerial image calculated with TEMPEST-HN, which is an approximate method to include the partial coherence, begins to deviate from the aerial images with TEMPEST-PCD when the magnitude of the diffracted orders is no longer independent of the angle of incidence. These diffracted orders were found to be independent of angle when the height of the feature is greater than 20% of the opening size and when σ multiplied by NA (σNA) is less than 0.5. In comparison to these TEMPEST simulation techniques, the scalar simulations with SPLAT, however, predict a larger intensity in the clear areas of the mask that is incorrect. This increase is due to the fact that the scalar SPLAT simulation neglects the induced currents on the mask. These currents short the vector fields near mask edges and decrease the effective width which reduces the amount of light transmitted. Since SPLAT simulation is much faster than TEMPEST simulation, SPLAT should be used as a first-order prediction of the aerial image and TEMPEST should be used to determine second-order effects due to these induced currents. Such studies are needed for sub-resolution optical proximity correction features.

By using a pre-patterned edge in resist to produce a 180^{o} phase shift, a line with sub-wavelength resolution can be patterned. The linewidth with the two layer process is approximately on the order of a wavelength while a phase-shift edge on a mask can produce a linewidth that is two-thirds the wavelength. The simulations also indicate that a σ less than 0.3 is needed to produce an intensity minimum that is less than 0.1. This intensity minimum is necessary in order to produce a well-defined photoresist line after development. The phase-shift edge of the top-layer resist causes light to refract into the bottom of the resist in a region directly below the top layer. This refraction is due to the critical angle effect of light incident at 90^{o} with respect to the phase edge. This introduces an unwanted asymmetry in the intensity profile and in the developed resist line. The refraction at the edge also causes a 25nm shift in alignment of the line with the top-layer resist edge. Finally, $\pm 10\%$ thickness variation in the top-layer resist causes as much as $\pm 5\%$ linewidth variation and 25nm shift in alignment. The slope of the phase edge was shown to produce little change in the linewidth and alignment.

One important reason underlying this research is to enhance our understanding of the influence of the partial coherence on wafer printing. TEMPEST-PCD simulations more accurately model the patterning of a gate over an active area trench. TEMPEST-PCD predicts that the photoactive compound (PAC) concentration in the unexposed gate area will be higher at low σ prior to the post-exposure bake (PEB). This would seem to indicate that higher σ have less reflecting notching problems than lower σ in the latent image. However, since a post-exposure bake is always performed to reduce standing waves, the effect of the PEB must be taken into account prior to reaching a conclusion on the role of σ on reflective notching. When the diffusion is Fickian and when the dose, D, multiplied by the photo speed, C, is much less than one, the generation of PAC is linearly proportional to intensity. This implies that the PEB diffusion can be taken into account by convolving the intensity with a Gaussian diffusion function rather than by convolving the PAC with this Gaussian function. This PEB diffusion in effect low pass filters the initial PAC concentration. When the diffusion length is near the interference distance in the resist, $\frac{\lambda}{4n}$, notching in the pattern is eliminated and the diffusion in effect removes the coherence between the interfering

beams. This diffusion process implies that simulation of the partial coherence may not be important in practice, and coherent illumination can be assumed when the diffusion length is near $\frac{\lambda}{4n}$.

Studies of wafer inspection lead to several important conclusions about light interaction in inspection methods. Thin-film layers induce spherical aberration which becomes significant for high NA inspection systems, as shown in Section 8.2. Through a ray tracing model, these spherical aberrations were calculated for a glass-and-air interface. These calculations indicate that a defect below 0.4μm of glass may be difficult to distinguish because of a loss in resolution due to these induced spherical aberrations. In Section 8.4, a method of designing a filter based on the reciprocity theorem and implemented in TEMPEST was presented to find the far fields radiated from a point source placed at the bottom of a contact hole. From these far fields, a filter was designed and forced to be phase only. This phase filter increases threefold the amount of light reaching the bottom of the hole. Evidence that this filter launches a leaky mode that propagates down the hole was also found. This propagation allows the light to interact with defects at the bottom of the hole and to propagate up the hole. The sensitivity of this filter was also assessed and found to have moderate dependence on hole size and oxide thickness.

With this TEMPEST extension to include the partial coherence, TEMPEST-PCD is effective in modeling two-dimensional and three-dimensional electromagnetic scattering problems in optical lithography and inspection. Hopefully, this extension will assist in the development of future lithographic and inspection technologies.

The main goal of this thesis was to extend topography scattering to include partial coherence effects. The optimal expansion method of Gamo was found to be the best approach. This problem has plagued the IC industry for some time. A problem that required 256 simulations in the past can now be reduced to 30 simulations with the decomposition technique. Availability of a tool to simulate partial coherence yielded much substantive information about the full effect of the partial coherence. While the partial coherence in topography scattering is not the major concern it was once thought to be, it is hoped that

the methods and code developed in this thesis will continue to provide a full check of the partial coherence when needed.

Bibliography

[1] E. Abbe, "Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," Archiv f. Mikroskopische Anat., vol. 9, pp. 413-468 (1873).

[2] S. André and A. Weill, "Variable Numerical Aperture and Partial Coherence studies: Process Window and Proximity Effects.

[3] C. Balanis, *Advanced Engineering Electromagnetics* (John Wiley & Sons, New York, 1989), pp. 314-323.

[4] ibid., pp. 323-325.

[5] P. H. Berning, "Theory and Calculations of Optical Thin Films," *Physics of Thin Films*, ed. George Hass (Academic Press, New York, 1963), pp. 931-934.

[6] E. Barouch, B. Bradie, G. Karniadakis, and S. Orszag, "Comprehensive 3D simulator with non-planar substrates," Proc. of SPIE, vol. 1264, pp. 334-342 (1990).

[7] M. Beran and G. Parrent, Theory of Partial Coherence (Prentice Hall, Englewood Cliffs, N. J., 1964).

[8] N. Bobroff and A. Rosenbluth, "Alignment errors from resist coating topography," J. Vac. Sci. Technol. B, vol. 6, no. 1, pp. 403-408 (1988).

[9] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, New York, 1980), p. 438.

[10] ibid., pp. 459-468.

[11] ibid., pp. 513-516.

[12] ibid., pp. 519-526.

[13] ibid., pp. 535-537.

[14] Y. Borodovsky, "Impact of Local Partial Coherence Variations on Exposure Tool Performance," Proc. of SPIE, vol. 2440, pp. 750-770 (1995).

[15] R. von Bünau et al., "Depth of focus enhancement in optical lithography," J. Vac. Sci. Technol. B, vol. 10, pp. 3047-3054 (1992).

[16] R. von Bünau et al., "Optimization of pupil filters for increased depth of focus," Jpn. J. Appl. Phys., vol. 32, pp. 5850-5855 (1993).

[17] R. von Bünau, "Depth of Focus Enhancement in Optical Lithography," Ph. D. Dissertation, Stanford University, 1995.

[18] C. B. Burckhardt, "Diffraction of a Plane Wave at a Sinusoidally Stratified Dielectric Grating," J. Opt. Soc. Am., vol. 56, no. 11, pp. 1502-1509 (1966).

[19] D. Canals-Frau and M. Rousseau, Optica Acta, vol. 5, p. 15 (1958).

[20] P. Canestrari, S. Carrera, G. A. Degiorgis, and V. Visentini, "Impact of Reticle Defects on Submicron 5x Lithography," Proc. of SPIE, vol. 1261, pp. 225-237 (1990).

[21] P. Canestrari, G. Degiorgis, P. De Natale, L. Gazzaruso, and G. Rivera, "Optimization of partial coherence for half micron i-line lithography," Proc. of SPIE, vol. 1463, pp. 446-455 (1991).

[22] J. R. Carson, "Reciprocal theorems in radio communications," Proc. IRE, vol. 17, pp. 952-956 (1929).

[23] W. Charman, "Some experimental measurements of diffraction images in low-resolution microscopy," J. Opt. Soc. Am. vol. 53, p. 410 (1963).

[24] N. Cobb and A. Zakhor, "Fast, Low-Complexity Mask Design," Proc. SPIE, vol. 2440, p. 313 (1995).

[25] N. Cobb, "Sum of Coherent Systems Decomposition by SVD," not published (1995).

[26] D. Cole *et al.*, "Extending scalar aerial image calculations to higher numerical apertures," Journal of Vacuum Science and Technology B, vol. 10, No. 6, pp. 3037-3041 (1992).

[27] P. Considine, "Effects of Coherence on Imaging Systems," J. Opt. Soc. Am., vol. 56, p. 1001 (1966).

[28] M. De and S. C. Som, "Diffraction images of circular openings with partially coherent illumination," Opt. Acta, vol. 9, p. 17 (1962).

[29] M. De and S. C. Som, "Diffraction Images of Circular Phase Objects in Partially Coherent Light," J. Opt. Soc. Am., vol. 53, no. 7, p. 779 (1963).

[30] J. W. Demmel, J. R. Gilbert, and X. S. Li, "SuperLU Users' Guide," University of California at Berkeley (1997).

[31] H. Dijkstra and C. Juffermans, "Optimization of Anti-Reflection Layers for Deep UV Lithography," Proc. of SPIE, vol 1927, pp. 275-286 (1993).

[32] F. H. Dill, "Optical Lithography," IEEE Trans. Electron Dev., vol. ED-22, no. 7, pp. 440-444 (1975).

[33] D. G. Flagello and R. Rogoff, "The influence of photoresist on the optical performance of high NA steppers," Proc. of SPIE, vol. 2440, pp. 340-348 (1995).

[34] H. Fukuda *et al.*, "Spatial filtering for depth of focus and resolution enhancement in optical lithography," J. Vac. Sci. Technol. B, vol. 9, pp. 3113-3116 (1991).

[35] G. M. Gallatin, J. C. Wedster, E. C. Kintner, and F. Wu, "Modeling the images of alignment marks under photoresist," Proc. of SPIE, vol. 772, pp. 193-201 (1987).

[36] J. Gamelin, "Simulation of Topography Scattering for Optical Lithography with the Connection Machine," M.S. Thesis, Memorandum No. UCB/ERL M89/71, University of California, Berkeley, May 1989.

[37] J. Gamelin, R. Guerrieri, and A. R. Neureuther, "Exploration of Scattering from Topography with Massively Parallel Computers," J. Vac. Sci. Technol. B, vol. 7, no. 6, pp. 1984-1990 (1989).

[38] H. Gamo, "Matrix Treatment of Partial Coherence," *Progress in Optics*, (North Holland, New York, 1964), vol. 3, p. 187.

[39] J. Garafalo, J. DeMarco, J. Bailey, J. Xiao, and S. Vaidya, "Reduction of ASIC Gate-level line-end shortening by Mask Compensation," Proc. of SPIE, vol. 2440, pp. 171-183 (1995).

[40] J. Goodman, "Statistical Properties of Laser Speckle Patterns," Laser Speckle and Related Phenomena, edited by J. Dainty, pp. 42- 46, (1984).

[41] J. Goodman, *Statistical Optics*, (Wiley, New York, 1985), pp. 109-111.

[42] J. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1995), p. 32.

[43] ibid., pp. 37-40.

[44] ibid., pp. 102-103.

[45] ibid., pp. 134-136.

[46] R. Guerrieri, K. Tadros, J. Gamelin, and A. Neureuther, "Massively Parallel Algorithms for Scattering in Optical Lithography," IEEE Trans. CAD, vol. 10, no. 9, pp. 1091-1100 (1991).

[47] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*, (McGraw-Hill, New York, 1961).

[48] W. H. Hayt, Jr., and J. E. Kimmerly, *Engineering Circuit Analysis*, Third Edition, (McGraw-Hill, New York, 1978).

[49] E. Hecht, *Optics*, (Addison-Wesley, Menlo Park, 1987). pp. 523-528.

[50] W. Henke, D. Mewes, M. Weiss, G. Czech, and R. Schiessl-Hoyler, "Simulation of Defects in 3-Dimensional Resist Profiles in Optical Lithography," Microelectronic Engineering, vol. 13, pp. 497-501 (1991).

[51] H. H. Hopkins, "The concept of partial coherence in optics," Proc. Roy. Soc. A, vol. 208, p. 263 (1951).

[52] H. H. Hopkins, "On the diffraction theory of optical images," Proc. Roy. Soc. A, vol. 217, p. 408 (1953).

[53] H. H. Hopkins, "The frequency response of a defocused optical system," Proc. Roy. Soc. A, vol. 231, p. 91 (1955).

[54] H. H. Hopkins, "Applications of coherence theory in microscopy and interferometry," Journal of the Optical Society of America, vol. 47, No. 6, pp. 508-526 (1957).

[55] J. C. Housley and D. J. Williams, "Dyes in Photoresist: Today's View," Semiconductor International, pp. 142-144, April 1988.

[56] B. Huynh, K. K. Toh, W. E. Haller and A. R. Neureuther, "Optical Printability of Defects in Two-dimensional Patterns," J. Vac. Sci. Technol. B, vol. 6, no. 6, p. 2207 (1988).

[57] D. F. Ilten and K. V. Patel, "Standing Wave Effects in Photoresist Exposure," Image Technology, pp. 9-14 (1971).

[58] G. Indebetouw and H. Bai, "Imaging with Fresnel zone pupil masks," Appl. Opt., vol. 23, pp. 4299-4302 (1984).

[59] F. G. Kasper, "Diffraction by thick, periodically stratified gratings with complex dielectric constant," J. Opt. Soc. Am., vol. 63, no. 1, pp. 37-45 (1973).

[60] M. C. King, "The emerging technology for 1:1 optical projection lithography," *COMPCON spring '80. VLSI: New Architectural Horizons* (IEEE Press, New York). pp. 294-297 (1980).

[61] E. C. Kintner, "Method for the calculation of partially coherent imagery," Applied Optics, vol. 17, no. 17, pp. 2747-2753 (1978).

[62] E. Korka, "Standing Waves in Photoresists," Applied Optics, vol. 9, no. 4, pp. 969-970 (1970).

[63] M. E. Law and R. W. Dutton, "Verification of Analytic Point Defect Models Using SUPREM-IV," IEEE Trans. CAD, vol. 7, pp. 181-190 (1988).

[64] D. Lee, D. Newmark, K. Toh, P. Flanner, and A. Neureuther, "SPLAT v5.0 Users' Guide," Memorandum No. UCB/ERL M95/13 University of California, Berkeley (1995).

[65] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving Resolution in Photolithography with a Phase-Shifting Mask," IEEE Trans. Elect. Dev., vol. ED-31, pp. 753-763 (1984).

[66] J. Lim, *Two-Dimensional Signal and Image Processing*, (PTR Prentice Hall, New Jersey, 1990), pp. 154-157.

[67] ibid., pp. 642-647.

[68] D. Lucas, A. Strojwas, and H. Tanabe, "Efficient 3D Phase Shifting Mask Lithography Simulation," Proc. of SPIE, vol. 2440, pp. 422-434 (1995).

[69] C. A. Mack, "PROLITH: A Comprehensive Optical Lithography Model," Proc. of SPIE, vol. 538, pp. 207-220 (1985).

[70] C. A. Mack, "An analytical expression for the standing wave intensity in photoresist," Applied Optics, vol. 25, no. 12, pp. 1958-1961 (1985).

[71] C. A. Mack, "An Algorithm for Optimizing Stepper Performance Through Image Manipulation," Proc. of SPIE, vol. 1264, pp. 71-82 (1990).

[72] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, (Cambridge University Press, Cambridge, 1995), pp. 340-373.

[73] A. S. Marathay, *Elements of Optical Coherence Theory* (John Wiley & Sons, New York, 1982), pp. 105-112.

[74] V. Mastromarco, A. R. Neureuther, and K. K. H. Toh, "Printability of Defects in Optical Lithography: Polarity and Critical Location Effects," J. Vac. Sci. Technol. B, vol. 6, no. 1, p. 224 (1988).

[75] T. Matsuzawa, A. Moniwa, N. Hasegawa, and H. Sunami, "Two-Dimensional Simulation of Photolithography on Reflective Stepped Substrate," IEEE Trans. CAD, vol. CAD-6, no. 3, pp. 446-451 (1987).

[76] C. W. McCutchen, "Generalized Aperture and the Three-Dimensional Diffraction Image," J. Opt. Soc. Am., vol. 54, no. 2, p. 240 (1964).

[77] S. Middlehoek, "Projection Masking, Thin Photoresist Layers and Interference Effects," IBM Jour. Res. Dev., vol. 14, pp. 117-124 (1970).

[78] K. Mikami, H. Mohri, H. Miyashita, N. Hayashi, H. Sano, "Development and Evaluation of Chromium-Based Attenuated Phase Shift Masks for DUV Exposure," Proc. of SPIE, vol. 2521, pp. 333-342 (1995).

[79] M. Mino and Y. Okano, "Improvement of the OTF of a defocused optical system through the use of shaded apertures," Appl. Opt., vol. 10, pp. 2219-2225 (1971).

[80] L.W. Nagel, "SPICE2 - A Computer Program to Simulate Semiconductor Circuits," *M. S. Thesis*, Memorandum No. ERL-M520, University of California, Berkeley, May 1975.

[81] A. Neureuther and K. Zaki, "Numerical Methods for the Analysis of Scattering from Non-Planar Periodic Structures," Intn'l URSI Symposium on Electromagnetic Waves, Stresa, Italy, pp. 282-285 (1969).

[82] A. Neureuther, P. Flanner, III, and S. Shen, "Coherence of defect interactions with features in optical imaging," J. Vac. Sci. Technol. B, vol. 5 (1), pp. 308-312 (1987).

[83] M. Neviere, P. Vincent, R. Petit, and M. Cadilhac, "Systematic Study of Resonances of Holographic Thin Film Couplers," Optics Communications, vol. 9, no. 1, pp. 48-53 (1973).

[84] C. Nolscher, L. Mader, and M. Schneegans, "High contrast single layer resists and antireflection layers - an alternative to multilayer resist techniques," Proc. of SPIE, vol. 1086, pp. 242-250 (1989).

[85] D. Nyyssonen and C. P. Kirk, "Optical microscope imaging of lines patterned in thick layers with variable edge geometry: theory," J. Opt. Soc. Am. A, vol. 5, no. 8, pp. 1270-1280 (1988).

[86] T. Ogawa, M. Kimura, T. Gotyo, Y. Tomo, and T. Tsumori, "Practical resolution enhancement effect by new complete anti-reflective layer in KrF excimer laser lithography," Proc. of SPIE, vol. 1927, pp. 263-274 (1993).

[87] H. Ohtsuka, K. Kuwahara, and T. Onodera, "Diagonal Phase Errors Affecting on Optical Intensity in Phase Defect Repair Elements," Jap. J. Appl. Physics, Part 1, vol. 32, no. 12B, pp. 5892-5899 (1993).

[88] J. Ojeda-Castañeda *et. al.*, "Line-spread function relatively insensitive to defocus," Opt. Lett., vol. 8, pp. 458-460 (1983).

[89] J. Ojeda-Castañeda *et. al.*, "Spatial filter for increasing depth of focus," Opt. Lett., vol. 10, pp. 520-522 (1985).

[90] J. Ojeda-Castañeda *et. al.*, "Annular apodizers for low sensitivity to defocus and to spherical aberration," Opt. Lett., vol. 11, pp. 487-489 (1986).

[91] J. Ojeda-Castañeda *et. al.*, "Bessel annular apodizers: imaging characteristics," Appl. Opt., vol. 26, pp. 2770-2772 (1987).

[92] J. Ojeda-Castañeda and A. Diaz, "High focal depth quasibifocus," Appl. Opt. vol. 27, pp. 4163-4165 (1988).

[93] J. Ojeda-Castañeda *et. al.*, "Apodization of annular apertures: Strehl ratio," Appl. Opt., vol. 27, pp. 5140-5145 (1988).

[94] W. G. Oldham, S. N. Nandgaonkar, A. R. Neureuther, and M. M. O'Toole, "A General Simulator for VLSI Lithography and Etching Process: Part I - Applications to Projection Lithography," IEEE Trans. Electron Devices, vol. ED-26, no. 4, pp. 717-722 (1979).

[95] M. M. O'Toole, "Simulation of Optically Formed Image Profiles in Positive Resist," Ph. D. Dissertation, U. Calif., Berkeley (1979).

[96] W. N. Partlo, S. G. Olson, C. Sparks, and J. E. Connors, "Optimizing NA and Sigma for Sub-Half Micrometer Lithography," Proc. of SPIE, vol. 1927, pp. 320-332 (1993).

[97] Y. Pati and T. Kailath, "Phase-shifting masks for microlithography: Automated design and mask requirements," J. Opt. Soc. Am. A, vol. 2438, p. 2438 (1994).

[98] J. Petersen, D. Herr, M. Lutz, and J. Bontrager, "An experimental method for the minimization of linewidth variation in photoresist patterns over polysilicon topography," Proc. of SPIE, vol. 921, pp. 326-336 (1988).

[99] R. Petit and M. Cadihac, "Sur la diffraction d'une onde plane parun réseau infiniment conducteur," C. R. Acad. Sci. B, vol. 262, no. 7, pp. 468-471 (1966).

[100] M. R. Pinto, C. S. Rafferty, and R. W. Dutton, "PISCES-II - Poisson and Continuity Equation Solver," Stanford Electronics Laboratory Technical Report, Stanford University, Sept. 1984.

[101] T. Pistor, "Generalizing the TEMPEST FDTD Electromagnetic Simulation Program," M.S. Thesis, University of California at Berkeley (1997).

[102] C. Progler and D. Byrne, "Merit functions for lithographic lens design," J. Vac. Sci. Technol. B, vol. 14, no. 6, pp. 3714-3718 (1996).

[103] C. Progler, H. Du, and G. Wells, "Potential causes of across field CD variation," Proc. of SPIE, vol. 3051 (1997).

[104] Q.-D. Qian and F. A. Leon, "Fast algorithms for 3D high NA lithography simulation," Proc. of SPIE, vol. 2440, pp. 372-380 (1995).

[105] Lord Rayleigh, "On the Dynamical Theory of Gratings," Proc. R. Soc. London A, vol. 79, pp. 399-416 (1907).

[106] B. Saleh and M. Rabbani, "Simulation of partially coherent imagery in the space and frequency domains and by modal expansion," Applied Optics, vol. 21, pp. 2770-2777 (1982).

[107] S. Selberherr, A. Schütz, and H. W. Pötzl, "MINIMOS - a Two-Dimensional MOS Transistor Analyzer," IEEE Trans. Electron Devices, vol. ED-27, pp. 1540-1550 (1980).

[108] S. Y. Shaw, S. Palmer, and S. J. Schuda, "Printability Study of Opaque and Transparent Defects Using Standard and Modified Illumination," Proc. of SPIE, vol. 2440, pp. 878-890 (1995).

[109] C. J. R. Sheppard and Z. S. Hegedus, "Axial behavior of pupil-plane filters," J. Opt. Soc. Am. A, vol. 5, pp. 643-647 (1988).

[110] B. J. Sheu, D. L. Scharfetter, P. K. Ko, and M. C. Jeng, "BSIM: Berkeley short-channel IGFET model for MOS transistors," IEEE J. Solid-State Circuits, vol. SSC-22, pp. 558-565 (1987).

[111] R. Socha, A. Neureuther, and R. Singh, "Printability of Phase Shift Defects Using a Perturbational Model," Proc. SPIE, vol. 2087, pp. 277-287 (1993).

[112] R. Socha, A. Wong, M. Cagan, Z. Krivokapic, and A. Neureuther, "Effects of Wafer Topography on the Formation of Polysilicon Gates," Proc. of SPIE, vol. 2440, p. 361 (1995).

[113] R. Socha and A. Neureuther, "The Role of Illumination and Thin-Film Layers on the Printability of Defects," Proc. of SPIE, vol. 2440, p. 532 (1995).

[114] R. Socha, A. Neureuther, and R. Singh, "Models for Characterizing Phase-Shift Defects in Optical Projection Printing," IEEE Trans. Semi. Man., vol. 8, no. 2, pp. 139-149 (1995).

[115] R. Socha and A. Neureuther, "Propagation effects of partial coherence in optical lithography," J. Vac. Sci. Technol. B, vol. 14, p. 3724 (1996).

[116] R. Socha, C. Progler, and A. Neureuther, "The Effect of Partial Coherence on Reflective Notching," Proc. of SPIE, vol. 3051 (1997).

[117] R. Socha and A. Neureuther, "Pupil Filter Design Methodology for Defect Detection in Wafer Inspection," J. Vac. Sci. Technol. B., vol. 15, (1997).

[118] S. C. Som, "Diffraction Images of Annular and Disk-like Objects under Partially Coherent Illumination," J. Opt. Soc. Am., vol. 57, no. 12, p. 1499 (1967).

[119] A. Sommerfeld, *Partial Differential Equations in Physics*, (Academic Press, New York, 1949), pp. 236-265.

[120] A. Sommerfeld, *Optics*, (Academic Press, New York, 1954), p. 197.

[121] W. Steel, "Effects of Small Aberrations on the Images of Partially Coherent Objects," J. Opt. Soc. Am., vol. 47, p. 405 (1957).

[122] S. Subramanian, "Rapid calculation of defocused partially coherent images," Applied Optics, Vol. 20, No. 10, pp. 1854-1857 (1854).

[123] K. Tadros, A. Neureuther, J. Gamelin, and R. Guerrieri, "Investigation of Reflective Notching with Massively Parallel Simulation," Proc. of SPIE, vol. 1264, pp. 322-332 (1990).

[124] Technology Modeling Associates, DEPECT-2, Technology Modeling Associates, 1990.

[125] E. Tejnil, Private Communication.

[126] K. K. Toh and A. Neureuther, "Identifying and monitoring effects of lens aberrations in projection printing", Proc. of SPIE, vol. 772, pp. 202-209 (1987).

[127] Y. Tomo, T. Kasuga, M. Saito, M. Someya and T. Tsumori, "0.35mm rule device pattern fabrication using high absorption type novolac photoresist in single layer deep UV lithography: Surface image transfer for contact hole fabrication," J. Vac. Sci. Technol., B, p. 2576 (1992).

[128] B. Thompson, "Image formation with partially coherent light," Progress in Optics, vol. 7, pp. 169-230 (1969).

[129] A. Tol, G. Maxwell, H. Urbach, and R. Visser, "Simulation of scattering effects in photolithography," Proc. of SPIE, vol. 1264, pp. 294-308 (1990).

[130] J. Tsujiuchi, "Correction of optical images by compensation of aberrations and by spatial frequency filtering," Progress in Optics, vol. 3, pp. 131-152 (1963).

[131] T. Uchiyama, S. Shioiri, T. Hashimoto, and K. Kasama, "Standing Wave Effect of Various Illumination Methods in 0.25μm KrF Excimer Laser Lithography," Jpn. J. Appl. Phys., vol. 34, pp. 6560-6564 (1995).

[132] H. Urbach and D. Bernard, "Modelling Latent Image Formation in Photolithography using the Helmholtz Equation," Proc. of SPIE, vol. 1264, pp. 278-293 (1990).

[133] C. Varamit and G. Indebetouw, "Imaging properties of defocused partitioned pupils," J. Opt. Soc. Am., vol. 6, pp. 799-802 (1985).

[134] S. Venit and W. Bishop, Elementary Linear Algebra, (PWS-Kent Publishing Company, Boston, 1989), p. 311.

[135] M. Vetterli, "Multidimensional subband coding: some theory and algorithms," Signal Processing, vol. 6, p. 97 (1984).

[136] H. Watanabe, Y. Todokoro, and M. Inoue, Jpn. J. Ser. 5, Proc. Micro-Process, pp. 28-32 (1991).

[137] H. Watanabe, E. Sugiura, T. Imoriya, Y. Todokoro, and M. Inoue, "Detection and printability of shifter defects in phase shifting masks II. Defocus characteristics," Jpn. J. Appl. Phys., vol. 31, p. 4155 (1992).

[138] H. Watanabe and Y. Todokoro, "Phase-shifting lithography: Mask making and its application," J. Vac. Sci. Technol. B, vol. 11, no. 6 (1993).

[139] M. Watts, "A high sensitivity two layer resist process for use in high resolution optical lithography," Proc. of SPIE, vol. 469, pp. 2-10 (1984).

[140] W. Weinstein, "Images of Incoherently Illuminated Bright and Opaque Disks," J. Opt. Soc. Am., vol. 45, p. 1006 (1955).

[141] D. W. Widmann, "Quantitative Evaluation of Photoresist Patterns in the 1μm Range," Applied Optics, vol. 14, no. 4, pp. 931-934 (1975).

[142] D. W. Widmann and H. Binder, "Linewidth Variations in Photoresist Patterns on Profiled Surfaces," IEEE Trans. Elec. Dev., vol. ED-22, no. 7, pp. 467-471 (1975).

[143] J. Wiley, "The Printability of 5x Reticle Submicron Defects," Bay Area Chrome Users Society (BACUS) Eight Annual Symposium, Microlithography News, pp. 15-22 (1988).

[144] J. Wiley, "Process Effects in 5x Reticle Defect Printability," Bay Area Chrome Users Society (BACUS) Ninth Annual Symposium, Microlithography News, pp. 168-175 (1989).

[145] J. Wiley, T. Fu, T. Tanaka, S. Takeuchi, S. Aoyama, J. Miyazaki, and Y. Watakabe, "Phase Shift Mask Pattern Accuracy Requirements and Inspection Technology," Proc. of SPIE, vol. 1464, pp. 346-355 (1991).

[146] J. Wiley and J. Reynolds, "Device yield and reliability by specification of mask defects," Solid-State Technol., July 1993.

[147] D. Williamson, "The Elusive Diffraction Limit," OSA Proc. on EUV Lithography, vol. 23, pp. 68-76 (1994).

[148] G. Wojcik, D. Vaughn, and L. Galbraith, "Calculation of Light Scatter from Structures on Silicon Surfaces," Proc. of SPIE, vol. 772, pp. 21-31 (1987).

[149] G. Wojcik, J. Mould, R. Ferguson, R. Martino, K.K. Low, "Some image modeling issues for I-line, 5x phase shifting masks," Proc. of SPIE, vol. 2197, pp. 455-465 (1994).

[150] E. Wolf, "New Spectral Representation of Random Sources and the Partially Coherent Fields that They Generate," Optics Communications, vol. 38 No. 1, pp. 3-6 (1981).

[151] S. Wolf and R. N. Tauber, *Silicon Processing, vol. 1*, (Lattice Press, Sunset Beach, CA, 1986), p. 192.

[152] A. Wong, "Rigorous Three-Dimensional Time-Domain Finite-Difference Electromagnetic Simulation," Ph.D. Dissertation, University of California at Berkeley (1994).

[153] A. Wong, Tempest Users' Guide, Memorandum No. UCB/ERL M95/14, University of California, Berkeley, March 1995.

[154] K. Yamanaka, H. Iwaski, H. Nozue, and K. Kasama, "NA and s Optimization for High NA I-line Lithography," Proc. of SPIE, vol. 1927, pp. 310-319 (1993).

[155] K. S. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media," IEEE Trans. Ant. Prop. vol. 14, pp. 302-307 (1966).

[156] M. Yeung, "Modeling high numerical aperture optical lithography," Proc. SPIE, vol. 922, pp. 149-167 (1988).

[157] M. Yeung, D. Lee, R. Lee, and A. Neureuther, "Extension of the Hopkin's theory of partially coherent imaging to include thin-film interference effects," Proc. SPIE, vol. 1927, pp. 452-462 (1993).

[158] M. Yeung, "Three-dimensional reflective-notching simulation using multipole accelerated physical-optics approximation," Proc. of SPIE, vol. 2440, pp. 395-409 (1995).

[159] M. Yeung, "Validity of the Classical Theory of Spontaneous Emission and The Fast Multipole Method for Electromagnetic Scattering," Ph. D. Dissertation, University of California, Berkeley, Memorandum No. UCB/ERL M95/112, (1995).

[160] M. Zuniga, G. Walraff, and A. Neureuther, "Reaction Diffusion Kinetics in Deep-UV Positive Tone Resist Systems," Proc. of SPIE, vol. 2438, pp. 113-124 (1995).

[161] L. S. Zurbrick and W. Henke, "The Significance and Detection of Transmissive Defects on 5x Reticles," 1993 IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop (IEEE Press, New York, 1993).