SAND2015-8794R

First Principle Simulation of Diffraction based Metrology Techniques

S. Sohail H. Naqvi, John R. McNeil, Richard Krukar, Paulo R. G. Franco
Center for High Technology Materials
Department of Electrical Engineering
University of New Mexico
Albuquerque, NM 87131

Robert Leland
Sandia National Laboratories
Albuquerque, NM

Ken P. Bishop
Company X
Albuquerque, NM

Abstract

Robert, I would appreciate your comments on this

Sohail

ROBERT LELAND
Div. 1424
Sandia National Lab.
Albuquerque, NM 87185

Introduction

Light scattered from periodic structures carrier useful information about the diffracting structure. The existence of the periodic structure provides an enhanced sensitivity to changes in the shape of the grating, and numerous metrology applications in microelectronics manufacturing have evolved over the years to exploit this behavior of the scattered fields. Diffraction based techniques now exist for alignment [1], overlay [2], temperature measurement [3], latent image focus [4'5] and exposure [6] measurement, post exposure bake monitoring [7] and numerous dimensional parameter control applications [8'9]. These applications exploit the sensitivity of the diffraction pattern to changes in the value of the parameter being measured. An attractive feature of these techniques is that they are rapid, non–contact, non–destructive and quantitative. An additional advantage is that they are greatly amenable to real–time and in–situ applications.

The continued shrinkage of dimensions is placing an unprecedented demand on the various metrology techniques. For diffraction based techniques it is clear that a fundamental understanding of the diffraction phenomena is essential to the continued use of these methods as critical dimensions approach a quarter micron. Using scalar diffraction theory, the far–field diffraction pattern can be calculated from the Fraunhoffer diffraction integral [10]. It is well know that as the dimension of the diffracting structure becomes comparable to the wavelength of the incident field, use of scalar diffraction theory can no longer be justified [11]. This is illustrated in Figure 1 which is a plot of the power diffracted into the reflected 0–order, as a function of grating depth. The diffracted powers are calculated for a monochromatic ($\lambda = 0.6328\ \mu m$) plane wave normally incident on an equal line and space etched silicon (n = 3.84) grating having a period of $1.5\mu m$. The three curves are for a scalar theoretic prediction and rigorous vector theory predictions for TE and TM polarization cases.

It can be shown using scalar theory that, for the grating under consideration, the fraction of incident power diffracted into the reflected 0–order obeys the following relationship [9],

$$P_0 \propto \left| 1 + Exp\left[j\frac{2\pi h}{\lambda/2} \right] \right|^2 , \qquad (1)$$

where h is the height of the grating and λ is the wavelength of incident radiation. The proportionality factor is found by noting that in the limit as the height of the grating becomes very small, P_0 is simply given by the Fresnel Formula [Born and wolf page 40]

$$P_0 \Big|_{h=0} = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2 = 0.344 \ . \qquad (2)$$

Here n_1 is the refractive index of the incident medium (air), and n_2 is the refractive index of the grating region (silicon). The vector theoretic calculations were done using rigorous coupled wave theory, described in the following section. Figure 1 illustrates that for grating depths less than $\lambda/4$, reasonable results are obtained using scalar theory. For deeper gratings, a vector theoretic approach is required. Figure 1 also illustrates how scalar theory may be used to describe some general characteristics of the diffracted signal especially for the case of TE polarization. In most cases, the diffraction behavior, for the TM polarization case, can only be described using a rigorous vector approach.

Rapid advances in computational capabilities, due, in part at least, to the shrinkage in device dimensions, have made it possible to obtain solution to many of the diffraction problems previously considered to be intractable. Numerous techniques have been developed over the years for the solution of the grating diffraction problem. An excellent overview of these techniques may be found in [12'13]. A large number of solution methods may be broadly classified into integral and differential methods. Both methods are equally rigorous, however, they differ widely in ease of implementation and applicability to a broad class of problems. Due to the mathematical complexity associated with integral methods, they are not widely used. Among the differential methods Rigorous Coupled Wave Theory (RCWT) developed by Moharram and Gaylord, [12] has proved to be particularly useful because of its ease of use and wide range of applicability. It will be discussed in detail in a later section. An equivalent differential technique, [14], (known as the modal method) that utilizes waveguide modes for the expansion of the fields inside the grating region was developed by Burchardt [15] and Kaspar [16]. This method has been used and extended by Nyyssonen and Kirk, [17] and by Yuan and Strojwas [18] for the rigorous simulation of images obtained in a microscope and simulation of alignment signals.

Both RCWT and modal method utilizes a Fourier expansion of the refractive index in the grating region. Thus, these techniques work well for gratings having a continuously varying refractive index (e.g. latent image gratings) and grating where the extent of the refractive index discontinuity is not large (e.g. dielectric and semiconductor gratings). In cases where the refractive index discontinuity is large, the analytical waveguide model developed by Botten et. al. [19] and extended by Davidson, [20] has been shown to have improved convergence characteristics [14]. Integral methods also do not utilize a Fourier based expansion. In these methods, the field is represented by a convolution product involving a Green function deduced from the propagation equation.

Two additional solution techniques should be mentioned here because of their versatility in solving nearly all singly and doubly periodic grating diffraction problems. The tradeoff is in the computationally intensive requirements of both techniques. The first technique utilizes Finite Element Analysis and has been developed by Wojck et. al. [21] for the simulation of metrology and alignment signals in IC–related applications. The second technique, TEMPEST developed at Berkeley, [22] is a massively parallel computer solution of the vector diffraction problem using a time–domain, finite–difference method. TEMPEST is implemented on the massively parallel architecture of the Thinking Machines Corp. Connection Machine. Recently there has been an effort to compare the various grating simulation techniques [23] and a good preliminary agreement was achieved between the Finite Element Method, the analytic waveguide method and the coupled wave approach for resist on silicon diffraction gratings.

In the sections that follow a description of RCWT is provided along with its application to various diffraction based metrology techniques. These applications include latent image diffraction analysis, phase–shift mask characterization, large pitch etched structure characterization and optical microscopy. Recently a novel technique has been proposed for the characterization of grating structures [24]. The grating pattern is illuminated with a laser beam and the power diffracted into a particular order is measured as the angle of incidence is varied over some range. Slight changes in dimensional parameters of the grating (e.g. linewidth, sidewall angle, height) change the shape of the angular scan curve. Assuming no two parameter sets produce the same curve, a prediction of grating

dimensional parameters can be made. The novel aspect of this work is that the prediction algorithm is trained on angular scan curves generated using RCWT [25]. A discussion of the technique as well as some recent results will be given in a later section.

Rigorous Coupled Wave Theory

Consider the diffraction grating shown in Figure 2 . The grating is assumed to be infinite dimensional with periodicity in the x–direction and no variation in the structure in the y–direction. Practically, 12 to 20 lines are considered to be a reasonable approximation to an infinite structure. The grating structure is divided into different regions. The parameters of interest are defined in Figure 2. The model considers an electromagnetic plane wave obliquely incident on a diffraction grating placed on a P layer dielectric stack. Region 1 is the incident region, region2 is the grating region and regions 3 represents the P–layer dielectric stack, while the substrate is region4. In order to consider arbitrary profiles, the grating region is sliced into L arbitrarily thin layers, each having a refractive index profile varying in a periodic manner in the x–direction and invariant in the z–direction.

The incident field is an obliquely incident TE polarized plane wave having the E–field vector in the direction parallel to the grating lines. The reader is referred to [26] for the TM and arbitrary polarization cases. The periodic structure produces both forward and backward diffracted waves. The field in the incident region, (region 1) can be described as the sum of the incident plane wave and multiple backward diffracted orders as

$$E_1(x,y,z) = Exp[-j(k_{0x}x + k_{0z}z)] + \sum_{i\,=\,-N}^{N} R_i Exp[-j(k_{ix}x - k_{1iz}z)], \qquad (3)$$

where $k_{0x} = k\sqrt{\epsilon_1}\sin\theta$, $k_{0z} = k\sqrt{\epsilon_1}\cos\theta$ and $k = 2\pi/\lambda$. Here θ is the angle of incidence, λ is the wavelength in free–space, and N is the number of positive and negative kept in the analysis. R_i are the coefficients of the i'th order backward diffracted waves. From Floquet's theorem it is well know that the diffracted field must have the same periodicity as the grating structure and

$$k_{ix} = k_{0x} - i\frac{2\pi}{d}, \qquad (4)$$

where k_{ix} is the x–component of the ith diffracted plane wave and d is the period of the grating. If the order is propagating, $k_{ix} = k\sqrt{\epsilon_1}\sin\theta_i$ and substituting for k_{0x} it can be seen that equation 4 is simply the general representation of the grating equation,

$$\sin\theta_i = \sin\theta - i\frac{\lambda}{d}. \qquad (5)$$

In region 1 the diffracted plane waves have wave vectors with magnitudes $|k_{1i}| = k\sqrt{\epsilon_1}$ and therefore, $k_{1iz} = \sqrt{(k^2\epsilon_1 - k_{ix}^2)}$. Applying coupled wave analysis, in the grating region (region 2), the field for each n'th slice may be expanded in terms of it's space harmonic components in the periodic structure and expressed as

$$E_{2n}(x,y,z) = \sum_{i=-N}^{N} S_{in}(z)Exp[-j(k_{ix}x + k_{0z}z)], \tag{6}$$

where n is the sliced layer index, i is the space–harmonic index and $S_{in}(z)$ are the space–harmonic field amplitudes for the n'th slice. Each wave in this expansion represents an inhomogenous plane wave inside the grating region corresponding to the diffracted orders outside the grating. The essential difference between RCWT and modal methods is in the expansion of the electromagnetic fields in the grating region. In any region, the total field must satisfy the wave equation. For TE polarization, we have

$$\nabla^2 E_{2n} + k^2 \epsilon_n(x,z) = 0, \tag{7}$$

where $\epsilon_n(x,z)$ is the Fourier series expansion of the index profile within each slice and is given by

$$\epsilon_n(x,z_n) = \epsilon_0 \sum_{l=-\infty}^{\infty} \hat{\epsilon}_{nl} Exp\left[j\frac{2\pi l x}{d}\right]. \tag{8}$$

Substituting equations (6) and (8) into (7), a set of state space equations are obtained whose solutions may be expressed as

$$S_{in}(z) = \sum_{m=1}^{2M} C_{mn} w_{imn} Exp(\lambda_{mn} z), \tag{9}$$

where M = 2N + 1.

Here w_{imn} and λ_{mn} are the eigenvectors and eigenvalues obtained from the solution of the differential equation for the n'th planar grating slab. C_{mn} are the unknown coefficients to be determined. Application of the boundary conditions at each slice interface requires the continuity of the tangential electric and magnetic fields that yields a set of linear equations. The grating is placed on top of the dielectric stack and the last boundary condition equation is obtained by matching the fields at the L'th grating slice and region 3 interface. The effect of field scattering in the multilayered dielectric stack layer is incorporated in the last boundary condition in the following manner. The field in the p'th dielectric layer, beneath the grating structure, consists of multiple backward and forward diffracted orders [27] which may be expanded as,

$$E_p(x,y,z) = \sum_{i=-N}^{N} (X_{ip}Exp[-jk_{ipz}(z-w_p) + Y_{ip}Exp[jk_{ipz}(z-w_p)])Exp[-j(k_{ix}x)], \tag{10}$$

for p = 3,1 .... 3,P. Here $k_{ipz} = (k^2 \epsilon_{\ p} - k_{ix}^2)^{1/2}$, and w_p is the height of the p'th region. For each p'th region, if we define $X_{ip} = D_{ip}Y_{ip}$ and enforce appropriate boundary conditions the following recursive relationship for D_{ip} is obtained [28],

$$D_{ip} = Exp(j2k_{ipz}w_p)\left[\frac{E_{ip}+1.0}{E_{ip}-1.0}\right], \qquad \text{for} \quad p = 3,1 \ \ldots \ 3,P \tag{11}$$

where

$$E_{ip} = \begin{cases} \dfrac{k_{ipz}}{k_{i(p+1)z}} \left[\dfrac{D_{i(p+1)} + 1.0}{D_{i(p+1)} - 1.0} \right] & \text{for } p = 3,1 \ \ldots \ 3,P-1 \\ \dfrac{k_{iPz}}{k_{i4z}} & \text{for } p = 3,P \end{cases} \tag{12}$$

The coefficient D_{ip} incorporates the effect of the dielectric stack and reduces the number of unknowns. Thus, the L'th grating slice, region 4 boundary can be effectively considered as the last interface. The total set of linear equations obtained may be solved using Gaussian Elimination to yield the reflection coefficients R_i.

An important consideration in the solution of the diffraction problem using RCWT is the choice of N, the maximum number of diffracted orders kept in the analysis. Approximate memory requirements using N diffracted modes is $16(2N + 1)^2$ bytes. The computer time required to solve one diffraction problem increases as N^3 and therefore it is important to choose the correct value of N to provide an accurate solution in a reasonable amount of time. The rate of convergence of solution with increasing N is primarily determined by the difference in refractive index of the line and space in the grating region, and to lesser extent by the incident field wavelength to grating period ratio, λ/d. A rule of thumb relationship that is useful in determining the correct value of N that should be used for dielectric gratings, is

$$N = \left| \left(n_2 - n_1 \right)^{1.6} \right| * \left| \left(\frac{d}{\lambda} \right)^{1.25} \right|, \tag{13}$$

where, |x| is the ceiling of x and is the smallest integer larger than x. n_2 and n_1 are the refractive indices of the line and space portions of the grating respectively. The value of N chosen should always be confirmed by increasing the value of N to see if a significant change in diffraction efficiencies occur.

Modeling of Optical Microscope Images

Optical microscopy has traditionally been the metrology technique of choice in semiconductor manufacturing, but with the continuing reduction in feature size of semiconductor components, the interpretation of the signal obtained in the microscope has become more and more difficult. Several researchers have addressed the issue of measurements of submicrometer features using optical imaging. A complete description of techniques used to model the image obtained in an optical microscope may be found in [29], and is summarized below:

1. Calculate the incident light pattern on the object's surface, using an imaging model. To perform this task Abbe's theory, a scalar approach, is used.

2. Using a scattering model, calculate the reflected or transmitted light pattern due to the presence of a grating surface .

3. Calculate the collected light pattern on the surface of the detector using the same imaging model as in step 1.

In a scalar scattering model the object is described by a planar complex reflectance and transmittance. This model was initially used by Nyyssonen et. al. [30] to determine the characteristics of the optical image of a line as observed in an optical microscope.

Since accuracy of the optical metrology tool should be 1/10 of the measured width and based upon the fact that the scalar theory can only be applied to objects that have thickness smaller than $\lambda/4$, Nyyssonen [31] adopted the waveguide model to calculate the reflected fields from semiconductors. Kirk and Nyyssonen [17] improved the model allowing it to predict the optical microscope images of line objects with arbitrary edge geometry. Yuan et al [18] extended the model to include arbitrary polarization and a large NA lens. This simulation software, known as METRO, is available from Carnegie Mellon University. The analytical waveguide model has also been used to simulate the scattered field and it is the basis for the optical microscope simulation software Metrologia that is distributed by Spectel Co. [20].

Recently the METRO package utilizing the modal approach for the scattered fields has been modified to use RCWT for the calculation of the scattered fields. Figure 3 is a plot of the image that would be obtained in an optical microscope using Kohler illumination. The object is seen through a bright field optical microscope with partially coherent illumination (S=2/3) and objective's NA=0.90. Illumination is done by an Arc Hg lamp where the center frequency is 480.0 nm with a bandwidth at FWHM of 150 nm. The sample under consideration is a developed photoresist (n = 1.656) grating on a thin layer of nitride (n = 2.044) placed on a Silicon substrate. The photoresist grating has a period of 1.2 μm, linewidth of 0.6 μm and is 1.0 μm thick. The three plots simulate the effect of nitride thickness variation on the optical microscope profile. It can be seen that if the same threshold was used in the three cases, a different value of linewidth would be obtained. Figure 4 is a plot of the linewidth that would be obtained if the threshold was defined using the microscope plot for the 75nm nitride thickness case. This is one of the reasons why optical microscope are not used to measure sub–micron photoresist linewidths on a layer of nitride.

Analysis of Diffraction from Latent Image Gratings

The latent image consists of the distribution of photoactive compound (PAC) concentration within the photoresist layer. This represents a change in chemical distribution inside the photoresist which is induced by the aerial image projected by the exposure tool, [32]. The index of refraction in the photoresist is linearly dependent on the PAC concentration, [33], and consequently it is a function of both exposure and focus. A latent image grating is formed by exposing the photoresist using a grating photomask in the exposure tool, or using holographic exposure techniques, thereby causing a periodic variation in the PAC concentration. If we consider two samples exposed with a grating image for different amounts of time or using different focus values, the three–dimensional distribution of refractive indices of the photoresist will be different for the different regions of the phase grating created inside the photoresist. The PAC concentration then forms the basis for the pattern of the photoresist which is subsequently formed as a result of post–exposure bake and development.

The ability to model diffraction from a latent image grating relies on the fact that the refractive index of the photoresist changes with PAC concentration. The PAC concentration is a function of the amount of light absorbed which depends on the aerial image, focus, exposure dose and the optical

properties of the resist and underlying films. Figure 5 is an example of the PAC concentration calculated using PROLITH. A value of 1 corresponds to unexposed resist, while a value of 0 corresponds to fully exposed resist. As a first approximation, the refractive index profile inside the photoresist can be assumed to be linearly dependent on the PAC concentration profile.

To calculate the amount of power diffracted into the various orders from a latent image grating, we combine the lithography simulation program with rigorous coupled wave analysis. A piecewise constant approximation to the continuously varying PAC profile in the z direction is created by slicing the PAC profile into very thin slices, along planes parallel to the x–y plane. In each slice, the PAC profile can be assumed be constant with depth z. This PAC concentration profile is incorporated into the diffraction analysis program using the procedure described in [34]. The imaged pattern is a grating periodic in the x–direction. Following the treatment in [34], the index variation in the n'th grating slice varies as a function of x only and is expanded in a Fourier series. The analysis procedure is slightly modified to accommodate the continuous variation in refractive index inside the photoresist due to a continuously varying PAC concentration profile. The Fourier series expansion of the refractive index profile is written as

$$\epsilon_n(x, z_n) = \epsilon_0(\epsilon_2 - \epsilon_3) \sum_{l=-\infty}^{\infty} \hat{\epsilon}_{nl} Exp\left[j\frac{2\pi l x}{d} \right] + \epsilon_0\epsilon_3. \tag{14}$$

In order to calculate the l'th Fourier series coefficients, $\hat{\epsilon}_{nl}$, of each n'th grating slice from the PAC concentration profile, a sample–and–hold method is used. An example of an arbitrary slice and its sampled approximation is shown in Figure 6 . Each Fourier coefficient is calculated using

$$\hat{\epsilon}_{l,n} = \begin{cases} \dfrac{1}{j2\pi l} \sum_{q=0}^{Q-1} f_q \left\{ \exp\left[-j\dfrac{2\pi q l}{Q} \right] - \exp\left[-j\dfrac{2\pi l(q+1)}{Q} \right] \right\} & l \neq 0 \\ \dfrac{1}{Q} \sum_{q=0}^{Q-1} f_q & l = 0, \end{cases} \tag{15}$$

where Q is the total number of samples in the x–direction and f_q is the value of each sample at each location as shown in Figure 3. No other modification is made to the analysis procedure. Additional details of the analysis procedure are provided in [35].

The shape of the developed photoresist structure is critically dependent on the focal position chosen during exposure. Optimal focus is defined to be the focal position which results in near vertical sidewalls in the developed photoresist. It is reasonable to assume that optimal focus conditions will be obtained when the contrast of the latent image is maximized. Specifically it is the PAC concentration contrast, $C_{PAC}(z_i)$, that needs to be maximized. At any location inside the photoresist this contrast may be defined in terms of the maximum, $m_{\max}(z_i)$, and minimum, $m_{\min}(z_i)$, PAC concentrations at z_i as

$$C_{PAC}(z_i) = \frac{m_{\max}(z_i) - m_{\min}(z_i)}{m_{\max}(z_i) + m_{\min}(z_i)} . \tag{16}$$

Because of the standing wave effects of the exposing radiation, C_{PAC} varies with z, and an average PAC concentration can be defined as,

$$C_{avg} = \frac{1}{N}\sum_{i=1}^{N} C_{PAC}(z_i) \quad . \tag{17}$$

This number, C_{avg}, represents a figure of merit for the overall contrast or quality of the latent image. The larger C_{avg} is, the larger the PAC gradient is in the x–direction of the line structure, and the greater the slope, or sharpness of the developed photoresist lines, which is the intended purpose of improving the lithography process.

Figure 7 is a plot of the average contrast as a function of focus for three different samples. In each case the photoresist was 1.1 μm thick, and C_{avg} was maximized at a focal position of –0.3 μm. This value is consistent with results obtained in [32] and [4]. Note that the focal position is defined such that 0.0μm corresponds to the top (surface) of the photoresist layer, and negative locations are inside the photoresist. The three wafer samples considered in Figure 6 were: photoresist on a bare Si wafer, photoresist on a silicon nitride layer 1.0 μm thick on a Si wafer, and photoresist on a layer of polycrystalline silicon 0.25 μm thick on a layer of silicon oxide 0.155 μm thick on a Si wafer. The latent image is a periodic structure (i.e. a grating) having a nominal linewidth of 1.0 μm and a pitch (period) of 2.0 μm. The exposure doses used to model the wafers were 150 mJ/cm2 for the first sample, 110 mJ/cm2 for the second sample, and 120 mJ/cm2 for the third sample. Simulations have shown that within a wide range of exposure values, the focus position for which C_{avg} is maximized remains unchanged.

Figure 8 illustrates the intensities of the 1st–order diffracted beams from latent images plotted versus exposure tool focus using rigorous coupled wave diffraction analysis, together with PROLITH for the same three sample structures considered in connection with Figure 6. The diffracted intensities reach a maximum at the focal position of approximately –0.3 μm, the same focal position which resulted in a maximum C_{avg} in Figure 6. This plot clearly demonstrates that optimal focus location can be determined by monitoring the diffracted signal from a latent image grating. Experimental verification of this optimal focus location technique is provided in [5].

The optimal exposure value for any lithographic process is defined as the exposure value which results in the developed photoresist line having the desired CD. It is well known that the optimal exposure value for each wafer depends on the optical properties of each sample, as discussed in [6]. Using the combination of PROLITH and RCWT discussed previously it is possible to associate a latent image diffraction value with each developed CD. Figure 9 is a theoretical plot of the fraction of incident power diffracted into the first order as a function of exposure dose. It can be seen that the diffracted power increases monotonically with exposure dose over a wide range of values which includes the optimum exposure range. Figure 10 is a theoretical plot of the CD that would have been obtained had the exposure been stopped at the indicated 1st–order diffracted power value. Thus, by monitoring the latent image diffraction signal "in–situ", exposure would be stopped when the desired diffraction level is reached: resulting in optimum CD of the developed line.

Towards Solving the Inverse Diffraction Problem:

The previous discussion has shown that the amount of incident power diffracted power diffracted into the various orders can be accurately predicted for an arbitrarily shaped grating placed

on top of a dielectric stack. The inverse of this problem is: ***Given the diffraction efficiencies, is it possible to predict the shape of the scatterer ?*** If no apriori assumptions are made about the shape of the grating, it appears doubtful that an answer to the above question can be found. For any practical problem, however, it should be possible to parametrize the grating. For example, the grating can be assumed to have a trapezoidal shape with unknown height, width, and sidewall angle. These parameters are defined in Figure 11 . Note that Figure 9 represents only one of numerous ways in which the grating may be parametrized. The period of the grating is assumed to be known, since it can be easily measured very accurately. In addition, the number of dielectric layers underneath the grating is known and the index of refraction of these layers may be assumed to be a known quantity. The practical problem of interest now becomes ***Given the diffraction efficiencies, is it possible to estimate various parameters of the grating ?*** This problem has been investigated by numerous researchers over the years; see ref 4. of [8] for a complete list of references. In all these applications scalar theory has been used to estimate the grating parameter of interest, from a measurement of the diffraction signal.

The grating parameter estimation problem may also be viewed as a pattern recognition problem, whereby a distinct diffraction signal is associated with a specific grating shape. In order for this technique to work a large calibration data base of diffraction signals must be generated. In a previous article, in the present scatterometry series, it was shown by Gottscho et. al. [36] that a neural network may be trained to classify the grating from its observed diffraction signal. For that experiment a large number of samples were created with varying sidewall angles. The samples were measured using cross–section SEM and classified to have undercut, overcut or vertical sidewalls. The diffraction signal from these calibration samples were then measured and used to train a neural network. With the availability of first principle diffraction simulation methods, the possibility arises of training the pattern recognition technique on theoretically generated diffraction data. This would eliminate the need for a physical calibration set with the associated time and expense for the calibration of the samples and measurement of the diffraction signals. In addition, the theoretically generated diffraction signals would by definition have no error since they utilize a first principle simulation.

Diffraction Scatter Analysis

The first use of rigorous diffraction theory in the development of a technique to measure grating parameters was for the linewidth determination of a chrome–on–glass photomask gratings [37]. Using RCWT it was observed that the fraction of incident power diffracted into the 0–transmitted order decreased monotonically with increasing linewidth. Assuming a rectangular line–shape of known height for the photomask grating, a one–to–one relationship existed between the linewidth of the grating and the 0–transmitted order power. The technique then simply consisted of shining a laser beam on the diffraction grating and measuring the fraction of incident power diffracted into the 0–transmitted order. Prediction of linewidth was then made using a look–up table that was generated using RCWT. This technique was tested on photomask gratings manufactured by different companies and excellent results were obtained.

In most cases the diffraction signal is a complicated function of the grating parameters and a simple interpretation of the signal is not possible. Consider a 32 μm period etched silicon grating illuminated with a He–Ne laser beam. This is the same problem as was considered in [8]. The

diffracted signal consists of nearly a 100 diffracted orders at angular locations specified by the grating equation. Since the number of diffracted is so large an envelope function may be obtained for the diffraction orders. The diffracted power distribution envelope is a very sensitive function of the grating parameters. Figure 12 is a theoretical plot of this envelope function for rectangular shaped gratings having a linewidth of 1.5 μm and varying height. The plots are displaced vertically, however, it can be clearly seen that the shape of the envelope changes significantly as the grating height is varied.

The strategy in solving the parameter identification problem is conceptually an iterative one. We hypothesize a geometry, compute the associated diffraction pattern, compare this with an actual measurement, adjust the hypothetical geometry, recompute, compare, etc. until satisfactory convergence. The practical approach is to decouple the expensive forward problem of computing the hypothetical diffraction patterns from the inexpensive inverse problem of matching to the actual diffraction pattern so that the recurring cost of monitoring the process is small. Hence a large set of diffraction patterns for likely perturbations of the grating parameters of interest are computed and used as input to a statistical analysis or neural network pattern matching software. In the problem considered above only the height of the grating was assumed to be unknown. The theoretical calibration data was therefore calculated for each height value between 0.6 μm and 1.2 μm, with a discretization step of 0.02 μm. The multivariate statistical technique, Partial Least Square (PLS), was used to train on the simulated scatter data. This method is very robust and not computationally intensive. In addition, in a comparison of similar techniques to the analysis of infrared spectroscopy data, [38], PLS was shown to be the method of choice. Prediction of height is an extremely simple operation and simply involves a matrix–vector multiplication.

Experimentally, the diffraction pattern from gratings etched into bulk silicon of 5 wafers, was measured and a prediction of the height done using PLS trained on the theoretical scatter data. 14 die locations on each wafer were measured. These gratings were known to have vertical sidewalls and a 1.5 μm linewidth. The height of the gratings were verified using a scanning force microscope (SFM). Figure 13 is a plot of the predicted heights and also measured using the SFM. It can be seen that the technique worked to perfection and ... 7nm +– 3nm.

When the period of the grating is much larger than the wavelength of incident radiation, a large number of diffraction orders is obtained and an envelope function can be defined. For short pitch structures only a few, if any, diffracted orders exist and a diffracted power envelope is obtained in a different manner. Now, the envelope function is obtained by measuring the fraction of incident power diffracted into a particular order as the angle of incidence is varied over a certain range. Figure 14 is a plot of the power diffracted into the first order for a 1 micron thick photoresist on Si grating for three different linewidths. The grating period is assumed to have a rectangular shape and has a period of 2 microns. Figure 15 is a plot of the power diffracted into the first order for trapezoidal shaped, 1 micron thick photoresist gratings having different sidewall angles. The top linewidth is 1 micron and again the period is 2 microns. Illumination is assumed to be from a He–Ne laser. These plots illustrate that a distinct change in the envelope functions are obtained as the grating parameters are varied.

To test the diffraction analysis strategy theoretical plots were generated over a larger parameter space. The height of the grating varied between ... and ... with a step size of ... The sidewall was varied from 80 to 90 degrees with a step size of 2 degrees and the linewidth was varied between ..

and ... with a step size of .... Thus a total of ... plots were generated and used to train the PLS routine. The parameter prediction of the routine is given in tabulated form in Table 1.

The forward problem is computationally intense because it requires essentially an integration of Maxwell's equations for the complex geometry associated with each point in a very large parameter space. We have therefore implemented this computation on a massively parallel supercomputer since they can be as much as an order of magnitude more cost–effective than traditional vector supercomputers.

The particular massively parallel machine in use is the 1024–processor nCUBE2 hypercube supercomputer at Sandia National Laboratories. The nCUBE2 is a MIMD machine, thus, unlike the connection machine, each processor acts independently. A d–dimensional hypercube multiprocessor consists of a set of $P = 2^d$ processors, identified by distinct binary numbers from 0 to $2^d - 1$. Information is transmitted between them by passing messages through a network in which wires connect processors whose binary values differ in a single bit. For example, a $d = 3$ hypercube could be represented by a standard cube in which the corners were processors and the edges were connecting wires.

Hypercube multiprocessors enjoy wide popularity because they have several attractive properties: The network is very regular and can be described concisely in a recursive fashion. Each processor is connected to just d communication wires, so larger systems can be engineered than if each processor were directly connected to all others. A message can be routed between any two processors by traversing at most d wires, so a broadcast to all processors can be completed in $O(\log P)$ time as compared with $O(P)$ time for architectures such as meshes which have a fixed number of connections per processor. Finally, to travel between two processors, a message simply uses one wire from each bit in which the two processor's numbers differ. The hypercube architecture therefore offers a well balanced, elegant approximation to the conceptually ideal machine in which each processor is directly connected to all others.

There are two basic goals in using any parallel machine. The first is to assign each processor a comparable amount of work – otherwise some processors will be idle while others are overloaded. The second is to minimize communication between processors during the course of the computation – otherwise processors must in general wait for the requisite data to arrive before proceeding with their computations. An efficient parallel computation must therefore have good load balance and low communication overhead.

Keeping these goals in mind, we have investigated various strategies for parallelizing the forward computation. For challenging problems the parameter space is large because the geometry is intricate, and the simplest and most efficient strategy is to exploit parallelism in this parameter space. That is, each processor is assigned a comparable portion of the parameter space which it integrates nearly independently of the other processors. The processors can in principle operate completely independently and thereby achieve perfect efficiency if input/output requirements are neglected. In practice, memory requirements may make it impossible to run a full integration on one processor, in which case each integration must be distributed across several processors. In addition, portions of the integrations for distinct parameter sets are in general redundant, and this can be exploited to net benefit at the cost of introducing extra communication and algorithmic complexity. We have implemented parallel forward scattering RCWT solvers based on these ideas for singly and double periodic structures and have achieved very high efficiency on the nCUBE hypercube [39].

References

1. S. Wittekoek, H. Linders, and et. al., "Precision wafer stepper aignment and metrology using diffraction gratings and laser interferometry," Proc. SPIE, p. 22, Aug. 1985.

2. R. Pforr, S. Wittekoek, R. Van Den Bosch, L. Van Den Hove, R. Jonckheere, T. Fahner, and R. Seltmann, "In Process image detecting technique for determination of overlay, and image quality for ASM–L wafer stepper," Optical/Laser Microlithography V, vol. SPIE 1674, pp. 594–608, 1992.

3. M. P. Lang, "Real–Time image processing techniques applied to optical non–contact temperature measure ments," M.S. Thesis, University of New Mexcio, 1992.

4. T. E. Adams, "Application of latent image metrology in microlithography," Integrated Circuit Metrology, Inspection and Process Control V, vol. Proc. SPIE 1464, pp. 294–312, March 1991.

5. L. M. Milner, K. P. Bishop, S. S. H. Naqvi, and J. R. McNeil, "Stepper focus characterization using diffraction from latent images," Journal of Vacuum Science and Technology, Submitted2.

6. K. C. Hickman, S. M. Gaspar, S. S. H. Naqvi, K. P. Bishop, J. R. McNeil, G. D. Tipton, B. L. Draper, and B. R. Stallard, "Use of diffraction from latent images to improve lithography control," Journal of Vacuum Sci ence and Technology B, vol. 10, No. 5, pp. 2259–2266, Sep/Oct 1992.

7. J. Sturtevant, S. Holmes, and P. Rabidoux, "Post exposure bake characteristics of a chemically amplified deep–ultravoilet resist," Advances in Resist Technology and Processing IX, vol. Proc. SPIE 1672, pp. 114–124, March 1992.

8. R. Krukar, A. Kornblit, L. Clark, J. Kruskal, D. Lam bert, E. A. Rietman, and R. A. Gottscho, "Reactive ion etching and depth characterization using statistical and neural network analysis of light scattering data," Journal of Applied Physics, p. Submitted.

9. P. Chapados, "In Situ critical dimension sensor for process control," Microlithography world, To be published (1993).

10. M. Born and E. Wolf, in Principles of optics, Pergam mon Press, New York, 1985..

11. N. C. Gallagher and S. S. H. Naqvi, "Diffractive optics: scalar and non–scalar design analysis," Holographic Optics: Optically and Computer Generated, vol. Proc. SPIE Vol. 1052, pp. 32–40.

12. T. K. Gaylord and M. G. Moharram, "Analysis and appli cations of optical diffraction by gratings," Proc. IEEE, vol. 73, pp. 894–937 , May 1985.

13. R. Petit, "Electromagnetic theory of gratings," in Topics in current physics, Springer–Verlag, Berlin, 1980.

14. R. Magnusson and T. K. Gaylord, "Equivalence of multiwave coupled–wave theory and modal theory for periodic–media diffractio," Journal of Optical Society of America, vol. 68, No. 12, pp. 1777–1779, Dec. 1978.

15. C. B. Burchardt, "Diffraction of electromagnetic waves by a sinusoidally stratified dielectric grating," J. Opt. Soc. Am., vol. 56, No. 11, pp. 1502–1509, Nov. 1966.

16. F. G. Kaspar, "Diffraction by thick periodically stra tified gratings with complex dielectric constant," J. Opt. Soc. Am., pp. 37–45, Jan. 1973.

17. D. Nyyssonen and C. P. Kirk, "Optical microscope imag ing of lines patterned in thick layers with variable edge geometry: theory," JOSA A, pp. 1270–1280, Aug. 1988.

18. C–M Yuan and A. J. Strojwas, "Modeling of optical microscope images of integrated–circuit structures," J. of Opt. Soc. of Am. A, vol. Vol. 8, No. 5, pp. 778–790, May 1991.

19. I. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams, and J. R. Adrewartha, "The dielectric lamellar grating," Optica Acta, vol. vol. 28, No. 3, 1981.

20. M. Davidson, "Analytic saveguide solutions and the coherence probe microscope," Microelectronic Engineering, vol. 13, pp. 523–526, 1991.

21. G. L. Wojcik, J. Mould, R. J. Monteverde, J. J. Pro chazka, and J. R. Frank Jr., "Numerical simulation of thick line width measurements by refelected light," Integrated Circuit Metrology, Inspection and Process Control IV, vol. SPIE Vol. 1463, 1991.

22. J. Gamelin, "Simulation of topography scattering for optical lithography with the connection machine," M.S. Thesis, University of California Berkeley, May 1989.

23. G. L. Wojcik, J. Mould Jr., E. Marx, and M. P. Davidson, "Numerical reference models for optical metrology simulation," Integrated Circuit Metrology, Inspection and Process Control VI, vol. Proc. SPIE 1473, pp. 70 82, March 1992.

24. K. P. Bishop, S. M. Gaspar, L. M. Milner, S. Sohail H. Naqvi, and J. R. McNeil, "Grating line shape characterization using scatterometry," International Conference on the Application and Theory of Periodic Structures, vol. SPIE Vol. 1545, pp. 64–73, 1991.

25. R. Krukar, "A methodology for the use of diffracted scatter analysis to measure the critical dimension of periodic structures," Ph.D Thesis, University of New Mexico, 1993.

26. M. G. Moharram and T. K. Gaylord, "Three–dimensional vector coupled–wave analysis of planar–grating diffraction," Journal of the Optical Society of America, vol. Vol. 73, No. 9, Sep. 1983.

27. J. P. Montgomery, "Scattering by an Infinite Periodic Array of Thin Conductors on a Dielectric Sheet," IEEE Trans. Antennas Propagat., vol. AP–23, pp. 70–75, Jan. 1975.

28. S. S. H. Naqvi, "Scattering from a strip grating placed on a multilayer dielectric stack," IEEE/AP–S Symposium, pp. 410–414, Dallas, Texas, May 1991.

29. C–M. Yuan, "Modeling of optical alignment and metrology in VLSI manufacturing," Ph.D Thesis, Carnegie Mellon, 1989.

30. D. Nyyssonen, "Linewidth measurement with an optical microscope: the effect of operating condition on the image profile," Appl. Optics, pp. 2223–2230, August, 1977.

31. D. Nyyssonen, "Theory of optical edge detection and imaging of thick layers," Journal of the Optical Society of America, vol. Vol. 72, pp. 1425–1436, 1982.

32. C. Mack, "Understanding focus effects in submicrometer optical lithography," Optical Engineering, vol. 27, No. 12, pp. 1093–1100, Dec. 1988.

33. C. Mack, private communications.

34. M. G. Moharram and T. K. Gaylord , "Diffraction analysis of dielectric surface–relief gratings," Journal of the Optical Society of America, vol. 72, pp. 1385–1392, Oct., 1982.

35. K. P. Bishop, Use of rigorous diffraction analysis for the characterization of latent–image gratings in photoresist, University of New Mexico, M.S Thesis, July, 1992.

36. R. A. Gottscho, "Reactive Ion Etched Grating profile determination," Microlithography world, To be published (1993).

37. S. S. H. Naqvi, S. M. Gaspar, K. C. Hickman, and J. R. McNeil, "Linewidth measurement of gratings on photomasks: A simple technique ," Applied Optics, pp. 1377–1384, April 1992.

38. E. V. Thomas and D. M. Haaland, "Comparison of multivariate calibration methods for quantitative spectral analysis," Analytical Chemistry, vol. 62, No. 9, pp. 1091–1099, May 1990.

39. S. S. H. Naqvi and R. Leland, "Massively parallel solution of scattering from doubly periodic structures," in OSA Annual Meeting Technical Digest, vol. Vol. 23, p. 182, 1992.

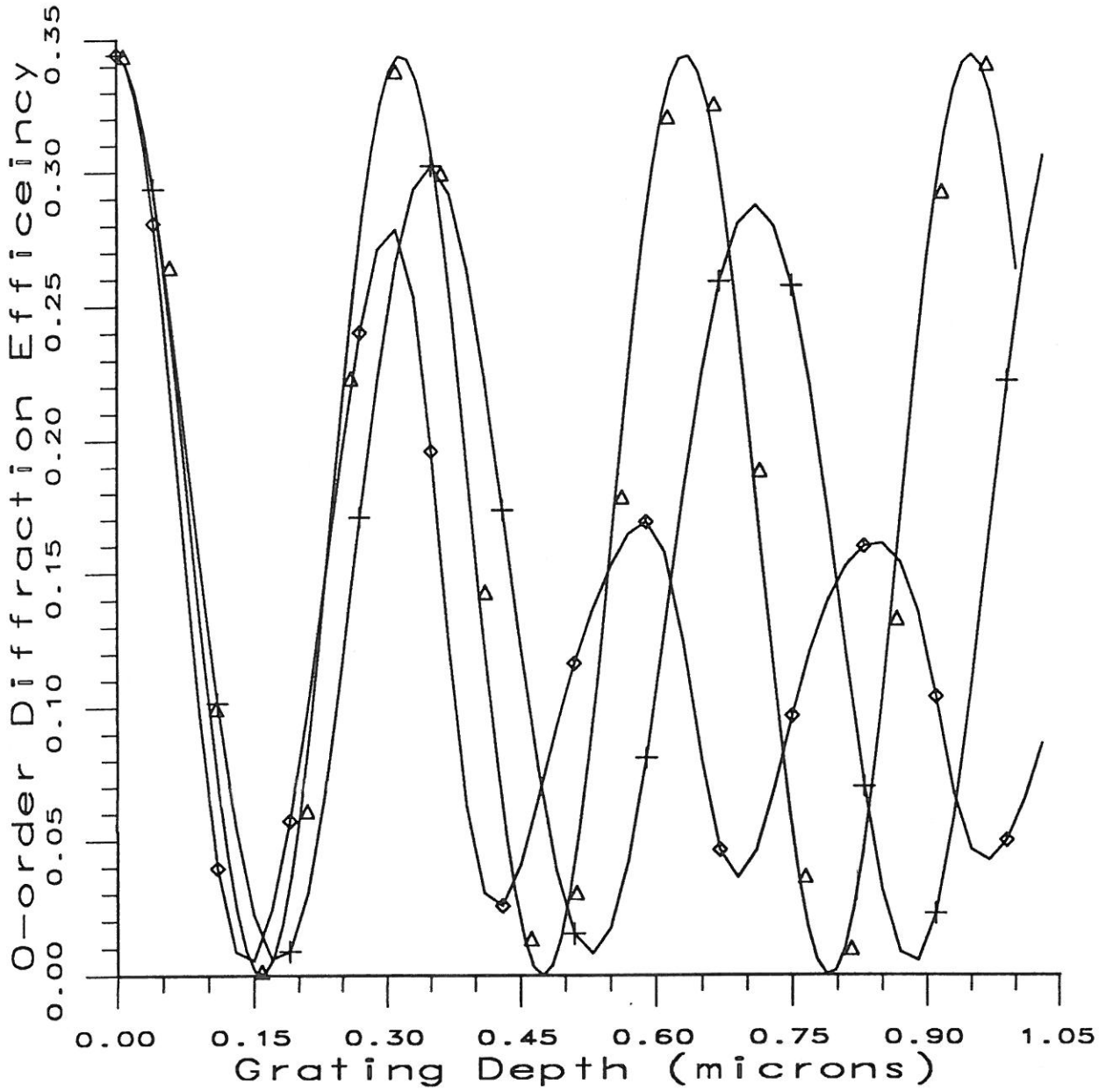

0-order Diffraction Efficiency
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.00
0.15
0.30
0.45
0.60
0.75
0.90
1.05
Grating Depth (microns)

Figure 1:

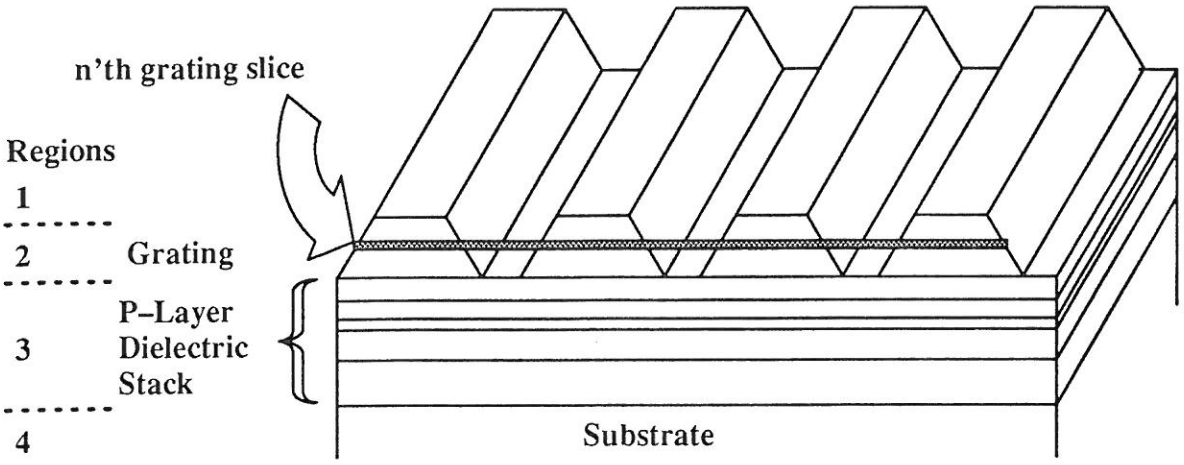
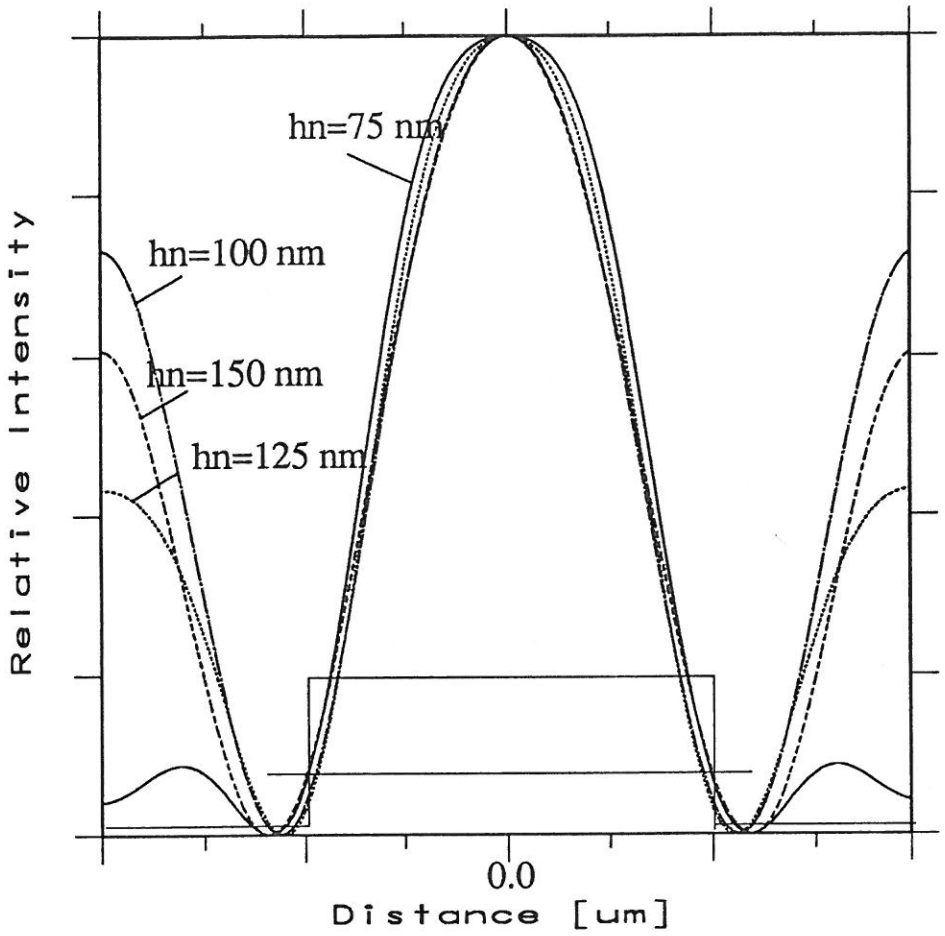

n'th grating slice
Regions
1
2
Grating
P–Layer
Dielectric
3
Stack
4
Substrate

FIGURE 2



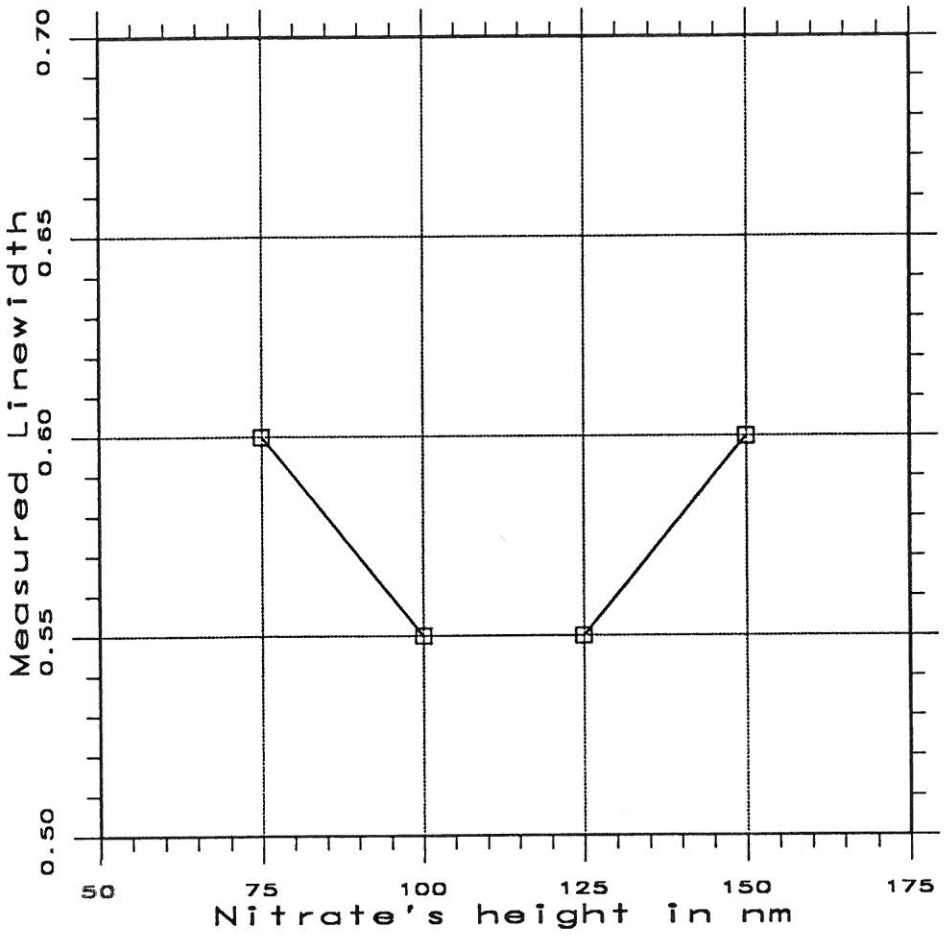
pitch=1.2 width-0.6 hpr=1.0 hnit=var

Kohler illumination:480 nm+/- nm

fullfield optical microscope:NA cond=0.60 NA obj=0.90

Simulation using CW model N=20

Figure 3


Measured Linewidth
0.70
0.65
0.60
0.55
0.50
50
75
100
125
150
175
Nitrate's height in nm

—— □ □ □

Figure 4

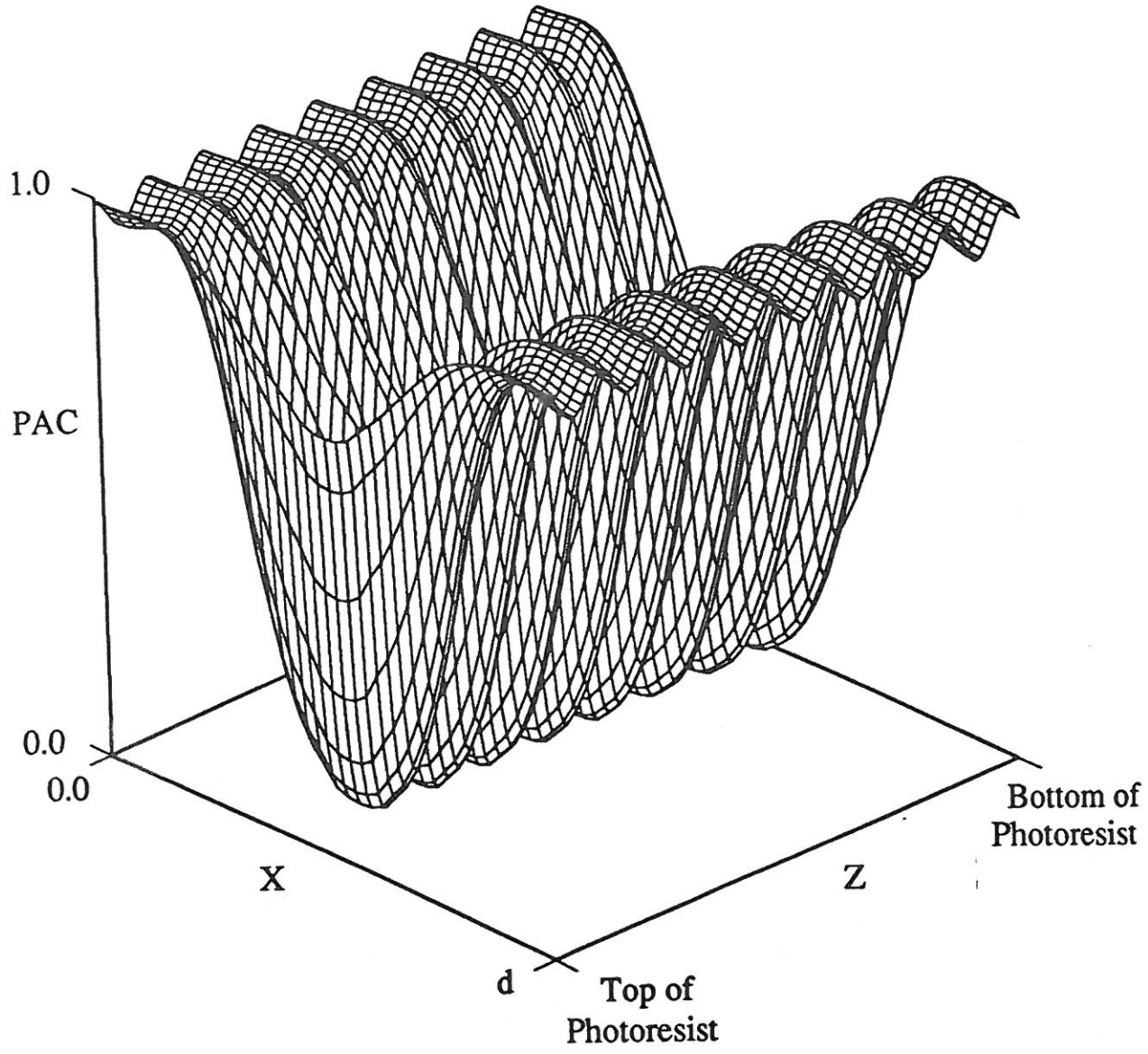


Figure 2.5

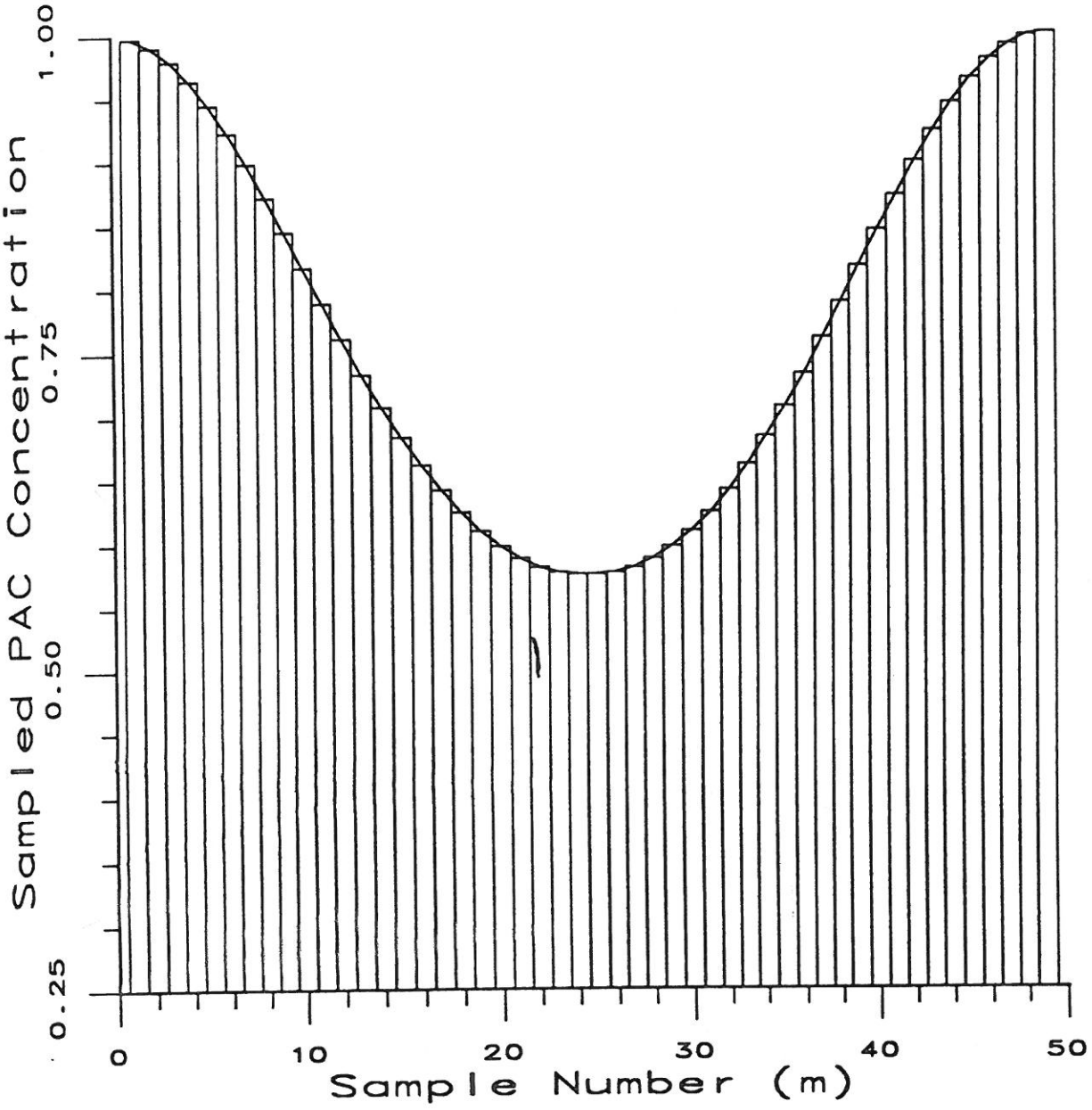

Sampled PAC Concentration
1.00
0.75
0.50
0.25
0
10
20
30
40
50
Sample Number (m)

Figure #6

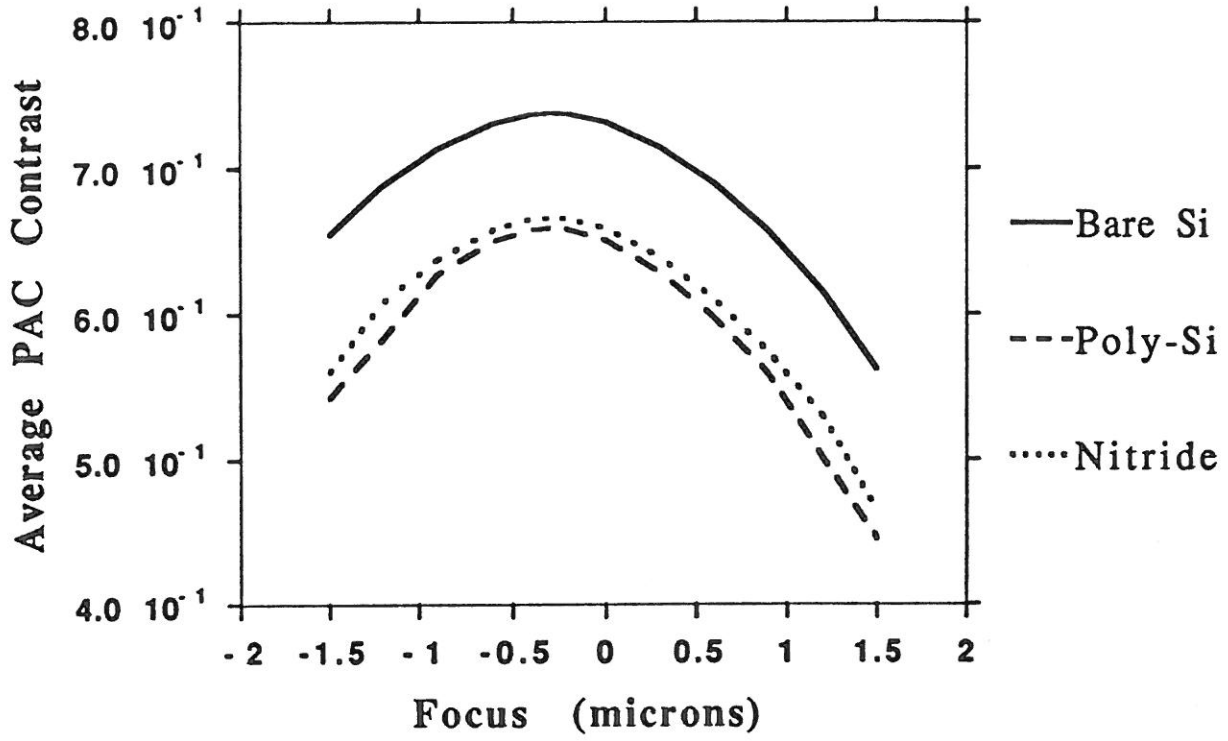

8.0 10⁻¹
7.0 10⁻¹
6.0 10⁻¹
5.0 10⁻¹
4.0 10⁻¹
Average PAC Contrast
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
Focus (microns)
Bare Si
Poly-Si
Nitride

Figure 67

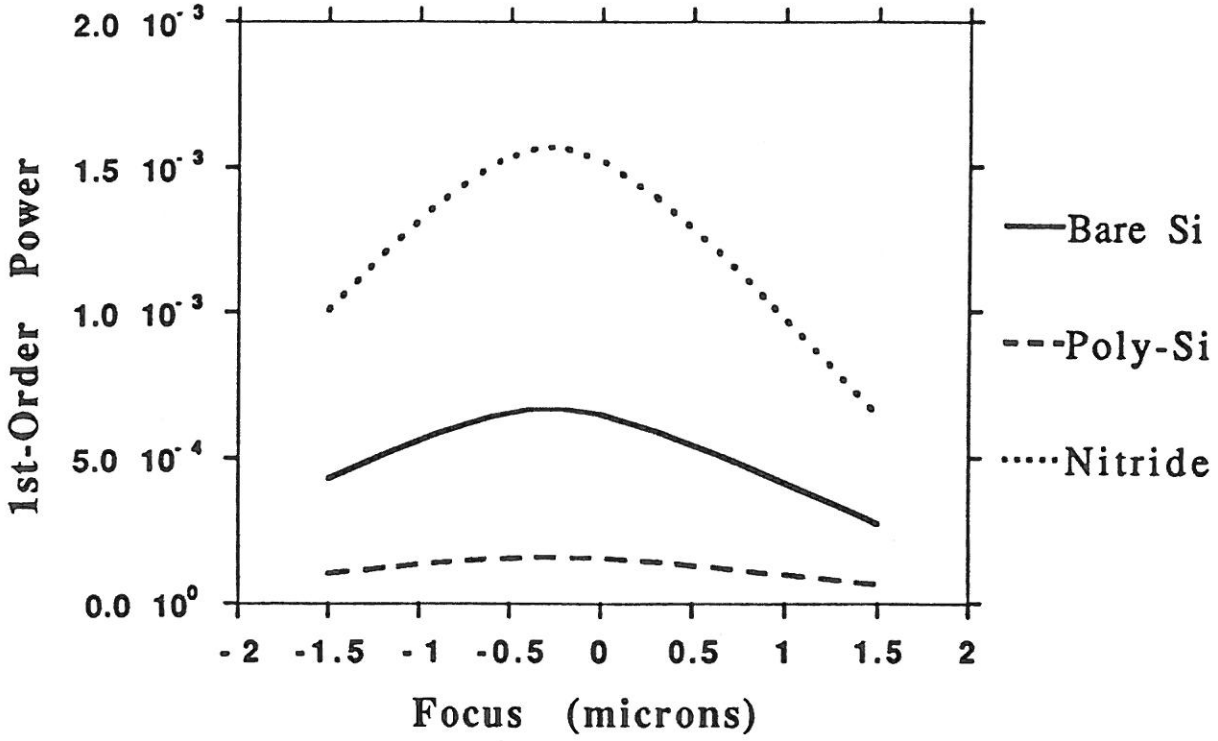

2.0 10-3
1.5 10-3
1.0 10-3
5.0 10-4
0.0 10^0
1st-Order Power
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
Focus (microns)
Bare Si
Poly-Si
Nitride

Figure # 8

Figure 9 from Lisa

Figure 10 from L15A.

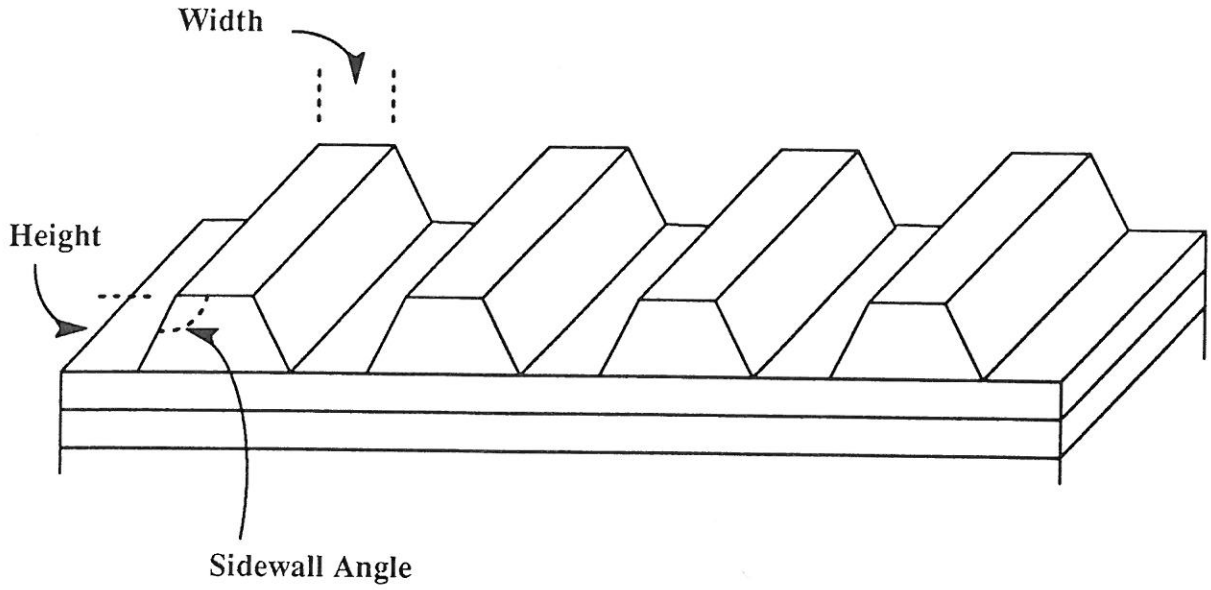
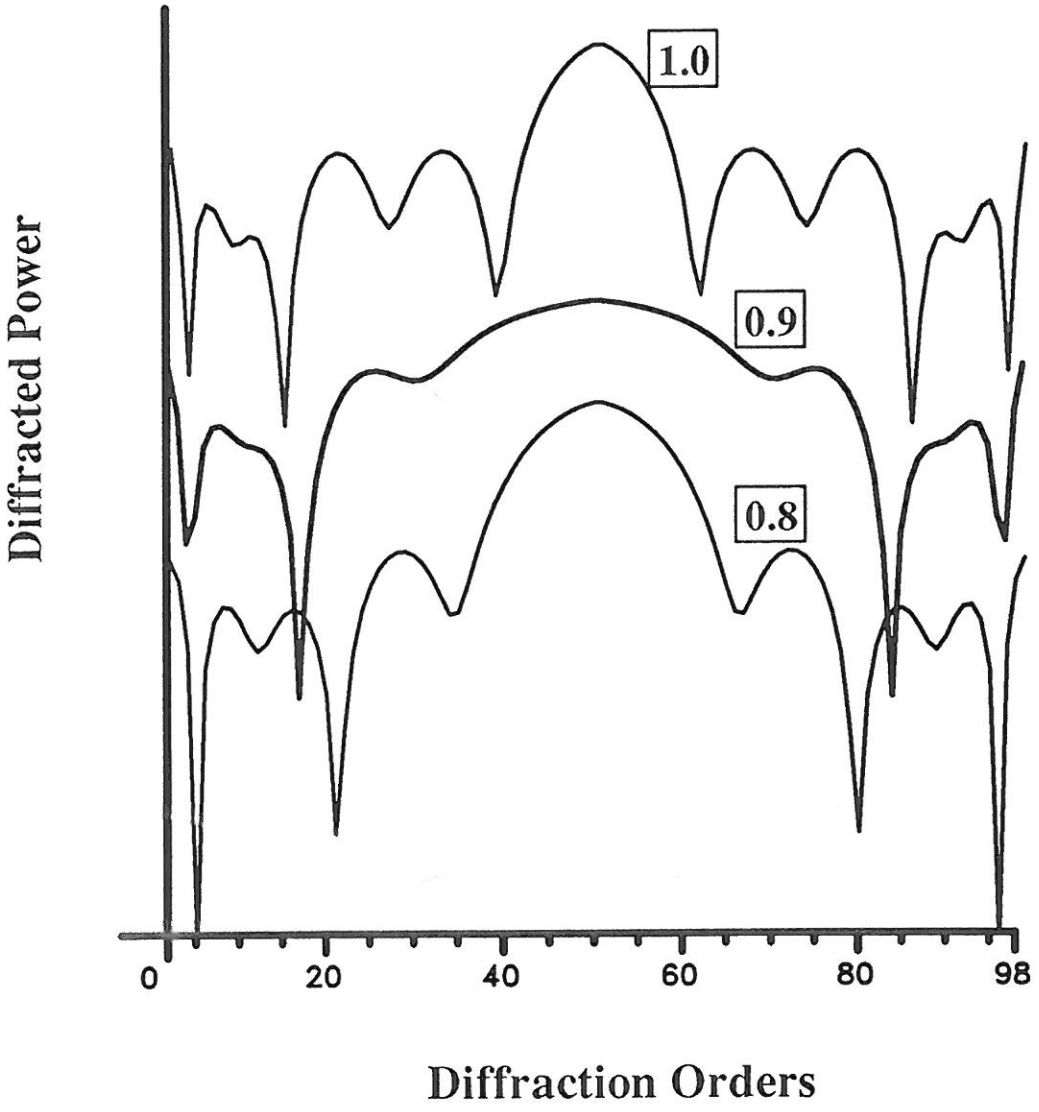


FIGURE 9

THEORETICAL PLOTS

EFFECT OF GRATING HEIGHT



Note: The three Plots are displaced vertically

Figure 12

Plot 13, 14, 15 from Richard.

From naqvi@shazam.eece.unm.edu Mon Jan 18 11:07:28 1993
To: rwlelan@cs.sandia.gov
Subject: SST Article intro

<|,1>Introduction

<Para02>

Light scattered from periodic structures carrier useful information about the diffracting structure. The existence of the periodic structure provides an enhanced sensitivity to changes in the shape of the grating. and numerous metrology applications in microlectronics manufacturing have evolved over the years to exploit this behaviour of the scattered fields. Diffraction based techniques now exist for alignment, overlay, temperture measurement, latent image focus and exposure measurement, post expost bake monitoring and numerous dimensional parameter control applications. An attractive feature of these techniques is that they are rapid, non-contact, non-destructive and quantitative. An additional advantage is that they are greatly amenable to real-time and in-situ applications.

<Para02>

The continued shrinkage of dimensions is placing an unprecedented demand on the various metrolgoy techniques. For diffraction based techniques it is clear that a fundamental understanding of the diffraction phenomena is essential to the contiued use of these methods as critical dimensions approach a quarter micron. Using scalar diffraction theory, the far-field diffraction pattern can be calculated from the fraunhoffer diffraction integral. It is well know that as the dimension of the diffracting structure becomes comparable to the wavelength of the incident field, use of scalar diffraction theory can no longer be justified. For example, let us consider an equal line and space etched silicon grating. According to Fourier theory the power diffracted into the various orders is a function of the linewidth to space ratio only. Specifically the diffracted powers are not a function of wavelngth, period of grating, and polarization. no power should be diffracted into the even orders. In addition, polarization is not accounted for. gratingWhen the size of the diffracting the case of a periodic structure illuminated with a laser beamThese applications exploit the sensitivity of the diffraction pattern to changes in the value of the parameter being measured. The existense of the periodic pattern provides an enhanced s

<Para02>

<|,2>Rapid advances in computational capabilities, due, in part at least, to the shrinkage in device dimensions, have made it possible to obtain solution to many of the diffraction problems previously considered to be intractable. Numerous techniques have been developed over the years for the solution of the grating diffraction problem. An excellent overview of these techniques may be found in [Gaylord Proc. paper and Petit Book]. A large number of solution methods may be broadly classified into integral and differential methods. Both methods are equally rigorous, however, they differ widely in ease of implementation and applicability to a broad class of problems. Due to the mathematical complexity associated with integral methods, they are not widely used. Among the differential methods Rigorous Coupled Wave Theory (RCWT) developed by Moharram and Gaylord has proved to be particularly useful because of its ease of use and wide range

of applicability. It will be discussed in detail in section *&^. An equivalent differential technique (known as the modal method) that utilizes waveguide modes for the expansion of the fields inside the grating region was developed by Burchardt and Kaspar. This method has been used and extended by Nyyssonen and Kirk and by Yuan and Wojck for the rigorous simulation of images obtained in a microscope and simulation of alignment signals.

<Para02>

Both RCWT and modal method utilizes a Fourier expansion of the refractive index in the grating region. Thus these techniques work well for gratings having a continuously varying refractive index (e.g. latent image gratings) and grating where the extent of the refractive index discontinuity is not large (e.g. dielectric and semiconductor gratings). In cases where the refreactive index discontinuity is large, the analytical waveguide model developed by Botten et. al. and extended by Davidson has been shown to have improved convergence characteristics.

<Para02>

Two additional solution techniques should be mentioned here because of their versatility in solving nearly all singly and doubly periodic grating diffraction problems. The tradeoff is in the computationally intensive requirements of both techniques. The first <|,3>technique utilizes Finite Element Analysis and has been developed by Wojck et. al. for the simulation of metrology and alignment signals in IC-related applications. The second technique, TEMPEST developed at Berkeley, is a massively paralle computer solution of the vector diffraction problem using a time-domain, finite-difference method. TEMPEST is implemented on the massively parallel architecture of the Thinking Machines Corp. Connection Machine.