

A Hybrid Analysis of Ellipsometry Data from Patterned Structures

Wei Kong, Hsu-ting Huang, and Fred L. Terry, Jr.

*Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI. 48109-2122*

Abstract. Rigorous coupled wave analysis (RCWA) has been used for modeling the polarization dependent reflection from periodic patterns for process monitoring and control. However, the computational load of this vector method is very heavy. In this paper, we will carefully examine a much simpler scalar method for reflection modeling. We also extend the application of the vector analysis to some special non-periodic structures by combining RCWA with the scalar model. We conclude that this hybrid approach is of significant promise for *in situ* IC production applications.

INTRODUCTION

Advanced semiconductor process development and control requires high accuracy, high speed and non-invasive wafer state monitors. Specular reflected light techniques have proven to be very successful for both etch and growth rate monitoring of vacuum processes on unpatterned substrates [1]. However, applications of these methods in actual production have been very limited due to the diffraction problems inherent in monitoring patterned structures. Multi-wavelength ellipsometry utilizing a scalar theory has been used to monitor patterned structures [2]. However, the scalar approach becomes inaccurate as the investigated feature size goes into micron and sub-micron regime due to strong diffraction effects. Recent success in the use of vector diffraction theory has been demonstrated [3,4,5,6] for modeling the polarization dependent reflection problem for periodic gratings; however, the numerical algorithm (RCWA) used for this modeling is inherently limited to periodic structures. The real IC structures are much more complicated. Also, the computational load is extremely heavy for large feature size patterns due to the higher diffraction orders needed for satisfactory accuracy. The typical computation time for 2 μm structures is \sim min/wavelength on a 300 MHz Pentium II™ system.

In this paper, we will first present a careful study of the scalar approach for modeling large feature size structures. We will show that the scalar model, while

simple and fast (computation time negligible compared with RCWA), can yield accurate results for structures with feature size above 10 μm . We find that the limited transverse (spatial) coherence length of the probing light should be considered for the analysis of relatively large structures. We will finally present a hybrid vector-scalar approach for modeling limited-area periodic gratings surrounded by large feature size objects. The investigation of these structures is of practical importance. Since the collimated probing beams typically are \sim 5mm in diameter, the illuminated sample area is often a mixture of small-sized pattern regions and some larger structures or open area. This work will facilitate the use of specular diffraction-based measurements to *in situ* IC production.

EXPERIMENTS

Two 7×7×0.12 inch quartz masks were designed to fabricate experimental samples. The CD tolerance is $\pm 0.1\mu\text{m}$. The first mask is a set of gratings of 1×1 cm^2 area with nominally equal dark and clear lines. The line-width of these gratings ranges from 1 μm to 170 μm . This mask was used mainly for coherent scalar model tests. The second mask includes 3 broad-line gratings (line-width = 300, 400, and 500 μm respectively) and some more complicated non-periodic structures (one of them is a 0.7 cm wide 2 μm -line grating region sandwiched in between two 50 μm -line

grating regions). The area of each of these structures is $1.5 \times 1.5 \text{ cm}^2$. They were used for the test of partial coherence model (broad-line gratings) and hybrid model (non-periodic structures).

Two types of samples were prepared using these masks. The first is a set of photoresist patterns on silicon substrate. They were fabricated by patterning a $1 \text{ }\mu\text{m}$ thick Shipley 1808 photoresist layer on 6 inch $\langle 100 \rangle$ orientation single crystal Si wafer using contact photolithography. The second type is a set of SiO_2 patterns on silicon substrate. The photoresist patterns were first created on a 6 inch 5000 \AA thermal oxide on Si substrate wafer. The patterns were then etched using a Lam 9400 SE TCP high-density plasma etch tool. The etch conditions were: 10 mTorr pressure, an etch gas of C_2F_6 (120 sccm), a TCP power of 200 W, and a bias power of 40W. With this recipe, the etch rate of SiO_2 is approximately $428 \text{ \AA}/\text{min}$, and the oxide vs. silicon etching selectivity is about 1:1. Ideally, the etching was targeted to stop at the oxide/silicon interface. The endpoint was determined by controlling the etch time. Later analysis showed that there was an approximately 100 \AA of over etch. The photoresist was striped after the oxide was etched.

All spectroscopic ellipsometry (SE) measurements were performed on a Sopra GESP-5 spectroscopic ellipsometer/photometer system.

MODELING AND SIMULATION

By definition, ellipsometry measures the ratio of parallel (p) to perpendicular (s) reflection coefficients:

$$\rho = r_p / r_s = \tan \psi \cdot \exp(i\Delta) \quad (1)$$

where ψ and Δ are so called ellipsometric parameters, corresponding to the amplitude and phase respectively.

In Sopra GESP-5 system, the incident light is modulated by a rotating polarizer. The reflected light intensity at the detector can be expressed as [7]:

$$I = I_0(1 + \alpha \cos(2P) + \beta \sin(2P)) \quad (2)$$

where P is the angular position of polarizer. The Fourier coefficients α and β are measured and Ψ and Δ can be obtained from the following expression [7]:

$$\tan \Psi = \sqrt{\frac{1+\alpha}{1-\alpha}} \tan A, \quad \cos \Delta = \frac{\beta}{\sqrt{1-\alpha^2}} \quad (3)$$

where A is the angular position of the analyzer. It is evident from the above expression that (Ψ , Δ) and (α , β) are two equivalent representations of the

ellipsometry measurement. Ψ and Δ are associated with the field amplitudes, while α and β are more directly related to the light intensity.

Both the scalar and vector models and the combination of the two analyses have been utilized to fit the optical data of the patterned wafers.

For scalar analysis, we are using the modified Heimann approach [8] of the Lucent group [2]. The model assumes that the profile of a patterned wafer is divided into separate uniform thin film regions. The total reflection coefficient of the structure is calculated as a complex combination of the individual reflection coefficients from the different regions, i.e.,

$$\begin{aligned} R_p &= \sum_{i=1}^2 af_i \cdot \exp(2j \frac{2\pi}{\lambda} \delta_i \cos \phi_0) R_{p_i} \\ R_s &= \sum_{i=1}^2 af_i \cdot \exp(2j \frac{2\pi}{\lambda} \delta_i \cos \phi_0) R_{s_i} \end{aligned} \quad (4)$$

where R_{p_i} and R_{s_i} are the reflection coefficients in the p and s polarization corresponding to the i-th region. The nonnegative real number af_i is the area fraction of the i-th region. In Eq. (4), δ_i is the thickness of a layer of vacuum added on top of the i-th region to consider the phase lag due to the different heights of the stacks. The ratio of R_p and R_s determines Ψ and Δ by Eq. (1).

The validity of this linear combination of the reflection coefficients in the above analysis is based on the assumption that the detecting light is coherent across the pattern features (one period of the grating, for example), i.e., the transverse (spatial) coherence length is greater than the pattern feature size. The logical opposite of a fully coherent field is an incoherent field. This will happen as the coherence length of the light is much smaller than the pattern feature size and negligible. In this case the field amplitudes at any two points across the pattern feature are not correlated. Therefore, the intensities, instead of the amplitudes, of the field from different regions should be added. From Eq. (2), we can obtain the following rule of combination:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \frac{\sum_{i=1}^2 I_{0i} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}}{\sum_{i=1}^2 I_{0i}} \quad (5)$$

where I_{0i} is the average reflected light intensity from the i-th region, and is proportional to the area fraction and the reflectivity of the corresponding region.

If the coherence length of the light is smaller than, but by no means negligible compared to the pattern feature size, the field over the pattern is partially coherent, which is an intermediate state between a fully coherent and an incoherent field. We

approximate the result of this case by dividing the grating profile of one period into a coherence region (which is defined by the coherence length and centered at the stack boundary) and incoherence regions. Within the range of the coherence region, the field amplitudes are superimposed; and then the intensity (and therefore α and β) of different incoherence regions are added by Eq. (5).

For the micron and sub-micron periodic patterns, the scalar model is not valid due to the strong diffraction. RCWA, a numerical vector diffraction method, has been successfully used for the optical data simulation in this regime. We extended this application to some special non-periodic patterns, i.e. micron-sized periodic pattern area surrounded by bigger structures or open area. The analysis of the data from these structures was performed by combining the vector and scalar models. For the small sized periodic pattern region, RCWA is used to calculate the reflection; for the larger structures or open area, scalar model or Fresnel's formulae is applied; and finally the total reflection is obtained by adding the reflections from different regions using scalar model.

EXPERIMENTAL RESULTS

The ellipsometry measurement and simulations with the scalar model and hybrid approach will be presented and discussed in this section.

Scalar Model

For the scalar analysis of the photoresist gratings on Si, the structure was divided into two regions: the photoresist film on Si and bare Si. The measured SE data and the scalar model fit of a 10 μ m line grating are shown in FIGURE 1. The grating depth and area fraction were used as unknown parameters to fit the data. The fit was optimized by Levenberg-Marquardt regression. We found in the simulations that the heights and the positions of the peaks in the curves were sensitive to the area fraction and grating thickness respectively. The grating depth yielded by the regression is $1.1074 \pm 0.0006 \mu\text{m}$, which is very close to the SE measurement of the blanket photoresist region on the same sample.

The data were collected with the grating aligned both normal and parallel to the plane of incidence. As expected, we observed stronger scattering induced structures in SE data for the case normal to the grating direction. The influence of scattering is serious for the gratings with line-widths smaller than 30~40 μm for the normal-to-the-grating measurement, and the scalar model cannot give satisfactory simulations in this case.

For the parallel-to-the-grating alignment, however, the scalar model is successful for gratings as small as below 10 μm (line-width).

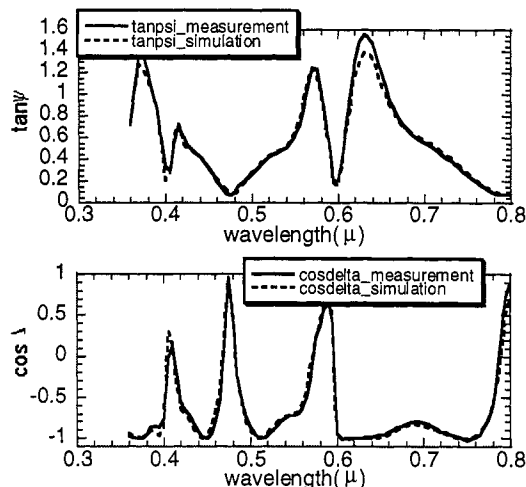


FIGURE 1. SE data (at 65°) and scalar simulation for the PR grating of a 10 μm line-width on Si substrate. The grating is parallel to the plane of incidence.

Besides the scattering, there are some geometrical limitations that may also affect the scalar simulation. For the normal-to-the-grating alignment, there is some area on the lower bare Si region that is shaded by the neighboring stack and cannot be illuminated by the light. This area does not contribute to the reflection and therefore the area fraction used as parameter of regression does not represent the real area fraction. In addition, the number of internal reflections (the beam bouncing between the interfaces) is finite due to the limited width of the grating line. This may affect the accuracy of the thin film reflection calculation, which assumes an infinite number of internal reflections. These geometrical effects, however, are not present in the parallel-to-the-grating measurement.

The shadowing effect can be detected in the angle-scanning ellipsometry measurement. The results are shown in FIGURE 2. The ellipsometry data was measured at $\lambda = 400\text{nm}$ with the angle of incidence varying from 7° to 75°. Since the shadow area changes with the angle of incidence, the measured curve can not be properly fitted by any fixed area fraction, which is shown in (a). If the shadowing effect is considered and a changing area fraction (as a function of angle of incidence) is used, the feature of the curve can then be captured, as seen in (b). For the parallel-to-the-grating measurement, this is not a problem, as seen in (c). It is noticed that the heights of the peaks of the two measured curves are different. This is due to the different area fraction in the two alignments as a result of the shadowing effect, which is unique to the normal-to-the-grating case.

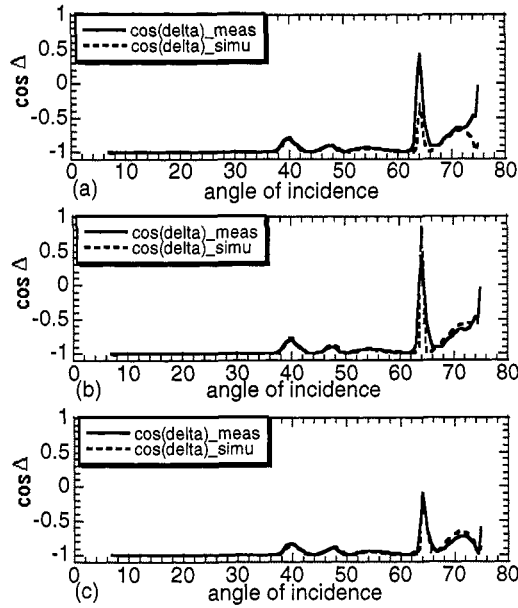


FIGURE 2. The angle-scanning ellipsometry measurement and simulation at $\lambda=400\text{nm}$ for $20\ \mu\text{m}$ line PR grating. (a) and (b): “perpendicular measurement” with normal scalar simulation in (a) and shadowing effect modified simulation in (b); (c): “parallel measurement” with normal scalar simulation.

We find that the scalar model can successfully fit the data from gratings as big as $120\ \mu\text{m}$ (line-width). For even bigger structures, however, the light is no longer coherent across the features. FIGURE 3 shows

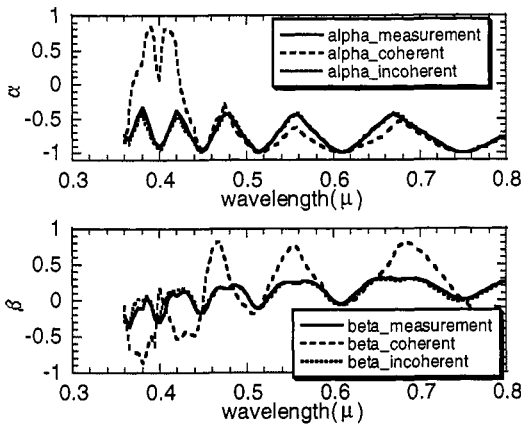


FIGURE 3. The coherent and incoherent fit of a 75° SE measurement from a Si wafer partially covered by photoresist. The area ratio of covered and bare area is 1:1.

the data measured from a silicon wafer partially covered by photoresist. The feature size of this structure is virtually the size of the probing beam spot. It is evident in the figure that the coherent model is not valid any more, while the incoherent simulation is in near-perfect agreement with the experimental data.

The oscillation of the incoherence curves is less strong due to the lack of interference. Since the intensities are added in the incoherence model, the (α, β) mode is more naturally chosen to represent the data instead of (Ψ, Δ) mode. The thickness of the photoresist obtained from the regression is $1.1372 \pm 0.0011\ \mu\text{m}$, which is close to the blanket photoresist measurement.

It is a natural inference from the above analysis that there must exist an intermediate state between the fully coherent and incoherent cases. We verified this by testing the broad-line ($300, 400, 500\ \mu\text{m}$) photoresist grating and SiO_2 grating samples. Neither the coherent model nor the incoherent model can successfully fit these experimental data. We developed a simple approximate partial coherence model, in which one repeat period of grating is divided into coherent and incoherent regions, defined by a transverse coherence length as an unknown parameter. This lateral coherence length (and therefore fraction) increases linearly with the light wavelength according to the theory of coherence optics. Typical data using the partial coherence model is shown in FIGURE 4. In addition to the area fraction and grating depth, the regression on this model also yielded the transverse coherence length. TABLE 1 lists the coherence lengths obtained from measuring different line-width gratings on both PR and SiO_2 samples at 60° and 75° . The transverse coherence length on the sample is \sim (transverse coherence length of the probe beam)/ $\cos(\theta)$, θ is the angle of incidence. With this geometric correction the approximate probe beam coherence length is $\sim 105\ \mu\text{m}$ for all cases. An experimental limitation of the broad-line grating measurement is that the higher diffraction modes are very close to the specular reflection and they may also enter the detector. If the intensity of those higher modes is strong and not negligible, the specular simulation will fail to match the measurement.

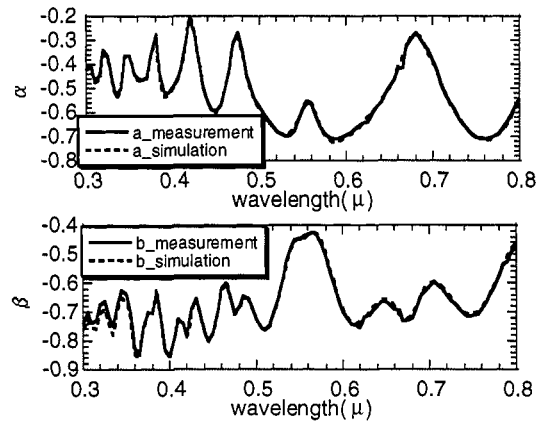


FIGURE 4. Partially coherent simulation of 60° SE data from $400\ \mu\text{m}$ line PR grating on Si.

TABLE 1. The Transverse Coherence Length (at $\lambda = 300\text{nm}$) Yielded by Fitting SE Data of Different Gratings

Angle Of Incidence	Grating Line-width	PR Gratings		SiO ₂ Gratings	
		Coherence Fraction	Coherence Length	Coherence Fraction	Coherence Length
60°	300 μm	0.3419	205.1 μm	0.3580	214.8 μm
	400 μm	0.2707	216.6 μm	0.2668	213.4 μm
	500 μm	0.2252	225.2 μm	0.2077	207.7 μm
75°	300 μm	0.6520	391.2 μm	Fitting not good due to the interference of higher diffraction modes entering the detector	
	400 μm	0.5220	417.6 μm		
	500 μm	0.4073	407.3 μm		

Hybrid Method

The typical feature size of IC samples is 1 micron or sub-micron. The scalar model can only be applied to structures above 10 μm . What makes it still important is, among others, that it can be combined with RCWA to solve less periodic structures. Figure 5 shows the hybrid simulation of the measurement of a limited area of 2 μm line gratings surrounded by a big structure (50 μm line). The upper figure is the best fit, and the lower figure shows how the data from the 2 μm grating middle region and the big-structure side region were incoherently combined into the total reflection data. This fit is not yet a completely inverse solution. The grating region and the big structure were first measured and fitted separately. The final fit was based on the best fit of each of these regions. In this way, we had a successful test of the physical model and it is of great promise that a complete regression process based on this model be achieved in the future.

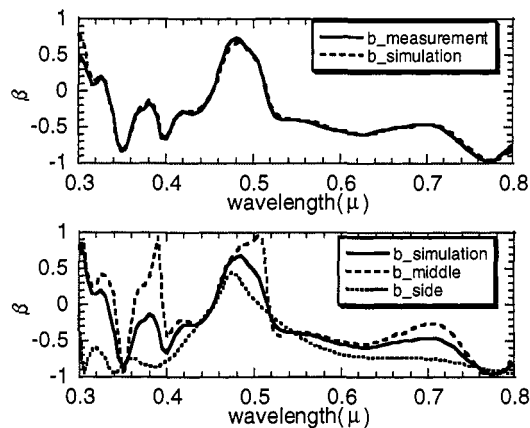


FIGURE 5. The best fit (upper) of SE data from non-periodic structure using hybrid approach, and the combination of SE data from different regions (lower)

CONCLUSIONS

The scalar model can be applied to the analysis of specular reflection data from patterns with feature

sizes above 10 μm . The coherence property of the light radiated from the source must be considered with respect to the feature size of the investigated patterns to ensure the proper modeling. A hybrid approach combining the vector and scalar models shows promise for analyzing less periodic patterns.

ACKNOWLEDGMENTS

This work was supported in part by the Semiconductor Research Corporation (contract 97-FC085), and AFOSR/DARPA MURI Center for Intelligent Electronics Manufacturing (AFOSR F49620-95-1-0524). The authors would also like to thank Mr. O. Srivannavit and Mr. H. Kim for assistance with sample fabrication.

REFERENCES

1. Vincent, T.L., Khargonekar, P.P., and Terry, F.L., Jr., *J. Electrochem. Soc.*, **144**, p2467 (1997)
2. Maynard, H.L., Laydi, A. N., Lee, J. T. C., *J. Vac. Sci. Tech.*, **B15**, p109 (1997)
3. Lee, M. E., Galarza, C., Kong, W., Sun, W., and Terry, F. L., Jr., International Conference on Characterization and Metrology for ULSI Technology, Gaithersburg, MD, March 23-27, 1998, AIP Conference Proceedings 449, pp. 331-5 (1998)
4. Kong, W., Huang, H. T., Lee, M. E., Galarza, C., Sun, W., and Terry, F. L., Jr., paper 19.2, SRC TECHCON, Las Vegas, Nevada, September 9-11, 1998.
5. Niu, X., Jakatdar, N., Bao, J., Spanos, C., and Yedur, S., Metrology, Inspection, and Process Control for Microlithography XIII, SPIE, March 1999.
6. Huang, H.-T., Kong, W., Kim, H., Sun, W., and Terry, F. L., Jr., Abs. 244, 195th Electrochemical Society Meeting, Seattle, Wa., May 2-5, 1999.
7. SOPRA, GESP-5 *User's Operating and Technical Manual*, June 1993
8. Heimann, P. A. and Schutz, R. J., *J. Electrochem. Soc.*, **131**, pp. 881-5 (1984).