

Effect of Bandwidth and Numerical Aperture in Optical Scatterometry

Thomas A. Germer and Heather J. Patrick

Optical Technology Division
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

ABSTRACT

We consider the effects of finite spectral bandwidth and numerical aperture in scatterometry measurements and discuss efficient integration methods based upon Gaussian quadrature in one dimension (for spectral bandwidth averaging) and two dimensions inside a circle (for numerical aperture averaging). Provided the wavelength is not near a Wood's anomaly for the grating, we find that the resulting methods converge very quickly to a level suitable for most measurement applications. In the vicinity of a Wood's anomaly, however, the methods provide rather poor behavior. We also describe a method that can be used to extract the effective spectral bandwidth and numerical aperture for a scatterometry tool. We find that accounting for spectral bandwidth and numerical aperture is necessary to obtain satisfactory results in scatterometry.

Keywords: optical critical dimension metrology, scatterometry

1. INTRODUCTION

Optical scatterometry, sometimes referred to as optical critical dimension (OCD) metrology, has become an attractive tool for dimensional metrology in the semiconductor industry, due in large part to its potential for providing in-line feedback information necessary for tight process control. In order to make the technique quantitative and traceable to the International System (SI) of units, a full uncertainty analysis is required, and the effects of various instrumental artifacts need to be understood, if not accounted for.¹ The simplest analysis of scatterometry data compares the measured data with the results of electromagnetic simulations, often from rigorous coupled wave (RCW) calculations, using the known wavelengths and incident angles.² However, in all real instruments, data for each recorded wavelength contains contributions from a span of wavelengths and incident directions.

In this paper, we investigate the effects of finite bandwidth and numerical aperture on the theoretical simulations. Since performing these averages necessitates an increase in computation time, and since scatterometry measurements are computationally intensive, we seek to find methods that are computationally efficient. We begin by describing a computationally efficient technique for performing wavelength and angle averages, which are parameterized by an order for which convergence can be tested. We then perform some simulations on a long period (period large compared to wavelength) and a short period (period small compared to wavelength) grating in order to determine how quickly convergence is obtained as the order is increased. We then suggest a method for determining the bandwidth and the numerical aperture that can be used for subsequent scatterometry simulations.

2. NUMERICAL METHODS

When taking into account spectral bandwidth and numerical aperture in modeling optical signatures, one generally needs to find an efficient method by which to average the optical signature over wavelength or incident direction. For a single dimension, such as for spectral bandwidth, we need to evaluate the mean,

$$\bar{f}(\lambda) = \int w(x)f(\lambda + x) dx, \quad (1)$$

where λ is the wavelength, $f(\lambda)$ is the function we are integrating, and $w(x)$ is a spectral bandwidth function. For this discussion, the function $w(x)$ is normalized as

$$\int w(x) dx = 1. \quad (2)$$

In order to program Eq. (1) in a computer, we must approximate it as the sum

$$\bar{f}(\lambda) \approx \sum_{i=1}^n w_i f(\lambda + x_i), \quad (3)$$

where the x_i and w_i are chosen appropriately. There are a number of ways that this choice can be made. The trapezoid and Simpson's rules,^{3,4} for example, are common when the function $f(\lambda + x)$ is relatively inexpensive to compute. However, in our application, each computation of $f(\lambda + x)$ requires a time consuming RCW simulation. Thus, we must choose the x_i and w_i judiciously.

Often, we can assume the function $f(\lambda + x)$ is relatively smooth. In this case, Gaussian quadrature offers an extremely efficient and accurate approximation for the integral.^{3,4} If $g_j(x)$ for $j = 1, \dots, n$ are approximation functions, then, for some set of x_i , we can solve the linear set of equations

$$\sum_{i=1}^n w_i g_j(x_i) = \int w(x) g_j(x) dx \quad (4)$$

for each j for the w_i , and then, the weighted average of any linear combination of the functions will be solved exactly. When the functions are chosen to be orthogonal polynomials, so that

$$\int w(x) p_i(x) p_j(x) dx = \begin{cases} i = j : & 1 \\ i \neq j : & 0 \end{cases}, \quad (5)$$

and if x_i are chosen to be the zeros of the p_{n+1} , then it can be shown that the approximation, Eq. (3), is not only exact for any linear combination of the first n polynomials, but for first $2n + 1$ polynomials. Therefore, for example, if the function $f(\lambda)$ is well approximated by a cubic polynomial, then only two evaluations of $f(\lambda)$ are required to obtain an accurate mean value $\bar{f}(\lambda)$. If $w(x)$ is rectangular [that is, $w(x)$ is constant within a specific domain, and zero otherwise], then the Legendre polynomials provide the appropriate set of polynomials. If the $w(x)$ is Gaussian, then the Hermite polynomials are the appropriate set of polynomials. For a general $w(x)$, we can generate polynomials by a recursive relationship.³

Table I gives a summary of the parameters x_i and w_i for rectangular, Gaussian, and triangular weighting functions for the first five and the ninth non-trivial orders. The values are tabulated in a manner that scale the x_i by the standard deviation σ_λ of $w(x)$. These values differ from other tabulations,⁴ which are often given for non-normalized weighting functions $w(x)$ with different standard deviations and normalizations different than Eq. (2).

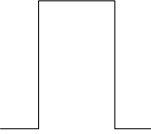

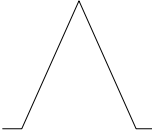
Integration over a circular aperture can also be performed using Gaussian quadrature. In this case, we make use of the Zernike polynomials,

$$Z_m^l(r, \phi) = \begin{cases} m \geq 0 : & R_m^l(r) \cos(m\phi) \\ m < 0 : & R_m^l(r) \sin(m\phi) \end{cases}, \quad (6)$$

defined on the interior of the unit circle. The radial function $R_m^l(r)$ is zero when $m - l$ is odd. We choose the coordinates r_i for sampling by choosing the zeros of $R_m^0(r)$ when m is even and $R_m^1(r)$ when m is odd. The ϕ_i are chosen evenly spaced, such that there are $2m$ sampled points azimuthally. Like Gaussian quadrature in one dimension, that in two dimensions yields exact results when the function is well approximated by Zernike polynomials having order less than twice the order used to determine the coordinates. Table II gives suitable values for r_i , ϕ_i , and w_i . The 1st, 2nd, and 4th orders can be found in Ref. [4, Sec. 25.4.61]. Again, the results are normalized to obtain a mean.

When we perform the averages over either wavelength or numerical aperture, we perform them on the entire Mueller matrix evaluated by RCW. In the case of numerical aperture averaging, we must ensure that we maintain the correct polarization basis when we consider directions outside of the central cone direction.⁵

Table 1. Parameters for averaging bandwidth.

	<u>Rectangular</u>		<u>Gaussian</u>		<u>Triangular</u>	
						
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
2	± 1.00000000	0.50000000	± 1.00000000	0.50000000	± 1.00000000	0.50000000
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
3	0.00000000	0.44444444	0.00000000	0.66666667	0.00000000	0.58333333
	± 1.34164079	0.27777778	± 1.73205081	0.16666667	± 1.54919334	0.20833333
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
4	± 0.58886444	0.32607258	± 0.74196378	0.45412415	± 0.64232931	0.40113370
	± 1.49153184	0.17392742	± 2.33441422	0.04587585	± 1.83938344	0.09886630
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
5	0.00000000	0.37393165	0.00000000	0.53333333	0.00000000	0.41773700
	± 0.93265620	0.31460345	± 1.35562618	0.22207592	± 1.10207529	0.23947324
	± 1.56954953	0.15573245	± 2.85697001	0.01125741	± 2.01211032	0.05165826
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
6	± 0.41330055	0.23395697	± 0.61670659	0.40882847	± 0.47584688	0.32302220
	± 1.14524825	0.18038079	± 1.88917588	0.08861575	± 1.41241206	0.14748123
	± 1.61508458	0.08566225	± 3.32425743	0.00255578	± 2.12061707	0.02949656
	⋮		⋮		⋮	
n	x_i/σ_λ	w_i	x_i/σ_λ	w_i	x_i/σ_λ	w_i
10	± 0.25785792	0.147762112	± 0.48493571	0.34464233	± 0.31493991	0.22945000
	± 0.75066284	0.13463336	± 1.46598909	0.13548370	± 0.95126258	0.15383706
	± 1.17677189	0.109543181	± 2.48432584	0.01911158	± 1.52795328	0.08123062
	± 1.49833370	0.074725675	± 3.58182348	0.00075807	± 1.99401518	0.03011326
	± 1.68685559	0.033335672	± 4.85946283	0.00000431	± 2.31027638	0.00536906

3. EXAMPLE PROFILES AND TOOL

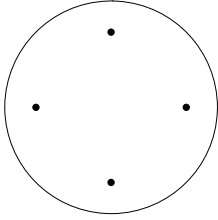
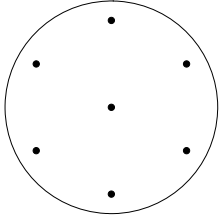
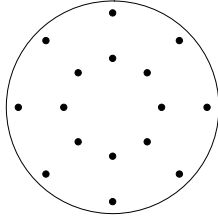
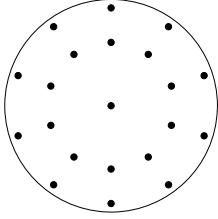
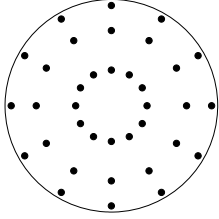
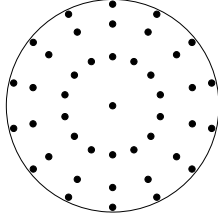
In this article, we use two virtual samples and a single virtual tool to illustrate the effects of finite bandwidth and finite numerical aperture. The target grating profiles have vertical-walled silicon lines of height 250 nm on top of a silicon substrate. The width and pitch of the large pitch sample are 100 nm and 800 nm, respectively, while the width and pitch of the small pitch sample are 45 nm and 90 nm, respectively.

The scatterometry tool is a rotating compensator spectroscopic ellipsometer operating from 250 nm to 1000 nm. The incident angle is 65°. The standard deviation of the bandwidth is assumed to be fixed at $\sigma_\lambda = 5$ nm across the spectrum. The numerical aperture, given by the sine of the cone halfangle, is assumed to be $NA = 0.05$. The incident light was linearly polarized at 45°. The results are presented in terms of the parameters N , S , and C , which are defined by the elements of the normalized reflected Stokes vector, $(1, N, S, C)$. The tool parameters were chosen to be similar to those of a physically existing tool, upon which measurements were performed and which is described later in the text.

4. SIMULATION RESULTS

Figure 1 shows the results for the averaging over bandwidth for the 800 nm and 90 nm gratings. The N , S , and C spectra (as a function of photon energy $E = hc/\lambda$, where h is Planck's constant and c is the speed of light) are shown in the lower graphs and are evaluated for the $n = 10$ approximation, assuming a rectangular distribution. The differences between the $n = 1$ approximation (single wavelength) and the $n = 10$ approximation

Table 2. Parameters for performing integrals on a circle.

$m = 2$	$m = 3$	$m = 4$
		
$i = 1, \dots, 4$ $r_{1-4} = 0.70710678$ $\phi_i = \pi i/2$ $w_{1-4} = 0.25000000$	$i = 1, \dots, 7$ $r_1 = 0.00000000$ $r_{2-7} = 0.81649658$ $\phi_i = \pi i/3$ $w_1 = 0.25000000$ $w_{2-7} = 0.12500000$	$i = 1, \dots, 16$ $r_{1-8} = 0.45970084$ $r_{9-16} = 0.88807383$ $\phi_i = \pi i/4$ $w_{1-16} = 0.06250000$
$m = 5$	$m = 6$	$m = 7$
		
$i = 1, \dots, 21$ $r_1 = 0.00000000$ $r_{2-11} = 0.59586158$ $r_{12-21} = 0.91921106$ $\phi_i = \pi i/5$ $w_1 = 0.11111111$ $w_{2-11} = 0.05124858$ $w_{12-21} = 0.03764031$	$i = 1, \dots, 36$ $r_{1-12} = 0.33571069$ $r_{13-24} = 0.70710678$ $r_{25-36} = 0.94196514$ $\phi_i = \pi i/6$ $w_{1-12} = 0.02314814$ $w_{13-24} = 0.03703704$ $w_{25-36} = 0.02314815$	$i = 1, \dots, 43$ $r_1 = 0.00000000$ $r_{2-15} = 0.46080423$ $r_{16-29} = 0.76846154$ $r_{30-43} = 0.95467902$ $\phi_i = \pi i/7$ $w_1 = 0.06250000$ $w_{2-15} = 0.02348888$ $w_{16-29} = 0.02772810$ $w_{30-43} = 0.01574730$

($\Delta N = N_1 - N_{10}$, $\Delta S = S_1 - S_{10}$, and $\Delta S = S_1 - S_{10}$) are given in the upper graphs of Fig. 1. The differences, while small in some parts of the spectra, approach and exceed 10% in a number of areas for the 800 nm grating. Because we are showing spectral data while considering spectral averaging, it comes as no surprise that the effects are largest in regions of the spectra where large curvature or structure exists. Some of these locations correspond to Wood's anomalies,⁷ apparent in the 800 nm pitch grating and marked in Fig. 1. When the period of the grating is Λ and the measurement geometry is conical, Wood's anomalies occur when

$$\frac{\Lambda}{\lambda} (\sin \theta_i + 1) \quad (7)$$

is an integer. For the 90 nm pitch grating, no Wood's anomalies are apparent or expected, but the effects of spectral averaging appear to exceed 10% in some wavelength regions. Since spectroscopic ellipsometry measurements typically have precisions on the 10^{-3} level, the effects of bandwidth could cause significant bias in their interpretation, if not accounted for.

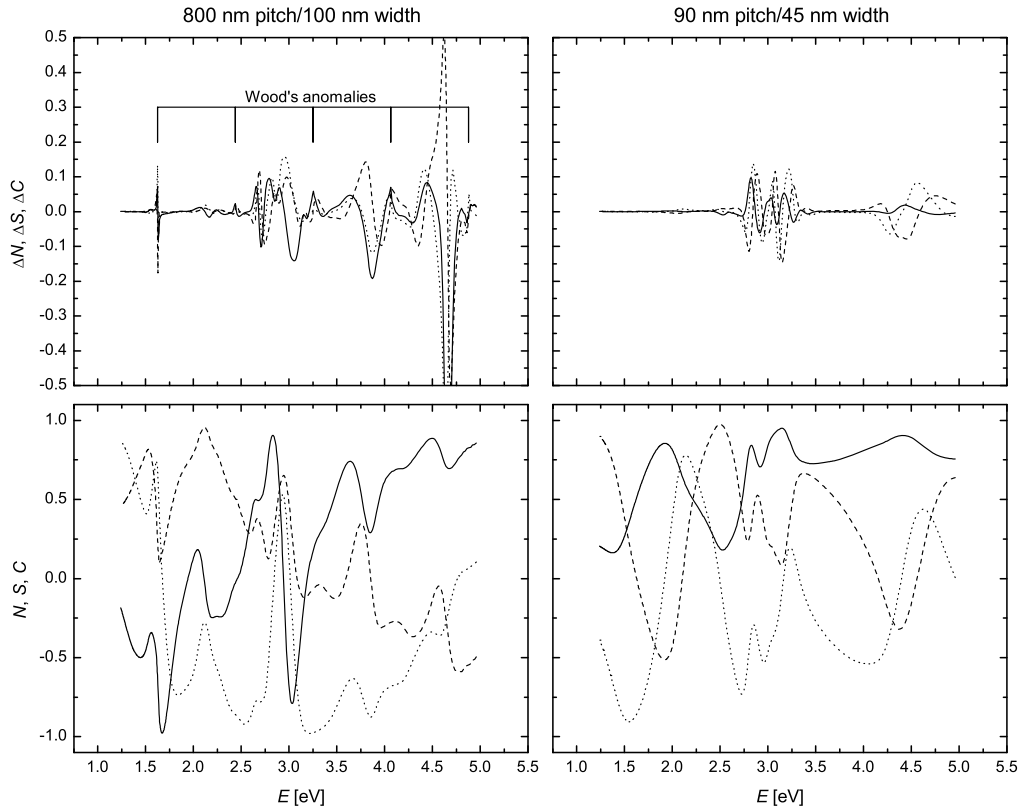


Figure 1. Spectra for the (left) 800 nm pitch, 100 nm width, and 250 nm high silicon binary grating and for the (right) 90 nm pitch, 45 nm width, and 250 nm high silicon binary grating: (bottom) the N , S , and C spectra calculated for a rectangular bandwidth with $\sigma_\lambda = 5$ nm and $n = 6$ rectangularly-weighted integration, and (top) the difference spectra, ΔN , ΔS , and ΔC , between that calculated by the $n = 6$ rectangularly-weighted integration and by evaluating at the central wavelength ($n = 1$). The curves represent (solid) N or ΔN , (dotted) S or ΔS , and (dashed) C or ΔC . Locations of Wood's anomalies are marked for the 800 nm pitch structure.

As the order n increases, we find that convergence occurs very quickly. We can assess the degree to which the average converges by evaluating the root mean square (rms) difference between the spectra calculated with the $n = 10$ approximation and those for $n \leq 6$. Figure 2 shows the convergence behavior for the $\sigma_\lambda = 5$ nm bandwidth averaging, assuming the rectangular band shape. The convergence behavior with the Gaussian and triangular bands are very similar, but not shown. For the 90 nm grating, convergence is extremely fast, dropping to less than 10^{-3} for $n = 3$, and perhaps even an acceptable level by $n = 2$. For the 800 nm grating, convergence appears to be much slower. There is an initial improvement, to better than 10^{-2} , but then the convergence appears to flatten out. Inspection of the difference spectra, however, shows that the residual convergence for large $n > 3$ is dominated by features very close to the Wood's anomalies. Outside of these features, convergence appears to be about as quick as that obtained for the 90 nm grating.

Figure 3 shows similar results for the averaging over the numerical aperture. The lower graphs show the spectra calculated using the $m = 7$ approximation, while the upper graphs show the differences calculated for $m = 1$ (single direction at the center of the aperture). For the numerical aperture chosen in this demonstration, the effects are smaller than that for bandwidth, but nonetheless are not negligible for a scatterometry measurement.

The convergence behavior for the numerical aperture average, shown in Fig. 4, also shows similar behavior compared to that for bandwidth. That is, convergence occurs very quickly for the 90 nm grating, such that at $m = 2$, the differences are about 10^{-6} , well below any expected measurement precision. Convergence for the

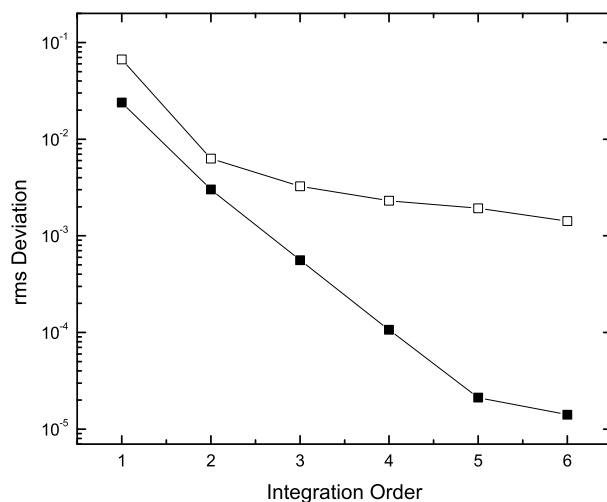


Figure 2. The rms difference between spectra calculated for $\sigma_\lambda = 5$ nm as a function of the integration order n and the $n = 10$ rectangularly-weighted integration. The points are for (open squares) the 800 nm pitch, 100 nm width, and 250 nm high silicon binary grating and for (solid squares) the 90 nm pitch, 45 nm width, and 250 nm high silicon binary grating.

800 nm grating in initially quick, but again levels out, dominated by the effects near the Wood's anomalies. Convergence away from the Wood's anomalies is very quick.

5. MEASUREMENT OF BANDWIDTH AND NUMERICAL APERTURE

In an effort to measure the bandwidth and the numerical aperture, we chose to perform a measurement of a nominally 1000 nm SiO_2 thick thermal oxide film on a Si substrate. Choosing as simple a sample as possible (no grating), but for which there is significant significant structure (thick film), gives us an ability to isolate the effects of finite bandwidth and numerical aperture from sample-specific artifacts. Our measurement was performed on a Woollam M-2000F focused spot spectroscopic ellipsometer,* which has a rotating incident compensator, a rotating detection polarizer, and a fixed angle of incidence of approximately 65° . The instrument measures data at 478 wavelengths, spanning the range from 247 nm to 1000 nm. The measurements were performed in a manner that obtained 11 of the 16 Mueller matrix elements. The parameters N , S , and C are determined from the m_{12} , m_{43} , and m_{33} elements of the normalized Mueller matrix reflectance, respectively. The measured N , S , and C spectra are shown in Fig. 5. To fit data to model parameters, we use a Levenberg-Marquardt search algorithm, which minimizes

$$\chi_r^2 = \frac{1}{3N_\lambda - M} \sum_{i=1}^{N_\lambda} \left[\frac{(N_i^{\text{meas}} - N_i^{\text{calc}})^2}{\sigma^2(N_i)} + \frac{(S_i^{\text{meas}} - S_i^{\text{calc}})^2}{\sigma^2(S_i)} + \frac{(C_i^{\text{meas}} - C_i^{\text{calc}})^2}{\sigma^2(C_i)} \right], \quad (8)$$

where X_i^{meas} is a measured value X , X_i^{calc} is a corresponding calculated value, and $\sigma(X_i)$ is the estimate of its standard deviation. The ellipsometer software supplied with the instrument provides an estimate of the standard deviation of the measured Mueller matrix elements $\sigma(X_i)$. The number of data points is N_λ and the number of fitting parameters is M .

*Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

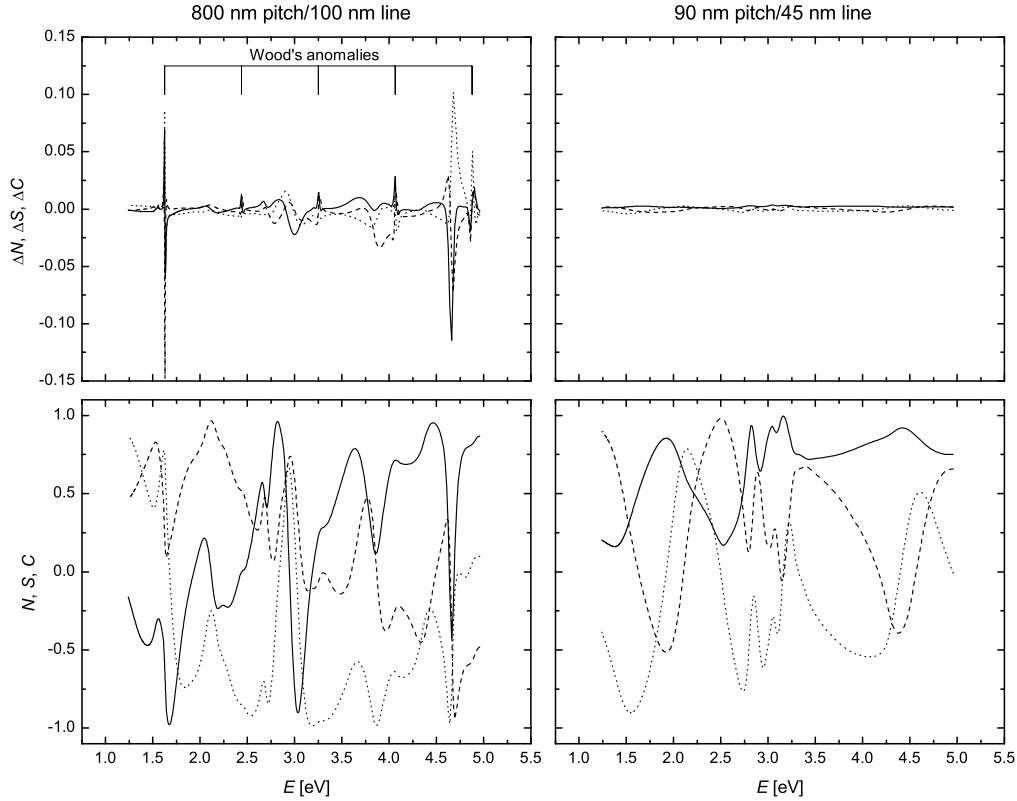


Figure 3. Spectra for the (left) 800 nm pitch, 100 nm width, and 250 nm high silicon binary grating and for the (right) 90 nm pitch, 45 nm width, and 250 nm high silicon binary grating: (bottom) the N , S , and C spectra calculated for $\text{NA} = 0.05$ with the $m = 7$ Gauss-Zernike integration, and (top) the difference spectra, ΔN , ΔS , and ΔC , between that calculated by the $m = 7$ Gauss-Zernike integration and by evaluating at the central direction ($m = 1$). The curves represent (solid) N or ΔN , (dotted) S or ΔS , and (dashed) C or ΔC . Locations of Wood's anomalies are marked for the 800 nm pitch structure.

The model we used to describe the film consisted of a SiO_2 layer of thickness t_{oxide} above an interface layer of thickness t_{inter} above the Si substrate. The optical properties (index and absorption coefficient) of each layer were taken from the literature.⁶ When we fit the data, letting just the thicknesses (t_{oxide} and t_{inter}) and the incident angle θ_i float, we obtain a minimum value $\chi_r^2 = 333$, $\text{rms} = 0.055$, and

$$\begin{aligned} t_{\text{oxide}} &= 1016.5 \text{ nm}, \\ t_{\text{inter}} &= 5.3 \text{ nm}, \\ \theta_i &= 64.64^\circ. \end{aligned}$$

This fit is shown in Fig. 5 as gray curves and is a relatively poor fit to the data, tending to overshoot the oscillations in the ellipsometric parameters, especially at the higher energies. We then let bandwidth σ_λ and the numerical aperture (NA) also be floating parameters, letting the bandwidth function be triangular with $n = 5$ and using the $m = 4$ approximation for the numerical aperture. When we include the bandwidth and numerical aperture, we find a significantly improved fit, with $\chi_r^2 = 21.7$, $\text{rms} = 0.0134$, and

$$\begin{aligned} t_{\text{oxide}} &= 1020.87 \text{ nm} \\ t_{\text{inter}} &= 4.2 \text{ nm} \\ \theta_i &= 65.17^\circ \end{aligned}$$

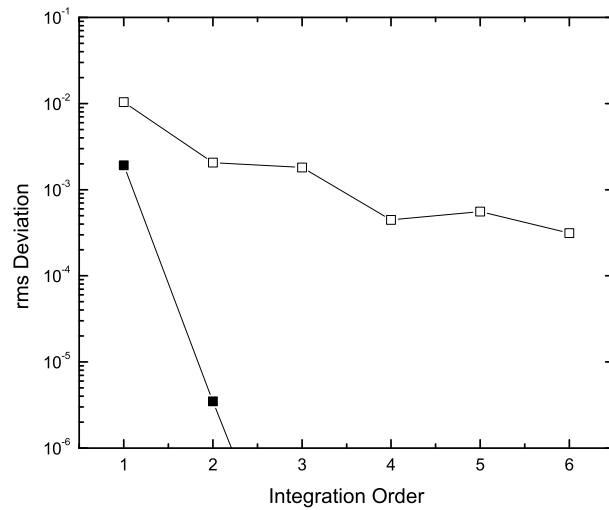


Figure 4. The rms difference between spectra calculated for NA = 0.05 as a function of the Gauss-Zernike integration order m and the $m = 7$ Gauss-Zernike integration. The points are for (open squares) the 800 nm pitch, 100 nm width, and 250 nm high silicon binary grating and for (solid squares) the 90 nm pitch, 45 nm width, and 250 nm high silicon binary grating.

$$\begin{aligned}\sigma_{\lambda} &= 1.6 \text{ nm} \\ \text{NA} &= 0.065\end{aligned}$$

This fit is shown as the black curves in Fig.5. The value for NA is close to that estimated by geometric measurement. The value for bandwidth is also very close to experimentally obtained results. In order to check the instrument's spectral spread function, we directed the light from a Hg-Ar source of narrow-band spectral lines into the instrument's spectrometer, and fit eight of the observed lines to Gaussian profiles. This measurement yielded an instrument bandwidth of $\sigma_{\lambda} = 1.7 \text{ nm}$.

Finally, the measurement using the Hg-Ar spectral line source revealed a small shift in the wavelength scale for the instrument of approximately 0.3 nm (the instrument read too high). If we let a wavelength offset $\Delta\lambda$ be an additional floating parameter, we obtained a slightly improved fit, with $\chi_r^2 = 19.5$, rms = 0.0122, and

$$\begin{aligned}t_{\text{oxide}} &= 1020.55 \text{ nm} \\ t_{\text{inter}} &= 3.9 \text{ nm} \\ \theta_i &= 65.17^\circ \\ \sigma_{\lambda} &= 1.5 \text{ nm} \\ \text{NA} &= 0.065 \\ \Delta\lambda &= -0.28 \text{ nm}\end{aligned}$$

where the negative $\Delta\lambda$ is consistent with the sign of the observed instrument shift. Thus, it appears that we can effectively measure the spectral bandwidth, the numerical aperture, and any wavelength shift of the instrument using the method described.

That the fit does not have χ_r^2 close to unity is not particularly bothersome, since the estimate of the standard deviations given by the instrument do not reflect all of the errors in the measurement. Furthermore, because of this, estimates of the errors in these parameters, obtained from the curvature of χ_r^2 , underestimated the true errors, which we do not report. As more parameters are added to those being fit, correlations between all the parameters increase.

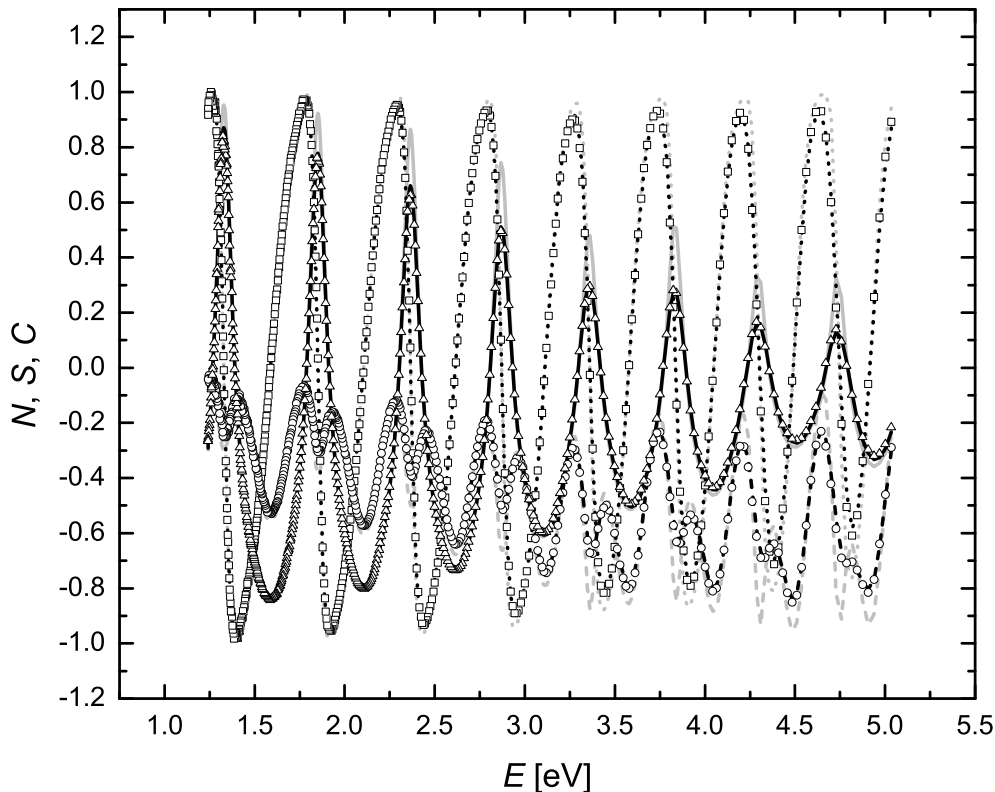


Figure 5. Ellipsometric spectra of a ≈ 1000 nm film on silicon. The measured data are (triangle) N , (squares) S , and (squares) C . The simulated data are (solid) N , (dotted) S , and (dashed) C . The fits ignoring finite bandwidth and numerical aperture are in gray, while the fits including the effects of bandwidth and numerical aperture are in black.

6. DISCUSSION

The results suggest that suitable averaging can be achieved with as few as three, or, possibly, two, samples in the spectral domain and as few as four in the directional domain. While averaging over wavelength scales linearly with the number of points sampled in wavelength, averaging over direction is slightly more complicated, because conical geometries (those in which the grating vector is in the plane of incidence) typically take about a factor of four longer to compute than non-conical ones. If the central direction of the incident numerical aperture is non-conical, then we can make use of symmetry to reduce the computation time. The quickest such sampling would be to choose two of the points in a non-conical geometry and one in the conical geometry, using symmetry to deduce the opposing conical geometry. Thus, the computational burden of directional averaging in a non-conical geometry can be as low as a factor of about six. This yields a total computational burden of 12 if two-point wavelength averaging is used and 16 if three-point wavelength averaging is used.

One can observe from Table 1 that to the lowest order shown ($n = 2$), the sampling points and their respective weights for rectangular, Gaussian, and triangular bandpasses are identical. This finding will, in fact, be true for any valid, symmetric $w(x)$. Since convergence of the average is very quick, one can conclude that knowing the precise spectral spread function $w(x)$ is much less important than knowing its standard deviation.

The convergence at wavelengths near the Wood's anomalies can be poor, because, while the Mueller matrix elements, and thus the parameters N , S , and C , are continuous across the boundaries, higher order derivatives are not. That is, there are cusps in the Mueller matrix elements at the Wood's anomalies. Thus, the assumption

that the function we are integrating can be approximated by a polynomial does not hold. For those wavelengths where part of the spectral bandwidth or numerical aperture overlaps a Wood's anomaly, we need an alternative method to integrate over bandwidth and numerical aperture. In the case of the bandwidth, we can split the bandwidth into parts which are above and below, generate the orthogonal polynomials for each side, and solve for the sampling points and their respective weights. For the numerical aperture integrals, this process may be much more complicated. In the end, it may be most appropriate and effective to employ a simpler approach, such as trapezoidal integration, which makes fewer assumptions about the functions being integrated, or to avoid such problems altogether by limiting the fit to regions of the spectrum where Wood's anomalies do not affect the spectra. Since this issue only occurs for large pitch gratings, it may be an issue for only a small number of applications.

7. CONCLUSIONS

We described methods that account for finite spectral bandwidth and numerical aperture in scatterometry measurements. Provided the wavelength is not near a Wood's anomaly, we find that the resulting methods converge very quickly to a level suitable for most measurement applications. We also described a method that can be used to extract the effective spectral bandwidth and numerical aperture for a scatterometry tool. We find that accounting for these effects are necessary to obtain satisfactory results in scatterometry.

REFERENCES

- [1] Germer, T.A., Patrick, H.J., Silver, R.M., and Bunday, B., "Developing an uncertainty analysis for optical scatterometry," in *Metrology, Inspection, and Process Control for Microlithography XXIII*, Allgair, J.A., and Raymond, C.J., Eds., Proc. SPIE **7272**, 72720T (2009).
- [2] Raymond, C.J., "Scatterometry for Semiconductor Metrology," in *Handbook of Silicon Semiconductor Metrology*, edited by Diebold, A.C., (Dekker, New York, 2001), pp. 477–514.
- [3] Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., *Numerical Recipes in C*, Second Edition, (Cambridge University Press, Cambridge, 1992).
- [4] Abramowitz, M., and Stegun, I.A., *Handbook of Mathematical Functions*, (U.S. Government Printing Office, Washington, D.C., 1972).
- [5] Germer, T.A., and Marx, E., "Simulations of Optical Microscope Images," in *Metrology, Inspection, and Process Control for Microlithography XX*, Archie, C.N., Ed., Proc. SPIE **6152**, 61520I (2006).
- [6] Herzinger, C.M., Johs, B., McGahan, W.A., Woollam, J.A., and Paulson, W., "Ellipsometric determination of optical constants for silicon and thermally grown silicon dioxide via a multi-sample, multi-wavelength, multi-angle investigation," *J. Appl. Phys.* **83**, 3323–3336 (1998).
- [7] Wood, R.W., "On a remarkable case of uneven distribution of light in a diffraction grating spectrum," *Philos. Mag.* **4**, 396–402 (1902).