Improved deep-etched multilayer grating reconstruction by considering etching anisotropy and abnormal errors in optical scatterometry

Jinlong Zhu,[1] Shiyuan Liu,[1,2,*] Hao Jiang,[1] Chuanwei Zhang,[1] and Xiuguo Chen[1]

[1]*State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

[2]*Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China*

**Corresponding author: shyliu@hust.edu.cn*

Received October 8, 2014; revised December 23, 2014; accepted December 24, 2014; posted January 7, 2015 (Doc. ID 224575); published February 4, 2015

Accurate, fast, and nondestructive reconstruction of the etched nanostructures is important for etching process control to achieve good fidelity, as well as high manufacturing yield. In this work, we have demonstrated the improved deep-etched multilayer grating reconstruction by simultaneously considering the model error of nonuniform side-wall angle (SWA) because of different etching anisotropies of various materials and by suppressing the abnormally distributed measurement errors using a robust statistics based method in optical scatterometry. More specifically, we introduce an additional parameter to perfect the profile model under measurement, and use a robust estimation procedure at the end of each iteration of the Gauss–Newton (GN) method to obtain the more accurate parameter departure vector. By applying the proposed methods, more accurate reconstructed results can be achieved.

OCIS codes: (120.0120) Instrumentation, measurement, and metrology; (120.2130) Ellipsometry and polarimetry; (290.3200) Inverse scattering; (050.1950) Diffraction gratings; (000.5490) Probability theory, stochastic processes, and statistics.

http://dx.doi.org/10.1364/OL.40.000471

As a primary feature pattern transfer method in integrated circuit (IC) manufacturing industry, etching directly affects the profile of the transferred film pattern [1,2]. To control the etching process for achievement of good fidelity, as well as high manufacturing yield, accurate, fast, and nondestructive reconstruction of the etched nanostructure is highly desirable. Among all the metrology techniques, the optics-based ones, such as specular x-ray reflectivity (SXR) [3], critical dimension small-angle x-ray scattering (CD-SAXS) [4] and ellipsometric scatterometry [5,6] are regarded as having the potential to meet these demands. Moreover, optical scatterometry is relatively ideal because of its low cost, high throughput, small hardware, and minimal sample damage [7].

Optical scatterometry is a model-based technique, and its success relies heavily on two main procedures involved, namely, the forward modeling and the inverse parameter extraction. For the forward modeling, no matter what kind of algorithm (such as the finite element method (FEM) [8] or the rigorous coupled-wave analysis (RCWA) [9–11]) is used, the parameterized profile model can be regarded as the key and prerequisite for the reconstruction of the measured nanostructure, since the model errors [12–14] between the profile model and the actual profile of the nanostructure directly affect the final parameter extraction accuracy. Besides the modeling errors, the inevitable measurement error is another source that may reduce the parameter extraction accuracy. If we know the statistical property of the measurement errors exactly, we can use the maximum likelihood estimation (MLE) to accurately extract the parameters. In particular, when the related inverse problem is discrete and the measurement errors are independent and normally distributed [15,16], the least square (LSQ) function based methods such as LSQ-based Gauss–Newton (GN) algorithm can be used to extract the parameters, and in this case, the least LSQ solution is the same as the MLE solution [17]. However, as emphasized in [18–20], the measurement errors in optical scatterometry are never normally distributed in consideration of the multiple system error sources, such as the finite bandwidth [21] and the collimation of the light source [22]. The abnormally distributed measurement errors will inevitably affect the parameter extraction accuracy when using the traditional LSQ-based methods. Hence, in order to improve the measurement accuracy, it is desirable to pick out the primary model error for the perfection of the profile model, as well as developing efficient algorithm to suppress the abnormally distributed measurement errors.

In this Letter, we will present the improved reconstruction of a deep-etched multilayer grating by perfecting its profile model via considering the model error of nonuniform side-wall angle (SWA) and by using the principle of robust statistics to suppress the abnormally distributed measurement errors in the measured signature.

First we present the model error of nonuniform SWA and its origin. The cross-section image of the deep-etched multilayer grating obtained by transmission electron microscopy (TEM) (TE20, TEM.FEI Co.) is presented in Fig. 1, in which we can find that the sample consists of Si, SiO_2, and nitride Si_3N_4 trapezoidal gratings from bottom to top. Figure 2(a) presents the etching for the Si_3N_4 layer, as well as its underneath thermal thin SiO_2 layer on a (100)-orientation single layer Si substrate in the first half of the etching procedure by a reactive ion etcher (KIYO, LAM Co.). In consideration of the ultrathin thickness of the SiO_2 layer (about 10 nm), its SWA

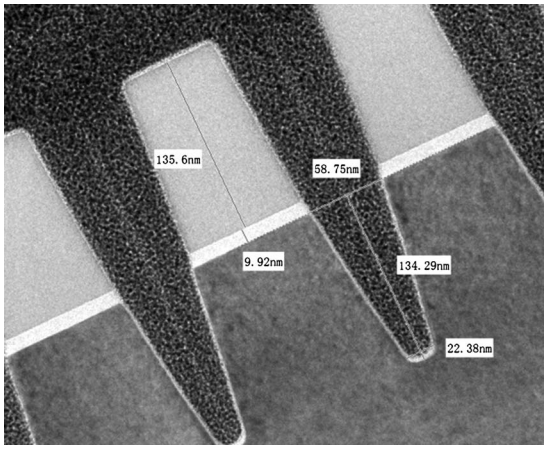

Fig. 1. Cross-section TEM image of the investigated deep-etched multilayer grating, which consists of Si, SiO_2, and Si_3N_4 trapezoidal gratings from bottom to top.

can be treated the same as that of the nitride layer, which is represented by SWA_1. With the pushing of the etching procedure, the Si substrate is etched with an SWA represented by SWA_2, as shown in Fig. 2(b). After removing the ArF photoresist, we can obtain the deep-etched multilayer grating. Traditionally, it is assumed that $SWA_1 = SWA_2$, since both of them are achieved in the same etching procedure, as shown in Fig. 2(c). However, we should point out that each material presents a specific anti-etching ability; hence, its etching anisotropy A [23], which is defined as

$$A = \frac{R_{\text{vertical}}}{R_{\text{lateral}}}, \tag{1}$$

is also specific. Here R_{vertical} and R_{lateral} are the vertical and lateral etch rates, respectively. It is, deservedly, to deduce that the anisotropy is equivalent to the tangent of the SWA of the trapezoidal profile, i.e., $A = \tan(\text{SWA})$. From the above discussion, it is valid to add an additional model parameter, namely, the SWA of Si grating represented by SWA_2, on the basis

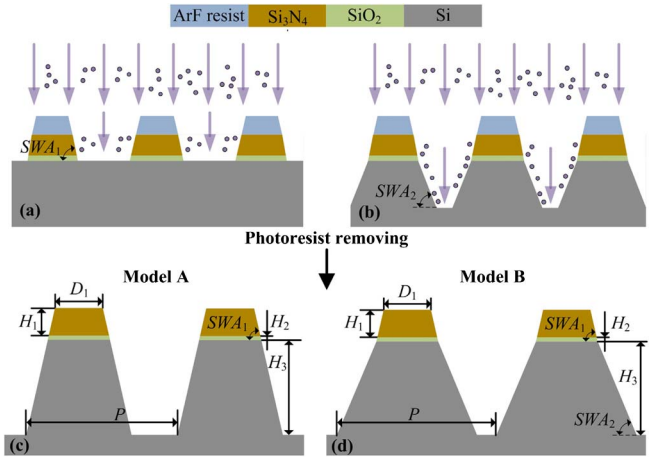

Fig. 2. Etching for (a) Si_3N_4/SiO_2 layers and (b) Si substrate in one etching procedure, and the two possible profile models of (c) uniform SWA and (d) nonuniform SWA after photoresist removing, respectively.

of the simple model presented in Fig. 2(c) (characterized by D_1, H_1, H_2, H_3, SWA_1 and pitch P), as shown in Fig. 2(d). For simplicity, the models in Figs. 2(c) and 2(d) are labeled as Model A and Model B, respectively.

Besides the model error, the abnormally distributed measurement errors will also reduce the measurement accuracy of the LSQ-based method, as emphasized above. To give a detailed explanation, we present the commonly used LSQ-based GN algorithm below. First, we formulate the inverse problem in optical scatterometry as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\Omega}\{[\mathbf{y} - \mathbf{f}(\mathbf{x})]^{\mathrm{T}}\mathbf{w}[\mathbf{y} - \mathbf{f}(\mathbf{x})]\}. \tag{2}$$

The formulation on the right side of Eq. (2) is the well-known LSQ function, which is used to quantify the difference between the measured signature $\mathbf{y}$ and the calculated signature $\mathbf{f}(\mathbf{x})$. Here, $\mathbf{x}$ represents the vector containing the parameters under measurement, $\hat{\mathbf{x}}$ is the solution of the inverse problem, $\mathbf{w}$ is the weight matrix, and Ω is the parameter domain. Then the GN algorithm, whose key point is the following iterative formulation:

$$\Delta\mathbf{x}^{(i)} = -[\mathbf{J}(\mathbf{x}^{(i)})^{\mathrm{T}}\mathbf{w}\mathbf{J}(\mathbf{x}^{(i)})]^{-1}\mathbf{J}(\mathbf{x}^{(i)})^{\mathrm{T}}\mathbf{w}\Delta\mathbf{y}^{(i)}, \tag{3}$$

can be used to converge iteratively to the solution $\hat{\mathbf{x}}$, provided an appropriate initialization has been chosen. Here the terms $\Delta\mathbf{x}^{(i)}$, $\mathbf{J}(\mathbf{x}^{(i)})$, and $\Delta\mathbf{y}^{(i)}$ represent the parameter departure vector, the Jacobian, and the residual column vector of the ith iterative step, respectively, and superscript T represents the transpose. Since $\mathbf{J}(\mathbf{x}^{(i)})$ and $\mathbf{w}$ are constants in the current ith iteration as shown in Eq. (3), the parameter departure vector $\Delta\mathbf{x}^{(i)}$ is unique determined by $\Delta\mathbf{y}^{(i)}$. In consideration of the fact that $\Delta\mathbf{y}^{(i)}$ contains the abnormal measurement errors, after by multiplying the magnification coefficient $-[\mathbf{J}(\mathbf{x}^{(i)})^{\mathrm{T}}\mathbf{w}\mathbf{J}(\mathbf{x}^{(i)})]^{-1}\mathbf{J}(\mathbf{x}^{(i)})^{\mathrm{T}}\mathbf{w}$, the measurement errors will inevitably deliver a bias to $\Delta\mathbf{x}^{(i)}$. To reduce the effect of abnormally distributed measurement errors on the extracted parameters, we propose to use a robust estimation procedure,

$$\begin{aligned}\Delta\hat{\mathbf{x}}^{*(i)} &= \arg\min_{\Delta\mathbf{x}^{*(i)}\in\Omega^{*}}\left\{\sum_{j=1}^{m}\rho[\Delta y_j^{(i)} + \mathbf{J}(\mathbf{x}^{(i)})_j\Delta\mathbf{x}^{*(i)}]\right\} \\ &= \arg\min_{\Delta\mathbf{x}^{*(i)}\in\Omega^{*}}\left\{\sum_{j=1}^{m}\rho[r_j^{(i)}]\right\},\end{aligned} \tag{4}$$

to take the place of Eq. (3) for obtaining the more accurate parameter departure vector $\Delta\hat{\mathbf{x}}^{*(i)}$. Here m, $\Delta y_j^{(i)}$, $\mathbf{J}(\mathbf{x}^{(i)})_j$, Ω^{*}, and ρ represent the number of elements in the measured signature, the jth element of $\Delta\mathbf{y}^{(i)}$, the jth row of $\mathbf{J}(\mathbf{x}^{(i)})$, the parameter domain of $\Delta\mathbf{x}^{*(i)}$, and a pre-defined robust objective function. The asterisk is used to distinguish the result obtained by using the robust estimation from that obtained by Eq. (3). Equation (4) is referred to as an M-estimate that is usually regarded as the most relevant class for model fitting [24], and the solution of Eq. (4) is called an M-estimator. According to the research achievement of Beaton and

Tukey [25], we can get the solution $\Delta\hat{\mathbf{x}}^{*(i)}$ by choosing a weight function $\omega(r)$, which is defined as

$$\omega\left(r_j^{(i)}\right) = \rho'\left(r_j^{(i)}\right)/r_j^{(i)}, \tag{5}$$

where $\rho'(r_j^{(i)})$ denotes the derivative with respect to $r_j^{(i)}$. Different weight functions present different abilities for the suppressing of abnormally distributed measurement errors. In this Letter, the weight function is chosen as

$$\omega\left(r_j^{(i)}\right) = \begin{cases} \dfrac{c_{\mathrm{A}}}{r_j^{(i)}} \sin\dfrac{r_j^{(i)}}{c_{\mathrm{A}}}, & |r_j^{(i)}| \le \pi c_{\mathrm{A}} \\ 0, & |r_j^{(i)}| > \pi c_{\mathrm{A}} \end{cases}, \tag{6}$$

where c_{A} is a scale that is adjusted in each iteration. In this work, c_{A} is automatically selected using the following equation:

$$c_{\mathrm{A}} = 1.4628 \times \mathrm{MAD}(\mathbf{r}^{(i)}), \tag{7}$$

where $\mathrm{MAD}(\mathbf{r}^{(i)}) = \mathrm{median}[|\mathbf{r}^{(i)} - \mathrm{median}|\mathbf{r}^{(i)}||]$ [26], and $\mathbf{r}^{(i)} = [r_1^{(i)}, r_2^{(i)}, \ldots, r_m^{(i)}]$. The use of Eq. (6) indicates that the weight of a specific data point is the function of its ith calculated residual $r_j^{(i)}$, and if the absolute value of $r_j^{(i)}$ is larger than πc_{A}, the weight is set as zero. After defining the weight function, the iteratively reweighted least squares (IRLSs) [27] can be applied to solve the weighted least-square problem presented in Eq. (4). By applying the above robust estimation, the more accurate extracted parameters can be achieved. In the following content we will experimentally demonstrate the improved grating reconstruction with consideration of the above profile modeling error and abnormally distributed measurement error. A house-built dual-rotating compensator Mueller matrix ellipsometer (DRC-MME) [28,29], together with an in-house optical modeling software package based on RCWA [11], was used to reconstruct the deep-etched multilayer grating as shown in Fig. 1. The Mueller matrices (normalized to m_{11}) were measured at 601 points over wavelengths ranging from 200 to 800 nm with the incidence angle θ fixed at 55°. We treated the azimuthal angle φ as a floating parameter since, in our DRC-MME prototype there is no positioning device on the rotating stage yet, and the zero angle cannot be guaranteed accurately by the manual mode. The pitch P is fixed at 154 nm for both Model A and Model B, and the profile parameters, as well as the azimuthal angle, φ are obtained using the traditional LSQ-based GN method. The initial values of geometrical parameters and azimuthal angle are chosen as the nominal ones and zero, respectively. The extracted parameters for Model A and Model B are presented in Table 1, in which we can find that the values of D_1, H_1, H_2, H_3, and SWA_1 given by Model B are closer to the TEM measured ones than that given by Model A.

Next we will demonstrate that the proposed robust estimation method can further improve the measurement accuracy by suppressing the abnormally distributed measurement errors in the measured signature on the basis of the corrected profile model (Model B). First, we will validate that some of the measurement errors in the measured Mueller matrix of the deep-etched multilayer grating are abnormally distributed. We calculate the weighted fitting differences of the measured and best fitted Mueller matrix elements, as shown in Fig. 3. Note that, in practice, to accurately obtain the measurement errors is impossible; thus the fitting differences between the measured Mueller matrix and the best fitted one are a compromise, but this compromise is still able to indirectly reflect the actual statistics property of the measurement errors in the measured signature. The best fitted Mueller matrix elements are calculated from the extracted parameters of Model B by the LSQ-based GN method. As seen in Fig. 3, many data points are obviously out of the range of −3 to 3, and a part of which are even larger than 300. Moreover, we could also notice that most of the large abnormally distributed data points locate in the Mueller matrix elements of m_{12}, m_{13}, m_{14}, m_{21}, m_{33}, m_{34}, m_{43}, and m_{44}.

Table 1. Comparison of Parameters Extracted from MME and TEM Measurement

	LSQ-Based GN			
Parameter	Model A	Model B	Robust Estimation	TEM
D_1 (nm)	65.52	74.63	75.32	75.01
H_1 (nm)	127.21	132.45	135.51	135.6
H_2 (nm)	26.44	13.44	9.12	9.92
H_3 (nm)	123.32	130.67	132.79	134.29
SWA_1 (°)	83.98	86.96	87.45	87.49
SWA_2 (°)	—	81.92	82.74	82.29
φ (°)	−1.89	−1.99	−2.64	—

We used the proposed robust estimation method to extract the parameters of Model B for the deep-etched multilayer grating, and presented the fitted results in Table 1. As expected, nearly all of the fitted geometrical parameters obtained by robust estimation are closer to the TEM measured ones than that obtained by the LSQ-based

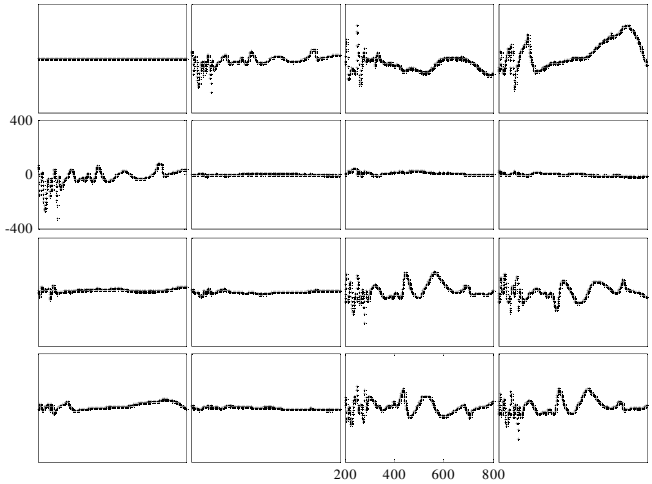

Fig. 3. Weighted fitting differences of Mueller matrix elements. The Mueller matrix elements are normalized to m_{11}. The horizontal axes, varying from 200 to 800 nm with an increment of 1 nm, denote the wavelengths; the vertical axes, varying from −400 to 400, denote the associate weighted fitting differences.

GN method. More specifically, we can find that the parameters H_1, H_2, H_3, and SWA_1 obtained by robust regression are 3.06, 4.32, 2.12 nm, and 0.49° closer to the TEM measured values than that by the LSQ-based GN method for Model B, respectively. The above results have demonstrated that the measurement errors in the measured signature cannot be treated as normally distributed, and it is necessary to use the principle of robust statistics to suppress the abnormally distributed measured errors.

In summary, we have demonstrated that, by considering the model error of nonuniform SWA, and by using the principle of robust statistics, the profile reconstruction of the deep-etched multilayer grating can be greatly improved. More specifically, the model error is dealt with by introducing an additional parameter SWA_2 to perfect the profile model, and the effect of abnormally distributed measured errors on the extracted parameters is suppressed by using a robust estimation procedure to replace the traditional GN iteration procedure.

This work was funded by the National Natural Science Foundation of China (Grant Nos. 51475191 and 51405172), the National Instrument Development Specific Project of China (Grant No. 2011YQ160002), and the Program for Changjiang Scholars and Innovative Research Team in University of China (Grant No. IRT13017). The authors would like to thank the facility support of the Center for Nanoscale Characterization and Devices, Wuhan National Laboratory for Optoelectronics (WNLO) (Wuhan, China).

References

1. G. Y. Zhang, P. F. Qi, and X. R. Wang, Science **314**, 974 (2006).
2. T. Lill and O. Joubert, Science **319**, 1050 (2008).
3. H.-J. Lee, C. L. Soles, H. W. Ro, R. L. Jones, E. K. Lin, W. Wu, and D. R. Hines, Appl. Phys. Lett. **87**, 263111 (2005).
4. R. L. Jones, T. Hu, C. L. Soles, E. K. Lin, R. M. Reano, S. W. Pang, and D. M. Casa, Nano Lett. **6**, 1723 (2006).
5. D. Fuard, C. Perret, V. Farys, C. Gourgon, and P. Schiavone, J. Vac. Sci. Technol. B **23**, 3069 (2005).
6. H. J. Patrick, T. A. Germer, Y. Ding, H. W. Ro, L. J. Richter, and C. L. Soles, Appl. Phys. Lett. **93**, 233105 (2008).
7. H. Huang and F. Terry, Jr., Thin Solid Films **455–456**, 828 (2004).
8. Y. Ohkawa, Y. Tsuji, and M. Koshiba, J. Opt. Soc. Am. A **13**, 1006 (1996).
9. M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, J. Opt. Soc. Am. A **12**, 1068 (1995).
10. L. Li, J. Opt. Soc. Am. A **13**, 1870 (1996).
11. S. Y. Liu, Y. Ma, X. G. Chen, and C. W. Zhang, Opt. Eng. **51**, 081504 (2012).
12. T. H. Ghong, S.-H. Han, J.-M. Chung, J. S. Byun, T. J. Kim, D. E. Aspnes, Y. D. Kim, I. H. Park, and Y.-W. Kim, Opt. Lett. **35**, 733 (2010).
13. A. Kato and F. Scholze, Appl. Opt. **49**, 6102 (2010).
14. J. L. Zhu, S. Y. Liu, C. W. Zhang, X. G. Chen, and Z. Q. Dong, J. Micro/Nanolith. MEMS MOEMS **12**, 013004 (2013).
15. M. A. Henn, H. Gross, F. Scholze, M. Wurm, C. Elster, and M. Bär, Opt. Express **20**, 12771 (2012).
16. M. A. Henn, S. Heidenreich, H. Gross, A. Rathsfeld, F. Scholze, and M. Bär, Opt. Lett. **37**, 5229 (2012).
17. R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems* (Academic, 2005).
18. R. C. Geary, Biometrika **34**, 209 (1947).
19. E. J. Schlossmacher, J. Am. Stat. Assoc. **68**, 857 (1973).
20. J. L. Zhu, S. Y. Liu, X. G. Chen, C. W. Zhang, and H. Jiang, Opt. Express **22**, 22031 (2014).
21. T. A. Germer and H. J. Patrick, Proc. SPIE **7638**, 76381F (2010).
22. X. G. Chen, C. W. Zhang, and S. Y. Liu, Appl. Phys. Lett. **103**, 151605 (2013).
23. C.-L. Cheng, H.-C. Chang, J.-C. Lin, and K.-J. Song, Phys. Rev. Lett. **78**, 3713 (1997).
24. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes–The Art of Science Computing* (Cambridge University, 2007).
25. A. E. Beaton and J. W. Tukey, Technometrics **16**, 147 (1974).
26. A. Douiri, M. Schweiger, J. Riley, and S. Arridge, Opt. Lett. **30**, 2439 (2005).
27. J. A. Scales, A. Gersztenkorn, and S. Treitel, J. Comput. Phys. **75**, 314 (1988).
28. S. Y. Liu, X. G. Chen, and C. W. Zhang, ECS Trans. **60**, 237 (2014).
29. X. G. Chen, S. Y. Liu, C. W. Zhang, H. Jiang, Z. C. Ma, T. Y. Sun, and Z. M. Xu, Opt. Express **22**, 15165 (2014).