Measurement configuration optimization for accurate grating reconstruction by Mueller matrix polarimetry

Xiuguo Chen
Huazhong University of Science and Technology
Wuhan National Laboratory for Optoelectronics
Wuhan 430074, China

Shiyuan Liu
Huazhong University of Science and Technology
Wuhan National Laboratory for Optoelectronics
Wuhan 430074, China
and
Huazhong University of Science and Technology
State Key Laboratory of Digital Manufacturing Equipment and Technology
Wuhan 430074, China
E-mail: shyliu@mail.hust.edu.cn

Chuanwei Zhang
Huazhong University of Science and Technology
State Key Laboratory of Digital Manufacturing Equipment and Technology
Wuhan 430074, China

Hao Jiang
University of Texas at Arlington
Department of Mechanical and Aerospace Engineering
Arlington, Texas 76019

Abstract. Due to the rich information provided by the Mueller matrices when the most general conical diffraction configuration is considered, the Mueller matrix polarimetry has demonstrated a great potential in semiconductor manufacturing. As the configurations of the incidence and azimuthal angles have different influences on the measurement accuracy, it is necessary to select an optimal one among the multitude of possible options. We introduce the norm of a configuration error propagating matrix to assess the measurement accuracy for different measurement configurations. The optimal configuration for a Si grating sample was achieved by minimizing the norm of the configuration error propagating matrix. Experimental results show the agreement between the theoretically predicted optimal configuration and the experimental exhibited one obtained by using a dual-rotating compensator Mueller matrix polarimeter and thus demonstrated the validity of the proposed optimization method. [DOI: 10.1117/1.JMM.12.3.033013]

Subject terms: Mueller matrix polarimetry; inverse problem; grating reconstruction; measurement accuracy; measurement configuration; optimization.

Paper 13023P received Mar. 13, 2013; revised manuscript received Jun. 6, 2013; accepted for publication Jul. 12, 2013; published online Aug. 19, 2013.

1 Introduction

As a nonimaging optical measurement technique, the Mueller matrix polarimetry (MMP) has been successfully introduced for critical dimension and overlay metrology recently.[1–3] Due to the rich information provided by the Mueller matrices when the grating lines are no longer perpendicular to the incidence plane but are positioned at different azimuthal angles, MMP has demonstrated a great potential in semiconductor manufacturing. Theoretically, we can obtain all the Mueller matrices by continuously varying the wavelength and the incidence and azimuthal angles to achieve high measurement precision and accuracy. However, in order to improve the efficiency of data acquisition and analysis, it is the common practice to choose a subset of the three measurement conditions from the available ranges. The combination of the selected wavelengths and incidence and azimuthal angles is defined as the measurement configuration. For example, we can fix the incidence and azimuthal angles in proper values while continuously varying the wavelengths in an available range. Similarly, we can also fix the wavelength and azimuthal angles in proper values while continuously varying the incidence angles in an available range. In general, a multitude of possible measurement configurations can be chosen by making different combinations of the three measurement conditions. It is worth while to point out that there are great discrepancies in the final measurement precision and accuracy in different configurations. Therefore, there is a need for MMP to choose an optimal one from the multitude of possible measurement configurations.

In the past decades, several approaches have been proposed to optimize the measurement configuration for conventional ellipsometric scatterometry. Logofatu proposed a sensitivity analysis for fitting method by defining the sensitivity as the estimated precision of the structural parameters to optimize the measurement configuration for angle-resolved rotating-analyzer and angle-resolved phase-modulation scatterometers.[4,5] Littau et al. investigated several optimal diffraction signature scan path selection techniques to improve scatterometry precision.[6] Gross et al. proposed an algorithm to determine the optimal measurement data set by minimizing the condition number of the Jacobian matrix, whose elements are defined as the partial derivatives of the diffraction signature with respect to the structural parameters.[7] Vagos et al. developed an uncertainty and sensitivity analysis package that can be used to guide the model and azimuthal angle optimization processes.[8] A recent study on spectroscopic MMP reported that the Mueller matrices obtained in some measurement configurations may help decorrelate the fitting structural parameters.[9] They further proposed to choose the measurement configuration with small parameter correlations and small estimated precision of the structural parameters.[10] It is also interesting to note from the previous works[4,5,10]

that the accuracy of the extracted structural parameters varies greatly in different measurement configurations. Typical minimization of the parameter correlations and estimated precision may achieve an optimal configuration with a higher measurement precision, but cannot guarantee the final measurement accuracy.

In this article, we propose a measurement configuration optimization method for spectroscopic MMP to find an optimal combination of the fixed incidence and azimuthal angles, with which more accurate measurement can be achieved. The proposed method is inspired by the theoretical analysis of error propagation in grating reconstruction by MMP. Error analysis for conventional ellipsometric scatterometry has been studied in recent years. Al-Assaad et al. investigated the propagation of different types of errors from the scatterometric data to the extracted structural parameters.[11] Germer et al. developed a scatterometry sensitivity analysis program, named OCDSense, to describe the propagation of measurement noise and to estimate systematic effects in measurement.[12,13] Motivated by these analyses, we first derive a generalized first order error propagating formula, which reveals the mechanism of error propagation in grating reconstruction by MMP. Based on this formula, a systematic error propagating formula is further derived, which relates the systematic errors in the extracted structural parameters with those error sources such as the configuration error as well as the systematic error in the measured Mueller matrices. Here, the configuration error is defined as the biases in the incidence and azimuthal angles and is typically induced by the mechanical positioning errors and the finite numerical apertures of the focusing lens in the measurement system.[14] The configuration error will induce systematic errors in the extracted structural parameters and thus degrade the final measurement accuracy. We then introduce the norm of the configuration error propagating matrix, which represents the maximum gain factor in the propagation of the configuration error, to assess the influences of the configuration error on the measurement accuracy for different measurement configurations. We further optimize the measurement configuration for MMP by minimizing the norm of the configuration error propagating matrix.

The reminder of this article is organized as follows. Section 2 first briefly introduces the inverse problem in grating reconstruction by MMP and then presents the error propagation in grating reconstruction by MMP and the measurement configuration optimization method in detail. Section 3 introduces the experimental setup, including the setup of a dual-rotating compensator Mueller matrix polarimeter as well as the geometric structure of the investigated Si grating sample. Section 4 provides the simulation and experimental results to examine the validity of the proposed optimization method. Finally, we draw some conclusions in Sec. 5.

2 Methods

2.1 *Inverse Problem in Grating Reconstruction by MMP*

Without loss of generality, we denote the structural parameters under measurement as an M-dimensional vector $\mathbf{x} = [x_1, x_2, \ldots, x_M]^{\mathrm{T}}$, where the superscript "T" represents the transpose. The vector $\mathbf{a} = [\theta, \varphi]^{\mathrm{T}}$ that consists of the incidence angle θ and azimuthal angle ϕ denotes the measurement configuration. The χ^2 function is usually applied to estimate the fitting errors between the measured and calculated Mueller matrix elements $m_{ij,k}^{\mathrm{meas}}$ and $m_{ij,k}^{\mathrm{calc}}(\mathbf{x}, \mathbf{a})$, which is defined as

$$\chi^2 = \sum_{k=1}^{N_\lambda} \sum_{i,j} \left[\frac{m_{ij,k}^{\mathrm{meas}} - m_{ij,k}^{\mathrm{calc}}(\mathbf{x}, \mathbf{a})}{\sigma(m_{ij,k})} \right]^2, \tag{1}$$

where k denotes the spectral point from the total number N_λ, and indices i and j show all the Mueller matrix elements except m_{11}. $\sigma(m_{ij,k})$ is the standard deviation associated with $m_{ij,k}$. For clarity, the measured Mueller matrix element $m_{ij,k}^{\mathrm{meas}}$ in Eq. (1) is marked as y_l with the three indices i, j, and k lumped into a single index l. The calculated Mueller matrix element $m_{ij,k}^{\mathrm{calc}}(\mathbf{x}, \mathbf{a})$ is correspondingly marked as $f_l(\mathbf{x}, \mathbf{a})$. Thus, Eq. (1) can be simply rewritten as

$$\chi^2 = \sum_{l=1}^{N} w_l [y_l - f_l(\mathbf{x}, \mathbf{a})]^2 = [\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a})]^T \mathbf{W} [\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a})], \tag{2}$$

where w_l is the weighting factor and is given by $w_l = 1/\sigma^2(y_l)$ and $N = 15N_\lambda$. $\mathbf{W}$ is an $N \times N$ diagonal matrix with diagonal elements w_l. The inverse problem in grating reconstruction is typically formulated as a least square regression problem such that

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \Omega} \{ [\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a}^*)]^{\mathrm{T}} \mathbf{W} [\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a}^*)] \}, \tag{3}$$

where $\hat{\mathbf{x}}$ is the solution of the inverse problem that contains the extracted structural parameters, and Ω is the associated parameter domain. $\mathbf{a}^*$ denotes the given value of vector $\mathbf{a}$ in the parameter extraction.

2.2 *Error Propagation in Grating Reconstruction by MMP*

We assume that the function $\mathbf{f}(\mathbf{x}, \mathbf{a})$ is sufficiently smooth and can be expanded in a Taylor series which, truncated to the first order, leads to a linear model at $(\hat{\mathbf{x}}, \mathbf{a}^*)$

$$\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) + \mathbf{J}_{\mathrm{x}} \cdot (\mathbf{x} - \hat{\mathbf{x}}) + \mathbf{J}_{\mathrm{a}} \cdot (\mathbf{a} - \mathbf{a}^*), \tag{4}$$

where $\mathbf{J}_{\mathrm{x}}$ and $\mathbf{J}_{\mathrm{a}}$ are the $N \times M$ and $N \times 2$ Jacobian matrices with respect to $\mathbf{x}$ and $\mathbf{a}$, respectively, whose elements are given by

$$[\mathbf{J}_{\mathrm{x}}]_{ij} = \left. \frac{\partial f_i(\mathbf{x}, \mathbf{a})}{\partial x_j} \right|_{\mathrm{x}=\hat{\mathrm{x}}, \mathrm{a}=\mathrm{a}^*}, \tag{5a}$$

$$[\mathbf{J}_{\mathrm{a}}]_{ij} = \left. \frac{\partial f_i(\mathbf{x}, \mathbf{a})}{\partial a_j} \right|_{\mathrm{x}=\hat{\mathrm{x}}, \mathrm{a}=\mathrm{a}^*}. \tag{5b}$$

Substitution of $\mathbf{x} = \mathbf{x}_0$ and $\mathbf{a} = \mathbf{a}_0$ into Eq. (4) gives

$$\mathbf{f}(\mathbf{x}_0, \mathbf{a}_0) = \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) + \mathbf{J}_{\mathrm{x}} \Delta\mathbf{x} + \mathbf{J}_{\mathrm{a}} \Delta\mathbf{a}, \tag{6}$$

where $\mathbf{x}_0$ and $\mathbf{a}_0$ are the true values of $\mathbf{x}$ and $\mathbf{a}$, respectively. $\Delta\mathbf{x}$ and $\Delta\mathbf{a}$ represent the error propagated into $\hat{\mathbf{x}}$ and the

configuration error in the parameter extraction, respectively, and are given by $\Delta\mathbf{x} = \mathbf{x}_0 - \hat{\mathbf{x}}$ and $\Delta\mathbf{a} = \mathbf{a}_0 - \mathbf{a}^*$. The measurement vector $\mathbf{y}$ will be the sum of the true signature $\mathbf{f}(\mathbf{x}_0, \mathbf{a}_0)$ of the grating sample and a deterministic offset vector $\boldsymbol{\mu}_{\Delta y}$ and a random noise vector $\boldsymbol{\varepsilon}_{\Delta y}$, i.e.,

$$\mathbf{y} = \mathbf{f}(\mathbf{x}_0, \mathbf{a}_0) + \boldsymbol{\mu}_{\Delta y} + \boldsymbol{\varepsilon}_{\Delta y}. \tag{7}$$

The vectors $\boldsymbol{\mu}_{\Delta y}$ and $\boldsymbol{\varepsilon}_{\Delta y}$ represent the intrinsic systematic and random errors in $\mathbf{y}$ induced by the measurement system.

Inserting Eqs. (6) and (7) into Eq. (2), we have

$$\begin{aligned}\chi^2_{\min} &= [\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*)]^{\mathrm{T}} \mathbf{W} [\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*)] \\ &= [\mathrm{J_x}\Delta\mathbf{x} + \mathrm{J_a}\Delta\mathbf{a} + \boldsymbol{\mu}_{\Delta y} + \boldsymbol{\varepsilon}_{\Delta y}]^{\mathrm{T}} \mathbf{W} [\mathbf{J_x}\Delta\mathbf{x} + \mathbf{J_a}\Delta\mathbf{a} + \boldsymbol{\mu}_{\Delta y} + \boldsymbol{\varepsilon}_{\Delta y}].\end{aligned} \tag{8}$$

By taking the derivative of both sides of Eq. (8) with respect to $\mathbf{x}$, we derive that

$$\tilde{\mathbf{J}}_{\mathrm{x}}\Delta\mathbf{x} + \tilde{\mathbf{J}}_{\mathrm{a}}\Delta\mathbf{a} + \tilde{\boldsymbol{\mu}}_{\Delta y} + \tilde{\boldsymbol{\varepsilon}}_{\Delta y} = 0, \tag{9}$$

where $\tilde{\mathbf{J}}_{\mathrm{x}} = \mathbf{W}^{1/2}\mathbf{J}_{\mathrm{x}}$, $\tilde{\mathbf{J}}_{\mathrm{a}} = \mathbf{W}^{1/2}\mathbf{J}_{\mathrm{a}}$, $\tilde{\boldsymbol{\mu}}_{\Delta y} = \mathbf{W}^{1/2}\boldsymbol{\mu}_{\Delta y}$, and $\tilde{\boldsymbol{\varepsilon}}_{\Delta y} = \mathbf{W}^{1/2}\boldsymbol{\varepsilon}_{\Delta y}$. We call Eq. (9) the generalized first-order propagating formula, which relates the error $\Delta\mathbf{x}$ in $\hat{\mathbf{x}}$ with the error sources such as the configuration error $\Delta\mathbf{a}$ as well as the systematic and random errors $\boldsymbol{\mu}_{\Delta y}$ and $\boldsymbol{\varepsilon}_{\Delta y}$ in $\mathbf{y}$. Assuming that the vector $\boldsymbol{\varepsilon}_{\Delta y}$ has a zero mean, we can derive the following equation by taking the mean value of both sides of Eq. (9) that

$$\boldsymbol{\mu}_{\Delta x} = \langle\Delta\mathbf{x}\rangle = \tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\Delta\mathbf{a} + \tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\boldsymbol{\mu}}_{\Delta y}, \tag{10}$$

where $\tilde{\mathbf{J}}_{\mathrm{x}}^{+} = (\tilde{\mathbf{J}}_{\mathrm{x}}^{\mathrm{T}}\tilde{\mathbf{J}}_{\mathrm{x}})^{-1}\tilde{\mathbf{J}}_{\mathrm{x}}^{\mathrm{T}}$ is the Moore-Penrose pseudo-inverse of the matrix $\tilde{\mathbf{J}}_{\mathrm{x}}$. Equation (10) is also called the systematic error propagating formula, which describes how the configuration error $\Delta\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in $\mathbf{y}$ lead to the systematic error $\boldsymbol{\mu}_{\Delta x}$ in $\hat{\mathbf{x}}$. According to Eq. (10), we further derive the following inequality that

$$\|\boldsymbol{\mu}_{\Delta x}\| \le \|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\| \cdot \|\Delta\mathbf{a}\| + \|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\| \cdot \|\tilde{\boldsymbol{\mu}}_{\Delta y}\|, \tag{11}$$

where the notation $\|\cdot\|$ represents the l_p $(p = 1, 2, \infty)$ vector norm and the l_p matrix norm that is induced by the associated vector norm.[15] Equation (11) can be derived according to the triangle inequality of the vector norm as well as the compatibility of the vector norm and the induced matrix norm. Equation (11) gives the upper limit of the systematic errors propagated into $\hat{\mathbf{x}}$. $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$ and $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\|$ represent the maximum gain factors in the propagation of $\Delta\mathbf{a}$ and $\boldsymbol{\mu}_{\Delta y}$, respectively.

2.3 *Measurement Configuration Optimization*

According to Eq. (10), we can estimate the systematic error $\boldsymbol{\mu}_{\Delta x}$ in $\hat{\mathbf{x}}$ if the configuration error $\Delta\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in $\mathbf{y}$ are known. We can use the estimate of the systematic error $\boldsymbol{\mu}_{\Delta x}$ as a correction term to correct the solution $\hat{\mathbf{x}}$ and obtain the corrected solution $\hat{\mathbf{x}}_0 = \hat{\mathbf{x}} + \boldsymbol{\mu}_{\Delta x}$, which will be much closer to the true value $\mathbf{x}_0$ and therefore improves the accuracy of parameter estimation. However, the configuration error $\Delta\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in $\mathbf{y}$ are usually difficult to obtain in practice, which makes the above correction unfeasible. As described in Eq. (10), the systematic error $\boldsymbol{\mu}_{\Delta x}$ in $\hat{\mathbf{x}}$ will be mainly induced by the configuration error $\Delta\mathbf{a}$ if we assume that the measurement system is well calibrated and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in $\mathbf{y}$ is small. The configuration error $\Delta\mathbf{a}$, which typically arises from the mechanical positioning errors and the finite numerical apertures of the focusing lens in the measurement system,[14] is approximately unvaried with the measurement configurations. However, the matrix $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$, which is called the configuration error propagating matrix, is a function of the measurement configuration. Similarly, $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$ is also varied with the measurement configurations. According to Eq. (11), we know that the systematic error $\boldsymbol{\mu}_{\Delta x}$ in $\hat{\mathbf{x}}$ will be small if we choose the measurement configuration with small $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$. Therefore, we can optimize the measurement configuration by

$$(\theta_{\mathrm{opt}}, \varphi_{\mathrm{opt}}) = \arg \min_{\theta\in\Theta, \varphi\in\Phi} [\max_{\mathrm{x}\in\Omega}(\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|)]. \tag{12}$$

Equation (12) needs some interpretations. Considering the local properties of the matrices $\tilde{\mathbf{J}}_{\mathrm{x}}$ and $\tilde{\mathbf{J}}_{\mathrm{a}}$ as described in Eq. (5) that is defined at $(\hat{\mathbf{x}}, \mathbf{a}^*)$, we first scan the values of $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$ in the given parameter domain Ω for the maximum. Then we scan all of the maximal values of $\|\tilde{\mathbf{J}}_{\mathrm{x}}^{+}\tilde{\mathbf{J}}_{\mathrm{a}}\|$ in the ranges of incidence and azimuthal angles (Θ and Φ) for the minimum. The combination of incidence and azimuthal angles corresponding to this minimum will be the final optimal measurement configuration. The former scan ensures that the measurement configuration is stable for the changes of structural parameters. The latter scan ensures the optimization of the overall measurement accuracy.

3 Experimental Setup

The experimental setup used in this article is a dual-rotating compensator Mueller matrix polarimeter (DRC-MMP) (RC2®, J. A. Woollam Co., Lincoln, Nebraska) with in-house forward modeling software based on rigorous coupled-wave analysis (RCWA).[16–18] As schematically shown in Fig. 1, the system setting of the dual-rotating compensator Mueller matrix polarimeter in order of light propagation is $\mathrm{PC_{r1}SC_{r2}A}$, where P and A stand for the fixed polarizer and analyzer, $\mathrm{C_{r1}}$ and $\mathrm{C_{r2}}$ refer to the first and second frequency-coupled rotating compensators, and S stands for the sample. With the light source used in this polarimeter, the wavelengths available are in the 193 and 1690 nm

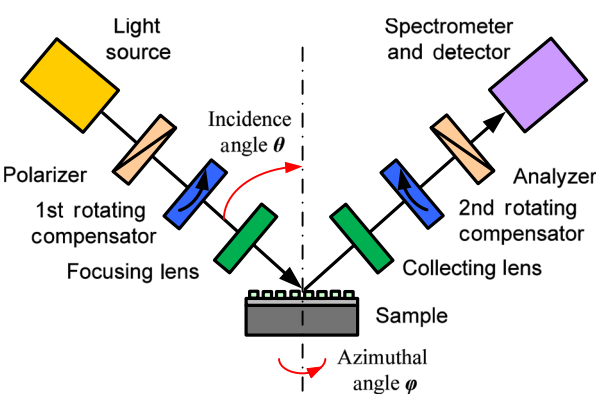

Fig. 1 Scheme of the dual-rotating compensator Mueller matrix polarimeter.

range, covering the spectral range of 200 to 800 nm used in this article. With this dual-rotating compensator setting, we can obtain the full Mueller matrix elements of the sample under measurement. See Refs. 19 and 20 for details on the data reduction.

The investigated sample is a one-dimensional Si grating, whose scanning electron microscope (SEM) cross-section image is shown in Fig. 2. The etched Si grating is chosen for this study due to its long-term dimensional stability, higher refractive index contrast, and relevance to the semiconductor industry. Optical properties of Si are taken from Ref. 21. As depicted in Fig. 2, a cross-section of the Si grating is characterized by a symmetrical trapezoidal model with top critical dimension TCD, bottom critical dimension BCD, grating height Hgt, and period pitch. Dimensions of the structural parameters obtained from Fig. 2 are TCD = 350 nm, Hgt = 472 nm, and BCD = 383 nm. In the following experiments, structural parameters of the Si grating that need to be extracted include TCD, Hgt, and BCD, while the grating period is fixed at its nominal dimension, i.e., pitch = 800 nm.

4 Results and Discussion

The procedure of measurement configuration optimization is time consuming, and it is a priority to reduce the search domain to minimize the calculation time. Since a regular grating has rotation symmetry C_{2z},[22] its Mueller matrices remain unchanged after 180 deg rotation in the azimuthal angle. In addition, the grating also has reflection symmetry relative to the plane that is perpendicular to the direction of grating period. In other words, replacing φ with $-\varphi$ changes nothing. Therefore, we can restrict the range of azimuthal angles from 0 to 90 deg. In the experiments, the incident angle is varied from 55 to 65 deg, and the spectral range is from 200 to 800 nm with an increment of 5 nm. When applying RCWA to calculate the Mueller matrices, the number of retained orders in the truncated Fourier series is 12, and the Si grating as shown in Fig. 2 is sliced into 15 layers along the vertical direction.

The systematic error propagating formula as described in Eq. (10) is the foundation for the proposed measurement configuration optimization method. It is necessary to validate this formula first before optimizing the configuration. To this

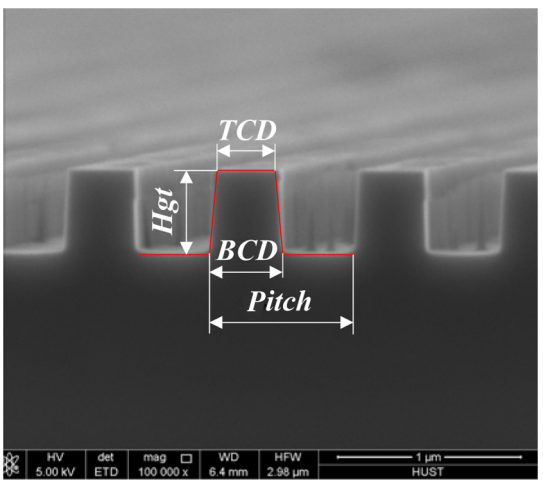

Fig. 2 Scanning electron microscope (SEM) cross-section image of the investigated Si grating.

end, we first calculate the sample Mueller matrices for a given group of structural parameters $\mathbf{x}$ and in a specific measurement configuration $\mathbf{a}^* + \Delta\mathbf{a}$. An error simulator that takes into account all the major error sources in the DRC-MMP was developed to simulate errors intrinsically induced by this polarimeter. The error sources in the DRC-MMP include the random noise in the measured light fluxes and the bias in a system-dependent vector $\mathbf{b}$, of which the former and the latter will induce random and systematic errors $\boldsymbol{\varepsilon}_{\Delta y}$ and $\boldsymbol{\mu}_{\Delta y}$ in the measured Mueller matrices, respectively. Without regard to the imperfection of the components in this polarimeter, the variables in vector $\mathbf{b}$ will be mainly the transmission axis angles of the polarizer and analyzer P and A, fast axis angles C_{S1} and C_{S2}, and phase retardances δ_1 and δ_2 of the two compensators, i.e., $\mathbf{b} = [\mathrm{P}, \mathrm{A}, C_{S1}, C_{S2}, \delta_1, \delta_2]^{\mathrm{T}}$, which are typically determined through a calibration process.[23] We then generate the "measured" Mueller matrices by adding the above calculated sample Mueller matrices with the simulated errors and use them to imitate the actually measured sample Mueller matrices in the procedure of parameter extraction. The Levenberg–Marquardt (LM) algorithm[24] is then applied to extract the structural parameters from the measured Mueller matrices with the measurement configuration fixed at $\mathbf{a}^*$. The LM algorithm typically converges rapidly to the global minimum if suitable initial values are provided. It is certain that there will exist errors $\Delta\mathbf{x}$ (including random and systematic errors) in the extracted structural parameters $\hat{\mathbf{x}}$, i.e., $\Delta\mathbf{x} = \hat{\mathbf{x}} - \mathbf{x}$. If we repeat the above procedure n times, we will obtain n groups of extracted structural parameters. The mean value of the errors $\Delta\mathbf{x}$ in the n groups of extracted structural parameters $(\Delta\mathbf{x})$ will be the statistically calculated systematic error in $\hat{\mathbf{x}}$, which is induced by both the configuration error $\Delta\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in the measured Mueller matrices (induced by the bias $\Delta\mathbf{b}$ in vector $\mathbf{b}$). The above statistically calculated systematic error is then compared with those theoretically estimated by Eq. (10) to examine its validity and to identify which error sources will contribute most in limiting the accuracy of the grating reconstruction.

Figure 3 depicts the comparison between the statistically calculated systematic errors in the extracted structural parameters TCD, Hgt, and BCD and those theoretically estimated by Eq. (10). The structural parameters corresponding to Fig. 3 are TCD = 350 nm, Hgt = 472 nm, and BCD = 383 nm, which are the results measured by SEM. The incidence angle θ is fixed at 60 deg, and the azimuthal angles ϕ are varied from 0 to 90 deg with an increment of 5 deg. In Fig. 3, the configuration error is $\Delta\mathbf{a} = [\Delta\theta, \Delta\varphi]^{\mathrm{T}} = [0.5\ , 1.0\ ]^{\mathrm{T}}$ deg, and the bias in vector $\mathbf{b}$ is $\Delta\mathbf{b} = [\Delta\mathrm{P}, \Delta\mathrm{A}, \Delta C_{S1}, \Delta C_{S2}, \Delta\delta_1, \Delta\delta_2]^{\mathrm{T}} = [0.5, 0.5, 0.5, 0.5, 1.0, 1.0\ ]^{\mathrm{T}}$ deg. The biases in the incidence angle θ and vector $\mathbf{b}$ were roughly estimated from measurements acquired for air as well as a SiO_2 film with known thickness on a Si substrate. The bias in the azimuthal angle φ was roughly estimated according to the symmetry in the measured Mueller matrices of the investigated Si grating sample. These values represent a high level of biases in measurements, probably exceeding most instrumental situations, and are used here to examine the validity of Eq. (10) and to investigate the influences of different error sources on the final extracted structural parameters. As observed from Fig. 3, the theoretically estimated systematic errors show a good agreement with those

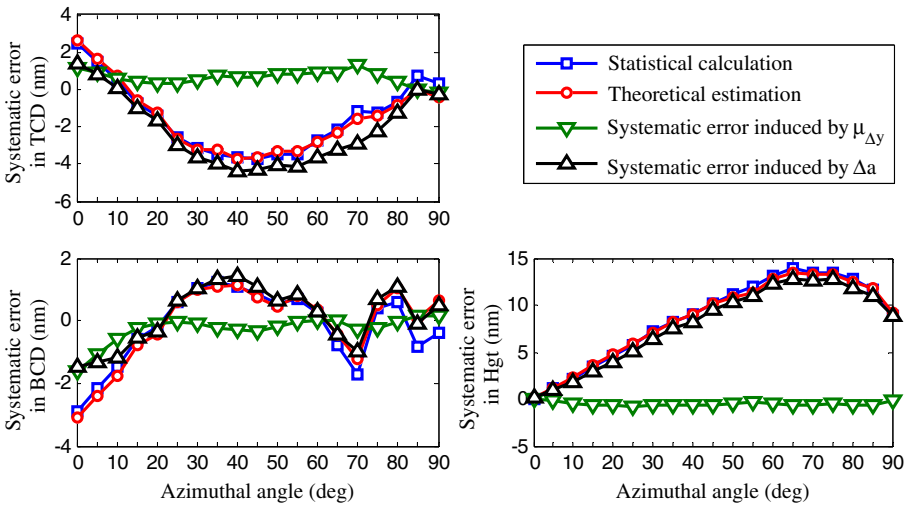

Fig. 3 Comparison between the statistically calculated systematic errors in the extracted structural parameters TCD, Hgt and BCD and those theoretically estimated by Eq. (10). The curves marked with upward and downward triangles are all the theoretically calculated systematic errors and are used to represent the systematic errors in the extracted parameters that are induced by the configuration error $\mathbf{\Delta a}$ and the systematic error $\boldsymbol{\mu}_{\Delta y}$ in the "measured" Mueller matrices, respectively.

statistically calculated ones, which therefore demonstrates the validity of the derived systematic error propagating formula given by Eq. (10). We also observe from Fig. 3 that the configuration error $\mathbf{\Delta a}$ has a greater influence on the systematic error $\boldsymbol{\mu}_{\Delta \mathbf{x}}$ in $\hat{\mathbf{x}}$ than the systematic error $\boldsymbol{\mu}_{\Delta y}$ in the measured Mueller matrix elements does. Thus, we can focus on the configuration error $\mathbf{\Delta a}$ in the following experiments and use the norm of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ as the cost function to further optimize the measurement configuration for MMP.

In order to optimize the measurement configuration for MMP, the l_2 norm of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ was calculated in a parameter domain with TCD varied from 345 to 355 nm, Hgt from 465 to 475 nm, and BCD from 378 to 388 nm. In the optimization procedure, the incidence and azimuthal angles were varied from 55 to 65 deg and from 0 to 90 deg, respectively, both with an increment of 5 deg. In the given parameter domain, the maximal norms $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ calculated in different measurement configurations are presented in Fig. 4. According to Eq. (12), we know that the minimum of all the maximal norms $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ corresponds to the final optimal measurement configuration. As observed from Fig. 4, the norm of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ calculated when the incidence angle $\theta = 55$ deg and azimuthal angle $\varphi = 20$ deg is smaller than those calculated in other configurations. It is therefore expected that the structural parameters extracted in $\theta = 55$ deg and $\varphi = 20$ deg will have smaller systematic errors and will be more accurate than other measurement configurations.

In order to examine the above prediction of the optimal measurement configuration, the Si grating sample as shown in Fig. 2 was measured by the RC2® polarimeter in the incidence and azimuthal angles varied from 55 to 65 deg and from 0 to 90 deg, respectively, both with an increment of 5 deg. The systematic errors in the structural parameters are defined as the differences between the structural parameters of the Si grating sample extracted by the LM algorithm and the results measured by SEM. The l_2 norm of the systematic errors $\|\boldsymbol{\mu}_{\Delta x}\|$ in the extracted structural parameters was then calculated for each measurement configuration and shown in Fig. 5. An examination of Fig. 5 shows that the variation of the norms of the systematic errors $\|\boldsymbol{\mu}_{\Delta x}\|$ is not in rigorous agreement with that of the norms of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ as depicted in Fig. 4. It may be because the relation given by Eq. (11) is not a rigorous equality but an inequality. However, we do have a qualitative agreement. For example, the greater the incidence and azimuthal angles are, the greater the norms of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ are, and then the greater the norms of the systematic errors $\|\boldsymbol{\mu}_{\Delta x}\|$ are. Importantly, the optimal measurement configuration shown in Fig. 5, whose corresponding norm of the systematic error $\|\boldsymbol{\mu}_{\Delta x}\|$ achieves the minimum value, is in accordance with the theoretical prediction given in Fig. 4. Figure 6 illustrates the fitting result of the calculated and the polarimeter measured Mueller matrices in the optimal measurement configuration $\theta = 55$ deg and $\varphi = 20$ deg. The extracted structural parameters with 95% confidence level are TCD $= 348.58 \pm 0.024$ nm,

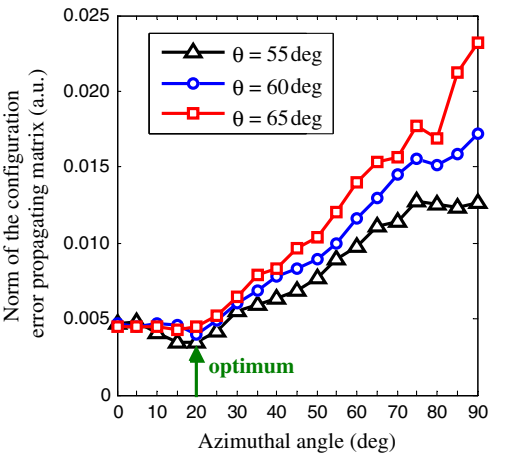

Fig. 4 The maximal norms of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ calculated in different measurement configurations.

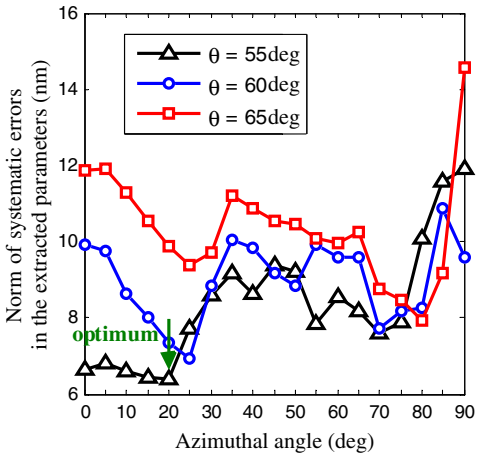

Fig. 5 The norms of the systematic errors $\|\boldsymbol{\mu}_{\Delta x}\|$ in the structural parameters of the Si grating sample extracted in different measurement configurations.

Hgt = 472.91 ± 0.017 nm, and BCD = 389.16 ± 0.026 nm. Here, the 95% confidence level uncertainties appended to the extracted parameter values are estimated by propagating the random uncertainties in the grating reconstruction procedure. The estimate of the standard deviations of the measured Mueller matrix elements is provided by the CompleteEASE™ software supplied with the RC2® instrument. As observed from Fig. 6, the Mueller matrices of the Si grating sample calculated in the optimal measurement configuration show a good agreement with the measured Mueller matrices. Consequently, we can conclude that the norm of the configuration error propagating matrix $\|\tilde{\mathbf{J}}_x^+ \tilde{\mathbf{J}}_a\|$ can be applied as an objective function to optimize the measurement configuration for MMP to achieve more accurate measurement.

5 Conclusions

In this article, the measurement configuration optimization for spectroscopic MMP was investigated to find an optimal combination of the incidence and azimuthal angles, with which more accurate measurement can be achieved. We derived a systematic error propagating formula which relates the systematic errors in the extracted structural parameters with the error sources such as the configuration error and the intrinsic systematic error in the measured Mueller matrices. Simulations performed on a Si grating sample have demonstrated the validity of this formula and also revealed that the configuration error has a greater influence on the systematic errors in the extracted structural parameters than the intrinsic systematic error in the measured Mueller matrices does. We then adopted the norm of the configuration error propagating matrix to assess the influence of the configuration error on the measurement accuracy for different configurations. The optimal configuration with the incidence angle $\theta = 55$ deg and azimuthal angle $\varphi = 20$ deg was achieved for the investigated Si grating sample by minimizing the norm of the configuration error propagating matrix. Experiments performed with a dual-rotating compensator Mueller matrix polarimeter show the agreement between the theoretically predicted optimal configuration and the experimentally exhibited one. It is also worthwhile to point out that the optimal configuration is sample dependent, but the proposed method is general and may be extended to other scatterometric techniques.

Acknowledgments

This work was funded by National Natural Science Foundation of China (Grant No. 91023032, 51005091, and 51121002) and National Instrument Development Specific Project of China (Grant No. 2011YQ160002).

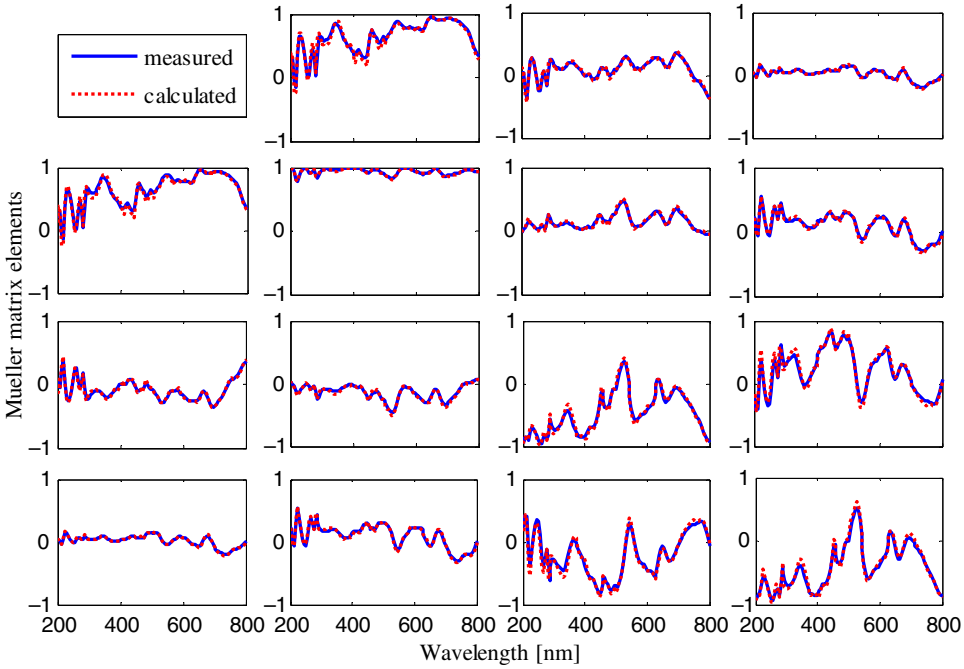

Fig. 6 Fitting result of the calculated and the polarimeter measured Mueller matrices of the Si grating sample in the optimal measurement configuration $\theta = 55$ deg and $\varphi = 20$ deg.

References

1. T. Novikova et al., "Metrology of replicated diffractive optics with Mueller polarimetry in conical diffraction," *Opt. Express* **15**(5), 2033–2046 (2007).
2. Y. N. Kim et al., "Device based in-chip critical dimension and overlay metrology," *Opt. Express* **17**(23), 21336–21343 (2009).
3. J. Li et al., "Mueller matrix measurement of asymmetric gratings," *J. Micro/Nanolith. MEMS MOEMS* **9**(4), 041305 (2010).
4. P. C. Logofatu, "Sensitivity analysis of grating parameter estimation," *Appl. Opt.* **41**(34), 7179–7186 (2002).
5. P. C. Logofatu, "Phase-modulation scatterometry," *Appl. Opt.* **41**(34), 7187–7192 (2002).
6. M. Littau et al., "Diffraction signature analysis methods for improving scatterometry precision," *Proc. SPIE* **6152**, 615236 (2006).
7. H. Gross and A. Rathsfeld, "Sensitivity analysis for indirect measurement in scatterometry and the reconstruction of periodic grating structures," *Wave Random Complex* **18**(1), 129–149 (2008).
8. P. Vagos et al., "Uncertainty and sensitivity analysis and its applications in OCD measurement," *Proc. SPIE* **7272**, 72721N (2009).
9. T. Novikova et al., "Application of Mueller polarimetry in conical diffraction for critical dimension measurements in microelectronics," *Appl. Opt.* **45**(16), 3688–3697 (2006).
10. M. Foldyna et al., "Critical dimension of biperiodic gratings determined by spectral ellipsometry and Mueller matrix polarimetry," *Eur. Phys. J. Appl. Phys.* **42**, 351–359 (2008).
11. R. M. Al-Assaad and D. M. Byrne, "Error analysis in inverse scatterometry. I. Modeling," *J. Opt. Soc. Am. A* **24**(2), 326–338 (2007).
12. T. A. Germer et al., "Developing uncertainty analysis for optical scatterometry," *Proc. SPIE* **7272**, 72720T (2009).
13. T. A. Germer, "The instruction manual for the OCDSense," ftp://ftp.nist.gov/pub/physics/germer/OCDSense/ (1 March 2013).
14. T. A. Germer and H. J. Patrick, "Effect of bandwidth and numerical aperture in optical scatterometry," *Proc. SPIE* **7638**, 76381F (2010).
15. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Chapter 5, Cambridge University Press, New York (1985).
16. M. G. Moharam et al., "Formulation for stable and efficient implementation of the rigorous coupled wave analysis of binary gratings," *J. Opt. Soc. Am. A* **12**(5), 1068–1076 (1995).
17. L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," *J. Opt. Soc. Am. A* **13**(9), 1870–1876 (1996).
18. S. Y. Liu et al., "Estimation of the convergence order of rigorous coupled-wave analysis for binary gratings in optical critical dimension metrology," *Opt. Eng.* **51**(8), 081504 (2012).
19. R. W. Collins and J. Koh, "Dual rotating-compensator multichannel ellipsometer: instrument design for real-time Mueller matrix spectroscopy of surfaces and films," *J. Opt. Soc. Am. A*, **16**(8), 1997–2006 (1999).
20. R. W. Collins, "Multichannel ellipsometry," Chapter 7 in *Handbook of Ellipsometry*, H. G. Tompkins and E. A. Irene, Eds., William Andrew Publishing & Springer-Verlag, New York (2005).
21. C. M. Herzinger et al., "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**(6), 3323–3336 (1998).
22. L. Li, "Symmetries of cross-polarization diffraction coefficients of gratings," *J. Opt. Soc. Am. A* **17**(5), 881–887 (2000).
23. L. Broch, A. En Naciri, and L. Johann, "Systematic errors for a Mueller matrix dual rotating compensator ellipsometer," *Opt. Express* **16**(12), 8814–8824 (2008).
24. W. H. Press et al., *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Chapter 15, Cambridge University Press, New York (2007).

Xiuguo Chen is currently a PhD candidate at Huazhong University of Science and Technology under the guidance of Prof. Shiyuan Liu. He received his MS degree from the School of Mechanical Science and Engineering of the same university in 2009. His research involves various issues in OCD metrology, including fast optical modeling and robust parameter extraction. He is a student member of SPIE, OSA, and IEEE.

Shiyuan Liu is a professor of mechanical engineering at Huazhong University of Science and Technology, China, leading his Nanoscale and Optical Metrology Group with research interest in metrology and instrumentation for nanomanufacturing. He also actively works in the area of optical lithography, including partially coherent imaging theory, wavefront aberration metrology, optical proximity correction, source mask optimization, and inverse lithography technology. He received his PhD in mechanical engineering from Huazhong University of Science and Technology, China, in 1998. He is a member of SPIE, OSA, AVS, IEEE, and CSMNT (Chinese Society of Micro/Nano Technology). He holds 30 patents and has authored or coauthored more than 100 technical papers.

Chuanwei Zhang is an assistant professor at Huazhong University of Science and Technology, China. He received his BE and ME in mechanical engineering from Wuhan University in 2004 and 2006, respectively, and then received his PhD in mechanical engineering from Huazhong University of Science and Technology, in 2009. He is currently working on optical techniques for critical dimension, overlay, and 3-D profile metrology for nanomanufacturing. He is a member of SPIE, OSA, and IEEE.

Hao Jiang is currently a postdoctoral research associate and adjunct instructor in the Department of Mechanical and Aerospace Engineering at University of Texas at Arlington. He earned his BS and MS degrees in Mechanical Engineering from Huazhong University of Science and Technology, China, in 2000 and 2004, respectively, and his PhD degree from Florida Institute of Technology, Melbourne, Florida, in 2011. His primary research interests are in micro/nano scale metrology, unpowered wireless acoustic emission sensing, and the development of state-of-the-art sensor technologies and mathematic models for structural health monitoring and biomedical applications.