Measurement configuration optimization for grating reconstruction by Mueller matrix polarimetry

Xiuguo Chen [a], Shiyuan Liu [*, a, b], Chuanwei Zhang [a], and Hao Jiang [c]

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

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

[c] Department of Mechanical and Aerospace Engineering, University of Texas at Arlington, Arlington 76019, USA

ABSTRACT

As a non-imaging optical measurement technique, spectroscopic Mueller matrix polarimetry (MMP) has been introduced for critical dimension (CD) and overlay metrology with recent great success. Due to the additional information provided by the Mueller matrices when the most general conical diffraction configuration is considered, MMP has demonstrated a great potential in semiconductor manufacturing. In order to make full use of the additional information provided by the Mueller matrices, it is of great importance for MMP to optimize the measurement configuration. In this paper, we introduce the norm of a configuration error propagating matrix as the cost function to optimize the measurement configuration for spectroscopic MMP with the aim of finding an optimal combination of fixed incidence and azimuthal angles, which provides higher measurement accuracy. The optimal measurement configuration can be achieved by minimizing the norm of the configuration error propagating matrix in the available ranges of incidence and azimuthal angles. Experiments performed on a silicon grating with a dual-rotating compensator Mueller matrix polarimeter have demonstrated the validity of the proposed measurement configuration optimization method.

Keywords: measurement configuration, grating reconstruction, Mueller matrix polarimetry, inverse problem

1. INTRODUCTION

The optics-based metrology tools have drawn more and more attention in semiconductor manufacturing due to their attractive advantages, such as low cost, non-destruction, and high throughput. Based on the conventional spectroscopic or angle-resolved ellipsometry, optical scatterometry, also referred to as optical critical dimension (OCD) metrology, has been widely used for critical dimension (CD) monitoring [1-3]. The success of optical scatterometry heavily depends on two key techniques. The first one involves the calculation of the optical signature from a diffractive structure (often grating) using accurate forward modeling approaches, such as rigorous coupled-wave analysis (RCWA) [4-6], finite element method (FEM) [7, 8], or finite-difference time-domain (FDTD) method [9]. Here, the general term *signature* contains the scattered light information from the diffractive structure, which can be in the form of reflectance, ellipsometric angles, Stokes vector elements, or Mueller matrix elements. The second one involves the reconstruction of the structural profile from the measured signature, which is a typical inverse diffraction problem with the objective of finding a profile whose calculated signature can best match the measured one.

Recently, the Mueller matrix polarimetry (MMP) has been successfully introduced for CD and overlay metrology [10-12]. Due to the rich information provided by the Mueller matrices when the grating lines are no longer perpendicular to the incidence plane but 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

* Corresponding author: shyliu@mail.hust.edu.cn; phone: +86 27 8755 9543; webpage: http://www2.hust.edu.cn/nom.

Metrology, Inspection, and Process Control for Microlithography XXVII, edited by Alexander Starikov,
Proc. of SPIE Vol. 8681, 86812B · © 2013 SPIE · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2011487

azimuthal angles is defined as the *measurement configuration*. For example, we can fix the incidence and azimuthal angles in proper values while continuously vary the wavelengths in an available range. Similarly, we can also fix the wavelength and azimuthal angle in proper values while continuously vary the incidence angles in an available range. In OCD metrology, the former configuration is usually called the spectrally resolved or spectroscopic scatterometry, while the latter is called the (incidence) angle resolved or angular scatterometry. We can obtain a multitude of possible measurement configurations by making different combinations of the three measurement conditions. It is worth 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 configurations, which provides the best measurement precision and accuracy.

In the past decades, several approaches have been proposed to optimize the measurement configuration for conventional ellipsometric scatterometry. Logofatu proposed a SAF (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 [13, 14]. Littau et al. investigated several optimal diffraction signature scan path selection techniques to improve scatterometry precision [15]. Gross et al. proposed an algorithm to determine the optimal measurement data sets by minimizing the condition numbers of the corresponding Jacobian matrices, which are defined as the partial derivatives of the diffraction signature with respect to the structural parameters [16]. Vagos et al. developed an uncertainty and sensitivity analysis package that can be used to guide the model and azimuthal angle optimization processes [17]. A recent study on spectroscopic MMP reported that the Mueller matrices obtained in some measurement configurations may help decorrelate the fitting structural parameters and make the solution of the inverse diffraction problem more robust [18]. They further proposed to choose the measurement configuration with small correlations and small estimated precision of the structural parameters [19]. However, it can be also observed from the previous works [13, 14, 19] that the extracted structural parameter values vary greatly in different measurement configurations. Typical minimization of the parameter correlations and estimated precision may achieve an optimal configuration with higher measurement precision but cannot make sure the final measurement accuracy.

In this paper, we propose a measurement configuration optimization method for spectroscopic MMP with the aim of finding an optimal combination of fixed incidence and azimuthal angles, which provides higher measurement accuracy. The biases in the incidence and azimuthal angles, also named the configuration error, which typically arise from the mechanical positioning errors as well as the finite numerical apertures of the focusing lens in the measurement system [20, 21], will induce systematic errors in the extracted structural parameters and thus influence the final measurement accuracy. We first derive a generalized first order error propagating formula to reveal the mechanism of the error propagation in grating reconstruction. Based on the generalized error propagating formula, a systematic error propagating formula is further derived, which relates the systematic error propagated into the extracted structural parameter with those error sources such as the configuration error and the intrinsic systematic error in the measured Mueller matrices. According to the systematic error propagating formula, we introduce the norm of the configuration error propagating matrix, which reveals the maximum gain factor in the propagation of the configuration error, to assess the influence of the biases in the incidence and azimuthal angles on the measurement accuracy. The configurations with small matrix norms imply minor influence of the configuration error on final measurement accuracy. Then we attempt to optimize the measurement configuration by minimizing the norm of the configuration error propagating matrix in the available ranges of incidence and azimuthal angles.

The reminder of this paper is organized as follows. Section 2 first introduces the measurement setup for the grating sample, including the configuration of a dual-rotating compensator Mueller matrix polarimeter as well as the geometric structure of the silicon grating sample. The inverse problem in grating reconstruction by MMP is also briefly revisited. Then the proposed measurement configuration optimization method is illustrated in detail. Section 3 provides the optimization results and the comparison with experimental results. Finally, we draw some conclusions in Section 4.

2. METHODS

2.1 Measurement setup for the grating sample

The experimental setup used in this paper is a dual-rotating compensator Mueller matrix polarimeter (RC2 ellipsometer, J. A. Woollam Co.) with in-house forward modeling software based on RCWA [4-6]. As schematically shown in Fig. 1(a), the system configuration of the RC2 ellipsometer in order of light propagation is $PC_{r1}SC_{r2}A$, where P and A stand

for the fixed polarizer and analyzer, C_{r1} and C_{r2} refer to the 1st and 2nd frequency-coupled rotating compensators, and S stands for the sample. With the light source used in the RC2 ellipsometer, the wavelengths available are in the 193~1690nm range, covering the spectral range of 200~800nm with a step of 5nm used in this paper. In this dual-rotating compensator configuration, we can obtain the full Mueller matrix elements of the sample, see Refs. [22, 23] for details on the data reduction.

The investigated sample is a one-dimensional silicon grating, whose scanning electron microscope (SEM) cross-section image is shown in Fig. 1(b). 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 silicon are taken from Ref. [24]. As depicted in Fig. 1(b), cross section of the Si grating is characterized by a symmetrical trapezoidal model with top critical dimension *TCD*, sidewall angle *SWA*, line height *Hgt*, and period *pitch*. Dimensions of the structural parameters obtained from Fig. 1(b) are *TCD* = 350nm, *Hgt* = 472nm, and *SWA* = 88°. In the following experiments, structural parameters of the Si grating that need to be extracted include *TCD*, *Hgt* and *SWA*, while the grating period is fixed at its nominal dimension, i.e. *pitch* = 800nm.

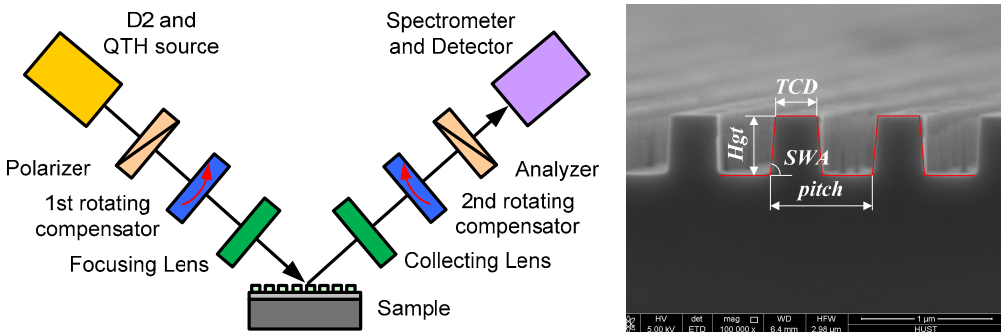

Figure 1. (a) Basic scheme of the dual-rotating compensator Mueller matrix polarimeter; (b) SEM cross-section image of the investigated silicon grating.

2.2 The inverse problem in grating reconstruction

Without loss of generality, we denote the structural parameters under measurement as an *M*-dimensional vector $\mathbf{x} = [x_1, x_2, \ldots, x_M]^T$, where the superscript "T" represents the transpose, and $x_1, x_2, \ldots, x_M$ can be the linewidth, sidewall angle, thickness of a grating sample. The vector $\mathbf{a} = [\theta, \varphi]^T$ denotes the combination of fixed incidence angle θ and azimuthal angle φ, and the given values of $\mathbf{a}$ in the parameter extraction is denoted as $\mathbf{a}^*$. The χ^2 function is applied to estimate the fitting errors between the measured and calculated Mueller matrix elements $m_{ij,k}^{\text{meas}}$ and $m_{ij,k}^{\text{calc}}(\mathbf{x}, \mathbf{a}^*)$, which is defined as

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

where k denotes the spectral point from the total number N_λ. Indices i, j show all the Mueller matrix elements except m_{11}. $\sigma(m_{ij,k})$ is the standard deviation associated with $m_{ij,k}$. The optimal estimation of the structural parameters under measurement can be achieved by solving

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \Omega} \sum_{k=1}^{N_\lambda} \sum_{i,j} \left[\frac{m_{ij,k}^{\text{meas}} - m_{ij,k}^{\text{calc}}(\mathbf{x}, \mathbf{a}^*)}{\sigma(m_{ij,k})} \right]^2. \tag{2}$$

For clarity, the measured Mueller matrix element $m_{ij,k}^{\text{meas}}$ in Eqs. (1) and (2) is marked as y_l with three indices i, j and k lumped into a single index l. The calculated Mueller matrix element $m_{ij,k}^{\text{calc}}(\mathbf{x}, \mathbf{a}^*)$ is correspondingly marked as $f_l(\mathbf{x}, \mathbf{a}^*)$. Therefore, Eq. (1) can be simply rewritten as

$$\chi^2 = \sum_{l=1}^{N} w_l \left[y_l - f_l(\mathbf{x}, \mathbf{a}^*) \right]^2, \tag{3}$$

where w_l is the weighting factor and is given by $w_l = 1/\sigma^2(y_l)$ and $N = 15N_\lambda$. Equation (3) can be further written as a matrix expression

$$\chi^2 = \left[\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a}^*) \right]^{\mathrm{T}} \mathbf{W} \left[\mathbf{y} - \mathbf{f}(\mathbf{x}, \mathbf{a}^*) \right], \tag{4}$$

where **W** is an $N \times N$ diagonal matrix with diagonal elements w_l.

2.3 Measurement configuration optimization

We assume that the function **f**(**x**, **a**) is sufficiently smooth and can be Taylor series expanded about $(\hat{\mathbf{x}}, \mathbf{a}^*)$

$$\mathbf{f}(\mathbf{x}, \mathbf{a}) \approx \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) + \mathbf{J}_\mathbf{x} \cdot (\mathbf{x} - \hat{\mathbf{x}}) + \mathbf{J}_\mathbf{a} \cdot (\mathbf{a} - \mathbf{a}^*), \tag{5}$$

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

$$\left[\mathbf{J}_\mathbf{x} \right]_{ij} = \left. \frac{\partial f_i(\mathbf{x}, \mathbf{a})}{\partial x_j} \right|_{\mathbf{x}=\hat{\mathbf{x}},\, \mathbf{a}=\mathbf{a}^*}, \tag{6a}$$

$$\left[\mathbf{J}_\mathbf{a} \right]_{ij} = \left. \frac{\partial f_i(\mathbf{x}, \mathbf{a})}{\partial a_j} \right|_{\mathbf{x}=\hat{\mathbf{x}},\, \mathbf{a}=\mathbf{a}^*}. \tag{6b}$$

According to Eq. (5), we will have

$$\mathbf{f}(\mathbf{x}_0, \mathbf{a}_0) \approx \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) + \mathbf{J}_\mathbf{x} \Delta\mathbf{x} + \mathbf{J}_\mathbf{a} \Delta\mathbf{a}, \tag{7}$$

where $\Delta\mathbf{x} = \mathbf{x}_0 - \hat{\mathbf{x}}$ and $\Delta\mathbf{a} = \mathbf{a}_0 - \mathbf{a}^*$. $\mathbf{x}_0$ and $\mathbf{a}_0$ are the true values of **x** and **a**, respectively. The measurement vector **y** will be the sum of the true signal $\mathbf{f}(\mathbf{x}_0, \mathbf{a}_0)$ and a deterministic offset vector $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ and a random vector $\boldsymbol{\varepsilon}_{\Delta\mathbf{y}}$, i.e.,

$$\mathbf{y} = \mathbf{f}(\mathbf{x}_0, \mathbf{a}_0) + \boldsymbol{\mu}_{\Delta\mathbf{y}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{y}}, \tag{8}$$

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

Inserting Eqs. (7) and (8) into Eq. (4), we will find, near the optimal estimation that

$$\chi^2_{\min} = \left[\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) \right]^{\mathrm{T}} \mathbf{W} \left[\mathbf{y} - \mathbf{f}(\hat{\mathbf{x}}, \mathbf{a}^*) \right] \approx [\mathbf{J}_\mathbf{x} \Delta\mathbf{x} + \mathbf{J}_\mathbf{a} \Delta\mathbf{a} + \boldsymbol{\mu}_{\Delta\mathbf{y}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{y}}]^{\mathrm{T}} \mathbf{W} [\mathbf{J}_\mathbf{x} \Delta\mathbf{x} + \mathbf{J}_\mathbf{a} \Delta\mathbf{a} + \boldsymbol{\mu}_{\Delta\mathbf{y}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{y}}]. \tag{9}$$

By taking the derivatives of both sides of Eq. (9) with respect to each element of **x**, we can derive

$$\tilde{\mathbf{J}}_\mathbf{x} \Delta\mathbf{x} + \tilde{\mathbf{J}}_\mathbf{a} \Delta\mathbf{a} + \tilde{\boldsymbol{\mu}}_{\Delta\mathbf{y}} + \tilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{y}} \approx 0. \tag{10}$$

where $\tilde{\mathbf{J}}_\mathbf{x} = \mathbf{W}^{1/2} \mathbf{J}_\mathbf{x}$, $\tilde{\mathbf{J}}_\mathbf{a} = \mathbf{W}^{1/2} \mathbf{J}_\mathbf{a}$, $\tilde{\boldsymbol{\mu}}_{\Delta\mathbf{y}} = \mathbf{W}^{1/2} \boldsymbol{\mu}_{\Delta\mathbf{y}}$, and $\tilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{y}} = \mathbf{W}^{1/2} \boldsymbol{\varepsilon}_{\Delta\mathbf{y}}$. We call Eq. (10) *the generalized first-order error propagating formula*, which relates the error $\Delta\mathbf{x}$ in $\hat{\mathbf{x}}$ with those error sources such as the configuration error $\Delta\mathbf{a}$ in the given $\mathbf{a}^*$ and the intrinsic systematic and random errors $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ and $\boldsymbol{\varepsilon}_{\Delta\mathbf{y}}$ in the measurement vector **y**. Assuming the random vector $\boldsymbol{\varepsilon}_{\Delta\mathbf{y}}$ has zero mean, we can derive the following equation by taking the mean values of both sides of Eq. (10), i.e.,

$$\boldsymbol{\mu}_{\Delta\mathbf{x}} = \langle \Delta\mathbf{x} \rangle \approx \tilde{\mathbf{J}}_\mathbf{x}^{+} \tilde{\mathbf{J}}_\mathbf{a} \Delta\mathbf{a} + \tilde{\mathbf{J}}_\mathbf{x}^{+} \tilde{\boldsymbol{\mu}}_{\Delta\mathbf{y}}, \tag{11}$$

where $\tilde{\mathbf{J}}_{\mathbf{x}}^{+} = (\tilde{\mathbf{J}}_{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{J}}_{\mathbf{x}})^{-1}\tilde{\mathbf{J}}_{\mathbf{x}}^{\mathrm{T}}$ is the Moore-Penrose pseudo-inverse of the matrix $\tilde{\mathbf{J}}_{\mathbf{x}}$. We call Eq. (11) the *systematic error propagating formula*, which describes how the configuration error $\Delta\mathbf{a}$ in vector $\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ in vector $\mathbf{y}$ lead to the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in $\hat{\mathbf{x}}$. According to Eq. (11), we can derive the relation that

$$\left\|\boldsymbol{\mu}_{\Delta\mathbf{x}}\right\| \le \left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\| \cdot \left\|\Delta\mathbf{a}\right\| + \left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\right\| \cdot \left\|\tilde{\boldsymbol{\mu}}_{\Delta\mathbf{y}}\right\|, \tag{12}$$

where the notation $\left\|\cdot\right\|$ represents the l_p (p = 1, 2, ∞) norm [25]. $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ and $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\right\|$ represent the maximum gain factors in the propagation of $\Delta\mathbf{a}$ and $\boldsymbol{\mu}_{\Delta\mathbf{y}}$.

If we assume that the measurement system is well calibrated and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ in $\mathbf{y}$ is a small quantity, thus the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in $\hat{\mathbf{x}}$ can be deemed to be mainly induced by the configuration error $\Delta\mathbf{a}$. The configuration error $\Delta\mathbf{a}$ typically arises from the mechanical positioning errors and the finite numerical apertures of the focusing lens in the measurement system [20, 21], which is approximately unvaried with the measurement configurations. However, the matrix $\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}$, which we call *the configuration error propagating matrix*, is a function of the measurement configuration. Similarly, $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ is also varied with the measurement configurations. Hence, we obtain different systematic errors $\Delta\mathbf{x}$ in $\hat{\mathbf{x}}$ in different measurement configurations. According to Eq. (11), we can estimate the systematic error $\Delta\mathbf{x}$ if we have known the configuration error $\Delta\mathbf{a}$ as well as the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ in $\mathbf{y}$. Thus, we can adjust the optimal estimation $\hat{\mathbf{x}}$ by $\hat{\mathbf{x}}_0 = \hat{\mathbf{x}} + \boldsymbol{\mu}_{\Delta\mathbf{x}}$, which will be closer to the actual parameter values $\mathbf{x}_0$, and therefore improve the accuracy of parameter estimation. However, the configuration error $\Delta\mathbf{a}$ and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{y}}$ in $\mathbf{y}$ are sometimes difficult to obtain, which will make this adjustment unfeasible. In this case, we can optimize the measurement configuration by

$$\left(\theta_{\mathrm{opt}}, \varphi_{\mathrm{opt}}\right) = \min_{\theta\in\Theta,\,\varphi\in\Phi}\left[\max_{\mathbf{x}\in\Omega}\left(\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|\right)\right]. \tag{13}$$

Equation (13) needs some interpretations. Considering the local properties of the matrices $\tilde{\mathbf{J}}_{\mathbf{x}}$ and $\tilde{\mathbf{J}}_{\mathbf{a}}$ as described in Eq. (6), which is defined at $\hat{\mathbf{x}}$ and $\mathbf{a}^{*}$, we first scan the values of $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ in the given parameter domain Ω for the maximum. Then we scan all of the maximal values of $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ 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. RESULTS

The procedure of measurement configuration optimization is time-consuming, and it is necessary to reduce the search domain to minimize the calculation time. Since a regular grating has rotation symmetry C_{2z} [26], its Mueller matrices remain unchanged after 180° rotation in the azimuthal angle. In addition, the grating also has reflection symmetry relative to the plane that 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 to 0~90°. The incidence angle is varied from 60° to 65° in the experiment. 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. 1(b) is sliced into 15 layers along the vertical direction.

In addition, the systematic error propagating formula as described in Eq. (11) is the foundation for the proposed measurement configuration optimization method. Thus, it is necessary to validate Eq. (11) first before optimizing the configuration. In order to validate this formula, we first calculate the spectral Mueller matrices for a given group of structural parameters $\mathbf{x}$ and in a specific measurement configuration $\mathbf{a}^{*} + \Delta\mathbf{a}$. The calculated Mueller matrices are then treated as the "measured" Mueller matrices in the process of parameter extraction. When extracting the structural parameters from the "measured" Mueller matrices, the measurement configuration is fixed at $\mathbf{a}^{*}$. It is certain that there will be a bias $\delta\mathbf{x}$ in the extracted parameters $\hat{\mathbf{x}}$, i.e. $\Delta\mathbf{x} = \hat{\mathbf{x}} - \mathbf{x}$, which will be the systematic errors induced by the

configuration error $\delta\mathbf{a}$. The "measured" systematic errors are then compared with those estimated by Eq. (11). The Levenberg-Marquardt (LM) algorithm [27] is applied to extract the structural parameters from the "measured" Mueller matrices, which typically converges rapidly to the global minimum if suitable initial values are provided.

Figure 2 depicts the comparison between the "measured" systematic errors in the extracted structural parameters *TCD*, *Hgt* and *SWA* with those estimated according to Eq. (11). The given parameter values in Fig. 2 are *TCD* = 350nm, *Hgt* = 472nm, and *SWA* = 88°, which are the results measured by SEM. The incidence angle is fixed at 65°, and the azimuthal angles are varied from 0° to 90° with a step size of 15°. In Fig. 2 (a) and (b), the biases in the incidence and azimuthal angles are 1.0° and 1.5°, respectively. In Fig. 2(b), we further add systematic errors $\boldsymbol{\mu}_{\Delta y}$ into the generated "measured" Mueller matrices, which arise from the inaccurate calibration of the angles of the transmission axes of the fixed polarizer and analyzer P and A, inaccurate calibration of the angles of the fast axes C_{S1} and C_{S2} and phase retardances δ_1 and δ_2 of the two compensators. The given offsets in P, A, C_{S1}, C_{S2}, δ_1 and δ_2 in Fig. 2(b) are $\Delta P = 0.5°$, $\Delta A = 0.5°$, $\Delta C_{S1} = 0.5°$, $\Delta C_{S2} = 0.5°$, $\Delta\delta_1 = 1.0°$, and $\Delta\delta_2 = 1.0°$, which is the typical level of a well calibrated dual-rotating compensator MMP. As can be observed from Fig. 2, the estimated systematic errors show a good agreement with those "measured" ones, which demonstrates the validity of the derived systematic error propagating formula given by Eq. (11). In addition, we can also observe from Fig. 2(b) that the configuration error $\Delta\mathbf{a}$ has greater influence on the systematic errors of the extracted structural parameters than the systematic error $\boldsymbol{\mu}_{\Delta y}$ in the measured Mueller matrix elements. Therefore, we can focus on the configuration error $\Delta\mathbf{a}$ in the following experiments and use the norm of the configuration error propagating matrix $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ as the cost function to further optimize the measurement configuration for MMP.

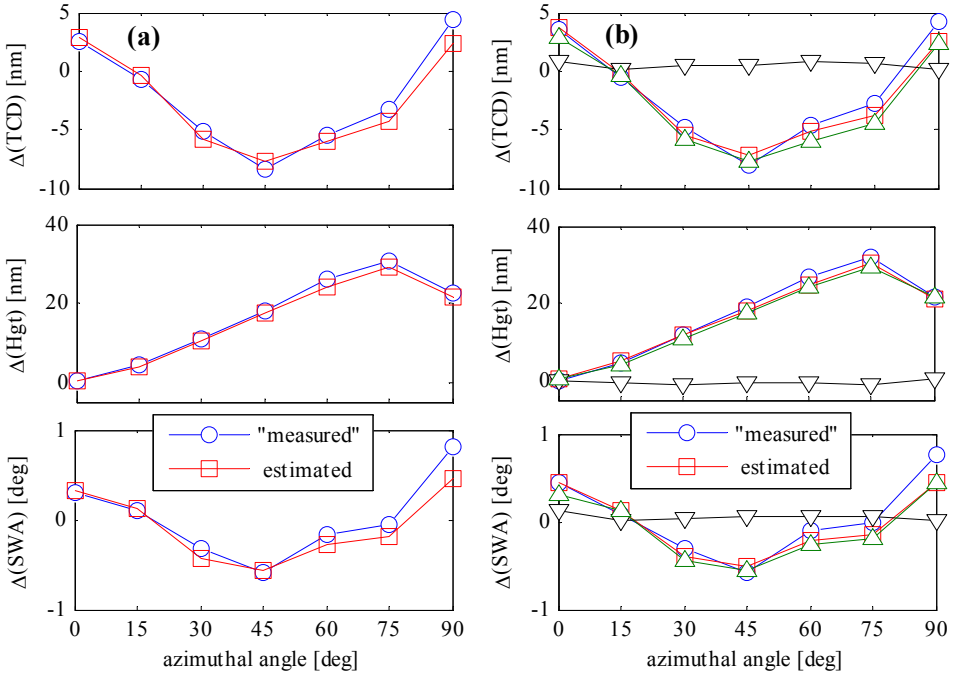

Figure 2. Comparison between the "measured" systematic errors in the extracted structural parameters *TCD*, *Hgt* and *SWA* with those estimated systematic errors by Eq. (11). The systematic errors in the extracted structural parameters are induced (a) only by the configuration error $\Delta\mathbf{a}$ and (b) by the configuration error $\Delta\mathbf{a}$ (marked with upward triangles) as well as the intrinsic systematic error $\boldsymbol{\mu}_{\Delta y}$ in the "measured" Mueller matrices (marked with downward triangles).

The l_2 norm of the configuration error propagating matrix $\left\|\tilde{\mathbf{J}}_{\mathbf{x}}^{+}\tilde{\mathbf{J}}_{\mathbf{a}}\right\|$ was calculated in a parametric domain, centered on the results measured by SEM as shown in Fig. 1(b), with *TCD* varied from 345nm to 355nm, *Hgt* from 465nm to 475nm, and *SWA* from 87° to 88°. The incidence and azimuthal angles are varied from 60° to 65° and from 0° to 90°,

respectively, both with a step size of 5°. In the given parametric domain, the maximal norms $\left\|\tilde{\mathbf{J}}_x^+\tilde{\mathbf{J}}_a\right\|$ calculated with different measurement configurations are presented in Fig. 3. As can be observed from Fig. 3, the norm of the configuration error propagating matrix $\left\|\tilde{\mathbf{J}}_x^+\tilde{\mathbf{J}}_a\right\|$ calculated with the incidence angle $\theta = 65°$ and azimuthal angle $\varphi = 70°$ is smaller than those calculated with other configurations. It is expected that the extracted structural parameters of $\theta = 65°$ and $\varphi = 70°$ will have smaller systematic errors and thus will be more accurate than other measurement configurations.

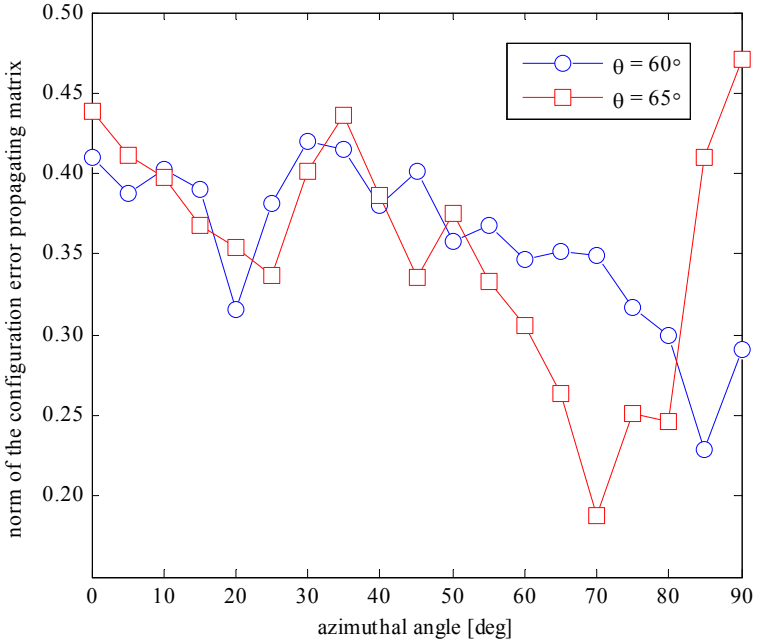

Figure 3. The l_2 norms of the configuration error propagating matrix $\left\|\tilde{\mathbf{J}}_x^+\tilde{\mathbf{J}}_a\right\|$ calculated in different measurement configurations.

In order to validate the above predicted optimal measurement configuration, the Si grating sample as shown in Fig. 1(b) was measured by the RC2 ellipsometer in incidence angles $\theta = 60°$ and $\theta = 65°$ and in azimuthal angles varied from 0° to 90° with a step size of 5°. The structural parameters of the Si grating sample extracted by the LM algorithm are compared with the results measured by SEM, and the differences between them were treated as the systematic errors in the extracted structural parameters. The l_2 norms of the systematic errors $\left\|\boldsymbol{\mu}_{\Delta x}\right\|$ in the extracted structural parameters were then calculated for the corresponding measurement configurations, as shown in Fig. 4. An examination of Fig. 4 shows that the experimental results are not in rigorous agreement with the theoretical predictions given in Fig. 3, which maybe because the relation given by Eq. (12) is not a rigorous equality but an inequality. However, we do have qualitative agreements between the theoretical predictions and the experimental results. Importantly, the optimal measurement configuration associated with the minimal norm of the systematic errors $\left\|\boldsymbol{\mu}_{\Delta x}\right\|$ in the extracted structural parameters given in Fig. 4 is in accordance with the theoretical prediction of the optimal measurement configuration given in Fig. 3. Figure 5 illustrates the fitting result of the measured and calculated Mueller matrices of the Si grating sample with the optimal measurement configuration $\theta = 65°$ and $\varphi = 70°$. The extracted structural parameters associated with Fig. 5 are $TCD = 355.1$nm, $Hgt = 467.7$nm, and $SWA = 87.9°$. As can be observed from Fig. 5, the Mueller matrices calculated with the optimal measurement configuration show a good agreement with the measured Mueller matrices. Therefore, we can conclude that the norm of the configuration error propagating matrix $\left\|\tilde{\mathbf{J}}_x^+\tilde{\mathbf{J}}_a\right\|$ can be applied as the cost function to optimize the measurement configuration for MMP to achieve higher measurement accuracy.

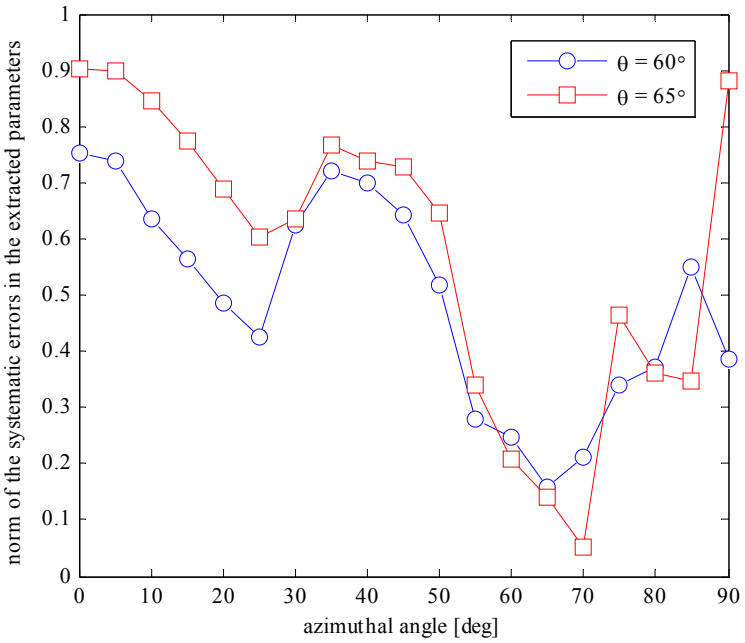

Figure 4. The l_2 norms of the systematic errors $\|\boldsymbol{\mu}_{\Delta x}\|$ in the structural parameters of the Si grating sample extracted in different measurement configurations.

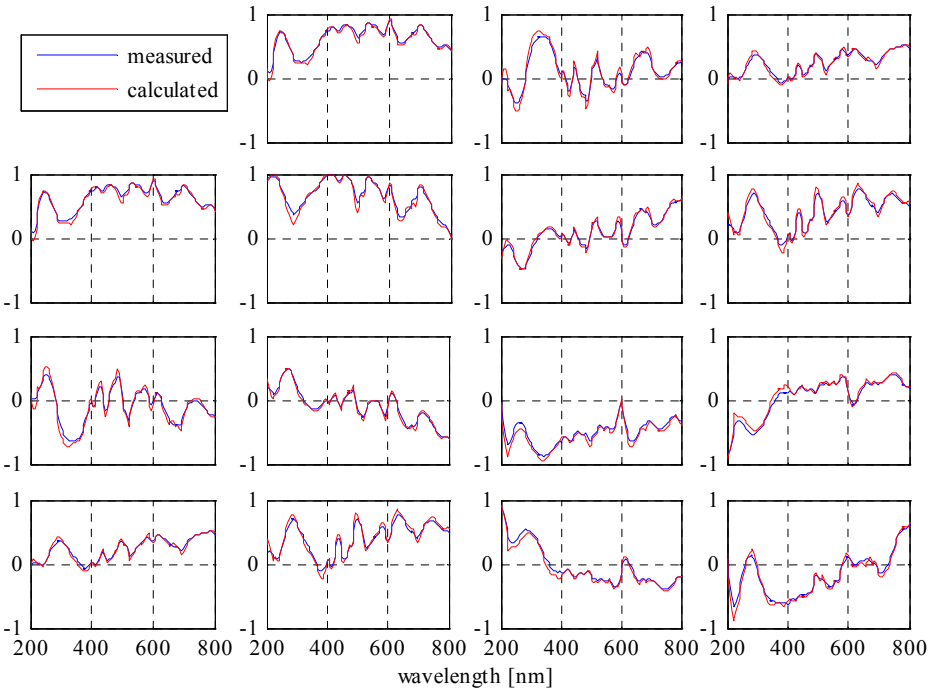

Figure 5. Fitting result of the measured and calculated Mueller matrices of the Si grating sample. The incidence and azimuthal angles are $\theta = 65°$ and $\varphi = 70°$, respectively.

4. CONCLUSIONS

The biases in the incidence and azimuthal angles, also named the configuration error in this paper, will induce systematic errors in the extracted structural parameters and thus influence the final measurement accuracy. We have derived a generalized first order error propagating formula, which reveals the general mechanism of error propagation in grating reconstruction. Based on the generalized error propagating formula, we then derived a systematic error propagating formula, which relates the systematic error propagated into the extracted structural parameter with those error sources such as the configuration error and the intrinsic systematic error in the measured Mueller matrices. According to the systematic error propagating formula, we introduced the norm of the configuration error propagating matrix to assess the influence of the biases in the incidence and azimuthal angles on the measurement accuracy. The optimal measurement configuration with higher measurement accuracy has been achieved by minimizing the norm of the configuration error propagating matrix in the available ranges of incidence and azimuthal angles. Experiments performed on a silicon grating with a dual-rotating compensator Mueller matrix polarimeter have demonstrated the validity of the proposed measurement configuration optimization method.

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).

REFERENCES

[1] B. K. Minhas, S. A. Coulombe, S. Sohail, H. Naqvi, and J. R. McNeil, "Ellipsometric scatterometry for the metrology of sub-0.1-μm-linewidth structures," Appl. Opt. **37**, 5112-5115 (1998).

[2] H. T. Huang, W. Kong, and F. L. Terry, "Normal-incidence spectroscopic ellipsometry for critical dimension monitoring," Appl. Phys. Lett. **78**, 3983-2985 (2001).

[3] X. Niu, N. Jakatdar, J. Bao, and C. J. Spanos, "Specular spectroscopic scatterometry," IEEE Trans. Semicond. Manuf. **14**, 97-111 (2001).

[4] M. G. Moharam, E. B. Grann, D. A. Pommet, and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled wave analysis of binary gratings," J. Opt. Soc. Am. A **12**, 1068-1076 (1995).

[5] L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," J. Opt. Soc. Am. A **13**, 1870-1876 (1996).

[6] S. Y. Liu, Y. Ma, X. G. Chen, and C. W. Zhang, "Estimation of the convergence order of rigorous coupled-wave analysis for binary gratings in optical critical dimension metrology," Opt. Eng. **51**, 081504 (2012).

[7] H. Gross, R. Model, M. Bär, M. Wurm, B. Bodermann, and A. Rathsfeld, "Mathematical modelling of indirect measurements in scatterometry," Measurement **39**, 782-794 (2008).

[8] J. Pomplun and F. Schmidt, "Accelerated a posterior error estimation for the reduced basis method with application to 3D electromagnetic scattering problems," SIAM J. Sci. Comput. **32**, 498-520 (2010).

[9] H. Ichikawa, "Electromagnetic analysis of diffraction gratings by the finite-difference time-domain method," J. Opt. Soc. Am. A **15**, 152-157 (1998).

[10] T. Novikova, A. De Martino, P. Bulkin, Q. Nguyen, B. Drevillon, V. Popov, and A. Chumakov, "Metrology of replicated diffractive optics with Mueller polarimetry in conical diffraction," Opt. Express **15**, 2033-2046 (2007).

[11] Y. N. Kim, J. S. Paek, S. Rabello, S. Lee, J. Hu, Z. Liu, Y. Hao, and W. McGahan, "Device based in-chip critical dimension and overlay metrology," Opt. Express **17**, 21336-21343 (2009).

[12] J. Li, J. J. Hwu, Y. Liu, S. Rabello, Z. Liu, and J. Hu, "Mueller matrix measurement of asymmetric gratings," J. Micro/Nanolith. MEMS MOEMS **9**, 041305 (2010).

[13] P. C. Logofatu, "Sensitivity analysis of grating parameter estimation," Appl. Opt. **41**, 7179-7186 (2002).

[14] P. C. Logofatu, "Phase-modulation scatterometry," Appl. Opt. **41**, 7187-7192 (2002).

[15] M. Littau, D. Forman, J. Bruce, C. J. Raymond, and S. G. Hummel, "Diffraction signature analysis methods for improving scatterometry precision," Proc. SPIE **6152**, 615236 (2006).

[16] H. Gross and A. Rathsfeld, "Sensitivity analysis for indirect measurement in scatterometry and the reconstruction of periodic grating structures," Wave Random Complex **18**, 129-149 (2008).

[17] P. Vagos, J. T. Hu, Z. Liu, and S. Rabello, "Uncertainty and sensitivity analysis and its applications in OCD measurement," Proc. SPIE **7272**, 72721N (2009).
[18] T. Novikova, A. De Martino, S. B. Hatit, and B. Drevillon, "Application of Mueller polarimetry in conical diffraction for critical dimension measurements in microelectronics," Appl. Opt. **45**, 3688-3697 (2006).
[19] M. Foldyna, A. De Martino, E. G. Caurel, R. Ossikovski, C. Licitra, F. Bertin, K. Postava, and B. Drevillon, "Critical dimension of biperiodic gratings determined by spectral ellipsometry and Mueller matrix polarimetry," Eur. Phys. J. Appl. Phys. **42**, 351-359 (2008).
[20] T. A. Germer, H. J. Patrick, R. M. Silver, and B. Bunday, "Developing uncertainty analysis for optical scatterometry," Proc. SPIE **7272**, 72720T (2009).
[21] T. M. Germer and H. J. Patrick, "Effect of bandwidth and numerical aperture in optical scatterometry," Proc. SPIE **7638**, 76381F (2010).
[22] 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**, 1997-2006 (1999).
[23] R. W. Collins, "Multichannel Ellipsometry," in *Handbook of Ellipsometry*, H. G. Tompkins and E. A. Irene, eds., (William Andrew Publishing & Springer-Verlag, 2005), Chap. 7.
[24] C. M. Herzinger, B. Johs, W. A. McGahan, J. A. Woollam, and W. Paulson, "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).
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*, (Cambridge University Press, 1985), Chap. 5.
[26] L. Li, "Symmetries of cross-polarization diffraction coefficients of gratings," J. Opt. Soc. Am. A **17**, 881-887 (2000).
[27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., (Cambridge University Press, 2007), Chap. 15.