Thin Solid Films 571 (2014) 653–659

Contents lists available at ScienceDirect

Thin Solid Films

journal homepage: www.elsevier.com/locate/tsf

Formulation of error propagation and estimation in grating reconstruction by a dual-rotating compensator Mueller matrix polarimeter

Xiuguo Chen [a], Shiyuan Liu [a,b,*], Honggang Gu [a], Chuanwei Zhang [b]

[a] Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China
[b] State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

ARTICLE INFO

Article history:
Received 2 June 2013
Received in revised form 17 December 2013
Accepted 10 January 2014
Available online 24 January 2014

Keywords:
Mueller matrix polarimetry
Grating reconstruction
Inverse problem
Error analysis

ABSTRACT

Recently, the Mueller matrix polarimeter (MMP) has been introduced for critical dimension and overlay metrology. In practice, the measurement process invariably has errors. These errors, which can be generally categorized into random and systematic errors, have great influences on the final precision and accuracy of the extracted structural parameters. In this paper, we present detailed formulations for the propagation and estimation of random and systematic errors in grating reconstruction using a dual-rotating compensator MMP (DRC-MMP). We derive a generalized first-order error propagating formula, which reveals the mechanism of error propagation in grating reconstruction using the DRC-MMP. According to this first-order error propagating formula and the measurement principle of the DRC-MMP, we then derive detailed formulations for the estimation of random and systematic errors that are propagated into the final extracted structural parameters. Simulations performed on silicon grating samples have demonstrated the validity of the present theoretical derivations.

1. Introduction

Process control in microelectronic manufacturing requires real-time monitoring techniques. Among the different techniques, optical scatterometry, also referred to as optical critical dimension (OCD) metrology, has been widely used for critical dimension (CD) monitoring [1–3]. The grating reconstruction process in optical scatterometry is essentially an inverse diffraction problem solving process with the objective of finding a profile whose theoretical signature can best match the measured one. Here, the general term *signature* contains the scattered light information from the grating structure, which can be in the form of reflectance, ellipsometric angles, Stokes vector elements, or Mueller matrix elements. In practice, the grating reconstruction process invariably has errors. These errors, which can be generally categorized into random and systematic errors, have great influences on the final precision and accuracy of the extracted structural parameters. In order to estimate how scatterometry will behave for nanostructures or to guide the design of future scatterometers, theoretical analysis of error propagation in grating reconstruction is of great importance for both scatterometer manufacturers and users. Error and uncertainty analysis for conventional ellipsometric scatterometry has been studied in recent years. Al-Assaad and Byrne investigated the propagation of different types of errors from the scatterometric data to the extracted structural parameters [4]. 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 [5,6]. Silver et al. introduced an approach for embedding reference metrology results directly in the uncertainty analysis and library-fitting process to reduce parametric uncertainties [7]. Vagos et al. developed an uncertainty & sensitivity analysis (U&SA) package for measurement optimization [8].

Recently, the Mueller matrix polarimeter (MMP), developed based on the coupled ferroelectric liquid crystal cell [9,10] or dual-rotating compensator [11–13] technique, has been introduced for CD and overlay metrology [14–16] and has demonstrated great potential in semiconductor manufacturing. Compared with the conventional ellipsometric scatterometer, the design and calibration of the MMP are much more complicated. There will be more error sources in grating reconstruction using the MMP, and therefore the corresponding error analysis will be much more challenging. In the past decades, error analysis for this instrument has been discussed in literature. Hauge analyzed errors that were due to diattenuation in the retardation elements [17]. Goldstein and Chipman discussed small errors in orientational alignment of three of the four polarization elements [18], and Chenault et al. extended this method to larger errors [19]. Twietmeyer and Chipman performed error analysis based on the covariance matrix method with the aim of optimizing the polarimeter design in the presence of error sources [20]. Broch et al. investigated systematic errors caused by the inaccurate

* Corresponding author at: Head of Nanoscale and Optical Metrology Group, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan, Hubei 430074, China. Tel.: +86 27 8755 9543; fax: +86 27 8755 8045.
E-mail address: shyliu@mail.hust.edu.cn (S. Liu).

0040-6090/$ – see front matter
http://dx.doi.org/10.1016/j.tsf.2014.01.049

azimuthal arrangement of optical components and the residual ellipticity introduced by imperfect optical components [21,22]. It is interesting to note that most of these analyses are focused on systematic errors in the Mueller matrix measurement. Although it is important from the point of view of instrument design, it is still desirable to quantify the relation between all of the possible error sources in measurement and the final solution of the inverse diffraction problem in order to further clarify the influences of different error sources.

The objective of this paper is to provide detailed formulations for the propagation and estimation of random and systematic errors in grating reconstruction by a dual-rotating compensator MMP (DRC-MMP). The primary features of the DRC-MMP are the mechanically rotating wave plates, which may cause vibration and beam wander and ultimately limit the maximum data acquisition speed. Despite these limitations, the DRC configuration based MMP still remains popular up to now, likely because it is inexpensive and straightforward to align and calibrate. In this paper, we analyze the possible error sources in grating reconstruction using the DRC-MMP. Then we derive a generalized first-order error propagating formula, which reveals the mechanism of error propagation in the grating reconstruction procedure. According to the first-order error propagating formula and the measurement principle of the DRC-MMP, we then derive detailed formulations for the estimation of random and systematic errors that are propagated into the final extracted structural parameters. Based on these analyses, we build the relation between the solution of the inverse diffraction problem and the error sources in both the DRC-MMP and the optical model. It is expected that our theoretical analyses will be used either to assess the measurement precision and accuracy for developing DRC-MMPs in CD and overlay metrology or to optimize the DRC-MMP design to make it satisfy the requirements of an advanced technology node.

The remainder of this paper is organized as follows. Section 2 first introduces error sources in grating reconstruction using the DRC-MMP. Next, the formalism of error propagation and estimation in grating reconstruction using the DRC-MMP is presented in Section 3. Section 4 provides simulations performed on two silicon grating samples to demonstrate the theoretical derivations of error propagation and estimation. Finally, we draw some conclusions in Section 5.

2. Errors in grating reconstruction using a DRC-MMP

Fig. 1 depicts the grating reconstruction procedure and the errors in this procedure using a DRC-MMP. In this procedure, an optical model corresponding to the grating sample is constructed in the beginning. Without loss of generality, we denote the structural parameters under measurement as an M-dimensional vector $\mathbf{x} = [x_1, x_2, \cdots, x_M]^{\mathrm{T}}$, where the superscript "T" represents the transpose. The vector $\mathbf{a} = [a_1, a_2, \cdots, a_L]^{\mathrm{T}}$ consists of all of the other L fixed parameters that are input into the optical model. The χ^2 function is usually applied to estimate the fitting errors between the calculated and measured Mueller matrix spectra, which is defined in a matrix form as

$$\chi^2 = \left[\mathbf{M}^{\mathrm{meas}} - \mathbf{M}^{\mathrm{calc}}(\mathbf{x},\ \mathbf{a})\right]^T \mathbf{W} \left[\mathbf{M}^{\mathrm{meas}} - \mathbf{M}^{\mathrm{calc}}(\mathbf{x},\ \mathbf{a})\right], \quad (1)$$

where $\mathbf{M}^{\mathrm{meas}}$ and $\mathbf{M}^{\mathrm{calc}}(\mathbf{x},\mathbf{a})$ are the N-dimensional column vectors that consist of the measured and calculated Mueller matrix elements, respectively. $\mathbf{W}$ is an $N \times N$ positive definite weighting matrix, which is usually chosen to be the inverse of the covariance matrix of the measured Mueller matrix elements. The optimal estimation of the structural parameters $\hat{\mathbf{x}}$ can be achieved by solving

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\Omega}{\arg\min} \left[\mathbf{M}^{\mathrm{meas}} - \mathbf{M}^{\mathrm{calc}}(\mathbf{x},\ \mathbf{a}^*)\right]^T \mathbf{W} \left[\mathbf{M}^{\mathrm{meas}} - \mathbf{M}^{\mathrm{calc}}(\mathbf{x},\ \mathbf{a}^*)\right], \quad (2)$$

where $\mathbf{a}^*$ denotes the priori value of $\mathbf{a}$ in grating reconstruction, and Ω is the parameter domain.

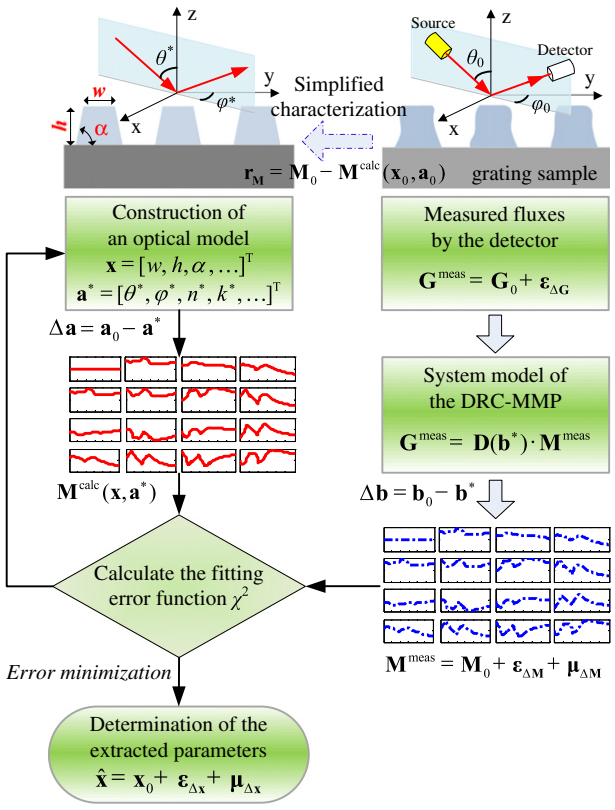

Fig. 1. The grating reconstruction procedure and the errors in this procedure using a dual-rotating compensator Mueller matrix polarimeter (DRC-MMP).

Generally, two types of error can be distinguished from the final extracted structural parameters $\hat{\mathbf{x}}$, i.e., random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$ and systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$. The former represents the amount of random uncertainties in the extracted parameters while the latter represents the deterministic difference between the true and the extracted parameter values. The extracted structural parameters $\hat{\mathbf{x}}$ will be the sum of the true parameter value $\mathbf{x}_0$, the random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$, and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$, i.e., $\hat{\mathbf{x}} = \mathbf{x}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{x}} + \boldsymbol{\mu}_{\Delta\mathbf{x}}$. The random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$ in $\hat{\mathbf{x}}$ arises from the random noise in measurements using the DRC-MMP, whereas the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ results from the deterministic offsets both in the DRC-MMP and in the applied optical model during grating reconstruction.

2.1. Errors in the optical model

The systematic errors in the optical model typically arise from the bias $\Delta\mathbf{a}$ in vector $\mathbf{a}$ and the simplified characterization of the grating sample under measurement. The variables in vector $\mathbf{a}$ can be any parameters that are fixed in grating reconstruction, such as the incidence and azimuthal angles θ and φ, the refractive indices n and k, or even the structural parameters. The bias $\Delta\mathbf{a}$ describes the difference between the priori value $\mathbf{a}^*$ and the true value $\mathbf{a}_0$ of vector $\mathbf{a}$, i.e., $\Delta\mathbf{a} = \mathbf{a}_0 - \mathbf{a}^*$. In addition, a priori knowledge of the grating sample, such as its periodicity and its rectangular or trapezoidal-like profile shape, is usually introduced to calculate its optical signature in practice. However, this knowledge is sometimes partial or inaccurate. For example, the actual grating sample may have round corners and/or line edge roughness, which are sometimes ignored in practical optical modeling. The systematic errors in the optical model that are induced by the simplified characterization of the grating sample can be quantified by the residual signature $\mathbf{r}_{\mathbf{M}} = \mathbf{M}_0 - \mathbf{M}^{calc}(\mathbf{x}_0,\mathbf{a}_0)$, where $\mathbf{M}_0$ denotes the true signature of the grating sample and $\mathbf{M}^{calc}(\mathbf{x}_0,\mathbf{a}_0)$ denotes the calculated signature associated with the true values $\mathbf{x}_0$ and $\mathbf{a}_0$.

2.2. Errors in the DRC-MMP

As schematically shown in Fig. 2, the system configuration of the DRC-MMP 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 [11–13]. The system model of the DRC-MMP can be formulated as [23,24]

$$\mathbf{G}^{\text{meas}} = \mathbf{D}(\mathbf{b}) \cdot \mathbf{M}^{\text{meas}}, \tag{3}$$

where $\mathbf{G}^{\text{meas}}$ is a K-element flux vector measured by the detector. $\mathbf{D}$ is a $K \times 16$ polarimetric matrix, which is a function of a system-dependent vector $\mathbf{b}$. The variables in $\mathbf{b}$ can be the transmission axis angles of the polarizer P and analyzer A, the fast axis angles C_{S1} and C_{S2}, the phase retardances δ_1 and δ_2 of the two compensators etc., i.e., $\mathbf{b} = [P, A, C_{S1}, C_{S2}, \delta_1, \delta_2, \ldots]^{\text{T}}$, which are typically determined from a calibration process.

The measured flux vector $\mathbf{G}^{\text{meas}}$ will be the sum of the true signal (noise-free signal) $\mathbf{G}_0$ and a random vector $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ representing measurement noise in $\mathbf{G}^{\text{meas}}$, i.e., $\mathbf{G}^{\text{meas}} = \mathbf{G}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$. We assume that the random vector $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ has a zero mean and any offset in the measured Mueller matrix $\mathbf{M}^{\text{meas}}$ is induced by the bias $\Delta\mathbf{b}$ in vector $\mathbf{b}$, which describes the difference between the calibrated value $\mathbf{b}^*$ and the true value $\mathbf{b}_0$ of vector $\mathbf{b}$, i.e., $\Delta\mathbf{b} = \mathbf{b}_0 - \mathbf{b}^*$. Therefore, the measured Mueller matrix $\mathbf{M}^{\text{meas}}$ determined by Eq. (3) will be the sum of the true signature of the grating sample $\mathbf{M}_0$, the random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$, and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{M}}$, i.e., $\mathbf{M}^{\text{meas}} = \mathbf{M}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}} + \boldsymbol{\mu}_{\Delta\mathbf{M}}$. The random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ in $\mathbf{M}^{\text{meas}}$ is rooted in the random noise $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ in $\mathbf{G}^{\text{meas}}$, whereas the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{M}}$ arises from the bias $\Delta\mathbf{b}$ in vector $\mathbf{b}$. As depicted in Fig. 1, the random and systematic errors $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ and $\boldsymbol{\mu}_{\Delta\mathbf{M}}$ in the measured Mueller matrix $\mathbf{M}^{\text{meas}}$ together with the systematic errors in the calculated Mueller matrix $\mathbf{M}^{\text{calc}}(\mathbf{x},\mathbf{a}^*)$ will finally propagate into the extracted structural parameters $\hat{\mathbf{x}}$ when solving Eq. (2).

3. Error propagation and estimation in grating reconstruction using a DRC-MMP

3.1. Generalized formulation for error propagation

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

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

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

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

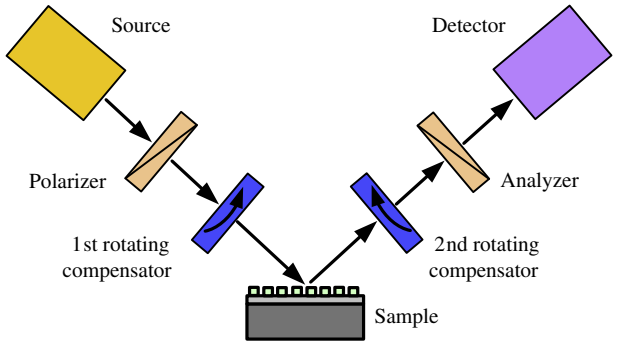

Fig. 2. Basic scheme of the dual-rotating compensator Mueller matrix polarimeter.

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

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

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

where Δx denotes the error (including random and systematic errors) that is propagated into the extracted parameters $\hat{x}$, and can be expressed by $\Delta\mathbf{x} = \mathbf{x}_0 - \hat{\mathbf{x}} = -\boldsymbol{\mu}_{\Delta\mathbf{x}} - \boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$.

Inserting Eq. (6) into Eq. (1) and noticing that $\mathbf{M}^{\text{calc}}(\mathbf{x}_0,\mathbf{a}_0) = \mathbf{M}_0 - \mathbf{r}_{\mathbf{M}}$ and $\mathbf{M}^{\text{meas}} = \mathbf{M}_0 + \boldsymbol{\mu}_{\Delta\mathbf{M}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ in the meanwhile, we have

$$\begin{aligned}\chi^2_{\min} &= \left[\mathbf{M}^{\text{meas}} - \mathbf{M}^{\text{calc}}(\hat{\mathbf{x}},\ \mathbf{a}^*)\right]^T \mathbf{W} [\mathbf{M}^{\text{meas}} - \mathbf{M}^{\text{calc}}(\hat{\mathbf{x}},\ \mathbf{a}^*)] \\ &= [\mathbf{J}_{\mathbf{x}}\Delta\mathbf{x} + \mathbf{J}_{\mathbf{a}}\Delta\mathbf{a} + \mathbf{r}_{\mathbf{M}} + \boldsymbol{\mu}_{\Delta\mathbf{M}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}}]^T \mathbf{W} [\mathbf{J}_{\mathbf{x}}\Delta\mathbf{x} + \mathbf{J}_{\mathbf{a}}\Delta\mathbf{a} + \mathbf{r}_{\mathbf{M}} + \boldsymbol{\mu}_{\Delta\mathbf{M}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}}].\end{aligned} \tag{7}$$

By taking the derivatives of both sides of Eq. (7) with respect to each element of $\mathbf{x}$, we obtain

$$\tilde{\mathbf{J}}_{\mathbf{x}}\Delta\mathbf{x} + \tilde{\mathbf{J}}_{\mathbf{a}}\Delta\mathbf{a} + \tilde{\mathbf{r}}_{\mathbf{M}} + \tilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}} + \tilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}} = 0, \tag{8}$$

where $\tilde{\mathbf{J}}_{\mathbf{x}} = \mathbf{W}^{1/2}\mathbf{J}_{\mathbf{x}}, \tilde{\mathbf{J}}_{\mathbf{a}} = \mathbf{W}^{1/2}\mathbf{J}_{\mathbf{a}}, \tilde{\mathbf{r}}_{\mathbf{M}} = \mathbf{W}^{1/2}\mathbf{r}_{\mathbf{M}}, \tilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}} = \mathbf{W}^{1/2}\boldsymbol{\mu}_{\Delta\mathbf{M}}$, and $\tilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}} = \mathbf{W}^{1/2}\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$. The detailed derivation of Eq. (8) is given in Appendix A. We call Eq. (8) the generalized *first-order error propagating formula*, which relates the error $\Delta\mathbf{x}$ in the extracted parameters $\hat{\mathbf{x}}$ with those error sources, such as the bias $\Delta\mathbf{a}$ in vector $\mathbf{a}$, the residual signature $\mathbf{r}_{\mathbf{M}}$ induced by the optical model, and the random and systematic errors $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ and $\boldsymbol{\mu}_{\Delta\mathbf{M}}$ in the measured Mueller matrix $\mathbf{M}^{\text{meas}}$. According to Eq. (8), we can further estimate the random and systematic errors $\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$ and $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ that are propagated into $\hat{\mathbf{x}}$.

3.2. Estimation of random errors

Random errors are typically described by using their variances or standard deviations. A signal-dependent noise model has been proposed to estimate random noise introduced by the optical imaging systems [25–27]. Assuming that the random noise $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ in the measured fluxes $\mathbf{G}^{\text{meas}}$ is the additive Gaussian white noise, and that the pixels of the detector are homogeneous, we extend the signal-dependent noise model to estimate the variances of $\mathbf{G}^{\text{meas}}$ by

$$\sigma^2(g_k) = \varepsilon_0 + \varepsilon_1 g_k + \varepsilon_2 g_k^2 + \varepsilon_3 g_k^3, \tag{9}$$

where g_k ($k = 1, 2, \ldots, K$) is the k-th measured flux of $\mathbf{G}^{\text{meas}}$. ε_i ($i = 0, 1, 2, 3$) are the coefficients of the noise model, which are determined by the performances of the instrument. The physical meanings of the four coefficients can be interpreted as follows: ε_0 is associated with the dark current noise, thermal noise, and background noise, ε_1 is associated with the shot noise, ε_2 is associated with the light source intensity noise and modulation noise, and ε_3 is associated with the low frequency noise. In a well-developed DRC-MMP, the second term of Eq. (9) will be predominant, while other terms should have small quantities.

According to Eq. (3), we derive the covariance matrix of the measured Mueller matrix $\mathbf{M}^{\text{meas}}$ resulting from the covariance matrix of the measured fluxes $\mathbf{G}^{\text{meas}}$ that

$$\mathbf{C}(\mathbf{M}^{\text{meas}}) = \mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}) = \mathbf{D}^+(\mathbf{b}^*) \cdot \mathbf{C}(\mathbf{G}^{\text{meas}}) \cdot \left[\mathbf{D}^+(\mathbf{b}^*)\right]^T, \tag{10}$$

where $\mathbf{D}^+(\mathbf{b}^*) = [\mathbf{D}(\mathbf{b}^*)^{\text{T}}\mathbf{D}(\mathbf{b}^*)]^{-1}[\mathbf{D}(\mathbf{b}^*)]^{\text{T}}$ is the Moore–Penrose pseudo-inverse of matrix $\mathbf{D}(\mathbf{b}^*)$. Assuming that the measured fluxes g_k ($k = 1, 2, \ldots, K$) are uncorrelated and noticing that the random vector

$\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ has a zero mean, the covariance matrix $\mathbf{C}(\mathbf{G}^{\text{meas}})$ can be expressed as a diagonal matrix with the diagonal elements $\sigma^2(g_k)$, i.e.,

$$\mathbf{C}(\mathbf{G}^{\text{meas}}) = \mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}) = \left\langle \boldsymbol{\varepsilon}_{\Delta\mathbf{G}} \boldsymbol{\varepsilon}_{\Delta\mathbf{G}}^T \right\rangle = \text{diag}\left(\sigma^2(g_1),\ \sigma^2(g_2),\ \ldots,\ \sigma^2(g_K)\right). \tag{11}$$

According to Eqs. (10) and (11), we derive the variances of the Mueller matrix elements m_{ij} that

$$\sigma^2\left(m_{ij}\right) = \sum_{k=1}^{K} \left(d_{rk}^{+}\right)^2 \sigma^2(g_k), \tag{12}$$

where d_{rk}^{+} is the (r, k)-th element of the matrix $\mathbf{D}^{+}(\mathbf{b}^{*})$. The index r corresponds to the serial number of m_{ij} in $\mathbf{M}^{\text{meas}}$.

The uncertainties in the measured Mueller matrix will then propagate into the extracted structural parameters $\hat{\mathbf{x}}$ in the process of solving Eq. (2). According to Eq. (8), we further derive the covariance matrix of $\hat{\mathbf{x}}$ that

$$\mathbf{C}(\hat{\mathbf{x}}) = \mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}) = \widetilde{\mathbf{J}}_{\mathbf{x}}^{+} \cdot \mathbf{C}(\widetilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}}) \cdot \left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right)^T = \widetilde{\mathbf{J}}_{\mathbf{x}}^{+} \cdot \mathbf{W}^{1/2} \cdot \mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}) \cdot \mathrm{W}^{1/2} \cdot \left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right)^T, \tag{13}$$

where $\widetilde{\mathbf{J}}_{\mathbf{x}}^{+} = \left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}\widetilde{\mathbf{J}}_{\mathbf{x}}\right)^{-1}\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}$ is the Moore–Penrose pseudo-inverse of matrix $\widetilde{\mathbf{J}}_{\mathbf{x}}$. If the weighting matrix $\mathbf{W}$ is chosen to be $\mathbf{W} = [\mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{M}})]^{-1}$, then Eq. (13) can be further simplified to

$$\mathbf{C}(\hat{\mathbf{x}}) = \mathbf{C}(\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}) = \widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right)^T = \left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}\widetilde{\mathbf{J}}_{\mathbf{x}}\right)^{-1}\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}\widetilde{\mathbf{J}}_{\mathbf{x}}\left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}\widetilde{\mathbf{J}}_{\mathbf{x}}\right)^{-1} = \left(\widetilde{\mathbf{J}}_{\mathbf{x}}^{T}\widetilde{\mathbf{J}}_{\mathbf{x}}\right)^{-1}. \tag{14}$$

In this case, the χ^2 function as defined in Eq. (1) will be the Mahalanobis distance [28]. The standard deviation of parameter x_i $(i = 1, 2, \ldots, M)$ can be estimated from the diagonal element of $\mathbf{C}(\hat{\mathbf{x}})$ and is given by

$$\sigma(x_i) = \sqrt{[\mathbf{C}(\hat{\mathbf{x}})]_{ii}}. \tag{15}$$

According to Eq. (15), the estimated uncertainty in parameter x_i with a desired confidence level is given by

$$u(x_i) = \kappa\sigma(x_i), \tag{16}$$

where κ is the coverage factor associated with the prescribed confidence level.

3.3. Estimation of systematic errors

The systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in the extracted structural parameters $\hat{\mathbf{x}}$ arises from the deterministic offset in the DRC-MMP as well as in the applied optical model during grating reconstruction. Taking the mean values of both sides of Eq. (8) and noticing that the random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ in $\mathbf{M}^{\text{meas}}$ has a zero mean, we obtain the following equation

$$\langle \Delta\mathbf{x} \rangle = -\boldsymbol{\mu}_{\Delta\mathbf{x}} = -\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\widetilde{\mathbf{J}}_{\mathbf{a}}\Delta\mathbf{a} - \widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\widetilde{\mathbf{r}}_{\mathbf{M}} - \widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\widetilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}}. \tag{17}$$

We call Eq. (17) as the *systematic error propagating formula*, which describes how the bias $\Delta\mathbf{a}$ in vector $\mathbf{a}$, the residual signature $\mathbf{r}_{\mathbf{M}}$ induced by the optical model, and the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{M}}$ in the measured Mueller matrix $\mathbf{M}^{\text{meas}}$ lead to the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in $\hat{\mathbf{x}}$. According to Eq. (3), a generalized formula is derived to estimate the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{M}}$ in $\mathbf{M}^{\text{meas}}$ induced by the bias $\Delta\mathbf{b}$ in vector $\mathbf{b}$, which is given by

$$\boldsymbol{\mu}_{\Delta\mathbf{M}} = \left[\mathbf{D}(\mathbf{b}^{*}) + \Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}}\right]^{+} \cdot \left(\Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}}\right) \cdot \langle \mathbf{M}^{\text{meas}} \rangle, \tag{18}$$

where $\left[\mathbf{D}(\mathbf{b}^{*}) + \Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}}\right]^{+}$ is the Moore–Penrose pseudo-inverse of matrix $\left[\mathbf{D}(\mathbf{b}^{*}) + \Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}}\right]$, and the term $\Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}}$ is defined by

$$\Delta\mathbf{b}^T \cdot \frac{\partial\mathbf{D}}{\partial\mathbf{b}} = \begin{bmatrix} \Delta\mathbf{b}^T \cdot \frac{\partial d_{1,1}}{\partial\mathbf{b}} & \Delta\mathbf{b}^T \cdot \frac{\partial d_{1,2}}{\partial\mathbf{b}} & \cdots & \Delta\mathbf{b}^T \cdot \frac{\partial d_{1,16}}{\partial\mathbf{b}} \\ \Delta\mathbf{b}^T \cdot \frac{\partial d_{2,1}}{\partial\mathbf{b}} & \Delta\mathbf{b}^T \cdot \frac{\partial d_{2,2}}{\partial\mathbf{b}} & \cdots & \Delta\mathbf{b}^T \cdot \frac{\partial d_{2,16}}{\partial\mathbf{b}} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta\mathbf{b}^T \cdot \frac{\partial d_{K,1}}{\partial\mathbf{b}} & \Delta\mathbf{b}^T \cdot \frac{\partial d_{K,2}}{\partial\mathbf{b}} & \cdots & \Delta\mathbf{b}^T \cdot \frac{\partial d_{K,16}}{\partial\mathbf{b}} \end{bmatrix}, \tag{19}$$

where $d_{i,j}$ is the (i, j)-th element of $\mathbf{D}(\mathbf{b})$. The partial derivatives in Eq. (19) are calculated at $\mathbf{b} = \mathbf{b}^{*}$. The detailed derivation of Eq. (18) is given in Appendix B.

Eqs. (17–19) relate the systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in the extracted structural parameters $\hat{\mathbf{x}}$ with the error sources in the optical model and DRC-MMP. According to Eqs. (17–19), we can identify which error sources will contribute most in limiting the accuracy of grating reconstruction. Consequently, useful strategies for controlling and improving the final measurement accuracy can be developed. In addition, according to Eq. (17), we can further derive the following inequality that

$$\left\|\boldsymbol{\mu}_{\Delta\mathbf{x}}\right\| \le \left\|\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\widetilde{\mathbf{J}}_{\mathbf{a}}\right\| \cdot \|\Delta\mathbf{a}\| + \left\|\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right\| \cdot \|\widetilde{\mathbf{r}}_{\mathbf{M}}\| + \left\|\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right\| \cdot \|\widetilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}}\|, \tag{20}$$

where the notation $\|\cdot\|$ denotes the l_p $(p = 1, 2, \infty)$ vector norm and the l_p matrix norm that is induced by the associated vector norm [29]. Eq. (20) gives the upper bound of the systematic errors that are propagated into $\hat{\mathbf{x}}$. $\left\|\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\widetilde{\mathbf{J}}_{\mathbf{a}}\right\|$ and $\left\|\widetilde{\mathbf{J}}_{\mathbf{x}}^{+}\right\|$ represent the maximum gain factors in the propagation of different error sources, which can be used as objective functions to optimize the measurement configuration of the MMP for more accurate grating reconstruction [30]. In order to make the above description much clearer, the important symbols adopted in the theoretical derivations are listed in Appendix C.

4. Results

To verify the theoretical derivations of the error propagation and estimation in grating reconstruction, simulations were performed on two silicon grating samples. Fig. 3 depicts cross-section images of the investigated silicon gratings. Sample (a) has a perfect trapezoidal profile and is characterized by top critical dimension *TCD*, sidewall angle *SWA*, line height *Hgt*, and period *pitch*. Sample (b) also has a trapezoidal-like profile but with top and bottom round corners R_1 and R_2 and is used to investigate the systematic errors induced by simplified characterization of the sample. Nominal dimensions of the two grating samples are: $TCD = 45$ nm, $Hgt = 100$ nm, $SWA = 82°$, and $pitch = 90$ nm. Top and bottom round corners of sample (b) are given by $R_1 = R_2 = 5$ nm. In the following simulations, structural parameters of the silicon grating samples that need to be extracted include *TCD*, *Hgt* and *SWA*, while others are fixed at their nominal values. Optical properties of silicon are taken from Ref. [31]. Parameter settings of the DRC-MMP are fixed at $P = A = 45°$, $C_{S1} = C_{S2} = 0°$, and $\delta_1 = \delta_2 = 90°$ (correspond to a quarter waveplate). Coefficients of the parametric noise model as described in Eq. (9) are taken as $\varepsilon_0 = 12.13$, $\varepsilon_1 = 4.31\text{e}-2$, $\varepsilon_2 = 6.89\text{e}-7$, and $\varepsilon_3 = 2.73\text{e}-11$, which is a group of calibration values determined by the performance of our instrument. The spectral range is varied from 200 to 800 nm with an increment of 5 nm.

Since the following comparisons are based on the "measured" Mueller matrix, it is necessary to describe how to generate it before presenting the comparison results. The "measured" Mueller matrix is defined as the theoretical sample Mueller matrix added with simulated errors, and is used to imitate the actually measured sample Mueller matrix in the process of grating reconstruction. To generate the "measured" Mueller matrix, we first calculate the theoretical sample Mueller matrix by using rigorous coupled-wave analysis [32–34]. The theoretical light

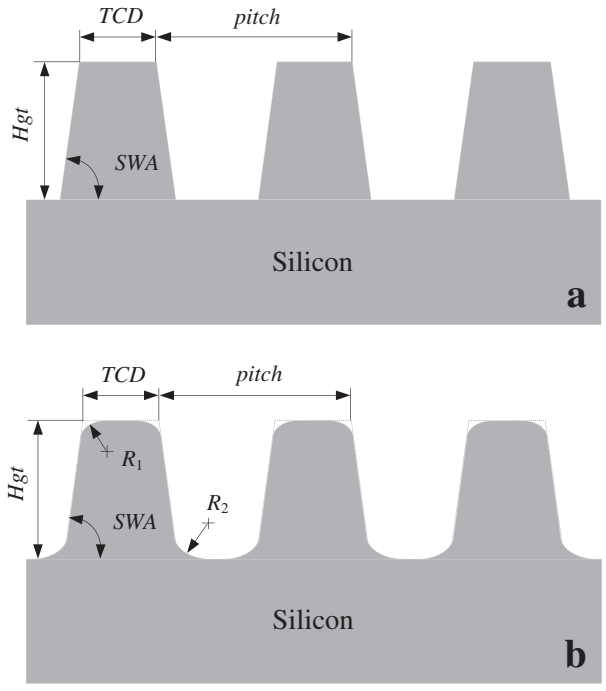

Fig. 3. Cross-section images of the investigated silicon gratings.

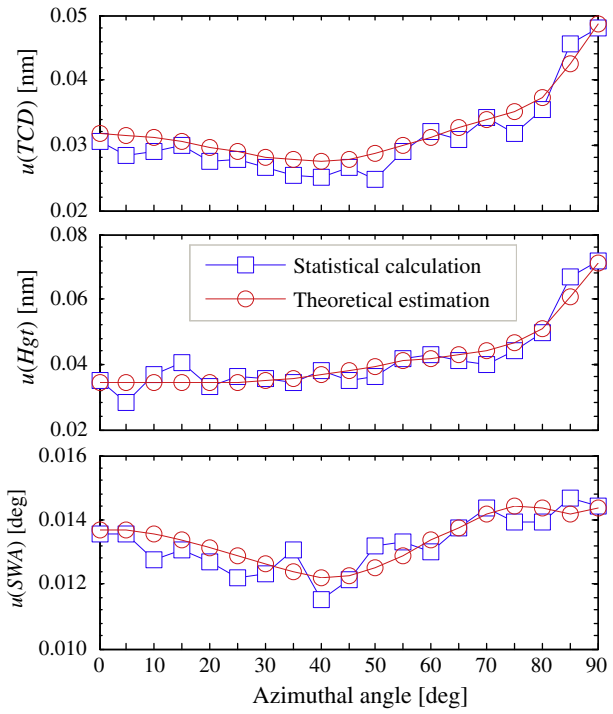

Fig. 4. Comparison between the statistically calculated uncertainties in the extracted structural parameters *TCD*, *Hgt* and *SWA* with those theoretically estimated by Eq. (16). In the calculations, the incidence angle is fixed at 65° and the azimuthal angles are varied from 0° and 90° with an increment of 5°.

fluxes g_k ($k = 1, 2, \ldots, K$) can be calculated by performing Fourier analysis on the light intensity waveform obtained from the first element of the emerging Stokes vector. Then we add random noise into the theoretical fluxes to imitate the actually measured fluxes $\mathbf{G}^{meas}$ given in Eq. (3), where the random noise is generated using a zero mean Gaussian distribution model with the signal-dependent variance $\sigma^2(g_k)$ calculated by Eq. (9). We further add biases into 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 to make the system-dependent vector $\mathbf{b}$ in Eq. (3) contain calibration errors. According to Eq. (3), we can obtain the "measured" sample Mueller matrix $\mathbf{M}^{meas}$. By continuously varying the wavelengths, we will further obtain the "measured" Mueller matrix spectrum. The systematic errors induced by the optical model, such as simplified characterization of the grating sample and biases in the incidence and azimuthal angles θ and φ, can be simulated by adding the corresponding errors into the inverse diffraction problem solving process. In the simulations, the solution of Eq. (2) is solved by using the Levenberg–Marquardt algorithm [35], which typically converges rapidly to the global minimum if suitable initial values are provided.

The uncertainties in the extracted structural parameters can be estimated using Eq. (16). In order to examine the validity of the estimated uncertainties, a group of "measured" Mueller matrix spectra are repeatedly generated first for sample (a). We then extract the structural parameters from the generated "measured" Mueller matrix spectra. Based on the extracted structural parameters, we can directly calculate the uncertainties in the extracted structural parameters. We then compare the above statistically calculated uncertainties with those theoretically estimated by Eq. (16). In the calculations, the uncertainties in the extracted structural parameters are estimated with a 95% confidence level, i.e., $\kappa = 1.96$. The incidence angle is fixed at 65°, and the azimuthal angles are varied from 0° to 90° with an increment of 5°. Fig. 4 depicts the comparison between the statistically calculated uncertainties in the extracted structural parameters *TCD*, *Hgt* and *SWA* and those theoretically estimated by Eq. (16). As observed from Fig. 4, the theoretically estimated uncertainties show good agreement with those statistically calculated ones for all the combinations of incidence and azimuthal angles. Minor differences in the results are possibly attributed to numerical errors. It therefore demonstrates the validity of the derived uncertainty estimation formula given in Eq. (16).

The systematic error $\boldsymbol{\mu}_{\Delta\mathbf{x}}$ in the extracted structural parameters $\hat{\mathbf{x}}$ can be estimated using Eqs. (17–19). In the simulations, we first examine the validity of the estimated systematic error in $\hat{\mathbf{x}}$ induced by the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ for sample (a). Then, we further examine the validity of the estimated systematic error in $\hat{\mathbf{x}}$ induced by the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ as well as the residual signature $\mathbf{r}_\mathbf{M}$ for sample (b), where the residual signature $\mathbf{r}_\mathbf{M}$ is induced by ignoring the top and bottom round corners and characterizing it with a perfect trapezoidal profile. To this end, we first generate the "measured" Mueller matrix spectrum for the structural parameters $\mathbf{x}_0$ and fixed values $\mathbf{a}_0 = \mathbf{a}^* + \Delta\mathbf{a}$ and $\mathbf{b}_0 = \mathbf{b}^* + \Delta\mathbf{b}$. Then, we extract the structural parameters from the "measured" Mueller matrix spectrum with the given values $\mathbf{a}^*$ and $\mathbf{b}^*$. It is certain that there will exist errors $\Delta\mathbf{x}$ (including the random and systematic errors $\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$ and $\boldsymbol{\mu}_{\Delta\mathbf{x}}$) in $\hat{\mathbf{x}}$, i.e., $\Delta\mathbf{x} = \mathbf{x}_0 - \hat{\mathbf{x}} = -\boldsymbol{\mu}_{\Delta\mathbf{x}} - \boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$. 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 will be the statistically calculated systematic error in $\hat{\mathbf{x}}$. Fig. 5(a) depicts the comparison between the theoretically estimated and statistically calculated systematic errors in *TCD*, *Hgt* and *SWA* of sample (a) that are induced by the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$. Fig. 5(b) depicts the comparison between the theoretically estimated and statistically calculated systematic errors in *TCD*, *Hgt* and *SWA* of sample (b) that are induced by the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ as well as the residual signature $\mathbf{r}_\mathbf{M}$. In both Fig. 5(a) and (b), the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ are given by $\Delta\mathbf{a} = [\Delta\theta, \Delta\varphi]^T = [0.5°, 1.0°]^T$ and $\Delta\mathbf{b} = [\Delta P, \Delta A, \Delta C_{S1}, \Delta C_{S2}, \Delta\delta_1, \Delta\delta_2]^T = [0.3°, 0.3°, 0.3°, 0.3°, 0.5°, 0.5°]^T$. As observed from Fig. 5, the theoretically estimated systematic errors exhibit good agreement with those statistically calculated ones for all the combinations of incidence and azimuthal angles, which therefore demonstrates the validity of the systematic error estimation formula given in Eqs. (17)–(19).

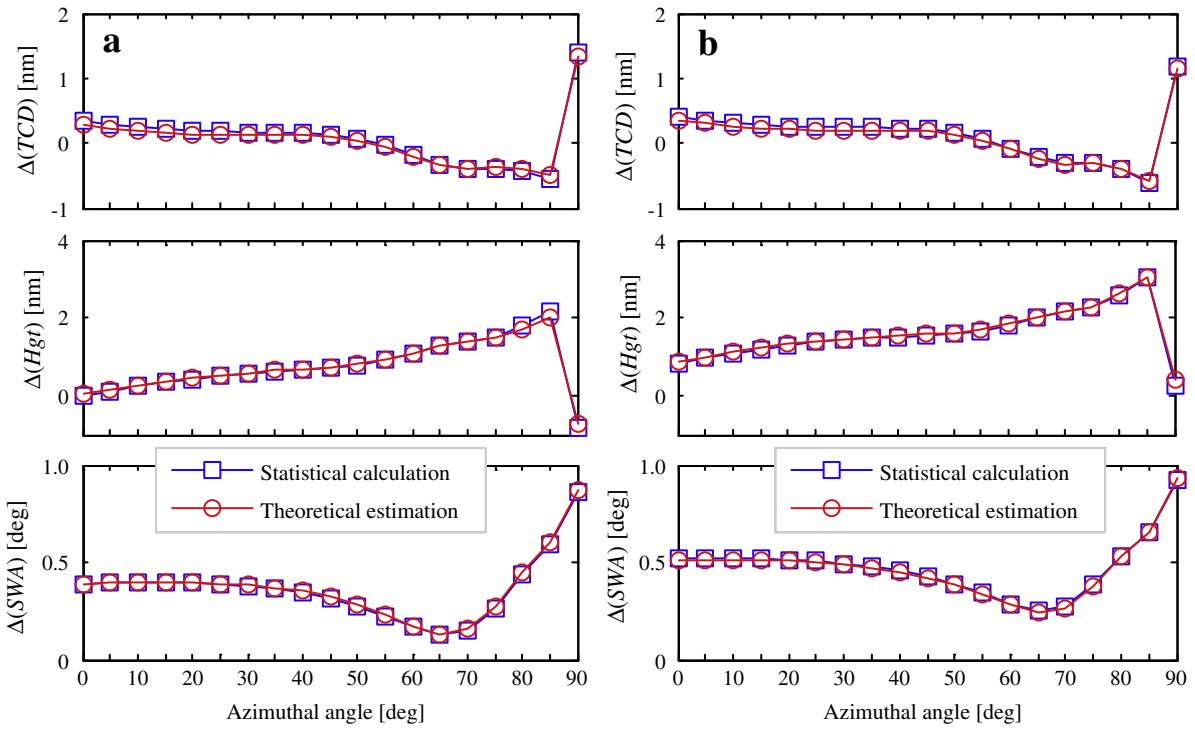

Fig. 5. Comparison between the statistically calculated systematic errors in the extracted structural parameters *TCD*, *Hgt* and *SWA* with those theoretically estimated by Eqs. (17–19). The systematic errors in the extracted structural parameters are induced by (a) the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ for sample (a), and (b) the biases $\Delta\mathbf{a}$ and $\Delta\mathbf{b}$ as well as the residual signature $\mathbf{r_M}$ for sample (b). In the calculations, the incidence angle is fixed at 65° and the azimuthal angles are varied from 0° to 90° with an increment of 5°.

5. Conclusions

In this paper, the propagation and estimation of random and systematic errors in grating reconstruction using a DRC-MMP were theoretically discussed. We derived a generalized first-order propagating formula, which reveals the mechanism of error propagation in grating reconstruction using the DRC-MMP. According to the first-order error propagating formula and the measurement principle of the DRC-MMP, we presented the detailed formulations for the estimation of random and systematic errors in the grating reconstruction procedure. Simulations performed on silicon grating samples have demonstrated that the derived formulations can offer good estimation of the uncertainties and systematic errors that are propagated into the final extracted structural parameters. It is expected that the derived formulations can be used either to assess the measurement precision and accuracy for developing DRC-MMPs in CD and overlay metrology or to guide the measurement configuration optimization for the MMP to achieve a more precise and accurate measurement.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (Grant Nos. 91023032, 51005091, and 51121002) and the National Instrument Development Specific Project of China (Grant No. 2011Y Q160002).

Appendix A. Derivation of Eq. (8)

In order to derive Eq. (8), we introduce the matrix $\mathbf{A}(\mathbf{x}) = \mathbf{J_x}\Delta\mathbf{x} + \mathbf{J_a}\Delta\mathbf{a} + \mathbf{r_M} + \boldsymbol{\mu}_{\Delta\mathbf{M}} + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ for the sake of brevity. Therefore, Eq. (7) can be rewritten as

$$\chi^2_{\min} = [\mathbf{A}(\mathbf{x})]^{\mathrm{T}}\mathbf{W}[\mathbf{A}(\mathbf{x})] = \left[\widetilde{\mathbf{A}}(\mathbf{x})\right]^{\mathrm{T}}\left[\widetilde{\mathbf{A}}(\mathbf{x})\right], \tag{A.1}$$

where $\widetilde{\mathbf{A}}(\mathbf{x}) = \mathbf{W}^{1/2}\mathbf{A}(\mathbf{x}) = \widetilde{\mathbf{J}}_{\mathbf{x}}\Delta\mathbf{x} + \widetilde{\mathbf{J}}_{\mathbf{a}}\Delta\mathbf{a} + \widetilde{\mathbf{r}}_{\mathbf{M}} + \widetilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}} + \widetilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}}$. By taking the derivatives of both sides of Eq. (A.1) with respect to each element of $\mathbf{x}$, we derive that

$$2\frac{\partial\left[\widetilde{\mathbf{A}}(\mathbf{x})\right]^{\mathrm{T}}}{\partial\mathbf{x}}\widetilde{\mathbf{A}}(\mathbf{x}) = 0, \tag{A.2}$$

$$\frac{\partial\left[\widetilde{\mathbf{A}}(\mathbf{x})\right]^{\mathrm{T}}}{\partial\mathbf{x}} = -\widetilde{\mathbf{J}}_{\mathbf{x}}^{\mathrm{T}}. \tag{A.3}$$

Inserting Eq. (A.3) into Eq. (A.2), we have

$$-2\widetilde{\mathbf{J}}_{\mathbf{x}}^{\mathrm{T}}\cdot\left[\widetilde{\mathbf{J}}_{\mathbf{x}}\Delta\mathbf{x} + \widetilde{\mathbf{J}}_{\mathbf{a}}\Delta\mathbf{a} + \widetilde{\mathbf{r}}_{\mathbf{M}} + \widetilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}} + \widetilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}}\right] = 0. \tag{A.4}$$

Considering that $\widetilde{\mathbf{J}}_{\mathrm{x}}$ is not always equal to a zero matrix, thereby we obtain

$$\widetilde{\mathbf{J}}_{\mathbf{x}}\Delta\mathbf{x} + \widetilde{\mathbf{J}}_{\mathbf{a}}\Delta\mathbf{a} + \widetilde{\mathbf{r}}_{\mathbf{M}} + \widetilde{\boldsymbol{\mu}}_{\Delta\mathbf{M}} + \widetilde{\boldsymbol{\varepsilon}}_{\Delta\mathbf{M}} = 0. \tag{A.5}$$

Appendix B. Derivation of Eq. (18)

According to the system model of the DRC-MMP given in Eq. (3), we have

$$\mathbf{G}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{G}} = \mathbf{D}(\mathbf{b}^*)\cdot(\mathbf{M}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}} + \boldsymbol{\mu}_{\Delta\mathbf{M}}). \tag{B.1}$$

As mentioned above, the random error $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ in $\mathbf{M}^{\mathrm{meas}}$ is rooted in the random noise $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ in $\mathbf{G}^{\mathrm{meas}}$, and both $\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$ and $\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$ have a zero mean. By taking the mean values of both sides of Eq. (B.1), we have

$$\mathbf{G}_0 = \mathbf{D}(\mathbf{b}^*)\cdot\langle\mathbf{M}^{\mathrm{meas}}\rangle = \mathbf{D}(\mathbf{b}^*)\cdot(\mathbf{M}_0 + \boldsymbol{\mu}_{\Delta\mathbf{M}}). \tag{B.2}$$

We assume that the polarimetric matrix $\mathbf{D}(\mathbf{b})$ is sufficiently smooth and can be expanded in a Taylor series which, truncated to the first order, leads to a linear model at $\mathbf{b} = \mathbf{b}^*$

$$\mathbf{D}(\mathbf{b}) = \mathbf{D}(\mathbf{b}^*) + (\mathbf{b}-\mathbf{b}^*)^{\mathrm{T}} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}. \tag{B.3}$$

Substitution of $\mathbf{b} = \mathbf{b}_0$ into Eq. (B.3) gives

$$\mathbf{D}(\mathbf{b}_0) = \mathbf{D}(\mathbf{b}^*) + \Delta\mathbf{b}^{\mathrm{T}} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}. \tag{B.4}$$

Inserting Eq. (B.4) into Eq. (B.2) and noticing that $\mathbf{G}_0 = \mathbf{D}(\mathbf{b}_0) \cdot \mathbf{M}_0$, we derive that

$$\mathbf{D}(\mathbf{b}_0) \cdot \boldsymbol{\mu}_{\Delta\mathbf{M}} = \left(\Delta\mathbf{b}^{\mathrm{T}} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}\right) \cdot \langle \mathbf{M}^{\mathrm{meas}} \rangle. \tag{B.5}$$

According to Eq. (B.5), we further obtain

$$\boldsymbol{\mu}_{\Delta\mathbf{M}} = [\mathbf{D}(\mathbf{b}_0)]^{+} \cdot \left(\Delta\mathbf{b}^{\mathrm{T}} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}\right) \cdot \langle \mathbf{M}^{\mathrm{meas}} \rangle. \tag{B.6}$$

Since the true value $\mathbf{b}_0$ of $\mathbf{b}$ is usually unknown, the polarimetric matrix $\mathbf{D}(\mathbf{b}_0)$ will be also unknown. We use Eq. (B.4) to approximate $\mathbf{D}(\mathbf{b}_0)$. By inserting Eq. (B.4) into Eq. (B.6), we derive that

$$\boldsymbol{\mu}_{\Delta\mathbf{M}} = \left[\mathbf{D}(\mathbf{b}^*) + \Delta\mathbf{b}^{T} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}\right]^{+} \cdot \left(\Delta\mathbf{b}^{T} \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{b}}\right) \cdot \langle \mathbf{M}^{\mathrm{meas}} \rangle. \tag{B.7}$$

Appendix C. List of important symbols

a_l	lth element of $\mathbf{a}$
$\mathbf{a}$	L-dimensional vector consisting of L fixed parameters that are input into the optical model
$\mathbf{a}_0$	true value of $\mathbf{a}$
$\mathbf{a}^*$	a priori value of $\mathbf{a}$ during grating reconstruction
$\Delta\mathbf{a}$	difference between $\mathbf{a}_0$ and $\mathbf{a}^*$ as defined by $\Delta\mathbf{a} = \mathbf{a}_0 - \mathbf{a}^*$
$\mathbf{b}$	system-dependent vector with elements being the transmission axis angles of the polarizer P and analyzer A, the fast axis angles C_{S1} and C_{S2} and the phase retardances δ_1 and δ_2 of the two compensators etc., i.e., $\mathbf{b} = [P, A, C_{S1}, C_{S2}, \delta_1, \delta_2, \ldots]^{\mathrm{T}}$
$\mathbf{b}_0$	true value of $\mathbf{b}$
$\mathbf{b}^*$	calibrated value of $\mathbf{b}$
$\Delta\mathbf{b}$	difference between $\mathbf{b}_0$ and $\mathbf{b}^*$ as defined by $\Delta\mathbf{b} = \mathbf{b}_0 - \mathbf{b}^*$
$\mathbf{C}(\cdot)$	covariance matrix of the corresponding vector
$\mathbf{D}(\mathbf{b})$	$K \times 16$ polarimetric matrix as a function of $\mathbf{b}$
g_k	kth element of $\mathbf{G}^{\mathrm{meas}}$
$\mathbf{G}^{\mathrm{meas}}$	K-dimensional flux vector measured by the detector and expressed by $\mathbf{G}^{\mathrm{meas}} = \mathbf{G}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$
$\mathbf{G}_0$	true value of $\mathbf{G}^{\mathrm{meas}}$
$\mathbf{J}_\mathbf{a}$	$N \times L$ Jacobian matrix with respect to $\mathbf{a}$
$\mathbf{J}_\mathbf{x}$	$N \times M$ Jacobian matrix with respect to $\mathbf{x}$
$\mathbf{M}^{\mathrm{meas}}$	N-dimensional vector consisting of N measured Mueller matrix elements of the grating sample as expressed by $\mathbf{M}^{\mathrm{meas}} = \mathbf{M}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{M}} + \boldsymbol{\mu}_{\Delta\mathbf{M}}$
$\mathbf{M}_0$	true value of $\mathbf{M}^{\mathrm{meas}}$
$\mathbf{M}^{\mathrm{calc}}(\mathbf{x},\mathbf{a})$	N-dimensional vector consisting of N calculated Mueller matrix elements of the grating sample with respect to $\mathbf{x}$ and $\mathbf{a}$
$\mathbf{r}_\mathbf{M}$	residual signature induced by the simplified characterization of the grating sample and defined by $\mathbf{r}_\mathbf{M} = \mathbf{M}_0 - \mathbf{M}^{calc}(\mathbf{x}_0,\mathbf{a}_0)$
$u(\cdot)$	estimated uncertainty in the corresponding variable
$\mathbf{W}$	$N \times N$ positive definite weighting matrix
x_m	mth element of $\mathbf{x}$
$\mathbf{x}$	M-dimensional vector consisting of M structural parameters under measurement
$\mathbf{x}_0$	true value of $\mathbf{x}$
$\hat{\mathbf{x}}$	optimal estimation of $\mathbf{x}$ achieved by solving the inverse diffraction problem and expressed by $\hat{\mathbf{x}} = \mathbf{x}_0 + \boldsymbol{\varepsilon}_{\Delta\mathbf{x}} + \boldsymbol{\mu}_{\Delta\mathbf{x}}$
ε_i	ith coefficient of the noise model with $i = 0, 1, 2, 3$
$\boldsymbol{\varepsilon}_{\Delta\mathbf{G}}$	measurement noise in $\mathbf{G}^{\mathrm{meas}}$
$\boldsymbol{\varepsilon}_{\Delta\mathbf{M}}$	random error in $\mathbf{M}^{\mathrm{meas}}$
$\boldsymbol{\varepsilon}_{\Delta\mathbf{x}}$	random error in $\hat{\mathbf{x}}$
$\sigma^2(\cdot)$	variance of the corresponding variable
$\sigma(\cdot)$	standard deviation of the corresponding variable
$\boldsymbol{\mu}_{\Delta\mathbf{M}}$	systematic error in $\mathbf{M}^{\mathrm{meas}}$
$\boldsymbol{\mu}_{\Delta\mathbf{x}}$	systematic error in $\hat{\mathbf{x}}$

References

[1] H.T. Huang, W. Kong, F.L. Terry, Appl. Phys. Lett. 78 (2001) 3983.
[2] R.M. Al-Assaad, S. Regonda, L. Tao, S.W. Pang, W.C. Hu, J. Vac. Sci. Technol. B 25 (2007) 2396.
[3] M. El Kodadi, S. Soulan, M. Besacier, P. Schiavone, J. Vac. Sci. Technol. B 27 (2009) 3232.
[4] R.M. Al-Assaad, D.M. Byrne, J. Opt. Soc. Am. A 24 (2007) 326.
[5] T.A. Germer, H.J. Patrick, R.M. Silver, B. Bunday, Proc. SPIE 7272 (2009) 72720T.
[6] T.A. Germer, (available online at) ftp://ftp.nist.gov/pub/physics/germer/OCDSense/ 2010.
[7] R.M. Silver, N.F. Zhang, B.M. Barnes, H. Zhou, A. Heckert, R. Dixson, T.A. Germer, B. Bunday, Proc. SPIE 7272 (2009) 727202.
[8] P. Vagos, J.T. Hu, Z. Liu, S. Rabello, Proc. SPIE 7272 (2009) 72721N.
[9] A. De Martino, Y.K. Kim, E. Garcia-Caurel, B. Laude, B. Drevillon, Opt. Lett. 28 (2003) 616.
[10] E. Garcia-Caurel, A. De Martino, B. Drevillon, Thin Solid Films 455–456 (2004) 120.
[11] R.M.A. Azzam, Opt. Lett. 2 (1978) 148.
[12] R.W. Collins, J. Koh, J. Opt. Soc. Am. A 16 (1999) 1997.
[13] H.G. Tompkins, E.A. Irene, Handbook of Ellipsometry, William Andrew Publishing & Springer-Verlag, New York, 2005. (chap. 7).
[14] T. Novikova, A. De Martino, P. Bulkin, Q. Nguyen, B. Drévillon, Opt. Express 15 (2007) 2033.
[15] Y. Kim, J. Paek, S. Rabello, S. Lee, J. Hu, Z. Liu, Y. Hao, W. McGahan, Opt. Express 17 (2009) 21336.
[16] J. Li, J.J. Hwu, Y. Liu, S. Rabello, Z. Liu, J. Hu, J. Micro/Nanolithogr. MEMS MOEMS 9 (2010) 041305.
[17] P.S. Hauge, J. Opt. Soc. Am. A 68 (1978) 1519.
[18] D.H. Goldstein, R.A. Chipman, J. Opt. Soc. Am. A 7 (1990) 693.
[19] D.B. Chenault, J.L. Pezzaniti, R.A. Chipman, Proc. SPIE 1746 (1992) 231.
[20] K.M. Twietmeyer, R.A. Chipman, Opt. Express 16 (2008) 11589.
[21] L. Broch, A. En Naciri, L. Johann, Opt. Express 16 (2008) 8814.
[22] L. Broch, A. En Naciri, L. Johann, Thin Solid Films 519 (2011) 2601.
[23] M. Bass, Handbook of Optics, Second ed. McGraw-Hill, New York, 1995. (vol. 2, chap. 2).
[24] M.H. Smith, Appl. Opt. 41 (2002) 2488.
[25] A.K. Jain, Fundamentals of Digital Image Processing, Prentice Hall, New Jersey, 1989. (chap. 8).
[26] F. Argenti, G. Torricelli, L. Alparone, Signal Process. 86 (2006) 2056.
[27] N. Acito, M. Diani, G. Corsini, IEEE Trans. Geosci. Remote Sens. 49 (2011) 2957.
[28] A.C. Rencher, Methods of Multivariate Analysis, second ed. John Wiley & Sons, Inc., New York, 2002.
[29] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985. (chap. 5).
[30] X.G. Chen, S.Y. Liu, C.W. Zhang, H. Jiang, J. Micro/Nanolithogr. MEMS MOEMS 12 (2013) 033013.
[31] C.M. Herzinger, B. Johs, W.A. McGahan, J.A. Woollam, W. Paulson, J. Appl. Phys. 83 (1998) 3323.
[32] M.G. Moharam, E.B. Grann, D.A. Pommet, T.K. Gaylord, J. Opt. Soc. Am. A 12 (1995) 1068.
[33] L. Li, J. Opt. Soc. Am. A 13 (1996) 1870.
[34] S.Y. Liu, Y. Ma, X.G. Chen, C.W. Zhang, Opt. Eng. 51 (2012) 081504.
[35] C.W. Zhang, S.Y. Liu, T.L. Shi, Z.R. Tang, J. Opt. Soc. Am. A 26 (2009) 2327.