Generalized measurement configuration optimization for accurate reconstruction of periodic nanostructures using optical scatterometry

Jinlong Zhu [*a,b], Yating Shi [b], Shiyuan Liu [b], and Lynford L. Goddard [a]

[a] *Photonic Systems Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[b] *State Key Laboratory of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China*

** Corresponding author: jinlongzhu.bryant@gmail.com*

ABSTRACT

Optical scatterometry is a model based technique, which conventionally requires minimization of a predefined least square function. This minimization relies heavily on the measurement configuration: wavelength, incident angle, azimuthal angle, and sample position, which brings up the question of how to find the configuration that maximizes measurement accuracy. We propose a general measurement configuration optimization method based on error propagation theory and singular value decomposition, by which the measurement accuracy can be approximated as a function of a Jacobian matrix with respect to the measurement configurations. Simulation and experiments for a one-dimensional trapezoidal grating establishes the feasibility of the proposed method.

Keywords: scatterometry, measurement configuration optimization, Mueller matrix ellipsometer.

1. INTRODUCTION

Optical scatterometry, also referred to as optical critical dimension (OCD) metrology, is a technique which uses polarized light as the probe to investigate the surface profiles and interior structures of nanoscale samples. Compared with the electron microscopy based techniques such as scanning electron microscopy (SEM) and transmission electron microscopy (TEM), optical scatterometry presents several typical advantages such as low cost, high throughput, small hardware, and minimal sample damage [1-3]. Hence, optical scatterometry is a critical technique for the in-line wafer-to-wafer process monitoring and control in lithography and etch processes [4-6].

Generally, optical scatterometry can be divided into several categories, including reflectance-based [1, 7], ellipsometric parameter-based [3], and Mueller matrix-based [8-10] ones. No matter what kind of apparatus the optical scatterometry relies on, the object of the periodic nanostructure reconstruction is conventionally formulated as the minimization problem of the least square (LSQ) function, which creates the requirement of abundant measurement information so that the minimization problem is overdetermined. Thus, the measurement signature should be a vector containing multiple data points, which directly requires the variability of measurement conditions such as wavelength, incident angle, azimuthal angle, and sample position in x, y, and z. In practice for the consideration of efficiency of data acquisition and analysis, it is common to choose a subset of the measurement conditions from available ranges. The combination of the selected wavelengths, incidence and azimuthal angles, and sample positions is defined as the measurement configuration. It is worthwhile to point out that the measurement precision and accuracy varies with the measurement configuration [8],

* Corresponding author: jinlongzhu.bryant@gmail.com;
webpage: http://www2.hust.edu.cn/nom and http://psl.mntl.illinois.edu/.

Metrology, Inspection, and Process Control for Microlithography XXX, edited by Martha I. Sanchez, Vladimir A. Ukrainstev
Proc. of SPIE Vol. 9778, 977823 · © 2016 SPIE · CCC code: 0277-786X/16/$18 · doi: 10.1117/12.2218988

which creates the issue of measurement configuration optimization for the selection of the optimal configuration corresponding to the highest measurement precision and accuracy. Ku et al. proposed a feature region signature matching method, which seeks the reflectance at some angles containing more information about the structure of the surface relief profile than the reflectance at other angles in a library data match process [11]. Vagos et al. proposed an uncertainty and sensitivity analysis scheme to guide the model and azimuthal angle optimization processes [12]. Gross et al. proposed a method to determine the optimal measurement data set by minimizing the condition number of the Jacobian matrix [13]. Littau et al. investigated several optimal diffraction signature scan path selection techniques to improve scatterometry precision [14]. Novikova et al. demonstrated that the Mueller matrices obtained at some measurement configurations may help to decouple the fitting structural parameters [8]. Chen et al. proposed a measurement configuration optimization method for spectroscopic Mueller matrix ellipsometry (MME) to find an optimal combination of the fixed incidence and azimuthal angles, whose core is transforming the issue of measurement configuration optimization into a max-min 2-norm of the dot product of Jacobians corresponding to fitting structural parameters and configurations respectively [15]. Overall, nearly all of the above methods just focus on a specific type of optical scatterometry; thus, a general measurement configuration optimization method that is suitable for all types of optical scatterometry is highly desirable.

In the present article, we propose a general measurement configuration optimization method arising from the theoretical analysis of the error propagation using the first order Taylor expansion of the LSQ function. Then by using the single value decomposition (SVD) technique and related matrix theory, we demonstrate that the measurement configuration optimization can be formulated as the max-min problem of Frobenius norm of the Jacobian matrix with respect to fitting structure parameters. We should mention that the proposed method is not only useful in OCD metrology, but also applicable in the closely related field of optical wafer defect detection. Two of its primary implementation methods, namely, interferometric [11-13] and through-focus microscopy [14-16], also rely on generating overdetermined measurement data using multiple measurement configurations.

The reminder of this article is organized as follows. Section 2 briefly introduces the inverse problem in optical scatterometry and then presents the error propagation theory and the measurement configuration optimization method in detail. Section 3 introduces the experimental setup, including the setup of a dual-rotating compensator MME as well as the profile detail of the investigated Si grating sample. Section 4 provides the simulations and experiments to examine the validity of the proposed method. Finally, we draw some conclusions in Section 5.

2. METHODOLOGY

2.1 Inverse problem in optical scatterometry

The inverse problem in optical scatterometry is conventionally described as an object to minimize an LSQ function, which can be generally expressed as [22]

$$G(\mathbf{x}) = \sum_{j=1}^{m} \omega_j \left[y_j - f_j(\mathbf{x},\mathbf{a}) \right]^2 = [\mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{a})]^{\mathrm{T}} \mathbf{w} [\mathbf{y} - \mathbf{f}(\mathbf{x},\mathbf{a})], \tag{1}$$

where y_j is the jth measured data point, and $\mathbf{y}$ is the measured signature as a vector containing m data points. $f_j(\mathrm{x})$ is the jth calculated data point with respect to the profile parameters under measurement as an n-dimensional vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ and the measurement configuration, which, in general consists of the wavelength, incident angle, azimuthal angle, and sample position. $\mathbf{f}(\mathbf{x},\mathbf{a})$ is the calculated signature as a vector containing m data points. w_j is the jth weight factor, and $\mathbf{w}$ is an $m \times m$ diagonal matrix with diagonal elements $\{w_j\}$. Consequently, without losing generality, the inverse problem in optical scatterometry can be formulated as

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

where $\hat{\mathbf{x}}$ is the solution of the inverse problem under a given measurement configuration $\mathbf{a}^*$, and Ω is the associated parameter domain.

2.2 Measurement configuration optimization

For most of the cases in optical scatterometry, the measurement configuration parameters are not included in the parameters under measurement. Hence error analysis involved here only focuses on the profile parameter errors of a

nanostructure. Assuming a structural parameter vector $\mathbf{x}$ is close enough to the true parameter vector $\mathbf{x}^*$ under a measurement configuration $\mathbf{a}$, and the function $\mathbf{f}(\mathbf{x},\mathbf{a})$ is sufficiently smooth, then the function value $\mathbf{f}(\mathbf{x}^*,\mathbf{a})$ can be expanded in the vicinity of $\mathbf{x}$ using the first order Taylor expansion formulation

$$\mathbf{f}(\mathbf{x}^*,\mathbf{a}) = \mathbf{f}(\mathbf{x},\mathbf{a}) - \mathbf{J}(\mathbf{x},\mathbf{a})\cdot(\mathbf{x}-\mathbf{x}^*) = \mathbf{f}(\mathbf{x},\mathbf{a}) - \mathbf{J}(\mathbf{x},\mathbf{a})\cdot\Delta\mathbf{x}, \tag{3}$$

where $\mathbf{J}(\mathbf{x},\mathbf{a})$ is the Jacobian matrix with respect to $\mathbf{x}$ and it is a function of $\mathbf{x}$ and $\mathbf{a}$, and $\Delta\mathbf{x}$ represents the error in $\mathbf{x}$ and is given by $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}^*$. For simplicity, we will use $\mathbf{J}$ to take the place of $\mathbf{J}(\mathbf{x},\mathbf{a})$ in the following content. While for the measurement vector $\mathbf{y}$, it is complicated since it can be regarded as the sum of the true signature $\mathbf{F}(\mathbf{x}^*,\mathbf{a})$ and a deterministic offset vector $\boldsymbol{\mu}$ and a random noise vector $\boldsymbol{\varepsilon}$ [10, 23],

$$\mathbf{y} = \mathbf{F}(\mathbf{x}^*,\mathbf{a}) + \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \tag{4}$$

Here we need to give some explanation about Eq. (4). $\boldsymbol{\mu}$ and $\boldsymbol{\varepsilon}$ are the additive intrinsic systematic and additive random errors induced by the measurement system. While for the operator $\mathbf{F}(\cdot)$ it is different from $\mathbf{f}(\cdot)$. The operator $\mathbf{f}(\cdot)$ presented in Eq. (1) ~ (3) is inherently the theoretical forward modeling operator, which usually represents a simplification of the actual measurement system. $\mathbf{F}(\cdot)$ is an extremely complicated forward modeling operator, which represents the actual measurement system and in practice cannot be accurately formulated by any mathematical method. Hence, Eq. (4) is the more realistic formulation than the commonly used expression $\mathbf{y} = \mathbf{f}(\mathbf{x}^*,\mathbf{a}) + \boldsymbol{\mu} + \boldsymbol{\varepsilon}$, which only considers the additive errors. Inserting Eqs. (3) and (4) into Eq. (1), we have

$$G(\mathbf{x}) = [\mathbf{F}(\mathbf{x}^*,\mathbf{a}) + \boldsymbol{\mu} + \boldsymbol{\varepsilon} - \mathbf{f}(\mathbf{x}^*,\mathbf{a}) + \mathbf{J}\Delta\mathbf{x}]^{\mathrm{T}}\mathbf{w}[\mathbf{F}(\mathbf{x}^*,\mathbf{a}) + \boldsymbol{\mu} + \boldsymbol{\varepsilon} - \mathbf{f}(\mathbf{x}^*,\mathbf{a}) + \mathbf{J}\Delta\mathbf{x}], \tag{5}$$

By calculating the derivative of both sides of Eq. (5) with respect to $\mathbf{x}$ and setting the derivative equal to zero, we have

$$\mathbf{w}^{1/2}\mathbf{J}\Delta\mathbf{x} + \mathbf{w}^{1/2}[\mathbf{F}(\mathbf{x}^*,\mathbf{a}) + \boldsymbol{\mu} + \boldsymbol{\varepsilon} - \mathbf{f}(\mathbf{x}^*,\mathbf{a})] = 0. \tag{6}$$

Using the Moore-Penrose pseudo-inverse of $\mathbf{w}^{1/2}\mathbf{J}$, we can further derive

$$\Delta\mathbf{x} = (\mathbf{J}^{\mathrm{T}}\mathbf{w}\mathbf{J})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{w}[\mathbf{f}(\mathbf{x}^*,\mathbf{a}) - \mathbf{F}(\mathbf{x}^*,\mathbf{a}) - \boldsymbol{\mu} - \boldsymbol{\varepsilon}] = \widetilde{\mathbf{J}}[\mathbf{f}(\mathbf{x}^*,\mathbf{a}) - \mathbf{F}(\mathbf{x}^*,\mathbf{a}) - \boldsymbol{\mu} - \boldsymbol{\varepsilon}] = \widetilde{\mathbf{J}}\cdot\widetilde{\mathbf{g}} \tag{7}$$

where $\widetilde{\mathbf{J}} = (\mathbf{J}^{\mathrm{T}}\mathbf{w}\mathbf{J})^{-1}\mathbf{J}^{\mathrm{T}}\mathbf{w}$ and $\widetilde{\mathbf{g}} = \mathbf{f}(\mathbf{x}^*,\mathbf{a}) - \mathbf{F}(\mathbf{x}^*,\mathbf{a}) - \boldsymbol{\mu} - \boldsymbol{\varepsilon}$. Eq. (7) indicates the error vector $\Delta\mathbf{x}$ is dominated by the term $\widetilde{\mathbf{J}}$ and $\widetilde{\mathbf{g}}$. However, each element in the term $\widetilde{\mathbf{g}}$ is unable to be achieved accurately. The feasible way for reducing the value of $\Delta\mathbf{x}$ is to reduce the "amplification factor" $\widetilde{\mathbf{J}}$. Considering the singular value decomposition (SVD) of $\widetilde{\mathbf{J}}$, we have

$$\Delta\mathbf{x} = \widetilde{\mathbf{J}}\cdot\widetilde{\mathbf{g}} = \mathbf{U}\mathrm{diag}(s_i)\mathbf{V}\mathbf{g} = \sum_{i=1}^{r} s_i(\mathbf{v}_i^{\mathrm{T}}\mathbf{g})\mathbf{u}_i, \tag{8}$$

where $\mathbf{U}$, $\mathbf{V}$ and r are the $n \times n$ orthogonal matrix, the $m \times m$ orthogonal matrix and the rank of $\widetilde{\mathbf{J}}$, respectively. $\mathbf{u}_i$ and $\mathbf{v}_i$ are the column vectors of $\mathbf{U}$ and $\mathbf{V}$, respectively. Instability arises due to multiplication by large singular values s_i, and that is also the origin of the bloom of regularization techniques [24, 25]. For the kth element Δx_k of $\Delta\mathbf{x}$, we have

$$\Delta x_k = \sum_{i=1}^{r} s_i(\mathbf{v}_i^{\mathrm{T}}\mathbf{g})u_{ik} = \sum_{i=1}^{r} s_i\widetilde{\mathrm{g}}_{ik}, \tag{9}$$

where u_{ik} is the kth element of $\mathbf{u}_i$, and $\widetilde{\mathrm{g}}_{ik} = (\mathbf{v}_i^{\mathrm{T}}\mathbf{g})u_{ik}$. Applying the Cauchy-Schwarz inequality, we can deduce

$$|\Delta x_k| = \sum_{i=1}^{r}|s_i\widetilde{\mathrm{g}}_{ik}| \le \sqrt{(\sum_{i=1}^{r} s_i^2)(\sum_{i=1}^{r}\widetilde{\mathrm{g}}_{ik}^2)} = \|\widetilde{\mathbf{J}}\|_{\mathrm{F}}\sqrt{\sum_{i=1}^{r}\widetilde{\mathrm{g}}_{ik}^2}, \tag{10}$$

further we have

$$\sum_{k=1}^{n}|\Delta x_k| \le \|\widetilde{\mathbf{J}}\|_{\mathrm{F}}\sum_{k=1}^{n}\sqrt{\sum_{i=1}^{r}\widetilde{\mathrm{g}}_{ik}^2}, \tag{11}$$

where symbol $|\bullet|$ indicates absolute value, and $\|\tilde{\mathbf{J}}\|_{\mathrm{F}}$ denotes the Frobenius norm of $\tilde{\mathbf{J}}$. Inequality (11) indicates that we can minimize $\sum_{k=1}^{n}|\Delta x_k|$ by searching for the measurement configuration $\mathbf{a}$ corresponding to the minimal value of $\|\tilde{\mathbf{J}}\|_{\mathrm{F}}\sum_{k=1}^{n}\sqrt{\sum_{i=1}^{r}\tilde{g}_{ik}^2}$. However, one should note that we can never know the value of $\sum_{k=1}^{n}\sqrt{\sum_{i=1}^{r}\tilde{g}_{ik}^2}$ exactly. Hence, we can only eclectically analyze the term $\tilde{\mathbf{J}}$ to approximate the true upper limit of $\sum_{k=1}^{n}\Delta x_k$. In this article, we define the measurement configuration optimization procedure as

$$\mathbf{A}=\{\mathbf{a}\left|\max_{\mathbf{x}\in\Omega}(\|\tilde{\mathbf{J}}\|_F)<\delta,\delta\in R^+,\mathbf{a}\in\Theta\right\}. \tag{12}$$

where Θ is the domain of measurement configuration $\mathbf{a}$. Expression (12) needs some interpretations. $\tilde{\mathbf{J}}(\mathbf{x},\mathbf{a})$ is equivalent to $\tilde{\mathbf{J}}$ in essence, and the reason for the use of $\tilde{\mathbf{J}}(\mathbf{x},\mathbf{a})$ is to emphasize that $\tilde{\mathbf{J}}$ is the function of $\mathbf{x}$ and $\mathbf{a}$. The term $\max_{\mathbf{x}\in\Omega}(\|\tilde{\mathbf{J}}\|_F)$ indicates we firstly search for the maximal $\|\tilde{\mathbf{J}}(\mathbf{x},\mathbf{a})\|_{\mathrm{F}}$ in a small vicinity Ω of the nominal parameter vector $\mathbf{x}$ under a given measurement configuration $\mathbf{a}$. At each measurement configuration we have a specific maximal $\|\tilde{\mathbf{J}}(\mathbf{x},\mathbf{a})\|_{\mathrm{F}}$, then by scanning the domain Θ and picking out those measurement configuration $\mathbf{a}$ that lead to the term $\max_{\mathbf{x}\in\Omega}(\|\tilde{\mathbf{J}}\|_F)$ smaller than a pre-selected positive real number δ, the optimal set $\mathbf{A}$ of measurement configuration can be obtained. Each element in $\mathbf{A}$ is a potential that corresponds to the relative small $\sum_{k=1}^{n}|\Delta x_k|$. The number of elements in $\mathbf{A}$ is a function of δ, the choice of which depends on the user's belief about the reliability of the measurement configuration optimization procedure.

3. EXPERIMENT SETUP

3.1 Measurement setup and sample description

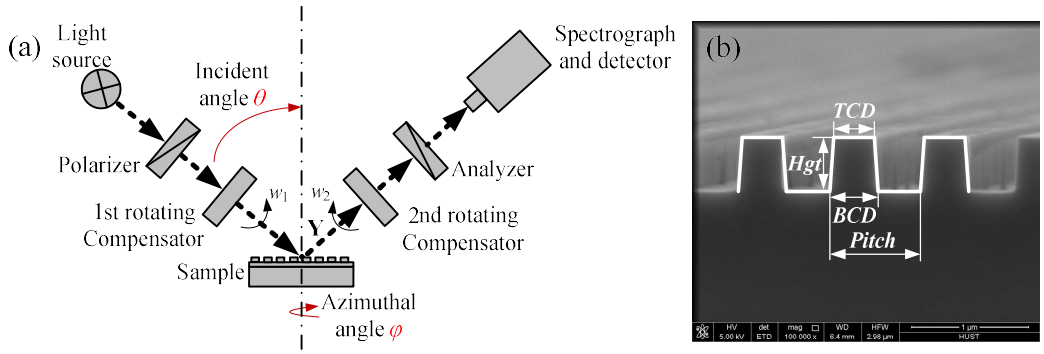

Fig. 1 (a) Measurement setup of the DRC-MME, and (b) cross-section SEM image of the Si grating.

A dual-rotating-compensator (DRC) ellipsometer (RC2 ellipsometer, J. A. Woollam Co.) suitable from ultraviolet to infrared spectrum as presented in Fig. 1(a) and a self-developed optical modeling software based on rigorous coupled-wave analysis (RCWA) [26, 27] is used for demonstration. As labeled in Fig. 1(a), the measurement configuration of the DRC-MME considered in the present paper consists of the incident angle θ and the azimuthal angle φ, while the wavelength λ is kept as constant in the range of 200 ~ 800 nm and the sample position is also fixed. The sample under measurement is a (100)-orientation single layer Si grating, whose cross-section image obtained by an SEM (X-SEM)

(Nova NanoSEM450, FEI Co.) is presented in Fig. 1(b). The grating parameters under measurement are the top critical dimension *TCD*, the height of grating *Hgt*, and the bottom critical dimension *BCD*, while the grating period *Pitch* is fixed at its nominal dimension, i.e., *Pitch* = 800 nm.

4. RESULTS

4.1 Numerical results

In this section, we firstly calculate the corresponding Mueller matrix of a "perfect" Si grating with whose profile parameters *TCD*, *Hgt* and *BCD* being respectively set as 350 nm, 472 nm and 383 nm, after which the simulated different types of errors will be added into the calculated pure signature. As described in Section 2.2, the errors include the random error $\boldsymbol{\varepsilon}$, the additive system error $\boldsymbol{\mu}$ and the non-additive system error contained in $\mathbf{F}(\mathbf{x}^*, \mathbf{a})$. We firstly simulate the above three types of errors by considering several main error sources. The random error $\boldsymbol{\varepsilon}$ and additive system error $\boldsymbol{\mu}$ mainly arise from the fluctuation of the measured light fluxes [22] and the bias in a system-independent vector respectively, and the standard deviation of $\boldsymbol{\varepsilon}$ at a specific wavelength and the specific element of $\boldsymbol{\mu}$ can both be set as a fraction of root-mean-square (rms) in the Mueller matrix over the full wavelength range of interest [28].The fractions of the wavelengths differ from each other, but are all set within the range of 0 ~5% in this article. The simulation of the non-additive system error contained in $\mathbf{F}(\mathbf{x}^*, \mathbf{a})$ is more complicated than that of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$, which in the present article is simulated by considering the finite numerical aperture of the RC2 ellipsometer, the mechanical positioning errors $\Delta\theta$ and $\Delta\varphi$ of θ and φ respectively, and the spectral resolution of the monochromator. The values of $\Delta\theta$ and $\Delta\varphi$ are set as 0.5 and 1.0 degree for all the measurement configurations respectively. At a specific measurement configuration $\mathbf{a} = [\theta, \varphi]$, we can calculate the corresponding contaminated Mueller matrix by following the above simulation process, after which we find the maximal $\left\|\tilde{\mathbf{J}}\right\|_F$ in the vicinity of the nominal values [350, 472, 383] nm of *TCD*, *Hgt* and *BCD*. Here we define the "vicinity" as the ranges of 340 ~ 360 nm, 462 ~ 482 nm and 373 ~ 383 nm for *TCD*, *Hgt* and *BCD* respectively. By picking out the set of $\left\|\tilde{\mathbf{J}}\right\|_F$ that are smaller than a pre-selected δ from all the maxima, the optimal measurement configuration is expected to be among the set of measurement configurations that correspond to the set of $\left\|\tilde{\mathbf{J}}\right\|_F$. The ranges of θ and φ considered in this article are 45 ~ 65 degree and 0 ~ 90 degree respectively, and the increments of each are set as 5 degree.

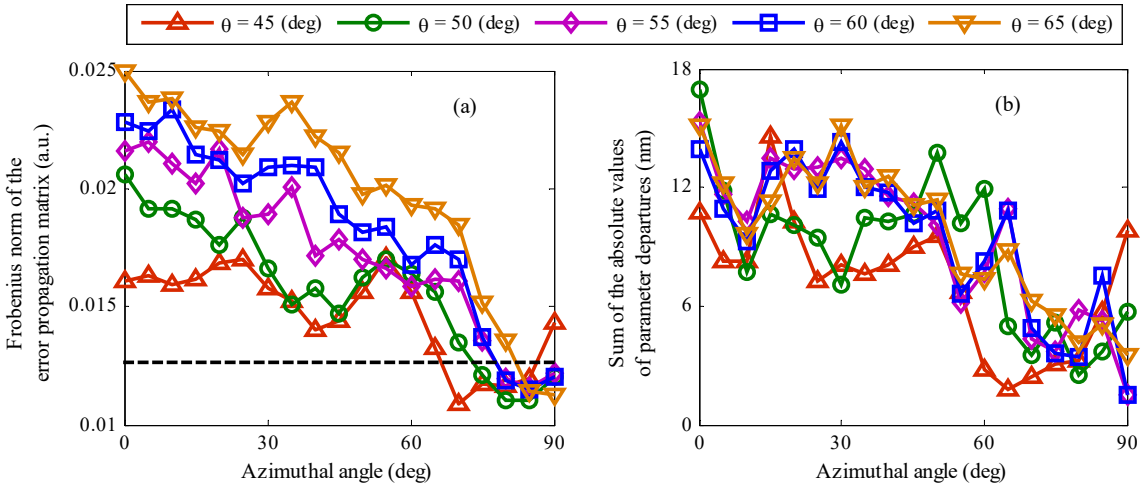

Fig. 2 The (a) Frobenius norm of the error propagation matrix and (b) sum of the absolute values of geometrical parameter departures under different incident and azimuthal angles, respectively. The black dashed line in Fig. 2(a) represents the chosen value of δ, which is 0.0125.

Figure 2(a) presents the Frobenius norm of the error propagation matrix $\|\tilde{\mathbf{J}}\|_F$ under different incident and azimuthal angles, while Fig. 2(b) presents the corresponding sum of absolute values of the geometrical parameter departures. As expected, Fig. 2(b) presents the same downward trend as that of Fig. 2(a) when the value of azimuthal angle increases. We set the value of δ as 0.0125, and pick out those measurement configurations that correspond to values of $\|\tilde{\mathbf{J}}\|_F$ smaller than 0.0125. These selected measurement configurations should ensure that the corresponding sums of the absolute values of parameter departures are as small as possible.

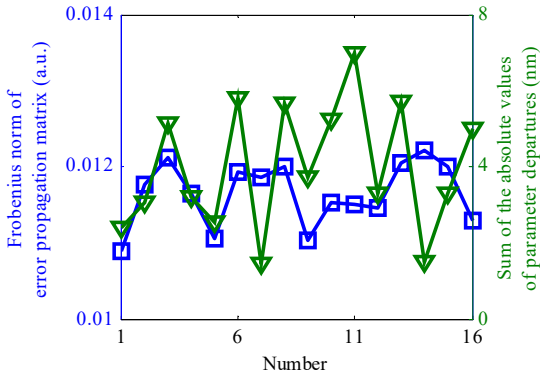

Fig. 3 The optimal set of $\|\tilde{\mathbf{J}}\|_F$ that are smaller than 0.0125 and the corresponding sums of the absolute values of parameter departures.

The left axis and right axis of Fig. 3 represent the Frobenius norm of $\|\tilde{\mathbf{J}}\|_F$ and sum of the absolute values of parameter departures, respectively. As expected, under all these small $\|\tilde{\mathbf{J}}\|_F$, the corresponding extracted error sums are all smaller than seven nanometers. While for those large $\|\tilde{\mathbf{J}}\|_F$, as can be seen in the left part of Fig. 2(b), nearly all of the extracted error sums are larger than eight nanometers. But, we should also note that the local trends of the two curves do not match with each other because of the effect of the term $\sum_{k=1}^{n}\sqrt{\sum_{i=1}^{r}\tilde{g}_{ik}^2}$.

4.2 Experimental results

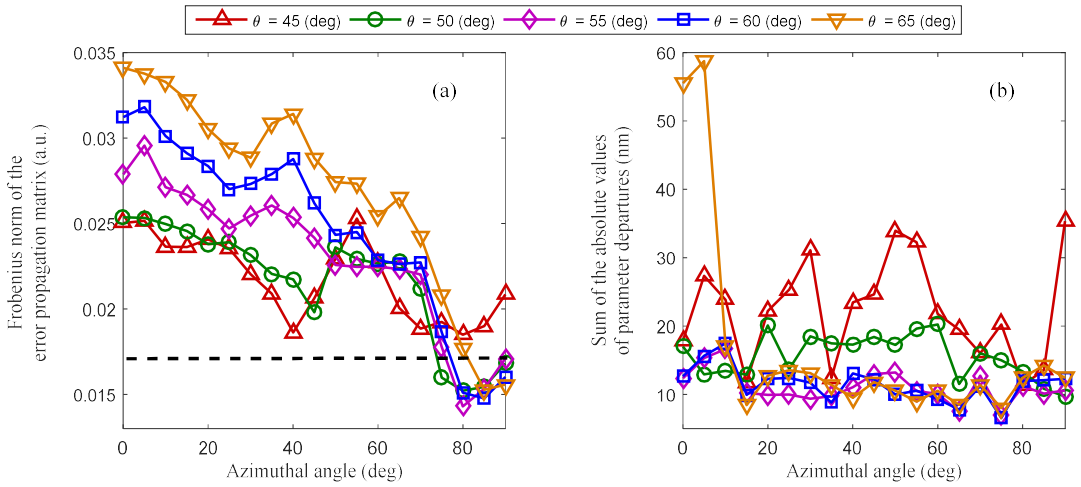

Fig. 4 The (a) Frobenius norm of the error propagation matrix and (b) sum of the absolute values of geometrical parameter departures under different incident and azimuthal angles, respectively. The black dashed line in Fig. 4(a) represents the chosen value of δ, which is 0.0180.

In this section we will experimentally demonstrate the effectiveness of our proposed method. The Mueller matrices (normalized to m11) were measured at 61 points over wavelengths ranging from 200 to 800 nm with the resolution set as 10 nm, and the incidence angle and azimuthal angle respectively vary in the range of 45 ~ 65 degree and 0 ~ 90 degree. It is not possible to know the actual geometrical values; thus, we make a tradeoff here to just treat the SEM measured ones, which are *TCD* = 350 nm, *Hgt* = 472, and *BCD* = 383 nm, as the actual values. The same procedure as presented in section 2.2 and 3.2 is performed to calculate the experimental Frobenius norm of the error propagation matrix, whose results are presented in Fig. 4(a). As can be seen from this figure, the five curves correspond to different incident angles show the similar trend as that of Fig. 2(a); this is not surprising since in the simulation section the actual geometrical values are set to be the same as the SEM measured ones, and besides we should also remember that the term $\left\|\tilde{\mathbf{J}}\right\|_{\mathrm{F}}$ mainly depends on the forward operator. With the increase of the azimuthal angle, the Frobenius norm of the error propagation matrix decreases. Typically, nearly all of the Frobenius norm of the five curves reach a global minimum around the azimuthal angle 80°, which are a little bit different from that in Fig. 2(a). Thus, it is natural to consider setting the threshold δ as a value that is just a little bit larger than the global minima of all the five curves. Here, we set δ as 0.018. Next, we present the sum of the absolute values of parameter departures for all the measurement signatures under different measurement configurations in Fig. 4(b). The parameter departure here represents the difference between the extracted geometrical values and the SEM measured ones.

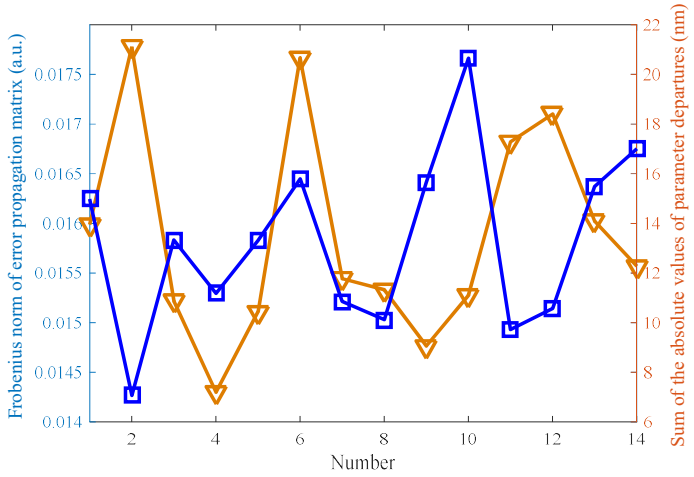

Fig. 5 The experimentally optimal set of $\left\|\tilde{\mathbf{J}}\right\|_{\mathrm{F}}$ that are smaller than 0.0180 and the corresponding sums of the absolute values of parameter departures.

When comparing Fig. 4(a) and Fig. 4(b), we can find that there is not the same decreasing trend. This is not surprising since in reality the measurement error sources are far more complex than that in the simulation environment. For example, in the one-dimensional silicon grating there are line-edge roughness and line-width roughness, which were not considered in the forward operator used in this article. We then picked those summations of absolute values of the parameter departures that correspond to the Frobenius norm value smaller than 0.018 and plot them in Fig. 5. As expected, all of the summations of the absolute values of parameter departures are smaller than 20 nm, in which the global minimum that corresponds to number 4 is even smaller than 8 nm. This demonstrates that by selecting an appropriate δ for the term $\left\|\tilde{\mathbf{J}}\right\|_{\mathrm{F}}$ we can pick out the candidates corresponding to the best measurement accuracy.

5. CONCLUSIONS

In summary, we performed the error propagation analysis in the inverse parameter extraction of scatterometry, by which we found that the measurement accuracy is partly guaranteed by the pseudo-inverse Jacobian with respect to the geometrical parameters. Accordingly, we further deduced that the candidates of the highest measurement accuracy are in close relation with the Frobenius norm of the pseudo-inverse Jacobian, which is the core of our proposed measurement configuration optimization method. By performing the simulation and experiment for a one-dimensional periodic grating,

we have demonstrated the effectiveness of the proposed method. The method is also applicable for optimizing the measurement configuration in optical microscopy based wafer inspection systems. We believe the present work will provide a different point of view for the accurate nanostructure reconstruction in IC manufacturing.

ACKNOWLEDGMENT

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

REFERENCES

[1] C. J. Raymond, M. R. Murnane, S. L. Prins, S. S. H. Naqvi, J. W. Hosch, and J. R. McNeil, "Multiparameter grating metrology using optical scatterometry," J. Vac. Sci. Technol. **15**, 361–368 (1997).

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

[3] H. Huang, and F. Terry Jr., "Spectroscopic ellipsometry and reflectometry from gratings (scatterometry) for critical dimension measurement and in situ, real-time process monitoring," Thin Solid Film **455**, 828–836 (2004).

[4] H. J. Patrick, T. A. Gemer, Y. F. Ding, H. W. Ro, L. J. Richter, and C. L. Soles, "Scatterometry for in situ measurement of pattern flow in nanoimprinted polymers," Appl. Phys. Lett. **93**, 233105 (2008).

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

[6] M. G. Faruk, S. Zangooie, M. Angyal, D. K. Watts, M. Sendelbach, L. Economikos, P. Herrera, and R. Wilkins, "Enabling scatterometry as an in-line measurement technique for 32nm BEOL application," IEEE Trans. Semicond. Manufact. **24**, 499-512 (2011).

[7] R. M. Al-Assaad, and D. M. Byrne, "Error analysis in inverse scatterometry. I. Modeling," J. Opt. Soc. Am. A **24**, 326-338 (2007).

[8] 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).

[9] S. Y. Liu, X. G. Chen, and C. W. Zhang, "Mueller matrix polarimetry: A powerful tool for nanostructure metrology," ECS Trans. **60**, 237-242 (2014).

[10] J. L. Zhu, S. Y. Liu, X. G. Chen, C. W. Zhang, and H. Jiang, "Robust solution to the inverse problem in optical scatterometry," Opt. Express **22**, 22031–22042 (2014).

[11] Y. Ku, S. Wang, D. Shyu, N. Smith, "Scatterometry-based metrology with feature region signatures matching," Opt. Express **14**, 8482–8491 (2006).

[12] P. Vagos et al, "Uncertainty and sensitivity analysis and its applications in OCD measurement," Proc. SPIE **7272**, 72721N (2009).

[13] 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).

[14] M. Littau at al., "Diffraction signature analysis methods for improving scatterometry precision," Proc. SPIE **6152**, 615236 (2006).

[15] X. G. Chen, S. Y. Liu, C. W. Zhang, and H. Jiang, "Measurement configuration optimization for accurate grating reconstruction by Mueller matrix polarimery," J. Micro/Nanolith. MEMS MOEMS **12**, 033013 (2013).

[16] R. Zhou, C. Edwards, A. Arbabi, G. Popescu, and L. L. Goddard, "Detecting 20nm wide defects in large area nanopatterns using optical interferometric microscopy," Nano Lett. 13, 3716-3721 (2013).

[17] R. Zhou, C. Edwards, G. Popescu, L. L. Goddard, "9nm node wafer defect inspection using visible light," Proc. SPIE 9050, 905017 (2014).

[18] R. Zhou, C. Edwards, C. Bryniarski, G. Popescu, and L. L. Goddard, "9nm node wafer defect inspection using three-dimensional scanning: a 405nm diode laser, and a broadband source," Proc. SPIE 9424, 942416 (2015).

[19] B. M. Barnes, M. Y. Sohn, F. Goasmat, H. Zhou, and R. M. Silver, "Scatterfield microscopy of 22nm node patterned defects using visible and DUV light," Proc. SPIE 8324, 83240F (2012).
[20] B. M. Barnes, F. Goasmat, M. Y. Sohn, H. Zhou, and R. M. Silver, "Enhancing 9nm node dense patterned defect optical inspection using polarization, angle, and focus," Proc. SPIE 8681, 86810E (2013).
[21] B. M. Barnes, M. Y. Sohn, F. Goasmat, H. Zhou, A. E. Vladar, R. M. Silver, and A. Arceo, "Three-dimensional deep sub-wavelength defect detection using λ = 193nm optical microscopy," Optics Express 21, 26219-26226 (2013).
[22] M. A. Henn, H. Gross, F. Scholze, M. Wurm, C. Elster, and M. Bar, "A maximum likelihood approach to the inverse problem of scatterometry," Opt. Express **20**, 12771-12786 (2012).
[23] A. Kato and F. Scholze, "Effect of line roughness on the diffraction intensities in angular resolved scatterometry," Appl. Opt. **49**, 6102-6110 (2010).
[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes – the art of science computing*, 3rd ed. (Cambridge University, 2007).
[25] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*, 1st ed. (Academic, 2005).
[26] 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).
[27] L. Li, "Use of Fourier series in the analysis of discontinuous periodic structures," J. Opt. Soc. Am. A **13**, 1870–1876 (1996).
[28] R. M. Al-Assaad, and D. M. Byrne, "Error analysis in inverse scatterometry. I. Modeling," J. Opt. Soc. Am. A **24**, 326-338 (2007).