Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Modeling of line roughness and its impact on the diffraction intensities and the reconstructed critical dimensions in scatterometry

Hermann Gross[1], Mark-Alexander Henn[1], Sebastian Heidenreich[1],

Andreas Rathsfeld[2], Markus Bär[1]

submitted: June 1, 2012

[1] Physikalisch-Technische Bundesanstalt
Braunschweig and Berlin
Fachbereich 8.4,
Mathematische Modellierung und Datenanalyse
Abbestr. 2–12
10587 Berlin
Germany
E-Mail: hermann.gross@ptb.de
mark-alexander.henn@ptb.de
sebastian.heidenreich@ptb.de
markus.baer@ptb.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: andreas.rathsfeld@wias-berlin.de

No. 1711

Berlin 2012

2010 *Mathematics Subject Classification.* 78A46, 65N30, 62P35.

Key words and phrases. diffraction gratings, metrology.

The presented results are part of the project IND17 Scatterometry within the European Metrology Research Programm (EMRP) supported by the European Union and implemented by EURAMET. The authors particularly thank A. Kato and F. Scholze from PTB for many fruitful discussions.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We investigate the impact of line edge and line width roughness (LER, LWR) on the measured diffraction intensities in angular resolved extreme ultraviolet (EUV) scatterometry for a periodic line-space structure designed for EUV lithography. LER and LWR with typical amplitudes of a few nanometers were previously neglected in the course of the profile reconstruction. The 2D rigorous numerical simulations of the diffraction process for periodic structures are carried out with the finite element method (FEM) providing a numerical solution of the two-dimensional Helmholtz equation. To model roughness, multiple calculations are performed for domains with large periods, containing many pairs of line and space with stochastically chosen line and space widths. A systematic decrease of the mean efficiencies for higher diffraction orders along with increasing variances is observed and established for different degrees of roughness. In particular, we obtain simple analytical expressions for the bias in the mean efficiencies and the additional uncertainty contribution stemming from the presence of LER and/or LWR. As a consequence this bias can easily be included into the reconstruction model to provide accurate values for the evaluated profile parameters. We resolve the sensitivity of the reconstruction from this bias by using the LER/LWR perturbed efficiency datasets for multiple reconstructions. If the scattering efficiencies are bias-corrected, significant improvements are found in the reconstructed bottom and top widths toward the nominal values.

1 Introduction

Together with classical metrology techniques such as electron and optical microscopy, scatterometry is widely used to evaluate the precision of diffractive elements in lithography [21]. Since the critical dimensions (CDs) of such elements decrease continuously with the progress in technology, extreme ultraviolet (EUV) scatterometry using light with wavelengths in a small range around 13.5 nm is an adequate tool for the characterization of photo-masks and wafers [12, 27, 19, 22].

However, the determination of the CDs from measured light diffraction patterns is a challenging task. The numerical simulation of the scattering process requires a rigorous modeling by Maxwell's equations [20]. Many methods have been developed to deal with this problem [3, 14, 15, 16, 25]. In this work the finite element method (FEM), that is also capable of computing highly oscillatory fields [5, 8, 2], is applied. Solving the inverse problem of scatterometry amounts to determining the geometry of an optical grating whose diffraction pattern fits a given set of measurement data best. Like many inverse problems, the inverse problem of scatterometry is ill-posed [24] and its treatment requires a priori information. A common approach for its regularization is to set up an equivalent low dimensional optimization problem with a weighted least squares function that is minimized using iterative algorithms [9, 1]. The a priori information in this case not only includes information about the geometrical profile of the investigated element, but also knowledge of the variances of the measured data.

Typically the surface structure is sought in a certain class of gratings that can be described by a small number of parameters. Their values are assumed to lie in certain intervals according to the known quality of the manufacturing process. The weight factors in the least squares function account for measurement uncertainties and therefore represent the a priori knowledge of the underlying measurement error model. In previous publications [10, 17, 9] the reconstructions have been obtained with fixed weight factors commonly representing a relative measurement uncertainty of about 1-3%. A comparison of the reconstructed profiles using EUV-scatterometry and the results obtained using atomic force and electron microscopy [26] has revealed that scatterometry underestimates the side-wall angle, an important feature of the EUV-mask, by several degrees. Imperfect modeling is supposed to be one of the main reasons for this [9, 18, 13]. In particular, to get reliable simulations and reconstructions, line edge roughness (LER) and line width roughness (LWR) have to be taken into account. Recently, Kato and Scholze [13] have suggested approximative analytical expressions for the systematic corrections of the scattered efficiencies stemming from LER and LWR. They have applied Fraunhofer's diffraction method, i.e., the Fourier transform of the reflectivity function of a perturbed binary grating. They also provide formulas for the resulting contributions of LER and LWR to the variances of the efficiencies. These results were, however, obtained for a simplified reflectivity function and not for a realistic geometry of a scatterometry mask. One main goal of the present paper is to resolve the correction to the mean values and the variances of the efficiencies stemming from the influences of LER and LWR based on a realistic geometry.

The present numerical study investigates the impact of LER and LWR on the rigorously calculated diffraction patterns for EUV scatterometry. We use a typical EUV mask composed of stacked TaO-TaN-SiO_2 absorber lines of 80 nm height, 93.33 nm width, and a pitch of 280 nm. To enable the reflection of the incident light a system of layers is coated beneath the line space structure of the EUV mask. This underlying multilayer with a total thickness of 360 nm consists of 49 periodically repeated groups of a Mo-layer and a Si-layer separated by two intermediate MoSi-layers and two capping layers (SiO_2, Si) on top of it. The line roughness LWR and LER are simulated by periodic line-space structures with large periods, consisting of many lines per period with stochastically chosen widths and center positions. A systematical decrease of the mean efficiencies for higher diffraction orders along with increasing variances is found and established for different degrees of roughness. This systematic bias has to be included in the reconstruction model to provide better results for the reconstructed profile parameters.

After a brief description of the mathematical model in Section 2, Section 3 presents the 2D model of LER and LWR used to simulate different degrees of line roughness (cf. [13]). Section 4 contains the results obtained for different numbers of lines per period in the FEM domain and for the different degrees of roughness, respectively. The efficiencies of the randomly created samples are computed for at least 100 realizations of a "rough"grating. The results allow to estimate the mean values, the variances and the deviations from the reference values obtained for the perfectly periodic ("non-rough") grating structure. Finally in Section 5, we are using the LER/LWR perturbed efficiency datasets for repeated reconstructions, i.e., we apply Monte Carlo sampling to study the impact of aperiodic perturbations on the resulting reconstructed profile parameters. The crucial impacts of line roughness on the reconstructed CDs are confirmed, especially the systematic bias and the implications of the increasing variances of the efficiencies with higher diffraction orders are demonstrated. Section 6 closes the paper with a discussion of the result and the conclusions.

2 Mathematical Model of Scatterometry

The mathematical model to describe the propagation of electromagnetic waves in matter is based on Maxwell's equations. From incident light data and characteristic parameters of the irradiated surface profile, the efficiencies and phase shifts for the different diffraction directions are calculated. The time-harmonic Maxwell equations reduce to the two-dimensional Helmholtz equation [20] if geometry and material properties are invariant in one direction

$$\Delta u\left(x,y\right)+k^{2}u\left(x,y\right)=0. \tag{1}$$

In this equation u is the transversal field component that oscillates in the groove direction and k is the wave number $k\left(x,y\right)=\omega\left(\mu_{0}\epsilon\left(x,y\right)\right)^{1/2}$ that is constant for areas filled with the same material. The boundary conditions that are imposed on this partial differential equation are the common transmission conditions on the interfaces between domains, quasi-periodic boundary conditions on the lateral boundaries due to the periodic nature of the structure and the usual outgoing wave conditions in the infinite regions [8]. This boundary value problem can be solved with the finite element method (FEM) for elliptic PDEs [4]. We use the software package DIPOG [6] as the workhorse for our investigations, developed by Weierstrass Institute for Applied Analysis and Stochastics (WIAS) in Berlin.

A measurement dataset, i.e., the diffraction pattern is usually given as a vector $\boldsymbol{y}=(y_1,\ldots,y_m)$, consisting of measured efficiencies or phase shift differences for different wavelengths, incident angles or polarization states. We assume that mask and illumination can be parameterized by a vector $\boldsymbol{p}=(p_1,\ldots,p_n,\lambda_k,\theta_l)$. Here p_i denote geometry parameters of the line-space structure, λ_k and θ_l are wave lengths and angles of incidence for the incoming light. By f we denote the nonlinear operator mapping $\boldsymbol{p}$ to the corresponding diffraction pattern $f\left(\boldsymbol{p}\right)=\left(f_1\left(\boldsymbol{p}\right),\ldots,f_m\left(\boldsymbol{p}\right)\right)$. Note that f is calculated by solving the PDE in Equation (1). The inverse problem of scatterometry can then be formulated as a regression problem minimizing the least squares functional

$$\chi^{2}\left(\boldsymbol{p}\right)=\left\|f\left(\boldsymbol{p}\right)-\boldsymbol{y}\right\|^{2}=\sum_{j=1}^{m}\omega_j\left[f_j\left(\boldsymbol{p}\right)-y_j\right]^{2}. \tag{2}$$

The ω_j's are weight factors that are usually chosen to be the reciprocal variances of the measurements y_j. In the present study we use simulated measurement datasets for the diffracted efficiencies, i.e., the weight factors, resp. the variances of the measurements are known exactly and the weighted LSQ method is suitable for solving the inverse problem. Note that accurate variance estimates are crucial for the application of weighted LSQ and alternative optimization methods like maximum likelihood estimation are appropriate in the case of unknown variances [11]. Supposing independent measurements and the absence of systematic errors, the measurement errors of the jth data point can be modeled as being normally distributed with zero mean, i.e., $\epsilon_j\sim\mathcal{N}\left(0,\sigma_j^2\right)$. Their variances σ_j^2 are composed of two independent random variables, such that

$$\sigma_j^{2}=\left(a\cdot f_j\left(\boldsymbol{p}\right)\right)^{2}+b^{2}. \tag{3}$$

The first term $(a\cdot f_j(\boldsymbol{p}))^2$ indicates the contribution of a linearly dependent noise. The second term b^2 is the contribution of the background noise independent of the measured light intensities. For real EUV measurements [19, 9] a is in the range of 0.01-0.03. The constant b has values of about 0.001-0.002%, if the efficiencies are given in per cent. Power fluctuations of the incidental beam during the recording of the diffraction patterns are the main reason to consider a relative contribution to the variance given by $a > 0$. The average time needed to record an EUV diffraction pattern is about one hour.

The jth measured value is considered as a sum of the value of the model function and the noise contribution ϵ_j with the above variance σ_j^2

$$y_j = f_j\left(\boldsymbol{p}\right) + \epsilon_j. \tag{4}$$

This means that we consider the measurements as noisy realizations of the model. In Section 6 we will come back to a modified form of Equations (3) and (4) considering the effects due to LER/LWR revealed by the results in Section 4 and 5.

3 Modeling Line Edge Roughness (LER) and Line Width Roughness (LWR)

To model line roughness, multiple calculations are performed for domains with large periods, containing many pairs of line and space with stochastically chosen line and space widths. Figure 1 shows this schematically. LER is modeled by random perturbations of the center positions x_i, $i = 1, \ldots, N$, of N neighboring absorber lines, whereas the line width and the pitch d of each profile in this chain are fixed to their nominal values. By the same token, LWR is presented in the scheme of N neighboring lines by randomly perturbed line widths CD_i, $i = 1, \ldots, N$, with unperturbed centers x_i and constant pitch. We assume normally distributed perturbations with variances σ_x^2 and σ_{CD}^2 around the undisturbed center positions and the nominal line width. For different lines, these perturbations are assumed to be independent. Obviously, the positions of left and right edges of the lines are correlated in this modeling concept, i.e., the correlation coefficient is +1 for LER and -1 for LWR, respectively. If both effects are superimposed independently, an uncorrelated roughness of the left and right edges is provided and the variance of each line edge position is given by $\sigma_{edge}^2 = \sigma_x^2 + \sigma_{CD}^2/4$. In the following we depict this situation as LEWR.

Imposing quasi-periodic boundary conditions for the FEM solution of the Helmholtz equation (1), the N neighboring absorber lines form a super-cell with a period of $P = N \cdot d$ for the FEM calculations and d as the pitch of the unperturbed line-space structure. Figure 1 depicts also the

Table 1: Complex indices of refraction; first three components belong to the absorber line structure and the last three ones to the underlying capping and multilayer system of the EUV mask; $MoSi$ depicts the intermediate layers between the Mo- and Si-layers; the values at a wavelength of λ_1=13.389 nm are itemized.

absorber line	$\mathbf{n}_{\lambda_1}$	$\mathbf{k}_{\lambda_1}$
TaO	0.94843	0.03100
TaN	0.94201	0.03416
SiO_2	0.97450	0.01531
substrate	$\mathbf{n}_{\lambda_1}$	$\mathbf{k}_{\lambda_1}$
Si	0.99967	0.00182
$MoSi$	0.97000	0.00424
Mo	0.92552	0.00621

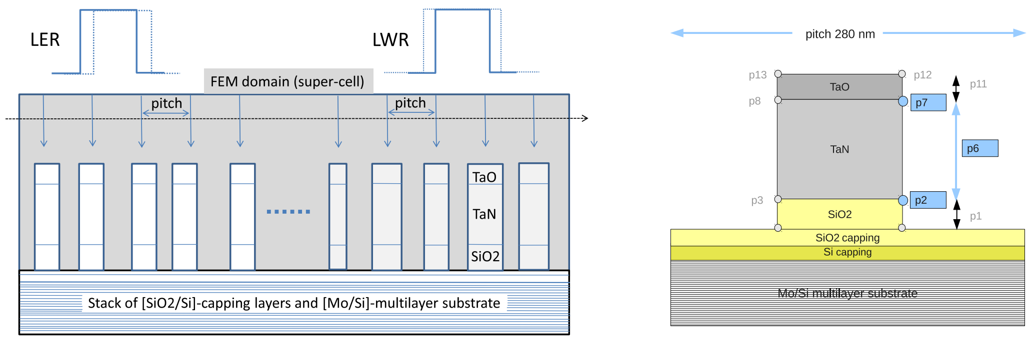

Figure 1: Super-cell containing many profile lines used for roughness modeling by randomly changed center positions for LER and randomly varied line widths for LWR (left); cross section of a single line whose profile parameters such as the horizontal coordinates of the corners $p2$, $p7$ and the height $p6$ have to be reconstructed (right).

detailed structure of the basic absorber line with the profile parameters like corners and heights characterizing the critical dimensions of the structure. These are the parameters which have to be reconstructed by the evaluation of the measured light diffraction pattern.

Regarding the impact of line roughness for EUV gratings, Kato et al. [13] have used the same approach of randomly distributed center positions or line widths, respectively, for their analytical considerations with the Fraunhofer's diffraction method. Germer [7] applied similar design principles for his profile variations of silicon lines investigated in the visible spectral range. Schuster et al. [23] have studied the impact of LER for silicon gratings on the basis of sinusoidal perturbations for the line positions with amplitudes in the range of 2-8 nm and for incident light with wave lengths of 400 nm and 250 nm, respectively.

4 Results

For the rigorous calculations presented here, FEM domains containing 24 and 48 rectangular absorber lines with a pitch of 280 nm and a line to space ratio of 0.5 are used, i.e., super-cells with periods P of 6.72 μm and 13.44 μm, respectively. Calculations with smaller FEM domains of 12 and 6 lines show similar systematic shifts of the mean of the perturbed efficiencies, but they exhibit significantly increased standard deviations for the same amount of samples. For the sake of brevity, they are not presented here.

4.1 Super-cell with 24 lines

About 1000 diffraction patterns for two different scenes of perturbations were calculated. Standard deviations of 2.8 nm and 5.6 nm, i.e., 1% and 2% relative to the pitch of d = 280 nm were used to create random samples of super-cells containing the normally distributed center positions and line widths, respectively. The included orders of diffraction extend from -9 to +8 and efficiencies smaller than 0.001 % were excluded. All calculations were performed with a high level of discretization to ensure convergent results. A wavelength of λ = 13.389 nm and an

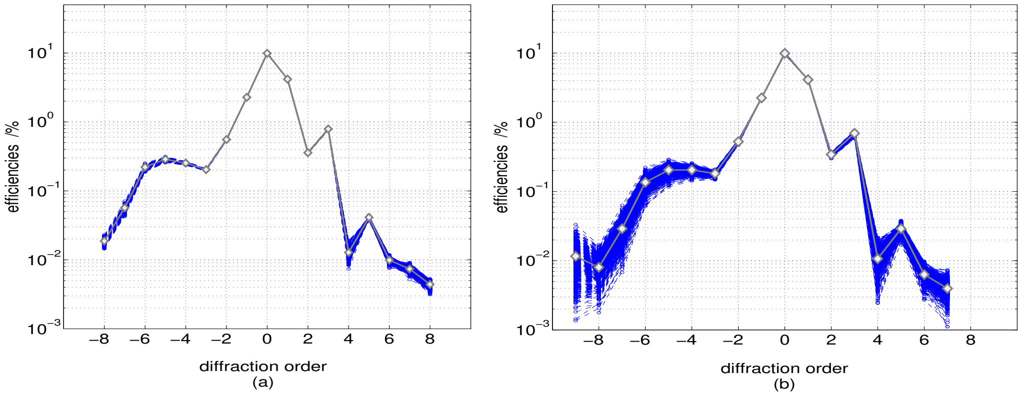

Figure 2: Simulated diffraction patterns for randomly perturbed line-space structures at a wavelength of 13.389 nm; blue circles depict the calculated efficiencies and diamonds the mean efficiencies of all 1000 samples; two different random perturbations of the center positions and the widths (LEWR): (a) $\sigma_x = \sigma_{CD}$= 2.8 nm and (b) $\sigma_x = \sigma_{CD}$= 5.6 nm.

angle of incidence of $\theta = 6°$ were applied, both values are typical for EUV scatterometry. The optical indices of the material components used for the computations are listed in Table 1. On a Linux workstation with 12 Intel Xeon processors (X5460@3.16 GHz) the computation for one simulation takes about 15 minutes. Hence, on a high performance cluster the computation of 1000 diffraction patterns representing a roughness example needs typically 1-2 days depending on the available cluster resources.

Figures 2, 3 and 4 reveal the details of these calculations. Looking at the simulated efficiencies as a function of the diffraction order in Figure 2, significantly increased variances for higher diffraction orders can be recognized. Furthermore one realizes that a doubling of line roughness give rise to a disproportional growth of the variances of the efficiencies. The mean efficiencies over all samples are depicted as diamonds in Figure 2.

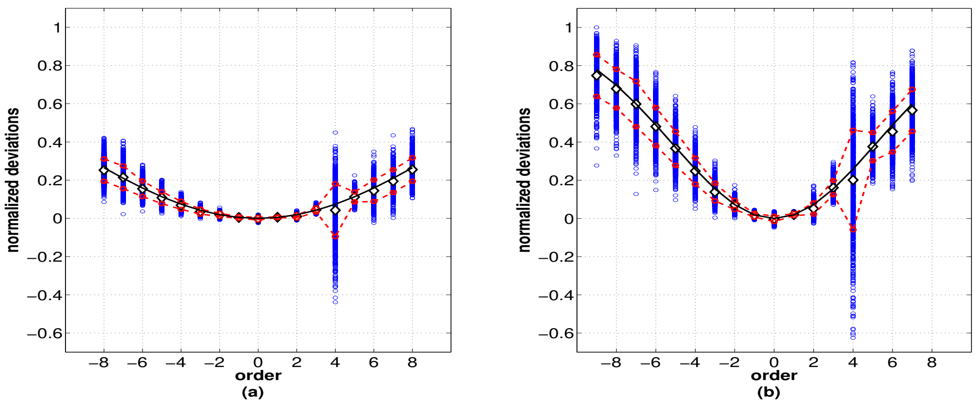

Figure 3: Normalized deviations from the efficiencies of the unperturbed reference line structure, depicted as circles; diamonds represent the mean over the deviations of all 1000 samples; dashed lines indicate the mean efficiency $\pm$ standard deviation; two different random perturbations of the center positions and the widths (LEWR): (a) $\sigma_x = \sigma_{CD}$= 2.8 nm and (b) $\sigma_x = \sigma_{CD}$= 5.6 nm.

A systematic nonlinear decrease of the mean efficiencies for higher diffraction orders along with

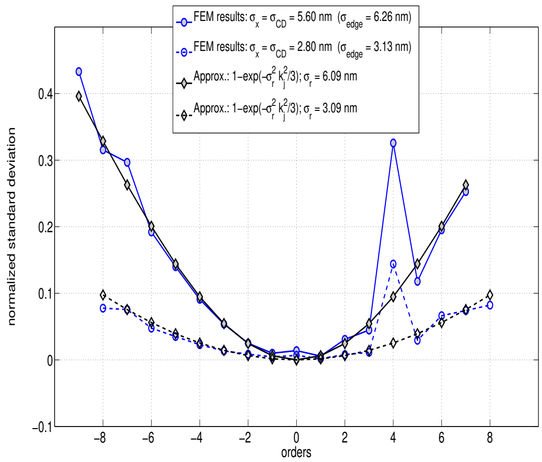

Figure 4: Standard deviations relative to the mean perturbed efficiencies, shown as circles for the two given examples in previous Figure 3; diamonds depict approximations by an exponential function.

increasing variances is observed for different degrees of roughness. The deviations between the unperturbed reference efficiencies $f_{j,ref}(\boldsymbol{p})$ and the mean of perturbed efficiencies $\overline{f_{j,pert}(\boldsymbol{p})}$ are always greater than zero. Figure 3 reveals this by comparison and normalization with the reference values of the efficiencies obtained from the unperturbed line-space structure.

Assume that the parameter σ_r characterizes a general aperiodic perturbation in the sense of the applied LEWR model (cf. Figure 1) and that σ_r scales with the imposed perturbations $\sigma_{edge} = \sqrt{\sigma_x^2 + \sigma_{CD}^2/4}$ of the given grating samples, i.e., $\sigma_r = \alpha \cdot \sigma_{edge}$. Then the mean normalized deviations relative to the references can be approximated by the following exponential function

$$\frac{f_{j,ref}(\boldsymbol{p}) - \overline{f_{j,pert}}}{f_{j,ref}(\boldsymbol{p})} \approx 1 - \exp(-\sigma_r^2 k_j^2) = 1 - \exp\big(-(\alpha\sigma_{edge})^2 k_j^2\big), \quad (5)$$

with σ_r= 3.09 nm ($\alpha = 0.98(7)$) and σ_r= 6.09 nm ($\alpha = 0.97(2)$), respectively. The diffraction order n_j is expressed by the corresponding x-component of the wave vector of the propagating plane wave mode for incidence angle $\theta = 0^\circ$ (cf. [8]), i.e., $k_j = 2\pi n_j/d$.

Equation (5) implies that random perturbations of line and space widths cause an exponential damping of the mean efficiencies. The exponent is proportional to the product of the squared diffraction orders n_j and a constant σ_r^2 which approximates the variance of the line centers and widths. These outcomes confirm the validity of the formula derived by Kato and Scholze [13] using the Fraunhofer approximation.

The increasing variances of the efficiencies with higher diffraction orders become also very clear. For the given two examples of LEWR perturbations, Figure 4 depicts the standard deviations of the efficiencies relative to their mean values. Note that they can be approximated by an exponential function too, just that its exponent is weighted by 1/3 compared to Equation (5) and using the determined values 3.09 nm ($\alpha = 0.98(7)$) and 6.09 nm ($\alpha = 0.97(2)$) for $\sigma_r = \alpha \cdot \sigma_{edge}$ characterizing the damping of the mean efficiencies.

The variances belonging to the diffraction order +4 are outliers in both examples. From the investigation in [13] for a similar EUV mask one can see that LWR is dominating the intensity fluctuations for orders where the form factor has local minima. The form factor of a grating describes the impact of the structure profile on the diffraction pattern by a $sinc^2$ function. For

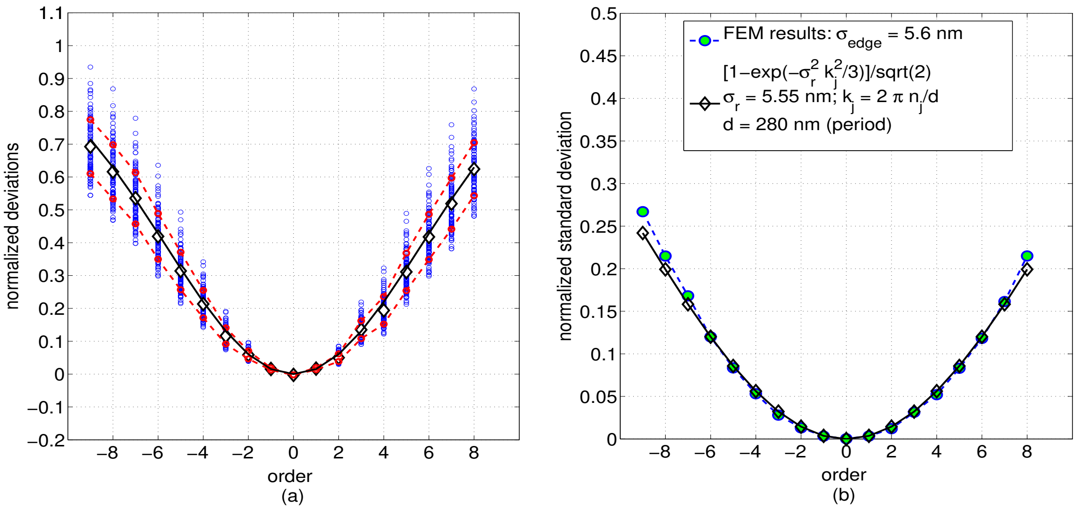

Figure 5: (a) Normalized deviations from the efficiencies of the unperturbed reference line structure, depicted as circles; diamonds represent the mean over the deviations of all 100 samples; dashed lines indicate the mean efficiency $\pm$ standard deviation; random perturbations of the center positions (LER): $\sigma_x = 5.6$ nm. (b) Standard deviations relative to the mean perturbed efficiencies shown as circles for the given example; diamonds depict best approximation by an exponential function.

our EUV grating and the applied angle of incidence of 6° this is the case for diffraction order +4 in the range of the considered orders.

4.2 Super-cell with 48 lines

In order to estimate the changes of the variances of the efficiencies in dependence on the number of lines in the super-cell, the calculations for a super-cell twice the size, i.e., for 48 lines per domain are repeated. As already mentioned we have tested 6 and 12 lines per super-cell too, but for brevity they are not presented here. Furthermore, LER and LWR perturbations are calculated separately in order to clarify their specific impacts. We impose again normally distributed perturbations with values for σ_x (LER) and σ_{CD} (LWR), respectively, of 2 percent relative to the period of 280 nm of the basic absorber line structure. The results of these calculations are given in Figures 5 and 6. Subsequently, LEWR perturbations with $\sigma_x = \sigma_{CD} = 5.6$ nm were imposed. The results of these simulations are shown in Figure 7. Besides the normalized deviations from the efficiencies of the unperturbed reference line structure, the standard deviations relative to the mean perturbed efficiencies as function of the diffraction order are given in these three figures. Due to the computational cost for such large domains containing 48 lines, only 100 diffraction patterns for each scene of line roughness were calculated.

We recognize once more an exponential damping for the mean values of the efficiencies, but now significantly reduced variances (cf. Figure 5) as it was expected due to the improved averaging by doubling the line numbers. Applying Equation (5) the systematic decrease of the mean efficiencies with higher diffraction orders can be approximated fairly well by $\sigma_r = 5.55$ nm ($\alpha = 0.99(1)$) for the LER example and by $\sigma_r = 2.74$ nm ($\alpha = 0.97(8)$) for the LWR simulations indicating a significantly reduced impact of LWR except for order +4 (cf. Figure 6). A comparison of Figures 5 and 6 shows that for equal amplitudes of LER and LWR, the LER

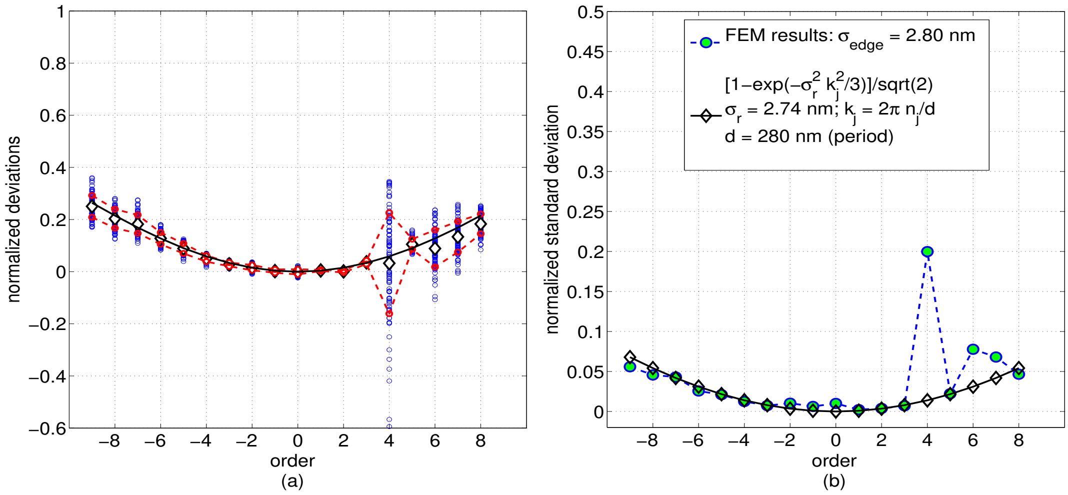

Figure 6: (a) Normalized deviations from the efficiencies of the unperturbed reference line structure, depicted as circles; diamonds represent the mean over the deviations of all 100 samples; dashed lines indicate the mean efficiency $\pm$ standard deviation; random perturbations of the line widths (LWR): $\sigma_{CD} = 5.6$ nm. (b) Standard deviations relative to the mean perturbed efficiencies shown as circles for the given example; diamonds depict an approximation by an exponential function.

influence both on the systematic shift of the mean (Figures 5a, 6a) and on the variance (Figures 5b, 6b) is substantially stronger. In addition, it becomes clear that the systematic damping due to LER is nearly symmetric.

The asymmetric behavior for the positive diffraction orders and especially for order +4 in the LEWR example (cf. Figure 7) is therefore arising from the random perturbations of the line widths as already mentioned in the previous Subsection 4.1. The order dependent decrease of the mean efficiencies for the LEWR scene with 48 lines is well approximated by $\sigma_r = 6.17$ nm ($\alpha = 0.98(5)$) applying Equation (5). Once more, this is in good accordance with the analytical considerations in [13] employing Fraunhofer diffraction method. The results of all presented line roughness examples and their approximations with the proposed exponential damping are summarized in Table 2.

Table 2: Results σ_r applying the proposed exponential approximations (eqs. (5) and (6)) for all calculated examples and comparison with the expected values $\sigma_{edge} = \sqrt{\sigma_x^2 + \sigma_{CD}^2/4}$; for $\sigma_x = \sigma_{CD} \Rightarrow \sigma_{edge} = 1.118 \cdot \sigma_x$; all σ values are given in nm;

example	σ_x	σ_{CD}	σ_{edge}	σ_r	$\alpha = \sigma_r/\sigma_{edge}$
LEWR, 24 lines	2.80	2.80	3.13	3.09	0.98(7)
LEWR, 24 lines	5.60	5.60	6.26	6.09	0.97(2)
LER, 48 lines	5.60	-	5.60	5.55	0.99(1)
LWR, 48 lines	-	5.60	2.80	2.74	0.97(8)
LEWR, 48 lines	5.60	5.60	6.26	6.17	0.98(5)

Note that the normalized standard deviations of the efficiencies for all examples produced by 48 lines per super-cell are fairly well approximated applying (5) and weighting of the exponent

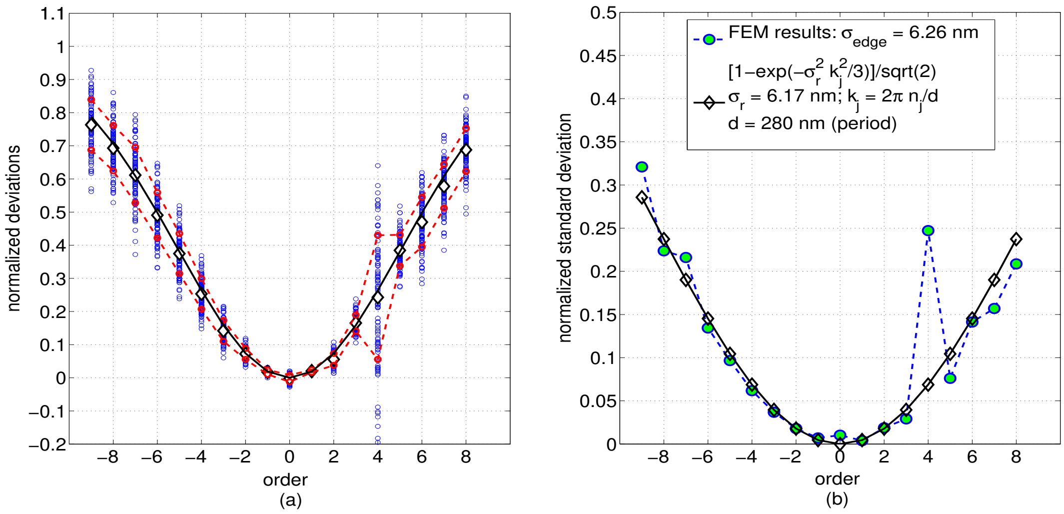

Figure 7: (a) Normalized deviations from the efficiencies of the unperturbed reference line structure, depicted as circles; diamonds represent the mean over the deviations of all 100 samples; dashed lines indicate the mean efficiency $\pm$ standard deviation; random perturbations of the center positions and widths (LEWR): $\sigma_x = \sigma_{CD} = 5.6$ nm. (b) Standard deviations of the efficiencies relative to their mean values shown as circles; diamonds depict an approximation by an exponential function.

of this equation with one third as in Subsection 4.1 with 24 lines (cf. Figure 4). But furthermore the whole function is scaled down by $1/\sqrt{2}$ which is simply caused by the doubling of the line numbers. Considering additional results for the normalized standard deviations obtained by 12 and 6 lines per super-cell which are not shown here, the dependency of the normalized standard deviations of the efficiencies on the number N of lines compared to the configuration with 24 lines is well described by

$$\Sigma(N, k_j, \sigma_r) \approx \sqrt{\frac{24}{N}} \left(1 - \exp(-\sigma_r^2 k_j^2/3)\right). \tag{6}$$

That means, even the normalized standard deviations Σ of the efficiencies stemming from the presence of line roughness can be summarized by an analytic expression.

5 Impact of roughness on reconstructed critical dimensions (CD)

The presented effects of line roughness on the efficiency pattern for the forward problem in scatterometry modeling are crucial for the obtained values and variances of the reconstructed profile parameters, i.e., the CDs of the examined mask. The impact of line roughness on the efficiencies will pass through the reconstruction algorithms. To clarify this connection we use the LEWR perturbed efficiency datasets for repeated reconstructions, i.e., we apply Monte Carlo sampling to study the impact of the aperiodic perturbations on the reconstructed profile parameters.

First, an LEWR perturbed efficiency dataset with $\sigma_x = \sigma_{CD} = 2.8$ nm composed of 100 different patterns is chosen as the measured input for reconstructions, i.e., for the minimization of the objective functional given in Equation (2). Three profile parameters, namely the two horizontal

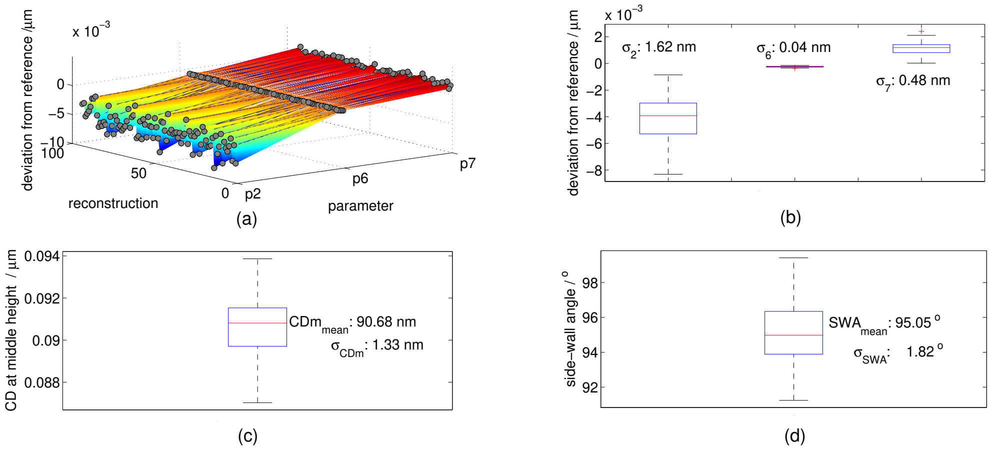

Figure 8: Impact of normally distributed random perturbations of the center positions and widths (1 % LEWR: $\sigma_x = \sigma_{CD} = 2.8$ nm; 100 samples) on the reconstructed parameters right lower and upper corner p_2, p_7 and height p_6 of the TaN absorber layer (a),(b); on width CDm of the absorber layer at middle height (c) and on the side-wall angle SWA of the absorber line (d).

coordinates of the corners p_2, p_7 and the height p_6 of the main absorber layer (cf. Figure 1), are forming the parameter vector $\boldsymbol{p}$ to be reconstructed from the given efficiencies. Second, we produce from the same sets of efficiencies a bias-corrected set by multiplying the efficiencies with the corresponding inverse exponential damping factor characterizing the applied LEWR perturbations, and then we repeat the MC reconstructions.

Figures 8 and 9 present the obtained results essentially as boxplots. The distribution of the reconstructed parameters is given by the deviations from their nominal values. The ranges between lower and upper quartiles are indicated by boxes, the medians by lines. The whiskers extend to the most extreme data points not considered as outliers, and the outliers are plotted individually as plus signs. The calculated standard deviations σ_i, $i = 2, 6, 7$, for the reconstructed parameters are given in nm.

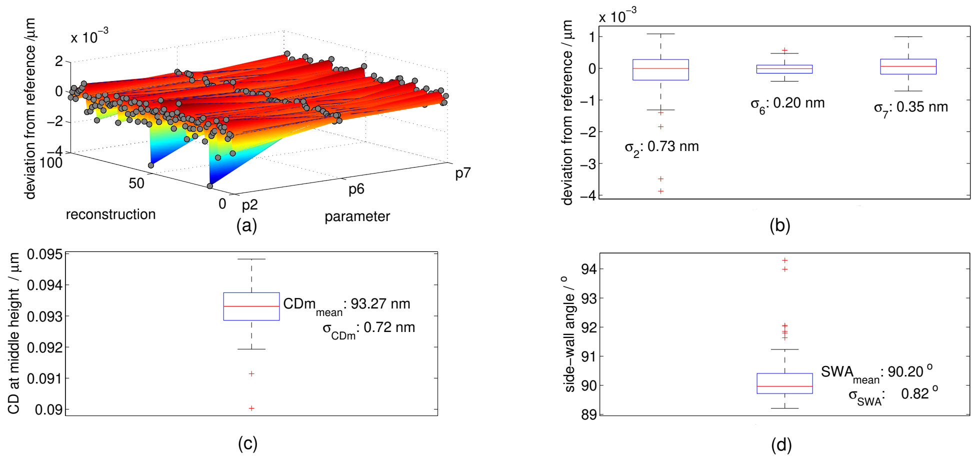

Figure 9: Impact of normally distributed random perturbations of the center positions and widths (1 % LEWR: $\sigma_x = \sigma_{CD} = 2.8$ nm; 100 samples) on the reconstructed parameters right lower and upper corner p_2, p_7 and height p_6 of the TaN absorber layer (a),(b); on width CDm of the absorber layer at middle height (c) and on the side-wall angle SWA of the absorber line (d); bias-corrected datasets of efficiencies are used.

As expected, the crucial implications of line roughness on the reconstructed CDs are confirmed, especially the systematic bias give rise to significant shifts in the corners of the absorber line (cf. Figure 8 (a),(b)). Note that the nominal value for CDm at the middle height of the absorber line is 93.33 nm and 90° for the side-wall angle of the examined structure (cf. Figure 8 (c),(d)). Only after compensating the LEWR perturbed efficiencies by the right inverse damping, the reconstructed parameters show a distribution around their nominal values. Figure 9 reveals this distinctly. The remaining variances are mainly due to the random perturbations of the efficiencies caused by line roughness.

6 Discussion and conclusions

The most important consequence from the presented numerical experiments is that the revealed LER/LWR-bias has to be included in the scatterometric measurement model by an order dependent damping factor. The representation of the measurements and their associated variances previously given in Section 2 extend at least to

$$\begin{aligned} y_j &= \exp\left(-\sigma_r^2 k_j^2\right) \cdot f_j(\boldsymbol{p}) + \epsilon_j \,, & (7)\\ \sigma_j^2 &= \Big(a \cdot \exp\big(-\sigma_r^2 k_j^2\big) \cdot f_j(\boldsymbol{p})\Big)^2 + b^2. & (8) \end{aligned}$$

Note that for the variances of the measurement errors in Equation (8) only the bias is taken into account adequately. The variances of the efficiencies due to LEWR however are not yet considered.

For EUV measurements taken from surface areas with a size in the range of 500 μm x 500 μm we expect that the contribution of the roughness to σ_j^2 (cf. Equation (2)) will be significantly reduced by spacial averaging compared to the calculated values in our investigations. Due to the required computation time for large FEM domains along with the small EUV wavelenghts the maximal lateral expansion of the super-cells used in this work is restricted to 13.44 μm. Applying Equation (6) with a value of N = 1785, i.e., an imaginary extension of our FEM super-cell period to 499.8 μm (= 0.280*1785), the normalized standard deviations of the efficiencies would be scaled down by a factor of about 0.116 compared to the values given in Figure 4 obtained with 24 lines. Nevertheless, for higher diffraction orders, LEWR based fluctuations of the efficiencies of several percent remain and correspond to variances as in the first term on the right-hand side of (3) with typical values for factor a in the range of 0.01-0.03 (cf. Section 2). Therefore equation (8) for the variances of the efficiencies may be supplemented by the estimated LEWR contributions

$$\sigma_j^2 = \left[a^2 + \Sigma^2(N, k_j, \sigma_r)\right] \cdot \Big(\exp\left(-\sigma_r^2 k_j^2\right) \cdot f_j(\boldsymbol{p})\Big)^2 + b^2, \tag{9}$$

with $\Sigma(N, k_j, \sigma_r)$ given by Equation (6) applied with a value of line numbers N at least ≥ 1785. A further reduction of the contribution of Σ to the variances should, however, be expected by the fact that no perturbations along the line direction is assumed in our two-dimensional approach. Further investigations such as 3D calculations for perturbed line-space structures are needed to estimate the expected further degradation of the LEWR contribution to the variances of the efficiencies.

As already mentioned in Section 2, accurate variance estimates are crucial for the application of weighted LSQ. In general, for real measurement datasets of efficiencies the roughness induced bias and the additional variances are unknown. Equations (6) to (9) represent a first approximation to describe these effects and to include them into the model. Finally, at least σ_r should be treated as an additional unknown parameter which has to be reconstructed too. To solve the indirect problem under such conditions, other optimization methods are more appropriate (cf. [11]) and will be discussed in further publications.

The proposed equations to approximate the roughness effects are symmetric with respect to the diffraction orders. In particular this holds for the presented LER simulations, i.e., the variations of the center positions of the absorber lines this is very well supported. Compared to LER the presented LWR calculations indicate a significantly reduced damping impact on the mean of the efficiencies for equal amplitudes of perturbations. The damping factor can still be approximated fairly well by a symmetric exponential formula. Only for some positive diffraction orders, where the form factor of the grating causes local minima in the efficiency pattern, strongly increased variances have been revealed.

The presented assessments of LER and LWR effects were provided for a typical EUV line-space structure by repeated calculations for large FEM computation domains with stochastically chosen line and space widths. We find acceptable results for the mean values of the efficiencies by using an exponential damping formula similar to that of the Fraunhofer approximation. Switching to the inverse problem, ignoring these dampings in the simulated measurement datasets leads to large errors of the parameters in the reconstruction algorithms. Including these dampings by providing bias-corrected datasets, we get acceptable results for the reconstruction.

Though our model is a simple two-dimensional assuming extremely slow changes of the widths and center locations along the line direction, we conjecture that damping will improve the case of faster changes too. Nevertheless rigorous 3D calculations for biperiodic grating structures including stochastical perturbations similar to the proposed ones are necessary to validate finally the impact of line roughness.

References

[1] Al-Assaad, R.M. and Byrne, D.M., Error analysis in inverse scatterometry. I. Modeling, *Journal of the Optical Society of America A*, **24 (2)**, pp. 326–338, 2007.

[2] Cessenat, O. and Despres, B., Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem, *SIAM journal on numerical analysis*, **35 (1)**, pp. 255–299, 1998.

[3] Chandezon, J., Raoult, G., and Maystre, D., A new theoretical method for diffraction gratings and its numerical application, *Journal of Optics*, **11 (4)**, pp. 235–241, 1980.

[4] Ciarlet, P.G., *The finite element method for elliptic problems*, North-Holland, 1978.

[5] Elschner, J., Hinder, R., and Schmidt, G., Finite element solution of conical diffraction problems, *Advances in Computational Mathematics*, **16 (2)**, Springer-Verlag, Berlin Heidelberg, pp. 139–156, 2002.

[6] Elschner, J., Hinder, R., Rathsfeld, A., and Schmidt, G., Homepage of the software package DIPOG, http://www.wias-berlin.de/software/DIPOG

[7] Germer, T.A., Effect of line and trench profile variation on specular and diffusive reflectance from periodic structure, *J. Opt. Soc. Am. A*, **24 (3)**, pp. 696–701, 2007.

[8] Gross, H., Model, R., Bär, M., Wurm, M., Bodermann, B., and Rathsfeld, A., Mathematical modelling of indirect measurements in scatterometry, *Measurement*, **39 (9)**, pp. 782–794, 2006.

[9] Gross, H., Rathsfeld, A., Scholze, F., and Bär, M., Profile reconstruction in extreme ultraviolet (EUV) scatterometry: modeling and uncertainty estimates, *Measurement Science and Technology*, **20**, p. 105102/1–105102/11, 2009.

[10] Gross, H., Rathsfeld, A., Scholze, F., Model R., and Bär, M., Computational methods estimating uncertainties for profile reconstruction in scatterometry, *Proc. of SPIE*, **6995** , pp. 6995OT-1–6995OT-9, 2008.

[11] Henn, M.-A., Gross, H., Elster, C., Scholze, F., Wurm, M., and Bär, M, A statistical approach to the inverse problem of scatterometry using maximum likelihood estimation, in preparation.

[12] Henn, M.-A., Model, R., Bär, M., Wurm, M., Bodermann, B., Rathsfeld, A., and Gross, H., On numerical reconstructions of lithographic masks in DUV scatterometry, *Proc. SPIE*, **7390**, p. 73900Q, 2009.

[13] Kato, A. and Scholze, F., Effect of line roughness on the diffraction intensities in angular resolved scatterometry, *Applied Optics*, **49 (31)**, pp. 6102–6111, 2010.

[14] Lalanne, P. and Morris, G.M., Highly improved convergence of the coupled-wave method for TM polarization, *Journal of the Optical Society of America A*, **13 (4)**, OSA, pp. 779–784, 1996.

[15] Li, L., New formulation of the Fourier modal method for crossed surface-relief gratings, *Journal of the Optical Society of America A*, **14 (10)**, OSA pp. 2758–2767, 1997.

[16] Moharam, M.G., Grann, E.B., Pommet, D.A., and Gaylord, T.K., Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings, *Journal of the Optical Society of America A*, **12 (5)**, OSA, pp. 1068–1076, 1995.

[17] Patrick, H., Germer, T., and Ding, Y., Ro, H., Richter, L., and Soles, C., In situ measurement of annealing-induced line shape evolution in nanoimprinted polymers using scatterometry, *Proc. SPIE*, **7271**, p. 727128, 2009.

[18] Patrick, H. and Germer, T. and Silver, R. and Bunday, B., Developing an uncertainty analysis for optical scatterometry, *Proc. SPIE*, **7272**, p. 72720T, 2009.

[19] Perlich, J., Kamm, F.M., Rau, J., Scholze, F., and Ulm, G., Characterization of extreme ultraviolet masks by extreme ultraviolet scatterometry, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, **22**, pp. 3059, 2004.

[20] Petit, R. and Botten, L.C., *Electromagnetic theory of gratings*, Springer-Verlag, Berlin, Heidelberg, 1980.

[21] Raymond, C.J., Murnane, M.R., Prins, S.L., Sohail, S., Naqvi, H., McNeil, J.R., and Hosch, J.W., Multiparameter grating metrology using optical scatterometry, *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, **15 (2)**, AVS, pp 361–368, 2009.

[22] Scholze, F. and Laubis, C., Use of EUV scatterometry for the characterization of line profiles and line roughness on photomasks, *EMLC 2008*, VDE VERLAG GMBH, Berlin und Offenbach, pp. 374–382, 2008.

[23] Schuster, T., Rafler S., Paz, V.F., Frenner, F., and Osten, W., Fieldstiching with Kirchhoff-boundaries as a model based description for line edge roughness (LER) in scatterometry, *Microelectric Engineering*, **86**, pp. 1029–1032, 2009.

[24] Tarantola, A., *Inverse problem theory*, Elsevier Amsterdam etc., 1987.

[25] Tavrov, A., Totzeck, M., Kerwien, N., and Tiziani, H.J., Rigorous coupled-wave analysis calculus of submicrometer interference pattern and resolving edge position versus signal-to-noise ratio, *Optical Engineering*, **41**, pp. 1886, 2002.

[26] Wurm, M., *Über die dimensionelle Charakterisierung von Gitterstrukturen auf Fotomasken mit einem neuartigen DUV-Scatterometer*, PhD thesis, Friedrich-Schiller-Universität Jena, 2008.

[27] Wurm, M., Bonifer, S., Bodermann, B., and Richter, J., Deep ultraviolet scatterometer for dimensional characterization of nanostructures: system improvements and test measurements, *Measurement Science and Technology*, IOP Publishing, **22 (9)**, p. 94024–94032, 2011.