Zuse Institute Berlin

MARTIN HAMMERSCHMIDT, MARTIN WEISER, XAVIER GARCIA SANTIAGO, LIN ZSCHIEDRICH, BERND BODERMANN,SVEN BURGER

Quantifying parameter uncertainties in optical scatterometry using Bayesian inversion

Quantifying parameter uncertainties
in optical scatterometry
using Bayesian inversion

Martin Hammerschmidt, [ab] Martin Weiser, [b] Xavier Garcia Santiago, [ac]
Lin Zschiedrich, [a] Bernd Bodermann, [d] Sven Burger [ab]

[a] JCMwave GmbH
Bolivarallee 22, D – 14 050 Berlin, Germany

[b] Zuse Institute Berlin (ZIB)
Takustraße 7, D – 14 195 Berlin, Germany

[c] Institut für Theoretische Festkörperphysik (KIT)
Wolfgang-Gaede-Straße 1, D – 76 131 Karlsruhe, Germany

[d] Physikalisch-Technische Bundesanstalt (PTB)
Bundesallee 100, D – 38 116 Braunschweig, Germany

ABSTRACT

We present a Newton-like method to solve inverse problems and to quantify parameter uncertainties. We apply the method to parameter reconstruction in optical scatterometry, where we take into account a priori information and measurement uncertainties using a Bayesian approach. Further, we discuss the influence of numerical accuracy on the reconstruction result.

Keywords: computational metrology, optical metrology, computational lithography, nanolithography, finite-element methods, nanooptics

1. INTRODUCTION

Optical scatterometry is a method to measure the size and shape of periodic micro- or nanostructures on surfaces.[1] For this purpose the geometry parameters of the structures are obtained by reproducing experimental measurement results through numerical simulations. Such simulations are typically performed using parameterized models and nonlinear optimization algorithms to find parameter settings which minimize differences between the measurements and the numerically obtained result. Often also knowledge about experimental measurement errors and model uncertainties is available, as well as prior knowledge obtained from additional measurements. This can be used to also quantify uncertainties of the reconstructed parameter values.[2–4]

Here we demonstrate an efficent method for parameter reconstruction and uncertainty quantification using a Newton method to solve the inverse problem, an efficient finite-element based solver for the forward-problem, and a Bayesian approach for relating measurement uncertainties and prior knowledge to the reconstruction results. The paper is structured as follows: An optical scatterometry setup which serves as application example is presented in Section 2. The reconstruction method is discussed in Section 3. Results are presented in Section 4. Limits and caveats of using difference quotients to compute partial derivatives are addressed in the Appendix.

2. SCATTEROMETRIC SETUP

In order to test our method for solving the inverse problem we use an experimental scatterometry data set obtained at PTB. Details of the measurement configuration have recently been reported.[5] Briefly, in the experiment, a silicon grating (1D periodic lines) with nominal pitch of $p_x = 50\,\text{nm}$ and nominal linewidth of $CD = 25\,\text{nm}$ is used as scattering target in a goniometric setup with an inspection wavelength of $\lambda = 266\,\text{nm}$. A schematic of the measurement is shown in Figure 1 (left). A collimated light beam at well defined polarization and angle of incidence (inclination angle θ, rotation angle ϕ) illuminates the target. The intensity of specular reflection is recorded. Spectra depending on inclination angle θ are recorded for S- and for P-polarization, and for rotation angles of $\phi = 0$ and $\phi = 90$.

The measured data set used in this study is plotted in Figure 2 (circles).
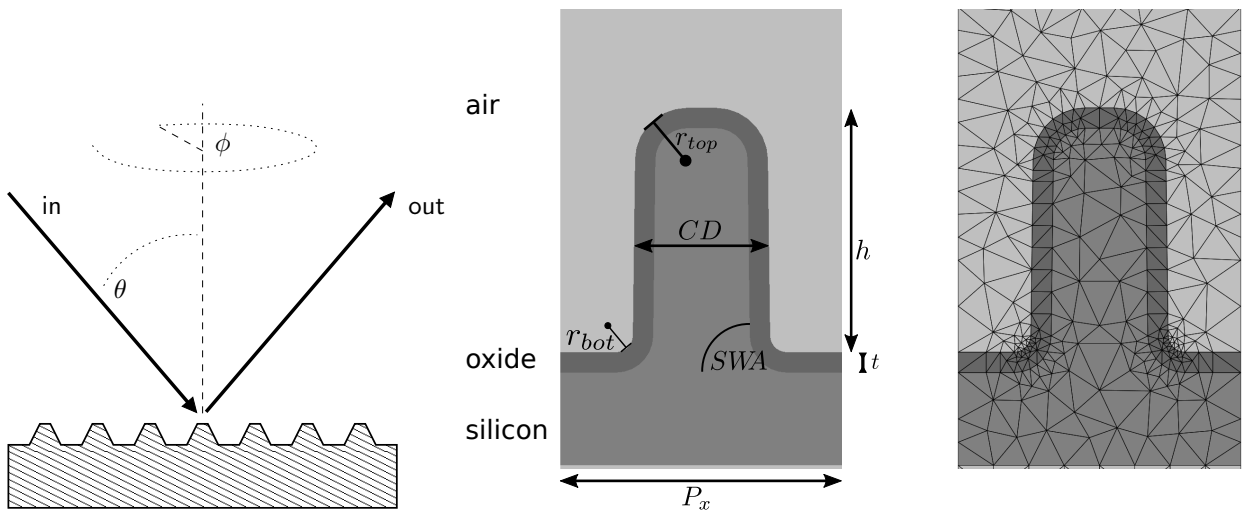
Figure 1. *Left:* Schematics of the experimental 2θ setup with incidence angle θ and azimuthal orientation ϕ. *Center:* Schematics of the model of a unit cell of the silicon line grating with free parameters line height, h, critical dimension at $h/2$, CD, oxide thickness, h_{ox}, sidewall angle SWA, top and bottom corner roundings, r_{top} and r_{bot}. *Right:* Visualization of the triangular mesh for the FEM discretization.

3. RECONSTRUCTION METHOD

The reconstruction method described in the following employs a Bayesian perspective[6] to compute not only the most likely geometrical parameters describing the shape of the line grating, but also to quantify the uncertainties inherent in the process. We start by describing the forward model, that is the modeling of the scattering signal given a set of parameters x. We address some specifics of the electromagnetic field solver and finally give a detailed insight into the Gauss-Newton method used to reconstruct the parameters.

Forward model. The forward model describes how the obtained measurements $y \in \mathbb{R}^m$ depend on the microstructure's parameters $x \in \mathbb{R}^n$. The computational model aims to describe the experimental setup outlined in Section 2. The measurement of the various combinations of incidence angles, polarizations, and azimuthal

orientations are serialized and form the model evidence vector y_M. Each entry corresponds to a solution of the forward model for one of the given combinations. We parameterize the shape of the silicon line with a total of $n = 6$ free parameters: the line height h, the line width (critical dimension) at $h/2$, CD, the oxide layer thickness, h_{ox}, the sidewall angle, SWA, and the top and bottom corner roundings, r_{top} and r_{bot}. Their definitions can be found in Figure 1. In the reconstruction we allow for parameter values within large intervals describing a wide range of line shapes. The bounds employed are listed as ranges in Table 1 for all free parameters. However, not all combinations of parameters out of these intervals lead to meaningful line shapes. For example, we demand that the corner rounding radius at the top of the line is smaller than half the width at the top. This restricts the choice of parameter configurations as the parameters are no longer independent of each other.

The line grating is then modeled as a 2D unit cell in the cross section of an infinitely periodic array by employing periodic boundary conditions at the boundaries in $\pm x$ direction, with a fixed periodicity of $p_x = 50$ nm. The silicon substrate and air superstrate are modeled as infinite half-spaces by employing transparent boundary conditions at the top and bottom of the computational domain shown in Figure 1.

The scattering of monochromatic light off the nanoscopic line grating is described by the linear Maxwell's equations in frequency domain. These lead to a single second order partial differential equation

$$\nabla \times \mu^{-1} \nabla \times E - \omega^2 \varepsilon E = 0 \tag{1}$$

where ε and μ are the permittivity and permeability tensors, and ω is the time-harmonic frequency. Here, impressed current sources are implicitly neglected. The materials are non-magnetic (μ_r=1.0) and refractive indices for silicon, the top oxide layer and air of $n_{\text{Si}} = 1.967 + 4.443i$, $n_{\text{ox}} = 1.7212 + 0.113i$ and $n_{\text{air}} = 1.0$, respectively, are used.

Maxwell solver. Analytical solutions of Maxwell's equations are not available on the complex parameterized geometries under consideration. Therefore, a numerical simulation of the measurement process is necessary. We employ the finite-element (FEM) electromagnetic field solver JCMsuite[7,8] which has been successfully used in scatterometric investigations ranging from the optical[9] to the EUV and X-ray regimes[10,11] on 2D (e.g., line masks) and 3D (e.g., FinFETs, contact holes) scattering targets.

To solve the scattering problem outlined above, the computational domain is split into interior domain $\Omega_{\text{CD}} \subset \mathbb{R}^3$ and an exterior domain Ω_{ext} which hosts the incoming electromagnetic fields and the outgoing scattered field. Waves incident on Ω_{CD} are added to the right hand side of Equation (1) in the appriopriate manner.[8] Transparent boundary conditions are realized by the perfectly matched layer (PML) method. The finite element method employs polynomials of order p with local support as ansatz functions over each element with a mesh size h. Transforming Equation (1) into a weak formulation by multiplication with a test function and subsequent integration leads to a linear equation

$$\mathcal{A}E = f \tag{2}$$

for the N coefficients E of the FEM ansatz functions. Here, $\mathcal{A}$ is a sparse matrix and f contains contributions of incoming fields.

We use a workflow for the Maxwell solver starting and ending in the scripting language Matlab. Within a typical Matlab function call, geometrical, material and source parameters are defined, program calls for mesh generation and solutions of the linear system, Eq. (2), are distributed to various computation cores of a workstation or cluster, and the results are automatically collected and post-processed.

Inverse problem. For reconstructing likely parameters and quantifying the uncertainty of the reconstruction, we take a Bayesian perspective.[12] In this framework, the conditional probability $\pi(x|y)$ of a certain parameter vector x given the acquired measurement vector y, also known as posterior probability, is given by Bayes' theorem as the product of likelihood $\pi(y|x)$ and prior probability $\pi(x)$. With a sufficiently parameterized model capable

of representing the actual geometry, the likelihood describes essentially the measurement errors. Here we assume independently normally distributed errors with zero mean and diagonal covariance Γ_l:

$$\pi(y|x) \sim \exp\left(-\frac{1}{2}(y - y(x))^T \Gamma_l^{-1}(y - y(x))\right) \tag{3}$$

The parameters defining the geometric shape of the microstructure have to stay in an admissible bounded region in order to avoid non-physical self-intersection. The admissible region can in general be defined as $X = \{x \in \mathbb{R}^m \mid g_i(x) \geq 0, i = 1, \dots, r\}$ with smooth scalar valued functions g_i. Imposing little bias within X, but assuming that parameters close to the boundary ∂X are less likely to occur in practice, we define a prior density as

$$\pi(x) \sim \exp\left(\sum_{i=1}^{r} \mu_i \log g_i(x)\right) \tag{4}$$

for penalty factors $\mu_i > 0$. The larger these factors are, the more bias towards the analytic center of the feasible region is imposed by the prior density. To facilite the computation of derivatives of the prior, we rely heavily on the use of automatic differentiation.[13]

From the resulting posterior density $\pi(x|y) = \pi(y|x)\pi(x)$ reasonable point estimates can be computed. We use the maximum posterior estimate $x_{\text{MAP}} = \arg\max_{x \in X} \pi(x|y)$. Taking the logarithm of $\pi(x|y)$, this can be computed by solving the minimization problem

$$x_{\text{MAP}} = \arg\min_{x \in X} F(x), \quad F(x) = \frac{1}{2}(y - y(x))^T \Gamma_l^{-1}(y - y(x)) - \sum_{i=1}^{r} \mu_i \log g_i(x). \tag{5}$$

Here we employ a Gauß-Newton like method[14] for computing local minimizers of F. The objective's first and second derivatives are

$$F'(x) = (y - y(x))^T \Gamma_l^{-1} y'(x) - \sum_{i=1}^{r} \frac{\mu_i}{g_i(x)} g'(x)$$

and

$$F''(x) = y'(x)^T \Gamma_l^{-1} y'(x) + (y - y(x))^T \Gamma_l^{-1} y''(x) + \sum_{i=1}^{r} \left(\frac{\mu_i}{g_i(x)^2} g_i'(x)^T g_i'(x) - \frac{\mu_i}{g_i(x)} g_i''(x)\right).$$

For good data reproduction (i.e., small $\|y - y(x)\|$) and parameters not close to the boundary of the admissible region (i.e., small $\mu_i/g_i(x)$) the second order terms can be neglected, yielding a positive definite Hessian approximation close to the minimizer:

$$\hat{F}''(x) = y'(x)^T \Gamma_l^{-1} y'(x) + \sum_{i=1}^{r} \frac{\mu_i}{g_i(x)^2} g_i'(x)^T g_i'(x) \approx F''(x)$$

Starting at $x_0 \in X$, the Gauß-Newton method suggests steps $\Delta x_k = -\hat{F}''(x_k)^{-1} F'(x_k)$ which due to positivity of $\hat{F}''(x)$ are descent directions, and performs updates $x^{k+1} = x_k + \alpha_k \Delta x_k$ with a step length $\alpha_k \in \,]0, 1]$ determined by line search such that monotone decrease of the objective is guaranteed.

The Taylor approximation

$$F(x) \approx \hat{F}(x) = F(x_{\text{MAP}}) + F'(x_{\text{MAP}})(x - x_{\text{MAP}}) + \frac{1}{2}(x - x_{\text{MAP}})^T \hat{F}''(x_{\text{MAP}})(x - x_{\text{MAP}}) \tag{6}$$

provides a local Gaussian approximation of the posterior by $\pi(x|y) \approx c \exp(-\hat{F}(x))$ and thus allows a local uncertainty quantification in terms of the covariance $\hat{\Gamma}_p = \hat{F}''(x_{\text{MAP}})^{-1}$. The corresponding variances $\hat{F}''(x_{\text{MAP}})_{ii}^{-1}$ of the marginal posterior distributions of the individual parameters x_i then give an idea of how reliably identified the parameters are. A more concise information about how well parameters are identified can be inferred from the eigenvalues and eigenvectors of the covariance $\hat{\Gamma}_p$, giving linear combinations of parameters that are more or less reliably estimated.

4. RECONSTRUCTION RESULTS

We present the results of a performed reconstruction with a random starting point in section 4.1. In section 4.2 we study the effect of numerical discretization accuracy on the reconstruction results.

4.1 Geometry reconstruction using a Gauß-Newton method

We start the reconstruction algorithm outlined in the previous section at a randomly chosen starting point x_0 with $CD = 23.8262$ nm, $h = 43.1662$ nm, $SWA = 89.2796°$, $t = 3.5963$ nm, $r_{\text{top}} = 9.2135$ nm and $r_{\text{bot}} = 2.9608$ nm. The measurement errors are assumed to be normally distributed with zero mean and relative standard deviation of 2%, thus fixing Γ_l and the likelihood in (3). The bounds employed in the prior are given as ranges in Table 1 for all parameters. The coupled constraints on the corner rounding radii are also included in the prior as outlined in Section 3, Equation (4).

The algorithm converges in just six Gauß-Newton iterations to the desired relative accuracy of $1 \cdot 10^{-4}$ in the differential. The resulting parameter values in the maximum posterior estimate x_{MAP} are $CD = 25.3785$ nm, $h = 48.0835$ nm, $SWA = 86.9803°$, $t = 4.9392$ nm, $r_{\text{top}} = 10.3685$ nm and $r_{\text{bot}} = 4.7940$ nm (cf., Table 1). In Figure 2 the experimental and simulated intensities in the 0th diffraction order are shown as function of inclination angle θ. The four different angular spectra refer to the different polarizations and azimuthal orientations of the illumiation. In Fig. 2 (left) the simulated data for the initial configuration x_0 is shown together with a plot of the differences between measured and simulated data points at the bottom. In Fig. 2 (right) the same is shown with the simulations results for x_{MAP}. We observe an almost perfect alignment of the simulated data for x_{MAP} and the measurements. Deviations are much smaller than in the initial case on the left. Note the different scales of the y-axis in the deviation plots at the bottom.

From the local Gaussian approximation of the posterior distribution (cf., Equation (6)) we can quantify the uncertainty of the reconstructed parameter values in terms of standard deviations. We find low standard deviations of 0.3984 nm and 0.1615 nm for the line width CD and the oxide layer thickness t. The side wall angle SWA has a standard deviation of 0.9988°. The height h and the two corner rounding radii r_{top} and r_{bottom} show larger uncertainties of 2.4838 nm, 4.2892 nm and 3.2171 nm respectively. These results agree also within the uncertainty limits with a recent, different evaluation using a slightly different model.[15]

4.2 Numerical convergence study

In this section we study the impact of numerical accuracy of the forward model on the reconstructed parameter values and their uncertainties. The numerical accuracy of the solution of the forward problem essentially depends on the mesh width and on the polynomial degree of the ansatz functions p. The small features sizes and layer thicknesses require a reasonably fine meshing of the unit cell geometry as shown in Figure 1 (right). This constitutes a lower limit on the mesh finesse provided we aim for a quality mesh which fulfills the Delaunay criterion. Hence in this study, we keep the mesh fixed and increase the polynomial degree p from 1 to 5. For each p we perform the reconstruction using the same starting point x_0 listed above and observe the computed x_{MAP}. The results of this study are shown in Table 1. The results for $p > 2$ seem very well converged and show only small variations in x_{MAP} as well as the standard deviations compared to $p = 3$.

	CD [nm] (15, 35)	h [nm] (40, 60)	SWA [deg] (84, 90)	t [nm] (2.5, 6.5)	r_{top} [nm] (3.5, 18)	r_{bot} [nm] (0, 10)
Range						
p=1	25.0365 (0.2121)	49.6280 (1.0519)	87.1857 (0.8990)	4.7531 (0.1295)	10.6567 (1.4282)	6.9962 (1.7699)
p=2	25.3741 (0.2840)	48.0863 (1.8411)	86.9918 (0.9447)	4.9368 (0.1367)	10.3858 (2.8578)	4.7964 (2.3854)
p=3	25.3786 (0.4006)	48.0828 (2.4917)	86.9801 (1.0002)	4.9392 (0.1621)	10.3688 (4.3045)	4.7927 (3.2360)
p=4	25.3785 (0.3993)	48.0833 (2.4873)	86.9803 (0.9994)	4.9392 (0.1618)	10.3687 (4.2965)	4.7933 (3.2252)
p=5	25.3785 (0.3984)	48.0835 (2.4838)	86.9803 (0.9988)	4.9392 (0.1615)	10.3685 (4.2892)	4.7940 (3.2171)

Table 1. Reconstruction results with estimated standard deviations (in parentheses) for numerical settings with increasing accuracy of the forward-problem ($p = 1$ to $p = 5$). The admissible parameter regions in the reconstruction (Range) are indicated for all parameters.

We distribute the independent forward problems on a parallel computing cluster. Hence the computation time for all 168 data points considered in each iteration of the Gauß-Newton reconstruction method is limited by
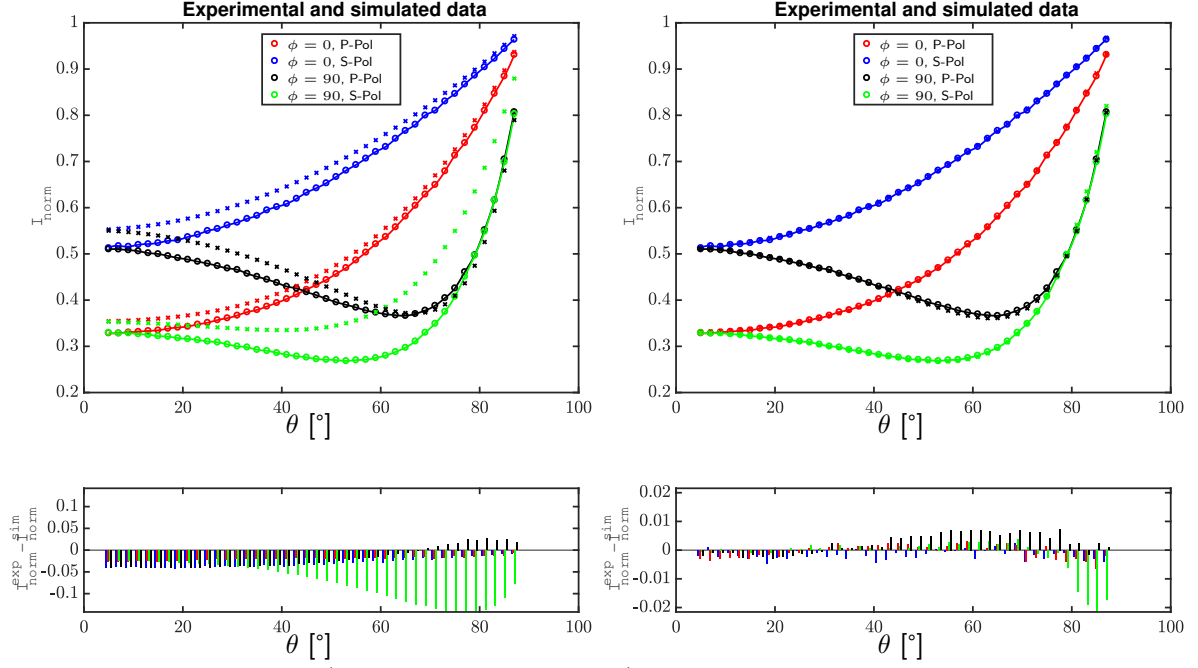
Figure 2. *Left:* Experimental data (circles and connecting lines) and simulated data for the random initial configuration x_0 (crosses). We observe a general qualitative alignment of simulated data and measured data. The plot at the bottom shows the difference between measured and simulated signals. Largest deviations are observed for the S-polarization, $\phi = 90°$ and large inclination angles θ. *Right:* Experimental data (circles and connecting lines) and simulated data for the x_{MAP} configuration (crosses). We observe a very good quantitative alignment of the data. Deviations are much smaller than in the initial case. Note the different scales of the y-axis in the deviation plots at the bottom.
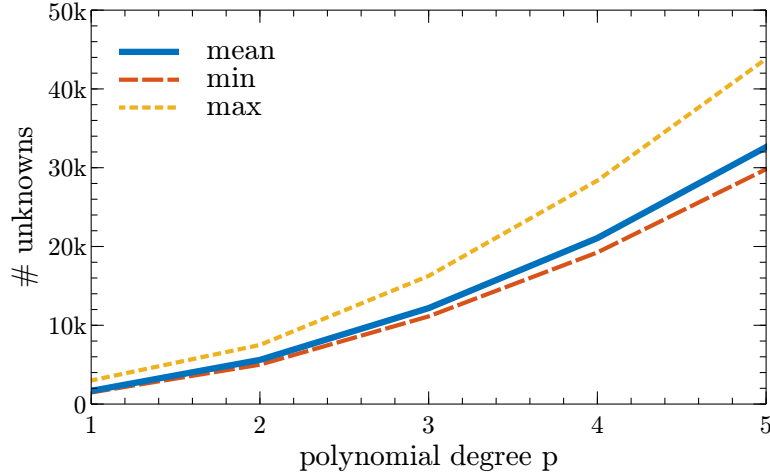

Figure 3. Number of unknowns for increasing polynomial degree used in the forward-problem computations ($p = 1$ to $p = 5$). The mean over all simulations in the reconstruction is shown in blue, yellow and orange mark the minimum and maximum encountered.

the time of a single FEM evaluation. The average numerical effort required by the forward problems increases from 2047 unknowns for $p = 1$ to 35863 for $p = 5$. In Figure 3 the number of unknowns is shown as a function of the polynomial degree p. The mean of all 168 simulation is given by the blue line. The PML thickness is automatically adjusted by the solver JCMsuite[7] by considering the angle of incidence, the material distribution in the exterior and the illumination wavelength. In combination with a remeshing for every iteration in the reconstruction, this leads to differences in the number of degrees of freedom over the angle of incidence scan. In Figure 3 orange and yellow lines mark the minimum and maximum the degrees of freedom observed in all

simulations performed during the reconstruction. The solution of these relatively small FEM problems including post processing takes at most 7 seconds of CPU time. Based on these findings it seems reasonable to choose
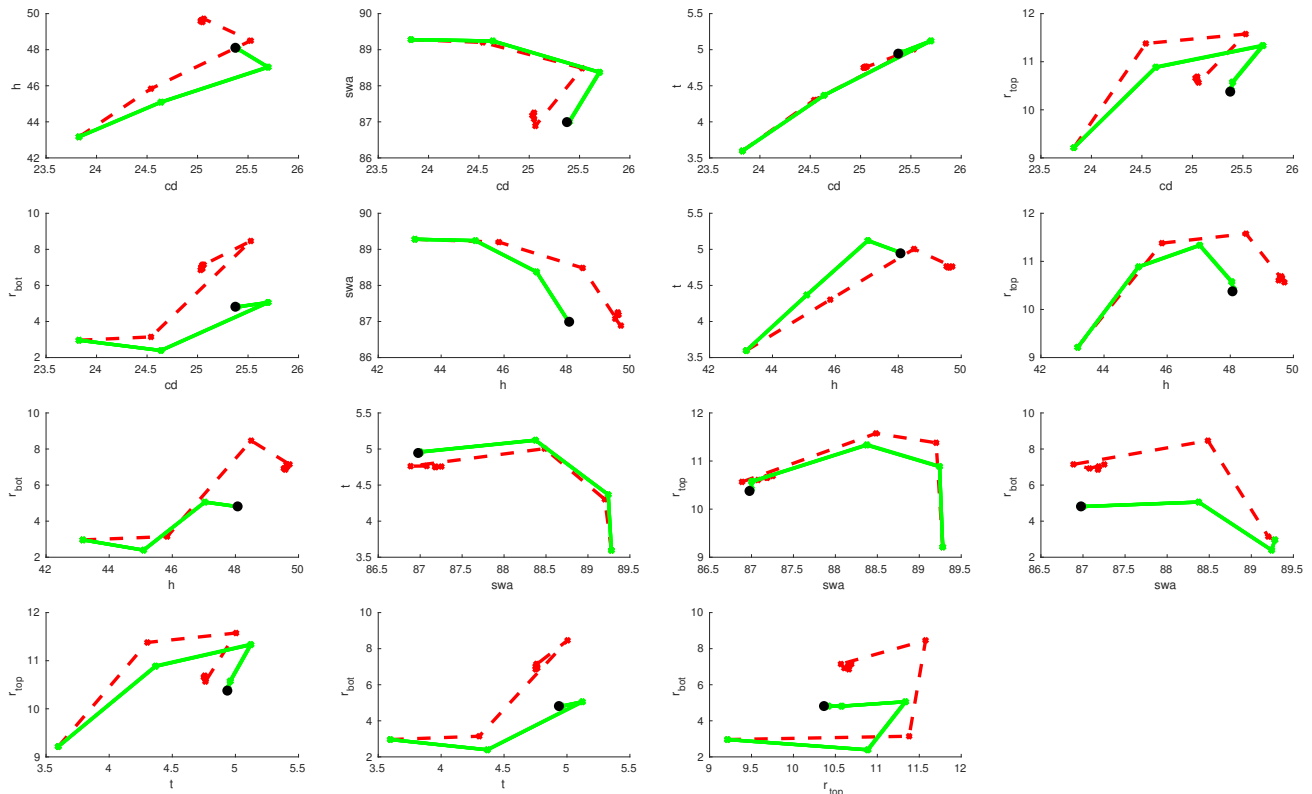

Figure 4. Pair-wise projections of the trajectories of the Gauß-Newton method in parameter space. Converged trajectories for $p >= 2$ are shown in green whereas the trajectory for $p = 1$ is shown in red. The iterates are indicated by crosses which for the green trajectories are almost identical. The location of x_{MAP} is indicated by a black dot covering the last iterates.

Investigating the influence of the polynomial degree in more detail, we look at the trajectories of the Newton method in parameter space. These are shown as pair-wise projections in Figure 4. The graphs of the converged trajectories for $p >= 2$ are shown in green whereas the trajectory for $p = 1$ is shown in red. We observe no visible distinction between the converged results. All four trajectories include six data points respectively, indicated by crosses in the plots. These are almost indistinguishable as all descent directions found by the Newton method are virtually identical and only differ slightly in step length. In most subplots the last two steps taken are too small to be resolved and are covered by the black circles marking $x_{\mathrm{MAP}}^{p=5}$. The trajectory for $p = 1$ (red, dotted line) is distinctly different as it requires 9 steps and converges toward a different estimate.

5. CONCLUSION

The parameters describing the shape of a silicon line grating and associated uncertainties have been found by means of an Bayesian inverse problem solved with a Gauß-Newton method. The agreement of experimental measurements and the numerical model is very good. The Bayesian perspective and a quadratic approximation of the posterior distribution allows for a quantification of the parameter uncertainties in terms of standard deviations of a Gaussian centered at the reconstructed parameters. The employed Gauß-Newton method to solve the inverse problem demonstrated good convergence properties and robustness. An in depth analysis of the numerical reconstruction method will follow in a subsequent paper.

Acknowledgments

REFERENCES

[1] Pang, L., Peng, D., Hu, P., Chen, D., He, L., Li, Y., Satake, M., and Tolani, V., "Computational metrology and inspection (CMI) in mask inspection, metrology, review, and repair," *Adv. Opt. Techn.* **1**, 299 (2012).

[2] Groß, H., Rathsfeld, A., Scholze, F., and Bär, M., "Profile reconstruction in extreme ultraviolet (EUV) scatterometry: Modelling and uncertainty estimates," *Meas. Sci. Technol.* **20**, 105102 (2009).

[3] Henn, M.-A., Gross, H., Scholze, F., Wurm, M., Elster, C., and Bär, M., "A maximum likelihood approach to the inverse problem of scatterometry," *Opt. Express* **20**, 12771 (2012).

[4] Klauenberg, K., Wübbeler, G., Mickan, B., Harris, P., and Elster, C., "A tutorial on bayesian normal linear regression," *Metrologia* **52**, 878 (2015).

[5] Wurm, M., Bonifer, S., Bodermann, B., and Richter, J., "Deep ultraviolet scatterometer for dimensional characterization of nanostructures: system improvements and test measurements," *Meas. Sci. Technol.* **22**, 094024 (2011).

[6] Sullivan, T., [*Introduction to Uncertainty Quantification*], Springer (2015).

[7] JCMwave GmbH, "Program package JCMsuite." http://www.jcmwave.com.

[8] Pomplun, J., Burger, S., Zschiedrich, L., and Schmidt, F., "Adaptive finite element method for simulation of optical nano structures," *phys. stat. sol. (b)* **244**, 3419 (2007).

[9] Potzick, J., Dixson, R., Quintanilha, R., Stocker, M., Vladar, A., Buhr, E., Häßler-Grohne, W., Bodermann, B., Frase, C., and Bosse, H., "International photomask linewidth comparison by NIST and PTB," *Proc. SPIE* **7122**, 71222P (2008).

[10] Scholze, F., Laubis, C., Dersch, U., Pomplun, J., Burger, S., and Schmidt, F., "The influence of line edge roughness and CD uniformity on EUV scatterometry for CD characterization of EUV masks," *Proc. SPIE* **6617**, 66171A (2007).

[11] Soltwisch, V., Haase, A., Wernecke, J., Probst, J., Schoengen, M., Burger, S., Krumrey, M., and Scholze, F., "Correlated diffuse x-ray scattering from periodically nanostructured surfaces," *Phys. Rev. B* **94**, 035419 (2016).

[12] Kaipio, J. and Somersalo, E., [*Statistical and Computational Inverse Problems*], Springer (2005).

[13] Griewank, A. and Walther, A., [*Evaluating Derivatives, Principles and Techniques of Algorithmic Differentiation, second edition*], SIAM, Philadelphia (2008).

[14] Deuflhard, P., [*Newton Methods for Nonlinear Problems. Affine Invariance and Adaptive Algorithms*], vol. 35 of *Computational Mathematics*, Springer (2006).

[15] Wurm, M., Endres, J., Probst, J., Schoengen, M., Diener, A., and Bodermann, B., "Metrology of nanoscale grating structures by UV scatterometry," *Opt. Express* **25**, 2460 (2017).

[16] Burger, S., Zschiedrich, L., Pomplun, J., Schmidt, F., and Bodermann, B., "Fast simulation method for parameter reconstruction in optical metrology," *Proc. SPIE* **8681**, 868119 (2013).

<h1 style="text-align:center">Appendix</h1>

Pitfalls in using difference quotients to compute partial derivatives

Computation of partial derivatives of the electric field and derived quantities is required for both, the Newton method for solving the inverse problem, and for local uncertainty quantification at x_{MAP}. Here we compare two methods of how to compute first and second order derivatives numerically, direct computation and computation using finite differences, and we show the impact of insufficient computation of these on the reconstruction results.

For direct computation the system matrix can be employed to compute partial derivatives $\frac{\partial E}{\partial x_i}$ efficiently as reported previously.[16] This can be seen easily by differentiating the linear system $AE = f$ which gives

$$\frac{\partial E}{\partial x_i} = A^{-1} \left(\frac{\partial f}{\partial x_i} - \frac{\partial A}{\partial x_i} E \right). \tag{7}$$

Similarly, higher-order partial derivatives are accessible recursively.

Alternatively, a finite-difference scheme can be used:

$$\frac{df}{dx_i} = \frac{f(x + \delta x_i) - f(x - \delta x_i)}{2\delta x_i}. \tag{8}$$

$$\frac{d^2 f}{dx_i^2} = \frac{f(x + \delta x_i) - f(x) + f(x - \delta x_i)}{(\delta x_i)^2}. \tag{9}$$

$$\frac{d^2 f}{dx_i dx_j} = \frac{f(x + \delta x_i + \delta x_j) - f(x - \delta x_i + \delta x_j) - f(x + \delta x_i - \delta x_j) + f(x - \delta x_i - \delta x_j)}{4\delta x_i \delta x_j}. \tag{10}$$

It can be shown that from a computational point of view it is generally more efficient to directly compute the partial derivatives than to compute additional data points for difference quotients. Irregardless of their computational efficiency we want to investigate the influence of finite difference approximations to the partial derivatives. To this extent, we have repeated the study of Section 4 by replacing the computation of the second order partial derivatives used to determine the standard deviations with a finite difference approximation.

	CD [nm]	h [nm]	SWA [deg]	t [nm]	r_{top} [nm]	r_{bot} [nm]
p=1	25.0106 (-)	49.7642 (0.8392)	86.7925 (-)	4.7325 (-)	10.3226 (0.1500)	6.4657 (1.2493)
p=2	25.3728 (0.3429)	48.0927 (1.2692)	86.9852 (1.3253)	4.9361 (0.1858)	10.3824 (1.9281)	4.7951 (2.2815)
p=3	25.3788 (0.3472)	48.0824 (1.2579)	86.9810 (1.3285)	4.9393 (0.1867)	10.3694 (1.9056)	4.7939 (2.3644)
p=4	25.3787 (0.3481)	48.0827 (1.2572)	86.9811 (1.3309)	4.9393 (0.1871)	10.3690 (1.9111)	4.7945 (2.3680)
p=5	25.3787 (0.3493)	48.0829 (1.2575)	86.9811 (1.3344)	4.9393 (0.1876)	10.3689 (1.9234)	4.7952 (2.3685)

Table 2. Reconstruction results with estimated standard deviations (in brackets) for numerical settings with increasing accuracy of the forward-problem. The standard deviations are not reliable in this case as they are computed in a not sufficiently converged finite difference scheme.

The results of this scan are listed in Table 2. Unsurprisingly we get the same values for x_{MAP} as the quasi Newton iterations are not influenced by our change. The standard deviations however are changing. In case of low numerical accuracy ($p = 1$) the standard deviations for the CD, the SWA and t are not reliably computed. Even though we seem to observe convergence in this quantity with increasing p, we note difference to the values found with exact second order partial derivatives listed in Table 1. The standard deviations for CD and t are relatively similar whereas the larger standard deviations for h, r_{top} and r_{bot} change drastically by factors of almost two.

Convergence of first order derivatives: We studied the convergence of first order partial derivatives of the 0th order intensity using equation (8) with respect to the ones obtained with equation (7) for one specific setting of the forward problem ($p = 5$, $\theta = 45°$). The results are shown in Figure 5. The convergence with respect to the perturbation δx_i is shown for both polarizations, the different colors indicate the parameters. For both polarizations exponential convergence can be observed for larger values of δx_i. At a relative accuracy of approximately $1 \cdot 10^{-5}$ the behavior changes and the relative errors remain at or sligthly below this value.
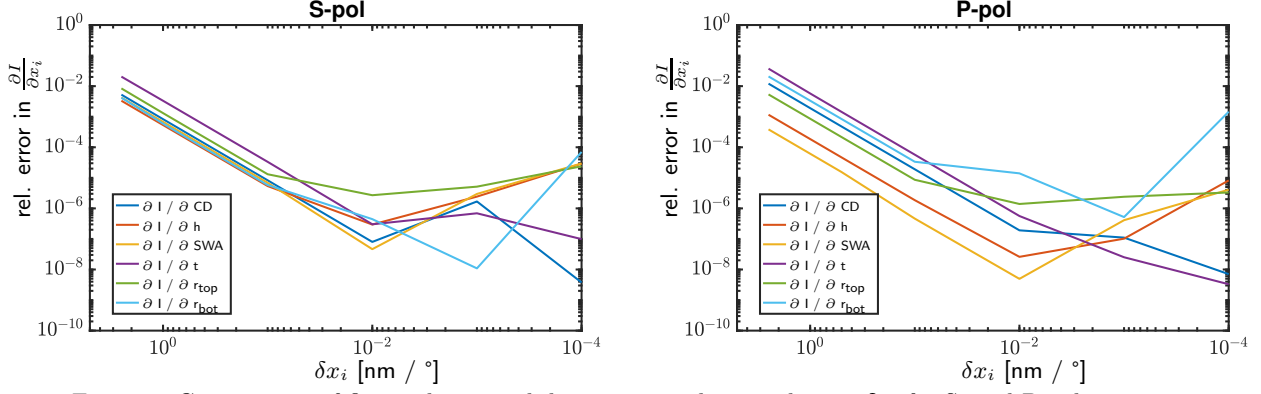
Figure 5. Convergence of first order partial derivatives with perturbation δx_i for S- and P-polarization.

Convergence of second order derivatives: We studied the convergence of second order derivatives in the same setup as for the first order derivatives. The convergence with δx_i is shown in Figure 5. All quantities exhibit exponential convergence provied the perturbation δx_i is larger than $1 \cdot 10^{-1}$. The relative error for this value is below $1 \cdot 10^{-4}$ except for $\frac{d^2 I}{dSWA^2}$ where it is sligthly larger for the P-polarization. However, instead of decreasing further the relative error increases exponentially with smaller δx_i. This can be explained by on one hand by the smaller and smaller changes introduced in the measured signal (0th order reflected intensity) which is captured by the numerators in equation (9) and (10). The scaling by the denominator then leads to the exponential increase observed in Figure 6.
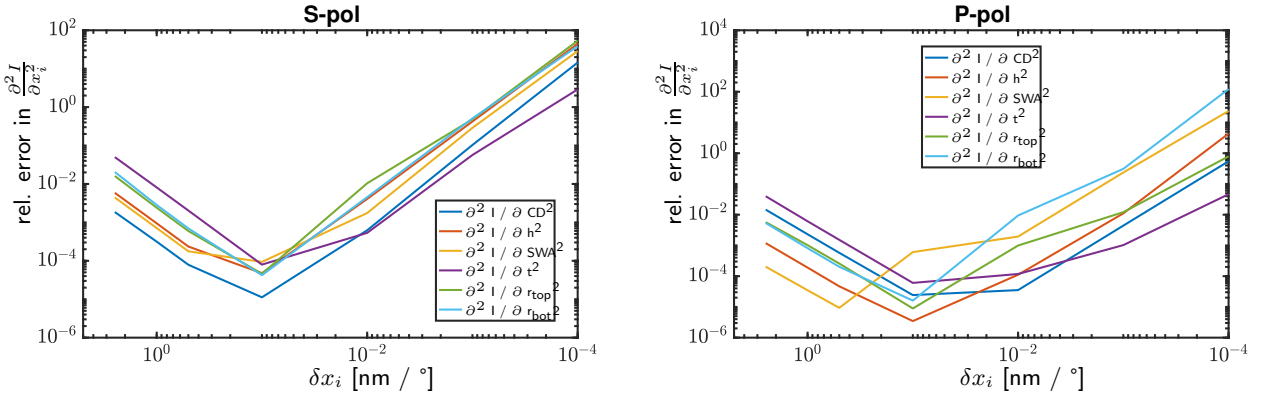

Figure 6. Convergence of second order partial derivatives with perturbation δx_i for S- and P-polarization.

These studies serve to explain the differences observed between the standard deviations in Tables 2 and 1: The employed perturbations ($\delta x_i \in [1 \cdot 10^{-2}, 4 \cdot 10^{-2}]$ depending on the magnitude of the different parameters) are large enough to provide accurate approximation of the first order partial derivatives thus ensuring the convergence of the method. However, the computational accuracy of the second order partial derivatives are insufficient for these perturbations, especially for those parameters with very small second order derivatives. The computed standard deviations in Table 2 thus deviate more significantly from those listed in Table 1.