Optimizing Nanoscale Quantitative Optical Imaging of Subfield Scattering Targets

MARK-ALEXANDER HENN[1,*], BRYAN M. BARNES[1], HUI ZHOU[1], MARTIN SOHN[1], AND RICHARD M. SILVER[1]

[1] Engineering Physics Division, National Institute of Standards and Technolgy (NIST), 100 Bureau Drive, Gaithersburg, MD, USA 20899
[*] Corresponding author: mark.henn@nist.gov

Compiled September 22, 2016

The full 3-D scattered field above finite sets of features has been shown to contain a continuum of spatial frequency information, and with novel optical microscopy techniques and electromagnetic modeling, deep-subwavelength geometrical parameters can be determined. Similarly, by using simulations, scattering geometries and experimental conditions can be established to tailor scattered fields that yield lower parametric uncertainties while decreasing the number of measurements and the area of such finite sets of features. Such optimized conditions are reported through quantitative optical imaging in 193 nm scatterfield microscopy using feature sets up to four times smaller in area than state-of-the-art critical dimension targets.

OCIS codes: *(120.3940) Metrology; (290.3700) Linewidth; (180.6900) Three-dimensional microscopy.*

http://dx.doi.org/10.1364/ao.XX.XXXXXX

Continuous advances in photolithographic technology, processes, and materials have led to a downward scaling of the critical dimensions (CDs) of semiconductor devices. These CDs, often correlating to the line width of the features of interest, are presently below 20 nm and given current trends will likely reach the atomic scale in the mid-2020s [1]. Current metrology techniques are being refined to meet the challenges presented by such small features.

The industrial workhorse for CD metrology is optical scatterometry [2], as optics provides lower cost, greater areal coverage, and non-destructive measurements. The measurement of the CDs for features of interest is interpolated by evaluating the optical scattering from several multi-line arrays positioned on the scribe lines, outside the active device area. When the incident light underfills the target, the array in simulations can be treated as a grating. When the beam overfills the target, "spurious" scattering and reflections arise that must be dealt with, although some metrology systems collect not only the 0^{th} order-scattering but also the $\pm 1^{st}$ orders to augment their CD measurements [2]. Research is leading to reductions in the size of scatterometry targets, with some recent projections of targets as small as 12 μm x 12 μm in area for CD scatterometry [3] and 10 μm x 10 μm for diffraction-based overlay metrology [4].

A recent paper [5] published by our group has demonstrated quantitative critical dimension measurements as small as 16 nm with parametric uncertainties as small as 1 nm or less. Three 30-line arrays with deep-subwavelength dimensions were measured quantitatively, with the narrowest of these lines being approximately 30 times smaller than the wavelength of the light, λ = 450 nm, used to measure them. Quantitative measurement of these features was enabled by choosing a geometrical model from limited prior information, completing several electromagnetic simulations, using an in-house implementation of the rigorous coupled-wave analysis (RCWA) [6], as functions of the parameters of that model, normalizing the simulated scattered fields using the observed tool functions, calculating images from those normalized fields, and estimating the model parameters and their uncertainties, including potential correlations, using nonlinear regression. Innovations in structured illumination [7], tool characterization and Fourier domain normalization [8], systematic error estimation, *a priori* information, and 3-D scattered light field analysis were all critical to unlocking deep-subwavelength information from sets of images acquired through-focus.

In this Letter we optimize this alternative optical methodology for $\lambda = 193$ nm, the shortest wavelength used with the National Institute of Standards and Technology's (NIST) scatterfield microscopes [7]. This technique uses the broad continuum of scattered spatial frequencies that is inherent to a finite grating in order to parametrically determine the CDs of a finite array of features. The software used in the present study is the commercially available JCMsuite[9], a solver for time-harmonic Maxwell's equations and other applications using the finite element method (FEM). The FEM approach together with the use of perfectly matched layers as absorbing boundaries makes it possible to investigate the scattering from a variety of non-periodic 2-D and 3-D geometries. For these simulations Köhler illumination is assumed, hence the source can be treated as a superposition of 12 plane waves, reducing simulation complexity. With real space imaging, pairing a low illumination numerical aperture (INA) with a high collection numerical aperture (CNA) (here, INA = 0.1, CNA = 0.95) yields spatially resolved intensity variations that are essential for determining the subwavelength

dimensions of subfield-of-view targets with this methodology.

Key considerations in choosing the geometric models for the simulation studies were simplicity and applicability to the semiconductor industry, as presented in Fig. 1. The first model is a rectangular, pure *Si* (n=0.85, k=2.61) cross-section with a native SiO_2 (n=1.56, k=0.0) coating, and only two geometry parameters, the rectangle's width and height. However, since that geometry is not reflective of actual manufacturing geometries, a second, more realistic, geometry model with fins buried in a SiO_2 layer was used. In the interests of simplicity again a single height and single width parameter were floated with the sidewall angle of the buried structure, the optical constants, and the SiO_2 oxide height all fixed. Generally speaking the parametric correlation between the height and width parameters should be low in both cases.

Subsequent imaging of the scattering structures requires taking only the far field data into account. Here, the Fourier transform corresponding to the returning part of the total field is determined and the Fourier spectrum falling within the CNA is used as input to propagate the field and calculate the images at different focus heights relative to the substrate. The sampling rate in the x direction corresponds to a pixel size of 25 nm with an overall range of 3 μm between -1.5 μm to 1.5 μm relative to the center of the structure. Together with the above specifications this configuration leads to an individual data set for each simulated intensity profile, at a given focus height that consists of 120 data points. A representative scattering intensity profile is shown as Fig. 1 (d). A library containing input files for JCMsuite and example data for the above simulation setup are shown in Code File 1 [10].

Often an experimental measurement of linear arrays is simulated using a 2-D geometry for faster computation, which implicitly assumes the target to be infinitely long. Optimizing the length of these subfield scattering targets requires testing of this approximation. Figure 2 shows the root-mean square deviation (RMSD) between the scattering intensity profiles for an infinite structure and four finite structures. For Fig. 2, we assume no natural oxide to be present for either structure to keep the 3-D simulations within our memory constraints (512 GB). The varied parameter in this study is the length of the finite line array, i.e., the extension in y direction (cf. Fig. 1), of 10 rectangular lines at two orthogonal polarizations and 21 equally spaced focus heights. Out-of-focus data were used in Ref. [5] over a 9 λ range which leads here to a focal range -0.86 μm to 0.86 μm for an infinite structure and four finite structures. The 3-D line lengths l_y vary from 1 μm to 4 μm.

Note that the distinctions for the finite line lengths are more apparent as the focus height increases. This is due primarily to an increase in the scattering interactions between the two ends of these finite lines that obscure the scattered intensity profile from the centers of the lines with increased defocus. One must consider l_y in determining the focal range $z(l_y)_{max}$ over which a finite length target can be used. Based on the simulation data from Fig. 2, for length $l_y = 2$ μm, the similarity between the finite and the infinite model is sufficient for simulations of the rectangular structure for $|z(l_y)_{max}| < 0.5$ μm, using a threshold for the RMSD of around 0.04. As a consequence we restrict the focal range in the simulations to 11 focus heights between -0.5 μm and +0.5 μm, leading to a total of 1320, i.e. 11·120, data points per polarization state.

Experimental data collected using NIST's 193 nm Microscope support that l_y can be as short as 2 μm at this wavelength. Figure 3 shows experimental images for a 2 μm and a 6 μm long

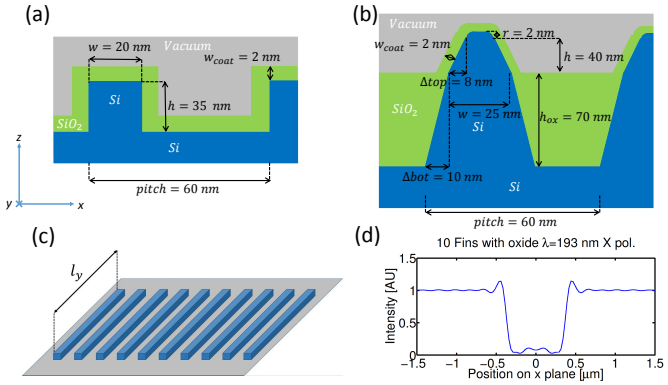

Fig. 1. Cross-sectional views of single lines in the geometry models used in the simulation studies. The coordinate system is defined at the far left. (a) the "rectangular" model with floated parameters w, h. (b) the "fins" model with floated parameters w, h. (c) 3-D schematic of the "rectangular" structure with 10 lines without an oxide layer. (d) An intensity profile for a single focus height and polarization perpendicular to the lines.

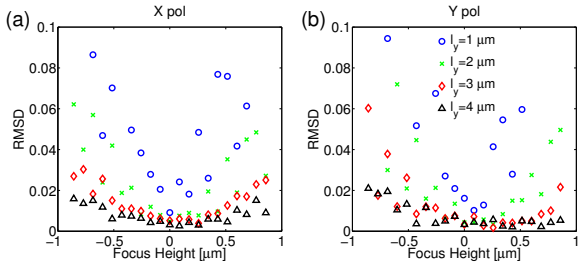

Fig. 2. Root-mean square deviation between the profiles of infinite line and finite line for different line lengths l_y and focus heights. (a) light polarized perpendicular to the lines, (b) light polarized along the lines.

structure and the RMSD of the 2 μm profile going through focus as compared to that of the 6 μm array. Here the physical microscope has a cadioptric objective lens ($0.11 < \mathrm{NA} < 0.74$) and slit illumination ($\mathrm{INA}_x < 0.15$, $\mathrm{INA}_y < 0.7$) measuring a *Si*-on-*Si* sample (cf. Fig. 1 (a)) with nominally 80 nm pitch, thus comparison to Fig. 2 is qualitative as the precise role of the INA upon defocused images requires further study. As RMSD decreases with increasing l_y through 4 μm in Fig. 2, this 6 μm line is expected to yield an intensity profile resembling that of an infinite line, while the 2 μm line may show a focal dependence. The distinction however between 2 μm and 6 μm is not as pronounced in this experiment, which may indicate that even smaller l_y or a larger $z\left(l_y\right)_{max}$ than those determined from the simulations may be possible.

Quantitative analyses of parametric uncertainties from simulation data are now used to study the capability of scatterfield microscopy to determine the geometrical parameters of the line/fin structures described above and more importantly, address the question of what an optimized combination of target and experimental conditions would be that would require the least area on the wafer while being sized large enough for an accurate determination of the CDs using scatterfield microscopy. Since in the regime where $\lambda \gg$ CD metrology is only possible using a model-based approach, we will review some basic facts

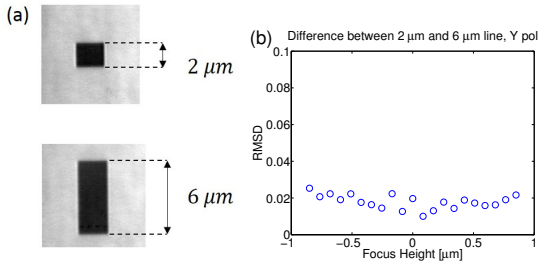

Fig. 3. (a) xy-plane images of two 30 line structures with different line lengths taken acquired using the NIST 193 nm Microscope, close to best focus with light polarized along the lines and spot size 30 μm in diameter, (b) root-mean square deviation between the intensity profiles going through focus.

about regression, before investigating the effects of changes in the measurement setup and in the measured targets. Note that the approach to these simulation studies can be generalized and is by no means tied to a particular instrument.

Following the nomenclature as provided in Ref. [11], the number of geometric parameters and the optical properties of the materials involved can be expressed mathematically as a vector of parameters $\mathbf{a} = \{a_1, \ldots, a_k\}$. In model-based metrology the parameters of interest, e.g., the height (h) and width (w) of a line structure, are determined by nonlinear regression. Given a vector of measurement data $\{y_1, \ldots, y_N\}$ and a physical model that yields simulation data $\{y(x_i, \mathbf{a})\}, i = 1, \ldots, N$ that depend on the parameters of interest $\mathbf{a} = \{w, h\}$, and the conditions x_i under which the data has been taken, we have a nonlinear regression for y_i and $y(x_i, \mathbf{a})$ given by $y_i = y(x_i, \mathbf{a}) + \epsilon_i$, with ϵ_i being the corresponding error on the i-th data point. The error model will be reduced to a simplified, uncorrelated random error that is scaled to the incident intensity, I_0, as there is no systematic error in the ideal microscope. Therefore, the 3σ uncertainties (coverage factor k=3) shown may be well below what is experimentally achievable. We assume the random vector $\boldsymbol{\epsilon} = \{\epsilon_1, \ldots, \epsilon_N\}$ to be Gaussian with zero mean and diagonal covariance matrix $\mathbf{V}$. Once the best fit value $\hat{\mathbf{a}}$ is found, its uncertainty can be estimated using the covariance matrix

$$Cov\left[\hat{\mathbf{a}}\right] = \left(\mathbf{D}\left(0\right)^T \mathbf{V}^{-1} \mathbf{D}\left(0\right)\right)^{-1}, \text{ with } \mathbf{D}\left(0\right) = \left(\frac{\partial y\left(x_i, \hat{\mathbf{a}}\right)}{\partial a_j}\right), \tag{1}$$

denoting the Jacobian matrix of the model function at the best fit value. The best fit value is usually found using gradient based optimization algorithms. Depending on the initial guess and the nonlinearity of the model function this might be a cumbersome process, for each step requires the rigorous simulation of the scattering process. We therefore generate a grid on which we interpolate the model function, decreasing the computation time to 0.05 s for a single evaluation of the interpolation. In order to prevent inverse crimes [12] we both generate the input data on a finer grid than the one we use for the regression and also add a $0.03 \cdot I_0$ uncorrelated random background noise to it, hence $\mathbf{V} = (0.03)^2 \cdot \mathbf{I}_N$, with $\mathbf{I}_N$ denoting the N-dimensional identity matrix. Note that for the sake of simplicity and reduction of computation time this error model is slightly less complicated compared to the full $\mathbf{V}$ matrix approach used in Ref. [5] that has been used to account for correlated errors.

We can now investigate how the parametric uncertainties change with respect to the structure itself. More precisely the total number of lines or fins is varied in this next simulation study. From an industrial point of view this directly addresses how large the target needs to be for metrological purposes. Several libraries were generated for increasing numbers of lines and fins, from 2 to 20 in steps of 2. The input data corresponds to a nominal line width of 20 nm and a height of 35 nm for the rectangular structure and a nominal line width of 25 nm and a height of 40 nm for the fins, with random background noise added in both cases. Once the best fit values were found their uncertainties were estimated using Eq. (1). From the results, found in Fig. 4, one can see the target size can be reduced to 12 lines or 12 fins without losing too much accuracy. Therefore from these simulations, the patterned area of an optimized, 193 nm target could be as small as 0.7 μm x 2 μm, however an unpatterned buffer area around the target would increase the actual size closer to 5 μm x 5 μm to minimize optical interactions with its surroundings. A more conservative estimate of 6 μm x 6 μm would still result in a four-fold reduction in area compared to the 12 μm x 12 μm CD scatterometry target from Ref. [3]. However there is a possible trade off in time as real space images are collected at multiple focus heights, compared to a single image of the pupil plane for that CD scatterometry target.

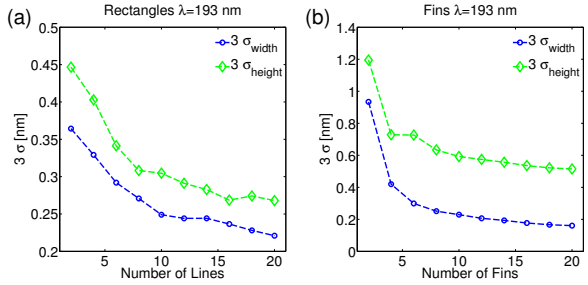

Fig. 4. Dependence of the estimated uncertainties on the number of lines (a) and fins (b) of the investigated structure.

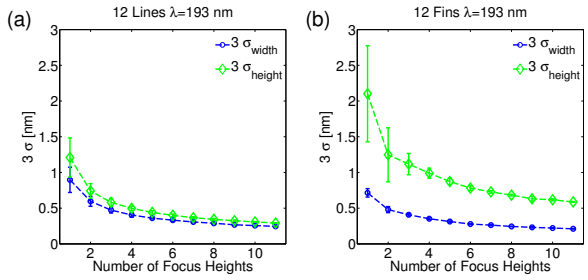

Fig. 5. Dependence of the estimated uncertainties on the number of focus heights for rectangular profiles (a) and fins (b) using both polarizations, error bars based on 10 different permutations for the drawing order of the focus heights.

Thus, we must investigate the dependence of the parametric uncertainties on the number of focus heights. The resulting parametric uncertainties as a function of the number of focus heights are presented in Fig. 5. Compared to using only a single focus height, uncertainties can be dramatically improved by adding as few as three or four focus heights. Of course the estimated uncertainties do not only depend on the number, but also on the actual focus heights selected, especially in the early stages of this process, when only one or two positions are considered. Since there are $\binom{n}{k}$ possibilities to choose k focus heights from a total

of n available focus heights it is nearly impossible to determine the optimal order for all possible permutations. We therefore report the mean and the standard deviation of the estimated uncertainties based on 10 different permutations of the drawing order. The general trends with respect to the number of focus heights appear to hold independent of the order in which they are drawn.

Underlying reasons that lead to these estimated uncertainties can be seen by examining the corresponding χ^2 surfaces, i.e., the difference between y_i and $y(x_i, \mathbf{a})$ in dependence on $\mathbf{a}$, of the different setups. The differences in the shapes in Fig. 6 illustrate the complex interplay of the parametric uncertainty, the sensitivity (slope) and parametric correlation (skewness), and as can be seen in the differences between fins and rectangles, defies generalization for arrays of arbitrary features, however a decrease in skewness, i.e. a smaller parametric correlation, with increasing number of lines/fins can be observed here.

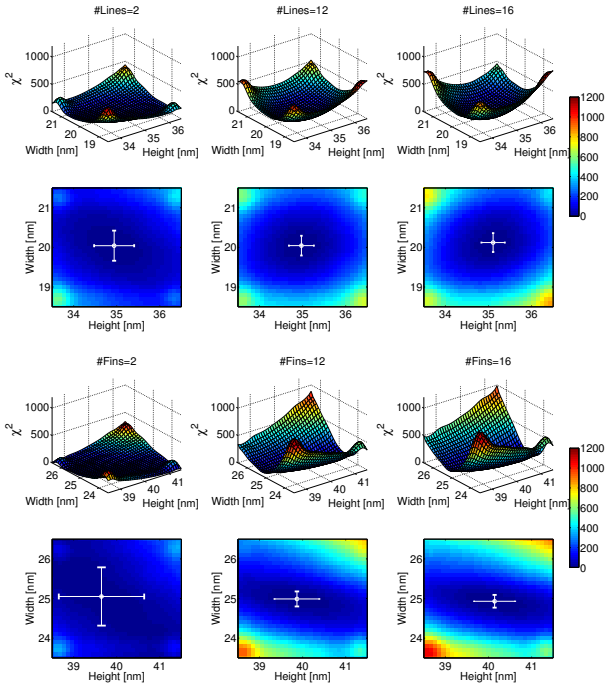

Fig. 6. Resulting χ^2-surfaces corrected for the minimum χ^2 value for three different numbers of lines and fins, note the increasing steepness due to a higher sensitivity, and decreasing skewness, due to decreasing parametric correlation.

The presented methodology can be applied to smaller critical dimensions, as are expected from the semiconductor industry, though is uncertainty in the design of future semiconductor generations, e.g. sub-7 nm nodes may include concepts such as gate-all-around structures [13]. As an example we present how the accuracy scales with a smaller CD using the fin geometry. Figure 7 shows the χ^2-surface for a scaled-down 12 fin structure with $w = 7$ nm, $h = 11$ nm, $\Delta top = 2$ nm, $\Delta bot = 2.5$ nm, $r = 0.5$ nm and a pitch of 16 nm. Without modifications to the simulated microscope setup, the parametric accuracy is almost identical even for a decrease in CD from 20 nm to 7 nm.

These simulation studies demonstrate that combinations of reduced line lengths, and reduced numbers of lines in a finite set of features can yield target sizes comparable to those in use in nanoelectronics manufacturing with a potential for up to a four-fold decrease in target area. For favorable quantification of even smaller deep-subwavelength features, these optimization methods should be repeated to enable more thoughtful choices in geometrical layout and experimental design.

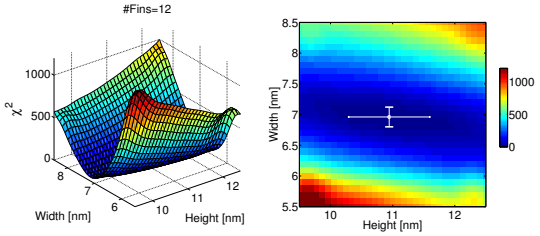

Fig. 7. Resulting χ^2-surfaces corrected for the minimum χ^2 value for the 7 nm wide fin structure.

ACKNOWLEDGEMENTS

Parts of this Letter appeared as a conference proceeding, Ref. [14]. Certain commercial materials are identified in this Letter in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials are necessarily the best available for the purpose.

REFERENCES

1. J. M. Shalf and R. LelandComputer (2015).
2. A. J. den BoefSurface Topography: Metrology and Properties **4** (2016).
3. H. Cramer, S. Petra, B. O. F. Auer, H.-J. Smilde, B. Wisse, S. Welch, S. Kruijswijk, P. Hinnen, B. Segers, C. Leewis, F. Staals, M. Escalante Marun, S. Young, W. Guo, and A. den BoefProc. SPIE **9424** (2015).
4. H.-J. H. Smilde, M. Jak, A. den Boef, M. van Schijndel, M. Bozkurt, A. Fuchs, M. van der Schaar, S. Meyer, S. Morgan, K. Bhattacharyya, G.-T. Huang, C.-M. Ke, and K.-H. ChenProc. SPIE **8788** (2013).
5. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, R. G. Dixson, and M.-A. HennLight Sci Appl **5**, e16038 (2016).
6. M. Moharam, T. Gaylord, E. B. Grann, and D. A. PommetJOSA A **12** (1995).
7. R. M. Silver, B. M. Barnes, R. Attota, J. Jun, M. Stocker, E. Marx, and H. J. PatrickApplied optics **46** (2007).
8. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, and F. GoasmatApplied optics **52** (2013).
9. S. Burger, L. Zschiedrich, J. Pomplun, and F. SchmidtIntegrated Photonics and Nanophotonics Research and Applications p. ITuE4 (2008).
10. M.-A. Henn and B. Barnes, "A library to enable the modeling of optical imaging of finite multi-line arrays," .
11. N. F. Zhang, R. M. Silver, H. Zhou, and B. M. BarnesAppl. Opt. **51** (2012).
12. J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems* (Springer Science & Business Media, 2006).
13. C. P. Auth and J. D. PlummerElectron Device Letters, IEEE **18** (1997).
14. B. Barnes, M.-A. Henn, M. Sohn, H. Zhou, and R. SilverProc. SPIE **9778** (2016).

REFERENCES

1. J. M. Shalf and R. Leland, “Computing beyond moore’s law,” Computer (2015).
2. A. J. den Boef, “Optical wafer metrology sensors for process-robust cd and overlay control in semiconductor device manufacturing,” Surface Topography: Metrology and Properties 4 (2016).
3. H. Cramer, S. Petra, B. O. F. Auer, H.-J. Smilde, B. Wisse, S. Welch, S. Kruijswijk, P. Hinnen, B. Segers, C. Leewis, F. Staals, M. Escalante Marun, S. Young, W. Guo, and A. den Boef, “Intra-field patterning control using high-speed and small-target optical metrology of cd and focus,” Proc. SPIE 9424 (2015).
4. H.-J. H. Smilde, M. Jak, A. den Boef, M. van Schijndel, M. Bozkurt, A. Fuchs, M. van der Schaar, S. Meyer, S. Morgan, K. Bhattacharyya, G.-T. Huang, C.-M. Ke, and K.H. Chen, “Sub-nanometer in-die overlay metrology: measurement and simulation at the edge of finiteness,” Proc. SPIE 8788 (2013).
5. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, R. G. Dixson, and M.-A. Henn, “Deep-subwavelength nanometric image reconstruction using fourier domain optical normalization,” Light Sci Appl 5, e16038 (2016).
6. M. Moharam, T. Gaylord, E. B. Grann, and D. A. Pommet, “Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings,” JOSA A 12 (1995).
7. R. M. Silver, B. M. Barnes, R. Attota, J. Jun, M. Stocker, E. Marx, and H. J. Patrick, “Scatterfield microscopy for extending the limits of image-based optical metrology,” Applied optics 46 (2007).
8. J. Qin, R. M. Silver, B. M. Barnes, H. Zhou, and F. Goasmat, “Fourier domain optical tool normalization for quantitative parametric image reconstruction,” Applied optics 52 (2013).
9. S. Burger, L. Zschiedrich, J. Pomplun, and F. Schmidt, “JCMsuite: An Adaptive FEM Solver or Precise Simulations in Nano-Optics ,” Integrated Photonics and Nanophotonics Research and Applications p. ITuE4 (2008)
10. M.-A. Henn and B. M. Barnes, "A Library to Enable the Modeling of Optical Imaging of Finite Multi-Line Arrays", http://doi.org/10.18434/T42C7D
11. N. F. Zhang, R. M. Silver, H. Zhou, and B. M. Barnes, “Improving optical measurement uncertainty with combined multitool metrology using a Bayesian approach,” Appl. Opt. 51 (2012).
12. J. Kaipio and E. Somersalo, Statistical and Computational Inverse Problems (Springer Science & Business Media, 2006).
13. C. P. Auth and J. D. Plummer, "Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's", Electron Device Letters 18 (1997),
14. B. Barnes, M.-A. Henn, M. Sohn, H. Zhou, and R. Silver, “Enabling quantitative optical imaging for in-die-capable critical dimension targets,” Proc. SPIE 9778 (2016).