Scanning effects in coherent fourier scatterometry

S. Roy
s.roy@tudelft.nl

Delft University of Technology, Optics Research Group, Imaging Science and Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

O. El Gawhary

Delft University of Technology, Optics Research Group, Imaging Science and Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

VSL Dutch Metrology Institute, 2600 AR Delft, The Netherlands

N. Kumar

Delft University of Technology, Optics Research Group, Imaging Science and Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

S. F. Pereira

Delft University of Technology, Optics Research Group, Imaging Science and Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

H. P. Urbach

Delft University of Technology, Optics Research Group, Imaging Science and Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands

Incoherent Fourier Scatterometry (IFS) is a successful tool for high accuracy nano-metrology. As this method uses only far field measurements, it is very convenient from the point of view of industrial applications. A recent development is Coherent Fourier Scatterometry (CFS) in which incoherent illumination is replaced by a coherent one. Through sensitivity analyses using rigorous electromagnetic simulations, we show that the use of coherence and multiple scanning makes Coherent Fourier Scatterometry (CFS) more sensitive than Incoherent Fourier Scatterometry (IFS). We also report that in Coherent Fourier Scatterometry it is possible to determine the position of the sample with respect to the optical axis of the system to a precision dependent only on the experimental noise.
[DOI: http://dx.doi.org/10.2971/jeos.2012.12031]

Keywords: scatterometry, grating, lithography inspection, coherence, nanopositioning

1 INTRODUCTION

Now-a-days, most semiconductor chips are designed to have small feature sizes, typically tens of nanometers. With increasing demand for more packing density, many innovative techniques, such as the use of extreme ultraviolet wavelength, immersion based optical system or their combinations, are being investigated to reach even lower size structures. However, fast and accurate quality control in volume production photolithography has always been an issue to be taken care of. Optical scatterometry has been one of the successful solutions to this problem, providing a non-invasive in-situ measurement with accuracy limited theoretically only by system noise. There are several other reasons for scatterometry to be one of the favoured methods in this problem. More about this can be found in several publications, for example, in [1]–[3].

In this technique measured far field intensity pattern generated by the interaction of an incident field and a scatterer is compared with simulated ones through rigorous analysis of Maxwell's equations. Generally, the final objective of this comparison is to retrieve certain properties of the scatterer by numerical optimization. Thus, scatterometry belongs to the class of inverse problems of electromagnetism [4]. However, in the present work we narrow down the objective to retrieve the shape of a one dimensional grating scatterer using focussed coherent optical illumination, termed as Coher-

ent Fourier Scatterometry (CFS), introduced in [5]. It is to be noted that the pitch of the grating is assumed to be known a priori and can be utilized to maximize sensitivity.

The working version of optical scatterometry commonly known as Incoherent Fourier Scatterometry (IFS) consists in illuminating the sample by spatially incoherent plane wavefront onto the sample [6], mainly because unwanted problems due to coherence, such as speckle etc., can be avoided while sufficiently accurate measurements are still possible. However, there are certain advantages if coherent illumination is used, mainly, a gain in sensitivity. An added feature in CFS is the ability of scanning the sample owing to the fact that coherent scatterometry is sensitive to shift of the object through a phase factor proportional to the shift [5]. Scanning plays the main role in the aforementioned gain in sensitivity. Secondly, using coherent illumination one can retrieve the phase information together with the intensity, without using any extra reference beam. Finally, given that coherent scatterometry is sensitive to lateral shift of the sample, the position of the object can also be retrieved along with the required shape parameters. This may also find some usefulness in various areas not directly related to metrology, for example, in nanopositioning without using a

reference beam [7]. Thus, in future, these added advantages may outplay the disadvantages of coherence.

The paper is organized as follows. In the next section we show a model for the typical grating we want to retrieve and we define our parameter vector, with and without considering the position of the grating as unknown. In the third section we introduce the relevant mathematical relations we need to perform a sensitivity analysis, followed by the fourth where we present the results. Here, at first with explicit examples we analyze the role of scanning and show how scanning helps to achieve better sensitivity in CFS. Then we discuss more about finding the position of the sample together with its shape.

2 EXPERIMENTAL CONDITIONS AND DEFINITION OF THE GRATING VECTOR

Let us consider a simple experiment in which a one-dimensional silicon grating having a profile of a trapezium is illuminated by the focussed field of a diffraction limited objective with spatially coherent incoming wavefront. When operating in reflection the same objective can be used to collect the reflected wavefront forming a so called epi-illumination arrangement (Figure 1(a)). Two polarizers, not shown in the figure, can be placed in paths of incoming and outgoing field. We will be considering ideal cases, i.e., a perfectly plane incident wavefront, absorption-free and diffraction-limited objective lens performance and ideal polarizers. Also, by coherent beam we mean complete spatial coherence and for incoherent beam complete spatial incoherence. The substrate and grating materials can have complex refractive indices, whereas, generally the medium surrounding the grating has real refractive index. In the plane above the lens, when the linearly polarized electric field is perpendicular to the groove (x in Figure 1(a)), we will refer to it as X polarization and for the electric field along the groove it will be Y polarization.

As mentioned before, the main difference between CFS and IFS is the phase shift occurring in CFS as the grating is moved by small amount $\Delta \mathbf{r} = (\Delta x, \Delta y, \Delta z)$. A proof of this can be found in [5] which can be rewritten in the following formula for this phase shift for a two dimensional planar grating with grating vector $\mathbf{\Lambda} = (\Lambda_x, \Lambda_y, 0)$

$$\mathbf{R}'_{lm} = \mathbf{R}_{lm} \exp \left[i2\pi \left(\frac{l\Delta x}{\Lambda_x} + \frac{m\Delta y}{\Lambda_y} \right) \right] \qquad (1)$$

where $\mathbf{R}'_{lm}$ is the diffraction amplitude for the l, m^{th} order after shifting $\Delta \mathbf{r}$ from its initial position $\mathbf{R}_{lm}$. There is no change when $l = m = 0$, i.e, for zero order. This means that coherent scatterometry works differently than incoherent scatterometry only when at least one nonzero order is present. Now we can define scanning in CFS as capture of several frames by applying small lateral shifts of the grating in the direction of its grating vector till the grating is displaced by the distance of one pitch. For the present context of one dimensional grating as in Figure 1(a), $\Lambda_y \rightarrow \infty$ and the scans are done along X. A typical CFS measurement will consist of several frames placed side by side, which, we call a superframe. This aspect
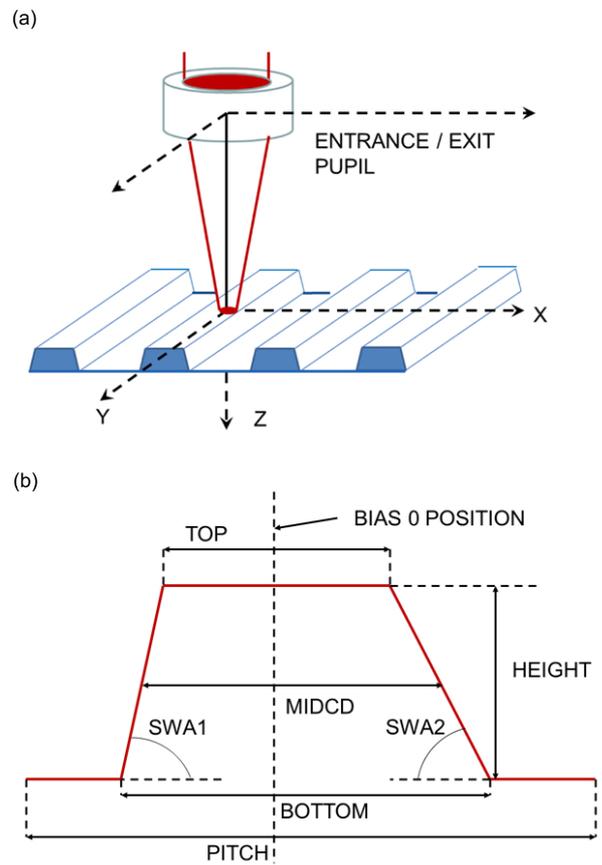

(a)

(b)

FIG. 1 The axis schematics (Figure 1(a)) and details of the grating profile at zero bias (Figure 1(b))

of coherent scatterometry encourages us to determine the position of the grating with respect to the optical axis. This is made through an additional parameter called *bias*.

In Figure 1(b), the parameters defining the shape of the grating are shown. *height* is the maximum height in the work cycle, *SWA* are side walled angles (left is *SWA*1 and right is *SWA*2) and *MIDCD* (MIDdle Critical Dimension) is the width of the trapezium at half the height. Other choices are also possible but nonetheless these make a sufficient set. The possible experimental lateral misalignment is taken care by *bias*, as mentioned before. The zero bias position is defined as the situation when the center of one period (as shown in Figure 1(b)), coincides with the optical axis of the objective. Any nonzero bias implies some lateral misalignment. The scans are symmetrically distributed around the *bias* axis spanning the length of one pitch. The separation between them, which depends on the number of scanning positions we choose and the actual value of the pitch, is called shift, corresponds to Δx in Eq. (1). If the number of scanning positions on one side of the *bias* axis including the one on the axis is given by S, then, with the configuration explained above we will have $2S - 1$ scan positions inside the pitch for a given S where the first and the final one have identical far field for being exactly one period away. This implies that we will have effectively $2S - 2$ scan positions. To be noted, in the same scheme, no scan implies $S = 1$.

Taking all this into account, the vector defining

the geometry of the grating was chosen to be $\mathbf{a} = [height, swa1, swa2, midcd, bias]$. For convenience, we will also assume symmetric grating with $swa1 = swa2$. This simplified assumption does not influence the general outcomes and leads to the reduced grating vector $\mathbf{a} = [height, swa, midcd, bias]$. In an even simplistic case the position of the grating is unimportant or known as a priori information, we can further reduce the vector to $[height, swa, midcd]$, which should be the easiest one to investigate.

3 MATHEMATICAL RALATIONS FOR SENSITIVITY ANALYSIS

In order to establish a scheme for the comparison of CFS from IFS, we have chosen to analyse the difference of sensitivities of each method and make a comparison between them. Given a merit function, the uncertainty matrix can be related to the Hessian of the function, and then elements of inverse of this matrix will correspond to the sensitivities of different parameters. This approach is used by many authors to analyse the precision of critical dimension metrology, for example, in [8] or in [9], one can find a brief or a detailed discussion respectively.

Let a function $f_p \equiv f(x_p, \mathbf{a})$ represent our model which transforms the input field into output intensity by simulating the reflection from the grating of the spot focussed by the objective on it. Referring to Figure 1(a), since waves coming from different incident angles may be diffracted into the same outgoing angle, this function maps many input waves to one output intensity value. $f(x_p, \mathbf{a})$ depends on pixel co-ordinate x_p and parameter vector $\mathbf{a} = (a_1, a_2, ..., a_N)$[1] when the incident wave is planar. If a least square merit function is defined by

$$\chi^2 = \sum_{p=1}^{(2S-1) \times M} \left[\frac{f_p^m - f_p}{\sigma_p} \right]^2, \tag{2}$$

where f_p^m is the measured intensity, M is the total number of pixels in a single fame, σ_p is the standard deviation of noise at pixel p assuming a normal distribution and, as we already know, one coherent superframe contains $2S - 1$ frames . The covariance matrix is defined by $\mathbf{C} = \mathbf{A}^{-1}$, where elements of $\mathbf{A}$ are given by the Hessian matrix

$$A_{jk} = \sum_{p=1}^{(2S-1) \times M} \frac{1}{\sigma_p^2} \left[\frac{\partial f_p}{\partial a_j} \frac{\partial f_p}{\partial a_k} \right]. \tag{3}$$

$\mathbf{C}$ gives us the variances and covariances of the parameters. We can find 3-sigma uncertainties from the diagonal elements of $\mathbf{C}$. The formula for uncertainty in parameter a_j is

$$\Delta a_j = 3\sqrt{(2S-1)MC_{jj}} \tag{4}$$

The multiplication with number of pixels in a superframe is to allow the results to be independent of the number of pixels used in a specific simulation. This is needed to make a

fair comparison between CFS and IFS owing to larger number of data in CFS. Eq. (4) gives the 3-sigma uncertainty per unit pixel per unit noise standard deviation for j^{th} parameter. The off-diagonal terms of $\mathbf{C}$ shows covariances between the parameters. To make the desired comparison between CFS and IFS, we define coherent sensitivity gain **csg**

$$csg_j = \frac{\Delta a_{j,incoherent}}{\Delta a_{j,coherent}} \tag{5}$$

An important variable which determines the difference between CFS and IFS is introduced as overlap variable F defined as

$$F = \frac{\lambda}{\Lambda \, NA}, \tag{6}$$

where λ is the illuminating wavelength, NA is the numerical aperture and Λ is the pitch of the grating. This overlap variable is important in coherent scatterometry as it takes into account system (illumination wavelength and NA) and sample (pitch) factors together. We want to concentrate on mostly the geometrical effects as F is varied, so it is preferable to vary F by changing the pitch only. This is justified since pitch is assumed to be known and we are free to adjust it to obtain maximum sensitivity. We keep the wavelength of illumination constant allowing us to keep the same refractive index throughout.

To fix the range, let us consider overlap variable varying from 0.7 to 2.2, since this range of F includes most interesting features. For $F > 2$ there is no order other than zeroth which is captured by the system, so, according to Eq. (1), no effect of scanning can be seen in the far field. If this is the case, then coherent and incoherent scatterometry gives the same far field and the sensitivities of both processes should be identical. Thus, this region sets the lower limit for the pitch, at $F = 2$, that is useful for obtaining the benefits of coherent scatterometry and can be used as a check for consistency of the simulations. The values of F such as $1 < F \le 2$ is the region where only zeroth and first orders are captured. The pupil starts to get populated by two beam interferences and difference between coherent and incoherent far field starts to build up. In a specific pixel, due to two-beam interference, there is some probability that the resultant intensity is zero and so without scanning this region shows some sharp sensitivity variations. This can be made somewhat stable with scanning, as with the additional phase in first diffracted order due to the shift, the probability of nearly destructive interference at a given pixel occurring in all non-redundant scans is lessened. By the same logic there may be some cases where no scan can give better *csg* for some parameter than experiments with more scans, due to coincidental constructive interference in more pixels. However, these are very special cases and general conclusions cannot be drawn from them, and, in most cases they will disappear or reappear in some random manner for a different NA, shape of the grating or *bias*. In the region $0.67 < F \le 1$ the second order starts to be captured and most of the pupil is now the result of three-beam interference. The coherence of light starts to play a strong role here and significant change in the far field is expected, so also any effect due to scanning. Further lowering of F will lead to even larger pitch and more interferences. This is to be avoided as this calls for rapid increase of numerical complexity and accordingly the compu-

---

[1]To be noted that x_p is a two dimensional vector as $x_p \Rightarrow (\theta_p, \phi_p)$, where θ and ϕ are planar and conical incidence angles from pixel p. Here real integer $p \in [1, M]$ denotes each pixel of a M-pixel frame.

tational time, making the overall optimization required at the final step of CFS, rather slow which is unacceptable in any practical environment.

4 RESULTS AND SENSITIVITY ANALYSIS

In the remaining part of this section, we consider a specific grating as our sample. This grating is assumed to be made of silicon on a silicon substrate and the medium surrounding the the grating is air. We will be using illumination wavelength of 633 nm, at which the complex refractive index of silicon of $3.882 - 0.019i$.

Regarding sensitivity analysis simulations, the derivatives in Eq. (3) are computed by finite differences where the size of the grids has been fixed to 0.1 nm and the angular grid is 0.1 degrees. The incident polarization is linear and can be X or Y, where we assume no polarizer at the output.

We chose a typical shape of the grating defined by parameter vector $[height, swa, midcd] = [150, 90, 0.5]$ (in nm, degree and fraction of pitch) where we analyze the sensitivity variation with pitch assuming that the *bias* is known. Later we remove this restriction. The *height* is chosen in the range as normally used in scatterometric measurements and 90 degree *SWA* is most commonly used for binary gratings. As the pitch is varying, the *MIDCD* is kept scaled to half of the pitch so that as the pitch decreases the grating profile is not becoming very small, and the process can still 'see' the refractive index variation. The *bias* is kept at zero, i.e., the optical axis divides the profile symmetrically as in Figure 1(b). The specific choice of *bias* does influence the sensitivity when no scanning is done but this influence becomes almost non existent with sufficient scanning, as is shown later. The results were calculated using RCWA [10, 11] with number of positive Fourier modes retained is 15, which has been tested to give sufficient convergence in the the range of F as mentioned earlier for silicon grating with 633 nm illumination and for both TE and TM polarization.

4.1 Sensitivity Gain in Coherent Fourier Scatterometry

Figure 2 shows the results for *csg* for height when no polarizer at the output and the input is X polarized (top) or Y polarized (bottom) for a numerical aperture of 0.4. The red line shows *csg* for no scan ($S = 1$) and the blue line is for minimal scan ($S = 2$) of two positions. As expected, *csg* starts to vary for $F \leq 2$. For no scan it oscillates till F is greater than 1, and starts to show a steady increase with smaller F. It may be noted that not only a more steady behaviour is observed when only one more scan is added, but *csg* always stays greater than unity. Thus we may conclude that coherent scanning in CFS is more sensitive than IFS. However, the gain may be made larger and more stable with more scanning, which is the next step to investigate.

Figure 3 shows how the *csg* for height is improved upon addition of more scans and how this effect of improvement is dependent on the polarization of the incident wave. Clearly
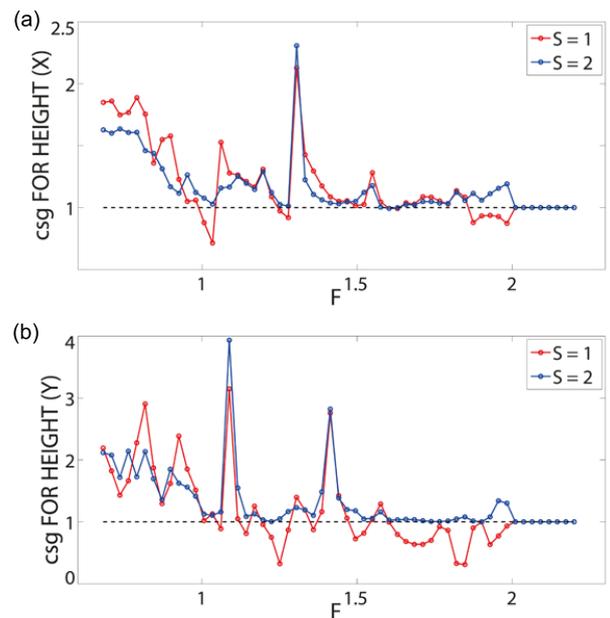

FIG. 2 *csg* for height is plotted for no scan (red) and minimal scan (blue) for two input polarization X (Figure 2(a)) and Y (Figure 2(b)) and no polarizer at the output. The *NA* is 0.4.
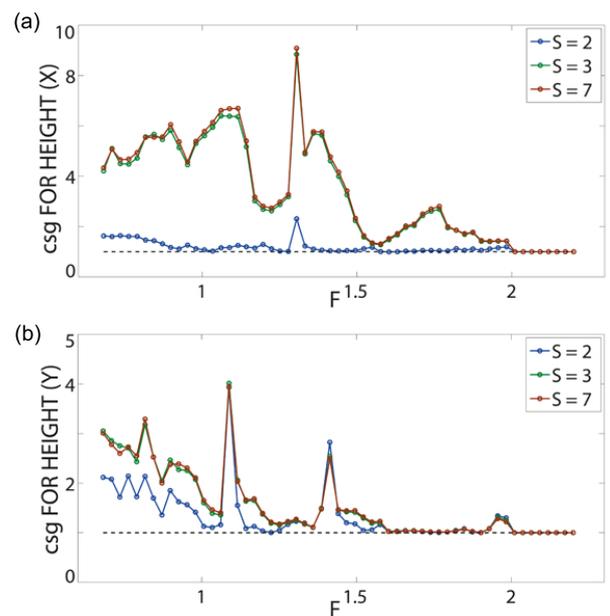

FIG. 3 *csg* for height is plotted for minimal scan (blue), $S = 3$ (green) and $S = 7$ (brown) for two input polarization X (Figure 3(a)) and Y (Figure 3(b)) and no polarizer at the output. The *NA* is 0.4.

addition of more scans improve the *csg* although naturally there is an optimum number of scanning beyond which the gain is marginal. Here a comparison between $S = 3$ and $S = 7$ reveals this optimum number to be $S = 3$, or in other words, 4 scans. However, this number should depend on overlap parameter because it should change as higher orders are being captured. Also in this example the effect of scanning seems to have a more dominant effect when the incident light is polarized along X.

To show whether these conclusions are true for higher NA, which is normally used for practical scatterometry applica-
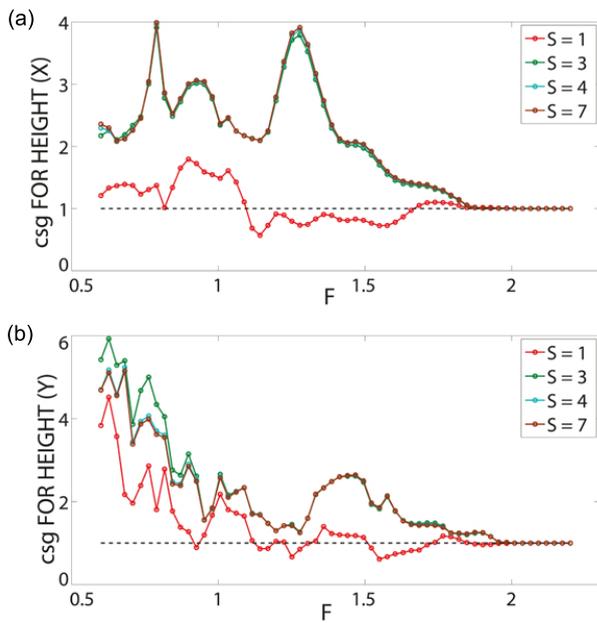
FIG. 4 *csg* for height is plotted for no scan (red), $S = 3$ (green), $S = 4$ (light blue) and $S = 7$ (brown) for two input polarization X (Figure 4(a)) and Y (Figure 4(b)) and no polarizer at the output. Note the change in the number of optimum scan positions, more apparent for input polarization Y, as F becomes less than 1 and the second order is captured. The NA is 0.9.
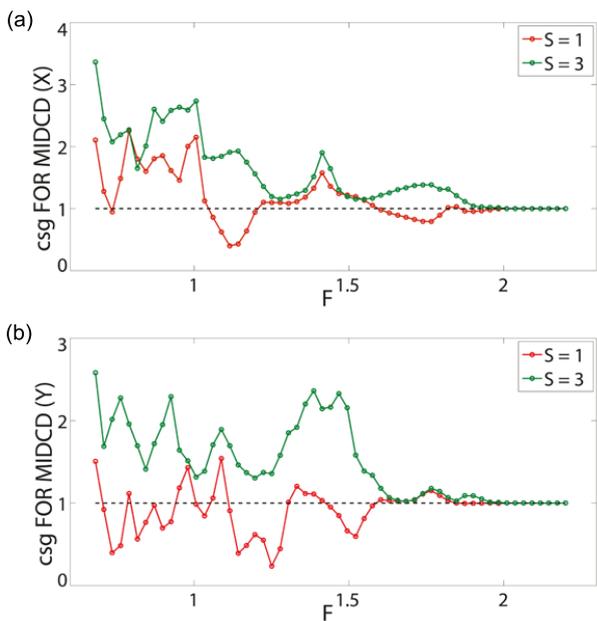

FIG. 5 *csg* for *MIDCD* is plotted for no scan (red) and $S = 3$ (green) for two input polarization X (Figure 5(a)) and Y (Figure 5(b)) and no polarizer at the output. The NA is 0.9.
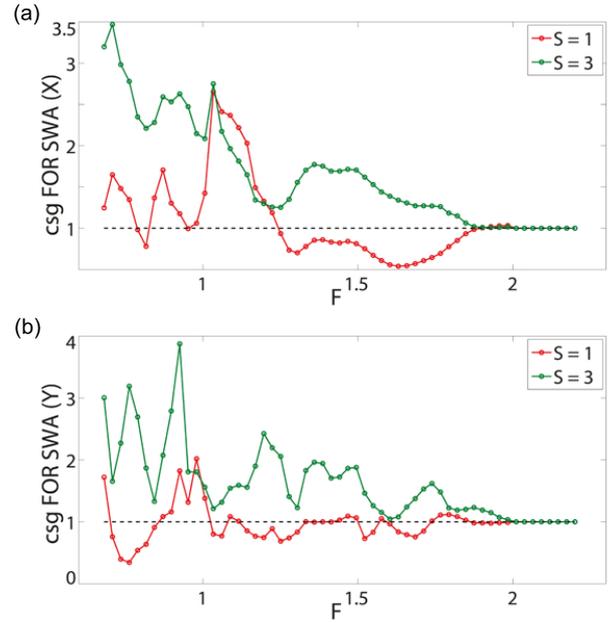

FIG. 6 *csg* for *SWA* is plotted for no scan (red) and $S = 3$ (green) for two input polarization X (Figure 6(a)) and Y (Figure 6(b)) and no polarizer at the output. The NA is 0.9.
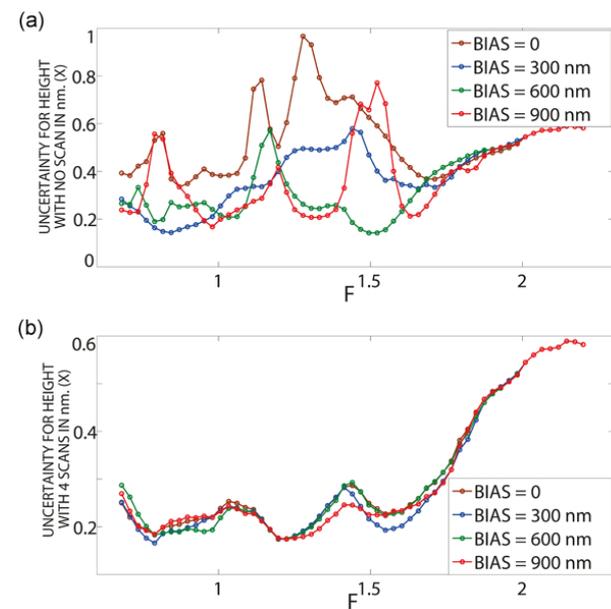

FIG. 7 3 sigma uncertainty per unit pixel with $\sigma_{noise} = 1 \times 10^{-4}$ for height is plotted for no scan (Figure 7(a)) and $S = 3$ (Figure 7(b)) with input polarization X and no polarizer at the output. In each plot sensitivities of different *bias* (0, 300 nm, 600 nm, and 900 nm) is shown in different colour. The NA is 0.9. Note the difference of the vertical axis as the scanning is introduced.

tions, the relevant plot of *csg* for height is shown in Figure 4 for $NA = 0.9$. The basic nature of the plot is similar to the previous results involving smaller NA. With the introduction of more scans, it can be noted that optimum number of scans is changed from 4 ($S = 3$) to 6 ($S = 4$), which is clearly visible for Y-polarization when F decreases below 1 and the second order comes inside the aperture. This effect is absent for X-polarization possibly due to smaller change in the far field for this case. Thus for the whole range of pitch we are interested,

we can say 6 scans are optimum, though 4 will be sufficient for most cases and can be considered optimum in practice.

The behaviour of other two shape parameters, namely *SWA* and *MIDCD*, can also be seen to be of similar nature. To avoid repetition we show only the results for no scan and $S = 3$, in Figure 5 and 6 for *MIDCD* and *SWA* respectively.

In Figure 7 we show the 3-sigma uncertainty per pixel, as defined in Eq. (4), for *height* with no scanning and 4 scans. The results are for input polarization X. For simplicity and a more
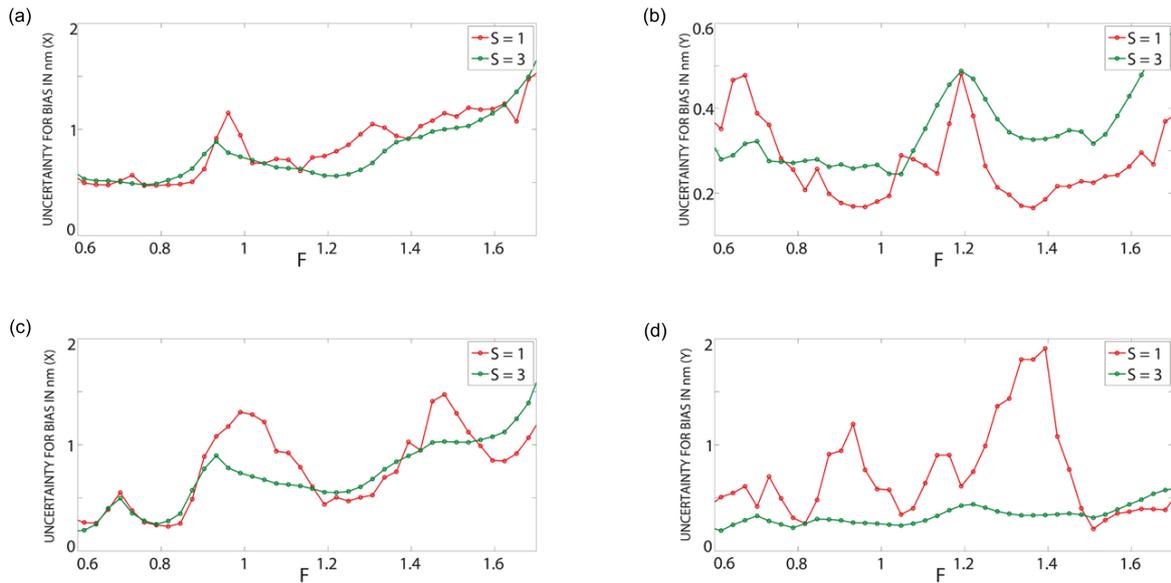
FIG. 8 3 sigma uncertainty per unit pixel with $\sigma_{noise} = 1 \times 10^{-4}$ for *bias* is plotted for no scan (red) and $S = 3$ (green) for two input polarization X (Figure 8(a) and Figure 8(b)) and Y (Figure 8(c) and Figure 8(d)) and no polarizer at the output. The top row is for *bias* = 0 and the bottom row is for *bias* = 900 nm. The NA is 0.9.

general approach, we assumed the standard deviation of the noise to be independent of pixel position having a value of $\sigma_{noise} = 1 \times 10^{-4}$ [12]. This level of noise is standard in IFS if some noise reduction image processing is done on experimental data. This gives an uncertainty in height of about 0.2 nm with scanning. From the values of the Y-axis in both plots it can be seen that the uncertainties are lowered and stabilized with scanning. This means that exact positioning of the sample is not important if sufficient scanning is done and a possible choice of F can be made without considering effects due to specific position of the grating. Conversely, it indicates sufficient scanning results in small correlation between *bias* with shape parameters and an independent determination of *bias* may be possible without retrieval of shape parameters.

To conclude this section, we may say that CFS provides better sensitivity than IFS under identical circumstances. The specific gain is dependent on the number of scans, NA and pitch and also on illumination wavelength and the grating material, however, given a specific realization, the pitch can be chosen to obtain a large ***csg***.

4.2 Dependence on *bias* and its retrieval

After establishing the gain in sensitivity in CFS compared to IFS the next step would be to extend the usefulness of CFS to take into account the additional parameter *bias* defining the position of the sample with respect to optical axis of the system. We should first check whether addition of this new parameter jeopardizes sensitivities of the shape parameters since that was our primary goal and cannot be compromised. If this is satisfied then we can check which level of uncertainty CFS gives in determination of *bias*.

There is a difference in the definition of uncertainty of the *bias*. Since zeroth order is invariant towards shift, the number of pixels whose intensity is dependent on *bias* increases as the F decreases, essentially being equal to the difference between

the total number of pixels and the pixels containing only zeroth order. If this is called M_{int} then Eq. (4) becomes

$$\Delta a_j = 3\sqrt{(2S-1)M_{int}C_{jj}} \qquad (7)$$

Unlike M, M_{int} is a function of F and goes to zero for $F \geq 2$. As $F \to 2^-$, these contributing pixels will start to be concentrated towards the edge of the aperture and are prone have lower SNR. Thus it is beneficial to keep F smaller so that sufficient samples exist with adequate SNR to allow for a practical determination of *bias*. In the simulation we avoid this case and also oversample NA to avoid any error arising due to the sharp reduction in number of data points.

To obtain the uncertainties in *bias* we may plot it for various scans. As shown in Figure 8, the uncertainty ranges between 1 to 2 nm and is often about 0.5 nm. This is for a noise SD of 1×10^{-4} as mentioned earlier. This, due to small correlation of *bias* with other parameters, can be independently optimized with a reasonable a-priori knowledge of the grating sample. We have taken two examples here at *bias* = 0 and at *bias* = 900 nm. To be noted, in *bias* = 0 case and Y polarization, no scan performs better than the multiple scanning for a large range of values of F. This is probably coincidental as it appears specifically for that *bias*. Also, since *bias* is not a shape parameter scanning may have a less dominant effect to reduce the uncertainty in this case. Nonetheless, With this level of uncertainty for positioning, CFS seems to be a good and convenient tool for nano-positioning.

5 CONCLUSION

In this paper we showed that Coherent Optical Scanning (CFS) as alternative and superior candidate in scatterometry than Incoherent Optical Scatterometry (IFS). Through sensitivity analysis we showed how scanning, which is at the heart of CFS, can lead to larger sensitivity in determination of shape parameters. This fact remains true for different cross sections

and different NA of which we showed only two examples here. There exists an optimum number of scan positions and this number can be directly related to phase shifting interferometry if we consider the scanning to be an operation similar to phase shifting of reference waves which requires at least three shifts for complete knowledge of interfering fields. By the same logic the number of scanning positions required increases to 6 as F decreases below 1 and the second order is captured, resulting in three-wave-interference in each pixel. We also showed that with sufficient scanning it is possible to reduce the correlation of bias and shape parameters and this has two advantages, namely the choice of specific value of overlap parameter F for which all the shape parameters have sufficiently high sensitivity can be done regardless of the specific position of the grating, and, independent determination of bias should be possible when shape parameters are not of interest. Both of these are desirable to make CFS a convenient and flexible tool. Finally, we showed that positioning of the grating with respect to axis of the optical system is possible to an accuracy of fraction of nanometers in CFS under standard practical conditions of optical nano-metrology. With a noise SD of $\sigma_{noise} = 1 \times 10^{-3}$, which is normally obtained without strict noise control, a positional uncertainty of 5 nm and height uncertainty of 2 nm should still be possible in CFS. Though we restricted ourselves to one-dimensional lateral misalignments because the grating in our example was one-dimensional, using Eq. (1), a two dimensional grating positioning should also be possible.

6 ACKNOWLEDGEMENTS

References

[1] P. C. Logofatu, "Sensitivity analysis of grating parameter estimation," Appl. Optics **41(34)**, 7179–7186 (2002).

[2] P. Thony, D. Harrison, D. Henry, E. Severgnini, and M. Vasconi, "Review of CD measurement and Scatterometry," in *Proceedings to the International Conference on Characterization and metrology of ULSI Technology,* 381–388 (AIP, Austin, 2003).

[3] I. Gereige, S. Robert, S. Thiria, F. Badran, G. Granet, and J. J. Rousseau, "Recognition of diffraction-grating profile using a neural network classifier in optical Scatterometry," J. Opt. Soc. Am. A **25**, 1661–1667 (2008).

[4] D. Colton, and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory* (Springer, Heidelberg, 1998).

[5] O. El Gawhary, N. Kumar, S. F. Pereira, W. M. J. Coene, and H. P. Urbach, "Performance Analysis of Coherent Optical Scatterometry," App. Phys. B: Lasers and Optics **105(4)**, 775–781 (2011).

[6] M. Wurm, and B. Bodermann, *Non-imaging optical metrology of structured surfaces: Prospects, Challenges and Limitations* (Scatterometry and ellipsometry on structured surfaces, Braunschweig, 18-19 March 2009).

[7] N. Li, W. Wu, and S. Y. Chou, "Sub-20-nm Alignment in Nanoimprint Lithography Using Moire' Fringe," Nano Letters **6(11)**, 2626–2629 (2006).

[8] R. Silver, T. Germer, R. Attota, B. M. Barnes, B. Bunday, J. Allgair, E. Marx, and J. Jun, "Fundamental Limits of Optical Critical Dimension Metrology: A Simulation Study," Proc. SPIE **6518**, 65180U (2007).

[9] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical recipes in C* (Second Edition, Cambridge University Press, Cambridge, 1992).

[10] M. G. Moharam, D. A. Pommet, and E. B. Grann, "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach," J. Opt. Soc. Am. A **12(5)**, 1077–1086 (1995).

[11] M. G. M. M. v. Kraaij, and J. M. L. Maubach, *A more efficient Rigorous Coupled-Wave Analysis algorithm*, in *Proceedings to Progress in Industrial Mathematics at ECMI,* 164–168 (Springer, Eindhoven, 2004).

[12] H. Gross, and A. Rathsfeld, "Sensitivity analysis for indirect measurement in scatterometry and the reconstruction of periodic grating structures," Wave. Random Complex **18(1)**, 129–149 (2008).