Improved ptychographic inspection of EUV reticles via inclusion of prior information

Ansuinelli, P.; Coene, W.M.J.M.; Urbach, Paul

DOI
10.1364/AO.395446

Publication date
2020

Document Version
Accepted author manuscript

Published in
Applied Optics

Improved ptychographic inspection of EUV reticles via inclusion of prior information

PAOLO ANSUINELLI,[1,*] WIM M.J. COENE,[1,2] H. PAUL URBACH[1]

[1] *Optics Research Group, Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, P.O. Box 5046, 2600GA Delft, The Netherlands*
[2] *ASML Netherlands B.V., De Run 6501, 5504 DR Veldhoven, The Netherlands*
[*] *p.ansuinelli@tudelft.nl*

Abstract: The development of actinic mask metrology tools represents one of the major challenges to be addressed on the roadmap of extreme ultra violet (EUV) lithography. Technological advancements in EUV lithography result in the possibility to print increasingly fine and highly resolved structures on a silicon wafer, however the presence of fine–scale defects, interspersed in the printable mask layout, may lead to defective wafer prints. Hence the development of actinic methods for review of potential defect sites becomes paramount. Here, we report on a ptychographic algorithm that makes use of prior information about the object to be retrieved, generated by means of rigorous computations, to improve the detectability of defects whose dimensions are of the order of the wavelength. The comprehensive study demonstrates that the inclusion of prior information as a regularizer in the ptychographic optimization problem results in a higher reconstruction quality and an improved robustness to noise with respect to the standard ptychographic iterative engine (PIE). We show that the proposed method decreases the number of scan positions necessary to retrieve an high quality image and relaxes requirements in terms of signal to noise ratio (SNR). The results are further compared with the state–of–art total variation based ptychographic imaging.

1. Introduction

Since EUV lithography is nowadays adopted for high volume manufacturing (HVM) in the semiconductor industry, it becomes relevant to address various technical and technological challenges on its roadmap [1]. One of the crucial concerns is mask defectivity [2]. While the technology matures and the technical advancements pave the way towards the printability of increasingly fine features, defects of smaller sizes may also become printable. This dictates the need for highly sensitive mask inspection tools that can generate highly resolved defect maps for defect review and inspection. The EUV mask can be inspected at a non actinic wavelength [3], however the images obtained by means of deep ultra violet (DUV) and e-beam tools can differ significantly from the aerial images of the scanner [4,5]. This, in turn, can result in the incapability to detect crucial defects that can lead to device failure. Hence, the actinic inspection of EUV masks is particularly important. Further, as mask manufacturers are considering phase shifting EUV mask absorbers for the next generation of EUV technology, methods for metrology and quantification of these phase shifts are important.

Ptychography [6] – a lensless imaging method that enables wide field–of–view, high resolution imaging via phase retrieval – is a possible candidate for the inspection of samples, in reflective mode at short wavelengths [7–9]. This coherent diffractive imaging (CDI) method has been introduced as a potential actinic inspection tool for EUV mask inspection [10–12]. In ptychography a probe sequentially illuminates a given scattering object at partially overlapping positions. The scattered light is usually detected in the far–field and the recorded diffraction patterns are computationally processed to image the transmission/reflection function of the scattering object. The translational diversity and the highly redundant information in the data set result in a

robust solution of the phase problem. Nevertheless, since the phase problem is a nonlinear and non–convex inverse problem, ill–defined solutions can still arise and there are neither theoretical guarantees on the uniqueness of the retrieved solution, nor certainties about its optimality [13]. In such cases including prior information in the optimization algorithm can be beneficial.

The use of prior information is ubiquitous in the inverse problems community [14,15] and it was proven to be important to unveil deep subwavelength details in optical imaging methods [16–18]. In the context of CDI, prior information about the amplitude of the ptychographic illumination function enabled subwavelength imaging at the edges of the lines of a periodic object [19]. Recent studies have further shown the benefits that stem from the inclusion of the total variation (TV) prior in the ptychographic algorithms [20–22].

In this work we have taken a different approach and instead of using a prior for the probe or a TV prior we have devised a way to generate a physically sound prior for the transmission/reflection function of the scattering object.

In what follows we show how one can make use of rigorous forward modeling with Maxwell solvers to compute such prior and how this can be included in the ptychographic method. We apply the algorithm to the problem of detection and imaging of extrusion and intrusion type defects in a patterned EUV mask layout. The method is shown to outperform the regular ptychographic iterative engine (PIE) and the total variation based ptychographic method.

2. Method

Ptychography can be framed as a cost functional minimization problem in which, considering the j–th probe position $P(\boldsymbol{r} - \boldsymbol{R}_j)$, one seeks a certain object $O_j(\boldsymbol{r})$ that best fits the j–th recorded diffraction intensities $I_j(\boldsymbol{k})$:

$$\begin{aligned} \mathcal{L}(O_j, I_j) &\coloneqq \left\| |\mathcal{F}(P(\boldsymbol{r} - \boldsymbol{R}_j) O_j(\boldsymbol{r}))| - \sqrt{I_j(\boldsymbol{k})} \right\|^2 \\ O_j(\boldsymbol{r}) &\coloneqq \arg\min_{O_j} \mathcal{L}(O_j, I_j) \end{aligned} \tag{1}$$

where $\mathcal{F}$ represents the Fourier transform, $P(\boldsymbol{r})$ is the probe function, $O_j(r)$ is the part of the guessed object which is illuminated when the probe is located at the j–th probe position and $I_j(\boldsymbol{k})$ is the intensity pattern measured at the j–th probe position. In (1) and in all that follows we assume the probe to be known. The PIE proceeds in a sequential fashion and the complete object, $O(\boldsymbol{r})$, is reconstructed by the sequential reconstruction of its j-th views [6]. Owing to the translational diversity and to the redundancy in the data the ptychographic reconstructions are very robust. However, as the phase problem is a nonconvex optimization problem – meaning that there can be multiple local minima – and because the reconstruction is obtained solely by processing the data set, one might wonder whether the inclusion of physically sound prior knowledge in the algorithm can lead to a more satisfactory solution. Prior information is commonly included introducing a regularization term to the data discrepancy functional. In this case:

$$O_j^\alpha(r) \coloneqq \arg\min\{\mathcal{L}(O_j, I_j) + R^\alpha(O_j)\} \coloneqq \arg\min\{\mathcal{L}^\alpha(O_j, I_j)\} \tag{2}$$

where $R^\alpha(O_j)$ is the regularizer, which depends upon the regularization parameter α, $\mathcal{L}^\alpha(O_j, I_j)$ is the regularized cost functional and $O_j^\alpha(r)$ is the associated object update. $R^\alpha(O_j)$ is intended as a penalty term in variational approaches or as the negative logarithm of a prior probability distribution from the Bayesian viewpoint [23]. Different regularizers (priors) have been studied in the field of image reconstruction, for instance structural priors [24], sparsity promoting priors [25] or edge–preserving priors. An important example of the latter is total variation (TV), in which $R^\alpha(O) = ||\nabla O(\boldsymbol{r})||_1$. TV is commonly employed due to its effectiveness in smoothing noise – by favouring images that have a sparse gradient – while preserving edges. In recent years there have

been a number of studies reporting implementations of TV regularized ptychography [20–22]. In this work we have taken a different approach and instead of considering a regularizer that promotes generic properties of the reconstructed image (like its sparsity or the sparsity of its gradient), the regularizer $R^{\alpha}(O)$ has been designed to promote adhesion to a given prior image (object):

$$R^{\alpha}(O_j) = \alpha \left\| O_j(\boldsymbol{r}) - O_{p,j}(\boldsymbol{r}) \right\|^2 \tag{3}$$

where $O_{p,j}(\boldsymbol{r})$ is the prior of the object at the j–th probe position. Crafting a "good" prior is challenging. This task requires one to properly account for the physics that contributes to the process of image formation. For instance it is known that at EUV wavelengths waveguiding effects and diffraction of light along the thickness of the absorber materials have an impact on image formation [26]. Further, the phase of the scattered wave $\Phi = \boldsymbol{k} \cdot \boldsymbol{r}$ accumulates linearly as photons pass through an object and the rate of accumulation depend on the refractive index which changes from material to material [27]. This is important to realize when one wants to compute an *a–priori* map for the phase of a wave scattered by nanostructures that consist of a sequence of layers of different materials and thicknesses. Such a situation applies to the multilayer Bragg–reflector and to the absorber layers of an EUV mask. These aspects should be properly encoded in the prior, in order for it to be reliable and accurate.

These physical aspects and the intrinsically 3D thickness effects can be duly accounted for recurring to three dimensional fully rigorous simulations. Forward Maxwell solvers can compute the complex field, which is a rigorous solution in terms of amplitude and phase of Maxwell's equations, for a given 3D scattering geometry [28, 29]. It is important to notice that light–matter interaction is modeled differently in the rigorous Maxwell solvers and in ptychography. On the one hand the rigorous electromagnetic solvers provide an accurate solution of Maxwell's equation, on the other hand ptychography models light–matter interaction in terms of the 2D "probe–times–object" approximation. This fundamental difference in the physical models could be a reason of concern when intermixing the use of the rigorous solvers with ptychographic algorithms. In other words, a certain rigorously computed complex–valued far–field Ψ^{Maxw} can be used in ptychographic algorithms only when it can be interpreted in terms of the 2D ptychographic approximation of light matter interaction:

$$\Psi^{Maxw}_{p,j} \approx \mathcal{F}(P(\boldsymbol{r} - \boldsymbol{R}_j) O_{p,j}(\boldsymbol{r})) \tag{4}$$

where $\Psi^{Maxw}_{p,j}$ is the far–field, amplitude and phase, as computed by the forward Maxwell solver, for the nominal – *a priori* known – scattering geometry on the mask. Notice that although ptychography assumes the object function to be two dimensional, $\Psi^{Maxw}_{p,j}$ is computed by the rigorous 3D simulations. In other words, the 2D object reflection function comprises genuine 3D information about the object. Bearing (4) in mind we proceed to write, at the j–th probe position, the ptychographic cost functional inclusive of the prior term. By Eqs. (1)–(3):

$$\begin{aligned} \mathcal{L}^{\alpha}(O_j, I_j) &\coloneqq \mathcal{L}(O_j, I_j) + R^{\alpha}(O_j) = \\ &= \left\| |\mathcal{F}(P(\boldsymbol{r} - \boldsymbol{R}_j) O_j(\boldsymbol{r}))| - \sqrt{I_j(\boldsymbol{k})} \right\|^2 + \alpha \left\| O_j(\boldsymbol{r}) - O_{p,j}(\boldsymbol{r}) \right\|^2 \end{aligned} \tag{5}$$

where α regulates the interplay among the two terms on the right hand side of Eq. (5). The second term in (5) penalizes large deviations of the reconstructed object from the prior object. Eq. (5) can be minimized analytically using Wirtinger calculus to yield the update rule that can be used in a gradient descent algorithm. The gradient of Eq. (5) with respect to O^*_j is:

$$\nabla \mathcal{L}^{\alpha}(O_j, I_j) = P^*_j \left[\mathcal{F}^{-1}(\Psi_j - \Psi_{c,j}) \right] + \alpha(O_j - O_{p,j}) = \nabla \mathcal{L}_1 + \nabla \mathcal{L}^{\alpha}_2 \tag{6}$$

In (6) $\Psi_j = \mathcal{F}(P_j O_j)$ is the guessed wavefront in the momentum space, $P_j = P(\boldsymbol{r} - \boldsymbol{R}_j)$ and $\Psi_{c,j}$ is the revised wavefront obtained by enforcing the measured amplitudes on the far–field guess. We focus now on $\nabla \mathcal{L}_2^{\alpha}$ and, with the use of Eq. (4), we obtain:

$$\begin{aligned} \nabla \mathcal{L}_2^{\alpha} &= \alpha(O_j - O_{p,j}) = \alpha \frac{P_j^*}{(|P_j|^2)}(P_j O_j - P_j O_{p,j}) = \\ &= \alpha \frac{P_j^*}{|P_j|^2} \mathcal{F}^{-1}(\mathcal{F}(P_j O_j) - \mathcal{F}(P_j O_{p,j})) = \alpha \frac{P_j^*}{|P_j|^2} \mathcal{F}^{-1}(\Psi_j - \Psi_{p,j}^{Mawx}) \end{aligned} \tag{7}$$

The steepest descent update, at given iteration n dictates:

$$O_{j,n+1} = O_{j,n} - \beta \nabla \mathcal{L}^a \tag{8}$$

therefore the part of the object that is illuminated by the probe at the j–th probe position, $O_j(\boldsymbol{r})$, is updated as follows:

$$\begin{aligned} O_{j,n+1} = O_{j,n} &+ \beta \frac{|P_j|}{|P_{j,max}|} \frac{P_j^*}{(|P_j|^2 + c)}(\Psi_{c,j,n} - \Psi_{j,n}) \\ &+ \alpha' \frac{|P_j|}{|P_{j,max}|} \frac{P_j^*}{(|P_j|^2 + c)}(\Psi_{p,j}^{Maxw} - \Psi_{j,n}) \end{aligned} \tag{9}$$

In Eq. (9) c is a small constant that prevents division by zero, $\alpha' = \beta\alpha$, and $|P_j|/|P_{j,max}|$ is a scaling factor – present in the PIE – that makes the update less reliable where the probe is dim. Notice that, besides α' and β, the scaling factors in (9) are the same for the two terms. This is in order to avoid overenforcing the prior in those pixels where the probe is dim. For $\alpha = 0$, Eq. (9) yields the same update rule of the PIE. The algorithm proceeds in a sequential fashion, and the complete object, $O(\mathbf{r})$, is reconstructed by the sequential reconstruction of its j–th views, according to Eq. (9).

A layout of this approach is given in Fig. 1.

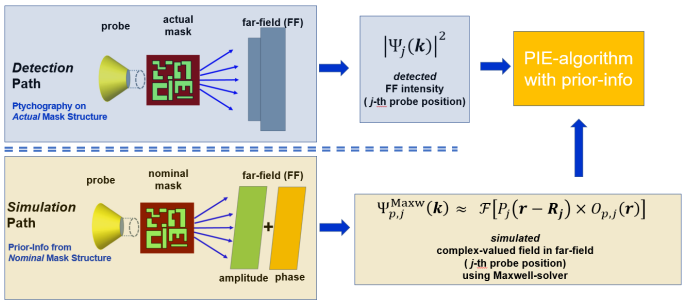

Fig. 1. A schematic that illustrates the approach of the proposed work.

The inclusion of prior information as a regularizer in Eq. (5) is preferable to the use of $O_p(\boldsymbol{r})$ as a starting guess in the optimization. There are various reasons for this:

- the presence of the quadratic term in (5) improves the conditioning of the problem. This stabilizes the inversion with respect to the noise and improves the performance of the iterative method. Further, the regularizer aids in creating a better model by achieving a balance in the bias–variance trade–off for a proper selection of α [30]. Values of α which are too small make the reconstruction too sensitive to the noise, however setting α to a value that is too high biases the reconstruction towards the prior, yielding a poor fit. A proper value of α achieves a balance among these two cases.

- The prior in Eq. (9) is present at every step of the iteration, therefore preventing the reconstruction from diverging towards an "unphysical" solution, and
- the regularization parameter α can be tuned to reflect the degree of belief in the prior.

In what follows, we will show that the inclusion of the regularization term of Eq. (3) via the use of the accurate physical–analytical models yields a better reconstruction with respect to the standard case in which the transmission function $O(\boldsymbol{r})$ is retrieved solely by processing the measured data–set.

3. Results

We have applied the method outlined above to reconstruct the patterned absorber of a 3D EUV mask. Table 1 reports the materials and the thicknesses of the layers used in this work.

Table 1. **Layers thicknesses and Materials at λ = 13.5 nm**

layer	thickness [nm]	n	k
ARC TaBO	2	0.952	0.026
Absorber TaBN	58	0.95	0.031
Ru	0.5	0.88586	0.01727
Ru (Capping layer)	2	0.88586	0.01727
Si	1.8968	0.99888	0.00183
MoSi2	0.7986	0.96908	0.00435
Mo	2.496	0.92347	0.00649
MoSi2	1.8908	0.96908	0.00435

Four EUV masks have been considered through this study:

- the "nominal" mask. This is the cell as given by prior information. This cell does not contain any information about the defects.
- The "actual" mask. This is the cell that mimics the "reference" mask which is close to the prior but not exactly the same. In order to account for this difference the actual cell has been generated from the prior cell, displacing the sides of the polygons of 1–5 nm. This cell is displayed in Fig. 2(a).
- The programmed defects mask. Consistently with the practice in EUV mask defectivity studies we have perturbed the actual mask, at known locations, with additive and subtractive features (extrusions and intrusion defects) (Figs. 2(b–c))

The size of the defects in Figs. 2(b,c) is the same on a given polygon, and it changes from polygon to polygon. The number and the side length of the squares that constitute the rough extrusions/intrusions on a certain polygon are the following: [*number of squares*, *side length*] = [3, 16 nm], [6, 12 nm], [7, 9 nm], [7, 6 nm]. Such sizes have been chosen in accordance to the theoretical Abbe resolution limit imposed by the NA (11 nm in this case). If the collection NA was to be smaller the defects would have been made bigger accordingly.

To understand whether the inclusion of prior information yields any benefits for our specific application we have carried out a computational die–to–database comparison [31]. This is done

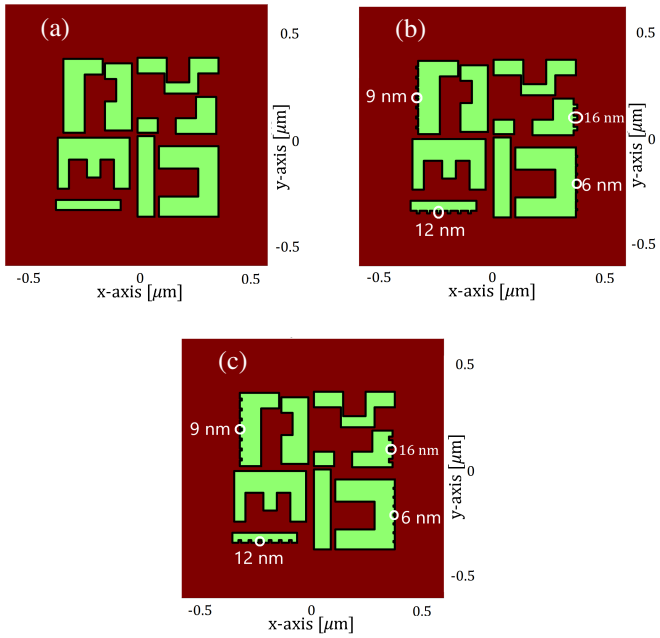

Fig. 2. Top views of EUV mask layouts. a) Actual cell , b) programmed defect mask (extrusions), c) programmed defect mask (intrusions). The sidelength of the single defect is specified in the figure.

comparing the reconstruction of the defected mask with the reconstruction of the defect–free actual mask. The two reconstructions are subtracted one from the other to identify the defects at their locations. The impact of the defects is quantified by a certain figure of merit. In what follows we will use the defect SNR defined as [31]:

$$SNR_{\delta} = \frac{\bar{A}_d - \bar{A}_a}{std(A_a)} \tag{10}$$

Where $\bar{A}_d$ is the average magnitude of the defected area, $\bar{A}_a$ is the average magnitude of the whole difference image – where the object is present – and $std(A_a)$ is the standard deviation of the latter area. The definition of the defect SNR is independent of the defect size. This investigation is done using the standard PIE and the PIE with prior, where we use Eq. (9) as update rule to reconstruct both the actual mask (Fig. 2(a)) and the defected masks (Fig. 2(b–c)).

As outlined above four data sets have been computed:

- one complex data set, in amplitude and phase, for the nominal mask. This is the far field $\Psi_{p,j}^{Maxw}$, related to the nominal mask, we referred to in Eqs. (4) and (9).
- One intensity–only data set for the actual defect–free mask in Fig. 2(a). This data–set has been corrupted with noise to emulate measured data.
- Two intensity–only data sets for the defected masks, one for the extrusions in Fig. 2(b) and one for the intrusions in Fig. 2(c). This data sets have also been corrupted with noise to emulated measured data.

The data sets mentioned above have been generated via fully rigorous 3D simulations using a volume–integral Maxwell solver [32, 33]. The solver is formulated for the problem of scattering

from periodic objects hence, in order to avoid cross–talk among adjacent cells, we have opted for a supercell approach. The cell is a square with lateral dimension $\Lambda = 3.5\mu m$. The sampling in the far field and in the illuminating NA equals $\frac{2\pi}{\Lambda}$. Although the lateral dimensions of the supercell are of the order of hundreds times the wavelength, the computational complexity and the memory requirement necessary to solve the forward problem are maintained relatively low, of the order of $O(N \log N)$. The probe is assumed to be a Gaussian beam with a 3σ amplitude of about 1.5 μm and it is described by its angular spectrum. The scattered far–field is evaluated, for each of the plane waves which compose the illumination, in parallel on a multicore HPC cluster. The output far–field that results from the interaction of the probe with the object is then given by the weighed coherent superposition of the individual contributions. The ptychographic scans are performed shifting the object of 0.2 μm, in 5 positions, inside the supercell. The probe is polarized in the x direction – parallel to the horizontal axis of the supercell – by proper linear combination of s and p polarization states. The collection NA is 0.6, close to the value (0.54) used for an identical wavelength in [19]. Fig. 3(a,b) illustrates the probe, its cross–section and one of the acquired diffraction patterns for the mask in Fig. 2(a).

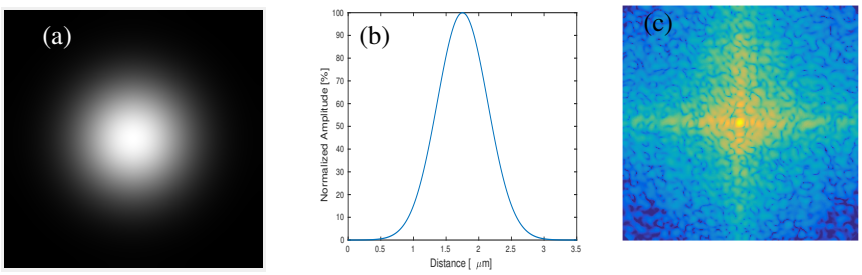

Fig. 3. a) Probe used in the computations. b) Cross–section of a. c) Diffraction pattern (log scale).

Fig. 4(a,b) shows the prior for the central probe position. In all that follows we have fixed the regularization parameter α to 2e-2.

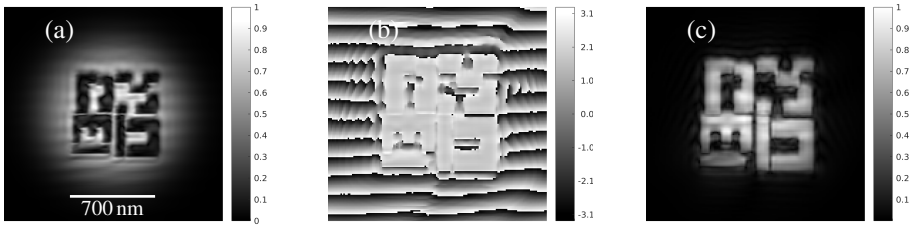

Fig. 4. Zoom of the prior object, a) magnitude, b) phase and c) magnitude after and removal of the strong unscattered beam to emulate the presence of a beam–stop in the experimental set–up. The fringes in (b) are caused by the projection of the term $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ over the x–y plane related to the 6° tilt of the probe with respect to the normal.

3.1. Extrusion defects

The magnitude of the ptychographic reconstruction of the patterned absorber depicted in Fig.2(b), without and with prior respectively, is depicted in Fig. 5(a,b), while the phase is shown in Fig.

5(d,e). White Gaussian noise has been added to the synthetically generated data for an SNR = 110 dB. The object error, at iteration n, has been computed as the deviation of the reconstructed object, $O_n(\boldsymbol{r})$, from the theoretical object $O_t(\boldsymbol{r})$:

$$E_{O,n} = \frac{\sum_{\mathbf{r}} |O_t(\mathbf{r}) - \gamma O_n(\mathbf{r})|^2}{\sum_{\mathbf{r}} |O_t(\mathbf{r})|^2} \tag{11}$$

where γ is a parameter that compensates for phase ambiguities [34].

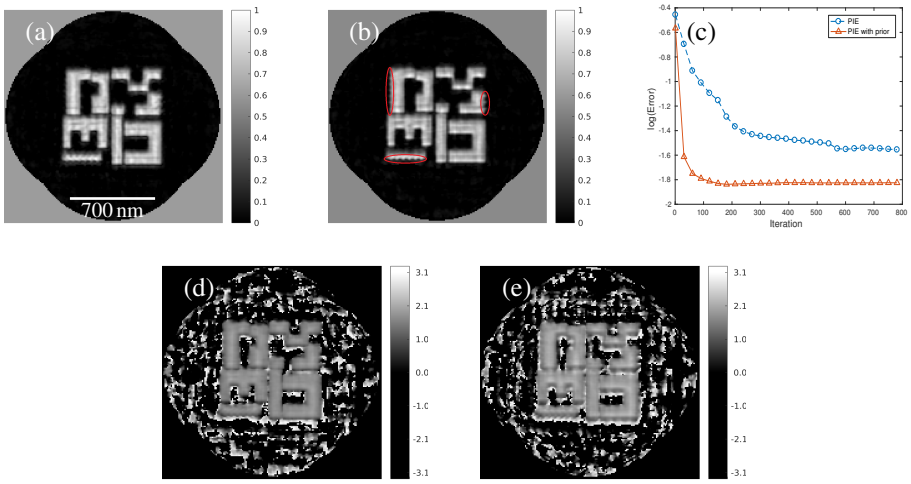

Fig. 5. Ptychographic reconstructions: a) magnitude as given by the PIE; b) magnitude given by the new algorithm PIE with prior term included. The rough extrusions are highlighted in the red circles;c) error; d) phase given by the PIE and e) phase given by the PIE with prior.

Result of the reconstructions is highlighted in Figs. 5(a–d). The rough extrusions, highlighted in Fig 5(b), are better resolved in both the amplitude and phase images in Fig. 5(b,d). The error, defined in (11) and shown in Fig 5(c) shows an overall better reconstruction and convergence when the prior is included in the optimization algorithm. Practically, we have observed the algorithm to converge in about a third of the iterations of the regular PIE. As stated before a quantitative assessment of the improvement in terms of defect inspection can be obtained in a die–to–database comparison by computing difference images. Those are defined as the magnitude of the difference among the retrieved ptychographic reconstruction of the objects in Figs. 2(a,b):

$$\Delta O(\mathbf{r}) = |O_a(\mathbf{r}) - O_\delta(\mathbf{r})| \tag{12}$$

with $O_a(\mathbf{r})$ and $O_\delta(\boldsymbol{r})$ being the reconstructed reflection functions relative to actual mask in Fig. 2(a) and to the defected mask in Fig. 2(b–c). In what follows we will refer to Eq. (12) as the object difference metric. Fig. 6 shows $\Delta O(\mathbf{r})$ defined in (12), obtained when reconstructing $O_a(\boldsymbol{r})$ and $O_\delta(\boldsymbol{r})$ using prior information – by the ptychographic update rule (9) – and the standard PIE. Here, and in all that follows, the objects have been aligned before their subtraction.

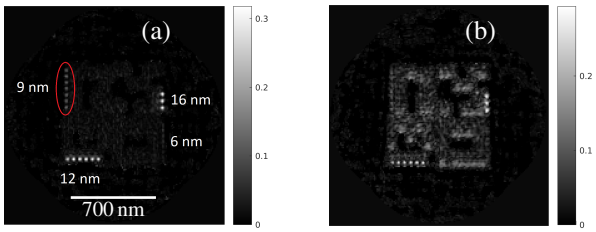

Fig. 6. a) and b) show the object difference metric ΔO in Eq. 12 obtained for extrusions. a) makes use of prior while in b) we employed the standard PIE. The sizes in a) are the side lengths of each of the blocks of the extrusions.

All the defects are better resolved in Fig. 6(a) than in Fig. 6(b), and their signature appear to be stronger in the difference image 6(a). Particularly, the rough defects of 9 nm size are not distinguishable in Fig. 6(b), however they are detectable in Fig. 6(a), as highlighted in the red circle. We found the finest details, the rough extrusions of 6 nm size, to be absent in the reconstruction in Figs. 5(a–b), and in the difference images in Figs. 6(a–b). The side of the polygon over which these defects are located appears to be smooth in the reconstructed image. This computational experiment highlights that subwavelength ptychographic imaging is possible to a certain extent, however the spatial periodicities below the theoretical limit of about $\lambda/2$ are not retained, in this case, in the reconstruction. Being the collection NA equal to 0.6 the Abbe limit is 11 nm which is slightly above the size, 9nm, of the smallest defect we managed to resolve. A comparison of the retrieved SNR for Figs. 6(a,b) is reported in Table 2:

Table 2. **Extrusion defects SNR**

Defect Size [nm]	Number of defects	SNR PIE with prior	SNR PIE
16	3	3.8	1.8
12	6	5.3	1.2
9	7	1.6	N/A

Table 2 highlights a steep improvement in the detectability of defects when incorporating the prior term in the reconstruction algorithm. The value N/A means that the defets are not visible in the difference image. A cross section of the reconstructions in Fig 5(a–b), taken on the location of the defects, is shown in Figure 7.

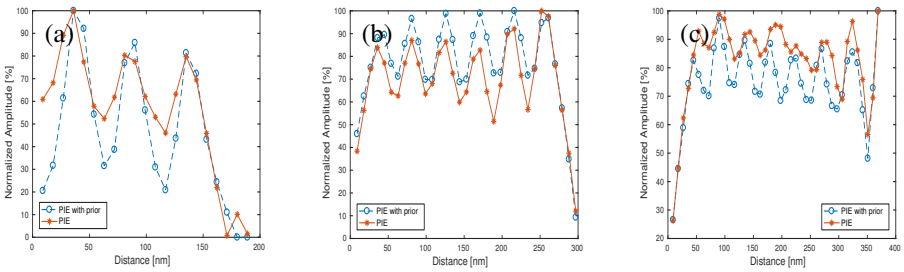

Fig. 7. Cross section of the reconstruction at the defect sites.a) 3 defects of 16 nm sidelength; b) 6 defects of 12 nm sidelength; c) 7 defects of 9 nm sidelength.

The cross sections in Fig. 7 generally have a move pronounced peak to valley ratio when the prior is included and, in the case of Fig. 7(c), the periodicity of the signal is more evident, while it is lost – in the central part of the plot – in the case of the standard PIE.

3.2. Intrusions

The magnitude of the ptychographic reconstruction of the patterned absorber layout in Fig. 2(c), without and with prior respectively, is depicted in Fig. 8(a,b), while the phase is shown in Figs. 8(d,e).

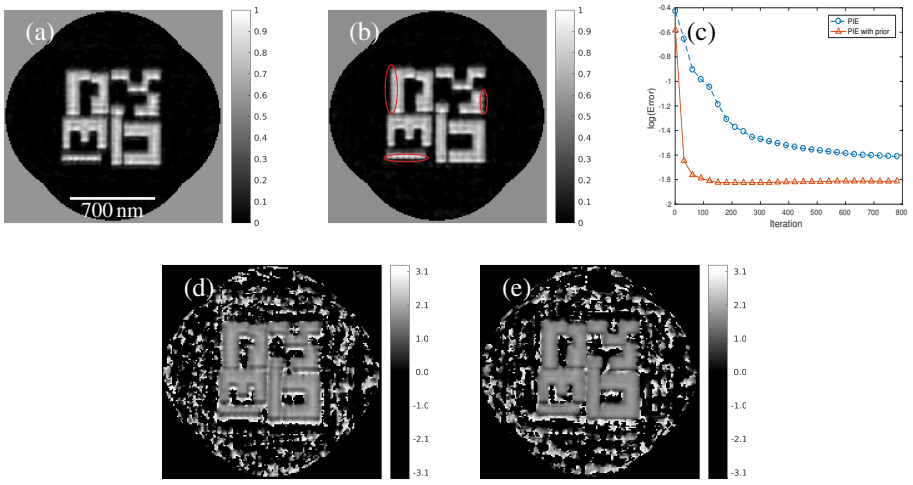

Fig. 8. Ptychographic reconstructions: a) magnitude as given by the PIE; b) magnitude given by the PIE with prior included. The rough inclusions are highlighted in the red circles; c) phase given by the PIE and d) phase given by the PIE with prior;e) error.

Fig. 9 shows the object difference metric $\Delta O(\mathbf{r})$ defined in (12), obtained using the update rule (9) and the standard PIE.

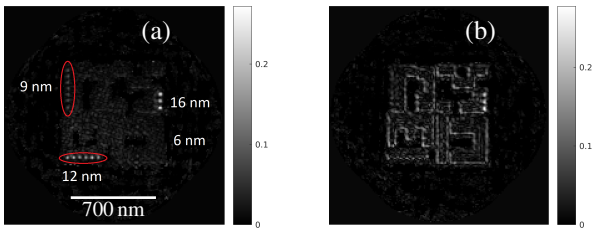

Fig. 9. a) and b) show ΔO in Eq. (12) obtained for intrusions. a) makes use of prior while in b) we employed the standard PIE. The sizes in a) are the side lengths of each of the blocks of the intrusions.

The comparison of the SNR for Figs. 8(a,b) is reported in Table 3:

All the intrusion defects have decreased SNR with respect to extrusion defects of the same size, and even with the inclusion of the prior term the 9 nm size intrusions are difficult to image. As guided modes propagate within the absorber, and as they are not immediately truncated at the edges of the structures, they can couple and interfere within the small intrusions, making

Table 3. **Intrusion defects SNR**

Defect Size [nm]	Number of defects	SNR with prior	SNR PIE
16	3	3.8	1.5
12	6	3.4	N/A
9	7	1	N/A

these defects harder to image with high contrast [26]. Fig. 10 illustrates the cross section of the reconstructions in Fig. 8(a–b), taken at the location of the defects:

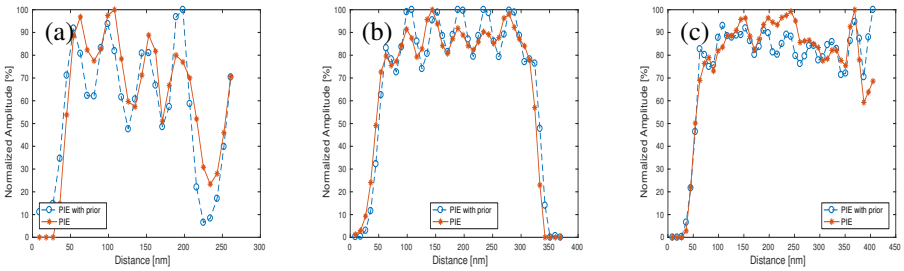

Fig. 10. Cross section of the reconstruction at the defect sites.a) 3 defects of 16 nm sidelength; b) 6 defects of 12 nm sidelength; c) 7 defects of 9 nm sidelength.

3.3. Effect of number of probe positions, SNR, initial guess and a comparison with TV regularization

3.3.1. Increasing the number of probe positions

Ptychography achieves a very robust reconstruction by exploiting translational diversity and redundancy in the data set. In the study presented earlier the data set included five probe positions. As the object is, in this case, entirely covered by the probe there is an high degree of redundancy in the data in spite of the few probe positions. However it can be interesting to see whether increasing the number of probe positions allows one to get to a reconstruction which is as good as in the case in which the prior term is included. We have performed this study for the case of the extrusions type defects. We have used 9 probe positions and the reconstruction has been carried out using the PIE. The 9 positions constitute a 3–by–3 grid which span, in steps of 200nm along the x and y directions, a square whose sidelength is 400nm. The reconstruction and the difference image is showed in Fig. 11.

Fig. 11(b) reports an overall better reconstruction of the defects with respect to Fig. 6(b). Still the reconstruction is not as satisfactory as the one obtained with 5 probe positions and the inclusion of the prior term. This is highlighted in Table 4.

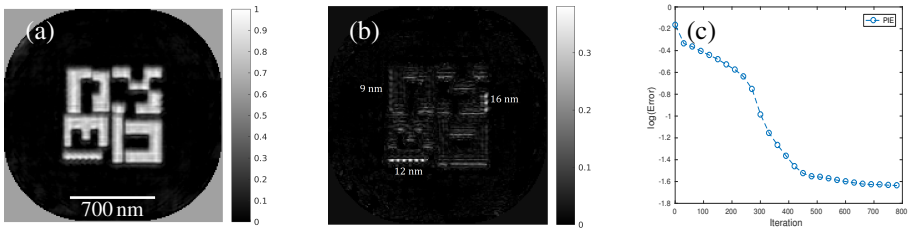

Fig. 11. Ptychographic reconstructions: a) magnitude as given by the PIE; b) difference image given by PIE; c) error.

Table 4. **Extrusions defects SNR**

Defect Size [nm]	SNR with prior (5 probe positions)	SNR PIE (9 probe positions)
16	3.8	3
12	5.3	2.5
9	1.6	0.6

3.3.2. Decreasing SNR

The effect of the noise is here studied. A decrease of the SNR constitutes a problem for the retrieval of fine features that weakly scatter light, because their signature could be below the noise floor. A workaround could be to increase the radiation dose, but this could in turn damage the sample. The role of the prior as a regularizer is helpful in this, as it stabilizes the inversion and enables one to achieve a better reconstruction when the SNR decreases. Here we have decreased the SNR from 110 dB to 100 dB and 90 dB. The results, for the extrusion defects, are given in Figs. 12 and 13.

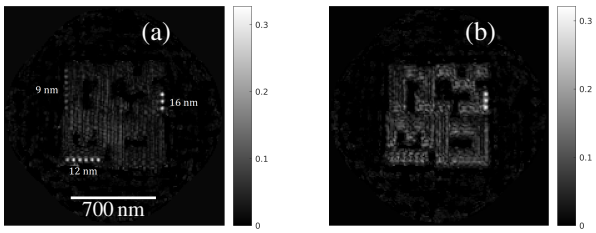

Fig. 12. a) and b) show the object difference metric ΔO in Eq. (12) obtained when SNR = 100 dB. a) PIE with prior; b) PIE.

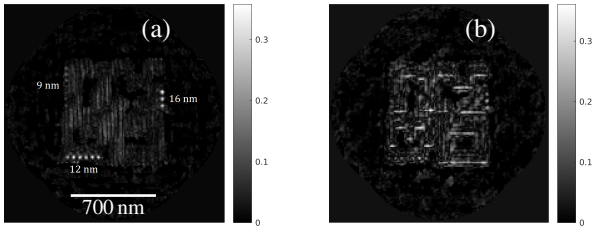

Fig. 13. a) and b) show the object difference metric ΔO in Eq. (12) obtained when SNR = 90 dB. a) PIE with prior; b) PIE.

All of the defects are still visible in Fig. 12(a), while in Fig. 13(b) only the bigger ones are visible. In Fig. 13 the 16 and 12 nm size defects are visible, but none of the defects can be detected by the PIE.

3.3.3. Using the prior as initial guess

The advantages of having a regularizer in Eq. (5) rather than using the prior only as a starting point were outlined in Section 2. Here we compare the results, for extrusion type defects, given using the update rule Eq. (9), the PIE and the PIE using the prior as starting guess. Fig. 14 highlights results for an SNR = 110 dB. Fig. 15 shows results for an SNR = 100 dB.

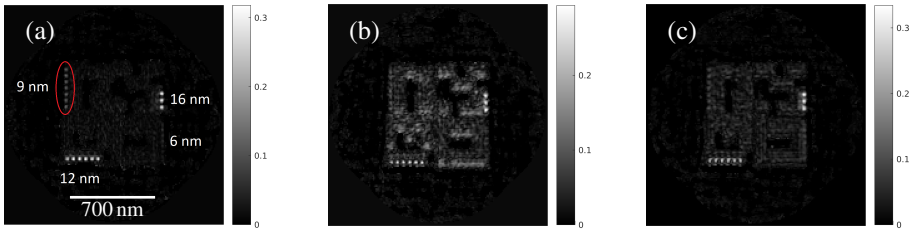

Fig. 14. Difference image SNR 110 dB: a) Update rule Eq. (9); b) PIE; c) PIE with prior as starting guess.

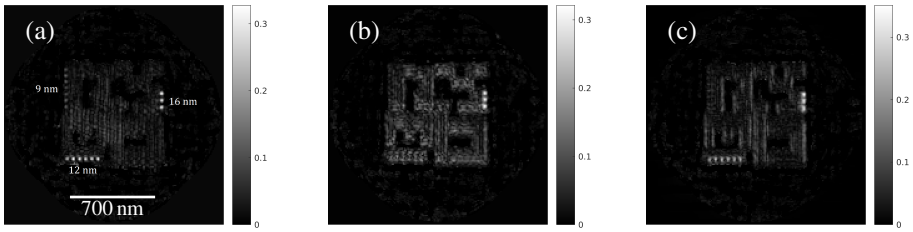

Fig. 15. Difference image SNR 100 dB: a) Update rule Eq. (9); b) PIE; c) PIE with prior as starting guess.

As it can be noticed a proper starting guess in the standard PIE yields a better reconstruction, however the absence of the regularizer – that stabilizes the reconstruction with respect to the noise and that promotes the retrieval of a better fit via a bias–variance trade–off – has a negative impact over the reconstruction.

3.3.4. Comparison with TV regularization

Recent works have reported the use of total variation as regularizer for denoising in ptychography [20–22]. Here we compare results obtained using the two different regularizers $R^{\alpha}(O_j) = \alpha||O_j(\boldsymbol{r}) - O_{p,j}(\boldsymbol{r})||^2$ and $R^{\alpha_{TV}}(O_j) = \alpha_{TV}||\nabla O_j(\boldsymbol{r})||_1$. Total variation is here solved via the alternating direction method of multipliers (ADMM). We solve the following problem at the j–th probe position:

$$\begin{aligned} \min_{O_j,p} \quad & \left\| |\mathcal{F}(P(\boldsymbol{r}-\boldsymbol{R}_j)O_j(\boldsymbol{r}))| - \sqrt{I_j(\boldsymbol{k})} \right\|^2 + \alpha_{TV}||p||_1 \\ \text{s.t.} \quad & p = \nabla O_j \end{aligned} \tag{13}$$

The augmented Lagrangian for real functional of complex variables can be written as [35]:

$$\begin{aligned} \mathcal{L}_{\rho}(O_j,p,y) = & \left\| |\mathcal{F}(P(\boldsymbol{r}-\boldsymbol{R}_j)O_j(\boldsymbol{r}))| - \sqrt{I_j(\boldsymbol{k})} \right\|^2 \\ & + 2\,\mathrm{Re}(\langle p - \nabla O_j, y\rangle) + \rho||p-\nabla O_j||^2 + \alpha_{TV}||p||_1 \end{aligned} \tag{14}$$

where $\langle\cdot,\cdot\rangle$ is the inner product over the complex space, $\mathrm{Re}(z)$ is the real part of z, y is the Lagrangian multiplier and ρ is the penalty parameter. Eq. (14) can be further simplified scaling the variable y by $1/\rho$ [36]. At iteration n the ADMM solves the following steps:

$$\begin{aligned} O_j^{n+1} &= \arg\min_{O_j} \mathcal{L}_{\rho}(O_j, p^n, y^n) \\ p^{n+1} &= \arg\min_{p} \mathcal{L}_{\rho}(O_j^{n+1}, p, y^n) \\ y^{n+1} &= y^n + \rho(p^{n+1} - \nabla O_j^{n+1}) \end{aligned} \tag{15}$$

the derivations necessary to minimize (15) are analogous to the ones that can be found elsewhere [21] and are here omitted. Problem (13) is minimized per probe position, the subproblem with respect to O_j is solved via a steepest descent, the subproblem with respect to p has a closed–form solution in the form of a soft–thresholding operator, and α_{TV} has been computed by the L–curve and it is equal to 1e-6 [37].

The object difference images for SNR=110 dB and for SNR=90 dB, for extrusions defects, are outlined in Fig. 16 and Fig. 17.

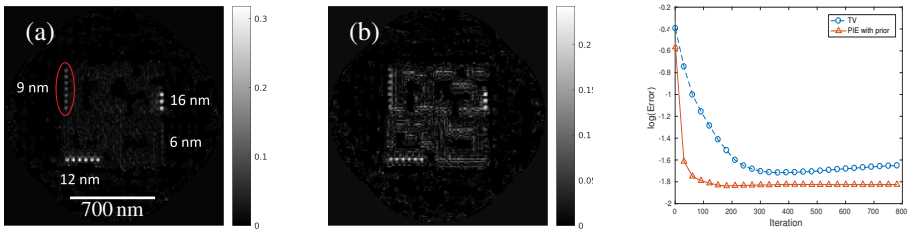

Fig. 16. Object difference image SNR 110 dB: a) Update rule (9); b) TV regularization; c) error.

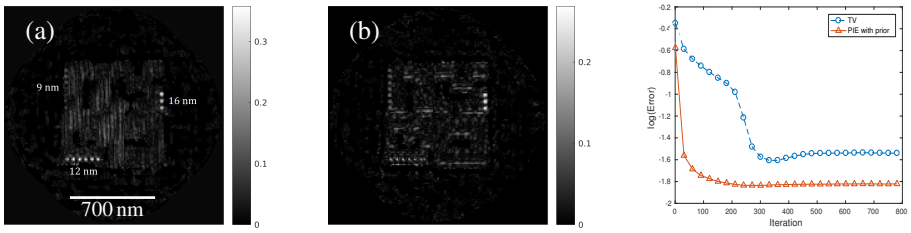

Fig. 17. Object difference image SNR 90 dB: a) Update rule (9); b) TV regularization; c) error.

Although TV regularization produces results that are qualitatively superior to the PIE (Figs. 6(b) and 13(b)) the proposed method that includes object prior in the optimization enables a more robust reconstruction.

4. Conclusions

We have presented a ptychographic algorithm that makes use of prior information about the transmission/reflection function of the object to improve the quality of the ptychographic reconstruction. Prior information is generated by employing the rigorous forward Maxwell solver – to properly account for the complex physics that contributes to the process of image formation – and by interpreting the outcome of the fully 3D rigorous simulations in terms of the "probe–times–object" ptychographic approximation of light–matter interaction. The method has been applied to the problem of actinic ptychographic inspection of extrusion and intrusion type defects in a patterned EUV mask layout. The numerical results indicate a steep improvement over the standard PIE in terms of quality of reconstruction and convergence. Smaller defects are detectable and all the defects are imaged with higher SNR. The method is shown to achieve better reconstructions using a smaller number of probe positions, to be more robust to the noise and to outperform total variation based ptychographic imaging. Although the algorithm has been applied to the specific technological problem of EUV mask inspection, this work can be interesting for the ones that employ lensless imaging for the reconstruction of nanostructures and that have prior information about the object structure available.

Funding

This project has received financial support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 675745.

Disclosures

The authors declare no conflicts of interest.

Acknowledgments

We are grateful to Mark Van Kraaij, ASML Research, for assistance with the use of the Maxwell solver and to the anonymous reviewers that have contributed to improve the quality of this manuscript.

References

1. V. Bakshi, *EUV lithography*, vol. PM2814 (SPIE Press, 2018).
2. D. Hellweg, S. Perlitz, K. Magnusson, R. Capelli, M. Koch, and M. Malloy, "Actinic review of euv masks: performance data and status of the aims euv system," in *Extreme Ultraviolet (EUV) Lithography VII,* vol. 9776, International Society for Optics and Photonics (SPIE, 2016), pp. 374 – 381.

3. K. Badger, E. Gallagher, K. Seki, G. McIntyre, T. Konishi, Y. Kodera, and V. Redding, "Evaluation of non-actinic EUV mask inspection and defect printability on multiple EUV mask absorbers," in *Photomask and Next-Generation Lithography Mask Technology XX,* vol. 8701, International Society for Optics and Photonics (SPIE, 2013), pp. 315 – 327.
4. M. Waiblinger, T. Bret, R. Jonckheere, and D. V. den Heuvel, "E-beam based mask repair as door opener for defect free EUV masks," in *Photomask Technology 2012,* vol. 8522, International Society for Optics and Photonics (SPIE, 2012), pp. 471 – 480.
5. J. Na, D. Lee, C. Do, H. seok Sim, J.-H. Lee, J. Kim, H.-S. Seo, H. Kim, and C. U. Jeon, "Application of actinic mask review system for the preparation of HVM EUV lithography with defect free mask," in *Metrology, Inspection, and Process Control for Microlithography XXXI,* vol. 10145, International Society for Optics and Photonics (SPIE, 2017), pp. 191 – 198.
6. J. M. Rodenburg and H. M. L. Faulkner, "A phase retrieval algorithm for shifting illumination," Appl. Phys. Lett. **85**, 4795–4797 (2004).
7. D. F. Gardner, B. Zhang, M. D. Seaberg, L. S. Martin, D. E. Adams, F. Salmassi, E. Gullikson, H. Kapteyn, and M. Murnane, "High numerical aperture reflection mode coherent diffraction microscopy using off-axis apertured illumination," Opt. Express **20**, 19050–19059 (2012).
8. C. L. Porter, M. Tanksalvala, M. Gerrity, G. Miley, X. Zhang, C. Bevis, E. Shanblatt, R. Karl, M. M. Murnane, D. E. Adams, and H. C. Kapteyn, "General-purpose, wide field-of-view reflection imaging with a tabletop 13 nm light source," Optica **4**, 1552–1557 (2017).
9. M. Odstrcil, J. Bussmann, D. Rudolf, R. Bresenitz, J. Miao, W. S. Brocklesby, and L. Juschkin, "Ptychographic imaging with a compact gas–discharge plasma extreme ultraviolet light source," Opt. Lett. **40**, 5574–5577 (2015).
10. P. Helfenstein, R. Rajeev, I. Mochi, A. Kleibert, C. A. F. Vaz, and Y. Ekinci, "Beam drift and partial probe coherence effects in euv reflective-mode coherent diffractive imaging," Opt. Express **26**, 12242–12256 (2018).
11. P. Ansuinelli, W. Coene, and P. Urbach, "EUV mask feature reconstruction via phase retrieval," in *Nanoengineering: Fabrication, Properties, Optics, Thin Films, and Devices XVI,* vol. 11089, International Society for Optics and Photonics (SPIE, 2019), pp. 299 – 305.
12. R. Rajendran, I. Mochi, P. Helfenstein, I. Mohacsi, S. Redford, A. Mozzanica, B. Schmitt, S. Yoshitake, and Y. Ekinci, "Towards a stand-alone high-throughput EUV actinic photomask inspection tool: RESCAN," in *Metrology, Inspection, and Process Control for Microlithography XXXI,* vol. 10145, International Society for Optics and Photonics (SPIE, 2017), pp. 199 – 210.
13. R. Horstmeyer, R. Y. Chen, X. Ou, B. Ames, J. A. Tropp, and C. Yang, "Solving ptychography with a convex relaxation," New J. Phys. **17**, 053044 (2015).
14. C. R. Vogel, *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, 2002).
15. A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation* (Society for Industrial and Applied Mathematics, 2005).
16. P. Sidorenko, O. Kfir, Y. Shechtman, A. Fleischer, Y. Eldar, M. Segev, and O. Cohen, "Sparsity-based super-resolved coherent diffraction imaging of one-dimensional objects," Nat. communications **6**, 8209 (2015).
17. A. Szameit, Y. Shechtman, E. Osherovich, E. Bullkich, P. Sidorenko, H. Dana, S. Steiner, E.-B. Kley, S. Gazit, T. Cohen-Hyams, S. Shoham, M. Zibulevsky, I. Yavneh, Y. Eldar, O. Cohen, and M. Segev, "Sparsity-based single-shot sub-wavelength coherent diffractive imaging," Nat. materials **11**, 455–9 (2012).
18. T. Zhang, C. Godavarthi, P. C. Chaumet, G. Maire, H. Giovannini, A. Talneau, M. Allain, K. Belkebir, and A. Sentenac, "Far-field diffraction microscopy at λ/10 resolution," Optica **3**, 609–612 (2016).
19. D. Gardner, M. Tanksalvala, E. Shanblatt, X. Zhang, B. Galloway, C. Porter, R. Karl, C. Bevis, D. Adams, M. Murnane, and G. Mancini, "Subwavelength coherent imaging of periodic samples using a 13.5 nm tabletop high-harmonic light source," Nat. Photonics **11** (2017).
20. H. Chang, P. Enfedaque, J. Zhang, J. Reinhardt, B. Enders, Y.-S. Yu, D. Shapiro, C. G. Schroer, T. Zeng, and S. Marchesini, "Advanced denoising for x-ray ptychography," Opt. Express **27**, 10395–10418 (2019).
21. H. Chang, P. Enfedaque, and S. Marchesini, "Iterative Joint Ptychography-Tomography with Total Variation Regularization," arXiv e-prints arXiv:1902.05647 (2019).
22. V. Nikitin, S. Aslan, Y. Yao, T. Biçer, S. Leyffer, R. Mokso, and D. Gürsoy, "Photon-limited ptychography of 3d objects via bayesian reconstruction," OSA Continuum **2**, 2948–2968 (2019).
23. S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb, "Solving inverse problems using data-driven models," Acta Numer. **28**, 1–174 (2019).
24. Q. Fang, R. Moore, D. Kopans, and D. Boas, "Compositional-prior-guided image reconstruction algorithm for multi-modality imaging," Biomed. optics express **1**, 223–235 (2010).
25. A. Mohammad-Djafari, "Bayesian approach with prior models which enforce sparsity in signal and image processing," EURASIP J. on Adv. Signal Process. **2012**, 52 (2012).
26. S. Zayko, E. Mönnich, M. Sivis, D. M. Mai, T. Salditt, S. Schäfer, and C. Ropers, "Coherent diffractive imaging beyond the projection approximation: waveguiding at extreme ultraviolet wavelengths." Opt. express **23 15**, 19911–21 (2015).
27. J. Rodenburg and A. Maiden, "Ptychography," in *Springer Handbook of Microscopy,* P. W. Hawkes and J. C. H. Spence, eds. (Springer International Publishing, Cham, 2019), pp. 819–904.
28. X. Wei, A. J. Wachters, and H. P. Urbach, "Finite-element model for three-dimensional optical scattering problems,"

J. Opt. Soc. Am. A **24**, 866–881 (2007).
29. J. Pomplun, S. Burger, L. Zschiedrich, and F. Schmidt, "Adaptive finite element method for simulation of optical nano structures," physica status solidi (b) **244**, 3419 (2007).
30. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics (Springer New York Inc., New York, NY, USA, 2001).
31. I. Mochi, K. A. Goldberg, B. L. Fontaine, A. Tchikoulaeva, and C. Holfeld, "Actinic imaging of native and programmed defects on a full-field mask," in *Extreme Ultraviolet (EUV) Lithography,* vol. 7636, International Society for Optics and Photonics (SPIE, 2010), pp. 425 – 433.
32. M. C. van Beurden, "Fast convergence with spectral volume integral equation for crossed block-shaped gratings with improved material interface conditions," J. Opt. Soc. Am. A **28**, 2269–2278 (2011).
33. M. C. van Beurden, "A spectral volume integral equation method for arbitrary bi-periodic gratings with explicit Fourier factorization," Prog. In Electromagn. Res. B **36**, 133–149 (2012).
34. A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," Ultramicroscopy **109**, 1256 – 1262 (2009).
35. L. Li, X. Wang, and G. Wang, "Alternating direction method of multipliers for separable convex optimization of real functions in complex variables," Math. Probl. Eng. **2015**, 1–14 (2015).
36. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Found. Trends Mach. Learn. **3**, 1–122 (2011).
37. P. C. Hansen, "The l-curve and its use in the numerical treatment of inverse problems," in *in Computational Inverse Problems in Electrocardiology, ed. P. Johnston, Advances in Computational Bioengineering,* (WIT Press, 2000), pp. 119–142.