Extended Nijboer-Zernike (ENZ) based mask imaging: Efficient coupling of electromagnetic field solvers and the ENZ imaging algorithm

Olaf T.A. Janssen*[a], Sven van Haver[a], Augustus J.E.M. Janssen[b], Joseph J.M. Braat[a], H. Paul Urbach[a] and Silvania F. Pereira[a]

[a]IST-Optics Research Groep, Delft University of Technology, Lorentzweg 1, NL-2628 CJ Delft, The Netherlands;
[b]Philips Research Europe, HTC 36, NL-5656 AE Eindhoven, The Netherlands

ABSTRACT

Results are presented of mask imaging using the Extended Nijboer-Zernike (ENZ) theory of diffraction. We show that the efficiency of a mask imaging algorithm, derived from this theory, can be increased. By adjusting the basic Finite Difference Time Domain (FDTD) algorithm, we can calculate the near field of isolated mask structures efficiently, without resorting to periodic domains. In addition, the calculations for the points on the entrance sphere of the imaging system can be done separately with a Fourier transformed Stratton-Chu near-to-far-field transformation. By clever sampling in the radial direction of the entrance pupil, the computational effort is already reduced by at least a factor of 4.

Keywords: mask imaging, Extended Nijboer-Zernike theory, pupil sampling, Stratton-Chu, FDTD

1. INTRODUCTION

Numerical simulations of mask imaging are vital for the lithographic community. In the quest of obtaining increasingly smaller structures on a wafer, the design of a typical lithographic mask contains features with dimensions in the order of the wavelength. In addition, the mask has become optically thick and hyper-NA immersion lithography is currently often applied, so that polarization effects influence now to a large degree the image. To correctly predict the image from a given mask design computational tools are required that are fully vectorial. In reverse, the mask design does not follow straightforwardly from a desired wafer image. The entire mask design process has become a painstakingly long iterative procedure.

Computational lithography is mature; efficient algorithms have been developed that often run on dedicated hardware. Rigorous electromagnetic field solvers have been around for several decades and can easily replace non-rigorous methods to simulate the complex interaction of illumination with a lithographic mask. Hopkins' approach[1] (Transmission Cross-Coefficients computation) is widely used to calculate aerial images. It is, in basis, a scalar treatment which has been extended to support vectorial calculations.[2] Nonetheless, the method is based on assumptions that are increasingly violated for future nodes. In addition, it requires the use of periodic mask patterns, which is uncommon for realistic masks.

As an alternative to existing algorithms, we present Extended Nijboer-Zernike (ENZ) based mask imaging.[3] The method is fully vectorial, rigorous and does not rely on assumptions made by Hopkins' method. It is a source point integration method (or Abbe's method[4]) which is generally considered to be computationally inefficient. However, we believe that ENZ-based mask imaging has the promise of being relatively efficient. In this paper, we present an efficient coupling between rigorous electromagnetic field solvers and the ENZ-imaging algorithm. This coupling enables the new mask imaging algorithm to fit seamlessly into existing mask design schemes.

In Section 2, the ENZ-based mask imaging algorithm is explained with a focus on connecting electromagnetic field solvers to the imaging algorithm. Section 3 discusses several computational results of the imaging algorithm, and our conclusions are summarized in Section 4.

*Olaf T.A. Janssen, E-mail: o.t.a.janssen@tudelft.nl, Telephone: +31 (0)15 2782116, Fax: +31 (0)15 2788105

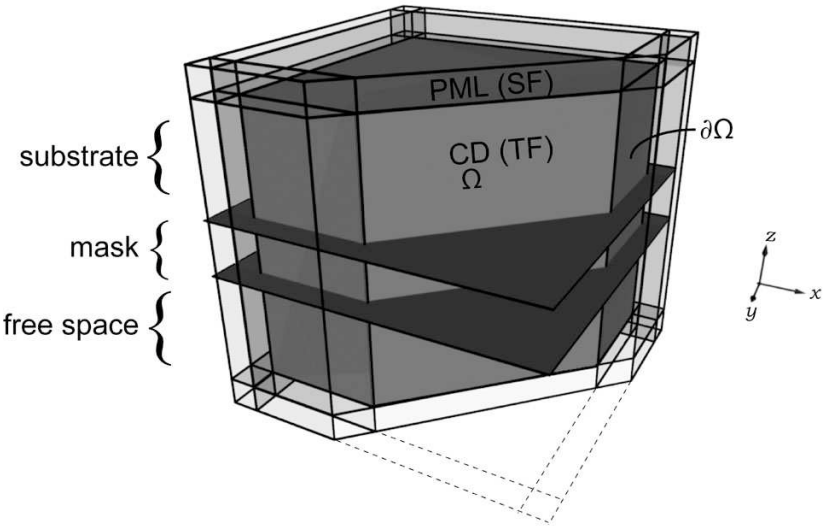

Figure 1. Cross-sectional representation of the computational domain (CD) surrounded by a perfectly matched layer (PML) and the interface $\partial\Omega$ that separates the CD and the PML, as used in the FDTD solver. It is shown that mask and substrate interfaces can cut through the interface $\partial\Omega$ to represent infinitely extending material.

2. METHODOLOGY

2.1 ENZ-based mask imaging

The mask imaging algorithm simulates the entire optical system that is typical of current lithographic set-ups. The source distribution is discretized into point sources, for which the image intensity is calculated independently. The interaction of the incident light with the absorbing mask is calculated with the FDTD model described below. Instead of an FDTD model, any rigorous solver can be used, as long as it can solve the electromagnetic near-field of the mask. Using a Stratton-Chu near-to-far-field transformation, which will also be described in more detail, the electric field at the spherical entrance pupil of the lens system is calculated. The field is then decomposed in a Zernike expansion. From this Zernike representation, the aerial images produced by an optical system with arbitrary aberrations follow by calculating the full vectorial Debye diffraction integral using ENZ diffraction theory.[5,6]

2.2 FDTD

The FDTD method is an extensively used rigorous electromagnetic field solver. Since the introduction of the basic algorithm,[7] many extensions have been proposed and separately to overcome the limitations of the basic model. A good overview of these can be found in the book of Taflove and Hagness.[8] Since the FDTD method operates in the time domain, it is well-suited for calculating the spectral or temporal response of an illuminated geometry. For the lithography application, we are only interested in steady-state solutions, so that a frequency domain method could be considered to be more appropriate. However, the relatively low memory requirements for large geometries and the extendibility of the FDTD method have made it a popular tool also in the lithographic community.

The FDTD implementation used is an in-house developed code at Delft University of Technology, which has previously been benchmarked.[9] A rectangular three-dimensional computational domain (CD) is defined which is the volume that encloses all non-trivial scatterers of the problem, see Fig. 1. In this case it encloses the active region of the mask. The orthogonal mesh in the CD consists of rectangular Yee-elements and is non-uniform. By aligning the non-uniform mesh with the material interfaces, staircasing errors of the in general non-conforming mesh are minimized. An absorbing layer surrounds the CD in all directions. This perfectly matched layer[10] (PML) is, in addition, non-reflective and thus fully transparent to outgoing waves. As a result, it appears numerically as if the CD is surrounded by homogeneous materials free of scatterers.

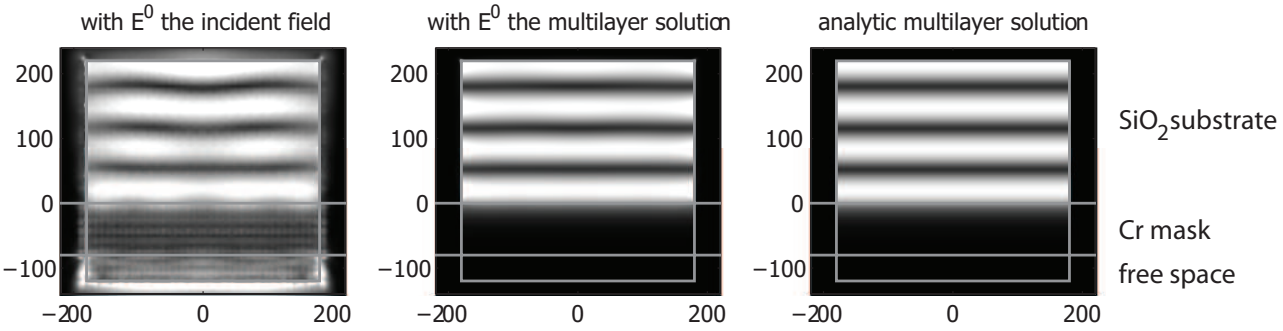

Figure 2. Near field (E_x) of an x-polarized orthogonally incident plane wave on a lithographic mask structure consisting of a featureless 80nm thick Cr absorbing mask on a SiO_2 substrate suspended in air. (left) Calculated using FDTD, with the incident plane wave as zero field E^0, (center) using FDTD, with the solution to the layered problem as zero field E^0, (right) the analytic solution to the layered problem. The gray thick lines denote the interface between the CD and the PML, as well as the material interfaces that cross into the PML.

In common FDTD implementations, letting material interfaces cross into the PML is problematic as illustrated in the left image of Fig. 2. This issue is usually circumvented by taking a very large, memory and computationally inefficient CD with (quasi)periodic boundary conditions. In practice, for Hopkins' method the simulation of a (quasi)periodic problem is required. Instead, the newly proposed ENZ-mask imaging algorithm can produce images of isolated mask features. Our FDTD implementation does allow the interfaces of a multilayered structure to cross into the PML. Because Maxwell's equations are linear we can write the total field (TF), $\mathbf{E}^T$, as the sum of two solutions to Maxwell's equations,

$$\mathbf{E}^T = \mathbf{E}^S + \mathbf{E}^0, \tag{1}$$

where $\mathbf{E}^0$ is usually the incident field in free space and $\mathbf{E}^S$ is called the scattered field (SF). As in other implementations, we calculate the total field in the CD and the scattered field in the surrounding PML region. In a standard FDTD algorithm, the fields in the two regions are coupled by adding or subtracting the incident field on the boundary of the computational domain and the PML. However, on this boundary in a multilayered geometry, the incident field is not a good solution to Maxwell's equations, because the incident field is defined for free space. The algorithm therefore produces incorrect results. Instead of using the incident field for $\mathbf{E}^0$, we use the steady-state solution to the layered geometry without mask features which does give correct results in the steady-state of the FDTD simulation as can be seen in Fig. 2. This multilayer solution can easily be calculated analytically for general incident fields.

The time domain simulation is stopped when the fields in the computational domain have become harmonic. While FDTD calculates with real field values, using the information from one harmonic cycle, complex field values can be deduced so that phase information is contained in the output.

Our FDTD method implements the PML as a convolutional PML[11] and allows for media with negative permittivity and non-zero absorption by implementing an auxiliary differential equations (ADE) technique for dispersive media.[12] Next to the absorbing PML boundary, quasi-periodic boundary conditions can be applied using the so-called sine-cosine technique[13] for oblique incident plane waves. Our model also includes the modules for calculating the analytic multilayer solution and the Stratton-Chu near-to-far-field calculation described below. The code is parallelized for efficient execution on multi-processor machines.

2.3 Near-to-far-field calculation

Because the distance between the mask and the entrance pupil is far larger than the wavelength and the dimensions of the features on the mask, the Fraunhofer approximation for the field on the entrance pupil surface is valid. Since the distance is comparable to the size of the entrance pupil, we have to consider a spherical entrance pupil.

In order to link the calculated near-fields to the entrance pupil of the optical system, we use the Stratton-Chu formula.[14] It relates the scattered fields $\mathbf{E}^S$ and $\mathbf{H}^S$ on the boundary, $\partial\Omega$, of a given domain to the field at any point in or outside the domain, and is given by

$$\mathbf{E}^S(\mathbf{r}') = -\iint_{\partial\Omega} (\mathbf{n}\times\mathbf{E}^S(\mathbf{r}))\underline{\underline{G}}_H(\mathbf{r},\mathbf{r}') - (\mathbf{n}\times\mathbf{H}^S(\mathbf{r}))\underline{\underline{G}}_E(\mathbf{r},\mathbf{r}')\,\mathrm{d}r^2. \tag{2}$$

Herein, $\underline{\underline{G}}(\mathbf{r},\mathbf{r}')$, is Green's tensor of the layered system, because we allow material interfaces to cross the Stratton-Chu integration surface. For a general layered system this term is difficult to calculate analytically. In addition, the Fraunhofer approximation follows from this inefficiently. First, we would need to calculate the field at many points on a large flat surface. Second, using a Fourier transform of the fields there we would obtain the Fraunhofer far field. Instead, we implement the Fourier transformed Stratton-Chu formula:[15]

$$\mathcal{F}[\mathbf{E}^S](k_x,k_y,z) = -\iint_{\partial\Omega} (\mathbf{n}\times\mathbf{E}^S(\mathbf{r}))\mathcal{F}[\underline{\underline{G}}_H](\mathbf{r},\mathbf{k}) - (\mathbf{n}\times\mathbf{H}^S(\mathbf{r}))\mathcal{F}[\underline{\underline{G}}_E](\mathbf{r},\mathbf{k})\,\mathrm{d}r^2, \tag{3}$$

which now uses the Fourier transformed Green's tensor, $\mathcal{F}[\underline{\underline{G}}](\mathbf{r},\mathbf{k})$, of a layered system. This tensor is efficiently calculated in the same algorithm that gives the analytic multilayer solution for the FDTD simulation. By using the Fourier transformed equation, the far field at the location on the sphere determined by the wavenumbers k_x, k_y and distance z is almost directly obtained:

$$\mathbf{E}^S_{\mathrm{far}}(k_x,k_y,z) = \mathcal{F}[\mathbf{E}^S]\frac{k_z}{|\mathbf{k}|}. \tag{4}$$

In order to do the numerical integration of Eq. (3), the rectangular boundaries of the FDTD computational domain are discretized to a uniform, orthogonal Stratton-Chu grid. Because field values on the FDTD grid are not collocated but staggered, the FDTD field values are linearly interpolated onto the Stratton-Chu grid.

An advantage of the described Stratton-Chu method is that separate points in the far field can be calculated, corresponding directly to points on the spherical entrance pupil of the optical system. It is therefore very suited for parallelized computation. More importantly, it gives the freedom to choose any kind of entrance pupil sampling, which is described in more detail below.

2.4 Sampling models

In the beginning of this section, it is explained that we require the Zernike representation of the field F in the entrance pupil, which is given by

$$F(\rho,\theta) = \sum_{n,m} \beta_n^m R_n^{|m|}(\rho)e^{im\theta} \qquad ,0\le\rho\le 1, \quad 0\le\theta\le 2\pi, \tag{5}$$

where ρ and θ are the radial and angular polar coordinates and $R_n^{|m|}$ are the radial Zernike polynomials.[4] The summation is over all integer n, m such that $n-|m|$ is even and ≥ 0. We assume that F is a relatively smooth function of ρ and θ, which implies that β_n^m rapidly decay in $|m|$ and n. By orthogonality of the Zernike terms $R_n^{|m|}(\rho)e^{im\theta}$, the β_n^m are given in integral form as

$$\beta_n^m = \frac{n+1}{\pi}\int_0^1\int_0^{2\pi} F(\rho,\theta)R_n^{|m|}(\rho)e^{-im\theta}\rho\,\mathrm{d}\rho\mathrm{d}\theta. \tag{6}$$

The goal is now to find a sampling of θ and ρ in the pupil, so that the non-polynomial functions are accurately integrated with as few points as possible. We shall restrict ourselves to separable sampling schemes. It is convenient to take the θ-sampling set T as

$$T = \{\theta_l = \frac{2\pi l}{L} \quad | \quad l = 0,1,...,L-1\}. \tag{7}$$

We now address the issue of choosing an appropriate ρ-sampling set R. We note that because our scheme is separable, the ρ-sampling is independent of m. If the sampling would be dependent on m, we would require a

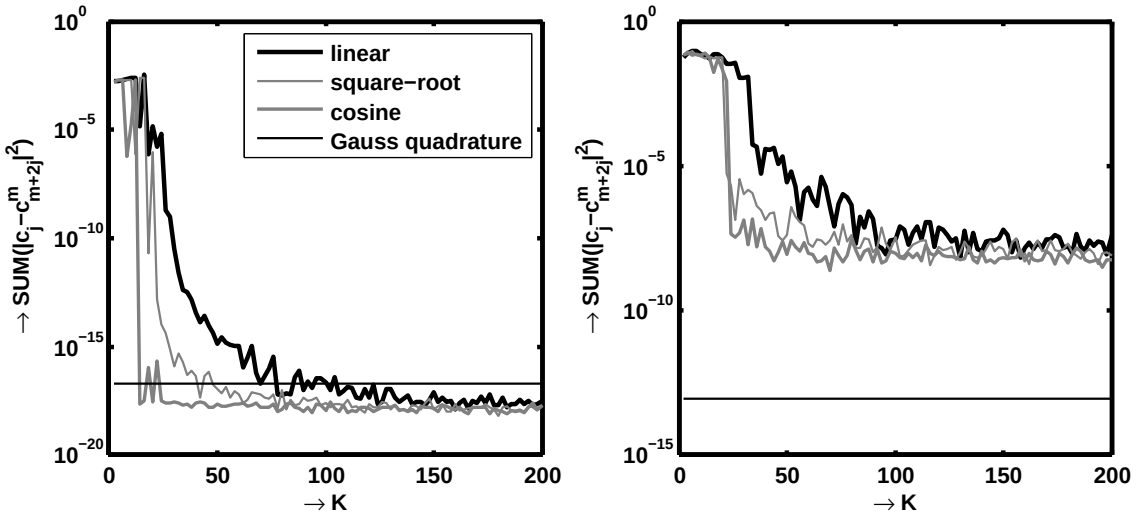

Figure 3. Fitting of the function $F^m(\rho) = J_m(2\pi a\rho)$ to a Zernike expansion with coefficients c_j, with $j = 0...20$, compared to analytical solution c^m_{m+2j}. The errors are shown of the linear, square-root and cosine sampling scheme using K sampling points, and the Gaussian quadrature error with fixed number of sampling points $K = j_{\max} = 21$. Left graph shows the errors for the case of a slowly varying function ($m = 10$, $a = 1$), while the right graph shows the errors for a violently oscillating function ($m = 20$, $a = 5$).

new set of sampling points for every value of m while we aim to decrease the number of sampling points needed. To that end, we consider the following sampling schemes:

$$\text{linear sampling:} \qquad \rho_k = \frac{k+\frac{1}{2}}{K} \qquad , k = 0, 1, ..., K-1, \tag{8}$$

$$\text{square-root sampling:} \qquad \rho_k = (\frac{k+\frac{1}{2}}{K})^{\frac{1}{2}} \qquad , k = 0, 1, ..., K-1. \tag{9}$$

$$\text{cosine sampling:} \qquad \rho_k = \cos\frac{k+\frac{1}{2}}{2K}\pi \qquad , k = 0, 1, ..., K-1. \tag{10}$$

The linear sampling scheme is commonly used in cases where the pupil function is only available in sampled form on circles of linearly increasing radii.[16] Such a restriction does not hold in our application. The square-root sampling scheme occurs in several publications.[17,18] Cosine sampling, also known as Chebyshev sampling, is advocated in general terms in Eisinberg et al.[19] A detailed study on developing a sampling scheme taylored to Zernike functions will be published in the near future.

3. COMPUTATIONAL RESULTS

We shall now compare the sampling schemes in (8-10) to one another and to a benchmarking method based on an appropriate Gaussian quadrature scheme based on Jacobi polynomials. Gaussian quadrature is not a separable method and is therefore not suited for practical use, as mentioned before. For a tilted Bessel test function of order m, optimal weighting coefficients are determined using least-squares fitting for the three separable sampling schemes. The reconstructed function is then compared to the true coefficients (that can be found explicitly), see Fig. 3. It is clear that for both the low and high order test function the cosine sampling gave more accurate results at low numbers of sampling points. Note also that Gauss quadrature fits give a fixed optimal number of sampling points for a given fit order. As can also be seen, in some cases the separable sampling schemes can give better accuracy than the Gaussian quadrature fit.

To show the coupling of the FDTD with the ENZ imaging algorithm and the influence of the entrance pupil sampling at work we ran the mask imaging algorithm for a few test cases. We simulated a 80nm thick Cromium

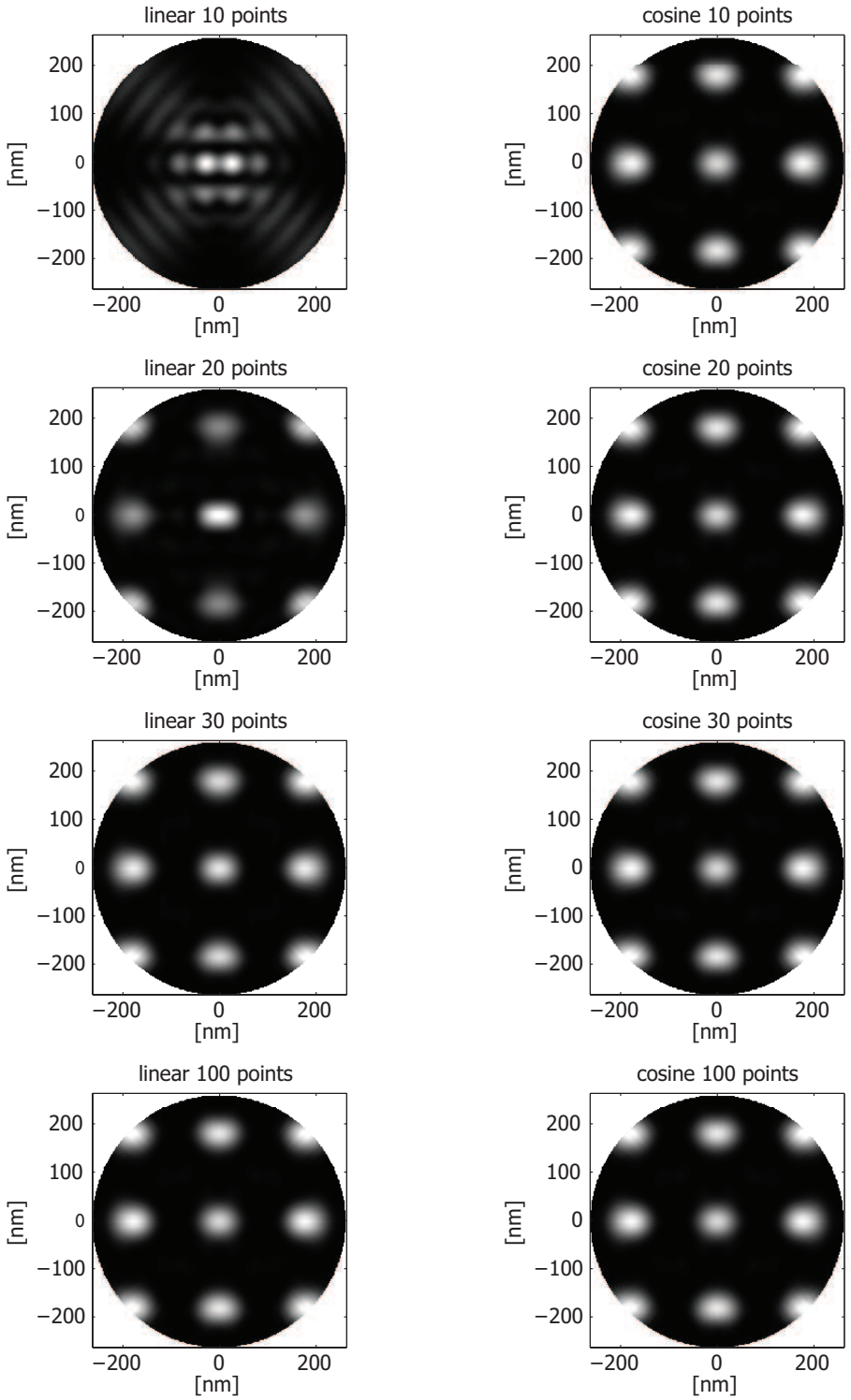

Figure 4. Intensity plots of aerial images in best focus produced by the ENZ-based mask imaging algorithm. All images are calculated from a single FDTD near field calculation of a 3x3 array of contact holes as described in the text. The far field is calculated on the entrance pupil with linear and cosine sampling and with various sampling densities in the ρ direction.

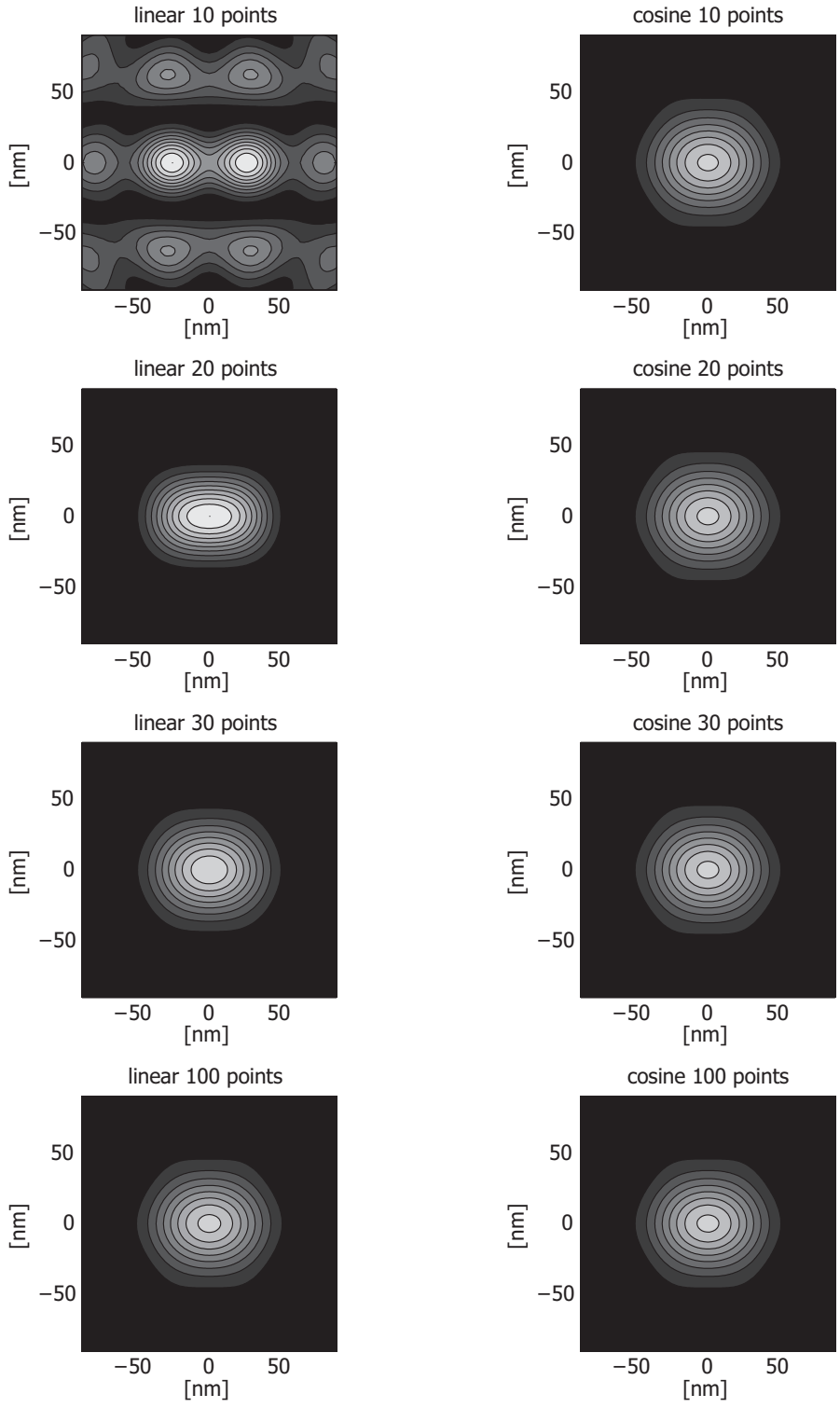

Figure 5. Same as Fig. 4, but zoomed in at the central contact hole.

mask ($\epsilon = -3.32 + 3.42i$) with a 3x3 array of square contact holes with a diameter of 180nm and a center-to-center separation of 720nm. The mask was placed on a SiO_2 ($\epsilon = 2.25$) substrate suspended in air. We assumed 193nm Köhler illumination originating from a single point source, which translated to an orthogonally incident x-polarized plane wave hitting the substrate side for the FDTD simulation. Note that the boundary conditions of the FDTD simulation were non-periodic.

The small mask features were imaged in a simulated immersion system with NA = 2.2 and magnification factor of 4, which leads to an effective numerial aperture on the object side of NA = 0.55. The optical lens system was assumed to be aberration-free.

Next, we defined eight test cases for the entrance pupil sampling. We used linear sampling with 10, 20, 30 and 100 points for ρ, and cosine sampling with 10, 20, 30, and 100 points. For all test cases we used 50 sampling points for θ, so that all errors can be assumed to originate from an insufficient sampling in ρ. The field values on the sampling points were used as input for the ENZ-imaging algorithm, which resulted in the aerial images shown in Fig. 4. Both images produced with the 100 point sampling look identical. For the images produced using linear sampling, it is clear that the image deteriorates if we use less sampling points. For 10 points the 3x3 array of contact holes does not show up correctly, for 20 points the central hole is much brighter than the outside holes and its image appears to be squeezed in the y-direction, finally for 30 sampling points there is no visible difference with the 100 sampling point image. The differences are more clearly visible in Fig. 5. By contrast, the image produced using cosine sampling of only 10 sampling points was already not visibly different from the 100 sampling point image. The Root Mean Square (RMS) error of the image using 10 point cosine sampling compared to the 100 point cosine sampling is around one percent (0.018). In comparison, the RMS error of the image using 30 point linear sampling with the 100 point linear sampling is 0.17. This RMS value only came near one percent for more than 40 linearly chosen sampling points. By switching from the linear sampling scheme to a cosine sampling scheme we could, in this example, reduce the number of points with at least a factor of 4, while retaining the same accuracy.

4. DISCUSSION AND CONCLUSIONS

While current mask imaging algorithms cannot be beaten in terms of speed, there may be need for rigorous mask imaging algorithms in the future. The fully vectorial ENZ-based mask imaging algorithm is such an algorithm. It can already image a given entrance pupil distribution quite efficiently, but it also requires an efficient coupling to rigorous electromagnetic solvers. We showed that, by using an FDTD and the Fourier transformed Stratton-Chu formula, we can choose which points on the entrance pupil to calculate. We have shown that we can already achieve a substantial computational reduction of at least a factor of 4 by switching to a simple cosine sampling scheme from a standard linear sampling scheme. The fit ultimately depends on the complexity of the entrance pupil distribution. A more detailed study on the relation between sampling schemes and the Zernike functions is under way.

REFERENCES

1. H.H. Hopkins. On the diffraction theory of optical images. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 217(1130):408–432, 1953.
2. D.G. Flagello, T. Milster, and A.E. Rosenbluth. Theory of high-NA imaging in homogeneous thin films. *Journal of the Optical Society of America A*, 13(1):53–64, 1996.
3. S. van Haver, O.T.A. Janssen, J.J.M. Braat, A.J.E.M. Janssen, H.P. Urbach, and S.F. Pereira. General imaging of advanced 3D mask objects based on the fully-vectorial Extended Nijboer-Zernike (ENZ) theory. *Proceedings of the SPIE*, 6924, 2008.
4. M. Born and E. Wolf. *Principles of Optics*. Cambridge University Press, New York, 2002.
5. J.J.M. Braat, P. Dirksen, and A.J.E.M. Janssen. Assessment of an extended Nijboer-Zernike approach for the computation of optical point-spread functions. *Journal of the Optical Society of America A*, 19(5):858–870, 2002.
6. J.J.M. Braat, P. Dirksen, A.J.E.M. Janssen, and A.S. van de Nes. Extended Nijboer-Zernike representation of the vector field in the focal region of an aberrated high-aperture optical system. *Journal of the Optical Society of America A*, 20(12):2281–2292, 2003.

7. K. Yee. Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Transactions on Antennas and Propagation*, 14(3):302–307, 1966.
8. A. Taflove and S.C. Hagness. *Computational electrodynamics*. Artech House, Boston, 2005.
9. P. Lalanne, M. Besbes, J.P. Hugonin, S. van Haver, O.T.A. Janssen, A.M. Nugrowati, M. Xu, S.F. Pereira, H.P. Urbach, A.S. van de Nes, P. Bienstman, G. Granet, A. Moreau, S. Helfert, M. Sukharev, T. Seideman, F. Baida, B. Guizal, and D. van Labeke. Numerical analysis of a slit-groove diffraction problem. *Journal European Optical Society-Rapid Publications*, 2:07022, 2007.
10. J.P. Berenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics*, 127(2):363–379, 1996.
11. J.A. Roden and S.D. Gedney. Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media. *Microwave and Optical Technology Letters*, 27:334–339, 2000.
12. M. Fujii, M. Tahara, I. Sakagami, W. Freude, and P. Russer. High-order FDTD and auxiliary differential equation formulation of optical pulse propagation in 2-D Kerr and Raman nonlinear dispersive media. *IEEE Journal of Quantum Electronics*, 40(2):175–182, 2004.
13. D.T. Prescott and N.V. Shuley. Extensions to the FDTD method for the analysis of infinitely periodic arrays. *IEEE Microwave and Guided Wave Letters*, 4(10):352–354, 1994.
14. J.A. Stratton and L.J. Chu. Diffraction theory of electromagnetic waves. *Physical Review*, 56(1):99–107, 1939.
15. X. Wei, A.J. Wachters, and H.P. Urbach. Finite-element model for three-dimensional optical scattering problems. *Journal of the Optical Society of America A*, 24(3):866–881, 2007.
16. A. Prata and W. Rusch. Algorithm for computation of Zernike polynomials expansion coefficients. *Applied Optics*, 28(4):749–754, 1989.
17. H.H. Hopkins. Diffraction theory of laser read-out systems for optical video discs. *Journal of the Optical Society of America A*, 69(1):4–24, 1979.
18. J.L. Rayces. Least-squares fitting of orthogonal polynomials to the wave-aberration function. *Applied Optics*, 31(13):2223–2228, 1992.
19. A. Eisinberg, G. Franzé, and P. Pugliese. Vandermonde matrices on Chebyshev points. *Linear Algebra and its Applications*, 283(1):205–219, 1998.