Vectorial aerial-image computations of three-dimensional objects based on the Extended Nijboer-Zernike theory

S. van Haver[1,*], J.J.M. Braat[1], A.J.E.M. Janssen[2], O.T.A. Janssen[1], S.F.Pereira[1]

[1] *Delft University of Technology, IST-Optics Research Group,*
Lorentzweg 1, NL-2628CJ Delft, The Netherlands.

[2] *Philips Research Europe,*
HTC 36, NL-5656AE Eindhoven, The Netherlands

[*] *Corresponding author: s.vanhaver@tudelft.nl*

In this paper we present details on a novel imaging algorithm based on the Extended Nijboer-Zernike (ENZ) theory of diffraction. We derive integral expressions relating the electric field distribution in the entrance pupil of an optical system to the electric field in its focal region. The evaluation of these integrals is made possible by means of a highly accurate and efficient series expansion similar to those occurring in standard ENZ-theory. Based on these results an ENZ-imaging scheme is constructed and evaluated in detail with attention for the convergence properties and computational complexity of the new method. © 2008 Optical Society of America

1. Introduction

In present day optics, the ability to obtain accurate images of a general object by means of computational methods is of great importance. Computational imaging not only provides valuable information on the image performance of proposed designs for optical systems, it also enables a detailed study of image formation itself. Although a large number of different imaging algorithms exist, all of them can be classified according to whether they operate in the spatial domain or in the frequency or Fourier domain. In the frequency domain approach, the assumption has to be made that the response of the system is space-invariant, implying that the imaging system generates an invariant point response to the coherent near-field issued from the object. When operating in the spatial domain, this assumption is not

explicitly needed but, once incorporated, it leads to an important reduction in calculation time. For this reason, although not strictly required, spatial invariance is also included in the spatial domain method. Regarding the frequency domain method, its origin is found in the Abbe microscope imaging theory [1]. It has been generalized to imaging problems by the work of Duffieux [2], Maréchal [3] and Hopkins. The latter author has extended the Fourier or frequency-domain approach to imaging with an extended source (partially coherent illumination) [4], including a defocussing of the image [5]. More recent extensions of the Fourier approach to imaging allow the inclusion of imaging aberrations and the effects of 'vector imaging' at high numerical aperture [6-8]. The Fourier approach has become very popular because of its possibility to employ the Fast Fourier Transform in numerical computations [9] and, more recently, to apply fast algorithms to obtain the so-called transmission cross-correlation coefficients [10]. The fast image calculation does however require an inherent periodicity in the object. If this periodicity is not present, it can be artificially introduced by periodically repeating the specific part of the object that has to be imaged. In this case the continuous frequency spectrum of the object is replaced by a discrete comb-like spectrum which introduces certain artifacts in the final image.

The alternative to frequency domain imaging is the calculation of the convolution of the object function with the point response function of the imaging system [11]. The first analytical result for the ideal response function or point-spread function goes back to Airy [12]. Through-focus imaging with an ideal system has been studied by Lommel [13]. Refinements including the influence of aberrations have been studied by Strehl [14], Conrady [15,16] and Picht [17]. A systematic study of imaging in the presence of aberrations has been carried out by Zernike and Nijboer [18-20], with the use of their circle polynomials on the circular exit pupil as the key contribution. The calculation of the point-spread function at high numerical aperture was initiated by Ignatowsky [21] and further developed by Hopkins [22] and Wolf [23]. In this latter publication, the high-numerical-aperture point-spread function through-focus was presented using a numerical evaluation of the Debye diffraction integral. This approach has become the standard one to treat point-spread function calculation using so-called 'vector diffraction'. Instead of the numerical evaluation of a diffraction integral, special functions have been proposed to represent the wave field in the exit pupil that yield a direct analytic solution in the image region, as for example, Gauss-Laguerre modal expansions [24], multipole functions [25], radial prolate spheroidal functions [26], spherical harmonics [27], Walsh functions [28]. Most of the special functions quoted above suffer from the fact that they cannot easily cope with the sharp edge that is encountered in optical diffraction problems, namely the circular 'hard' diaphragm that sharply limits the lateral extent of the imaging bundles in an optical system. For that reason, we use a recent extension of the method for high-numerical-aperture point-spread calculation according to [23]

2

that allows the analytic treatment of the influence of aberrations and defocusing. In a series of papers, the so-called Extended Nijboer-Zernike (ENZ) theory was presented that provides the user with semi-analytic expressions for the through-focus aberrated point-spread function in the low-aperture scalar diffraction case [29,30] and for the vector diffraction case [31]. Although the calculation of the image of a single point has been drastically accelerated by the availability of an accurate semi-analytic series expansion instead of a diffraction integral to be numerically evaluated, the application of this new tool for convolution calculations to obtain complete images is still too slow in practice.

In this paper, we present details on a recently proposed novel approach to point-spread-function-based image calculation [32]. The basic tool used in this new imaging approach is the ENZ theory for through-focus point-spread function calculation of general imaging systems with aberrations and spatial lens transmission variations. However, instead of using a point-source in the object plane, providing us with a quasi-uniform wave at the entrance pupil and slowly varying phase in the exit pupil due to imaging aberrations, we start with an extended object that is many equivalent point-spread functions large in the object plane. The wave field produced in the entrance pupil by this coherently illuminated extended object field is then expanded in Zernike-Nijboer circle polynomials, including its amplitude and phase variations following from the diffraction of the coherent incident wave (in general a plane wave) by the extended object. The wave field in the exit pupil is the product of the wave field in the entrance pupil multiplied by the amplitude and phase (aberration) of the lens transmission function. This wave field with its Zernike polynomial expansion then immediately yields the through-focus aberrated image.

In Section 2 of this paper, we exploit the general tool for through-focus point-spread function calculation provided by the ENZ-formalism to compute the image of more general objects. To this goal we construct the Zernike expansion of the wave field in the entrance pupil of the system. We give the detailed derivation of imaging from object space to image space with different refractive indexes to include immersion imaging and show how the Zernike expansion of the entrance pupil field is incorporated in the formalism. The immersion imaging mode is encountered not only in microscopy but also in high-resolution optical lithography systems with numerical aperture values higher than unity [33]. Prior derivations were not always exact in calculating the absolute field strengths in image space given the object field components. In Section 3, we present our new method using an extended object as input for our point-spread function based imaging method. The input for the imaging method is the optical near-field of the object given the illuminating wave. Rigorous electromagnetic diffraction is applied to obtain this near field, using a specially developed FDTD (Finite Difference Time Domain) method [34]. In Section 4, we analyze the accuracy and the typical computational complexity of our method. Finally, in Section 5, we present our conclusions.

2.   ENZ-theory for advanced imaging

The extended Nijboer-Zernike (ENZ) theory of diffraction has been first described in references [29] and [30]. In these papers, a semi-analytic solution of the Debye diffraction integral for the imaging of a point source by a general optical system was introduced. The development of ENZ-theory was intended to provide a method to characterize optical systems by means of intensity measurements in the focal region. In order to achieve this, it is fundamental to have both an accurate and fast algorithm to compute the point-spread function for a general aberrated system. In terms of computations, this comes down to computing the point-spread function from the otherwise uniform exit-pupil, for an exit-pupil that is influenced by a general aberration. Although the non-uniformities, present in the pupil due to aberrations, are usually small and should be relatively small in order to make ENZ-characterization possible, the pupil imaging method itself is not limited by the size of the deformations and can be applied to general pupils. Recognizing the fact that in general imaging one also needs to image an exit-pupil distribution that can be very complex, both in amplitude and phase, we believe that this appealing feature of the ENZ-theory is well suited to be exploited for general imaging.
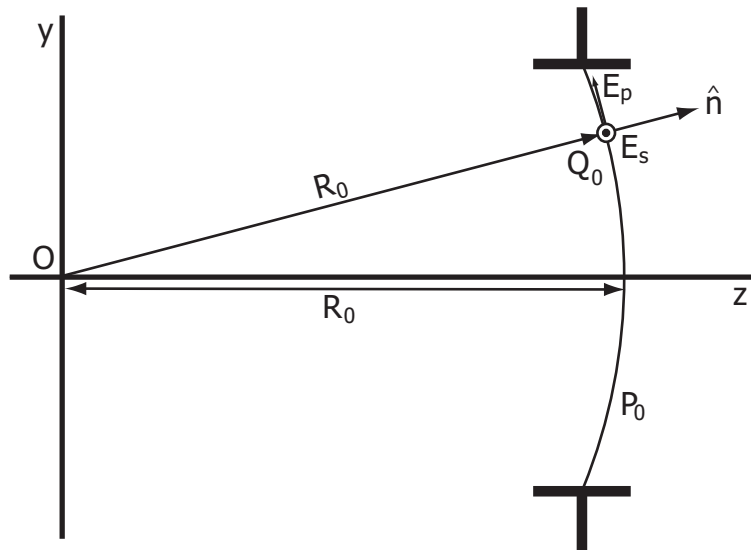

Fig. 1. Schematic representation of the geometry under consideration. O is the origin, R_0 is the radius of the entrance pupil sphere P_0, $\hat{n}$ is the normal to the pupil surface and Q_0 is a general point in the entrance pupil where we locally define the orthogonal linear p- and s-polarization states.

Here, we will introduce ENZ-based imaging by deriving an ENZ-based imaging method for advanced lithographic systems. For such systems, the distance, R_0, between the object

(in this case, a lithographic mask) and the entrance pupil of the optical system, is much larger than the transverse dimensions of the object being imaged. In this case, it is allowed to apply the Fraunhofer approximation which considers all light coming from the mask to be originating from a single point O. However, the radius of the entrance pupil is of the same order of magnitude as R_0, and therefore the entrance pupil can not be considered a flat surface, but is a spherical surface P_0 with its origin at O (see Fig. 1). We now take a general point $Q_0(\rho, \theta)$ on the entrance pupil sphere P_0, where ρ and θ are polar coordinates with ρ normalized with respect to the lateral half diameter of the entrance pupil (see Fig. 1). In a very good approximation, the electric field vector of the light arriving at a point $Q_0(\rho, \theta)$ will be perpendicular to the normal $\hat{n}$ of the spherical surface and it is thus possible to locally represent the field at Q_0 by a linear combination of two orthogonal polarization states perpendicular to $\hat{n}$. Here, we will adhere to s- (perpendicular) and p- (parallel) states of linear polarization as they can be conveniently mapped between the entrance and exit pupil of an aplanatic imaging system. In our case, this mapping is uniquely defined as lithographic systems very accurately satisfy Abbe's sine condition, within a tolerance better than 10^{-5}.
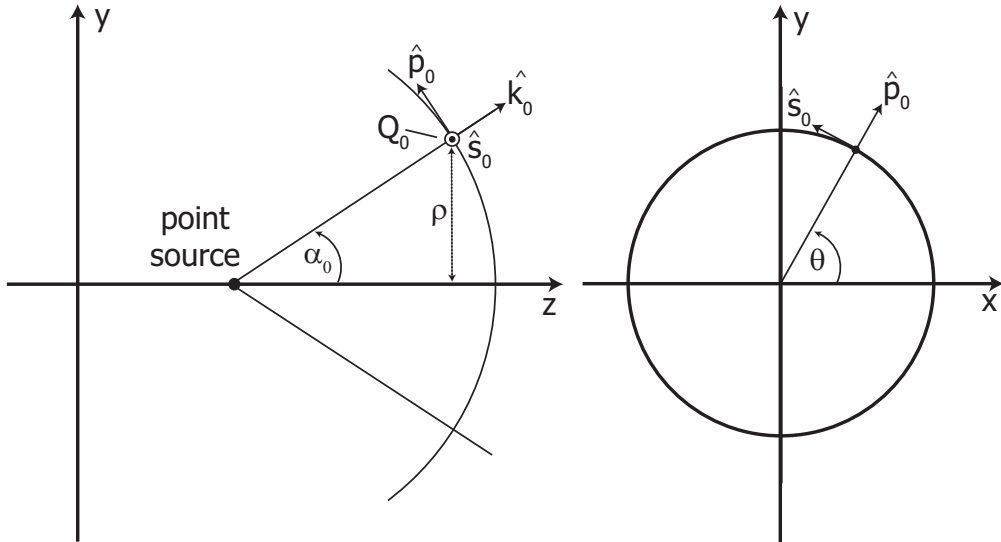

Fig. 2. Definition of the local basis for a general point Q_0 on the entrance pupil sphere with an axial cross-section (left-hand graph) and a cross-section perpendicular to the z-axis (right-hand graph).

We now introduce a new coordinate base in $Q_0(\rho, \theta)$ according to

$$\hat{\mathbf{k}}_0 = \sin \alpha_0 \cos \theta \, \hat{\mathbf{x}} + \sin \alpha_0 \sin \theta \, \hat{\mathbf{y}} + \cos \alpha_0 \, \hat{\mathbf{z}}, \tag{1}$$

$$\hat{\mathbf{p}}_0 = \cos \alpha_0 \cos \theta \, \hat{\mathbf{x}} + \cos \alpha_0 \sin \theta \, \hat{\mathbf{y}} - \sin \alpha_0 \, \hat{\mathbf{z}}, \tag{2}$$

$$\hat{\mathbf{s}}_0 = -\sin\theta\,\hat{\mathbf{x}} + \cos\theta\,\hat{\mathbf{y}}, \tag{3}$$

where $\alpha_0 = \arcsin(\frac{\rho}{R_0})$ and the right-handed unit vector bases $(\hat{\mathbf{p}}_0, \hat{\mathbf{s}}_0, \hat{\mathbf{k}}_0)$ and $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ are depicted in Fig. 2. We suppose that we have available the Cartesian electric field components at a general point Q_0, for example, by means of rigorous electromagnetic computation. From the components E_x and E_y, we can find E_z by applying the orthogonality of field components with respect to the propagation unit vector $\hat{\mathbf{k}}_0$ yielding

$$E_{0,z} = -(E_{0,x}\cos\theta + E_{0,y}\sin\theta)\tan\alpha_0. \tag{4}$$

The field components on the new basis are then given by

$$E_{0,p}(\rho,\theta) = \mathbf{E}_0(\rho,\theta)\cdot\hat{\mathbf{p}}_0 = \frac{E_{0,x}\cos\theta + E_{0,y}\sin\theta}{\cos\alpha_0}, \tag{5}$$

$$E_{0,s}(\rho,\theta) = \mathbf{E}_0(\rho,\theta)\cdot\hat{\mathbf{s}}_0 = -E_{0,x}\sin\theta + E_{0,y}\cos\theta. \tag{6}$$

Note that the component in the direction of $\hat{\mathbf{k}}$, $E_{0,k}$, is by definition equal to zero because of the Fraunhofer approximation. Next, we have to relate the field at a general point Q_1 in the exit pupil to the incident field in the corresponding point Q_0 on the entrance pupil sphere P_0. The field in the entrance pupil equals the far-field produced by diffraction of the incident field at the mask structure. For the transition of the far-field from the entrance pupil to the exit pupil, we need to consider the energy transport between the pupil surfaces. The basic relationship between the energy flow through the entrance and the exit pupil follows from paraxial optics. It provides us with the ratio between elementary areas on the pupil surfaces by means of the square of the paraxial magnification. The exact mapping of an elementary surface area from entrance to exit pupil outside the paraxial domain needs an extra condition; in our case, for large-field imaging systems, the Abbe sine condition has to be satisfied. Following the arguments in [6] and allowing for different refractive indices n_0 and n_1 in object and image space, we have the following relationship between the energy flow through corresponding areas in entrance and exit pupil (in the absence of absorption)

$$\epsilon v|E|^2 dS = \text{constant}\,,$$

or,

$$n_0|E_0|^2 dS_0 = n_1|E_1|^2 dS_1\,, \tag{7}$$

where $\epsilon = n^2$ is the relative electric permittivity of the medium, $v = c/n$ is the propagation velocity and dS_0 and dS_1 are the corresponding flow cross-sections on the pupil surfaces. Denoting the distances of entrance and exit pupil from the object and image plane by R_0 and R_1, respectively, and using the expression for the elementary solid angles in object and image space, we find

$$n_0 R_0^2 \frac{dk_{x,0} dk_{y,0}}{k_0 k_{z,0}}\,|E_0|^2 = n_1 R_1^2 \frac{dk_{x,1} dk_{y,1}}{k_1 k_{z,1}}\,|E_1|^2, \tag{8}$$

with $k_{x,i}$, $k_{y,i}$ and $k_{z,i}$ being the wavevector components in the object and image space.

Using $k_0/n_0 = k_1/n_1$, we write the electric field at the image side as

$$|E_1| = \frac{R_0}{R_1} \sqrt{\frac{k_{z,1}}{k_{z,0}}} \sqrt{\frac{dk_{x,0}dk_{y,0}}{dk_{x,1}dk_{y,1}}} \, |E_0|. \tag{9}$$

Denoting the paraxial magnification by M and writing the Abbe sine-condition according to

$$k_{x,0} = Mk_{x,1}, \qquad k_{y,0} = Mk_{y,1}, \tag{10}$$

we obtain

$$
\begin{aligned}
|E_1| &= \left(\frac{MR_0}{R_1}\right) \sqrt{\frac{n_1}{n_0}} \sqrt{\frac{k_{z,1}}{k_1} \frac{k_0}{k_{z,0}}} \, |E_0| \\
&= \left(\frac{MR_0}{R_1}\right) \sqrt{\frac{n_1}{n_0}} \frac{(1 - s_0^2\rho^2)^{1/4}}{(1 - \frac{n_1^2}{n_0^2}M^2 s_0^2\rho^2)^{1/4}} \, |E_0|.
\end{aligned} \tag{11}
$$

Finally, we use the paraxial relation between pupil object and image distances $R_1/R_0 = (n_1/n_0)MM_p$ with M_p the pupil magnification and use Newton's paraxial imaging equation in putting $|M_p| = |R_p/f_1|$ with R_p being the distance from the image side focal point to the exit pupil and f_1 the image side focal distance. The field in the exit pupil is then given by

$$|E_1| = \left|\frac{f_1}{R_p}\right| \sqrt{\frac{n_0}{n_1}} \frac{(1 - s_0^2\rho^2)^{1/4}}{(1 - \frac{n_1^2}{n_0^2}M^2 s_0^2\rho^2)^{1/4}} \, |E_0| = \left|\frac{f_1}{R_p}\right| \sqrt{\frac{n_0}{n_1}} T_R(\rho)|E_0|, \tag{12}$$

where the transmission factor

$$T_R(\rho) = \frac{(1 - s_0^2\rho^2)^{1/4}}{(1 - \frac{n_1^2}{n_0^2}M^2 s_0^2\rho^2)^{1/4}} \tag{13}$$

can be ascribed to the radiometric effect for a system satisfying the Abbe sine-condition in our general case with unequal refractive indices in object and image space. In the more common case with $n_0 = n_1$, we can put $|f_1| = |f_0| = f_L$, the focal distance of the imaging system.

As already mentioned above, we limit ourselves to systems satisfying Abbe's sine condition. In this case a point Q_0 at position (ρ, θ) is mapped directly to a point $Q_1(\rho, \theta)$ on the exit pupil sphere with the property that the normalized radial coordinate ρ and the azimuthal coordinate remain unchanged. In addition to the radiometric effect described by the transmission factor T_R, the light travelling through the imaging system will also experience some deformations introduced by imperfections of the imaging system. For lithographic systems these aberrations are generally small but, nevertheless, we include them in the formalism.

We describe both the amplitude and phase aberrations of the imaging system by its complex transmission function T_I. In fact, in the most general case we should consider T_I^p and T_I^s, being the transmission function for p- and s-polarization components respectively, but here we will adhere to imaging systems that are free of birefringence, allowing a single transmission function for all field components. The transmission function of the imaging system is then given as

$$T_I = A(\rho, \theta) \exp\left[i\Phi(\rho, \theta)\right] \quad , \tag{14}$$

where $A(\rho, \theta)$ is the amplitude transmission function and $\Phi(\rho, \theta)$ is the phase aberration. In contrast to basic ENZ-theory, where the field $\mathbf{E_0}$ in the entrance pupil is uniform, we are dealing here with a general non-uniform field $\mathbf{E_0}$. We can now write the $E_{1,s}$ and $E_{1,p}$ components as

$$
\begin{aligned}
E_{1,s}(\rho, \theta) &= \frac{f_1 \, T_I \, T_R}{R_p} \sqrt{\frac{n_0}{n_1}} \, E_{0,s}(\rho, \theta) \\
&= \frac{f_1 \, T_I \, T_R}{R_p} \sqrt{\frac{n_0}{n_1}} \left(-E_{0,x}(\rho, \theta)\sin\theta + E_{0,y}(\rho, \theta)\cos\theta\right) \quad ,
\end{aligned} \tag{15}
$$

$$
\begin{aligned}
E_{1,p}(\rho, \theta) &= \frac{f_1 \, T_I \, T_R}{R_p} \sqrt{\frac{n_0}{n_1}} \, E_{0,p}(\rho, \theta) \\
&= \frac{f_1 \, T_I \, T_R}{R_p} \sqrt{\frac{n_0}{n_1}} \frac{E_{0,x}(\rho, \theta)\cos\theta + E_{0,y}(\rho, \theta)\sin\theta}{(1 - n_1^2 M^2 s_0^2 \rho^2 / n_0^2)^{\frac{1}{2}}} \quad .
\end{aligned} \tag{16}
$$

The x-, y- and z-components are given by

$$E_{1,x}(\rho, \theta) = E_{1,p}(\rho, \theta)\cos(\alpha_1)\cos\theta - E_{1,s}(\rho, \theta)\sin\theta \quad , \tag{17}$$

$$E_{1,y}(\rho, \theta) = E_{1,p}(\rho, \theta)\cos(\alpha_1)\sin\theta + E_{1,s}(\rho, \theta)\cos\theta \quad , \tag{18}$$

$$E_{1,z}(\rho, \theta) = E_{1,p}(\rho, \theta)\sin(\alpha_1) \quad , \tag{19}$$

or, by using $\cos(\alpha_1) = \sqrt{1 - s_0^2 \rho^2}$ and inserting Eqs.(15)-(16), we get

$$
\begin{aligned}
E_{1,x}(\rho, \theta) = \frac{f_1 \, T_I(\rho, \theta) \, T_R(\rho)}{R_p} \sqrt{\frac{n_0}{n_1}} & \left[\frac{(1 - s_0^2 \rho^2)^{\frac{1}{2}}}{(1 - n_1^2 M^2 s_0^2 \rho^2 / n_0^2)^{\frac{1}{2}}} \times \right. \\
& \left\{ E_{0,x}(\rho, \theta)\cos^2\theta + E_{0,y}(\rho, \theta)\cos\theta\sin\theta \right\} + \\
& \left. \left\{ E_{0,x}(\rho, \theta)\sin^2\theta - E_{0,y}(\rho, \theta)\cos\theta\sin\theta \right\} \right],
\end{aligned} \tag{20}
$$

$$
\begin{aligned}
E_{1,y}(\rho, \theta) = \frac{f_1 \, T_I(\rho, \theta) \, T_R(\rho)}{R_p} \sqrt{\frac{n_0}{n_1}} & \left[\frac{(1 - s_0^2 \rho^2)^{\frac{1}{2}}}{(1 - n_1^2 M^2 s_0^2 \rho^2 / n_0^2)^{\frac{1}{2}}} \times \right. \\
& \left\{ E_{0,x}(\rho, \theta)\cos\theta\sin\theta + E_{0,y}(\rho, \theta)\sin^2\theta \right\} + \\
& \left. \left\{ -E_{0,x}(\rho, \theta)\cos\theta\sin\theta + E_{0,y}(\rho, \theta)\cos^2\theta \right\} \right],
\end{aligned} \tag{21}
$$

8

$$E_{1,z}(\rho,\theta) = \frac{f_1 \, T_I(\rho,\theta) \, T_R(\rho) \, s_0 \, \rho}{R_p(1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}}} \, \sqrt{\frac{n_0}{n_1}} \left\{ E_{0,x}(\rho,\theta) \cos\theta + E_{0,y}(\rho,\theta) \sin\theta \right\}.$$

(22)

We now choose to represent the field components $E_{0,x}$ and $E_{0,y}$, including possible deformations introduced by the optical system and represented by $T_I(\rho,\theta)$, by means of a Zernike expansion

$$E_{0,x}(\rho,\theta) \, T_I(\rho,\theta) = \sum_{n,m} \beta_{n,x}^m R_n^{|m|}(\rho) \exp(im\theta),$$

(23)

$$E_{0,y}(\rho,\theta) \, T_I(\rho,\theta) = \sum_{n,m} \beta_{n,y}^m R_n^{|m|}(\rho) \exp(im\theta).$$

(24)

Using this expansion and after a somewhat lengthy manipulation, the field components in the exit pupil can finally be written as

$$E_{1,x}(\rho,\theta) = \frac{f_1 \, T_R(\rho)}{2R_p \, (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}}} \sqrt{\frac{n_0}{n_1}} \sum_{n,m} R_n^{|m|}(\rho) \exp(im\theta) \times$$

$$\left\{ \beta_{n,x}^m \left[\left\{ (1 - s_0^2\rho^2)^{\frac{1}{2}} + (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} \right\} \right. \right.$$

$$\left. - \left\{ \frac{(1 - n_1^2 M^2/n_0^2)s_0^2\rho^2 \cos 2\theta}{\left[(1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} + (1 - s_0^2\rho^2)^{\frac{1}{2}} \right]} \right\} \right]$$

$$\left. - \beta_{n,y}^m \left[\frac{(1 - n_1^2 M^2/n_0^2)s_0^2\rho^2 \sin 2\theta}{\left[(1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} + (1 - s_0^2\rho^2)^{\frac{1}{2}} \right]} \right] \right\},$$

(25)

$$E_{1,y}(\rho,\theta) = \frac{f_1 \, T_R(\rho)}{2R_p \, (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}}} \sqrt{\frac{n_0}{n_1}} \sum_{n,m} R_n^{|m|}(\rho) \exp(im\theta) \times$$

$$\left\{ \beta_{n,y}^m \left[\left\{ (1 - s_0^2\rho^2)^{\frac{1}{2}} + (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} \right\} \right. \right.$$

$$\left. + \left\{ \frac{(1 - n_1^2 M^2/n_0^2)s_0^2\rho^2 \cos 2\theta}{\left[(1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} + (1 - s_0^2\rho^2)^{\frac{1}{2}} \right]} \right\} \right]$$

$$\left. - \beta_{n,x}^m \left[\frac{(1 - n_1^2 M^2/n_0^2)s_0^2\rho^2 \sin 2\theta}{\left[(1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}} + (1 - s_0^2\rho^2)^{\frac{1}{2}} \right]} \right] \right\},$$

(26)

$$E_{1,z}(\rho,\theta) = \frac{f_1 \, T_R(\rho) \, s_0}{R_p \, (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}}} \sqrt{\frac{n_0}{n_1}} \times$$

$$\sum_{n,m} \rho \, R_n^{|m|}(\rho) \exp(im\theta) \left\{ \beta_{n,x}^m \cos\theta + \beta_{n,y}^m \sin\theta \right\}.$$

(27)

The final step to obtain the image produced by the optical system involves computing the Debye diffraction integral with the aid of the exit pupil field distribution defined by Eqs. (25)-(27). As we are interested in systems with a large numerical aperture, we have to deal with the full vectorial version of the Debye diffraction integral as formulated in the well-known papers by Ignatowsky [21], Wolf [23] and Richards and Wolf [35]. The expressions they introduced were intended to determine the image of a point-source, but, in a very good approximation for sufficiently large exit pupil diameter, they can be equally well applied to more general fields in the exit pupil. The general expression for an arbitrary field distribution on the exit pupil sphere in image space can be found in [36],

$$
\begin{aligned}
\mathbf{E_2}(r, \phi, f) & = \frac{-i s_0^2}{\lambda} \exp\left(\frac{-if}{u_0}\right) \iint_C \frac{\mathbf{E_i}(\rho, \theta + \pi)}{(1 - s_0^2 \rho^2)^{1/2}} \times \\
& \exp\left\{\frac{if}{u_0}\left[1 - (1 - s_0^2\rho^2)^{1/2}\right]\right\} \exp\{i2\pi r\rho\cos(\theta - \phi)\}\rho d\rho d\theta,
\end{aligned}
\tag{28}
$$

with C being the scaled integration area on the exit pupil which in our case is equal to the unit circle. For the vector quantity $\mathbf{E_i}(\rho, \theta + \pi)$ in the integrand of the Debye-integral, one should take the field-invariant $r\mathbf{E_1}$ along a propagation direction defined by (ρ, θ), with r the distance along the propagation direction towards or away from the image point. This quantity has also been denoted by 'ray strength'. Once we have obtained the field components $\mathbf{E_1}$ on a sphere with its midpoint at the image center and with a radius R_1, the quantity $\mathbf{E_i}$ is simply given by $R_1\mathbf{E_1}$. Multiplication of R_1 with the geometrical amplitude factor f_1/R_p in Eqs.(25)-(27) allows us to write the complete scaling factor $f_1 s_0^2 R_1/\lambda R_p$ as $s_0^2 f_1/(\lambda[1 - Mf_1/R_1])$ using the paraxial property $R_p = R_1 - Mf_1$. Special cases arise when one of the pupils or both are located at infinity. For an exit pupil at infinity, the geometrical scaling factor reduces to $f_1 s_0^2/\lambda$. An entrance pupil at infinity poses a problem because the field components $E_{0,s}$ and $E_{0,p}$ tend to zero and the evaluation of the Zernike coefficients of Eqs.(23)-(24) would be indefinite. In this case, the field components can be measured or calculated, far enough from the object itself, on a sphere with a predefined radius R_0 that will replace the actually infinite value of the radius. With the finite energy flow from the infinitely distant entrance pupil, the field components on the chosen sphere with radius R_0 can be obtained. Using the corresponding R_1-value for the exit pupil position, the quantity $r\mathbf{E_1}$, a ray invariant, is then uniquely obtained. The coordinates (r, ϕ, f) used in Eq.(28) are normalized cylindrical co-ordinates in the image space with the origin located at the geometrical focus, see Fig.2, with f being the normalized axial coordinate and (r, ϕ) the lateral polar coordinates. The normalization has been carried out with respect to the diffraction unit $\lambda/NA = \lambda/(n_1 s_0)$ in the radial r-direction and $f = -2\pi n_1 u_0 z/\lambda$ in the axial direction with $u_0 = 1 - \sqrt{1 - s_0^2}$.

To evaluate Eq.(28), we insert the exit pupil field given by Eqs.(25)-(27) and the radiometric factor T_R into Eq.(28). Using the expression $\int_0^{2\pi} \exp(im\theta) \exp(i2\pi r\rho\cos(\theta - \phi))d\theta = $

10

$2\pi i^m J_m(2\pi r\rho) \exp(im\phi)$ with $J_m(x)$ the m-th order Bessel function of the first kind, we obtain for the vector field in the focal region using column vector notation,

$$\mathbf{E_2}(r,\phi,f) = \frac{-i\pi f_1 s_0^2}{\lambda\left[1 - M f_1/R_1\right]}\sqrt{\frac{n_0}{n_1}}\exp\left(\frac{-if}{u_0}\right)\sum_{n,m}(-i)^m \exp[im\phi]\times$$

$$\left[\beta_{n,x}^m\begin{pmatrix} V_{n,0}^m + s_0^2\left(\frac{n_0^2-n_1^2 M^2}{2n_0^2}\right)\left\{V_{n,+2}^m\exp[+2i\phi] + V_{n,-2}^m\exp[-2i\phi]\right\} \\ -is_0^2\left(\frac{n_0^2-n_1^2 M^2}{2n_0^2}\right)\left\{V_{n,+2}^m\exp[+2i\phi] - V_{n,-2}^m\exp[-2i\phi]\right\} \\ -is_0\left\{V_{n,+1}^m\exp[+i\phi] - V_{n,-1}^m\exp[-i\phi]\right\}\end{pmatrix}\right.$$

$$\left.+ \beta_{n,y}^m\begin{pmatrix} -is_0^2\left(\frac{n_0^2-n_1^2 M^2}{2n_0^2}\right)\left\{V_{n,+2}^m\exp[+2i\phi] - V_{n,-2}^m\exp[-2i\phi]\right\} \\ V_{n,0}^m - s_0^2\left(\frac{n_0^2-n_1^2 M^2}{2n_0^2}\right)\left\{V_{n,+2}^m\exp[+2i\phi] + V_{n,-2}^m\exp[-2i\phi]\right\} \\ -s_0\left\{V_{n,+1}^m\exp[+i\phi] + V_{n,-1}^m\exp[-i\phi]\right\}\end{pmatrix}\right],$$

(29)

where, for integer $j = -2,\cdots,+2$, we have

$$V_{n,j}^m(r,f) = \int_0^1 \rho^{|j|}\frac{\left\{(1-n_1^2 M^2 s_0^2\rho^2/n_0^2)^{\frac{1}{2}} + (1-s_0^2\rho^2)^{\frac{1}{2}}\right\}^{-|j|+1}}{(1-s_0^2\rho^2)^{\frac{1}{4}}(1-n_1^2 M^2 s_0^2\rho^2/n_0^2)^{\frac{3}{4}}}\times$$

$$\exp\left[\frac{if}{u_0}\left(1-\sqrt{1-s_0^2\rho^2}\right)\right]R_n^{|m|}(\rho)J_{m+j}(2\pi r\rho)\rho d\rho.$$

(30)

In contrast with earlier publications, for instance [37-39], Eq. 29 has the factor $(-i)^m$ instead of i^m. The minus sign is due to the fact that we have the same reference direction for the polar angles θ and ϕ in object, pupil and image space. In the earlier publications, there was a rotation of π between the polar coordinate systems (ρ,θ) and (r,ϕ). It is also important to note that former results pertained to imaging from infinity $(M = 0)$ and to equal indices in object and image space, yielding a geometrical prefactor $-i\pi s_0^2 R/\lambda$ with R denoting the focal distance of the imaging system. The results presented here yield the same limiting value for this special case. Also, in accordance with former results from ENZ-analysis (see [36], appendix D, where a point-source at infinity was considered), one can devise a series expansion to quickly obtain accurate values of the integral above, that applies to imaging at finite distances. The functions that are used in the expansion and the values of the new expansion coefficients are given in Appendix A.

In this section, we have identified the implications of considering imaging of an extended object at a finite distance from the entrance pupil, within the frame-work of the ENZ-formalism. We have shown how to adjust the standard ENZ-formalism to accommodate for this new range of applications, and we have introduced a new expression for the field in the focal region of a general imaging system. In this new formulation, the field in the focal region is directly related to the field in the entrance pupil. The only approximation

used in our considerations is the Fraunhofer approximation that requires the object to be small compared to the distance between object and entrance pupil. For the use of these new expressions in mask imaging, this condition is generally satisfied.

3. Computation scheme

In this section, we translate the imaging method devised in Section 2 into a computation scheme that simulates mask imaging by a lithographic system. When considering such a full lithographic system simulation, a choice should be made what model to use in order to describe the finite extent of the illumination source correctly. Here, we will adhere to the Abbe approach instead of the more commonly used Hopkins approximation (see [40], Sections 8.2 and 8.3). Although the Hopkins approach is generally faster, it has its limitations when relatively thick mask objects are considered. On the other hand, the Abbe approach imposes no limitations on the structure of the mask and has therefore our preference.
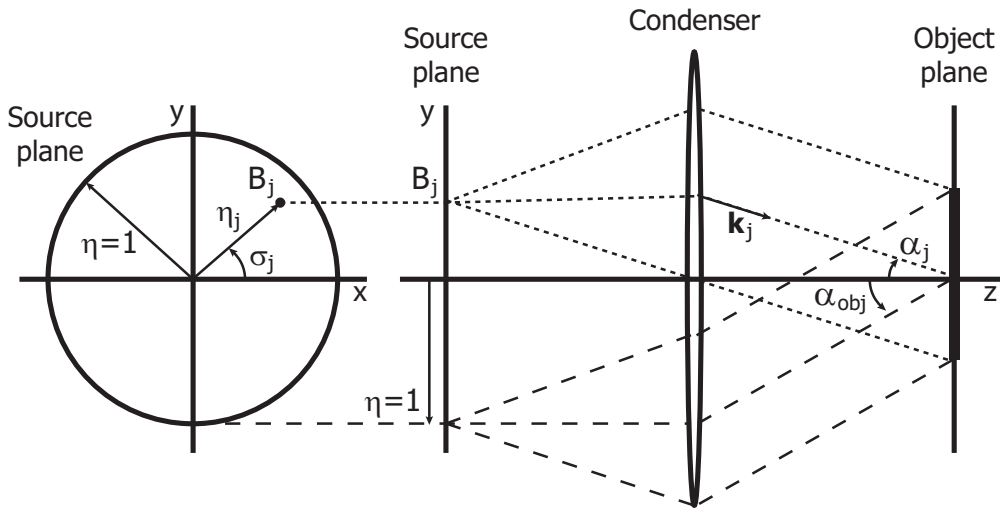

Fig. 3. Schematic representation of the Köhler illumination.

In the Abbe approach, the finite-size illumination source approximated by a sampled equivalent source with constant values for its strength and far-field radiation pattern over each sampling subarea. In advanced lithography, the illumination system, based on Köhler illumination, has been extended with smoothing elements, for instance, a fly-eyes lens array or an assembly of light-guiding quartz rods. These measures are meant to improve the uniformity of the effective source of the illumination system, both regarding its near-field and far-field properties. Suppose that we can measure the radiance function $B(x, y; \zeta_x, \zeta_y)$ of the effective source, so that we are then left with calculating the corresponding electromagnetic field

strengths in the object plane where the mask is located. In the following, we limit ourselves to a centered object with very limited extent. With the Köhler illumination arrangement, the axial object collects the forward directed radiation of each source point in the effective source. In this case, we can use the value of the radiance function in this direction, given by $B(x, y; 0, 0)$. The power flow from an elementary source element dS_S towards the object element with size dS_O is given by

$$
\begin{aligned}
d^2P &= B(x, y; 0, 0)dS_S d\Omega_O \\
&= B(x, y; 0, 0)dS_S dS_O \cos(\alpha)/f_C^2 \ ,
\end{aligned}
\tag{31}
$$

where we have used that the solid angle $d\Omega_O$ subtended by the object is given by its projected surface divided by the square of the focal length f_C of the condensor lens (α is the projection angle). In the case of the sampled source, we suppose a constant behavior of the source over each sampling area and we are then allowed to integrate the elementary power flow using the value of the radiance function in each sampling point. For a sampled source area with index j we find

$$
P_j = B(x_j, y_j; 0, 0)S_j S_O cos(\alpha_j/f_C^2) \ ,
\tag{32}
$$

where S_j is the area of the j-th source element. The power contained in the parallel beam that is issued by the sampling area with index j_S is given by

$$
P_j = \epsilon_0 c n_0 |\mathbf{E}_O|^2 S_O cos(\alpha_j) \ .
\tag{33}
$$

From Eqs.(32) and (33) we obtain for the electric field strength

$$
|\mathbf{E}_O| = \sqrt{\frac{B(x_j, y_j; 0, 0)S_j}{\epsilon_0 c n_0 f_C^2}} \ .
\tag{34}
$$

In the expression above, we have taken the radiance function in a medium with refractive index n_0, the value in the object space. If the radiance function of the source is only available in vacuum or air, one can apply the relation $B = n_0^2 B_V$ which follows from the conservation law of radiance (B_V is the measured radiance function in vacuum or air).

In what follows, the Cartesian co-ordinates (x_j, y_j) in the effective source will be replaced by reduced coordinates, with the normalized co-ordinates in the entrance pupil of the imaging system as the reference. To this goal, we start with the numerical aperture in the object space, given by $s_{obj} = \sin(\alpha_{obj}) = |n_1 M s_0/n_0|$ with $n_1 s_0$ the numerical aperture in the image space. The angular measure of the source is normalized with respect to s_{obj}, such that the lateral reduced coordinate η of the source is unity if the source completely fills the entrance pupil. The relationship between the Cartesian coordinates (x_j, y_j) of a general source point and its reduced coordinates is given by

$$
\begin{aligned}
x_j/f_C &= \eta_j \cos(\sigma_j)s_{obj} \ , \\
y_j/f_C &= \eta_j \sin(\sigma_j)s_{obj} \ ,
\end{aligned}
\tag{35}
$$

where we have used polar coordinates (η, σ) in the reduced domain because of the frequently occurring rotational symmetry in illumination systems. The state of polarization of the light issued by the light source has not been discussed so far. In the case of an unpolarized source, the source power is equally distributed over two orthogonally polarized states of polarization. In more specific cases, the power has to be distributed over the x-, y- and z-field components of the light according to the measured state of (partial) polarization of the source.

In the Abbe approach, for each source region, we also have at our disposal the average propagation direction that is needed to analyze the diffraction process at the mask and the propagation of the diffracted light through the imaging system. A general source point with reduced polar coordinates (η_j, σ_j) gives rise to a fictitious plane wave whose normalized propagation vector is given by

$$\hat{\mathbf{k}}_j = \left(-\eta_j s_{obj} \cos(\sigma_j)\,\hat{\mathbf{x}}, -\eta_j s_{obj} \sin(\sigma_j)\,\hat{\mathbf{y}}, \sqrt{1 - \eta_j^2 s_{obj}^2}\,\hat{\mathbf{z}} \right)\ . \tag{36}$$

We now have available, at the location of the mask, the electric field strength and propagation direction of the fictitious wave produced by each sampling area of the source. The intensity in a selected plane in the image space can be calculated, carrying out the full chain of diffraction at the mask, diffraction at the diaphragm or pupils of the imaging system and the wave propagation towards the image region. The intensities due to each sampled region in the incoherently radiating source are added to obtain the total intensity distribution in the image region.

The interaction of each plane wave with an object is simulated separately with a rigorous electromagnetic solver based on the Finite Difference Time-Domain (FDTD) method. The FDTD implementation used here is an in-house developed code at Delft University of Technology. It uses absorbing boundary conditions based on a convolutional Perfectly Matched Layer (PML) implementation [41, 42]. Next to absorbing PML boundaries, quasi-periodic boundary conditions can be applied using the so-called Sine-Cosine technique [43]. Media with negative permittivity and non-zero absorption are simulated via the Auxiliary Differential Equations (ADE) technique for dispersive media [44]. The code is parallelized for efficient execution on multi-processor machines.

The objects simulated for this paper are relatively small isolated structures. The established approaches and simulation tools to treat such structures almost always involve periodic boundary conditions, where the isolated structures are padded by structureless object space to approximate the isolated case. Since the ENZ formalism does not require periodic boundary conditions for the rigorous calculation, PML absorbing boundary conditions are employed on all sides of the computational domain. The PML regions are non-reflective and thus fully transparent to outgoing waves. As a result, it appears numerically as if the object extends homogeneously in all directions. The rectangular three-dimensional computational

domain (CD) is then defined by the smallest volume that encloses the structured region of the object.

In common FDTD implementations, letting material interfaces cross the PML boundary is problematic. Our FDTD implementation does allow the interfaces of a multi-layer structure to cross into the PML. Because Maxwell's equations are linear we can write the total field (TF), $\mathbf{E}_{tot}$, as the sum of two solutions to Maxwell's equations,

$$\mathbf{E}_{tot} = \mathbf{E}_{sca} + \mathbf{E}_{zero}, \tag{37}$$

where $\mathbf{E}_{zero}$ is a known solution such as the incident field in free space and $\mathbf{E}_{sca}$ is called the scattered field. As in other implementations, we calculate the total field in the CD and the scattered field in the surrounding PML region. 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 layered geometry, the incident field is not the correct solution to Maxwell's equations, because the incident field is defined for homogeneous free space. The algorithm produces, therefore, incorrect results. Instead of using the incident field for $\mathbf{E}_{zero}$, we use the analytical steady-state solution of a plane wave incident on a structureless object. This approach was shown to give correct results for objects that include a multi-layer configuration [45].

The output of an FDTD simulation consists of near-field values in the computational domain. To simulate the propagation of the field from the mask to the entrance pupil of the optical system a near-to-far-field transformation is applied. For this purpose, we use the Stratton-Chu formula [46], which relates the scattered fields $\mathbf{E}_{sca}$ and $\mathbf{H}_{sca}$ on the boundary, $\partial\Omega$, of a given domain to the field at any point in or outside the domain at $\mathbf{r}'$

$$\mathbf{E}_{sca}(\mathbf{r}') = - \iint_{\partial\Omega} (\mathbf{n} \times \mathbf{E}_{sca}(\mathbf{r}))\underline{\underline{G}}_H(\mathbf{r}, \mathbf{r}') - (\mathbf{n} \times \mathbf{H}_{sca}(\mathbf{r}))\underline{\underline{G}}_E(\mathbf{r}, \mathbf{r}') \, \mathrm{d}r^2. \tag{38}$$

Herein, $\underline{\underline{G}}(\mathbf{r}, \mathbf{r}')$, is the Green's tensor of the layered system (subscripts, E and H, indicate the electric and magnetic Green's tensor). The layered Green's tensor is needed because we allow material interfaces to cross the Stratton-Chu integration surface. For a general layered system, these terms are difficult to calculate analytically. Instead, we implement the Fourier transformed Stratton-Chu formula [47]:

$$\mathcal{F}[\mathbf{E}_{sca}](k_x, k_y, z') = - \iint_{\partial\Omega} (\mathbf{n} \times \mathbf{E}_{sca}(\mathbf{r}))\mathcal{F}[\underline{\underline{G}}_H](\mathbf{r}, \mathbf{k}_{sc})$$
$$-(\mathbf{n} \times \mathbf{H}_{sca}(\mathbf{r}))\mathcal{F}[\underline{\underline{G}}_E](\mathbf{r}, \mathbf{k}_{sc}) \, \mathrm{d}r^2, \tag{39}$$

which now uses the Fourier transformed Green's tensor, $\mathcal{F}[\underline{\underline{G}}](\mathbf{r}, \mathbf{k}_{sc})$, of a layered system. Here $\mathbf{k}_{sc}$ is the wave vector in the direction of the far-field observation point. $\mathcal{F}[\underline{\underline{G}}](\mathbf{r}, \mathbf{k}_{sc})$ is efficiently calculated in the same algorithm that gives the analytic multi-layer solution

for the FDTD simulation. By using the Fourier transformed equation for the Stratton-Chu formula, the Fraunhofer far-field at the location on the spherical entrance pupil determined by the wavenumbers $k_{x,sc}$, $k_{y,sc}$ and distance z' is almost directly obtained:

$$\mathbf{E_0}(k_{x,sc}, k_{y,sc}, z') = \mathcal{F}[\mathbf{E}_{sca}]\frac{k_{z,sc}}{|\mathbf{k}_{sc}|}. \tag{40}$$

This quantity is the field $\mathbf{E_0}(\rho, \theta)$ that is used in the ENZ-imaging algorithm and has been introduced in Eqs. (4-6). In order to do the numerical integration of Eq. (39), the rectangular boundaries of the FDTD computational domain are discretized to a uniform, orthogonal boundary grid. Because field values on the FDTD grid are not collocated but staggered, the FDTD field values are linearly interpolated.

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 can severely reduce the computational burden of the far-field calculation and the ENZ-imaging algorithm [45].

It must be noted that for small isolated objects, a method is needed that uses the field on all boundaries of the domain to obtain rigorous far-field results, such as the Stratton-Chu formula. In practice, where for example a large mask area is considered, a simple Fourier-based near-to-far-field transformation may be more efficient and equally accurate.

Once the field in the entrance pupil is available one can proceed with applying the ENZ-based expressions that were derived in Section 2. We perform a Zernike expansion of the computed field components $E_{0,x}$ and $E_{0,y}$, including possible transmission defects T_I, according to Eqs.(23-24). In our case, the coefficients $\beta_{n,x}^m$ and $\beta_{n,y}^m$ are obtained through a least-squares fitting operation. This could in principle also be done using inner-product evaluation of the function to be fitted and the Zernike orthogonal functions. Nevertheless, we have observed that a least-squares approach, for a set of well chosen data points, is far more efficient. More information on the sampling schemes applied in ENZ-imaging can be found in [45].

In the final step, to obtain the electric field contribution to the image from a single point in the illumination source, we evaluate Eq.(29) for the sets of Zernike coefficients $\beta_{n,x}^m$ and $\beta_{n,y}^m$ that result from the optimal least-squares fit discussed above. Note that Eq. (29) relies on the basic function $V_{n,j}^m$, which is defined in Eq. (30) and can conveniently be computed using the recipe given in Appendix A.

Following the procedure given above, the computation of the image contribution from a single source-point, should be repeated for every point comprising the light source. The total aerial image produced by the optical system then follows after incoherent summation of all contributions.

16

The full computational scheme described in this section can be summarized as follows

1. A single source point gives rise to illumination of the mask by a plane wave, where the angle of incidence of the plane wave is directly related to the spatial position of the considered source point.

2. The interaction between the mask and the incident plane wave is computed by means of a rigorous electromagnetic solver (in our case an in-house FDTD implementation).

3. An adapted version of the Stratton-Chu method is applied to obtain the field in the entrance pupil from the near-field at the mask.

4. The field in the exit pupil is obtained by applying the optical transfer function between the entrance and exit pupil of the optical system.

5. The field in the exit pupil is represented as a Zernike expansion after which the aerial image contribution, due to illumination by a single source point, follows directly from the ENZ-imaging algorithm.

6. Finally, step 1-5 should be repeated for each elementary point in the light source after which their intensity contribution should be summed incoherently to obtain the aerial-image of the mask produced by the lithographic system. In practice, a discrete sampling of the source will be carried out to keep the computational effort within reasonable bounds.

In this section, we have given a detailed description of the ENZ-based imaging scheme. In the remainder of this paper, we will discuss its characteristics and highlight some of its main features.

4. Evaluation of the ENZ-based method

This section discusses the anticipated accuracy and convergence of the imaging method proposed in this paper. We will limit ourselves to the imaging part of the algorithm, which includes the computations from entrance pupil to image region. We will not go into detail on the rigorous electromagnetic computations of the near field at the object, because, in principle, the proposed method allows for any rigorous solver to be used. For more details on the FDTD implementation developed in-house at TU Delft as used for the examples in this paper we refer to [48].

4.A. Representation of $\mathbf{E_0}T_I$ as a Zernike expansion

In the following, we assume that the arbitrary field $\mathbf{E_0}$ in the entrance pupil of the imaging system is known. In order to compute the image resulting from this field distribution we

should apply Eqs.(23)-(24) and the resulting sets of Zernike coefficients, $\beta_{n,x}^m$ and $\beta_{n,y}^m$, are subsequently inserted into Eq.(29). Note that T_I represents the complex transmission function of an imaging system that we assume to be free of birefringence and that $T_I \equiv 1$ for an aberration-free system.

As already mentioned in Section 3, we perform a least-squares fitting operation to obtain $\beta_{n,x}^m$ and $\beta_{n,y}^m$. The number of Zernike functions needed to accurately describe $\mathbf{E_0}T_I$ strongly depends on the object being imaged. This number is therefore determined iteratively. The maximum azimuthal order, m_{max}, and radial order, n_{max}, of the Zernike functions are increased until the desired fitting accuracy is reached. In Fig. 4, we show the residual RMS fitting error, for some particular objects (nine regularly arranged contact holes, an elbow structure and a hammerhead structure, respectively), versus m_{max} and n_{max}. Note that by definition of the Zernike functions $m_{max} \leq n_{max}$.
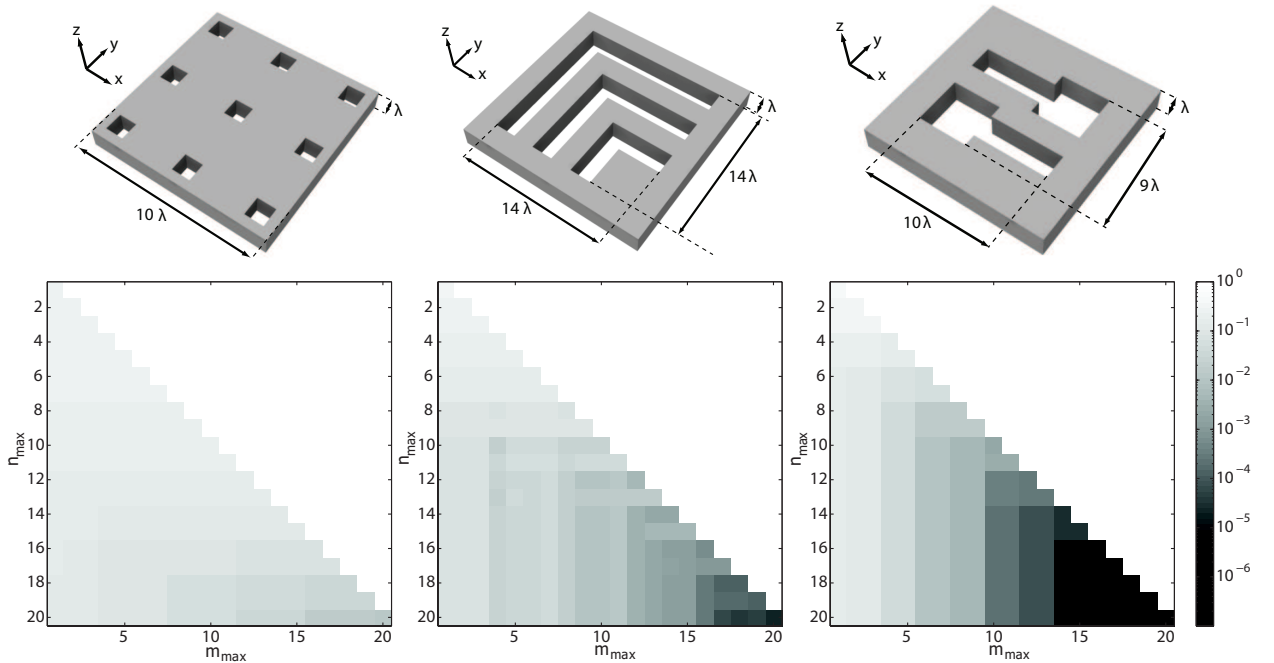

Fig. 4. For three different objects (top row), the Root Mean Square (RMS) error present in the Zernike approximation of the electric field components in the entrance pupil (bottom row) is shown. The system settings for the three objects are, from left to right: normal incidence TM polarized plane wave illumination with an object side numerical aperture of 0.525 and for the middle and right object normal incidence TE-polarization with an object side numerical aperture of 0.2375.

The simulation results in Fig. 4 clearly illustrate that the number of Zernike polynomials required to obtain a certain degree of fitting accuracy for the field in the entrance pupil, strongly varies between different objects. In general, one would expect the required number of Zernike polynomials to depend on the size and complexity of the objects. Nevertheless, a complex object doesn't necessary imply the need for a large number of Zernike functions for an accurate fit. A complex object can produce a relatively smooth field distribution in the pupil which would then require a limited number of Zernike polynomials to be fitted accurately. On the other hand, a fairly simple structure with a high degree of periodicity, can produce a sharply modulated pupil distribution that closely resembles the diffraction pattern that would arise in the case of a purely periodic object. In this case, the number of Zernike polynomials required to obtain an accurate fit is fairly large. Thus, in general, the number of Zernike functions needed grows according to the size and degree of periodicity of the object under consideration.

Once we have acquired a Zernike expansion of sufficient accuracy for the field in the entrance pupil we can proceed with applying the approach described in Section 3. In the next subsection we describe the convergence properties of the expressions that relate the entrance pupil field distribution to the field in the image region.

4.B. Computation of the basic integral $V_{n,j}^m(r,f)$

In Section 2, we have derived an expressions for $\mathbf{E_2}$, the field in the image region of the optical system (Eq.(29)). This expression depends on the basic integral, $V_{n,j}^m(r,f)$, which is defined in Eq.(30). Similar to standard ENZ-theory, a series expansion has been devised to evaluate $V_{n,j}^m(r,f)$ efficiently (Appendix A). It is important to recognize that $V_{n,j}^m(r,f)$ solely depends on the specifications of the imaging system and is independent of the object being imaged. Therefore, the $V_{n,j}^m(r,f)$-functions should only be computed once for a specific magnification and aperture setting of the imaging system, after which the $V_{n,j}^m(r,f)$-function values can be stored and subsequently reused in future image simulations.

What remains, is the calculation of $V_{n,j}^m(r,f)$, for $j = -2,-1,0,1,2$ and for a range of values of m and n, in a both accurate and efficient manner. When we study the expansion of $V_{n,j}^m(r,f)$ as found in Eq.(A23) we find that it contains an infinite sum over the parameter t. In practice, t has to be cut off at some finite value t_{max}. The relation between t_{max} and the Root Mean Square (RMS) error present in the computed values for $V_{n,j}^m(r,f)$ has been displayed in Fig.5. One can see that a modest amount of terms, say $t_{max} = 25$, already results in an accuracy of 10^{-4}. As $\mathbf{E_2}$ is linear in $V_{n,j}^m(r,f)$ the error in $\mathbf{E_2}$ introduced by $V_{n,j}^m(r,f)$ is also expected to be of the order 10^{-4}. Furthermore, we see a steady decrease in the RMS, for increasing t_{max}, down to 10^{-12}. From this, we can conclude that the $V_{n,j}^m(r,f)$-functions can always be computed down to the accuracy required by a particular application. Therefore,
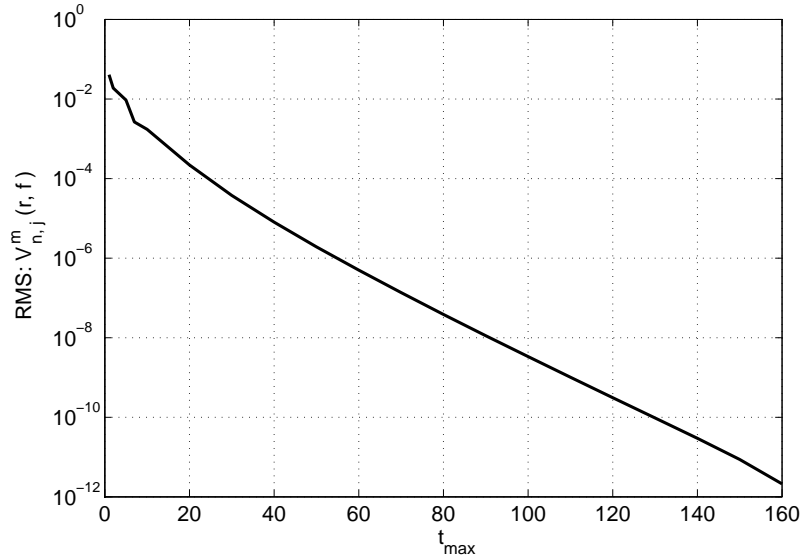
Fig. 5. The Root Mean Square (RMS) error present in the $V_{n,j}^m(r, f)$-functions when approximated with the series expansion in which the infinite sum is replaced by a summation over $t = 1, 2...t_{max}$ (see Appendix A).

the $V_{n,j}^m(r, f)$ functions do not pose a limit on the overall accuracy of the ENZ-method.

In this and the previous subsection, we have discussed the accuracy of the two computation modules that together comprise the ENZ-method. As discussed above, we can always compute the $V_{n,j}^m(r, f)$-functions down to the desired accuracy. A higher accuracy, of course, requires a larger computational burden, but in the case of the $V_{n,j}^m(r, f)$-functions this is of limited interest as the functions can be computed in advance and can be stored in a look-up table. Thus, the overall accuracy of the method is mainly determined by the quality of the Zernike expansions for the field in the entrance pupil. In the next subsection, we will evaluate the resulting accuracy for the field in the image region given the accuracy of the Zernike expansions in the entrance pupil.

4.C.   Field accuracy in the image region

As discussed in the previous subsections, we can assume that the overall accuracy of the ENZ-method is determined by the fitting accuracy for the fields in the entrance pupil. In Fig. 6, we have plotted the RMS error in the intensity in the image region versus the RMS error in the Zernike expansions of the fields in the entrance pupil. The figure clearly illustrates a similar behavior for the RMS errors of all three objects studied. Note that the lines representing the contact holes and elbow structure do not cover the full range of the RMS errors in entrance

pupil expansions, because for the maximum number of Zernike terms used in these examples ($m_{max} = n_{max} = 20$), the best obtained accuracies were 10^{-2} and 10^{-4} respectively.
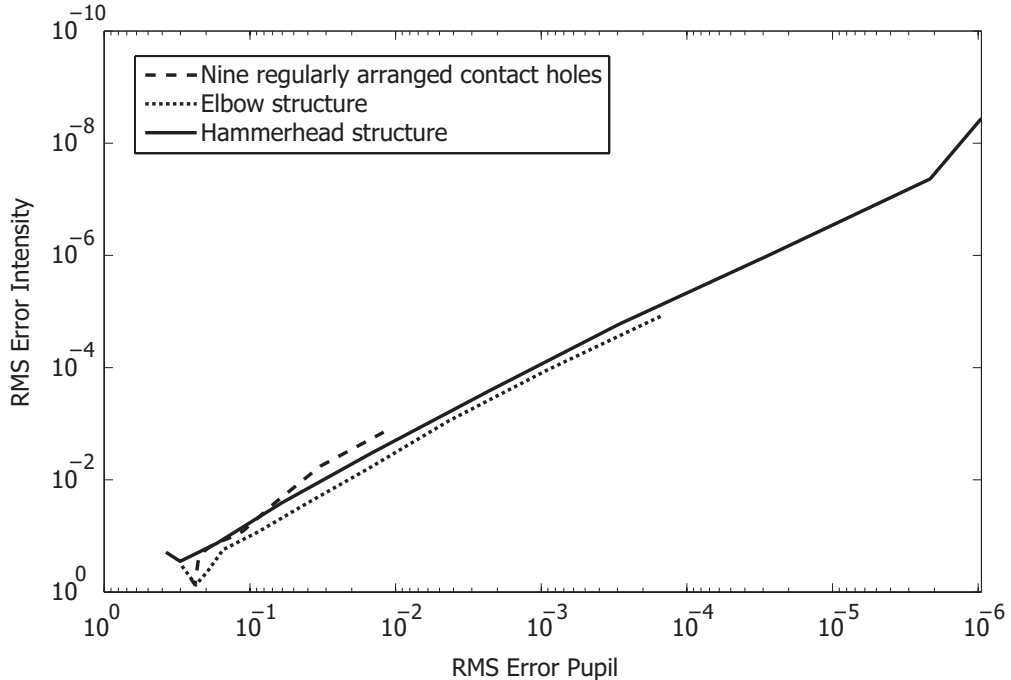

Fig. 6. For the same three objects as in Fig. 4, the RMS error for the intensity in the image region is shown as a function of the RMS error in the Zernike expansion of the pupil fields.

It is interesting to note that a pupil field accuracy of 10^{-3} already yields an intensity fidelity in the image region that is better than 10^{-4}, despite the nonlinearities that arise in the creation of the intensity distribution in the image.

Now that we have acquired a clear view on the expected accuracy of the ENZ-based imaging method, we should relate this to the computational burden that is associated with it. This will be main topic of the next subsection.

4.D. *Computational considerations*

The accuracy of a method should always be discussed in relation to the computational burden involved with it. As the ENZ-method is constructed from several largely independent modules, the computational complexity should be evaluated likewise. Here we will limit the discussion of the computational complexity to those modules already discussed in the previous subsections. We do not go into detail on the near-field computations and the propagation into the entrance pupil as in principle any rigorous electromagnetic solver could be used here.

As we apply a least-squares fitting operation to obtain the Zernike expansion coefficients for the entrance pupil fields, the computational complexity will largely be determined by this operation. In our case we use the least-squares fitting algorithm included in Matlab (Standard Matlab function *mldivide.m* [49]), which is based on QR factorization with column pivoting, to obtain the Zernike coefficients. The computational complexity of this algorithm is $O(N_Z^3)$, where N_Z is the total number of Zernike coefficients to be fitted.
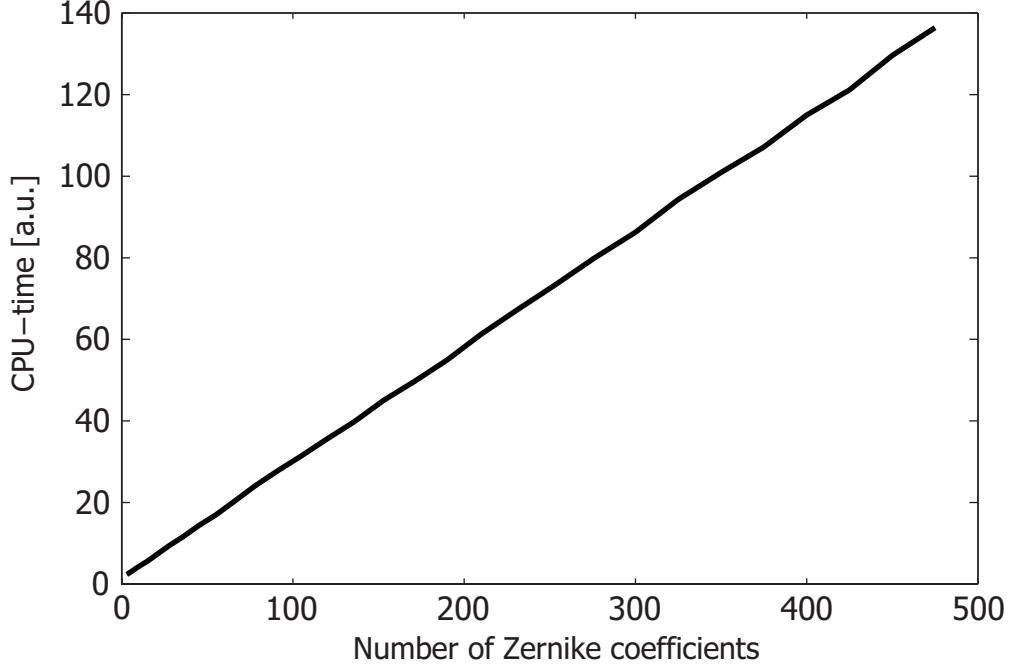

Fig. 7. The relation between the computation time and the total number of Zernike coefficients used in the expansion of the entrance pupil field is shown.

The computational complexity of the second module, which computes the field in the focal region given two sets of Zernike coefficients, $\beta_{n,x}^m$ and $\beta_{n,y}^m$, can be deduced from Eq.(29). Recall that the $V_{n,j}^m$ functions are independent of the object and can therefore be calculated in advance. As a result, all terms between the large parentheses on the second and third line of Eq. (29) are fixed for a given set of (r, ϕ, f). The computational task is thus reduced to computing

$$\mathbf{E_2}(r, \phi, f) = C_1(f) \sum_{n,m} \left[\beta_{n,x}^m \mathbf{C}_2(m, n, r, \phi, f) + \beta_{n,y}^m \mathbf{C}_3(m, n, r, \phi, f) \right], \quad (41)$$

where C_1, $\mathbf{C}_2$ and $\mathbf{C}_3$ all represent data stored in a look-up table. Consequently, the computational complexity is proportional to

$$N_f \times N_r \times N_\phi \times 2N_Z \quad , \quad (42)$$

22

where N_f, N_r and N_ϕ are the number of sampling points in the f, r and ϕ direction of our cylindrical coordinate system in the focal volume. The expression in Eq. (42) thus predicts a linear relation between the CPU-time and N_Z, the number of Zernike coefficients used in the computation. This behavior is clearly visible in Fig. 7, where we have plotted the CPU-time versus the number of Zernike coefficients.

On comparing the computational complexities determined above, we can conclude that both modules can be dominant. Whenever $N_f \times N_r \times N_\phi \ll N_Z^2$, the least-squares fitting module will dominate the required computational workload, while in case that $N_f \times N_r \times N_\phi \gg N_Z^2$, the computational contribution of the second module will be more important.

In this subsection we have discussed the computational complexity of the novel simulation modules introduced by the ENZ-method. It was shown that both the least-squares square fitting operation as the field construction from Zernike coefficients can be dominant in the total required computational burden for these two modules. Nevertheless, the contribution to the computational workload of the complete ENZ-method will generally remain limited. The computational burden for the complete ENZ-method is dominated by the Abbe treatment of the illumination source in combination with the rigorous near field computations.

5. Conclusions and discussion

In this paper, we have introduced a new imaging method based on the ENZ-theory. Although the standard ENZ-theory is meant to provide the through-focus point-spread function of a general system, we have shown that its range of application can be extended to include imaging of general objects.

The main result presented in this paper is a semi-analytic expression relating a general field in the entrance pupil to the resulting field in the focal region of an optical system. The field in the entrance pupil can, in principle, be obtained using any rigorous electromagnetic solver and follows from the interaction between an incident plane wave and the object. In general, the light source illuminating the object will be of finite extent. In this case the source is considered as a collection of (weighted) incoherent point sources and the total image is obtained by incoherent summation of the image intensity contributions of all point sources (the Abbe approach).

In Section 3 we have shown that our results can be efficiently implemented into a computation scheme to compute the image produced by a general isolated object. For the most part, our scheme is similar to standard Abbe-type imaging methods, but it uses the ENZ-based algorithms for image formation instead of the more conventional approach based on Fourier optics. Although the ENZ-algorithm is fundamentally different from the Fourier based approach, it does not alter the overall computation scheme significantly. As a result, a conventional piece of Abbe-type imaging software developed modularly can be easily

equipped with our new method by simply replacing the conventional imaging module by its ENZ-based counterpart.

In Section 4 the accuracy and convergence of the ENZ-method is evaluated. It is shown that the series expansion used to generate the ENZ basic functions is accurate all the way down to machine precision. Considering that the ENZ basic functions solely depend on the properties of the imaging system and the location of the image region, and at the same time are independent of the object being imaged, they can be calculated and stored in advance. Based on these observations we can conclude that the accuracy of the ENZ-imaging algorithm is limited by the residual RMS error present in the Zernike expansion of the exit pupil distribution. In principle, as the Zernike polynomials constitute a complete set, the residual error in the expansion can be made arbitrarily small. However, a higher accuracy will require a larger number of expansion coefficients to be determined and this will, of course, increase the computational burden.

For simulation methods it is most relevant to know the relation between the accuracy and the computational burden or complexity. In Subsection 4.4 the theoretical lower boundary for the computational complexity is determined. It is shown that two distinct tasks in the ENZ imaging scheme can both be dominant in terms of computational burden. Whenever an object generates a strongly oscillating pupil distribution, a large number of Zernike coefficients is required for an accurate fit. In this case the least-squares fitting operation requires substantial computational effort. On the other hand, if the pupil distribution is relatively smooth and a fairly large number of image points are of interest, the field construction in the focal region will be the dominant task.

Altogether, we believe that ENZ-based imaging provides an appealing addition to the available arsenal of image simulators. The method can generate extremely accurate results and does so totally independent of other existing methods. As a result, the ENZ-method is an excellent choice to benchmark and this will be one of the focus points of our future work. In addition to this, the ENZ-method works especially well for isolated structures. While other methods have to rely on progressive zero padding to approximate a (quasi) isolated structure, the ENZ-method truly treats the smallest computational domain which fits the isolated object of interest. Of course, this has a beneficial effect on the computational burden in general but especially reduces the amount of memory required for the rigorous treatment of the light interactions with the object.

References

1. E. Abbe, "eiträge zur theorie des mikroskops und der mikroskopischen wahrnemung", Arch. Mikrosk. Anat. Entwichlungsmech. **9**, 413-468 (1873).

2. P.M. Duffieux, *L'Intégrale de Fourier et ses Applications à l'Optique* (privately published, 1946).

3. A. Maréchal, "Study of the combined effects of diffraction and geometrical aberrations on the image of a luminous point," Rev. Opt. **26**, 257-277 (1947).

4. H.H. Hopkins, "On the diffraction theory of optical images," Proc. R. Soc. London, Ser. A **217**, 408-432 (1953).

5. H.H. Hopkins, "The frequency response of a defocused optical system," Proc. R. Soc. London, Ser. A **231**, 91-103 (1955).

6. D. Flagello, T. Milster and A.E. Rosenbluth, "Theory of high-NA imaging in homogeneous thin films," J. Opt. Soc. Am. A **13**, 53-64 (1996).

7. C. J. R. Sheppard and H. J. Matthews, "Imaging in high aperture optical systems. J. Opt. Soc. Am. A **4**, 1354-1360 (1987).

8. C. J. R. Sheppard, "Aberrations in high aperture conventional and confocal imaging systems," Appl. Opt. **27**, 4782-4786 (1988).

9. J. W. Goodman, *Introduction to Fourier optics* (Roberts & Co. Publishers, 2004).

10. K. Adam, Y. Granik, A. Torres and N. Cobb, "Improved modeling performance with an adapted vectorial formulation of the Hopkins imaging equation," Proc. SPIE **5040**, pp. 78-91 (2003).

11. M. Born and E. Wolf, *Principles of Optics*, 7th (expanded) edition (Cambridge University Press, 1999).

12. G. B. Airy, "On the diffraction of an object-glass with circular aperture," Trans. Camb. Phil. Soc. **5**, 283-291 (1834).

13. E. Lommel, "Die beugungserscheinungen einer kreisrunden oeffnung und eines kreisrunden schirmschens theoretisch und experimentell bearbeitet," Abh. Bayer. Akad. **15**, 233-328 (1885).

14. K. Strehl, *Theorie des fernrohrs* (J.A. Bart, 1894).

15. A. E. Conrady, "Star discs," Mon. Not. R. Astr. Soc. **79**, p. 575-593 (1919).

16. A. E. Conrady, "The five aberrations of lens-systems," Mon. Not. R. Astr. Soc. **79**, p. 60-66 (1918).

17. J. Picht, "Über den schwingungsvorgang, der einem beliebigen (astigmatischen) strahlenbündel entspricht," Ann. Phys. **77**, 685-782 (1925).

18. F. Zernike, "Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode." Physica **1**, 689-704 (1934).

19. B. R. A. Nijboer, *The diffraction theory of aberrations*, Ph. D. Thesis (University of Groningen, Groningen, The Netherlands, 1942).

20. F. Zernike and B. R. A. Nijboer, "Théorie de la diffraction des aberrations," in La Théorie des Images Optiques, P. Fleury, A. Marchal, and C. Anglade, eds. (Editions de la Revue d'Optique, Paris, 1949), pp. 227-235.

21. V. S. Ignatowsky, "Diffraction of a lens of arbitrary aperture," Trans. Opt. Inst. **1**, 1-36 (1919).

22. H. H. Hopkins, "The airy disc formula for systems of high relative aperture," Proc. Phys. Soc. **55**, 116-128 (1943).

23. E. Wolf, "Electromagnetic diffraction in optical systems. I. An integral representation of the image field," Proc. R. Soc. London, Ser. A **253**, 349-357, (1959).

24. P. Török and P. Munro, "The use of Gauss-Laguerre vector beams in STED microscopy," Opt. Express 12, 3605-3617 (2004).

25. C. J. R. Sheppard and P. Török, "Efficient calculation of electromagnetic diffraction in optical systems using a multipole expansion," J. Mod. Opt. **44**, 803-818 (1997).

26. B. R. Frieden, "Evaluation, design and extrapolation methods for optical signals, based on the prolate functions," in *Progress in Optics* **9**, E. Wolf ed., (Pergamon, 1971), pp. 311-407.

27. C. J. R. Sheppard, "Aberrations in high aperture conventional and confocal imaging systems," Appl. Opt. **27**, 4782-4786 (1988).

28. C. Colautti, B. Ruiz, E. E. Sicre and M. Garavaglia, "Walsh functions: Analysis of their properties under Fresnel diffraction," J. Mod. Opt. **34**, 1385-1391 (1987).

29. A. J. E. M. Janssen, "Extended Nijboer-Zernike approach for the computation of optical point-spread functions," J. Opt. Soc. Am. A **19**, 849-857 (2002).

30. 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," J. Opt. Soc. Am. **A 19**, 858-870 (2002).

31. J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen and A. van de Nes, "Extended Nijboer-Zernike representation of the field in the focal region of an aberrated high-aperture optical system," J. Opt. Soc. Am. A **20**, 2281-2292 (2003).

32. S. van Haver, O.T.A. Janssen, A.J.E.M. Janssen, J.J.M. Braat, H.P. Urbach and S.F. Pereira, "General imaging of advanced 3D mask objects based on the fully-vectorial extended Nijboer-Zernike (ENZ) theory," Proc. SPIE **6924**, 69240U, 1-8 (2008).

33. H. J. Levinson, *Principles of lithography* (2nd edition) (SPIE, 2001).

34. K. Yee. "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media," IEEE T. Antenn. Propag. **14**, 302-307 (1966).

35. B. Richards and E. Wolf, "Electromagnetic diffraction in optical systems. II. Structure

of the image field in an aplanatic system," Proc. R. Soc. London, Ser. A **253**, 358-379, (1959).

36. J.J.M. Braat, S. van Haver, A.J.E.M. Janssen and P. Dirksen, "Assessment of optical systems by means of point-spread functions," in *Progress in Optics* **51**, E. Wolf ed., (Elsevier, 2008), pp. 349-468.

37. J.J.M. Braat, P. Dirksen, A.J.E.M. Janssen, A.S. van de Nes and S. van Haver, "Extended Nijboer-Zernike approach to aberration and birefringence retrieval in a high-numerical-aperture optical system," J. Opt. Soc. Am. A. 22 (2005), pp. 2635-2650.

38. S. van Haver, J.J.M. Braat, P. Dirksen and A.J.E.M. Janssen, "High-NA aberration retrieval with the extended Nijboer-Zernike vector diffraction theory," J. Europ. Opt. Soc. Rap. Public. **1**, 06004 (2006).

39. S. van Haver, J.J.M. Braat, P. Dirksen and A.J.E.M. Janssen, "High-NA aberration retrieval with the extended Nijboer-Zernike vector diffraction theory - Erratum," J. Europ. Opt. Soc. Rap. Public. **2**, 07011e (2007).

40. A.K. Wong, *Optical Imaging in Projection Microlithography*, Tutorial Text Series **TT66** (SPIE Press, 2005).

41. J.P. Berenger, "Three-dimensional perfectly matched layer for the absorption of electromagnetic waves," J. Comput. Phys. **127**, 363-379 (1996).

42. J.A. Roden and S.D. Gedney, "Convolutional PML (CPML): An efficient FDTD implementation of the CFS-PML for arbitrary media," Microw. Opt. Tech. Lett. **27**, 334-339 (2000).

43. D.T. Prescott and N.V. Shuley, "Extensions to the FDTD method for the analysis of infinitely periodic arrays," IEEE Microw. Guid. Wave Lett. **4**, 352-354 (1994).

44. 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 J. Quant. Electron. **40**, 175-182 (2004).

45. O.T.A. Janssen, S. van Haver, A.J.E.M. Janssen, J.J.M. Braat, H.P. Urbach and S.F. Pereira, "Extended Nijboer-Zernike (ENZ) based mask imaging: efficient coupling of electromagnetic field solvers and the ENZ imaging algorithm," Proc. SPIE **6924**, 692410, (2008).

46. J.A. Stratton and L.J. Chu, "Diffraction theory of electromagnetic waves," Phys. Rev. **56**, 99-107 (1939).

47. 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).

48. P Lalanne, M. Besbes, J.P. Hugonin, S. van Haver, O.T.A. Janssen, A.M. Nugrowati, M. Xu, S.F. Pereira, HP 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," J. Europ. Opt. Soc. Rap. Public. **2**, 07022 (2007).

49. MathWorks Matlab, Version 7.5.0.342 (R2007b), www.mathworks.com.

A. Series expansion for $V_{n,j}^m$

In this appendix we present a method for obtaining a series expansion for the integral $V_{n,j}^m(r, f)$ given by

$$V_{n,j}^m(r, f) = \int_0^1 \rho^{|j|} \frac{\left\{(1 - s_0{}^2\rho^2)^{\frac{1}{2}} + (1 - n_1^2 M^2 s_0^2 \rho^2/n_0^2)^{\frac{1}{2}}\right\}^{-|j|+1}}{(1 - s_0{}^2\rho^2)^{\frac{1}{4}} (1 - n_1^2 M^2 s_0{}^2\rho^2/n_0^2)^{\frac{3}{4}}} \times$$
$$\exp\left[\frac{if}{u_0}\left(1 - \sqrt{1 - s_0^2\rho^2}\right)\right] R_n^{|m|}(\rho) J_{m+j}(2\pi r\rho)\rho\, d\rho. \tag{A1}$$

We follow a similar approach as in [36], Appendix D, to transform the integral in (A1) into a tractable form. We write

$$\exp\left[\frac{if}{u_0}\left(1 - \sqrt{1 - s_0^2\rho^2}\right)\right] \frac{\left\{(1 - s_0{}^2\rho^2)^{\frac{1}{2}} + (1 - n_1^2 M^2 s_0^2\rho^2/n_0^2)^{\frac{1}{2}}\right\}^{-|j|+1}}{(1 - s_0{}^2\rho^2)^{\frac{1}{4}} (1 - n_1^2 M^2 s_0{}^2\rho^2/n_0^2)^{\frac{3}{4}}}$$
$$= \exp\left\{g' + if'\rho^2\right\} \sum_{t=0}^\infty B_t \rho^{2t}\ . \tag{A2}$$

Here the coefficients g' and f' are defined by requiring the best fit for the constant and the quadratic term in ρ in the ln of the function (A2). So let

$$F(\rho) = \frac{if}{u_0}\left(1 - \sqrt{1 - s_0^2\rho^2}\right)$$
$$+ (-|j| + 1)\ln\left[(1 - s_0{}^2\rho^2)^{\frac{1}{2}} + \left(1 - n_1^2 M^2 s_0^2\rho^2/n_0^2\right)^{\frac{1}{2}}\right]$$
$$- \frac{1}{4}\ln\left(1 - s_0{}^2\rho^2\right) - \frac{3}{4}\ln\left(1 - n_1^2 M^2 s_0^2\rho^2/n_0^2\right)\ , \tag{A3}$$

and define

$$A(\rho) = 1 - \sqrt{1 - s_0^2\rho^2} = \sum_{n=0}^\infty a_n \rho^{2n}\ , \tag{A4}$$

$$B(\rho) = \ln\left[(1 - s_0{}^2\rho^2)^{\frac{1}{2}} + \left(1 - n_1^2 M^2 s_0^2\rho^2/n_0^2\right)^{\frac{1}{2}}\right] = \sum_{n=0}^\infty b_n \rho^{2n}\ , \tag{A5}$$

$$C(\rho) = \ln\left(1 - s_0{}^2\rho^2\right) + 3\ln\left(1 - n_1^2 M^2 s_0^2\rho^2/n_0^2\right) = \sum_{n=0}^\infty c_n \rho^{2n}\ , \tag{A6}$$

so that

$$F(\rho) = \frac{if}{u_0}A(\rho) + (-|j| + 1)B(\rho) - \frac{1}{4}C(\rho) = \sum_{n=0}^\infty f_n \rho^{2n}\ , \tag{A7}$$

$$f_n = \frac{if}{u_0}a_n + (-|j| + 1)b_n - \frac{1}{4}c_n\ , \quad n = 0, 1, \dots\ . \tag{A8}$$

We shall determine a_n, b_n and c_n.

a_n: We have by Taylor expansion

$$a_0 = 0 \;\; ; \;\; a_n = -\binom{\frac{1}{2}}{n}(-1)^n s_0^{2n} \;\; , \;\; n = 1, 2, \dots . \tag{A9}$$

b_n: Consider the function

$$f(x) = \ln\left[(1-x)^{\frac{1}{2}} + (1-ax)^{\frac{1}{2}}\right] . \tag{A10}$$

Then

$$
\begin{aligned}
f'(x) &= \frac{1}{2}\frac{1}{(1-x)^{\frac{1}{2}} + (1-ax)^{\frac{1}{2}}}\left(\frac{-1}{(1-x)^{\frac{1}{2}}} + \frac{-a}{(1-ax)^{\frac{1}{2}}}\right) \\
&= -\frac{1}{2}\frac{(1-x)^{\frac{1}{2}} - (1-ax)^{\frac{1}{2}}}{(1-x)-(1-ax)}\frac{(1-ax)^{\frac{1}{2}} + a(1-x)^{\frac{1}{2}}}{(1-x)^{\frac{1}{2}}(1-ax)^{\frac{1}{2}}} \\
&= \frac{1}{2x}\left(1 - (1-ax)^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}\right) \\
&= -\frac{1}{2}\sum_{t=1}^{\infty}\left[(-1)^t\sum_{r=0}^{t}\binom{-\frac{1}{2}}{r}\binom{-\frac{1}{2}}{t-r}a^r\right]x^{t-1} ,
\end{aligned}
\tag{A11}
$$

where the third identity is obtained by working out the numerator and denominator and simplifying, and the fourth identity is obtained by multiplying the Taylor expansions of $(1-ax)^{-1/2}$ and of $(1-x)^{-1/2}$. Therefore, by integrating from 0 to x,

$$f(x) = \ln 2 - \sum_{t=1}^{\infty}\left[\frac{(-1)^t}{2t}\sum_{r=0}^{t}\binom{-\frac{1}{2}}{r}\binom{-\frac{1}{2}}{t-r}a^r\right]x^t . \tag{A12}$$

Using (A12) with $a = n_1^2 M^2/n_0^2$ and $x = s_0^2\rho^2$, we see that

$$b_0 = \ln 2 \;\; ; \;\; b_n = -\frac{(-1)^n s_0^{2n}}{2n}\sum_{r=0}^{n}\binom{-\frac{1}{2}}{r}\binom{-\frac{1}{2}}{n-r}\frac{n_1^{2r}M^{2r}}{n_0^{2r}} \;\; , \;\; n = 1, 2, \dots . \tag{A13}$$

c_n: We have by Taylor expansion

$$c_0 = 0 \;\; ; \;\; c_n = -\frac{(s_0)^{2n}(1 + 3(\frac{n_1 M}{n_0})^{2n})}{n} \;\; , \;\; n = 1, 2, \dots . \tag{A14}$$

Thus with a_n, b_n, c_n from (A9), (A13), (A14), we can compute the f_n of $F(\rho) = \sum_{n=0}^{\infty} f_n \rho^{2n}$ according to (A8).

Next, we proceed by writing

$$F(\rho) = g' + if'\rho^2 + \sum_{n=0}^{\infty} A_n \rho^{2n} , \tag{A15}$$

30

where $g' + if'\rho^2$ is the best quadratic approximation of $F(\rho)$ using $\rho d\rho$ on $[0, 1]$ as weight function. To this end, we convert the Taylor expansion $\sum_{n=0}^{\infty} f_n \rho^{2n}$ of $F(\rho)$ into a Zernike expansion $\sum_{k=0}^{\infty} \beta_{2k}^0 R_{2k}^0$ according to the formula

$$\beta_{2k}^0 = \sum_{n=k}^{\infty} \frac{2k+1}{k+1} \frac{\binom{n}{k}}{\binom{n+k+1}{n}} f_n \ , \quad k = 0, 1, \dots , \tag{A16}$$

see [29], Eq. (10). Now

$$\beta_0^0 R_0^0(\rho) + \beta_2^0 R_2^0(\rho) = (\beta_0^0 - \beta_2^0) + 2\beta_2^0 \rho^2 \tag{A17}$$

is the desired best approximation of $F(\rho)$. Hence

$$g' = \beta_0^0 - \beta_2^0 \ , \quad f' = \frac{2}{i} \beta_2^0 \ , \tag{A18}$$

and

$$A_0 = f_0 - g' \ , \quad A_1 = f_1 - if' \ ; \quad A_n = f_n \ , \quad n = 2, 3, \dots . \tag{A19}$$

The final step in achieving (A2) consists of writing

$$\exp\left(\sum_{n=0}^{\infty} A_n \rho^{2n} \right) = \sum_{t=0}^{\infty} B_t \rho^{2t} \ . \tag{A20}$$

As in [36], Appendix D this is done recursively according to

$$B_0 = \exp\left(A_0 \right) \ , \quad B_{t+1} = \sum_{j=0}^{t} \frac{t+1-j}{t+1} A_{t+1-j} B_j \ , \quad t = 0, 1, \dots . \tag{A21}$$

From (A2) one can now compute the $V_{n,j}^m$ in (A1) as in [36], Appendix D.2. Thus one writes (with some minor corrections on [36], Appendix D.2)

$$R_n^{|m|}(\rho) = \rho^{|m|} \sum_{s=0}^{p} C_s \rho^{2s} \ , \quad C_s = (-1)^{p-s} \binom{q+s}{p} \binom{p}{s} \ , \tag{A22}$$

where $p = (n - |m|)/2$, $q = (n + |m|)/2$. Then

$$V_{n,j}^m(r, f) = \sum_{s=0}^{p} \sum_{t=0}^{\infty} C_s B_t T_{|j|+|m|+2s+2t}^{m+j}(r, f') \ , \tag{A23}$$

where for integer k, l with $l - |k|$ even and ≥ 0

$$T_l^k(r, f') = \int_0^1 \rho^l e^{if'\rho^2} J_k(2\pi r \rho) \rho d\rho \ . \tag{A24}$$

These T_l^k have been computed in [29], Eqs. (14-16) in the form of a power-Bessel series.