

Microscopy image resolution improvement by deconvolution of complex fields

Yann Cotte,* M. Fatih Toy, Nicolas Pavillon, and Christian Depeursinge

Ecole Polytechnique Fédérale de Lausanne (EPFL), Microvision and Microdiagnostics Group, 1015 Lausanne, Switzerland

*yann.cotte@a3.epfl.ch

Abstract: Based on truncated inverse filtering, a theory for deconvolution of complex fields is studied. The validity of the theory is verified by comparing with experimental data from digital holographic microscopy (DHM) using a high-NA system (NA=0.95). Comparison with standard intensity deconvolution reveals that only complex deconvolution deals correctly with coherent cross-talk. With improved image resolution, complex deconvolution is demonstrated to exceed the Rayleigh limit. Gain in resolution arises by accessing the objects complex field - containing the information encoded in the phase - and deconvolving it with the reconstructed complex transfer function (CTF). Synthetic (based on Debye theory modeled with experimental parameters of MO) and experimental amplitude point spread functions (APSF) are used for the CTF reconstruction and compared. Thus, the optical system used for microscopy is characterized quantitatively by its APSF. The role of noise is discussed in the context of complex field deconvolution. As further results, we demonstrate that complex deconvolution does not require any additional optics in the DHM setup while extending the limit of resolution with coherent illumination by a factor of at least 1.64.

© 2010 Optical Society of America

OCIS codes: (100.1830) Deconvolution; (030.1670) Coherent optical effects; (090.1995) Digital holography; (100.6640) Superresolution; (100.5070) Phase retrieval; (110.0180) Microscopy.

References and links

1. S. V. Aert, D. V. Dyck, and A. J. den Dekker, "Resolution of coherent and incoherent imaging systems reconsidered—classical criteria and a statistical alternative," *Opt. Express* **14**, 3830–3839 (2006).
2. J. G. McNally, T. Karpova, J. Cooper, and J. A. Conchello, "Three-dimensional imaging by deconvolution microscopy," *Methods* **19**, 373–385 (1999).
3. C. Vonesch, "Fast and automated wavelet-regularized image restoration in fluorescence microscopy," Ph.D. thesis, EPFL, LIB Laboratoire d'imagerie biomédicale (2009).
4. W. Wallace, L. H. Schaefer, and J. R. Swedlow, "A workingperson's guide to deconvolution in light microscopy," *Biotechniques* **31** (2001).
5. B. Colicchio, O. Haerberl, C. Xu, A. Dieterlen, and G. Jung, "Improvement of the lls and map deconvolution algorithms by automatic determination of optimal regularization parameters and pre-filtering of original data," *Opt. Commun.* **244**, 37–49 (2005).
6. F. Aguet, S. Geissbühler, I. Märki, T. Lasser, and M. Unser, "Super-resolution orientation estimation and localization of fluorescent dipoles using 3-d steerable filters," *Opt. Express* **17**, 6829–6848 (2009).
7. P. Sarder and A. Nehorai, "Deconvolution methods for 3-d fluorescence microscopy images," *IEEE Signal Process. Mag.* **23**, 32–45 (2006).

8. E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of fresnel off-axis holograms," *Appl. Opt.* **38**, 6994–7001 (1999).
9. C. Depeursinge, P. Jourdain, B. Rappaz, P. Magistretti, T. Colomb, and P. Marquet, "Cell biology explored with digital holographic microscopy," *Biomed. Opt. p. BMD58* (2008).
10. C. J. Sheppard, "Fundamentals of superresolution," *Micron* **38**, 165–169 (2007).
11. D. Mendlovic, A. W. Lohmann, N. Konforti, I. Kiryuschev, and Z. Zalevsky, "One-dimensional superresolution optical system for temporally restricted objects," *Appl. Opt.* **36**, 2353–2359 (1997).
12. A. Shemer, D. Mendlovic, Z. Zalevsky, J. Garcia, and P. G. Martinez, "Superresolving optical system with time multiplexing and computer decoding," *Appl. Opt.* **38**, 7245–7251 (1999).
13. Z. Zalevsky and D. Mendlovic, *Optical superresolution*, vol. 91 (Springer, 2004).
14. E. N. Leith, D. Angell, and C. P. Kuei, "Superresolution by incoherent-to-coherent conversion," *J. Opt. Soc. Am. A* **4**, 1050–1054 (1987).
15. R. Gerchberg and W. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 227–246 (1972).
16. V. Mico, Z. Zalevsky, C. Ferreira, and J. García, "Superresolution digital holographic microscopy for three-dimensional samples," *Opt. Express* **16**, 19260–19270 (2008).
17. G. Indebetouw, Y. Tada, J. Rosen, and G. Brooker, "Scanning holographic microscopy with resolution exceeding the rayleigh limit of the objective by superposition of off-axis holograms," *Appl. Opt.* **46**, 993–1000 (2007).
18. V. Lauer, "New approach to optical diffraction tomography yielding a vector equation of diffraction tomography and a novel tomographic microscope," *J. Microsc.* **205**, 165–176 (2002).
19. M. Debailleul, V. Georges, B. Simon, R. Morin, and O. Haeberle, "High-resolution three-dimensional tomographic diffractive microscopy of transparent inorganic and biological samples," *Opt. Lett.* **34**, 79–81 (2009).
20. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**, 82–87 (2000).
21. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968).
22. M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, 1987), 6th ed.
23. M. Gu, *Advanced Optical Imaging Theory* (Springer-Verlag, 2000).
24. Y. Cotte and C. Depeursinge, "Measurement of the complex amplitude point spread function by a diffracting circular aperture," in "Focus on Microscopy," (2009), Advanced linear and non-linear imaging, pp. TU–AF2–PAR–D.
25. X. Heng, X. Q. Cui, D. W. Knapp, J. G. Wu, Z. Yaqoob, E. J. McDowell, D. Psaltis, and C. H. Yang, "Characterization of light collection through a subwavelength aperture from a point source," *Opt. Express* **14**, 10410–10425 (2006).
26. A. Marian, F. Charrière, T. Colomb, F. Montfort, J. Kühn, P. Marquet, and C. Depeursinge, "On the complex three-dimensional amplitude point spread function of lenses and microscope objectives: theoretical aspects, simulations and measurements by digital holography," *J. Microsc.* **225**, 156–169 (2007).
27. N. Pavillon, C. S. Seelamantula, J. Kühn, M. Unser, and C. Depeursinge, "Suppression of the zero-order term in off-axis digital holography through nonlinear filtering," *Appl. Opt.* **48**, H186–H195 (2009).
28. C. J. R. Sheppard and M. Gu, "Imaging by a high aperture optical-system," *J. Mod. Opt.* **40**, 1631–1651 (1993).
29. M. Leutenegger, R. Rao, R. A. Leitgeb, and T. Lasser, "Fast focus field calculations," *Opt. Express* **14**, 11277–11291 (2006).
30. D. E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning* (Addison-Wesley, 1989).
31. V. Torczon, "On the convergence of pattern search algorithms," *SIAM J. Optim.* **7**, 125 (1997).
32. Y. Cotte, M. F. Toy, E. Shaffer, N. Pavillon, and C. Depeursinge, "Sub-rayleigh resolution by phase imaging," *Opt. Lett.* **35**, 2176–2178 (2010).
33. M. Totzeck and H. J. Tiziani, "Phase-singularities in 2d diffraction fields and interference microscopy," *Opt. Commun.* **138**, 365–382 (1997).
34. C. J. Sheppard and K. Larkin, "Vectorial pupil functions and vectorial transfer functions," *Optik (Stuttg.)* **107**, 79–87 (1997).
35. H. Guo, S. Zhuang, J. Chen, and Z. Liang, "Imaging theory of an aplanatic system with a stratified medium based on the method for a vector coherent transfer function," *Opt. Lett.* **31**, 2978–2980 (2006).

1. Introduction

Coherently illuminated imaging systems suffer from an inferior lateral resolution compared to their incoherent counterpart [1]. This aspect is further intensified by a variety of post-processing methods to improve the image quality of incoherent light microscopy [2,3]. Many 2D deconvolution methods, like deblurring, can be applied to improve image quality of incoherent imaging systems [4] and 3D deconvolution techniques give rise to enhanced optical sectioning capability [2]. Based on iterative expectation-maximization algorithms for maximum-likelihood deconvolution

lution of incoherent images, even enhanced resolution has been demonstrated [5,6] at the cost of computational power. All such efforts make deconvolution a common post-processing method for biological applications such as deconvolution of fluorescence microscopy images [7].

On the other hand, the capability of DHM to image and measure amplitude and quantitative phase simultaneously, makes it an attractive research tool in many fields of biological research [8,9], since it is marker free, non-invasive regarding the light intensity, and only camera acquisition rate limited. Consequently, attempts have been made to bring the conveniences of improved resolution to coherent light systems, too.

The general idea is to use degrees of freedom that are deemed unnecessary [10]. For example they can be in real space [11], in the temporal domain [12], in the spectral domain [13], or in the polarization [14]. Generally, these methods require alteration of experimental setup with additional modifications e.g. gratings or mechanically moving parts, giving rise to practical issues. By using the phase-retrieval method of Gerchber-Saxton [15], attempts have been made to improve time multiplexing [13]. Nonetheless, DHM offers already the intrinsic advantage of providing the amplitude A as well as the phase Φ from the reconstructed complex field U . Time multiplexing methods combined with DHM methods have been demonstrated to work with low-NA systems [16] but still demand of scalability to high-NA. For 'midrange' systems of NA=0.42, a resolution improvement of nearly 2 is possible with a synthetic aperture requiring, however, the use of a detection scan [17]. For aperture synthesis by beam scanning [18], tomographic diffractive microscopy (TDM) recently showed high-resolution with a 'highrange' NA system [19]. Other coherent light methods like structured illumination microscopy (SIM) [20] use coherent excitation for intensity based fluorescence imaging. Despite demonstration of sub-wavelength resolution [extended to Abbe's resolution limit $\lambda/(2NA)$] by phase structuring, the complex detection is only partially used in excitation.

For the first time to our knowledge, we demonstrate coherent sub-wavelength resolution [$\lambda/(2NA)$] with high-NA (NA=0.95) by directly using the information content available from amplitude as well as from phase in DHM. By adapting mentioned standard deconvolution post processing methods to coherent illumination imaging conditions, the phase imaging process does not need to be compromised. No additional optical components nor scanning procedures are required since the method is applied at a step posterior to the experiment itself. We also show that the usual trade-off between precision in object localization and Rayleigh's resolution criterion [1] can be overcome by the presented method.

The structure of the paper is to give first a theoretical overview of conventional deconvolution, following referred to as 'intensity deconvolution'. From that, the theory of 'complex deconvolution' in its physical model is presented in the last part of the 2nd section. In the 3rd section experimental details are provided and the 4th section shows how to treat the test target's and the APSF's data. Furthermore, the adaption of the synthetic CTF is outlined and used for the determination of the influence of noise in the 5th section. Also in this section, the final results are presented and compared to intensity deconvolution. A discussion of the implementation of complex deconvolution concludes in the 6th section.

2. Theory

Coherent and incoherent imaging systems behave very differently. One simple attribute of image properties is the spectrum of spatial frequencies ν

$$\nu_{c,inc} = 2\nu_{c,coh}, \quad (1)$$

which allows double the frequency for an incoherent (inc) system compared to the coherent (coh) counterpart [21]. Apart from Eq. (1), the frequency transmission is differently shaped, triangular for incoherent and rectangular for coherent cases of standard imaging. The respective

shape results in better imaging contrast for coherent systems and a smaller limit of resolution for the incoherent counterpart. The limit of resolution according to Rayleigh's criterion of resolution is given by:

$$d_{min} = \alpha \frac{\lambda}{NA}, \quad (2)$$

with $\alpha_{coh} = 0.82$ and $\alpha_{incoh} = 0.61$ [22]. Likewise, the wavenumber k is related to spatial frequency ν and wave vector $\vec{k} = (k_x, k_y, k_z)$ by

$$k = |\vec{k}| = 2\pi\nu = \frac{2\pi}{d}, \quad (3)$$

where d corresponds to a spatial distance and k_c is related to d_{min} by

$$k_{c,coh} = \frac{2\pi}{d_{min,coh}}. \quad (4)$$

2.1. Inverse filter deconvolution of intensity fields

For a 2D incoherently illuminated imaging system with magnification M , the intensity function $I(x_2, y_2)$ in the image plane is presented as a convolution integral (following notation is based on [23])

$$I(x_2, y_2) = \iint_{-\infty}^{\infty} |h(x_1 + Mx_2, y_1 + My_2)|^2 o_i(x_1, y_1) dx_1 dy_1, \quad (5)$$

where $o_i(x_1, y_1)$ is the intensity function in the object plane and $h(x, y)$ is the complex point spread function (APSF). Because of the use of the intensity point spread function (IPSF) $|h(x, y)|^2$, no phase term is included. One can express this in k -space as

$$J(k_x, k_y) = C(k_x, k_y) O_i(k_x, k_y),$$

and
$$I(x_2, y_2) = \iint_{-\infty}^{\infty} J(k_x, k_y) \exp[-i2\pi(k_x Mx_2 + k_y My_2)] dk_x dk_y, \quad (6)$$

where J , O_i , and C are the 2-D Fourier transform of I , o_i , and $|h|^2$, respectively, such as

$$C(k_x, k_y) = \iint_{-\infty}^{\infty} |h(x_1, y_1)|^2 \exp[i2\pi(k_x x_1 + k_y y_1)] dx_1 dy_1. \quad (7)$$

Conventionally, C is called the optical transfer function (OTF), J the intensity image spectrum, and O_i the intensity object spectrum. For incoherently illuminated imaging systems, the standard deconvolution approach, namely inverse filtering [7], consists in inverting Eq. (6):

$$o_i(x_1, y_1) = \iint_{-\infty}^{\infty} O_i(k_x, k_y) \exp[-i2\pi(k_x x_1 + k_y y_1)] dk_x dk_y = \mathcal{F}^{-1} \left\{ \frac{\tilde{J}(k_x, k_y)}{C(k_x, k_y)} \right\}. \quad (8)$$

Instead of dividing $J(k_x, k_y)$ itself, $\tilde{J}(k_x, k_y)$ is introduced with a low-pass filtered spectrum and permits to suppress noise related spatial frequencies above a maximal wavenumber k_{max} .

$$\tilde{J}(k_x, k_y) = J(k_x, k_y) \Gamma_{k_{max}}(k_x, k_y) \quad \text{where} \quad \Gamma_{k_{max}}(k_x, k_y) = \begin{cases} 1 & \sqrt{k_x^2 + k_y^2} \leq k_s \\ f_s & k_s \leq \sqrt{k_x^2 + k_y^2} \leq k_{max} \\ 0 & \sqrt{k_x^2 + k_y^2} > k_{max} \end{cases} \quad (9)$$

The interplay of k_c , k_{max} , and k_s is discussed in detail in section 5.2. The basic idea of the function f_s is to linearly decrease the frequency values from $f_s(k_s) = 1$ to $f_s(k_{max}) = 0$ within a small smoothing kernel $k_{max} - k_s$. The purpose is to smooth the mask's rim borders in order to suppress aliasing effects.

Apart from this truncated inverse filtering deconvolution, there exists a variety of more elaborated methods [7]. However, for the sake of generality, complex deconvolution is derived accordingly to Eq. (8). It is conventionally referred as the most 'physical' method since it consists only in inverting the imaging process [7].

2.2. Inverse filter deconvolution of complex fields

For a 2D imaging system based on coherent illumination, the complex image function $U(x_2, y_2)$ can be expressed as the convolution of the complex object function $o(x_1, y_1)$ and the APSF [23]:

$$U(x_2, y_2) = \iint_{-\infty}^{\infty} h(x_1 + Mx_2, y_1 + My_2) o(x_1, y_1) dx_1 dy_1, \quad (10)$$

We express this relation in the Fourier domain again:

$$\begin{aligned} G(k_x, k_y) &= c(k_x, k_y) O(k_x, k_y), \\ \text{and } U(x_2, y_2) &= \iint_{-\infty}^{\infty} G(k_x, k_y) \exp[-i2\pi(k_x Mx_2 + k_y My_2)] dk_x dk_y, \end{aligned} \quad (11)$$

where G , O , and c are the 2-D Fourier transform of U , o , and h , respectively, such as

$$c(k_x, k_y) = \iint_{-\infty}^{\infty} h(x_1, y_1) \exp[i2\pi(k_x x_1 + k_y y_1)] dx_1 dy_1, \quad (12)$$

hence, c is called the coherent transfer function (CTF), G the complex image spectrum and O the complex object spectrum.

Following the idea of Eqs. (8) and (11), the complex deconvolution can be recast in an easier expression as a deconvolution with an experimental or synthetic CTF :

$$o(x_1, y_1) = \iint_{-\infty}^{\infty} O(k_x, k_y) \exp[-i2\pi(k_x x_1 + k_y y_1)] dk_x dk_y = \mathcal{F}^{-1} \left\{ \frac{\tilde{G}(k_x, k_y)}{c(k_x, k_y)} \right\}. \quad (13)$$

The inverse filtering can be performed directly by dividing the two complex fields of G and c . However, just as the intensity based approach [2], the inverse filtering method in the complex domain suffers from noise amplification for small values of the denominator of $\tilde{G}(k_x, k_y)/c(k_x, k_y)$, especially at high spatial frequencies. Consequently, \tilde{G} is defined accordingly to Eq. (9):

$$\tilde{G}(k_x, k_y) = G(k_x, k_y) \Gamma_{k_{max}}(k_x, k_y). \quad (14)$$

The mask $\Gamma_{k_{max}}$ can be real numbered since zero amplitude annihilates the phase contribution in Eq. (14). The spectrum of c is accordingly bandwidth limited by dividing \tilde{G} so that only 'unphysical' frequencies related to noise are truncated. In intensity deconvolution, however, there is a need for more sophisticated methods [2, 3, 7]: small values within the bandwidth are still sensitive to noise. We expect complex deconvolution to be less noise sensitive. As mentioned, the CTF is rectangularly shaped, hence constant in transmission. The autocorrelation of the CTF results in a triangular OTF with a linearly decreasing transmission. Therefore, noise is less amplified in Eq. (13) because division by a very small intensity value in the denominator of Eq. (13) is improbable, contrary to in the division by the OTF in Eq. (8). Often, additional assumptions (noise's origin or amplitude) [2, 3, 7] or iterative methods serve to improve intensity

based deconvolution [5, 6]. The presented complex deconvolution foregoes any assumptions since it is simply based on inverting image formation from Eq. (10).

3. Experiment

To demonstrate the effectiveness of complex deconvolution, we use a test target consisting of a thin opaque aluminum film (thickness=100nm) on a conventional coverslip [24]. Nano-metric apertures ($\varnothing_{nominal}=80\text{nm}$) are drilled with focused ion beam (FIB) milling in the coating and are placed at very close pitches η . The fabricated pitch is controlled and measured by scanning electron microscopy (SEM), as shown in Fig. 1.

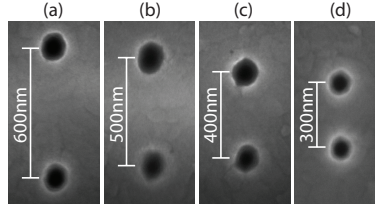


Fig. 1. SEM image of pair of nano-holes drilled by FIB in aluminum film at 100 000 \times magnification. The images show nominal center-to-center pitches η of 600nm (a), 500nm (b), 400nm (c) and 300nm (d) with according scale bars.

The SEM measurements of the diameter \varnothing reveals marginally bigger diameters ($\varnothing_{real} \approx 90\text{nm}$) than the nominal ones due to their slightly conical shapes. The real pitch η , however, varies only within $\pm 5\text{nm}$ from the nominal specifications.

A single nano-metric aperture can serve as an experimental complex point source for h_{exp} [24, 25] and its imaged amplitude and phase have been shown to be characteristic [26]. The differently pitched double hole series (cf. Fig. 1) serve as experimental test targets for U . The light source is a YAG laser at $\lambda=532\text{nm}$. The used microscope objective is a Zeiss $\times 63$ NA=0.95 in air (refractive index $n_m = 1$) in combination with a relay magnification to reach a lateral sampling of $\delta x = 56\text{nm}$.

4. Processing

The complex field is provided by digital holographic microscopy (DHM) [8] in transmission configuration. Thus, the amplitude $A(x,y)$ as well as the phase $\Phi(x,y)$ of the complex field emitted by nano-holes can be extracted by following the methods of [8] which is expressed as

$$U(x,y) = a_n A(x,y) \exp[i\Phi(x,y)], \quad (15)$$

where a_n is a real normalization constant.

4.1. Experimental APSF

DHM's feature of digital refocusing is used to propagate the recorded holograms in the focal plane from which the CTF and OTF can be calculated by Eqs. (12) and (7), respectively. The reconstructed experimental CTF of the single aperture is called c_{exp} and is illustrated in Figs. 2(a) and 2(c).

The modulus of c_{exp} describes the system's transmittance as function of the related wavenumber. The high value for the transmittance of the carrier wave (seen for $|c_{exp}(k_x = 0, k_y = 0)|$), i.e. the background brightness of the image, is specific to the DHM's dynamic range, ideally from zero to the sum of the reference and object wave intensities.

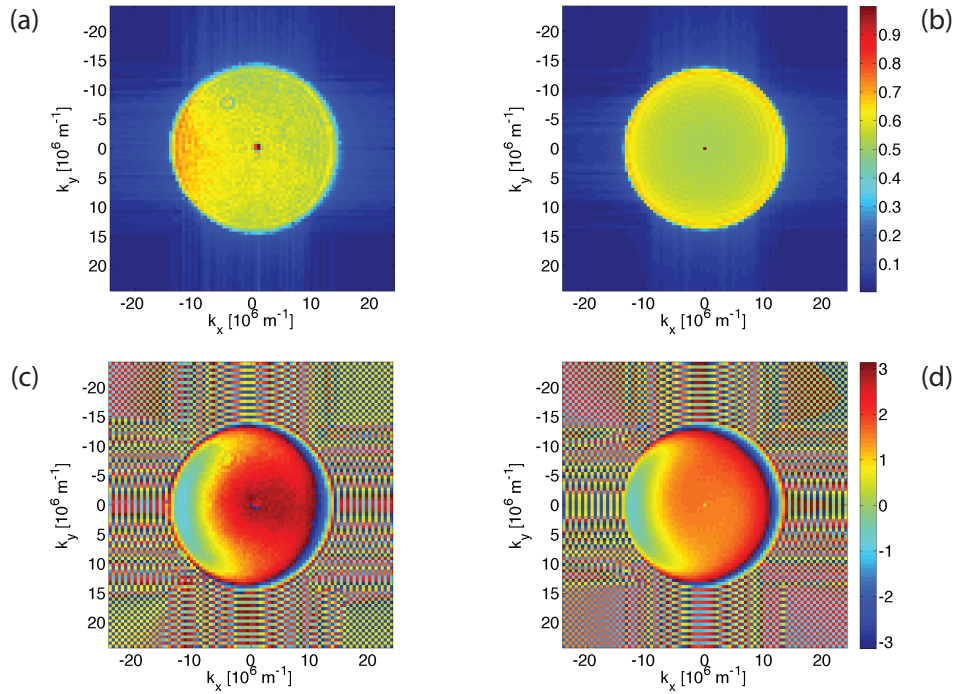


Fig. 2. Experimental and synthetic transfer functions in focal plane at $\lambda=532\text{nm}$ and $\text{NA}=0.95$. The experimental amplitude CTF $|c_{exp}|$ (a) and phase CTF $\arg[c_{exp}]$ (c) are imaged from a single nano-metric aperture. According to Eq. (22), (b) shows the fitted synthetic amplitude CTF $|c_{syn}|$ and (d) its phase part $\arg[c_{syn}]$.

The phase of the transmitted wavefronts are shown in parts (c) and (d) of Fig. 2. In the focal plane, the representation of the wavefront in k -space results in the phase of the focal spot image. Vice versa, defocusing results in de-phasing of the transmitted wavefront. The effects of aberration on the system can be seen e.g. the coma as a deformation of the wavefront phase in Fig. 2(c).

To ensure optimal optical imaging conditions the effective system's NA must be estimated from the experimentally reconstructed CTF. The effective NA relates to the discrete spectral support as [27]

$$NA = \frac{n_m \lambda}{N \delta x} m_{px}, \quad (16)$$

with a square recording zone of $N \times N$ pixels, immersion refractive index n_m , and uniform lateral sampling δx . The discrete spatial bandwidth of the microscope objective corresponds to m_{px} . By combining Eq. (3) with Eq. (16), the axes of Fig. 2 are recast into k -space and the effective NA can be directly read out to be in accordance with the nominal one.

4.2. Synthetic APSF

The complex deconvolution process by the experimental CTF: c_{exp} ought to be compared to a reference system. This system is based on a synthetic CTF: c_{syn} transformed by Eq. (12) from a synthetic APSF. The APSF represents a synthesis since the following scalar Debye theory is computed with experimentally assessed parameters of the optical imaging system.

For a single point object $o(x_1, y_1) = \delta(x_1, y_1)$, the image field $U_\delta(x_2, y_2)$ is the APSF $h(Mx_2, My_2)$. Therefore, the APSF can be experimentally measured with a sufficiently small object diameter ($\varnothing \ll d_{min}$) or synthesized by a theoretical description. A synthetic h for high-aperture systems can be approximated by the scalar Debye theory expressed in a spherical coordinate system in Eq. (17) of θ and Φ within the object space

$$\begin{cases} x_1 = f \sin \theta \cos \phi, \\ y_1 = f \sin \theta \sin \phi, \\ z_1 = -f \cos \theta, \end{cases} \quad \text{which satisfies} \quad f^2 = x_1^2 + y_1^2 + z_1^2, \quad (17)$$

and thus a polar coordinate system in Eq. (18) within the image space originating in the focus

$$\begin{cases} x_2 = r_2 \cos \Psi, \\ y_2 = r_2 \sin \Psi, \\ z_2, \end{cases} \quad \text{which satisfies} \quad r_2^2 = x_2^2 + y_2^2. \quad (18)$$

The Debye integral results in [23]:

$$U_\delta(r_2, \Psi, z_2) = \frac{i}{\lambda} \int_0^{2\pi} \int_0^\alpha P(\theta, \phi) \exp[-ikr_2 \sin \theta \cos(\phi - \Psi) - ikz_2 \cos \theta - ik\Phi(\theta, \phi)] \sin \theta d\theta d\phi, \quad (19)$$

where $P(\theta, \phi)$ is the apodization function according to Eq. (20) and $\Phi(\theta, \phi)$ the aberration function. Generally, the sine condition holds for an aplanatic imaging system within the field of view

$$P(\theta, \phi) = \sqrt{\cos \theta}. \quad (20)$$

Aberrations in high aperture $\Phi(\theta, \phi)$ may be developed as spherical harmonics in a complete orthogonal set [28] and are included in our model for the primary aberrations. The 2D APSF affected by aberration can be calculated at a certain distance z_2 relative to the focal plane

$$h(z_2) = \iint_{\Omega} U_\delta(r_2, \Psi, z_2) dr_2 d\Psi. \quad (21)$$

Finally, the CTF can be synthesized as the 2-D Fourier transform of a synthesized 2D APSF.

In order to correctly model the aberration-affected system, the coefficients $A_{n,m}$ of the aberration function Φ have to be adapted in a fitting process. In this optimization process, each calculation of the synthetic APSF is performed by FFT of the pupil function as presented in [29]. For the critical angle of the imaging system of NA=0.95, the vectorial Debye model, also suggested in [29], would be more appropriate since it describes polarization dependent effects [23]. However, for the computation-intensive fitting procedure, the faster scalar model is chosen.

As already pointed out, aberrations will most prominently affect the phase of the CTF. Therefore, the fitting process is performed in the phase domain and compares the experimental data with the synthetic phase CTF in a least-square cost functional f :

$$f(A_{n,m}) = \sum_{k_x, k_y} |\arg[c_{exp}] - \arg[c_{syn}(A_{n,m})]|^2, \quad (22)$$

where the synthetic model c_{syn} is a function of the spherical harmonics with amplitude factor $A_{n,m}$ according to [28]. The global minimum of f is found by a genetic algorithm [30] and the fine fitting is performed by pattern research [31]. The results are listed in Table 1.

Table 1. Results of fit of experimental data from optical system at $\lambda=532\text{nm}$ and $\text{NA}=0.95$

aberration	spherical harmonics $A_{n,m}$	fitted amplitude [10^{-9}]
defocus	$A_{2,0}$	45
primary spherical aberration	$A_{4,0}$	350
secondary spherical aberration	$A_{6,0}$	-600
primary coma	$A_{3,1}$	400
primary astigmatism	$A_{2,2}$	25

The CTF synthesized with the parameters from Table 1 is shown in Figs. 2(b) and 2(d). The modulus spectra show the same transmittance values and the bandwidths' sizes agree. The most prominent difference consists in an asymmetry favoring negative k_x wavenumbers and therefore a direction dependent effect. The model does not include direction dependent intensity responses which would be suggested by a vectorial influence.

On the other side, the phase part of the fitted spectrum in Fig. 2(d) shows a good agreement with the experimental one in Fig. 2(c). Since the experimental phase CTF was fitted, the scalar model manages to create a similar wavefront which does not affect the modulus. In reality, though, the direction dependent transmittance, seen in Fig. 2(a), is connected to its phase in part (c) of Fig. 2. Consequently, the estimated spherical harmonics' prefactors $A_{n,m}$ are likely to be overestimated.

The synthetic CTF allows to model a synthetic test target based on Eq. (21) corresponding to the experimental test target depicted in Fig. 1:

$$U_{syn} = h_{syn}(x, y + \eta/2) + h_{syn}(x, y - \eta/2). \quad (23)$$

Most importantly, U_{syn} and c_{syn} serve as a reference system to determine the influence of noise on the deconvolution process. Without loss of generality, noise can be added to the synthetic APSF as a gaussian probability distribution n_g and yields the estimation of U_{noise} by computing Eq. (23).

4.3. Test target

The fields of the test target have been recorded for all η of Fig. 1 and processed according to the procedure described in section 4.1. As an illustration, Figs. 3(a) and 3(b) ($\eta=400\text{nm}$) show the complex field spectrum \tilde{G} calculated by Fourier transforming the reconstructed complex image field U .

The amplitude $|\tilde{G}|$ in Fig. 3(a) shows the image spectrum accompanied by two frequency filters (seen as minimum transmittance) in k_y direction. These minimum transmittance filters can be understood as spectral presentation of the destructive interference between the waves emitted by the two holes, reported as phase singularities in [32]. The discontinuities in $\arg[\tilde{G}]$ [cf. Fig. 3(b)] occurs at spatial positions where the spherical waves emitted from each hole are out of phase. As reported in [32] the orientation angle θ of those lines of singularities varies systematically with the pitch η of the two holes. Consequently, the $\eta - \theta$ relationship corresponds closely to a $\eta - k_y$ relationship meaning that the position of the transmittance minima varies again as a function of the pitch. Figure 3(c) illustrates the behavior within the bandwidth. It can be seen that the minimum transmittance position of the filter shifts to higher frequencies as η decreases. Finally, for $\eta=300\text{nm}$, the minimum almost reaches $k_{y,c}$ and $|\tilde{G}|$ barely features higher frequency content. The exact maximally possible shift of the minimum

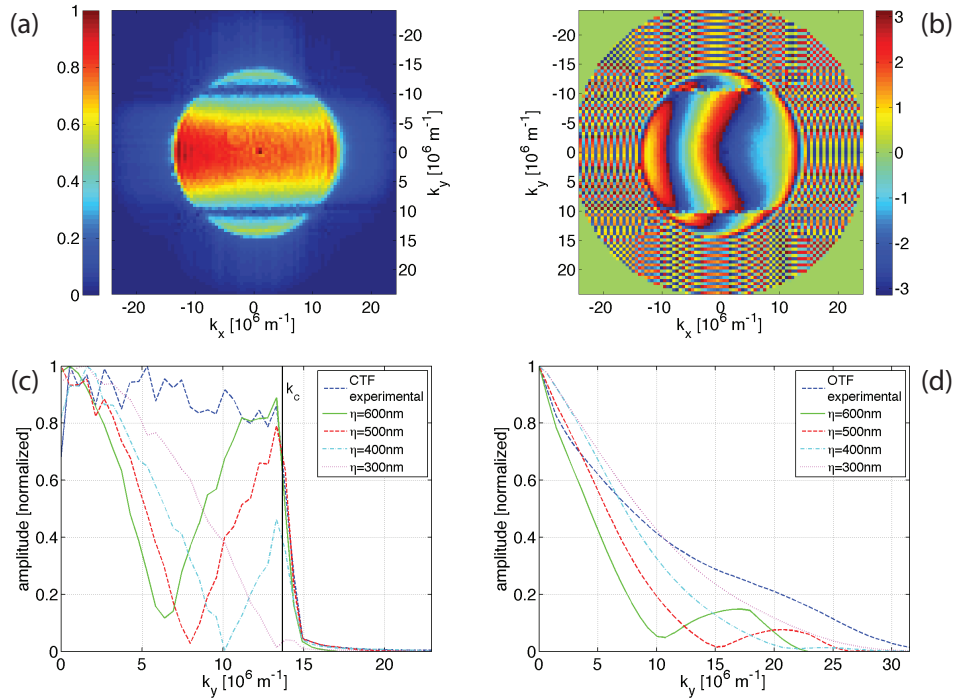


Fig. 3. Experimental transfer functions in focal plane at $\lambda=532\text{nm}$ and $\text{NA}=0.95$ of the test target (cf. Fig. 1). The Γ -masked amplitude spectrum $|\tilde{G}|$ (a) and phase spectrum $\arg[\tilde{G}]$ (b) are illustrated for $\eta=400\text{nm}$. (c) compares $|\tilde{G}|$ cross-sections in k_y for $k_x = 0$ with the experimental CTF of a single nano-metric hole. (d) shows the same comparison of the experimental OTF and $|\tilde{J}|$.

transmittance in Fig. 3(c) matches with the largest observable angle θ of the phase singularities. The corresponding limit of resolution is derived to be [32]

$$d_{min}^{cd} = \min \left[\frac{\lambda}{2} \left(\frac{1}{\text{NA}} \pm \frac{\Delta\phi}{\pi} \right) \right]. \quad (24)$$

Note that $\Delta\phi$ indicates an offset phase difference for instance through a longitudinal displacement $\Delta z = \Delta\phi\lambda / (2\pi)$. In the case of the used test target of Fig. 1 $\Delta\phi$ yields 0. As a consequence, the relevant d_{min}^{cd} corresponds to a minimal distance 64% smaller than suggested by Eq. (2) for the coherent case and still 24% superior to the corresponding equation for the incoherently illuminated case. In the event of $\Delta\phi \neq 0$, asymmetric phase singularities would appear in the spatial phase map [32]. As a consequence, the transmittances minima (respectively discontinuities in $\arg[\tilde{G}]$) shifts on one spectral side to lower and on the other spectral side to higher wavenumbers.

Similarly, the spectra $|\tilde{J}|$ are compared with the experimental OTF in Fig. 3(d). It is important to note that $|\tilde{J}|$ is calculated by Fourier transforming the intensity fields I of Eq. (25),

$$I(x, y) = |U(x, y)|^2, \quad (25)$$

albeit gained by coherent imaging system according to Eq. (15). That is to say that for the computation of $|\tilde{J}|$ the data are processed as if they originated from an incoherent imaging system as in Eq. (5). The purpose of this approximation is to compare the performance of this

'improper' intensity deconvolution to proper complex deconvolution, and in particular to erode the phase's role.

5. Results

5.1. Complex and intensity deconvolution of test targets

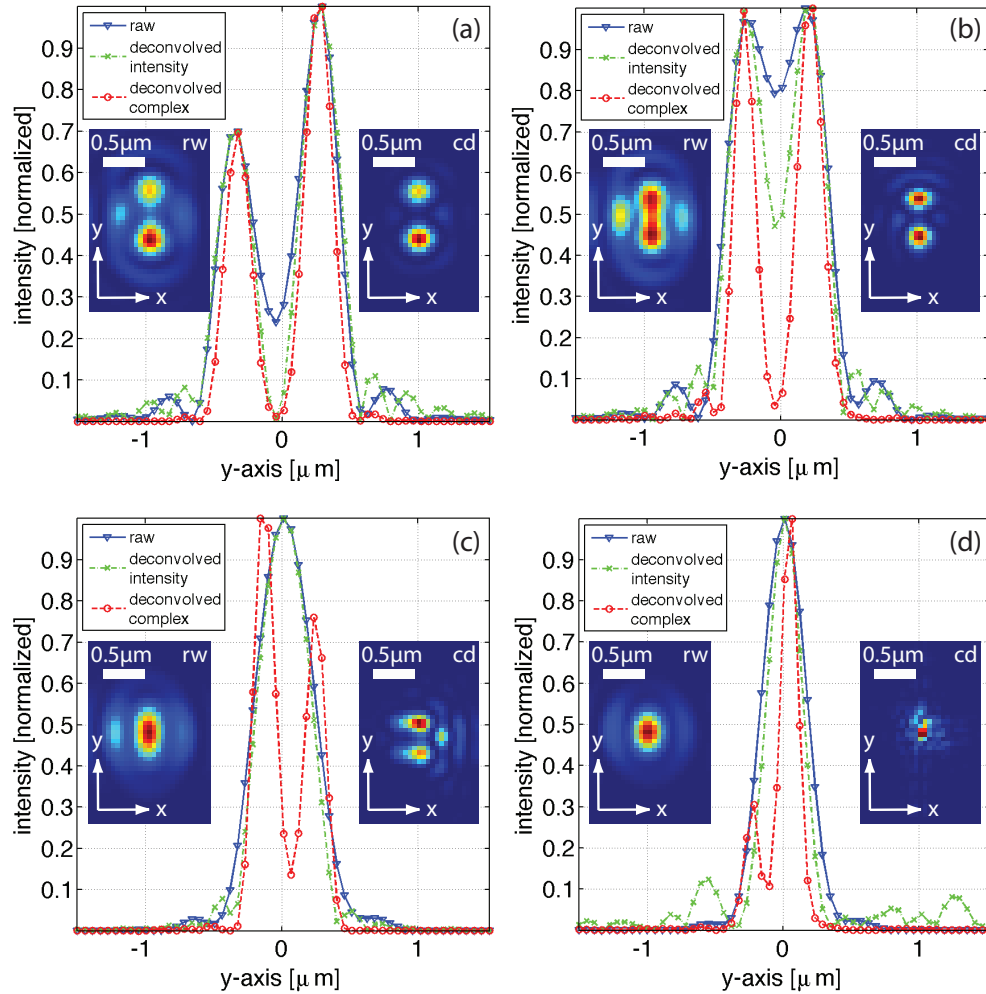


Fig. 4. Comparisons of unresolved and super-resolved profiles of two nano-holes of test target (cf. Fig. 1) with center-to-center distances $\eta=600\text{nm}$ in (a), $\eta=500\text{nm}$ in (b), $\eta=400\text{nm}$ in (c), and $\eta=300\text{nm}$ in (d). The raw data images I are reconstructed in the focal plane at $\lambda=532\text{nm}$ and $\text{NA}=0.95$ ($d_{\text{min,coh}}=460\text{nm}$). The 'raw' profile shows the central y cross-section of the resolution limited raw data I (cf. 'rw' insert). The 'deconvolution complex' profile shows the corresponding section of $|o|^2$ (cf. 'cd' insert) resulting from complex deconvolution by the experimental CTF. Additionally, 'deconvolution intensity' compares the profile of o_i resulting from intensity deconvolution by the experimental OTF.

The raw images show the intensities I of two PSF above [cf. Figs. 4(a)–4(b), inserts 'rw'] and beneath [cf. Figs. 4(c)–4(d), inserts 'rw'] the coherent limit of resolution $d_{\text{min,coh}}=460\text{nm}$. It can

Table 2. Results of peak-to-peak (p-t-p) distance measurements of the test target at $\lambda=532\text{nm}$ and $\text{NA}=0.95$. The standard precision is based on the lateral sampling of 56nm , the complex deconvolution is determined in Fig. 5(d)

SEM measurement	$(\eta \pm 5)[\text{nm}]$	600nm	498nm	403nm	304nm
raw	$(p-t-p \pm 56)[\text{nm}]$	629	448	–	–
	contrast [%]	60	22	–	–
	FWHM [nm]	230	285	–	–
intensity deconvolution [vide Eq. (8)]	$(p-t-p \pm 56)[\text{nm}]$	629	452	–	–
	contrast [%]	77	50	–	–
	FWHM [nm]	223	233	–	–
complex deconvolution [vide Eq. (13)]	$(p-t-p \pm 25)[\text{nm}]$	594	482	372	275
	contrast [%]	80	96	69	25
	FWHM [nm]	224	196	182	141

be seen that the PSF pairs beneath $d_{\min,coh}$ converge and cannot be distinguished anymore by the contrast criterion. The inserts labeled 'cd', though, show the intensity images $|o|^2$ complex deconvolved with k_{\max} (deduced in section 5.2). In the profile plot, I and $|o|^2$ are compared to intensity deconvolved images o_i along the y cross-sections through the pitch centers. The exact results are listed in Table 2.

It shows that both deconvolution methods manage to improve the contrast of the $\eta=500\text{nm}$ and $\eta=600\text{nm}$ image. The contrast is higher for complex deconvolution whilst holding a more accurate match on the actual pitch η . For the case of $\eta=400\text{nm}$ the intensity deconvolution fails to resolve individual peaks. Despite being beyond $d_{\min,coh}=460\text{nm}$, the coherent deconvolution method results in a correct localization within 25nm while holding a contrast of 69% . However, the deconvolved image suffers from a residual artifact which is caused by a mismatch during the DHM reconstruction procedure [8]. The final test target measurement of 300nm lies just at the edge of the new limit of resolution derived in Eq. (24). The peak is clearly observable in the cross-section, however, in the image itself the mentioned DHM reconstruction mismatch causes too many artifacts, so that the two peaks are not clearly recognizable any more.

5.2. Determination of k_{\max} and noise influence

The choice of correct mask radius k_{\max} is crucial to avoid cropping information or adding noise. The mask radius is related to Eq. (4) meaning that it can be chosen based on the minimal structure d_{\min} to be resolved which must fulfill at least $k_{\max} \leq k_c$. It is not trivial, though, to decide to which extend the filter's diameter can be enlarged.

In order to address this question, the fields of the test target were deconvolved while varying the filter's diameter k_{\max} [expressed as the according minimal resolvable distance d_{\min} by Eq. (4)]. Note that the smoothing is fixed to a small value $k_s = 2\pi/(d_{\min} - 60\text{nm})$. For the results analysis, the deconvolved amplitude fields (cf. Fig. 4) are interpolated and fitted by two gaussian curves:

$$g(x) = a_1 \exp\left[-\frac{(x-\mu_1)^2}{2b_1^2}\right] + a_2 \exp\left[-\frac{(x-\mu_2)^2}{2b_2^2}\right]. \quad (26)$$

The positions of μ_i provide the peak-to-peak (p-t-p) distance of the holes' images. Assuming equivalent transmittance of the imaged holes' pairs, the effective full width at half maximum

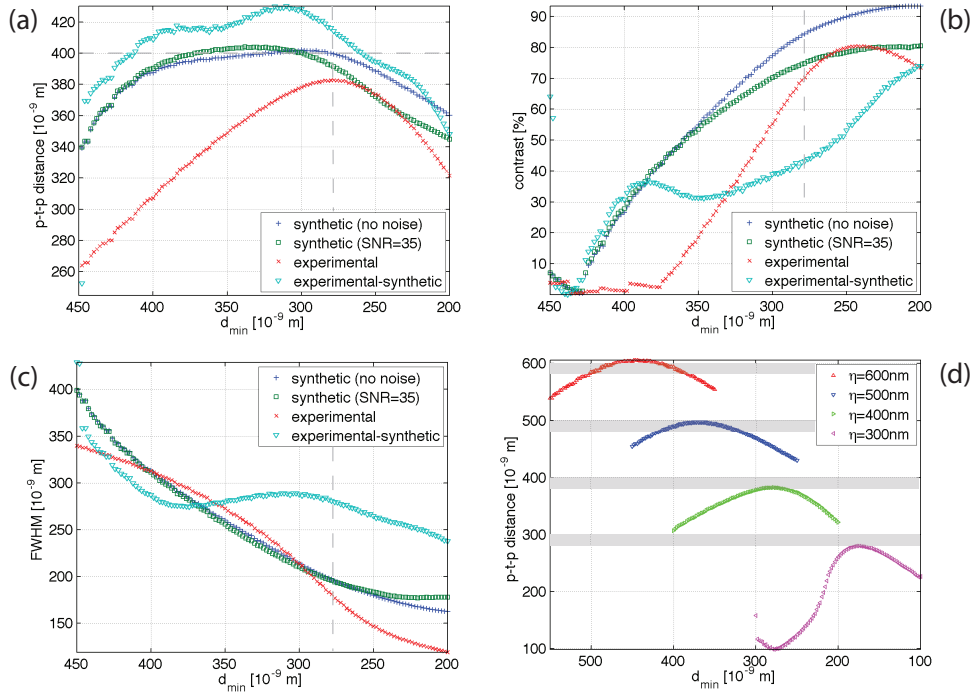


Fig. 5. Influence of $k_{max}(d_{min})$ on complex deconvolution results according to Eq. (4). (a-c) statistics for $\eta=400$ nm for deconvolution of U_{exp} with c_{exp} 'experimental', for deconvolution of U_{exp} with c_{syn} 'experimental-synthetic', for deconvolution of U_{syn} with c_{syn} 'synthetic (no noise)', and deconvolution of U_{noise} with c_{noise} 'synthetic (SNR=35)'. (d) statistics of p-t-p in dependence of d_{min} for all targets η . The grey bars indicate error margin of 25nm.

(FWHM) is averaged for b_1 and b_2 and determined from Eq. (26) as Eq. (27):

$$FWHM = \sqrt{2 \ln 2} (b_1 + b_2). \quad (27)$$

The contrast is calculated by the ratio of the minimum value between the two maxima values of a_1 and a_2 .

The results are shown in Figs. 5(a)–5(c) for the exemplary case of $\eta=400$ nm. The notation is as following: The legend 'experimental' indicates complex deconvolution of U_{exp} with c_{exp} . Contrarily, the legend 'synthetic(...)' indicates the usage of synthesized fields and CTFs according to section 4.2. For indicator '...(no noise)' Eqs. (19) and (23) are free of noise, whereas gaussian noise was added successively for the indicator '...(SNR=35)' with the according signal-to-noise ration (SNR). Finally, the case 'experimental-synthetic' represents a hybrid, the complex deconvolution of the experimental fields by the synthetic noise free CTF.

First of all, Fig. 5(a) demonstrates the impact of noise. The bending of the p-t-p curve indicates a dependence of the measured p-t-p distance on $k_{max}(d_{min})$. This trend is strongest for the 'experimental' plot. On the contrary, the fully 'synthetic (noise free)' deconvolution shows a weak dependence which suggests noise as a source of the dependence trend. A stronger bending of the p-t-p curve for smaller d_{min} can be created by adding noise to the 'synthetic(SNR=35)' deconvolution. Vice versa, the 'experimental' dependence becomes weaker for the 'experimental-synthetic' deconvolution but suffers from a vertical upward shift which may result from a mod-

eling mismatch of h_{syn} . Finally, the filter's radius dependence can be partially decoupled by using a noise free synthetic CTF, as expected for the use of synthetic OTF in intensity deconvolution [2, 7].

It is in principle desirable to choose the filters' radii as high as possible, in the limit of a 'correct' p-t-p distance. Figure 5(b) shows a general contrast trend of the deconvolved images in which all cases have in common: the higher the frequency content, the better the contrast. This observation is in accordance with Fig. 5(c) showing clearly the inverse trend of FWHM: peaks become narrower with bigger filter radii. For the 'experimental-synthetic' case the trend of FWHM and contrast is perturbed. At about $d_{min} \approx 380nm$, the trend of increasing contrast is damped [cf. Fig. 5(b)] since the FWHM's trend of narrowing stagnates [cf. Fig. 5(c)]. It is most likely that artifacts caused by the modeling mismatch lead to the (trend opposed) broadening.

In our case, we define a precision of $\pm 25nm$ (about half δx) of the p-t-p distance as acceptable in order to minimize the FWHM or, in other words, to maximize the contrast. This error margin is indicated in Fig. 5(d), which shows the p-t-p distances for the 4 'experimental' cases of the test target. According to the defined criterion, it can be seen that a unique filter diameter can be assessed. Moreover, a trend of k_{max} is clearly observable: the optimal filter diameter corresponds to wavenumbers corresponding roughly to $(150 \pm 30)nm$ beneath the minimal structure η to be observed. This trend responds to the hypothesis $k_{max} \approx k(\eta - 150nm)$ if $\eta < d_{min}$.

6. Discussion

In this section, effects and mechanisms of complex deconvolution are discussed and its limitations are under debate. The method is classified and the effectiveness of different approaches is compared. Finally, an outlook of generality is given.

At a first glance, the deconvolved phase spectrum [cf. Figs. 6(a) and 6(b)] appears more even than compared to the original spectrum [cf. Figs. 3(a) and 3(b)]. In particular, the phase in k_x direction indicates an aberration correction. The mean deconvolved spectrum transmittance drops down to ~ 0.6 , therefore appearing to wane. Division, in the frequency domain, by very small values of the CTF gives rise to large transmission values [cf. Fig. 6(a)]. Those 'bad' pixels can also be seen in the phase spectrum [cf. Fig. 6(b)] as local pixel phase jumps. Since these local pixel phase jumps occur randomly or close to weak signal strengths, they can be considered as an effect of noise on the CTF. Just as it is known from intensity deconvolution [2], the bigger k_{max} is chosen, in order to resolve smaller d_{min} , the better the SNR must be.

The origin of the gain in resolution is a stronger spectral support at high frequencies in the spectrum after deconvolution [cf. Figs. 6(a) and 6(b)]. The cross section in Fig. 6(c) compares the profiles of spectra before [cf. Figs. 3(a) and 3(b)] and after [cf. Figs. 6(a) and 6(b)] deconvolution. The same is shown for intensity deconvolution in Fig. 6(d). The comparison between Figs. 6(c) and 6(d) highlights that only the complex deconvolved spectral support in Fig. 6(c) is increased and frequencies above the cutoff are added giving rise to the gain in resolution. However, plot (c) of Fig. 6 is not a sufficient proof of improved resolution since already adding a constant frequency could lead to a similar result. The 'real' resolution improvement results from the frequency pattern in Fig. 6(a) which exhibit an accurately restored continuous frequency content. The reason for this behavior lies in the phase discontinuity seen in Fig. 6(b) that shifts the phase spectrum by π . This shift extends the spectrum continuously to higher frequencies.

Apparently, this mechanism fails for the intensity deconvolution [cf. Fig. 6(d)]. Moreover, the frequencies related to resolution exceeding the Rayleigh limit are located at higher frequencies thus making them more sensitive to noise. Vice versa, due to the shift toward lower frequencies, the complex spectrum is less sensitive to noise. Consequently, as long as the image spectrum contains sufficient spectral support higher than the minimal transmittance [cf. Figs. 3(c) and

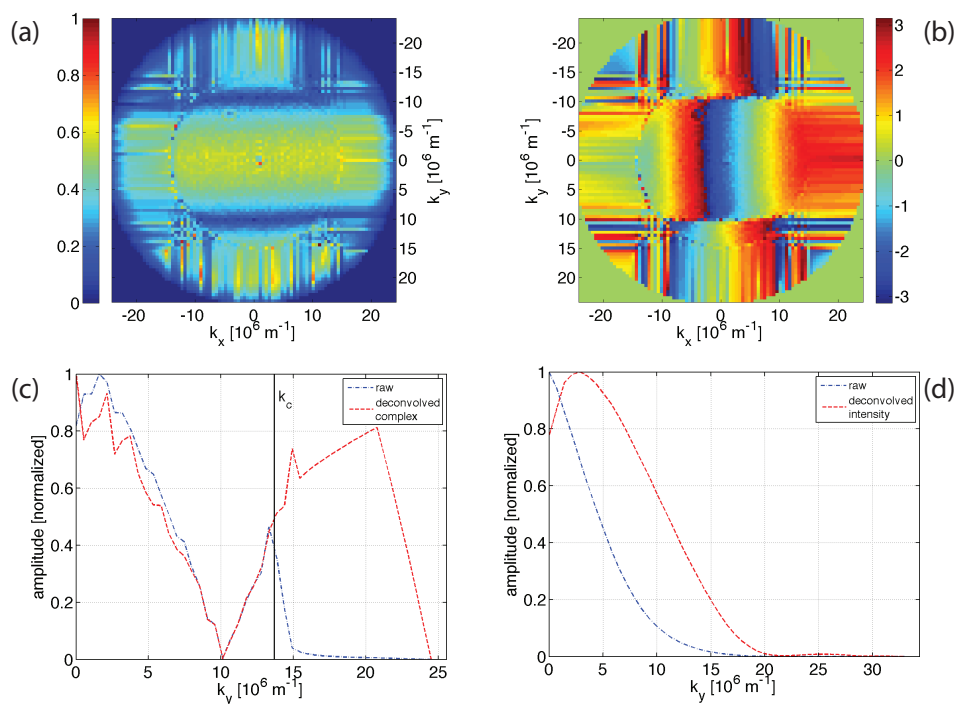


Fig. 6. Experimental transfer functions in focal plane for $\lambda=532\text{nm}$ and $\text{NA}=0.95$ of test target (cf. Fig. 1) after deconvolution. The amplitude spectrum $|O|$ (a) and phase spectrum $\arg[O]$ (b) are illustrated for $\eta=400\text{nm}$ after division by CTF. (c) compares $|O|$ cross-section in k_y for $k_x = 0$ with \tilde{G} for the $\eta=400\text{nm}$ case. (d) shows the same comparison for $|O_i|$ and $|\tilde{J}|$.

3(d)] the image is notably improved in resolution and contrast (cf. Fig. 4). For the $\eta=300\text{nm}$ case the higher frequencies are almost cut off. Only a 'small' peak is recovered which may be subject to noise and to model mismatch. The model sensitivity is reflected in the accuracy of the related p-t-p distance (see Table 2) which decreases towards $\eta=300\text{nm}$. In the same way, it holds for the decreasing contrast.

The effectivity of the complex deconvolution is summarized in Fig. 7. First of all, the merged PSF of the resolution limited system is shown in Fig. 7(b). Even if intensity deconvolution [cf. Fig. 7(c)] results in narrower FWHM, the resolution is not improved. On the other hand, complex deconvolution by the experimental CTF succeed in resolving the individual peaks [cf. Fig. 7(e)]. Complex deconvolution by a noise free synthetic CTF [cf. Fig. 7(d)] is capable of resolving them as well. The mismatch of the model with experimental data, however, results in an overestimation of η . In Fig. 7(d) it appears that the FWHMs tend to be very slim but artifacts cause a broadening and worse contrast as confirmed by Figs. 5(b) and 5(c). Thus, no real image improvement is achieved since the model mismatch causes severe artifacts. Finally, complex deconvolution by the experimental CTF holds the best match on η .

The customized test target of Fig. 1 has been specially designed to prove feasibility and to analyze the functionality of complex deconvolution. While the theory is derived from a general coherent imaging point of view, the used test target allows studying the practical role of sample internal coherent cross-talk as an essential part of the coherent imaging process. It is observed that the occurrence of discontinuities in the phase spectrum greatly supports resolu-

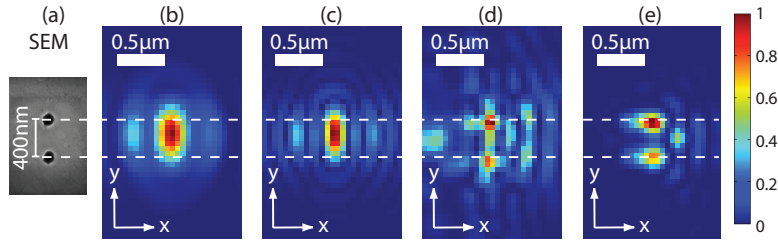


Fig. 7. XY images in focal plane of test target with sub-resolution pitch $\eta=400\text{nm}$ [cf. insert (a) imaged by SEM]. Insert (b) shows the unresolved test target's raw image I at $\lambda=532\text{nm}$ and $\text{NA}=0.95$. Insert (c) shows o_i resulting from intensity deconvolution. Insert (d) shows $|o|^2$ resulting from complex deconvolution by the synthetic CTF and insert (e) the according result for deconvolution by the experimental CTF.

tion improvements through complex deconvolution. In the sample specific case of $\Delta\phi = 0$, the limit of resolution is shown to be extended by a factor of 1.64 as anticipated by Eq. (24).

In a more general consideration, any non-transparent object could be thought to be composed of a three dimensional distribution of scatterers. For instance, through a longitudinal displacement Δz , an arbitrary phase offset difference $\Delta\phi \neq 0$ between point-scatterers would be created. As underlined in test target specific section 4.3, any $\Delta\phi$ would result in at least one spectrally lower shifted phase discontinuity. We suspect that it could give rise to even higher resolution, scaling accordingly to Eq. (24). Similarly, for the case of phase objects, the occurrence of phase singularities has been reported [33] which may also result in the appearance of spectral phase discontinuities and eventually image resolution improvements.

However, since the presented theory is not limited to spectral phase discontinuities, resolution beyond Eq. (24) is imaginable. The behavior of $\eta = 300\text{nm}$ sample's complex deconvolution suggests that the use of a noise free synthetic CTF should be advantageous since the SNR becomes much more crucial in the spectral sub-discontinuity range. As pointed out in section 4.2, a vectorial CTF [34] may be more suitable to effectively avoid the discussed artifacts. A direct calculation of its vectorial components [35] may allow a fast implementation.

7. Conclusion

We have discussed a general theory for complex deconvolution that excludes noise by truncating high spatial frequencies without any further assumptions on the noise source. The theoretical consideration is based on the assumption that each sub-Rayleigh object acts as a spherical wave emitter accordingly to Huygens' principle. In the classical imaging formalism for incoherent light, those emitters do not interact and are simply imaged as an ensemble of PSFs. Consequently, the intensity deconvolution mechanism works well as long as the interference interaction between the scatterers is small. Beyond that limitation, intensity deconvolution results in an incorrect evaluation of the scatterers' positions and their contrasts vanish.

Nonetheless, within the original passband of the imaging system, there exists information that originates from the objects' scattered light and lies outside the bandpass: the phase discontinuities. Coherent imaging bears the capability of recovering such intrinsic data derived from interferences in order to improve the resolution beyond the Rayleigh limit. This paper demonstrates experimentally the effectiveness of complex deconvolution for the developed test target of known structure. The results indicate that using complex deconvolution with an experimental CTF can increase resolution whilst localizing the objects within $\pm 25\text{nm}$. Moreover, image improvement for complex deconvolution using a synthetic CTF is, in principle, possible as well. However, a very exact adaptation of the model to the experimental data is crucial. The phase

fitted CTF allows to characterize the imaging system and demonstrates the noise's influence during the complex deconvolution.

In conclusion, the complex deconvolution method results in comparable or better resolution for coherent optical systems than normally achieved for incoherent optical system. The coherent optical system's lateral resolution is demonstrated to be almost doubled through complex deconvolution. It is a post-processing method that does not require any modification of the setup and is best suitable for methods providing complex fields such as DHM.

Acknowledgements

The authors want to acknowledge Swiss National Science Foundation (SNSF), grant #205 320-120 118, for its financial support and the Center of MicroNanoTechnology (CMI) for the cooperation on its research facilities.