

Phase retrieval with transverse translation diversity: a nonlinear optimization approach

Manuel Guizar-Sicairos and James R. Fienup

The Institute of Optics, University of Rochester, Rochester, New York, 14627, USA

mguizar@optics.rochester.edu, fienup@optics.rochester.edu

Abstract: We develop and test a nonlinear optimization algorithm for solving the problem of phase retrieval with transverse translation diversity, where the diverse far-field intensity measurements are taken after translating the object relative to a known illumination pattern. Analytical expressions for the gradient of a squared-error metric with respect to the object, illumination and translations allow joint optimization of the object and system parameters. This approach achieves superior reconstructions, with respect to a previously reported technique [H. M. L. Faulkner and J. M. Rodenburg, *Phys. Rev. Lett.* **93**, 023903 (2004)], when the system parameters are inaccurately known or in the presence of noise. Applicability of this method for samples that are smaller than the illumination pattern is explored.

© 2008 Optical Society of America

OCIS codes: (100.5070) Phase retrieval; (100.3010) Image reconstruction techniques; (110.3200) Inverse scattering; (110.7440) X-ray imaging; (340.7480) X-rays, soft x-rays, extreme ultraviolet (EUV).

References and links

1. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
2. J. R. Fienup, "Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint," *J. Opt. Soc. Am. A* **4**, 118–123 (1987).
3. P. S. Idell, J. R. Fienup and R. S. Goodman, "Image synthesis from nonimaged laser-speckle patterns," *Opt. Lett.* **12**, 858–860 (1987).
4. J. N. Cederquist, J. R. Fienup, J. C. Marron and R. G. Paxman, "Phase retrieval from experimental far-field speckle data," *Opt. Lett.* **13**, 619–621 (1988).
5. J. Miao, P. Charalambous, J. Kirz and D. Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature (London)* **400**, 342–344 (1999).
6. S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall and J. C. H. Spence, "X-ray image reconstruction from a diffraction pattern alone," *Phys. Rev. B* **68**, 140101 (2003).
7. J. R. Fienup, "Lensless coherent imaging by phase retrieval with an illumination pattern constraint," *Opt. Express* **14**, 498–508 (2006).
8. H. N. Chapman *et al.*, "High-resolution *ab initio* three-dimensional x-ray diffraction microscopy," *J. Opt. Soc. Am. A* **23**, 1179–1200 (2006).
9. M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with Fourier-weighted projections," *J. Opt. Soc. Am. A* **25**, 701–709 (2008).
10. E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *J. Opt. Soc. Am.* **52**, 1123–1130 (1962).
11. J. W. Goodman, *Introduction to Fourier Optics, 3rd Ed.* (Roberts & Company, Englewood, 2005).
12. O. Hellwig, S. Eisebitt, W. Eberhardt, W. F. Schlotter, J. Lüning and J. Stöhr, "Magnetic imaging with soft x-ray spectroholography," *J. Appl. Phys.* **99**, 08H307 (2006).
13. S. G. Podorov, K. M. Pavlov and D. M. Paganin, "A non-iterative reconstruction method for direct and unambiguous coherent diffractive imaging," *Opt. Express* **15**, 9954–9962 (2007).

14. M. Guizar-Sicairos and J. R. Fienup, "Holography with extended reference by autocorrelation linear differential operation," *Opt. Express* **15**, 17592–17612 (2007).
15. J. R. Fienup and C. C. Wackerman, "Phase-retrieval stagnation problems and solutions," *J. Opt. Soc. Am. A* **3**, 1897–1907 (1986).
16. H. Takajo, T. Takahashi, H. Kawanami and R. Ueda, "Numerical investigation of the iterative phase-retrieval stagnation problem: territories of convergence objects and holes in their boundaries," *J. Opt. Soc. Am. A* **14**, 3175–3187 (1997).
17. V. Elser, "Phase retrieval by iterated projections," *J. Opt. Soc. Am. A* **20**, 40–55 (2003).
18. H. H. Bauschke, P. L. Combettes and D. R. Luke, "Hybrid projection-reflection method for phase retrieval," *J. Opt. Soc. Am. A* **20**, 1025–1034 (2003).
19. H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm," *Phys. Rev. Lett.* **93**, 023903 (2004).
20. J. M. Rodenburg and H. M. L. Faulkner, "A phase retrieval algorithm for shifting illumination," *Appl. Phys. Lett.* **85**, 4795–4797 (2004).
21. J. M. Rodenburg, A. C. Hurst and A. G. Cullis, "Transmission microscopy without lenses for objects of unlimited size," *Ultramicroscopy* **107**, 227–231 (2007).
22. J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs and I. Johnson, "Hard-x-ray lensless imaging of extended objects," *Phys. Rev. Lett.* **98**, 034801 (2007).
23. H. M. L. Faulkner and J. M. Rodenburg, "Error tolerance of an iterative phase retrieval algorithm for moveable illumination microscopy," *Ultramicroscopy* **103**, 153–164 (2005).
24. R. A. Gonsalves and R. Childlaw, "Wavefront sensing by phase retrieval," *Proc. SPIE* **207**, 32–39 (1979).
25. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," *Opt. Eng.* **21**, 829–832 (1982).
26. R. G. Paxman and J. R. Fienup, "Optical misalignment sensing and image reconstruction using phase diversity," *J. Opt. Soc. Am. A* **5**, 914–923 (1988).
27. R. G. Paxman, T. J. Schulz and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," *J. Opt. Soc. Am. A* **9**, 1072–1085 (1992).
28. R. G. Paxman, "Diversity imaging," in *Signal Recovery and Synthesis*, 2001 OSA Technical Digest Series (Optical Society of America, 2001), paper SWA1.
29. G. R. Brady and J. R. Fienup, "Nonlinear optimization algorithm for retrieving the full complex pupil function," *Opt. Express* **14**, 474–486 (2006).
30. M. R. Bolcar and J. R. Fienup, "Method of phase diversity in multi-aperture systems utilizing individual sub-aperture control," *Proc. SPIE* **5896**, 58960G (2005).
31. B. J. Thelen, M. F. Reiley and R. G. Paxman, "Fine-resolution multispectral imaging using wavelength diversity," in *Signal Recovery and Synthesis*, Vol. 11 of 1995 OSA Technical Digest Series (Optical Society of America 1995), pp. 44–46, paper RTuD3.
32. H. R. Ingleby and D. R. McGaughey "Parallel multiframe blind deconvolution using wavelength diversity," *Proc. SPIE* **5562**, 58–64 (2004).
33. W. Hoppe, "Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden - und Phasenbestimmung bei unperiodischen Objekten," *Acta Crystallogr. Sect. A* **25**, 508–514 (1969).
34. J. M. Rodenburg and R. H. T. Bates, "The theory of super-resolution electron microscopy via Wigner-distribution deconvolution," *Phil. Trans. R. Soc. Lond. A* **339**, 521–553 (1992).
35. N. Nakajima, "Phase retrieval from Fresnel zone intensity measurements by use of Gaussian filtering," *Appl. Opt.* **37**, 6219–6226 (1998).
36. S. T. Thurman, The Institute of Optics, University of Rochester, Rochester, New York, USA, R. T. DeRosa and J. R. Fienup are preparing a manuscript to be called "Amplitude metrics for field retrieval with hard-edge and uniformly-illuminated apertures."
37. J. R. Fienup, "Invariant error metrics for image reconstruction," *Appl. Opt.* **36**, 8352–8357 (1997).
38. M. Guizar-Sicairos, S. T. Thurman and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.* **33**, 156–158 (2008).
39. Image modified from an original courtesy of Universidad Autónoma de Nuevo León, Mexico.

1. Introduction

Image reconstruction by phase retrieval, also referred to as coherent diffraction (or diffractive) imaging (CDI), is a lensless imaging technique in which conventional image forming optics (e.g. lenses, mirrors or holographic optical elements) are substituted by computational image reconstruction [1–9]. In this approach a coherent optical beam is incident on a reflective or transmissive object and the intensity of the field diffracted by the object is collected with an intensity detector array at some distance (usually in the far field). For this approach to work, we need an illuminating beam with a transverse coherence that is at least as large as the object is

wide and the pixel pitch of the detector array should be fine enough that the measured intensity is Nyquist sampled. These requirements place an upper bound on the maximum object extent (field of view) that can be imaged with this technique.

In this configuration the measured far-field intensity is used, along with some *a priori* information about the object (object support constraint and/or nonnegativity constraint), to computationally retrieve the phase of the field at the detector plane by using an iterative phase retrieval algorithm [1].

Due to the difficulty in fabricating diffractive optical elements (for conventional imaging) or unresolved point sources (for Fourier transform holographic reconstruction [10–12]) with sufficient precision for the x-ray regime, there has been an increasing interest in the development and application of phase retrieval to coherent x-ray imaging [5, 6, 8]. For CDI the final image resolution depends on the largest scattered angle collected with a moderately good signal-to-noise ratio (SNR) and the wavelength of the illuminating beam. Furthermore, the sampling and transverse coherence requirement for CDI is half of that required for holographic reconstructions, either with a point-source reference or a boundary-wave reference from an extended object [13, 14], because these techniques require measurement of the interference of the field diffracted from the object with that originating from a reference point source or sharp feature that is adequately separated from the object (holographic separation condition [11, 14]).

Although algorithms that are robust and capable of escaping local minima have been proposed and successfully used [1–9, 15–18], image reconstruction by phase retrieval still may suffer from stagnation modes that can be very persistent. The reconstruction becomes even more difficult when the image is complex-valued [2, 7] (as occurs in many important applications of coherent x-ray and optical imaging, and wavefront sensing), in which case a nonnegativity constraint cannot be used.

The use of multiple diverse measurements can make phase retrieval more robust. A very practical form of diversity for CDI, along with a robust reconstruction algorithm, was introduced by Faulkner and Rodenburg [19, 20] and was experimentally demonstrated for the optical and x-ray regime [21, 22]. This form of diversity, applicable when the sample can be illuminated repeatedly without significant change or damage, consists of taking a sequence of far-field intensity patterns, for which the object is displaced transversely relative to an illumination pattern that is known *a priori*. It was shown that an iterative phase retrieval technique, named ptychographical iterative engine (PIE), could be used in this case to increase the field of view (FOV) of the reconstruction, make the algorithms more robust, and achieve superior reconstructions. The success of this technique relies, however, on an accurate knowledge of the illumination pattern and the transverse displacements of the object. In practice, these parameters might be inaccurately known due, for example, to the limited precision of the translating stages and inaccurate or incomplete knowledge of features of the aperture (or focusing optics) that generates the illuminating beam [23].

In this paper we develop and test a nonlinear optimization routine to solve the phase retrieval problem with transverse translation diversity. By using a gradient-descent-based algorithm, we jointly optimize over the object, the illumination beam and the translation parameters, so that not only is the object estimation improved but also our initial estimate of the illumination pattern and the translations are refined. We provide analytic expressions for the gradients of a squared-error metric with respect to the object, illumination and translation parameters, that are crucial for efficient implementation of the algorithm.

The nonlinear optimization approach produces results superior to those of the PIE when the illumination pattern and translations are known inaccurately. Furthermore, in contrast to the PIE, which uses the measured intensity patterns sequentially, our technique uses all the measurements simultaneously when refining the object estimate, thus achieving reconstructions

with reduced noise artifacts.

Whereas PIE was only shown to work for objects larger than the diameter of the illumination pattern, we show that this form of diversity also provides very robust reconstructions (compared with single-measurement phase retrieval) when the object of interest is smaller than the illumination pattern. In this case the diverse measurements are obtained by moving the object within the transverse extent of the illumination pattern.

2. Object translation diversity

Algorithms that overcome some of the common stagnation problems and ambiguities of phase retrieval by introducing diversity have been developed for optical metrology applications and telescope alignment. They rely on measuring different intensity patterns, after introducing a known modification to the optical setup. A known amount of defocus is typically introduced by moving the detector along the optical axis [24–29]. Other forms of diversity that have been successfully used are sub-aperture piston diversity [30] and wavelength diversity [24, 25, 28, 31, 32].

The set of measurements, along with the knowledge of the diversity function, are used to find a field that agrees with all the measurements by using some form of iterative algorithm. Stagnation problems are overcome by providing a set of measurements that more robustly constrain the phase retrieval problem.

In x-ray CDI, the intensity measurement is typically taken in the far-field regime, so that the field intensity is proportional to the squared-magnitude of the Fourier transform of the object [11]. In the far-field regime, a longitudinal shift of the detector (an increased propagation distance) will only change the measured data by a multiplicative constant and a transverse magnification; this will not provide suitably diverse measurements.

For CDI, the intensity measurements are taken at a prescribed distance along the optical axis, and are sampled in the transverse coordinates (u, v) . The phase retrieval problem then reduces to finding the object field, $f(x, y)$, given a sampled measurement of the Fourier intensity, $I(u, v) = |F(u, v)|^2$. We approximate the continuous Fourier transform by the discrete Fourier transform,

$$F(u, v) = |F(u, v)| \exp[i\theta(u, v)] = \text{DFT}\{f(x, y)\} = \frac{1}{\sqrt{MN}} \sum_{x,y} f(x, y) \exp\left[-i2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)\right], \quad (1)$$

computed using the fast Fourier transform (FFT) algorithm, where N and M are the array dimensions, (x, y) and (u, v) are integer pixel coordinates in object and Fourier domain respectively, and we have dropped unimportant multiplicative scaling factors. Throughout the remainder of this manuscript, upper-case letters will refer to the discrete Fourier transforms of their lower-case counterparts as given by Eq. (1).

Introducing a known transverse translation of the object with respect to a known illumination pattern is a practical way of introducing diversity into the phase retrieval problem for x-ray CDI. An early form of this approach, developed originally for scanning transmission electron microscopy [33] and referred to as ptychography, required shifting the specimen by small amounts to many positions along a Cartesian grid. Although a non-iterative reconstruction procedure can be performed, this approach required obtaining and processing of very large four-dimensional data sets because the translation shift spacing is required to be of the same scale as the final spatial resolution of the reconstruction [34].

Nakajima proposed a reconstruction method that requires three far-field intensity measurements taken with the object translated with respect to a Gaussian illumination pattern [35], thus dramatically reducing the amount of data that needs to be collected and processed. For this approach the translations do not need to be on the scale of the final image resolution. However, the method suffers from severe disadvantages for x-ray CDI. Although for optical wavelengths

Gaussian beams appear naturally as optical resonator modes, this is not the case for coherent x-ray beams, where the beam must go through gratings and small apertures to achieve temporal and spatial coherence, and significantly departs from the Gaussian form. Furthermore, since the reconstruction method is based on the solution of simultaneous equations and is performed on a row-by-row basis, the results are found to be very sensitive to noise in the measured data.

Phase retrieval with translation diversity was brought to fruition by Faulkner and Rodenburg with the introduction of a robust iterative algorithm [19–22]. Their technique is capable of reconstructing a complex-valued image using an arbitrary illumination pattern and any number of translation-diverse far-field measurements, with a resolution that is only limited by the largest scattered angle collected with a moderately good SNR and the wavelength of the illuminating beam.

Following the work by Faulkner and Rodenburg, we can state the translation-diverse phase retrieval problem as follows. If we denote $o(x,y)$ and $p(x,y)$ as the object transmittance and illumination fields, respectively, the n -th resultant field is then given by

$$f_n(x,y) = o(x-x_n, y-y_n)p(x,y), \quad (2)$$

provided that the specimen is thin. Within the paraxial approximation, the collected measurements in the far field are

$$I_n(u,v) = |F_n(u,v)|^2. \quad (3)$$

The problem solved by Faulkner and Rodenburg is then to find a single object field $o(x,y)$ that generates the fields $f_n(x,y)$ that agree with all the intensity measurements $I_n(u,v)$, where the illumination $p(x,y)$ and the coordinate translations (x_n, y_n) are assumed known *a priori*.

It is important to note that although we refer to $p(x,y)$ as an illumination function, this formulation does not put any constraints on the form or nature of $p(x,y)$ so that it may be given by either a complex-valued field that originates from the diffraction of an aperture, a focusing lens, or be the transmittance of an aperture itself.

As was already discussed in [19], if the fields, $f_n(x,y)$, have no overlap with one another, the task is reduced to solving the conventional phase retrieval problem several times with no diversity. It is then important to have a substantial amount of overlap between the different field realizations. This constrains the problem very robustly and dramatically increases the success of phase retrieval. It is also worth noting that because a translation in all fields will produce a linear phase in the Fourier domain that cannot be detected, the technique is not sensitive to the absolute position of either the object or the illumination function. This implies that only the relative displacements of the object and illumination are important, and the technique is insensitive to whether it is the object or the illumination that moves [23]. An additional advantage is that a reconstruction of an extended FOV can be recovered with this technique.

Having translation diverse measurements can be shown to overcome two of the ambiguity problems commonly associated with phase retrieval: defocus and twin image.

The defocus ambiguity is exclusive of complex-valued image reconstruction (when a non-negativity constraint cannot be used). If only a support constraint is available for reconstruction, and the estimation of the support is larger than the object support, then a reconstruction of the object or a slightly out of focus image could fit inside the support and satisfy the far-field intensity constraint [8]. This results in a blurred image of the original object. This ambiguity is removed if translation diversity is used because there is only one plane where the translating object and illumination fields can be expressed as a product of one another.

The twin-image problem [15] is a well-known inherent ambiguity in CDI: both $f(x,y)$ and $f^*(-x,-y)$ [where $(*)$ denotes complex conjugation] have the same Fourier intensity pattern since

$$F^*(u,v) = |F(u,v)| \exp[-i\theta(u,v)] = \text{DFT} \{f^*(-x,-y)\}, \quad (4)$$

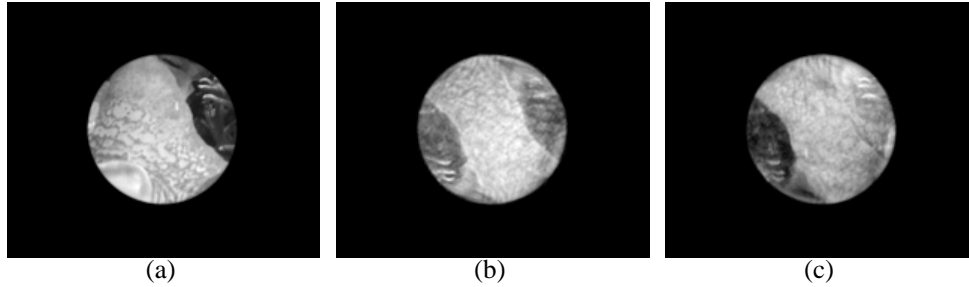


Fig. 1. Reconstruction examples using single-measurement phase retrieval with a support constraint. (a) Amplitude of object-space field, $f(x,y)$. The amplitude of two reconstructions that exhibit the twin-image problem, to different extents, are shown in (b) and (c), their final support errors are $E_S = 0.0162$ and $E_S = 0.0092$, respectively. A 172×210 portion of the 256×256 array is shown in (a)-(c).

so if the support constraint is symmetric about the origin, either one is a valid solution, or in other words, either the far-field phase or its negative could be retrieved. This inherent ambiguity can lead to a very persistent stagnation problem that commonly appears when reconstructing objects with a centrosymmetric support. The problem is characterized by the simultaneous appearance of features from both $f(x,y)$ and $f^*(-x,-y)$ inside the support. The fact that this stagnation mode has more energy outside the support than the actual solution identifies it as a local minima stagnation mode.

Figures 1(b) and 1(c) show reconstruction examples that exhibit the aforementioned twin-image problem to different extents. The 256×256 complex-valued field $f(x,y)$, amplitude shown in Fig. 1(a), was obtained by multiplying an extended object [amplitude and phase shown in Figs. 2(a) and 2(b), respectively] by a real-valued circular aperture of 50 pixel radius, shown in Fig. 2(c). The simulated Fourier intensity data, $I(u,v)$, was obtained by Fourier transforming the field, $f(x,y)$, and taking the squared modulus of the result. Making the extent of the field, $f(x,y)$, less than half of the entire array ensures that the intensity will be Nyquist sampled. The square-root of the Fourier intensity was then used along with the object support (support constraint) as constraints for the phase retrieval algorithm. Because the image is complex-valued, a nonnegativity constraint cannot be used. A sequence of 45 iterations of the hybrid input-output algorithm (HIO) [1], with a feedback parameter of $\beta = 0.7$, followed by 5 of the error-reduction algorithm were used, totaling 1000 iterations (20 sets of 50 iterations each). Because the support is centrosymmetric and the image is complex-valued, this reconstruction is particularly difficult for single-measurement phase retrieval.

This procedure was repeated for 10 different random starting guesses, thus obtaining 10 independent reconstructions. Three of the 10 reconstructions exhibited a very pronounced twin-image problem, one of which is shown in Fig. 1(b), where the reconstruction shows a high symmetry through the origin that is not present in $f(x,y)$. Five of the 10 reconstructions exhibited a mild twin-image problem, such as that shown in Fig. 1(c). Notice that although the twin image dominates the reconstruction shown in Fig. 1(c), and the image is recognizable, faint features of the upright image are also discernable, for example the three coral branches on the upper-right part of the image. For the remaining two reconstructions, a faithful image of the object shown in Fig. 1(a) was obtained, with no discernable twin-image problem.

Convergence of the iterative transform algorithm with a support constraint is typically mon-

itored by computing the support error, given by

$$E_S^2 = \frac{\sum_{(x,y) \notin S} |\hat{f}(x,y)|^2}{\sum_{x,y} |\hat{f}(x,y)|^2}, \quad (5)$$

where S is the set of points where the reconstruction, $\hat{f}(x,y)$, satisfies the support constraint. The two aforementioned reconstructions with no visually discernable twin-image problem had a final support error of $E_S = 0.0026$ and $E_S = 0.0060$, respectively. For the five reconstructions exhibiting a mild twin-image problem, $E_S = 0.0034, 0.0074, 0.0076, 0.0092, 0.0105$, respectively. Notice that on average E_S is an indicator of a mild twin-image stagnation problem. Finally, for the three reconstructions with a pronounced twin-image problem, the final support error was $E_S = 0.0150, 0.0162, \text{ and } 0.0192$, respectively. A severe twin-image problem will have large value of E_S , which identifies it as a stagnated partially-reconstructed image. If the final value of E_S for a given reconstruction is high, then the reconstruction should be started again with a different starting guess. Alternatively there are techniques that can be used to escape stagnation or combine different reconstructions to arrive at an improved composite that can be used for further iterations [15], these techniques are, however, outside of the scope of this paper.

This twin-image problem is removed by translation diversity by the knowledge of the transverse shifts that were imposed on the object. Although $f_n(x,y) = o(x-x_n, y-y_n)p(x,y)$ and $f_n^*(-x, -y) = o^*(-x-x_n, -y-y_n)p^*(-x, -y)$ have the same Fourier magnitude, we can differentiate them from the sign of the object shift, (x_n, y_n) , between two diversity images if the two fields overlap. Thus translation diversity eliminates the ambiguity between the image and its twin, making the simultaneous appearance of features of the regular and twin image in the reconstruction much less likely.

3. Ptychographical iterative engine

The PIE approach [19–21], is to iteratively correct the object field function $o(x,y)$ to sequentially match all of the Fourier intensity measurements. For a single iteration, the current object estimate, $\hat{o}_n(x,y)$, is multiplied by a translated version of the known illumination function, to obtain an estimate of the n -th diversity field,

$$\hat{h}_n(x,y) = \hat{o}_n(x,y)p(x+x_n, y+y_n) = \hat{f}_n(x+x_n, y+y_n), \quad (6)$$

then the Fourier transform of $\hat{h}_n(x,y)$ is computed:

$$\hat{H}_n(u,v) = |\hat{H}_n(u,v)| \exp [i\hat{\phi}_n(u,v)] = \text{DFT} \{ \hat{h}_n(x,y) \}, \quad (7)$$

then the measured Fourier amplitude (the square root of the measured intensity for the n -th diversity position) is imposed while preserving the phase

$$G_n(u,v) = \sqrt{I_n(u,v)} \exp [i\hat{\phi}_n(u,v)]. \quad (8)$$

The inverse Fourier transform of $G_n(u,v)$ is used to update the object estimate by

$$\hat{o}_{n+1}(x,y) = \hat{o}_n(x,y) + \frac{|p(x+x_n, y+y_n)|}{\max_{x,y} |p(x,y)|} \frac{p^*(x+x_n, y+y_n)}{|p(x+x_n, y+y_n)|^2 + \alpha} \beta [g_n(x,y) - \hat{h}_n(x,y)], \quad (9)$$

where β is a parameter that will be shown to be related to the step size of a steepest descent search and α is a constant that prevents numerical problems in dividing out the aperture function

when $p(x,y) \simeq 0$. A single iteration of this algorithm is achieved by repeating the outlined procedure until all of the diversity measurements, $I_n(u,v)$, have been used to update the object estimate, *i.e.* $n = 1, 2, \dots, q$, for q diversity images. An important feature of this procedure is that it sequentially uses the diversity images to update the current estimate for the object field. Notice that in this definition a single iteration requires using all q measurements.

The relation of the PIE algorithm and a steepest descent search can be more readily seen by multiplying Eq. (9) by $p(x+x_n, y+y_n)$,

$$\begin{aligned} \hat{f}_{n+1}(x+x_n, y+y_n) - \hat{f}_n(x+x_n, y+y_n) \\ = \frac{|p(x+x_n, y+y_n)|}{\max_{x,y} |p(x,y)|} \beta [g_n(x,y) - \hat{f}_n(x+x_n, y+y_n)], \end{aligned} \quad (10)$$

where we used Eq. (6) and

$$\frac{|p(x+x_n, y+y_n)|^2}{|p(x+x_n, y+y_n)|^2 + \alpha} \simeq 1. \quad (11)$$

Following a procedure analogous to that outlined in [1], we can compute the direction of steepest descent for a squared-error metric for $\sqrt{I_n(u,v)}$ with respect to $\hat{f}(x+x_n, y+y_n)$, given by

$$\begin{aligned} - \left[\frac{\partial}{\partial \hat{f}_n^R(x+x_n, y+y_n)} + i \frac{\partial}{\partial \hat{f}_n^I(x+x_n, y+y_n)} \right] \left\{ \sum_{u,v} \left[|\hat{F}_n(u,v)| - \sqrt{I_n(u,v)} \right]^2 \right\} \\ = 2 [g_n(x,y) - \hat{f}_n(x+x_n, y+y_n)], \end{aligned} \quad (12)$$

where the superscripts R and I indicate the real and imaginary part respectively ($\hat{f} = \hat{f}^R + i\hat{f}^I$).

Upon comparing the right hand side of Eq. (10) with the direction of steepest descent, given by Eq. (12), it becomes evident that the update step of the PIE algorithm, Eq. (9), can be interpreted as a steepest descent algorithm with a spatially variant step size given by

$$\frac{|p(x+x_n, y+y_n)|}{\max_{x,y} |p(x,y)|} \beta. \quad (13)$$

Notice that the step is proportional to the (normalized) illumination function, thus providing a longer step where the beam amplitude was higher and reducing the step size on pixels that were poorly illuminated and might then have a lower SNR, or not illuminated at all. Since other gradient search algorithms, *e.g.* conjugate gradient, are superior to steepest descent, this result indicates that it should be possible to improve on the PIE.

Figure 2 shows a reconstruction using the PIE algorithm. A 256×256 complex-valued object, whose amplitude and phase are shown in Figs. 2(a) and 2(b) respectively, was multiplied by a real-valued aperture $p(x,y)$ with a 50 pixel radius, shown in Fig. 2(c). The resulting field was then Fourier transformed and the squared-magnitude of the result was used as the measured data. This procedure was repeated for 12 different positions of the aperture, (x_n, y_n) , indicated by circles in Fig. 2(d), thus obtaining twelve two-dimensional far-field intensity measurements, $I_n(u,v)$. The aperture was shifted in intervals of 14 and 35 pixels in the x and y directions, respectively.

The PIE algorithm, with $\beta = 1$ and $\alpha = 0.01$, was used for reconstruction of $o(x,y)$ assuming perfect knowledge of $p(x,y)$ and (x_n, y_n) . The algorithm was initialized with $\hat{o}(x,y)$ as a real-valued, uniform function. The amplitude and phase of the reconstruction after 200 iterations are shown in Figs. 2(e) and 2(f), respectively. Notice that a very good estimation of the original object is obtained within the support sampled by $p(x,y)$.

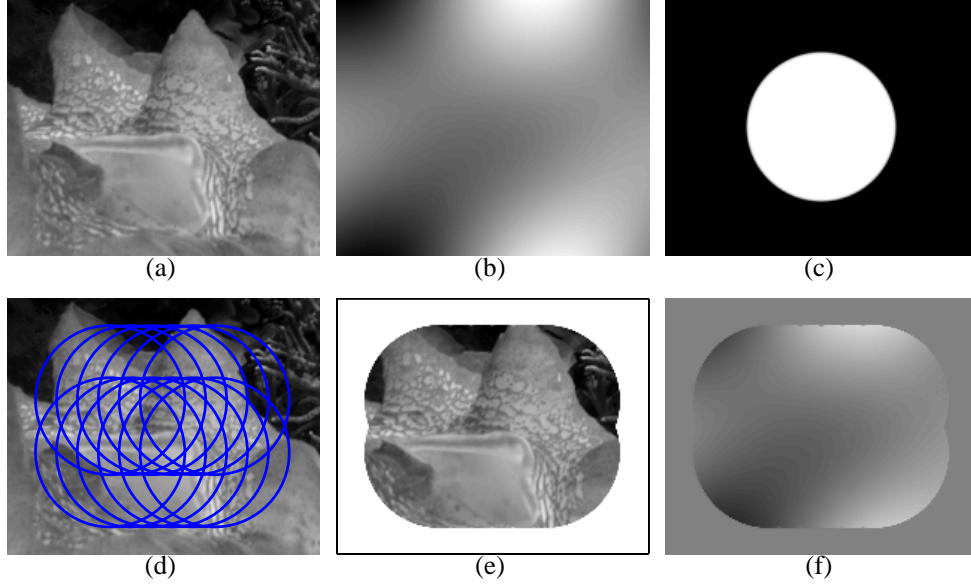


Fig. 2. Object, $o(x,y)$, (a) amplitude and (b) phase. (c) Amplitude of $p(x,y)$. (d) Circles indicate the twelve positions of $p(x,y)$ that were used for the reconstruction. (e) Amplitude and (f) phase of the reconstruction using the PIE. Phase is shown from -0.4 to 0.4 radians in (b) and (f). A 172×210 portion of the 256×256 array is shown in (a)-(f).

In agreement with Faulkner and Rodenburg, we have found this approach to be superior to single-measurement phase retrieval in terms of robustness and convergence speed. Nevertheless it still relies on accurate knowledge of the object displacements (x_n, y_n) and the aperture function $p(x,y)$.

To illustrate the effect of errors in the knowledge of system parameters, an additional numerical simulation was performed where we introduced errors in the *a priori* known information, *i.e.* $p(x,y)$ and (x_n, y_n) . The error in the translation parameters can be characterized by the global-shift-invariant root-mean-squared error (RMSE) of the translations, Δr , defined by

$$(\Delta r)^2 = \min_{a,b} \frac{1}{q} \sum_{n=1}^q (x_n - \hat{x}_n - a)^2 + (y_n - \hat{y}_n - b)^2, \quad (14)$$

where (x_n, y_n) and (\hat{x}_n, \hat{y}_n) are the true and estimated object translations, respectively, for the n -th diversity image. We introduced minimization with respect to global additive constants (a and b) to account for the fact that the reconstruction is unaffected by a global displacement of all the fields. In other words, since the reconstruction procedure is not affected by the absolute position of the object (only the translations are important), then addition of a constant to every displacement (in either x or y) should not increase our measure of the shift error. Upon minimizing Δr with respect to a and b we find

$$(\Delta r)^2 = \langle (x_n - \hat{x}_n)^2 \rangle - \langle (x_n - \hat{x}_n) \rangle^2 + \langle (y_n - \hat{y}_n)^2 \rangle - \langle (y_n - \hat{y}_n) \rangle^2, \quad (15)$$

where

$$\langle x_n \rangle = \frac{1}{q} \sum_{n=1}^q x_n. \quad (16)$$

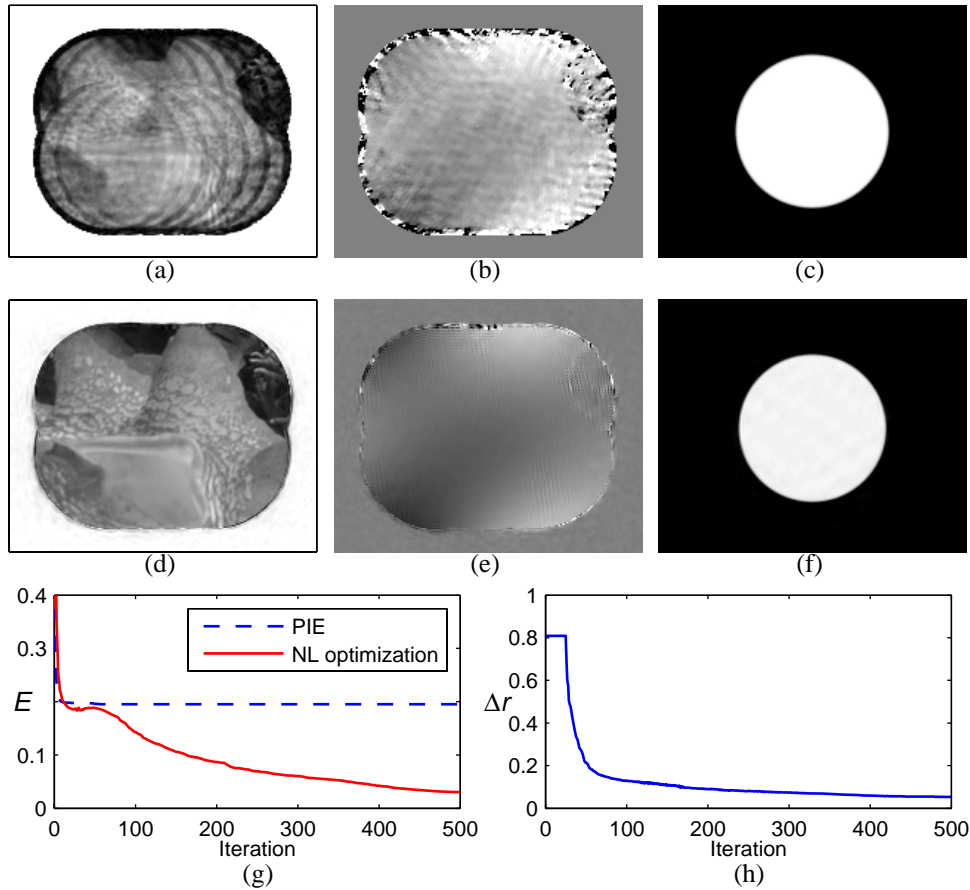


Fig. 3. Reconstruction results. Amplitude and phase after 500 iterations are shown in (a) and (b) for the PIE and (d) and (e) for the nonlinear optimization approach. (c) and (f) show the amplitude of the initial estimate of $p(x,y)$ and the result after nonlinear optimization respectively. (g) Normalized invariant error metric, E , vs. iteration number. (h) Shift error, Δr vs. iteration number for the nonlinear optimization algorithm. Phase is shown from -0.4 to 0.4 radians in (b) and (e). A 172×210 portion of the 256×256 array is shown in (a)-(f).

During the reconstruction the aperture was assumed to have a radius of 52 pixels (instead of 50), as shown in Fig. 3(c). Notice that an error of this nature would occur not only from an inaccurate measurement of the aperture, but also from an inaccurate knowledge of the distance from the sample to the detector, since the sampling in object domain scales proportionally to the propagation distance. Additionally, zero-mean Gaussian-distributed noise was added to the actual translations of the object, thus yielding a shift error of $\Delta r = 0.8$ pixel. The amplitude and phase of the reconstruction with 500 iterations of the PIE algorithm is shown in Figs. 3(a) and 3(b), respectively. Notice that the quality of the image has degraded significantly.

4. Nonlinear optimization approach

For phase retrieval techniques that use diversity, where the problem is robustly constrained, nonlinear optimization algorithms have been shown to be a valuable tool for reconstruction. The principal benefit of this approach is that we can straightforwardly include any form of non-

ideality that may affect the measurements in our forward model [36] (inaccurate knowledge of *a priori* information, finite pixel size, transverse coherence, detector misalignments) and optimize over inaccurately known parameters as well, provided that the gradients of the error metric with respect to the unknown parameters can be computed efficiently.

For the problem at hand we define the squared-error metric

$$\varepsilon = \sum_{n=1}^q \sum_{u,v} W_n(u,v) \left\{ \left[|\hat{F}_n(u,v)|^2 + \delta \right]^\gamma - \left[I_n(u,v) + \delta \right]^\gamma \right\}^2, \quad (17)$$

where

$$\hat{F}_n = \text{DFT} \{ \hat{f}_n(x,y) \} = \text{DFT} \{ \delta(x - \hat{x}_n, y - \hat{y}_n) \hat{p}(x,y) \}, \quad (18)$$

δ is a small constant that prevents problems in the gradient computation where I_n is close to zero, γ is a real-valued constant, and $W_n(u,v)$ is a weighting function that can be used to emphasize regions with high SNR or set to zero for regions with poor SNR or where no signal was measured. This can be used, for example, to eliminate the effects of a beam stop or dead detector pixels by setting $W_n(u,v) = 0$ for those pixels. For the simulation results shown in this paper we used a uniform unity weighting function for all pixels.

The gradient of ε with respect to the real and imaginary parts of the object, $\hat{o}(x,y) = \hat{o}_R(x,y) + i\hat{o}_I(x,y)$, is obtained by computing the expression

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \hat{o}_R(x,y)} + i \frac{\partial \varepsilon}{\partial \hat{o}_I(x,y)} = 4 \sum_{n=1}^q \hat{p}^*(x + \hat{x}_n, y + \hat{y}_n) \text{IDFT} \left\{ W_n \left[\left(|\hat{F}_n|^2 + \delta \right)^\gamma - \left(I_n + \delta \right)^\gamma \right] \right. \\ \left. \times \gamma \left(|\hat{F}_n|^2 + \delta \right)^{\gamma-1} \hat{F}_n \exp \left[i2\pi \left(\frac{u\hat{x}_n}{M} + \frac{v\hat{y}_n}{N} \right) \right] \right\}, \quad (19) \end{aligned}$$

where $\text{IDFT}\{\cdot\}$ is the inverse discrete Fourier transform.

The gradient with respect to the real and imaginary parts of $\hat{p}(x,y)$ can be computed in a similar fashion

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \hat{p}_R(x,y)} + i \frac{\partial \varepsilon}{\partial \hat{p}_I(x,y)} = 4 \sum_{n=1}^q \hat{o}^*(x - \hat{x}_n, y - \hat{y}_n) \text{IDFT} \left\{ W_n \left[\left(|\hat{F}_n|^2 + \delta \right)^\gamma - \left(I_n + \delta \right)^\gamma \right] \right. \\ \left. \times \gamma \left(|\hat{F}_n|^2 + \delta \right)^{\gamma-1} \hat{F}_n \right\}. \quad (20) \end{aligned}$$

Finally, the gradient with respect to the translation shifts (\hat{x}_n, \hat{y}_n) is given by

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \hat{x}_n} = \frac{8\pi}{M} \sum_{u,v} W_n \left[\left(|\hat{F}_n|^2 + \delta \right)^\gamma - \left(I_n + \delta \right)^\gamma \right] \gamma \left(|\hat{F}_n|^2 + \delta \right)^{\gamma-1} \\ \times \text{Im} \left[\hat{F}_n^* \text{DFT} \left(\hat{p}(x,y) \text{IDFT} \left\{ u' \hat{O}(u',v') \exp \left[-i2\pi \left(\frac{u'\hat{x}_n}{M} + \frac{v'\hat{y}_n}{N} \right) \right] \right\} \right) \right] \quad (21a) \end{aligned}$$

$$\begin{aligned} \frac{\partial \varepsilon}{\partial \hat{y}_n} = \frac{8\pi}{N} \sum_{u,v} W_n \left[\left(|\hat{F}_n|^2 + \delta \right)^\gamma - \left(I_n + \delta \right)^\gamma \right] \gamma \left(|\hat{F}_n|^2 + \delta \right)^{\gamma-1} \\ \times \text{Im} \left[\hat{F}_n^* \text{DFT} \left(\hat{p}(x,y) \text{IDFT} \left\{ v' \hat{O}(u',v') \exp \left[-i2\pi \left(\frac{u'\hat{x}_n}{M} + \frac{v'\hat{y}_n}{N} \right) \right] \right\} \right) \right]. \quad (21b) \end{aligned}$$

The same data used for the reconstruction with the PIE algorithm [with results shown in Figs. 3(a) and 3(b)] was fed to a conjugate-gradient routine that utilizes the expressions for the gradients as given by Eqs. (19)-(21). Better results were obtained by starting with a few iterations where we only optimized over $\hat{o}(x,y)$ (25 iterations for this reconstruction) and following with joint estimation of $\hat{o}(x,y)$, $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n) in subsequent iterations. We define one iteration of the algorithm to be every time the gradient needs to be computed. The amplitude and phase of the object reconstruction after 500 iterations (with $\gamma = 0.5$) are shown in Figs. 3(d) and 3(e) respectively, where a substantial improvement over the PIE reconstruction is evident. The amplitude of the final estimate of $\hat{p}(x,y)$ is shown in Fig. 3(f). Although for estimation of the object (and for the PIE), as few as 3 or 4 diversity measurements are typically sufficient for proper convergence of the algorithm, we observed that as many as 10 measurements are sometimes necessary to refine an inaccurate initial estimate for the illumination pattern and the translations. This is to be expected because the illumination pattern is optimized on a point-by-point basis and the number of free parameters is significantly increased.

Although the error metric given in Eq. (17) is a good measure of convergence to a solution (for both algorithms) it only measures how good our estimate matches the measured data by comparing the Fourier intensities. Thus this metric is insensitive to errors in the estimated phase of the far-field, which is the quantity we attempt to retrieve. In numerical simulations (where we know the correct far-field phase) we can define a metric which will not only measure the agreement with the measured data but also indicate convergence to the correct solution. We define the normalized invariant field RMSE as [37]

$$E^2 = \frac{1}{q} \sum_{n=1}^q \left[\min_{\rho_n, x', y'} \frac{\sum_{x,y} |\rho_n \hat{f}_n(x-x', y-y') - f_n(x,y)|^2}{\sum_{x,y} |f_n(x,y)|^2} \right]. \quad (22)$$

Notice that this error metric compares the estimated and true complex-valued fields, while individually minimizing the error over a translation (x', y') and a complex-valued constant ρ_n for each diversity field. Because a reconstruction that is translated and multiplied by a constant is still considered successful, this minimization procedure is important to make E invariant to these operations.

Computation of E , as given in Eq. (22), requires finding the peak of the cross-correlation of \hat{f}_n and f_n to within a small fraction of a pixel [37, 38]. For the results shown in this paper this fraction was 1/25 of a pixel, computed by means of an efficient algorithm for subpixel image registration by cross-correlation [38].

Figure 3(g) shows E versus iteration number for the reconstruction with PIE and nonlinear optimization for the case of inaccurate knowledge of both $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n) . Notice that while the PIE stagnates after a few iterations, the nonlinear optimization approach continues making progress in the reconstruction while refining the estimation of $\hat{o}(x,y)$, $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n) . The shift error, Δr , is shown versus iteration number in Fig. 3(h). During the first 25 iterations this error remains constant because we optimize only over $\hat{o}(x,y)$, subsequent iterations show a gradual improvement of the estimation of (\hat{x}_n, \hat{y}_n) .

An additional advantage of the nonlinear optimization algorithm over the PIE is that in the latter the measured intensities are imposed sequentially, so that every time the object field estimate $\hat{o}_n(x,y)$ is updated, noise artifacts are introduced by the n -th diversity image. In contrast, nonlinear optimization improves the object estimation on every iteration by using all of the intensity measurements simultaneously, which makes it more robust to noise in the detected intensity patterns. To demonstrate that effect, we performed a computer simulation with noisy data, but with perfectly known $p(x,y)$ and (x_n, y_n) .

After normalizing the simulated intensity patterns (used for the reconstruction shown in Fig. 2) to have about 10^4 photons on their brightest pixel, we applied poisson distributed noise.

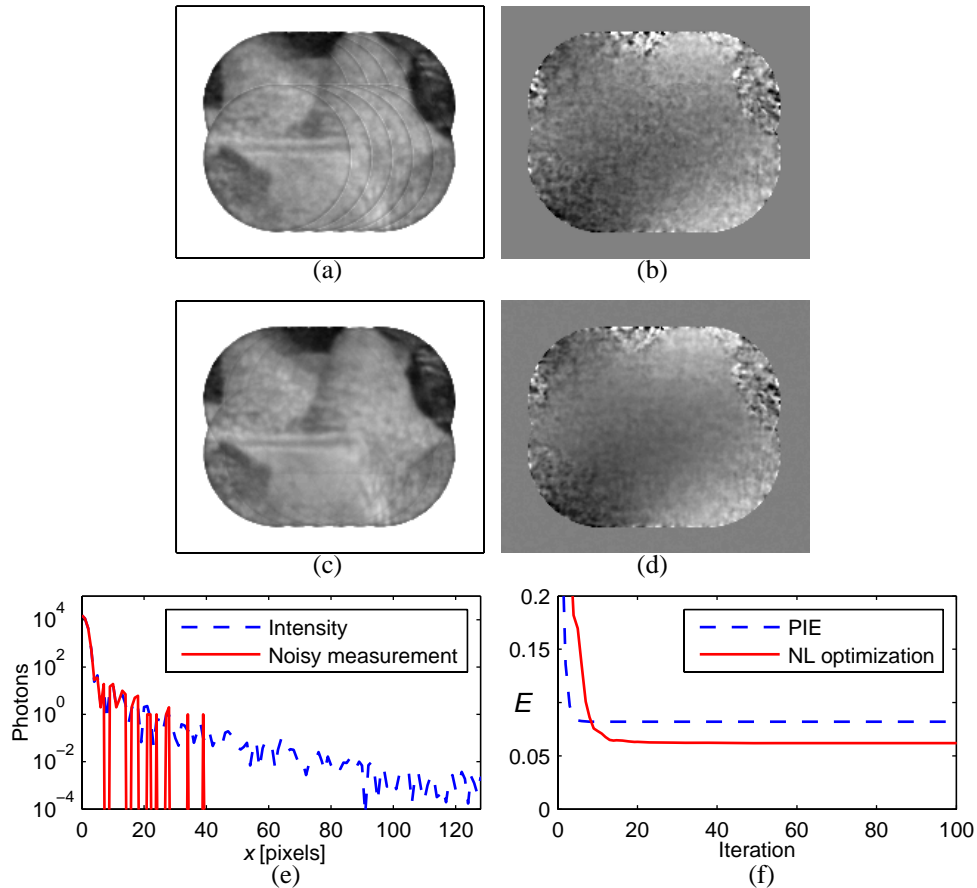


Fig. 4. Reconstructions from noisy data. Recovered amplitude and phase with the PIE, (a) and (b) respectively, and the nonlinear optimization algorithm, (c) and (d). (e) Cut through a measured Fourier intensity pattern (before and after applying noise). (f) Normalized invariant error metric, E , vs. iteration number for reconstructions from noisy data. Phase is shown from -0.4 to 0.4 radians in (b) and (d). A 172×210 portion of the 256×256 array is shown in (a)-(d).

Figure 4(e) shows a cut through a sample intensity pattern before and after noise was applied. Notice that after about 20 pixels from the brightest pixel, the intensity falls below an average of 1 photon/pixel; this results in the loss of high frequency components and we thus expect the reconstruction to have a reduced resolution.

The amplitude and phase of the reconstructions after 100 iterations are shown in Figs. 4(a) and 4(b) for the PIE, and Figs. 4(c) and 4(d) for the nonlinear optimization algorithm. Notice that the artifacts of sequentially imposing the far-field intensity measurements, that are noticeable in Fig. 4(a), are substantially reduced by the use of nonlinear optimization, as shown in Fig. 4(c). Figure 4(f) shows that a smaller normalized invariant error metric, E , is obtained with the nonlinear optimization approach over the PIE.

Although we achieved a reconstruction with reduced noise artifacts with the nonlinear optimization algorithm, we found both approaches to be very robust in the presence of noise. A large amount of noise needed to be introduced before these artifacts were visible in either

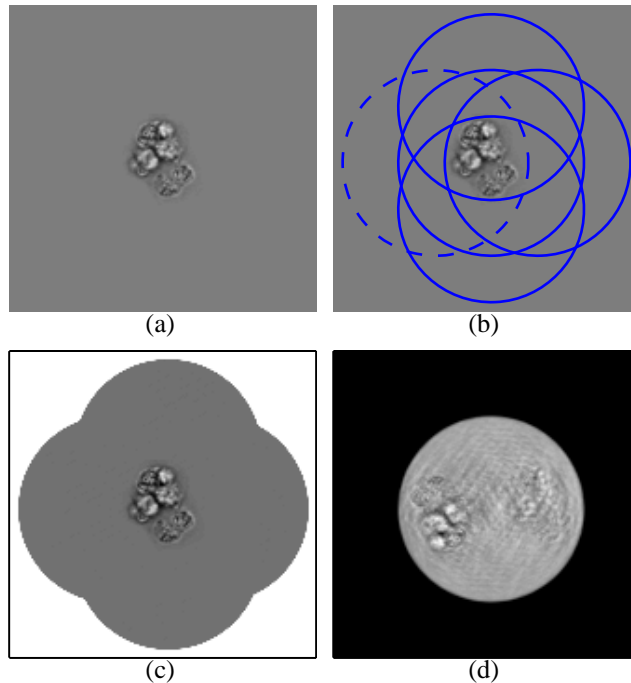


Fig. 5. (a) Object $o(x,y)$. (b) Circles indicate the five positions of $p(x,y)$ that were used for the nonlinear optimization reconstruction. (c) Reconstruction after 100 iterations of the nonlinear optimization algorithm. (d) Example reconstruction from a single intensity measurement using phase retrieval with support [dashed line in (b)] and nonnegativity constraints. A 165×165 portion of the 256×256 array is shown in (a)-(d).

reconstruction.

5. Object contained within the illumination pattern

Since an important application of x-ray CDI is imaging small specimens, we can encounter the situation where the aperture or illumination function overfills the entire extent of the object. In this situation, no new information lies outside of the support of $p(x,y)$ and diversity is introduced by translating the object within the extent of $p(x,y)$ and measuring the resulting far-field intensity patterns. It is then important to investigate whether in this situation translation diversity would work as well as it does with objects that are larger than the aperture.

Five far-field intensity measurements were simulated by translating the 256×256 object (an image of a group of cells [39]), shown in Fig. 5(a), multiplying by the aperture shown in Fig. 2(c) and computing the intensity of the Fourier transform of the resulting function. The circles in Fig. 5(b) show the support of $p(x,y)$ for each of the five positions that were used for reconstruction.

Figure 5(c) shows the reconstructed object after 100 iterations of the nonlinear optimization algorithm, it is clear that a faithful reconstruction of the original object was achieved. Similar results were obtained with the PIE algorithm. Through this and other numerical simulations, we have found that the advantages of using translation diversity (robustness and convergence speed) are maintained even when the object is small and can be contained within the aperture (or illumination pattern). This approach, however, requires the illumination beam to be coherent

over the extent of $p(x,y)$ and not just the object width as would be for a single-measurement image reconstruction. Faster convergence was observed by translating the object close to the edge of the aperture.

Figure 5(d) shows an example reconstruction after 1000 iterations of the single-measurement phase retrieval iterative transform algorithm as outlined for the reconstructions shown in Fig. 1. Only the far-field intensity produced by the leftmost position of $p(x,y)$, indicated by a dashed line in Fig. 5(b), was used for this reconstruction. Additional to the support constraint we used a nonnegativity constraint during the algorithm iterations. In this reconstruction the twin image dominates, but because the support is centrosymmetric, the reconstruction is susceptible to the twin-image problem and faint features of the upright image can also be observed.

6. Conclusion

Transverse translation diversity provides a very practical solution to many of the difficulties that can be encountered with image reconstruction by phase retrieval that is easy to implement experimentally. In this approach the diversity images are collected by translating the object of interest with respect to a known illumination pattern. In agreement with previously reported results [19, 20], we have found that this form of diversity very robustly constrains the image reconstruction problem and leads to superior reconstructions and a significant increase in convergence speed over phase retrieval using a single intensity measurement.

We have shown that the PIE algorithm is closely related to a steepest descent search with a spatially variable step length. This algorithm, developed by Faulkner and Rodenburg [19, 20], leads to very good reconstructions if the illumination, $p(x,y)$, and the translations, (x_n, y_n) , are accurately known *a priori*. However, an experimental setup is bound to have some degree of uncertainty on the characterization of these parameters. It is shown that the reconstruction with the PIE suffers from severe artifacts when the system parameters are inaccurately known.

We have developed and tested a nonlinear optimization approach to solve the phase retrieval problem with translation diversity. Using analytical expressions for the gradient of a squared-error metric, we enable joint optimization over the object, illumination, and translations [$\hat{\delta}(x,y)$, $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n) respectively], so that the initial estimate of the parameters is refined as the reconstruction progresses. We have shown that this optimization approach leads to better reconstructions when the system is inaccurately characterized. In particular, the ability to optimize over the object translations, relaxes the required setup stability and accuracy of calibration for the translating stages in an experimental setup.

Although a single iteration of the nonlinear optimization approach is about $4.5\times$ longer than an iteration of the PIE [$\sim 6\times$ if optimizing over $\hat{\delta}(x,y)$, $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n)], the nonlinear optimization approach is more robust in the presence of inaccurate system parameters and yields reduced noise artifacts. For imaging simulations we have observed comparable convergence for the PIE and nonlinear optimization if the illumination and translations are known accurately. Furthermore, the number of diversity images required for proper convergence was found to be increased when the system parameters are known inaccurately and we jointly optimize over $\hat{\delta}(x,y)$, $\hat{p}(x,y)$ and (\hat{x}_n, \hat{y}_n) .

We have additionally shown that translation diversity also yields very robust reconstructions when the illumination pattern is larger than the object of interest. In this case the object is translated within the support of $p(x,y)$ and no additional information about the object is introduced by the translation.