Fresnel Coherent Diffractive Imaging

G. J. Williams,[1] H. M. Quiney,[1] B. B. Dhal,[1] C. Q. Tran,[1] K. A. Nugent,[1] A. G. Peele,[2] D. Paterson,[3,*] and M. D. de Jonge[3]

[1]*School of Physics, University of Melbourne, Victoria 3010, Australia*

[2]*Physics Department, La Trobe University, Victoria 3086, Australia*

[3]*Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois, 60439, USA*

(Received 30 March 2006; published 14 July 2006)

We present an x-ray coherent diffractive imaging experiment utilizing a nonplanar incident wave and demonstrate success by reconstructing a nonperiodic gold sample at 24 nm resolution. Favorable effects of the curved beam illumination are identified.

DOI: 10.1103/PhysRevLett.97.025506 PACS numbers: 61.10.Nz, 41.50.+h, 42.30.Rx

X-ray diffraction has been a powerful tool for structure determination for decades but nanoscale noncrystallographic imaging was not possible until the advent of efficient focusing optics, which can be used to create an x-ray microscope. A range of applications from structure determination for rational drug design to the imaging of defects and strain in novel materials would benefit from such an efficient, high resolution device. Unfortunately, the resolution of these microscopes is limited by the spot size produced by the optic and is many times the wavelength of the x rays. Within the last decade, several groups have reported success in x-ray coherent diffractive imaging (CDI)—a lensless microscopy—which, in principle, provides wavelength limited resolution.

X-ray CDI was first demonstrated by Miao *et al.* [1] in a transmission geometry and later by Robinson *et al.* [2] in a reflection geometry. It has since been applied to biological samples [3,4]. Compared to traditional x-ray microscopy, the ultimate resolution is not limited by the lens, but rather by the wavelength, λ. Since the interaction of x rays with matter is weak compared to that of electrons, multiple scattering can be neglected and the 3D structure may be determined without physical slicing of the sample [5,6].

CDI requires that the intensity be sampled at the Nyquist frequency [7] and some *a priori* sample information is known. Typically, an iterative procedure to is used to find the phase of the scattered wave in the detector plane. The exit surface wave (ESW) at the sample is related to the wave at the detector by free space propagation. The routine generally utilizes an algorithm related to that created by Gerchberg and Saxton (GS) [8], wherein a solution is estimated by sequential application of constraints to an iterate in the detector and sample planes. These constraints are commonly the magnitude of the diffracted wave, as given by the measured intensity (modulus) and some knowledge of the extent of the object (support).

An intrinsic problem is the uniqueness of the reconstruction. One can prove that it is pathologically rare [9] for a nonunique solution to exist, but many ''trivial ambiguities'' will exist in any case. These ambiguities include phase offsets and conjugation errors, as well as variations in position of the reconstructed object. On the other hand, a unique solution does exist if Fresnel diffraction is utilized by providing an incident wave with known, finite curvature [10,11]. Modern optics produce fields with curvature of a few tens of nanometers [12], allowing Fresnel diffraction from large biomolecules to be measured.

In this Letter, we describe the experimental determination of a gold object through a phase retrieval scheme that includes the reconstruction of the complex amplitude of the incident wave, and describe the behavior of the error reduction (ER) [13] algorithm under these conditions.

The experimental geometry is demonstrated schematically in Fig. 1. At Sector 2-ID-B of the Advanced Photon Source (APS), the x-ray beam was monochromatized by a Si spherical grating monochromator (SGM) and used to illuminate a Fresnel zone plate (FZP), providing a beam with finite phase curvature in the vicinity of the focal plane. The sample was placed within this region. The combination of a central stop upstream of the FZP and the 10 μm order sorting aperture (OSA) removed not only the higher-order contributions from the FZP, but also the direct beam so that no beamstop was required on the CCD detector, whose pixel size was 24 μm. The FZP had a nominal outer

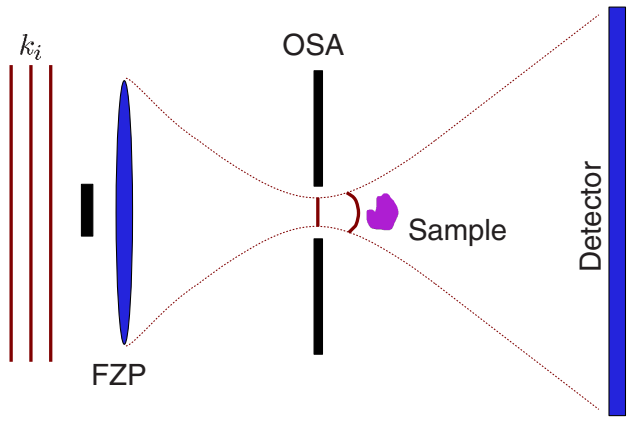

FIG. 1 (color online). A schematic of the experiment. A plane wave with wave number k_i is incident upon a FZP creating a curved wave front at the sample location.

zone width of 50 nm and a focal length of 12 mm at 1.83 keV. In this experiment, the sample was placed about 1 mm downstream of the focus. The x-ray beam remained *in vacuo* until the central stop and reentered vacuum immediately after transiting the sample, whereupon it entered a 0.5 m evacuated flight path. The transverse coherence of the beam has been studied [14] and the exit slit of the SGM was chosen to provide a coherence length of about 120 μm at the FZP in both transverse directions, producing a coherence length larger than the sample at the sample plane. The focusing produced a flux density at the sample that was at least an order of magnitude greater than the unfocused beam.

Figure 2(a)–2(c) are various regions within the data set used for the reconstruction presented in this Letter. This data set comprised 89 frames, each with a 3s exposure time. Figure 2(a) demonstrates the positional registration between the object in the beam and its diffraction. This registration leads to a translational sensitivity in the experiment, complicating the combination of many data sets to form a single long-time exposure. Figure 2(b) is the "holographic" region that results from interference of the incident beam with the sample's ESW. It was possible to recover an image from this hologram by backpropagating from the detector to the sample. This image was resolution limited by the properties of the lens and used to generate an initial guess for the algorithm and a support constraint. Figure 2(c) contains the topmost fringe of the pattern and shows that it is well sampled for the purposes of CDI. There is a ratio of more than 10 000:1 in the analog to digital units of the detector between the average value in a pixel in the central region [Fig. 2(b)] and the average value over the highest angle fringe measured [topmost in Fig. 2(c)].

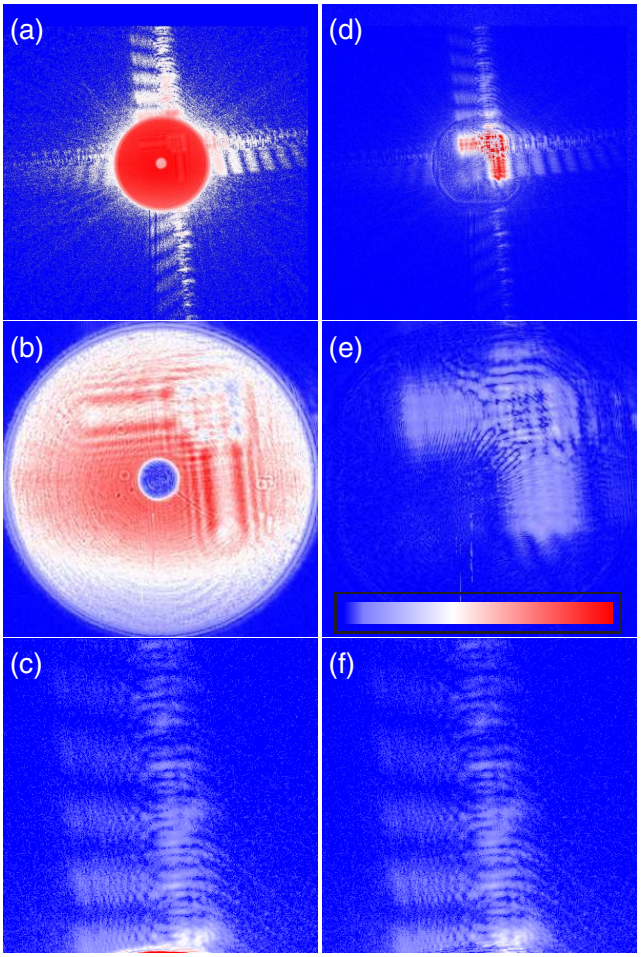

FIG. 2 (color online). Comparison of $\sqrt{I^m}$ to the reconstruction. (a), (d) are the data and the ESW propagated to the detector plane, respectively (logarithmically scaled). (b), (e) are close-ups of central region plotted on the same linear scale. (c), (f) are the topmost flare, with mapping is as in (b) and (e), except thresholded at 15% of the maximum value in those panels.

The method of finding a set of phases consistent with the measured intensity can be described as a "projections onto sets" method. In this case, operators are defined for each of the constraints to be applied [15]. For example, if a support constraint is defined to set all amplitudes to zero outside some region known to contain the sample, an operator, π_s, can be said to act on an iterate, ρ, giving a quantity, $\pi_s\rho$, which obeys the constraint. Elser's Difference Map [16], Fienup's Hybrid Input-Output [13], and the original GS [8] algorithm can all be expressed using this notation. We use ER, for which the $(k+1)$th iterate is given by

$$\rho_{k+1} = \pi_s \pi_m \rho_k, \tag{1}$$

where π_m is a modulus constraint. Typically, π_m will contain an implicit propagation to the detector and back. For example, if we define the propagation operator as $\mathcal{F}$, $\pi_m = \mathcal{F}^{-1}\tilde{\pi}_m\mathcal{F}$, where operators decorated with a tilde operate explicitly in the space of the detector and $\tilde{\pi}_m$ imposes the modulus constraint in that space. Henceforth, we regard ρ_k as an iterate and $\pi_m\rho_k$ as an estimate of the sample's ESW, i.e., a potential solution.

Equation (1) formed the basis of the iterative technique. After an initial attempt to guess the support from the hologram, it was found that the support had to be altered to allow additional features to form. Rather than this tight support, we began with a very loose constraint and iteratively tightened it by generating a new support on each cycle. To do this, the iterate was convolved with a 1 pixel variance Gaussian function and thresholded at a progressively higher level, up to 11% of the maximum value of the image. This is similar to the "shrinkwrap" algorithm [17]. The final threshold was determined empirically by monitoring the error metric, defined in the detector plane by

$$\chi^2 = \frac{\sum_{n=0}^{N-1}[|\sqrt{I^m(n)}| - |[\mathcal{F}\rho_k](n)|]^2}{\sum_{n=0}^{N-1} I^m(n)}, \tag{2}$$

where ρ_k is the iterate on the kth iteration and $I^m(n)$ is the measured intensity in the nth pixel. The upper limit on the threshold corresponded to a minimum of the metric, which

monotonically increased with increasing threshold. After the 30th iteration, the support was fixed at its final shape.

To deemphasize the fitting of spurious scatter in the data, we used the support to place a constraint on the spatial frequency spectrum in the data [18]. This constraint improved the convergence of the iterative algorithm. If we let z be the coordinate in and $\mathbf{r}$ be the 2D vector perpendicular to the direction of propagation, the wave at the detector is $\psi_{\mathrm{tot}}(\mathbf{r}_2, z_2) = \mathcal{F}\{[1 + T(\mathbf{r}_1)]\psi_{\mathrm{inc}}(\mathbf{r}_1, z_1)\}$, where we want $T(\mathbf{r}_1)\psi_{\mathrm{inc}}(\mathbf{r}_1, z_1)$, the ESW at the sample. It was first necessary to recover the incident illumination on the sample, which was accomplished by means of the procedure detailed in Ref. [19]. At each iteration, the modulus constraint then consists of the following steps: (i) propagate from the sample plane to the detector plane ($\mathcal{F}$), (ii) add the complex illumination determined as described above ($\tilde{\pi}_{WF}$), (iii) enforce consistency with the modified measured intensity ($\tilde{\pi}_M^k$), (iv) subtract the illumination ($\tilde{\pi}_{WF}^{-1}$), (v) backpropagate to the sample plane ($\mathcal{F}^{-1}$). Equation (1) then becomes:

$$\rho_{k+1} = \pi_s \pi_m \rho_k = \pi_s \mathcal{F}^{-1} \tilde{\pi}_{WF}^{-1} \tilde{\pi}_M^k \tilde{\pi}_{WF} \mathcal{F} \rho_k. \quad (3)$$

In essence, this is the algorithm described in Ref. [20].

To test the accuracy of our reconstruction, we propagate the ESW back to the detector plane for comparison with the measured data. The scattered amplitude is represented in Figs. 2(d)–2(f). This has been isolated from the corresponding panels of Fig. 2(a)–2(c), respectively—which contain both the illuminating and scattered fields whose relationship is described by the formalism of Eq. (3). In panel (a), we see that the scatter is reconstructed to high angle, far beyond the holographic region. Based on Abbe Theory [21] and the numerical aperture, NA—defined to be one half of the detector acceptance angle—we find a theoretical resolution of $\Gamma = 0.82\lambda/\mathrm{NA} = 24$ nm. The signal to noise ratio remained favorable at the edges of the detector. It is necessary to recover the incident illumination with resolution at least as high as that of the ESW, if the transmission is to be decoupled from the illumination. We accomplished this by keeping the numerical aperture unchanged in separate measurements of the beam and sample scattering. In Fig. 2(e), one can see vertical lines near the bottom of the image. These are evidently the result of detector defects and persisted in the reconstruction even if the modulus constraint was not enforced in this region. The remaining panels show that the fine detail in the fringes is reconstructed. The final error metric value for this fit was $\chi^2 = 1.99 \times 10^{-4}$ and the reconstruction is in excellent agreement with the data. This reconstruction was one of 20 trials, each starting with a random set of phases. The propagation between sample and detector planes was accomplished by means of a fast Fourier transform (FFT) with array size 1024×1024.

The target object is a nested grouping of chevrons, shown in Fig. 3(a). This image is a scanning transmission

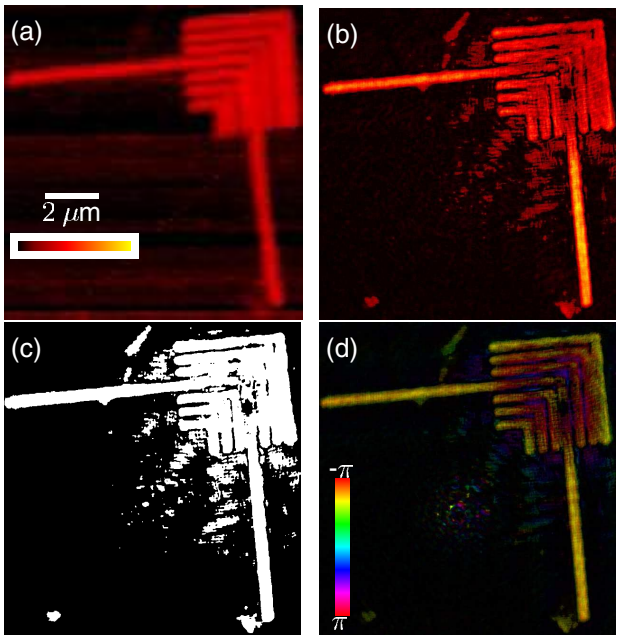

FIG. 3 (color online). A STXM image, (a), was collected following CDI data acquisition. The ESW reconstruction in (b)—$T(\mathbf{r}_1)\psi(\mathbf{r}_1)$—is of higher resolution and shows features that are recondite in the STXM image. (c) is the final support. $T(\mathbf{r}_1)$ is represented in (d) as the product of its amplitude and phase.

x-ray micrograph (STXM) taken immediately after the data collection. The sample was constructed of gold by electron beam lithography on a Si_3N_4 window, with approximate height 150 nm. The ESW reconstructed from the coherent scatter is shown in panel (b) and its final support in Fig. 3(c). The oversampling ratio, defined to be one over the fraction of the support occupied by the object [1], as judged by the extent of the final support, is 12 here. We see that the shape is well recovered and the object has very sharp edges as expected. It is interesting to note that the reconstruction contains some elements not due to the lithographic object. The long, thin object near the center of the upper long leg and the patches at the end of the lower leg and the bottom left corner appear in the same locations in each of the 20 reconstructions. These are also present in the micrograph; however, they were not evident on the sample by scanning electron or atomic force microscopy after the experiment. We believe these to be particles of dust or imperfections in the substrate. Their presence serendipitously demonstrates that small objects of lower atomic number are amenable to imaging in this way.

The remaining amplitude outside the main body of the chevrons is thought to be an artifact. While it appears in a similar location in all reconstructions, the details are not constant. Further, while 89 frames of data were summed, a total of 800 were collected. The frames used here were selected on the basis that they are well correlated with one another. If a larger subset of the data is used, these artifacts grow in amplitude relative to the chevrons. The lack of

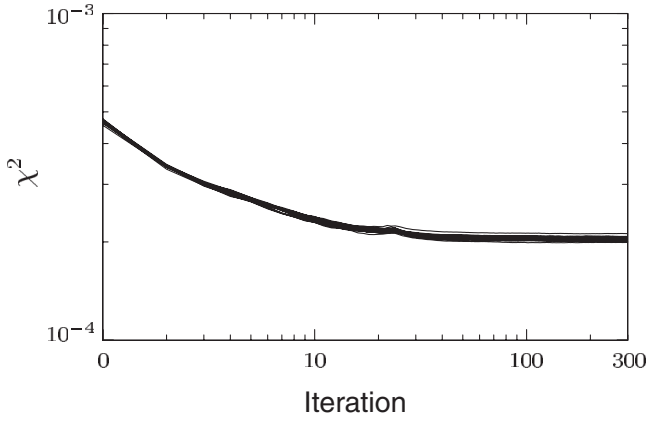

FIG. 4. The behavior of χ^2 during the reconstruction displays none of the stagnation common in plane-wave CDI using ER.

correlation is believed to be due to sample movement. The phase of the incident wave at a particular z_0 is proportional to $e^{i\pi r^2/\lambda R}$. R is the curvature of the incident beam given by Gaussian optics and in this limit is equal to the distance from the focus. This factor is such that a shift of 100 nm can change the phase of the illumination on the sample by as much as π rad. We anticipate the algorithm will become unreliable with such large shifts in the phase of the incident illumination.

Figure 3(d) is the product of the magnitude and phase of $T(\mathbf{r}_1)$, with the phase and amplitude given by hue and value, respectively. The bright artifacts just below and left of center are due to division by the incident illumination and mark the center of the FFT array. The complicated phase structure is not believed to be indicative of the sample structure and may be due to sample motion.

A pivotal concern is the reproducibility of the iterative fitting. In the current case, a relatively small number of iterations resulted in convergence. The error metric ranged from $\chi^2 = 1.99 \times 10^{-4}$ to $\chi^2 = 2.12 \times 10^{-4}$, a difference of 6.5%. Figure 4 is a plot of the trajectory of χ^2 for all 20 reconstructions. As expected from a curved illumination experiment [20]—and in contrast to planar illumination—all reconstructions appear at the same position in the array and have no phase offset relative to one another. Most importantly with regard to reconstructions using ER, the twin image never forms and so stagnation is avoided. In the simulations of Ref. [20], a "Fresnel number," N_F, of 5 was found to provide excellent and reliable convergence. In this case, $N_F = 28$ across the largest dimension of the sample.

We have demonstrated experimentally a CDI result from an object illuminated by a beam of x rays with curvature. The ability to characterize the incident beam from an independent measurement removes any arbitrary phase offset between reconstructions that may otherwise occur, raising the possibility of quantitative analysis of the ESW. To achieve the latter, the geometrical dimensions of the experiment must be accurately and precisely determined, as the reconstruction is sensitive to errors in the propagation distances or sample translations and rotations.

This sensitivity provides a self-consistency check on the reconstruction, as the propagation distances must be the same in the reconstruction of the illumination and the ESW. Two striking benefits of the method are the improved convergence of the iterative algorithm and the ability to perform the experiment without a beamstop, which traditionally results in a loss of low spatial frequency information in the analogous plane-wave experiment.

The authors acknowledge the support of the Australian Research Council Centre of Excellence for Coherent X-ray Science and the Australian Synchrotron Research Program. Use of the APS was supported by the U.S. Department of Energy, under Contract No. W-31-109-ENG-38. We also thank Matteo Altissimo for providing the sample.

*Present Address: Australian Synchrotron, 800 Blackburn Road, Clayton, VIC 3168, Australia.

[1] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, Nature (London) **400**, 342 (1999).
[2] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney, Phys. Rev. Lett. **87**, 195505 (2001).
[3] J. Miao *et al.*, Proc. Natl. Sci. U.S.A. **100**, 110 (2003).
[4] D. Shapiro *et al.*, Proc. Natl. Sci. U.S.A. **102**, 15 343 (2005).
[5] G. J. Williams *et al.*, Phys. Rev. Lett. **90**, 175501 (2003).
[6] J. Miao *et al.*, Phys. Rev. Lett. **89**, 088303 (2002).
[7] D. Sayre, Acta Crystallogr. **5**, 843 (1952).
[8] R. W. Gerchberg and W. O. Saxton, Optik (Jena) **35**, 237 (1972).
[9] R. H. T. Bates, Optik (Jena) **61**, 247 (1982).
[10] K. A. Nugent *et al.*, Acta Crystallogr. Sect. A **61**, 373 (2005).
[11] T. A. Pitts and J. F. Greenleaf, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **50**, 1035 (2003).
[12] B. D. Harteneck *et al.*, Nature (London) **435**, 1210 (2005).
[13] J. R. Fienup, Appl. Opt. **21**, 2758 (1982).
[14] C. Q. Tran *et al.*, Opt. Lett. **30**, 1 (2005).
[15] H. H. Bauschke, P. L. Combettes, and D. R. Luke, J. Opt. Soc. Am. A **19**, 1334 (2002).
[16] V. Elser, J. Opt. Soc. Am. A **20**, 40 (2003).
[17] S. Marchesini *et al.*, Phys. Rev. B **68**, 140101(R) (2003).
[18] G. J. Williams, H. M. Quiney, and K. A. Nugent (to be published).
[19] H. M. Quiney, A. G. Peele, Z. Cai, D. Paterson, and K. A. Nugent, Nature Phys. **2**, 101 (2006).
[20] H. M. Quiney, K. A. Nugent, and A. G. Peele, Opt. Lett. **30**, 1638 (2005).
[21] M. Born and E. Wolf, *Principles of Optics* (Cambridge University Press, Cambridge, England, 1999), 7th ed.