Fresnel diffractive imaging: Experimental study of coherence and curvature

L. W. Whitehead, G. J. Williams, H. M. Quiney, and K. A. Nugent
School of Physics, The University of Melbourne, Victoria 3010, Australia

A. G. Peele
Department of Physics, La Trobe University, Bundoora, Victoria 3086, Australia

D. Paterson
Australian Synchrotron, Clayton, Victoria 3168, Australia

M. D. de Jonge and I. McNulty
Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, USA
(Received 9 October 2007; published 14 March 2008)

A Fresnel coherent diffractive imaging experiment is performed using a pinhole as a test object. The experimental parameters of the beam curvature and coherence length of the illuminating radiation are varied to investigate their effects on the reconstruction process. It is found that a sufficient amount of curvature across the sample strongly ameliorates the effects of low coherence, even when the sample size exceeds the coherence length.

DOI: 10.1103/PhysRevB.77.104112 PACS number(s): 61.05.cp, 42.30.Rx, 87.59.−e

I. INTRODUCTION

Coherent diffractive imaging (CDI) seeks to recover an image of a noncrystalline sample from its far-field diffraction pattern. Iterative recovery methods, such as the Gerchberg-Saxton algorithm,[1] as generalized by Fienup,[2] and other variants, are employed and the methods have produced striking images of manufactured gold samples,[3–5] nanocrystals,[6] and cellular samples.[7,8] It has been suggested that in the future the technique may be of use in imaging individual protein molecules.[9]

A central assumption is that the illuminating radiation is fully coherent. Although this is a good approximation for lasers, it is less appropriate for current x-ray sources. While full coherence is clearly desirable, it seems that a fully coherent beam is not strictly required for the iterative algorithms to converge. It has been previously suggested that the coherence length should be at least twice the lateral dimensions of the object,[10] spanning the extent of its autocorrelation function.

CDI has typically been performed using plane-wave illumination. Recent experiments, however, have successfully applied the technique using a curved beam.[11] This new extension, known as Fresnel coherent diffractive imaging (FCDI), so called because the data are now a Fresnel—rather than a Fraunhofer—diffraction pattern, is a powerful new addition to the field. The curvature creates a unique solution to the inversion of the diffraction pattern[12] and stabilizes the reconstruction algorithms, largely preventing the problem of stagnation.

Although Fresnel coherent diffractive imaging makes use of optics to introduce phase curvature to the illumination, the resolution of the reconstruction is not optic limited. Instead, as with plane-wave CDI, the resolution is limited by the angle to which the diffraction pattern can be measured.

In this paper we first present a review of the relevant theory of the propagation of partially coherent x rays. Then we describe an experiment examining the effects of both the curvature and the partial coherence of the illumination on the reconstruction of a 4-μm-diam pinhole. Finally we show that, although lower degrees of coherence in the beam entering the experiment have a detrimental effect on the reconstructions, the effects of poorer coherence can be overcome through an application of greater curvature in the illuminating field.

II. THEORY

The far-field intensity $I(\mathbf{s})$ is related to the complex wave field at the detector, $F(\mathbf{s})$, by $I(\mathbf{s})=F^*(\mathbf{s})F(\mathbf{s})$, where $\mathbf{s}$ is a two-dimensional vector in the detector plane. This wave field is obtained by propagation from the exit surface of the object, where the wave field is denoted by $f(\mathbf{r})$ and $\mathbf{r}$ is a vector in the object plane. In the far field, the wave field $F(\mathbf{s})$ is related to $f(\mathbf{r})$ by a Fourier transform relation $F(\mathbf{s})=\hat{F}[f(\mathbf{r})]$, where $\hat{F}$ represents the Fourier transform operator. From experiment, however, we know only $|F(\mathbf{s})|=\sqrt{I(\mathbf{s})}$. *A priori* information about the sample, typically information about the area over which the sample is nonzero (its support), is built into the reconstruction algorithm. All of the missing information is contained in the phase of $F(\mathbf{s})$, so that the process of forming the image of a diffracting object by coherent diffractive imaging is a form of phase retrieval. The iterative methods employed by coherent diffractive imaging use the fast-Fourier-transform algorithm to propagate between the exit surface and the far field to find a function that satisfies the support constraint and whose Fourier transform predicts the correct intensity distribution at the detector. These algorithms rely on the fundamental assumption that the object is effectively finite so that the far-field intensity distribution can be sampled at or above its Nyquist frequency.

The simplest of these algorithms, proposed by Fienup and known as error reduction (ER),[2] uses the dimensions and

shape of the object as a constraint (the support constraint) in the sample plane. The algorithm proceeds by taking an initial guess of the complex wave at the object plane, propagating to the far field, retaining the calculated phase information but replacing the amplitude with the square root of the measured diffraction pattern (the modulus constraint), propagating back to the object plane, reapplying the support constraint, propagating back to the detection plane, reapplying the modulus constraint, and so on. This process is repeated until a consistent solution is achieved.

ER often stagnates, and so a number of alternate algorithms have been proposed that include some form of feedback parameter that allow the algorithm to break stagnation and converge more readily.[13]

The progress of these algorithms can be tracked by monitoring the error metric:

$$\chi^2 = \frac{(\sum_i |A_i^{calc}| - \sqrt{I_i^{meas}})^2}{\sum_i I_i^{meas}}, \tag{1}$$

where i is the pixel number, A_i^{calc} is the calculated amplitude in the far field of the current iterate, and I_i^{meas} is the measured intensity data. A lower metric indicates a better fit to the far-field amplitude data, and it has been shown that in the absence of noise an accurate fit to the data produces a faithful reconstruction of the original object through back-propagation of the wave field.[14]

An interplay between the experimental parameters of curvature and degree of coherence in Fresnel coherent diffractive imaging was suggested in an earlier work.[15] There, it was demonstrated that the presence of sufficient curvature in the incoming beam reduced the negative effects of partially coherent illumination. We review the relevant theory here.

The exit surface wave of a thin object illuminated by a curved x-ray beam with wavelength λ is

$$\Psi(\mathbf{r}) = t(\mathbf{r})A(\mathbf{r})\exp\left(\frac{-i\pi r^2}{\lambda R}\right), \tag{2}$$

where $A(\mathbf{r})$ is the amplitude of the illumination, $t(\mathbf{r})$ is the (possibly complex) transmission function of the object, and R is the radius of curvature of the wave field. In the experiment described below, the object is a pinhole and so the effective exit surface wave is the apertured illumination, $a(\mathbf{r})\exp(-i\pi r^2/\lambda R)$. For most x-ray experiments, the radius of curvature can be approximated by the propagation distance from the focus of a lens, mirror, or zone plate.

The partially coherent far-field intensity is related to the exit surface wave by means of the integral

$$I(\mathbf{s}) = \int\int \exp[-2\pi i\mathbf{s}\cdot(\mathbf{r}_2 - \mathbf{r}_1)]J(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2. \tag{3}$$

We choose a coherence function of the form

$$J(\mathbf{r}_1,\mathbf{r}_2) = \Psi(\mathbf{r}_1)\Psi^*(\mathbf{r}_2)g(\mathbf{r}_1 - \mathbf{r}_2), \tag{4}$$

depending on two in-plane two-dimensional (2D) vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, and

$$g(\mathbf{r}_1 - \mathbf{r}_2) = \exp\left(\frac{-|\mathbf{r}_1 - \mathbf{r}_2|^2}{2\sigma^2}\right). \tag{5}$$

We note that if $g(\mathbf{r}_1 - \mathbf{r}_2)$ is a constant, and hence $\sigma \to \infty$, we recover the expected result for coherent illumination. When the coherence is not perfect (finite σ), the intensity is blurred by convolution with the Fourier transform of this function. The parameter σ is defined in the plane of the sample and is referred to here as the coherence length of the incident radiation.

To make explicit the connection of Eq. (3) to the expected intensity distribution when the illumination is perfectly coherent, we perform a variable substitution $\boldsymbol{\rho} = \mathbf{r}_1$ and $\boldsymbol{\eta} = \mathbf{r}_2 - \mathbf{r}_1$, giving

$$I(\mathbf{s}) = \int \Gamma(\boldsymbol{\eta})\exp[-2\pi i(\mathbf{s}\cdot\boldsymbol{\eta})]\exp\left[-\eta^2\left(\frac{1}{2\sigma^2} + \frac{i\pi}{\lambda R}\right)\right]d\boldsymbol{\eta}, \tag{6}$$

where we have introduced the modified autocorrelation function

$$\Gamma(\boldsymbol{\eta}) = \int a(\boldsymbol{\rho})a^*(\boldsymbol{\eta} + \boldsymbol{\rho})\exp\left(\frac{i\pi\boldsymbol{\eta}\cdot\boldsymbol{\rho}}{\lambda R}\right)d\boldsymbol{\rho}. \tag{7}$$

It is now obvious that in the case of perfectly coherent illumination ($\sigma \to \infty$), $I(\boldsymbol{s})$ describes a Fresnel far-field distribution. If the radius of curvature also approaches infinity ($R \to \infty$), we recover the Fraunhofer result; $I(\boldsymbol{s})$ then describes the Fourier transform of the unmodified autocorrelation.

The third term in the integrand of Eq. (6) is a Gaussian function of complex argument, the real and imaginary parts of which depend only on the coherence length and the curvature, respectively, of the incident wave. It is the competition between these two quantities that provides Fresnel CDI with its resilience to partial coherence: when R is small, the function oscillates so rapidly that only components of the autocorrelation for which $\boldsymbol{\eta} \approx 0$ contribute to the far-field intensity. This analysis can be developed into a heuristic for estimating how much curvature is required to overcome the adverse effects of partial coherence. A satisfactory estimate of the amount of curvature required for a given σ is

$$R \ll \frac{2\sigma^2}{\lambda}, \tag{8}$$

and we will reference this relation in the context of our experimental results.

III. EXPERIMENT

The theoretical discussion reviewed in the previous section assumed an isotropic coherence length. The experimental arrangement described in this section does not allow for the vertical coherence length to be adjusted, and so all the results here concern a change in the coherence length in the horizontal direction. However, we note that the conclusions of the previous section also obtain where the coherence length varies only in one direction. Accordingly, for the remainder of this paper, the term coherence length will refer to the horizontal coherence length.

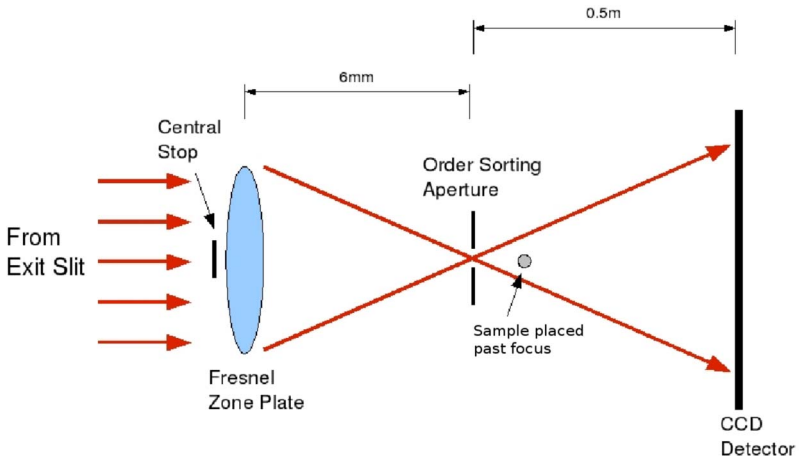

FIG. 1. (Color online) Diagram of experimental setup. The conditioned beam emerges from the left, is brought to a focus by the Fresnel zone plate, and propagated to the charge-coupled-device (CCD) detector. The sample was placed at a variable distance from the focus, allowing us to vary the amount of phase curvature incident upon it, while the coherence length in the conditioned beams is altered by adjusting the horizontal exit slit.

The experimental setup is shown in Fig. 1. The experiment was performed at beamline 2-ID-B at the Advanced Photon Source. After emerging from the undulator the beam is focused onto an entrance slit by means of a spherical mirror. A spherical grating monochromator selected the desired wavelength and projects the image of the entrance slit onto its exit slit.[16] The image of the source at the exit slit of the monochromator then becomes our effective source.

The beam propagated 8 m after emerging from the exit slit, whereupon it was focused using an 80-μm-diam Fresnel zone plate (FZP). Photons with an energy of 1.83 keV were used, resulting in a focal length of 6 mm. A 28-μm-diam central stop was placed on a silicon nitride window approximately 3 mm upstream of the FZP, and an order-sorting aperture was placed at the focal plane to select only the first-order beam. The sample was a pinhole with a nominal diameter of 4 μm. To explore the effects of curvature R, the sample was placed at one of two positions, either 1.75 or 6.75 mm, from the focus.

The detector was a 1024×1024 square pixel CCD detector with a pixel size of 24 μm. The x rays propagate 0.5 m from the sample to the detector under vacuum. The requirement that the scattering be sampled at or above its Nyquist frequency means that the sample must be less than 7 μm across, a condition satisfied here. Two of the data sets are shown in Fig. 2.

Table I shows the approximate horizontal coherence lengths of the illuminating field at the sample for each of the four data sets. The values of σ were extrapolated from previous measurements on the 2-ID-B beamline[17,18] and have an uncertainty of $\pm 20\%$. Also shown are the values of R and $\frac{2\sigma^2}{\lambda}$ for each data set, which are related via Eq. (8). The data sets are labeled according to their relative curvatures and coherence values. The uppercase letters refer to the curvatures (high for the 1.75-mm sample-focus distance and low for the 6.75-mm sample-focus distance) while the lowercase letters refer to the relative coherence lengths (long or short). For example, HC-l is high curvature with the longer coherence length, which was obtained when the sample was at a point closer to the focus and the exit slit was at the smaller setting.

IV. ANALYSIS

The data were analyzed using the error-reduction algorithm. The modulus constraint was applied in the detector plane and the support constraint applied in the sample plane.

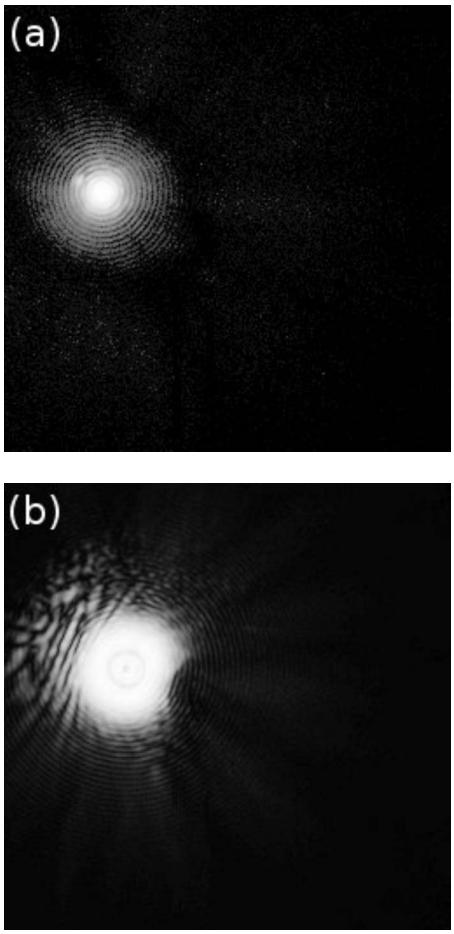

FIG. 2. High coherence data sets of diffraction from a 4-μm pinhole, cropped to 256×256 pixels and log scaled, centered on the optical axis. (a) is the low-curvature data set, $\sigma = 63$ μm. (b) is the high-curvature set, $\sigma = 16$ μm.

TABLE I. When referring to specific data sets we use the labels outlined here. The uppercase letters refer to the curvature (high or low) and the lowercase letters indicate the relative coherence length (long or short). The value of the curvature constraint is set by the experimental parameter R.

Label	Description	R [mm]	σ [μm]	$\frac{2\sigma^2}{\lambda}$ [m]
HC-l	High curvature, long σ	1.75	$\approx$16	0.75
HC-s	High curvature, short σ	1.75	$\approx$0.88	0.0023
LC-l	Low curvature, long σ	6.75	$\approx$63	12
LC-s	Low Curvature, short σ	6.75	$\approx$3.4	0.034

A support was chosen to have straight sides and right angles and be slightly larger than the pinhole.

The sample can be treated as a real positive object and the wave leaving the sample as the product of this object with the complex spherically curved incident wave field. The incident field was not precisely known and was assumed to have a perfectly uniform amplitude distribution and a curvature given by the nominal sample-focus distance. This is known not to be correct, and so the reconstructed exit wave field is expected to reproduce amplitude modulations present in the incident field.

The solution for the sample was constrained to be real and positive for the first 500 iterations, after which this constraint on the curvature of the incident field was removed; a complex sample implicitly allows for inaccuracies in the initial estimate of the curvature of the incident field to be compensated by the algorithm. The ER algorithm was then allowed to run for a further 1500 iterations. Upon the constraint removal, the error metric was found to drop rapidly, but the curvature estimate did not vary greatly from that enforced during the first 500 iterations, indicating that the initial estimate of the curvature of the incident field was consistent with both the known pinhole diameter and the experimental parameters used in the propagator employed in ER.

Each data set was reconstructed 100 times with each reconstruction using a different random complex array as the starting point.

V. RESULTS

Typical reconstructions from all four data sets are shown in Fig. 3. Figure 4 shows an image of the amplitude of the detected field, scaled to the correct size in the sample plane, with the low-curvature–high-coherence (LC-l) reconstruction superimposed. It can be seen that the reconstruction has reproduced details in the variation of the illuminating field as well as the hard edges of the sample. The direct comparison with the zone plate data is possible in this instance because the sample is more than twice the focal distance from the zone plate, so that, to an excellent approximation, the illuminating field can be represented by a geometric scaling of the field at the detector. This approximation is much less appropriate for the higher-curvature data, but the comparison in Fig. 4 confirms that the reconstructions are being obtained correctly and, as expected, retain some of the features of the illuminating field.

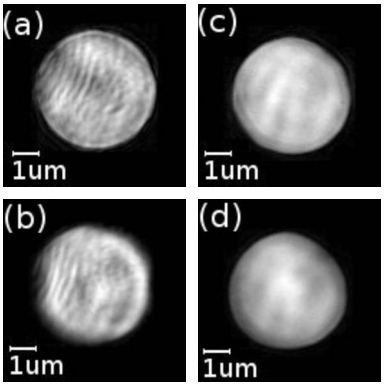

FIG. 3. Amplitude reconstructions from a 4-μm-diam pinhole. The support takes up approximately one-third of the array. These images have been cropped to a region of interest of 128×128 pixels. (a) is the HC-l reconstruction, (b) HC-s, (c) LC-l, and (d) LC-s. The LC-s reconstruction (d) appears to be deformed slightly along the diagonal, which is not consistent with reconstructions from the other data sets. This represents a lack of reproducibility at low curvature which is not present with high curvature.

The reconstructions return an object of the approximate size and shape of the sample. We note, however, that the low-curvature–short-coherence-length (LC-s) reconstruction does not reproduce the circular shape of the sample.

A comparison of the values and distribution of the final value of the error metric was used to assess the reliability of the reconstruction. Figure 5 shows the trajectory of the error metric for the four data sets. The large drop in the error metric at iteration 500 corresponds to the relaxation of the requirement that the sample be real and positive, and so allows the algorithm to effective adjust the estimate of the incident wave curvature. It is apparent that the high-curvature–long-coherence-length data set (HC-l) always converged to essentially the same error metric and therefore the same solution. The high-curvature–short-coherence-length

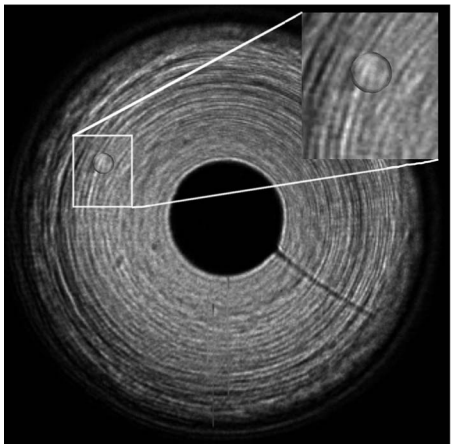

FIG. 4. Illumination with overlay. The inset shows the reconstructed intensity structure of the LC-l data and agrees well with the variations of intensity due to the zone plate. Note that the intensity has been scaled to the maximum amplitude of the zone plate, rather thanthe maximum amplitude of the reconstruction as in Fig. 3.

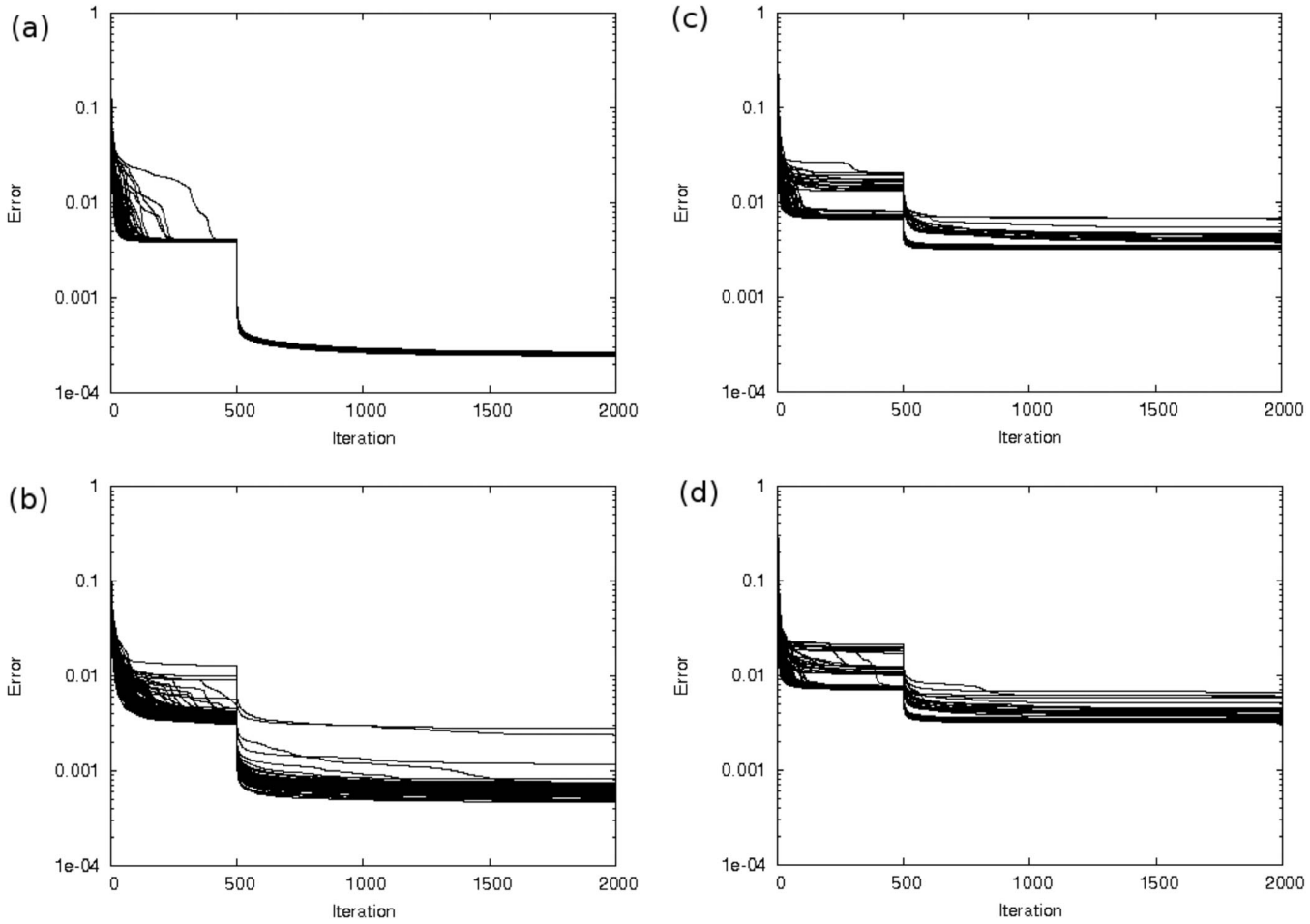

FIG. 5. Trajectory of error metrics. The significant drop in error at iteration 500 is where the curvature constraint was removed and a standard support constraint was applied. This phenomenon is often seen when changing constraint or algorithms. (a) is the error metric of HC-l, (b) HC-s, (c) LC-l, and (d) LC-s. The HC-l set (a) far outperforms all others in terms of both minimizing the metric and reproducing the same value.

(HC-s) data set also converges to a comparable error metric in most cases, though the consistency is far poorer than for the case with higher curvature. The low-curvature reconstructions are of comparable overall error metric but the shorter-coherence-length data (LC-s) showed a relatively even distribution of final error metrics over a band, while the longer-coherence-length reconstructions (LC-l) displayed a clustering around two end points.

The reproducibility of the four data sets was measured by comparing all 100 reconstructions of each data set with the mean of selected successful reconstructions. The reproducibility is defined as the normalized sum of the squares of the differences between the reconstruction and the mean reconstruction, where the mean reconstruction was the average over the 20 reconstructions with the lowest χ^2 value. Figure 6 shows the average reproducibility of each data set on the abscissa and the χ^2 values on the ordinate. The origin of Fig. 6 therefore represents an ideal in which all reconstructions are entirely reproducible and agree with the measured data. By this measure the high-curvature sets are considerably more reproducible for both coherence settings, attaining a reproducibility error value a factor of 2 less than that of

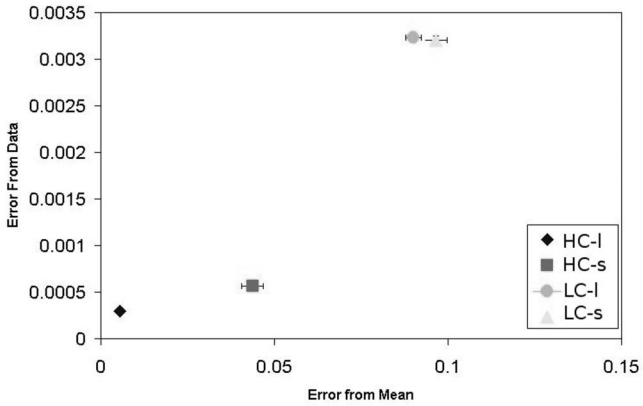

FIG. 6. Combination error metrics. The origin represents an ideal reconstruction that perfectly fits the data and is entirely reproducible. The distance from the origin is therefore a good estimate of the ranking of the data sets from most to least trustworthy. Error bars are the standard deviation values for each set; the key refers to Table I.

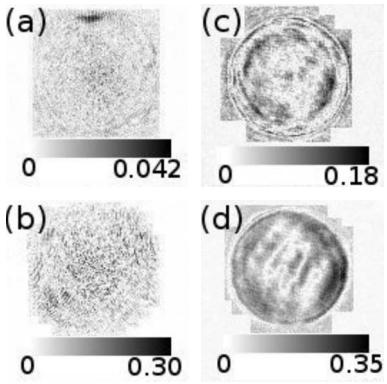

FIG. 7. 2D standard deviations. Both the high-curvature sets [(a) HC-l and (b) HC-s] show an essentially random distribution, whereas the low-curvature sets [(c) LC-l and (d) LC-s] show significant structure, indicating they are reconstructing different portions of the solution in a way that is not consistent. The amplitudes have been normalized and can therefore be directly compared.

the low-curvature sets (0.04 for HC-s compared to 0.098 for LC-s).

Figure 7 shows the average two-dimensional variation of each reconstruction from the mean, normalized to the average pixel value. Both low-curvature sets, panels (c) and (d), show significant structure, indicating that the algorithms consistently introduce artifacts into the image, such as errors in the reconstruction of the sample shape. The high-curvature reconstructions (a) and (b) show what appears to be an entirely random deviation from the mean, indicating less tendency to introduce artifacts.

VI. CONCLUSION

This paper has explored the role of curvature and coherence on coherent diffractive imaging. The reconstruction algorithms published so far all implicitly assume perfect spatial and temporal coherence. We have found that experimental departures from this limit have a detrimental effect on the reliability of the reconstructions and on the ability of the iterative reconstruction algorithms to successfully converge to a solution.

We have also shown that the introduction of curvature into the illumination significantly increases the reliability of the reconstruction and gives rise to an experimental configuration that is able to tolerate significantly lower levels of spatial coherence.

ACKNOWLEDGMENTS

The authors acknowledge the Australian Research Council Centre of Excellence and Fellowship programs. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

[1] R. Gerchberg and W. Saxton, Optik (Stuttgart) **35**, 237 (1971).
[2] J. Fienup, Appl. Opt. **21**, 2758 (1982).
[3] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, Nature (London) **400**, 342 (1999).
[4] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. H. Spence, Phys. Rev. B **68**, 140101(R) (2003).
[5] H. Chapman, A. Barty, S. Marchesini, A. Noy, S. Hau-Riege, C. Cui, M. Howells, R. Rosen, H. He, J. Spence, *et al.*, J. Opt. Soc. Am. A **23**, 1179 (2006).
[6] I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, and J. A. Pitney, Phys. Rev. Lett. **87**, 195505 (2001).
[7] D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A. Neiman, *et al.*, Proc. Natl. Acad. Sci. U.S.A. **102**, 15343 (2005).
[8] J. Miao, K. Hodgson, T. Ishikawa, C. Larabell, M. LeGros, and Y. Nishino, Proc. Natl. Acad. Sci. U.S.A. **100**, 110 (2003).
[9] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, Nature (London) **406**, 752 (2000).
[10] J. C. H. Spence, U. Weierstall, and M. Howells, Ultramicroscopy **101**, 149 (2004).
[11] G. J. Williams, H. M. Quiney, B. B. Dhal, C. Q. Tran, K. A. Nugent, A. G. Peele, D. Paterson, and M. D. de Jonge, Phys. Rev. Lett. **97**, 025506 (2006).
[12] K. Nugent, A. Peele, H. M. Quiney, and H. N. Chapman, Acta Crystallogr., Sect. A: Found. Crystallogr. **61**, 373 (2005).
[13] J. Fienup, Opt. Lett. **3**, 27 (1978).
[14] R. Bates, Optik (Stuttgart) **61**, 247 (1982).
[15] G. Williams, H. Quiney, A. Peele, and K. Nugent, Phys. Rev. B **75**, 104102 (2007).
[16] I. McNulty, A. Khounsary, Y. Feng, Y. Qian, J. Barraza, C. Benson, and D. Shu, Rev. Sci. Instrum. **67**, 3372 (1996).
[17] C. Q. Tran, A. G. Peele, A. Roberts, K. A. Nugent, D. Paterson, and I. McNulty, Opt. Lett. **30**, 204 (2005).
[18] D. Paterson, B. E. Allman, P. J. McMahon, J. Lin, N. Moldovan, K. A. Nugent, I. McNulty, C. T. Chantler, C. C. Retsch, T. H. K. Irving, *et al.*, Opt. Commun. **195**, 79 (2001).