Fresnel coherent diffractive imaging: treatment and analysis of data

G J Williams¹,²,³,⁴, H M Quiney¹,³, A G Peele²,³ and K A Nugent¹,³

¹ School of Physics, The University of Melbourne, Melbourne, VIC, Australia
² Department of Physics, La Trobe University, Bundoora, VIC, Australia
³ Australian Research Council Centre of Excellence for Coherent X-ray Science, School of Physics, The University of Melbourne, Melbourne, VIC, Australia

E-mail: gjwillms@slac.stanford.edu

New Journal of Physics 12 (2010) 035020 (18pp)
Received 28 November 2009
Published 31 March 2010
Online at http://www.njp.org/
doi:10.1088/1367-2630/12/3/035020

Abstract. Fresnel coherent diffractive imaging (FCDI) is a relatively recent addition to the suite of imaging tools available at third generation x-ray sources. It shares the strengths of other coherent diffractive techniques: resolution limits that are independent of focusing optics, single-plane measurement and high dose efficiency. The more challenging experimental geometry and detailed reconstruction algorithms of FCDI provide enhanced numerical stability and convergence properties to the iterative algorithms commonly used. Experimentally, a diverging beam is utilized, which facilitates sample alignment and allows the imaging of extended samples. We describe the underlying physics and assumptions that give rise to the FCDI iterative reconstruction algorithms, as well as their implications for the design of a successful FCDI experiment.

4 Current address: SLAC National Accelerator Laboratory, Menlo Park, CA, USA.
5 Author to whom any correspondence should be addressed.
1. Introduction

Coherent diffractive imaging (CDI) is a method whereby one plane of diffraction data may be transformed into an image of the sample. The practical limit to the resolution of images so acquired is based on the numerical aperture—the highest angle at which scattered probe particles can be measured—of the detector, not on the image-forming lens. This is generally achieved by an iterative algorithm that forces the wave leaving the object to be mutually consistent with its diffraction pattern and the experimenter’s knowledge of the sample.

The method was first demonstrated by Miao et al [1] on an Au test object and has since been applied to a wide range of samples, including biological materials [2]–[5] and foams [6]. While these demonstrations have used plane-wave illumination in a forward-scattering geometry, the method is not limited in this way; it has also been implemented in a Bragg-angle geometry [7]–[9], in a scanning-mode geometry [10], with partially coherent illumination [11] and with phase structure in the illumination of the sample [12]. The focus of this paper will be on the application of the described technique in the last case: Fresnel CDI (FCDI).

FCDI has a number of intriguing properties. Firstly, it can be shown that, in the absence of noise, there exists only one solution [13] to the FCDI problem. This is in contrast to the plane-wave case, where there is a translational invariance of the sample underlying the emergence of the far-field diffraction pattern. Practically, this means that the ambiguities identified by Bates [14] are not present and that—even in the presence of noise—iterative algorithms converge quickly and consistently. Secondly, when implemented as a finite, diverging beam illuminating a sample, the only region of a large sample contributing to the far-field intensity falls within the beam. This means that a region of interest on a large sample can be easily...
accessed [15, 16]. Finally, in certain instances, it is possible to derive quantitative information about the thickness [17] of the sample from the recovered ‘exit surface wave’ (ESW).

In this paper, we aim to explain the concepts underlying the FCDI approach and give a detailed account of the data treatment procedures used in and an interpretation of the results in [12].

2. Theory

2.1. Propagation of light and the origin of the ESW

An explicit manipulation of the wave leaving the sample lies at the heart of FCDI. Whereas, in plane-wave CDI, it is often sufficient to use only the Fourier transform, \( \mathcal{F} \), to propagate between the sample and detector planes, the phase structure introduced into the illuminating wave requires a more careful treatment of this quantity in FCDI. We consider the paraxial propagation of a monochromatic wave of wavelength \( \lambda \), \( \psi(r_n, z_n) \), contained in a plane at \( z_n \) to some other plane at \( z_m \) to be defined by

\[
\psi(r_m, z_m) = \frac{-i}{\lambda z_{nm}} \exp \left( \frac{i2\pi z_{nm}}{\lambda} \right) \exp \left( \frac{i\pi r_n^2}{\lambda z_{nm}} \right) \exp \left( \frac{-2\pi i r_n \cdot r_m}{\lambda z_{nm}} \right) \times \int dr_n \psi(r_n, z_n) \exp \left( \frac{i\pi r_n^2}{\lambda z_{nm}} \right) \exp \left( \frac{-2\pi i r_n \cdot r_m}{\lambda z_{nm}} \right) = \Theta(r_m, z_{nm}) \times \mathcal{F} \left[ \mathcal{E}(r_n, z_{nm}) \psi(r_n, z_n) \right],
\]

where, in the final line, we have introduced the functions \( \Theta(r_n, z_{nm}) \) and \( \mathcal{E}(r_n, z_{nm}) \) containing the prefactors to the integral and the spherical phase term in the integrand, respectively, and recognizing that this integral can be expressed as a Fourier transform. Henceforth, the sample plane will be indicated with s, the detector plane with d and the focal plane with f, e.g. \( z_{fd} \) denotes the focus-to-detector distance. We note that the primary departure of the propagation in FCDI to that of plane-wave CDI is that in the latter \( \mathcal{E}(r_d, z_{sd}) = 1 \), as \( z_{sd} = \infty \) is normally assumed.

The physical quantity of interest in FCDI is the transmission function of the sample. In the case of thin, singly scattering objects, the projection approximation [19] holds, i.e. the thickness, \( \tau(r) \), of the sample must obey the following relation with respect to the desired resolution \( \delta \) and the wavelength

\[
\tau(r) \ll \frac{2}{k \theta_{\text{max}}^2}
\]

\[
\ll \frac{4\delta^2}{\pi \lambda},
\]

where \( \theta_{\text{max}} \) is the largest angle to which significant scattering is measured and assumed to be small. For a two-dimensional (2D) measurement, we recognize that the physical ESW leaving an object is related to the incident wave \( \psi(r, z_n) \) and the object’s 3D refractive index map, \( n(r, z) \), projected onto an exit plane by \( \psi_{\text{ESW}}(r, z_s) = \mathcal{F}(r_s) \psi(r_s, z_s) \), where

\[
\mathcal{F}(r) = \exp \left[ \frac{i2\pi \lambda}{\lambda} \int_{\text{sample}} dz n(r, z) \right] \approx \exp \left\{ \frac{i2\pi }{\lambda} [1 - \delta(r) + i\beta(r)] \tau(r) \right\}.
\]
These three approximations—paraxiality, single scattering and projection—are normally comfortably accommodated by x-ray FCDI experiments.

In FCDI, neglecting intensity variations in the pupil function, one introduces a wave of the form
\[ \psi(r_s, z_s) \approx \exp \left( -i\pi r_s^2 / \lambda z_s \right) \]
to the sample. In the case where this is exactly true, the input of the propagation operator in equation (1) becomes \( \Xi(r_s, z_{sd})\mathcal{F}(r_s) \), where \( (z_{sd})^{-1} = (z_{fd}^{-1} - z_{fs}^{-1}) \) is the difference in the reciprocals of the focus-to-detector and focus-to-sample distances. This important realization allows the use of fast Fourier transform (FFT) in the implementation of the algorithm of section 2.2.2.

2.2. Algorithms

Generally speaking, the goal of CDI is to recover the complete complex wave that has given rise to a measured diffraction pattern in order to arrive at an image of the sample, which is a manifestation of \( n(r, z) \) or its projection. This is usually accomplished by means of an iterative process, which propagates the trial wave—called the ‘iterate’—between the sample and detector planes. In each plane, the iterate is constrained to display any properties of the object that may be known \textit{a priori}. The most obvious constraint is that the square modulus of the iterate in the detector plane must match the measured intensity distribution. This is commonly called a ‘modulus’ constraint [20]. We call the result of applying the modulus constraint to an iterate as the current ‘estimate’ of the wave. Due to the sampling requirement on the intensity distribution, the sample—or the portion of the sample contributing to the diffraction—must be finite in extent. This gives rise to the commonly applied ‘support’ constraint in the sample plane. Numerous additional constraints have been suggested; however, we will here restrict ourselves to these two.

The iterative algorithms common in CDI can be compactly expressed in a projector notation [21, 22], where an iterate, \( \rho_k \), is regarded as a vector in an \( N \)-dimensional vector space. The task of enforcing the constraints is assigned to projection operators, e.g. we call the modulus constraint \( \pi_m \) and the support constraint \( \pi_s \). For simplicity, these operators act explicitly in the sample plane, so that the propagation must be built into the modulus constraint:
\[ \pi_m = \mathcal{F}^{-1} \tilde{\pi}_m \mathcal{F}, \]
where the tilde indicates the projector as operating in the Fourier space. We will obey the following convention: \( \rho_k \) is the iterate and \( \pi_m \rho_k \) is the estimate of the target wave on the \( k \)th iteration.

Using this notation, we can state the algorithms most commonly used in FCDI. The \( k \)th iterate subjected to error reduction (ER) [20, 23] is
\[ \rho_{k+1} = (\pi_s \pi_m) \rho_k, \]
while for hybrid input/output (HIO) [23],
\[ \rho_{k+1} = [1 + (1 + \beta) \pi_s \pi_m - \pi_s - \beta \pi_m] \rho_k, \]
where \( \beta \) is the familiar, scalar HIO parameter. The experimental demonstration presented in sections 4 and 5 was achieved using only ER and a support constraint that was successively updated, similar to the shrinkwrap algorithm of Marchesini \textit{et al} [24].

The support generated here was generated by a two-step method: (i) find all regions within the recovered wave that have magnitude above a certain value and assign these to be unity in the support array; and (ii) convolve the support array with a Gaussian function of narrow width. The operation \( \pi_s \) is then accomplished by multiplying this array by the estimate.
2.2.1. The beam recovery algorithm. The wave illuminating the object must be known in order to correctly separate its features from those of the sample’s ESW. We achieve this by means of the algorithm demonstrated in Quiney et al [25]. This algorithm requires that one make a measurement of the far-field intensity arising from a lens and know the size of the pupil function of the lens the focal length of the lens, and the focal-plane-to-detector distance.

With this information in hand, we can iteratively recover the phase of the diverging beam in the plane of the detector, which can, of course, be propagated to any other plane numerically. This procedure requires propagation between three planes, each perpendicular to the propagation direction: the lens plane, \( z_l \); the focal plane, \( z_f \); and the detector plane, \( z_d \). Unfortunately, the propagator in the \( z_d \)- and \( z_l \)-planes oscillates too quickly to be properly sampled in the discrete array. This is overcome by recognizing that the wave is formed through the multiplication of a slowly varying component with a quickly varying one. By analytically propagating the quickly varying component and numerically propagating the other, the iterative algorithm retains its phase-retrieval properties and remains numerically stable.

Let the wave in the lens plane be \( \psi(r_l, z_l) = P(r_l) \exp(i \pi r_l^2/\lambda z_l) \), where \( P(r_l) \) is the complex pupil function of the lens and \( z_l \) is its focal length. Similarly, \( \psi(r_d, z_d) = \tilde{P}(r_d) \exp(-i \pi r_d^2/\lambda z_d) \) is the wave in the detector plane. The steps in the algorithm are

1. Propagate from \( z_l \) to \( z_d \): \( \psi_k(r_l, z_l) = \Theta(r_l, z_d) \hat{\mathcal{F}} \{ P_k(r_l) \} \)
2. Propagate from \( z_d \) to \( z_l \): \( \tilde{P}_k(r_d) = -i \exp(i 2 \pi z_d/\lambda) \hat{\mathcal{F}} \{ \psi_k(r_l, z_l) \xi(r_l, z_l) \} \)
3. Apply modulus constraint: \( \tilde{P}_k(r_d) = \tilde{P}_{\exp}(r_d) \times \frac{\tilde{P}_k(r_d)}{|\tilde{P}_k(r_d)|} \)
4. Propagate from \( z_d \) to \( z_l \): \( \psi_k(r_l, z_l) = \Theta(r_l, z_d) \hat{\mathcal{F}} \{ \tilde{P}_k(r_d) \} \)
5. Propagate from \( z_l \) to \( z_d \): \( P'_k(r_l) = -i \exp(i 2 \pi z_d/\lambda) \hat{\mathcal{F}} \{ \psi_k(r_l, z_l) \xi(r_l, z_d) \} \)
6. Enforce support constraint on pupil: \( P_{k+1}(r_l) = \pi_s P'_k(r_l) \)
7. Use \( P_{k+1}(r_l) \) in 1

Once convergence has been obtained, the algorithm is halted at step 2, which constitutes the best estimate of the pupil function in the detector plane. The estimate of the wave in the detector plane on the \( N \)th iteration is \( \psi(r_d, z_d) = \tilde{P}_N(r_d) \exp(-i \pi r_d^2/\lambda z_d) \).

To establish a connection to the operator notation, we introduce projectors for the various multiplications in the algorithm: \( \pi_{\psi}^{nm} \) for the multiplication by \( \Theta(r_m, z_{nm}) \) in step 1, \( \pi_{\xi}^{nm} \) for that by \( \xi(r_m, z_{nm}) \) and \( \pi_p^{nm} \) for that by \( -i \exp(i 2 \pi z_{nm}/\lambda) \). The \((k+1)\)th iterate is then

\[
P_{k+1} = \left[ \pi_s \pi_p^{\Psi} \hat{\mathcal{F}} \pi_{\xi}^{\Psi} \pi_{\psi}^{\Psi} \hat{\mathcal{F}} \pi_m \pi_p^{f} \hat{\mathcal{F}} \pi_{\xi}^{f} \hat{\mathcal{F}} \right] P_k.
\] (7)

2.2.2. The ESW recovery algorithm. While one might be tempted to simply include a propagation to the sample plane in the algorithm of section 2.2.1, one must remember that \( P(r_l) \) is constrained to be slowly varying, which cannot be assumed in the case of the final estimate of the sample’s ESW. Therefore, it becomes necessary to implement the algorithm in two parts: first recover the phase of the incident illumination and then recover the phase of the sample’s diffraction. As described at the end of section 2.1, a ‘curvature component’ must be reintroduced into the output of the beam recovery algorithm before using it in the recovering of the ESW.
The algorithm for recovering the ESW itself was described by Williams et al [12]. There are a few key differences between the Fresnel and plane-wave variants of CDI. Among these is the question of how the incident beam is treated in each case. In plane-wave CDI, the sample is small compared to the extent of the beam, but since the interference between the beam and light scattered by the sample is occluded by a beam stop—which protects the detector from the undiffracted beam—a transmission function can be easily recovered using Babinet’s principle, i.e. for regions sufficiently far away from the origin, \( \mathcal{F}[f(r)] = \mathcal{F}[1 - f(r)] \). By contrast, in FCDI, we measure the intensity arising from \( \mathcal{I}(r_s)\psi(r_s, z_s) \) propagated to the detector plane.

An important consideration is that while the beam may be used to define a region of interest in the sample [15, 16], the only requirement is that the object be illuminated with sufficient phase curvature [26]. In other words, the beam may be so large in the plane of the sample that a numerical propagation back to that plane will result in the overfilling of that unit of the discrete Fourier space.

A further consideration concerns the ability to impose the support constraint. In principle, one wishes to apply this to the transmission function; however, this requires a division operation in the plane of the sample. In the early stages of the reconstruction, this can prove to be quite unstable numerically, as many regions of the illumination may approach zero magnitude. To address these concerns, it is useful to perform a subtraction of the known illuminating wave from the iterate of the far-field ESW. In essence, this means that the quantity dealt with in the iterative scheme is \( \mathcal{I}(r_s) - 1 \) \( \psi(r_s, z_s) \). To account for this subtraction in the context of ER and HIO, we introduce a new projection operator: \( \tilde{\pi}_{WF} \). The modified algorithm is then

\[
\rho_{k+1} = \pi_s \pi_m \rho_k = \pi_s \mathcal{F}^{-1} \tilde{\pi}_{WF} \pi_m \mathcal{F} \rho_k.
\]

### 3. Experiment

The Fresnel CDI experimental geometry used here is superficially very similar to that of a scanning x-ray microscope, as can be seen in figure 1. There are three notable differences: the order-sorting aperture of the zone plate was placed in the focal plane and made as small as possible, the sample was displaced to some defocus distance where significant curvature has developed in the illuminating field emerging from the focus, and a 2D integrating detector was used. In this section, we describe the rationale for choosing the stated experimental parameters and detail the data collection.

#### 3.1. Important criteria

We introduce two useful heuristics that have a strong influence on FCDI experiment design, the far-field condition

\[
1 \gg \frac{d^2}{\lambda L}
\]

and the discrete Fourier transform (DFT) relation for pixel size in the sample and detector planes

\[
\Delta_s = \frac{\lambda L}{N \Delta_d}.
\]
where $\lambda$ is the wavelength of light, $L$ is the sample–detector distance, $d$ is the linear size of the sample, $\Delta_i$ is the linear pixel size in two planes linked by the DFT and $N$ is the side length of the discrete array.

By satisfying equation (9), we are assured that the sampling requirement on the detector-plane ESW can be estimated by equation (10). This is important because we must sample the intensity distribution due to the ESW at better than the Nyquist frequency. If the far-field condition is not met, one must be very careful when estimating the sampling requirements, as the intensity in near- or mid-field planes may require significantly more strenuous sampling due to the evolution of the wave. For example, if we wish to image an object with $d = 10 \, \mu m$ whose diffraction pattern fills a detector with $1000 \times 1000$, $20 \, \mu m$ pixels, we have a bound, i.e. $\lambda L \gg 10^{-10}$, which limits the resolution that can be achieved, because $\Delta_i \gg 10^{-10} \, m^2/0.02 \, m \approx 5 \, nm$.

3.2. Stability

We introduce a functional of the transmission function that is the square modulus of the sample-plane wave propagated to the detector plane: $\mathcal{I}[\mathcal{T}(r_s)] = |\psi(r_d, z_d)|^2$, where $\psi(r_d, z_d)$ is related to $\mathcal{T}(r_s)$ by equation (1). To express the change in the diffraction pattern by a small movement $\delta r$, we substitute $r_s + \delta r$ into equation (1) and discover that

$$\mathcal{I}[\mathcal{T}(r_s + \delta r)] = \left| \psi \left( r_d + \frac{z_{sd}}{R} \delta r - \delta r, z_d \right) \right|^2,$$

where $R$ is the radius of curvature of the field in the plane of the sample and $z_{sd}$ is the sample–detector distance. To recover the Fraunhofer limit, we note that both $R$ and $z_{sd}$ must be infinite. Only in this limit is the imaging geometry truly transversely invariant. The impact
upon CDI is immediately obvious: the effect of a moving sample is to ‘blur’ the diffraction pattern in the detector plane with a scaling parameter that goes as the magnification of the imaging system.

In FCDI, this poses a serious problem. As a consequence of the requirement that the beam at the sample have a Fresnel number that is greater than $5 \times 26$, the magnification factors for these experiments are typically 500–1000. Pragmatically, the sample stability requirement reduces to a question of how much blurring of the diffraction pattern can be tolerated in the iterative scheme or compensated for after-the-fact. As a heuristic, one might require that the blurring due to motion be much less than the point spread function of the CCD, which is typically $20 \mu m$ or more. The geometry of the experiment described here, which ultimately gives a resolution $24 \text{ nm}$, provides a bound:

$$\delta r \ll \left( \frac{z_{sd}}{R} - 1 \right) \times 20 \mu m \ll \left( \frac{0.5 \text{ m}}{1.0 \text{ mm}} - 1 \right) \times 20 \mu m \ll 40 \text{ nm}. \quad (12)$$

Although this is quite a stringent requirement, we note that it is not directly dependent upon the resolution of any image derived from the final reconstruction of the ESW.

Equation (11) applies equally to plane-wave CDI. In order for the experimental geometry to be transversely invariant, $z_{sd}/R \simeq 1$ and $z_{sd}$ must satisfy the far-field condition given by equation (9). These two conditions are probably not met in many plane-wave CDI experiments, but equation (12) indicates that sample motion would become problematic only as its magnitude approaches the point spread function of the detector.

### 3.3. Equipment

The iterative algorithms used in CDI typically depend upon the coherent propagation of light, as described in section 2. As a result, most such experiments are conducted at third-generation x-ray sources. The experiment described here was conducted at the Advanced Photon Source on beamline 2-ID-B [27], which provided 1.8 keV photons using a spherical grating monochromator. In this beamline geometry, the monochromator’s exit slit may be closed to improve the transverse coherence of the beam. This highly coherent, quasi-monochromatic beam exits the transport pipe by traversing a $200 \text{ nm}$ thick, $700 \mu m^2$ Si$_3$N$_4$ window. This window is approximately $8 \text{ m}$ from the exit slit.

A Fresnel zone plate was used to create the diverging beam required by FCDI. In this case, the zone plate had diameter $160 \mu m$ and an outermost-zone radius of $50 \text{ nm}$—which, in conjunction with a $20 \mu m$ central stop and $15 \mu m$ order-sorting aperture, gave rise to a diverging beam of sufficient quality to be used in this experiment. An identical Si$_3$N$_4$ window was used on the detector’s evacuated flight path. By placing the window very close to the sample, its scattering was minimally affected by absorption or rescattering, while the total air path seen by the beam was limited to about $1.5 \text{ cm}$. The detector used for this experiment was a Princeton Instruments direct-read CCD with $24 \mu m$, square pixels in an array of $1340 \times 1300$. The CCD was thermoelectrically cooled against a liquid N$_2$ bath to $-50^\circ \text{C}$. With a sample of largest dimension $8 \mu m$ and the detector $0.6 \text{ m}$ from the sample plane, the conditions described by equations (9) and (10) were met.

The data were collected using CCDImageServer software developed at the APS. The program can be configured to poll and record any process variables available to the EPICS-based control system. This facilitates the later analysis and troubleshooting of any inconsistent data.

*New Journal of Physics 12 (2010) 035020 (http://www.njp.org/)*
The sample used for this experiment was seven nested chevrons patterned in Au by lithography. One of the chevrons had side length of approximately 8 µm, while the others were contained in a square with ∼1 µm side length. The Au was approximately 150 nm thick and supported on a thick Si$_3$N$_4$ membrane.

4. Result

Broadly, two types of measurement are required. The ‘white-field’ data are collected without a sample in the path of the beam and are used to recover the complex illuminating wave. The ‘ESW’ data are collected when the sample is placed in the x-ray beam and are used in the iterative scheme in conjunction with the recovered complex illumination to phase the ESW in the detector.

4.1. Properties of the white-field data

These data are normally taken at intervals throughout the longer sample measurement, so that a change in the illumination does not unduly interfere with data analysis. Figure 2(a) shows a typical far-field ‘doughnut’ arising from the first-order focus of a zone plate. We note that this is essentially a low-resolution image of the pupil function, complete with the shadow due to the central stop. The only scattering outside this donut is due to the order-sorting aperture or the air path, i.e. we expect no features smaller than the diffraction-limited spot of the zone plate to be present in the reconstruction. This is a reflection of the near-ideal, thin-lens behavior of the optical configuration.
4.2. Properties of the ESW data

Examining figure 3(a), there are two obvious differences from the white-field data of figure 2(a). First, a ‘shadow’ of the sample appears inside the beam due to the zone plate. If one were to truncate the data at the edge of the beam, this would form an in-line-holography data set. Experimentally, the shadow is quite useful, as it allows one to place the sample in a preferred region of the illumination. For example, it is normally wise to avoid placing an interesting region of the sample on the optical axis, as the illuminating wave is changing quite rapidly and its reconstruction tends to work less well there. As we will describe in section 5.1, this shadow image provides direct input on the motion of the sample over time. Figure 3(a) shows the treated intensity used for the lower quality reconstruction, whereas figure 3(b) shows the intensity used for the highest quality reconstruction.

The second notable departure is the presence of strong x-ray scattering at high angles. This is the primary indication that the reconstructed ESW will yield high-resolution sample information. A close examination of the fringes at high angle will also reveal how well sampled the diffraction pattern is and give an estimate of the visibility, which is a critical indicator of whether the illuminating wave is sufficiently coherent to be accurately propagated by the method in section 2.

It is useful to perform a ‘magnification check’ immediately after positioning the sample in the illumination. This allows the sample–focal plane distance to be measured under the assumptions of geometrical optics. This is accomplished by collecting one frame of data, moving a motor perpendicular to the beam by a known amount—for example, 1 μm—and acquiring one more frame. The motion of the shadow image in the detector plane in microns...
is then divided by the distance moved by the sample. This yields the magnification, which can be used in the familiar similar-triangle argument to relate the sample–focus distance to the detector–focus distance. This latter quantity can be estimated quite easily as the focal length of the lens is well known and the magnification is determined by the ratio of the size of the beam in the detector to the size of the known pupil function.

5. Discussion

5.1. Data handling

Due to the quickly varying phase structure in the illuminating wave, the FCDI is quite sensitive to the relative positioning of the sample and the optic. During a 4 h series of acquisitions, this displacement is often much greater than the resolution desired from the reconstructed ESW. The cross-correlation of data frames $I^{(i)}$ and $I^{(j)}$ is

$$R_{ij} = \frac{\sum_{n=1}^{N} I^{(i)}(x_n) I^{(j)}(x_n)}{\sum_{n=1}^{N} I^{(i)}(x_n) \sum_{n=1}^{N} I^{(j)}(x_n)}, \quad (13)$$

where $x_i$ are the collection of $N$ pixels in the array. To pre-select a subset of the data that is self-consistent, $R_{ij}$ is calculated with respect to one frame across the data set. A typical cross-correlation plot is shown in figure 4. In the data treatment below, we will demonstrate reconstructions with two subsets of the whole data set: figure 3(a) includes all frames with $R > 0.9959$ and figure 3(b), containing a factor of three greater frames, with $R > 0.9957$. [Figure 4. The cross correlation of 800 frames of FCDI data against one reference frame collected at 4311 s. The gap starting at 3000 s reflects a period of instability in the apparatus or beamline. Data were collected only sporadically during this time. There are two plateaus in the plot, corresponding to periods of stability in the experimental system. That the plateau at an earlier time has a lower average $R$-value is an indication that there was a net change in the position of the sample or illumination.]
These values seem quite high, but it is important to remember that no background subtraction has yet been performed.

The data were collected frame by frame on the direct-read CCD discussed in section 3.3. The dark frames, $I_{i}^{\text{dark}}$, collected during the experiment were averaged and used in a background subtraction for each data frame, whether it contained ESW or WF-only scattering. The subtraction was carried out on a pixel-by-pixel basis according to

$$I_{\text{treat}}(x_i) = I(r)(x_i) - \alpha I_{i}^{\text{dark}}(x_i), \forall x_i.$$  \hfill (14)

The real number $\alpha$ is chosen so that the peak corresponding to the noise floor in the data is centered about zero in a histogram of $I_{\text{treat}}$. In general, the stochastic nature of the noise—whether it be CCD dark current or electronic noise—prevents this subtraction from being complete. We therefore establish a threshold, $\Lambda$, below which we assert that the readout of the CCD does not contain signal. This is implemented pixel by pixel according to

$$I_{\text{meas}}(x_i) = \begin{cases} I_{\text{treat}}(x_i), & \text{if } I_{\text{treat}}(x_i) \geq \Lambda \\ 0, & \text{otherwise.} \end{cases} \quad \forall x_i.$$  \hfill (15)

It is $I_{\text{meas}}(x_i)$ that is used in imposing the modulus constraint, $\pi_m$.

Figures 5 and 6 contain histograms of typical data before and after each stage of the procedure described above. In figure 5, we compare the histograms of a white field and a dark frame. The most common values in both frames are those that result from the detector itself, as evidenced by the tall peak at about 80 analogue-to-digital units (ADUs). Our aim is to reduce the effect of the detector on the distribution of energy in the frame. The value of $\alpha$ in equation (14) is determined by subtracting a dark frame from a data frame, calculating a new histogram and performing a fractional subtraction of the dark frame if the new histogram is not centered about
Figure 6. The histograms of one frame of data at each of the three stages described in section 5.1. For the basic subtraction, we see that the histogram is centered around zero, indicating that the uncertainty in the data is mostly random. Obviously, we cannot allow negative intensity data to be used, so all pixels with value below zero are set to zero (‘non-negative subtraction’). However, we find it helpful to set a higher threshold and so re-assign the value of all pixels below a threshold of 75 ADU to be zero.

zero ADUs. The red ‘+’ points in figure 6 show the histogram of \( I^{\text{treat}}(x_i) \) at this stage. The green ‘×’ and blue ‘∗’ symbols denote data thresholded at zero and 75 ADU, respectively. Once the subtraction is complete, the selected frames are summed.

The phase structure of the illuminating wave leads to the establishment of a well-defined optical axis. It is critical that the relationship between the optical axes of the experiment and the propagation are known and, in practice, it is easiest to simply center the data in the DFT array. This is most readily accomplished by applying one half-iteration of the algorithm to the data, which yields an estimate in the plane of the lens. The exterior boundary of the beam in these two planes should coincide, which provides a quick check on the centering of the data. Using this newly centered array, the ESW data are easily aligned. This operation normally produces an array that is no longer square. Here, we choose to pad the array with the value ‘−1’, as can be seen in figures 3(a) and (b), where the top and the right-hand side appear black. In the implementation of the iterative scheme, any pixel with this value in the modulus constraint is not constrained. In other words, these pixels are allowed to ‘float’, as we have no measurement of them. This is a compromise that permits the use of all measured data without imposing additional prior knowledge of the modulus.

5.2. Beam recovery

To recover the phase of the illuminating wave, we require accurate estimates of the lens–focus and focus–detector distances. We rely on our a priori knowledge of the optical system—the
zone plate diameter and outermost zone width, as well as the photon energy are known—and the magnification argument presented in section 4.2. We strictly enforce the known size of the zone plate through the application of the support constraint in the lens plane in the algorithm described in section 2.2.1.

The experimental parameters corresponding to the data here were

- wavelength: $\lambda = 6.78$ Å
- zone plate diameter: $D = 160 \mu$m
- outermost zone width: $\Delta r = 50$ nm
- focal length: $f = \frac{D \Delta r}{\lambda} = 11.8$ mm
- detector distance: $L = 0.509$ m.

In figure 2(b), we display the magnitude of the reconstructed, illuminating wave in the plane of the sample. This is the result of propagating the estimate that resulted from 300 iterations of the algorithm. It is immediately obvious that the wave overfills the DFT unit cell and, due to the periodic boundary conditions, ‘folds over’ on the edges of the discrete array. This is one of the problems that can be avoided by performing the subtraction of the complex illuminating wave as described in section 2.2.2.

5.3. ESW recovery

In addition to a knowledge of the geometry discussed above, we now require the result of the previous section to proceed with the reconstruction of the ESW. The phased far-field illumination is multiplied by an additional spherical phase factor so that propagation between the detector and sample planes may be accomplished by means of the FFT. The sample–focus distance is estimated, as discussed in section 3.3, to be 1.13 mm. This and other geometrical factors can be varied to choose the plane in which the sample-plane reconstruction is formed and it is often useful to do so within the experimental error of their measurement. With this information in hand, we apply the algorithm of section 2.2.2.

The estimate of the wave $\left[ \tilde{T}(r_s) - 1 \right] \psi(r_s, z_s)$ on the $k$th iteration can then be constrained by a support. The shadow image of the sample provides an immediate estimate of the fraction of the beam it occupies and its shape. In the current example, we chose to employ a support that starts with the size and shape estimated from the shadow image and gradually shrink it by thresholding at 11% of its value and convolving the result with a Gaussian function of variance 1 pixel. After 30 iterations, this procedure was halted and the support remained fixed for the next 270 iterations, although acceptable convergence was probably obtained after 100.

The magnitude of $\left[ \tilde{T}(r_s) - 1 \right] \psi(r_s, z_s)$ is shown for the two subsets of the data discussed above in figure 7. By accepting data with $R > 0.9959$, 89 of 800 frames are summed and reconstructed to yield figure 7(a), the magnitude, and 7(b), the phase. With $R = 0.9957$, 272 frames are summed and the magnitude of the recovered wave is shown in figure 7(c). Use of the latter data clearly results in an image that is of lower quality than the former. We attribute the difference between the two results to a change in the experimental geometry. The most likely cause is a slow drift of the sample with respect to the optic, as these are not thermally controlled. It is also possible that some property of the beamline or the source has changed, but these are not obvious from an examination of white-field data taken before and after the ESW data.

The reconstruction itself is of high quality. Interestingly, in addition to the expected chevron pattern, three other objects have appeared in the reconstructed wave. We believe these to be dust.
Figure 7. Images resulting from the FCDI experiment and reconstruction. (a) The magnitude of $|\mathcal{F}(r_s) - 1|\psi(r_s, z_s)$ as described in the text; (b) the phase of this quantity. The result of the more permissive selection criterion, $R > 0.9957$, is shown in (c). We can see that while (c) is more uniform than (a), it possesses less-sharp edges. To judge the appropriateness of the support, we examine the energy distributed outside the support region in the iterate, shown in (d), for the quantity in (a). From (a) and (b), we may derive $\mathcal{F}(r_s)$, whose magnitude and phase are shown in (e) and (f), respectively. All scales are linear. In (e), black and white represent complete opacity and transmission. An arbitrary constant, $3\pi/4$, has been added to the phase of the transmission function before plotting it in (f) to increase the contrast, but the scale is otherwise the same as in (b).
or other, weakly scattering contaminants on the sample. They are dimly visible in a scanning-
x-ray micrograph, but not present in a scanning-electron one. We assign a resolution to this
image of 24nm, which is derived from Abbe theory \cite{18} as described in \cite{12}.

It is instructive to ask what features of the estimated ESW are being suppressed by the
application of the support. For this purpose, we examine that portion of the estimate that falls
within the complement to the support. Figure 7(d) is the magnitude of this portion of the estimate
on the penultimate iteration of the procedure, whose result is shown in figure 7(a). The scale bar
shows that the largest single pixel in this array is 1% of the greatest value in the estimate. From
this we observe that the threshold has been set so high as to forbid the contribution of some
weakly scattering areas to the ESW on each iteration. In particular, the fiber along the upper
arm of the chevron was rendered discontinuous by the shrinking support, whereas some energy
is clearly present in the unconstrained estimate. This is consistent with figure 7(c), where these
objects have not been cut into by the thresholding, probably because the energy of the estimate
is more evenly distributed in this blurry image.

In order to make a connection between the object reconstructed using this approach and its
physical properties via the index of refraction, it is necessary to process the estimate to recover
$\mathcal{F}(\mathbf{r})$. The magnitude and phase of this quantity are shown in figures 7(e) and (f). The expected
result is that all areas outside the object of interest should have unit magnitude and zero phase
shift. In figure 7(e), all pixels of unity or greater are white. All areas outside the region occupied
by the illuminating wave, shown in figure 2(b), have no physical meaning. Similarly, near the
optical axis, which passes through the center of the array, the illumination approaches zero
and oscillates rapidly. In these two regions, the magnitude of the pixel may exceed unity by
two orders of magnitude; however, within the region well illuminated, the distribution of pixel
values outside the object narrowly peaks at unity. On this color scale, black would represent no
transmission and hence we see that the Au object appears gray and the weakly scattering objects
are difficult to discern. For ease of display in figure 7(f), we have added a phase of $3\pi/4$ because
the unaltered phase represented on the scale bar in figure 7(b) provides little contrast in print.

6. Conclusion

We have provided a detailed description of the algorithmic and experimental considerations
specific to FCDI. The technique has algorithmic and experimental advantages over its plane-
wave progenitor. The uniqueness relation \cite{13} between the detector-plane wave and its
diffraction pattern manifests itself in the fast and reliable convergence of the iterative algorithms.
The requirement for achieving these benefits is that one must understand the detailed nature of
the illuminating wave and the experimental geometry.

Because the wave illuminating the sample is diverging, the intensity due to the incident
radiation is spread over a large region of the detector. This has two practical benefits: there is no
missing data as the experiment does not require a beam stop to protect the 2D detector, and the
interference between the in-line holographs can be sampled in the detector. By establishing an
optical axis in the experiment, one destroys the translational invariance enjoyed by plane-wave
CDI. We believe that this is a small cost for the benefits provided by the experimental geometry.

The underlying principle of FCDI—that one should exploit explicit knowledge of the wave
incident upon the sample and the experimental geometry—is likely to become increasingly
important in CDI as new sources, such as x-ray-free-electron lasers \cite{28,29}, become
available.

\newblock New Journal of Physics 12 (2010) 035020 (http://www.njp.org/)
Acknowledgments

We thank Ian McNulty, Brian Abbey, Mark Pfeifer, David Paterson and Martin DeJonge for useful discussions and assistance with experimental work. All the authors of this work acknowledge the support of the Australian Research Council Centres (ARC) of Excellence program. KAN acknowledges the support of an ARC Federation Fellowship. We acknowledge travel funding provided by the International Synchrotron Access Program (ISAP) managed by the Australian Synchrotron. The ISAP is funded by a National Collaborative Research Infrastructure Strategy grant provided by the Federal Government of Australia Use of the Advanced Photon Source at Argonne National Laboratory was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. DE-AC02-06CH11357.

References

[1] Miao J, Charalambous P, Kirz J and Sayre D 1999 Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens Nature 400 342–4
[2] Miao J, Hodgson K O, Ishikawa T, Larabell C A, LeGros M A and Nishino Y 2003 Imaging whole Escherichia coli bacteria by using single-particle x-ray diffraction Proc. Natl Acad. Sci. USA 100 10–2
[3] Shapiro D et al 2005 Biological imaging by soft x-ray diffraction microscopy Proc. Natl Acad. Sci. USA 102 15343–6
[4] Williams G J et al 2008 High-resolution x-ray imaging of plasmodium falciparum-infected red blood cells Cytometry A 73 949–57
[5] Nishino Y, Takahashi Y, Imamoto N, Ishikawa T and Maeshima K 2009 Three-dimensional visualization of a human chromosome using coherent x-ray diffraction Phys. Rev. Lett. 102 018101
[6] Barty A et al 2008 Three-dimensional coherent x-ray diffraction imaging of a ceramic nanofoam: determination of structural deformation mechanisms Phys. Rev. Lett. 101 055501
[7] Robinson I K, Vartanyants I A, Williams G J, Pfeifer M A and Pitney J A 2001 Reconstruction of the shapes of gold nanocrystals using coherent x-ray diffraction Phys. Rev. Lett. 87 195505
[8] Williams G J, Pfeifer M A, Vartanyants I A and Robinson I K 2003 Three-dimensional imaging of microstructure in Au nanocrystals Phys. Rev. Lett. 90 175501
[9] Pfeifer M A, Williams G J, Vartanyants I A, Harder R and Robinson I K 2006 Three-dimensional mapping of a deformation field inside a nanocrystal Nature 442 63–6
[10] Thibault P, Dierolf M, Menzel A, Bunk O, David C and Pfeiffer F 2008 High-resolution scanning x-ray diffraction microscopy Science 321 379–82
[11] Whitehead L W, Williams G J, Quiney H M, Vine D J, Dilanian R A, Flewett S, Nugent K A, Peele A G, Balaur E and McNulty I 2009 Diffractive imaging using partially coherent x-rays Phys. Rev. Lett. 103 243902
[12] Williams G J, Quiney H M, Dhal B B, Tran C Q, Nugent K A, Peele A G, Paterson D and de Jonge M D 2006 Fresnel coherent diffractive imaging Phys. Rev. Lett. 97 025506
[13] Pitts T A and Greenleaf J F 2003 Fresnel transform phase retrieval from magnitude IEEE Trans. Ultrason. Ferroelectr. Freq. Control 8 1035–45
[14] Bates R H T 1982 Fourier phase problems are uniquely solvable in more than one dimension. I: Underlying theory Optik 61 247–62
[15] Abbey B, Nugent K A, Williams G J, Clark J N, Peele A G, Pfeifer M A, de Jonge M and McNulty I 2008 Keyhole coherent diffractive imaging Nat. Phys. 4 394–8
[16] Abbey B, Williams G J, Pfeifer M A, Clark J N, Putkonen C T, Torrance A, McNulty I, Levin T M, Peele A G and Nugent K A 2008 Quantitative coherent diffractive imaging of an integrated circuit at a spatial resolution of 20 nm Appl. Phys. Lett. 93 214101
Clark J N, Williams G J, Quiney H M, Whitehead L, de Jonge M D, Hanssen E, Altissimo M, Nugent K A and Peele A G 2008 Quantitative phase measurement in coherent diffraction imaging Opt. Express 16 3342–8

Born M and Wolf E 1999 Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light 7th edn (Cambridge: Cambridge University Press)

Berry M V 1971 Diffraction in crystals at high energies J. Phys. C: Solid State Phys. 4 697–722

Gerchberg R W and Saxton W O 1972 A practical algorithm for the determination of phase from image and diffraction plane pictures Optik 35 237–46

Bauschke H H, Combettes P L and Luke D R 2002 On the structure of some phase retrieval algorithms Proc. IEEE Int. Conf. Image Processing vol 2 pp 841–4

Elser V 2003 Phase retrieval by iterated projections J. Opt. Soc. Am. A 20 40–55

Fienup J R 1982 Phase retrieval algorithms: a comparison Appl. Opt. 21 2758–69

Marchesini S, He H, Chapman H N, Hau-Riege S P, Noy A, Howells M R, Weierstall U and Spence J C H 2003 X-ray image reconstruction from a diffraction pattern alone Phys. Rev. B 68 140101

Quiney H M, Peele A G, Cai Z, Paterson D and Nugent K A 2006 Diffractive imaging of highly focused x-ray fields Nat. Phys. 2 101–6

Quiney H M, Nugent K A and Peele A G 2005 Iterative image reconstruction algorithms using wave-front intensity and phase variation Opt. Lett. 30 1638–40

McNulty I, Khounsary A, Feng Y P, Qian Y, Barraza J, Benson C and Shu D 1996 A beamline for 1–4 keV microscopy and coherence experiments at the advanced photon source Rev. Sci. Instrum. 67 3372–2

Feldhaus J, Arthur J and Hastings J B 2005 X-ray free-electron lasers J. Phys. B: At. Mol. Opt. Phys. 38 S799–819

Chapman H N et al 2006 Femtosecond diffractive imaging with a soft-x-ray free-electron laser Nat. Phys. 2 839–43