Near-field limitations of Fresnel-regime coherent diffraction imaging

Benjamin A. Pound, John L. Barber, Kimberly Nguyen, Matthew C. Tyson, and Richard L. Sandberg*
Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
(Received 27 October 2016; revised manuscript received 7 July 2017; published 4 August 2017)

Coherent diffraction imaging (CDI) is a rapidly developing form of imaging that offers the potential of wavelength-limited resolution without image-forming lenses. In CDI, the intensity of the diffraction pattern is measured directly by the detector, and various iterative phase retrieval algorithms are used to "invert" the diffraction pattern and reconstruct a high-resolution image of the sample. However, there are certain requirements in CDI that must be met to reconstruct the object. Although most experiments are conducted in the "far-field"—or Fraunhofer—regime where the requirements are not as stringent, some experiments must be conducted in the "near field" where Fresnel diffraction must be considered. According to the derivation of Fresnel diffraction, successful reconstructions can only be obtained when the small-angle number, a derived quantity, is much less than one. We show, however, that it is not actually necessary to fulfill the small-angle condition. The Fresnel kernel well approximates the exact kernel in regions where the phase oscillates slowly, and in regions of fast oscillations, indicated by large An, the error between kernels should be negligible due to stationary-phase arguments. We experimentally verify this conclusion with a helium neon laser setup and show that it should hold at x-ray wavelengths as well.

DOI: 10.1103/PhysRevB.96.054104

I. INTRODUCTION

For many scientific fields, the "ideal" microscope should be capable of producing high-resolution images of nanoscale samples with femtosecond time resolution. Pulsed x-ray sources in particular present a promising method for high-speed high-resolution imaging. Direct imaging techniques require a lens to form the sample image on the detector, but x-ray lenses are often chromatic, inefficient, and difficult to manufacture. Coherent diffraction imaging (CDI) does not use an image-forming optic, thus avoiding the drawbacks of x-ray lenses. The sample is illuminated with coherent light, and the diffraction pattern is measured by a large-area pixelated detector. Since the detector can only measure the amplitude of the diffraction pattern, iterative phase retrieval (IPR) algorithms are used to reconstruct the sample amplitude and phase.

The standard CDI experiment is conducted in the so-called "far-field" (*Fraunhofer*) regime where the distance between the sample and the detector is much greater than the sample size. However, there are some circumstances in which far-field techniques cannot be used, such as at very high photon energies or with larger objects because the detector stand-off distance would be impracticably long. As novel and planned light sources, such as diffraction limited storage rings [1] and x-ray free-electron lasers [2] (FELs) dramatically increase the available coherent hard x-ray photon flux, including the proposed Matter and Radiation in Extremes Facility at Los Alamos National Laboratory that will possess a very hard x-ray($\sim$50-keV) FEL [3], these issues will need to increasingly be taken into consideration.

The distinguishing feature of near-field CDI is the significant phase curvature of the diffraction pattern when recorded by the detector. If the sample is illuminated by a plane wave (as was performed for our experimental work), the placement of the detector is key; the closer the detector is to the sample, the more phase curvature the diffraction pattern will possess. Another method for inducing such curvature is to focus the beam on the sample, creating phase curvature in the incident beam itself [4–6]. Studies using such a structured illumination have shown that the reconstructions are more robust to spatial incoherence in the beam [7,8], can allow imaging of extended objects [9,10], and reduce how exactly the support must be known [11]. These factors contribute to faster more reliable convergence of the reconstruction algorithm [12]. Most of these studies have used highly focused beams with spherical wave fronts, but it has been shown that astigmatic beams could allow for even better reconstructions [13]. The biggest downsides to focused-beam Fresnel imaging is that the beam must be characterized and positional errors can play a large role [14], which are not concerns for plane-wave CDI in the Fresnel regime.

There has been some effort to evaluate the validity of the Fresnel approximation itself. Many analyses mention a value Barber called the "small-angle number" An, that arises in the derivation of Fresnel diffraction in the case of plane-wave (not focused) wave fronts [3]. This value, derived in detail in Sec. I A, is the result of limiting the contributions of high-order terms in the Fresnel approximation; the "required" condition is that $An \ll 1$. This condition can be difficult to satisfy in practice, especially at low photon energies, but several authors have suggested that fulfilling the small-angle condition may not be necessary. Goodman stated that when $An > 1$ the diffraction integral could be understood via stationary-phase arguments, meaning that any errors incurred by violating the condition are averaged out by rapid oscillations in the kernel [15]. Veerman *et al.* agreed with Goodman's assessment and proposed that the small-angle condition could be relaxed somewhat in the particular case of paraxial beams [16]. Southwell provided the justification for needing paraxial beams by showing that it was required by the stationary-phase treatment [17]. Finally, Forbes examined the error of the Fresnel approximation by observing the effect of including or discarding higher-order terms in various approximations [18]. However, Forbes restricted his analysis to two different forms of the Fresnel approximation

*sandberg@lanl.gov

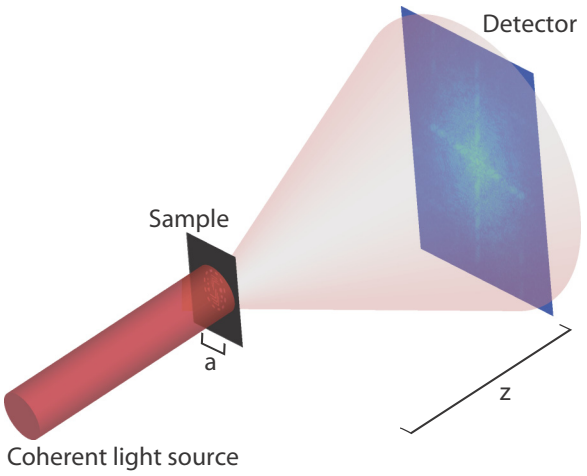

FIG. 1. Basic CDI setup; a coherent light source illuminates the sample of diameter a, and the resulting diffraction pattern is measured at a detector located a distance z from the sample.

and did not consider the error as compared to the exact solution. Here we show theoretically that the small-angle condition may not be the limiting factor in near-field CDI experiments and verify this conclusion experimentally by successfully reconstructing near-field CDI data with high An obtained with a visible light experimental setup. We predict that the small-angle number condition can be violated to yield high-resolution reconstructions at x-ray wavelengths as well.

A. Theoretical background

In this section, we will analyze the pertinent scattering theory of near-field CDI and the sensitivity to the large-angle number An. We begin by considering a sample exit wave propagating in the z direction through free space to a detector some distance z away as in Fig. 1. As can be derived directly from Maxwell's equations, the electric-field $\vec{E}(x,y,z,t)$ in a wave front traveling in a vacuum obeys the vector wave equation,

$$\frac{\partial^2}{\partial t^2}\vec{E} = c^2\nabla^2\vec{E}. \tag{1}$$

We make two simplifications: (1) We make the approximation that the polarization $\hat{\varepsilon}$ of the electric field is constant over the space of propagation so that $\vec{E}(x,y,z,t) = \hat{\varepsilon}E(x,y,z,t)$. (2) We assume that the wave has a "carrier frequency" ω so that its time dependence can be represented as $E(x,y,z,t) = E(x,y,z)e^{\pm i\omega t}$. Then the spatial variation of the electric-field magnitude $E(x,y,z)$ obeys the *Helmholtz equation*,

$$\frac{\partial^2}{\partial z^2}E = -k^2E - \nabla_\perp^2 E, \tag{2}$$

where $k = \omega/c = 2\pi/\lambda$ is the vacuum wave number of the source and $\nabla_\perp^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the part of the Laplacian transverse to the direction of the propagation. It is not an exaggeration to say that the Helmholtz equation *defines* diffraction for a propagating wave front.

For a complex wave front starting at $z = 0$, called the "object plane" for CDI and propagating a distance z to the "detector plane," the relevant solution to Eq. (2) is given by

$$E(\vec{r},z) = \int d\vec{r}\,'\,E(\vec{r}\,',0)H(\vec{r}-\vec{r}\,',z), \tag{3}$$

where the kernel H is

$$H(\vec{r},z) = \frac{z}{2\pi(|\vec{r}\,|^2+z^2)}\left(\frac{1}{\sqrt{|\vec{r}\,|^2+z^2}} - ik\right) \times \exp\left(ik\sqrt{|\vec{r}\,|^2+z^2}\right). \tag{4}$$

Here $\vec{r} = (x,y)$, $\vec{r}\,' = (x',y')$, and z is the distance between the sample and the detector. The two-dimensional vectors $\vec{r}\,'$ and $\vec{r}$ span planes perpendicular to the propagation direction at longitudinal positions 0 and z, respectively. Together Eqs. (3) and (4) constitute the Rayleigh-Sommerfeld solution to the Helmholtz equation [15].

We now outline several approximations which are often applied to the Rayleigh-Sommerfeld kernel. The magnitude of the quantity $\vec{r}-\vec{r}\,'$ in Eq. (3) is the *transverse* distance between a point on the sample and a point on the detector. This is typically *much* smaller than the sample-to-detector distance z, and so in Eq. (4) we ignore any terms of magnitude $r = |\vec{r}\,|$ compared to z (outside the exponent). Furthermore, any term of order $k = 2\pi/\lambda$ dwarfs terms of order $1/z$ as long as $z \gg \lambda$. Under these approximations, the kernel becomes

$$H(\vec{r},z) \approx -\frac{ik}{2\pi z}\exp\left(ik\sqrt{|\vec{r}\,|^2+z^2}\right). \tag{5}$$

The term in the exponents of (4) and (5) requires greater care. Expanding the exponential term differently leads to the far-field and the near-field approximations. The far-field approximation starts with the kernel H as in Eq. (3) with first argument $\vec{r}-\vec{r}\,'$,

$$H(\vec{r}-\vec{r}\,',z) \approx -\frac{ik}{2\pi z}\exp\left(ik\sqrt{|\vec{r}-\vec{r}\,'|^2+z^2}\right). \tag{6}$$

Define $\vec{R} = \vec{r} + z\hat{e}_z$, $R = |\vec{R}|$, $\hat{R} = \vec{R}/R$, $\hat{r}\,' = \vec{r}\,'/|\vec{r}\,'|$, and $\epsilon = |\vec{r}\,'|/R$. $\vec{R}$ is the three-dimensional vector from the origin (usually taken at the center of the exit wave) to a point on the detector. The quantity ϵ has a magnitude roughly equal to the scale length of the exit wave or sample, divided by the characteristic sizes of the detector and sample-to-detector distance z. We assume that this is a small quantity and manipulate the square root in the exponent above as follows:

$$\begin{aligned}
&\sqrt{|\vec{r}-\vec{r}\,'|^2+z^2} \\
&= \sqrt{|\vec{r}+z\hat{e}_z-\vec{r}\,'|^2}, && \text{(7a)}\\
&= \sqrt{|\vec{R}-\vec{r}\,'|^2} \\
&= R\sqrt{|\hat{R}-\epsilon\hat{r}'|^2} \\
&= R\sqrt{1-2\epsilon\hat{R}\cdot\hat{r}\,'+\epsilon^2} \\
&= R\left\{1-\epsilon\hat{R}\cdot\hat{r}\,'+\tfrac{1}{2}\epsilon^2[1-(\hat{R}\cdot\hat{r}\,')^2]+O(\epsilon^3)\right\}, && \text{(7b)}\\
&= R-\hat{R}\cdot\vec{r}\,'+\frac{1}{2}\frac{|\vec{r}\,'|^2}{R}[1-(\hat{R}\cdot\hat{r}\,')^2]+O(\epsilon^3), && \text{(7c)}
\end{aligned}$$

where the substitution in Eq. (7a) is possible because $\vec{r}$ and $\vec{r}\,'$ are both $\perp$ to $\hat{e}_z$, Eq. (7b) follows from Taylor expanding the preceding line in ϵ, resulting in Eq. (7c) after substituting $\epsilon = |\vec{r}\,'|/R$. The third term in Eq. (7c) (and all higher terms) can be ignored if it contributes a phase much smaller than 2π to the exponent, a condition which is obtained if

$$\frac{1}{2}k\frac{|\vec{r}\,'|^2}{R}[1-(\hat{R}\cdot\hat{r}\,')^2] \leqslant \frac{1}{2}\frac{2\pi}{\lambda}\frac{|\vec{r}\,'|^2}{|\vec{r}\,'+z\hat{e}_z|} \leqslant \frac{1}{2}\frac{2\pi}{\lambda}\frac{|\vec{r}\,'|^2}{z} \ll 2\pi.$$

Since several approximations were performed above and because the maximum value of $|\vec{r}\,'|$ in reality depends upon the shape of the sample, numerical factors of order unity usually are omitted, and the above condition is written simply as

$$Fr = \frac{|\vec{r}\,'|^2_{\max}}{\lambda z} \ll 1, \tag{8}$$

where Fr is the *Fresnel number*.

With these approximations, Eq. (3) becomes

$$E(\vec{r},z) = -\frac{ik}{2\pi z}\exp\left(ik\sqrt{|\vec{r}\,|^2+z^2}\right)\int d\vec{r}\,' E(\vec{r}\,',0)e^{-ik\hat{R}\cdot\vec{r}\,'}. \tag{9}$$

This has the form of a Fourier transform with spatial frequency $k\hat{R}$, a fact which can be leveraged to achieve efficient numerical evaluation of the propagated wave front. Note that the steps leading to Eq. (9) are somewhat distinct from those used in Ref. [3] to approximate the kernel for scattered light, although they arrive at a mathematically similar result.

The Fresnel approximation—which is the focus of this paper—proceeds along a different tack. Taking the first argument of the kernel to be $\vec{r}$ as in Eqs. (4) and (6), we Taylor expand the exponential term for small $|\vec{r}\,|/z$ as follows:

$$k\sqrt{|\vec{r}\,|^2+z^2} = kz\left(1+\frac{|\vec{r}\,|^2}{z^2}\right)^{1/2} \approx kz\left(1+\frac{1}{2}\frac{|\vec{r}\,|^2}{z^2}-\frac{1}{8}\frac{|\vec{r}\,|^4}{z^4}+\cdots\right). \tag{10}$$

The last term (and all subsequent terms) above can be ignored if it contributes a phase much smaller than 2π, a condition which implies that

$$An \equiv \frac{r^4_{\max}}{8\lambda z^3} \ll 1, \tag{11}$$

where $r_{\max} = \max|\vec{r}-\vec{r}\,'|$ is the maximum transverse distance between points on the sample and points on the detector in Eq. (3). The small-angle number An is a dimensionless number which determines when the Fresnel approximation is valid [3]. *If $An > 1$, the error incurred by discarding the higher-order terms could be substantial and indicates that the Fresnel approximation may not be valid.* Under these approximations, the Rayleigh-Sommerfeld kernel becomes the *Fresnel kernel*,

$$H_{Fr}(r,z) = -\frac{ik}{2\pi z}\exp\left(ikz+i\frac{kr^2}{2z}\right), \tag{12}$$

which can also be derived as the kernel for the paraxial approximation to the Helmholtz equation [15].

A significant property shared by *both* the Rayleigh-Sommerfeld and the Fresnel kernels follows from Parseval's theorem, which, if we take $\hat{f}(\vec{q}) \equiv \int d\vec{r}\, f(\vec{r})e^{-i\vec{q}\cdot\vec{r}}$ as our convention for the Fourier transform of the exit wave, states that

$$\int d\vec{r}\,|f(\vec{r})|^2 = \frac{1}{(2\pi)^2}\int d\vec{q}\,|\hat{f}(\vec{q})|^2$$

for functions in two dimensions. The Fourier transforms with respect to $\vec{r}$ of these two kernels are

$$\hat{H}(\vec{q},z) = \exp\left(iz\sqrt{k^2-|\vec{q}\,|^2}\right), \tag{13}$$

and

$$\hat{H}_{Fr}(\vec{q},z) = \exp\left(ikz-i\frac{z}{2k}|\vec{q}\,|^2\right), \tag{14}$$

where Eq. (14) can be derived either directly from Eq. (12) or by performing a Taylor expansion in $\vec{q}$ of the exponent in Eq. (13). Note that $|\hat{H}_{Fr}(\vec{q},z)| = 1$ always and $|\hat{H}(\vec{q},z)| = 1$ when $|\vec{q}| \leqslant k$. Since $|\vec{q}| \leqslant k$ for the parameters of any reasonable experiment that we envision, it is essentially always the case that $|\hat{H}(\vec{q},z)| = 1$ as well.

The intensity at a given propagation distance z and lateral position $\vec{r}$ is given (to within irrelevant multiplicative constants) by $I(\vec{r},z) = |E(\vec{r},z)|^2$. The following derivation then holds for *either* kernel:

$$\begin{aligned} P(z) &= \int d\vec{r}\, I(\vec{r},z) \\ &= \int d\vec{r}\,|E(\vec{r},z)|^2 \\ &= \frac{1}{(2\pi)^2}\int d\vec{q}\,|\hat{E}(\vec{q},z)|^2 \\ &= \frac{1}{(2\pi)^2}\int d\vec{q}\,|\hat{E}(\vec{q},0)\hat{H}(\vec{q},z)|^2 \\ &= \frac{1}{(2\pi)^2}\int d\vec{q}\,|\hat{E}(\vec{q},0)|^2 \\ &= \int d\vec{r}\,|E(\vec{r},0)|^2 \\ &= P(0). \end{aligned} \tag{15}$$

Here the integrated intensity $P(z)$ (which has units of power) is the total amount of energy passing through a plane at z. This result shows that for either kernel $P(z) = P(0)$, i.e., that both kernels exactly preserve the total propagated intensity. Only the *distribution* of intensity changes. Although this is to be expected for the exact kernel, it is of note that the approximate Fresnel kernel achieves this as well.

B. Theoretical findings

We wish to quantify the error in the kernel incurred by approximating Eq. (4) as Eq. (12). In order to accomplish this in a scale-invariant sense, it is convenient to define dimensionless variables and dimensionless, scaled versions of

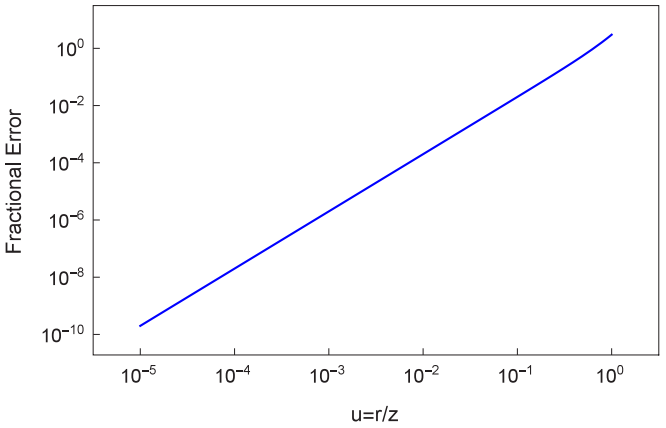

FIG. 2. The fractional error $||h_{Fr}| - |h||/|h|$ in the *magnitude* of the kernel incurred by making the Fresnel approximation. This curve is calculated using $\alpha = 10^9$, although it would look essentially unchanged for any value of α greater than approximately 10^4.

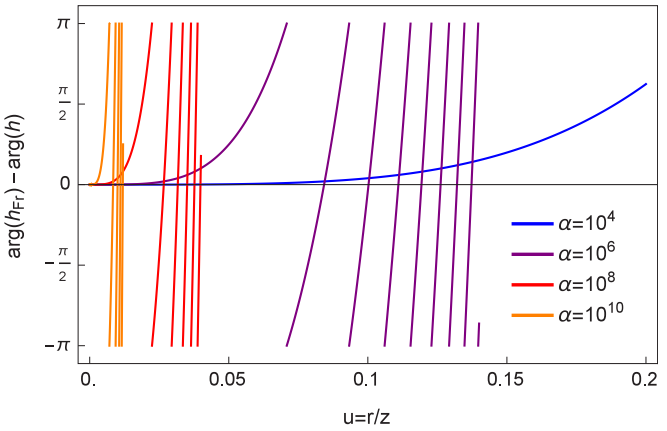

FIG. 3. The difference in the phase (i.e., the complex argument) between $h_{Fr}(u,\alpha)$ and $h(u,\alpha)$ for various values of α. Note that—in order to not clutter the image—the curves for $\alpha = 10^{10}$ and $\alpha = 10^8$ stop at $u = 0.012$ and 0.04, respectively.

these kernels. We define $h(u,\alpha) \equiv z^2 H(zu, \frac{\alpha}{k})$ and $h_{Fr}(u,\alpha) \equiv z^2 H_{Fr}(zu, \frac{\alpha}{k})$ so that

$$h(u,\alpha) = \frac{1}{2\pi(u^2+1)}\left(\frac{1}{\sqrt{u^2+1}} - i\alpha\right)\exp\left(i\alpha\sqrt{u^2+1}\right), \tag{16}$$

$$h_{Fr}(u,\alpha) = -\frac{i\alpha}{2\pi}\exp\left(i\alpha + i\frac{\alpha u^2}{2}\right). \tag{17}$$

Here $u \equiv r/z$ ("dimensionless transverse detector size") compares the size of the detector and sample to that of the sample-to-detector distance and is typically a fairly small number, whereas $\alpha \equiv kz$ ("dimensionless detector distance") is roughly the number of wavelengths that fit into a distance z and is therefore typically a very large number. An can be represented approximately in terms of u and α as

$$An \approx \frac{\alpha u^4}{16\pi}. \tag{18}$$

Here we consider the two different regimes of visible and x-ray wavelengths; the experiment for this paper was conducted using visible light, but the overall results will most likely be useful for hard x-ray experiments. Whereas the dependence on r makes u somewhat sample dependent, we give here some typical values of α and u to give a sense of scale. The data presented in the next section suggest that for 632.8-nm wavelength photons (1.96 eV), $\alpha \sim 8.8 \times 10^4$ and $u \sim 0.83$ when $z = 0.89$ cm and the half-width of the pattern on the detector is $r \sim 7400$ μm [Fig. 6(b)], and $\alpha \sim 4.3 \times 10^4$ and $u \sim 0.87$ when $z = 0.43$ cm and $r \sim 3700$ μm [Fig. 6(c)]. When considering 10-keV photons, $\alpha \sim 5 \times 10^{10}$ when $z =$ 1 m and $\alpha \sim 5 \times 10^8$ when $z = 1$ cm. Values around $u = 0.1$ would be considered fairly large for x-ray experiments at this energy range.

The fractional error in the *magnitude* displayed by the Fresnel kernel (compared to the exact kernel) is shown in Fig. 2 as a function of u for visible light and 10-keV photons. The difference in *phase* (i.e., complex argument) between h and h_{Fr} is shown in Fig. 3 for various values of α and u. It is immediately obvious that the greater the value of α, the lower the value of u when the phase error begins oscillating quickly.

We see that for our visible light experiment the amplitude error can be very large, ranging from 10% to 200%, and the phase error oscillates quickly at high u (a better sense of this can be gleaned from the larger range plotted in Fig. 4). The same analysis at x-ray wavelengths reveals slightly different behavior: While the phases of the two kernels decorrelate at small u from very rapid oscillations, the magnitude error will typically be less than about 1%.

Although the phase and amplitude errors of the Fresnel kernel can be quite large, the oscillatory properties of both kernels themselves should be examined. The phase difference of the kernels for $\alpha = 4.27 \times 10^4$ and $\alpha = 10^9$ are plotted in Fig. 4 along with the phase of the exact kernel. Although the scales of the two regimes are quite different, the behavior is the same—the frequency of the exact kernel's oscillation increases much more rapidly than the frequency of the phase error. This is not the default behavior; when α is less than about 50 (not shown) we observe the opposite behavior. This means, however, from the definition of u that the distance between sample and detector is comparable to the sample size [19]; this situation will rarely, if ever, occur.

We see in Fig. 4 that the phase error changes slowly when u is small (where the scale of small u is set by α), signifying that the Fresnel kernel approximates the Rayleigh-Sommerfield kernel well in that region. To further our analysis, we seek to know the effect of An on the kernel phase and phase error. We know that when $An = 1$ the phase difference should already be observable, and using Eq. (11) we find that for $\alpha = 4.27 \times 10^4$, $u \approx 0.185$, and for $\alpha = 10^9$, $u \approx 0.015$; a significant phase difference is observed at those u values as expected. We also see that the kernels themselves have oscillated many times before reaching these u values. In this region we can qualitatively apply the stationary phase argument already alluded to in Refs. [15–17], which suggests that the rapid oscillations of the kernel average the phase error to zero. Since the phase is multiplied by the amplitude in the integral, the amplitude error also is averaged out to zero. This

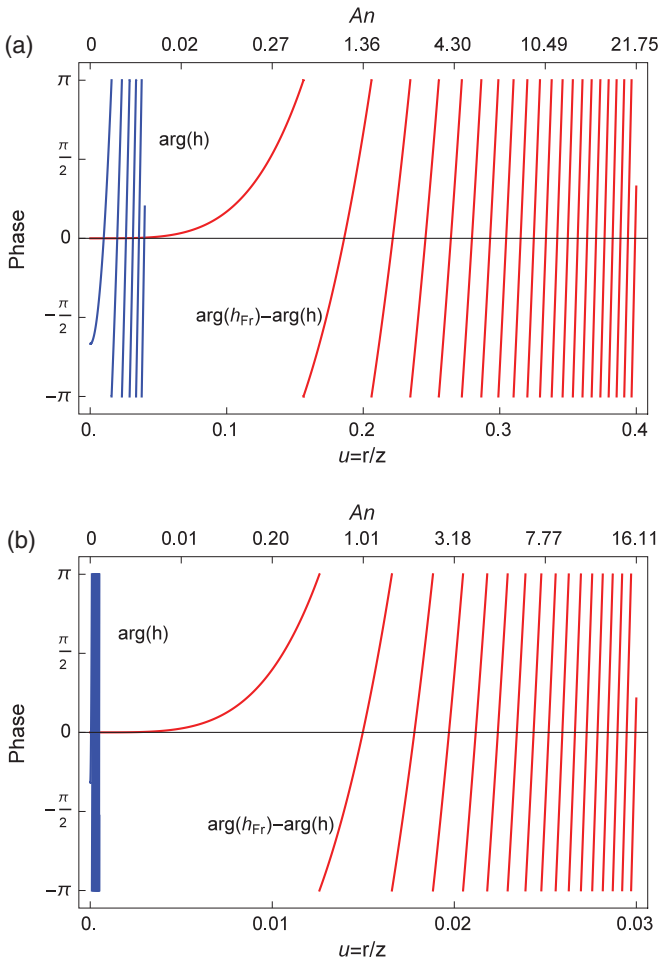

FIG. 4. The difference in the phase (i.e., the complex argument) between $h_{Fr}(u,\alpha)$ and $h(u,\alpha)$ plotted with the exact kernel $h(u,\alpha)$ at (a) $\alpha = 4.27 \times 10^4$ and (b) $\alpha = 10^9$ as a function of u and An as calculated by Eq. (18). The curves for arg(h) were stopped at $u = 0.04$ and $u = 0.0005$, respectively. There are approximately 20 cycles of arg(h) plotted in (b).

result implies that the diffraction patterns given by the Fresnel approximation and Rayleigh-Sommerfield exact solution are nearly identical in all situations that could be envisioned in an actual experiment, and so an IPR that uses the Fresnel approximation should be able to handle any appropriately oversampled diffraction dataset.

After the point where $An \geqslant 1$, the situation becomes murkier. It seems reasonable, after examination of Fig. 4, that the stationary phase argument would hold for somewhat larger u. It is not clear, however, at *what* point the stationary phase argument ceases to hold, or alternatively, if the kernels oscillate quickly enough so that the stationary phase argument *always* holds. We do not present any theoretical findings to answer these questions; in the next section, however, we present data with visible wavelengths (corresponding to $\alpha = 4.27 \times 10^4$) that suggest that even at low-α values and high-An values the stationary phase argument still holds. We can then make the conjecture that if the stationary phase argument always holds in the visible light regime, then the plots in Fig. 4 suggest that it should also always hold in x-ray experiments as well.

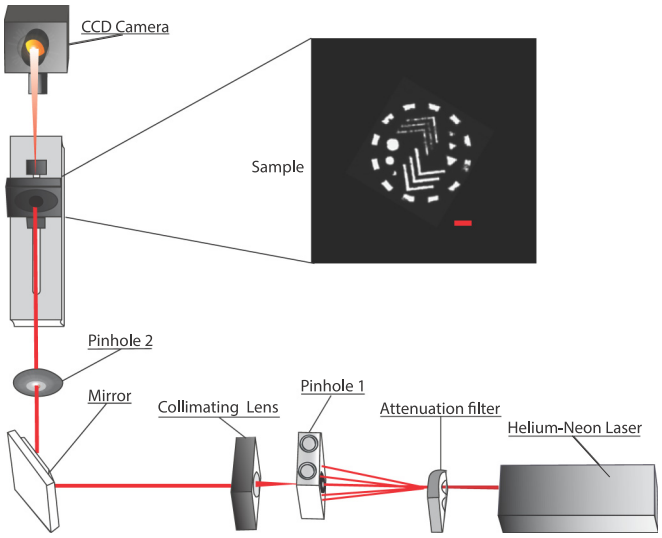

FIG. 5. Diagram of the experimental setup used to provide a coherent beam and record the diffraction pattern from the sample. An optical microscope image of the sample is shown in the inset where the red scale bar is 25-μm wide.

II. EXPERIMENTAL METHODS

In this section, we present the results from a simple near-field CDI experiment performed on a tabletop helium neon-based CDI system.

A. Data acquisition

As shown in Fig. 5, a 5-mW helium neon laser was used to produce a monochromatic 632-nm beam. Two adjustable attenuators reduce the beam intensity, permitting longer exposure times and reducing the risk of saturation or damage to the detector. The attenuated light then passes through a beam-defining pinhole about 150 μm in diameter. Next, the beam is collimated by a lens and redirected toward the sample by a mirror. An iris, tuned to be slightly larger than the central Airy disk, blocks any fringes created by diffraction from the pinhole. A CCD camera (ThorLabs 4070M) then records the diffraction pattern. The distance from the sample to detector z was varied to measure diffraction patterns at different Fr's and An's but was chosen to ensure that the oversampling ratio $Os = z\lambda/PN$ was greater than two. P and N are the detector pixel size and the number of collinear detector pixels with good signal to noise, respectively. The reconstruction of the sample was conducted as discussed in subsequent sections. The total width of the group of geometric apertures is 150 μm.

B. Data processing

In order to maximize the amount of light diffracted to high angles (i.e., near the edge of the detector) while avoiding saturation issues caused by the low dynamic range of our CCD camera, we measured 500 images at three exposure times, ensuring that the lowest-exposure data set did not include oversaturated pixels. The 500 images were summed and background subtracted, then the datasets from the three exposure times were stitched together so that the low-exposure

image provided detailed information of the low-angle scatter, whereas the high-exposure image provided improved signal to noise at higher scattering angles; the process used was similar to that described in Ref. [20]. The stitched image was then centered and thresholded to filter residual noise.

C. Phase retrieval

There are several phase retrieval algorithms that could be applied to reconstruct the images (see Ref. [21] for a good overview of the main variations). They all use constraints in Fourier and object spaces, but they differ in how they use those constraints. We used the hybrid input-output (HIO) and error reduction (ER) algorithms for the results in this paper as this combination has been shown to be one of the best in terms of stability and precision [22]. We define $f(\vec{r})$ as the actual image, $g_k(\vec{r})$ as an estimate of the actual image, and ϕ_k as an estimate of the real phase ψ. $F(\vec{u})$ and $G_k(\vec{u})$ are the Fourier transforms of $f(\vec{r})$ and $g_k(\vec{r})$, respectively, which means that $|F(\vec{u})|$ is the square root of the measured diffraction pattern. The basic steps of such a phase retrieval algorithm are shown below.

(1) The algorithm starts with a seed image $g_k(\vec{r})$, usually assigned random amplitudes and phases.

(2) The Fourier transform of the object $g_k(\vec{r})$ gives $G_k(\vec{u})$, which is an estimate of the measured diffraction pattern $F(\vec{u})$.

(3) The Fourier constraint is applied; the amplitude of $G_k(\vec{u})$ is replaced by the measured diffraction pattern amplitude $|F(\vec{u})|$, but the phase ϕ_k is left unchanged. The result is $G'_k(\vec{u})$.

(4) Inverse Fourier transforming $G'_k(\vec{u})$ gives an improved estimate of the image, called $g'_k(\vec{r})$.

(5) Applying the support constraint results in $g_{k+1}(\vec{r})$. Defining S as the known support of the sample, we see that for ER this constraint takes the form

$$g_{k+1}(\vec{r}) = \begin{cases} g'(\vec{r}), & \vec{r} \in S, \\ 0, & \vec{r} \notin S, \end{cases}$$

and for HIO,

$$g_{k+1}(\vec{r}) = \begin{cases} g'(\vec{r}), & \vec{r} \in S, \\ g_k(\vec{r}) - \beta g'(\vec{r}), & \vec{r} \notin S, \end{cases}$$

where β is a constant, normally chosen between 0.5 and 1. We used $\beta = 0.7$ for our results.

(6) Compute the error (this can be performed in Fourier or object space). If the error is below a certain threshold, quit. If not, go back to step 2.

As originally formulated, the HIO and ER algorithms only work in the far field, but they can be modified to work in the Fresnel regime by employing the "distorted object" approach [23], which modification we employed in our reconstruction algorithm. We used the shrink-wrap algorithm to dynamically generate the support region. We did not enforce any other constraints. Each reconstruction was allowed to run for 125 iterations where one iteration consisted of 10 iterations of ER, 20 iterations of HIO, and one application of shrink wrap. Other optimizations used in our phase retrieval code that aid convergence are found in Refs. [24,25].

III. RESULTS AND DISCUSSION

We took data at three sample-to-detector distances ($z =$ 30, 8.9, and 4.3 mm) that correspond to three different values of An ($An = 175, 11\,400, 41\,600$). We hypothesize that if there are significant errors in the Fresnel kernel because of large An that we will be unable to reconstruct the object in a satisfactory manner. For our data we also calculated the respective values of Fr ($Fr = 0.3, 1.0, 2.1$) where higher values of Fr denote that the data were taken in a nearer field than for lower values of Fr. The dataset with the lowest Fr could be considered to be in the Fraunhofer regime; reconstructions attempted with and without the distorted phase were equally successful with $Fr =$ 0.3 data. The data taken at $Fr = 2.1$ have an oversampling of about 2.5, so it was not possible to explore appreciably higher Fr's and An's, but despite this limitation the values of An span three orders of magnitude. The logarithmic-scale diffraction patterns and their reconstructions are found in Fig. 6, and a summary table listing various important values is found in Table I.

The first observation is that good quality reconstructions (reconstructing even many of the defects and dust particles consistently) are attainable at large An; indeed, $An \gg 1$ for all attempted reconstructions. We see in Fig. 7 that the 10%–90% rule gives a resolution between 2.7 and 2.9 μm for the $Fr = 0.3$ and $Fr = 2.1$ reconstructions, respectively, which is somewhat larger than the smallest sample features (which have dimensions around 2 μm). This explains why the smallest features, i.e., the smallest chevrons, triangles, or circles, are not cleanly resolved in any of the reconstructions, although this can also be attributed in part to dust obscuring some of the features in the test pattern. The pixel sizes in the datasets are 1.29, 0.37, and 0.18 μm for $Fr = 0.3$, 1.0, and 2.1, respectively, so we see that the good signal out to the edge of the detector in the $Fr = 0.3$ data resulted in near diffraction-limited full-pitch resolution, whereas the poorer signal-to-noise ratio in the other two datasets resulted in relatively poorer achieved resolution. A detector with greater sensitivity and lower noise would allow reconstructions at higher Fr to have resolutions closer to the laser wavelength.

We conducted a numerical study of the differences between diffraction patterns propagated via both kernels with x rays with a resolution of twice the wavelength to see if the difference between kernels is significant in high-resolution x-ray experiments. We used a 300-nm diameter object, 1-nm light, and 2-nm object pixel size and computed the diffraction patterns via convolution [26]. Depending on the desired experimental parameters the convolution method can require significant resources—we used a 96-core computing node with about 1.5 Tbytes of RAM for our simulations, operating on $120\,000 \times 120\,000$ element arrays. Such large arrays were required to reduce diffraction "leaking" around the edges of the computational array, although it is difficult to eliminate leaking completely. We simulated the diffraction of both kernels at two propagation distances of 1 μm and 0.2 mm. Memory and computing constraints dictated the longer distance. The propagation of 0.2 mm and 1 μm resulted in Fr's of 0.11, 22.5 and An's of 140 000,40 300, respectively.

The validity of the Fresnel kernel depends only on An, not Fr, and since $An \gg 1$ we might expect to see a difference

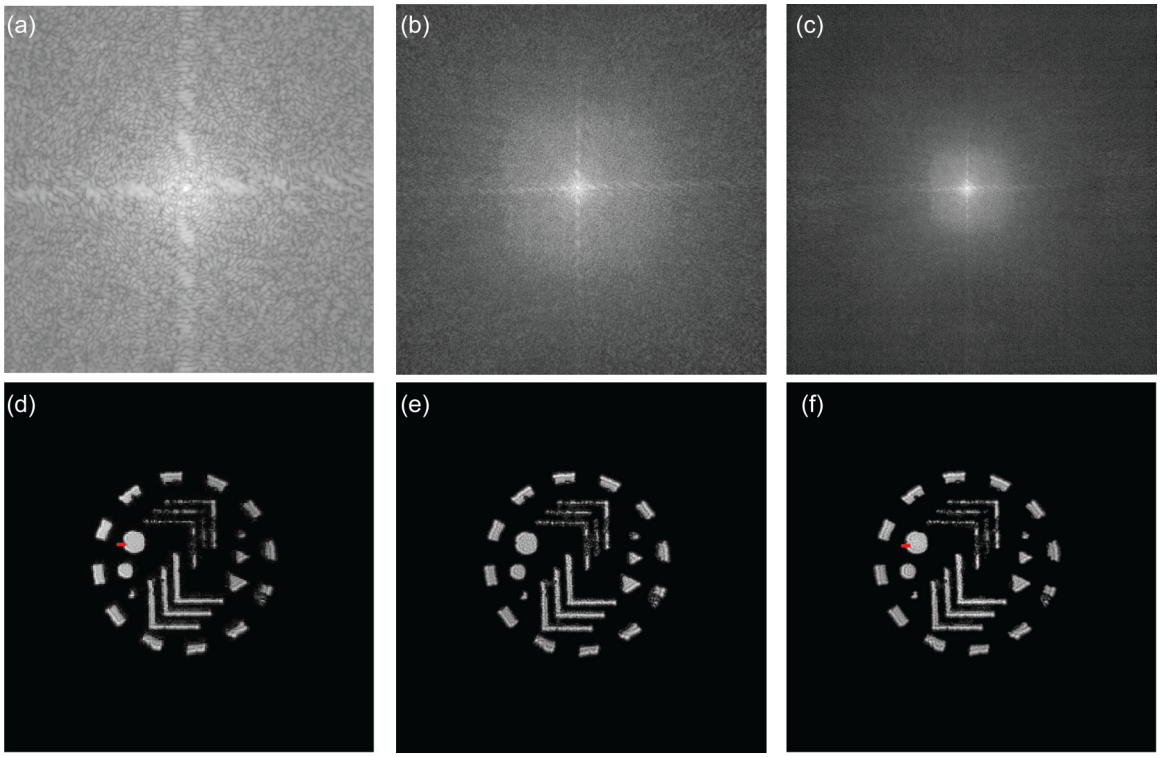

FIG. 6. Experimentally obtained diffraction patterns and reconstructions at different An's. (a)–(c) are uncropped logarithmic-scale diffraction patterns for Fresnel numbers of $Fr = 0.3$, 1.0, and 2.1, and (d)–(f) are the corresponding reconstructions. The data from $Fr = 0.3$ were binned by a factor of 4 to speed up the reconstruction algorithm; the other datasets were reconstructed without binning. The red dashes in (d) and (f) indicate the contours of the line-outs used in Fig. 7.

using the Fresnel and exact kernels. The 0.2-mm propagation yields an error of about 5% between kernels, computed by summing the absolute values of the diffraction pattern intensities and dividing by the sum of the Rayleigh-Sommerfield propagated diffraction pattern intensity. Some of this error can be attributed to diffraction leaking around the detector edges. The greater part of the error is simply due to differences between the Rayleigh-Sommerfield and the Fresnel kernels. When considering all sources for error in an actual x-ray experiment, however, such as detector imperfections, the use of beam stops and Poisson noise, 5% error seems tolerable. We note that at a propagation distance of 1 μm the kernel error is only slightly higher at 8%. These results support the conclusion that the value of An is not important when the Fresnel kernel is used without regard to photon energy or resolution. Situations involving larger samples and higher energies are certainly useful to consider, but those simulations are even more computationally expensive than the results we have presented here; additionally, there is no reason to believe that the results would be different.

TABLE I. Important values in the experiment at each Fr.

z (mm)	Os	Fr	An	u	α
30	17.1	0.3	175	0.25	2.98×10^5
8.9	5.1	1	11400	0.83	8.84×10^4
4.3	2.5	2.1	41600	0.87	4.27×10^4

It was mentioned in a previous section that it is difficult to determine at what point the kernel's oscillations overcome the error between the Fresnel and the Rayleigh-Sommerfield kernels. Although we do not present a good method for determining that bound here, we can say something about it empirically from these results. In Fig. 2 we see that the magnitude error for this setup should be very large (around 100%) for the $Fr = 2.1$ data. One would expect that a reconstruction of that data using the Fresnel kernel in the distorted object approach would not be possible, but we have demonstrated otherwise. Thus it is reasonable to assume that the error in the Fresnel kernel is not as great as Fig. 2 would

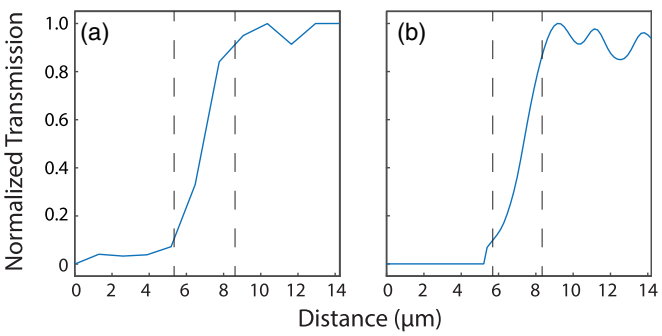

FIG. 7. Line-outs indicated in Fig. 6 for the $Fr = 0.3$,2.1 data. The dashed lines represent the values at 10% and 90% of the maximum value, resulting in a resolution of approximately 2.9 μm when $Fr = 0.3$ and 2.7 μm when $Fr = 2.1$.

suggest, and thus it seems that the kernel oscillations start early and fast enough to effectively minimize error accumulation. Encouragingly, we see in Fig. 4 that the kernel difference manifests much earlier and more rapidly in comparison to the exact kernel oscillations at x-ray wavelengths than for the helium neon setup. Therefore, since acceptable reconstructions were obtained for visible wavelengths, we would expect the phase error to play even less of a role for x-ray experiments.

We have shown theoretically that the condition $An \ll 1$, once thought to be a limiting factor in near-field reconstructions, actually only describes one small effect of an approximation step of the diffraction integral that usually is overshadowed by the rapid oscillations in the diffraction integral that average out to zero. As An becomes large the difference between the Fresnel and the Rayleigh-Sommerfield kernels becomes negligible compared to the oscillations of either kernel, and both the amplitude and the phase error get averaged to zero while evaluating the integral. We have demonstrated experimentally that there seems to be no appreciable restriction due to An at visible wavelengths and have argued based on our results that the restrictions would be even less of a concern at x-ray wavelengths; this conclusion, however, still needs to be verified by experiment.

ACKNOWLEDGMENTS

This work was performed under Contract No. DE-AC52-06NA25396 with the U.S. Department of Energy. We acknowledge Los Alamos National Laboratory LDRD funding under Program No. 20120278ER.

[1] M. Eriksson, J. F. van der Veen, and C. Quitmann, J. Synchrotron Radiat. **21**, 837 (2014).
[2] E. Weckert, IUCrJ **2**, 230 (2015).
[3] J. L. Barber, C. W. Barnes, R. L. Sandberg, and R. L. Sheffield, Phys. Rev. B **89**, 184105 (2014).
[4] 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).
[5] G. J. Williams, H. M. Quiney, A. G. Peele, and K. A. Nugent, New J. Phys. **12**, 035020 (2010).
[6] H. Quiney, J. Mod. Opt. **57**, 1109 (2010).
[7] G. J. Williams, H. M. Quiney, A. G. Peele, and K. A. Nugent, Phys. Rev. B **75**, 104102 (2007).
[8] L. W. Whitehead, G. J. Williams, H. M. Quiney, K. A. Nugent, A. G. Peele, D. Paterson, M. D. de Jonge, and I. McNulty, Phys. Rev. B **77**, 104112 (2008).
[9] B. Abbey, K. A. Nugent, G. J. Williams, J. N. Clark, A. G. Peele, M. A. Pfeifer, M. de Jonge, and I. McNulty, Nat. Phys. **4**, 394 (2008).
[10] B. Zhang, M. D. Seaberg, D. E. Adams, D. F. Gardner, E. R. Shanblatt, J. M. Shaw, W. Chao, E. M. Gullikson, F. Salmassi, H. C. Kapteyn, and M. M. Murnane, Opt. Express **21**, 21970 (2013).
[11] K. A. Nugent, A. G. Peele, H. N. Chapman, and A. P. Mancuso, Phys. Rev. Lett. **91**, 203902 (2003).
[12] H. M. Quiney, K. A. Nugent, and A. G. Peele, Opt. Lett. **30**, 1638 (2005).
[13] K. A. Nugent, A. G. Peele, H. M. Quiney, and H. N. Chapman, Acta Crystallogr., Sect. A: **A61**, 373 (2005).
[14] H. M. Quiney, A. G. Peele, Z. Cai, D. Paterson, and K. A. Nugent, Nat. Phys. **2**, 101 (2006).
[15] J. W. Goodman, in *Introduction to Fourier Optics*, 2nd ed., edited by S. W. Director *et al.*, McGraw-Hill Series in Electrical and Computer Engineering (McGraw-Hill, New York, 2005).
[16] J. A. C. Veerman, J. J. Rusch, and H. P. Urbach, J. Opt. Soc. Am. A **22**, 636 (2005).
[17] W. H. Southwell, J. Opt. Soc. Am. **71**, 7 (1981).
[18] G. W. Forbes, J. Opt. Soc. Am. A **13**, 1816 (1996).
[19] M. R. Teague, J. Opt. Soc. Am. **73**, 1434 (1983).
[20] J. Steinbrener, J. Nelson, X. Huang, S. Marchesini, D. Shapiro, J. J. Turner, and C. Jacobsen, Opt. Express **18**, 18598 (2010).
[21] S. Marchesini, Rev. Sci. Instrum. **78**, 011301 (2007).
[22] J. R. Fienup, Appl. Opt. **21**, 2758 (1982).
[23] X. Xiao and Q. Shen, Phys. Rev. B **72**, 033103 (2005).
[24] C. Guo, S. Liu, and J. T. Sheridan, Appl. Opt. **54**, 4698 (2015).
[25] M. Köhl, A. A. Minkevich, and T. Baumbach, Opt. Express **20**, 17093 (2012).
[26] T. Shimobaba, J. Weng, T. Sakurai, N. Okada, T. Nishitsuji, N. Takada, A. Shiraki, N. Masuda, and T. Ito, Comput. Phys. Commun. **183**, 1124 (2012).