Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 2 December 2005
Accepted 4 May 2006

Reconstruction of a yeast cell from X-ray diffraction data

Pierre Thibault,[a] Veit Elser,[a]* Chris Jacobsen,[b,c] David Shapiro[b] and David Sayre[b]

[a]Department of Physics, Cornell University, Ithaca, NY 14853-2501, USA, [b]Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794-3800, USA, and [c]Center for Functional Nanomaterials, Brookhaven National Laboratory, Upton, NY 11973, USA. Correspondence e-mail: ve10@cornell.edu

Details are provided of the algorithm used for the reconstruction of yeast cell images in the recent demonstration of diffraction microscopy by Shapiro, Thibault, Beetz, Elser, Howells, Jacobsen, Kirz, Lima, Miao, Nieman & Sayre [*Proc. Natl Acad. Sci. USA* (2005), **102**, 15343–15346]. Two refinements of the iterative constraint-based scheme are developed to address the current experimental realities of this imaging technique, which include missing central data and noise. A constrained power operator is defined whose eigenmodes allow the identification of a small number of degrees of freedom in the reconstruction that are negligibly constrained as a result of the missing data. To achieve reproducibility in the algorithm's output, a special intervention is required for these modes. Weak incompatibility of the constraints caused by noise in both direct and Fourier space leads to residual phase fluctuations. This problem is addressed by supplementing the algorithm with an averaging method. The effect of averaging may be interpreted in terms of an effective modulation transfer function, as used in optics, to quantify the resolution. The reconstruction details are prefaced with simulations of wave propagation through a model yeast cell. These show that the yeast cell is a strong-phase-contrast object for the conditions in the experiment.

1. Introduction

The proposal for using oversampled X-ray diffraction patterns as the basis of a new type of microscopy (Sayre, 1980) was advanced by many recent experiments. After its first demonstration with a manufactured two-dimensional specimen by Miao *et al.* (1999), the method was successfully used on other two-dimensional objects (Marchesini *et al.*, 2003), on three-dimensional engineered structures (Miao *et al.*, 2002; Chapman *et al.*, 2006), on stained bacteria (Miao *et al.*, 2003) and on microcrystals (Williams *et al.*, 2003). A recent advance was made by Shapiro *et al.* (2005) by the measurement of a highly detailed diffraction pattern produced by a single unstained freeze-dried yeast cell with soft X-rays (750 eV). An integral part of the proposed 'diffraction microscope' is the reconstruction algorithm that interprets the diffraction pattern; this action is performed in a conventional microscope by physical apparatus (lenses *etc.*). We report here on the development of this algorithmic component of the instrument that resulted from efforts directed at present-day two-dimensional data from a yeast cell.

Reconstruction of the image of an isolated object is made possible if its diffraction pattern is measured on a fine enough grid. In direct space, this oversampling results in a field of view empty but for a small region occupied by the object. If this region, called the 'support', is small enough, the phase problem has an overwhelmingly high probability of having a unique solution. For a more rigorous treatment of these questions, see Miao & Sayre (2000) and Elser & Millane (2006). All recent work in diffraction microscopy makes use of iterative reconstruction algorithms. The most popular to date is Fienup's (1982) hybrid input–output (*HIO*) algorithm, which was initially inspired by Gerchberg & Saxton's (1972) algorithm. Since then, many variations have been proposed (Marchesini *et al.*, 2003; Elser, 2003*a*; Wu *et al.*, 2004; Russel Luke, 2005). The algorithm used in this work is the difference map (Elser, 2003*a*), a generalization of the hybrid input–output, which makes it applicable to a wide variety of problems (Elser, 2003*c*; Elser & Rankenburg, 2006). An overview of this algorithm is given in §4.

Before reconstruction is attempted, some important attributes of the object to be reconstructed can be determined directly from the data or with only a minimum amount of processing. One of the data sets is reproduced in Fig. 1 (Shapiro *et al.*, 2005). The streaks of intensity at variance with two-dimensional Friedel symmetry call for an explanation even before the data are phased and the image reconstructed. Is the asymmetry with respect to 180° rotation due to the curvature of the Ewald sphere or evidence of strong contrast? By simulating the wave propagation through a model cell, we

show in §2 that, although both mechanisms do in principle contribute, by far the strongest source of asymmetry is the large optical thickness of the specimen. From this exercise, we learn that the contrast mechanism is far from satisfying the Born approximation and that the data correspond to the far-field diffraction pattern of a strongly phase shifted exit wave. This information is important when formulating the constraints used by the phasing algorithm (§4).

The detail of the diffraction pattern (Fig. 1, inset) shows a continuous speckle pattern – qualitative evidence that the oversampling of the diffraction pattern may be sufficient for phasing. A more quantitative test is provided by the autocorrelation of the object/exit wave, as obtained by Fourier transforming the intensity distribution in the diffraction pattern. Because of the non-Friedel asymmetry, this will be complex-valued; the real part is shown in Fig. 2. We see evidence of a rather sharply defined object support. The outline of the oval shape has perfect symmetry with respect to 180° rotation and corresponds to the autocorrelation of the object support. The diameter is twice that of the actual yeast cell. We have direct evidence of the cell from the faint 'ghosts' surrounding the symmetric central oval. These are probably the result of phase interference between the yeast cell and isolated point-like scatterers located a few cell diameters distant, thus forming something like a collection of Fourier-transform holograms (Stroke & Falconer, 1964; McNulty *et al.*, 1992; Eisebitt *et al.*, 2004). The autocorrelation image thus provides very detailed information about the support constraint used by the phasing algorithm (Fienup, 1978). Estimating the area in pixels of the support, we arrive at an oversampling ratio of about 25 (Miao & Sayre, 2000), or about 5 in each dimensions. A quantity more relevant to the difficulty of a reconstruction is the overdetermination ratio (Elser & Millane, 2006), given by

$$\Omega = \tfrac{1}{2}\frac{A_{\rm auto}}{A_{\rm object}}, \qquad (1)$$

where $A_{\rm object}$ and $A_{\rm auto}$ are the number of pixels in the object's and the autocorrelation's support. This number can be seen as the ratio of the number of independent constraints to the number of degrees of freedom: the solution is more overdetermined when Ω is larger. For the yeast cell problem, the overdetermination ratio is almost exactly 2 because the support is convex and nearly centrosymmetric. This represents the most challenging case in terms of computation since Ω cannot be less than 2 in two dimensions. Non-convex supports [as is seen in Chapman *et al.* (2006) for instance] have a higher overdetermination ratio and are known to be easier to solve (Fienup, 1987).

Data at the center of the diffraction pattern are missing because of the beamstop; this beamstop is required in present experiments to avoid saturation and damage of the X-ray detector by the very strong undiffracted beam and saturation by the stronger diffraction intensity at low spatial frequency. In crystallographic phasing, the region interior to the innermost Bragg spots is free of data and the central beam can be

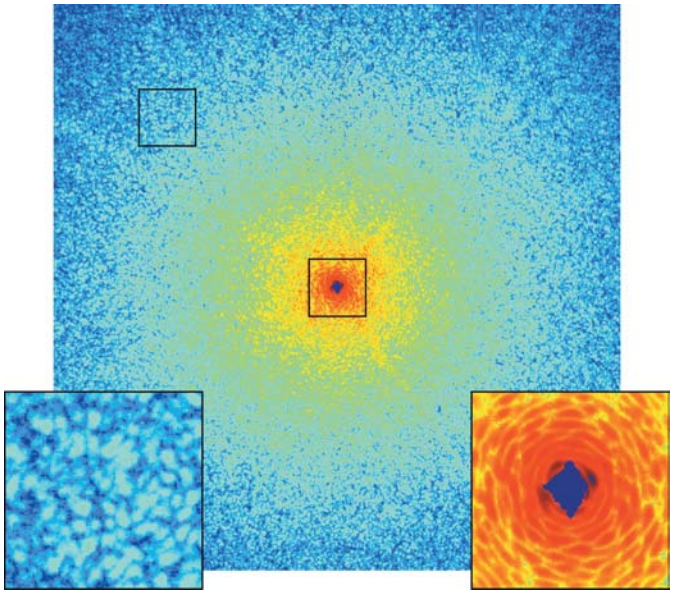

Figure 1
Soft X-ray ($\lambda = 1.65$ nm) diffraction pattern of a freeze-dried yeast cell on a logarithmic scale (Shapiro *et al.*, 2005). This 1200×1200 array extends to $(20.7\ \text{nm})^{-1}$ on the sides, giving the corresponding real-space array 10.3 nm wide pixels. Inset, left: Magnified portion of the diffraction pattern showing the speckles. Inset, right: Magnified central region showing the diamond-shaped missing data region.

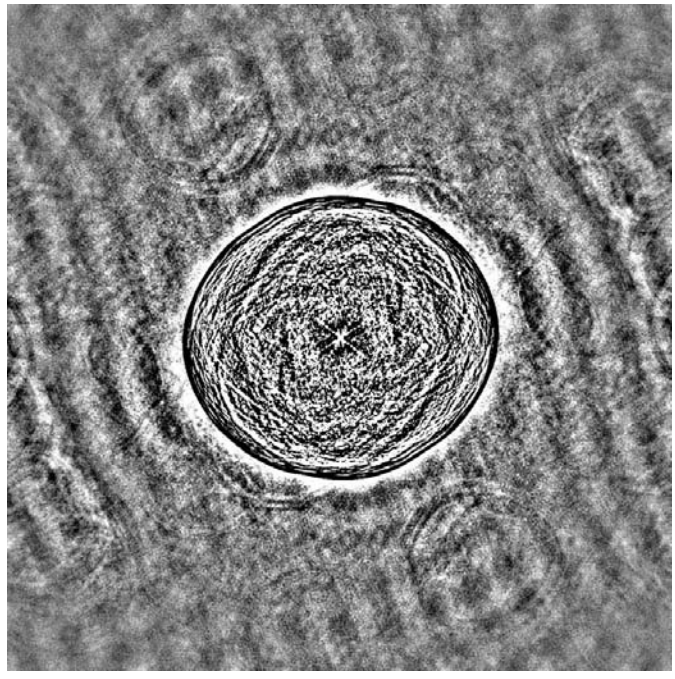

Figure 2
Real part of the autocorrelation. This image is the high-pass filtered inverse Fourier transform of the diffraction pattern shown in Fig. 1 (the high-pass filter reduces the effect of the sharp discontinuity due to missing central data). The central oval-shaped structure is the autocorrelation of the cell. Its contour is used to determine the size of the support and to calculate the overdetermination ratio. The surrounding shapes are faint images of the cell caused by the interference with point-like scatterers around the cell.

allowed to extend somewhat into that region without data loss. In the non-crystal diffraction experiment, the situation is different, and confinement of the beam to a minimal number of detector pixels becomes important. This is now being worked on. We meanwhile address the handling of significant amounts of data loss in §3.

2. Multislice simulations

To arrive at a better understanding of the contrast mechanism in this experiment, we simulated the propagation of X-rays through a model freeze-dried yeast cell. Our model cell is a 3 μm sphere made of lipids and proteins. All the lipids are concentrated in the 50 nm thick cell membrane. The protein material inside the cell is modeled as a binary-valued distribution occupying 25% of the cell volume. The exact spatial distribution of the protein material is random, but the envelope of its power spectrum was forced to follow a power law of the form $|F(q)| \propto q^{-\alpha}$. The exponent of this power law controls the relative contributions of high and low spatial frequencies. Fig. 3(*a*) shows the protein material in one slice of the cell when generated with our choice of exponent, $\alpha = 1.6$. For the simulations, the refractive index of the material inside the cell was sampled on a grid with $(25\,\text{nm})^3$ voxels.

The refractive indices $n = 1 - \delta - i\beta$ of a model lipid (62.3% H, 31.4% C and 6.3% O as number fractions) and protein (48.6% H, 32.9% C, 8.9% N, 8.9% O and 0.6% S) were calculated using the data of Henke *et al.* (1993). At the 750 eV X-ray energy of the experiment, these are

	δ	β
lipid	4.23×10^{-4}	6.88×10^{-5}
protein	5.44×10^{-4}	1.05×10^{-4}

From this we see that the wave acquires a phase shift of $\Delta\varphi = 2\pi t\delta/\lambda$ or about $\pi/2$ as it passes through the center of a 3 μm diameter cell with a quarter of its volume filled with protein so that $t = 0.75$ μm. The magnitude of this phase shift already implies that the Born approximation will not be valid in the interpretation of the diffraction pattern.

We used a propagation formula derived in Appendix *A* to calculate exit waves for our model yeast cell. The formulation is known as the 'beam propagation method' (Feit & Fleck, 1978; Van Roey & Lagasse, 1981) and is very similar to the multislice method (Cowley & Moodie, 1957), the former being traditionally used in optics and the latter in electron microscopy. This method is based on a discretization of the wave propagation along the direction of the incident wave (set along the z axis in this work) from one plane to another. The main result derived in Appendix *A* is the following formula:

$$\widetilde{\Psi}_{\mathbf{q}}(z+\Delta z) = \left\{\widetilde{\Psi}_{\mathbf{q}}(z) + \frac{ik\Delta z}{(1-q^2/k^2)^{1/2}}[\widetilde{\delta n}(z) * \widetilde{\Psi}(z)]_{\mathbf{q}}\right\} \times \exp\{i\Delta z[(k^2-q^2)^{1/2} - k]\}, \quad (2)$$

where $\widetilde{\Psi}_{\mathbf{q}}(z)$ is the two-dimensional Fourier transform of the wavefield in the transverse plane at position z, $k = 2\pi/\lambda$, $*$ is the convolution operation and $\widetilde{\delta n}_{\mathbf{q}}$ is the Fourier transform (in the transverse plane) of $\delta n(\mathbf{r}) = \frac{1}{2}(n^2-1) \approx -\delta(\mathbf{r}) - i\beta(\mathbf{r})$. This formula was preferred to Cowley & Moodie's (1957) standard multislice formula [equation (28) in Appendix *A*] because it does not rely on the small-scattering-angle approximation. Equation (2) becomes exact for arbitrarily large Δz when $\delta n = 0$, resulting in the simple free-space propagation equation:

$$\widetilde{\Psi}_{\mathbf{q}}(z+\Delta z) = \widetilde{\Psi}_{\mathbf{q}}(z)\exp\{i\Delta z[(k^2-q^2)^{1/2} - k]\}. \quad (3)$$

Fig. 3(*b*) shows a typical exit wave obtained with (2) and rendered such that hue and brightness correspond to phase and magnitude, respectively. The distribution of the exit-wave values in the complex plane, shown in Fig. 3(*d*), is itself useful in that it might be used to constrain the reconstruction of the object/exit wave. That these values lie on an approximate spiral is explained by the Eikonal approximation (25) and is obtained from (2) in the limit $q \ll k$. The distribution of exit-wave values Ψ is simply related to the distribution, transverse to the beam, of projected protein thicknesses t by

$$\Psi \propto \exp(ikt\delta n). \quad (4)$$

Broadening of the spiral is due to Ewald sphere lift-off effects, that is, terms in the propagation equations [equation (22) in Appendix *A*] of higher order in the transverse spatial frequency, q/k.

The spirals in Figs. 3(*c*) and 3(*d*) are a clear indication that the exit-wave amplitude is not proportional to the refractive

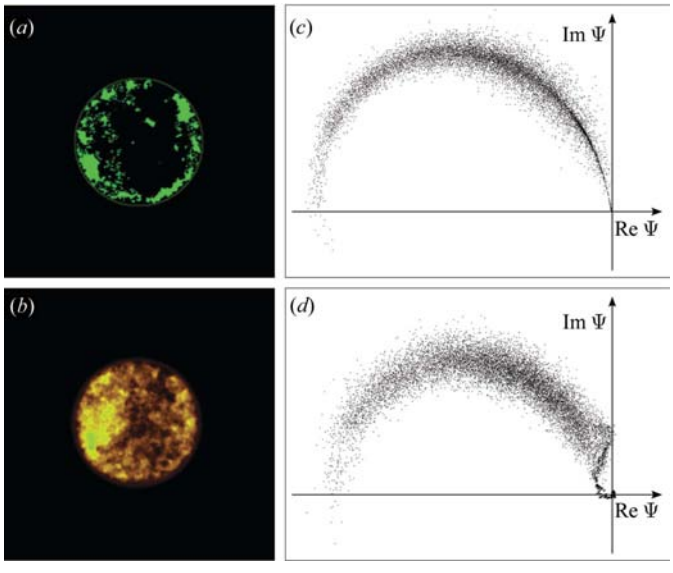

Figure 3
Yeast cell model. (*a*) Slice of the model through the origin. (*b*) Simulated wavefield at the focus plane (*i.e.* exit wave back-propagated to the center of the cell). The image is indistinguishable from the exit wave (not shown). (*c*) Spiral distribution of the values in the complex plane at the focus plane. (*d*) Value distribution at the exit plane. The structure close to the origin is caused by free propagation just outside the support region (see Fig. 4). (*a*) and (*b*) represent complex-valued objects. The colors are defined such that phase is mapped to hue and magnitude is mapped to brightness.

index of the specimen. In this context, the Born approximation is violated, and three-dimensional reconstructions cannot be attempted in the usual way suggested by traditional crystallography, that is through the accumulation into a three-dimensional Fourier data set of two-dimensional measurements on Ewald spheres, as was done by Chapman *et al.* (2006) on an engineered specimen. The two-dimensional reconstruction presented here involves focus effects related to the curvature of the Ewald sphere. By definition, features in the exit waves are more in focus if they are close to the exit plane, while they are slightly dispersed by the wave propagation if they are farther from that plane. For any two-dimensional object, the plane for which the wavefield has the smallest support is the plane in which the object is in focus [see *e.g.* Fig. 9 of Spence *et al.* (2002)]. For a spherical cell, this corresponds to the plane going through its center, corresponding to the exit wave back-propagated [in vacuum, using equation (3)] by half the diameter of the cell. We define this plane as the 'focal plane'.

Fig. 4(*a*) illustrates the simulated perturbation of the incident wave as it propagates through the cell. At the scale of this image, the lateral dispersion of the wavefield is unnoticeable. Fig. 4(*b*) shows the wavefield propagated (in free space) both forward and backward from the exit plane, with the axis in the direction of propagation compressed by a factor 5. As emphasized in Fig. 4(*c*), there is a noticeable spread of the field perturbation. This observation is very important for the understanding of the reconstruction method described in §4.

It is the exit wave back-propagated to the focal plane that the constraints of the reconstruction algorithm are applied to. This two-dimensional object is subject both to the support constraint and, because its forward propagation leads to the measured diffraction pattern, to the Fourier constraint (Fienup, 1982; Elser, 2003*a*). The non-negligible optical depth of the cell, in combination with the curvature of the Ewald sphere, will lead to a blurring of the elements out of the focus plane. How defocus affects the resolution is discussed in more detail in §4 and in Appendix *B*.

The missing data at the center of the diffraction pattern, which account for most of the power in the exit wave, preclude the direct use of value constraints in the reconstruction, and the spiral distribution in particular. With reduced missing data, and use of the unconstrained mode analysis described in §3, it may be possible to exploit value constraints in the future.

3. Unconstrained modes

Any reconstruction effort that uses a support constraint and is faced with a significant region of missing data at the center of the diffraction data must contend with the following form of ambiguity. Consider adding a broad Gaussian feature at the center of the support. If the tails of the Gaussian are small at the support boundary, then the added feature is still consistent with the support constraint. On the other hand, if the Gaussian is sufficiently broad, its counterpart in Fourier space will be narrow enough to fit inside the region of missing data and not upset the constrained diffraction intensities along that region's boundary. Clearly there is an optimal Gaussian mode, with respect to constraints in both direct and Fourier space. The amplitude of this mode is negligibly constrained by the data and support, and as such represents a source of ambiguity in the reconstruction.

There may be other modes, in addition to the Gaussian just described, that contribute to the ambiguity. In Appendix *C*, we develop a formalism for identifying all the 'unconstrained modes' for any particular support/missing-data combination. This analysis provides a set of eigenmodes of a 'constrained power operator', ranked by their corresponding eigenvalues. The constrained power of a mode ranges, by definition, between 0 and 2, where 2 would apply to a mode having most of its power in the measured diffraction pattern and, in direct space, outside the object support. We are concerned here with modes having negligible constrained power. A useful estimate of the number of such modes is derived in Appendix *C* and is given by the formula $M = N_C/\sigma$, where N_C is the number of missing Fourier samples at the center of the diffraction pattern and σ is the oversampling ratio (Miao & Sayre, 2000). For the data shown in Fig. 1, $N_C = 362$; this together with our earlier estimate $\sigma = 25$ gives $M = 14.5$. A heuristic approach consists in describing the unconstrained degrees of freedom as 'missing

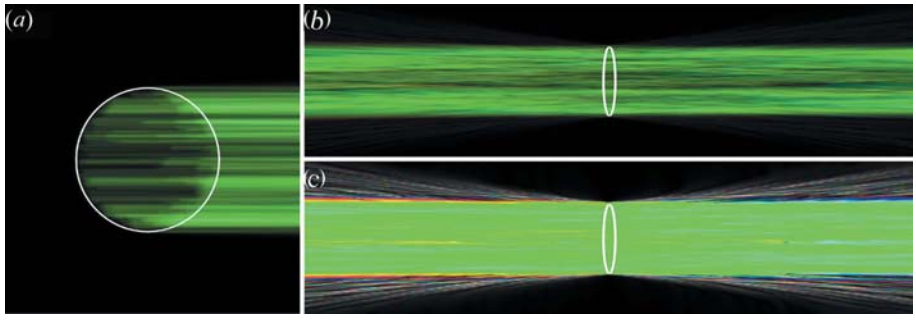

Figure 4
Wavefield propagation simulations. (*a*) Longitudinal section through the center of the cell (outlined in white) of the simulated wavefield perturbation. (*b*) Longitudinal section of the simulated free-space forward and backward propagation of the wavefield. (*c*) Same as (*b*) but with the higher-amplitude values truncated to see the perturbation propagating out of the cell. The coloring scheme is the same as in Fig. 3. In (*b*) and (*c*), the propagation axis has been compressed by a factor of 5.

speckles'; in the present work, M has essentially the same significance as the 'number of speckles' introduced recently by Miao *et al.* (2005).

Fig. 5 shows the results of a detailed unconstrained mode analysis. The number of modes (four) with constrained power less than 0.02 gives some sense of the number of unconstrained degrees of freedom in the reconstruction. Since all of these modes lie in Fourier space at the center of the diffraction pattern, their actual power is large. This makes value constraints, such as positivity, problematic. In the case of complex-valued reconstructions with pixel values having special distributions (for example Fig. 3*c*), the use of value constraints can in principle be beneficial. We chose not to use such constraints for the yeast cell reconstruction.

How does one resolve this M-parameter ambiguity in the reconstruction? Without the benefit of supplementary data or additional *a priori* knowledge, these mode amplitudes are completely unconstrained. Since the reconstruction algorithm we use is based on the available constraints, the amplitudes of the unconstrained modes drift only very slowly with iteration number and are therefore strongly subject to the random start that initiates the algorithm. To eliminate these very slow degrees of freedom and to improve the reproducibility of results, our algorithm orthogonalizes the reconstruction with respect to a set of unconstrained modes that are computed in advance. For the yeast reconstruction, we used four modes, the highest having constrained power 0.0162. In the resulting reconstructions, the artificial zeroing of these mode amplitudes is very noticeable, owing to the fact that their relative power should be quite large. To mitigate this largely aesthetic problem, we use an *ad hoc* prescription for restoring the mode amplitudes that still leads to reproducible results. This is simply to determine the amplitudes by the condition that the variance of the distribution of complex pixel values within the support is minimized. If Ψ_0 is the reconstruction that has been orthogonalized with respect to the M modes, the complex amplitude a_i of the ith mode χ_i is given by the minimization of

$$(\Delta\Psi)^2 = \left\langle \left| \Psi_0 + \sum_{i=1}^{M} a_i \chi_i - \left\langle \Psi_0 + \sum_{i=1}^{M} a_i \chi_i \right\rangle_{\mathrm{S}} \right|^2 \right\rangle_{\mathrm{S}}, \qquad (5)$$

where $\langle\rangle_{\mathrm{S}}$ is the complex-valued pixel average inside the support.

We find that reconstructions appear more 'plausible' when unconstrained modes are restored by this procedure than if their amplitudes are set to zero. Fig. 6 illustrates the effect of the mode replacement method. Fig. 6(*a*) is an example of a reconstruction obtained when the mode amplitudes are allowed to vary freely. Fig. 6(*b*) shows the result of zeroing the mode amplitudes. The final reconstruction, with the modes restored with the *ad hoc* rule, is shown in Fig. 6(*c*).

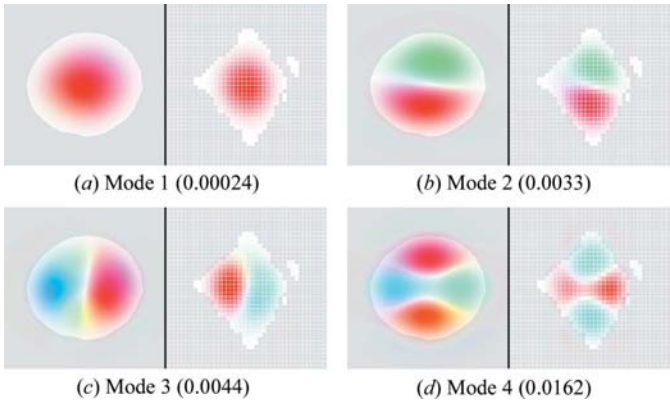

Figure 5
The four most weakly constrained modes (constrained power in parentheses). Each mode is shown in direct and Fourier space. The phase of the mode is illustrated by the hue and the amplitude by the saturation. Both the support (in real space) and region of measured data (in Fourier space) are superimposed in gray.

4. Reconstruction algorithm

There are a number of different algorithms that can be used for speckle diffraction reconstructions (Fienup, 1982; Bauschke *et al.*, 2003). Our reconstruction used the difference-map algorithm (Elser, 2003*a*,*c*) with support and Fourier magnitude as the two constraints. The support projector P_S zeros the reconstruction outside the support,

$$P_S(\Psi_{\mathbf{r}}) = \begin{cases} \Psi_{\mathbf{r}} & \text{if } \mathbf{r} \in \mathbf{S}, \\ 0 & \text{otherwise}, \end{cases} \qquad (6)$$

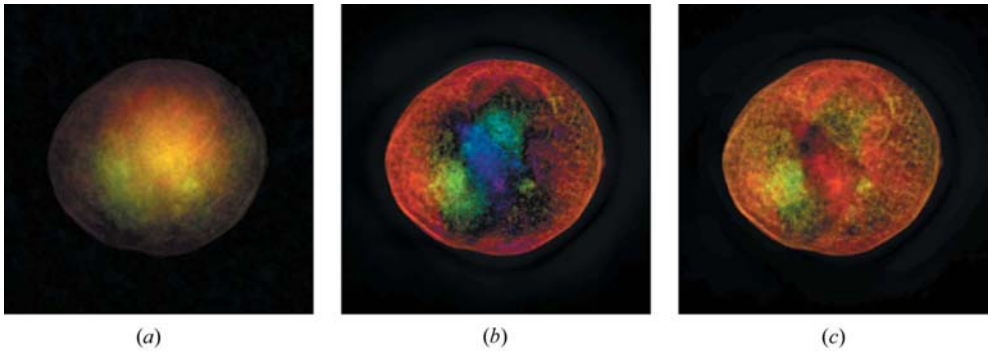

Figure 6
Illustration of the effect of the unconstrained modes. (*a*) Reconstruction with the freely varying mode amplitudes, (*b*) after the four least-constrained modes have been projected out, (*c*) after restoration of the modes using the *ad hoc* variance minimization rule. Both (*b*) and (*c*) are reproducible but (*b*) has a somewhat distracting and unnatural appearance.

while P_F rescales the corresponding Fourier transform to have the measured magnitudes $F_{\mathbf{q}}$ wherever the latter are known (and leaves others unchanged):

$$P_F(\widetilde{\Psi}_{\mathbf{q}}) = \begin{cases} F_{\mathbf{q}} \dfrac{\widetilde{\Psi}_{\mathbf{q}}}{|\widetilde{\Psi}_{\mathbf{q}}|} & \text{if } F_{\mathbf{q}} \text{ is known and } |\widetilde{\Psi}_{\mathbf{q}}| \neq 0, \\ \widetilde{\Psi}_{\mathbf{q}} & \text{otherwise.} \end{cases} \quad (7)$$

As before, $\mathbf{q}$ is the Fourier-space coordinate of the Fourier transform of $\Psi_{\mathbf{r}}$. In terms of directly measured quantities, $|\mathbf{q}|$ is equal, as usual, to $2\pi \sin\theta/\lambda$, where θ is the angle at which a given diffraction intensity is measured.

The difference map is defined by the iteration relation

$$\Psi_{n+1} = \Psi_n + \beta(\Psi_{Fn} - \Psi_{Sn}), \quad (8)$$

where

$$\begin{aligned} \Psi_{Fn} &= P_F\Big((\gamma_S + 1)P_S(\Psi_n) - \gamma_S \Psi_n\Big), \\ \Psi_{Sn} &= P_S\Big((\gamma_F + 1)P_F(\Psi_n) - \gamma_F \Psi_n\Big) \end{aligned} \quad (9)$$

are the nth Fourier and support estimates, respectively. A near-optimal choice for the γ parameters (Elser, 2003*b*) is $\gamma_S = \beta^{-1}$ and $\gamma_F = -\beta^{-1}$. With $\beta = 1$, this algorithm reduces to Fienup's hybrid input–output algorithm. For the yeast cell reconstruction, we used the simple alternative $\beta = -1$, for which the estimates become[1]

$$\Psi_{Fn} = P_F\Big(2P_S(\Psi_n) - \Psi_n\Big) \quad \text{and} \quad \Psi_{Sn} = P_S(\Psi_n). \quad (10)$$

It is clear in equation (8) that the iteration reaches a fixed point if the two estimates are equal, in which case the reconstruction is equal to either of these estimates. The Fourier and support estimates differ in general and the difference-map error, used to monitor convergence, is defined as

$$\epsilon_n = \|\Psi_{n+1} - \Psi_n\| = \|\Psi_{Fn} - \Psi_{Sn}\|. \quad (11)$$

The reconstruction iterate Ψ_n is complex-valued since the Fourier magnitudes in the data lack Friedel symmetry.

Details about the implementation of the algorithm that are special to the yeast reconstruction and deserve elaboration are the treatment of the support constraint and the averaging procedure used to achieve reproducible results in the presence of noisy constraints.

The support was initially defined as a rectangle of dimensions half the size of the oval shape observed in the autocorrelation (Fig. 2). It was then refined with various methods, going from simple thresholding to pixel-by-pixel editing in a drawing program. A more systematic approach has been demonstrated by Marchesini *et al.* (2003) who use a combination of low-pass filtering and thresholding to automatically generate a tightening support at each iteration in a procedure called 'shrink wrap'. We did not use this approach in our present work. Instead, each time a new support was obtained, a new reconstruction was attempted (without most of the refinements described below) for a few hundred iterations only. The new reconstruction attempt was then used to define a yet tighter support. At each step, the cell membrane became more and more sharply defined, which made the definition of the support outline more and more accurate. Excessively tight supports were avoided by noticing sharp upturns in the difference-map error.

A new procedure was found to be necessary for obtaining reproducible results. The noise in the Fourier data creates incompatibility in the constraints with the result that the reconstruction iterate never reaches a fixed point. Instead, the algorithm enters a steady state, characterized by an error that fluctuates around a constant, non-zero, value. The relatively small error in this regime suggests that the region explored by the algorithm contains the best solution, that is, the image that would be reconstructed in the absence of noise in the data. In the absence of convergence, however, the definition of 'solution' becomes ambiguous. We decided to adopt an averaging method to obtain a unique solution, independent of both the initial conditions and the number of iterations. The final reconstruction is defined as the average of many Fourier estimates in the steady-state regime. We take advantage of the chaotic dynamics of the algorithm to form averages out of estimates taken in a single run; we did not notice any benefit from the averaging of estimates from different runs, although it has been observed recently by Chapman *et al.* (2006) that using different runs could help to identify reconstruction pathologies (phase vortices). A tight support is needed to avoid compromising the average by a drift of the reconstruction within the support. Cross correlation between estimates could also be used to improve alignment before computing the average; this procedure was not used in the present case, but will be used in future work. Appendix *B* describes the effect of averaging over many estimates translated with respect to each other.

Besides reproducibility, another benefit of the averaging process is a systematic cancellation of the highly fluctuating (thus not well determined) degrees of freedom. Since the phase fluctuation is larger in the high spatial frequency region of the Fourier data, the global effect of averaging is a decrease in the detail seen in the reconstruction. Apart from achieving reproducibility, we believe that reconstructions not averaged over residual fluctuations convey a misleading degree of detail (resolution). As described in §5, the power reduction at high spatial frequencies, as a result of averaging, provides a way to measure the spatial resolution of the final reconstruction.

Averaging a complex-valued reconstruction introduces additional complications. Both projectors are blind to a constant phase factor in the reconstruction iterate [that is, $P_S(e^{i\theta}\Psi) = e^{i\theta}P_S(\Psi)$ and $P_F(e^{i\theta}\Psi) = e^{i\theta}P_F(\Psi)$]. Prior to averaging, we therefore reset the global phase factor such that the sum of the pixel values in Ψ_F lies on the real axis. This rotation in the complex plane has to be done only after the unconstrained modes are projected out of the image, since these free amplitudes bias the evaluation of the global phase factor.

[1] Note that $2P_S(\Psi) - \Psi$, appearing here with our particular choice of parameters, is called in other work the 'reflector' (Bauschke *et al.*, 2003) or the 'charge-flipping' operation (Wu *et al.*, 2004).

Noise and support size play an even more subtle role in the particular case of the reconstruction of a complex-valued image. In fact, it has often been observed (Fienup, 1987; Spence *et al.*, 2002; Faulkner & Rodenburg, 2004) that very tight supports are needed for the reconstruction of complex-valued objects, especially when noise is present in the data. This can be explained by the defocus ambiguity inherent to any reconstruction lacking value constraints (such as reality or positivity). The measurement of the cell's diffraction pattern gives the magnitude of the in-plane Fourier transform of the exit wave, $|\widetilde{\Psi}_{\mathbf{q}}|$. By virtue of equation (3), $|\widetilde{\Psi}_{\mathbf{q}}(z)| = |\widetilde{\Psi}_{\mathbf{q}}(z + \Delta z)|$ for any Δz. The Fourier constraint is therefore insensitive to the focus plane of the reconstruction. As illustrated in Fig. 4, if Δz is small enough (see Appendix *B* for a more precise definition of Δz), an exit wave propagated backward or forward using equation (3) still almost satisfies the support constraint as well. Hence, the reconstruction will never be uniquely determined since a whole ensemble of defocus planes will nearly satisfy both constraints. If there was no noise in the diffraction pattern and no unknown scatterers outside the support, the reconstruction algorithm should in principle converge to what we defined earlier as the focal plane. Therefore, in any realistic situation and in the absence of value constraints in the support, the reproducibility of the reconstruction is compromised not only by the non-convergence of the algorithm but also by this defocus ambiguity. This problem is not present in the case of real-valued exit waves since the exponential factor in (3) violates Friedel symmetry if $\Delta z \neq 0$. Very similar conclusions about this defocus ambiguity were recently published by Chapman *et al.* (2006).

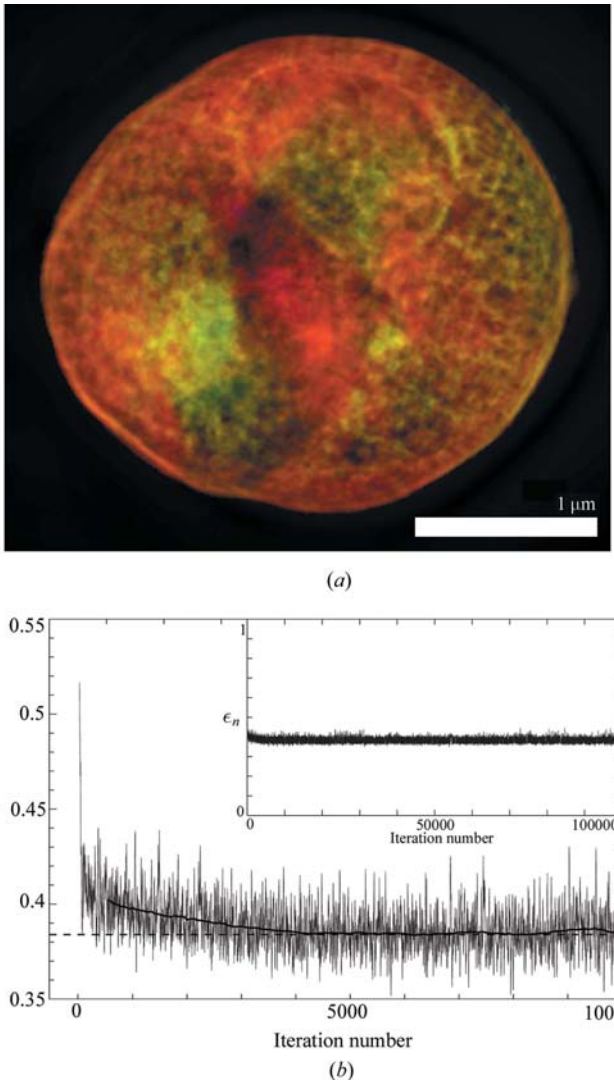

Figure 7
(*a*) Yeast cell reconstruction, based on an average of 1000 Fourier estimates. The weakly constrained modes have been restored using the variance minimization rule. As in previous figures, the phase is represented by the hue and the magnitude by the brightness. (*b*) Difference-map error (ϵ_n) for the first 10000 iterates (thin line). The bold line is a running average (over a 1000 iterate window) to emphasize the decay of the transient. The dotted line is the overall average, $\overline{\epsilon} = 0.384$. The sharp decrease of the error, usual indicator of near-convergence, occurs within the first 50 iterations. Inset: the difference-map error for the whole run. As explained in the text, the overall scale of the error can be considered as arbitrary because of the missing central data.

For the defocus issues as well, the averaging method provides a unique reconstruction, at the cost of a resolution decrease. It is shown in Appendix *B* that the averaging procedure on the defocus ensemble results in an effective low-pass filter on the reconstruction. This source of resolution loss should become dominant for loose supports and/or noisy data.

Future applications of the averaging method could depend less on a tight support if one instead translates the reconstruction to a unique position using the transformation rules in Fourier space. In addition to translations **r** transverse to the beam, we must also consider 'defocus' translations z along the beam. The corresponding phase shifts in Fourier space are given by

$$\phi'_{\mathbf{q}} = \phi_{\mathbf{q}} + \phi_0 + \mathbf{r} \cdot \mathbf{q} + z[(k^2 - q^2)^{1/2} - k], \qquad (12)$$

where we have also included the global phase ambiguity, ϕ_0. Arbitrary values of the parameters ϕ_0, **r** and z result in a new Fourier estimate also consistent with the diffraction data. Recently, Marchesini *et al.* (2005) have pointed out that this phase correction should also include other low-order phase aberrations (including astigmatism, coma, and even the possibility of a phase vortex centered at $\mathbf{q} = \mathbf{0}$). The values of **r**, z and other low-order coefficients are normally discovered by the algorithm in the process of satisfying the support constraint. However, as discussed above, lower-order terms are more sensitive to the tightness of the support. Alternatively, once the reconstruction has entered the averaging stage, one can relax the support constraint and instead fix these coefficients. An improved version of the algorithm implementing such a phase-error correction is currently in development.

5. Results

The reconstructed yeast cell image is shown in Fig. 7(*a*). It is the average of 1000 Fourier estimates, taken every 50 iterations in a single run. We let the averaging begin only after

50000 iterations to avoid transient effects. Fig. 7(*b*) shows the difference-map error of the first 10000 iterations. As suggested by this plot, 5000 iterations would have been sufficient to avoid transients. The inset of Fig. 7(*b*) shows the error evolution for the full 100000 iteration run.

One important point has to be made regarding normalization: prior to the reconstruction, the diffraction data were normalized in the standard way, that is by setting the average intensity per pixel equal to 1. However, the missing central data seriously underestimate the total intensity. The absolute scale of the difference-map error thus does not have a straightforward interpretation. One can compare the mean value of the error in the steady-state regime $\overline{\epsilon}$ with the Fourier magnitudes. The resolution shell where the Fourier magnitudes are of the order of $\overline{\epsilon}$ is independent of the normalization and evaluates to 42 nm in the current reconstruction. We expect this number to be proportional to the resolution of the reconstruction although, as explained below, we judge that it is an overestimate in the present case.

Defining the resolution of the reconstruction is at first not obvious. While resolution is limited by the apparatus in conventional microscopy and by the quality of the specimen in X-ray crystallography, both effects seem to be playing a role in X-ray diffraction microscopy. For a biological specimen, one has additionally a source of measurement noise associated with dose limits. The resolution of a particular reconstruction is therefore a function of the specimen observed. As explained above, the averaging method seems very well suited to reveal the effect of noise on the resolution.

The dimension of the smallest features in Fig. 7(*a*) is about 3 pixels, leading to a rough evaluation of the resolution of 30 nm. Fig. 8 shows a more rigorous way of determining the resolution. For any $\mathbf{q}$ in Fourier space, the ratio of the reconstructed intensity to the observed intensity, $|\widetilde{\Psi}_{\rm rec}(\mathbf{q})|^2/|F_{\rm obs}(\mathbf{q})|^2$, is the result of averaging over residual phase fluctuations in the steady state of the algorithm. The solid curve on the graph is the azimuthal average of this ratio. This curve can be regarded as being analogous to an optical transfer function for the imaging experiment. Of course, if we knew the true phase of the diffraction data and averaged the complex quantity $\Psi_{\rm rec}(q)/F_{\rm obs}$ in different annular resolution shells, we would have something analogous to a coherent transfer function for the coherent imaging system. However, since we have only magnitude data in the Fourier plane, it is more sensible to compare the reconstruction magnitudes (following averaging) with the measured magnitudes, in a manner more closely resembling a modulation transfer function (MTF). In the diffraction reconstruction, this figure indicates the fraction of power in the object against power in the data as a function of spatial frequency, which is quite analogous to the information presented in a MTF curve. For comparison, two classical optics MTF curves (dashed lines) have been added to the graph, showing that the 30 nm resolution estimate is sensible.

We verified that the finest details in the final reconstruction were reproduced when the algorithm was given different starting phases. When the algorithm is given the opposite sign of the difference-map parameter β, the agreement is only

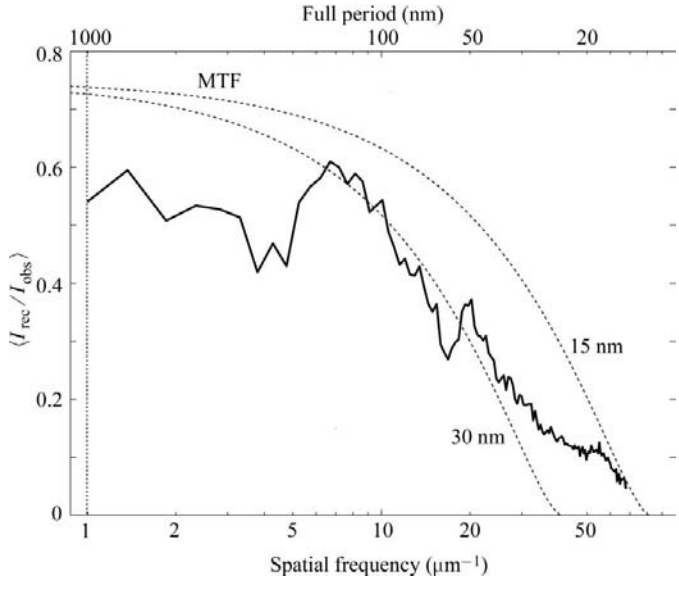

Figure 8
Resolution decrease caused by averaging. The black line shows the relative decrease of the reconstructed intensity as a result of averaging over residual phase fluctuations. The two dashed lines show the classical MTF for an incoherent imaging system with 75% efficiency and a Rayleigh resolution of 15 and 30 nm, respectively. The vertical dotted line indicates the approximate boundary of the beam stop. [From Shapiro *et al.* (2005).]

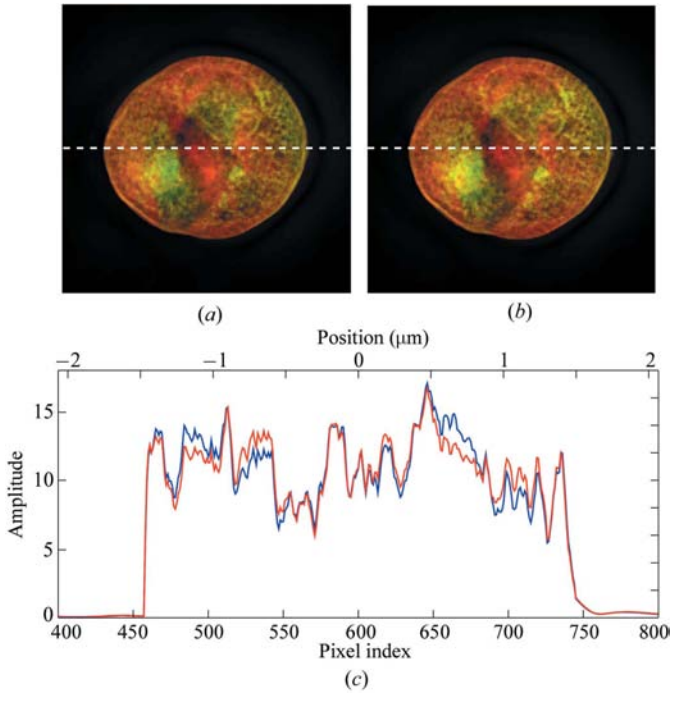

Figure 9
Comparison of the reconstructions obtained with (*a*) $\beta = -1$ and (*b*) $\beta = 1$. The dotted lines show the position of the cross section plotted below. (*c*) Cross section of the absolute value of the two reconstructions for $\beta = -1$ (blue) and $\beta = 1$ (red).

partial, though unnoticeable to the eye, as illustrated in Figs. 9(*a*) and 9(*b*). Fig. 9(*c*) shows a cross section of the absolute value of the reconstruction. This figure shows that the high spatial frequency features are very well reproduced and that the disagreement lies mostly in low spatial frequencies. That the reconstructions are not identical for different sets of parameters is to be expected since the averaging process depends on the ensemble of estimates sampled, which depends on the dynamics of the algorithm. Hence, the high similarity between the two reconstructions in Fig. 9 increases our confidence in our results. Contamination by the reconstruction twin (enantiomorph) could be a source of concern, especially, as in our case, when the support only weakly favors one twin over the other. Certainly, to get optimal results, the averaging of reconstructions should commence only after obvious signs of transients, such as seen in the difference-map error, have passed.

As suggested by the autocorrelation (Fig. 2), weak scatterers were present around the yeast cell during the measurement of the diffraction pattern. Fig. 10 shows the result of our attempt to reconstruct the 'dust' surrounding the cell. This reconstruction was realized by a simple relaxation of the support constraint: the modified support projector sets to zero only those pixels outside the support with amplitude below a predetermined threshold c:

$$P_{S_c}(\Psi_{\mathbf{r}}) = \begin{cases} \Psi_{\mathbf{r}} & \text{if } \mathbf{r} \in S \text{ or } |\Psi_{\mathbf{r}}| > c \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

In the present work, we found that setting this threshold to about six times the amplitude of the error was appropriate. Of course, too low a value for the threshold weakens the constraint and slows down the dynamics of the algorithm. This modified projector can be unstable if most of the power is not already inside the support. The reconstruction shown in Fig. 10 was made by starting with the final iterate of a previous run with the regular support projection. A total of 20000 iterations was needed to generate this average of 400 estimates. In future experiments, a small quantity of strong scatterers could be placed around the main specimen (see for instance Eisebitt *et al.*, 2004). The resulting increase in the overdetermination ratio should help the reconstruction algorithm and also resolve the defocus ambiguity if the specimen is expected to be complex valued. In the present work, these scatterers were too weak to be useful (less than 0.02% of the total power comes from the reconstructed dust).

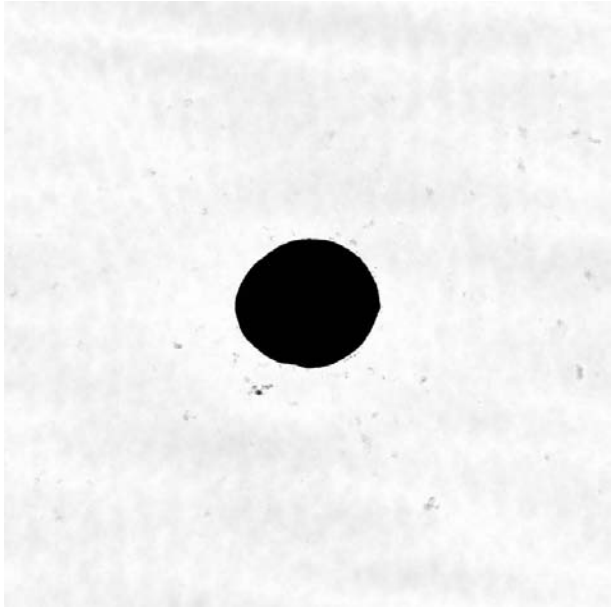

Figure 10
Reconstruction magnitudes of the small scatterers surrounding the cell, saturated to allow weaker points to be seen. This image is the average of 400 Fourier estimates. We have checked that this result is independent of the starting iterate.

6. Conclusions

The reconstruction of a real-space image from the X-ray diffraction pattern of a single unstained cell has stimulated several developments in reconstruction algorithms. Special attention was devoted to the reproducibility of the reconstruction. The missing data in the center of the diffraction pattern give rise to very weakly constrained degrees of freedom (modes) that we identified as the eigenvectors of a constrained power operator. Very weakly constrained modes were projected out of the reconstruction and replaced using a well defined *ad hoc* rule. The lack of convergence was addressed by adopting an averaging method. The final reconstruction therefore does not depend on the starting phases or the number of iterations. The averaging results in an improvement of image reliability along with an effective decrease of the resolution. This can be traced to noise in the data and, in the special case of a complex-valued object without value constraint, the ambiguity in the defocus-plane position.

A number of issues still stand in the way of the application of diffraction microscopy to the three-dimensional imaging of optically thick biological specimens. While missing central data poses a problem even for reconstructions of optically thin specimens (valid Born approximation), the problem is more acute when value constraints (positivity *etc.*) cannot be applied directly to the reconstructed object (exit wave). Present experiments are working at minimizing the region of missing data, and it may soon be possible to obtain two-dimensional data that can be reconstructed without the benefit of value constraints. Merging a series of two-dimensional reconstructions tomographically, to yield a three-dimensional map of the specimen's refractive index, poses another challenge. For thick specimens, this is probably feasible only at shorter wavelengths or when the relationship between the three-dimensional refractive-index contrast and the exit wave is simple (Rytov, 1937). An attractive possibility is to carry out the full set of two-dimensional reconstructions collectively, rather than individually. We are optimistic that some version of these strategies will succeed in the three-dimensional problem, since animated sequences of two-dimensional reconstructions of the yeast cell over a few degrees of tilt (Shapiro *et al.*, 2005) already convey a fair degree of three-dimensional structure.

APPENDIX A
Multislice propagation formula

We derive the propagation formula used in §2 to propagate an incident plane wave through an object with refractive-index distribution $n(\mathbf{r})$. The formalism and final formula are similar in essence to the multislice (Cowley & Moodie, 1957) and the beam propagation methods (Feit & Fleck, 1978; Van Roey & Lagasse, 1981). The radiation field Φ is treated in the scalar wave approximation:

$$\nabla^2\Phi + k^2 n^2 \Phi = 0. \tag{14}$$

Here k is the wavenumber of the incident and elastically scattered waves. We let $\widetilde{\Phi}_{\mathbf{q}}(z)$ denote the Fourier transform of Φ in the dimensions transverse to the incident radiation direction, given by z. Introducing the refractive contrast $2\delta n = n^2 - 1$, equation (14) becomes

$$\partial_z^2 \widetilde{\Phi}_{\mathbf{q}}(z) + (k^2 - q^2)\widetilde{\Phi}_{\mathbf{q}}(z) + 2k^2[\widetilde{\delta n} * \widetilde{\Phi}(z)]_{\mathbf{q}} = 0, \tag{15}$$

where $*$ denotes convolution in the transverse Fourier space. The modes $\mathbf{q}$ are sampled on a periodic grid for efficient computation of the convolution by two-dimensional FFTs. In free space ($\delta n = 0$), the most general solution of (15) has the form

$$\widetilde{\Phi}_{\mathbf{q}}(z) = A_{\mathbf{q}} \exp[iz(k^2 - q^2)^{1/2}] + B_{\mathbf{q}} \exp[-iz(k^2 - q^2)^{1/2}]. \tag{16}$$

An incident wave corresponds to $A_{\mathbf{q}} = \delta(\mathbf{q})$ (the Dirac delta function) and $B_{\mathbf{q}} = 0$.

The scattering is calculated in the modulation approximation

$$\widetilde{\Phi}_{\mathbf{q}}(z) = A_{\mathbf{q}}(z) \exp[iz(k^2 - q^2)^{1/2}], \tag{17}$$

where $A_{\mathbf{q}}(z) = \delta(\mathbf{q})$ for $z < z_0$, and is slowly varying for $z > z_0$. This precludes calculation of the back-scattered waves [second term in equation (16)] or waves with large $\mathbf{q}$. Mathematically, the approximation corresponds to the neglect of the third term in

$$\begin{aligned}\partial_z^2 \widetilde{\Phi}_{\mathbf{q}}(z) = [&-(k^2 - q^2)A_{\mathbf{q}}(z) + 2i(k^2 - q^2)^{1/2}\partial_z A_{\mathbf{q}}(z) \\ &+ \partial_z^2 A_{\mathbf{q}}(z)] \exp[iz(k^2 - q^2)^{1/2}].\end{aligned} \tag{18}$$

Substituting just the first two terms into equation (15), we obtain

$$\begin{aligned}&\partial_z A_{\mathbf{q}}(z) \exp[iz(k^2 - q^2)^{1/2}] \\ &\quad \approx i\frac{k^2}{(k^2 - q^2)^{1/2}} \{\widetilde{\delta n} * A(z) \exp[iz(k^2 - q^2)^{1/2}]\}_{\mathbf{q}},\end{aligned} \tag{19}$$

a set of first-order differential equations, one for each $\mathbf{q}$, coupled by the convolution term. The first-order structure of the equations enables the multislice approach, where the incident wave is propagated unidirectionally through the refractive medium. The exit wave will have z-independent modulation amplitudes $A_{\mathbf{q}}(z)$ and (17) ensures that the scattered waves have the correct wavenumber k.

To cast (19) into a form more amenable to computations, we define

$$\widetilde{\Psi}_{\mathbf{q}}(z) = \widetilde{\Phi}_{\mathbf{q}}(z) \exp(-ikz), \tag{20}$$

and substitute

$$\widetilde{\Psi}_{\mathbf{q}}(z) \exp\{-iz[(k^2 - q^2)^{1/2} - k]\} = A_{\mathbf{q}}(z) \tag{21}$$

into (19). The result is

$$\partial_z \widetilde{\Psi}_{\mathbf{q}}(z) = i[(k^2 - q^2)^{1/2} - k]\widetilde{\Psi}_{\mathbf{q}}(z) + i\frac{k^2}{(k^2 - q^2)^{1/2}}[\widetilde{\delta n} * \widetilde{\Psi}(z)]_{\mathbf{q}}. \tag{22}$$

In the limit $q \ll k$, we recover from (22) the Eikonal approximation since

$$\partial_z \widetilde{\Psi}_{\mathbf{q}}(z) = ik[\widetilde{\delta n} * \widetilde{\Psi}(z)]_{\mathbf{q}} \tag{23}$$

implies

$$\partial_z \Psi(z) = ik\delta n \Psi(z) \tag{24}$$

with solution

$$\Psi(z) = \Psi(z_0) \exp\left(ik \int_{z_0}^{z} \delta n \, \mathrm{d}z\right). \tag{25}$$

Our numerical formula is given by the finite difference integration of (22):

$$\begin{aligned}\widetilde{\Psi}_{\mathbf{q}}(z + \Delta z) = \widetilde{\Psi}_{\mathbf{q}}(z) + ik\,\Delta z \Big\{&[(1 - q^2/k^2)^{1/2} - 1]\widetilde{\Psi}_{\mathbf{q}}(z) \\ &+ \frac{1}{(1 - q^2/k^2)^{1/2}}[\widetilde{\delta n} * \widetilde{\Psi}(z)]_{\mathbf{q}}\Big\}.\end{aligned} \tag{26}$$

This is equivalent, to first order in Δz, to the form that was implemented in our computations, equation (2), restated here for completeness:

$$\begin{aligned}\widetilde{\Psi}_{\mathbf{q}}(z + \Delta z) = &\left\{\widetilde{\Psi}_{\mathbf{q}}(z) + \frac{ik\Delta z}{(1 - q^2/k^2)^{1/2}}[\widetilde{\delta n} * \widetilde{\Psi}(z)]_{\mathbf{q}}\right\} \\ &\times \exp\{i\Delta z[(k^2 - q^2)^{1/2} - k]\}.\end{aligned}$$

When $\delta n = 0$, this equation becomes of the same form as equation (17), making it valid for arbitrary large Δz. The exponential factor to the right is the free-space propagator. The simulations in this work involve a specimen for which the condition $k \gg |\nabla \delta n|$ is always satisfied, so that the Rytov approximation (Rytov, 1937; Davis, 1994) of equation (14) should be valid. We have verified that this is indeed the case.

A1. Implementation details

Practical implementations make use of the fast Fourier transform (denoted as $\mathcal{F}$ here) to compute the convolution appearing in (2), so that this equation can be written as the iteration of

$$\begin{aligned}\Delta\Psi_j &= ik\Delta z \delta n_j \times \mathcal{F}^{-1}\{\widetilde{\Psi}_j\}, \\ \widetilde{\Psi}_{j+1} &= \left[\widetilde{\Psi}_j + \frac{1}{(1 - q^2/k^2)^{1/2}} \times \mathcal{F}\{\Delta\Psi_j\}\right] \\ &\quad \times \exp\{i\Delta z[(k^2 - q^2)^{1/2} - k]\},\end{aligned} \tag{27}$$

with $\delta n_j = \delta n(\mathbf{r}, z_j)$. This particular formulation emphasizes the similarity with the standard multislice (Cowley & Moodie, 1957; Ishizuka & Uyeda, 1977) and beam propagation (Feit & Fleck, 1978; Van Roey & Lagasse, 1981; Thylén, 1983) formulae, which do not conserve the wavenumber of the scattered radiation:

$$\begin{aligned}\Psi'_j &= \exp[ik\Delta z\delta n_j] \times \mathcal{F}^{-1}\{\widetilde{\Psi}_j\},\\ \widetilde{\Psi}_{j+1} &= \mathcal{F}\{\Psi'_j\}\exp\{i\Delta z[(k^2 - q^2)^{1/2} - k]\}.\end{aligned} \tag{28}$$

In this small-angle, or paraxial, approximation, the square root in the second term's denominator of equation (2) is set to 1. Often, the propagator term is approximated by its first non-zero order in q, becoming the Fresnel propagator.

APPENDIX *B*
The effect of averaging on the reconstruction resolution

We consider here the effect of the three translational symmetries of the specimen on the averaging method. The first calculation shows that the average over an ensemble of translated estimates in a plane perpendicular to the propagation direction can be modeled as a Gaussian low-pass filtering operation. The second calculation shows how the resolution can be further affected by an averaging over many defocus planes. This defocus occurs only as a combined effect of the noise and complex-valuedness of the reconstruction.

B1. Averaging over translation in the transverse plane

We assume for simplicity that the distance between the support and the boundary of the object is the same along its contour. Let r_0 be this distance, expressing the tightness of the support. The averaging is made over the family of estimates having Fourier components that differ by a phase tilt $\exp(i\mathbf{r}\cdot\mathbf{q})$, with $|\mathbf{r}| < r_0$. Then, assuming that the distribution of $\mathbf{r}$ is a Gaussian of standard deviation r_0, we find that

$$\begin{aligned}\langle\widetilde{\Psi}_\mathbf{q}\exp(i\mathbf{r}\cdot\mathbf{q})\rangle &= \widetilde{\Psi}_\mathbf{q}\langle\exp(i\mathbf{r}\cdot\mathbf{q})\rangle\\ &= \widetilde{\Psi}_\mathbf{q}\exp(-\tfrac{1}{2}r_0^2q^2).\end{aligned} \tag{29}$$

This results in an effective Gaussian transfer function.

B2. Averaging over defocus planes

We model one slice of the specimen being reconstructed as a uniform disc of radius R. The goal of the first part of this calculation is to find the approximate amplitude of the scattered field at distance z downstream from the disc and a distance r from the axis passing through its center. The Fourier transform of the wavefield at the plane z is given by

$$\widetilde{\Psi}_\mathbf{q}(z) = \widetilde{\Psi}_\mathbf{q}(0)\exp\{iz[(k^2 - q^2)^{1/2} - k]\}, \tag{30}$$

where $\widetilde{\Psi}_\mathbf{q}(0)$ is the Fourier transform of the exit wave at $z = 0$. Since the specimen is assumed to be a uniform disc,

$$\Psi(\mathbf{r}, z=0) = \begin{cases}\Psi_0 & \text{if } |\mathbf{r}| < \mathbf{R}\\ 0 & \text{otherwise.}\end{cases} \tag{31}$$

Then $\widetilde{\Psi}_\mathbf{q}(0)$ is just an Airy disc:

$$\widetilde{\Psi}_\mathbf{q}(0) = \Psi_0R^2\frac{J_1(qR)}{qR}. \tag{32}$$

The inverse Fourier transform of (30) can be written as

$$\begin{aligned}\Psi(r, z) &= \frac{1}{2\pi}\int\widetilde{\Psi}_\mathbf{q}(z)\exp(i\mathbf{r}\cdot\mathbf{q})\,\mathrm{d}^2q\\ &\approx \frac{1}{2\pi}\int\left[\Psi_0R^2\frac{J_1(Rq)}{qR}\exp(-\tfrac{1}{2}izq^2/k)\right]\exp(i\mathbf{r}\cdot\mathbf{q})\,\mathrm{d}^2q\\ &= \Psi_0kR\int_0^\infty J_0(krx)J_1(kRx)\exp(-\tfrac{1}{2}ikzx^2)\,\mathrm{d}x.\end{aligned} \tag{33}$$

Using the method of stationary phase, we find that, when $R \gg (r - R)$ and $z \gg (r - R)$,

$$\Psi(r, z) \sim \Psi_0\frac{z}{k(r-R)^2}, \tag{34}$$

which is also, as it turns out, the asymptotic behavior of the Fresnel integral for a semi-infinite plane (that is, Fresnel diffraction of a plane wave by a straight edge).

As explained in §4, in the presence of noise in the Fourier data, the reconstruction algorithm enters in a steady-state regime. In the absence of convergence, the reconstruction is defined as an average over the Fourier estimates [equation (10)]. Because of noise, the region outside the support in these estimates is never zero but fluctuates with a r.m.s. value ε which scales like the difference-map error. ε can then be seen as the tolerance of the algorithm to fluctuations outside the support. The goal of this Appendix is to point out that this finite tolerance allows in turn a range of defocus values, as long as $\varepsilon > |\Psi/\Psi_0|$, that is:

$$\varepsilon > |\Psi/\Psi_0| \sim \frac{z}{k(R_S - R)^2}, \tag{35}$$

where we have set $r = R_S$, the radius of the circular support used in a reconstruction. It is then reasonable to assume that the extent of the defocus region $|z| < z_0$ tolerated by the algorithm is given by $z_0 \sim \varepsilon k(R_S - R)^2$.

The last step in this calculation consists in computing the effect of averaging on the ensemble of defocus planes. Taking z_0 as the standard deviation of the distribution of planes, we have

$$\begin{aligned}\langle\widetilde{\Psi}_\mathbf{q}\exp(-\tfrac{1}{2}izq^2/k)\rangle &= \widetilde{\Psi}_\mathbf{q}\langle\exp(-\tfrac{1}{2}izq^2/k)\rangle\\ &= \widetilde{\Psi}_\mathbf{q}\exp[-\tfrac{1}{8}(z_0q^2/k)^2]\\ &= \widetilde{\Psi}_\mathbf{q}\exp[-\alpha\varepsilon^2(R_S - R)^4q^4].\end{aligned} \tag{36}$$

The averaging over defocus planes results in a low-pass filter dependent on both the noise (through ε) and the tightness of the support, $|R_S - R|$. It is argued at the end of §5 that an effective alternative consists in translating the defocus planes, thereby avoiding the construction of a highly detailed support.

APPENDIX C
Unconstrained mode analysis

In direct space, there is no constraint on the value of pixels (voxels) $\mathbf{r} \in S$, where S is the support. Similarly, there is no Fourier space constraint at frequencies $\mathbf{q}$ that lie within the region of missing data C, usually at the center of the diffraction pattern. This motivates the following definition for the constrained power in the reconstructed image Ψ:

$$\langle\Psi|W|\Psi\rangle = \int\limits_{\mathbf{r}\notin S} \mathrm{d}\mathbf{r}\,|\Psi(\mathbf{r})|^2 + \int\limits_{\mathbf{q}\notin C} \mathrm{d}\mathbf{q}\,|\widetilde{\Psi}(\mathbf{q})|^2, \tag{37}$$

where

$$\widetilde{\Psi}(\mathbf{q}) = (2\pi)^{-D/2} \int \mathrm{d}\mathbf{r} \exp(-i\mathbf{q}\cdot\mathbf{r})\Psi(\mathbf{r}) \tag{38}$$

is the Fourier transformed image. W is a bounded self-adjoint operator on the space of square-integrable functions. Its spectrum, bounded by 0 and 2, enables us to identify modes that are negligibly constrained in both direct and Fourier spaces. The eigensystem to be solved can be written as an integral equation:

$$\begin{aligned}\lambda_n\Psi_n(\mathbf{r}) &= W\Psi_n(\mathbf{r}) \\ &= X_S(\mathbf{r})\Psi_n(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\, K_C(\mathbf{r}-\mathbf{r}')\Psi_n(\mathbf{r}'),\end{aligned} \tag{39}$$

where $X_S(\mathbf{r})$ is 1 inside S and 0 elsewhere and

$$K_C(\mathbf{r}) = (2\pi)^{-3D/2} \int\limits_{\mathbf{q}\in C} \mathrm{d}\mathbf{q}\, \exp(i\mathbf{q}\cdot\mathbf{r}). \tag{40}$$

This formulation emphasizes the similarity with the important question of how much information can be transmitted in a band-limited optical system (see for instance Gabor, 1975). In this case, one seeks modes that satisfy (Slepian & Pollack, 1961)

$$\lambda_n\Psi_n(\mathbf{r}) = \int\limits_{\mathbf{r}\in S} \mathrm{d}\mathbf{r}'\, K_C(\mathbf{r}-\mathbf{r}')\Psi_n(\mathbf{r}'). \tag{41}$$

As will be shown below, the eigenvalues of W have a behavior very similar to those of the system (41). The current problem is however mathematically different, since a support is known in both direct and Fourier spaces, so that the modes cannot be strictly band-limited. To our knowledge, these modes cannot be expressed in terms of the solutions of (41), the generalized prolate spheroidal wavefunctions.

When the regions S and C are sufficiently large, the relevant features of the spectrum yield to a semi-classical wavepacket analysis. Consider wavepacket modes inside the region S in direct space. The density of these modes in Fourier (momentum) space is given by the well known expression (Rayleigh, 1900; Ashcroft & Mermin, 1976)

$$\mathrm{d}M = \frac{V(S)}{(2\pi)^D}\,\mathrm{d}\mathbf{q}, \tag{42}$$

where $V(S)$ is the area/volume of the region S. To count the modes that also have negligible power in the region C of Fourier space, we integrate (42) over C:

$$M = \int_{\mathbf{q}\in C} \mathrm{d}M = \frac{V(S)\widetilde{V}(C)}{(2\pi)^D}. \tag{43}$$

When $M \gg 1$, the condition for semiclassical analysis to be valid, there will be approximately M negligibly constrained modes with $W \approx 0$. We adopt the simpler terminology 'unconstrained' for these modes from now on.

The structure of the spectrum of W, for modes with higher constrained power, is also simple. There is an infinite near-degeneracy of modes with $W \approx 1$, and relatively few modes with $0 < W < 1$. This too can be understood in semiclassical terms. Consider wavepackets of very small width in the region S. For suitably small widths, the number of independent wavepacket modes within S can be made arbitrarily large. On the other hand, all of these modes will have Fourier transforms with widths so broad that close to the maximum penalty, $W \approx 1$, is incurred by the second term in (37).

The number of unconstrained modes can be estimated from the number of missing Fourier data samples N_C in the region C, and the oversampling ratio, defined as

$$\sigma = \frac{L^D}{V(S)}, \tag{44}$$

where L is the linear size of the direct-space field of view. Since the density of Fourier samples for this field of view is $(L/2\pi)^D$, we have

$$N_C = \widetilde{V}(C)\left(\frac{L}{2\pi}\right)^D. \tag{45}$$

Combining (43), (44) and (45), we obtain

$$M = \frac{N_C}{\sigma}. \tag{46}$$

As in Weyl's formula, for the density of modes of the Laplacian on a bounded region, there are corrections associated with the boundary that diminish the density of modes. We therefore expect (46) to be an overestimate when the regions S and C have large surface-to-volume ratios.

To complete the correspondence between constraints in direct and Fourier space, we can define a missing data ratio

$$\tilde{\sigma} = \frac{\Lambda^D}{\widetilde{V}(C)}, \tag{47}$$

where Λ is the range of spatial frequencies in the diffraction pattern. The number of pixels (voxels) in the support S is given by

$$N_S = V(S)\left(\frac{\Lambda}{2\pi}\right)^D \tag{48}$$

and

$$M = \frac{N_S}{\tilde{\sigma}}. \tag{49}$$

Writing $M = V(C)V(S)(2\pi)^{-D}$ makes it apparent that the number of modes is essentially what is called the 'Shannon number' or the number of degrees of freedom (Gori & Guattari, 1973; Gabor, 1975) of a coherent optical system in the classical optics eigenproblem (41).

C1. Numerical methods for unconstrained mode computations

The two ratios associated with constraints in direct and Fourier space typically satisfy the relationship $1 \ll \sigma \ll \tilde{\sigma}$. From (46) and (49), we then have $M \ll N_C \ll N_S$, and neither the pixels in the support nor the missing samples in the diffraction pattern are well matched in number to serve as efficient bases for the unconstrained modes. Moreover, since the matrix elements of the operator W are not sparse in either of these bases, their computation would be costly.

A better basis, and one that exploits the symmetry between direct and Fourier space, is provided by the quantum harmonic oscillator modes. These modes have the property of having some number of oscillations within a classically limited area, yet they rapidly taper off outside the classical limit and thus can make a nice transition between missing and measured data regions. We illustrate this in one dimension, for support and missing data regions given by

$$S = \{x: |x| < \alpha b\}, \quad C = \{k: |k| < \alpha/b\}. \tag{50}$$

In two dimensions and three dimensions, we would construct a product basis from one-dimensional bases specified by widths along the Cartesian axes. The properly scaled harmonic oscillator modes in direct space are

$$\psi_n(x) = \left(\frac{1}{b}\right)^{1/2} H_n(x/b) \exp[-(x/b)^2/2], \tag{51}$$

where H_n is the nth Hermite polynomial. For computing the constraint penalty in Fourier space, we use the Fourier transform

$$\tilde{\psi}_n(k) = i^n b^{1/2} H_n(kb) \exp[-(kb)^2/2]. \tag{52}$$

Since at the boundaries of S and C we have

$$\psi_n(\alpha b) \sim \tilde{\psi}_n(\alpha/b) \sim H_n(\alpha) \exp(-\alpha^2/2), \tag{53}$$

the constraint penalty is equally divided between direct and Fourier space.

To determine the unconstrained modes, the constrained power operator W is evaluated in the basis above and diagonalized. Since one is only interested in the M least-constrained modes, the basis can be truncated at a small multiple of M. In typical applications, $N_C \ll N_S$, and it makes sense to evaluate the unconstrained modes on the smaller number of Fourier samples within and surrounding the region C. The corresponding modes in direct space may then be computed efficiently using the FFT.

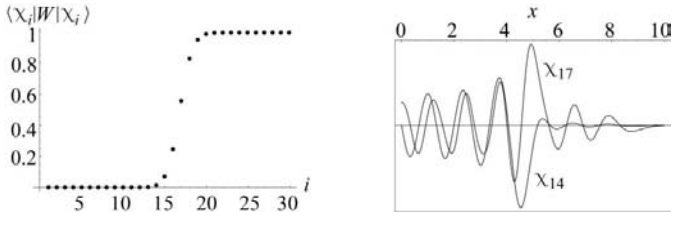

Figure 11
Left: Spectrum of the constrained power operator W in one dimension for support $|x| < 5$ and missing frequencies $|k| < 5$. Right: Comparison of modes χ_{14} and χ_{17} shows the onset of constraint as power invades the region $x > 5$.

C2. Example of a mode calculation

For the one-dimensional regions (50), we use the mode expansion

$$\Psi(x) = \sum_{n=0}^{N-1} c_n \psi_n(x), \tag{54}$$

where $\Psi_n(x)$ is given by (51) and the mode number cut-off N is chosen to be a few times the number of unconstrained modes M. Using (43), we obtain $M = (2/\pi)\alpha^2$ in our one-dimensional example. The truncated mode expansion of the constrained power operator takes the form

$$\langle \Psi | W | \Psi \rangle \approx \sum_{m=0}^{N-1} \sum_{n=0}^{N} w_{mn} c_m^* c_n, \tag{55}$$

where

$$w_{mn} = \begin{cases} \int_{|\xi|>\alpha} H_m(\xi) H_n(\xi) \exp(-\xi^2)\, \mathrm{d}\xi & \text{if } m \equiv n \,(\mathrm{mod})\, 4 \\ 0 & \text{otherwise.} \end{cases} \tag{56}$$

By diagonalizing the matrix w_{mn}, we obtain constrained power eigenvalues w_p and corresponding mode amplitudes c_{np} for the W eigenmodes χ_p $(p = 1, \ldots, N)$:

$$w_p c_{mp} = \sum_{n=0}^{N-1} w_{mn} c_{np}, \tag{57}$$

$$\chi_p(x) = \sum_{n=0}^{N-1} c_{np} \psi_n(x). \tag{58}$$

Fig. 11 shows the results of a computation for $\alpha = 5$, for which $M \approx 15.9$.

C3. Numerical implementation

Application of the mode analysis on experimental data requires that both S and C are well known. These sets are represented as masks in logical arrays (that is, characteristic functions sampled on a grid):

$$X_S(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \in S \\ 0 & \text{otherwise,} \end{cases} \qquad X_C(\mathbf{q}) = \begin{cases} 1 & \text{if } \mathbf{q} \in C \\ 0 & \text{otherwise.} \end{cases}$$

For simplicity, we assume that the arrays are two-dimensional $N \times N$ arrays. The main steps are:

1. *Evaluation of the number of missing modes*. The upper bound to the number of unconstrained modes can be written as

$$M = \frac{N_S N_C}{N^2}, \tag{59}$$

where $N_S = \sum_{\mathbf{r}} X_S(\mathbf{r})$, $N_C = \sum_{\mathbf{q}} X_C(\mathbf{q})$.

2. *Recentering*. If needed, recentering of the sets simply requires a translation of the indices to bring to $\mathbf{0}$ the position of the 'center of mass', given by

$$\bar{\mathbf{r}} = \frac{1}{N_S}\sum_{\mathbf{r}} \mathbf{r} X_S(\mathbf{r}), \tag{60}$$

with the obvious counterpart in Fourier space.

3. *Quantum harmonic oscillator modes*. As explained above, the eigenstates of a quantum harmonic oscillator form a well matched basis for the expansion of the modes of the power operator. The properly normalized two-dimensional version of (51) above is

$$\psi_{n,m} = \frac{1}{(\pi\sigma_x\sigma_y)^{1/2}} H_n(x/\sigma_x) H_m(y/\sigma_y) \exp[-\tfrac{1}{2}t(x^2/\sigma_x^2 + y^2/\sigma_y^2)] \tag{61}$$

and, in Fourier space,

$$\widetilde{\psi}_{n,m} = i^{(n+m)}\left(\frac{\sigma_x\sigma_y}{\pi}\right)^{1/2} H_n(q_x\sigma_x) H_m(q_y\sigma_y) \times \exp[-\tfrac{1}{2}(q_x^2\sigma_x^2 + q_y^2\sigma_y^2)]. \tag{62}$$

σ_x and σ_y can be defined as

$$\sigma_x = \left(\frac{\Delta x}{\Delta q_x}\right)^{1/2}, \quad \sigma_y = \left(\frac{\Delta y}{\Delta q_y}\right)^{1/2}, \tag{63}$$

where

$$(\Delta x)^2 = \frac{1}{N_S}\sum_{\mathbf{r}} x^2 X_S(\mathbf{r})$$

$$(\Delta y)^2 = \frac{1}{N_S}\sum_{\mathbf{r}} y^2 X_S(\mathbf{r})$$

and the analogous definition for $(\Delta q_x)^2$ and $(\Delta q_y)^2$.

4. *Computation of the unconstrained modes*. The expansion of the unconstrained modes in the basis (61) should in general include all $\psi_{k,l-k}$, for all possible k and all l up to a cutoff l_0. The matrix elements of the constrained power operator are given by

$$w_{i,j} = \sum_{\mathbf{r}} X_S \psi_i^* \psi_j + \sum_{\mathbf{q}} X_C \widetilde{\psi}_i^* \widetilde{\psi}_j, \tag{64}$$

where $\psi_i = \psi_{n_i,m_i}$ is a single-index relabeling of the expanding functions. Diagonalization of w yields the sought modes [equation (58) above].

We thank one of the referees for his useful comments and in particular for bringing to our attention the elements of the literature on wave propagation methods and on band-limited coherent imaging systems. This work was supported by Department of Energy grants DoE-FG02-05ER46198 and DoE-FG02-04ER46128 and National Institute of Health grant 1 R01 GM64846-01.

References

Ashcroft, N. W. & Mermin, N. D. (1976). *Solid State Physics*, p. 35. New York: Harcourt College.
Bauschke, H. H., Combettes, P. L. & Russel Luke, D. (2003). *J. Opt. Soc. Am. A*, pp. 1025–1034.
Chapman, H. N., Barty, A., Marchesini, S., Noy, A., Cui, C., Howells, M. R., Rosen, R., He, H., Spence, J. C. H., Weierstall, U., Beetz, T., Jacobsen, C. & Shapiro, D. (2006). *J. Opt. Soc. Am. A*, **23**, 1179–1200.
Cowley, J. M. & Moodie, A. F. (1957). *Acta Cryst.* **10**, 609–619.
Davis, T. J. (1994). *Acta Cryst.* A**50**, 686–690.
Eisebitt, S., Luning, J., Schlotter, W. F., Lorgen, M., Hellwig, O., Eberhardt, W. & Stohr, J. (2004). *Nature (London)*, **432**, 885–888.
Elser, V. (2003*a*). *J. Opt. Soc. Am. A*, **20**, 40–55.
Elser, V. (2003*b*). *J. Phys. A: Math. Gen.* **36**, 2995–3007.
Elser, V. (2003*c*). *Acta Cryst.* A**59**, 201–209.
Elser, V. & Millane, R. D. (2006). In preparation.
Elser, V. & Rankenburg, I. (2006). *Phys. Rev. E*, **73**, 026702.
Faulkner, H. M. L. & Rodenburg, J. M. (2004). *Phys. Rev. Lett.* **93**, 023903.
Feit, M. D. & J. A. Fleck, J. (1978). *Appl. Opt.* **17**, 3990–3998.
Fienup, J. R. (1978). *Opt. Lett.* **3**, 27–29.
Fienup, J. R. (1982). *Am. Opt.* **21**, 2758–2769.
Fienup, J. R. (1987). *J. Opt. Soc. Am. A*, **4**, 118–123.
Gabor, D. (1975). *J. Phys. E Sci. Instrum.* **8**, 73–78.
Gerchberg, R. W. & Saxton, W. O. (1972). *Optik (Stuttgart)*, **35**, 237.
Gori, F. & Guattari, G. (1973). *Opt. Commun.* **7**, 163–165.
Henke, B. L., Gullikson, E. M. & Davis, J. C. (1993). *Atom. Data Nucl. Data Tables*, **54**, 181–342.
Ishizuka, K. & Uyeda, N. (1977). *Acta Cryst.* A**33**, 740–749.
McNulty, I., Kirz, J., Jacobsen, C., Anderson, E. H., Howells, M. R. & Kern, D. P. (1992). *Science*, **256**, 1009–1012.
Marchesini, S., Chapman, H. N., Barty, A., Howells, M. R., Spence, J. C. H., Cui, C., Weierstall, U. & Minor, A. M. (2005). arXiv:physics/0510033. Japan: IPAP.
Marchesini, S., He, H., Chapman, H. N., Hau-Riege, S. P., Noy, A., Howells, M. R., Weierstall, U. & Spence, J. C. H. (2003). *Phys. Rev. B*, **68**, 140101.
Miao, J., Charalambous, P., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
Miao, J., Hodgson, K. O., Ishikawa, T., Larabell, C. A., LeGros, M. L. & Nishino, Y. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 110–112.
Miao, J., Ishikawa, T., Johnson, B., Anderson, E. H., Lai, B. & Hodgson, K. O. (2002). *Phys. Rev. Lett.* **89**, 088303-1–4.
Miao, J., Nishino, Y., Kohmura, Y., Johnson, B., Song, C., Risbud, S. H. & Ishikawa, T. (2005). *Phys. Rev. Lett.* **95**, 085503-1–4.
Miao, J. & Sayre, D. (2000). *Acta Cryst.* A**56**, 596–506.
Rayleigh, J. W. S. (1900). *Philos. Mag.* **49**, 539–540. Republished by Dover (1964).
Russel Luke, D. (2005). *Inverse Problems*, **21**, 37–50.
Rytov, S. M. (1937). *Izv. Akad. Nauk. SSSR Ser. Fiz.* **2**, 223–259.
Sayre, D. (1980). *Imaging Processes and Coherence in Physics*, edited by J. Schlenker, M. Fink, J. P. Goedgebuer, V. Malgrange, J. C. Viénot & R. H. Wade, Vol. 112, pp. 229–235. Berlin: Springer-Verlag.
Shapiro, D., Thibault, P., Beetz, T., Elser, V., Howells, M. R., Jacobsen, C., Kirz, J., Lima, E., Miao, H., Nieman, A. M. & Sayre, D. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 15343–15346.
Slepian, D. & Pollack, H. O. (1961). *Bell Syst. Tech.* **40**, 43–64.
Spence, J. C. H., Weierstall, U. & Howells, M. R. (2002). *Philos. Trans. R. Soc. London Ser. A*, **360**, 875–895.
Stroke, G. W. & Falconer, D. G. (1964). *Phys. Lett.* **13**, 306–309.
Thylén, L. (1983). *Opt. Quant. Electron.* **15**, 433–439.
Van Roey, J. & Lagasse, P. E. (1981). *J. Opt. Soc. Am. A*, **71**, 803–810.
Williams, G. J., Pfeifer, M. A., Vartanyants, I. A. & Robinson, I. K. (2003). *Phys. Rev. Lett.* **90**, 175501-1–4.
Wu, J. S., Weierstall, U., Spence, J. C. H. & Koch, C. T. (2004). *Opt. Lett.* **29**, 2737–2739.