Quantitative structure retrieval using scanning transmission electron microscopy

S. D. Findlay

School of Physics, University of Melbourne, Victoria 3010, Australia. Correspondence e-mail: s.findlay@physics.unimelb.edu.au

A method is described that reconstructs the projected object potential using data recorded in the coherent imaging mode of a scanning transmission electron microscope. The technique is applicable in the presence of multiple scattering. It is not required that the thickness is known. Model examples exploring the nature of the data set required, the stability of the algorithm and the limitations on resolution are provided.

1. Introduction

Electron probes offer many advantages in investigating materials on the nanoscale. At high energies, they can penetrate specimens, thus not restricting interactions to the surface. By virtue of their charge, electron beams may be manipulated and focused more easily than either X-rays or neutrons. Yet there arise limitations precisely because of the strength of the interaction between electrons and material solids. High-energy electrons are liable to undergo multiple scattering in all but the thinnest specimens and images resulting from strong multiple scattering are difficult to interpret in terms of the sample structure. Thus quantitative electron microscopy is limited either by the need to prepare very thin samples or through the difficulties of interpreting images formed from multiply scattered electrons.

Phase retrieval is a vital step in most structure-retrieval procedures in electron microscopy since knowledge of the *complex* wavefield represents complete knowledge of the wavefield beyond the specimen. The most direct forms of structure retrieval rely on very thin samples and assume either a single scattering approximation, in which there is a direct correspondence between the complex exit wave and the projected potential of the specimen, or a phase-object approximation, in which there is a direct correspondence between the phase of the exit wave and the projected potential of the specimen. Early attempts to move beyond these approximations by iterative methods include the work of Gribelyuk (1991) and Beeching & Spargo (1993). However, the stability of these algorithms remained limited to small thicknesses.

Iterative techniques, such as iterative phase retrieval (Allen *et al.*, 2004*a*), may be motivated on the basis of information: a quantity of known data is used to constrain and determine a quantity of unknown data. This balance between known and unknown data suggests that the situation can also be viewed as a minimization problem. Lentzen & Urban use this concept, applying simulated-annealing and maximum-likelihood techniques to complex wavefunction data to retrieve complex potential data (Lentzen & Urban, 1996, 2000). O'Leary & Allen (2005) recast this problem in Bloch-wave form to use non-linear-equation-solving methods on the Fourier coefficients of the wavefunction to obtain the Fourier coefficients of the potential.

Such methods become even more reliable when the problem is over-determined. One way to over-determine the problem is to use additional *a priori* information about the specimen. The channelling model of Van Dyck and co-workers assumes that crystals have a columnar structure and so reduces the number of unknowns to the column locations and types (Van Dyck & Chen, 1999; Geuens & Van Dyck, 2002). Another way to over-determine the problem is to obtain additional experimental data. Rez (1999) considered data from different accelerating voltages and Allen, Koch *et al.* (2001) considered data from different thicknesses. The over-determined approach was taken still further by Allen and co-workers (Allen *et al.*, 1999, 2000; Allen & Oxley, 2001), extending previous work (Allen *et al.*, 1998; Spence, 1998), who used complex wavefunctions from multiple tilts of an incident plane wave to reduce the problem to an over-determined set of linear equations, ensuring uniqueness of solution. The method handles all orders of multiple scattering but performing the accurate series of tilts is difficult to implement experimentally. We seek to retain the strengths of this approach to structure retrieval while circumventing the obstacle of performing the tilt series.

Geometries other than that of the conventional transmission electron microscope are available. One is scanning transmission electron microscopy, which is particularly tantalizing for our purposes because the convergent probe contains a range of incident directions (Spence & Cowley, 1978). The scanning transmission electron microscope is generally used to measure incoherent signals, as in high-angle annular dark-field imaging and in electron-energy-loss spectroscopy. Such techniques are useful because the images are often directly interpretable but this has reduced the tendency for theoretical simulations to be seen as a necessary adjunct to experiment. Inversion techniques have been presented for these imaging

modes assuming the incoherent imaging model, whereby the image is taken to be the convolution of the probe intensity with an object function for the specimen (Nellist & Pennycook, 1998a) but the assumptions on which this model is based may cease to be valid as probe size decreases (Ishizuka, 2001). Inversion techniques using coherent imaging have been considered. In the phase-object approximation, Rodenburg & Bates (1992) showed that diffraction pattern data recorded over a range of probe positions are sufficient to determine the transmission function and thus the projected potential. Plamann & Rodenburg (1998) took this a step further in describing dynamical effects with thickness presenting an inversion using what they called the improved phase-object approximation, where the approximation is taken to second order in thickness by making the reference plane of the phase object the centre of the crystal rather than the exit face. Convergent-beam electron diffraction offers other possible approaches to structure retrieval. For example, Vincent *et al.* (1999) applied a phase-retrieval-style algorithm to the intensity distribution within a dark-field disc to obtain thickness information about the displacement of reflecting planes.

A new approach is presented here that uses the scanning transmission electron microscope to collect data sets like those used by Rodenburg and co-workers but which goes beyond the assumption of the phase-object approximation to tackle the multiple-scattering problem after the manner of Allen and co-workers. This requires both an increased size for the data set and the separate implementation of a set of phase retrievals. However, it retains the attractive properties previously emphasized by Allen and co-workers: it is able to handle multiple scattering and it does not require that the sample thickness be known.

2. The scattering matrix and the structure matrix

Consider a slab-like crystal with surface area greatly exceeding the thickness. Define a coordinate system such that the reciprocal-lattice vectors in the set $\{\mathbf{G}\}$ are parallel to the plane of the crystal surface and the z direction is normal to the crystal surface. Moreover, we shall identify the optical axis with this z direction irrespective of the direction of the incident beam, thus working in a basis fixed in the crystal. Let $u_{\mathbf{G}}$ and $v_{\mathbf{G}}$ denote the Fourier coefficients of the wavefunction in the plane of the entrance and exit surfaces, respectively. The scattering matrix $\mathcal{S}(t)$, where t is the crystal thickness, is defined to be the matrix operator connecting the incident and exit waves. [The scattering-matrix approach was introduced to electron microscopy by Sturkey (1962), though the nomenclature has since changed somewhat.] Thus, in matrix notation,

$$v_{\mathbf{G}} = \sum_{\mathbf{H}} S_{\mathbf{G},\mathbf{H}}(t) u_{\mathbf{H}}, \qquad (1)$$

where the entrance wave has the form

$$\psi_u(\mathbf{r}_\perp) = \sum_{\mathbf{G}} u_{\mathbf{G}} \exp(2\pi i \mathbf{G} \cdot \mathbf{r}_\perp), \qquad (2)$$

in which $\mathbf{r}_\perp$ is a position vector in the plane of the crystal surface. The exit wave has an analogous form.

In conventional transmission electron microscopy, where a single plane wave is incident upon the crystal, the symmetric Laue position is described by $u_{\mathbf{G}} = \delta_{\mathbf{G},\mathbf{0}}$. It follows immediately from equation (1) that the Fourier coefficients of the exit-surface wave are contained in the $\mathbf{G} = \mathbf{0}$ column of the scattering matrix.

The scattering matrix is related to the projected potential of the specimen *via*

$$\mathcal{S}(t) = \exp\left[\frac{i\pi t}{K}\mathcal{A}\right], \qquad (3)$$

where $K = 1/\lambda$ is the refraction-corrected wavevector and the structure matrix $\mathcal{A}$ is given by

$$\mathcal{A} = \begin{pmatrix} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & -\mathbf{H}^2 + iU'_{\mathbf{0}} & W_{\mathbf{H-G}} & W_{\mathbf{H}} & W_{\mathbf{H+G}} & W_{\mathbf{2H}} & \cdots \\ \cdots & W_{\mathbf{G-H}} & -\mathbf{G}^2 + iU'_{\mathbf{0}} & W_{\mathbf{G}} & W_{\mathbf{2G}} & W_{\mathbf{G+H}} & \cdots \\ \cdots & W_{\mathbf{-H}} & W_{\mathbf{-G}} & iU'_{\mathbf{0}} & W_{\mathbf{G}} & W_{\mathbf{H}} & \cdots \\ \cdots & W_{\mathbf{-G-H}} & W_{\mathbf{-2G}} & W_{\mathbf{-G}} & -\mathbf{G}^2 + iU'_{\mathbf{0}} & W_{\mathbf{-G+H}} & \cdots \\ \cdots & W_{\mathbf{-2H}} & W_{\mathbf{-H-G}} & W_{\mathbf{-H}} & W_{\mathbf{-H+G}} & -\mathbf{H}^2 + iU'_{\mathbf{0}} & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}, \qquad (4)$$

in which $W_{\mathbf{G}} = U_{\mathbf{G}} + iU'_{\mathbf{G}}$, where $U_{\mathbf{G}}$ and $U'_{\mathbf{G}}$ are the Fourier coefficients of the elastic and absorptive potentials, respectively (Allen *et al.*, 2000). Note that the same Fourier coefficients of potential occur in several positions in the structure matrix.

Expanding equation (3) in a Taylor series is informative:

$$\mathcal{S}(t) \simeq \mathcal{I} + \frac{i\pi t}{K}\mathcal{A} + \left(\frac{i\pi t}{K}\right)^2 \mathcal{A}^2 + \dots. \qquad (5)$$

The different powers of the structure matrix may be interpreted as different orders of scattering. For instance, the $\mathcal{A}^2$ term may be interpreted as describing the contribution due to double scattering. In the single-scattering approximation, the series is truncated to first order in $\mathcal{A}$ and there is a one-to-one correspondence between the elements of the scattering matrix $\mathcal{S}$ and the structure matrix $\mathcal{A}$. For a sufficiently large structure matrix, all coefficients of potential will appear in the central column of $\mathcal{A}$ and it suffices in this approximation to determine only the central column of the scattering matrix $\mathcal{S}$ – precisely what is done in conventional transmission electron microscopy.

The situation is complicated when one moves beyond the single-scattering approximation because the higher-order contributions introduce non-linear terms in the object potential. However, Allen and co-workers (Allen *et al.*, 1999, 2000) showed that if the entire scattering matrix is obtained then the inversion problem may be reduced to a linear problem, which is therefore unique and well defined. This is possible because the scattering matrix and the structure matrix have the same eigenvector spectrum. The structure matrix may be written in terms of the matrix of eigenvectors $\mathcal{C}$ and the eigenvalues λ^i *via* the spectral representation

$$\mathcal{A} = \mathcal{C}[\lambda^i]_D \mathcal{C}^{-1}. \qquad (6)$$

Using equation (3), the scattering matrix may thus be written as

$$\mathcal{S}(t) = \mathcal{C} \left[\exp\left(\frac{i\pi t}{K}\lambda^i\right) \right]_D \mathcal{C}^{-1}. \quad (7)$$

So the eigenvectors of the structure matrix are obtained immediately on finding those of the scattering matrix. Obtaining the eigenvalues of the structure matrix from those of the scattering matrix is hampered because there is a uniqueness problem in taking the complex logarithm. Instead, we use the determined eigenvectors together with the known diagonal elements and symmetries of the structure matrix to generate from equation (6) a set of linear equations satisfied by the eigenvalues (Allen *et al.*, 2000). It is interesting to note that this approach does not require the thickness to be known, since it appears only with the unused eigenvalues of the scattering matrix.

A method for obtaining the entire scattering matrix using scanning transmission electron microscopy will be explained in the next section. For comparison purposes, a method for obtaining the entire scattering matrix using conventional transmission electron microscopy is given here. Suppose that we have an incident plane wave with tangential momentum component **H**. In reciprocal-space notation, equation (1) gives that the exit-surface wavefunction has Fourier coefficients

$$\Psi_{\mathbf{H}}(\mathbf{G}, t) = \mathcal{S}_{\mathbf{G},\mathbf{H}}(t). \quad (8)$$

This wavefunction is then transmitted through a lens, which is described *via* the transfer function

$$T_{\Delta f}(\mathbf{q}) = A(\mathbf{q}) \exp[-i\chi(\mathbf{q}, \Delta f)], \quad (9)$$

where $A(\mathbf{q})$ is the pupil aperture function and $\chi(\mathbf{q}, \Delta f)$ is the aberration function. The defocus value is explicitly emphasized because its variation is key to the phase-retrieval method to be used. Phase retrieval may be effected by varying any of the coherent aberrations (Allen, Oxley & Paganin, 2001) but for simplicity we consider only defocus.

The measured intensity is then

$$I_{\Delta f,\mathbf{H}}(\mathbf{r}_\perp, t) = \left| \Psi_{\Delta f,\mathbf{H}}(\mathbf{r}_\perp, t) \right|^2$$

$$= \left| \sum_{\mathbf{G}} \mathcal{S}_{\mathbf{G},\mathbf{H}}(t) T_{\Delta f}(\mathbf{G}) \exp(2\pi i \mathbf{G} \cdot \mathbf{r}_\perp) \right|^2. \quad (10)$$

Phase retrieval, the umbrella term for a slew of techniques that use intensity information to determine the phase, is necessary to obtain the wavefunction $\Psi_{\Delta f,\mathbf{H}}(\mathbf{r}_\perp, t)$ from the measured intensity. One method of phase retrieval is the through-focal-series method, which uses the intensities recorded at a range of defocus values to determine the phase *via* the requirement of consistent propagation between these images. Further details may be found in Allen *et al.* (2004*a*). From the complex wavefunction so determined, one can calculate the values of $\mathcal{S}_{\mathbf{G},\mathbf{H}}$. Since phase retrieval only determines the phase up to an arbitrary, additive, constant and the retrievals for different tilts are independent, different columns in the scattering matrix are not correctly phased with respect to one another. However, the relative phasing can be

accomplished because the symmetry of the structure matrix across the anti-diagonal, visible in equation (4), carries over to the scattering matrix and provides a link between the columns.

3. Measuring the scattering matrix with the scanning transmission electron microscope

Allen and co-workers (Allen *et al.*, 1999, 2000) proposed that the scattering matrix should be obtained with a conventional transmission electron microscope by performing a series of experiments at different tilts of the incident beam. Spence (1998) suggested using overlapping discs in convergent-beam electron diffraction geometry to determine the phases of diffraction beams pair-wise, in the manner of ptychography (Nellist & Rodenburg, 1998). A method for obtaining these data with the scanning transmission electron microscope using a through-focal-series approach to phase retrieval will be presented. This offers no saving in either the size of the data set recorded or the need to perform phase retrieval but it does circumvent the need to perform experiments for a range of different incident tilts – the convergent probe of the scanning transmission electron microscope contains a range of tilts already.

In the scanning transmission electron microscope, the objective lens is positioned before the specimen and used to form a convergent probe (Spence & Cowley, 1978). Thus, the incident wavefunction upon the crystal is, in reciprocal space, given by the lens transfer function of equation (9) with the addition of a phase term for the position **R** of the probe:

$$u_{\mathbf{H}} = T_{\Delta f}(\mathbf{H}) \exp(-2\pi i \mathbf{H} \cdot \mathbf{R}). \quad (11)$$

From equation (1), it readily follows that the exit-surface wavefunction is given by

$$\psi_{\Delta f}(\mathbf{r}_\perp, t, \mathbf{R}) = \sum_{\mathbf{H},\mathbf{G}} S_{\mathbf{G},\mathbf{H}}(t) T_{\Delta f}(\mathbf{H}) \exp(-2\pi i \mathbf{H} \cdot \mathbf{R})$$

$$\times \exp(2\pi i \mathbf{G} \cdot \mathbf{r}_\perp) \quad (12)$$

and the intensity in the diffraction plane is thus

$$I_{\Delta f}(\mathbf{G}, t, \mathbf{R}) = \left| \sum_{\mathbf{H}} S_{\mathbf{G},\mathbf{H}}(t) T_{\Delta f}(\mathbf{H}) \exp(-2\pi i \mathbf{H} \cdot \mathbf{R}) \right|^2. \quad (13)$$

Note that in these equations only the set $\{\mathbf{G}\}$ is used, *i.e.* the physical reciprocal-lattice vectors. The convergent probe contains a continuum of transverse momenta, however it has been shown that for perfect crystals the wavefunction de-couples into a set of wavefunctions each pertaining to a different tilt within the first Brillouin zone and each containing only vectors differing by physical reciprocal-lattice vectors (Findlay *et al.*, 2003). Thus in the crystalline case one may, without loss of generality, consider just one such set, provided measurements are restricted to the diffraction plane where this set is distinguishable from the others by spatial separation once the angular calibration has been made.

Note also that this intensity is effectively four-dimensional: for each *physical* probe position **R** in a two-dimensional scan, a *physical* diffraction pattern with points labelled by the two-dimensional reciprocal-lattice vector **G** is recorded. 'Physical'

has been emphasized because the experiment dictates an obvious interpretation of $I_{\Delta f}(\mathbf{G}, t, \mathbf{R})$: view $\mathbf{R}$ as a parameter and, for each value of this parameter, view $\mathbf{G}$ as the dynamic variable describing the image. In principle, the machine could be shut down and restarted between recording such a data set and recording that for the next position $\mathbf{R}$ as there is no dynamical connection between measurements at different probe positions.

However, comparing the general form of equation (13) with equation (10) suggests that, mathematically, another interpretation is possible. One might instead treat $\mathbf{G}$ as a parameter and $\mathbf{R}$ as the dynamical variable, forming intensity 'images' as a function of $\mathbf{R}$ for a fixed value of $\mathbf{G}$. This is not new. Taking $\mathbf{G} = 0$ and forming an image as a function of $\mathbf{R}$ is known as a bright-field image and is often related *via* reciprocity arguments to measurements in conventional transmission electron microscopy (Cowley, 1969).

Thinking about equation (13) in this way motivates an alternative option for phase retrieval. The obvious option is to perform phase retrieval on the physical wavefunction, on the images or diffraction patterns formed for a fixed value of $\mathbf{R}$. But having appreciated the similarity between equations (10) and (13), one may use phase retrieval on the latter, taking images as a function of $\mathbf{R}$ for a fixed value of $\mathbf{G}$. This corresponds to no physical wavefunction; we are not retrieving the phase of anything that propagates. However, the 'image' so described behaves and changes with defocus as if it did and the technique allows for the phasing of the coefficients involved in exact analogy to the components of a wavefunction.

To draw out the similarity with equation (10), let us rewrite equation (13) as

$$I_{\Delta f, \mathbf{G}}(\mathbf{R}, t) = \left| \sum_{\mathbf{H}} S^*_{\mathbf{G},\mathbf{H}}(t) T^*_{\Delta f}(\mathbf{H}) \exp(2\pi i \mathbf{H} \cdot \mathbf{R}) \right|^2. \quad (14)$$

The phase-retrieval algorithm used tends to be more stable if the dummy summation variable $\mathbf{H}$ is changed to $\mathbf{H} + \mathbf{G}$:

$$I_{\Delta f, \mathbf{G}}(\mathbf{R}, t) = \left| \sum_{\mathbf{H}} S^*_{\mathbf{G},\mathbf{H}+\mathbf{G}}(t) T^*_{\Delta f}(\mathbf{H} + \mathbf{G}) \exp(2\pi i \mathbf{H} \cdot \mathbf{R}) \right|^2. \quad (15)$$

Having retrieved the phase, shifting back to the original basis is straightforward.

For a fixed value of $\mathbf{G}$, the through-focal-series technique is used to retrieve the reciprocal-space 'wavefunction' $S^*_{\mathbf{G},\mathbf{H}}(t) T^*_{\Delta f}(\mathbf{H})$, which we aberration-correct to yield $S_{\mathbf{G},\mathbf{H}}(t)$. This then gives, correctly phased, the elements within one row of the scattering matrix (*cf.* within a column, as was obtained in conventional transmission electron microscopy). Repeating the procedure for different values of $\mathbf{G}$ determines the full scattering matrix. The different rows are not automatically phased correctly but this may be accomplished *via* the aforementioned symmetry across the anti-diagonal of the scattering matrix.

Thus we have a method of determining the entire scattering matrix, and therefore carrying out the inversion procedure, within the scanning transmission electron microscope. The requisite data are a set of diffraction patterns for different positions of the probe on the surface for a range of defocus values. The full method will be illustrated through simulation. But first we shall discuss further some of the experimental details that will affect the retrieval.

4. Experimental factors

4.1. Truncation

The inversion procedure, which requires the determination of eigenvectors, must be applied to a square matrix. One may in principle collect images in the form of equation (13) out to very large values of $\mathbf{G}$, but in performing the aberration-correction step it becomes clear that only those values of $S_{\mathbf{G},\mathbf{H}}$ for which $T_{\Delta f}(\mathbf{H}) \neq 0$, for which $\mathbf{H}$ fits within the probe-forming aperture, may be obtained. The largest square scattering matrix is limited to this resolution. As per equation (3), the scattering matrix is formed from the matrix exponential of the structure matrix. The structure matrix is formally of infinite order and therefore so too is the scattering matrix. Numerical modelling requires treatment of finite matrices but it may prove that adequate convergence requires them to be of large order. By performing the inversion on a truncated scattering matrix, we treat a model problem with scattering and structure matrices of relatively small order. The results will only be valid if the elements in the retrieved finite structure matrix are good approximations to those of the full structure matrix. This will be the case if the scattering matrix formed by the truncated structure matrix almost converges to the truncated portion of the scattering matrix as calculated from the full structure matrix. If this is not the case, then aliasing effects serve to distort many of the Fourier coefficients of potential retrieved.

This problem, called the truncation problem, has been discussed in the systematic row case by Allen & Oxley (2001). The systematic row case converges much faster than the general zone-axis case but the qualitative conclusions drawn remain the same. It was found that the aliasing acted from the off-diagonal corners inwards, being more significant for more severe truncations. Truncation tended to significantly overestimate the magnitude of higher-order Fourier coefficients and so, with the expectation that such Fourier coefficients should tend to decrease with increasing spatial frequency, it may be possible to identify suspect Fourier coefficients of the retrieved potential by inspection. Truncation will prove to be the most severe limitation on the inversion procedure in practice and will be discussed after presenting the simulations.

4.2. Phase retrieval

Determination of the scattering matrix as discussed in §3 requires a suitable phase-retrieval algorithm. This algorithm should be reasonably robust in the presence of noise, able to handle the possibility of vortices in the phase and ideally be able to correct for (or failing that to cope with) the presence of partial coherence. A through-focal-series method for phase retrieval was chosen, since it is known to satisfy the first two conditions (Allen, Faulkner *et al.*, 2001). The global iterative

method of Allen and co-workers (Allen *et al.*, 2004*a*,*b*) is another possibility, which may handle noise and partial coherence more rigorously, though this more elaborate course will not be pursued here.

While incoherent imaging in scanning transmission electron microscopy seems to be fairly robust in the presence of partial coherence (Nellist & Pennycook, 1998*b*), the coherent imaging mode is affected by it. It is readily shown from equation (14) that the effect of the spread in defocus values on the intensity images is identical to its effect in conventional transmission electron microscopy (Allen *et al.*, 2004*a*). Specifically, averaging equation (14) over defocus using a Gaussian distribution with a $1/e$ value of Δ, centred about Δf, gives

$$
\begin{aligned}
I_{\Delta f, \mathbf{G}}(\mathbf{R}, t) = \sum_{\mathbf{H}, \mathbf{H}'} & S^*_{\mathbf{G}, \mathbf{H}}(t) S_{\mathbf{G}, \mathbf{H}'}(t) T^*_{\Delta f}(\mathbf{H}) T_{\Delta f}(\mathbf{H}') \\
& \times \exp[2\pi i(\mathbf{H} - \mathbf{H}') \cdot \mathbf{R}] \\
& \times \exp[-\pi^2 \lambda^2 \Delta^2 (\mathbf{H}^2 - \mathbf{H}'^2)^2/4].
\end{aligned} \tag{16}
$$

As a consequence of the final exponential, this result is no longer exactly expressible as the modulus squared of a single wavefunction. The full form above is used to perform the simulation but when attempting to correct for the spread in defocus, we shall make the factorization approximation

$$
\begin{aligned}
\exp[-\pi^2 \lambda^2 \Delta^2 (\mathbf{H}^2 - \mathbf{H}'^2)^2/4] \\
\simeq \exp(-\pi^2 \lambda^2 \Delta^2 \mathbf{H}^4/4) \exp(-\pi^2 \lambda^2 \Delta^2 \mathbf{H}'^4/4),
\end{aligned}
$$

and deconvolve to remove the resultant envelope function (Coene *et al.*, 1996; Allen *et al.*, 2004*a*).

The effect of the finite source size is somewhat different in scanning transmission electron microscopy, being equivalent to the effect of finite detector pixel size in conventional transmission electron microscopy (Nellist & Rodenburg, 1994). This can be described as a convolution over probe position of the image with a distribution resulting from the finite source size. Provided it can be quantified, this issue can be corrected by deconvolution applied to the images prior to phase retrieval. As such, this effect is not included in the simulations presented.

There is an issue relating to our ability to determine the precise locations of the diffraction beams $\mathbf{G}$: the possibility of some systematic offset, measuring beams $\mathbf{G} + \mathbf{q}$ rather than $\mathbf{G}$. In this case, we would record an image

$$
I_{\Delta f, \mathbf{G}+\mathbf{q}}(\mathbf{R}) = \left| \sum_{\mathbf{H}} S^*_{\mathbf{G}+\mathbf{q}, \mathbf{H}+\mathbf{q}}(t) T^*_{\Delta f}(\mathbf{H}+\mathbf{q}) \exp[2\pi i(\mathbf{H}+\mathbf{q}) \cdot \mathbf{R}] \right|^2. \tag{17}
$$

The scattering matrix $S_{\mathbf{G}+\mathbf{q}, \mathbf{H}+\mathbf{q}}(t)$ is formed from equation (3) but using a structure matrix with diagonal entries shifted as $\mathbf{G} \rightarrow \mathbf{G} + \mathbf{q}$. This does not affect the symmetries off the diagonal of the structure matrix. But it does mean our assumption of its diagonal elements will be in error, and these are the inhomogeneous equations used to find the eigenvalues. It also means that the symmetry across the anti-diagonal in the scattering matrix, used to phase the different rows, will no longer be precise. Because of the offset $\mathbf{q}$ in the transfer function, the phase retrieval is much less successful. We shall

not simulate this possibility, assuming either that the diffraction pattern is adequately calibrated or that the success of the phase retrieval can be used to determine the most favourable value of $\mathbf{q}$ to use.

4.3. Validity check

Thus far, we have suggested factors that might limit the accuracy of the retrieved potential. What validity check might be applied to the results? The simplest course would use the retrieved result to solve the direct problem – to simulate the scattering matrix – and compare with the experimental data. That the scattering-matrix inversion does not require the thickness to be known was an advantage in the inverse problem. However, this deficit of knowledge now prevents solving the direct problem.

From a converged retrieval, one might try using minimization to determine the thickness by comparing simulated and experimental scattering matrices. Let us define a measure-of-fit function for the unknown thickness *via*

$$
F(t) = \sum_{\mathbf{G}, \mathbf{H}} \left| \frac{S^{\text{sim}}_{\mathbf{G}, \mathbf{H}}(t)}{S^{\text{sim}}_{\mathbf{0}, \mathbf{0}}(t)} - \frac{S^{\text{exp}}_{\mathbf{G}, \mathbf{H}}}{S^{\text{exp}}_{\mathbf{0}, \mathbf{0}}} \right|^2, \tag{18}
$$

where $S^{\text{sim}}_{\mathbf{G}, \mathbf{H}}(t)$ denotes an element of the simulated scattering matrix using the retrieved potential and guessed thickness t, while $S^{\text{exp}}_{\mathbf{G}, \mathbf{H}}$ denotes the corresponding scattering-matrix element of the experimental data. Both quantities have been normalized by the central scattering matrix element because the absolute magnitude and phase of the experimental data are likely to be unknown. This was not a problem for our inversion method: multiplicative constants do not affect the eigenvectors, only the unused eigenvalues.
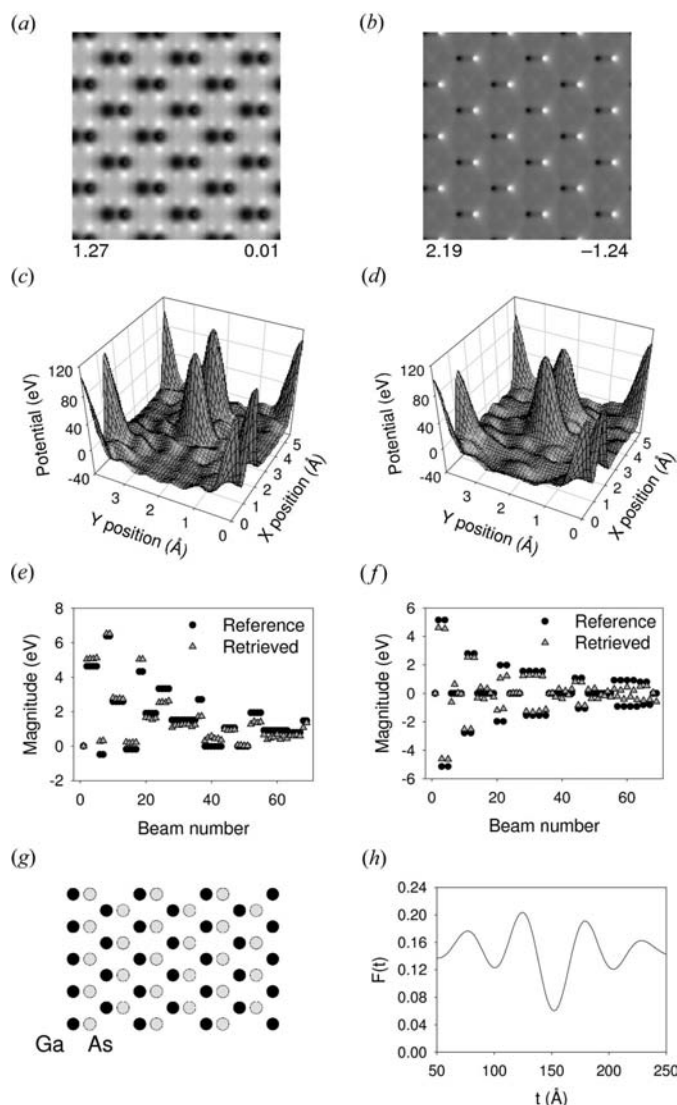
As a prescription for limiting the effects of aliasing resulting from truncation, we shall only take from the retrieval those Fourier coefficients of potential that lie in the central column of the structure matrix. For the purposes of forming $S^{\text{sim}}_{\mathbf{G}, \mathbf{H}}(t)$, we shall therefore make a scattering matrix from these Fourier coefficients only; elements that are not by symmetry identical to any in the central column are set to zero. We shall also force the known diagonal elements but set the mean absorption $U'_{\mathbf{0}}$ to zero since the normalization will eliminate it anyway.

5. Case study

Consider a 152 Å thick specimen of GaAs in the [110] orientation to be probed with 200 keV electrons. The scattering matrix was simulated using 261 beams but was then truncated to 69 beams, corresponding to a 36 mrad aperture, prior to retrieval. While this is a larger aperture than generally used in scanning transmission electron microscopy, it should be appreciated that the technique here does not require an especially fine probe; provided that the aberrations are well characterized, they need not be perfectly balanced out to the edge of the aperture, thus allowing for the use of larger apertures.

Figs. 1(*a*) and (*b*) show the intensity and phase of the exit-surface wave assuming normal plane-wave incidence. The

differences in form between the image and phase and the projected potential show that we are well beyond the range of single-scattering or phase-object approximations. For comparison, Fig. 1(g) shows schematically the projected structure. Five defocus values, in even steps from 0 to 400 Å, were used to generate the data to which through-focal series phase retrieval was applied. Prior to the phase retrieval and inversion, noise was added to the diffraction-pattern images using Poisson counting statistics with the images normalized such that the maximum number of counts is 10000 and will thus have 1% noise added. This means that pixels with an order of magnitude fewer counts will have noise at the 3%

level. Furthermore, temporal incoherence was simulated assuming a 0.3 eV full-width-at-half-maximum spread at the tip and a chromatic aberration coefficient of $C_c = 1.5$ mm, leading to a $1/e$ value of 13.5 Å in defocus spread. Subsequent to the phase retrieval, we shall attempt to correct for this using the factorized envelope approximation (Coene *et al.*, 1996; Allen *et al.*, 2004a). The through-focal series of five images was generated for each of the 69 beams in the objective aperture. Each image was 64×45 pixels, corresponding to the raster of probe positions within the repeating unit. On an Intel XEON 2.00 GHz machine, the complete phase retrieval took about 2 min, using 200 iterations for each series. For matrices of order less than 100, the inversion step takes a negligible amount of computation time. The slowest part of the calculation was the direct simulation accounting fully for chromatic aberration – this took several hours.

Figs. 1(c) and (d) show the projected potential for the input and retrieved data using only the 69 Fourier coefficients on the central column of the structure matrix. The retrieved potential is in fair agreement with the reference structure, though the retrieved potential erroneously has the As columns less pronounced than the Ga columns. Figs. 1(e) and (f) show, for the real and imaginary parts of the Fourier coefficients of the potential, the comparison between the input and retrieved data. The effects of noise are manifest in differences between coefficients that in the perfect structure would be identical. Fig. 1(h) shows $F(t)$ as defined by equation (18) and the global minimum at a thickness of 150 Å is in excellent agreement with the true thickness of 152 Å.
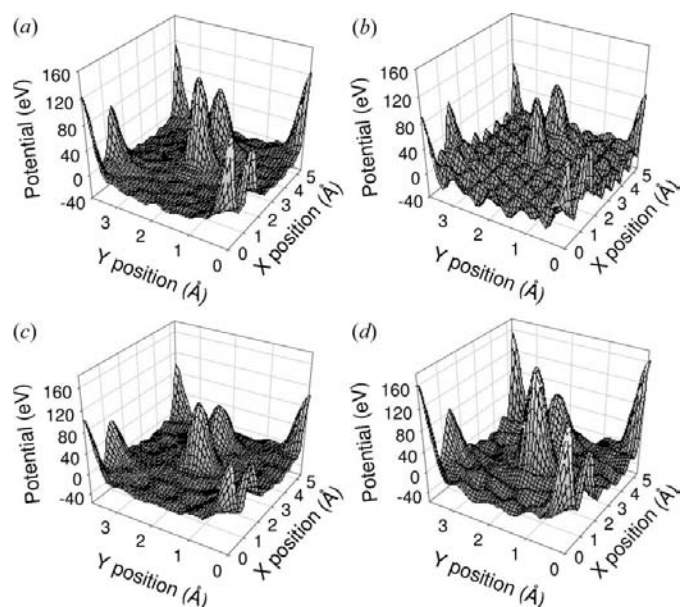
6. Discussion

It was suggested earlier that the greatest limitation on the method was the error involved in the truncation step. In the Bloch-wave method, the number of beams is selected according to some convergence criterion. For plane-wave incidence, the number of beams should be large enough that the central column, which contains the Fourier coefficients of the exit-surface wavefunction, is converged. The inversion procedure requires not a single column but a central submatrix. If we simulate a 69 beam matrix, then the inversion applied to that matrix is a self-consistent problem. But the converged calculation would require simulating for a matrix sufficiently large that the central 69-by-69 beam submatrix is essentially converged. This was approximated in the previous section by truncating a 261 beam matrix. Such a truncated problem would only be exactly consistent if the submatrix appeared in a block-diagonal structure, since only in such a case would the selected elements be independent of the omitted ones. This will never quite happen in practice because the repetition of elements in the structure matrix ensures that low-order Fourier coefficients recur in structure matrices of all orders.

That truncation is unavoidable begs the question of whether the inversion is stable. Unfortunately this is not guaranteed. The advantages garnered by converting the problem into a set of linear equations for the unknown eigenvalues come

Figure 1
(a) Intensity of exit-surface wavefunction for 152 Å thick GaAs in [110] orientation. Maximum and minimum intensities are given below the image. (b) Phase of the same exit-surface wavefunction. Maximum and minimum phase values are given below the image. (c) True potential using the first 69 Fourier coefficients. (d) Potential retrieved by 69 beam truncation of a 261 beam scattering matrix, including 1% noise and temporal incoherence. (e) Real component of Fourier coefficients. (f) Imaginary component of Fourier coefficients. (g) Sketch of projected GaAs structure. (h) $F(t)$ as defined by equation (18).

through the assumption that the eigenvectors of the scattering matrix are identically those of the structure matrix. The inversion is therefore unavoidably limited by how well the eigenvectors of the scattering matrix, subsequent to truncation and containing the errors incurred by noise and incoherent effects, match those of the structure matrix.

To demonstrate this possible instability in the inversion, we simulate the inversion step for a 300 Å thick sample of ZnS in the [110] orientation using 300 keV electrons. The different sample and parameters were chosen to illustrate the instability, which does not occur in the case study of the previous section. The scattering matrix was simulated using 459 beams. Fig. 2(a) shows the true projected potential using the 103 Fourier coefficients (corresponding to a 36 mrad aperture) on the central column of the structure matrix. Fig. 2(b) shows the potential retrieved by inversion of the scattering matrix truncated to 103 beams. Figs. 2(c) and (d) show similar results for 63 beams (corresponding to a 26 mrad aperture). It is apparent that the retrieval from 103 beams is worse than that from 63 beams, containing significant fluctuations and little differentiation between the two column types. The 63 beam inversion, while notably over-estimating the magnitude of the potential, is in good qualitative agreement with the true structure.

Strict rules have not been established as to when one should expect an inversion to succeed or when it might struggle. One indicator is the spread of significant magnitudes of scattering-matrix elements away from the diagonal. For a submatrix to behave as though it came from a block-diagonal matrix, one desires the intensity near the top left and bottom right corners of the matrix to be concentrated near the diagonal. This may provide that the elements in a single column discarded by the truncation are small. Cases have been found for an Si_3N_4 simulated inversion, in which there was significant spread away from the diagonal axis in the scattering matrix, that showed unstable oscillatory behaviour persisting to sub-matrices of order much larger than corresponds to realistic objective aperture sizes. That said, a variety of samples and thicknesses was found that were amenable to this inversion procedure and which possessed a tendency for the inversion to improve, though not monotonically, with increased aperture size.

Thus far, we have determined a fair estimate for a range of lower-order Fourier coefficients – range depending on aperture size – and also the thickness. The first of these stages was deterministic, the inversion procedure produces a unique result, though this does not give an assurance of high precision in the retrieved structure because of the effects of truncation. The second stage was based on minimization for a single variable but in all our simulations this minimum was well defined. The quantity of experimental data suggests that one might now switch to an optimization procedure to refine our values, thus improving the good qualitative agreement of Fig. 2(d) to good quantitative agreement and even extend the resolution (i.e. obtain some potential coefficients off the truncated central column). Previous optimization approaches were limited because they required a reasonable starting guess, which was obtained via either single scattering or channelling approaches, both of which have a thickness-limited domain of validity (Lentzen & Urban, 1996, 2000; O'Leary & Allen, 2005). The result of our current method could thus be combined with the experimental data and the techniques of previous authors to extend the precision of the retrieved potential.

7. Conclusions

Using simulations that include realistic constraints such as counting noise and temporal incoherence, we have demonstrated a structure retrieval method that uses diffraction pattern data collected using the scanning transmission electron microscope in coherent imaging mode to reconstruct the structure matrix. Though requiring the collection of a large data set, the method can handle multiple scattering and does not require that the thickness be known – indeed, it is able to provide an excellent estimate for the thickness after the fact. This approach removes the need to take data at carefully controlled incident orientations, which has hindered the implementation of the previous inversion method along these lines.

Figure 2
(a) True potential for ZnS using first 103 Fourier coefficients. (b) Potential retrieved by 103 beam truncation of a 459 beam scattering matrix. (c) True potential using first 63 Fourier coefficients. (d) Potential retrieved by 63 beam truncation of a 459 beam scattering matrix.

References

Allen, L. J., Faulkner, H. M. L. & Leeb, H. (2000). *Acta Cryst.* A**56**, 119–126.

Allen, L. J., Faulkner, H. M. L., Oxley, M. P. & Paganin, D. (2001). *Ultramicroscopy*, **88**, 85–97.

Allen, L. J., Josefsson, T. W. & Leeb, H. (1998). *Acta Cryst.* A**54**, 388–398.

Allen, L. J., Koch, C., Oxley, M. P. & Spence, J. C. H. (2001). *Acta Cryst.* A**57**, 473–474.

Allen, L. J., Leeb, H. & Spargo, A. E. C. (1999). *Acta Cryst.* A**55**, 105–111.

Allen, L. J., McBride, W., O'Leary, N. L. & Oxley, M. P. (2004*a*). *Ultramicroscopy*, **100**, 91–104.

Allen, L. J., McBride, W., O'Leary, N. L. & Oxley, M. P. (2004*b*). *J. Microsc.* **216**, 70–75.

Allen, L. J. & Oxley, M. P. (2001). *Ultramicroscopy*, **88**, 195–209.

Allen, L. J., Oxley, M. P. & Paganin, D. (2001). *Phys. Rev. Lett.* **87**, 123902.

Beeching, M. J. & Spargo, A. E. C. (1993). *Ultramicroscopy*, **52**, 243–247.

Coene, W. M. J., Thust, A., Op de Beeck, M. & Van Dyck, D. (1996). *Ultramicroscopy*, **64**, 109–135.

Cowley, J. M. (1969). *Appl. Phys. Lett.* **15**, 58–59.

Findlay, S. D., Allen, L. J., Oxley, M. P. & Rossouw, C. J. (2003). *Ultramicroscopy*, **96**, 65–81.

Geuens, P. & Van Dyck, D. (2002). *Ultramicroscopy*, **93**, 179–198.

Gribelyuk, M. A. (1991). *Acta Cryst.* A**47**, 715–723.

Ishizuka, K. (2001). *J. Electron Microsc.* **50**, 291–305.

Lentzen, M. & Urban, K. (1996). *Ultramicroscopy*, **62**, 89–102.

Lentzen, M. & Urban, K. (2000). *Acta Cryst.* A**56**, 235–247.

Nellist, P. D. & Pennycook, S. J. (1998*a*). *J. Microsc.* **190**, 159–170.

Nellist, P. D. & Pennycook, S. J. (1998*b*). *Phys. Rev. Lett.* **81**, 4156–4159.

Nellist, P. D. & Rodenburg, J. M. (1994). *Ultramicroscopy*, **54**, 61–74.

Nellist, P. D. & Rodenburg, J. M. (1998). *Acta Cryst.* A**54**, 49–60.

O'Leary, N. L. & Allen, L. J. (2005). *Acta Cryst.* A**61**, 252–259.

Plamann, T. & Rodenburg, J. M. (1998). *Acta Cryst.* A**54**, 61–73.

Rez, P. (1999). *Acta Cryst.* A**55**, 160–167.

Rodenburg, J. M. & Bates, R. H. T. (1992). *Philos. Trans. R. Soc. London Ser. A*, **339**, 521–553.

Spence, J. C. H. (1998). *Acta Cryst.* A**54**, 7–18.

Spence, J. C. H. & Cowley, J. M. (1978). *Optik (Stuttgart)*, **50**, 129–142.

Sturkey, L. (1962). *Proc. Phys. Soc.* **80**, 321–345.

Van Dyck, D. & Chen, J. H. (1999). *Acta Cryst.* A**55**, 212–215.

Vincent, R., Walsh, T. D. & Pozzi, M. (1999). *Ultramicroscopy*, **76**, 125–137.