ACE: Automated CTF Estimation

Satya P. Mallick[a,b,*], Bridget Carragher[b], Clinton S. Potter[b], David J. Kriegman[c]

[a]*Department of Electrical and Computer Engineering, University of California, San Diego, USA*
[b]*National Resource for Automated Molecular Microscopy, Department of Cell Biology, The Scripps Research Institute, La Jolla, USA*
[c]*Department of Computer Science and Engineering, University of California, San Diego, USA*

Abstract

We present a completely automated algorithm for estimating the parameters of the contrast transfer function (CTF) of a transmission electron microscope. The primary contribution of this paper is the determination of the astigmatism prior to the estimation of the CTF parameters. The CTF parameter estimation is then reduced to a 1D problem using elliptical averaging. We have also implemented an automated method to calculate lower and upper cutoff frequencies to eliminate regions of the power spectrum which perturb the estimation of the CTF parameters. The algorithm comprises three optimization subproblems, two of which are proven to be convex. Results of the CTF estimation method are presented for images of carbon support films as well as for images of single particles embedded in ice and suspended over holes in the support film. A MATLAB implementation of the algorithm, called ACE, is freely available.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

One of the most exciting challenges for biology today is in understanding the molecular machinery of the cell as a working, dynamic system. The technique of cryo electron microscopy (cryoEM)

has a unique role to play in addressing this challenge as it can provide structural information on large macromolecular complexes in a variety of conformational and compositional states while preserved under close to physiological conditions. Traditionally the methods for cryoEM have been time consuming and labor intensive, involving data acquisition, analysis and averaging of thousands to hundreds of thousands of images of the individual macro-molecular complexes. Thus, over the last few years there has been considerable interest and substantial efforts

*Corresponding author. Department of Electrical and Computer Engineering, University of California, San Diego, USA. Tel.: +1 858 6920238; fax: +1 858 5347029.

E-mail address: spmallick@graphics.ucsd.edu (S.P. Mallick).

devoted to developing automated methods to improve the ease of use and throughput of cryoEM [1–5].

A critical step in the processing and analysis of cryoEM images involves the estimation of a variety of factors that modulate the image of the specimen and which must be corrected in order to generate an accurate 3D reconstruction of the specimen. Principal among these is the contrast transfer function (CTF) of the microscope. The effect of the CTF is to introduce spatial frequency-dependent oscillations into the Fourier space representation of the image. These effects can be readily observed using an image of an amorphous carbon film where the power spectrum of the image exhibits a series of concentric ripples called Thon rings [6]. The precise location of the zeroes in the CTF is determined by the accelerating voltage, defocus and spherical aberration of the microscope (see the review by Wade [7], for example) whereas the overall shape of the pattern is determined by the amount of axial astigmatism in the objective lens of the microscope. The Thon rings are circular when the astigmatism is zero and progressively change to elliptical, parabolic and hyperbolic patterns as the astigmatism is increased. Estimating the parameters of the CTF and correcting the image for the CTF is essential in interpreting any image beyond a resolution corresponding to the first zero of the CTF.

A further decrease in the signal strength as a function of frequency arises as a result of a variety of factors (finite electron source size, energy spread of the beam, drift, etc.) which limits the 3D reconstruction resolution that can be achieved. These combined effects can be modeled using an envelope function. Finally, the image contains a noise component which is normally modeled as an additive linear component.

The theory of contrast transfer in the electron microscope [8,9] provides a parametric form for the CTF, the envelope function and the background noise. Our objective is thus to automatically recover these parameters, which can then be used to restore the images. Given the necessity of correcting the CTF when attempting to reconstruct structures to high resolutions, many methods for estimating the parameters of the CTF have

been proposed, and several of these have been automated to some degree. However, most of the proposed solutions for estimating the astigmatism in the image are somewhat ad hoc and for the most part the astigmatism is simply assumed to be negligible. In our method we have been focusing on providing a completely automated solution that is compatible with ongoing efforts to improve the overall automation and throughput of the entire process of cryoEM structural analysis. In order to account for astigmatism we first estimate the elliptical shape of the CTF rings using edge detection methods and then use elliptical averaging to improve the signal to noise ratio for the final estimate of the CTF, the envelope function and the noise. The algorithm has been implemented as a MATLAB routine and is freely available.

2. Related work

In this section we discuss prior work on automated CTF estimation. Early works on Contrast Transfer Theory are attributed to Hanszen [10] and Thon [6]. CTF estimation was initially performed manually. Initial automated work on CTF estimation includes work by Frank et al. [11] and Henderson et al. [12], who tried to estimate all parameters of the CTF in one pass by minimizing an error function based on the squared difference between modulus of the Fourier Transform and the theoretical CTF. The envelope function was modeled by multiplying $1/s$ to the CTF, where s was the frequency. The drawback of these approaches was that the background noise spectrum was not accounted for.

Zhou et al. [13] averaged the power spectrum along concentric circles about the origin (i.e rotational averaging). They estimated the background by interpolating the values between local minima. Zhu et al. [14] used a similar approach but assumed a Gaussian distribution for the background. Rotational averaging reduced the 2D parameter estimation problem to a 1D one which improved computational efficiency of estimation. The drawback of this approach was the assumption of no astigmatism.

Tani et al. [15] pointed out that the rotationally averaged 1D estimate of the power spectrum when re-sampled as a function of the square of the frequency (which they referred to as the "q^2 plot") is periodic, assuming that the value of the spherical aberration (C_s) is negligible. They filtered the Fourier transform of the "q^2 plot" to reduce noise. Astigmatism was calculated by dividing the image into small sectors and rotational averaging was done inside each of the sectors. In essence, the elliptical Thon rings were approximated by small circular arcs. As pointed out by Tani et al. [15], averaging should ideally be done on points in the power spectrum with the same CTF value. Our approach of elliptical averaging achieves this goal.

In the last few years some new approaches were introduced. Fernández et al. [16] used autoregressive (AR) modeling for estimating the power spectrum, while the background noise was modeled as an exponential of a polynomial of the frequency. Following the same line of work, Valázquez-Muriel et al. [17] used autoregressive moving averages (ARMA) to model the power spectrum. This series of work was significant because for the first time, effort was put into the estimation of the power spectrum before estimating the CTF parameters. However, their method of parameter adjustment is essentially an exhaustive search for parameters and does not provide any formal guarantees of convergence. Hence the program needs to run several times to obtain the actual solution. Secondly, while adjusting the defocus parameter, astigmatism was assumed to be absent. Under such an assumption, rotational averaging would give much better estimates of the defocus.

Sander et al. [18] used multivariate statistical analysis to group power spectra having similar CTF parameters and used class averages to get an estimate of the power spectrum. They used an iterative scheme to determine the parameters of the CTF. This iterative scheme is in essence an exhaustive search for parameters which minimize the correlation between the theoretical and actual CTF. The exhaustive search is performed over a user-defined region in the parameter space. It becomes computationally intensive if the user chooses a large region. On the other hand, it

becomes inaccurate if the user chooses a small region. Even this exhaustive search does not guarantee convergence to the global correlation maxima because for each step only one parameter is being varied.

Mindell et al. [19] used a smoothed version of the 2D estimate of the power spectrum as an estimate of the background. They sought to maximize the cross-correlation between a theoretical CTF^2 and the background subtracted 2D estimate of the power spectrum by doing a exhaustive search for the two defoci[1] defining the CTF and the angle of astigmatism.

Huang et al. [20] reduced the problem of background and envelope estimation to a constrained optimization problem. They solved the constrained optimization problem using the simplex algorithm of linear programming. They obtained an estimate of the CTF^2 by compensating the 1D power spectrum for the background and envelope functions. A lower and higher cutoff frequency to exclude some parts of the power spectrum were also defined. Astigmatism was calculated by dividing the image into small sectors (as many as 60) and rotationally averaging in each sector as described by Tani et al. [15]. The method proposed by Huang et al. [20] is both elegant and mathematically convincing and we have used the idea of constrained optimization in our own implementation.

In the various methods proposed till date astigmatism is either ignored or the elliptical Thon rings are approximated by circular arcs. The main contribution of this paper, which distinguishes it from all previous work, is an elegant method to calculate the parameters of astigmatism using edge detection, followed by elliptical averaging to boost signal to noise ratio.

3. Theory

In this section we briefly describe the image formation equation based on the contrast transfer theory [8,9,21]. Under the linear model of contrast

---

[1]It is explained in the next section that in the presence of astigmatism the CTF is defined by two defoci and not a single defocus.

transfer in an electron microscope the following assumptions are made:

- The CTF (c) and the envelope function (e) are both spatially invariant.
- The noise (n) is independent and additive.

With the above assumptions the image formation equation can be written as

$$i(x, y) = c(x, y) \otimes e(x, y) \otimes f(x, y) + n(x, y), \qquad (1)$$

where (x, y) are the spatial domain variables. i, c, e, and n represent the image, the CTF, the envelope function, and the noise in spatial domain respectively. f is the projection of the particle being imaged. The Fourier transform of the projection is called the structure factor. $\otimes$ refers to the convolution operator.

Taking the Fourier transform on both sides of Eq. (1), we get the image formation equation in the frequency domain.

$$I(s_x, s_y) = \mathscr{C}(s_x, s_y)\mathscr{E}(s_x, s_y)\mathscr{F}(s_x, s_y) + \mathscr{N}(s_x, s_y), \qquad (2)$$

where (s_x, s_y) denote the frequency domain variables, while I, $\mathscr{C}$, $\mathscr{E}$, $\mathscr{F}$ and $\mathscr{N}$ denote the image, the CTF, the envelope function, the structure factor and the noise in frequency domain respectively.

In polar coordinates, Eq. (2) can be rewritten as

$$I(s, \theta) = \mathscr{C}(s, \theta)\mathscr{E}(s, \theta)\mathscr{F}(s, \theta) + \mathscr{N}(s, \theta), \qquad (3)$$

where $s = \sqrt{s_x^2 + s_y^2}$ and $\theta = \arctan(s_y/s_x)$.

In the weak phase and weak amplitude approximation, the following parametric form for the CTF was proposed by Wade [7]:

$$\mathscr{C}(s, \theta) = \sqrt{1 - C_a^2}\sin(\gamma(s, \theta)) + C_a\cos(\gamma(s, \theta)), \qquad (4)$$

where C_a is the amplitude contrast. The argument γ is given by

$$\gamma(s, \theta) = 2\pi\left(\frac{-C_s\lambda^3 s^4}{4} + \frac{z(\theta)\lambda s^2}{2}\right), \qquad (5)$$

where λ is the wavelength of electrons in the microscope, C_s is the spherical aberration of the lens, and z is the defocus.

In the presence of astigmatism, the defocus is dependent on θ and is governed by the following equation proposed by Henderson et al. [12]:

$$z(\theta) = \frac{z_2 + z_1}{2} + \frac{z_2 - z_1}{2}\sin(2(\theta - \phi)), \qquad (6)$$

where z_1 and z_2 are the minimum and maximum defoci respectively. ϕ is the angle made by the major axis of the elliptical Thon rings with the x-axis.

The envelope function is a result of several phenomena such as the coherence of the electron beam, the lens current instability, specimen drift, etc. Models for the envelope function have been proposed to account for the above phenomena [13,22–27]. An approximation of the envelope function, called the "B-factor" parametric form [26,27], is given by

$$\mathscr{E}(s) = e^{-Bs^2}. \qquad (7)$$

In recent published work [20], the "B-factor" parametric form was found to be inadequate in describing the envelope function. Our experiments indicate that a few additional parameters, as well as the "B-factor", reasonably describe the envelope function. Based on our experiments, we introduce an empirical model for the envelope function, which has a constant, a linear and a sublinear term in the exponent in addition to the familiar quadratic term in the "B-factor" representation:

$$\mathscr{E}(s) = e^{-(k_1 + k_2\sqrt{s} + k_3 s + k_4 s^2)}. \qquad (8)$$

One of the empirical parametric forms for the noise spectrum which encompasses a wide range of different physical effects, including incoherent scattering, film noise and scanner noise [28], is given by

$$\mathscr{N}^2(s) = e^{-(n_1 + n_2\sqrt{s} + n_3 s + n_4 s^2)}. \qquad (9)$$

Our experiments show that the noise spectrum in a large number of test images fit the above noise model very accurately even for images recorded using a CCD camera. Therefore, we chose to use the above noise model.

4. The algorithm

As described above, there are 12 parameters which describe the power spectrum of an image taken using a TEM; the two defoci (z_1 and z_2), the amplitude contrast C_a, the astigmatism angle ϕ, the four parameters of the envelope function (k_1, k_2, k_3 and k_4) and the four parameters of the noise function (n_1, n_2, n_3 and n_4). Before we proceed to estimate these parameters we do a re-parameterization. If we have the knowledge of one of the defoci, say z_2, we can estimate the other defocus, say z_1, based on the knowledge of the ratio r of the major and minor axes of the elliptical Thon rings. Hence, we can replace parameter z_1 with a new parameter r. The new set of parameters are r, ϕ, z_2, C_a, k_1, k_2, k_3, k_4, n_1, n_2, n_3 and n_4.

It is not practically feasible to set up this parameter estimation problem as a single step optimization because of the large number of parameters. However, we can estimate certain parameters without any knowledge of the others. We propose a sequential algorithm in which a few parameters are calculated at each step. The sequence of steps are as follows:

(1) *Determination of astigmatism parameters* (r, ϕ) *and elliptical averaging*. The benefit of the new parameterization is that we can estimate the parameters r and ϕ without any knowledge of the other parameters. Based on the estimated r and ϕ, elliptical averaging of the power spectrum then averages all points at same CTF value.

(2) *Determination of a lower cutoff frequency* (s_l). A lower cutoff frequency is calculated to remove a region of the power spectrum which is dominated by the structure factor.

(3) *Determination of noise parameters* (n_1, n_2, n_3 and n_4). Having estimated r and ϕ, the noise parameters are estimated independent of all remaining parameters.

(4) *Determination of upper cutoff frequency* (s_u). The upper cutoff frequency is determined based on the energy contained in the noise subtracted power spectrum.

(5) *Determination of envelope parameters* (k_1, k_2, k_3 and k_4). The envelope function parameters can then be calculated independent of the defocus (z_2) and amplitude contrast (C_a).

(6) *Determination of CTF parameters* (z_1, z_2 and C_a). Finally z_2 and C_a are calculated. Based on the values of z_2 and r, z_1 can be calculated.

(7) *Noise and envelope parameter refinement*. The noise and envelope functions, unlike the CTF, are circularly symmetric. In steps 3 and 5, the parameters are obtained using elliptically averaged envelope and noise functions which introduce errors in their estimates. These parameters are refined using a circularly averaged power spectrum.

Below we describe each of these steps in detail.

4.1. Determination of astigmatism parameters and elliptical averaging

Astigmatism is an imaging artifact caused by a variety of reasons like inhomogeneity of the pole piece material, limited precision of machining, charging of the objective aperture and misalignment of the apertures. The presence of astigmatism can be seen in different ways. In the image it appears as the streaking in the graininess of the support film. A hole in the support film of an astigmatic image appears to have both light and dark fringes at its boundary. The effect is most easily observed in the power spectrum of the image where the Thon rings are no more circular. Fig. 1 illustrates power spectra with increasing levels of astigmatism. In an image with little or no astigmatism, the Thon rings are circular. Under the assumption of negligible spherical aberration, the shape of the Thon rings changes from circular to elliptic to parabolic and finally to hyperbolic with increasing astigmatism.

From a purely mathematical standpoint, if the spherical aberration is neglected, the rings (level sets of the power spectrum) can be any conic section depending on the values of defoci z_1 and z_2.

$z_1 = z_2 \Rightarrow$ the conic is a circle; see Fig.1(a),

$\frac{z_1}{z_2} > 0 \Rightarrow$ the conic is an ellipse; see Fig.1(b–d),

$z_1 = 0$ or $z_2 = 0$

$\quad \Rightarrow$ the conic is a parabola; see Fig.1(e),

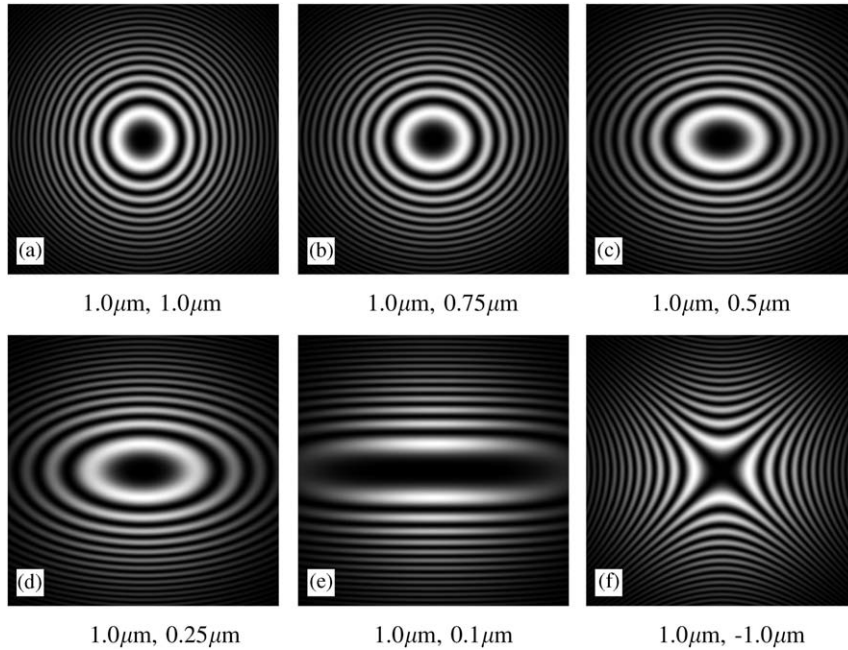$\frac{z_1}{z_2} < 0 \Rightarrow$ the conic is a hyperbola; see Fig.1(f).

Fig. 1. Simulated power spectra with increasing astigmatism are shown. The caption shows the two defocus values corresponding to the power spectrum. The Thon rings can distort to an ellipse, a parabola or a hyperbola.

We present a method for calculating astigmatism when the Thon rings are elliptical. The images with parabolic or hyperbolic Thon rings are never used in practice and should be automatically rejected by the algorithm. Automatic rejection of such images is an important feature required for complete automation of CTF estimation.

In the absence of astigmatism, the Thon rings are circular. Hence, rotational averaging of the power spectrum can be done, so that points with the same CTF value are averaged. Rotational averaging gives a 1D power spectrum which has a higher signal to noise ratio than the 2D estimate of the power spectrum. This in turn leads to better estimation of the noise, envelope and CTF parameters.

In the presence of astigmatism, the Thon rings are elliptical. If the parameters of the family of the concentric ellipses could be estimated, averaging along ellipses could improve signal to noise ratio. In previous approaches [13–15,20] elliptical averaging was approximated by dividing the power spectrum into several sectors, doing rotational

averaging in the small sectors and estimating the defocus separately for each sector. This procedure amounts to approximating the elliptical ring with a number of circular arcs. If the number of sectors is chosen to be large, the approximation should be good. However, the number of points which are averaged decreases and this leads to inferior estimates of defocus and other CTF parameters. On the other hand, if the sectors are large, much of the elliptical boundary is approximated using a single circular arc which leads to a bad approximation. This was the primary criticism of using a 1D power spectrum instead of doing parameter estimation on the 2D power spectrum. In our algorithm we estimate the parameters of the family of ellipses first and then take an average along the elliptical boundary to get a 1D power spectrum.

The two parameters which define the family of elliptical Thon rings are

(1) r: the ratio of the major and minor axes of the Thon rings;

(2) ϕ: the angle which the major axis makes with the x-axis.

The design of the algorithm is motivated by the following observation. Near the origin, the Thon rings are most prominent and the spacing between them is large. Far from the origin, the Thon rings start fading away and the spacing between them decreases drastically. Higher frequency rings can be blurred out by convolving the power spectrum estimate with a 2D Gaussian filter of large width. The most prominent gradient in the power spectrum is between the points where the power spectrum falls from a very high value to its first minimum. Edge detection with a high threshold value successfully recovers a single Thon ring. Edges are points that are local maxima of the magnitude of the gradient (above a given threshold) along the direction of the gradient. A higher value of threshold leads to a lesser number of edge pixels. In practice, we use the Canny edge detector [29] to detect the edges with a suitable threshold. All edges very close to the center are removed because if the defocus is very large, the first edge ring is very small and therefore not reliable for estimating r and ϕ. The result of applying the edge detection algorithm on the power spectrum of a typical micrograph is shown in Fig. 2(a). The location of the edge is different from the location of the first dark ring which corresponds to a local minimum of the power spectrum.

Following the edge detection we fit an ellipse to the detected edges. Any conic section can be represented by the following parametric form:

$$ax^2 + bxy + cy^2 + dx + ey = 1. \qquad (10)$$

The conic given by Eq. (10) is a hyperbola, parabola or an ellipse depending on the following conditions:

$b^2 - 4ac > 0 \implies$ the conic is a hyperbola,

$b^2 - 4ac = 0 \implies$ the conic is a parabola,

$b^2 - 4ac < 0 \implies$ the conic is an ellipse or a circle.

One of the advantages of using this general conic parameterization is that we can detect images with

parabolic and hyperbolic Thon rings and reject them. Additionally, if the conic exhibits reflection symmetry about the origin (which means that if a point (x_1, y_1) satisfies Eq. (10), then point $(-x_1, -y_1)$ also satisfies Eq. (10)), then the equation of the conic reduces to

$$ax^2 + bxy + cy^2 = 1. \qquad (11)$$

A simple translation of the detected edge coordinates shifts the center of the ellipse to the origin which allows us to use Eq. (11) as the equation of our ellipse. Also note that Eq. (11) is linear in parameters a, b and c, and so it is natural to consider linear least-squares estimation. However, a linear estimator which minimizes the least-square error might not be the appropriate choice for estimating the parameters because spurious edges (outliers) may be detected if the threshold of the Canny edge detector is not chosen properly; see Fig. 3(a). One approach would be to filter out these spurious edges based on connectivity and length of the edges. Our experiments showed that such an approach was not robust enough. To deal with outliers we use the robust estimation technique of Random Sample Consensus (RANSAC) [30].

RANSAC, unlike least-squares estimate, is robust to outliers. See Fig. 3 for an illustration. A brief description of the RANSAC algorithm follows. Assume that the minimum number of data points needed to estimate a parameter vector $\overrightarrow{p}$ is N and there are M data points in all. RANSAC has the following steps:

(1) Randomly select N data points out of M data points.
(2) Estimate the parameter vector $\overrightarrow{p}$ using the N selected points. The estimation procedure for this step can be any non-robust method.
(3) Find how many data points (of M) fit the model with parameter vector $\overrightarrow{p}$ within a user-defined tolerance. Call this K.
(4) If K is big enough, accept fit and exit with success.
(5) Repeat the above steps for a user-defined number of times.
(6) Exit with failure.
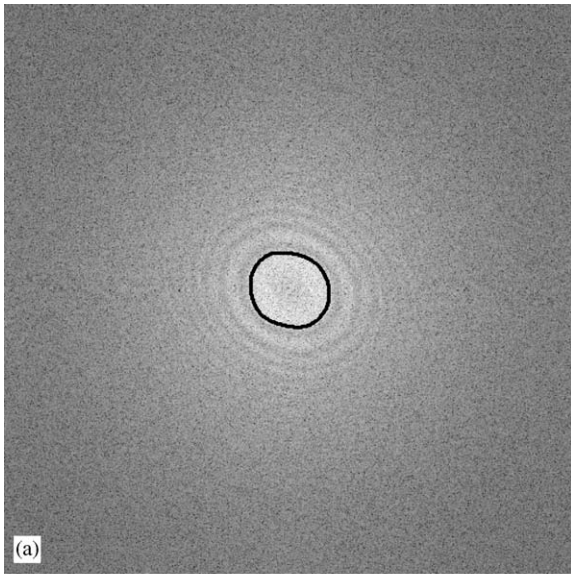
Edge detection
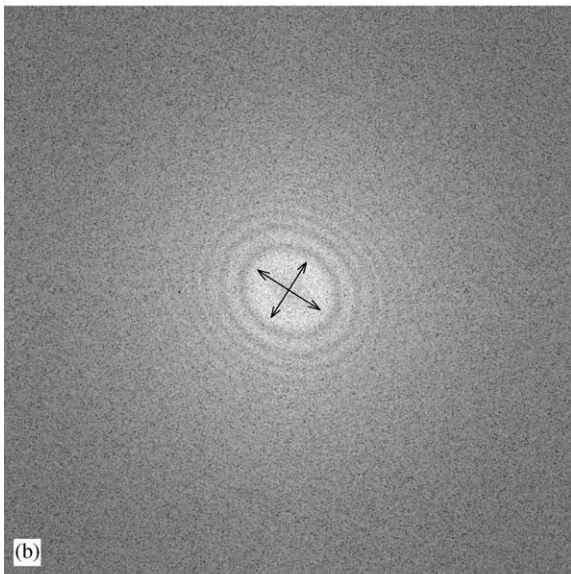
RANSAC estimate

Fig. 2. A typical result of edge detection and RANSAC estimation is shown. In (b) the two double headed arrows represent the estimated length and orientation of the major and minor axes.

For step 2 of RANSAC we use the linear least-squares estimate. The parameters (r, ϕ) can be calculated from the ellipse parameters a, b

and c by

$$l_1 = \frac{a+c}{2} - \frac{1}{2}\sqrt{(a-c)^2 + b^2},$$

$$l_2 = \frac{a+c}{2} + \frac{1}{2}\sqrt{(a-c)^2 + b^2},$$

$$\phi = \arctan\left(\frac{2(l_2 - a)}{b}\right), \tag{12}$$

$$r = \sqrt{\frac{l_1}{l_2}}. \tag{13}$$

Consider a power spectrum with elliptical Thon rings parameterized by r and ϕ. All the points along an ellipse have the same CTF value. Hence we can average all points along an ellipse to generate a 1D power spectrum. The elliptically averaged power spectrum is given by

$$I^2(s) = \int_0^{2\pi} I^2(s_x(\theta), s_y(\theta))\, d\theta$$
$$= \int_0^{2\pi} I^2(rs\cos(\theta - \phi), s\sin(\theta - \phi))\, d\theta. \tag{14}$$

Elliptical averaging is explained in Fig. 4. The slice of the power spectrum along the major axis is replaced by the averaged value along the ellipse shown. Hence, the 1D power spectrum is actually a slice along the major axes of the elliptical Thon rings. The defocus calculated using this 1D power spectrum would correspond to the smaller of the two defoci (underfocus). The other defocus can be calculated by knowing the ratio between the major and minor axes of the elliptical Thon rings. Notice that we could have chosen the slice of the power spectrum along the minor axis of the elliptical Thon rings and the defocus would correspond to the larger of the two defoci (underfocus).

4.2. Determination of the lower cutoff frequency

At low frequency, the structure factor dominates the power spectrum. This region of the power spectrum adversely affects the estimation accuracy of the CTF parameters. To filter out the effects of the structure factor, a lower cutoff frequency needs to be calculated.
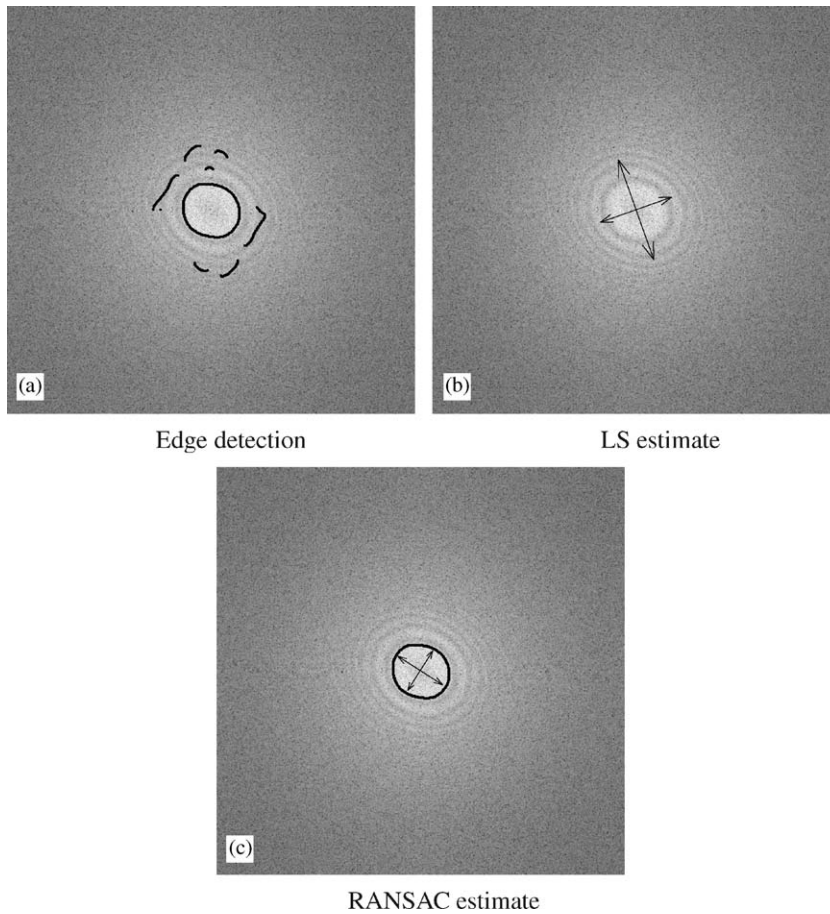
Fig. 3. In (a) the result of a noisy edge detection is shown. In (b) and (c) the two double headed arrows represent the estimated length and orientation of the major and minor axes using least-squares fit and RANSAC respectively. In (c) the points represent the inliers picked by RANSAC.

The power spectrum below this cutoff frequency is not used in the estimation of the CTF parameters.

As mentioned earlier, the edge ring corresponds to the region of maximum gradient in the power spectrum. The first dark Thon ring contains this edge ring. The region inside the first edge ring is dominated by the structure factor. The location (frequency) of the local maxima of the 1D power spectrum which is closest to the edge location is used as the lower cutoff frequency. Fig. 5 shows a 1D power spectrum. The first dotted line represents the lower cutoff frequency and we denote it by s_l.

4.3. Determination of the noise parameters

The estimate of the power spectrum is elliptically averaged to get a 1D power spectrum. Using the cutoff frequency calculated in the previous section, the 1D power spectrum corresponding to the lower frequency is removed. We denote this new 1D power spectrum by $\mathscr{P}(s)$. Hence, using Eq. (3), assuming that the data is uncorrelated with background noise and that elliptical averaging has been completed, we get

$$\mathscr{P}(s) = \mathscr{C}^2(s)\mathscr{E}^2(s)\mathscr{F}^2(s) + \mathscr{N}^2(s). \tag{15}$$

The above equation can be discretized such that $\mathscr{P}(\mathbf{s})$ is a vector containing samples of $\mathscr{P}(s)$ at
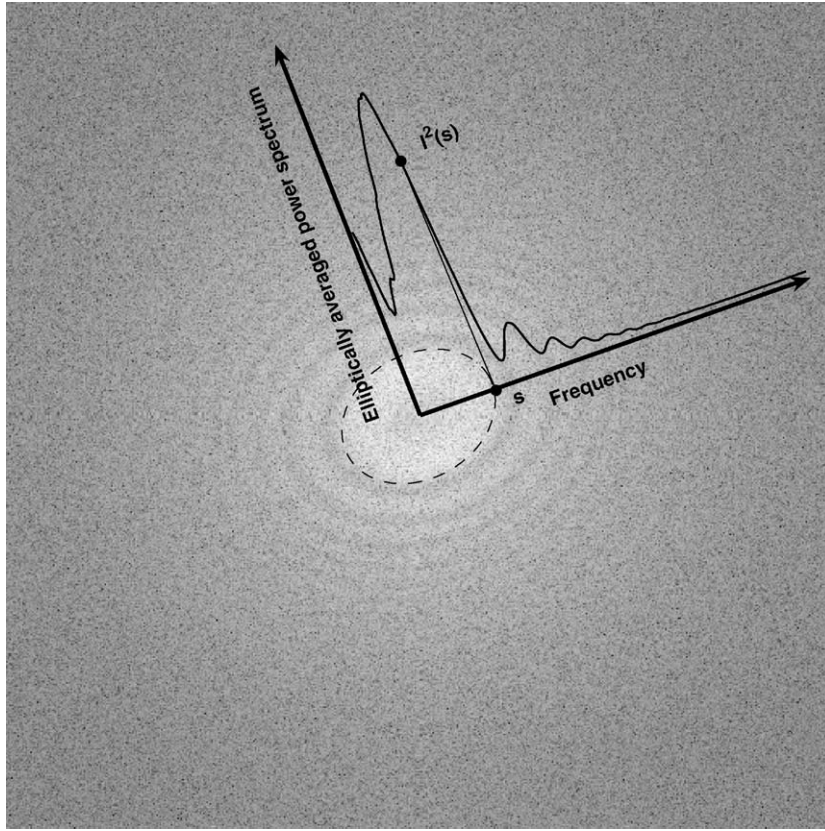
Fig. 4. Elliptical averaging. The 1D power spectrum is a slice of the 2D power spectrum along the direction of major axis (we could have used the 1D power spectrum along the minor axis as well). The signal to noise ratio of the 1D power spectrum is boosted using elliptical averaging. For a particular frequency s, the 1D power spectrum $I^2(s)$ is calculated by averaging the points in the 2D power spectrum having the same CTF value (which is shown using the ellipse).

discrete frequencies $\mathbf{s} = [s_1 s_2, \ldots, s_T]^\mathrm{T}$. A set of new notations are introduced to handle the equations in the discrete domain.

- *Point-wise multiplication* (.∗): Given two vectors $\mathbf{v} = [v_1 v_2 \ldots v_k]^\mathrm{T}$ and $\mathbf{u} = [u_1 u_2 \ldots u_k]^\mathrm{T}$

  $\mathbf{v}. \ast \mathbf{u} = [v_1 u_1 \; v_2 u_2 \; \cdots \; v_k u_k]^\mathrm{T}.$

- *Point-wise exponent operator* (.^{n}): Given a vector $\mathbf{v} = [v_1 v_2 \ldots v_k]^\mathrm{T}$,
  $\mathbf{v}.^n = [v_1^n \; v_2^n \; \cdots \; v_k^n]^\mathrm{T}.$
  Given a function f operating on a vector $\mathbf{v}$,
  $f.^n(\mathbf{v}) = [f^n(v_1) \; f^n(v_2) \; \cdots \; f^n(v_k)]^\mathrm{T}.$

All subsequent analysis will be in the discrete domain.

In the frequency range under consideration, if the frequency response due to structure is assumed to be white, then Eq. (15) reduces to

$$\mathscr{P}(\mathbf{s}) = \mathscr{C}.^2(\mathbf{s}). \ast \mathscr{E}.^2(\mathbf{s}) + \mathscr{N}.^2(\mathbf{s}). \qquad (16)$$

Let $\hat{\mathscr{N}}.^2(\mathbf{s})$ denote an estimate of $\mathscr{N}.^2(\mathbf{s})$ and with a parametric form given by Eq. (9). Under the assumption that the noise spectrum $\mathscr{N}.^2(\mathbf{s})$ changes slowly as compared to the function $\mathscr{C}.^2(\mathbf{s})$, the local minima of $\mathscr{P}(\mathbf{s})$ correspond (approximately) to the zero crossings of the CTF. At a zero crossing of the CTF, the function $\mathscr{P}(\mathbf{s})$ has contribution from the noise spectrum only. To get an estimate of the noise spectrum $\mathscr{N}.^2(\mathbf{s})$, we fit a curve $\hat{\mathscr{N}}.^2(\mathbf{s})$ to $\mathscr{P}(\mathbf{s})$ such that $\hat{\mathscr{N}}.^2(\mathbf{s})$ is less than $\mathscr{P}(\mathbf{s})$ at each frequency.
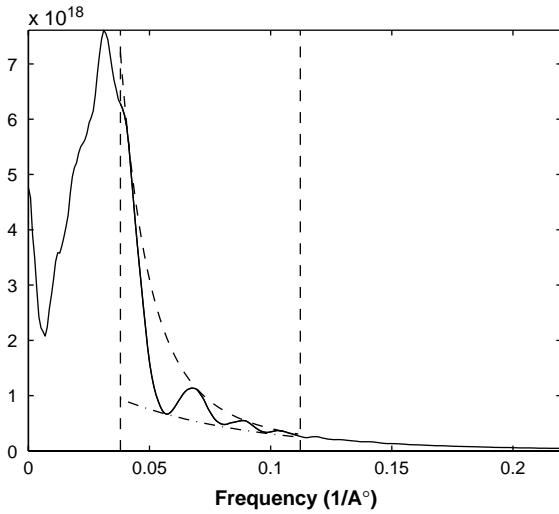
Fig. 5. Determination of lower and upper cutoff frequencies and the fitting of noise spectrum and envelope function. The solid curve represents the elliptically averaged 1D power spectrum of the carbon support film. The lower cutoff frequency and the higher cutoff frequency are shown using dashed vertical lines. The sum of noise spectrum and square of the envelope function is shown by the dashed curve above the power spectrum. The noise spectrum is shown by a dot-dash curve below the power spectrum. The 1D power spectrum shown above is a typical power spectrum of carbon support film.

Formally, to calculate $\hat{\mathcal{N}}.^2(\mathbf{s})$ we minimize the objective function given by

$$O_N(n_1, n_2, n_3, n_4)$$
$$= \| \log(\mathscr{P}(\mathbf{s})) - n_1 - n_2\mathbf{s}.^{0.5} - n_3\mathbf{s} - n_4\mathbf{s}.^2 \|^2 \quad (17)$$

under the constraints

$$\log(\mathscr{P}(\mathbf{s})) - n_1 - n_2\mathbf{s}.^{0.5} - n_3\mathbf{s} - n_4\mathbf{s}.^2 \geqslant 0. \quad (18)$$

In vector notation we can restate the above problem as

$$\hat{\mathbf{n}} = \arg \min_{\mathbf{n}} \|\mathbf{An} - \mathbf{b}\|^2 \quad (19)$$

under the constraint

$$\mathbf{An} - \mathbf{b} \geqslant 0, \quad (20)$$

where

$$\mathbf{A} = [\mathbf{1}_{T \times 1} \quad \mathbf{s}.^{0.5} \quad \mathbf{s} \quad \mathbf{s}.^2], \quad \mathbf{b} = \log(\mathscr{P}(\mathbf{s})),$$
$$\mathbf{n} = [n_1 \quad n_2 \quad n_3 \quad n_4]^T$$

($\mathbf{1}_{T \times 1}$ is a column vector of length T with all elements 1).

This constrained least-squares problem reduces to the form of quadratic programming given by

$$\hat{\mathbf{n}} = \arg \min_{\mathbf{n}}(\mathbf{n}^T\mathbf{Bn} + \mathbf{c}^T\mathbf{n}), \quad (21)$$

where $\mathbf{B} = \mathbf{A}^T\mathbf{A}$ and $\mathbf{c} = -2\mathbf{A}^T\mathbf{b}$ under the constraint

$$\mathbf{An} - \mathbf{b} \geqslant 0. \quad (22)$$

The convexity of a quadratic programming problem depends on whether or not the matrix $\mathbf{B}$ is positive semi-definite. The matrix $\mathbf{B}$ is positive semi-definite in the case of a constrained least-squares problem, and so the quadratic programming problem is convex. Hence, it is guaranteed to converge to a global minimum.

The constrained linear least-squares problem is solved using an algorithm by Coleman and Li [31].

4.4. Determination of upper cutoff frequency

The upper cutoff frequency is determined based on the energy contained in the signal. The energy contained in the power spectrum after removing the contribution from the background noise and removing the region dominated by the structure factor is given by

$$E = \sum_{s=s_l}^{s_T}[I.^2(s) - \mathcal{N}.^2(s)], \quad (23)$$

where s_l is the lower cutoff frequency. We define an upper cutoff frequency (s_u), so that 95% of the energy E is contained between s_l and s_u. Formally,

$$\sum_{s=s_l}^{s_u}[I.^2(s) - \mathcal{N}.^2(s)] = 0.95 \sum_{s=s_l}^{s_T}[I.^2(s) - \mathcal{N}.^2(s)].$$
$$(24)$$

4.5. Determination of envelope function parameters

Several phenomena such as the coherence of the electron beam, the lens current instability and specimen drift lead to an exponential decrease in signal strength with frequency. The effects of all these phenomena are modeled using a single function called the envelope function. As discussed

earlier, we use the following empirical form for the envelope function:

$$\hat{\mathscr{E}}.^2(\mathbf{s}) = e^{-(k_1+k_2\mathbf{s}.^{0.5}+k_3\mathbf{s}+k_4\mathbf{s}.^2)}, \tag{25}$$

where $\hat{\mathscr{E}}.^2(\mathbf{s})$ represents an estimate of the envelope function. Let $\mathscr{M}(\mathbf{s}) = \mathscr{P}(\mathbf{s}) - \mathscr{N}.^2(\mathbf{s})$ denote the background subtracted power spectrum. The same approach that was used to find the parameters of $\mathscr{N}.^2(\mathbf{s})$ can be used to find the parameters of $\hat{\mathscr{E}}.^2(\mathbf{s})$ with a minor modification. We fit $\hat{\mathscr{E}}.^2(\mathbf{s})$ to $\mathscr{M}(\mathbf{s})$ such that at each frequency s, $\hat{\mathscr{E}}.^2(\mathbf{s})$ is greater than $\mathscr{M}(\mathbf{s})$. Hence, the objective function is given by

$$O_{\mathscr{E}}(k_1, k_2, k_3, k_4) = \| \log(\mathscr{M}(\mathbf{s})) - k_1 - k_2\mathbf{s}.^{0.5} \\ - k_3\mathbf{s} - k_4\mathbf{s}.^2 \|^2 \tag{26}$$

under the constraint

$$\log(\mathscr{M}(\mathbf{s})) - k_1 - k_2\mathbf{s}.^{0.5} - k_3\mathbf{s} - k_4\mathbf{s}.^2 \leqslant 0. \tag{27}$$

The parameters k_1, k_2, k_3 and k_4 can be calculated in the same way as the parameters n_1, n_2, n_3 and n_4.

4.6. Determination of the CTF parameters

In this section we discuss the method used to calculate the parameters of the CTF, namely the defocus z_2 and the amplitude contrast C_a. An estimate of the CTF is given by

$$\hat{\mathscr{C}}.^2(\mathbf{s}) = \frac{\mathscr{P}(\mathbf{s}) - \hat{\mathscr{N}}.^2(\mathbf{s})}{\hat{\mathscr{E}}.^2(\mathbf{s})}. \tag{28}$$

A straightforward method would be to calculate the parameters using any constrained non-linear optimization method which minimizes the objective function

$$O_z = \| \hat{\mathscr{C}}.^2(\mathbf{s}) - \mathscr{C}.^2(\mathbf{s}, z_2, C_a) \|^2 \tag{29}$$

under the constraint

$$z_2 < 0 \text{ and } 0 \leqslant C_a \leqslant 0.2, \tag{30}$$

where $\mathscr{C}(\mathbf{s}, z_2, C_a)$ is the theoretical CTF obtained by using the elliptically averaged version of Eq. (4). This problem has enough local minima to prevent any guarantee of convergence to the actual solution. However, if we provide the algorithm with an initial guess which is sufficiently close to the solution, we can hope for convergence. A reasonable initial guess for the amplitude

contrast is 0. We use a simple but robust technique to calculate an initial estimate of the defocus as an initial estimate for the constrained non-linear optimization algorithm. $\hat{\mathscr{C}}.^2(s)$ is smoothed using a moving average low pass filter and then the derivative is taken to find local minima of $\hat{\mathscr{C}}.^2(s)$. Let, $m_1 \ldots m_r$ be the r local minima. In an ideal situation, m_i would correspond to the ith zero crossing of the CTF. However, we need to consider the possibility that some of the r minima could be spurious, and some of the actual minima could have been missed. We therefore define a set of defoci values z_{ij} given by

$$z_{ij} = \frac{2j + C_s\lambda^3 m_i^4}{2\lambda m_i^2}, \tag{31}$$

where z_{ij} is the defocus obtained by assuming that the ith minimum corresponds to the jth zero crossing of the CTF. Let $\mathscr{C}(s, z_{ij})$ be the theoretical CTF obtained with the defocus values z_{ij} using Eq. (4). The initial value of defocus (say z_{init}) is obtained by

$$z_{init} = \arg\min_{z_{ij}} \| \hat{\mathscr{C}}.^2(s) - \mathscr{C}.^2(s, z_{ij}) \|^2. \tag{32}$$

The z_{init} so obtained is close to the true defocus. However for greater accuracy, we use this value as an initial condition for a constrained non-linear optimization algorithm [32] which minimizes the objective function given by Eq. (29) and under the constraint given by inequality (30).

4.7. Noise and envelope parameter refinement

In Sections 4.3 and 4.5, the noise and envelope function parameters were calculated using an elliptically averaged power spectrum. It must be noted that the Thon rings are elliptical because of the CTF. The noise and the envelope functions are circularly symmetric. This fact is also evident from Eqs. (8) and (9) in which the envelope function and the noise spectrum are independent of θ. Elliptical averaging of circularly symmetric noise and envelope functions introduces errors in the estimated parameters.

An important observation is that as long as the parametric models suggested in Eqs. (8)

and (9) fit the elliptically averaged noise and envelope functions respectively, no errors are introduced in the estimation of CTF parameters like defocus and amplitude contrast.

To obtain a more accurate estimate of the noise and envelope functions a refinement step is introduced. The power spectrum is circularly averaged. Since all the parameters of the CTF are known, the circularly averaged power spectrum can be modeled as

$$\mathscr{P}_c(\mathbf{s}) = \mathscr{E}_c.^2(\mathbf{s}). * \mathscr{C}_c.^2(\mathbf{s}) + \mathscr{N}_c.^2(\mathbf{s}). \tag{33}$$

The subscript c is used to denote circular averaging. $\mathscr{C}_c.^2(\mathbf{s})$ represents the circularly averaged CTF^2 and is accurately known from previous sections. Approximate values of $\mathscr{E}_c.^2(\mathbf{s})$ and $\mathscr{N}_c.^2(\mathbf{s})$ are also known but the parameters need to be refined to account for the fact that the noise and envelope functions are circularly symmetric.

We set up a unconstrained non-linear optimization problem which minimizes the objective function given by

$$\sum_{s_l}^{s_u} [1 - (\mathscr{E}_c.^2(\mathbf{s}). * \mathscr{C}_c.^2(\mathbf{s}) + \mathscr{N}_c.^2(\mathbf{s})). / \mathscr{P}_c(\mathbf{s})]^2$$
$$\tag{34}$$

with the noise and envelope function parameters as variables. The noise and envelope function parameters calculated in Sections 4.3 and 4.5 are supplied as initial value to the non-linear optimizer. The above cost function is the normalized squared error between the actual power spectrum and the modeled power spectrum between the lower and upper cutoff frequencies. The normalization is required so that the fitting is not biased by large values of the power spectrum at certain frequencies at the cost of fitting poorly at frequencies where the value of the power spectrum is lower.

5. Results

The algorithm described above was implemented in MATLAB. The program is called Automated CTF Estimation (ACE) and requires the MATLAB image processing and optimization toolboxes. The user can choose to turn astigmatism estimation on or off. In the case when astigmatism estimation is turned off, circular rather than elliptical averaging is performed. The run time varies with the size of micrograph. For a 512×512 image, the run time of the algorithm is typically 15 s, while for a 4096×4096 image the run time is typically 1 min and 30 s on a 2 GHz Pentium 4 processor.

First, as a proof of concept, we show detailed results for each step of the algorithm using carbon support films. We next show the accuracy of the algorithm on a set of images of carbon support films in which defocus has been systematically varied from 0.6 μm under-focus to 6 μm under-focus in steps of 0.2 μm. This experiment was designed to prove the accuracy of the algorithm up to a constant bias. In both the above experiments the imaging conditions were: voltage = 200 kV, pixel size = 2.26 Å/pixel and spherical abberation = 2 mm. Finally we show the practical utility of the algorithm by testing it on images of particles embedded in ice and suspended over holes in the carbon film. The imaging conditions were: voltage = 120 kV, dose = 10.84 e⁻/Å², pixel size = 2.26 Å/pixel and spherical aberration = 2 mm.

5.1. Carbon support

The Thon rings are far more prominent in images of carbon support films than in images containing only protein specimens embedded in vitreous ice [7]. Thus, as a proof of concept, we first show detailed results of the performance of the algorithm on images of carbon support films. We start with a demonstration of the ellipse fitting algorithm. A typical result of the edge detection algorithm is shown in Fig. 2(a). To demonstrate the robustness of the algorithm toward outliers, we changed the parameters of the Canny edge detector to force spurious edges to be detected (Fig. 3(a)). A comparison of the least-squares estimate and a RANSAC estimate, shown in Fig. 3(b) and (c), illustrates that whereas the least-squares estimate is incorrect, RANSAC is able to reject the outliers to give the correct
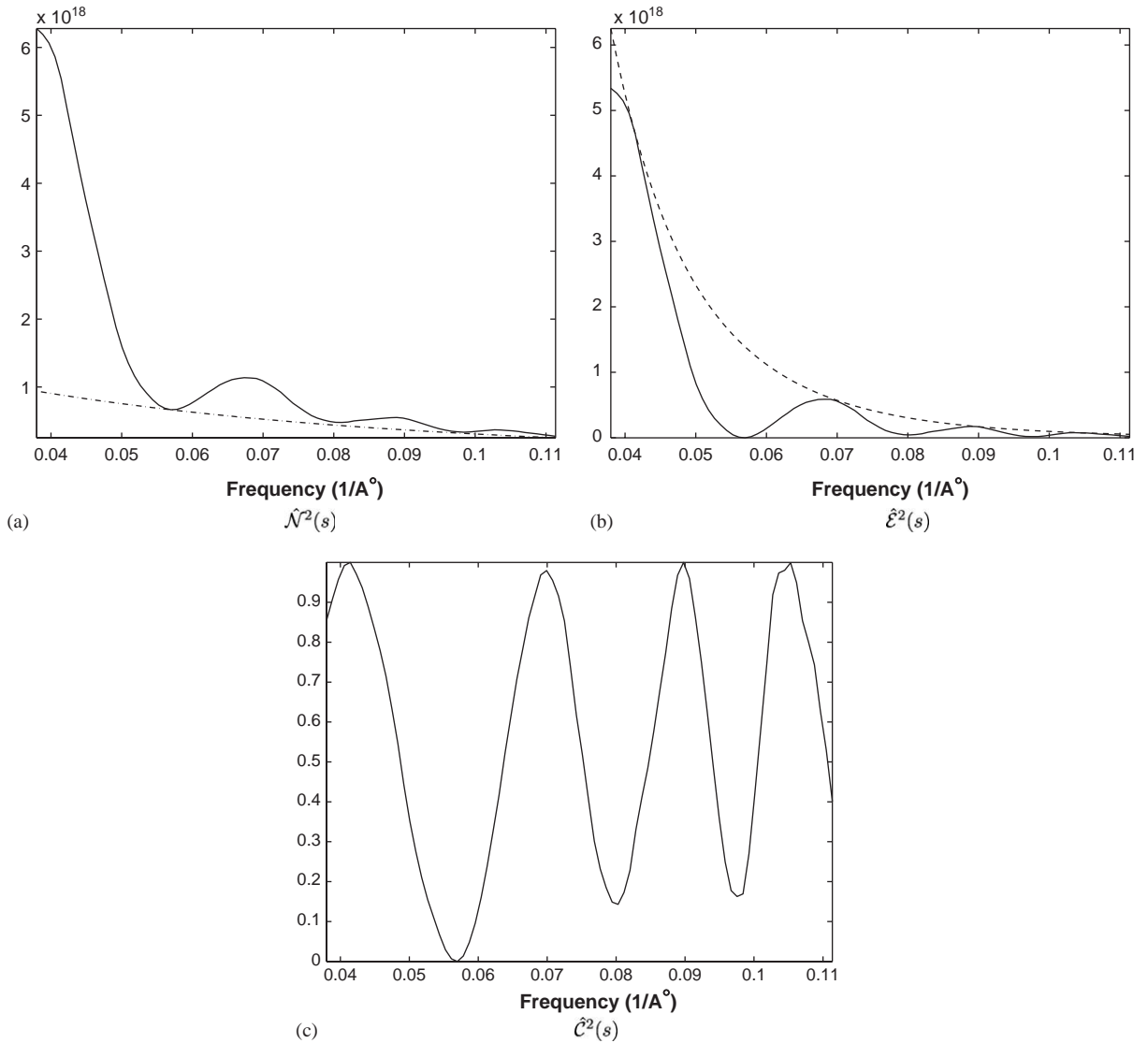
(a) $\hat{\mathcal{N}}^2(s)$

(b) $\hat{\mathcal{E}}^2(s)$

(c) $\hat{\mathcal{C}}^2(s)$

Fig. 6. Estimation of the CTF^2. In (a) the solid curve shows the part of the power spectrum between the lower and upper cutoff frequency. The dotted curve shows the estimated noise spectrum. In (b) the background subtracted 1D profile of the estimated power spectrum is shown using the solid curve. The estimated square of the envelope function is shown using the dashed curve. In (c) an estimate of the CTF^2 is shown.

result. A typical result of the estimates of the lower and upper cutoff frequencies is shown in Fig. 5. All subsequent results are shown in the frequency region between the upper and lower cutoff frequency indicated by the dashed lines.

A typical fit of noise spectrum $\hat{\mathcal{N}}^2(s)$ is shown in Fig. 6(a). Notice that the noise spectrum passes through the local minima and is strictly below the power spectrum. A typical fit of the square of the envelope function $\hat{\mathcal{E}}^2(s)$ is shown in Fig. 6(b). The square of the envelope function passes

through the local maxima of the noise subtracted power spectrum and is strictly above it. Fig. 6(c) shows an estimate of the square of the CTF (CTF^2) found using Eq. (28). The CTF^2 corresponding to a crude estimate of the defocus value is shown in Fig. 7(a) using a dashed curve. The crude estimate is obtained from the local minima shown using dots. Note that even with the crude estimate of defocus, many of the low frequency minima of the theoretical and estimated CTF^2 are aligned. This estimate is close to the real solution. This crude estimate is the initial value supplied to the non-linear constrained optimization algorithm used to refine the estimate of defocus and amplitude contrast. The dashed curve in Fig. 7(b) shows the CTF^2 corresponding to refined estimates of defocus and amplitude contrast. Notice that with the refined estimate of defocus and amplitude contrast, the minima of the estimated and theoretical CTF^2 are aligned more closely, especially at higher frequencies. Finally we show a composite of the true power spectrum and the estimated power spectrum in Fig. 8 for a 2D visualization of the estimated CTF.

5.2. Defocus series experiment

We designed a defocus series experiment to test the accuracy of the algorithm. The defocus of the microscope was changed from $0.6\,\mu m$ under-focus to $6\,\mu m$ under-focus in steps of $0.2\,\mu m$ without changing the astigmatism. Corresponding to each defocus setting an image of the carbon support film was recorded using a CCD camera. The motivation for designing such an experiment is as follows.

First, the accuracy with which defocus can be measured depends to an extent on the value of defocus itself. For example, it is easier to estimate defocus in the far from focus images as compared to near to focus images. Hence, the accuracy of an algorithm should be tested on a wide range of defocus values.

Second, the microscope can set defocus increments very precisely, but the zero defocus, which is used as a reference to set defocus values, is difficult to set to a high degree of accuracy. In other words, the zero defocus contains a zero error which
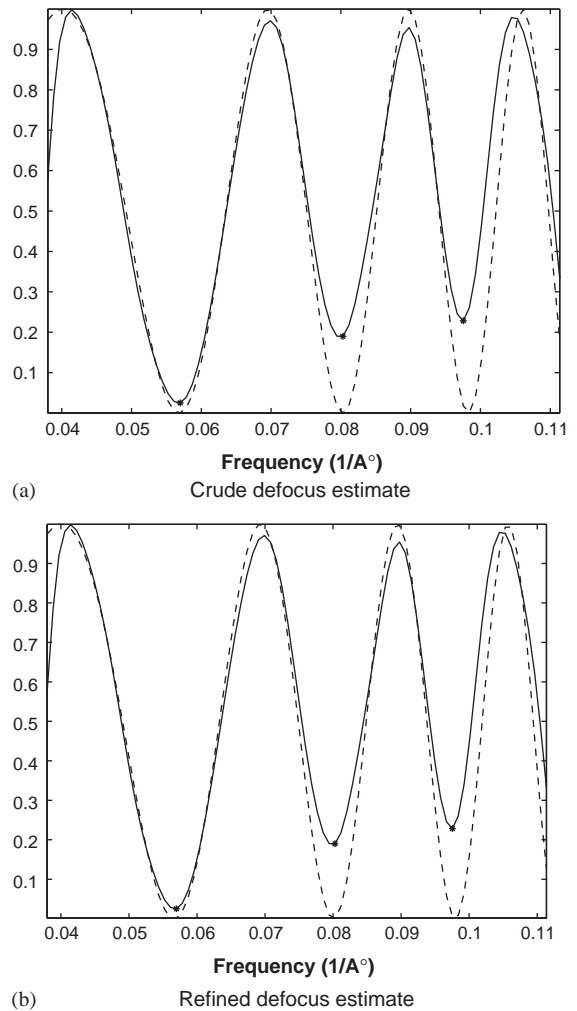

(a)   Crude defocus estimate

(b)   Refined defocus estimate

Fig. 7. The result of defocus and amplitude contrast estimation is shown. The solid curve is an estimate of the CTF^2 ($\hat{\mathscr{C}}^2(s)$). The dashed curve in (a) is the theoretical CTF^2 obtained using the crude defocus estimate. The crude estimate of defocus is based on the local minima shown using dots. In (b) defocus and amplitude contrast estimates are refined. Notice the local minima at higher frequencies are better aligned with the minima of the theoretical CTF^2 after refinement.

introduces a bias in the defocus set by the microscope. Therefore, to test the accuracy of an algorithm, comparing the estimated defocus and the nominal defocus set by the microscope is not appropriate. However, the difference between two defocus values set by the microscope
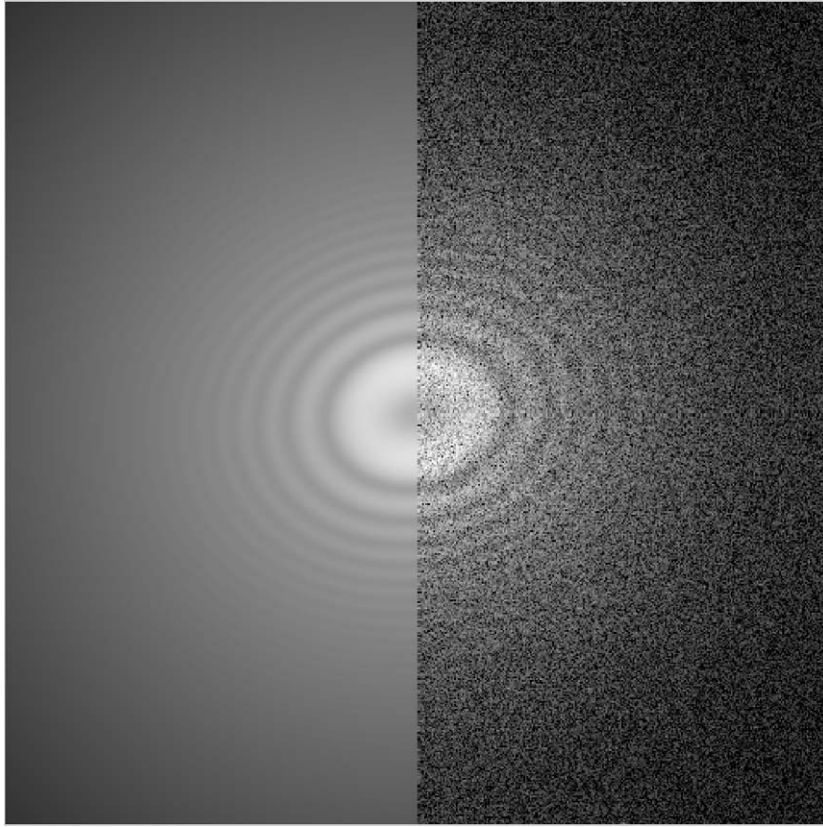
Fig. 8. A composite of the estimated (left) and observed power spectrum (right) in 2D is shown.

is independent of the zero error and is a suitable quantity to assume as ground truth for testing the efficiency of the algorithm. Comparing the nominal and estimated change in defocus would prove the algorithm to be correct up to a constant bias.

With the above considerations in mind, a large (and practical) defocus range of 0.6 μm under-focus to 6 μm under-focus was chosen with small increments of 0.2μm. The defocus was estimated for each image. The mean of the difference between successive defocus estimates was 0.203 μm and the standard deviation was 0.016 μm. The plot of the calculated defocus versus the nominal defocus is shown in Fig. 9. The solid line in Fig. 9 shows a straight line fit through the data. The zero error, which is given by the y-axis offset, was found to be 0.085 μm.

5.3. GroEL embedded in ice

The algorithm was next tested using images of single particles of GroEL embedded in vitreous ice and suspended over holes in the carbon support film. A typical image is shown in Fig. 10. The 2D power spectrum of each image was obtained by cropping out overlapping regions of the image and averaging their power spectra. Fig. 11 shows the result of edge detection and ellipse fitting for the image in Fig. 10 and the subsequent further stages of the algorithm are depicted in Fig. 12. The local minima in the CTF which are barely perceptible in Fig. 11 become much more prominent after the elliptical averaging step. Fig 12(e) shows the theoretical CTF^2 based on the initial estimate of defocus. The mean squared error between the theoretical CTF^2 (shown using dashed curve) and
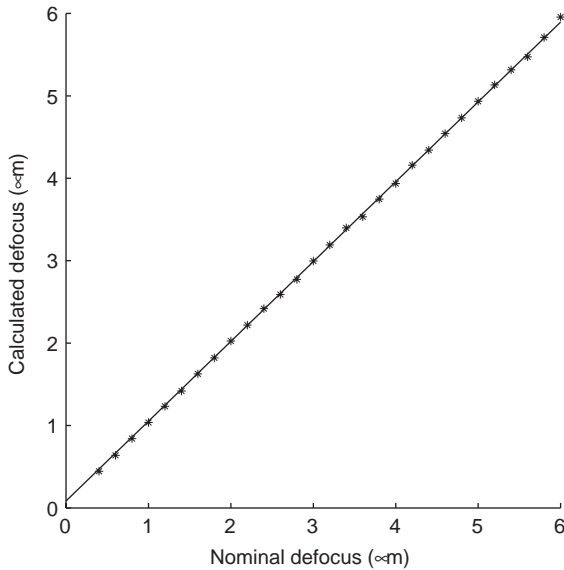
Fig. 9. The plot of the calculated defocus versus the nominal defocus set by a microscope is shown using dots. A line fit to the above data points is also shown. The zero error, which can be seen in the y-axis offset, was found to be 0.085 μm.

the estimate of CTF^2 (shown using solid curve) was calculated to be 0.049. The mean squared error reduces to 0.041 when the refined estimates of defocus and amplitude contrast are used for the calculation of the theoretical CTF^2 (Fig. 12(f)). A visual comparison of Fig. 12(e) and (f) also shows that the CTF^2 based on the refined estimate of defocus and amplitude contrast fits the estimate of the CTF^2 better.

The algorithm can fail due to a few reasons. On a small percentage of images, the algorithm does not converge to the true solution due to low signal to noise ratio. With the astigmatism estimation turned on, the failure of edge detection and ellipse fitting can also lead to incorrect results. Therefore, with astigmatism turned on the failure rate is higher because of the additional edge detection step involved. It is important to note that if the edges are properly detected, the algorithm is always more accurate when the astigmatism estimation is turned on. Hence, success rate should not be confused with accuracy.

To test the success rate, the algorithm was applied to a total of 540 images of GroEL particles suspended over holes. With astigmatism estimation turned off, the algorithm produced correct estimates (as verified by visual inspection of the fitted 1D power spectrum) of the CTF in 527 of the images (97.59% success rate). In 10 images (1.85%) the calculated parameters were incorrect and 3 images (0.56%) could not be processed at all. When the astigmatism estimation was turned on, 494 out of 540 images could be processed correctly (91.48% success rate). In 46 images (7.96%), ACE calculated incorrect results while 3 images (0.56%) could not be processed at all. As mentioned earlier the success rate falls due to the failure of the edge detection step.

The mean of the amplitude contrast calculated using the GroEL dataset of 540 images was found to be 7.6% with a variance of 0.4%. This result is in close agreement with the 7% amplitude contrast reported by Toyoshima et al. [33].

6. Discussion

We have implemented an automated algorithm for estimating the parameters of the CTF including the determination of astigmatism. The parameter estimation is very accurate when applying the algorithm to images of carbon support films. The problem becomes much more challenging when trying to estimate the CTF directly from images of protein specimens embedded in vitreous ice and suspended over holes in the carbon support film. One approach to this problem that has sometimes been taken is to estimate the defocus of the specimen using a nearby image of the carbon support film. However, we can show that the assumption that the carbon support film and the specimen in the nearby hole are at the same defocus is not always valid. For example, in one particular experiment in which we measured both the defocus of the specimen over the hole as well as from the carbon support film outside of the hole we found that the differences can be as large as 0.5 μm. These results are shown in Fig. 13. On a subset of the images used in Fig. 13, where the difference in the two defoci was found to be large,
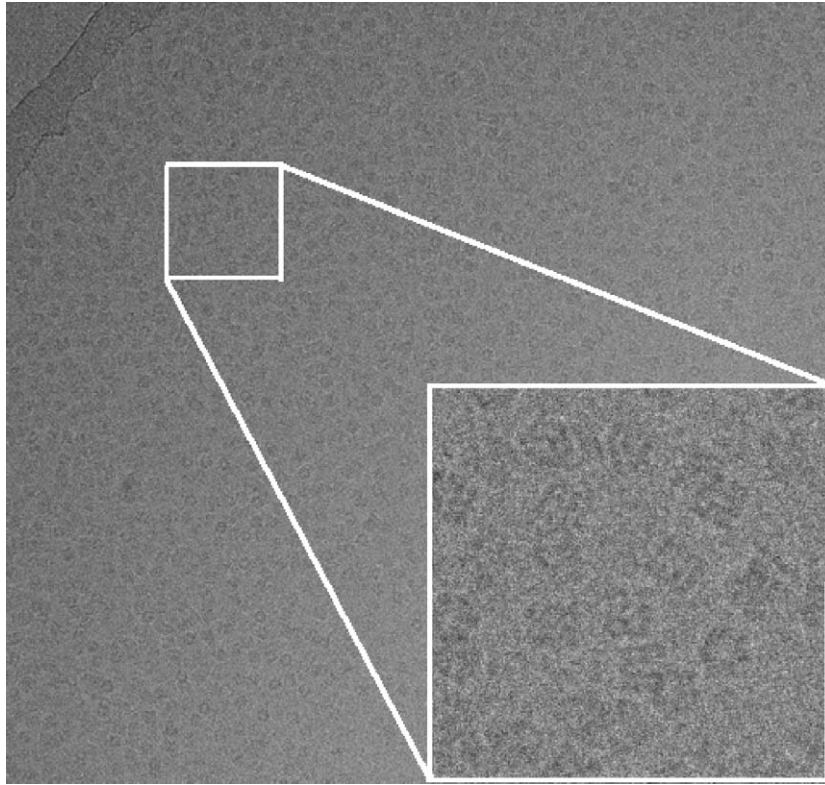
Fig. 10. The figure shows a typical image of GroEL particles embedded in ice suspended in a hole. A section of the micrograph is zoomed in to show the density of the particles in the hole.

the results were manually verified. Particles from a single image were boxed and the defocus was estimated using the CTFIT program in EMAN [28]. The defocus found manually was in close agreement with the defocus found automatically using ACE.

The calculation of astigmatism is very robust for carbon images. However, when the two defoci are nearly the same (i.e. $r \approx 1$), the direction of the major axis is arbitrary. This is as expected because when $r \approx 1$, the Thon rings are circular. This does not affect the 1D averaging because no matter what direction of major axis is chosen, the value of $r \approx 1$ will result in circular averaging. For ice-embedded particle images, the accuracy of astigmatism estimation depends on the kind of particle being imaged. For small particles having a dense population in ice, like most of the GroEL images described

in the Results section above, the estimation can be fairly reliable. However, for images with low densities of particles spurious edge detection results in a failure of the astigmatism estimation. For particles which are highly ordered, for example helical filaments like Tobacco Mosaic Virus (TMV), the ordered arrangement of the subunits results in a series of strong diffraction layer lines which interfere with the edge detection algorithm. For these situations, a possible approach is to use the carbon images near the location of the particles to estimate the astigmatism, as unlike defocus, this does not depend on the location of the imaged object. Once the astigmatism parameters have been estimated, the other parameters can be estimated directly from the particle images following elliptical averaging. This approach was used to estimate the CTF for 95 images of TMV embedded in ice with
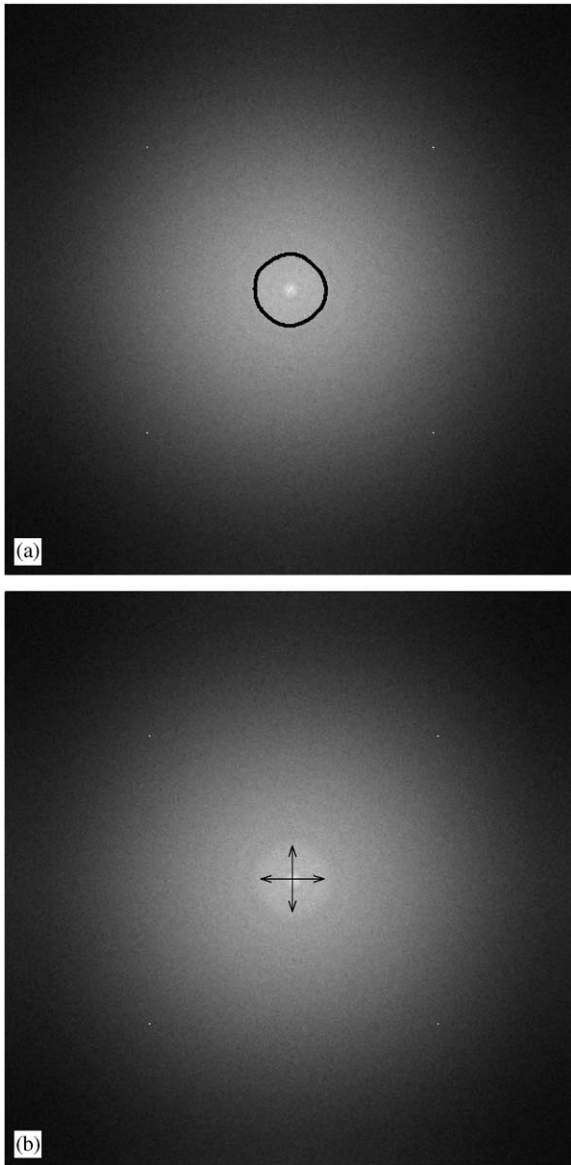
Fig. 11. The result of edge detection and ellipse fitting is shown for the power spectrum of an image of GroEL embedded in ice.

100% success rate (as judged by visual inspection of the fitted CTF).

The issue of assuming the structure factor to be flat also needs to be addressed. It is clear that having a structure factor would increase the accuracy of the estimation procedure. The present algorithm can be modified to take structure factor into account. The noise estimation does not change but before estimating the envelope function, the structure factor can be compensated for. The effect of structure factor on the accuracy of algorithm depends largely on the kind of particle being imaged. For the micrographs in which the particles are asymmetrical and are randomly oriented, the estimation is reliable. For micrographs which contain particles with repetitive patterns or particles which are aligned in a fixed orientation, the estimation can be inaccurate. As a rule of thumb, if the Thon rings are more pronounced in the power spectrum as compared to the structure factor, it is reasonable to assume that the estimation procedure will be accurate.

Without the ability to accurately estimate the astigmatism of acquired images, it is necessary to acquire images with the astigmatism set as close to zero as possible or to reject images exhibiting any noticeable astigmatism. However, if the parameters of astigmatism can be determined accurately, then Wiener filtering can in principle restore astigmatic images as accurately as non-astigmatic images. A few people in the past have raised the point that using astigmatic images for single particle reconstruction might be valuable in terms of avoiding resolution gaps resulting from the zeroes in the CTF. However as far as we are aware these discussions have not been published anywhere in the literature. This is perhaps an area for further investigation.

7. Conclusion

We have presented a completely automated method for the estimation of the CTF, the envelope function and the noise spectrum parameters of an image taken using a TEM. The method incorporates a novel way of estimating astigmatism, and reduces the problem of CTF estimation to a truly 1D estimation problem using elliptical averaging. The accuracy of the algorithm was demonstrated using images of carbon support film as well as on large datasets
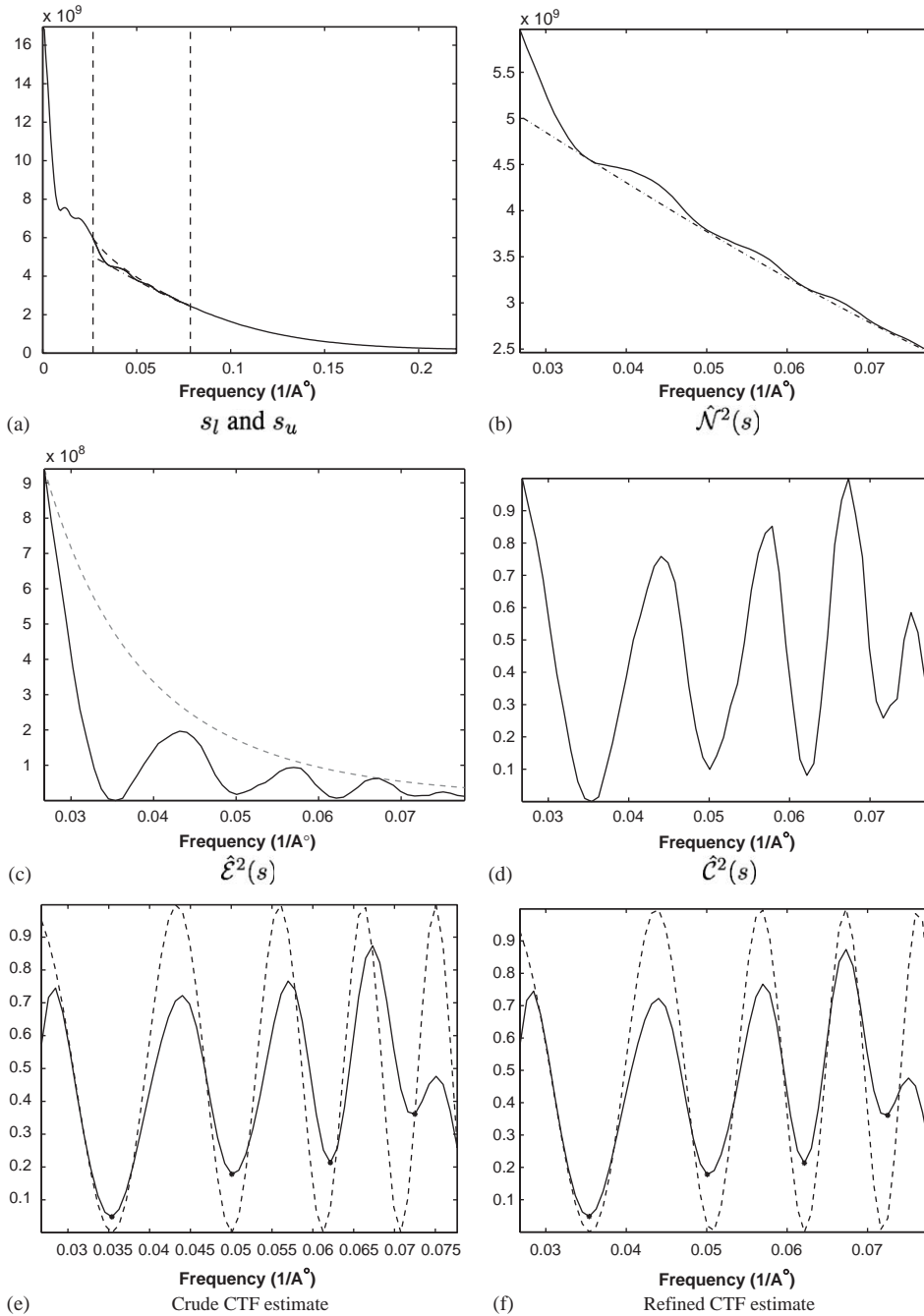
Fig. 12. This figure shows the various stages of the algorithm tested on GroEL embedded in ice. (a) shows the elliptically averaged power spectrum. The vertical dashed lines in (a) show the estimated lower and upper cutoff frequency. In (b), the solid curve represents the part of the power spectrum between the lower and upper cutoff frequency. The dashed curve represents the estimate of noise spectrum ($\hat{\mathcal{N}}^2$). In (c) the solid curve represents the noise subtracted power spectrum and the dashed curve represents the estimate of the square of the envelope function ($\hat{\mathcal{E}}^2$). In (d) the estimate of the square of the CTF (CTF^2) recovered from the power spectrum is shown. In (e) the solid curve shows the estimated CTF^2 based on the power spectrum. The dashed curve shows the theoretical CTF^2 based on an initial estimate of defocus. The initial estimate of the defocus was based on the local minima (shown using dots) of the estimated CTF^2 and was calculated using Eq. (32). The theoretical CTF^2 based on a refined estimate of defocus and amplitude contrast is shown in (f).
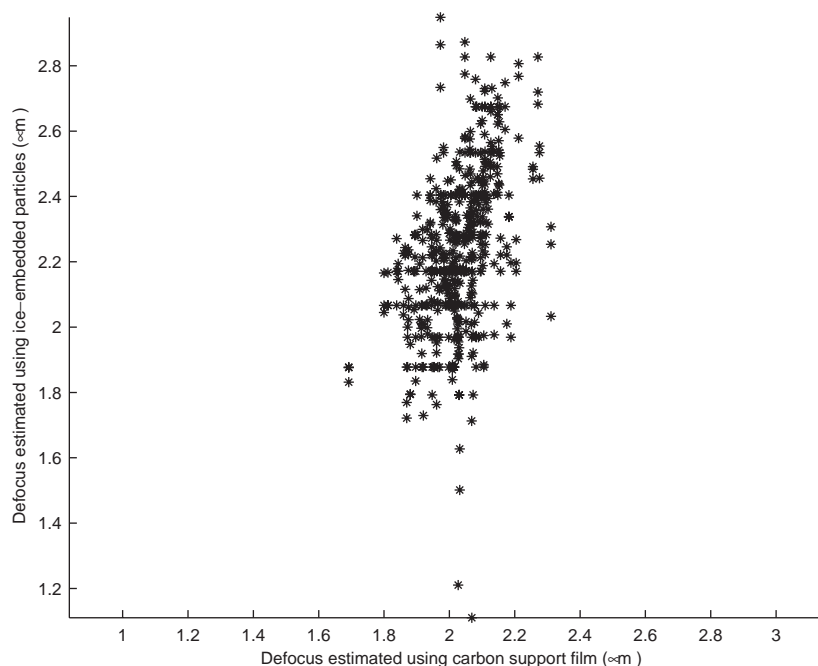
Fig. 13. The plot of calculated defocus of the specimen in a hole versus defocus calculated using carbon support film outside of the hole is shown for a dataset containing 540 images (4096 × 4096) of GroEL embedded in ice. Notice that the defocus calculated using the images of carbon support film can be very different from the actual defocus at the location of the specimen.

of single particles embedded in ice. A MATLAB implementation of the algorithm called ACE (Automated CTF Estimation) is freely available at the following web pages: http://nramm. scripps.edu/software/ace http://graphics.ucsd.edu/ ~spmallick/research/ace

References

[1] C.S. Potter, H. Chu, B. Frey, C. Green, N. Kisserberth, T.J. Madden, K.L. Miller, K. Nahrstedt, J. Pulokas, A. Reilein, et al., Ultramicroscopy 77 (1999) 153.
[2] B. Carragher, N. Kisserberth, D. Kriegman, R.A. Milligan, C.S. Potter, J. Pulokas, A. Reilein, J. Struct. Biol. 132 (2000) 33.
[3] Y. Zhu, B. Carragher, D. Kriegman, C.S. Potter, J. Struct. Biol. 135 (2001) 302.
[4] C.R. Booth, W. Jiang, W. Baker, Z. Hong Zhou, S.J. Ludtke, W. Chiu, J. Struct. Biol. 147 (2004) 116.
[5] P. Zhang, A. Beatty, J.L.S. Milne, S. Subramaniam, J. Struct. Biol. 135 (2001) 251.
[6] F. Thon, Electron Microsc. Mater. Sci. (1971) 570.
[7] R.H. Wade, Ultramicroscopy 46 (1992) 145.
[8] P.W. Hawkes, Electron Tomogr. (1992) 17.
[9] P.W. Hawkes, E. Kasper, Principles of Electron Optics: Wave Optics, vol. 3, Academic Press, London, 1994.

[10] K.J. Hanszen, The optical transfer theory of the electron microscope: fundamental principles and applications, Adv. Opt. Electron Microsc. 4 (1971) 1.

[11] J. Frank, Biophys. J. 12 (1972) 484.

[12] R. Henderson, J.M. Baldwin, K.H. Downing, J. Lepault, F. Zemlim, Ultramicroscopy 65 (1996) 31.

[13] Z.H. Zhou, W. Chiu, Ultramicroscopy 49 (1993) 407.

[14] J. Zhu, J. Frank, Accurate retrieval of transfer function from defocus series, Proceedings of the 13th International Congress on Electron Microscopy, vol. 1, 1994, pp. 465–466.

[15] K. Tani, H. Sasabe, C. Toyoshima, Ultramicroscopy 65 (1996) 31.

[16] J.A. Valázquez-Muriel, C.O.S. Sorzano, J.J. Fernández, J. Carazo, Ultramicroscopy 96 (2003) 17.

[17] J. Fernández, J. Sanjurjo, J. Carazo, Ultramicroscopy 68 (1997) 267.

[18] B. Sander, M. Golas, H. Stark, J. Struct. Biol. 142 (2003) 392.

[19] J.A. Mindell, N. Grigorieff, J. Struct. Biol. 142 (2003) 334.

[20] Z. Huang, P.R. Baldwin, S. Mullapudi, P. Penczek, J. Struct. Biol. 144 (2003) 79.

[21] J. Frank, Three-dimensional Electron Microscopy of Macromolecular Assemblies, Academic Press, New York, 1996.

[22] J. Frank, Optik 38 (1973) 519.

[23] R.H. Wade, J. Frank, Optik 49 (1977) 81.

[24] J.M. Kenney, J. Hantula, S.D. Fuller, L. Mindich, P.M. Ojala, D.H. Bamford, Virology 190 (1992) 635.

[25] K.H. Downing, D.A. Grano, Ultramicroscopy 7 (1982) 381.

[26] J. Zhu, P. Penczek, R. Schröder, J. Frank, J. Struct. Biol. 118 (1997) 197.

[27] A. Saad, S.J. Ludtke, J. Jakana, F.J. Rixon, H. Tsuruta, W. Chiu, J. Struct. Biol. 133 (2001) 32.

[28] S.J. Ludtke, P.R. Baldwin, W. Chiu, J. Struct. Biol. 128 (1997) 82.

[29] J. Canny, IEEE Trans. Pattern Anal. Machine Intell. 8 (1986) 679.

[30] M.A. Fischler, R.C. Bolles, Comm. ACM 24 (1981) 381.

[31] T.F. Coleman, Y. Li, SIAM J. Optim. 6 (1996) 1040.

[32] T.F. Coleman, Y. Li, SIAM J. Optim. 6 (1996) 418.

[33] C. Toyoshima, N. Unwin, Ultramicroscopy 24 (1988) 279.