Transfer Function Restoration in 3D Electron Microscopy via Iterative Data Refinement

C.O.S. Sorzano[(1,2)], R. Marabini[(2,3)], G.T. Herman[(4)],

Y. Censor[(5)] and J.M. Carazo[(2,3)]

4th December 2003

[(1)] Escuela Politécnica Superior, Univ. San Pablo-CEU, Campus Urb. Montepríncipe, s/n, 28668 Boadilla del Monte, Madrid, Spain.

[(2)] Centro Nacional de Biotecnología, Campus Universidad Autónoma s/n, 28049 Cantoblanco, Madrid, Spain.

[(3)] Escuela Politécnica Superior, Univ. Autónoma de Madrid, Campus Universidad Autónoma s/n, 28049 Cantoblanco, Madrid, Spain.

[(4)] The Graduate Center, The City University of New York, 365 Fifth Avenue, New York, NY 10016-4309, USA.

[(5)] Dept. of Mathematics, University of Haifa, Mt. Carmel, Haifa 31905, Israel.

Abstract

Three-dimensional electron microscopy (3D-EM) is a powerful tool for visualizing complex biological systems. As any other imaging device, the electron microscope introduces a transfer function (called in this field the Contrast Transfer Function, CTF) into the image acquisition process that modulates the various frequencies of the signal. Thus, the 3D reconstructions performed with these CTF-affected projections are also affected by an implicit 3D transfer function. For high resolution electron microscopy, the effect of the CTF is quite dramatic and limits severely the achievable resolution. In this work we make use of the *Iterative Data Refinement* (IDR) technique to ameliorate the effect of the CTF. It is demonstrated that the approach can be successfully applied to noisy data.

1 Introduction

The analysis of macromolecular complexes and their dynamics is one of the most interesting challenges in molecular biology. A promising future is awaiting the electron microscopist due to the possibilities of visualizing molecular machines, reconstructing unique (as opposed to averaged) objects, and imaging dynamic processes. The road to achieving these possibilities is via three-dimensional reconstruction from electron-microscopic images of the macromolecular complexes.

There are many methods for reconstructing a three-dimensional object from its line integrals (Herman, 1980; Natterer and Wübbeling, 2001). Typically, the line integrals are estimated for a set of parallel lines from a projection image that

is obtained by some instrument. A difficulty that arises in electron microscopy is that the image that is produced by the instrument corresponds to the convolution of the ideal projection image with a Point Spread Function (PSF). The PSF is usually described by its Fourier transform that is commonly called the Contrast Transfer Function (CTF); for examples see Figures 1 and 2.

The CTF severely limits the achievable resolution in the three-dimensional reconstruction. In particular, it filters both the high and the low frequencies, introduces zones of alternate contrast and eliminates all information at certain frequencies. It is, therefore, desirable to replace the reconstruction obtained by a "real" microscope by a reconstruction that would be obtained from images that would be produced by an ideal, aberration-free microscope. In order to achieve this goal several methods have been proposed: Frank and Penczek (1995) applied Wiener filtering in the three-dimensional space to the reconstructed volume; Zhu et al. (1997) incorporated a three-dimensional PSF into the data model and used a regularized steepest-descent technique; Stark et al. (1997) applied inverse CTF filtering to the reconstructed volume; Skoglund et al. (1996) incorporated a two-dimensional CTF, particular to each projection to the projection, model in a maximum-entropy reconstruction algorithm; Grigorieff (1998) provided a Fourier reconstruction algorithm in which the CTF for each projection is considered in a Wiener-like fashion; Ludtke et al. (1999) proposed a CTF correction applied to the individual projections with a weighting function in the Fourier space computed from a set of images sharing a common CTF; Ludtke et al. (2001) added a Wiener filter to the weighting function defined in Ludtke et al. (1999). An alterna-

tive is to explicitly introduce the effect of the CTF in the reconstruction equations; this was done by Zubelli et al. (2003), who then reformulated the problem so that Chahine's method became applicable to it. The existence of these multiple approaches is indicative of the fact that there is no agreed standard technique for the correction of CTF effects in 3D-EM of single particles and the search for superior methods is still active.

In this work we apply the technique of Iterative Data Refinement (IDR) – introduced in Censor et al. (1985) and further studied in Herman (1989), Herman and Ro (1990), Losada and Navarro (1998) and Ro et al. (1989) – to reduce the effect of the CTF and, thus, to obtain high-resolution structural information about the macromolecules under study. As opposed to many of the approaches discussed in the previous paragraph, our proposed algorithm can handle the case of differing CTFs in the projections and does not require estimation of the Signal-to-Noise Ratio (SNR). The potential benefit of the method is illustrated by an experiment that involves realistic simulation of the electron microscopic imaging of a biological macromolecule.

2 Mathematical Background

Contrast Transfer Function

Image formation by an electron microscope is due to several physical processes of electron interaction with the specimen. These effects combine to produce a single CTF, see Frank (1996, Chapter 2.II). A parametric model of this transfer function

has been used in the simulations presented in this work. This parametric model accounts for the various effects involved in the CTF (Zhu et al. (1997) and Frank (1996, Chapter 2.II)) and was also used in Zhou et al. (1996).

Basically, the microscope transfer function is a real-valued function in Fourier space formed by a damped harmonic function. The "sine" part of this function comes from the phase change that electrons undergo when interacting with the sample specimen. A detailed study of the electronic interaction in the image formation plane shows that the transfer function of an electron microscope can be usefully approximated by

$$CTF(\omega) = E(\omega)\left(sin(\pi\Delta f\,|\omega|^2) - Q_0 cos(\pi\Delta f\,|\omega|^2)\right), \tag{1}$$

where ω is the spatial frequency, Δf is the defocus, and Q_0 is a factor accounting for the loss of electrons during the image formation process. Usually Q_0 is a small number, which implies that the DC component of the projection Fourier transform is nearly removed and, thus, the absolute density values in the projections are not meaningful. For this reason, usually only the relative values are taken into consideration when interpreting a 3D-EM reconstruction. The damping envelope, $E(\omega)$, models microscope imperfections such as chromatic aberration, spherical aberration, current and voltage instabilities, angular aperture, etc.; see, e.g., Zhou et al. (1996) or Frank (1996, Chapter 2.II).

The model explained so far defines the shape of the profile of the CTF (for a typical example see Figure 1). Many studies assume that this profile is radially

symmetric, although this is not necessarily true. Astigmatism is a well-known effect which turns the circles produced by the radial symmetrization of the CTF profile into ellipses (see Figure 2). This results in a different defocus along every radial line of the Fourier space.

Phase Flipping

Notice (in Figure 1) that the sign changes at the zero-crossings result in a contrast inversion in the projection image and cause the complete elimination of the information at certain frequencies. This is a very limiting factor in electron microscopy, since without CTF correction all reconstructions are unreliable at frequencies beyond the one where the CTF first becomes zero. Nevertheless, biologically useful results can sometimes be obtained even without CTF correction (Bárcena et al., 2001; Sorzano et al., 2001). However, this can only be the case if the important biological information is not in the high-frequency part of the reconstruction, since (as illustrated in our experiment reported below) in a reconstruction without CTF correction, the information regarding frequencies beyond the first zero-crossing of the CTF is incorrect.

A simple method to alleviate this problem consists of multiplying the projection Fourier transform by the sign of the CTF (this approach is named *phase flipping* (Frank, 1996, page 45)) to produce the corrected projection data to which the reconstruction algorithm is then applied. Thus the correction of the CTF sign is simple, as it only needs to adjust the sign at those frequencies where it is flipped. However, amplitude correction is more difficult, as it requires either dividing by

the transfer function (avoiding zeroes by using, for instance, a Wiener filter) or the incorporation of the CTF operator into the reconstruction algorithm, allowing each projection to have its own CTF. The IDR approach of this paper addresses this more difficult problem. It will be set up based on the assumption that the data had been already corrected by phase flipping.

Computational Representation of Volumes and Projections

For the computational procedures of this paper we need to establish conventions for representing volumes and projections by finite sets of numbers. In this subsection we present our conventions and explain the operators which are incorporated into our algorithms.

We approximate arbitrary volumes by finite series expansions of the general form

$$\sum_{j=1}^{J} c_j b(\mathbf{r} - \mathbf{r}_j). \tag{2}$$

In this formula $\mathbf{r}$ is the point at which the volume is being approximated, the $\mathbf{r}_j$ are fixed points in space, b is a function of three variables, and the c_j are the coefficients of the expansion. In any application, b and the $\mathbf{r}_j$ are fixed, it is the c_j that distinguish one volume from another. Following Lewitt (1990, 1992) and Matej and Lewitt (1995, 1996), we use a generalized Kaiser-Bessel window function (also called a *blob*) for b and a finite subset of the Body Centered Cubic grid for the points $\mathbf{r}_j$. Such a representation was found useful in electron microscopy applications; see Bárcena et al. (2001); Marabini et al. (1997); Marabini et al.

(1998); Sorzano et al. (2001). The specific choice that we adopt for the blob and the grid is the one referred to as the "standard blob" in Matej and Lewitt (1996).

We approximate a projection by a two-dimensional array of numbers, each representing a projection value at a point of a square grid. To bring this into the electron microscopy context, we think of the square grid as lying on a projection plane that is perpendicular to the direction of the electrons. Assuming that we have m different projection planes, we use $\mathbf{g}_i$ ($1 \leq i \leq m$) to denote the array of numbers associated with the ith projection. We use $\mathbf{g}$ to denote the complete set of the m projections, meaning that $\mathbf{g}$ is the concatenated vector of all the $\mathbf{g}_i$s.

Given a volume representation as in (2) it is easy to calculate the ideal projection (line integrals along lines perpendicular to the projection plane and passing through the points of the square grid). This is so because the integration can be brought inside the summation and can be analytically evaluated for the known blob b and grid point $\mathbf{r}_j$. For $1 \leq i \leq m$, we define an *ideal projection operator* P_i that associates with the J-dimensional vector $\mathbf{c}$ (whose jth component is c_j) the vector representing the ideal projection of the volume onto the ith projection plane. Note that, in practice, P_i is a matrix of J columns and as many rows as the number of grid points in the ith projection plane.

Each projection also has its own *CTF operator* that we denote by H_i. In practice, given the projection $\mathbf{g}_i$, $H_i\mathbf{g}_i$ is computed by taking the discrete Fourier transform of $\mathbf{g}_i$, multiplying it point-wise by the phase-flipped CTF associated with the ith projection, and then taking the inverse discrete Fourier transform.

Iterative Data Refinement

The measuring device (the electron microscope) provides data that only approximate what we intend to measure. The discrepancy between the *actual data* (under our assumptions, corrupted by the phase-flipped CTF) and data that are idealized (uncorrupted by the CTF, henceforth called *ideal data*) can be estimated from the actual data and knowledge of the measuring process, leading to a better approximation of the ideal data. This new approximation can then be used to estimate the new discrepancy, and the process can be repeated. Our knowledge of the measurement process is insufficient to obtain the ideal data exactly, but the original discrepancy is significantly reduced by just a few of such iterative steps. This process is accomplished by the *iterative data refinement* (IDR) methodology of Censor et al. (1985). Here we briefly review the fundamentals of IDR and describe our specific implementation of it for the CTF removal problem. We then supply a short discussion that puts the approach in perspective and relates it to current literature.

In the following we use R to denote a recovery operator (in our case a three-dimensional reconstruction algorithm) that produces, for a complete set of projections $\mathbf{g}$, a vector $\mathbf{c}$ that represents a volume using (2). For now it is not important to specify our choice of R, we will do so below. An important assumption about R (well-justified by the known behavior of reconstruction algorithms (Herman, 1980)) is that if it is applied to the ideal data $\hat{\mathbf{g}}$ (that is, the concatenated vector of all the $P_i\mathbf{c}$s), then $\hat{\mathbf{c}} = R\hat{\mathbf{g}}$ is an acceptable approximation to $\mathbf{c}$. The problem is that, in practice, the actual data $\tilde{\mathbf{g}}$ is corrupted by the CTF and so $R\tilde{\mathbf{g}}$ is not a

satisfactory approximation to $\mathbf{c}$. The IDR approach aims at estimating the ideal data $\hat{\mathbf{g}}$ from the actual data $\tilde{\mathbf{g}}$, assuming knowledge of the CTFs.

IDR produces a sequence of vectors $\mathbf{g}^k$ ($k = 0, 1, 2, \cdots$); the aim is that they should be improving estimates of the ideal data. We denote by $\mathbf{g}_i^k$ the part of $\mathbf{g}^k$ that is associated with the ith projection. With this notation, our version of the IDR algorithm is formulated as the following iterative process. (Figure 3 depicts a block diagram of the algorithm.)

Algorithm: Iterative Data Refinement (IDR) for 3D Electron Microscopy

Initialization: *Take $\mathbf{g}^0 = \tilde{\mathbf{g}}$, the actual data (i.e., the experimentally measured data corrected by phase flipping).*

Iterative Step: *Given the current iterate $\mathbf{g}^k = (\mathbf{g}_i^k)_{i=1}^m$, calculate the next iterate $\mathbf{g}^{k+1} = (\mathbf{g}_i^{k+1})_{i=1}^m$ by using, for all $i = 1, 2, \cdots, m$, the formula*

$$\mathbf{g}_i^{k+1} = \mu_k \mathbf{g}_i^0 + (P_i - \mu_k H_i P_i) R \mathbf{g}^k, \tag{3}$$

where $\{\mu_k\}_{k=0}^{\infty}$ is a sequence of so-called relaxation parameters.

This algorithm generates an iterative sequence $\{\mathbf{g}^k\}_{k=0}^{\infty}$ which is guaranteed to converge to the ideal data $\hat{\mathbf{g}}$ under some stringent conditions, see Censor and Zenios (1997, Proposition 10.5.8). However, in practice, it has been shown (see Censor et al. (1985) and references therein) that, even when convergence cannot be guaranteed, the early iterates produced by the IDR algorithm are closer to the ideal data vector $\hat{\mathbf{g}}$ than $\mathbf{g}^0$ is. This property of IDR is the basis of the present study.

The underlying idea of IDR is to refine the *data* iteratively in a way that bridges the gap between an accurate model of data collection (but one for which we do not have a reconstruction algorithm) and an approximate model that leads to a reconstruction algorithm (which would work if the model were correct). This is quite different from just being another reconstruction method, since it iterates on the *data* rather then on the *unknowns* of the reconstruction problem. The IDR method is a member of the family of "iterative defect-correction methods" much used in the field of differential equations, see, e.g., Böhmer et al. (1984); Stetter (1978). For additional applications of the IDR approach, consult Section 2 of Censor et al. (1985), where beam hardening correction in x-ray computerized tomography, attenuation correction in single photon emission computed tomography (SPECT), and image reconstruction with incomplete data are described along with references to the original studies. See also Herman and Ro (1990) for a study of the connection between IDR and phase retrieval algorithms.

Reconstruction Algorithm

We now return to the choice of the recovery operator R that was left unspecified in the previous subsection. We emphasize that the basic approach of IDR is independent of this choice, any good reconstruction algorithm could be used. In our study we used for the recovery operator R the reconstruction algorithm called *block-ART with blobs* in which each block corresponds to one projection. We note that this algorithm has been found efficacious for 3D reconstruction from electron microscopic data (Bárcena et al., 2001; Marabini et al., 1998; Marabini et al., 1997;

Sorzano et al., 2001).

Block-ART with blobs is an iterative algorithm. Given a data vector $\mathbf{g}$, it produces a sequence of iterates $\{\mathbf{c}^k\}_{k=0}^{\infty}$, each of which defines a volume using (2).

Algorithm: Iterative Block-ART for Volume Recovery

Initialization: *Take* $\mathbf{c}^{(0)} = \mathbf{0}$*, the zero vector.*

Iterative Step: *Given the current iterate* $\mathbf{c}^k$*, calculate the next iterate* $\mathbf{c}^{k+1}$ *by the formula*

$$\mathbf{c}^{k+1} = \mathbf{c}^k + P_i^T \Sigma_k (\mathbf{g}_i - P_i \mathbf{c}^k), \tag{4}$$

where i=k mod m + 1.

Here P_i^T is the transpose of the ideal projection operator P_i, and each Σ_k is a relaxation matrix whose exact nature is explained by Eggermont et al. (1981), who also provide convergence results for the Iterative Block-ART algorithm (their Theorem 1.3). In the experiments, reported below, we do not run the algorithm to convergence, but use only one full cycle of it; i.e., we define $R\mathbf{g}$ to be $\mathbf{c}^m$. (This is justified by previous experience in this application area; see, e.g., Marabini et al. (1998).) Also, we selected each Σ_k to be a diagonal matrix, the value of each entry on the diagonal is a constant λ divided by the square of the norm of the corresponding row of P_i; see (2.18) in Eggermont et al. (1981). (This is also justified by previous experience. It was also found that for the data collection that we used for the experiments reported below, $\lambda = 0.05$ is a good choice.)

The recovery operator R is used not only in the IDR process, but also to produce the final reconstruction from the (possibly corrected) projection data. For this purpose, it needs to be extended by an additional step, since further analysis of the reconstructions requires that they should be evaluated at the points of a cubic grid. Such an evaluation is done using (2), yielding a three-dimensional array of numbers that we consider to be *the* reconstruction.

There is an interesting alternative to be considered here: why not incorporate the effect of the CTF into the model and then use Iterative Block-Art for Volume Recovery directly on the actual data? (In practice this means that in (4) P_i has to be replaced by H_iP_i in both places where it occurs, with corresponding changes in the calculations of the entries of the diagonal matrix Σ_k. Computationally this is some additional burden, but not too much: since H_i is symmetric, $H_i^T = H_i$ and the computation can be carried out by the method described at the end of the Subsection on Computational Representation of Volumes and Projections.) In the current paper the choice of the recovery operator is not essential: it can be replaced in the IDR process by any good reconstruction algorithm. We leave to the future the investigation of the efficacy of IDR relative to applying directly to the actual data either Iterative Block-Art for Volume Recovery or the alternative algebraic approach proposed by Zubelli et al. (2003).

3 Evaluation Methodology

To compare the performance of IDR with no CTF correction and with phase flipping, we adapted the methodology proposed in Furuie et al. (1994) and previously applied to electron microscopy in Marabini et al. (1998); Marabini et al. (1997); Sorzano et al. (2001). In the experiments described below, many sets of simulated electron microscopy projections of a particular molecule are taken and reconstructions are produced from each of these sets using the approaches to be compared. The success of the approaches are determined by the use of figures of merit (FOMs). We now provide details of this outline.

Projection Data Generation

The volume used in our tests was created from an atomic structure deposited in the Protein Data Bank (PDB), see Berman et al. (2000), namely the *Halobacterium halobium* bacteriorhodopsin (PDB id: 1BRD, Henderson et al. (1990)). For comparison purposes this volume was evaluated for points of a $64 \times 64 \times 64$ cubic grid, with distance 3.5 Å between neighboring grid points. We refer to the resulting three-dimensional array as the *phantom*.

Several sets of 2,000 projections were created with signal-to-noise ratio 1/3, this resembles cryomicroscopy conditions. This noise in the measurements was combined with other sources of inconsistency in the form of random translations (by moving the projection plane parallel to itself by a distance randomly selected from a zero-mean Gaussian distribution with standard deviation 7 Å) and random

rotations (by adding a zero-mean 5 degrees standard deviation Gaussian noise to each of the Euler angles that defined the orientation of the projection planes).

In addition to noise, the projections were convolved with a CTF. The parameters for a circularly-symmetric simulated CTF, see (1), were given values typically found in experimental conditions: $Q_0 = -0.06$ and $\Delta f = -20{,}000$ Å, and the factors that influence the damping envelope $E(\omega)$ were selected as acceleration voltage = 100 kV, spherical aberration = 5.5 mm, chromatic aberration = 6 mm, energy loss = 9.9 eV, convergence cone = 0.2 mrad, longitudinal displacement = 80 Å (see Velázquez-Muriel et al. (2003) for a description of these parameters). The radial profile of this CTF is shown in Figure 1. The same CTF was used in all projections, since this is the worst case that can occur because it makes it difficult to compensate for missing information in one projection by data from other projections.

A surface rendering of the ideal volume and a selection of projections are shown in Figure 4 and 5, respectively.

Figures of Merit

Figures of merit (FOMs) are numerical measures of the reconstruction quality that are based on specific aspects. A simple but often used measure is the sum of the squares of the differences between the individual values in the reconstructions and the corresponding values in the phantom. In Sorzano et al. (2001) this measure was divided by the number of points in the cubic grid and the result was called the FOM *scL*2.

More sophisticated measures are provided by the *Fourier Shell Correlation* (FSC), as described in Equation (3.65) of Frank (1996) with F_1 the phantom and F_2 the reconstruction. The FSC indicates, for every shell of frequencies (determined by a frequency k and a shell-thickness Δk), how well the reconstruction correlates with the phantom for all frequencies within that shell. Thus the FSC provides a separate FOM for every shell. The FSC can also be used to provide the additional single FOM of *resolution* by defining it as the frequency at which the FSC falls below 0.5.

The presence of many reasonable FOMs (Sorzano et al. (2001) lists 24 of them) makes exhaustive comparisons difficult. We have developed a methodology that applies multivariate statistics to obtain a single FOM that in some senses captures the essence of what is provided by the full set of FOMs; it is described in detail in the doctoral dissertation of Sorzano (2002). Here we give a brief description of the use of this methodology for selecting an optimal range for a parameter μ, such as one of the relaxation parameters μ_k in (3). To do this, a number of training data sets are created and each one of them is processed using a number of values of the parameter μ. Then the method for producing a single representative FOM proceeds in five steps: first, each of the FOMs that we wish to consider in producing the single FOM is normalized to have mean 0 and standard deviation 1; second, those FOMs that show no dependency (as indicated by a 1-way Analysis of Variance, ANOVA for short) with μ are removed; third, all FOMs showing a similar dependency with μ are clustered by a hierarchical classification; fourth, the dimensionality of the clusters is reduced using a Principal Component Anal-

ysis (PCA) and a single representative is selected for each cluster; and fifth, the cluster representatives are combined into a single FOM. We refer to it in this paper as the *training FOM*. The *optimal range* for μ is considered to be the maximal range within which the performance (as measured by the training FOM applied to the results obtained from the training data sets) is not statistically significantly different from the optimal performance.

Training

We first give the details of the training methodology for selecting the optimal range of μ_0. Preliminary tests indicated that we should not be looking outside the range [1.4, 2.4]. Within this range we investigated values of μ_0 at 0.1 increments. For each of this 11 values, ten complete actual (noisy) data sets were generated by the method described above and a training FOM was produced based on these 110 reconstructions. We found that the corresponding optimal range for μ_0 is [1.7,1.9].

To select μ_1, we essentially repeat this process, but now the evaluation is based on the reconstructions produced from $\mathbf{g}^2$. (For each of the 110 data sets, μ_0 was randomly selected from its optimal range.) The results were similar, namely the optimal range for μ_1 turned out to be the same as for μ_0. In fact, repeating this process for μ_k, for $2 \leq k \leq 15$, the same optimal range was found each time.

To determine the stopping criterion, 30 new complete actual data sets were generated and the IDR algorithm was run for 15 iterations. For each of the 30 actual data sets $\tilde{\mathbf{g}}$, reconstructions were produced from $\mathbf{g}^k$, for $0 \leq k \leq 15$. The FOM *scL2* was calculated for each reconstruction and, for $1 \leq k \leq 15$, the differ-

ence between *scL*2 values for that iteration and the previous iteration (from the same data set) was calculated. The results, plotted in Figure 6, show that there appear to be no significant changes produced by the IDR iterations beyond the sixth one. In this case, the *scL*2 value to which the algorithm converged was 0.998.

4 Results

Figure 7 illustrates our results: four central slices of the phantom and of reconstructions (from a new data set generated by the previously described rules) are shown. In Figure 8 we plot the associated Fourier Shell Correlations (see Section 3) for assessing the reconstruction quality. The improvement by any of the corrections over the uncorrected reconstruction is highly noticeable, even in the case of only phase flipping. Note that using only the FOM "resolution" (as defined in Section 3) there is not a significant difference between phase flipping and IDR (regardless of the number of iterations), since all the CTF corrected reconstructions have a resolution at around 0.06 Å^{-1} (as opposed to the resolution of the reconstruction from uncorrected data that is around 0.035 Å^{-1}). However, examining the full FSC curves instead of concentrating only on resolution reveals a clear improvement in the IDR-corrected reconstructions over the phase-flipped reconstruction in the range in which the CTF is inverted (between 0.037 Å^{-1} and 0.052 Å^{-1}). Furthermore, we can see the improvements produced by additional IDR iterations. All this is visually confirmed in Figure 7.

The discussion in the previous paragraph is anecdotal: it is based on recon-

structions from a single data set. To be able to assign statistical significance to our claim of superiority of the IDR reconstruction over the phase-flipped reconstruction we generated thirty additional actual data sets (using the same rules as before) and compared the phase-flipped reconstructions with the reconstructions after the sixth iteration of the IDR algorithm. The results are summarized in Figure 9, which plots for nineteen shells the average (over the thirty data sets) of the FSC value for IDR after six iterations less the FSC value for phase flipping. Standard deviations of these differences over the thirty experimental outcomes are also indicated. The standard deviation of the average difference is $1/\sqrt{30}$ times the standard deviation indicated in the figure. Hence it is clear that, for every one of the nineteen shells, one can extremely confidently reject the null hypothesis that the expected value of the FSC for the IDR reconstruction after six iteration is the same as the expected value of the FSC for the phase-flipped reconstruction in favor of the alternative hypothesis that the expected value of the FSC is higher for the IDR reconstruction. In fact, for each of the nineteen shells, the value of the average difference is more than ten times the standard deviation of the average difference (thus providing us with a P value less than 10^{-23}).

IDR's ability of removing the effect of an instrumental transfer function was first demonstrated by Ro et al. (1989) for magnetic resonance imaging. We have shown here that IDR is also efficacious in electron microscopy, which presents much nastier transfer functions and extremely noisy images. Our results demonstrate that it is possible to recover much of the information that is lost near the zeros of the CTF and that the quality of the reconstruction from electron micro-

scopic data can be significantly improved by iterative data refinement.

Acknowledgments

Partial support is acknowledged to the Comisión Interministerial de Ciencia y Tecnología of Spain through projects BIO98-0761 and BIO2001-1237 and to National Institutes of Health through grant HL70472. The work of Y. Censor was done in part at the Center for Computational Mathematics and Scientific Computation (CCMSC) at the University of Haifa and supported by Research Grant 592/00 from the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

The authors would like to thank Dr. J. J. Fernández for fruitful discussions and advice on the manuscript.

References

Bárcena, M., Donate, L., Ruiz, T., Dixon, N., Radermacher, M., Carazo, J. M., 2001. The Dnab-Dnac complex: A structure based on interactions among assymetric dimers. EMBO Journal 20, 1462–1468.

Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I., Bourne, P., 2000. The protein data bank. Nucleic Acids Research 28, 235–242.

Böhmer, K., Hemker, P., Stetter, H., 1984. The defect correction approach. Computing Suppl. 5, 1–32.

Censor, Y., Elfving, T., Herman, G. T., 1985. A method of iterative data refinement and its applications. Mathematical Methods in the Applied Sciences 7, 108–123.

Censor, Y., Zenios, S. A., 1997. Parallel Optimization: Theory, Algorithms, and Applications. Oxford University Press, New York, NY, USA.

Eggermont, P. P. B., Herman, G. T., Lent, A., 1981. Iterative algorithms for large partitioned linear systems with applications to image reconstruction. Linear Algebra Appl. 40, 37–67.

Frank, J., 1996. Three dimensional electron microscopy of macromolecular assemblies. Academic Press, San Diego, CA.

Frank, J., Penczek, P., 1995. On the correction of the contrast transfer function in biological electron microscopy. Optik 98, 125–9.

Furuie, S. S., Herman, G. T., Narayan, T. K., Kinahan, P. E., Karp, J. S., Lewitt, R. M., Matej, S., 1994. A methodology for testing for statistically significant differences between fully 3D PET reconstruction algorithms. Physics in Medicine & Biology 39, 341–354.

Grigorieff, N., 1998. Three-dimensional structure of bovine NADH : Ubiquinone oxidoreductase (complex I) at 22 Å in ice. J. Molecular Biology 277, 1033–1046.

Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E., Down-

ing, K. H., 1990. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. J. Molecular Biology 213, 899–929.

Herman, G. T., 1980. Image Reconstruction from Projections: The Fundamentals of Computerized Tomography. Academic Press, New York.

Herman, G. T., 1989. Standard image recovery methods in the Iterative Data Refinement framework. Proc. SPIE 1199, 1501–1511.

Herman, G. T., Ro, D., 1990. Image recovery using iterative data refinement with relaxation. Optical Engineering 29, 513–523.

Lewitt, R. M., 1990. Multidimensional digital image representations using generalized Kaiser-Bessel window functions. J. Optical Society of America A 7, 1834–1846.

Lewitt, R. M., 1992. Alternatives to voxels for image representation in iterative reconstruction algorithms. Physics in Medicine & Biology 37, 705–716.

Losada, M., Navarro, R., 1998. Point spread function of the human eye obtained by a dual double-pass method. Pure Appl. Opt. 7, L7–L13.

Ludtke, S. J., Baldwin, P. R., Chiu, W., 1999. EMAN: Semiautomated software for high-resolution single-particle reconstructions. J. Structural Biology 128, 82–97.

Ludtke, S. J., Jakana, J., Song, J., Chuang, D. T., Chiu, W., 2001. A 11.5 Å single

particle reconstruction of GroEL using EMAN. J. Molecular Biology 314, 253–262.

Marabini, R., Herman, G. T., Carazo, J. M., 1998. 3D reconstruction in electron microscopy using ART with smooth spherically symmetric volume elements (blobs). Ultramicroscopy 72, 53–65.

Marabini, R., Rietzel, E., Schröder, R., Herman, G. T., Carazo, J. M., 1997. Three-dimensional reconstruction from reduced sets of very noisy images acquired following a single-axis tilt schema: Application of a new three-dimensional reconstruction algorithm and objective comparison with weighted backprojection. J. Structural Biology 120, 363–371.

Matej, S., Lewitt, R. M., 1995. Efficient 3D grids for image reconstruction using spherically-symmetric volume elements. IEEE Trans. Nuclear Science 42, 1361–1370.

Matej, S., Lewitt, R. M., 1996. Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. IEEE Trans. Medical Imaging 15, 68–78.

Natterer, F., Wübbeling, F., 2001. Mathematical methods in image reconstruction. SIAM, Philadelphia.

Ro, D. W., Herman, G. T., Joseph, P. M., 1989. Resolution enhancement of magnetic resonance images using an iterative data refinement technique. Proc. SPIE 1199, 952–962.

Skoglund, U., Ofverstedt, L. G., Burnett, R., Bricogne, G., 1996. Maximum-entropy three-dimensional reconstruction with deconvolution of the contrast transfer function: A test application with adenovirus. J. Structural Biology 117, 173–188.

Sorzano, C. O. S., 2002. Algoritmos iterativos de reconstrucción tridimensional en microscopía electrónica de transmisión. Ph.D. thesis, E.T.S.I. Telecomunicación (Univ. Politécnica de Madrid).

Sorzano, C. O. S., Marabini, R., Boisset, N., Rietzel, E., Schröder, R., Herman, G. T., Carazo, J. M., 2001. The effect of overabundant projection directions on 3D reconstruction algorithms. J. Structural Biology 133, 108–118.

Stark, H., Orlova, E. V., Rinke-Appel, J., Junke, N., Müller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R., van Heel, M., 1997. Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. Cell 88, 19–29.

Stetter, H., 1978. The defect correction principle and discretization methods. Numerische Mathematik 29, 425–443.

Velázquez-Muriel, J. A., Sorzano, C. O. S., Fernández, J. J., Carazo, J. M., 2003. A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjusting. Ultramicroscopy 96, 17–35.

Zhou, Z. H., Hardt, S., Wang, B., Sherman, M. B., Jakana, J., Chiu, W., 1996.

CTF determination of images of ice-embedded single particles using a graphics interface. J. Structural Biology 116, 216–222.

Zhu, J., Penczek, P. A., Schröder, R., Frank, J., 1997. Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedure and application to the 70S *Escherichia coli* ribosome. J. Structural Biology 118, 197–219.

Zubelli, J. P., Marabini, R., Sorzano, C. O. S., Herman, G. T., 2003. Three-dimensional reconstruction by Chahine's method from electron microscopic projections corrupted by instrumental aberrations. Inverse Problems 19, 933–949.

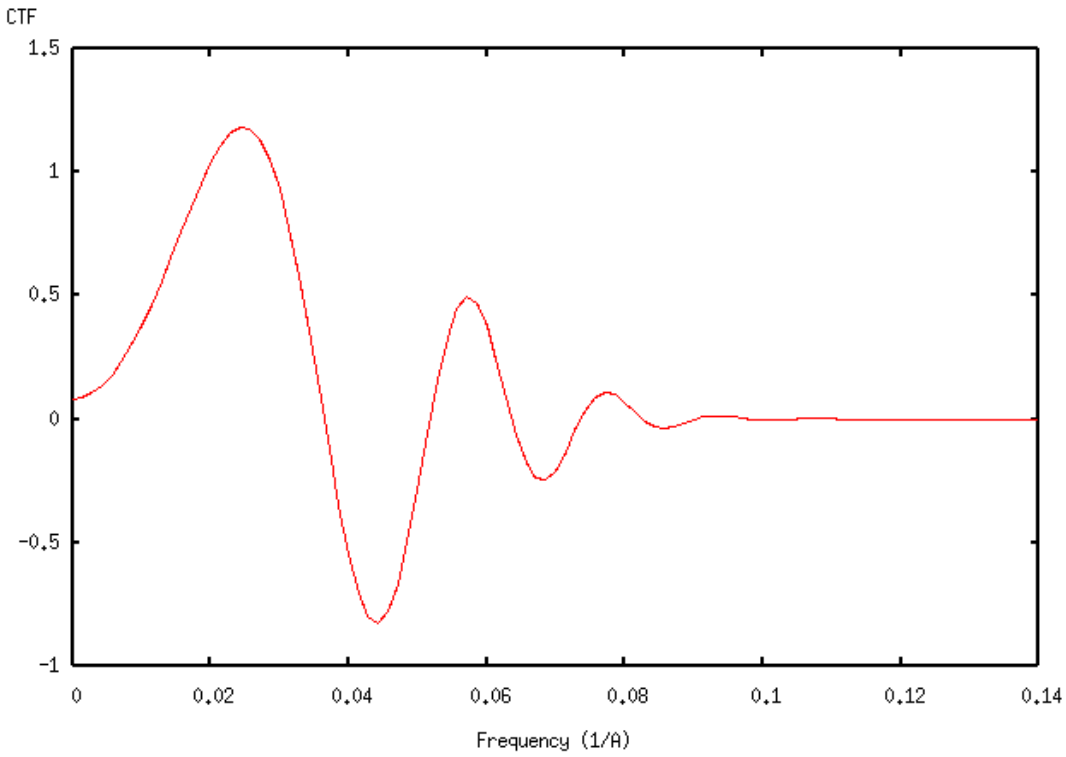

Figure 1: Radial profile of the CTF used in the cryomicroscopy simulations.

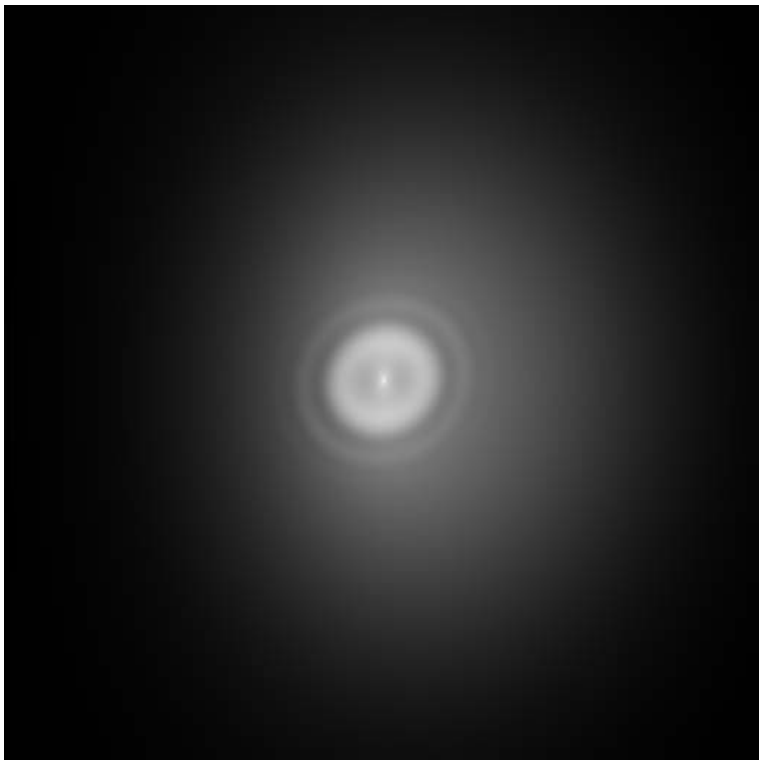

Figure 2: Amplitude of a typical astigmatic CTF.

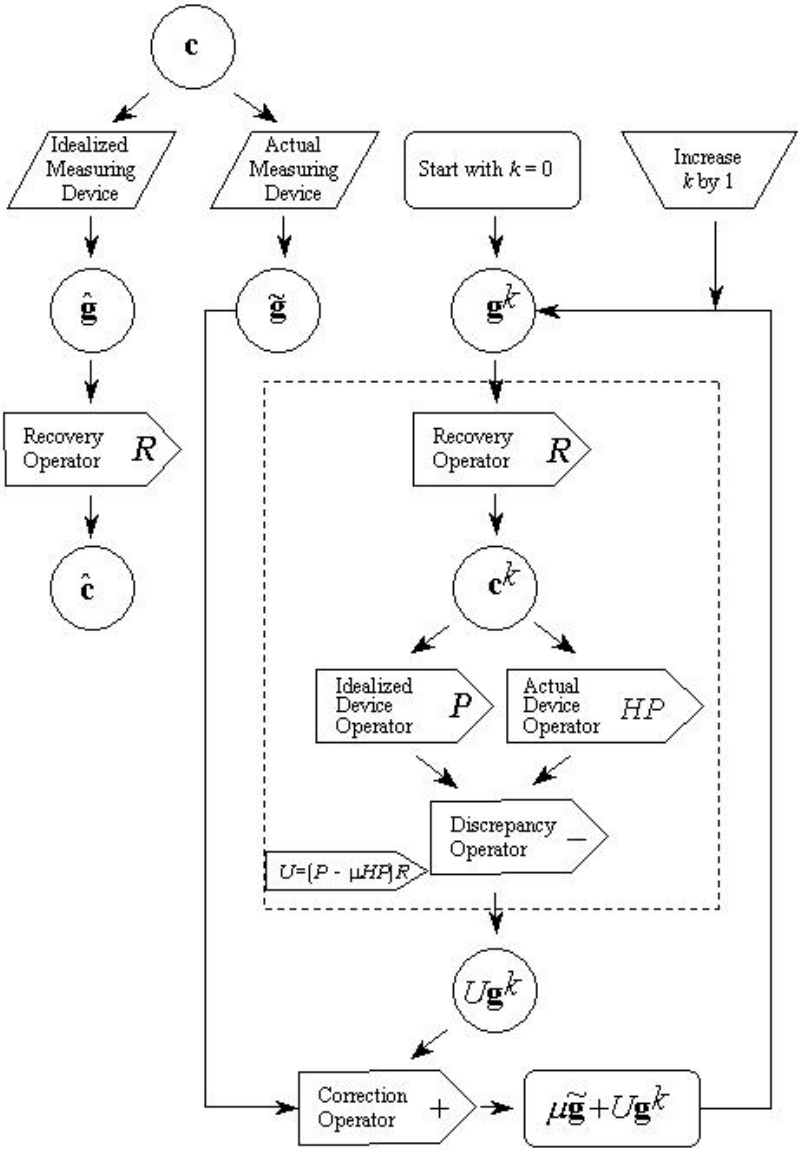

Figure 3: Block diagram for the IDR algorithm (based on the one originally published by Censor et al. (1985)).

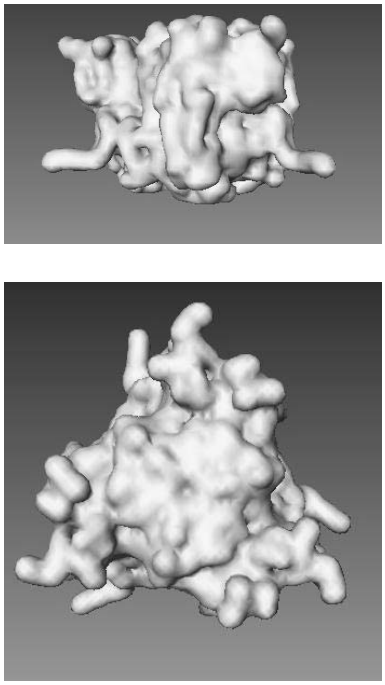

Figure 4: Side and top view of the isosurface of the bacteriorhodopsin phantom.

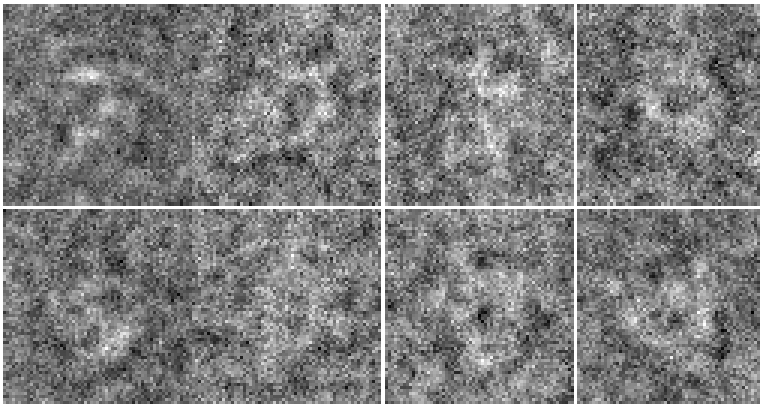

Figure 5: A selection of projections simulating cryomicroscopy images from the bacteriorhodopsin phantom.

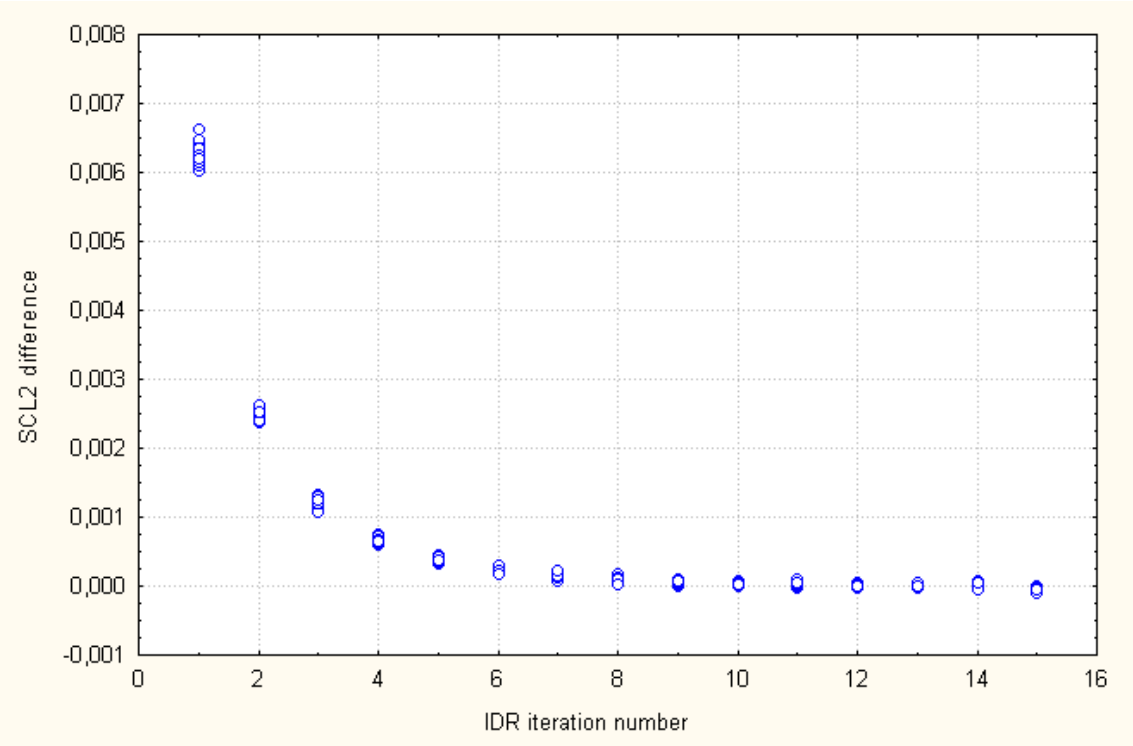

Figure 6: Difference in the scL2 FOM between the reconstruction after kth and $(k-1)$st IDR iterations for 30 tests.

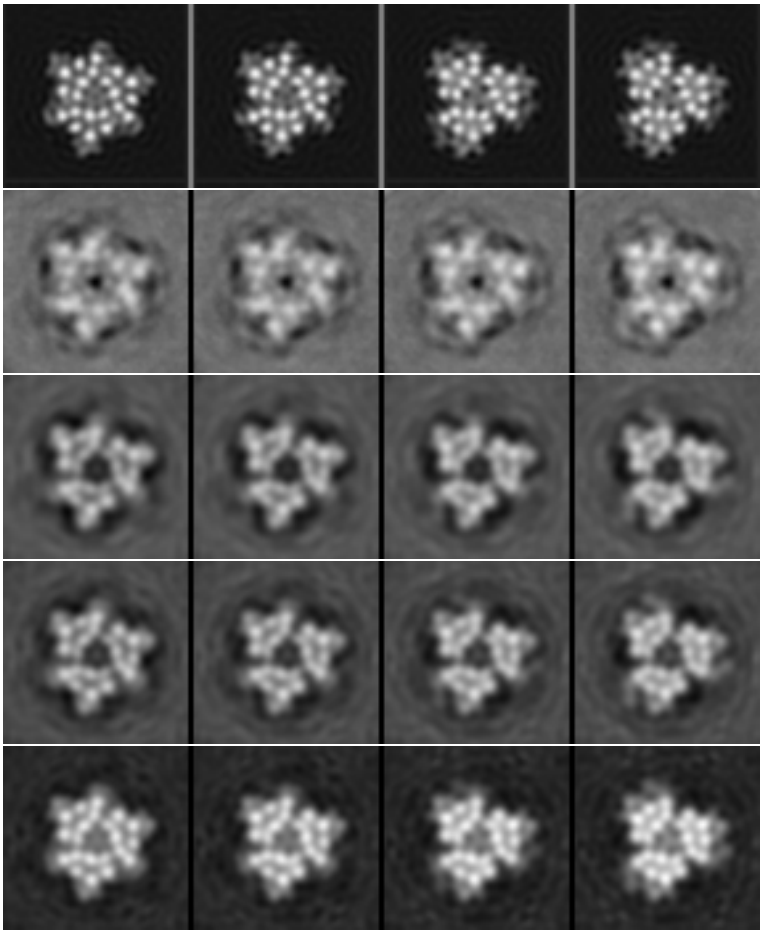

Figure 7: From top to bottom, slices corresponding to central sections of: bacteriorhodopsin phantom, reconstruction without CTF correction, reconstruction with phase flipping, reconstruction with IDR after 1 and 6 iterations.

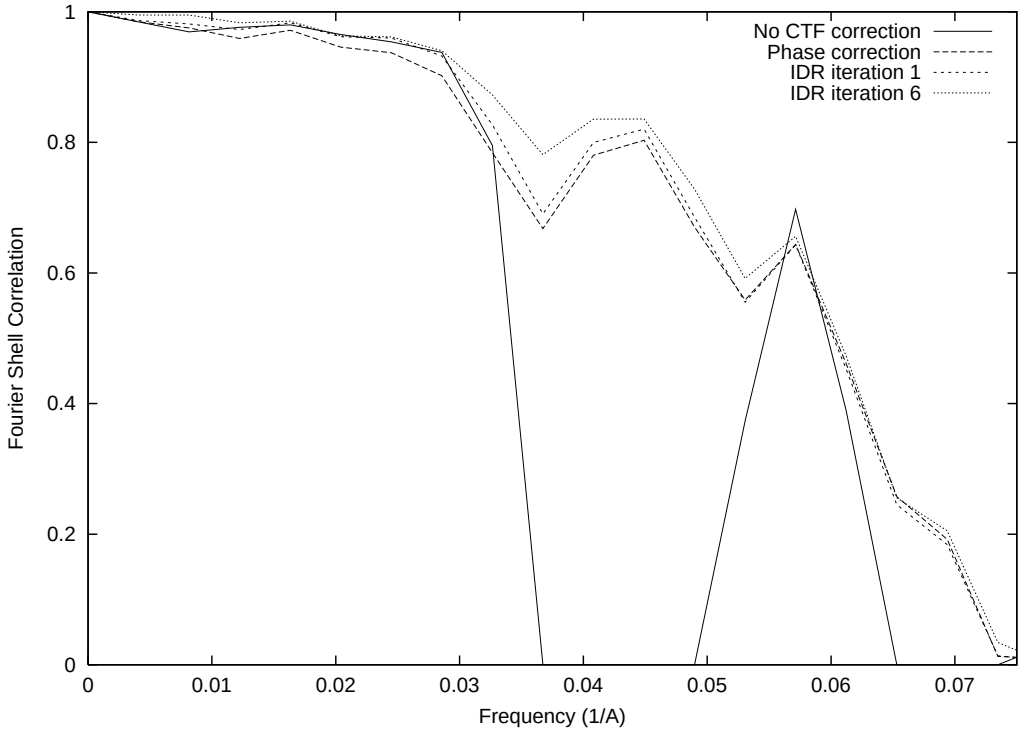

Figure 8: Fourier Shell Correlation curves for the reconstructions without CTF correction, with phase flipping and IDR after 1 and 6 iterations.

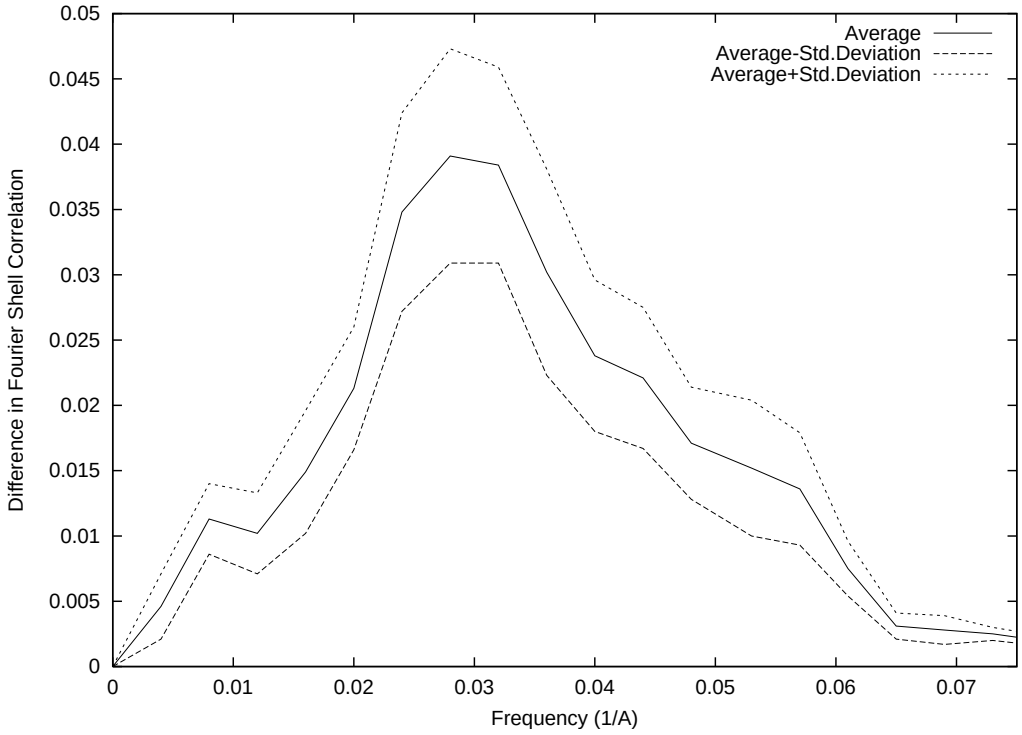

Figure 9: Plot of the values for 19 shells of the average over 30 actual data sets (plus and minus one standard deviation) of the FSC for IDR after 6 iterations less the FSC for phase flipping.