Three-Dimensional Reconstruction by Chahine's Method from Electron Microscopic Projections Corrupted by Instrumental Aberrations

J. P. Zubelli†§, R. Marabini‡§, C. O. S. Sorzano‡, and G. T. Herman¶

† zubelli@impa.br IMPA, Est. D. Castorina 110, RJ 22460-320, Brazil
‡ coss@cnb.uam.es CNB, Campus Universidad Autónoma s/n. 28049 Cantoblanco, Madrid, Spain.
§ roberto@cnb.uam.es EPS, Universidad Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain.
¶ GHerman@gc.cuny.edu Computer Science Program, Graduate Center-CUNY, 365 Fifth Av., New York, NY 10016-4309. USA.

Abstract. Three-dimensional reconstruction of nano-scale objects (such as biological macromolecules) can be accomplished using data recorded with a transmission electron microscope. An image obtained by a transmission electron microscope can be conceived of as an "ideal" projection subjected to a contrast transfer function (CTF), which attenuates most frequencies, reverses the phase of others, and even eliminates some of them. Such instrumental aberrations makes the problem of reconstruction from such data difficult. We reformulate the problem so that Chahine's method becomes applicable to it. We demonstrate the performance of our approach with numerical evidence using both simulated and actual electron microscopy data.
Keywords: Three-Dimensional Reconstruction, Electron Microscopy, CTF, Chahine's method.

AMS classification scheme numbers: 92C55, 35R30, 65F10

Submitted to: *Inverse Problems*

§ To whom correspondence should be addressed (zubelli@impa.br)

1. Introduction

Many important cellular functions are carried out by protein complexes that act as molecular machines. Three-dimensional electron microscopy (3D EM) is a powerful technique for imaging these complex biological macromolecules in order to understand their function. Unfortunately, the 3D EM reconstruction process cannot be made truly quantitative unless the aberrations introduced by the electron microscope are taken into account. We follow a model that is well accepted within the electron microscopy community [8, 28, 29, 31, 32] which incorporates the aberrations and assumes a linear relation between the data and the measurements as described below. This article presents a way of explicitly incorporating these aberrations into a version of the algorithm due to Chahine [2] for solving certain types of linear equation systems.

The microscope aberrations are described (in Fourier space) by the so-called *contrast transfer function* (CTF). The CTF filters both the high and the low frequencies, introduces zones of alternate contrast and eliminates all information at certain frequencies; see, for example, [32]. Figure 1 shows the radial profiles of three CTFs used in our simulations.

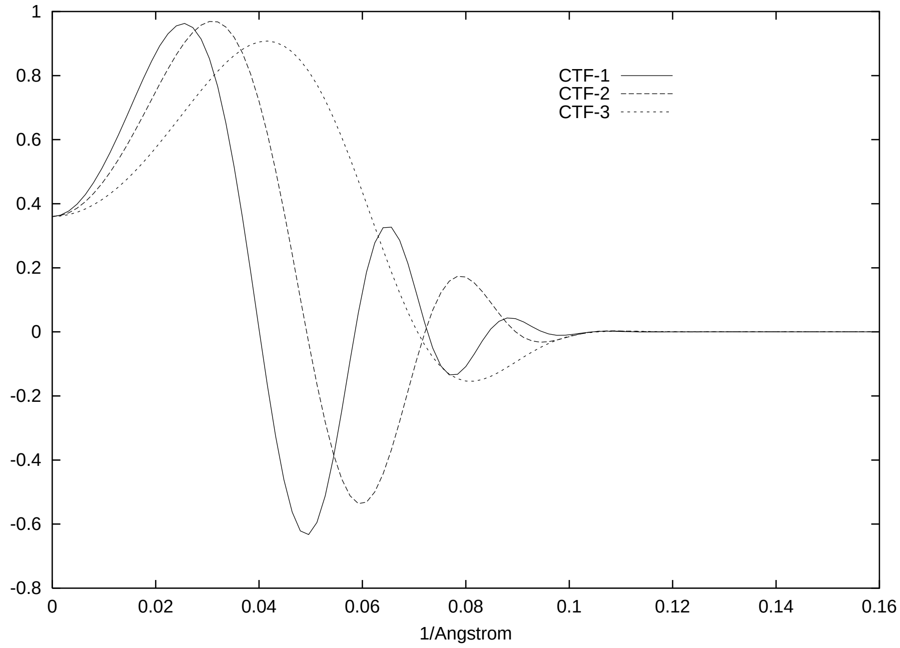

Figure 1. Radial profiles of the contrast transfer functions (CTFs) used in the experiments.

The model we consider [32] is the following: The measured data for the i-th projection y_i consist of the line integrals parallel to a chosen direction of the original compactly supported volume density $\rho : \mathbb{R}^3 \to \mathbb{R}_{\geq 0}$ followed by the convolution with the point spread function (PSF) h_i. The PSF h_i is the inverse Fourier transform of the CTF H_i. In practical applications, such projections y_i would be affected also by noise.

The problem is to reconstruct ρ from the collection of projections y_i, where

$i = 1, \cdots, I$. Here, by "to reconstruct" we mean to find a vector $c \in \mathbb{R}^J$ whose components are the coefficients of a sufficiently good approximation ρ_c to ρ in a finite dimensional space generated by suitable basis functions. One example is the representation of ρ_c using a voxel (volume element) basis, in which case the coefficients should be near to the averages of ρ over the corresponding voxels. Following [16] we do not use voxels as the basis functions, but spherical symmetric functions which are not only spatially limited but can also be chosen to be smooth. They are the generalized Kaiser-Bessel window functions, also known as *blobs*. The property of blobs that makes them very promising for the formulation of many algorithms is that they are smooth both in the real and in the Fourier domain. These blobs are placed in space so that they are centered on the points of the so-called body-centered cubic grid (BCC, obtained by interlacing two cubic grids in such a way that the points of either grid are central to eight points forming a cube in the other grid). For a detailed description of blobs and of their placement in space see [15, 19, 20, 21, 22].

From now on, we consider all the objects as defined on discrete and finite domains. Hence, our convolutions will be discrete ones (defined using wrap-arounds) and our Fourier transforms will be discrete Fourier transforms. We write each (noiseless) projection y_i as

$$y_i = h_i * \mathcal{P}_i c \ , \tag{1}$$

where $\mathcal{P}_i$ is the so-called *projection operator* along the i-th direction. It maps the volume coefficients into the two-dimensional array p_i whose entries correspond to the line integrals of the density ρ_c along the set of lines in the i-th direction. We will use J to denote the number of basis functions (and hence the dimensionality of the vector c) and N to denote the number of parallel lines in one projection (and hence the dimensionality of the vectors h_i and y_i); consequently $\mathcal{P}_i$ is an $N \times J$ matrix of nonnegative entries (a fact of which we make use later). In a typical 3D EM experiment for single-particle reconstruction [9], we have a total of 10^3-10^5 projections and each vector y_i has of the order of 64^2-315^2 components. Furthermore, the signal to noise ratio of each measured y_i is typically less than 1, and can be as small as 1/3.

We denote by $\mathcal{H}_i$ the two-dimensional convolution operators (in our approach $N \times N$ matrices), defined for $z \in \mathbb{R}^N$, by

$$\mathcal{H}_i z = h_i * z \ , \ i = 1, \cdots, I. \tag{2}$$

For convenience we introduce the auxiliary block-matrices $\mathcal{H}$, $\mathcal{P}$, and the vector y, consisting of

$$\mathcal{H} \stackrel{\text{def}}{=} \text{diag}[\mathcal{H}_1, \cdots, \mathcal{H}_I], \quad \mathcal{P} \stackrel{\text{def}}{=} \begin{bmatrix} \mathcal{P}_1 \\ \vdots \\ \mathcal{P}_I \end{bmatrix}, \text{ and } \quad y \stackrel{\text{def}}{=} \begin{bmatrix} y_1 \\ \vdots \\ y_I \end{bmatrix} . \tag{3}$$

In the sequel we also make use of the (real) matrix $\mathcal{B} = \mathcal{P}^T$, which is known in the reconstruction literature as the (discrete) *back-projection* operator. The sum of all components of a given vector z will be denoted by $\overline{z}$.

In an ideal situation, which is far from reality in microscopy, the CTF would be identically one and there would be no noise. In this case, our problem would reduce to recovering c from the data

$$p \stackrel{\text{def}}{=} \mathcal{P}c \; . \tag{4}$$

This problem has attracted a significant amount of attention; see [13, 23, 24] and references therein. The presence of the CTF, however, introduces further difficulties. In the microscopy literature it was recently considered by several authors: Frank and Penczek [10] applied Wiener filtering in the three-dimensional space to the reconstructed volume; Zhu et al. [33] incorporated a three-dimensional PSF into the data model and used a regularized steepest-descent technique; Stark et al. [27] applied inverse CTF filtering to the reconstructed volume; Skoglund et al. [26] incorporated a two-dimensional CTF particular to each projection to the projection model in a maximum-entropy reconstruction algorithm; Grigorieff [12] provided a Fourier reconstruction algorithm in which the CTF for each projection is considered in a Wiener-like fashion; Ludtke et al. [17] 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. [18] added a Wiener filter to the weighting function defined in [17].

One natural approach is to form the normal equations

$$\mathcal{B}\mathcal{H}^T\mathcal{H}\mathcal{P}c = d \stackrel{\text{def}}{=} \mathcal{B}\mathcal{H}^T y \; . \tag{5}$$

We remark that the operations of projection and back-projection are relatively fast, and the convolution can be efficiently implemented by the Fast Fourier Transform (FFT), so in practice we do not need to store the (prohibitively large) matrix

$$\mathcal{A} \stackrel{\text{def}}{=} \mathcal{B}\mathcal{H}^T\mathcal{H}\mathcal{P} \; . \tag{6}$$

Due to its size and the non-negativity of ρ, one approach to solving (5) is to try an algorithm due to Chahine [2] that has been used in many applications including atmospheric science [1, 2] and medical imaging [3, 4]. This method, due to its convergence properties [5] and simplicity, has proved to be competitive when applied to positron emission tomography. We discuss further the choice of Chahine's method in Section 2. It turns out, as we shall see below, that not all the necessary conditions for the published convergence results of Chahine's method are satisfied by the formulation of our problem given in (5).

The first goal of the present work is to present an answer to the following question: Based on the knowledge of the CTFs and the data y, how to adapt the problem (5) so that Chahine's algorithm becomes applicable to it? The second goal is to show, by numerical examples, that such method indeed produces good results, at least when compared to no CTF correction or to a simple phase correction.

The plan for the paper is the following. In Section 2 we describe Chahine's method and give some background information on 3D EM. In Section 3 we present our answer to the question stated in the previous paragraph. In Section 4 we describe some of the implementation details. In Section 5 we discuss the numerical results. We close in Section 6 with some conclusions and suggestions for further research.

2. Previous Work and Background

We first describe Chahine's method for solving systems of linear equations of the form

$$Mx = b \ , \tag{7}$$

where $x, b \in \mathbb{R}^J$ and M is a square $J \times J$ matrix satisfying some further hypotheses that will be spelled out in the sequel. The solution method goes as follows.

Consider an initial guess (say a constant vector of ones) $x^{(0)}$ in the positive orthant $\mathbb{R}^J_{>0}$ and iterate:

For $k = 0, 1, 2, \cdots$ do:
$e^{(k)} = Mx^{(k)}$;
$x_j^{(k+1)} = x_j^{(k)} \cdot b_j / e_j^{(k)}, \quad j = 1, \cdots, J.$

Chahine's method has been used in the medical imaging literature where it is sometimes referred to as the Image Space Reconstruction Algorithm (ISRA) [3, 4]. In fact, it has been independently proposed by Gold and Scofield in 1960 for the solution of certain linear integral equations [11]; see [5] and references therein. A fairly strong result concerning the method is given by the following.

Theorem 1 (De Pierro [4]) *If M is a nonnegative definite matrix with positive diagonal elements and if the entries of b are all positive, then the Chahine iteration generates a sequence $\left(x^{(k)}\right)_{k=0,1,\cdots}$ that possesses a convergent subsequence. Furthermore, if M is nonsingular, then the sequence converges.*

It turns out [5] that the limit of any convergent subsequence is a solution of the symmetric linear complementarity problem

$$Mx - b \geq 0, \quad x \geq 0, \quad x^T(Mx - b) = 0 \ . \tag{8}$$

The latter are the Kuhn-Tucker conditions associated to minimizing the quadratic functional

$$F(x) \stackrel{\text{def}}{=} \frac{1}{2} x^T M x - x^T b \tag{9}$$

subject to $x \geq 0$. The conditions in (8) are necessary and sufficient for x to be a minimizer of (9) subject to $x \geq 0$.

Other results are known [1, 4, 5, 6, 7]. In particular, P. Eggermont [6] strengthened Theorem 1 showing convergence even without the condition that M is nonsingular. In any case, the known results seem to rely on the hypotheses that M is nonnegative definite with positive diagonal elements and that all the entries of b are positive. The simplicity of the method and its demonstrated effectiveness for a number of large-scale problems plagued by low signal-to-noise ratio makes it an attractive candidate for our three-dimensional electron microscopy problem. Indeed, the number of operations necessary to compute the matrix $\mathcal{A}$ applied to a vector c representing the volume's density is surprisingly small. All that needs to be done is to compute projections of the current estimate (for blobs this can be efficiently performed by "footprint" techniques [22]), a fast Fourier transform (FFT), a product by the squared CTF, an inverse FFT, and a backprojection (that can also be done using footprints). To complete an iteration, this is followed by one multiplication and one division per component of the unknown c. Furthermore, for some bases (for example, the pixel basis) the volume has a *nonnegative* coefficient expansion. This means that the assumption that $c \geq 0$ is naturally satisfied. For other bases, such as blobs, this corresponds to a regularization of the problem since it imposes smoothness on the approximate ρ. For the applicability of Chahine's method we need to discuss whether or not the above mentioned hypotheses are satisfied.

We start with some technical assumptions:

(i) The PSFs are even functions (or equivalently the CTFs are real functions). This implies that the adjoint to the convolution operator $\mathcal{H}$ is itself, and so we may omit the super-index T in $\mathcal{H}^T$.

(ii) Every voxel (or blob) in the decomposition of image is intersected by at least one of the projecting rays.

(iii) For every i, $\overline{h}_i \geq 0$, (recall that $\overline{h}_i$ denotes the sum of all components of h_i). This assumption is harmless because if $\overline{h}_i < 0$, then we can replace our data y_i by $-y_i$ and proceed with h_i replaced by $-h_i$.

It is easy to check that the matrix $\mathcal{A}$ in (6) is nonnegative definite since the h_i's are real and even and $\mathcal{B} = \mathcal{P}^T$. The diagonal entries of $\mathcal{A}$ are positive, except for degenerate cases. Indeed, each entry $\mathcal{A}_{jj}$ corresponds to the ℓ^2 norm of the j-th column of $\mathcal{HP}$. Such ℓ^2 norm would be zero only if, for all i, the zero set of the i-th CTF would contain all the nonzero points of the Fourier transform of $\mathcal{P}_i e_j$, where e_j is the vector of zeros except for a one in the j-th position. Though possible, this is unlikely. Furthermore, if the number of projections is large, the CTFs are "mildly" behaved, and a good angular coverage is available, then one would expect $\mathcal{A}$ to be invertible (although we cannot say much a priori about its conditioning).

However, in the case of 3D EM in the presence of the CTFs, a major difficulty arises: the PSFs h_i change sign. Furthermore, their Fourier transforms have zeros and decay at infinity. The change in sign of the PSFs may cause the components of the vector d to change sign. Because of the decay of the CTFs, the matrix $\mathcal{A}$ may be singular or nearly singular (and thus highly ill-conditioned). Indeed, our first attempts to apply directly Chahine's method to the 3D EM problem with realistic CTFs did not lead to acceptable results.

3. Methodology

We now discuss a technique to circumvent the above problems in the case of 3D EM or in any other field where a similar data formation model holds. Our idea is to pre-process the data in a way that is equivalent to replacing the PSFs by some *virtual* PSFs. This pre-processing consists of adding a suitable multiple of a constant-component vector to each projection. For notational convenience take

$$\mathbf{1} \stackrel{\text{def}}{=} [1, 1, \cdots, 1]^T \in \mathbb{R}^N \ .$$

We propose the following idea: Add the vector $v_i \stackrel{\text{def}}{=} \lambda_i \mathbf{1}$ to y_i before and after it is convolved with h_i, where λ_i will be suitably chosen. We set v to be the vector

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_I \end{bmatrix} \ .$$

We shall show that the above mentioned pre-processing corresponds to solving a modification of problem (5) with a different set of PSFs and the right-hand side (RHS) vector

$$\widetilde{d} = \mathcal{B}\left(\mathcal{H}(y+v)+v\right) \ . \tag{10}$$

Note that for any vector $z \in \mathbb{R}^N$,

$$z * \mathbf{1} = \overline{z}\mathbf{1} \ . \tag{11}$$

The following claims are fundamental to the sequel.

Claim 1 *If we choose the numbers λ_i, $i = 1, \cdots, I$, appropriately, then all components of the vector $\widetilde{d}$ are positive.*

Proof: Fix $i \in \{1, \cdots, I\}$. Because of (11), $h_i * v_i = \overline{h}_i v_i$. So,

$$h_i * (y_i + v_i) + v_i = h_i * y_i + (\overline{h}_i + 1) v_i \ .$$

Now we choose λ_i so that

$$\lambda_i(\overline{h}_i + 1) > -\min_l \{0, (h_i * y_i)_l\} \ . \tag{12}$$

Since the back-projection matrix has only nonnegative entries, the resulting vector $\widetilde{d}$ has only nonnegative entries. Furthermore, none of the entries in $\widetilde{d}$ can be zero since by the choice of the λ_i's all the entries of $\widetilde{y} \stackrel{\text{def}}{=} (\mathcal{H}(y+v)+v)$ are positive. Therefore, the only way an entry of $\mathcal{B}\widetilde{y}$ would be zero is if the corresponding row of $\mathcal{B}$ vanishes. But this contradicts the second of our technical assumptions in Section 2 since it would correspond to a column of zeros in $\mathcal{P}$, in other words a voxel (or blob) which is not intersected by any of the line integrals that produce the totality of projections. *Q. E. D.*

Our next goal is to show that the pre-processing in Claim 1 leads to a set of virtual PSFs of the form

$$\widetilde{h}_i = h_i + \alpha_i v_i \ , \tag{13}$$

where the constants α_i will be chosen appropriately. Such a virtual PSF may not be realizable in a real microscope, but it leads to a set of equations that is amenable to Chahine's algorithm, as we shall see below. Associated with these virtual PSFs we define the operators

$$\widetilde{\mathcal{H}}_i z \stackrel{\text{def}}{=} \widetilde{h}_i * z \quad \text{and} \quad \widetilde{\mathcal{H}} \stackrel{\text{def}}{=} \operatorname{diag}[\widetilde{\mathcal{H}}_1, \cdots, \widetilde{\mathcal{H}}_I] \ . \tag{14}$$

Claim 2 *Let $c \in \mathbb{R}^J_{\geq 0} \setminus \{0\}$ be such that $\mathcal{A}c = d$ and suppose that $y = \mathcal{H}\mathcal{P}c$. Take $\widetilde{d}$ as in (10) with λ_i satisfying (12), for $i = 1, \cdots, I$. Then, there exists $\alpha = (\alpha_1, \cdots, \alpha_I) \in \mathbb{R}^I$ such that c is a solution of $\widetilde{\mathcal{A}}c = \widetilde{d}$ where $\widetilde{\mathcal{A}} = \mathcal{B}\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}\mathcal{P}$, and $\widetilde{\mathcal{H}}$ is associated with the PSFs $\widetilde{h}_i$ as defined in (13) and (14).*

Proof: For $i = 1, \cdots, I$, let $p_i = \mathcal{P}_i c$, and consequently $\overline{p}_i > 0$. We will show how to choose the α_i's so that $\mathcal{B}\widetilde{\mathcal{H}}\widetilde{\mathcal{H}}\mathcal{P}c = \widetilde{d}$. By (10) it is sufficient to prove that, for $i = 1, \cdots, I$,

$$\widetilde{\mathcal{H}}_i \widetilde{\mathcal{H}}_i \mathcal{P}_i c = \mathcal{H}_i(y_i + v_i) + v_i \ ; \tag{15}$$

or, using equations (13), (14), and the assumption that $y = \mathcal{H}\mathcal{P}c$, that

$$(h_i + \alpha_i\lambda_i\mathbf{1}) * (h_i + \alpha_i\lambda_i\mathbf{1}) * p_i = h_i * (h_i * p_i + \lambda_i\mathbf{1}) + \lambda_i\mathbf{1} \ . \tag{16}$$

Using that for $\mathbf{1} \in \mathbb{R}^N$ we have $\mathbf{1} * \mathbf{1} = N\mathbf{1}$ and Equation (11), it is easy to check that Equation (16) is a consequence of

$$(\overline{p}_i N)(\alpha_i\lambda_i)^2 + 2\overline{h}_i\overline{p}_i(\alpha_i\lambda_i) - \lambda_i(\overline{h}_i + 1) = 0 \ . \tag{17}$$

We must have by (12) and our third technical assumption that $\overline{h}_i \geq 0$ that $\lambda_i > 0$. Consequently, we can find the following real-valued solutions α_i of Equation (17). If $\overline{h}_i = 0$, then we can take $\alpha_i = \pm 1/\sqrt{\lambda_i \overline{p}_i N}$. If $\overline{h}_i \neq 0$, then using the fact that $\overline{y}_i = \overline{h}_i\overline{p}_i$, we can take,

$$\alpha_i = \frac{\overline{h}_i}{\lambda_i N}\left(\pm\sqrt{1 + \frac{N(\overline{h}_i + 1)\lambda_i}{\overline{h}_i\overline{y}_i}} - 1\right) \ . \tag{18}$$

This concludes the proof. *Q. E. D.*

Remark 1 *For each i, if $\overline{h}_i \neq 0$, then α_i can be chosen to be a nonnegative number that depends only on $\overline{h}_i$, $\overline{y}_i$, and N. If $\overline{h}_i = 0$, then we can take $\alpha_i = \pm 1/\sqrt{\lambda_i \overline{p}_i N}$.*

Remark 2 *The technical assumption (iii) of Section 2 gives us that for each $i = 1, \cdots, I$, the DC component of the PSF h_i is nonnegative. For reasons of numerical stability it is sound to choose the a positive value of α_i in our computations, so as to avoid loss of significant digits in the computations. In this case a more numerically sound way of computing α_i is by the expression*

$$\alpha_i = \frac{\left[(\overline{h}_i + 1)/\overline{y}_i\right]}{1 + \left(1 + \frac{\lambda_i\left[(\overline{h}_i+1)/\overline{y}_i\right]}{\overline{h}_i/N}\right)^{1/2}} \ . \tag{19}$$

From now on, we take α_i to be the positive root of Equation (17).

It would be desirable for any solution to our modified equations $\widetilde{\mathcal{A}}c = \widetilde{d}$ to satisfy the original equations $\mathcal{A}c = d$. This however, may not be the case. We have instead the following weaker result.

Claim 3 *Let $\widetilde{c}$ be a solution of $\widetilde{\mathcal{A}}\widetilde{c} = \widetilde{d}$, where α_i and $\widetilde{h}_i$ are computed according to (19) and (13), for $i = 1, \cdots, I$, and $\widetilde{d}$ is computed by (10). Then, $\widetilde{c}$ satisfies*

$$\mathcal{A}\widetilde{c} = d + \sum_{i=1}^{I} r_i \mathcal{P}_i^T \mathbf{1} \ , \tag{20}$$

where

$$r_i = (1 + \overline{h}_i)\lambda_i - \overline{\mathcal{P}_i \widetilde{c}} \left(N(\alpha_i \lambda_i)^2 + 2\overline{h}_i(\alpha_i \lambda_i) \right) \ . \tag{21}$$

Proof: If $\widetilde{c}$ satisfies $\widetilde{\mathcal{A}}\widetilde{c} = \widetilde{d}$, then by (5) we get that

$$\sum_{i=1}^{I} \mathcal{P}_i^T \left((h_i + \alpha_i \lambda_i \mathbf{1}) * (h_i + \alpha_i \lambda_i \mathbf{1}) * \mathcal{P}_i \widetilde{c} \right) = d + \sum_{i=1}^{I} \mathcal{P}_i^T \left(h_i * v_i + v_i \right) \ . \tag{22}$$

Expanding the product $(h_i + \alpha_i \lambda_i \mathbf{1}) * (h_i + \alpha_i \lambda_i \mathbf{1})$, and using that $\mathbf{1} * \mathbf{1} = N\mathbf{1}$, we have that (20) and (21) follow. *Q. E. D.*

In our numerical experiments, we found that the norm of $\sum_{i=1}^{I} r_i \mathcal{P}_i^T \mathbf{1}$ was small as compared to the norm of d, and so in practice the solutions of $\widetilde{\mathcal{A}}\widetilde{c} = \widetilde{d}$ yielded "near-solutions" of $\mathcal{A}c = d$.

4. Practical Implementation

In view of Claims 1 – 3 we transform our problem of seeking a solution to Equation (5) into searching for a solution to

$$\widetilde{\mathcal{A}}c = \widetilde{d} \ , \tag{23}$$

where now $\widetilde{d}$ has positive components and $\widetilde{\mathcal{A}}$ is a nonnegative definite matrix with positive diagonal entries. Indeed, $\widetilde{d}$ has positive components because of Claim 1. Furthermore, $\widetilde{\mathcal{A}}$ is a nonnegative definite matrix since $\widetilde{\mathcal{H}}$ is symmetric and thus we can write $\widetilde{\mathcal{A}} = (\widetilde{\mathcal{H}}\mathcal{P})^T \widetilde{\mathcal{H}}\mathcal{P}$. Finally, the diagonal entries of $\widetilde{\mathcal{A}}$ are positive because each entry $\mathcal{A}_{jj}$ corresponds to the ℓ^2 norm of the j-th column of $\widetilde{\mathcal{H}}\mathcal{P}$. Such ℓ^2 norm is not zero. To see this claim, notice that it is the sum for $i = 1, \cdots, I$ of the ℓ^2 norm of $h_i * \mathcal{P}_i e_j$. By Plancherel's theorem this is a (positive) multiple of the ℓ^2 norm of $\widetilde{\mathsf{H}}_i \mathcal{F}(\mathcal{P}_i e_j)$, where $\widetilde{\mathsf{H}}_i$ is the Fourier transform of the virtual PSF $\widetilde{h}_i$. The claim follows now because the DC components of all the *virtual* PSFs are nonzero.

Thus all the hypotheses in Theorem 1 hold and Chahine's method is applicable to our reformulated problem.

The above remarks lead to the following practical implementation of Chahine's method in 3D EM:

Given the point spread function h_i, and the data y_i, $i = 1, \cdots, I$.

(i) For each $i = 1, \cdots, I$, compute $\overline{h}_i$ and $\overline{y}_i$, and make sure (by changing (h_i, y_i) into $(-h_i, -y_i)$, if necessary) that they are both nonnegative. (See Remark 3 below.)

(ii) Choose λ_i so that Equation (12) holds for $i = 1, \cdots, I$.

(iii) For $i = 1, \cdots, I$, determine $\widetilde{h}_i$ by computing α_i according to Claim 2 and Equation (19). (Equivalently, add to the DC value of the i-th CTF the value of $N\alpha_i$.)

(iv) Apply Chahine's method to solve the equation

$$\widetilde{\mathcal{A}}c = \widetilde{d}\ ,$$

where $\widetilde{d}$ is given by (10).

Remark 3 *The effect of noise, mismatch on the estimation of the angles, and on the estimation of the CTF, can lead to further corruption of the data y_i, $i = 1 \cdots, I$. One problem that could arise is that the averages $\overline{y_i}$ and $\overline{h}_i$ may have different signs, which could lead to complex roots in equation (19). In such case, the matrix $\widetilde{\mathcal{A}}$ may not satisfy the hypotheses of Theorem 1 and Chahine's method would no longer be applicable. This cannot happen in Claim 2 itself, because one of its hypotheses is that $y = \mathcal{HP}c$, which implies that the data are perfect except for the CTF aberration. One practical way of dealing with this is by adding a constant volume to the sought value c. To wit, we choose a $\mu \in \mathbb{R}$ so that the data satisfy the condition that $\operatorname{sgn}(\overline{y_i} + \mu\overline{\mathcal{H}_i\mathcal{P}_i\mathbf{1}}) = \operatorname{sgn}(\overline{h}_i)$, for $i \le i \le I$.*

5. Numerical Results

In this section we describe some of the numerical experiments that we used to validate the technique described in Section 4. We compared our results with two implementations of the algebraic reconstruction techniques (ART); see [13]. They are both based on the *block-ART* variant of ART, which can be described as follows.

Starting from a J-dimensional zero vector as an estimate of c, we update this estimate iteratively. In an iterative step we make use of data from one projection only; we repeatedly cycle through all the projections in the complete algorithm. The update of the estimate is done in a computationally efficient manner [22] so as to reduce the discrepancy between the measured data provided by the projection used in that iterative step and the matching *pseudo projection data* that is obtained from the current estimate of c. A more detailed description of this implementation, including convergence properties, is given in [14].

In all the implementations (both ART and Chahine's method) we used blobs as the basis functions. Following [19], we selected blobs with parameters: order m=2, shape $\alpha = 13.3633059$, radius $a = 3.394212$, and grid sample separation 2.0. The initial estimate of c is constant-valued (1-valued for Chahine's method and 0-valued for ART).

In one implementation of ART, we apply a simple CTF phase correction, and in the other we do not take the CTF into account at all. By CTF phase correction, we mean that the algorithm makes use of the sign of the CTF so that the processed measurements (i.e., the input to the algorithm) are obtained by the inverse Fourier transform of the Fourier transform of the original measurements multiplied by -1 whenever the CTF is negative. All of our reconstructions were implemented using the XMIPP software available at http://www.cnb.uam.es/~bioinfo/.

In the sequel, we show the evolution of the objective function

$$F(c) = \frac{1}{2} c^T \widetilde{\mathcal{A}} c - c^T \widetilde{d}\ . \tag{24}$$

Notice that for the special case $\widetilde{\mathcal{H}} = \mathcal{H}$ and $d = \widetilde{d}$, a simple computation gives that

$$\frac{1}{2} ||\mathcal{H}\mathcal{P}c - y||^2 = F(c) + \frac{1}{2}||y||^2\ .$$

$F(c)$ is a way of monitoring the evolution of the iterates. It is known that $F(c^{(k)})$ is a non-increasing sequence [4]. We will also show the evolution of the squared relative residual error

$$\eta^{(k)} \stackrel{\text{def}}{=} \frac{||e^{(k)} - \widetilde{d}||^2}{||\widetilde{d}||^2}\ , \tag{25}$$

where $e^{(k)} = \widetilde{\mathcal{A}} c^{(k)}$.

In the synthetic-data examples, we used a total of $I = 1000$ projections with uniformly distributed random angles and $N = 128^2$ pixels each. The CTFs used were randomly chosen from the three examples displayed in Figure 1. The number of blob coefficients was $J = 34327$.

In the real data experiments, we used a total of $I = 570$ projections, with $N = 128^2$ pixels each, and the the number of blobs $J = 130869$.

The gray-level scale of the displayed results of the reconstructions was adjusted so that the minimum value was mapped to 0 and the maximum level was mapped to 255. In the plots of $F(c^{(k)})$ and $\eta^{(k)}$ we do not display the values for $k = 0$, since they would change the scale of the graphs and reduce the distinguishability of the remaining ordinates.

The experiments were performed using the XMIPP software running under the Red Hat Linux 7.2 operating system, on an Intel Xeon CPU of 1.70GHz. In the synthetic data experiments, the typical (system) time for one iterative step was of the order of 10 minutes. In the real data experiments, the typical (system) time for one iterative step was of the order of 25 minutes. In typical 3D EM experiments, the specimen preparation and data gathering exceed such times by several orders of magnitude.

5.1. Synthetic Data: The Bacteriophage ϕ-29 Connector

We chose as a phantom for our simulations a rather realistic one. It was produced from the structure of the bacteriophage ϕ-29 connector. The atomic coordinates of this protein can be found in the "Protein Data Base" (http://www.pdb.org) under the PDB-ID: 1FOU. In our model, the atoms are replaced by small blobs. Such blobs, however have different parameters from the ones used in the reconstructions, more precisely the blob radius was 12, the shape 0.1, and the order 2. The surface of the resulting phantom is displayed in Figure 2.

Figure 2. Surface rendering of the three dimensional ϕ-29 phantom.

We first present the reconstruction by Chahine's method of the original phantom in the (unrealistic) noiseless and no-CTF case. As can be seen in Figures 3 and 4, the reconstruction and the original phantom are in excellent agreement. In Figure 4 we display the values of $F(c^{(k)})$ and $\eta^{(k)}$ for $k = 1, 2, \cdots$. The corresponding values for the initial guess are $F(c^{(0)}) = -1.067 \times 10^{+11}$ and $\eta^{(0)} = 0.8089$. This example of a comparison between Chahine's method and ART in the unrealistic noiseless and aberrationless case is representative of a large set of such comparisons that we have performed.

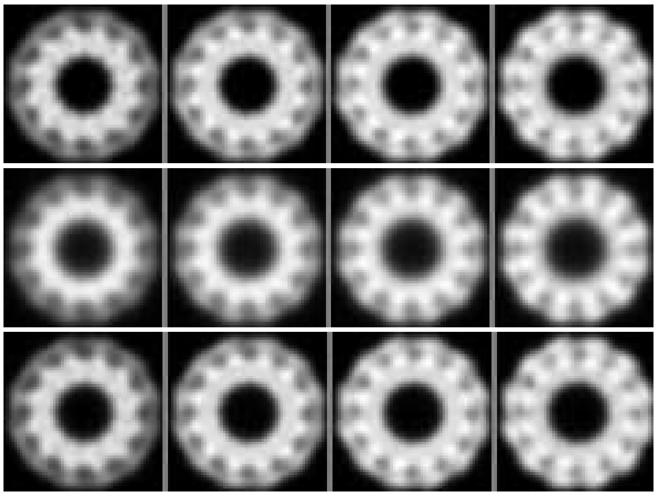

Figure 3. Central slices of the original ϕ-29 phantom (top), the reconstruction after 6 iterations (middle) and after 49 iterations (bottom), using Chahine's method.

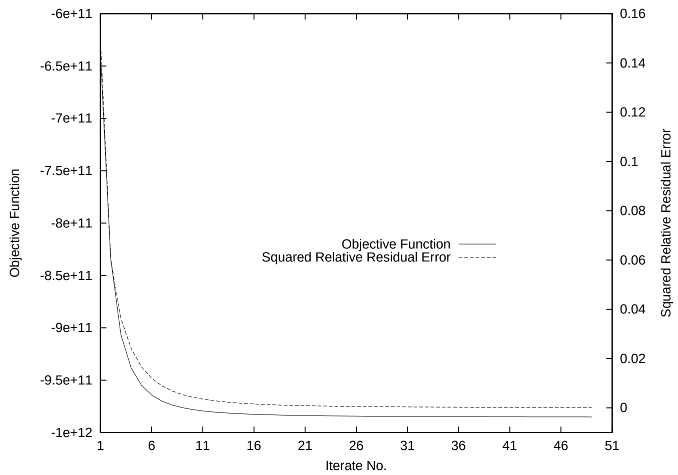

Figure 4. Evolution of the objective function $F(c^{(k)})$ of (24) and of the squared relative residual error $\eta^{(k)}$ of (25), with $k = 1, 2, \cdots$ for the ϕ-29 example by Chahine's method in the no-noise and no-CTF case.

Next we present reconstructions from noisy data in the presence of CTFs using three different methods: ART with no CTF correction, ART with CTF phase correction, and Chahine's method. The synthetic data were generated from projections from a random set of directions that was subjected to the CTFs described in Figure 1 and further corrupted by noise. The SNR is $1/3$ on the average. The relaxation parameter used for ART is 0.06. ART results are reported after two cycles through all the projections, while the results of Chahine's method are shown after the sixth iteration.

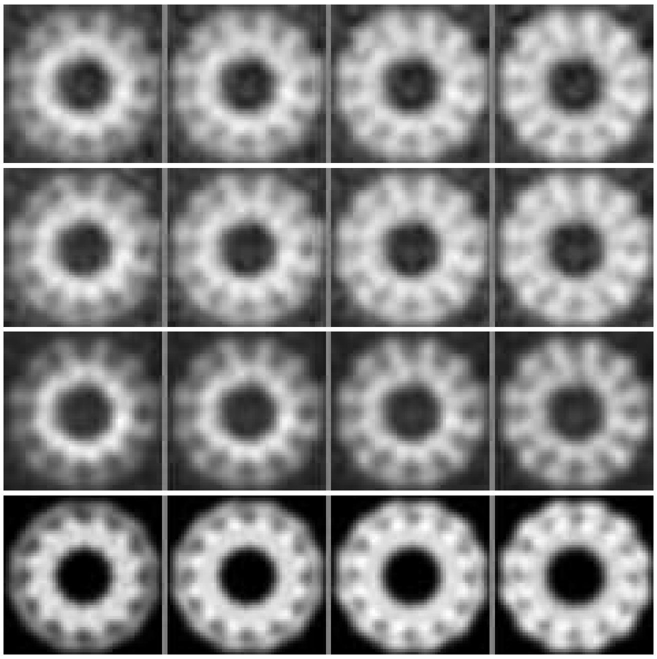

Figure 5. Central slices of the ϕ-29 phantom and its reconstructions from noisy data using ART, ART with phase correction, and the sixth iteration of Chahine's method. From top to bottom: ART, ART with phase correction, Chahine's method, phantom.

The evolution of the objective function and of the squared relative residual error in Chahine's method is presented in Figure 6. The corresponding values for the initial 1-valued estimate $c^{(0)}$ are $F(c^{(0)}) = -4.448 \times 10^{+11}$ and $\eta^{(0)} = 0.4229$, which means that most of the improvement was achieved in the first iteration.

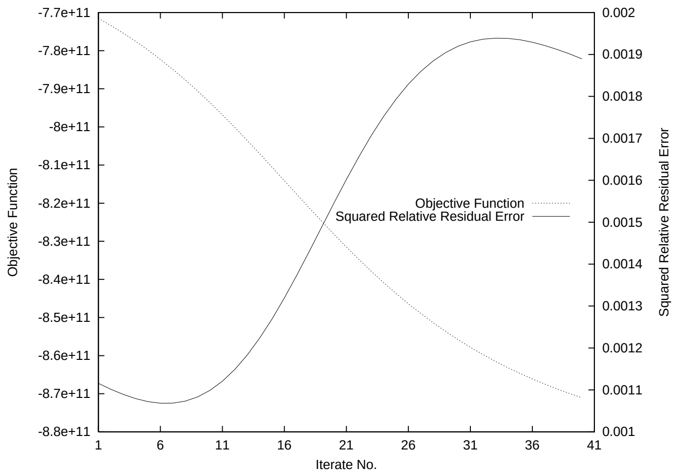

Figure 6. Evolution of the objective function $F(c^{(k)})$ and of the squared relative residual error $\eta^{(k)}$, with $k = 1, 2, \cdots$ for the noisy ϕ-29 example by Chahine's method.

In order to assess objectively the advantage of using Chahine's method over the other two algorithms, we present in Figure 7 a plot of the Fourier shell correlation between the phantom and the reconstructions. The *Fourier shell correlation* (FSC) between two volumes is defined as follows. Given the Fourier transform of the phantom $\mathcal{F}[p](\mathbf{R})$ and of the reconstruction $\mathcal{F}[r](\mathbf{R})$, the FSC at a given frequency $\mathbf{R}$ with a shell width $\Delta\mathbf{R}$ is

$$\mathrm{FSC}(\mathbf{R}, \Delta\mathbf{R}) = \frac{\sum_{\mathbf{R} \le \mathbf{R}' \le \mathbf{R}+\Delta\mathbf{R}} \mathcal{F}[p](\mathbf{R}')\mathcal{F}^*[r](\mathbf{R}')}{\left(\sum_{\mathbf{R} \le \mathbf{R}' \le \mathbf{R}+\Delta\mathbf{R}} |\mathcal{F}[p](\mathbf{R}')|^2\right)^{1/2} \left(\sum_{\mathbf{R} \le \mathbf{R}' \le \mathbf{R}+\Delta\mathbf{R}} |\mathcal{F}[r](\mathbf{R}')|^2\right)^{1/2}} .$$

See [30] for more information about the FSC.

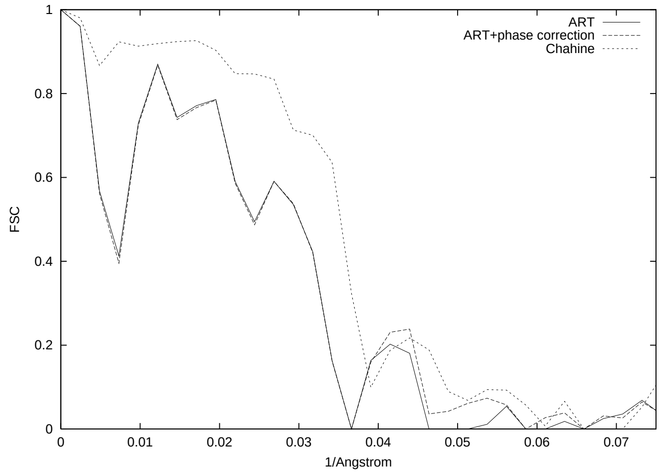

Figure 7. Correlation between the Fourier transform of the phantom and of the reconstruction as a function of frequency shell radius.

As can be seen in Figure 7, the correlation between the Chahine reconstruction and the phantom is better than the correlation between either of the ART-based reconstructions and the phantom for all the frequencies at which the phantom and the reconstructions are consistent. (The term “consistent” does not have a standard definition in the field of electron microscopy, but usually a reconstruction is considered consistent with a phantom at those frequencies at which the FSC is greater than 0.5.)

5.2. Real Data Reconstructions

We now present central slices of reconstructions using experimental data from cryomicroscopic projections of the Large T-antigen of the virus SV40. The angle information was obtained by *model based orientation search* [25] and the CTF information was obtained using the technique in [31]. We compare the reconstruction obtained by Chahine's method with the results obtained by ART with no CTF correction and with CTF phase correction. The results are reported in Figures 8 and 9.

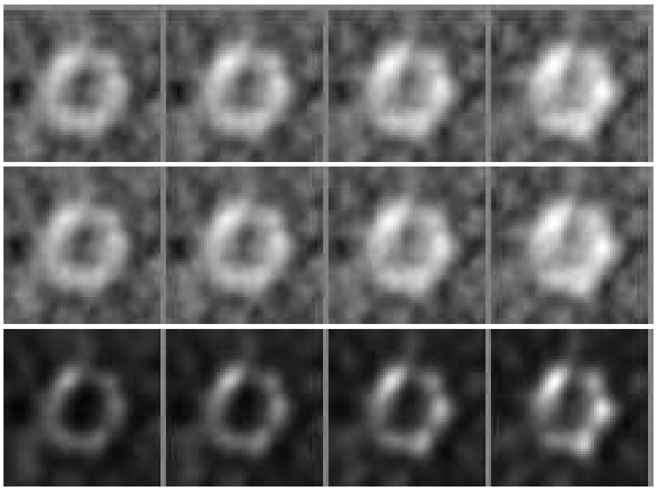

Figure 8. Central slices of reconstructions from experimental data. Top: ART reconstruction from uncorrected measurements. Middle: ART with CTF phase correction. Bottom: Chahine's method after 11 iterations.

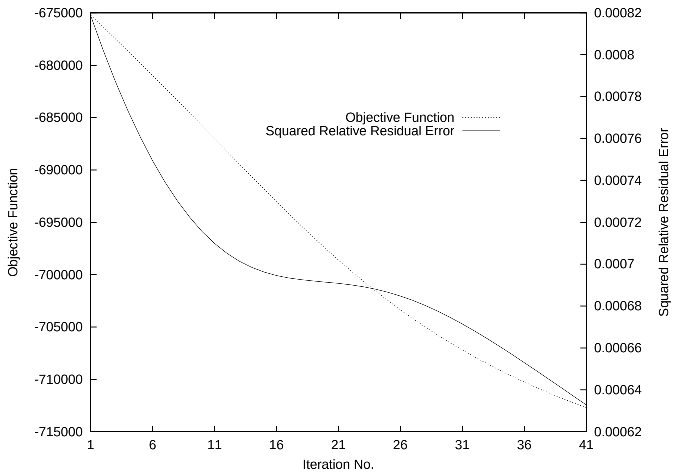

Figure 9. Evolution of the objective function $F(c^{(k)})$ and of the squared relative residual error $\eta^{(k)}$, with $k = 1, 2, \cdots 41$ for the real data reconstruction using Chahine's method. The corresponding values for the initial 1-valued estimate were $F(c^{(0)}) = -242114$ and $\eta^{(0)} = 0.641133$.

6. Conclusions and Suggestions for Further Research

We have shown that it is possible to replace the normal equations

$$\mathcal{A}c = \mathcal{B}\mathcal{H}^T y \quad \text{with} \quad \mathcal{A} = \mathcal{B}\mathcal{H}^T\mathcal{H}\mathcal{P}$$

by processing the data y in such a way that the resulting system

$$\widetilde{\mathcal{A}}c = \widetilde{d}$$

satisfies the hypotheses of Theorem 1, thus ensuring the convergence of the Chahine's algorithm iterates. The resulting method was implemented and was successfully tested with simulated as well as with experimental data. The reconstructed images that we presented here are typical of a much larger set of experiments we performed. Based on the experiments and on the FSC curves we have a strong indication that Chahine's algorithm is a promising technique for three-dimensional image reconstruction in electron microscopy when implemented with the data processing described in Section 3 and 4. Furthermore, the results in Figure 7 indicate that the correlation between the Chahine reconstruction and the phantom is better than the correlation between any of the ART based reconstructions for a substantial frequency range, including those for which the phantom and the reconstructions are consistent.

In future work, we plan address in detail the issue of regularization. We point our, however, that Chahine's method is naturally amenable to reguralization with a minimal additional cost. Indeed, following the standard procedure of regularizing the problem $\mathcal{A}c = b$ by means of finding $c^{(\epsilon)}$ such that $(\mathcal{A} + \epsilon I)c^{(\epsilon)} = b$, a moment's thought indicates that the iteration $c_j^{n+1} = b_j/(\epsilon + ((\mathcal{A}c^n)_j)/c_j^n))$ performs such regularization with only one additional operation per coefficient.

Concerning the need of finding a point where to stop the iterative procedure, we did not make it automatic in the code. We instead followed the plot of $F(c^{(n)})$. This is the function that is being minimized by Chahine's method. As it turns out a few iterations sufficed as shown in Figures 4 and 6 as well as in a large number of other experiments we performed.

The technique described in Section 3 can also be applied to other reconstruction problems of similar structure; for example, if the data y were treated by Wiener filtering. In this case, the PSF h_i would be substituted by

$$\mathcal{F}^{-1}\left(\frac{H_i}{\sum_j H_j^2 + \epsilon}\right),$$

where ϵ is the signal to noise ratio. This approach is a natural one for following up the present work.

Acknowledgments

This work was supported by CNPq (Brazil) through grant 30.1003, by the Comission Interministerial de Ciencia y Tecnologia (Spain) through grants BIO 98-0761 and BIO 2001-1237, and by the National Institute of Health (USA) through grants HL67465 and HL740472. The implementation in the present work made extensive use of the X-Windows-based Microscopy Image Processing Package (XMIPP) which is available from http://www.cnb.uam.es/~bioinfo/. The authors thank Prof. Alvaro De Pierro for providing a number of useful references. The experimental data used in the reconstructions was kindly provided by Dr. M. Gómez-Lorenzo at the Biocomputing Unit, Centro Nacional de Biotecnología, CSIC, Spain.

Bibliography

[1] V. Barcilon, *On Chahine's relaxation method for the radiative transfer equation*, J. Atmospherical Sciences, 32 (1975), pp. 1626–1630.

[2] M. Chahine, *Inverse problems in radiative transfer*, J. Atmospherical Sciences, 27 (1975), pp. 960–967.

[3] M. Daube-Witherspoon and G. Muehllehner, *An iterative image space reconstruction algorithm for volume ECT*, IEEE Transactions on Medical Imaging, MI-5 (1987), pp. 61–66.

[4] A. R. De Pierro, *On the convergence of the iterative image space reconstruction algorithm for volume ECT*, IEEE Transactions on Medical Imaging, MI-6 (1987), pp. 174–175.

[5] ——, *Nonlinear relaxation methods for solving symmetric linear complementarity problems*, J. Optim. Theory Appl., 64 (1990), pp. 87–99.

[6] P. P. B. Eggermont, *Multiplicative iterative algorithms for convex programming*, Linear Algebra Appl., 130 (1990), pp. 25–42. Linear algebra in image reconstruction from projections.

[7] ——, *On Volterra-Lotka differential equations and multiplicative algorithms for monotone complementarity problems*, in Mathematical methods in tomography (Oberwolfach, 1990), vol. 1497 of Lecture Notes in Math., Springer, Berlin, 1991, pp. 141–152.

[8] J. Frank, *Three-dimensional electron microscopy of macromolecular assemblies*, Academic Press, New York, NY, 1996.

[9] J. Frank, *Single-particle imaging of macromolecules by cryo-electron microscopy*, Annu. Rev. Biophys. Biomol. Struct., (2002), pp. 303–319.

[10] J. Frank and P. Penczek, *On the correction of the contrast function in biological electron microscopy*, Optik, 98 (1995), pp. 125–129.

[11] R. Gold and N. Scofield, *Iterative solution for matrix representation of detection systems*, Bulletin of the American Physical Society, 2 (1960), p. 276.

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

[13] G. T. Herman, *Image reconstruction from projections: The fundamentals of computerized tomography*, Academic Press Inc., New York, 1980.

[14] G. T. Herman, H. Levkowitz, S. McCormick, and H. K. Tuy, *Multilevel image reconstruction*, in Multiresolution Image Processing and Analysis, A. Rosenfeld, ed., Springer-Verlag, New York, 1984.

[15] R. M. Lewitt, *Multidimensional digital image representations using generalized Kaiser-Bessel window functions*, J. Opt. Soc. Am. A, 7 (1990), pp. 1834–1846.

[16] ——, *Alternatives to voxels for image representation in iterative reconstruction algorithms*, Phys. Med. Biol., 37 (1992), pp. 705–716.

[17] S. Ludtke, P. R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions*, J. Structural Biology, 128 (1999), pp. 82–97.

[18] S. Ludtke, J. Jakana, J. Song, D. Chuang, and W. Chiu, *A 11.5 A single-particle reconstruction of GroEL using EMAN*, J. Molecular Biology, 314 (2001), pp. 82–97.

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

[20] R. Marabini, E. Rietsel, R. Schröder, G. T. Herman, and J. M. Carazo, *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. Struct. Biol., 120 (1997), pp. 363–371.

[21] S. Matej and R. M. Lewitt, *Efficient 3D grids for image reconstruction using spherically-symmetric volume elements*, IEEE Trans. Nucl. Sci., 42 (1995), pp. 1361–1370.

[22] ——, *Practical considerations for 3-D image reconstruction using spherically symmetric volume elements*, IEEE T. Med. Imaging, 15 (1996), pp. 68–78.

[23] F. Natterer, *The mathematics of computerized tomography*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Reprint of the 1986 original.

[24] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[25] M. Radermacher, *3-dimensional reconstruction from random projections - orientational alignment via Radon transforms*, Ultramicroscopy, 53 (1994), pp. 121–136.

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

[27] H. Stark, E. Orlova, J. Rinke-Appel, N. Junke, F. Mueller, M. Rodnina, W. Wintermeyer, R. Brimacomber, and M. van Heel, *Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy*, Cell, 88 (1997), pp. 19–29.

[28] F. Thon, *Zur defokussierungsabhngigkeit des phasenkontrastes bei der elektronenmikroskopischen abbildung*, Z. Naturforsch, 21a (1966), pp. 476–478.

[29] P. N. T. Unwin, *Phase contrast electron microscopy of biological materials*, J. Microscopy, 98 (1973), pp. 299–312.

[30] M. van Heel, *Similarity measures between images*, Ultramicroscopy, 21 (1987), pp. 95–100.

[31] J. A. Velázquez-Muriel, C. O. S. Sorzano, J. J. Fernández, and J. M. Carazo, *A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjusting*, Ultramicroscopy, (2002). (accepted).

[32] R. Wade, *A brief look at imaging and contrast transfer*, Ultramicroscopy, 46 (1996), pp. 145–156.

[33] J. Zhu, P. Penczek, R. Schröder, and J. Frank, *Three-dimensional reconstruction with contrast transfer function correction from energy-filtered cryoelectron micrographs: Procedurre and application to the 70s* escherichia coli *ribosome*, J. Structural Biology, 118 (1997), pp. 197–219.