Published in final edited form as:
Ultramicroscopy. 2010 August ; 110(9): 1128–1142. doi:10.1016/j.ultramic.2010.04.002.

Fully Three-Dimensional Defocus-Gradient Corrected Backprojection in Cryoelectron Microscopy

Ivan G. Kazantsev[a,1], **Joanna Klukowska**[b], **Gabor T. Herman**[b,*], and **Laslo Cernetic**[c]

Ivan G. Kazantsev: kazantsev.ivan6@gmail.com; Joanna Klukowska: jklukowska.gc@gmail.com; Gabor T. Herman: gabortherman@yahoo.com; Laslo Cernetic: cslaszlo@inf.u-szeged.hu

[a] RISØ, Materials Research Department, Technical University of Denmark, DK-4000, Roskilde, Denmark

[b] Department of Computer Science, Graduate Center, City University of New York, New York, NY 10016, USA

[c] Department of Image Processing and Computer Graphics, University of Szeged, Szeged, Hungary

Abstract

Recognizing that the microscope depth of field is a significant resolution-limiting factor in three-dimensional cryoelectron microscopy, Jensen and Kornberg proposed a concept they called defocus-gradient corrected backprojection (DGCBP) and illustrated by computer simulations that DGCBP can effectively eliminate the depth of field limitation. They did not provide a mathematical justification for their concept. Our paper provides this, by showing (in the idealized case of noiseless data being available for all projection directions) that the reconstructions obtained based on DGCBP from data produced with distance-dependent blurring are essentially the same as what is obtained by a classical method of reconstruction of a 3D object from its line integrals. The approach is general enough to be applicable for correcting for any distance-dependent blurring during projection data collection. We present a new implementation of the DGCBP concept, one that closely follows the mathematics of its justifications, and illustrate it using mathematically-described phantoms and their reconstructions from finitely-many distance-dependently blurred projections.

Keywords

Cryoelectron microscopy; Contrast transfer function; Reconstruction; Correction; Distance dependence; Stationary phase approximation

1. Introduction

Three-dimensional cryoelectron microscopy (3D cryoEM) is an increasingly powerful tool for solving the structure of macromolecular complexes, providing resolution on the order of a nanometer. To increase resolution to subnanometer scale, reconstruction methods have to

[*]Corresponding author. Tel.: 1 212 817 8193; fax: 1 212 817 1510.
[1]Present address: Institute of Computational Mathematics and Mathematical Geophysics, Lavrentieva 6, Novosibirsk 630090, Russia

take further image formation model features into account. In 3D cryoEM, 2D projection images, called *micrographs*, of a 3D mass distribution (e.g., a macromolecule) are affected by many factors that modify the amplitudes and phases of the image of the specimen and which must be corrected for in order to reconstruct the true object [1]. One of the most important among these factors is the contrast transfer function of the microscope. The *contrast transfer function* (CTF) is the Fourier transform of a point spread function that describes the response of the imaging system to a point object. It affects various frequencies by modulating the magnitude and sign of their amplitude. CTF depends on many parameters of the imaging system, among them defocus. In electron microscopes, the defocus varies with the distance from the electron source. Thus, given a three-dimensional specimen, each layer (defined as a plane perpendicular to the electron beam) is blurred by a slightly different transfer function. Most of the methods of correction for CTF ignore this dependence on distance from the electron source. As the technology of electron microscopy improves (as achievable resolution increases) and the need for imaging larger specimens emerges, this imperfection of electron microscopes, which has not been considered important in the past, is likely to become an essential limitation. The difference between two reconstructions, one that uses the same CTF function for each layer of the specimen and one that takes distance dependence into consideration is illustrated in Fig. 1. An extensive discussion of how Fig. 1 was obtained is presented in Section 6.

There are several approaches in the literature that address the distance-dependent CTF issue [2,3,4,5,6,7,8]. Jensen and Kornberg's [4] approach makes CTF correction an integral part of the reconstruction procedure. In this paper we revisit the concept behind their method in order to provide a mathematical justification for it and to put it into the context of traditional computerized tomography techniques.

The method proposed by Jensen and Kornberg [4], based on the concept of defocus-gradient corrected backprojection (DGCBP), is an approach that operates on micrographs taken from arbitrary directions. The method exploits features of the forward model for 3D cryoEM together with the general structure of the weighted backprojection technique [9]. However, the authors of [4] did not elaborate how their method relates to other reconstruction and/or correction approaches, and they provided only a heuristic (rather than mathematical) justification as to why the method should work.

In our recent work [8] we provided a mathematical verification of the DGCBP concept for the case in which projections are obtained from a single axis rotation mode of data collection. We demonstrated that, for that geometry, DGCBP and the frequency distance relation method described by Dubowy and Herman [6] are equivalent in the sense that the mathematical formulas that describe a 2D object reconstructed by the two methods from its distance-dependently distorted 1D projections from all directions around the axis of rotation are in fact the same.

In this paper we generalize that proof to 3D objects to be reconstructed from 2D projections taken from arbitrary directions, again with the assumption that data for all directions are available. We make use of stationary phase approximation, which was introduced to the field by Edholm and Lewitt [10] and Xia *et al.* [11] and then used by Dubowy and Herman [6] in the frequency distance relation method mentioned above. (For a brief review of stationary phase approximation, see Appendix C.) We show that results obtained using the DGCBP concept are equivalent to the results produced by a classical reconstruction method from ideal projection data. This completes the mathematical justification of the DGCBP approach.

Our paper is organized as follows. In the next section we present a nonmathematical overview of our ideas. The rest of the paper contains mathematical discussions (including proofs of our claims) and simulation results. In Sec. 3 we provide the background and introduce the notation used throughout the paper. In Sec. 4 we review the principles of image formation in an electron microscope and the model for the CTF. In Sec. 5 we outline the DGCBP concept in terms of integral equations and explore its effect on the 3D delta function as a test object. We also derive the main result of our work regarding the equivalence of the DGCBP concept and deblurred (weighted) backprojection applied to undistorted projection data. In Sec. 6 we present numerical test results. For clarity, only essential mathematical formulas are included in the text. Details of mathematics and implementation are delayed until the appendices, so as not to interrupt the flow of the main ideas.

2. Overview

This section provides an overview of the material presented in this paper. We describe operators that are defined in the following sections and their use in modeling of and in correcting for distance-dependent CTF blurring in cryoEM, without any mathematical derivations. All claims made here are proven in the following sections and the appendices.

We first model mathematically the process of projection taking in cryoEM. Such a model needs to incorporate the distance-dependent nature of the CTF. We then mathematically model the correction for distance-dependent CTF that Jensen and Kornberg [4] incorporated into the weighted backprojection reconstruction algorithm.

We start by defining several operators which we first use to describe ideal projections (line integrals with no blurring at all). This projection operator $\mathcal{P}$ is composed of a rotation operator $\mathcal{R}$ and a compression operator $\mathcal{C}$. The imaged molecule is represented by a function of three variables v, $v(x_1, x_2, x_3)$ is the density of this molecule at the point $(x_1, x_2, x_3)^T$. Given a function v and two angles θ and φ, the operator $\mathcal{R}$ gives us a function $[\mathcal{R}v](\theta, \varphi, x_1, x_2, x_3)$ that represents the density value of the molecule at a point (x_1, x_2, x_3) after it has been rotated by the angle θ around X_3-axis and then by the angle φ around X_2-axis. Note that the coordinate system is attached to the microscope, so that the values of the molecule at a point $(x_1, x_2, x_3)^T$ before and after the rotation are different. To obtain the projection of the molecule rotated in this fashion, we compute line integrals through rotated molecule $\mathcal{R}v$ along lines parallel to the X_3-axis. This is modeled by the compression operator $\mathcal{C}$. The projection $\mathcal{P}v = \mathcal{C}\mathcal{R}v$ is a function of four variables: x_1 and x_2, which describe positions in the two dimensional projection plane, and θ and φ, which describe how the molecule was rotated before the compression. In order to model the reconstruction from projection data we define the backprojection operator $\mathcal{B}$, which is again composed of two simpler operators that, so to speak, reverse the actions of the compression and rotation operators. Given a projection image, the spreading back operator $\mathcal{S}$ returns a function of five variables: x_1, x_2, and x_3, which describe a point's position in 3D space, and θ and φ, which keep the information about direction from which projection was obtained. The values of the function that results from this spreading back operation are independent of x_3, the planes perpendicular to the X_3-axis contain copies of the projection image. The totaling operator $\mathcal{T}$ combines the functions that are provided by the spreading back operator by rotating them into appropriate positions and adding them together. The backprojection operator is then $\mathcal{B} = \mathcal{T}\mathcal{S}$ To obtain a final reconstruction, we need a deblurring operator $\mathcal{D}$ that compensates for the blurring introduced by the projection followed by backprojection. For these operators we show that the function v can be recovered exactly from all its projections by applying backprojection followed by deblurring, i.e., $v = \mathcal{D}\mathcal{B}\mathcal{P}v$.

We then incorporate the distance-dependent CTF into the model described above by redefining the compression and spreading back operators. In order to add a distance-dependent CTF to the compression operator, each slice of the imaged object v perpendicular to the X_3-axis is two-dimensionally convolved with a blurring function before taking the integral along the X_3 direction. The blurring, which depends on x_3, is specified by a function h and the resulting operator is called the distance-dependent compression operator $\mathcal{C}_h$. The distance-dependent projection operator is $\mathcal{P}_h = \mathcal{C}_h \mathcal{R}$. To compensate for the distance-dependent CTF, we make use of a distance-dependent spreading back operator $\mathcal{S}_h$ that deconvolves the image that is spread in the direction of X_3-axis in a way that is appropriate to compensate for the CTF blurring that changes from plane-to-plane perpendicular to the X_3-axis. Consequently, the function produced by the distance-dependent spreading back operator depends on x_3; each plane perpendicular to the X_3-axis contains the projection image that has been deconcolved in a manner that is dependent on x_3. The distance-dependent backprojection operator is $\mathcal{B}_h = \mathcal{T}\ \mathcal{S}_h$.

We show that $\mathcal{B}_h \mathcal{P}_h v \approx \mathcal{B}\mathcal{P}v$, which leads to our main result: $v \approx \mathcal{D}\mathcal{B}_h \mathcal{P}_h v$. This implies that the molecule, described by the function v, can be approximately reconstructed from its distance-dependently blurred projections using the distance-dependent backprojection followed by deblurring. This provides a mathematical justification for the proposed reconstruction procedure.

Finally, using numerical experiments we demonstrate that the proposed method that makes use of distance-dependent backprojection recovers more details from various types of noisy projections than a reconstruction method that is based on the assumption that the CTF is not dependent on distance.

3. Mathematical background

In this section we describe the necessary background material, notational conventions and the operators used in the rest of the paper.

3.1. Notation

Let X_1, X_2, and X_3 be the axes of a Cartesian coordinate system. It is common in the electron microscopy literature to attach the coordinate system to the object to be reconstructed. Then the microscope is treated as if it were rotating around this object to obtain projections from various directions. In this paper, we follow a different, but equivalent, convention. Our coordinate system is attached to the microscope, with the X_3-axis parallel to the electron beam. Thus, the projections are always taken parallel to the X_3-axis and it is the molecule that is rotated around the origin of the coordinate system (assumed to be fixed) to obtain various projections. This allows a simpler mathematical description of distance-dependent blurring than what would be needed using the more common convention.

We use four vector spaces, $\mathbb{R}^3$, $\mathbb{S} \times \mathbb{R}^3$, $\mathbb{R}^2$, and $\mathbb{S} \times \mathbb{R}^2$, where $\mathbb{R}$ is a set of all real numbers, $\mathbb{R}^2$ and $\mathbb{R}^3$ are short for $\mathbb{R} \times \mathbb{R}$ and $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$, respectively, and $\mathbb{S} = [0, 2\pi) \times [0, \pi)$ is the set of directions on the unit sphere in $\mathbb{R}^3$.

$\mathbb{R}^3$ is the vector space of the molecule, i.e., the object to be reconstructed. We represent points in $\mathbb{R}^3$ using vectors, an arbitrary point is given by $(x_1, x_2, x_3)^T$. Similarly, we use $(\xi_1, \xi_2, \xi_3)^T$ to represent points in the frequency domain. We denote by v the object being imaged. The value $v(x_1, x_2, x_3)$ of the function v at a point $(x_1, x_2, x_3)^T$ is the value of the molecule at the appropriate place. We assume that v is a square integrable function with an origin-centered ball as its finite support.

$\mathbb{S} \times \mathbb{R}^3$ is the space of rotated molecules. A point is specified by $(\theta, \varphi, x_1, x_2, x_3)^T$, where θ and φ specify how the molecule was rotated inside the microscope and x_1, x_2, and x_3 refer to the microscope's coordinate system. The value of a rotated molecule at $(\theta, \varphi, x_1, x_2, x_3)^T$ for a fixed x_1, x_2, and x_3 is different for different values of θ and φ.

$\mathbb{R}^2$ is the space of single micrographs. $\mathbb{S} \times \mathbb{R}^2$ is the space of projection data, i.e., the set of all micrographs. An arbitrary point is given by $(\theta, \varphi, x_1, x_2)^T$, where x_1 and x_2 specify a location in the micrograph and θ and φ specify how the molecule was rotated before the micrograph was taken.

We use four function spaces:

$$\mathbf{V}=\{v:\mathbb{R}^3 \to \mathbb{R}\}, \tag{1}$$

$$\mathbf{W}=\{w:\mathbb{S} \times \mathbb{R}^3 \to \mathbb{R}\}, \tag{2}$$

$$\mathbf{P}=\{p:\mathbb{R}^2 \to \mathbb{R}\}. \tag{3}$$

$$\mathbf{G}=\{g:\mathbb{S} \times \mathbb{R}^2 \to \mathbb{R}\}. \tag{4}$$

We do not define these function spaces with complete mathematical rigor; such treatment can be found in Herman and Tuy [12] or in Natterer and Wübbeling [13]. We use spaces of functions that are general enough to include ordinary functions as well as generalized functions, such as the impulse functions ι and κ in $\mathbf{V}$, defined by Eqs. (A.2) and (26) based on the Dirac delta δ of Eq. (A.1).

We make use of the following rotation matrices:

$$D_\theta=\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D_\varphi=\begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix}. \tag{5}$$

D_θ is a right-hand rotation by θ in the X_1X_2-plane and $D_\varphi$is a right-hand rotation by φ in the X_1X_3-plane. For shorter and clearer notation we make the following definitions:

$$\begin{pmatrix} x_1^F(\theta,\varphi) \\ x_2^F(\theta,\varphi) \\ x_3^F(\theta,\varphi) \end{pmatrix}=D_\theta^{-1}D_\varphi^{-1}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} x_1\cos\theta\cos\varphi+x_2\sin\theta-x_3\cos\theta\sin\varphi \\ -x_1\sin\theta\cos\varphi+x_2\cos\theta+x_3\sin\theta\sin\varphi \\ x_1\sin\varphi+x_3\cos\varphi \end{pmatrix} \tag{6}$$

and

$$\begin{pmatrix} x_1^B(\theta,\varphi) \\ x_2^B(\theta,\varphi) \\ x_3^B(\theta,\varphi) \end{pmatrix} = D_\varphi D_\theta \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1\cos\theta\cos\varphi - x_2\sin\theta\cos\varphi + x_3\sin\varphi \\ x_1\sin\theta + x_2\cos\theta \\ -x_1\cos\theta\sin\varphi + x_2\sin\theta\sin\varphi + x_3\cos\varphi \end{pmatrix}. \tag{7}$$

Given a point $(x_1, x_2, x_3)^T$ in a rotated object, $(x_1^F(\theta,\varphi), x_2^F(\theta,\varphi), x_3^F(\theta,\varphi))^T$ is this point's location before rotation. Similarly, given a point $(x_1, x_2, x_3)^T$ in a not rotated object, $(x_1^B(\theta,\varphi), x_2^B(\theta,\varphi), x_3^B(\theta,\varphi))^T$ is this point's location after rotation. *F* in Eq. (6) stands for forward, since it is used in forward model of electron microscopy data collection. *B* in Eq. (7) stands for backward, since it is used in the reconstruction.

3.2. Operators

An *operator* is a mapping that acts on a function or functions and produces another function. We use capital script letters to denote operators; for example, $\mathcal{P}$ denotes the projection operator. In the remainder of this section we define several operators that are used later.

An arbitrary rotation of the function v can be represented by a rotation around the X_3-axis (multiplication by a matrix D_θ in Eq. (5)), followed by a rotation around the X_2-axis (multiplication by the matrix D_φ in Eq. (5)), followed by another rotation around the X_3-axis. The last rotation is known as an *in-plane rotation* and without loss of generality can be ignored, as it is just a rotation in the projection plane. See Fig. 2 for illustrations of the first two rotations.

The *rotation operator* $\mathcal{R}$: $\mathbf{V} \to \mathbf{W}$ associates with a function of three variables a function of five variables. $[\mathcal{R}v](\theta, \varphi, x_1, x_2, x_3)$ is the value of a molecule v at a point (x_1, x_2, x_3) after it is rotated by θ and φ as illustrated in Fig. 2. The rotation operator is defined by

$$[\mathcal{R}v](\theta,\varphi,x_1,x_2,x_3) = v(x_1^F(\theta,\varphi), x_2^F(\theta,\varphi), x_3^F(\theta,\varphi)). \tag{8}$$

The operator $\mathcal{R}$ maps functions in $\mathbf{V}$ into functions in $\mathbf{W}$, but there are functions $w \in \mathbf{W}$ for which there is no $v \in \mathbf{V}$ such that $w = \mathcal{R}v$.

The *compression operator* $\mathcal{C}$: $\mathbf{W} \to \mathbf{G}$ takes a function of five variables $(\theta, \varphi, x_1, x_2, x_3)$ and produces a function of four variables $(\theta, \varphi, x_1, x_2)$ by integrating along the X_3-axis. The first two variables are included to denote the angles by which the molecule was rotated before the compression, and the other two variables indicate the location of a point in the X_1X_2-plane. The compression operator $\mathcal{C}$ is defined by

$$[Cw](\theta,\varphi,x_1,x_2) = \int_{\mathbb{R}} w(\theta,\varphi,x_1,x_2,x_3)\,dx_3. \tag{9}$$

The operator $\mathcal{C}$ maps any function in $\mathbf{W}$ (not just those that are in the range of the operator $\mathcal{R}$) into functions in $\mathbf{G}$. In electron microscopy we apply $\mathcal{C}$ only to functions that are in the range of $\mathcal{R}$.

The composition $\mathcal{P} = \mathcal{C}\mathcal{R}$ of these two operators defines the *projection operator* $\mathcal{P}$: $\mathbf{V} \to \mathbf{G}$, such that, for all $v \in \mathbf{V}$ and $(\theta, \varphi, x_1, x_2)$,

$$[\mathcal{P}v](\theta,\varphi,x_1,x_2)=[C[\mathcal{R}v]](\theta,\varphi,x_1,x_2). \tag{10}$$

Data that are provided to an algorithm for 3D reconstruction consist of imperfect samples of $\mathcal{P}v$ [14]. Geometries of data collection used in electron microscopy (as described, for example, in Subsection 3.3 of [14]) can be identified by considering the ranges over which θ and φ are sampled. For example, if θ ranges over the whole of $[0, 2\pi)$ and φ has any fixed value in $(0, \pi)$ other than $\frac{\pi}{2}$, then we have the so-called conical tilt mode of data collection that provides us (via the central section theorem) with information about the Fourier transform of v outside a missing cone. The choice $\varphi=\frac{\pi}{2}$ is special: when combined with θ ranging over the whole of $[0, 2\pi)$, it provides us with a complete single-axis tilt series and, hence, with information about the Fourier transform of v without any missing region. In view of this, it is not altogether surprising that $\varphi=\frac{\pi}{2}$ plays a special role also in the mathematical derivation provided in Appendix C.

During the reconstruction process we make use of operators that, so to speak, undo the actions of the operators described above. One of these is $\mathcal{S}$: $\mathbf{G} \to \mathbf{W}$, which we call the *spreading back operator*. If $g \in \mathbf{G}$, then $\mathcal{S}g$ is commonly referred to as a ridge function resulting from spreading back g (*ridge functions* are functions whose values do not change in one direction, in this case it is the direction of the X_3-axis). The value of $\mathcal{S}g$ at the point $(\theta, \varphi, x_1, x_2, x_3)$ is

$$[\mathcal{S}g](\theta,\varphi,x_1,x_2,x_3)=g(\theta,\varphi,x_1,x_2). \tag{11}$$

$\mathcal{T}$: $\mathbf{W} \to \mathbf{V}$ is the *totaling operator* that performs integration of functions in $\mathbf{W}$ over the unit sphere $\mathbb{S}$. The value of $\mathcal{T}w$, which is sometimes referred to as an *integral image*, at the point (x_1, x_2, x_3) is

$$[\mathcal{T}w](x_1,x_2,x_3)=\int_{\mathbb{S}} w(\theta,\varphi,x_1^B(\theta,\varphi),x_2^B(\theta,\varphi),x_3^B(\theta,\varphi))\sin\varphi d\varphi d\theta, \tag{12}$$

where the term $\sin\varphi$ is required to achieve uniform integration over all directions. The operator $\mathcal{T}$ maps any function in $\mathbf{W}$ into $\mathbf{V}$. In our applications we apply $\mathcal{T}$ to ridge functions that are in the range of $\mathcal{S}$. $\mathcal{T}$ unrotates the ridge functions $w(\theta, \varphi, x_1, x_2, x_3)$ by the angles θ and φ (the same angles by which the molecule was rotated before the projection was obtained), and then the unrotated ridge functions for all θ and φ are totaled together.

The composition $\mathcal{T}\mathcal{S}$ defines the *backprojection operator* $\mathcal{B}$: $\mathbf{G} \to \mathbf{V}$, such that, for all $g \in \mathbf{G}$ and (x_1, x_2, x_3),

$$[\mathcal{B}g](x_1,x_2,x_3)=[\mathcal{T}[\mathcal{S}g]](x_1,x_2,x_3). \tag{13}$$

In Appendix A we give precise mathematical definitions of additional operators including both 2D and 3D Fourier transforms (both denoted by $\mathcal{F}$), their inverses (denoted by $\mathcal{F}^{-1}$), special Fourier transforms and their inverses operating on functions in $\mathbf{V}$, $\mathbf{W}$, and $\mathbf{G}$ (denoted by $\mathcal{F}_{\mathbf{V}}$, $\mathcal{F}_{\mathbf{V}}^{-1}$, $\mathcal{F}_{\mathbf{W}}$, $\mathcal{F}_{\mathbf{W}}^{-1}$, $\mathcal{F}_{\mathbf{G}}$, and $\mathcal{F}_{\mathbf{G}}^{-1}$, respectively), 2D and 3D convolutions, and even an operator that convolves a function from $\mathbf{W}$ with one from $\mathbf{V}$ (all three denoted by *), and an operator that divides a function in $\mathbf{G}$ by a function in $\mathbf{V}$.

Using the inverse 3D Fourier transform, Eq. (A.16), and the 3D convolution, Eq. (A.18), operators, we define the *3D deblurring operator* $\mathcal{D}$: $\mathbf{V} \to \mathbf{V}$ as follows. Let $\hat{r} \in \mathbf{V}$, be the function defined by $\widehat{r}(\xi_1, \xi_2, \xi_3) = \sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}$. Then, for any $v \in \mathbf{V}$,

$$\mathcal{D}v = \frac{1}{16\pi^3 \sqrt{2\pi}} [\mathcal{F}^{-1}\widehat{r}] * v. \tag{14}$$

The operators and spaces of functions on which they act are summarized in Fig. 3.

The mathematical idealization of data collection in 3D cryoEM (no noise, no blurring, data from all projection directions available) gets us from the molecule $v \in \mathbf{V}$ to its projection data $\mathcal{P}v$. A well-known significant fact is that the molecule can be recovered from $\mathcal{P}v$ by application of backprojection followed by deblurring:

$$\mathcal{D}\mathcal{B}\mathcal{P}v = v. \tag{15}$$

A justification for this claim is given in Appendix B.

4. Image formation model in 3D cryoEM

The image formation by an electron microscope includes a blurring, which is different in each layer (plane perpendicular to the electron beam) of the specimen during the projection generation in 3D cryoEM. A projection data set that consists of all distorted 2D projections (micrographs) is defined as a compression of $\mathcal{R}v \in \mathbf{W}$ after it has been convolved with a point spread function $h \in \mathbf{V}$. Mathematically, the *distance-dependent compression operator* $\mathcal{C}_h$: $\mathbf{W} \to \mathbf{G}$ is defined by

$$C_h w = C[w * h] \tag{16}$$

(it follows from Eq. (A.20) that $\mathcal{C}_1 = \mathcal{C}$), and the *distance-dependent projection* operator $\mathcal{P}_h$: $\mathbf{V} \to \mathbf{G}$ is defined by

$$\mathcal{P}_h = C_h \mathcal{R}. \tag{17}$$

The Fourier transform of the distance-dependent projection data satisfies the formula

$$[\mathcal{F}_{\mathbf{G}} \mathcal{P}_h v](\theta, \varphi, \xi_1, \xi_1) = 2\pi \int_{\mathbb{R}} [\mathcal{F}_{\mathbf{W}} \mathcal{R} v](\theta, \varphi, \xi_1, \xi_2, x_3) H(\xi_1, \xi_2, x_3) dx_3, \tag{18}$$

where

$$H = \mathcal{F}_{\mathbf{V}} h. \tag{19}$$

This is easily proved by starting with its left hand side and then applying Eqs. (17), (A.12), (A.6), (A.5), (16), (9), a change in the order of integration, (A.19), (A.6), (A.21), (A.10), and (A.8). Equation (18) indicates an efficient layer-by-layer method of calculating $\mathcal{P}_h v$.

The approach proposed in this paper can be used to correct for *any* kind of distance-dependent blurring, specified by any point spread function $h \in \mathbf{V}$. In 3D cryoEM, the Fourier transform H of h is called the *contrast transfer function* (CTF). It can be assumed to be radially symmetric in the (ξ_1, ξ_2) plane (even when this is not the case, our mathematics is often still applicable after a simple change of variables) and has the form [1, Ch. 3]

$$H(\xi_1, \xi_2, x_3) = H_{CTF}(\xi, x_3)\, E_{spat}(\xi, x_3)\, E_{temp}(\xi), \tag{20}$$

where

$$\begin{aligned} H_{CTF}(\xi, x_3) &= (1-a)\sin(D(\xi, x_3)) - a\cos(D(\xi, x_3)), \\ D(\xi, x_3) &= 2\pi\lambda\xi^2\left(-\frac{\Delta f(x_3)}{2} + \frac{\lambda^2\xi^2 C_s}{4}\right), \\ E_{spat}(\xi, x_3) &= e^{-\pi^2 q_0^2 (C_s\lambda^3\xi^3 - \Delta f(x_3)\lambda\xi)^2}, \\ E_{temp}(\xi) &= e^{-(\frac{1}{2}\pi F_s \lambda\xi^2)^2}, \end{aligned} \tag{21}$$

and the parameters involved are:

- $\xi \equiv \sqrt{\xi_1^2 + \xi_2^2}$ is a spatial frequency,
- a is a fraction of the amplitude contrast, $0 \le a \le 1$,
- λ is the electron wavelength,
- C_s is the lens spherical aberration coefficient,
- $\Delta f(x_3)$ is the value of the defocus,
- q_0 is a quantity of dimension 1/length specifying the size of the source as it appears in the back focal plane,
- F_s is the lens focal spread coefficient.

Note that the defocus $\Delta f(x_3)$ depends explicitly on the distance from the electron source.

5. Defocus-gradient corrected backprojection

Irrespective of the choice of the point spread function h, our mathematical version of the DGCBP concept of [4] has the following essence. A distance-dependent backprojection operator $\mathcal{B}_h$: $\mathbf{G} \to \mathbf{V}$ is specified so that it has the property that, for all $v \in \mathbf{V}$,

$$\mathcal{B}\mathcal{P}v \approx \mathcal{B}_h\mathcal{P}_h v, \tag{22}$$

where ≈ stands for *approximately equal*. This, combined with Eq. (15), implies that

$$\mathcal{D}\mathcal{B}_h\mathcal{P}_h v \approx v, \tag{23}$$

and so the required volume can be approximated from the distance-dependently blurred projection data $\mathcal{P}_h v$ by applying to it first the distance-dependent backprojection operator and then the deblurring operator.

More specifically, we define the *distance-dependent spreading back operator* $\mathcal{S}_h$: $\mathbf{G} \rightarrow \mathbf{W}$ by

$$\mathcal{S}_h g = \mathcal{F}_{\mathrm{W}}^{-1} \frac{[\mathcal{F}_{\mathrm{G}} g]}{[\mathcal{F}_{\mathrm{V}} h]}. \tag{24}$$

This is a generalization of the spreading back operator, as defined in Eq. (11), because $\mathcal{S}_h = \mathcal{S}$, as can be seen by applying to the right hand side of Eq. (24), for this case, Eqs. (A.11), (A.7), (A.4), (A.22), (A.14), (A.12), (A.6), (A.5), the fact that the 2D inverse Fourier transform is in fact the inverse of the 2D Fourier transform, and (11). The *distance-dependent backprojection operator* is

$$\mathcal{B}_h = \mathcal{T} \mathcal{S}_h. \tag{25}$$

(Note: $\mathcal{B}_h = \mathcal{B}$.) To complete our mathematical justification we need to show that Eq. (22) holds for all $v \in \mathbf{V}$.

By the linearity of the operators this follows from the special case when v is the 3D impulse function centered at an arbitrary point $(\hat{x}_1, \hat{x}_2, \hat{x}_3) \in \mathbb{R}^3$, defined by

$$\kappa(x_1, x_2, x_3) = \delta(\widehat{x_1} - x_1)\delta(\widehat{x_2} - x_2)\delta(\widehat{x_3} - x_3). \tag{26}$$

According to Eq. (8), the rotated version of this function, for any pair of angles θ and φ, is

$$[\mathcal{R}\kappa](\theta, \varphi, x_1, x_2, x_3) = \delta(\widehat{x_1} - x_1^F(\theta, \varphi))\, \delta(\widehat{x_2} - x_2^F(\theta, \varphi))\, \delta(\widehat{x_3} - x_3^F(\theta, \varphi)). \tag{27}$$

Notice that the expression on the right hand side of Eq. (27) is equivalent to

$$\delta(\widehat{x_1^B}(\theta, \varphi) - x_1)\, \delta(\widehat{x_2^B}(\theta, \varphi) - x_2)\, \delta(\widehat{x_3^B}(\theta, \varphi) - x_3). \tag{28}$$

In the rest of the paper we use Eq. (28) to express $[\mathcal{R}\kappa](\theta, \varphi, x_1, x_2, x_3)$.

We first observe that

$$[\mathcal{F}_{\mathrm{G}} \mathcal{P}_h \kappa](\theta, \varphi, \xi_1, \xi_2) = H(\xi_1, \xi_2, \widehat{x_3^B}(\theta, \varphi))\, e^{-i(\xi_1 \widehat{x_1^B}(\theta, \varphi) + \xi_2 \widehat{x_2^B}(\theta, \varphi))}. \tag{29}$$

This can be derived by using Eqs. (18), (A.10), (A.6), (A.4), (28), and (A.1).

To get an integral expression for $\mathcal{B}_h\mathcal{P}_h\kappa$, we apply Eqs. (25), (24), (12), (A.11), (A.7), (A.4), (A.22), and (29) and obtain

$$[\mathcal{B}_h\mathcal{P}_h\kappa](x_1,x_2,x_3)=\frac{1}{2\pi}\int_{\mathbb{S}}\int_{\mathbb{R}^2}\frac{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}e^{i\xi_1(x_1^B(\theta,\varphi)-\widehat{x}_1^B(\theta,\varphi))}e^{i\xi_2(x_2^B(\theta,\varphi)-\widehat{x}_2^B(\theta,\varphi))}d\xi_1 d\xi_2\sin\varphi\, d\varphi\, d\theta. \tag{30}$$

To work with the outermost two integrals in Eq. (30), we rewrite $\mathcal{B}_h\mathcal{P}_h\kappa$ by introducing two functions

$$\Psi_1(\theta,\varphi)=x_1^B(\theta,\varphi)-\widehat{x}_1^B(\theta,\varphi), \tag{31}$$

$$\Psi_2(\theta)=x_2^B(\theta,0)-\widehat{x}_2^B(\theta,0). \tag{32}$$

In definition of Ψ_2, we set arbitrarily $\varphi = 0$ in x_2^B and $\widehat{x}_2^B$ because, by Eq. (7), they do not depend on φ. After changing the order of integration, Eq. (30) becomes

$$[\mathcal{B}_h\mathcal{P}_h\kappa](x_1,x_2,x_3)=\frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{S}}\frac{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}\sin\varphi e^{i\xi_1\Psi_1(\theta,\varphi)}d\varphi e^{i\xi_2\Psi_2(\theta)}d\theta d\xi_1 d\xi_2 \tag{33}$$

Consider first the special case when $(x_1, x_2, x_3)^T = (\hat{x}_1, \hat{x}_2, \hat{x}_3)^T$. In this case it follows trivially from Eq. (7) that $\widehat{x}_3^B(\theta,\varphi)=x_3^B(\theta,\varphi)$, for all (θ, φ), and so the value of the fraction inside the integral in Eq. (33) is always 1, independently of the choice of h. This means that $[\mathcal{B}_h\mathcal{P}_h\kappa](\hat{x}_1, \hat{x}_2, \hat{x}_3) = [\mathcal{B}_\iota\mathcal{P}_\iota\kappa](\hat{x}_1, \hat{x}_2, \hat{x}_3) = [\mathcal{B}\mathcal{P}\kappa](\hat{x}_1, \hat{x}_2, \hat{x}_3)$ and Eq. (22) is satisfied exactly at this point.

Our proof that $[\mathcal{B}_h\mathcal{P}_h\kappa](x_1, x_2, x_3) \approx [\mathcal{B}\mathcal{P}\kappa](x_1, x_2, x_3)$ in all other cases as well requires making use of the stationary phase approximation and therefore we placed it into Appendix C.

This gives us our main mathematical result: When projections from all the directions on the unit sphere are available and no noise is present during projection generation process, $\mathcal{B}_h\mathcal{P}_h\nu$ is approximately equal to $\mathcal{B}\mathcal{P}\nu$ The meaning of this is that the integral image produced from the distance-dependently blurred projection data of an object using the DGCBP concept is approximately the same as the integral image produced by a standard backprojection that uses true mathematical projection data of the same object. This gives a mathematical verification of the method of correction for distance-dependent blurring proposed by Jensen and Kornberg [4].

6. Numerical experiments

For experiments we selected the following parameters for the forward model of 3D cryoEM (see Section 4):

- $a = 0$,
- $\lambda = 0.033487$ Å,

- C_s = 22, 000, 000 Å,
- $\Delta f \in [1000, 3000]$ (in Å),
- $q_0 = 0.00746558$ Å^{-1}
- F_s = 141.35 Å.

We computed the distance-dependently blurred projection data $g(\theta, \varphi, x_1, x_2)$ of mathematically-defined phantoms, for randomly-selected directions (θ, φ) and 128 × 128 values of (x_1, x_2) in each direction, based on Eq. (17).

From these values we reconstructed the phantoms by numerically approximating $[\mathcal{D}\mathcal{B}_h g](x_1, x_2, x_3)$ at 128 × 128 × 128 values of (x_1, x_2, x_3). We see from Eqs. (25) and (24) that this involves a division, see Eq. (A.22). To avoid dividing by zero, we used a Tikhonov filter approximation

$$\frac{[\mathcal{F}_{\mathbf{G}} g](\theta, \varphi, \xi_1, \xi_2)}{H(\xi_1, \xi_2, x_3)} \approx \frac{[\mathcal{F}_{\mathbf{G}} g](\theta, \varphi, \xi_1, \xi_2)\, H(\xi_1, \xi_2, x_3)}{(H(\xi_1, \xi_2, x_3))^2 + \alpha(\xi_1^2 + \xi_2^2)}, \tag{34}$$

with $\alpha = 0.01$ ï¿œ^2; see [16]. More details regarding implementation can be found in Appendix D.

We first report on an experiment in which the phantom consists of seven identical spheres (Fig. 4) with centers located in a horizontal plane. We chose this phantom to illustrate the effects of distance-dependent blurring in the projection data. A 128 ×128 digital approximation of the central slice of the phantom and its 3D plot can be compared with matching images of the DGCBP reconstruction in Fig. 5.

A single tilted view (from one of the 5,000 generated directions) was chosen to illustrate the projection $\mathcal{P}v$ (Fig. 6(a)) and the micrograph $\mathcal{P}_h v$ (Fig. 6(b)). The variation in distortion (for example, for the spheres marked as A and B in Fig. 6(b)) caused by change in defocus is clearly seen. Profiles of the upper half of column number 59 and bottom half of column 53, as indicated in Fig. 6(a), are shown in Fig. 6(c) and (d).

We used a more complex phantom, shown in Fig. 7, to demonstrate the difference between reconstructions that take the distance-dependence of the CTF into account and those that ignore the distance-dependence of the CTF and assume a constant defocus appropriate only for the center layer of the phantom, which in our case is 2000 Å. We ran several simulations with this phantom with varying amounts of noise in the data to examine the sensitivity of our DGCBP-like correction. In the rest of this section we refer to a reconstruction that takes distance-dependence into account as *a DD backprojection* and to a reconstruction that corrects for the CTF corresponding to the center layer of the phantom as a *CL backprojection*.

Figure 1 demonstrates cross sections through the phantom, row (a), corresponding cross sections through the DD backprojection, row (b), cross sections through the CL backprojection, row (c), and the reconstruction that does not correct for the CTF at all, row (d). The projection data used for these reconstruction consisted of 5000 projection images that were affected only by the distance-dependent CTF blurring with no noise added to the distance-dependently blurred projections. In the CL backprojection the smallest spheres are blurred to the point that they do not appear to be separate from each other and there is also more cross-slice blurring than in the DD backprojection.

Next we modified our experiment by introducing different types of noise. We simulated structural noise, shot noise and digitization noise based on the guidance of Baxter *et al*. [15]. We needed to modify the model of noise suggested there in order to incorporate distance-dependent blurring into the structural noise. The structural noise originates from ice and often carbon film surrounding the molecule during imaging. As such it is different for each molecule and is subject to distance-dependent blurring. As a rough simulation of this phenomenon, prior to the distance-dependent projection taking we added to each voxel value in the rotated molecule a random sample from a zero-mean Gaussian distribution with standard deviation σ_1. This was done independently for each projection direction. This noise is convolved during the projection taking with the point spread function resulting in correlated noise in the projection images. Then, we simulated the shot noise and the digitization noise by adding, this time to each of the pixel values in the projection images, a random sample from a zero-mean Gaussian distribution with standard deviation σ_2. This was also done independently for each projection direction. Sample projection images for different values of σ_1 and σ_2 are compared to the ideal and noise-less projections in Fig. 8. We used the projection data sets obtained in this manner to reconstruct using DD backprojection and CL backprojection. The results for different values of σ_1 and σ_2 are shown in Fig. 9. The second and third columns are cross-sections of reconstructions from 5,000 projections with noise described by $\sigma_1 = 0.3052$ and $\sigma_2 = 2.99$. This value of σ_1 results in equal values of standard deviation of signal and noise after projections are taken. The value of σ brings down the signal to noise ratio to ½. The fourth and fifth columns are cross sections of reconstructions from 10,000 projections with noise described by $\sigma = 0.6103$ and $\sigma_2 = 6.0$ (which results in signal to noise ratio of ¼). The surface rendering of reconstructions obtained using the second set of standard deviation values are also shown in Fig. 10. Because surface rendering is done for a particular voxel threshold value, a lot of information is lost in such representations. For the rest of the simulations we only show cross-section images.

The parameters of the CTF need to be estimated from the projection images before reconstructions can be performed. To evaluate how DD backprojection is affected by incorrectly determined CTF parameters, we simulated projections in which the defocus parameters were not the same as the ones used in reconstruction. In all our experiments the defocus varies from 1000 Å to 3000 Å according to the function $\Delta f(x_3) = m(x_3 - 1/2) + b$, where $m = 2000/n$, $b = 1000$, and n = 128 is the number of discrete layers into which the molecule is subdivided. (Here x_3 is the layer index that goes from 1 to 128.) For each projection direction, we introduced a random variation to the distance-dependent CTF, by using a defocus function in which we added to m a sample from a zero-mean Gaussian distribution with standard deviation σ_3 and added to b a sample from a zero-mean Gaussian distribution with standard deviation σ_4. The reconstructions were performed by correcting for the CTF with the unperturbed values of m and b. We demonstrate the results for two different values of σ_3 and σ_4. The cross-sections of reconstructions are shown in Fig. 11. The DD backprojection results are again less blurred than the reconstructions obtained using CL backprojection on the same projection data.

In cryoEM the direction of each projection has to be estimated before the reconstruction can be performed. We repeated our experiment introducing, independently for each projection, a small difference between the direction from which the particular projection was actually obtained and the direction that was used during the reconstruction. The difference was introduced by adding to θ and φ used in the projection simulation two different samples from a zero-mean Gaussian distribution with standard deviation σ_5 and using the modified angles in the reconstruction. The resulting cross sections for two different values of σ_5 are shown in Fig. 12. For $\sigma_5 = 1$ the DD backprojection is once again less blurred than the CL backprojection from the same data. When we increased σ_5 to 3, both reconstruction are

blurred and even the larger spheres become indistinguishable. For comparison we also show a reconstruction obtained from projection data unaffected by the CTF, just by the incorrectly determined projection angles.

7. Conclusions

In this work we have shown that distance-dependent backprojection applied to distance-dependently blurred projections found in micrographs and backprojection applied to data produced by the true (no blurring) projection operator produce similar integral images. Theoretically, the DGCBP reconstruction is obtained by deblurring the integral image with the $\mathcal{D}$ filter. In practice, where only a limited number of projections are available, one can use a numerical approximation to the mathematical formula; we have demonstrated that such an approximation can produce satisfactory results under both noiseless and noisy conditions.

Acknowledgments

This work began during the stay of the first and fourth authors with the Discrete Imaging and Graphics group of the Graduate Center of the City University of New York. Financial support by Award Number R01HL070472 from the National Heart, Lung, And Blood Institute is acknowledged. The authors gratefully acknowledge the Danish National Research Foundation for supporting the Center for Fundamental Research: Metal Structures in Four Dimensions at the RISØ National Laboratory of Technical University of Denmark, where the work was finalized. They are particularly grateful to Søren Schmidt of that laboratory for useful discussions. The authors also acknowledge the helpful comments and suggestions of the reviewers. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, And Blood Institute or the National Institutes of Health.

References

1. Frank, J. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. 2. Oxford University Press; 2006.
2. Cohen HA, Schmid MF, Chiu W. Estimates of validity of projection approximation for three-dimensional reconstructions at high resolution. Ul-tramicroscopy 1984;14:219–226.
3. DeRosier DJ. Correction of high-resolution data for curvature of the Ewald sphere. Ultramicroscopy 2000;81:83–98. [PubMed: 10998793]
4. Jensen GJ, Kornberg RD. Defocus-gradient corrected back-projection. Ultramicroscopy 2000;84:57–64. [PubMed: 10896140]
5. Wan, Y.; Chiu, W.; Zhou, ZH. Full contrast transfer function correction in 3D cryo-EM reconstruction. International Conference on Communications, Circuits and Systems; 2004. p. 960-964.
6. Dubowy, JN.; Herman, GT. An approach to the correction of distance-dependent defocus in electron microscopic reconstruction. IEEE International Conference on Image Processing, ICIP; 2005; 2005. p. 748-751.
7. Wolf M, DeRosier DJ, Grigorieff N. Ewald sphere correction for single-particle electron microscopy. Ultramicroscopy 2006;106:376–382. [PubMed: 16384646]
8. Kazantsev, IG.; Herman, GT.; Cernetic, L. Backprojection-based reconstruction and correction for distance-dependent defocus in cryoelectron microscopy. Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro; 2008. p. 133-136.
9. Radermacher, M. Weighted back-projection methods. In: Frank, J., editor. Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell. 2. Springer; Berlin: 2006. p. 245-274.
10. Edholm, PR.; Lewitt, RM. Novel properties of the Fourier decomposition of the sinogram. International Workshop on Physics and Engineering of Computerized Multidimentional Imaging and Processing; 1986. p. 8-18.

11. Xia W, Lewitt RM, Edholm PR. Fourier correction for spatially variant collimator blurring in SPECT. IEEE Trans Med Imaging 1995;14:100–115. [PubMed: 18215814]
12. Herman, GT.; Tuy, HK. Image reconstruction from projections: An approach from mathematical analysis. In: Sabatier, PC., editor. Basic Methods of Tomography and Inverse Problems. Institute of Physics Publishing; 1988. p. 11-24.
13. Natterer F, Wöbbeling F. Mathematical Methods in Image Reconstruction. SIAM. 2001
14. Carazo, J-M.; Herman, GT.; Sorzano, COS.; Marabini, R. Algorithms for thee-dimensional reconstruction from the imperfect projection data provided by electron microscopy. In: Frank, J., editor. Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell. 2. Springer; Berlin: 2006. p. 217-244.
15. Baxter WT, Grassucci RA, Gao H, Frank J. Determination of signal-to-noise ratios and spectral SNRs in cryo-EM low-dose imaging of molecules. J Struct Biol 2009;166:126–132. [PubMed: 19269332]
16. Eldar YC, Unser M. Nonideal sampling and interpolation from noisy observations in shift-invariant spaces. IEEE Trans Signal Proc 2006;54:2636–2651.
17. Guillemin V, Sternberg S. Geometric Asymptotics. AMS. 1977
18. Defrise M. A factorization method for the 3D X-ray transform. Inverse Probl 1995;11:983–994.
19. Defrise M, Kinahan PE, Townsend DW, Michel Ch, Sibomana M, Newport DF. Exact and approximate rebinning algorithms for 3-D PET data. IEEE Trans Med Imaging 1997;16:145–158. [PubMed: 9101324]
20. Varslot T, Yarman CE, Cheney M, Yazici B. A variational approach to waveform design for synthetic-aperture imaging. Inverse Probl Imaging 2007;1:577–592.
21. Brigham, EO. The Fast Fourier Transform and Its Applications. Prentice-Hall; 1988.
22. Welling JS, Eddy WF, Young TK. Rotation of 3D volumes by Fourier-interpolated shears. Graphical Models 2006;68:356–370.

Appendix

A. Impulse functions, Fourier transforms, convolutions, and division of functions

The *Dirac delta* δ is a (generalized) function on $\mathbb{R}$ whose defining property is that, for any function f: $\mathbb{R} \to \mathbb{R}$ and for any $x \in \mathbb{R}$,

$$\int_{\mathbb{R}} f(x')\delta(x - x')\,dx' = f(x). \tag{A.1}$$

Using δ we are able to define various *impulse functions*, such as the κ of Eq. (26) and the ι in $\mathbf{V}$ that is defined by

$$\iota(x_1, x_2, x_3) = \delta(x_1)\delta(x_2). \tag{A.2}$$

The *slicing operators* $\mathcal{K}$ are used for defining various Fourier transform and convolution operators. For $v \in \mathbf{V}$, $w \in \mathbf{W}$, and $g \in \mathbf{G}$, respectively, they are defined by

$$[\mathcal{K}_{x_3} v](x_1, x_2) = v(x_1, x_2, x_3), \tag{A.3}$$

$$[\mathcal{K}_{\theta,\varphi,x_3} w](x_1, x_2) = w(\theta, \varphi, x_1, x_2, x_3), \tag{A.4}$$

$$[\mathcal{K}_{\theta,\varphi} g](x_1, x_2) = g(\theta, \varphi, x_1, x_2). \tag{A.5}$$

The *2D Fourier transform operator* $\mathcal{F}$: **P** → **P** and its inverse $\mathcal{F}^{-1}$: **P** → **P** are defined by

$$[\mathcal{F} p](\xi_1, \xi_2) = \frac{1}{2\pi} \int_{\mathbb{R}^2} p(x_1, x_2) e^{-i(\xi_1 x_1 + \xi_2 x_2)} dx_1 dx_2, \tag{A.6}$$

$$[\mathcal{F}^{-1} p](x_1, x_2) = \frac{1}{2\pi} \int_{\mathbb{R}^2} p(\xi_1, \xi_2) e^{i(\xi_1 x_1 + \xi_2 x_2)} d\xi_1 d\xi_2. \tag{A.7}$$

The 2D Fourier transform operators for other function spaces $\mathcal{F}_{\mathbf{V}}$: **V** → **V**, $\mathcal{F}_{\mathbf{W}}$: **W** → **W**, and $\mathcal{F}_{\mathbf{G}}$: **G** → **G** and their inverses $\mathcal{F}_{\mathbf{V}}^{-1}$:**V** → **V**, $\mathcal{F}_{\mathbf{W}}^{-1}$:**W** → **W**, and $\mathcal{F}_{\mathbf{G}}^{-1}$:**G** → **G** are defined by

$$[\mathcal{F}_{\mathbf{V}} v](\xi_1, \xi_2, x_3) = [\mathcal{F} \mathcal{K}_{x_3} v](\xi_1, \xi_2), \tag{A.8}$$

$$[\mathcal{F}_{\mathbf{V}}^{-1} v](x_1, x_2, x_3) = [\mathcal{F}^{-1} \mathcal{K}_{x_3} v](x_1, x_2), \tag{A.9}$$

$$[\mathcal{F}_{\mathbf{W}} w](\theta, \varphi, \xi_1, \xi_2, x_3) = [\mathcal{F} \mathcal{K}_{\theta,\varphi,x_3} w](\xi_1, \xi_2), \tag{A.10}$$

$$[\mathcal{F}_{\mathbf{W}}^{-1} w](\theta, \varphi, x, x_2, x_3) = [\mathcal{F}^{-1} \mathcal{K}_{\theta,\varphi,x_3} w](x_1, x_2), \tag{A.11}$$

$$[\mathcal{F}_{\mathbf{G}} g](\theta, \varphi, \xi_1, \xi_2) = [\mathcal{F} \mathcal{K}_{\theta,\varphi} g](\xi_1, \xi_2), \tag{A.12}$$

$$[\mathcal{F}_{\mathbf{G}}^{-1} g](\theta, \varphi, x_1, x_2) = [\mathcal{F}^{-1} \mathcal{K}_{\theta,\varphi} g](x_1, x_2). \tag{A.13}$$

An example of how such definitions are used in our work, we note that combining Eqs. (A. 8), (A.6), (A.3), (A.2), and (A.1) yields that, for all (ξ_1, ξ_2, x_3),

$$[\mathcal{F}_{\mathbf{V}} \iota](\xi_1, \xi_2, x_3) = \frac{1}{2\pi}. \tag{A.14}$$

The *3D Fourier transform operator* $\mathcal{F}$: $\mathbf{V} \to \mathbf{V}$ and its inverse $\mathcal{F}^{-1}$: $\mathbf{V} \to \mathbf{V}$ are defined by

$$[\mathcal{F}v](\xi_1,\xi_2,\xi_3)=(2\pi)^{-3/2}\int_{\mathbb{R}^3} v(x_1,x_2,x_3)\, e^{-i(\xi_1 x_1+\xi_2 x_2+\xi_2 x_2)}dx_1dx_2dx_3, \tag{A.15}$$

$$[\mathcal{F}^{-1}v](x_1,x_2,x_3)=(2\pi)^{-3/2}\int_{\mathbb{R}^3} v(\xi_1,\xi_2,\xi_3)\, e^{i(\xi_1 x_1+\xi_2 x_2+\xi_3 x_3)}d\xi_1d\xi_2d\xi_3. \tag{A.16}$$

Note that we use the same symbol for the 2D and the 3D Fourier transform operators, the actual operator that is being referred to is determined by the context.

The *2D convolution operator* *: $\mathbf{P} \times \mathbf{P} \to \mathbf{P}$ is defined by

$$[f_1 * f_2](x_1,x_2)=\int_{\mathbb{R}^2} f_1(x_1',x_2')f_2(x_1-x_1',x_2-x_2')\,dx_1'dx_2' \tag{A.17}$$

The *3D convolution operator* *: $\mathbf{V} \times \mathbf{V} \to \mathbf{V}$ is defined by

$$[f_1 * f_2](x_1,x_2,x_3)=\int_{\mathbb{R}^2} f_1(x_1',x_2',x_3')f_2(x_1-x_1',x_2-x_2',x_3-x_3')\,dx_1'dx_2'dx_3'. \tag{A.18}$$

We also need a convolution operator * that acts on two functions, $w \in \mathbf{W}$ and $v \in \mathbf{V}$. This is used for modeling distance-dependent blurring that occurs during data collection by an electron microscope, see Eq. (16). A function in **W** represents the collection of the rotations of an imaged object and the distance-dependent blurring of the electron microscope can be specified by a point spread function h in **V**. This convolution operator (whose outcome represents the distance-dependently blurred rotations of the molecule) is defined by

$$[w * h](\theta,\varphi,x_1,x_2,x_3)=[[\mathcal{K}_{\theta,\varphi,x_3}w] * [\mathcal{K}_{x_3}h]](x_1,x_2). \tag{A.19}$$

In the special case when h is the ι of Eq. (A.2), we get that, for $w \in \mathbf{W}$,

$$w * \iota = w. \tag{A.20}$$

This is easily seen using Eqs. (A.19), (A.17), (A.4), (A.3), and (A.1). We use the same symbol for the three different convolution operators, but the meaning is clear from the functions on which they are acting.

The *convolution theorem* provides a well-known relationship between the convolution operator in the spatial domain and multiplication in the frequency domain. For all $f_1 \in \mathbf{P}$, $f_2 \in \mathbf{P}$, and $(\xi_1, \xi_2) \in \mathbb{R}^2$, [13, Eq. (1.7)] implies that

$$[\mathcal{F}[f_1 * f_2]](\xi_1,\xi_2)=2\pi[\mathcal{F}f_1](\xi_1,\xi_2)[\mathcal{F}f_2](\xi_1,\xi_2). \tag{A.21}$$

The *division operator*/: $\mathbf{G} \times \mathbf{V} \to \mathbf{W}$ is defined by

$$[/(g,v)](\theta,\varphi,x_1,x_2,x_3)=\frac{g(\theta,\varphi,x_1,x_2)}{v(x_1,x_2,x_3)}. \tag{A.22}$$

Instead of/(g, v) we use the notation $\frac{g}{v}$.

B. Inversion of the projection operator

The purpose of this appendix is to justify the claim made in Eq. (15). We do this by appealing to known results presented in the book by Natterer and Wübbeling [13]. Essentially, all we need to do is to establish the relationship between our notation and that used in [13].

Let $S^2=\left\{(\beta_1,\beta_2,\beta_3)^T|\beta_1^2+\beta_2^2+\beta_3^2=1\right\}$ (the unit sphere in $\mathbb{R}^3$). For $\beta \in S^2$, let $\beta^\perp = \{\mathbf{x} \in \mathbb{R}^3 | \langle \mathbf{x}, \beta\rangle = 0\}$, where $\langle \mathbf{x}, \beta\rangle = x_1\beta_1 + x_2\beta_2 + x_3\beta_3$, for $\mathbf{x} = (x_1, x_2, x_3)^T$ and $\beta = (\beta_1, \beta_2, \beta_3)^T$ ($\beta^\perp$ is the plane through the origin perpendicular to β). Let $\mathbb{L} = \{(\beta, \mathbf{x})|\beta \in S^2, \mathbf{x} \in \beta^\perp\}$ (the set of all lines in $\mathbb{R}^3$). We introduce a new function space

$$\mathbf{L}=\{\ell:\mathbb{L} \to \mathbb{R}\}. \tag{B.1}$$

The *3D ray transform operator* $\mathcal{N}$: $\mathbf{V} \to \mathbf{L}$ is defined by Natterer and Wübbeling [13, Eq. (2.28)] by

$$[\mathcal{N}v](\beta,\mathbf{x})=\int_{\mathbb{R}} v(\mathbf{x}+t\beta)\,dt, \tag{B.2}$$

for all $\beta \in S^2$ and $\mathbf{x} \in \beta^\perp$. The *backprojection operator for the 3D ray transform* $\mathcal{W}$: $\mathbf{L} \to \mathbf{V}$ is defined by Natterer and Wübbeling [13, Eq. (2.31)] by

$$[\mathcal{W}\ell](\mathbf{x})=\int_{S^2} \ell(\beta,\mathbf{x}-\langle \mathbf{x},\beta\rangle\beta)\,d\beta, \tag{B.3}$$

for $\mathbf{x} \in \mathbb{R}^3$. We now prove that

$$\mathcal{BP}=\mathcal{WN}. \tag{B.4}$$

We first show that the operators $\mathcal{P}$ and $\mathcal{N}$ are in some sense equivalent. In Eq. (B.2) the integral through the object is taken along the line determined by a direction $\beta \in S^2$ and a point $\mathbf{y} \in \beta^\perp$; the line consist of the points in the set $\{\mathbf{y} + t\beta: t \in \mathbb{R}\}$. In Eq. (10) the object has been rotated by angles θ and φ and the integral is taken along a line parallel to the X_3-

axis. The two lines are really the same line through the object provided that β and $\mathbf{y}$ are appropriately defined in terms of θ, φ, and $(x_1, x_2)^T$.

To be mathematically precise, we show that, for all $v \in \mathbf{V}$, $\theta \in [0, 2\pi)$, $\varphi \in [0, \pi)$, and $(x_1, x_2)^T \in \mathbb{R}^2$, if

$$\beta=(-\cos\theta\sin\varphi, \sin\theta\sin\varphi, \cos\varphi)^T \tag{B.5}$$

and

$$\mathbf{y}=(x_1\cos\theta\cos\varphi+x_2\sin\theta, -x_1\sin\theta\cos\varphi+x_2\cos\theta, x_1\sin\varphi)^T, \tag{B.6}$$

then

$$[\mathcal{P}v](\theta, \varphi, x_1, x_2)=[\mathcal{N}v](\beta, \mathbf{y}). \tag{B.7}$$

This is done by first observing that, for β and $\mathbf{y}$ defined by Eqs. (B.5) and (B.6), respectively, we have by Eq. (6)

$$\mathbf{y}+x_3\beta=\begin{pmatrix} x_1^F(\theta, \varphi) \\ x_2^F(\theta, \varphi) \\ x_3^F(\theta, \varphi) \end{pmatrix}. \tag{B.8}$$

Eq. (B.7) follows, by starting with its right hand side and applying Eqs. (B.2), (B.8), (8), (9), and (10).

We now prove Eq. (B.4) by showing that, for all $v \in \mathbf{V}$ and $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3$,

$$[\mathcal{B}\mathcal{P}v](\mathbf{x})=[\mathcal{W}\mathcal{N}v](\mathbf{x}). \tag{B.9}$$

We first observe that, for β defined by Eqs. (B.5), it follows from Eq. (7) that

$$\mathbf{x} - \langle \mathbf{x}, \beta\rangle\beta=\begin{pmatrix} x_1^B(\theta, \varphi)\cos\theta\cos\varphi+x_2^B(\theta, \varphi)\sin\theta \\ -x_1^B(\theta, \varphi)\sin\theta\cos\varphi+x_2^B(\theta, \varphi)\cos\theta \\ x_1^B(\theta, \varphi)\sin\varphi \end{pmatrix}. \tag{B.10}$$

By applying Eqs. (13), (12), and (11), we get that

$$[\mathcal{B}\mathcal{P}v](\mathbf{x})=\int_{\mathbb{S}}[\mathcal{P}v](\theta, \varphi, x_1^B(\theta, \varphi), x_2^B(\theta, \varphi))\sin\varphi d\varphi d\theta. \tag{B.11}$$

To prove Eq. (B.9), we change the variables as indicated in Eq. (B.5) on the right hand side of Eq. (B.11) and combine Eqs. (B.6), (B.7), (B.10), and (B.3) to yield that $[\mathcal{B}\mathcal{P}v]$ (**x**) is equal to

$$\int_{S^2} [\mathcal{N}v](\beta, \mathbf{x} - \langle \mathbf{x}, \beta \rangle \beta)\, d\beta = [\mathcal{W}\mathcal{N}v](\mathbf{x}), \tag{B.12}$$

From Theorem 2.14 (with $\alpha = 1$) and Eqs. (1.15) and (1.7) of [13], it follows that, for all $v \in \mathbf{V}$,

$$v = \frac{1}{16\pi^3\sqrt{2\pi}} [\mathcal{F}^{-1}\widehat{r}] * \mathcal{W}\mathcal{N}v, \tag{B.13}$$

where, as before, $\widehat{r}(\xi_1, \xi_2, \xi_3) = \sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}$. Then Eq. (15) follows from Eqs. (B.13), (B.4), and (14).

C. Stationary phase approximation

The method of stationary phase [17, Chapter 1] is used for evaluation of highly oscillatory integrals of the form

$$I(\xi) = \int_{c_1}^{c_2} G(\sigma) e^{i\xi F(\sigma)}\, d\sigma. \tag{C.1}$$

If G is a smooth function, F is twice differentiable, and all critical points of F (i.e., points at which the first derivative F' is zero-valued) are nondegenerate (i.e., the second derivative F'' is not zero-valued), then as $\xi \to \infty$

$$I(\xi) = \sum_{s \in S} e^{i\xi F(s)} G(s) \sqrt{\frac{2}{\xi |F''(s)|}} e^{i\frac{\pi}{4} sgn(F''(s))} + O\left(\xi^{-1/2}\right), \tag{C.2}$$

where S is the set of critical points of F and sgn (a) denotes the sign of the argument a. This result implies that I (ξ) can be well approximated by the sum on the right hand side, as long as ξ is sufficiently large. However, it has been shown in many practical applications [6,11,18,19,20] that the sum approximates the original integral very well even for small values of ξ.

We consider the following a reasonable working tool based on the method of stationary phase. If two integrals of the kind shown in Eq. (C.1) differ only in G (they use the same ξ and F) and the values of G at the critical points of F are approximately the same, then the values of the integrals are also approximately the same.

Consider the two innermost integrals of Eq. (33),

$$\int_0^{2\pi}\int_0^{\pi} \frac{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}\sin\varphi e^{i\xi_1\Psi_1(\theta,\varphi)}d\varphi e^{i\xi_2\Psi_2(\theta)}d\theta. \tag{C.3}$$

Let us assume that ξ_1, ξ_1, x_1, x_2, x_3 are fixed. We define

$$I_1^\theta = \int_0^{\pi} G_1^\theta(\varphi) e^{i\xi_1 F_1^\theta(\varphi)} d\varphi, \tag{C.4}$$

where

$$G_1^\theta(\varphi) = \frac{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}{H(\xi_1,\xi_2,\widehat{x}_3^B(\theta,\varphi))}\sin\varphi, \tag{C.5}$$

$$F_1^\theta(\varphi) = \Psi_1(\theta,\varphi), \tag{C.6}$$

and

$$I_2 = \int_0^{2\pi} G_2(\theta) e^{i\xi_2 F_2(\theta)} d\theta, \tag{C.7}$$

where

$$G_2(\theta) = I_1^\theta, \tag{C.8}$$

$$F_2(\theta) = \Psi_2(\theta). \tag{C.9}$$

The iterated integral of Eq. (C.3) can be written, using these functions, as I_2. What we are going to show is that I_2 is (essentially) independent of the blurring function h or, equivalently, of its Fourier transom H; see Eq. (19). Our way of doing this relies on the following two facts. First, the critical points θ of $F_2 = \Psi_2$ satisfy

$$(x_1 - \widehat{x}_1)\cos\theta - (x_2 - \widehat{x}_2)\sin\theta = 0; \tag{C.10}$$

this follows from Eqs. (32) and (7). Second, for any $\theta \in [0, 2\pi)$, the critical points φ of F_1^θ satisfy

$$-(x_1-\widehat{x}_1)\cos\theta\sin\varphi+(x_2-\widehat{x}_2)\sin\theta\sin\varphi+(x_3-\widehat{x}_3)\cos\varphi=0; \tag{C.11}$$

see Eqs. (C.6), (31), and (7). From Eq. (7) follows that

$$x_3^B(\theta,\varphi)-\widehat{x}_3^B(\theta,\varphi)=-(x_1-\widehat{x}_1)\cos\theta\sin\varphi+(x_2-\widehat{x}_2)\sin\theta\sin\varphi+(x_3-\widehat{x}_3)\cos\varphi, \tag{C.12}$$

which, combined with Eq. (C.11), yields that, for all $\theta\in[0,2\pi)$, if φ is a critical point of F_1^θ, then $x_3^B(\theta,\varphi)-\widehat{x}_3^B(\theta,\varphi)=0$. We consider three cases.

Case 1. $x_1=\hat{x}_1$, $x_2=\hat{x}_2$, and $x_3=\hat{x}_3$.

This case has been dealt with in Section 5.

Case 2. $x_1=\hat{x}_1$, $x_2=\hat{x}_2$, and $x_3\neq\hat{x}_3$.

In this case it follows from Eq. (C.11) that the only critical point $\varphi\in[0,\pi)$ is $\varphi=\frac{\pi}{2}$. Furthermore, this critical point is nondegenerate, since $(x_3-\widehat{x}_3)\sin\frac{\pi}{2}\neq 0$. Observing Eq. (C. 5) and the statement following Eq. (C.12), we see that the stationary phase approximation applied to the integral in Eq. (C.4) implies that, for all $\theta\in[0,2\pi)$, I_1^θ is (essentially) independent of H. That I_2 is (essentially) independent of H follows from Eqs. (C.8) and (C. 7).

Case 3. Either $x_1\neq\hat{x}_1$, or $x_2\neq\hat{x}_2$, or both.

In this case there are exactly two critical points $\theta\in[0,2\pi)$ of F_2, they both satisfy Eq. (C. 10). We now show that they are both nondegenerate, by demonstrating that the alternative leads to a contradiction. If θ is a degenerate critical point of F_2, then by taking the second derivative of F_2 we get that

$$(x_1-\widehat{x}_1)\sin\theta+(x_2-\widehat{x}_2)\cos\theta=0. \tag{C.13}$$

Multiplying Eq. (C.10) by $\cos\theta$ and Eq. (C.13) by $\sin\theta$ and adding them together yields $x_1=\hat{x}_1$. Multiplying Eq. (C.10) by $-\sin\theta$ and Eq. (C.13) by $\cos\theta$ and adding them together yields $x_2=\hat{x}_2$. These together contradict the assumption of Case 3.

By the stationary phase approximation applied to the integral in Eq. (C.7), we see that if we could prove that $G_2(\theta)$ is (essentially) independent of H for the two critical points $\theta\in[0,2\pi)$ of F_2, then it would follow that I_2 is (essentially) independent of H. This is what we are going to do now to complete our proof.

It follows from Eqs. (C.10) and (C.11) that if θ is a critical point of F_2, then φ is a critical point of F_1^θ if, and only if,

$$(x_3-\widehat{x}_3)\cos\varphi=0. \tag{C.14}$$

Subcase 3a. $x_3=\hat{x}_3$.

In this subcase every $\varphi \in [0, \pi)$ is a critical point of F_1^{θ}. Hence Eqs. (C.8), (C.4), and (C.5), combined with the remark after Eq. (C.12), yield that $G_2(\theta)$ is (essentially) independent of H for the critical points $\theta \in [0, 2\pi)$ of F_2.

Subcase 3b. $x_3 \neq \hat{x}_3$.

In this subcase the only critical point $\varphi \in [0, \pi)$ of F_1^{θ} is $\varphi = \frac{\pi}{2}$. Furthermore, this critical point is nondegenerate, as can be seen by taking the second derivative of F_1^{θ}, applying Eq. (C.10) and observing that $(x_3 - \widehat{x_3})\sin\frac{\pi}{2} \neq 0$. Hence we can apply the stationary phase approximation to the integral in Eq. (C.4), and then Eqs. (C.8), (C.4), and (C.5), combined with the remark after Eq. (C.12), yield that $G_2(\theta)$ is (essentially) independent of H for the two critical points $\theta \in [0, 2\pi)$ of F_2.

This completes our proof that I_2 is (essentially) independent of the blurring function h or, equivalently, of its Fourier transform H; see (19). In other words, the integral in Eq. (C.3) and, hence, the integral in Eq. (33), is (essentially) independent of h, proving that $\mathcal{B}_h \mathcal{P}_h \kappa \approx \mathcal{B} \mathcal{P} \kappa$. This implies the validity of the concept of DGCBP, as expressed in Eq. (23).

D. Implementation

Jensen and Kornberg [4] altered the weighted backprojection method [9] to introduce the distance-dependent CTF correction in their implementation. Our implementation follows the mathematical developments of this paper and, hence, it is quite different in its details from the implementation in [4]

We make repeated use of the fast Fourier transform (FFT) and its inverse [21]. In using FFTs attention has to be paid to the scaling factor in Eq. (14) that is dependent on our definitions of Fourier transforms given in Eqs. (A.6)–(A.16), which may be scaled differently from the chosen FFT. Also, in order to reduce effects of various approximations in the implementation, intermediate working arrays could be set larger than the final array that needs to be produced. This can be done by padding the arrays representing projection data with zeros and working with such larger arrays through all the steps described below. In our implementation, we did not make use of this approach, since we were able to obtain acceptable results without it.

While our own actual implementation has been done in MATLAB, that choice is not essential and, most likely, is not optimal. Alternative computing environments can also be used to carry out the same sequence of logical steps, a description of which now follows.

We start with a projection data set $[\mathcal{P}_h v](\theta, \varphi, x_1, x_2)$ for a finite set M of (θ, φ) and, in each case, for an $N \times N$ square array of (x_1, x_2) with sampling distance Δ Å. Based on this projection data set, we need to estimate $[\mathcal{D}\mathcal{B}_h \mathcal{P}_h v](x_1, x_2, x_3)$ for an $N \times N \times N$ cubic array of (x_1, x_2, x_3) with sampling distance Δ Å. This is done in three stages.

Distance-dependent spreading back

In the first stage $[\mathcal{S}_h \mathcal{P}_h v](\theta, \varphi, x_1, x_2, x_3)$ is estimated as the $N \times N \times N$ cubic array of (x_1, x_2, x_3) with sampling distance Δ Å for all $(\theta, \varphi) \in M$. In our description we assume that x_3 is fixed; in practice, the process has to be carried out for each value of x_3.

First we calculate $H(\xi_1, \xi_2, x_3)$ for a square array of (ξ_1, ξ_2). Since we have an expression for $H(\xi_1, (\xi_2, x_3)$ in Eqs. (20) and (21), we can calculate its values for an $N \times N$ array with sampling distance $1/\Delta N$ Å^{-1}.

Applying the 2D FFT to the measured $N \times N$ array of samples of $[\mathcal{P}_h v](\theta, \varphi, x_1, x_2)$, we estimate $[\mathcal{F}_G \mathcal{P}_h v](\theta, \varphi, \xi_1, \xi_2)$ on an $N \times N$ array of (ξ_1, ξ_2) with sampling distance $1/\Delta N$ Å^{-1}. For each point (ξ_1, ξ_2) of this array, we define

$$\frac{[\mathcal{F}_G \mathcal{P}_h v](\theta, \varphi, \xi_1, \xi_2) H(\xi_1, \xi_2, x_3)}{(H(\xi_1, \xi_2, x_3))^2 + \alpha(\xi_1^2 + \xi_2^2)}. \tag{D.1}$$

We see from Eqs. (34) and (24) that, for our fixed (θ, φ) and x_3, the 2D inverse FFT of the array in Eq. (D.1) provides estimates of $[\mathcal{S}_h \mathcal{P}_h v](\theta, \varphi, x_1, x_2, x_3)$ for an $N \times N$ square array of (x_1, x_2) with sampling distance Δ Å.

By repeating this for all x_3, we obtain estimates of $[\mathcal{S}_h \mathcal{P}_h v](\theta, \varphi, x_1, x_2, x_3)$ as an $N \times N \times N$ cubic array of (x_1, x_2, x_3) with sampling distance Δ Å. This needs to be repeated for all $(\theta, \varphi) \in M$.

Distance-dependent backprojection via totaling

In the second stage $[\mathcal{B}_h \mathcal{P}_h](x_1, x_2, x_3) = [\mathcal{T} \mathcal{S}_h \mathcal{P}_h v](x_1, x_2, x_3)$ is estimated on an $N \times N \times N$ cubic array of (x_1, x_2, x_3) with sampling distance Δ Å by a discrete implementation of the totaling operator of Eq. (12) applied to the arrays obtained in the previous subsection.

First, for each value of $(\theta, \varphi) \in M$, we find estimates of

$$\rho_{\theta,\varphi}(x_1, x_2, x_3) = [\mathcal{S}_h \mathcal{P}_h v](\theta, \varphi, x_1^B(\theta, \varphi), x_2^B(\theta, \varphi), x_3^B(\theta, \varphi)), \tag{D.2}$$

for points of an $N \times N \times N$ cubic array with sampling distance Δ Å. According to Eq. (7), $\rho_{\theta,\varphi}(x_1, x_2, x_3)$ is obtained by unrotation by (θ, φ) of $[\mathcal{S}_h \mathcal{P}_h v](\theta, \varphi, x_1, x_2, x_3)$, which is the array defined at the end of the last sub-section. In our actual implementation the unrotation of the 3D voxel array is done by MATLAB's imrotate function with bicubic interpolation. There are potentially more accurate approaches; see, e.g., [22].

The integration in Eq. (12) is approximated by

$$[\mathcal{T} \mathcal{S}_h \mathcal{P}_h v](x_1, x_2, x_3) = \sum_{(\theta,\varphi) \in M} \rho_{\theta,\varphi}(x_1, x_2, x_3)\, d_{\theta,\varphi}, \tag{D.3}$$

where $d_{\theta,\varphi}$, can be chosen to be simply $4\pi/|M|$, which is valid if the sampling in (θ, φ) is reasonably uniform, or proportional to the size of the Voronoi neighborhood on the sampling point (θ, φ) on the unit sphere. Equation (D.3) is evaluated for an $N \times N \times N$ cubic array with sampling distance Δ Å.

Deblurring

The last stage is the implementation of Eq. (14) applied to the array of values obtained in the last section. A way to do this is to apply the 3D FFT and then to multiply the values in the resulting array at the point (ξ_1, ξ_2, ξ_3) by $\sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}$. To suppress the (less reliable) values at the higher frequencies we use a window function; in our implementation we used a window that sets the values to zero at points (ξ_1, ξ_2, ξ_3) whose distance from the origin in

Fourier space is greater than $^{0.9}/_{2\Delta}$. The required final $N \times N \times N$ cubic array of reconstructed values at points (x_1, x_2, x_3) with sampling distance Δ Å is obtained by an application of the inverse 3D FFT.

Display of the result

The described method looses information regarding the absolute (as opposed to relative) values in the object to be reconstructed due to its Fourier transform being set to zero at the origin. This problem is not purely mathematical: the CTF also causes the same if the value of a is (near) zero, see Eqs. (20) and (21). In order to obtain biologically useful information, the reconstructed values need to be mapped into displayed values in a careful manner.

To produce the illustrations for this paper, we applied an affine transformation to the reconstructed values obtained as in the last subsection in such a way that after the transformation the mean and standard deviation in the reconstruction are the same as in the phantom. After this the two arrays (phantom and reconstruction) are mapped into displayed values by the same rule.

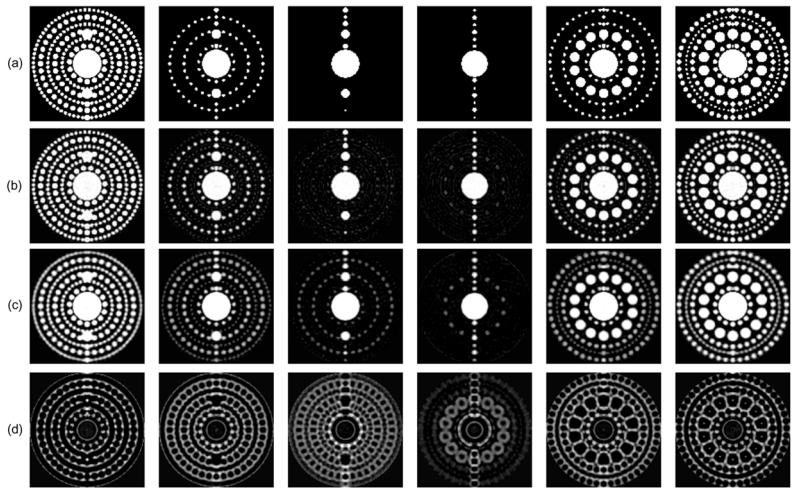

Figure 1.
Cross sections of a phantom (a) and of three reconstructions from projection data that were calculated with distance-dependent blurring: with correction for the distance-dependent contrast transfer function (b), with correction appropriate for the central layer of the specimen (c), and with no correction for the contrast transfer function (d).

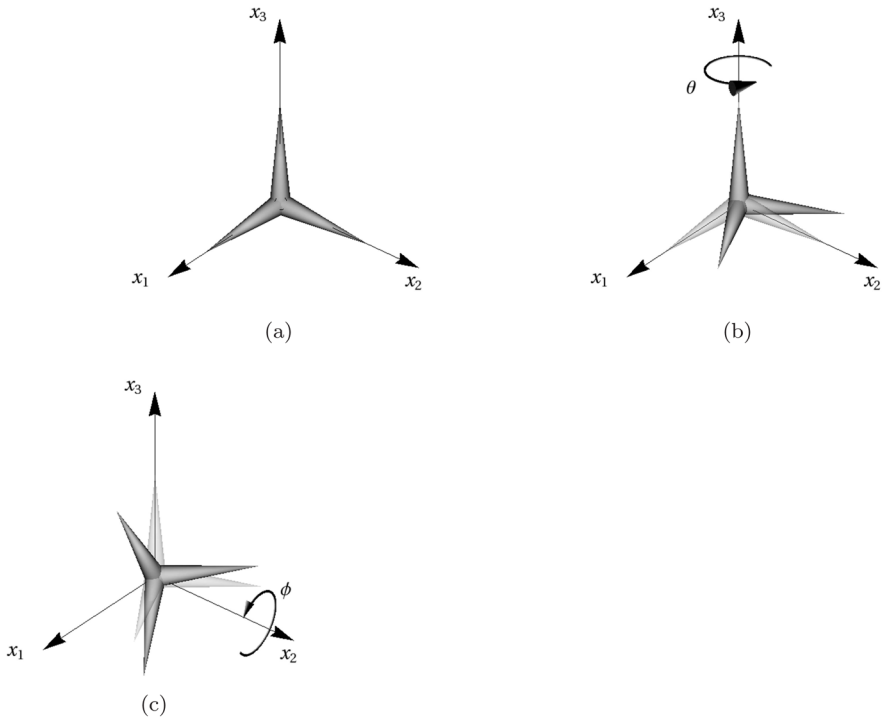

Figure 2.
Rotation of a molecule inside the microscope. (a) Molecule in standard position. (b) First rotation: by angle θ around the X_3-axis. (c) Second rotation: by angle φ around the X_2-axis.

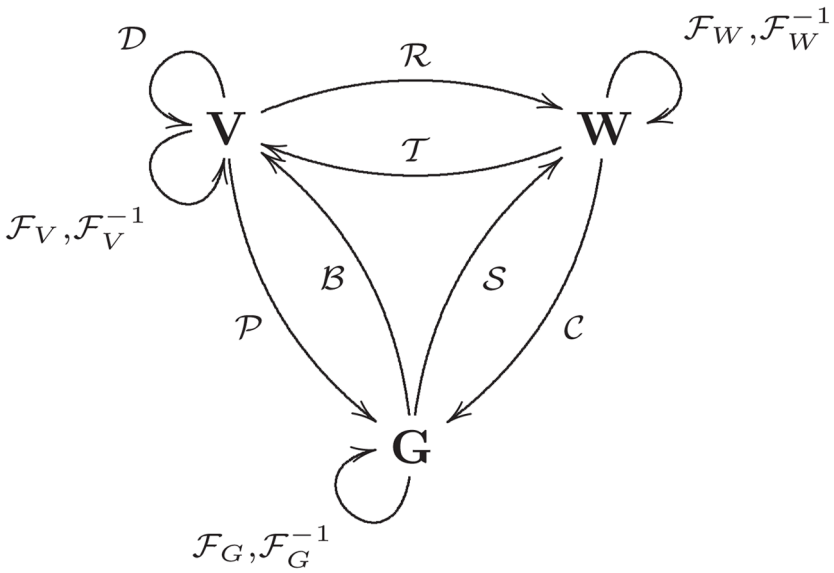

Figure 3.
Operators and the spaces of functions on which they act.

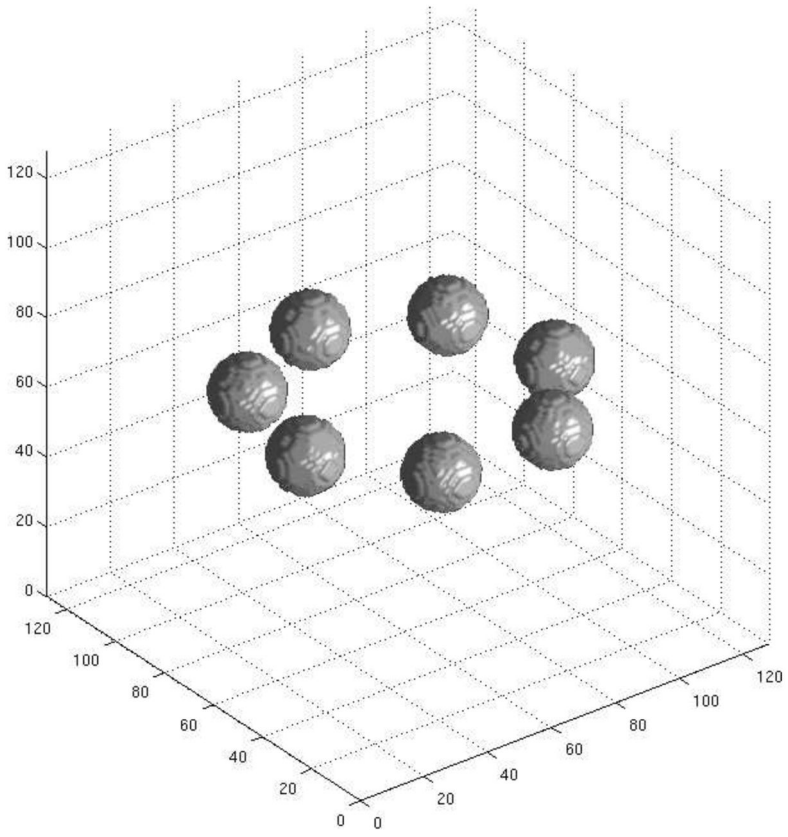

Figure 4.
Numerical test phantom.

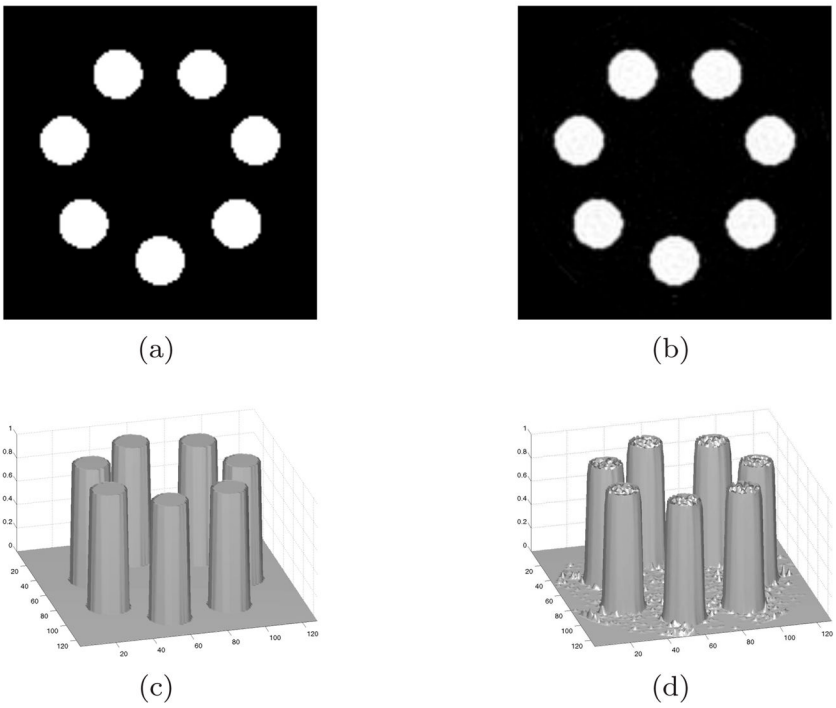

Figure 5.
(a) Image of a section of the phantom in Fig. 4. (b) Image of matching section of the DGCBP reconstruction from 5,000 distance-dependently blurred micrographs. (c) MATLAB 3D plot of (a). (d) MATLAB 3D plot of (b).

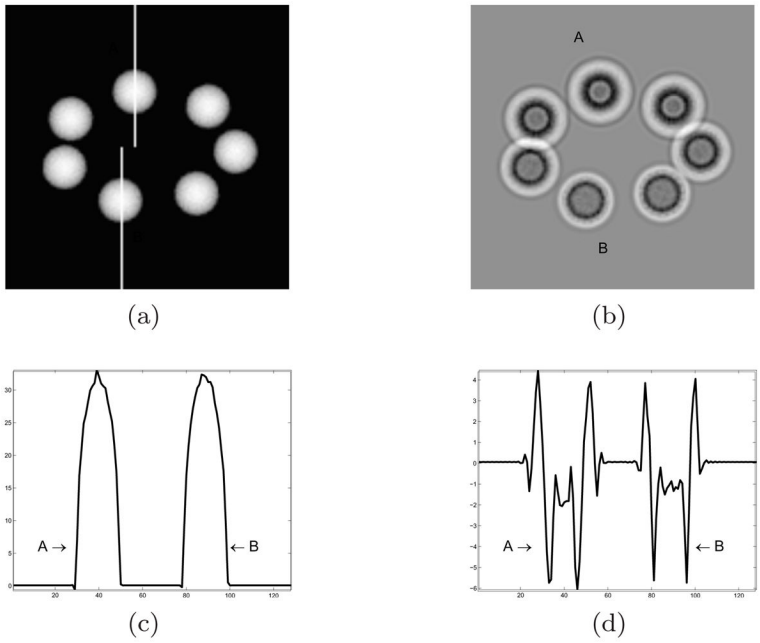

Figure 6.
(a) Image of a single unblurred projection of the phantom shown in Fig. 4; white lines indicate the pixels for which profiles are plotted. (b) Image of the matching distance-dependently blurred micrograph of the phantom. (c) Profiles for lines in (a). (d) Matching profiles in (b).

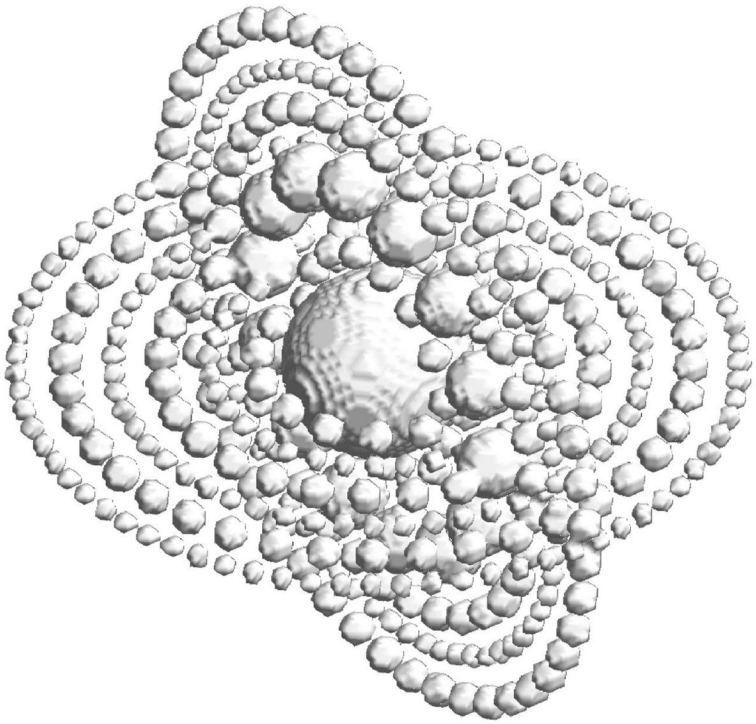

Figure 7.
Numerical test phantom composed of concentric rings of spheres of various sizes, digitized into a 128 × 128 × 128 voxel array.

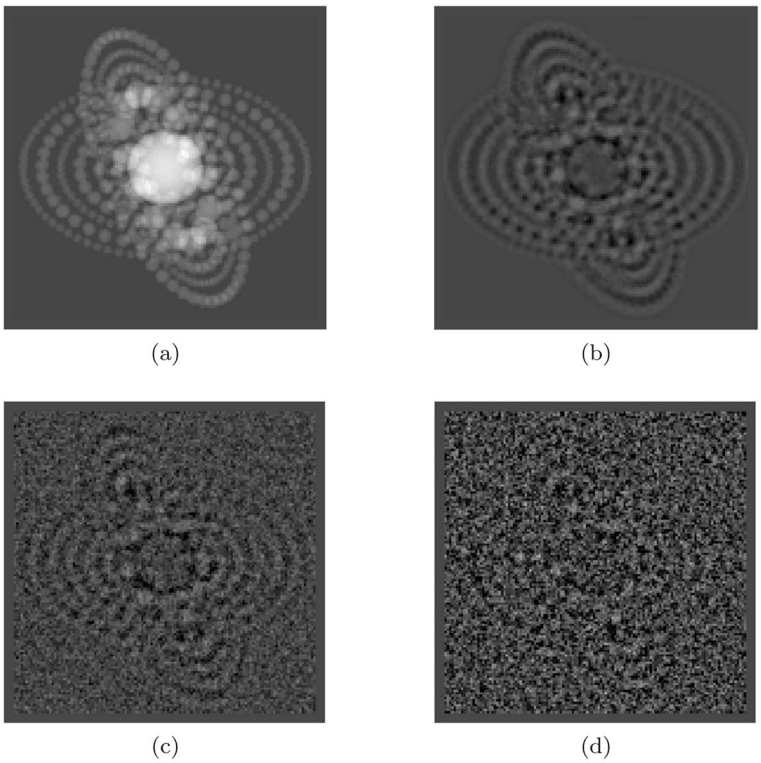

Figure 8.
A single projection of the phantom in Fig. 7: (a) ideal projection with no CTF blurring, (b) distance-dependently blurred projection, (c) distance-dependently blurred projection with added noise using $\sigma_1 = 0.3052$ and $\sigma_2 = 2.99$, (d) distance-dependently blurred projection with added noise using $\sigma_1 = 0.6103$ and $\sigma_2 = 6$.

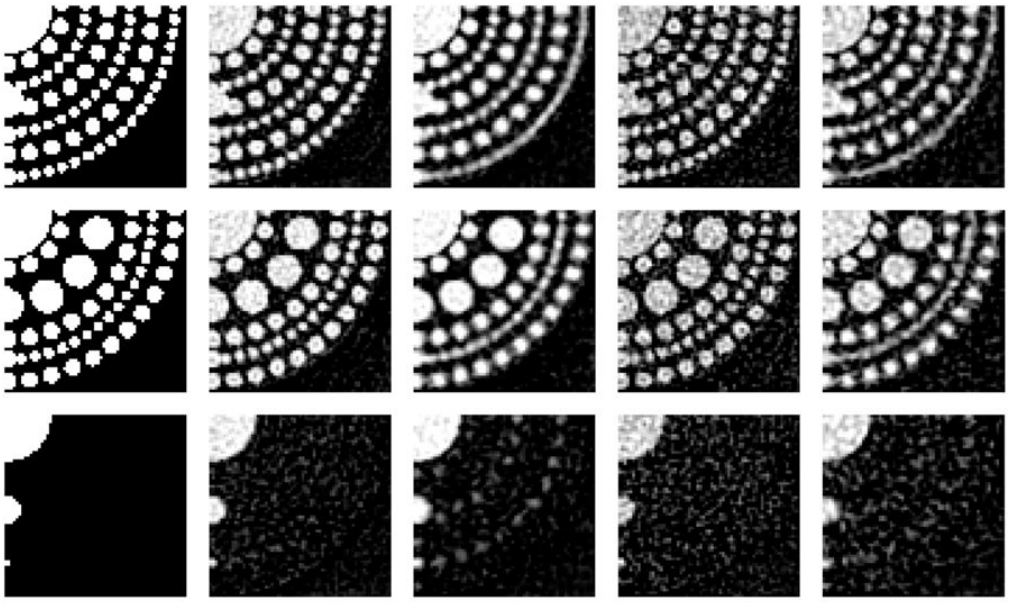

Figure 9.
Three different cross-sections of the phantom (first column), of the reconstructions from noisy projection data generated using $\sigma_1 = 0.3052$ and $\sigma_2 = 2.99$ obtained by DD backprojection (second column) and by CL backprojection (third column) and of the reconstructions from noisy projection data generated using $\sigma_1 = 0.6103$ and $\sigma_2 = 6$ obtained by DD backprojection (fourth column) and by CL backprojection (fifth column).

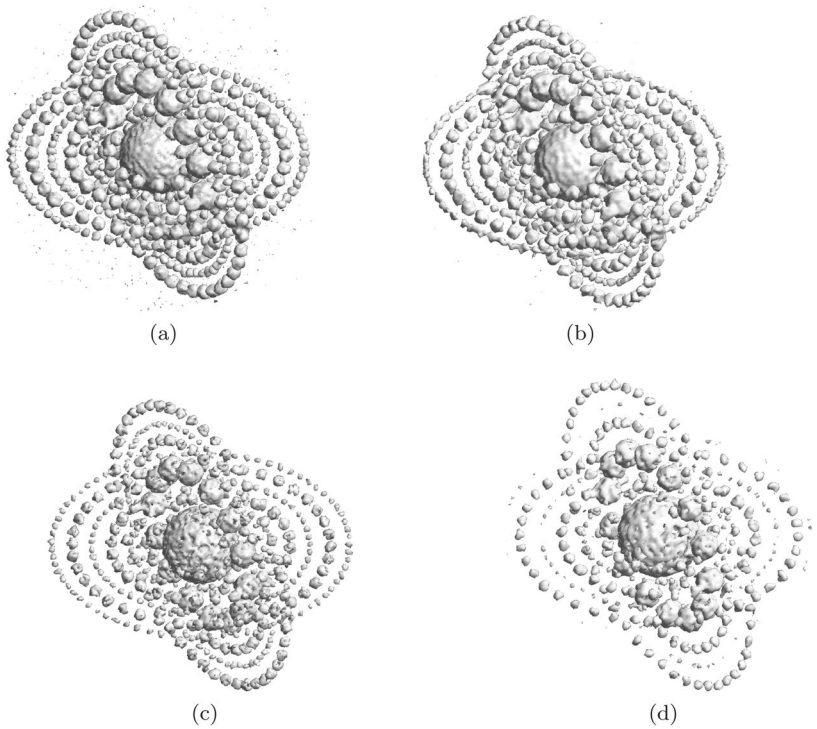

Figure 10.
Surface renderings of two reconstructions from the noisy projection data generated using $\sigma_1 = 0.6103$ and $\sigma_2 = 6$. (a) and (b) are rendered for voxel values thresholded at 0.5; (c) and (d) are rendered for voxel value thresholded at 0.9. (a) and (c) were obtained using DD backprojection; (b) and (d) were obtained using CL backprojection.

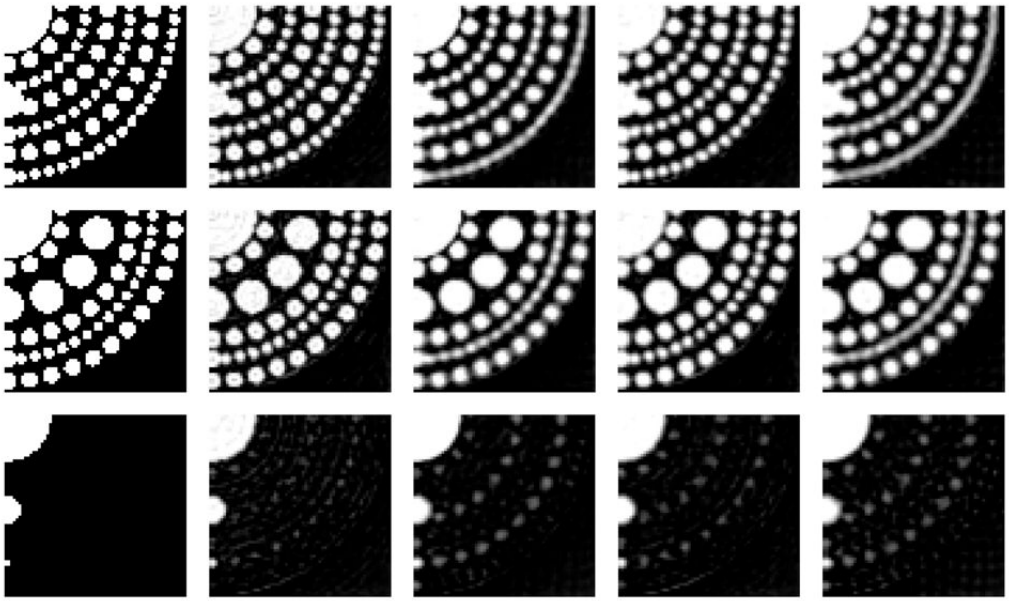

Figure 11.
Three different cross-sections of the phantom (first column), of the reconstructions from projection data with incorrectly determined defocus parameters using $\sigma_3 = 1$ and $\sigma_4 = 50$ obtained by DD backprojection (second column) and by CL backprojection (third column) and of the reconstructions from projection data with incorrectly determined defocus parameters using $\sigma_3 = 5$ and $\sigma_4 = 100$ obtained by DD backprojection (fourth column) and by CL backprojection (fifth column).

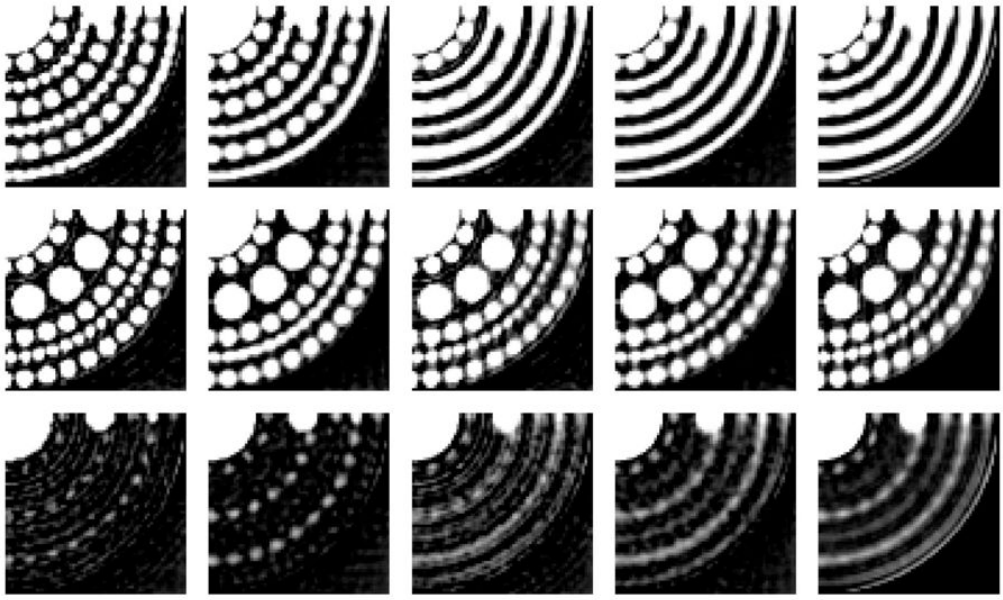

Figure 12.
Three different cross-sections of the reconstructions from projection data with incorrectly estimated projection angles using $\sigma_5 = 1$ obtained by DD backprojection (first column) and by CL backprojection (second column) and of the reconstructions from projection data with incorrectly estimated projection angles using $\sigma_5 = 3$ obtained by DD backprojection (third column), by CL backprojection (fourth column) and by backprojection from the data unaffected by a CTF (fifth column).