A memory efficient method for fully three-dimensional object reconstruction with HAADF STEM

A memory efficient method for fully three-dimensional object reconstruction with HAADF STEM

W. Van den Broek[a], A. Rosenauer[b], S. Van Aert[c], J. Sijbers[d], D. Van Dyck[c]

[a] *Electron and Ion Microscopy (ELIM), Ulm University, Albert-Einstein-Allee 11, 89081 Ulm, Germany*

[b] *Institut für Festkörperphysik (IFP), Universität Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany*

[c] *Electron Microscopy for Materials Science (EMAT), University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium*

[d] *iMinds - Vision Lab, Department of Physics, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium*

Abstract

The standard approach to object reconstruction through electron tomography is to reduce the fully three-dimensional problem to a series of two-dimensional slice-by-slice reconstructions. However, if the image of a single atom is to be incorporated as prior knowledge no such reduction is possible since this single atom image extends in two dimensions normal to said slice-by-slice reconstructions. Also for depth sectioning such reduction is impossible. Unfortunately, the size of the three-dimensional projection operator scales highly unfavorably with object size and readily exceeds the available computer memory. In this paper, it is shown that for incoherent image formation the memory requirement can be reduced to the fundamental lower limit of the object size, both for tomography and depth sectioning. Furthermore, it is shown through multislice calculations that high angle annular dark field scanning transmission electron microscopy can be sufficiently incoherent for the reconstruction of single element nanocrystals, but that dynamical diffraction effects can cause classification problems if more than one element is present.

Keywords: HAADF STEM, Depth sectioning, Multislice simulations, Object reconstruction, Tomography
PACS: 07.05.Pj, 42.25.Bs, 42.25.Fx, 42.30.Va, 42.30.Wb

1. Introduction

In tomography a three-dimensional object is retrieved from a set of two-dimensional images recorded at various tilt angles. The underlying assumption of conventional tomography is that the images must fulfill the projection requirement [1] which states that the image intensities must be a linear (or at least monotonic) function of a local property of the object integrated along the beam direction. In [2, 3, 4] the fact that conventional transmission electron microscopy does not obey this requirement is circumvented by redesigning the reconstruction algorithm. However, in this paper the for materials science more common approach is investigated where the microscope is operated in HAADF STEM mode (high angle annular dark field scanning transmission electron microscopy) and a conventional tomographic algorithm, that relies on the projection requirement, is used for reconstruction.

At nanometer resolution the HAADF STEM image intensity scales (approximately [5]) linearly with the line integral of the atomic number raised to a power of about 1.7 [6, 7]. In [8] it is argued that for HAADF STEM at atomic resolution the projection requirement demands that the image of any two atoms equals the sum of the image of each atom; which is another way of saying that the imaging process must be incoherent. That atomic resolution HAADF STEM can be approximately incoherent is shown theoretically by Hartel et al. in [9, 10]. Recent advancements, for example in [11, 12, 13], have clearly shown that HAADF STEM tomography at atomic resolution is also experimentally feasible.

The image of a single atom is often dominated by the electron beam shape and the shape of the atoms in a reconstruction will therefore reflect the electron beam conditions rather than the atomic potential. In [8], Van den Broek et al. have shown that this problem can be alleviated by including the shape of the electron beam and atomic potential as prior knowledge in the projection process. As a result the object to be reconstructed becomes sparse and is composed of Dirac-delta functions on the atom positions and zero in the vacuum. This sparsity can subsequently be exploited by an iterative reconstruction algorithm in which low values are forced to zero in each iteration. In this way this work ties in with compressed sensing [14, 15, 16].

As shown in [17] the same reconstruction technique can be applied to image modes like depth sectioning [18, 19, 20] and scanning confocal electron microscopy (SCEM) [21], where three-dimensional information is extracted solely by varying the defocus. By including these techniques in the realm of

reconstruction techniques, their potential can be increased dramatically.

Although including the shape of the electron beam and atomic potential as prior knowledge in the projection process yields significant benefits, it is difficult to realize for systems larger than a few tenths of nanometers because of the prohibitive computer memory load that comes with it. In this paper we show in Section 2 how the memory load can be reduced to the fundamental lower bound of the mere storage of the object. This is illustrated in Section 3 with reconstructions of objects of several nanometers wide from simulated tomographic data sets and a simulated depth sectioning data set. Furthermore, because the tomographic data sets are obtained through a multislice algorithm [22, 23, 24, 25, 26], they reveal that HAADF STEM is sufficiently incoherent for the reconstruction of single element nanocrystals, but that dynamical diffraction effects can cause severe classification problems when more than one element is present. Finally, in Section 4 the conclusions are drawn.

2. Memory efficient algebraic reconstruction

Algebraic reconstruction techniques (ART) [27] are often preferred over faster methods like filtered back projection because they can yield results with lower noise and less artifacts and can more easily handle incomplete data like missing wedges [28, 29, 30]. However, besides being slower due their iterative nature they come with the disadvantage of a heavy memory load: The projection process in ART is described as a matrix product of the weight matrix A with the unknown object vector f, yielding the measurement vector p, where the dimensions of A scale highly unfavorable with the object size.

Usually this three-dimensional problem is reduced to a series of two-dimensional problems by dividing the two-dimensional projections in lines perpendicular to the tilt axis from which the object is reconstructed slice-by-slice. The memory load is reduced even further by exploiting the fact that most elements of A are zero. However, if A is to take the shape of the image of a single atom into account—which has been shown to be highly advantageous in [8]—no slice-by-slice reconstruction is possible because these atomic images extend well beyond the width of the lines the projections would be split into. Furthermore, the sparsity of A is severely reduced as well.

In [17], ART is shown to be applicable to depth sectioning too, but the size of the matrix A is even larger than in the case of three-dimensional

tomography and prohibits the reconstruction of all but the smallest objects.

In the next two sections it will be shown how the matrix multiplications occurring in the multiplicative simultaneous iterative reconstruction technique (SIRT) [28, 29, 31] can be carried out implicitly, thereby reducing the load on the memory to the fundamental lower limit of the size of the object to be reconstructed.

2.1. Tomography

Following [28], the multiplicative SIRT algorithm can be written as

$$f_{k+1} = f_k A^T \left[\frac{q}{A f_k} \right], \tag{1}$$

where k indicates the iteration number and q is the experimental projection corrupted by Poisson noise; arithmetic operators between vectors are elementwise and each column sum of A equals 1. It is shown in [28] that when initiated with a constant value for f_0, (1) does converge to the maximum likelihood estimate, albeit slowly.

It is shown in [8] that the image of a single atom can be accounted for by the projection matrix A. As a consequence, the object f can be written as a collection of Dirac delta-functions located at the atom positions and surrounded by vacuum. This image formation process is illustrated in Figure 1. An underlying assumption is that the image formation is incoherent, i.e. the image of the object equals the sum of the single atom images. The validity of this assumption is elaborated upon further in Section 3. For a better understanding in the following A will describe a single projection angle and the more general case of multiple projection angles will be inferred later on.

In Figure 2 the size of A is derived. Suppose the object has $N \times N \times N$ voxels and the projection measures $N \times N$ pixels, then A will have a dimensionality of $N^2 \times N^3$. Fortunately, the number of non-zero elements scales with N^3 instead of N^5. The proportionality factor, however, equals the amount of pixels in the single atom image. To keep this number down, the single atom image is set to zero outside a disc with diameter d_p, containing $p\%$ of the total single atom image intensity, where typical values for p would be 80, 90 or 95. In Table 1 it is shown that for typical HRSTEM settings the proportionality factor ranges from about 200 to about 2000, rendering A unpractically large and forcing one to make a trade-off between model accuracy and memory load.

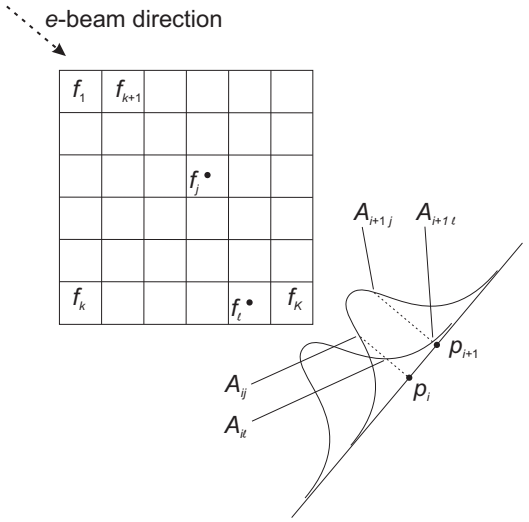

Figure 1: Illustration of the image formation process in tomography when the projection matrix A takes the atomic images into account. The object contains two Dirac delta-functions in voxels j and ℓ while the projection p contains extended atomic images.

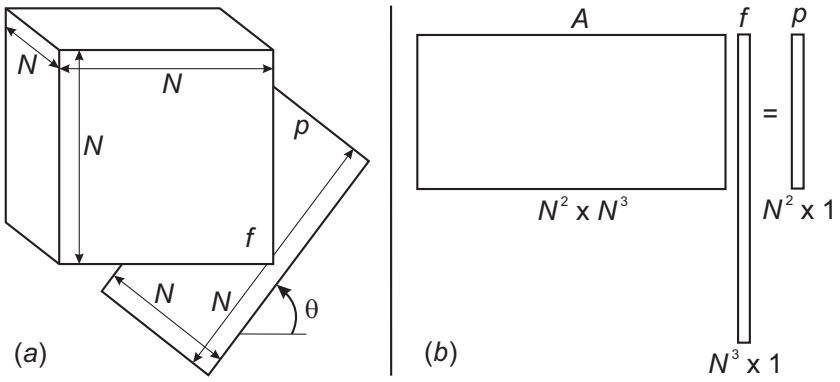

Figure 2: Derivation of the size of the projection matrix A. *2.a*: Relative orientations of the $N \times N \times N$ object f and the $N \times N$ projection p; the tilt angle is θ. *2.b*: Mathematical implementation of the projection process; f and p are written as column vectors related by the projection matrix A.

Table 1: Number of pixels (No. pix.) within the disc of diameter d_p containing p% of the total intensity of the image of a single Au atom. The pixels are 25 pm wide. See Tab. 2 for the simulation parameters.

p	50	80	90	95	99
d_p (nm)	0.14	0.43	0.78	1.3	4.4
No. pix.	25	232	765	2124	24328

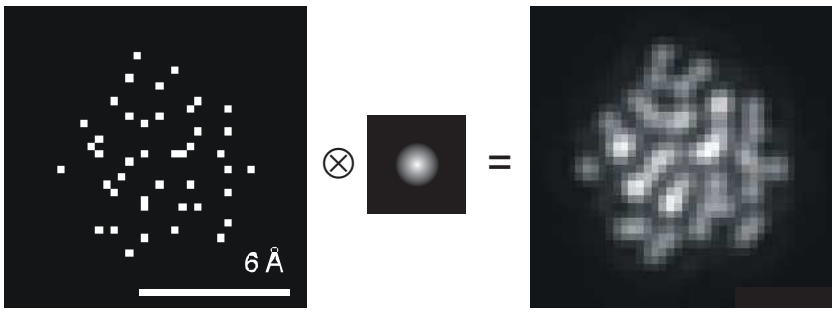

Figure 3: Illustration of the implicit matrix-vector product Af_k in (1). The object f (an amorphous Si-cluster taken from [8]) is projected slice by slice and the resulting two-dimensional projection is convolved with the two-dimensional single image intensity.

Calculation of the large matrix A can be avoided by carrying out the matrix-vector multiplications in (1) implicitly if one can work out what physical processes these multiplications describe.

The operation Af_k can rather straightforwardly be interpreted as a slice by slice projection of the object, followed by a convolution of the two-dimensional projection with the two-dimensional single atom image, as illustrated in Figure 3. Let A' denote the matrix that projects one $N \times N$ slice of f onto one $N \times 1$ column of p. The matrix A' thus has size $N^2 \times N$ with $\propto N^2$ non-zero elements, where the proportionality factor is approximately 1. For the subsequent two-dimensional convolution, only two $N \times N$ images need to be held in memory: the two-dimensional projection and the single atom image. An additional and important benefit is that the single atom image does not need to be cut-off beyond d_p.

For the interpretation of the matrix-vector multiplication $A^T[q/(Af_k)]$ in (1), consider that every row of A^T contains the single atom image centered on the projected position of the corresponding object voxel. This can be inferred from multiplying A with a vector with just one non-zero entry. The multipli-

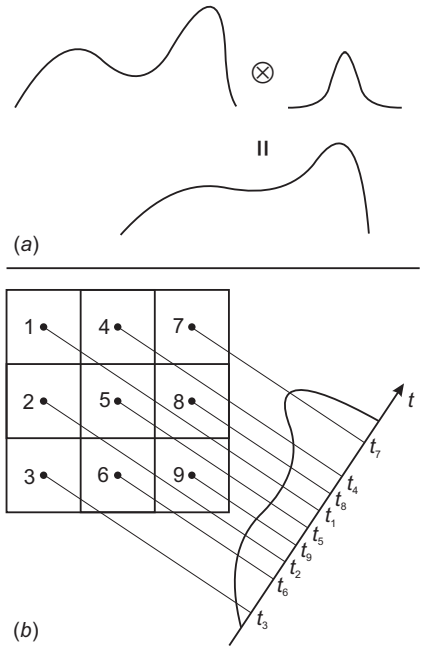

Figure 4: Illustration of the implicit matrix-vector product $A^T[q/(Af_k)]$ in (1). *4.a*: The image $q/(Af_k)$ is convolved with the single image intensity. *4.b*: The result from the first step is interpolated in the projected positions t_j of each of the N^3 object voxels j.

cation of A^T with the $N^2 \times 1$ vector $q/(Af_k)$ can therefore be interpreted as a convolution of an $N \times N$ image with the single atom image, followed by an interpolation in the projected positions of each of the N^3 object voxels. This is illustrated in Figure 4. In practice, a nearest neighbor interpolation proves to be sufficient. For the convolution step, again only two $N \times N$ images need to be held in memory, and the interpolation can be done consecutively over each of the projected positions of the N^3 object voxels.

The memory load is now brought back from up to a few thousand times N^3 to just N^3, corresponding to the mere storage of the object, which surely is a fundamental lower bound. And no trade-off between model accuracy and memory load exists anymore since no cut-off of the single atom image is necessary.

The more general case of more than one projection angle can be treated by computing $A^T[q/(Af_k)]$ implicitly for each projection. The results are summed and divided by the number of projections before elementwise multiplication with f_k. Division by the number of projections is necessary to keep

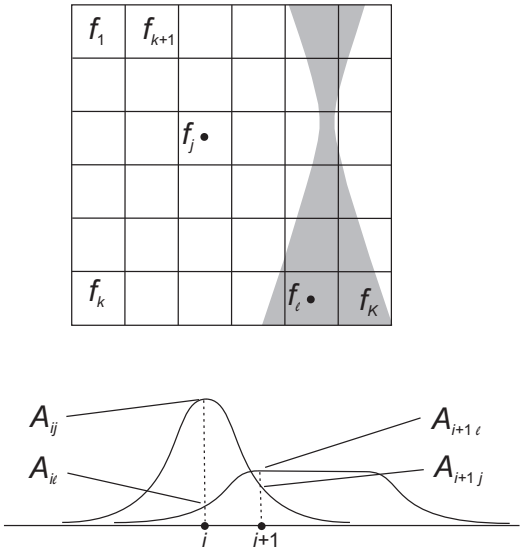

Figure 5: Illustration of the image formation process in depth sectioning. The object contains two Dirac delta-functions in voxels j and ℓ while the projection p contains extended atomic images brought about by a focused and a non-focused beam, respectively. The projection matrix A accounts for both the focused and the non-focused images.

the column sums of the total projection matrix A equal to one.

2.2. Depth sectioning

In depth sectioning the object is illuminated with a beam with a small depth of field, caused by a large beam convergence angle. Features of the object close to the beam crossover are fairly well localized in the resulting image while features further away are smeared out and will constitute a non-distinct background. Depth information is unlocked by varying the defocus, thereby bringing other regions of the object in focus. This image formation process is illustrated in Figure 5. Like before, an underlying assumption is that the image formation is incoherent, i.e. the image of the object equals the sum of the single atom images.

In [17], it is shown that the SIRT algorithm applies to depth sectioning as well and atom positions can in principle be retrieved. Preliminary results have shown that for depth sectioning an additive SIRT algorithm is preferable. Adapted from [31], we arrive at

$$f_{k+1} = f_k + \frac{A^T\left[\left(q - Af_k\right)/A\mathbb{I}_f\right]}{A^T\mathbb{I}_p}, \tag{2}$$

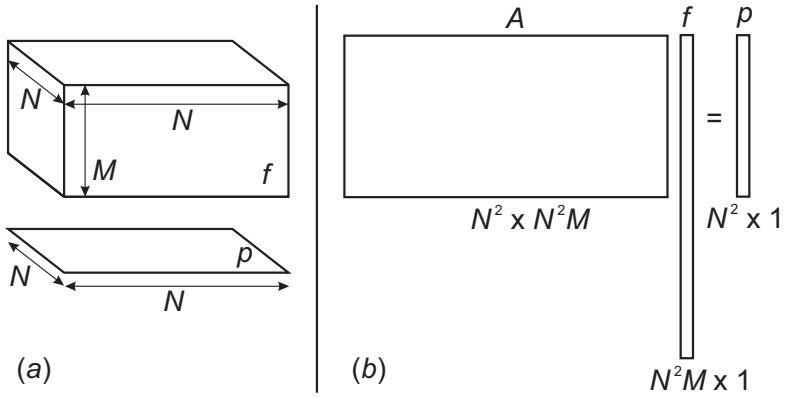

Figure 6: Derivation of the size of the projection matrix A in depth sectioning. *6.a*: Relative orientations of the $N \times N \times M$ object f and the $N \times N$ projection p. *6.b*: Mathematical implementation of the projection process; f and p are written as column vectors related by the projection matrix A.

where k indicates the iteration number and q the experimental projection, arithmetic operators between vectors are elementwise and $\mathbb{I}_p$ and $\mathbb{I}_f$ each denote a vector of which each element is 1 with a length equal to that of the projection and the object, respectively.

In the following A denotes the projection matrix corresponding to a single defocus value. In this paper the more usual case of several defoci is dealt with through so-called online optimization [32] where the solution is updated after each defocus instead of being updated with the mean of the corrections obtained from the complete set of defoci. As suggested in [28] the defoci are addressed in random order to speed up convergence.

It can be seen in Figure 5 that the image of a single atom depends on its depth relative to the beam crossover (i.e. the defocus) and is predominantly (apart from the atomic potential) given by the beam profile at that specific depth in the sample. If the object measures $N \times N \times M$ voxels and the projection measures $N \times N$ pixels, then the projection matrix A has dimensions $N^2 \times N^2 M$, see Figure 6. The matrix A has $\propto N^2 M$ non-zero values and the proportionality constant is the average number of pixels in the single atom image; as the beam widens away from focus, the number of pixels in the single atom image will be even higher than in the tomographic case, rendering A even less sparse.

The matrix-vector multiplication Af_k can now be done implicitly by a two-dimensional convolution of each horizontal layer of the object with its

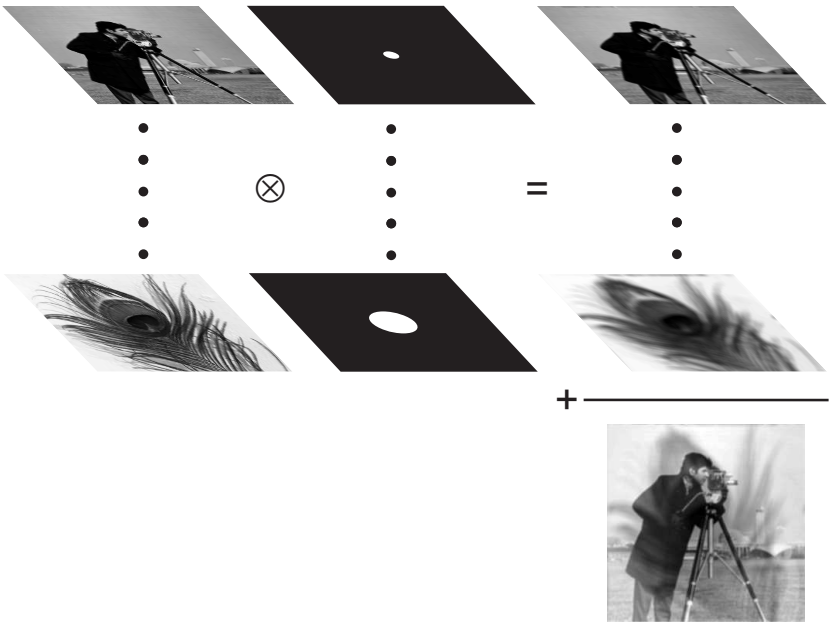

Figure 7: Illustration of the implicit matrix-vector product Af_k in (1) for depth sectioning. Each horizontal layer of the object is convolved with the corresponding single atom image, in this case the upper layer is in focus (convolved with a narrow image) and the lower layer (image taken from [33]) is out of focus (convolved with a wide image). The result is summed in the vertical direction and the image is dominated by the upper layer.

corresponding single atom image, followed by a sum in the vertical direction over all horizontal layers. This is illustrated in Figure 7. The two-dimensional convolutions can be done consecutively and the respective single atom images could be calculated on-the-fly, so that the memory capacity is limited only by the $N \times N \times M$ object. Since $\mathbb{I}_f$ has the same length as the object f_k, implementation of the matrix-vector multiplication $A\mathbb{I}_f$ is completely analogous to Af_k.

For the interpretation of the matrix-vector multiplication $A^T[(q-Af_k)/A\mathbb{I}_f]$, consider that every row of A^T contains the single atom image corresponding to the height in the sample of the corresponding object voxel, centered on this voxel's horizontal position. This can be inferred from the multiplication of A with a vector with just one non-zero entry. The multiplication of A^T with the $N^2 \times 1$ vector $(q - Af_k)/A\mathbb{I}_f$ can therefore be carried out implicitly by stacking the $N \times N$ image into an $N \times N \times M$ array and convolving each of the M layers with the corresponding single atom image. The explicit

interpolation that occurred in the tomographic case is not necessary here because the positions of the object voxels in the horizontal directions coincide with those of the image pixels. The matrix-vector multiplication $A\mathbb{I}_p$ can be implemented analogously because $\mathbb{I}_p$ has the same length as the projection q.

The memory load is now brought back to N^3, the mere storage of the object. Furthermore, no trade-off between model accuracy and memory load has to be made since the single atom image does not need to be cut-off.

D'Alfonso et al. show in [34] that the image formation in incoherent SCEM can be approximated as a convolution of the object function with an effective SCEM probe, thus making it clear that by simply choosing a different expression for the image forming beam, the formalism for depth sectioning can be applied to SCEM as well.

3. Simulations

In this section three simulations are carried out, serving two purposes, firstly, to show that the implicit matrix-vector multiplications as explained in Section 2 are correct, and secondly, to investigate if the image formation in HAADF STEM is sufficiently incoherent to allow for object reconstruction at the atomic scale.

3.1. Tomography

Since the projection matrix A accounts for the single atom image, the object f is sparse, i.e. only voxels containing an atom differ from zero. This sparsity is enforced by putting to zero every gray value below a certain threshold in every iteration of the reconstruction process. Because the expression in (1) is multiplicative, these voxels remain zero in subsequent iterations. For small objects this proved to be sufficient to retrieve the correct non-zero voxels and their values [8, 17]. For larger objects like in this study, this algorithm converges too slowly to be practical. Therefore, a heuristic adaptation is made. Since after a certain number of iterations the atoms are already distinguishable as clusters that have not condensed into a single voxel yet, a watershed algorithm [35] is used to locate the voxels containing these clusters' centers of mass, all other voxels are set to zero. The algorithm then runs until the gray values converge. The convergence criterion is that the maximum of the absolute differences between corresponding non-zero voxels of successive iterations is smaller than one thousandth of the minimum non-zero voxel of the current iteration.

Table 2: Microscope parameters used in the simulations in Secs. 2.1 and 3.1. U is the acceleration voltage, C_1 the focus value, C_3 the spherical aberration, α the convergence semi-angle, and θ_{in} and θ_{out} the inner and outer radius of the annular detector.

U	C_1	C_3
300 kV	-49 nm	1.2 mm
α	θ_{in}	θ_{out}
9.0 mrad	33 mrad	200 mrad

The projections in this section are computed with a multislice calculation by the STEMsim software [36] which has a dedicated module for tilt series. The program reads a text file in which the positions and types of the atoms are defined. For each tilt angle the nanocluster is rotated and cut into slices used in the multislice simulation. The thermal diffuse scattering is taken into account with an absorptive potential [37]. The single atom image is taken as the intensity of the convolution of the atomic potential with the probe's wave function. The reconstructions serve the double purpose of checking whether the implicit matrix multiplications are correct and if the HAADF STEM image formation allows for a good reconstruction.

3.1.1. Gold nanocrystal

The Au particle is shaped like a cuboctahedron, see Fig. 8, consists of 921 Au atoms arranged in an FCC lattice and has a vacancy in the center and on one of the surfaces. The lattice parameter is 0.408 nm and the Debye-Waller factor is 0.6331 Å^2. The microscope parameters are given in Table 2 and the pixels of the images and voxels of the reconstruction are 25 pm wide. The angles of the 45 projections range from 2° to 178° in steps of 4°. Poisson noise was applied to the images, causing a signal-to-noise ratio of 3, see Figure 9.

All atom positions are retrieved without falsely attributing an atom to one of the vacancies. The watershed algorithm was applied after 45 iterations and the reconstruction converged after 92 iterations. Non-zero voxels in the reconstruction were attributed to the closest atom positions in the original object. The accuracy of the atom positions was assessed by the root mean squared difference (RMSD) between the true atom position and the center of the voxel the atom is attributed to by the reconstruction. The RMSD is $0.58d$, with $d = 25$ pm the voxel size. If each atom were assigned to the

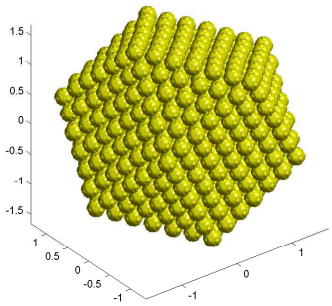

Figure 8: Visualization of the Au particle. The axes are calibrated in nm.

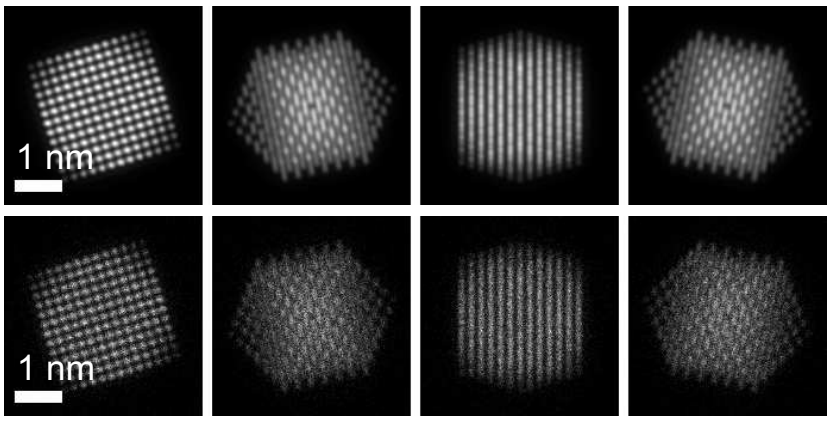

Figure 9: *Upper row*: Four typical noiseless projections of the Au particle. From left to right the tilt-angles are 2°, 46°, 90° and 134°. *Lower row*: The same images with Poisson noise, the signal-to-noise ratio is 3.

Table 3: Microscope parameters for the PbSe-CdSe particle. C_5 is the fifth order spherical aberrations, the other symbols are defined in Tab. 2.

U	C_1	C_3	C_5
300 kV	22.3 nm	−0.298 mm	75 cm
α	θ_{in}	θ_{out}	
20 mrad	50 mrad	115 mrad	

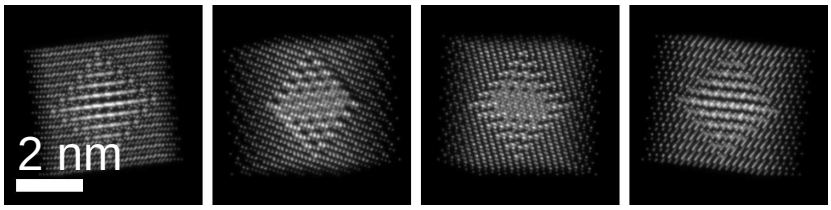

Figure 10: Four typical images of the PbSe-CdSe core-shell particle with a signal-to-noise ratio of 30. From left to right the tilt-angles are 10°, 60°, 110° and 160°.

correct voxel, the RMSD would be $d/2$.

3.1.2. PbSe-CdSe core-shell particle

The particle is derived from a CdSe rock-salt structure and has a cubic shape with sides of 3.7 nm. The unit-cell parameter a equals 0.61 nm. The Cd atoms in the cube's interior octahedron are replaced by Pb atoms and shifted by a vector $[-a/4,\ a/4,\ a/4]$; the structure is explained further in [38]. The total number of atoms is 1963 and the Debye-Waller factors are 1.465 Å^2, 1.570 Å^2 and 5.093 Å^2 for Se, Cd and Pb.

The microscope parameters are given in Table 3, the pixels of the projection and the voxels of the reconstruction are 16.4 pm wide. The angles of the 18 projections range from 0° to 170° in steps of 10°. Poisson noise with a signal-to-noise ratio of 30 was applied to the projections, see Figure 10. The image of a single atom, needed as prior knowledge, was set to the average of the image of a Se atom, a Cd atom and a Pb atom.

The reconstruction algorithm retrieves all atom positions with a RMSD of $0.74d$. The watershed algorithm was applied after 181 iterations and the reconstruction converged after 2896 iterations. The histogram of the gray values of the non-zero voxels exhibits three peaks corresponding to Se, Cd and Pb, see Figure 11. Atom types were assigned to the non-zero voxels by

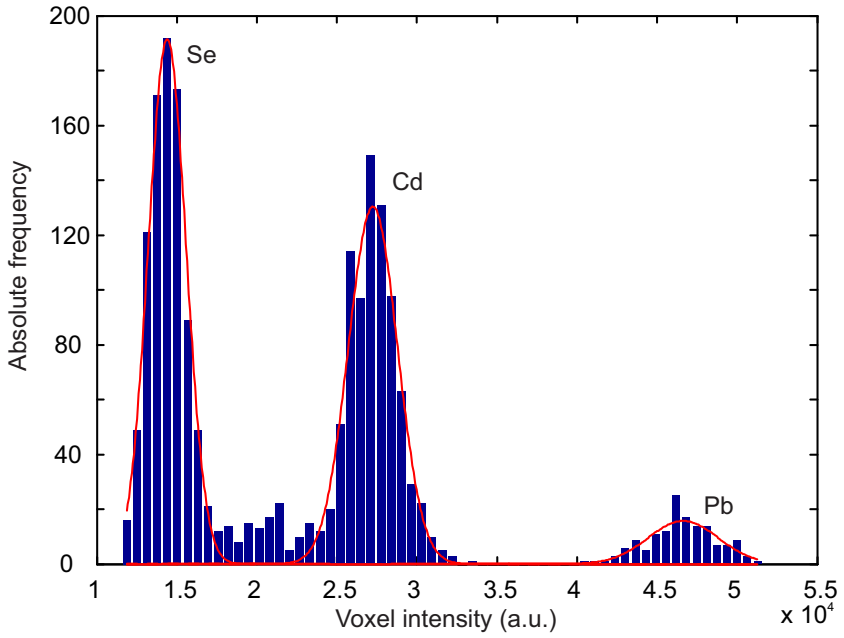

Figure 11: Histogram (blue bars) of the intensities of the non-zero voxels in the reconstruction of the PbSe-CdSe core-shell particle. To each of the three peaks a Gaussian peak has been fitted (red line), revealing a strong deviation from Gaussianity between the Se-peak and the Cd-peak.

thresholding the gray values. In Table 4 it is shown that if the thresholds are chosen as the values half-way between successive peaks 13% of Se-atoms are misclassified as Cd-atoms, while for the other atom types no misclassification occurs. In principle, the classification error for Se can be reduced by choosing a larger Se-Cd threshold, but at the cost of an increased classification error for Cd. In practice however, one would have no reason to choose the higher threshold since in an experimental situation the true object and thus the classification error are unknown.

In [8], it is shown that the estimated gray values follow a Gaussian distribution when the imaging is fully incoherent. For comparison, a Gaussian profile has been fitted to each of the peaks, and between the Se-peak and the Cd-peak a significant deviation from Gaussianity can be observed, thus proving that dynamical scattering still plays an important role. Further, notice that the classification error is not reciprocal: The amount of Se that is misclassified is by no means compensated by the amount of misclassified Cd. This means that, rather counterintuitively, the gray values of the histogram

Table 4: Table showing the classification errors for two Se-Cd thresholds, the Cd-Pb threshold was fixed at 3.70×10^4.

Se-Cd threshold: 2.10×10^4				Se-Cd threshold: 2.53×10^4			
	Se	Cd	Pb		Se	Cd	Pb
Se	0.87	0.13	0.00	Se	0.93	0.07	0.00
Cd	0.00	1.00	0.00	Cd	0.07	0.93	0.00
Pb	0.00	0.00	1.00	Pb	0.00	0.00	1.00

in Fig. 11 that range between approximately 1.8×10^4 and 2.3×10^4 belong exclusively to Se and are not shared between Se and Cd. If one were to determine the particle's stoichiometry from this reconstruction the concentration of Se would be underestimated and that of Cd would be overestimated.

3.2. Depth sectioning

For depth sectioning only the validity of the proposed reconstruction algorithm, based on the interpretation of the implicit matrix-vector multiplication with A^T in Eq. 2, is investigated. For this goal the projections need not be computed with a full multislice simulation, a simple convolution followed by a vertical summation, as explained in Section 2.2, suffices.

Furthermore, a technique analogous to charge flipping [39] proved to yield good results: As proposed in [40], after each iteration of the reconstruction algorithm, voxels with a density below a certain positive threshold f_{tr} have the sign of their density reversed. In this case the threshold was reset at each iteration at ten percent of the mean of the values above the threshold.

The microscope parameters are summarized in Table 5. The object to be reconstructed, f, measures $1000 \times 1000 \times 125$ voxels of $8 \times 8 \times 210$ pm^3 each. In its center resides an Au cuboctahedron that measures approximately 1.6 nm in diameter and consists of 309 atoms arranged in an FCC lattice with unit cell parameters a, b and c equal to 0.408 nm. The particle is tilted away from the $[00\bar{1}]$ direction by two subsequent tilts of 4° around the crystal's c and a axis. The voxels of the object f containing an atom center are designated a value of 1 and the others a value of 0. The single atom images encoded in the projection matrix A are a convolution of the projected atom potential (taken from [26]) and the probe intensity.

A total of 125 images with a defocus step of 0.21 nm is simulated and used as input for the reconstruction, the resulting data set is represented in

Table 5: Microscope parameters used for depth sectioning. C_7 is the seventh order spherical aberrations, the other symbols are defined in Tab. 3. These settings are taken from [41] and are optimal for a C_7-limited microscope. A visual representation is given in Figure 12.

U	C_1	C_3
200 kV	-2.67 nm	3.54 μm
C_5	C_7	α
-1.13 mm	10 cm	86.8 mrad

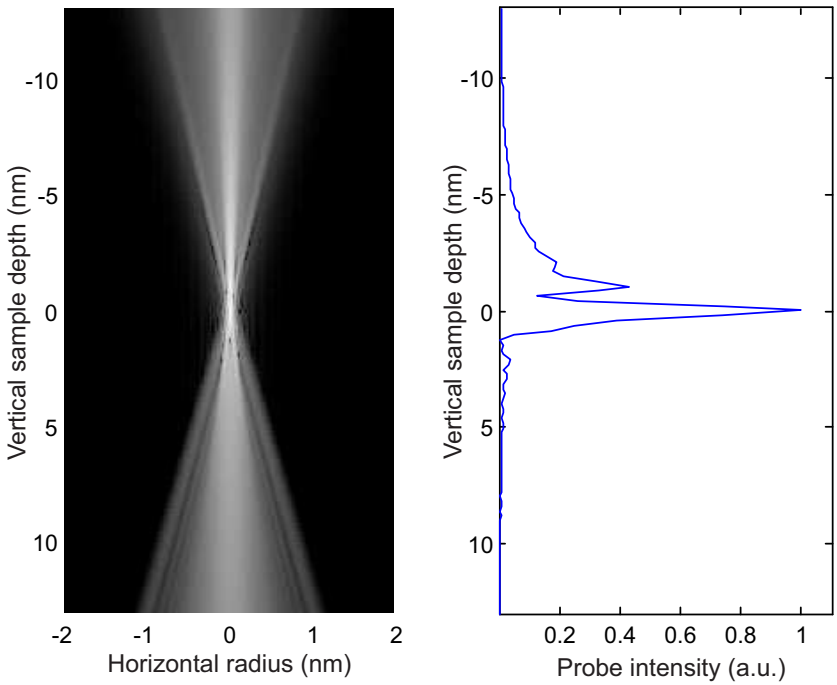

Figure 12: Electron probe intensity for microscope parameters given in Table 5. *Left*: Probe intensity on a logarithmic scale as a function of horizontal and vertical distance. *Right*: Vertical profile along the central column of voxels of the probe intensity, note the pronounced asymmetry.

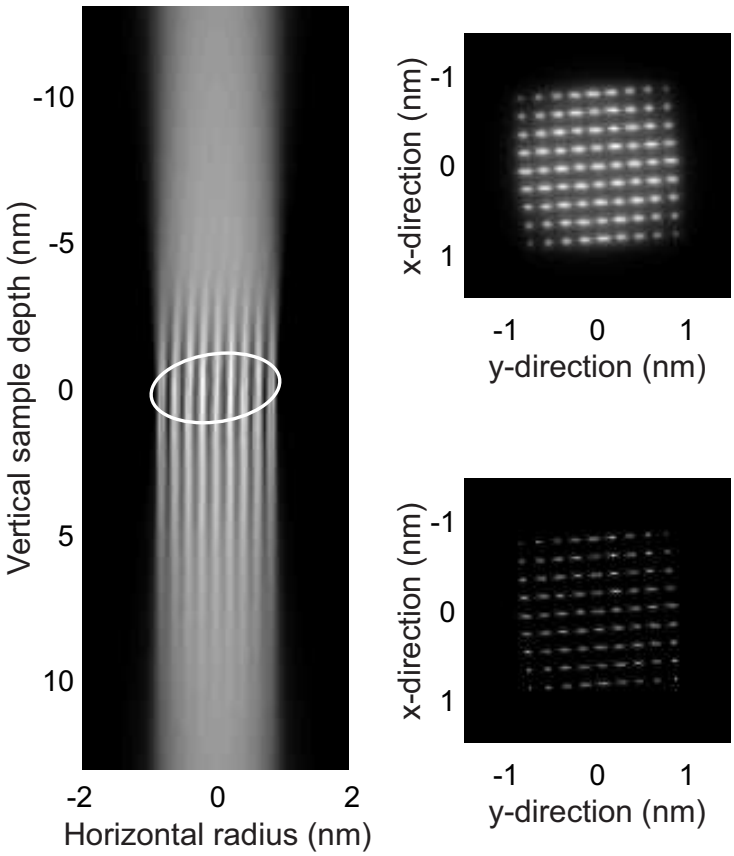

Figure 13: Representation of the depth sectioning data set. *Left*: Average of the vertical slices through the data set. The white oval indicates the particle's position. The gray scale is linear, note how the interpretation of this image is far from straightforward. *Upper right*: Average of the horizontal slices through the data set. *Lower right*: Horizontal slice through the middle of the particle.

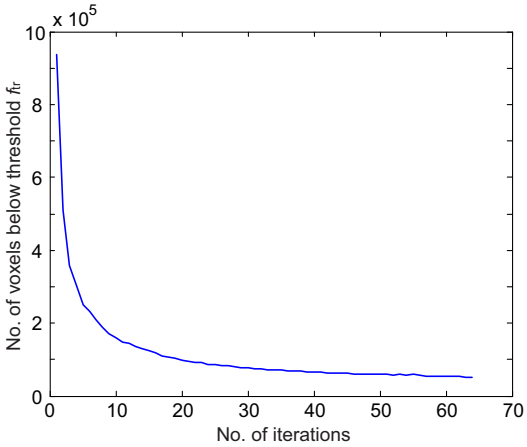

Figure 14: Number of voxels in the reconstruction from depth sectioning data with a value above the threshold f_{tr} as a function of iteration number of the reconstruction algorithm.

Figure 13. The reconstruction was run for 64 iterations and, as shown in Figure 14, by then the number of gray values above f_{tr} has leveled off. In Figure 15 depth-profiles averaged over a disc of radius 1.6 nm centered on the particle are shown for the measurement data set, the reconstructed object and the original object. It is clear that the severe elongation and asymmetry in the vertical direction, both so characteristic for depth sectioning, are overcome and that the full-width-half-maximum (FWHM) of the profile of the reconstruction now approximates that of the original object (2.07 nm and 1.85 nm, respectively). In contrast, the FWHM of the measurements cannot even be determined since it is much larger than the 26 nm spanned by the range of defoci.

Furthermore, it was noticed that 64 iterations were not sufficient for the vertical resolution of the reconstruction to allow atomic imaging.

3.3. Experimental considerations

In these simulations many experimental effects that are to be expected in practice have not been taken into consideration. Examples include, but are not limited to, the background caused by the particles' support film, the effective source size [42], a limited tilt-range for tomography and the beam damage to the nanoparticles that is to be expected at 300 kV. These simulations serve as a proof of concept and to explore the possible benefits of these reconstruction schemes, thereby providing an incentive to invest in tackling these practical problems.

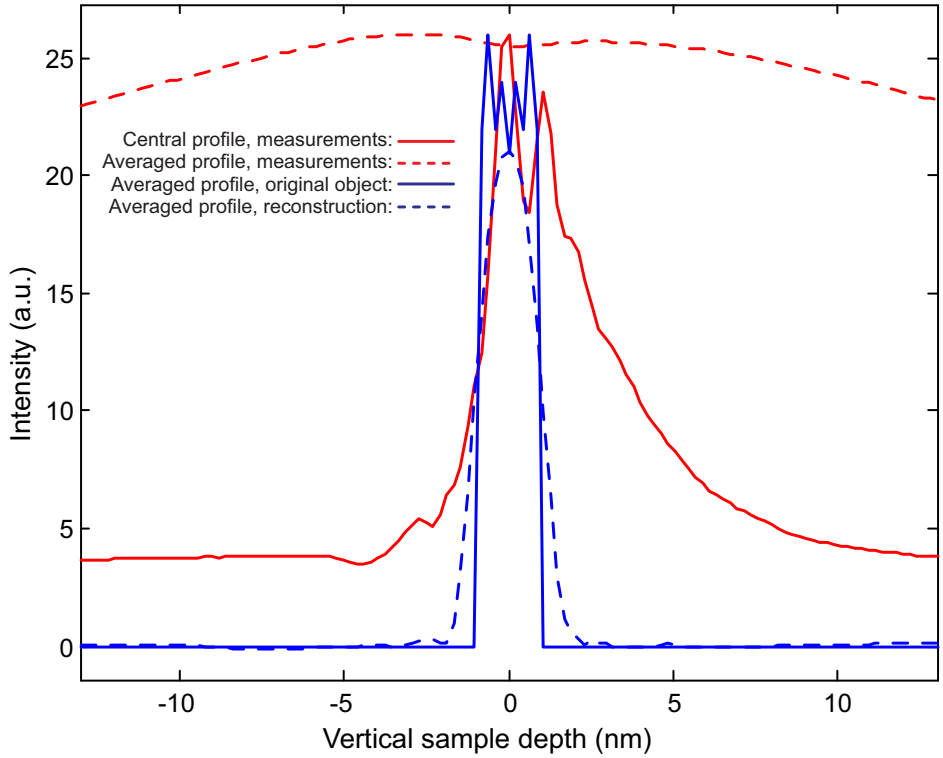

Figure 15: Representation of depth sectioning reconstruction by intensities as a function of sample depth. *Full red line*: Profile through the center of the measured data set, note the asymmetry. *Dashed red line*: Profile of the measured data set averaged over a disc of radius 1.6 nm centered on the particle, note that this profile is extremely broad. *Solid blue line*: Profile of the original object averaged over a disc of radius 1.6 nm centered on the particle (FWHM = 1.85 nm). *Dashed blue line*: Profile of the reconstructed object averaged over a disc of radius 1.6 nm centered on the particle (FWHM = 2.07 nm).

Whereas in conventional HAADF STEM tomography the state of the microscope is sufficiently well characterized by the set of tilt angles of the individual projections alone, in the scheme proposed in this paper that set must be expanded with the objective lens aberrations because the beam shape needs to be known accurately. The aberrations are measured and stored during corrector alignment, but [43, 44] suggest that these might drift too quickly for these figures to be reliable. A possible solution is to reconstruct the aberrations along with the object, as is demonstrated in [4] where it was proven necessary and possible to estimate the defocus of each image along with the reconstructed object.

4. Conclusion

In HAADF STEM tomography and depth sectioning, the shape of the electron beam and the atomic potential can usually be considered as prior knowledge. To include this knowledge in the reconstruction process of the object requires a fully three dimensional treatment of the data. This, however, causes too heavy a load on the computer memory, especially for depth sectioning.

In this paper we show that if the image formation is incoherent—i.e. the image of the object equals the sum of the single atom images—the memory load can be reduced to just the storage of the object, which can be considered as a fundamental lower bound. The validity of this approach was confirmed through numerical simulations for HAADF STEM tomography and depth sectioning.

It was further confirmed through the use of multislice calculations that the image formation process of nanocrystals in HAADF STEM indeed is sufficiently incoherent to allow for atomic resolution retrieval of single element nanocrystals. However, dynamical diffraction effects can cause classification problems when more than one element is present, thus impeding an accurate determination of the particle's stoichiometry.

Acknowledgments

W. Van den Broek acknowledges financial support by the Carl Zeiss Foundation as well as the German Research Foundation (DFG, Grant No. KO 2911/7-1). A. Rosenauer acknowledges the German Research Foundation (DFG, Grant No. AR 2057/8-1). S. Van Aert, J. Sijbers, and D. Van Dyck

gratefully acknowledge financial support from the Research Foundation Flanders (FWO, G.0393.11, G.0064.10, and G.0374.13). The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/20072013] under grant agreement no. 312483 (ESTEEM2).

References

[1] P. W. Hawkes, The electron microscope as a structure projector, in: J. Frank (Ed.), Electron Tomography: Three-Dimensional Imaging with the Transmission Electron Microscope, Plenum Press, New York and London, 1992.

[2] C. T. Koch, Using dynamically scattered electrons for three-dimensional potential reconstruction, Acta Cryst. A65 (2009) 364–370.

[3] W. Van den Broek, C. T. Koch, Method for retrieval of the three-dimensional object potential by inversion of dynamical electron scattering, Physical Review Letters 109 (2012) 245502.

[4] W. Van den Broek, C. T. Koch, General framework for quantitative three-dimensional reconstruction from arbitrary detection geometries in tem, Physical Review B 87 (2013) 184108.

[5] W. Van den Broek, A. Rosenauer, B. Goris, G. T. Martinez, S. Bals, S. Van Aert, D. Van Dyck, Correction of non-linear thickness effects in HAADF STEM electron tomography, Ultramicroscopy 116 (2012) 8 – 12.

[6] L. Reimer, Transmission Electron Microscopy; Physiscs of Image Formation and Microanalysis, Springer-Verlag, New York, 1993.

[7] D. B. Williams, C. B. Carter, Transmission Electron Microscopy: A Textbook for Materials Science, Plenum Press, New York and London, 1996, Ch. 22.

[8] W. Van den Broek, S. Van Aert, D. Van Dyck, A model based atomic resolution tomographic algorithm, Ultramicroscopy 109 (2009) 1485–1490.

[9] P. Hartel, H. Rose, C. Dinges, Conditions and reasons for incoherent imaging in STEM, Ultramicroscopy 63 (1996) 93–114.

[10] P. D. Nellist, S. J. Pennycook, Subangstrom resolution by underfocused incoherent transmission electron microscopy, Phys. Rev. Lett. 81 (1998) 4156–4159.

[11] S. Van Aert, K. J. Batenburg, M. Rossell, R. Erni, G. Van Tendeloo, Three- dimensional atomic imaging of crystalline nanoparticles, Nature 470 (2011) 374–377.

[12] B. Goris, S. Bals, W. Van den Broek, E. Carbó-Argibay, S. Gómez-Graña, L. M. Liz-Marzán, G. Van Tendeloo, Atomic-scale determination of surface facets in gold nanorods, Nature Materials 11 (2012) 930 935.

[13] C.-C. Chen, C. Zhu, E. R. White, C.-Y. Chiu, M. C. Scott, B. C. Regan, L. D. Marks, Y. Huang, J. Miao, Three-dimensional imaging of dislocations in a nanoparticle at atomic resolution, Nature 496 (2013) 74–77.

[14] R. Baraniuk, Compressive sensing [lecture notes], Signal Processing Magazine, IEEE 24 (4) (2007) 118–121.

[15] E. Candes, M. Wakin, An introduction to compressive sampling, Signal Processing Magazine, IEEE 25 (2) (2008) 21–30.

[16] J. Romberg, Compressive sampling, introduction to compressive sampling and recovery via convex programming, IEEE Signal Processing Magazine 25 (2008) 14–20.

[17] W. Van den Broek, S. Van Aert, D. Van Dyck, A model based reconstruction technique for depth sectioning with scanning transmission electron microscopy, Ultramicroscopy 110 (2010) 548–554.

[18] K. van Benthem, A. R. Lupini, M. P. Oxley, S. D. Findlay, L. J. Allen, S. J. Pennycook, Three-dimensional ADF imaging of individual atoms by through-focal series scanning transmission electron microscopy, Ultramicroscopy 106 (2006) 1062 – 1068.

[19] A. Y. Borisevich, A. R. Lupini, S. J. Pennycook, Depth Sectioning with the Aberration-Corrected Scanning Transmission Electron Microscope, Proc. Natl. Acad. Sci. 103 (2006) 3044–304.

[20] J. E. Allen, E. R. Hemesath, D. E. Perea, J. L. Lensch-Falk, Z. Y. Li, F. Yin, M. H. Gass, P. Wang, A. L. Bleloch, R. E. Palmer, L. J. Lauhon, High-resolution detection of Au catalyst atoms in Si nanowires, Nature Nanotechnology 3 (2008) 168–173.

[21] S. P. Frigo, Z. H. Levine, N. J. Zaluzec, Submicron imaging of buried integrated circuit structures using scanning confocal electron microscopy, Appl. Phys. Lett. 81 (2002) 2112–2114.

[22] J. M. Cowley, A. F. Moodie, The scattering of electrons by atoms and crystals. I. A new Theoretical approach, Acta Cryst. 10 (10) (1957) 609–619.

[23] J. G. Allpress, E. A. Hewat, A. F. Moodie, J. V. Sanders, *n*-Beam lattice images. I. experimental and computed images from $W_4Nb_26O_77$, Acta Cryst. A28 (1972) 528–536.

[24] D. F. Lynch, M. A. O'Keefe, *n*-Beam lattice images. II. methods of calculation, Acta Cryst. A28 (1972) 536–548.

[25] P. Goodman, A. F. Moodie, Numerical evaluation of N-beam wave functions in electron scattering by the multislice method, Acta Cryst. A30 (1974) 280–290.

[26] E. J. Kirkland, Advanced Computing in Electron Microscopy, 2nd Edition, Springer, New York Dordrecht Heidelberg London, 2010.

[27] R. Gordon, R. Bender, G. Herman, Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography, Journal of Theoretical Biology 29 (1970) 471–481.

[28] F. Natterer, F. Wübbeling, Mathematical Methods in Image Reconstruction, SIAM, Philadelphia, 2001.

[29] A. C. Kak, M. Slaney, Principles of Computerized Tomographic Imaging, Society of Industrial and Applied Mathematics, 2001, Ch. 4.

[30] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado, G. Van Tendeloo, 3D imaging of nanomaterials by discrete tomography, Ultramicroscopy 109 (2009) 730–740.

[31] J. Gregor, T. Benson, Computational Analysis and Improvement of SIRT, IEEE transactions on medical imaging 27 (2008) 918 – 924.

[32] R. Rojas, Neural Networks—A Systematic Introduction, Springer-Verlag, Berlin, New-York, 1996.

[33] C. T. Koch, A flux-preserving non-linear inline holography reconstruction algorithm for partially coherent electrons, Ultramicroscopy 108 (2) (2008) 141–150.

[34] A. J. D'Alfonso, E. C. Cosgriff, S. D. Findlay, G. Behan, A. I. Kirkland, P. D. Nellist, L. J. Allen, Three-dimensional imaging in double aberration-corrected scanning confocal electron microscopy, Part II: Inelastic scattering, Ultramicroscopy 108 (12) (2008) 1567 – 1578.

[35] F. Meyer, Topographic distance and watershed lines, Signal Process. 38 (1994) 113 – 125.

[36] A. Rosenauer, M. Schowalter, in: Springer Proc. Phys., Vol. 120, Cambridge, MA, 2007, p. 169.

[37] K. Ishizuka, A practical approach for stem image simulation based on the fft multislice method, Ultramicroscopy 90 (2002) 71–83.

[38] S. Bals, M. Casavola, M. A. van Huis, S. Van Aert, K. J. Batenburg, G. Van Tendeloo, D. Vanmaekelbergh, Three-dimensional atomic imaging of colloidal core-shell nanocrystals, Nano Letters 11 (8) (2011) 3420–3424.

[39] G. Oszlányi, A. Sütő, The charge flipping algorithm, Acta Cryst. A64 (2008) 123–134.

[40] W. Van den Broek, C. T. Koch, Feasibility study of the application of the charge flipping algorithm to atomic resolution tomography, in: The Proceedings of the 15th European Microscopy Congress, Royal Microscopical Society, Manchester Central, UK, 2012, p. PS2.5.

[41] V. Intaraprasonk, H. L. Xin, D. A. Muller, Analytic derivation of optimal imaging conditions for incoherent imaging in aberration-corrected electron microscopes, Ultramicroscopy 108 (11) (2008) 1454 – 1466.

[42] J. Verbeeck, A. Béché, W. Van den Broek, A holographic method to measure the source size broadening in STEM, Ultramicroscopy 120 (2012) 35–40.

[43] S. M. Schramm, S. J. van der Molen, R. M. Tromp, Intrinsic instability of aberration-corrected electron microscopes, Phys. Rev. Lett. 109 (2012) 163901.

[44] J. Barthel, A. Thust, On the optical stability of high-resolution transmission electron microscopes, Ultramicroscopy 134 (2013) 6–17.