Correction of non-linear thickness effects in HAADF STEM electron tomography

W. Van den Broek a, n, A. Rosenauer b, B. Goris a, G.T. Martinez a, S. Bals a, S. Van Aert a, D. Van Dyck a

a Electron Microscopy for Materials Science (EMAT), University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium
b Institut für Festkörperphysik (IFP), Universität Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany

ABSTRACT

In materials science, high angle annular dark field scanning transmission electron microscopy is often used for tomography at the nanometer scale. In this work, it is shown that a thickness dependent, non-linear damping of the recorded intensities occurs. This results in an underestimated intensity in the interior of reconstructions of homogeneous particles, which is known as the cupping artifact. In this paper, this non-linear effect is demonstrated in experimental images taken under common conditions and is reproduced with a numerical simulation. Furthermore, an analytical derivation shows that these non-linearities can be inverted if the imaging is done quantitatively, thus preventing cupping in the reconstruction.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Tomography is the discipline of retrieving the three-dimensional structure of an object from a set of two-dimensional projections of this object. A prerequisite is that the image intensities of the projections are proportional to an integral of a local property of the object along the projection direction, this is known as the projection requirement. For this reason, projections are often recorded with high angle annular dark field scanning transmission electron microscopy (HAADF STEM) in materials science. According to Rutherford scattering, an atom contributes to the image intensity proportionally to its atomic number \(Z\) raised to a power \(\gamma\) [1,2]. To a first approximation, the image intensities are therefore proportional to the integral of \(Z^\gamma\) along the beam direction, and the intensity levels in the reconstruction discriminate between the different elements.

In this paper, it is shown that in HAADF STEM a thickness dependent, non-linear damping of the intensities occurs, even under imaging conditions that are common practice. This damping causes the cupping artifact. Due to this artifact, the intensity in the interior of the reconstruction of a homogeneous particle is underestimated. A profile of the reconstructed intensity through the middle of the particle would be shaped like a downward arc instead of a straight line [3]. The cupping artifact can impede interpretation when the reconstructed intensities are needed to discriminate between different elements in neighboring regions. This is the case, for example, when the distribution of elements in a core–shell particle is investigated.

The cupping artifact has been observed in HAADF STEM before [4,5] and has been attributed correctly to the non-linearity of the contrast. In the field of incoherent bright field (IBF) STEM [6–8], the non-linearities in the recorded signal are corrected routinely. In this paper, it is shown that a similar correction is possible in HAADF STEM, provided that the image intensities are measured on an absolute scale [9–11].

The outline of this paper is as follows. In Section 2, the amount of damping is assessed for common recording conditions with an analytical derivation and with image simulations. Furthermore, a method to correct this excessive damping is proposed and a comparison to IBF STEM is made. In Section 3, it is explained how the image intensities are transformed to an absolute scale. Also, the influence of the measurement noise on the non-linear thickness correction is investigated. In Section 4, the method is successfully tested on a tilt series of a 30 nm gold particle, showing that cupping can indeed be prevented. In Section 5, conclusions are drawn.

2. Image formation

Each layer of a uniform sample scatters the same portion of the electrons that enter the layer to angles above \(\theta_{int}\), the inner semi-
angle of the HAADF detector. However, a portion of the electrons being scattered by a layer near the bottom of the sample already had been scattered towards the detector by the preceding layers. The net contribution of the lower layers to the HAADF STEM signal is therefore less than that of layers higher up. This has been formalized in [2,12] where the part \( I \) of the beam intensity that is scattered to angles smaller than \( \theta_{in} \) is shown to decrease exponentially as a function of object thickness \( t \)

\[
I = I_0 \exp(-\mu t) \quad \text{with} \quad \mu = n \sigma,
\]

(1)

where \( I_0 \) is the intensity of the beam as it enters the object, \( n \) is the number of atoms per unit of volume, \( \sigma \) is the Rutherford scattering cross-section to angles larger than \( \theta_{in} \) and, following Ref. [3], \( \mu \) is defined as the attenuation coefficient and equals the inverse of the mean free path between scattering events.

In this derivation, single scattering was assumed. In practice, however, previously scattered electrons scatter back to angles below \( \theta_{in} \), resulting in an intensity that is higher than predicted by \( I \). This expression is therefore only valid for thin specimens with a thickness of the order of the mean free path.

The derivations in [2,12] can easily be generalized to cases where the attenuation coefficient varies along the beam trajectory (as is the case in core–shell particles for example) to yield

\[
I = I_0 \exp\left(-\int \mu(t) \, dt\right).
\]

(2)

To keep the notations simple all subsequent derivations start from Eq. (1), keeping in mind that all of the results can easily be generalized to apply to this expression as well.

2.1. Thickness dependent damping in HAADF STEM

In HAADF STEM, electrons scattered to angles larger than the inner radius \( \theta_{in} \) of the detector are detected. Therefore, the recorded signal \( I \) is complementary to \( I \):

\[
I = I_0 - I,
\]

(3)

\[
I = I_0(1-\exp(-\mu t)),
\]

(4)

\[
I \simeq I_0\mu t - \frac{1}{2} I_0(\mu t)^2.
\]

(5)

The expansion in Eq. (5) shows that to first order, the HAADF STEM signal is proportional to the integrated attenuation coefficient along the beam direction and thus fulfills the projection requirement. However, for increased thicknesses the second term of the expansion is no longer negligible and a damping occurs, thereby violating the projection requirement.

This thickness dependent damping results in the so-called cupping artifact where the intensity in the interior of the reconstruction of a homogeneous particle is underestimated. This artifact is well known in X-ray tomography, as shown, for example, in [3]. In HAADF STEM, this artifact has been noticed as well, see, for example, Refs. [4,5,13], but has not been corrected so far.

The damping with thickness can be corrected by transforming the measurements \( I \) to a projection of the attenuation coefficient \( \mu \) by simply applying the inverse of Eq. (4)

\[
\mu t = -\ln\left(1 - \frac{I}{I_0}\right).
\]

(6)

By writing down Eq. (3), the outer radius \( \theta_{out} \) of the detector has been tacitly assumed to be \( \pi \) rad, which is not the case in practice. In Refs. [8,14], it is shown that a finite detector size causes contrast reversal for thicknesses above a few hundreds of nanometers. This effect is more pronounced for small \( \theta_{out} \). Since in this paper the specimen is thin (30 nm) and \( \theta_{out} \) is more than six times larger than \( \theta_{in} \), this effect will be neglected.

2.2. Relation with incoherent bright field STEM

In incoherent bright field STEM (IBF STEM) tomography [6–8], the signal is collected by a disc-shaped detector in the diffraction plane with the optical axis running through its center and a radius several times larger than the beam’s convergence semi-angle. By choosing the radius equal to the inner radius of the annular detector, the IBF STEM signal \( I \) is made complementary to the HAADF STEM signal. The projection of the attenuation coefficient can be obtained by inverting Eq. (1)

\[
\mu t = -\ln\left(\frac{I}{I_0}\right)
\]

(7)

as is routinely done in IBF STEM. The incoming beam intensity \( I_0 \) can easily be estimated from the vacuum intensity and the image contrast does not suffer from the non-monotonicity brought about by the finite detector size in HAADF STEM for thick samples [8].

However, IBF STEM suffers from low contrast when the specimen is thin and/or consists of light elements. In Appendix A, the signal-to-noise ratios \( \text{SNR}_{DF} \) and \( \text{SNR}_{BF} \) of \( \mu t \), retrieved from HAADF STEM and IBF STEM measurements respectively, are derived. The results are plotted in Fig. 1. For specimen thicknesses below

\[
\frac{\ln(2)}{\mu t}
\]

(8)

\( \text{SNR}_{DF} \) is considerably higher than \( \text{SNR}_{BF} \), showing that for samples thinner than 0.7 times the mean free path between scattering events HAADF STEM is preferable over IBF STEM. This conclusion is reflected in the fact that IBF STEM papers focus on thick specimens.

2.3. Image simulation with STEMsim

So far, quite a crude model has been used for the image formation process. For comparison, the image intensity as a function of thickness is calculated in a quantum mechanically correct manner in STEMsim [15]. The object is an Au slab of \( 40 \times 2.5 \times 2.5 \) nm\(^3\) tilted away by 11° from the \([100]\) direction to avoid a zone-axis orientation and consequent channeling. The Debye–Waller factors of the atoms are set to 0.6331 Å\(^2\) [16]. An average over eight frozen phonon configurations is taken. The microscope parameters are summarized in Table 1. It must be
noted that in STEMsim the intensity is given as a fraction of the incoming beam, i.e. \( I_0 = 1 \).

The intensity profile generated by STEMsim approximately follows the non-linear behavior with thickness as described by Eq. (4). This is demonstrated in Fig. 2 where Eq. (6) is used to transform the STEMsim profile to an approximate straight line representing the projected attenuation coefficient \( \mu t \).

### 3. Quantitative HAADF STEM

In quantitative HAADF STEM, information is extracted from the absolute intensities in the images [9–11]. It is therefore important to express the measured intensities as a fraction of the intensity of the incoming beam. Furthermore, the thickness correction in Eq. (6) might produce biased results in the presence of noise. Both problems are considered in this section.

#### 3.1. Detector calibration

In most HAADF STEM experiments, the image intensities \( I \) are not known on an absolute scale because the detector is not characterized well. As a result, only intensity differences within the same recording are meaningful and comparison with image simulations is problematic. In order to correct the damping through Eq. (6), the intensities \( I \) need to be known accurately as fractions of the incoming beam intensity \( I_0 \).

In Refs. [9–11], it is shown that a linear transformation suffices to recalibrate \( I \) as a fraction of \( I_0 \). The procedure to find this linear transformation for a commercial HAADF detector is given in Ref. [11] and has been applied in this paper.

#### 3.2. Bias due to noise

The transformation in Eq. (6) might introduce a bias because the intensities \( I \) suffer from Poisson noise and the non-linear logarithmic function treats negative deviations from the mean differently from positive deviations.

The investigation in Appendix B shows that a bias indeed occurs

\[
\langle -\ln(1 - \frac{I}{I_0}) \rangle = \langle -\ln(1 - \frac{\langle I \rangle}{I_0}) \rangle + B(\langle I \rangle),
\]

where \( \langle \cdot \rangle \) denotes the expectation value and \( B \) is the bias term. In the appendix, it is shown that the ratio of both terms on the right-hand side typically is below 1.44/\( I_0 \). For a beam current of 10 pA and a dwell time of 15 \( \mu s \), the number of electrons \( I_0 \) is approximately 950. The bias can therefore safely be ignored, as will be done in the remainder of this paper.

### 4. Experiment

The procedure is tested on a HAADF STEM tilt series of a gold particle of approximately 30 nm in diameter, supported on amorphous carbon. The tilt angles range from \(-70^\circ\) to \(+70^\circ\) in increments of \(2^\circ\). The probe sampling distance \( \Delta \) is 0.26 nm. The images have been recorded on a TECNAI G2 with a spherical aberration constant of 1.2 mm, an acceleration voltage of 200 kV, an inner HAADF detector radius \( \theta_{in} \) of 36 mrad and an outer radius \( \theta_{out} \) of 235 mrad. In practice, this inner radius proved to be large enough to prevent diffraction contrast; this is also suggested by the simulations in Section 2.3 and by Hartel et al. in Ref. [17], where an inner radius of at least three times the convergence semi-angle is recommended.

Two reconstructions with two different input signals are made:

1. the signal \( I/I_0 \),
2. the corrected signal \( -\ln(1-I/I_0) \).

The reconstruction is carried out in MATLAB using the multiplicative simultaneous iterative reconstruction technique (SIRT). Both the accuracy and the noise of the SIRT reconstruction grow with the number of iterations and that number is therefore commonly used to make a trade-off between both effects. In order to demonstrate the cupping artifact, accurately reconstructed gray values are more important than a low noise level; therefore, the number of iterations is set to 45.

In Fig. 3, a surface rendering of reconstruction 2 is shown. The different surface facets are clearly visible. Furthermore, ortho-slices through the center of reconstructions 1 and 2 are compared, thereby demonstrating the cupping artifact. In order to display the artifact clearly, the gray values of reconstructions 1 and 2 are clipped between \( 2.1 \times 10^{-3} \) and \( 3.5 \times 10^{-3} \), and \( 2.3 \times 10^{-3} \) and \( 4.1 \times 10^{-3} \), respectively. However, since such clipping overemphasizes the noise, the reconstructions are filtered with a \( 5 \times 5 \times 5 \) average filter for the purpose of display.

A three-dimensional rotational average of both, unfiltered, reconstructions is shown in Fig. 4. The cupping artifact is clearly visible in reconstruction 1, whereas it is absent in reconstruction 2. The intensity at the center is given by the value of just one voxel, while the number of voxels used in the average increases rapidly with the distance from the center, the reduced intensity at zero can therefore be attributed to noise. Furthermore, since the damping has been corrected, the gold in reconstruction 2 has a higher intensity.
The reconstructed quantity is the attenuation coefficient $\mu$ multiplied with the voxel size, which in this case is equal to the sampling distance $A$ of 0.26 nm. From the profile in Fig. 4, it is deduced that for this experiment $\mu = 1.3 \times 10^7$ m$^{-1}$, corresponding to a mean free path of 77 nm. The particle’s diameter of about 30 nm is considerably shorter than that justifying the use of the thickness correction. Furthermore, the diameter is well below the limit of $\ln(2)/\mu = 53$ nm for which HAADF STEM performs better than IBF STEM in terms of noise.

5. Conclusion

In this paper, it is shown that a non-linear thickness effect occurs in HAADF STEM imaging in the form of a non-linear damping of the intensities of the thicker parts of the object. This effect is explained from the point of view of Rutherford scattering and confirmed with advanced image simulations. The damping causes the so-called cupping artifact, where the intensity in the interior of the reconstruction of a homogeneous particle is underestimated. A strategy to correct this non-linear damping has therefore been proposed in our work. These results are of great importance for those studies where three-dimensional quantitative results at the nanoscale are aimed for.

This has been tested experimentally with a tilt series of a 30 nm gold particle, recorded under conditions that are common practice. A reconstruction using the uncorrected intensities showed a considerable cupping artifact, while a reconstruction with the corrected intensities showed none.

Acknowledgments

W. Van den Broek was working under the Condor project, a project under the supervision of the Embedded Systems Institute (ESI) and with FEI company as the industrial partner. This project is partially supported by the Dutch Ministry of Economic Affairs under the BSIK program. The work was supported by the Flemish Fund for Scientific Research (F.W.O. Vlaanderen) through project fundings for S. Van Aert and S. Bals (G.0064.10N, G.0393.11 and G.0247.08) and through a PhD research grant to B. Goris.

Appendix A. Signal-to-noise ratio

In this appendix, two expressions for the SNR of $\mu t$ will be derived, $\text{SNR}_{DF}$ and $\text{SNR}_{BF}$, depending on whether $\mu t$ has been retrieved from HAADF STEM or IBF STEM projections, respectively. Use is made of standard error propagation theory and the results are plotted in Fig. 1.

A.1. HAADF STEM

For HAADF STEM projections

$$\mu t = -\ln \left( 1 - \frac{l}{I_0} \right).$$  \hspace{1cm} (A.1)

Suppose that the intensity $l$ follows a Poisson distribution with variance $\sigma_l^2 = l$ and that $I_0$ has been estimated with negligible error, the variance $\sigma_{DF}^2$ of $\mu t$ is then given by

$$\sigma_{DF}^2 = \sigma_l^2 \left( -\ln \left( 1 - \frac{l}{I_0} \right) \right)^2. $$  \hspace{1cm} (A.2)

$$\sigma_{DF}^2 = \frac{l}{(I_0-l)^2}. $$  \hspace{1cm} (A.3)
The signal-to-noise ratio SNR$_{BF}$ is defined as

$$\frac{\mu t}{\sigma_{BF}^2}$$

and applying Eq. (4) leads to

$$\text{SNR}_{BF} = \sqrt{I_0} \frac{\mu t \exp(-\mu t/2)}{\sqrt{\exp(\mu t) - 1}}.$$  

(A.5)

A.2. IBF STEM

For IBF STEM projections

$$\mu t = -\ln \frac{I}{I_0}.$$  

(A.6)

Suppose that the intensity $I$ follows a Poisson distribution with variance $\sigma_{BF}^2 = I$, and that $I_0$ has been estimated with negligible error, the variance $\sigma_{BF}^2$ of $\mu t$ is then given by

$$\sigma_{BF}^2 = \sigma_{BF}^2 \left[ \frac{\partial}{\partial I} \left( -\ln \frac{I}{I_0} \right) \right]^2.$$  

(A.7)

$$\sigma_{BF}^2 = \frac{1}{I_0}.$$  

(A.8)

The signal-to-noise ratio SNR$_{BF}$ is defined as

$$\frac{\mu t}{\sigma_{BF}}.$$  

and applying Eq. (1) leads to

$$\text{SNR}_{BF} = \sqrt{I_0} \frac{\mu t \exp(-\mu t/2)}{\sqrt{\exp(\mu t) - 1}}.$$  

(A.9)

The bias term in Eq. (B.4) follows a Poisson distribution with $I_0$, and the following equalities have been used \[18\]:

$$\sum_{i=0}^\infty \frac{x^i}{i!} = \exp(x), \quad \sum_{i=0}^\infty \frac{x^i}{i!} = x \exp(x).$$  

(B.5)

The ratio $r$ of the first term and the bias term in Eq. (B.4) increases with $\langle I \rangle$. From Fig. 2, it can be seen that $\langle I \rangle$ typically is lower than 0.5$I_0$, which means that $r$ typically is lower than 1.44/I_0. For $\langle I \rangle$ converging to zero, $r$ approaches 0.5/I_0. For a beam current of 10 pA and a dwell time of 15 μs, the number of electrons $I_0$ is approximately 950. The bias can therefore safely be ignored, as has been done in this paper.

References