

**Automatic correction of nonlinear damping effects in HAADF–STEM tomography for nanomaterials of discrete compositions**

Zhong, Zhichao; Aveyard, Richard; Rieger, Bernd; Bals, Sara; Palenstijn, Willem Jan; Batenburg, K. Joost

**DOI**

[10.1016/j.ultramic.2017.10.013](https://doi.org/10.1016/j.ultramic.2017.10.013)

**Publication date**

2018

**Document Version**

Accepted author manuscript

**Published in**

Ultramicroscopy

**Citation (APA)**

Zhong, Z., Aveyard, R., Rieger, B., Bals, S., Palenstijn, W. J., & Batenburg, K. J. (2018). Automatic correction of nonlinear damping effects in HAADF–STEM tomography for nanomaterials of discrete compositions. *Ultramicroscopy*, 184, 57-65. <https://doi.org/10.1016/j.ultramic.2017.10.013>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Automatic Correction of Nonlinear Damping Effects in HAADF-STEM Tomography for Nanomaterials of Discrete Compositions

Zhichao Zhong<sup>a</sup>, Richard Aveyard<sup>b</sup>, Bernd Rieger<sup>b</sup>, Sara Bals<sup>c</sup>, Willem Jan Palenstijn<sup>a</sup>, K. Joost Batenburg<sup>a,d</sup>

<sup>a</sup>*Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*

<sup>b</sup>*Department of Imaging Physics, Delft University of Technology, The Netherlands*

<sup>c</sup>*EMAT, University of Antwerp, Antwerp, Belgium*

<sup>d</sup>*Mathematical Institute, Universiteit Leiden, Leiden, The Netherlands*

---

## Abstract

HAADF-STEM tomography is a common technique for characterizing the three-dimensional morphology of nanomaterials. In conventional tomographic reconstruction algorithms, the image intensity is assumed to be a linear projection of a physical property of the specimen. However, this assumption of linearity is not completely valid due to the nonlinear damping of signal intensities. The nonlinear damping effects increase w.r.t the specimen thickness and lead to so-called “cupping artifacts”, due to a mismatch with the linear model used in the reconstruction algorithm. Moreover, nonlinear damping effects can strongly limit the applicability of advanced reconstruction approaches such as Total Variation Minimization and discrete tomography.

In this paper, we propose an algorithm for automatically correcting the nonlinear effects and the subsequent cupping artifacts. It is applicable to samples in which chemical compositions can be segmented based on image gray levels. The correction is realized by iteratively estimating the nonlinear relationship between projection intensity and sample thickness, based on which the projections are linearized. The correction and reconstruction algorithms are tested on simulated and experimental data.

---

## 1. Introduction

In materials science, electron tomography (ET) is commonly used to characterize the three-dimensional (3D) structural and compositional information of nanomaterials. The 3D image is usually reconstructed from a tilt series of two-dimensional (2D) projections (projection images). The projection images should have a monotonic relationship between the measurement intensity and the integrated physical property of the specimen, which is referred to as the *projection requirement* in ET [1, 2]. Strictly speaking, the relationship should be linear, as most tomographic reconstruction algorithms are based on a linear mathematical model – the line integral model. It assumes that the projection is a measurement of a physical property integrated along the projection orientation (see

Chapter 3 in [3]).

High angle annular dark field (HAADF) scanning transmission electron microscopy (STEM) is commonly used for ET [1, 4] under the implicit assumption that the projection requirement can be approximately satisfied. The image intensity approximates to be proportional to the mass-thickness weighted by  $Z^{\sim 2}$ , where  $Z$  is the atomic number [4]. However, this approximation is not always valid. One example is that when projections of a crystalline material are acquired at zone-axis orientations, fringes and large overall intensity differences can be observed. Thus the tilts at zone-axis are usually excluded from the tomographic reconstruction step [5]. Another example is that the image intensity damps at high sample thickness due to the multiple scattering events redirecting electrons

1 outside the annular detector, which can occur in all  
2 projection orientations. While the zone-axis effects  
3 can be easily identified, intensity damping is not  
4 easily seen in individual projections. In this pa-  
5 per, we aim at addressing the nonlinear effects of  
6 intensity damping for tomographic reconstruction.

7 The consequence of intensity damping appears as  
8 the cupping artifact in tomographic reconstruction:  
9 the gray levels in the center of the reconstructed  
10 sample are underestimated while overestimated on  
11 the exterior [6]. In Fig. 1(a), an example of the  
12 cupping artifact is given. It is a 2D cross section of  
13 an Au-Ag core-shell nanoparticle [7], reconstructed  
14 using the SIRT algorithm [8]. If we look at the  
15 line-profile of the 2D image (Fig. 1(b)), the curve  
16 appears in a concave “cup” shape, while ideally it  
17 should be flat. The cupping artifacts are caused by  
18 the strong damping effects of Au at large thickness,  
19 which is illustrated by the simulated relationships  
20 between measurement intensity and sample thick-  
21 nesses using the multislice simulation method [5] in  
22 Fig. 2. In this example, the linear approximation  
23 is only valid for thickness smaller than 8 nm due to  
24 the clear damping effect for larger thickness.

25 It is important to correct the nonlinear effects  
26 and the subsequent cupping artifacts for three rea-  
27 sons. First of all, compositional analysis based on  
28 gray levels becomes difficult when the cupping arti-  
29 facts occur, as gray levels are not proportional any-  
30 more to density and atomic numbers. Second, mor-  
31 phological analysis based on segmentation of recon-  
32 struction images is hindered by the cupping arti-  
33 facts. Some straightforward segmentation meth-  
34 ods, e.g. Otsu’s method [9], require that for each  
35 chemical composition there should be one constant  
36 gray level. Third, the nonlinear effects limit apply-  
37 ing advanced reconstruction algorithms to address  
38 a critical issue of ET – the missing wedge artifacts  
39 caused by the limited tilt range of the sample. Al-  
40 gorithms such as total variation minimization [10]  
41 reduce the missing wedge artifacts by incorporating  
42 prior knowledge i.e. sparsity of the unknown sam-  
43 ple. Nevertheless, these algorithms have an even  
44 stronger requirement for the linear forward model  
45 which is inaccurate due to the nonlinear effects.

46 Despite these shortcomings of using uncorrected  
47 data, there are few publications addressing the non-  
48 linearity issue in ET [5, 6]. Nonlinear effects are

1 usually ignored or mitigated during image acquisi-  
2 tion by increasing the inner angle of the HAADF  
3 detector but at the cost of losing signal strength  
4 [5]. An alternative to adjusting the acquisition pa-  
5 rameters is to correct the measured data in a post-  
6 processing step by linearizing the projection data,  
7 provided that the incident beam intensity is known  
8 [6]. The method described here requires only the  
9 HAADF signal, consequently, it can be applied to  
10 correct cupping artifacts in many existing datasets  
11 acquired in a conventional manner. The mathe-  
12 matical model of nonlinearity and the concept of  
13 linearization in [6] are also used in this paper, which  
14 will be explained in Section 2.1.

15 Here, we propose an iterative algorithm to au-  
16 tomatically correct the nonlinear effects and the  
17 cupping artifacts. It does not require the extra  
18 measurement of the incident beam intensity as in  
19 [6]. Instead, it automatically models the nonlinear  
20 effects given the projection data. The algorithm  
21 iteratively searches for the minimal distance be-  
22 tween the acquired projections and the nonlinear  
23 re-projections of chemical compositions by varying  
24 the nonlinear model and the reconstruction image,  
25 so as to estimate a nonlinear relationship between  
26 the measured HAADF-STEM intensities and sam-  
27 ple thickness for all chemical compositions. The  
28 algorithm contains the following steps in every it-  
29 eration: first a reconstruction image with contin-  
30 uous gray levels is made; then the image is seg-  
31 mented into several binary images, each of which  
32 corresponds to a chemical composition; after that,  
33 the nonlinear effects are modeled by minimizing the  
34 projection distance; based on the model, the pro-  
35 jection data is linearized at last. The concept of  
36 iterative correction has been used to correct beam  
37 hardening artifacts for X-ray computed tomogra-  
38 phy, which is similarly caused by nonlinear intensi-  
39 ties [11, 12, 13].

40 Our approach is only applicable to samples con-  
41 sisting of several chemical compositions with uni-  
42 form densities, such as homogeneous or core-shell  
43 particles. It is assumed that for these samples the  
44 volumetric distributions of the compositions can be  
45 approximated well by segmenting the reconstructed  
46 image based on gray levels and that this segmen-  
47 tation improves as the correction model applied to  
48 the measured data becomes more accurate. In fact,  
49 these kinds of samples are commonly studied in

1 materials science. For example, the samples typi-  
 2 cally studied in the context of discrete tomography  
 3 [14, 15] match the requirements.

4 In Section 2, the correction algorithm is ex-  
 5 plained in detail. In section 3, we demonstrate  
 6 how the nonlinear effects are corrected using this  
 7 algorithm for real experimental data and phantom  
 8 simulations.

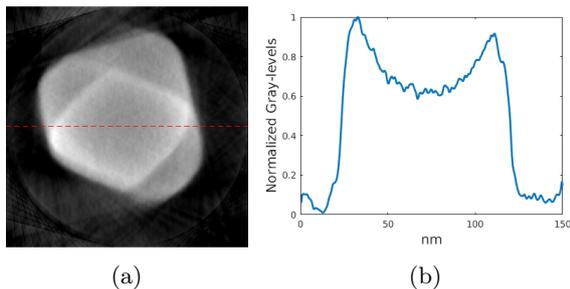


Figure 1: (a): 2D slice of the SIRT reconstruction of an Au-Ag nanoparticle. (b): Gray levels of the line-profile located at the dash line of the 2D slice.

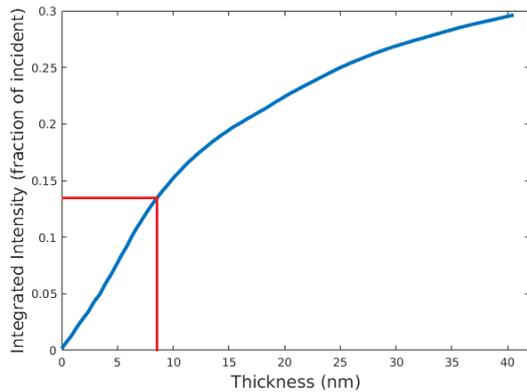


Figure 2: Normalized HAADF signal intensity w.r.t the thickness of Au slabs mistilted 10 degrees from the [100] zone axis about the  $\langle 100 \rangle$  axis, simulated using the multislice method [5]. The accelerating voltage is 200 kV, the convergence angle is 10 mrad and the detector angular range is 50 - 250 mrad. The intensities are scaled by the incident beam intensity. The red lines indicate the region where intensity is approximately linear to thickness.

## 2. The Nonlinear Model and the Correction Algorithm

### 2.1. The Nonlinear Model

To linearize the projections, we first need to define a model that describes the nonlinear relationship mathematically. A precise mathematical model is possible but does not fit as a subroutine of the correction algorithm. The computation of a sophisticated model, such as the one used in multi-slice simulations which take into consideration the multiple scattering of electrons [5], is extremely time-consuming and costly. Therefore, a simple model is preferred here.

Here, we choose a model that has already been used for describing the nonlinear relationship. In [6, 16], it is illustrated we can assume that the HAADF detector collects electrons complementary to the electrons scattered to angles smaller than its inner detector angle. The electrons can also be scattered to angles beyond the outer detector angle, but the proportion is negligibly small. By pragmatically applying a simple Beer-Lambert description of electron scattering we can state that the number of electrons scattered to small angles  $p_t$  decreases exponentially to the sample thickness  $t$  along the beam direction. The  $p_t$ - $t$  relationship is

$$p_t = I_0 \exp\left(-\sum_e^K \mu_e t\right), \quad (1)$$

where  $I_0$  is the incident beam intensity,  $e$  is the index of chemical composition,  $K$  is the total number of chemical compositions,  $\mu_e$  is the attenuation coefficient of chemical composition  $e$ . Therefore, the complementary HAADF signal intensity  $p$  at sample thickness  $t$  is:

$$p = I_0(1 - \exp(-\sum_e^K \mu_e t)) + p_b, \quad (2)$$

where  $p_b$  is the bias signal, which is influenced by the dark current, carbon grid, and possibly other factors.

This mathematical model has been used to correct the cupping artifacts successfully in [6], which is applicable only if the incident beam intensities can be measured. An advantage of this simple

1 model is that it can easily be transformed into a  
 2 linear relationship by taking logarithms so that we  
 3 can avoid solving nonlinear least-squared problems  
 4 for tomographic reconstruction.

5 In the practice of ET, a series of projections are  
 6 taken at different angles. The image intensity of  
 7 each pixel corresponds to the electrons scattered for  
 8 an electron beam transmitting through the sample,  
 9 which is called a line projection here. In total, there  
 10 are  $M$  pixels for all the images. The image intensity  
 11 of the  $i_{th}$  pixel is now written as an entry  $p_i$  in  
 12  $\mathbf{p} \in \mathbf{R}^M$ . In addition, the space of reconstruction  
 13 is a cubic volume partitioned into  $N$  voxels.

14 We also assume the chemical compositions are  
 15 not mixed and voxels are small enough to resolve  
 16 every chemical composition, which means that in  
 17 each voxel only one element is present. As stated in  
 18 the introduction, this algorithm is applied to sam-  
 19 ples with uniform density. Thus we assume that  
 20 each chemical composition is either present (1) or  
 21 absent (0) in each voxel. The distribution of chem-  
 22 ical composition  $e$  is described by binary variables  
 23  $s_{ej}$ , where  $j = 1, \dots, N$  is the index of voxel.

24 Now we define the nonlinear relationship in the  
 25 discrete form. For pixel  $i$ , the corresponding sample  
 26 thickness of chemical composition  $e$  is now written  
 27 as the ray-sum  $\sum_{j=1}^N w_{ij}s_{ej}$ , where the factor  $w_{ij}$   
 28 is determined by the area of intersection between  
 29 the  $i_{th}$  line projection and the  $j_{th}$  voxel. The rela-  
 30 tionship between projection intensities and binary  
 31 volumes are:

$$p_i = I_0(1 - \exp(-\sum_{e=1}^K \mu_e \sum_{j=1}^N w_{ij}s_{ej})) + p_b, \quad (3)$$

32 where  $i = 1, \dots, M$ .

### 33 2.2. The Correction Algorithm

34 The basis of the correction algorithm is to esti-  
 35 mate the nonlinear relationship of Eq. 3 based on  
 36 the reconstructed distributions of chemical compo-  
 37 sitions. The procedures of the automatic correc-  
 38 tion algorithm are given in the flowchart (Fig. 3).  
 39 The correction is realized iteratively through the  
 40 following steps: (1) make a reconstruction image  
 41 based on the linear model from the projections; (2)  
 42 segment the reconstruction into a series of binary

1 images, one for each chemical composition; (3) esti-  
 2 mate the parameters of the nonlinear model in Eq.  
 3 3 given the projections and the binary images; (4)  
 4 reduce the nonlinearities in the projections through  
 5 a rescaling of the intensities based on the nonlinear  
 6 model.

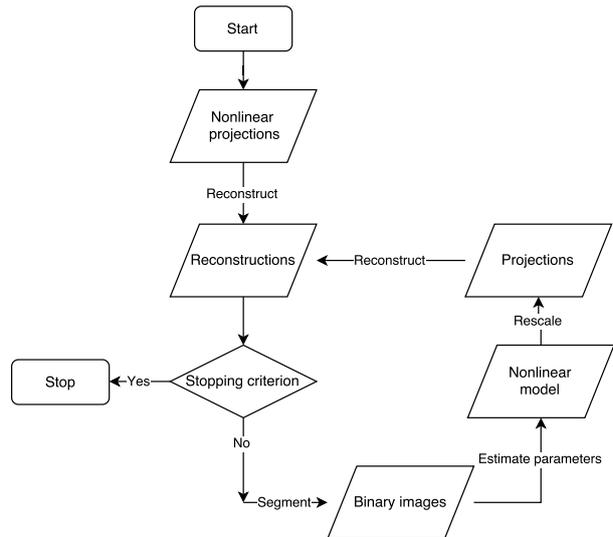


Figure 3: Flowchart of the correction algorithm

Before we explain the steps explicitly, we estab-  
 lish an objective function which will be used to  
 guide the optimization in the correction algorithm.  
 We define it as the  $l_2$  norm of the distance between  
 the acquired projections and the re-projection of  
 binary images based on our nonlinear model:

$$\mathcal{C}(I_0, p_b, \boldsymbol{\mu}, \mathbf{S}) = \|\mathbf{p} - I_0(1 - \exp(-\mathbf{W} \sum_{e=1}^K \mu_e \mathbf{s}_e)) - p_b\|_2^2, \quad (4)$$

where  $\mathbf{W} = \{w_{ij}\}$ ,  $\boldsymbol{\mu} = \{\mu_e\}$  and  $\mathbf{S} = \{s_{ej}\}$ .

We also define a stopping criterion. The cost  
 value at the  $r_{th}$  iteration is denoted as the  $c^r$ . The  
 loop is terminated if the cost is stable, which is  
 when the following criterion is met:

$$\frac{c^r + c^{r-1}}{c^{r-2} + c^{r-3}} > t, \quad (5)$$

where  $0 < t < 1$  is a thresholding value. Note that  
 although we minimize the cost function in some  
 steps of the algorithm, the cost is not guaranteed  
 to reach a global minimum in the end.

1     *Step 1: Reconstruction*

2     As the first step, a reconstruction with continu-  
3     ous gray levels is made for determining the binary  
4     images in the next step. Though it is possible to  
5     reconstruct binary images directly using some dis-  
6     crete tomography algorithms (e.g. [14]), these algo-  
7     rithms will possibly not give better results than ba-  
8     sic algorithms given an inaccurate forward model.  
9     Thus, we choose to first make a reconstruction  $\mathbf{x}$   
10    with continuous gray levels based on a linear model  
11    and then segment the reconstruction into binary  
12    images  $\mathbf{S}$ .

13    The reconstruction is computed using the simul-  
14    taneous iterative reconstruction technique (SIRT)  
15    [8] which solves the following least-squares prob-  
16    lem:

$$\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{p}_{lin} - \mathbf{W}\mathbf{x}\|_2^2. \quad (6)$$

17    The widely used SIRT algorithm is chosen for its  
18    robustness to noise and its easy implementation.

19    The input for this step is a set of “linearized”  
20    projections  $p_{lin}$ . For the first iteration, they are  
21    just the acquired projections. For the other itera-  
22    tions, they are adopted as the projections that have  
23    been rescaled in the previous iteration, which will  
24    be explained in Step 4.

25    *Step 2: Segmentation*

26    The binary images are then determined by seg-  
27    menting the reconstruction image  $\mathbf{x}$ . As gray lev-  
28    els are related to atomic numbers, we segment the  
29    SIRT reconstruction by global thresholding. The  
30    thresholds for the segmentation are determined by  
31    solving the following optimization problem:

$$\mathbf{S}^* = \operatorname{argmin}_{\mathbf{S} \in \mathcal{S}} \mathcal{C}(I_0, p_b, \boldsymbol{\mu}, \mathbf{S}). \quad (7)$$

32    The solution of this problem is found by straightfor-  
33    ward (brute-force) sampling of the space of thresh-  
34    olds, each time evaluating the cost function. In  
35    practice, the thresholds are sampled from the min-  
36    imum to the maximum of gray levels of the SIRT  
37    reconstruction in Step 1.

38    The first iteration is again an exception since pa-  
39    rameters have not yet been estimated and the ob-  
40    jective function cannot be computed. Thus, the  
41    above segmentation method is not applicable. In-  
42    stead, the thresholds are determined using Otsu’s

method which finds optimal thresholds based on  
the gray level histograms [9].

3     *Step 3: Nonlinear parameters estimation*

4     Given the binary images, we can update the free  
5     parameters of the nonlinear model  $I_0, p_b, \boldsymbol{\mu}$  by min-  
6     imizing the objective function, which is a nonlinear  
7     regression problem. This nonlinear regression prob-  
8     lem is solved using the Nelder–Mead method [17].  
9     To improve the stability of the regression, the three  
10    parameters are estimated separately and iteratively  
11    in an inner loop:

For  $l = 1 : L$

$$\begin{aligned} p_b^{l+1} &= \operatorname{argmin}_{p_b} \mathcal{C}(I_0^l, p_b, \boldsymbol{\mu}^l, \mathbf{S}^*); \\ \boldsymbol{\mu}^{l+1} &= \operatorname{argmin}_{\boldsymbol{\mu} > 0} \mathcal{C}(I_0^l, p_b^{l+1}, \boldsymbol{\mu}, \mathbf{S}^*); \\ I_0^{l+1} &= \operatorname{argmin}_{I_0 > \max(\mathbf{p})} \mathcal{C}(I_0, p_b^{l+1}, \boldsymbol{\mu}^{l+1}, \mathbf{S}^*). \end{aligned} \quad (8)$$

here  $l$  is the iteration number of the inner loop.  
The estimation algorithm requires initial parame-  
ter values. In the experiments, we found that the  
initial values have little influence on the conver-  
gence result but proper initial values help to con-  
verge faster. Since we know that the beam inten-  
sity  $I_0$  should be at least the maximal image inten-  
sity and that the attenuation coefficients  $\mu$  and the  
bias intensity  $p_b$  are very small, we can start from  
 $I_0^1 = \max(\mathbf{p})$ ,  $p_b^1 = 0$  and  $\boldsymbol{\mu}^1 = \mathbf{0}$ , which were used  
in all the experiments in the paper.

23    *Step 4: Projection intensities rescaling*

24    Given the parameters, we rescale the measured  
25    projections  $\mathbf{p}$  to reduce nonlinear damping effects  
26    using:

$$\mathbf{p}_{lin} = \log \frac{I_0 + p_b - \mathbf{p}}{I_0}, \quad (9)$$

where  $\mathbf{p}_{lin}$  is the rescaled projections and is used  
as the input data for Step 1. At the last iteration,  
the rescaled projections are returned as the output  
 $\mathbf{p}_{lin}$ . These correspond to the linearly projected  
sum of the attenuation coefficients.

### 3. Experiments and Simulations

We report the correction of cupping artifacts for  
two sets of experimental data and three phantom

1 simulations. The experimental data show strong  
 2 nonlinear effects because the samples consist of  
 3 thick metallic materials. Two phantom simulations  
 4 resembling the experimental data were performed,  
 5 as ground-truth is missing for quality assessment of  
 6 the reconstruction image due to the lack of other  
 7 measurement methods. In addition, a phantom of  
 8 four chemical compositions was simulated to investigate  
 9 the robustness of the algorithm when more  
 10 chemical compositions are present, as the experimental  
 11 samples consist of only one or two chemical  
 12 compositions.

### 13 3.1. Experiments

14 The first experimental sample is an assembly  
 15 consisting of 16 Pt nanoparticles, each of which  
 16 has a diameter of about 60 nm (Fig. 4(a)) [18]. It  
 17 has only one chemical composition and a relatively  
 18 more complex structure than the second sample.

19 The second sample is a hetero-nanoparticle,  
 20 which is an Ag nanoparticle with a diameter of approximately  
 21 110 nm with an embedded Au octahedron [7]. It is studied  
 22 as a case where the cupping artifacts reduce the image contrast  
 23 between different chemical compositions. The specifications of  
 24 data acquisition are listed in Table 1.  
 25

26 This dataset has been used to investigate  
 27 HAADF-EDS bimodal tomography (HEBT) in [7].  
 28 In that study, the authors have noticed that the  
 29 raw data had strong intensity damping which not  
 30 only limited straightforward segmentation of the  
 31 HAADF reconstructions but also undermined the  
 32 validity of HEBT based on linear models. Therefore,  
 33 in [7] the data has been linearized in the data  
 34 preprocessing as mentioned in section 3.2 of [7].

#### 35 3.1.1. Results: Nanoparticle Assembly

36 Fig. 5 (a) is the initial SIRT reconstruction,  
 37 based on which a binary image (Fig. 5(c)) was segmented  
 38 using Otsu's method. Fig. 5 (b) and (d) are the  
 39 reconstruction and the binary image acquired after  
 40 applying the correction algorithm. To obtain  
 41 morphological information which is difficult to observe  
 42 in the reconstruction images, we plotted their edges  
 43 (Fig. 5 (e)) which are detected using a Sobel filter  
 44 that depends on the derivatives of gray levels.

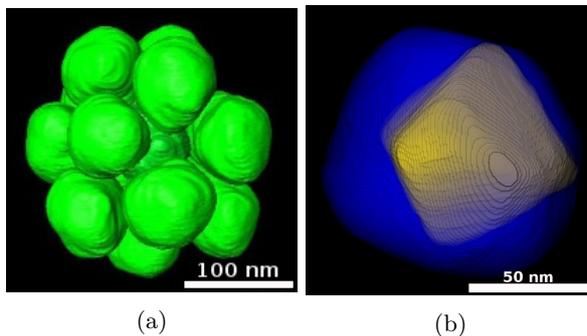


Figure 4: (a): 3D volume rendering of the Pt nanoparticle assembly. (b): 3D volume rendering of the Au-Ag nanoparticle.

In addition, the fidelity of the nonlinear regression for the nonlinear model was investigated. The fitted nonlinear model w.r.t thickness is plotted in Fig 6, where the thickness was computed as the forward projection of the binary image after correction. The error bars indicate the mean intensities and the standard deviations of the projection intensity.

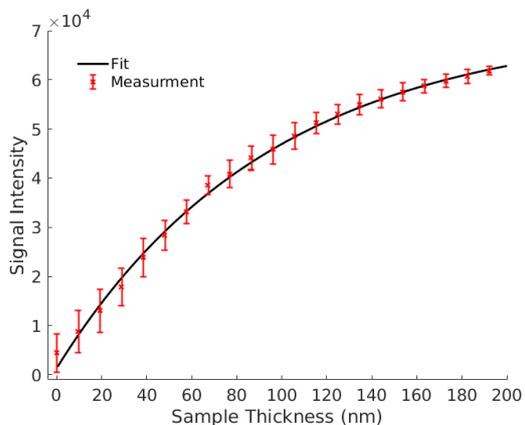


Figure 6: The nonlinear damping model fitted for projection signal intensity w.r.t. sample thickness of the nanoparticle assembly. The error bars indicate mean intensities and the standard deviations of the projection data.

#### 3.1.2. Results: Au-Ag Core-shell Nanoparticle

For this experimental data, the SIRT reconstructions and segmented binary images before and after correction are shown in Fig. 7. In addition, the line profiles across the reconstruction images for some

Table 1: Data acquisition specifications.

specimen	nanoparticle assembly	core-shell nanoparticle
electron microscope	Tecnai G2, FEI company	Tecnai Osiris, FEI company
accelerating voltage	200 kV	120 kV
convergence angle	16 mrad	18 mrad
HAADF detector range	82-180 mrad	54-230 mrad
projection angles range	$-74^\circ$ to $74^\circ$	$-75^\circ$ to $75^\circ$
projection angle increment	$2^\circ$	$5^\circ$

1 iterations are plotted in Fig. 8 to demonstrate how  
 2 gray levels evolve during a run of the correction  
 3 algorithm.

4 As discussed in the introduction, the nonlinear  
 5 effects also hinder adopting prior knowledge to re-  
 6 duce missing wedge artifacts. In this data, the pro-  
 7 jections were only acquired from  $-75^\circ$  to  $75^\circ$ . We  
 8 thus compared reconstructions using advanced re-  
 9 construction algorithms: total-variation minimiza-  
 10 tion (TV-min) [10], discrete algebraic reconstruc-  
 11 tion technique (DART) [14] and total variation regu-  
 12 larized DART (TVR-DART) [15], which incorpo-  
 13 rate the prior knowledge of image sparsity, discrete  
 14 gray levels and image sparsity combined with dis-  
 15 crete gray levels respectively. The images recon-  
 16 structed from the nonlinear projections and the cor-  
 17 rected projections are given in In Fig. 9.

18 Finally, we plotted the normalized residuals of  
 19 the cost function w.r.t. iterations for the two ex-  
 20 perimental data (Fig. 10). For the first and sec-  
 21 ond experiments, the cost values converge to stable  
 22 minimums after 16 and 12 iterations respectively.

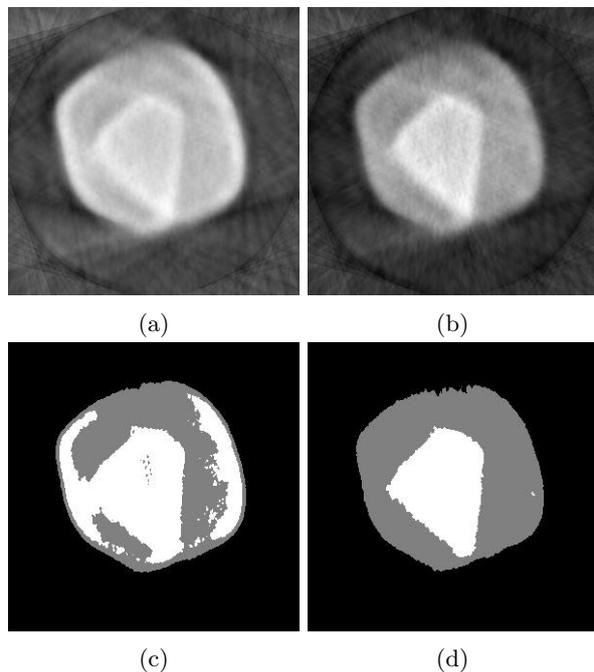


Figure 7: (a) and (b): SIRT reconstructions of the Au-Ag nanoparticle from the nonlinear projections and corrected projections. (c) and (d): Binary images segmented based on the reconstruction images (a) and (b) respectively.

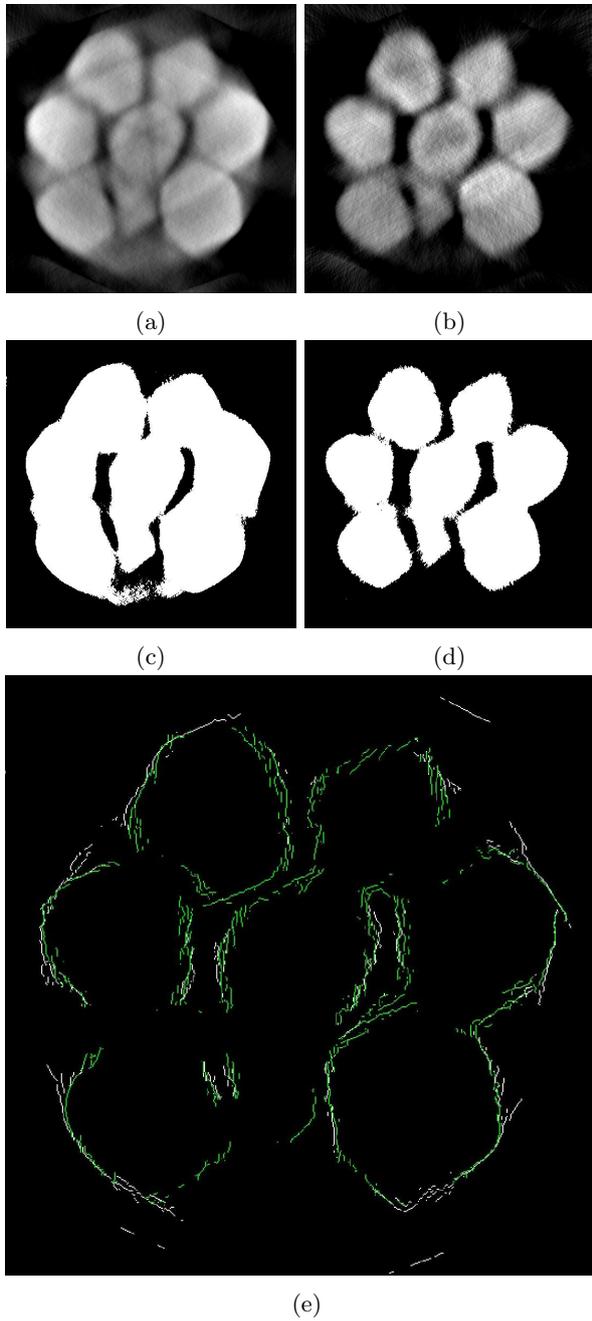


Figure 5: (a) and (b): SIRT reconstructions of the Pt nanoparticle assembly from the nonlinear projections and corrected projections respectively. (c) and (d): Binary images obtained by segmenting (a) and (b) respectively. (e) Edges of reconstructions before (white) and after correction (green).

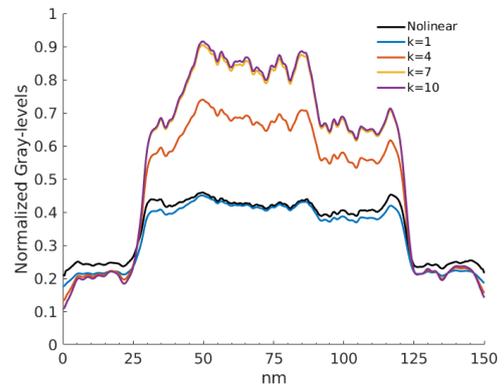


Figure 8: Cross-section line profiles of the SIRT reconstructions of the Au-Ag nanoparticle at different iterations.

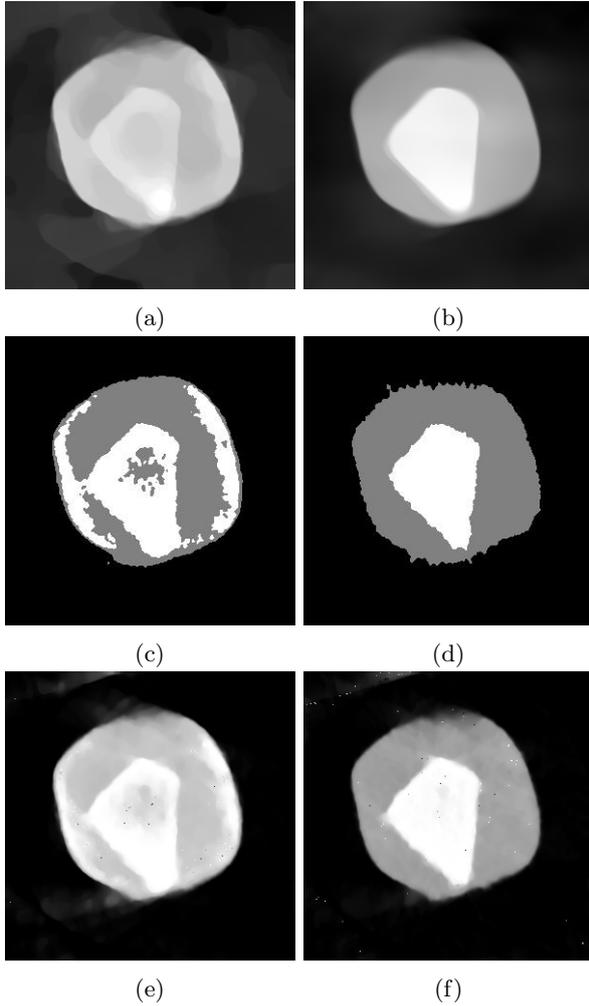


Figure 9: (a)/(b), (c)/(d) and (e)/(f) are the TV-min, DART and TVR-DART reconstructions of the Au-Ag nanoparticle from projections before/after the correction respectively.

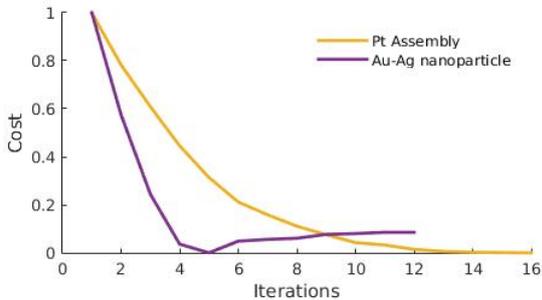


Figure 10: The residuals of cost function (Eq. 10) w.r.t. iterations for the two experimental datasets.

### 3.2. Phantom Simulations

First of all, two phantom simulations were made resembling the two experimental datasets. Note that the purpose of the simulation is not to validate the nonlinear model, but to assess the quality of nonlinear correction assuming the nonlinear forward model is accurate once all model parameters have been accurately obtained. For each sample, we first applied the correction algorithm to the experimental data to obtain binary images and nonlinear forward models. Afterwards, projections were simulated by projecting the binary images based on the nonlinear model. In addition, Gaussian noise was added to the projections to make the simulation more realistic.

The simulations provide ground-truth to quantify the quality of reconstructions. Here, the error metric is defined as the mean difference between the reconstructed and the ground-truth binary images:

$$err = \frac{1}{K} \sum_e^K \sum_j^N \|s_{ej} - g_{ej}\| / \sum_j^N g_{ej}, \quad (10)$$

where  $\{g_{ej}\}$  are the ground-truth binary images.

The third phantom simulation, focused on the correction for more than two chemical compositions, was made using the same shapes as the nanoparticle assembly phantom. What is different is that instead of having one composition for all particles, there are particles of four different compositions, each having a different atomic number. Then projections were made by projecting the particles based on the nonlinear model.

#### 3.2.1. Results of Simulations

The first phantom resembles the nanoparticle assembly, whose contours are plotted in Fig. 11 (c) and (d). Fig. 11 (a) is the initial SIRT reconstruction before correction, based on which a binary image (Fig. 1(c)) was segmented. Fig. 11 (b) and (d) show the SIRT reconstruction and the binary image after applying the correction algorithm. The error metrics of the binary images are respectively 5% and 2% before and after correction.

The results of the second phantom simulation are shown in Fig. 12, where (a) and (b) are the SIRT

1 reconstructions before and after correction respec-  
 2 tively. The binary images in Fig. 12 (c) and (d)  
 3 were segmented from the SIRT reconstruction im-  
 4 ages. The ground-truth phantom is plotted using  
 5 red and green contours for Au and Ag respectively.  
 6 The error metrics of the binary images are respec-  
 7 tively 56% and 1% before and after correction.

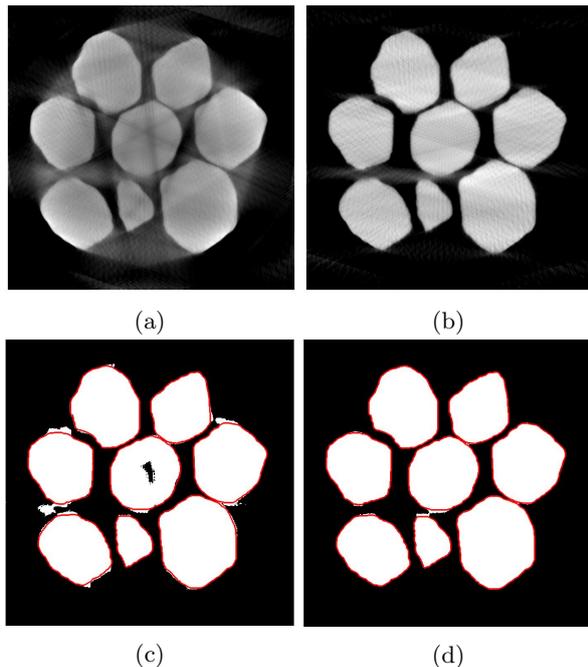


Figure 11: (a) and (b): SIRT Reconstruction images of the nanoparticle assembly phantom simulation before and after the nonlinearity correction. (c) and (d): Binary images segmented based on (a) and (b) respectively. The red contour shows the shape of the phantom.

8 The third phantom simulation presents the case  
 9 when four chemical compositions exist in the same  
 10 phantom. The SIRT reconstruction images before  
 11 and after correcting the nonlinearity are shown in  
 12 Fig. 13 (a) and (b) respectively, while the corre-  
 13 sponding binary images are given in Fig. 13 (c)  
 14 and (d). The error metrics of the binary images  
 15 are respectively 69% and 20% before and after cor-  
 16 rection.

### 17 3.3. Discussion

18 In the initial reconstruction of the nanoparticle  
 19 assembly (Fig. 5(a)), the artifacts appear, on one

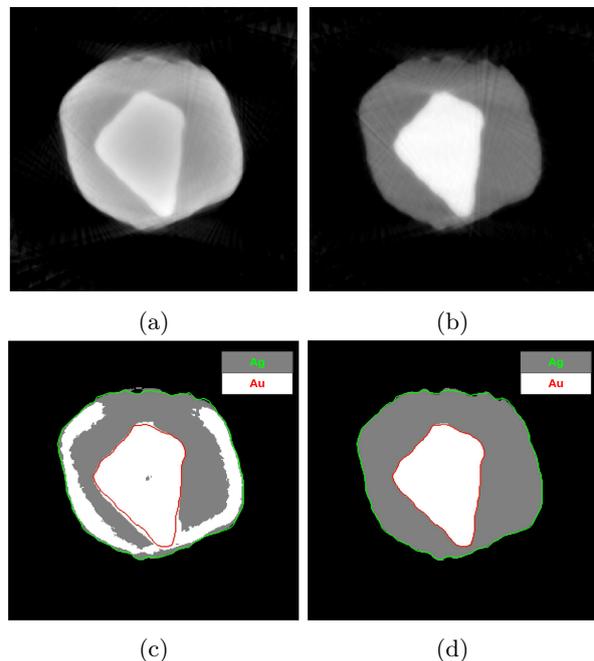


Figure 12: (a) and (b): SIRT reconstructions of the Au-Ag nanoparticle phantom simulation before and after the nonlinearity correction. (c) and (d): Binary images segmented based on (a) and (b) respectively. The red and green contours show the shape of the phantoms of Au and Ag respectively.

hand, as dark streaks elongated from the gaps be-  
 2 tween particles. On the other, they appear as un-  
 3 derestimated gray levels in the interior, for which  
 4 we see missing pixels in the binary image (Fig.  
 5 5(c)).

6 The correction algorithm successfully reduced  
 7 these artifacts and produced images easier to in-  
 8 terpret. The correction algorithm also changed  
 9 the morphology of the reconstruction image (Fig.  
 10 5(b)), as can be seen from the plot of edges. The  
 11 change may be due to the removal of the overes-  
 12 timated gray levels on the background. The plot  
 13 of fitting (Fig. 6) shows that the experimental  
 14 data matches our nonlinear model, demonstrating  
 15 a damping effect following the exponential rule. It  
 16 is also noticeable that the standard deviations de-  
 17 crease at large thickness, which can be explained by  
 18 noting that the errors introduced by segmentation  
 19 are relatively smaller at larger thickness.

20 In the initial SIRT reconstruction image of the  
 21 Au-Ag nanoparticle (Fig. 7(a)), the cupping arti-

Table 2: Errors Metrics of Binary Images.

	before correction	after correction
nanoparticle assembly phantom	5%	2%
Au-Ag nanoparticle phantom	56%	1%
phantom of four chemical compositions	69%	20%

1 facts caused the loss of contrast between Au and  
 2 Ag, even though Au and Ag have a large difference  
 3 in atomic number. As a result, many pixels were  
 4 misclassified in the binary images (Fig. 7(b)). The  
 5 algorithm corrected the experimental data and en-  
 6 hanced the contrast between Au and Ag. Demon-  
 7 strated in Fig. 8, the contrast between Au(center)  
 8 and Ag(outskirts) was enhanced step by step. At  
 9 last, the Au and Ag particles were segmented cor-  
 10 rectly based on gray levels.

11 The Au-Ag nanoparticle should be suitable for  
 12 incorporating prior knowledge to correct missing  
 13 wedge artifacts. It contains two distinct compo-  
 14 sitions with uniform densities, and thus the recon-  
 15 struction image should be sparse and have constant  
 16 gray levels. However, before the correction, incor-  
 17 porating different variants of prior knowledge in the  
 18 reconstruction actually appears to be detrimental  
 19 to the image quality, as can be seen in Fig. 9. Espe-  
 20 cially the tip of the Au particle was expanded. The  
 21 expanded tip probably is a mixture of cupping arti-  
 22 facts and missing wedge artifacts. After correcting  
 23 the nonlinear effects, the linearized projection data  
 24 was suitable for using the advanced algorithms as  
 25 the reconstructions show.

26 The first two phantom simulations show artifacts  
 27 (in Fig. 11(a) and Fig. 12(a)) very similar to those  
 28 from the experimental data, which indicates that  
 29 the modeling of nonlinear effects is accurate. Both  
 30 reconstructions after correction are free of these arti-  
 31 facts, and are in good agreement with the ground-  
 32 truth phantom, as the error metrics were reduced  
 33 (Table 2).

34 For the third simulation, we see cupping artifacts  
 35 (Fig. 13(a)) with features observed in the previ-  
 36 ous two cases. First, there are dark streaks and  
 37 underestimated gray levels. Second, the contrast

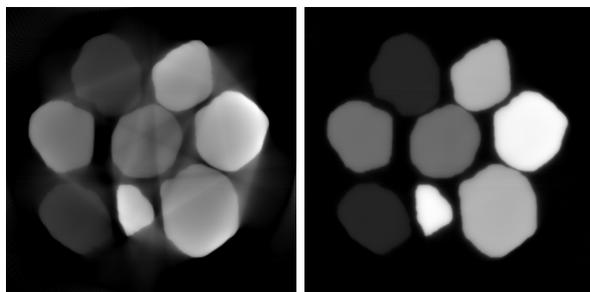
between different chemical compositions is blurred.  
 These artifacts were corrected after applying the  
 correction algorithm (Fig. 13(b)).

The segmented binary images after correction  
 (Fig. 13(d)) show a stack of different chemical com-  
 positions at the borders of some particles. How-  
 ever, these misclassified pixels are not caused by  
 the cupping artifacts, but due to the limitation of  
 the global thresholding [19]. The gray levels in the  
 reconstruction image are continuously dropping at  
 the borders. These pixels were classified into par-  
 ticles of smaller gray levels. Despite the imperfect  
 segmentation, the correction algorithm converged  
 to a result free from cupping artifacts, which also  
 indicates the good robustness of the algorithm.

#### 4. Conclusion

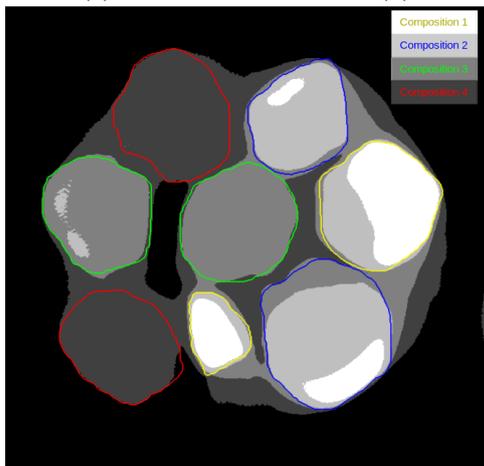
In this paper, we proposed an iterative algorithm  
 to automatically correct the cupping artifacts in  
 tomographic reconstructions from HAADF-STEM  
 projections with nonlinearly damping intensities  
 using only the projection data. The correction is  
 based on modeling the nonlinear relationship be-  
 tween projection intensities and sample thickness  
 as an exponential function.

We showed that the algorithm is an effective tool  
 in achieving better tomographic reconstructions. It  
 successfully corrected the nonlinear damping effects  
 and the subsequent cupping artifacts in three cases  
 where one, two and four chemical compositions are  
 present respectively. The correction is useful for  
 improving the accuracy of morphological analysis  
 and compositional analysis for 3D nanostructures  
 and nanomaterials. In addition, users can benefit  
 from it in enhancing the Z-contrast between chem-  
 ical compositions as well as in facilitating incorpo-  
 rating prior knowledge to correct the missing wedge  
 artifacts.

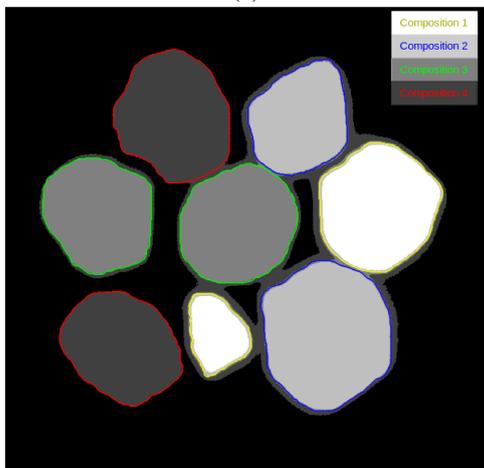


(a)

(b)



(c)



(d)

Figure 13: (a) and (b): SIRT reconstructions of the phantom simulation with four chemical compositions before and after correcting the nonlinear effects. (c) and (d): Binary images segmented based on (a) and (b) respectively. The colorful contours show the shape of the phantom particles of four different chemical compositions.

For limited data (e.g. with only a small range of tilts), the correction algorithms may fail due to the inaccurate segmentation caused by the dominant missing wedge artifacts. Potentially, this issue may be addressed by replacing SIRT and possibly the segmentation step by an advanced reconstruction algorithm (e.g. TVR-DART). However, it is still an unsolved question how to automatically set the parameters of the reconstruction algorithms, which has to be done in each iteration of the correction algorithm.

Note that the algorithm is only applicable to samples consist of several chemical compositions with homogeneous densities that can be segmented based on images gray levels. This is because the graylevel-based segmentation method fails easily when the chemical compositions are mixed or have similar atomic numbers. Moreover, this segmentation method is a global thresholding method. It may lead to poor initial segmentation results and consequently failed corrections when the cupping artifacts are very strong. Consequently, the next step of improving the algorithm is to incorporate advanced segmentation methods or spectroscopic techniques to determine the distributions of chemical compositions.

## 5. Acknowledgement

This research is supported by the Dutch Technology Foundation STW (<http://www.stw.nl/>), which is part of the Netherlands Organization for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs, Agriculture and Innovation under project number 13314. Funding from the European Research Council (Starting Grant No. COLOURATOMS 335078) is acknowledged by S. Bals. The authors would like to thank Dr. Thomas Altantzis and Dr. Bart Goris for providing the experimental data, and Prof. Dr. Luis M. Liz-Marzan for providing the investigated samples.

- 1 [1] P. A. Midgley, M. Weyland, J. M. Thomas, B. F. G. Johnson, Z-Contrast tomography: a technique in three-dimensional nanostructural analysis based on Rutherford scattering, *Chem. Commun.* (2001) 907–908.
- 2  
3  
4  
5
- 6 [2] C. Kübel, A. Voigt, R. Schoenmakers, M. Otten, D. Su, T.-C. Lee, A. Carlsson, J. Bradley, Recent advances in electron tomography: TEM and HAADF-STEM tomography for materials science and semiconductor applications, *Microscopy and Microanalysis* 11 (05) (2005) 378–400.
- 7  
8  
9  
10  
11  
12
- 13 [3] A. C. Kak, M. Slaney, *Principles of computerized tomographic imaging*, IEEE press, 1988.
- 14
- 15 [4] P. A. Midgley, M. Weyland, 3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography, *Ultramicroscopy* 96 (3) (2003) 413–431.
- 16  
17  
18
- 19 [5] R. Aveyard, Z. Zhong, K. J. Batenburg, B. Rieger, Optimizing experimental parameters for the projection requirement in HAADF-STEM tomography, *Ultramicroscopy* 177 (2017) 84 – 90.
- 20  
21  
22  
23
- 24 [6] 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.
- 25  
26  
27  
28
- 29 [7] Z. Zhong, B. Goris, R. Schoenmakers, S. Bals, K. J. Batenburg, A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM, *Ultramicroscopy* 174 (2017) 35 – 45.
- 30  
31  
32  
33
- 34 [8] J. Gregor, T. Benson, Computational Analysis and Improvement of SIRT, *IEEE Transactions on Medical Imaging* 27 (7) (2008) 918–924.
- 35  
36
- 37 [9] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics* 9 (1) (1979) 62–66.
- 38  
39  
40
- 41 [10] B. Goris, W. Van den Broek, K. J. Batenburg, H. Heidari Mezerji, S. Bals, Electron tomography based on a total variation minimization reconstruction technique, *Ultramicroscopy* 113 (2012) 120–130.
- 42  
43  
44  
45
- [11] G. Van Gompel, K. Van Slambrouck, M. De-frise, K. J. Batenburg, J. de Mey, J. Sijbers, J. Nuyts, Iterative correction of beam hardening artifacts in CT, *Medical Physics* 38 (S1) (2011) S36.
- [12] J. Hsieh, R. C. Molthen, C. a. Dawson, R. H. Johnson, An iterative approach to the beam hardening correction in cone beam CT, *Medical physics* 27 (1) (2000) 23–29.
- [13] E. V. de Castele, D. V. Dyck, J. Sijbers, E. Raman, An energy-based beam hardening model in tomography, *Physics in Medicine & Biology* 47 (23) (2002) 4181.
- [14] 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 (6) (2009) 730–740.
- [15] X. Zhuge, W. J. Palenstijn, K. J. Batenburg, TVR-DART: a more robust algorithm for discrete tomography from limited projection data with automated gray value estimation, *IEEE Transactions on Image Processing* 25 (1) (2016) 455–468.
- [16] D. B. Williams, C. B. Carter, *The transmission electron microscope*, Springer, 1996.
- [17] J. Lagarias, J. Reeds, H. Wright, P. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions, *Journal of Optimization* 9 (1) (1998) 112–117.
- [18] A. Sánchez-Iglesias, M. Grzelczak, T. Altantzis, B. Goris, J. Pérez-Juste, S. Bals, G. Van Tendeloo, S. H. Donaldson, B. F. Chmelka, J. N. Israelachvili, L. M. Liz-Marzán, Hydrophobic interactions modulate self-assembly of nanoparticles, *ACS Nano* 6 (12) (2012) 11059–11065.
- [19] K. Batenburg, J. Sijbers, Adaptive thresholding of tomograms by projection distance minimization, *Pattern Recognition* 42 (10) (2009) 2297 – 2305.