Reconstruction methods for combined HAADF-STEM and EDS tomography

Zhong, Z.

Reconstruction Methods for Combined HAADF-STEM and EDS Tomography

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op maandag 10 december 2018
klokke 11:15 uur

door

Zhichao Zhong

geboren te Guangdong, China
in 1987

Eerste Promotor:	Prof. dr. K.J. Batenburg	(Universiteit Leiden)
Tweede Promotor:	Prof. dr. B. Rieger	(TU Delft)

Samenstelling van de promotiecommissie:

Voorzitter:	Prof. dr. A.W. van der Vaart	(Universiteit Leiden)
Secretaris:	Prof. dr. S.J. Edixhoven	(Universiteit Leiden)
Overige leden:	Prof. dr. S. Bals	(Universiteit Antwerpen)
	Prof. dr. ir. F.J. Verbeek	(Universiteit Leiden)
	Dr. T. van Leeuwen	(Universiteit Utrecht)

Reconstruction Methods for Combined HAADF-STEM and EDS Tomography

Zhichao Zhong

The research in this thesis is partly financed by the Netherlands Organisation for Scientific Research (NWO), research programme "Automated multi-modal tomography for sub-22nm IC nodes", project number 13314. The research was carried out at the Centrum Wiskunde & Informatica (CWI), Amsterdam, the Netherlands.

The photo on the cover was taken by Xinyi Zhao in Beijing. The clear but colorless view above water represents HAADF-STEM imaging, while the colorful but fuzzy reflection in the water represents EDS imaging.

Printed by: Gildeprint - Enschede
ISBN: 978-94-6323-417-7

Contents

1 Introduction and outline

The PhD research that is collected in this thesis is dedicated to the development of novel tomographic reconstruction methods for characterizing both structural and compositional information of nanomaterials. In this chapter, first an introduction is given to the two imaging modalities in electron tomography that are central to this thesis: high angle annular dark field scanning transmission electron microscopy (HAADF-STEM) and energy dispersive X-ray spectroscopy (EDS). Second, we formulate the mathematics for the problem of tomographic reconstruction. Finally, we discuss the challenges for performing accurate tomographic reconstructions for individual chemical elements based on these modalities, and give an overview for several methods developed during the PhD research.

1.1 Electron microscopy and tomography

An electron microscope (EM) uses accelerated electrons instead of visible light to image materials at the nanoscale. Conventional EM is based on the same principle as optical microscopy, but with an electron source and electromagnetic lenses. As illustrated in Figure 1.1, a parallel electron beam is formed and used to illuminate the sample. An image is then formed and projected on the camera. This conventional imaging mode is called transmission electron microscopy (TEM). At present, the highest resolution realized in high-resolution TEM is around 0.5 [Kis+08].

It is also possible to focus the electron beam to an atomic-size probe, and use the probe to scan across the sample. The image is formed by measuring the intensity of transmitted or scattered electrons at every scanning position. This imaging mode is called scanning transmission electron microscopy (STEM). One popular technique to detect the intensity of electrons in STEM is high angle annular dark

field (HAADF). As Figure 1.2 shows, an annular detector collects the electrons that are scattered to high collection angles (> 60 mrad). Ideally, the intensity of the electrons scattered to such high angles increases monotonically with respect to the mass and thickness of materials. In comparison, when imaging crystalline materials, TEM images are affected by diffraction of electons in addition to the sample thickness, which makes interpretation difficult [Küb+05].

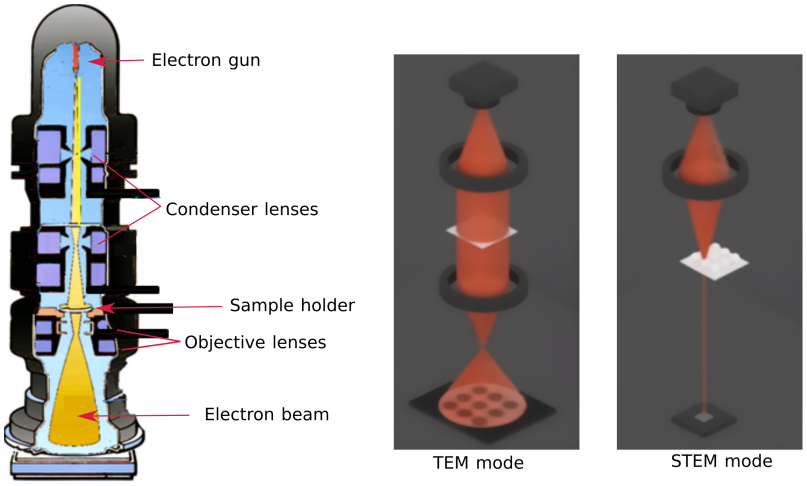


Figure 1.1: Overview of a TEM. Accelerated electrons are generated by the electron gun. In TEM mode, a parallel electron beam is formed by the condenser lenses and used to illuminate the sample placed on the sample holder. The objective lenses form an image of the sample. In STEM mode, the electron beam is focused to a small probe by a set of condenser lenses and an objective lens. The probe is used to scan across the sample. The intensity of transmitted or scattered electrons is measured as the image intensity at each scanning position. (The pictures are adapted from wikipedia.org.)

The monotonic relationship between signal intensities and sample thickness is an important property of HAADF-STEM [Mid+01]. Consider a sample of thickness t with homogeneous attenuation coefficient μ as illustrated in Figure 1.3. Assume that the HAADF detector collects almost all the electrons scattered to angles higher than the inner collection angle of the annular detector, which are complementary to the electrons passing through the hole. According to the Beer-Lambert law, the intensity of electrons that are transmitted or scattered to small angles is approximately given by:

$$I_t = I_0 \exp(-\mu t), \tag{1.1}$$

where I_0 is the intensity of the incident electron beam. As a result, the complementary HAADF signal intensity I is:

$$I = I_0 - I_0 \exp(-\mu t). \tag{1.2}$$

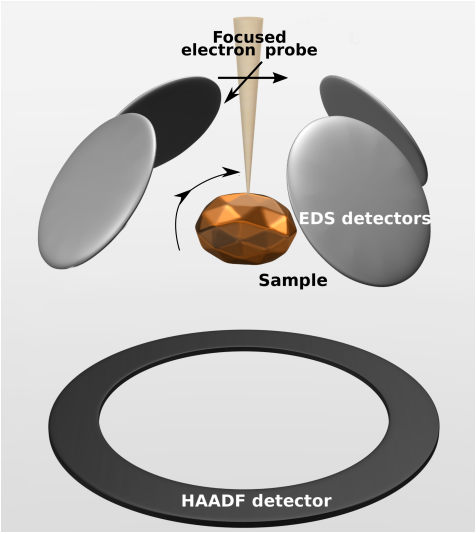


Figure 1.2: Schematic of simultaneous HAADF and EDS imaging in STEM mode. The focused electron probe moves in directions perpendicular to the beam direction. The HAADF detector collects the electrons scattered to high angles. The four EDS detectors are positioned symmetrically around the sample, and collect the X-rays emitted from the sample. The sample can be tilted by rotating the sample holder (not shown in this figure) for tomographic experiments.

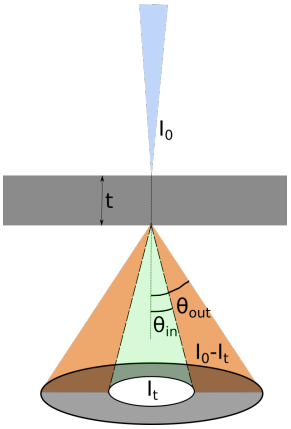


Figure 1.3: Electrons scattered by a thin specimen. The specimen has a thickness of t and a homogeneous density. An incident beam with intensity I_0 is focused on the sample. I_t, which is the intensity of electrons transmitted or scattered to angles smaller than the inner collection angle θ_{in}, can be modeled by the Beer-Lambert law. Assuming that the electrons scattered to angles higher than the outer collection angle θ_{out} are few, the intensity of electrons collected by the detector is given by $I_0 - I_t$.

For small μt, a *linear* approximation of Eq.1.2 can be derived using the Taylor

expansion:

$$I \approx I_0 \mu t, \tag{1.3}$$

which indicates that the signal intensity is approximately linearly proportional to the sample thickness weighted by the attenuation coefficient [MW03]. In practice, the validity of the linearity assumption is affected by a variety of factors, such as the atomic number and thickness of sample, the inner and outer collection angle of the annular detector as well as the accelerating voltage [AR16].

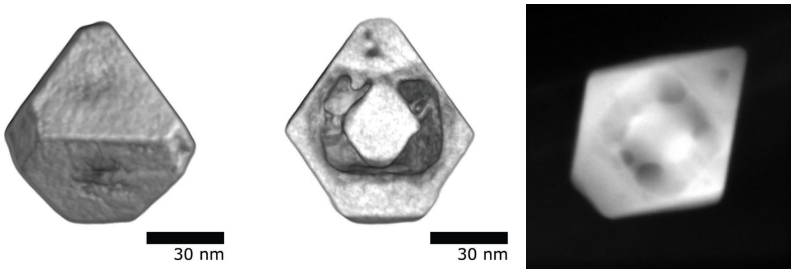


Figure 1.4: HAADF-STEM for a Au-Ag nanoparticle. Left: volume rendering of the tomographic reconstruction of the sample. Middle: inner structure of the reconstruction. Right: a HAADF-STEM projection image for the nanoparticle. (The pictures are adapted from [Zan+16a]. The sample is prepared by the Bionanoplasmonics Laboratory, CIC biomaGUNE, Spain. The HAADF-STEM projection data is provided by EMAT, University of Antwerp, Belgium.)

Another property of HAADF-STEM is that the signal intensity also depends on the atomic number Z of the projected materials, which is referred to as the *Z-contrast*. Based on empirical studies, the attenuation coefficient μ is proportional to Z^α, where α approaches 2 for an ideal detector [Mid+01; Tre11]. The image contrast depending on Z can show the variation of chemical composition and be used for compositional analysis.

However, when the difference in Z is small or chemical elements are mixed, compositional analysis based on HAADF-STEM is difficult. For example, Figure 1.4 shows the HAADF-STEM image of a nanoparticle composed of Au ($Z = 79$) and Ag ($Z = 47$) that are alloyed. It is difficult to separate the Au and Ag in the projection image, while knowing their concentrations is essential for understanding the sample.

In these situations, it is possible to apply spectroscopic techniques that resolve chemical information based on analyzing the energy of radiation (electrons, X-rays, etc.). The spectroscopic techniques that can be combined with STEM include energy dispersive X-ray spectroscopy (EDS) [Gen+12; Lep+13; CM17] and electron energy loss spectroscopy (EELS) [Jar+09; Yed+12; Hab+14]. Both techniques can be used for chemical mapping since the signal intensities are related to the concentrations of chemical elements, and have their own strengths and weaknesses. For example, EDS is more suitable for elements with high atomic numbers compared to EELS, and vice versa. In this thesis we focus on EDS, motivated by the

challenge of imaging semiconductor components in 3D at the nanoscale. In this application domain the key chemical elements match well with the applicability scope of EDS. Moreover, the methods developed in this thesis potentially can be adapted for EELS, as both techniques yield element-specific images.

Figure 1.2 also shows a modern EDS detection system with four energy-resolved X-ray detectors positioned symmetrically around the sample. For each scanning position in STEM mode, the energy-resolved detectors collect the X-rays emitted by the atoms that are excited by the electron beam. The detectors also measure the energy of every incoming X-ray photon, and generate raw data as a three-dimensional (3D) data cube that is illustrated in Figure 1.5. In each pixel of the 2D array of the data cube, there is a spectrum of X-ray counts for over 1000 energy channels. Such a data cube is called a *spectral image*. Figure 1.6 shows examples of EDS spectra for one pixel and for all X-rays integrated over all pixels. From the spectral image, characteristic X-rays, which are emitted from transitions between different electronic shells of a certain chemical element, are integrated over a narrow band of energy channels to form an image. The image can be seen as the 2D projection of the concentration of an element, and is referred to as the *elemental map* [WC16, Chapter 16]. Figure 1.7 shows the elemental maps for Au and Ag of the nanoparticle in Figure 1.4.

While HAADF-STEM gives the projection of all atoms, EDS yields multiple elemental maps, each showing the concentration of a single element. For a thin-film sample with a uniform thickness, assuming that the variation of electron intensity and the interaction between the generated X-rays and the sample are negligible, the image intensity I_a, which is the intensity of characteristic X-rays for element a, is proportional to the sample thickness t and the concentration of the element [WW06]. This relationship is expressed as:

$$I_a = \zeta_a D C_a \rho t, \tag{1.4}$$

where ζ_a is the sensitivity factor, D is the total electron dose, C_a is the concentration of chemical element a, ρ is the density. The mass-thickness of element a is given by $C_a \rho t$.

The sensitivity factor ζ is defined in the so-called ζ-factor method for quantifying the elemental compositions in a sample based on the EDS signal intensity [WW06]. Theoretically, the ζ factor is determined by the ionization cross-section, the fluorescence yield, the relative transition probability, the atomic weight, the detector collection-angle, and the detector efficiency. In practice, the value for a specific element can be estimated using a pure-element standard sample, given that the signal intensity, the total electron dose, and the mass-thickness are known. Quantifying the elemental compositions based on the ζ-factor is a relatively recent development. A more common approach is the so-called Cliff-Lorimer approach, which relates the the signal intensities to the elemental compositions but not to the concentration of a specific element [CL75]. Moreover, the accuracy of this method is limited when lacking calibration samples of which the compositions are accurately known. The accuracy of the ζ-factor method was also limited in the past

due to the difficulty in precisely measuring the thickness of the standard sample. Zanaga *et al.* propose to estimate the ζ factors based on thickness measured using electron tomography which makes the ζ-factor method more reliable and feasible [Zan+16b].

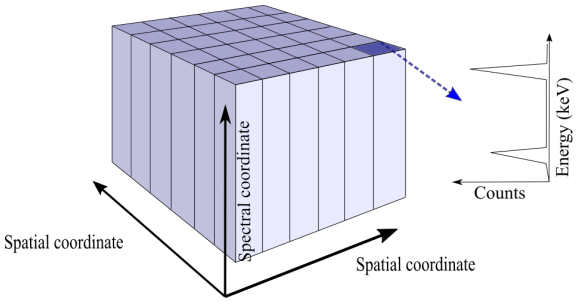


Figure 1.5: Schematic of the EDS data cube acquired in STEM mode. The data cube (spectral image) consists of a full spectrum in each pixel in a 2D array. Each spectrum contains the X-ray intensities at 1000 or more energy channels.

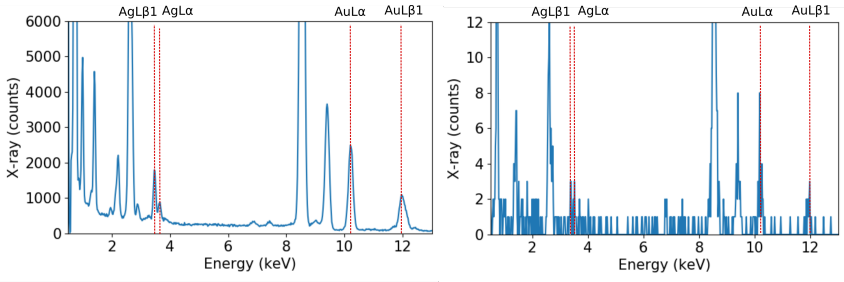


Figure 1.6: EDS X-ray spectra for the Au-Ag nanoparticle showing in Figure 1.4. Left: all counts integrated over all pixels. Right: X-ray counts of one pixel of the 128×128×1024 spectral image. The dashed lines indicate the characteristic energies for Au and Ag. Elemental maps are extracted by integrating X-ray counts over the energy channels near the characteristic lines. (The EDS data is provided by EMAT, University of Antwerp.)

The most critical issue for EDS mapping is the strong noise, as the examples in Figure 1.7 show. The noise is mainly due to the small number of X-ray photons being detected (e.g. less than 10 counts per pixel) [WC16, Chapter 16]. A key factor limiting the X-ray detection is the small area covered by detectors. While the possibilities of X-ray emission in all directions are the same, the maximal total solid angle covered by the four detectors is about 0.7 sr, which means only about 6% of the emitted X-rays can be collected [Kra+17] compared to the full $4\ \pi$ emission solid angle. The signal-to-noise ratio can be enhanced by applying a higher electron dose. This can be realized by increasing the scanning time and/or the beam current. However, a long scanning time is often accompanied by spatial drift of the sample, and increasing the beam current is limited by how much current

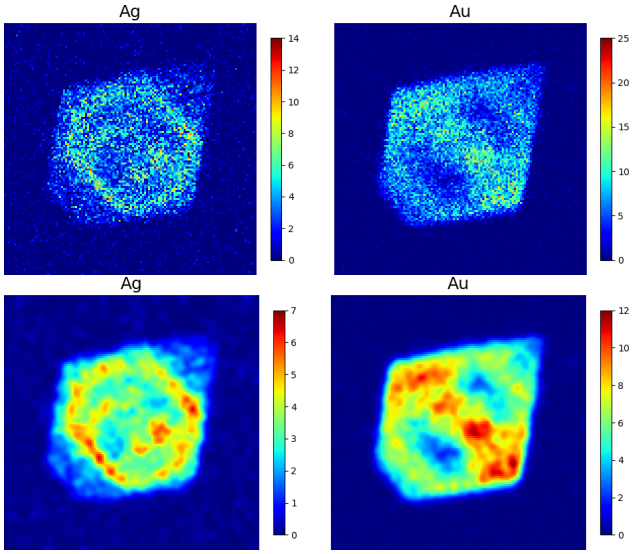


Figure 1.7: EDS elemental maps for Au (right) and Ag (left) of the Au-Ag nanoparticle showing in Figure 1.4, before (upper row) and after (lower row) being smoothed by a Gaussian filter. Note that the X-ray intensities are low (less than 25 counts), and therefore strong Poisson noise is present. (The EDS data is provided by EMAT, University of Antwerp.)

the sample can withstand.

So far, we have discussed acquiring 2D projection images using HAADF-STEM and EDS. In many research domains, such as research on the production of semiconductors, characterizing the 3D structure is crucial for understanding the physical properties. 3D imaging for nanomaterials is often performed by electron tomography (ET) – a technique to reconstruct 3D structures from a series of 2D projection images taken in different directions [SM12], [WC16, Chapter 12]. In an experiment of ET, the sample is placed on a holder which can be tilted to a certain range of angles, as illustrated in Figure 1.2 and Figure 1.8. At a certain tilt angle, a projection/spectral image is acquired using HAADF-STEM/EDS/EELS, after which the sample is rotated, and another image is acquired. At the end of the experiment, one or more tilt series of images are acquired. From a tilt series, a 3D object can be obtained using a reconstruction algorithm. Reconstruction algorithms are based on the assumption that the image intensities are proportional to the integration of some properties of the sample. The assumption is referred to as the *projection requirement*. The projection requirement is satisfied for HAADF-STEM and EDS if the assumptions for the monotonic relationships are valid (Eq.1.3 for HAADF-STEM and Eq.1.4 for EDS).

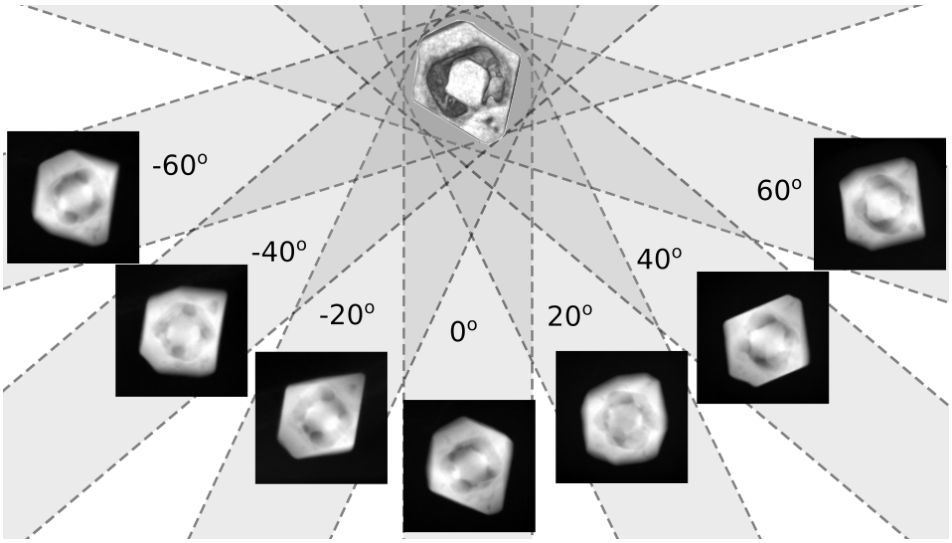


Figure 1.8: A series of HAADF-STEM images for a tomographic experiment in which a sample is tilted over an angular range of ±60°. For HAADF-STEM tomography, the increment of tilt is usually 1° or 2°

1.2 Mathematics of tomography

In this section, we will introduce the mathematical model for tomographic imaging, which is fundamental for developing novel reconstruction algorithms. Consider an electron probe with a small convergence angle (e.g. 10 mrad) and a large depth of field (e.g. 25 nm). We can approximate the probe as a ray and consider the STEM imaging process as a *parallel beam* tomographic model. When the sample is rotated over a single axis, the data collected for each slice orthogonal to the rotation axis is independent from the other slices. In this case, the 3D parallel beam model can be considered as a stack of 2D parallel beam models. In this section, we will discuss the 2D model.

1.2.1 The Radon transform

The parallel beam model can be mathematically described by the *Radon transform*. The Radon transform for 2D parallel beam is illustrated in Figure 1.9. Consider the object to be reconstructed as a function $f : \mathbb{R}^2 \to \mathbb{R}$. The projection data for the sample tilted by angle θ and the ray at scanning position u is a function $P : [0, \pi] \times \mathbb{R} \to \mathbb{R}$. Radon transform maps f to P as the integral of f along the line $l_{\theta,u}$ described by $u = x\cos\theta + y\sin\theta$, where x and y correspond to the spatial coordinates centered at the rotating axis. The projection $P_\theta(u)$ of $f(x, y)$ is given by:

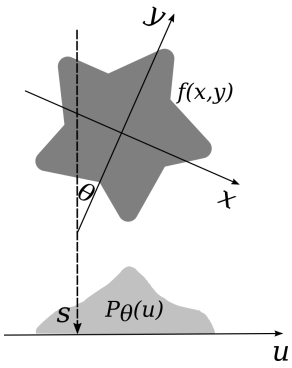


Figure 1.9: Schematic of the Radon transform of function f in 2D. The sample is tilted by angle θ. The ray at scanning position u is described by the line $u = x\cos\theta + y\sin\theta$. The projection data $P_\theta(u)$ of $f(x, y)$ is given by the line integral of $f(x, y)$ along the line.

$$P_\theta(u) = \int_{l_{\theta,u}} f(x, y)\mathrm{d}s, \tag{1.5}$$

where s is the length along line $l_{\theta,u}$.

In practice the data are measured as a discrete sampling of the continuous model, which is illustrated in Figure 1.10. The projection data are expressed as a vector $\mathbf{p} \in \mathbb{R}^M$, where M denotes the total number of pixels for all tilt angles combined. Consider the sample to be located in a 2D space discretized into N pixels as an image. The reconstructed unknowns are then expressed as a vector $\mathbf{x} \in \mathbb{R}^N$, each entry of which corresponds to a pixel value of the 2D image. For each ray i, the projection data p_i is then modeled as the weighted sum of the pixel values x_j along the ray, which is expressed as:

$$p_i = \sum_j^N w_{ij} x_j. \tag{1.6}$$

Each weight w_{ij} is determined by the area of pixel j intersected with ray i. The full set of equations for all rays is:

$$\mathbf{p} = \mathbf{W}\mathbf{x}. \tag{1.7}$$

The multiplication of matrix $\mathbf{W} = \{w_{ij}\}$ and $\mathbf{x}$ is called the *forward projection* of $\mathbf{x}$. The goal of tomographic reconstruction is to estimate the unknowns $\mathbf{x}$ from the data $\mathbf{p}$.

1.2.2 Reconstruction algorithms

One group of reconstruction algorithms is based on inverting the Radon transform Eq.1.5 to find an analytical expression for $f(x, y)$, which is known as the analytical

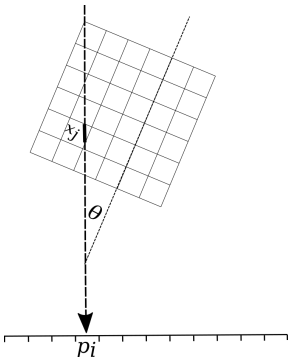


Figure 1.10: Schematic of the discrete linear projection model in 2D. The unknown corresponding to pixel j is denoted by x_j. The measurement data corresponding to the i_{th} ray is denoted by p_i. The contribution of x_j to p_i is given by $w_{ij}x_j$, where w_{ij} is determined by the area of pixel j intersected with ray i (indicated in bold).

reconstruction methods. A widely used analytical reconstruction method is *filtered backprojection* (FBP) [KS88, Chapter 3]. The FBP method has the advantage of being computationally efficient, as it only consists of a convolution operation and an integration operation. However, the analytical methods are based on the assumption that there is an infinite number of projections available. When only a limited number of projections are acquired, artifacts will appear in the reconstruction. Another group of reconstruction methods which are based on inverting the discrete model Eq.1.7 is called algebraic reconstruction methods. These methods tend to handle the problem of limited number of projections better than the analytical methods, as they do not assume that infinite number of projections are available. The unknown $\mathbf{x}$ is usually determined such that if we compute the projection of it, the *data discrepancy* $\mathfrak{D}(\mathbf{Wx}, \mathbf{p})$ between the measured $\mathbf{p}$ and the reprojected data $\mathbf{Wx}$ is minimized, which is expressed as:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}}\, \mathfrak{D}(\mathbf{Wx}, \mathbf{p}). \tag{1.8}$$

The data discrepancy is often defined as the squared l_2 norm of the difference between $\mathbf{p}$ and $\mathbf{Wx}$: $\| \mathbf{Wx} - \mathbf{p} \|_2^2$. The minimization problem can be solved by iterative algorithms, such as algebraic reconstruction technique (ART) [GBH70] and simultaneous iterative reconstruction technique (SIRT) [GB08]. Compared to the analytical algorithms, the algebraic algorithms result in fewer artifacts when data are missing or noisy. Over-fitting to noise can be prevented by stopping the algorithm early (e.g. at tens of iterations), however, at the cost of not being able to reconstruct fine structures [ENH10].

In addition to the quadratic functional, non-quadratic functionals such as Student's t, Kullback–Leibler divergence and negative log-likelihood have been examined extensively for tomographic reconstruction considering the statistical properties of the measurement data. In particular, the Kullback–Leibler divergence and

the negative log-likelihood are often used for data that are Poisson noise limited [HW16; ER07], while the quadratic functional is based on Gaussian noise model. Minimizing the data discrepancy often converges to the same result as maximizing the likelihood for measuring the data. Examples of the latter approach are the maximum likelihood estimation (MLE) and the maximum a posteriori (MAP) estimation [SV82; Gre90], which are often applied on emission tomography assuming that the measured data are random variables obeying certain statistical models.

To yield more accurate reconstructions, a popular approach is to incorporate prior knowledge about the reconstructed object [BO13; Bat+09]. The prior knowledge is often implemented as a regularization term added to the minimization problem:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}}\, \mathfrak{D}(\mathbf{W}\mathbf{x}, \mathbf{p}) + \lambda \mathfrak{R}(\mathbf{x}). \tag{1.9}$$

The regularization term $\mathfrak{R}(\mathbf{x})$ is a functional promoting some properties of the reconstructed image, and λ indicates the weight of the regularization term. A well-known example is the total variation (TV) regularization [BO13]. The reconstructed image with penalized TV tends to possess piecewise constant features.

1.3 Challenges and overview

After introducing the mathematics of tomography, we now look at the actual tomographic experiment based on simultaneously acquired HAADF-STEM and EDS. Both imaging modalities have their own advantages and disadvantages. The noise in EDS mapping is a critical issue for computing accurate tomographic reconstructions, as real measurement data deviate from the ideal projection data defined in Eq.1.7. In many EDS tomographic experiments, the noise issue is addressed by image filtering before and after reconstruction at the cost of reducing resolution [Zan+16a; Sla+16a]. The noise can also be reduced using multivariate statistical analysis methods which separate the noise from the spectral image, but with limited effectiveness when the noise dominates the data [Bur+16; Jol02].

Moreover, the number of tilts in EDS tomography is usually limited (e.g. 29 tilts for $\pm 70°$), as it takes a long time to acquire enough X-rays in every scan (about 5 to 10 mins). Consequently, it is difficult to reconstruct accurately with information from many angles missing. Furthermore, the projection matrix $\mathbf{W}$ is often ill-conditioned. This means that even small noise in measurement data can cause large errors in the reconstruction. The strong noise, together with the limited number of tilts and the ill-conditioned matrix, poses strong challenges for EDS tomographic reconstruction. Other challenges in EDS tomography include detector shadowing [Sla+16b; Sla+16a] as well as X-ray absorption and fluorescence [WW06]. The sample holder causes shadowing of X-rays on the detector, and consequently the signal intensity varies as a function of the tilt angles. To addressed the influence on the reliability of the tomographic reconstruction, novel correction methods have been proposed, such as adjusting the acquisition time at different tilts based on the detector geometry [Kra+17] and compensating the

angle-dependency in combination with HAADF-STEM tomography [Zan+16a]. In addition, an X-ray absorption correction method for EDS tomography has been proposed [Bur+16], showing improved reconstruction results. Although these effects are not the focus of this thesis, they should be carefully considered especially when conducting quantitative EDS tomography.

Compared to EDS, HAADF-STEM can yield a larger number of tilt images (e.g. 141 tilts for $\pm 70°$) with much higher signal-to-noise ratios, and consequently can lead to reconstructions with less noise and higher resolution. In addition, variation of chemical elements can be characterized based on the HAADF reconstruction as long as the difference in contrast is mainly contributed by the difference in Z. Thus, it is possible to obtain element-specific 3D objects by segmenting the HAADF-STEM tomographic reconstruction based on the Z contrast. However, this is only applicable to samples of segmentable compositions with large difference in Z.

Table 1.1 compares some properties of HAADF-STEM and EDS tomography in terms of the chemical information, noise, and the number of tilts. These properties of the two simultaneously performed modalities are complementary to each other. Attempts have been made to combine HAADF-STEM tomography and EDS tomography. For instance, HAADF-STEM tomography has been combined with EDS tomography for projection alignment, joint analysis and thickness estimation [Gor+14; Bur+16; Kra+17]. In particular, an approach has been proposed to combine HAADF-STEM and EDS tomography to acquire quantified element-specific reconstructions [Zan+16a]. In this approach, the HAADF-STEM reconstruction is computed to estimate the sample thickness which is then combined with the ratio maps of elements obtained by EDS. Given these promising results, **the overarching goal of this thesis is to explore novel methods that combine HAADF-STEM and EDS tomography to acquire accurate and element-specific reconstructions.**

Table 1.1: HAADF-STEM tomography vs EDS tomography

	HAADF-STEM	EDS
Chemical information	mixed	element-specific
Noise level	low	high
Number of tilts	large	small

As it turns out, to successfully combine the strengths of HAADF-STEM and EDS tomography, many challenges must be solved, including developing novel algorithms that can combine both modalities, making these advanced approaches generally applicable, as well as solid modeling of the image formation. This thesis mainly addresses four challenges as follows.

First, knowing that EDS maps each chemical element and HAADF-STEM maps all elements combined, one can combine the two modalities in one reconstruction process.The simultaneous reconstruction process needs to be based on a

tomographic model that is consistent with both modalities.

In Chapter 2, we introduce a technique to reconstruct for all elements simultaneously from EDS elemental maps and HAADF-STEM projection images, which is named HAADF-EDS bimodal tomography (HEBT). We assume that the HAADF-STEM projection images are the weighted sum of EDS maps for all present chemical elements. We therefore introduce a linear sum constraint to a reconstruction process for all elements, which can be solved using an iterative algorithm.

Second, while HEBT requires EDS mapping for all the chemical elements imaged by HAADF-STEM, which may be impossible sometimes, we need a different combining strategy which can be applied to one element each time. A clear combining strategy is to encourage the Z-contrast reconstruction and the element-specific reconstruction to preserve consistent image features such as common edges. The remaining questions include what feature to exploit and how to incorporate the consistency in the reconstruction process.

In Chapter 3, HAADF-STEM and EDS is combined in a way different to HEBT. We penalize so-called total nuclear variation (TNV) of a Z-contrast reconstruction made from HAADF-STEM data and an element-specific reconstruction made from EDS data, to encourage common edge locations and parallel/antiparallel gradients. This combining approach can be applied to the reconstruction for one element each time.

Third, in addition to combining modalities, many advanced reconstruction methods can also be applied to improve the accuracy of reconstruction. However, these methods are based on different assumptions for the sample and the imaging process, and a clear guideline for deciding which algorithms to use is still missing in the field of HAADF-STEM + EDS tomography. It is also possible to combine algorithms with different strengths and weaknesses in one reconstruction framework to obtain optimal results. Therefore, a framework for applying the algorithms needs to be developed.

In Chapter 4, we propose the framework to combine different advanced reconstruction algorithms for HAADF-STEM + EDS tomography. Algorithmic recipes composed of different ingredients can be applied to augment tomographic reconstruction. The ingredients mainly belong to three modules: statistical modeling, variational regularization, and HEBT. To incorporate the correct prior knowledge and physical constraints, we also provide guidelines to tailor recipes based on the experimental conditions and the samples.

Finally, the linear integral model is fundamental for combined tomographic reconstruction. However, the linearity assumption can be invalid for thick samples. For EDS, the nonlinear signals are mainly caused by the X-ray absorption of the sample. For HAADF-STEM, when the sample is thick, the signal intensity damps at large thickness as illustrated in Eq.1.2. While numerical methods to correct the absorption in EDS tomography have been proposed [Bur+16], numerical methods that require no extra experimental step to correct the intensity damping in HAADF-STEM tomography are missing.

To address the nonlinear damping effects in HAADF-STEM data, we propose an automatic correction algorithm for samples consisting of homogeneous compo-

sitions, which is described in Chapter 5. The correction algorithm only requires the projection images as input. A nonlinear model is estimated based on the reconstructed structure as well as the errors between linearly re-projected data and measurement data. It is possible to use the correction algorithm together with the correction for X-ray absorption to improve the accuracy of HAADF-STEM + EDS tomography.

2
HAADF-EDS bimodal tomography

2.1 Introduction

Electron tomography (ET) is nowadays commonly used in materials science to characterize the three-dimensional (3D) structure and composition of nanomaterials starting from a tilt series of two-dimensional (2D) projection images [Her09]. Typically, the projection images for ET in materials science are obtained using high angle annular dark field (HAADF) scanning transmission electron microscopy (STEM) [Mid+01; MW03]. Images acquired using HAADF-STEM are called Z-contrast images because the projected intensity is related to the average atomic number that is integrated along the projection direction [Mid+01; Küb+05]. Consequently, the chemical composition can be characterized in 3D. However, when investigating heteronanostructures with small differences in Z, spectroscopic techniques are required to investigate the 3D distributions of the different chemical elements.

Previously, both energy dispersive X-ray spectroscopy (EDS) [Sag+07; Gen+12; Lep+13; Sla+16b] and electron energy loss spectroscopy (EELS) [Jar+09; Yed+12; Yed+14] have been used in combination with tomographic reconstruction techniques. Both techniques require similar computational steps to produce element-

This chapter is based on:
Z. Zhong, B. Goris, R. Schoenmakers, S. Bals, and K. J. Batenburg. "A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM". *Ultramicroscopy* 174 (2017), pp. 35–45.

specific images (elemental maps) that give the 2D projections of a chemical element, which also satisfy the projection requirement for tomography under certain circumstances [Sla+16b; Yed+12]. In this study, we only focus on EDS-STEM tomography.

HAADF-STEM tomography and EDS-STEM tomography are highly complementary techniques that each have advantages and disadvantages. The major advantage of HAADF-STEM tomography in comparison to EDS-STEM tomography is that it yields reconstructions with a relatively high signal to noise ratio (SNR). However, the reconstructed image intensities contain only aggregate information of all elements, while the EDS-STEM technique yields element-specific reconstructions. So far, HAADF-STEM has been combined with EDS-STEM in ET in terms of tilt series alignment [Gor+14], density estimation [Bur+16] or thickness estimation [Zan+16a]. It is highly desirable to develop reconstruction techniques that can exploit the favorable properties of these complementary techniques simultaneously. The concept of "multimodal imaging" has been introduced in the field of medical imaging, where the data from several imaging modalities such as PET, SPECT, CT and MRI are combined in a single joint reconstruction procedure [ZMA10].

In this chapter, we introduce the multi-modal imaging concept to ET, by proposing a novel HAADF-EDS bimodal tomographic (HEBT) reconstruction technique that simultaneously reconstructs from projection images acquired by two complementary imaging modalities. In this method, chemical elements are linked in the reconstruction process but separated in the final output. The aim of our algorithm is to keep the element-specific feature of elemental maps while preserving the high SNR of Z-contrast images.

Section 2.2 will begin with discussing the mathematical models of HAADF-STEM tomography and EDS-STEM tomography. A new approach to link the models will be proposed and the HAADF-EDS bimodal tomographic reconstruction technique will be explained. In Section 2.3 and Section 2.4, we will investigate the performance of the new technique using both simulated and experimental data. In Section 2.5, the advantages and the outlook of HEBT will be discussed.

2.2 Projection models and the reconstruction method

2.2.1 HAADF-STEM and EDS-STEM imaging models

Suppose there are K chemical elements in a specimen, we have K volumetric objects as the unknowns to be reconstructed, so the distribution of each chemical element is represented by a voxel image. Images formed by HAADF-STEM and EDS-STEM are related to the density distributions of these chemical elements.

For HAADF-STEM projection images, it is known that the intensity is proportional to the number of electrons scattered at high angles. For a single atom, the number of these electrons is proportional to the scattering cross section which depends on its atomic number [Tre11; Pen89; Wal06]. For thin-film specimens

in which multiple scattering and absorption is negligible, the number of scattered electrons p^h equals the sum of scattering cross sections of all the atoms probed by the electron beam:

$$p^h = \sum_{e=1}^{K} \sigma^e N^e = \sum_{e=1}^{K} \sigma^e \frac{\int \rho^e(t)dt}{M^e}, \tag{2.1}$$

where $e = 1, 2, \dots K$ are the indices denoting the type of chemical element, σ^e is the scattering cross section, N^e is the number of atoms, $\rho^e(t)$ is the mass-thickness and M^e is the atomic weight. By defining the *HAADF-STEM response factor* $z^e = \frac{\sigma^e}{M^e}$, the image grayscale is concisely expressed as the weighted sum of mass-thickness of all atoms:

$$p^h = \sum_{e=1}^{K} z^e \int \rho^e(t)dt. \tag{2.2}$$

For the sake of numerical computation, the volume to be reconstructed is often discretized into N equally-spaced voxels. Thus, the density distribution of chemical element e is written as a vector $\boldsymbol{\rho}^e \in \mathbb{R}^N$, $e = 1, 2, \dots K$. The Z-contrast images used as tomographic reconstruction inputs are taken at different tilt angles, where every pixel specifically corresponds to a beam position and a tilted angle of the specimen. In total there are M pixels for all the tilted images. The grayscale on the i_{th} pixel is now written as an entry p_i^h in $\mathbf{p}^h \in \mathbb{R}^M$. Now the continuous line integral in Eq. 2.2 is replaced by the discrete ray-sum as:

$$p_i^h = \sum_{e=1}^{K} z^e \sum_{j=1}^{N} w_{ij} \rho_j^e, \tag{2.3}$$

where the factor w_{ij} is determined by the area intersected between the i_{th} ray integral and the j_{th} voxel [KS88, Chapter 7]. Note that in the conventional HAADF-STEM tomography where the reconstruction models are defined by $p_i^h = \sum_{j=1}^{N} w_{ij} x_j$, the reconstructed quantity is actually $\sum_{e=1}^{K} z^e \rho_j^e$, which describes the distribution of the weighted sum of densities.

Unlike in HAADF-STEM tomography where projection images contain information about all atom types simultaneously, in EDS-STEM tomography each chemical element has its own series of tilted element-specific images, which depicts the projection of the chemical element and are usually called *elemental maps* (see more in [Sch+10; Hab+14] and [WC16, Chapter 16]). Their grayscales correspond to the photon counts of the characteristic X-ray of a chemical element. Under the thin-film approximation in which X-ray absorption and fluorescence is negligible, the characteristic X-rays counts p^e for the e_{th} chemical element is proportional to the mass-thickness of this chemical element probed by the electron beam (discussed in [WW06; WC16]), which we define here as:

$$p^e = \zeta^e \int \rho^e(t)dt, \tag{2.4}$$

where ζ^e is the *EDS-STEM response factor* that characterizes how many characteristic X-ray counts are collected for a unit amount of the chemical element. Using the same notations as Eq. 2.3, the line integral relationship can be written in a discrete form as:

$$p_i^e = \zeta^e \sum_{j=1}^{N} w_{ij} \rho_j^e. \tag{2.5}$$

Based on the model, each chemical element can be characterized independently. Please note that in EDS-STEM tomography as in [Sag+07; Gen+12; Lep+13], the reconstructed quantity is the weighted density distribution $\zeta^e \rho_j^e$.

2.2.2 Linking HAADF-STEM and EDS-STEM

An obvious and internal connection between the two types of imaging techniques is that their projection images are both related to density distributions. However, the relations to density are based on different response factors (z^e and ζ^e) which are difficult to estimate. To estimate these factors, special pure-element specimens need to be prepared and measured with extra labor and cost. Moreover, estimated factors are often not reusable since their values vary for different experimental setups.

Instead, we estimate the ratio of response factors $r^e = \frac{z^e}{\zeta^e}$, which we refer to here as the *response ratio factors*, to link the two types of images. They can be estimated based on the assumption that both types of images are linearly related to the projection of density distribution. To be more specific, if we replace the $\sum_{j=1}^{N} w_{ij} \rho_j^e$ by $\frac{p_i^e}{\zeta^e}$ (according to Eq. 2.5) in Eq. 2.3, we have:

$$p_i^h = \sum_{e=1}^{K} r^e p_i^e, \tag{2.6}$$

where there are K unknowns r^e. For M pixels in the HAADF-STEM and EDS-STEM images, there is an overdetermined system of M linear equations for the K unknowns. By solving this system of linear equations (e.g. using the linear least squares method), we can estimate the response ratio factors. This can be done using only the tomographic projection images and without measuring extra specimens, and is the first step to incorporate HAADF-STEM and EDS-STEM in a simultaneous reconstruction process.

2.2.3 HAADF-EDS bimodal tomographic reconstruction

By making the substitution $x_j^e = z^e \rho_j^e$, the HAADF-STEM model of Eq. 2.3 and the EDS-STEM model of Eq. 2.5 can be rewritten as:

$$p_i^h = \sum_{e=1}^{K} \sum_{j=1}^{N} w_{ij} x_j^e, \tag{2.7}$$

$$r^e p_i^e = \sum_{j=1}^{N} w_{ij} x_j^e. \tag{2.8}$$

In a full system of equations, containing an equation for each measured value in each projection image, the above equations are written as $\mathbf{p}^h = \sum_{e=1}^{K} \mathbf{W}\mathbf{x}^e$ and $r^e \mathbf{p}^e = \mathbf{W}\mathbf{x}^e$. We see that both systems now have the same unknowns, the images $\mathbf{x}^e$ for all chemical elements. The unknowns $\mathbf{x}^e$ have the same unit as the intensities reconstructed from HAADF-STEM projections, but they can also be transformed into the quantitative distributions of the individual elements for each voxel when EDS response factors (ζ factors) are provided.

To obtain reconstructions that are maximally consistent with both HAADF-STEM and EDS-STEM data, we should minimize the following residuals for EDS-STEM and HAADF-STEM simultaneously:

$$\mathbf{x}^* = \underset{\mathbf{x}=(\mathbf{x}^{1T}\ldots\mathbf{x}^{eT}\ldots\mathbf{x}^{kT})^T}{\operatorname{argmin}} \alpha^2 \parallel \mathbf{p}^h - \sum_{e=1}^{K} \mathbf{W}\mathbf{x}^e \parallel_2^2 + (1-\alpha)^2 \sum_{e=1}^{K} \parallel r^e \mathbf{p}^e - \mathbf{W}\mathbf{x}^e \parallel_2^2, \tag{2.9}$$

where $0 < \alpha < 1$ is introduced here to balance between the HAADF-STEM and EDS-STEM terms. The square terms are weighted by α^2 so that α corresponds to the image intensity. This weighting factor determines the weight of the HAADF-STEM term in the reconstruction process and should be chosen depending on the noise level of the elemental maps. In principle, α can be arbitrarily chosen between 0 and 1. However, in practice, if α is too small, the influence from the HAADF-STEM data will be hardly observable. Our empirical studies show that a number between 0.7 and 0.9 yields consistent results that balance the influences of the two modalities for our experimental data. In Section 3, we will discuss more about how the weighting factor influences reconstruction results.

The minimization problem in Eq. (9) can be formulated as a least squares problem:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{p}^a - \mathbf{W}^a \mathbf{x} \parallel_2^2, \tag{2.10}$$

where

$$\mathbf{p}^a = \begin{pmatrix} (1-\alpha)r^1\mathbf{p}^1 \\ \vdots \\ (1-\alpha)r^e\mathbf{p}^e \\ \vdots \\ (1-\alpha)r^K\mathbf{p}^K \\ \alpha\mathbf{p}^h \end{pmatrix},$$

$$\mathbf{W}^a = \begin{pmatrix} (1-\alpha)W & \dots & \emptyset & \dots & \emptyset \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \emptyset & \dots & (1-\alpha)W & \dots & \emptyset \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \emptyset & \dots & \emptyset & \dots & (1-\alpha)W \\ \alpha W & \dots & \alpha W & \dots & \alpha W \end{pmatrix},$$

$$\text{and } \mathbf{x} = \begin{pmatrix} \mathbf{x}^1 \\ \vdots \\ \mathbf{x}^e \\ \vdots \\ \mathbf{x}^K \end{pmatrix}.$$

This least square problem can be solved using an iterative algorithm. In this chapter, the widely used simultaneous iterative reconstruction technique (SIRT) [GB08] is adopted in the experiments. To incorporate the physical constraint that the elemental composition should not have negative values, we apply a non-negativity constraint to SIRT by thresholding negative values in every iteration. We refer to the complete method as *HAADF-EDS bimodal tomography (HEBT)*.

The SIRT algorithm is more robust to noisy data than the common weighted backprojection algorithm, as it computes a weighted least-squares solution, which effectively averages the noise over all projection angles, assuming that the noise follows a normal distribution. We point out that there are tomography reconstruction algorithms that are even more robust with respect to noisy data: (i) statistical reconstruction algorithms that model the statistical distribution of the noise and (ii) algorithms that incorporate prior knowledge such as discreteness or smoothness of the image. As the noise in the EDS data is Poisson distributed, but the noise in the resulting elemental maps follows a different distribution that is difficult to model in detail, we consider the Gaussian model to be a solid choice.

2.3 Experimental design and data

We design three experiments for different purposes. In the first experiment, the HEBT algorithm is applied to reconstruct 2D images from simulation data. In

this simulation experiment, the stability of the HEBT reconstruction technique as a function of the response ratio factors can be investigated.

In the second experiment, we investigated cubic Au-Ag nanoparticles using HEBT. As the two compositions (Au and Ag) are well separated in the particle and have a substantial difference in atomic number, the 3D distribution of the different chemical elements can be investigated using HAADF-STEM tomography and does not require EDS. Here, this HAADF-STEM reconstruction can be used as ground truth to compare the quality of the HEBT reconstructions in comparison to conventional EDS reconstructions.

The key advantage of HEBT with respect to conventional HAADF-STEM reconstruction only becomes clear if the HAADF-STEM reconstruction does not allow for straightforward segmentation of the elements, either because the difference in Z-contrast between the elements is low, or because the elements are mixed at a sub-voxel resolution. In such cases, HEBT can potentially reconstruct the individual 3D elemental volumes (not possible by HAADF-STEM), while achieving a more faithful reconstruction at lower noise level compared to conventional EDS reconstructions. This advantage is illustrated by the results of the third experiment, applying the HEBT algorithm to another nanoparticle in which an alloy of Au and Ag is present.

2.3.1 Phantom simulation

The first experiment is based on a 2D phantom image shown in Figure 2.1, which was created to resemble a slice of the non-alloyed Au-Ag nanoparticle (see Figure 2.3). Figure 2.1 (a) and (b) are the Au and Ag phantom objects with homogeneous density. Figure 2.1 (c) is a Z-contrast phantom image of Au and Ag phantoms weighted by HAADF-STEM response factors that are assumed to be $z^{(Au)} = 79^{1.5}$ and $z^{(Ag)} = 47^{1.5}$ [Tre11]. To simulate projection images, tilt series of projections were computed using the ASTRA toolbox [Aar+15]. The projection geometry has 512 pixels and 31 tilt angles from $-75°$ to $75°$ with a step size of $5°$.

The HAADF-STEM sinogram (Figure 2.1 (f)), which is assumed to be low-noise, is simply assigned as the tilt series of the Z-contrast phantom. For EDS-STEM, two sinograms (Figure 2.1 (d) and (e)) were generated by applying Poisson noise to the tilt series of Au and Ag phantom objects. The EDS-STEM mapping process was simulated in a way that the X-ray count on each pixel is rendered as a random integer generated from the Poisson distribution. Based on the EDS-STEM models, the mean parameters of the Poisson distributions were assigned as the tilt series multiplied by the response factors. The response factors were selected as $\zeta^{(Au)} = 1.88 \times 10^{-2}$ and $\zeta^{(Ag)} = 2.4 \times 10^{-2}$ so that the mean expected numbers of X-ray counts approximate the mean X-ray counts in the elemental maps of the first sample (Figure 2.3 (b) and (c)). A filtering operation using an 8-pixel 1D Gaussian filter was applied to the EDS-STEM sinograms as an easy implementation of noise smoothing [Gen+13; Sla+16b]. The intensity of the HAADF-STEM sinogram is at a much larger order of magnitude than the EDS-STEM sinograms.

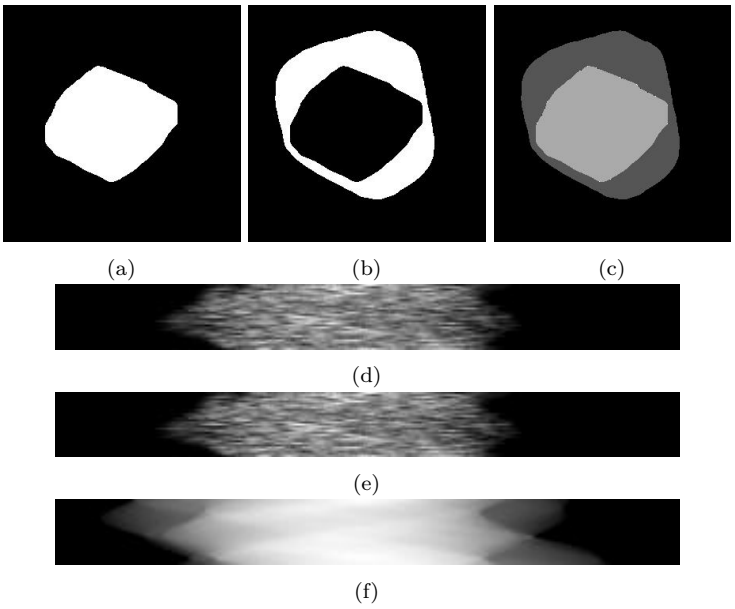


Figure 2.1: (a) The Au phantom object, (b) the Ag phantom object and (c) the Au-Ag Z-contrast phantom image. (d) The Au EDS-STEM sinogram, (e) the Ag EDS-STEM sinogram and (f) the HAADF-STEM sinogram.

2.3.2 Au-Ag nanoparticles

For the real-world experiments, tilt series of projection images were acquired using the same procedures for both Au-Ag samples. First, the sample was mounted on the tomographic holder placed in an electron microscope (Tecnai Osiris, FEI company) equipped with four silicon drift detectors (SuperX system, FEI company). During the tilt series, the sample was tilted from $-75°$ to $75°$ with a tilt increment of $5°$ for the first sample. At each tilt, a Z-contrast image was first recorded by the HAADF detector. The sample was then scanned with an acquisition time of 300 seconds to record X-rays spectrum images over 2048 energy channels. In order to reduce the shadowing effect of SDD detectors, the detectors on one side to which X-rays were blocked were turned off, while the other two detectors on the other side were turned on [Sla+14]. A tilt series of the second sample was acquired using almost the same procedures except that the sample was tilted over 29 steps from $-70°$ to $70°$. The raw data were then processed before being used as tomographic reconstruction input data. For HAADF-STEM, the tilt series of Z-contrast projection images were aligned using the cross-correlation method. The intensity damping has also been corrected by linearizing the nonlinear intensity-thickness relation [van+12]. For EDS-STEM, the spectrum images were denoised using principal component analysis (PCA) decomposition/reconstruction [Bur+16; Luc+13].

The high peaks near 8.040 keV and 8.904 keV come from Cu in the holder, which

will overwhelm and dominate the other components in PCA if they are included. To avoid this, we only took out the energy channels near the Au and Ag peaks for PCA decomposition (Figure 2.2 (b)) (the characteristic peaks are $M_\alpha = 2.15$ keV, $M_\beta = 2.20$ keV and $L_\alpha = 9.70$ keV for Au, and $L_\alpha = 2.98$ keV and $L_\beta = 3.19$ keV for Ag). After PCA decomposition, we examined every component and selected the first 15 components for PCA reconstruction and abandoned the remaining components as noise. Next, the denoised spectrum images near characteristic channels were extracted and summed up to the elemental maps (Figure 2.2 (b)).

Note that since the X-ray counts are very low for such a high resolution, even after PCA denoising the elemental maps remain very noisy. Therefore, we applied an averaging image filter with a 12 $\times$12 pixel Gaussian kernel (rotational-symmetric) to the elemental maps. Finally, the elemental maps were again aligned to match the Z-contrast images using the cross-correlation method. For each sample, the data processing steps resulted in three tilt series of projection images for each sample: two tilt series of elemental maps and one tilt series of Z-contrast images (see examples in Figure 2.3).

2.4 Experimental results

In addition to the HEBT reconstructions, we also computed HAADF-STEM tomographic reconstructions from Z-contrast projection images and EDS-STEM tomographic reconstructions from elemental maps. All the reconstructions were computed using the SIRT algorithm with non-negativity constraints unless indicated otherwise. The number of iterations is chosen to be large enough to assure convergence of HEBT as well as not too large to avoid over-fitting the least square problem. The weighting factor α was chosen as 0.7 unless indicated, which we found to be a good value in our experiments that balances the influence of the EDS-STEM and HAADF-STEM data. The response ratio factors used in HEBT were estimated by fitting the linear models of Eq. 2.6 using the non-zero pixels in the tomographic input data using the NNLS (Non-negative least squares) algorithm [LH74, Chapter 23].

We can assess the image quality of reconstructions with reference images in the first two cases. For the simulation, we can compare reconstructions with the phantom images; for the non-alloy Au-Ag nanoparticle, we use the segmentations acquired from the Z-contrast reconstructions as the ground-truth references. Here we use three types of image quality metrics. (i) Structural similarity index (SSIM, [Wan+04]) computes structural similarity between images, which aligns with image quality perceived by human eyes. Since image intensities are different for HAADF-STEM and EDS-STEM, we exclude the luminance and contrast terms for SSIM, and only compute the structure term. (ii) Mean-squared error (MSE) simply computes the difference between reference images ($\mathbf{x}$) and reconstructions ($\mathbf{y}$) which were scaled by scaling factors that give minimal MSE. The computation is formulated as: $\mathrm{MSE}(\mathbf{x}, \mathbf{y}) = \min_c \parallel \mathbf{x} - c\mathbf{y} \parallel_2^2$, where c is the scaling factor. (iii) The difference in pixels (DP) is computed as the l_1 norm of the difference between

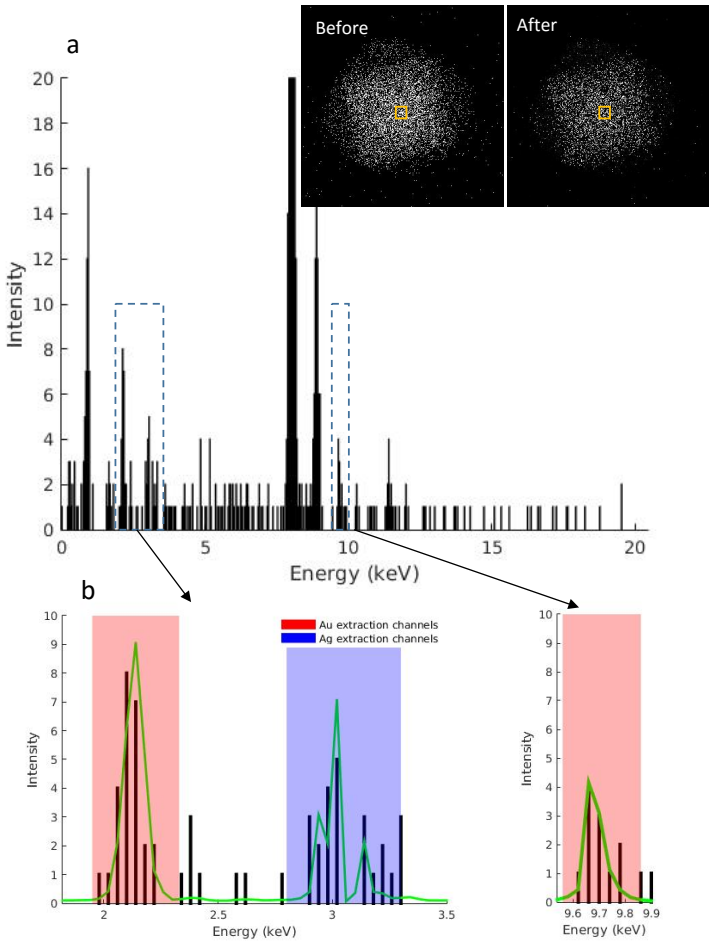


Figure 2.2: (a) Spectrum of the Au-Ag nanoparticle before PCA denoising. On the top-right corner show the Au elemental maps before and after PCA denoising. The yellow boxes indicates where the intensities of the spectrum were extracted. (b) Zoom-in to the spectrum (black) and the denoised spectrum (green). The colored regions indicated at which channels the denoised spectrum images were extracted to elemental maps.

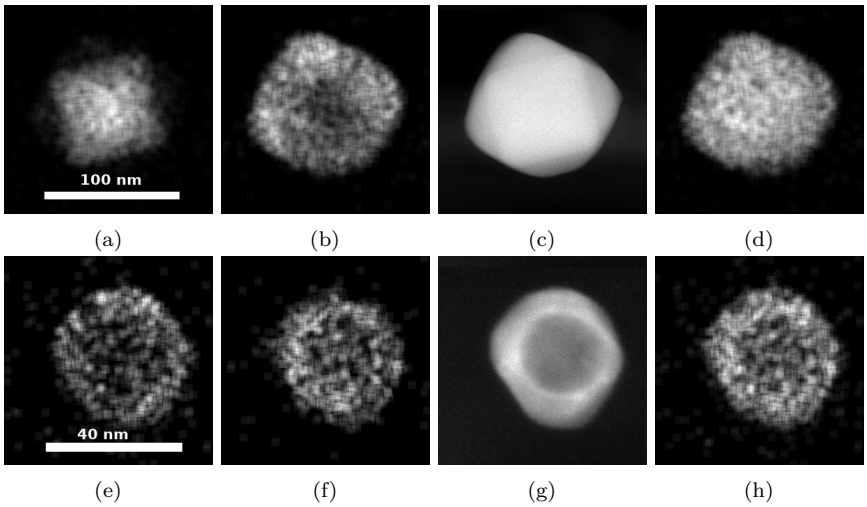


Figure 2.3: The upper and bottom rows correspond to the Au-Ag nanoparticle at tilt angle -75° and the alloyed Au-Ag nanoparticle at tilt angle 30° respectively. Figures (a) and (e) are Au elemental maps. Figures (b) and (f) are Ag elemental maps. Figures (c) and (g) are Z-contrast projection images. Figures (d) and (f) are the sum of elemental maps weighted by the estimated response ratio factors . The image sizes are 300×300 pixels.

two binary images. The reference images are already binary, while reconstructions for elements have continuous intensity. Given the knowledge that elements have homogeneous density, we binarize the reconstructions with thresholds, which are chosen as the ones giving minimal DP. Mathematically this can be written as $\mathrm{DP}(\mathbf{x}, \mathbf{y}) = \min_b \parallel \mathbf{x} - B_b(\mathbf{y}) \parallel$, where $B_b(\mathbf{y})$ means binarizing an image with the threshold b.

2.4.1 Phantom objects

Estimating the response ratio factors is the first step of HEBT. The response ratio factors for Au and Ag were estimated to be $\mathbf{r}_{est} = [3.27 \times 10^4, 1.68 \times 10^4]$, while the ground truths are $\mathbf{r}_{gt} = [3.66 \times 10^4, 1.43 \times 10^4]$ based on the given response factors. The goodness of how the data matches the linear model is indicated by the coefficient of determination $R^2 = 0.91$, which can be interpreted as 91% of the data can be explained by the linear model.

The simulation study aims at studying the stability of HEBT when errors are present in the estimated response ratio factors. Here, the estimated response ratio factors differ from the ground-truth by -10.66% and 17.48% respectively. Thus, reconstructions were made by HEBT with estimated and ground-truth response ratio factors respectively (Figure 2.4). Both were computed with weighting factor $\alpha = 0.7$ and for 200 iterations. First of all, intuitively we see both results show less noise and sharper contrast compared to EDS-STEM tomographic reconstructions.

Secondly, the deviation of response ratio factors only results in slightly different distributions of noise between elemental reconstructions. We can observe that "streaks" are more suppressed in Figure 2.4 (a) than (b) since r_{Au} is overestimated compared to the true values. On the other hand, Figure 2.4 (d) looks noisier than (e) because r_{Ag} is underestimated. From the image quality metrics, HEBT with $\mathbf{r}_{est}$ produces nearly the same image quality as HEBT with $\mathbf{r}_{gt}$. One noticeable result is that although HEBT with $\mathbf{r}_{gt}$ outperforms HEBT with $\mathbf{r}_{est}$ in terms of SSIM and DP, for MSE the result with $\mathbf{r}_{est}$ is better. This can be explained as that noise has been taken into account when estimating the response ratio factors, consequently yielding better statistical fitting for reconstructions. In conclusion, the HEBT algorithm output shows good stability *w.r.t.* errors in the estimation of the response ratios.

Table 2.1: Image quality metrics of reconstructions.

Methods	HEBT with $\mathbf{r}_{est}$	HEBT with $\mathbf{r}_{gt}$	EDS-STEM
SSIM Au	0.9923	0.9917	0.9437
SSIM Ag	0.9837	0.9835	0.8739
MSE Au	0.0260	0.0260	0.0801
MSE Ag	0.0449	0.0469	0.0740
DP Au	1664	1940	6936
DP Ag	4915	4762	7318

2.4.2 Non-alloy Au-Ag nanoparticle

The first sample that is experimentally investigated is an Ag nanoparticle with a diameter of approximately 110 nm with an embedded Au octahedron. Examples of Z-contrast images and elemental maps are given in Figure 2.3, indicating that Ag and Au are well separated.

The response ratio factors $r^{(Au)}$ and $r^{(Ag)}$ were estimated from all the non-zero pixels using the NNLS algorithm. The fitting results are $\mathbf{r} = [5.31 \times 10^4, 8.64 \times 10^4]$ with a coefficient of determination $R^2 = 0.95$. The example of Figure 2.3 (d) shows that the sum of elemental maps weighted by $\mathbf{r}$ closely but not perfectly matches the HAADF-STEM projection image due to noise. After the estimation, the reconstructions were computed slice by slice in a volume of $300 \times 300 \times 300$ voxels by solving the least square problem of Eq. 2.9.

Figure 2.5 shows the 2D reconstruction images at different slices. Compared to EDS-STEM reconstructions, HEBT reconstructions demonstrate smoother intensity distributions, suppressed noise levels and clearer boundaries. Especially for the Ag reconstructions, morphological analysis becomes easier as exterior boundaries show a sharper contrast to the background after being regularized by the HAADF-STEM term. The HAADF term also regularizes intensities of noise to

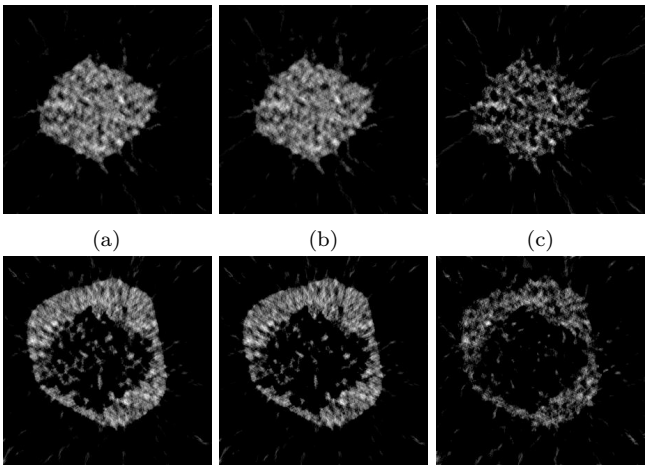
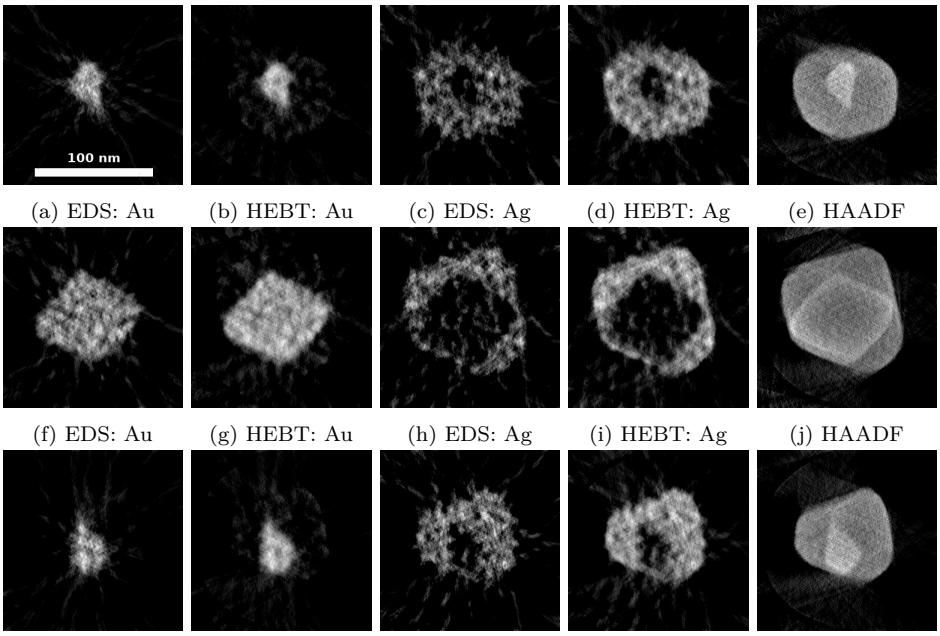


Figure 2.4: Figures (a), (b) and (c) are reconstructions of Au distribution. Figures (d), (e) and (f) are reconstructions of Ag. The left and middle columns are HEBT reconstructions respectively with estimated response ratio factors $\mathbf{r}_{est}$ and with ground-truth response ratio factors $\mathbf{r}_{gt}$; the right column is reconstructions from only EDS-STEM elemental maps.

concentrate within the particle and not spread across the background. For example, in the HEBT reconstructions for Au, we can see noise forming a 'shadow' of the entire particle on the background. Fortunately, the 'shadow' noise is rather weak and can be removed by thresholding or smoothing.

The HAADF-STEM reconstructions have clear boundaries between Au and Ag in this case. Therefore, we can easily segment the two particles, and use the segmentation as the ground truth for reconstruction quality assessment. In Figure 2.6, the HAADF-STEM reconstruction was segmented into two parts by manually recognizing the boundaries in every slice using the FEI Amira 6.0 software, which are considered as the ground truths of compositional distributions. Meanwhile, we also demonstrate the 3D volume rendering of EDS-STEM reconstructions and HAADF reconstructions for comparison. The image quality metrics were computed in 3D and listed in Table 2.2. The metrics show that the image quality of HEBT reconstructions is intrinsically enhanced in comparison to conventional EDS reconstructions.

Based on the ground truth from the segmented HAADF-STEM reconstructions, the influence of two parameters for HEBT can be investigated: the weighting factor α and the number of iterations. Here, we sampled the weighting factor from 0.01 to 0.99 for HEBT reconstructions with different numbers of iterations for one slice. Figure 2.7 plots the MSE indices at each weighting factor. It first indicates a decrease of MSE as α grows, as the noise is increasingly suppressed by the HAADF-STEM term. When α gets close to 1, MSE starts to increase rapidly after reaching a minimum. To understand this phenomenon, we plot the



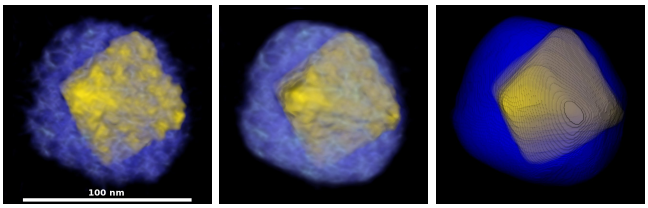
(a) EDS: Au (b) HEBT: Au (c) EDS: Ag (d) HEBT: Ag (e) HAADF

(f) EDS: Au (g) HEBT: Au (h) EDS: Ag (i) HEBT: Ag (j) HAADF

(k) EDS: Au (l) HEBT: Au (m) EDS: Ag (n) HEBT: Ag (o) HAADF

Figure 2.5: 2D reconstruction images for the non-alloyed nanoparticle at slice number 80, 150 and 220 corresponding to the up, middle and bottom rows respectively. The left two columns are the distributions of Au reconstructed by conventional EDS tomography and by HEBT respectively. The middle two columns are the reconstructions of Ag. The right column shows the reconstructions from HAADF Z-contrast projection images.

reconstructions at $\alpha = 0.7$ for 50/100/500 iterations. It shows that for 50 iterations, Ag appears in the reconstruction of Au (Figure 2.8 (b)). The explanation is that a too large α makes minimizing residuals for EDS-STEM terms become very inefficient due to their small weights. If the residuals of EDS-STEM terms remain large while the residual of HAADF-STEM has already been minimized, backprojection from HAADF-STEM projection images will show up in the reconstructions. In such a case, we can see appearance from the other compositions. The appearance can be reduced by increasing the number of iterations. In this case, Ag disappears in the Au reconstruction as the number of iterations grows. On the other hand, if the number of iterations is chosen very large, this may lead to over-fitting of the least-square problem, which results in the presence of noise in reconstructions. The over-fitting also explains why – for small weighting factors – the MSE metric decreases as the number of iterations increases (see Figure 2.7). In a word, the weighting factor influences the noise suppression and the convergence of least square problem; to guarantee convergence for large weighting factors, a large number of iterations should be adopted. From Figure 2.7, we conclude that $\alpha = 0.7$ and 100 iterations are close-to-optimal settings for this example.



(a) EDS (b) HEBT (c) HAADF

Figure 2.6: 3D volume rendering of Au (yellow, interior) and Ag (blue, exterior) distributions in the non-alloyed nanoparticle reconstructed using (a) EDS-STEM tomography and (b) HEBT. The 3D volume rendering of Au and Ag segmented from HAADF-STEM reconstructions (ground-truth) is shown in (c).

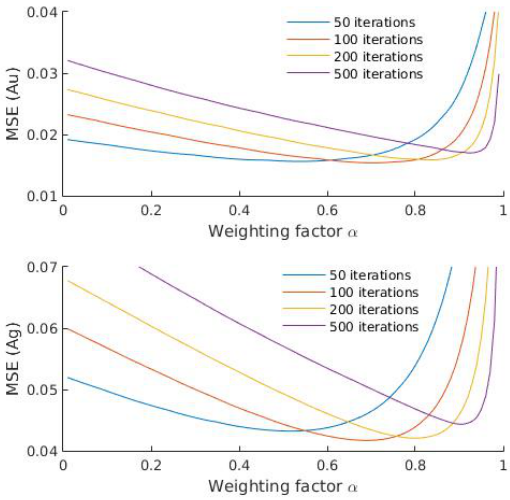


Figure 2.7: Mean-squared errors for Au and Ag reconstructions under different iterations. The reconstructions are sampled for weighting factors from 0.01 to 0.99.

To investigate whether HEBT leads to improved ability to spatially resolve the chemical composition of nanomaterials in comparison to separate EDS-STEM reconstructions, we have conducted two additional validation experiments. In the first experiment, a binary mask is created from the HAADF-STEM reconstruction, which is then enforced during each iteration step of the SIRT reconstruction from elemental maps. For the second experiments, a binary mask is created based on the Z-contrast projection images which are subsequently applied to the elemental maps prior to tomographic reconstruction. The results of these experiments are shown in Figures 2.9 (a) to (d). It can be seen that the results are qualitatively similar to the reconstructions without the masks in the sense that the chemical composition is no better spatially localized than in the unmasked case. This can be contrasted

Table 2.2: Image quality metrics.

Methods	EDS-STEM	HEBT
SSIM Au	0.9661	0.9680
SSIM Ag	0.9024	0.9097
MSE Au	0.0093	0.0069
MSE Ag	0.0368	0.0229
DP Au	233805	134861
DP Ag	1213822	674403

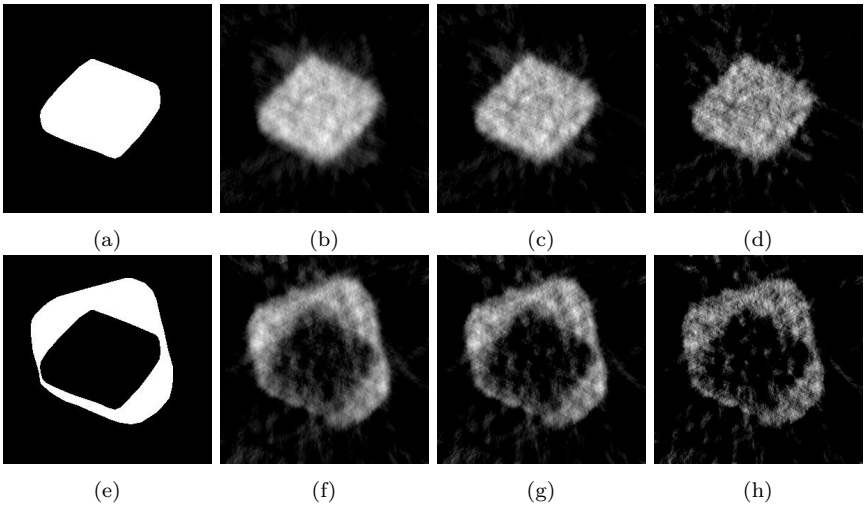


Figure 2.8: 2D reconstruction images at slice number 150. The upper row is for Au, and the bottom row is for Ag. (a) and (e) are the reference images for computing SSIM, which are segmented from the HAADF-STEM reconstruction. (b)/(f), (c)/(g) and (d)/(h) are respectively HEBT reconstructions with weighting factor $\alpha = 0.7$ under 50, 100, 500 iterations.

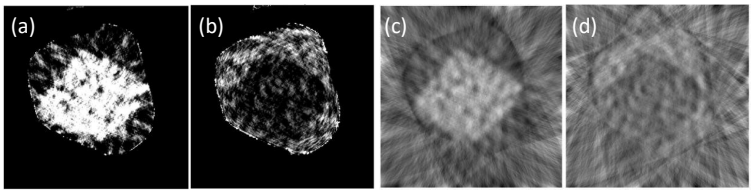


Figure 2.9: Figures (a) and (b) are distributions of Au and Ag in the non-alloy nanoparticle reconstructed from elemental maps. During the reconstruction process, reconstruction volumes were masked by the binarized HAADF-STEM reconstruction. Figures (c) and (d) are distributions of Au and Ag in the non-alloy nanoparticle reconstructed from elemental maps that have been masked by binarized Z-contrast images.

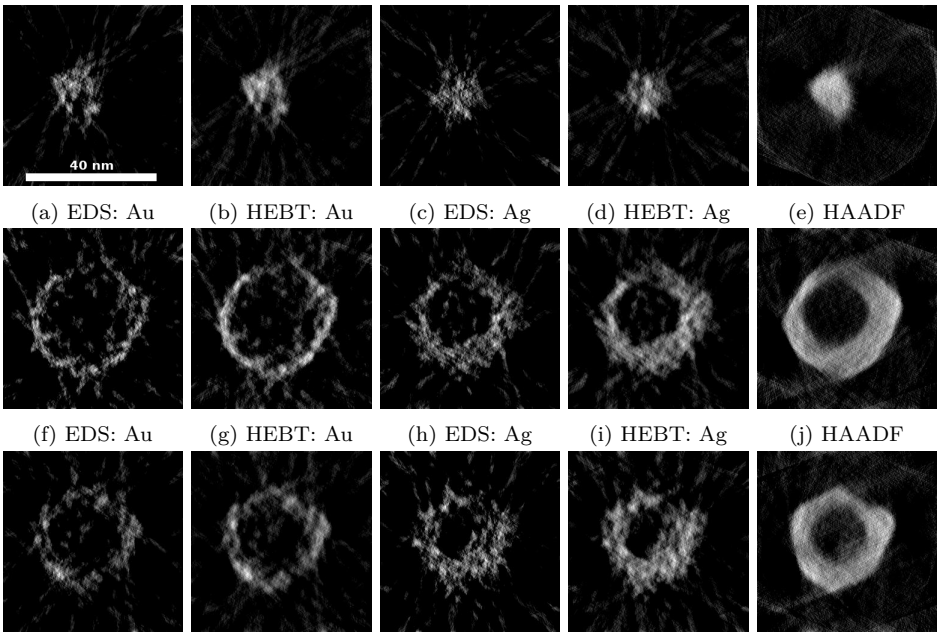
to Figure 2.5, where the better localization is visible, clearly demonstrating the advantage of our HEBT reconstruction technique.

2.4.3 Alloyed Au-Ag nanoparticle

In this case, we demonstrate the application of HEBT on data for which 3D compositional analysis is difficult for both EDS-STEM tomography and HAADF-STEM tomography. The sample is an Au-Ag alloy nanoparticle with a diameter about 30 nm. As suggested by the Z-contrast images in Figure 2.3 (d), segmentation cannot be made based on HAADF-STEM reconstructions since no clear boundary exists between the two compositions. Although elemental distributions can be reconstructed from elemental maps, the elemental maps are very noisy (Figure 2.3 (e) and (f)) and lead to strong noise in the EDS-STEM tomographic reconstruction results.

The HEBT reconstructions were computed using $\alpha = 0.7$ for 200 iterations. The response ratio factors were estimated to be $\mathbf{r} = [5.63 \times 10^4, 6.52 \times 10^4]$ with a coefficient of determination $R^2 = 0.79$. The values for the same elements differ from the first experimental case. This is likely due to an intensity rescaling that was applied when storing the HAADF-STEM data. As our response ratio factors are automatically scaled, this does not affect the final results.

Compared with EDS-STEM tomography, HEBT gives more interpretable results with less noise and stronger contrast to the background as shown in the 2D slices of Figure 2.10. Here, since we no longer have ground-truth images, we cannot compute image quality metrics. Figure 2.11 shows that the HEBT reconstructions provide more information in 3D on the concentration of the different elements compared to the EDS-STEM reconstructions. The elemental distributions with reduced noise indicate that the Au is more concentrated in the exterior than Ag.

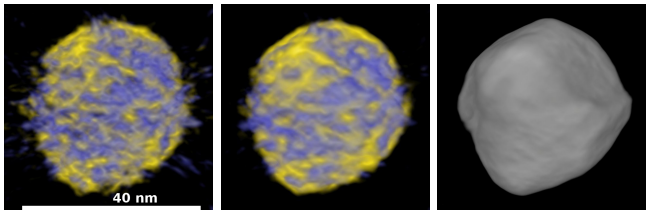


(a) EDS: Au (b) HEBT: Au (c) EDS: Ag (d) HEBT: Ag (e) HAADF

(f) EDS: Au (g) HEBT: Au (h) EDS: Ag (i) HEBT: Ag (j) HAADF

(k) EDS: Au (l) HEBT: Au (m) EDS: Ag (n) HEBT: Ag (o) HAADF

Figure 2.10: 2D reconstruction images for the alloyed nanoparticle at slice number 80, 150 and 220 corresponding to the up, middle and bottom rows respectively. The left two columns are the distributions of Au reconstructed by conventional EDS tomography and by HEBT respectively. The middle two columns are the distributions of Ag. The right column shows the reconstructions from Z-contrast projection images.



(a) EDS (b) HEBT (c) HAADF

Figure 2.11: 3D volume rendering of Au (yellow) and Ag (blue) distributions in the alloyed nanoparticle reconstructed using (a) EDS-STEM tomography and (b) HEBT. The 3D volume rendering of the HAADF-STEM reconstruction is shown in (c).

2.5 Conclusion

In this study, we have developed *HAADF-EDS bimodal tomography* for the 3D characterization of the chemical composition at the nanometer scale. This technique first links elemental maps with Z-contrast images that are recorded simultaneously in STEM mode and contain complementary information. The linking is made by estimating response ratio factors that give the linear relation of two types of images and by scaling their intensities to the same unit. Simultaneously from two types of projection images, 3D elemental distributions are reconstructed. The reconstruction process results in a simultaneous minimization of the projection errors of both EDS-STEM and HAADF-STEM and is carried out using an iterative method such as SIRT.

HEBT has first been tested on a phantom object that is based on hetero-nanoparticles. We specifically demonstrated that HEBT is robust *w.r.t.* errors in the response ratio factor estimation. Subsequently, we used HEBT to reconstruct the 3D elemental distributions of two different nanoparticles. To investigate the image quality enhancement of HEBT, we first reconstructed an Au-Ag nanoparticle where the different elements could be distinguished based on Z-contrast. Taking the Z-contrast reconstruction results as the ground truth, we see that reconstructions computed by HEBT are improved in comparison to EDS-STEM tomographic reconstructions in terms of image quality. In this case, we also demonstrated that HEBT with a large weighting factor requires a large number of iterations to converge and separate between elements. In the second experimental case, Ag and Au are alloyed, and thus it is impossible to investigate the 3D distributions of the chemical elements based on HAADF-STEM tomography. Using the HEBT techniques, we are able to investigate the spatial distribution of Ag and Au inside the particle. The interpretation of the final result is more straightforward in comparison to conventional EDS-STEM tomography, for which the results contain more noise.

The HEBT algorithm is based on the assumption that both HAADF-STEM projection images and EDS-STEM elemental maps can be modeled as perfect linear projections of the structure. In practice, this assumption is not completely valid as nonlinear phenomena such as X-ray absorption and electron channelling may break the projection requirement [Sla+16b; Bur+16]. In addition, the EDS noise follows a Poisson distribution, while the least squares problem in Eq. 2.9 is based on the assumption that the noise follows a Gaussian distribution. Our purpose here is to demonstrate the feasibility of HEBT, while recognizing these sources of inaccuracy. In future work, we plan to incorporate more sophisticated models for self-absorption (similar to [Bur+16]) and elemental map extraction (similar to [Luc+13]), as well as to adopt advanced denoising reconstruction algorithm based on the Poisson noise model such as the EM method.

Conventionally, quantitative analysis based on EDS-STEM measurements suffers from the high noise level in these measurements. By combining EDS and HAADF, especially by imposing the 3D information obtained by HAADF, the

improvement in the reconstructions (compared to pure EDS-STEM reconstruction) will lead to more reliable quantification, provided that the corresponding zeta factors are known. This application also requires an accurate estimation of the response ratio factors. Therefore, we are developing a new estimation method which is based on the Poisson noise model rather than the Gaussian noise model.

In conclusion, the newly developed HEBT technique is a promising technique to analyze chemical compositions of nanomaterials in 3D. By exploiting more complete information from two complementary types of images, it can characterize the elemental distribution even when it is not straightforward using HAADF-STEM tomography or EDS-STEM tomography. This advantage means that the 3D characterization of chemical composition can be pushed to materials with smaller dimensions and more complex compositions.

Acknowledgment

The author thank Dr. Bart Goris and Prof. dr. Sara Bals (EMAT, University of Antwerp, Antwerp, Belgium) for providing the experimental data in this chapter, as well as Prof. dr. Luis M. Liz-Marzan (Bionanoplasmonics Laboratory, CIC biomaGUNE, San Sebastián, Spain) for providing the nanoparticle samples.

3

Joined HAADF-EDS reconstructions regularized by total nuclear variation

3.1 Introduction

X-ray energy-dispersive spectroscopy (EDS) tomography is an electron tomography (ET) technique for 3D compositional characterization. It refers to making a tomographic reconstruction for the distribution of a specific chemical element from a tilt series of images called elemental maps. It is based on the linearity assumption that the image intensities, which correspond to X-ray counts, are proportional to the mass-thickness of the chemical element [WW06; Lep+13]. Unfortunately, the number of X-ray counts is often low due to small emission probabilities and small detection angles. Consequently, the signal-to-noise ratios (SNRs) are low and the number of tilts is limited, which leads to poorly reconstructed images. To obtain sufficient X-ray counts, a high electron dose is often applied by setting a large beam current or a long acquisition time. However, this is then limited by how much dose the sample can survive.

This chapter is based on:
Z. Zhong, W. J. Palenstijn, J. Adler, and K. J. Batenburg. "EDS tomographic reconstruction regularized by total nuclear variation joined with HAADF-STEM tomography". *Ultramicroscopy* 191 (2018), pp. 34–43.

Using advanced reconstruction algorithms, it is possible to make an accurate reconstruction from a small number of tilts or from data with low SNRs. Total variation (TV) regularization algorithms are a widely-used method for reconstructing from a small number of tilts, which find a solution with sparse gradients to the ill-posed inverse problem [BO13; SKP06; LSP08]. In ET, it is adopted for reducing the missing wedge artifacts [Gor+12]. In EDS tomography with low-count data, it is used for the effect of suppressing noise and encouraging piecewise constant structures [Bur+16; Zan+16a]. However, additional difficulties arise when TV regularization is applied to EDS tomography. The noise aggravates a significant issue of TV – the staircase effect that produces small flat regions separated by edges. In addition, as the TV regularization tends to preserve sharp discontinuities, the noisy edges become sharp and saw-like. When the number of tilts is small or the X-ray counts are low, these effects become even more severe as a result.

In this chapter, while TV regularization encourages the information of sharp edges in the reconstructed image, we instead use total nuclear variation (TNV) regularization which also encourages common edges in multiple images [RL15; Hol14; Dur+16]. Using TNV regularization, it is possible to augment the available data with extra information from another imaging modality. The other modality being exploited is high-angle annular dark-field (HAADF) STEM. Its image contrast depends on the atomic numbers of probed atoms. Therefore, the tomographic reconstruction based on HAADF-STEM shows a (weighted) sum of distributions for all chemical elements [MW03; Mid+01], which sometimes also contains sharp edges showing the variance of distributions. More importantly, due to a strongly reduced time constraint, more HAADF-STEM tilt images can be measured with higher SNRs in a relatively short time. Thus the HAADF-STEM reconstruction is usually more accurate and less noisy.

Our proposed method performs the EDS reconstruction together with a HAADF-STEM reconstruction with joint TNV regularization, from a tilt series of EDS maps and a tilt series of HAADF-STEM images. As an extension of TV imposed on multiple images, the TNV regularization also promotes the sparsity of gradients for each image. Hence, it has similar effects in terms of suppressing noise and preserving sharp discontinuities. Additionally, TNV regularization promotes joint reconstructions that have common edge locations and gradients in the parallel/antiparallel directions. The TNV regularization can penalize the staircase effects and saw-like edges in the EDS reconstruction for not having coinciding edges in the HAADF-STEM reconstruction image.

Note that the staircase effects can also be reduced using total generalized variation (TGV) regularization which incorporates smoothness information of the reconstructed image. However, the TGV regularization incorporates no additional information from extra data. Thus, saw-like edges may be still present in the TGV regularized reconstruction. It is also noteworthy that the proposed method can be seen as a bimodal tomography approach. In Chapter 2 and our paper [Zho+17], we have proposed another bimodal method which is named HAADF-EDS bimodal tomography (HEBT). That method incorporates a different prior, exploiting that the HAADF data should be a linear combination of the EDS data for all chemical

elements in the specimen. Therefore, HEBT can only be used if EDS data for all chemical elements in this sample have been acquired. In contrast, the TNV-regularized method only uses the data for a single chemical element in addition to the HAADF-STEM data.

The rest of this chapter consists of the following sections. Section 3.2 illustrates the notations for EDS tomography, HAADF-STEM tomography, TNV regularization as well as the TNV-regularized joint reconstruction method. Section 3.3 demonstrates simulation studies and an experimental study. Lastly, in Section 3.4 we discuss the experimental results and draw a conclusion.

3.2 Method

Consider a specimen located in a 3D volume space discretized into N voxels. The reconstruction unknowns for a single given chemical element in the specimen are expressed as a vector $\mathbf{x}^e \in \mathbb{R}^N$. The reconstruction $\mathbf{x}^e$ is referred to as the *EDS reconstruction*. The input data are a tilt series of elemental maps, which are expressed as a vector $\mathbf{p}^e \in \mathbb{R}^{M^e}$, where M^e denotes the total number of pixels. The image intensities correspond to the characteristic X-ray counts, which are proportional to the linear projection of the corresponding chemical element probed by the focused beam under thin film approximation. This linear relationship is modeled by the following system of equations:

$$p_i^e = \sum_{j=1}^{N} w_{ij}^e x_j^e, \tag{3.1}$$

for $i = 1, \cdots, M^e$ and $j = 1, \cdots, N$. Each pixel position i corresponds to a ray determined by the beam position and the tilt angle of the specimen. The weight factor w_{ij}^e is determined by the area of the jth voxel intersected by the ith ray integral. The matrix $\mathbf{W}^e \in \mathbb{R}^{M^e \times N}$ is referred to as the EDS projection matrix.

The volume space for the HAADF-STEM reconstruction is defined as the same for joining the reconstruction with the EDS reconstruction. Similarly, the reconstruction unknowns for the same sample are expressed as a vector $\mathbf{x}^h \in \mathbb{R}^N$, with h denoting HAADF-STEM. The tilt series of projection images are expressed as a vector $\mathbf{p}^h \in \mathbb{R}^{M^h}$, where M^h is the total number of pixels for HAADF-STEM acquisition. Note that M^h may be different from the number of pixels for EDS acquisition (M^e). In particular, this means it is possible to record HAADF-STEM data for more tilts than for which EDS data are acquired. Following the linear integral model, the linear relationship between the tilt series $\mathbf{p}^h$ and reconstruction unknowns $\mathbf{x}^h$ is:

$$p_k^h = \sum_{j=1}^{N} w_{kj}^h x_j^h, \tag{3.2}$$

where $k = 1, \cdots, M^h$ is the pixel index, w_{kj}^h is the entry of the HAADF-STEM

projection matrix $\mathbf{W}^h \in \mathbb{R}^{M^h \times N}$ determined by the area of voxel j intersected by the ray integral k.

Reconstruction algorithms can be divided into two categories: analytical algorithms (e.g. filtered backprojection (FBP) [KS88]) and iterative algorithms. Here we focus on iterative algorithms for their capability of implementing regularization. Classically, the iterative algorithm minimizes a data cost function, based on the above linear systems, to find solutions for the inverse problems

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}} \parallel \mathbf{p}^e - \mathbf{W}^e\mathbf{x}^e \parallel^2, \tag{3.3}$$

$$\mathbf{x}^{h*} = \underset{\mathbf{x}^h}{\operatorname{argmin}} \parallel \mathbf{p}^h - \mathbf{W}^h\mathbf{x}^h \parallel^2, \tag{3.4}$$

for EDS and HAADF-STEM tomography respectively.

To regularize the reconstruction, a regularization term is added to the cost function, resulting in a new minimization problem. Taking TV-regularized EDS tomography as an example, the reconstruction is computed by solving the minimization problem of:

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}} \parallel \mathbf{p}^e - \mathbf{W}^e\mathbf{x}^e \parallel^2 + \lambda \mathrm{TV}(\mathbf{x}^e), \tag{3.5}$$

where the term $\mathrm{TV}(\mathbf{x}^e)$ gives the total variation of the reconstruction image, and λ is the factor determining the strength of the TV regularization.

The TV term for an arbitrary 3D gray-scale image $\mathbf{u} \in \mathbb{R}^N$ is defined as:

$$\mathrm{TV}(\mathbf{u}) = \sum_j^N \parallel \bigtriangledown u_j \parallel, \tag{3.6}$$

where

$$\bigtriangledown u_j = \begin{pmatrix} \bigtriangledown^x u_j \\ \bigtriangledown^y u_j \\ \bigtriangledown^z u_j \end{pmatrix} \tag{3.7}$$

is a discrete approximation of the gradient for the jth voxel. The operators $\bigtriangledown^x$, $\bigtriangledown^y$ and $\bigtriangledown^z$ approximate gradients in the X, Y and Z directions respectively by taking the forward difference between voxels. The norm $\parallel \cdot \parallel$ is usually chosen as l_1-norm or l_2-norm. In this chapter, we use the l_2-norm, for which the TV regularization is also called isotropic TV. Using l_1-norm TV tends to encourage horizontal and vertical edges, which is a drawback. In contrast, l_2-norm TV is rotationally invariant and thus is preferable in the application of EDS tomography [BO13; RL15].

Next, we describe the notation for TNV regularization. TNV is usually imposed on images with multiple channels such as RGB images. In this chapter, we join an EDS reconstruction image and a HAADF-STEM reconstruction as the two channels for one image. Consider an arbitrary L-channel image $\mathbf{u}$. The image

intensity on the jth pixel can be expressed as $\mathbf{u}_j = (u_j^{(1)}, \dots u_j^{(2)} \dots, u_j^{(L)})^T$. The TNV of $\mathbf{u}$ is then defined as the nuclear norm of the Jacobian matrix:

$$\mathrm{TNV}(\mathbf{u}) = \sum_j \| (\mathbf{J}(\mathbf{u}))_j \|_\star, \tag{3.8}$$

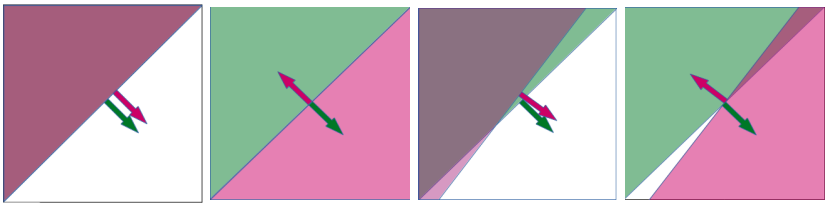
where the $(\mathbf{J}(\mathbf{u}))_j$ is the Jacobian matrix defined by:

$$(\mathbf{J}(\mathbf{u}))_j = \begin{pmatrix} \bigtriangledown^x u_j^{(1)}, \bigtriangledown^y u_j^{(1)}, \bigtriangledown^z u_j^{(1)} \\ \vdots \\ \bigtriangledown^x u_j^{(L)}, \bigtriangledown^y u_j^{(L)}, \bigtriangledown^z u_j^{(L)} \end{pmatrix}, \tag{3.9}$$

and the nuclear norm $\| \cdot \|_\star$ is defined as the l_1-norm of the Jacobian matrix' singular values. Note that the l_1-norm here does not encourage horizontal and vertical edges. For a one-channel image, the TNV is reduced to the isotropic TV.

TNV regularization introduces the following effects by encouraging the rank-sparsity in the Jacobian matrix. First of all, TNV regularization leads to the similar effects as TV regularization in terms of promoting the sparsity of image gradients, preserving sharp discontinuities and suppressing noise. The TNV is also rotationally invariant like the isotropic TV. Second, the TNV regularization gives preference to the images that have common edge locations and parallel or antiparallel gradient vectors, while it does not introduce false features between channels [RL15].

Figure 3.1 illustrates examples of TNV computed for two-channel images with gradients pointing in parallel, antiparallel, almost parallel and almost antiparallel directions. For the parallel and antiparallel examples, one of the two singular values of the Jacobian matrix will be zero, while for the other two examples, the two singular values will be non-zero and unique. Suppose the norms of all gradients are equal to 1, the TNV values for the pixels where two edges cross for figures (a), (b), (c) and (d) are 2, 2, 2.14 and 2.14 respectively. Therefore, minimizing TNV gives preference to the parallel or antiparallel gradients.



(a) Parallel gradients (b) Antiparallel gradients (c) Almost parallel gradients (d) Almost antiparallel gradients

Figure 3.1: Illustration of TNV for two-channel images and gradients.

In our application of jointly regularizing the EDS and HAADF-STEM reconstructions, the data fit terms for EDS and HAADF-STEM are minimized together with a TNV term:

$$\mathbf{x}^{e*}, \mathbf{x}^{h*} = \underset{\mathbf{x}^e, \mathbf{x}^h}{\operatorname{argmin}} \parallel \mathbf{p}^e - \mathbf{W}^e \mathbf{x}^e \parallel^2 + \parallel \mathbf{p}^h - \mathbf{W}^h \mathbf{x}^h \parallel^2 + \lambda \mathrm{TNV}(\mathbf{x}^e, \mathbf{x}^h) \qquad (3.10)$$

where $\mathrm{TNV}(\mathbf{x}^e, \mathbf{x}^h)$ represents the TNV for a two-channel image consisting of $\mathbf{x}^e$ and $\mathbf{x}^h$, which means $L = 2$ and $\mathbf{u}_j = (x_j^e, x_j^h)^T$ for Eq. 3.9. λ is the regularization parameter determining the strength of TNV regularization. In practice, the value of λ should be chosen carefully to obtain a desired reconstruction result. A too large value may lead to an over-regularized image with blurred edges, while a too small value may lead to insufficient regularization effects.

In practice, we have to consider the magnitude of these two types of input data. The magnitudes of EDS elemental maps and HAADF-STEM images can be tremendously different. Intensities of the EDS elemental map, which correspond to the X-ray counts, are usually of the order of magnitude 1 or 2, while intensities of the HAADF-STEM image usually have different magnitudes. This difference of magnitude may cause a biased TNV term. To avoid this, we scale the image intensities of both types of images to a range from 0 to 1 before the reconstruction step. Afterwards the reconstructed image can be re-scaled by the same value so that quantitative characterization is still feasible.

3.3 Experiments

In this section, we first investigate the proposed method based on a phantom simulation dataset and a multi-slice simulation dataset, for which ground-truth images are available. After that, we apply the method to a real experimental dataset.

For the simulation datasets, noiseless EDS maps are first computed. Based on the assumption that the noise is Poisson-distributed with expected values given by the noiseless maps, we generate maps with Poisson noise similarly to the noise generation in [MHD18] as follows: for each pixel on each map, a random number is generated for a Poisson distribution taking the corresponding noiseless image intensity as the expected value. This is then taken as the noise-corrupted image intensity. For a Poisson distribution the SNR is given by the square root of the expected value. The magnitude (and therefore the noise level) of the EDS maps could be calculated from assumed values of probe current, dwell time, fluorescence yield, solid angle and detection efficiency [MHD18; Che+16]. In this chapter, for simplicity, we set the magnitudes to levels similar to real experimental data (up to $\sim$100 counts).

For comparison, reconstructions are also made for a commonly used non-regularized reconstruction method – the simultaneous iterative reconstruction technique (SIRT) [GB08], and the TV regularization method defined by Eq. 3.5. We use the SIRT implementation in the ASTRA Toolbox [PBS13], and use the Douglas-Rachford

primal-dual splitting algorithm [BH13] implemented in the Operator Discretization Library (ODL) [AKÖ17] in Python to compute the regularized reconstructions. SIRT is computed for 50 iterations to avoid over-fitting to noise, while the regularized reconstruction is computed until convergence. We only perform 2D reconstructions for slices, but the results can be generalized to the third dimension for the same effects. For conciseness, we refer to the EDS reconstructed image jointly made using the TNV regularization as the TNV reconstruction, the reconstructed image regularized by TV as the TV reconstruction and the reconstructed image using SIRT as the SIRT reconstruction.

We use the linear correlation coefficient to measure image quality, which determines the extent to which the reconstructed image $\mathbf{u}$ is linearly related to the ground-truth image $\mathbf{g}$. The correlation coefficient r is calculated as:

$$r = \frac{\sum_i (u_i - \bar{u})(g_i - \bar{g})}{\sqrt{\sum_i (u_i - \bar{u})^2 \sum_i (g_i - \bar{g})^2}}, \tag{3.11}$$

where $\bar{u}$ and $\bar{g}$ are the mean values of $\mathbf{u}$ and $\mathbf{g}$ respectively.

We also compute segmentation errors which can indicate the accuracy of simple quantitative characterization. The segmentation error is calculated for a binary image segmented from a reconstructed image. It is defined as the proportion of incorrectly segmented pixels to the total number of non-zero pixels in the ground-truth binary segmentation. For the binary image $\mathbf{s}$ and the ground-truth binary image $\mathbf{t}$, the segmentation error e is:

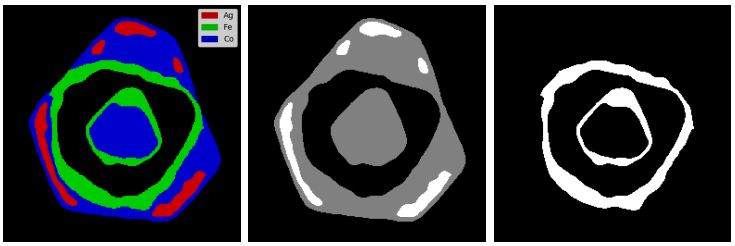
$$e = \frac{\sum_i |s_i - t_i|}{\sum_i t_i}, \tag{3.12}$$

for $s_i \in \{0, 1\}$ and $t_i \in \{0, 1\}$. The reconstruction image is segmented by thresholding the image intensities. To find the optimal thresholds, we calculate the segmentation errors for a set of thresholds between the minimal and maximal image intensities of the reconstructed image. Then the one corresponding to the minimal segmentation error is chosen as the optimal threshold, and the corresponding segmentation is adopted as the optimal segmentation.

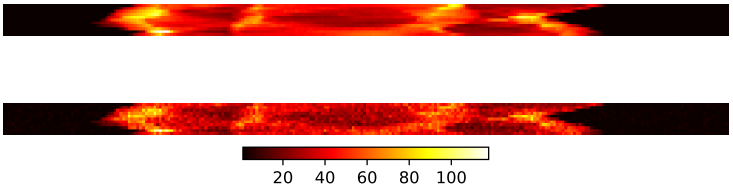
3.3.1 Phantom simulation

Data simulation and preparation

The 2D phantoms are created to simulate the single slice of a core-shell particle composed of three chemical elements: Ag, Fe and Co. Figure 3.2 (a) shows the distributions for these elements. Fe and Co have very similar atomic numbers and consequently show the same contrast in the HAADF-STEM reconstruction image. Au, which is embedded in the outer shell, has a relatively large atomic number and a strong contrast in the HAADF-STEM image. A purpose of this phantom study is to investigate whether the strong Ag features in the HAADF-STEM reconstruction will introduce false features to the EDS reconstruction.



(a) Chemical distributions (b) HAADF-STEM phantom (c) Fe phantom



(d) Top and bottom: noiseless and noise-corrupted EDS maps simulated for the Fe phantom

Figure 3.2: Illustration of the core-shell phantom: (a) Distribution of chemical elements Ag, Fe and Co; (b) HAADF-STEM phantom with the Z-contrast; (c) Fe phantom for simulating EDS maps. (d) simulated noiseless EDS maps (top) and noise-corrupted EDS maps (bottom) for Fe; each row and each column correspond to a tilt angle and a beam position respectively.

There are two phantoms in this example: a HAADF-STEM phantom and an EDS phantom for Fe, which are respectively shown in Figure 3.2 (b) and Figure 3.2 (c). We simulated the HAADF-STEM projection data using the ASTRA Toolbox for every 1° from 20° to 160° for a limited angular range to create a missing wedge. After that, we added Gaussian noise into the data. In addition, we used the ASTRA Toolbox to simulate projection data for the Fe phantom for a small number of tilts within the same angular range, which are for every 14° from 20° to 160° (Figure 3.2 (d)). The intensities of the noiseless maps were scaled and then used to generate the EDS maps corrupted by Poisson noise. In this case, the average image intensity on the non-background pixels is 43.1. In the last step, the maps were filtered using a Gaussian filter ($\sigma = 0.8$ pixel) as a denoising process.

Reconstruction results

First of all, we performed TNV-regularized reconstructions for a set of regularization parameter λ values to inspect the impact of this factor. We then plotted the correlation coefficients for these reconstructions in Figure 3.3. Similarly, we performed TV reconstructions for various λ values and a SIRT reconstruction. The correlation coefficients are also plotted in Figure 3.3. The TNV reconstructions corresponding to different λ values are shown in Figure 3.4. In addition, Figure 3.4 shows the HAADF-STEM reconstructions, which demonstrate little difference

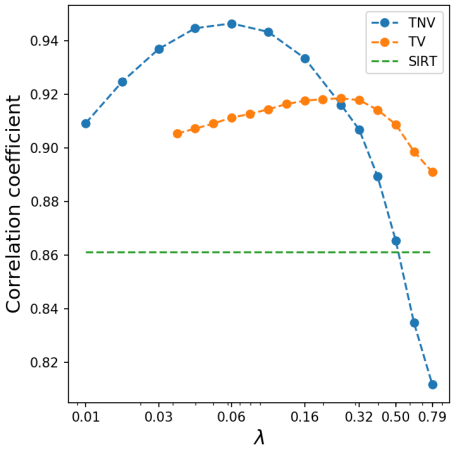
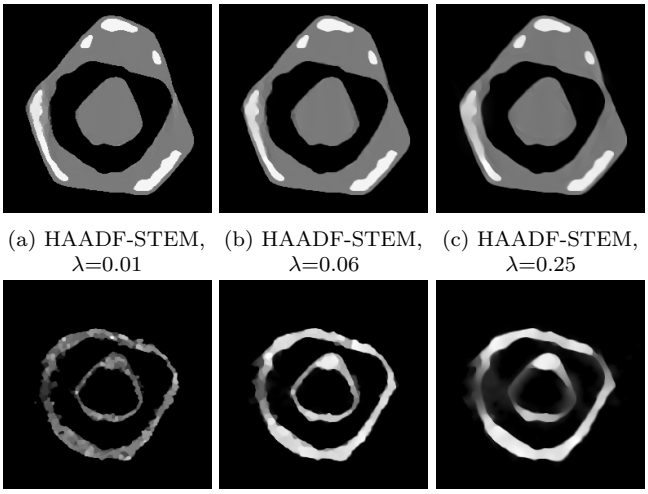


Figure 3.3: Correlation coefficients w.r.t. regularization parameter values (λ) in the core-shell phantom simulation study computed for different reconstruction methods. The SIRT reconstruction is performed for a fixed relaxation parameter set to 1 [GB08].

when different λ values are applied.



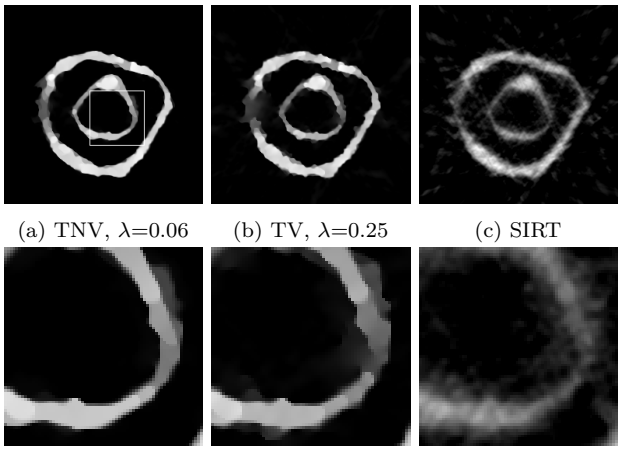
(a) HAADF-STEM, λ=0.01 (b) HAADF-STEM, λ=0.06 (c) HAADF-STEM, λ=0.25

(d) EDS, λ=0.01 (e) EDS, λ=0.06 (f) EDS, λ=0.25

Figure 3.4: TNV-regularized joint reconstructions for the core-shell nanoparticle phantom simulation: (a) - (c) HAADF-STEM reconstructions corresponding to different values of regularization parameter λ; (d) - (f) EDS reconstructions corresponding to different values of regularization parameter λ.

The correlation coefficient for TNV reaches the maximum when λ equals 0.06.

Compared to the TNV reconstruction with a smaller λ, the reconstruction is less noisy and shows fewer staircase effects. For a larger λ, the image starts to show blurred edges as a result of over-regularization. These effects are similar for the TV regularization. Therefore, it is reasonable to take the reconstruction corresponding to the maximal correlation coefficients as the optimal reconstruction for the regularization method.



(a) TNV, λ=0.06 (b) TV, λ=0.25 (c) SIRT

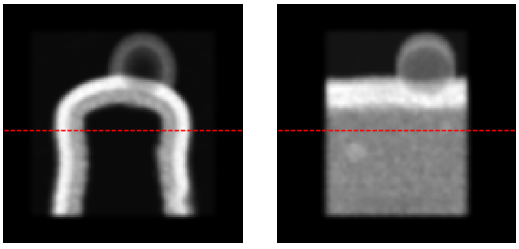
(d) RIO in (a) (e) RIO in (b) (f) RIO in (c)

Figure 3.5: Reconstructions for the core-shell nanoparticle phantom simulation: (a) - (c) EDS reconstructions using TNV, TV and SIRT corresponding to the optimal λ values; (d) - (f) Regions of interest (RIO) in (a), (b) and (c) respectively. The white box in (a) indicates where the RIO is, which is the same for (b) and (c).

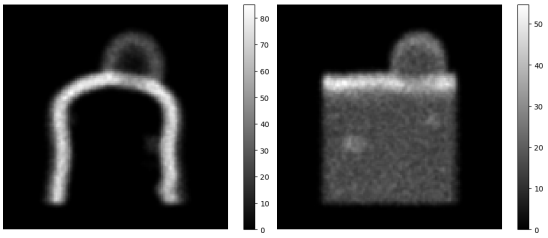
Figure 3.5 compares the TNV reconstruction, the TV reconstruction and the SIRT reconstruction corresponding to the maximal correlation coefficients. The region-of-interest images show that the TNV reconstruction has more effective noise suppression effects and less staircase effects compared to the TV reconstruction. Both regularization methods reduce the missing wedge artifacts in the horizontal direction, while the TNV regularization is more effective and accurate. This is due to the augmentation by the more accurate reduction for missing wedge artifacts in the HAADF-STEM reconstruction from a large number of tilts. In addition, note that the TNV reconstruction does not show false shadows of the Ag structure from the HAADF-STEM reconstruction. Figure 3.6 shows the segmented images and the segmentation errors. The plot indicates that the TNV reconstruction is more consistent with the ground-truth phantom.

(a) TNV (b) TV (c) SIRT (d) Segmentation errors

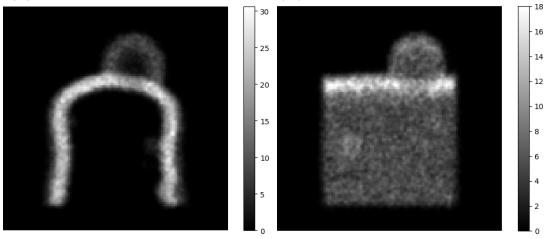
Figure 3.6: (a) - (c) Segmented images corresponding to the TNV, TV and SIRT reconstruction images in Figure 3.5 (a) - (c); (d) Segmentation errors w.r.t the EDS phantom for (a) - (c).



(a) HAADF-STEM image at 2° (b) HAADF-STEM image at 92°

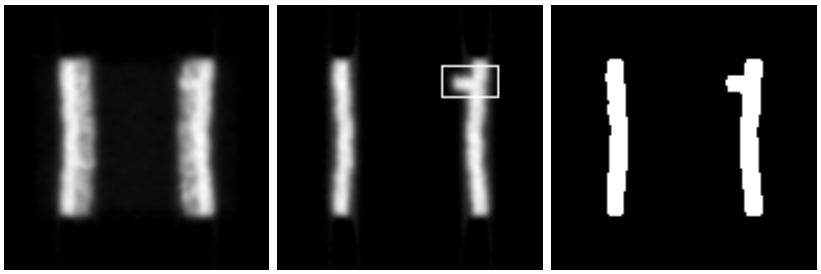


(c) EDS map for Ta at 2° (d) EDS map for Ta at 92°



(e) EDS map for Ta at 2°, 1/3 image intensities (f) EDS map for Ta at 92°, 1/3 image intensities

Figure 3.7: Examples of simulated tilt images for the multislice simulation. The elemental maps for Ta shown here are after Gaussian smoothing. The dashed lines indicate the position of the reconstructed slice. The colorbars indicate the image intensities. The tilt axis is vertical.



(a) HAADF reconstruction (b) EDS reconstruction for Ta (c) Segmented image for (b)

Figure 3.8: Ground-truth for the multi-slice simulation. The white frame in (b) indicates where the "tip" feature is.

3.3.2 Multi-slice simulation study

Data simulation and pre-processing

This study is based on a multi-slice simulation dataset of a 25×25×25 nm semiconductor described in [AR16]. Compared to the phantom simulation, the multi-slice simulation is more realistic since it includes modeling of the real physics and is based on a specimen model at the atomic level. The multi-slice simulation was performed for a high-tension of 200 keV, a convergence angle of 10 mrad and a focused beam on the uppermost point of the specimen. The HAADF detector has an inner angle of 90 mrad and an outer angle of 230 mrad. The focused electron probe was sampled by a 256×256 pixel array over 9.3×9.3 Å. The EDS elemental maps were generated by summing the probability of characteristic emission.

The model is a region of a PMOS finFET. It consists of several layers placed on top of a Si region sequentially: O, HfO_2, Ta, and $TiAlN_2$. Between the HfO_2 layer and the Ta layer there is a carbon nanoparticle contaminant. There are pinholes in the HfO_2 layer which are filled with Ta. We focus on the EDS reconstruction for Ta. The distribution of amorphous Ta is homogeneous in the specimen, for which the total-variation regularization can be applied. Also, the HAADF-STEM data have a large Z-contrast for Ta compared to the other elements. Therefore, the HAADF-STEM reconstruction can be used to augment the EDS reconstruction.

The raw dataset is noise-free and is at atomic resolution (≈ 0.049 nm/pixel), where atomic-scale structures are clearly present. When we tried to add strong noise to the images, the atomic-scale structures were strongly affected by the noise. Therefore, to make the data more suitable for studying noisy data, we down-sampled the simulation data to a lower resolution. Hence, the simulated images were binned from 512 × 512 pixels to 128 × 128 pixels. After that, the images were filtered by a Gaussian filter ($\sigma = 1.0$ pixel) to simulate the effect of a less tightly focused beam. Figure 3.7 (a) and (b) show two examples of the HAADF-STEM projection images. The process of binning approximated a low-resolution simulation. In real experiments, for lower resolution a less tightly

focused beam should be used so that the entire area after binning pixel will be filled by the beam. In the end, we selected the tilt series of HAADF-STEM images from 2° to 178° every 2°, while the tilts at zone-axis (0° and 90°) were excluded due to the channeling effects.

Similar to the phantom simulation, we generated the maps corrupted by Poisson noise from the noiseless simulation data. In this simulation, we focus on comparing the reconstructions made from different cases of data. Thus we made three data cases for different EDS datasets combined with the same HAADF-STEM data. In the first case, the fully sampled EDS data consists of 30 tilts from 2° to 176° for every 6°, of which two examples are shown in Figure 3.7 (c) and Figure 3.7 (d). The average image intensity is about 16.4. Both the second and third cases simulate the scenarios when the total electron dose is reduced by a factor of 3. The second one is a limited-tilt case, where the number of tilts is decreased from 30 to 8. The angular range is from 2° to 170° for every 24°. The third case is a low-SNR case. The number of tilts is again 30, while the noiseless maps were scaled by a factor of 3 to simulate reducing the dose. In this case, the average image intensity is about 5.8. For this case, two examples are given in Figure 3.7 (e) and Figure 3.7 (f). To all the maps, a Gaussian filter ($\sigma = 0.8$ pixel) was applied as a denoising pre-processing step.

Figure 3.8 (a) and Figure 3.8 (b) respectively show the SIRT reconstructions from the noiseless HAADF-STEM and EDS data, for the 2D slice indicated by the dashed line in Figure 3.7. This EDS reconstruction is regarded as the ground-truth. In addition, Figure 3.8 (c) shows the ground-truth segmented image for the EDS reconstruction. Note that there is a "tip" feature on the top right of Ta reconstruction, which is a simulated semiconductor defect.

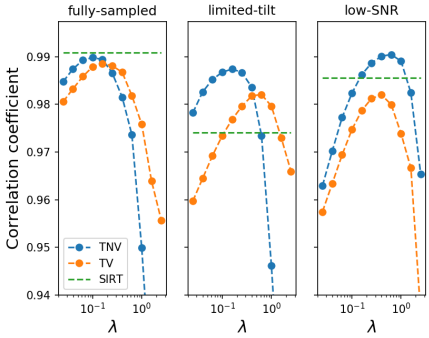
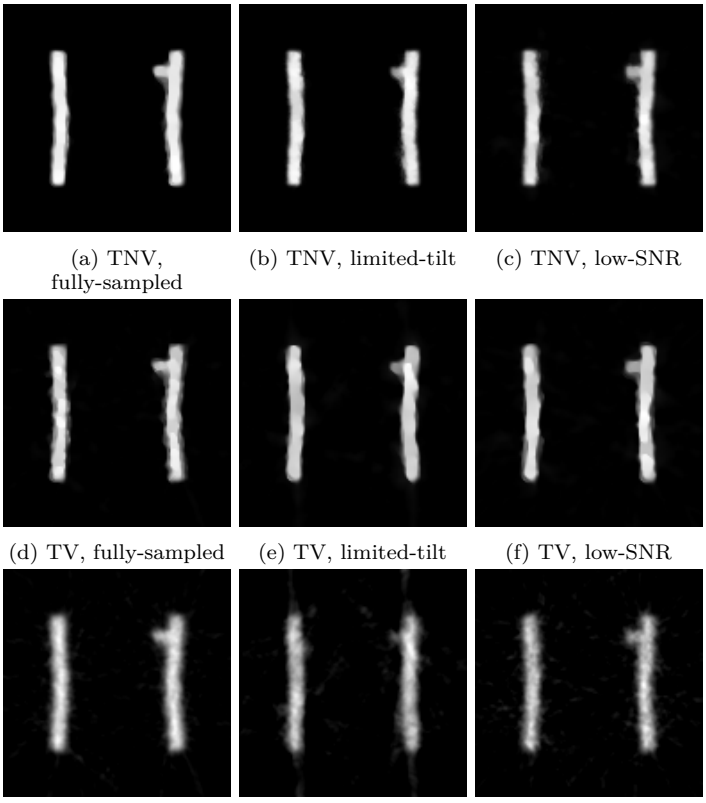


Figure 3.9: Correlation coefficients w.r.t. regularization parameter λ for different numbers of tilts and magnitudes of image intensities in the multi-slice simulation study. The SIRT reconstructions are performed for a fixed relaxation parameter set to 1.



(a) TNV, fully-sampled (b) TNV, limited-tilt (c) TNV, low-SNR

(d) TV, fully-sampled (e) TV, limited-tilt (f) TV, low-SNR

(g) SIRT, fully-sampled (h) SIRT, limited-tilt (i) SIRT, low-SNR

Figure 3.10: EDS reconstructions for the multi-slice simulation: (a) - (c) TNV reconstructions corresponding to the optimal λ values; (d) - (f) TV reconstructions corresponding to the optimal λ values; (g) - (i) SIRT reconstructions. The EDS data in the left, middle and right columns are respectively the fully-sampled, limited-tilt and low-SNR.

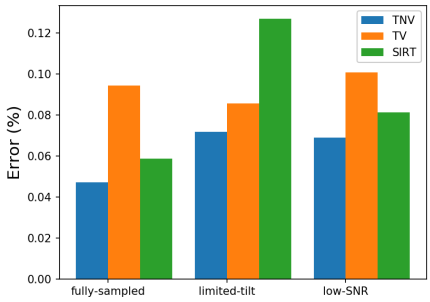


Figure 3.11: Segmentation errors for the reconstructions in Figure 3.10

Reconstruction results

Similar to the phantom study, we performed reconstructions with different regularization parameters and calculated the correlation coefficients w.r.t. the ground-truth, which are plotted in Figure 3.9. Figure 3.10 shows the TNV and TV reconstructions corresponding to the maximal correlation coefficient values as well as the SIRT reconstructions. We segmented these reconstructions using the thresholding values corresponding to minimal segmentation error. Figure 3.11 shows the segmentation errors.

In the fully-sampled case, for which the data quality is relatively good, the maximal correlation coefficients are close for each method. The optimal TNV reconstruction image is less noisy than the TV reconstruction. The SIRT reconstruction is smooth while showing less sharper edges. Thus, although the SIRT reconstruction has slightly higher correlation coefficient, the segmentation error is larger than the TNV reconstruction as the edge information is less accurate.

For the limited-tilt case, the maximal correlation coefficient for TNV reconstruction is larger than the coefficient for TV reconstruction, which is in turn larger than the coefficient for SIRT reconstruction. The SIRT reconstruction loses the "tip" defect feature and corresponds to a large segmentation error, while the TV reconstruction is still affected by staircase effects and has rounded ends. In contrast, the TNV reconstructions show edges that are more accurate and overlap with the HAADF-STEM reconstruction. Also, the "tip" feature is clearer. For the low-SNR case, the SIRT and TV reconstructions have more accurate structure compared to the limited-tilt case, however, are also more noisy. In comparison, the TNV reconstruction is both smooth and has accurate structure.

In general, the TNV reconstruction method outperforms the TV method and SIRT when the number of tilts is reduced or the data SNR is reduced. Comparing the TNV reconstruction image for the limited-tilt case and the low-SNR case, we notice that the latter one is slightly better than the former one. In particular, the "tip" feature is more accurately reconstructed in the latter one. However, we should not draw a general conclusion based on this, as the result may vary

depending on the structure or noise levels of reconstructions.

3.3.3 Real experiment

Data acquisition and preparation

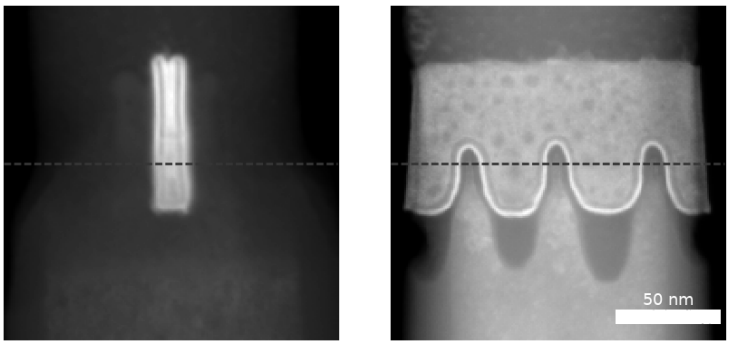
In the last part of this section, we apply the TNV-regularized reconstruction method to real experimental data. The sample is a pillar-shape semiconductor consisting of 8 chemical elements (N, O, Al, Si, Ti, Co, Hf and Ta) [Qiu+15]. The data were acquired using a FEI Titan electron microscope equiped with the SUPER-X system with 4 SDD X-ray detectors. The sample was placed on a Fishione on-axis rotation tomography holder, which allows 360° rotation and avoids detector shadowing [Sla+16b]. For HAADF-STEM, a tilt series of 221 images were taken from 0° to 220° every 1° at a 120 keV high-tension. The angular range was chosen as the maximal angular range allowed by the sample holder in a single smooth acquisition sequence. Two projection images at orthogonal angles are shown in Figure 3.12 (a) and Figure 3.12 (b). For EDS, a tilt series of 47 spectrum-image data-cubes were acquired for an angular range from 0° to 216° at approximately every 5°. The convergence angle was 10 mrad, the high-tension was 120 keV and the probe current was $\sim$ 280 pA. The data acquisition time was about 270 seconds per tilt, which corresponds to a dwell time of 4.11 ms/pixel for the image of 256×256 pixels. The images were later cropped to 192×192 pixels. In addition, another tilt series of HAADF-STEM projection images aligned to the EDS measurement were taken simultaneously.

Figure 3.13 plots the spectrum of total X-ray counts integrated over the spectrum images at 0°, where characteristic lines of all the presenting elements are indicated, among which we focus on the chemical element Ti. In this case, Ti is suitable for applying the TNV-regularized reconstruction method. It has a relatively low Z-contrast compared to the heavy element Ta that surrounds it. This property can be utilized to augment the EDS reconstruction. Figure 3.12 (c) and Figure 3.12 (d) shows two elemental maps of Ti. The elemental maps were extracted by integrating the spectrum images at the Ti-$K\alpha$ line (4.51 keV) for an integration window of 0.25 keV. Neither background subtraction nor PCA denoising has been applied. The maps were then smoothed using a Gaussian filter for $\sigma = 0.8$ pixel.

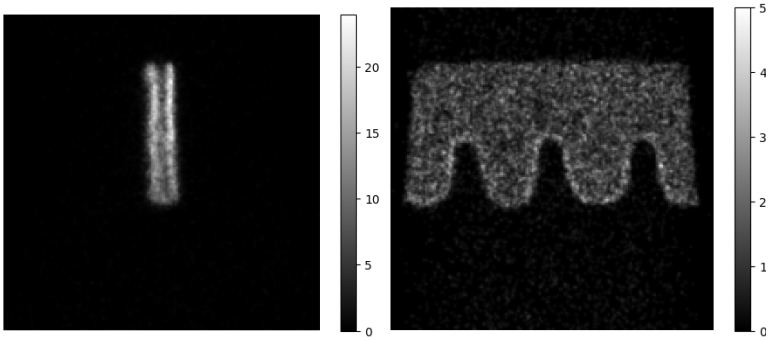
For the alignment between tilt images, the tilt series of HAADF-STEM images were aligned using the cross-correlation approach implemented in the FEI Inspect3D software. After that, the HAADF-STEM images co-acquired with EDS were aligned to the images at the same tilt in the large HAADF tilt series. Lastly, the same alignment settings were applied to the Ti elemental maps.

Reconstructions and results

In addition to the EDS data of 47 tilts, we also removed some tilts to investigate the TNV-regularized reconstruction method for 26 and 14 tilts respectively. Note



(a) HAADF-STEM image at 0° (b) HAADF-STEM image at 90°



(c) EDS map for Ti at 0° (d) EDS map for Ti at 90°

Figure 3.12: Examples of tilt images for the semiconductor sample. The elemental maps shown here are after Gaussian smoothing. The dashed lines indicate the position of the reconstructed slice. The colorbars indicate the image intensities corresponding to the X-ray counts. The tilt axis is vertical.

that in this case, as there is no ground truth measurement, we cannot compute the variation of correlation coefficients w.r.t. the parameters like in the simulation studies. Instead, we first performed reconstructions for different λ values from the 47-tilt data, and chose one that shows the strongest noise-suppression and no over-regularization effect. For TNV and TV, we selected the reconstructions for $\lambda = 0.04$ and $\lambda = 0.10$ respectively. After that, we applied the same regularization parameters to the 26-tilt and 12-tilt datasets. The reconstructions are shown in Figure 3.14.

For the 47-tilt reconstructions, the TNV and TV reconstruction are similar to each other. Although for both the noise is suppressed and sharp edges are enhanced, the edges in the TNV reconstruction are slightly less noisy compared to the TV reconstruction. This is more obvious for the 26-tilt case when the TV reconstruction shows strong staircase-like patterns, while the TNV reconstruction is still smooth. For the 12 tilt case, the TV reconstruction and SIRT reconstruction are very noisy, while the TNV reconstruction is smooth and still shows structural

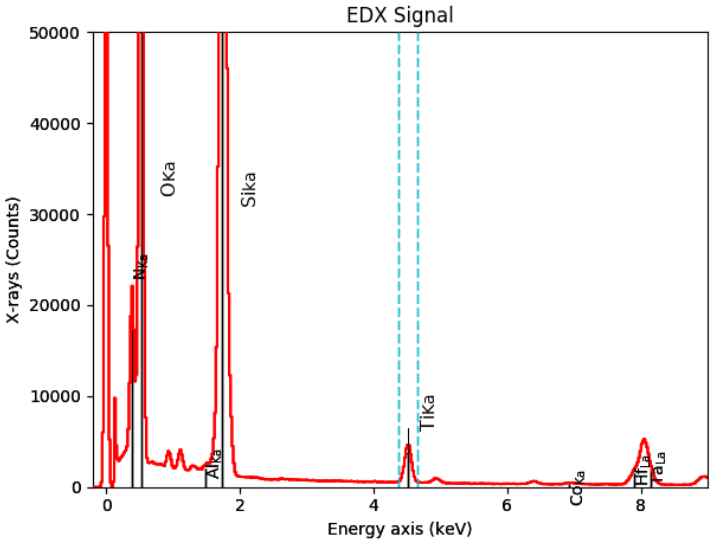


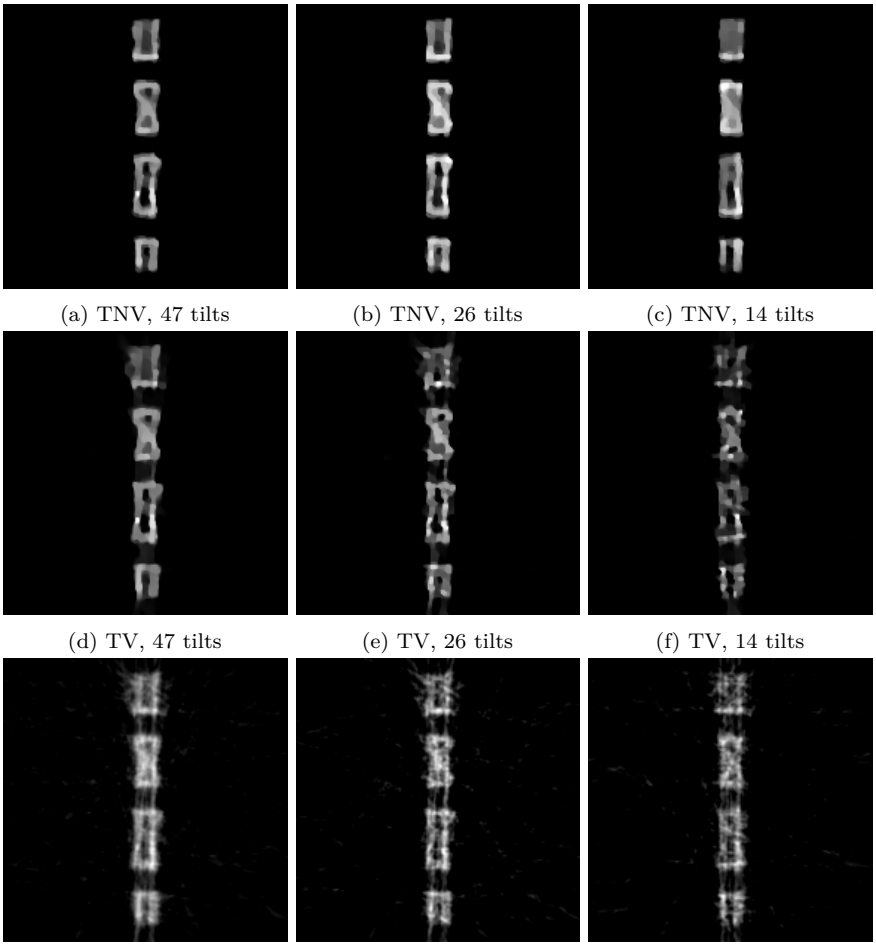
Figure 3.13: An example of the X-ray spectrum for total counts integrated over the spectrum image at 0°. The Y axis corresponds to the total number of X-ray counts in the spectrum image at each energy channel. The black lines indicate the characteristic lines for chemical elements in the sample (N-$K\alpha$, O-$K\alpha$, Al-$K\alpha$, Si-$K\alpha$, Ti-$K\alpha$, Co-$K\alpha$, Hf-$L\alpha$, Ta-$L\alpha$). The dashed lines indicate the integration windows (0.25 keV) for extracting the Ti-$K\alpha$ elemental maps at 4.51 keV.

information similar to the 47-tilt reconstruction.

Figure 3.15 (a) shows the correlation coefficients of the 26-tilt and 14-tilt reconstructions w.r.t. the 47-tilt reconstructions corresponding to the same method. For the 26-tilt case, the correlation coefficients show that the TNV reconstructions are more linearly related to the 47-tilt reconstruction, which indicates that this reconstruction is more accurate even when some EDS data is missing. Figure 3.15 (b) shows the segmentation errors w.r.t. the segmented image in Figure 3.15 (c), which was segmented based on the reconstruction image Figure 3.14 (a). For the 26-tilt case, quantitative information based on the TNV reconstruction is more accurate compared to the others. Nevertheless, for the 14-tilt case, although the correlation coefficient and segmentation error for the TNV reconstruction are slightly smaller than those for the TV reconstructions, they are nearly the same as for the SIRT reconstruction. This may be because the image is so blurred that the metrics fail.

3.4 Discussion and conclusion

In this chapter, a new tomographic reconstruction approach based on EDS is proposed. The EDS reconstruction is performed together with a HAADF-STEM

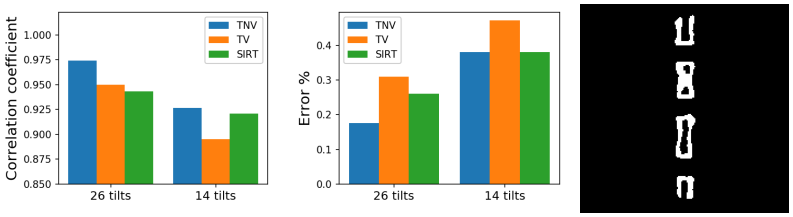


(a) TNV, 47 tilts (b) TNV, 26 tilts (c) TNV, 14 tilts

(d) TV, 47 tilts (e) TV, 26 tilts (f) TV, 14 tilts

(g) SIRT, 47 tilts (h) SIRT, 26 tilts (i) SIRT: 14 tilts

Figure 3.14: EDS reconstructions for the real experimental data. The EDS data corresponding to the left, middle and right columns consists of 47, 26 and 14 tilts respectively; (a) - (c) TNV reconstructions; (d) - (f) TV reconstructions; (g) - (i) SIRT reconstructions.



(a) Correlation coefficients (b) Segmentation errors (c) Binary image used to compute the segmentation error.

Figure 3.15: Measurements for the 26-tilt and 14-tilt EDS reconstructions in Figure 3.14.

reconstruction, and they are jointly regularized by a shared TNV term. Using TNV regularization, it is possible to suppress noise and accurately reconstruct from a limited number of tilts, similar to TV regularization. TNV regularization further incorporates that the reconstruction images have sparse gradients that point in the same directions. Therefore, it encourages the EDS reconstruction to form edges overlapping with the more accurate and less noisy HAADF-STEM reconstruction.

The proposed method has been investigated on a phantom simulation, a multi-slice simulation and a real experimental dataset. For all the datasets, we used a large number of high-quality HAADF-STEM tilt images to augment the EDS reconstruction from only a small number of tilts. Note that even if the number of HAADF-STEM tilt images is smaller, the EDS reconstruction should still be augmented by the method due to the high SNR of HAADF-STEM. Also note that it is still an unsolved question how to set the regularization parameter automatically.

For the phantom simulation, the TNV method demonstrates noise suppression effects and correction of missing wedge artifacts. The reconstructed edges match those of the HAADF-STEM reconstruction. In addition, the reconstructed image shows no false structures resembling the structures in the HAADF-STEM reconstruction. For the multi-slice simulation, we compared the reconstructions for reduced dose. The reconstructions regularized by TNV are consistently more accurate compared to those regularized by TV or without regularization. Lastly, the TNV regularization was applied to real experimental data, and also shows effective noise suppression and sharp-edge enhancement. The TNV reconstruction made from 26 tilts still shows structures similar to the many-tilt reconstruction, while quantitative characterization may fail if the number of tilts is decreased to 14.

As an alternative to the method proposed here, the TNV regularization could also be incorporated in a sequential way: first perform the HAADF-STEM reconstruction, then use the reconstruction as a TNV prior to perform the EDS reconstruction. The sequential approach is more computationally efficient if the same HAADF-STEM reconstruction can be repeatedly used. Also, it gives the opportunity to perform the HAADF-STEM reconstruction using different reconstruction methods with various priors. However, these need further research and

experimental validation and will be studied in the future. Also, we will explore to combine multiple elements together with one TNV term.

In conclusion, the proposed method can be used to obtain more accurate reconstructions for data acquired using the conventional data acquisition scheme or for existing datasets. Particularly, it can be used to augment the noisy EDS reconstruction for elements of low concentrations, which may be most interesting to characterize. Moreover, using the proposed method, it is possible to adopt shorter data acquisition time or smaller beam currents while keeping nearly the same image quality. It also paves the way for developing a faster EDS tomography pipeline when the next-generation EDS detectors with higher detection rates are available.

Acknowledgment

The author would like to thank Dr. Richard Aveyard and Prof. dr. Bernd Rieger (TU Delft, Delft, the Netherlands) for providing the multi-slice simulation data. The author also thank Dr. Yang Qiu (imec, Leuven, Belgium - presently at South University of Science and Technology of China) and Dr. Hugo Bender (imec, Leuven, Belgium) for providing the experimental dataset of the semiconductor pillar sample.

4

Algorithmic recipes of numerical methods

4.1 Introduction

In materials science, the compositional characterization in three dimensions (3D) is important for understanding the properties of nanomaterials. Energy-dispersive X-ray spectroscopic (EDS) STEM allows mapping of the distributions of chemical elements in 2D by detecting the X-rays emitted from the specimen. A 3D volumetric image of these chemical distributions can then be reconstructed from a tilt series of the 2D maps [Sag+07; Lep+13; WC16]. Such a technique is referred to as *EDS tomography*. However, EDS tomography is limited by many practical issues [Sla+16b; Kra+17; Bur+16]. One of the most significant issues is the limited number of detected X-ray counts caused by low emission rates and small solid angles of detectors. As a result, strong Poisson noise is present in the tilt series of elemental maps, which leads to reconstructions with low signal-to-noise ratios (SNRs). In addition, the number of tilt images is often small due to the long data acquisition time. The limited number of tilts results in an ill-posed inverse problem, which, together with the high levels of noise, strongly limits the accuracy of the reconstructed volume. The possibilities for improving the quality of the measured data are often limited by the electron dose that the sample can withstand.

This chapter is based on:
Z. Zhong, W. J. Palenstijn, N. R. Viganò, and K. J. Batenburg. "Numerical methods for low-dose EDS tomography". *Ultramicroscopy* 194 (2018), pp. 133–142.

Ill-posed inverse problems have been studied extensively in (electron) tomography, and various reconstruction methods have been developed ([WC16, Chapter 7], [BL08]). However, choosing the most appropriate algorithm in the context of a specific sample and specific imaging conditions is currently problematic for practitioners in EDS tomography.

The aim of this chapter is to provide guidelines for using and combining three different types of methods: statistical modeling, variational regularization and bimodal tomography. These modules are chosen based on the assumptions made for data statistics, sample structures and instrumental setups respectively. As a result, we provide the possibility to tailor the reconstruction algorithm as a *recipe* composed of *ingredients* chosen for each module.

First of all, we describe how the tomographic reconstruction process can be modeled as an inverse problem with Poisson statistics, whereas the conventional alternative is based on Gaussian-statistic data. For instance, the simultaneous iterative reconstruction technique (SIRT) [GB08], used in [Lep+13; Sla+16b], actually solves an inverse problem assuming Gaussian noise. Poisson noise is addressed in a separate denoising step and by the smoothing effects introduced by SIRT. However, smoothing blurs the images and reduces the resolution. Also, inaccurate modeling may introduce artifacts in the reconstructed images. For EDS mapping with low X-ray counts, it is reasonable to assume the image intensities as measurements of Poisson processes like in many other photonic imaging modalities, e.g. positron emission tomography (PET). Image reconstruction with Poisson statistics has already been studied extensively [SV82; HW16].

Secondly, we present the module for variational regularization methods. These have been developed to address the issue of overfitting (to noise) present in direct modeling methods such as maximum likelihood estimation (MLE) [SV82] in situations with extremely low counts [HW16; YF02; Bar10]. For instance, total variation (TV) regularization is widely adopted. It encourages sparsity of gradients, which helps to suppress noise, promote piecewise constant structures and reduce the artifacts caused by missing data [Gor+12; BO13].

In addition, for EDS tomography, the reconstructions for different chemical elements often share image features, such as edges. Total nuclear variation (TNV) regularization – an expansion of TV – encourages such common edge locations of correlated reconstructions in addition to promoting sparse gradients [Hol14; Dur+16]. There are many other regularization methods such as total generalized variation which encourages piecewise smooth structures [BKP10]. In this chapter, we focus on TV and TNV as our ingredients for the purpose of demonstration.

Thirdly, even with regularization, the reconstructions may still be highly inaccurate when lacking accurate data. In situations with strong noise, TV regularization may introduce staircasing artifacts in the reconstruction [BO13]. The third module augments the reconstruction with additional accurate data by combining EDS tomography with other imaging modalities. Here, we use the HAADF-EDS bimodal tomography (HEBT) technique that was proposed in Chapter 2 and our paper [Zho+17]. HEBT considers the HAADF-STEM projection images, which usually have higher SNRs and resolution, to be the weighted sum of the EDS

maps for all present elements. In Chapter 2, HEBT is based on a Gaussian noise model. In this chapter, we contribute to HEBT by introducing the formulation for reconstruction with Poisson statistics.

All these ingredients can be implemented as solving minimization problems in the reconstruction process. In this chapter, we combine the ingredients of different modules into a *single* optimization problem that can be solved by a generic algorithm. Choosing the right combination of ingredients can lead to complementary effects. For example, HEBT implies a constraint that may suppress the staircasing artifacts introduced by the variational regularization.

The remainder of this chapter is structured as follows. In Section 4.2, we illustrate the theory and the guidelines for choosing ingredients of an algorithmic recipe. In Section 4.3, we investigate and compare the performance of different recipes on simulation and experimental data. In the last section, we draw a conclusion for this chapter. We do not discuss the pre-processing steps in the spectral domain, while in practice these should be carefully considered for the influence on the data statistics. Also, other issues e.g. detector shadowing effects and X-ray self-absorption strongly affect the reconstruction results, but are addressed in other papers [Sla+16b; Kra+17; Bur+16].

4.2 Method

In this section, we will describe the notation for the inverse problem with Gaussian or Poisson statistics, the regularization methods as well as the adapted HEBT method. After that, we will discuss the guidelines for constructing recipes.

4.2.1 Notation of EDS tomography

In EDS tomography, the tilt series of projection images, called elemental maps, are extracted from tilt series of spectrum images, which contain a spectrum of X-ray counts for every pixel position. The intensities of the elemental map correspond to the detected X-ray counts emitted from the chemical element.

We first formulate the relationship between the reconstructed image and the ideal measurement data without noise corruption. Under the thin-film assumption, the ideal data are proportional to the expected numbers of X-ray counts that are in turn proportional to the concentration of the corresponding element probed by the focused beam [WW06]. Thus the ideal data are proportional to the linear projection of the reconstructed quantities, which are expressed as a vector $\mathbf{g}^e \in \mathbb{R}^{M^e}$. Here M^e denotes the total number of pixels for all angles for element e ($e = 1, \cdots, L$). Consider the specimen to be located in a 3D volume space discretized into N voxels. The reconstructed quantities, which are proportional to the concentration of the element, are expressed as a vector $\mathbf{x}^e \in \mathbb{R}^N$. This linear relationship is modeled by the system of equations:

$$g_i^e = \sum_{j=1}^{N} w_{ij}^e x_j^e. \tag{4.1}$$

The i_{th} pixel position is determined by the beam position and the tilt angle of the specimen. The weight factor w_{ij}^e is determined by the area of voxel j intersected by the focused beam of pixel i. The matrix $\mathbf{W}^e = (w_{ij}^e)$ describes the EDS imaging setup.

The real data, which are corrupted by noise, are expressed as a vector $\mathbf{p}^e \in \mathbb{R}^{M^e}$. The reconstruction problem is then to determine the unknown $\mathbf{x}^e$ such that if we compute the projection of $\mathbf{x}^e$, the discrepancy between the real and computed data is minimized. It is common to assume that the real data are ideal data corrupted by Gaussian distributed noise, which is a valid approximation when the number of X-ray counts is large. In this case, we take the sum of squared errors between the measurement data and the ideal data as the data discrepancy, expressed as:

$$\mathscr{D}_{\mathrm{L2}}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e) = \| \mathbf{W}^e\mathbf{x}^e - \mathbf{p}^e \|_2^2, \tag{4.2}$$

which is denoted as *L2 data discrepancy* in this chapter, named after the l_2 norm ($\| \cdot \|_2$).

However, when the number of X-ray counts is small, the Gaussian model is not an accurate approximation anymore. A more solid assumption is to consider the real data as Poisson distributed measurements taking the ideal data as the expected values. We then use the *Kullback-Leibler (KL) divergence* [Ber+10; Csi91] to define the data discrepancy, which is expressed as:

$$\begin{aligned}\mathscr{D}_{\mathrm{KL}}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e) &= \sum_{i=1}^{M^e} (p_i^e - g_i^e + g_i^e \log(\frac{g_i^e}{p_i^e})) \\ &= \sum_{i=1}^{M^e} (p_i^e - \sum_{j=1}^{N} w_{ij}^e x_j^e + \sum_{j=1}^{N} w_{ij}^e x_j^e \log(\frac{\sum_{j=1}^{N} w_{ij}^e x_j^e}{p_i^e})),\end{aligned} \tag{4.3}$$

for $\mathbf{x}^e \succeq 0$.

Given the data discrepancy $\mathscr{D}$ defined by either Eq. 4.2 or Eq. 4.3, the reconstruction is computed by minimizing the discrepancy:

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}}\, \mathscr{D}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e). \tag{4.4}$$

Minimizing KL divergence $\mathscr{D}_{\mathrm{KL}}$ is equivalent to maximizing the log-likelihood of the Poisson distributions for $\mathbf{p}^e$ [Csi91], while minimizing L2 discrepancy $\mathscr{D}_{\mathrm{L2}}$ corresponds to solving a least-squares problem. The popular reconstruction algorithm SIRT in fact solves the problem of minimizing a weighted version of L2 discrepancy [GB08].

4.2.2 Variational regularization

To incorporate TV regularization, we add a regularization term to the minimization problem:

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}}\, \mathscr{D}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e) + \lambda \mathscr{R}_{\mathrm{TV}}(\mathbf{x}^e), \tag{4.5}$$

where λ is the parameter determining the strength of regularization. $\mathscr{R}_{\mathrm{TV}}(\mathbf{x}^e)$ is a regularization term giving the total variation of image $\mathbf{x}^e$, defined as:

$$\mathscr{R}_{\mathrm{TV}}(\mathbf{x}^e) = \sum_{j=1}^{N} \parallel \bigtriangledown x_j^e \parallel_2, \tag{4.6}$$

where $\bigtriangledown$ is the discrete approximation of the gradient operator. If the reconstruction image is 3D, $\bigtriangledown$ approximates the gradients in the X, Y and Z directions respectively using the forward difference as $\bigtriangledown x_j = (\bigtriangledown^X x_j, \bigtriangledown^Y x_j, \bigtriangledown^Z x_j)^T$. Note that reconstructions can also be performed by stacking 2D reconstructions of each slice, for which the gradients are only computed in the X and Y directions. In practice, it is more preferable to directly reconstruct in 3D to also incorporate regularization in the Z direction. The TV defined in this chapter is called isotropic TV [BO13], for which the gradient magnitude at pixel location j is computed as the l_2 norm of the gradient.

In addition to sparse gradients, we can use TNV regularization to incorporate the correlation between reconstructions, such as the reconstructions for multiple elements in the same sample. It is an extension of TV regularization from one-channel images to multi-channel images, which encourages the images in multiple channels to have common edge locations and parallel/antiparallel gradient directions. Suppose there are Q reconstructions that share the same volume space, we can formulate them as a single multi-channel image $\{\mathbf{x}^k\}$ $(k = 1, \ldots, Q)$. The TNV regularization term is defined as the nuclear norm of the Jacobian matrix of the multi-channel image:

$$\mathscr{R}_{\mathrm{TNV}}(\{\mathbf{x}^k\}) = \sum_{j=1}^{N} \parallel (\mathbf{J}\{x^k\})_j \parallel_\star . \tag{4.7}$$

The Jacobian matrix at pixel position j is given by:

$$(\mathbf{J}\{x^k\})_j = \begin{pmatrix} \bigtriangledown^X x_j^1 & \bigtriangledown^Y x_j^1 & \bigtriangledown^Z x_j^1 \\ & \vdots & \\ \bigtriangledown^X x_j^Q & \bigtriangledown^Y x_j^Q & \bigtriangledown^Z x_j^Q \end{pmatrix}, \tag{4.8}$$

where the nuclear norm $\parallel \cdot \parallel_\star$ is given by the l_1-norm of the vector consisting of the matrix' singular values. Minimizing TNV encourages the rank-sparsity of the Jacobian matrix, which leads to parallel or anti-parallel gradient vectors.

To apply TNV regularization on the EDS reconstructions for all chemical elements, we can set $\{\mathbf{x}^k\} = \{\mathbf{x}^e\}$ for $e = 1, \ldots, L$. In this case, the optimization problem is:

$$\{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} \sum_{e=1}^{L} \mathscr{D}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e) + \lambda \mathscr{R}_{\mathrm{TNV}}(\{\mathbf{x}^e\}), \tag{4.9}$$

where the reconstructions for all elements are computed simultaneously.

TNV allows to correlate multiple reconstructions in a flexible manner. In addition to promoting common features between multiple elemental volumes, it is also possible to relate EDS tomography to other tomographic modalities, such as HAADF-STEM tomography. This is subject to having the HAADF reconstruction sharing common edges with the EDS reconstructions. More details are discussed in Chapter 3 and [Zho+18b]. It is even possible to use the TNV regularization to correlate with EELS-STEM tomography [Hab+14]. Despite the many possibilities to apply TNV, in this chapter, we focus on the TNV regularization defined by Eq. 4.9 that correlates the EDS reconstructions for all elements.

4.2.3 HAADF-EDS bimodal tomography

HEBT is used to perform reconstructions simultaneously from the EDS data and the HAADF-STEM data. The elemental reconstructions are made by minimizing the sum of HAADF-STEM data discrepancy and EDS data discrepancy, based on the assumption that the HAADF-STEM projection data are the weighted sum of the EDS maps for all present elements. The weights are referred to as the *response ratio factors*.

The HEBT method in Chapter 2 is defined for least-squares. Here, we modify the formula so that the KL divergence can be used. The reconstruction problem of HEBT is expressed as:

$$\begin{aligned} \{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} \quad & \alpha \mathscr{D}_{\mathrm{L2}}(\sum_{e=1}^{L} \mathbf{W}^h r^e \mathbf{x}^e; \mathbf{p}^h) + (1-\alpha) \sum_{e=1}^{L} \mathscr{D}(\mathbf{W}^e\mathbf{x}^e; \mathbf{p}^e), \\ & \text{subject to } \mathbf{p}^h = \sum_{e=1}^{L} r^e \mathbf{p}^e, \\ & \mathbf{x}^e \succeq 0, \quad e = 1, \ldots, L, \end{aligned} \tag{4.10}$$

where the first term is the L2 data discrepancy for the tilt series of HAADF-STEM images $\mathbf{p}^h \in \mathbb{R}^{M^h}$, and M^h denotes the total number of pixels for all HAADF-STEM tilt images. The matrix $\mathbf{W}^h \in \mathbb{R}^{M^h \times N}$ is the HAADF-STEM projection matrix that describes the HAADF-STEM imaging setup. r^e's are the response ratio factors for different chemical elements. The second term is the sum of EDS data discrepancies for all the elements. The EDS data discrepancy can be chosen between KL divergence and L2 discrepancy, depending on how the

noise is modeled. The parameter $\alpha \in [0, 1)$ is the trade-off weight between the HAADF-STEM and EDS data discrepancies.

Note that in Chapter 2, the EDS map intensities are scaled by the response ratio factors r^e, which changes the EDS data statistics. Here, instead we move the response ratio factors to the HAADF-STEM term so that the EDS maps remain unchanged. The response ratio factors r^e can be estimated based on the linear equations $\mathbf{p}^h = \sum_{e=1}^{L} r^e \mathbf{p}^e$ using least-squares regression [Zho+17]. Since r^e's are assumed to be spatially invariant, we can bin the images p^h and p^e's to increase the SNRs and improve the accuracy of estimated values.

4.2.4 Preparing the recipe

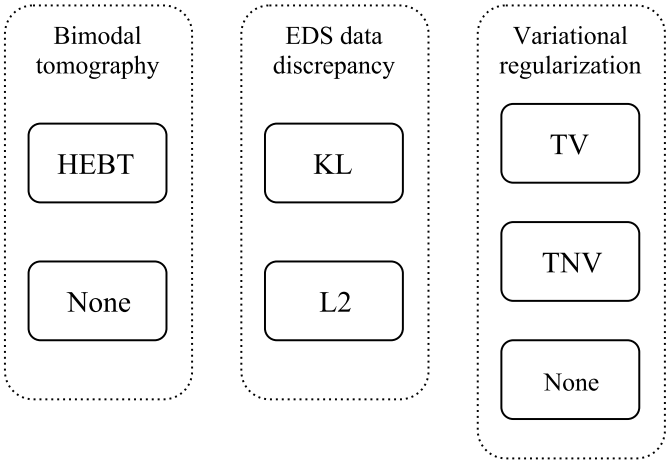


Figure 4.1: Ingredients of the algorithmic recipes.

Based on the above discussions, we can summarize a generic optimization problem that includes the three modules:

$$\{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} \ (1-\alpha) \sum_{e=1}^{L} \mathscr{D}_e(\mathbf{x}^e) + \alpha \mathscr{D}_h(\sum_{e=1}^{L} r^e \mathbf{x}^e) + \lambda \mathscr{R}(\{\mathbf{x}^e\}), \tag{4.11}$$

where the EDS data discrepancy $\mathscr{D}_e(\mathbf{x}^e)$ is *always required*, while the HAADF data discrepancy $\mathscr{D}_h(\sum_{e=1}^{L} r^e \mathbf{x}^e)$ and the regularization term $\mathscr{R}(\{\mathbf{x}^e\})$ are optional. To construct a recipe, we first choose an ingredient for each module according to the list in Figure 4.1, then make an instance of this optimization problem by setting the minimization terms.

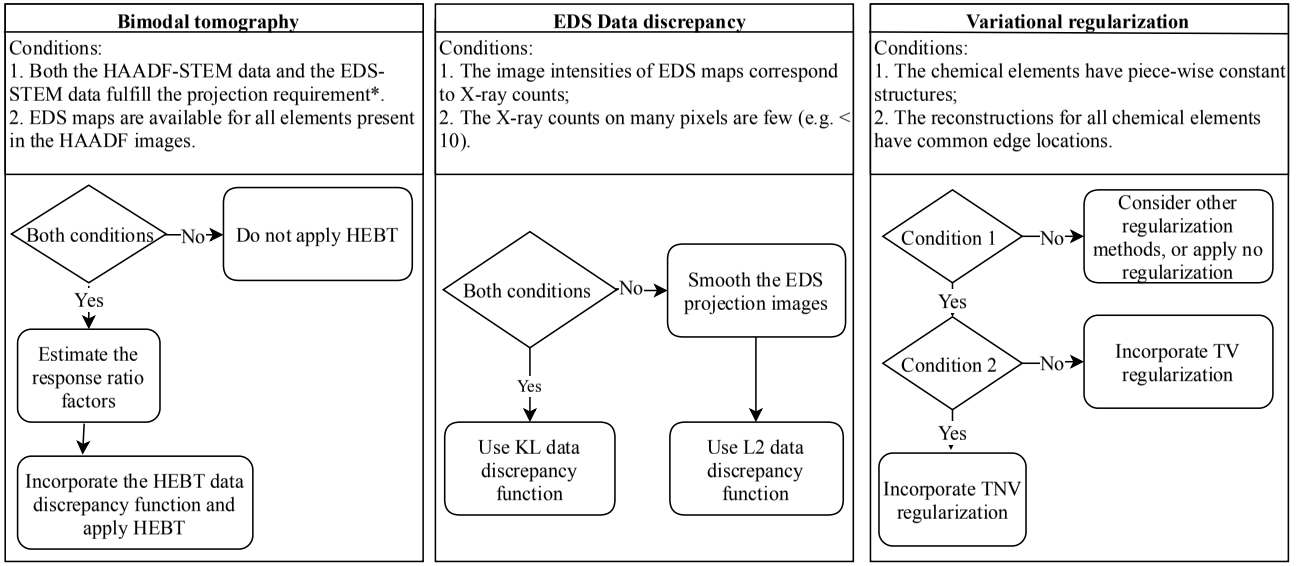


Figure 4.2: Flowchart for making the recipe. (* The projection requirement refers to that the images intensities are linearly related to some integrated physical properties of the sample.)

When constructing a recipe, one should carefully consider the validity of the assumptions behind ingredients. In Figure 4.2, we provide a flowchart as guidelines for choosing ingredients and the conditions w.r.t. the properties of the data and the sample. These conditions are based on mathematical assumptions summarized below:

- *HEBT*: the HAADF-STEM projection images are the linear sum of the EDS maps for all present elements.
- *KL data discrepancy*: the image intensities of EDS maps correspond to X-ray counts that follow Poisson distributions.
- *L2 data discrepancy*: the image intensities of EDS maps approximately follow Gaussian distributions.
- *TV regularization*: the reconstruction has sparse gradients, piecewise constant features and sharp discontinuities.
- *TNV regularization*: in addition to the assumption for TV, multiple reconstructions have common edge locations and parallel/antiparallel gradients.

For instance, HEBT should not be included in the recipe when not all the chemical elements present in the HAADF-STEM images are mapped by EDS, or when the HAADF-STEM projection images are strongly affected by nonlinear damping effects.

4.2.5 Solving the reconstruction problem

After making an instance of Eq. 4.11 for the recipe, a numerical algorithm is needed for solving the optimization problem. We use the *Douglas-Rachford primal-dual splitting algorithm* (DR) [BH13] to compute the solution, which is a broadly applicable algorithm for solving convex optimization problems. For our application, the DR algorithm solves the mathematical problem of the following general form:

$$\mathbf{v} = \underset{\mathbf{v}}{\operatorname{argmin}} \quad f(\mathbf{v}) + \sum_{k=1}^{R} g_k(\mathbf{A}_k \mathbf{v}), \tag{4.12}$$

where $f(\cdot)$ and $g_k(\cdot)$'s are proper, convex and lower semicontinuous functions and $\mathbf{A}_k$'s are linear operators.

In fact, all our data discrepancy and regularization terms can be cast into the form of $g_k(\mathbf{A}_k \mathbf{v})$. Therefore, different optimization problems derived from Eq. 4.11 can be solved using the same DR algorithm. In Appendix, we provide more details for fitting our optimization problems into Eq. 4.12.

4.3 Experiments

In this section, we investigate the performance of different recipes on simulation data as well as real experimental data. We use the DR algorithm implemented in the Operator Discretization Library (ODL) [AKÖ17].

It is necessary to measure the quality of reconstruction to compare reconstructions made using different recipes, or based on different HEBT weights α and regularization parameters λ. In this chapter, the quality of reconstruction is measured by the linear correlation coefficient which determines the linear relation between the reconstruction and the ground truth. For the real experimental data, the ground truth is obtained by segmenting the HAADF-STEM reconstruction. The correlation coefficient is computed by:

$$r = \frac{\sum_i (x_i - \bar{x})(v_i - \bar{v})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (v_i - \bar{v})^2}}, \tag{4.13}$$

where $\bar{x}$ and $\bar{v}$ are the mean values of the reconstruction $\mathbf{x}$ and the ground truth $\mathbf{v}$ respectively.

4.3.1 Non-mixed Phantom simulation

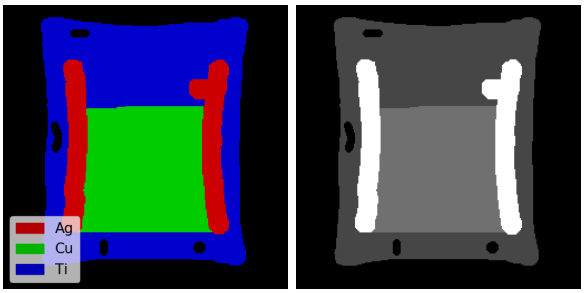
Data simulation

The 2D phantom resembles a structure that contains three homogeneous compositions, which are shown in different colors in Figure 4.3 (a). We assume that the image contrast scales are respectively $z_{Ag} = 47^{1.7}$, $z_{Cu} = 29^{1.7}$ and $z_{Ti} = 22^{1.7}$ given the corresponding atomic numbers Z of these elements, so that the contrast scales as Z^α with α chosen as 1.7 [Tre11]. The HAADF phantom is shown in Figure 4.3 (b).

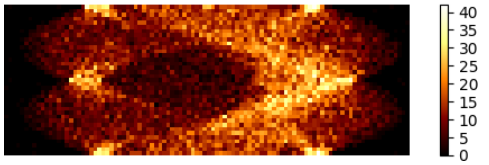
We simulated a tilt series of 1D projection images for the HAADF-STEM phantom for every 5° from 0° to 180° using the ASTRA Toolbox [PBS13]. In addition, we simulated tilt series of 1D maps for each individual element. A realistic value for the image intensity can be determined by considering the incident beam current, the probe live time, the fraction of incident electrons causing ionization, the fluorescence yield, the detector solid angle, and the detector efficiency [Che+16]. In this chapter, we simply set the intensities to absolute scales close to real experimental data for the brevity of the chapter. We then applied Poisson noise by drawing random numbers for expected values given by the noiseless map intensities. Figure 4.3 (c) shows the simulated maps with noise for Ti. The mean image intensity on non-background pixels is 11.76.

Reconstruction results

For this dataset we can choose KL data discrepancy, TNV regularization and HEBT as the ingredients for our preferable recipe (KL-TNV-HEBT), based on



(a) Elemental distribution (b) HAADF Phantom



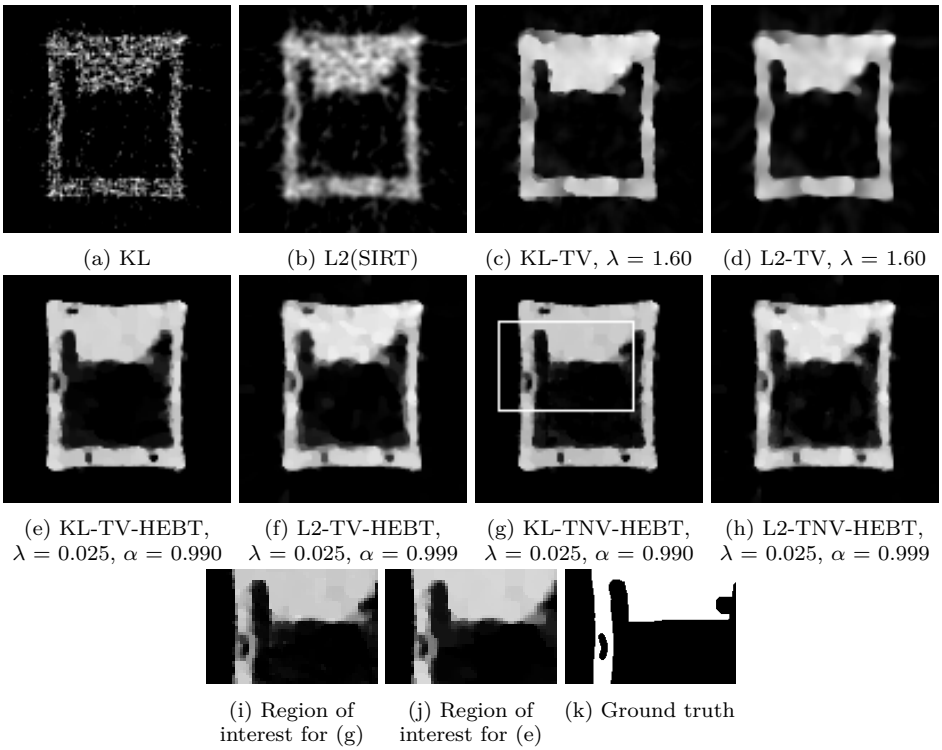
(c) Simulated noisy maps for Ti. The row and column correspond to the tilt angle and the beam position respectively. The mean image intensity on non-background pixels is 11.76.

Figure 4.3: The non-mixed phantom simulation data.

the observation that the Poisson noise is strong and the individual elements have homogeneous structures that share edges. Additionally, reconstructions based on other recipes were also performed for comparing the effects. Although reconstructions were made for all elements, only the reconstructions for Ti are shown for the brevity of this section.

For comparison, we first show the non-regularized reconstructions. The reconstruction (Figure 4.4 (a)) based on the KL divergence was computed by solving Eq. 4.4 using the DR algorithm. Figure 4.4 (b) shows the reconstruction based on L2 data discrepancy computed using the SIRT algorithm for 50 iterations, combined with a pre-smoothing using a Gaussian filter ($\sigma = 1.0$). In fact, SIRT also incorporates implicit regularization on the image smoothness, which is determined by the number of iterations. The L2 reconstruction is less noisy than the KL reconstruction due to the smoothing effect. However, SIRT strongly blurs the small structures.

Second, we performed EDS reconstructions with TV-regularization with the KL or the L2 data discrepancy (KL-TV/L2-TV). The reconstructions were made for different values of regularization parameter λ, for which the correlation coefficients were computed and plotted in Figure 4.6 (a). Figure 4.4 (c) and (d) respectively show the optimal KL-TV or L2-TV reconstructions that correspond to the largest correlation coefficients. Compared to the non-regularized reconstructions, these reconstructions are more homogeneous with sharper edges. We observe that small structures have also been smoothed by the TV regularization.



(a) KL (b) L2(SIRT) (c) KL-TV, $\lambda = 1.60$ (d) L2-TV, $\lambda = 1.60$

(e) KL-TV-HEBT, $\lambda = 0.025$, $\alpha = 0.990$ (f) L2-TV-HEBT, $\lambda = 0.025$, $\alpha = 0.999$ (g) KL-TNV-HEBT, $\lambda = 0.025$, $\alpha = 0.990$ (h) L2-TNV-HEBT, $\lambda = 0.025$, $\alpha = 0.999$

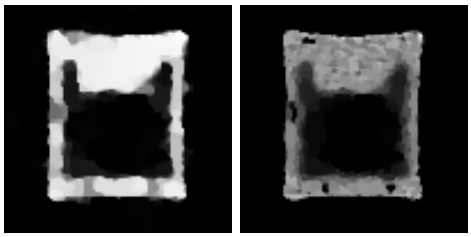
(i) Region of interest for (g) (j) Region of interest for (e) (k) Ground truth

Figure 4.4: Reconstructions for Ti in the non-mixed phantom using various recipes.

Also, the TV regularization introduces obvious staircasing artifacts.

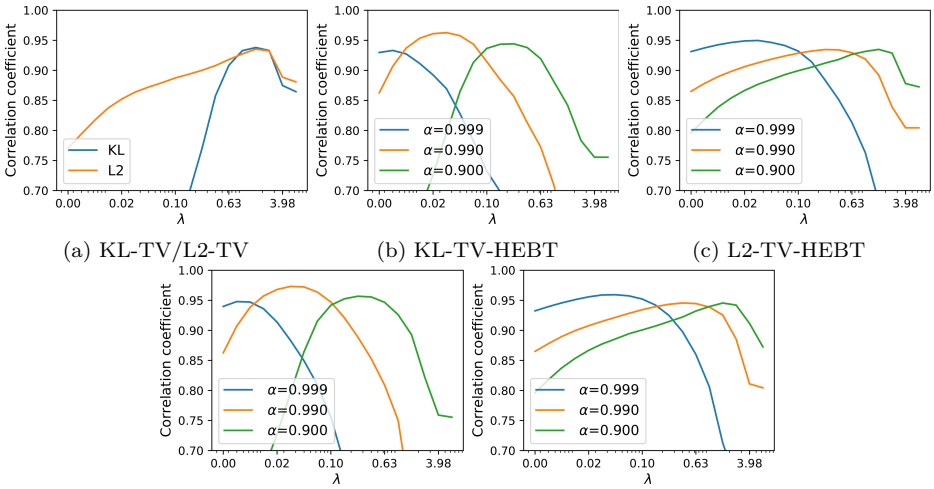
Third, we introduce HEBT to reduce the staircasing artifacts. We performed the recipe of KL-TV-HEBT and L2-TV-HEBT for a range of regularization parameter λ and HEBT weight α. The corresponding correlation coefficients are plotted in Figure 4.6 (b) and (c) respectively. Figure 4.4 (e) and (f) respectively show the optimal reconstructions. We see that the KL-TV-HEBT reconstruction shows clearly reconstructed features at smaller scales (i.e. the "holes") and less staircase artifacts compared to the L2-TV-HEBT reconstruction, due to the proper assumption of data discrepancy. Figure 4.5 (a) and (b) show the optimal KL-TV-HEBT reconstructions for a smaller and a larger α respectively. For the smaller α the reconstruction is similar to the KL-TV reconstruction as the HAADF-STEM data discrepancy is not given with a substantial weight, while for the larger α the reconstruction is more noisy.

Finally, we replaced the TV regularization by TNV regularization to promote the common edges of different elements. Figure 4.4 (g) and (h) are the optimal KL-TNV-HEBT and L2-TNV-HEBT reconstructions. Figure 4.6 shows the zoom-in images for the regions of interest (ROI). Compared with the TV-regularized reconstruction, the TNV-regularized reconstruction is more accurate for areas near



(a) KL-TV-HEBT, $\lambda = 0.251$, $\alpha = 0.900$ (b) KL-TV-HEBT, $\lambda = 0.004$, $\alpha = 0.999$

Figure 4.5: KL-TV-HEBT reconstructions for different α values optimized w.r.t. λ values.



(a) KL-TV/L2-TV (b) KL-TV-HEBT (c) L2-TV-HEBT

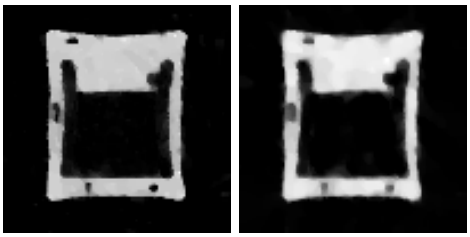
(d) KL-TNV-HEBT (e) L2-TNV-HEBT

Figure 4.6: Correlation coefficients sampled for different values of HEBT weight α and regularization parameter λ.

the common edge locations. The improvement of accuracy is also indicated by the correlation coefficients (see Figure 4.6 (b) and (d)).

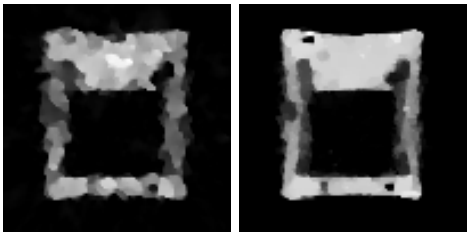
For these regularized HEBT reconstructions (Figure 4.4 (e)-(h)), the KL data discrepancy leads to more homogeneous gray values. However, if the SNRs are high enough, the Gaussian distribution assumed by the L2 data discrepancy can also form a close approximation even though the noise is Poisson distributed. For instance, the L2-TNV-HEBT reconstruction in Figure 4.7 demonstrates little staircasing artifacts when the image intensities are increased by 400% (and therefore the SNRs by 200% for Poisson noise).

In addition, the TV regularization has been shown to reduce the artifacts introduced by the missing wedge [Gor+12], which is a common issue in electron tomography. Figure 4.8 (a) shows the TV-KL reconstruction for data with an an-



(a) KL-TNV-HEBT, $\lambda = 0.025$, $\alpha = 0.990$ (b) L2-TNV-HEBT, $\lambda = 0.063$, $\alpha = 0.999$

Figure 4.7: HEBT-KL-TNV and HEBT-L2-TNV reconstructions for data with 200% SNRs.



(a) TV-KL $\lambda = 1.0$ (b) KL-TNV-HEBT, $\lambda = 0.063$, $\alpha = 0.999$

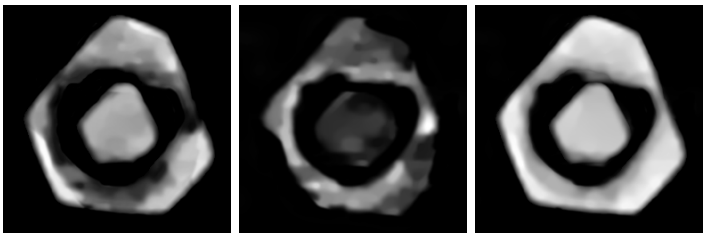
Figure 4.8: reconstructions from data with the missing wedge.

gular range from $-80°$ to $80°$. In comparison, the KL-TNV-HEBT reconstruction (Figure 4.8 (b)) shows more clear structures in the horizontal direction. Therefore, an algorithmic recipe combining proper ingredients might also better reduce missing wedge artifacts.

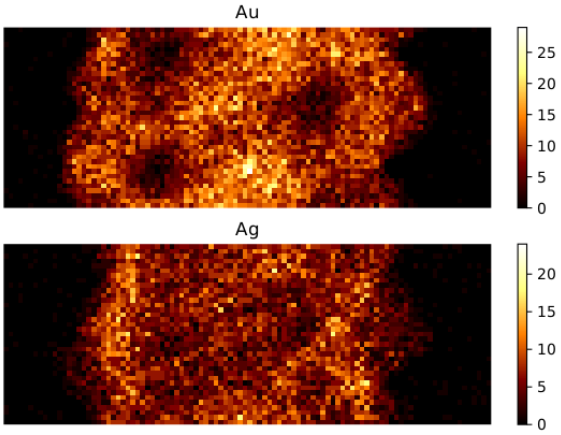
4.3.2 Mixed phantom simulation

Data simulation

The purpose of this simulation is to study the reconstruction methods on inhomogeneous structures liked alloyed materials, as opposed to the homogeneous structures used in the first simulation. The phantom was created resembling the nano-rattle sample investigated in [Zan+16a]. The alloyed nanoparticle consists of Au and Ag components, which have inhomogeneous concentrations. Figure 4.9 (a) and (b) show the Au and Ag phantoms respectively. We created the HAADF phantom as the weighted sum of these two phantoms for $z_{Ag} = 47^{1.7}$ and $z_{Au} = 79^{1.7}$, which is shown in Figure 4.9 (c). Unlike the non-mixed phantom, the structures of Au and Ag components can be hardly distinguished in this image. We simulated the tilt series of 1D EDS maps and HAADF projection data for every $5°$ from $0°$ to $180°$ and added the Poisson noise to the EDS maps following the same procedures as in the first simulation.



(a) Au (b) Ag (c) HAADF



(d) EDS maps for Au and Ag. The row and column correspond to the tilt angle and the beam position respectively. The mean image intensities on non-background pixels are respectively 8.77 and 5.71.

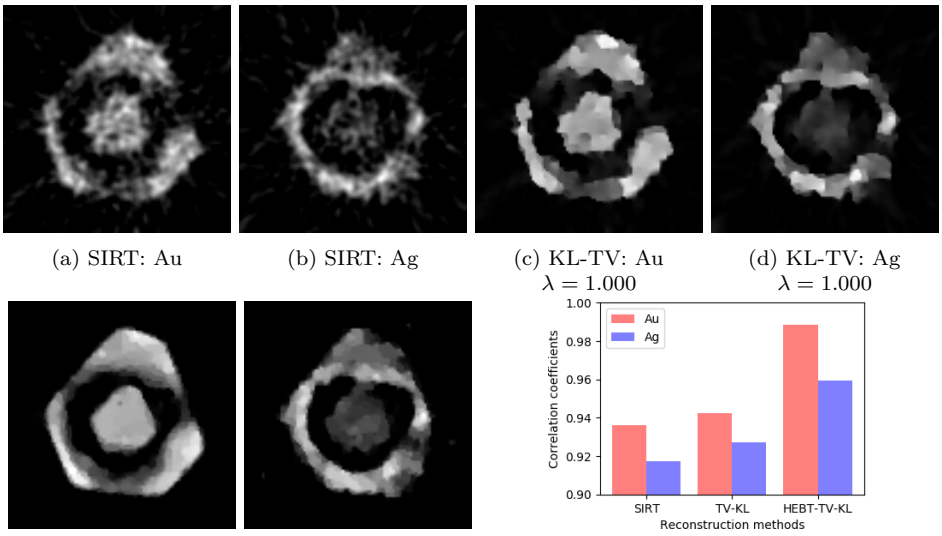
Figure 4.9: The mixed phantom simulation data.

Reconstruction results

TNV regularization is not applicable in this case since the reconstructions for Au and Ag do not necessarily share the same edge locations. We apply TV regularization for noise reduction. Therefore, we consider a recipe of KL-TV-HEBT that satisfies the conditions in Figure 4.2.

For comparison, we first performed SIRT reconstructions (for 50 iterations). The results are shown in Figure 4.10 (a) and (b), which demonstrate low SNRs. Also, we performed KL-TV reconstructions. The reconstructions corresponding to maximal correlation coefficients are shown in Figure 4.10 (c) and (d), which show significant staircasing artifacts due to the strong noise.

The optimal KL-TV-HEBT reconstructions are shown in Figure 4.10 (e) and (f). As a result, the combination of HEBT and TV effectively improves the quality of reconstruction. In particular, HEBT reduces the staircasing artifacts and results in more interpretable reconstructed images. The improvement of image quality is verified by the correlation coefficients in Figure 4.10 (g).



(a) SIRT: Au (b) SIRT: Ag (c) KL-TV: Au $\lambda = 1.000$ (d) KL-TV: Ag $\lambda = 1.000$

(e) KL-TV-HEBT: Au $\lambda = 0.040$, $\alpha = 0.99$ (f) KL-TV-HEBT: Ag $\lambda = 0.025$, $\alpha = 0.99$ (g) Correlation coefficients of the reconstructions

Figure 4.10: Reconstructions for Au and Ag in the mixed phantom using various recipes.

4.3.3 Real experimental data

Data acquisition

Table 4.1: Data acquisition specifications

Electron microscope	Tecnai Osiris FEI company
X-rays detectors	SuperX system, FEI company
Scanning time	300 seconds
Accelerating voltage	120 kV
Projection angles range	$-75°$ to $75°$
Projection angle increment	$5°$
Number of tilts	31
Image size	300×300 pixels
Image size after binning	100×100 pixels

We now investigate the proposed method on a real experimental dataset. The sample is a core-shell nanoparticle of an Au cube embedded in an Ag particle, which has been investigated in Chapter 2 and paper [Zho+17]. The two components have clear boundaries, homogeneous densities and different Z-contrasts. Thus, the core-shell nanoparticle is suitable for applying a TNV regularization.

The experimental data, which consist of a tilt series of spectrum images and correlated HAADF-STEM projection images, were acquired using an electron microscope equipped with four silicon drift detectors. The specifications of the EDS data acquisition are listed in Table 4.1. During the tilt acquisition, only the X-ray detectors on one side were turned on so that the detector shadowing effects were compensated. However, this approach also limited the number of X-ray counts that could be acquired. After PCA denoising, elemental maps were extracted by integrating the spectrum images near the characteristic peaks (Au: $M_\alpha = 2.15$ keV, $M_\beta = 2.20$ keV and $L_\alpha = 9.70$ keV; Ag: $L_\alpha = 2.98$ keV and $L_\beta = 3.19$ keV) as described in [Zho+17]. The HAADF-STEM tilt series were aligned using the cross-correlation method. The EDS elemental maps were then aligned using the same alignment settings. The intensity damping in the HAADF-STEM data was corrected using the correction algorithm [Zho+18a]. Finally, all the images were binned to 100×100 pixels so as to increase the SNRs to reasonable levels. Figure 4.11 shows two examples of the elemental maps. Au and Ag have distinct

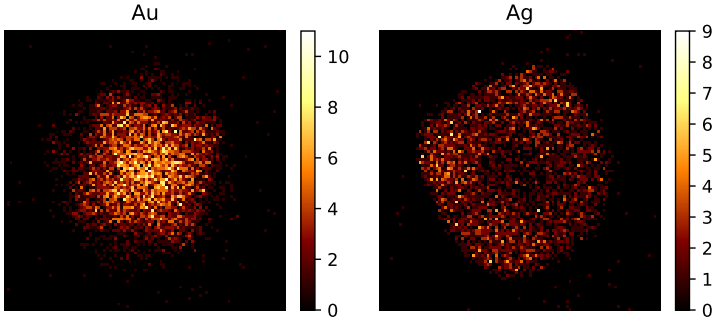


Figure 4.11: Au and Ag elemental maps at 0°

Z-contrasts in the HAADF-STEM images. Therefore, ground truth for evaluating the EDS reconstructions can be obtained by segmenting the HAADF-STEM reconstruction into Au and Ag components. Figure 4.12 (a) shows a slice of the 3D reconstruction for HAADF-STEM, which was made with TV regularization to promote piecewise constant structures and to facilitate the subsequent segmentation. Figure 4.11 (b) and (c) show the subsequent segmented images for Au and Ag respectively.

Figure 4.12 (d) and (e) show the SIRT reconstructions for Au and Ag from the EDS maps, which are indeed noisy and inaccurate. We hope to use a tailored recipe to make more accurate reconstructions. Given the low X-ray counts, the sample structure and the correlated HAADF-STEM data, we apply a KL-TNV-HEBT recipe.

(a) HAADF reconstruction (b) Au segmentation (c) Ag segmentation (d) SIRT: Au (e) SIRT: Ag

Figure 4.12: Reconstructions for slice number 50. The HAADF reconstruction was performed in 3D with TV regularization using the DR algorithm.

Results and Discussion

We first searched for the optimal α and λ parameters for the KL-TNV-HEBT recipe. Since it is time-consuming to compute 3D reconstructions, we selected a 2D slice to sample reconstructions. Here we used the slice at the center of the sample (number 50), which should give a good estimation for the SNRs of the entire volume. Figure 4.13 (a) and (b) show the correlation coefficients with the segmented HAADF-STEM reconstruction, computed for the 2D reconstructions at slice 50. The correlation coefficients for Ag reach maximum at $\alpha = 0.9900$ and $\lambda = 0.10$, at which the correlation coefficient for Au is also close to maximal.

We then applied the λ and α to the entire volume. We performed the reconstruction for the entire volume with regularization in 3D. Figure 4.14 (a)-(f) show some slices of the 3D reconstruction. For comparison, we also performed 2D regularized reconstruction for each slice, some of which are shown in Figure 4.14 (g)-(l). Figure 4.14(m)-(r) show the ground truth for evaluating these reconstructions, which were obtained by segmenting the TV-regularized HAADF reconstruction. The 3D reconstructions are smoother and more accurate compared to the 2D reconstructions, since the large variation in the direction of rotation axis was penalized. Figure 4.13 (c) compares the correlation coefficients for 3D and 2D reconstructions for every slice in the volume. Once again, we conclude that 3D reconstructions are to be preferred when regularizations are applied.

4.4 Conclusion

When characterizing the chemical structure of nanomaterials in 3D by EDS tomography, the limited number of tilt EDS maps, each having a limited signal-to-noise ratio, often leads to noisy and inaccurate EDS tomographic reconstructions. In this chapter, we show that the reconstruction can be improved by using an algorithmic recipe that combines several sophisticated methods for modeling the reconstruction problem. We also provide guidelines for tailoring the recipes based on the specific sample/dataset.

Different algorithmic recipes have been used to reconstruct from both simula-

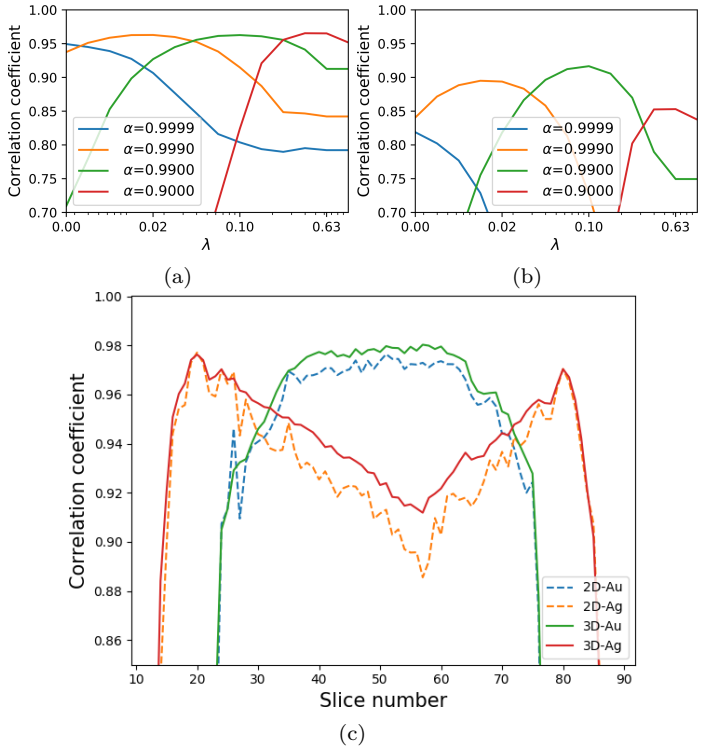


Figure 4.13: (a) and (b): correlation coefficients for the KL-TNV-HEBT reconstructions for slice number 50, sampled for different λ and α values. (c): correlation coefficients for all slices of 2D or 3D reconstructions.

tion and real experimental data. We evaluated the accuracy of reconstructions based on the correlation coefficients w.r.t. ground truth. For all these experiments, the algorithms lead to more accurate reconstruction compared to more naive algorithms when they are tailored for the dataset and sample.

In conclusion, even with very limited data, EDS tomographic reconstruction can still be made accurately using the right recipe. This is useful for characterizing samples sensitive to large dose, or for data measured in a short time. Moreover, it has the flexibility to include other modeling or regularization methods, which allows to extend the options of ingredients. In the future, we will also explore automatic mechanisms for selecting parameters to make the advanced algorithms more accessible.

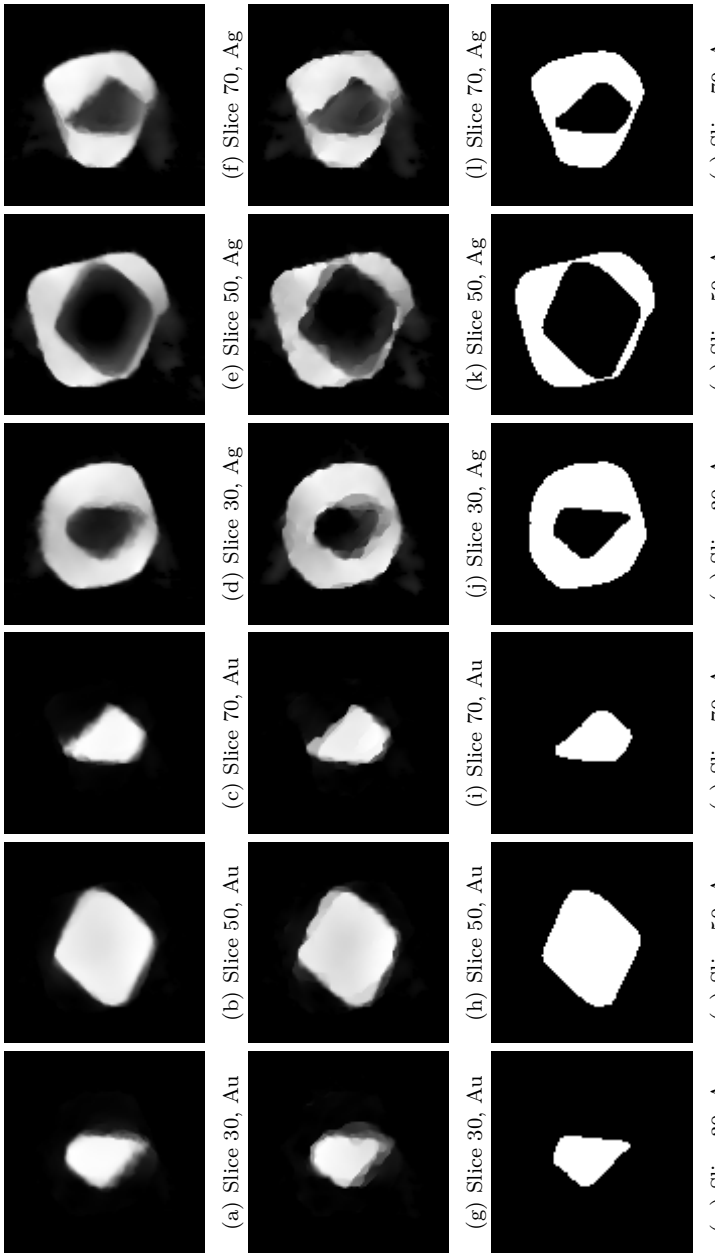


Figure 4.14: KL-TNV-HEBT reconstructions. (a)-(f): slices of 3D reconstructions; (g)-(l): 2D reconstructions; (m)-(r): ground truth (GT) for evaluating the reconstructions.

4.5 Appendix

In this chapter we use the Douglas-Rachford primal-dual splitting algorithm to solve the optimization problem as a sum of multiple objective functions, which is a broadly applicable algorithm for solving the following convex optimization problem from [BH13]:

$$\min_{v} \quad f(v) + \sum_{k=1}^{R} (g_k \square l_k(\mathbf{A}_k v - \langle v, z \rangle), \tag{14}$$

where $f(\cdot)$, $g_k(\cdot)$'s and $l_k(\cdot)$'s are proper, convex and lower semicontinuous functions and $\mathbf{A}_k$'s are linear operators. The infimal convolution $g_k \square l_k(\cdot)$ is defined as:

$$g_k \square l_k(v) = \inf_{y} g(y) + l(v - y). \tag{15}$$

By setting $z = 0$ and

$$l_k(v) = 0 \text{ if } x = 0, \infty \text{ if } x \neq 0, \tag{16}$$

we simplify the mathematical problem to Eq. 4.12.

To construct a optimization problem based on Eq. 4.12 given a reconstruction recipe, we set $f(\cdot) = 0$, and map $g_k(\cdot)$'s and $\mathbf{A}_k$'s to our functions and operators. For example, for a KL-TNV-HEBT recipe, the optimization problem can be made from:

$$\begin{aligned}
&f(\{\mathbf{x}^e\}) = 0, \\
&g_1(\{\mathbf{v}^e\}) = (1-\alpha) \sum_{e=1}^{L} D_{\mathrm{KL}}(\mathbf{v}^e; \mathbf{p}^e), \\
&\mathbf{A}_1 = \{\mathbf{W}^e\}, \\
&g_2(\{\mathbf{v}^e\}) = \alpha \parallel \sum_{e=1}^{L} \mathbf{v}^e - \mathbf{p}^e \parallel_2^2, \\
&\mathbf{A}_2 = \{\mathbf{W}^h\}, \\
&g_3(\{\mathbf{Z}_j\}) = \lambda \sum_{j=1}^{N} \parallel \mathbf{Z}_j \parallel_\star, \\
&\mathbf{A}_3 = \mathbf{J} \otimes \mathbf{I}_{\mathrm{N}}.
\end{aligned}$$

The matrix $\mathbf{I}_{\mathrm{N}}$ is the $N \times N$ identity matrix, and $\otimes$ denotes the Kronecker product. Table 2 lists the instances for all functions and linear operators used in this chapter. Another key to deriving the particular DR algorithm instances is to derive $prox_{\sigma}[g_k^*](y)$, which is the proximal operator for the convex conjugate of $g_k(\cdot)$.

Table 2: Mathematical instances for the ingredients

Ingredient	Minimization term	DR algorithm instances
L2	$\mathscr{D}_e(\mathbf{x}^e) = \parallel \mathbf{W}^e\mathbf{x}^e - \mathbf{p}^e \parallel^2$	$g(\{\mathbf{v}^e\}) = (1-\alpha)\sum_{e=1}^{L} \parallel \mathbf{v}^e - \mathbf{p}^e \parallel_2^2$, $\mathbf{A} = \{\mathbf{W}^e\}$
KL	$\mathscr{D}_e(\mathbf{x}^e) = \mathscr{D}_{\mathrm{KL}}(W^e\mathbf{x}^e; \mathbf{p}^e)$	$g(\{\mathbf{v}^e\}) = (1-\alpha)\sum_{e=1}^{L} \mathscr{D}_{\mathrm{KL}}(\mathbf{v}^e; \mathbf{p}^e)$, $\mathbf{A} = \{\mathbf{W}^e\}$
HEBT	$\mathscr{D}_h(\{\mathbf{x}^e\}) = \parallel \sum_{e=1}^{L} \mathbf{W}^h r^e \mathbf{x}^e - \mathbf{p}^h \parallel_2^2$	$g(\{\mathbf{v}^e\}) = \alpha \parallel \sum_{e=1}^{L} r^e\mathbf{v}^e - \mathbf{p}^h \parallel_2^2$, $\mathbf{A} = \{\mathbf{W}^h\}$
None	$\mathscr{D}_h(\{\mathbf{x}^e\}) = 0$	
TV	$\mathscr{R}(\{\mathbf{x}^e\}) = \sum_{e=1}^{L}\sum_{j=1}^{N} \parallel \bigtriangledown x_j^e \parallel_2$	$g(\{\mathbf{Y}^e\}) = \lambda\sum_{e=1}^{L}\sum_{j=1}^{N} \parallel \mathbf{y}_j^e \parallel_2$, $\mathbf{A} = \{\bigtriangledown\}$
TNV	$\mathscr{R}(\{\mathbf{x}^e\}) = \sum_{j=1}^{N} \parallel (\mathbf{J}\{x^e\})_j \parallel_\star$	$g(\{\mathbf{Z}_j\}) = \lambda\sum_{j=1}^{N} \parallel \mathbf{Z}_j \parallel_\star$, $\mathbf{A} = \mathbf{J} \otimes \mathbf{I}_\mathrm{N}$
None	$\mathscr{R}(\{\mathbf{x}^e\}) = 0$	

The exact forms of the proximal operators for the functions in Table 2 are derived and provided in [Dur+16; RL15; SJP12].

Acknowledgment

The author would like to thank Dr. Bart Goris and Prof. dr. Sara Bals (EMAT, University of Antwerp, Antwerp, Belgium) for providing the experimental data in this chapter, as well as Prof. dr. Luis M. Liz-Marzan (Bionanoplasmonics Laboratory, CIC biomaGUNE, San Sebastián, Spain) for providing the nanoparticle samples.

5

Correcting nonlinear damping effects in HAADF-STEM tomography

5.1 Introduction

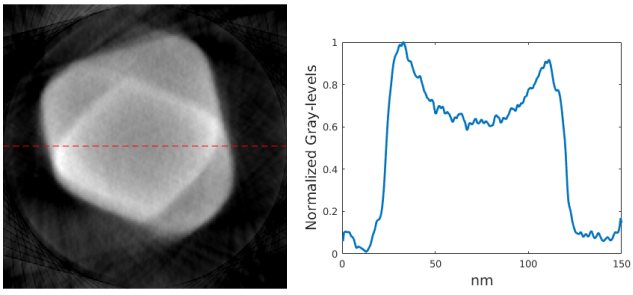
In materials science, electron tomography (ET) is commonly used to characterize the three-dimensional (3D) structural and compositional information of nanomaterials. The 3D image is usually reconstructed from a tilt series of two-dimensional (2D) projections (projection images). The projection images 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 [Mid+01; Küb+05]. 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 [KS88, Chapter 3].

High angle annular dark field (HAADF) scanning transmission electron microscopy (STEM) is commonly used for ET [Mid+01; MW03] under the implicit

This chapter is based on:
Z. Zhong, R. Aveyard, B. Rieger, S. Bals, W. J. Palenstijn, and K. J. Batenburg. "Automatic correction of nonlinear damping effects in HAADF-STEM tomography for nanomaterials of discrete compositions". *Ultramicroscopy* 184.Part B (2018), pp. 57–65.

assumption that the projection requirement can be approximately satisfied. The image intensity approximates to be proportional to the mass-thickness weighted by $\sim Z^2$, where Z is the atomic number [MW03]. 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 [Ave+17]. Another example is that the image intensity damps at high sample thickness due to the multiple scattering events redirecting electrons outside the annular detector, which can occur in all projection orientations. While the zone-axis effects can be easily identified, intensity damping is not easily seen in individual projections. In this chapter, we aim at addressing the nonlinear effects of intensity damping for tomographic reconstruction.

The consequence of intensity damping appears as the cupping artifact in tomographic reconstruction: the gray levels in the center of the reconstructed sample are underestimated while overestimated on the exterior [van+12]. In Figure 5.1 (a), an example of the cupping artifact is given. It is a 2D cross section of an Au-Ag core-shell nanoparticle [Zho+17], reconstructed using the SIRT algorithm [GB08]. If we look at the line-profile of the 2D image (Figure 5.1 (b)), the curve appears in a concave "cup" shape, while ideally it should be flat. The cupping artifacts are caused by the strong damping effects of Au at large thickness, which is illustrated by the simulated relationships between measurement intensity and sample thicknesses using the multislice simulation method [Ave+17] in Figure 5.2. In this example, the linear approximation is only valid for thickness smaller than 8 nm due to the clear damping effect for larger thickness.



(a) (b)

Figure 5.1: (a): 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.

It is important to correct the nonlinear effects and the subsequent cupping artifacts for three reasons. First of all, compositional analysis based on gray levels becomes difficult when the cupping artifacts occur, as gray levels are not proportional anymore to density and atomic numbers. Second, morphological analysis based on segmentation of reconstruction images is hindered by the cupping artifacts. Some straightforward segmentation methods, e.g. Otsu's method [Ots79],

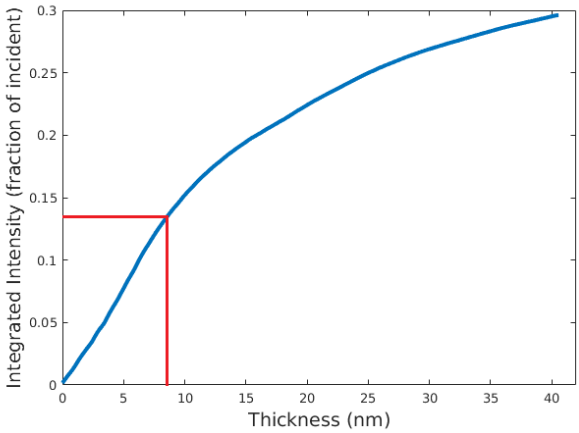


Figure 5.2: Normalized HAADF signal intensity w.r.t the thickness of Au slabs mistilted 10 degrees from the [100] zone axis about the <100> axis, simulated using the multislice method [Ave+17]. The accelerating voltage is 200 kV, the convergence angle is 10 mrad and the detector angular range is 50 - 250 mrad. The intensities are scaled by the incident beam intensity. The red lines indicate the region where intensity is approximately linear to thickness.

require that for each chemical composition there should be one constant gray level. Third, the nonlinear effects limit applying advanced reconstruction algorithms to address a critical issue of ET – the missing wedge artifacts caused by the limited tilt range of the sample. Algorithms such as total variation minimization [Gor+12] reduce the missing wedge artifacts by incorporating prior knowledge i.e. sparsity of the unknown sample. Nevertheless, these algorithms have an even stronger requirement for the linear forward model which is inaccurate due to the nonlinear effects.

Despite these shortcomings of using uncorrected data, there are few publications addressing the nonlinearity issue in ET [Ave+17; van+12]. Nonlinear effects are usually ignored or mitigated during image acquisition by increasing the inner angle of the HAADF detector but at the cost of losing signal strength [Ave+17]. An alternative to adjusting the acquisition parameters is to correct the measured data in a post-processing step by linearizing the projection data, provided that the incident beam intensity is known [van+12]. The method described here requires only the HAADF signal, consequently, it can be applied to correct cupping artifacts in many existing datasets acquired in a conventional manner. The mathematical model of nonlinearity and the concept of linearization in [van+12] are also used in this chapter, which will be explained in Section 5.2.1.

Here, we propose an iterative algorithm to automatically correct the nonlinear effects and the cupping artifacts. It does not require the extra measurement of the incident beam intensity as in [van+12]. Instead, it automatically models the nonlinear effects given the projection data. The algorithm iteratively searches for the minimal distance between the acquired projections and the nonlinear re-projections

of chemical compositions by varying the nonlinear model and the reconstruction image, so as to estimate a nonlinear relationship between the measured HAADF-STEM intensities and sample thickness for all chemical compositions. The algorithm contains the following steps in every iteration: first a reconstruction image with continuous gray levels is made; then the image is segmented into several binary images, each of which corresponds to a chemical composition; after that, the nonlinear effects are modeled by minimizing the projection distance; based on the model, the projection data is linearized at last. The concept of iterative correction has been used to correct beam hardening artifacts for X-ray computed tomography, which is similarly caused by nonlinear intensities [van+11; Hsi+00; van+02].

Our approach is only applicable to samples consisting of several chemical compositions with uniform densities, such as homogeneous or core-shell particles. It is assumed that for these samples the volumetric distributions of the compositions can be approximated well by segmenting the reconstructed image based on gray levels and that this segmentation improves as the correction model applied to the measured data becomes more accurate. In fact, these kinds of samples are commonly studied in materials science. For example, the samples typically studied in the context of discrete tomography [Bat+09; ZPB16] match the requirements.

In Section 5.2, the correction algorithm is explained in detail. In section 5.3, we demonstrate how the nonlinear effects are corrected using this algorithm for real experimental data and phantom simulations.

5.2 The nonlinear model and the correction algorithm

5.2.1 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 [Ave+17], 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 [van+12; WC16], 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

$$p_t = I_0 \exp(-\sum_{e}^{K} \mu_e t), \tag{5.1}$$

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

$$p = I_0(1 - \exp(-\sum_{e}^{K} \mu_e t)) + p_b, \tag{5.2}$$

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 [van+12], which is applicable only if the incident beam intensities can be measured. An advantage of this simple model is that it can easily be transformed into a linear relationship by taking logarithms so that we can avoid solving nonlinear least-squared problems for tomographic reconstruction.

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

We also assume the chemical compositions are not mixed and voxels are small enough to resolve every chemical composition, which means that in each voxel only one element is present. As stated in the introduction, this algorithm is applied to samples with uniform density. Thus we assume that each chemical composition is either present (1) or absent (0) in each voxel. The distribution of chemical composition e is described by binary variables s_{ej}, where $j = 1, \ldots, N$ is the index of voxel.

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

$$p_i = I_0(1 - \exp(-\sum_{e=1}^{K} \mu_e \sum_{j=1}^{N} w_{ij} s_{ej})) + p_b, \tag{5.3}$$

where $i = 1, \ldots, M$.

5.2.2 The correction algorithm

The basis of the correction algorithm is to estimate the nonlinear relationship of Eq. 5.3 based on the reconstructed distributions of chemical compositions. The procedures of the automatic correction algorithm are given in the flowchart (Figure 5.3). The correction is realized iteratively through the following steps: (1) make a

reconstruction image based on the linear model from the projections; (2) segment the reconstruction into a series of binary images, one for each chemical composition; (3) estimate the parameters of the nonlinear model in Eq. 5.3 given the projections and the binary images; (4) reduce the nonlinearities in the projections through a rescaling of the intensities based on the nonlinear model.

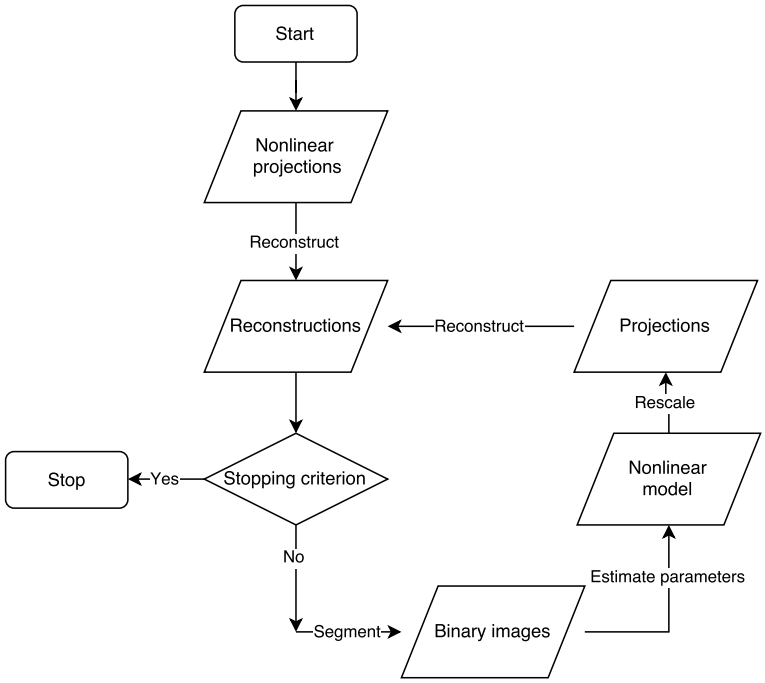


Figure 5.3: Flowchart of the correction algorithm

Before we explain the steps explicitly, we establish 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:

$$\mathscr{C}(I_0, p_b, \boldsymbol{\mu}, \mathbf{S}) = \| \mathbf{p} - I_0(1 - \exp(-\mathbf{W}\sum_{e=1}^{K} \mu_e \mathbf{s}_e)) - p_b \|_2^2, \tag{5.4}$$

where $\boldsymbol{W} = \{w_{ij}\}$, $\boldsymbol{\mu} = \{\mu_e\}$ 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:

$$\frac{c^r + c^{r-1}}{c^{r-2} + c^{r-3}} > t, \tag{5.5}$$

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.

Step 1: Reconstruction

As the first step, a reconstruction with continuous gray levels is made for determining the binary images in the next step. Though it is possible to reconstruct binary images directly using some discrete tomography algorithms (e.g. [Bat+09]), these algorithms will possibly not give better results than basic algorithms given an inaccurate forward model. Thus, we choose to first make a reconstruction $\mathbf{x}$ with continuous gray levels based on a linear model and then segment the reconstruction into binary images $\mathbf{S}$.

The reconstruction is computed using the simultaneous iterative reconstruction technique (SIRT) [GB08] which solves the following least-squares problem:

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmin}} \parallel \mathbf{p}_{lin} - \mathbf{W}\mathbf{x} \parallel_2^2 . \tag{5.6}$$

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

The input for this step is a set of "inearized" projections p_{lin}. For the first iteration, they are just the acquired projections. For the other iterations, they are adopted as the projections that have been rescaled in the previous iteration, which will be explained in Step 4.

Step 2: Segmentation

The binary images are then determined by segmenting the reconstruction image $\mathbf{x}$. As gray levels are related to atomic numbers, we segment the SIRT reconstruction by global thresholding. The thresholds for the segmentation are determined by solving the following optimization problem:

$$\mathbf{S}^* = \underset{\mathbf{S}\in\mathscr{S}}{\operatorname{argmin}} \mathscr{C}(I_0, p_b, \boldsymbol{\mu}, \mathbf{S}). \tag{5.7}$$

The solution of this problem is found by straightforward (brute-force) sampling of the space of thresholds, each time evaluating the cost function. In practice, the thresholds are sampled from the minimum to the maximum of gray levels of the SIRT reconstruction in Step 1.

The first iteration is again an exception since parameters have not yet been estimated and the objective function cannot be computed. Thus, the above segmentation method is not applicable. Instead, the thresholds are determined using Otsu's method which finds optimal thresholds based on the gray level histograms [Ots79].

Step 3: Nonlinear parameters estimation

Given the binary images, we can update the free parameters of the nonlinear model $I_0, p_b, \boldsymbol{\mu}$ by minimizing the objective function, which is a nonlinear regression problem. This nonlinear regression problem is solved using the Nelder–Mead method [Lag+98]. To improve the stability of the regression, the three parameters

are estimated separately and iteratively in an inner loop:

For $l = 1 : L$

$$
\begin{aligned}
{p_b}^{l+1} &= \underset{p_b}{\operatorname{argmin}}\, \mathscr{C}(I_0^l, p_b, \boldsymbol{\mu}^l, \mathbf{S}^*);\\
\boldsymbol{\mu}^{l+1} &= \underset{\boldsymbol{\mu}>0}{\operatorname{argmin}}\, \mathscr{C}(I_0^l, p_b^{l+1}, \boldsymbol{\mu}, \mathbf{S}^*);\\
{I_0}^{l+1} &= \underset{I_0>\max(\mathbf{p})}{\operatorname{argmin}}\, \mathscr{C}(I_0, p_b^{l+1}, \boldsymbol{\mu}^{l+1}, \mathbf{S}^*).
\end{aligned}
\tag{5.8}
$$

here l is the iteration number of the inner loop. The estimation algorithm requires initial parameter values. In the experiments, we found that the initial values have little influence on the convergence result but proper initial values help to converge faster. Since we know that the beam intensity I_0 should be at least the maximal image intensity and that the attenuation coefficients μ 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 this chapter.

Step 4: Projection intensities rescaling

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

$$
\mathbf{p}_{lin}^{,} = \log \frac{I_0 + p_b - \mathbf{p}}{I_0}, \tag{5.9}
$$

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.

5.3 Experiments and simulations

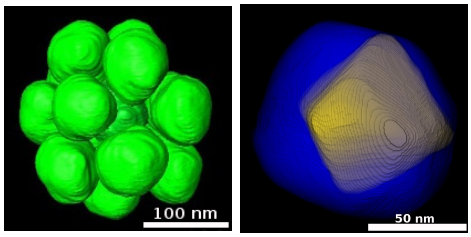
We report the correction of cupping artifacts for two sets of experimental data and three phantom simulations. The experimental data show strong nonlinear effects because the samples consist of thick metallic materials. Two phantom simulations resembling the experimental data were performed, as ground-truth is missing for quality assessment of the reconstruction image due to the lack of other measurement methods. In addition, a phantom of four chemical compositions was simulated to investigate the robustness of the algorithm when more chemical compositions are present, as the experimental samples consist of only one or two chemical compositions.

5.3.1 Experiments

The first experimental sample is an assembly consisting of 16 Pt nanoparticles, each of which has a diameter of about 60 nm (Figure 5.4 (a)) [Sán+12]. It has

Table 5.1: 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^o to 74^o	-75^o to 75^o
Projection angle increment	2^o	5^o



(a) (b)

Figure 5.4: (a): 3D volume rendering of the Pt nanoparticle assembly. (b): 3D volume rendering of the Au-Ag nanoparticle.

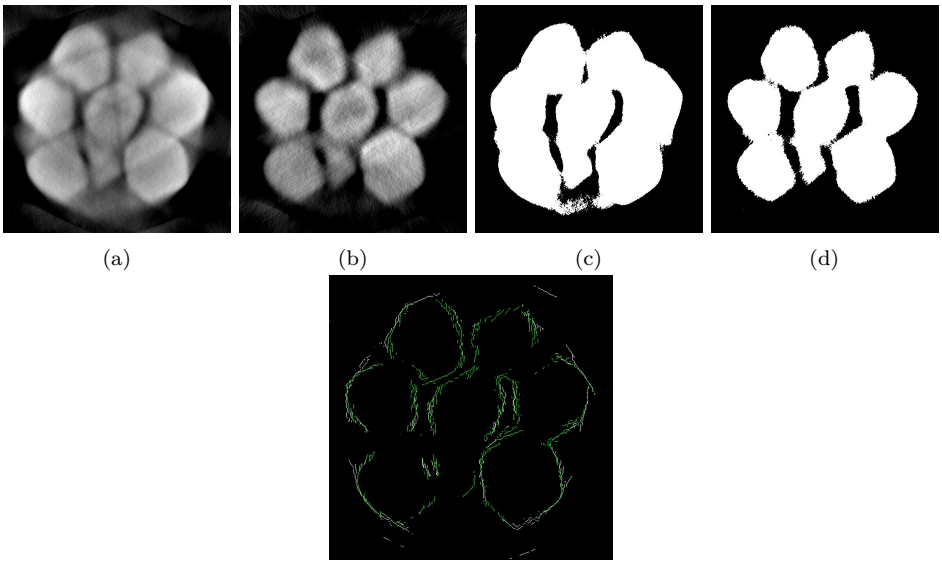
only one chemical composition and a relatively more complex structure than the second sample.

The second sample is a hetero-nanoparticle, which is an Ag nanoparticle with a diameter of approximately 110 nm with an embedded Au octahedron [Zho+17]. It is studied as a case where the cupping artifacts reduce the image contrast between different chemical compositions. The specifications of data acquisition are listed in Table 5.1.

This dataset has been used to investigate HAADF-EDS bimodal tomography (HEBT) in Chapter 2 and [Zho+17]. In that study, the authors have noticed that the raw data had strong intensity damping which not only limited straightforward segmentation of the HAADF reconstructions but also undermined the validity of HEBT based on linear models. Therefore, in [Zho+17] the data has been linearized in the data preprocessing as mentioned in the paper.

Results: nanoparticle assembly

Figure 5.5 (a) is the initial SIRT reconstruction, based on which a binary image (Figure 5.5 (c)) was segmented using Otsu's method. Figure 5.5 (b) and (d) are the reconstruction and the binary image acquired after applying the correction algorithm. To obtain morphological information which is difficult to observe in the reconstruction images, we plotted their edges (Figure 5.5 (e)) which are detected using a Sobel filter that depends on the derivatives of gray levels.



(a) (b) (c) (d)

(e)

Figure 5.5: (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).

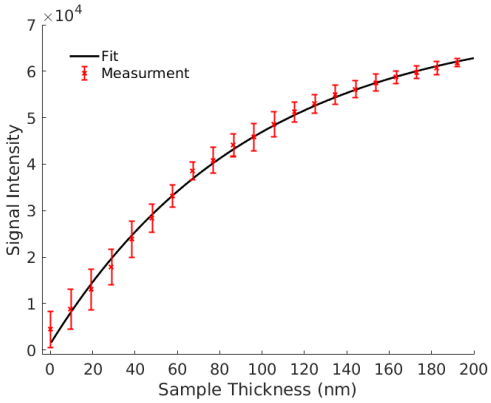


Figure 5.6: The nonlinear damping model fitted for projection signal intensity w.r.t. sample thickness of the nanoparticle assembly. The error bars indicate mean intensities and the standard deviations of the projection data.

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 Figure 5.6,

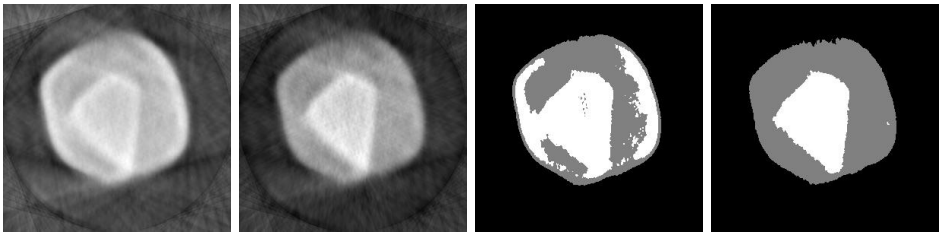
where the thickness was computed as the forward projection of the binary image after correction. The error bars indicate the mean intensities and the standard deviations of the projection intensity.

Results: Au-Ag core-shell nanoparticle

For this experimental data, the SIRT reconstructions and segmented binary images before and after correction are shown in Figure 5.7. In addition, the line profiles across the reconstruction images for some iterations are plotted in Figure 5.8 to demonstrate how gray levels evolve during a run of the correction algorithm.

As discussed in the introduction, the nonlinear effects also hinder adopting prior knowledge to reduce missing wedge artifacts. In this data, the projections were only acquired from -75^o to 75^o. We thus compared reconstructions using advanced reconstruction algorithms: total-variation minimization (TV-min) [Gor+12], discrete algebraic reconstruction technique (DART) [Bat+09] and total variation regularized DART (TVR-DART) [ZPB16], which incorporate the prior knowledge of image sparsity, discrete gray levels and image sparsity combined with discrete gray levels respectively. The images reconstructed from the nonlinear projections and the corrected projections are given in In Figure 5.9.

Finally, we plotted the normalized residuals of the cost function w.r.t. iterations for the two experimental data (Figure 5.10). For the first and second experiments, the cost values converge to stable minimums after 16 and 12 iterations respectively.



(a) (b) (c) (d)

Figure 5.7: (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.

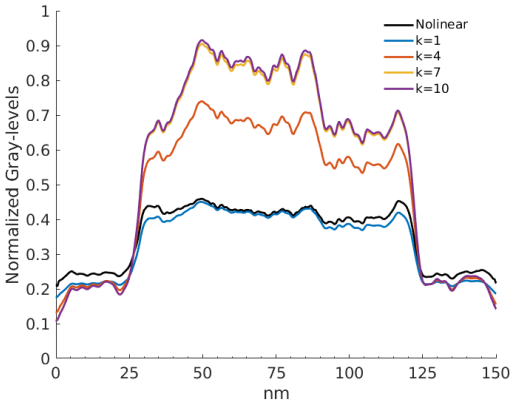


Figure 5.8: Cross-section line profiles of the SIRT reconstructions of the Au-Ag nanoparticle at different iterations.

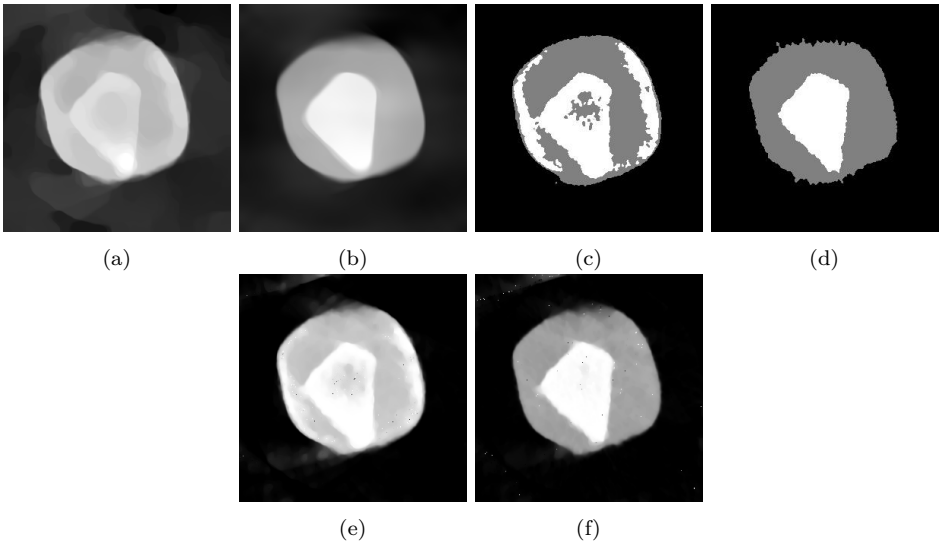


Figure 5.9: (a)/(b), (c)/(d) and (e)/(f) are the TV-min, DART and TVR-DART reconstructions of the Au-Ag nanoparticle from projections before/after the correction respectively.

5.3.2 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.

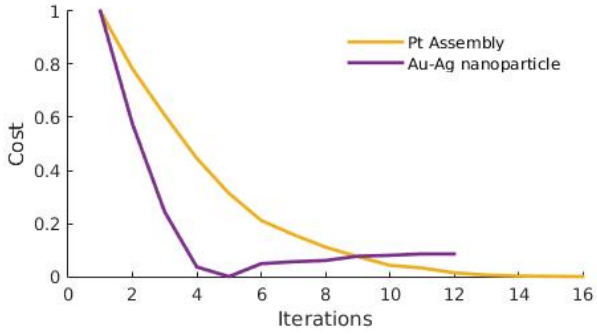


Figure 5.10: The residuals of cost function (Eq. 5.10) w.r.t. iterations for the two experimental datasets.

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:

$$err = \frac{1}{K}\sum_{e}^{K}\sum_{j}^{N} \parallel s_{ej} - g_{ej} \parallel / \sum_{j}^{N} g_{ej}, \tag{5.10}$$

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.

Results of simulations

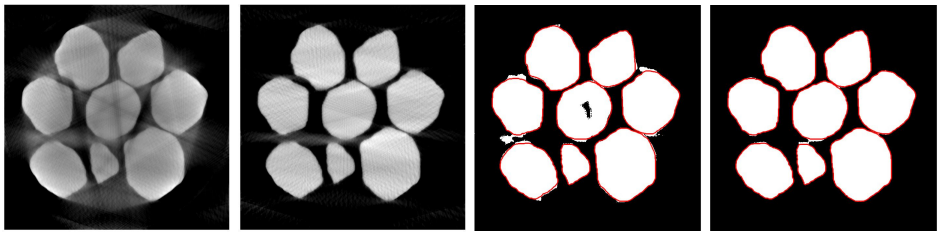
The first phantom resembles the nanoparticle assembly, whose contours are plotted in Figure 5.11 (c) and (d). Figure 5.11 (a) is the initial SIRT reconstruction before correction, based on which a binary image (Figure 5.1 (c)) was segmented. Figure 5.11 (b) and (d) show the SIRT reconstruction and the 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 Figure 5.12, where (a) and (b) are the SIRT reconstructions before and after correction respectively.

The binary images in Figure 5.12 (c) and (d) were segmented from the SIRT reconstruction images. The ground-truth phantom is plotted using red and green contours for Au and Ag respectively. The error metrics of the binary images are respectively 56% and 1% before and after correction.

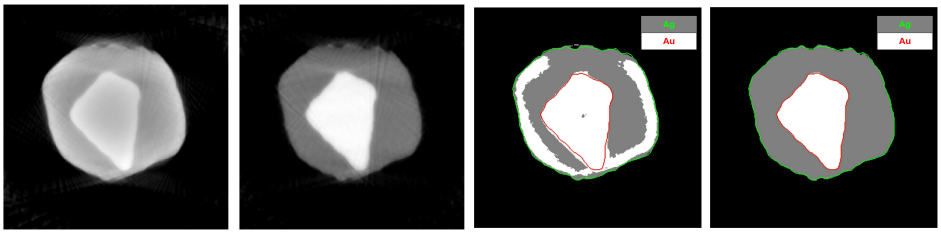
Table 5.2: Errors Metrics of Binary Images.

	Before correction	After correction
Nanoparticle assembly phantom	5%	2%
Au-Ag nanoparticle phantom	56%	1%
Phantom of four chemical compositions	69%	20%



(a) (b) (c) (d)

Figure 5.11: (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.



(a) (b) (c) (d)

Figure 5.12: (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.

The third phantom simulation presents the case when four chemical compositions exist in the same phantom. The SIRT reconstruction images before and after correcting the nonlinearity are shown in Figure 5.13 (a) and (b) respectively,

while the corresponding binary images are given in Figure 5.13 (c) and (d). The error metrics of the binary images are respectively 69% and 20% before and after correction.

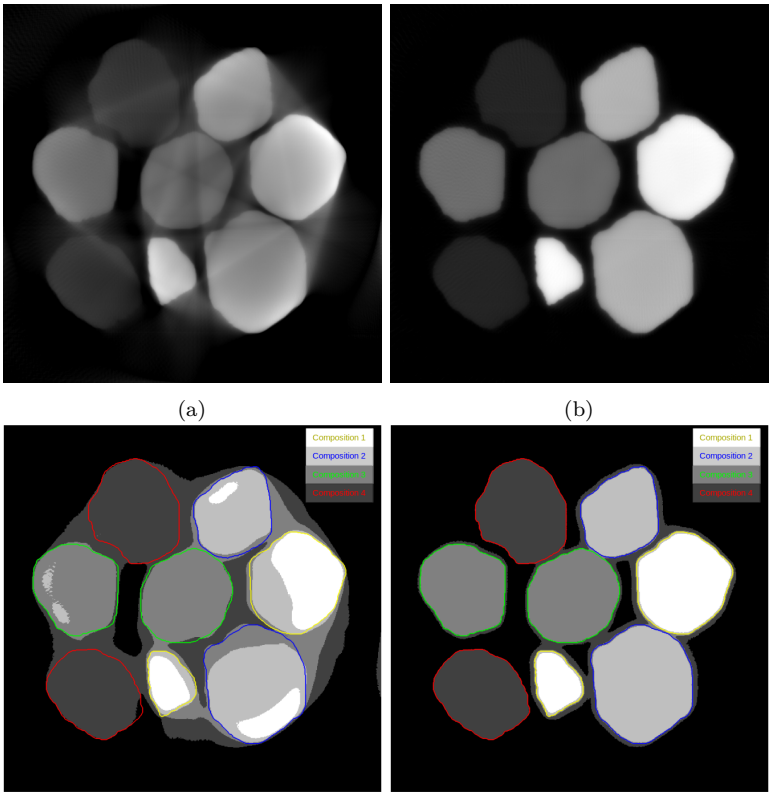


Figure 5.13: (a) and (b): SIRT reconstructions of the phantom simulation with four chemical compositions before and after correcting the nonlinear effects. (c) and (d): Binary images segmented based on (a) and (b) respectively. The colorful contours show the shape of the phantom particles of four different chemical compositions.

5.3.3 Discussion

In the initial reconstruction of the nanoparticle assembly (Figure 5.5 (a)), the artifacts appear, on one hand, as dark streaks elongated from the gaps between particles. On the other, they appear as underestimated gray levels in the interior, for which we see missing pixels in the binary image (Figure 5.5 (c)).

The correction algorithm successfully reduced these artifacts and produced

images easier to interpret. The correction algorithm also changed the morphology of the reconstruction image (Figure 5.5 (b)), as can be seen from the plot of edges. The change may be due to the removal of the overestimated gray levels on the background. The plot of fitting (Figure 5.6) shows that the experimental data matches our nonlinear model, demonstrating a damping effect following the exponential rule. It is also noticeable that the standard deviations decrease at large thickness, which can be explained by noting that the errors introduced by segmentation are relatively smaller at larger thickness.

In the initial SIRT reconstruction image of the Au-Ag nanoparticle (Figure 5.7 (a)), the cupping artifacts caused the loss of contrast between Au and Ag, even though Au and Ag have a large difference in atomic number. As a result, many pixels were misclassified in the binary images (Figure 5.7 (b)). The algorithm corrected the experimental data and enhanced the contrast between Au and Ag. Demonstrated in Figure 5.8, the contrast between Au(center) and Ag(outskirts) was enhanced step by step. At last, the Au and Ag particles were segmented correctly based on gray levels.

The Au-Ag nanoparticle should be suitable for incorporating prior knowledge to correct missing wedge artifacts. It contains two distinct compositions with uniform densities, and thus the reconstruction image should be sparse and have constant gray levels. However, before the correction, incorporating different variants of prior knowledge in the reconstruction actually appears to be detrimental to the image quality, as can be seen in Figure 5.9. Especially the tip of the Au particle was expanded. The expanded tip probably is a mixture of cupping artifacts and missing wedge artifacts. After correcting the nonlinear effects, the linearized projection data was suitable for using the advanced algorithms as the reconstructions show.

The first two phantom simulations show artifacts (in Figure 5.11 (a) and Figure 5.12 (a)) very similar to those from the experimental data, which indicates that the modeling of nonlinear effects is accurate. Both reconstructions after correction are free of these artifacts, and are in good agreement with the ground-truth phantom, as the error metrics were reduced (Table 5.2).

For the third simulation, we see cupping artifacts (Figure 5.13 (a)) with features observed in the previous two cases. First, there are dark streaks and underestimated gray levels. Second, the contrast between different chemical compositions is blurred. These artifacts were corrected after applying the correction algorithm (Figure 5.13 (b)).

The segmented binary images after correction (Figure 5.13 (d)) show a stack of different chemical compositions at the borders of some particles. However, these misclassified pixels are not caused by the cupping artifacts, but due to the limitation of the global thresholding [BS09]. The gray levels in the reconstruction image are continuously dropping at the borders. These pixels were classified into particles 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.

5.4 Conclusion

In this chapter, 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 between 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 chemical compositions as well as in facilitating incorporating prior knowledge to correct the missing wedge artifacts.

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.

Acknowledgment

The author would like to thank Dr. Thomas Altantzis, Dr. Bart Goris and Prof. dr. Sara Bals (EMAT, University of Antwerp, Antwerp, Belgium) for providing the experimental data in this chapter, as well as Prof. dr. Luis M. Liz-Marzan (Bionanoplasmonics Laboratory, CIC biomaGUNE, San Sebastián, Spain) for providing the nanoparticle samples.

6
Conclusion

In this PhD thesis, we propose several approaches to pave the way for HAADF-STEM + EDS tomography: (1) the HAADF-EDS bimodal tomographic reconstruction technique, which is based on jointly modeling the consistence of the two imaging modalities; (2) TNV-regularized joined reconstruction which allows to incorporate the prior knowledge that common edges exist in the reconstructions from HAADF and EDS data respectively; (3) a set of algorithmic recipes to tailor various reconstruction algorithms for given experimental conditions and sample properties; (4) an algorithm for automatically correcting the nonlinear damping effects in HAADF-STEM tomographic data.

Experimental results of the HEBT algorithm in Chapter 2 show that HEBT enables investigating the structure of chemical elements with lower noise levels compared to element-wise EDS reconstruction. By promoting consistency between the forward projections of the element-specific reconstructions and the low-noise HAADF-STEM data, reconstructions are obtained that are more accurate in comparison to pure EDS tomography, resulting in a lower discrepancy between the reconstructions and the ground-truth.

Chapter 3 shows that using TNV regularization, it is possible to encourage a reconstruction to have edges overlapping with another reconstruction. The EDS element-specific reconstruction can be augmented by joining it with the HAADF-STEM reconstruction that has higher signal-to-noise ratios and is computed from a larger number of tilts. Alternatively, by promoting shared edges between the EDS and HAADF reconstructions, the number of tilts required for EDS can be reduced compared to element-wise EDS reconstruction, while maintaining similar reconstruction quality.

In Chapter 4, the algorithmic recipes enable to tailor a set of reconstruction algorithms based on the actual experiments, sample, and data. In the experimental section, we demonstrate how to choose suitable ingredients and algorithmic pa-

rameters for a variety of data cases. In all these cases, the tailored recipes result in more accurate reconstruction results compared to more naive algorithms.

The above chapters are based on the simple assumption that the projection models are perfectly linear. In practice, this is often not completely valid. The algorithm proposed in Chapter 5 can correct the nonlinear damping effects in HAADF-STEM data. The experimental results show that the algorithm can reduce the cupping artifacts as well as improve further interpretation and segmentation of the reconstruction. For HEBT, correcting the nonlinearity can make the HAADF data consistent with the linear model.

Overall, this thesis provides several approaches to improve the accuracy of element-specific reconstructions made by combining HAADF-STEM and EDS data. An important potential application of this work is the quality inspection in the semiconductor industry. Besides the need of high image quality and the ability to resolve chemical elements, being able to carry out a full ET experiment and the subsequent computational steps in very short time (less than an hour) is crucial for the adoption of these techniques in an industrial R&D setting. In the future, a robust and automated pipeline for the complete tomographic process needs to be developed.

We envision that by extending our proposed methodology with a data acquisition scheme that selects the projection angles in an optimized manner, the time required for acquiring sufficiently many EDS images can be substantially reduced. For further automation of the computational part of the imaging pipeline, automated procedures will need to be developed for setting the values of the various parameters in the reconstruction algorithms in an objective and repeatable manner.

Bibliography

[Aar+15] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers. "The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography". *Ultramicroscopy* 157 (2015), pp. 35–47.

[AKÖ17] J. Adler, H. Kohr, and O. Öktem. *Operator Discretization Library*. 2017. url: `https://github.com/odlgroup/odl`.

[AR16] R. Aveyard and B. Rieger. "Tilt series STEM simulation of a 25×25×25 nm semiconductor with characteristic X-ray emission". *Ultramicroscopy* 171 (2016), pp. 96–103.

[Ave+17] R. Aveyard, Z. Zhong, K. J. Batenburg, and B. Rieger. "Optimizing experimental parameters for the projection requirement in HAADF-STEM tomography". *Ultramicroscopy* 177 (2017), pp. 84–90.

[Bar10] J. M. Bardsley. "A theoretical framework for the regularization of Poisson likelihood estimation problems". *Inverse Problems & Imaging* 4.1 (2010), pp. 11–17.

[BL08] J. M. Bardsley and A. Luttman. "Total variation-penalized Poisson likelihood estimation for ill-posed problems". *Advances in Computational Mathematics* 31.1 (2008), p. 35.

[Bat+09] K. J. Batenburg, S. Bals, J. Sijbers, C. Kübel, P. A. Midgley, J. C. Hernandez, U. Kaiser, E. R. Encina, E. A. Coronado, and G. van Tendeloo. "3D imaging of nanomaterials by discrete tomography". *Ultramicroscopy* 109.6 (2009), pp. 730–740.

[BS09] K. Batenburg and J. Sijbers. "Adaptive thresholding of tomograms by projection distance minimization". *Pattern Recognition* 42.10 (2009), pp. 2297–2305.

[Ber+10] M. Bertero, P. Boccacci, G. Talenti, R. Zanella, and L. Zanni. "A discrepancy principle for Poisson data". *Inverse Problems* 26.10, 105004 (2010), p. 105004.

[BH13] R. I. Bot and C. Hendrich. "A Douglas–Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators". *SIAM Journal on Optimization* 23.4 (2013), pp. 2541–2565.

[BKP10] K. Bredies, K. Kunisch, and T. Pock. "Total generalized variation". *SIAM Journal on Imaging Sciences* 3.3 (2010), pp. 492–526.

[Bur+16] P. Burdet, Z. Saghi, A. N. Filippin, A. Borrás, and P. A. Midgley. "A novel 3D absorption correction method for quantitative EDX-STEM tomography". *Ultramicroscopy* 160 (2016), pp. 118–129.

[BO13] M. Burger and S. Osher. “A guide to the TV zoo”. In: *Level Set and PDE Based Reconstruction Methods in Imaging*. Cham: Springer International Publishing, 2013, pp. 1–70.

[Che+16] Z. Chen, M. Weyland, X. Sang, W. Xu, J. Dycus, J. LeBeau, A. D’Alfonso, L. Allen, and S. Findlay. “Quantitative atomic resolution elemental mapping via absolute-scale energy dispersive X-ray spectroscopy”. *Ultramicroscopy* 168 (2016), pp. 7–16.

[CL75] G. Cliff and G. W. Lorimer. “The quantitative analysis of thin specimens”. *Journal of Microscopy* 103.2 (1975), pp. 203–207.

[CM17] S. M. Collins and P. A. Midgley. “Progress and opportunities in EELS and EDS tomography”. *Ultramicroscopy* 180 (2017), pp. 133–141.

[Csi91] I. Csiszar. “Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems”. *The Annals of Statistics* 19.4 (1991), pp. 2032–2066.

[Dur+16] J. Duran, M. Moeller, C. Sbert, and D. Cremers. “Collaborative total variation: A general framework for vectorial TV models”. *SIAM Journal on Imaging Sciences* 9.1 (2016), pp. 116–151.

[ER07] R. Elena and A. S. Robert. “Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems”. *Mathematical Methods in the Applied Sciences* 30.13 (2007), pp. 1527–1544.

[ENH10] T. Elfving, T. Nikazad, and P. C. Hansen. “Semi-convergence and relaxation parameters for a class of SIRT algorithms”. *Electronic Transactions on Numerical Analysis* 37 (2010), pp. 321–336.

[Gen+12] A. Genç, H. Cheng, J. Winterstein, L. Pullan, and B. Freitag. “3D chemical mapping using tomography with an enhanced XEDS system”. *Microscopy and Analysis* 116 (2012), pp. 23–25.

[Gen+13] A. Genç, L. Kovarik, M. Gu, H. Cheng, P. Plachinda, L. Pullan, B. Freitag, and C. Wang. “XEDS STEM tomography for 3D chemical characterization of nanoscale particles”. *Ultramicroscopy* 131 (2013), pp. 24–32.

[GBH70] R. Gordon, R. Bender, and G. T. Herman. “Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography”. *Journal of Theoretical Biology* 29.3 (1970), pp. 471–481.

[Gor+12] B. Goris, W. van den Broek, K. J. Batenburg, H. Heidari Mezerji, and S. Bals. “Electron tomography based on a total variation minimization reconstruction technique”. *Ultramicroscopy* 113 (2012), pp. 120–130.

[Gor+16] B. Goris, M. Meledina, S. Turner, Z. Zhong, K. J. Batenburg, and S. Bals. “Three dimensional mapping of Fe dopants in ceria nanocrystals using direct spectroscopic electron tomography”. *Ultramicroscopy* 171 (2016), pp. 55–62.

[Gor+14] B. Goris, L. Polavarapu, S. Bals, G. van Tendeloo, and L. M. Liz-Marzán. "Monitoring galvanic replacement through three-dimensional morphological and chemical mapping ". *Nano Letters* 14.6 (2014), pp. 3220–3226.

[Gre90] P. J. Green. "Bayesian reconstructions from emission tomography data using a modified EM algorithm". *IEEE Transactions on Medical Imaging* 9.1 (1990), pp. 84–93.

[GB08] J. Gregor and T. Benson. "Computational analysis and improvement of SIRT". *IEEE Transactions on Medical Imaging* 27.7 (2008), pp. 918–924.

[Hab+14] G. Haberfehlner, A. Orthacker, M. Albu, J. Li, and G. Kothleitner. "Nanoscale voxel spectroscopy by simultaneous EELS and EDS tomography". *Nanoscale* 6 (2014), pp. 14563–14569.

[Her09] G. T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Second. Springer, 2009.

[HW16] T. Hohage and F. Werner. "Inverse problems with Poisson data: statistical regularization theory, applications and algorithms". *Inverse Problems* 32.9 (2016), p. 093001.

[Hol14] K. M. Holt. "Total nuclear variation and Jacobian extensions of total variation for vector fields". *IEEE Transactions on Image Processing* 23.9 (2014), pp. 3975–3989.

[Hsi+00] J. Hsieh, R. C. Molthen, C. a. Dawson, and R. H. Johnson. "An iterative approach to the beam hardening correction in cone beam CT". *Medical physics* 27.1 (2000), pp. 23–29.

[Jar+09] K. Jarausch, P. Thomas, D. N. Leonard, R. Twesten, and C. R. Booth. "Four-dimensional STEM-EELS: Enabling nano-scale chemical tomography". *Ultramicroscopy* 109.4 (2009), pp. 326–337.

[Jol02] I. Jolliffe. *Principal Component Analysis*. Springer-Verlag New York, 2002.

[KS88] A. C. Kak and M. Slaney. *Principles of computerized tomographic imaging*. IEEE press, 1988.

[Kis+08] C. Kisielowski, B. Freitag, M. Bischoff, H. van Lin, S. Lazar, G. Knippels, P. Tiemeijer, M. van der Stam, S. von Harrach, M. Stekelenburg, and et al. "Detection of single atoms and buried defects in three dimensions by aberration-corrected electron microscope with 0.5-Å information limit". *Microscopy and Microanalysis* 14.5 (2008), pp. 469–477.

[Kra+17] J. Kraxner, M. Schfer, O. Rschel, G. Kothleitner, G. Haberfehlner, M. Paller, and W. Grogger. "Quantitative EDXS: Influence of geometry on a four detector system". *Ultramicroscopy* 172 (2017), pp. 30–39.

[Küb+05] C. Kübel, A. Voigt, R. Schoenmakers, M. Otten, D. Su, T.-C. Lee, A. Carlsson, and J. Bradley. "Recent advances in electron tomography: TEM and HAADF-STEM tomography for materials science and semiconductor applications". *Microscopy and Microanalysis* 11.05 (2005), pp. 378–400.

[Lag+98] J. Lagarias, J. Reeds, H. Wright, and P. Wright. "Convergence properties of the Nelder–Mead simplex method in low dimensions". *Journal of Optimization* 9.1 (1998), pp. 112–117.

[LSP08] S. J. LaRoque, E. Y. Sidky, and X. Pan. "Accurate image reconstruction from few-view and limited-angle data in diffraction tomography". *Journal of the Optical Society of America* 25.7 (2008), pp. 1772–1782.

[LH74] C. L. Lawson and R. J. Hanson. *Solving least squares problems*. SIAM, 1974.

[Lep+13] K. Lepinay, F. Lorut, R. Pantel, and T. Epicier. "Chemical 3D tomography of 28nm high K metal gate transistor: STEM XEDS experimental method and results". *Micron* 47 (2013), pp. 43–49.

[Luc+13] G. Lucas, P. Burdet, M. Cantoni, and C. Hbert. "Multivariate statistical analysis as a tool for the segmentation of 3D spectral data". *Micron* 52-53 (2013), pp. 49–56.

[MHD18] K. E. MacArthur, M. Heggen, and R. E. Dunin-Borkowski. "Differentiating the structure of PtNi octahedral nanoparticles through combined ADF-EDX simulations". *Advanced Structural and Chemical Imaging* 4.1 (2018), p. 2.

[Mid+01] P. A. Midgley, M. Weyland, J. M. Thomas, and B. F. G. Johnson. "Z-Contrast tomography: a technique in three-dimensional nanostructural analysis based on Rutherford scattering". *Chemical Communications* (10 2001), pp. 907–908.

[MW03] P. A. Midgley and M. Weyland. "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography". *Ultramicroscopy* 96.3 (2003), pp. 413–431.

[Ots79] N. Otsu. "A threshold selection method from gray-level histograms". *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), pp. 62–66.

[PBS13] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers. "The ASTRA Tomography Toolbox". *13th International Conference on Computational and Mathematical Methods in Science and Engineering* (2013), pp. 1–7.

[Pen89] S. J. Pennycook. "Z-contrast STEM for materials science". *Ultramicroscopy* 30.1 (1989), pp. 58–69.

[Qiu+15] Y. Qiu, P. van Marcke, O. Richard, H. Bender, and W. Vandervorst. "3D imaging of Si FinFET devices by combined HAADF-STEM and EDS tomography". In: *Microscopy of Semiconductor Materials*. Cambridge UK. 2015.

[RL15] D. S. Rigie and P. J. La Rivière. "Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization". *Physics in Medicine & Biology* 60.5 (2015), p. 1741.

[SM12] Z. Saghi and P. A. Midgley. "Electron tomography in the (S)TEM: From nanoscale morphological analysis to 3D atomic imaging". *Annual Review of Materials Science* 42 (2012), pp. 59–79.

[Sag+07] Z. Saghi, X. Xu, Y. Peng, B. Inkson, and G. Möbus. "Three-dimensional chemical analysis of tungsten probes by energy dispersive X-ray nanotomography". *Applied Physics Letters* 91.25 (2007), p. 251906.

[Sán+12] 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, and L. M. Liz-Marzán. "Hydrophobic interactions modulate self-assembly of nanoparticles". *ACS Nano* 6.12 (2012), pp. 11059–11065.

[Sch+10] P. Schlossmacher, D. Klenov, B. Freitag, and H. Von Harrach. "Enhanced detection sensitivity with a new windowless XEDS system for AEM based on silicon drift detector technology". *Microscopy today* 18.04 (2010), pp. 14–20.

[SV82] L. Shepp and Y. Vardi. "Maximum likelihood reconstruction for emission tomography". *IEEE transactions on medical imaging* 1.2 (1982), pp. 113–122.

[SJP12] E. Y. Sidky, J. H. Jorgensen, and X. Pan. "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm". *Physics in Medicine & Biology* 57.10 (2012), pp. 3065–91.

[SKP06] E. Y. Sidky, C.-M. Kao, and X. Pan. "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT". *Journal of X-Ray Science and Technology* 14 (2006), pp. 119–139.

[Sla+14] T. J. Slater, P. H. C. Camargo, M. G. Burke, N. J. Zaluzec, and S. J. Haigh. "Understanding the limitations of the Super-X energy dispersive X-ray spectrometer as a function of specimen tilt angle for tomographic data acquisition in the S/TEM". *Journal of Physics: Conference Series* 522.1 (2014), p. 012025.

[Sla+16a] T. J. Slater, Y. Chen, G. Auton, N. Zaluzec, and S. J. Haigh. "X-ray Absorption Correction for Quantitative Scanning Transmission Electron Microscopic Energy-Dispersive X-Ray Spectroscopy of Spherical Nanoparticles". *Microscopy and Microanalysis* 22.2 (2016), pp. 440–447.

[Sla+16b] T. J. Slater, A. Janssen, P. H. Camargo, M. G. Burke, N. J. Zaluzec, and S. J. Haigh. "STEM-EDX tomography of bimetallic nanoparticles: A methodological investigation". *Ultramicroscopy* 162 (2016), pp. 61–73.

[Tre11] M. M. Treacy. "Z dependence of electron scattering by single atoms into annular dark-field detectors". *Microscopy and Microanalysis* 17.06 (2011), pp. 847–858.

[van+02] E. van de Casteele, D. van Dyck, J. Sijbers, and E. Raman. "An energy-based beam hardening model in tomography". *Physics in Medicine & Biology* 47.23 (2002), p. 4181.

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

[van+11] G. van Gompel, K. van Slambrouck, M. Defrise, K. J. Batenburg, J. de Mey, J. Sijbers, and J. Nuyts. "Iterative correction of beam hardening artifacts in CT". *Medical Physics* 38.S1 (2011), S36.

[Wal06] T. Walther. "A new experimental procedure to quantify annular dark field images in scanning transmission electron microscopy". *Journal of Microscopy* 221.2 (2006), pp. 137–144.

[Wan+04] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. "Wavelets for Image Image quality assessment: From error visibility to structural similarity". *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612.

[WW06] M. Watanabe and D. B. Williams. "The quantitative analysis of thin specimens: A review of progress from the Cliff-Lorimer to the new ζ-factor methods". *Journal of Microscopy* 221 (2006), pp. 89–109.

[WC16] D. B. Williams and C. B. Carter. *Transmission Electron Microscopy: Diffraction, Imaging, and Spectrometry*. Springer, 2016.

[Yed+12] L. Yedra, A. Eljarrat, R. Arenal, E. Pellicer, M. Cabo, A. Lopez-Ortega, M. Estrader, J. Sort, M. D. Baro, S. Estradé, et al. "EEL spectroscopic tomography: Towards a new dimension in nanomaterials analysis". *Ultramicroscopy* 122 (2012), pp. 12–18.

[Yed+14] L. Yedra, A. Eljarrat, J. M. Rebled, L. López-Conesa, N. Dix, F. Sánchez, S. Estradé, and F. Peiró. "EELS tomography in multiferroic nanocomposites: from spectrum images to the spectrum volume". *Nanoscale* 6.12 (2014), pp. 6646–6650.

[YF02] D. F. Yu and J. A. Fessler. "Edge-preserving tomographic reconstruction with nonlocal regularization". *IEEE Transactions on Medical Imaging* 21.2 (2002), pp. 159–173.

[ZMA10] H. Zaidi, M.-L. Montandon, and A. Alavi. "The clinical role of fusion imaging using PET, CT, and MR Imaging". *Magnetic Resonance Imaging Clinics of North America* 18.1 (2010), pp. 133–149.

[Zan+16a] D. Zanaga, T. Altantzis, L. Polavarapu, L. M. Liz-Marzán, B. Freitag, and S. Bals. "A new method for quantitative XEDS tomography of complex heteronanostructures". *Particle & Particle Systems Characterization* 33.7 (2016), pp. 396–403.

[Zan+16b] D. Zanaga, T. Altantzis, J. Sanctorum, B. Freitag, and S. Bals. "An alternative approach for ζ-factor measurement using pure element nanoparticles". *Ultramicroscopy* 164 (2016), pp. 11–16.

[Zho+18a] Z. Zhong, R. Aveyard, B. Rieger, S. Bals, W. J. Palenstijn, and K. J. Batenburg. "Automatic correction of nonlinear damping effects in HAADF-STEM tomography for nanomaterials of discrete compositions". *Ultramicroscopy* 184.Part B (2018), pp. 57–65.

[Zho+17] Z. Zhong, B. Goris, R. Schoenmakers, S. Bals, and K. J. Batenburg. "A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM". *Ultramicroscopy* 174 (2017), pp. 35–45.

[Zho+18b] Z. Zhong, W. J. Palenstijn, J. Adler, and K. J. Batenburg. "EDS tomographic reconstruction regularized by total nuclear variation joined with HAADF-STEM tomography". *Ultramicroscopy* 191 (2018), pp. 34–43.

[Zho+18c] Z. Zhong, W. J. Palenstijn, N. R. Viganò, and K. J. Batenburg. "Numerical methods for low-dose EDS tomography". *Ultramicroscopy* 194 (2018), pp. 133–142.

[ZPB16] X. Zhuge, W. J. Palenstijn, and 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), pp. 455–468.

List of publications

- Publications that are part of this thesis:

  [1] Z. Zhong, B. Goris, R. Schoenmakers, S. Bals, and K. J. Batenburg. "A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM". *Ultramicroscopy* 174 (2017), pp. 35–45.

  [2] Z. Zhong, R. Aveyard, B. Rieger, S. Bals, W. J. Palenstijn, and K. J. Batenburg. "Automatic correction of nonlinear damping effects in HAADF-STEM tomography for nanomaterials of discrete compositions". *Ultramicroscopy* 184.Part B (2018), pp. 57–65.

  [3] Z. Zhong, W. J. Palenstijn, J. Adler, and K. J. Batenburg. "EDS tomographic reconstruction regularized by total nuclear variation joined with HAADF-STEM tomography". *Ultramicroscopy* 191 (2018), pp. 34–43.

  [4] Z. Zhong, W. J. Palenstijn, N. R. Viganò, and K. J. Batenburg. "Numerical methods for low-dose EDS tomography". *Ultramicroscopy* 194 (2018), pp. 133–142.

- Publications that are not part of this thesis:

  [5] B. Goris, M. Meledina, S. Turner, Z. Zhong, K. J. Batenburg, and S. Bals. "Three dimensional mapping of Fe dopants in ceria nanocrystals using direct spectroscopic electron tomography". *Ultramicroscopy* 171 (2016), pp. 55–62.

  [6] R. Aveyard, Z. Zhong, K. J. Batenburg, and B. Rieger. "Optimizing experimental parameters for the projection requirement in HAADF-STEM tomography". *Ultramicroscopy* 177 (2017), pp. 84–90.

Samenvatting

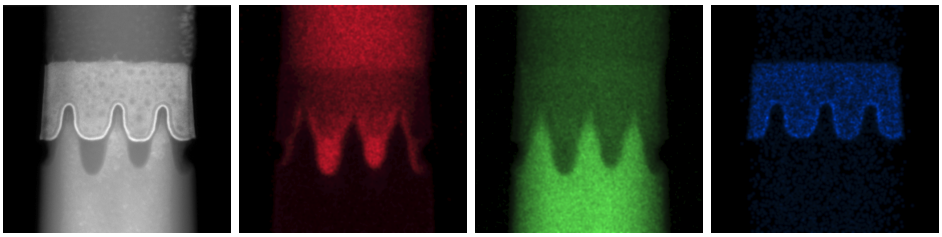
(*The English summary follows after the Dutch summary.*)

Tomografie is een niet-invasieve techniek waarmee beelden gemaakt kunnen worden van doorsnedes van een object. Deze doorsnedes worden berekend aan de hand van projectiebeelden die onder verschillende hoeken worden opgenomen. Deze techniek staat ook wel bekend als computertomografie (CT). Het principe achter tomografie kan toegepast worden op projectiebeelden die met verschillende beeldvormingstechnieken zijn opgenomen. Standaard CT, gebaseerd op Röntgenstraling, kan alleen structuren onderscheiden van tenminste honderd micrometer in grootte. In levenswetenschappen en materiaalwetenschappen is het van belang om structuren in beeld te brengen die slechts een paar atomen groot zijn. Denk hierbij bijvoorbeeld aan een computerchip, die structuren heeft van ongeveer 10 nanometer. Voor deze toepassingen wordt het principe van tomografie gecombineerd met elektronenmicroscopie (EM). Deze combinatie staat bekend als elektronentomografie (ET). EM maakt gebruikt van versnelde elektronen, en met deze techniek kunnen atomaire structuren in beeld gebracht worden.

Het onderzoek in dit proefschrift richt zich op de tomografische reconstructie op basis van projectiebeelden die afkomstig zijn van twee verschillende EM modaliteiten. De eerste modaliteit is *high angle annular dark field scanning transmission microscopy* (HAADF-STEM), en de tweede modaliteit is *energy-dispersive X-ray spectroscopy* (EDS). Figuur S1 toont voorbeelden van projectiebeelden opgenomen met respectievelijk HAADF-STEM en EDS. Tegenwoordig is HAADF-STEM een standaard modaliteit in EM, terwijl EDS soms gebruikt wordt als een aanvullende techniek om extra informatie te verzamelen over de chemische samenstelling van het onderzochte object. HAADF-STEM resulteert in beelden met een enkel kanaal die informatie van verschillende chemische elementen door elkaar mengt. Wanneer men gebruik maakt van standaard beeldvormingsmethodes, hebben HAADF-STEM beelden relatief lage ruisniveaus. In tegenstelling tot HAADF-STEM, leidt EDS tot een verzameling beelden, één voor elk van de afzonderlijke elementen. Deze EDS beelden hebben typisch hogere ruisniveaus. Het toepassen van tomografie op EDS afbeeldingen is uitdagender vergeleken met HAADF-STEM. Dit komt onder andere door de sterke Poissonruis, een beperkte hoeveelheid projectiehoeken, en een lange opnametijd.

Rekeninghoudend met de complementaire eigenschappen van de twee modaliteiten, stellen we in Hoofdstuk 2 een aanpak voor die elementspecifieke reconstructies uitvoert op basis van projectiebeelden afkomstig van zowel HAADF-STEM als EDS.

In deze aanpak worden de elementafhankelijke projectiebeelden opgenomen voor alle chemische elementen die in het object aanwezig zijn. In de praktijk kan het moeilijk zijn om al de hiervoor benodigde beelden op te nemen, door beperkingen van EDS opnametechnieken. Daarom stellen wij in Hoofdstuk 3 een

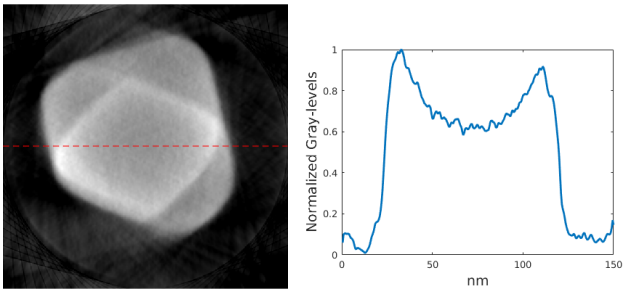


(a) HAADF-STEM (b) O (c) Si (d) Ti

Figuur S1: (a): Een HAADF-STEM projectiebeeld van een halfgeleiderobject. (b) – (d): EDS projectiebeelden voor verschillende elementen uit hetzelfde object. [English: (a): HAADF-STEM projection image for a semiconductor sample. (b) - (d): EDS projection images for some chemical elements in the same sample.]

andere aanpak voor die projectiebeelden van de twee modaliteiten combineert. Deze techniek verbetert de elementspecifieke reconstructie, door het aanmoedigen van gedeelde randlocaties met de reconstructie gemaakt op basis van HAADF-STEM data.

Een sleutelprobleem bij het inbrengen van dergelijke informatie die *a priori* over het object bekend is, is hoe kan worden bepaald welke methode goed werkt voor een gegeven object. In Hoofdstuk 4 presenteren we een raamwerk voor het construeren van geavanceerde reconstructiemethodes, in de vorm van een *recept* dat rekening houdt met eigenschappen van het object en het scanproces. Dit leidt tot een verzameling richtlijnen voor het bepalen van een geschikte methode, afhankelijk van experimentele condities en het af te beelden object zelf.



(a) reconstruction (b) line profile

Figuur S2: Bekervormige artefacten: (a) een gereconstrueerde doorsnede van een nanodeeltje. (b) intensiteiten overeenkomstig met de stippellijn in (a). De intensiteiten in het midden worden onderschat. [English: Cupping artifacts: (a) a slice of the reconstruction for a nanoparticle. (b) image intensities on the dashed line in (a). The image intensities in the center are underestimated.]

Het toepassen van ET op relatief dikke microscopie-samples is uitdagend, omdat in dit geval niet-lineaire effecten sterke invloed hebben op de HAADF-STEM data. Een consequentie is dat bekervormige artefacten zichtbaar zijn in de re-

constructiebeelden, wat er toe leidt dat intensiteiten in het midden van het reconstructiebeeld onderschat worden. Figuur S2 toont een dergelijk artefact in de reconstructie van een nanodeeltje. In Hoofdstuk 5 stellen we een algoritme voor dat dergelijke niet-lineaire effecten automatisch corrigeert, zodat de HAADF-STEM data gelineariseerd kan worden. Hierna kan deze data gebruikt worden voor tomografische reconstructie op basis van HAADF-STEM + EDS.

Samenvattend wordt in dit proefschift een aantal technieken voorgesteld die een aanzienlijke verbetering kunnen geven ten opzicht van de gangbare reconstructiemethoden. De numerieke methoden kunnen toegepast worden op HAADF en EDS data die volgens gangbare technieken zijn opgenomen. Het is daarnaast mogelijk om op basis van de voorgestelde technieken de benodigde opnametijd van EDS data te verminderen. Een mogelijke toepassing van tomografie op basis van gecombineerde HAADF-STEM en EDS data, is het in beeld brengen van structuren in driedimensionale halfgeleidermaterialen. Dit maakt het onder andere mogelijk fabricagefouten te identificeren. Dit kan de verdere ontwikkeling van geavanceerde geïntegreerde schakelingen faciliteren.

Summary

Tomography is a non-destructive technique for imaging slices (sectional images) of an object. The slices are computed from a series of projection images taken from different angles. This computation is known as tomographic reconstruction. The principle of tomography can be applied to different techniques for acquiring the projection images. For instance, the widely-used computed tomography (CT) for medical diagnostics is based on X-ray imaging. Standard X-ray CT can only resolve structures that are at least about one hundred micrometers in size. In the life sciences and materials science, the size of critical structures can sometimes be as small as a few atoms. For instance, the size of structures in computer chips is nowadays around 10 nanometers. In these fields, the principle of tomography is combined with electron microscopy (EM), a combination which is known as electron tomography (ET). EM uses accelerated electrons for imaging and can resolve structures at the atomic scale.

The research in this thesis is focused on tomographic reconstruction based on two imaging modalities in EM. The first modality is *high angle annular dark field scanning transmission microscopy* (HAADF-STEM), and the second modality is *energy-dispersive X-ray spectroscopy* (EDS). Figure S1 shows examples of projection images acquired by HAADF-STEM and EDS. While HAADF-STEM is a standard modality in EM nowadays, EDS is sometimes used as a supplementary technique to resolve more chemical information. HAADF-STEM yields a single-channel image that mixes the information for all chemical elements. Using

a standard imaging scheme, HAADF-STEM images have relatively low noise levels. EDS, on the other hand, yields multiple element-specific images. However, the resulting projection images have much higher noise levels. Tomography based on EDS is more challenging compared to HAADF-STEM, due to the strong Poisson noise, the limited number of projection angles and the long data acquisition time.

Considering the complementary properties of the two modalities, in Chapter 2 we propose an approach to perform element-specific reconstructions from HAADF-STEM and EDS tomographic data that are simultaneously acquired. Using this technique, it is possible to obtain element-specific reconstructions with better image quality compared to only using EDS data.

In this approach, element-specific projection images are required for all chemical elements present in the sample, which may be not possible in practice due to limitations of the EDS imaging technique. In Chapter 3, we propose a different approach for combining HAADF-STEM and EDS, which improves the element-specific reconstruction from EDS data by encouraging the reconstructed images to have common edge locations with the reconstruction from HAADF-STEM data.

A key problem with incorporating such prior knowledge is to know which method works well on which type of sample. In Chapter 4 we present a framework for constructing advanced reconstruction methods as a "recipe" that is tailored to the particular sample. We present guidelines for determining which method should be used depending on the experimental conditions and sample properties.

Applying ET on samples with large thicknesses is challenging due to the dominant nonlinear effects in HAADF-STEM data. As a consequence, cupping artifacts can appear in the reconstructed tomographic images, which means that the image intensities being underestimated in the center of reconstruction. Figure S2 shows the cupping artifact present in the reconstructed image for a nanoparticle. In Chapter 5, we propose an algorithm for correcting these nonlinear effects automatically, so that the HAADF-STEM data can be linearized and subsequently used for HAADF-STEM + EDS tomographic reconstruction.

Overall, several approaches are given in this thesis for improving the image quality of element-specific tomographic reconstructions. The numerical methods can be applied to HAADF and EDS data acquired in a conventional manner using the existing electron microscopic technique. It is also possible to design novel data acquisition-processing pipelines based on them, to reduce the time for EDS data acquisition. A potential application of HAADF-STEM + EDS tomography is imaging structures and faults in 3D semiconductor materials, which can facilitate the development of advanced integrated circuits.

Curriculum Vitae

Zhichao Zhong was born in Guangdong, China in 1987. From 2008, he studied Optical Information Science and Technology at Sun Yat-sen University, China. In 2012, he obtained his bachelor's degree. From 2012, he studied Optics and Photonics at Karlsruhe Institute of Technology (Germany) and Aix-Marseille University (France) as a student of the Erasmus Mundus Master program *Europhotonics*. In 2014, he obtained master's degrees from the two universities after writing a master thesis entitled "Stair Detection and Estimation From Stereo Vision". After the master study, he joined the Centrum Wiskunde & Informatica (CWI) in Amsterdam as a PhD candidate under the supervision of Prof. dr. K.J. Batenburg (Leiden University). At CWI he worked on developing advanced tomographic reconstruction techniques based on electron microscopy. He closely collaborated with researchers from University of Antwerp, TU Delft, imec, and Thermo Fisher Scientific. In the last year, he conducted an internship at Thermo Fisher Scientific. In November 2018, he finished his PhD research. Now he works as a data scientist at Wehkamp B. V. in the Netherlands.

Acknowledgments

This thesis could not have been completed without the support of many people. Here I would like to acknowledge some of them in particular.

First and foremost, I would like to express my gratitude to my supervisor Joost Batenburg. From you I learned not only how to do research but also how to be a researcher. Thanks to your useful comments I learned to write scientific articles; thanks to many discussions with you I learned to look at problems from a bigger picture; and thanks to your support in networking, I learned to collaborate with people.

I would like to thank all my colleagues at CWI. I have very happy memories about barbecues, summer trips and gaming nights with you. I especially thank my officemates Allard Hendriksen, Rien Lagerwerf and Jan-Willem Buurlage. Your jokes, whining and genuine concerns made everyday in the office lovely. Also thank you for teaching me table tennis. I thank Willem Jan Palenstijn for answering countless questions and helping to polish my writings. I thank Nicola Vigano and Xiaodong Zhuge for the helpful supervision. I also would like to thank the other colleagues in the Computational Imaging group: Adriaan Graas, Alexander Kostenko, Daniël Pelt, Folkert Bleichrodt, Francien Bossema, Felix Lucka, Giulia Colacicco, Henri Der Sarkissian, Holger Kohr, Maureen van Eijnatten, Mathé Zeegers, Poulami Ganguly, Rob Bisseling, Robert van Liere, Richard Schoonhoven and Sophia Coban. I am very proud that you are making tomography fun and exciting. I also want to express my gratitude to the supporting staff at CWI, especially Nada Mitrovic and Duda Tepsic for your excellent works.

I also would like to thank people outside CWI who have helped me during the research. I thank Bernd Rieger, Richard Aveyard and Yan Guo for the intriguing discussions and the useful comments. I thank Bart Goris and Sara Bals for sharing the experimental data and for hosting me at EMAT. I thank Remco Schoenmakers for comments from the industrial perspective and for arranging the internship. I also thank Hugo Bender, Jonas Adler, Johannes Schmidt-Hieber, Tristan van Leeuwen and the many other people that I have had the pleasure to interact with over the past four years.

This thesis would not have been possible without the support of my family. 感谢我所有的家人朋友，特别是我的父母，哥哥和姐姐。感谢你们支持追求我的理想。虽然我们相隔千山万水，每个周末跟你们视频的时间都是我最温暖的时候，因为你们一直都是世界上最关心我的人。希望我的小小成就能让你们感到自豪。愿天堂的奶奶可以看看我拿到博士学位的时刻，多么希望你能够亲眼看到这本论文，再听我讲讲这背后的故事。

Lastly, I would like to give special thanks to Xinyi Zhao. Like many PhD students, for a time I felt very confused about myself. But that has never happened again since I met you. You always care about my feelings, understand my thoughts and support my decisions. Thanks to you, I could write this thesis with faith in myself. I hope we will make many more achievements together in the future.