Image quality assessment and image fusion for electron tomography

Guo, Yan

DOI
10.4233/uuid:ef8223e4-5c89-4960-8d6e-0603fd514368

Publication date
2020

Document Version
Final published version

Citation (APA)
Guo, Y. (2020). *Image quality assessment and image fusion for electron tomography*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:ef8223e4-5c89-4960-8d6e-0603fd514368

IMAGE QUALITY ASSESSMENT AND IMAGE FUSION FOR ELECTRON TOMOGRAPHY

IMAGE QUALITY ASSESSMENT AND IMAGE FUSION FOR ELECTRON TOMOGRAPHY

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
woensdag 24 juni 2020 om 10:00 uur

door

Yan GUO

Master of Science in Communication Engineering,
RWTH Aachen University, Aachen, Germany,
geboren te Beijing, China.

Dit proefschrift is goedgekeurd door de

promoter: Prof. dr. ir. B. Rieger

Samenstelling promotiecommissie bestaat uit:

Rector Magnificus,	voorzitter
Prof. dr. ir. B. Rieger,	Technische Universiteit Delft, promoter
Onafhankelijke leden:	
Prof. dr. ir. M.J.T. Reinders,	Technische Universiteit Delft
Prof. dr. I. Nyström,	Uppsala Universitet, Zweden
Prof. dr. K.J. Batenburg,	Universiteit Leiden
Dr. R. Schoenmakers,	Thermo Fisher Scientific, Eindhoven
Dr. F.M. Vos,	Technische Universiteit Delft
Prof. dr. S. Stallinga,	Technische Universiteit Delft, reservelid

The work in this dissertation was conducted at the Computational Imaging Group, Faculty of Applied Sciences, Delft University of Technology.

Keywords: image quality assessment, multimodal image fusion, electron tomography, HAADF-STEM, X-ray spectroscopy, EDS, nanomaterials

Printed by: Ridderprint BV

ISBN 000-00-0000-000-0

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

CONTENTS

SUMMARY

Electron tomography is a powerful tool in materials science to characterize nanostructures in three dimensions (3D). In scanning transmission electron microscopy (STEM), the sample under study is exposed to a focused electron beam and tilted to obtain two-dimensional (2D) projections at different angles; many imaging modes are available such as high-angle annular dark-field (HAADF). In tomography, the collection of projections is called a tilt-series, from which we can reconstruct a 3D image that represents the sample. While HAADF tomography can clearly reveal the inner structure of the sample, it cannot directly provide compositional information. To better understand nanomaterials with more types of elements, spectral imaging techniques like energy dispersive X-ray spectroscopy (EDS) must be pursued. EDS tomography, however, is currently hampered by slow data acquisition, resulting in a small number of elemental maps with low signal-to-noise ratio (SNR).

Electron tomography, especially EDS tomography, is an ill-posed inverse problem whose solution is not stable and unique. Although advanced reconstruction techniques may yield a more accurate result by incorporating prior knowledge, they also involve fine-tuning parameters that highly influence the reconstruction quality. Furthermore, while great efforts have been dedicated to developing tomography techniques for image enhancement, directly combining reconstruction volumes at hand has still not been widely considered to the best of our knowledge.

In this thesis, we first pave the way for automatic no-reference single-parameter selection to prevent the tedious and time-consuming hand-tuning for advanced reconstruction techniques. To begin with, we propose an image quality metric, which quantifies the reconstruction quality through cross-atomic contamination and noise, to determine the optimal weighting factor for HAADF-EDS bimodal tomography. Then, we move on to find the optimal regularization parameter for two regularized reconstruction algorithms: total variation (TV) and total nuclear variation (TNV). The regularization strength is measured by the sum of the Gaussian gradient magnitude for TV and the oriented structure strength, which is described by the highest two responses in orientation space, for TNV. Simulation and/or experimental results of nanoparticles show that all three metrics can replace the full-reference metrics in practice. Therefore, they can be used by non-experts if a reference is unaccessible.

Then, we present a HAADF-EDS cross-modal fusion framework that has no fine-tuning parameter to simultaneously achieve high SNR and high elemental specificity. Based on simulated and experimental datasets of semiconductor devices that have more complex structures and more types of elements, we demonstrate that our algorithm can produce continuous edges, homogeneous foreground and clean background in its element-specific reconstructions. More importantly, it stably delivers results with high fidelity even for limited and noisy EDS datasets. This property is highly desired in the semiconductor industry where acquisition time and sample damage are essential.

SAMENVATTING

Elektronen tomografie is een krachtig techniek in materiaalkunde om nanostructuren te karakteriseren in drie dimensies (3D). In scannende transmissie elektronen microscopie (STEM) wordt de sample blootgesteld aan een gefocuste elektronen bundel en gekanteld om twee dimensionele (2D) projecties te verkrijgen onder verschillende hoeken; er zijn veel verschillende beeldvorming modi beschikbaar zoals high-angle annular dark-field (HAADF). In tomografie wordt de verzameling van projecties een kanteling-serie genoemd, waarvan we een representatieve 3D afbeelding van de sample kunnen reconstrueren. Hoewel HAADF tomografie duidelijk de innerlijke structuur van de sample weergeeft, is het niet mogelijk om direct informatie over de atomaire compositie te verkrijgen. Om een beter begrip van nanomaterialen te krijgen met meer verschillende elementen zijn spectrale beeldvormings technieken zoals energy dispersive X-ray spectroscopy (EDS) nodig. EDS tomografie wordt echter momenteel belemmerd door trage data acquisitie, dit resulteert in een lage hoeveelheid afbeeldingen per element met een lage signaal ruis verhouding (SNR).

Elektronen tomografie, en vooral EDS tomografie, is een ondergedetermineerd invers probleem waarvan de oplossing niet stabiel en uniek is. Alhoewel geavanceerde reconstructie technieken een meer accurate resultaat kunnen opleveren door voorkennis te gebruiken, omvatten ze ook fine-tuning parameters die een grote invloed hebben op de reconstructie kwaliteit. Bovendien, ondanks dat er grote inspanningen zijn geleverd voor het ontwikkelen van tomografische technieken voor beeldverbetering, is het direct combineren van reconstructievolumes nog niet eerder gedaan, naar ons beste weten.

In dit proefschrift ontwikkelen we eerst een methode voor automatische geen-referentie enkele-parameter selectie om de langdradig en tijdrovende taak van handmatig parameter tuning te voorkomen voor geavanceerde reconstructie technieken. Om te beginnen stellen we een beeld kwaliteit metriek voor, deze kwantificeert de reconstructie kwaliteit door middel van kruis-atomisch contaminatie en ruis, om te bepalen wat de optimale wegingsfactoren zijn voor HAADF-EDS bimodale tomografie. Waarna we de optimale regularisatie parameter proberen te vinden voor twee geregulariseerde reconstructie algoritmes; totale variatie (TV) en totale nucleaire variatie (TNV). De regularisatie sterkte wordt gemeten door de sommatie van de Gausische gradiënt magnitude voor TV en de georiënteerde structuur sterkte, welke worden beschreven door de hoogste twee reacties in oriëntatie ruimte. Simulatie en/of experimentele resultaten van nanodeeltjes laten zien dat in de praktijk alle drie de metrieken de volle-referentie metrieken kunnen vervangen. Daardoor kunnen ze gebruikt worden door niet-experts als een referentie ontoegankelijk is.

Daarna presenteren we een HAADF-EDS kruis-modaal fusie geraamte, dat geen afstembare parameter heeft, om simultaan een hoge SNR en hoge element specificiteit te behalen. Gebaseerd op gesimuleerde en experimentele datasets van halfgeleider apparaten die meer complexe structuren en meer verschillende elementen hebben demon-

streren we dat onze algoritme continue randen, homogene voorgrond en een schone achtergrond kan produceren in element-specifieke reconstructies. Belangrijker is dat het stabiele resultaten biedt met een hoge betrouwbaarheid, zelfs voor gelimiteerde en ruizige EDS datasets. Deze kenmerken zijn zeer gewenst in de halfgeleider industrie, waar acquisitietijd en sample schade essentieel zijn.

1

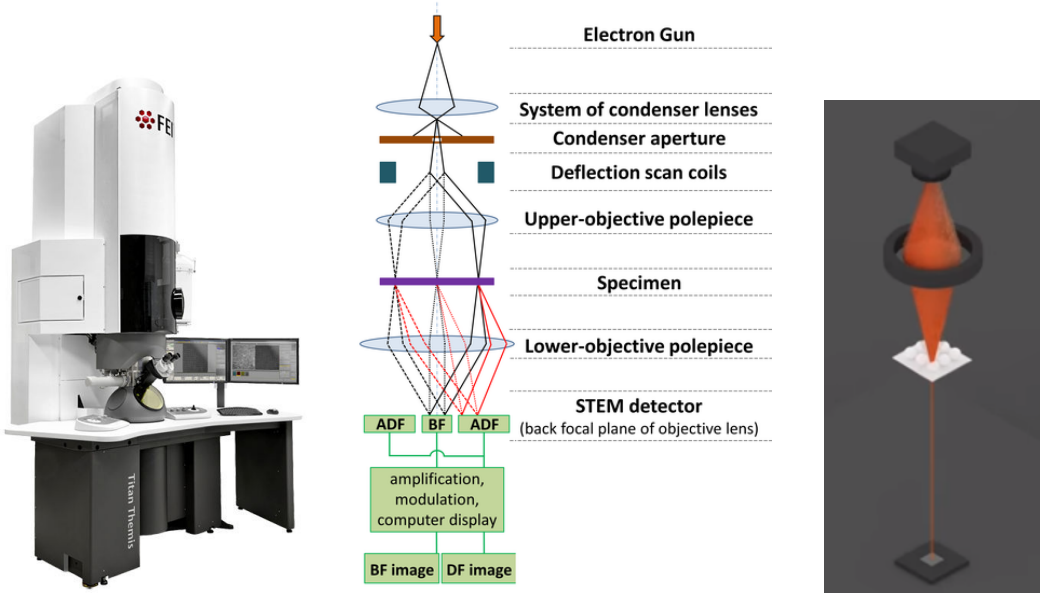
INTRODUCTION

1

The PhD research conducted in this thesis focuses on developing image processing techniques to characterize hard nanomaterials, such as heterogeneous catalysts and semiconductor devices. In this chapter, we first introduce scanning transmission electron microscopy, which is a popular two-dimensional (2D) imaging technique in materials science. Then, we move on to its three-dimensional (3D) extension: electron tomography. Finally, we state the objectives of the PhD research and give an overview of the thesis.

1.1. INSIDE SCANNING TRANSMISSION ELECTRON MICROSCOPE

Scanning transmission electron microscopy (STEM[1]) enables materials scientists to analyze a wide range of inorganic specimens on the nanoscale. As shown in Fig. 1.1, a typical STEM is a conventional transmission electron microscope (TEM) equipped with additional deflection scan coils and detectors [1]. The scan coils help to focus the electron beam to an atomic-sized probe where the beam is parallel to the optical axis in a raster scan; hence, STEM imaging can be considered as a serial recording using parallel beams. The detectors form images in an annular dark-field (ADF) or a bright-field (BF) mode, depending on the scattering angle of the transmitted electron beam.



(a) (b) (c)

Figure 1.1: (a) A FEI Titan S/TEM, from [2]. (b) Schematic of a STEM and (c) STEM mode, adapted from wikipedia.org. Scan coils focus the electron beam to an atomic-sized probe, and detectors form images.

The ADF detector collects scattered electrons. Moreover, one can use an ADF detector with a sufficiently large inner angle (> 50 mrad $\approx 3^\circ$ [1]) to collect only the incoherently scattered electrons. Images formed by such a high-angle annular dark-field (HAADF) detector are also called Z-contrast images, because the intensity of HAADF-STEM approximately varies with the atomic number Z of a chemical element as $\sim Z^n$.

[1] In this thesis, we use the same acronym or initial to denote both the technique (microscopy) and the instrument (microscope).

Theoretically, $n = 2$ [3] but practically $n = 1.6 - 1.9$, because the Coulomb potential of the bare nucleus is screened by the electron cloud [4]; the exact value of n also depends on the inner and outer angles of the detector, see [4, Fig. 6]. Since our eyes can detect intensity changes larger than 5% – 10% [1, Chapter 22], HAADF images can yield contrast in samples that consist of high- and low-Z elements (e.g., platinum nanoparticles supported on a carbon grid [5], $Z_{\mathtt{Pt}} = 78$, $Z_{\mathtt{C}} = 6$). However, when the atomic numbers of these elements become closer (e.g., nickel and manganese in lithium ion batteries [6], $Z_{\mathtt{Ni}} = 28$, $Z_{\mathtt{Mn}} = 25$), HAADF images may no longer be able to discriminate one element from another.

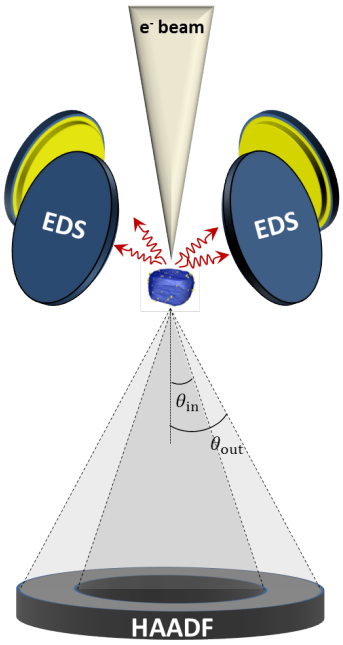


Figure 1.2: Schematic of simultaneous HAADF and EDS imaging in STEM. A focused electron beam scans over the sample. A HAADF detector with an inner angle $\theta_{\mathtt{in}}$ (> 50 mrad) and outer angle $\theta_{\mathtt{out}}$ collects highly scattered electrons; four symmetrically deployed EDS detectors collect emitted characteristic X-rays.

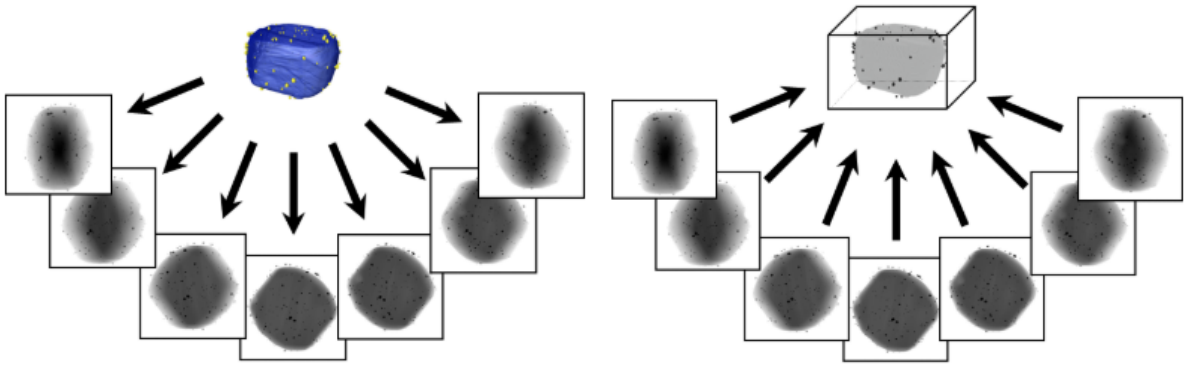
To better understand nanomaterials with more complex compositions, STEM can be further coupled with two spectroscopic techniques: energy dispersive X-ray spectroscopy (EDS) and electron energy loss spectroscopy (EELS). The former uses an X-ray spectrometer to detect characteristic X-rays emitted by the atoms in the sample upon electron ionization [6]; the latter uses an electron spectrometer to measure the amount of energy that electrons lose after inelastic scattering [7]. Although both techniques are adopted to analyze chemical compositions, EDS is more suitable for heavy elements with high Z because it has a more constant (and higher) peak-to-background ratio than EELS [6, 8].

Throughout this thesis, we focus on HAADF and EDS because they are more commonly employed to study hard nanomaterials [6, 9–11]. These two imaging modalities can be simultaneously acquired in STEM, see Fig. 1.2. A HAADF detector collects the highly scattered electrons, and four symmetrically positioned EDS detectors collect the emitted characteristic X-rays [6]. To improve collection efficiency, these individual spectral images are later combined into one spectrum, from which the final EDS elemental

maps are extracted by integrating the number of X-ray counts (above background) over a certain energy range [12]. HAADF and EDS provide complementary information of the sample. The former characterizes the structure with high signal-to-noise ratio (SNR) but little chemical information, and the latter, conversely, reveals compositions with high elemental specificity but low SNR. While the HAADF projections and EDS maps can reveal valuable structural and compositional information, they lack sensitivity in the third dimension [6]. Therefore, electron tomography must be pursued to fully analyze (thicker) nanomaterials with more complex morphology and chemistry [3].

1.2. ELECTRON TOMOGRAPHY FOR THREE-DIMENSIONAL CHARACTERIZATION

As a powerful tool in materials science for 3D characterization, electron tomography involves many steps between (the first) sample preparation and (the last) data interpretation [13]. In this section, we mainly focus on data acquisition and tomographic reconstruction, see Fig. 1.3. During data acquisition, the sample is tilted to obtain a series of 2D HAADF projections (and/or EDS maps) at different angles. With tomographic reconstruction, we can compute a 3D image that represents the sample from the collection of (aligned) projections, namely a tilt-series.



(a) (b)

Figure 1.3: Simplified schematic of electron tomography, adapted from [14]. During (a) data acquisition, a 3D sample is tilted to obtain 2D projections at different angles. With (b) tomographic reconstruction, a 3D image that represents the sample is computed from the collection of projections.

1.2.1. DATA ACQUISITION

In STEM, the sample is placed in a specialized tomography holder whose type determines the maximum tilt angle α [15]. For instance, a traditional single-axis holder limits α to $\sim \pm 70^\circ$ due to mechanical constraints. Since information beyond α is missing, reconstructions are elongated in the direction perpendicular to the tilt axis; this problem is termed “missing wedge" [15]. Compared to the single-axis holder, an improved dual-axis holder has a second tilt axis orthogonal to the first [15]; therefore, one can reduce the “missing wedge" to a “missing pyramid" by recording a second (orthogonal) tilt-series, see Fig. 1.4. Moreover, 360° data acquisition can be achieved through on-axis

rotation tomography holders and pillar-shaped samples [16]; these advanced rotation holders will eventually replace the traditional ones to avoid losing any information. Fig.

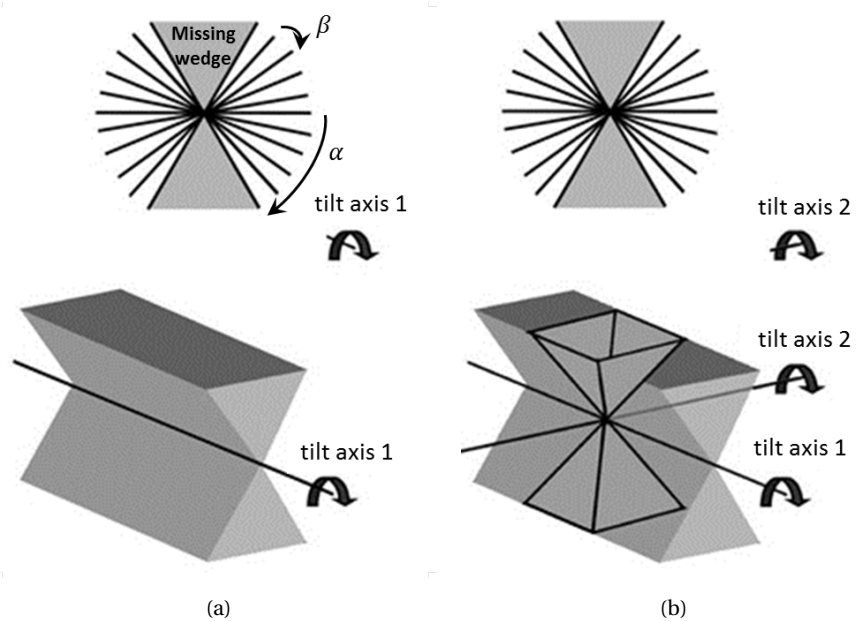


Figure 1.4: Graphical illustration of data sampling in Fourier space, adapted from [17]. α is the maximum tilt angle, and β is the increment between consecutive projections. The “missing wedge" in (a) single-axis tomography can be reduced to a “missing pyramid" in (b) dual-axis tomography.

1.5 shows examples of the three aforementioned holders.

Figure 1.5: Examples of tomography holders, adapted from [18]: (a) single-axis with missing wedge, (b) dual-axis with missing pyramid, and (c) on-axis that allows 360° rotation.

The total amount of electron dose delivered to the sample should be as low as possible to avoid damaging the sample, even though a high value is preferred to increase the SNR of projections. For the simultaneous HAADF and EDS imaging, the overall electron dose is usually limited to $\sim 10^9$ e/nm^2 [6, 9, 10, 19]. Given a fixed electron dose, one can either increase the number of tilts with a small increment β yet only collect noisy projections, or increase the signal at each angle with a large β yet limit the number of projections. Both might be problematic for an accurate alignment in the subsequent step. Furthermore, a small β may take longer acquisition time, because after each tilt one needs to readjust the sample, refocus and wait for the sample to stop drifting; a large β may reduce the reconstruction quality, because data might be too insufficiently sampled in the Fourier space to achieve accuracy. As a rule of thumb, $\beta \in [1^\circ, 10^\circ]$ in the

conventional step-by-step acquisition process [6, 10, 13, 19, 20]. Note that an alternative continuous approach has also gained attention in TEM tomography to reduce time [21–23], in which the sample is continuously rotated during data acquisition and the electrons are detected by an advanced direct (electron) detection camera [22, 23]. However, this continuous tilting scheme is not considered in this thesis where we focus on STEM as the method of choice to enable EDS spectroscopic mapping.

Mathematically, tomographic reconstruction implies that each pixel in the tilt-series is an integral (or a sum) through the sample [24]. However, such linear requirement may not be satisfied in STEM, because the resulting signal not only depends on the nuclei (or say, the atomic number Z) but also on the form of the beam reaching a voxel [25]. Fortunately, in practice if the deviation from linearity is known, one only needs to guarantee that the detected signal is a monotonic function with some physical properties of the sample and use the deviation information to correct the nonlinearity in post-processing [26]. In the rest of this section, we will describe when the HAADF projections and EDS maps fulfill the projection requirement, when they fail, and what to do for the correction.

Let us assume that the sample has thickness t, (homogeneous) density ρ, and a number of E different chemical elements making up an element set $\mathscr{E}$. Moreover, it is thin enough ($t < 100$ nm) to assume that most electrons either undergo a single-scattering event or are not scattered [1, Chapter 3]. Let I_0 be the total electron dose. The number of electrons directed and scattered to low angles $< 3^\circ$ is [1]

$$I = I_0 \cdot e^{-N\sigma t}, \tag{1.1}$$

in which N is the number of atoms per unit volume and σ is the single-atom scattering cross section [1, Eq. (22.6)]. Assuming that electrons rarely scatter to an angle higher than the outer angle of the HAADF detector (i.e., $\theta_{\mathtt{out}}$ in Fig. 1.2), the intensity of HAADF projections is

$$I^h = I_0 - I = I_0 \cdot (1 - e^{-N\sigma t}), \tag{1.2}$$

which, for small $N\sigma t < 1$, is

$$I^h \approx I_0 \cdot N\sigma t. \tag{1.3}$$

While Eq. (1.3) shows a linear relationship between I^h and $N\sigma t$, the HAADF signal contains nonlinearities. For instance, the contrast of light elements (e.g., crystalline silicon, $Z_{\mathtt{Si}} = 14$) may be substantially enhanced in zone-axis projections due to electron channeling [27]; in practice, these projections are excluded for reconstruction to avoid the channeling effect [28]. Moreover, the detected signal at greater thickness may be damped, because multiple scattered electrons may not fall onto the HAADF detector. The damping effect and associated cupping artifacts in the reconstruction can be automatically corrected by an iterative linearization scheme [29]. In [25], Aveyard et al. demonstrated how the microscope parameters (e.g., accelerating voltage, detector geometry) influence the linear relationship and explored the optimal experimental design for a broad range of nanomaterials with different types, atomic numbers and thickness.

For a thin sample where X-ray absorption and fluorescence are negligible (i.e., thin-foil criterion [1]), the intensity of EDS maps for element $e \in \mathscr{E}$ is [12]

$$I^e = I_0 \cdot \frac{C^e}{\zeta^e \rho t}, \tag{1.4}$$

in which I_0 is the total electron dose, C^e is the weight fraction of e in the sample, and ρt is the mass-thickness. ζ^e is a proportional factor connecting I^e to C^e and ρt, which is dependent on the ionization cross-section, the fluorescence yield, the relative transition probability, the atomic weight, the detector collection-angle, and the detector efficiency [12, Eq. (5)]. Since ζ^e is independent of I_0, C^e and ρt, it can be experimentally determined by measuring the X-ray intensities of standard thin-films with known compositions and thickness, namely the pure-element thin-film standards [12].

Similar to I^h, the above I^e is also sensitive to nonlinearities. For example, the integrated EDS signal varies with the tilt angle while using a single-axis holder, because X-rays emitted from the sample are partially blocked by the penumbra of the holder (especially at lower angles) and hence “shadowed" from the detector [30]. To make the total signal of each map the same, one can normalize the intensities of EDS maps or manually adjust the acquisition time at each angle [10]. In addition, Zanaga et al. used HAADF projections as a supplement to compensate for the shadowing effect, thereby enabling a more reliable EDS quantification [11]. Linearity is also invalid due to X-ray absorption, which arises if one or more of the characteristic X-rays has an energy less than 3 keV, or if the sample is too thick to validate the thin-foil criterion [1]. This can be corrected in many ways [12, 31, 32], such as the conventional Cliff-Lorimer method or a newly proposed ζ-factor method.

1.2.2. TOMOGRAPHIC RECONSTRUCTION

Before reconstruction, one needs to align the acquired tilt-series, that is, correct the lateral shift and rotation among HAADF projections and/or EDS maps [13]. A common method to this end is image correlation, which uses a cross-correlation function to match the patterns within images. Since this method does not need fiducial markers, it can be applied to various samples; however, it may accumulate errors if the correlation function exhibits a broad peak or high background noise [13].

After data alignment, let us move on to the mathematics and prior art on tomographic reconstruction. Note that we only use a 2D object and its one-dimensional (1D) parallel beam projections to explain the mathematical basis. In fact, if we scan a 3D sample using parallel beam and rotate it on a single axis such as z-axis, the projection data of its xy-slices are mutually independent. Consequently, the 2D reconstruction scheme can be easily extended to 3D by stacking a set of xy-slices at different heights along the z-axis [24].

Theoretically, a parallel projection is a collection of (parallel) line integrals [24]. Denote the object as a continuous function $f(x,y)$ in a xy-Cartesian coordinate system centered at the rotation axis, see Fig. 1.6. Given an arbitrary line $L(\theta,t)$: $t = x\cos\theta + y\sin\theta$ in the system, the line integral $P_\theta(t)$ of $f(x,y)$ for a constant θ is

$$P_\theta(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - t)\,dx\,dy, \tag{1.5}$$

which is also known as the Radon transform.

Practically, signals are detected in a discrete manner. To formulate this, we discretize $f(x,y)$ into a total number of N pixels and assume that each has a constant intensity value x_n, $n = 1,\cdots,N$, see Fig. 1.7. Since now a ray is defined as a “fat" line running

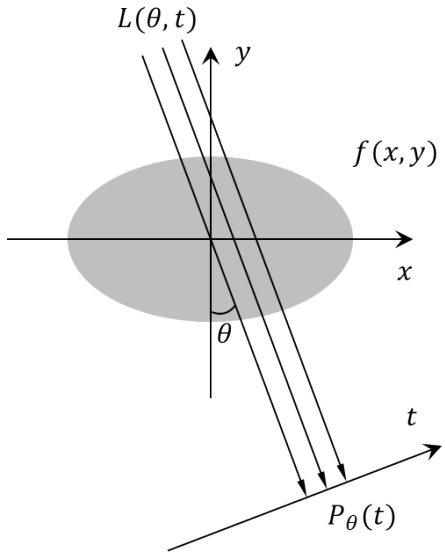


Figure 1.6: Schematic of the continuous forward projection. Given a 2D object $f(x, y)$, its forward projection for an arbitrary line $L(\theta, t)$ is modeled as the line integral $P_\theta(t)$ of f along the line, i.e., the Radon transform.

through the xy-plane, the (continuous) line integral becomes a (discrete) ray-sum [24]:

$$p_m = \sum_{n=1}^{N} w_{mn} x_n, \quad m = 1, \cdots, M, \tag{1.6}$$

in which M is the total number of rays and w_{mn} is the intersected area between the m-th ray and n-th pixel. Specifically, for the HAADF projection and EDS map of $e \in \mathscr{E}$

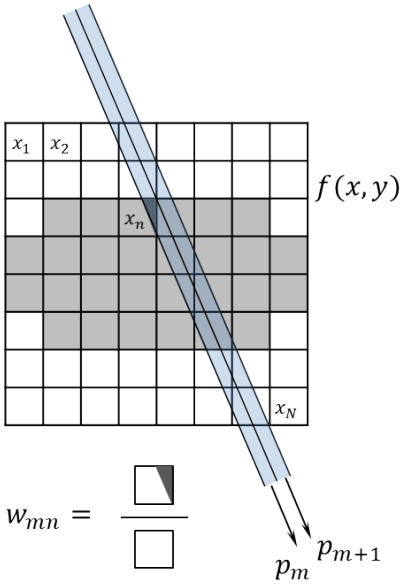


Figure 1.7: Schematic of the backward modeling for tomographic reconstruction. Discretize the unknown 2D object $f(x, y)$ into a number of N pixels, and assume that each has a constant intensity value x_n, $n = 1, \cdots, N$. The ray-sum p_m models the measured intensity corresponding to the m-th ray at the detector.

(explained in 1D in this section), this discrete modeling is

$$p_m^h = \sum_{n=1}^{N} w_{mn}^h x_n^h, \quad m = 1, \cdots, M^h, \tag{1.7}$$

$$p_m^e = \sum_{n=1}^{N} w_{mn}^e x_n^e, \quad m = 1, \cdots, M^e, \tag{1.8}$$

respectively. Note that M^h is not equal to M^e if the number of projections in the HAADF tilt-series is different from the EDS. Let us define $\mathbf{p}^u = [p_1^u, \cdots, p_{M^u}^u]^T \in \mathbb{R}^{M^u \times 1}$ for the projections, $\mathbf{x}^u = [x_1^u, \cdots, x_N^u]^T \in \mathbb{R}^{N \times 1}$ for the sample, and $\mathbf{W}^u = \{w_{mn}^u\} \in \mathbb{R}^{M^u \times N}$ for the projection matrix. Then, we can rewrite Eq. (1.7) and Eq. (1.8) to

$$\mathbf{p}^u = \mathbf{W}^u \mathbf{x}^u, \tag{1.9}$$

in which $u \in \{h, e\}$ symbolizes which modality is being used.

Tomographic reconstruction aims to estimate the unknown $\mathbf{x}^u$ from the projection data $\mathbf{p}^u$. This inverse problem, however, is ill-posed because $\mathbf{p}^u$ is inevitably corrupted by noise (e.g., Poisson noise, readout noise, etc. [24]) during the measurement; such ill-posedness becomes even severer for electron tomography where the number of projections is much smaller than the number of pixels to be reconstructed. Accordingly, the solution of $\mathbf{x}^u$ may not be stable and unique. In this section, we refer to algorithms that concurrently couple multiple tilt-series in a joint reconstruction framework as “multi-channel" and “single-channel" otherwise.

SINGLE-CHANNEL

Single-channel approaches having been developed so far are either analytical or algebraic. Analytical algorithms such as the classical filtered backprojection are based on the Fourier slice theorem, which relates the Fourier transform of a projection to the Fourier transform of the object along a radial line [24], see Fig. 1.8. Since these algorithms rely on interpolation in the Fourier space and directly calculate the reconstruction in a single step, they produce thin streaks if the number of projections is small, or if the projections are not uniformly distributed over 180° [24]. Therefore, analytical algorithms are inferior for electron tomography because the projections are (often) noisy and limited. Algebraic algorithms, on the contrary, iteratively minimize the data discrepancy $\mathscr{D}(\mathbf{W}^u\mathbf{x}^u, \mathbf{p}^u)$ between the reprojected $\mathbf{W}^u\mathbf{x}^u$ and the original $\mathbf{p}^u$

$$\mathbf{x}^\star = \arg\min_{\mathbf{x}^u} \mathscr{D}(\mathbf{W}^u\mathbf{x}^u, \mathbf{p}^u), \quad u \in \{h, e\}, \tag{1.10}$$

assuming that the projection requirement is fulfilled. In practice, they produce less artifacts for noisy and/or limited datasets [33].

The definition of $\mathscr{D}(\mathbf{W}^u\mathbf{x}^u, \mathbf{p}^u)$ depends on the statistical properties of $\mathbf{p}^u$ [34]. For instance, if the number of detected X-ray/photon counts is so low that the dominant noise in $\mathbf{p}^u$ is Poisson-distributed, then $\mathscr{D}$ is modeled as a nonquadratic functional, such as Kullback-Leibler divergence and negative log-likelihood [35]. While most algorithms in this category were originally proposed for emission tomography where the unknown

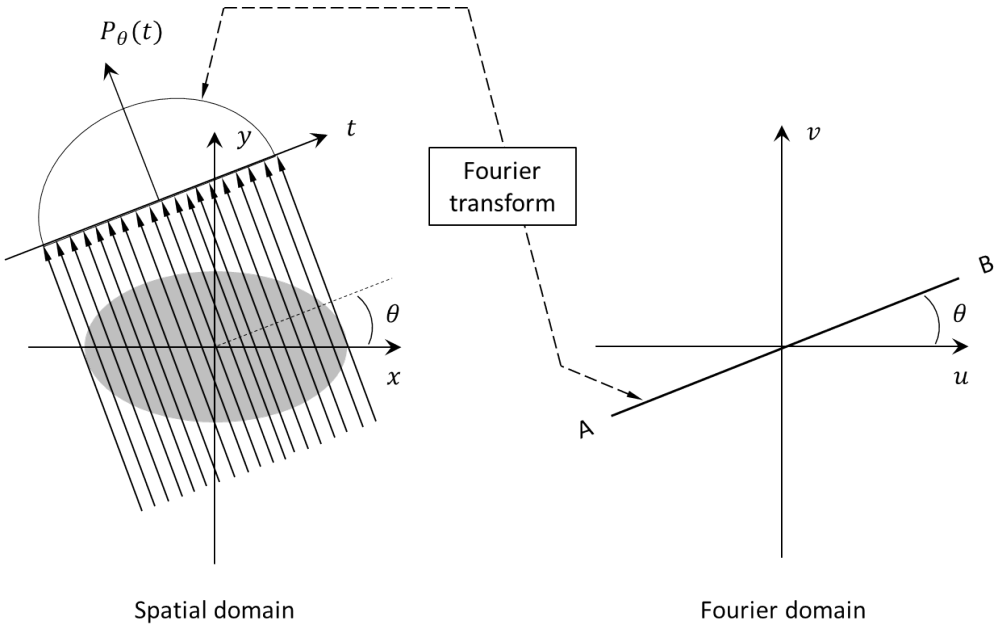


Figure 1.8: Graphical illustration of the Fourier slice theorem, which relates the Fourier transform of a parallel projection to the Fourier transform of the object along a radial line. In other words, the Fourier transform of $P_\theta(t)$ gives the values of $F(u,v)$ along line AB.

emitter density follows a Poisson distribution (e.g., the maximum likelihood reconstruction [36], the maximum a posteriori estimation [37]), they have also been adapted to electron tomography [34]. Moreover, if $\mathbf{p}^u$ is mainly limited by Gaussian noise, then $\mathscr{D}$ is often defined as a quadratic functional: $\left\|\mathbf{W}^u\mathbf{x}^u-\mathbf{p}^u\right\|_2^2$, i.e., the squared l_2 norm of $\mathbf{W}^u\mathbf{x}^u-\mathbf{p}^u$. Many reconstruction techniques have been proposed for this category, such as the classical ray-by-ray algebraic reconstruction technique (ART) and all-inclusive simultaneous iterative reconstruction technique (SIRT). Other variants also exist (e.g., simultaneous algebraic reconstruction technique (SART), ordered subset convex algorithm, etc. [38]), which aim for a much faster convergence within only a few iterations.

To further improve the reconstruction quality, one can incorporate various types of prior knowledge into the (conventional) algebraic algorithms. For example, Batenburg and Sijbers added extra discretization steps to the normal SART routine and developed discrete algebraic reconstruction technique (DART), assuming that the sample under study has only a few chemical elements and hence the reconstructed image has only a few discrete intensity levels [39]. In addition, l_1 regularization techniques were also investigated because they could approximate the unknown $\mathbf{x}^u$ by only using a small set of the basis functions [40–43]. Regularized algebraic reconstruction algorithms are formulated as

$$\mathbf{x}^\star=\arg\min_{\mathbf{x}^u}\mathscr{D}(\mathbf{W}^u\mathbf{x}^u,\mathbf{p}^u)+\lambda\mathscr{R}(\mathbf{x}^u),\quad u\in\{h,e\},\tag{1.11}$$

in which the regularization term $\mathscr{R}(\mathbf{x}^u)$ is weighted by an extra parameter λ. A common choice of $\mathscr{R}(\mathbf{x}^u)$ is the popular total variation regularization (TVR). For instance, Goris et al. [40] linked TVR in SIRT to compensate for the missing wedge, and Zhuge et al. [41] combined TVR and DART in TVR-DART to reduce the number of fine-tuning parameters.

MULTICHANNEL

In electron tomography, multichannel algorithms are either “unimodal" or “bimodal", depending on whether they use more than one modality (i.e., either h or e, or both) for reconstruction. One example of the (multichannel) unimodal algorithms is total nuclear variation (TNV) regularized EDS tomography, in which Zhong et al. [34] employed TNV to couple multiple EDS datasets so as to encourage common edges in the joint EDS reconstructions. Bimodal algorithms in electron tomography mostly combine the HAADF-STEM projections and EDS maps for integrating their complementary information [11, 19, 34, 44]. For instance, a manual parameter was introduced to weigh the HAADF and EDS channels for effectively suppressing noise and enhancing contrast [19]. Furthermore, regularizations such as the aforementioned TVR and TNV have also been considered [34, 44]. Since these regularization-aided (bimodal) techniques are so complicated that their performance highly depends on the assumptions and conditions of different datasets, it is necessary to provide a comprehensive guideline [34] for helping users to prepare for a suitable algorithmic recipe.

1.3. CHALLENGES AND OBJECTIVES

As mentioned in Section 1.1, if the HAADF projections have low Z-contrast, then EDS mapping becomes necessary to provide the valuable compositional information of the sample. However, this technique has been hampered by a low detection efficiency and slow data acquisition. For instance, the Super-X configuration depicted in Fig. 1.2 can only detect approximately 7% of the generated X-rays with its large solid angle detectors ($\Omega = 0.9$ sr [6] versus $4\pi \approx 12.6$ sr of the entire sphere). Consequently, in order to achieve an acceptable SNR for the subsequent reconstruction, acquisition usually takes 300 s [6] and may vary from 236 s to 895 s [10] per spectrum. A very high total amount of electron dose may increase the X-ray detection, but it also damages the sample. As a result, the overall dose is practically limited to $\sim 10^9$ e/nm^2 [6, 9, 10, 19, 20], resulting in a small number of elemental maps with low SNR (e.g., 47 EDS maps with max. 30 X-ray counts per pixel versus 221 HAADF projections with up to 10^5 electron counts per pixel [20]).

Many post-processing steps are available to increase the SNR of the spectral images and/or EDS maps. For instance, binning helps to increase the number of X-ray counts per pixel but at the cost of decreasing the spatial resolution [10]. Moreover, one can employ smoothing filters for denoising (e.g., moving average filter [6], edge-preserving filter [9]) but may also degrade the resolution or even produce negative intensities [10]. While multivariate statistical analysis can be an effective tool to separate the noise from the spectral data [32, 45], it becomes ineffective if the noise is dominant.

Given a limited set of noisy EDS maps, advanced techniques (see Section 1.2.2) may yield a more accurate reconstruction thanks to the incorporation of prior knowledge. However, they also involve extra parameters that highly influence the reconstruction quality. For example, the TV-, HOTV- and TNV-regularized techniques require additional parameters to determine the strength of regularization. Since over-regularization blurs fine structures and under-regularization leaves the image too noisy, it is very important to properly choose these parameters. While one may compute reconstructions for a large set of candidate values and compare them to a reference image, such method becomes infeasible with limited computational resource and/or without a reference. In

addition, although great efforts have been dedicated to developing advanced tomography techniques, directly combining reconstruction volumes at hand has still not been broadly considered to the best of our knowledge.

To address the issues discussed above, in this thesis we aim to:

1. automatically assess the image quality of the reconstructed volume so as to determine the parameters for advanced tomography techniques without a reference image;
2. design fine-tuning-free fusion algorithms to combine the HAADF and EDS reconstructions for higher image quality;
3. further reduce the number of EDS maps and/or the amount of X-ray counts to (implicitly) decrease the acquisition time and/or electron dose.

1.4. THESIS OUTLINE

In **Chapter 2**, we introduce an image quality metric, which assesses the reconstruction quality through cross-atomic contamination and noise, to automatically determine the optimal weighting factor for HAADF-EDS bimodal tomography [19]. Moreover, we effectively reduce the computational time[2] to 10%[3] by approximating the curve of metric versus weighting factor to a parabola using polynomial fitting.

Chapter 3 provides another image quality metric to measure the strength of regularization for TNV-based bimodal tomography [44]. Specifically, we use the oriented structure strength, which is described by the highest two responses in orientation space, to concurrently quantify the sharpness and noisiness of reconstructions. Similarly, observing that the curve of metric versus regularization parameter is unimodal with a distinct maximum, we adopt golden section search for the optimum of the regularization parameter and reduce the computational time by 85%.

Since reconstructions in electron tomography are preferably done in 3D [34], a 3D metric that can incorporate the information in all directions is favored to quantify the reconstruction quality. In **Chapter 4**, we use the easy-to-implement 3D Gaussian gradient magnitude to measure the regularization strength for the popular TV-aided reconstruction technique. In this case, reconstruction-assessment pipeline is accelerated by approximating the curve of the sum of gradient magnitude versus regularization parameter via spline interpolation.

Chapters 5 and 6 present a HAADF-EDS cross-modal fusion framework that does not require any fine-tuning to simultaneously achieve high SNR and high elemental specificity. In **Chapter 5**, we demonstrate our initial algorithm on simulated and experimental datasets of semiconductor devices and show that it can produce continuous edges, homogeneous foreground and clean background in its element-specific reconstructions. More importantly, it stably delivers results with high fidelity even for limited and noisy EDS datasets. In **Chapter 6**, we further add a weighted mask to the earlier framework to enable light element fusion. This modified algorithm is evaluated on an experimental dataset of a nanowire device that has more complex structures and more types of ele-

[2] Computational time refers to the amount of time spent on iterative reconstruction and quality assessment.

[3] Time reduction is compared between exhaustive search (for the optimum of the weighting factor or regularization parameter) and polynomial fitting/golden section search/spline interpolation.

ments.

Finally, **Chapter 7** summarizes the work presented in this thesis and provides recommendations for future research.

1

REFERENCES

[1] D. B. Williams and C. B. Carter, *Transmission Electron Microscopy: A Textbook for Materials Science*. Springer US, 2009.

[2] "FEI Titan Series," accessed 08 Apr. 2019. [Online]. Available: http://www.directindustry.com/prod/fei/product-123345-1401183.html

[3] P. A. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, pp. 413–431, 2003.

[4] P. Hartel, H. Rose, and C. Dinges, "Conditions and reasons for incoherent imaging in STEM," *Ultramicroscopy*, vol. 63, pp. 93–114, 1996.

[5] L. Jones *et al.*, "Rapid estimation of catalyst nanoparticle morphology and atomic-coordination by high-resolution Z-contrast electron microscopy," *Nano Letters*, vol. 14, p. 6336–6341, 2014.

[6] A. Genc *et al.*, "XEDS STEM tomography for 3D chemical characterization of nanoscale particles," *Ultramicroscopy*, vol. 131, pp. 24–32, 2013.

[7] L. Yedra *et al.*, "EEL spectroscopic tomography: towards a new dimension in nanomaterials analysis," *Ultramicroscopy*, vol. 122, pp. 12–18, 2012.

[8] H. von Harrach *et al.*, "Comparison of the detection limits of EDS and EELS in S/TEM," *Microscopy and Microanalysis*, vol. 16, pp. 1312–1313, 2010.

[9] K. Lepinay, F. Lorut, R. Pantel, and T. Epicier, "Chemical 3D tomography of a 28 nm high K metal gate transistor: STEM XEDS experimental method and results," *Micron*, vol. 47, pp. 43–19, 2013.

[10] T. J. A. Slater and et al., "STEM-EDX tomography of bimetallic nanoparticles: a methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[11] 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*, vol. 33, 2016.

[12] 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*, vol. 221, pp. 89–109, 2006.

[13] M. Hayashida and M. Malac, "Practical electron tomography guide: recent progress and future opportunities," *Micron*, vol. 91, pp. 49–74, 2016.

[14] "Electron Tomography Developments," accessed 08 Apr. 2019. [Online]. Available: https://www.int.kit.edu/1731.php

[15] I. Arslan, J. R. Tong, and P. A. Midgley, "Reducing the missing wedge: high-resolution dual axis tomography of inorganic materials," *Ultramicroscopy*, vol. 106, pp. 994–1000, 2006.

[16] X. Ke *et al.*, "Three-dimensional analysis of carbon nanotube networks in interconnects by electron tomography without missing wedge artifacts," *Microscopy and Microanalysis*, vol. 16, pp. 210–217, 2010.

[17] "Projects: Electron Tomography," accessed 15 Apr. 2019. [Online]. Available: https://www.rug.nl/research/electron-microscopy/emeritus_website/projects

[18] "Fischione Tomography Holders," accessed 15 Apr. 2019. [Online]. Available: https://www.fischione.com/products/holders

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

[20] Y. Qiu, P. van Marcke, O. Richard, H. Bender, and V. Wilfried, "3D imaging of Si FinFETs by combined HAADF-STEM and EDS tomography," in *18th International Microscopy Congress*, Prague, Czech Republic, 2014, pp. 975–976.

[21] M. Gemmi *et al.*, "Fast electron diffraction tomography," *Journal of Applied Crystallography*, vol. 48, pp. 718–727, 2015.

[22] V. Migunov *et al.*, "Rapid low dose electron tomography using a direct electron detection camera," *Scientific Reports*, vol. 5, p. 14516, 2015.

[23] S. Koneti *et al.*, "Fast electron tomography: applications to beam sensitive samples and in situ TEM or operando environmental TEM studies," *Materials Characterization*, vol. 151, pp. 480–495, 2019.

[24] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. IEEE Press, 1988.

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

[26] G. T. Herman, "Correction for beam hardening in computed tomography," *Physics in Medicine & Biology*, vol. 24, pp. 81–106, 1979.

[27] R. Aveyard and B. Rieger, "Tilt series STEM simulation of a 25x25x25 nm semiconductor with characteristic X-ray emission," *Ultramicroscopy*, vol. 171, pp. 96–103, 2016.

[28] M. C. Scott, C. C. Chen, M. Mecklenburg, C. Zhu, R. Xu, P. Ercius, U. Dahmen, B. C. Regan, and J. Miao, "Electron tomography at 2.4-ångström resolution," *Nature*, vol. 483, pp. 444–447, 2012.

[29] Z. Zhong *et al.*, "Automatic correction of nonlinear damping effects in HAADF-STEM tomography for nanomaterials of discrete compositions," *Ultramicroscopy*, vol. 184, pp. 57–65, 2018.

[30] T. J. A. Slater *et al.*, "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," *J. Phys: Conf. Ser.*, vol. 522, p. 012025, 2014.

[31] G. Cliff and G. W. Lorimer, "The quantitative analysis of thin specimens," *Journal of Microscopy*, vol. 103, pp. 203–207, 1975.

[32] P. Burdet *et al.*, "A novel 3D absorption correction method for quantitative EDX-STEM tomography," *Ultramicroscopy*, vol. 160, pp. 118–129, 2016.

[33] M. Weyland and P. A. Midgley, "Electron tomography," *Materials Today*, vol. 7, no. 12, pp. 32–40, 2004.

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

[35] T. Hohage and F. Werner, "Inverse problems with Poisson data: statistical regularization theory, applications and algorithms," *Inverse Problems*, vol. 32, p. 093001, 2016.

[36] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. 1, pp. 113–122, 1982.

[37] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging*, vol. 9, pp. 84–93, 1990.

[38] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica*, vol. 28, pp. 94–108, 2012.

[39] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2542–2553, 2011.

[40] B. Goris, W. van den Broek, K. J. Batenburg, H. H. Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012.

[41] 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 Trans. Image Process.*, vol. 25, no. 1, pp. 455–468, 2016.

[42] T. Sanders, "Open access codes for imaging problems," 2016. [Online]. Available: http://www.toby-sanders.com/software

[43] T. Sanders and R. B. Platte, "Multiscale higher order TV operators for l_1 regularization," *Advanced Structural and Chemical Imaging*, vol. 4, pp. 12–29, 2018.

[44] 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*, vol. 191, pp. 34–43, 2018.

[45] G. Lucas, P. Burdet, M. Cantoni, and C. Hébert, "Multivariate statistical analysis as a tool for the segmentation of 3D spectral data," *Micron*, vol. 52-53, p. 49–56, 2013.

2

No-Reference Weighting Factor Selection for Bimodal Tomography

Bimodal tomography introduces a weighting factor α to incorporate X-ray data into projection images acquired from scanning transmission electron microscope (STEM) so as to achieve an atom-specific three-dimensional (3D) reconstruction of an object on the nanoscale. Currently its value is chosen by computing reconstructions for a large range of $\alpha \in (0,1)$ and comparing them to a hand-segmented ground truth with the mean square error (MSE). Since this is infeasible for an industrial application, in this paper we propose an image quality metric to quantify the quality of tomograms in terms of cross-atomic contamination and noise for selecting the weighting factor without a ground truth. Numerical results demonstrate that our framework can determine the close-to-optimal weighting factor within an accuracy of ± 0.03. Moreover, approximating the shape of the minimum by a parabola effectively reduces the computational time by 90%.

Published in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1243–1247, 2018 [1].

2.1. INTRODUCTION

Electron tomography (ET) is essential for studying specimens in materials science, as it reveals the 3D structure of an object from a series of its two-dimensional (2D) projections on the nanoscale [2]. In STEM, projections formed by a high-angle annular dark-field (HAADF) detector have high signal-to-noise ratio (SNR) but only contain aggregated information of all chemical elements along the projection direction [2]; projections obtained from energy dispersive X-ray spectroscopy (EDS) accomplish an atom-specific reconstruction but suffer from low SNR [3]. In order to simultaneously exploit these two complementary techniques, HAADF-EDS bimodal tomography (HEBT) proposed in [4] introduces a weighting factor α to link both modalities into one reconstruction. The choice of α depends on the noise level and influences the reconstruction result. However, there is no a priori way to determine the "best" value. In [4], the optimal α is found by computing reconstructions over the whole range of $\alpha \in (0,1)$ and comparing them to a hand-segmented ground truth with the MSE. Since this is inapplicable for an industrial application, a quantitative quality control for reconstructions in the absence of a reference image is desired.

In recent years, no-reference image quality assessment has been widely investigated for different application scenarios [5, 6]. Proposed algorithms can be generally divided into two categories: (*i*) distortion-specific, that is, algorithms are designed specifically for one distortion. For instance, the framework presented in [7] uses Gabor filter to evaluate the streak (ringing) artifacts resulting from the iterative image restoration; (*ii*) non-distortion-specific, i.e., algorithms are generic and can respond to multiple degradations. Besides applications in computer vision, a lot of efforts have also been dedicated to developing assessment algorithms in the field of ET, such as evaluating the performance of tomographic reconstruction algorithms and/or the quality of tomograms. In [8], the length of phase boundary was treated as a quantitative morphological image characteristic to compare the commonly adopted filtered backprojection algorithm and the DIRECTT technique. In [9], Okariz et al. statistically analyzed the intensity profiles at the edge of objects in the reconstructed volume to set the number of iterations used for the simultaneous iterative reconstruction technique.

In this paper, we propose an image quality metric to choose the close-to-optimal weighting factor α for HEBT by means of quantifying the reconstruction quality of a core-shell nanoparticle consisting of gold (Au) and silver (Ag). It can replace the MSE adopted in [4] if no ground truth is available. To begin with, Section 2.2 introduces the HEBT reconstruction technique and the methodology for deciding the optimal α with a hand-segmented ground truth. Related image quality assessment algorithms are briefly reviewed in Section 2.3 as prior work, followed by our proposed quality metric, and results are presented in Section 2.4. Section 2.5 summarizes our work and discusses possible future extensions.

We use the following notations throughout this paper. Bold uppercase $\mathbf{W}$ and lowercase $\mathbf{w}$ represent matrices and column vectors, respectively, while non-bold letters W and w are scalars. Operators $(\cdot)^T$ and $*$ stand for transpose and convolution. $\mathbb{R}^{m\times n}$ denotes the space of all $m \times n$ matrices with real-valued elements.

2.2. HAADF-EDS BIMODAL TOMOGRAPHY

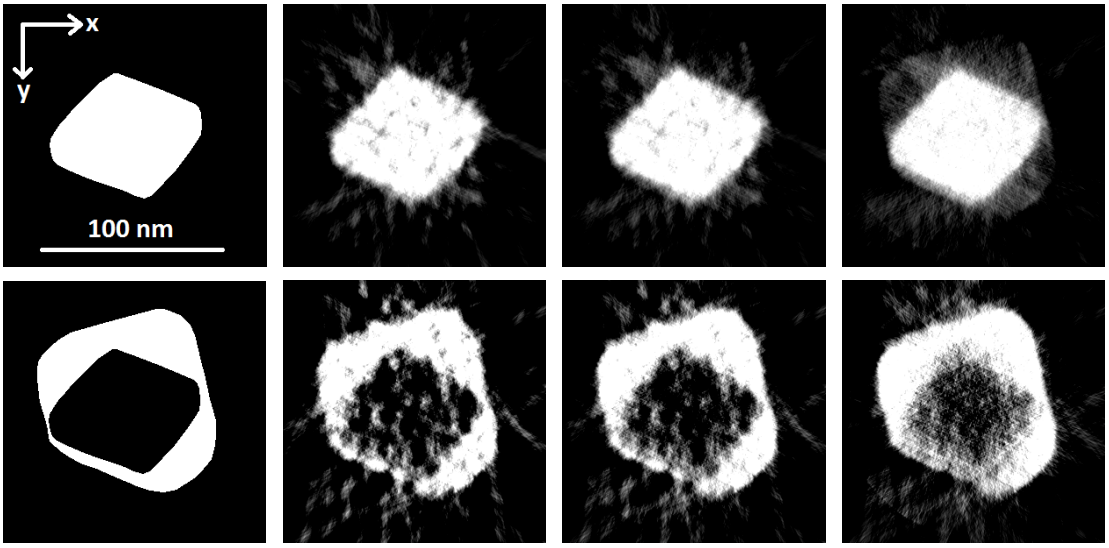
Let us assume that a specimen has a number of E different chemical elements. Each element $e = 1, \cdots, E$ is associated with an unknown volumetric object $\mathbf{x}^{(e)} \in \mathbb{R}^{N\times 1}$ where N is the total number of equally-spaced voxels to be reconstructed. Let HAADF-STEM and EDS-STEM images be $\mathbf{p}^h \in \mathbb{R}^{M\times 1}$ and $\mathbf{p}^{(e)} \in \mathbb{R}^{M\times 1}$, respectively, in which M is the total number of pixels in one projection image. In [4], HAADF-EDS bimodal tomographic reconstruction is defined as a least squares minimization problem

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \alpha^2 \left\| \mathbf{p}^h - \sum_{e=1}^{E} \mathbf{W}\mathbf{x}^{(e)} \right\|_2^2 + (1-\alpha)^2 \sum_{e=1}^{E} \left\| r^{(e)}\mathbf{p}^{(e)} - \mathbf{W}\mathbf{x}^{(e)} \right\|_2^2, \tag{2.1}$$

in which $\mathbf{x} = [\mathbf{x}^{(1)T}, \cdots, \mathbf{x}^{(E)T}]^T$, and each entry w_{mn} in $\mathbf{W} \in \mathbb{R}^{M\times EN}$ is determined by the intersected area between the m-th ray integral and n-th voxel [10]. The response ratio factor $r^{(e)}$ for element e is calculated as

$$p_m^h = \sum_{e=1}^{E} r^{(e)} p_m^{(e)}, \quad m = 1, \cdots, M. \tag{2.2}$$

In Eq. (2.1), a weighting factor $\alpha \in (0,1)$ is introduced to balance the residue terms of HAADF-STEM and EDS-STEM. In principle, α can be arbitrarily chosen between 0 and 1, whereas in practice it can neither be too small nor too large. The former makes the influence from HAADF-STEM hardly observable, and the latter leads the minimization of EDS-STEM residue term to become inefficient.



(a) GT (b) $\alpha = 0.5$ (c) $\alpha = 0.7$ (d) $\alpha = 0.9$

Figure 2.1: Au (upper) and Ag (lower) images at slice 150. The size of the reconstructions is 300×300×300 voxels. (a) ground truth (GT); (b)-(d) HEBT reconstruction results with $\mathcal{N} = 100$ iterations and weighting factors $\alpha \in \{0.5, 0.7, 0.9\}$, respectively. For better visualization, we perform percentile contrast stretching from 0 to 87%.

We consider the same core-shell nanoparticle as in [4] that consists of Au (core) and Ag (shell). Fig. 2.1 depicts HEBT reconstruction results for $\mathcal{N} = 100$ iterations and different weighting factors $\alpha \in \{0.5, 0.7, 0.9\}$ at slice 150 along the z-axis. The size of the

reconstructions is $300 \times 300 \times 300$ voxels. Two binary images in the first column are the hand-segmented ground truth with homogeneous intensity. For $\alpha = 0.5$, $\mathcal{N} = 100$ introduces overfitting, that is, the least squares optimization fits to the noise rather than true patterns; thus, reconstructions in Fig. 2.1(b) are noisier with perceptible streaks showing up. In Fig. 2.1(d), $\alpha = 0.9$ is too large and hence Au leaks into the background of Ag reconstruction and vice versa. In order to find the "best" value of α beforehand, Zhong et al. computed the reconstruction $\mathbf{x}$ for a large range of $\alpha \in (0,1)$ and compared it to the hand-segmented ground truth $\mathbf{x}_\mathrm{r}$ (see Fig. 2.1(a)) via

$$\mathrm{MSE}(\mathbf{x}_\mathrm{r}, \mathbf{x}) = \min \|\mathbf{x}_\mathrm{r} - c\mathbf{x}\|_2^2 \tag{2.3}$$

where c is a scaling factor [4]. Since this is not feasible for an industrial application, an image quality metric to quantify the quality of reconstructions is desired, such that α can be determined without a hand-segmentation.

2.3. DETERMINE WEIGHTING FACTOR WITHOUT REFERENCE

According to Fig. 2.1(a), ideal reconstructions of the core-shell nanoparticle should be binary with homogeneous foreground and zero-valued background. Inspired by the analysis of Fig. 2.1, we build our non-distortion-specific quality metric on concurrently assessing: (*i*) cross-atomic contamination, that is, how much Au is showing up in Ag regions and vice versa; (*ii*) inhomogeneity of the extracted fore- and background, and (*iii*) noise. In this section, we first present metrics that separately evaluate the aforementioned three, followed by our quality metric for choosing the close-to-optimal α for Au in the absence of its ground truth. The analysis of Ag follows the same principle.

2.3.1. CROSS-ATOMIC CONTAMINATION METRIC Q_{CC}

In order to measure the cross-atomic contamination, we first generate a binary mask B_{Au} for Au slice by slice based on the edge candidate points that are found in its volumetric reconstruction. In [11], edges are extracted by a scale-normalized differential entity

$$\mathscr{G}_\mathrm{g}^{\sigma_\mathrm{g}} = \sigma_\mathrm{g}(L_x^2 + L_y^2) \tag{2.4}$$

with $L = f * g_\mathrm{g}(\cdot;\sigma_\mathrm{g})$, such that the scale at which an edge being detected can be automatically selected. Edge strength is defined as the gradient magnitude of a smoothed image L, which is obtained by convolving the input image $f(x,y)$ with a Gaussian kernel $g_\mathrm{g}(\cdot;\sigma_\mathrm{g})$ whose standard deviation is σ_g. Finally, we calculate the cross-atomic contamination metric Q_{CC} by averaging the intensity of pixels outside the mask.

2.3.2. INHOMOGENEITY METRICS $Q_{\mathrm{IH},1}$ AND $Q_{\mathrm{IH},2}$

We evaluate the inhomogeneity of a non-ideal gray-scaled Au reconstruction by comparing it to its binary mask B_{Au}. In [12], similarity between two images f_1 and f_2 is measured by the Pearson coefficient

$$\mathrm{PC} = \frac{\sum_i (f_{1,i} - \bar{f}_1)(f_{2,i} - \bar{f}_2)}{\sqrt{\sum_i (f_{1,i} - \bar{f}_1)^2 \sum_i (f_{2,i} - \bar{f}_2)^2}} \tag{2.5}$$

where $f_{1,i}$ and $f_{2,i}$ are the intensity values of the i-th pixel, $\bar{f}_1$ and $\bar{f}_2$ the average intensities over all pixels in f_1 and f_2, respectively. When $\bar{f}_1$ and $\bar{f}_2$ are not subtracted, a new coefficient, the so-called overlap coefficient

$$\mathrm{OC} = \frac{\sum_i f_{1,i} f_{2,i}}{\sqrt{\sum_i {f_{1,i}}^2 \sum_i {f_{2,i}}^2}} \tag{2.6}$$

is defined. We denote our two inhomogeneity metrics as $Q_{\mathtt{IH},1} = 1 - \mathrm{PC}$ and $Q_{\mathtt{IH},2} = 1 - \mathrm{OC}$, respectively.

2.3.3. NOISE METRICS $Q_{\mathtt{N},1}$ AND $Q_{\mathtt{N},2}$

We investigate the noise level of Au reconstructions by computing the amount of streaks and oriented structures they contain. It is based on the previous work in [7] and [13].

In [7], streak artifacts are analyzed by a 2D Gabor filter, which can be regarded as modulating a Gaussian envelope by a sinusoidal wave with fixed frequency. Given a specific orientation θ, the corresponding Gabor response for an input image $f(x,y)$ is

$$\mathscr{G}^\theta = f * g(\cdot;\varphi,\gamma,\sigma,F_g,\theta), \tag{2.7}$$

in which φ is the phase offset, γ and σ the spatial aspect ratio and standard deviation of the Gaussian envelope, and F_g and θ the central frequency and orientation of the Gabor filter, respectively. Method in [7] works as follows:

1. decompose $f(x,y)$ using Gabor filter w.r.t. different orientations to obtain $\mathscr{G}^\theta(u,v)$;
2. for each row i (or column j) in $\mathscr{G}^\theta$, calculate the maximum oscillation strength S_i (or S_j), which is defined as the response difference between the local maximum and its neighboring local minimum;
3. compute the overall metric value for streak artifacts by finding the maximum oscillation strength S among all rows and columns and averaging over all orientations θ.

Since we do not have a priori knowledge of the width of streaks, we further extend the original module to a filter bank-based version. Its design involves two important parameters: F_g and θ. The former is determined by the central frequency of the filter at the highest frequency (F_M), the ratio between two neighboring central frequencies (F_r) and the number of frequencies (N_F), and the latter by the number of orientations (N_θ). We calculate our first noise metric $Q_{\mathtt{N},1}$ by modifying the algorithm proposed in [7] as: oscillation strength calculation and maximum value extraction are performed not only over all rows and columns but also over all frequency bands.

In [13], oriented structures are extracted by a Gaussian profile with orientation selectivity. A linear orientation space for a specific angle ϕ is defined as

$$\mathscr{H}^\phi = f * h(\cdot;N_h,F_h,B_h,\phi) \tag{2.8}$$

where $h(\cdot;N_h,F_h,B_h,\phi)$ is obtained by rotating the orientation selective template filter $h(\cdot;N_h,F_h,B_h)$ over ϕ. N_h, which relates to the orientation selectivity, is the number of filters along the ϕ-axis, and F_h and B_h the central frequency and bandwidth of the Gaussian profile, respectively. After constructing the orientation space, we further find the maximum response over ϕ and denote it as our second noise metric $Q_{\mathtt{N},2}$.

2.3.4. PROPOSED METHOD: A COMBINATION OF INDIVIDUALS

To find a proper combination for the aforementioned individual metrics, we first evaluate their own properties for different $\alpha \in [0.1, 0.9]$ with a step size of 0.01 at slice 150. Table 2.1 lists all important parameters, for choosing which we follow the guideline in [7, 13, 14] without fine-tuning.

Table 2.1: Parameters for reconstruction quality assessment

Ref.	Parameter	Symbol	Value
[7]	Phase offset	φ	0
	Spatial aspect ratio	γ	0.5
[14]	Central frequency of filter at the highest frequency	F_M	$\sqrt{2}/4$
	Frequency ratio	F_r	$\sqrt{2}$
	Number of orientations	N_θ	8
	Number of frequencies	N_F	3
[13]	Number of filters	N_h	33
	Central frequency of Gaussian profile	F_h	0.15
	Bandwidth of Gaussian profile	B_h	$0.5F_h$

As illustrated in Fig. 2.2, the background of Au reconstruction gets more contamination from Ag when α is increasing as it increases the ratio of HAADF-STEM term that contains aggregated information. Moreover, inhomogeneity and noise metrics have a clear unique minimum. We define our quality metric Q as the product of all individual metrics, namely

$$Q = Q_{\mathrm{CC}} \times Q_{\mathrm{IH},1} \times Q_{\mathrm{IH},2} \times Q_{\mathrm{N},1} \times Q_{\mathrm{N},2}. \tag{2.9}$$

Note that we do not normalize the individual metrics to $[0, 1]$; otherwise, the minimum of each curve at zero would automatically dictate the minima of the multiplication. Fig. 2.3 depicts MSE and the combinational quality metric Q for Au w.r.t. different weighting factor α and number of iterations $\mathcal{N}$, in which Fig. 2.3(a) is the same as Fig. 7 in [4]. It can be observed that there is a relatively large range of $\alpha(\sim 0.18)$ within an uncertainty of ± 0.03, see red dash-dot lines in Fig. 2.3(a). Although parabolic curves in Fig. 2.3(b) are slightly different from the ones in Fig. 2.3(a), the optimal values of α around the minima of parabolas are almost the same.

2.4. RESULTS

Since HEBT takes only one α value as the input but reconstructs both Au and Ag as the output, we consider them simultaneously in this section by summing up their MSE (or Q) values to find the minima. Note that it takes around two hours to generate one parabolic curve in Fig. 2.3(b), during which CPU time is mainly occupied for noise analysis. Therefore, we choose 8 equidistant samples for α from 80 points in total and perform polynomial fitting to predict parabolas and reduce the computational time by 90%.

Fig. 2.4 depicts the optimal values of α for different number of iterations $\mathcal{N}$ at slice 150, which are found by MSE, true and predicted quality metric Q, respectively. It demon-

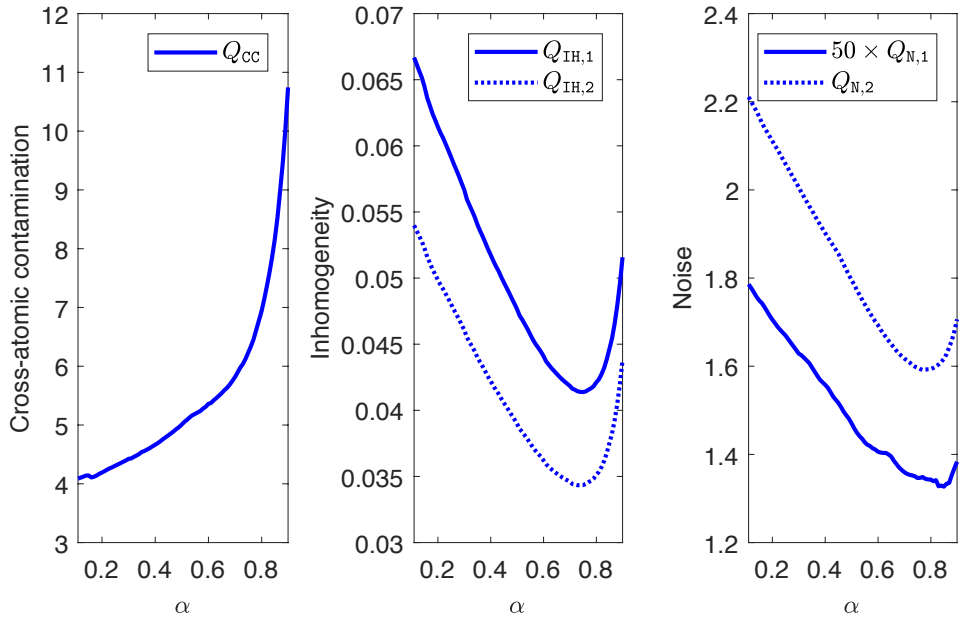
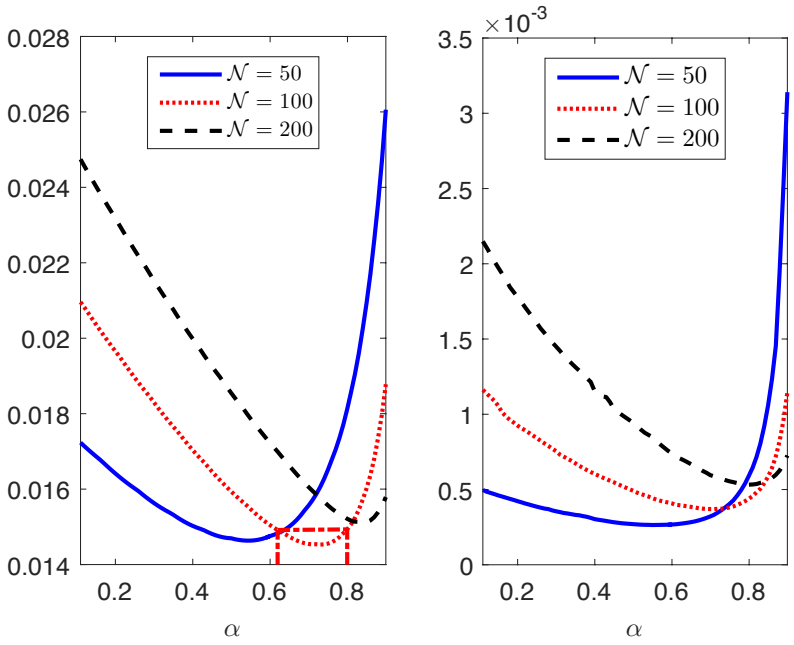


Figure 2.2: Metric values of cross-atomic contamination, inhomogeneity and noise versus weighting factor $\alpha \in [0.1, 0.9]$ for Au with 100 iterations adopted for HEBT at slice 150.



(a) MSE

(b) Quality metric Q

Figure 2.3: MSE and quality metric Q versus weighting factor $\alpha \in [0.1, 0.9]$ with different number of iterations $\mathcal{N}$ for Au at slice 150.

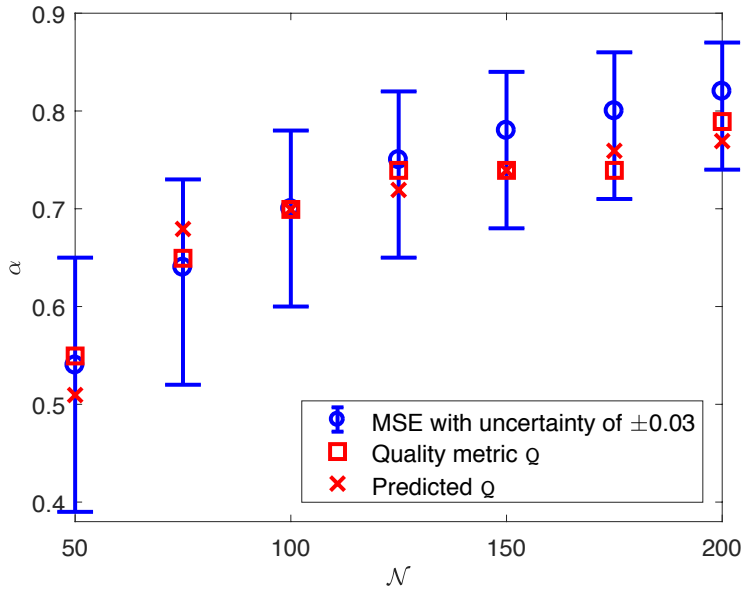


Figure 2.4: Weighting factor α versus number of iterations $\mathcal{N}$ adopted for HEBT at slice 150.

strates that besides α obtained from MSE, the other two also tend to increase with the increment of $\mathcal{N}$. The reason is that for large $\mathcal{N}$, large α guarantees that it converges to the true pattern rather than noise. Moreover, α from both true and predicted quality metric values achieve an uncertainty of ± 0.03 independent of the number of iterations $\mathcal{N}$ adopted for HEBT.

Fig. 2.5 illustrates the consistency of the optimal α w.r.t. different slices while Au and Ag are being considered simultaneously. Note that our quality metric Q is the closest to MSE at slice 150 because it is in the middle of the reconstruction stack ($300 \times 300 \times 300$) and thus suffers the least from boundary artifacts. However, even in the worst case where Q being the furthest to MSE, i.e., slices 80 and 170, α calculated and/or predicted by our quality metric still achieves an accuracy of ± 0.03.

2.5. CONCLUSION

In this paper, we propose a no-reference quality metric for HEBT to automatically determine its weighting factor α by quantitatively evaluating the quality of tomograms. Furthermore, approximating the parabola by polynomial fitting reduces the computational time to 10%, which makes our quality metric more promising. As for future work, we consider embedding the proposed assessment module into a learning system, such that α can be chosen in real-time without the need to perform reconstruction.

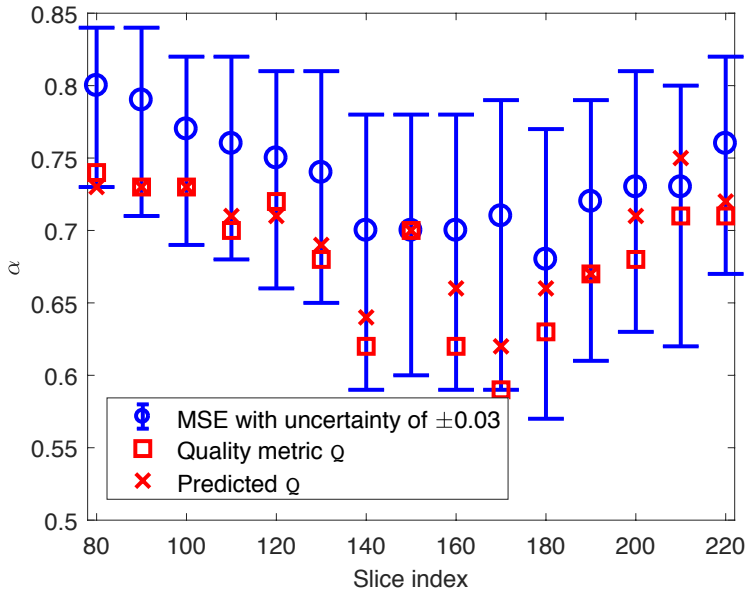


Figure 2.5: Weighting factor α versus slice index with 100 iterations adopted for HEBT.

REFERENCES

[1] Y. Guo and B. Rieger, "No-reference weighting factor selection for bimodal tomography," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018, pp. 1243–1247.

[2] P.A. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, pp. 413–431, 2003.

[3] T.J.A. Slater et al., "STEM-EDX tomography of bimetallic nanoparticles: a methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[4] Z. Zhong et al., "A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM," *Ultramicroscopy*, vol. 174, pp. 35–45, 2017.

[5] V. Kamble and K.M. Bhurchandi, "No-reference image quality assessment algorithms: a survey," *Optik*, vol. 126, pp. 1090–1097, 2015.

[6] M. Shahid et al., "No-reference image and video quality assessment: a classification and review of recent approaches," *EURASIP Journal on Image and Video Processing*, vol. 40, 2014.

[7] B. Zuo, J. Tian, and D. Ming, "A no-reference ringing metrics for images deconvolution," in *Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition*, 2008, vol. 1, pp. 96–101.

[8] S. Lück et al., "Statistical analysis of tomographic reconstruction algorithms by morphological image characteristics," *Image Analysis and Stereology*, vol. 29, pp. 61–77, 2010.

[9] A. Okariz, T. Guraya, M. Iturrondobeitia, and J. Ibarretxe, "A methodology for finding the optimal iteration number of the SIRT algorithm for quantitative electron tomography," *Ultramicroscopy*, vol. 173, pp. 36–46, 2017.

[10] A.C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[11] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. J. of Computer Vision*, vol. 30, pp. 117–156, 1998.

[12] E.M.M. Manders, F.J. Verbeek, and J.A. Aten, "Measurement of co-localization of objects in dual-color confocal images," *Journal of Microscopy*, vol. 169, pp. 375–382, Mar. 1993.

[13] M. van Ginkel, *Image analysis using orientation space based on steerable filters*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2002.

[14] F. Bianconi and A. Fernandez, "Evaluation of the effects of Gabor filter parameters on texture classification," *Pattern Recognition*, vol. 40, pp. 3325–3335, 2007.

3

Parameter Selection for TNV-Regularized Electron Tomography

Regularization has been introduced to electron tomography for enhancing the reconstruction quality. Since over-regularization smears out sharp edges and under-regularization leaves the image too noisy, finding the optimal regularization strength is crucial. To this end, one can either manually tune regularization parameters by trial and error, or compute reconstructions for a large set of candidate values and compare them to a reference image. Both are cumbersome in practice. In this paper, we propose an image quality metric Q to quantify the reconstruction quality for automatically determining the optimal regularization parameter λ without a reference image. Specifically, we use the oriented structure strength, which is described by the highest two responses in orientation space, to simultaneously measure the sharpness and noisiness of reconstruction images. We demonstrate the usefulness of Q on a recently introduced total nuclear variation regularized reconstruction technique using simulated and experimental datasets of core-shell nanoparticles. Results show that it can replace the full-reference correlation coefficient to find the optimal λ. Moreover, observing that the curve of Q versus λ has a distinct maximum attained for the best quality, we adopt the golden section search for the optimum to effectively reduce the computational time by 85%.

Published in *Proceedings of the Scandinavian Conference on Image Analysis*, pp. 452–464, 2019 [1].

3.1. INTRODUCTION

Electron tomography enables materials scientists to characterize nanoparticles in three dimensions (3D) [2]. Scanning transmission electron microscopy (STEM) has many imaging modes such as high-angle annular dark-field (HAADF) [2], in which the sample under study is exposed to a focused electron beam and tilted to obtain two-dimensional (2D) projections at different angles. In tomography, the collection of projections is called a tilt-series, from which we can reconstruct a 3D image that represents the sample. Although HAADF tomography can clearly reveal the inner structure of the sample, it cannot explicitly provide the compositional information. To better understand samples with more complex chemical compositions, spectral imaging techniques like energy dispersive X-ray spectroscopy (EDS) [3] must be pursued. EDS tomography, however, is currently hampered by slow data acquisition, resulting in a limited number of elemental maps with low signal-to-noise ratio (SNR) [3].

Electron tomography is an ill-posed inverse problem whose solution is not stable and unique. Therefore, l_1 regularizations (e.g., total variation (TV) [4], higher order total variation (HOTV) [5, 6]) have been introduced to enhance the reconstruction quality. However, regularizations, especially the common TV, inevitably aggravate jaggy edges and staircase artifacts when being applied to the (noisy) EDS datasets. To alleviate such artifacts but still benefit from regularization, Zhong et al. incorporated the HAADF-STEM projections with high SNR into EDS maps using total nuclear variation (TNV) to enforce anti-/parallel gradients and common edges in joint reconstructions [7]. Like other regularization-based approaches, TNV also requires a fine-tuning parameter λ to determine the strength of regularization. The “best" λ is now chosen by computing reconstructions for a large set of candidate values and comparing them to a reference image with the correlation coefficient [7]. Since this is infeasible if the reference is unavailable, we need to automatically measure the reconstruction quality for determining the optimal λ.

So far, many no-reference quality assessment algorithms have been proposed to set appropriate parameters for inverse problems. For instance, Zhu and Milanfar developed a structure tensor based image content index to optimize denoising algorithms [8]. Since this index is easy to compute, it has also been adopted to determine the optimal regularization parameter for the TV reconstruction technique [9]. Applications dedicated to electron tomography also exist [10, 11]. For example, Okariz et al. derived the optimal number of iterations for simultaneous iterative reconstruction technique (SIRT) by statistically analyzing the edge profile of reconstructions [11]. Furthermore, we recently proposed a non-distortion-specific image quality metric to quantify the cross-atomic contamination and noise so as to select the optimal weighting factor for bimodal tomography [10]. However, automatically selecting parameters for regularized electron tomography has still not been widely investigated to the best of our knowledge.

In this paper, we aim to automatically find the optimal regularization parameter λ for TNV in the absence of a reference image. Specifically, we extend the concept of image content index [8] to orientation space (OS) [12], in which we develope a metric Q to assess the reconstruction quality regarding the sharpness and noisiness. We demonstrate our Q on simulated and experimental datasets of core-shell nanoparticles containing gold and silver. Results show that this OS-based Q is more robust to noise than

the original tensor-based version. Moreover, it can replace the full-reference correlation coefficient used in [7] to determine the optimal λ. In Section 3.2, we introduce the TNV-regularized reconstruction technique and its relations to TV. Section 3.3 elaborates the orientation space as prior work, followed by our quality assessment framework for parameter determination. We present the experiments and results in Section 3.4 and summarize our work in Section 3.5.

3.2. TNV-REGULARIZED ELECTRON TOMOGRAPHY

Originally proposed for color images [13], total nuclear variation (TNV) has later been applied to multi-channel spectral CT data for encouraging common edge locations and a shared gradient direction among different channels [14]. Let us assume that an arbitrary 3D image $\mathbf{A}$ has a number of L channels, in which $\mathbf{A}_n = [A_n^{(1)}, \cdots, A_n^{(L)}]^T \in \mathbb{R}^{L\times 1}$ is the intensity value tuple of its n-th voxel. Given the Jacobian matrix of $\mathbf{A}$ as

$$\mathbf{J}_n\mathbf{A} = \begin{bmatrix} \nabla^x A^{(1)}, & \nabla^y A^{(1)}, & \nabla^z A^{(1)} \\ & \vdots & \\ \nabla^x A^{(L)}, & \nabla^y A^{(L)}, & \nabla^z A^{(L)} \end{bmatrix}, \tag{3.1}$$

then TNV of $\mathbf{A}$ is

$$\mathrm{TNV}(\mathbf{A}) = \sum_n \|\mathbf{J}_n\mathbf{A}\|_\star \tag{3.2}$$

where $\|\mathbf{J}_n\mathbf{A}\|_\star$, the nuclear norm of $\mathbf{J}_n\mathbf{A}$, is the sum of its singular values [14]. When $L = 1$, TNV reduces to the isotropic (l_2-norm) TV [7].

We consider a specimen with a number of E different chemical elements. Each element $e = 1, \cdots, E$ has its EDS map $\mathbf{p}^{(e)} \in \mathbb{R}^{M^e\times 1}$ and is associated with one unknown reconstruction volume $\mathbf{x}^{(e)} \in \mathbb{R}^{N\times 1}$. M^e is the number of pixels in the map and N the number of discretized voxels to be reconstructed. Similarly, let $\mathbf{p}^h \in \mathbb{R}^{M^h\times 1}$ and $\mathbf{x}^h \in \mathbb{R}^{N\times 1}$ be the projection and volumetric reconstruction of HAADF, respectively. Note that M^h, the number of pixels in the HAADF projection, is not equal to M^e if the HAADF tilt-series has more acquisition angles than the EDS.

Given $\mathbf{A}_n$ as a two-channel image $\mathbf{A}_n = [x_n^{(e)}, x_n^h]^T$, i.e., one element of interest plus HAADF, the TNV-regularized EDS and HAADF joint tomography is [7]

$$\mathbf{x}^{(e)*}, \mathbf{x}^{h*} = \underset{\mathbf{x}^{(e)}, \mathbf{x}^h}{\arg\min} \left\|\mathbf{p}^{(e)} - \mathbf{W}^{(e)}\mathbf{x}^{(e)}\right\|_2^2 + \left\|\mathbf{p}^h - \mathbf{W}^h\mathbf{x}^h\right\|_2^2 + \lambda\mathrm{TNV}(\mathbf{x}^{(e)}, \mathbf{x}^h). \tag{3.3}$$

Extending $\mathbf{A}_n$ to multiple channels with more than one element is also possible, as long as they share common edges [7]. In Eq. (3.3), $\mathbf{W}^{(e)} \in \mathbb{R}^{M^e\times N}$ and $\mathbf{W}^h \in \mathbb{R}^{M^h\times N}$ are the projection matrices of the EDS and HAADF, respectively, whose entries $w_{mn}^{(e)}$ and w_{mn}^h are determined by the intersected area between the m-th ray integral and n-th voxel. When the HAADF term is removed and $A_n = x_n^{(e)}$, Eq. (3.3) reduces to the TV-regularized EDS tomography [4].

The parameter λ in Eq. (3.3) determines the strength of TNV regularization. A large λ may blur sharp edges and produce an over-smoothed reconstruction, whereas a small one may make the regularization ineffective. To choose this crucial parameter, Zhong

et al. computed the reconstructions $\mathbf{x}^{(e)*}$ for a large set of λ (e.g., 100 values uniformly sampled from 10^{-3} to 10^{1} on the logarithmic scale) and compared them to a noise-free image using the correlation coefficient [7]. Since this is infeasible in industry, we need a no-reference quality metric to quantify the reconstruction quality so as to (blindly) determine the optimal λ.

3.3. NO-REFERENCE REGULARIZATION PARAMETER DETERMINATION

Considering that the effect of regularization varies spatially, we propose to use the local oriented structure strength (OSS) to measure the image quality; it has large values for well structured patches containing lines and edges and small values for blurry/noisy ones. In this section, we first introduce the concept of orientation space [12] and then present our OSS-based quality assessment framework.

3.3.1. ORIENTATION SPACE

The linear orientation space of a 3D input image $I(\mathbf{x})$ can be constructed as

$$I_h(\mathbf{x},\phi,\theta) = I(\mathbf{x}) * h(\mathbf{x};\phi,\theta) \tag{3.4}$$

where $\mathbf{x}$ is the Cartesian coordinate tuple containing x, y and z. Operator $*$ denotes convolution. $h(\mathbf{x};\phi,\theta)$ is obtained by rotating an elongated template filter $h(\mathbf{x})$ over angles ϕ and θ in a unit sphere. $\phi \in [0,2\pi)$ is the counter-clockwise angle measured from the positive x-axis in the xy-plane; $\theta \in [0,\pi)$ is the angular distance from the positive z-axis [12]. One promising candidate for $h(\mathbf{x})$ is a Gabor filter [15]; however, it cannot produce a zero response to a constant signal. Therefore, we use a similar filter which is zero for a constant signal [12].

According to van Ginkel et al., the choice of the template filter $h(\cdot)$ is largely free, as long as the scale and orientation can be dealt with separately [16]. To this end, Faas and van Vliet constrained the Fourier transform of $h(\mathbf{x})$ to have separable radial and angular parts, that is,

$$\mathscr{F}\{h(\mathbf{x})\} = H(\mathbf{f}) = H_{\mathtt{rad}}(f)\,H_{\mathtt{ang}}(\phi,\theta) \tag{3.5}$$

where $\mathbf{f}$ is the polar coordinate tuple containing f, ϕ and θ in the Fourier domain [12]. The radial component $H_{\mathtt{rad}}(f;f_c,b_f)$ is a Gaussian-like bandpass filter where f_c and b_f are the central frequency and bandwidth of the Gaussian profile, respectively. It reaches its maximum for $f = f_c$ and goes to zero for $f = 0$. The angular component $H_{\mathtt{ang}}(\phi,\theta;N)$ relies on a parameter N to control the orientation selectivity, which is the number of orientations in the upper half of the unit sphere formed by ϕ and θ. For details of mathematical expressions see [12]. When θ is removed, $H(\mathbf{f})$ becomes the 2D filterbank presented in [16].

$I_h(\mathbf{x},\phi,\theta)$ has a number of peaks. The amplitude of the strongest peak

$$A_1(\mathbf{x}) = \max_{\phi,\theta} |I_h(\mathbf{x},\phi,\theta)| \tag{3.6}$$

captures highly regular regions with one single orientation; the amplitude of the second strongest peak $A_2(\mathbf{x})$ highlights special patterns such as deformation and bifurcation;

the remaining peaks and noise are described by a residue term $R(\mathbf{x},\phi,\theta)$ which reflects chaotic regions [16]. Intuitively, a large A_1 and a small A_2 indicate a prominent elongated structure.

3.3.2. RECONSTRUCTION QUALITY ASSESSMENT USING ORIENTATION SPACE

Our patch-based quality assessment algorithm consists of three steps: *(i)* construct an orientation space; *(ii)* compute the local and *(iii)* global quality metrics, see Fig. 3.1. Note that this method is currently implemented and discussed here in 2D in a first result.

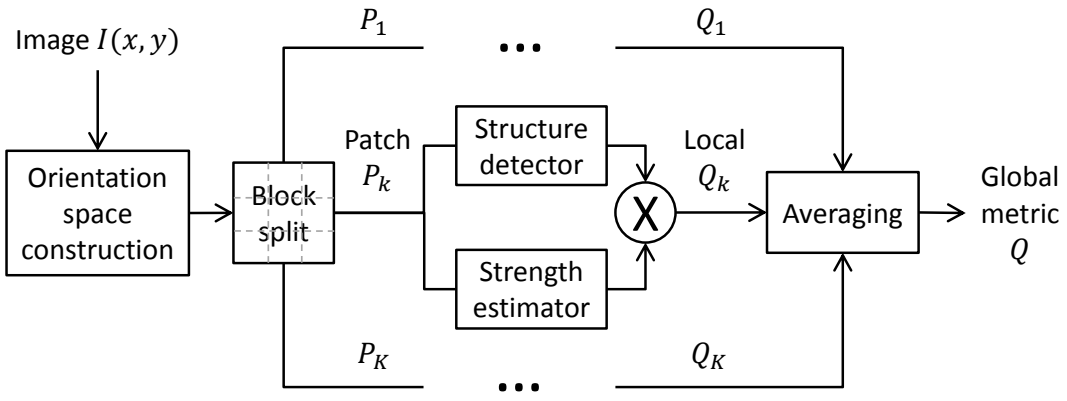


Figure 3.1: Framework for reconstruction quality assessment. Details in Section 3.3.2.

CONSTRUCT ORIENTATION SPACE

For each reconstruction slice $I(x,y)$, we first construct its orientation space $I_h(x,y,\phi)$ using Eq. (3.4). Then, we extract the amplitudes of the two strongest peaks $A_1(x,y)$ and $A_2(x,y)$. Throughout this paper, $I_h(x,y,\phi)$, $A_1(x,y)$ and $A_2(x,y)$ are computed with the open source DIPimage toolbox [17]. Moreover, we set $f_c = 0.25$, $b_f = 0.8 f_c$ and $N = 8$, so that the template filter $h(x,y;f_c,b_f,N)$ behaves like a line/edge detector [16].

COMPUTE LOCAL METRIC

Divide $I(x,y)$ into a number of K non-overlapped rectangular patches P_k, $k = 1,\cdots,K$; each goes through two modules: structure detector and strength estimator. The *structure detector* determines whether P_k contains any prominent structure (e.g., edges) by measuring its contrast. To eliminate outliers such as noise, we define the contrast of P_k as the interquartile range (75 percentile minus 25 percentile) rather than the full range (maximum minus minimum) of its pixel intensities. We set $\text{is\_stru}_k = 1$ if the contrast of P_k is larger than the average intensity of $I(x,y)$ and $\text{is\_stru}_k = 0$ otherwise. The *strength estimator* quantifies the saliency of patch P_k, for which the gradient structure tensor has been considered earlier. For instance, Zhu and Milanfar proposed the image content index [8]

$$q = s_1 \frac{s_1 - s_2}{s_1 + s_2} \tag{3.7}$$

where s_1 and s_2 are the singular values of the 2×2 tensor matrix. In this paper, we replace s_1 and s_2 by the amplitudes A_1 and A_2, because the latter two are more sensitive to fine

structures under noise [18]. Consequently, the oriented structure strength of P_k is given by

$$\mathrm{OSS}_k = 1 - \frac{\mathrm{geomean}\{q_i\}}{\mathrm{mean}\{q_i\}}, \quad q_i = A_{1,i}\frac{A_{1,i} - A_{2,i}}{A_{1,i} + A_{2,i}}, \quad i \in P_k, \tag{3.8}$$

in which geomean$\{\cdot\}$ and mean$\{\cdot\}$ represent the geometric and arithmetic mean, respectively. The underlying rationale is that the more "spiky" q is, the stronger the oriented structure in P_k will be. Moreover, if P_k is constant or exactly at the boundary between two orientation fields (i.e., $A_{1,i} = A_{2,i} = 0$ or $A_{1,i} = A_{2,i}$, $\forall i \in P_k$ [16]), we set $\mathrm{OSS}_k = 0$. Finally, we compute the local quality metric by multiplying the outputs of the two independent modules: $Q_k = \mathrm{is\_stru}_k \times \mathrm{OSS}_k$.

3

COMPUTE GLOBAL METRIC

We define the global quality metric Q as the geometric mean of all nonzero Q_k, that is,

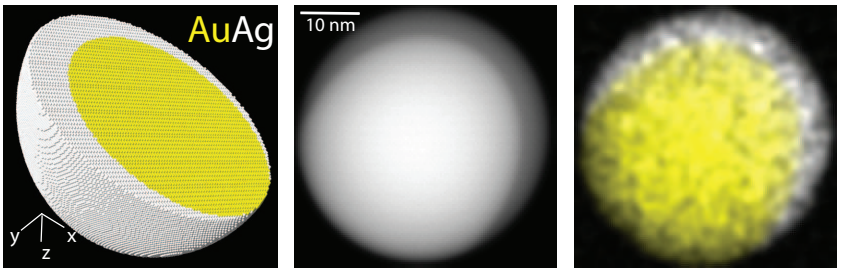
$$Q = \mathrm{geomean}\{Q_k\}, \quad Q_k \neq 0, \quad k = 1, \cdots, K. \tag{3.9}$$

We do not consider the arithmetic mean because it (unwanted) gives higher weight to Q_k with larger numeric range.

3.4. EXPERIMENTS AND RESULTS

In this section, we demonstrate that our quality metric Q can select a close-to-optimal λ for the TNV-regularized reconstruction technique. Hereinafter, we consider simulated and experimental datasets of core-shell nanoparticles containing gold (Au) in the core and silver (Ag) in the shell. These two chemical elements have distinct atomic numbers ($Z_{\mathrm{Au}} = 79$, $Z_{\mathrm{Ag}} = 47$) and hence can produce high Z-contrast HAADF-STEM projections for the TNV to augment EDS maps. Moreover, the TNV-regularized tomography was realized by Douglas-Rachford primal-dual splitting algorithm with the operator discretization library [19]. We set 400 iterations to guarantee convergence and sampled 100 points for λ, which were uniformly distributed between 0.001 and 1.0 on the logarithmic scale.

3.4.1. SIMULATED DATASET



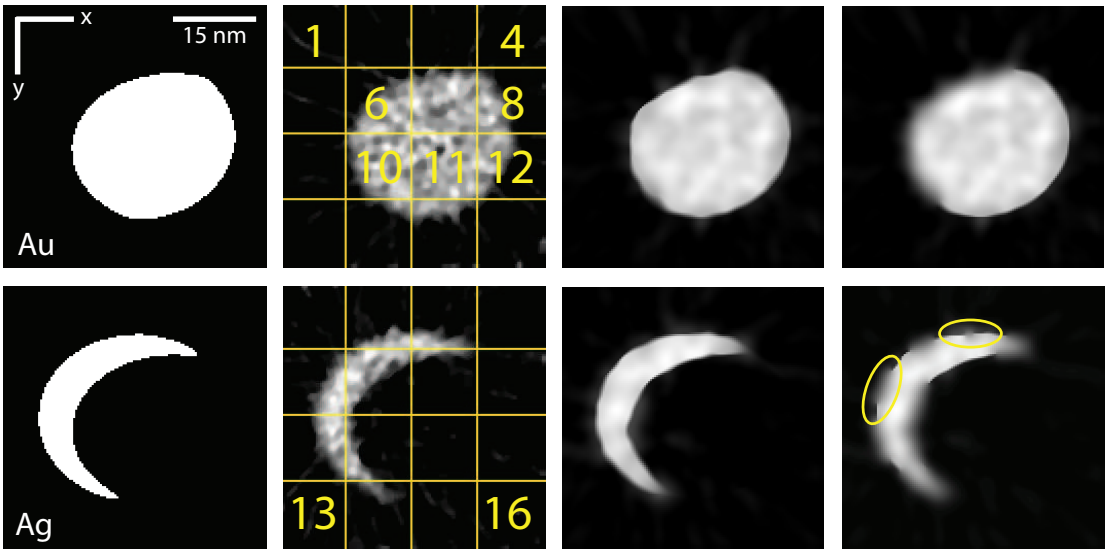
(a) Atomic design (b) HAADF (c) EDS map

Figure 3.2: (a) Atomic design of a core-shell nanoparticle consisting of gold (Au, yellow) and silver (Ag, white). (b) Simulated HAADF-STEM projection and (c) superposed EDS map at 7.5°.

To begin with, we simulated a noise-free multislice dataset using an Au-Ag nanoparticle in a box with a size of 40 nm × 40 nm × 40 nm, see Fig. 3.2(a). For details of simulation

see [20]. HAADF-STEM projections and EDS maps with a size of 128 pixel × 128 pixel (≈ 4Å/pixel) were simulated in every 2.5° over [0°, 180°). We used a focused electron beam normalized to an intensity of 1, a convergence angle of 10 mrad, and a detector with an inner angle of 90 mrad and outer 230 mrad. Since we did not include any (spherical) aberration, we set the accelerating voltage to 120 kV rather than 200 kV [3] for a broader beam.

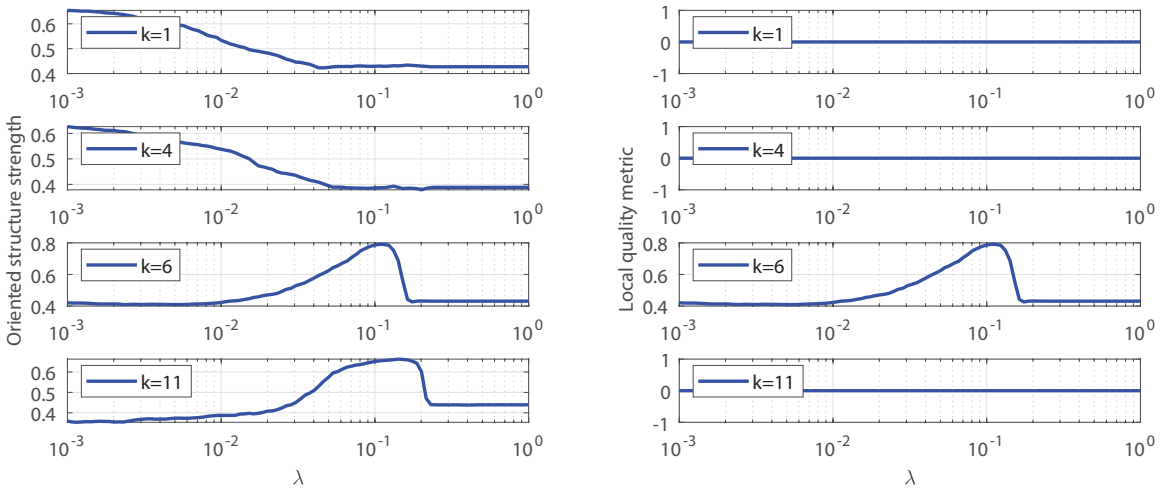
Then, we introduced several post-processing steps to make this noiseless dataset more realistic. HAADF-STEM projections were blurred by Gaussian smoothing ($\sigma = 1.0$ pixel) and corrupted by Poisson noise with a mean of the number of electron counts (up to 10^5 per pixel) and Gaussian noise with a standard deviation of 0.2. Projections suffering from the channeling effect were removed [21]. For EDS maps, we set the maximum X-ray count per pixel to 4 for Au and 3 for Ag, so that the total number of X-ray counts per angle were comparable to real experimental data [3]. Since EDS maps were much noisier, we employed a Gaussian filter ($\sigma = 1.0$ pixel) for denoising. Finally, we subsampled the EDS tilt-series by a factor of 2, as in practice the number of EDS maps is typically smaller than HAADF projections due to acquisition time. The resulting HAADF-STEM and EDS data are shown in Fig. 3.2.



(a) Ground truth (b) $\lambda = 0.001$ (c) $\lambda = 0.1233$ (d) $\lambda = 0.1748$

Figure 3.3: Au (upper) and Ag (lower) xy-slices for the simulated dataset at $z = 24$. (a) Ground truth, segmented from SIRT reconstructions with 100 iterations using 72 elemental maps between $[0, 180^\circ)$; (b)-(d) TNV reconstructions with regularization parameter $\lambda \in \{0.001, 0.1233, 0.1748\}$. The size of the reconstructions is $128 \times 128 \times 128$ voxels.

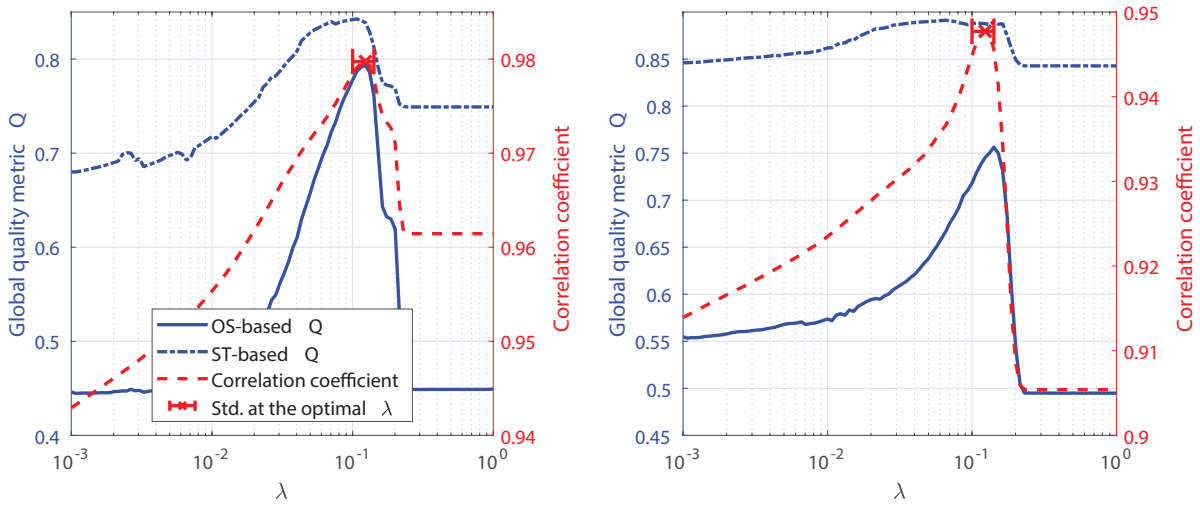
Fig. 3.3 illustrates the xy-slices of Au and Ag at $z = 24$, which are reconstructed with TNV using different λ. Two binary images in Fig. 3.3(a) are the ground truth segmented from SIRT reconstructions with 100 iterations given the full-view noiseless EDS maps. Fig. 3.3(b) shows 16 patches with four different types of structures: object (P_{11}), background (P_4, P_{16}), background with streak artifacts (P_1, P_{13}), and edge (P_6, P_8, P_{10}, P_{12}). For $\lambda = 0.001$, a weak regularization leads to an overall noisy reconstruction. However, when λ is increased up to a certain level (e.g., $\lambda = 0.1748$), strong regularization starts to nonuniformly degrade the sharp edges, see yellow circles in Fig. 3.3(d).



(a) OSS_k (b) Q_k

Figure 3.4: Oriented structure strength OSS_k and local quality metric Q_k versus λ for four patches P_k in Fig. 3.3. P_1: background with streak artifacts; P_4: background; P_6: edge; P_{11}: object. Results are averaged over ten noise realizations.

Fig. 3.4(a) plots the oriented structure strength (OSS) as a function of λ for four patches selected from Fig. 3.3. OSS curves of the background patches P_1 and P_4 (w/ and w/o perceptible streak artifacts) are decreasing when λ is increasing, because stronger regularization can more effectively suppress the noise. In addition, OSS curves of the edge (P_6) and object (P_{11}) patches are similar, whereas the former has a clearer unique maximum. Fig. 3.4(b) shows the corresponding local quality metrics, in which only Q_6 with $\mathrm{is\_stru}_6 = 1$ is nonzero.



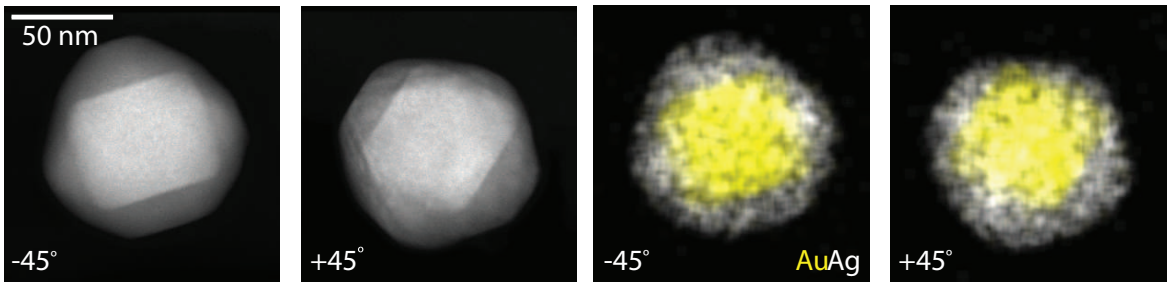
(a) TNV, Au (b) TNV, Ag

Figure 3.5: Global quality metric Q and correlation coefficient versus λ for the simulated dataset at $z = 24$. Q is derived either from the orientation space (OS) or structure tensor (ST); CC is obtained by comparing reconstructions to the ground truth in Fig. 3.3(a). Results are averaged over ten noise realizations.

Fig. 3.5 depicts our main result, in which we plot the global quality metric Q and correlation coefficient (CC) as a function of λ. Q is derived either from our orientation space (OS) or from the structure tensor (ST) [8]; CC is calculated by comparing reconstructions to the binary segmentation of the noiseless SIRT reconstruction. It can be observed that our OS-based Q has a very good agreement with CC for the optimal λ, i.e., λ values around the maxima of OS-based Q and CC are almost the same. Moreover, our OS-based Q has a higher dynamic range than the ST-based version especially for Ag, see Fig. 3.5(b). As a result, it would be more robust to small fluctuations such as noise in practice.

Note that TNV is an iterative technique that takes significant amount of time for reconstruction. For example, it took 10 hours to compute reconstructions for 100 different λ. Many efficient one-dimensional search algorithms are available for time reduction, and we choose the golden section search [22]. This algorithm assumes that the objective function is unimodal within a certain range and evaluates it at triples of points whose values form the golden ratio [22]. Since the golden section search can narrow the original 100 values of λ down to no more than 15, it would effectively reduce the total computational time by approximately 85%.

3.4.2. EXPERIMENTAL DATASET

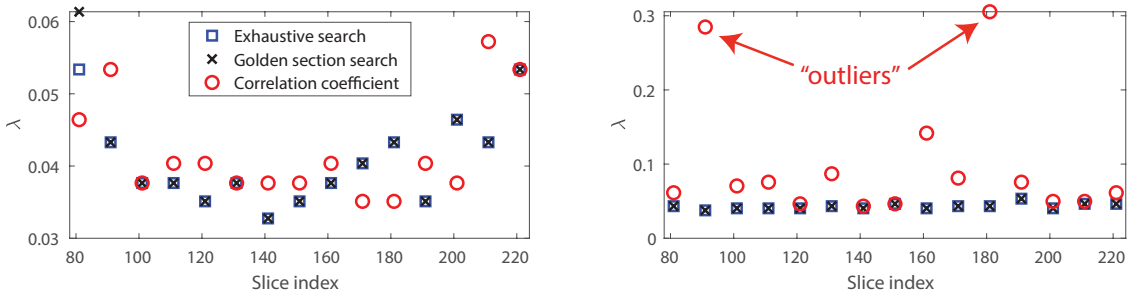


(a) HAADF-STEM (b) EDS map

Figure 3.6: Experimental (a) HAADF-STEM projections and (b) superposed EDS maps of a Au-Ag core-shell nanoparticle at $-45°$ and $+45°$.

Our experimental Au-Ag core-shell nanoparticle was scanned in a FEI Tecnai Osiris microscope, which was operated with an accelerating voltage of 120 kV and equipped with four Super-X energy dispersive silicon drift detectors [23]. HAADF-STEM projections with a size of 300 pixel × 300 pixel were acquired at 31 tilt angles, ranging from $-75°$ to $+75°$ with an increment of $5°$. In addition, one X-ray spectral image has also been recorded at each angle for 300 seconds. The raw dataset was then processed before reconstruction. The HAADF-STEM tilt-series was aligned using cross-correlation; X-ray spectral images were denoised by principal component analysis and deconvolved into two equi-sized elemental maps, one for Au and the other for Ag [23]. Fig. 3.6 gives an example of the post-processed experimental projections, for which we hand-segmented the HAADF reconstruction to obtain the ground truth of EDS.

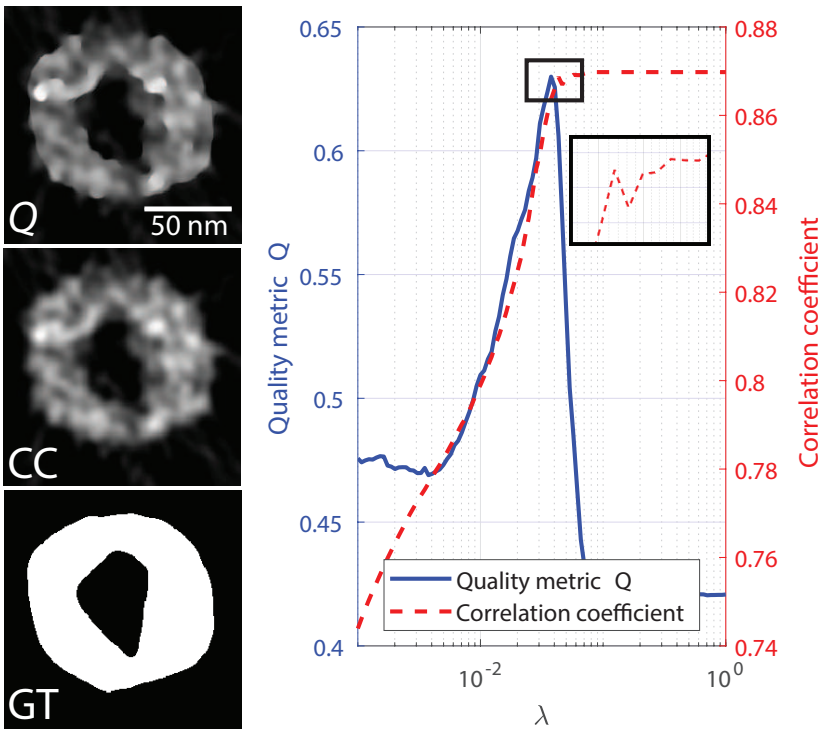
Fig. 3.7 shows the variation of the optimal λ w.r.t. different slices. λ values found by our no-reference metric Q and the full-reference metric CC are comparable, considering that the search space spans over 3 orders of magnitude. Moreover, golden section search



(a) Au (b) Ag

Figure 3.7: Optimal regularization parameter λ versus slice index for the experimental dataset. The size of the reconstructions is $300 \times 300 \times 300$ voxels.

and exhaustive search lead to the same λ most of the time, though the former may terminate at the local maximum before reaching the global one (e.g., slice number 81 in Fig. 3.7(a)). Note that Fig. 3.7(b) has two "outliers", and we show the details of slice number 91 in Fig. 3.8. It is obvious that the Ag reconstruction computed from Q maintains finer structure than the one from CC, especially at the edges of the outer ring. From Fig. 3.8(d) we can see that the curve of CC strangely "jumps" after a certain λ, even though the underlying structures have already been smeared out. This shows that even using CC



(a) (b)

Figure 3.8: (a) Ag TNV reconstructions at $z = 91$ with λ found by quality metric Q and correlation coefficient (CC), compared to the hand-segmented ground truth (GT). The corresponding curves of Q and CC versus λ are shown in (b).

as a metric to choose the optimal λ for the TNV-regularized electron tomography is not always reliable.

3.5. DISCUSSION AND CONCLUSION

In this paper, we developed a no-reference quality metric Q to score the oriented structure strength of reconstruction images for detecting over- and under-regularization. Based on simulated and experimental datasets of Au-Ag core-shell nanoparticles, we demonstrated that our Q can replace the full-reference correlation coefficient to automatically determine the optimal regularization parameter λ for the recently proposed TNV reconstruction technique. Since the original experimental dataset was noisy, we further binned the tilt-series by a factor of 3 to increase the SNR. Consequently, the size of the reconstructions was reduced from $300 \times 300 \times 300$ voxels to $100 \times 100 \times 100$ voxels, for which Q still achieved a relatively high accuracy in terms of parameter determination. More interestingly, the optimal λ found in this case became larger, probably because the dataset with less noise did not produce severe paintbrush/staircase artifacts under a stronger regularization.

Compared to the iterative TNV reconstruction, time spent for the quality assessment is minor, e.g., 10 hours versus 5 minutes for 100 different λ on a desktop equipped with eight Intel Xeon X5550 CPU cores (24 GB RAM) and one NVIDIA GeForce GTX670 GPU (4 GB memory). Considering that the curve of reconstruction quality versus λ is unimodal with a distinct maximum, we adopted the golden section search to “predict" the optimal λ, which effectively reduced the total computational time (reconstruction plus assessment) by 85%.

As for future work, we consider testing the applicability of our quality metric on other iterative reconstruction techniques with (e.g., TV and HOTV) and/or without (e.g., SIRT) regularizations. Moreover, we will also extend the current framework to 3D.

REFERENCES

[1] Y. Guo and B. Rieger, "Parameter selection for regularized electron tomography without a reference image," in *Proceedings of the 2019 Scandinavian Conference on Image Analysis*, 2019, pp. 452–464.

[2] P. A. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, pp. 413–431, 2003.

[3] T. J. A. Slater et al., "STEM-EDX tomography of bimetallic nanoparticles: a methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[4] B. Goris et al., "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012.

[5] T. Sanders, A. Gelb, R. B. Platte, I. Arslan, and K. Landskron, "Recovering fine details from under-resolved electron tomography data using higher order total variation l_1 regularization," *Ultramicroscopy*, vol. 174, pp. 97–105, 2017.

[6] T. Sanders and R. B. Platte, "Multiscale higher order TV operators for l_1 regularization," *Advanced Structural and Chemical Imaging*, vol. 4, pp. 12–29, 2018.

[7] 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*, vol. 191, pp. 34–43, 2018.

[8] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on Image Processing*, vol. 19, pp. 3116–3132, 2010.

[9] H. Liang and D. S. Weller, "Regularization parameter trimming for iterative image reconstruction," in *49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 755–759.

[10] Y. Guo and B. Rieger, "No-reference weighting factor selection for bimodal tomography," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2018, pp. 1243–1247.

[11] A. Okariz, T. Guraya, M. Iturrondobeitia, and J. Ibarretxe, "A methodology for finding the optimal iteration number of the SIRT algorithm for quantitative electron tomography," *Ultramicroscopy*, vol. 173, pp. 36–46, 2017.

[12] F. G. A. Faas and L. J. van Vliet, "3D-orientation space; filters and sampling," in *Proceedings of the 13th Scandinavian Conference on Image Analysis*, 2003, pp. 36–42.

[13] K. M. Holt, "Total nuclear variation and jacobian extensions of total variation for vector fields," *IEEE Transactions on Image Processing*, vol. 23, pp. 3975–3989, 2014.

[14] D. S. Rigie and P. J. La Riviére, "Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization," *Physics in Medicine & Biology*, vol. 60, pp. 1741–1771, 2015.

[15] J. Chen, Y. Sato, and S. Tamura, "Orientation space filtering for multiple orientation line segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 417–430, 2000.

[16] M. van Ginkel, L. J. van Vliet, and P. W. Verbeek, "Applications of image analysis in orientation space," *Fourth Quinquennial Review 1996-2001 Dutch Society for Pattern Recognition and Image Processing*, pp. 355–370, 2001.

[17] "DIPlib & DIPimage," http://www.diplib.org, Last accessed 16 Dec 2018.

[18] M. van Ginkel, *Image analysis using orientation space based on steerable filters*, Ph.D. thesis, Delft University of Technology, Delft, The Netherlands, 2002.

[19] J. Adler, H. Kohr, and O. Öktem, "Operator discretization library (ODL)," https://github.com/odlgroup/odl, 2017.

[20] R. Aveyard and B. Rieger, "Tilt series STEM simulation of a 25x25x25 nm semiconductor with characteristic X-ray emission," *Ultramicroscopy*, vol. 171, pp. 96–103, 2016.

[21] M. C. Scott et al., "Electron tomography at 2.4-ångström resolution," *Nature*, vol. 483, pp. 444–447, 2012.

[22] W. H. Press et al., *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge, England, 3 edition, 2007.

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

4

Parameter Selection for TV-Regularized Electron Tomography

In this chapter, we propose an efficient approach to automatically select the regularization parameter λ for the popular total variation (TV) minimization reconstruction technique [1]. Since reconstructions in electron tomography are preferably done in three dimensions (3D) [2], a 3D metric that can incorporate the information in all directions is also favored to quantify the reconstruction quality. To this end, we use the 3D Gaussian gradient magnitude because this easy-to-implement metric is mathematically related to the TV regularization term, namely the l_1- or l_2-norm of the gradient of reconstruction [3]. Furthermore, considering that 3D TV reconstruction is time-consuming, and that the reconstructed images change smoothly with respect to λ [4], we adopt spline interpolation to approximate the curve of the sum of gradient magnitude versus λ so as to reduce the computational time[1].

We evaluate our method on five experimental HAADF tilt-series of different nanoparticles, three of which consist of gold and the other two gold (Au) and silver (Ag). Since Au and Ag are well-separated in the latter two nanoparticles and have substantially different atomic numbers ($Z_{\mathrm{Au}} = 79$, $Z_{\mathrm{Ag}} = 47$), HAADF-STEM tomography with its high Z-contrast is good enough to discriminate Au from Ag without EDS [5]. Results show that our method is able to quantify the strength of TV regularization and find a balance between noise suppression and structure blurring without a reference image. The aforementioned spline interpolation helps to accelerate the reconstruction-assessment pipeline.

This chapter is organized as follows. In Section 4.1, we introduce the TV regularized reconstruction technique and a few existing methods for automatic regularization parameter selection. Then, we present our own method in Section 4.2, followed by experiments and results in Section 4.3. Finally, Section 4.4 summarizes the work and draws conclusions.

4.1. TV-REGULARIZED ELECTRON TOMOGRAPHY

Total variation (TV) minimization in compressed sensing assumes that the image to be reconstructed has a sparse representation in a specific base [1, 6]. Observing that the gradient of nanostructured materials is often sparse, Goris et al. incorporated TV regularization in simultaneous iterative reconstruction technique (SIRT) to reduce missing wedge artifacts [1]. Let $\mathbf{p} \in \mathbb{R}^{M\times 1}$ and $\mathbf{x} \in \mathbb{R}^{N\times 1}$ be the (vectorized) projection and reconstruction, respectively, in which M is the number of rays and N the number of discretized unknown voxels. The TV regularized reconstruction technique extends the standard l_2-norm reconstruction in the following form [1]

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \left\{ \frac{1}{2} \left\| \mathbf{W}\mathbf{x} - \mathbf{p} \right\|_2^2 + \lambda \mathrm{TV}(\mathbf{x}) \right\} \tag{4.1}$$

where $\mathbf{W} \in \mathbb{R}^{M\times N}$ is the projection matrix, whose entries w_{mn} are determined by the intersected area between the m-th ray integral and n-th voxel. Since Eq. (4.1) simultaneously minimizes the data fidelity term (i.e., the distance between the projection $\mathbf{p}$ and reconstruction $\mathbf{x}$) and the total variation of $\mathbf{x}$, it prefers sharp transitions between spe-

[1] Throughout this chapter, computational time refers to the amount of time spent on reconstruction and assessment.

cific gray values over gradual changes in $\mathbf{x}^*$ [1]. The regularization term TV($\mathbf{x}$) is defined as the l_1- or l_2-norm of the gradient of $\mathbf{x}$, that is,

$$\begin{aligned} \mathrm{TV}_{l_1}(\mathbf{x}) &= \|\nabla\mathbf{x}\|_1 = \sum_n |\nabla x_n| \\ \text{or} \quad \mathrm{TV}_{l_2}(\mathbf{x}) &= \|\nabla\mathbf{x}\|_2 = \sqrt{\sum_n (\nabla x_n)^2}, \end{aligned} \tag{4.2}$$

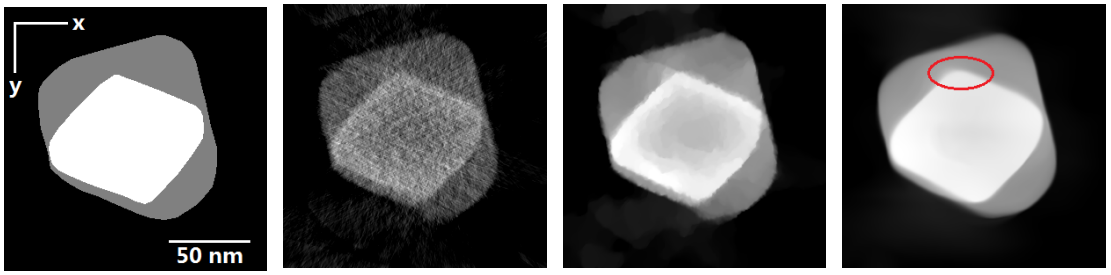
in which $\nabla x_n = \left[\partial x_n/\partial x, \partial x_n/\partial y, \partial x_n/\partial z\right]^T$ is the discrete gradient of the n-th voxel [3]. Note that the l_1- and l_2-norm TV are also called anisotropic and isotropic TV, because the former encourages horizontal and vertical edges and the latter, conversely, is rotation invariant.

The parameter λ in Eq. (4.1) determines the regularization strength of TV($\mathbf{x}$). A large λ may result in over-regularization that blurs high frequency details, whereas a small one may lead to under-regularization that lets noise and missing wedge artifacts pass [1]. A few approaches exist to automatically set this crucial parameter [4, 7–9]. For instance, Hansen and O'Leary chose the "optimal" λ such that the log-log curve of the data fidelity term and regularization term has the maximal curvature [9]. Moreover, if a reference image such as hand-segmented ground truth is accessible, one may compare the computed reconstructions to the reference using correlation coefficient, root mean squared error (RMSE), or segmentation error [3, 4, 10]. However, these methods are hindered for industrial applications because: (*i*) reference images are mostly unavailable; (*ii*) TV, as an iterative algorithm, takes a long time to compute reconstructions $\mathbf{x}^*$ for a large set of λ; (*iii*) λ values found by the methods mentioned above can vary a lot if application-specific properties of the reconstruction are not taken into account. Therefore, in the next section we present an effective no-reference 3D image quality metric to automatically find a proper value for λ.

4.2. AN EFFECTIVE REGULARIZATION PARAMETER SELECTION METHOD

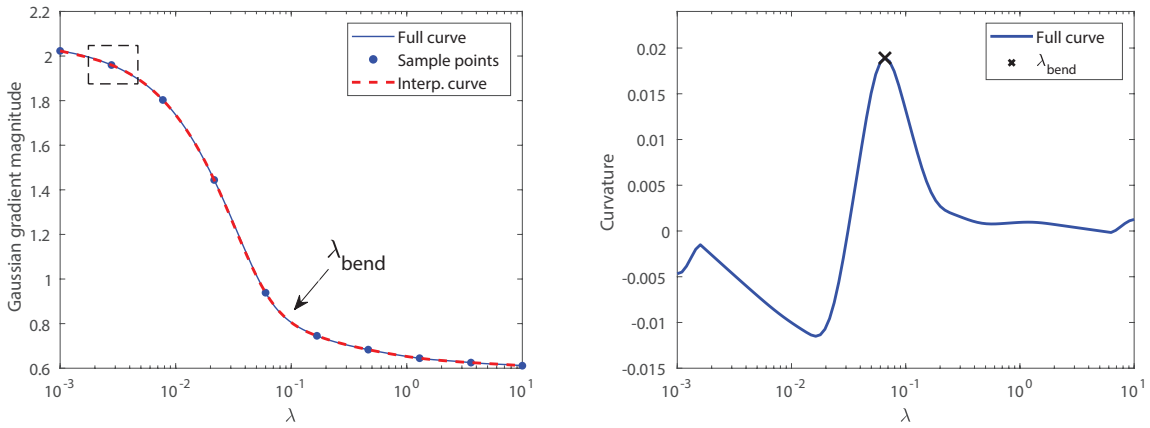
To describe our method, we use the same experimental dataset as in [11, Sec. 4.2], which is a core-shell nanoparticle consisting of Au and Ag. Fig. 4.1 shows the central orthoslices of its TV reconstructions computed with 300 iterations, l_2-norm, and different λ values. The image in the first column is the hand-segmented ground truth that has only three intensity values [5]: 0 (background), 1 (Ag), and 2 (Au). Note that these values are not weighed by the HAADF-STEM response factors of Au and Ag [12]. For $\lambda = 10^{-3}$, TV reconstruction resembles SIRT; hence, the discontinuous foreground may make automatic segmentation far from straightforward [13]. For $\lambda = 10^{0.31}$, a strong regularization deforms the interface between Au and Ag, see the red circle in Fig. 4.1(d).

In Fig. 4.2(a), the blue line plots the sum of the 3D Gaussian gradient magnitude ($\sigma = 1$ voxel) of TV reconstructions for 100 λ values, which are uniformly distributed between 10^{-3} and 10^{1} on the logarithmic scale. This continuous curve has a clear bending point λ_{bend} with the maximal curvature. Therefore, we empirically define λ_{bend} as the "ideal" regularization parameter because when $\lambda > \lambda_{\text{bend}}$, regularization starts to blur the edges as no noise is left. Intuitively, one may choose the second derivative to com-



(a) Ground truth (b) $\lambda = 10^{-3.0}$ (c) $\lambda = 10^{-0.86}$ (d) $\lambda = 10^{0.31}$

Figure 4.1: Central orthoslices for the Au-Ag core-shell nanoparticle in [11, Sec. 4.2]. (a) Ground truth, hand-segmented from SIRT reconstructions [5]; (b)–(d) TV_{l_2} reconstructions computed with 300 iterations and different regularization parameter λ. The size of the image is $300 \times 300 \times 300$ voxels.



(a) Sum of 3D Gauss. grad. mag. (b) Curvature

Figure 4.2: (a) Sum of 3D Gaussian gradient magnitude of TV_{l_2} reconstruction versus regularization parameter λ. Full curve is plotted with 100 λ values uniformly distributed on the logarithmic scale, ten of which are used for spline interpolation. We define the bending point λ_{bend} as the "ideal" λ. (b) Curvature of the full curve computed using the proposed sliding window approach.

pute the curvature and find λ_{bend}. Nevertheless, this simple technique is not suitable for our case because it only considers the current point λ_i and the adjacent λ_{i-1} and λ_{i+1}. To include more neighbors, we use a sliding window approach that has a size of $2w+1$. For each λ_i, we first rotate the arc in $[\lambda_{i-w}, \lambda_{i+w}]$ so that it either opens upwards or downwards depending on its convexity. Then, we approximate it to a parabola $p_2 x^2 + p_1 x + p_0$ and define p_2 as the curvature of the arc; the result is shown in Fig. 4.2(b) where the fluctuation near 10^{-3} and 10^{1} is due to boundary artifact. Finally, we find λ_{bend} that has the highest p_2.

λ_{bend} is dependent on the window size: the larger w is, the higher $\lambda^{w}_{\text{bend}}$ will be (e.g., $\lambda^{w=5}_{\text{bend}} = 10^{-1.18}$ and $\lambda^{w=35}_{\text{bend}} = 10^{-0.66}$). To make our algorithm window size-independent, we modify it as follows:

1. set $[w_L, w_H]$ as the range for w and compute the corresponding $\lambda^{w_L}_{\text{bend}}$ and $\lambda^{w_H}_{\text{bend}}$;
2. compute the distance from each $\lambda_i \in [\lambda^{w_L}_{\text{bend}}, \lambda^{w_H}_{\text{bend}}]$ to the line segment that connects $\lambda^{w_L}_{\text{bend}}$ and $\lambda^{w_H}_{\text{bend}}$;
3. return λ_i with the maximum distance as the desired λ_{bend}.

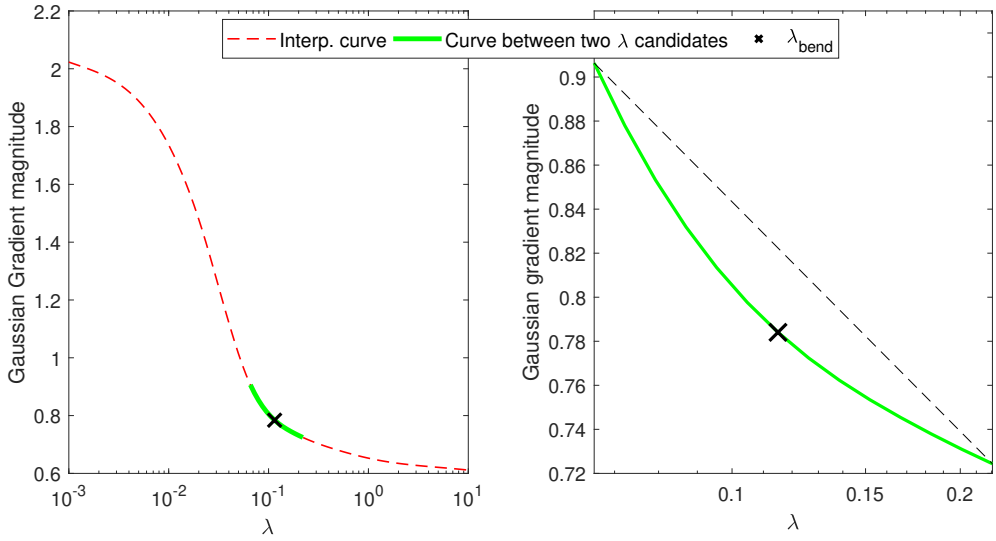
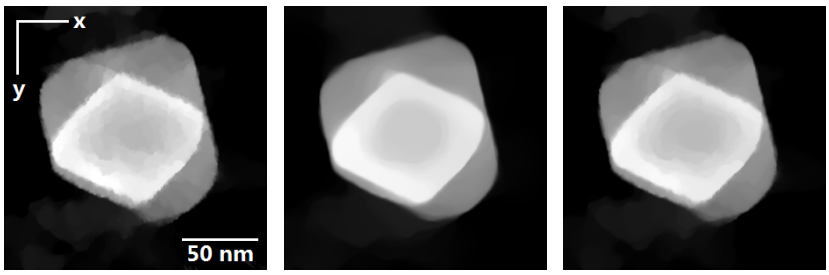


Figure 4.3: (Left) Interpolated approximation of the sum of 3D Gaussian gradient magnitude in Fig. 4.2(a) and (right) zoom-in between $[\lambda_{\text{bend}}^{w_L}, \lambda_{\text{bend}}^{w_H}]$.

Moreover, considering that 3D TV reconstruction is time-consuming especially for a large set of λ, and that the curve of the sum of Gaussian gradient magnitude versus λ is smooth, we adopt interpolation methods to speed up the reconstruction-assessment process. The red dashed line in Fig. 4.2(a) shows an example of spline interpolation that uses only ten equidistant samples; normalized RMSE between the full curve and its interpolation is 1.36×10^{-4}. Fig. 4.3 illustrates how our algorithm works.



(a) Our method (b) CC (c) Seg. error

Figure 4.4: Central orthoslices of TV_{l_2} reconstructions for the Au-Ag core-shell nanoparticle in [11, Sec. 4.2]. "Optimal" regularization parameter $\lambda \in \{10^{-0.94}, 10^{-0.17}, 10^{-0.70}\}$, which are respectively found by (a) our method, (b) correlation coefficient (CC), and (c) segmentation error (Seg. error).

We compare our no-reference metric to two full-reference metrics: correlation coefficient [3]

$$\text{CC} = \frac{\sum_n (x_n - \bar{x})(g_n - \bar{g})}{\sqrt{\sum_n (x_n - \bar{x})^2 \sum_n (g_n - \bar{g})^2}} \tag{4.3}$$

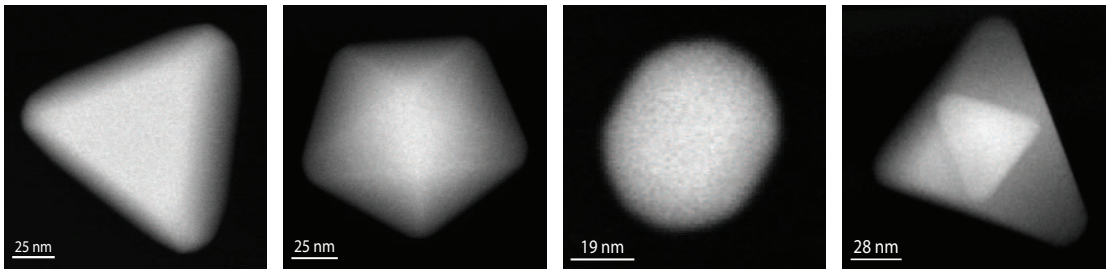
and segmentation error [10, 14]

$$\text{Seg. error} = \frac{\sum_n |s_n - g_n|}{\sum_n g_n}, \tag{4.4}$$

in which x_n, g_n and s_n denote the n-th voxel of the gray-scaled reconstruction, hand-segmentation and the image segmented from $\mathbf{x}$ using (multilevel) Otsu's method [15]. We do not consider the RMSE because intensities of the reference image in Fig. 4.1(a) does not involve the interaction between the atoms and the incident electron beam. Fig. 4.4 shows the reconstruction slices computed with the optimal λ found by the three metrics; visually, our result is comparable with the other two. However, different methods return different λ values, depending on whether they consider application-specific properties of the reconstruction (e.g., segmentation error); this is also in line with the observation in [4].

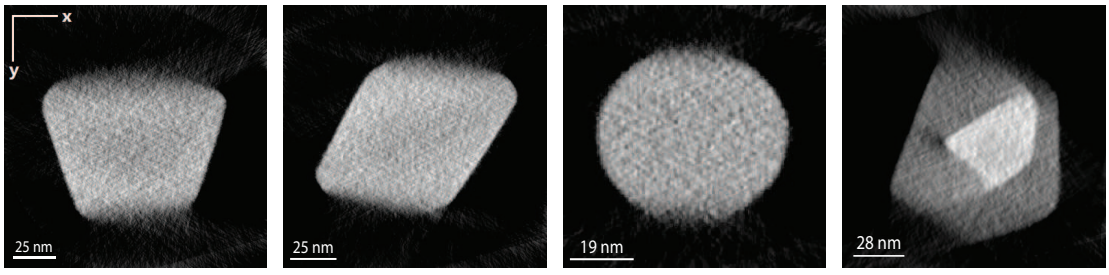
4.3. EXPERIMENTS AND RESULTS

In this section, we evaluate our method on four other HAADF tilt-series. An Au tetrahedron, an Au pentagon platelet, an Au sphere, and an Au-Ag core-shell tetrahedron were scanned in an aberration-corrected Thermo Fisher Scientific Titan S/TEM, which was operated with an accelerating voltage of 300 kV, a beam current of 50 pA and a dwell time of 1 μs. HAADF-STEM projections were acquired from -75° to $+75^\circ$ with an increment of 3°. Fig. 4.5 gives an example of the experimental projections, and Fig. 4.6 shows their xy-slices at different z reconstructed with SIRT using 100 iterations. Since the sample rotation is limited to $\pm 75^\circ$, SIRT reconstructions are more elongated in the direction of the missing wedge.



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

Figure 4.5: Experimental HAADF-STEM projections of four nanoparticles at 0°: (a) Au tetrahedron, (b) Au pentagon platelet, (c) Au sphere, and (d) Au-Ag core-shell tetrahedron. The size of the projections is 384×384, 352×352, 192×192, and 320×320 pixels (3.86 Å/pixel).



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

Figure 4.6: Central orthoslices of SIRT reconstructions for the four new datasets (see Fig. 4.5). The size of the image is (a) $384 \times 384 \times 384$, (b) $352 \times 352 \times 352$, (c) $192 \times 192 \times 192$, and (d) $320 \times 320 \times 320$ voxels.

4.3.1. IMPLEMENTATION DETAILS

Throughout this chapter, we use a primal-dual hybrid gradient (PDHG) method [16] provided in the operator discretization library [17] to compute TV reconstructions. In order to ensure convergence, we set the primal and dual step sizes as $\tau = {}^{0.1}/_{\|L\|}$ and $\sigma = {}^{0.99}/_{\tau\|L\|^2}$, respectively, in which L is an operator related to the PDHG implementation [4]. Moreover, we normalize the regularization parameter λ to $\hat{\lambda} = \lambda\|\nabla\|/\|\mathbf{W}\|$ for scaling the data fidelity term and regularization term to the same range [4]; hence, Eq. (4.1) becomes

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \left\{ \frac{1}{2} \frac{\left\|\mathbf{W}\mathbf{x} - \mathbf{p}\right\|_2^2}{\|\mathbf{W}\|} + \hat{\lambda} \frac{\|\nabla\mathbf{x}\|_{1/2}}{\|\nabla\|} \right\}. \tag{4.5}$$

This normalized $\hat{\lambda}$ will be referred to as the regularization parameter hereinafter.

Given the number of iterations $\mathcal{N} = 300$, this implementation takes ~16 h to compute 100 reconstructions with a size of $320 \times 320 \times 320$ voxels, using Python 3.6.0 on a high performance compute server equipped with two Intel E5-2670 v3 CPUs[2] and four NVIDIA Tesla K40c GPUs[3]. Therefore, the idea of interpolation (see Section 4.2) is certainly attractive to reduce the computational time.

4.3.2. RESULTS

Fig. 4.7 depicts the sum of the Gaussian gradient magnitude of l_2-norm regularized TV reconstructions as a function of λ, in which the full curve computed with 100 λ points and the spline interpolation with only 10 samples are almost the same. Quantitatively, the normalized RMSE between the two curves is 7.84×10^{-4} for Au tetrahedron, 8.68×10^{-4} for Au platelet, 5.95×10^{-4} for Au sphere, and 3.91×10^{-4} for Au-Ag tetrahedron.

We use the interpolated approximation to find λ_{bend}. Fig. 4.8 gives an example of ten TV reconstructions for an Au-Ag tetrahedron, which are computed with the first nine λ samples in Fig. 4.7(d) and λ_{bend}, respectively. Visually, we can see that the "optimal" λ should lie in $[10^{-1.22}, 10^{-0.78}]$ where noise is gradually decreasing and structure maintaining; $\lambda_{\text{bend}} = 10^{-1.14}$ found by our approach has this trade-off between noisiness and sharpness.

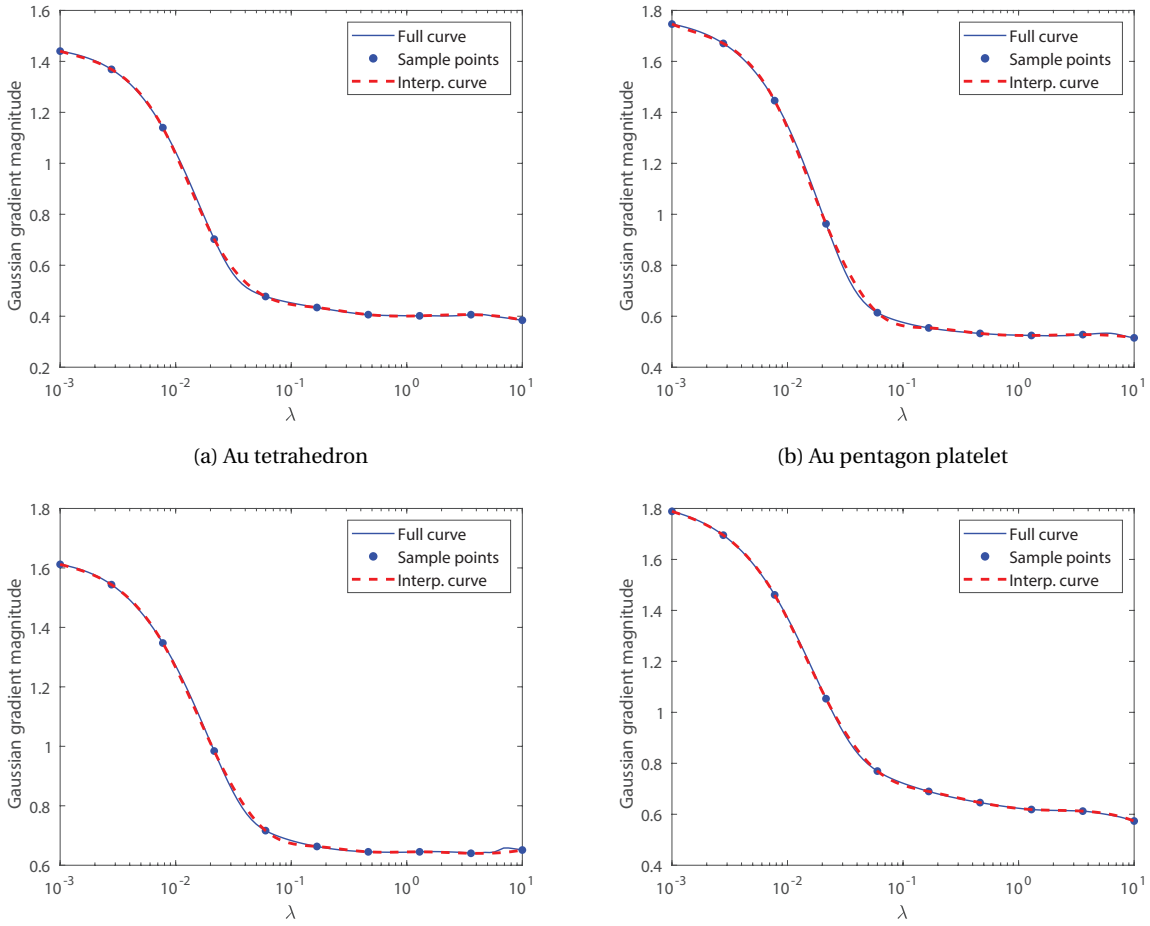
Fig. 4.9 shows the central orthoslices of TV reconstructions for four datasets. The resulting reconstructions, whether are based on l_1- or l_2-norm TV, do not have perceivable noise nor blurry edges. Furthermore, the anisotropic l_1-norm generally results in smaller λ because it has stronger regularization strength given the same λ, see Eq. (4.2). Since no reference image is available, we cannot compute the correlation coefficient and segmentation error for comparison.

4.4. DISCUSSION AND CONCLUSION

In this chapter, we used the simple sum of the 3D Gaussian gradient magnitude as an effective metric to automatically find a proper regularization parameter for the TV-based reconstruction technique. We demonstrated it on five experimental HAADF tilt-series of different nanoparticles and showed that it can find the balance between noise reduction

[2] 48 cores in total and 256 GB RAM, with one core being used to 100% and ~1.6% memory occupied.

[3] Each GPU has 12 GB dedicated memory, only used to create the forward operator with ASTRA CUDA [18], which is supported in ODL [17] as back-end.



(a) Au tetrahedron

(b) Au pentagon platelet

(c) Au sphere

(d) Au-Ag tetrahedron

Figure 4.7: Sum of 3D Gaussian gradient magnitude of TV_{l_2} reconstructions versus λ and the spline interpolation for (a) Au tetrahedron, (b) Au pentagon platelet, (c) Au sphere, and (d) Au-Ag tetrahedron.

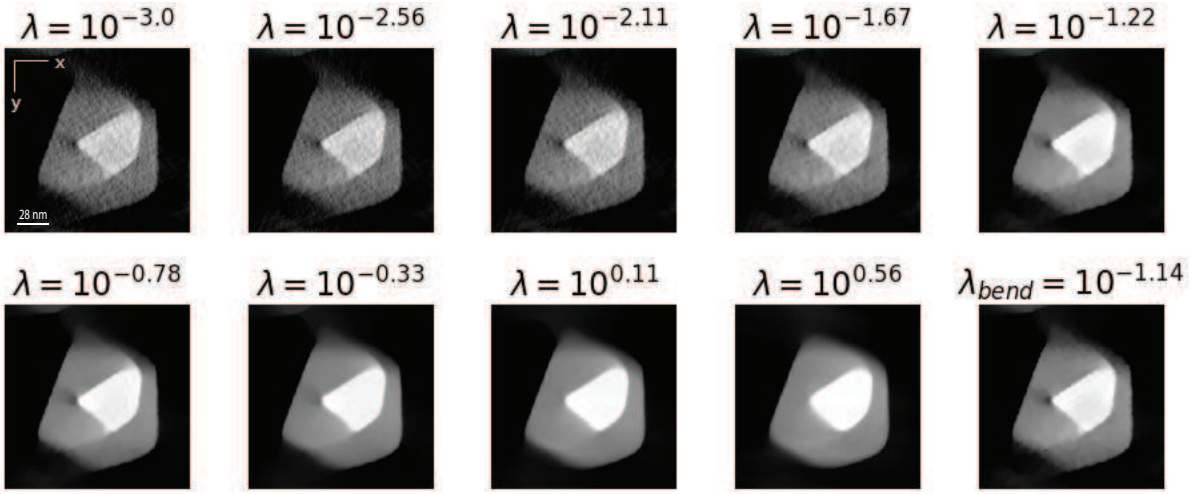
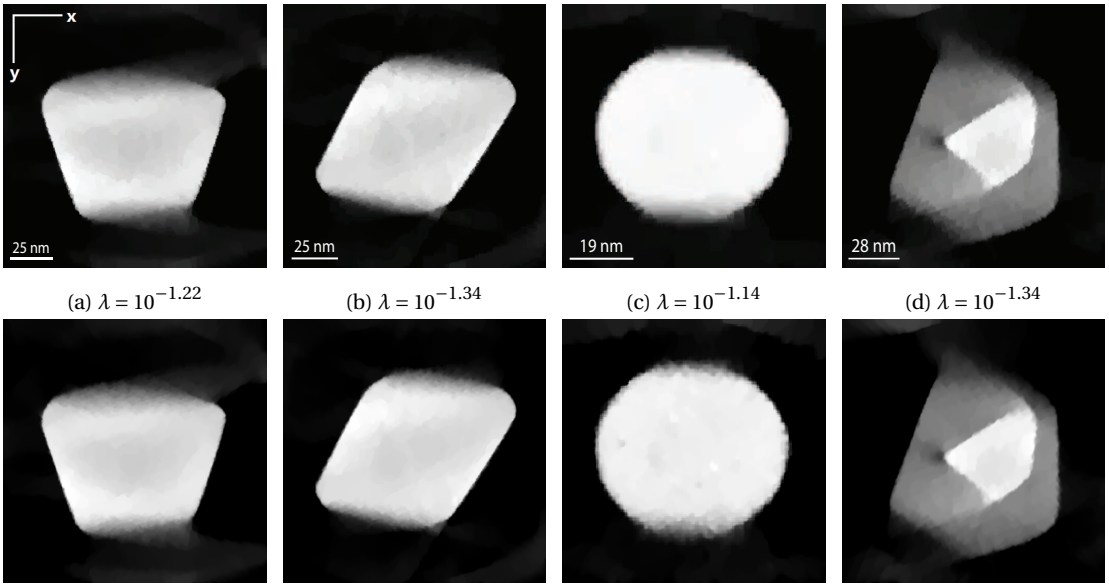


Figure 4.8: Central orthoslices of ten TV_{l_2} reconstructions for an Au-Ag tetrahedron; computed with the first nine λ samples in Fig. 4.7(d) and λ_{bend}, respectively.



(a) $\lambda = 10^{-1.22}$ (b) $\lambda = 10^{-1.34}$ (c) $\lambda = 10^{-1.14}$ (d) $\lambda = 10^{-1.34}$

(e) $\lambda = 10^{-1.02}$ (f) $\lambda = 10^{-1.10}$ (g) $\lambda = 10^{-1.14}$ (h) $\lambda = 10^{-1.14}$

Figure 4.9: TV reconstructions regularized with (upper) l_1-norm and (lower) l_2-norm; computed with 300 iterations and the λ_{bend} found by our method. Volume size see Fig. 4.6.

and edge blurring without any reference. Compared to the 3D iterative reconstruction, time spent for magnitude quantification is minor, e.g., 16 h versus 5 min (details see Section 4.3.1); hence, we adopted spline interpolation to approximate the curve of the sum of gradient magnitude versus regularization parameter to accelerate the reconstruction-assessment pipeline.

While the first results look promising, we note that our method should be further tested on a broader scale of experimental data. On the one hand, despite different geometries, all samples are homogeneous with very few high spatial frequency components; on the other hand, they were imaged under the same or similar microscope settings. As a result, it is very important to evaluate our method on other experimental datasets of inhomogeneous samples with more high spatial frequency components and/or acquired under different microscope settings.

REFERENCES

[1] B. Goris et al., "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012.

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

[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*, vol. 191, pp. 34–43, 2018.

[4] M. J. Lagerwerf, W. J. Palenstijn, F. Bleichrodt, and K. J. Batenburg, "An efficient interpolation approach for exploring the parameter space of regularized tomography algorithms," *Fundamenta Informaticae*, vol. 172, pp. 143–167, 2020.

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

[6] T. Sanders, A. Gelb, R. B. Platte, I. Arslan, and K. Landskron, "Recovering fine details from under-resolved electron tomography data using higher order total variation l_1 regularization," *Ultramicroscopy*, vol. 174, pp. 97–105, 2017.

[7] G. M. Vainikko, "The discrepancy principle for a class of regularization methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 22, pp. 1–19, 1982.

[8] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Review*, vol. 34, pp. 561–580, 1992.

[9] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 14, pp. 1487–1503, 1992.

[10] H. Vanrompay et al., "Experimental evaluation of undersampling schemes for electron tomography of nanoparticles," *Particle & Particle Systems Characterization*, vol. 36, pp. 1900096, 2019.

[11] Y. Guo and B. Rieger, "Parameter selection for regularized electron tomography without a reference image," in *Proceedings of the 2019 Scandinavian Conference on Image Analysis*, 2019, pp. 452–464.

[12] M. M. J. Treacy, "Z dependence of electron scattering by single atoms into annular dark-field detectors," *Microscopy and Microanalysis*, vol. 17, pp. 847–858, 2011.

[13] B. Goris et al., "Advanced reconstruction algorithms for electron tomography: from comparison to combination," *Ultramicroscopy*, vol. 127, pp. 40–47, 2013.

[14] E. Bladt, D. M. Pelt, S. Bals, and K. J. Batenburg, "Electron tomography based on highly limited data using a neural network reconstruction technique," *Ultramicroscopy*, vol. 158, pp. 81–88, 2015.

[15] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, 1979.

[16] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120–145, 2011.

[17] J. Adler, H. Kohr, and O. Öktem, "Operator discretization library (ODL)," https://github.com/odlgroup/odl, 2017.

[18] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA toolbox: a platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.

5

HAADF-EDS Cross-Modal Fusion for Electron Tomography

In this paper, we present a multichannel cross-modal fusion algorithm to combine two complementary modalities in electron tomography: X-ray spectroscopy and scanning transmission electron microscopy (STEM). The former reveals compositions with high elemental specificity but low signal-to-noise ratio (SNR), while the latter characterizes structure with high SNR but little chemical information. We use multivariate regression to build a cross-modal fusion framework for these two modalities to simultaneously achieve high elemental specificity and high SNR for a target element chosen from the sample under study. Specifically, we first compute three-dimensional tomograms from tilt-series datasets of X-ray and STEM using different reconstruction algorithms. Then, we generate many feature images from each tomogram. Lastly, we adopt partial least squares regression to assess the connection between these feature images and the reconstruction of the target element. Based on simulated and experimental datasets of semiconductor devices, we demonstrate that our algorithm can not only produce continuous edges, homogeneous foreground and clean background in its element-specific reconstructions, but also more accurately preserve fine structures than state-of-the-art tomography techniques. Moreover, we show that it can deliver results with high fidelity even for X-ray datasets with limited tilts or low counts. This property is highly desired in the semiconductor industry where acquisition time and sample damage are essential.

Published in *IEEE Trans. Image Process.*, vol. 28, pp. 4206–4218, 2019 [1].

5.1. INTRODUCTION

Electron tomography is a powerful tool in materials science to characterize the complex three-dimensional (3D) structure of inorganic specimens on the nanoscale [2]. In transmission electron microscopy (TEM), the sample under study is exposed to an electron beam and tilted to obtain two-dimensional (2D) projection images at different angles. Several imaging modalities exist, such as bright-field TEM [3] and high-angle annular dark-field scanning transmission electron microscopy (HAADF-STEM) [4]. In tomography, these projections are called a tilt-series, from which we can reconstruct a volume representing the sample [2]. Since the intensity of HAADF-STEM scales with the atomic number Z of the element ($\sim Z^n$, $n = 1.6 - 1.9$, depending on the inner and outer detector angles [5]), it can also indirectly reveal the compositional information of the sample. However, when the sample has elements with close atomic numbers, HAADF-STEM images may no longer be distinctive for these elements. To better understand more complex compositions, spectral imaging techniques like energy dispersive X-ray spectroscopy (EDS) [5] must be pursued.

Tomographic reconstruction is an ill-posed inverse problem because of the inevitable noise in the measurements such as Poisson noise and readout noise [6]; reconstruction becomes even more problematic for electron tomography where the number of projections is much smaller than the 3D volume [7]. Consequently, its solution might not be stable and unique. So far, dozens of reconstruction techniques have been proposed, and the classical filtered backprojection (FBP) is still frequently applied in practice thanks to its simplicity and speed [8]. Alternatively, iterative algorithms (e.g., simultaneous iterative reconstruction technique (SIRT) and its variants [9]) have also attracted large attention as they produce less artifacts for noisy datasets [2]. Moreover, prior knowledge has been incorporated to further enhance the reconstruction quality [7, 10–13]. For instance, assuming that the sample of interest has piecewise constant structures, Goris et al. incorporated total variation regularization into SIRT and effectively reduced missing wedge artifacts in the reconstruction [7]. While conventionally only a single modality and/or tilt-series is used in a reconstruction algorithm, advanced approaches tend to combine two (or more) datasets from multiple modalities for integrating complementary information [14–16]. Bimodal tomography [15], for example, links HAADF-STEM and EDS projections into a joint reconstruction scheme. The former modality is at atomic resolution with high signal-to-noise ratio (SNR) but not intrinsically element-specific, and the latter, conversely, is rich in chemical information but suffers from low SNR. Although great efforts have been dedicated to improving reconstruction techniques, directly combining reconstruction volumes at hand has still not been widely considered, to the best of our knowledge.

To generate a composite image benefiting from different modalities and/or reconstruction algorithms, one can employ various statistical approaches that project high-dimensional inputs onto low-dimensional outputs. With the capability of reducing redundancies yet highlighting similarities and differences, statistical methods have been broadly investigated and applied in multimodal image fusion, that is, multiple input images of different modalities are fused into a single output. For example, many infrared and visible image fusion frameworks have incorporated principal component analysis (PCA) for decorrelation [17]. As part of these methods, regression techniques can not

only decompose two sets of input images into lower dimensions but also assess their connections. For instance, with partial least squares (PLS) regression, van de Plas et al. built a linear model to fuse data obtained from mass spectrometry and optical microscopy for studying protein, peptide, lipid, and drug distributions in tissues [18].

Recently, we extended the concept of multimodal fusion to electron tomography [19]. Specifically, we adopted and modified the regression-based cross-modality modeling of van de Plas [18] to fuse X-ray and electron tomograms for reconstructing bi-elemental nanomaterials. We demonstrated our method on an experimental dataset of a core-shell nanoparticle (consisting of gold and silver with distinct atomic numbers of 79 and 47, respectively) and showed that it enabled reconstructions with sharper edges and smoother fore- and background than bimodal tomography [15]. However, the extension to more complex multi-elemental nanostructures has still been lacking thus far. Therefore, in this paper, we first present a 3D multichannel cross-modal fusion algorithm based on our preliminary work in 2D [19]. Then, we validate it on simulated and experimental datasets of semiconductor devices. In particular, we investigate the fusion quality for a small number of EDS acquisition angles and low-dose EDS maps. EDS maps with high SNR are currently hampered by slow data acquisition in practice, and hence algorithms that can deal with lower SNR are desired.

The rest of this paper is organized as follows. In Section 5.2, we briefly review the classical and state-of-the-art electron tomography along with statistical methods that have been employed in multimodal image fusion. Section 5.3 introduces our six-step cross-modal fusion framework and Section 5.4 the experimental setup, including simulation and experimental datasets, and quantitative assessment procedures. Our results are presented in Section 5.5 and discussed in Section 5.6. Finally, we draw conclusions in Section 5.7.

5.2. CONTEXT AND PRIOR ART

5.2.1. ELECTRON TOMOGRAPHY

In this section, we refer to algorithms that jointly reconstruct a volume from multiple modalities as “multichannel" and “single-channel" otherwise [20].

SINGLE-CHANNEL

Single-channel approaches are either analytical or iterative. Analytical algorithms are based on the Fourier slice theorem and directly calculate the reconstruction in a single step. While FBP is the most commonly employed, it aggravates thin streaks in the reconstruction if the number of projections is low. To alleviate such artifacts yet preserve desirable fine structures, Jin et al. trained a convolutional neural network to regress the FBP results towards a ground truth image [21]. Although the authors demonstrated the proposed FBPConvNet on sparse-view X-ray projections, they stressed that it could be generalized to other modalities. Iterative algorithms, however, formulate the reconstruction problem as a large under-determined linear system and solve it iteratively [6]. Common methods include algebraic reconstruction technique (ART), simultaneous iterative reconstruction technique (SIRT), simultaneous algebraic reconstruction technique (SART), and maximum likelihood expectation-maximization (ML-EM) [9]. With the capability of incorporating various types of prior knowledge, iterative methods are more

robust to deal with ill-posed inverse problems. For instance, assuming that the sample under study only consists of a few elements, that is, the reconstruction only has a few discrete intensity levels, Batenburg and Sijbers developed discrete algebraic reconstruction technique (DART) to achieve a more accurate reconstruction from limited and/or noisy projections [10]. Furthermore, inspired by l_1 regularization in compressive sensing, Goris et al. combined the popular total variation regularization (TVR) with SIRT and proposed total variation minimization (TVM) reconstruction to compensate for the missing wedge in electron tomography [7]. TVR was also combined with DART, such that the TVR-DART would require less tuning parameters [11]. Alternative l_1 regularization approaches, such as higher order total variation (HOTV) [12, 13], have also been investigated. Since the solution of HOTV-based reconstruction algorithm is not limited to a piecewise constant function, it could more effectively recover fine features than the common TV [12].

MULTICHANNEL

In general, multichannel algorithms that simultaneously couple datasets from multiple sources are applied either in multispectral or multimodality reconstruction. In electron tomography, most methods combine the complementary information of HAADF-STEM and EDS projections. For example, Zanaga et al. used HAADF-STEM to supplement EDS for improving its shadowing effects and lower spatial resolution, thereby enabling a more reliable EDS quantification [14]. Zhong et al. introduced a manual parameter to weigh the HAADF-STEM and EDS channels in bimodal tomography for effectively suppressing noise and enhancing contrast [15]. Regularizations, such as total nuclear variation (TNV) derived from TV, have also been considered [16]. Different from TV that only promotes sparse gradients in the EDS reconstructions, TNV further incorporates HAADF-STEM to encourage anti-/parallel gradients for enforcing common edges in the joint reconstructions [16].

Compared with zero- and single-parameter algorithms (e.g., FBP and SIRT), advanced techniques may yield a more accurate result using a limited set of noisy projections. Many of them, however, involve extra parameters that need to be carefully tuned. For instance, the weighting factor α in bimodal tomography [15] is currently chosen by comparing reconstructions over the whole range of $\alpha \in (0,1)$ to a hand-segmented ground truth, which is very impractical.

5.2.2. STATISTICAL METHODS APPLIED TO MULTIMODAL IMAGE FUSION

According to Sui et al., statistical approaches applied to multimodal image fusion are either driven by data or by hypotheses [22].

DATA-DRIVEN METHODS

Data-driven statistical methods include, but are not limited to, principal component analysis (PCA), independent component analysis (ICA), and nonnegative matrix factorization (NMF). PCA has been a common initializer for fusion algorithms in brain imaging to effectively remove redundancies [22]. It was also improved to robust PCA [23] and adopted to suppress noise yet retain useful information for infrared and visible image fusion [17]. As an extension of PCA, ICA can separate correlated input images

into independent components. ICA-aided fusion algorithms usually involve other techniques like multiscale transforms. For instance, Ghahremani and Ghassemian chose ICA and curvelet transform to reduce the spectral distortion of pan-sharpened multispectral bands [24]. Moreover, incorporating ICA into training-based algorithms is also popular, in which a set of ICA bases are trained from patches with similar contents as the source images [17]. Different from PCA, NMF only allows additive, not subtractive, combinations due to its nonnegativity constraints [25]. As a result, it represents parts of the objects and corresponds better to the human perception mechanism. Wang et al. further extended the traditional NMF to a nonnegative sparse representation (NNSR) model for fusing infrared and visible images [26]. As the NNSR emphasizes not only the nonnegativity but also sparsity of the coefficients, it can achieve a rational (only with nonnegative intensities) and convenient (with just a few sparse components) image interpretation.

HYPOTHESES-DRIVEN METHODS

Hypotheses-driven statistical methods, such as regression, can characterize the relationship among source images. In remote sensing, for instance, multivariate regression has been a powerful tool to merge multispectral (MS) and panchromatic (PAN) images. The former captures visible light in a small number of spectral bands at low resolution, and the latter is sensitive to all wavelengths of the light at high resolution. In this case, regression is adopted to estimate the weights between the MS channels and PAN image at both the reduced [27] and full scale [28] for pansharpening. As a well-studied member in the regression family, partial least squares (PLS) establishes a linear multivariate model to relate the inputs [29]. It was first applied to multimodal fusion in neuroimaging by Martínez-Montes et al. to concurrently analyze electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) data [30]. Since then, PLS has been broadly employed for multimodal fusion not only in brain imaging [22] but also in biomedical [18] and chemical imaging [31]. For instance, van de Plas et al. [18] chose this multivariate regression technique to fuse mass spectrometric and microscopic images. With a series of case studies, they showed that the proposed modeling could maintain both the high chemical specificity and high spatial resolution.

5.3. METHOD

Let us assume that a sample has a number of E different chemical elements that make up the element set $\mathscr{E}$, and that its HAADF-STEM projections are acquired at a number of P_{H} angles. Meanwhile, X-ray spectral projections are recorded at a number of P_{E} angles and deconvolved into a number of E EDS maps; each corresponds to one chemical element $e \in \mathscr{E}$. Note that the spectral images usually have fewer tilts than the HAADF projections ($P_{\mathrm{E}} < P_{\mathrm{H}}$) and suffer from much lower SNR. Binning may effectively increase the number of X-ray counts collected in EDS maps but also degrade their spatial resolution [32].

The proposed fusion algorithm has six steps, where the first five are illustrated in Fig. 5.1:

1. compute multiple tomograms from HAADF and each EDS tilt-series using a number of N different reconstruction algorithms;
2. *i*) check visibility of all elements and select a target element $e^\star$ for fusion;
   ii) denoise the EDS reconstructions of $e^\star$;

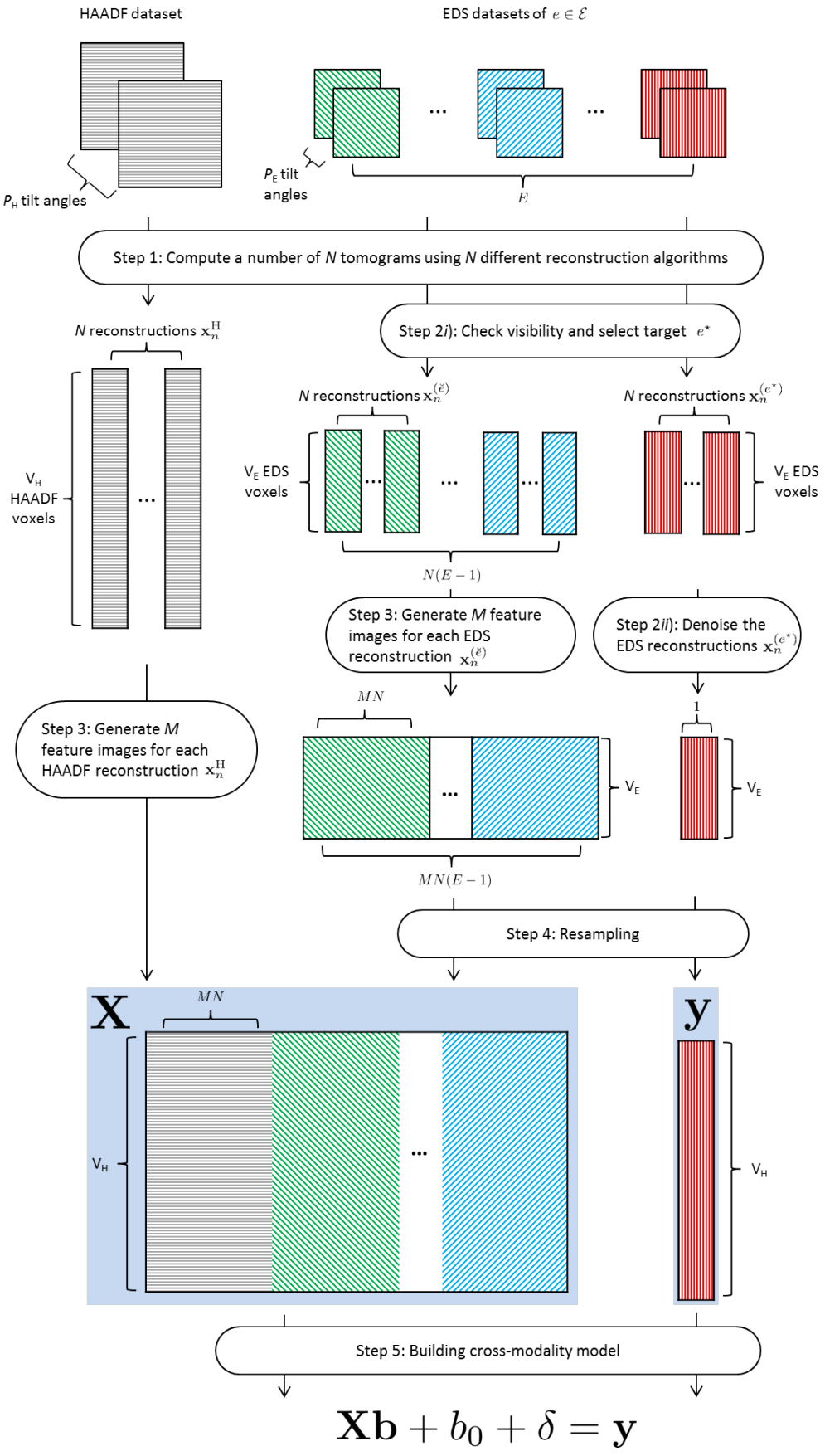


Figure 5.1: The first five steps of fusion algorithm. Assume the sample of interest is composed of a set of chemical elements $\mathscr{E}$. Denote the target element chosen for fusion as $e^\star$ and the rest $\check{e} : e \in \mathscr{E} \setminus e^\star$. H symbolizes HAADF and E EDS. Details in Section 5.3.

3. generate a number of M feature images for each HAADF reconstruction, and each EDS reconstruction of $\check{e} : e \in \mathscr{E} \setminus e^\star$;
4. (optional) upsample all EDS-related images if the X-ray spectral images have been binned;
5. build a cross-modality model between the feature images and the denoised EDS reconstruction of $e^\star$; apply the model for fusion;
6. evaluate the reliability of the fusion result.

5.3.1. STEP 1: COMPUTING TOMOGRAMS

Our fusion framework starts with tomographic reconstruction. To date, a range of software packages have been issued for electron tomography, such as the open source ASTRA toolbox [33]. Given a HAADF tilt-series and a number of N available reconstruction algorithms, we can compute N volumetric images $\mathbf{x}_n^{\mathrm{H}} \in \mathbb{R}^{V_{\mathrm{H}} \times 1}$, $n = 1, \cdots, N$ where V_{H} is the total number of voxels being reconstructed. Similarly, for each element $e \in \mathscr{E}$, we can also reconstruct $\mathbf{x}_n^{(e)} \in \mathbb{R}^{V_{\mathrm{E}} \times 1}$ with $n = 1, \cdots, N$. Note that $V_{\mathrm{E}} \neq V_{\mathrm{H}}$ if the original X-ray spectral images have been binned.

5.3.2. STEP 2: CHECKING VISIBILITY AND DENOISING

In principle, one can choose any element $e \in \mathscr{E}$ as the fusion target $e^\star$; in practice, however, $e^\star$ should be visible in the HAADF reconstructions for building a representative cross-modality model. Our approach to measuring the visibility of $\mathbf{x}_n^{(e)}$ in $\mathbf{x}_n^{\mathrm{H}}$ is taken from [27]. First, we upsample $\mathbf{x}_n^{(e)}$ to $\tilde{\mathbf{x}}_n^{(e)} \in \mathbb{R}^{V_{\mathrm{H}} \times 1}$ if $V_{\mathrm{E}} \neq V_{\mathrm{H}}$ using trilinear interpolation. Then, we calculate visibility weights w_e, $\forall e \in \mathscr{E}$ as

$$\mathbf{x}_n^{\mathrm{H}} = w_0 + w_1 \tilde{\mathbf{x}}_n^{(1)} + \cdots + w_E \tilde{\mathbf{x}}_n^{(E)} \tag{5.1}$$

with ordinary least squares regression. Note that this only needs to be done once for any reconstruction algorithm. Since the intensity of HAADF-STEM scales with the atomic number of the element, we suggest that the weight of the chosen $e^\star$ should be at least three times as high as the light ones (e.g., N, O, etc.)

Once we select $e^\star$, we have N volumes $\mathbf{x}_n^{(e^\star)}$ at hand. Although they are computed by N different reconstruction algorithms, they share common patterns. To capture the most dominant structure among these reconstructions and reduce their pixel-specific variations, one can choose any dimension reduction technique mentioned in Section 5.2.2. Since all intensity values in the output image are inherently nonnegative, we perform NMF over all the reconstructions of $e^\star$ to obtain one denoised image $\mathbf{x}_{\mathrm{d}}^{(e^\star)}$.

5.3.3. STEP 3: GENERATING FEATURE IMAGES

The third step aims to generate more information for building the model. In particular, feature images of the N HAADF and $N(E-1)$ EDS reconstructions (with $e^\star$ being excluded) are generated with several texture filters. In our case, we use the entropy, range, standard deviation and Gaussian gradient magnitude filters; Table 5.1 lists the filter parameters. More filters could be included and provide the model with more feature images, namely higher degrees of freedom, to describe the denoised EDS reconstruction; however, they may simultaneously introduce unwanted patterns. For example, an ori-

entation filter may reproduce undesirable star-shaped streaks in the fusion result. One can also extend the filtering operation to various scale spaces, and we choose a two-level Gaussian scale space. Note that more filters and deeper scale spaces would lead to higher computational costs both in time and memory.

Table 5.1: Filters and related parameters in step 3

Name of filter	Parameter
Local entropy	$3 \times 3 \times 3$ neighborhood
Local range	$3 \times 3 \times 3$ neighborhood
Local standard deviation	$3 \times 3 \times 3$ neighborhood
Gaussian gradient magnitude	Standard deviation $\sigma = 1$

5.3.4. STEP 4: RESAMPLING

EDS reconstruction of $e^\star$ denoised in step (2) and feature images generated in step (3) should be of the same size for establishing the final model. Similarly, if the spectral images have been binned, we use the trilinear interpolation to upsample all EDS-related images at this stage.

5.3.5. STEP 5: BUILDING CROSS-MODALITY MODEL

We formulate the cross-modality modeling as a linear regression task

$$\mathbf{y} = \mathbf{X}\mathbf{b} + b_0 + \delta \tag{5.2}$$

where $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_P]$ and $\mathbf{b} = (b_1, \cdots, b_P)^T$ with $P = M \cdot N \cdot E$. Each predictor variable $\mathbf{x}_p$ denotes one (vectorized) feature image, and the response $\mathbf{y}$ is the denoised EDS reconstruction of $e^\star$. b_p, $p = 0, \cdots, P$ are the scalar regression coefficients to be found, and δ is the mismatch term. We employ partial least squares (PLS) regression, which is different from the ordinary least squares that directly establishes a linear model in the original data space. By first performing PCA to project both the predictor and response variables to $N_{\texttt{comp}}$ components in another space [29], PLS can produce stable results with low variability even if the correlation among predictor variables is high. Since the variance explained in the response variable $\mathbf{y}$ increases with the number of PLS components $N_{\texttt{comp}}$, we set $N_{\texttt{comp}}$ to its maximum value $N_{\texttt{comp}} = P - 1$ and solve this regression problem by `plsregress()` in MATLAB. Once we find all coefficients b_p, we fuse the image for $e^\star$ as

$$\mathbf{x}_{\texttt{f}}^{(e^\star)} = b_0 + b_1\mathbf{x}_1 + \cdots + b_P\mathbf{x}_P, \tag{5.3}$$

which is guaranteed to be the closest to the denoised EDS reconstruction $\mathbf{x}_{\texttt{d}}^{(e^\star)}$.

5.3.6. STEP 6: EVALUATING RELIABILITY

Lastly, we evaluate the reliability of the fusion result, considering that PLS will always build a model regardless of whether there is a linear relationship, and that the PLS output

is actually a prediction only. Specifically, we quantify the correspondence between the fused image $\mathbf{x}_{\mathtt{f}}^{(e^\star)}$ and the denoised $\mathbf{x}_{\mathtt{d}}^{(e^\star)}$ using the proportion of variance explained and the (Pearson) correlation coefficient. The former is returned by `plsregress()`, which is the ratio between the sum of squares of $\mathbf{y}$ explained by the $N_{\mathtt{comp}}$ PLS components and the total sum of squares of $\mathbf{y}$ [34]. The latter is calculated as

$$\mathrm{CC} = \frac{\sum_i (x_{\mathtt{f},i} - \overline{x_{\mathtt{f}}})(x_{\mathtt{d},i} - \overline{x_{\mathtt{d}}})}{\sqrt{\sum_i (x_{\mathtt{f},i} - \overline{x_{\mathtt{f}}})^2 \sum_i (x_{\mathtt{d},i} - \overline{x_{\mathtt{d}}})^2}}, \tag{5.4}$$

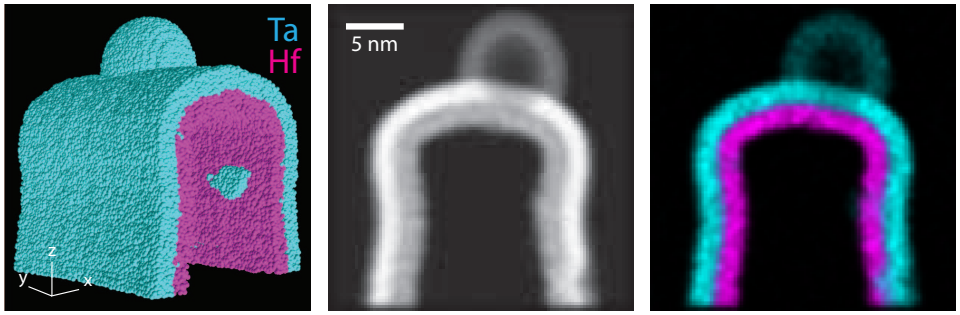
which measures how well the relative intensity distribution of $\mathbf{x}_{\mathtt{f}}$ matches $\mathbf{x}_{\mathtt{d}}$ [18]. $x_{\mathtt{f},i}$ and $x_{\mathtt{d},i}$ are the intensity values of the i-th voxel, $\overline{x_{\mathtt{f}}}$ and $\overline{x_{\mathtt{d}}}$ the average intensities over all voxels in $\mathbf{x}_{\mathtt{f}}$ and $\mathbf{x}_{\mathtt{d}}$, respectively.

5.4. EXPERIMENTS

Based on simulated and experimental datasets, we investigated the performance of our method and compared it to other state-of-the-art reconstruction techniques.

5.4.1. SIMULATION AND EXPERIMENTAL DATASETS

We start with two noise-free multislice simulation datasets that were generated from two semiconductor models. The first is a defective finFET structure with a size of 25 nm ×

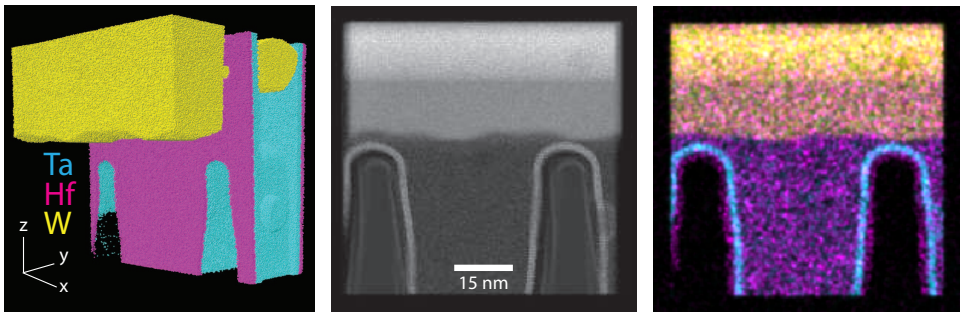


(a) Atomic design (b) HAADF-STEM (c) EDS maps

Figure 5.2: (a) Atomic design of a defective finFET structure. Tantalum (Ta, cyan) layer and hafnium (Hf, magenta) in HfO_2 layer are shown for clarity. (b) HAADF-STEM projection and (c) superposed EDS maps at 2°. Since the intensity of HAADF-STEM scales with the atomic number Z, only Ta ($Z_{\mathtt{Ta}} = 73$) and Hf ($Z_{\mathtt{Hf}} = 72$) are visible but O ($Z_{\mathtt{O}} = 8$) is not. Details in Section 5.4.1.

25 nm × 25 nm [35]. It has crystalline silicon (Si) as its source-drain fin, on top of which are oxygen (O), hafnium dioxide (HfO_2), tantalum (Ta), and titanium aluminum nitride ($TiAlN_2$). A few defects have been introduced, such as three pinholes in the HfO_2 layer with diameters of 1, 2 and 3 nm allowing Ta to contact the fin. Moreover, a 7 nm ellipsoid carbon contaminant is trapped between the HfO_2 and Ta layers. The atomic design of Ta and Hf is shown in Fig. 5.2(a). HAADF-STEM images were simulated with an accelerating voltage of 200 kV, a focused electron probe normalized to a total intensity of 1, a convergence angle of 10 mrad, and a detector with an inner angle of 90 mrad and outer angle 230 mrad. For details of the simulation see [35]. X-ray maps were generated by summing up the probability of characteristic emission. The raw tilt-series of this finFET

structure consists of 180 projections, ranging from 0° to 358° with an increment of 2° between consecutive projections. We only used the first 90 (unique) projections for reconstruction; each contains one HAADF-STEM image with a size of 128 pixel × 128 pixel (≈ 2Å/pixel) and eight elemental EDS maps of the same size and resolution. The second model is a larger region of PMOS with a size of 70 nm × 70 nm × 70 nm. Besides two Si fins, Ta metal gate and HfO_2 layer, a tungsten (W) contact was also added, see Fig. 5.3(a). Projections of this PMOS device were simulated over [0°, 180°) in every 2.5° with the same parameters set for the previous case. At each angle, one HAADF-STEM image with a size of 256 pixel × 256 pixel (≈ 3.8Å/pixel) and the equi-sized EDS maps of all chemical elements were recorded.

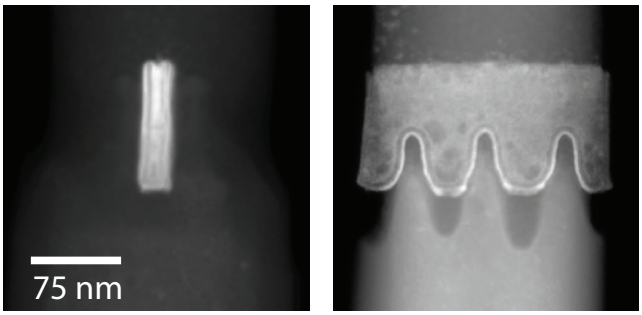


(a) Atomic design (b) HAADF-STEM (c) EDS maps

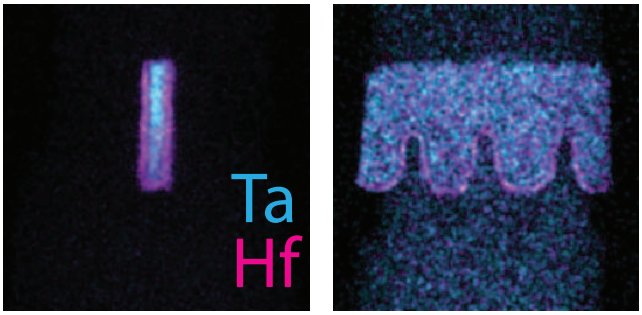
Figure 5.3: (a) Atomic design of a PMOS model where tantalum (Ta, cyan), hafnium (Hf, magenta) and tungsten (W, yellow) are shown. (b) HAADF-STEM projection and (c) superposed EDS maps at 2.5°. Ta ($Z_{\mathtt{Ta}} = 73$), Hf ($Z_{\mathtt{Hf}} = 72$) and W ($Z_{\mathtt{W}} = 74$) are clearly visible in the HAADF-STEM image. Details in Section 5.4.1.

To produce projections that are more comparable to real experimental data, we performed some post-processing steps on the simulated tilt-series. For HAADF-STEM images, we first applied Gaussian smoothing ($\sigma = 1.0$ pixel) to simulate a less focused lens system. Then, we added Poisson noise with a mean value of the HAADF intensity, and Gaussian noise with a standard deviation of 0.2 to corrupt the noiseless dataset. Finally, we excluded zone-axis projections (0° and 90°) because of the significant channeling effect present, as is custom in practice [36]. Compared to HAADF-STEM images, X-ray counts collected in the spectral images are usually fewer (max. 30 per pixel [5]), resulting in much noisier EDS maps. Therefore, we employed a Gaussian filter ($\sigma = 1.0$ pixel) for denoising. Moreover, since the number of EDS projections is always smaller than the STEM projections (due to time) in real experiments [32], we subsampled the two EDS tilt-series by factors of 3 and 2, respectively, assuming that they were recorded in every 6° and 5° over [0°, 180°). Post-processed projections of the defective finFET structure and PMOS model are shown in Fig. 5.2 and Fig. 5.3.

Our experimental dataset is a pillar-shaped semiconductor device comprised of eight chemical elements: N, O, Al, Si, Ti, Hf, Ta and Co [37]. It was placed on a Fischione on-axis rotation tomography holder allowing a 360° projection acquisition and scanned in a FEI Titan electron microscope equipped with four Super-X energy dispersive silicon drift detectors. The microscope was operated at an accelerating voltage of 120 kV with a beam current of 280 pA and a convergence angle of 10 mrad. HAADF-STEM projection images were acquired at 221 angles uniformly distributed between 0° and 220°. In addition, 47



(a) HAADF-STEM



(b) Superposed EDS maps of Ta and Hf

Figure 5.4: Experimental (a) HAADF-STEM projections and (b) EDS maps of a pillar-shaped semiconductor device at 0° (left) and 90° (right). Titanium with an atomic number of $Z_{\mathtt{Ti}} = 22$ is surrounded by tantalum (cyan, $Z_{\mathtt{Ta}} = 73$) and hafnium (magenta, $Z_{\mathtt{Hf}} = 72$). Details in Section 5.4.1.

full spectral images were recorded from 0° to 216° in approximately every 5°; each had a constant acquisition time of 270 s. They were later deconvolved into eight EDS maps corresponding to the eight aforementioned components in the sample. An example of the experimental tilt-series is given in Fig. 5.4.

5.4.2. FUSION FRAMEWORK

Besides the widely used reconstructions via FBP and SIRT, we also added SART and ML-EM to our fusion framework. The former combines the best of ray-by-ray ART and all-inclusive SIRT and can yield reconstructions of good quality along with high numerical accuracy in only a few iterations; the latter assumes the Poisson nature of electron/photon noise (rather than Gaussian) and inherently includes the nonnegativity constraint. For the EDS datasets, we chose Hanning window (rather than Ram-Lak filter) for FBP to deemphasize high frequencies, and 50 iterations for SIRT to avoid reconstructions overfitting to the noise. Moreover, we set 5 and 50 iterations for SART and ML-EM, respectively. For the less-noisy HAADF datasets, the number of iterations used for all iterative algorithms were doubled. Although we only incorporated four reconstruction techniques, we stress that our framework is general, and algorithms can be easily added (or removed) if necessary. Throughout this paper, visibility weights were calculated using the SIRT reconstructions of EDS and HAADF.

5.4.3. TILT AND X-RAY COUNT SETTINGS

We used the simulated datasets to study the fusion accuracy of our method for two cases: “limited-tilt" and “low-count". Compared to HAADF-STEM images, EDS spectra usually require much longer acquisition time at each tilt angle to achieve an acceptable SNR. Consequently, either case can effectively reduce the amount of time for nanomaterials being exposed to the electron beam and hence limit the sample damage. For the “limited-tilt" case, we reduced the number of EDS elemental maps in the two datasets from 30 and 36 to 8 and 9, respectively, guaranteeing that the remaining angles were uniformly distributed between $[0°, 180°)$. For the “low-count" case, we fixed the number of tilt angles for both EDS datasets to 30 and 36 and lowered their elemental X-ray counts by decreasing the maximal counts per pixel by three quarters. For the experimental dataset, we only considered the “limited-tilt" case by subsampling the original EDS tilt-series with 47 projections to 25 and 14, respectively.

5.4.4. BENCHMARK ALGORITHMS AND EVALUATION METRICS

We compared our proposed scheme to the classical FBP (with Hanning window) and SIRT (50 iterations), and two more advanced regularized approaches: TV and TNV, for which we set 200 and 400 iterations to guarantee convergence [16]. Moreover, since reconstruction qualities of TV and TNV highly depend on the regularization coefficient, we manually tuned their inputs for reasonable outputs, following the guidelines in [16].

We chose the structural similarity index (SSIM) and correlation coefficient to evaluate the reconstruction quality. SSIM is defined as

$$\mathrm{SSIM}(f,g) = l(f,g) \times c(f,g) \times s(f,g) \tag{5.5}$$

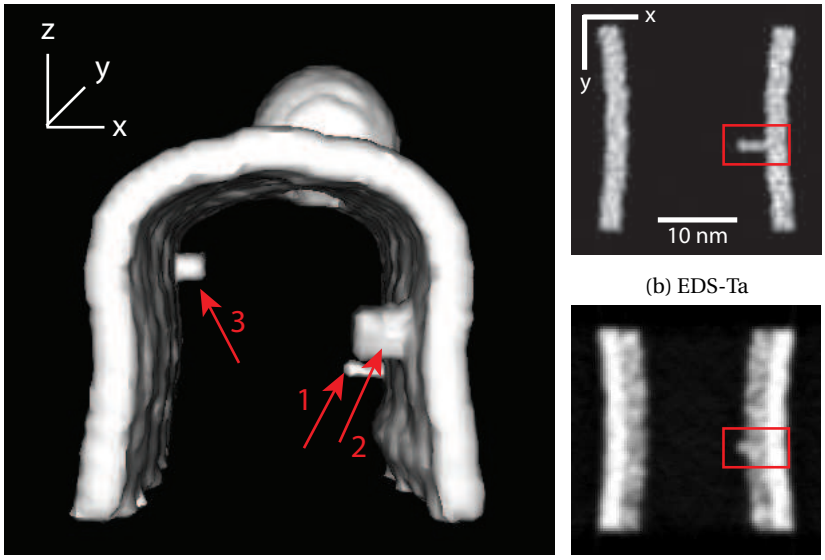
where

$$l(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}$$
$$c(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2}$$
$$s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3}$$

which measures the similarity between the reconstructed image f and ground truth g in three aspects: luminance (l), contrast (c) and structure (s) [38]. A higher SSIM value corresponds to a better reconstruction. μ_f and μ_g are the average intensity of f and g; σ_f and σ_g are the standard deviation; σ_{fg} is the covariance between f and g. Moreover, C_1, C_2 and C_3 are the constants introduced to avoid denominators being close to zero. By default, $C_1 = (0.01L)^2$, $C_2 = (0.03L)^2$ and $C_3 = C_2/2$ with L denoting the dynamic range of f and g [38]. Note that $s(f,g)$ reduces to the correlation coefficient in Eq. (5.4) if $C_3 = 0$. For the simulation datasets, the ground truth g was computed by SIRT with 100 iterations given the full-view noiseless (element-wise) EDS maps. We did not use the mask generated from the atoms' coordinates as reference image because it does not involve the interaction between the atoms and the incident electron beam.

5.5. RESULTS

5.5.1. SIMULATED FINFET DATASET



(a) 3D visualization of Ta

(b) EDS-Ta

(c) HAADF

Figure 5.5: (a) Volume rendering of the noiseless Ta (tantalum) reconstruction for the simulated finFET dataset; SIRT with 100 iterations using 90 elemental maps between $[0°, 180°)$. Three arrows point to the Ta penetration caused by the pinhole defects in the inner HfO_2 layer. (b) and (c) are the Ta and HAADF reconstruction xy-slices at $z = 37$, corresponding to the first arrow in (a). Red rectangles indicate the defect.

For the first simulated dataset, we select Ta as the target $e^\star$, which has the highest visibility weight $w_{\mathtt{Ta}} = 0.84$. Fig. 5.5(a) shows a volume rendering of its noiseless reconstruction, in which penetrations (indicated by red arrows) result from pinhole defects in the inner HfO_2 layer. We consider this 3D volume as the ground truth. Fig. 5.5(b) shows an orthoslice at location 1 and the red rectangle highlights an 1 nm defect. Since the atomic numbers of Ta and Hf are close ($Z_{\mathtt{Ta}} = 73$, $Z_{\mathtt{Hf}} = 72$), they yield similar Z-contrast in the HAADF reconstruction and make discrimination difficult, see Fig. 5.5(c).

Fig. 5.6 depicts the xy-slices of Ta reconstructions at $z = 37$, which are generated by FBP, SIRT, TV, TNV and our fusion algorithm. Images in the last column are the (NMF denoised) response variable $\mathbf{y}$ in Eq. (5.2). In the first row, the number of used projections and the maximal X-ray counts per pixel in the EDS datasets are 30 and 20, respectively. We refer to this as the “normal" case for brevity because it is comparable to a typical experiment. Moreover, the second and third rows show the “limited-tilt" and “low-count" cases, in which either the number of projections or X-ray counts is reduced by three quarters.

Although FBP successfully reproduces the penetration defect in the normal case (Fig. 5.6(a), top row), it also introduces visible line artifacts in the background. This can be largely suppressed by SIRT (Fig. 5.6(b)). Tomograms from regularized TV and TNV are visually indistinguishable regarding the continuity of edges and smoothness of foreground. Our algorithm can not only get rid of background noise as TV and TNV, but also keep the structural details of the Ta layer (e.g., rectangular ends) that are somewhat smeared by the other two techniques. When the number of projections is reduced from

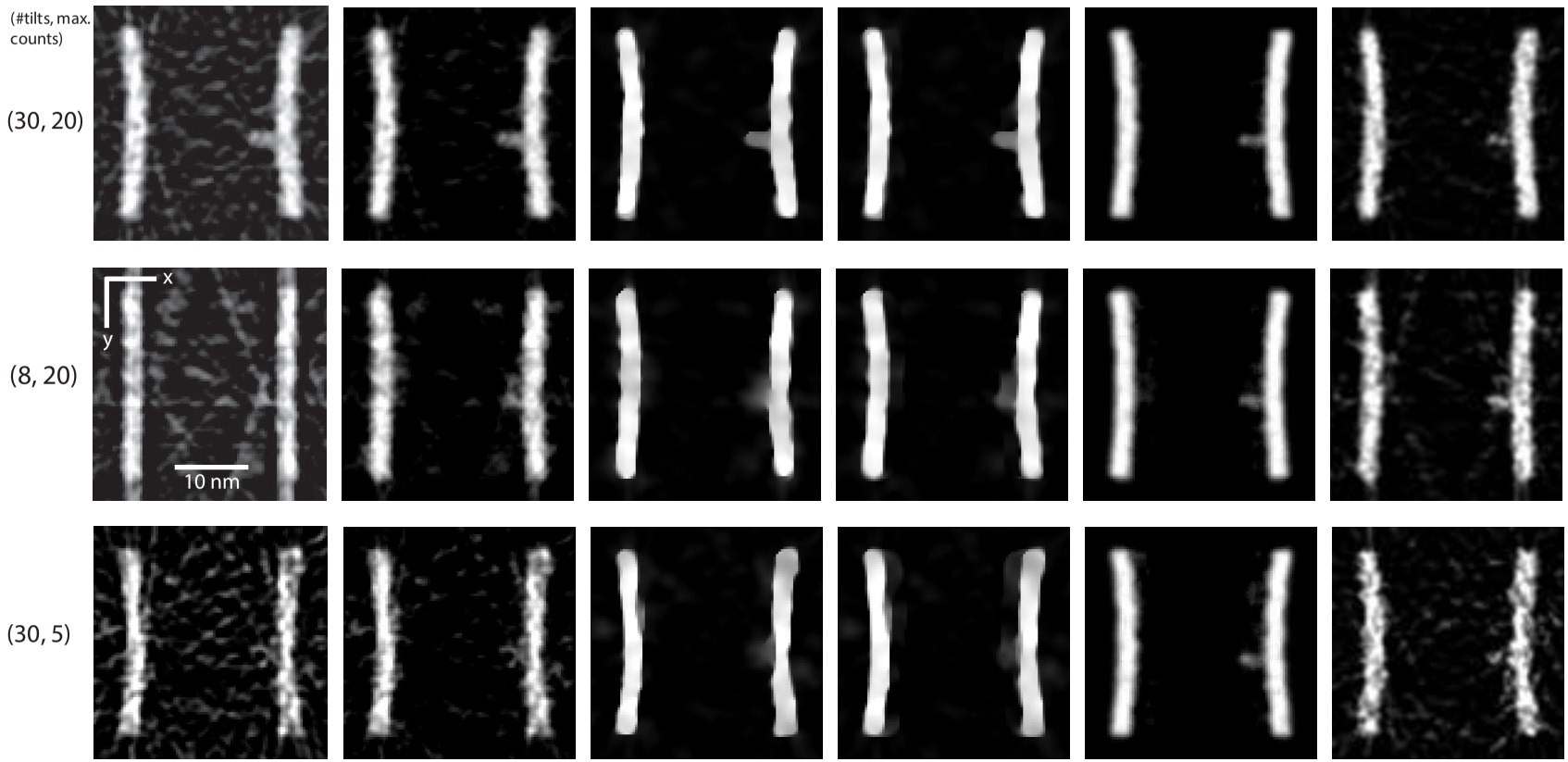


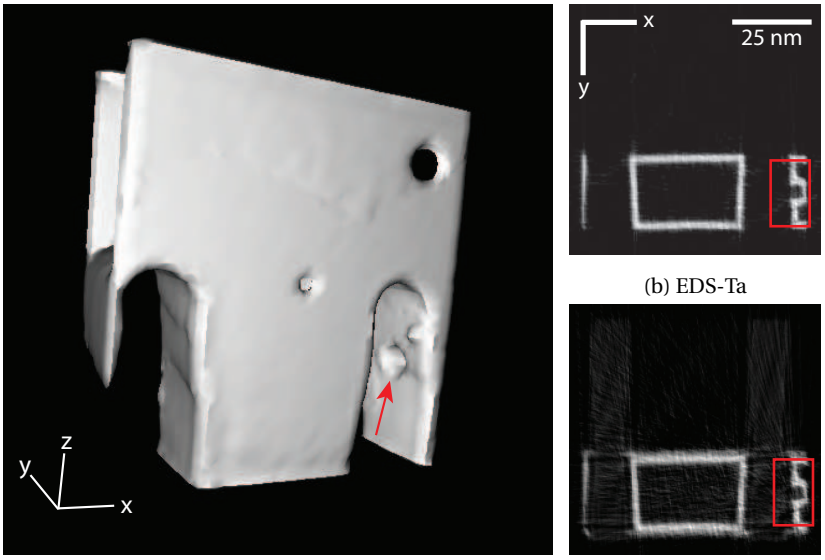
Figure 5.6: Ta xy-slices of (a) FBP, (b) SIRT, (c) TV, (d) TNV, (e) fusion and (f) NMF denoising for the simulated finFET at $z = 37$ (the height of the first defect in Fig. 5.5(a)). The top row is the “normal" case where the number of EDS projections and the maximal X-ray counts per pixel are 30 and 20. Middle and bottom rows are the limited-tilt case (with 8 projections) and low-count case (max. 5 X-ray counts). For better visualization, we perform percentile contrast stretching from 0 to 99%.

30 to 8 (Fig. 5.6, middle row), FBP, SIRT and NMF deform the curved Ta layer because of the limited tilts. Regularized TV and TNV manage to retain its shape but simultaneously introduce jaggy edges and paintbrush artifacts as both encourage piecewise constant structures [16]. However, TNV better preserves the rectangular ends of the Ta layer thanks to the augmentation from HAADF-STEM projections. The image computed by our algorithm is visually pleasing due to its continuous edges and homogeneous foreground. Additionally, it maintains the curved Ta layer to some extent even though it highly relies on the NMF denoised reconstruction. The tiny defect can be seen in TV, TNV, NMF and our fusion results, whereas it is distorted from a tip to a bulb in the former two reconstructions and contaminated by streak artifacts in the third. For the low-count case (Fig. 5.6, bottom row), classical FBP and SIRT, and the resulting NMF create a lot of undesirable lines in their fore- and background due to the low signal. To significantly restrain such artifacts, we set high regularization coefficients λ for TV ($\lambda_{\mathrm{TV}} = 0.045$) and TNV ($\lambda_{\mathrm{TNV}} = 0.1$), but inevitably distort the shape of the Ta layer and fail to satisfactorily reconstruct the defect. Although the penetration in our fused image is mixed with noise along the edges, one can still see it thanks to its relatively high intensity (at least three times higher than the noise). From top to bottom conditions in Fig. 5.6, the corresponding proportion of variance explained and correlation of the fusion to the denoised image are (0.95, 0.98), (0.91, 0.96) and (0.83, 0.91), respectively.

Table 5.2 summarizes the SSIM and correlation coefficient (CC) values of the five aforementioned algorithms on the three settings in Fig. 5.6. Different noise realizations do not change the values in Table 5.2 to the shown digits. Our scheme ranks the best in all three cases. FBP falls far behind in SSIM due to the line artifacts visible in Fig. 5.6. The two regularized techniques achieve similar quantitative performance, but are surpassed by SIRT in terms of CC because regularizations may oversmooth the underlying structure and make reconstructions less accurate. Note that our fusion algorithm can produce stable results regardless of limited or noisy datasets, as demonstrated in Fig. 5.6 and Table 5.2.

Table 5.2: Comparison of SSIM and correlation coefficient (CC) for Ta reconstruction of simulated finFET dataset

(#tilts, counts)	Metrics	FBP	SIRT	TV	TNV	Fusion
(30, 20)	SSIM	0.33	0.81	0.87	0.88	**0.96**
	CC	0.95	0.97	0.90	0.90	**0.98**
(8, 20)	SSIM	0.29	0.79	0.85	0.86	**0.95**
	CC	0.80	0.95	0.88	0.89	**0.98**
(30, 5)	SSIM	0.22	0.70	0.80	0.80	**0.95**
	CC	0.86	0.93	0.90	0.91	**0.98**



(a) 3D visualization of Ta

(b) EDS-Ta

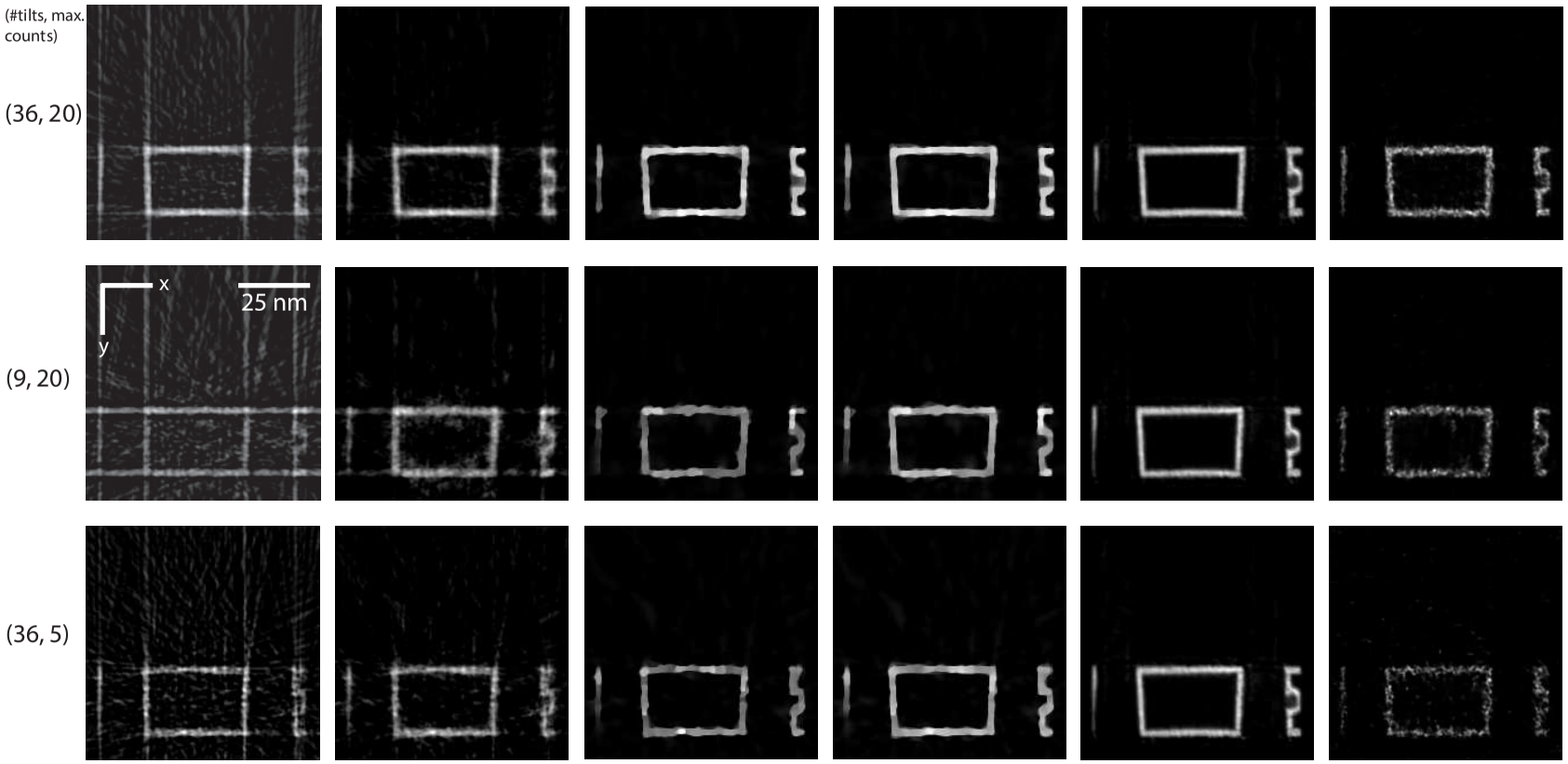
(c) HAADF

Figure 5.7: (a) Volume rendering of the noiseless Ta reconstruction for the simulated PMOS dataset; SIRT with 100 iterations using 72 elemental maps between $[0^\circ, 180^\circ)$. (b) and (c) are the Ta and HAADF reconstruction xy-slices at $z = 94$, in which red rectangles encompass a fine structure indicated by the arrow in (a).

5.5.2. SIMULATED PMOS DATASET

The noiseless Ta reconstruction of the simulated PMOS dataset is rendered in 3D in Fig. 5.7(a), and its xy-slice at $z = 94$ depicted in Fig. 5.7(b). This uniformly distributed element ranks second in the visibility check, surpassed by W ($w_{\mathtt{Ta}} = 0.42$ and $w_{\mathtt{W}} = 0.85$). Fig. 5.7(c) is the associated HAADF reconstruction where thin Ta/Hf layers, and two Si fins are visible. Red rectangles encompass a fine structure that is pointed out by the red arrow in Fig. 5.7(a).

Fig. 5.8 shows the corresponding xy-slices reconstructed by all techniques. As before, the top row is the normal case where the number of projections and the maximal X-ray counts per pixel in the EDS tilt-series are 36 and 20, respectively; the middle row is the limited-tilt case using only 9 projections, and the bottom the low-count case with no more than 5 X-ray counts per pixel. As illustrated in Fig. 5.8(a), background of FBP reconstructions are severely contaminated by lines if the sample of interest contains laminate. SIRT can suppress such artifacts but fails to deliver clear edges. Despite that TV and TNV achieve cleaner background than FBP and SIRT, they also noticeably smear the Ta layer especially for the limited-tilt and low-count cases: rectangular corners are rounded; laminate becomes thicker; line segments on the left are much shorter than they are supposed to be, see Fig. 5.7(b). In general, such smearing effects are more pronounced in TV than TNV because the latter can incorporate the edge information from HAADF-STEM. Our fusion algorithm can generate a volume that is qualitatively comparable to the ground truth for the normal case, see Fig. 5.7(b) and Fig. 5.8(e). In the other two cases, fine structures are still kept intact even though they are corrupted by noise in the regression targets; however, line artifacts in the background (with intensities no more than 15% of the foreground) show up. Quantitative results in Table 5.3 confirm this, in which our algorithm is in the top rank except for SSIM in the limited-tilt case. Outputs



(a) FBP (b) SIRT (c) TV (d) TNV (e) Fusion (f) NMF

Figure 5.8: Ta xy-slices of (a) FBP, (b) SIRT, (c) TV, (d) TNV, (e) fusion and (f) NMF denoising for the simulated PMOS at $z = 94$ (the height of the arrow in Fig. 5.7(a)). The top row is the “normal" case where the number of EDS projections and the maximal X-ray counts per pixel are 36 and 20. Middle and bottom rows are the limited-tilt case (with 9 projections) and low-count case (max. 5 X-ray counts), respectively.

of the reliability validation for the aforementioned three cases are (variance explained, correlation): (0.86, 0.94), (0.74, 0.88) and (0.61, 0.8), respectively. The low variance explained for the low-count case is due to the noisy NMF, but our fusion remains relatively smooth and clear.

Table 5.3: Comparison of SSIM and correlation coefficient (CC) for Ta reconstruction of simulated PMOS dataset

(#tilts, counts)	Metrics	FBP	SIRT	TV	TNV	Fusion
(36, 20)	SSIM	0.48	0.86	**0.91**	**0.91**	0.90
	CC	0.92	0.97	0.94	0.94	**0.98**
(9, 20)	SSIM	0.37	0.80	**0.90**	**0.90**	0.83
	CC	0.68	0.92	0.90	0.92	**0.93**
(36, 5)	SSIM	0.25	0.76	0.86	0.85	**0.88**
	CC	0.80	0.92	0.90	0.91	**0.96**

5.5.3. EXPERIMENTAL DATASET

Fig. 5.9 depicts an orthoslice of the HAADF reconstruction for the experimental dataset, which clearly reveals the structural information of this pillar-shaped semiconductor device. For example, the dark contrast highlighted by the second red arrow possibly results from a defect (void inside the laminate). However, since the innermost Ta layer and the outermost Hf yield similar Z-contrast, directly discerning them from the HAADF reconstruction without any chemical information would be challenging.

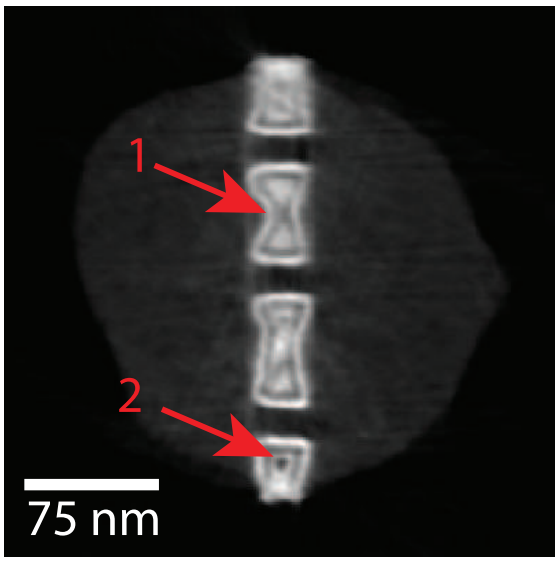
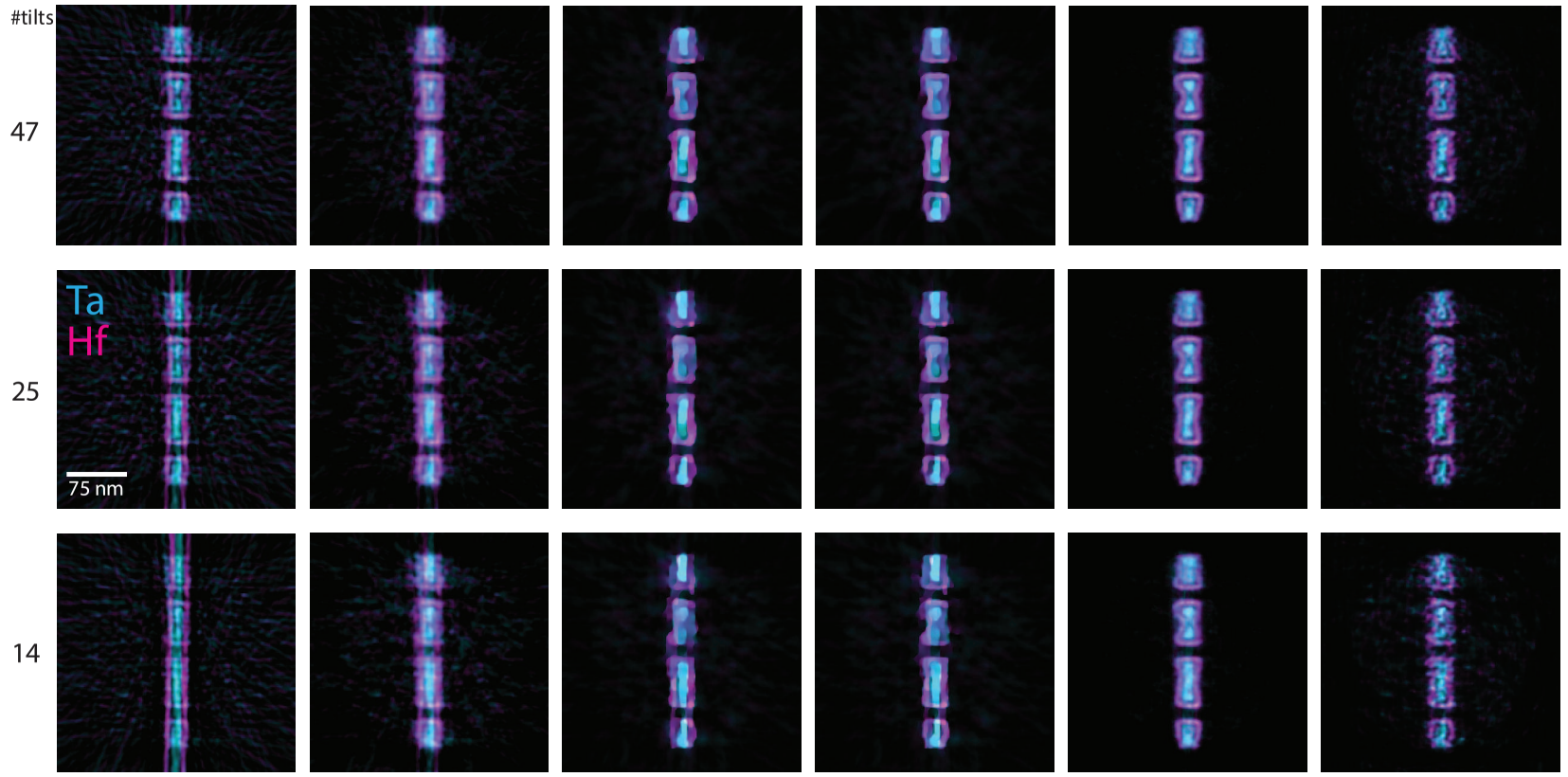


Figure 5.9: An orthoslice of HAADF reconstruction for the experimental dataset. The first red arrow indicates a squeezing feature, and the second a dark contrast possibly results from some void inside the laminate.



(a) FBP (b) SIRT (c) TV (d) TNV (e) Fusion (f) NMF

Figure 5.10: Superposed xy-slices of Ta (cyan) and Hf (magenta) reconstructions computed by (a) FBP, (b) SIRT, (c) TV, (d) TNV, (e) fusion and (f) NMF denoising for the experimental dataset using 47 (top), 25 (middle) and 14 (bottom) EDS projections.

Fig. 5.10 illustrates the Ta (cyan) and Hf (magenta) reconstructions generated by FBP, SIRT, TV, TNV and our fusion algorithm using 47, 25 and 14 EDS projections; the last column shows the response variable $\mathbf{y}$ in Eq. (5.2). Visibility weights of Ta and Hf are 0.2 and 0.25, respectively, ranking behind Si with $w_{\mathtt{Si}} = 0.41$. From top to bottom, the variance explained and the correlation to $\mathbf{y}$ are (0.88, 0.92), (0.84, 0.91) and (0.82, 0.90) for Ta, and (0.89, 0.93), (0.84, 0.91) and (0.82, 0.90) for Hf. When the number of projections is decreasing, thin streaks in FBP drastically degrade its reconstruction quality, because the structure under study aligns with the projection direction, see Fig. 5.10(a); moreover, the background of SIRT reconstructions is also obviously getting noisier. Note that such noise is much lower in the resulting NMF. TV and TNV are also able to suppress the background noise; however, they simultaneously smear the Ta and Hf layers and make their boundary almost indistinguishable, especially for the 14-tilt case. Despite the spotty foreground in NMF, our fusion algorithm still achieves the best performance in noise suppression and structure preservation (e.g., the squeezing pattern indicated by the first arrow in Fig. 5.9) regardless of the number of projections. Moreover, it also maintains the void observed in the HAADF reconstruction to some extent, which is almost unrecognizable in the other five reconstructions. Therefore, our method can enable easier and more accurate fault analysis in the subsequent process. As no ground truth is available for this experimental dataset, we cannot compute the SSIM and correlation coefficient.

5.6. DISCUSSION

In the previous section, we demonstrated that our fusion algorithm is more robust for the limited and noisy datasets than other state-of-the-art tomography techniques. For the simulated datasets, we reduced either the number of EDS elemental maps or the maximal X-ray counts per pixel to one quarter of the initial values and showed that our method could deliver stable outputs which were visually consistent with the ground truth of the Ta layer. We also fused Hf (for both finFET and PMOS) and W (only for PMOS) and found that the homogeneously distributed W resulted in visually better fusion. Because Hf was mixed with the light element O ($Z_{\mathtt{O}} = 8$), its signals in the HAADF-STEM projections were weaker ($w_{\mathtt{Hf}} = 0.42$ for finFET and $w_{\mathtt{Hf}} = 0.18$ for PMOS). Consequently, fused Hf images were less smooth in the foreground and contaminated by the heavier Ta (and W) in the background; such contamination could also be inferred from the low correlation to the denoised Hf (e.g., only 0.68 in the low-count case for PMOS) because this metric focuses on the relative pattern similarity [18]. For the experimental dataset, we subsampled the EDS tilt-series to simulate the “limited-tilt" scenario. Our method could reproduce fine structures and maintain clear boundaries between the Hf and Ta layers even if only 14 tilts were available.

Since both TV and TNV noticeably distort fine details in their reconstructions, we further tested HOTV [12] (online available [39]) on the simulated finFET dataset for reconstructing Ta. We found that the performance of HOTV (second- or third-order with regularization parameters as Eq. (15) in [12]) lies between SIRT and TV. That is, it could better recover the penetration defect in Fig. 5.5(b) but could not adequately suppress noise especially in the foreground. Consequently, desired fine features were mixed with (undesired) noise and hence did not stand out.

All datasets we used have no missing wedge (sample rotation was not limited within a

certain range in the electron microscope due to mechanical constraints, e.g., $\pm 70^\circ$ [32]); these pillar-shaped samples and rotation holders will eventually replace the traditional tomography holders. However, we tested the robustness of the precedent 2D version [19] on the missing wedge artifact using an experimental dataset of a core-shell nanoparticle consisting of gold (Au) and silver (Ag), which only contained 31 tilts ranging from -75° to $+75^\circ$ with an increment of 5° between the consecutive projections. Qualitatively, our method outperformed HAADF-EDS bimodal tomography [15] in terms of the sharpness of edges and smoothness of fore- and background. Quantitatively, it achieved higher correlation coefficient between the reconstruction and the hand-segmented ground truth both for Au and Ag.

One prerequisite for our cross-modality framework to work is that chemical element chosen for fusion should be visible in the HAADF-STEM images; this is quantified by the visibility weight w_e in Eq. (2). Since light elements such as N and O never show up in the HAADF, we set the threshold for the target $e^\star$ three times as high as the weights of the light ones (e.g., O, $3 \times w_0 = 0.15$ for finFET). This limitation may be overcome by incorporating other imaging modalities, such as bright-field and dark-field STEM with different detector geometries.

In addition, we also validated whether the resulting model is representative (i.e., whether fused images can be relied upon) through the proportion of variance explained and the correlation to the denoised image. For instance, if we try to fuse the carbon contaminant in Fig. 5.2 whose visibility weight is almost zero (even though its structure is clearly sketched by the surrounding HfO_2 and Ta layers, see Fig. 5.2(b)), the corresponding variance explained is only 0.17 and correlation coefficient 0.5. As a rule of thumb, these two metrics should be above 0.6 and 0.8, respectively, to achieve a reliable fusion.

We further investigated how the choice of three popular 3D upsampling techniques: nearest-neighbor, trilinear and tricubic, influences the fusion quality. Since the resolution of EDS reconstructions decreases rapidly with an increased binning ratio r, we only considered $r = 2$ with $V_E = 1/8 V_H$. The nearest-neighbor interpolation produced the least favorable result; therefore, we left it out even though it has the highest speed. Moreover, we found that the trilinear and tricubic interpolation are comparable in terms of their visually indistinguishable fusion results, and similar SSIM and CC values. We chose the trilinear because it is slightly faster than tricubic interpolation.

Computational time of our fusion algorithm is mainly spent on computing tomograms (step 1), generating feature images (step 2) and building cross-modality model (step 5). In our case, for instance, it took 300 s for tomographic reconstruction, 25 s for feature image generation, and another 300 s for cross-modal modeling for Ta in the simulated finFET structure using MATLAB 2017a on a desktop equipped with eight Intel Xeon X5550 CPU cores (24 GB memory) and NVIDIA GeForce GTX670 GPU (4 GB memory). Note that only the last 300 s (for building the model) are required to fuse any other element in this sample. Throughout this paper, FBP, SIRT, SART and ML-EM reconstructions were computed with the ASTRA toolbox [33]; regularized TV and TNV algorithms were realized by a Douglas-Rachford primal-dual splitting algorithm provided in the operator discretization library [40]. Although our method is slower than FBP (3 s), SIRT (12 s) and TV (273 s), it is still much faster than TNV (~80 mins); manually tuning regularization coefficients for TV and TNV to suppress noise yet avoiding over-regularization also

consumes a lot of time. Moreover, considering that the acquisition time for each spectral image varies from 236 s to 895 s [32], our algorithm is still quite appealing.

5.7. CONCLUSION

In this paper, we have presented a regression-based cross-modal fusion framework for electron tomography, which does not require any fine-tuning parameter. We have adopted it to combine the EDS and HAADF reconstructions and investigated its performance using simulated and experimental datasets of semiconductor devices that contain chemical elements with close atomic numbers. Results have shown that our fusion algorithm can consistently yield more accurate reconstructions than the conventional yet widely employed FBP and SIRT, and the advanced regularization-based TV and TNV. Furthermore, it can still restore fine structures and achieve a high reconstruction quality even for limited and noisy EDS datasets. Such properties are highly desired in the semiconductor industry where the number of EDS maps is limited aiming for a shorter acquisition time, and the incident electron dose is usually low to minimize the sample damage.

REFERENCES

[1] Y. Guo, R. Aveyard, and B. Rieger, "A multichannel cross-modal fusion framework for electron tomography," *IEEE Transactions on Image Processing*, vol. 28, pp. 4206–4218, 2019.

[2] M. Weyland and P. A. Midgley, "Electron tomography," *Materials Today*, vol. 7, no. 12, pp. 32–40, 2004.

[3] A. J. Koster, U. Ziese, A. J. Verkleij, A. H. Janssen, and K. P. de Jong, "Three-dimensional transmission electron microscopy: a novel imaging and characterization technique with nanometer scale resolution for materials science," *Journal of Physical Chemistry B*, vol. 104, no. 40, pp. 9368–9370, 2000.

[4] P. A. Midgley and M. Weyland, "3D electron microscopy in the physical sciences: the development of Z-contrast and EFTEM tomography," *Ultramicroscopy*, vol. 96, pp. 413–431, 2003.

[5] A. Genc, 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*, vol. 131, pp. 24–32, 2013.

[6] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[7] B. Goris, W. van den Broek, K. J. Batenburg, H. H. Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012.

[8] P. Ercius, O. Alaidi, M. J. Rames, and G. Ren, "Electron tomography: a three-dimensional analytic tool for hard and soft materials research," *Advanced Materials*, vol. 27, pp. 5638–5663, 2015.

[9] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica*, vol. 28, pp. 94–108, 2012.

[10] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2542–2553, 2011.

[11] 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 Trans. Image Process.*, vol. 25, no. 1, pp. 455–468, 2016.

[12] T. Sanders, A. Gelb, R. B. Platte, I. Arslan, and K. Landskron, "Recovering fine details from under-resolved electron tomography data using higher order total variation l_1 regularization," *Ultramicroscopy*, vol. 174, pp. 97–105, 2017.

[13] T. Sanders and R. B. Platte, "Multiscale higher order TV operators for l_1 regularization," *Advanced Structural and Chemical Imaging*, vol. 4, pp. 12–29, 2018.

[14] 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*, vol. 33, 2016.

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

[16] 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*, vol. 191, pp. 34–43, 2018.

[17] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: a survey," *Information Fusion*, vol. 45, pp. 153–178, 2019.

[18] R. van de Plas, J. Yang, J. Spraggins, and R. M. Caprioli, "Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping," *Nature Methods*, vol. 12, pp. 366–374, 2015.

[19] Y. Guo and B. Rieger, "Image fusion of X-ray and electron tomograms," in *Proceedings of the 2018 IEEE International Conference on Image Processing*, 2018, pp. 863–867.

[20] D. S. Rigie and P. J. L. Rivire, "Joint reconstruction of multi-channel spectral CT data via constrained total nuclear variation minimization," *Physics in Medicine & Biology*, vol. 60, pp. 1741–1762, 2015.

[21] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, 2017.

[22] J. Sui, T. Adali, Q. Yu, J. Chen, and V. D. Calhoun, "A review of multivariate methods for multimodal fusion of brain imaging data," *Journal of Neuroscience Methods*, vol. 204, pp. 68–81, 2012.

[23] E. J. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM*, vol. 58, no. 3, 2011.

[24] M. Ghahremani and H. Ghassemian, "Remote-sensing image fusion based on curvelets and ICA," *International Journal of Remote Sensing*, vol. 36, pp. 4131–4143, 2015.

[25] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[26] J. Wang, J. Peng, X. Feng, G. He, and J. Fan, "Fusion method for infrared and visible images by using non-negative sparse representation," *Infrared Physics & Technology*, vol. 67, pp. 477–489, 2014.

[27] B. Aiazzi, S. Baronti, and M. Selva, "Improving component substitution pansharpening through multivariate regression of MS+Pan data," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3230–3239, 2007.

[28] G. Vivone, R. Restaino, and J. Chanussot, "Full scale regression-based injection coefficients for panchromatic sharpening," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3418–3431, 2018.

[29] R. D. Tobias, "An introduction to partial least squares regression," in *Proceedings of the Twentieth Annual SAS Users Group International Conference*, Cary, USA, 1995, pp. 1250–1257.

[30] E. Martínez-Montes, P. A. Valdés-Sosa, F. Miwakeichi, R. I. Goldman, and M. S. Cohen, "Concurrent EEG/fMRI analysis by multiway partial least squares," *NeuroImage*, vol. 22, pp. 1023–1034, 2004.

[31] A. A. Gowen and R. M. Dorrepaal, "Multivariate chemical image fusion of vibrational spectroscopic imaging modalities," *Molecules*, vol. 21, no. 7, pp. 870–888, 2016.

[32] T. J. A. Slater, A. Janssen, P. H. C. Camargo, M. G. Burke, N. J. Zaluzec, and S. J. Haigh, "STEM-EDX tomography of bimetallic nanoparticles: a methodological investigation," *Ultramicroscopy*, vol. 162, pp. 61–73, 2016.

[33] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA toolbox: a platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.

[34] S. de Jong, "SIMPLS: an alternative approach to partial least squares regression," *Chemometrics and Intelligent Laboratory Systems*, vol. 18, pp. 251–263, 1993.

[35] R. Aveyard and B. Rieger, "Tilt series STEM simulation of a 25x25x25 nm semiconductor with characteristic X-ray emission," *Ultramicroscopy*, vol. 171, pp. 96–103, 2016.

[36] M. C. Scott, C. C. Chen, M. Mecklenburg, C. Zhu, R. Xu, P. Ercius, U. Dahmen, B. C. Regan, and J. Miao, "Electron tomography at 2.4-ångström resolution," *Nature*, vol. 483, pp. 444–447, 2012.

[37] Y. Qiu, P. van Marcke, O. Richard, H. Bender, and V. Wilfried, "3D imaging of Si FinFETs by combined HAADF-STEM and EDS tomography," in *18th International Microscopy Congress*, Prague, Czech Republic, 2014, pp. 975–976.

[38] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[39] T. Sanders, "Open access codes for imaging problems," http://www.toby-sanders.com/software, 2016.

[40] J. Adler, H. Kohr, and O. Öktem, "Operator discretization library (ODL)," https://github.com/odlgroup/odl, 2017.

6

Improved Fusion for Cross-Elemental Contamination Suppression

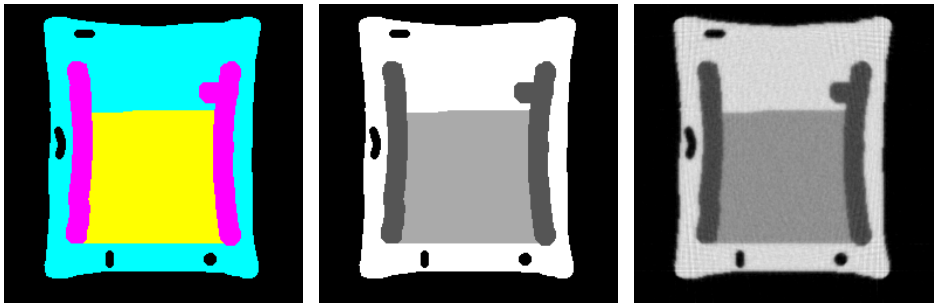
In Chapter 5, we presented a HAADF-EDS cross-modal fusion framework to simultaneously achieve high signal-to-noise ratio (SNR) and high elemental specificity. While the simulation and experimental results of semiconductor devices look promising, we noted that if the target element has a low weight fraction in the sample (i.e., low ratio and/or low atomic number), the fused image might be contaminated by other elements in the background [1, Sec. VI].

In this chapter, we add a weighted mask to the earlier framework so as to suppress the “cross-elemental contamination". We evaluate the modified algorithm on an experimental dataset of a gate-all-around nanowire device that has more complex structures and more element types [2, 3]. Fusion results are improved in three aspects: (*i*) reduced contamination artifact; (*ii*) enabled three-dimensional (3D) interpretation; (*iii*) light element fusion becomes possible if it has a distinguishable shape in the sample.

This chapter is structured as follows. In Section 6.1, we illustrate the cross-elemental contamination using a two-dimensional (2D) phantom. Section 6.2 and 6.3 present the improved fusion framework, and experiments and results, respectively. Finally, we discuss our work and draw conclusions in Section 6.4.

6.1. WHAT IS CROSS-ELEMENTAL CONTAMINATION?

We use a simplified continuous two-dimensional (2D) phantom [4, Sec. 3.1.1] consisting of Ag, Cu and Ti to elaborate on the cross-elemental contamination, see Fig. 6.1(a). The intensity of the element in the HAADF phantom is weighed by Z^n, in which Z is the atomic number ($Z_{\mathtt{Ag}} = 47$, $Z_{\mathtt{Cu}} = 29$, $Z_{\mathtt{Ti}} = 22$) and $n = 1.7$ [4]. Note that the composition and/or structure of the phantom may not be realistic.

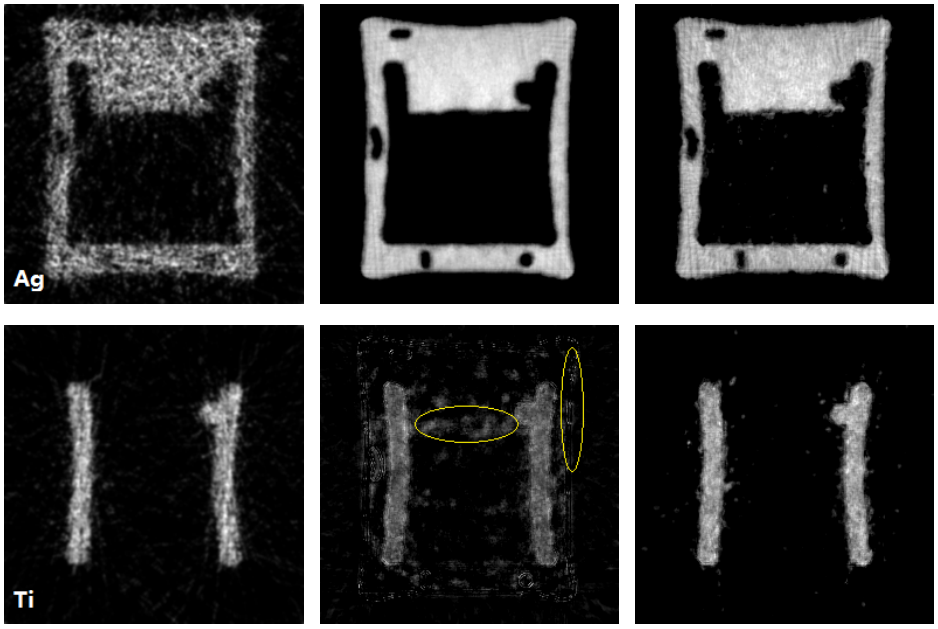


(a) Elemental distribution (b) HAADF phantom (c) HAADF reconstruction

Figure 6.1: Simplified continuous two-dimensional phantom, adapted from [4]. (a) Elemental distribution of Ag, Cu and Ti; the intensity of the element in (b) HAADF phantom is scaled by Z^n where Z is the atomic number ($Z_{\mathtt{Ag}} = 47$, $Z_{\mathtt{Cu}} = 29$, $Z_{\mathtt{Ti}} = 22$) and $n = 1.7$ [4]. (c) HAADF reconstruction. Details on page 80.

One-dimensional HAADF projections and EDS maps were generated with a simplified forward projection model. Assuming that the imaging system allows a 360° projection acquisition, we simulated one HAADF and three EDS tilt-series over $[0°, 180°)$ in every 2° and 6°, respectively, using the ASTRA toolbox [5]. Then, we scaled the intensity values of the HAADF and EDS to be comparable to real experimental data. Finally, the noiseless dataset was corrupted by Poisson noise with a mean value of the intensity and Gaussian noise with a standard deviation of 0.2 [1].

Fig. 6.2(a) and 6.2(b) depict the Ag and Ti reconstructions computed by SIRT and our



(a) SIRT (b) w/o mask (c) w/ mask

Figure 6.2: Ag (upper) and Ti (lower) reconstructions; computed with (a) SIRT, (b) fusion without and (c) with a weighted mask. Yellow circles indicate the contamination by Ag, namely cross-elemental contamination.

(earlier) fusion algorithm [1]; the corresponding HAADF reconstruction is in Fig. 6.1(c). Since Ag is heavier than Cu and Ti, it yields the highest Z-contrast in HAADF; thus, the fine structures of Ag (e.g., the four holes) are well preserved in fusion but almost invisible in SIRT. Compared with Ag, Ti has a much lower weight fraction in the phantom and produces much weaker signals in HAADF; consequently, the fused image is severely contaminated by Ag, see the yellow circles in Fig. 6.2. We name this artifact cross-elemental contamination.

6.2. MODIFIED SEVEN-STEP FUSION FRAMEWORK

Considering that regions without the target element should also not be contained in feature images, we add a weighted mask to the earlier fusion framework [1] in order to suppress the cross-elemental contamination. The modified fusion algorithm has seven steps; changed ones are in italics:

1. compute multiple tomograms from the HAADF and EDS tilt-series using different reconstruction algorithms;
2. check visibility *block by block* for all chemical elements and select a target element $e^\star$ for fusion;
3. denoise the EDS reconstructions of $e^\star$ and *generate its weighted mask*;
4. generate multiple feature images for each HAADF reconstruction, and each EDS reconstruction except for $e^\star$;
5. (optional) upsampling;
6. build a cross-modality model between the *masked* feature images and the denoised EDS reconstruction of $e^\star$ using partial least squares regression; apply the model for fusion;

7. *evaluate the reliability of the fusion result.*

STEP 3: GENERATING WEIGHTED MASK

The weighted mask is generated out of the EDS reconstruction of $e^\star$. Many options are available: the reconstruction of a single algorithm (e.g., SIRT), the denoised image, an over-regulated TV reconstruction, etc. We take the denoised image as an example. First, we employ a Gaussian filter ($\sigma = 1$ voxel) to smooth the denoised image even further. Then, we compute its binary mask using hysteresis thresholding [6]; this technique considers areas above a low threshold τ_L to be the foreground if they are also connected to areas above a higher and more stringent threshold τ_H. Finally, we smooth the binary mask with a Gaussian filter ($\sigma = 1$ voxel) to generate a gray-scaled one.

STEP 7: EVALUATING RELIABILITY

We now quantify the correspondence between the fusion result and the *masked* denoised image using the proportion of variance explained and the Pearson correlation coefficient [1]. We choose the masked image instead of the original reconstruction to reduce the interference from background noise.

Fig. 6.2(c) shows the fusion results with mask. On the one hand, this newly added mask effectively suppresses the cross-elemental contamination and enhances the image contrast for Ti; on the other hand, it inevitably decreases the amount of information for building the model and hence has a negative impact on the visual quality of the fused Ag image.

6.3. EXPERIMENTS AND RESULTS

In this section, we evaluate the modified fusion algorithm on a new experimental dataset of a gate-all-around nanowire (NW) device [2, 3]. The volume of interest contains two fins (i.e., two sets of NWs), two source/drain, and one gate [2, Fig. 2]. It was mounted on a Fischione 2050 on-axis rotation tomography holder and scanned in a double-corrected FEI Titan^3 G2 60-300 S/TEM. The microscope was operated with an accelerating voltage of 120 kV, a beam current of 250 pA, and a convergence angle of 10 mrad. Projections of the device were collected from -90° to $+90^\circ$ with an increment of 3°; at each angle, one HAADF-STEM image with a size of 512×512 pixels (7.8 Å/pixel) and the equi-sized EDS maps (net counts) of all chemical elements were recorded. Fig. 6.3 gives an example of the experimental tilt-series.

6.3.1. IMPLEMENTATION DETAILS

Throughout this chapter, we implemented the modified fusion framework in Python 3.7. For the unchanged steps, we used the same parameters as in [1, Sec. IV.B]; FBP, SIRT, SART, and EM were computed with the ASTRA toolbox [5], feature images were generated with the scikit-image package [7], and the regression problem was solved by `PLSRegression()` in scikit-learn [8]. For the newly introduced mask, we set the low and high threshold τ_L and τ_H to half of the maximum intensity of the Gaussian-filtered denoised image and the threshold computed by isodata algorithm [9], considering that all the reconstructions of $e^\star$ were stretched to $[0, 255]$ before denoising. Note that a lower

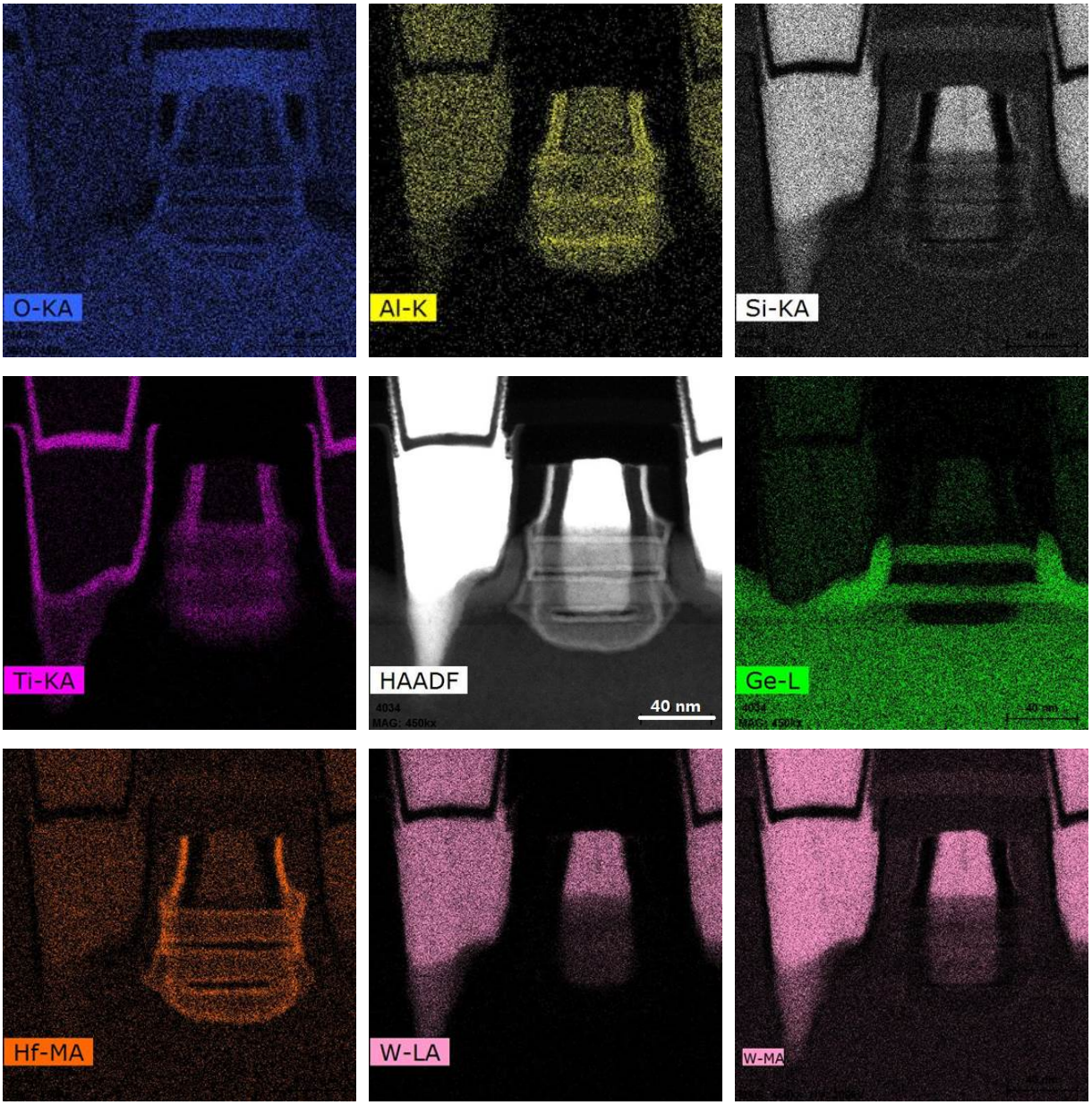


Figure 6.3: Experimental HAADF projection and EDS maps of a gate-all-around nanowire device [2, 3]. Germanium (Ge), the nanowire, is surrounded by hafnium dioxide (HfO_2), which is a high-κ dielectric material.

τ_L may include more fine details but at the expense of letting cross-elemental contamination pass.

6.3.2. RESULTS

Fig. 6.4–6.6 plot the reconstruction slices of HAADF, Ge, Hf, W and Ti in three directions: across fins, along fins, and perpendicular to the z-axis. Since these four elements do not spread across the sample but concentrate in a certain range, we divide the SIRT reconstructions into 10 blocks in the z-direction and quantify the visibility weight block by block, see Table 6.1. As shown in Fig. 6.4–6.6, fusion algorithm delivers more visually pleasing results (regarding, for example, edge continuity, foreground homogeneity,

Table 6.1: Visibility weight of chemical elements chosen for fusion; computed by dividing SIRT reconstructions with a size of 512 × 512 × 512 voxels into 10 blocks in the z-direction. Bold numbers indicate where the corresponding element is present in the sample.

Ti	-0.03	-0.03	-0.02	0.02	**0.20**	**0.22**	**0.23**	**0.27**	**0.20**	**0.20**
Ge	**0.91**	**0.91**	**0.83**	**0.78**	**0.47**	0.07	-0.11	-0.07	-0.03	0.01
Hf	0.05	0.05	0.11	0.19	**0.42**	0.15	0.05	0.07	0.02	0.20
W	0.05	0.05	0.06	0.07	**0.33**	**0.71**	**0.89**	**0.84**	**0.85**	**0.74**

background cleanness, etc.) than SIRT for all the elements expect for Ti due to its low visibility weight (max. 0.27). Moreover, it reveals some fine structures that are almost hidden in SIRT. For instance, arrow 4 points to a brightness inversion between SIRT (bright) and fusion (dark). Since dark contrast also appears in HAADF, we may infer that there is a void in W.

While the first results look promising, the newly added mask cannot distinguish features from noise. As a result, it may lead to missing low-contrast structures and/or close small gaps. For example, W pointed by arrows 2 and 6 has moderate intensity in HAADF and SIRT but is masked out in fusion; such artifacts can also be observed in Ti (arrow 3). Furthermore, HfO_2 layers (arrows 1 and 5), which should be thin and well-separated, are connected in our results due to the low EDS spatial resolution. Note that these local details cannot be reflected from the global reliability evaluation metrics (variance explained, correlation coefficient), which are (0.95, 0.97) for Ge, (0.89, 0.94) for Hf, (0.96, 0.98) for W, and (0.89, 0.95) for Ti.

6.4. DISCUSSION AND CONCLUSION

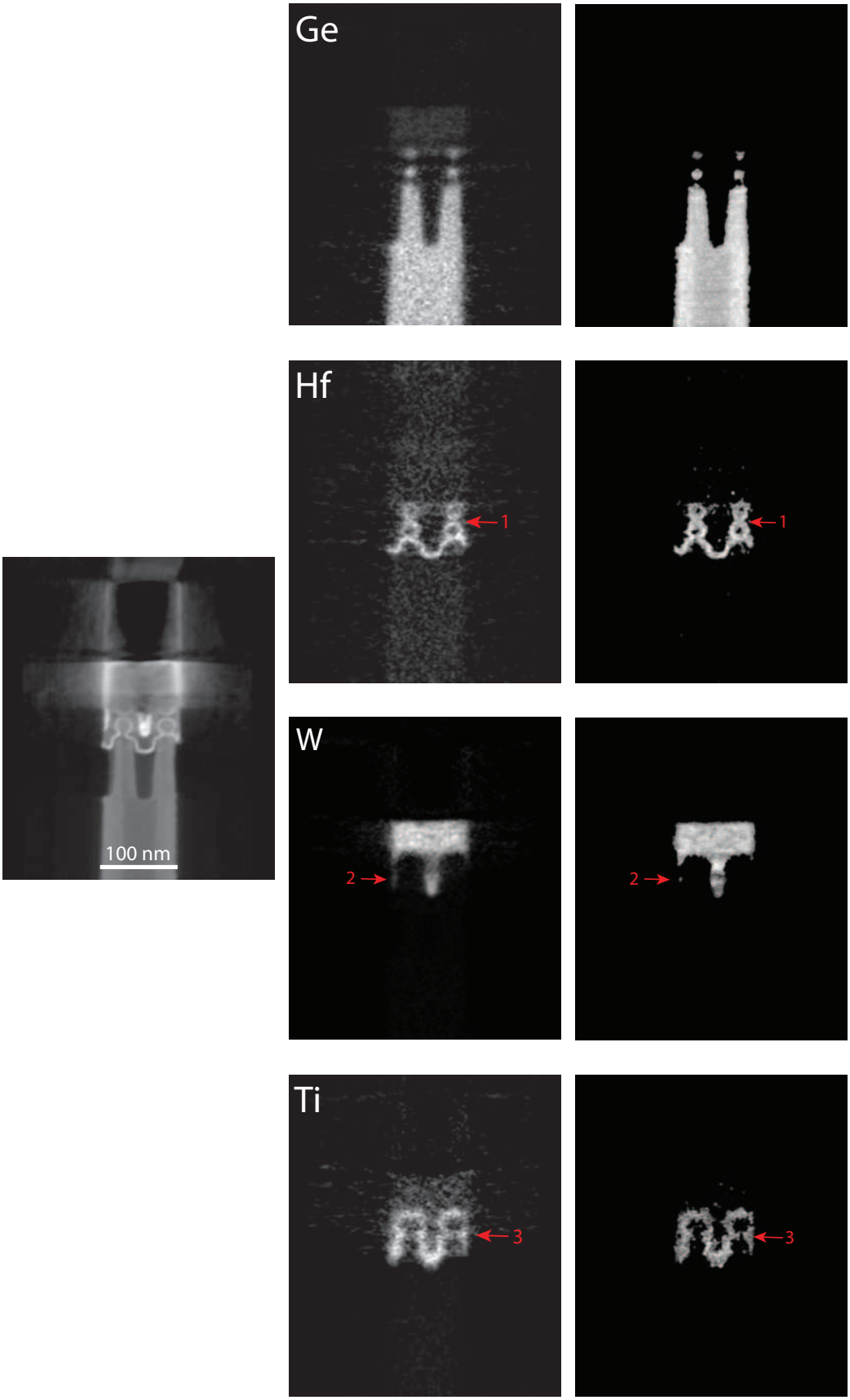
In this chapter, we added a weighted mask to our earlier fusion framework and evaluated its pros and cons on an experimental dataset of a complex Ge nanowire device. Results showed that masking can enable light element fusion by suppressing the cross-elemental contamination. However, it may also mask out desired features that have low contrast in the EDS reconstructions.

As for computational time, it took ~15 mins for tomographic reconstruction and ~80 mins on average for the following fusion, using a high performance compute server equipped with two Intel Gold 6148 CPUs[1] and four Tesla P100 GPUs[2]. Since computational resources (time and memory) were mainly spent on model building, namely partial least squares regression, we suggest to divide large image volumes into smaller blocks if the target element does not spread across the sample but concentrate in a certain range. Moreover, code optimization such as parallelism would greatly reduce the computational complexity.

While we used the same parameters as before [1] for the modified framework, other options are also feasible. For instance, if the EDS tilt-series has only a small number

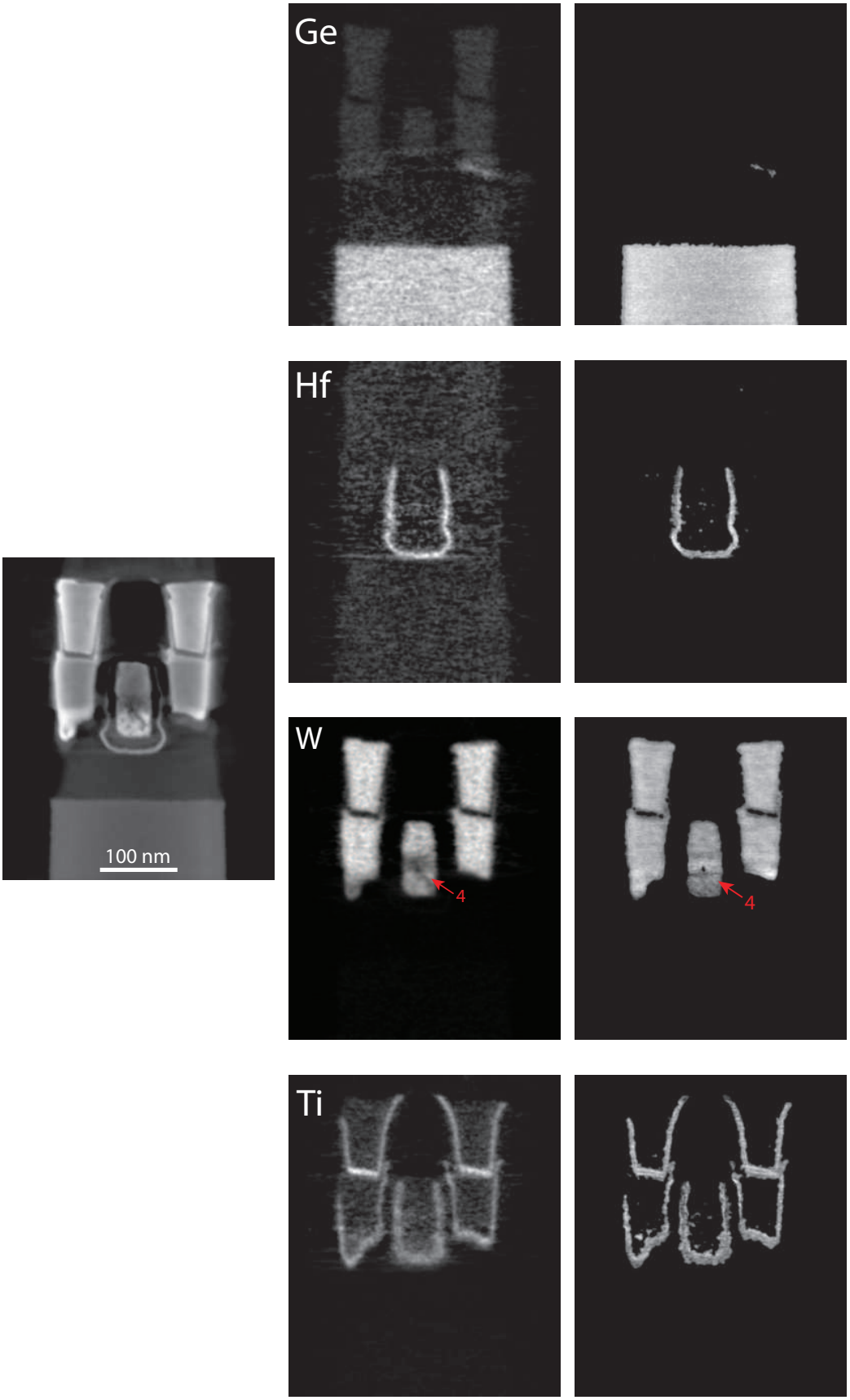
[1] 80 cores in total and 384 GB RAM, with one core being used to 100%, ~5.3% memory occupied for feature generation and ~32% for model building.

[2] Each GPU has 16 GB dedicated memory, only used for tomographic reconstruction with the ASTRA toolbox [5].



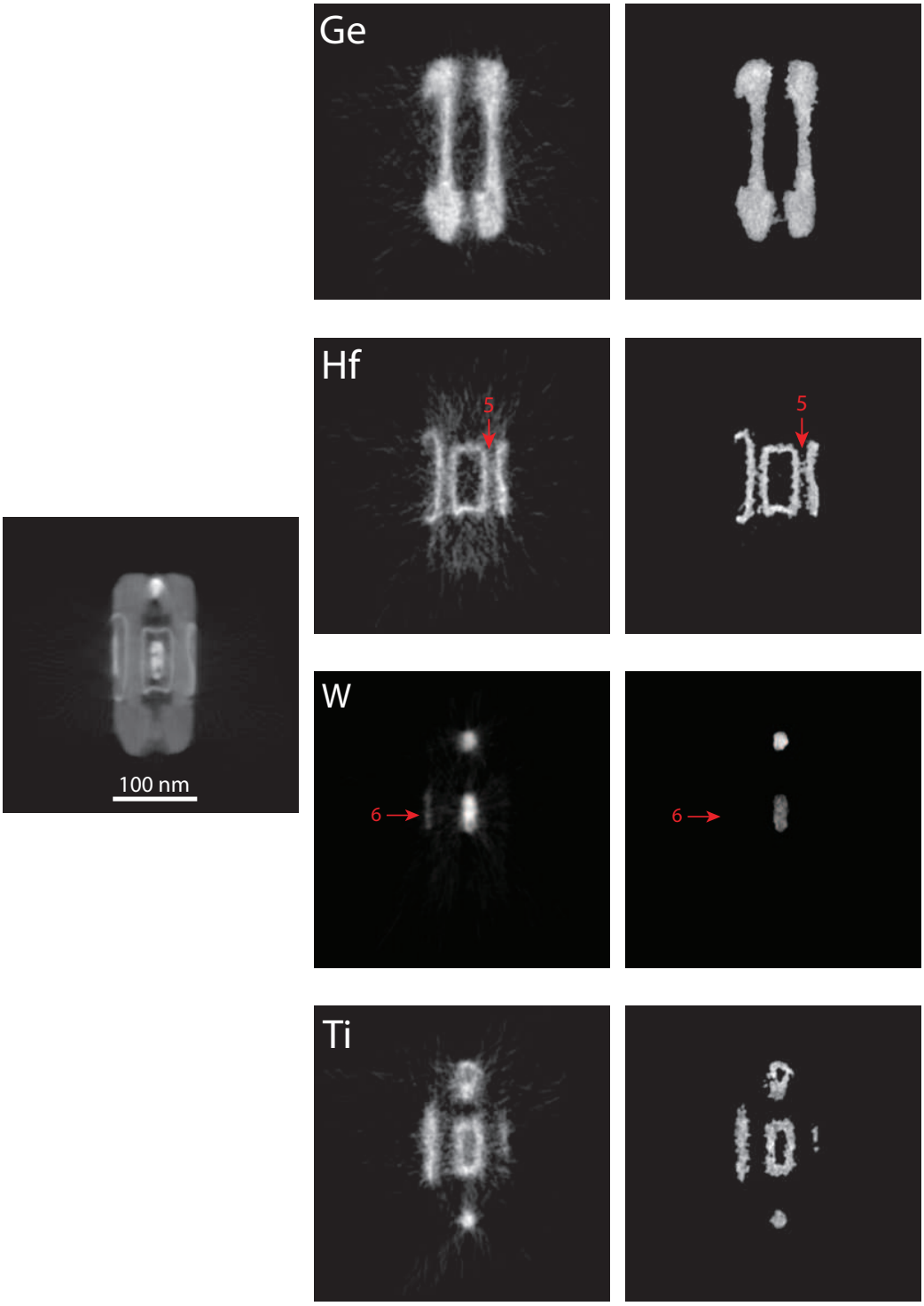
(a) HAADF (b) SIRT (c) Fusion

Figure 6.4: (a) A HAADF reconstruction slice across the fins. (b) and (c) are the corresponding EDS reconstructions computed with SIRT and fusion. The newly added mask may close small gaps (arrow 1) and lead to missing low-contrast structures (arrows 2 and 3).



(a) HAADF (b) SIRT (c) Fusion

Figure 6.5: (a) A HAADF reconstruction slice along the fins. (b) and (c) are the corresponding EDS reconstructions computed with SIRT and fusion. A dark contrast can be observed in HAADF and fusion (arrow 4) but is missing in SIRT.



(a) HAADF (b) SIRT (c) Fusion

Figure 6.6: (a) A HAADF reconstruction slice perpendicular to the z-axis. (b) and (c) are the corresponding EDS reconstructions computed with SIRT and fusion. Masking may connect thin layers (arrow 5) and mask out low-contrast features (arrow 6).

of elemental maps and/or low-dose, one can opt for an over-regulated TV reconstruction instead of the denoised image to generate a continuous mask. Furthermore, if the tomography software does not offer multiple reconstruction algorithms, users can add new filters, principal component analysis, and non-negative matrix factorization to step (4) for feature expansion; note that this operation may put extra burden on computation.

The newly added mask is extracted by hysteresis thresholding, which is simple and purely data-driven. Consequently, the result may be jaggy and not be able to distinguish features from noise. To improve this, one can turn to model-driven segmentation approaches such as active contour, or user-interaction. Since users, in general, have considerable prior knowledge of the sample under study, they can help to include or crop certain areas in the initial mask.

REFERENCES

[1] Y. Guo, R. Aveyard, and B. Rieger, "A multichannel cross-modal fusion framework for electron tomography," *IEEE Transactions on Image Processing*, vol. 28, pp. 4206–4218, 2019.

[2] P. Favia et al., "TEM investigations of gate-all-around nanowaire devices," *Semicond. Sci. Technol.*, vol. 34, pp. 124003, 2019.

[3] H. Bender et al., "3D characterization of nanowire devices with STEM based modes," *Semicond. Sci. Technol.*, vol. 34, pp. 114001, 2019.

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

[5] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA toolbox: a platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.

[6] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679–698, 1986.

[7] S. van der Walt et al., "Scikit-image: image processing in Python," *PeerJ*, vol. 2, pp. e453, 2014.

[8] F. Pedregosa et al., "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[9] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, pp. 630–632, 1978.

7

Conclusion

7.1. TECHNICAL DEVELOPMENTS AND FOLLOW-UP RESEARCH

In materials science, HAADF projections and EDS maps can be simultaneously acquired in STEM to reveal the structural and compositional information of hard nanomaterials. The latter, however, suffer from a low SNR and hence require advanced tomography techniques with hand-tuning parameters to enable a 3D element-wise reconstruction. In this work, we first paved the way for automatic weighting factor (or regularization parameter) selection to prevent the tedious and time-consuming hand-tuning. Then, we developed a fine-tuning-free approach that integrates the HAADF and EDS to concurrently accomplish high SNR and high elemental specificity.

7.1.1. PARAMETER SELECTION FOR RECONSTRUCTION TECHNIQUES

In **Chapter 2**, we introduced an image quality metric,which quantifies the quality of tomograms in terms of cross-atomic contamination and noise, to automatically find the optimal weighting factor in HEBT. We evaluated our metric on an experimental dataset of an Au-Ag core-shell nanoparticle and showed that it can replace the mean squared error (MSE) used in [1] if no ground truth is available. Instead of exhaustive search (for the optimum of the weighting factor), we adopted polynomial fitting to approximate the curve of metric versus weighting factor; this results in the shortening of computational time[1] to 10%.

Chapter 3 presented another no-reference image quality metric to detect under- and over-regularization of TNV. Specifically, we used the oriented structure strength described by the highest two responses in orientation space to simultaneously measure the sharpness and noisiness of reconstruction images. Based on simulated and experimental datasets of Au-Ag core-shell nanoparticles, we showed that our metric can replace the full-reference correlation coefficient in practice. In this case, we opted for golden section search rather than brute-force search to reduce the computational time to 15%, as the curve of metric versus regularization parameter has a distinct maximum attained for the best reconstruction quality.

In **Chapter 4**, we considered the popular TV-aided reconstruction technique, for which we used Gaussian gradient magnitude and spline interpolation to quantify the regularization strength and reduce the computational time, respectively. Although experimental results of homogeneous samples (e.g., Au nanotriangle and nanosphere [2]) showed that our method is effective to find the sweet spot between noise suppression and structure blurring, we noted that its robustness should be further explored on inhomogeneous samples with more high spatial frequency components.

All the metrics presented in Chapter 2–4 can select the close-to-optimal parameter for advanced reconstruction techniques without any reference. Furthermore, polynomial fitting, golden section search, and spline interpolation also help to accelerate the reconstruction-assessment pipeline. While the first results look promising, a thorough follow-up study is still needed to guarantee their reliability for industrial applications.

1. As we mainly considered homogeneous samples with very few high spatial frequency components, the proposed quality metrics must be validated on more experimental datasets of nanomaterials with different geometries and/or acquired

[1] Computational time refers to the amount of time spent on reconstruction and assessment.

under different microscope settings.

2. Parameter trimming [3] can help to reduce the computational time as well. By embedding the developed assessment modules into the iterative reconstruction procedure, one may directly monitor the image quality after each iteration and then trim values that are unlikely to achieve the best quality upon convergence.
3. A comprehensive guideline is desired to support (image assessment) algorithm developments because the “optimal" parameters are highly dependent on user requirements, including but not limited to higher (gray-scaled) reconstruction quality, lower (binary) segmentation error, or both.

7.1.2. HAADF-EDS CROSS-MODAL FUSION

Although the proposed quality assessment algorithms can automatically select the weighting factor or regularization parameter in an objective and repeatable manner, a fine-tuning-free approach is still preferred. To this end, in **Chapter 5** we presented a regression based cross-modal fusion framework that has no hand-tuning parameter. HAADF and EDS were linked as follows: (*i*) compute multiple tomograms from HAADF and EDS tilt-series using different reconstruction techniques; (*ii*) generate many feature images from each tomogram; (*iii*) adopt partial least squares regression to assess the connection between these feature images and the reconstruction of the target element. Based on simulated and experimental datasets of semiconductor devices consisting of chemical elements with close atomic numbers, we demonstrated that our algorithm can not only produce continuous edges, homogeneous foreground, and clean background in its element-specific reconstructions, but also more accurately preserve fine structures than state-of-the-art tomography techniques.

In **Chapter 6**, we further added a weighted mask to the earlier framework and applied it to a new experimental dataset of Ge nanowire device with more complex structures and more element types. Fusion results were improved in three aspects: (*i*) alleviated cross-elemental contamination; (*ii*) enabled interpretation across and along fins; (*iii*) light element fusion became possible if it has a distinguishable shape in the sample.

While the modified algorithm can deal with more complex multi-elemental nanostructures, there is still room for improvement. First, the newly added mask is computed by a simple and purely data-driven segmentation approach; hence, it may be jaggy and mask out low-contrast structures or close small gaps. To improve this, one can opt for model-driven segmentation (e.g., active contour) or user-interaction. Since users typically have substantial prior knowledge of the device under study, they can roughly tell whether a certain area should be included from the initial mask. Moreover, the current framework uses reconstructions computed by the ASTRA toolbox [4] as its input for fusion. Although we made it compatible with the outputs from Inspect3D [5], its portability still needs to be tested on other open and commercial 3D tomography softwares.

7.2. OUTLOOK

According to Moore’s Law, the number of transistors in an integrated circuit (IC) would double around every two years. To meet these expectations, electronic components used to fabricate ICs have been dramatically decreased in size yet increased in complexity.

While the innovative fin-based field-effect transistor (FinFET) is mainstream at present, it will eventually be extended to gate-all-around (GAA) FET where lateral GAA (LGAA) will be implemented first, followed by vertical GAA (VGAA) [6]. Moreover, other devices have also been considered as candidates to go beyond the scaling limits of complementary metal-oxide-semiconductor (CMOS), namely "beyond CMOS" [6–8]. Although their structures are not necessarily more complex than the VGAA, they are made of new materials, including graphene, carbon nanotubes and molybdenum disulfide (MoS_2) [6, Fig. 1].

As stated above, future semiconductor devices will be more complex in shape and/or contain more (new) types of materials. In that case, the combined HAADF-EDS strategy [9–11] is not sufficient for metrology [6] because (*i*) light elements can only produce very low contrast in HAADF; (*ii*) high electron dose required to compensate the low SNR of EDS cannot be performed on beam-sensitive materials. In this section, we first introduce some emerging STEM imaging techniques that could meet the new demands (e.g., low electron dose to prevent beam damage, low-contrast materials characterization). Then, we discuss what opportunities they would bring to the subsequent image processing.

7.2.1. FUTURE OF IMAGING IN STEM

Undersampling (e.g., tilt- and image-undersampling [12]) that exploits the concept of compressed sensing could lower the electron dose. For instance, Béché et al. measured a subset of pixels in projections by placing a beam blanker to pseudorandomly deflect the beam during raster scan [13]. While undersampling can be conveniently incorporated in data acquisition, open questions still need to be answered. First, compressed sensing assumes that the original image has a sparse representation in a specific base [13]. Since the desired sparsity can be achieved in various ways such as random sampling [13] and sparse gradient [14], it is very important to evaluate which method is more suitable to what imaging modalities and samples. Moreover, undersampling may have a negative effect on the subsequent tomography steps. For example, a very large increment between consecutive projections would result in inaccurate alignment and/or insufficient amount of information for reliable reconstruction.

To reveal weakly scattering objects in STEM, coherent phase-contrast imaging has been investigated [15–19]. However, the conventional annular bright-field mode [15, 16] is experimentally demanding because it is intolerant to residual aberrations and sample mistilt [19]. To generate linear phase contrast with high efficiency, Ophus et al. placed a patterned phase plate in the probe-forming aperture and a pixelated detector below the sample to record the transmitted beam at each scanning point [17]. This MIDI-STEM (matched illumination and detector interferometry) approach was demonstrated on a heterogeneous sample consisting of gold nanoparticles and amorphous carbon and delivered significantly more linear measurements than the traditional BF- and ADF-STEM [17].

Electron ptychography also showed its great potential to produce strong contrast for both heavy and light elements with much lower electron dose [18–20]. For instance, Yang et al. achieved concurrent Z- and phase-contrast imaging by inserting a fast pixelated camera in the ADF detector plane [18]. The authors reconstructed phase images from the collected diffraction pattern and maximized its contrast through a post-acquisition

correction for the lens aberrations [18]. Since this framework allows for simultaneous coherent and incoherent imaging, it may be compatible with spectroscopy techniques like EDS.

7.2.2. FUTURE OF IMAGE PROCESSING FOR STEM TOMOGRAPHY

Over the past few years, deep learning has been vividly researched in the image processing community to tackle a broad variety of problems, including but not limited to recognition, segmentation, and multimodal fusion. It uses multiple processing layers to progressively extract higher-level features from the raw pixel values [21]. Thus, it depends less on feature extraction and selection than conventional machine learning [22].

Deep learning requires a large amount of data for training [22]; for instance, ImageNet designed for object recognition contains more than 15 million labeled images belonging to over 22,000 categories [23]. However, datasets in the field of STEM tomography are often confidential and have inadequate quantities. This problem might be overcome by transfer learning, that is, reusing one network developed for a base dataset as the starting point of another network to be trained on a target dataset [24]. Data augmentation can also effectively multiply the amount of data via rotations, cropping, and so on [22].

Deep learning is promising to deal with limited STEM data. In fact, it has already been tested and validated in computed tomography where images are formed by X-rays [25–27]. For instance, Pelt et al. evaluated their mixed-scale dense convolutional neural network on various common limitations, including limited number of projections, missing wedge, and low-dose [27]. Nevertheless, the implementation for electron tomography has not caught up [2]. Note that deep learning can be performed at different stages of tomography, for the purpose of preprocessing noisy projections, learning prior knowledge, and enhancing the image quality of reconstructions computed by standard algorithms.

Furthermore, deep learning may play an important role to automate the reconstruction process, i.e., blindly predicting parameters for advanced reconstruction techniques through pre-built quality assessment models. This also raises a few questions and concerns. First, conventional learning-based quality predictors designed for pictures [22] are not suitable for STEM images. Second, collecting adequate quantities of expert-labeled data is very difficult and time-consuming. Such paucity issue may be addressed by transfer learning and data augmentation. Since STEM projections and reconstructions are not natural images, transformations (e.g., rotation, reflection, etc. [22]) would not significantly degrade the perceived quality. Last but not least, how to build a general assessment model that works for more reconstruction algorithms should be carefully investigated.

Finally, applying deep learning to multimodal fusion in STEM tomography is certainly attractive, considering its great success in many other application fields such as medical imaging and remote sensing [28]. This can be implemented in two ways. On the one hand, the element-wise information provided by one modality can be used to separate the mixed signals from another modality (e.g., EDS and HAADF), namely unmixing. On the other hand, ADF and phase images [18], which respectively contain the information of heavy and light elements, can be combined into a composite scene (i.e.,

integration) to fully determine the sample structure. Although the image size in STEM tomography is typically large (e.g., $512 \times 512 \times 512$ voxels), we believe that parallel computing and user-friendly libraries would make deep learning possible.

REFERENCES

[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*, vol. 174, pp. 35–45, 2017.

[2] E. Bladt, D. M. Pelt, S. Bals, and K. J. Batenburg, "Electron tomography based on highly limited data using a neural network reconstruction technique," *Ultramicroscopy*, vol. 158, pp. 81–88, 2015.

[3] H. Liang and D. S. Weller, "Regularization parameter trimming for iterative image reconstruction," in *Proceedings of the 49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 755–759.

[4] W. van Aarle, W. J. Palenstijn, J. D. Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The ASTRA toolbox: a platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.

[5] Thermo Fisher Scientific, "3D Tomography with Inspect 3D Software," 2019. [Online]. Available: https://www.thermofisher.com/nl/en/home/electron-microscopy/products/software-em-3d-vis/inspect-3d-software.html

[6] N. G. Orji *et al.*, "Metrology for the next generation of semiconductor devices," *Nature Electronics*, vol. 1, pp. 532–547, 2018.

[7] S. B. Desai *et al.*, "MoS_2 transistors with 1-nanometer gate lengths," *Science*, vol. 354, pp. 99–102, 2016.

[8] T. Roy *et al.*, "Dual-gated MoS_2/WSe_2 van der Waals tunnel diodes and transistors," *ACS Nano*, vol. 9, pp. 2071–2079, 2015.

[9] J. Zhu, Y. Shen, and S. Zhao, "Chemical analysis of semiconductor devices using TEM energy-dispersive X-ray spectroscopy (EDS) and electron energy-loss spectroscopy (EELS)," in *Proceedings of the 23rd International Symposium on the Physical and Failure Analysis of Integrated Circuits*, July 2016, pp. 86–93.

[10] B. Fu and M. Gribelyuk, "3D analysis of semiconductor devices: a combination of 3D imaging and 3D elemental analysis," *Journal of Applied Physics*, vol. 123, p. 161554, 2018.

[11] H. Bender *et al.*, "Combined STEM-EDS tomography of nanowire structures," *Semiconductor Science and Technology*, vol. 34, p. 114002, 2019.

[12] H. Vanrompay *et al.*, "Experimental evaluation of undersampling schemes for electron tomography of nanoparticles," *Particle & Particle Systems Characterization*, vol. 36, p. 1900096, 2019.

[13] A. Béché, B. Goris, B. Freitag, and J. Verbeeck, "Development of a fast electromagnetic beam blanker for compressed sensing in scanning transmission electron microscopy," *Applied Physics Letters*, vol. 108, p. 093103, 2016.

[14] B. Goris, W. van den Broek, K. J. Batenburg, H. H. Mezerji, and S. Bals, "Electron tomography based on a total variation minimization reconstruction technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012.

[15] S. Findlay *et al.*, "Robust atomic resolution imaging of light elements using scanning transmission electron microscopy," *Appl. Phys. Lett.*, vol. 95, p. 191913, 2009.

[16] L. Gu *et al.*, "Direct observations of lithium staging in partially delithiated $LiFePO_4$ at atomic resolution," *J. Am. Chem. Soc.*, vol. 133, pp. 4661–4663, 2011.

[17] C. Ophus *et al.*, "Efficient linear phase contrast in scanning transmission electron microscopy with matched illumination and detector interferometry," *Nature Communications*, vol. 7, p. 10719, 2016.

[18] H. Yang *et al.*, "Simultaneous atomic-resolution electron ptychography and Z-contrast imaging of light and heavy elements in complex nanostructures," *Nature Communications*, vol. 7, p. 12532, 2016.

[19] J. G. Lozano *et al.*, "Low-dose aberration-free imaging of Li-rich cathode materials at various states of charge using electron ptychography," *Nano Letters*, vol. 18, pp. 6850–6855, 2018.

[20] Y. Jiang *et al.*, "Electron ptychography of 2D materials to deep sub-ångström resolution," *Nature*, vol. 559, pp. 343–349, 2018.

[21] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, 2015.

[22] J. Kim *et al.*, "Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment," *IEEE Signal Processing Magazine*, vol. 34, pp. 130–141, 2017.

[23] J. Deng *et al.*, "ImageNet: A large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[24] J. Yosinski *et al.*, "How transferable are features in deep neural networks," *Advances in Neural Information Processing Systems*, vol. 27, pp. 3320–3328, 2014.

[25] K. H. Jin *et al.*, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, pp. 4509–4522, 2017.

[26] X. Yang *et al.*, "Low-dose X-ray tomography through a deep convolutional neural network," *Scientific Reports*, vol. 8, p. 2575, 2017.

[27] D. M. Pelt, K. J. Batenburg, and J. A. Sethian, "Improving tomographic reconstruction from limited data using mixed-scale dense convolutional neural networks," *Journal of Imaging*, vol. 4, p. 128, 2018.

[28] Y. Liu *et al.*, "Deep learning for pixel-level image fusion: recent advances and future prospects," *Information Fusion*, vol. 42, pp. 158–173, 2018.

ACKNOWLEDGEMENTS

To be able to finish this thesis, I am in debt to ...

- to Bernd for meeting all my expectations for a super-supervisor;
- to Richard for guiding me to the field of electron microscopy;
- to Zhichao for answering me tons of questions starting with a “Hey, bro";
- to Willem Jan, Joost, Remco and Hugo for the illuminating discussions;
- to Hans and Sara for sharing the experimental datasets of nanoparticles;
- to Holger for lending me dozens of “five minutes";
- to Yuchen for giving me an interactive TEM tutorial at Thermo Fisher;
- to Tian and Boling for helping me to adapt myself to the new environment;
- to Anna, Leon, Tom and Jos W. for creating a good atmosphere in the room;
- to Robert, Lena, Christiaan, Jelle H. and Jelle S. for arranging the summer BBQ, winter hotpot, Sinterklaas, pizza nights, board games and other off-work events;
- to Shan for the chewing gums, the bird, the bear and, of course, the laughter;
- to Ronald for bringing a playful vibe every single day;
- to Annelies and Angela for being supportive and responsible;
- to Leo for making my trip worry-free by taking over Signals and Systems;
- to all the QI/CI members for making at least one third of my day joyful;
- to Kaifeng for teaching me how to conduct research at the very beginning;
- to Yifeng for always pairing the homemade goodies with Baileys;
- to Yun for being the icing on our annual travel cake;
- to my family, who had a front row seat to everything. 我想感谢我的家人,特别是父母,表姐和大姨。谢谢你们对我的鼓励，以及对我每一个小小成就的肯定。

About the Author

Yan Guo

Yan Guo was born in Beijing, China on May 28, 1990. She received the B.E. degree in telecommunications engineering from the Beijing University of Posts and Telecommunications, China, in 2012, and the M.Sc. degree in communications engineering from the RWTH Aachen University, Germany, in 2016. From 2012 to 2015, she was with the Institute for Communication Technologies and Embedded Systems, RWTH Aachen University, as a Student Assistant within the framework of FP7 project ICT-317669 METIS. From 2015 to 2016, she was an Assistant Research Engineer within the framework of Horizon 2020 FANTASTIC-5G with the European Research Center, Huawei Technologies Düsseldorf GmBH, Germany. In 2016, she started as a Ph.D. candidate in the Department of Imaging Physics, Delft University of Technology, The Netherlands. Her research focused on image processing for the application field of electron microscopy, of which the most important findings are presented in this thesis.

LIST OF PUBLICATIONS

1 **Y. Guo**, and B. Rieger, “Parameter selection for regularized electron tomography without a reference image," in *Proceedings of the Scandinavian Conference on Image Analysis (SCIA)*, pp. 452–464, Apr. 2019.

2 **Y. Guo**, R. Aveyard, and B. Rieger, "A multichannel cross-modal fusion framework for electron tomography," *IEEE Transactions on Image Processing*, vol. 28, pp. 4206–4218, Mar. 2019.

3 **Y. Guo**, and B. Rieger, "Image fusion of X-ray and electron tomograms," in *Proceedings of the 2018 IEEE International Conference on Image Processing*, pp. 863–867, Oct. 2018.

4 **Y. Guo**, and B. Rieger, "No-reference weighting factor selection for bimodal tomography," in *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1243–1247, Apr. 2018.

Propositions

accompanying the dissertation

IMAGE QUALITY ASSESSMENT AND IMAGE FUSION FOR ELECTRON TOMOGRAPHY

by

Yan GUO

1. User expectation is the key to developing robust image quality assessment algorithms. (this thesis)
2. Universally applicable tomographic reconstruction techniques do not exist.
3. No comment is better than incorrect comment.
4. While facing a problem, scientists want to understand it as deep as possible, whereas engineers want to solve it as quick as possible.
5. Perfection should be pursued pragmatically if resources are limited.
6. The best friend is the most unlikely friend.
7. Too many choices lead to nothing.
8. “Save the earth" is a cover-up of “save ourselves".
9. “The first wealth is health." (Ralph Waldo Emerson)
10. The greatest risk of global catastrophe is a highly infectious virus. (Bill Gates, 2015)

These propositions are regarded as opposable and defendable, and have been approved as such by the supervisor prof. dr. ir. B. Rieger.

Stellingen

behorende bij het proefschrift

IMAGE QUALITY ASSESSMENT AND IMAGE FUSION FOR ELECTRON TOMOGRAPHY

door

Yan GUO

1. De verwachting van de gebruiker is de sleutel tot het ontwikkelen van robuuste algoritmes die de kwaliteit van een afbeelding beoordelen. (deze dissertatie)
2. Universeel toepasbare tomografische reconstructietechnieken bestaan niet.
3. Geen commentaar is beter dan incorrect commentaar.
4. Geplaatst voor een probleem willen wetenschappers het zo goed mogelijk begrijpen, terwijl ingenieurs het zo snel mogelijk willen oplossen.
5. Perfectie zou pragmatisch nagestreefd moeten worden als bronnen gelimiteerd beschikbaar zijn.
6. De beste vriend is de meest onwaarschijnlijke vriend.
7. Teveel keuze leidt tot niets.
8. “Red de aarde" is een verhulling van “red onszelf".
9. Gezondheid is het grootste goed. (no source)
10. Het grootste risico op een globale catastrofe is een zeer besmettelijk virus. (Bill Gates, 2015)

Deze stellingen worden opponeerbaar en verdedigbaar geacht en zijn als zodanig goedgekeurd door de promotor prof. dr. ir. B. Rieger.