

Numerical methods for low-dose EDS tomography

Zhichao Zhong^{*,a}, Willem Jan Palenstijn^a, Nicola Roberto Viganò^a, K. Joost Batenburg^{a,b}

^a Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

^b Mathematical Institute, Universiteit Leiden, Leiden, The Netherlands



ABSTRACT

Energy-dispersive X-ray spectroscopic (EDS) tomography is a powerful three-dimensional (3D) imaging technique for characterizing the chemical composition and structure of nanomaterials. However, the accuracy and resolution are typically hampered by the limited number of tilt images that can be measured and the low signal-to-noise ratios (SNRs) of the energy-resolved tilt images. Various sophisticated reconstruction algorithms have been proposed for specific types of samples and imaging conditions, yet deciding on which algorithm to use for each new case remains a complex problem. In this paper, we propose to tailor the reconstruction algorithm for EDS tomography in three aspects: (1) model the reconstruction problem based on an accurate assumption of the data statistics; (2) regularize the reconstruction to incorporate prior knowledge; (3) apply bimodal tomography to augment the EDS data with a high-SNR modality. Methods for the three aspects can be combined in one reconstruction procedure as three modules. Therefore, a reconstruction algorithm can be constructed as a ‘recipe’. We also provide guidelines for preparing the recipe based on conditions and assumptions for the data. We investigate the effects of different recipes on both simulated data and real experimental data. The results show that the preferred recipe depends on both acquisition conditions and sample properties, and that the image quality can be enhanced using a properly tailored recipe.

1. Introduction

In materials science, the compositional characterization in three dimensions (3D) is important for understanding the properties of nanomaterials. Energy-dispersive X-ray spectroscopic (EDS) STEM allows mapping of the distributions of chemical elements in 2D by detecting the X-rays emitted from the specimen. A 3D volumetric image of these chemical distributions can then be reconstructed from a tilt series of the 2D maps [1–3]. Such a technique is referred to as *EDS tomography*. However, EDS tomography is limited by many practical issues [4–6]. One of the most significant issues is the limited number of detected X-ray counts caused by low emission rates and small solid angles of detectors. As a result, strong Poisson noise is present in the tilt series of elemental maps, which leads to reconstructions with low signal-to-noise ratios (SNRs). In addition, the number of tilt images is often small due to the long data acquisition time. The limited number of tilts results in an ill-posed inverse problem, which, together with the high levels of noise, strongly limits the accuracy of the reconstructed volume. The possibilities for improving the quality of the measured data are often limited by the electron dose that the sample can withstand.

Ill-posed inverse problems have been studied extensively in (electron) tomography, and various reconstruction methods have been developed ([3, Chapter 7] and [7]). However, choosing the most appropriate algorithm in the context of a specific sample and specific imaging conditions is currently problematic for practitioners in EDS tomography.

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

First of all, we describe how the tomographic reconstruction process can be modeled as an inverse problem with Poisson statistics, whereas the conventional alternative is based on Gaussian-statistic data. For instance, the simultaneous iterative reconstruction technique (SIRT) [8], used in [2,4], actually solves an inverse problem assuming Gaussian noise. Poisson noise is addressed in a separate denoising step and by the smoothing effects introduced by SIRT. However, smoothing blurs the images and reduces the resolution. Also, inaccurate modeling may introduce artifacts in the reconstructed images. For EDS mapping with low X-ray counts, it is reasonable to assume the image intensities as measurements of Poisson processes like in many other photonic imaging modalities, e.g. positron emission tomography (PET). Image reconstruction with Poisson statistics has already been studied extensively [9,10].

Secondly, we present the module for variational regularization methods. These have been developed to address the issue of overfitting (to noise) present in direct modeling methods such as maximum likelihood estimation (MLE) [9] in situations with extremely low counts [10–12]. For instance, total variation (TV) regularization is widely

* Corresponding author.

E-mail address: zhong@cwi.nl (Z. Zhong).

adopted. It encourages sparsity of gradients, which helps to suppress noise, promote piecewise constant structures and reduce the artifacts caused by missing data [13,14].

In addition, for EDS tomography, the reconstructions for different chemical elements often share image features, such as edges. Total nuclear variation (TNV) regularization – an expansion of TV – encourages such common edge locations of correlated reconstructions in addition to promoting sparse gradients [15,16]. There are many other regularization methods such as total generalized variation which encourages piecewise smooth structures [17]. In this paper, we focus on TV and TNV as our ingredients for the purpose of demonstration.

Thirdly, even with regularization, the reconstructions may still be highly inaccurate when lacking accurate data. In situations with strong noise, TV regularization may introduce staircasing artifacts in the reconstruction [14]. The third module augments the reconstruction with additional accurate data by combining EDS tomography with other imaging modalities. Here, we use the bimodal HAADF-EDS tomography (HEBT) technique that was proposed in our previous paper [18]. HEBT considers the HAADF-STEM projection images, which usually have higher SNRs and resolution, to be the weighted sum of the EDS maps for all present elements. In the original paper, HEBT is based on a Gaussian noise model. In this paper, we contribute to HEBT by introducing the formulation for reconstruction with Poisson statistics.

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

The remainder of this paper is structured as follows. In Section 2, we illustrate the theory and the guidelines for choosing ingredients of an algorithmic recipe. In Section 3, we investigate and compare the performance of different recipes on simulation and experimental data. In the last section, we draw a conclusion for this paper. We do not discuss the pre-processing steps in the spectral domain, while in practice these should be carefully considered for the influence on the data statistics. Also, other issues e.g. detector shadowing effects and X-ray self-absorption strongly affect the reconstruction results, but are addressed in other papers [4–6].

2. Method

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

2.1. Notation of EDS tomography

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

We first formulate the relationship between the reconstructed image and the ideal measurement data without noise corruption. Under the thin-film assumption, the ideal data are proportional to the expected numbers of X-ray counts that are in turn proportional to the concentration of the corresponding element probed by the focused beam [19]. Thus the ideal data are proportional to the linear projection of the reconstructed quantities, which are expressed as a vector $\mathbf{g}^e \in \mathbb{R}^{M^e}$. Here M^e denotes the total number of pixels for all angles for element e ($e = 1, \dots, L$). Consider the specimen to be located in a 3D volume space discretized into N voxels. The reconstructed quantities, which are proportional to the concentration of the element, are expressed as a

vector $\mathbf{x}^e \in \mathbb{R}^N$. This linear relationship is modeled by the system of equations:

$$\mathbf{g}_i^e = \sum_{j=1}^N w_{ij}^e x_j^e. \quad (1)$$

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

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

$$\mathcal{D}_{L_2}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e) = \|\mathbf{W}^e \mathbf{x}^e - \mathbf{p}^e\|_2^2, \quad (2)$$

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

However, when the number of X-ray counts is small, the Gaussian model is not an accurate approximation anymore. A more solid assumption is to consider the real data as Poisson distributed measurements taking the ideal data as the expected values. We then use the *Kullback-Leibler (KL) divergence* [20,21] to define the data discrepancy, which is expressed as:

$$\begin{aligned} \mathcal{D}_{\text{KL}}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e) &= \sum_{i=1}^{M^e} \left(p_i^e - g_i^e + g_i^e \log \left(\frac{g_i^e}{p_i^e} \right) \right) \\ &= \sum_{i=1}^{M^e} \left(p_i^e - \sum_{j=1}^N w_{ij}^e x_j^e + \sum_{j=1}^N w_{ij}^e x_j^e \log \left(\frac{\sum_{j=1}^N w_{ij}^e x_j^e}{p_i^e} \right) \right), \end{aligned} \quad (3)$$

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

Given the data discrepancy \mathcal{D} defined by either Eq. (2) or Eq. (3), the reconstruction is computed by minimizing the discrepancy:

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}} \mathcal{D}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e). \quad (4)$$

Minimizing KL divergence \mathcal{D}_{KL} is equivalent to maximizing the log-likelihood of the Poisson distributions for \mathbf{p}^e [21], while minimizing L2 discrepancy \mathcal{D}_{L_2} corresponds to solving a least-squares problem. The popular reconstruction algorithm SIRT in fact solves the problem of minimizing a weighted version of L2 discrepancy [8].

2.2. Variational regularization

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

$$\mathbf{x}^{e*} = \underset{\mathbf{x}^e}{\operatorname{argmin}} \mathcal{D}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e) + \lambda \mathcal{R}_{\text{TV}}(\mathbf{x}^e), \quad (5)$$

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

$$\mathcal{R}_{\text{TV}}(\mathbf{x}^e) = \sum_{j=1}^N \left\| \nabla x_j^e \right\|_2, \quad (6)$$

where ∇ is the discrete approximation of the gradient operator. If the reconstruction image is 3D, ∇ approximates the gradients in the X, Y and Z directions respectively using the forward difference as $\nabla x_j = (\nabla^X x_j, \nabla^Y x_j, \nabla^Z x_j)^T$. Note that reconstructions can also be performed by stacking 2D reconstructions of each slice, for which the gradients are only computed in the X and Y directions. In practice, it is

more preferable to directly reconstruct in 3D to also incorporate regularization in the Z direction. The TV defined in this paper is called isotropic TV [14], for which the gradient magnitude at pixel location j is computed as the l_2 norm of the gradient.

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

$$\mathcal{R}_{\text{TNV}}(\{\mathbf{x}^k\}) = \sum_{j=1}^N \left\| \left(\mathbf{J}\{\mathbf{x}^k\} \right)_j \right\|_{\star}. \quad (7)$$

The Jacobian matrix at pixel position j is given by:

$$\left(\mathbf{J}\{\mathbf{x}^k\} \right)_j = \begin{pmatrix} \nabla^X x_j^1 \nabla^Y x_j^1 \nabla^Z x_j^1 \\ \vdots \\ \nabla^X x_j^Q \nabla^Y x_j^Q \nabla^Z x_j^Q \end{pmatrix}, \quad (8)$$

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

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

$$\{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} \sum_{e=1}^L \mathcal{D}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e) + \lambda \mathcal{R}_{\text{TNV}}(\{\mathbf{x}^e\}), \quad (9)$$

where the reconstructions for all elements are computed simultaneously.

TNV allows to correlate multiple reconstructions in a flexible manner. In addition to promoting common features between multiple elemental volumes, it is also possible to relate EDS tomography to other tomographic modalities, such as HAADF-STEM tomography. This is subject to having the HAADF reconstruction sharing common edges with the EDS reconstructions. More details are discussed in our paper [22]. It is even possible to use the TNV regularization to correlate with EELS-STEM tomography [23]. Despite the many possibilities to apply TNV, in this paper, we focus on the TNV regularization defined by Eq. (9) that correlates the EDS reconstructions for all elements.

2.3. HAADF-EDS bimodal tomography

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

The HEBT method in the original paper [18] is defined for least-squares. Here, we modify the formula so that the KL divergence can be used. The reconstruction problem of HEBT is expressed as:

$$\begin{aligned} \{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} & \alpha \mathcal{D}_{\text{L2}} \left(\sum_{e=1}^L \mathbf{W}^{h,e} \mathbf{x}^e; \mathbf{p}^h \right) + (1 - \alpha) \sum_{e=1}^L \mathcal{D}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e), \\ \text{subject to } & \mathbf{p}^h = \sum_{e=1}^L r^e \mathbf{p}^e, \\ & \mathbf{x}^e \geq 0, \quad e = 1, \dots, L, \end{aligned} \quad (10)$$

where the first term is the L2 data discrepancy for the tilt series of

HAADF-STEM images $\mathbf{p}^h \in \mathbb{R}^{M^h}$, and M^h denotes the total number of pixels for all HAADF-STEM tilt images. The matrix $\mathbf{W}^h \in \mathbb{R}^{M^h \times N}$ is the HAADF-STEM projection matrix that describes the HAADF-STEM imaging setup. r^e 's are the response ratio factors for different chemical elements. The second term is the sum of EDS data discrepancies for all the elements. The EDS data discrepancy can be chosen between KL divergence and L2 discrepancy, depending on how the noise is modeled. The parameter $\alpha \in [0, 1)$ is the trade-off weight between the HAADF-STEM and EDS data discrepancies.

Note that in the original HEBT paper, the EDS map intensities are scaled by the response ratio factors r^e , which changes the EDS data statistics. Here, instead we move the response ratio factors to the HAADF-STEM term so that the EDS maps remain unchanged. The response ratio factors r^e can be estimated based on the linear equations $\mathbf{p}^h = \sum_{e=1}^L r^e \mathbf{p}^e$ using least-squares regression [18]. Since r^e 's are assumed to be spatially invariant, we can bin the images \mathbf{p}^h and \mathbf{p}^e 's to increase the SNRs and improve the accuracy of estimated values.

2.4. Preparing the recipe

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

$$\{\mathbf{x}^{e*}\} = \underset{\{\mathbf{x}^e\}}{\operatorname{argmin}} (1 - \alpha) \sum_{e=1}^L \mathcal{D}_e(\mathbf{x}^e) + \alpha \mathcal{D}_h \left(\sum_{e=1}^L r^e \mathbf{x}^e \right) + \lambda \mathcal{R}(\{\mathbf{x}^e\}), \quad (11)$$

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

When constructing a recipe, one should carefully consider the validity of the assumptions behind ingredients. In Fig. 2, we provide a flowchart as guidelines for choosing ingredients and the conditions w.r.t. the properties of the data and the sample. These conditions are based on mathematical assumptions summarized below:

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

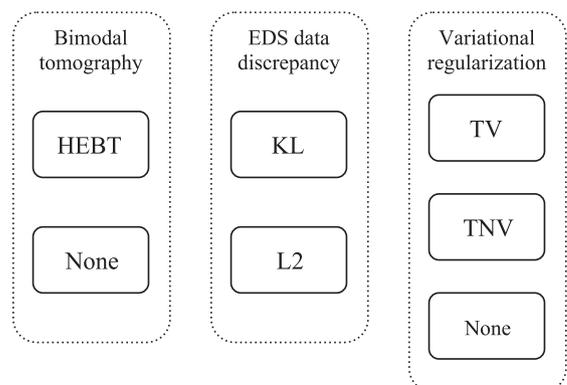


Fig. 1. Ingredients of the algorithmic recipes.

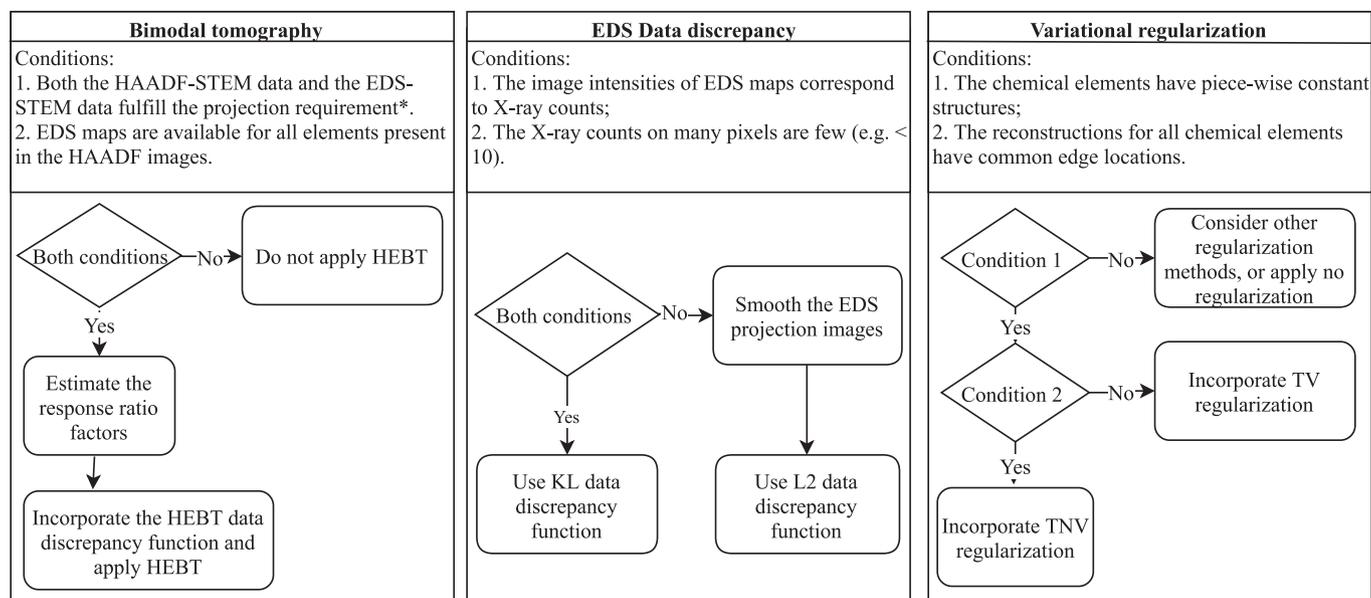


Fig. 2. Flowchart for making the recipe. (* The projection requirement refers to that the images intensities are linearly related to some integrated physical properties of the sample.)

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

2.5. Solving the reconstruction problem

After making an instance of Eq. (11) for the recipe, a numerical algorithm is needed for solving the optimization problem. We use the *Douglas-Rachford primal-dual splitting algorithm* (DR) [24] to compute the solution, which is a broadly applicable algorithm for solving convex optimization problems. For our application, the DR algorithm solves the mathematical problem of the following general form:

$$\mathbf{v} = \underset{\mathbf{v}}{\operatorname{argmin}} f(\mathbf{v}) + \sum_{k=1}^R g_k(\mathbf{A}_k \mathbf{v}), \quad (12)$$

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

In fact, all our data discrepancy and regularization terms can be cast into the form of $g_k(\mathbf{A}_k \mathbf{v})$. Therefore, different optimization problems derived from Eq. (11) can be solved using the same DR algorithm. In Appendix, we provide more details for fitting our optimization problems into Eq. (12).

3. Experiments

In this section, we investigate the performance of different recipes on simulation data as well as real experimental data. We use the DR algorithm implemented in the Operator Discretization Library (ODL) [25].

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

$$r = \frac{\sum_i (x_i - \bar{x})(v_i - \bar{v})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (v_i - \bar{v})^2}}, \quad (13)$$

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

3.1. Non-mixed phantom simulation

3.1.1. Data simulation

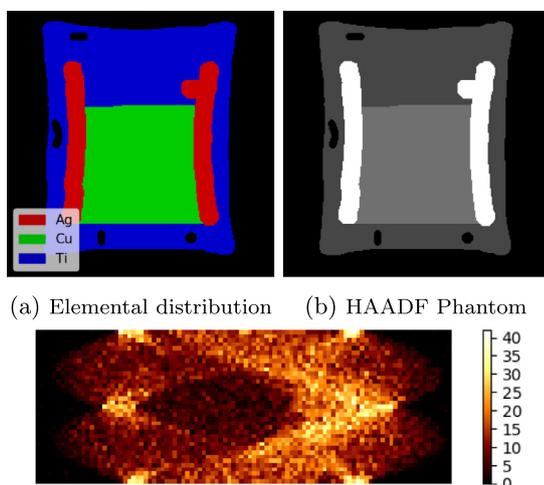
The 2D phantom resembles a structure that contains three homogeneous compositions, which are shown in different colors in Fig. 3(a). We assume that the image contrast scales are respectively $z_{Ag} = 47^{1.7}$, $z_{Cu} = 29^{1.7}$ and $z_{Ti} = 22^{1.7}$ given the corresponding atomic numbers Z of these elements, so that the contrast scales as Z^α with α chosen as 1.7 [26]. The HAADF phantom is shown in Fig. 3(b).

We simulated a tilt series of 1D projection images for the HAADF-STEM phantom for every 5° from 0° to 180° using the ASTRA Toolbox [27]. In addition, we simulated tilt series of 1D maps for each individual element. A realistic value for the image intensity can be determined by considering the incident beam current, the probe live time, the fraction of incident electrons causing ionization, the fluorescence yield, the detector solid angle, and the detector efficiency [28]. In this paper, we simply set the intensities to absolute scales close to real experimental data for the brevity of the paper. We then applied Poisson noise by drawing random numbers for expected values given by the noiseless map intensities. Fig. 3(c) shows the simulated maps with noise for Ti. The mean image intensity on non-background pixels is 11.76.

3.1.2. Reconstruction results

For this dataset we can choose KL data discrepancy, TNV regularization and HEBT as the ingredients for our preferable recipe (KL-TNV-HEBT), based on the observation that the Poisson noise is strong and the individual elements have homogeneous structures that share edges. Additionally, reconstructions based on other recipes were also performed for comparing the effects. Although reconstructions were made for all elements, only the reconstructions for Ti are shown for the brevity of this section.

For comparison, we first show the non-regularized reconstructions. The reconstruction (Fig. 4(a)) based on the KL divergence was computed by solving Eq. (4) using the DR algorithm. Fig. 4(b) shows the reconstruction based on L2 data discrepancy computed using the SIRT



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

Fig. 3. The non-mixed phantom simulation data.

algorithm for 50 iterations, combined with a pre-smoothing using a Gaussian filter ($\sigma = 1.0$). In fact, SIRT also incorporates implicit regularization on the image smoothness, which is determined by the number of iterations. The L2 reconstruction is less noisy than the KL reconstruction due to the smoothing effect. However, SIRT strongly blurs the small structures.

Second, we performed EDS reconstructions with TV-regularization with the KL or the L2 data discrepancy (KL-TV/L2-TV). The reconstructions were made for different values of regularization parameter λ , for which the correlation coefficients were computed and plotted in Fig. 5(a). Fig. 4(c) and (d) respectively show the optimal KL-TV or L2-TV reconstructions that correspond to the largest correlation coefficients. Compared to the non-regularized reconstructions, these reconstructions are more homogeneous with sharper edges. We observe that small structures have also been smoothed by the TV regularization. Also, the TV regularization introduces obvious staircasing artifacts.

Third, we introduce HEBT to reduce the staircasing artifacts. We performed the recipe of KL-TV-HEBT and L2-TV-HEBT for a range of regularization parameter λ and HEBT weight α . The corresponding correlation coefficients are plotted in Fig. 5(b) and (c) respectively. Fig. 4(e) and (f) respectively show the optimal reconstructions. We see that the KL-TV-HEBT reconstruction shows clearly reconstructed features at smaller scales (i.e., the ‘holes’) and less staircase artifacts compared to the L2-TV-HEBT reconstruction, due to the proper assumption of data discrepancy. Fig. 6(a) and (b) show the optimal KL-TV-HEBT reconstructions for a smaller and a larger α respectively. For the smaller α the reconstruction is similar to the KL-TV reconstruction as the HAADF-STEM data discrepancy is not given with a substantial weight, while for the larger α the reconstruction is more noisy.

Finally, we replaced the TV regularization by TNV regularization to promote the common edges of different elements. Fig. 4(g) and (h) are the optimal KL-TNV-HEBT and L2-TNV-HEBT reconstructions. Fig. 5 shows the zoom-in images for the regions of interest (ROI). Compared with the TV-regularized reconstruction, the TNV-regularized reconstruction is more accurate for areas near the common edge locations. The improvement of accuracy is also indicated by the correlation coefficients (see Fig. 5 (b) and (d)).

For these regularized HEBT reconstructions (Fig. 4(e)–(h)), the KL data discrepancy leads to more homogeneous gray values. However, if the SNRs are high enough, the Gaussian distribution assumed by the L2 data discrepancy can also form a close approximation even though the

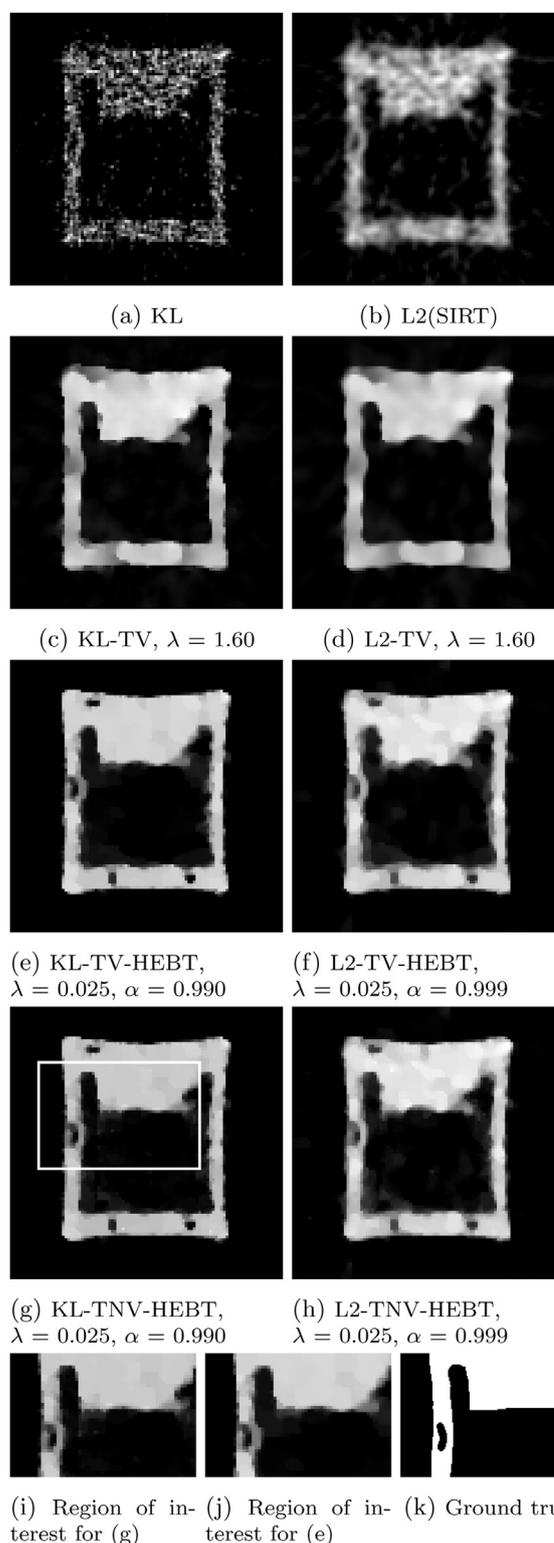
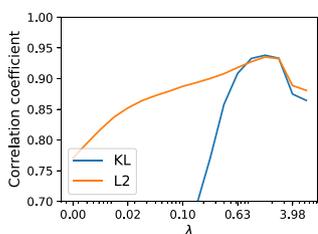


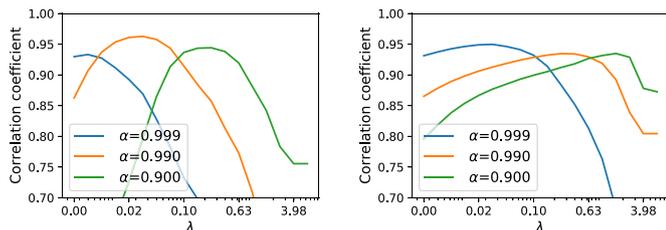
Fig. 4. Reconstructions for Ti in the non-mixed phantom using various recipes.

noise is Poisson distributed. For instance, the L2-TNV-HEBT reconstruction in Fig. 7 demonstrates little staircasing artifacts when the image intensities are increased by 400% (and therefore the SNRs by 200% for Poisson noise).

In addition, the TV regularization has been shown to reduce the artifacts introduced by the missing wedge [13], which is a common issue in electron tomography. Fig. 8(a) shows the TV-KL reconstruction

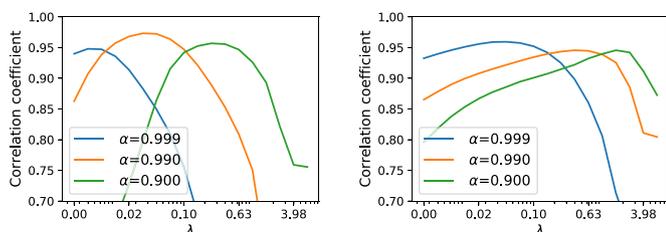


(a) KL-TV/L2-TV



(b) KL-TV-HEBT

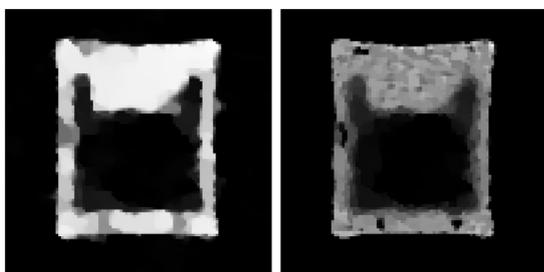
(c) L2-TV-HEBT



(d) KL-TNV-HEBT

(e) L2-TNV-HEBT

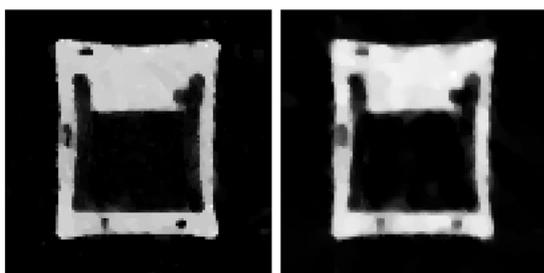
Fig. 5. Correlation coefficients sampled for different values of HEBT weight α and regularization parameter λ .



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

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

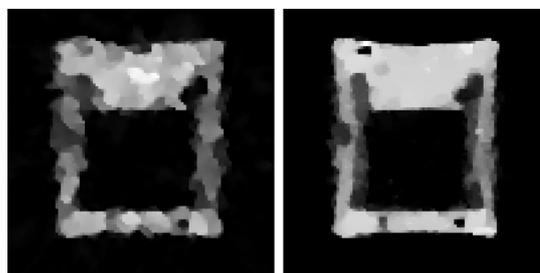
Fig. 6. KL-TV-HEBT reconstructions for different α values optimized w.r.t. λ values.



(a) KL-TNV-HEBT, $\lambda = 0.025, \alpha = 0.990$

(b) L2-TNV-HEBT, $\lambda = 0.063, \alpha = 0.999$

Fig. 7. HEBT-KL-TNV and HEBT-L2-TNV reconstructions for data with 200% SNRs.



(a) TV-KL $\lambda = 1.0$

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

Fig. 8. Reconstructions from data with the missing wedge.

for data with an angular range from -80° to 80° . In comparison, the KL-TNV-HEBT reconstruction (Fig. 8(b)) shows more clear structures in the horizontal direction. Therefore, an algorithmic recipe combining proper ingredients might also better reduce missing wedge artifacts.

3.2. Mixed phantom simulation

3.2.1. Data simulation

The purpose of this simulation is to study the reconstruction methods on inhomogeneous structures like alloyed materials, as opposed to the homogeneous structures used in the first simulation. The phantom was created resembling the nano-rattle sample investigated in [29]. The alloyed nanoparticle consists of Au and Ag components, which have inhomogeneous concentrations. Fig. 9(a) and (b) show the Au and Ag phantoms respectively. We created the HAADF phantom as the weighted sum of these two phantoms for $z_{Ag} = 47^{1.7}$ and $z_{Au} = 79^{1.7}$, which is shown in Fig. 9(c). Unlike the non-mixed phantom, the structures of Au and Ag components can be hardly distinguished in this image. We simulated the tilt series of 1D EDS maps and HAADF projection data for every 5° from 0° to 180° and added the Poisson noise to the EDS maps following the same procedures as in the first simulation.

3.2.2. Reconstruction results

TNV regularization is not applicable in this case since the reconstructions for Au and Ag do not necessarily share the same edge locations. We apply TV regularization for noise reduction. Therefore, we consider a recipe of KL-TV-HEBT that satisfies the conditions in Fig. 2.

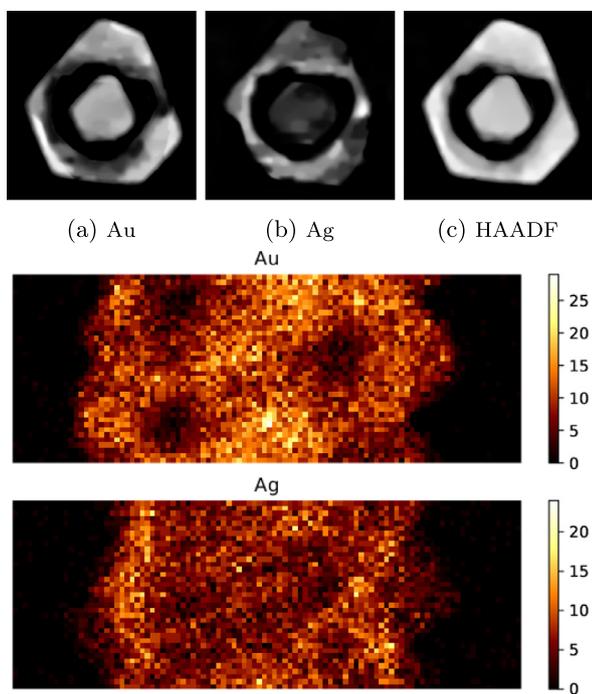
For comparison, we first performed SIRT reconstructions (for 50 iterations). The results are shown in Fig. 10(a) and (b), which demonstrate low SNRs. Also, we performed KL-TV reconstructions. The reconstructions corresponding to maximal correlation coefficients are shown in Fig. 10(c) and (d), which show significant staircasing artifacts due to the strong noise.

The optimal KL-TV-HEBT reconstructions are shown in Fig. 10(e) and (f). As a result, the combination of HEBT and TV effectively improves the quality of reconstruction. In particular, HEBT reduces the staircasing artifacts and results in more interpretable reconstructed images. The improvement of image quality is verified by the correlation coefficients in Fig. 10(g).

3.3. Real experimental data

3.3.1. Data acquisition

We now investigate the proposed method on a real experimental dataset. The sample is a core-shell nanoparticle of an Au cube embedded in an Ag particle, which has been investigated in our previous paper [18]. The two components have clear boundaries, homogeneous densities and different Z-contrasts. Thus, the core-shell nanoparticle is suitable for applying a TNV regularization.



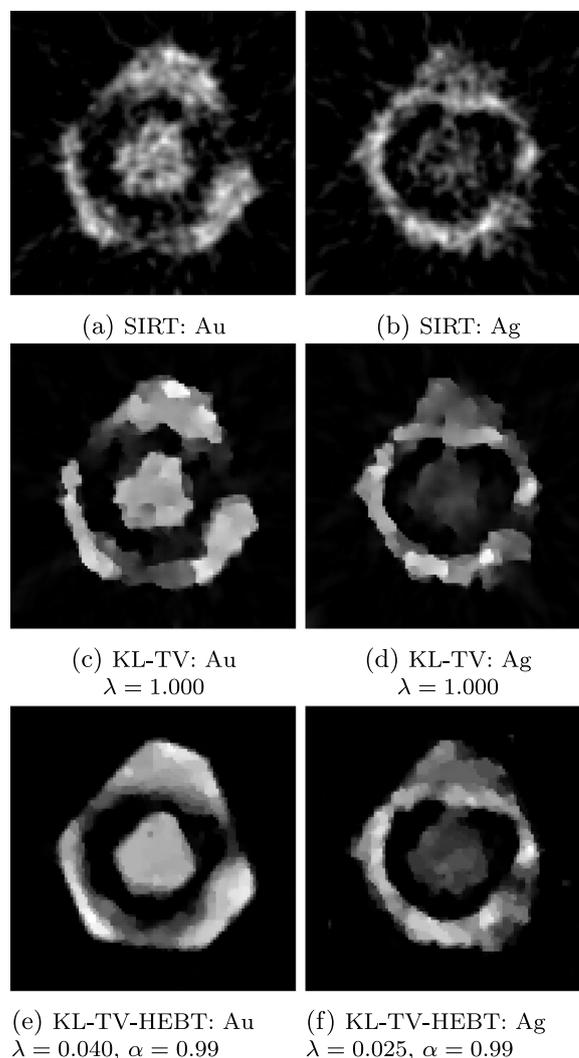
(d) EDS maps for Au and Ag. The row and column correspond to the tilt angle and the beam position respectively. The mean image intensities on non-background pixels are respectively 8.77 and 5.71.

Fig. 9. The mixed phantom simulation data.

The experimental data, which consist of a tilt series of spectrum images and correlated HAADF-STEM projection images, were acquired using an electron microscope equipped with four silicon drift detectors. The specifications of the EDS data acquisition are listed in Table 1. During the tilt acquisition, only the X-ray detectors on one side were turned on so that the detector shadowing effects were compensated. However, this approach also limited the number of X-ray counts that could be acquired. After PCA denoising, elemental maps were extracted by integrating the spectrum images near the characteristic peaks (Au: $M_\alpha = 2.15$ keV, $M_\beta = 2.20$ keV and $L_\alpha = 9.70$ keV; Ag: $L_\alpha = 2.98$ keV and $L_\beta = 3.19$ keV) as described in [18]. The HAADF-STEM tilt series were aligned using the cross-correlation method. The EDS elemental maps were then aligned using the same alignment settings. The intensity damping in the HAADF-STEM data was corrected using the correction algorithm [30]. Finally, all the images were binned to 100×100 pixels so as to increase the SNRs to reasonable levels. Fig. 11 shows two examples of the elemental maps.

Au and Ag have distinct Z-contrasts in the HAADF-STEM images. Therefore, ground truth for evaluating the EDS reconstructions can be obtained by segmenting the HAADF-STEM reconstruction into Au and Ag components. Fig. 12(a) shows a slice of the 3D reconstruction for HAADF-STEM, which was made with TV regularization to promote piecewise constant structures and to facilitate the subsequent segmentation. Fig. 11(b) and (c) show the subsequent segmented images for Au and Ag respectively.

Fig. 12(d) and (e) show the SIRT reconstructions for Au and Ag from the EDS maps, which are indeed noisy and inaccurate. We hope to use a



(g) Correlation coefficients of the reconstructions

Fig. 10. Reconstructions for Au and Ag in the mixed phantom using various recipes.

tailored recipe to make more accurate reconstructions. Given the low X-ray counts, the sample structure and the correlated HAADF-STEM data, we apply a KL-TNV-HEBT recipe.

Table 1
Data acquisition specifications.

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

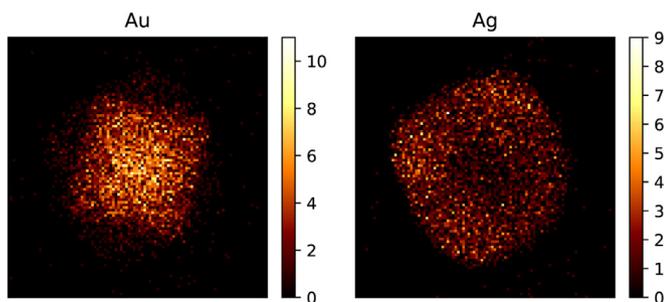


Fig. 11. Au and Ag elemental maps at 0°.

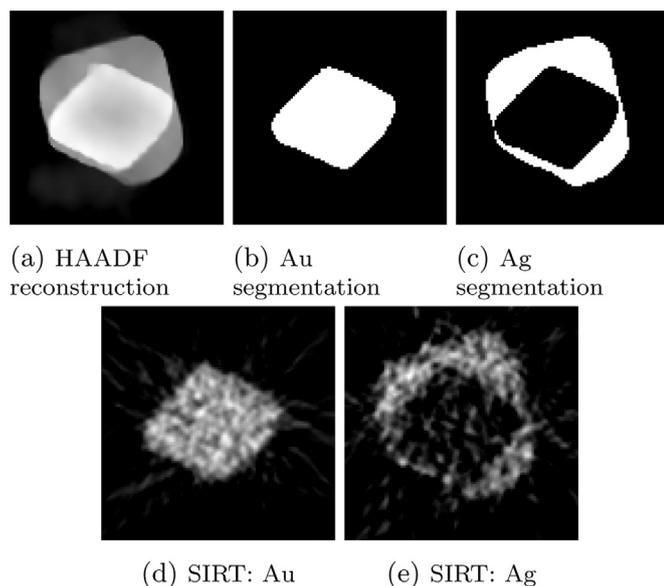


Fig. 12. Reconstructions for slice number 50. The HAADF reconstruction was performed in 3D with TV regularization using the DR algorithm.

3.3.2. Results and discussion

We first searched for the optimal α and λ parameters for the KL-TNV-HEBT recipe. Since it is time-consuming to compute 3D reconstructions, we selected a 2D slice to sample reconstructions. Here we used the slice at the center of the sample (number 50), which should give a good estimation for the SNRs of the entire volume. Fig. 13(a) and (b) show the correlation coefficients with the segmented HAADF-STEM reconstruction, computed for the 2D reconstructions at slice 50. The correlation coefficients for Ag reach maximum at $\alpha = 0.9900$ and $\lambda = 0.10$, at which the correlation coefficient for Au is also close to maximal.

We then applied the λ and α to the entire volume. We performed the reconstruction for the entire volume with regularization in 3D. Fig. 14(a)–(f) show some slices of the 3D reconstruction. For comparison, we also performed 2D regularized reconstruction for each slice,

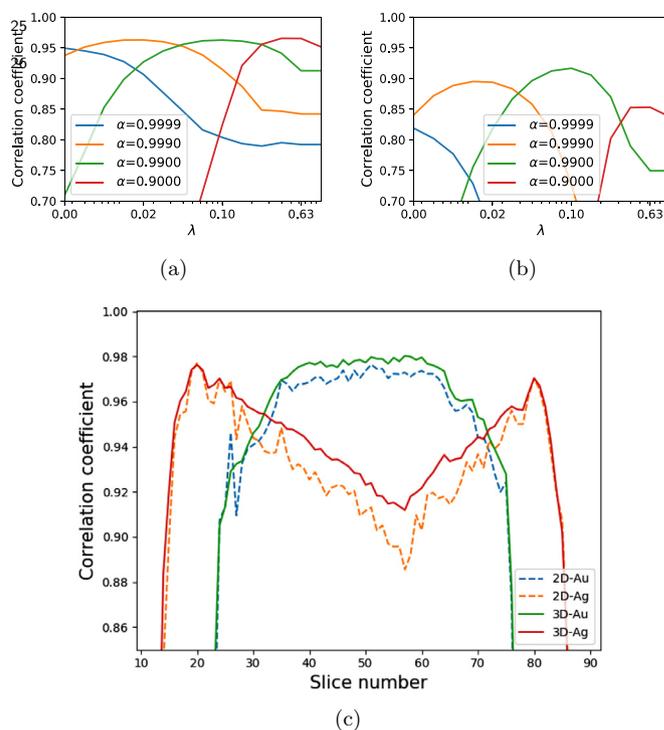


Fig. 13. (a) and (b): correlation coefficients for the KL-TNV-HEBT reconstructions for slice number 50, sampled for different λ and α values. (c): correlation coefficients for all slices of 2D or 3D reconstructions.

some of which are shown in Fig. 14(g)–(l). Fig. 14(m)–(r) show the ground truth for evaluating these reconstructions, which were obtained by segmenting the TV-regularized HAADF reconstruction. The 3D reconstructions are smoother and more accurate compared to the 2D reconstructions, since the large variation in the direction of rotation axis was penalized. Fig. 13(c) compares the correlation coefficients for 3D and 2D reconstructions for every slice in the volume. Once again, we conclude that 3D reconstructions are to be preferred when regularizations are applied.

4. Conclusion

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

Different algorithmic recipes have been used to reconstruct from both simulation and real experimental data. We evaluated the accuracy of reconstructions based on the correlation coefficients w.r.t. ground truth. For all these experiments, the algorithms lead to more accurate reconstruction compared to more naive algorithms when they are tailored for the dataset and sample.

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

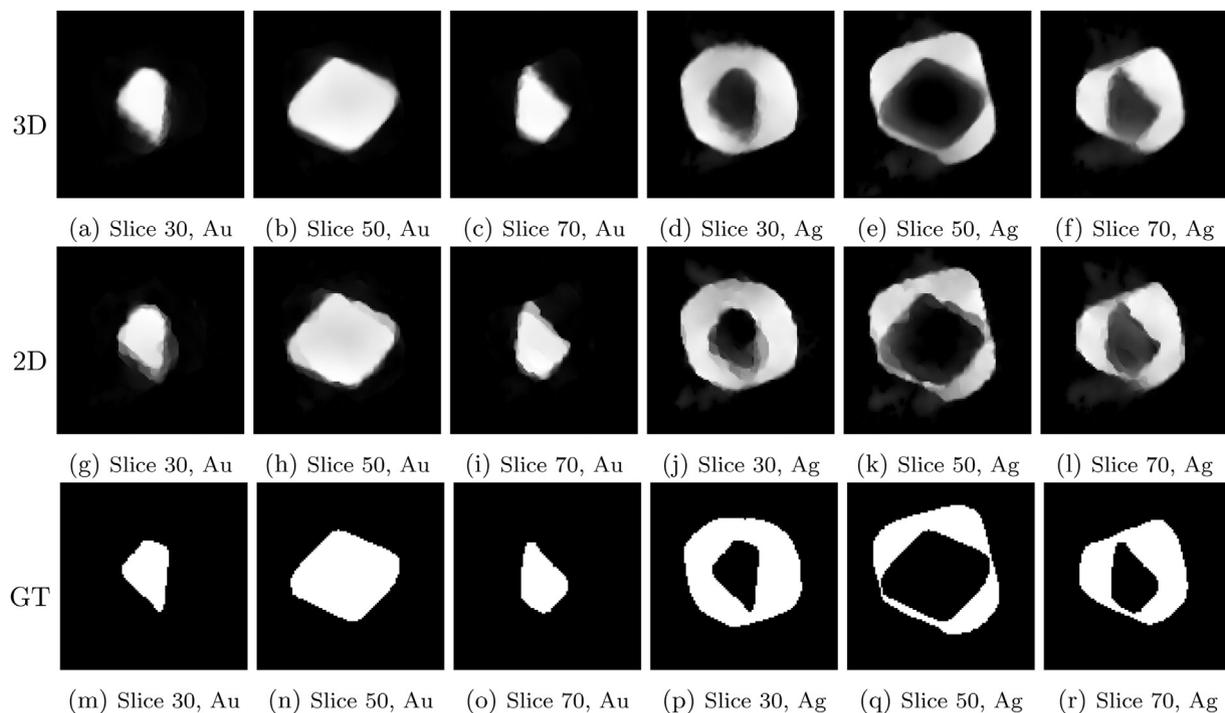


Fig. 14. KL-TNV-HEBT reconstructions. (a)–(f): slices of 3D reconstructions; (g)–(l): 2D reconstructions; (m)–(r): ground truth (GT) for evaluating the reconstructions.

Acknowledgement

This work is part of the research programme “Automated multimodal tomography for sub-22 nm IC nodes” with project number 13314 which is partly financed by the Netherlands Organisation for Scientific

Research (NWO). NV acknowledges financial support from the European Union’s Horizon 2020 research and innovation programme (VOXEL H2020-FETOPEN-2014-2015-RIA GA 665207). The authors would like to thank Dr. Bart Goris and Prof. Sara Bals for providing the investigated experimental data.

Appendix

In this paper we use the Douglas-Rachford primal-dual splitting algorithm to solve the optimization problem as a sum of multiple objective functions, which is a broadly applicable algorithm for solving the following convex optimization problem from Bot and Hendrich [24]:

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

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

$$g_k \square l_k(v) = \inf_y g_k(y) + l(v - y). \quad (15)$$

By setting $z = 0$ and

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

we simplify the mathematical problem to Eq. (12).

To construct a optimization problem based on Eq. (12) given a reconstruction recipe, we set $f(\cdot) = 0$, and map $g_k(\cdot)$ ’s and \mathbf{A}_k ’s to our functions and operators. For example, for a KL-TNV-HEBT recipe, the optimization problem can be made from:

$$f(\{\mathbf{x}^e\}) = 0,$$

$$g_1(\{\mathbf{v}^e\}) = (1 - \alpha) \sum_{e=1}^L D_{\text{KL}}(\mathbf{v}^e; \mathbf{p}^e),$$

$$\mathbf{A}_1 = \{\mathbf{W}^e\},$$

$$g_2(\{\mathbf{v}^e\}) = \alpha \left\| \sum_{e=1}^L \mathbf{v}^e - \mathbf{p}^e \right\|_2^2,$$

$$\mathbf{A}_2 = \{\mathbf{W}^h\},$$

$$g_3(\{\mathbf{Z}_j\}) = \lambda \sum_{j=1}^N \left\| \mathbf{Z}_j \right\|_{\star},$$

$$\mathbf{A}_3 = \mathbf{J} \otimes \mathbf{I}_N.$$

Table 2
Mathematical instances for the ingredients.

Ingredient	Minimization term	DR algorithm instances
L2	$\mathcal{D}_e(\mathbf{x}^e) = \ \mathbf{W}^e \mathbf{x}^e - \mathbf{p}^e\ _2^2$	$g(\{\mathbf{v}^e\}) = (1 - \alpha) \sum_{e=1}^L \ \mathbf{v}^e - \mathbf{p}^e\ _2^2, \mathbf{A} = \{\mathbf{W}^e\}$
KL	$\mathcal{D}_e(\mathbf{x}^e) = \mathcal{D}_{\text{KL}}(\mathbf{W}^e \mathbf{x}^e; \mathbf{p}^e)$	$g(\{\mathbf{v}^e\}) = (1 - \alpha) \sum_{e=1}^L \mathcal{D}_{\text{KL}}(\mathbf{v}^e; \mathbf{p}^e), \mathbf{A} = \{\mathbf{W}^e\}$
HEBT	$\mathcal{D}_h(\{\mathbf{x}^e\}) = \ \sum_{e=1}^L \mathbf{W}^{hr^e} \mathbf{x}^e - \mathbf{p}^h\ _2^2$	$g(\{\mathbf{v}^e\}) = \alpha \ \sum_{e=1}^L r^e \mathbf{v}^e - \mathbf{p}^h\ _2^2, \mathbf{A} = \{\mathbf{W}^h\}$
None	$\mathcal{D}_h(\{\mathbf{x}^e\}) = 0$	
TV	$\mathcal{R}(\{\mathbf{x}^e\}) = \sum_{e=1}^L \sum_{j=1}^N \left\ \nabla_{x_j^e} \right\ _2$	$g(\{\mathbf{Y}^e\}) = \lambda \sum_{e=1}^L \sum_{j=1}^N \left\ \mathbf{y}_j^e \right\ _2, \mathbf{A} = \{\nabla\}$
TNV	$\mathcal{R}(\{\mathbf{x}^e\}) = \sum_{j=1}^N \left\ \left(\mathbf{J}(\mathbf{x}^e) \right)_j \right\ _{\star}$	$g(\{\mathbf{Z}_j\}) = \lambda \sum_{j=1}^N \left\ \mathbf{Z}_j \right\ _{\star}, \mathbf{A} = \mathbf{J} \otimes \mathbf{I}_N$
None	$\mathcal{R}(\{\mathbf{x}^e\}) = 0$	

The matrix \mathbf{I}_N is the $N \times N$ identity matrix, and \otimes denotes the Kronecker product. Table 2 lists the instances for all functions and linear operators used in this paper. Another key to deriving the particular DR algorithm instances is to derive $\text{prox}_{g_k^*}[g_k^*](y)$, which is the proximal operator for the convex conjugate of $g_k(\cdot)$. The exact forms of the proximal operators for the functions in Table 2 are derived and provided in [16,31,32].

References

- Z. Saghi, X. Xu, Y. Peng, B. Inkson, G. Möbus, Three-dimensional chemical analysis of tungsten probes by energy dispersive X-ray nanotomography, *Appl. Phys. Lett.* 91 (25) (2007) 251906.
- K. Lepinay, F. Lorut, R. Pantel, T. Epicier, Chemical 3D tomography of 28 nm high k metal gate transistor: STEM XEDS experimental method and results, *Micron* 47 (2013) 43–49.
- D.B. Williams, C.B. Carter, *Transmission Electron Microscopy: Diffraction, Imaging, and Spectrometry*, Springer, 2016.
- T.J. Slater, A. Janssen, P.H. Camargo, M.G. Burke, N.J. Zaluzec, S.J. Haigh, STEM-EDX Tomography of bimetallic nanoparticles: a methodological investigation, *Ultramicroscopy* 162 (2016) 61–73.
- J. Kraxner, M. Schöfer, O. Röschel, G. Kothleitner, G. Haberfehlner, M. Paller, W. Grogger, Quantitative EDXS: influence of geometry on a four detector system, *Ultramicroscopy* 172 (2017) 30–39.
- P. Burdet, Z. Saghi, A.N. Filippin, A. Borrás, P.A. Midgley, A novel 3D absorption correction method for quantitative EDX-STEM tomography, *Ultramicroscopy* 160 (2016) 118–129.
- J.M. Bardsley, A. Luttmann, Total variation-penalized Poisson likelihood estimation for ill-posed problems, *Adv. Comput. Math.* 31 (1) (2008) 35.
- J. Gregor, T. Benson, Computational analysis and improvement of SIRT, *IEEE Trans. Med. Imaging* 27 (7) (2008) 918–924.
- L. Shepp, Y. Vardi, Maximum likelihood reconstruction for emission tomography, *IEEE Trans. Med. Imaging* 1 (2) (1982) 113–122.
- T. Hohage, F. Werner, Inverse problems with Poisson data: statistical regularization theory, applications and algorithms, *Inverse Probl.* 32 (9) (2016) 093001.
- D.F. Yu, J.A. Fessler, Edge-preserving tomographic reconstruction with nonlocal regularization, *IEEE Trans. Med. Imaging* 21 (2) (2002) 159–173.
- J.M. Bardsley, A theoretical framework for the regularization of poisson likelihood estimation problems, *Inverse Probl.* 4 (1) (2010) 11–17.
- B. Goris, W. Van den Broek, K.J. Batenburg, H. Heidari Mezerji, S. Bals, Electron tomography based on a total variation minimization reconstruction technique, *Ultramicroscopy* 113 (2012) 120–130.
- M. Burger, S. Osher, *A guide to the TV zoo*, Springer International Publishing, Cham, pp. 1–70.
- K.M. Holt, Total nuclear variation and Jacobian extensions of total variation for vector fields, *IEEE Trans. Image Process.* 23 (9) (2014) 3975–3989.
- J. Duran, M. Moeller, C. Sbert, D. Cremers, Collaborative total variation: a general framework for vectorial TV models, *SIAM J. Imaging Sci.* 9 (1) (2016) 116–151.
- K. Bredies, K. Kunisch, T. Pock, Total generalized variation, *SIAM J. Imaging Sci.* 3 (3) (2010) 492–526.
- Z. Zhong, B. Goris, R. Schoenmakers, S. Bals, K.J. Batenburg, A bimodal tomographic reconstruction technique combining EDS-STEM and HAADF-STEM, *Ultramicroscopy* 174 (2017) 35–45.
- M. Watanabe, D.B. Williams, The quantitative analysis of thin specimens: a review of progress from the Cliff-Lorimer to the new ζ -factor methods, *J. Microsc.* 221 (2) (2006) 89–109.
- M. Bertero, P. Boccacci, G. Talenti, R. Zanella, L. Zanni, A discrepancy principle for Poisson data, *Inverse Probl.* 26 (10) (2010) 105004.
- I. Csizsar, Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems, *Annals Stat.* 19 (4) (1991) 2032–2066.
- Z. Zhong, W.J. Palenstijn, J. Adler, K.J. Batenburg, EDS Tomographic reconstruction regularized by total nuclear variation joined with HAADF-STEM tomography, *Ultramicroscopy* 191 (2018) 34–43.
- G. Haberfehlner, A. Orthacker, M. Albu, J. Li, G. Kothleitner, Nanoscale voxel spectroscopy by simultaneous EELS and EDS tomography, *Nanoscale* 6 (2014) 14563–14569.
- R.I. Bot, C. Hendrich, A Douglas–Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators, *SIAM J. Optim.* 23 (4) (2013) 2541–2565.
- J. Adler, H. Kohr, O. Öktem, *Operator Discretization Library*, 2017, <https://doi.org/10.5281/zenodo.556409>, <https://github.com/odlgroup/odl>.
- M.M. Treacy, Z dependence of electron scattering by single atoms into annular dark-field detectors, *Microsc. Microanal.* 17 (06) (2011) 847–858.
- W.J. Palenstijn, K.J. Batenburg, J. Sijbers, *The ASTRA tomography toolbox*, *Proceedings of the Thirteenth International Conference on Computational and Mathematical Methods in Science and Engineering* (2013) 1–7.
- Z. Chen, M. Weyland, X. Sang, W. Xu, J. Dycus, J. LeBeau, A. D'Alfonso, L. Allen, S. Findlay, Quantitative atomic resolution elemental mapping via absolute-scale energy dispersive X-ray spectroscopy, *Ultramicroscopy* 168 (2016) 7–16.
- D. Zanaga, T. Altantzis, L. Polavarapu, L.M. Liz-Marzán, B. Freitag, S. Bals, A new method for quantitative XEDS tomography of complex heterostructures, *Particle Particle Syst. Charact.* 33 (7) (2016) 396–403.
- Z. Zhong, R. Aveyard, B. Rieger, S. Bals, W.J. Palenstijn, K.J. Batenburg, Automatic correction of nonlinear damping effects in HAADF-STEM tomography for nano-materials of discrete compositions, *Ultramicroscopy* 184 (Part B) (2018) 57–65.
- D.S. Rigie, P.J.L. Rivière, Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization, *Phys. Med. Biol.* 60 (5) (2015) 1741.
- E.Y. Sidky, J.H. Jorgensen, X. Pan, Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm, *Phys. Med. Biol.* 57 (10) (2012) 91–3065.