

Denoising Atomic Resolution 4D Scanning Transmission Electron Microscopy Data with Tensor Singular Value Decomposition

Chenyu Zhang¹, Rungang Han², Anru R. Zhang², Paul. M. Voyles¹

¹ Department of Materials Science and Engineering, University of Wisconsin-Madison

² Department of Statistics, University of Wisconsin-Madison

Abstract

Tensor singular value decomposition (SVD) is a method to find a low-dimensional representation of data with meaningful structure in three or more dimensions. Tensor SVD has been applied to denoise atomic-resolution 4D scanning transmission electron microscopy (4D STEM) data. On data simulated from a SrTiO_3 [100] perfect crystal and a Si [110] edge dislocation, tensor SVD achieved an average peak signal-to-noise ratio (PSNR) of ~ 40 dB, which matches or exceeds the performance of other denoising methods, with processing times at least 100 times shorter. On experimental data from SrTiO_3 [100] and LiZnSb [$11\bar{2}0$]/GaSb [110] samples, tensor SVD denoises multiple GB 4D STEM data sets in ten minutes on a typical personal computer. Denoising with tensor SVD improves both convergent beam electron diffraction patterns and virtual-aperture annular dark field images.

Keywords: scanning transmission electron microscopy, convergent beam electron diffraction, image denoising, low-rank tensor

1. Introduction

Denoising 4D STEM¹ data is challenging because of its large size² and complex noise distribution. A 16-bit depth 4D STEM dataset with a 100-by-100 beam raster scan and 512-by-

512 px on each convergent beam electron diffraction (CBED) pattern is 5 GB. With recent developments in ultrafast detectors³⁻⁴, the data will be generated faster with a smaller signal and more noise in each CBED pattern and even larger data sets. To be applied to multi-GB 4D STEM data using moderate computing resources, a denoising algorithm has to be both time- and memory-efficient. Another challenge for denoising 4D STEM data is the potential for complex noise distributions. A perfect counting mode detector would acquire data with only Poisson noise^{5,6}, but errors from a real counting-mode detector may introduce more complex statistics. An active pixel sensor type detector operating in linear mode generates signal in response to single electrons that obey a Landau distribution^{7,8}. As a result, each electron generates a different intensity and the noise is again not pure Poisson noise.

A common approach to denoising complex data, including data from electron microscopy, is to find a low-rank representation of the noisy data. The data are viewed as vectors in a high-dimensional space. The true, noiseless data occupy only a restricted subspace with lower dimensionality. Noise shifts the points out of the subspace. If we can find appropriate basis vectors that span the noiseless subspace, then use them to project the data into the subspace, the noise is removed. Those basis vectors can also be used to gain understanding of the data by, for example, spectral unmixing (see ⁹ for additional discussion and *e.g.* ¹⁰ for rigorous mathematics). The most widely-applied rank reduction tool for electron microscopy data is principal component analysis (PCA)^{10,11}. PCA looks for a set of orthogonal basis vectors that can describe the original data without losing significant information. PCA has been modified to account for Poisson noise statistics specifically to create Poisson PCA or exponential PCA¹². PCA is a powerful tool for denoising electron energy loss spectroscopy (EELS)^{13,14,15} and energy-dispersive X-ray spectroscopy (EDS)^{16,17} data collected in STEM, and it has been implemented in a variety of

software packages including a DigitalMicrograph plugin¹⁸, temDM¹⁹, and Hyperspy²⁰. Various other component analysis methods exist which differ in how they define the optimal basis vectors. For example, independent component analysis (ICA)²¹ seeks to split the data into a set of components, but unlike PCA, ICA does not require components to be orthogonal to each other. ICA has also been used for denoising²², which requires signal and noise to be independent and noise variance cannot be dependent on the signal. Another example is non-negative matrix factorization (NMF)²³, which can only be applied to non-negative matrices and imposes a non-negativity constraint on both the basis vectors and the corresponding coefficients, with orthogonality as an option for the basis vectors²⁴. NMF has been applied to both hyperspectral data denoising^{25,26,27} and unmixing²⁸. The non-negative constraint can be desirable when searching for hidden basis vectors with physical meanings that require them to be non-negative. PCA and its cousins work only on 2D matrices, so 3D data like spectrum images must be unwrapped into 2D before processing, losing some of their structure, usually in one of the two spatial dimensions.

Outside electron microscopy, the many current state-of-the-art denoising methods are non-local²⁹. These approaches find self-similarity within the data, even if the similar regions are separated in one or more dimensions, and combine information from similar regions to remove random noise. As a first step, the image is divided into small regions called patches, which often overlap with one another. Common non-local methods for general images include non-local means³⁰ and the block-matching and 3D filtering³¹. In electron microscopy, Mevenkamp et al. applied non-local means³² and adapted block-matching and 3D-filtering (BM3D)³³ to denoise single frame atomic-resolution STEM image. Both methods assumed a 2D periodic structure in the image to help locate similar patches, then noiseless ground truth for each group of patches

are determined by maximizing the Poisson maximum-likelihood³⁴ of finding the noisy patches with given truth, or applying a Wiener filter respectively. Yankovich et al. applied non-local principle component analysis (NLPCA) to denoise STEM EDS spectrum image data^{35,36}, NLPCA utilizes k-means clustering³⁷ to group similar 3D (x , y , and energy) patches, then Poisson PCA to denoise each group. Maggioni et al. proposed the block-matching and 4D-filtering (BM4D) method³⁸ as an extension of BM3D to denoising 3D data in the form of movies (x , y , and time). BM4D searches for similar 3D patches within a local search window which is usually smaller than the total data size using the photometric distance, then sets a distance threshold to group the 3D patches. A Wiener filter was applied within groups of similar 3D patches to remove image noise. For all of these methods, the initial patchifying of the data requires substantial memory. The set of all possible overlapping 10x10 pixel patches in a 100x100 pixel image occupies 81 times the memory of the original image. The periodic and local window search methods are, in part, ways to avoid this maximum patch-ifying while still finding similar patches. For larger data sets like 4D STEM, the memory and / or processing time can easily exceed typical computing resources. Besides memory consumption, the non-local method could possibly result in a loss of detail or introduce artifacts when denoising the group of similar patches.

Tensor singular value decomposition (SVD) is a low-rank denoising method for high-order data in three or more dimensions.^{39–43} Here we apply the tensor SVD method of Zhang and Xia⁴¹ to denoise atomic-resolution STEM data, focusing on 4D STEM but also testing EDS spectrum images. Unlike matrix rank-reduction methods like PCA, tensor SVD can maintain the full structure of data, giving it the potential to exploit, for example, structure in both spatial dimensions. We compare tensor SVD to matrix PCA and to the state-of-the-art non-local

methods NLPCA and BM4D. Overall, the denoising performance is as good or better than all the other methods, and the computation time and memory requirements are much more similar to PCA than to NLPCA or BM4D. Single-digit GB 4D STEM data sets can be denoised in a few minutes on a typical desktop computer.

2. Methods

2.1 4D STEM Experiment

A SrTiO_3 (STO) [100] single crystal sample was prepared by wedge polishing at a 1.6° angle using an Allied MultiPrep polishing system with diamond lapping films, followed by final thinning in a Fishione 1050 ion-mill using Ar ion beam. A LiZnSb $[11\bar{2}0]$ thin film sample was grown on a GaSb [110] substrate as described in Ref. ⁴⁴, then prepared for TEM using a Zeiss Auriga focused ion beam (FIB) with a final FIB polishing step with a 5 kV 100 pA Ga-ion beam. The sample surface was polished in a Fishione 1040 Nanomill with a 900 eV Ar ion beam. Both samples were stored under vacuum and cleaned in an Ibss GV10x DS Asher plasma cleaner operated under 20 W for 10 minutes to remove contaminations before being inserted into the TEM column.

A Thermo-Fisher Scientific Titan STEM equipped with a CEOS probe corrector operated at 200 kV was used to collect 4D STEM data. A 24.5 mrad semi convergence angle and an 18.9 pA current probe was used, which was optimized for the highest spatial resolution. A Direct Electron DE-16 camera operating in linear mode (as opposed to counting mode) was used to collect CBED patterns with camera acquisition synchronized to the electron beam scan using a customized scan controller. 60 mm camera length was used for CBED pattern acquisition to achieve about 37 mrad maximum collection angle in CBED patterns, collecting the whole bright

field disk and low angle dark-field signal. Details of parameter selection and scan controller to synchronize camera acquisition and beam scan can be found in Ref⁴⁵. For each 4D STEM acquisition, 150×150 probe positions were used with 512×512 pixel CBED patterns acquired under 0.9 ms exposure time at each probe position. The scan step sizes were 0.04 nm on the LiZnSb sample and 0.02 nm on the STO sample, small enough to resolve the lattice structures on both samples.

2.2 Multislice Simulation

Two structures were used to simulate 4D STEM data: a STO single crystal along the [100] zone axis, and a Si [110] edge dislocation core generated by molecular dynamic simulations with distorted structure around the dislocation core⁴⁶. CBED patterns from both structures were simulated using a frozen-phonon multislice algorithm in the graphic processing units (GPU)-accelerated Prismatic simulation package⁴⁷. Probe convergence angle was selected to match the experimental conditions. A 2D symmetric Gaussian function with 110 pm full width half maximum was convolved into the simulated data to account for the incoherent source size^{48,49}. All computations were performed on a computing cluster equipped with quadcore Nvidia Tesla M10 GPUs. A full list of simulation parameters for the two models can be found in table 1.

Parameter name	SrTiO ₃ [100] single crystal	Si [110] edge dislocation
Model size	49 Å × 49 Å × 250 Å	100 Å × 100 Å × 250 Å
Probe step size	0.175 Å	
Potential space sampling	0.06 Å	
Slice thickness	1.9525 Å	1.90 Å

Frozen phonon configurations	20	
Root mean square thermal displacements ^{50,46}	Sr: 0.0887 Å Ti: 0.0746 Å O: 0.0947 Å	Si: 0.076Å

Table 1 Multislice simulation parameters

Simulated CBEDs were cropped and resampled to mimic the reciprocal space pixel size and detection limit in the experiments and converted from intensity units of the fraction of the incident beam to the number of detected electrons using the frame rate and measured beam current from our Titan STEM. Different frame rates were used to generate simulated data with different noise levels. Random, Poisson-distributed noise was introduced into simulated CBED patterns using the number of electrons in each pixel as the mean value for Poisson distributed random numbers. Simulated scan distortion or other artifacts were not added to the data.

2.3 Tensor Singular Value Decomposition

The concept of the tensor SVD approach^{41,42} is illustrated in Figure 1 using simulated 4D STEM data on the STO structure. Figure 1 (a) shows the noisy observation tensor $\mathbf{Y}$ with the size $p_1 \times p_2 \times p_3$. The observation tensor is the sum of a ground truth tensor $\mathbf{X}$ (Figure 1(b)) and an additive noise tensor $\mathbf{Z}$ (Figure 1(c)), such that

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}; \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathbb{R}^{p_1 \times p_2 \times p_3}. \quad [1]$$

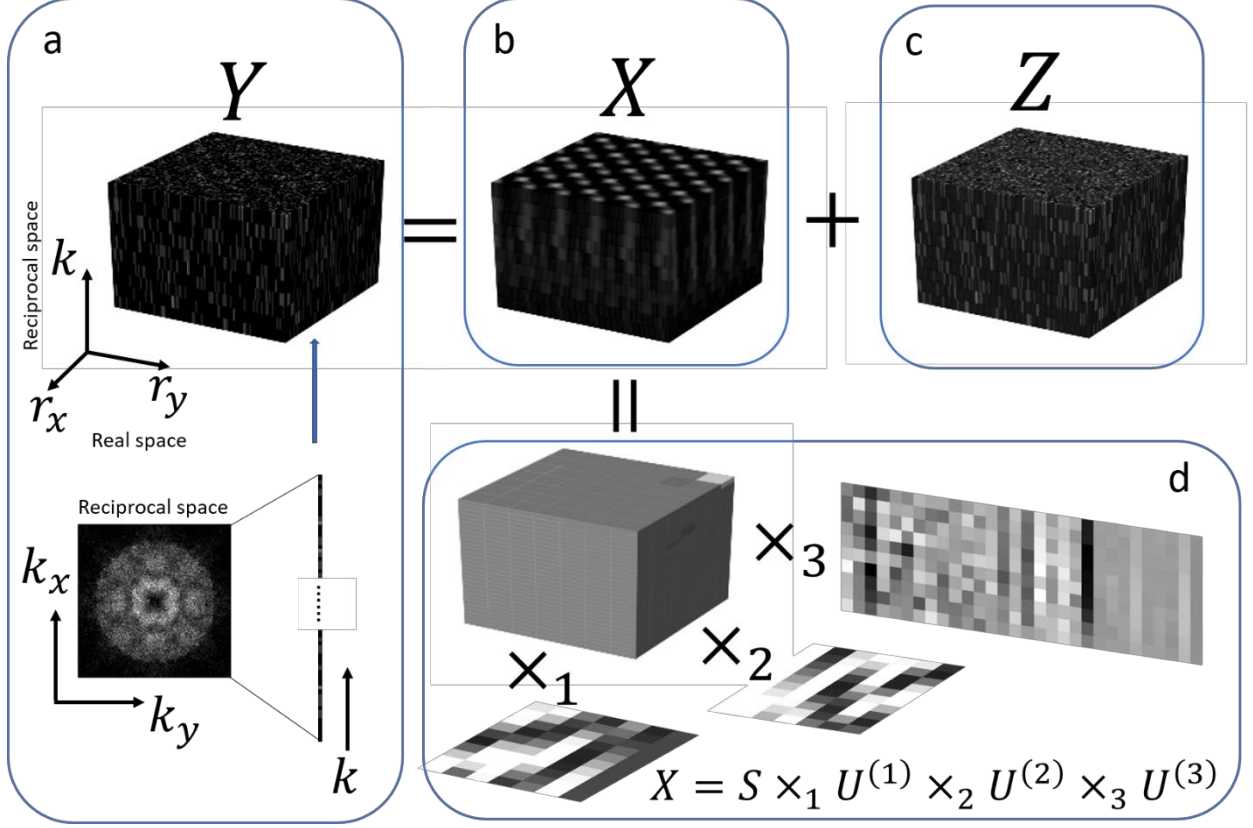


Figure 1 Illustration of the tensor SVD problem. (a) noisy observation tensor Y constructed from CBED patterns by unfolding the reciprocal space into one single dimension, (b) noiseless ground truth tensor X with low rank, (c) additive noise tensor Z , (d) Tucker decomposition of the low-rank tensor into a core tensor and three unitary matrices.

Additive noise is not a good model for Poisson noise at low mean counts, but the data could be transformed using the Anscombe variance-stabilizing transform to create pure additive noise,⁵¹ then un-transformed after denoising. The transform was not applied here. X is assumed to have Tucker low-rank (r_1, r_2, r_3) . In other words, it can be decomposed into a product of core tensor S with size $r_1 \times r_2 \times r_3$, and three unitary matrices $U^{(1)}$, $U^{(2)}$, and $U^{(3)}$, following Tucker decomposition⁵²,

$$X = S \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)}, S \in \mathbb{R}^{r_1 \times r_2 \times r_3}, U^{(k)} \in \mathbb{O}^{p_k \times r_k}, \quad [2]$$

as shown in Figure 1(d). Here, the tensor-matrix product is defined as:

$$\mathbf{X}_{ijk} = \sum_{a,b,c} \mathbf{S}_{abc} U_{ia}^{(1)} U_{ib}^{(2)} U_{ic}^{(3)} \quad [3].$$

The problem of finding the lower-rank decomposition shown on the right side of eq [2] from the noisy observation $\mathbf{Y}$ is what we call a tensor SVD problem.

For three-dimensional spectrum image data with dimensions x , y , and energy, there is an analogy between tensor SVD and the output of PCA. The $U^{(3)}$ matrix in eq [2] can be interpreted as the PCA loading matrix, which contains a total of p_3 spectra each with r_3 elements, and $\mathbf{S} \times_1 U^{(1)} \times_2 U^{(2)}$ gives the 3D tensor equivalent of the score matrix. In this “score tensor”, a total of p_3 spatial maps of size $p_1 \times p_2$ are created, showing the amplitude of the corresponding loading spectrum along the first two dimensions. Thus, in addition to denoising, tensor SVD can also be applied to decompose high dimensional data and find characteristic signals, such as different sets of diffraction patterns from different structures or with different rotational symmetry^{53,54}. This is not possible for non-local-method-based tools such as BM4D³⁸ and NLPCA³⁶.

For tensor SVD problems, unlike their counterparts in matrices, the best rank- (r_1, r_2, r_3) approximation cannot be obtained from truncations of a full decomposition⁵⁵, and ranks need to be determined first before solving the tensor SVD problem. We have set the rank along each dimension to the number of principal components determined by scree tests on 2D matrices formed by unfolding 3D tensor along that dimension⁴². For example, when unfolding the 3D tensor along the first dimension, the elements in the resulting 2D matrix with size $p_1 \times (p_2 p_3)$ are arranged as

$$[\mathcal{M}_1(\mathbf{Y})]_{i,(j-1)p_3+k} = \mathbf{Y}_{ijk}, \mathbf{Y} \in \mathbb{R}^{p_1 \times p_2 \times p_3}, \mathcal{M}_1(\mathbf{Y}) \in \mathbb{R}^{p_1 \times p_2 p_3}. \quad [4]$$

Eigenvalues are then calculated on the product of this 2D matrix with its transpose to generate a scree plot. Following previous work applying PCA on hyperspectral data^{15,56,57,17}, the number of principal components is determined by finding the point where the curve becomes linear on a log-linear plot. This method for rank determination is justified below in section 2.4 using simulated data. The computation time to generate the scree plots is about 1 minute for 4D STEM data with $150 \times 150 \times 16384$ data size.

To solve the tensor SVD problem given a rank (r_1, r_2, r_3) , we applied the higher-order orthogonal iteration (HOOI) method, first proposed to find low-rank tensor in ref⁵⁵. The HOOI method iteratively refines the unitary matrix U along each dimension by performing SVD on a related 2D matrix, then calculating the core tensor S and noiseless tensor X in the end with refined U matrices. The HOOI method requires that the product of any two ranks must be larger than the third rank.

In our implementation,^{41,42} the HOOI method seeks the best rank- (r_1, r_2, r_3) approximation $\hat{X}$ for noisy input Y by minimizing the difference between the tensors X and Y

$$\hat{X} = \underset{rank(X)=r_1, r_2, r_3}{\operatorname{argmin}} \|Y - X\|_F, \quad [5]$$

where the difference between the two tensors is measured by the Frobenius norm, $\|\cdot\|_F$. $\hat{X}$ is then the estimate of the denoised data. Additional details of the mathematical approach may be found in Ref⁴¹. The Frobenius norm as a similarity measure for Y and $\hat{X}$ is not tuned for any particular noise distribution, since the noise in our 4D STEM data acquired in the linear mode has a complicated distribution. As a (probably superior) alternative to the Anscombe transform, a Poisson negative log-likelihood could be used as a similarity measurement for counting mode data^{3,58}, which contains closer to pure Poisson noise. A previous report³⁶ also suggests that the

application of Poisson negative log-likelihood performs better under cases with extremely limited counts than the Anscombe transform.

This implementation finds a low-rank representation $\hat{\mathbf{X}}$ that represents a large fraction of the total variance of the noisy data $\mathbf{Y}$, under the assumption that the signal should be more significant and contribute more to the total variance of the data than noise. This method could be less effective if the noise contributes significantly to the total variance, for example, if the data contains some very high-intensity noise points.

2.4 Data Unfolding and Parameter Selection

To generate a 3D tensor for tensor SVD from 4D STEM data, the reciprocal (k_x, k_y) dimensions were unfolded into a single dimension. Other ways to rearrange the data are discussed in section 3.2. For experimental data, each CBED pattern was downsampled by a factor of 4 to 128×128 pixels, then unfolded into a one-dimensional array to generate a noisy signal tensor with $150 \times 150 \times 16384$ points. For simulated data, the unfolded noisy signal tensor contained $114 \times 114 \times 16384$ points.

For d -dimensional signals, the tensor SVD method has $d+1$ tunable parameters: the d ranks of the low-rank approximation and a parameter controlling the number of iterations in HOOI. The tensor ranks are determined with scree tests on 2D matrices, which is justified by tests on simulated data in Section 3.1. The number of HOOI iterations was fixed at ten for all the calculations presented here. Tests on the simulated Si dislocation data show no improvement in peak signal to noise ratio (PSNR)⁵⁹ after denoising after 4-5 iterations, even on input data with a high noise level and an input PSNR ~ 2 dB. On simulated STO data with high periodicity and low noise level, the HOOI method converged after only one iteration. Compute time increases linearly with the number of iterations, so the reported times could be shortened by using fewer

iterations, especially when the input data has a low noise level or high periodicity. Another way to control the iteration is to calculate the difference between the low-rank tensor $\mathbf{X}$ from two consecutive iterations and stop iterating when the difference is below a threshold. This scheme has the disadvantage of either doubling the memory consumption if two copies of $\mathbf{X}$ are saved in order to calculate the difference or doubling the compute time if the tensor $\mathbf{X}$ is recalculated after each iteration.

2.5 Other Denoising Methods and Implementations

To benchmark the denoising performance of tensor SVD, we compare the denoising results from tensor SVD against the denoising results from state-of-the-art methods NLPCA, BM4D, and the more widely used matrix PCA. NLPCA and BM4D were applied to the unwrapped 3D tensor generated from 4D data. Application of NLPCA on 3D atomic resolution STEM EDS spectrum image data has been reported before³⁵, and the parameters optimized for STEM EDS data were used to denoise 4D STEM data. BM4D was proposed and widely applied to MRI data in Ref. ³⁸, and adapted V-BM4D which was designed to handle time sequences⁶⁰ has been applied to denoise both for real-life photos⁶¹ and microscopy images⁶². Considered that our data has different feature sizes and periodicity than MRI data, we have optimized the denoising parameters of BM4D on our own data. The parameters of the BM4D denoising method were optimized one at a time, starting from the Ref. ³⁸ values using simulated STO and Si dislocation data to achieve the highest PSNR after denoising for an input PSNR of 20 dB for STO and 22 dB for Si dislocation. Table 2 presents the parameters for both approaches. Matrix PCA was applied to 4D STEM data using the decomposition method implemented in Hyperspy²⁰. The four-dimensional data was unfolded into a 2D matrix with the real space dimension used as the

navigation dimension, and the reciprocal space was used as the signal dimension in the Hyperspy signal. A scree plot⁶³ was used to determine the number of principal components to use in the denoised data.

NLPCA, BM4D, and tensor SVD were implemented as single thread programs in Matlab. All three Matlab scripts and the Hyperspy Python software package run on a typical Windows desktop with dual Intel Xeon 1.6 GHz processors and 32 GB accessible memory. Matrix SVD has also been implemented in Hyperspy with similar performance to the MATLAB results reported here.

NLPCA		BM4D	
Number of components	10	Method	Wiener filtering
Number of clusters	10	Cube size	4
Patch size along r_x and r_y (px)	12	Group size	32
Patch size along k (px)	Same as the k dimension size	Step	2
		Search-cube size	23 (STO)/31 (Si dislocation)
		Search Similarity threshold	4

Table 2. Denoising parameters for NLPCA and BM4D

3. Results

3.1 Tensor SVD Denoising of Simulated 4D STEM Data

Tensor SVD has been applied to simulated data from the STO structure and Si dislocation structure to test its performance and compare denoised results to the ground truth. Figure 2 shows the denoising performance on simulated STO data with noise appropriate for acquisition

at 1,000 frames per second (FPS) on our Titan with 18.9 pA current and $3.85 \times 10^6 \text{ e}^-/\text{\AA}^2 \cdot \text{ms}$ dose rate. The same beam current and dose rate have been used for all simulations under different frame rate. Denoising performance is quantified in terms of PSNR, an appropriate measure for Poisson-noise corrupted data,⁵⁹ and defined as

$$PSNR = 10 \times \log_{10} \left(\frac{MAX^2}{MSE} \right), \quad [6]$$

where MSE in eq [6] is the mean squared error between the ground truth and noisy image that measures the averaged squared intensity difference, and MAX is the maximum intensity in the ground truth. The ranks of the low-rank tensor representation determined from scree plots were 7 for the real-space dimensions r_x and r_y and 30 for the unwrapped k dimension. The tensor SVD denoised CBED pattern shown in Figure 2 (d) is indistinguishable from the ground truth shown in Figure 2 (f) by eye, as well as the virtual ADF images reconstructed from CBED patterns using denoised result and ground truth, as shown in Figure 2 (c) and (e). The average PSNR for the data set shown in Figure 2 (g) suggests that tensor SVD improves PSNR by about 20 dB under various noise levels. The same tensor ranks were used for all the different noise levels, so the denoising performance might be improved if the parameters were optimized at each specific noise level. All the CBED patterns from different positions on the sample are denoised with higher than 44 dB PSNR, as shown in the spatial map of PSNR in Figure 3 (h), with a variation of less than 5 dB. The spatial variation in PSNR occurs because some CBED patterns show more significant features and are better represented in the low-rank tensor.

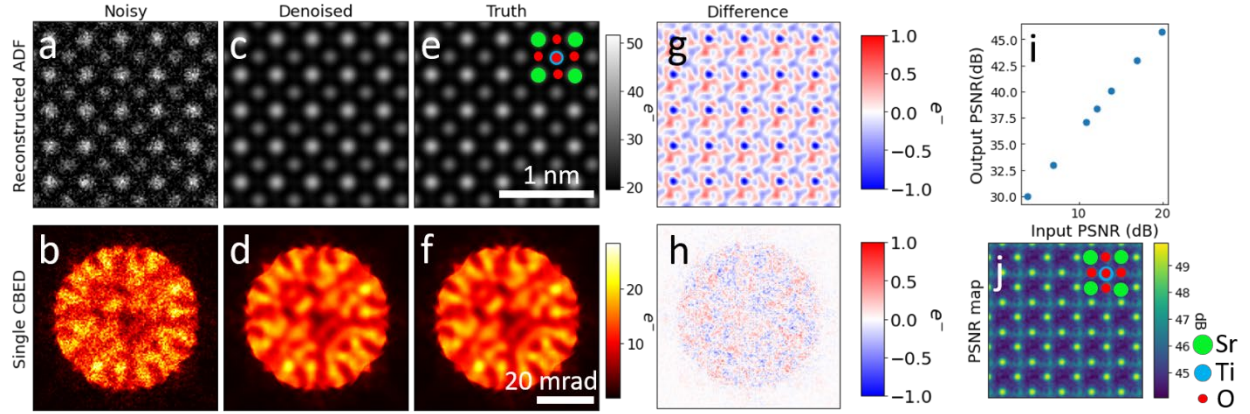


Figure 2. Denoising 4D STEM simulated from a perfect STO. crystal: virtual ADF image and a single CBED pattern pairs from (a)-(b) simulated 1000 FPS noisy data, (c)-(d) tensor SVD denoised data, (e)-(f) ground truth, and (g)-(h) difference between denoised data and ground truth. (i) Output PSNR from SVD denoised data vs. input PSNR from noisy data calculated under different noise levels. (j) Spatial distribution of denoising PSNR for 1000 FPS data.

Figure 3 shows tensor SVD's denoising performance on simulated data from the aperiodic Si dislocation structure. The ranks of the low-rank tensor were 32 in r_x and r_y and 180 for the unwrapped k dimension, significantly higher than the ranks used on STO data because the aperiodic structure data contains more unique features that require higher rank to represent. Average PSNR values are improved by about 15 dB under different noise levels, as shown in Figure 3 (i). The spatial distribution of PSNR in Figure 3 (k) shows about 7 dB lower PSNR around the dislocation core compared to the periodic lattice structure. The lower PSNR in the aperiodic region can also be observed in the denoised CBED pattern from the dislocation core shown in Figure 3 (f), which is noisier than the pattern from the periodic region in Figure 3 (h). Both the PSNR and CBED patterns suggest that the non-regular structures around the dislocation core are not well represented by the low-rank tensor. It also suggests that more cautions should be used when extracting information from non-regular structures in the denoised data.

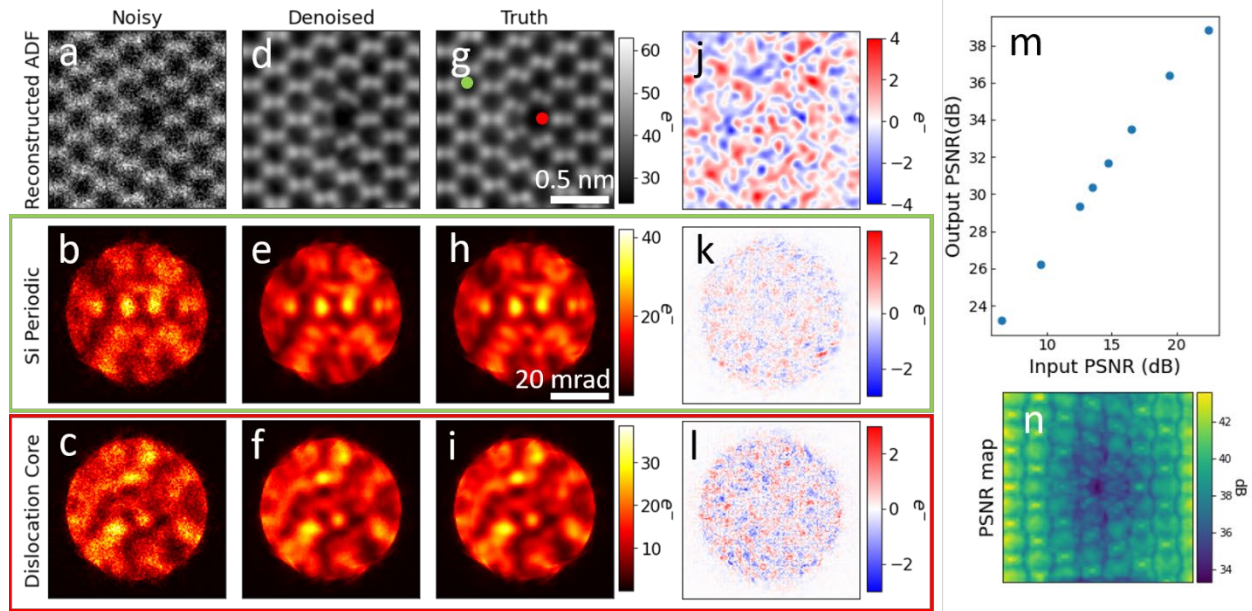


Figure 3. Denoising simulated 4D STEM data from a Si dislocation core: The virtual ADF image, a single CBED pattern from the periodic region of the Si structure at the position of the green dot in (g), and a single CBED pattern from the dislocation core region at the position marked by the red dot in (g) extracted from: (a)-(c) simulated 1000 FPS noisy data, (d)-(f) tensor SVD denoised data, (g)-(i) ground truth, and (j)-(l) the difference between the ground truth and denoised image. (m) Output PSNR from SVD denoised data vs. input PSNR from noisy data at different noise levels. (n) Spatial distribution of denoising PSNR for 1000 FPS data.

Tensor SVD denoising also improves more complicated signals calculated from the 4D STEM data. For example, symmetry STEM is a recently proposed method in which an image is created from the normalized cross correlation of between each CBED pattern and the same pattern after application of a symmetry transformation like a rotation of mirror.⁵⁴ It is intended as an atomically-resolved measure of crystallographic symmetries in the sample. Figure 4 shows

the symmetry STEM image calculated using the method described in ref⁵⁴ from the simulated SrTiO₃ 4D STEM data 180° and 90° rotation transformations. The noiseless ground truth shows the expected symmetries with a high degree of correlation from every atomic site in the [001] unit cell. However, with noise added to mimic 10000 FPS acquisition, the high symmetry peak on Sr sites are missing, creating clear artifacts in the image, and the normalized symmetry signal is overall significantly lower for the entire image. The symmetry STEM images calculated from tensor SVD denoised 4D STEM data both recover all the peaks, removing the noise-induced artifacts, and show a similar symmetry signal level as the ground truth.

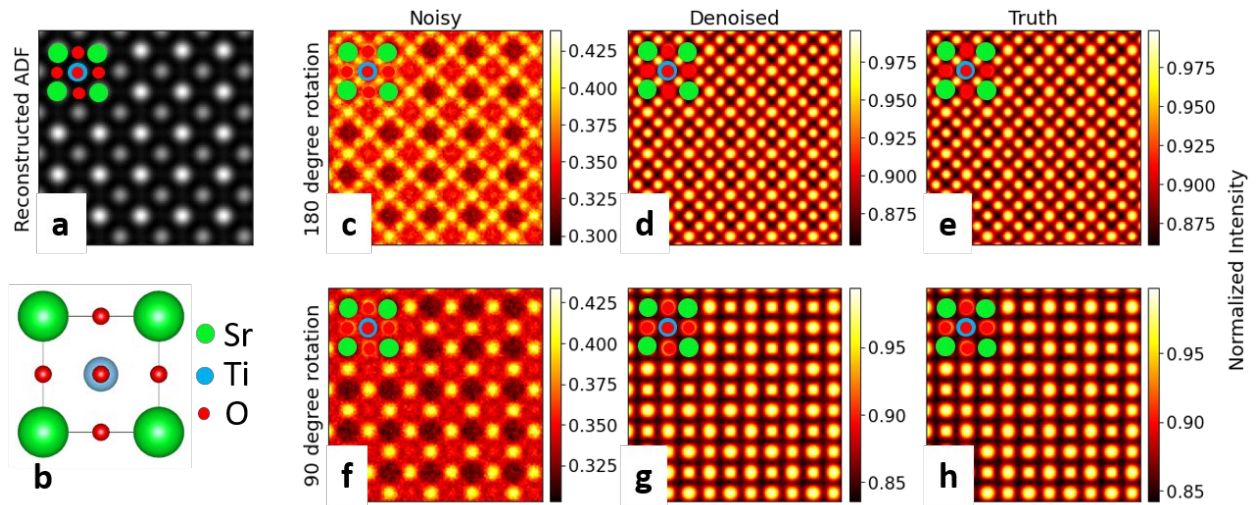


Figure 4 (a) Reconstructed ADF image from 4D STEM data, (b) schematic of SrTiO₃ unit cell in [001] orientation, (c) – (e) symmetry STEM images for a 180° rotation, and (f) – (h) symmetry STEM images for a 90° rotation from noisy data, tensor SVD denoised data, and the ground truth.

Figure 5 shows the accuracy of Si atom positions in ADF images reconstructed from noisy and denoised 4D STEM data (as shown in Figure 3 (a) and (b)). Each Si dumbbell is fitted to the sum of two 2D Gaussian functions, and the accuracy is calculated as the displacement of the atom image position with respect to the true atom position derived from the ground truth. The displacement scatters plot shown in Figure 5 (a) shows a clear improvement after tensor SVD denoising, as all the points gather more closely to the zero displacement point at the center. In Figure 5 (b), accuracy is calculated as the mean displacement of all the Si atoms in the noisy or denoised data, and tensor SVD denoised data shows a clear improvement in accuracy when the frame rate is 2000 FPS or higher. At 1000 FPS, tensor SVD shows little improvement in peak fitting results. A similar scheme has been used to fit the Sr and Ti positions in reconstructed ADF images from simulated 4D STEM data on the SrTiO_3 model, followed by precision calculated as the standard deviation of Sr-Sr bond length and Sr-Ti bond length. The results show about 80% improvement after denoising in the precision calculated from Sr-Ti bond length, and over 90% improvement in the precision calculated from Sr-Sr bond length over various simulation frame rate between 1000 and 10000 FPS.

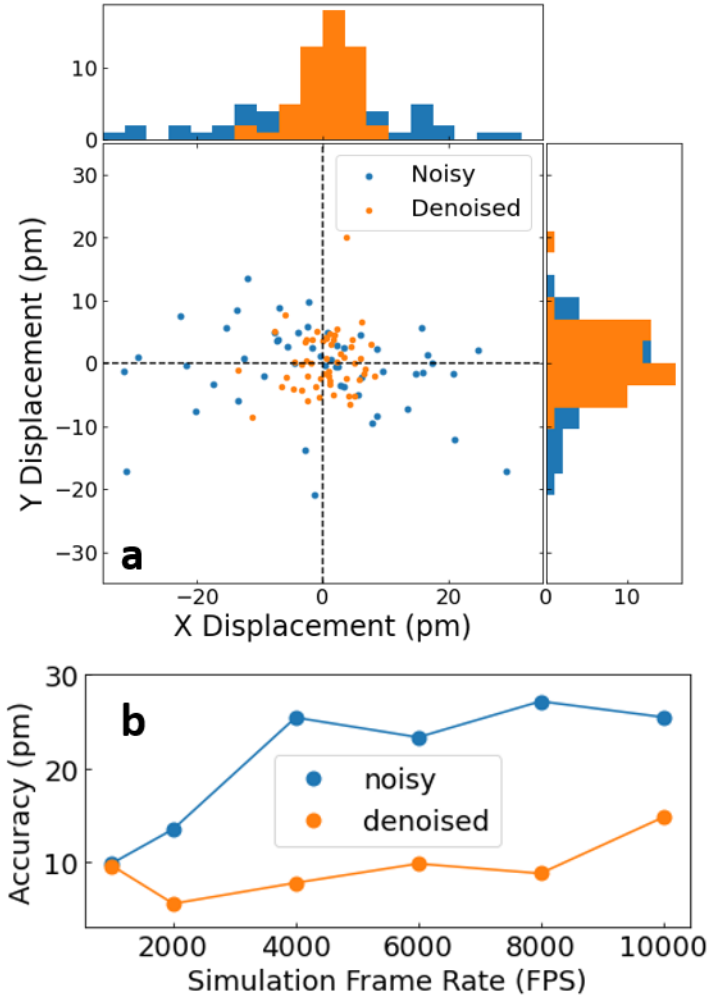


Figure 5 (a) Displacement of every Si site between the reconstructed ADF image using 2000 FPS noisy or denoised data and the ground truth and the histogram along the x and y direction, (b) accuracy calculated from all Si atom positions vs. frame rate used in the simulations

In general, we expect tensor SVD denoising to be helpful for 4D STEM applications that are depend on limited angular range integration or fine features in the CBED patterns, such as the lattice symmetry detection. Tensor SVD will be less helpful for the applications such as electric field⁶⁴ and magnetic field⁶⁵ detection from the center of mass of the CBED patterns,

since these applications involve integrating over a large group of pixels. Moreover, the dimensionality reduction step in tensor SVD does not conserve the total intensity in the CBED pattern, so tensor SVD will create errors for applications that rely on the absolute diffraction intensity..

Figure 6 shows that ranks determined with scree tests from 2D matrices are close to the optimal denoising ranks for 3D tensor. Denoised PSNR at an input PSNR of 20 dB for STO and 22 dB for Si dislocation was computed as a function of the ranks over a wide range (0-100 for r_x and r_y , 0-250 for k dimension) with coarse step size (5 for r_x and r_y , 10 for k dimension), then optimized with a narrow search of increment of 1 along all three dimensions. Figure 4 compares the ranks determined by scree plots (blue dots) and the ranks with the highest PSNR (red lines) for the simulated STO data and Si dislocation data. For all three dimensions of both structures, optimal ranks marked by the red lines are very close to the number of principal components determined by the scree plot. This procedure does not, however, consider all possible combinations of ranks along all dimensions.

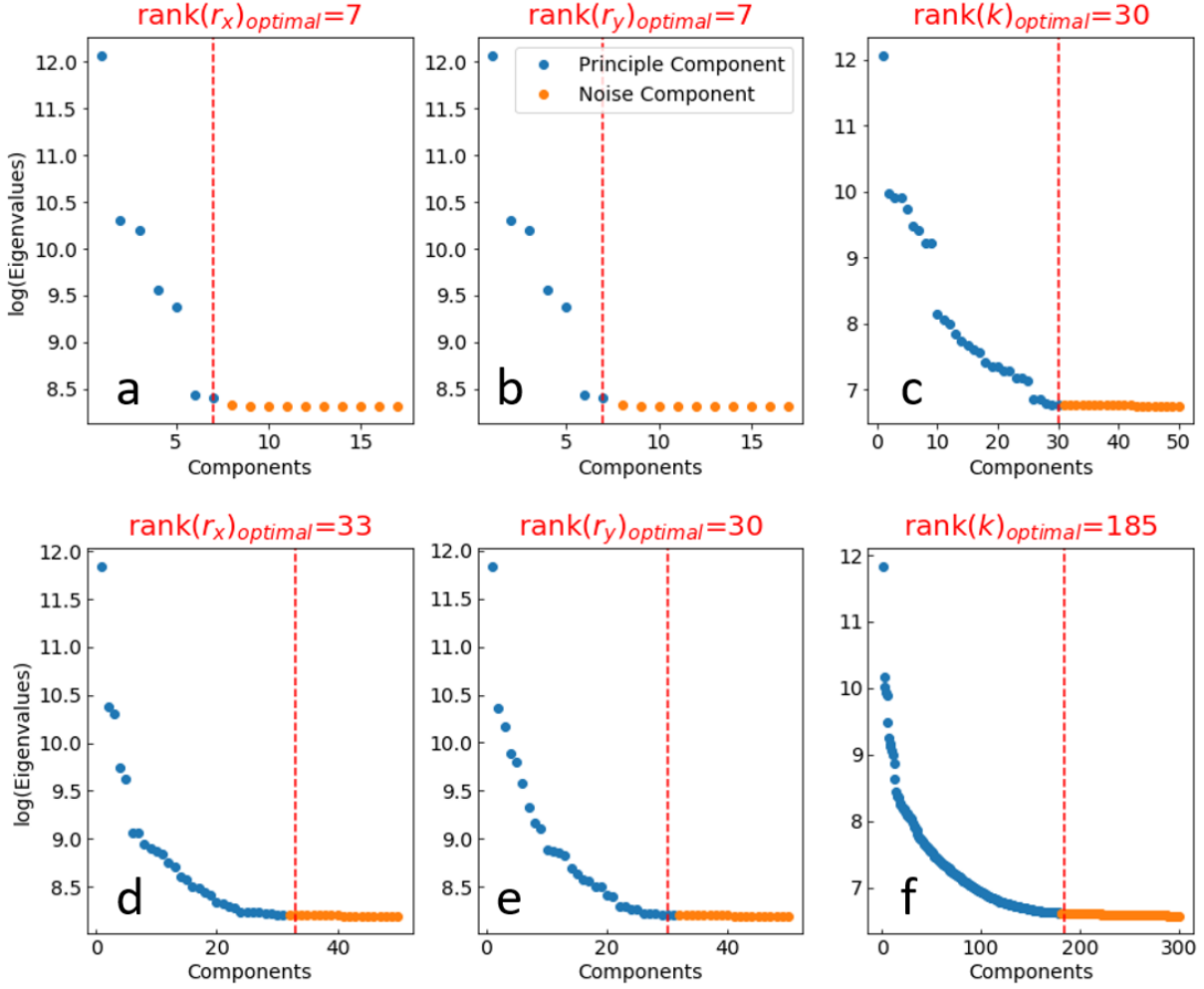


Figure 6. Scree plots of r_x , r_y , and k dimension for (a) - (c) simulated STO data and (d) - (f) simulated Si edge dislocation data. Ranks for each dimension that generate optimal denoising results are reported above each graph and marked with the red dashed lines. Optimal rank from the scree plot is the dividing point between blue data and orange data.

Computing the optimal rank for output PSNR as a function of input PSNR shows that as the noise level increases, the optimal number of principal components decreases. The number of principal components on the Si dislocation data along the k dimension decreases to 50 from 180 with simulated noise corresponding to 10,000 FPS and an input PSNR of 13 dB , and further

decreases to 20 at 40,000 FPS and an input PSNR of 6 dB . A scree test on noiseless data suggests that there are actually 10,500 k dimension components that contribute to the Si dislocation data, corresponding to the complexity and variability of the CBED data, but most of them are overwhelmed once noise is introduced and cannot be identified nor recovered by denoising.

3.2 Alternate Data Shapes

We have explored other ways to generate a 3D input tensor from the 4D STEM data using the simulated 4D STEM data from Si dislocation core. One alternate rearrangement is to unfold the real space dimensions r_x, r_y into one single dimension r to create a 3D tensor that is a stack of CBED patterns. The denoising ranks determined from scree tests are 24, 24, and 170 along the k_x, k_y , and r dimensions, respectively. The denoised ADF image and CBED patterns using this alternate input tensor show improved PSNR, as shown in Figure 5 (c) and (d), but both are blurry compared to the ground truth and to the denoising result by unfolding reciprocal space dimensions (Figure 3 (d), (e), and (f)). The average PSNR from the denoised data is 34.11 dB, about 4 dB lower than the denoised result after unfolding reciprocal space. Unfolding in real

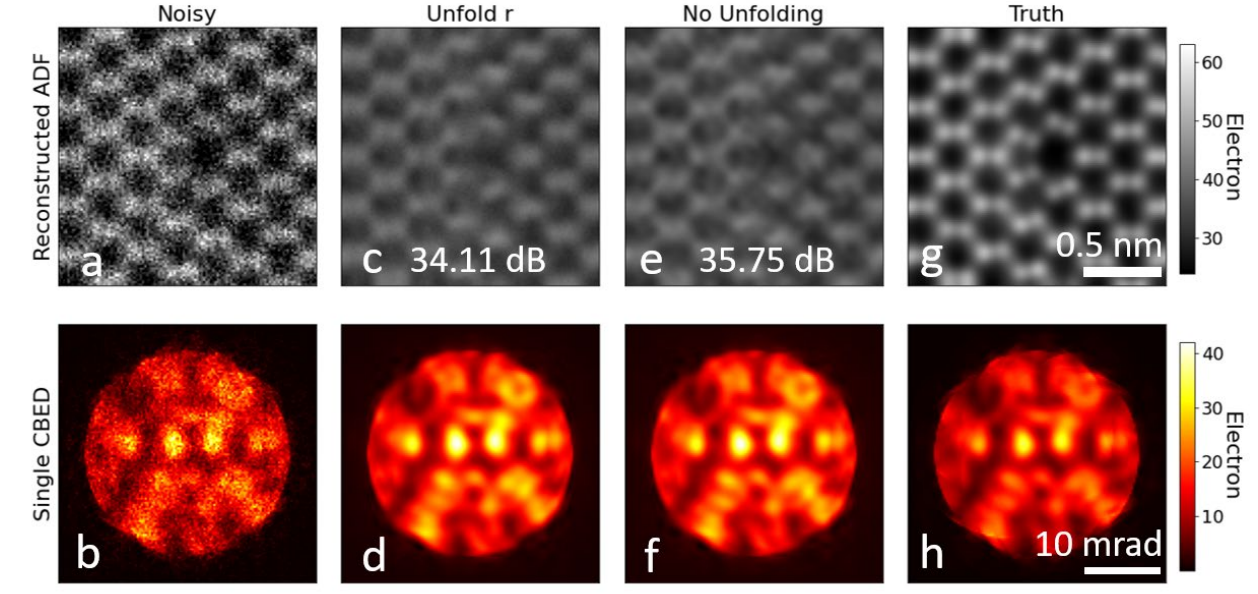


Figure 7. Virtual ADF image and CBED pattern from the periodic region of the Si structure (the position of the green dot in Figure 4 (g)) from (a), (b) noisy data, (c), (d) tensor SVD denoised data with real space dimensions unfolded into one dimension, (e), (f) tensor SVD denoised data without unfolding, (g), (h) ground truth.

space reduces information about the atomic lattice periodicity in the data, making the low-rank representation less effective. The data in reciprocal space are less periodic, so unfolding the reciprocal space is more effective.

Tensor SVD can also treat the four-dimensional data directly, without unfolding, to find the best rank- (r_1, r_2, r_3, r_4) approximation. For Si dislocation data, denoising ranks determined from the scree test are 32, 32, 24, and 24 for two real space dimensions and two reciprocal space dimensions, respectively. Tensor SVD denoised 4D data, as shown in Figure 7 (e) and (f), shows a blurry virtual ADF image and CBED pattern, and the average PSNR is 35.75 dB. Meanwhile, the processing time for 4D input data was 30 times longer than the processing time on 3D input data (see section 4.3). As a result, we do not recommend using the original 4D data as the input

for tensor SVD. From the denoised results using data with unfolded real space dimensions and with no unfolding, we suspect that some blurring may occur when trying to reduce the rank for dimensions without an effective low-rank representation.

3.3 Tensor SVD Denoising of Experimental 4D STEM Data

Tensor SVD has been used to denoise experimental 4D STEM data collected on an STO [100] single crystal sample and at the interface of a LiZnSb [11 $\bar{2}$ 0] hexagonal Heusler alloy thin film grown on GaSb [110] substrate. Experimentally collected data has a more complex noise distribution that does not obey Poisson statistics, and it suffers from scan distortion during data acquisition, both of which introduce added complexity compared to simulated data.

Figure 8 (a) - (d) show tensor SVD's denoising performance on the experimental STO data. The denoised ADF image is significantly improved, with clear atomic sites and higher SNR, and the denoised CBED pattern shows clear arc-shaped features from the Bragg diffraction disks. Scan distortions and other imperfections also are preserved in the denoised image. The top part of the denoised ADF image shows sheared atoms as a result of sample drift. The tensor ranks for STO data are determined to be 30, 32, and 180 for r_x , r_y , and k , respectively, from the scree plots shown in Figure 8 (i) - (k). The ranks are significantly higher than the ranks used for simulated STO data, probably as a result of non-ideal data collection conditions like drift and probe jitter. Tensor SVD denoising took 538.9 sec to compute on this 1.6 GB data set, once the ranks were determined. This includes computing the low-rank tensor representation and the reconstructed data set.

Figure 8 (e) - (h) show tensor SVD's denoising performance on the LiZnSb/GaSb interface data. The bottom left part of the sample has been damaged in the FIB, and the sample is significantly thicker than the STO sample. The virtual ADF image shown in Figure 8 (g) from denoised data shows higher SNR with clear atomic sites compare to the noisy ADF image shown in Figure 8 (e). The denoised CBED pattern shows smooth intensity variation with fewer high-intensity pixels generated by random noise, especially in the dark field region. There are no arc

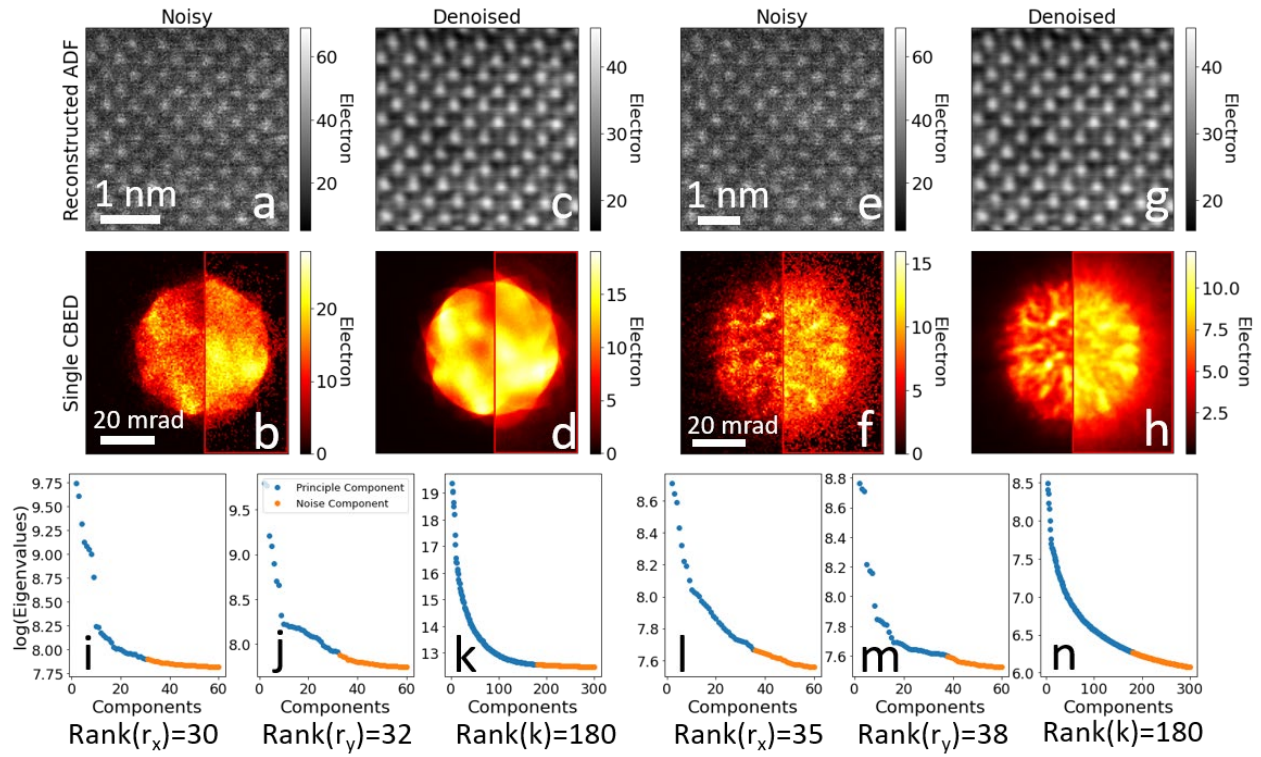


Figure 8 Denoising experimental 4D STEM data: (a) – (h) Virtual ADF images and single CBED patterns pairs. The right half of each CBED is shown as the square root of intensity to emphasize weak features. The color scales are for the linear left half of each image. (a) – (d) Data from STO [100]: (a), (b) as acquired, (c), (d) after tensor SVD denoising. (e) – (f) Data from a LiZnSb [11 $\bar{2}$ 0] / GaSb [110] interface: (e), (f) as acquired, (g), (h) after tensor SVD denoising. (i) – (n) Scree plots used to determine the tensor rank for denoising. Components in blue are retained and components

in orange are discarded in the denoised reconstruction of the data. The 0 component which represents the mean of the data is omitted from the graphs. (i) – (k) STO. scree plots. (l) – (n) LiZnSb / GaSb scree plots.

features from Bragg disks because of the sample thickness. The tensor ranks are 35, 38, and 180 for r_x , r_y , and k , determined using the scree plots shown in Figure 6 (l) - (n). The denoising ranks are very close to the ranks used for STO data, even though the LiZnSb/GaSb sample consists of two different lattice structures, confirming that the rank is heavily influenced by non-sample factors like scan distortion. It took 575.0 sec for tensor SVD to denoise the 2.7 GB LiZnSb/GaSb experimental data. The processing time was mostly determined by the tensor ranks, so the processing times for LiZnSb data and STO data were similar even though the LiZnSb data is about 60% larger than the STO data (see section 4.1 for a detailed discussion).

3.4 Applications to 3D Spectrum Image Data Sets

We have applied tensor SVD to denoise EDS spectrum images acquired from [001] zone axis of a Ca-stabilized $\text{Nd}_{2/3}\text{TiO}_3$ sample (previously published in Ref. ^{35,66}). This data set has a much lower signal than the 4D STEM data, with a mean value of 0.015 counts across the entire data set. The EDS spectrum images from 1000 energy channels were integrated into seven elemental maps corresponding to seven elemental x-ray bands (O $K\alpha$, Ca $K\alpha$, Ti $K\alpha$, Ti $K\beta$, Nd $L\alpha$, Nd $L\beta$, and Nd $L\beta_4$), which increases the mean signal to 1.8 counts. Scree tests show that tensor SVD ranks should be 10, 10, and 7 along x , y , and the energy dimensions, respectively. For comparison, NLPCA and matrix PCA denoised results were replotted from ref ³⁵.

Figure 9 shows the tensor SVD results compared to NLPCA and matrix PCA results replotted from Ref³⁵. The tensor SVD performance is better than matrix PCA but not as good as NLPCA. Tensor SVD shows improved SNR for both the Nd and Ti maps and is able to distinguish the difference in the atomic sites occupied by Ti and Nd. The denoised Ca map from tensor SVD is smoother than the raw data with reduced noise. However, the lattice structures shown in the Ca map does not agree with the NLPCA denoised result and does not show separate atomic sites. We suspect the features in the denoised Ca map could be mostly denoising artifacts. Overall, NLPCA denoising results in clearer atomic features, especially on the low signal Ca map. This is perhaps unsurprising, given that NLPCA was developed for extreme noisy data³⁶, including features like the non-local approach and Poisson maximum-likelihood similarity measurement. Tensor SVD computing time for this small data set are very short (0.1 sec in this case).

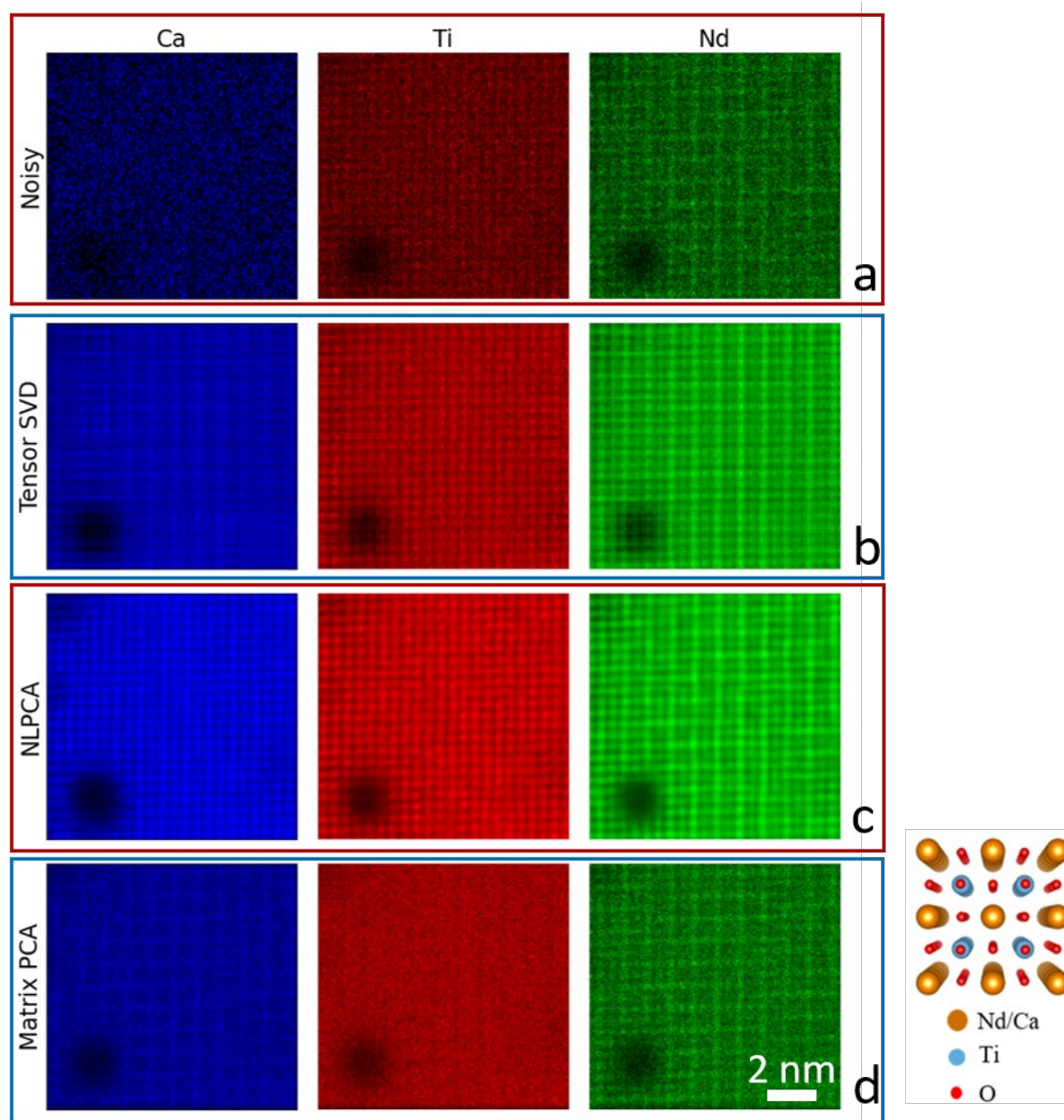


Figure 9, EDS elemental maps from (a) noisy data, (b) tensor SVD denoised data, (c) NLPCA denoised data, and (d) matrix PCA denoised data displayed in the order of Ca, Ti, Nd, from left to right. (a), (c) and (d) are replotted from the data in ref³⁵.

3.5 Comparison to Other Methods for 4D STEM Denoising

Subsets of the simulated 4D STEM data with reduced size in reciprocal space were constructed from both STO and Si dislocation simulations to compare the performance of tensor SVD, NLPCA, BM4D, and matrix PCA. Reduced data size was required in order to accommodate the massive memory requirements of NLPCA. The subset was selected by keeping all the pixels in real space but using a 10-by-10 pixel patch covering a square area with 5 mrad side length around the optical axis in reciprocal space. The final 3D tensor input with reciprocal space unfolded into one dimension has $114 \times 114 \times 100$ data size.

Figure 10 shows the comparison between tensor SVD, NLPCA, BM4D, and matrix PCA using the reduced data. Due to a small region of reciprocal space involved, the contrast on the virtual real space image is different from what we would typically expect from a round-shaped bright-field detector covering a range of collection angles. For example, the virtual image does not show structural details such as the dumbbell shape Si atom pairs. All four methods effectively remove image noise and improve SNR. For example, the virtual image and CBED pattern subset from tensor SVD and NLPCA are visually indistinguishable from the ground truth. The BM4D denoised CBED patterns show some visible mismatch to the ground truth and some pixels with visible noise. Matrix PCA denoised CBED pattern looks smooth without significant noisy pixels, but the intensity is not a good match to the ground truth. The PSNR results in Figure 10 (e) and (f) show that across a wide range of input PSNRs, tensor SVD and NLPCA improve PSNR by 20 - 25 dB, and BM4D and matrix PCA improve PSNR by 10 - 15 dB. Tensor SVD has either the best performance or close to the best performance among the four methods in terms of PSNR. A similar conclusion can be drawn using other metrics to measure the denoising performance, including SSIM⁵⁹ and MSE without the normalization step used in PSNR. Figure 8

also shows the elapsed computing required for the various algorithms for this 10 MB data set. Matrix PCA is the fastest, followed by tensor SVD, and both of those approaches are at least 100 times faster than BM4D or NLPCA. Optimizing the denoising rank for each noise level improves the PSNR on the Si dislocation data, but has little improvement on STO data.

NLPCA is much slower than the rank-reduction methods but has the best denoising performance at high input noise level, as the method was designed and optimized to handle extremely low signal cases with a sophisticated clustering method for similarity detection and a Poisson noise specific similarity measure. The BM4D and matrix PCA methods do not denoise well on either the STO or Si dislocation structure compared to NLPCA and tensor SVD, but matrix PCA is the fastest one among the four different methods. We suspect one of the reasons behind BM4D's poor performance is that similar cubes are searched only within the search-cube size, which does not capture the overall periodicity from the lattice structure.

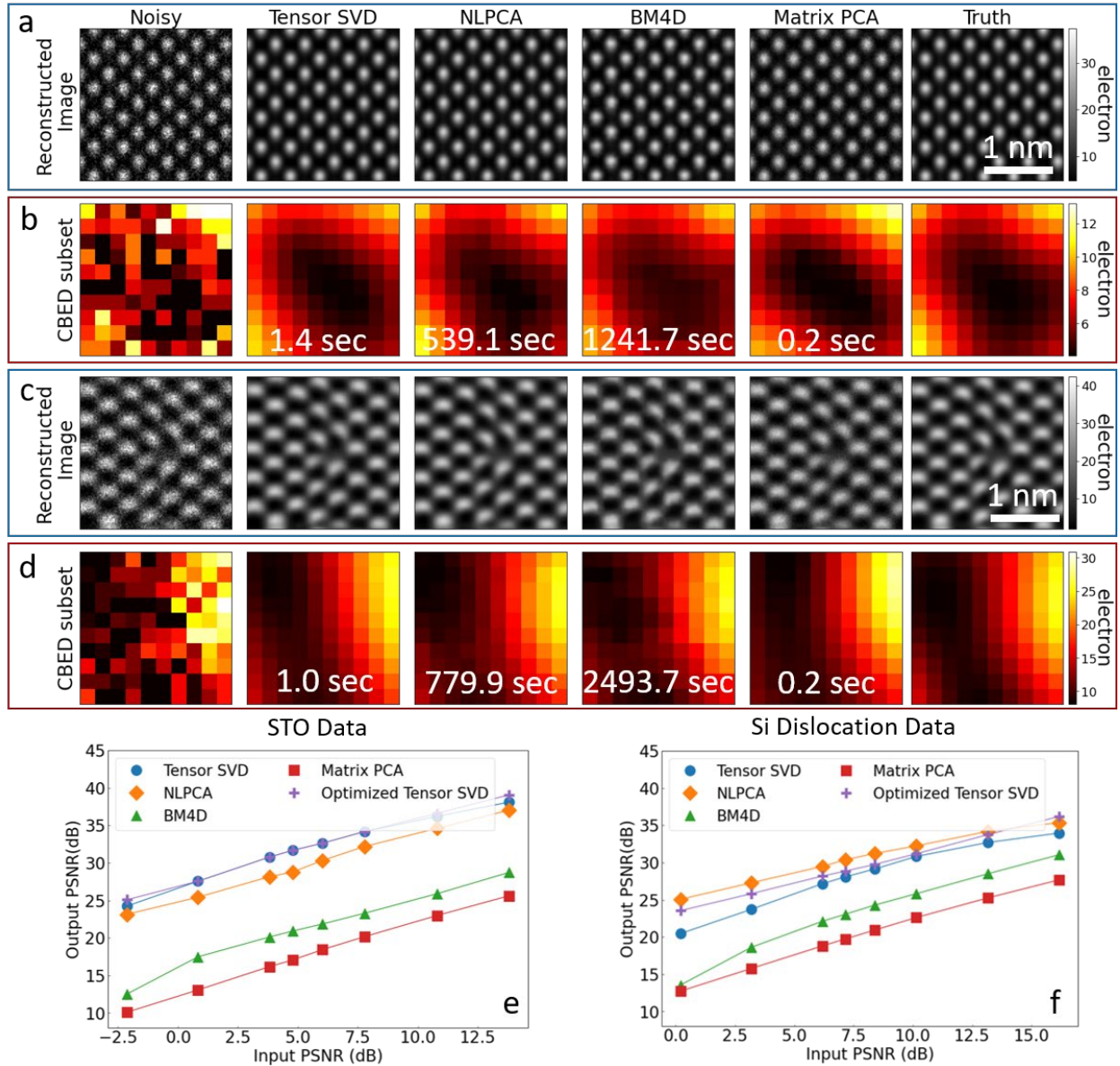


Figure 10. Comparison of tensor SVD to other denoising methods: (a) virtual images from simulated STO data, (b) CBED pattern patches from STO. data, (c) virtual images from simulated Si dislocation data, (d) CBED pattern subset from Si dislocation data extracted. (a) – (d) use the signal level of 1000 fps experimental data, and the numbers in (b) and (d) are the computing time required for denoising. (e)-(f) PSNR of the denoised data as a function of input noise level for (e) STO data and (f) Si dislocation data.

4. Discussion

Tensor SVD provides similar or better performance than the other state-of-the-art methods we have tested, NLPCA and BM4D, at a fraction of the computing time. The majority of computing time in tensor SVD problem is spent on matrix SVD calculations as the HOOI algorithm goes through each dimension to update the corresponding unitary matrix $U^{(k)}$. For each dimension, an eigenvalue problem is solved on a square matrix with the size equal to the product of ranks for all the rest of the dimensions. When using 4D data instead of 3D data, the algorithm loops over one more dimension within each iteration of HOOI, and the eigenvalue problem for each dimension must be solved on a much larger matrix. As a result, the processing time on 4D data is tens of times longer than the processing time on a rearranged 3D data for the data sets tested here and will scale non-linearly with the data dimensions.

The computing time for tensor SVD does not vary for different data sets when the denoising ranks are fixed, which makes it easy to predict the computation time. In contrast, both NLPCA and BM4D can have varying computing time for different data sets even of the same size because different data can result in different clustering of the non-local patches. For NLPCA, if the majority of the patches are grouped inside one cluster, then the PCA step on this cluster will be slow, and the denoising process will be slower compared to the cases when the cubes are uniformly distributed over multiple clusters.

Another advantage of tensor SVD is that it occupies a comparatively small amount of memory while denoising, especially compared to the non-local methods. Our implementation of tensor SVD has a peak memory consumption that is two times the original data size, which makes it possible to handle larger data sets. The computing time and memory use for matrix PCA are both smaller than for tensor SVD, but the performance is worse.

We realize that computing time and memory requirements are implementation-dependent. All the results from tensor SVD, NLPCA, and BM4D presented here are based on a single-threaded implementation in Matlab using a personal computer to allow an “apples to apples” comparison. A single-threaded implementation of tensor SVD in python made available as part of the Hyperspy package has similar performance. Processing time for tensor SVD probably could be shortened using GPU computing with careful consideration of the use of memory.

Although we have not tested it, tensor SVD denoising of low-resolution 4D STEM data using a small convergence angle and large electron probe for strain mapping^{67,68} or to measure the structure of an amorphous material^{1,69} should be possible. Tensor SVD may be less effective than for atomic resolution 4D STEM data, however, as the lower periodicity in the data may mean that the intrinsic tensor rank is higher. For orientation mapping, for example, unwrapping in real space to preserve the small number of distinct patterns in reciprocal space may generate better results. Tensor SVD should also be capable of identifying prototype diffraction patterns from large data sets for phase and orientation mapping.

5. Conclusions

Tensor SVD is a method to find a low-rank representation of tensor data. It is analogous to component analysis of matrix data using PCA, but it preserves the full three or four (or more) dimensional structure of the data set. We have used tensor SVD to denoise atomic-resolution 4D STEM data. The method can be directly applied to multiple GB 4D STEM data sets with the adjustable parameters determined by scree tests. Computing times on a typical desktop computer is a few minutes, and the required memory is twice the size of the data. Tests on simulated data

show that tensor SVD can benefit applications include symmetry STEM and atom position fitting, especially under a limited electron dose. Measured by PSNR of the denoised data, tensor SVD has the best or close to the best performance compared to other state-of-the-art methods which require substantially more computation time and memory. Tests on experimental data show that tensor SVD significantly improves the signal to noise ratio, and the denoised data shows clearer structures in both real space images and CBED patterns. Tensor SVD can also be applied to other high dimensional data sets, such as EDS spectrum images.

Acknowledgments

We thank Rebecca Willett for introducing the authors and suggesting this line of research, and our colleague Carter Francis for help on python implementation and integrating with Hyperspy package.

Competing interests

The authors declare that they have no known competing for financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials

Computer codes for tensor SVD denoising, data processing, and data visualization are available on DOI:10.5281/zenodo.3911801. All the computing codes and simulated data, experimental data, and denoised data (total 152 GB) are available via the Materials Data Facility <https://doi.org/10.18126/vh9q-i1l6>. A Python version of tensor SVD runs on top of Hyperspy is available on https://github.com/paul-voyles/tensor_svd.

Funding

This work was primarily supported by the US Department of Energy, Basic Energy Sciences (DE-FG02-08ER46547), for work by C. Z. and P. M. V. for 4D STEM experiments and simulations, denoising calculations, and implementation of tensor SVD in python. Development of tensor SVD by R. H. and A. R. Z. was supported by National Science Foundation (DMS-1811868, CAREER-1944904), National Institutes of Health (R01-GM131399). Experiments were conducted using facilities supported by the Wisconsin MRSEC (DMR-1720415).

Author's contributions

CZ collected experimental and simulated 4D STEM dataset and performed the denoising. RH, and ARZ developed the tensor SVD method. CZ developed the python implementation for Hyperspy and wrote the manuscript. PMV supervised the project. All authors together revised the manuscript.

1. Ophus, C. Four-Dimensional Scanning Transmission Electron Microscopy (4D-STEM): From Scanning Nanodiffraction to Ptychography and Beyond. *Microsc. Microanal.* **25**, 563–582 (2019). DOI: 10.1017/S1431927619000497
2. Nord, M., Webster, R. W. H., Paton, K. A., Mcvitie, S., Mcgrouter, D. & Paterson, G. W. Fast Pixelated Detectors in Scanning Transmission Electron Microscopy. Part I: Data Acquisition, Live Processing and Storage. 1–14 (2019).
3. Phys, A. & Liberti, E. Phase reconstruction using fast binary 4D STEM data Phase reconstruction using fast binary 4D STEM data. **124101**, (2020). DOI: 10.1063/1.5143213
4. Paterson, G. W., Lamb, R. J., Ballabriga, R., Maneuski, D., O'Shea, V. & McGrouther, D. Sub-100 nanosecond temporally resolved imaging with the Medipix3 direct electron detector. *Ultramicroscopy* **210**, 112917 (2020). DOI: 10.1016/j.ultramic.2019.112917
5. Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D. A. & Cheng, Y. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013). DOI: 10.1038/nmeth.2472
6. Tate, M. W., Purohit, P., Chamberlain, D., Nguyen, K. X., Hovden, R., Chang, C. S., Deb, P., Turgut, E., Heron, J. T., Schlom, D. G., Ralph, D. C., Fuchs, G. D., Shanks, K. S., Philipp, H. T., Muller, D. A. & Gruner, S. M. High Dynamic Range Pixel Array Detector for Scanning Transmission Electron Microscopy. *Microsc. Microanal.* 1–13 (2016). doi:10.1017/S1431927615015664 DOI: 10.1017/S1431927615015664
7. McMullan, G., Faruqi, A. R. & Henderson, R. *Direct Electron Detectors. Resolut. Revolut. Recent Adv. cryoEM* **579**, (Elsevier Inc., 2016). DOI:

10.1016/bs.mie.2016.05.056

8. McMullan, G., Faruqi, A. R., Clare, D. & Henderson, R. Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy* **147**, 156–163 (2014). DOI: 10.1016/j.ultramic.2014.08.002
9. Voyles, P. M. Informatics and data science in materials microscopy. *Curr. Opin. Solid State Mater. Sci.* **21**, 141–158 (2017). DOI: 10.1016/j.cossms.2016.10.001
10. Wold, S; Esbensen, K; Geladi, P. Principle component analysis. *Chemom. Intell. Lab. Syst.* **2**, 37–52 (1987). DOI: 10.5455/ijlr.20170415115235
11. Jolliffe, I. T. *Principal Component Analysis*. Springer (Springer-Verlag, 2002). doi:10.1007/b98835 DOI: 10.1007/b98835
12. Collins, M. & Schapire, R. E. A Generalization of Principal Component Analysis to the Exponential Family. *Adv. Neural Inf. Process. Syst.* (2001).
13. Trebbia, P. & Bonnet, N. EELS elemental mapping with unconventional methods I. Theoretical basis: Image analysis with multivariate statistics and entropy concepts. *Ultramicroscopy* **34**, 165–178 (1990). DOI: 10.1016/0304-3991(90)90070-3
14. Bosman, M., Watanabe, M., Alexander, D. T. L. & Keast, V. J. Mapping chemical and bonding information using multivariate analysis of electron energy-loss spectrum images. *Ultramicroscopy* **106**, 1024–1032 (2006). DOI: 10.1016/j.ultramic.2006.04.016
15. Allen, F. I., Watanabe, M., Lee, Z., Balsara, N. P. & Minor, A. M. Chemical mapping of a block copolymer electrolyte by low-loss EFTEM spectrum-imaging and principal component analysis. *Ultramicroscopy* **111**, 239–244 (2011). DOI:

10.1016/j.ultramic.2010.11.035

16. Burke, M. G., Watanabe, M., Williams, D. B. & Hyde, J. M. Quantitative characterization of nanoprecipitates in irradiated low-alloy steels: Advances in the application of FEG-STEM quantitative microanalysis to real materials. *J. Mater. Sci.* **41**, 4512–4522 (2006). DOI: 10.1007/s10853-006-0084-x
17. Yaguchi, T., Konno, M., Kamino, T. & Watanabe, M. Observation of three-dimensional elemental distributions of a Si device using a 360 degree-tilt FIB and the cold field-emission STEM system. *Ultramicroscopy* **108**, 1603–1615 (2008). DOI: 10.1016/j.ultramic.2008.06.003
18. Watanabe, M., Okunishi, E. & Ishizuka, K. Analysis of Spectrum-Imaging Datasets in Atomic-Resolution Electron Microscopy. *Microsc. Anal.* **23**, 5–7 (2009).
19. MSA. at <<http://temdm.com/web/msa/>>
20. Peña, F. de la, Prestat, E., Fauske, V. T., Burdet, P., Jokubauskas, P., Nord, M., Ostasevicius, T., MacArthur, K. E., Sarahan, M., Johnstone, D. N., Taillon, J., Lähnemann, J., Migunov, V., Eljarrat, A., Caron, J., Aarholt, T., Mazzucco, S., Walls, M., Slater, T., *et al.* hyperspy/hyperspy: HyperSpy v1.5.2. (2019). doi:10.5281/ZENODO.3396791 DOI: 10.5281/ZENODO.3396791
21. Hyvarinen, A. & Oja, E. Independent component analysis: algorithms and applications. *Neural Networks* **13**, 411–430 (2000). DOI: 10.7819/rbgn.v19i63.1905
22. Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L. & Smith, S. M. Automatic denoising of functional MRI data: Combining independent component

- analysis and hierarchical fusion of classifiers. *Neuroimage* **90**, 449–468 (2014). DOI: 10.1016/j.neuroimage.2013.11.046
23. Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P. & Plemmons, R. J. Algorithms and applications for approximate nonnegative matrix factorization. *Comput. Stat. Data Anal.* **52**, 155–173 (2007). DOI: 10.1016/j.csda.2006.11.006
 24. Wilson, K. W., Raj, B., Smaragdis, P. & Divakaran, A. Speech denoising using nonnegative matrix factorization with priors. in *2008 IEEE Int. Conf. Acoust. Speech Signal Process.* 4029–4032 (IEEE, 2008). doi:10.1109/ICASSP.2008.4518538 DOI: 10.1109/ICASSP.2008.4518538
 25. Chen, Y., Huang, T. Z., Zhao, X. Le & Deng, L. J. Hyperspectral image restoration using framelet-regularized low-rank nonnegative matrix factorization. *Appl. Math. Model.* **63**, 128–147 (2018). DOI: 10.1016/j.apm.2018.06.044
 26. Ye, M., Qian, Y. & Zhou, J. Multitask sparse nonnegative matrix factorization for joint spectral-spatial hyperspectral imagery denoising. *IEEE Trans. Geosci. Remote Sens.* **53**, 2621–2639 (2015). DOI: 10.1109/TGRS.2014.2363101
 27. Karami, A., Yazdi, M. & Zolghadre Asli, A. Noise reduction of hyperspectral images using kernel non-negative tucker decomposition. *IEEE J. Sel. Top. Signal Process.* **5**, 487–493 (2011). DOI: 10.1109/JSTSP.2011.2132692
 28. He, W., Zhang, H. & Zhang, L. Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **55**, 3909–3921 (2017). DOI: 10.1109/TGRS.2017.2683719

29. Buades, a. & Coll, B. A non-local algorithm for image denoising. *Comput. Vis. Pattern* **2**, 60–65 (2005). DOI: 10.1109/CVPR.2005.38
30. Shyjila, P. A. & Wilscy, M. Non local means image denoising for color images using PCA. *Commun. Comput. Inf. Sci.* **131 CCIS**, 288–297 (2011). DOI: 10.1007/978-3-642-17857-3_29
31. Dabov, K., Foi, A. & Katkovnik, V. Image denoising by sparse 3D transformation-domain collaborative filtering. *IEEE Trans. Image Process.* **16**, 1–16 (2007). DOI: 10.1109/TIP.2007.901238
32. Mevenkamp, N., Yankovich, A. B., Voyles, P. M. & Berkels, B. Non-local Means for Scanning Transmission Electron Microscopy Images and Poisson Noise based on Adaptive Periodic Similarity Search and Patch Regularization. in *Vision, Model. Vis.* 63–70 (2014). doi:10.2312/vmv.20141277 DOI: 10.2312/vmv.20141277
33. Mevenkamp, N., Binev, P., Dahmen, W., Voyles, P. M., Yankovich, A. B. & Berkels, B. Poisson noise removal from high-resolution STEM images based on periodic block matching. *Adv. Struct. Chem. Imaging* **1**, 3 (2015). DOI: 10.1186/s40679-015-0004-8
34. Deledalle, C.-A., Tupin, F. & Denis, L. Poisson NL means: Unsupervised non local means for Poisson noise. in *2010 IEEE Int. Conf. Image Process.* 801–804 (IEEE, 2010). doi:10.1109/ICIP.2010.5653394 DOI: 10.1109/ICIP.2010.5653394
35. Yankovich, A. B., Zhang, C., Oh, A., Slater, T. J. A., Azough, F., Freer, R., Haigh, S. J., Willett, R. & Voyles, P. M. Non-rigid registration and non-local principle component analysis to improve electron microscopy spectrum images. *Nanotechnology* **27**, 364001 (2016). DOI: 10.1088/0957-4484/27/36/364001

36. Salmon, J., Deledalle, C., Willett, R. & Harmany, Z. Poisson noise reduction with non-local PCA. *2012 IEEE Int. Conf. Acoust. Speech Signal Process.* 1109–1112 (2012). doi:10.1109/ICASSP.2012.6288081 DOI: 10.1109/ICASSP.2012.6288081
37. Likas, A., Vlassis, N. & J. Verbeek, J. The global k-means clustering algorithm. *Pattern Recognit.* **36**, 451–461 (2003). DOI: 10.1016/S0031-3203(02)00060-2
38. Maggioni, M., Katkovnik, V., Egiazarian, K. & Foi, A. Nonlocal Transform-Domain Filter for Volumetric Data Denoising and Reconstruction. *IEEE Trans. Image Process.* **22**, 119–133 (2013). DOI: 10.1109/TIP.2012.2210725
39. Chen, J. & Saad, Y. On the tensor svd and the optimal low rank orthogonal approximation of tensors. *SIAM J. Matrix Anal. Appl.* **30**, 1709–1734 (2008). DOI: 10.1137/070711621
40. Zhang, Z., Ely, G., Aeron, S., Hao, N. & Kilmer, M. Novel methods for multilinear data completion and de-noising based on tensor-SVD. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 3842–3849 (2014). doi:10.1109/CVPR.2014.485 DOI: 10.1109/CVPR.2014.485
41. Zhang, A. & Xia, D. Tensor SVD: Statistical and Computational Limits. *IEEE Trans. Inf. Theory* **64**, 7311–7338 (2018). DOI: 10.1109/TIT.2018.2841377
42. Zhang, A. & Han, R. Optimal Sparse Singular Value Decomposition for High-Dimensional High-Order Data. *J. Am. Stat. Assoc.* **114**, 1708–1725 (2019). DOI: 10.1080/01621459.2018.1527227
43. Han, R., Willett, R. & Zhang, A. An Optimal Statistical and Computational Framework for Generalized Tensor Estimation. (2020). at <<http://arxiv.org/abs/2002.11255>>

44. Du, D., Strohbeen, P. J., Paik, H., Zhang, C., Genser, K., Rabe, K. M., Voyles, P. M., Schlom, D. G. & Kawasaki, J. K. Control of polymorphism during epitaxial growth of hyperferroelectric candidate LiZnSb on GaSb (111)B. 1–6 (2020). at <http://arxiv.org/abs/2001.05611>
45. Zhang, C., Feng, J., DaCosta, L. R. & Voyles, P. M. Atomic resolution convergent beam electron diffraction analysis using convolutional neural networks. *Ultramicroscopy* **210**, 112921 (2020). DOI: 10.1016/j.ultramic.2019.112921
46. Yankovich, A. B., Berkels, B., Dahmen, W., Binev, P., Sanchez, S. I., Bradley, S. a, Li, A., Szlufarska, I. & Voyles, P. M. Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts. *Nat. Commun.* **5**, 4155 (2014). DOI: 10.1038/ncomms5155
47. Pryor, A., Ophus, C. & Miao, J. A Streaming Multi-GPU Implementation of Image Simulation Algorithms for Scanning Transmission Electron Microscopy. *Adv. Struct. Chem. Imaging* (2017). doi:10.1186/s40679-017-0048-z DOI: 10.1186/s40679-017-0048-z
48. Dwyer, C., Erni, R. & Etheridge, J. Method to measure spatial coherence of subangstrom electron beams. *Appl. Phys. Lett.* **93**, 021115 (2008). DOI: 10.1063/1.2957648
49. Dwyer, C., Erni, R. & Etheridge, J. Measurement of effective source distribution and its importance for quantitative interpretation of STEM images. *Ultramicroscopy* **110**, 952–957 (2010). DOI: 10.1016/j.ultramic.2010.01.007
50. Yamanaka, T., Hirai, N. & Komatsu, Y. Structure change of $\text{Ca}_{1-x}\text{Sr}_x\text{TiO}_3$ perovskite with composition and pressure. *Am. Mineral.* **87**, 1183–1189 (2002). DOI:

10.2138/am-2002-8-917

51. Anscombe, F. J. The Transformation of {P}oisson, Binomial, and Negative-Binomial Data. *Biom* **35**, 246–254 (1948).
52. Tucker, L. SOME MATHEMATICAL NOTES ON THREE-MODE FACTOR ANALYSIS. *Z. Psychol.* **31**, 279–311 (1966).
53. Im, S., Chen, Z., Johnson, J. M., Zhao, P., Yoo, G. H., Park, E. S., Wang, Y., Muller, D. A. & Hwang, J. Direct determination of structural heterogeneity in metallic glasses using four-dimensional scanning transmission electron microscopy. *Ultramicroscopy* **195**, 189–193 (2018). DOI: 10.1016/j.ultramic.2018.09.005
54. Krajnak, M. & Etheridge, J. A symmetry-derived mechanism for atomic resolution imaging. 1–7 (2019). at <<http://arxiv.org/abs/1903.04780>>
55. Lathauwer, L. D. E., Moor, B. D. E. & Vandewalle, J. On The Best Rank-1 and Rank-(R1, R2, ..., Rn) Approximation of Higher-Order Tensors. **21**, 1324–1342 (2000).
56. Cueva, P., Hovden, R., Mundy, J. A., Xin, H. L. & Muller, D. A. Microscopy Microanalysis Data Processing for Atomic Resolution Electron Energy Loss Spectroscopy. *Microsc. Microanal.* **18**, 667–675 (2012).
57. Titchmarsh, J. M. & Dumbill, S. Multivariate statistical analysis of FEG-STEM EDX spectra. *J. Microsc.* **184**, 195–207 (1996). DOI: 10.1046/j.1365-2818.1996.1400698.x
58. Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D. A. & Cheng, Y. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013). DOI:

10.1038/nmeth.2472

59. Horé, A. & Ziou, D. Image quality metrics: PSNR vs. SSIM. *Proc. - Int. Conf. Pattern Recognit.* 2366–2369 (2010). doi:10.1109/ICPR.2010.579 DOI: 10.1109/ICPR.2010.579
60. Maggioni, M., Boracchi, G., Foi, A. & Egiazarian, K. Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms. *IEEE Trans. Image Process.* **21**, 3952–3966 (2012). DOI: 10.1109/TIP.2012.2199324
61. Kuang, Y., Zhang, L. & Yi, Z. An adaptive rank-sparsity K-SVD algorithm for image sequence denoising. *Pattern Recognit. Lett.* **45**, 46–54 (2014). DOI: 10.1016/j.patrec.2014.03.003
62. Furnival, T., Leary, R. K. & Midgley, P. A. Denoising time-resolved microscopy image sequences with singular value thresholding. *Ultramicroscopy* **178**, 112–124 (2017). DOI: 10.1016/j.ultramic.2016.05.005
63. D’agostino, R. B. & Russell, H. K. Scree Test. *Encycl. Biostat.* 1–3 (2005). doi:10.1002/0470011815.b2a10082 DOI: 10.1002/0470011815.b2a10082
64. Müller-Caspary, K., Krause, F. F., Béché, A., Schowalter, M., Galioit, V., Löffler, S., Verbeeck, J., Zweck, J., Schattschneider, P. & Rosenauer, A. Atomic electric fields revealed by a quantum mechanical approach to electron picodiffraction. *Nat. Commun.* **5**, 5653 (2014). DOI: 10.1038/ncomms6653
65. Matsumoto, T., So, Y.-G., Kohno, Y., Sawada, H., Ikuhara, Y. & Shibata, N. Direct observation of S7 domain boundary core structure in magnetic skyrmion lattice. *Sci. Adv.* **2**, e1501280–e1501280 (2016). DOI: 10.1126/sciadv.1501280

66. Azough, F., Kepaptsoglou, D., Ramasse, Q. M., Schaffer, B. & Freer, R. On the origin of nanochessboard superlattices in a-site-deficient ca-stabilized Nd₂/3TiO₃. *Chem. Mater.* **27**, 497–507 (2015). DOI: 10.1021/cm5036985
67. Ozdol, V. B., Gammer, C., Jin, X. G., Ercius, P., Ophus, C., Ciston, J. & Minor, A. M. Strain mapping at nanometer resolution using advanced nano-beam electron diffraction. *Appl. Phys. Lett.* **106**, (2015). DOI: 10.1063/1.4922994
68. Latychevskaia, T., Woods, C. R., Wang, Y. B., Holwill, M., Prestat, E., Haigh, S. J. & Novoselov, K. S. Convergent beam electron diffraction of multilayer Van der Waals structures. *Ultramicroscopy* **212**, 112976 (2020). DOI: 10.1016/j.ultramic.2020.112976
69. Voyles, P. M. & Muller, D. A. Fluctuation microscopy in the STEM. *Ultramicroscopy* **93**, 147–159 (2002). DOI: 10.1016/S0304-3991(02)00155-9