Microscopy and Microanalysis (2022), **28**, 1998–2008
doi:10.1017/S1431927622012259

Microscopy AND Microanalysis

Original Article

Analysis of Interpretable Data Representations for 4D-STEM Using Unsupervised Learning

Alexandra Bruefach[1], Colin Ophus[2] and Mary C. Scott[1,2]*

[1]Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA and [2]National Center for Electron Microscopy, Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Abstract

Understanding the structure of materials is crucial for engineering devices and materials with enhanced performance. Four-dimensional scanning transmission electron microscopy (4D-STEM) is capable of mapping nanometer-scale local crystallographic structure over micron-scale field of views. However, 4D-STEM datasets can contain tens of thousands of images from a wide variety of material structures, making it difficult to automate detection and classification of structures. Traditional automated analysis pipelines for 4D-STEM focus on supervised approaches, which require prior knowledge of the material structure and cannot describe anomalous or deviant structures. In this article, a pipeline for engineering 4D-STEM feature representations for unsupervised clustering using non-negative matrix factorization (NMF) is introduced. Each feature is evaluated using NMF and results are presented for both simulated and experimental data. It is shown that some data representations more reliably identify overlapping grains. Additionally, real space refinement is applied to identify spatially distinct sample regions, allowing for size and shape analysis to be performed. This work lays the foundation for improved analysis of nanoscale structural features in materials that deviate from expected crystallographic arrangement using 4D-STEM.

Key words: 4D-STEM, diffraction, machine learning, nanomaterials, scanning transmission electron microscopy

(Received 18 March 2022; revised 30 May 2022; accepted 27 June 2022)

Introduction

To design materials with desired physical behaviors, it is crucial to understand how local disorder and deviations from ideal crystallographic arrangements impact the functionality and properties of a material. It can be challenging to gain reliable measurements of local structural features at the nanometer scale, especially for low-dimensional materials with defective or disordered structures that can deviate significantly from a perfect crystal. While powder X-ray diffraction (PXRD) (Negishi et al., 2014) and other X-ray techniques (Yao et al., 2010; Duan et al., 2016) have been successfully applied to quantify the presence and solve the structure of various materials, mapping the heterogeneity and spatial distribution of the nanoscale structural components that make up a material can give more insight on the features that allow for enhanced performance (Greer, 2006).

Transmission electron microscopy (TEM) is well positioned to elucidate nanometer-scale material structure due to the ability to focus the electron probe down to the sub-nanometer scale (Ponce et al., 2021). Convergent beam electron diffraction (CBED) allows for a diffraction pattern to be collected after a converged electron probe interacts with a specimen. Due to recent advances in fast and direct electron detection (Pelz et al., 2022), four-dimensional scanning transmission electron microscopy (4D-STEM) has become a widely popular technique to probe materials' structure. The method consists of collecting a series of CBED patterns at multiple probe positions over a 2D scan space. The experimental 4D-STEM setup is shown in Figure 1a. At each probe position, a diffraction pattern is collected (Figs. 1b, 1c), which can be processed to create real space maps representing some component or structurally distinct region of the dataset, termed "virtual images" (Gammer et al., 2015). There are several data transformations that can be performed on each diffraction pattern to allow information to be more readily extracted (Ophus, 2019; Kacher et al., 2021). The polar elliptical transformation (Fig. 1e) is frequently applied to the Cartesian data (over the x and y probe coordinates) to visualize the data in polar coordinates as a function of the radial direction (r) and annular direction (θ). This representation makes the application of statistical measurements with regards to either r or θ straightforward (Kacher et al., 2021). From here, each pattern may be represented as a line profile, or 1D vector. For example, the summed intensities in each row of the polar elliptical transformation is known as the radial integral (Fig. 1f). Similar statistical measurements such as the average or standard deviation can be applied as a function of r or θ in polar space. For a more detailed explanation of 4D-STEM and the data representations, the reader is referred to Ophus (2019) and Kacher et al. (2021).

In previous work, 4D-STEM has been used to determine the spatial distribution of crystallographic orientations (Panova et al., 2019; Allen et al., 2021; Londono-Calderon et al., 2021), phases (Shukla et al., 2018), and relative strain (Mahr et al., 2021) within crystalline (Mukherjee et al., 2020), polycrystalline

*Corresponding author: Mary C. Scott, E-mail: mary.scott@berkeley.edu

Cite this article: Bruefach A, Ophus C, Scott MC (2022) Analysis of Interpretable Data Representations for 4D-STEM Using Unsupervised Learning. *Microsc Microanal* **28**, 1998–2008. doi:10.1017/S1431927622012259

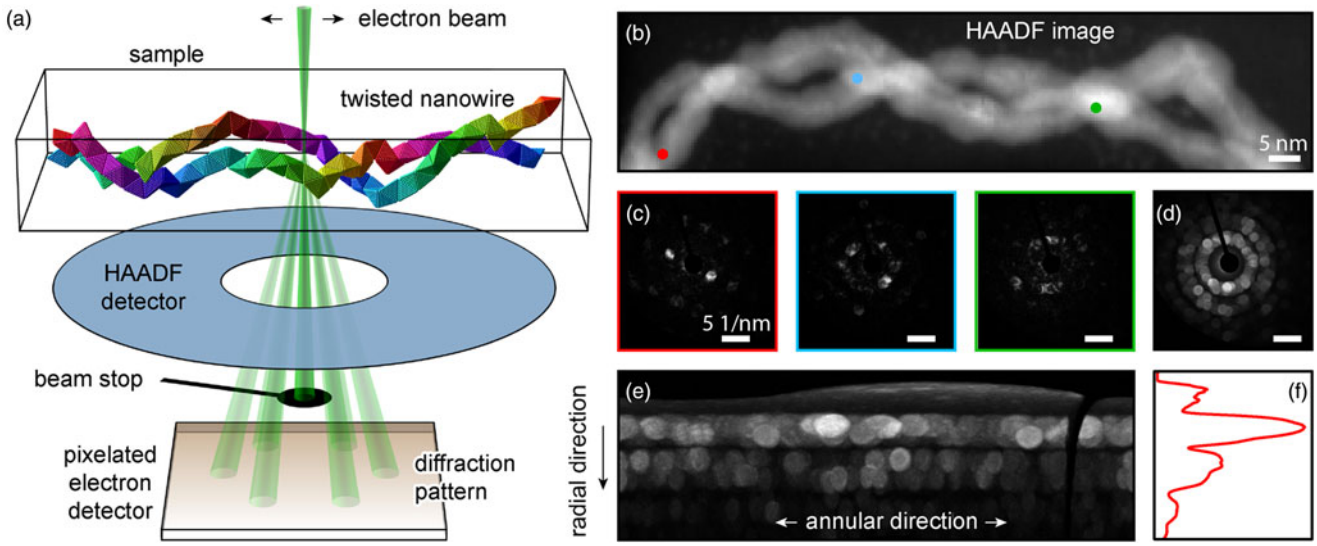

Fig. 1. 4D-STEM workflow and data representations. (**a**) Visual depiction of 4D-STEM experiment on twisted nanowires dataset. (**b**) High-angle annular dark-field (HAADF) image of the twisted nanowire and (**c**) patterns from select regions. (**d**) Cartesian, (**e**) polar-elliptical, and (**f**) radial Integral representations for the maximum CBED pattern from the 4D-STEM dataset. The radial direction (r) label holds true for (**e**) and (**f**), while the magnitude of the radial integral in (**f**) is the sum of the intensity along the annular direction (θ) from (**e**).

(Grimley et al., 2018), and amorphous materials (Pekin et al., 2019; Mu et al., 2021; Yang et al., 2021). The fields of image processing, data science, and machine learning have been applied harmoniously to discover trends within the 4D-STEM datasets (Rauch et al., 2010; Thati et al., 2015; Izadi et al., 2017; Shukla et al., 2018; Mehta et al., 2020; Zintler et al., 2020; Yuan et al., 2021; Cautaerts et al., 2022; Deng et al., 2022). However, these approaches all require prior knowledge of the crystal structure (s) present in the sample and can only compare the dataset of interest to crystallographic orientations of known or expected structures. Supervised learning has been used in the microscopy community, however, this requires generating labeled data by either simulating data (Munshi et al., 2022) or manually annotating experimental data (Groschner et al., 2021). Simulation is a viable option but may not capture structures that deviate from ideal crystals. Manual annotation is time-consuming for 4D-STEM, as each dataset may contain tens of thousands of diffraction patterns. Thus, unsupervised learning is a natural approach for 4D-STEM data analysis.

While there is some work focused on unsupervised learning to evaluate 4D-STEM data by either performing fuzzy C-means (Martineau et al., 2019), non-negative matrix factorization (NMF) (Uesugi et al., 2021) or applying hierarchical models to the raw images (Shi et al., 2021), none to our knowledge have shown utility in the analysis of 4D-STEM datasets containing polycrystalline or highly disordered structures at the single pattern level. Approaches based on the full dataset by nature contain a large number of features and therefore are prone to the curse of dimensionality (Bellman, 1957), which occurs due to the exponential increase in subspace volume with the addition of a dimension. The dimension of the classification subspace is determined by the number of input features. Workflows based on the full, high-dimensional dataset have only been successful on nearly ideal structures with little disorder or overlap among crystalline regions (Ma et al., 2020). Therefore, these techniques are limited in their ability to capture information in datasets with low signal or more structural complexity at the single pattern level.

There are only a few literature examples of 4D-STEM analysis of ultrasmall (<10 nm) nanomaterials (Bruma et al., 2016; DaSilva et al., 2020; Ophus et al., 2021), and none to our knowledge that area have been established without prior knowledge of the material structure. Since this class of materials often contain significant levels of disorder at the single pattern level compared with their bulk counterparts, conventional methods designed for bulk materials often fall short. A large barrier to widespread use of 4D-STEM for resolving the structure of ultrathin low-dimensional materials is the lack of progress in data analysis methods that can capture deviations from expected or ideal crystal structures.

Developing workflows with the ability to reduce the information contained in 4D-STEM datasets to a smaller feature set that still capture the key features that make the patterns distinct from one another is crucial for analysis via unsupervised learning. Dimensionality reduction allows for the largely complex dataset to be distilled down to a more manageable size (Guccione et al., 2021). Advances in feature representations for 4D-STEM pattern classification have been primarily limited to Bragg disks (BD) (Meng & Zuo, 2017; Pekin et al., 2017; Zeltmann et al., 2020). Recently, principal component analysis (PCA) and NMF were applied to cluster BD from polycrystalline gold nanoparticles into feature sets (Allen et al., 2021). In this work, it was shown that while both NMF and PCA have the ability to reasonably discern grains within nanoparticles, only NMF was able to output results that were directly interpretable as specific orientations due to the non-negativity constraint (Allen et al., 2021). This aligns with work in image analysis fields comparing NMF and PCA, where the sparse solution provided by NMF is more readily interpretable than the holistic output that PCA typically provides due to the presence of both additive and subtractive combinations (Lee and Seung, 1999; Guillamet et al., 2002).

While the BD representation has been successful in many applications in addition to those mentioned (Panova et al., 2016, 2019; Han et al., 2018), there are several shortcomings of this method. The first is that this method requires a highly accurate detection of BD, which is not always possible in low signal or

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

highly disordered datasets (Cooper et al., 2016; Pekin et al., 2017). Second, even with advancements in crystallographic orientation mapping, accurate mapping of structures that are off-zone axis is a challenge. Finally, this representation does not provide an avenue for differentiation between highly disordered regions from amorphous materials, as BD may not be robustly detected in either of these sample types. Thus, advances in inclusive featurization of CBED patterns could expand the utility of this method to a wider variety of materials. While other featurization approaches have been demonstrated (Ophus, 2019; Rauch & Veron, 2019; Mehta et al., 2020; Mu et al., 2021; Thornsen et al., 2021), methods to evaluate the information carried in these representations has not been reported.

This work addresses this knowledge gap by evaluating three engineered feature representations for CBED patterns within 4D-STEM datasets using the NMF approach described in prior work (Allen et al., 2021). The BD representation was analyzed due to the widespread interest and development of methods to precisely extract this feature from 4D-STEM data (Pekin et al., 2017; Han et al., 2018; Mukherjee et al., 2020; Zeltmann et al., 2020; Allen et al., 2021; Kacher et al., 2021; Yang et al., 2021). However, recovering information that is lost during the BD detection can provide an avenue for improved classification, especially for regions of a sample that may not have strong Bragg scattering. Virtual imaging, when applied directly to the raw patterns, has the ability to recover information that may have been lost during disk detection. Here, the angular average (AA) is introduced as a set of features based on mean virtual imaging of the raw data that does not assume that the pattern can be described purely from the detected BD. A statistical approach based on the variance, which is referred to as the radial variance (RV) throughout the article, is also analyzed to develop an understanding of how feature vectors based not on spatial but on statistical variations throughout the dataset capture information within the patterns.

An analysis of the information transferred in the BD, AA, and RV features extracted from three simulated 4D-STEM datasets of polycrystalline Ag films is performed, which provide insights on the application and limitations of each representation. The insights gained from the simulated data are then applied to an experimental dataset of a double helical polycrystalline nanowire containing Palladium, Gold, and Silver (Pd@AuAg NWs). The pipeline (Fig. 2) begins with the collection and calibration of a 4D-STEM dataset, then leads to the extraction of feature representations, and uses NMF to evaluate the data. Each engineered feature leads to interpretable results that relate to the structure of the underlying material, and specific use cases in which the AA and RV representations capture specific material structures better than the BD representation for specific sample types are discussed. This work sets the foundation for improved understanding of materials structure by expanding upon semi-automated data analysis methods in the microscopy community.

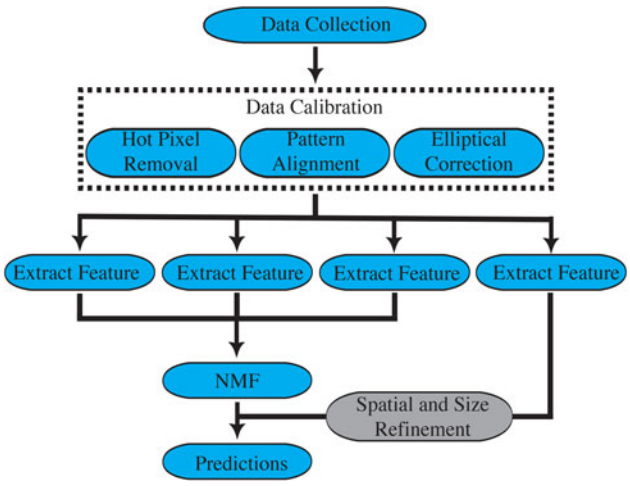

Fig. 2. Diagram of unsupervised learning workflow.

Materials and Methods

Ag Film 4D-STEM Simulation

The Ag films were created using a custom MATLAB code by randomly distributing 100 grain centers with random crystalline face-centered cubic (FCC) orientations in a 1,000 Å × 1,000 Å × 300 Å cell. Each grain consisted of an approximately spherical Wulff shape consisting of [001], [011], and [111] facets. Atoms which in any grain which were closer to another grain's origin were removed to prevent overlapping grain regions, with an additional 2 Å separation between grains. The structures were not energetically relaxed. The diameters of the grains in Ag1, Ag2, and Ag3 were restricted to three different sizes (35 Å × 41 Å × 40 Å in Ag1, 52 Å × 62 Å × 60 Å in Ag2, 70 Å × 82 Å × 80 Å in Ag3). The 4D-STEM simulations were performed over the first quarter of the cell (250 Å × 250 Å × 300 Å) using custom MATLAB scripts that implement the multislice algorithm (Cowley & Moodie, 1957) and methods defined by Kirkland (2020) and the plane wave reciprocal space interpolated scattering matrix (PRISM) algorithm (Ophus, 2017). The patterns were generated using an acceleration potential of 300 keV, a probe semiconvergence angle of 1.05 mrad, a 5 Å pixel size in real space, and a 0.01 $Å^{-1}$ pixel size in reciprocal space. After the 4D-STEM simulation was performed, a Gaussian kernel was applied over the real space probe scanning dimensions with a standard deviation of 2 pixels to replicate to account for a non-zero source size and coherence limitations, which is expected to be the dominant aberration (Dwyer et al., 2010). There were 31, 36, and 42 distinct grains in the field of view for the Ag1, Ag2, and Ag3 datasets, respectively.

Nanowire Synthesis

Polyvinylpyrollidone (PVP, 40,000 MW), dimethylformamide (DMF), silver nitrate ($AgNO_3$), and Gold(III) chloride trihydrate ($HAuCl_4 \cdot 3H_2O$, 49.0% Au Basis) were purchased from Sigma-Aldrich. When specified, solutions were prepared in deionized (DI) water (resistivity >18 MΩ/cm).

Samples were prepared as described in a previous study with slight variations (Wang et al., 2011). Solutions of PVP (5.55 mg, 500 mM) in DMF (100 mL), $HAuCl_4$ (0.394 mg, 50 mM) in DMF (20 μL), $AgNO_3$ (0.170 mg, 50 mM) in DI water (20 μL), and L-ascorbic acid (7.045 mg, 400 mM) in DI water (100 μL) were prepared ahead of time. The reaction solution was created in a 4 mL vial (washed 3 × with water and acetone) by mixing DMF (800 μL), PVP solution (100 μL), $HAuCl_4$ solution (20 μL), and $AgNO_3$ (20 μL) solution. The reaction mixture was then vortexed for 2 s. The L-ascorbic acid solution (100 μL) was added drop-wise to the mixture while gently swirling. At this point, the reaction solution changed color from pale yellow to clear. The solution was left at room temperature for 7 days and appeared light brown/purple after incubation.

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

To create the Pd shell layer on the underlying ultrathin AuAg NWs, solutions of L-ascorbic acid (0.211 mg, 1.875 mM) and H_2PdCl_2 (0.03 mg, 2 mM) were prepared in DI water ahead of time. In a 4 mL vial (3 × washed with DI Water/acetone), an unpurified aliquot of the AuAg reacted solution (50 μL) was added to the L-ascorbic acid solution (640 μL). Finally, the H_2PdCl_4 (60 μL) was added and allowed to incubate for at least 30 min. The reaction solution was then purified by centrifuging the product at 7,500 rpm for 4 min. The supernatant was decanted, the reaction was rinsed with DI water, and the centrifugation procedure was repeated three times. The final purified product was re-dispersed in DI water. The samples were prepared for imaging by depositing 10 μL of purified NW solution on 400 mesh formvar/ultrathin carbon grids.

Data Collection

The dataset was collected from a double aberration-corrected modified FEI Titan 80-300 microscope (TEAM I at the National Center for Electron Microscopy within the Molecular Foundry at Lawrence Berkeley National Laboratory) equipped with a Gatan K3 detector set to collect with a pixel time of 0.05 s, 4 × bin, and full range. The accelerating voltage was set to 300 kV with a spot size of 6. A 10 μm aperture was used (semiconvergence angle of 2 mrad) with a camera length of 1.05 m. A step size of 5 Å was used, with a total real space sampling of 231 positions in x and 58 images in y, leading to a dataset with 13,398 CBED patterns. The beamstop was used because it was found that the scattering signal to be significantly stronger for this sample with the beamstop in.

Data Calibration

The 4D-STEM dataset was imported and preprocessed using the publicly available methods in py4DSTEM (Savitzky et al., 2021). First, the experiemental CBED patterns were binned by 2, leading to the dimension of 432 × 432 in diffraction space. The dataset was then cropped to 220 × 50 in real space (11,000 diffraction images) to reduce the amount of images containing only the carbon film, leading to an analysis dataset of dimensions 220 × 50 × 432 × 432. In the experimental dataset, hot pixels were filtered in diffraction space by detecting any pixels that were 0.55 larger than the local sorted intensity values and replacing these with their 3 × 3 local median.

The diffraction patterns were aligned and any detected elliptical distortions were corrected for by fitting an ellipse to the Bragg Vector Map, which was made after detecting the disks in the patterns and distributing them into an image. Both the image alignments and the elliptical distortion corrections were applied to all feature representations.

Feature Extraction

Bragg Disks

There are a number of workflows that can be used to detect BD (Pekin et al., 2017; Han et al., 2018). Here, the correlative disk detection method available in py4DSTEM is followed (Savitzky et al., 2021). Briefly, this method consists of detecting the local maxima of the cross-correlation between the diffraction image and the probe kernel. False positive disks surrounding the beam stop in the experimental data were removed using a mask surrounding the beam stop. Elliptical distortions were detected and corrected in both the simulated and experimental data. The aligned and elliptically corrected BD were rasterized as described by Allen et al. (2021) based on their detected locations using a bin size of 3 × 3, further reducing the diffraction space dimensions to 84 × 84 (7,056 pixels) in the simulated dataset and 144 × 144 (20,736 pixels) in the experimental dataset. Figures 3a, 3b show an example of the mean virtual image (VI) and the maximum diffraction pattern of a simulated dataset. Figure 3c reveals the 2D Bragg disk position histogram, known as the "Bragg Vector Map," across the entire simulated dataset. Figures 3d, 3e demonstrate the process of transforming an individual pattern, while Figure 3f shows what an individual pattern looks like after the complete transformation using the BD described in a prior study (Allen et al., 2021). While other methods of representing BD have been established, such as adaptive binning based on the Voronoi diagram (Savitzky et al., 2021), this Cartesian gridding representation was chosen because it can be streamlined to analyze multiple datasets in which different Bragg reflections appear.

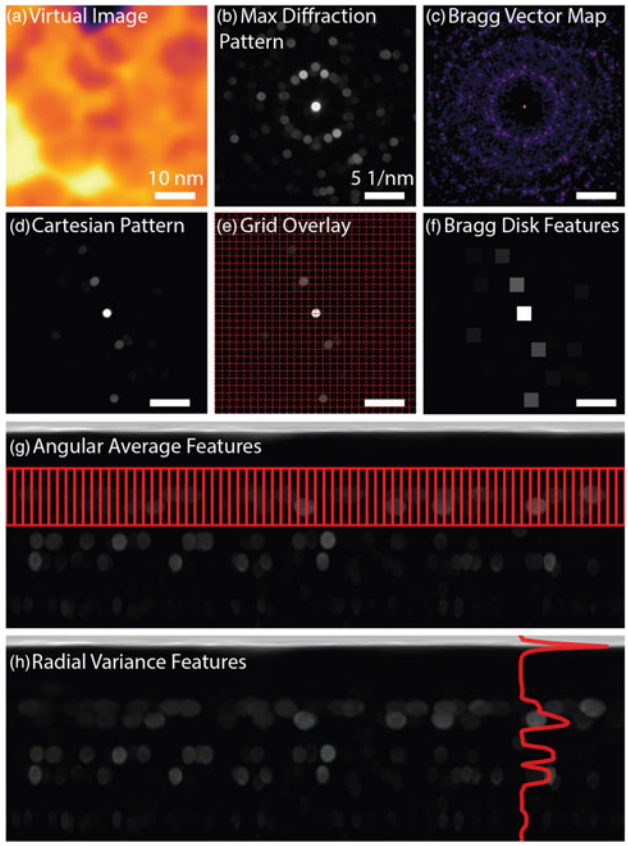

Fig. 3. Feature representations on simulated Ag polycrystalline film. (**a**) Virtual image created by averaging the intensity of each diffraction pattern. (**b**) Maximum diffraction pattern created by taking the maximum intensity in diffraction space. (**c**) Bragg vector map, which indicates the positions and intensities of the detected BD. Color scale for (**a**) and (**c**) range from low (black) to high (yellow) intensity. (**d**–**f**) Workflow for preparing detected disks into features for unsupervised learning. (**g**) Visualization of the AA featurization over the first detected radial ring in the maximum diffraction pattern. (**h**) Representation of the RV profile overlaid on the maximum diffraction pattern. Each colored box in (**e**) and (**g**) represents a region in which a single input feature is generated, while the red line in (**h**) represents one input feature per row in the polar elliptically transformed image.

Angular Averaging

By taking the average intensity every 5° over each radial ring within the pattern, information related to the spatial distribution

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

of intensities can be recovered. This approach effectively applies a series of mean virtual images as inputs, as shown in Figure 3g. Each colored box in Figure 3 represents a region in which the pixels were averaged to create a single input feature. While virtual imaging has been commonly utilized to understand 4D-STEM datasets (Shukla et al., 2018; Mehta et al., 2020; Thornsen et al., 2021), virtual apertures are traditionally applied by identifying reflections of interest within the maximum diffraction pattern. Instead, this representation is expanded by creating virtual apertures that cover the entirety of the detected radial rings, reducing information loss from biased placement of virtual apertures. The AA was extracted by determining the positions of the radial rings by finding the local minima in the radial integral of the maximum diffraction pattern, then averaging every five columns for each pattern within this range. The output of this step was a 2D array of dimensions $10{,}000 \times 288$ for the simulated dataset, which had four rings present, and $11{,}000 \times 142$ for the experimental dataset, which had two rings present. Two features were removed per ring from the experimental dataset due to the presence of the beamstop in the region.

Radial Variance

In addition to spatial features, statistical metrics that can numerically capture the degree of disorder or variation present in a particular radial ring may also allow for more robust separation of crystalline from noncrystalline regions within a dataset. The RV is a statistical metric based on variance (equation (1)) that measures the numerical deviation of the individual values at each θ from the average value in each radial index:

$$\mathrm{var}(x[i]) = \frac{\sum_{\theta=1}^{\theta=n} (x_\theta[i] - \bar{x}[i])}{n}, \tag{1}$$

Here $x_0[i]$ is the set of intensities at the radial index i, θ is the angular distribution, which spans from 1 to n. In this case, the span of θ is from 1° to 360°. The RV value will be larger in more crystalline patterns where there are large intensity fluctuations in a radial ring, and effectively zero in perfectly amorphous materials. Methods based off of intensity variances within diffraction patterns have been frequently applied to discern trends in medium range order in amorphous materials using a technique called fluctuation electron microscopy (FEM) (Treacy & Gibson, 1996). Although the RV presented in this article is not the same as those used in FEM, this featurization consolidates the entire 2D pattern into a single line profile, drastically decreasing the dimensionality.

To extract the RV feature, the polar elliptical transformation was first applied to the cartesian diffraction patterns. The variance was then calculated over each row of pixels, which corresponds to a radial distance from the incident beam. For the experimental dataset, the beamstop was masked and these pixels were not included in the variance calculation. The variance profile was cropped to remove effects from the center beam and detector edges. The output of this approach was a 2D array of dimension $10{,}000 \times 80$ for the simulated dataset and $11{,}000 \times 100$ for the experimental dataset. This feature set is unable to capture distinct high or low intensity regions across the θ direction and is rather only sensitive to the intensity variation along θ over all r values. The RV line profile of the maximum diffraction pattern has been overlaid along the polar elliptically transformed maximum diffraction pattern in Figure 3h.

Feature Preprocessing

All of the feature vectors were scaled in a range from 0 to 1 using the MinMaxScaler in scikit-learn (Pedregosa et al., 2011) which follows the equation (2):

$$X_{\mathrm{scaled}} = \frac{X - X_{\mathrm{min}}}{X_{\mathrm{max}} - X_{\mathrm{min}}} * (\,\mathrm{max} - \mathrm{min}\,) + \mathrm{min}, \tag{2}$$

where X_{scaled} is the scaled feature vector, X is the value at each position, X_{min} and X_{max} are the minimum and maximum value of the feature vector, respectively, and max, min are the range in which the features are scaled (from 0 to 1 in this work).

Dimensionality Reduction on Single Features

NMF solves the matrix factorization problem equation (3) with a non-negativity constraint:

$$\|V - WH\|_F, \quad \text{subject to } W \geq 0,\ H \geq 0. \tag{3}$$

NMF effectively reduces the input matrix, V, into a series of weighted linear combinations of the columns in V. If V is an $n \times m$ matrix, W will become $n \times c$, while H will become $c \times m$, where c is the number of chosen components. By convention, the columns of the W matrix become the basis feature sets, while the H matrix contains the weights, or contributions, of each basis set to the individual patterns (Lee & Seung, 1999).

Iterative NMF was performed to cluster the individual features into regions of similarity. This was performed as described in a previous report (Allen et al., 2021), where PCA was first performed to determine the optimal number of clusters (determined from the Scree plot). The NMF clustering cycle to reduce the dataset into at least five times the number of optimal clusters determined by PCA. Each cycle, features with correlations above a predefined threshold (typically between 0.15 and 0.5) were merged and a new cycle began. The selected merge threshold was highly dependent on the input feature type, with RV (0.3–0.5) and AA (0.4–0.5) representations typically having a higher threshold than the BD (0.18–0.25) representations. It was found that lower merge threshold values for AA and RV only led to one feature representation, which is why larger values were selected. NMF was performed until no features with correlations above this threshold were found. This was repeated 25 times with different random seeds, and the model with the smallest weighted reconstruction error was selected. The parameters (including initial cluster number and merge threshold) and runtimes of each model are presented in Supplementary Table S1.

Spatial and Size Refinement and Measurements

Since each pattern within a 4D-STEM dataset is located in a real space position, a spatial and grain size refinement was performed to separate spatially distinct grains and remove grains below a specified size after clustering (Supplementary Fig. S1). This allowed for additional analysis of the size, diameter, and shape of detected grains. Spatial and size separation was performed by first transforming each column of the NMF output components into 2D arrays of the size of the real space scans (200×200 in the simulated dataset, 220×50 in the experimental dataset). The yen threshold function from scikit-image (van der Walt et al., 2014) was applied to each image produced from the

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

NMF output, which allowed for the identification of spatially independent clusters within the cluster image. All values below the threshold in each image were set to 0. These spatially independent components were filtered out by size (labeled regions less than 25 or 10 pixels were removed for the simulated and experimental data, respectively). The clusters were separated based on the detected labeled region. The area and diameters of regions labeled were measured using the regionprops function in scikit-image (van der Walt et al., 2014). The area was calculated by summing the number of pixels in each cluster. The max diameter was determined using the Feret method (Walton, 1948). While the spatial separation step allows for a workflow to analyze cluster shape and size, this step prevents spatially separated clusters of similar orientation from having the same cluster label.

Scoring Metrics

The labels were generated for the simulated datasets based on the Ag structures at each probe position. For each independent grain, a single binary image was created.

The area-based precision metric (precision_A) was calculated by first determining where in the field of view grains were located from the binary label images. The labels for precision_A were generated by setting the pixels that contained a grain of any identity to 1, and pixels with no grain to zero. Thus, the precision_A true label set only contains on image. The cluster maps were similarly binarized based on whether or not a grain of any identity was detected. The precision_A can simply be thought of as the precision at which the model can detect a crystalline region over a vacuum region. The precision_A score was calculated from each binary cluster image based on the formula for precision, shown in equation (4):

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}. \quad (4)$$

Other area-based metrics, such as the True Positive Rate (TPR_A), True Negative Rate (TNR_A), False Positive Rate (FPR_A), and False Negative Rate (FNR_A) were similarly calculated based on the steps described for precision_A.

The second metric discussed is the true positive rate (TPR_G) for each cluster that was detected in the AA and BD featurizations. This is calculated by finding the best match grain in the cluster labels and calculating the TPR for each match, which is shown in equation (5):

$$\text{TPR} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}. \quad (5)$$

To determine the best match grain, the sum of the weights were determined over the area in which each label was 1. After the best match was found, the TPR, True Negative Rate (TNR), Type I Error (False Positives), and Type II Error (False Negatives) were calculated for each pair.

For the experimental data, there were no labels produced. Instead, the cluster dissimilarity score was determined for each cluster by first calculating the weighted average cluster diffraction pattern of each cluster after spatial separation of the NMF results. The cluster pattern was scaled to 0–1 using the using the MinMaxScaler in scikit-learn (Pedregosa et al., 2011) (equation (2)) and the sum of differences between the weighted cluster pattern and scaled contributing cluster patterns was calculated. This is written mathematically in equation (6):

$$\text{dissimilarity}(\text{cluster}[i]) = \frac{\sum_{j=1}^{j=n} (w_j[i] * DP_j[i] - \overline{DP}[i])}{\sum_{j=1}^{j=n} w_j[i]}, \quad (6)$$

where ($DP_j[i]$) and ($\overline{DP}[i]$) are the contributing and weighted average patterns for each cluster, respectively. The weights per pattern (j) within a cluster (i) are represented by $w_j[i]$. The dissimilarity score is more useful in discerning relative performance of different features on the same dataset, as the score is highly dependent on the noise within the dataset. Thus, dissimilarity score for simulated data will be significantly lower than for experimental data, although the trend may still provide insight on performance. Since multiple clusters were present in each dataset, the distribution of dissimilarities were visualized in a box plot for the experimental data.

Results and Discussion

Silver Film Simulations

NMF was applied independently to RV, AA, and BD feature sets extracted from three simulated polycrystalline Ag film datasets with different sized grains, and thus varying extent of overlap between adjacent grains. Figure 4a shows the cluster labels for the three simulated datasets, with the first row having almost no overlap between neighboring grains and the last row having significant overlap. We will refer to the dataset in the top row of Figure 4a as Ag1, the middle row as Ag2, and the bottom row as Ag3. The results of applying NMF to the RV, AA, and BD features for the three datasets are shown in the panels of Figures 4b, 4c, and 4d, respectively. The colors of the clusters are indicative of the order in which the clusters are stored, but not necessarily the cluster match across the images.

As the overall area of Ag grains increase in the dataset, the precision_A (shown in Figure 4e for the three featurizations and datasets) is expected to increase because there are less patterns that do not contain a Ag grain, thus lowering the possibility of false positives. The precision_A is the only metric discussed for the RV feature, since it lacks the sensitivity to 2D spatial features within the patterns thus does not allow for accurate orientation or phase mapping. It is most useful to consider the RV feature as being sensitive to extent of crystallinity rather than specific orientations. The RV feature set consistently underestimates the total area containing grains in all three simulations, leading to a high FNR_A (70% for Ag1 and Ag2, 72% for Ag3). However, Figure 4e reveals that the RV is more precise at detecting crystalline regions in Ag1 than the other two featurizations for this dataset. Thus, when the model based on RV feature predicts a Ag region in a dataset with more sparsely distributed crystalline regions, it is most likely to contain a Ag grain. The AA had the lowest area-based precision for Ag1 (75.3%). The clusters determined from the AA show expansion relative to the cluster labels, indicating that while the precision_A is lowest, the FPR_A is lower than in the RV representation. This could be a result of the AA feature representation responding to the Gaussian blurring performed on the simulated data. A full list of area-based metrics are presented in Supplementary Table S2.

For applications in which the crystalline regions in the dataset are expected to be sparse and it is crucial to avoid false positive

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

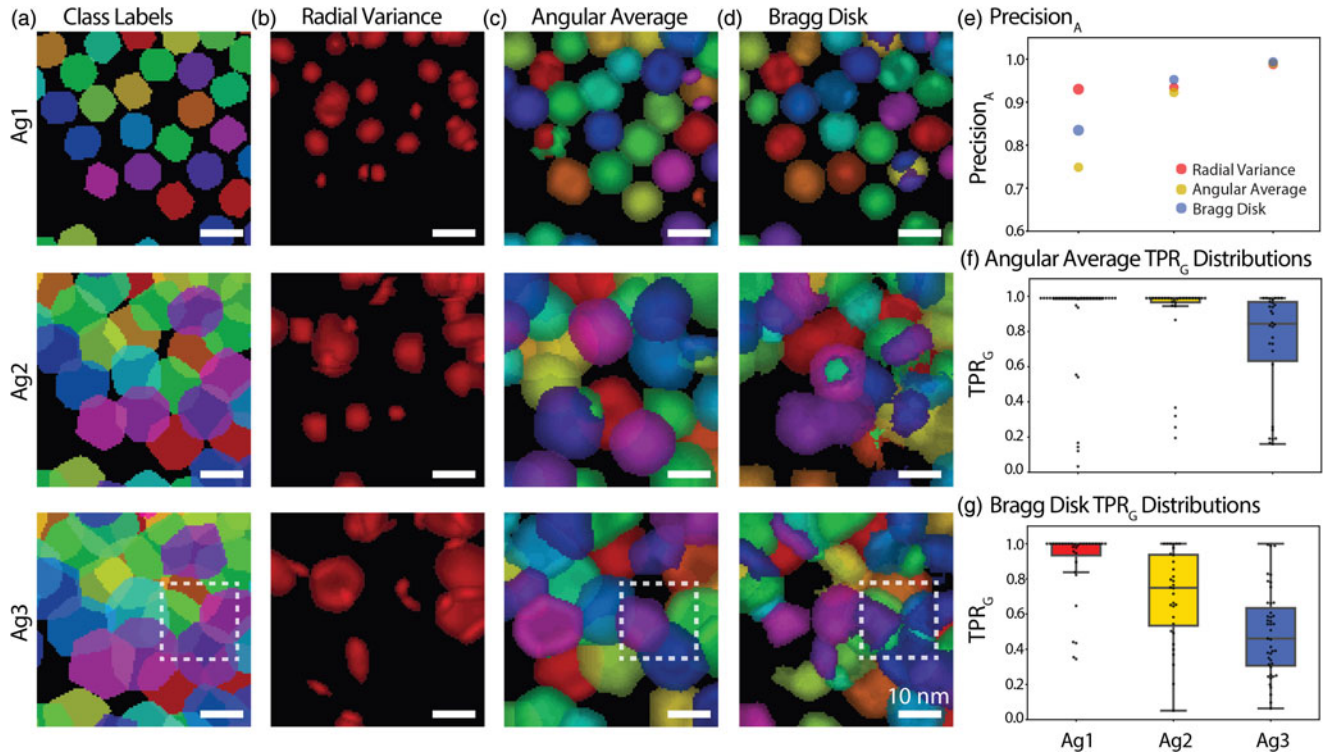

Fig. 4. Results of NMF applied to simulated Ag Films with varying grain sizes and extents of overlap. Colors indicate distinct labeled or detected clusters. (**a**) Cluster labels for the Ag1 (row 1), Ag2 (row 2), and Ag3 (row 3) simulated datasets. (**b**) Cluster maps created from applying NMF to the (**b**) RV, (**c**) AA, and (**d**) BD features. The colors of the clusters are indicative of the order in which the clusters are stored, but not necessarily the cluster match across the images in (**a**), (**b**), (**c**), and (**d**). (**e**) Precision$_A$ of the nine predictions. (**f**) TPR$_G$ distribution of the grain matches for the clusters determined by AA, which retain the highest TPR$_G$ as extent of overlap increases. (**g**) TPR$_G$ distribution of the grain matches for the clusters determined by the BD, which has a deteriorating TPR$_G$ as extent of overlap increases. The white dotted box in the bottom row (Ag3) highlights a region in which the BD feature overfits to overlapping grains, while the AA feature performs well.

detection, RV may outperform other representations. One example of an application in which the RV may be useful is detecting small crystalline precipitates in an amorphous matrix, such as confirming the presence of crystalline regions in precipitation hardened amorphous alloys similar to those reported in Roy et al. (2009). This is underpinned by the fact that the speed of this model is faster than the model based on the BD, which has the next highest precision. However, the application of the RV feature is somewhat limited to specific cases in which high variance is expected in the crystalline regions. Models based on the RV feature will not be suitable for orientation mapping since variance across patterns is not consistent among all crystalline regions and it is not sensitive to specific grains. Additionally, as the extent of grain overlap increases, the variance along a radial ring is expected to decrease due to the increased number of reflections in the ring, thus it is not well suited for situations in which overlapping grains are expected to be prevalent. As the overall area covered by the Ag grains increases, the precision$_A$ of the AA and BD features improves, which also reduces the use cases for the RV feature.

Figure 4f shows the TPR$_G$ distribution for the AA feature. The TPR$_G$ was relatively high for all of the detected clusters using the AA feature in Ag1, with an average value of 86.4%. Furthermore, this representation retains the highest TPR$_G$ as the extent of grain overlap increases, holding an average value of 89.3% for Ag2 and 73.6% for Ag3. As the overall area containing crystalline regions increases, the precision$_A$ for the AA feature also increases (92.7% for Ag2 and 99.5% for Ag3), further supporting the use of this representation over the BD representation for datasets expected to contain overlapping grains.

The BD representation outperforms the AA representation for mapping regions of distinct grains when the clusters have little overlap as both the precision$_A$ (83.9%) and the average TPR$_G$ (89.5%) are higher for the BD representation in Ag1. However, it is observable from the cluster maps (Fig. 4d) that this performance is not retained as extent of grain overlap increases. The cluster maps indicate that the BD representation is prone to overfitting to the regions with overlapping grains (see white dotted box in the bottom row of Fig. 4). These regions are consistently detected as independent from the parent grains instead of being allocated into multiple clusters. This lowers the average TPR$_G$ for the model based on the BD features (Fig. 4g), which decreased to 70.2% for Ag2, and to 48.8% for Ag3. Thus, for datasets in which grains are expected to overlap, the AA feature is expected to outperform the BD representation in detecting distinct grains. This can be more readily observed in maps of the number of detected clusters per pattern (Supplementary Fig. S3). For applications in which it is desired to map the size and shape of overlapping grains, the AA is likely to be more reliable than the BD representation. Since the BD feature tends to overfit regions of overlapping grains, clusters detected from the BD feature may relate to an interface or boundary between neighboring grains or orientations within a dataset. Thus, in applications where it is desirable to analyze overlapping regions or boundaries independently from the parent regions, such as in some strain mapping applications, the BD feature may be applied, even if there is

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

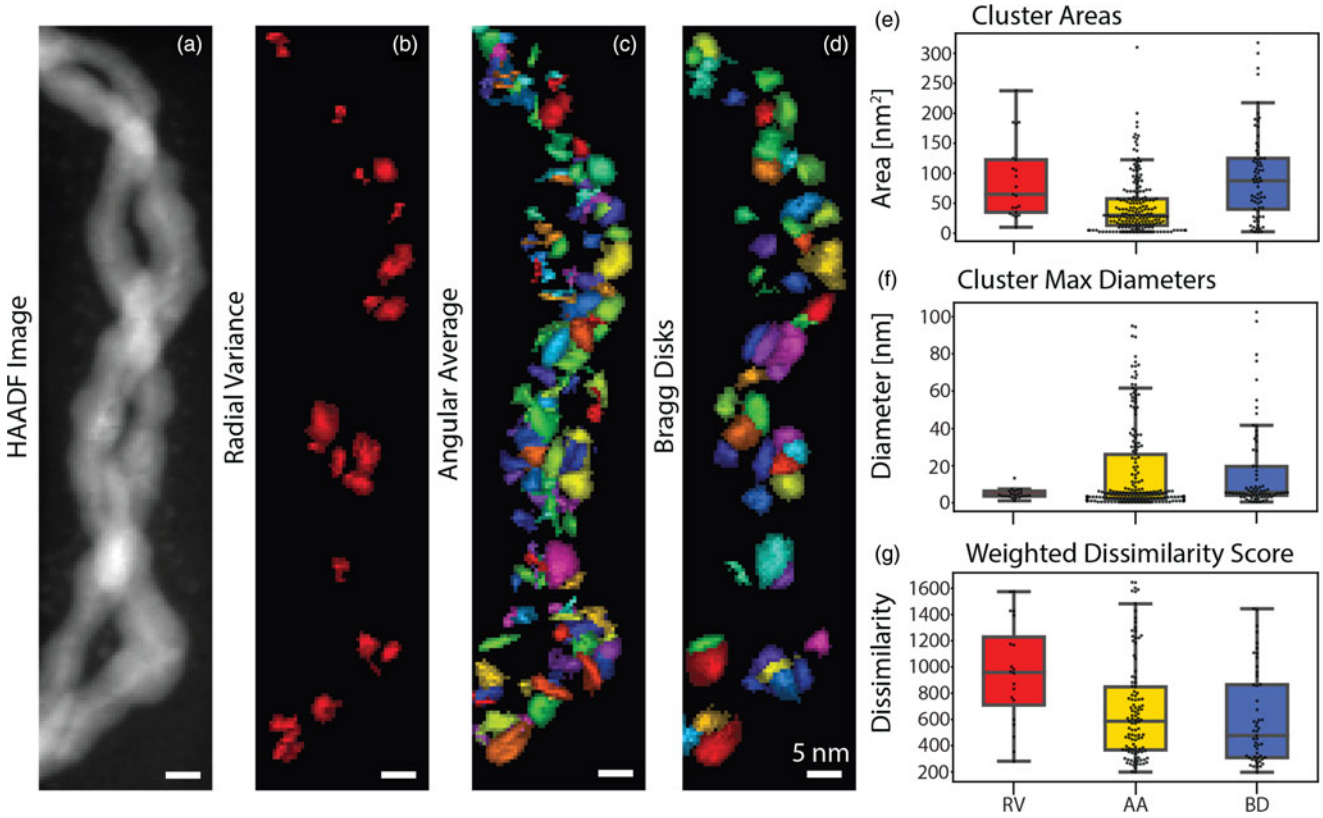

Fig. 5. Experimental dataset and results of unsupervised learning performed on individual feature representations. (**a**) Simultaneous high-angle annular dark-field (HAADF) image of scan region. Clusters from the models based on the (**b**) RV, (**c**) AA, and (**d**) BD. Colors indicate distinct detected clusters. Box plots of the (**e**) cluster areas (averages 87.9, 45.3, and 95.0 nm^2 for RV, AA, and BD), (**f**) cluster max diameters (averages 4.6, 17.1, and 16.8 nm for RV, AA, and BD), and (**d**) weighted dissimilarity score. (Averages 955.5, 676.2, and 605.9 for RV, AA, and BD).

significant grain overlap. It should be noted that the average TNR from the grain matches TPR_G is at least 98% for all of the applied featurizations (Supplementary Table S3), indicating that the model is sensitive to specific regions. A full list of the overall grain-based average rates across all datasets are reported in Supplementary Table S4. The overall average TPR_G across all three datasets was 83.2% for AA and 67.4% for the BD representation, indicating that if no prior knowledge about the extent of grain overlap or disorder is known about the sample beforehand, the AA representation may give more reliable results. The dissimilarity score (reported in Supplementary Table S3) was also calculated for all featurizations and it was found that the average score was typically higher for the RV feature (17.0 for Ag1, 30.6 for Ag2, and 58.1 for Ag3) than for the AA (12.2 for Ag1, 19.0 for Ag2, and 34.4 for Ag3) and BD (12.6 for Ag1, 21.7 for Ag2, and 32.0 for Ag3) in each dataset. The AA and BD dissimilarity scores were similar, however it is likely that the dissimilarity score for the AA in Ag3 surpasses the BD value due to the inclusion of overlapping regions within a grain. Since may of the overlapping regions in Ag3 from the BD feature are detected in their own clusters, the dissimilarity score will be lower since these will be self similar. Although the AA had high TPR_G for the grains that were detected, there were still grains that were not detected using the AA and BD feature representations. It can be observed from the cluster maps that both the AA and BD feature are missing some clusters on the bottom region of the field of view and in between neighboring grains in datasets Ag2 and Ag3. The detection of grains overall may be limited by grain orientation or number of observations in which these grains are the only grain present.

Pd@AuAg Twisted Nanowires

The featurization and machine learning workflow was also applied to an experimental dataset of a colloidally synthesized Pd@AuAg double helical NW. The final twisted NWs typically have a diameter of 5–10 nm and contain large amounts of structural disorder, making them an ideal candidate for testing novel featurization protocols. Additionally, it is proposed that these structures may contain icosahedral packing (Velazquez-Salazar et al., 2011; Wang et al., 2011), which is challenging to discern using conventional 4D-STEM analysis methods due to the grains overlapping in specific orientations. The feature representations for this dataset are shown in Supplementary Figure S4.

The results of applying iterative NMF to the dataset are revealed in Figure 5. Figure 5a shows the HAADF image of the analyzed region, while Figures 5b, 5c, and 5d show the results of applying iterative NMF to the RV, AA, and BD representations, respectively. Analysis reveals that the model based on the RV feature recovers the least information, while the model based on the AA feature recovers information that is missed by both the RV and BD representations. The AA feature was able to best recover information from patterns that are in the bottom left-hand corner of the NW.

As was revealed in the simulated data, the RV provides an incomplete understanding of the sample. There are only five

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

clusters initially detected using the RV feature, and many of the regions within each cluster are distributed broadly throughout the structure. With spatial and size refinement, the number of clusters increased to 20, which is approximately one fifth of the number detected in the AA (109) representation and approximately one half of the number detected by the BD (51) representation after spatial and size refinement. This result substantiates the earlier claim that RV is not an appropriate representation for datasets in which multiple crystalline clusters must be detected and distinguished from each other.

After spatial and size refinement, analysis on the cluster areas and diameters was performed. The average cluster areas (Fig. 5e) and max cluster diameters (Fig. 5f) detected were 87.9, 45.3, and 95.0 nm^2 and 4.6, 17.1, and 16.8 nm for the RV, AA, and BD representations, respectively. Based on these values, it is likely that the clusters detected are based on signal from the NW shell rather than the core, as the core is expected to have grain sizes with diameters on the order of 2 nm (Wang et al., 2011). This is expected as the shell is much thicker and likely more ordered than the core structure (Wang et al., 2011).

The RV detects clusters with a similar area distribution as the BD representation, but much narrower distribution of max diameters. The model based on the AA detects clusters with smaller areas, but a wider distribution of max diameters. It is observed that the clusters detected from the AA feature set are elongated relative to the other two feature sets. The elongated shape discerned by the AA feature may be a result of an oriented attachment growth mechanism, where the AA feature classifies the similar orientations within the same cluster while the BD feature clusters these regions distinctly. Nanowire growth via oriented attachment has been reported in these material systems at similar length scales (Wang et al., 2014, 2016).

The dissimilarity score distribution was also measured for each cluster (Fig. 5g). This is a relative, unitless metric. A larger magnitude dissimilarity score can be interpreted as patterns in a cluster deviating more highly from the average, while a smaller score can be interpreted as the patterns in the cluster being more self-similar. For datasets without labels, only the self-similarity within a cluster can be used to determine the relative accuracy of that cluster. The average dissimilarity scores were 955.5, 676.2, and 605.9 for the RV, AA, and BD representations, respectively. The weighted dissimilarity for the RV is larger than both the AA and BD values providing further indication that this featurization workflow does not have the ability to discern individual grains but rather crystalline areas. These trends are similar to what was calculated from the simulated results. The AA and BD have a similar distribution of weighted dissimilarity scores, indicating that either of these representations may be reasonable for this dataset. Since previous work has described multiple overlapping grains within these NWs (Ophus et al., 2021; Wang et al., 2011), the AA feature may provide better overall measurements of the cluster areas and diameters.

As was observed in the simulated dataset, the AA feature is able to capture more overall regions of overlapping grains than the BD (Supplementary Fig. S5). Thus, it is likely that the average weighted cluster dissimilarity score for the AA is larger than the BD because the AA clusters are more likely to contain overlapping regions. While this leads to more reliable measurement of grain shape and size in the AA clusters, the extent of overlap between grains may make it difficult to discern the parent grain structure from the weighted cluster pattern itself. In order to directly interpret the structure within these grains, it may be beneficial to use only the patterns that are detected in a single cluster to represent and interpret the full region. Categorization of the detected clusters into crystallographic orientations or phases can be completed by combining the NMF clustering workflow reported in this work with the ACOM workflow presented by Ophus et al. (2021), for example, by consolidating the BD detected throughout the dataset into sets based on the detected clusters. The performance on experimental data may be validated by comparing the clusters detected through these methods and the results of performing the full ACOM method.

Conclusion

In this work, a pipeline for engineering and analyzing features for 4D-STEM data was presented. The features analyzed are representative of different types of information present in electron diffraction images and the application of these features is underpinned by analysis of both simulated and experimental data. It was revealed that the RV feature has the highest precision in detecting crystalline regions in a Ag simulation with sparsely distributed grains, but is not able to detect regions of specific orientations. The AA feature is most robust at detecting independent grains when there is significant grain overlap, yet the BD representation still outperforms the AA when there is no grain overlap. These featurization protocols were applied to an experimental dataset of a Pd@AuAg NW and it was demonstrated that the AA featurization allowed for more grains within the NW to be detected. However, the BD representation still provided valuable insight on distinct regions within the experimental dataset.

This work opens the door to more complete classification of disordered and defect-rich structures, and we have shown the utility of our workflow for ultrathin double helical NWs with large amounts of structural diversity. There is an opportunity to further develop this method by exploring new feature representations or applying supervised methods to the cluster patterns in order to identify the crystallographic structure within the region.

Supplementary Material. To view supplementary material for this article, please visit https://doi.org/10.1017/10.1017/S1431927622012259.

Acknowledgments. AB and py4DSTEM development are supported by the Toyota Research Institute. Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research was partially supported by the National Science Foundation under award number 1848079.

Conflict of Interest. The authors declare that they have no competing interest.

References

Allen FI, Pekin TC, Persaud A, Rozeveld SJ, Meyers GF, Ciston J, Ophus C & Minor AM (2021). Fast grain mapping with sub-nanometer resolution using 4D-STEM with grain classification by principal component analysis and non-negative matrix factorization. *Microsc Microanal* **1**, 1–10.

Bellman RE (1957). *Dynamic Programming*. Princeton, New Jersey: Princeton University Press.

Bruma A, Santiago U, Alducin D, Plascencia Billa G, Whetten RL, Ponce A, Mariscal M & Jose-Yacaman M (2016). Structure determination of superatom metallic clusters using rapid scanning electron diffraction. *J Phys Chem C* **120**, 1092–1908.

Cautaerts N, Crout P, Anes HW, Prestat E, Jeong J, Dehm G & Liebscher CH (2022). Free, flexible and fast: Orientation mapping using the multi-

core and GPU-accelerated template matching capabilities in the python-based open source 4D-STEM analysis toolbox Pyxem. *Ultramicroscopy* **237**, 113517.

Cooper D, Denneulin T, Bernier N, Beche A & Rouviere J-L (2016). Strain mapping of semiconductor specimens with nm-scale resolution in a transmission electron microscope. *Micron* **80**, 145–165.

Cowley JM & Moodie AF (1957). The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Crystallogr* **10**, 609–619.

DaSilva JC, Smeaton MA, Dunbar TA, Xu Y, Balazs DM, Kourkoutis LF & Hanrath T (2020). Mechanistic insights into superlattice transformation at a single nanocrystal level using nanobeam electron diffraction. *Nano Lett* **20**, 5267–5274.

Deng HD, Zhao H, Jin NL, Hughes L, Savitzsky B, Ophus C, Fraggedakis D, Borbely A, Yu Y-S, Lomeli E, Yan R, Liu J, Shapiro DA, Cai W, Bazant MZ, Minor AM & Chueh WC (2022). Correlative image learning of chemo-mechanics in phase-transforming solids *Nat Mater* **21**, 547–554.

Duan X, Yang F, Antono E, Yang W, Pianetta P, Ermon S, Mehta A & Liu Y (2016). Unsupervised data mining in nanoscale X-ray spectro-microscopic study of NdFeB magnet. *Sci Rep* **6**, 1–8.

Dwyer C, Erni R & Etheridge J (2010). Measurement of effective source distribution and its importance for quantitative interpretation of STEM images. *Ultramicroscopy* **110**, 952–957.

Gammer C, Ozdol VB, Liebscher CH & Minor AM (2015). Diffraction contrast imaging using virtual apertures. *Ultramicroscopy* **155**, 1–10.

Greer JR (2006). Bridging the gap between computational and experimental length scales: A review on nano-scale plasticity. *Rev Adv Mater Sci* **13**, 59–70.

Grimley ED, Frisone S, Schenk T, Park MH, Fancher CM, Mikolajick T, Jones JL, Schroeder U & LeBeau JM (2018). Insights into texture and phase coexistence in polycrystalline and polyphasic ferroelectric HfO_2 thin films using 4D-STEM. *Microsc Microanal Conf Proc* **24**, 184–185.

Groschner CK, Choi C & Scott MC (2021). Machine learning pipeline for segmentation and defect identification from high-resolution transmission electron microscopy data. *Microsc Microanal* **27**, 549–556.

Guccione P, Lopresti M, Milanesio M & Callandro R (2021). Multivariate analysis applications in X-ray diffraction. *Crystals* **11**, 12.

Guillamet D, Schiele B & Vitria J (2002). Analyzing non-negative matrix factorization for image classification. *International Conference on Pattern Recognition* **2**, 116–119.

Han Y, Nguyen K, Cao M, Cueva P, Xie S, Tate MW, Purohit PH, Gruner SM, Park J & Muller DA (2018). Strain mapping of two-dimensional heterostructures with subpicometer precision. *Nano Lett* **18**, 3746–3751.

Izadi E, Darbal A, Sarkar R & Rajagopalan J (2017). Grain rotations in ultrafine-grained aluminum films studied using in situ TEM straining with automated crystallogrphic orientation mapping. *Mater Des* **113**, 186–194.

Kacher J, Xie Y, Viogt SP, Zhu S, Yuchi H, Key J & Kalidindi SR (2021). In situ transmission electron microscopy: Signal processing challenges and examples. *IEEE Signal Process Mag* **39**, 89–113.

Kirkland EJ (2020). *Advanced Computing in Electron Microscopy*, 3rd ed. New York: Springer Science & Business Media.

Lee DD & Seung HS (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791.

Londono-Calderon A, Williams DJ, Schneider MM, Savitzsky BH, Ophus C, Ma S, Zhu H & Pettes MT (2021). Intrinsic helical twist and chirality in ulrathin tellurium nanowires. *Nanoscale* **13**, 9606–9614.

Ma J, Jiang X, Fan A, Jiang J & Yan J (2020). Image matching from handcrafted to deep features: A survey. *Int J Comput Vis* **129**, 23–79.

Mahr C, Muller-Caspary K, Grieb T, Krause FF, Schowalter M & Rosenauer A (2021). Accurate measurement of strain at interfaces in 4D-STEM: A comparison of various methods. *Ultramicroscopy* **221**, 113196.

Martineau BH, Johnstone DJ, van Helvoort AT, Midgley PA & Eggeman AS (2019). Unsupervised machine learning applied to scanning precession electron diffraction data. *Adv Struct Chem Imag* **5**, 1–14.

Mehta AN, Gauquelin N, Nord M, Orekhov A, Bender H, Cerbu D, Verbeeck J & Vandervorst W (2020). Unravelling stacking order in epitaxial bilayer mx2 using 4D-STEM with unsupervised learning. *Nanotechnology* **31**, 445702.

Meng T & Zuo J-M (2017). Improvements in electron diffraction pattern automatic indexing algorithms. *Eur Phys J Appl Phys* **80**, 107901.

Mu X, Chen L, Mikut R, Hahn H & Kubel C (2021). Unveiling local atomic bonding and packing of amorphous nanophases via independent component analysis facilitated pair distribution function. *Acta Mater* **212**, 116932.

Mukherjee D, Gamler JTL, Skrabalak SE & Unocic RR (2020). Lattice strain measurement of core@shell electrocatalysts with 4D scanning transmission electron microscopy nanobeam electron diffraction. *ACS Catal* **10**, 5529–5541.

Munshi J, Rakowski A, Savitsky BH, Zeltmann SE, Ciston J, Henderson M, Cholia S, Minor AM, Chan MK & Ophus C (2022). Disentangling multiple scattering with deep learning: application to strain mapping from electron diffraction patterns. Preprint arXiv:2202.00204.

Negishi Y, Nakazaki T, Malola S, Takano S, Niihori Y, Kurashige W, Yamazoe S, Tsukuda T & Hakkinen H (2014). A critical size for emergence of nonbulk electronic and geometric structures in dodecanethiolate-protected Au clusters. *J Am Chem Soc* **137**, 1206–1212.

Ophus C (2017). A fast image simulation algorithm for scanning transmission electron microscopy. *Adv Struct Chem Imag* **3**, 1–11.

Ophus C (2019). Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microsc Microanal* **25**, 563–582.

Ophus C, Zeltmann SE, Bruefach A, Rakowski A, Savitzsky BH, Minor AM & Scott M (2021). Automated crystal orientation mapping in py4dSTEM using sparse correlation matching. *Microsc Microanal* **28**, 390–403.

Panova O, Chen XC, Bustillo KC, Ophus C, Bhatt M, Balsara N & Minor AM (2016). Orientation mapping of semicrystalline polymers using scanning electron nanobeam diffraction. *Micron* **88**, 30–36.

Panova O, Ophus C, Takacs CJ, Bustillo KC, Balhorn L, Salleo A, Balsara N & Minor AM (2019). Diffraction imaging of nanocrystalline structures in organic semiconductor thin films. *Nat Mater* **18**, 860–865.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M & Duchesnay E (2011). Scikit-learn: Machine learning in Python. *J Mach Learn Res* **12**, 2825–2830.

Pekin TC, Ding J, Gammer C, Ozdol B, Ophus C, Asta M, Ritchie RO & Minor AM (2019). Direct measurement of nanostructural change during in situ deformation of a bulk metallic glass. *Nat Commun* **10**, 1–7.

Pekin TC, Gammer C, Ciston J, Minor AM & Ophus C (2017). Optimizing disk registration algorithms for nanobeam electron diffraction strain mapping. *Ultramicroscopy* **176**, 170–176.

Pelz PM, Johnson I, Ophus C, Ercius P & Scott MC (2022). Real-time interactive 4D-STEM phase contrast imaging from electron event representation data: Less computation with the right representation. *IEEE Signal Process Mag* **39**, 25–31.

Ponce A, Aguilar JA, Tate J & Jose Yacaman M (2021). Advances in the electron diffraction characterization of atomic clusters and nanoparticles. *Nanoscale Adv* **3**, 311–325.

Rauch EF & Veron M (2019). Methods for orientation and phase identification of nano-sized embedded secondary phase particles by 4D scanning precession electron diffraction. *Acta Crystallogr B: Struct Sci, Cryst Eng Mater* **75**, 505–511.

Rauch ER, Portillo J, Nicolopoulos S, Bultreys D, Rouvimov S & Moeck P (2010). Automated nanocrystal orientation and phase mapping in the transmission electron microscope on the basis of precession electron diffraction. *Z Kristallogr Cryst Mater* **225**, 103–109.

Roy D, Sing SS, Mitra R, Rosinski M, Michalski A, Lojkowski W & Manna I (2009). Synthesis and characterization of precipitation hardened amorphous matrix composite by mechanical alloying and pulse plasma sintering of $Al_{65}Cu_{20}Ti_{15}$. *Philos Mag* **89**, 1051–1061.

Savitzky BH, Zeltmann SE, Hughes LA, Brown HG, Zhao S, Pelz PM, Pekin TC, Barnard ES, Rangel DaCosta L, Kennedy E, Xie Y, Janish MT, Schneider MM, Herring P, Gopal C, Anapolsky A, Dhall R, Bustillo KC, Ercius P, Scott MC, Ciston H, Minor AM & Ophus C (2021). py4DSTEM: A software package for four-dimensional scanning transmission electron microscopy data analysis. *Microsc Microanal* **27**, 1–32.

Shi C, Cao M, Muller D & Han Y (2021). Rapid and semi-automated analysis of 4D-STEM data via unsupervised learning. *Microsc Microanal Conf Proc* **27**, 58–59.

Shukla AK, Ramasse QM, Ophus C, Kepaptsoglou DM, Hage FS, Gammer C, Bowling C, Gallegos PAH & Venkatachalam S (2018). Effect of

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press

composition on the structure of lithium- and manganese-rich transition metal oxides. *Energy Environ Sci* **11**, 830–840.

Thati SK, Ding J, Zhang D & Hu XH (2015). Feature selection and analysis of diffraction images. *IEEE International Conference on Software Quality, Reliability and Security - Companion (QRS) 2017*.

Thornsen E, Frafjord J, Friis J, Marioara C, Wenner S, Andersen S & Holmestad R (2021). Studying GPI zones in Al-Zn-Mg alloys by 4D-STEM. *Mater Charact*, 111675.

Treacy MMJ & Gibson JM (1996). Variable coherence microscopy: A rich source of structural information from disordered materials. *Acta Crystallogr A* **52**, 212–220.

Uesugi F, Koshiya S, Kikkawa J, Nagai T, Mitsuishi K & Kimoto K (2021). Non-negative matrix factorization for mining big data using four-dimensional scanning transmission electron microscopy. *Ultramicroscopy* **221**, 113168.

van der Walt S, Schonberger JL, Nunez-Iglesias J, Boulgne F, Warner JD, Yager N, Gouilart E, Yu T & the scikit-image contributors (2014). scikit-image: Image processing in Python. *PeerJ* **2**, e453.

Velazquez-Salazar JJ, Esparza R, Mejia-Rosales SJ, Estrada-Salas R, Ponce A, Deepak FL, Castro-Guerrero C & Jose-Yacaman M (2011). Experimental evidence of icosahedral and decahedral packing in one-dimensional nanostructures. *ACS Nano* **5**, 6272–6278.

Walton W (1948). Feret's statistical diameter as a measure of particle size. *Nature* **162**, 329–330.

Wang Q-L, Fang R, He L-L, Feng J-J, Yuan J & Wang A-J (2016). Bimetallic PdAu alloyed nanowires: Rapid synthesis via oriented attachment growth and their high electrocatalytic activity for methanol oxidation reaction. *J Alloys Compd* **684**, 379–388.

Wang Y, Choi S-II, Zhao X, Peng H-C, Chi M, Huan CZ & Xia Y (2014). Polyol synthesis of ultrathin Pd nanowires via attachment based growth and their enhanced activity towards formic acid oxidation. *Adv Funct Mater* **21**, 131–139.

Wang Y, Wang Q, Sun H, Zhang W, Chen G, Wang Y, Shen X, Han Y, Lu X & Chen H (2011). Chiral transformation: From single nanowire to double helix. *J Am Chem Soc* **133**, 20060–20063.

Yang N, Ophus C, Savitzsky BH, Scott MC, Bustillo K & Lu K (2021). Nanoscale characterization of crystalline and amorphous phases in silicon oxycarbide ceramics using 4D-STEM. *Mater Charact* **181**, 1111512.

Yao T, Sun Z, Li Y, Pan Z, Wei H, Xie Y, Nomura M, Niwa Y, Yan W, Wu Z, Jiang Y, Liu Q & Wei S (2010). Insights into initial kinetic nucleation of gold nanocrystals. *J Am Chem Soc* **132**, 7696–7701.

Yuan R, Zhang J, He L & Zuo J-M (2021). Training artificial neural networks for precision orientation and strain mapping using 4D electron diffraction datasets. *Ultramicroscopy* **231**, 113256.

Zeltmann SE, Muller A, Bustillo KC, Savitsky B, Hughes L, Minor AM & Ophus C (2020). Patterned probes for high precision 4D-STEM Bragg measurements. *Ultramicroscopy* **209**, 112890.

Zintler A, Eilhardt R, Wang S, Krajnak M, Schramowski P, Stammer W, Petzold S, Kaiser N, Kersting K, Alff L & Molina-Luna L (2020). Machine learning assisted pattern matching: Insight into oxide electronic device performance by phase determination in 4D-STEM datasets. *Microsc Microanal Conf Proc* **26**, 1908–1909.

https://doi.org/10.1017/S1431927622012259 Published online by Cambridge University Press