Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Unsupervised Machine Learning for Automated Crystal Orientation Mapping on Noisy 4D-STEM Data

Kai Kamijo
Tohoku University
Motoki Shiga
motoki.shiga.b4@tohoku.ac.jp

Tohoku University
Shusuke Kanomi
Tohoku University
Tomohiro Miyata
Tohoku University
Hiroshi Jinnai
Tohoku University

Article

Keywords:

Posted Date: April 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9374144/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at npj Computational Materials on October 3rd, 2026. See the published version at https://doi.org/10.1038/s41524-026-02337-x.

Unsupervised Machine Learning for Automated Crystal Orientation Mapping on Noisy 4D-STEM Data

Kai Kamijo[1], Motoki Shiga[1,2,3,4,*, †], Shusuke Kanomi[5,6], Tomohiro Miyata[5], and Hiroshi Jinnai[5]

*Corresponding author
†motoki.shiga.b4@tohoku.ac.jp

[1] Graduate School of Information Sciences, Tohoku University, 6-3-09 Aoba, Aramaki-aza Aoba-ku, Sendai, 980-8579, Japan

[2] Unprecedented-scale Data Analytics Center, Tohoku University, 468-1 Aramaki-Aza-Aoba, Aoba-ku, Sendai 980-8578, Japan

[3] Center for Basic Research on Materials, National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan

[4]RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

[5]Institute of Multidisciplinary Research for Advanced Materials, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan

[6]Science & Innovation Center, Mitsubishi Chemical Corporation, 1000 Kamoshida-cho, Aoba-ku, Yokohama-shi, Kanagawa, 227-8502, Japan

Abstract

Four-dimensional scanning transmission electron microscopy (4D-STEM) is a high-throughput measurement used to comprehensively acquire diffraction patterns across a region of interest in a specimen. While the measurement is useful for identifying crystallographic microstructures, the data analysis requires significant effort owing to the complexity and high dimensionality of the data. Furthermore, the analysis of low signal-to-noise data from beam-sensitive materials such as polymers is particularly challenging. This study proposes an automatic analysis pipeline that leverages denoising and unsupervised crystallographic analysis for 4D-STEM data. The denoising method is based on a self-supervised approach that eliminates the need for clean training images. The developed method is extended to 4D-STEM data by utilizing diffraction patterns observed at neighboring probe positions. The subsequent unsupervised analysis is designed to be equivariant to the crystal in-plane rotations, enabling the effective identification of crystal components. The efficacy of the proposed pipeline is evaluated using both synthetic and experimental 4D-STEM data. Quantitative evaluation for synthetic data across various noise levels demonstrates the robustness of our approach. The result from the experimental data shows good agreement with expert analysis in the previous study; furthermore, the pipeline identifies minor crystallographic components that were previously overlooked. These results demonstrate that the pipeline enables robust and automated analysis with significantly reduced analysis cost.

Introduction

Electron microscopy is an essential tool in materials science for identifying material and device structures from multiple perspectives. Using electron microscopy, we can obtain

various data, such as spatial images, spectra, and diffraction patterns, to unveil new insights into materials. Advancements in microscopic instrumentation have enabled comprehensive and simultaneous measurements across various techniques. For instance, spectral imaging techniques, such as electron energy loss spectroscopy (EELS) and energy dispersive X-ray spectroscopy (EDS), have facilitated the comprehensive acquisition of spectral data across a region of interest (ROI). The resulting spectral images can then be used to determine the spatial distributions of chemical states. Another useful technique is 4D scanning transmission electron microscopy (4D-STEM), which enables the comprehensive acquisition of diffraction patterns across an ROI [1]. This technique records the diffraction pattern at each probe position on a 2D spatial grid, resulting in a 4D dataset. The observed data are used to analyze crystal microstructures by identifying the crystallographic orientations at each spatial point [2–5]. However, the high dimensionality and sheer volume of datasets generated by these techniques render manual analyses impractical. Hence, the need for automated analysis pipelines to comprehensively characterize underlying microstructures has steadily grown.

Machine learning is an effective approach for automatically discovering hidden rules and relationships within observed data. With the increasing number of publicly available datasets and developments in measurement technologies, machine learning has become an indispensable component in both academic studies and commercial industries. Machine learning has been applied to microscopy data for denoising (image restoration), 3D structure estimation, and chemical component identification [6]. In noisy microscopy data, denoising is an essential data-processing step for reliable structure analysis. In 4D-STEM of electron-beam sensitive materials, such as polymers targeted in this study, the measurements are necessarily conducted under low-dose conditions. This yields low-count diffraction patterns dominated by severe statistical fluctuations following a Poisson

distribution, which hampers accurate microstructural analysis. In scientific research fields, a self-supervised denoising approach has emerged as a powerful tool because it enables the training of denoising models using only noisy observations. Therefore, this approach does not rely on clean ground-truth data, which is often difficult to obtain in experimental studies. A seminal self-supervised denoising method is Noise2Void [7], which requires only a single noisy image by leveraging the blind-spot architecture. During model training, Noise2Void masks randomly selected pixels as denoising targets and trains the model to accurately predict noise-free intensities from neighboring pixels. The training of the model parameters was optimized by minimizing the error between the predictions and the observed noisy pixels. Theoretically, the predictions of the model are statistically consistent with those performed by models trained using supervised clean data. Thus, self-supervised models based on this architecture can achieve comparable performance to supervised models using clean ground-truth data. A primary drawback of Noise2Void is its low training efficiency because only a limited number of pixels are available for model training. By incorporating a masking mechanism into the denoising model, Laine's proposed architecture enables the model to use all the pixels for training [8]. This denoising model was further extended to video data, which are represented as a 3D numerical array featuring temporal correlations in addition to a 2D space [9]. By utilizing temporal correlations from adjacent frames, this model achieves robust denoising performance, as demonstrated by the successful acquisition of high-speed videos of catalytic reactions at atomic resolution from noisy TEM image observations [10]. In contrast, the development of efficient denoising methods for 4D-STEM data by leveraging correlations between diffraction patterns remains a significant challenge.

Another useful machine learning approach is unsupervised learning such as non-negative matrix factorization (NMF) [11] and vertex component analysis (VCA) [12]. These methods

have been adapted for spectral image datasets such as STEM-EELS/EDS, which are represented as a 3D numerical array, to automatically identify the spatial distribution of chemical components [13–15]. Moreover, a high-throughput analysis pipeline was developed by leveraging dimensional reduction and unsupervised learning techniques to identify the 3D microstructures of chemical components from spectral imaging data combined with a computed-tomography technique, which is represented as a 4D array [16]. These approaches facilitate high-throughput analysis while eliminating errors arising from manual analysis. Furthermore, they allow the extraction of complex but essential structural features from massive noisy datasets based on quantitative criteria. Following these approaches, unsupervised methods for 4D-STEM data have been developed to analyze crystallographic microstructures [17–20]. However, these methods failed to account for the rotation-equivariant property in diffraction patterns, which is crucial for interpreting the patterns. Consequently, these conventional methods cannot efficiently analyze 4D-STEM data containing diverse orientation variants.

To address these challenges, we propose a robust analysis pipeline consisting of two machine learning methods for denoising and crystallographic component analysis. For denoising, we propose a self-supervised method for 4D-STEM data that leverages correlations underlying in both the reciprocal space of each diffraction pattern and the real space related to the probe position. For the component analysis, we identify the crystallographic components as groups of probe positions observed equivalent diffraction patterns, invariant to in-plane rotation. Since the diffraction pattern of each component can be indexed to specific electron beam directions defined by Miller indices, the resulting spatial mapping of the components effectively reveals the underlying crystallographic microstructure. To identify these components, we adapt NMF to ensure rotational equivariance by utilizing a Fourier transformation in the polar coordinate representations

of the diffraction patterns. We evaluate the efficacy of the proposed pipeline using both synthetic and experimental 4D-STEM data [5]. The quantitative evaluation for synthetic data across various noise levels demonstrates the robustness of our approach. The result from the experimental data shows good agreement with expert analysis in the previous study [5]. Furthermore, the pipeline identifies minor crystallographic components that were previously overlooked. These results demonstrate that the pipeline enables robust and automated analysis with significantly reduced analysis cost.

Results

Fig. 1a illustrates our proposed pipeline for generating a spatial map of crystal microstructures from 4D-STEM data. In this pipeline, the observed diffraction patterns are first denoised using a self-supervised model that equips a blind-spot architecture (**Fig. 1b**). Based on the denoised patterns, the spatial distribution of the crystallographic components is estimated using an unsupervised method (**Fig. 1c**). The following sections describe each pipeline component in detail.

Self-Supervised Denoising for 4D-STEM Data

We propose a self-supervised denoising method for 4D-STEM data based on a blind-spot neural network architecture, generalizing models developed for images (2D data) [8] and videos (3D data) [9]. The proposed model leverages correlations not only within the reciprocal space of each diffraction pattern but also across the real space related to the probe position. Hereafter, we refer to our proposed denoising method as self-supervised denoising for 4D-STEM data (4D-SSD), whereas the baseline model applied to individual 2D images (diffraction patterns) is denoted as 2D-SSD.

Fig. 1b shows the 4D-SSD and 2D-SSD models, both of which enables an efficient computation of a blind-spot architecture. To facilitate understanding, we first describe the blind-spot architecture in the 2D-SSD model. Using an input diffraction pattern, the model first generates four diffraction patterns (2D images) by rotating the input pattern by 0°, 90°, 180° and 270° and then the patterns are processed independently by a directional U-Net (**Supplementary Fig. 1a**). By utilizing upward-shift operation after convolutional and pooling layers in the U-Net, each pixel in the U-Net output is computed using only the pixels located at the same vertical positions or in the rows above in the input pattern, as a portion of the architecture is illustrated in **Supplementary Fig. 1b**. It results in lower pixels in the input are blinded for the output computation. To exclude the input pixel at the same position for the output computation, the output pattern is further shifted downward by one pixel relative to the input. As shown in **Supplementary Fig. 1c**, this operation is implemented by padding on the top and cropping on the bottom row in the pattern. The four patterns processed via the U-Net and downward-shift operation are rotated back to their original rotations and subsequently integrated using convolutional layers with 1×1 kernels to generate the denoised pattern. Overall, each pixel in the 2D-SSD model output is computed without reference to the input pixel at the same position, thereby realizing a blind-spot architecture (**Supplementary Fig. 1d**).

As shown in **Fig. 1b**, the 4D-SSD model is a generalization of the 2D-SSD model for leveraging the real spatial correlations associated with the probe positions. To utilize the correlation, the 4D-SSD model inputs diffraction patterns observed at probe positions adjacent to the target position (x, y) in addition to the denoising target pattern. These patterns are combined by concatenating them along the channel dimension. For the computational space efficiency, instead of processing all adjacent patterns at once, the model processes the subsets by dividing them into four triplets containing three patterns

(the target and two adjacent patterns): $\{(x-1,y+1),(x,y),(x+1,y-1)\},\{(x,y+1),(x,y),(x,y-1)\}$, $\{(x-1,y),(x,y),(x+1,y)\}$, and $\{(x+1,y+1),(x,y),(x-1,x-1)\}$. Similar to the 2D-SSD framework, four rotated patterns are generated from each triplet, which are subsequently processed independently by the first directional U-Net. From the 16 output patterns of the U-Net, those rotated by the same angles are concatenated; these concatenated patterns are processed by the second directional U-Net to integrate the information from all adjacent patterns. Consistent with the data processing scheme in the 2D-SSD model, the four output patterns via the second U-Net are shifted downward and subsequently rotated back to their original rotations. Consequently, these outputs are integrated using convolutional layers with 1×1 kernels to generate the final denoised pattern. Similar to 2D-SSD, 4D-SSD retains a blind-spot architecture, even though it incorporates adjacent diffraction patterns.

The parameters of the 4D-SSD model are trained by minimizing the mean squared errors between the model outputs and the noisy patterns. From a 4D-STEM dataset, each training sample, which serves as the model input, consisted of a 3×3 diffraction patterns centered at the target probe position. In numerical experiments, the samples were split into training and validation subsets at an 8:2 ratio, and the validation loss was used for the learning rate adaptation and early stopping, as described in section **DNN Model Training** of **Methods**. Because the blind-spot architecture prevents direct access to the target pixel intensities for noise-free intensity prediction during training, the same dataset can be used for both training and evaluation without information leakage.

Because noisy patterns follow a Poisson distribution, we applied the Anscombe transform to the observed diffraction patterns (inputs of the denoising models) and an inverse transform to the model outputs, to approximate the noise as the standard Gaussian noise.

Despite the proposed denoising, a low-frequency background component that is distinct from the observed random noise but irrelevant to the crystal structure remain. Therefore, a Gaussian high-pass filter was applied via a 2D Tukey window to remove this component and emphasize the diffraction spots as a postprocessing step.

Crystallographic Component Analysis

Based on the denoised diffraction patterns, the crystallographic microstructure was estimated using the unsupervised learning method developed in this study. Our unsupervised method identifies crystallographic components as groups of probe positions observed equivalent diffraction patterns, invariant to in-plane rotation. These patterns can be indexed to specific electron beam directions defined by Miller indices by comparing them with simulated patterns. As shown in **Fig. 1c**, this method consists of four steps: (1) transformation of the denoised patterns into polar coordinates for angular analysis, (2) Fourier transformation along the angular axis of the polar data representation, (3) unsupervised component analysis for the transformed data, and (4) estimations of the local crystalline orientations for the identified components based on the Fourier phase analysis. The estimated components and orientations are subsequently mapped back onto the real space of the specimen. The details of each step are described below.

Data Transformation for Estimating Crystallographic Microstructures

The rotational symmetry of the diffraction patterns is analyzed by transforming each denoised pattern from Cartesian coordinates in the 2D reciprocal space (q_x, q_y) into polar coordinates represented by a radial distance r and a rotation angle θ. This step enables the diffraction patterns to be functions of the radial distance and azimuthal angle. To determine the rotational orders regardless of the rotational angles, we perform a 1D

Fourier transform along the angular axis for each radial position. This procedure converts the polar image at each probe position $I_{xy}(r,\theta)$ into a 2D Fourier representation in the (r,m) domain as follows:

$$F_{xy}(r,m) = FT_{\theta}\{I_{xy}(r,\theta)\},$$

where m denotes the rotational order.

Unsupervised Crystallographic Components Analysis

The diffraction patterns represented by the magnitude of the Fourier representation $|F_{xy}(r,m)|$ are decomposed into representative crystallographic components. We apply NMF [11] for this task because all values in the magnitude are non-negative, allowing us to extract physically interpretable components. For NMF, the data matrix $\boldsymbol{X}$ is first generated by stacking each diffraction pattern represented by the Fourier magnitude as a row vector. Next, NMF is used to factorize the matrix $\boldsymbol{X}$ into a coefficient matrix $\boldsymbol{W}$ and a basis matrix $\boldsymbol{H}$ as follows:

$$\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H}$$

Each column vector in $\boldsymbol{W}$ represents component weights of diffraction patterns for a component, while each row vector in $\boldsymbol{H}$ represents the Fourier magnitude pattern of a component. The spatial distributions of the identified components are generated by rearranging the elements of $\boldsymbol{W}$ according to the corresponding probe positions. The Fourier magnitude images of the identified components are generated by rearranging each row vector in $\boldsymbol{H}$. Unknown matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ are obtained by minimizing the mean squared error between $\boldsymbol{X}$ and its reconstruction by $\boldsymbol{W}\boldsymbol{H}$ under the nonnegativity constraint for all elements. In this study, we implemented the minimization based on the hierarchical alternative least-squares (HALS) algorithm [21] using the Python package malspy [22]. Unlike the existing method that applies NMF directly to the observed diffraction patterns [20], we

implemented it on the Fourier magnitudes of the patterns. Although this approach is straightforward, our data transformation is essential for achieving an in-plane rotation-invariant analysis of the crystallographic components.

In-Plane Rotation Estimation of Crystallographic Components

To identify the major rotational order of each component, the proposed pipeline first evaluates the presence of peaks at each radial distance r in the Fourier magnitude images of the identified components. For the evaluation, a robust z-score is computed for the profile obtained by summing the Fourier magnitudes across the radial-distance axis. The presence of peaks is determined by whether the z-score of the radial distance exceeded a predefined threshold, which was fixed at 2.0 in our analysis.

Next, 1D profiles are extracted along rotational-order axis at the radial distances evaluated as having peaks. We then average the extracted profiles at consecutive radial distances into a profile. In the Fourier transform of non-sinusoidal signals, such as the diffraction patterns considered here, higher-order harmonics at integer multiples of the fundamental rotational order can also exhibit strong intensities. Therefore, explicitly accounting for the presence of these harmonic components is necessary to correctly extract the dominant rotational order of the electron diffraction spots. Two-, three-, four-, and six-fold rotational symmetries can occur in electron diffraction patterns of crystals. Among these, the three-fold rotational symmetry appears as six-fold symmetry in electron diffraction patterns owing to centrosymmetry. Consequently, only the even rotational orders must be considered. For automatic rotational order estimation, the peak detection based on a robust z-score is performed at the positions with even rotational orders along the axis. The threshold for the detection was fixed at 1.5 in all experiments. The rotational orders in which a peak is detected are considered candidates for the subsequent in-plane rotation

estimation. If no peak is identified, the signal is interpreted as being distributed across all even rotational orders, and the minimum twofold rotational symmetry is considered the candidate. For the diffraction pattern at probe position (x, y), this step generated the candidate peak positions $\left(R_{ci}^{(x,y)}, K_{ci}^{(x,y)}\right), i = 1,2,\ldots$, where $R_{ci}^{(x,y)}$ and $K_{ci}^{(x,y)}$ are the radial distance and rotational order of the i-th peak in the c-th component, respectively.

Using each selected position $\left(R_{ci}^{(x,y)}, K_{ci}^{(x,y)}\right)$, the phase computed by Fourier transformation at probe position (x, y) was converted into an in-plane rotational angle as follows:

$$\phi_c^{(x,y)} = \frac{1}{K_{ci}^{(x,y)}} \arg\left(F_{xy}\left(R_{ci}^{(x,y)}, K_{ci}^{(x,y)}\right)\right) \quad (1).$$

The rotational angles of all components at all probe positions are computed to generate a spatial map of the crystallographic in-plane orientation. Component weights represented by matrix $\boldsymbol{W}$ are used as spatial masks so that orientation values are assigned only to regions associated with the major component.

Performance Evaluation using Synthetic Data

We evaluated our proposed pipeline using synthetic data. The assumed crystal structure was isotactic polystyrene (iPS), which was the target material for our experimental 4D-STEM observations. The microstructure assumed in the synthetic data is shown in **Fig. 2**. This structure contains four crystallographic components whose diffraction patterns are obtained with the electron-bean incidence along the crystallographic orientation: [0 0 1], [1 $\bar{1}$ $\bar{1}$], [2 $\bar{2}$ 1] and [2 $\bar{5}$ 1]. **Fig. 2a** shows theoretical electron diffraction patterns computed using software ReciPro [23] as described in section **Computation of Ground-Truth Diffraction Patterns** of **Methods**. The structure was discretized on a 100×100 grid

(probe positions) in the real space and partitioned into three distinct regions: the top-left and bottom-left squares and the right half-region (**Fig. 2b**). While the first two regions each contain a single component of the crystallographic orientation [0 0 1] or [1 $\bar{1}$ $\bar{1}$], the latter includes two components of [2 $\bar{2}$ 1] and [2 $\bar{5}$ 1], the ratios of which vary linearly along the vertical axis, producing a smooth transition between them. **Fig. 2c** shows the spatial distribution of the in-plane rotational angles. In the region of component [1 $\bar{1}$ $\bar{1}$], the angle varies linearly along the vertical axis, whereas the angles remain constant in the other regions. The ground-truth data were computed from the simulated diffraction patterns of the four components using the component intensities (**Fig. 2b**) and in-plane rotational angles (**Fig. 2c**). Synthetic noisy diffraction patterns were then generated using a Poisson distribution. For each pixel, the mean parameter of the distribution was set as the product of the ground-truth intensity and weight parameter w_s. The parameter w_s was introduced to adjust the signal-to-noise (S/N) ratio in synthetic datasets.

We applied our proposed denoising method (4D-SSD) to the synthetic 4D-STEM datasets with $w_s = 0.1, 0.2\ 0.3, 0.5,$ and 1.0, and compared the results with those obtained from a conventional Gaussian filter and 2D-SSD. **Fig. 3** shows the ground-truth diffraction patterns, observed pattens in synthetic data with $w_s = 0.2$ where contain only a single component and denoised patterns. The denoising results in this figure are the direct outputs from the denoising methods, prior to the high-pass filter. **Fig. 3** shows that the synthetic observed patterns contain heavy stochastic noise, which obscures the diffraction spots and hinders the identification of their symmetry. Although the Gaussian filter blurred the intensities to reduce noise, the filtered images remain too noisy to accurately detect diffraction spots. By contrast, 2D-SSD and 4D-SSD successfully recovered the diffraction spots while suppressing noise. However, 2D-SSD generated artifacts, which are indicated by the magenta arrows. The artifacts appear to originate from the component

[0 0 1], which is the major component of the data. These artifacts occurred because 2D-SSD directly captures representative parts of the individual diffraction patterns during training and outputs them as denoised images. Conversely, 4D-SSD successfully suppressed these artifacts. In addition, 4D-SSD restored the weak spots observed in the regions indicated by the white arrows for $[1\ \bar{1}\ \bar{1}]$. These improvements were achieved by leveraging the real space correlation associated with probe positions, which was not utilized by 2D-SSD. Even when diffraction spots are hidden in individual noisy patterns, 4D-SSD can accurately restore these spots by utilizing neighboring diffraction patterns that share similar crystallographic information. Overall, the patterns denoised using 4D-SSD exhibit the highest similarity to the ground truth patterns among the compared methods. This improvement in denoising performance is of substantially importance for crystallographic component analysis, as will be demonstrated later.

Next, these denoising methods were compared using quantitative performance metrics: PSNR and an unsupervised PSNR (uPSNR) [24]. Despite not using ground-truth data, uPSNR is a statistically consistent estimator of the PSNR. The detailed definition is described in section **Unsupervised Metric for Evaluating Denoising Results** in **Methods**. The denoised patterns were normalized to match to the ground-truth patterns and enable a comparison of the uPSNRs among synthetic datasets with different noise levels; here, the intensity range of the ground-truth pattern was used as M for the computation. **Table 1** summarizes the metrics of the compared methods for the synthetic data with $w_s = 0.2$. The results for other noise levels $w_s = 0.1, 0.3, 0.5, \text{and } 1.0$ are summarized in **Supplementary Table 1**. Consistent with the results shown in **Fig. 3**, 2D-SSD and 4D-SSD achieved higher PSNRs than the Gaussian filter, indicating the most faithful restoration of the ground-truth patterns. However, no statistically significant difference was observed between 2D-SSD and 4D-SSD because the diffraction spots

occupied only a small fraction of the total image area. Notably, the same performance ranking was maintained when evaluations were conducted using the unsupervised metric uPSNR. The methods that yielded higher PSNRs also yielded higher uPSNRs. This close agreement between PSNR and uPSNR demonstrates that uPSNR provides a reliable surrogate for supervised evaluation, even without access to clean ground-truth patterns.

Fig. 4 shows analysis results of the crystallographic components obtained by the subsequent procedure. As a reference, **Figs. 4a–4c** show the ground-truth diffraction patterns of the four components (**Fig. 4a**), the patterns transformed into the polar representation (**Fig. 4b**), and their 1D Fourier transforms along the angular axis (**Fig. 4c**). The Fourier-transformed data were band-limited in [2, 14]. The changes in these transforms owing to an in-plane rotation are shown in **Supplementally Fig. 2**. According to this figure, patterns in the polar representation shifted along the angular axis according to the in-plane rotational angles, whereas the magnitudes of the Fourier transforms remained invariant. This observation demonstrates that this representation successfully extracts rotation-invariant features while discarding in-plane rotation-dependent phase information.

To identify the crystallographic components, we applied NMF to the Fourier magnitude images obtained from the diffraction patterns denoised by 4D-SSD. The number of components in NMF was set to four, corresponding to the ground truth. **Fig. 4d** shows the Fourier magnitudes identified by NMF, while **Fig. 4e** displays the spatial distributions of the corresponding components. The results are in good agreement with the ground-truth distributions shown in **Fig. 2b**. It exhibited that NMF successfully recovered the underlying structural components by effectively disentangling the spatial mixing of components and achieving an in-plane rotation invariant estimation.

Next, we estimated the spatial distributions of the in-plane rotational angles of the components. **Fig. 4f** presents representative results of the peak-detection procedure applied to the Fourier magnitudes of the components estimated by NMF. For Fourier magnitudes shown in the top-right panels in **Fig. 4f**, the procedure first performed the summations along the angular axis to compute the robust z-scores. The computed z-scores left shown in the left panels. Using these z-scores, radial distances containing peaks were identified as highlighted by the orange and light-blue rectangles. We then computed the summations along the radial-distance axis within the range of the radial distances detected in the previous step, as shown in the bottom panels. Using these results and considering symmetric constraints, the informative peaks were detected, as highlighted by the magenta rectangles. The procedure revealed that the component [0 0 1] had sixfold rotational symmetry. In contrast, although two radial peaks were detected for the component $[2\ \bar{5}\ 1]$, no significant rotational order was identified for either peak. Therefore, the rotational symmetry of the component was estimated to be twofold. **Fig. 4g** highlights the ranges of radial distances determined in this step, using magenta dashed circles. These data are used for the in-plane rotational angle estimation.

Using the estimated rotational order and peaks, we estimated the in-plane rotational angles of all probe positions from the Fourier phase using the positions identified by the peak detection. **Fig. 4h** presents the spatial distribution of the in-plane rotational angles estimated using the estimation method obtained via 2D-SSD and 4D-SSD, along with the ground-truth data. The spatial distributions generated using 4D-SSD highly matched with the ground-truth data, whereas the 2D-SSD result is significantly degraded. For instance, 4D-SSD clearly captured the absence of the in-plane rotational variation in the component [0 0 1] and the systematic rotational variation in the component $[1\ \bar{1}\ \bar{1}]$. For regions in which the components $[2\ \bar{2}\ 1]$ and $[2\ \bar{5}\ 1]$ overlapped, the method captured a distinct

change in the estimated orientation observed across the central region. In the overlapping regions, angle estimation was primarily governed by the NMF-based component extraction. The accuracy of NMF was also influenced by the S/N ratio of the input data, which depended on the denoising performance. This result demonstrates the importance of effective denoising for the reliable crystallographic orientation analysis of 4D-STEM data.

We further compared these methods using additional synthetic data with S/N parameters $w_s = 0.1, 0.2, 0.3, 0.5, \text{and } 1.0$. **Supplementary Fig. 3** shows examples of the denoising results while **Supplementary Fig. 4** displays the analysis results of spatial distribution of components and their orientations obtained via 2D-SSD and 4D-SSD denoising. The results demonstrated that 4D-SSD is more robust than 2D-SSD for noisy observations. This robustness was further confirmed from quantitative performance metrics of the Pearson correlation coefficients between the estimation and ground-truth data (**Fig. 5**). The estimations of the component intensities (**Fig. 5a**) and rotational angles (**Fig. 5b**) obtained using 4D-SSD outperformed those obtained using 2D-SSD in all cases. These results confirm that the NMF can robustly identify crystallographic components from 4D-STEM data denoised by 4D-SSD even when these components are spatially mixed and rotated within the observation regions. Overall, the analysis results of the synthetic data demonstrate that the proposed pipeline provides robust estimations of the crystallographic microstructure from noisy 4D-STEM data.

Microstructure Analysis of the Experimental Data of Isotactic Polystyrene

Our proposed pipeline was applied to the experimental 4D-STEM data of an isotactic polystyrene (iPS) obtained in the previous study [5]. As shown in **Fig. 6a**, an iPS spherulite comprises flat-on and edge-on lamellar crystals coexisting with each other. Therefore, this

417 data provides a suitable testbed for evaluating the applicability of the proposed pipeline
418 under realistic experimental conditions involving complex structures and heavy noise. **Fig.**
419 **6b** shows ADF-STEM image of the observed region on iPS. The image was computed by
420 averaging intensities within an annular region between 1.27 nm^{-1} and 4.43 nm^{-1} from the
421 center of the diffraction patterns, as indicated in **Supplementary Fig. 5**. For the
422 experimental data, we first denoised using a Gaussian filter, 2D-SSD, and 4D-SSD. **Fig. 7**
423 shows observed and denoised diffraction patterns at the position (1) – (4) in **Fig. 6b**.
424 Consistent with the results for the synthetic data shown in **Fig. 3**, the 2D-SSD and 4D-
425 SSD methods reduced the observation noise more effectively than the Gaussian filter,
426 yielding clearer diffraction spots. This finding was also observed for the quantitative
427 metric uPSNRs, as summarized in **Table 1**. The 2D-SSD and 4D-SSD methods improved
428 the uPSNR by over 10 dB compared with the Gaussian filter. Furthermore,
429 **Supplementary Fig. 6** demonstrates that the postprocessing using a high-pass filter
430 emphasizes diffraction spots in the results obtained via 2D-SSD and 4D-SSD methods but
431 highlights both the spots and noise in the case of the Gaussian filter.

432 Next, the patterns denoised by 4D-SSD and the high-pass filter were Fourier transformed,
433 followed by the crystallographic component estimation using NMF with seven
434 components. To emphasize diffraction spots located far from the center, we conducted the
435 log transformation for pre-processing. Small intensities leading to underflow were
436 replaced with zeros after the transformation. The results for the data denoised using the
437 proposed method are presented in **Supplementary Fig. 7**. **Supplementary Figs. 7a** and
438 **7b** show the Fourier magnitudes and spatial intensity distributions of components
439 identified by the NMF, respectively. **Supplementary Fig. 7c** shows the distribution of the
440 in-plane rotational angles. The angles were computed using the peak frequencies indicated
441 by the red arrows in **Supplementary Fig. 7a**. In **Supplementary Fig. 7c**, the angles are

displayed only in the map of the component where yields the maximum intensity at each position. Representative diffraction patterns extracted from the spatial positions in which only specific components are dominant are summarized in **Fig. 8** and **Supplementary Fig. 8**. In these figures, the images in (b) and (c) correspond to the observed patterns and those denoised by the proposed method, respectively. Patterns (1) and (2) in component I, shown in **Fig. 8c**, exhibit a clear six-fold symmetry resembling the pattern of the [0 0 1] orientation, although they are hidden in images of **Fig. 8b** owing to the observation noise. Patterns (3) and (4) in components II appeared to exhibit the same symmetry, although a diffraction spot is not observed because of the beam stopper. Thus, these components highly matched the flat-on lamella. On the other hand, components III and IV aree considered as the edge-on lamella because their patterns are matched the diffraction patterns from either of $[1\ \bar{1}\ \bar{1}]$, $[2\ \bar{2}\ 1]$, or $[2\ \bar{5}\ 1]$. While component VII is the background, components V and VI are new components, neither of which are analyzed in the previous study. However, identifying their crystallographic orientations is difficult because these components are mixtures of various minor components.

Fig. 6c shows the flat-on region (the sum of components I and II) and edge-on region (the sum of components III and IV) regions resembling the left half of Fig. 4c in the previous work [5], which was analyzed by experts. This result demonstrates that the NMF-based automatic analysis successfully captured the dominant structural components in the experimental data. **Figs. 6d** and **6e** show the spatial distributions of the in-plane rotational angles estimated using our pipeline. The flat-on region in **Fig. 6d** exhibited a slight variation in orientation, indicating a nearly uniform in-plane alignment. By contrast, the edge-on region shown in **Fig. 6e** exhibited a clear spatial variation, reflecting the presence of locally varying lamellar orientations. **Fig. 9** shows a magnified view of spatial distribution of in-plane rotational angles of component III (an edge-on lamellar region)

indicated in **Fig. 6e**. The main long lamellar indicated in (3), (5), and (6) are differentiated from the branched lamella in (4). This result is consistent with Fig. 6 in the paper [5], which relied on an expert-guided interpretation of the selected diffraction patterns.

Discussion

We developed a self-supervised denoising and unsupervised learning methods for the automatic crystallographic analysis of 4D-STEM data. The effectiveness of these methods was demonstrated using both synthetic and experimental real datasets. The proposed pipeline offers a new approach for analyzing low S/N 4D-STEM datasets acquired from beam-sensitive materials. In future work, we plan to automatically interpretate diffraction patterns generated by the proposed method. This study analyzed only major patterns to identify the material microstructure; the orientation identification, such as Miller index estimation, which is necessary for analyzing more complicated microstructures remains difficult. Developing a robust supervised learning approach using comprehensive simulation datasets is an important challenge to realize fully automated analysis of noisy 4D-STEM data.

Methods

DNN Model Training

The code was created using Python. The deep learning models used for 2D-SSD and 4D-SSD denoising were implemented using the PyTorch package (ver. 2.9.0). The model parameters were optimized using the Adam algorithm [24] with an initial learning rate of 10^{-4}. The learning rate was adaptively reduced when the validation loss plateaued using function “torch.optim.lr_scheduler.ReduceLROnPlateau”. In addition, early stopping was

applied by monitoring the validation loss to prevent overfitting. The model checkpoint that achieved the lowest validation loss was saved and used for the final evaluation.

From a 4D-STEM dataset, each training sample, which served as the model input, consisted of a 3×3 neighborhood of nine diffraction patterns centered at the target probe position. For the synthetic data, we used all valid target positions from a 100×100 probe-position grid, which yielded 10,000 samples. For the real data, we randomly selected 10,000 target probe positions from a full 500×500 scan (250,000 positions) and constructed a dataset by extracting the corresponding 3×3 neighborhoods. In both cases, the resulting dataset was split into training and validation subsets at an 8:2 ratio, and the validation loss was used to control the training.

Unsupervised Metric for Evaluating Denoising Results

An issue in self-supervised denoising learning is its evaluation. As clean ground-truth images are inherently unavailable in real experimental settings, standard supervised metrics such as the conventional peak signal-to-noise ratio (PSNR) cannot be applied for the performance evaluation. To enable an evaluation without the need for ground-truth images, we employed an unsupervised PSNR (uPSNR) metric [25]. Although this metric is computed solely from noisy images, it serves as a statistically consistent estimator of the conventional supervised PSNR. Thus, uPSNR enables a quantitative comparison of denoising methods using only experimental 4D-STEM noisy data, obviating the need for clean ground-truth data. The uPSNR metric for a diffraction pattern at probe position (x, y) is defined using the unsupervised mean squared error (uMSE) as follows:

$$\mathrm{uPSNR}(x,y) = 10\log_{10}\frac{M^2}{\mathrm{uMSE}(x,y)},$$

where M is the maximum possible pixel intensity. For each diffraction pattern at probe position (x, y), the unsupervised mean squared error (uMSE) is defined as

$$\mathrm{uMSE}(x,y) = \frac{1}{n}\sum_{i,j}\left(D_{ij}^{(x,y)} - N_{ij}^{(x,y)}\right)^2 - \frac{1}{2nK}\sum_{k=1}^{K}\sum_{i,j}\left(N_{ij}^{(x_{ak},y_{ak})} - N_{ij}^{(x_{bk},y_{bk})}\right)^2$$

where $D_{ij}^{(x,y)}$ denotes the pixel intensity at pixel position (i, j) in the denoised diffraction pattern at probe position (x, y), $N_{ij}^{(x,y)}$ is the corresponding noisy observation, n is the number of pixels in a diffraction pattern, and K is the number of neighboring pairs used for the evaluation. The second term is an estimation of the noise variance involving pairs of noisy reference patterns $N_{ij}^{(x_{ak},y_{ak})}$ and $N_{ij}^{(x_{bk},y_{bk})}$. For 4D-STEM data, we selected 8 horizontally and vertically adjacent pairs from the positions adjacent to (x, y) for the evaluation. For a special case using a Gaussian denoising filter, the second term in uMSE must be additionally weighted by $(1 - 2w_g)$, where w_g is the coefficient at the center of the filter [26]. This additional weight is introduced to account for variance shrinkage resulting from the inclusion of the target pixels in the filter.

Computation of Ground-Truth Diffraction Patterns

Theoretical diffraction patterns were computed for the four observed crystallographic orientations of the material: [0 0 1], [1 $\bar{1}$ $\bar{1}$], [2 $\bar{2}$ 1] and [2 $\bar{5}$ 1], using an electron-beam energy of 200 kV. Diffraction spots were blurred using a Gaussian filter with a width parameter of 0.18 nm^{-1} to generate realistic diffraction patterns. The generated diffraction patterns were cropped to extract the central regions of 140×140 pixels and then superimposed with a background signal. A direct beam was subsequently superimposed onto each pattern, which was modeled as a 2D Gaussian function centered at the diffraction pattern center with a standard deviation of ~0.8 nm^{-1} (12 pixels). The peak

amplitude of the Gaussian was set to approximately 10 times the maximum intensity of the original diffraction pattern to reproduce the strong intensity contrast between the direct beam and the diffraction spots observed in the experimental data. The effect of the beam stopper was simulated by applying a circular mask with a radius of eight pixels at the center of the pattern and filling the masked region with a constant background intensity. After the direct beam and beam stopper were added, the intensities of these diffraction patterns were scaled by a factor of two.

Data availability

All the data necessary to validate the conclusions in the paper are presented in the paper and/or Supplementary Information. The datasets generated during this study are available from the corresponding author upon reasonable request.

Code availability

Our developed code for the denoising and crystallographic component analysis of 4D-STEM data is available in our GitHub repository (the URL will be made available after acceptance).

References

1. Ophus, C. Four-dimensional scanning transmission electron microscopy (4D-STEM): From scanning nanodiffraction to ptychography and beyond. *Microscopy and Microanalysis* **25**, 563–582 (2019).
2. Rauch, E. F. & Véron, M. Automated crystal orientation and phase mapping in TEM. *Materials Characterization* **98**, 1–9 (2014).

3. Panova, O. et al. Orientation mapping of semicrystalline polymers using scanning electron nanobeam diffraction. *Micron* **88**, 30–36 (2016).

4. Kanomi, S., Marubayashi, H., Miyata, T. & Jinnai, H. Reassessing chain tilt in the lamellar crystals of polyethylene. *Nat. Commun.* **14**, 5531 (2023).

5. Kanomi, S., Azuma, K., Miyata, T., Toda, A. & Jinnai H., Orientation distribution and branching mechanism of lamellar crystals inside an isotactic polystyrene spherulite. *Polymer* **326**, 128335 (2025).

6. Kalinin, S. V. et al. Machine learning in scanning transmission electron microscopy. *Nat. Rev. Methods Primers* **2**, 11 (2022).

7. Krull, A., Buchholz, T., & Jug, F. Noise2Void - Learning denoising from single noisy images. *Proc. of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2129–2137 (2019).

8. Laine, S., Karras, T., Lehtinen, J. & Aila, T. High-quality self-supervised deep image denoising. *Proc. of the 33rd International Conference on Neural Information Processing Systems*, 6970–6980 (2019).

9. Sheth, D. Y., Mohan, S. & Vincent, J. L. Unsupervised deep video denoising. *Proc. of the 2021 IEEE/CVF International Conference on Computer Vision*, 1739–1748 (2021).

10. Crozier, P. A. et al. Visualizing nanoparticle surface dynamics and instabilities enabled by deep denoising. *Science* **387**, 949–954 (2025).

11. Lee, D. D. & Seung, H. S. Algorithms for non-negative matrix factorization. *Proc. of the 14th International Conference on Neural Information Processing Systems*, 535–541 (2000).

12. Nascimento, J. M., & Dias, J. M. Vertex component analysis: A fast algorithm to unmix hyperspectral data, *IEEE Trans. Geoscience and Remote Sensing* **43**, 898–910 (2005).

13. Dobigeon, N., & Brun, N. Spectral mixture analysis of EELS spectrum-images. *Ultramicroscopy* **120**, 25–34 (2012).

14. Shiga, M. et al. Sparse modeling of EELS and EDX spectral imaging data by nonnegative matrix factorization. *Ultramicroscopy* **170**, 43–59, (2016).

15. Muto, S. & Shiga, M., Application of machine learning techniques to electron microscopic/spectroscopic image data analysis. *Microscopy* **69**, 110–122 (2020).

16. Shiga, M., Ono, T., Morishita, K., Kuno, K. & Moriguchi, N., Fast computational approach with prior dimension reduction for three-dimensional chemical component analysis using CT data of spectral imaging. *Microscopy* **73**, 488–498 (2024).

17. Shi, C. et al. Uncovering material deformations via machine learning combined with four-dimensional scanning transmission electron microscopy. *npj Computational Materials* **8**, 114 (2022).

18. Kimoto, K. et al. Unsupervised machine learning combined with 4D scanning transmission electron microscopy for bimodal nanostructural analysis. *Scientific Reports* **14**, 2901 (2024).

19. Yoo, T. et al. Unsupervised machine learning and cepstral analysis with 4D-STEM for characterizing complex microstructures of metallic alloys. *npj Computational Materials* **10**, 223 (2024).

20. Kimoto, K. et al. Nonnegative matrix factorization incorporating domain specific constraints for four dimensional scanning transmission electron microscopy. *Scientific Reports* **15**, 39143 (2025).

21. Cichocki, A. & Phan, A.-H. Fast local algorithms for large scale nonnegative matrix and tensor factorizations. *IEICE Trans. Fundamentals* **92**, 708–721 (2009).

22. https://github.com/MotokiShiga/malspy

23. Seto, Y. & Ohtsuka, M. ReciPro: Free and open-source multipurpose crystallographic software integrating a crystal model database and viewer, diffraction and microscopy simulators, and diffraction data analysis tools. *J. Appl. Cryst.* **55**, 397–410 (2022).

24. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv*, arXiv:1412.6980 (2014).

25. Morales, A., M. et al. Evaluating unsupervised denoising requires unsupervised metrics, *Proc. of the 40th International Conference on Machine Learning*, 23937–23957 (2023).

26. Blu, T. & Luisier, F. The SURE-LET approach to image denoising, *IEEE Trans. Image Process.* **16**, 2778–2786 (2007).

Acknowledgments

This study was partially supported by grants from JST BOOST (grant number JPMJBS2421 [K.K.]), JST CREST (grant number JPMJCR2435 [M.S.], JPMJCR1993 [H.J.], JPMJCR24S3 [H.J.], JPMJCR25Q3 [H.J.], and JPMJCR25S2 [T.M.]), and JSPS KAKENHI (grant number JP23H02017 [T.M.], JP24K01153 [T.M.], JP25K23517 [S.K]).

Author contributions

All the authors designed the study. K.K and M.S. developed the denoising and analysis methods. K.K created the computational codes. K.K and M.S. drafted of the manuscript. All the authors contributed to the writing and editing of the manuscript.

Competing interests

The authors declare that they have no competing interests.

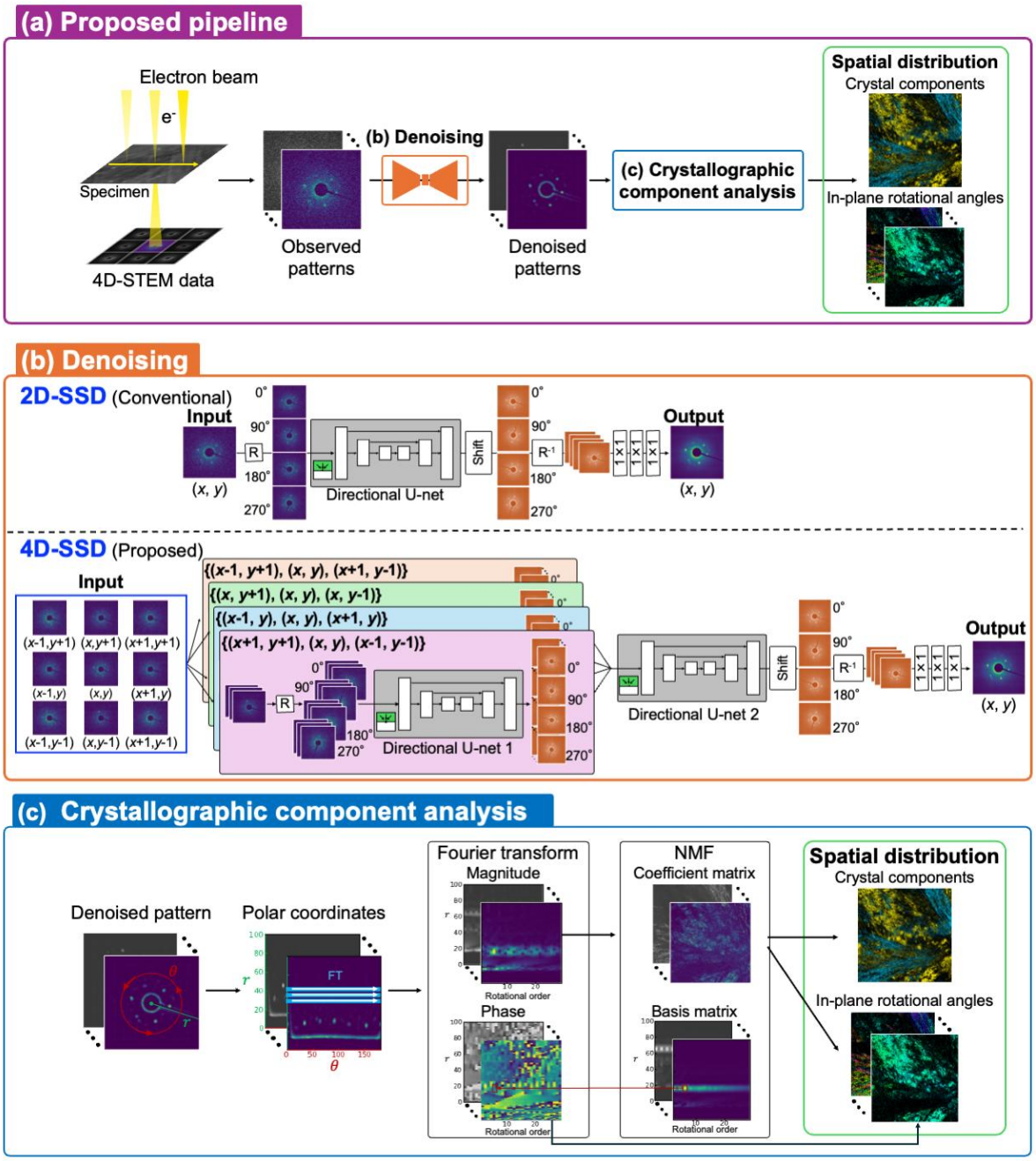

Fig. 1. Overview of the proposed pipeline.

(a) Schematic illustration of the proposed pipeline. Diffraction patterns in 4D-STEM data are first denoised using a neural network model (4D-SSD). The denoised data are then analyzed by unsupervised method to obtain the spatial distribution of the crystallographic components and their in-plane rotational angles. **(b)** Self-supervised denoising models (2D-SSD and 4D-SSD) that equips a blind-spot architecture. The architecture is performed based on directional U-Nets, rotation and shift operators. The detailed mechanism is shown in Supplementary Fig. 1. **(c)** Crystallographic component analysis, which consists of polar coordinate transformation, Fourier transform, and nonnegative matrix factorization.

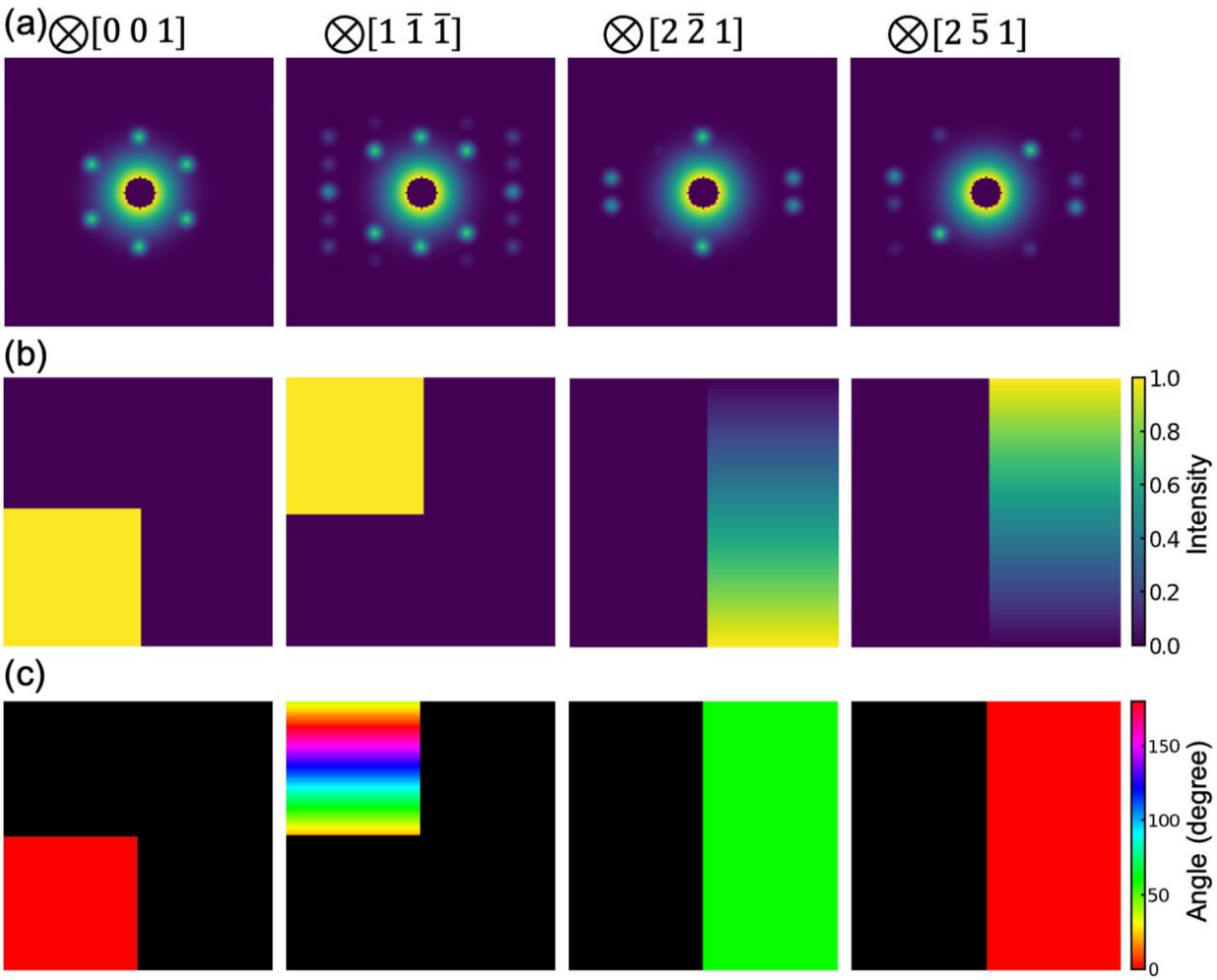

Fig. 2. Synthetic 4D-STEM data.
(a) Simulated diffraction patterns. (**b)** The real-space intensity distributions of each crystallographic component. **(c)** The distributions of the in-plane rotational angles of each crystallographic component.

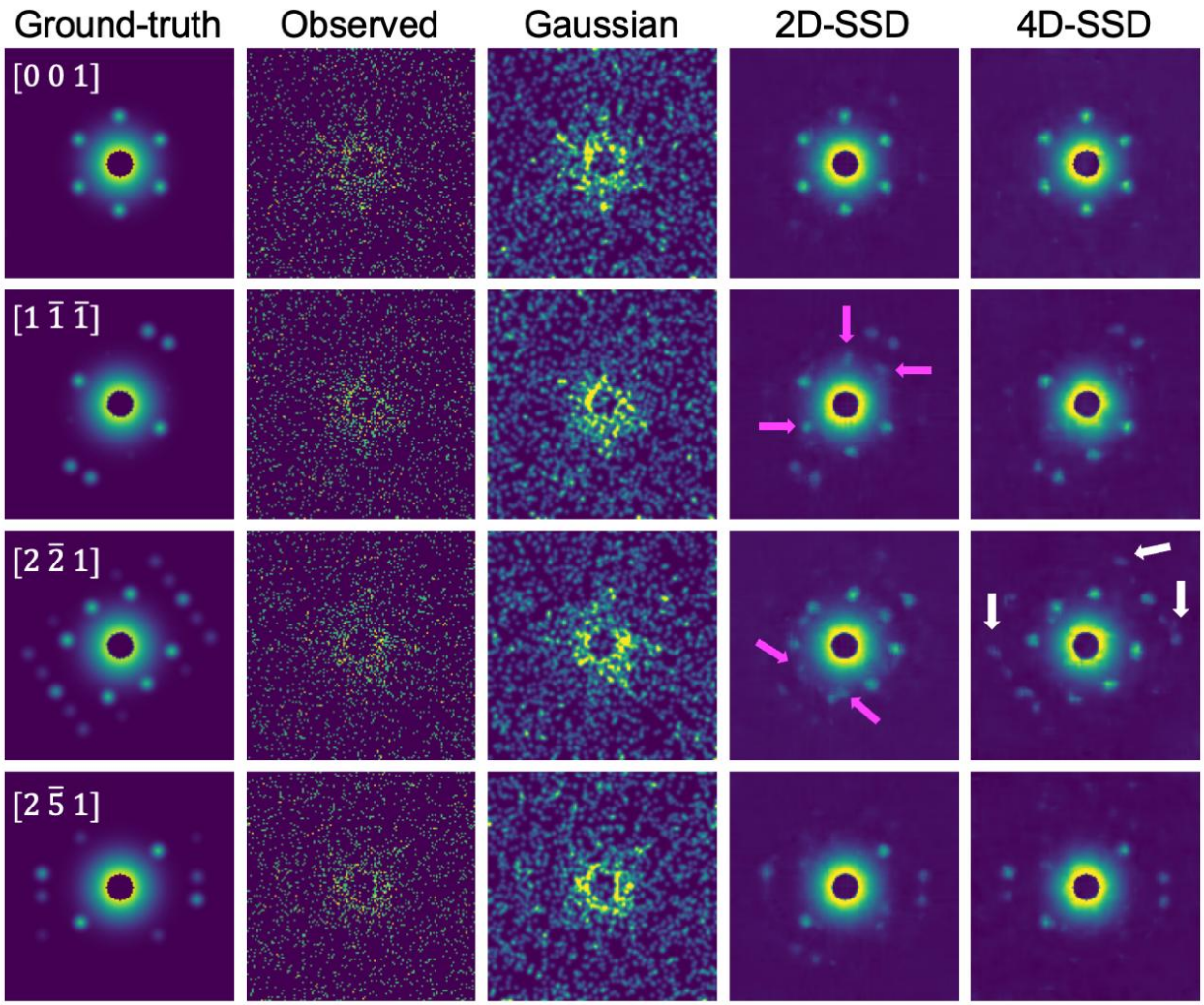

Fig. 3. Denoising results for the diffraction patterns in the synthetic dataset ($w_s = 0.2$) The ground-truth, observed, and denoised (Gaussian filter, 2D-SSD and 4D-SSD (the proposed method) patterns. The intensity range of each image was normalized for visualization. The magenta arrows indicate artifacts generated by 2D-SSD and 4D-SSD. White arrows indicate weak spots detected only by 4D-SSD.

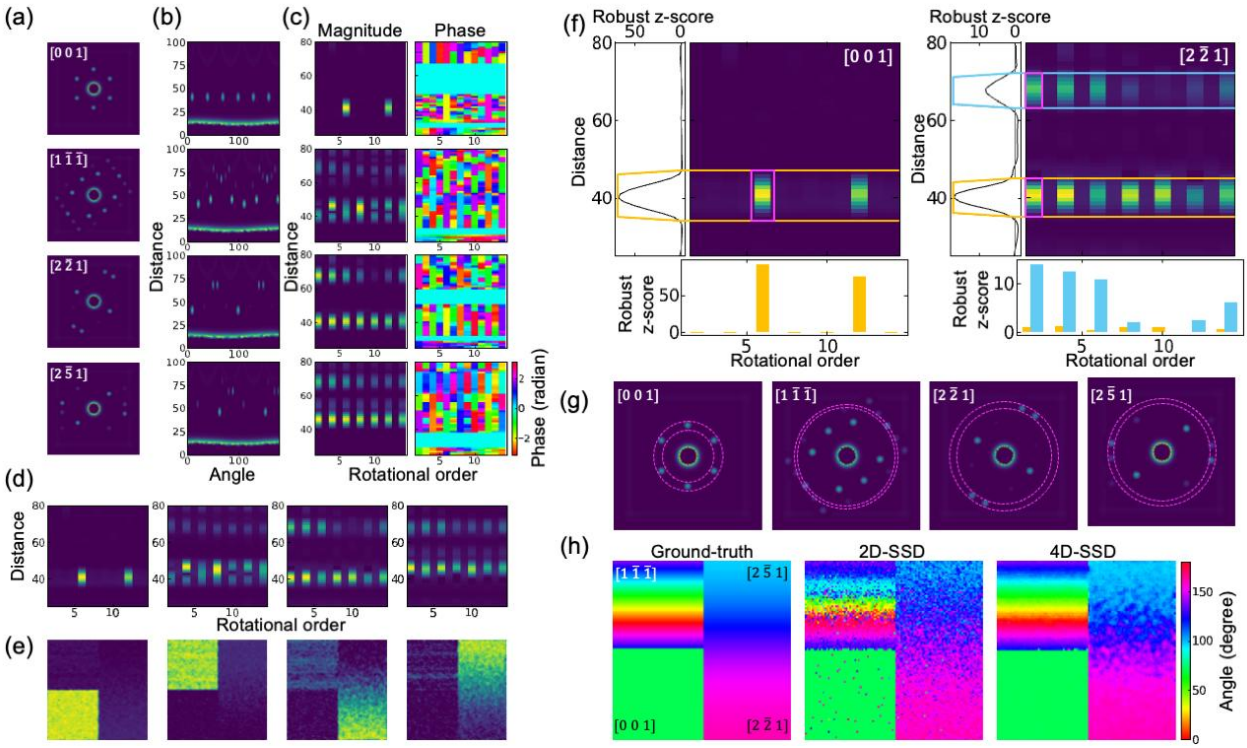

Fig. 4. Analysis results of crystal microstructural components

(a) Diffraction patterns of the ground-truth components: [0 0 1], [1 $\bar{1}$ $\bar{1}$], [2 $\bar{2}$ 1] and [2 $\bar{5}$ 1]. **(b)** Polar coordinate representations of the patterns in (a). **(c)** Magnitudes and phases of 1D Fourier transformation along the angular axis for the polar representations. **(d)** Fourier magnitudes of the components identified by NMF from pattern denoised by 4D-SSD. **(e)** Real-space distributions of the component coefficients estimated by NMF. **(f)** Two examples of detected peaks (magenta rectangles) with Fourier magnitudes estimated by the two-step detection procedure. The top-right panels show that Fourier magnitudes of components estimated by NMF. The left panels show the summations along the angular (horizontal) axis, where the detected radial distances containing peaks are highlighted with orange and light-blue rectangles. The bottom panels show the summations along the radial-distance (vertical) axis within the radial distances detected by the first step. **(g)** The areas indicated by magenta dashed circles are identified by the peak detection and used for the in-plane orientation estimation. **(h)** In-plane rotational angle distributions of the ground-truth data and those estimated via 2D-SSD and 4D-SSD.

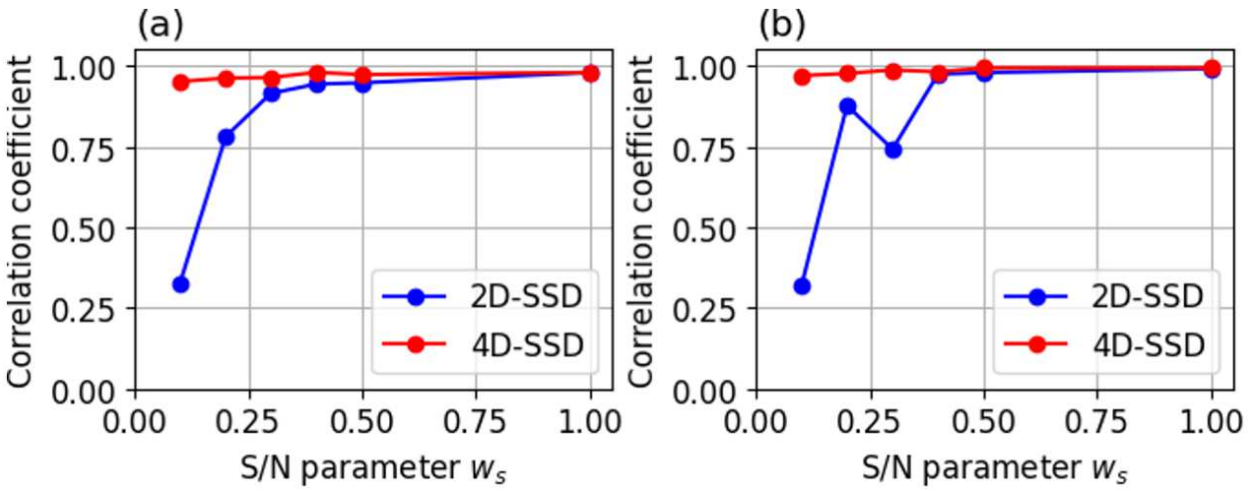

Fig. 5. Performance comparison using the synthetic datasets

Correlation coefficients between estimations by NMF and the ground-truth data for **(a)** crystallographic components and **(b)** in-plane rotational angles. The parameter w_s refers to the S/N ratio in the synthetic 4D-STEM data and is inversely related to the noise level.

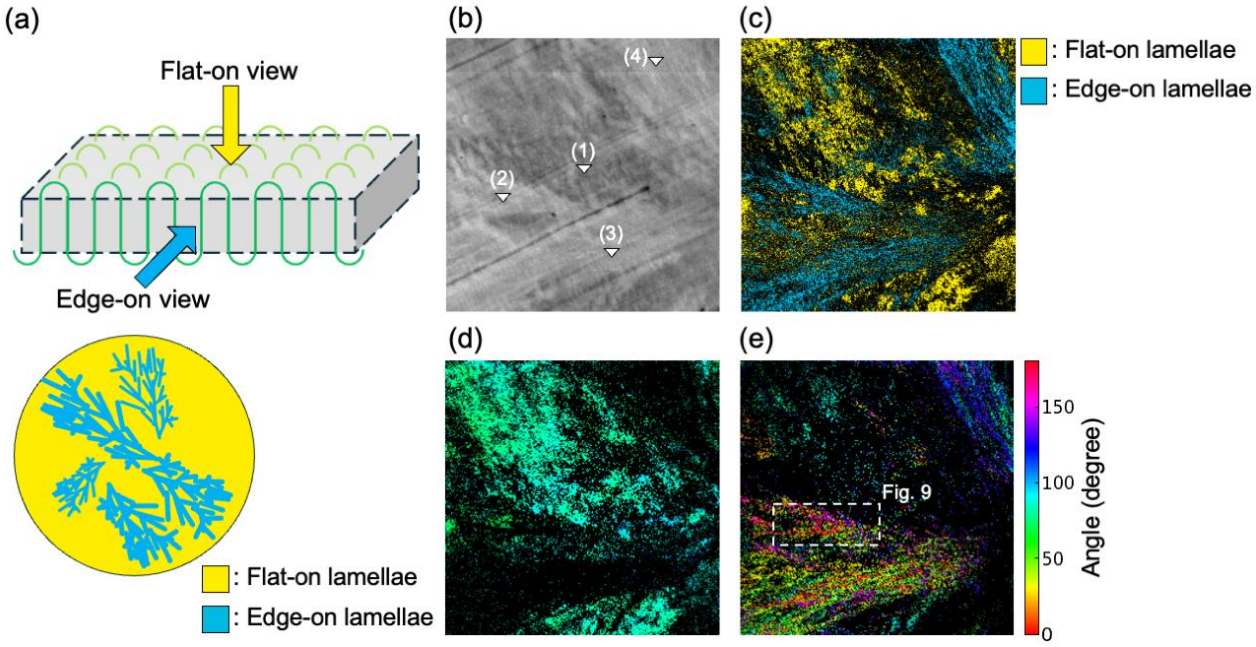

Fig. 6. Real experimental data and crystallographic analysis result

(a) Structure of an isotactic polystyrene (iPS) spherulite. **(b)** ADF-STEM image reconstructed from the 4D-STEM data. Numbers (1) – (4) indicate the observation positions of the diffraction patterns in **Fig. 7**. **(c)** Spatial distribution of major components identified by our NMF. **(d, e)** Spatial distributions of in-plane rotational angles in flat-on (d) and edge-on lamellae (e) estimated by our proposed pipeline.

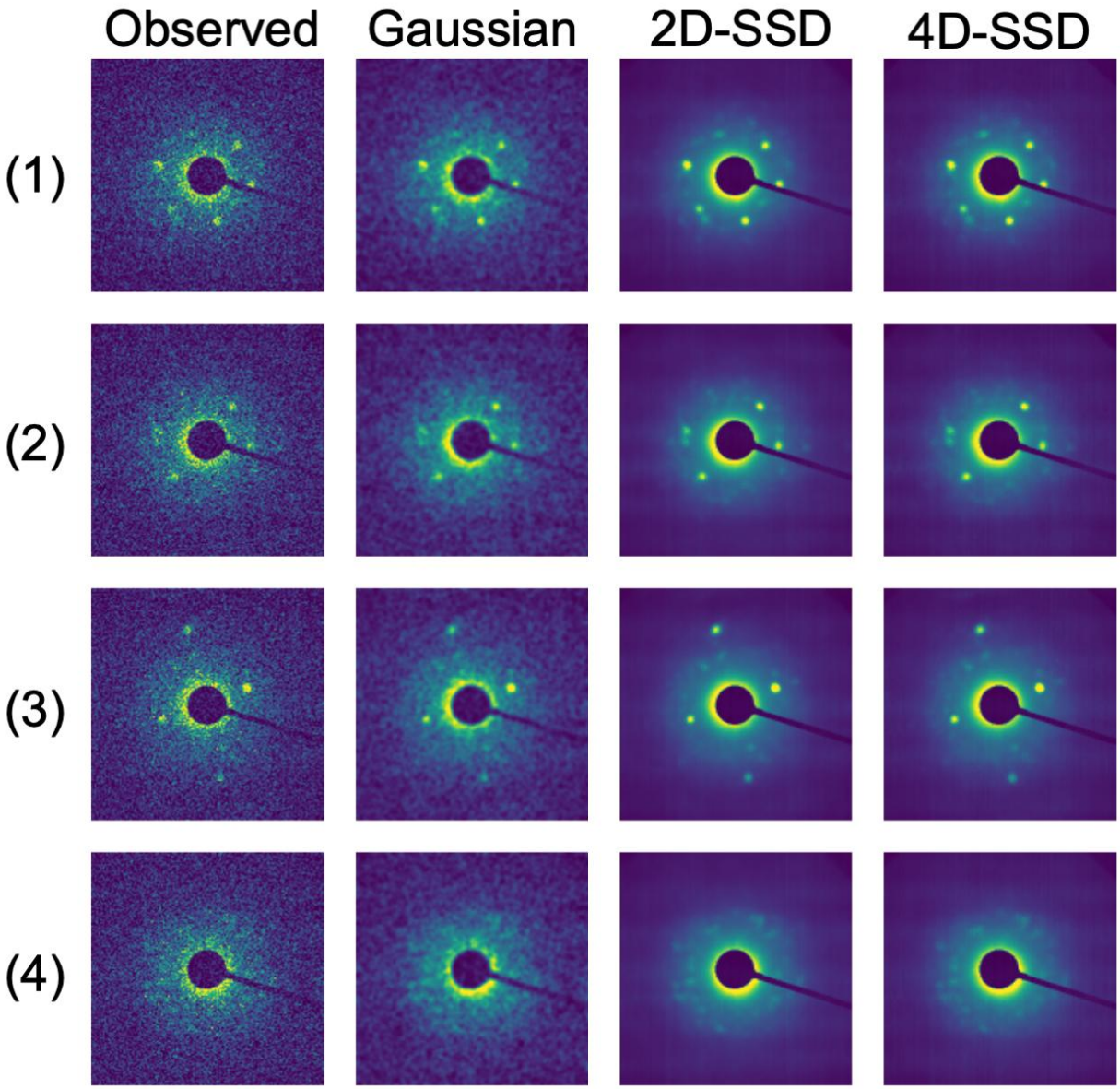

Fig. 7. Denoising results for the diffraction patterns in the real dataset
The observed and denoised (Gaussian filter, 2D-SSD, and 4D-SSD (the proposed method)) diffraction patterns are presented. The intensity range of each image was normalized for visualization. The observed positions (1)–(4) are indicated in Fig. 6b. According to Fig. 6b**,** position (1) was estimated as a flat-on lamella while positions (2) and (3) were identified as edge-on lamellae. Position (4) was estimated as a non-crystal region.

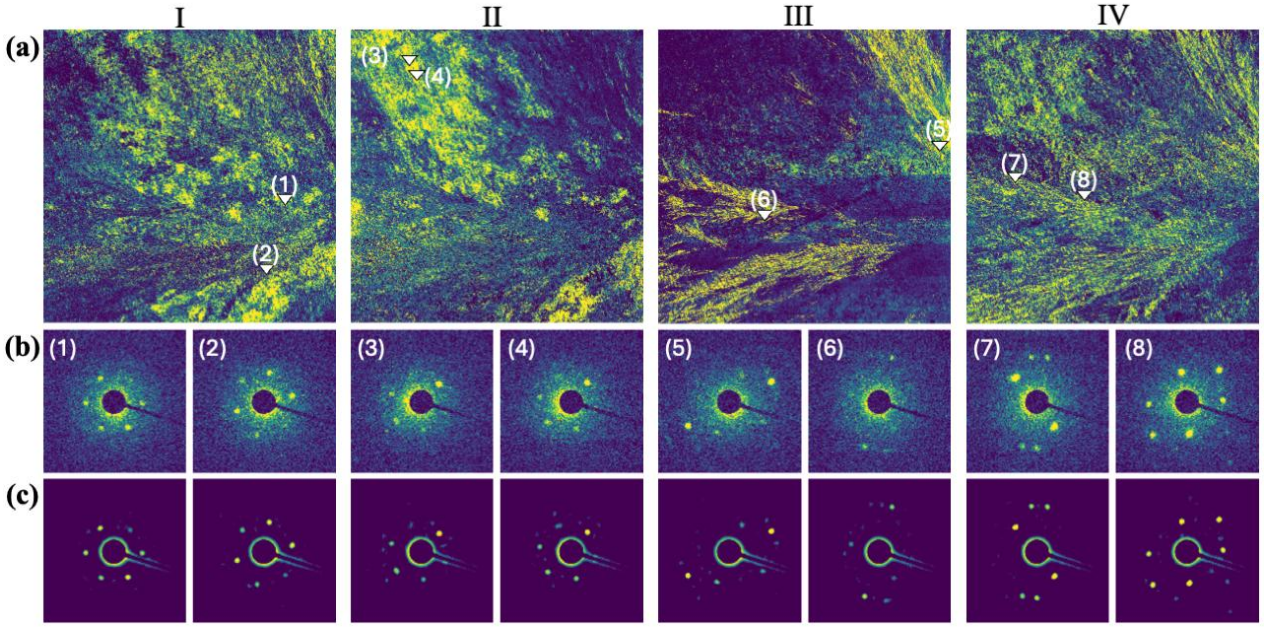

Fig. 8. Spatial distribution of components identified by NMF from real data
(a) Component distribution identified as flat-on lamellae (I and II) and edge-on lamellae (III and IV). The flat-on lamellae correspond to the pattern from the $[0\ 0\ 1]$ orientation while the edge-on lamellae correspond to the pattern from either the $[1\ \bar{1}\ \bar{1}]$, $[2\ \bar{2}\ 1]$ or $[2\ \bar{5}\ 1]$ orientation. **(b)** The diffraction patterns observed at the positions (1)–(8) indicated in the top panels. **(c)** Denoised diffraction pattens. The patterns were processed using the proposed denoising method, followed by a high-pass filter.

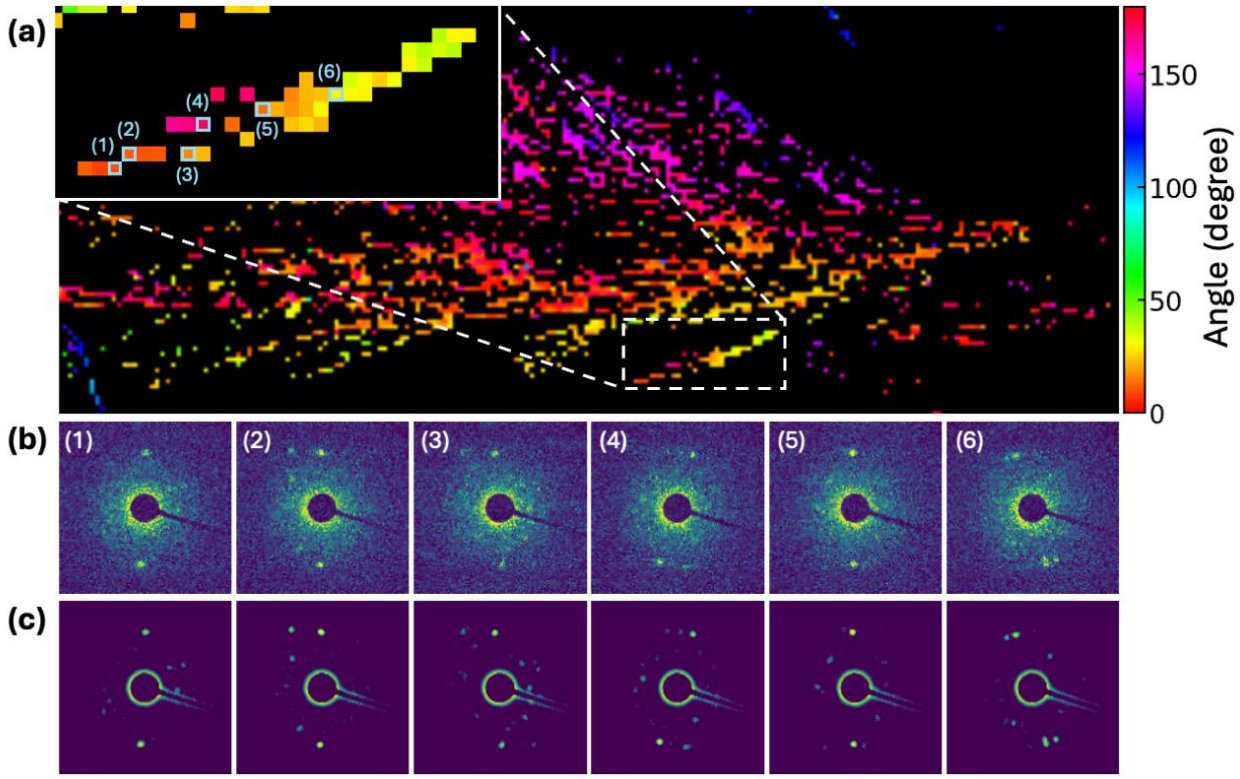

Fig. 9. (a) Spatial distribution of in-plane rotational angles of component III (an edge-on lamellar region) indicated in Fig. 6d. **(b)** The diffraction patterns observed at the spatial points indicated in the top panels. **(c)** Denoised diffraction patterns.

Table 1. Denoising performance metrics.

Mean ± standard deviation of PSNRs and uPSNRs. The PSNRs were computed with respect to the ground-truth images assumed in the synthetic data, whereas the uPSNRs were computed without the ground-truth data.

	Synthetic data ($w_s = 0.2$)		**Real data**
	PSNR [dB]	**uPSNR [dB]**	**uPSNR [dB]**
Observed	-1.06 ± 0.24	-	-
Gaussian filter	11.27 ± 0.20	9.38 ± 0.57	37.49 ± 0.52
2D-SSD	13.47 ± 0.28	16.52 ± 4.00	50.56 ± 3.45
4D-SSD	13.56 ± 0.22	16.56 ± 4.09	50.74 ± 3.56

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- manus4DSTEMSI.pdf