

SUPPLEMENTARY INFORMATION

Unsupervised machine learning for automated crystal orientation mapping on noisy 4D-STEM data

Kai Kamijo¹, Motoki Shiga^{1,2,3,4,*,†}, Shusuke Kanomi^{5,6}, Tomohiro Miyata⁵,
and Hiroshi Jinnai⁵

*Corresponding author

†motoki.shiga.b4@tohoku.ac.jp

¹ Graduate School of Information Sciences, Tohoku University, Sendai, Japan

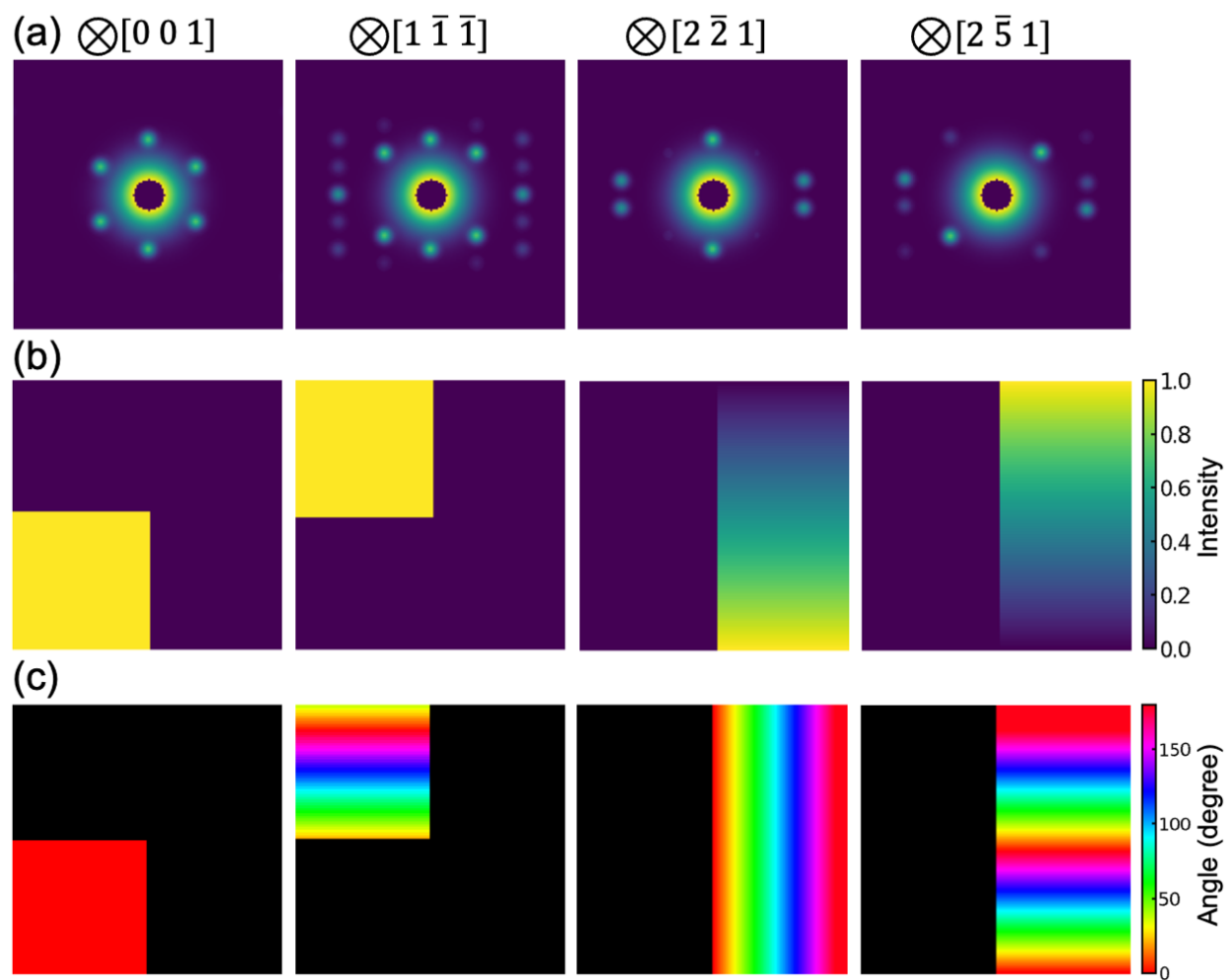
² Unprecedented-scale Data Analytics Center, Tohoku University, Sendai, Japan

³ Center for Basic Research on Materials, National Institute for Materials Science,
Tsukuba, Japan

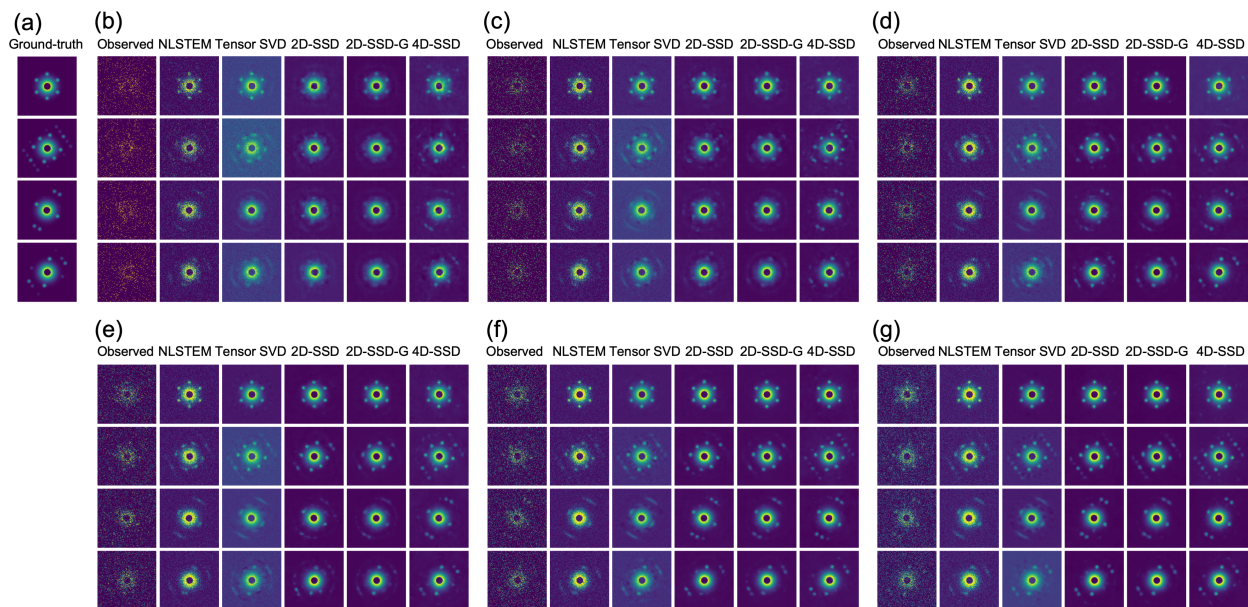
⁴RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

⁵Institute of Multidisciplinary Research for Advanced Materials, Tohoku University,
Sendai, Japan

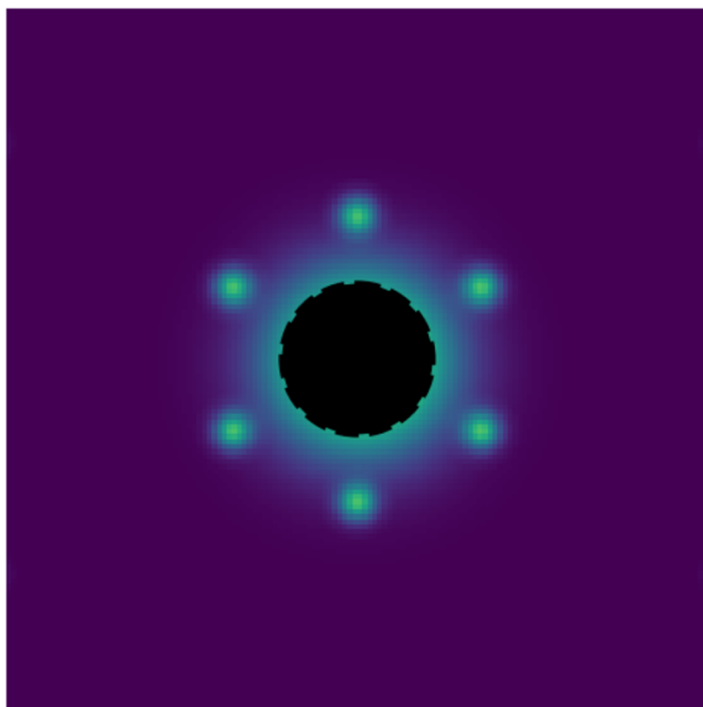
⁶Science & Innovation Center, Mitsubishi Chemical Corporation, Yokohama, Japan



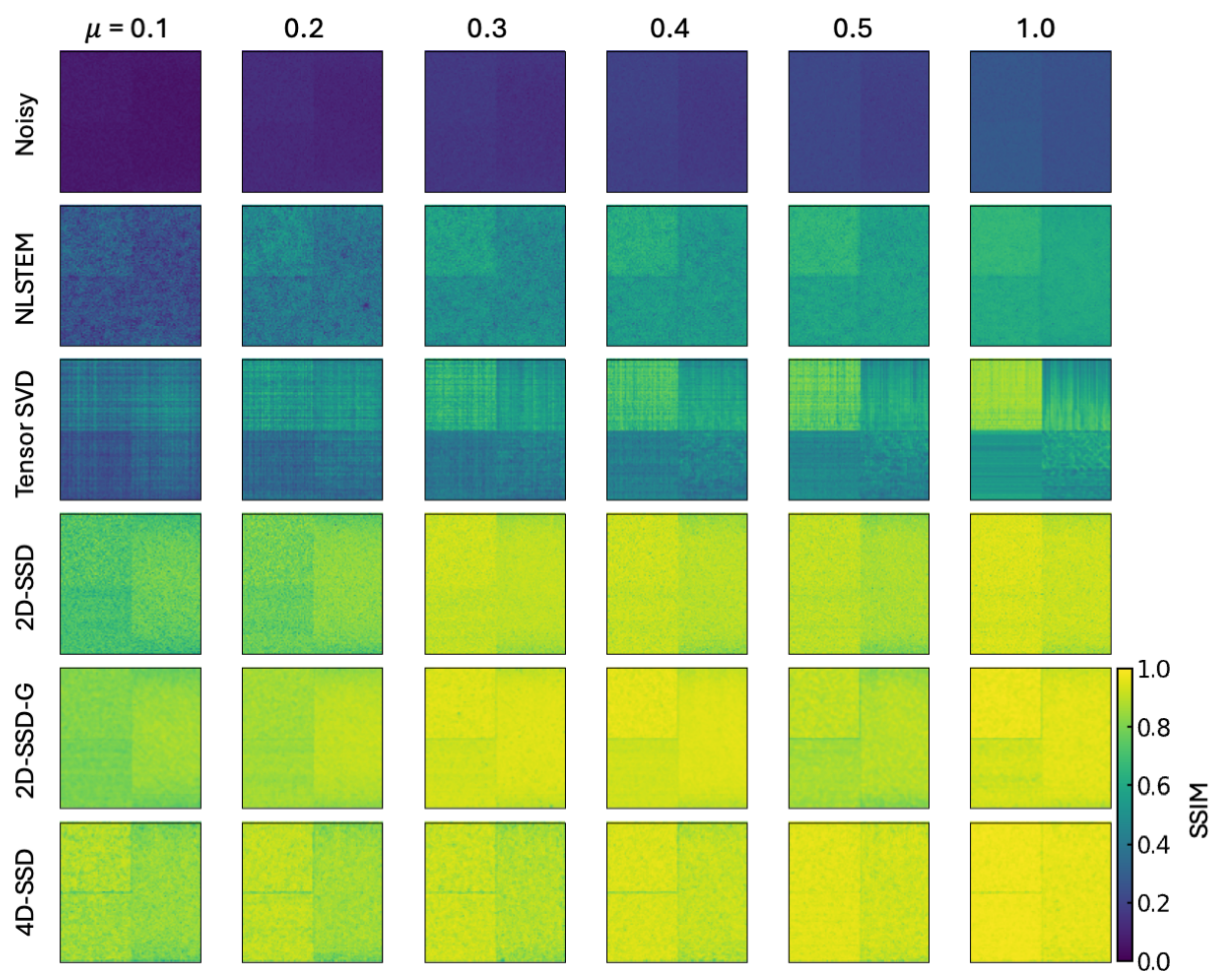
Supplementary Fig. 2. Synthetic 4D-STEM data. (a) Simulated diffraction patterns. (b) The real-space intensity distributions of each crystallographic component. (c) The distributions of the in-plane rotational angles of each crystallographic component.



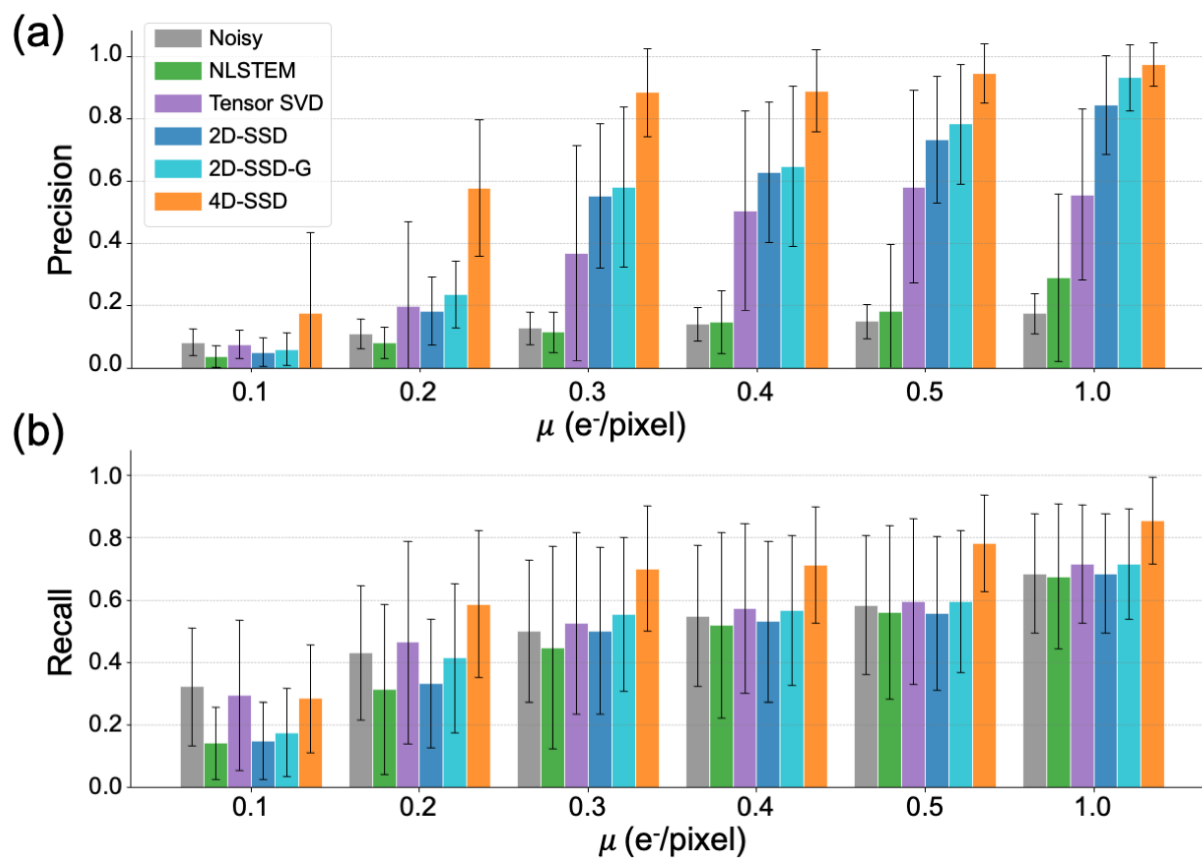
Supplementary Fig. 3. Denoised patterns obtained by various compared methods for synthetic data. (a) the ground-truth patterns and denoising results from synthetic data with beam-dose level of (b) $\mu = 0.1$, (c) 0.2, (d) 0.3, (e) 0.4, (f) 0.5, and (g) 1.0. The intensity range of each image was normalized for visualization.



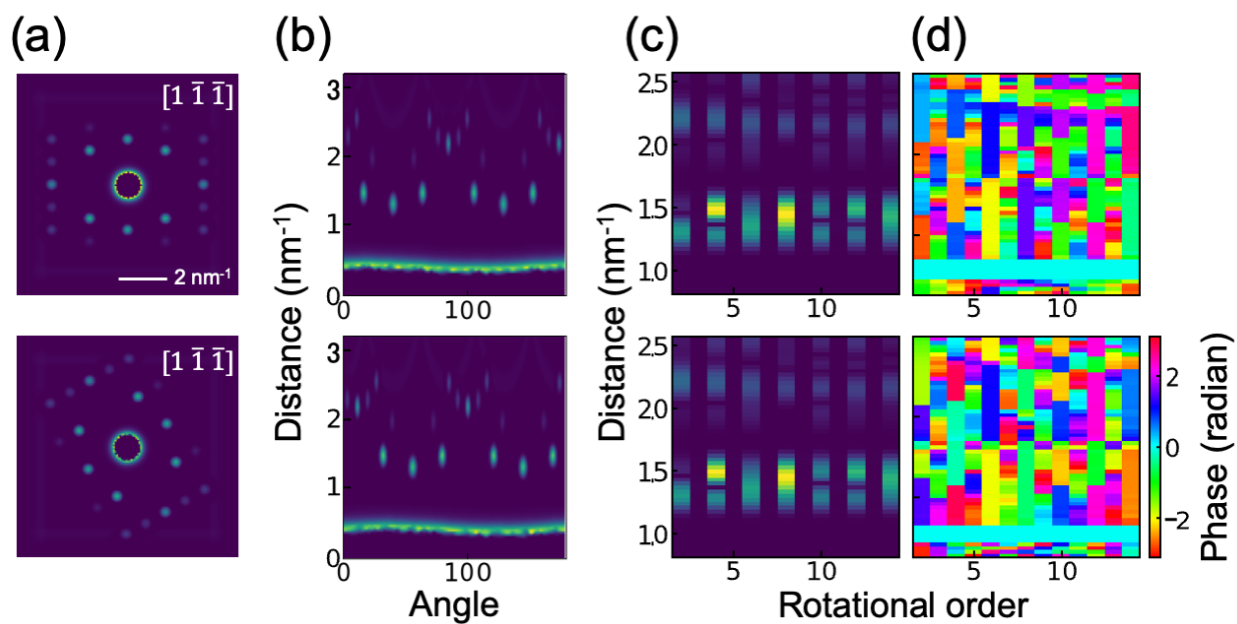
Supplementary Fig. 4. Masked region applied during SSIM computation. The masked region is indicated by the central black circle. This region contains only the direct beam and lacks any crystallographic information.



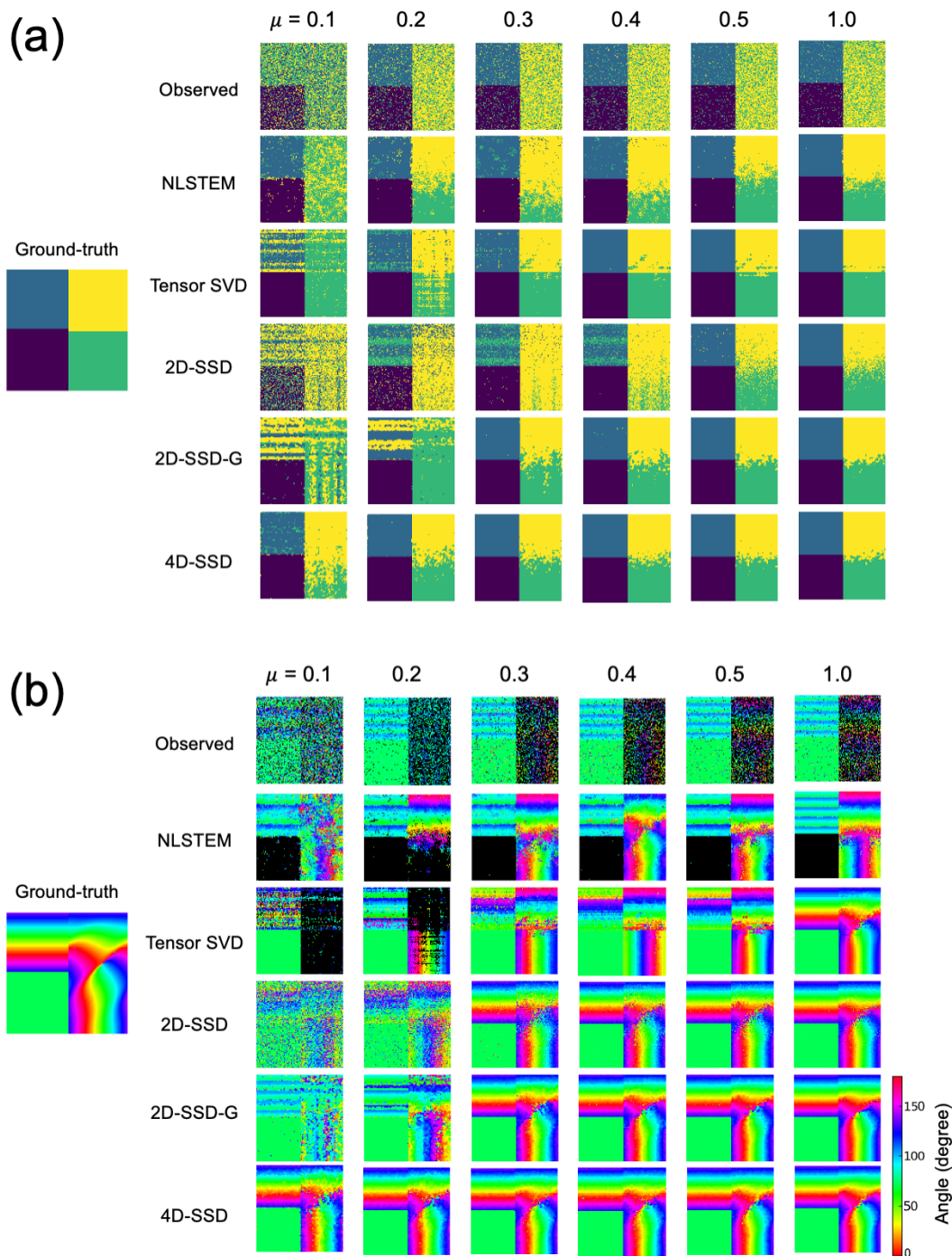
Supplementary Fig. 5. Spatial maps of SSIMs between denoised diffraction patterns and the ground-truth patterns. The variable μ indicates the beam dose level. Brighter pixel regions exhibited better denoising performance.



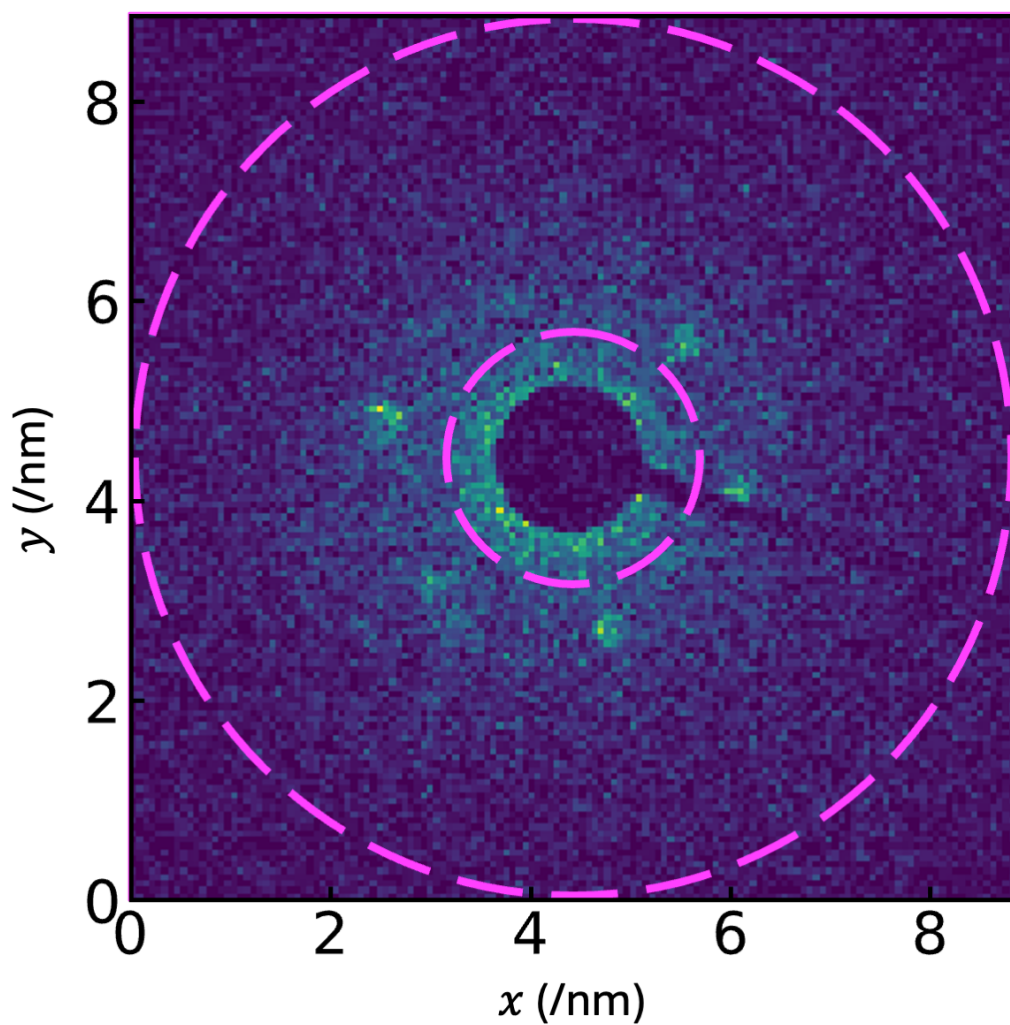
Supplementary Fig. 6. Performance metrics of diffraction spots recovery from denoised patterns: (a) precision and (b) recall.



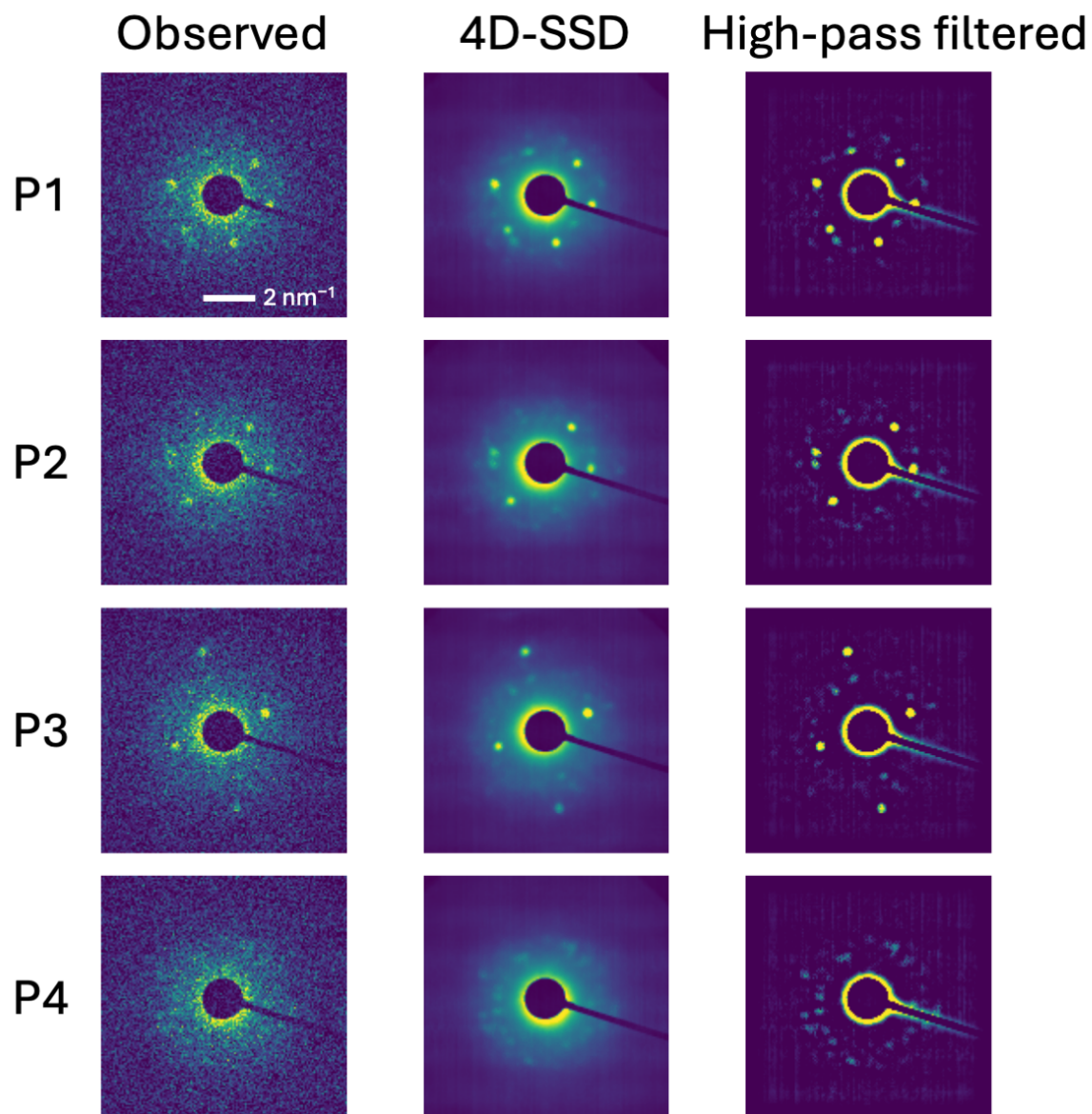
Supplementary Fig. 7. Data transformations of diffraction patterns from in-plane rotated crystals in component $[1 \bar{1} \bar{1}]$. (a) Diffraction patterns with rotational angles of $\theta = 0^\circ$ (top) and 60° (bottom). (b) Polar representations of the diffraction patterns. (c, d) Magnitudes (c) and phases (d) of the Fourier transform of the polar representations.



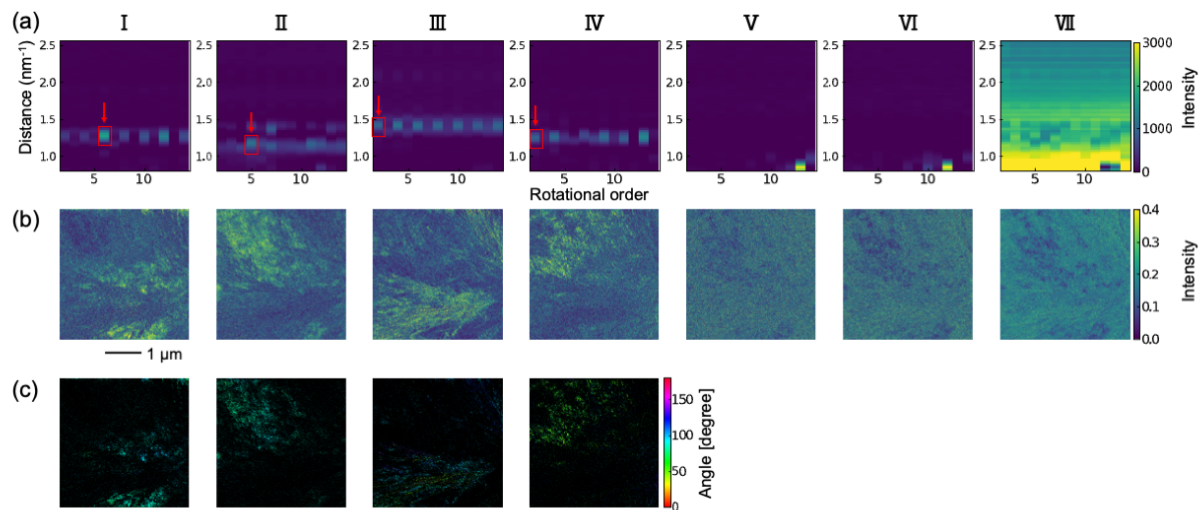
Supplementary Fig. 8. Crystallographic component analysis results for synthetic data with various beam-dose levels. (a) Spatial distributions of components denoised by 4D-SSD and conventional methods. For visualization, each position was assigned to a single component with the maximum intensity. (b) Spatial distributions of the in-plane rotational angles estimated from the Fourier phase images. The variable μ indicates the dose level.



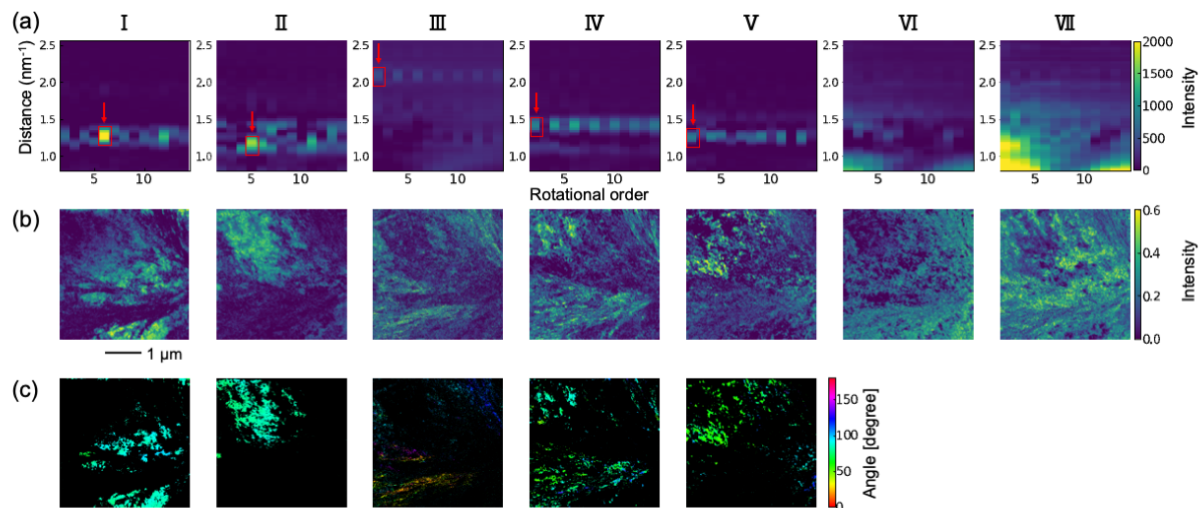
Supplementary Fig. 9. The annular range used for generating the ADF-STEM image (Fig. 6b) from the 4D-STEM data. The range is highlighted by the magenta dashed circles.



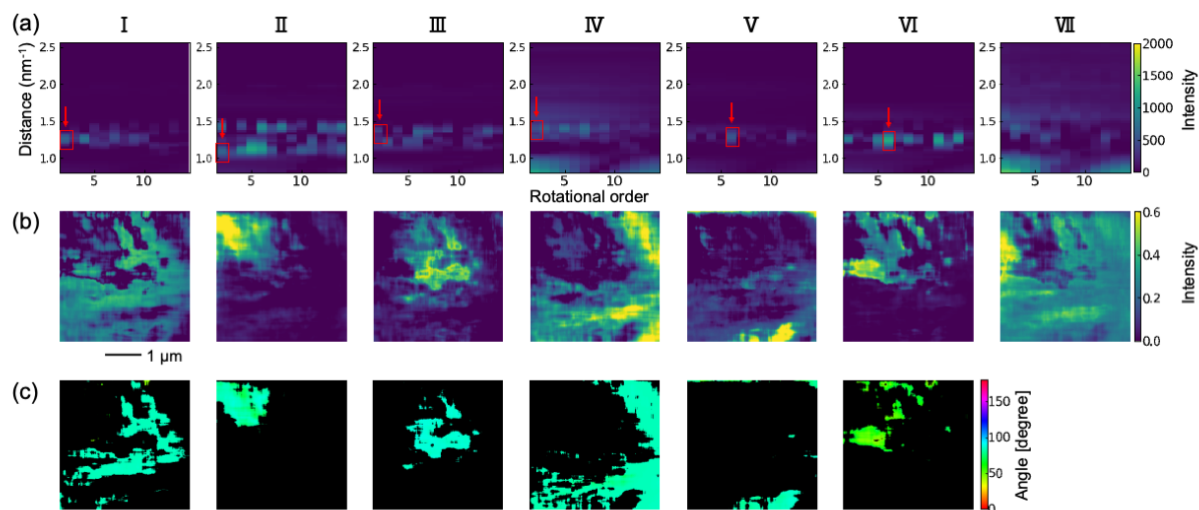
Supplementary Fig. 10. Diffraction patterns for experimental 4D-STEM data after denoising followed by the high-pass filtering. Observation positions P1–P4 are indicated in Fig. 6b.



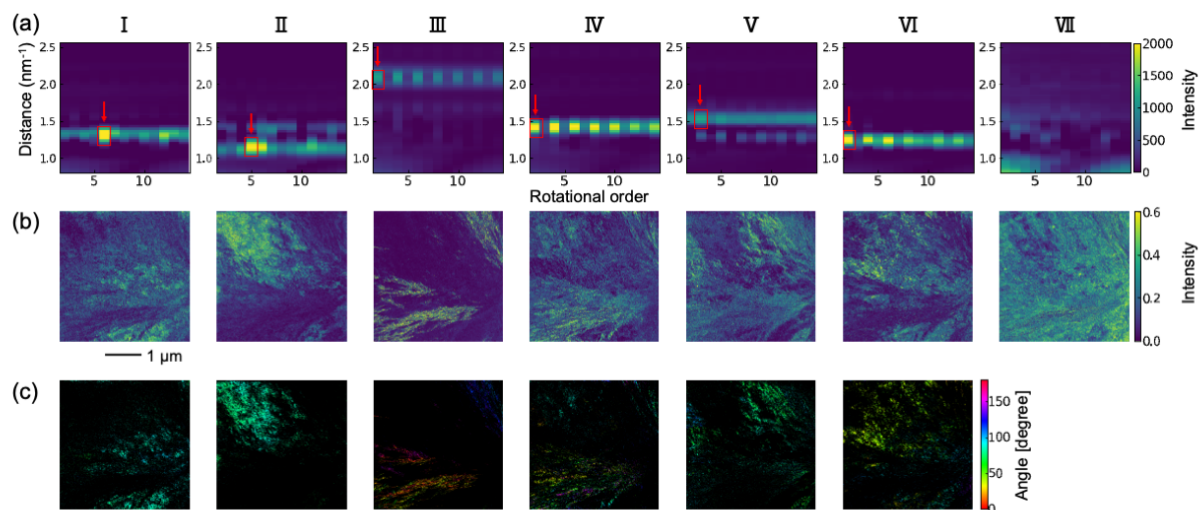
Supplementary Fig. 11. Crystallographic component analysis results obtained by NMF for the experimental data (raw data). (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



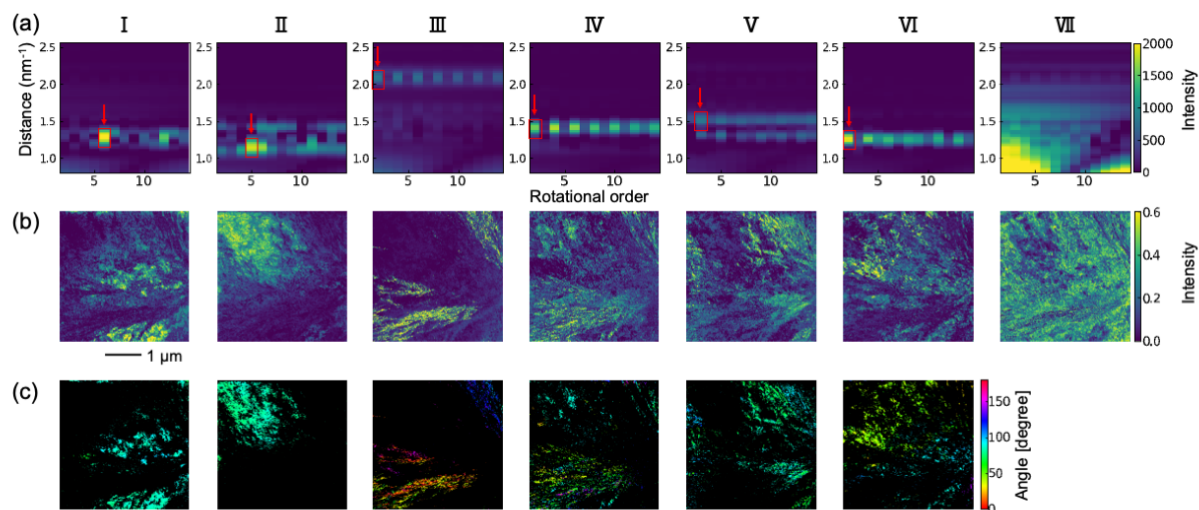
Supplementary Fig. 12. Crystallographic component analysis results obtained by NMF for the experimental data denoised by NLSTEM. (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



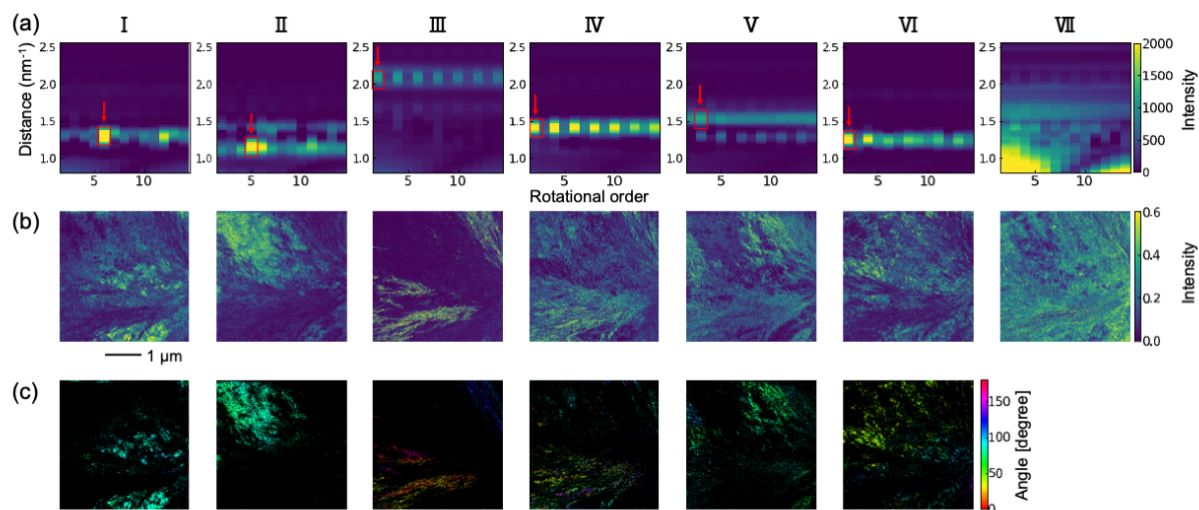
Supplementary Fig. 13. Crystallographic component analysis results obtained by NMF for the experimental data denoised by Tensor SVD. (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



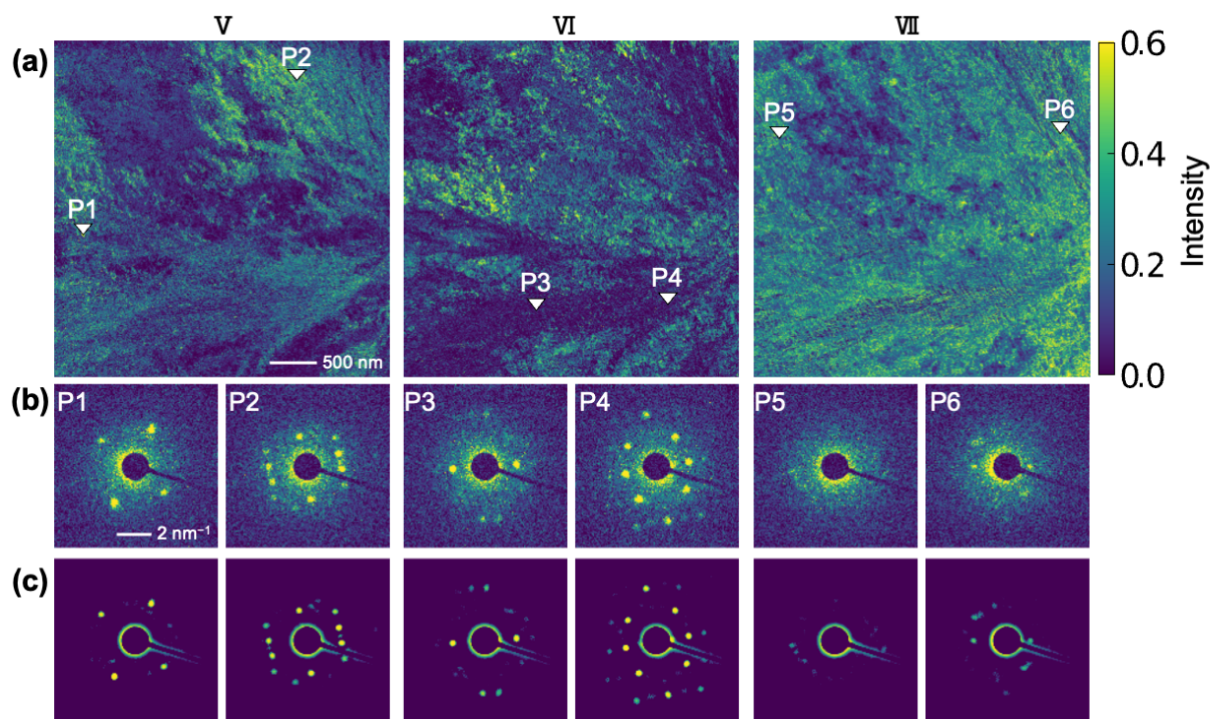
Supplementary Fig. 14. Crystallographic component analysis results obtained by NMF for the experimental data denoised by 2D-SSD. (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



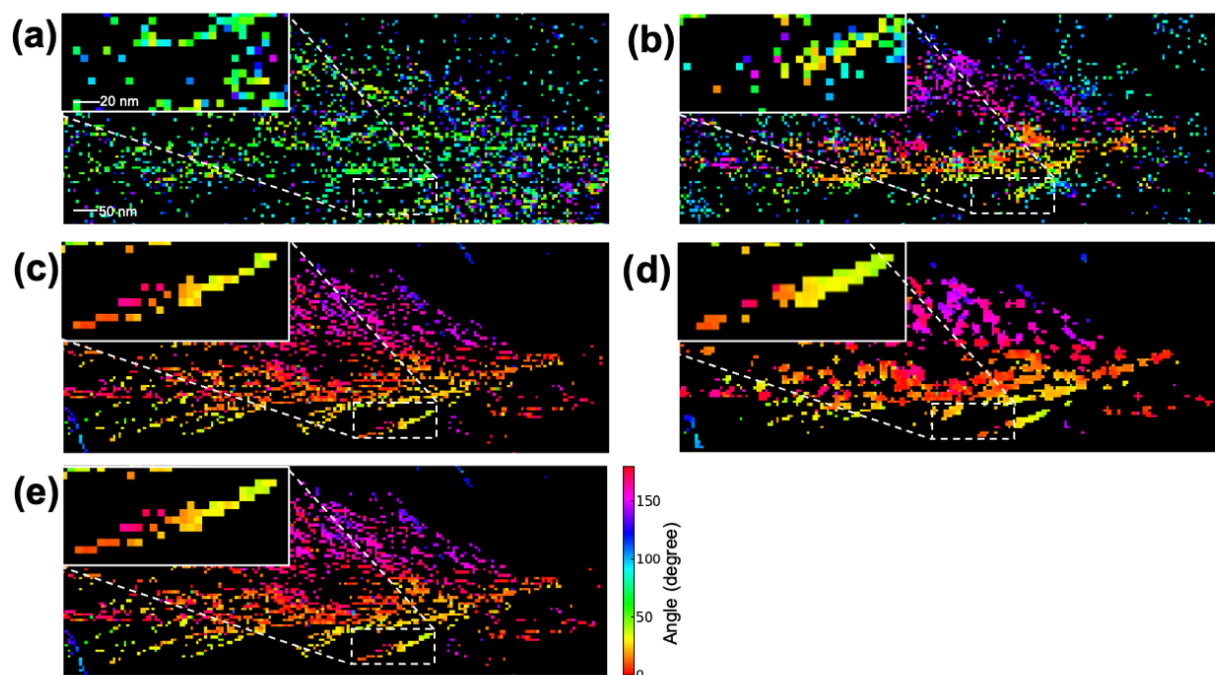
Supplementary Fig. 15. Crystallographic component analysis results obtained by NMF for the experimental data denoised by 2D-SSD-G. (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



Supplementary Fig. 16. Crystallographic component analysis results obtained by NMF for the experimental data denoised by 4D-SSD. (a) Fourier magnitudes of the crystal components identified using NMF. (b) Spatial maps of the components. (c) Spatial map of the in-plane rotational angles estimated using our pipeline. The peak positions indicated by red arrows in (a) were used for angle estimation. At each position in these maps, the angle of the component corresponding to the maximum intensity is displayed.



Supplementary Fig. 17. Spatial distribution of other components identified by NMF from real data denoised by 4D-SSD. (a) Spatial distribution of components not identified as flat-on or edge-on lamellae. (b) Diffraction patterns observed at the positions P1–P6 indicated in the top panels. (c) Denoised diffraction patterns. The patterns were processed using the proposed denoising method, followed by high-pass filtering.



Supplementary Fig. 18. Spatial maps of estimated in-plane rotational angles of component III (an edge-on lamellar region indicated in Fig. 6b). These maps represent the analysis results obtained from (a) the raw noisy data and the data denoised using (b) NLSTEM, (c) 2D-SSD, (d) 2D-SSD-G, and (e) 4D-SSD. Map (e) is the same as that shown in Fig. 10a.