



OPEN Unsupervised learning of spatially-resolved ARPES spectra for epitaxially grown graphene via non-negative matrix factorization

Masaki Imamura  & Kazutoshi Takahashi

This study proposed an unsupervised machine-learning approach for analyzing spatially-resolved ARPES. A combination of non-negative matrix factorization (NMF) and k-means clustering was applied to spatially-resolved ARPES spectra of the graphene epitaxially grown on a SiC substrate. The Dirac cones of graphene were decomposed and reproduced fairly well using NMF. The base and activation matrices obtained from the NMF results reflected the detailed spectral features derived from the number of graphene layers and growth directions. The spatial distribution of graphene thickness on the substrate was clearly visualized by the clustering using the activation matrices acquired via NMF. Integration with k-means clustering enables clear visualization of spatial variations. Our method efficiently handles large datasets, extracting spectral features without manual inspection. It offers broad applicability beyond graphene studies to analyze ARPES spectra in various materials.

Keywords Photoemission spectroscopy, Machine learning, Non-negative matrix factorization, Graphene

Angle-resolved photoemission spectroscopy (ARPES), is a powerful tool for investigating material properties by directly observing band dispersion in energy and momentum spaces. Recently, advancements in focusing optics have led to the development of spatially-resolved ARPES, significantly decreasing the beam size to a few micrometers or several ten nanometers^{1–3}. These advances have enabled the measurement of local electronic structures previously difficult to measure, such as tiny flake samples and the local electronic structures for inhomogeneous materials, which consist of multiple domains with distinct electronic structures^{4–11}.

In conventional spatially-resolved ARPES measurements, obtained data are visualized by creating spatial maps with the integrated intensity of specific energy and momentum windows of the spectra at each measurement point. Subsequently, representative spectra with distinct features were analyzed in detail at some selected points. This process involves arbitrary decisions in the selection of integrated windows and the extraction of characteristic points based on knowledge and experience. However, as the spot size of the light source reduced to the nanometer scale owing to the recent development of equipment, the amount of data generated through grid scanning with automated equipment and data acquisition increased, making it practically impossible to examine all data in detail. Consequently, there is a risk of overlooking fine electronic structures, such as weak bands or shoulder structures that appear in small regions. Therefore, it is crucial to apply a machine learning-based approach to reduce the workload and extract features for handling large volumes of data. Recently, various machine learning approaches, such as PCA¹², k-means method^{13,14}, and convolutional neural network^{15,16}, have been reported for analyzing the ARPES data. Non-negative matrix factorization (NMF) is a widely-used algorithm for dimensionality reduction and feature extraction that decomposes a matrix into the product of two matrices, such that all three matrices have no negative elements^{17–21}. The resulting matrices served as basis and activation matrices. Because of this non-negativity, the physical meanings of the resulting matrices can be readily interpreted. The original matrix was reproduced using a linear combination of the basis and activation matrices. Applying NMF to spatially-resolved ARPES spectra is expected to decompose and group vast amounts of spectra to elucidate the underlying physics with less effort than in the conventional analysis.

In this study, we demonstrated an unsupervised approach for obtaining spatially-resolved photoemission spectra. NMF was performed on spatially-resolved ARPES spectra of graphene grown on a SiC substrate. The experimentally obtained spectra were decomposed into several features and well reproduced using the basis and activation matrices. The basis and activation matrices obtained using NMF revealed the spectral feature

Synchrotron Light Application Center, Saga University, 1 Honjo, Saga 840-8502, Japan. ✉email: mimamura@cc.saga-u.ac.jp

modification depending on the measurement location. In addition, the distribution of graphene on the substrate was visualized by classification using activation vectors obtained with NMF.

Methods

The dataset used in this study included a spatial map of the ARPES spectra of a graphene sample grown on a SiC substrate. Graphene was grown on the n-doped 6 H-SiC(0001) using a face-to-face method²². The miscut angle of the wafers used in this study was $< 0.05^\circ$. The thickness of the graphene layer was controlled by varying the annealing temperature. The thermal distribution of the substrate during annealing led to the spatial distribution of the thickness of the resulting graphene. The ARPES measurements were performed at the Saga University beamline (BL13) of the SAGA Light Source²³, and the incident energy for ARPES was 40 eV. Thickness-dependent dispersions of monolayer (ML), bilayer (BL), and trilayer (TL) graphene around the K point, consistent with a previous report^{11,24–26}, were used for the analysis.

The ARPES maps were acquired by sampling approximately 400 points in an (x, z) grid spanning the sample. All ARPES spectra were indexed simultaneously with the measured locations (x and z). The measurement locations during spatially-resolved measurement were changed with steps of 0.5 and 0.2 mm for the x and z directions, respectively. These steps were sufficiently large compared with the spot size of incident beam (approximately $100 \mu\text{m} \times 150 \mu\text{m}$). The observed ARPES images were flattened into one-dimensional vectors and stacked to convert them into a two-dimensional matrix for applying NMF. As a preprocessing step, a binning of spectra to suppress the data size and the effect of the spectral noise were carried out. The photoelectron intensities were normalized. The background intensity was removed from the spectra to avoid reproducing the background noise in the NMF. NMF and clustering using the k-means method were performed using the scikit-learn library, distributed for machine learning in Python²⁷. In NMF, the matrix constructed from experimentally obtained spectra is decomposed into the product of a basis matrix W and an activation matrix H. Each column of W represents the frequent patterns within the dataset, while each row of H corresponds to the contribution of these patterns to the original spectra. In the subsequent sections, we will refer to the columns or rows extracted from the W and H matrices as vectors. The initial W and H matrices of NMF were set as non-negative random matrices.

Results and discussions

Figure 1a shows the basis vector W obtained from the NMF results for the ARPES dataset of graphene. The vertical and horizontal axes correspond to binding energies and angles, respectively. The angle is output as a detected angle in the measured data and is then converted to a momentum for a general analysis, but in this study it was treated as a dimensionless quantity. The basis vectors correspond to the common patterns in the relevant dataset. Because the dataset consists of Dirac cone structures observed in the ARPES spectra of graphene, which vary with the number of graphene layers at various locations on the substrate, the basis vectors reflect the thickness-dependent spectral features. The ARPES spectrum of ML graphene exhibits a linearly dispersed Dirac cone, whereas that of BL graphene shows two linearly dispersed Dirac cones^{24–26}. For TL graphene, the spectrum show three π bands that differs depending on stacking sequence of layers¹¹.

The components referred to as W0 and W1 exhibit features of the ML graphene. A difference in the photoelectron intensity exists along the energy direction between components W0 and W1. Specifically, component W0 exhibited a strong intensity near the Fermi level, whereas component W1 exhibited a weaker intensity in the same energy region. The components labeled W2, W3, and W4 were similar to the spectrum of TL graphene, which consisted of two linear dispersions and one hole-like dispersion, consistent with ABA stacking¹¹. These spectral features were nearly same, but the center of the spectrum in W3 shifts slightly to a smaller angle side relative to other components. In addition, component W4 exhibited weaker intensity at higher energies than W2 and W3. Although the spectra of BL graphene were also included in the dataset, the BL graphene-derived spectral features were not displayed as basis vectors.

The number of base vectors is a hyperparameter for NMF. For too many base vectors, NMF attempts to reproduce slight differences in the spectrum, such as the difference derived from spectral noise. This results in basis vectors that have nearly identical appearances, differing only in noise intensity. In addition, an increase in the basis vector complicates the interpretation of the results. Therefore, the number of basis vectors must be selected carefully. To determine the number of base vectors, the summation of the squared error (SSE) between the experimental and reconstructed spectra with the product of the basis and activation vectors acquired from NMF was calculated. Figure 1b shows SSE as a function of the number of basis vectors. The SSE decreased as more basis vectors were used. Because the decrease in the SSE slowed down over the five basis vectors applied, the number of basis vectors used was set to five in this study, based on the elbow method.

Because the activation vectors of each spectrum differ depending on the spectral shape, these differences in activation vectors enabled us to distinguish spectra. We performed clustering of ARPES spectra using the k-means method with the activation vectors as the feature values. Optimization of the number of clusters is based on a plot of the within-cluster sum of squared errors as a function of the number of clusters (Fig. 2). If the number of clusters was determined using the elbow method, it was set to $n = 5$. However, a further increase in the number of clusters enabled the classification of slight differences in the spectra. The number of clusters was determined to be $n = 8$ for the following discussion.

Figure 3a–h show representative NMF results. The left panel shows the experimental ARPES spectra of the dataset. The numbers in the left panel represent labels obtained by clustering using the k-means method. The middle panel shows the spectrum reconstructed using the product of the base vector W and activation vector H obtained by NMF. The left panel shows the ratio of activation vector H to each basis vector W used to calculate the spectrum in the middle panel. As shown in Fig. 3, the measured spectra are well reproduced by the products

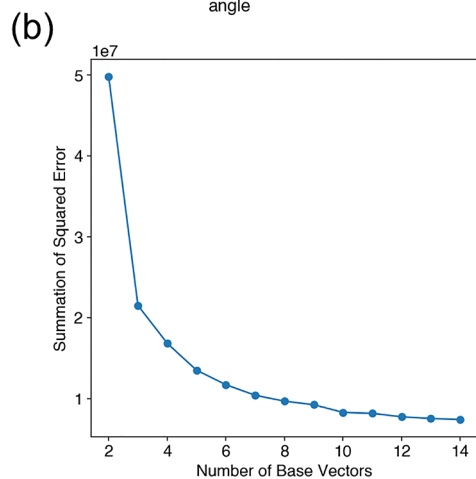
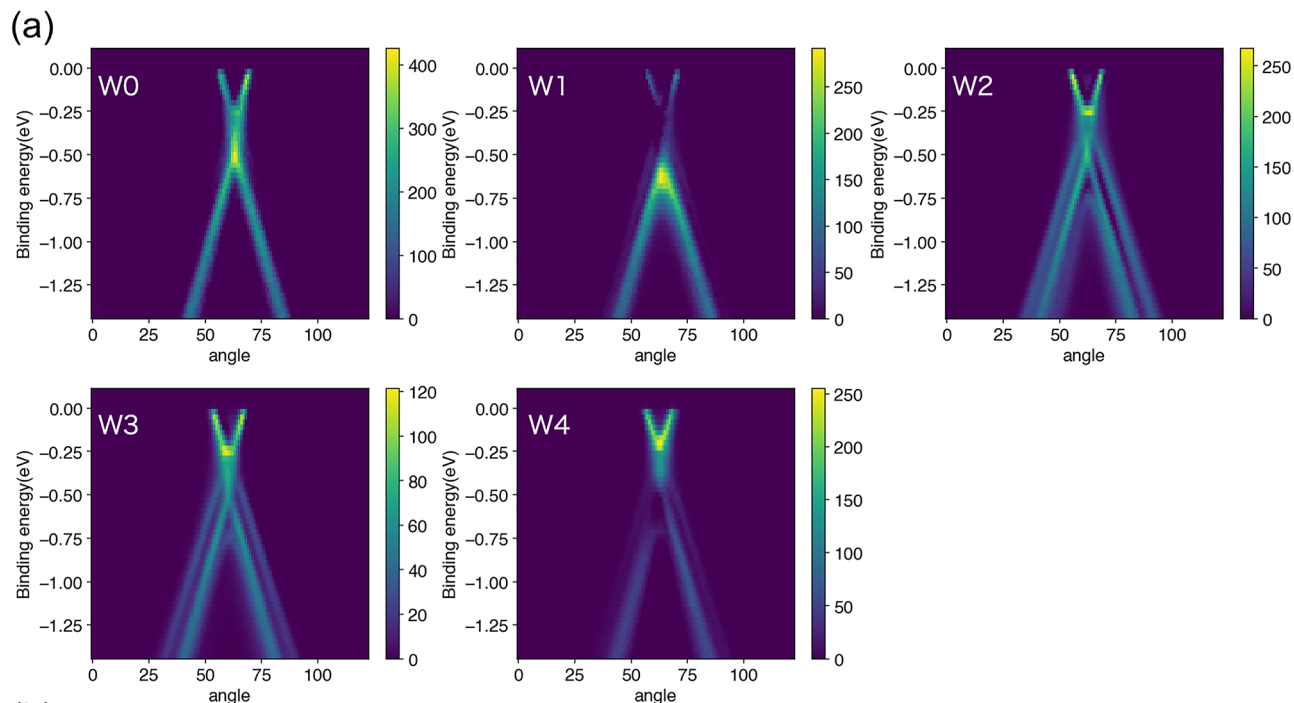


Fig. 1. (a) Basis vectors W obtained using non-negative matrix factorization (NMF) are shown. Angle axis are arbitrary. (b) Summation of mean squared errors between experimental ARPES spectra and reproduced spectra by NMF are plotted as a function of number of basis vector W .

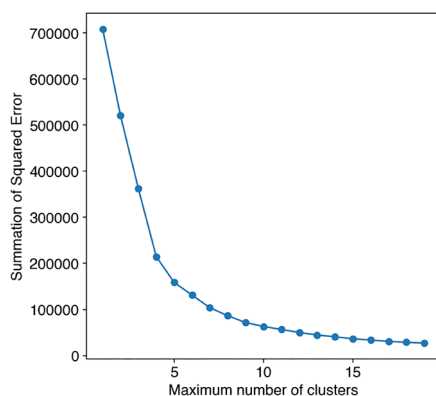


Fig. 2. Within-cluster summation of sum of squared error is plotted as a function of the number of cluster.

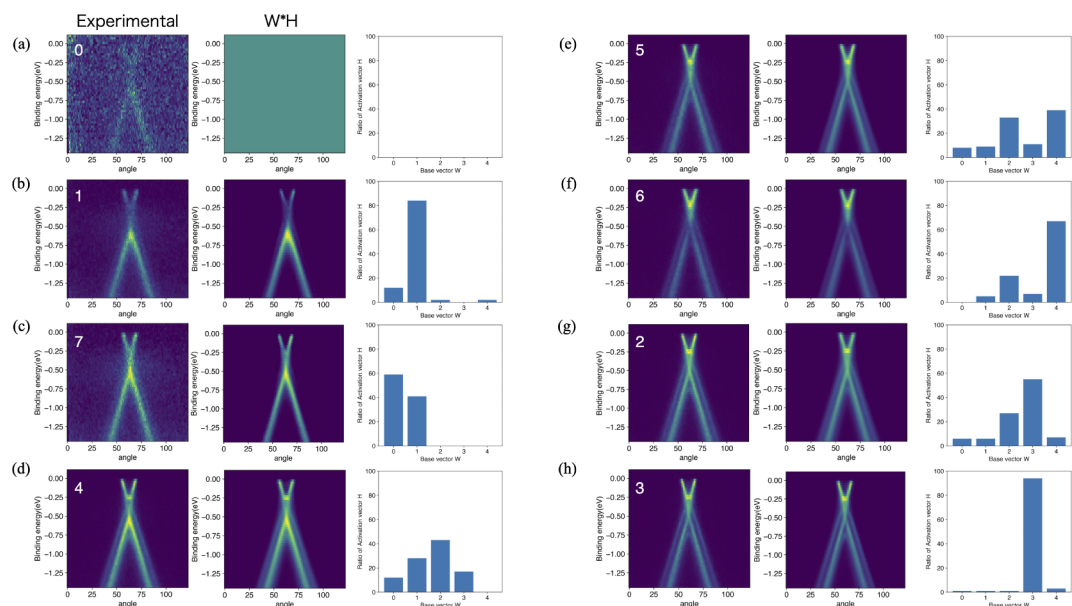


Fig. 3. NMF results for graphene ARPES dataset. Representative of the cluster based on the clustering with activation vector H are shown. The numbers indicated in left panels are the labels acquired from the clustering. Experimentally observed spectrum, reconstructed spectrum by the product of basis and activation vectors, and the ratio of the activation vectors for reconstructions are indicated from left to right panels. Angle axis are arbitrary.

of the basis and activation vectors. Spectral features that varied depending on the measured locations were reproduced well by adjusting the activation vector H, which determines the weights of the basis vector W.

Figure 3a shows the results of the spectrum labeled as zero via clustering. The experimental spectra of label zero show weak or no graphene-derived spectral features, indicating that minimal amount of graphene grew at that measured position. When NMF was performed on the ARPES dataset without the background removal as a preprocessing step, the basis vectors attempted to reproduce the background region. Consequently, multiple basis vectors derived from the background were included in the results and affected the SSE value, which led to a misunderstanding of the results. Therefore, in this study, background removal was performed by setting the intensity below the maximum intensity in the background region to zero. This process does not affect the spectral features derived from graphene in regions of strong spectral intensity originating from graphene. However, because of the background processing, noisy spectra with weak intensities derived from graphene were also removed as background noise. Consequently, the spectra labeled zero were treated as having no spectral intensity, and the resulting reconstructed spectra and activation vector H also exhibited an empty image and an array of zeros.

Figure 3b,c show the spectra of ML-graphene labeled 1 and 7, respectively. Dirac cone features were observed in both spectra, whereas the spectral intensities differed along the energy direction. Whereas the spectrum labeled 1 (Fig. 3b) is reproduced primarily with component W1, the spectrum labeled 7 (Fig. 3c) is reproduced with a combination of the components W0 and W1. Although only one spectrum is shown here, the intensity in the energy direction differed depending on the position of each spectrum of ML graphene, and the spectrum was reproduced by adjusting the values of the basis vectors W0 and W1.

Figure 3d–f are spectra labeled 4, 5, and 6. Features derived from BL graphene were observed in these spectra. Similar to the ML spectra, the intensities along the energy direction in these spectra were different. In Fig. 3d, the spectral intensities are similar at the high- and low-binding energy sides, whereas in Fig. 3e,f, the intensities of the high-binding energy side are weaker than those in the vicinity of the Fermi level. The contribution of component W4 was small for label 4 (Fig. 3d), whereas the contribution of component W4 was large for label 5 (Fig. 3e) and label 6 (Fig. 3f). The difference between Fig. 3e,f was attributed to the contribution of W2 and W4. NMF can fairly well reproduce the experimentally observed spectra that exhibit various spectral features with the modification of the activation vector.

Conversely, the basis vectors W2–W4 used for the primary reconstruction of the experimental spectra of BL graphene exhibited features consistent with the spectrum of TL graphene, but the TL-derived dispersions were unclear in the reconstructed spectra. This was achieved by the band distribution of TL becoming blurred by the addition of other components, such as W1, during the reproduction of the spectra for BL graphene.

Figure 3g,h show the spectra attributed to TL graphene. The reproduced spectra have a larger contribution from component W3 and are labeled 2 and 3 by clustering. The center position of component W3 in Fig. 1 shifts to a lower angle compared to the other components. The experimental spectra in Fig. 3g,h also shifted to lower angle, and the spectrum in Fig. 3h show larger shift than that in Fig. 3g. To reproduce the center shifts of the Dirac cone, the contribution from W3 was adjusted for these spectra.

The intensity modulation along the energy direction the ML and BL graphene was observed in Fig. 3. These intensity modulations were also observed when the measurement angle was rotated along the polar direction, deviating from normal emission. In this study, the measurement angle was fixed during spatially-resolved measurements of the dataset. The modulation of the spectral intensity along the energy direction might be attributed to the terrace structure of the substrate with domains of graphene growing in a non-normal direction^{28,29}, or the SiC substrate being twisted when it was fixed to the holder.

Figure 4 shows a spatial map of the labels acquired using the k-means method for the activation matrix. The vertical and horizontal axes represent the setting values of the x- and z-axes of the ARPES measurement manipulator, respectively. These setting values of locations correspond to the relative locations of graphene grown on the SiC substrate. The spatial distribution of the labels shows that the number of graphene layers increases from right to left, as shown in Fig. 4. The thickness of graphene grown on a substrate depends on the temperature^{22,30,31}. Therefore, the left-hand side of the spatial map with a higher number of layers exhibited a higher substrate temperature during annealing than the right-hand side. The labels were distributed in the spatial map without mixing labels, indicating that NMF and clustering by the activation vectors obtained by NMF worked well for the relevant sample. However, clustering is based on the similarity of activation vectors that reflect the intensity distributions in the spectra. The neighboring spectra were similar, but a slight difference in the activation vectors resulted in the assignment of different clustering labels. Therefore, for convenience, the spectra in this study were clustered into eight clusters; however, in reality, the spectral shape changed gradually.

The shifts of center angle of the spectra are observed in Fig. 3g,h. The spectra labeled 2 and 3 are located at the periphery of the substrate, as shown in Fig. 4. At ARPES measurements, the graphene grown on SiC substrate was fixed onto the sample holder with a Mo foil. Since the SiC substrate and Mo foil have different work functions, a slight electric field is generated at the edge of the sample. The generated electric field bends the emitted photoelectrons, affecting the detected emission angle. This is consistent with the fact that the amount of shift was larger for label 3 compared to label 2.

In general, since NMF is an NP-hard problem, its solution is not unique and is sensitive to initial values of matrices W and H . To investigate the dependence on initial values, we performed NMF while varying the number of basis vectors W and changing the initial matrices. In this work, even though more than 20 different initializations were tested, the resulting basis vectors W showed negligible difference. In addition, differences among the outputs from tested initializations appeared only in slight variations in the intensity of the basis vectors W and values of the activation vectors H . Therefore, the clustering results based on the activation vectors H were not affected by different initializations. In the case of two-dimensional materials, where the spectral shape changes significantly with the number of layers, similar results are expected even if NMF is sensitive to the initial matrices. Nevertheless, it is important to keep in mind and verify whether the NP-hard problem causes issues. NMF uses a two-dimensional matrix constructed by stacking the experimentally obtained spectra which were transformed into one-dimensional arrays as input. The input spectra contain information about the initial and final states as well as matrix elements. As a result, the generated basis matrix W displays the asymmetric band-dispersion derived from the intensity of experimental spectra that are affected by the initial, final states, and the matrix element. Even when using this method, it is necessary to consider the matrix elements effects, just as in standard ARPES measurements. Optimization of the experimental setup parameters, such as the incident angle of photon, emission angle of electron, photon energy, and photon polarization so that target bands can be observed clearly on the measurement, will be helpful. In addition, the basis matrix W , that is obtained as common patterns involved in many spectra, are treated as common images without considering their physical meanings of photoelectron intensities in NMF. Therefore, it is essential to carefully compare and evaluate the reconstructed spectra by NMF with the experimentally obtained ones individually, while also examining the effect of initialization and appropriately adjusting hyperparameters such as the number of components. Moreover, validating the significance of the results based on physical insights is crucial. In this context, comparing the calculated spectral images obtained as basis vectors with bands that include the matrix elements effects from theoretical calculations can be particularly effective.

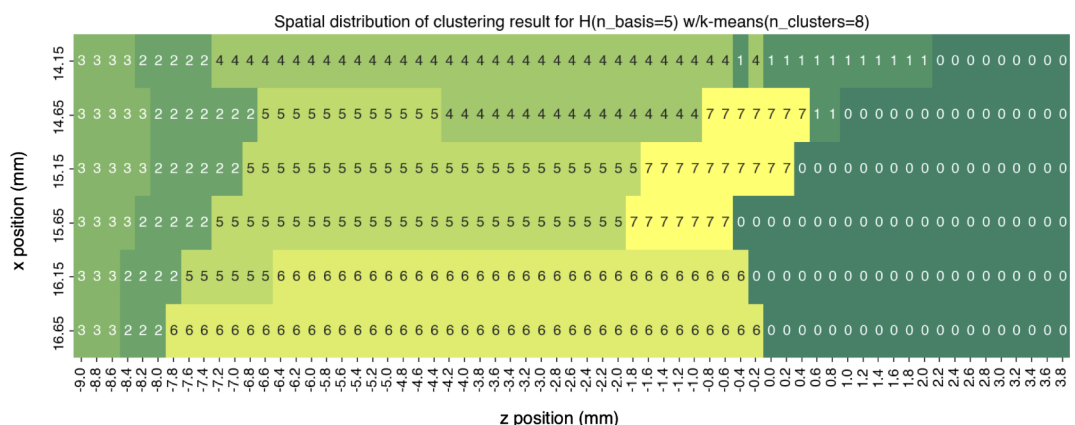


Fig. 4. Spatial distribution of the labels acquired by the clustering based on the activation vectors H .

At NMF, the basis matrix W and the activation matrix H are optimized to minimize the reconstruction error across the entire dataset using the given shape of W and H . When the number of basis vectors is small and there are few spectra containing subtle structures, including subtle structures in the basis matrix results in a larger overall error. To ensure that subtle structures are included and extracted as the resulting basis matrix, it is necessary that a sufficient proportion of the spectra contain these subtle structures. Alternatively, using a sufficiently large number of basis vectors might allow for the incorporation of subtle spectral features, thereby reducing the overall dataset error and potentially including components associated with subtle structures in the results. The spatially-resolved ARPES are often conducted to observe subtle spectral features that appear only in specific regions using sufficiently small spot. The present NMF and k-means analysis would be applicable to resolve subtle spectral features that commonly appear in specific regions even when insufficiently small spot is used.

Recent development of the instruments enables ‘data-heavy’ ARPES measurement, such as k_y - or k_z -resolved measurements using deflection-type electron lens or varying the excitation photon energy. Moreover, dependence on the polarization of excitation light, spin-resolved measurement, and time-resolved measurement with pump-probe method generates data-heavy multi-dimensional spectra. Although the computational load will increase, we believe that situations with increased data volume are precisely when automated methods like proposed in this article are most needed.

The application of k-means clustering to the ARPES spectra has been reported previously^{13,14}, and compared with previous reports, the NMF-based method used in this study can extract specific band dispersions as basis vectors and screen regions of physical interest. This feature enables the detection of spectra with characteristic band structures and even eliminates spectral components arising from contamination or defects. In ARPES, the spectra sometimes contain mixed-phase information, depending on the measured location, relative sizes of the incident beam, and domain size. The ARPES spectra from various locations reflected the differences in the ratios of the mixed phases. By distinguishing these mixed features using NMF, it is possible to extract the spectral features originating from a single phase without arbitrariness as a hyper-spatial resolution that exceeds the spot size. By performing the same analysis on datasets with worse signal-to-noise ratios than the dataset used in this study, we expect to provide comprehensive information on the sample in a shorter scan time. In the future, we intend to apply this procedure to datasets with poor signal-to-noise ratios to evaluate its robustness.

Conclusions

In this study, an unsupervised approach was proposed to efficiently analyze large amounts of data obtained using spatially-resolved photoelectron spectroscopy. NMF was applied to a dataset of ARPES spectra of graphene on a SiC substrate. The spatial distributions of graphene with different numbers of layers on the substrate were visualized using k-means clustering on the activation matrix. NMF can extract frequent patterns from a large number of ARPES spectra using various features as the basis vectors. The activation vectors were adjusted to reproduce slight differences between the spectra. The spatial distribution of the graphene grown on the substrate was also revealed by applying k-means clustering to the obtained activation vectors. The proposed method can be applied to any ARPES spectrum and the spectral features can be extracted without examining the individual spectra.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 25 June 2024; Accepted: 20 September 2024

Published online: 15 October 2024

References

- Avila, J., Lorcy, S. & Dudin, P. ANTARES: Space-resolved electronic structure. *J. Electron. Spectrosc. Relat. Phenom.* **266**, 147362 (2023).
- Bostwick, A., Rotenberg, E., Avila, J. & Asensio, M. C. Zooming in on electronic structure: NanoARPES at SOLEIL and ALS. *Synchrotron Radiat. News* (2012).
- Dudin, P. et al. Angle-resolved photoemission spectroscopy and imaging with a submicrometre probe at the SPECTROMICROSCOPY-3.2L beamline of Elettra. *J. Synchrotron Radiat.* **17**, 445–450 (2010).
- Johansson, L. I. et al. Multiple π -bands and Bernal stacking of multilayer graphene on C-face SiC, revealed by nano-angle resolved photoemission. *Sci. Rep.* **4**, 4157 (2014).
- Wilson, N. R. et al. Determination of band offsets, hybridization, and exciton binding in 2D semiconductor heterostructures. *Sci. Adv.* **3**, e1601832 (2017).
- Cattelan, M. & Fox, N. A perspective on the application of spatially resolved ARPES for 2D materials. *Nanomaterials* **8**, 284 (2018).
- Peng, H. et al. Substrate doping effect and unusually large angle Van Hove singularity evolution in twisted bi- and multilayer graphene. *Adv. Mater.* **29**, 1606741 (2017).
- Mo, S. K. Angle-resolved photoemission spectroscopy for the study of two-dimensional materials. *Nano Converg.* **4**, 6 (2017).
- Lv, B., Qian, T. & Ding, H. Angle-resolved photoemission spectroscopy and its application to topological materials. *Nat. Rev. Phys.* **1**, 609–626 (2019).
- Bao, C. et al. Spatially-resolved electronic structure of stripe domains in IrTe₂ through electronic structure microscopy. *Commun. Phys.* **4**, 229 (2021).
- Bao, C. et al. Stacking-dependent electronic structure of trilayer graphene resolved by nanospot angle-resolved photoemission spectroscopy. *Nano Lett.* **17**, 1564–1568 (2017).
- Ekahana, S. A. et al. Transfer learning application of self-supervised learning in ARPES. *Mach. Learn. Sci. Technol.* **4**, 035021 (2023).
- Melton, C. N. et al. K-means-driven gaussian process data collection for angle-resolved photoemission spectroscopy. *Mach. Learn. Sci. Technol.* **1**, 045015 (2020).

14. Iwasawa, H., Ueno, T., Masui, T. & Tajima, S. Unsupervised clustering for identifying spatial inhomogeneity on local electronic structures. *Npj Quantum Mater.* **7**, 24 (2022).
15. Peng, H. et al. Super resolution convolutional neural network for feature extraction in spectroscopic data. *Rev. Sci. Instrum.* **91**, 033905 (2020).
16. Liu, J., Huang, D., Yang, Y. & Qian, T. Removing grid structure in angle-resolved photoemission spectra via deep learning method. *Phys. Rev. B* **107**, 165106 (2023).
17. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).
18. Long, C. J., Bunker, D., Li, X., Karen, V. L. & Takeuchi, I. Rapid identification of structural phases in combinatorial thin-film libraries using X-ray diffraction and non-negative matrix factorization. *Rev. Sci. Instrum.* **80**, 103902 (2009).
19. Stanev, V. et al. Unsupervised phase mapping of X-ray diffraction data by nonnegative matrix factorization integrated with custom clustering. *Npj Comput. Mater.* **4**, 43 (2018).
20. Tanimoto, H. et al. Non-negative matrix factorization for 2D-XAS images of lithium ion batteries. *J. Phys. Commun.* **5**, 115005 (2021).
21. Shiga, M. & Muto, S. Non-negative matrix factorization and its extensions for spectral image data analysis. *E-J. Surf. Sci. Nanotechnol.* **17**, 148–154 (2019).
22. Yu, X. Z. et al. New synthesis method for the growth of epitaxial graphene. *J. Electron. Spectrosc. Relat. Phenom.* **184**, 100–106 (2011).
23. Takahashi, K. et al. Upgrade of Saga-university beamline in SAGA-LS. *J. Phys. Conf. Ser.* **425**, 072007 (2013).
24. Ohta, T., Bostwick, A., Seyller, T., Horn, K. & Rotenberg, E. Controlling the electronic structure of bilayer graphene. *Science* **313**, 951–954 (2006).
25. Jin, S. et al. Epitaxial growth of uniform single-layer and bilayer graphene with assistance of nitrogen plasma. *Nanomaterials* **11**, 3217 (2021).
26. Riedl, C., Coletti, C. & Starke, U. Structural and electronic properties of epitaxial graphene on SiC(0 0 1): a review of growth, characterization, transfer doping and hydrogen intercalation. *J. Phys. Appl. Phys.* **43**, 374009 (2010).
27. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
28. Norimatsu, W. & Kusunoki, M. Formation process of graphene on SiC (0001). *Phys. E Low-Dimens. Syst. Nanostruct.* **42**, 691–694 (2010).
29. Sakakibara, R. et al. Step unbunching phenomenon on 4H-SiC (0001) surface during hydrogen etching. *Appl. Phys. Lett.* **123**, 031603 (2023).
30. Hass, J., De Heer, W. A. & Conrad, E. H. The growth and morphology of epitaxial multilayer graphene. *J. Phys. Condens. Matter* **20**, 323202 (2008).
31. Zebardastan, N. et al. High quality epitaxial graphene on 4H-SiC by face-to-face growth in ultra-high vacuum. *Nanotechnology* **34**, 105601 (2023).

Acknowledgements

This work was supported by JSPS KAKENHI (Grant Numbers: 20K03821, 24K06927) and the Partnership Project for Fundamental Technology Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author contributions

Masaki Imamura: Conceived and designed the experiments; analyzed the data; wrote the manuscript. Kazutoshi Takahashi: Conducted the experiments; Contributed to data interpretation; reviewed and edited the manuscript. All authors have read and approved the final manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024