




Removing grid structure in angle-resolved photoemission spectra via deep learning method

Junde Liu ^{1,2}, Dongchen Huang ^{1,2}, Yi-feng Yang ^{1,2,3,*} and Tian Qian ^{1,2,3,†}

¹Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³Songshan Lake Materials Laboratory, Dongguan, Guangdong 523808, China



(Received 1 November 2022; revised 22 March 2023; accepted 22 March 2023; published 5 April 2023)

Spectroscopic data may often contain unwanted extrinsic signals. For example, in the angle-resolved photoemission spectroscopy (ARPES) experiment, a wire mesh is typically placed in front of the charge-coupled device to block stray photoelectrons but could cause a gridlike structure in the spectra during quick measurement mode. In the past, this structure was often removed using the mathematical Fourier filtering method by erasing the periodic structure. However, this method may lead to information loss and vacancies in the spectra because the grid structure is not strictly linearly superimposed. Here, we propose a deep learning method to overcome this problem effectively. Our method takes advantage of the self-correlation information within the spectra themselves and can greatly optimize the quality of the spectra while removing the grid structure and noise simultaneously. It has the potential to be extended to all spectroscopic measurements to eliminate other extrinsic signals and enhance the spectral quality based on the self-correlation of the spectra solely.

DOI: [10.1103/PhysRevB.107.165106](https://doi.org/10.1103/PhysRevB.107.165106)

I. INTRODUCTION

In past decades, angle-resolved photoemission spectroscopy (ARPES) has driven the research of novel quantum materials with its incredible ability to directly probe the electronic structures [1–7]. Owing to the rapid development of experimental techniques, highly rapid data acquisition modes (“fixed” or “dithered” modes) are frequently required in many application scenarios. For example, the fast scanning mode is often used in spatial-resolved ARPES experiments to map the energy band spectra of a wide area of the sample [8,9]. It is also preferred when the spot size of the laser beam is tiny, in which case the measurement should be limited to a very short time to avoid sample damage [10]. In addition, the fast scanning mode can save a lot of acquisition time for measuring higher-dimensional data, such as band structures in two-dimensional momentum space or dynamical electronic structures in time-resolved ARPES [11–15].

However, because of the metal mesh in front of the analyzer charge-coupled device (CCD), the spectra obtained using the fast scanning mode typically have a gridlike structure as shown in Fig. 1(a) [7,16], which hinders the direct observation of energy band features. Although the grid structure may be averaged out using the swept mode, measurements in this mode often cover a large portion of unwanted energy ranges, causing a significant time waste in opposition to the original purpose. Post-spectral processing techniques and methods are hence needed to remove this grid structure.

A traditional method is to use Fourier filtering by converting the real-space spectra to the Fourier domain [16]. But

simply erasing the peaks in Fourier space may lose intrinsic information because the grid structure is not strictly periodic and linearly superimposed. For instance, the peaks and energy bands in Fourier space may merge together when their widths are comparable, which makes it difficult to remove the peaks without losing information about the energy bands. It may become even more challenging when the data quality is noisy or not good enough so that the peaks in Fourier space cannot be well identified. Therefore, the traditional Fourier filtering method requires the peaks in Fourier space to be sharp and distinguishable from the intrinsic bands, which essentially limits the range of its application scenarios.

Fortunately, machine learning methods have shown strong capabilities in spectra processing [17–20]. Deep learning based methods can achieve better performance than traditional mathematical Gaussian smoothing methods in removing noise from spectra [20]. More surprisingly, we have shown that a noisy spectra image can be decomposed into an image of clean spectra and an image corresponding to noise, provided that the noise is sparse and not so coherent with the intrinsic energy bands. This inspired us to view the grid structure as an extrinsic signal, so that the spectra may be decomposed into a clean part and a grid part.

Following the above line of thought, we propose a deep learning based method to remove the grid structure and identify the intrinsic signal in the measured ARPES spectra in this work. Our method utilizes the self-correlation information of the spectra themselves and parametrizes the observed ARPES data by two convolutional neural networks (CNNs). One network is designed to extract the clean spectra, and the other aims to extract the grid texture. We show that this enables us to preserve the signal of the energy bands in the spectra even if the grid width and energy bandwidth are comparable or the spectral quality is not so good. As a result, our method can

*Corresponding author: yifeng@iphy.ac.cn

†Corresponding author: tqian@iphy.ac.cn

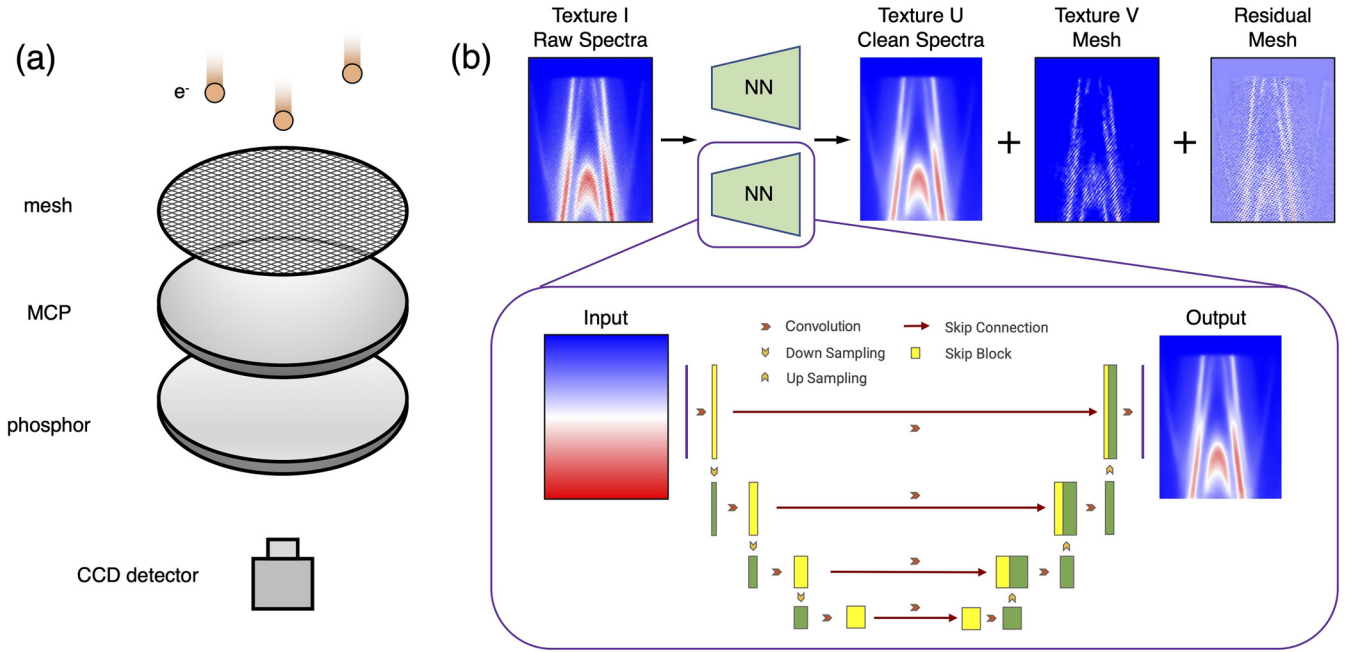


FIG. 1. (a) An illustration of the ARPES analyzer. (b) The raw spectra image is parametrized as a mixture of two textures represented by two individual neural networks of the same structure. The first texture aims to recover the spectra and the other finds the grid structure somewhat attached to the energy bands. The residual is the general grid structure since the neural network does not learn all the features in the input image.

nicely remove the grid structure and extend the application of the fast scanning mode, which is essential for ARPES experiments. Our idea can be applied to any extrinsic structures from other multidimensional data, thus improving the quality of the spectra obtained in more general scenarios.

II. METHODS

We start by assuming that a complex image can be represented by a mixture of two or more simple components or textures, each of which can be viewed as an image. At first glance, this seems impossible and underdetermined because the unknown parameters are more than the known ones. But this problem has been proved solvable in the high-dimensional analysis if one component is incoherent with another [21,22]. In other words, pixels or patches are correlated within each texture but independent of those within another texture. This insight has led to successful algorithms showing good performance in computer vision [23,24] and ARPES denoising [25] even without requiring a training set.

Motivated by these successes, we extend the above idea and develop an algorithm for degriding the ARPES data. As shown in Fig. 1(b), we attempt to decompose a raw spectra image I into two textures U and V and expect that U gives the clean spectra and V approximates the grid structure. All the images I, U, V may be viewed as matrices, and the entries in U or V should be correlated. We parametrize each texture via a U-shape neural network [26], which has shown good performance in modeling correlated data and real-world tasks in computer vision [26]. U and V are then the outputs of two individual neural networks of the same structure elaborated in Appendix B.

The summation of both textures U, V should give the raw spectra I , which we take as the loss function. This yields

$$\min_{\theta, \phi} \|I - U_{\theta} - V_{\phi}\|_F^2, \quad (1)$$

where θ, ϕ are the parameter collections for each neural network, and $\|\cdot\|_F^2$ denotes the square of the Frobenius norm, which is defined as the summation of all entries of a matrix. The parameters are then optimized by stochastic gradient descent with respect to the loss function given in Eq. (1). Our algorithm is implemented via Pytorch [27] and the detailed training hyperparameters are listed in Appendix B.

Noisy data. In practice, noise is unavoidable in the raw data. To make the algorithm more practical, we combine denoising techniques into this framework. The noise s is assumed to be sparse because an image of measured spectra should not be corrupted everywhere. Following similar techniques [28], we decompose the raw spectra into three textures: the clean spectra U , the grid structure V , and the noise s . As discussed in [25], the noise can be further parametrized as $s = a \circ a - b \circ b$, where a, b are two vectors with learnable entries to be optimized, and the symbol \circ denotes their elementwise product. The loss function now becomes

$$\min_{\theta, \phi, a, b} \|I - U_{\theta} - V_{\phi} - a \circ a + b \circ b\|_2^2, \quad (2)$$

where every matrix is flattened as a vector and $\|\cdot\|_2^2$ denotes the square of ℓ^2 norm. We find quite stable training with the form in Eq. (2).

III. APPLICATIONS

Figure 2(a) shows the raw data and the degrided results using different methods for the ARPES data of MnBi_2Te_4

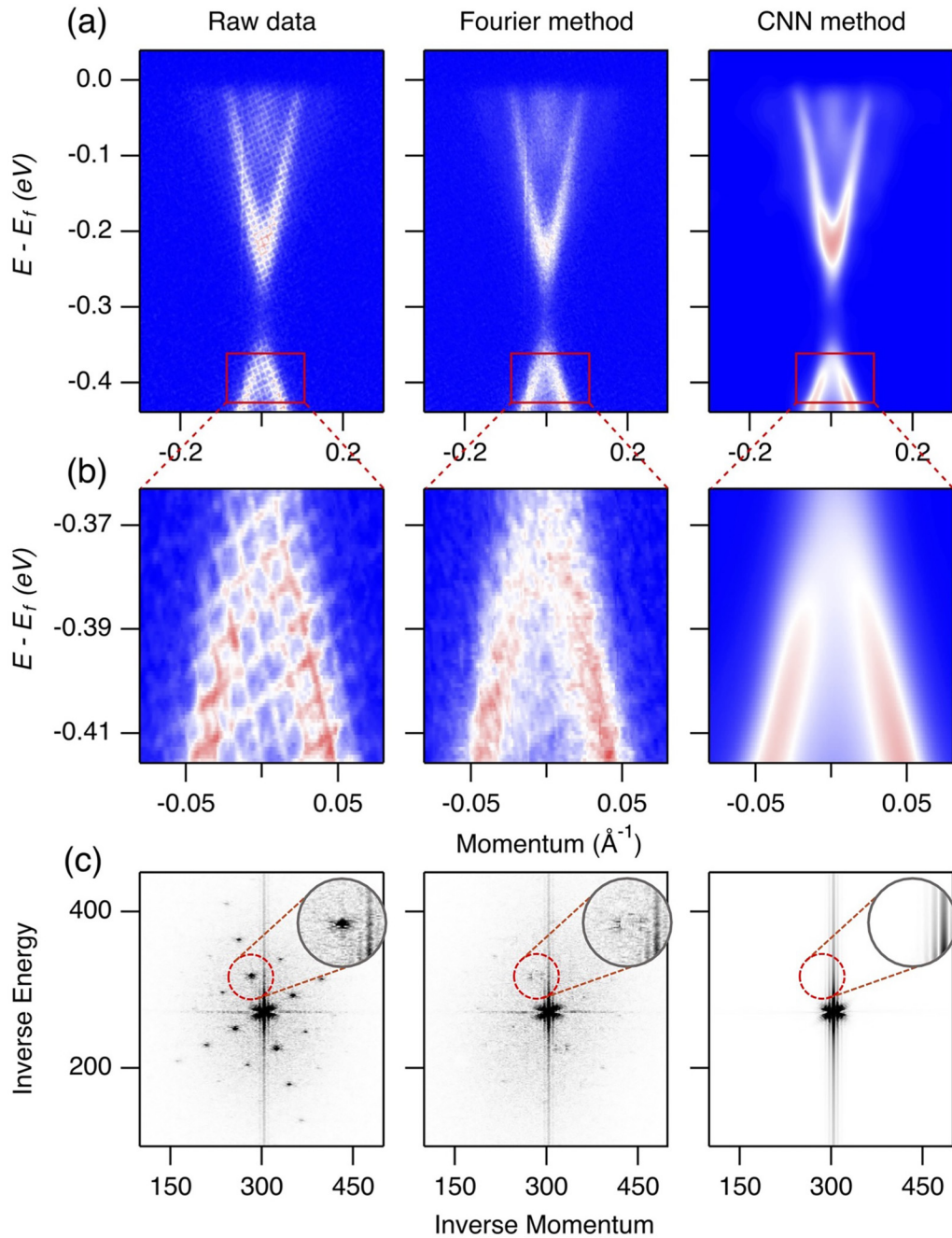


FIG. 2. Comparison of the degripped results using different methods. (a) MnBi_2Te_4 ARPES data along the cut through $\bar{\Gamma}$ and the degripped results based on the Fourier filtering method and our deep-learning method. (b) Close-up showing the area in the red box region indicated in panel (a). (c) Corresponding Fourier transform results of panel (a). The inset shows the close-up of the first-order peak and energy band signals indicated with a red circle.

along the cut through $\bar{\Gamma}$. The raw data were measured in the fast scanning mode. The quasiperiodic grid structure is clearly visible in the raw data, as may be best viewed in the enlarged plot Fig. 2(b). The grid structures overlap with the energy band features, which may hinder the observation and analysis of some fine details.

Figure 2(c) shows the spectra in the Fourier space after applying the Fourier transform to the data in Fig. 2(a). We see that the energy band signals are concentrated in the central region and along the lines due to its nonperiodic nature, while

the signals of the quasiperiodic grid structure yield the Fourier peaks distributed symmetrically near the diagonals of the image, whose locations represent the reciprocal vectors of the grid. The inset (red circle) shows the close-up of the first-order Fourier peak and energy band signals. They are quite close to each other, making it extremely difficult to erase the grid structure without losing intrinsic information.

The degripped results by using the Fourier filter based method and the deep learning based method are compared in Fig. 2(b). While both methods can remove the grid structure

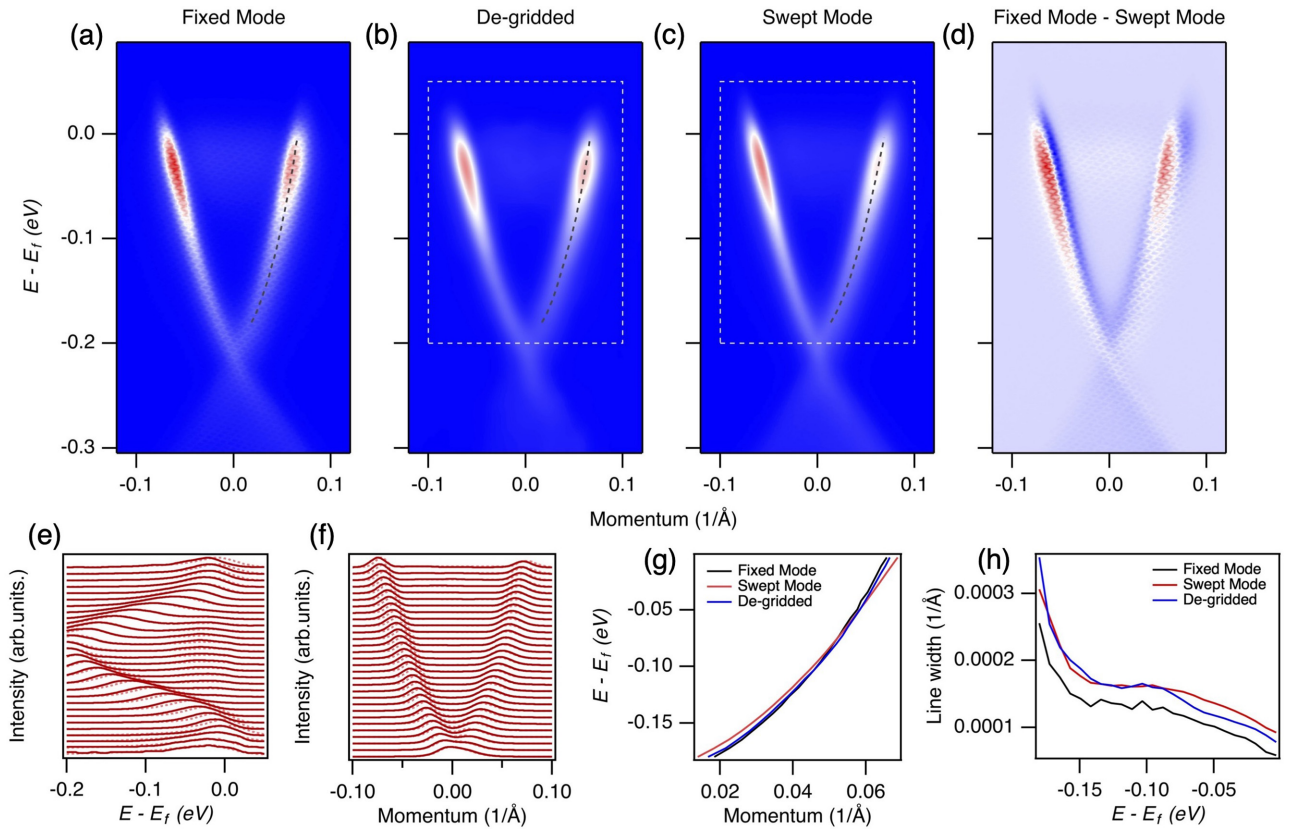


FIG. 3. Quantitative analysis of the degraded results. (a) Bi_2Se_3 ARPES data along the cut through Γ measured with fixed mode. (b) The corresponding degraded results of panel (a) based on our deep-learning method. (c) Same as panel (a) but measured with swept mode. (d) Differential spectra of panels (a) and (b). (e) Energy distribution curves (EDCs) corresponding to the white box region indicated in panels (b) and (c). The dark-red solid lines represent the degraded results and the light-red dashed lines represent the swept mode results. (f) Corresponding momentum distribution curves (MDCs). (g) Peak position obtained by fitting the MDCs curve. The band dispersions of fixed mode spectra, degraded spectra, and swept mode spectra are indicated by gray dashed lines in panels (a)–(c), respectively. (h) Peak width obtained by fitting the MDCs curves of fixed mode spectra, degraded spectra, and swept mode spectra.

from the original spectra, the quality of the resulting spectra processed by the deep learning based method is obviously much better. We see that the deep learning based spectra are very clean and smooth, but the Fourier filter based ones are still rough and even contain some empty spaces. This is because the existence of the grid structure interferes with the acquisition of the intrinsic energy band signals in the measurement process. As a result, simply erasing the Fourier peaks fails to fill the missing intrinsic signal affected by the grid structure and thus loses the corresponding energy band features. Indeed, as shown in Fig. 2(c), it is difficult for the Fourier filter based method to completely remove the Fourier peaks without affecting the intrinsic energy band signal. By contrast, deep learning based methods can naturally and directly extract the complete intrinsic energy band signal through local correlation without losing critical information. It not only eliminates the Fourier peaks entirely but also completes the intrinsic energy band signal, making the spectra quite smooth and comprehensive.

To evaluate the fidelity and reliability of our deep-learning method, we conducted a quantitative analysis on the ARPES data of the topological insulator Bi_2Se_3 along the cut through Γ . As we can see, the data measured under fixed mode in Fig. 3(a) show clear grid structures, while those under swept

mode in Fig. 3(c) only show clean energy band information. A straightforward comparison shows that our degraded results in Fig. 3(b) have a good performance and show very comparable energy band structures, indicating that the convolutional neural network method removes the grid structure while retaining the intrinsic energy band information to a great extent. It should be noted that our method assumes a linear superposition of the energy band and grid structure in the spectra. From the differential spectra of fixed mode and swept mode shown in Fig. 3(d), the extracted part contains both features of the grid structure and of the energy band, indicating that the spectra are not exactly their linear superposition. Nevertheless, the naive linear superposition assumption may not be a bottleneck in practice. Its good performance indicates that the correlation between grid structure and energy band is weak so the correlation information can be assembled into the extracted grid texture under the linear superposition assumption.

To further show the good performance of our deep-learning method and the linear superposition assumption, we have also conducted a quantitative line shape analysis. Figures 3(e) and 3(f) compare the energy distribution curves (EDCs) and momentum distribution curves (MDCs) of the degraded data and the swept mode data in the white box region indicated

in Figs. 3(b) and 3(c). It can be seen that the curves of the degraded data (dark-red solid lines) and the swept mode data (light-red dashed lines) are more closely matched, indicating that the deep-learning method can yield comparable results with the swept mode, which preserves a clean energy band signal to a large extent without losing intrinsic information.

In addition, since important physics often occurs in the energy bands near the Fermi level, we have made further analysis of band features near the Fermi level to show the performance of our algorithm. Specifically, we extracted the peak position and width of the energy bands near the Fermi level region by fitting the MDCs. As compared in Figs. 3(g) and 3(h), the fitting results from the fixed mode, degraded, and swept mode data are almost the same, demonstrating that our deep-learning method preserves the intrinsic information of the energy bands near the Fermi level region. Note that the peak widths of the fixed mode data are smaller, probably because the grid structure enhances the local intensity and reduces the peak widths. More examples are given in Appendix A to illustrate the universality and robustness of our degraded method.

However, there are still some scenarios in which our method may not be as applicable. Based on our assumption of linear superposition, the energy band signal is required to occupy a significant proportion of the spectra so that its correlation information is sufficient for the algorithm to extract. For example, when the acquisition time is too short, the signal-to-noise ratio of the spectra may be too low for the algorithm to distinguish genuine signals from spurious signals such as grid and noise. Moreover, for some extremely fine structures, the presence of a grid may obscure a significant amount of intrinsic information such that the algorithm cannot accurately reconstruct the signals. Additionally, since our method relies on the assumption of sparse noise, our algorithm is not strictly applicable when the noise is particularly dense or when there are bad channels in the spectra. Therefore, while our approach has demonstrated good performance in many scenarios, it is important to be aware of these limitations and to evaluate its applicability carefully in each specific case.

IV. DISCUSSION

A. The training dynamics

As described above, the deep learning based method can remove the weak-correlated extrinsic grid structure and extract the strong-correlated intrinsic energy band signals. The reason for its good performance may be ascribed to the correlation of neighboring pixel values in the spectra of ARPES experiments. Specifically, because of the finite linewidth of the energy band, the regions of band features generally span a finite range and contain a number of MCP events. Since a single event on the MCP is a spot that occupies multiple pixels in the CCD detector, the energy band usually covers many CCD pixels so that the intensity of neighboring CCD pixels does not exhibit large fluctuations. By contrast, the size of the grid structure in most cases is much smaller than the linewidth of the energy band and contains fewer MCP events so that the changes of CCD pixel values brought by the grid structure are more localized and their global correlation is much weaker.

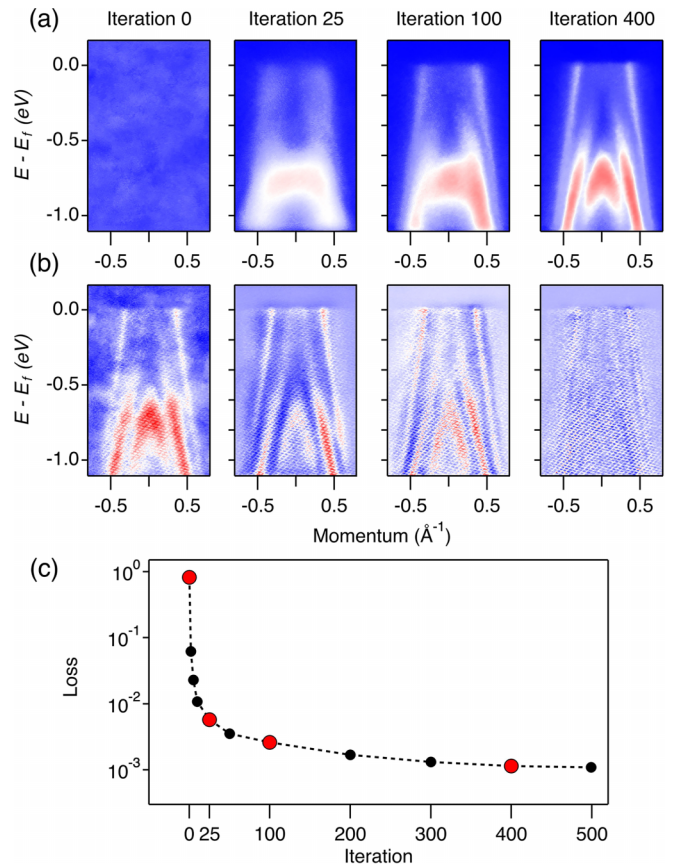


FIG. 4. Intermediate results during the training process dependent on iteration. (a) $\text{Ir}_{1-x}\text{Pt}_x\text{Te}_2$ ARPES data and iteration-dependent degraded results for the iteration steps 0, 25, 100, and 400, which denotes the number of loops in the computing process. (b) Corresponding residual spectra of (a). (c) Loss $L(x, f(x'))$ as a function of the iteration. The corresponding iteration numbers plotted in (a) and (b) are highlighted with enlarged red circles.

Therefore, even if the grid has contaminated the information at several CCD pixels, the value of those CCD pixels can be retrieved by inferring the most statistically probable value from the CCD pixel values in a larger area nearby while retaining the intrinsic information from the energy band.

To demonstrate how the algorithm extracts the intrinsic energy band signal, we present the dynamical processing results of $\text{Ir}_{1-x}\text{Pt}_x\text{Te}_2$ ARPES data in Fig. 4. It can be seen that as the iteration of the computing process increases, the intrinsic energy band signals are gradually extracted [Fig. 4(a)], leaving only the grid structure [Fig. 4(b)]. Specifically, at zero iteration, the extracted signal is the random initial input. As the number of iterations increases to 25, the contours of the two energy bands start to emerge. Although the band structures are very crude at this stage, they have already covered all the areas where the real energy bands are located. As the iteration increases to 100 and 400, more detailed structures are revealed and we eventually obtain the clear spectra. The computing process tends to extract the signals [Fig. 4(a)] with more local and stronger correlations from the remaining spectra [Fig. 4(b)] to produce the intrinsic band structures.

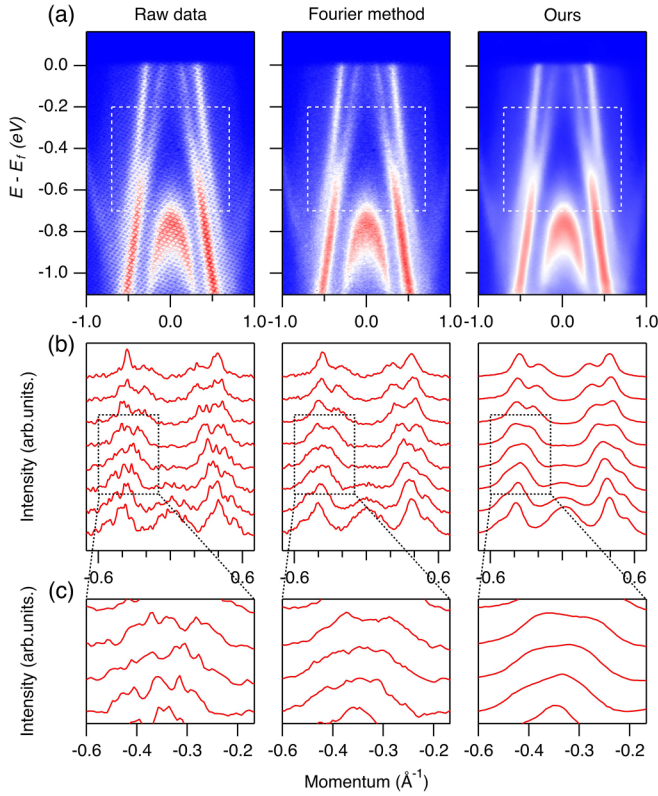


FIG. 5. Comparison analysis of different methods. (a) The raw data of ARPES spectra for $\text{Ir}_{1-x}\text{Pt}_x\text{Te}_2$ and the degraded results based on the Fourier filtering method and the deep-learning method. (b) Corresponding momentum distribution curves of the area in the white box region indicated in panel (a). (c) Close-up showing of the area in the black box region indicated in panel (b).

Figure 4(c) plots the evolution of the loss function $L(x, f(x'))$ with the training iteration. We see that it decreases rapidly from the initial point to iteration 25. As shown in Fig. 4(a), the algorithm first captures the general shape of the energy bands, indicating that the weak global correlations can be extracted quite easily. As the correlation information in the energy bands to be extracted becomes stronger and more local, the loss decreases more slowly. After approximately 300 iterations, the loss converges to a small value. The convergence shows that the algorithm can no longer extract more local correlation information and the model is not overfitted to the grid structure, thus ensuring that the energy band information of the intrinsic signals may be largely preserved. From this point of view, we may conclude that the neural network first learns low-frequency long-range features and then adds high-frequency local details during the training.

B. Effect of denoising

As the grid structure is removed, the spurious noise is also seen to be removed from the spectra and the quality of the resulting spectra is greatly improved. To quantitatively evaluate the performance of denoising, we use line shape analysis on $\text{Ir}_{1-x}\text{Pt}_x\text{Te}_2$ ARPES data. As shown in Fig. 5(a), our method can preserve the intrinsic band structures while removing the noise compared to the original data and the

degraded results using the Fourier filtering method. This is more clearly seen in the smoother MDCs shown in Figs. 5(b) and 5(c) for the area indicated by the white box region in Fig. 5(a).

The reason why the noise can be simultaneously removed should be attributed to the foundation of our method, which respects the locally correlated intrinsic energy bands in noisy and gridlike spectra. It has been proved [21] that if a signal is a superposition of some incoherent components, each component can be identified by proper parametrization and optimization. In the analysis of spectroscopic measurements, both the spurious noise signals and the grid structures are weakly correlated with the intrinsic energy bands. The weak correlation of the noise may be understood since human eyes are able to identify the intrinsic energy band from a noisy measurement, thus indicating that the noise does not overlap the full information of the intrinsic energy band. The superposition of the grid and noise, combined with the intrinsic energy bands, gives the raw spectra. We may hence view them as three nearly independent parts. Thus, by parametrizing them via three corresponding neural networks in Eq. (2), we can separate them from the total spectra and identify the intrinsic energy bands as shown in Fig. 5(a). This setting is easily generalizable, so our method can be extended to many application scenarios to remove different extrinsic signals simultaneously and achieve clear and desired spectra. Although, our algorithm does not assume Poisson statistics about the noise. We have tested the robustness of our algorithm in removing Poisson distribution noise and found that its variance is consistently minimal across different spectra. This indicates that our algorithm is highly reliable and does not introduce additional errors in the quantitative analysis, thus providing a robust method for improving the quality of the spectra. In addition, we mention that it might be difficult to extract error bars when analyzing the denoised data because we do not explicitly deal with the Poisson statistics.

V. CONCLUSION

In this work, we propose a deep learning based method to eliminate the gridlike structure in fast scanning mode for ARPES spectroscopy. This is quite helpful for saving measurement time and enhancing spectral quality. Compared to the traditional Fourier filtering methods, our deep-learning method can extract intrinsic information directly by making use of the local correlation nature of the energy band signals. It can remove extrinsic features including the grid structure and noise simultaneously, and thus greatly improve the spectra. Our algorithm may also be extended to other spectroscopic measurements for removing unwanted signals, including quasiperiodic structures, spurious signals, noise, etc.

ACKNOWLEDGMENTS

We thank M.-J. Pan, F.-M. Chen, J.-R. Huang, and B. Jiang for useful discussions, and Z.-C. Rao, S.-Y. Gao, and W.-H. Fan for data support. This work was supported by the National Natural Science Foundation of China (NSFC Grants No. U1832202, No. 11888101, No. 11974397, and No. 12174429), the Chinese Academy of Sciences

(Grants No. QYZDB-SSW-SLH043, No. XDB33000000, and No. XDB28000000), the Informatization Plan of Chinese Academy of Sciences (Grant No. CAS-WX2021SF-0102), and the Synergetic Extreme Condition User Facility (SE-CUF).

J.L. and D.H. contributed equally to this work.

APPENDIX A: MORE EXAMPLES

Figures 6(a) and 6(b) show more degrided examples based on the deep-learning method. Each degrided

spectrum is normalized according to the total intensity of its raw spectrum, while the range of the colorbar is set to the same. Regardless of sharp or diffuse energy bands in the spectra, our method can always remove the grid structure and extract well the intrinsic energy band information. This indicates that our method based on CNNs is quite stable and has a strong potential for generalization.

Figure 6(c) shows the distribution of the spectral intensity obtained from Figs. 6(a) and 6(b), which can represent the dynamic range of the spectra. The distribution based on the degrided spectra covers most of that of the raw spectra,

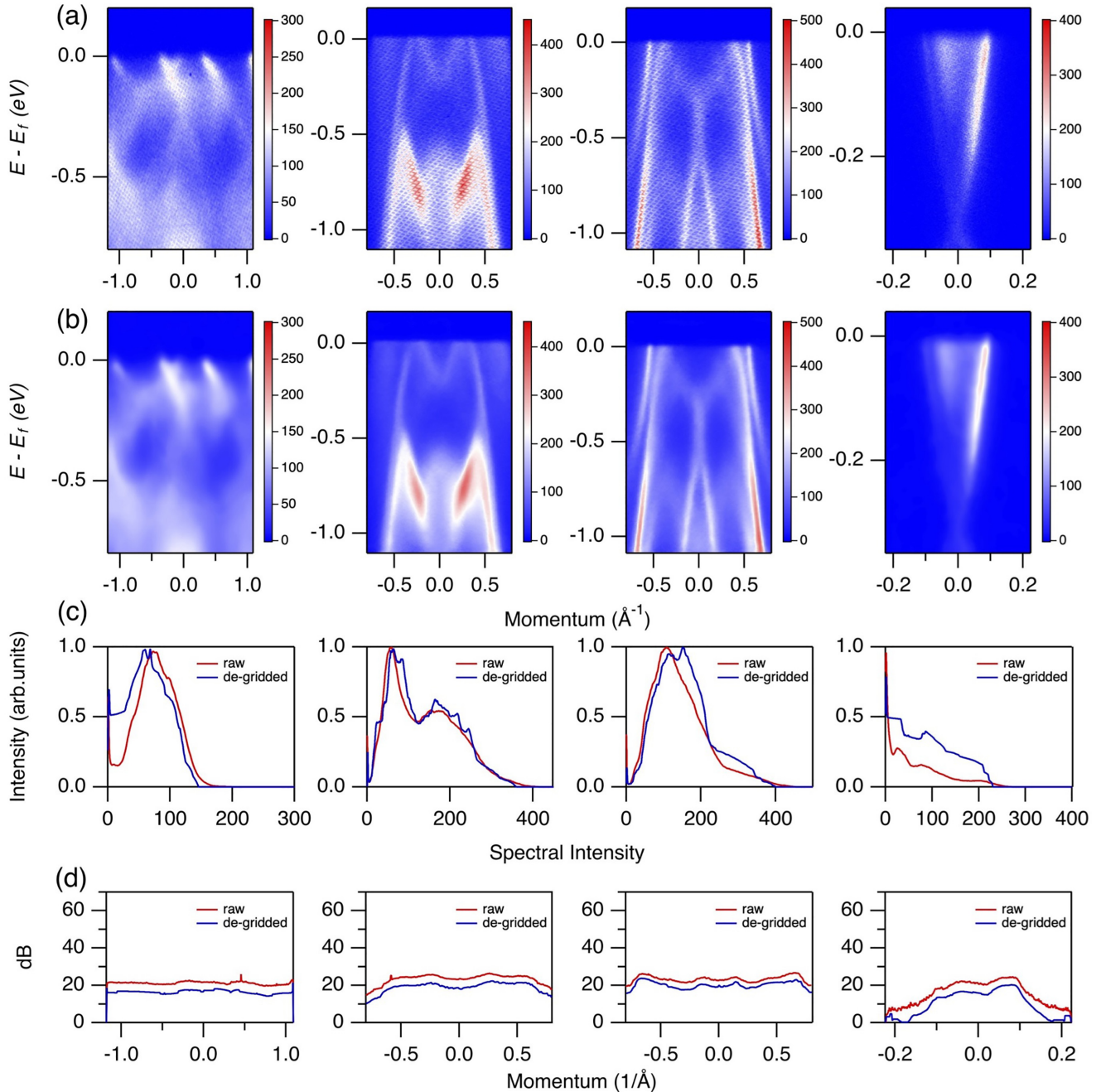


FIG. 6. The degrided results for more ARPES spectra. (a) The raw data with gridlike structure. (b) The corresponding degrided results based on our deep-learning method. (c) The distribution on the spectral intensity of the raw data and degrided data. (d) The dynamic range of EDCs for different momentum positions based on the raw data and degrided data.

TABLE I. Neural network architecture.

Encoder network	
Input spectra x	
Conv2d, BN, $5 \times 5 \times 16$, stride = 2, lReLU	
Conv2d, BN, $5 \times 5 \times 32$, stride = 2, lReLU	
Conv2d, BN, $5 \times 5 \times 64$, stride = 2, lReLU	
Conv2d, BN, $5 \times 5 \times 128$, stride = 2, lReLU	
Conv2d, BN, $5 \times 5 \times 128$, stride = 2, lReLU	
Decoder network	
Conv2d, BN, $5 \times 5 \times 128$, stride = 1, lReLU	
Upsampling, Conv2d, BN, $5 \times 5 \times 128$, stride = 1, lReLU	
Upsampling, Conv2d, BN, $5 \times 5 \times 64$, stride = 1, lReLU	
Upsampling, Conv2d, BN, $5 \times 5 \times 32$, stride = 1, lReLU	
Upsampling, Conv2d, BN, $5 \times 5 \times 16$, stride = 1, lReLU	
Conv2d, $1 \times 1 \times 1$, stride = 1, sigmoid	

indicating that our degriding method sets no artificial limit on the extracted spectra. However, it is worth noting that the maximum spectral intensity in each distribution of the degrided data is smaller than that in the distribution of the raw data. This is because the intensity is enhanced by the grid structure in the raw data, so the maximum spectral intensity becomes lower when the grid is removed.

To verify the effect of the degrided method more clearly, we calculate the dynamic range of each EDC curve at different momentum positions of the raw data and the degrided data according to the formula in dB:

$$\text{dB} = 10 \times \log_{10} \left(\frac{I_{\max}}{1 + I_{\min}} \right), \quad (\text{A1})$$

where I_{\max} and I_{\min} denote the maximum and minimum intensity of each EDC curve, respectively. The results are compared in Fig. 6(d). The dynamic range of the degrided data is not much different from that of the raw data, indicating that our degriding method does not have a significant effect on

the resolution of the spectra and the degrided results are acceptable. The dynamic range after degriding is smaller at different momentum positions, probably because of the removal of the grid structure signal, which reduces the maximum intensity and hence the dynamic range.

APPENDIX B: IMPLEMENTATION DETAILS

The neural network parameters are listed in Table I. We use batch normalization (BN) [29] every layer, and set the parameter of leaky ReLU to 0.2. This algorithm is not sensitive to the input, so we use similar inputs as reported in previous work [23,28], plus a small fraction (~ 0.01) of the raw data in this work. To achieve better results, discrepant learning rates should be used for the three neural networks. We find (0.2, 2, 600) is an appropriate parameter choice for many ARPES images.

APPENDIX C: PULSE-COUNTING MODE SPECTRA

To verify the generalizability of our algorithm, we applied the degriding method to the ARPES spectra of Bi_2Se_3 under pulse-counting mode. In this mode, an MCP event will not occupy several CCD pixels, but only a single CCD pixel, i.e., the CCD pixels are independent of each other. Figure 7 shows the degrided results of pulse-counting mode spectra, where each spectrum is normalized according to the total intensity.

Even though the signal intensities are independent between CCD pixels, the degrided results in Fig. 7 show that our algorithm can still remove the grid signal nicely and yield comparable results as the swept mode. This indicates that although the CCD pixels are acquired independently, the energy band has a certain linewidth that can occupy a finite range of CCD pixels, so the algorithm can use this correlation to extract clean energy band information. This suggests that our algorithm only depends on the correlation between energy band signals and hence may be universally applied in different modes of ARPES applications.

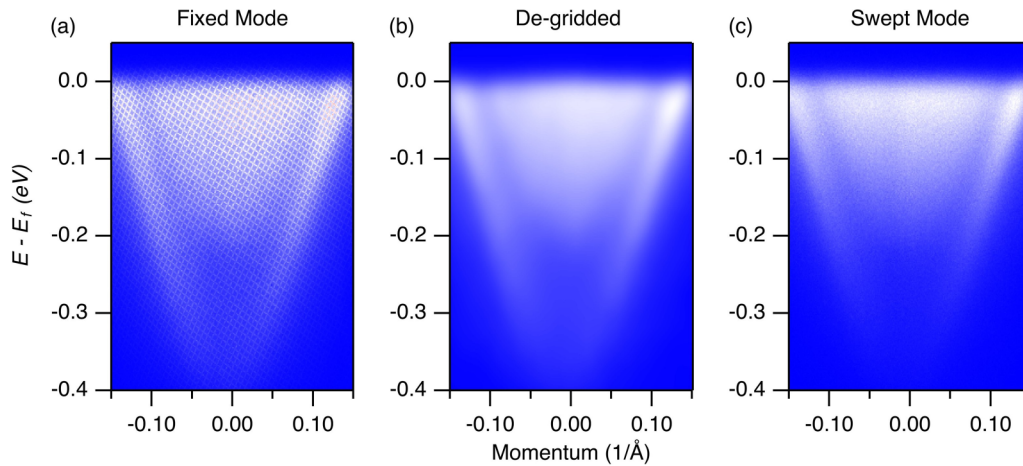


FIG. 7. Degriding results of pulse-counting mode spectra. (a) The pulse-counting mode ARPES spectra of Bi_2Se_3 under fixed mode. (b) Corresponding degrided results of (a) based on our deep-learning method. (c) The pulse-counting mode ARPES spectra of Bi_2Se_3 under swept mode.

- [1] A. Damascelli, Z. Hussain, and Z.-X. Shen, Angle-resolved photoemission studies of the cuprate superconductors, *Rev. Mod. Phys.* **75**, 473 (2003).
- [2] M. Hashimoto, I. M. Vishik, R.-H. He, T. P. Devereaux, and Z.-X. Shen, Energy gaps in high-transition-temperature cuprate superconductors, *Nat. Phys.* **10**, 483 (2014).
- [3] H. Yang, A. Liang, C. Chen, C. Zhang, N. B. M. Schroeter, and Y. Chen, Visualizing electronic structures of quantum materials by angle-resolved photoemission spectroscopy, *Nat. Rev. Mater.* **3**, 341 (2018).
- [4] C. Zhang, Y. Li, D. Pei, Z. Liu, and Y. Chen, Angle-resolved photoemission spectroscopy study of topological quantum materials, *Annu. Rev. Mater. Res.* **50**, 131 (2020).
- [5] B. Lv, T. Qian, and H. Ding, Angle-resolved photoemission spectroscopy and its application to topological materials, *Nat. Rev. Phys.* **1**, 609 (2019).
- [6] B. Q. Lv, T. Qian, and H. Ding, Experimental perspective on three-dimensional topological semimetals, *Rev. Mod. Phys.* **93**, 025002 (2021).
- [7] J. A. Sobota, Y. He, and Z.-X. Shen, Angle-resolved photoemission studies of quantum materials, *Rev. Mod. Phys.* **93**, 025006 (2021).
- [8] H. Iwasawa, E. F. Schwier, M. Arita, A. Ino, H. Namatame, M. Taniguchi, Y. Aiura, and K. Shimada, Development of laser-based scanning μ -ARPES system with ultimate energy and momentum resolutions, *Ultramicroscopy* **182**, 85 (2017).
- [9] M. Cattelan and N. A. Fox, A perspective on the application of spatially resolved ARPES for 2D materials, *Nanomaterials* **8**, 284 (2018).
- [10] M. E. Grass, P. G. Karlsson, F. Aksoy, M. Lundqvist, B. Wannberg, B. S. Mun, Z. Hussain, and Z. Liu, New ambient pressure photoemission endstation at advanced light source beamline 9.3.2, *Rev. Sci. Instrum.* **81**, 053106 (2010).
- [11] J. Orenstein, Ultrafast spectroscopy of quantum materials, *Phys. Today* **65**(9), 44 (2012).
- [12] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Nonequilibrium dynamical mean-field theory and its applications, *Rev. Mod. Phys.* **86**, 779 (2014).
- [13] C. L. Smallwood, R. A. Kaindl, and A. Lanzara, Ultrafast angle-resolved photoemission spectroscopy of quantum materials, *Europhys. Lett.* **115**, 27001 (2016).
- [14] C. Bao, P. Tang, D. Sun, and S. Zhou, Light-induced emergent phenomena in 2D materials and topological materials, *Nat. Rev. Phys.* **4**, 33 (2022).
- [15] A. de la Torre, D. M. Kennes, M. Claassen, S. Gerber, J. W. McIver, and M. A. Sentef, *Colloquium*: Nonthermal pathways to ultrafast control in quantum materials, *Rev. Mod. Phys.* **93**, 041002 (2021).
- [16] S. Liu, E. Kotta, Y. Xu, J. Mutch, J.-H. Chu, M. Hoesch, S. K. Mahatha, J. D. Denlinger, and L. A. Wray, Fourier-based methods for removing mesh anomalies from angle resolved photoemission spectra, *J. Electron Spectrosc. Relat. Phenom.* **260**, 147255 (2022).
- [17] P. Zhang, P. Richard, T. Qian, Y.-M. Xu, X. Dai, and H. Ding, A precise method for visualizing dispersive features in image plots, *Rev. Sci. Instrum.* **82**, 043712 (2011).
- [18] Y. He, Y. Wang, and Z.-X. Shen, Visualizing dispersive features in 2D image via minimum gradient method, *Rev. Sci. Instrum.* **88**, 073903 (2017).
- [19] H. Peng *et al.*, Super resolution convolutional neural network for feature extraction in spectroscopic data, *Rev. Sci. Instrum.* **91**, 033905 (2020).
- [20] Y. Kim *et al.*, Deep learning-based statistical noise reduction for multidimensional spectral data, *Rev. Sci. Instrum.* **92**, 073901 (2021).
- [21] E. J. Candès, X. Li, Y. Ma, and J. Wright, Robust principal component analysis? *J. Assoc. Comput. Mach.* **58**, 1 (2011).
- [22] J. Wright and Y. Ma, Dense error correction via ℓ^1 -minimization, *IEEE Trans. Inf. Theory* **56**, 3540 (2010).
- [23] D. Ulyanov, A. Vedaldi, and V. Lempitsky, Deep image prior, *Int. J. Comput. Vis.* **128**, 1867 (2020).
- [24] Y. Gandelsman, A. Shocher, and M. Irani, “Double-DIP”: unsupervised image decomposition via coupled deep-image-priors, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (IEEE, Los Alamitos, 2019).
- [25] D. Huang, J. Liu, T. Qian, and Y.-F. Yang, Spectroscopic data de-noising via training-set-free deep learning method, [arXiv:2210.10494](https://arxiv.org/abs/2210.10494).
- [26] O. Ronneberger, P. Fischer, and T. Brox, U-net: convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, edited by N. Navab, J. Hornegger, W. M. Wells, and Alejandro F. Frangi (Springer International Publishing, Cham, 2015).
- [27] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshain, L. Antiga *et al.*, PyTorch: an imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox and R. Garnett (Curran Associates, Inc., San Francisco, 2019).
- [28] C. You, Z. Zhu, Q. Qu, and Y. Ma, Robust recovery via implicit bias of discrepant learning rates for double over-parameterization, in *Advances in Neural Information Processing Systems*, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Curran Associates, Inc., San Francisco, 2020).
- [29] S. Ioffe and C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on Machine Learning*, edited by Francis Bach and David Blei (PMLR, Lille, 2015).