

# An Adaptive Symmetric Zero-Area Peak Finding Method Suitable for Micro-Area X-ray Spectroscopy

Yuhan Fang<sup>1</sup>, Bohua Yin<sup>2, \*</sup>, and Zhengjie Li<sup>2</sup>

<sup>1</sup> University of Jinan, Jinan, Shandong, China

<sup>2</sup> Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing, China

\*Corresponding author: yinbh@mail.iee.ac.cn

---

## Abstract

In micro-area spectroscopic analysis, the precise identification of characteristic X-ray peak positions is of paramount importance. Analyzing the energy information from these characteristic X-ray peaks enables the determination of the types of elements present in the sample and their relative abundances. However, spectral graphs often present intricate challenges, such as overlapping and weak spurious peaks, which significantly complicate the peak-finding process. To tackle these hurdles in micro-area X-ray spectroscopy, this paper proposes an adaptive threshold symmetric zero-area peak-finding algorithm. Compared to traditional methods such as the simple comparison method, derivative method, and Gaussian fitting method, the symmetric zero-area peak-finding algorithm demonstrates significant advantages in detecting overlapping, weak, and spurious peaks. Through peak-finding experiments conducted on original spectral data from various samples, and by comparing the results with standard characteristic X-ray element tables and other established algorithms, it is evident that the symmetric zero-area algorithm not only accurately distinguishes each individual peak but also effectively differentiates real peaks from noise-induced false peaks, thus proving the feasibility of the symmetric zero-area peak-finding algorithm in micro-area spectroscopy.

## Keywords

Peak-finding Algorithm; Gaussian Fitting; Symmetric Zero-area; Micro-area Spectroscopy; Adaptive.

---

## 1. Introduction

Spectroscopy technology serves as a vital analytical tool in electron microscopy, complements complementing electron morphological imaging to concurrently characterize the microstructure and chemical composition of materials. This synergy significantly broadens the application spectrum and research depth of electron microscopy. Through spectroscopic analysis, not only can elemental qualitative and quantitative analyses be conducted at the nanoscale, but information about the spatial distribution, chemical states, and electronic structure of elements can also be obtained. As spectroscopy technology continues to advance, enhancements in spectrometer performance, refinement of quantitative analysis methods, and the integration of big data and machine learning, the synergistic utilization of spectroscopy and electron microscopy will deepen further. This synergy promises to offer more robust support for scientific research across diverse disciplines. The burgeoning significance of spectroscopy within electron microscopy underscores its status as an indispensable tool for material characterization and analysis. Micro-area spectroscopy multi-channel analysis stands as the cornerstone of is the core of spectroscopy technology, providing an

indispensable approach for material characterization and analysis. This method, non-destructive in nature, discerns the types and concentrations of elements by detecting the characteristic X-ray energy emitted when an electron beam interacts with a sample's atoms. The processing of spectroscopic data is crucial, encompassing tasks like data smoothing, automatic peak finding, peak energy determination, and elemental qualitative and quantitative analysis, with precision of peak calculations emerges as a pivotal factor in determining spectroscopic quality. Presently, prevalent peak-finding methodologies include the simple comparison method, Gaussian function fitting method, and derivative method[1].

This paper introduces an adaptive symmetric zero-area peak-finding algorithm, addressing the limitations of conventional peak-finding methods through experiments and simulation analyses. The algorithm employs a convolution sliding transformation technique with a window function featuring a "zero-area" characteristic and experimental spectra. Since the selected transformation function is the standard deviation of the Gaussian function, and it adjusts its number of points according to the FWHM of the spectrum, thereby achieving adaptability. This algorithm outperforms traditional peak-finding methods, particularly in discerning weak and overlapping peaks.

## 2. Spectral Data Processing

Micro-area X-ray spectroscopy technology serves extremely important in material analysis, but its original spectral data usually exhibit high background and noise. These undesirable elements can significantly interfere with subsequent peak identification processes, thereby compromising the accuracy of analysis outcomes. Moreover, precise element identification and quantitative analysis hinge upon meticulous energy calibration, which establishes the correspondence between channel numbers and characteristic X-ray energy. Therefore, prior to commencing peak identification and analysis, a series of preprocessing steps, including energy calibration and fast Fourier filtering, must be executed to effectively eliminate noise and interference from the original data. This foundational preprocessing lays the groundwork for accurate peak identification in subsequent stages. The algorithm flowchart is illustrated below.

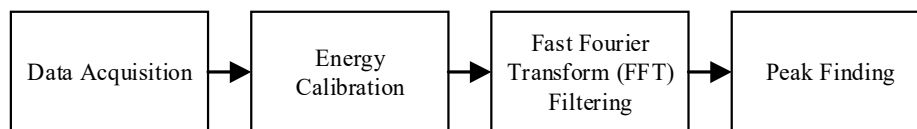
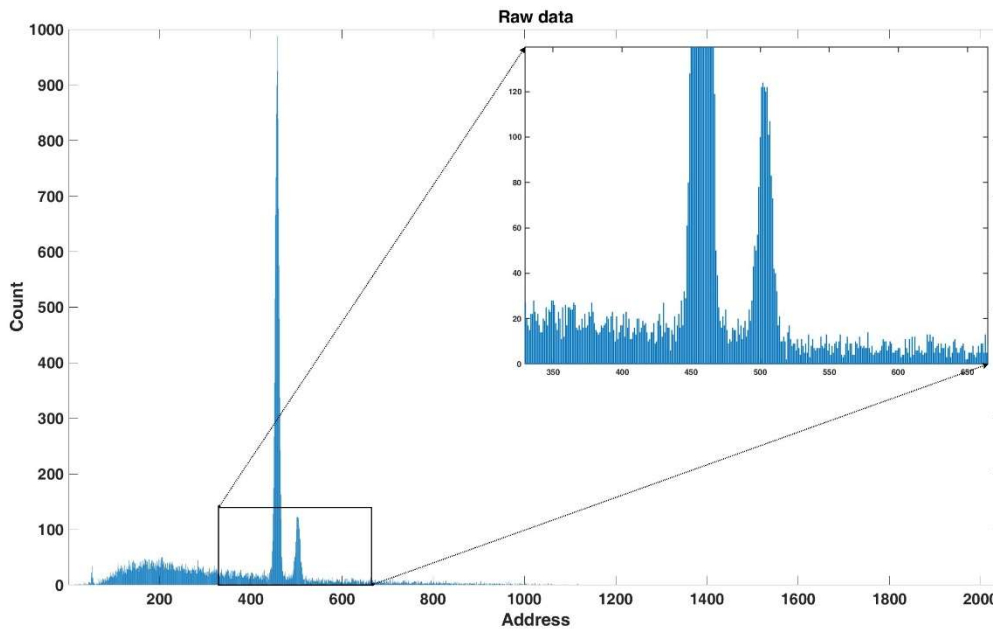


Fig. 1 Flowchart of the algorithm

### 2.1 Section Data Smoothing

Micro-area spectroscopy involves irradiating specific material samples with X-rays or other particles. Upon interaction with the sample, these particles trigger the emission or absorption of energy-specific X-rays by the atoms within. Detectors capture and record these X-rays, forming the energy spectra observed. Each characteristic peak in the spectrum indicates the presence of a specific element within the sample; the peak's position corresponds to the element's energy, while its height reflects the element's abundance. The Fig. 2 shows an energy spectrum obtained from bombarding a pure zinc (Mn) element sample with a scanning electron microscope at 25KV high voltage, with a collection time of 200 seconds. Given that the spectral collection process relies on the statistics of individual X-ray events, inherent randomness leads to fluctuations in peak heights within the graph. Even under ideal conditions, repeated measurements will exhibit variance. This statistical fluctuation phenomenon interferes with precise spectrum analysis, particularly when signals are weak.

To enhance the accuracy and reliability of analysis, spectral data typically undergo a smoothing process. This crucial step aids in diminishing random noise, thereby enhancing the signal-to-noise ratio. Consequently, peaks become more distinct, facilitating easier identification and quantification. This, in turn, enables more precise qualitative and quantitative analysis of sample elements. Common denoising methods for spectral data include mean filtering, median filtering, Savitzky-Golay filtering, Fourier transform filtering, and wavelet transform filtering.



**Fig. 2** Elemental energy spectroscopy data of Mn element and local enlarged image

This paper chooses to use Fast Fourier Transform (FFT) filtering, which offers considerable advantages in micro-area spectroscopy peak analysis. Notably, FFT filtering boasts high computational efficiency, enabling swift processing of sizable datasets, and is well-suited for real-time analysis applications. Leveraging FFT's capabilities in the frequency domain allows for precise noise control while effectively retaining valuable signals. Moreover, FFT exhibits superior suppression of periodic noise and adeptly preserves signal details while eliminating noise, rendering it adept at handling datasets characterized by complex background noise. Furthermore, FFT can be seamlessly integrated with other algorithms to augment peak-finding capabilities, rendering it particularly well-suited for micro-area spectroscopy analysis[2][1].

The operational procedure of FFT filtering is as follows: Initially, the spectral data are subjected to a fast Fourier transform, leading to frequency data:

$$X_k = DFT[x(n)] = \sum_{n=0}^{N-1} x(n)W_N^{kn}, \quad k = 0, 1, \dots, N-1 \quad (1)$$

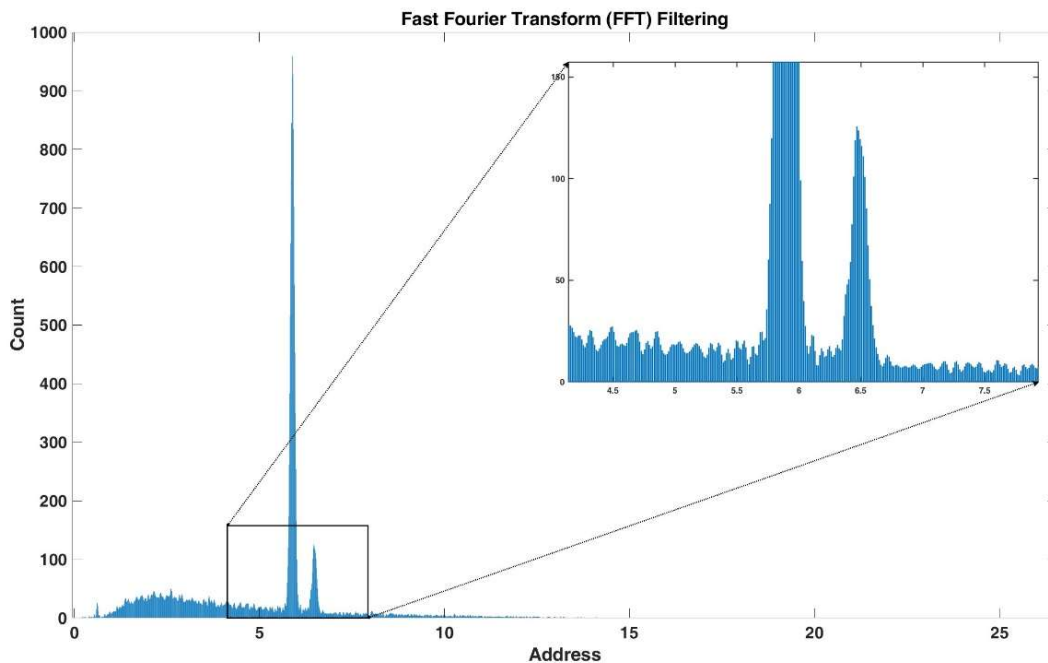
Then, a bandwidth threshold, denoted as  $T$ , is established, and frequencies exceeding this threshold in the spectrum are assigned a value of zero:

$$X_k = \begin{cases} X_k & X(k) > T \\ 0 & X(k) < T \end{cases} \quad (2)$$

Lastly, the revised spectral data undergo a fast Fourier inverse transform, yielding the updated spectrum data.

$$x(n) = IDFX[X(k)] = \sum_{k=0}^{N-1} X(k)W_N^{-kn}, \quad n = 0, 1, \dots, N-1 \quad (3)$$

Following FFT filtering, as depicted in the Fig. 3, the energy spectral data typically exhibit smoother characteristics, exhibit smoother characteristics. This refinement facilitates clearer differentiation and recognition of peak values, particularly those previously obscured by noise. The FFT-filtered spectrum offers substantial advantages for qualitative and quantitative elemental analysis, given the enhanced signal-to-noise ratio. Moreover, features such as the shape, width, and height of the peaks may more accurately reflect the underlying physical processes involved.



**Fig. 3** Energy spectrum after FFT filtering and its local enlargement

## 2.2 Peak Finding Algorithm Comparison

In micro-area spectroscopy analysis, the significance of peak finding lies chiefly in the capability to precisely identify and quantitatively analyze the elemental constituents within a sample. Through accurate localization of peak positions in the spectrum, it becomes feasible to discern the elements present and their respective concentrations. This holds profound implications for fields such as materials science, chemical analysis, and physical research. Reliable peak-finding techniques enhance the reliability and accuracy of analyses, furnishing indispensable tools probing the composition, structure, and interactions of materials at the microscopic scale. Common peak-finding algorithms include the simple comparison method, derivative method, Gaussian fitting, and symmetric zero-area method. Presented below is an exhaustive comparison of these algorithms, as shown in Table 1.

**Table 1.** The advantages and disadvantages of each filtering algorithm and its applicability

Peak-finding Method	Advantages	Disadvantages	Applicability
Simple Comparison	Computationally simple, suitable for preliminary analyses.	Limited resolution; poor robustness against random noise and systemic noise.	Initial peak identification in data with minimal noise.
Derivative Method	Improves resolution, effectively detects closely adjacent peaks.	Noise significantly impacts results;	Analysis requires edge detection where peaks are close together.
Gaussian Fitting	High precision in peak value and shape analysis, suitable for complex overlapping peaks.	Model adaptability and accuracy highly depend on reasonable parameter initialization.	Precise quantitative analysis, especially in complex and overlapping peak scenarios.
Symmetric Zero-Area	Adaptive threshold, efficiently handles overlapping peaks, enhances analysis sensitivity and specificity.	Implementation is relatively complex, requiring expertise for optimization.	Especially suitable for complex spectra and weak signal analysis, such as micro-area spectroscopy.

After conducting a thorough comparison of aforementioned methods, this paper selects the symmetric zero-area method as the peak-finding technique. As a relatively novel technology, the symmetric zero-area peak-finding method primarily utilizes the principle of symmetric zero-area to effectively mitigate background noise interference and refine the accuracy of peak detection. Compared to traditional methods, it streamlines the data processing workflow and optimizes the identification of overlapping peaks, demonstrating high adaptability across various spectroscopic datasets. Additionally, its computational efficiency renders it suitable for rapid analysis of expansive datasets. These attributes collectively position it as an efficient and precise solution for handling intricate data characterized by significant noise and overlapping peaks. The subsequent section provides a detailed exposition of the symmetric zero-area peak-finding method:

The characteristic of a linear function is a second derivative of zero; however, when the second derivative near a point is negative, a non-linear function can achieve a local maximum at that point, indicating that the second derivative near a characteristic peak is less than zero. Therefore, using the second derivative method can effectively suppress base levels exhibiting nearly linear changes, thereby enhancing peak identification capabilities. Empirical evidence shows the second derivative method's prowess in identifying small peaks adjacent to larger ones. Since spectroscopic data comprise equidistant discrete data, the second difference method serves as an approximation for detecting peak presence. The negative second difference at point  $i$ ,  $\tilde{y}(i)$  is defined as follows:

$$\tilde{y}(i) = - \sum_{j=i-n}^{i+n} [y(j+H) - y(j)] - [y(j) - y(j-H)] \quad (4)$$

Where:  $H=2n+1$ . Throughout the second-order differencing process, a negative sign precedes the summation operation to orient the extremum of the second-order difference in alignment with the peak direction. Consequently, if the negative second-order difference in proximity to the peak's location surpasses zero, it signifies a local maximum. Deriving from equation(4), it can be deduced that:

$$\tilde{y}(i) = \sum_{j=-n+H}^{n+H} c(j)y(i+j) \quad (5)$$

Here:

$$c(j) = \begin{cases} 2 & -n \leq j \leq n \\ -1 & n < |j| \leq n+H \end{cases} \quad (6)$$

This signifies a symmetric zero-area "window" function, enabling the transformation of the second-order difference of the spectral data into the convolution with function (6). Following further computations and setting  $m = n+H$ , the general expression for the symmetric zero-area transformation can be derived as:

$$\tilde{y}(i) = \sum_{j=-m}^m c(j)y(i+j) \quad (7)$$

$$\sum_{j=-m}^m c(j) = 0, \quad c(j) = c(-j) \quad (8)$$

Where:  $W=2m+1$  denotes the "window width," expressed in channel numbers;  $y(i)$  and  $\tilde{y}(i)$  respectively represent the data of the  $i$ -th channel of the experimental spectrum and the transformed spectrum;  $c(i)$  is the symmetric zero-area "window" function, which is the transformation function. Therefore, the essence of the symmetric zero-area transformation lies in convolution transformation of the symmetric "window" function with zero area alongside the experimental spectrum. It can be readily demonstrated that the "window" function satisfying equation (8) results in zero for the convolution transformation with a linear base, and positive values are only produced when peaks are present, with the maximum transformation value indicating the peak position. The key to the algorithm is selecting an appropriate symmetric zero-area transformation function, and this paper adopts a peak-like function as the transformation function for the symmetric zero-area method.

This type of function can be represented as a difference between a specific symmetric function  $G(i)$  and a constant  $d$ , denoted as:

$$c(j) = G(j) - d \quad (9)$$

To ensure that the condition of equation (8) is satisfied, the following derivation is obtained:

$$d = \frac{1}{W} \sum_{j=-m}^m G(j) \quad (10)$$

Based on the spectral characteristics, the function  $G(j)$  chosen by this algorithm is the Gaussian function.

$$G(i) = \exp[-4\ln 2(j/H)^2] \quad (11)$$

Summarizing the above, the symmetric zero-area transformation function utilized in this paper adopts the peak-like function formula as follows:

$$c(j) = \exp[-4\ln 2(j/H)^2] - \frac{1}{W} \sum_{j=-m}^m \exp\left[-4\ln 2\left(\frac{j}{H}\right)^2\right] \quad (12)$$

In equation (12),  $W=2m+1$  represents the window width, which refers to the variable width.  $H$  is the half-peak width, which corresponds to the channel address.

Given that  $\sigma = FWHM / 2\sqrt{2\ln 2}$  is the standard deviation of the Gaussian function, and  $FWHM$  is the full width at half maximum of the peak, the following relationship can also be derived:

$$c(j) = \frac{8\ln 2}{\sigma^4} [(8\ln(2j^2) - \sigma^2) \exp\{-4\ln[2\left(\frac{j}{\sigma}\right)^2]\}] \quad (13)$$

Initially, the relationship between the full width at half maximum ( $FWHM$ ) and energy undergoes calibration and is stored in memory. During the transformation process, distinct points employ varying  $FWHM$  values to construct the transformation function. The length of the transformation function can be dynamically determined since the standard deviation of the Gaussian function and the number of points in the transformation function adjust according to the  $FWHM$  of the spectral data. This mechanism facilitates an adaptive peak-finding effect.

Typically, the peak value corresponds to the channel address where the ratio of the transformed spectrum to its standard deviation attains its maximum value. The expression for this ratio at that juncture is as follows:

$$SS_i = \frac{\tilde{y}_i}{y_i} = \frac{\sum_{i=-m}^m c(j)y(i+j)}{\left(\sum_{i=-m}^m c(j)y(i+j)\right)^{1/2}} > f \quad (14)$$

Where  $f$  represents a sensitivity parameter or threshold value, a constant. The magnitude of  $f$  inversely affects the number of detected peaks: the larger  $f$ , the fewer peaks detected, with a higher likelihood of missed detections; the smaller  $f$ , the greater the number of detected peaks, with an increased risk of false alarms. Typically,  $f$  is chosen within the range of 1 to 3, contingent upon the spectrum's quality and analytical requisites.

### 3. Experimental and Results Analysis

To evaluate the efficacy of various peak-finding algorithms in spectroscopic analysis, experiments were conducted using a Zeiss XL-30 scanning electron microscope paired with an AMETAK Silicon Drift Detector (SSD) to gather original energy spectra from a diverse range of metals including  $Mg$ ,  $Al$ ,  $Au$ ,  $Ag$ ,  $Zn$ ,  $Fe$ ,  $Cu$ ,  $Mn$ ,  $Sn$ , and non-metals like  $C$ ,  $Si$ , along with several common compounds ( $CeF_3$ ,  $PbS$ ). Subsequently, algorithmic programs based on different peak-finding principles were developed using the MATLAB software platform to process the original spectral data and perform peak-finding analysis.

### 3.1 Characteristic Peak Finding Experiment Analysis

**Table 2.** AMEKEK standard element characteristic peak position

Sample Element	Peak	$K_{\alpha}$	$K_{\beta}$
	C	0.282keV	0keV
	F	0.677keV	0keV
	Mg	1.25keV	0keV
	AL	1.49keV	0keV
	Si	1.74keV	0keV
	S	2.31keV	2.46keV
	Ag	2.98keV	3.15keV
	Sn	3.44keV	3.66keV
	Ce	4.84keV	5.26keV
	Mn	5.9keV	6.49keV
	Fe	6.4keV	7.06keV
	Cu	8.05keV	8.9keV
	Zn	8.64keV	9.57keV
	Au	9.71keV	11.44keV
	Pb	10.55keV	12.61keV

To ensure consistency in algorithm comparison, the experiment set standardized peak-finding parameters: threshold factor ( $f$ ) set at 1, peak-finding window width ( $w$ ) at 15 channels, minimum peak height at 10 counts, and minimum peak width at 5 channels. The  $f$  regulates the influence of noise on peak identification, with higher  $f$  values leading to more conservative peak finding and reducing the misidentification of noise as peaks. The  $w$  defines the range of adjacent data points considered when determining a peak. The minimum peak height and minimum peak width serve as criteria for validating a peak; peaks falling below these predetermined thresholds are disregarded to prevent noise from being misconstrued as characteristic peaks. Table 2 below shows the standard peak positions for some elements from the AMETEK official elemental energy chart. This paper compares the element peaks found by different algorithms with the standard peak positions. Fig. 4 shows the differences between the  $K_{\alpha}$  peaks and the standard  $K_{\alpha}$  peaks, while Fig. 5 shows the differences between the  $K_{\beta}$  peaks and the standard  $K_{\beta}$  peaks.

Upon comparing the peak positions identified by the algorithms with the theoretical peak positions in the AMETEK standard energy spectrum data table, it was observed that all algorithms successfully identified the characteristic peak positions of the target elements. For example, considering pure *Au*, its standard  $K_{\alpha}$  and  $K_{\beta}$  peaks are located at 9.71 keV and 11.44 keV, respectively. The derivative method detected peak values at 9.7291 keV (a difference of 0.0191 keV) and 11.5092 keV (a difference of 0.06922 keV). In contrast, the peak values identified by the derivative method, Gaussian fitting method, and symmetric zero-area method were consistent, at 9.7133 keV (a difference of 0.0033 keV) and 11.492 keV (a difference of 0.052 keV), showing higher peak-finding precision. However, when distinguishing adjacent dense peaks of elements such as *Sn* and *Ag*, the simple comparison method struggled to accurately identify them, whereas the other methods effectively differentiated these dense peaks.

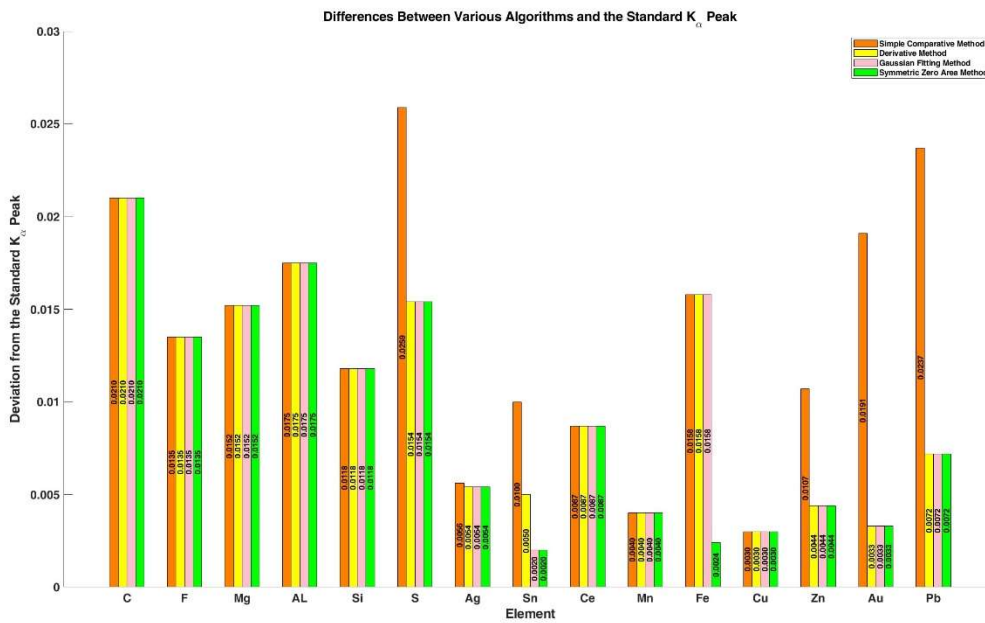


Fig. 4 Differences Between Various Algorithms and the Standard  $K_{\alpha}$  Peak

To visually compare the peak-finding outcomes of different methods, pure Au was chosen as an example, and the peak-finding results of the simple comparison method, derivative method, Gaussian fitting method, and symmetric zero-area method (as shown in Fig. 6 below). As evident from the figure, the first three methods successfully identified the target peaks but struggled to effectively distinguish the false peaks present in the spectrum. Conversely, the symmetric zero-area method adeptly identified and eliminated these false peaks.

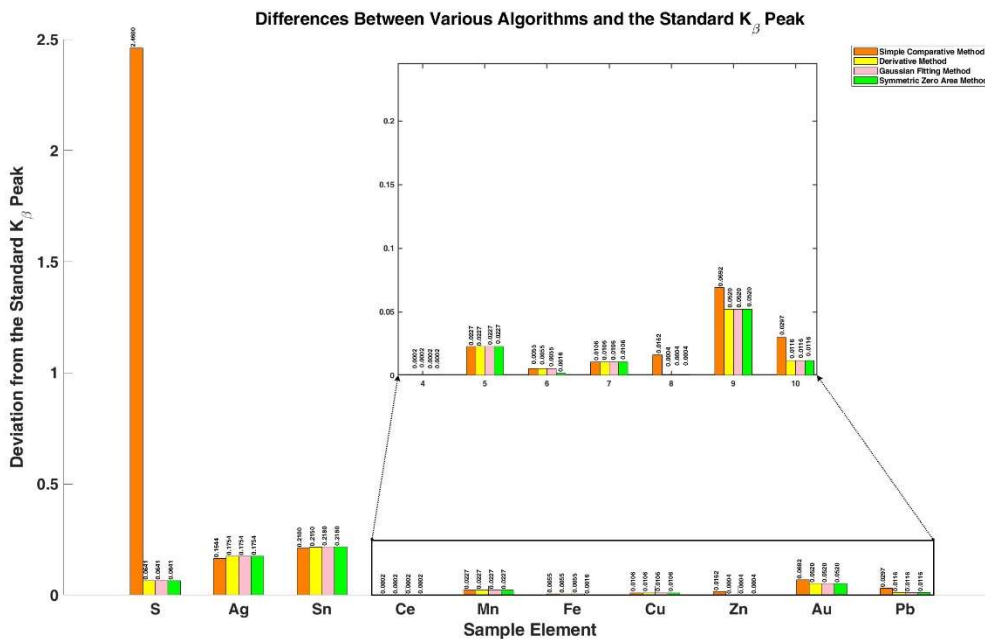


Fig. 5 Differences Between Various Algorithms and the Standard  $K_{\beta}$  Peak

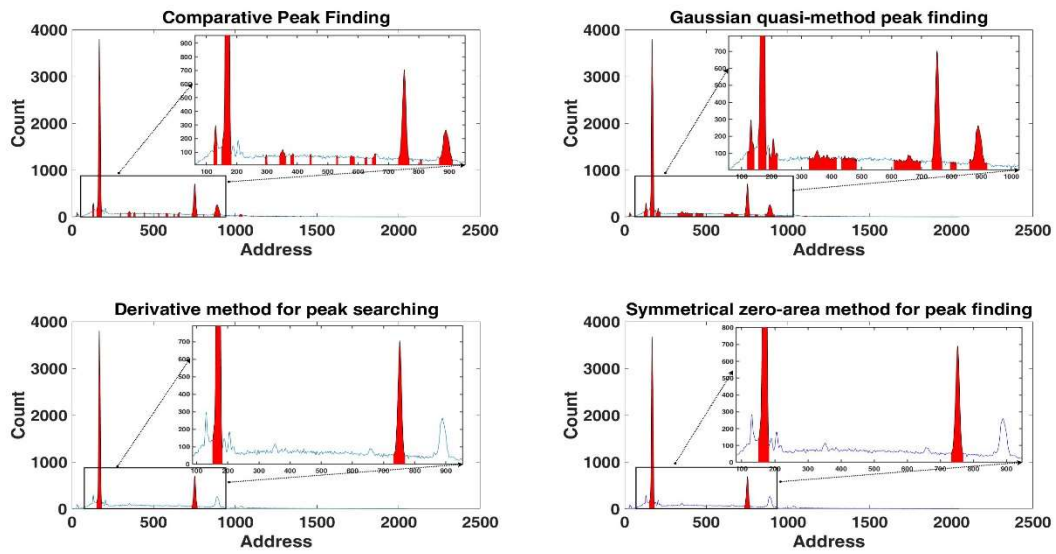


Fig. 6 Comparison of Mn elemental peak search of different algorithms

### 3.2 False Peak Finding Experiment Analysis

To further assess the efficacy of the peak-finding algorithm in real-world scenarios, we conducted false peak finding experiments using various pure metals such as *Mg*, *Al*, *Au*, *Ag*, *Zn*, *Fe*, *Cu*, *Mn*, and non-metallic *Si*, as well as several common compounds (*CeF3*, *SnO2*, and *PbS*). The corresponding data is presented in the table 3 below.

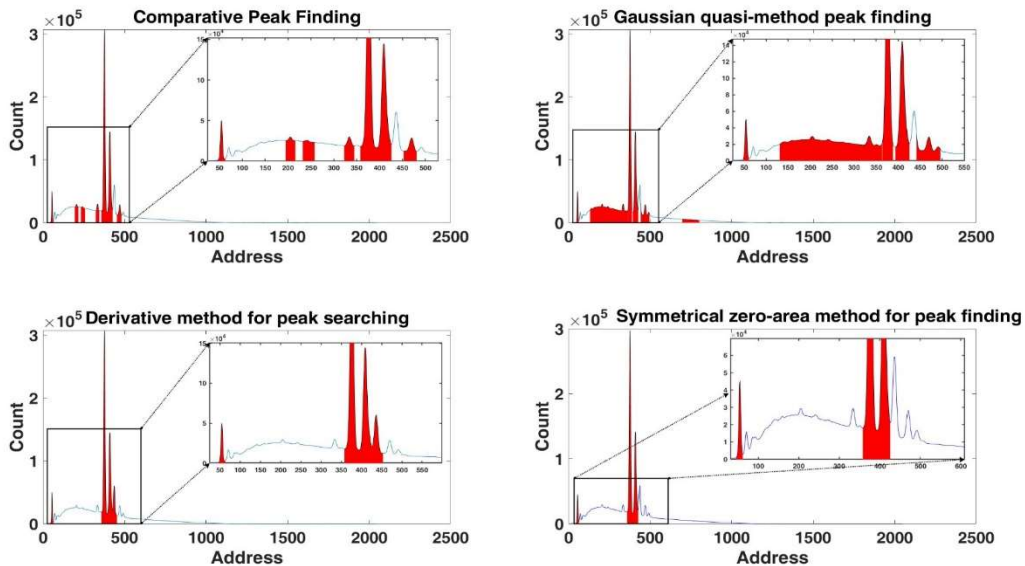


Fig. 7 Comparison of peak finding for compound  $CeF_3$  by different algorithms

The fig. 7 shows the peak detection experimental results using  $CeF_3$  compound as an example. From the figure, it is evident that all algorithms consistently identify the sample's characteristic peaks, achieving a 100% detection rate. However, there are notable discrepancies in the accuracy of peak

positioning and the discernment of false peaks among the four algorithms. The symmetric zero-area method not only accurately located all element characteristic peaks but also exhibits no false positive detections, demonstrating exceptionally high accuracy and specificity. While the simple comparison method and derivative method also correctly identified all characteristic peaks, they mistakenly identified some energy values of *Si*, *Ag*, *Mn*, *PbS*, and *Al* as false peaks, showing slightly poorer interference resistance. The Gaussian fitting method, although accurately locating characteristic peaks, detected more false peaks, especially evident in elements like *Si*, *Ag*, *Al*, *Mn*, *Mg*, and *CeF3*, indicating its susceptibility to noise and background interference.

**Table 3.** Comparison of the number of false peaks found by different peak-finding algorithms

Sample	Number of False Peaks Detected by Simple Comparison Method	Number of False Peaks Detected by Derivative Method	Number of False Peaks Detected by Gaussian Fitting Method	Number of False Peaks Detected by Symmetric Zero Area Method
<i>Si</i>	11	0	3	0
<i>Mg</i>	6	0	2	0
<i>Al</i>	13	11	12	0
<i>Au</i>	24	2	11	0
<i>Ag</i>	21	0	6	0
<i>Zn</i>	22	0	5	0
<i>Fe</i>	12	3	4	0
<i>Cu</i>	9	0	3	0
<i>Mn</i>	4	0	1	0
<i>CeF3</i>	10	1	11	0
<i>SnO2</i>	17	0	5	0
<i>PbS</i>	15	1	4	0

#### 4. Summary

This paper introduces an adaptive peak-finding algorithm rooted in the symmetric zero-area method, leveraging a Gaussian function and convolution window to achieve adaptive optimization of peak-finding thresholds. To thoroughly evaluate the algorithm's performance, peak-finding experiments were conducted across a spectrum of samples, and the results were compared horizontally with those obtained using other classical algorithms. The primary conclusions drawn from this study are as follows:

The symmetric zero-area method employs a convolution window applied to energy spectrum curve, dynamically optimizing the window function's shape and width to adapt to diverse peak characteristics, thereby facilitating precise delineation of peak boundaries. Consequently, in contrast to traditional peak-finding algorithms, the symmetric zero-area method exhibits superior capabilities in identifying dense peaks and eliminating false peaks. When multiple peaks closely adjoin or partially overlap, the symmetric zero-area method can adeptly discern each individual peak, preventing misidentification as a single broad peak and, thereby augmenting peak-finding precision. This algorithm effectively segregates real peaks from noise-induced false peaks, thereby enhancing the reliability of the peak-finding outcomes. Comparative experiments with traditional algorithms such as the derivative method and Gaussian fitting have further corroborate the superior accuracy of the symmetric zero-area peak-finding algorithm in identifying dense and weak peaks. In summation, the symmetric zero-area method offers a refined solution for peak identification in micro-area

spectroscopy, surmounting the constraints of conventional algorithms and laying the foundation for accurate analysis of material chemical components.

## Acknowledgments

The authors gratefully acknowledge the financial support from Research Instrument and Equipment Development Project of the Chinese Academy of Sciences funds.

## References

- [1] Pang Jushou, Zheng Guifang, Hou Xiaofeng. Peak Finding with Symmetric Zero-Area Transformation Method [J]. *Journal of Atomic Energy Science and Technology*, 1987, (03): 270-279.
- [2] Alfonso Fernández-González, Jose Manuel Montejo-Bernardo, Natural Logarithm Derivative Method: A novel and easy methodology for finding maximums in overlapping experimental peaks, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Volume 74, Issue 3, 2009, Pages 714-718, ISSN 1386-1425.
- [3] Ryo Murakami, Yoshitomo Harada, Yutaka Sonobayashi, et al. Correlation analysis with measurement conditions and peak structures in XPS spectral round-robin tests on MnO powder sample, *Journal of Electron Spectroscopy and Related Phenomena*, Volume 264, 2023, 147298, ISSN 0368-2048.
- [4] Li Zhengjie, Yin Bohua, Chu Mingzhang, et al. Study on Real-Time Trapezoidal Shaping Algorithm for Micro-area X-ray Spectroscopy Signals [J]. *Journal of Electron Microscopy*, 2022, 41(03): 234-240.
- [5] Dong Xin, Xu Ruimei, Zhao Wenxia. Study on Interfering Peaks in Micro-area Energy Dispersive X-ray Fluorescence Spectroscopy [J]. *Journal of Electron Microscopy*, 2017, 36(03): 229-233.