An improved SNIP denoising algorithm applied in EDXRF

Liangquan Ge\textsuperscript{a,b}, Hui Li\textsuperscript{b}, Chuanfeng Tang \textsuperscript{b}, Fei Li\textsuperscript{a,b,*}, Kai Zhu\textsuperscript{b}, Jin Yang\textsuperscript{b}

\textsuperscript{a} Applied Nuclear Technology in Geosciences Key Laboratory of Sichuan Province, China
\textsuperscript{b} College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, China

Corresponding author and e-mail: Fei Li, lifeimvp@sina.com

Abstract. Based on the phenomenon of peak height reduction in existing denoising algorithms, this paper modifies the traditional SNIP algorithm, proposes an improved SNIP algorithm and applies it to EDXRF denoising. Two evaluation parameters RMSE and PCC are used to compare the improved SNIP algorithm with the traditional SNIP algorithm. The experimental results show that the improved SNIP algorithm has great advantages both in similarity and relevance, and the results of the improved SNIP algorithm are more consistent with the characteristics of the original data peak than the original SNIP algorithm, which is beneficial to the qualitative and quantitative analysis of the energy dispersive X-ray fluorescence spectrum.

1. Introduction
The energy dispersive X-ray fluorescence spectrum is the characteristic X-ray peak and continuous background formed by a series of effects produced by the interaction between photons and substance. Before EDXRF can be used for qualitative and quantitative analysis of the samples to be measured, it is necessary to denoising, background subtraction and overlapping peak separation. Among them, denoising is one of the key links of EDXRF, which has a direct impact on the analysis results. The algorithms commonly used for denoising are the least square method \cite{1}, the wavelet transform method \cite{2}, the Fourier transform method \cite{3}, and so on. The experimental results show that these algorithms can effectively eliminate the statistical fluctuations, and can largely retain the energy spectrum information while eliminating the noise. SNIP algorithm commonly used in background subtraction can effectively suppress the influence of the substrate, and has a strong weakening effect on the influence of background factors such as background peak and scattering ray peak \cite{4}. At the same time, these algorithms also have the disadvantage of reducing the spectral peak height, which is not conducive to the subsequent quantitative analysis. And most of these algorithms are gamma spectrum denoising, but few of them are used in X-ray spectrum denoising, because the peaks of X-ray spectrum are relatively dense and the matrix effect can not be avoided. In this paper, we improve the SNIP algorithm and apply it to X-ray spectrum denoising. The experimental results show that it can effectively improve the phenomenon of energy spectrum peak height reduction.

2. Material and Method

2.1. Principles and disadvantages of the traditional SNIP algorithm
The idea of snip algorithm applied to background subtraction of gamma energy spectrum was first proposed by Ryan et al. \cite{5-7}.
The actual steps to deduct the background of gamma energy spectrum are as follows:

1) carry out logarithmic transformation (LLS transformation) for each count value, and the transformation formula is as follows:

\[ y(i) = \ln(\ln(x(i) + 1) + 1) \]  \hspace{1cm} (1)

Where \( i \) is the corresponding count, \( x(i) \) is the channel address, and \( y(i) \) is the operation result. The natural logarithm operator in the formula is mainly used for processing high count peaks.

2) iterate over the energy spectrum data according to the SNIP algorithm, and the formula is as follows:

\[ y_r(i) = \min \left( y(i), \frac{y(i-m) + y(i+m)}{2} \right) \]  \hspace{1cm} (2)

Where \( m \) is the artificially set parameter, which is related to the filter window width \( W \) (\( W=2m+1 \)). In the traditional SNIP algorithm, it is 4 times of the peak area half height width (FWHM). The smaller value of \( y(i) \) with the current channel address \( i \) and the mean value of channel address \( y(i \pm m) \) at both ends of the filter window is taken as the result \( y_r(i) \).

3) when all \( y_r(i) \) are obtained, the background spectrum of the full spectrum can be obtained by inverse LLS operation.

The traditional SNIP algorithm usually uses a single window width when filtering and stripping the energy spectrum. The filter width is determined by the width of the spectrum peak, and the width of the spectrum peak generally increases with the increase of energy. Choosing different window width will lead to different results. For example, the window width suitable for the low energy band of gamma peak may cause the net count in the high energy band to be deducted as the background, while the window width suitable for the high energy band of gamma peak may cause the background in the low energy band to be treated as the net count without being completely deducted [8]. Therefore, using a single filter width will reduce the effect of the algorithm.

2.2. The existing improved SNIP algorithm and its shortcomings

At present, there are some shortcomings in the existing improved SNIP algorithms. Their differences lie in the calculation method of filter window width:

1) the width of the initial filter window is 3 times of the maximum value of the half height width of the energy spectrum, and then decreases to 1 [9];

2) the filter window width of each peak area is the channel value corresponding to the peak width, and the peak area of the non-omnipotent peak is 0 [10];

3) according to the characteristics of Gauss distribution, the left and right boundaries of all-around peaks are redetermined, the approximate peak boundary is determined by the first derivative method, and then the energy spectrum is stripped [11];

4) the peak width is obtained by using the functional relationship between peak width and energy, and the adaptive filter width is obtained [12].

The above methods are mainly focused on the improvement of filter width, but the improvement of filter width has no effect on the situation that the peak value of SNIP algorithm decreases after iterations.

2.3. Improved SNIP algorithm

The purpose of this paper is to find the functional relationship between the peak decline ratio of energy spectrum, the channel address (energy) and the number of iterations, and then replace it to achieve peak reduction without affecting the algorithm effect, so as to effectively avoid the situation that the peak decline ratio will gradually increase with the increase of energy after multiple iterations of SNIP.

After several iterations of SNIP algorithm, compared with the decline ratio of the original data to the processed data, the fitting is performed (see Figure 1). The energy spectrum address (energy) is used as the abscissa and the descent ratio as the ordinate. The fitting curve conforms to the exponential function distribution, so the exponential fitting is adopted.
It can be seen from Figure 1 that the decline proportion is more and more close to 1 with the increase of channel address (energy). The general EDXRF spectrum channel address is generally 1024 or 2048, the decline ratio is 40% to 70%, and the peak shape will have obvious distortion. After multiple iterations, the decline ratio is fitted (see Figure 2 and figure 3). Among them, the number of iterations is abscissa and the descending scale is ordinate.

![Graph](image1.png)

**Figure 1.** The change of decline ratio with channel

![Graph](image2.png)

**Figure 2.** The change of decline ratio with denoising times
Figure 3. (a) The amplification area with 0-10 of denoising times, (b) The amplification area with 10-20 of denoising times.

The comparison shows that if the number of iterations is less than 10, the influence factor is more in line with the exponential relationship, and if the number of iterations is more than 10, it is more in line with the linear relationship.

After comprehensive consideration, the following formula is obtained:

\[
f(i)' = A(M)(1 + B(i))f(i)
\]

The iterative influence formula: 
\[
A(M) = \begin{cases} 
-9.605241(1 - 1.12927^M) & M \leq 10 \\
-7.65184 + 3.03598M & M > 10 
\end{cases}
\]

The channel address influence formula: 
\[
B(f(i)) = 0.9303(1 - \exp(-8.6101e - 4 * f(i)))
\]

Where \(i\) is the energy spectrum channel address, \(f(i)\) is the original signal, \(f(i)\)' is the signal after denoising, and \(M\) is the number of iterations.

3. Results and discussion

The evaluation parameters selected for the improved SNIP algorithm are root mean square error (RMSE) and Pearson correlation coefficient (PCC). RMSE is related to the similarity of the two signals, while PCC is related to the correlation. The new improved SNIP algorithm proposed in this paper is used to reduce the noise of energy spectrum, and its denoising effect is compared with that of traditional SNIP algorithm.

A set of X-ray fluorescence data was iterated for 15 times, and then the average RMSE and PCC were compared and tested successively (see Table 1).
Table 1. The comparison between the traditional SNIP algorithm and the improved SNIP algorithm

| 传统 SNIP 算法 | 改进 SNIP 算法 |
|-------------|-------------|
| 迭代次数 | RMSE     | PCC      | RMSE     | PCC      |
| 1          | 140.1989  | 0.9962   | 35.2699  | 0.9938   |
| 2          | 215.0684  | 0.9928   | 84.7045  | 0.9942   |
| 3          | 256.1275  | 0.9862   | 96.48    | 0.9902   |
| 4          | 280.0102  | 0.9766   | 105.2399 | 0.9819   |
| 5          | 294.8753  | 0.9647   | 118.7876 | 0.9704   |
| 6          | 304.6651  | 0.9511   | 136.3097 | 0.9567   |
| 7          | 311.4143  | 0.9362   | 156.3736 | 0.9416   |
| 8          | 316.2471  | 0.9205   | 178.5169 | 0.9256   |
| 9          | 319.8148  | 0.9042   | 203.0683 | 0.9089   |
| 10         | 322.5162  | 0.8874   | 230.4694 | 0.8918   |
| 11         | 324.6074  | 0.8704   | 244.6847 | 0.8746   |
| 12         | 326.2595  | 0.8535   | 259.3096 | 0.8573   |
| 13         | 327.5879  | 0.8366   | 273.4805 | 0.8401   |
| 14         | 328.6718  | 0.8198   | 287.3485 | 0.8231   |
| 15         | 329.5669  | 0.8032   | 301.111  | 0.8062   |

RMSE: root mean square error  
PCC: Pearson correlation coefficient

It can be concluded from the Table 1 that after different iterations, the RMSE obtained by the improved SNIP algorithm is far smaller than that obtained by the traditional SNIP algorithm, and the PCC obtained by the improved SNIP algorithm is also significantly larger than that obtained by the traditional SNIP algorithm. It can be concluded that the improved SNIP algorithm is superior to the traditional SNIP algorithm in both similarity and correlation. With the increase of iterations, RMSE and PCC tend to be stable.

The traditional SNIP algorithm and the improved SNIP algorithm are used to reduce the original data (Figure 4. (a)) 5 times respectively (see Figure 4). It can be seen intuitively from the Figure 4 that with the increase of iteration times, the peak height obtained by the traditional SNIP algorithm drops more and more obviously. Even after the fifth iteration, the weak peak is almost lost, and the feature peak is also greatly affected, which is not conducive to the extraction of the feature peak information. Although the improved SNIP algorithm reduces the peak height to some extent, it has little influence and retains the weak peak relatively completely, which is beneficial to the subsequent qualitative and quantitative analysis.
Figure 4. (a) The original data, (b) The result of denoising once, (c) The result of denoising twice, (d) The result of denoising three times, (e) The result of denoising four times, (f) The result of denoising five times.

4. Conclusion
In this paper, a new and improved SNIP algorithm is proposed and applied to EDXRF denoising simultaneously with the traditional SNIP algorithm. RMSE and PCC of energy spectrum processed by the two methods are calculated respectively for comparison. The results show that after several iterations, the improved SNIP algorithm is obviously superior to the traditional SNIP algorithm in terms of peak retention, making the results more consistent with the characteristics of the original data peak. Therefore, the denoising of energy spectrum can make the characteristic peak of the element more obvious, and the recognition effect can be improved obviously, which is more conducive to accurately finding weak peak and heavy peak.
Acknowledgement
This work was funded by the National Natural Science Foundation of China (No.41774147), the National Key R&D Project (No.2017YFC0602100) and Science and technology activity project for overseas personnel in Sichuan province.

References
[1] ZHU Meng-Hua, LIU Liang-Gang, QI Dong-Xu, YOU Zhong, XU Ao-Ao. Least square fitting of low resolution gamma ray spectra with cubic B-spline basis functions. Chinese Physics C. 2009, 33(1):24-30.
[2] Davide Bianchia, Claudia Lenauera, Gerhard Betzb, András Vernes. A wavelet filtering method for cumulative gamma spectroscopy used in wear measurements. Applied Radiation and Isotopes. 2017,120:51-59.
[3] M.J. Safari, F. Abbasi Davani, H. Afarideh, S. Jamili, E. Bayat. Discrete Fourier Transform Method for Discrimination of Digital Scintillation Pulses in Mixed Neutron-Gamma Fields. IEEE Transactions on Nuclear Science. 2016, 63(1):325-332.
[4] WU Hexi, LIU Qingcheng, YANG Bo, LIU Yujuan. SNIP algorithms applied to analyze γ-ray spectra of natural radioactivity nuclides. He Jishu/Nuclear Techniques. 2010, 33(7):513-516.
[5] Omer M. Analysis of SNIP algorithm for background estimation in spectra measured with LaBr3:Ce detectors [C]. Zero-Carbon Energy Kyoto 2012, Tokyo. Springer: 245-252.
[6] Miroslav Morháč. An algorithm for determination of peak regions and baseline elimination in spectroscopic data. Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment. 2009, 600(2): 478 - 487.
[7] WU Hexi, LIU Qingcheng, YANG Bo, et al. SNIP algorithms applied to analyze γ-ray spectra of natural radioactivity nuclides[J]. Nuclear Techniques, 2010, 33(7): 513-516.
[8] LIU Yonggang. Research on decomposition method of γ spectrometer[D]. Beijing: China University of Geosciences, 2011.
[9] LONG Bin, FENG Tiancheng, SU Chuanying, et al. a self-adaptive method for the clipping of scatter background of γ spectrum[J]. Nuclear Electronics & Detection Technology, 2013, 33(10): 1293-1296.
[10] ZHANG Yupeng, XIAO Xiaoling, ZHANG Xiang, et al. Research on elimination background by extending clipping window of snip algorithm[J]. China Energy and Environment Protection, 2017, 39(3): 143-146.
[11] WANG Yiming, WEI Yixiang. Baseline elimination method for γ - ray spectra based on improved snip algorithm[J]. Nuclear Electronics & Detection Technology, 2012, 32(12): 1356-1359.
[12] GAO Yan, ZENG Guoqiang, GU Min, YAN Lei, LUO Mingtao, HOUYang, ZHU Zhu ,TANG Wei. Development of adaptive width SNIP algorithm in nuclides identification instrument[J]. Nuclear Techniques,2019,42(6):060401.