Improved Wavelet Packet Noise Reduction for Microseismic Data via Fuzzy Partition

Zhiqiang Lan, *Graduate Student Member, IEEE*, Peng Gao, Peng Wang, Yaojun Wang, Jiandong Liang, and Guangmin Hu

***Abstract*—It is crucial to retain as many signal components as possible, while noise is eliminated for wavelet packet noise reduction algorithms. Conventional thresholding functions take coefficients smaller than a given threshold as the noise component and get them removed in many ways. In this letter, considering the potential probability of being the signal-related component of coefficients whose value is smaller than the given threshold and the nonnegligible fact that microseismic (MS) event is sparse compared with the noise, we proposed a new wavelet packet-based denoising method via fuzzy partition. First, instead of a hard partition from a given threshold, wavelet packet coefficients get reduced by a fuzzy partition to retain more potential signal elements and suppress the noise. Second, we employ fuzzy c-means (FCM) clustering to identify the interval period of the MS event further to remove the residual noise and additional resonance in processed time series. We tested our method on synthetic datasets and real-field data from an MS monitoring experiment in a coal mine in Sichuan Basin, China. We utilize Pearson correlation coefficient and root-mean-square error between ideal signal and denoised data as performance indicators in synthetic tests, while sample entropy and kurtosis of denoised data are involved in the real field dataset. Test results from synthetic datasets and the real field dataset demonstrate that the proposed noise reduction method is superior to traditional hard-, soft-, and garrote-thresholding and is more applicable and effective in MS data processing.**

***Index Terms*—Clustering, denoising, fuzzy partition, microseismic (MS).**

I. Introduction

PASSIVE microseismic (MS) monitoring is an effective way to understand underground structures and processes by analyzing weak signals caused by subsurface activities. It gradually becomes one of the most potent tools used for safe mining and geological hazard early-warning during the past few decades [1]. The weak signals mentioned earlier, which characterize the underground status and changes, are so-called MS events. Typically, MS data are contaminated by background noise and have a low signal-to-noise ratio (SNR). This tricky situation poses enormous difficulty on MS event identification, imaging, and other relevant analysis [2]. Due to the benefits of noise reduction to MS event localization and further applications, noise attenuation has become an important research field in MS data processing [3].

Different from fast fourier transform (FFT) [4] and empirical mode decomposition (EMD)-based methods [5], wavelet transform is a computation-efficient and powerful nonstationary signal time-frequency analysis tool with complete theory support [6]. As the extension of the wavelet transform, wavelet packet transform (WPT) simultaneously decomposes the approximate and detail component to provide richer time-frequency information [7]. WPT organizes decomposition coefficients into disparate leaf nodes representing different frequency subbands. A proper threshold is then applied to coefficients by a certain way called thresholding function to remove noise elements with small impacts on signal parts. Finally, reconstruction is invoked to rebuild data with noise reduced.

Hard- and soft-thresholding are two of the most widely used thresholding methods and commonly lead to satisfactory results. However, the former causes additional resonance in processed data because of discontinuous coefficient thresholding, and the latter brings about information loss due to excessive shrinkage of signal-related coefficients. Various tradeoff approaches have been proposed to minimize noise residual and signal information loss with diverse continuous strategies [8]. However, they all depend on the hard partition of coefficients and ignore potential signal-related coefficients smaller than the given threshold. Moreover, it is difficult to doubtlessly determine the coefficient with a value closer to the given threshold to be either signal-related or noise-related part.

This letter proposed an improved noise reduction method for MS data based on WPT and fuzzy partition, mainly consisting of three steps: node classification, signal-node fuzzy shrinkage, and time-domain identification. Node classification divides WPT nodes into pure noise nodes or signal nodes by a decomposition-level-dependent threshold. We directly set the pure noise node to a zero vector. In the second step, instead of a hard partition with a given threshold, we introduce a fuzzy degree function to indicate the probability of the coefficient being the signal element and perform nonlinear shrinkage according to the corresponding probabilistic. Finally, considering the MS event is sparse in the time domain and the adverse impacts of imbalanced cluster size, we further remove noise elements in processed data by identifying the interval period of the MS event with an improved cluster contribution weighted fuzzy c-means (FCM) clustering algorithm.

Manuscript received May 24, 2021; revised June 29, 2021; accepted July 14, 2021. Date of publication August 2, 2021; date of current version December 31, 2021. *(Corresponding author: Peng Gao.)*

Zhiqiang Lan, Peng Gao, Peng Wang, Yaojun Wang, and Jiandong Liang are with the School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: lanzhiqiang@live.com; penggao@uestc.edu.cn).

Guangmin Hu is with the School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China.

Digital Object Identifier 10.1109/LGRS.2021.3098057

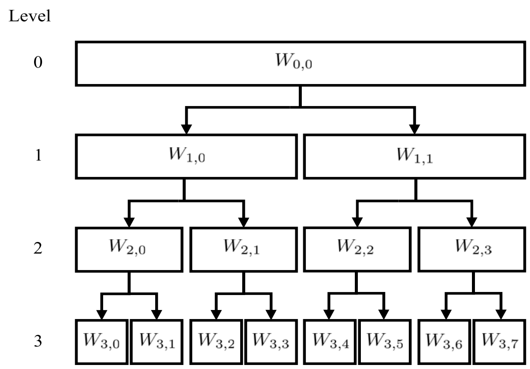

Fig. 1. Three-layer wavelet packet decomposition.

The structure of this article is summarized as follows. Section II first introduces the background of the WPT-based noise suppression method and the improved FCM, and it then details the proposed noise reduction method. Performance evaluation is given in Section III. Finally, Section IV concludes this article.

II. Methodology

A. Background of Wavelet Packet Noise Reduction

As the extension of wavelet transform, WPT simultaneously decomposes low- and high-frequency parts to provide better time-frequency resolution. Let f_s be the sample frequency, subband of node $W_{i,j}$, the j-th node in level i, is then represented by $[(f_{\mathrm{sj}}/2^{i+1}), (f_s(j+1)/2^{i+1})]$. Fig. 1 shows the overview of a three-layer WPT, where $W_{0,0}$ is actually the raw time series. For noisy data X and a certain decomposition level L, we assume that $F(\cdot)$ is the operator of WPT, while $F^{-1}(\cdot)$ is the corresponding inverse operation, and the denoised result $\hat{X}$ is obtained from following steps:

$$\Phi = F(X) \tag{1}$$

$$\hat{\Phi} = \mathrm{THR}(\Phi, \Lambda) \tag{2}$$

$$\hat{X} = F^{-1}(\hat{\Phi}) \tag{3}$$

where Φ is the set of WPT nodes in decomposition level L, while $\hat{\Phi}$ denotes the set of corresponding processed coefficient nodes, $\mathrm{THR}(\cdot,\cdot)$ represents the shrinkage operation, and Λ is the set of threshold $\lambda_{i,j}$ utilized for processing of corresponding node $W_{i,j}$.

Since signal elements produce larger coefficients in wavelet domain than the noise after wavelet transformation, thresholding functions regard coefficients smaller than a given threshold as noise components and eliminate noise by reducing the magnitude of noise-related coefficients. Here, we take the hard-thresholding function as an example for illustration of wavelet coefficients thresholding, and it is defined as follows:

$$\mathrm{THR}_h(\Phi, \lambda) = \begin{cases} \Phi, & |\Phi| > \lambda \\ 0, & \text{others} \end{cases}. \tag{4}$$

A universal threshold for noise suppression in WPT is

$$\lambda = \sigma\sqrt{2\log\left(N \log_2 N\right)} \tag{5}$$

where $\sigma = \mathrm{MAD}_{1,1}/0.6745$ is the appropriate estimate of noise energy under the assumption that noise is Gaussian distributed and dominantly contributes to the high-frequency response of the measured data [9], $\mathrm{MAD}_{1,1}$ represents the absolute median estimated at $W_{1,1}$, and N is the number of data points in sequence X.

B. Preliminary of Fuzzy C-Means Clustering

FCM is an objective function-based clustering method, and its clustering procedure is equivalent to minimizing the objective function. To reduce the adverse impact from the imbalanced size of clusters, we introduced a cluster contribution weighted-based FCM from [10]. For samples with n members $X = \{x_1, x_2 \ldots, x_n\}$, let $\nu = \{\nu_1 \ldots, \nu_c\}$ represent cluster centers and μ denote the membership degree matrix, objective function of the improved FCM is given as

$$J_m = \sum_{i=1}^{n}\sum_{j=1}^{c} \frac{\mu_{\mathrm{ij}}^m d_{\mathrm{ij}}^2}{\eta_j} \tag{6}$$

where $m > 1$ is the fuzziness index, $c > 1$ denotes the number of clusters, η_j represents size of cluster ν_j, and d_{ij} is the distance between x_i and ν_j. We can easily obtain update formulas to minimize the objective function with the Lagrange multiplier method, and they are

$$\nu_j = \frac{\sum_{i=1}^{n} \mu_{\mathrm{ij}}^m x_i}{\sum_{i=1}^{n} \mu_{\mathrm{ij}}^m} \tag{7}$$

$$\mu_{\mathrm{ij}} = \frac{\left(\frac{\eta_j}{d_{\mathrm{ij}}^2}\right)^{\frac{1}{m-1}}}{\sum_{k=1}^{c}\left(\frac{\eta_k}{d_{\mathrm{ik}}^2}\right)^{\frac{1}{m-1}}}. \tag{8}$$

Since it is difficult to know size of clusters previously, η_j is estimated and updated from

$$\eta_j = \sum_{i=1}^{n} \mu_{\mathrm{ij}}. \tag{9}$$

Therefore, procedures of clustering with the improved FCM are summarized as follows.

1) Initialize cluster centers ν.
2) Initialize all members of cluster size η with value one.
3) Update membership degree matrix μ with (8).
4) Estimate cluster size with (9) and update cluster centers ν with (7).
5) Repeat steps 3) and 4) until $|J_m(R) - J_m(R-1)| < \epsilon$ or $R > R_{\max}$, where ϵ is a predefined small positive value, R represents iteration counts, and $R_{\max}$ is the preset upper limitation of iteration round.

C. Proposed Noise Reduction Method

Fig. 2 shows the workflow of the proposed noise reduction method. As the common parts of the WPT-based noise reduction approach, WPT decomposes noisy raw data $X = \{x_1, x_2, \ldots, x_N\}$ into multiple nodes containing coefficients representing information of different frequency bands, while IWPT is to reconstruct the time domain data with processed coefficients. Details of the main three procedures in the proposed method: node classification, signal-node fuzzy shrinkage, and time identification will be introduced in the following.

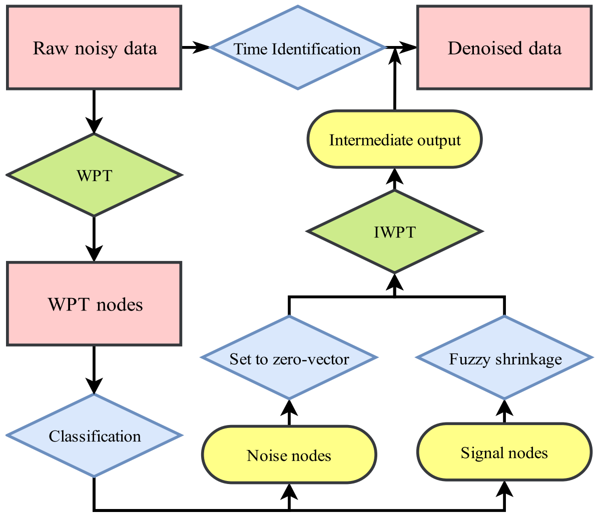

Fig. 2. Workflow of proposed noise reduction method.

1) Node Classification: Node classification is to identify whether a WPT node contains signal components or is pure noise. The division is conducted by the level-dependent threshold from [11], and its definition is

$$\lambda_i = \frac{\sigma\sqrt{2\log(N)}}{\log(i+1)} \tag{10}$$

where N is exactly number of data points in X. Node $W_{i,j}$ is classified as a signal node only when the maximum value of absolute $W_{i,j}$ is not less than λ_i. Otherwise, it is regarded as a pure noise node and is directly set to zero vector.

2) Signal-Node Fuzzy Shrinkage: Conventionally, only the coefficient greater than the given threshold is regarded as the signal component. However, from the perspective of fuzzy partition, coefficients whose value is smaller than the given threshold are also potential signal elements. Directly dropping these potential useful elements can lead to information loss in reconstructed data.

In this article, inspired by fuzzy entropy [12], we proposed a new fuzzy shrinkage method for noise removing in the WPT node. For a given fuzzy tolerance β, the fuzzy shrinkage factor θ of coefficient x is defined as

$$\theta(x,\beta) = 1.0 - \exp\left(-\frac{|x|^\alpha}{\beta}\right) \tag{11}$$

where $\alpha > 1$ is the fuzziness index. In this article, we use $\alpha = 2.0$, and β is the same to the threshold for simplicity. We then easily obtain the processed signal node $\hat{W_{\mathrm{ij}}}$ from

$$\hat{W_{i,j}} = \theta\left(W_{i,j},\lambda_i\right)W_{i,j}. \tag{12}$$

3) Time Identification: After coefficients shrinkage, IWPT is called to construct the denoised intermediate data X_{im}. We then employ the improved FCM mentioned in Section II-B to obtain the time interval the MS event exists. Instead of the intermediate output, noisy raw data are involved in time identification to avoid causing more information loss. Data points out of the MS event interval from the intermediate data are reduced. Therefore, noise and additional resonance get removed further.

In detail, characteristic factors: kurtosis K, skewness S, and Δ, get involved for input information enrichment in FCM. Their definitions are given as follows, respectively:

$$K_i = \sum_{i-w}^{i+w}\frac{\left(x_i-\bar{X}\right)^4}{\delta^4} \tag{13}$$

$$S_i = \sum_{i-w}^{i+w}\left|\frac{\left(x_i-\bar{X}\right)^3}{\delta^3}\right| \tag{14}$$

and

$$\Delta_i = \sqrt{\sum_{i-w}^{i+w}(x_i-A_i)^2} \tag{15}$$

where $A_i = \sum_{i-w}^{i+w} x_i$ is the windowed average amplitude, and w represents half-width of the computation window.

Apparently, parameter in FCM is set with $c = 2$, and the fuzziness index we used here is 2.0 from [13]. A membership matrix μ is obtained from X after clustering. Since the MS event is much shorter than noise, we can easily obtain the vector μ^s, which describes the membership degree of data points belonging to signal cluster, from

$$\mu^s = \left\{\mu_{\mathrm{ik}}|i\in[1,N], k = \arg\min_j \eta_j\right\} \tag{16}$$

where η_j is the estimated cluster size from (9). We then obtain the time-domain shrinkage coefficients μ_s of the intermediate data via a threshold from [13], that is

$$\mu_s(i) = \begin{cases}\mu^s(i), & \mu^s(i) < 0.25\\ 1.0, & \text{others.}\end{cases} \tag{17}$$

Therefore, the final data $\hat{X}$ with noise removed is obtained from

$$\hat{X} = \mu_s X_{\mathrm{im}} \tag{18}$$

III. Performance Evaluation

In this section, we carried out tests on synthetic datasets and the real-field dataset to evaluate the reliability of the proposed noise reduction method. The real-field dataset is from the MS monitoring experiment conducted in Jiayang Coal Mine, Sichuan Basin, China. If not specially stated, traces from neither synthetic datasets nor the real-field dataset use a sample interval of 0.001 s and have the same length of 1000 data points. Considering the character of MS events, the mother wavelet used here is *db*8, and we fix the decomposition level to 3.

A. Synthetic Data

In synthetic datasets, due to the advantage of known ideal signal, both the absolute Pearson correlation coefficient P and root-mean-square error (RMSE) between the denoised data and the ideal signal are utilized as performance indicators [14]. Their definitions are given as follows, respectively:

$$P(V_1,V_2) = \frac{|E(V_1V_2)-E(V_1)E(V_2)|}{\sqrt{E\left(V_1^2\right)-E^2(V_1)}\sqrt{E\left(V_2^2\right)-E^2(V_2)}} \tag{19}$$

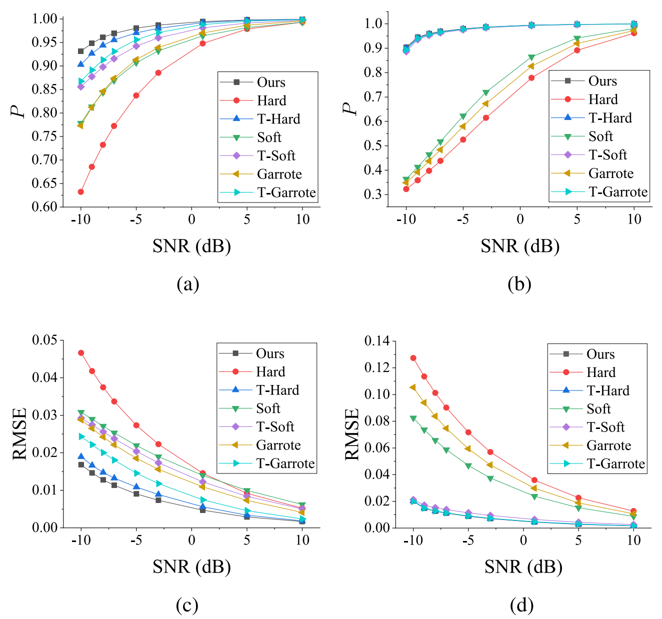

Fig. 3. Noise reduction performance comparison on synthetic data. *Ours* denotes the proposed method, *Hard*, *Soft*, and *Garrote* represent hard-thresholding, soft-thresholding, and garrote-thresholding, respectively. The prefix $T-$ means that of proposed *time-domain identification* is extra applied. (a) p in WGN test. (b) p in pink noise test. (c) RMSE in WGN test. (d) RMSE in pink noise test.

and

$$\mathrm{RMSE}(V_1, V_2) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(V_1(i) - V_2(i))^2} \tag{20}$$

where V_1 and V_2 are two time series with the same length of N, and $E(\cdot)$ denotes the operator of expectation. For noisy data X and corresponding denoised result $\hat{X}$, a larger $P(X, \hat{X})$ or a smaller $\mathrm{RMSE}(X, \hat{X})$ obviously represents a better noise reduction performance.

Synthetic datasets are composed for reliability verification of the proposed noise reduction method first. Ricker wavelet with a dominant frequency of 150 Hz is used as the ideal signal here. There are two test datasets composed of the ideal signal added with noise at a certain SNR. To better simulate real environment noise, White Gaussian noise (WGN) is employed as the noise source in one dataset, while in another is pink noise [15]. In this article, SNR is defined as

$$\mathrm{SNR} = 10\log_{10}\frac{\sum_t |s(t)|^2}{\sum_t |x(t) - s(t)|^2} \tag{21}$$

where $s(t)$ denotes the ideal signal, while $x(t)$ is the noisy data. There are 1000 noisy traces in each dataset at each SNR. We also conducted the same tests on hard-, soft- [16], and garrote-thresholding [17] for comparison.

Fig. 3 shows the results of synthetic data experiments. In both WGN and pink noise tests, the proposed step of time-domain identification is effective in noise reduction performance improvement, and our denoising method produces higher P and lower RMSE compared with other methods. Meanwhile, we noticed that our method nearly reaches a result of P of at least 0.9 and RMSE of at most 0.02 in tests of WGN and pink noise, respectively, which is better than others.

TABLE I

RESULT COMPARISON ON REAL-FIELD DATA

Method	K_g	SampEn
None	15.065	0.887
Ours	**57.364**	**0.008**
Hard Thresholding	19.448	0.214
Soft Thresholding	31.013	0.119
Garrote Thresholding	26.606	0.145

In conclusion, the proposed noise reduction method outperforms traditional thresholding function-based approaches.

B. Real-Field Data

In the real-field dataset, absolute Pearson correlation coefficient and root-mean-square error are inappropriate for performance evaluation due to the unknown pure signal. Similar to information entropy, sample entropy is another potent tool to evaluate the randomness of data. Its magnitude is positively related to noise energy intensity in data [18]. Sample entropy of denoised data is utilized as the performance indicator here. For data $X = \{x_1, x_2 \ldots, x_n\}$ with n data points and a given slice length m, the sample entropy can be obtained from following steps.

1) Obtain $n - m$ slice vectors Y_1 to Y_{n-m}, where $Y_i = \{x_i \ldots, x_{i+m-1}\}$.
2) Calculate the Chebyshev distance d_{ij} between Y_i and Y_j.
3) Define $B_i^m(r)$ for each Y_i, $B_i^m(r) = (\|d_{\mathrm{ij}} < r\|_0 \; n - m + 1)$, where $j \in [1, n-m]$, $\|\cdot\|_0$ is the operator of $l_{0-\mathrm{norm}}$, r is a threshold, and its value is usually 0.1–0.25 times standard deviation of X.
4) Calculate average $B_i^m(r)$ for each Y_i from $B^m(r) = (1/n - m)\sum_{i=1}^{n-m} B_i^m(r)$.
5) Let $m = m + 1$, repeat steps 1)–3) to obtain $B^{m+1}(r)$.
6) Calculate the estimated sample entropy of X from $\mathrm{SampEn}(X, m, r) = -\ln(B^{m+1}(r)/B^m(r))$.

A lower sampEn means that the noise is removed more clearly. Besides, since the MS event is a kind of pulse signal, we also employ kurtosis to indicate its intensity. Similar to (13), global kurtosis K_g of data X is defined as

$$K_g = \frac{1}{\delta^4}E\left((X - E(X))^4\right) - 3 \tag{22}$$

where $E(\cdot)$ also denotes the operator of average, and δ is the standard deviation of X. Because of enhanced pulse tendency after noise reduction, a higher K_g represents a better denoising performance.

There are 258 noisy traces analyzed in the real-field dataset. Table I shows the experiments results, where *None* means no noise reduction is applied. As we can see, our method obtains the maximum K_g and minimum SampEn in the real-field dataset. Fig. 4 gives an example from the real field dataset for detail illustration, where the left part shows time series, while the right shows the frequency-domain information.

In detail, through our noise reduction procedure, K_g of the example data increased from 25.985 to 61.194, and SampEn decreased from 0.759 to 0.004.

Through the test results from both synthetic datasets and the real-field dataset, it is not difficult to find that the noise reduction performance of the proposed method outperforms

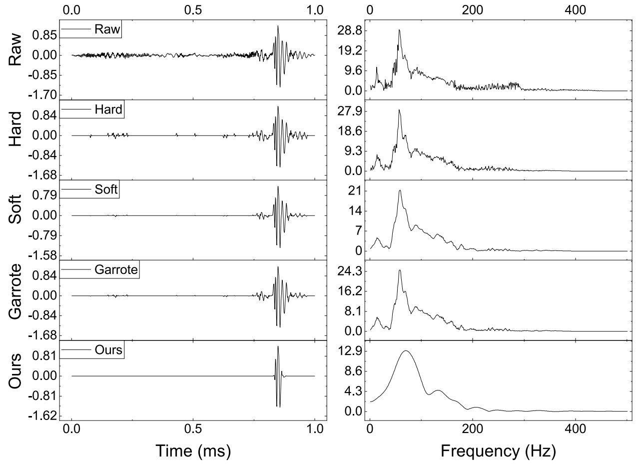

Fig. 4. Detail comparison from real-field data. From top to bottom are noisy raw data and result from hard-, soft-, garrote-thresholding, and the proposed method, respectively.

TABLE II
TIME COST OF EACH METHOD FOR VARIOUS DATA LENGTHS (IN MILLISECOND)

Data Length	*Hard*	*Soft*	*Garrote*	*Ours(CPU)*	*Ours(CUDA)*
100	1.00	0.94	0.96	6.82	6.68
1000	2.45	2.33	2.35	27.19	16.42
5000	2.72	2.56	2.98	74.76	45.51
10000	3.07	2.88	2.89	167.14	95.41
20000	3.77	3.69	3.66	220.26	137.44
50000	6.25	6.27	6.05	402.24	325.76

other approaches mentioned earlier and our method is more applicable and effective for noise reduction in MS data processing.

C. Computation Cost

Finally, we discussed the computational cost of the proposed method via simulations running on a computer with an AMD Ryzen 2700 (3.2 GHz) processor, an NVIDIA GeForce GTX 1060 (5 GB) GPU, and Windows 10 Professional (19043.1081). The runtime environment of simulations is Python (3.8.10) and Pytorch (1.9.0) with the support of CUDA (11.1.102). We recorded the time cost of the four noise reduction methods mentioned in Table I for different data lengths in Table II.

We can see that though our method is more time consuming than conventional wavelet-based noise suppression approaches due to the extra computation cost in clustering, it still takes much less time for noise removal than data production. In general, the proposed method is applicable for practical data processing applications, especially with the help of GPU. All code used in this letter is available in https://github.com/dean-lan/FuzzyWaveletPacketNoiseReduction.git.

IV. CONCLUSION

To improve the performance of noise reduction in MS data processing, we proposed a new WPT-based denoising method via fuzzy partition. It employs the fuzzy coefficient shrinkage and MS event time interval identification. The fuzzy coefficient shrinkage changes the way to suppress noise components to keep more potential signal information. We then utilize interval period identification of the MS event in noisy raw data to remove additional resonance and noise in the nonsignal interval. Results from both synthetic datasets and the real-field dataset show that the proposed noise reduction method is superior to the traditional thresholding function-based methods mentioned. Moreover, though the proposed noise reduction method is more time consuming, it is still applicable and effective for processing time-domain sparse data like MS traces.

ACKNOWLEDGMENT

The authors would like to thank the Sichuan Safety Technology Center, Chengdu, China, and Jiayang Coal Mine, Sichuan Basin, China, for help in experiments. They would also like to thank the valuable comments and suggestions from reviewers and editors.

REFERENCES

[1] S. J. Gibowicz, "Seismicity induced by mining: Recent research," in *Advances in Geophysics*, vol. 51. Amsterdam, The Netherlands: Elsevier, 2009, pp. 1–53.
[2] S. M. Mousavi, C. A. Langston, and S. P. Horton, "Automatic microseismic denoising and onset detection using the synchrosqueezed continuous wavelet transform," *Geophysics*, vol. 81, no. 4, pp. V341–V355, Jul. 2016.
[3] W. Chen and S. Song, "Automatic noise attenuation based on clustering and empirical wavelet transform," *J. Appl. Geophys.*, vol. 159, pp. 649–665, Dec. 2018.
[4] R. N. Bracewell and R. N. Bracewell, *The Fourier Transform and Its Applications*, vol. 31999. New York, NY, USA: McGraw-Hill, 1986.
[5] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.
[6] M. Sifuzzaman, M. R. Islam, and M. Z. Ali, "Application of wavelet transform and its advantages compared to Fourier transform," *J. Phys. Sci.*, vol. 13, pp. 121–134, 2009.
[7] A. Fathi and A. R. Naghsh-Nilchi, "Efficient image denoising method based on a new adaptive wavelet packet thresholding function," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3981–3990, Sep. 2012.
[8] N. A. Golilarz, N. Robert, and J. Addeh, "Survey of image denoising using wavelet transform combined with thresholding functions," *Comput. Res. Prog. Appl. Sci. Eng.*, vol. 3, no. 4, pp. 132–135, 2017.
[9] C. Beale, C. Niezrecki, and M. Inalpolat, "An adaptive wavelet packet denoising algorithm for enhanced active acoustic damage detection from wind turbine blades," *Mech. Syst. Signal Process.*, vol. 142, Aug. 2020, Art. no. 106754.
[10] Y. Liu, T. Hou, and F. Liu, "Improving fuzzy c-means method for unbalanced dataset," *Electron. Lett.*, vol. 51, no. 23, pp. 1880–1882, 2015.
[11] Y. Liu, X. Lu, G. Bei, and Z. Jiang, "Improved wavelet packet denoising algorithm using fuzzy threshold and correlation analysis for chaotic signals," *Trans. Inst. Meas. Control*, vol. 43, no. 6, pp. 1394–1403, 2021.
[12] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, Jun. 2007.
[13] D. Zhu, Y. Li, and C. Zhang, "Automatic time picking for microseismic data based on a fuzzy C-means clustering algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1900–1904, Dec. 2016.
[14] J. Benesty, J. Chen, and Y. Huang, "On the importance of the Pearson correlation coefficient in noise reduction," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 4, pp. 757–765, May 2008.
[15] Y. Ghanbari and M. R. Karami-Mollaei, "A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets," *Speech Commun.*, vol. 48, no. 8, pp. 927–940, Aug. 2006.
[16] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[17] H.-Y. Gao, "Wavelet shrinkage denoising using the non-negative garrote," *J. Comput. Graph. Statist.*, vol. 7, no. 4, pp. 469–488, 1998.
[18] C. Liu, H. Duan, and L. Zhao, "An improved wavelet packet denoising algorithm based on sample entropy for IoT," in *Proc. 15th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2019, pp. 1524–1528.