Article

EEMD and Multiscale PCA-Based Signal Denoising Method and Its Application to Seismic P-Phase Arrival Picking

Kang Peng [1,2,3], **Hongyang Guo** [1] **and Xueyi Shang** [1,*]

1   State Key Laboratory of Coal Mine Disaster Dynamics and Control, School of Resources and Safety Engineering, Chongqing University, Chongqing 400044, China; pengkang@cqu.edu.cn (K.P.); guohongyang@cqu.edu.cn (H.G.)
2   School of Resources and Safety Engineering, Central South University, Changsha 410083, China
3   State Key Laboratory of Coal Resources in Western China, Xi'an University of Science and Technology, Xi'an 710054, China
*   Correspondence: shangxueyi@cqu.edu.cn; Tel.: +86-18711050540

Abstract: Signal denoising is one of the most important issues in signal processing, and various techniques have been proposed to address this issue. A combined method involving wavelet decomposition and multiscale principal component analysis (MSPCA) has been proposed and exhibits a strong signal denoising performance. This technique takes advantage of several signals that have similar noises to conduct denoising; however, noises are usually quite different between signals, and wavelet decomposition has limited adaptive decomposition abilities for complex signals. To address this issue, we propose a signal denoising method based on ensemble empirical mode decomposition (EEMD) and MSPCA. The proposed method can conduct MSPCA-based denoising for a single signal compared with the former MSPCA-based denoising methods. The main steps of the proposed denoising method are as follows: First, EEMD is used for adaptive decomposition of a signal, and the variance contribution rate is selected to remove components with high-frequency noises. Subsequently, the Hankel matrix is constructed on each component to obtain a higher order matrix, and the main score and load vectors of the PCA are adopted to denoise the Hankel matrix. Next, the PCA-denoised component is denoised using soft thresholding. Finally, the stacking of PCA- and soft thresholding-denoised components is treated as the final denoised signal. Synthetic tests demonstrate that the EEMD-MSPCA-based method can provide good signal denoising results and is superior to the low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, EEMD-Hankel–SVD, and wavelet-MSPCA-based denoising methods. Moreover, the proposed method in combination with the AIC picking method shows good prospects for processing microseismic waves.

Keywords: signal denoising; principal component analysis; ensemble empirical mode decomposition; microseismic signal; P-phase arrival picking

1. Introduction

Signals from various practical applications are usually corrupted by unwanted noise, such as background noise, electronic noise, and instrumental noise. Consequently, to improve signal detection, classification, and phase arrival picking, it is necessary to remove unwanted noise. Currently, signal denoising is widely used in the fields of medical signals [1], speech signals [2], mechanical diagnosis [3], structural health monitoring [4–6], and seismic monitoring networks [7–9]; however, this study focused specifically on microseismic (MS) signal denoising.

Spectral filtering is frequently applied to denoise a signal, and a Butterworth-based filter (low-pass filter, high-pass filter, and band-pass filter) is usually adopted [10], which is simple but may filter out a frequency band that contains useful information when the noise and signal share the same frequency band. In addition, the adopted frequency band may vary for different signals; in other words, it cannot denoise a signal adaptively. Therefore,

more effective filtering methods have been proposed for signal denoising, especially based on multiscale features (e.g., wavelet analysis and empirical mode decomposition).

For wavelet analysis-based signal denoising, Beenamol et al. [7] used Shannon and Tsallis entropies to select important wavelet coefficients to denoise a signal; however, noise may still exist in the selected wavelet coefficients. Thus, Tang et al. [8], Mousavi et al. [9], and Chen et al. [11] denoised MS signals based on the combination of wavelet transformation and adaptive thresholding, synchro squeezing wavelet transformation and custom thresholding of single-channel data, and empirical wavelet transformation and automatic noise attenuation thresholding algorithms. However, redundant information in the low-frequency band is not removed by filtering out the high-frequency components and thresholding. Patil [12] combined wavelet transformation and singular value decomposition (SVD) for noise reduction, whereas Bakshi [13] decomposed signals by wavelet transformation and applied principal component analysis (PCA) to denoise the wavelet coefficients of multiple signals containing similar noises and obtained good denoising results. SVD and PCA can efficiently separate useful signals from noise. Wavelet analysis decomposes signals from high- to low-frequency components through wavelet function expansion and translation, which has limited adaptive decomposition abilities for complex signals.

Empirical mode decomposition (EMD), proposed by Huang et al. [14], has better adaptability than wavelet decomposition, and EMD and its improvement have been widely used in signal processing. Mao et al. [15], Dao et al. [16], and Li et al. [17] proposed signal denoising methods based on EMD and thresholding; however, some redundant information could not be denoised by thresholding. Qiao et al. [18] adopted EMD with multiscale eigenvalues of SVD for denoising, which could further reduce noise. However, similar components may exist in different scales for EMD, making it difficult to separate useful components and noise. To overcome this problem, Wu and Huang [19] proposed ensemble empirical mode decomposition (EEMD). On this basis, Jia et al. [20], Han et al. [21], and Li et al. [17] took advantage of permutation entropy, adaptive thresholding, and soft thresholding to denoise signals, attaining a better denoising performance than that based on EMD. Moreover, variational mode decomposition (VMD) has been applied for signal denoising [22], which successfully avoids mode mixing and is robust for both noise and sampling, unlike EMD.

Subsequently, SVD was combined with the Hankel matrix to denoise a signal [23], where the Hankel matrix can expand a signal into a high-dimensional matrix and is very beneficial for feature extraction. Recently, coda wave interferometry (CWI) has been successfully applied to monitor stress distribution and damage detection in engineering [4–6], where wavelet transformation [5], normal band-pass filter, narrow band-pass filter, and double-notch filter [6] were used to extract useful signal features. These techniques provided a good research direction for noise reduction and damage detection. In addition, deep neural networks have been proposed for signal denoising; for example, DeepDenoiser [24] achieves impressive denoising of seismic signals, even when the signal and noise share a common frequency band. However, perfect separation of signal and noise in the time–frequency domain requires recovering two complex spectra for the signal and noise from the complex spectra of the noisy signal.

In this study, we focused on the PCA denoising concept proposed by Bakshi [13], which can enhance useful signal and noise separation, and the highly adaptive EEMD, Hankel matrix, and multiscale PCA (MSPCA) are combined to remove noise from a single signal. The main innovations and contributions of this research are as follows: (1) The EEMD was introduced into MSPCA denoising to achieve improved signal decomposition adaptability compared to wavelet decomposition. (2) The Hankel matrix was used to transform a single EEMD component into high-dimensional data that are suitable for conducting PCA denoising on only one signal. (3) Soft thresholding was selected for further denoising of the PCA-denoised component. (4) The denoising results are superior to those of the low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, EEMD-Hankel–SVD, and wavelet-MSPCA-based denoising methods. The remainder of

this paper is organised as follows. Section 2 introduces the EEMD, Hankel matrix, and MSPCA-based denoising methods. The synthetic and application tests are described in Sections 3 and 4, respectively. Section 5 compares the proposed method with other denoising methods. The research results and prospects are presented in Section 6.

2. Theoretical Fundamentals

2.1. Basis of EEMD

Wu and Huang [19] proposed a noise-assisted EEMD based on EMD by adding white noise to the signals. The steps of the EEMD algorithm are as follows.

(1) By adding a group of white Gaussian noise $\omega_j(t)$ into a noisy signal $x(t)$, a new signal $x'_j(t)$ is obtained as follows:

$$x'_j(t) = x(t) + \omega_j(t) \tag{1}$$

where t is the signal time, j is the jth time of adding white Gaussian noise, and $j = 1, 2, \ldots, M$, where M is the number of times white Gaussian noise is added.

(2) Through EMD of $x'_j(t)$, n intrinsic mode functions (IMFs), $c_{j,i}$, and one residual component $c_{j,n+1}$ are obtained as follows:

$$x'_j(t) = \sum_{i=1}^{n} c_{j,i}(t) + c_{j,n+1} \tag{2}$$

where i is the IMF identity, $i = 1, 2, \ldots, n$, and n is the number of IMFs.

(3) By adding different white Gaussian noises [$\omega_j(t)$] into the noisy signal $x(t)$ and repeating steps (1) and (2), the ith IMF component $c_{j,i}(t)$ and the residual component $c_{j,n+1}$ through the jth decomposition are obtained.

(4) Based on the zero-mean principle of the spectra of white Gaussian noise, the effects of white Gaussian noise as the reference structure of the time-domain distribution are eliminated. The average IMF component $c_i(t)$ and the residual component $c_{j,n+1}$ corresponding to noisy signals can be expressed as follows:

$$\begin{cases} c_i(t) = \dfrac{1}{M} \sum_{j=1}^{M} c_{j,i}(t) \\ c_{n+1}(t) = \dfrac{1}{M} \sum_{j=1}^{M} c_{j,n+1}(t) \end{cases} \tag{3}$$

The relationship between the amplitude, times, and denoising effect of the added white Gaussian noise into the noisy signal $x(t)$ is represented as follows:

$$\varepsilon_i = \frac{\varepsilon}{\sqrt{M}} \tag{4}$$

where ε_i is the denoising effect of the added white Gaussian noise, and ε is the amplitude of the white Gaussian noise.

As shown in Formula (4), when the noise amplitude is fixed, the larger the number of times white Gaussian noise is added, the closer the final decomposition result is to the true values. If the amplitude of the added noise is too small, noise does not affect the selection of the extreme much, thus playing no small role in complementing the scales.

(5) Finally, the noisy signal $x(t)$ can be decomposed as follows:

$$x(t) = \sum_{i=1}^{n} c_i(t) + c_{n+1}(t) \tag{5}$$

2.2. Basis of Hankel Matrix and MSPCA

For a single signal, only one non-zero singular value corresponds to an SVD. Therefore, a singular value cannot be used for denoising. Hence, we constructed a $p \times q$ order Hankel matrix H_i from c_i:

$$H_i = (c_{i,km})_{p \times q} = \begin{bmatrix} c_{i,11} & c_{i,12} & \cdots & c_{i,1q} \\ c_{i,21} & c_{i,22} & \cdots & c_{i,2q} \\ \vdots & \vdots & \vdots & \vdots \\ c_{i,p1} & c_{i,p2} & \cdots & c_{i,pq} \end{bmatrix} \tag{6}$$

where $c_{i,\ km} = c_i\ (k + m - 1)$, and $k + m - 1 = N$, k and m are the row and column identities, respectively, and N is the total sampling number of c_i.

Based on the SVD, matrix H_i can be decomposed as follows: $H_i = U \sum V^T$, ($U \in \mathrm{R}^{p \times p}$, $\sum \in \mathrm{R}^{p \times q}$ and $V \in \mathrm{R}^{q \times q}$). We marked $T = U\sum$ and $P = V$. H can be then treated as the sum of the outer products of the q vectors:

$$H_i = TP^T = \sum_{k=1}^{q} t_k p_k{}^T = t_1 p_1{}^T + t_2 p_2{}^T + \cdots + t_q p_q{}^T \tag{7}$$

where t_k and p_k indicate the score vector and load vector, respectively, $k = 1, 2, \ldots, q$, and T represents the matrix transposition.

The t_k length reflects the H_i coverage in direction p_k: a longer t_k represents greater importance, while a shorter t_k is mainly noise and placed in E. Subsequently, matrix H can be rewritten as

$$H_i = t_1' p_1{}^T + t_2' p_2{}^T + \ldots + t_{q'}' p_{q'}{}^T + E \tag{8}$$

where $t_k'(k = 1, 2, \ldots \ldots, q')$ is the score vector; a larger t_k remains unaltered, whereas a shorter t_k is set as a zero vector.

Considering that the PCA of the matrix, H_i is equivalent to the eigenvector analysis of the covariance matrix $H_i{}^T H_i$, and the load vector of matrix H_i is the same as the eigenvector of $H_i{}^T H_i$. Therefore, we determine t_j' depending on the eigenvalues of the matrix $H_i{}^T H_i$, and the descending ordered eigenvalues are

$$\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_q \tag{9}$$

In general, 85% of the cumulative ratio between an eigenvalue and the total sum of the eigenvalues is selected as the threshold to determine t_k'. The eigenvalues corresponding to the abandoned eigenvectors are set as zero vectors, thus providing the reconstructed Hankel matrix as

$$\hat{H}_i = \hat{T} P^T = t_1' p_1{}^T + t_2' p_2{}^T + \ldots + t_{q'}' p_{q'}{}^T \tag{10}$$

The PCA-denoised component c_i' is obtained from the data of the first row and last column in $\hat{H}_i$.

2.3. EEMD-MSPCA-Based Denoising Method

The first few components obtained by EEMD are mainly high-frequency noises, and noise may exist in the low-frequency components. Therefore, we proposed an EEMD-MSPCA-based seismic signal denoising method (Figure 1), and the specific steps are as follows.
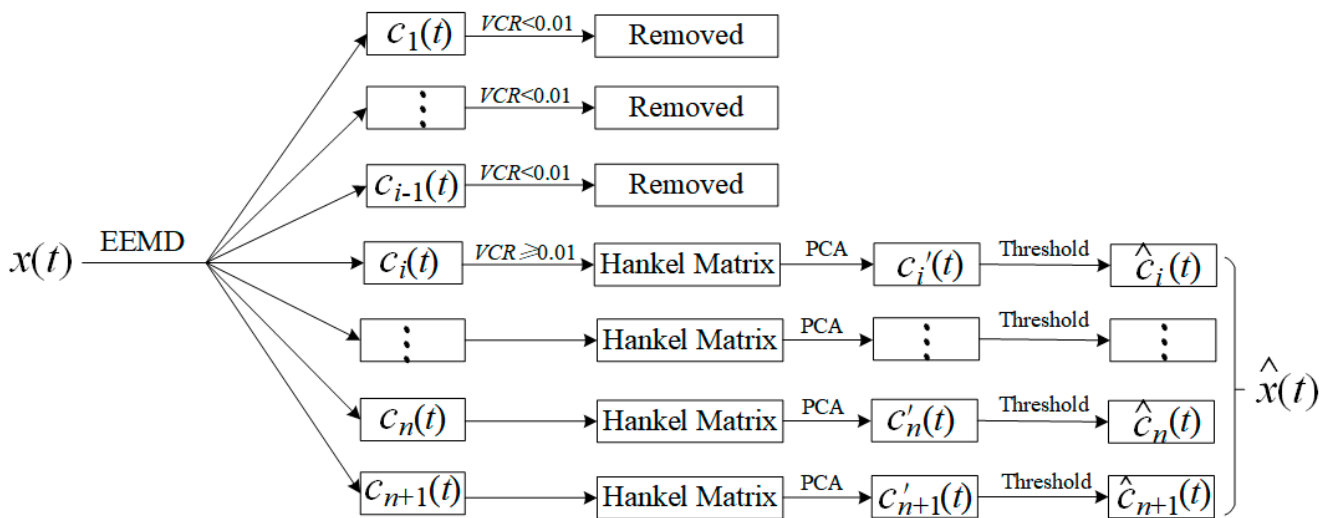
Figure 1. Principle of EEMD-MSPCA-based MS signal denoising.

(1) The noisy signal $x(t)$ is decomposed by EEMD described in Section 2.1, obtaining n IMFs $c_i(t)$ and one residual component $c_{n+1}(t)$.

(2) The variance contribution rates (VCR) of the IMFs are calculated, and the first few high-frequency components with VCRs < 0.01 are removed. The VCR of the ith IMF $c_i(t)$ is defined as

$$VCR_i = \left[\frac{1}{N} \sum_{t=1}^{N} c_i(t)^2 - \left(\frac{1}{N} \sum_{t=1}^{N} c_i(t) \right)^2 \right] \bigg/ \sum_{i=1}^{n+1} \left[\frac{1}{N} \sum_{t=1}^{N} c_i(t)^2 - \left(\frac{1}{N} \sum_{t=1}^{N} c_i(t) \right)^2 \right] \quad (11)$$

(3) The Hankel matrix $\boldsymbol{H}_i$ is constructed for each remaining IMF $c_i(t)$, which is treated as the data matrix for PCA denoising, and its score vector and load vector are calculated through SVD.

(4) Eigenvalues and eigenvectors of the covariance matrix $\boldsymbol{H}_i^{\mathrm{T}} \boldsymbol{H}_i$ are calculated, and the eigenvalues are ranked in descending order. The score vector that corresponds to an eigenvalue with a cumulative eigenvalue percentage larger than 85% is set as a zero vector, and a detailed illustration is shown in Section 3. Subsequently, the Hankel matrix is reconstructed with the new score vector and load vector to obtain the denoised signal $c_i'(t)$.

(5) The signal $c_i'(t)$ is further denoised with soft thresholding and calculated as follows:

$$Th_s = \begin{cases} \mathrm{sign}(c_i')(|c_i'| - T), & |c_i'| > T \\ 0, & |c_i'| \le T \end{cases} \quad (12)$$

where $|\cdot|$, $\mathrm{sign}(\cdot)$, and T indicate the absolute value, sign function, and threshold value, respectively. A fixed threshold $T = \sigma \cdot \sqrt{2 \log(N)}$ (σ denotes data variance) was used in this study.

(6) For the remaining IMFs, denoised components $\hat{c}_i(t)$ can be obtained by repeating steps (3–5), and the final denoised signal $\hat{x}(t)$ is obtained by stacking the denoised components $\hat{c}_i(t)$, as follows:

$$\hat{x}(t) = \hat{c}_i(t) + \cdots + \hat{c}_n(t) + \hat{c}_{n+1}(t) \quad (13)$$

3. Synthetic Tests of EEMD-MSPCA-Based Denoising

Blocks, ECG, Bumps, Doppler shift, and Heavy sine signals with white Gaussian or random noise are usually adopted to test a denoising method, because it is easy to calculate the quantitative parameter signal to noise ratio (SNR) compared with a real signal. Thus, it is convenient for other researchers to reproduce or compare their results with ours. The Blocks signal [$s(t)$ in Figure 2a] is selected to demonstrate the proposed denoising

algorithm. The number of signal sampling points was set as $N = 1024$, and the added noise was expressed as $n(t) = 0.2 \times$ randn(1, 1024), i.e., the noise is 0.2 times the standard normal distribution, and the noisy signal $x(t)$ is shown in Figure 2b.
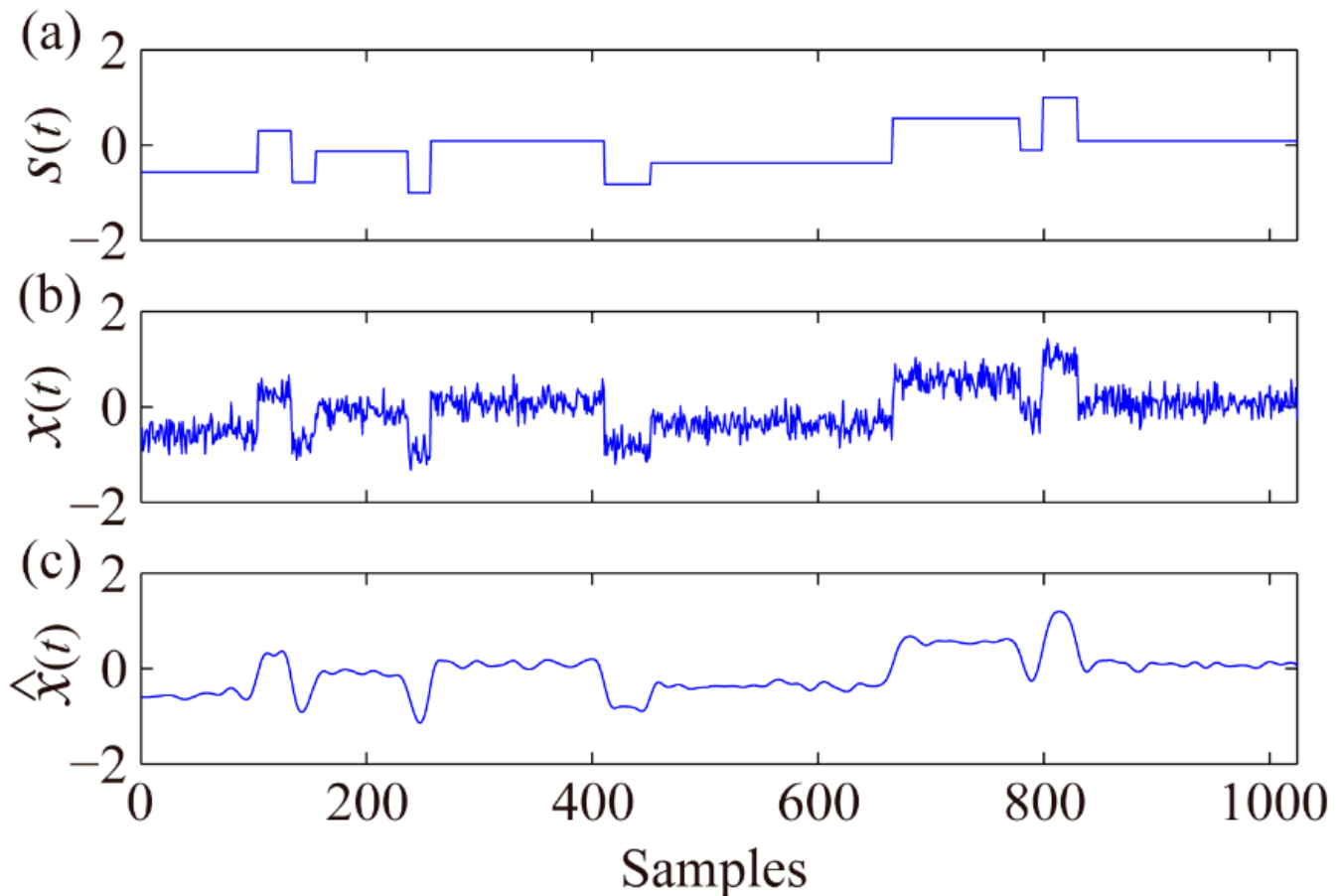

Figure 2. Real (**a**), noisy (**b**), and EEMD-MSPCA-denoised (**c**) Blocks signals.

With EEMD, $x(t)$ is decomposed into nine IMF components, c_1–c_9, and one residual component, c_{10} (Figure 3a). These data indicate that the first few IMF components are mainly high-frequency noises, and the results obtained via Formula (11) show that the VCRs of the first two IMFs are <0.01. Therefore, they are eliminated (marked as zero vectors in Figure 3b) before PCA denoising.

c_3 was selected to illustrate the PCA-based denoising process. First, the Hankel matrix H_3 constructed with c_3 using Formula (6) is decomposed into $H_3 = U \sum V^T$ through SVD, with $T = U \sum$ and $P = V$. Eigenvalues and eigenvectors of the covariance matrix $H_3{}^T H_3$ are calculated, and the eigenvalues are ranked in descending order (red points in Figure 4). The percentages of cumulative eigenvalues are shown in blue points in Figure 4. The score vectors corresponding to the cumulative eigenvalue percentages <85% are used for PCA-based c_3 denoising, and the denoised signal c'_3 is obtained by the reconstructed Hankel matrix $\hat{H}_3$, using Formula (10). Finally, the soft thresholding-denoised $\hat{c}_3$, derived from Formula (12), is shown in Figure 3b. By utilising the same process, the remaining components are denoised to obtain $\hat{c}_4 \sim \hat{c}_{10}$.
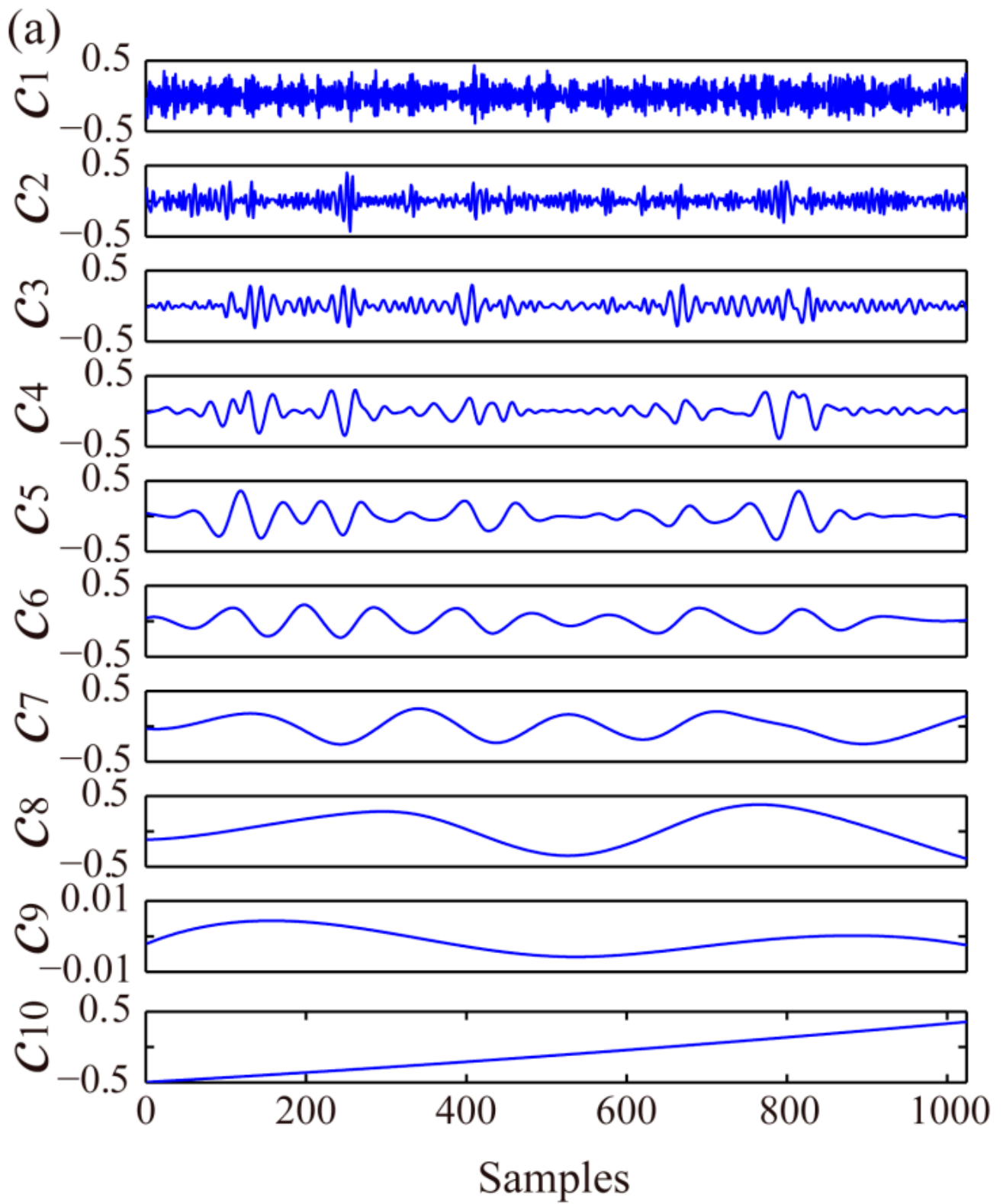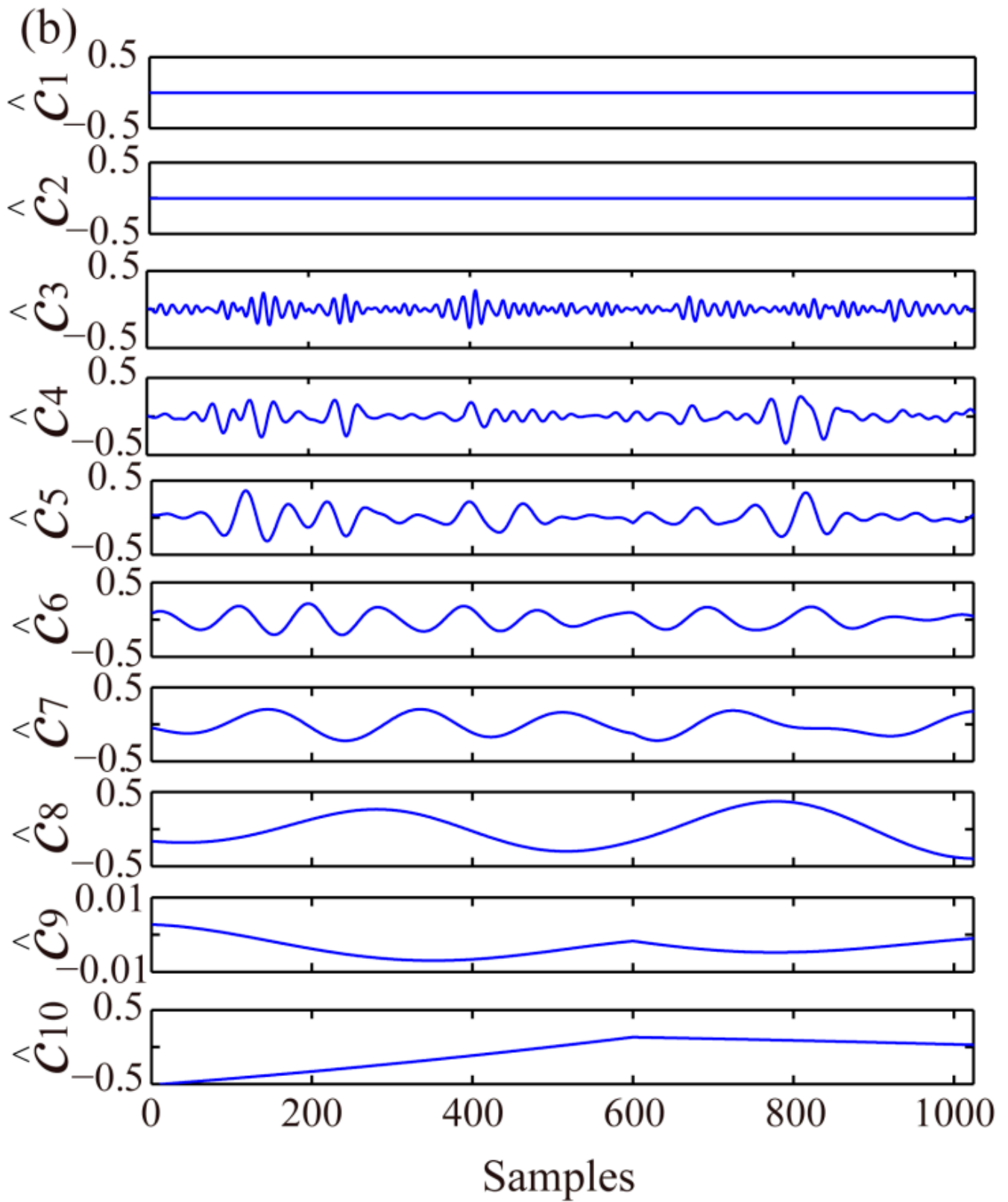
Figure 3. *Cont.*

Figure 3. Original (**a**) and EEMD-MSPCA-denoised (**b**) IMFs for the noisy Blocks signal (Figure 2b).
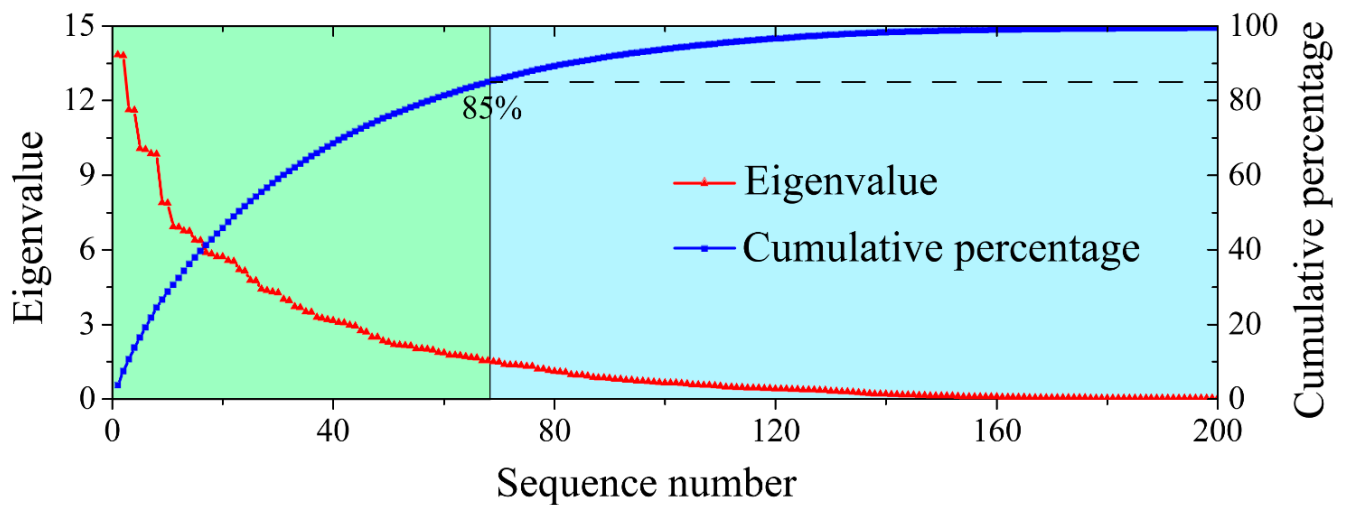
Figure 4. Descending ordered eigenvalues and cumulative eigenvalue percentages of the covariance matrix $H_i^{\mathrm{T}}H_i$. The red points represent the series of eigenvalues arranged in descending order; the blue points indicate the series of cumulative eigenvalue percentages.

As illustrated in Figure 3b, the denoised $\hat{c}_3$ enhanced the local amplitude characteristics of $N = 100$–250 and $N = 300$–450 and suppressed the noise amplitudes of other bands to an extent. Similar denoising results were obtained for the other IMFs, demonstrating the effectiveness of the EEMD-MSPCA-based denoising method. The final denoised signal (Figure 2c) was obtained using $\hat{x}(t) = \sum_{i=3}^{9} \hat{c}_i(t) + \hat{c}_{10}(t)$. SNR was selected to quantitatively evaluate the denoising performance, as defined in Formula (14). The SNRs increased from 6.99 to 12.55, and the denoised signal effectively retained the signal characteristics.

$$SNR = 10\lg[\sum_{i=1}^{N} s_i^2 / \sum_{i=1}^{N} n_i^2] \tag{14}$$

where s_i is the amplitude of the signal without noise and n_i is the amplitude of noise, which is equal to the difference between s_i and the denoised signal.

Furthermore, the ECG, Bumps, Doppler shift, and Heavy sine signals (separately shown top in Figure 5a–d) were selected to test the proposed denoising method. The noise is added using the same steps as those used in the Blocks signal test. The obtained noisy signals and EEMD-MSPCA-denoised signals are shown in the middle and bottom of Figure 5a–d, respectively, and their SNRs are listed in Table 1. Figure 5 depicts the local characteristics of the denoised signals, which are obviously enhanced, and Table 1 displays the increased SNRs of the denoised signals. A comparison of the proposed method with other denoising methods will be shown in the Discussion section.

Table 1. SNRs of the five noisy signals with different denoising methods.

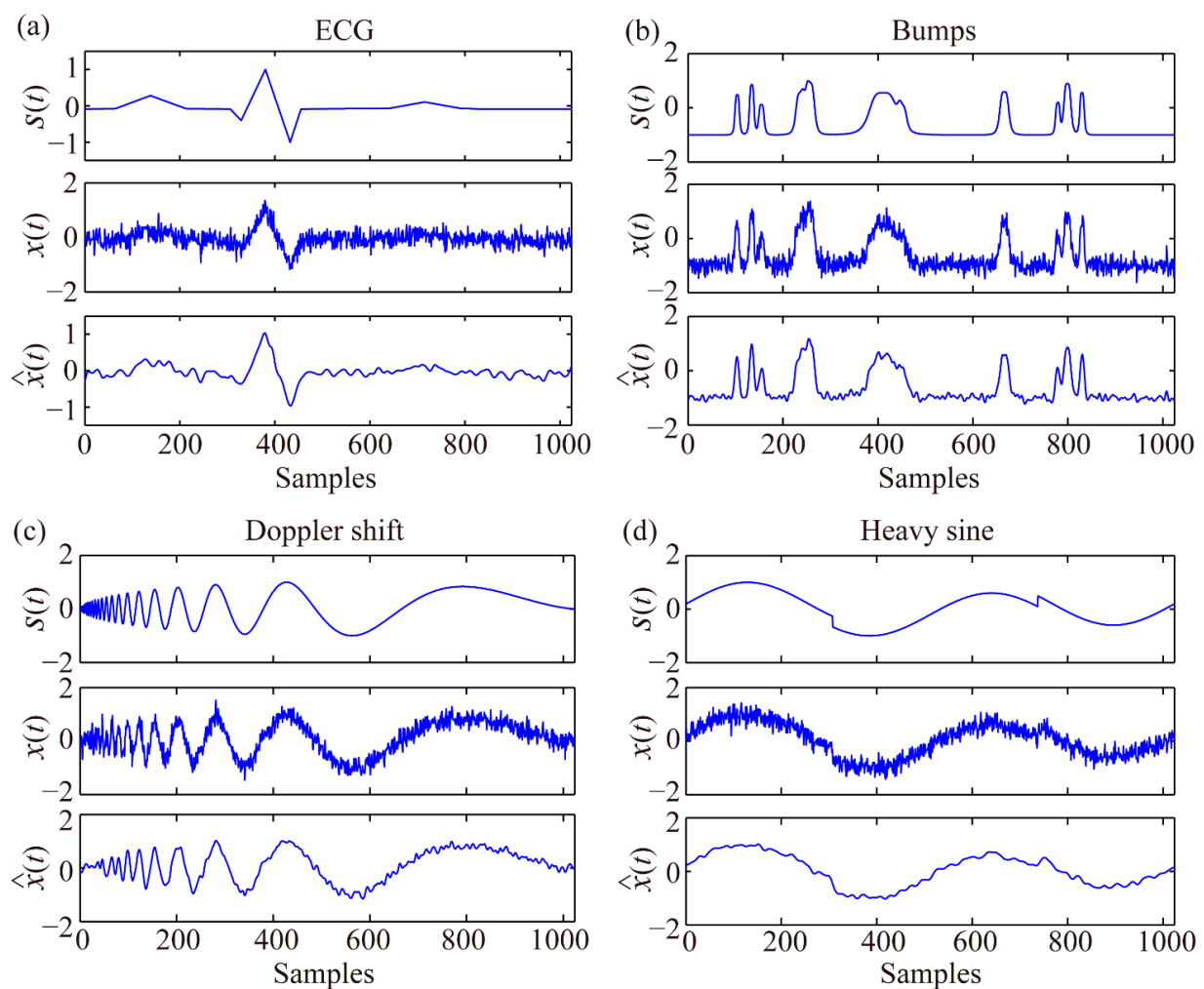| Signal Name	Noisy Signal	Low-Pass Filter	Wavelet Reconstruction	EEMD Reconstruction	Hankel–SVD	EEMD-Hankel–SVD	Wavelet-MSPCA	EEMD-MSPCA
Blocks	6.99	4.67	10.32	10.45	8.69	10.78	11.83	12.55
ECG	0.49	4.46	8.32	8.63	2.93	9.68	9.26	9.43
Bumps	12.55	6.29	17.83	18.38	14.93	19.23	19.88	20.13
Doppler shift	9.32	6.16	15.18	15.59	11.46	16.51	16.58	16.84
Heavy sine	9.51	15.95	18.52	18.67	14.87	19.18	19.04	19.28

Figure 5. Four typical signals without noise (**top**), after adding noise (**middle**), and after EEMD-MSPCA denoising (**bottom**).

4. MS Signal EEMD-MSPCA Denoising and P-Phase Arrival Picking

To further verify the effectiveness of the EEMD-MSPCA-based denoising method, 500 MS signals with a sampling frequency of 6000 Hz were randomly selected from the MS monitoring system in the Yongshaba mine (China) as the application dataset, which usually contains different noises (e.g., rock drilling, loco transportation, electrical noise, mine drawing noise, and blower vibration noise) and acts as a good dataset to test the denoising method. Because the real MS signal is unknown, we used the automatic P-phase arrival picking error instead of the signal SNR to evaluate the denoising performance. The STA/LTA, PAI-K, and AIC picking methods described previously [25] were employed, and the picking error was defined as the difference between the automatic and manual picking points; a smaller picking error represents better denoising performance.

Three typical MS signals before and after denoising and the P-phase arrival picking results are shown in Figure 6. The original waveforms (Figure 6a–c) demonstrate that there is heavy noise in the recorded signals with low SNRs, especially for the third signal (Figure 6c), which includes high-frequency noise, power frequency noise, and spike noise. The EEMD-MSPCA-based denoising method can effectively remove the above-mentioned noises and retain the characteristics of P-phase arrival, which is favourable for P-phase arrival picking. The STA/LTA, kurtosis, and AIC values of the denoised signals are sharper than those without denoising, verifying the improvement of signal SNRs and the effectiveness of the proposed denoising method.
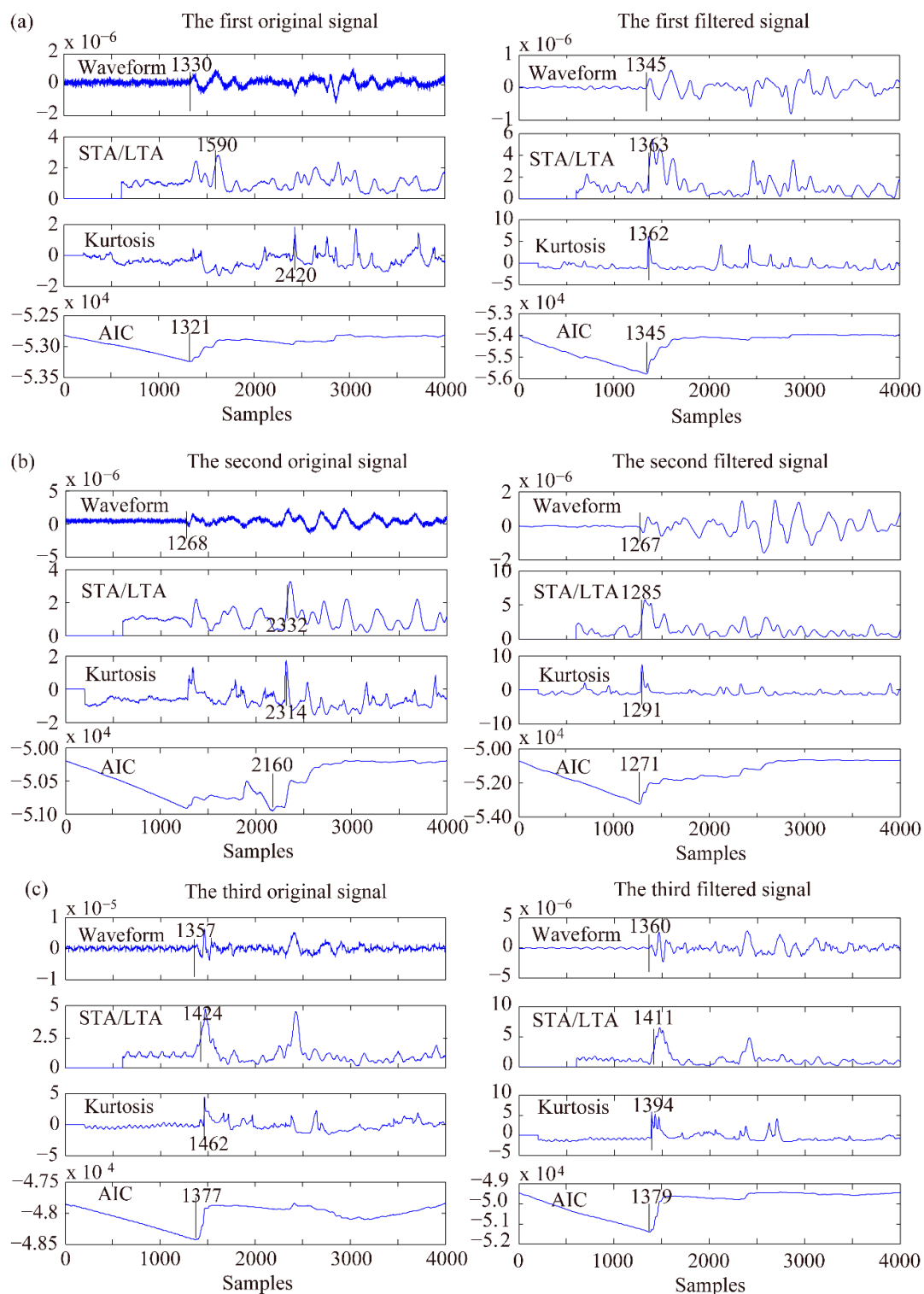
Figure 6. Three typical MS signals before and after EEMD-MSPCA denoising and their STA/LTA, PAI-K, and AIC pickings.

The picking error statistics of the STA/LTA picking, PAI-K picking and AIC picking for the 500 MS signals are shown in Figure 7a–c, respectively. The data indicate that there are more small picking error signals after the EEMD-MSPCA denoising, and the signal number of picking errors that are larger than 30 ms or no pickings is reduced, suggesting an improvement in SNR by the proposed denoising method. Furthermore, the total cost function (TCF) proposed by Li et al. [26] was selected for quantitative comparison. The cost function (CF_i) and TCF are defined in Formulas (15) and (16), respectively, where the

TCF is obtained through the summation of cost functions corresponding to P-phase arrival picking errors, and a small cost is incurred for a small picking error. Therefore, a smaller TCF indicates better picking results. Figure 7 shows that the TCFs of the denoised signals are smaller than those based on the recorded signals, and the TCF based on denoised signals and the AIC method showed the best picking results.

$$CF_i = \begin{cases} 0.0, & \text{accuracy} < 5\,\text{ms} \\ 0.2, & 5\,\text{ms} \leq \text{accuracy} < 10\,\text{ms} \\ 0.4, & 10\,\text{ms} \leq \text{accuracy} < 15\,\text{ms} \\ 0.6, & 15\,\text{ms} \leq \text{accuracy} < 20\,\text{ms} \\ 0.8, & 20\,\text{ms} \leq \text{accuracy} < 25\,\text{ms} \\ 1.0, & 25\,\text{ms} \leq \text{accuracy} < 30\,\text{ms} \\ 1.5, & \text{accuracy} \geq 30\,\text{ms} \end{cases} \tag{15}$$
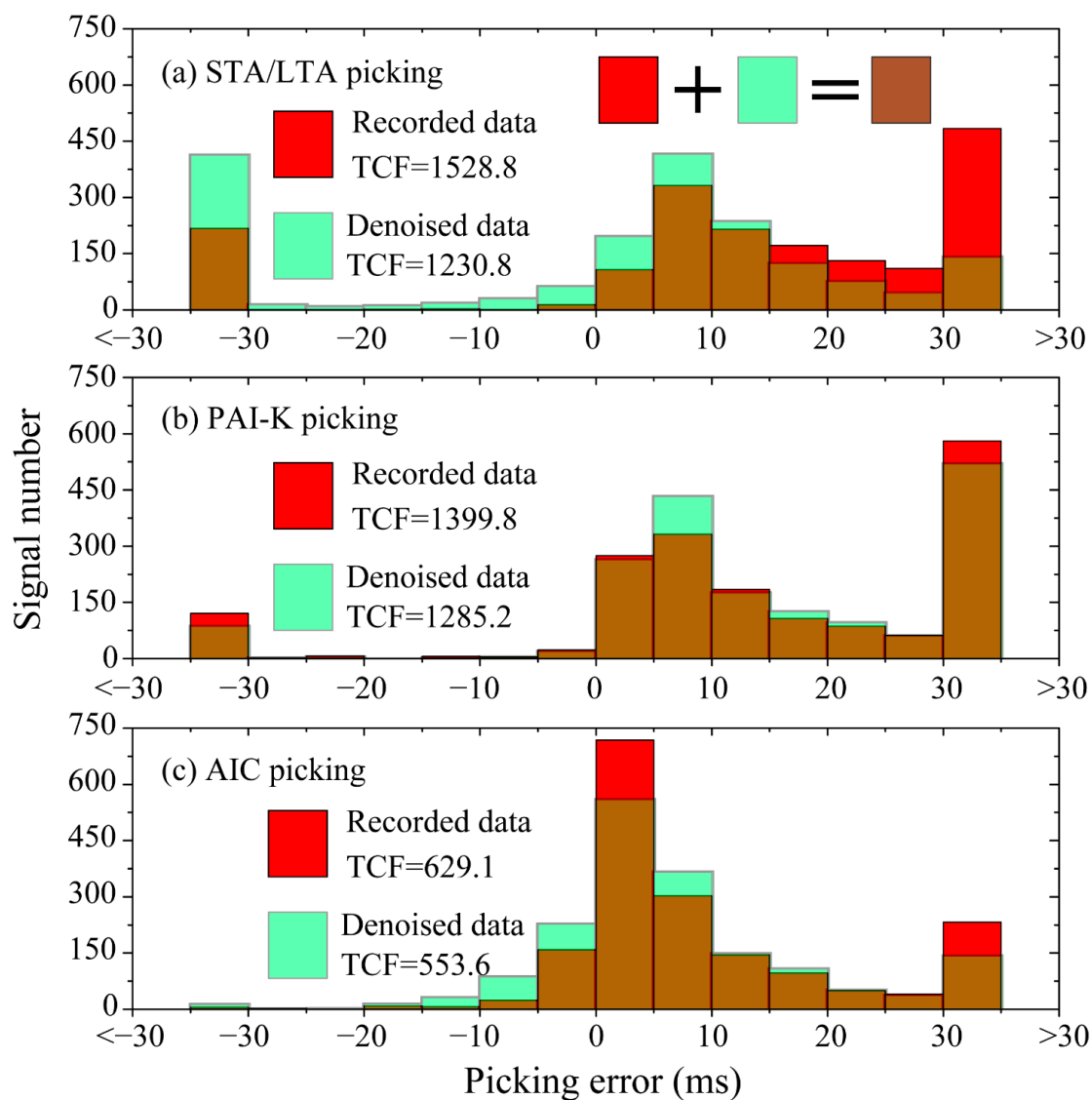
$$TCF = \sum_i CF_i \tag{16}$$

Figure 7. Picking error histogram and TCFs of different methods.

5. Discussion

To better validate the proposed denoising method, low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, EEMD-Hankel–SVD, and wavelet-MSPCA-based denoising methods were selected for comparison. The filter used is a four-order low-pass Butterworth filter with cutoff frequency of 200 Hz, wavelet reconstruction and EEMD reconstruction both remove the first few detail coefficients with VCRs smaller than 0.01, Hankel–SVD was proposed by Jiang et al. [23], EEMD-Hankel–SVD was proposed by Li et al. [27], and the wavelet-MSPCA-based denoising was proposed by Bakshi [13]. Synthetic tests were performed on a Thinkpad laptop with 8 GB RAM and a 1.6 GHz Intel Core i5 Dual Core Processor. The results showed that the computational costs of the low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, EEMD-Hankel–SVD, wavelet-MSPCA, and EEMD-MSPCA-based denoising methods were 0.01, 0.02, 0.25, 0.18, 0.15, 2.96, and 3.03 s, respectively. The EEMD has a relatively large computational cost because of the addition of white Gaussian noise in each EMD step. The EEMD-Hankel–SVD and proposed EEMD-MSPCA-based denoising methods require a much longer computation time than wavelet-MSPCA-based denoising, because the EEMD/wavelet decomposition times are proportional to the row number of the Hankel matrix shown in Formula (6). The denoising performance of the seven methods for the noisy Blocks signal (Figure 2b) is shown in Figure 8, and the SNRs of different denoising methods for the five noisy signals used in the synthetic tests are listed in Table 1.
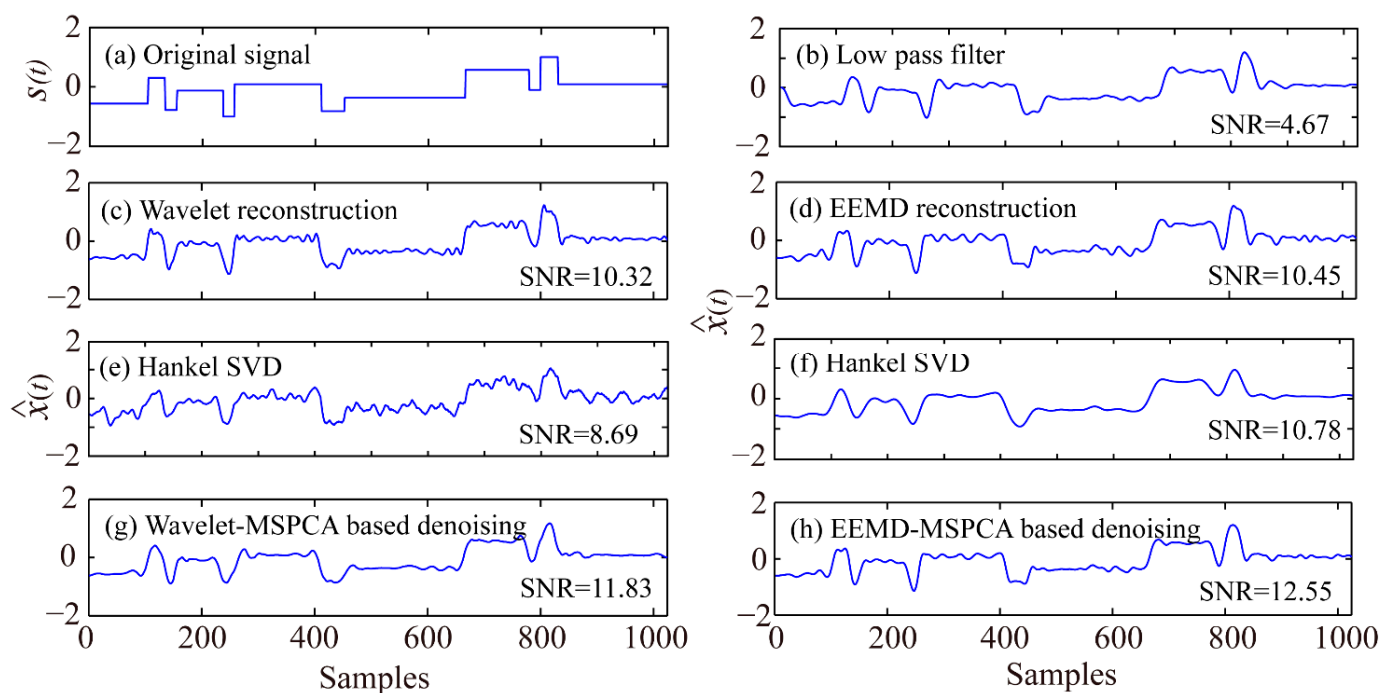

Figure 8. Denoising results of the noisy Blocks signal (Figure 2b).

Figure 8 demonstrates that the low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, and wavelet-MSPCA-based denoising methods are inferior to the EEMD-MSPCA-based denoising method in some respects. Moreover, Table 1 shows that The SNR of the ECG signal improved from 0.49 to >8. Figure 5a shows that many amplitudes in the time series of the ECG signal are zero, causing the $\sum_{i=1}^{N} s_i^2$ and $\sum_{i=1}^{N} n_i^2$ in the SNR definition to be relatively small and large, respectively, after including noises, and a small SNR can then be obtained. Furthermore, there is a coefficient of 10 for the SNR definition. Thus, denoising can significantly improve the SNR in this case. Thus, the low-pass filter improves the waveform quality; however, it can decrease the SNR due

to waveform shift, which is harmful for P-wave arrival picking. All the other denoising methods improved the SNRs, with the EEMD-Hankel–SVD and EEMD-MSPCA obtaining the best denoising performance. The underlying reasons are as follows: Hankel–SVD denoising removes information corresponding to low eigenvalues, while the other denoising methods eliminate high-frequency components in the denoising process, which illustrates the importance of removing high-frequency noise. Subsequently, EEMD reconstruction can reduce the mode mixing effect with a slightly superior denoising result compared to the wavelet reconstruction method. Finally, PCA- and soft thresholding-based denoising can denoise the remaining low-frequency noise [28], as well as retain the signal characteristics.

6. Conclusions

In this study, an EEMD-MSPCA-based signal denoising method was proposed to overcome the difficulty of wavelet-MSPCA in denoising a single signal. This method was verified using synthetic tests and MS signal denoising. The main conclusions are summarised as follows: (1) The proposed method eliminates high-frequency noise through EEMD and the variance contribution rate; subsequently, a Hankel matrix is constructed with each IMF for PCA denoising. PCA and soft thresholding can denoise low-frequency components, as well as retain signal characteristics. (2) EEMD-MSPCA-based denoising performs better than the low-pass filter, wavelet reconstruction, EEMD reconstruction, Hankel–SVD, EEMD-Hankel–SVD, and wavelet-MSPCA-based denoising methods. (3) EEMD-MSPCA-denoised signals effectively improve the P-phase arrival picking accuracy of the STA/LTA, PAI-K, and AIC methods, and the EEMD-MSPCA and AIC picking method achieves a good P-phase arrival picking result, which plays an important role in the MS parameter calculation (e.g., source location and focal mechanism). Further denoising improvements can consider other adaptive decomposition methods (e.g., VMD, local mean decomposition (LMD)), noisy component selection parameters (e.g., signal entropy and coefficient), and combining with other denoising methods.

Author Contributions: K.P. modified the paper and provided some useful suggestions; H.G. wrote the paper; X.S. provided the original idea and wrote the program code. All authors have read and agreed to the published version of the manuscript.

References

1. Chatterjee, S.; Thakur, R.S.; Yadav, R.N.; Gupta, L.; Raghuvanshi, D.K. Review of noise removal techniques in ECG signals. *IET Signal Process.* **2020**, *14*, 569–590. [CrossRef]
2. Alamdari, N.; Azarang, A.; Kehtarnavaz, N. Improving deep speech denoising by noisy2noisy signal mapping. *Appl. Acoust.* **2021**, *172*, 107631. [CrossRef]
3. Yin, X.; Xu, Y.; Sheng, X.; Shen, Y. Signal denoising method using AIC–SVD and its application to micro-vibration in reaction wheels. *Sensors* **2019**, *19*, 5032. [CrossRef]
4. Niederleithinger, E.; Wang, X.; Herbrand, M.; Müller, M. Processing ultrasonic data by coda wave interferometry to monitor load tests of concrete beams. *Sensors* **2018**, *18*, 1971. [CrossRef]
5. Chakraborty, J.; Wang, X.; Stolinski, M. Damage detection in multiple RC structures based on embedded ultrasonic sensors and wavelet transform. *Buildings* **2021**, *11*, 56. [CrossRef]
6. Wang, X.; Chakraborty, J.; Niederleithinger, E. Noise reduction for improvement of ultrasonic monitoring using coda wave interferometry on a real bridge. *J. Nondestruct. Eval.* **2021**, *40*, 14. [CrossRef]

7. Beenamol, M.; Prabavathy, S.; Mohanalin, J. Wavelet based seismic signal de-noising using Shannon and Tsallis entropy. *Comput. Math. Appl.* **2012**, *64*, 3580–3593. [CrossRef]
8. Tang, N.; Zhou, X.; Li, Y.; Zhu, D. Adaptive threshold shearlet transform for surface microseismic data denoising. *J. Appl. Geophys.* **2018**, *153*, 64–74. [CrossRef]
9. Mousavi, S.M.; Langston, C.A.; Horton, S.P. Automatic microseismic denoising and onset detection using the synchro squeezed continuous wavelet transform. *Geophysics* **2016**, *81*, v341–v355. [CrossRef]
10. You, L.Y.; Man, J.H.; Yan, K.Z.; Wang, D.C.; Li, H.L. Combined Fourier-wavelet transforms for studying dynamic response of anisotropic multi-layered flexible pavement with linear-gradual interlayers. *Appl. Math. Model.* **2020**, *81*, 559–581. [CrossRef]
11. Chen, W.; Song, H. Automatic noise attenuation based on clustering and empirical wavelet transform. *J. Appl. Geophys.* **2018**, *159*, 649–665. [CrossRef]
12. Patil, R. Noise reduction using wavelet transform and singular vector decomposition. *Procedia Comput. Sci.* **2015**, *54*, 849–853. [CrossRef]
13. Bakshi, B.R. Multiscale PCA with application to multivariate statistical process monitoring. *AICHE J.* **2004**, *44*, 1596–1610. [CrossRef]
14. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.A.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A* **1998**, *454*, 903–995. [CrossRef]
15. Mao, Q.J.; Fang, X.; Hu, Y.P.; Li, G.N. Chiller sensor fault detection based on empirical mode decomposition threshold denoising and principal component analysis. *Appl. Therm. Eng.* **2018**, *144*, 21–30. [CrossRef]
16. Dao, X.; Gao, M.; Li, C. Adaptive modulation interval filtering algorithm based on empirical mode decomposition. *Measurement* **2019**, *141*, 277–286. [CrossRef]
17. Li, X.; Dong, L.; Li, B.; Lei, Y.; Xu, N. Microseismic signal denoising via empirical mode decomposition, compressed sensing, and soft-thresholding. *Appl. Sci.* **2020**, *10*, 2191. [CrossRef]
18. Qiao, N.; Wang, L.H.; Liu, Q.Y.; Zhai, H.Q. Multi-scale eigenvalues Empirical Mode Decomposition for geomagnetic signal filtering. *Measurement* **2019**, *146*, 885–891. [CrossRef]
19. Wu, Z.H.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [CrossRef]
20. Jia, R.S.; Liang, Y.Q.; Hua, Y.C.; Sun, H.M.; Xia, F.F. Suppressing non-stationary random noise in microseismic data by using ensemble empirical mode decomposition and permutation entropy. *J. Appl. Geophys.* **2016**, *133*, 132–140. [CrossRef]
21. Han, J.J.; Mirko, V.D.B. Microseismic and seismic denoising via ensemble empirical mode decomposition and adaptive threshold-ing. *Geophysics* **2015**, *80*, 1942–2156. [CrossRef]
22. Liu, W.; Duan, Z. Seismic signal denoising using f-X variational mode decomposition. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1313–1317. [CrossRef]
23. Jiang, H.; Chen, J.; Dong, G.; Liu, T.; Chen, G. Study on Hankel Matrix-Based SVD and Its Application in Rolling Element Bearing Fault Diagnosis. *Mech. Syst. Signal Process.* **2015**, *52*, 338–359. [CrossRef]
24. Zhu, W.; Mousavi, S.M.; Beroza, G.C. Seismic signal denoising and decomposition using deep neural networks. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9476–9488. [CrossRef]
25. Li, X.B.; Shang, X.Y.; Wang, Z.W.; Dong, L.J.; Wang, L. Identifying P -phase arrivals with noise: An improved Kurtosis method based on DWT and STA/LTA. *J. Appl. Geophys.* **2016**, *133*, 50–61. [CrossRef]
26. Li, X.B.; Shang, X.Y.; Morales-Esteban, A.; Wang, Z.W. Identifying P-phase arrival of weak events: The Akaike Information Criterion picking application based on the Empirical Mode Decomposition. *Comput. Geosci.* **2017**, *100*, 57–66. [CrossRef]
27. Li, W.; Jiang, X.B.; Chen, H.B.; Jin, Z.P.; Liu, Z.J.; Li, X.W.; Lin, J.X. Denosing method of mine microseismic signal based on EEMD_Hankel_SVD. *J. China Coal Soc.* **2018**, *43*, 1910–1917.
28. Kopsinis, Y.; McLaughlin, S. Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding. *IEEE Trans. Signal Process.* **2009**, *57*, 1351–1362. [CrossRef]