Seismic Data Separation Based on the Equidistant-Spectral Constrained Morphological Component Analysis

Xiaokai Wang, *Member, IEEE*, Chunmeng Cui, Dawei Liu, Pu Liu, Zhensheng Shi, and Wenchao Chen

Abstract— During seismic acquisition, the received seismic data typically comprise many components, such as effective reflections and various interferences. Some components, such as industrial electrical interference and traffic vibrations, manifest as the equidistant narrowband discrete spectra (ENBD-spectra) in the frequency domain. Morphological component analysis (MCA) is widely used for separating different component from complicated seismic data. Therefore, it has been successfully used to extract the narrowband components from seismic data. However, the conventional MCA method overlooks equidistant feature of ENBD-spectra component in seismic data separation. In this study, we propose an improved MCA method that uses the interval between neighboring spectrum peaks as a constraint to separating the data with ENBD-spectra component. Two types of seismic datasets are used to show the proposed MCA's effectiveness. The first type of dataset contains industrial electrical interference, while another type of dataset contains high-speed train (HST)-induced seismic signals. Both synthetic data examples and real data examples show that the proposed method has better performance in separating the seismic data with ENBD-spectra component and keeping the fidelity of separation compared with the conventional MCA method.

Index Terms— Continuous wavelet transform (CWT), discrete Fourier transform (DFT), equidistant narrowband discrete spectra (ENBD-spectra), morphological component analysis (MCA).

I. INTRODUCTION

SEISMIC exploration is one of the important means to detect underground structures, which has been extensively used in oil and gas exploration. However, the received seismic dataset comprises a lot of components, including effective reflections and other interferences which may seriously affect seismic data processing and interpretation. Therefore, seismic data separation is one vital step before seismic data processing and interpretation.

In seismic exploration, the component that can reflect the underground geological structure and rock physical characteristics is commonly referred to as the effective signal. Conversely, the component that obstructs seismic exploration and covers the effective signal energy is treated as the noise [1]. There are a lot of seismic noises in seismic data, which can be divided into noncoherent noise and coherent noise. Noncoherent noise, most of which is known as random noise, lacks coherency between different seismic traces and has no obvious regularity. Typically, the distinction between random noise and effective signals is exploited to suppress random noise in seismic data. For instance, time–frequency transform, low-rank estimation method [2], [3], [4], and their combinations are used to suppress random noise and extract seismic signals with low-rank characteristics. In contrast, coherent noise displays regularity with a fixed range of frequency or apparent velocity, such as ground roll, refracted wave, and multiples. These types of noise had been thoroughly studied and suppressed with many methods [5], [6], [7], [8], [9].

However, there are some common existing interferences, such as industrial electrical interference and traffic vibrations, exhibiting special spectrum characteristics. Industrial electrical interference often annihilates effective signals, seriously affecting seismic data quality. Fig. 1(a) and (b) shows the waveform and amplitude spectrum of one seismic trace affected by strong industrial electrical interference. This interference is mainly centered around 50 Hz and its harmonic components (100, 150 Hz, and so on). In Fig. 1(b), the frequency distance between two neighboring peaks is equal and each frequency peak's band is narrow. Therefore, this kind of spectra is called the equidistant narrowband discrete spectra (ENBD-spectra). High-speed train (HST) is one of the fastest and most convenient ways to travel in China, and its common commercial operation speeds are 250, 300, and 350 km/h. The high-speed railway network exceeded 45 000 km by the end of 2023. The running HST in the extensive railway network would induce seismic waves and affect the geophone near high-speed railway. Fig. 1(c) and (d) shows the waveform and amplitude spectrum of one seismic trace affected by running HST. The amplitude spectrum of the seismic signal generated by this source also exhibits distinct ENBD-spectra characteristic with 3.3-Hz peak interval [10]. Industrial electrical interference and HST-induced signal in Fig. 1 show strong amplitude and
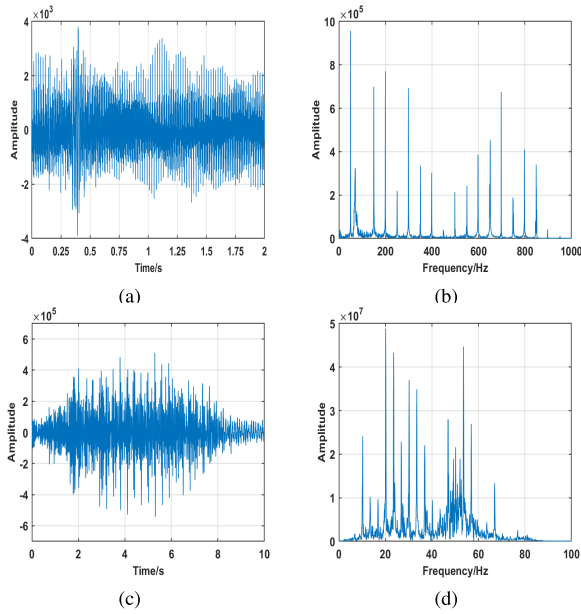
Fig. 1. Typical real seismic data with ENBD-spectra component. (a) Waveform and (b) amplitude spectrum of the seismic data with strong industrial electrical interference. (c) Waveform and (d) amplitude spectrum of the seismic data with strong HST-induced signal.

Fig. 2. Schematic of signal sparse representation.

dominate the amplitude spectrum. ENBD-spectra feature can serve as a crucial foundation for signal separation, which can be reflected in many seismic signals. Therefore, it is valuable to find ways to separate the seismic data with ENBD-spectra component, providing a robust data basis for seismic data interpretation.

To separate the seismic data with ENBD-spectra component, conventional methods usually treat the ENBD-spectra component as the summation of some narrowband components. Therefore, the signal can be separated according to the distribution difference in the signal in some special domains, such as the time domain [11], [12] and the frequency domain [13]. Among them, the cosine function approximation method [14] in the time domain estimates the narrowband component by function approximation. These methods adaptively adjust parameters, such as amplitude, frequency, and phase based on data, to generate the narrowband component and then subtract it from the dataset. Nevertheless, these methods are difficult to extract weak narrowband components because their features are hard to discern. The notch filtering method [15] is a simple approach to suppress the narrowband component in the frequency domain. However, these methods need spectral peak position priors to design multinotches' filter. Therefore, they degrade other components within the notch bands and may distort other components. Moreover, we can obtain better processing results by comprehensively using time–frequency-domain features through short-time Fourier transform [16], S-transform [17], [18], synchrosqueezing transform, and other methods.

Morphological component analysis (MCA) [19], [20], [21] is one classical signal separation method based on sparse representation. It assumes that a signal can be modeled as a linear combination of several components with different morphologies. Each component has a sparse representation
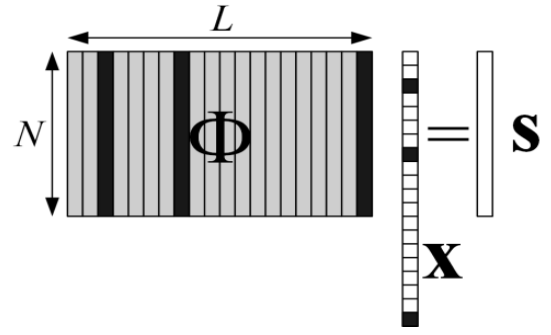
dictionary which is not sparse in representing other components. This method exploits the diverse morphologies of components to separate different components. The conventional mathematical transforms, such as Fourier transform, wavelet transform, curvelet transform [21], [22], and shearlet transform [23], [24], can be selected as sparse representations according the focused component's morphology. However, the commonly used Fourier transform and wavelet transform only consider the ENBD-spectra component as several narrowband components and do not fully use the equal frequency distance between neighboring frequency peaks [25], [26].

Therefore, we fully consider the ENBD-spectra characteristic and use the equal frequency distance between frequency peaks as one constraint for MCA. We propose a method for seismic data separation based on ENBD-spectra feature within the framework of MCA. The sections of this study are organized as follows. Section II outlines the theory of MCA and presents the proposed method. Section III illustrates the effectiveness and superiority of the proposed method through synthetic and real seismic data. We choose two different real seismic datasets to further show the proposed method's effectiveness. One seismic dataset is acquired from a coal mine area with industrial electrical interference, while another one is the seismic dataset containing the moving HST-induced seismic component. Finally, Section IV presents the conclusion.

II. METHOD

A. Morphological Component Analysis

Mallat and Zhang [27] proposed the sparse representation theory in 1993. The sparse representation's schematic is shown in Fig. 2. A signal $s \in \mathbb{R}^N$ can be represented by a linear combination of dictionary elements marked in black. And the coefficient in matrix X participating in the sparse representation is also marked in black, indicating that the atom is nonzero. When the number of nonzero coefficients is small enough, a sparse representation of the signal can be obtained. Besides, the optimization problem to find sparse representation of signals can be expressed as

$$\min_{\Phi, X} \|s - \Phi X\|_2^2 + \lambda \|X\|_0 \tag{1}$$

where X is the representation coefficient under the dictionary Φ. And λ is a parameter to balance sparsity and

constraint conditions. When the error is small enough and the representation coefficient has few nonzeros values, the signal can be represented sparsely.

The sparse representation dictionary generally takes some existing classical transforms. However, the conventional single dictionary is difficult to sparsely represent the complex signal with several components with different morphology. With the development of sparse representation theory, Starck et al. [28] and Bobin et al. [29] proposed the principal theory of MCA, which provides a novel approach for seismic data separation. MCA makes the specific assumption that the signal to be separated can be modeled as a linear combination of different morphological components. And each component is associated with a specific dictionary that is sparse in representing the corresponding component and exhibits nonsparsity in representing other components. This deliberate design ensures that sparse representations are designed according to the specific morphological features of each component, contributing to the effective decomposition and reconstruction process.

In seismic signal processing, it can be assumed that the complex signal, denoted as s, is a combination of two distinct components with different morphological features and random noise, which can be represented as

$$s = s_1 + s_2 + n \tag{2}$$

where s_1 and s_2 are two components with different morphologies. And n represents random noise, which is often disregarded in practical application.

It also assumes that each component $s_i (i = 1, 2)$ has a dictionary Φ_i that is sparse in representing s_i and not sparse in representing other components. Furthermore, these dictionaries should have fast implementations denoted as T_1 and T_2. And R_1 and R_2 represent fast reconstruction forms that are the inverse transformations of dictionary. Therefore, MCA aims to separate these different signal components by solving the optimization problem

$$\{X_1^{\text{opt}}, X_2^{\text{opt}}\} = \underset{\{X_1, X_2\}}{\arg\min} \|X_1\|_1 + \|X_2\|_1$$
$$\text{s.t. } s = R_1 X_1 + R_2 X_2 \tag{3}$$

where X_1 and X_2 are the sparse representation coefficient vectors for two components, respectively.

The block coordinate relaxation (BCR) algorithm proposed by Bruce and Sardy [30] in 1998 aims to solve the optimization problem in (3). The core idea of the BCR algorithm is to set a reasonable threshold strategy and update the sparse coefficient alternately at each iteration until the iterative termination condition is reached. The basic process of the BCR algorithm [31] is as follows.

1) Initialization: The first step is to set the maximum iterations $k_{\max}$, the initial threshold λ_{initial}, and the termination threshold λ_{final}. The choice of maximum iterations should be considered based on factors such as problem complexity, computational resources, and algorithm convergence properties. The initial threshold needs to make the dictionary atoms adequately match the component in the initial stage of the iteration. This helps ensure that the representation coefficient of the signal shows significantly more sparsity during the

early stage of the iteration, thus leading the algorithm to converge more efficiently. The termination threshold ensures that the iterative operation can be stopped at a reasonable threshold. In addition, The initial sparsity coefficients are set to $X_1^0 = 0$ and $X_2^0 = 0$.

2) Iteration: As iteration number k increases, the sparse coefficient is updated alternately as

$$X_1^k = H_{\lambda_1}^k \left(T_1 \left(s - R_2 X_2^{k-1} \right) \right)$$
$$X_2^k = H_{\lambda_2}^k \left(T_2 \left(s - R_1 X_1^{k-1} \right) \right) \tag{4}$$

where $H_{\lambda_1}^k$ and $H_{\lambda_2}^k$ are the threshold processing using the threshold parameters λ_1^k and λ_2^k. The thresholds λ_1^k and λ_2^k can be updated linearly or exponentially with the iteration steps k [26]. The linear threshold in the kth iteration can be expressed as

$$\lambda^k = \lambda_{\text{initial}} - k \left(\frac{\lambda_{\text{initial}} - \lambda_{\text{final}}}{k_{\max} - 1} \right) \tag{5}$$

and the exponential iteration threshold can be expressed as

$$\lambda^k = \lambda_{\text{initial}} \left(\left(\frac{\lambda_{\text{initial}}}{\lambda_{\text{final}}} \right)^{\frac{1}{k_{\max} - 1}} \right)^{-k}. \tag{6}$$

3) Output Results: The iteration terminates when the iteration number reaches the maximum iterations or the result of further iterations is very similar to the former. The optimal coefficients X_1^{opt} and X_2^{opt} can be obtained after stopping the iteration. And the different component can be reconstructed as follows:

$$s_1 = R_1 X_1^{\text{opt}}$$
$$s_2 = R_2 X_2^{\text{opt}} \tag{7}$$

where R_1 and R_2 represent the inverse transformations of dictionary. Therefore, the BCR algorithm offers a solution for separating the two components from the complex signal using the optimal sparse representation coefficients.

B. Equidistant-Spectral Constrained MCA

1) Peak Interval Estimation for ENBD-Spectra: In the conventional MCA method, the ENBD-spectra component is usually modeled as the superposition of a series of narrowband components, and conventional approaches often use this modeling to select sparse representation dictionaries [32]. Therefore, the common overcomplete dictionary construction for the ENBD-spectra component does not fully consider the equal interval between two neighboring peaks. The frequency spacing between adjacent spectral peaks is a crucial characteristic of this component. A more suitable sparse dictionary for ENBD-spectra component should make full use of the equal interval between two neighboring peaks. Therefore, one important step is to precisely estimate of the peak interval between two neighboring peaks. Fig. 1(b) and (d) shows that a periodic feature spectrum in the frequency domain dominates the amplitude spectrum, and therefore, the cross-correlation can be used to estimate the spectral interval between two neighboring peaks. Here, a series of equidistant spectra with variable intervals are generated. The cross-correlations between the
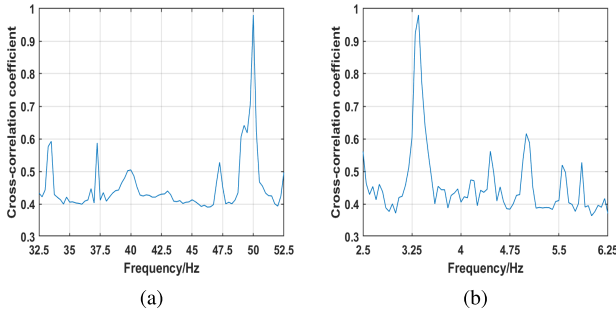
Fig. 3. Relationship between cross-correlation value and frequency interval. (a) Seismic data with industrial electrical interference. (b) Seismic data with HST-induced seismic signal.

Fig. 4. Process of constructing DFT with interval constraints.

real ENBD-spectrum and these preset equidistant spectra are calculated, and thus, the curve of cross-correlation value and interval can be obtained. Finally, the interval corresponding to the maximum cross-correlation value is considered as an estimation of the peak interval between two neighboring peaks. The results of peak interval estimation shown in Fig. 3 are obtained by processing the seismic data in Fig. 1(b) and (d). Fig. 3(a) illustrates the spectral peak estimation results for the seismic data with industrial electrical interference, where the interval is estimated to be 50 Hz which is equal to the standard frequency of China power grid. Meanwhile, Fig. 3(b) presents spectral peak estimation results for the seismic data with HST-induced seismic signal. The spectral interval is estimated to be 3.33 Hz which is equal to the ratio between trace speed (83.33 m/s) and carriage length (25 m).

2) Sparse Dictionary Selection: If the seismic data s contain industrial electrical interference or HST-induced seismic signal, it can be modeled as a linear summation of seismic signal s_1 and ENBD-spectra component s_2. According to the basic assumption, one key step is to find sparse dictionary for seismic signal and ENBD-spectra component. The seismic signal s_1 can be expressed as the convolution of the reflection coefficients and the source wavelet. Therefore, the continuous wavelet transform is usually chosen as the sparse dictionary for component s_1 [33], [34], which can be defined as

$$\text{WT}_x(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t - \tau}{a}\right)dt \quad (8)$$

$$x(t) = \frac{1}{C_\Psi} \int_0^\infty a^{-\frac{5}{2}} da \int_{-\infty}^\infty \text{WT}_x(a, \tau)\psi\left(\frac{t - \tau}{a}\right)d\tau \quad (9)$$

where $\text{WT}_x(a, \tau)$ is the wavelet transform coefficient of the signal $x(t)$. And $\psi(t)$ is the wavelet function with the scale factor a and the time shift factor τ, whose Fourier transform form is $\Psi(w)$. Besides, $C_\Psi = \int_0^\infty (|\Psi(aw)|^2/a)da < \infty$ is the admissibility condition of wavelet function.

For the ENBD-spectra component, such as industrial electrical interference and HST-induced signal in Fig. 1, each frequency peak in the frequency domain is very similar to amplitude spectrum of single-frequency signal, which is very similar to the basic atom of Fourier transform. Thus, conventional methods [35] usually choose the discrete Fourier transform (DFT) or discrete cosine transform (DCT) as the sparse dictionary for the single-frequency component or
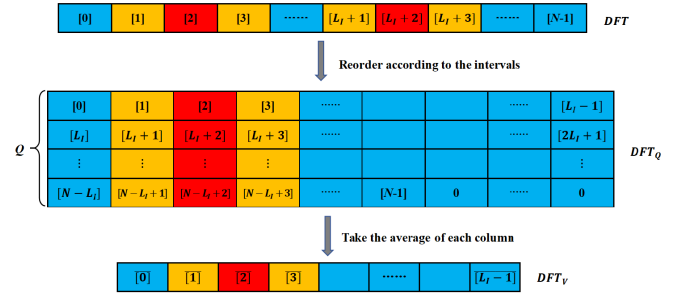
narrowband component. Among them, the DFT of signal $x[n]$ is defined as

$$\text{DFT}_x[m] = \sum_{n=0}^{N-1} x[n]W_N^{nm}, \quad m = 0, 1, \ldots, N-1 \quad (10)$$

Meanwhile, the inverse DFT that is used for signal reconstruction can be defined as

$$x[n] = \frac{1}{N} \sum_{m=0}^{N-1} \text{DFT}_x[m]W_N^{-nm}, \quad n = 0, 1, \ldots, N-1 \quad (11)$$

where $\text{DFT}_x[m]$ is the DFT coefficient of signal $x[n]$. And the exponential term $W_N^{nm} = e^{-j(2\pi/N)nm}$ serves as the basis function of DFT in this formula. Different values of k facilitate the extraction of signal components at distinct frequencies.

3) Equidistant-Spectral Constraint and Threshold Adjustment: However, DFT only regards ENBD-spectra component as the summation of some narrowband components, but ignores the obvious feature that the interval between two neighboring frequency peaks is equal. Therefore, it does not fully use the unique ENBD-spectra component's feature. To make DFT adapt to ENBD-spectra component, we reshape the N-points DFT coefficient vector into a matrix (DFT_Q) with Q rows and L_I columns (as shown in Fig. 4). L_I is the peak interval and Q can be calculated by the roundup of N/L_I. This matrix can be averaged along column direction to form an L_I-points vector DFT_V.

The sparse coefficients of ENBD-spectra component are more concentrated in the corresponding positions of the basic frequency and its multiple frequency (as shown the red and orange regions in Fig. 4). Consequently, we can reduce the iterative threshold λ_2^k for screening these DFT coefficients, which can enhance the ability to extract the ENBD-spectra component. To achieve this, we set a corresponding threshold strategy for each DFT coefficient as follows:

$$\tilde{\lambda}_2^k[m] = \begin{cases} \lambda_2^k/M, & \text{DFT}_V[i] \geq \lambda_2^k \text{ and } m = i + q * L_I \\ \lambda_2^k * M, & \text{else} \end{cases} \quad (12)$$

where $\tilde{\lambda}_2^k$ is the adjusted iteration threshold and M is the adjustment coefficient that determines the extent to reduce or amplify the threshold. DFT_V is the reshaped L_I-points coefficient vector. Besides, i represents the position of sparse coefficients. And q is the index used to indicate the frequency of harmonic multiples, which ranges from 0 to $Q-1$. Besides, M is an important parameter that affects the quality and
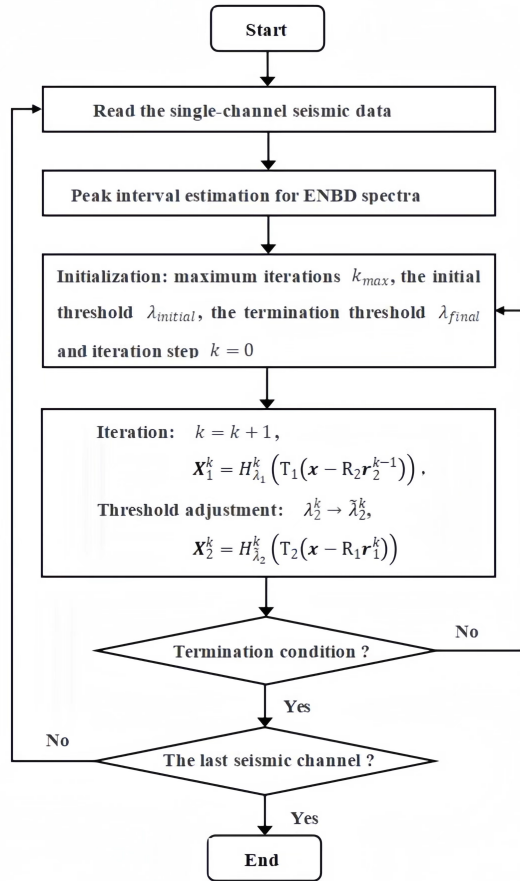
Fig. 5. Process of separating seismic signal based on the equidistant-spectral constrained morphological component.

Fig. 6. Synthetic seismic data of single trace. (a) Waveform and (b) amplitude spectrum.

accuracy of the final result. In real situation, a suboptimal value can be obtained by selecting parameter M based on a similar synthetic signal. Among them, we set the value range of this parameter according to the sampling length and spectral interval of the data. In this range, the signal-to-noise ratio (SNR) is used as the standard to select the most suitable M. This approach reduces the threshold for separating seismic signals with the ENBD-spectra component, corresponding to the red and orange regions in Fig. 4. And it also increases the threshold for separating other components which corresponds to the blue regions in Fig. 4. It can improve the ability of signal discrimination and separation accuracy. So the process of separating seismic data based on the equidistant-spectral constrained morphological component can be shown as Fig. 5.

III. DATA EXAMPLES

Among the numerous seismic data containing ENBD-spectra component, we chosen two kinds of seismic dataset, one contains industrial electrical interference and another contains HST-induced seismic signal to validate the proposed method's effectiveness.

A. Industrial Electrical Interference Suppression Examples

1) Synthetic Data Example: The synthetic data in Fig. 6 are composed of effective seismic signal, industrial electrical interference, and random noise. The effective seismic
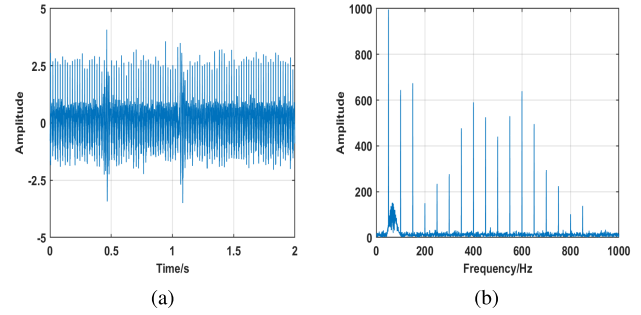
signal is generated by two Morlet wavelets with dominant frequencies of 61 and 73 Hz. The simulated industrial electrical interference is generated by superposition of multiple single-frequency signals whose frequencies are 50 Hz and its multiple frequency. The simulated industrial electrical interference in Fig. 6 shows typical ENBD-spectra feature in the frequency domain. The seismic signal and interference's frequency band overlap with each other, which makes it difficult to separate different components.

Here, we use notch filtering [36] and conventional MCA method [26] as the compared method. The separation results by the notch filtering method, the conventional MCA method, and our proposed method are shown in Figs. 7 and 8. It is obvious that all the three methods can effectively suppress industrial electrical interference in seismic data. Furthermore, as shown in Fig. 7(c)–(f), two MCA methods can suppress random noise at the same time. However, the results in Fig. 7(d) show that conventional MCA can suppress industrial electrical interference while may damage some effective signals, as pointed with red arrow. Our proposed method significantly reduces this damage to effective signals, which is shown in Fig. 7(f).

For the extracted signals in Fig. 8, it is evident that the notch filtering method residues a lot of random noise and causes significant damage to the effective signal within the notch band, particularly at the position indicated by three dotted rectangles in Fig. 8(b). As shown in Fig. 8(c)–(f), both the conventional MCA method and our method can remove random noise and recover two wavelets. However, the conventional MCA method still exhibits some damages to two wavelets in Fig. 8(c) and (d), while our method can reduce such damage by introducing ENBD-spectra feature as one constraint [Fig. 8(e) and (f)].

The comparison of SNR in Table I further shows the performance of three methods. Two MCA methods have obvious higher SNR than notch filtering, while the proposed method can have 3-dB improvement over the conventional MCA method. These comparisons in waveform, amplitude, and SNR show the proposed method's effectiveness in separating the seismic data with industrial electrical interference.

2) Real Seismic Data: To further evaluate the proposed method, we use one 2-D real seismic dataset acquired from a coal mining area in China and show this dataset in Fig. 9. The effective seismic signal is covered by strong industrial electrical interference. The average amplitude spectrum in
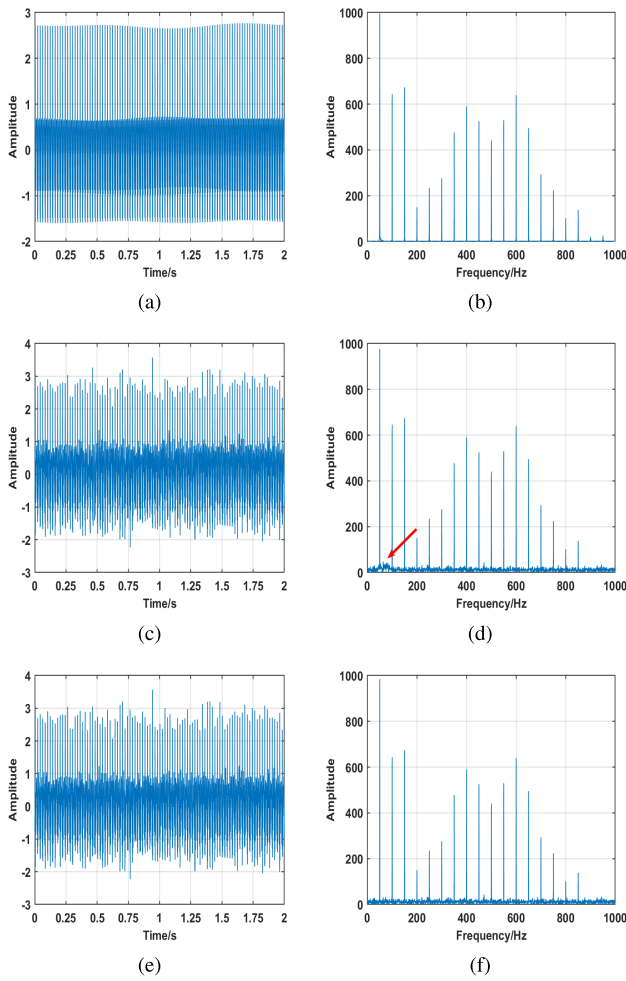
Fig. 7. Suppressed interference of single trace. (a) Waveform and (b) amplitude spectrum of notch filtering method. (c) Waveform and (d) amplitude spectrum of the conventional MCA method. (e) Waveform and (f) amplitude spectrum of our proposed method.
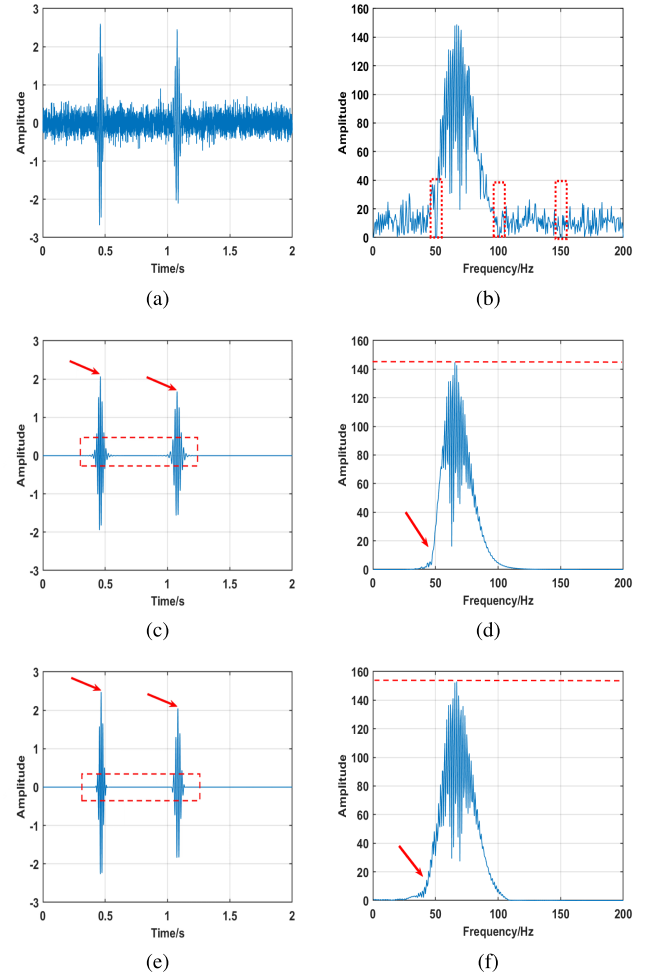

Fig. 8. Extracted effective signal of single trace. (a) Waveform and (b) amplitude spectrum of notch filtering method. (c) Waveform and (d) amplitude spectrum of the conventional MCA method. (e) Waveform and (f) amplitude spectrum of our proposed method.

Fig. 9(b) along with the amplitude spectrum of the 23rd seismic trace shown in Fig. 9(c) exhibit typical ENBD-spectra feature, and the effective seismic signal's amplitude spectrum is significantly weaker than that of the interference. Therefore, recovering the weak seismic signal is one challenging work.

Notch filtering, the conventional MCA method, and the proposed method are used to separate this dataset and further suppress strong industrial electrical interference. The separation results of three methods are shown in Fig. 10. Three methods can suppress the major industrial electrical interference and recover the effective seismic signals. However, notch filtering remains a considerable amount of interference, as shown from Fig. 10(a). The conventional MCA method can gain better separation performance. As shown in Fig. 10(b), the conventional MCA method can recover the effective seismic signal clearly, but has some visible interference leakages near the up bound. The proposed method can provide a good recovery result for effective seismic signal and does not have interference leakage near the bound [Fig. 10(c)]. In addition, the 53rd trace not only contains industrial electrical interference signals (50 Hz and its frequency doubled harmonic components) but also contains a component concentrated

TABLE I
COMPARISON OF SNR OF INDUSTRIAL ELECTRICAL INTERFERENCE

Method	SNR	Method	SNR
Unprocessed data	-8.475 dB	The conventional MCA	10.682 dB
The notch filtering	4.765 dB	Our proposed method	13.712 dB

at 84.5 Hz. The conventional MCA method extracted all the narrowband components and does not discriminate industrial electrical interference and other components. The proposed method can improve the signal separation accuracy through the interval constraint, which also leads to some difference in Fig. 10(b) and (c).

We calculate the separation results' average amplitude spectra and shown them in Fig. 11. The notch filtering method can effectively remove the interference within the filter band, but it causes serious damage within the effective signal's frequency band [Fig. 11(b)]. Moreover, the conventional MCA method can overcome the shortcomings of notch filtering and gain better separation performance at the frequency band near 50 Hz and its multiples [Fig. 11(d)]. However, the periodic small peaks, which are pointed by some red arrows in Fig. 11(d), indicate that the conventional method has a
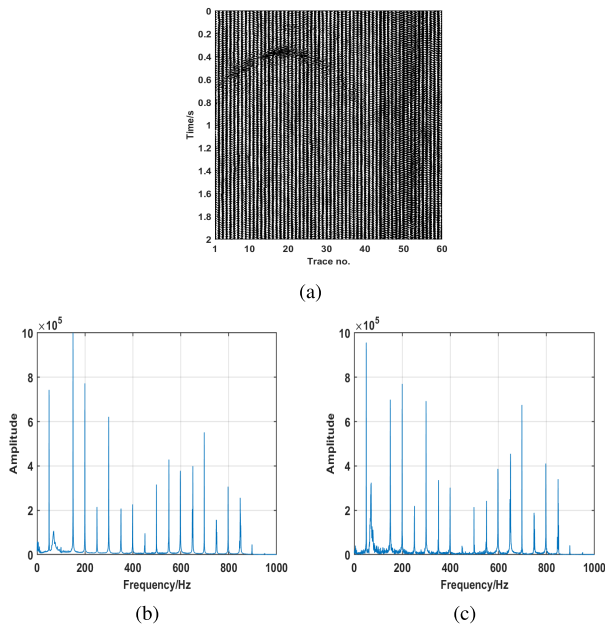
(a)

(b)

(c)

Fig. 9. Real seismic data with industrial electrical interference. (a) Seismic profiles. (b) Average amplitude spectrum. (c) Amplitude spectrum of seismic data of 23rd trace.
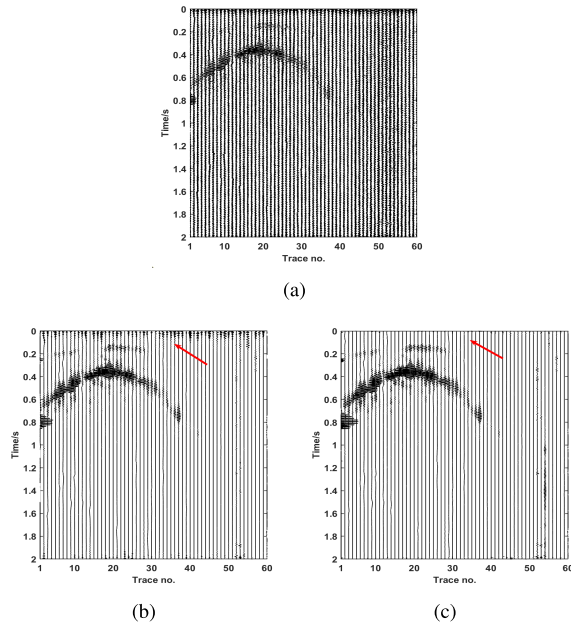

(a)

(b)

(c)

Fig. 10. Results of seismic profiles. (a) Notch filtering method. (b) Conventional MCA method. (c) Our proposed method.

little interference leakage to signal part. Compared with notch filtering and conventional MCA methods, the proposed method can achieve more thorough suppression of interference and keep the fidelity of the effective signal [Fig. 11(f)]. Taking the 23rd seismic data [Fig. 10(c)] as an example, the effective signal waveform and amplitude spectrum results shown in Fig. 12 can also be used to verify the performance of the proposed method.

B. HST-Induced Seismic Signal Extraction Examples

1) Synthetic Data Example: HST-induced seismic signal is mainly generated by the contact between the wheels and
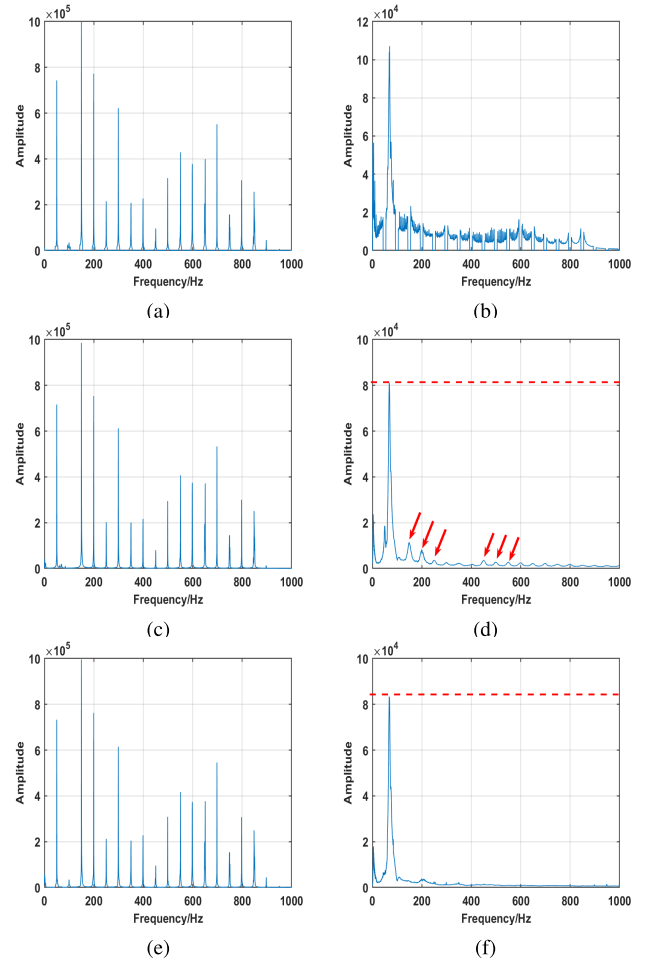

Fig. 11. Comparison of average amplitude spectrum. (a) Interference and (b) effective signals of notch filtering method. (c) Interference and (d) effective signals of the conventional MCA method. (e) Interference and (f) effective signals of our proposed method.

the track during the operation of high-speed railway trains (HSTs) [37]. Fig. 13 shows one typical structure for China's HST. It has M carriages and each carriage has four pairs of wheels. The length of each carriage l is generally 25 m. The interval between the first and the second pair of wheels d_1 is 2.5 m, while the interval between the first and third pairs of wheels d_2 is 14.5 m [38], [39]. When an HST travels at a constant speed, the carriages of the HST would periodically press the track through wheels. Therefore, the HST-induced seismic signal also exhibits periodic features, and thus it has ENBD-spectra feature in the frequency domain [40], [41]. HST-induced seismic signal has been used in train speed estimation, underground structure inversion, etc. [42], [43]. In this example, we take HST-induced seismic signal as the effective component to be separated.

The peak interval of the HST-induced signal's ENBD-spectra is equal and determined by the ratio between HST speed and carriage length [32]. If the HST speed is set to 100 m/s, the periodicity caused by the same carriage structure is 0.25 s, and thus the peak interval of ENBD-spectrum is 4 Hz.

To show the proposed method in extracting the HST-induced signal, a 1-D synthetic dataset is generated with one effective

(a)   (b)

Fig. 14. Synthetic seismic data of single trace. (a) Waveform and (b) amplitude spectrum.

(a)   (b)

(c)   (d)

Fig. 15. Extracted effective signal of single trace. (a) Waveform and (b) amplitude spectrum of the conventional MCA method. (c) Waveform and (d) amplitude spectrum of our proposed method.
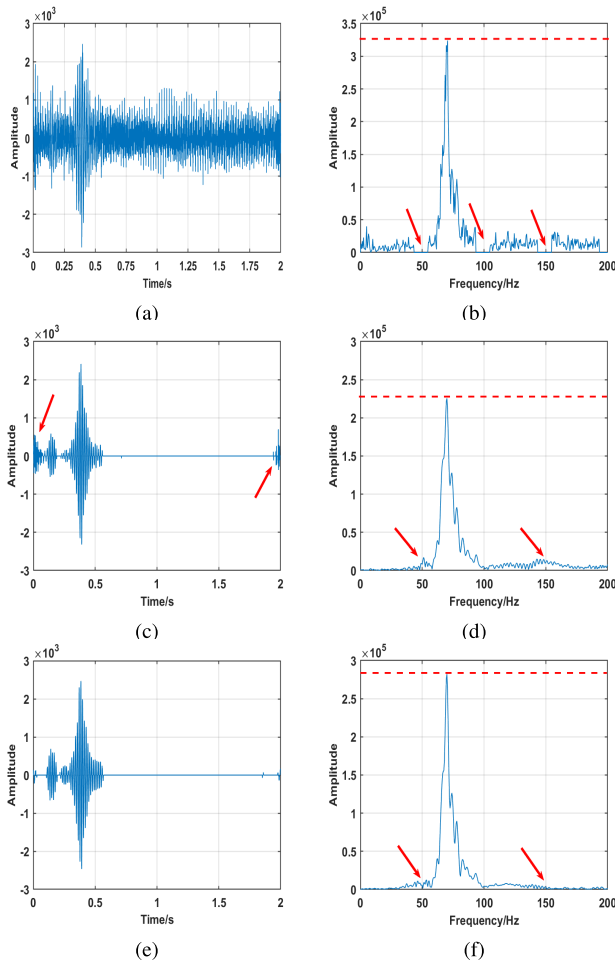

(a)   (b)

(c)   (d)

(e)   (f)

Fig. 12. Extracted effective signal of the 23rd trace. (a) Waveform and (b) amplitude spectrum of the notch filtering method. (c) Waveform and (d) amplitude spectrum of the conventional MCA method. (e) Waveform and (f) amplitude spectrum of our proposed method.
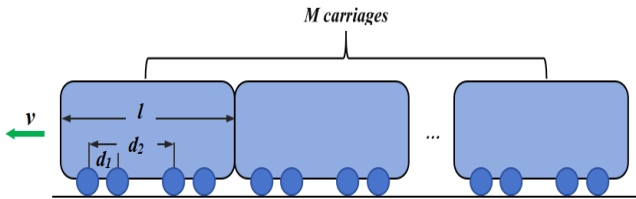

Fig. 13. Schematic of a uniformly moving HST source.

HST-induced signal with 4-Hz peak interval, some Ricker wavelets (their dominate frequencies are 7, 15, 30, and 40 Hz), and random noise. Fig. 14 shows this synthetic dataset and its amplitude spectrum.

The conventional MCA method and the proposed method are applied to this synthetic dataset and further extract the HST-induced signal. The HST-induced signals extracted by the conventional MCA method and the proposed method are shown in Fig. 15. Comparing the results in the time domain, the signal extracted by the proposed method [Fig. 15(c)] can recover better periodicity and gain better anti-noise performance than the conventional method [Fig. 15(a)], where the residuals of Ricker can be observed at the red ellipse in Fig. 15(a). For the frequency domain comparison, the extracted
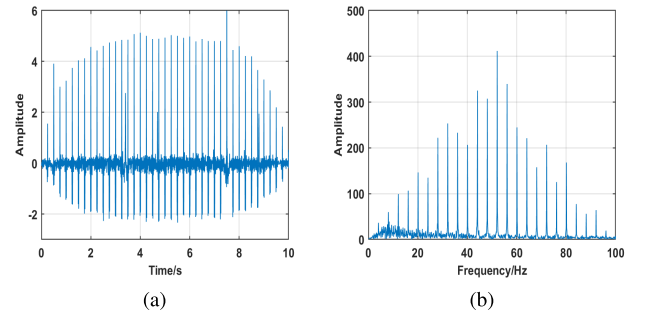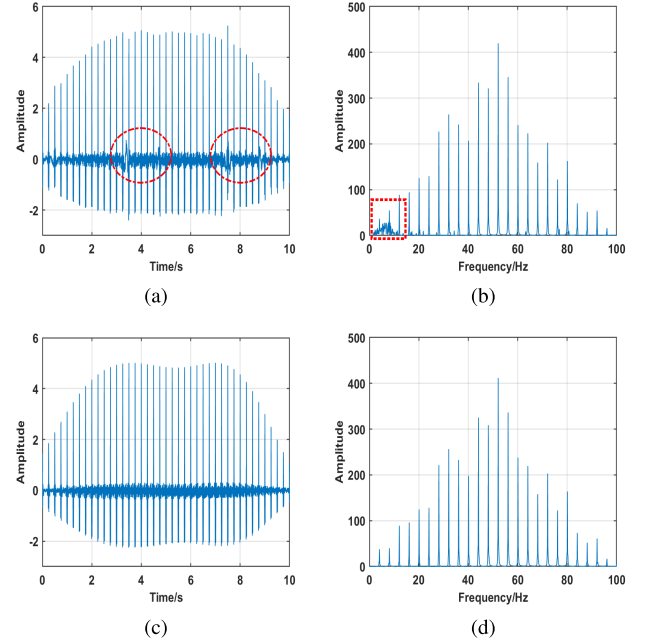
HST-induced signal by the proposed method [Fig. 15(d)] has a more obvious ENBD-spectra feature than the conventional MCA's result [Fig. 15(b)] which still contains some unwanted energy in the red dotted rectangle area. Fig. 16 shows the noise separated by the conventional MCA method and the proposed method, which contains the random noise and some Ricker wavelets. Generally, two waveforms provided by the two methods [Fig. 16(a) and (c)] are similar. However, the conventional method provides an amplitude spectrum where some low-frequency components are missing [Fig. 16(b)], while the proposed method yields a continuous amplitude spectrum [Fig. 16(d)], which also confirmed that the low-frequency Ricker wavelets are separated and classified to the right part.

To quantitatively evaluate the conventional MCA method and the proposed method, SNR is used. Here, SNR is defined as the ratio between HST-induced signal and other parts. The SNR of the synthetic dataset is 11.446 dB. The SNR comparison is shown in Table II, the proposed method can gain better SNR performance than the conventional MCA method. Through the comparison of these results, it is evident
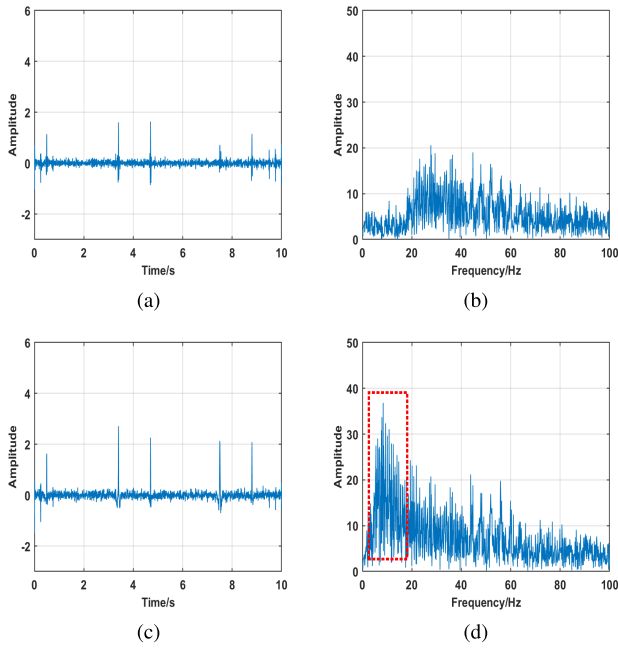
Fig. 16.   Separated interference of single trace. (a) Waveform and (b) amplitude spectrum of the conventional MCA method. (c) Waveform and (d) amplitude spectrum of our proposed method.

TABLE II
COMPARISON OF SNR OF THE HST-INDUCED SEISMIC SIGNALS

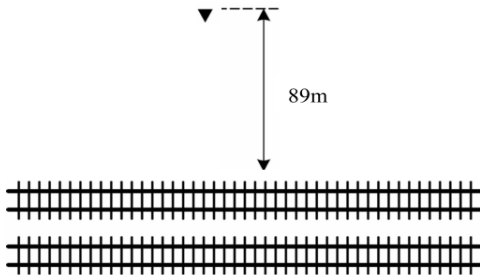| Method	SNR
Unprocessed data	11.446 dB
The conventional MCA	13.962 dB
Our proposed method	19.286 dB

Fig. 17.   Schematic of seismic data acquisition for a high-speed railway in southern China.

that the proposed method can provide better data separation performance.

2) Real Data Example: To further test the performance, the seismic data collected by one seismic geophone near the high-speed railway in southern China are used (the position of seismic receiver is shown in Fig. 17).

Fig. 1(c) shows the received seismic data by receiver when one HST passed the geophone. Fig. 1(d) proves this dataset is dominated by the ENBD-spectra component which is caused by the moving HST. Fig. 18 shows the extracted HST-induced seismic signals' waveforms and amplitude spectra. Compared with the conventional MCA method [Fig. 18(a)], the proposed method recovers a more obvious periodic feature in the time domain [as the red arrows pointed in Fig. 18(c)], which
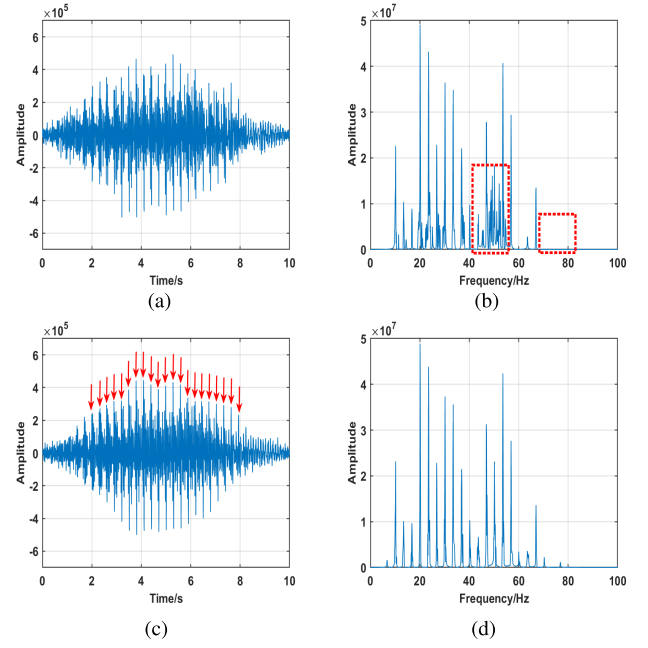

Fig. 18.   Extracted effective signal of single trace. (a) Waveform and (b) amplitude spectrum of the conventional MCA method. (c) Waveform and (d) amplitude spectrum of our proposed method.
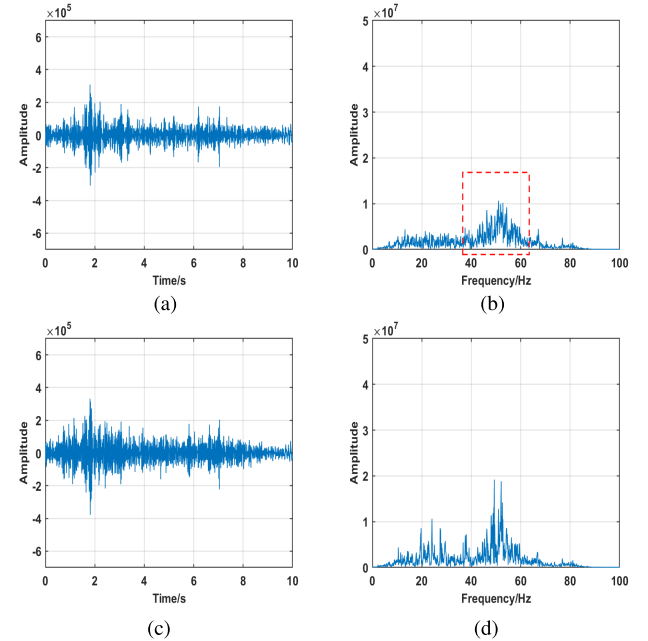
Fig. 19.   Separated interference of single trace. (a) Waveform and (b) amplitude spectrum of the conventional MCA method. (c) Waveform and (d) amplitude spectrum of our proposed method.

conforms to the fact that all the carriages of one HST have a similar structure. Moreover, the conventional MCA method has noise residual [Fig. 18(b)], while the proposed method recovers can recover the clear ENBD-spectra feature in the frequency domain [Fig. 18(d)].

Fig. 19 shows the noise part separated by the conventional MCA method and the proposed method. Compared with the conventional MCA method's noise part in Fig. 19(a) and (b), the noise part separated by the proposed method has a stronger energy. We also compare the separation results of

(a)
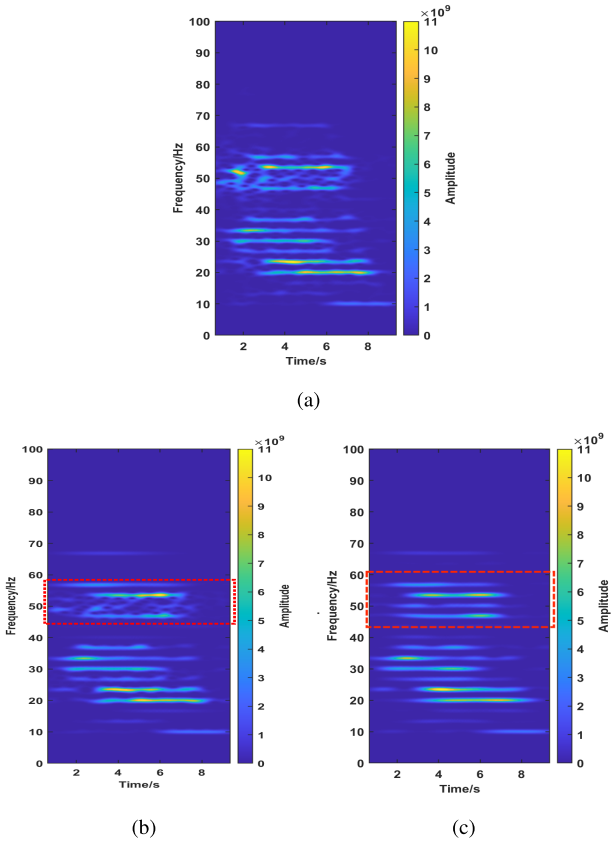

(b)

(c)

Fig. 20. Results of time–frequency spectrum. (a) Original signal. (b) Effective signals of the conventional MCA method. (c) Effective signals of our proposed method.
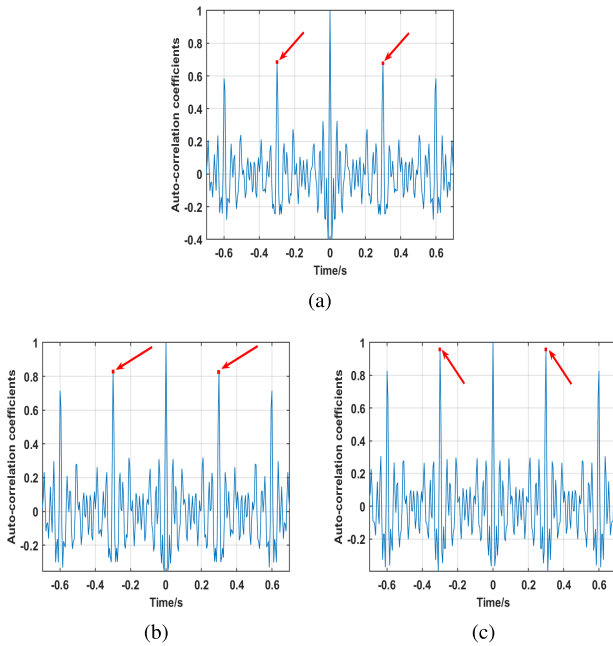

(a)

(b)

(c)

Fig. 21. Autocorrelation coefficients of HST source seismic signals. (a) Unprocessed real data. (b) Conventional MCA method. (c) Our proposed method.

HST-induced seismic signals in the time–frequency domain and show the time–frequency spectra in Fig. 20. Fig. 20(c) proves the proposed method can extract the ENBD-spectra

feature component better than the conventional MCA method [Fig. 20(b)].

As Fig. 13 shows, when the HST with a periodic structure maintains a constant velocity, the HST-induced signal should have a periodic feature. To further evaluate the extracted HST-induced seismic signals, we compute their autocorrelation functions and show them in Fig. 21. When the time lag is −0.3 and 0.3 s, these autocorrelation functions reach the second largest maximum. Specifically, the second largest maximum for the autocorrelation function of received data is 0.68. After separation, the second largest maxima for the autocorrelation function of the extracted HST signal by the conventional method and the proposed method are 0.83 and 0.95, respectively. The autocorrelation function evaluation verifies that the HST-induced signal extracted by the proposed method has a higher periodicity than that of the conventional method, which shows the proposed method can provide a more complete separation result than the conventional MCA method for the complex data with the ENBD-spectra component.

IV. CONCLUSION

This study proposes an improved MCA method to separate seismic data with the ENBD-spectra component. The seismic data are modeled as a linear superposition of two components: one exhibiting ENBD-spectra feature and the other lacking this characteristic. Based on the morphological differences in these components, the DFT and the CWT are selected as the dictionary of the ENBD-spectra component and other components. Besides, the ENBD-spectra feature is used as constraints to introduce MCA, which can effectively and accurately separate the seismic data with ENBD-spectra feature and other components. Both the synthetic seismic examples and the real seismic examples show that our proposed method can effectively reduce the damage and distortion of the target signal caused by the conventional methods. Moreover, this method can be extended to the separation of other signals with ENBD-spectra component.

REFERENCES

[1] L. Li and Z. Li, *Principles, Method, and Interpretation of Seismic Exploration*. Bath, U.K.: Geological Press, 2007.

[2] W. Chen et al., "3-D seismic diffraction separation and imaging using the local rank-reduction method," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 4507110.

[3] R. Anvari, A. R. Kahoo, M. S. Monfared, M. Mohammadi, R. M. D. Omer, and A. H. Mohammed, "Random noise attenuation in seismic data using Hankel sparse low-rank approximation," *Comput. Geosci.*, vol. 153, Aug. 2021, Art. no. 104802.

[4] J. Mafakheri, A. R. Kahoo, R. Anvari, M. Mohammadi, M. Radad, and M. S. Monfared, "Expand dimensional of seismic data and random noise attenuation using low-rank estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2773–2781, Mar. 2022.

[5] G. Li, Y. Li, and B. Yang, "Seismic exploration random noise on land: Modeling and application to noise suppression," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4668–4681, Aug. 2017.

[6] Y. Chen, S. Jiao, J. Ma, H. Chen, Y. Zhou, and S. Gan, "Ground-roll noise attenuation using a simple and effective approach based on local band-limited orthogonalization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2316–2320, Nov. 2015.

[7] Z. Feng, "Robust ground roll noise suppression based on dictionary learning and bandpass filtering," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[8] B. Zhang, Y. Yang, B. Yu, and Y. Zhang, "Direct wave and clutter cancellation algorithm based on noise signal and compressed sensing," in *Proc. IEEE 6th Int. Conf. Signal Image Process. (ICSIP)*, Oct. 2021, pp. 768–772.

[9] B. Wu, J. Yu, H. Ren, Y. Lou, and N. Liu, "Seismic traffic noise attenuation using l_p -Norm robust PCA," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 11, pp. 1998–2001, Nov. 2020.

[10] J. Shao, Y. Wang, and L. Chen, "Near-surface characterization using high-speed train seismic data recorded by a distributed acoustic sensing array," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 5912911.

[11] S. Gao, B. Zhao, S. Zhu, G. Luo, and Z. He, "Identification and elimination of monofrequency interference in seismic data by an autocorrelation algorithm," *Chin. J. Geophys.*, vol. 54, no. 3, pp. 854–861, 2011.

[12] F. Fang and Z. Yang, "Time domain suppression scheme of narrowband interference based on narrowband OFDM power line communications," in *Proc. Int. Conf. Commun. Electron. Syst. (ICCES)*, Oct. 2016, pp. 1–6.

[13] D. C. Henley, "Spectral clipping: A ProMAX module for attenuating strong monochromatic noise," *CREWES Calg. AB Can*, vol. 13, pp. 311–320, Jan. 2001.

[14] H. Xia, "Elimination of 50hz power interference via wavelet packet transform and sine function simulation," *Prog. Explor. Geophys.*, vol. 28, pp. 341–344, 2005.

[15] W. Wei, C. Liao, T. Deng, W. Ding, and D. Wang, "Research on frequency domain filtering algorithm for power line interference in high speed data acquisition system," *J. Geodesy Geodyn.*, vol. 31, no. 1, pp. 100–104, Apr. 2011.

[16] C. Lu, P. Guo, L. Feng, A. Yang, J. Wang, and C. Xing, "An intentional acoustic interference approach to control output signals of MEMS gyroscope based on short-time Fourier analysis," in *Proc. 20th Int. Conf. Electron. Packag. Technol.*, 2019, pp. 1–4.

[17] C.-C. Huang, S.-F. Liang, M.-S. Young, and F.-Z. Shaw, "A novel application of theS-transform in removing powerline interference from biomedical signals," *Physiolog. Meas.*, vol. 30, no. 1, pp. 13–27, Jan. 2009.

[18] Y. Yang, X. Cai, and X. Li, "The denosing of seismic signal based on generalized S transform," in *Proc. 6th Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, vol. 1, Aug. 2014, pp. 275–278.

[19] Y. Li, Y. Zhang, and X. Xu, "Advances and perspective on morphological component analysis based on sparse representation," *Acta Electronica Sinica*, vol. 37, no. 1, pp. 146–152, 2009.

[20] E. Temlioglu and I. Erer, "Clutter removal in ground-penetrating radar images using morphological component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1802–1806, Dec. 2016.

[21] D. Liu, W. Wang, X. Wang, Z. Shi, M. D. Sacchi, and W. Chen, "Improving sparse representation with deep learning: A workflow for separating strong background interference," *Geophysics*, vol. 88, no. 1, pp. WA253–WA266, Jan. 2023.

[22] Z. Hu, J. Gao, and N. Liu, "Separation of blended seismic data using the synchrosqueezed curvelet transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 4, pp. 711–715, Apr. 2020.

[23] W.-Q. Lim, "The discrete shearlet transform: A new directional transform and compactly supported shearlet frames," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1166–1180, May 2010.

[24] D. Liu, H. Zhang, X. Wang, W. Chen, Z. Shi, and Z. Zhao, "Separation of seismic multiple reflection-refraction based on morphological component analysis with high-resolution linear radon transform," *Geophysics*, vol. 87, no. 4, pp. V367–V379, Jul. 2022.

[25] X. Wang, J. Chen, W. Chen, Y. Jiang T. Bao, and J. Ning, "Sparse modeling of seismic signals produced by high-speed trains," *Chin. J. Geophys.*, vol. 62, no. 6, pp. 2336–2343, Jun. 2019.

[26] J. Xu, W. Wang, J. Gao, and W. Chen, "Monochromatic noise removal via sparsity-enabled signal decomposition method," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 533–537, May 2013.

[27] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.

[28] J.-L. Starck, M. Elad, and D. Donoho, "Redundant multiscale transforms and their application for morphological component separation," in *Advances in Imaging and Electron Physics*, vol. 132. USA: Elsevier, Dec. 2004, pp. 287–348.

[29] J. Bobin, J.-L. Starck, J. M. Fadili, Y. Moudden, and D. L. Donoho, "Morphological component analysis: An adaptive thresholding strategy," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2675–2681, Nov. 2007.

[30] A. G. Bruce and S. Sardy, "Block coordinate relaxation methods for nonparamatric signal denoising," *Proc. SPIE Wavelet Appl. V*, vol. 3391, pp. 75–86, Mar. 1998.

[31] Z. Shao, J. He, and S. Feng, "Separation of multicomponent chirp signals using morphological component analysis and fractional Fourier transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1343–1347, Aug. 2020.

[32] W. Xiaokai, W. Baoli, C. Wenchao, and L. Jiaqi, "Using the data from one receiver to estimate running velocity of high-speed train," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 55, no. 5, pp. 798–804, 2019.

[33] X. Li, W. Chen, W. Wang, X. Wang, and Y. Zhou, "Harmonic noise removal in united time-frequency domain via sparse promotion," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2015, pp. 4725–4729.

[34] C. Zhao, L. Jiang, X. Wang, D. Liu, Z. Shi, and W. Chen, "Prestack seismic noise attenuation based on 3D CWT," in *Proc. 1st Int. Meeting Appl. Geosci. Energy Expanded Abstr.*, Sep. 2021, pp. 2834–2838.

[35] J. Bobin, Y. Moudden, and J.-L. Starck, "Enhanced source separation by morphological component analysis," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 5, Jan. 2006, p. 5.

[36] H. Özdemir and R. Saatçilar, "Efficient multichannel filtering of seismic data," *Geophys. Prospecting*, vol. 38, no. 1, pp. 1–22, Jan. 1990.

[37] Y. Liu, Y. Yue, Y. Li, and Y. Luo, "On the retrievability of seismic waves from high-speed-train-induced vibrations using seismic interferometry," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[38] H. Wang and J. Chen, "Wavefield modeling and characteristic analysis for a uniformly moving high-speed-train source under a two-layered model with applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Jan. 2023, Art. no. 5919314.

[39] L. Zhang, "Research on train load transfer characteristics and mechanism in ballastless track of high-speed railway," Ph.D. dissertation, Beijing Jiaotong Univ., Beijing, China, Tech. Rep. 14115339, 2020.

[40] X. Wang, B. Wang, and W. Chen, "The second-order synchrosqueezing continuous wavelet transform and its application in the high-speed-train induced seismic signal," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 6, pp. 1109–1113, Jun. 2021.

[41] J. Luo, X. Wang, and W. Chen, "Imaging the subsurface with the high-speed train seismic data-based elastic reverse time migration," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4505409.

[42] G. Kouroussis, D. P. Connolly, and M. C. Forde, "Train speed calculation using ground vibrations," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 229, no. 5, pp. 466–483, 2015.

[43] J. Luo, X. Wang, and W. Chen, "Subsurface elastic parameter reconstruction based on seismic data from the high-speed trains using full waveform inversion," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–8, Feb. 2022.

Xiaokai Wang (Member, IEEE) received the B.S. degree in information engineering and the Ph.D. degree in communication engineering from Xi'an Jiaotong University (XJTU), Xi'an, China, in 2006 and 2012, respectively.

He was a Post-Doctoral Fellow at the Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, from 2012 to 2014. He joined the Department of Computational Geophysics, XJTU, in 2015. He was a Visiting Scholar at the Bureau of Economic Geology, The University of Texas at Austin, Austin, TX, USA, from 2016 to 2017. Since 2018, he has been with the School of Information and Communication Engineering, XJTU, where he is currently a Professor. His research interests include seismic attributes extraction, time–frequency analysis, and high-speed rail seismology.

Chunmeng Cui received the B.S. degree in information engineering from Xi'an Jiaotong University (XJTU), Xi'an, China, in 2021, where he is currently pursuing the master's degree with the School of Information and Communication Engineering.

His research interests include sparse representation theory and seismic signal processing.

Zhensheng Shi received the M.S. degree in electromagnetic fields and microwave technology from Xi'an Jiaotong University (XJTU), Xi'an, China, in 1989.

He is currently a Senior Engineer with the School of Information and Communication Engineering, XJTU. His research interests include RF, microwave circuit design, and radar signal processing and imaging.

Dawei Liu received the bachelor's degree in communication engineering from Chang'an University, Xi'an, China, in 2013, the master's degree in electronic and communication engineering from Xi'an Jiaotong University (XJTU), Xi'an, in 2018, and the Ph.D. degree from the School of Information and Communication Engineering, XJTU, in 2022.

From 2022 to 2023, he was a Post-Doctoral Scholar with the Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA. He is currently a Post-Doctoral Fellow with the Department of Physics, University of Alberta, Edmonton, AB, Canada. His research interests include tensor decomposition, deep learning, and time–frequency analysis and their applications in seismic data processing.

Dr. Liu is a reviewer for IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Geophysics*, *Acta Geophysica*, and *Petroleum Science*.
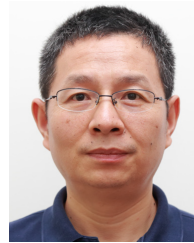
Wenchao Chen received the B.S. and M.S. degrees in seismic exploration and information technology from Chang'an University, Xi'an, China, in 1993 and 1996, respectively, and the Ph.D. degree in electromagnetic field and microwave technology from Xi'an Jiaotong University, Xi'an, in 2000.

From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Computation Science, Northwestern Polytechnical University, Xi'an. Since 2002, he has been with Xi'an Jiaotong University, where he is currently a Professor. His research interests include high-speed rail seismology, seismic and GPR signal processing, blind signal processing, and sparse representation.

Dr. Chen is a member of SEG and the Chinese Geophysical Society. He is also a referee for several journals, including IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, *Geophysics*, *Interpretation*, and *Journal of Applied Geophysics*.

Pu Liu, photograph and biography not available at the time of publication.