Research Article

A Feature Extraction Method of Wheelset-Bearing Fault Based on Wavelet Sparse Representation with Adaptive Local Iterative Filtering

Zhan Xing, Jianhui Lin , Yan Huang, and Cai Yi

State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China

Correspondence should be addressed to Jianhui Lin; lin13008104673@126.com

Received 15 December 2019; Revised 14 May 2020; Accepted 10 June 2020; Published 25 July 2020

Academic Editor: Zhixiong Li

Copyright © 2020 Zhan Xing et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The feature extraction of wheelset-bearing fault is important for the safety service of high-speed train. In recent years, sparse representation is gradually applied to the fault diagnosis of wheelset-bearing. However, it is difficult for traditional sparse representation to extract fault features ideally when some strong interference components are imposed on the signal. Therefore, this paper proposes a novel feature extraction method of wheelset-bearing fault based on the wavelet sparse representation with adaptive local iterative filtering. In this method, the adaptive local iterative filtering reduces the impact of interference components effectively and contributes to the extraction of sparse impulses. The wavelet sparse representation, which adopts L1-regularized optimization for a globally optimal solution in sparse coding, extracts intrinsic features of fault in the wavelet domain. To validate the effectiveness of this proposed method, both simulated signals and experimental signals are analyzed. The results show that the fault features of wheelset-bearing are sufficiently extracted by the proposed method.

1. Introduction

As the core system of high-speed trains, the bogie frame plays an extremely crucial role in the operation process. Among the components of the bogie frame, wheelset-bearing is the core component for the connection between the wheelset and the frame. When a high-speed train operates on a rail, the wheelset-bearing plays an important role in the power transmission. Compared with the other common bearings on the static mechanical equipment, the service conditions of the wheelset-bearing are quite different. The wheelset-bearing bears not only the static pressure of high-speed train but also the unstable dynamic load caused by the radial acceleration during operating time. A higher speed naturally causes a greater vibration and dynamic force to the wheelset-bearing. In addition, the wheelset-bearing will bear a large axial force when the train passes through a curve.

When a high-speed train operates on a rail, various operation characteristics could cause the wheelset-bearing fault. Since the distance between adjacent stations is

usually not long for high-speed train, the acceleration and the braking of high-speed train occur frequently. This causes the dynamic load range of the wheelset-bearing to change greatly and frequently. In addition, due to the effect of overwhelming impact from track-vehicle system, which is caused by the polygonal wear of wheels, track irregularities, and irregular turnout, various faults (such as peeling, flaw) might appear on the wheelset-bearing. Once these faults appear under the condition of high-speed rotation, the service conditions of wheelset-bearing will deteriorate rapidly, which will eventually affect the safety of high-speed train. Therefore, it is of great significance to detect wheelset-bearing fault [1]. As for the traditional fault detection of wheelset-bearing, it cannot be conducted until the wheelset-bearing is disassembled during the level-three maintenance of a high-speed train. This indicates that fault detection of wheelset-bearing cannot be executed before the level-three maintenance. Therefore, the fault diagnosis, based on vibration signal, becomes a feasible technique for wheelset-bearing fault in the early stage.

In general, when the bearing fault occurs, the periodical impulses are generated. Therefore, the vibration signals are collected to determine whether the fault exists on the bearing. During the operation of high-speed train, the defect inevitably appears on the wheel-tread. Compared with the traditional bearing, the energy of the rotation frequency for wheelset, generated by the defect on the wheel-tread, is larger because of the interaction between wheel-tread and rail. Therefore, the vibration information of both rotation frequency and harmonics for wheelset is also evidently contained in the collected vibration signals, which makes the frequency components contained in the vibration signals more complex [2]. This is the most important difference between the vibration signals of traditional bearing and wheelset-bearing. If the rotation frequency information of wheelset cannot be handled appropriately in the analytical process, the mistaken fault features might be extracted. In addition, the collected vibration signals usually contain nonstationary component. They are also interfered with noise component, which causes extraordinarily challenges for the detection of fault [3, 4].

As for the fault detection in a general bearing, many fault diagnosis methods, including empirical model decomposition and its variants [5, 6], empirical wavelet decomposition [7, 8], variational mode decomposition [9, 10], minimum entropy deconvolution [11, 12], local mean decomposition [13, 14], deep learning [15, 16], and sparse representation [17, 18], have been proposed for bearing fault detection. Among these techniques, the sparse representation might be an advanced method for feature extraction of bearing fault. Recently, many scholars have devoted themselves to applying sparse representation methods to bearing fault detection. Chen et al. [17] proposed a method named SpaEIAD for fault extraction, which showed a good performance in denoising of signal. In [19], feature-sign search was adopted for sparse representation which also obtained a desired extraction result. Sun et al. [20] designed a parametric impulsive dictionary and improved the stopping criteria of the OMP for sparse representation. Ding [21] proposed a shock response convolutional sparse coding technique and achieved the extraction of shock response based on time location coefficients. Qin [22] proposed a new sparse representation method based on a family of model-based impulsive wavelets, which was able to accurately represent the bearing fault impulses. In [23], a bearing fault extraction method, based on the adaptive OMP algorithm and improved K-SVD with adaptive transient dictionary, was proposed and it achieved a good effect in fault detection and computation speed.

In the traditional sparse representation, the power levels of different features will affect the results extracted by the sparse representation [24]. When the energy of interference, usually expressed as the nonstationary component, is stronger than the energy of fault features, it will detect and extract the interference component instead of the fault features. Considering this condition, Qin separated the harmonics and modulated components from the vibration signal of gearbox bearing successfully with the improved OMP and Fourier dictionary in [23]. This method achieved

an outstanding performance in fault extraction through two experiments. However, the service conditions of the wheelset-bearing are more complex than those of the gearbox bearing in [23]. As analysis showed before, apart from the influence of nonstationary component, the wheelset's rotation frequency component also should be considered to be removed. Due to this reason, the aforementioned separation algorithm in [23] should be executed additionally, which increases the complexity of the algorithm. In addition, the Fourier dictionary might not be a suitable dictionary for the separation of the wheelset's rotation frequency component. Therefore, Qin's algorithm might not be fully suitable for the application of fault extraction for the wheelset-bearing. According to the analysis in [2], two resonance frequencies, excited by the defect of wheel-tread and the wheelset-bearing fault, are more likely at low frequency and high frequency, respectively. Additionally, the frequency of nonstationary component, which is distributed throughout the signal, is also at low frequency relative to the resonance frequency excited by the wheelset-bearing fault. In order to remove the impact of these two components conveniently and effectively, adaptive local iterative filtering (ALIF) [25] is introduced. The ALIF is suitable to process the separation of the components belonging to different frequency. Therefore, the ALIF can reduce or eliminate the influence of aforementioned two components.

When sparse representation is solely applied in time domain, it will lead to the inadequate feature extraction. This indicates that the acquired fault features cannot be extracted thoroughly. The wavelet domain is another scale representation of signal [26]. The most useful information of the signal in time domain can be compressed and represented in wavelet domain without losing local information. During the process of wavelet decomposition, the noise component can be partially separated. This highlights the local features of signal. In addition, the initial atoms of constructed dictionary, obtained in wavelet domain, are more fitted to the signal. Therefore, the wavelet sparse representation, which indicates that the sparse representation is applied in the wavelet domain, is able to extract intrinsic features of signals.

To diagnose wheelset-bearing fault more effectively, a novel feature extraction method, namely, ALIF-SBAKW, based on the wavelet sparse representation (Split Bregman for sparse coding and approximate K-SVD for dictionary learning) with the adaptive local iterative filtering (ALIF), is proposed in this paper. The paper is organized as follows. Section 2 elaborates the details of wavelet sparse representation. Section 3 describes the main principle of ALIF-SBAKW for feature extraction. The proposed ALIF-SBAKW is verified by simulations and experiments, respectively, in Sections 4 and 5. Section 6 concludes the paper.

2. Wavelet Sparse Representation

2.1. Sparse Representation. According to the sparsity of fault impulses, the observed signals can be represented sparsely by combining the dictionary with the sparse coefficient, as shown in

$$\mathbf{y}_o = \mathbf{D}\mathbf{Z} + \mathbf{e}, \quad (1)$$

where $\mathbf{y}_o \in R^n$ denotes the observed signal, $\mathbf{e} \in R^n$ is the measurement noises, $\mathbf{D} \in R^{n \times m}$ denotes dictionary matrix, and $\mathbf{Z} \in R^m$ denotes sparse coefficient vector. The extracted signal can be constructed by the multiplication of dictionary $\mathbf{D}$ and sparse coefficient $\mathbf{Z}$.

It can be observed that (1) is an underdetermined equation, which means (1) has infinite solutions. To solve this kind of problem, it can be turned into the problem of finding the L0-regularized optimization and L1-regularized optimization. Compared with L0-regularized optimization, L1-regularized optimization is more suitable for solving an NP-hard problem [19]. Therefore, L1-regularized optimization is adopted. The equation of L1-regularized optimization is given in

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \|\mathbf{Z}\|_1, \\ \text{s.t.} \quad & \|\mathbf{y}_o - \mathbf{D}\mathbf{Z}\|_2^2 \leq T^2, \end{aligned} \quad (2)$$

where $T = c\sigma$ and c denotes adjusted gain of standard deviation σ of noise.

To solve this kind of optimization problem, a penalty factor ρ can be introduced to reduce the constraint. A new objective function can be obtained:

$$\{\hat{\mathbf{D}}, \hat{\mathbf{Z}}\} = \arg \min \left(\frac{\rho}{2} \|\mathbf{y}_o - \mathbf{D}\mathbf{Z}\|_2^2 + \|\mathbf{Z}\|_1 \right). \quad (3)$$

In fact, the length of signal, i.e., n , is usually a large number, which means that a very large amount of computing resources is consumed in solving the problem. In order to reduce the computational burden, the observed signal can be segmented into a series of truncated-signals $\mathbf{y}_i$ with a certain overlap; as shown in Figure 1, dataset $\mathbf{Y}$ comprised of the truncated-signals can be obtained. Equation (3) can be solved in the form of matrix $\mathbf{Y}$. Assuming that the dataset is $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_l] \in R^{b \times l}$ after segmentation, accordingly dictionary matrix and sparse coefficient matrix are defined as $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_k] \in R^{b \times k}$ and $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_l] \in R^{k \times l}$, respectively. Therefore, (3) is transformed into

$$\begin{aligned} \{\hat{\mathbf{D}}, \hat{\mathbf{Z}}\} &= \arg \min \left(\frac{\rho}{2} \|\mathbf{y}_i - \mathbf{D}\mathbf{z}_i\|_2^2 + \|\mathbf{z}_i\|_1 \right), \\ \|\mathbf{d}_j\|_F^2 &= 1, \\ j &= 1, 2, \dots, k; i = 1, 2, \dots, l. \end{aligned} \quad (4)$$

To solve the objective optimization problem in (4), two optimization steps containing both sparse coding and dictionary learning are executed. In sparse coding, a sparse approximation is used to find sparse coefficients with a fixed dictionary. The dictionary learning step is used to update the dictionary with the obtained sparse coefficients.

2.2. Wavelet Sparse Representation. In order to better extract the intrinsic features of signals, the wavelet decomposition can be adopted in sparse representation. As for a one-

FIGURE 1: The signal is segmented into a series of truncated-signals with a certain overlap.

dimensional signal, there are mainly two forms of coefficients after wavelet decomposition: approximation coefficients (CAs) and detail coefficients (CDs). The CAs contain the main information of the original signal. On the contrary, the CDs contain the subordinate components of the original signal. The wavelet decomposition of the signal is given in Figure 2. It can be observed that the number of decomposed subbands is related to the decomposed level q . These subbands mainly consist of a CA band and q CD bands. Accordingly the number of subbands is $q + 1$.

The fault features of bearing are mainly hidden in the CAs after the wavelet decomposition of signal. Accordingly, they can be called impulse wavelet coefficients (IWCs). In addition, the CDs mainly contain the noise component of signal. According to the performance of wavelet decomposition, the sparse representation can be accomplished by the obtained IWCs. The objective function can be transformed into

$$\begin{aligned} \{(\hat{\mathbf{D}})_I, (\hat{\mathbf{z}}_i)_I\} &= \arg \min \left(\frac{\rho}{2} \|(\mathbf{y}_i)_I - (\mathbf{D})_I(\mathbf{z}_i)_I\|_2^2 + \|(\mathbf{z}_i)_I\|_1 \right), \\ \|(\mathbf{d}_j)_I\|_F^2 &= 1, \quad j = 1, 2, \dots, k; i = 1, 2, \dots, l, \end{aligned} \quad (5)$$

where $(\mathbf{y}_i)_I$ denotes the i -th column of dataset $(\mathbf{Y})_I$, which is composed of the segmented IWCs in the wavelet domain. Equation (5) suggests that the dictionary matrix $(\mathbf{D})_I$ and sparse coefficient matrix $(\mathbf{Z})_I$ can be obtained after the wavelet sparse representation. As for the CDs, they can be directly set to zero because they mainly contain noise component. It should be noticed that the level of wavelet decomposition has great significance for the IWCs. The wavelet decomposition of the one-dimensional signal is a process of downsampling. If there are too many required levels, the IWCs will be distorted, causing the fault features, hidden in the IWCs, to be weakened or to disappear. Therefore, the choice of level for wavelet decomposition is extraordinarily important.

FIGURE 2: The wavelet decomposition of signal.

At the beginning of wavelet sparse representation, k columns of dataset $\mathbf{Y}$ are randomly selected as the initialization dictionary ($k < l$). During the calculation, the row b of the designed dictionary is always much smaller than the column k of the designed dictionary. This makes the dictionary redundant. With the feature of redundancy, the advantage of this kind of dictionary is that it is more conducive to expressing a highly diversified signal and reconstructing the local features of the signal. Furthermore, the noise is generally not considered to be sparse. Therefore, the fault signal, which needs to be extracted from the noise, will become much sparser and more stable by using the representation of redundant dictionary.

2.3. Split Bregman for Sparse Coding. Split Bregman (SB) iteration is one of the effective methods of L1-regularized optimization. Due to its ability to solve a very wide class of L1-regularized problems by using alternating iteration, SB has been widely used in the field of image processing [27, 28]. It is also suitable for sparse coding. In this paper, the sparse coding problem needs to be solved in the form of (6) and (7) based on the principle of SB:

$$\min \quad \|(\mathbf{f}_i)_I\|_1 + \frac{\rho}{2} \|(\mathbf{y}_i)_I - (\mathbf{D})_I (\mathbf{z}_i)_I\|_2^2, \quad (6)$$

$$\text{s.t.} \quad (\mathbf{f}_i)_I = (\mathbf{z}_i)_I,$$

$$\min \quad \|(\mathbf{f}_i)_I\|_1 + \frac{\rho}{2} \|(\mathbf{y}_i)_I - (\mathbf{D})_I (\mathbf{z}_i)_I\|_2^2 + \frac{\lambda}{2} \|(\mathbf{f}_i)_I - (\mathbf{z}_i)_I\|_2^2, \quad (7)$$

where $(\mathbf{D})_I \in R^{b \times k}$ and $(\mathbf{y}_i)_I \in R^b$ are given. $(\mathbf{z}_i)_I \in R^k$ and $(\mathbf{f}_i)_I \in R^k$ are the unknown quantities. Defining $F((\mathbf{z}_i)_I, (\mathbf{f}_i)_I) = \|(\mathbf{f}_i)_I\|_1 + (\rho/2) \|(\mathbf{y}_i)_I - (\mathbf{D})_I (\mathbf{z}_i)_I\|_2^2$, an elegant form of iteration, based on the principle of SB, can be obtained by the simplification to (7), as seen in (8) and (9).

$$((\mathbf{z}_i)_I^{r+1}, (\mathbf{f}_i)_I^{r+1}) = \min F((\mathbf{z}_i)_I, (\mathbf{f}_i)_I) + \frac{\lambda}{2} \|(\mathbf{f}_i)_I - (\mathbf{z}_i)_I - (\mathbf{w}_i)_I^r\|_2^2, \quad (8)$$

$$(\mathbf{w}_i)_I^{r+1} = (\mathbf{w}_i)_I^r + ((\mathbf{z}_i)_I^{r+1} - (\mathbf{f}_i)_I^{r+1}). \quad (9)$$

In (8), the item $F((\mathbf{z}_i)_I, (\mathbf{f}_i)_I)$ contains L1 and L2 components with two different independent variables. When the value of one of the variables is updated, the other variable can be seen as a constant. Due to this alternating characteristic, (8) can be split into two steps:

$$(\mathbf{z}_i)_I^{r+1} = \min \frac{\rho}{2} \|(\mathbf{y}_i)_I - (\mathbf{D})_I (\mathbf{z}_i)_I\|_2^2 + \frac{\lambda}{2} \|(\mathbf{f}_i)_I^r - (\mathbf{z}_i)_I - (\mathbf{w}_i)_I^r\|_2^2, \quad (10)$$

$$(\mathbf{f}_i)_I^{r+1} = \min \|(\mathbf{f}_i)_I\|_1 + \frac{\lambda}{2} \|(\mathbf{f}_i)_I - (\mathbf{z}_i)_I^{r+1} - (\mathbf{w}_i)_I^r\|_2^2. \quad (11)$$

When solving (10), the solution $(\mathbf{z}_i)_I^{k+1}$ can be obtained by differentiation with corresponding independent variable and setting result to zero. Therefore, the computed equation of $(\mathbf{z}_i)_I^{k+1}$ is shown below:

$$(\mathbf{z}_i)_I^{r+1} = (\rho(\mathbf{D})_I^T (\mathbf{D})_I + \lambda \mathbf{I})^{-1} (\rho(\mathbf{D})_I^T (\mathbf{y}_i)_I + \lambda ((\mathbf{f}_i)_I^r - (\mathbf{w}_i)_I^r)). \quad (12)$$

In addition, $(\mathbf{f}_i)_I^{k+1}$ can be obtained by using a shrinkage operator at the second iteration, as shown in (13) and (14). $\odot$ denotes elementwise multiplication.

$$(\mathbf{f}_i)_I^{r+1} = \text{shrink}\left((\mathbf{z}_i)_I^{r+1} + (\mathbf{w}_i)_I^r, \frac{1}{\lambda}\right), \quad (13)$$

$$\text{shrink}(\mathbf{e}, v) = \text{sign}(\mathbf{e}) \odot \max(|\mathbf{e}| - v, \mathbf{0}). \quad (14)$$

2.4. Approximate K-SVD for Dictionary Learning. After the sparse coefficient matrix $(\mathbf{Z})_I$ is calculated, the dictionary $(\mathbf{D})_I$ will be considered updated. K-SVD is an available method for dictionary learning, which can effectively reduce the sparsity of the corresponding sparse coefficient matrix of dictionary [23]. The objective function can be modified as follows:

$$\begin{aligned} (\hat{\mathbf{D}})_I &= \arg \min \|(\mathbf{Y})_I - (\mathbf{D})_I (\mathbf{Z})_I\|_2^2 = \arg \min \left\| (\mathbf{Y})_I - \sum_{j=1}^k (\mathbf{d}_j)_I (\mathbf{z}_T^j)_I \right\|_2^2 \\ &= \arg \min \left\| \left((\mathbf{Y})_I - \sum_{j=1, j \neq i}^k (\mathbf{d}_j)_I (\mathbf{z}_T^j)_I \right) - (\mathbf{d}_i)_I (\mathbf{z}_T^i)_I \right\|_2^2 \\ &= \arg \min \|(\mathbf{E}_i)_I - (\mathbf{d}_i)_I (\mathbf{z}_T^i)_I\|_2^2, \quad \|(\mathbf{d}_j)_I\|_F^2 = 1, \quad j = 1, 2, \dots, k, \end{aligned} \quad (15)$$

where $(\mathbf{z}_T^j)_I \in R^{1 \times l}$ denotes the j -th row of $(\mathbf{Z})_I$. Based on the principle of K-SVD, the optimization of (15) is equivalent to optimizing the nonzero elements in $(\mathbf{z}_T^i)_I$:

$$\arg \min \left\| (\mathbf{E}_i)_I \Phi_i - (\mathbf{d}_i)_I (\mathbf{z}_T^i)_I \Phi_i \right\|_2^2 = \arg \min \left\| (\mathbf{E}_i^R)_I - (\mathbf{d}_i)_I (\mathbf{z}_T^{iR})_I \right\|_2^2, \quad (16)$$

where matrix Φ_i is a size of $l \times n_0(\varphi_i)$ with ones on $(\varphi_i(\alpha), \alpha)$ and zeros elsewhere $(\alpha = 1, 2, \dots, n_0(\varphi_i))$. φ_i denotes a group of indices for the elements of nonzero in $(\mathbf{z}_T^i)_I$, i.e., $\varphi_i = \{v \mid (\mathbf{z}_T^i)_I(v) \neq 0, 1 \leq v \leq l\}$, and $n_0(\varphi_i)$ denotes the number of nonzero in $(\mathbf{z}_T^i)_I$.

This problem can be solved directly by SVD decomposition (i.e., $\mathbf{E}_k^R = \mathbf{U}\mathbf{S}\mathbf{V}^T$). $(\mathbf{d}_i)_I$ and $(\mathbf{z}_T^{iR})_I$ are defined as the first column of $\mathbf{U}$ and the first column of $\mathbf{V}$ multiplied by $\mathbf{S}(1, 1)$, respectively.

Based on the K-SVD, approximate K-SVD (AK-SVD) is introduced to improve the dictionary by iteration and reduce the computational burden simultaneously [29]. In this case, the following two steps can be alternately iterated to obtain an approximate solution, as shown below:

$$(\mathbf{d}_i)_I = \frac{(\mathbf{E}_i^R)_I ((\mathbf{z}_T^{iR})_I)^T}{\left\| (\mathbf{E}_i^R)_I ((\mathbf{z}_T^{iR})_I)^T \right\|_2}, \quad (17)$$

$$((\mathbf{z}_T^{iR})_I)^T = ((\mathbf{E}_i^R)_I)^T (\mathbf{d}_i)_I. \quad (18)$$

3. Proposed ALIF-SBAKW

3.1. The ALIF Algorithm. In order to reduce the impact of the nonstationary and the wheelset's rotation frequency components, adaptive local iterative filtering (ALIF) is used to process signals. ALIF is a novel time-frequency analysis algorithm, which is inspired by EMD [25]. The flowchart of ALIF is shown in Figure 3.

$$\mathcal{L}_n(s)(x) = \int_{-l_n(x)}^{l_n(x)} s(x+t) \omega_n(x, t) dt, \quad (19)$$

$$SD = \frac{\|H_{i,n} - H_{i,n-1}\|_2}{\|H_{i,n-1}\|_2}, \quad SD \leq \varepsilon, \quad (20)$$

$$H_{i,n} = s_n(x) - \int_{-l_n(x)}^{l_n(x)} s_n(x+t) \omega_n(x, t) dt = s_{n+1}(x). \quad (21)$$

In this algorithm, the operator $\mathcal{L}_n(s)(x)$ denotes the moving average of the signal $s(x)$, as shown in (19). In (19), $\omega_n(x, t)$ denotes the low pass filter constructed by the solution of Fokker-Planck (FP) equations with a mask length of $2l_n(x)$. The main idea of computing mask length is to compute a multiple of the distance of subsequent local minima and maxima of $s(x)$. Obtaining a continuously varying and smooth $l_n(x)$ can be achieved by interpolating the values of the distance of the subsequent local extrema of $s(x)$ and subtracting the high frequency from the interpolated line. It can be observed in Figure 3 that the ALIF consists of two loops: the outer loop and inner loop. The outer loop mainly derives the IMFs captured by the inner

Given a signal $s(x)$;
Initial a set of IMF : IMF = {}.

Outer Loop: While the number of extrema of $s(x) > 2$:

$n = 1 \quad s_n(x) = s(x)$

Inner Loop: While the stopping criterion $SD > \varepsilon$:

① Compute the filter length $l_n(x)$

② $s_{n+1}(x) = s_n(x) - \int_{-l_n(x)}^{l_n(x)} s_n(x+t) \omega_n(x, t) dt$

③ $n = n + 1$

End While

IMF = IMF $\cup \{s_n(x)\}$

$s(x) = s(x) - s_n(x)$

End While

IMF = IMF $\cup \{s(x)\}$

FIGURE 3: The algorithm of ALIF.

loop. It determines whether the process of decomposition can stop with the number of extrema. The inner loop mainly captures a single IMF component with a stopping criterion. The stopping criterion in the inner loop usually requires $n \rightarrow \infty$, which is difficult to apply in practice. Therefore, an exact threshold ε can be set as a stopping criterion shown in (20), where $H_{i,n}$ is the n -th step of the i -th inner loop shown in (21).

The ALIF performs much better when processing the separation of the aforementioned two components. ALIF follows the iterative framework of the EMD algorithm. The moving average in ALIF is the convolution between the signal and the low pass filter. The low pass filter, constructed by the solution of FP equation, is compactly supported and is tending to zero smoothly at both ends, which ensures the nonexistence of artificial oscillations. In addition, the length $l_n(x)$ of filters in ALIF is adapted accordingly. This ensures that the nonstationary changes in signal can be captured more effectively. Therefore, ALIF is more stable under perturbation.

3.2. The Choice of Impulse-IMF. As presented in Section 1, in order to make the conducting of fault extraction more conducive, the nonstationary and the wheelset's rotation frequency components that are caused by the service conditions of the wheelset-bearing should be removed firstly. After that, an appropriate signal component, which mainly contains fault information, should be chosen as the impulse-IMF for wavelet sparse representation. According to the algorithm of ALIF, a series of IMFs for different frequency can be obtained. These IMFs are arranged in a frequency order from high to low. According to the different frequency of the signal components, the IMF1 (fault information with high frequency) is empirically chosen as the impulse-IMF. This can effectively reduce or eliminate the influence of the aforementioned two components (low frequency), which is beneficial to the feature extraction of the wheelset-bearing fault.

3.3. Proposed ALIF-SBAKW. A novel feature extraction method of wheelset-bearing fault, ALIF-SBAKW, is proposed in this paper. The flowchart of ALIF-SBAKW is shown

in Figure 4. ALIF-SBAKW mainly consists of the following steps:

Step 1: the collected vibration signal is decomposed into a series of IMFs by ALIF. The impulse-IMF containing the impulses with noise can be selected from the IMFs.

Step 2: the IWCs can be obtained by the wavelet decomposition of impulse-IMF. Generally, the number of decomposed levels can be set to 1 or 2 for the signal with a short length. The noise level σ of IWCs is estimated, and the adjusted gain c should be given. The IWCs can be segmented, with the maximal overlap, into a series of truncated-signals $(\mathbf{y}_i)_I$ whose length is b . Let $(\mathbf{Y})_I = [(\mathbf{y}_1)_I, (\mathbf{y}_2)_I, \dots, (\mathbf{y}_l)_I] \in R^{b \times l}$ and k columns ($k < l$) of $(\mathbf{Y})_I$ be randomly selected as an initial dictionary $(\mathbf{D})_I = [(\mathbf{d}_1)_I, (\mathbf{d}_2)_I, \dots, (\mathbf{d}_k)_I] \in R^{b \times k}$. The sparse coefficient matrix $(\mathbf{Z})_I = [(\mathbf{z}_1)_I, (\mathbf{z}_2)_I, \dots, (\mathbf{z}_l)_I] \in R^{k \times l}$ is initialized to zeros.

Step 3: sparse representation can be applied to the given $(\mathbf{Y})_I$, $(\mathbf{D})_I$, $(\mathbf{Z})_I$, and σ . $(\mathbf{D})_I$ and $(\mathbf{Z})_I$ can be updated by the application of SB and AK-SVD. $(\mathbf{Y})_I$ can be reconstructed by the multiplication $(\mathbf{D})_I(\mathbf{Z})_I$ after sparse representation. The IWCs is restored by tiling $(\mathbf{Y})_I$.

Step 4: in order to achieve the purposes of strengthening the IWCs and enhancing the impulse response in the original signal, the convolution between the IWCs and a typical impulse is applied. The extracted impulses of the vibration signal can be reconstructed by the wavelet reconstruction with convolutional IWCs and zero-setting CDs.

4. Simulation Validation

In order to illustrate and verify the effect of the proposed ALIF-SBAKW, a simulation validation is designed in this section. As the analysis in Section 1 showed, a simulated signal generated by bearing rotation can be constructed as

$$x(t) = I(t) + x_0(t) + \eta(t), \quad (22)$$

where $I(t)$ denotes the impulse component generated by the fault of bearing. It can be expressed by (23), where $u(t)$ is a unit step function, J is the total number of generated impulses, Am_j is the amplitude of the j -th impulse, B denotes the structure damper coefficient, f_r denotes the resonance frequency, and T_g , whose reciprocal is the fault characteristic frequency ($f_g = T_g^{-1}$), is the time interval between the two adjacent impulses. The simulated parameters of (23) are shown in Table 1. The time-domain waveform of $I(t)$ is shown in Figure 5(a):

$$I(t) = \sum_{j=1}^J Am_j e^{-B(t-jT_g)} \cos(2\pi f_r(t-jT_g)) u(t-jT_g). \quad (23)$$

The simulated $x_0(t)$ denotes some unknown interferences with nonstationary component which are generated

during the process of measurement. It can be described by a cosine function and its corresponding modulation function. Here, $x_0(t)$ is directly written out, as shown in (24). In addition, $\eta(t)$ denotes noise component of $x(t)$, which can be described by Gaussian random noise, whose standard deviation σ_0 can be set to 1. The power ratio between the impulse component $I(t)$ and the total noisy simulated signal $x(t)$ is -18.7008 dB:

$$\begin{aligned} x_0(t) = & 4(t-0.5)^2 + (2(t-0.5)^2 + 0.2) \\ & \times \sin((20\pi + 0.2 \cos(40\pi t))t) \\ & + 3.6 \cos(160\pi t) \\ & + 3.5 \left(1 + \cos\left(16\pi t + \frac{\pi}{6}\right)\right) \cos\left(480\pi t + \frac{\pi}{3}\right). \end{aligned} \quad (24)$$

When the sampling frequency and sampling time are set to 10000 Hz and 1 s, respectively, a simulated signal $x(t)$, as shown in Figure 5(b), can be obtained. Figure 5(c) is the Hilbert envelope spectrum of $x(t)$. It can be observed from Figure 5(b) that the impulses have been completely covered because of the interference components. The features of bearing fault cannot be identified in time domain. In addition, the fault characteristic frequency $f_g = T_g^{-1}$ also cannot be detected in its corresponding Hilbert envelope spectrum.

4.1. The Setting Rule of Adjusted Gain c . In order to obtain a good performance of fault extraction in ALIF-SBAKW, the selection of suitable adjusted gain c is very crucial. If c is set too small, more wrong impulses and noises will be extracted in the results. Instead, if c is set too large, the number of fault impulses will be extracted insufficiently. In the fault diagnosis of bearing, the extracted performance of fault characteristic frequency and its harmonics in the Hilbert envelope spectrum is an important evaluation. Envelope spectrum kurtosis (ESK) is an effective measure index of extracted fault features in the Hilbert envelope spectrum. A higher value of ESK implies clearer fault features with larger amplitudes in their envelope spectrum and a larger number of harmonics for fault characteristic frequency [2]. Therefore, in order to select c more reasonably, the ESK of different extracted signals with different c in a particular range is calculated. The adjusted gain c with the highest ESK is selected as the most suitable parameter.

4.2. The Choice of Wavelet Base. It should be noted that, except the adjusted gain c , the choice of wavelet base also should be reasonable to obtain a good performance of wavelet decomposition and reconstruction. Daubechies wavelet is compact support and it has a wonderful regularity. It can not only retain the peak feature of the impulse, but also obtain a smoothing bearing fault signal. Therefore, it is very suitable for wavelet transform of bearing fault signal [30]. In addition, the choice of the support length of wavelet also has an impact on the analytical results. In order to avoid the boundary problem and low order vanishing moments

FIGURE 4: The flowchart of ALIF-SBAKW.

TABLE 1: Parameters of signal $I(t)$.

Am_j	T_g^{-1} (Hz)	B	f_r (Hz)
4	46	800	1800

caused by the overlong and short support length of wavelet, respectively, the filter length 8 (i.e., Daubechies 8-tap wavelet) is selected for wavelet decomposition in this paper.

4.3. Simulation Results. According to the flowchart of ALIF-SBAKW, a series of crucial calculated parameters are given. The threshold ε of stopping criterion in ALIF is set to 0.001. The decomposed level of wavelet decomposition is 1. The chosen wavelet basis is Daubechies 8-tap wavelet. The adjusted gain c of noise standard deviation is determined as 4.2. The size of the dictionary is set to 10×40 , which means that the length of segmented signal is 10. The parameters, ρ and λ in SB, are set to 100 and 1000, respectively. The chosen impulse-IMF is IMF1. Finally, the simulated analysis of (22) is implemented based on the procedure of the ALIF-SBAKW method. The decomposed IMFs of (22) by ALIF are shown in Figure 6. The simulated result of feature extraction of bearing fault is shown in Figure 7(a) and the corresponding Hilbert envelope spectrum is shown in Figure 7(b). It can be observed from the results that the fault features are clearly

detected in Figure 7(a), in which the interference components are totally eliminated. The fault characteristic frequency f_g and its harmonics are explicitly illustrated in Figure 7(b). The simulated results show that the fault features are sufficiently extracted by the proposed method no matter whether in time domain or in the Hilbert envelope spectrum.

4.4. Performance Comparison. In order to further illustrate the advancement of proposed method, two comparative methods are applied to analyze the same simulated signal. Firstly, wavelet sparse representation, i.e., SBAKW, is used to process the simulated signal directly. The adjusted gain c of noise standard deviation is set to 24, and the other parameters are determined as above. The simulated result using SBAKW is shown in Figure 8(a), and its Hilbert envelope spectrum is shown in Figure 8(b). Secondly, the OMP-KSVD algorithm, a well-known sparse representation algorithm, is applied to handle the IMFs decomposed by ALIF, i.e., ALIF-OMPK. In general, there are two different iteration stopping criteria in OMP algorithm: target sparsity and error goal. The size of the dictionary is set to 64×256 in ALIF-OMPK. The number of iterations is set to 10, and the chosen impulse-IMF also is IMF1. When the target sparsity is adopted as the stopping criterion and it is set to 2, the simulated result is shown in Figure 9(a) and its Hilbert

FIGURE 5: The simulated signal: (a) impulse components $I(t)$, (b) simulated signal $x(t)$, and (c) Hilbert envelope spectrum of $x(t)$.

envelope spectrum is shown in Figure 9(b). When the error goal is adopted as the stopping criterion and it is set to 54.7, the simulated result is shown in Figure 10(a) and its Hilbert envelope spectrum is shown in Figure 10(b).

Making a comparison between SBAKW and ALIF-SBAKW, features of interferences are extracted by using SBAKW in Figure 8(a). This is because the sparse representation is more likely to detect the components having stronger energy. In this simulated case, the energy of nonstationary component is much stronger than that of fault features. The extracted features obviously are not fault features. The information reflected in Hilbert envelope spectrum also cannot extract the fault characteristic frequency f_g effectively either. Therefore, SBAKW cannot detect fault features of bearing directly.

The results extracted by ALIF-OMPK under two different iteration stopping criteria can also identify the information of fault characteristic, as shown in Figures 9 and 10. However, the extracted results are not desired compared with the results extracted by the ALIF-SBAKW. The ALIF-SBAKW method significantly removes noise between the two impulses, which has better performance in denoising than that of the ALIF-OMPK. Additionally, although the fault characteristic frequency f_g and its harmonics can also be identified in Hilbert envelope spectrum by using ALIF-OMPK, the amplitude of f_g is weaker than that extracted by the ALIF-SBAKW. The number of harmonics extracted by the ALIF-SBAKW is also much more fruitful. In order to further conduct the comparison, the value of each ESK is calculated as a comparative indicator. The values of ESK in Figures 7(b), 9(b), and 10(b) are 511.7, 336.4, and 477.8, respectively. This implies that the results, extracted by the ALIF-SBAKW, are superior. Therefore, compared with the

ALIF-OMPK, the proposed ALIF-SBAKW method can extract more fruitful information of fault no matter whether in time domain or in the Hilbert envelope spectrum.

5. Experimental Validation

In order to further validate the effect of the proposed ALIF-SBAKW, the experimental data of wheelset-bearing fault has been obtained through the testing rig shown in Figure 11(a). The experiment was conducted by the project between Southwest Jiaotong University and CRRC Corporation. The axle box bearing running on the testing rig is from China Railway High-speed (CRH) vehicle, and the double-row tapered roller bearing is adopted for CRH vehicle axle box. The testing rig consists of a motor, a loading device, a pair of driving wheels, a testing wheelset, and an axle box. The testing wheelset, which is driven by the driving wheels in the bottom, is supported by the axle box bearing. The driving power is delivered by the motor, and it can be conveyed to the driving wheels through the rubber belts. The accelerometer for collecting vibration signals is mounted on the axle box, as shown in Figure 11(b). In Figures 11(c) and 11(d), two typical bearing faults are introduced into the experiment: outer-race fault and roller fault. The fault characteristic frequency of these two typical faults can be calculated using (25) and (26), respectively. The bearing parameters d_B , d_P , N_Z , and ϕ denote roller-ball diameter, pitch diameter, the number of balls, and the contact angle of balls, respectively. The values of bearing parameters are listed in Table 2; f_{ro} denotes the rotation frequency. When the speed is 100 km/h, the corresponding rotation frequency is 10.3 Hz. Therefore, it can be easily calculated that the fault characteristic frequency of outer-race fault f_{BPFO} and roller

FIGURE 6: The decomposed IMFs using the ALIF: (a–e) IMF1–IMF5.

FIGURE 7: The simulated results using the ALIF-SBAKW method: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

fault f_{BSF} are 83.23 Hz and 33.69 Hz, respectively. The accelerometer is installed on the axle box, and the fault signal is collected at a sampling rate of 10 kHz.

In this section, the proposed ALIF-SBAKW method is used to analyze the collected vibration signal. In order to

further validate the effect of proposed ALIF-SBAKW and highlight its superiority, four different comparative methods, namely, SBAKW, ALIF-OMPK (using target sparsity), EWT, and fast kurtogram [7], are applied to process the analyzed signal:

FIGURE 8: The simulated results using the SBAKW: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

FIGURE 9: The simulated results using ALIF-OMPK with target sparsity: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

FIGURE 10: The simulated results using ALIF-OMPK with error goal: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

$$f_{\text{BPFO}} = \frac{N_z}{2} \left(1 - \frac{d_B}{d_P} \cos \phi \right) f_{\text{ro}}, \quad (25)$$

$$f_{\text{BSF}} = \frac{d_P}{2d_B} \left(1 - \left(\frac{d_B}{d_P} \cos \phi \right)^2 \right) f_{\text{ro}}. \quad (26)$$

5.1. Outer-Race Fault Experiment. In the experiment of outer-race fault, the collected signal is shown in Figure 12(a) and its Hilbert envelope spectrum is shown in Figure 12(b). It can be seen that the fault characteristic frequency f_{BPFO}

and its harmonics cannot be discovered clearly due to the existence of power-line interference and noises.

The proposed ALIF-SBAKW method is used to analyze the collected signal in Figure 12(a). The calculated parameters, used in ALIF-SBAKW, are mostly identical with those set in the simulation validation. The adjusted gain c and the segmented length b are set to 3 and 10, respectively. The number of IMFs, decomposed by ALIF, is 6, as shown in Figures 13(a)–13(d) (only the first 4 IMFs are shown). Similarly, the IMF1 is still set as the impulse-IMF. The results of fault extraction are shown in Figures 13(e) and 13(f). It can be seen that the impulses in time domain are indeed extracted. In addition, the fault characteristic frequency of

FIGURE 11: Experimental device: (a) testing rig, (b) accelerometer, (c) outer-race fault, and (d) roller fault.

TABLE 2: Parameters of testing bearing.

d_B (mm)	d_p (mm)	N_Z	ϕ (rad)
26.9	180	19	0.1571

FIGURE 12: The collected outer-race fault signal and its Hilbert envelope spectrum: (a) time-domain waveform and (b) Hilbert envelope spectrum of signal in (a).

outer-race fault f_{BPF0} and its harmonics are clearly displayed in Hilbert envelope spectrum, which indicates that there certainly exists fault on the surface of the outer race.

In order to further validate the effect of proposed ALIF-SBAKW, four comparative methods are conducted. In SBAKW, the adjusted gain c of noise standard deviation is set to 5.5, and the other parameters are identical with those set in ALIF-SBAKW. The extracted results using SBAKW are

shown in Figure 14. In ALIF-OMPK, the size of the dictionary is set to 128×512 , and the sparsity is 2. The number of iterations is set to 10. The extracted signal using ALIF-OMPK is shown in Figure 15(a), and its Hilbert envelope spectrum is shown in Figure 15(b). In EWT, the maximum number of segmented bands is set to 14. In the results obtained by the EWT, most subband signals are unable to contain useful information due to the segment of narrow

FIGURE 13: Results obtained by the proposed ALIF-SBAKW: (a–d) IMF1-IMF4, (e) extracted signal, and (f) Hilbert envelope spectrum of signal in (e).

FIGURE 14: Results obtained by SBAKW: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

frequency bands. Among all bands, only the 13th subband signal can extract some feature information. The analytical results are shown in Figure 16. In fast kurtogram, the highest level of decomposition is set to 6. The kurtogram is obtained as shown in Figure 17(a). The centre frequency and optimal

level of filter are set to 3945.3125 and 6, respectively. The extracted results are shown in Figures 17(b) and 17(c).

Making comparison between SBAKW and ALIF-SBAKW, although some kinds of vibration features can be extracted by using SBAKW, they mainly derive from the

FIGURE 15: Results obtained by ALIF-OMPK: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

FIGURE 16: Results obtained by EWT: (a) detected boundaries of Fourier spectrum, (b) the 13th subband signal, and (c) Hilbert envelope spectrum of signal in (b).

energy of power-line interference. The characteristic frequency of power-line interference f_{PLI} can be detected in Hilbert envelope spectrum, whereas the f_{BPFO} is almost submerged in other frequency components and its harmonics cannot be extracted. This shows that the interference seriously affects the extraction of correct features for impulses. Compared with ALIF-OMPK, EWT, and fast kurtogram, the extracted signals are not purified, and the fault impulses are nearly drowned in the noise component in time domain. Furthermore, the whole amplitude of extracted signals, analyzed by EWT and fast kurtogram, is greatly reduced compared with the signal extracted by the ALIF-SBAKW, which means that the extraction of fault is affected. Compared with the Hilbert envelope spectrum, the most significant difference is that the amplitude of f_{BPFO} and its harmonics, extracted by ALIF-OMPK, EWT, and fast kurtogram, are generally weaker than those extracted by the

ALIF-SBAKW. As for the ALIF-OMPK, although the number of its harmonics increases, the f_{BPFO} is inconspicuous. For EWT, the number of harmonics of f_{BPFO} is far less than the number of harmonics in Figure 13(f). By contrast, the f_{BPFO} of fast kurtogram and its harmonics cannot be discovered directly in Figure 17(c). Therefore, the proposed ALIF-SBAKW can be used to analyze the fault extraction of outer race effectively, and it performs better than the other four comparative methods.

5.2. Roller Fault Experiment. In the roller fault experiment, the collected signal is shown in Figure 18(a) and its Hilbert envelope spectrum is shown in Figure 18(b). It should be noted that the even harmonics of f_{BSF} are often dominant in Hilbert envelope spectrum. This is because the defect on rolling element surface impacts both the inner race and

FIGURE 17: Results obtained by fast kurtogram: (a) kurtogram, (b) extracted signal, and (c) Hilbert envelope spectrum of signal in (b).

FIGURE 18: The collected roller fault signal and its Hilbert envelope spectrum: (a) time-domain waveform and (b) Hilbert envelope spectrum of signal in (a).

outer race, which excites two impulses and results in two shocks per basic period [10, 31]. Therefore, the $2f_{BSF}$ and its harmonics should be used as the diagnostic indices for roller fault. It can be observed from Figure 18(a) that the impulses of roller fault in time-domain waveform are almost overwhelmed. Additionally, although $2f_{BSF}$ can be reluctantly discovered, its harmonics are not purified. They are almost overwhelmed in complex frequency components and thus cannot be detected.

The collected signal is firstly processed by the proposed ALIF-SBAKW method. Similarly, the adjusted gain c and the

segmented length b are set to 4 and 10, respectively. The IMFs of collected signal, decomposed by ALIF, are shown in Figures 19(a)–19(d) (similarly the first 4 IMFs are shown). The result of fault extraction is shown in Figure 19(e), and the corresponding Hilbert envelope spectrum is shown in Figure 19(f). It can be observed that the impulses, caused by the roller fault, are directly extracted. The double fault characteristic frequency of roller $2f_{BSF}$ and its harmonics are clearly discovered in Hilbert envelope spectrum.

Similarly, four comparative methods are used to process the collected signal. In SBAKW, the adjusted gain c of noise

FIGURE 19: Results obtained by the proposed ALIF-SBAKW: (a–d) IMF1-IMF4, (e) extracted impulses, and (f) Hilbert envelope spectrum of signal in (e).

standard deviation is set to 6, and the other parameters remain unchanged as those set in ALIF-SBAKW. The analytical results are shown in Figure 20. In ALIF-OMPK, the related parameters are consistent with the parameter settings of ALIF-OMPK in Section 5.1. The extracted results are shown in Figure 21. In EWT, the detected boundaries of Fourier spectrum are shown in Figure 22(a). According to the results, only the 17th subband signal is able to extract some feature information, as shown in Figures 22(b) and 22(c). In fast kurtogram, the centre frequency and the optimal level of filter are set to 1875 and 5, respectively. The analytical results are shown in Figure 23.

The reason why the double fault characteristic frequency $2f_{\text{BSF}}$ can be detected in Hilbert envelope spectrum of the original collected signal is that the original energy of fault impulses in the roller fault experimental data is extraordinarily strong. However, the quality of results, extracted by four comparative methods, is still not as good as that extracted by ALIF-SBAKW. Compared with SBAKW, although some kinds of vibration features can be extracted by

using SBAKW, other different features are simultaneously extracted compared with the signal in Figure 19(e). This is due to the existence of power-line interference and noises, which lead to the inconspicuousness of $2f_{\text{BSF}}$ directly in Figure 20(b). Making comparison among the left three methods, we find that the extracted signals can only observe the peaks of impulses and there still exist noises at the interval between the adjacent impulses compared with signal in Figure 19(e).

According to the comparison in Hilbert envelope spectrum, it can be observed from Figures 20(b) and 22(c) that the frequency, extracted by SBAKW and EWT, is nearly unable to detect the harmonics of $2f_{\text{BSF}}$ directly. In Figures 21(b) and 23(c), although the amplitude of partly harmonics, extracted by the ALIF-OMPK and fast kurtogram, is not the most prominent, a certain number of harmonics still can be identified. In order to further highlight the superiority of the ALIF-SBAKW, some feature indicators are introduced and calculated to make a more direct comparison between ALIF-SBAKW, ALIF-OMPK,

FIGURE 20: Results obtained by SBAKW: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

FIGURE 21: Results obtained by ALIF-OMPK: (a) extracted signal and (b) Hilbert envelope spectrum of signal in (a).

FIGURE 22: Results obtained by EWT: (a) detected boundaries of Fourier spectrum, (b) the 17th subband signal, and (c) Hilbert envelope spectrum of signal in (b).

FIGURE 23: Results obtained by fast kurtogram: (a) kurtogram, (b) extracted signal, and (c) Hilbert envelope spectrum of signal in (b).

TABLE 3: Comparison among the three methods.

Indicator	Method		
	ALIF-SBAKW	ALIF-OMPK	Fast kurtogram
Number of harmonics	9	8	5
CF	14.47	11.08	5.62
IF	57.77	17.29	7.58
Kurtosis	59.26	20.44	5.57
ESK	22.43	20.00	19.88

and fast kurtogram. The introduced feature indicators include crest factor (CF), impulse factor (IF), kurtosis [32], and envelope spectrum kurtosis (ESK). The related parameters and calculated values are shown in Table 3.

Theoretically, when the values of CF, IF, and kurtosis are higher, the features of impulses extracted in time domain are relatively stronger. In addition, the ESK mainly reflects the richness of fault information in Hilbert envelope spectrum. As shown in Table 3, obviously each of the values of the indicator, calculated by the results of the ALIF-SBAKW, is higher than those calculated by the results of the other two methods. Consequently, the comparative results indicate that the ALIF-SBAKW is still superior to the ALIF-OMPK and fast kurtogram. Overall, it can be concluded that, compared with SBAKW, ALIF-OMPK, EWT, and fast

kurtogram, the proposed ALIF-SBAKW has better performance in some degree.

6. Conclusions

The fault diagnosis of wheelset-bearing has great significance to the safety of high-speed train. Sparse representation is an advanced method for bearing fault extraction. However, it is hard for the traditional sparse representation to conduct fault extraction under severe service conditions, especially under the complicated track-vehicle system. If the energy of interference is stronger than the energy of fault features, the interference component, instead of the fault features, will be detected and extracted. Therefore, the ALIF-SBAKW is proposed in this paper. There are two reasons why this new method can solve this problem. On the one hand, the ALIF can effectively reduce or eliminate the nonstationary and the wheelset's rotation frequency components (caused by severe service conditions of the high-speed train), which is conducive to realizing fault extraction. On the other hand, the wavelet sparse representation can deeply find the intrinsic features of signal and extract the wheelset-bearing fault. The ALIF-SBAKW method is validated by simulated and experimental signals. The results show that the ALIF-SBAKW method is extraordinarily suitable for the fault feature extraction of wheelset-bearing signals, especially compared with SBAKW, ALIF-OMPK, EWT, and fast kurtogram in

experimental validation. To some extent, this novel method can effectively complete the fault diagnosis of wheelset-bearing.

Finally, although the ALIF-SBAKW method can effectively extract the fault feature, this method cannot be effectively applied to the separation of multiple faults now. Therefore, further research should be made to solve the considered problems. In addition, the fault extraction method, proposed by Qin et al. in [23], is an excellent technique. In the future research of separation for multiple

faults, a more comprehensive comparison and research with the method in [23] will be conducted in terms of the effect of component separation and fault extraction.

Appendix

The derivation process of simplification from (7) to (8) and (9) is presented briefly in this part. According to the principle of SB [28], (7) is solved by

$$((\mathbf{z})_I^{r+1}, (\mathbf{f})_I^{r+1}) = \min_{\mathbf{z}, \mathbf{f}} F((\mathbf{z})_I, (\mathbf{f})_I) - F((\mathbf{z})_I^r, (\mathbf{f})_I^r) - \langle \mathbf{p}_z^r, (\mathbf{z})_I - (\mathbf{z})_I^r \rangle - \langle \mathbf{p}_f^r, (\mathbf{f})_I - (\mathbf{f})_I^r \rangle + \frac{\lambda}{2} \|(\mathbf{f})_I - (\mathbf{z})_I\|_2^2, \quad (\text{A.1})$$

$$\mathbf{p}_z^{r+1} = \mathbf{p}_z^r - \lambda((\mathbf{z})_I^{r+1} - (\mathbf{f})_I^{r+1}), \quad (\text{A.2})$$

$$\mathbf{p}_f^{r+1} = \mathbf{p}_f^r - \lambda((\mathbf{f})_I^{r+1} - (\mathbf{z})_I^{r+1}), \quad (\text{A.3})$$

where $\mathbf{p}_z^r \in R^k$ and $\mathbf{p}_f^r \in R^k$ are in the subgradient of F at $(\mathbf{z})_I^r$ and $(\mathbf{f})_I^r$, respectively. $\langle \mathbf{o}_1, \mathbf{o}_2 \rangle$ denotes the inner

product between $\mathbf{o}_1$ and $\mathbf{o}_2$. Defining $(\mathbf{z})_I^0 = (\mathbf{f})_I^0 = \mathbf{p}_z^0 = \mathbf{p}_f^0 = \mathbf{0}$, (A.2)–(A.3) can be written as

$$\mathbf{p}_z^{r+1} = \mathbf{p}_z^0 - \lambda((\mathbf{z})_I^1 - (\mathbf{f})_I^1) - \lambda((\mathbf{z})_I^2 - (\mathbf{f})_I^2) \cdots - \lambda((\mathbf{z})_I^{r+1} - (\mathbf{f})_I^{r+1}) = -\lambda \sum_{i=1}^{r+1} ((\mathbf{z})_I^i - (\mathbf{f})_I^i), \quad (\text{A.4})$$

$$\mathbf{p}_f^{r+1} = \mathbf{p}_f^0 - \lambda((\mathbf{f})_I^1 - (\mathbf{z})_I^1) - \lambda((\mathbf{f})_I^2 - (\mathbf{z})_I^2) \cdots - \lambda((\mathbf{f})_I^{r+1} - (\mathbf{z})_I^{r+1}) = -\lambda \sum_{i=1}^{r+1} ((\mathbf{f})_I^i - (\mathbf{z})_I^i). \quad (\text{A.5})$$

Then, both of the above equations can be replaced by $(\mathbf{w})_I^{r+1} = \sum_{i=1}^{r+1} ((\mathbf{z})_I^i - (\mathbf{f})_I^i)$ and are simplified as the same

iterative equation, i.e., (9) in Section 2.2. Finally, the simplification can be simplified as

$$\begin{aligned} ((\mathbf{z})_I^{r+1}, (\mathbf{f})_I^{r+1}) &= \min_{\mathbf{z}, \mathbf{f}} F((\mathbf{z})_I, (\mathbf{f})_I) - F((\mathbf{z})_I^r, (\mathbf{f})_I^r) - \langle \mathbf{p}_z^r, (\mathbf{z})_I - (\mathbf{z})_I^r \rangle - \langle \mathbf{p}_f^r, (\mathbf{f})_I - (\mathbf{f})_I^r \rangle + \frac{\lambda}{2} \|(\mathbf{f})_I - (\mathbf{z})_I\|_2^2, \\ &= \min_{\mathbf{z}, \mathbf{f}} F((\mathbf{z})_I, (\mathbf{f})_I) - F((\mathbf{z})_I^r, (\mathbf{f})_I^r) + \lambda \langle (\mathbf{w})_I^r, (\mathbf{z})_I - (\mathbf{z})_I^r \rangle - \lambda \langle (\mathbf{w})_I^r, (\mathbf{f})_I - (\mathbf{f})_I^r \rangle + \frac{\lambda}{2} \|(\mathbf{f})_I - (\mathbf{z})_I\|_2^2, \end{aligned} \quad (\text{A.6})$$

$$= \min_{\mathbf{z}, \mathbf{f}} F((\mathbf{z})_I, (\mathbf{f})_I) - \lambda \langle (\mathbf{w})_I^r, (\mathbf{f})_I - (\mathbf{z})_I \rangle + \frac{\lambda}{2} \|(\mathbf{f})_I - (\mathbf{z})_I\|_2^2 + C_1,$$

$$= \min_{\mathbf{z}, \mathbf{f}} F((\mathbf{z})_I, (\mathbf{f})_I) + \frac{\lambda}{2} \|(\mathbf{f})_I - (\mathbf{z})_I - (\mathbf{w})_I^r\|_2^2 + C_2,$$

where C_1 and C_2 denote relevant constants involved in simplification. In fact, these constants can be ignored in minimal optimization. Therefore, (8) is obtained by ignoring C_2 .

Data Availability

The experimental data are based on the CRRC project and are confidential.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51905453) and the China Postdoctoral Science Foundation (no. 2019M663899XB).

References

- [1] J. Ding, W. Zhao, B. Miao, and J. Lin, "Adaptive sparse representation based on circular-structure dictionary learning and its application in wheelset-bearing fault detection," *Mechanical Systems and Signal Processing*, vol. 111, pp. 399–422, 2018.
- [2] J. Ding, J. Zhou, and Y. Yin, "Fault detection and diagnosis of a wheelset-bearing system using a multi-Q-factor and multi-level tunable Q-factor wavelet transform," *Measurement*, vol. 143, pp. 112–124, 2019.
- [3] W. Guo and P. W. Tse, "A novel signal compression method based on optimal ensemble empirical mode decomposition for bearing vibration signals," *Journal of Sound and Vibration*, vol. 332, no. 2, pp. 423–441, 2013.
- [4] Z. Feng, H. Ma, and M. J. Zuo, "Amplitude and frequency demodulation analysis for fault diagnosis of planet bearings," *Journal of Sound and Vibration*, vol. 382, pp. 395–412, 2016.
- [5] C. Y. Yang and T. Y. Wu, "Diagnostics of gear deterioration using EEMD approach and PCA process," *Measurement*, vol. 61, pp. 75–87, 2015.
- [6] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: a suitable tool for biomedical signal processing," *Biomedical Signal Processing and Control*, vol. 14, pp. 19–29, 2014.
- [7] Z. Liu, J. Ding, J. Lin, and Y. Huang, "A rolling bearing fault diagnosis-optimized scale-space representation for the empirical wavelet transform," *Shock and Vibration*, vol. 2018, Article ID 2749689, 22 pages, 2018.
- [8] Z. Qiao, Y. Liu, and Y. Liao, "An improved method of EWT and its application in rolling bearings fault diagnosis," *Shock and Vibration*, vol. 2020, Article ID 4973941, 13 pages, 2020.
- [9] C. Yi, Y. Lv, and Z. Dang, "A fault diagnosis scheme for rolling bearing based on particle swarm optimization in variational mode decomposition," *Shock and Vibration*, vol. 2016, Article ID 9372691, 10 pages, 2016.
- [10] Y. Huang, J. Lin, Z. Liu, and W. Wu, "A modified scale-space guiding variational mode decomposition for high-speed railway bearing fault diagnosis," *Journal of Sound and Vibration*, vol. 444, pp. 216–234, 2019.
- [11] Z. Wang, J. Zhou, J. Wang et al., "A novel fault diagnosis method of gearbox based on maximum kurtosis spectral entropy deconvolution," *IEEE Access*, vol. 7, pp. 29520–29532, 2019.
- [12] Z. Wang, W. Du, J. Wang et al., "Research and application of improved adaptive MOMEDA fault diagnosis method," *Measurement*, vol. 140, pp. 63–75, 2019.
- [13] Z. Wang, J. Wang, W. Cai et al., "Application of an improved ensemble local mean decomposition method for gearbox composite fault diagnosis," *Complexity*, vol. 2019, Article ID 1564243, 17 pages, 2019.
- [14] X. Zhang, Z. Song, D. Li, W. Zhang, Z. Zhao, and Y. Chen, "Fault diagnosis for reducer via improved LMD and SVM-RFE-MRMR," *Shock and Vibration*, vol. 2018, Article ID 4526970, pp. 1–13, 2018.
- [15] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Transactions on Instrumentation and Measurement*, p. 1, 2020.
- [16] H. Zhao, J. Zheng, W. Deng, and Y. Song, "Semi-supervised broad learning system based on manifold regularization and broad network," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 3, pp. 983–994, 2020.
- [17] X. Chen, Z. Du, J. Li, X. Li, and H. Zhang, "Compressed sensing based on dictionary learning for extracting impulse components," *Signal Processing*, vol. 96, pp. 94–109, 2014.
- [18] H. Tang, J. Chen, and G. Dong, "Sparse representation based latent components analysis for machinery weak fault detection," *Mechanical Systems and Signal Processing*, vol. 46, no. 2, pp. 373–388, 2014.
- [19] F. Hou, J. Chen, and G. Dong, "Weak fault feature extraction of rolling bearings based on globally optimized sparse coding and approximate SVD," *Mechanical Systems and Signal Processing*, vol. 111, pp. 234–250, 2018.
- [20] R.-B. Sun, Z.-B. Yang, Z. Zhai, and X.-F. Chen, "Sparse representation based on parametric impulsive dictionary design for bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 122, pp. 737–753, 2019.
- [21] J. Ding, "Fault detection of a wheelset bearing in a high-speed train using the shock-response convolutional sparse-coding technique," *Measurement*, vol. 117, pp. 108–124, 2018.
- [22] Y. Qin, "A new family of model-based impulsive wavelets and their sparse representation for rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2716–2726, 2018.
- [23] Y. Qin, J. Zou, B. Tang, Y. Wang, and H. Chen, "Transient feature extraction by the improved orthogonal matching pursuit and K-SVD algorithm with adaptive transient dictionary," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 215–227, 2020.
- [24] Z.-H. Zhang, J.-M. Ding, and J.-H. Lin, "Train wheelset bearing multifault impulsive component separation using hierarchical shift-invariant dictionary learning," *Shock and Vibration*, vol. 2019, Article ID 5697137, 14 pages, 2019.
- [25] A. Ciccone, J. Liu, and H. Zhou, "Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis," *Applied and Computational Harmonic Analysis*, vol. 41, no. 2, pp. 384–411, 2016.
- [26] B. Ophir, M. Lustig, and M. Elad, "Multi-scale dictionary learning using wavelets," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 1014–1024, 2011.
- [27] V. P. Gopi, P. Palanisamy, K. A. Wahid, and P. Babyn, "MR image reconstruction based on iterative split bregman algorithm and nonlocal total variation," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 985819, 16 pages, 2013.
- [28] T. Goldstein and S. Osher, "The split bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [29] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," 2008.
- [30] J. P. Patel and S. H. Upadhyay, "Comparison between artificial neural network and support vector method for a fault diagnostics in rolling element bearings," *Procedia Engineering*, vol. 144, pp. 390–397, 2016.

- [31] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics-A tutorial," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485–520, 2011.
- [32] A. Moosavian, G. Najafi, B. Ghobadian, M. Mirsalim, S. M. Jafari, and P. Sharghi, "Piston scuffing fault and its identification in an IC engine by vibration analysis," *Applied Acoustics*, vol. 102, pp. 40–48, 2016.