Research Article

Fault Detection and Behavior Analysis of Wheelset Bearing Using Adaptive Convolutional Sparse Coding Technique Combined with Bandwidth Optimization

Liu He 1, Cai Yi 1, Jianhui Lin 1, and Andy C.C. Tan 2

1 State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu, Sichuan, China
2 Tunku Abdul Rahman University, Jalan Bandar Sungai Long, Cheras, Kajang 43000, Selangor, Malaysia

Correspondence should be addressed to Cai Yi; yicai@swjtu.edu.cn

Received 7 May 2020; Revised 11 September 2020; Accepted 27 October 2020; Published 18 November 2020

Wheelset bearing is a critical and easily damaged component of a high-speed train. Wheelset bearing fault diagnosis is of great significance to ensure safe operation of high-speed trains and realize intelligent operation and maintenance. The convolutional sparse coding technique based on the dictionary learning algorithm (CSCT-DLA) provides an effective algorithm framework for extracting the impulses caused by bearing defect. However, dictionary learning is easily affected by foundation vibration and harmonic interference and cannot learn the key structure related to fault impulses. At the same time, the detection performance of fault impulse heavily depends on the selection of parameters in this approach. Union of convolutional dictionary learning algorithm (UC-DLA) is an efficient algorithm in CSCT-DLA. In this paper, UC-DLA is introduced and improved for wheelset bearing fault detection. Finally, a novel bearing fault detection method, adaptive UC-DLA combined with bandwidth optimization (AUC-DLA-BO), is proposed. The mathematical formulation of AUC-DLA-BO is a sort of constrained optimization problem, which can overcome foundation vibration and harmonic interference and adaptively determine parameters related to UC-DLA. The proposed method can detect the fault resonance band adaptively, eliminate the noise with the same frequency band as the fault resonance band, and highlight the bearing fault impulses. Simulated signals and bench tests are used to verify the effectiveness of the proposed method. The results show that AUC-DLA-BO can effectively detect bearing faults and realize the refined analysis of fault behavior.

1. Introduction

The high-speed train is a typical representative of modern advanced transportation tools and has been paid attention to by many countries (such as Japan, Germany, Spain, China, etc.). High-speed train realizes high-speed movement by driving wheelset rotation by the traction motor. The power system of a high-speed train is a typical rotating mechanical system. Wheelset bearing is a critical and easily damaged mechanical component in the power system of the high-speed train. It is used to support the rotating system of high-speed trains to resist static and dynamic external forces and to transfer power. Unexpected wheelset bearing faults can lead to reduced efficiency of a power system and equipment shutdown and damage and in some cases may lead to safety accidents [1–3]. Therefore, a powerful wheelset bearing fault diagnosis technique is very important to avoid catastrophic failures in high-speed trains.

The strong nonlinear and nonstationary vibration signal containing impulses will be generated when a local defect appears on wheelset bearing [4, 5]. The generated impulses usually exhibit modulation phenomenon [6, 7] and are often submerged in strong background noise and irrelevant interferences [8–10]. Thus, it is a challenge to extract impulses from the vibration signal for bearing fault diagnosis.

In recent years, the compressed sensing (CS) technique has attracted widespread attention and attention made remarkable achievements in information theory, image processing, pattern recognition, and biomedical engineering [11–13]. CS is also used for bearing fault feature extraction.
recently with its features of flexibility, sparsity, and super-resolution [10, 12, 14, 15]. The essence of processing bearing vibration signals with CS is to obtain sparse representation of the vibration signal in the redundant dictionary. The redundancy dictionary design is the key issue of sparse representation [14, 16], which mainly includes analytic dictionary and learning dictionary [10, 17–20]. Analytic dictionary needs to be constructed with existing predefined dictionary in advance. In most cases, the impulses related to bearing fault have different vibration waveform features. Even for the same bearing, the vibration waveform features excited by different fault modes are different. There is always inevitable dissimilarity between atoms of analytic dictionary and real vibration signal. So, sparse representation based on analytic dictionary might fail to capture fault features in the target signal. Learning dictionary (including regular dictionary learning and shift-invariant dictionary learning (SIDL)) are learned directly from target signals by machine learning techniques and could match the local structures of signals [19]. Many advanced and effective dictionary learning methods such as K-SVD [14], shift-invariant sparse coding-based feature-sign search (SISC-FSS) [17], shift-invariant K-SVD (SI-K-SVD) [21], and convolutional sparse coding based on ADMM (CSC-ADMM) [22] have been proposed and successfully applied to the learning of bearing fault feature structures. A recent contribution is that SISC-FSS was introduced into mechanical fault diagnosis for adaptive feature extraction of bearing fault in [23]. SI-K-SVD was used to identify the double-impact characteristics of rolling bearing vibration signal in [19]. CSC-ADMM was introduced to extract impulse responses and time location coefficients in [24, 25]. Although sparse representations based on learning dictionary have been applied successfully in the field of fault detection, there are still two major shortcomings which are difficult to solve in practical engineering application:

1. In the process of dictionary learning and sparse coding, the number of dictionaries, length of dictionary atom, and target sparsity seriously affect the performance of fault diagnosis, and it is hard to interpret the physical meaning of the learned waveform in the dictionary.

2. Dictionary learning is very susceptible to strong interference signals including foundation vibration and harmonic interference [26]. If the vibration signals contain strong interference signals, the impulse responses related to the fault might not be learned from the vibration signals.

A union of convolutional dictionary learning algorithm (UC-DLA) was proposed by Cristian Rusu [27], which can efficiently implement the convolutional sparse representations of vibration signals. UC-DLA has the clear physical meaning of representing the impulses generated by the circular interactions of faults and their matching surfaces. The learned circular structured dictionaries can describe the shift-invariance of the fault-related impulses; the sparse representations matrix can describe the time and intensity of fault-related impulses occurrence. As such, this paper introduces UC-DLA to the field of wheelset bearing fault diagnosis. Like other dictionary learning methods, UC-DLA does not work well for vibration signals with strong interference signals and inappropriate selections of the related parameters also seriously affect the performance of wheelset bearing fault diagnosis. Therefore, a novel fault detection method, adaptive UC-DLA combined with bandwidth optimization (AUC-DLA-BO), is proposed for wheelset bearing fault diagnosis in this paper. AUC-DLA-BO can eliminate interference signals and adaptively determine parameters related to UC-DLA. The new method not only realizes fault detection of wheelset bearing but also realizes fault behavior analysis.

2. Theoretical Background

2.1. The Convolutional Sparse and Circular Structure Characteristics of Impulses. When a local defect appears on wheelset bearing, bearing rollers will strike the defect reciprocally, and the impulses $I(t)$, shown in Figure 1(a), are repeatedly generated.

Taking a defect of the outer race as an example to illustrate the dynamic generating process of fault impulses $I(t)$. When the roller passes over the defect on the outer race for the first time, the first impulse $I(t_1)$ is generated. Due to the continuous rotation of the wheelset, other impulses, $I(t_1), I(t_2), I(t_3), I(t_4), I(t_5)$, and so on, will be generated in turn. Considering the influence of acquisition noises $N(t)$ and foundation vibration $E(t)$ of the mechanical system, the vibration signals $s(t)$ (plotted in Figure 1(a)) collected by accelerometer will be the mixtures of impulses $I(t)$, acquisition noises $N(t)$, and foundation vibration $E(t)$. The shapes of different impulses in Figure 1 are similar. If the types and sizes of defects, the rotating speeds of wheelset bearings, and the type of wheelset bearings and loads are different, the shapes of impulses will be different. The time when locations of the impulses appear depends on the rotating speed and the geometric parameters of wheelset bearing. Therefore, impulses $I(t)$ can be regarded as the translation of a general shock response with different intensities along the time axis. The convolutional operation can well describe the translation-invariant characteristics of impulses $I(t)$. Impulses $I(t)$, plotted in Figure 1(a), are expressed as

$$I(t) = c(t) \ast x(t),$$

where $\ast$ is the convolution operator, $c(t)$ is the shock responses, and $x(t)$ denotes the locations where the shock responses appear.

Impulses $I(t)$, plotted in Figure 1(a), are the result of convolution of shock responses $c(t)$ and location coefficients $x(t)$. Shock responses $c(t)$ and location coefficients $x(t)$ are shown in Figures 1(b) and 1(c), respectively. Thus, the convolution operator well describes the dynamic generating process of wheelset bearing fault impulse. The natural frequency and damping characteristics of the system can be derived from shock responses $c(t)$. The amplitudes of lo-
cation coefficients $x(t)$ are the intensities of shock responses, which are related to rotation speeds and loads of wheelset bearing. The interval between nonzero coefficients is related to rotation speeds and geometric parameters of wheelset bearing. In practical engineering, only impulses $I(t)$ can be collected. Shock responses $c(t)$ and location coefficients $x(t)$ can only be obtained by solving the inverse problems of equation (1).

Convolution operator is not conducive to the expression of matrix operation. Convolution calculation is converted to matrix multiplication by employing the Toeplitz matrix.

$$y = c \ast x = \text{tope}(c)x = \begin{bmatrix}
    c_1 & 0 & \cdots & 0 & 0 \\
    c_2 & c_1 & \cdots & 0 & 0 \\
    \vdots & \ddots & \ddots & \vdots & \vdots \\
    c_n & \cdots & c_2 & c_1 & 0 \\
    0 & \cdots & c_n & c_2 & c_1 \\
    \vdots & \ddots & \cdots & \ddots & \ddots \\
    0 & \cdots & 0 & \cdots & c_n
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n \\
x_{n+1}
\end{bmatrix},$$  \hspace{1cm} (2)
where the vector \( c \in \mathbb{R}^n \) is shock responses, the vector \( x \in \mathbb{R}^m \) is location coefficients, and the vector \( y \in \mathbb{R}^p \) is the result of their convolution, where \( n \leq m \) and \( p = m + n - 1 \). \( n \) and \( m \) denote the data lengths of shock responses and convolution result, respectively. \( \text{tope}(\cdot) \) denotes the Toeplitz operator.

Circular matrices \( C \in \mathbb{R}^{m \times m} \) are completely defined by their first column \( c \in \mathbb{R}^m \): every column is a circular shift of the first one. The example circular matrix \( C \) can be expressed as

\[
C = \text{circ} \left( c \right) \overset{\text{def}}{=} \begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_2 & c_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_n \\ c_n & \cdots & \cdots & c_1 \end{bmatrix},
\]

where \( \text{circ} \) is the operator that constructs vectors \( c \) into circular matrices. The main property of circular matrices \( C \) is their eigenvalue factorization:

\[
C = F^T \Lambda F, \Lambda = \text{diag} \left( \sigma \right) \in \mathbb{C}^{m \times m},
\]

where \( F \) denotes the unitary Fourier matrix \( (F^H F = I) \) and the diagonal \( \sigma = \sqrt{n} F c, \sigma \in \mathbb{C}^m \).

Convolutional matrices can be reduced to circulant matrices by observing, and equation (2) can be rewritten as

\[
\begin{aligned}
&\text{y} = c \ast x = \text{tope}(c)x = \text{circ}(\text{pad}(c))\text{pad}(x) = C_{\text{conv}} x_{\text{conv}}, \\
&\text{pad}(c) = \begin{bmatrix} c \\ 0_{(m-1) \times 1} \end{bmatrix}, \text{pad}(x) = \begin{bmatrix} x \\ 0_{(m-1) \times 1} \end{bmatrix},
\end{aligned}
\]

where \( C_{\text{conv}} \in \mathbb{R}^{p \times p} \) are convolutional matrices, \( \text{pad}(\cdot) \) is zero padding operator, and \( x_{\text{conv}} \in \mathbb{R}^p \) is location coefficients (also called sparse representations) over convolutional matrices.

In equation (5), the fault impulses of wheelset bearing are described with the convolution matrix multiplied by the locations coefficients. By using dictionary learning and sparse coding technique, the inverse problem of equation (1) can be solved, and the fault information of the wheelset bearing can be well revealed. Therefore, in this paper, UC-DLA is introduced into the field of wheelset bearing fault diagnosis.

2.2. Basic Theory and Discussion of UC-DLA. UC-DLA provides heuristics that approximate solutions to the following problem: given a signal \( Y \in \mathbb{R}^{p \times N} \), a target sparsity \( k \), and the number of dictionaries \( L \), a dictionary \( D_{\text{conv}} \in \mathbb{R}^{p \times pL} \) and a sparse representation \( X_{\text{conv}} \in \mathbb{R}^{p \times N} \) are computed so that the following equation holds [27]:

\[
\begin{aligned}
&\minimize_{D_{\text{conv}}, X_{\text{conv}}} \left\| Y - D_{\text{conv}} X_{\text{conv}} \right\|_F^2, \\
&\text{subject to: } \left\| \text{vec}(X_{\text{conv}}) \right\|_1 \leq k \cdot N,
\end{aligned}
\]

where \( Y = [y^1, y^2, \ldots, y^N] \), \( D_{\text{conv}} = [C_{\text{conv}}^1, C_{\text{conv}}^2, \ldots, C_{\text{conv}}^L], \ X_{\text{conv}} = [x_{\text{conv}}^1, x_{\text{conv}}^2, \ldots, x_{\text{conv}}^L] \), \( C_{\text{conv}}^l \) is \( \text{circ}(\text{pad}(c^l)) \), and \( x_{\text{conv}}^l = \text{pad}(x^l) \). \( \text{vec}(\cdot) \) is an operator that transforms a matrix into a vector, and \( p \) is the length of the target signal. Since the target signal \( Y \) has \( N \) columns, the sparse coefficients \( x^l = [x_1^l, x_2^l, \ldots, x_N^l] \). If the shock response length is \( n \), then \( c^l \in \mathbb{R}^n \). Equation (5) shows that the solution of equation (6) is the solution of the following convolution sparse problem:

\[
\begin{aligned}
&\minimize_{\{c^l\}, \{x^l\}} \sum_{l=1}^N \left\| y^l - \sum_{i=1}^L c^l \ast x_i^l \right\|_F^2, \\
&\text{subject to: } \left\| x_i^l \right\|_1 \leq k.
\end{aligned}
\]

An iterative alternating optimization algorithm could be used to solve this optimization problem: keep dictionary \( D_{\text{conv}} \) fixed, update sparse representation \( X_{\text{conv}} \), then keep sparse representation fixed, update dictionary \( D_{\text{conv}} \), and perform this update process in a loop until convergence.

2.2.1. Updating the Convolutional Dictionary. When both target signals \( \hat{Y} \) and the associated sparse representation \( X_{\text{conv}} \) are fixed, the updating process of a convolutional dictionary is an optimization problem:

\[
D_{\text{conv}} = \arg \min_{D_{\text{conv}}} \left\| Y - D_{\text{conv}} X_{\text{conv}} \right\|_F^2.
\]

Using equation (4), we can get

\[
\left\| Y - D_{\text{conv}} X_{\text{conv}} \right\|_F^2 = \left\| Y - \sum_{l=1}^L C_{\text{conv}}^l x_{\text{conv}}^l \right\|_F^2 = \left\| Y - \sum_{l=1}^L (F^T Y_{\text{conv}} F x_{\text{conv}}^l) \right\|_F^2,
\]

where \( C_{\text{conv}}^l = F^T \Sigma_{\text{conv}}^l F, \ A_l = \text{diag}(\sigma_{\text{conv}}^l) \), and \( \sigma_{\text{conv}}^l = F \cdot \text{pad}(c^l) \). According to the invariance of the Frobenius norm under unitary transformations, equation (9) could be rewritten as
where $\bar{Y}$ is Fourier transform of $Y$, $\bar{x}_l$ and $\bar{c}$ are Fourier transform of $x_{\text{conv}}$ and $c$, respectively. $\otimes$ denotes the Kronecker product, $c = [c^1, \ldots, c^n]^T \in \mathbb{R}^{nL}$ and $F_{1:p}$ is the first $n$ columns of the $p \times p$ Fourier matrix. Due to $\text{vec}(F_{1:p} \bar{c} \bar{x}^l) = ((\bar{x}^l)^T \otimes F_{1:n}) \text{vec}(c^l)$, the matrix approximation problem in equation (8) could be converted to linear least squares problem in equation (10). The solution of equation (10) is given as

$$c = (B^T B)^{-1} B^T \bar{y}. \quad (11)$$

The final, convolutional dictionary $D_{\text{conv}}$ is $[\text{circ}(\text{pad}(c^1)), \ldots, \text{circ}(\text{pad}(c^n))]$.

### 2.2.2. Updating the Sparse Representation.

When convolutional dictionaries are known, updating sparse representation in equation (6) is a typical sparsity-constrained optimization problem, which has been widely discussed in compressed sensing theory, and many methods can accomplish this optimization task. This sparse representation problem is expressed as

$$\begin{align*}
X_{\text{conv}} &= \arg\min_{X_{\text{conv}}} \| Y - D_{\text{conv}} X_{\text{conv}} \|_F^2, \\
\text{subject to: } &\| \text{vec}(X_{\text{conv}}) \|_1 \leq k \cdot N. \quad (12)
\end{align*}$$

The orthogonal matching pursuit (OMP) algorithm [28] is used to solve this optimization problem, but any other sparse approximation method could be chosen. Repeatedly update dictionaries and sparse representations until $\varepsilon^t/\varepsilon^{t-1} \leq \zeta$, the UC-DLA converges, where $\varepsilon^t = \| Y - D_{\text{conv}} X_{\text{conv}} \|_F$ is the error of the $t$th loop and $\zeta$ is the convergence condition. In summary, the algorithm flow of UC-DLA is described in Figure 2.

### 2.3. Discussion of UC-DLA.

UC-DLA provides a good framework for learning convolution dictionary related to shock responses, obtaining the locations where coefficients of shock responses appear and detecting fault impulses of wheelset bearing. In bearing fault signals, impulses are always the key structure indicating defective faults. The convolution sparse coding based on UC-DLA is used to detect the impulse structure in the signal, which provides a good idea for wheelset bearing fault diagnosis.

Taking the simulation signal in Figure 1 as an example, the first signal $x(t)$ contains impulses $I(t)$ and Gaussian noise $N(t)$, and the second signal $s(t)$ is processed by UC-DLA. Results of the UC-DLA processing signal $x(t)$ and signal $s(t)$ are plotted in Figures 3 and 4, respectively. UC-DLA can effectively recover the impulse signal when the impulse signal only contains Gauss noise, as shown in Figures 3(a) and 3(b). The shock responses and location coefficients obtained from the signal $x(t)$ are shown in Figures 3(c) and 3(d); the shape of learned shock responses plotted in Figure 3(c) is very similar to the real shock responses and almost coincides with the first three oscillations of the responses. The time of active nonzero location coefficients inferred from simulation signals completely corresponds to the real situation (see Figure 3(d)). However, if the impulses contain periodic fundamental vibration signal with large energy, a large error will occur when the signal is processed by UC-DLA. The processing results are shown in Figure 4. There are great differences in the shape of shock responses learned from signal $s(t)$ in Figure 4(c) and real shock response, especially in the low oscillation region. Compared with Figures 3(c) and 4(c), the low oscillation part is most susceptible to noise interference because of its small amplitude. However, due to the independent and identical distribution of each sampling point of Gauss noise, the dictionary learning method can better fit the original low-frequency oscillation from the signal with Gauss noise. However, the sampling points of the base vibration do not satisfy the distribution characteristics, so the learned shock response includes the fluctuation components of the foundation vibration, as shown in Figures 3(a), 3(b), 4(a), and 4(b). Because of the large errors in the shock responses learned from the signal, the location coefficients learned from the signal and the real location coefficients are dislocated, as shown in Figure 4(d). In practice, although the whole time domain of shock responses is always disturbed by noise and foundation vibration, the signal-to-noise ratio (SNR) in the high-oscillation region should be much larger than that in the low-oscillation region. In order to describe the shock responses completely, the high-oscillation region of shock responses should contain at least two oscillation periods.

According to the theory of UC-DLA, there are three parameters, length of shock response $n$, a target sparsity $k$, and the number of dictionaries $L$, that need to be predefined and the quality of the algorithm is closely related to these parameters. The too large parameter $n$ will lead to more noise, and the too small one cannot fully describe the shock response characteristics. Parameter $k$ reflects
the number of impulses included in the target signal, which is related to the type of bearing fault and rotation speed of wheelset. If the parameter $k$ is too large, the false impulses will be absorbed into the reconstructed impulses. If the parameter $k$ is too small, some of the impulses with smaller oscillation amplitude are thrown away and the reconstructed impulses cannot represent all the impulses in the target signal. The parameter $L$ indicates the number of resonance frequencies induced by defects of the wheelset bearing. In practical engineering, the measured...
vibration signals of bearing may contain different shock response types, shock response numbers, and shock response lengths. Therefore, improper parameter selection will have adverse effects on the extraction of shock responses caused by wheelset bearing. In conclusion, the convolution sparse coding method used in bearing fault diagnosis needs to overcome the following two kinds of problems:

1. In dictionary learning, how to overcome the influence of foundation vibration and harmonic interference on a dictionary. In practical engineering, these interferences are inevitable, such as low-frequency vibration caused by track irregularity and power-line interference.

2. In dictionary updating and sparse representation updating, how to automatically select the length of shock responses, target sparsity, and number of a dictionary to obtain optimal impulses detection performance.

These two problems are very significant for the application of sparse representation based on the learning dictionary in bearing fault diagnosis.

3. The Proposed AUC-DLA-BO Method

In this section, automatic parameter determination UC-DLA method for fault diagnosis of wheelset bearing is proposed. Firstly, we will discuss the solutions to the two problems mentioned in Section 2.3. To distribute the operation process of the proposed method, we establish a wheelset bearing fault model based on the wheelset bearing system plotted in Figure 5. This model considers the periodic impulses caused by wheelset defects, harmonic interference generated by wheelset polygon, power-line interference, and Gaussian noise generated by measurement. A set of simulation signals generated by the wheelset bearing fault model are designed and expressed as follows: where \( s_1(t) \) denotes the impulses caused by wheelset defects, characteristic frequency \( f_1 \), \( s_2(t) \) means the impulses caused by the bearing outer race fault with characteristic frequency \( f_2 \), \( s_3(t) \) denotes power-line interference with characteristic frequency 50 Hz, \( s_4(t) \) is harmonic interference, and \( N_s(t) \) is Gaussian noise generated at a −5 dB SNR of the sum of \( s_3(t), s_1(t), s_2(t), s_3(t), \) and \( s_4(t) \) are, respectively, expressed as follows:
\[ s_1(t) = \sum_{n=0}^{N_1} A_1 e^{-\beta_1 (t-nT_1)} \cos(2\pi \omega_1 (t-nT_1))u(t-nT_1), \]
\[ s_2(t) = \sum_{n=0}^{N_2} A_2 e^{-\beta_2 (t-nT_2)} \cos(2\pi \omega_2 (t-nT_2))u(t-nT_2), \]
\[ s_3(t) = 0.5 \cos(100\pi t), \]
\[ s_4(t) = 0.3 \sin(2\pi f_1 t) + 0.1 \sin(4\pi f_2 t) + 0.1 \sin(6\pi f_3 t), \]
\[ s_s(t) = s_1(t) + s_2(t) + s_3(t) + s_4(t) + N_r(t), \]
\[ s_f(t) = s_1(t) + s_2(t) + s_3(t) + s_4(t). \]

where \( A_1 \) and \( A_2 \) are the amplitudes of the impulses, \( \beta_1 \) and \( \beta_2 \) are the structural damping coefficients, \( \omega_1 \) and \( \omega_2 \) denote the excited resonance frequencies, \( N_1 \) and \( N_2 \) denote the number of impulses, \( u(t) \) is a unit step function, \( f_1 = 10.29 \) Hz denotes the rotational frequency, \( T_1 = f_1^{-1} \) and \( T_2 = f_2^{-1} \). The parameters related to \( s_1(t) \) and \( s_2(t) \) are listed in Table 1. The sampling frequency of the simulation signal \( s_s(t) \) is 10 kHz.

### 3.1. Elimination Method of Foundation Vibration and Harmonic Interference

If there is a large amount of fundamental vibration and harmonic interference in the analyzed signal, the impulse response learned by UC-LDA method contains part of the fluctuation component of the fundamental vibration (shown in Figures 4(a)–4(c)), which is obviously not conducive to the application of sparse representation method in bearing fault impulses extraction. Ideally, the fundamental vibration and harmonic interference should be eliminated as much as possible before the dictionary learning. As we all know, signals with different vibration modes have different frequencies. It is a natural idea to distinguish different vibration signals by the difference of frequency spectrum of different vibration signals. Figure 6(a) is the simulation signal of wheelset bearing fault represented by equation (13), and Figure 6(b) is the Fourier spectrum of the simulation signal. Different frequency interval (FI) corresponds to different vibration components in Figure 6(b). FI1, FI2, FI3, and FI4 represent the frequency interval of signals \( s_1(t) \), \( s_2(t) \), \( s_3(t) \), and \( s_4(t) \), respectively. Therefore, it is an essential process before the dictionary learning that filtering the signal can eliminate the influence of fundamental vibration and harmonic interference on the algorithm.

In practical engineering applications, the frequency interval of the real fault signal is not constant, and the bandwidth and center frequency of the filter should be adjusted arbitrarily to suit different application scenarios. In this paper, empirical wavelet transform is used as a variable filter, but any other wavelet transforms with similar properties can be selected such as Morlet wavelet [29], harmonic wavelet [30], and tunable Q-factor wavelet transform [31]. For any given frequency interval \([f_l, f_h]\), the empirical scaling function \( \phi(t) \) or the empirical wavelet function \( \psi(t) \) can be used to filter the signal. \( \phi(t) \) is used when \( f_1 = 0, \psi(t) \) is used in other cases. Frequency response function of scale function and wavelets are, respectively, designed as follows [32]:
where $0 \leq f_i < f_h \leq 0.5 f_s$, $f_i$ is sampling frequency, $\chi(f) = f^4 (35 - 84 f + 70 f^2 - 20 f^3)$; and $\gamma < (f_h - f_i)/(f_h + f_i)$ should be satisfied. It is very easy to construct an appropriate filter with EWT when given the cut-off frequency. The filtered result can be calculated by the following formula:

$$\phi(f) = \begin{cases} 1, & |f| \leq (1 - \gamma)f_h, \\ \cos\left(\frac{\pi}{2} \chi\left(\frac{1}{2}f_h \left(|f| - (1 - \gamma)f_h\right)\right)\right), & (1 - \gamma)f_h \leq |f| \leq (1 + \gamma)f_h, \\ 0, & \text{otherwise}, \end{cases}$$

$$\psi(f) = \begin{cases} 1, & |f| \leq (1 + \gamma)f_i, \\ \cos\left(\frac{\pi}{2} \chi\left(\frac{1}{2}f_i \left(|f| - (1 - \gamma)f_i\right)\right)\right), & (1 - \gamma)f_i \leq |f| \leq (1 + \gamma)f_i, \\ 0, & \text{otherwise}. \end{cases}$$

3.2. Automatic Selection of the Length of Shock Responses. The length of shock responses is closely related to the structural damper coefficient. To obtain parameters $\beta_1$, $\alpha_1$ or $\alpha_2$, $\omega_2$ of the impulse response, the previous two maximum values $a(t_1)$ and $a(t_2)$ must be known in advance, as seen in Figures 7(a) and 7(b). The time difference between the two maximum values is $\Delta t = t_1 - t_2 = 1/\omega_2$ (take the shock response in Figure 7(b) as an example). Assuming that the shock response length is $n$, $\omega_2 = 1/\Delta t = f_i/n$, and the resonance frequency $\omega_2$ should satisfy the following relationship:

$$f_{\min} \leq \omega_2 = \frac{f_i}{n} \leq \frac{f_s}{2}, \quad (17)$$

where $f_{\min}$ is the minimum frequency of shock responses. Equation (17) can be converted into the following form:

$$2 \leq n \leq \frac{f_s}{f_{\min}}. \quad (18)$$

According to equation (18), the higher vibration frequency implies the shorter shock response length $n$. The value range determined by equation (18) is too wide to be effectively used in practical projects. That is because we suggest that the target signal should be filtered before the dictionary learning, and the frequency range of the filter is $[f_i, f_h]$. $\omega_2$ should satisfy the following conditions:
\[ f_1 \leq \omega_2 = \frac{f_s}{n} \leq f_h, \]  

Finally, we can determine that the length of the shock response should meet the following equation:

\[ \frac{f_s}{f_h} \leq n \leq \frac{f_s}{f_1} \]  

In general, the central frequency of the resonance band \((f_1 + f_h)/2\) is equal to the resonance frequency \(\omega_2\), so \(n = 2f_s/(f_1 + f_h)\) is a reasonable estimate of shock response length. Note that this value is the minimum value of the shock response length. For a resonance frequency of 1 kHz \((f_s = 10 \text{ kHz})\), \(n\) exceeds 20. Therefore, \(n\) can be initialized to 20. When the central frequency of the filter is less than 1 kHz, the length \(n\) should be increased to represent the low-frequency resonance frequency. Finally, \(n = \max(20, 2f_s/(f_1 + f_h))\) is the reliable value of the length of shock responses.

### 3.4. Automatic Selection of the Number of Dictionaries

According to the physical meaning of shock response, a shock response is mainly determined by resonance frequency and damping coefficient. The number of dictionaries \(L\) depends on the number of resonance frequencies excited by the fault. In order to analyze conveniently, the UC-DLA method is used to process signal \(s_1(t) + s_2(t) + N(t)\), the length of shock responses \(n = 20\), target sparsity \(K = 12\), and the number of dictionaries \(L = 2\). The simulation signal (blue line) and sparse reconstruction signal (red line) are plotted in Figure 9(a). UC-DLA method can accurately extract the impulses of two different structures and has very strong denoising ability. The shock responses learned from the signal in Figure 9(a) are plotted in Figures 9(c) and 9(d). Obviously, the shape of the shock responses is similar to real shock responses. If the number of dictionaries is appropriate, UC-DLA has a good effect on impulses extraction. In fact, we do not know in advance the number of resonance frequencies excited by the fault, which makes it very difficult to choose the number of dictionaries. The spectrums of shock responses are plotted in Figure 9(b). Because the structure of the shock response is related to the resonance frequency, the dominant frequencies of the shock responses can be used to identify different shock responses. It can be seen from Figure 9(b) that the dominant frequencies of shock responses coincide with the center frequency of resonance band of simulation signal. Because different shock response has different resonance band (such as Filter1 and Filter2), a suitable filter can always be found, so the filtered signal only contains one type of shock response. So, the number of dictionaries can always be set to \(L = 1\).

### 3.5. AUC-DLA-BO Method of Fault Impulses Detection

Operator \(CD(\cdot)\) is defined as UC-DLA processing of the signal \(s(t)\), and operator \(M(\cdot)\) is defined as empirical wavelet filtering of the signal \(s(t)\). According to the discussion in the above four sections, a new method of fault impulses detection for wheelset bearing named AUC-DLA-BO can be expressed as the following constrained optimization problem:
\[
\begin{align*}
\begin{bmatrix}
\bar{f}_1, \bar{f}_h \\
\end{bmatrix} &= \arg\max_{f_1, f_h} \text{metric} \left( C D \left( M \left( s \right) \right) \right), \\
\text{subject to:} & \quad 0 \leq f_1 < f_h \leq \frac{f_1}{2}, 6f_{bs} < f_h - f_1,
\end{align*}
\]  \hspace*{1cm} \text{(21)}

where \( \text{metric}(\cdot) \) is an index used for quantification of the fault signal with impulses characteristics. In this paper, the Hilbert envelope spectrum fault feature ratio (HESFFR) is used to determine whether a measured signal is informative, which is given as [33]

\[
\text{metric}(\text{ES}) = \frac{\text{ES}(f_{bs}) + \text{ES}(2f_{bs}) + \text{ES}(3f_{bs})}{\text{ES}},
\]  \hspace*{1cm} \text{(22)}

where \( \text{ES}(f_{bs}), \text{ES}(2f_{bs}), \) and \( \text{ES}(3f_{bs}) \) are envelope spectrum amplitudes corresponding to frequency points \( f_{bs}, 2f_{bs}, \) and \( 3f_{bs}, \) and \( \text{ES} \) is the total amplitudes of the envelope spectrum. Any of the following metaheuristic optimization algorithms, such as genetic algorithm [34], particle swarm optimization [35, 36], quantum inspired differential evolution [37, 38], etc., is capable of solving the constrained optimization problem in equation (21). The new method is a coupling method of filter optimization and convolutional sparse coding and its procedure is illustrated in Figure 10. To sum up, the main procedure of AUC-DLA-OB is as follows:

1. Use EWT filter with the band-pass \([f_1, f_h]\) to filter the original signal.
2. The length of shock response is determined according to the pass-band of the filter, the number of dictionaries is set to 1, then the filtered signal is processed by UC-DLA, and the convolution sparse reconstruction signal is obtained.
3. The filter parameters \([f_1, f_h]\) are continuously updated to maximize the HESFFR value of the reconstructed signal and obtain the corresponding parameters \(\tilde{f}_1, \tilde{f}_h\).
4. Repeat steps 1 and 2 according to parameters \(\tilde{f}_1, \tilde{f}_h\) to detect fault impulses.

4. A Case Study: Fault Simulation Signal

To illustrate the effectiveness of the proposed method, AUC-DLA-OB is used to process the simulation signal in Figure 6(a). First, we try to detect the fault impulses with the fault characteristic frequency of 100 Hz, in which the fault characteristic frequency is a prior parameter. In practical engineering applications, it can be calculated according to structural parameters of bearing and rotation speed. The results obtained using AUC-DLA-OB are shown in Figure 8.
It can be seen from Figure 11(b) that the proposed method accurately finds out the resonance frequency band of shock response 2 and recovers the fault impulses from the resonance band, as shown in Figure 11(d). The shock responses and location coefficients obtained by AUC-DLA-OB are plotted in Figures 11(e) and 11(f). It is obvious that the learned shock response and the real shock response have similar vibration shapes, and the time interval between the shock coefficients is exactly the real fault impulses interval. The signal in Figure 11(c) is a filtered directly without sparse coding. Compared with Figures 11(c) and 11(d), the characteristics of the impulses in AUC-DLA-OB processed signals are more obvious and contain fewer noise signals. The Fourier spectrums and envelope spectrums of the signals in Figures 11(c) and 11(d) are shown in Figure 12. The noise energy in Figure 12(b) is significantly smaller than that in Figure 12(a), and the harmonic components in the envelope spectrum of Figure 12(d) are significantly richer and more obvious than those in Figure 12(c). Figure 12(e) further confirms this conclusion and the envelope spectrum of signal in Figure 12(d) is closest to the real envelope spectrum.

Second, we try to detect the fault impulses with the fault characteristic frequency of 26 Hz. Obviously, this shock response corresponding to this fault feature is shock response 1. The results obtained using AUC-DLA-OB are shown in Figure 13. The proposed method can accurately detect the shock response 1, and the learned shock response 1 can be regarded as the translation of the real shock response 1, as shown in Figure 13(e). Comparing Figures 13(c) and 13(d), and the corresponding Fourier spectrum and envelope spectrum (plotted in Figures 14(a)–14(d)), obviously, the proposed method not only accurately finds the resonance band where the fault feature, but also directly extracts the fault impulses from it, which greatly improves the SNR of the fault impulses. Figure 14(e) shows that the envelope spectrum obtained by AUC-DLA-OB is the closest to the envelope spectrum of real impulses.

It can be seen from the simulation signal composition in Figure 6(a) that the signal contains harmonic interference and power-line interference signals. The proposed method avoids the influence of these interference signals and successfully detects two fault impulses signals. In conclusion, the proposed method can learn the fault shock response and extract the impulses signal related to the learned shock response under the condition of strong interference.

In order to illustrate the effectiveness of the proposed method, the UC-DLA method and the well-known
The spectral kurtosis method are used to analyze the simulation signals in Figure 6(a). The results obtained using the UC-DLA method are plotted in Figure 15. The parameters used in UC-DLA method are \( n = 20, K = 13, L = 2 \) and the original signals are divided into five segments, which form the input matrix of dictionary learning. Because the impulse signal corresponding to impulse response 1 has large energy, so it is less affected by other interference signals, UC-DLA method can effectively extract the relevant impulse signal, as shown in Figures 15(a) and 15(c). The impulses signal corresponding to shock response 2 can hardly be detected directly, as shown in Figures 15(b) and 15(c), but the proposed method can detect the impulse signal accurately, so the AUC-DLA-OB has strong adaptability. UC-DLA still has the disadvantage that parameters cannot be estimated adaptively, which seriously limits the application of this method, and overcoming this disadvantage is exactly one of the remarkable highlights of the proposed method in this paper.

Figure 16 shows the results of spectral kurtosis processing, where \( K_{\text{max}} \) is the maximum of kurtosis. Compared with Figures 13(d) and 16(b), the proposed method can significantly eliminate the noise between adjacent impulses. By comparing the envelope spectrum in Figures 14(d) and 16(c), the amplitude and the number of harmonics of the proposed AUC-DLA-OB are better than those of the spectral kurtosis method. At the same time, the spectral kurtosis method can only detect the impulse signal corresponding to shock response 1, but not impulse signal corresponding to impulse response 2. The comparison results show that the proposed method is more suitable for fault impulses detection.
5. A Case Study: Experimental Data of Wheelset Bearings

In this section, the experimental data collected on a wheelset bearing test bench are used to illustrate the advancement of the proposed AUC-DLA-OB. The test bench is composed of a driving wheel, motor, loading device, test wheelset, and acceleration measuring equipment. The acceleration sensors installed on wheelset axle box are used to collect vibration signals of wheelset bearings. The wheelset bearing test bench is shown in Figures 17(a)–17(c). Two types of fault bearings are selected for testing which are, respectively, shown in Figures 18(a) and 18(b). Outer race defect and roller defect are conducted in fault experiments. In all experiments, the rotation frequency of the wheelset is 10.28 Hz and sampling frequency of the vibration signal is \( f_s = 10 \text{ kHz} \). The structural parameters and fault characteristic frequencies of the bearing are shown in Table 2.

Firstly, the proposed AUC-DLA-OB is used to process the outer race faults of wheelset bearing. The collected vibration signals with outer race faults are shown in Figure 19(a) and its corresponding Fourier spectrum is shown in Figure 19(b). Because the test environment is a strong electromagnetic environment, the collected signal has serious power-line interference. The results of wheelset bearing outer race fault signal processed by AUC-DLA-OB are plotted in Figure 20. The filter obtained by the proposed method is shown in Figure 20(b). Obviously, the filter in Figure 20(b) can effectively eliminate power-line interference and capture fault signals. The signal filtered by the filter in Figure 20(b) is shown in Figure 20(c). The filtered signal contains a lot of fault impulses, but there is still a lot of noise. The shock responses, convolutional sparse reconstruction signal, and corresponding location coefficients obtained through AUC-DLA-OB are shown in Figures 20(d), 20(e), and 20(g), respectively. Compared with Figures 20(c) and 20(d), the proposed method successfully captures the fault...
impulses in the signal. The convolution sparse reconstruction signal has obvious impact characteristics, and the noise in the impulses is effectively suppressed.

In order to explain the action behavior between the rollers and the outer race defects, the dynamic action relationship of the rollers and outer race are plotted in Figure 21. Three defects are evenly distributed on the outer race of the wheelset bearing. When the rollers pass over each defect, a strong impact will be generated. When the roller 1 passes over the defect A in Figure 21, the first impulse $1A$ will be generated. At the same time, the rollers 7 and 13 near the defects B and C will pass over defects C and B, and the impulses $7B$ and $13C$ will be successively generated. Finally, as the rotation of bearing, the impulse sequences generated by the outer ring defect (such as $1A$, $7B$, $13C$, $19A$, $6B$, $12C$, $18A$, $5B$, and $17C$) will be obtained. In the actual test, only one defect is in the bearing region; assuming that the defect is A, the impulse sequences related to defect A (such as $1A$, $2A$, $3A$, $4A$, $5A$, $6A$, $7A$, $8A$, $9A$, etc.) should be emerged in the collected signal. Figures 20(f) and 20(h) are locally zoomed figures of Figures 20(e) and 20(g), respectively. Compared with Figures 20(f) and 22, the arrangement of impulse sequences detected by the proposed AUC-DLA-OB is consistent with the theoretical impulse sequences. At the same time, the time when the nonzero location coefficient occurs corresponds to the time of the fault impulse that occurs one by one, as shown in Figures 20(f) and 20(h). The time interval of the nonzero location coefficient is exactly equal to $f_{BPFo}$. In conclusion, AUC-DLA-OB can well describe the dynamic behavior and active relationship between the rollers and outer race.

The envelope spectrums of the original signal and convolutional sparse reconstruction signal in Figure 20(e) are shown in Figures 22(a) and 20(b), respectively. By comparing Figures 22(a) and 22(b), the characteristic frequency $f_{BPFo}$ and its harmonics of the outer race fault of the wheelset bearing is found in Figures 22(a) and 22(b). However, compared with Figure 22(a), the amplitude of the envelope spectrum in Figure 22(b) is significantly enhanced and more harmonics are obtained. Therefore, the results show that the proposed AUC-DLA-OB can effectively detect the bearing outer race faults.

Secondly, the proposed AUC-DLA-OB is used to process the roller faults of wheelset bearing. The roller fault signal is shown in Figure 23(a) and its corresponding Fourier spectrum is shown in Figure 23(b). There are a lot of low-frequency components in Figure 23(b) that are not related to
the fault, and these low-frequency components are mainly composed of power-line interference and foundation vibration of the test bench. In the process of fault analysis, these interferences should be eliminated as much as possible to highlight the fault characteristics, and the proposed AUC-DLA-OB can deal with this problem effectively. Figure 24 shows the processing results of roller fault signal with AUC-DLA-OB. The optimized filter effectively avoids the interference signal and detects the resonance frequency band of the fault, as shown in Figure 24(b). The impulse characteristic of the signal filtered by the optimized filter in Figure 24(b) is more obvious than that of the original signal, as shown in Figure 24(c). The shock responses, convolutional sparse reconstruction signal, and corresponding location coefficients obtained through AUC-DLA-OB are shown in Figures 24(d), 24(e), and 24(g), respectively. According to the comparison between Figures 24(c) and 24(e), convolutional sparse coding can detect the impulses in the signal accurately and suppress the noise between the two adjacent impulses utmost.

In order to analyze the dynamic behavior between roller fault and races, the interaction relationship between roller fault and races is shown in Figure 25. With the rotation of the bearing, the defective roller rolls between the inner and outer races of the bearing. When the roller rotates for a circle, the roller fault collides with the outer race and produces impulses. Suppose that the starting position of the defective roller is 1A, after a half-rotation of the roller, the roller enters the bearing region (BR) from the non-bearing region (NBR). At this time, the defective roller collides with the inner race and produces impulses 1B. With the rotation of the roller, the impulse sequences (such as 1C, 1D, 1E, 1F, 1G, 1H, 1J, 1L, 1K, 1M, and 1N) will be generated in turn. Because the roller is only subjected to a small force in NBR, the amplitudes of the impulses generated in the NBR are very

![Figure 13: AUC-DLA-OB for detection of fault impulses with the fault characteristic frequency of 26 Hz.](image)

- (a) Convergence process of solving constrained optimization problems;
- (b) filtering frequency band related to fault impulses;
- (c) filtered signal by filters in Figure 13(b);
- (d) sparse reconstruction signal;
- (e) learned shock response 1; and
- (f) location coefficients obtained from simulation signals.
Figure 14: (a) Fourier spectrum of signal in Figure 13(c); (b) Fourier spectrum of signal in Figure 13(d); (c) and (d) envelope spectrums of signals in Figure 13(c) and in Figure 13(d), respectively; and (e) comparison between envelope spectrum of real impulses and obtained envelope spectrum (envelope spectrums in Figures 14(c) and 14(d)).

Figure 15: Processing results obtained using UC-DLA. (a) Reconstruction impulses corresponding to the learned shock response 1; (b) reconstruction impulses corresponding to the learned shock response 2; (c) learned shock response 1; and (d) learned shock response 2.
$K_{\text{max}} = 14.6$ @ Level 2.5, $Bw = 833.3333 \text{Hz}$, $f_c = 1250 \text{Hz}$

Figure 17: (a) Schematic sketch of the wheelset bearing test bench; (b) photographs of test bench; and (c) installation position of acceleration sensor.

Figure 18: Photos of defects. (a) Three artificial outer race defects and (b) an artificial roller defect.

<table>
<thead>
<tr>
<th>$d$ (mm)</th>
<th>$D$ (mm)</th>
<th>$Z$</th>
<th>$\alpha$ (rad)</th>
<th>$f_{\text{BPFO}}$ (Hz)</th>
<th>$f_{\text{BSF}}$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>180</td>
<td>19</td>
<td>$\pi/20$</td>
<td>83.23</td>
<td>33.93</td>
</tr>
</tbody>
</table>
Figure 19: (a) Bearing outer race fault signal and (b) its Fourier spectrum.

Figure 20: AUC-DLA-OB for identification of wheelset bearing outer race fault. (a) Convergence process of solving constrained optimization problems; (b) filtering frequency band related to fault impulses; (c) filtered signal by filters in Figure 20(b); (d) learned shock responses; (e) sparse reconstruction signal; (f) locally zoomed figure of the impulses shown in Figure 20(e); (g) obtained location coefficients; and (h) locally zoomed figure of the location coefficients shown in Figure 20(g).
small, which is often submerged in the noise and difficult to
detect. By comparing Figures 24(f) and 26, the arrangement
of impulse sequences detected by the proposed AUC-DLA-
OB is consistent with the theoretical impulse sequences in
NR.

Figures 25(a) and 25(b) show the envelope spectrums
of the original signal in Figure 23(a) and the convolution
sparse reconstruction signal in Figure 24(e), respectively.
Compared with Figures 25(a) and 25(b), the envelope
spectrum of the original signal can find the fault character-
istic frequency $f_{BPFO}$ and its twice harmonic com-
ponent. However, due to a large number of interference
components in the whole envelope spectrum, it is easy to
misjudge the bearing fault based on the envelope spec-
trum in Figure 25(a). The characteristic frequency and its
higher harmonics of bearing roller fault are very obvious
in the envelope spectrum in Figure 25(b), and the spec-
trum line is very clean and easy to distinguish, which is
very conducive to the judgment of fault type. So, the
results show that the proposed AUC-DLA-OB can ef-
fectively detect the bearing roller fault.

In order to verify the advantages of the proposed AUC-
DLA-OB, traditional convolutional sparse coding (UC-
DLA) and spectral kurtosis are used to process the measured

---

**Figure 21:** Dynamic action relationship between outer race faults and rollers.

**Figure 22:** (a) Envelope spectrum of original bearing outer race fault signal and (b) envelope spectrum of signal in Figure 20(e).

**Figure 23:** (a) Bearing roller fault signal and (b) its Fourier spectrum.
Figure 24: AUC-DLA-OB for identification of wheelset bearing roller fault. (a) Convergence process of solving constrained optimization problems; (b) filtering frequency band related to fault impulses; (c) filtered signal by filters in Figure 24(b); (d) learned shock responses; (e) sparse reconstruction signal; (f) locally zoomed figure of the impulses shown in Figure 24(e); (g) obtained location coefficients; and (h) locally zoomed figure of the location coefficients shown in Figure 24(g).

Figure 25: (a) Envelope spectrum of original bearing roller fault signal. (b) Envelope spectrum of signal in Figure 24(e).
Figure 26: Dynamic action relationship between races and roller fault.

Figure 27: UC-DLA with parameters $n = 20, k = 17, L = 1$ for identification of wheelset bearing outer race fault. (a) Sparse reconstruction signal; (b) locally zoomed figure of the impulses shown in Figure 27(a); (c) learned shock responses; and (d) envelope spectrum of signal in Figure 27(a).

Figure 28: Continued.
Figure 28: UC-DLA with parameters $n = 20, k = 34, L = 2$ for identification of wheelset bearing outer race fault. (a) Learned shock response 1; (b) learned shock response 2; (c) sparse reconstruction signal corresponding to shock response 1; (d) sparse reconstruction signal corresponding to shock response 2; (e) locally zoomed figure of the impulses shown in Fig. (c); (f) locally zoomed figure of the impulses shown in Figure 28(d); (g) envelope spectrum of signal in Figure 28(c); and (h) envelope spectrum of signal in Figure 28(d).

Figure 29: Spectral kurtosis for identification of wheel bearing outer race fault. (a) Fast Kurtogram; (b) filtered signal; and (c) envelope spectrum of signal in Figure 29(b).
Figure 30: UC-DLA with parameters $n = 20, k = 13, L = 1$ for identification of wheelset bearing roller fault. (a) Sparse reconstruction signal; (b) locally zoomed figure of the impulses shown in Figure 30(a); (c) learned shock responses; and (d) envelope spectrum of signal in Figure 30(a).

Figure 31: Continued.
signals shown in Figures 19(a) and 23(a). The results of UC-DLA processing the outer race fault are shown in Figures 27 and 28. It can be seen from Figure 27 that when the number of dictionaries is 1, the fault impulses cannot be detected directly. When the number of dictionaries is selected as 2, the outer race fault feature can be detected, as shown in Figure 28(g). Compared with Figures 28(a) and 20(d), it can be seen that the mean value of the learned shock response is not zero, which is contrary to the definition of shock response. The reason for this result is that the fault impulses deviates from zero due to the interference of power-line and foundation vibration. It can also be seen from Figure 28(e) that the impulses in the sparse reconstructed signal deviate from the zero point. Comparing Figures 28(g) and 21(b), it is obvious that the envelope spectrum obtained by AUC-DLA-OB is better than that by UC-DLA. It should be noted that Figure 28 is the result of parameters fine-tuning, but in actual engineering application, good parameter selection is difficult to achieve. The results of spectral kurtosis processing the outer race fault are shown in Figure 29. The essence of spectral kurtosis is a filtering method, which is equivalent to step 4 in our proposed method. Although the spectral kurtosis can find the fault characteristics of the bearing outer race, as shown in Figure 29(c), it is obviously worse than the result in Figure 22(b). According to the time domain waveform analysis in Figures 20(e) and 29(b), the SNR of the impulses signal detected by the proposed method is significantly higher than that of spectral kurtosis. The results of UC-DLA and spectral kurtosis processing bearing roller fault are shown in Figures 30–32. Comparing the results, similar conclusions in the comparative analysis of outer race fault processing can be obtained.

Compared with UC-DLA, the proposed method can get good results without parameters fine-tuning. Compared with the spectral kurtosis, the proposed method detects the fault resonance band by filtering method and then
reconstructs the filtered signal by convolutional sparse representation. The reconstructed signal can suppress the noise between two adjacent fault impulses, as shown in Figures 21(e) and 24(e). Because the noise in the resonance band is in the same frequency band as the fault signal, the filtering method, such as spectral kurtosis, cannot eliminate this noise. It should be noted that the spectral kurtosis is developed for fast fault diagnosis task, while the proposed method focuses on more refined fault behavior analysis (the waveform of the impulse response, time location, and intensity of impulses). Refined analysis leads to a large amount of calculation consumption. Therefore, the computational complexity of the proposed method is higher than the spectral kurtosis, and it is suitable for offline diagnosis with abundant computational resources.

The purpose of selecting spectral kurtosis as a comparison method is to explain the difference between the filtering method and the proposed method.

6. Conclusion

In this paper, adaptive filter technique based on EWT and convolutional sparse coding technique based on UC-DLA were introduced to detect faults impulses in wheelset bearings. A novel fault detection method, named AUC-DLA-OB, was proposed. The new method has the following innovations:

(1) The bearing fault impulses are expressed as the convolution of the shock response and the location coefficients, in which the shock response is related to the resonance frequency and damping coefficient, and the location coefficients characterize the time location and intensity of the impulses. AUC-DLA-OB provides an effective algorithm framework for the description of bearing fault signal as shock response-location coefficients.

(2) The proposed AUC-DLA-OB can eliminate the influence of harmonic interference signal on dictionary learning and realize the adaptive selection of parameters (i.e., the length of shock response n, target sparsity k, and the number of dictionaries L) related to convolutional sparse coding technique.

(3) Compared with other advanced bearing fault detection methods, AUC-DLA-OB can suppress the noise between the two adjacent impulses utmost and enhance the fault impulses. Therefore, the proposed method is very suitable for refined analysis of bearing fault behavior.

Finally, the high computational complexity of the proposed method has to be mentioned. In the proposed method, the GA algorithm combined with EWT is used to detect the fault resonance band to eliminate the fundamental vibration and harmonic interference and to ensure that the filtered signal has a single fault type to facilitate the selection of parameter L. Therefore, the proposed method has high computational complexity, which is suitable for offline diagnosis tasks requiring refined analysis. If the fundamental vibration and harmonic interference are in low-frequency band and the number of fault types in the signal is known, a good result can be obtained by using high-pass filtering method combined with UC-DLA. Otherwise, it is recommended to use the AUC-DLA-OB to detect bearing fault.

Data Availability

The bearing test adopts the double-row tapered roller bearings with three types of faults, including three defects on the surface of outer race, two defects on the surface of two rolling elements, and their combination. The bearing test is conducted in CRRC, and the data are confidential.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (no. 51905453) and the China Post-doctoral Science Foundation (no. 2019M663899XB).

References


