Research Article

Wheelset-Bearing Fault Detection Using Adaptive Convolution Sparse Representation

Jianming Ding [iD],[1,2] Zhaoheng Zhang [iD],[2] and Yanli Yin [iD][3]

[1]*The State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044, China*
[2]*The State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China*
[3]*School of Mechatronics & Vehicle Engineering, Chongqing University, Chongqing 40074, China*

Correspondence should be addressed to Jianming Ding; fdingjianming@126.com

Wheelset bearings are crucial mechanical components of high-speed trains. Wheelset-bearing fault detection is of great significance to ensure the safety of high-speed train service. Convolution sparse representations (CSRs) provide an excellent framework for extracting impulse responses induced by bearing faults. However, the performance of CSR on extracting impulse responses is fairly sensitive to inappropriate selection of method-related parameters, and a convolution model for representing the impulse responses has not been discussed. In view of these two unsolved problems, a convolutional representation model of the impulse response series is developed. A novel fault detection method, named adaptive CSR (ACSR), is then proposed based on combinations of CSR and methods for estimating three parameters related to CSR. Finally, the effectiveness of the proposed ACSR method is validated via simulation, bench testing, and a real-life running test employing a high-speed train.

1. Introduction

Wheelset bearings are crucial mechanical components of high-speed trains, and their major roles are to transform the rotational motion of wheelsets to the linear motion of high-speed trains, transmit driving motor torques, and bear the vertical loads of frames and car bodies. During the long-term running process of a high-speed train, the complex dynamic actions on wheelset bearings inevitably lead to the initiation and further extension of wheelset-bearing faults and finally endanger train operational safety [1]. Therefore, it is of great significance to detect wheelset-bearing faults to ensure the safety of high-speed train service.

Vibration-based analyses, as feasible and effective tools for the detection of wheelset-bearing faults, can provide fruitful feature information regarding the working status of a monitored bearing [2]. Once a defect appears on the surface of a bearing component, a series of impulse responses induced by the defect will be generated as the wheelset rotates. However, when the defect enters and leaves the bearing zone, the amplitudes of those impulse responses will be

modulated. When bearing rollers slip, the repetitive frequencies of the impulse responses are modulated [3]. Under certain conditions, impulse response series with different resonance frequencies will be excited [4]. In addition, measured noise and strong wheel-rail interference submerge or pollute the weak impulse responses [5]. The spectra of the fault signals will be smeared. As a result, such problems make the detection of wheelset-bearing faults complex and difficult.

To effectively resolve those difficulties in detecting wheelset-bearing faults, many advanced signal processing methods have been proposed, which primarily include the filter-based high-frequency resonance technique [6], short-time Fourier transform (STFT) [7], Wigner–Ville distribution (WVD) [8], empirical mode decomposition (EMD) [9], wavelet transform (WT) [10, 11], and compressive sensing (CS) [12, 13]. However, the inconvenient selection of centre frequency and bandwidth hampers the wide application of the filter-based high-frequency resonance technique [14]. STFT is inherently unsuited to analyse time-varying signals [15]. The cross terms of WVD on the

multicomponent signals cause unexpected interference when detecting faults [16]. As a result, EMD, WT, and CS have been widely applied in the field of rotational machine fault detection and have become the leading algorithms for fault detection.

EMD, as an adaptive signal processing method, is suitable for analysing nonlinear and nonstationary signals and can be used to decompose analysed signals into sets of intrinsic mode functions (IMFs) and residuals [17]. Hence, EMD has been employed quite successfully in the fields of fault diagnosis, failure detection, damage identification, and health monitoring [18]. However, EMD suffers from the lack of a theoretical foundation, and the definition of its IMF is still controversial [19]. In addition, measured noise and sampling errors easily result in the incorrect placement of signal extremes, and EMD-sifting processes based on the envelopes determined by incorrect extremes inevitably will generate inaccurate or even erroneous IMFs. Thus, EMD performs poorly when used for fault detection in low signal-to-noise ratio (SNR) or weak transient situations [20]. An IMF could include several-mode resonance responses induced by bearing faults because of EMD mode mixtures, and the single-mode resonance responses induced by the bearing faults could be divided into the different IMFs because of the mode break. Mode problems caused by using EMD adversely affect fault detection performance [21]. To alleviate mode problems, variants of EMD (such as EEMD [22], BEMD [23], CEMD [24], and MEMD [25]) have been developed. However, there is still no theory that guarantees mode problems will be avoided.

WT can provide time-frequency information on the impulse response induced by bearing faults. The continuous wavelet transform (CWT) and discrete wavelet transform (DWT) methods have been successfully applied to fault detection for rotational machinery [26] given their multi-resolution merits. The redundancy coefficients and huge computational costs of CWT restrict its wide application in practical engineering. With the fast iterative algorithm, the DWT has gained fruitful application in mechanical fault detection [27]. To improve the decomposition performance of WT for high-frequency bands containing rich fault modulation information, the wavelet packet transform was proposed [28]. However, the decomposition quality of a DWT heavily depends on the selection of the mother wavelet [29]. The shift-variance characteristic of most DWTs causes the impulse responses to be distorted [30]. The fixed dyadic frequency partitioning of DWTs easily generates scale mixtures and scale breaks [31]. The low oscillation of the wavelet basis in a DWT weakens its ability to sparsely represent impulses with highly oscillatory characteristics [32]. To well match vibration signals containing multimode resonance responses induced by composite faults, the multiwavelet packet was proposed [20]. To realize the shift-invariant function of WTs, the dual-tree complex wavelet [30] and higher-density dyadic wavelet transform [33] were developed. To improve the flexibility of the frequency partitioning of WTs, overcomplete rational dilation discrete wavelet transform was proposed [31]. To adjust WT oscillations, the wavelet transform with tunable Q-factor [32] and

ensemble superwavelet transform [34] were proposed. Therefore, there are no WTs that have comprehensive performance (such as self-adaption, shift invariance, flexible partitioning of frequency bands, and tunable oscillations).

Sparse representation mainly consists of sparse coding and dictionary design. Sparse coding models the analysed signals as linear combinations of atoms in a redundant dictionary. Dictionary design is, as much as possible, adapted to the features of the vibration signals to well match the high-level structures of the impulse responses embedded in the vibration signals. Sparse representation is widely employed and has yielded state-of-the-art results in multiple fields of machine learning, neuroscience, signal processing, image and audio processing, classification, and statistics [35, 36]. In terms of sparse coding, its exact resolution is usually an NP-hard problem [37]. Thus, some pursuit algorithms are considered instead, mainly including matching pursuit [38] and basis pursuit [39]. For dictionary design, the methods for constructing dictionaries include those for explicit and implicit dictionaries. An implicit dictionary is a dictionary that is directly inferred from input data using machine learning techniques, including regular dictionary learning [40] and shift-invariant dictionary learning (SIDL) [41, 42]. By virtue of the excellent performance of sparse representation for representing signals (such as flexibility, sparsity, and superresolution), sparse representation-based fault detection has become increasingly popular in the field of mechanical fault detection. Sparse representation based on matching pursuit and explicit dictionaries has been used to extract impulse responses induced by rotational machine faults for gear and bearing fault detection [43]. Sparse representation, through the combination of basis pursuit and explicit dictionaries, is employed to capture impulse patterns for rotating machine fault diagnosis [44–46]. Dictionary learning has advantages and potential for mining high-level structures embedded in signals. A nonlocal sparse model based on regular dictionary learning has been proposed [47]. SIDL can obtain single-value and jointly optimized results over the entirety of vibration signals, unlike regular dictionary learning. Sparse representation based on pursuit algorithms and SIDL can be exploited to extract impulse responses submerged in the vibration signals of rotating machine systems [41, 42, 48, 49]. Those findings show that the fault detection performance of sparse representation is superior to those of EMD and WT. In addition, group sparsity is applied to bearing fault detection [50] but requires prior knowledge of the impulsive periods.

Convolution sparse representation is another name for SIDL-based sparse representation. A new approach based on the alternating direction method of multipliers was proposed in 2016 [51] and is called CSR (convolution sparse representation based on the alternating direction method of multipliers) in this paper. CSR realizes not only SIDL but also shift-invariant sparse representations (SISC) of vibration signals. The strategy of the interleaved optimization between SIDL and SISC rather than the alternating optimization of SIDL and basis pursuit in traditional SIDL-based sparse representation leads to higher computational efficiencies and more accurate convolution

sparse representations [52]. CSR has been applied to the detection of faults in wheelset bearings [5]. Although CSR has obtained satisfied fault detection results, convolution sparse representation framework or model for representing impulse response series has still not been discussed, and its fault detection performance is sensitive to inappropriate selections of method-related parameters. In view of these two unsolved problems, a convolutional representation model of impulse response series induced by bearing faults is proposed. A novel fault detection method, which is named adaptive CSR (ACSR), is then proposed in this paper.

This paper is organized as follows. The convolutional representation model for characterizing impulse response series induced by bearing defects is proposed in Section 2. Section 3 introduces the basic theory of CSR and discusses CSR-related parameters. A novel fault detection method, ACSR, is proposed in Section 4. A simulation-based verification of ACSR is conducted in Section 5. An experimental validation of ACSR is performed in Section 6. Section 7 concludes the paper.

2. Convolutional Representation Model of Impulse Response Series

When there is a defect on the surface of a wheelset-bearing component, an impulse response series (IRS) will be generated as the wheelset rotates. An impulse response caused by the defect can be modelled as the impulse response of a single degree of freedom mass-spring-damper system [3]:

$$r(t) = Ae^{-\beta t}\sin(2\pi f_r t)u(t). \tag{1}$$

Hence, the IRS with fault-characteristic frequency of T_p^{-1} can be represented as [3]

$$
\begin{aligned}
s(t) &= \sum_{m=0}^{M-1} r_m\left(t - mT_p - \sum_{i=0}^{m}\tau_i\right) \\
&= \sum_{m=0}^{M-1} A_m e^{-\beta\left(t-mT_p-\Sigma_{i=0}^{m}\tau_i\right)}\sin\left(2\pi f_r\left(t - mT_p\right.\right. \\
&\quad \left.\left. - \sum_{i=0}^{m}\tau_i\right)\right)u\left(t - mT_p - \sum_{i=0}^{m}\tau_i\right),
\end{aligned}
\tag{2}
$$

where A_m is the amplitude of the mth impulse response, $u(t)$ is a unit step function, T_p is the time period corresponding to the fault-characteristic frequency, β is the structure damper coefficient, f_r is the excited resonance frequency, τ_i represents the effects of the random slippage of the rollers and is the ith realization of a zero mean, uniformly distributed random variable with standard deviations of $0.01T_p \sim 0.02T_p$, and $s(t)$ is the IRS with M impulse responses.

Because such vibrations are often measured using an accelerometer, the measured vibration signal can be described in an acceleration format [3], i.e.,

$$r''(t) = Le^{-\beta t}\cos(2\pi f_r t + \theta)u(t), \tag{3}$$

where

$$
\begin{aligned}
\theta &= \pi - \tan\left(A\left(\beta^2 - (2\pi f_r)^2\right)\left(-4\pi A\beta f_r\right)^{-1}\right), \\
L &= \sqrt{A^2\left(\beta^2 - (2\pi f_r)^2\right)^2 + 4A^2\beta^2(2\pi f_r)^2}.
\end{aligned}
\tag{4}
$$

According to equations (2) and (3), the acceleration version of the IRS, $a(t)$, can be expressed as

$$
\begin{aligned}
a(t) &= \sum_{m=0}^{M} L_m \exp\left(-\beta\left(t - mT_p - \sum_{i=0}^{m}\tau_i\right)\right)\cos \\
&\quad \cdot \left(2\pi f_r\left(\left(t - mT_p - \sum_{i=0}^{m}\tau_i\right)\right)\right)u\left(t - mT_p - \sum_{i=0}^{m}\tau_i\right).
\end{aligned}
\tag{5}
$$

Define impulse responses $d(t)$ as

$$d(t) = \exp(-\beta t)\cos(2\pi f_r t + \theta)u(t). \tag{6}$$

Define time-location coefficients $x(t)$ as

$$
x(t) = \begin{cases} L_m, & t = mT_p + \sum_{i=0}^{m}\tau_i, \\ 0, & \text{other.} \end{cases}
\tag{7}
$$

The acceleration version of the IRS in equation (5) can be modelled as the convolution of the defined impulse response $d(t)$ and associated defining time-location coefficients $x(t)$:

$$a(t) = d(t) * x(t). \tag{8}$$

If the measured vibration signals contain C kinds of impulse responses with different resonance frequencies, the measure-version IRS can be modelled as

$$a(t) = \sum_{c=1}^{C} d_c(t) * x_c(t) + N, \tag{9}$$

where N denotes the measured noises, $d_c(t)$ is the cth impulse response type, and $x_c(t)$ is the time-location coefficients related to $d_c(t)$.

The convolution representation model for representing the IRS in equation (9) can clearly represent the dynamic interaction procedures of the defects of wheelset-bearing components and their matching surfaces and the vibration characteristics. To illustrate the physical meanings of the convolution representation model, an example is shown in Figure 1.

In Figure 1(b), impulse responses $d_c(t)$, $c \in [1, C]$ can describe the dynamic behaviour of the single strike of a defect with its matching surface and are related to the resonance frequencies and damping coefficients of the resonances excited by the strike. The nonzero elements of the time-location coefficients can determine the time of fault action by aligning the time coordinates in Figures 1(a) and 1(c). The amplitudes of the nonzero elements of the time-location coefficients can represent the amplitude-modulation behaviour caused by alternately entering and leaving the
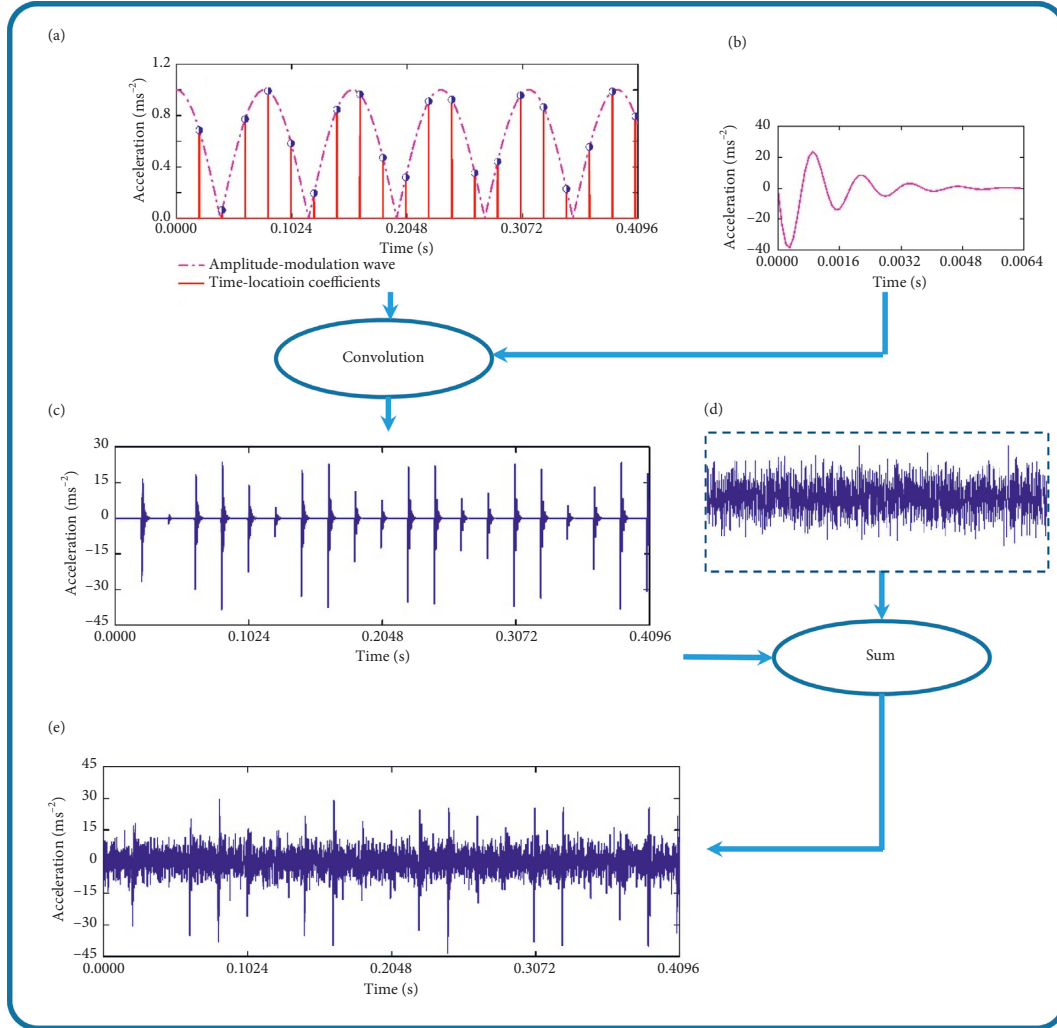
Figure 1: Convolutional representation model of impulse response series. (a) Time-location coefficients. (b) An impulse response. (c) IRS. (d) Measured noises. (e) Measure-version IRS.

wheelset-bearing load zone of the defective inner race or rolling elements. The time span between the nonzero elements of the time-location coefficients can characterize the frequency modulation induced by the slippage of the rollers or fluctuations in rotation speed. Those zero elements of the time-location coefficients properly can eliminate the measured noises. Therefore, the convolutional representation model is fairly suitable for representing the impulse response series caused by bearing faults. If there is a technique that can directly infer the impulse responses $d_c(t)$ and time-location coefficients $x_c(t)$ measured IRS in Figure 1(e). Obviously, the information for detecting bearing faults will be obtained.

3. Basic Theory of CSR

Convolutional sparse representation properly provides a framework for inferring the impulse responses and time-location coefficients from the measured vibration signals $a(t)$ in equation (9) and is expressed as [51]

$$
\left\{d_c(t), x_{k,c}(t)\right\} = \underset{d_c, x_{k,c}}{\arg\min} \frac{1}{2} \sum_{k=1}^{K} \left\| a_k - \sum_{c=1}^{C} d_c(t) * x_{k,c}(t) \right\|_F^2
$$

$$
+ \lambda \sum_{k=1}^{K} \sum_{c=1}^{C} \left\| x_{k,c}(t) \right\|_1,
$$

$$
\left\| d_c(t) \right\|_2^2 = 1, \quad \forall c,
$$

(10)

where $a_k(t) \in R^n$ denotes the kth set of analysed signals with length n, $d_c(t)$ denotes different kinds of impulse responses, $x_{k,c}(t) \in R^{n-P+1}$ denotes time-location coefficients associated with the analysed signal $a_k(t)$ and impulse response $x_{k,c}(t)$, p is the length of impulse response, $\lambda \in R^+$ is a regularization parameter, and the constraint on the norms of shock responses $d_c(t)$ avoids the scaling ambiguity between the impulse response and time-location coefficients.

The solution methods to the optimization problem in equation (10) can be divided into two categories: alternating

optimization and interleaved optimization. Alternating optimization methods solve for impulse responses and time-location coefficients using different optimal techniques, e.g., feature-sign search [52] or fast iterative shrinkage thresholding [53] are used to optimize the time-location coefficients, and the Lagrange multiplier method [48] is employed to optimize impulse responses. After many calculation steps when updating the time-location coefficients, switch to the impulse responses, and alternating execution of the two optimization processes until convergence. However, the interleaved optimization can simultaneously update the impulse responses and time-location coefficients in a calculation step using the same optimization method [51]. ADMM-based CSR is a kind of interleaved version of optimizing the convolution sparse representation in equation (10). Interleaved optimization has much higher optimization efficiency and obtains even more accurate results than alternating optimization [51]. ADMM-based interleaved optimization includes shift-invariant sparsity coding and shift-invariant dictionary learning in an optimization calculation step.

After SIDL and SISC are performed on a set of signals $\left(a_1(t) \; a_2(t) \; \cdots \; a_K(t) \right)$, different types of impulse responses $d_c^L(t)$ are obtained. The time-location coefficients $x_c^L(t)$ are inferred by the obtained impulse response-based SISC of the signals $a(t) = \left(a_1^T(t) \; a_2^T(t) \; \cdots \; a_K^T(t) \right)^T$ and are expressed as

$$x_c^L(t) = \underset{x_c^L(t)}{\arg\min} \frac{1}{2} \left\| a(t) - \sum_{c=1}^{C} d_c^L(t) * x_c^L(t) \right\|_F^2 + \lambda_c \sum_{c=1}^{C} \left\| x_c^L(t) \right\|_1.$$

(11)

Therefore, IRSs embedded in measured signals are extracted using the following formula:

$$\text{IRS}_c = d_c^L(t) * x_c^L(t), \quad c \in [1, C].$$

(12)

Although convolution sparse representation provides a framework for extracting IRSs caused by wheelset-bearing faults, and ADMM can effectively and efficiently solve CSR-related optimization problems, practical research has discovered that inappropriate selections of CSR-related parameters adversely influence the extraction of IRSs. CSR-related parameters can be divided into two categories: the boundary condition-related parameters listed in Table 1 and signal feature-related parameters listed in Table 2.

Considering the CSR Fourier transform and sampling frequency, the length of a single set signal n and the number of analysed signal sets K can be set to 1024 and 8, respectively. The convergence conditions for CSR (primal residuals of SISC r_x, dual residuals of SISC s_x, primal residuals of SIDL r_d, and dual residuals of SIDL s_d) are previously set to 0.001.

However, four parameters related to signal features cannot be set beforehand and should be adaptively tuned because the different measured signals contain different types of impulse responses. The types of impulse responses are closely related to the orders of the excited resonance frequencies induced by the wheelset bearing. The impulse

responses with different resonance frequencies and damping coefficients naturally have different impulse response lengths. The CSR-related regularization parameter reflects the sparsity of the impulse responses and is tightly related to the rotational speed and geometry parameters of the wheelset bearing. The selection of a suitable ADMM-related penalty parameter ρ is critical to obtaining a good convergence rate. There are two strategies for selecting ρ: the increasing parameter scheme [52] and the adaptive method [54]. Due to the good convergence performance of the adaptive method, the adaptive method is employed in this paper and can be described as [54]

$$\rho^{(j+1)} = \begin{cases} \tau \rho^{(j)}, & \text{if } \left\| r^{(j)} \right\|_2 > \mu \left\| s^{(j)} \right\|_2, \\ \tau^{-1} \rho^{(j)}, & \text{if } \left\| s^{(j)} \right\|_2 > \mu \left\| r^{(j)} \right\|_2, \\ \rho^{(j)}, & \text{otherwise}, \end{cases}$$

(13)

where $r^{(j)} = x^{(j)}(t) - y^{(j)}(t)$, $s^{(j)} = \rho^{(j)} \left(y^{(j-1)}(t) - y^{(j)}(t) \right)$ on the optimization process of equations (14) and (15), and $r^{(j)} = d^{(j)}(t) - g^{(j)}(t)$, $s^{(j)}(t) = \rho^{(j)} \left(\left(g^{(j-1)}(t) - g^{(j)}(t) \right) \right)$ are residuals of dual variables, respectively. τ and μ are constants, the typical values of which in [52] are $\tau = 2$ and $\mu = 10$. Thus, selecting the ADMM-related parameter ρ becomes a determination of the initial value $\rho^{(0)}$. According to the test in [52], $\rho^{(0)} = \xi \lambda$. For this paper, ξ was set to 100. Next, the other three parameters will be discussed in detail. To illustrate the influences caused by the different regularization parameters, the IRS vs. λ were extracted and are shown in Figure 2.

4. The Proposed ACSR

A novel fault detection named adaptive CSR (ACSR) is proposed in this paper. The procedure of the proposed method is shown in Figure 3. It mainly contains four steps:

(1) Estimating the types of impulse responses C

(2) Estimating the length of impulse responses P

(3) Estimating the regularization parameter λ

(4) Extracting IRS using ACSR with optimal parameters (C_o, P_o, λ_o)

4.1. Estimating the Types of Impulse Responses. According to the physical definition of an impulse response in equation (6), impulse responses are a function of resonance frequency ω_r and damping coefficient β. The resonance frequency can be used to identify different types of impulse responses [55]. Thus, the dominant frequencies of the inferred impulse responses from measured vibration signals can be used to estimate the types of impulse responses C. When the analysed signals are executed by CSR with c, c impulse responses can be obtained. Their amplitude-frequency spectra are obtained by Fourier transform taken on c learned impulse responses, respectively. The frequencies which the maximal amplitudes point to are the extracted main frequencies in the amplitude-frequency spectra. Therefore, the real value of C can be estimated by repeatedly executing CSR with $C = 2$,

TABLE 1: Boundary condition-related parameters.

Parameters	n	K	r_x	s_x	r_d	s_d
Physical meaning	The length of the single set signal	The number of signal sets	Primal residuals	Dual residuals	Primal residuals	Dual residuals
Preset values	1024	4	0.001	0.001	0.001	0.001

TABLE 2: Signal feature-related parameters.

Parameters	C	P	λ	ρ
Physical meaning	Types of impulse responses	The length of impulse responses	CSR-related regularization parameter	ADMM-related penalty parameter
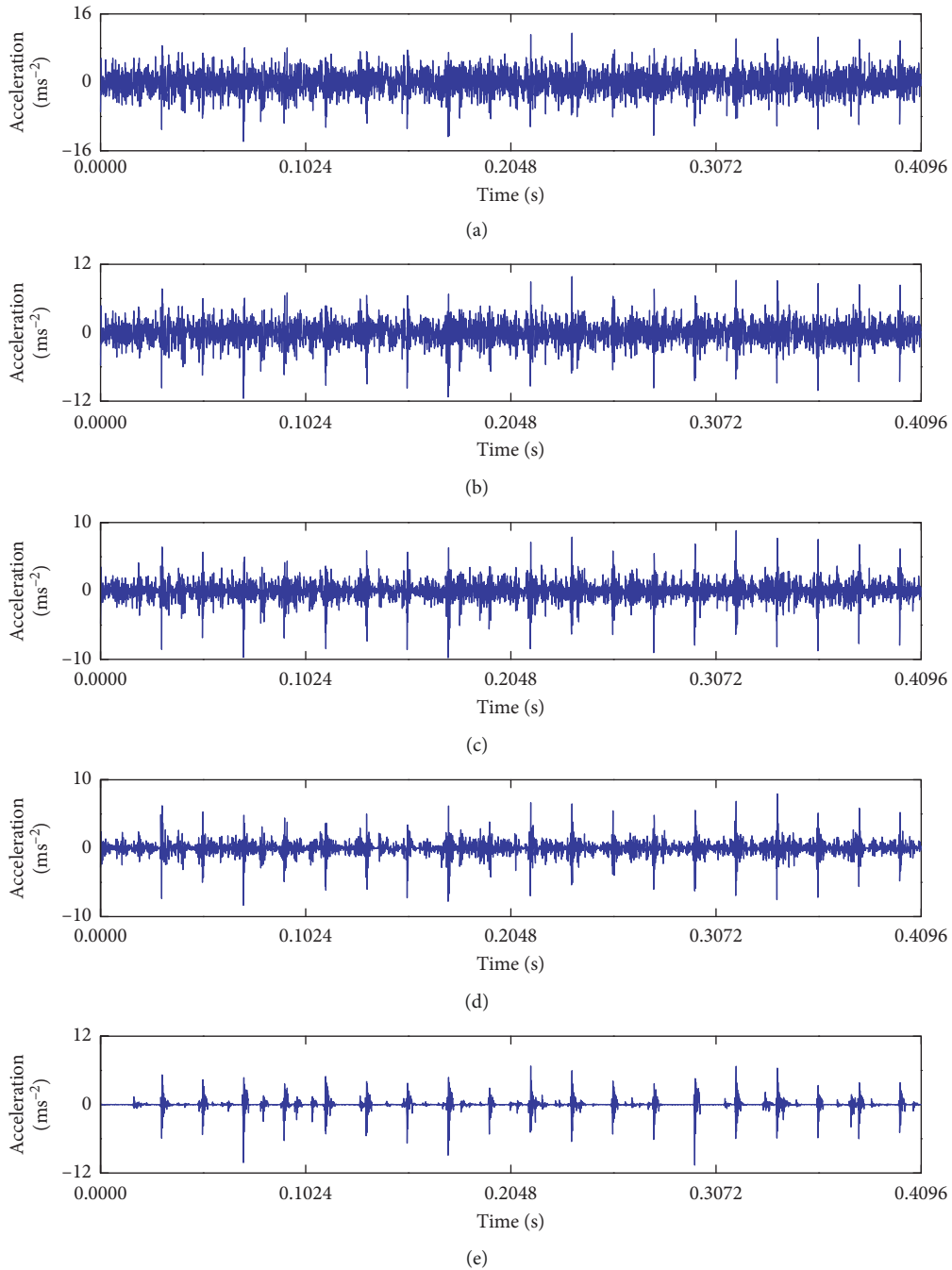
FIGURE 2: IRSs extracted using CSR for different parameters λ. (a) Original sound, kurtosis value = 40.3023. (b) $\lambda = 1$, kurtosis value = 69.9564. (c) $\lambda = 2$, kurtosis value = 98.9707. (d) $\lambda = 3$, kurtosis value = 135.1575. (e) $\lambda = 5$, kurtosis value = 183.7083.
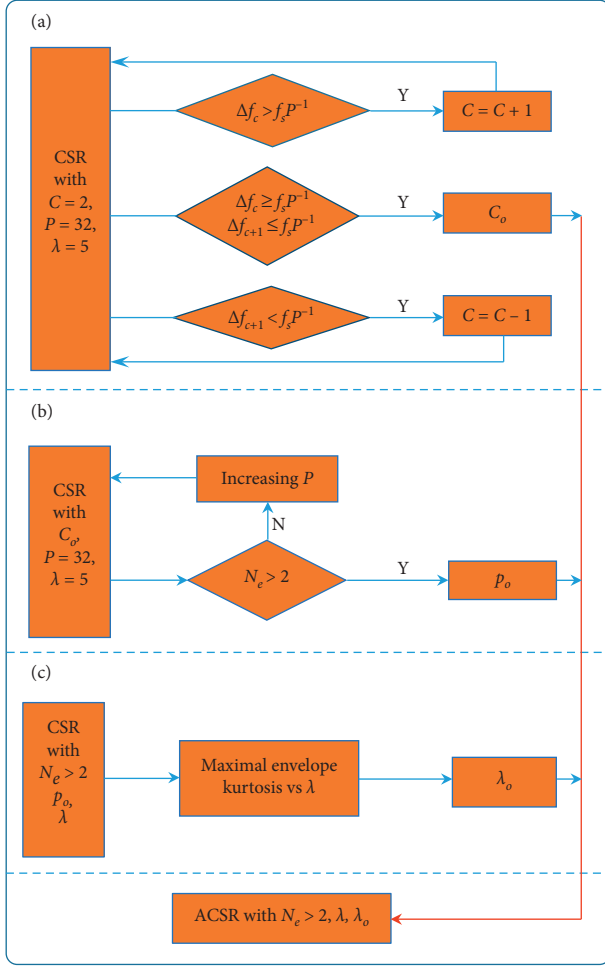
FIGURE 3: The procedure for the proposed ACSR. (a) Estimating the number of impulse responses. (b) Estimating the lengths of the impulse responses. (c) Estimating the regularization parameter.

FIGURE 4: Amplitude-frequency spectra for estimating the types of impulse responses in Case 1.

$P = 32$, and $\lambda = 5$. If the differences between any two dominant frequencies Δf_C are less than $f_s p^{-1}$ ($f_s p^{-1}$ means the frequency resolution of Fourier transform, f_s denotes the sampling frequency and is 10000 Hz, p is set to 32 in this paper, and $f_s p^{-1}$ is equal to 312.5 Hz), there are less than C types of impulse responses. Then, $C = C - 1$, CSR is repeatedly executed to learn $(C - 1)$ types of impulse responses, and the resulting dominant frequencies Δf_{C-1} are calculated until the frequency differences of any two dominant frequencies are more than the frequency resolution. Otherwise, $C = C + 1$, CSR is repeatedly executed to learn $(C + 1)$ types of impulse responses, their dominant frequencies are calculated until at least one frequency difference Δf_{C+1} is less than the frequency resolution, and $C = C - 1$ (in this paper, the initial value of C is set to 2). The final value of C is output as the estimated real values of the types of impulse responses embedded in the measured vibration signals. In Section 5, the detailed procedures for calculating the main frequencies when estimating C are shown in Figures 4 and 5 for Case 1 and Case 2, respectively.

According to the above-discussed rules for estimating the types of impulse responses C, in Case 1, when C was

initialized to 2, the difference between the two main frequencies $f_1 = 2508$ Hz and $f_2 = 2439$ Hz, Δf_2, was equal to 69 Hz and was less than the frequency resolution in Figure 4. Thus, $C = C - 1$. The type of impulse responses embedded in the simulation signals of Case 1 was 1 and was identical to the simulation setting. Similarly, there were two types of impulse response for Case 2, i.e., there were two IRSs for Case 2.

4.2. Estimating the Lengths of the Impulse Responses.
Although the full time-domain waveform of an impulse response is contaminated by measured noise, the high amplitude zone of an impulse response should have a larger SNR than the low amplitude zone in Figure 6.

If only two extreme values $y(t_1)$ and $y(t_2)$ can be precisely obtained in Figure 6, the two parameters: resonance frequency f_r and damping coefficient β for describing an impulse response are indirectly computed and are expressed as [55]

$$f_r = \frac{1}{t_n} = \frac{1}{t_2 - t_1},$$

$$\beta = \frac{1}{2\pi} \ln\left(\frac{y(t_1)}{y(t_1)}\right). \tag{14}$$

Therefore, extracting information on an impulse response should contain at least two extreme values, i.e., $N_e \geq 2$.

Ideally, when an impulse response properly has two extreme values, and its length is p, the time difference between two extreme values $t_n = t_1 - t_2 = p(f_s)^{-1}$. The resonance frequency associated with two extreme values f_r can be computed by the following formulation:
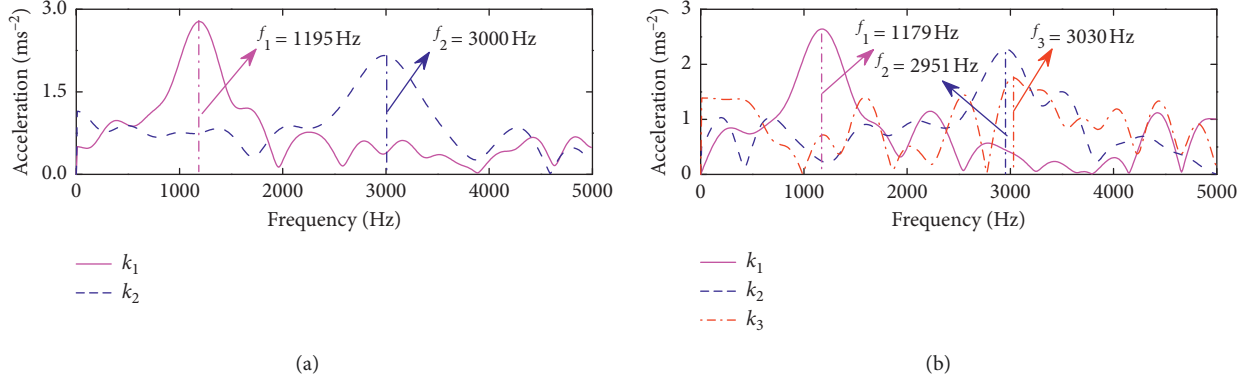
(a)

(b)

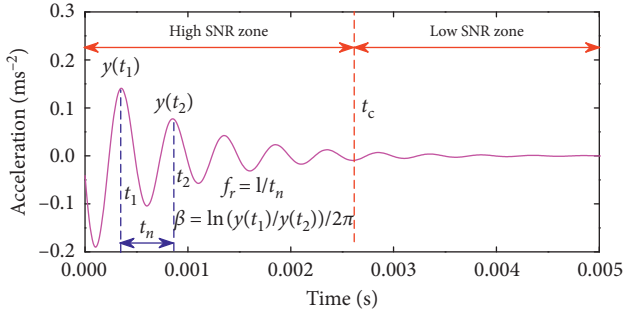Figure 5: Amplitude-frequency spectra for estimating the types of impulse responses in Case 2.

Figure 6: SNR distribution in an impulse response.

$$f_r = \frac{1}{t_n} = \frac{f_s}{P}. \tag{15}$$

Therefore, the resonance frequency of the leaned impulse response f_r should satisfy the following conditions:

$$f_{\min} \leq f_r = f_s p^{-1} \leq 0.5 f_s. \tag{16}$$

As a result, a possible length of p should satisfy the following relationship:

$$2 \leq P \leq f_s f_{\min}^{-1}, \tag{17}$$

where $f_{\min}$ is the minimum resonance frequency and is considered to be 300 Hz in this paper. Higher vibration frequencies necessitate smaller kernel function p lengths. When the resonance frequency is equal to $0.5 f_s$, the length of kernel P is only larger than 2. When the resonance frequency is 350 Hz ($f_s = 10$ kHz), P exceeds 28.5. Therefore, p is initially set to 32 in this paper. If there are not two complete maximum values in the learned kernel function, the lengths of the impulse responses P should be increased to represent the lower resonance frequency.

To illustrate the influence of different P on extracting IRSs, impulse responses with different P were learned and are shown in Figure 7, and the resulting IRSs were extracted and are shown in Figure 8. Through a careful analysis of Figures 8(a), 8(c), 8(e), and 8(g), the amplitudes and number of extracted impulse responses were hardly influenced by the length of impulse response. However, in Figures 8(b), 8(d), 8(f), and 8(h), the amplitudes of the envelope spectra of the IRS with

longer impulse responses are larger than those for the shorter impulse response. This is beneficial for detecting weak faults. Therefore, it is reasonable to initially set P to 32 in this paper.

4.3. Estimating Regularization Parameter. In Figure 2, it can be seen that the IRSs with different regularization parameters had different envelope spectra kurtoses and contained different noise levels. The number of impulse responses decreased with increasing regularization parameter. Conversely, the noise contained in the extracted IRS increases. Therefore, it is critical to determine a rational regularization parameter value suitable for the analysed vibration signals. If the quality of the extracted IRS can be measured by an index, automatic parameter selection is feasible.

Envelope spectrum kurtosis, as an effective measure index of impulsive feature distribution, could be used to accurately evaluate the information capacity of bearing faults [46]. As the number of periodic impulse responses increases, the envelope spectrum kurtosis of the extracted IRS becomes larger. However, when the number of periodic impulse responses exceeds the actual value, the kurtosis of its envelope spectrum will decrease with continuous increases in the regularization parameter λ. Therefore, the variance of the envelope spectrum kurtosis can reflect the impulsive feature distribution caused by different values of the regularization parameter. The maximum kurtosis of the envelope spectrum points to the desired result of the regularization parameter. The envelope spectrum kurtosis vs. λ in Section 5 for Case 1 and Case 2 were calculated and are shown in Figures 9 and 10, respectively.

4.4. Extracting IRSs Using ACSR with Optimal Parameters. After estimating the three parameters, an adaptive version of CSR, which is called the ACSR with optimal parameters (C_o, P_o, λ_o), was obtained. IRS can then be adaptively extracted by ACSR. Its concrete steps are listed as follows:

(1) Learning the C types of the impulse responses from the partition signals a_k, $k \in [1, K]$ using equations (11) and (17) with the optimal parameters (C_o, P_o, λ_o)

(2) Inferring the time-location coefficients of the measured signals a using the sparse representation based on the learned impulse responses using equation (12) with the optimal parameters (C_o, P_o, λ_o)
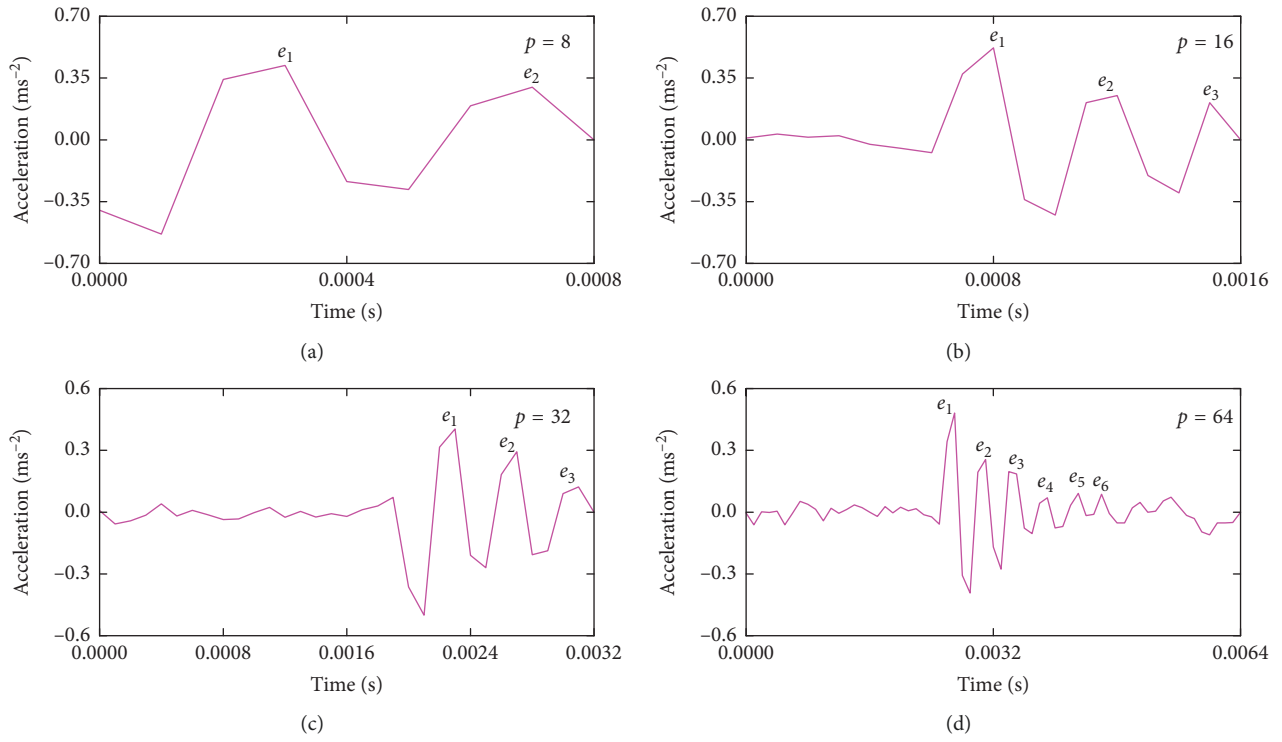
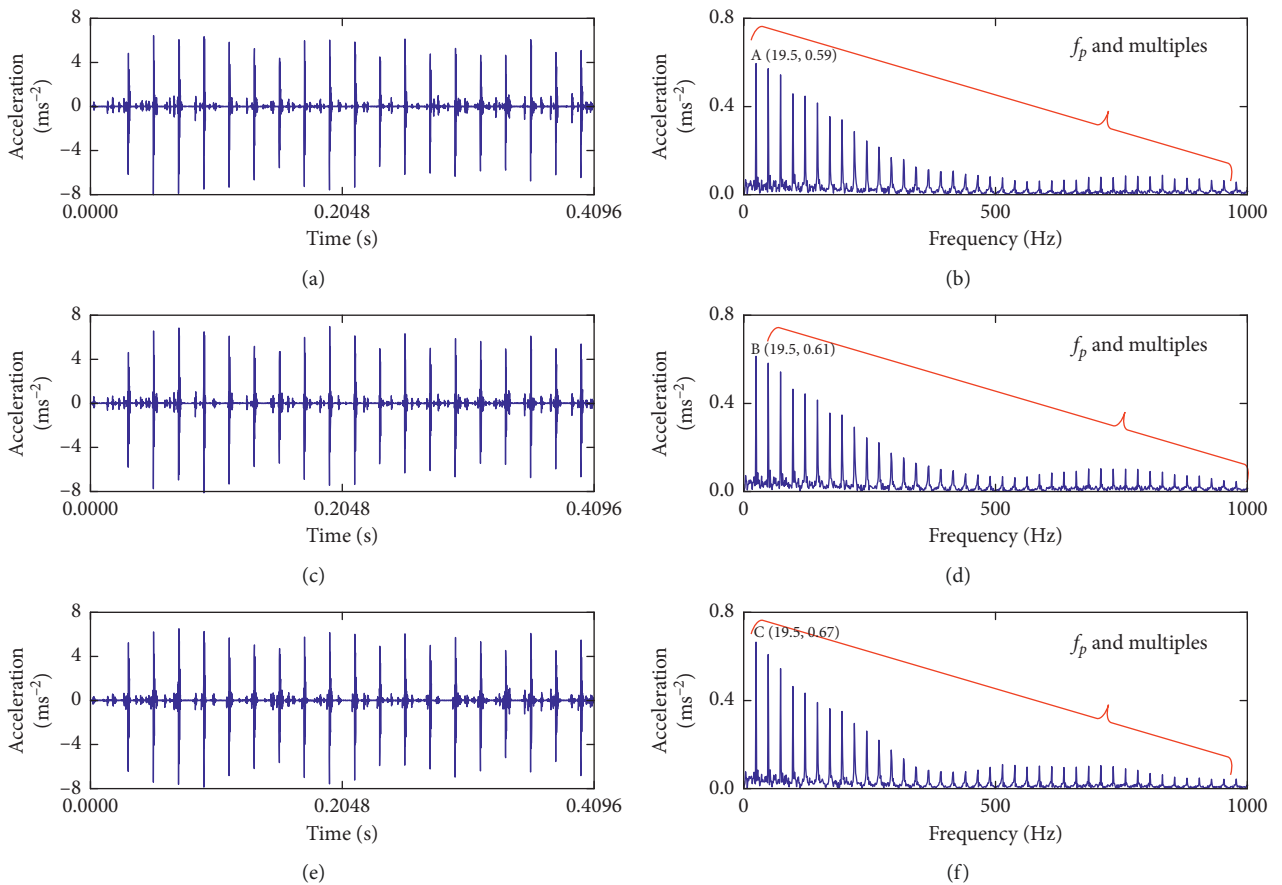FIGURE 7: Impulse responses with different lengths P: (a) $P = 8$; (b) $P = 16$; (c) $P = 32$; (d) $P = 64$.
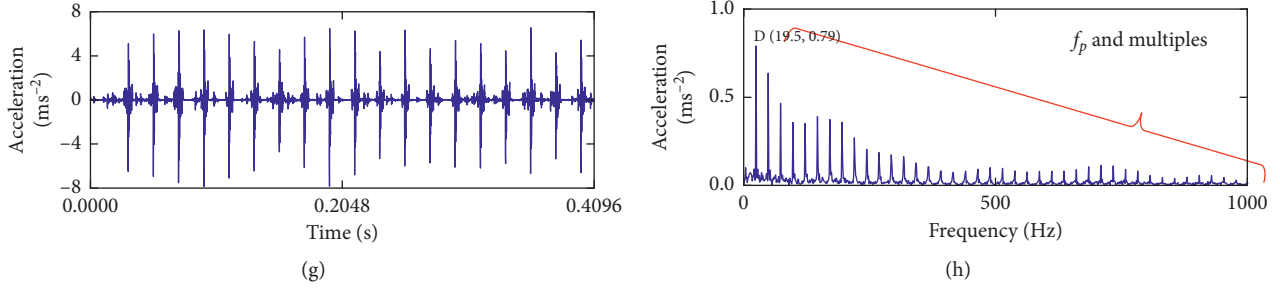

FIGURE 8: Continued.

(g)

(h)

Figure 8: IRSs with different impulse lengths P: (a) $P = 8$; (b) $P = 16$; (c) $P = 32$; (d) $P = 64$.
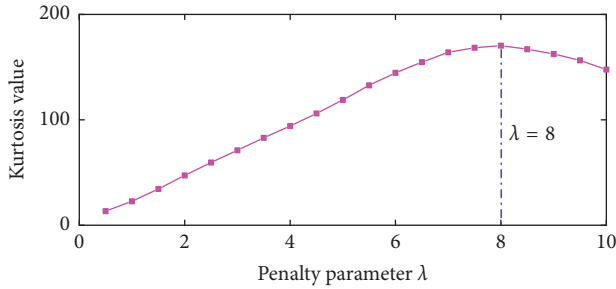

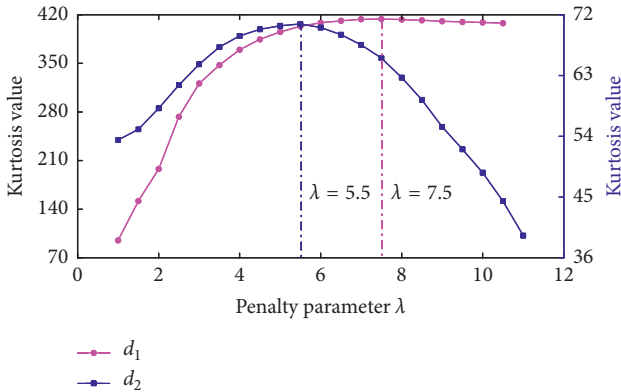Figure 9: Envelope spectra kurtosis for different values of regularization parameter for Case 1.

Figure 10: Envelope spectra kurtosis for different values of regularization parameter for Case 2.

(3) Extracting IRSs convolving the learned impulse responses and resulting inferred time-location coefficients using equation (12).

5. Simulation Validation

To illustrate the effectiveness of the proposed method, two classes of simulation signals with different resonances frequencies are introduced in this section.

5.1. Case 1: Simulation Signals with One Type of Impulse Response. The simulation signals with a type of impulse responses contained an IRS. IRS can be realized by the instantiation of the related parameters in equation (9). The parameters are listed in Table 3.

Table 3: Simulation parameters for Case 1.

A_m (m)	β	f_r (Hz)	f_p (Hz)	M	C
0.00000005	1300	2500	49.1	19	1

The simulated IRS at a SNR of $-10\,dB$ is shown in Figure 11. According to Figure 4 and the rules for determining the number of types of impulse responses, there should be one type of impulse response ($C = 1$). When P was initially set to 32, the number of maximum values of the learned impulse response shown in Figure 12 was more than 2. Therefore, the length of the kernel function in Case 1, P, was set to 32. Finally, the envelope spectra kurtosis vs. λ are shown in Figure 9, and the optimal target sparsity λ_o was 8. The learned impulse response and IRS extracted using ACSR are shown in Figure 12. Fourier and Hilbert envelope spectra of the extracted IRS are shown in Figure 13.

To illustrate the effectiveness of the proposed extracted IRS method, two well-known fault detection methods, spectral kurtosis [56] and EEMD [18], were used to analyse the same signal in Figure 11(a), and the obtained results are shown in Figures 14 and 15, respectively (to save space, the first intrinsic mode function (IMF), which was much more impulsive, is shown).

By comparing the time-domain waveform of the extracted IRS in Figure 12(b) with Figures 14(a) and 15(a), the proposed ACSR was able to extract the IRS containing 19 impulse responses from the simulation's noisy signals. However, the other two comparative methods failed to extract distinct IRSs, and there was strong noise between adjacent impulse responses. This shows that the proposed ACSR surely can characterize the kinematic process of the bearing faults. By comparing the of envelope spectra in Figures 13(b), 14(b), and 15(b), it can be seen that both the amplitude and harmonic number of the fault-characteristic frequency were larger than for the other two methods. In addition, the fault-characteristic frequency harmonics obtained by EEMD and spectra kurtosis were confused by some uncorrelated spectra lines, indicating that the proposed ACSR had good performance when extracting IRSs caused by wheelset bearings.

5.2. Case 2: Simulation Signals with Two Types of Impulse Responses. The simulation signals with two types of impulse responses contained two IRSs. Each of the IRSs can be
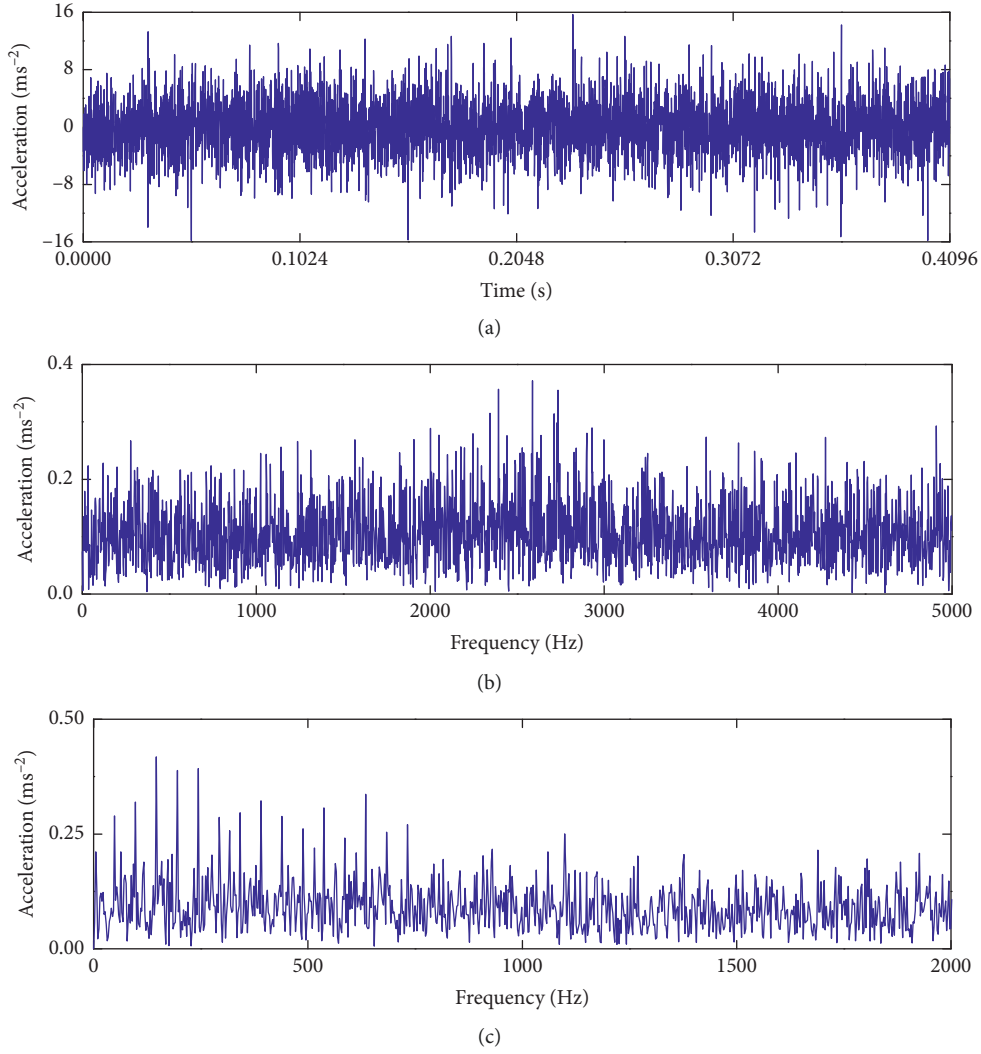
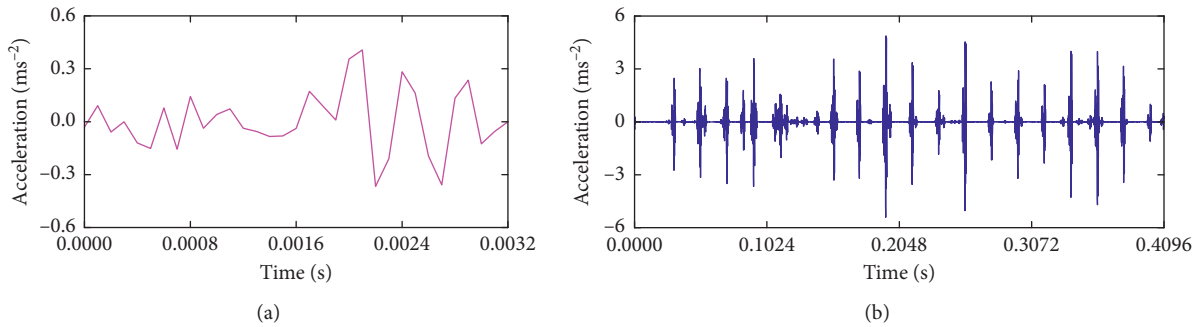FIGURE 11: Simulated IRS: (a) original signals; (b) its Fourier spectra; (c) its envelope spectrum.

FIGURE 12: Extracted IRS: (a) learned impulse response; (b) extracted IRS.

individually realized by instantiating the related parameters in equation (9). The parameters are listed in Tables 4 and 5.

The simulated IRSs at an SNR of −1 dB are shown in Figure 16(a). The Fourier and envelope spectra of the simulated signals are shown in Figures 16(b) and 16(c), respectively. The fault frequency f_1 was obtained, but fault frequency f_2 almost cannot be seen in the envelope spectra in Figure 16(c). The proposed ACSR was used to process the simulation signals in Figure 16(a). When C was set as 2, two main frequencies can be seen in Figure 5(a); f_1 was equal to
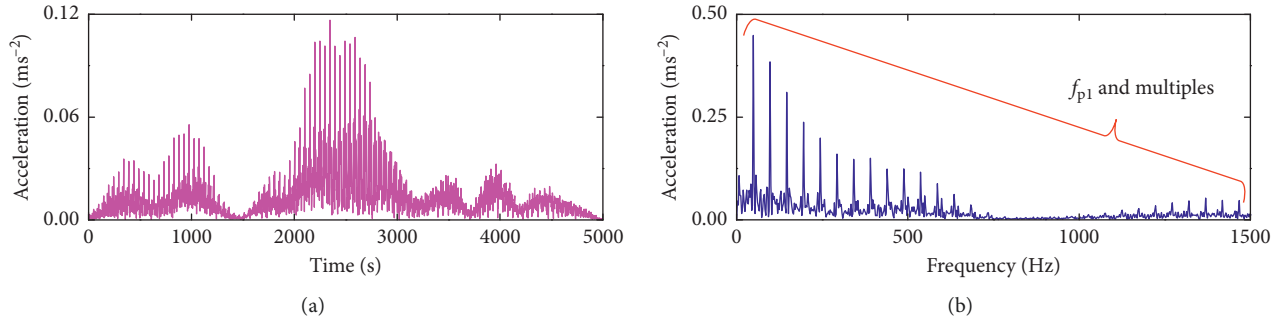
FIGURE 13: (a) Fourier spectra of the extracted IRS; (b) Hilbert envelope spectra of the extracted IRS.
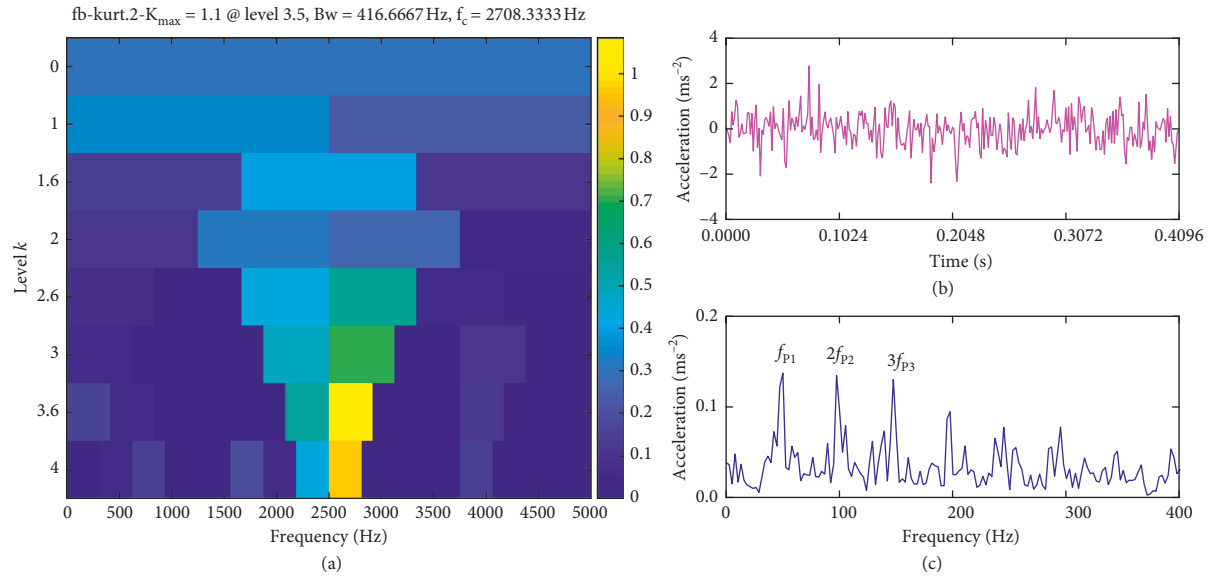
FIGURE 14: Results obtained by (a) spectra kurtosis: (b) purified signal; (c) Hilbert envelope spectra.
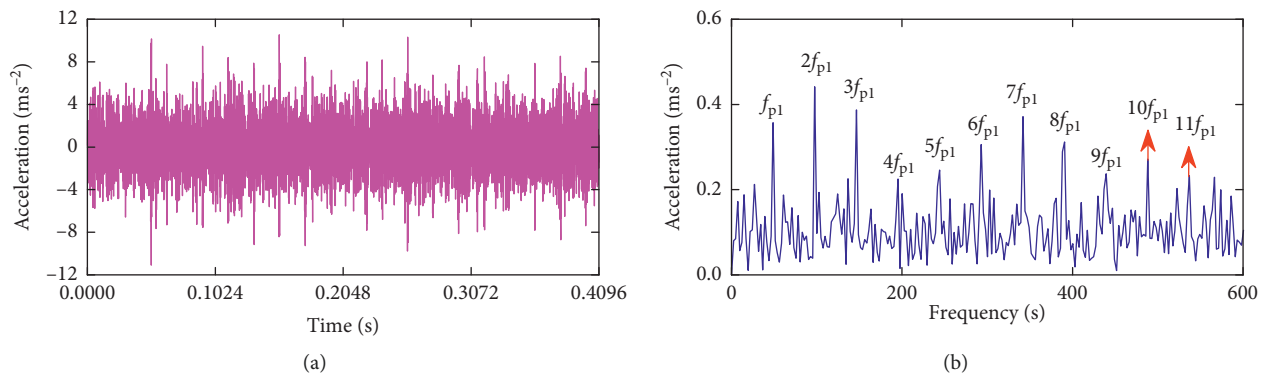
FIGURE 15: Results obtained by EEMD: (a) first IMF; (b) envelope spectra of the first IMF.

TABLE 4: Simulation parameters in the first IRS.

A_1 (m)	β_1	f_{r1} (Hz)	f_1 (Hz)	M_1	C
0.00000005	1500	3000	32.7	12	2

TABLE 5: Simulation parameters in the second IRS.

A_2 (m)	β_2	f_{r2} (Hz)	f_2 (Hz)	M_2	C
0.00000003	1000	1200	81	31	2

1195 Hz and f_2 was equal to 3000 Hz. The difference in the two main frequencies was 1805 Hz, which was greater than the frequency resolution of 312.5 Hz. According to the rule

for estimating the number of impulse responses, $C = C + 1$ and was set to 3, and three main frequencies can be seen in Figure 5(b). The difference between the main frequency f_2
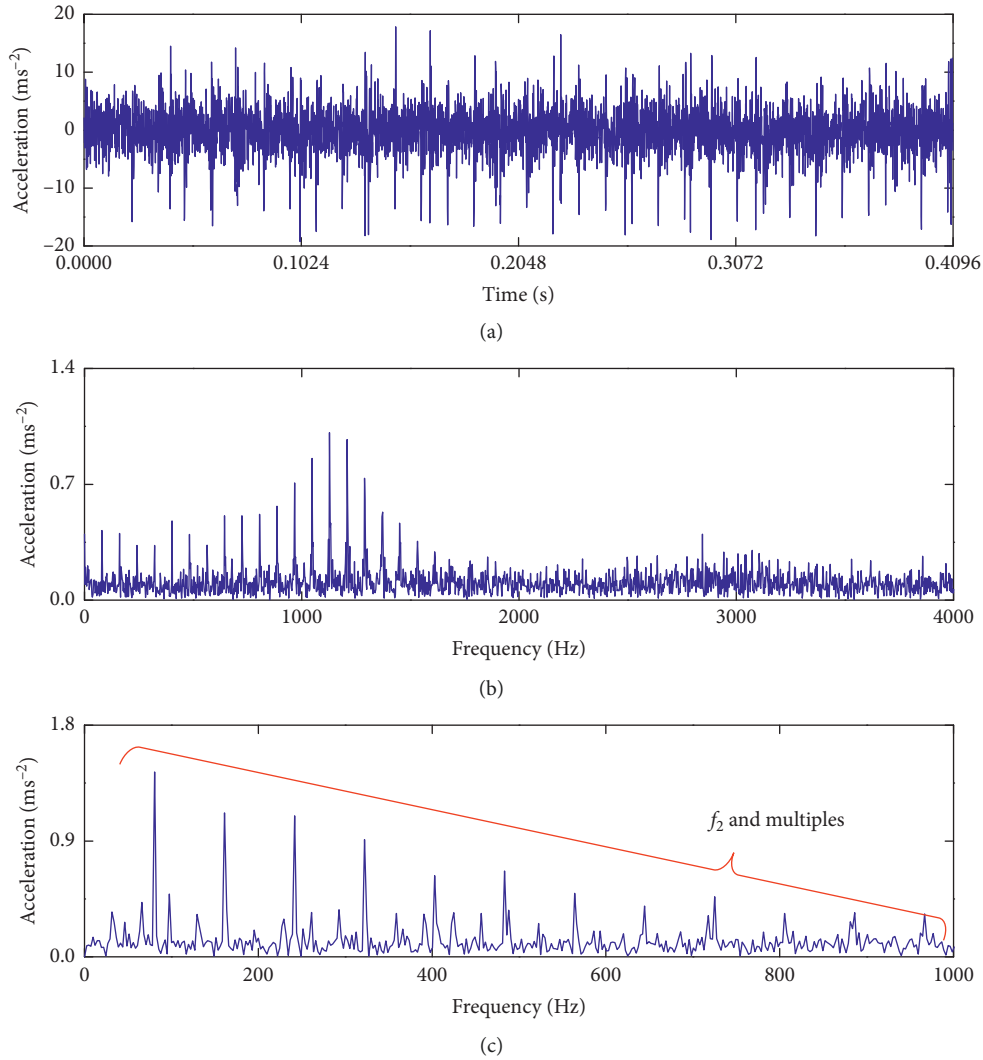
FIGURE 16: Simulated IRSs: (a) original signal; (b) its Fourier spectra; (c) its envelope spectra.

(2951 Hz) and f_3 (3030 Hz) was less than the frequency resolution of 312.5 Hz. Therefore, the number of the type of impulse responses should have been 2. When P was set to 32, the number of maximum values of the learned impulse responses shown in Figure 17 exceeded 2. Therefore, P was set to 32 for Case 2. The curves of the envelope spectra kurtosis vs. different values of the regularization parameter are shown in Figure 10. The regularization parameters λ_{1o} and λ_{2o} for impulse responses 1 and 2 were 7.5 and 5.5, respectively. The IRSs extracted using ACSR are shown in Figure 17. The Fourier and envelope spectra of IRS1 and IRS2 are shown in Figure 18.

To illustrate the advantage of the proposed ACSR, two comparative methods (spectra kurtosis and EEMD) were employed to analyse the simulation signals in Figure 16(a), and the results are shown in Figures 19 and 20, respectively (the previous two IMFs are shown due to their higher impulsiveness). On the one hand, by comparing the extracted time-domain waveforms in Figures 17(b) and 17(d), 19(a), and 20(a) and 20(c), it can

be seen that the proposed ACSR can clearly extract IRSs for which there is almost any amount of noise in two adjacent impulse responses and can isolate two ISRs. This is advantageous for characterizing the actions of wheelset-bearing faults. On the other hand, from the envelope spectra in Figures 18(b) and 18(d), 19(b), and 20(b) and 20(d), the amplitudes and harmonic numbers of the fault-characteristic frequencies of the IRSs extracted using the proposed ACSR were larger than those for the other two comparative methods, and the envelope spectra of the IRSs extracted using the proposed ACSR were almost unconfused by noise, indicating that the fault detection performance of ACSR was good.

6. Experimental Verification

To further test the effectiveness and fault detection performance of the proposed ACSR for practical vibration signals, bench and running tests, respectively, were carried out.
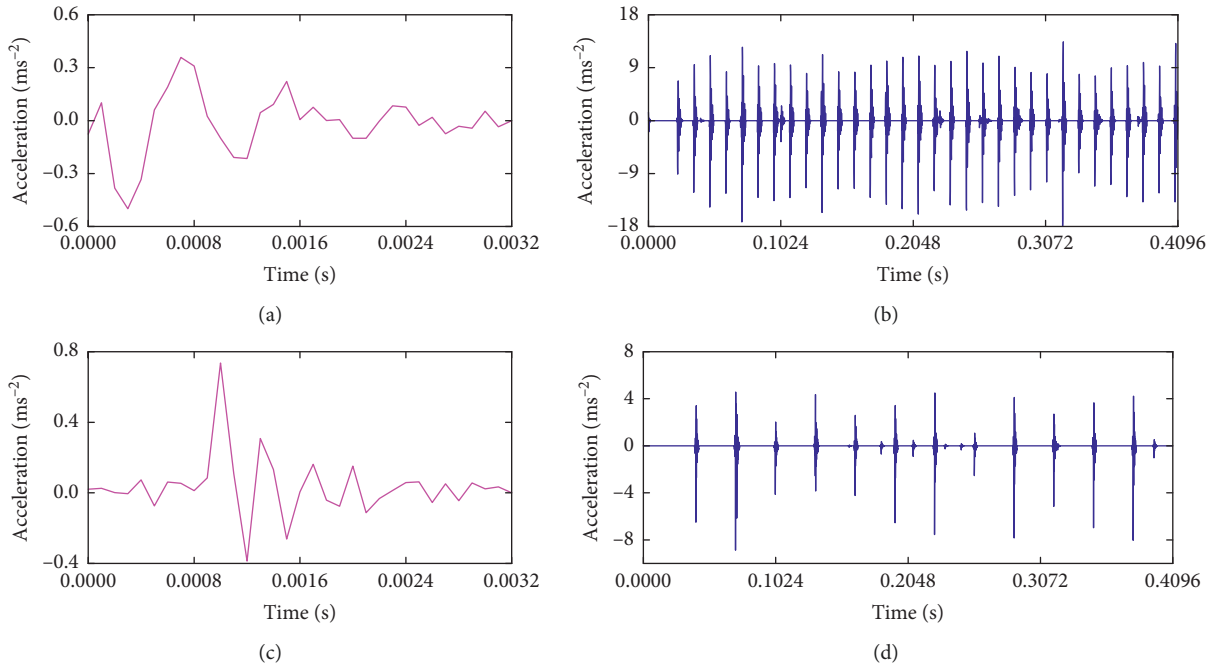
FIGURE 17: IRSs extracted using ACSR: (a) learned impulse response 1; (b) extracted IRS1; (c) learned impulse response 2; (d) extracted IRS2.
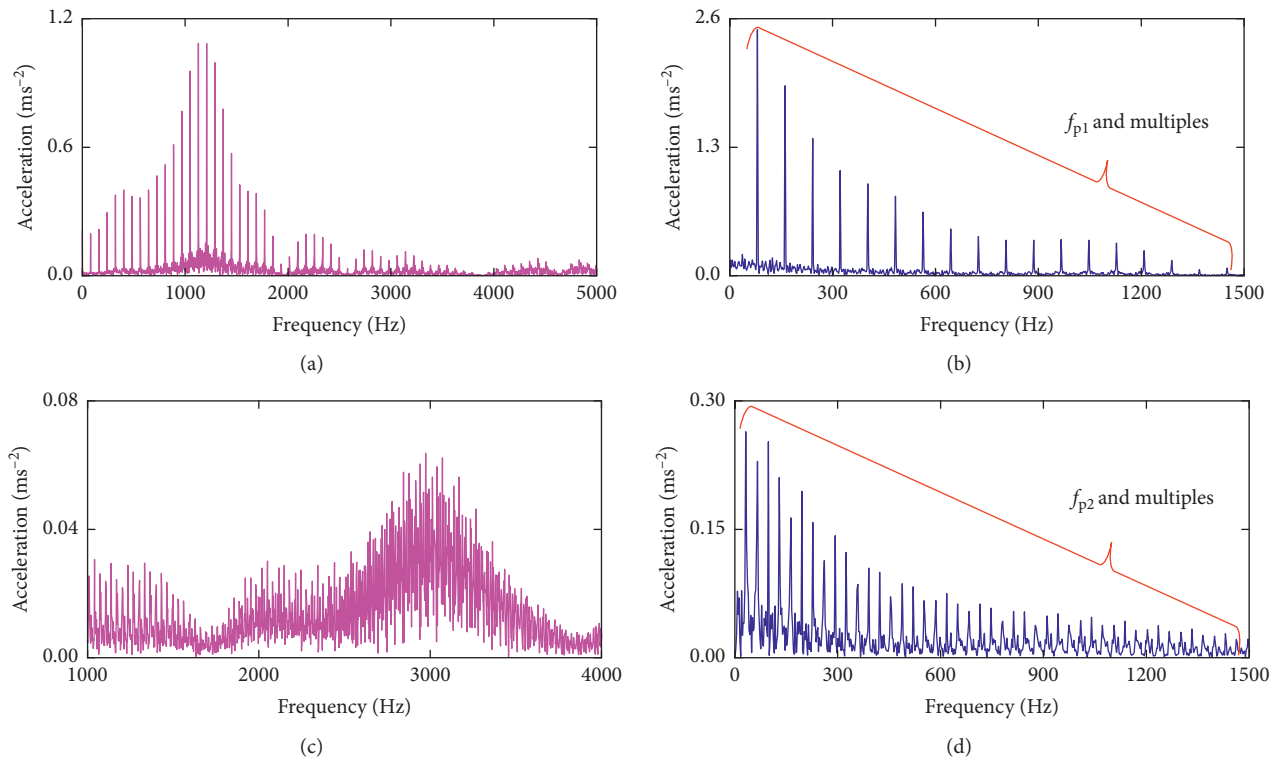

FIGURE 18: (a) Fourier spectra of the extracted IRS1; (b) Hilbert envelope spectra of the extracted IRS1; (c) Fourier spectra of the extracted IRS2; (d) Hilbert envelope spectra of the extracted IRS2.

6.1. Bench Test. The wheelset-bearing test bench used for the practical test of the proposed method is shown in Figure 21. The test bench consisted of a motor, driving wheel, loading device, wheelset, and axle box. The motor delivered driving power at different motor speeds. The driving power was conveyed to the driving wheel through rubber belts. The traction power of the driving wheel was then transmitted to the wheelset. Faults on the outer and inner races, which are
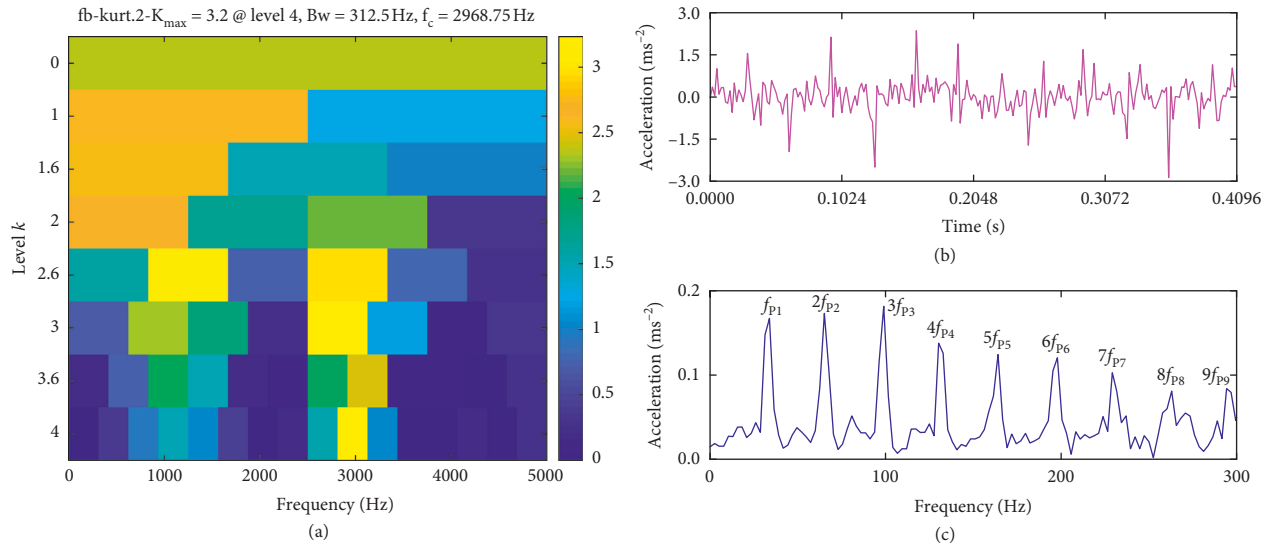
FIGURE 19: Processing results obtained using spectra kurtosis: (a) fast kurtogram; (b) purified signal; (c) Hilbert envelope spectra.
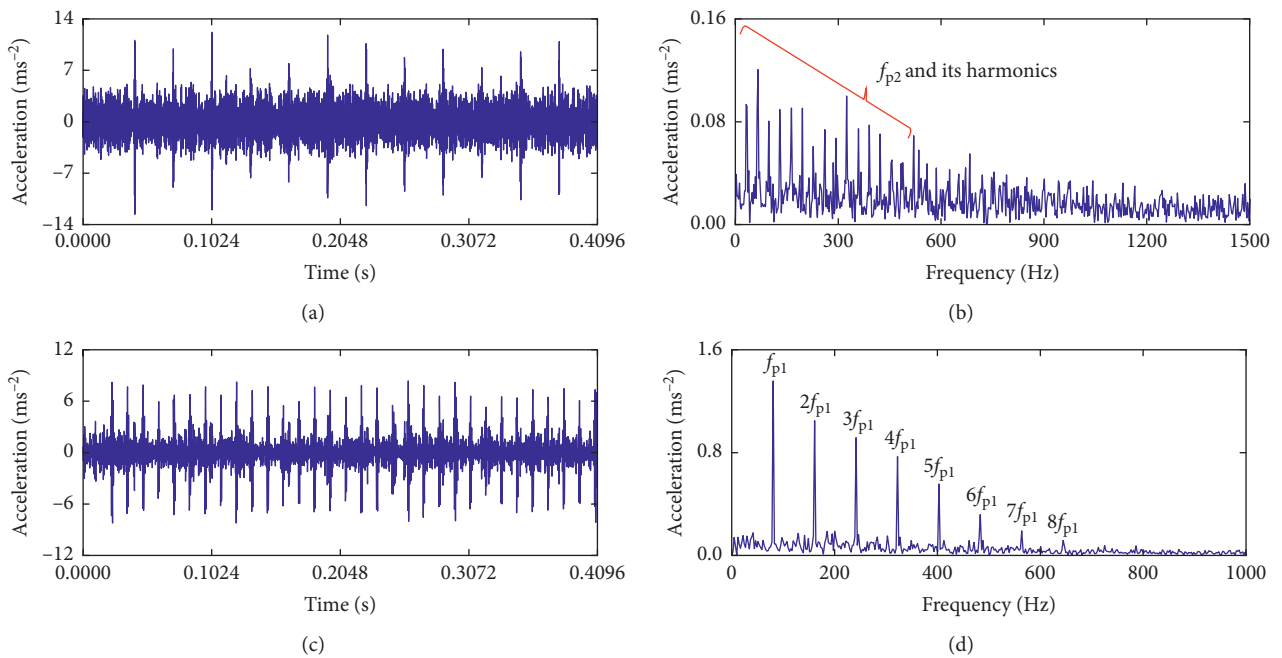
FIGURE 20: Processing results obtained using EEMD: (a) IMF1; (b) envelope spectra of IMF1; (c) IMF2; (d) envelope spectra of IMF2.
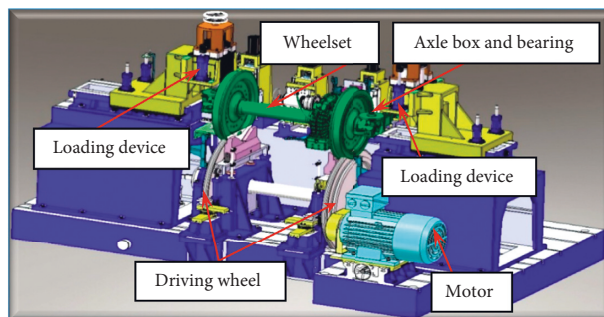

FIGURE 21: Schematic sketch of the wheelset-bearing test bench.

(a)   (b)
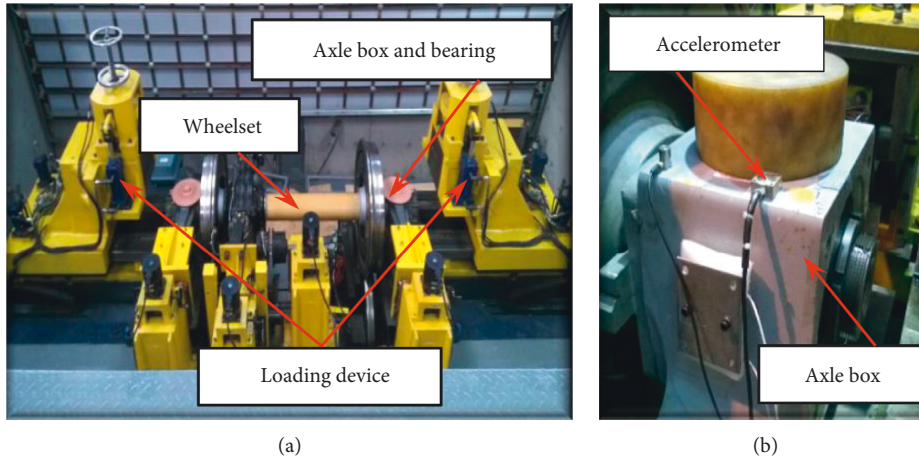
FIGURE 22: Photo of the faults (a) on the outer race and (b) on the inner race.

(a)   (b)

FIGURE 23: Photos of (a) test bench and (b) measurement sensor.

shown in Figure 22, were introduced to the roller bearing installed in the axle box (see Figure 23(b) for the axle box and the accelerometer mounted on it). Figure 23(a) also shows a photo of the test bench. The fault bearing parameters are listed in Table 6.

The vibration signals collected from the wheelset-bearing system test bench are shown in Figure 24 when the outer race of the wheelset bearing had the faults shown in Figure 22. The rotational frequency of the wheelset, f_o, was 15.4 Hz and corresponded to the running speed of 150 kmh^{-1}. Its sampling frequency was 10 kHz. In the envelope spectra of the measured signals, there was only the basic frequency of the inner race fault, f_{BPFI}, but there were no spectral lines about the outer-race fault-characteristic frequency f_{BPFI}. f_{BPFI} and f_{BPFO} are expressed as follows:

$$f_{BPFI} == \frac{N_b}{2}\left(1 + \frac{B_d}{P_d}\cos(\phi)\right)f_o,$$

$$f_{BPFO} = \frac{N_b}{2}\left(1 - \frac{B_d}{P_d}\cos(\phi)\right)f_o, \tag{18}$$

where N_b is the roller number, P_d is the pitch diameter, B_d is the roller diameter, ϕ is the contact angle, and f_o is the rotational speed of the wheelset. According to the parameters listed in Table 6, the outer-race fault-characteristic

frequency f_{BPFO} was 124.9 Hz, and the inner-race fault-characteristic frequency, f_{BPFI}, was 168.1 Hz.

The proposed ACSR was used to analyse the measured vibration signals in Figure 24(a). The main frequencies for estimating the types of impulse responses embedded in measured vibration signals were calculated and are shown in Figure 25.

According to the rule for estimating the types of impulse responses, there should have been 2 types of impulse responses. When P was initially set to 32, the extreme points of the learned impulse responses in Figures 26(a) and 26(c) were larger than 2. Therefore, P was set to 32. The curve of the envelope spectra kurtosis with different values of the regularization parameters is shown in Figure 27. The two optimal regularization parameters related to the impulse responses d_1 and d_2 were $\lambda_{1o} = 19.5$ and $\lambda_{2o} = 20.5$, respectively.

The learned impulse responses and extracted IRSs are shown in Figure 26. The envelope spectra of IRS1 and IRS2 are shown in Figures 28 and 29, respectively. As seen in Figures 26(b) and 26(d), IRSs induced by the outer fault and inner fault were clearly extracted using the proposed ACSR. The fault-characteristic frequencies of the outer and inner faults and their harmonics can be seen in Figures 28(a) and 28(b), respectively. To illustrate the increased ability of the proposed ACSR for detecting wheelset-bearing faults, spectra kurtosis and EEMD were employed to analyse the

TABLE 6: Rolling bearing test parameters.

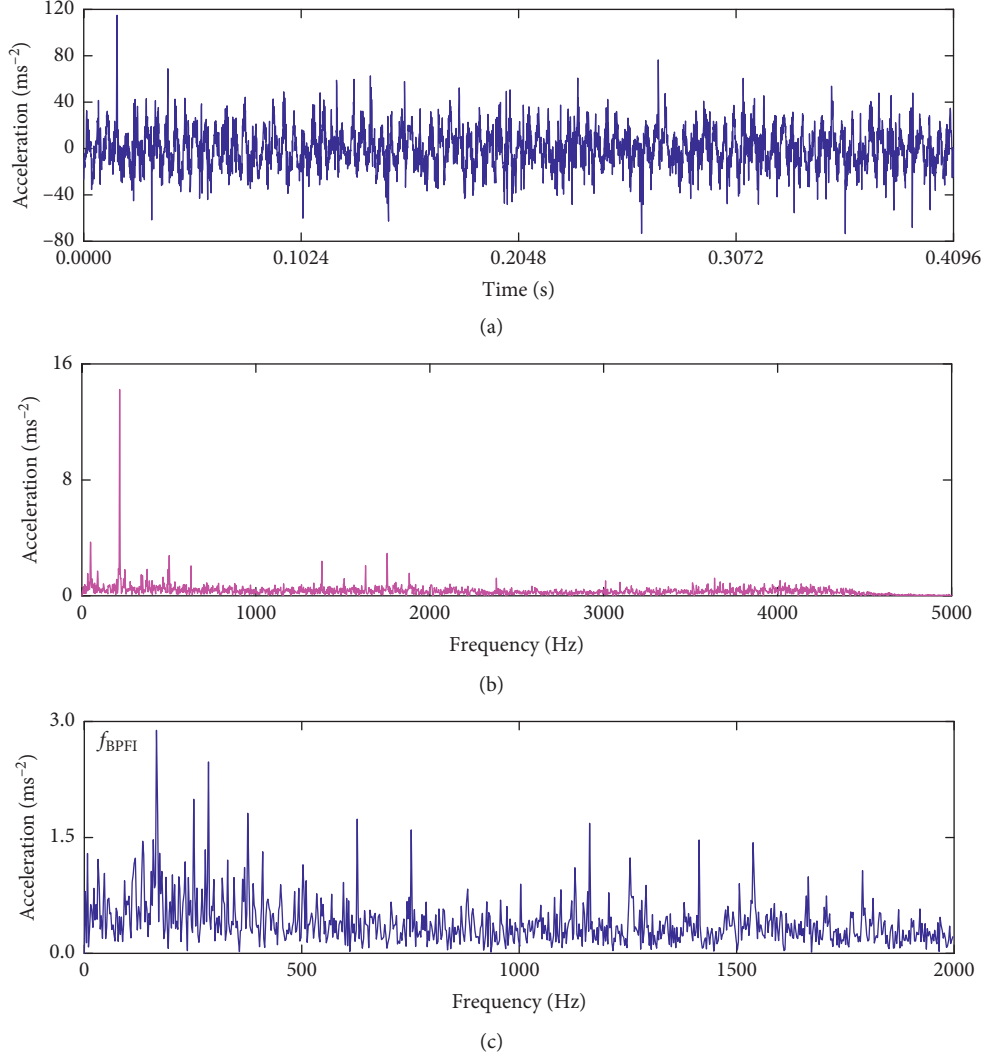| Roller number	Roller diameter (mm)	Pitch diameter (mm)	Contact angle (rad)
19	26.9	180	0.1571

(a)

(b)

(c)

FIGURE 24: Measured signals: (a) time-domain waveform; (b) Fourier spectra of (a); (c) envelope spectra of (a).

same vibration signals in Figure 24(a), and the results of the analyses are shown in Figures 30 and 31, respectively. For EEMD, IMF1, IMF3, and most of the impulsive features are shown. The two comparative methods only found the fault on the outer race.

6.2. Real-Life Running Test in a High-Speed Train. To test the fault detection performance of ACSR under the practical running conditions of a high-speed train, an accelerometer was installed on the axle box cover of a high-speed train, as shown in Figure 32. The tested wheelset bearing used with the high-speed train was identical to that used with the bench test, and its parameters are listed in Table 6. When the roller of the wheelset bearing had a defect, the vibration acceleration signals were collected from the axle box cover of

the running high-speed train at a speed of $200 \, \text{kmh}^{-1}$ (the corresponding rotational frequency of wheelset was 20.6 Hz) and are shown in Figure 33. The sampling frequency of the measured vibration signals was 10 kHz. The roller-characteristic frequency f_{BSF} was 134.6 Hz. The roller-characteristic frequency is defined as

$$f_{\text{BSF}} = \frac{P_d}{B_d} \left(1 - \frac{B_d}{P_d} (\cos \phi)^2 \right) f_o. \tag{19}$$

As seen in Figure 33(c), the roller-characteristic frequency and its harmonics could not be discovered from the envelope spectra of the original measured signals for fault detection.

ACSR was therefore used to process the measured signals in Figure 33(a). The main frequencies for estimating the types of impulse responses were calculated
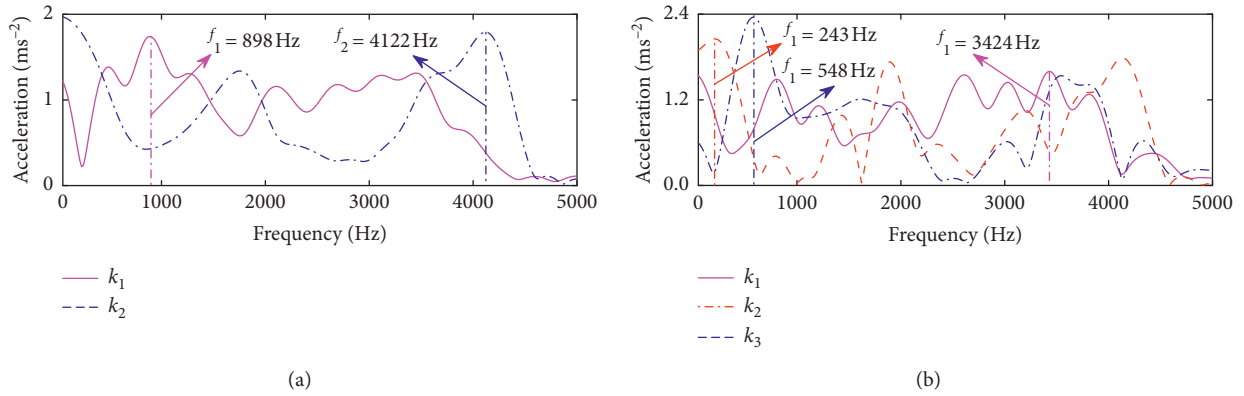
(a)

(b)

FIGURE 25: Main frequencies for estimating the impulse response types.
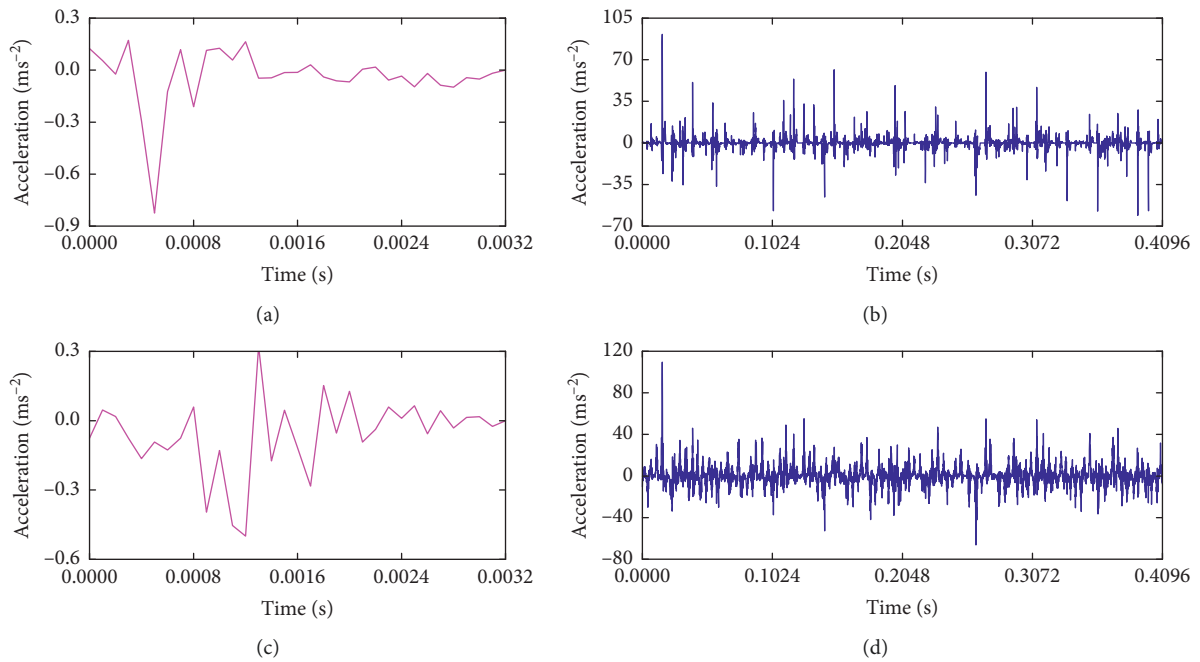

(a)

(b)

(c)

(d)

FIGURE 26: Learned impulse responses and extracted IRSs: (a) the first type of impulse response (IR1); (b) IRS1 of IR1; (c) the second type of impulse response IR2; (d) IRS2 of IR2.
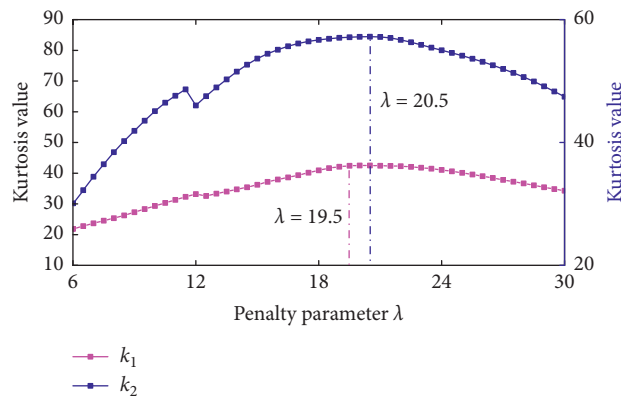

FIGURE 27: Envelope spectra kurtoses for different values of the regularization parameter.
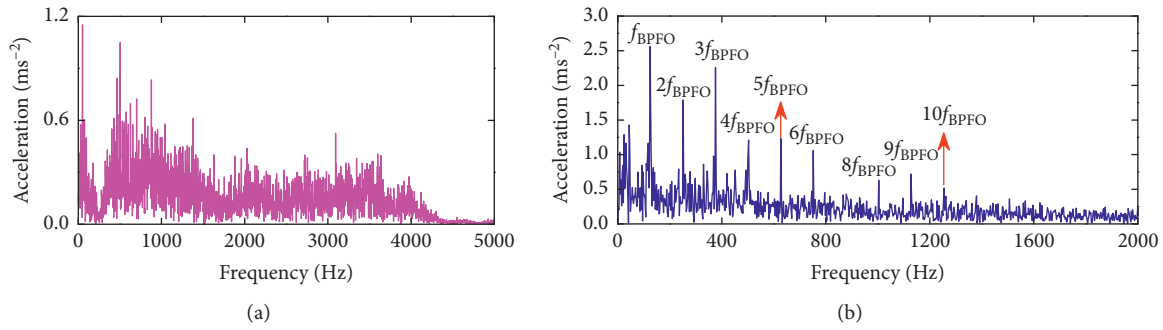
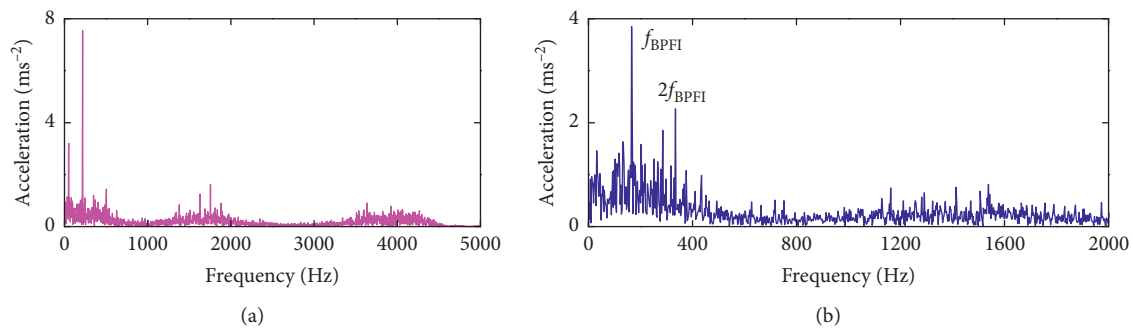FIGURE 28: (a) Fourier spectra of IRS1; (b) envelope spectra of IRS1.

FIGURE 29: (a) Fourier spectra of IRS2; (b) envelope spectra of IRS2.
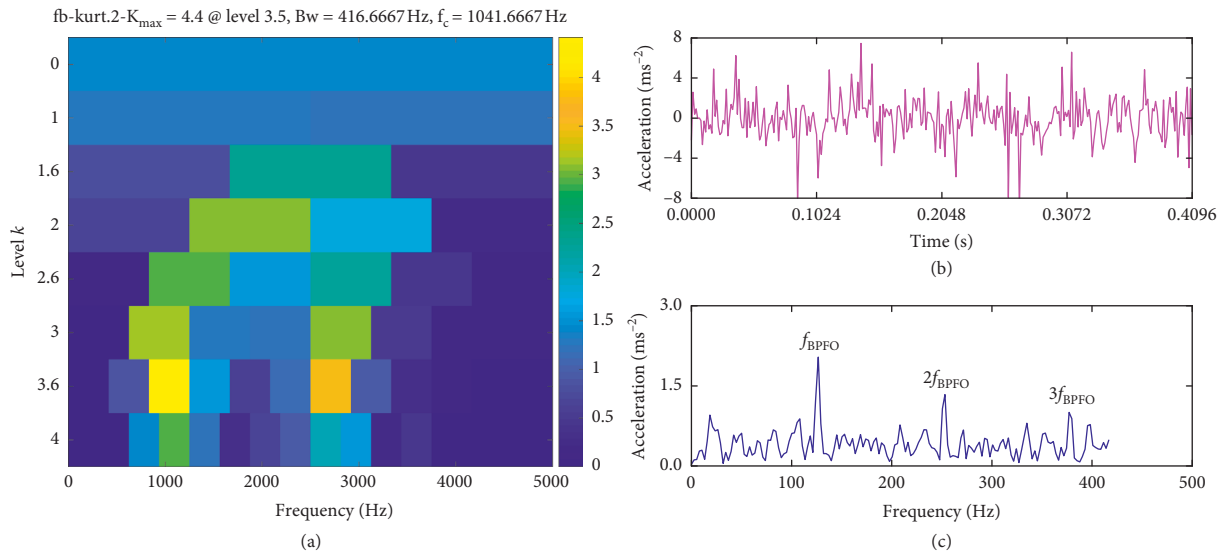

FIGURE 30: Processing results obtained using (a) spectra kurtosis: (b) purified signal; (c) Hilbert envelope spectra.

and are shown in Figure 34. According to the rule for estimating the optimal number of types of impulse responses, C was equal to 2. In the learned impulse response in Figure 35, the number of maximum extreme values for each type of learned impulse responses exceeded 2. The length of the kernel function was therefore set to 32. The curves for selecting the optimal values of the regularization parameter were calculated and are shown in Figure 36. The regularization parameter

related to IRS1, λ_{1o}, was 20, and that related to IRS2, λ_{2o}, was 13. The learned impulse responses and extracted IRSs obtained using ACSR are shown in Figure 35. The Fourier and envelope spectra of the extracted IRSs are shown in Figure 37. In Figure 35(d), the vibration behaviour when the faulty roller alternatively entered and left the bearing zone (BZ) or nonbearing zone can be distinctly seen. The fault-characteristic frequency of the roller and its harmonics were extracted for detecting the wheelset-bearing
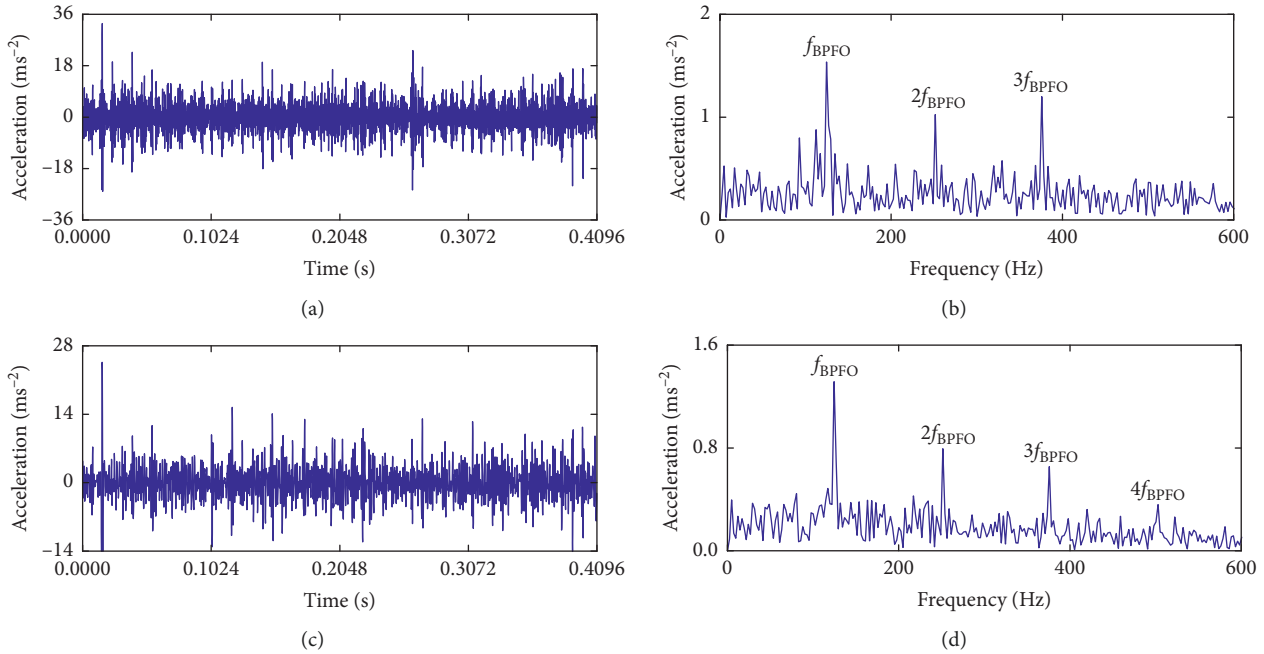
(a)

(b)

(c)

(d)

Figure 31: Processing results obtained using EEMD: (a) IMF1; (b) envelope spectra of IMF1; (c) IMF3; (d) envelope spectra of IMF3.
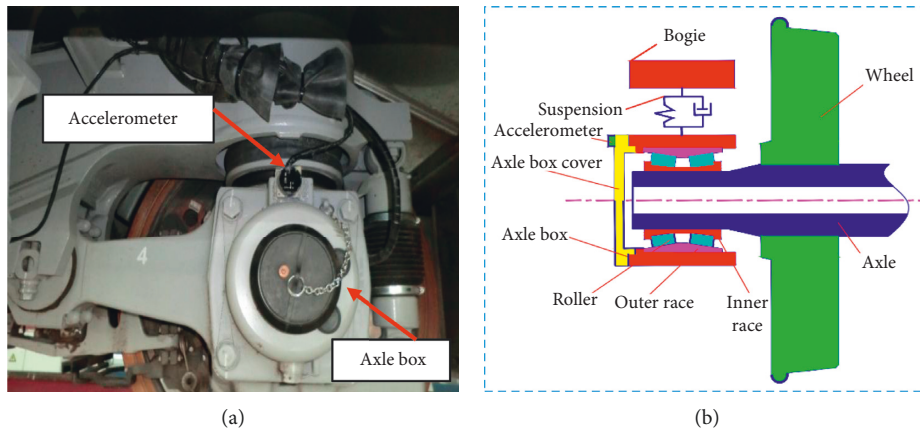

(a)

(b)

Figure 32: (a) Measurement sensor installed on the axle box cover; (b) structure diagram of the wheelset bearing.
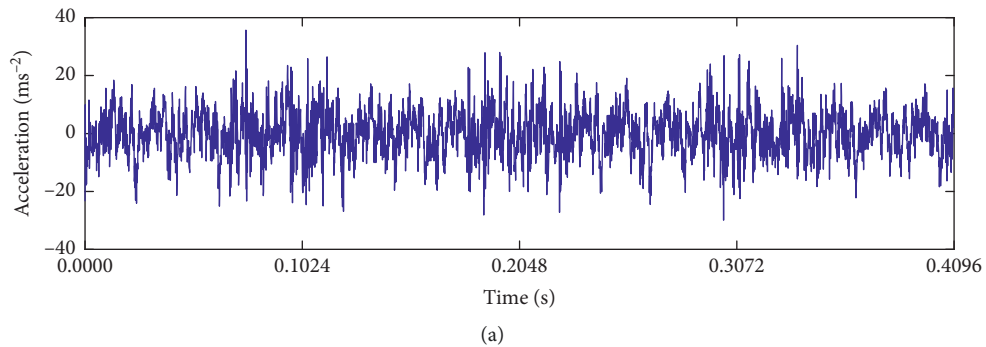

(a)

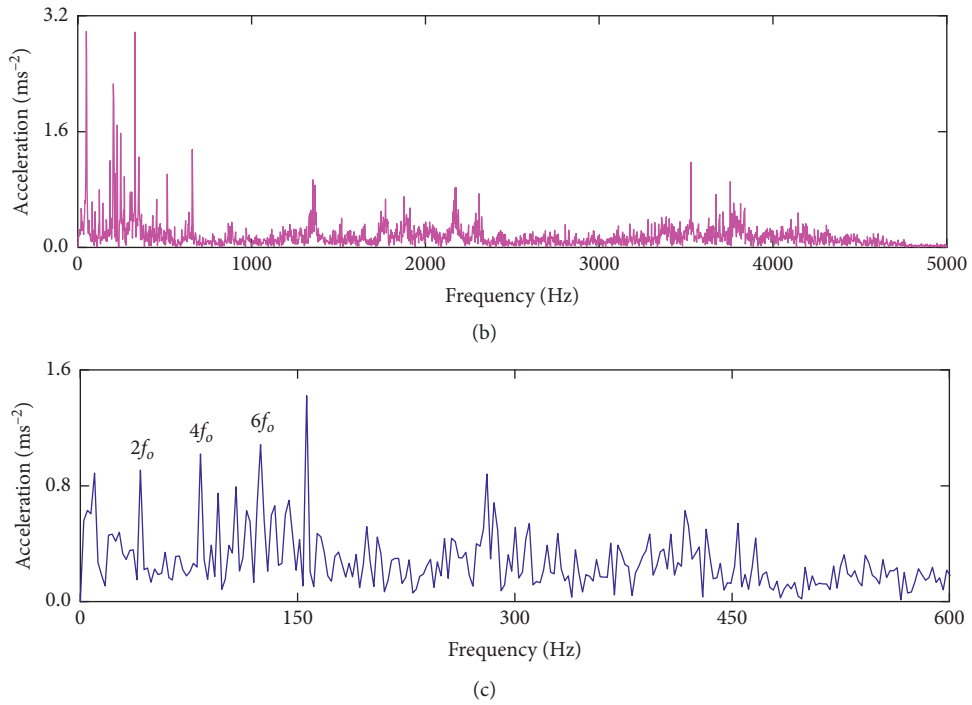Figure 33: Continued.

(b)

(c)

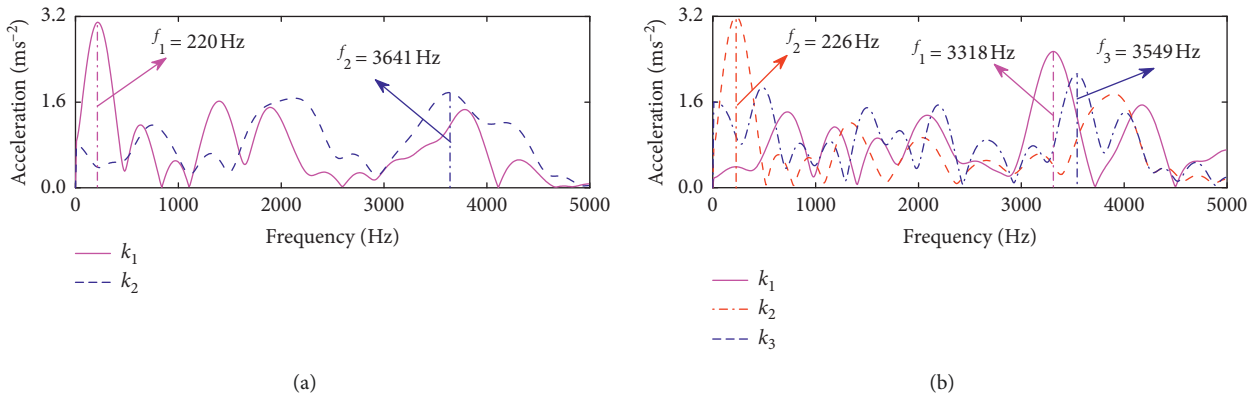FIGURE 33: Measured vibration signals: (a) original signal; (b) Fourier spectra of (a); (c) envelope spectra of (a).

(a)

(b)

FIGURE 34: Main frequencies for estimating the types of impulse responses.
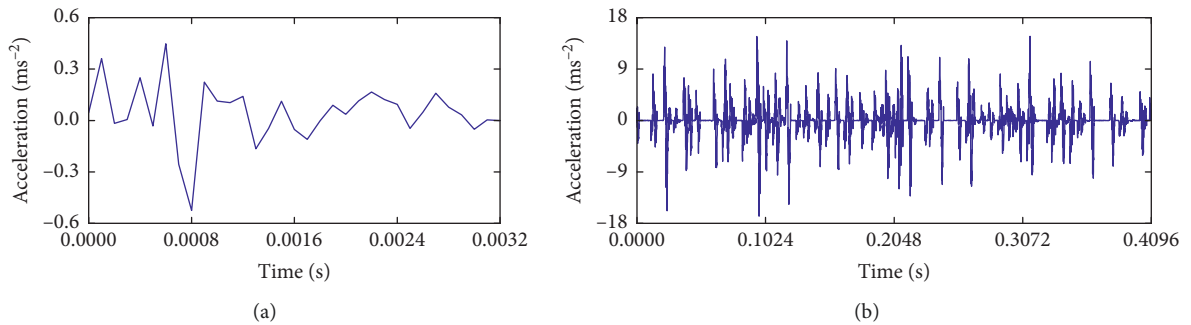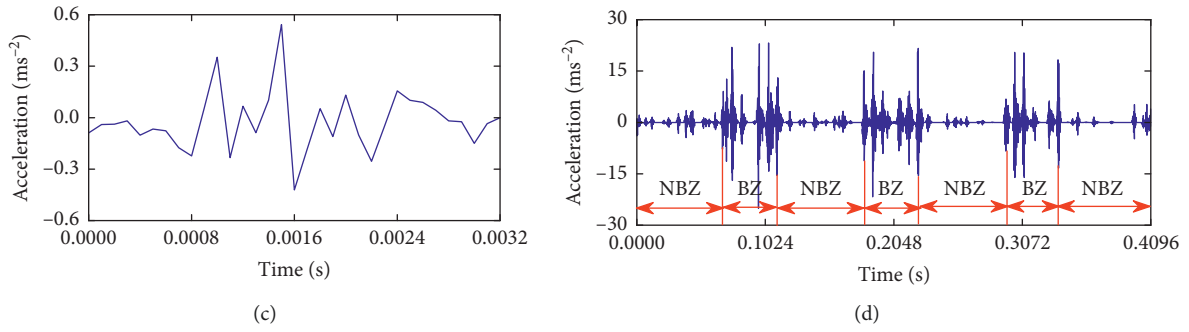

(a)

(b)

FIGURE 35: Continued.

(c)

(d)

FIGURE 35: Learned impulse responses: (a) IR1; (b) IRS1 of IR1; (c) IR2; (d) IRS2 of IR2.
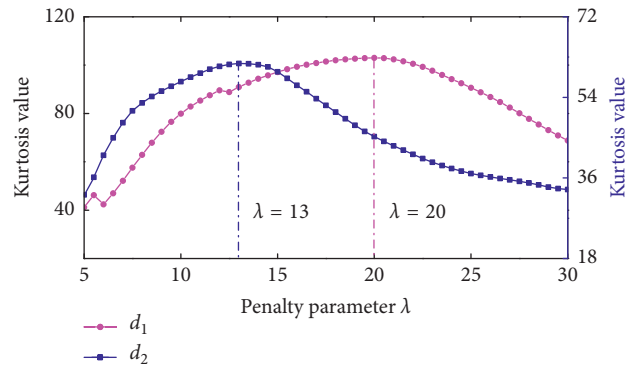
FIGURE 36: Envelope spectra kurtoses for different penalty values.

(a)

(b)

(c)

(d)
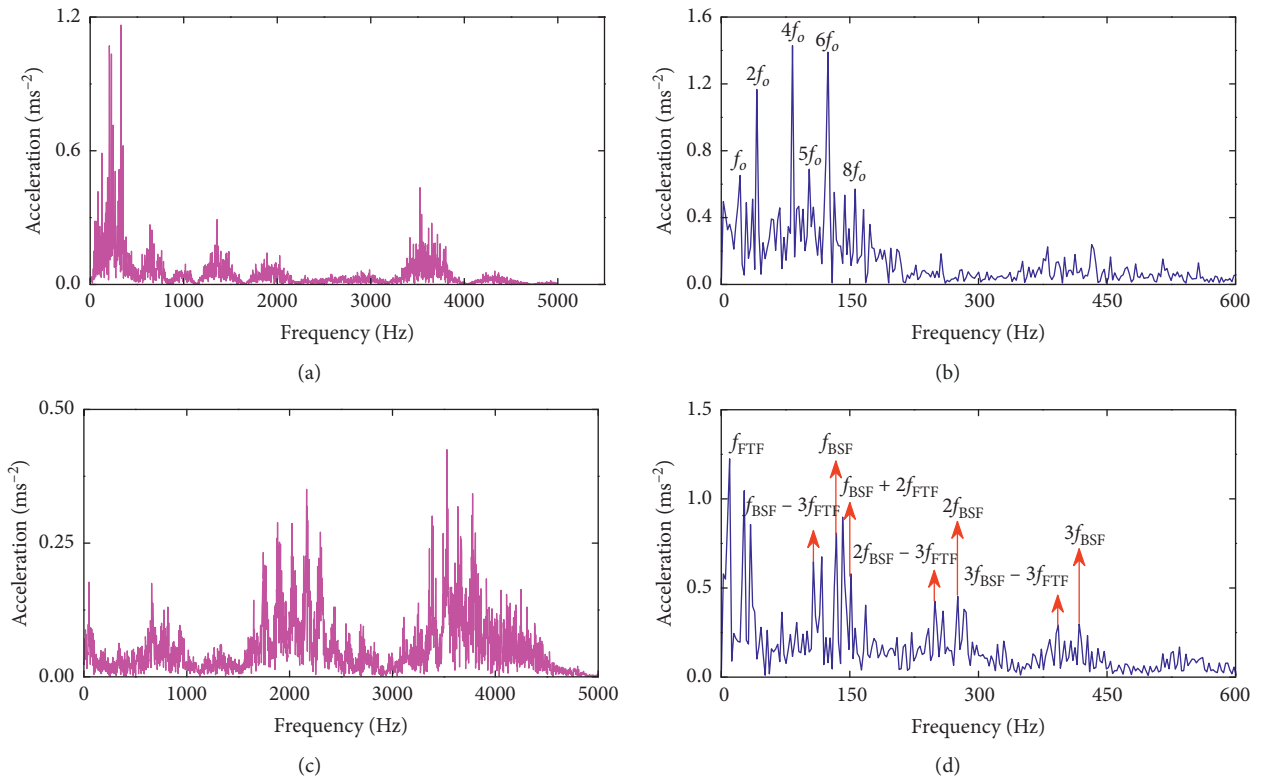
FIGURE 37: Fourier and envelope spectra: (a) Fourier spectra of IRS1; (b) envelope spectra of IRS1; (c) Fourier spectra of IRS2; (d) envelope spectra of IRS2.

FIGURE 38: The defect on the tested wheel tread.

FIGURE 39: Results obtained using (a) spectra kurtosis: (b) purified signal; (c) Hilbert envelope spectra.

FIGURE 40: Results obtained using EEMD: (a) IMF1; (b) envelope spectra of IMF1; (c) IMF3; (d) envelope spectra of IMF3.

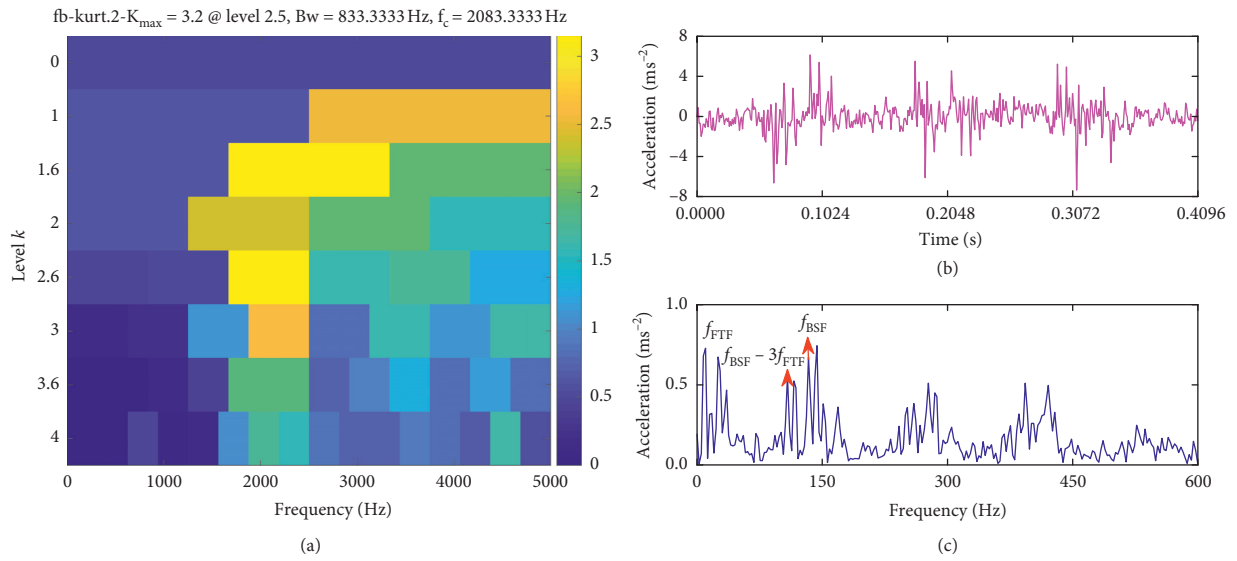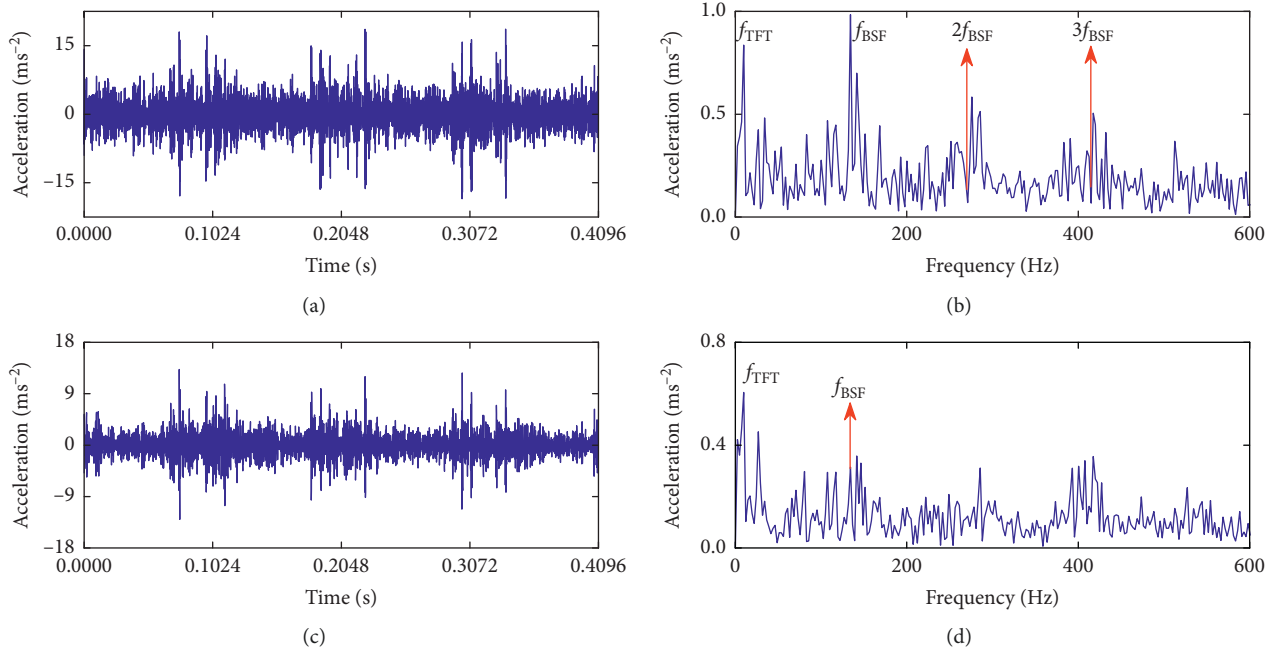fault, as shown in Figure 37(d). The rotational frequency f_o of the wheelset and its harmonics can be observed in Figure 37(b). In accordance with the vibrational feature, there should have been a defect on the wheel tread. When the tested wheelset was carefully checked, the wheel tread was found to have the defect shown in Figure 38.

To compare the performances of the proposed ACSR with the well-known spectra kurtosis and EEMD methods, two comparative methods were used to analyse the measured vibration signals in Figure 33(a). The analysed results are shown in Figures 39 and 40. As seen in Figure 39(a), the extracted IRSs were distorted. From Figure 39(b), it can be seen that the spectra kurtosis only found the basic frequency of the roller-fault-characteristic frequency. In Figures 40(a) and 40(c), there was strong noise in the two impulse responses. As a result, the fault action of the roller could not be characterized. As shown in Figures 40(b) and 40(d), the basic frequency of the roller-fault-characteristic frequency could only be detected from the envelope spectra of the IMFs, indicating that the proposed ACSR performed better when detecting wheelset-bearing faults under the running conditions of the high-speed train.

7. Conclusion

In this paper, a novel fault detection method named ACSR was proposed:

(1) A convolutional representation mode for characterizing impulse response series was developed. An IRS induced by bearing faults can be described as the convolution of an impulse response and the resulting time-location coefficients.

(2) ACSR was proposed based on the combination of CSR and a method for estimating three parameters (the number of types of impulse responses C, length of impulse response P, and regularization parameter related to the convolution sparse representation λ).

(3) The capacity of ACSR for characterizing and detecting bearing faults was validated by simulation and bench and real-line tests. Compared to spectra kurtosis and EEMD, ACSR can not only extract full IRSs but also isolate IRSs induced by multiple faults in a wheelset-bearing system. Therefore, ACSR's performance for detecting wheelset-bearing faults is superior to that for two other comparative methods.

In addition, ACSR provides a route for estimating the states of wheel treads using the measured vibration signals of axle boxes. This will be further investigated in the future. Using the Matlab.2014a platform on a Dell Inspiron 14 laptop computer, the mean calculation time of SIDL and SISC for the tested signals in Figure 33(a) was approximately 52 seconds. The real-time performance of ACSR needs to be improved.

Data Availability

The data used to support the findings of the study have not been made available because they are confidential.

Conflicts of Interest

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

Acknowledgments

References

[1] J. Ding and C. Ding, "Automatic detection of a wheelset bearing fault using a multi-level empirical wavelet transform," *Measurement*, vol. 134, pp. 179–192, 2019.

[2] J. Ding, F. Li, J. Lin, B. Miao, and L. Liu, "Fault detection of a wheelset bearing based on appropriately sparse impulse extraction," *Shock and Vibration*, vol. 2017, no. 71, Article ID 7853918, 17 pages, 2017.

[3] M. Liang and I. Soltani Bozchalooi, "An energy operator approach to joint application of amplitude and frequency-demodulations for bearing fault detection," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1473–1494, 2010.

[4] G. He, K. Ding, and H. Lin, "Fault feature extraction of rolling element bearings using sparse representation," *Journal of Sound and Vibration*, vol. 366, pp. 514–527, 2016.

[5] 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.

[6] Y.-T. Sheen, "An envelope analysis based on the resonance modes of the mechanical system for the bearing defect diagnosis," *Measurement*, vol. 43, no. 7, pp. 912–934, 2010.

[7] S. R. Qin and Y. M. Zhong, "Research on the unified mathematical model for FT, STFT and WT and its applications," *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1335–1347, 2004.

[8] Y.-K. Park and Y.-H. Kim, "A method to minimise the cross-talk of the wigner-ville distribution," *Mechanical Systems and Signal Processing*, vol. 11, no. 4, pp. 547–559, 1997.

[9] D. Yu, J. Cheng, and Y. Yang, "Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings," *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 259–270, 2005.

[10] W. Su, F. Wang, H. Zhu, Z. Zhang, and Z. Guo, "Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement," *Mechanical Systems and Signal Processing*, vol. 24, no. 5, pp. 1458–1472, 2010.

[11] X. Wang, Y. Zi, and Z. He, "Multiwavelet denoising with improved neighboring coefficients for application on rolling bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 285–304, 2011.

[12] L. Cui, J. Wang, and S. Lee, "Matching pursuit of an adaptive impulse dictionary for bearing fault diagnosis," *Journal of Sound and Vibration*, vol. 333, no. 10, pp. 2840–2862, 2014.

[13] 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.

[14] X. Chiementin, F. Bolaers, and J.-P. Dron, "Early detection of fatigue damage on rolling element bearings using adapted wavelet," *Journal of Vibration and Acoustics*, vol. 129, no. 4, pp. 495–506, 2007.

[15] X. Fan and M. J. Zuo, "Gearbox fault detection using Hilbert and wavelet packet transform," *Mechanical Systems and Signal Processing*, vol. 20, no. 4, pp. 966–982, 2006.

[16] 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.

[17] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[18] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[19] S. A. Bagherzadeh and M. Sabzehparvar, "A local and online sifting process for the empirical mode decomposition and its application in aircraft damage detection," *Mechanical Systems and Signal Processing*, vol. 54-55, pp. 68–83, 2015.

[20] H. Jiang, C. Li, and H. Li, "An improved EEMD with multiwavelet packet for rotating machinery multi-fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 225–239, 2013.

[21] Z. K. Peng, P. W. Tse, and F. L. Chu, "An improved Hilbert-Huang transform and its application in vibration signal analysis," *Journal of Sound and Vibration*, vol. 286, no. 1-2, pp. 187–205, 2005.

[22] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[23] G. Rilling, P. Flandrin, P. Gonalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 12, pp. 936–939, 2007.

[24] T. Tanaka and D. P. Mandic, "Complex empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 101–104, 2007.

[25] J. Fleureau, A. Kachenoura, L. Albera, J.-C. Nunes, and L. Senhadji, "Multivariate empirical mode decomposition and application to multichannel filtering," *Signal Processing*, vol. 91, no. 12, pp. 2783–2792, 2011.

[26] J. Lin and L. Qu, "Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis," *Journal of Sound and Vibration*, vol. 234, no. 1, pp. 135–148, 2000.

[27] J. Lin and M. J. Zuo, "Gearbox fault diagnosis using adaptive wavelet filter," *Mechanical Systems and Signal Processing*, vol. 17, no. 6, pp. 1259–1269, 2003.

[28] Y. Lei, J. Lin, Z. He, and Y. Zi, "Application of an improved Kurtogram method for fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1738–1749, 2011.

[29] Y. Lei, Z. He, and Y. Zi, "Application of the EEMD method to rotor fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 23, no. 4, pp. 1327–1338, 2009.

[30] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 123–151, 2005.

[31] B. Chen, Z. Zhang, C. Sun, B. Li, Y. Zi, and Z. He, "Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors," *Mechanical Systems and Signal Processing*, vol. 33, pp. 275–298, 2012.

[32] I. W. Selesnick, "Wavelet transform with tunable Q-factor," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3560–3575, 2011.

[33] Y. Qin, B. Tang, and J. Wang, "Higher-density dyadic wavelet transform and its application," *Mechanical Systems and Signal Processing*, vol. 24, no. 3, pp. 823–834, 2010.

[34] W. He, Y. Zi, B. Chen, F. Wu, and Z. He, "Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform," *Mechanical Systems and Signal Processing*, vol. 54-55, pp. 457–480, 2015.

[35] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.

[36] R. Jiang, H. Qiao, and B. Zhang, "Efficient Fisher discrimination dictionary learning," *Signal Processing*, vol. 128, pp. 28–39, 2016.

[37] A. M. Tillmann, "On the computational intractability of exact and approximate dictionary learning," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 45–49, 2015.

[38] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[39] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Transactions on Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.

[40] K. Skretting and K. Engan, "Recursive least squares dictionary learning algorithm," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2121–2130, 2010.

[41] H. Liu, C. Liu, and Y. Huang, "Adaptive feature extraction using sparse coding for machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 558–574, 2011.

[42] 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.

[43] L. Cui, N. Wu, C. Ma, and H. Wang, "Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a new step-impulse dictionary," *Mechanical Systems and Signal Processing*, vol. 68-69, pp. 34–43, 2016.

[44] Y. Qin, Y. Mao, and B. Tang, "Vibration signal component separation by iteratively using basis pursuit and its application in mechanical fault detection," *Journal of Sound and Vibration*, vol. 332, no. 20, pp. 5217–5235, 2013.

[45] W. Fan, G. Cai, Z. K. Zhu, C. Shen, W. Huang, and L. Shang, "Sparse representation of transients in wavelet basis and its application in gearbox fault feature extraction," *Mechanical Systems and Signal Processing*, vol. 56-57, pp. 230–245, 2015.

[46] H. Zhang, X. Chen, Z. Du, and R. Yan, "Kurtosis based weighted sparse model with convex optimization technique for bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 80, pp. 349–376, 2016.

[47] H. Zhang, X. Chen, Z. Du, X. Li, and R. Yan, "Nonlocal sparse model with adaptive structural clustering for feature extraction of aero-engine bearings," *Journal of Sound and Vibration*, vol. 368, pp. 223–248, 2016.

[48] H. Zhou, J. Chen, G. Dong, and R. Wang, "Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 65–79, 2016.

[49] Z. Feng and M. Liang, "Complex signal analysis for planetary gearbox fault diagnosis via shift invariant dictionary learning," *Measurement*, vol. 90, pp. 382–395, 2016.

[50] W. He, Y. Ding, Y. Zi, and I. W. Selesnick, "Sparsity-based algorithm for detecting faults in rotating machines," *Mechanical Systems and Signal Processing*, vol. 72-73, pp. 46–64, 2016.

[51] B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 301–315, 2016.

[52] A. Chambolle, R. A. De Vore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 319–335, 1998.

[53] R. Chalasani, J. C. Principe, and N. Ramakrishnan, "A fast proximal method for convolutional sparse coding," in *Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN)*, Dallas, TX, USA, August 2013.

[54] B. S. He, H. Yang, and S. L. Wang, "Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities," *Journal of Optimization Theory and Applications*, vol. 106, no. 2, pp. 337–356, 2000.

[55] 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.

[56] J. Yan and R. B. Randall, "The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines," *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 308–331, 2006.