Hindawi
Shock and Vibration
Volume 2020, Article ID 8823102, 19 pages
https://doi.org/10.1155/2020/8823102

Research Article

Sparsity-Oriented Nonconvex Nonseparable Regularization for Rolling Bearing Compound Fault under Noisy Environment

Xiaocheng Li,[1] Jingcheng Wang ,[1,2] and Hongyuan Wang[1]

[1]*Department of Automation, Shanghai Jiaotong University, Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai, China*
[2]*Autonomous Systems and Intelligent Control International Joint Research Center, Xi'an Technological University, Xi'an, China*

Correspondence should be addressed to Jingcheng Wang; jcwang@sjtu.edu.cn

Received 20 March 2020; Revised 9 May 2020; Accepted 4 June 2020; Published 30 June 2020

Academic Editor: Rafał Burdzik

Rolling bearing is widely used in rotating machinery and, at the same time, it is easy to be damaged due to harsh operating environments and conditions. As a result, rolling bearing is critical to the safe operation of the machinery devices. Compound fault of rolling bearing is not a simple superimposition of multiple single faults, but the coupling of multiple fault features, making the vibration signal, becomes complicated. In our study, sparsity-oriented nonconvex nonseparable regularization (SONNR) method is proposed to rolling bearing compound fault diagnosis under noisy environment. Firstly, a theoretical model of rolling bearing compound fault is established, and the vibration characteristics of rolling bearing compound fault are analyzed. Secondly, four-layer structure of the SONNR method is proposed: input layer, nonconvex sparse regularization layer, signal reconstruction layer, and compound faults isolation layer. Finally, the validity of the method is verified by simulation data and actual data, and it is compared with the traditional time domain diagnostic methods and artificial intelligence methods.

1. Introduction

Rolling bearing is widely used in many mechanical devices. With increase in complexity of machinery, the requirements of reliability for rolling bearing are increasing. Bearing damage, even minor damage, will affect the normal function of the devices [1–3]. A rolling bearing is often composed of an inner ring, an outer ring, a number of rolling elements, and a cage. The inner ring acts to match the shaft and rotates around the shaft. The outer ring acts to cooperate with the bearing seat to provide support. The cage evenly distributes the rolling elements, prevents the rolling elements from falling off, and guides the rolling elements to rotate. Fortunately, the vibration signal generated by a single fault of the rolling bearing has been extensively studied and a series of powerful diagnostic methods have been made. However, in the course of the rolling bearing, there is not only single fault; with the change of the operating environment and running time increasing, bearing often has two or more faults that are named compound fault of the rolling bearing. Compound fault is not a simple superimposition of multiple single faults, but the coupling of multiple fault features, making the vibration signal complicated. In addition, rolling bearings are often disturbed by various background noises during operation. When there are early weak faults in the bearing, the fault characteristics are often covered by strong noise, making it difficult to obtain the fault features. The background noise is often assumed to be additive white Gaussian noise (AWGN) [4]. In the classical techniques, the assumption of Gaussian noise is taken under consideration [5–7]. The research signal in this paper already contains noise signal. In order to verify the effectiveness of the method, we artificially added Gaussian white noise. Therefore, the study of compound fault signals detection under noisy environment technologies becomes urgent to overcome.

To monitor and detect bearing faults and damage, vibration signals are often used. In the past decades, various signal processing techniques were developed for extracting fault information about vibration signals. At present, the methods of compound fault diagnosis of rolling bearing are mainly empirical mode decomposition (EMD) method, wavelet analysis method, and sparsity-assisted filtering methods [8–11]. With the rapid development of artificial

intelligence and computer technology, intelligent diagnostic methods based on deep learning (DL) [16–19] were developed, including convolutional neural network (CNN), stacked auto encoder (SAE), deep belief network (DBN), deep Boltzmann machine (DBM), and recurrent neural network (RNN). However, the existing methods have the following limitations. The intrinsic mode function (IMF) component extracted by the EMD method is essentially a band-amplified signal that has a sharp cut in the impact. In addition, wavelet analysis and EMD are both signal processing methods based on band division. When the center frequency of the impact signal overlaps between the center frequencies of other signals, the impact component cannot be effectively extracted. L1 norm regularization methods have become widely used in various applications due to the convexity of the L1 norm. However, L1 norm solutions often underestimate the high-amplitude components, yet these comprise the signal of interest in most cases. Thus, L1 norm solutions are not ideal for certain applications, including bearing fault diagnosis. A vast number of deep models had been proposed and verified with various types of bearing fault signals. For artificial intelligence, all the above publications stopped at the problem of a single fault, while the more challenging case, the multiple fault bearing diagnosis, has not been verified yet.

In general, it is challenging to extract features of vibration signals to indicate bearing faults, especially for compound faults. The coupling characteristics of rolling bearing compound fault signals are mainly manifested in three aspects: (1) coupling characteristic caused by the change of structural parameters, that is, when the rolling bearing fails, the geometric parameters of the bearing structure change due to the fault, which causes a certain difference between different frequency components. The ratio relationship, the ratio, and the phase corresponding to the frequency is consistent. (2) Coupling occurring between different frequency components. Since different rolling bearing fault signals have corresponding characteristic frequencies and their harmonic components, and the phases corresponding to different signal components have a certain correlation, some signal frequencies may have a certain correlation with other harmonic frequencies. (3) Convex spare solutions often underestimate the high-amplitude components, yet these comprise the signal of interest in most cases. Thus, convex spare solutions are not ideal for certain applications, including bearing fault diagnosis.

In order to overcome these issues, sparsity-oriented nonconvex nonseparable regularization (SONNR) for rolling bearing compound fault under noisy environment is studied in this paper. The SONNR method uses a nonconvex penalty and simultaneously maintains the convexity of the least squares cost function to be minimized. Nonconvex sparse regularization (NSR) that maintains the convexity of the cost function has been recently studied, to capture the advantages of both nonconvex regularization and convex optimization [20, 21]. The SONNR method has a four-layer structure: input layer, nonconvex sparse regularization layer, signal reconstruction layer, and compound faults isolation layer. The measured time-domain signal is input layer. In NSR decomposition layer, nonconvex sparse regularization by K-singular value decomposition (K-SVD) is used to improve denoising performance of raw signal. In signal reconstruction layer, more sparse and noise-reduced signals are reconstructed. In compound faults isolation layer, the measured time-domain signal is decomposed into high and low components using overcomplete rational-dilation wavelet transforms (RADWT).

The main contributions of this paper are summarized as follows:

(1) The physical model of the compound fault is established, and the fault impact signal, the envelope signal, and the envelope signal spectrum of the compound fault signal are derived. The modulation characteristics of the compound fault signal are obtained from the analysis of the envelope spectrum, which lays a theoretical foundation for the subsequent fault information extraction.

(2) The nonconvex optimization method for compound faults under noisy environment has seldom been considered in previous studies. The problem of the characteristic coupling of the rolling bearing compound faults signals and the interference caused by the noisy environment makes this work a great challenge. Exploiting the high-amplitude components property, a K-SVD-based nonconvex nonseparable regularization denoising algorithm is presented to extract bearing fault features.

(3) The RADWT decomposes a complex signal into the sum of an oscillatory component and a transient component. For compound faults, RADWT can well reveal the amplitude modulation frequency of the fault shock response.

(4) We present a verification based on real damage (accelerated lifetime) data set to validate the effectiveness of the proposed method in the diagnosis of compound faults in rolling bearing. In particular, we have artificially added noise based on the acquired signals.

The main advantage compared to conventional filtering methods is that the proposed method can better preserve the bearing fault signal while reducing noise and other interference components; thus it can significantly improve the estimation accuracy of the bearing fault signal. In addition, the compound fault signal can be separated to extract more obvious fault features.

The rest of this paper is organized as follows: Section 2 introduces the characteristics of the rolling bearing fault signal. Section 3 details the proposed SONNR method. Section 4 presents a simulation study and a real damage study, which is followed by concluding remarks in Section 5.

2. Analysis of Modulation Characteristics of Rolling Bearing Compound Fault Signal

We detect the signals from the fault signals collected by the sensors. The detection is based on the modulation

characteristics of the vibration signals. Therefore, before detecting the fault signals, we need to understand the signal modulation characteristics in the theoretical analysis.

2.1. Compound Fault Theory Model. The response of the impact sequence caused by the spalling fault at the sensor is a series of unilateral exponentially decaying oscillations. "Resonant demodulation" of this series of exponentially decaying sequences will obtain its envelope spectrum, which is the characteristic spectrum of rolling bearing fault. When a compound bearing has a compound fault, there is more than one location of the spalling fault. The position of the spalling point relative to the acceleration sensor changes continuously with the rotation. The spalling point will enter the load area and exit the load area. The magnitude and direction of the received series of vibration impact forces also change. In this case, the vibration signal of the rolling bearing compound fault collected by the acceleration sensor is more complicated, and the corresponding envelope signal and its spectrum are also more complicated. In order to make a more realistic interpretation of the envelope signal and the envelope spectrum and improve the accuracy of the diagnosis, a model for analyzing the envelope signal and its frequency spectrum of the compound bearing compound fault is proposed in the following.

In this chapter, the theoretical model is studied by taking the compound fault of the inner ring spalling and the rolling element spalling as an example. Figure 1 shows the theoretical model of compound fault of inner ring spalling and rolling element spalling.

When $t = 0$, the position of the spalling point of the inner ring is exactly at the peak of the load distribution density curve, that is, the lowest position in the axis direction, and it just collides with the rolling element. Since the fit between the inner ring and the shaft is an interference fit during assembly, the inner ring will rotate as the shaft rotates. The spalling point on the inner ring also rotates with the rotation of the bearing, so that the position of the spalling point does not always fall within the load distribution area as shown in Figure 1; most of the time it will fall outside the load distribution area.

Due to the different load, the resulting pulse force is different. When the spalling point of the inner ring falls within the load distribution area as shown in the figure, the magnitude and direction of the vibration pulse force are related to the position of the spalling point and the angle of the axis. When the spalling point is outside the load area shown in Figure 1, it can be known from the force analysis that if the rotational inertia force generated when the rolling bearing rotates is ignored at this time, no vibration pulse force is generated in this area. In summary, when the internal ring has a spalling fault, the vector value of the generated vibration pulse force is not only affected by the load distribution but also related to the location of the spalling fault point.

From this analysis, it can be seen that the magnitude and direction of the impulse force caused by the damage point of the inner ring and rolling element are affected by the load distribution and the position of the damage point. These factors are discussed in the following sections.

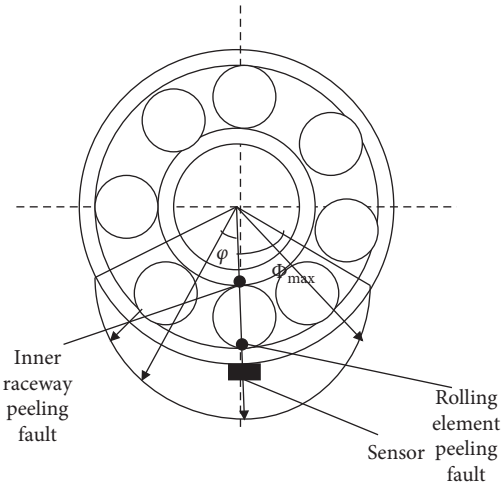

Figure 1: Compound fault theory model of internal ring spalling and rolling element spalling.

2.1.1. Regardless of Load Distribution and Location of Spalling Point. When the spalling point of the inner ring appears, its vibration characteristics appear as pulse vibrations in equal time intervals. If the impact of the load distribution and the position of the spallation point on the pulse force is ignored, it can be determined that each collision between the defect and the rolling element produces one vibration pulse, and the magnitude of each pulse force is equal. The expression of a series of pulses caused by spalling of the inner ring is shown in

$$\Delta_i(t) = \sum_{k=-\infty}^{+\infty} d_i \delta(t - kT_i), \tag{1}$$

where d_i is intensity of pulse force caused by inner ring damage, $\delta(t)$ is unit pulse function, $T_i = 1/f_i$ is interval between pulses, and f_i is inner ring fault characteristic frequency.

2.1.2. Load Distribution Expression. Rolling bearing is mainly subject to radial load forces during operation. When a rolling bearing is subjected to radial loads, the load distribution is shown in

$$q(\varphi) = q_{\max}\left[1 - \frac{1}{2\varepsilon}(1 - \cos\varphi)\right]^n, \tag{2}$$

where $\varphi_{\max} = \cos^{-1}(C_d/(2\delta_{\max} + C_d))$, $\varepsilon = (1/2)(1 - (C_d/(2\delta_{\max} + C_d)))$, C_d is radial clearance, and $\delta_{\max}$ is maximum contact deformation $n = 1.1$, $|\varphi| \le \varphi_{\max}$.

2.1.3. Expression of Influence Coefficient of Pulse Force Acting Direction. Assuming that the magnitude of the pulse force is proportional to the load distribution density at the position

where it is applied, the magnitude of the pulse force acting on the sensor axis is

$$f_i(t) = \Delta_i(t) q(2\pi f_s t) p(2\pi f_s t). \tag{3}$$

Since the phases of $q(\varphi)$ and $p(\varphi)$ are the same, let

$$c_i(t) = q(2\pi f_s t) p(2\pi f_s t). \tag{4}$$

The vibration envelope signal due to spalling damage of the inner ring is ($*$ indicates convolution)

$$v_i(t) = A_i[\Delta_i(t) c_i(t)] * e(t). \tag{5}$$

The envelope spectrum corresponding to $v_i(t)$ is

$$V_i(f) = A_i[\Delta_i(f) * C_i(f)] E(f). \tag{6}$$

For the rolling element, the amplitude spectrum is shown in

$$|V_b(f)| = \left[V_{b0}(f)^2 + V_{bi}(f)^2 + 2|V_{b0}(f)||V_{bi}(f)|\cos\Delta\varphi_b\right]^{(1/2)}. \tag{7}$$

Then, the envelope spectrum of the compound fault of the inner ring and the rolling element can be processed by the superimposition principle, and the vector sum of the total amplitude is shown as follows:

$$|V(f)| = \left[V_i(f)^2 + V_0(f)^2 + 2|V_i(f)||V_0(f)|\cos\Delta\phi\right]^{(1/2)}, \tag{8}$$

$$\Delta\phi = -m\Delta\theta - n\Delta\varphi, \quad (m, n \text{ are integers}). \tag{9}$$

According to (8) and (9), the envelope amplitude spectrum of the combined fault of the inner ring spalling and the rolling element spalling can be obtained. The envelope spectrum diagram is shown in Figure 2. The frequency component is the same as in the case of a single fault: the characteristic spectrum of the inner ring spalling is a series of main spectral lines with the inner circle fault characteristic frequency f_i as the distance, and at the same time, the two sides of each main spectral line are separated by the frequency f_r and its multiples. The amplitudes of the main and sidelines decrease with increasing frequency. It should be noted that there is a peak at the frequency f_r. The frequency spectrum is a series of main spectral lines with the characteristic frequency f_b of the rolling element as the interval. At the same time, each of the main spectral lines is separated from the revolution frequency f_f and its side frequency with smaller amplitude side spectral lines. In addition, there is a peak at the revolution frequency f_f.

By analyzing the modulation characteristics of the compound fault, we can know that the fault characteristics of the inner ring and the rolling element need to be execrated in order to prevent missed diagnosis and misdiagnosis.

2.2. Coupling Characteristics and Challenges of Rolling Bearing Compound Fault Signals. The coupling characteristics of rolling bearing compound fault signals are mainly manifested in four aspects: (1) Due to the difference in the location and size of the fault, the response waveforms generated by different faults will also affect each other, coupling multiple fault characteristics and interfering with each other. (2) The coupling caused by changes in structural parameters, that is, when the rolling bearing fails, the fault causes the geometric parameters of the bearing structure to change, which causes a certain degree of difference between different frequency components. The relationship between this ratio and the phase corresponding to the frequency is the same. (3) The coupling between different frequency components, that is, different rolling bearing fault signals generally appear to have corresponding characteristic frequencies and their harmonic components. Because the phases corresponding to different signal components have a certain correlation, some signal frequencies may have a certain correlation with other harmonic frequencies. (4) For rolling bearing compound fault under noisy environment, the signal modulation characteristics of the time domain signal are not obvious. It will be strengthened, and the impact component with low energy may be overwhelmed by the fault characteristics with high energy. These four aspects will lead to the incomplete diagnosis of fault characteristics when the single fault diagnosis method is used to diagnose the rolling bearing compound fault.

This section illustrates these challenges using actual signals as examples.

Figure 3 is the acceleration vibration signal and the frequency spectrum of the bearing when the inner and outer raceways and rolling elements have spalling fault at the same time. The shaft speed is 30.01 Hz, the load is 5 Kg, the sampling frequency is 51.2 KHz, and the number of sampling points is 65536. The bearing's outer ring characteristic frequency is 91.47 Hz, the inner ring characteristic frequency is 148.55 Hz, the rolling element characteristic frequency is 59.78 Hz, and the cage rotation frequency is 11.34 Hz. As can be seen from Figure 4, the vibration of time-domain signal of the rolling bearing compound fault is relatively chaotic. Although several impacts can be observed, it is difficult to determine the fault location. Envelope analysis is performed on this signal, and the envelope spectrum obtained is shown in Figure 5.

As can be seen in Figure 4(b), there is a spectral peak of 91.41 Hz and its multiples, which is consistent with the characteristic frequency of the outer ring fault. The analysis can only detect the inner ring fault, while other simultaneous outer ring and rolling element faults are omitted. This determines that the traditional single-point fault method can only be used to diagnose the compound bearing of a rolling bearing, and it is difficult to analyze and diagnose it comprehensively and accurately.

The autocorrelation processing is performed on the envelope signal of Figure 4(b) to obtain the self-power spectrum shown in Figure 5. Figure 5 eliminates noise glitches and makes the fault frequency in the spectrum more prominent. We can see that there is a spectral peak of 91.41 Hz and its frequency doubling, which can more significantly determine that the bearing has an outer ring spalling fault. The energy caused by the impact of the fault is distributed in the entire frequency domain. The overall analysis can only display the characteristics of

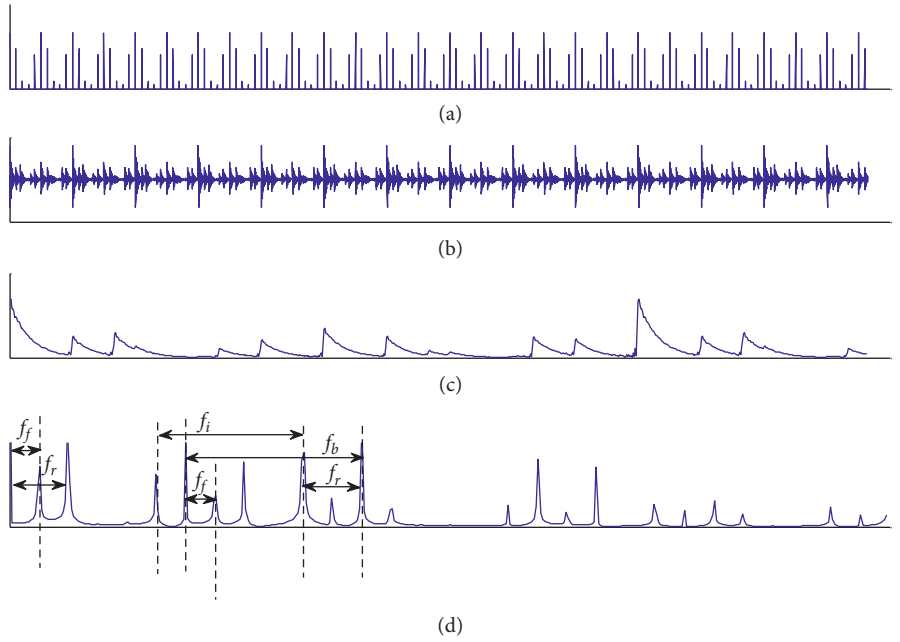

Figure 2: Modulation characteristics of vibration signal of compound fault. (a) Fault shock. (b) Fault shock response. (c) Envelope signal. (d) Envelope spectrum.

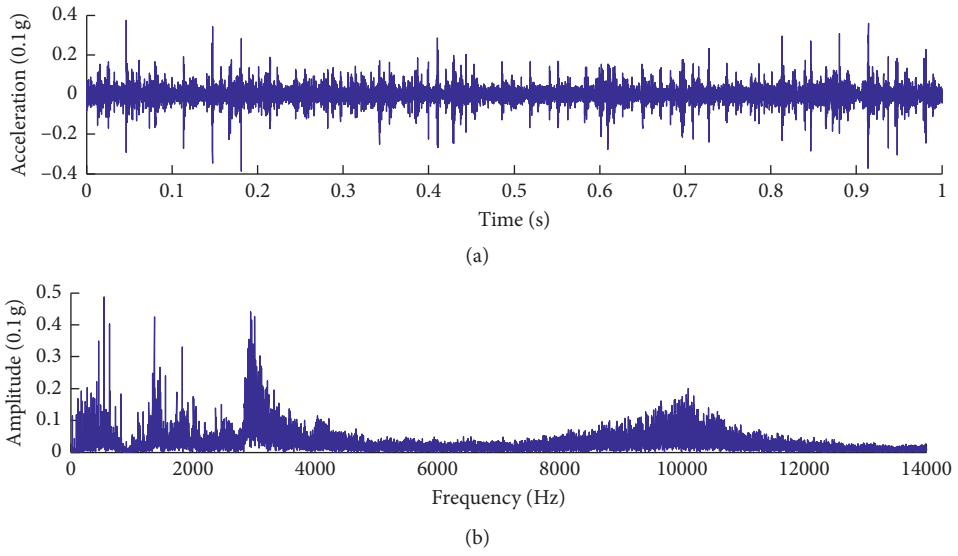

Figure 3: Vibration acceleration signal of compound fault bearing. (a) Time-domain diagram of acceleration signal. (b) Frequency-domain diagram of acceleration signal.

the compound fault in the same result. The fault features with lower energy may be overwhelmed by the fault features with higher energy. Therefore, compound faults should be expressed in different channels as much as possible.

3. The Proposed Algorithm SONNR

Figure 6 gives a flowchart of the proposed methodology. The SONNR method has a four-layer structure: input layer, nonconvex sparse regularization layer, signal reconstruction layer, and compound faults isolation layer. The measured time-domain signal is input layer. In NSR decomposition layer, nonconvex sparse regularization by K-SVD is used to improve the denoising performance of raw signal. In signal reconstruction layer, more sparse and noise-reduced signals are reconstructed. In compound faults isolation layer, the measured time-domain signal is decomposed into high and low components using RADWT. The high-Q component contains a sustained oscillation cycle signal; the low-Q component contains the rolling fault impact transient signal.

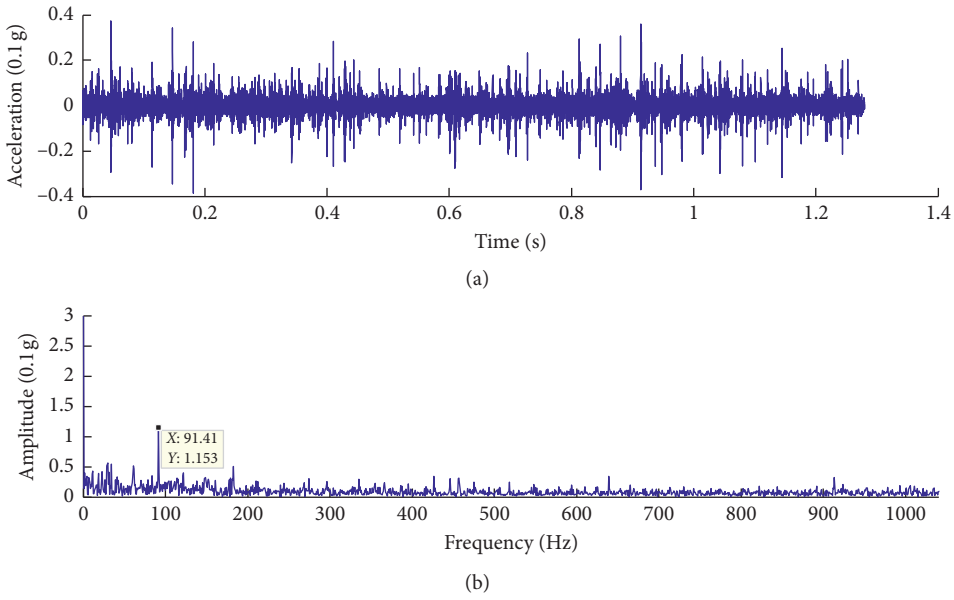

Figure 4: Fault rolling bearing vibration acceleration. (a) Time-domain acceleration signal. (b) Spectrum of acceleration signal.

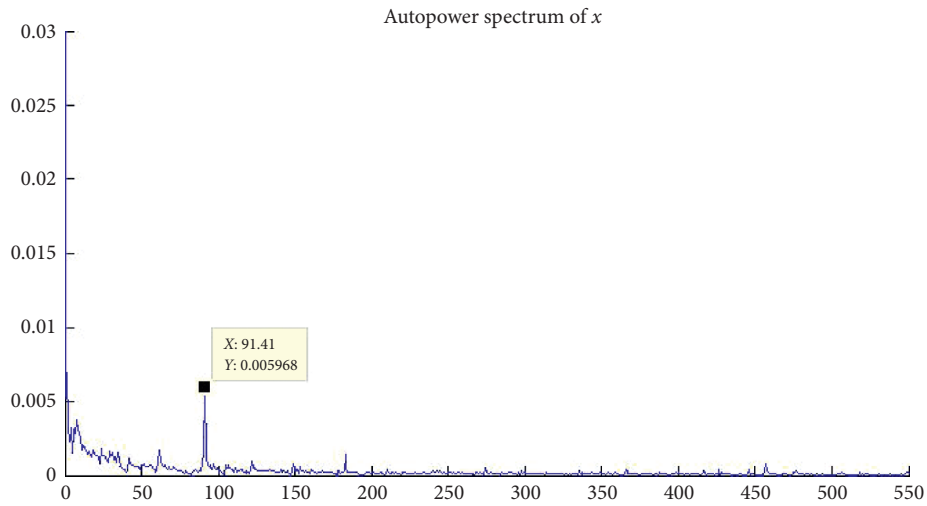

Figure 5: Self-power spectrum of the envelope signal.

3.1. Nonconvex Sparse Regularization by K-SVD. In order to enhance sparsity, many NSR methods have been proposed, and a general way to model the problem is the nonconvex regularization least squares problem:

$$\min_x \left\{ F(x) = \frac{1}{2}\|y - Ax\|_2^2 + \lambda\psi(x) \right\}, \tag{10}$$

where ψ: $\mathbb{R}^N \longrightarrow \mathbb{R}$ is a nonconvex sparsity-inducing penalty. In this section, we use the nonconvex generalized minimax-concave (GMC) penalty in [22].

In our SONNR method, we use the GMC penalty ψ_{GMC}, which is defined as

$$\psi_{\text{GMC}} = \|x\|_1 - \min_v \left\{ \|v\|_1 + \frac{\gamma}{2\lambda}\|A(x - v)\|_2^2 \right\}. \tag{11}$$

Using the GMC penalty in (11), the cost function in (10) can be expressed as

$$\begin{aligned} F(x) &= \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1 - \min_v \left\{ \lambda\|v\|_1 + \frac{\gamma}{2}\|A(x - v)\|_2^2 \right\} \\ &= \max_v \left\{ \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1 - \lambda\|v\|_1 - \frac{\gamma}{2}\|A(x - v)\|_2^2 \right. \\ &= \max_v \left\{ \frac{1}{2}(1 - \gamma)\|Ax\|_2 + \lambda\|x\|_1 + g(x, v) \right\} \\ &= \frac{1}{2}(1 - \gamma)\|Ax\|_2 + \lambda\|x\|_1 + \max_v \{g(x, v)\}. \end{aligned} \tag{12}$$

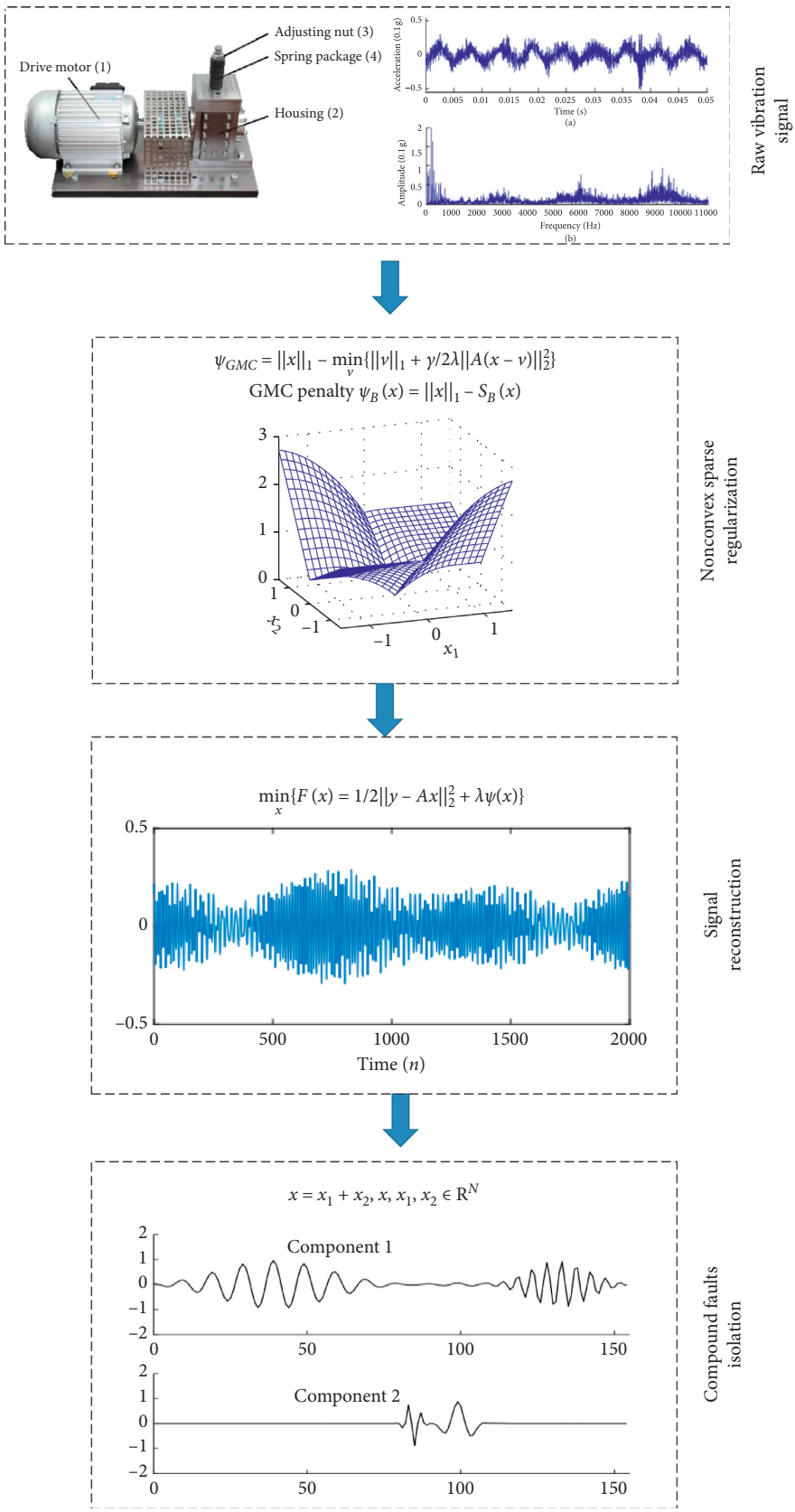

Figure 6: The flowchart of the proposed SONNR.

In our SONNR method, we use the GMC penalty ψ_{GMC}, which is defined as

$$\psi_{\mathrm{GMC}}(x) = \psi_B(x) = \|x\|_1 - S_B(x), \tag{13}$$

where $S_B(x) \coloneqq \inf_{v \in R^N} \left\{ \|v\|_1 + (1/2)\|B(x-v)\|_2^2 \right\}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$. $S_B(x)$ is generalized Huber function. The GMC penalty ψ_{GMC} is shown in Figure 7 and the generalized Huber function $S_B(x)$ is shown in Figure 8.

The minimization of the GMC-regularized least squares problem can be written as a saddle-point problem:

$$\left(x^{\mathrm{opt}}, v^{\mathrm{opt}}\right) = \arg\min_x \max_v F(x, v), \tag{14}$$

where

$$F(x, v) = \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1 - \lambda\|v\|_1 - \frac{\lambda}{2}\|A(x-v)\|_2^2, \tag{15}$$

with $0 < \lambda < 1$, $0 < \mu < 2/\max\{1, \lambda/(1-\lambda)\}\|A\|_2^{-2}$, which ensures that the iterative algorithm converges.

The linear transformation A in (11) is important for sparse representation. An effective transformation can promote the sparsity of coefficients. For bearing fault diagnosis, we use K-SVD as the linear transformation A in this paper. The detailed description of K-SVD can be found in [23]. In K-SVD, the linear transformation A can be solved by

$$A = \arg\min_A \sum_{i=1}^{N} \min_{x_i} \left\{ \left\|Ax_i - y_i\right\|^2 + \lambda \left\|x_i\right\|_1 \right\}. \tag{16}$$

K-SVD can be regarded as a generalized form of K-means. In the K-means algorithm, each semaphore can only be approximated by one atom, while each signal in K-SVD is a linear combination of multiple atoms. Solve the sparse representation coefficients based on the given signal y and the initial dictionary D. This problem can be solved by K-SVD, and the closest solution A can be found by the basis pursuit (BP) algorithm.

Iterative algorithm for nonconvex sparse regularization by K-SVD problem is shown in Algorithm 1.

3.2. Principle of Overcomplete Rational-Dilation Wavelet Transforms. The authors in [24] proposed an overcomplete rational expansion from the perspective of the frequency domain, wavelet transform (RADWT), overcomplete rational expansion wavelet transform by changing the upsampling of the signal p and downsampling the value of q to adjust the quality factor of the wavelet, greatly expanding the range of the quality factor and increasing. The flexibility of quality factor selection: the quality factor Q of this method is determined by the following equation:

$$Q = \sqrt{\frac{p}{q}}\,\frac{1}{(1-p)/q}. \tag{17}$$

The quality factor is a measure of the degree of frequency aggregation of a signal. The higher the quality factor, the more concentrated the frequency of the signal, and the higher the number of oscillations in the time domain.

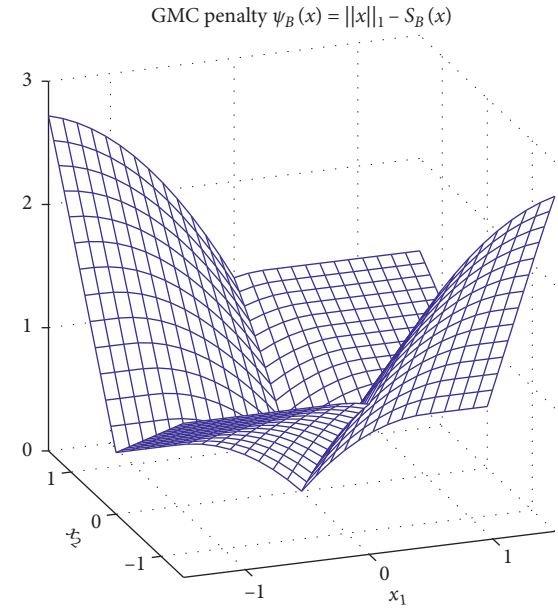

Figure 7: GMC penalty ψ_{GMC}.

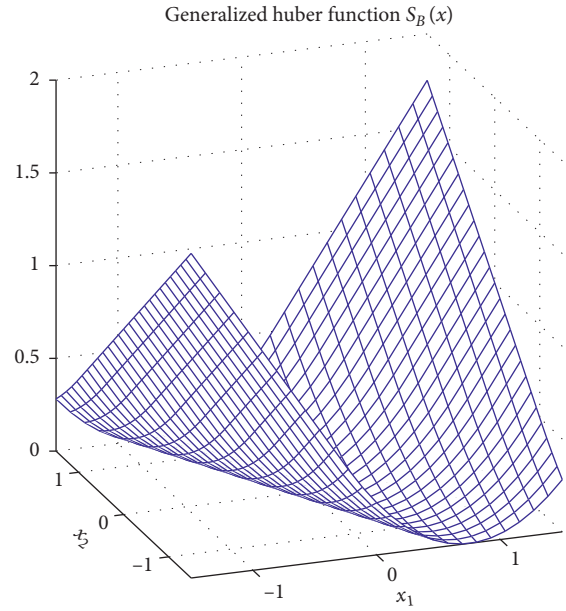

Figure 8: Generalized Huber function $S_B(x)$.

The RADWT decomposes a complex signal into the sum of an oscillatory component and a transient component [25]. The RADWT nonlinearly separates the components in the signal according to the oscillation characteristics and establishes the optimal sparse representation of each of the high resonance components and the low resonance components.

It is assumed that the observed signal x can be expressed as the sum of two signals x_1 and x_2:

(1) **initialize** $x^{(0)}, v^{(0)}, 0 < \lambda < 1, 0 < \mu < 2/\max\{1, \lambda/(1-\lambda)\}\|A\|_2^{-2}$
(2) $i = 0, 1, 2, \ldots$ **do**
(3) calculate the coefficients x_i using a simple gradient descent procedure
(4) update the dictionary $\mathbf{A}^{(n+1)} = \mathbf{A}^{(n)} - \eta \sum_{i=1}^{N} (\mathbf{A}^{(n)} x_i - y_i) x_i^T$
(5) $\omega^{(i)} = x^{(i)} - \mu \mathbf{A}^T (\mathbf{A} x^{(i)} - y) + \mu\gamma \mathbf{A}^T \mathbf{A} (x^{(i)} - v^{(i)})$
(6) $u^{(i)} = v^{(i)} + \mu\gamma \mathbf{A}^T \mathbf{A} (x^{(i)} - v^{(i)})$
(7) $x^{(i+1)} = soft(w^{(i)}; \lambda\mu)$
(8) $v^{(i+1)} = soft(u^{(i)}; \lambda\mu)$
(9) **end for**
(10) **return** x^{i+1}

Algorithm 1: Iterative algorithm for *nonconvex sparse regularization by K-SVD* problem.

$$\left.\begin{aligned} &x = x_1 + x_2 \\ &x, x_1, x_2 \in R^N \end{aligned}\right\}. \tag{18}$$

The purpose of the RADWT analysis is to estimate the source signals x_1 and x_2 from the observed signal x, respectively.

3.3. Solving SONNR Problem. In our study, sparsity-oriented nonconvex nonseparable regularization (SONNR) method is proposed to rolling bearing compound fault diagnosis under noisy environment. Four-layer structure of SONNR method is proposed: input layer, nonconvex sparse regularization layer, signal reconstruction layer, and compound faults isolation layer.

Iterative algorithm for our SONNR problem is shown in Algorithm 2.

In view of the coupling characteristics of rolling bearing compound faults and the problem that traditional fault diagnosis methods cannot comprehensively diagnose compound faults, the coupling signals need to be decoupled in order to more effectively diagnose the fault characteristics one by one. As shown in Algorithm 2, The SONNR method has a four-layer structure: input layer, nonconvex sparse regularization layer, signal reconstruction layer, and compound faults isolation layer. The SONNR method can better preserve the bearing fault signal while reducing noise and other interference components; thus it can significantly improve the estimation accuracy of the bearing fault signal. In addition, the compound fault signal can be separated to extract more obvious fault features.

4. Applications in Bearing Compound Fault Diagnosis

4.1. Simulated Compound Fault under Noisy Environment. In order to verify the effectiveness of the SONNR extraction fault response, this method is first used to analyze a simulated compound fault signal.

The rolling bearing fault model is used to simulate the compound fault of the inner ring and the outer ring under the strong noise background. The mathematical model of the inner ring spalling fault of the rolling bearing is shown in

$$\begin{cases} x_i(t) = s(t) + n(t) = \sum_i A_i h(t - iT - \tau_i) + n(t), \\ A_i = A_0 \cos 2(2\pi f_r t + \phi_A) + C_A, \\ h(t) = \exp(-Bt)\cos(2\pi f_n t + \phi_\omega). \end{cases} \tag{19}$$

The mathematical model of the rolling bearing outer ring spalling fault is shown in

$$\begin{cases} x_0(t) = s(t) + n(t) = \sum_i A_i h(t - iT - \tau_i) + n(t), \\ A_i = cons \tan t, \\ h(t) = \exp(-Bt)\cos(2\pi f_n t + \phi_\omega). \end{cases} \tag{20}$$

The mathematical model of the compound fault of the inner ring and the outer ring is shown in

$$x(t) = x_i(t) + x_0(t), \tag{21}$$

where τ_i is a small fluctuation of the i-th impact with respect to the average period T. Assume that the bearing fault is a steel peeled off from the external raceway; the rotating speed of the bearing is 3660 rpm. It is also assumed that the fault characteristic frequency of inner ring is $f_i = 600$ Hz and fault characteristic frequency of outer ring is $f_0 = 500$ Hz. The natural frequency f_d of the rolling bearing system is 9225 Hz, the random slip is assumed to be positively distributed over a standard deviation from 0.5% of the rotational speed, and the sampling frequency f_s is 60 KHz. The simulated signal without noise is shown in Figure 9.

In Figure 10, the random white noise distributed over the positive is added (the signal-to-noise ratio is −10 dB), and the time domain waveform after adding noise is shown in Figure 10. As can be seen from Figure 10, the fault feature is almost overwhelmed due to the background noise variable.

Determine the decomposition parameters of the SONNR: $\lambda = 0.8$, $Q_1 = 4.5$, $Q_2 = 1$, $J_1 = 60$, $J_2 = 15$, $r_1 = r_2 = 3.5$. SONNR sparse decomposition of the fault signal is shown in Figures 11 and 12. It can be seen from Figure 11(b) that there is a spike in the outer ring fault frequency of 500 Hz and its multiplier (1000 Hz, 1500 Hz, 2000 Hz, etc.), and it can be diagnosed that there is an outer ring fault. The inner ring fault frequency is 600 Hz and its multiple frequency 1200 Hz. It can be seen from Figure 12(b) that there is a main peak at 1800 Hz, etc., and a side band appears at the frequency shift (200 Hz) on both sides of the main peak, indicating that

(1) **initialize** $x^{(0)}, v^{(0)}, 0 < \lambda < 1, 0 < \mu < 2/\max\{1, \lambda/(1-\lambda)\}\|A\|_2^{-2}, Q_1, Q_2$
(2) $i = 0, 1, 2, \ldots$ **do**
(3) calculate the coefficients x_i using a simple gradient descent procedure
(4) update the dictionary $\mathbf{A}^{(n+1)} = \mathbf{A}^{(n)} - \eta \sum_{i=1}^{N} (\mathbf{A}^{(n)} x_i - y_i) x_i^T$
(5) $\omega^{(i)} = x^{(i)} - \mu \mathbf{A}^T (\mathbf{A} x^{(i)} - y) + \mu\gamma \mathbf{A}^T \mathbf{A} (x^{(i)} - v^{(i)})$
(6) $u^{(i)} = v^{(i)} + \mu\gamma \mathbf{A}^T \mathbf{A} (x^{(i)} - v^{(i)})$
(7) $x^{(i+1)} = soft(\omega^{(i)}; \lambda\mu)$
(8) $x_1^{(i+1)}, x_2^{(i+1)} = RADWT(x^{(i+1)}, Q_1, Q_2)$
(9) $v^{(i+1)} = soft(u^{(i)}; \lambda\mu)$
(10) **end for**
(11) **return** $x_1^{(i+1)}, x_2^{(i+1)}$

Algorithm 2: Iterative algorithm for SONNR problem.

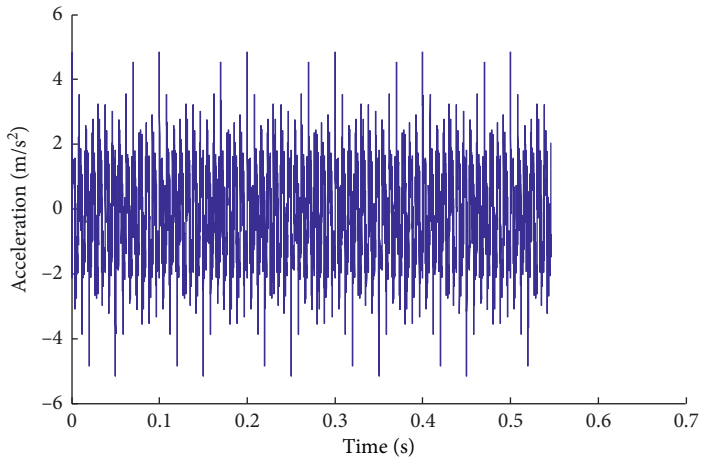

Figure 9: Compound fault signal.

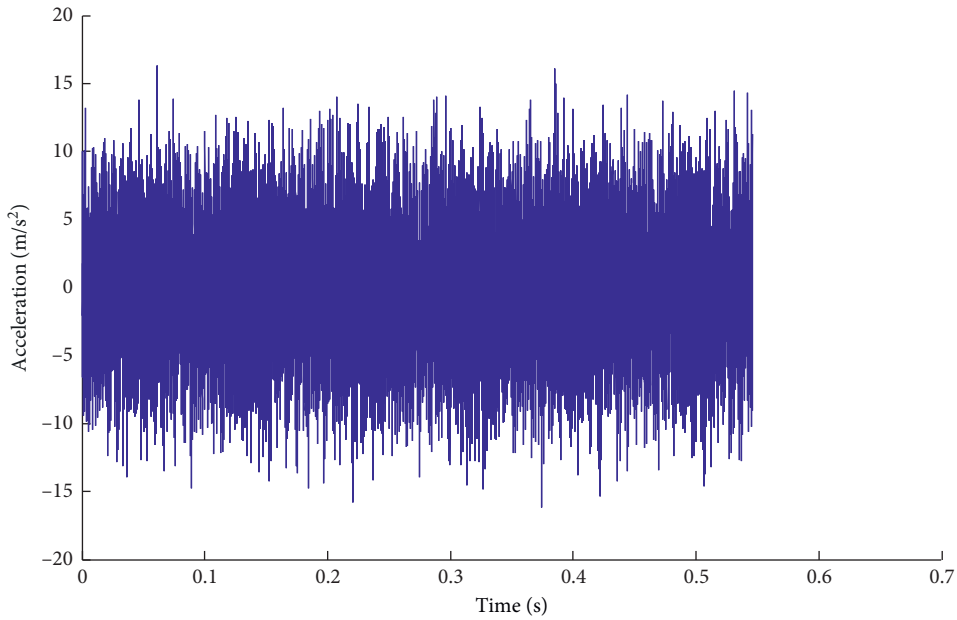

Figure 10: Compound fault signal with noise.

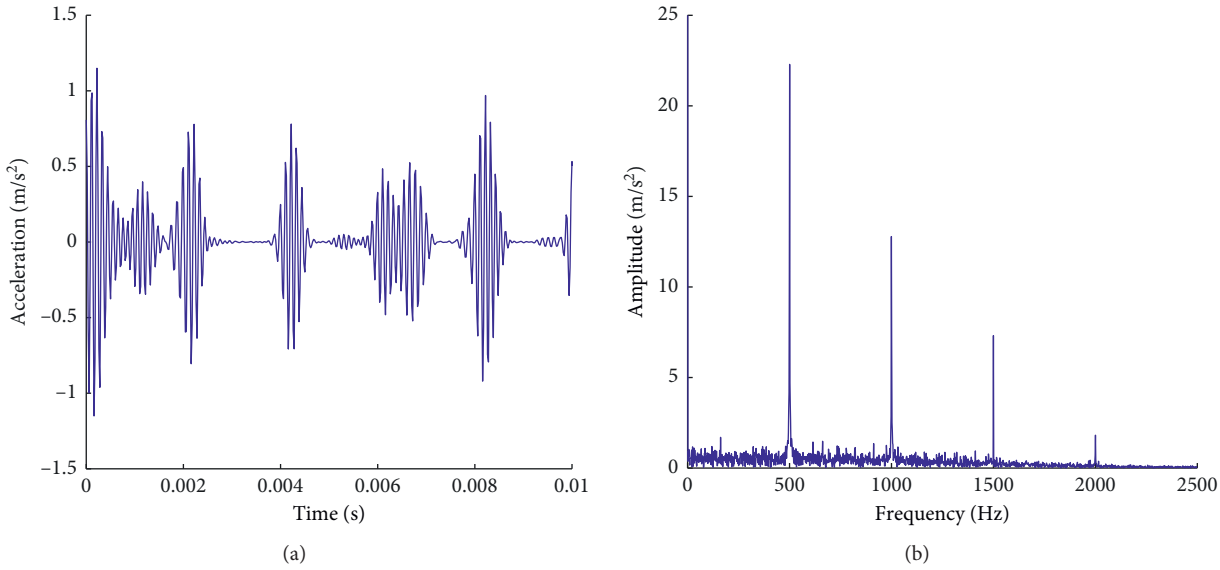

Figure 11: High-quality factor components and their envelope spectra. (a) High-quality factor component waveform. (b) High-quality factor component envelope spectrum.

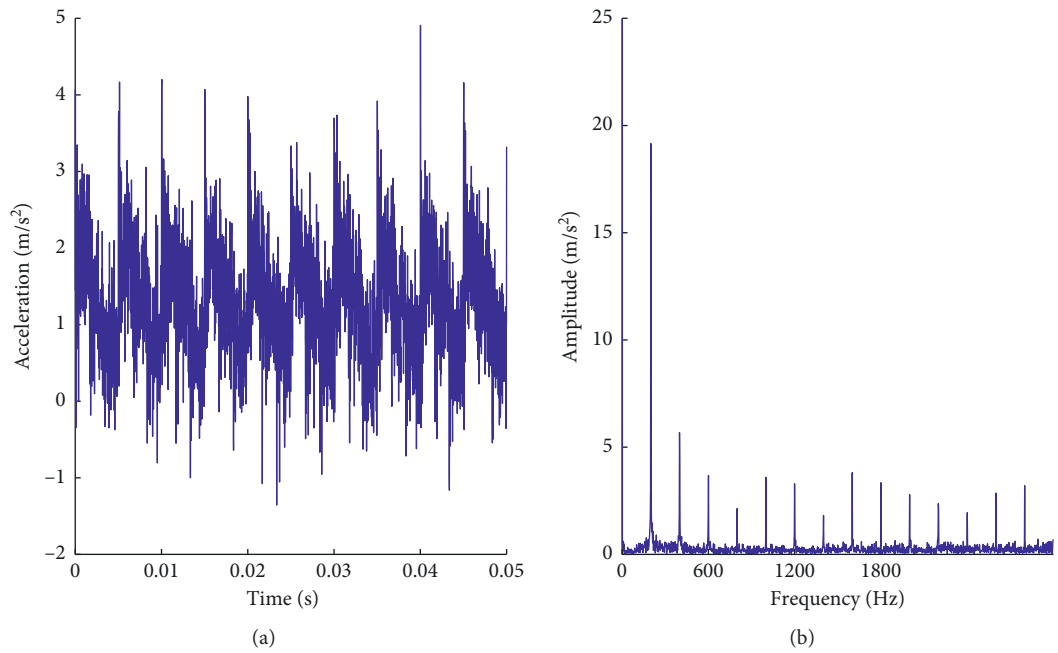

Figure 12: Low-quality factor components and their envelope spectra. (a) Low-quality factor component waveform. (b) Low-quality factor component envelope spectrum.

there is an inner ring fault in the rolling bearing. The results show that it is effective to extract the compound fault information of rolling bearing by the SONNR method.

4.2. Results Based on Accelerated Lifetime (Real Damage) Data Set. The proposed method is validated by benchmark data set of Paderborn University [26], which involves compound damage to various operating conditions. This benchmark generates real bearing damage samples by accelerating lifetime tests. The option to compound fault with measurement data onto real bearing damage: different damage symptoms occur in the bearing or identical damage symptoms occur on different bearing components. Our test was carried out by including data sets from bearing with multiple damage at both raceways, at inner and outer rings (KB23). The compound fault form of KB23 is shown in Figure 13. The computational complexity is $O(n{\cdot}d)$, the length of the sequence n is 3840, and the computational dimension d is 1.

The manufacturer specific information about bearing KB23 is shown in Table 1.

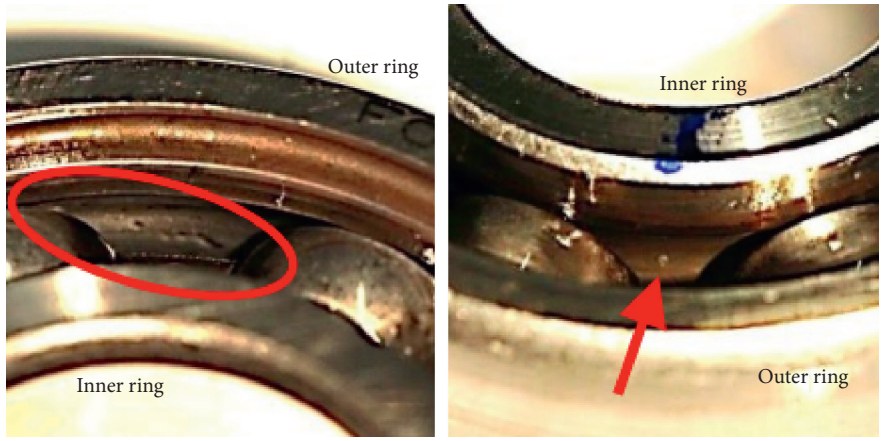

Figure 13: KB23: spalling at the ring of the inner ring and outer ring [26].

4.2.1. Step 1: Noise Signal Preparation. In order to verify the effectiveness of the proposed method of fault classification in noisy environments, we add Gaussian white noise to each data set. The noise n is white Gaussian noise with standard deviation $\sigma = 0.6$. As shown in Figure 14, the bearing KB23 fault signal, the additive white Gaussian noise signal, and the frequency domain signal of the KB23 fault signal are shown from top to bottom. It can be seen that the signal is seriously polluted by noise, and the time domain signal and the frequency domain signal cannot distinguish the vibration characteristics of the original model. Therefore, it is very difficult to extract effective fault information about a signal with noise.

4.2.2. Step 2: Nonconvex Sparse Regularization by K-SVD. We use GMC with a fixed regularization parameter λ to analyze the simulation signal of inner fault with noisy signal. In order to explore the effect of GMC on the performance of the sparse regularized method, we compare this method with the L1 regularization method.

The L1-regularized least squared problem is

$$\min_x \left\{ J(x) = \frac{1}{2}\|y - Ax\|_2^2 + \lambda \|x\|_1 \right\}. \tag{22}$$

In this case, the parameters of GMC nonconvexity and L1 regularization are set to $\gamma = 0.8$. A is an oversampled inverse discrete Fourier transform and the result of Fourier coefficients is shown in Figure 15(c) by K-SVD. The performance is evaluated by the average root-mean-squared error (RMSE). Sparse result by GMC and L1 is shown in Figures 15(a) and 15(b). The RMSE of the GMC and L1-norm methods is 0.129 and 0.149. The comparison results show that the GMC is more accurate and uses fewer coefficients and yields more accurate denoising result.

4.2.3. Step 3: Compound Fault Isolation. The fault characteristic frequency of the outer ring is

$$f_0 = \frac{z}{2} f_r \left(1 - \frac{B_d}{P_d} \cos\phi \right). \tag{23}$$

Table 1: Manufacturer specific information about bearing.

Criterion	Parameters
Bearing type	Ball bearing 6203 (KB23)
Diameter of inner raceway	24.0 mm
Diameter of outer raceway	33.1 mm
Pitch circle diameter	28.55 mm
Number of rolling elements	8 pc
Rolling element diameter	6.75 mm
Length of rolling element	6.75 mm
Nominal pressure angle	0°N
Static load rating	3800 N
Dynamic load rating	3800 N
Speed limit	12000 rpm
Number of load cycles	2769500
Lifetime	15:01 h:min
Damage method	Acc. lifetime test
Sampling rate	64 kHz

The fault characteristic frequency of the inner ring is

$$f_i = \frac{z}{2} f_r \left(1 + \frac{B_d}{P_d} \cos\phi \right). \tag{24}$$

Determine the decomposition parameters of the RADWT: $Q_1 = 4$, $Q_2 = 1$, $J_1 = 30$, $J_2 = 15$, $r_1 = r_2 = 3.5$. Taking the parameters of the bearing into (16) and (17), we can get $f_i = 123.3$, $f_o = 76.25$.

From Figure 16, we can know that the RADWT decomposes a complex signal into the sum of two different components. The RADWT nonlinearly separates the components in the signal according to the oscillation characteristics and establishes the optimal sparse representation of each of the high resonance components (Figure 16(b)) and the low resonance components (Figure 16(c)). We perform envelope analysis on Figures 16(b) and 16(c), respectively, and we can get Figures 17 and 18.

It can be seen from Figure 17 that there is a spike in the outer ring fault frequency of 76.25 Hz and its multiplier, and it can be diagnosed that there is an outer ring fault. Figure 18 shows the fundamental rotation frequency of the shaft and its harmonics (25 Hz), the ball pass frequency of the inner race (123.3 Hz), and its sidebands as well as its

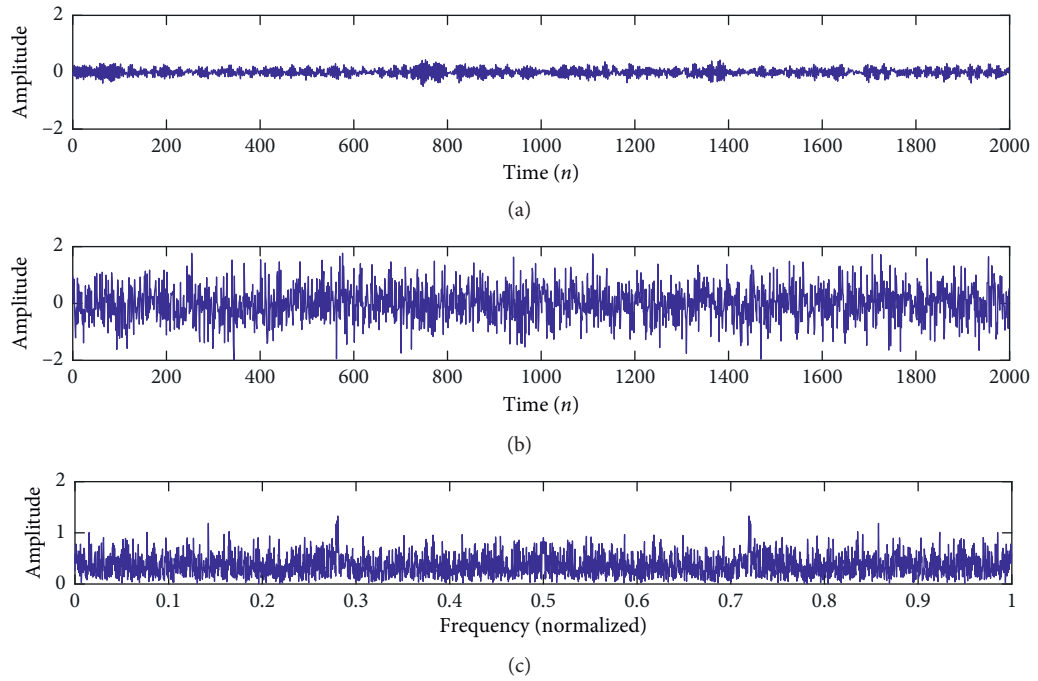

Figure 14: KB23 fault signal with noise. (a) KB23 fault signal. (b) Additive white Gaussian noise with $\sigma = 0.6$. (c) FFT of noisy inner fault signal.

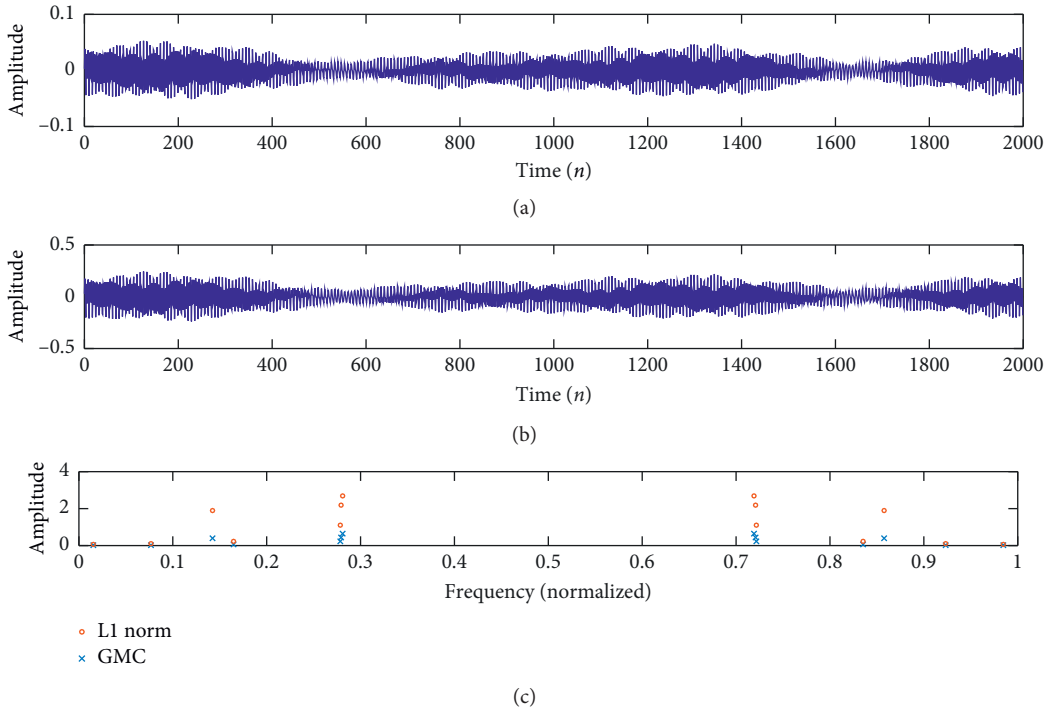

Figure 15: Sparse result by GMC and L1. (a) Denoising using GMC penalty. (b) Denoising using L1 penalty. (c) Optimized coefficients.

corresponding harmonics. Figure 18 indicates that there is an inner ring fault.

The results show that it is effective to extract the compound fault information of rolling bearing by the SONNR method.

4.3. Comparison with Time-Domain Diagnostic Methods. Envelope analysis, EMD, EEMD, and EEMD-PCA are commonly used to extract the rolling bearing fault information. The traditional analysis techniques are applied to the vibration signal of KB23.

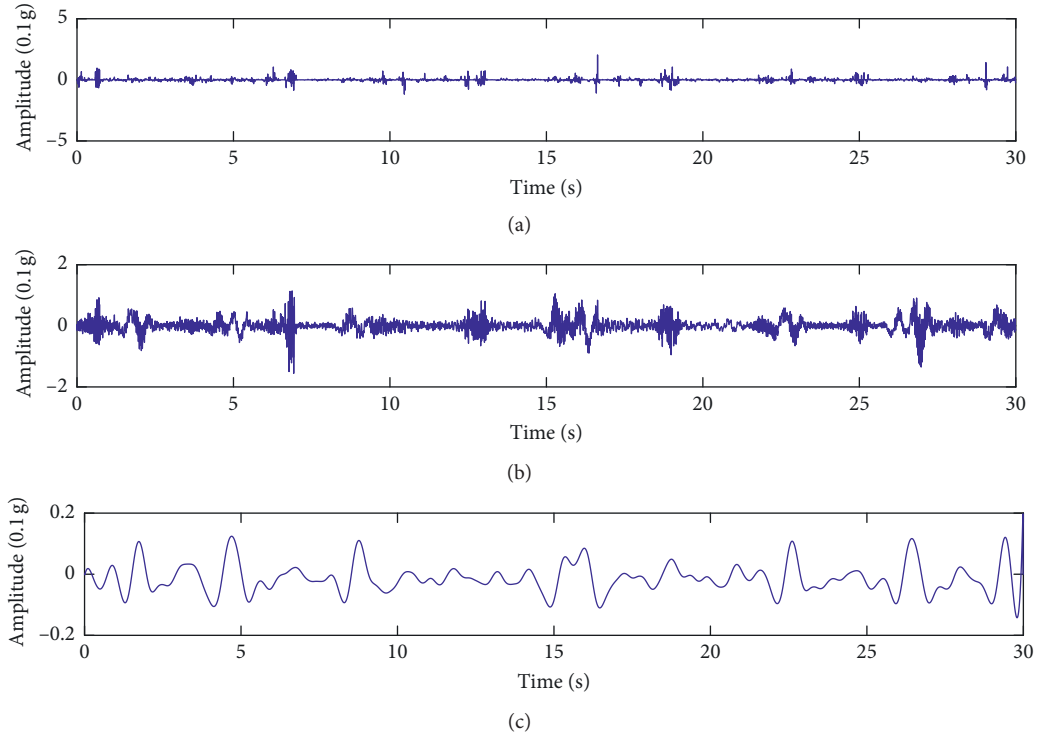

Figure 16: Compound fault isolation result by RADWT. (a) Reconstructed signal using GMC penalty. (b) High-quality factor components. (c) Low-quality factor components.

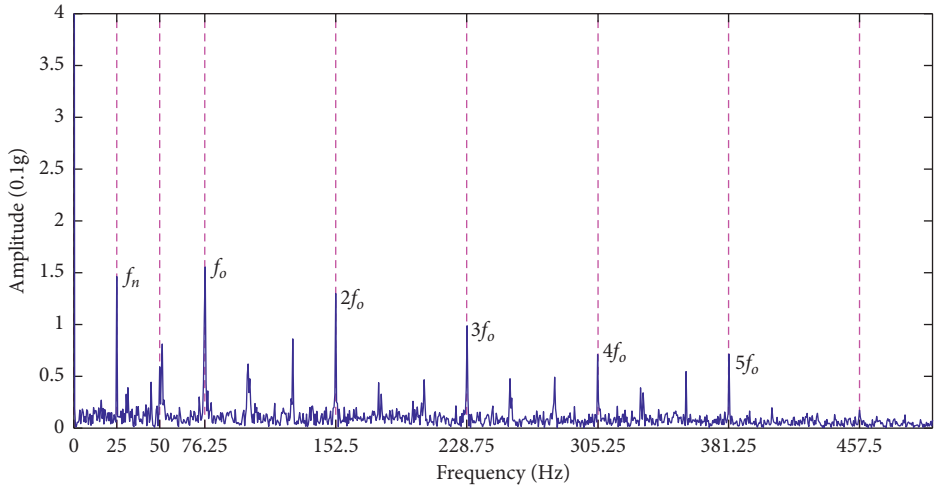

Figure 17: High-quality factor component envelope spectrum.

Envelope analysis result of KB23 is shown in Figure 19. From Figure 19, it is difficult to find the characteristic frequency of the fault, so the traditional envelope analysis is difficult to reveal the characteristic frequency of the compound fault shock response.

EMD and EEMD analysis results of KB23 are, respectively, shown in Figures 20 and 21. Through EMD decomposition, we can get a series of IMF with different frequencies, which can be used to extract the characteristics of the original signal. From Figures 20(b) and 21(b), it is difficult to find the characteristic frequency of the fault, so the EMD and EEMD analysis is difficult to reveal the characteristic frequency of the compound fault shock response.

EEMD-PCA analysis results of KB23 are, respectively, shown in Figures 22 and 23. In this section, we reconstruct the original compound fault signal KB23 by EEMD-PCA method. The reconstructed picture is shown in Figure 22. From Figure 22, we can find that the EEMD-PCA method plays a role in noise reduction. Next, spectrum analysis is

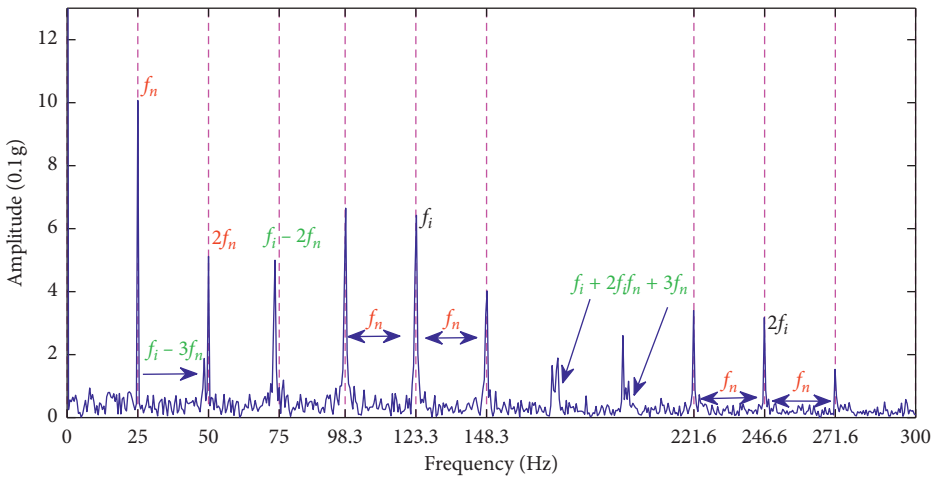

Figure 18: Low-quality factor component envelope spectrum.

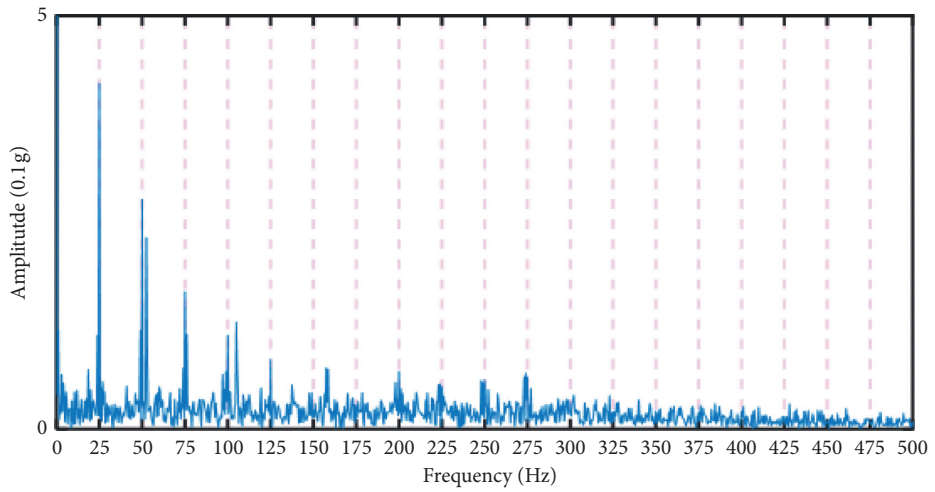

Figure 19: Envelope analysis of KB23.

performed on the reconstructed signal, and the result is shown in Figure 23. From Figure 23, it is difficult to find the characteristic frequency of the fault, so the EEMD-PCA analysis is difficult to reveal the characteristic frequency of the compound fault shock response.

The comparison experiments with traditional envelope analysis and EMD, EEMD, and EEMD-PCA analysis show that the proposed technique SONNR of fault information extraction has fewer spectral glitches and a higher signal-to-noise ratio. And this method can reveal the amplitude modulation frequency of fault shock response more than traditional methods.

4.4. Comparison with Artificial Intelligence Methods. Further tests were compared with artificial intelligence methods. The test was carried out by including data sets from bearings with compound damage at both raceways, at inner and outer ring (KB23, KB24, and KB27) in [22]. The benchmark data set included 26 damaged bearing states (12 outer faults, 3 compound faults, and 11 inner faults) and 6 undamaged (healthy) states for reference. Each bearing state included 80 measurements. Therefore, our training types are divided into 4 categories (class 1: healthy; class 2: outer fault; class 3: inner fault; class 4: compound fault), 70% of which are used to train the network, and the remaining 30% are used for testing. The experiments are tested in Python 3.7 software platform using a laptop equipped with Intel-i5 3.2 GHz CPU, 8 GB memory. At the end of this section, experimental results are verified by comparing with commonly used diagnostic methods including stacked auto encoder (SAE), another version of stacked auto encoder DBN, and convolutional neural network (CNN). For the neural networks, parameters such as the number of hidden layers and neurons were not tuned and this explains the poor performance.

It can be seen from Table 2 that the artificial intelligence methods get the poor performance. It can be seen that the artificial intelligence algorithm is not very effective for the classification of compound faults. This also makes the

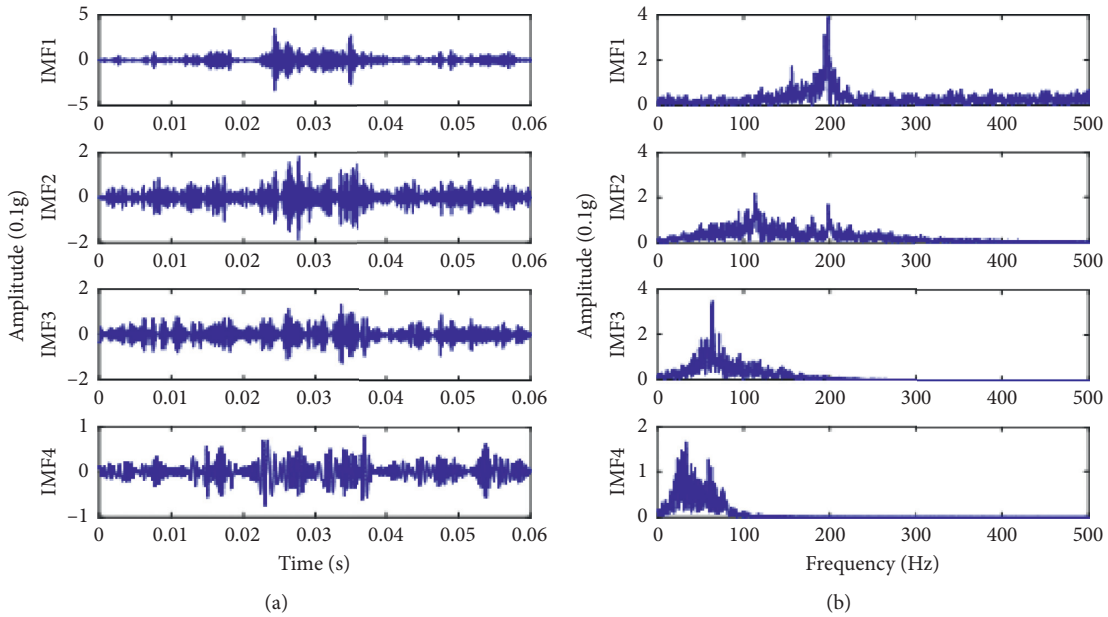

Figure 20: EMD analysis of KB23. (a) IMFs. (b) Spectrum of IMFs.

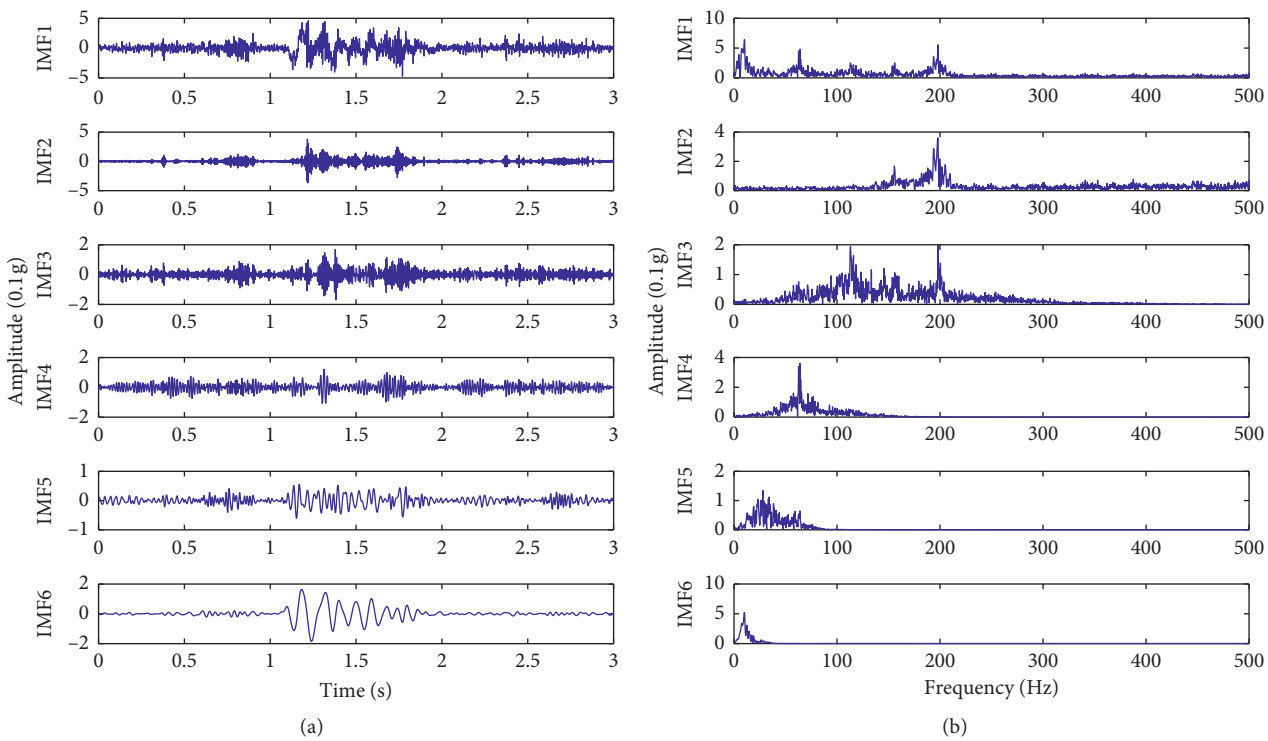

Figure 21: EEMD analysis of KB23. (a) IMFs. (b) Spectrum of IMFs.

application of artificial intelligence algorithms in compound fault diagnosis a challenging topic in the future.

5. Conclusions

In this paper, we present the SONNR method for rolling bearing compound fault diagnosis. The main advantage of the SONNR over conventional fault diagnosis methods is that SONNR can better preserve the bearing fault signal while reducing the interference of noise and other components; thus it can reveal the periodic fault impact mode in the bearing signal. The physical model of the compound fault is established, and the modulation characteristics of the compound fault signal are obtained from the analysis of the envelope spectrum, which lays a theoretical foundation for the subsequent fault information extraction. The nonconvex

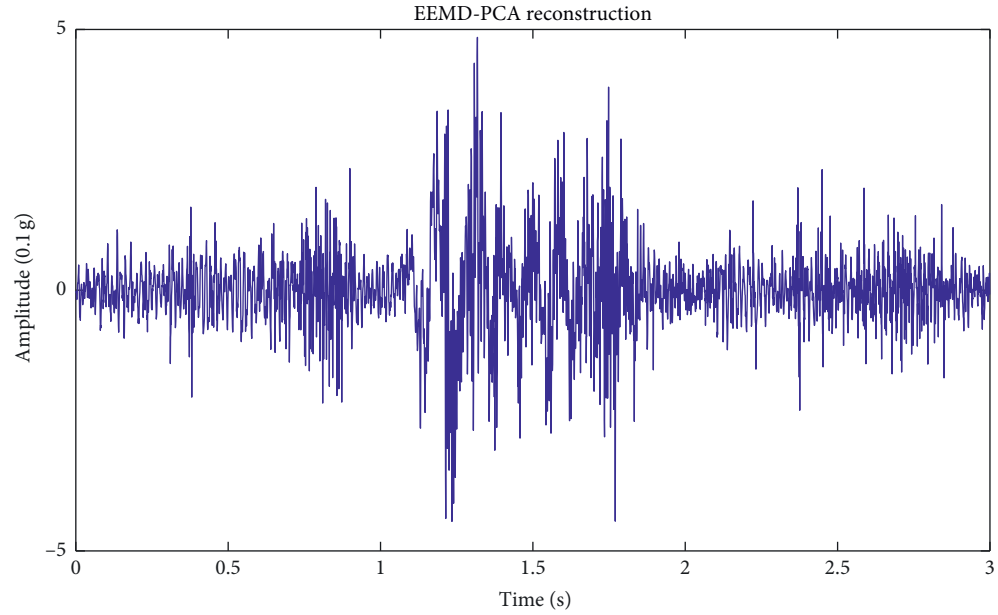

Figure 22: EEMD-PCA reconstruction analysis of KB23.

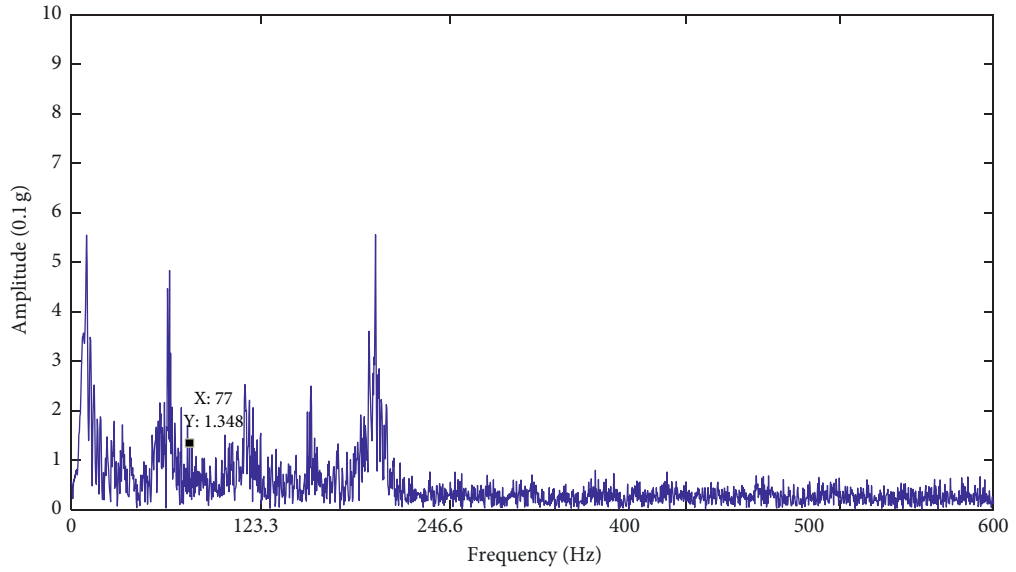

Figure 23: Spectrum of reconstruction signal.

Table 2: Classification accuracy on MFS data set.

Method	Accuracy (%)	Training time (s)
SAE	63.4272	20.41032
DBN	64.8357	20.53823
CNN	65.3333	11.54596

optimization method for compound faults under noisy environment has been considered in our studies. Exploiting the high-amplitude components property, a K-SVD-based nonconvex nonseparable regularization denoising algorithm is presented to extract bearing fault features. For compound faults, RADWT can well reveal the amplitude modulation frequency of the fault shock response. Finally, we present a verification based on real damage (accelerated lifetime) data set to validate the effectiveness of the proposed method in the diagnosis of compound faults in rolling bearing. In particular, we have artificially added noise based on the acquired signals. The comparison experiments with

traditional envelope analysis and EMD, EEMD, and EEMD-PCA analysis show that the proposed SONNR method has more advantages in complex fault diagnosis than traditional time domain method. Further tests were compared with artificial intelligence methods. The result shows that the artificial intelligence methods are not very effective for the classification of compound faults. This also brings challenges to future research.

Data Availability

The data (a benchmark data set for condition monitoring of rolling bearings) used to support the findings of this study have been deposited in the KAtDataCenter website of the Chair of Design and Drive Technology, Paderborn University, Germany, repository (http://mb.uni-paderborn.de/kat/datacenter).

Conflicts of Interest

The authors declare no potential conflicts of interest.

Acknowledgments

This work was supported by National Natural Science Foundation of China (nos. 61533013 and 61633019), Shaanxi Provincial Key Project (2018ZDXMGY-168), and Shanghai Project (17DZ1202704).

References

[1] J. Liu and Y. Shao, "Overview of dynamic modelling and analysis of rolling element bearing with localized and distributed faults," *Nonlinear Dynamics*, vol. 93, no. 4, pp. 1765–1798, 2018.

[2] F. Cong, J. Chen, G. Dong, and M. Pecht, "Vibration model of rolling element bearings in a rotor-bearing system for fault diagnosis," *Journal of Sound and Vibration*, vol. 332, no. 8, pp. 2081–2097, 2013.

[3] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Vibrations based performance prediction of ball bearings caused by localized defects," *Nonlinear Dynamics*, vol. 69, no. 3, pp. 847–875, 2012.

[4] I. Gondal, M. Amar, and C. Wilson, "Vibration spectrum imaging: a novel bearing fault classification approach," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 494–502, 2014.

[5] J. Lundén, S. A. Kassam, and V. Koivunen, "Robust nonparametric cyclic correlation-based spectrum sensing for cognitive radio," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 38–52, 2009.

[6] V. Girondin, K. M. Pekpe, H. Morel, and J.-P. Cassar, "Bearings fault detection in helicopters using frequency readjustment and cyclostationary analysis," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 499–514, 2013.

[7] P. Borghesani and M. R. Shahriar, "Cyclostationary analysis with logarithmic variance stabilisation," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 51–72, 2016.

[8] B. Pang, G. Tang, T. Tian et al., "Rolling bearing fault diagnosis based on an improved HTT transform," *Sensors*, vol. 18, no. 4, pp. 1203–1208, 2018.

[9] Y. Lv, Q. Zhu, and R. Yuan, "fault diagnosis of rolling bearing based on fast nonlocal means and envelop spectrum," *Sensors*, vol. 15, no. 1, pp. 1182–1198, 2015.

[10] X. Chen, F. Feng, and B. Zhang, "weak fault feature extraction of rolling bearings based on an improved kurtogram," *Sensors*, vol. 16, no. 9, p. 1482, 2016.

[11] L. Yong, H. Zhang, and C. Yi, "Trivariate empirical mode decomposition via convex optimization for rolling bearing condition identification," *Sensors*, vol. 18, no. 7, pp. 2325–2330, 2018.

[12] X. Chen, D. Liu, G. Xu, K. Jiang, and L. Liang, "Application of wavelet packet entropy flow manifold learning in bearing factory inspection using the ultrasonic technique," *Sensors*, vol. 15, no. 1, pp. 341–351, 2014.

[13] 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.

[14] 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.

[15] Y. Qin, "A new family of model-based impulsive wavelets and their sparse representation for rolling bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2716–2726, 2018.

[16] R. A. Kanai, R. G. Desavale, and S. P. Chavan, "Experimental-based fault diagnosis of rolling bearing using artificial neural network," *Journal of Tribology*, vol. 138, no. 3, 2016.

[17] Q. Fu, B. Jing, P. He, S. Si, and Y. Wang, "Fault feature selection and diagnosis of rolling bearing based on EEMD and optimized Elman_AdaBoost algorithm," *IEEE Sensors Journal*, vol. 18, no. 12, 2018.

[18] H. Cao, L. Niu, S. Xi, and X. Chen, "Mechanical model development of rolling bearing-rotor systems: a review," *Mechanical Systems and Signal Processing*, vol. 102, pp. 37–58, 2018.

[19] X. Li, W. Zhang, and Q. Ding, "A robust intelligent fault diagnosis method for rolling element bearing based on deep distance metric learning," *Neurocomputing*, vol. 310, pp. 77–95, 2018.

[20] I. Selesnick and M. Farshchian, "Sparse signal approximation via nonseparable regularization," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2561–2575, 2017.

[21] I. Selesnick, "Sparse regularization via convex analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4481–4494, 2017.

[22] S. Wang, I. W. Selesnick, G. Cai et al., "Nonconvex sparse regularization and convex optimization for bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, p. 7332, 2018.

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

[24] I. Bayram and I. W. Selesnick, "Frequency-domain design of overcomplete rational-dilation wavelet transforms," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 2957–2972, 2009.

[25] I. W. Selesnick and I. Bayram, "Oscillatory + transient signal decomposition using overcomplete rational-dilation wavelet transforms," in *Proceedings of SPIE the International Society for Optical Engineering*, p. 7446, San Diego, CA, USA, August 2009.

[26] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: a benchmark data set for data-driven classification," in *Proceedings of the 3rd European Conference of the Prognostics and Health Management Society*, Bilbao, Spain, July 2016.