



An improved spectral amplitude modulation method for rolling element bearing fault diagnosis

Danchen Zhu¹ · Bolong Yin¹ · Chengxin Teng¹

Received: 19 December 2022 / Accepted: 28 March 2023 / Published online: 13 April 2023
© The Author(s), under exclusive licence to The Brazilian Society of Mechanical Sciences and Engineering 2023

Abstract

The fault signatures in the real bearing vibration signals are relatively weak and are usually buried by strong background noise due to the influence of multiple structures and the complex transmission path of the equipment, which brings many difficulties to the accurate and effective fault diagnosis of rolling element bearing. To address this problem, an improved spectral amplitude modulation method was proposed in this paper. First, in the process of SAM calculation, signal distortion and interference components enhancement due to the existence of transmission path were suppressed in advance to improve the analysis effect. Second, the selection range of magnitude order (MO) was determined from 0 to 2, and an optimal MO selection strategy while the maximum value of the harmonics significant index was employed as the target. Finally, the complexity of the bearing vibration signal was taken into consideration, the whole frequency band was divided and the maximum value of the signals' sparsity was set as the target for signal reconstruction, which further suppressed the interference of irrelevant components and the fault characteristics were highlighted. Abundant fault features could be extracted in the envelope spectrum of the reconstruction signal. The effectiveness of this approach was verified using the simulation bearing inner race fault signal, the measured bearing inner and outer race fault signals. The results show that the method in this manuscript can extract the bearing fault signatures accurately from the strong background interference.

Keywords Spectral amplitude modulation · Transmission path · Resonance band selection · Rolling element bearing · Fault diagnosis

1 Introduction

Rolling element bearings (REBs) are widely employed in rotating machinery as an important part. However, as REBs usually work in high-temperature and high-pressure conditions, the probability of their failure is high, which greatly threatens the stable operation of the whole equipment. The bearing failures are often accompanied by the generation of periodic impact components in the vibration signal, whose frequency is known as the bearing fault characteristic frequency (FCF). As a result, to ensure the safe and stable operation of the equipment and avoid unnecessary losses, it is of great importance to research the fault diagnosis of REB, which key point is to extract a series of defect features.

However, the fault characteristics of REB are weak, and there are always multi-structures in the machinery, they are also the vibration sources and lead to the existence of complex transmission path, which can make the actual bearing vibration signals buried by strong interference parts. Moreover, the signal distortion may also appear. Therefore, how to suppress the interference of background noise and signal distortion, how to select the appropriate analysis method to highlight the fault characteristics components, and improve the signal-to-noise ratio have become important research contents.

To solve these problems, scholars have conducted extensive research, and one of the most classical methods is resonance demodulation. Fast Kurtogram [1] is a fast signal analysis method proposed by Antoni, which quickly determines the location of the optimal analysis frequency band through the spectral kurtosis index. Since then, many research has been carried out on how to select the resonance bands more effectively, and numerous improvements have been made, mainly in the parameters selection and band division

Technical Editor: Jarir Mahfoud.

✉ Danchen Zhu
bryantzdc@sina.com

¹ Naval Petty Officer Academy, Bengbu 233012, China

approaches. Wang [2] proposes the TIEgram method, the traversal segmentation model is used to enhance the 1/3 binary tree structure segmentation model in Fast Kurtogram, meanwhile, a fusion index which combines the kurtosis, the correlation coefficient and the spectral negative entropy is constructed for optimal band selection. The harmonic spectral kurtosis index is proposed by Zhang [3] and applied to the periodic impact components detection, which achieves good results. The fault energy of correntropy is used by Ni [4] to determine the location of the optimal frequency band. The Enfigram technique is proposed by Xu [5], while the division rule of the whole frequency band is improved and the envelope Fourier index is also employed. Zhang [6] proposes the Ewtfergram technique, the improved spectrum segmentation method, and the weighted time–frequency energy ratio are utilized to reconstruct the band segmentation and optimal band selection. In addition, other methods such as adaptive correlated kurtogram [7] and subband average kurtogram [8], can also improve the accuracy of effective analysis frequency band selection. However, problems are still not solved as the heavy noise can also affect the determination of the appropriate analysis frequency band.

Signal decomposition techniques have received great attention as the complex components of real bearing vibration signals. The signals are decomposed adaptively into mode functions which contain different frequency components according to the characteristics of the signal, and the fault signatures can be enhanced by selecting the effective components and discarding the irrelevant ones. Some usually used signal decomposition techniques are the empirical mode decomposition (EMD) [9, 10], the local mean decomposition (LMD) [11], the variational mode decomposition (VMD) [12–15], the singular value decomposition (SVD) [16, 17], the empirical wavelet decomposition (EWT) [18–20], and so on. The application of blind deconvolution algorithms considers the suppression of the transmission path's influence, for example, the minimum entropy deconvolution (MED) [21, 22], the maximum correlated kurtosis deconvolution (MCKD) [23, 24] and the maximum second-order cyclostationarity blind deconvolution (CYCBD) [25, 26]. All the above blind deconvolution approaches have been widely employed in bearing fault diagnosis and have achieved good results. For either the signal decomposition approaches or the blind deconvolution algorithms, the proper selection of calculation parameters is extremely important. Numerous improvements are conducted relying on the optimization algorithms, and though the ideal results are achieved, the analysis efficiency is hard to be guaranteed.

The spectral amplitude modulation (SAM) method is proposed to effectively extract the different frequency components of the signal [27], which is calculated mainly by applying the magnitude order (MO) to the amplitude spectrum of the original signal, changing the distribution characteristics

of different frequency components, and then obtaining the improved signal with the help of the phase information of the original signal and the inverse Fourier transform. The SAM method has been shown to achieve good results in fault feature extraction of REB [28, 29], where the calculation process of the SAM is optimized with the MO value is selected more properly; however, influence of the transmission path is not taken into consideration, and it is still worth discussing how to further improve its analysis effect.

According to the description of some existing approaches, especially the advantages and disadvantages of the newly proposed SAM-based approaches, an ISAM approach is proposed in this manuscript for the diagnosis of REB to solve the problem of extracting bearing fault features under strong background noise. First, based on the characteristics of the signal itself, the influence of the transmission path is suppressed in the analysis process of SAM to reduce the signal distortion and initially remove the noise interference. Then the maximum harmonics significant index (HSI) is utilized as the target to determine the appropriate value of the MO. Second, the optimal resonance band is determined and employed for signal reconstruction with the sparsity value as the index. Finally, the fault signatures can be detected in the envelope spectrum of the reconstruction signal. The main contributions of this manuscript are described as follows.

- (1) The influence of the transmission path is suppressed during the SAM calculations, which reduces the signal distortion and preliminarily removes the noise interference.
- (2) A selection strategy of MO is constructed to avoid blindness chosen of the parameter in the analysis based on the principle of maximum HSI value.
- (3) An optimal resonance band selection and signal reconstruction strategy based on the maximum sparsity index is proposed to further highlight the fault characteristics of the signal and remove irrelevant interference.

The rest of this paper is organized as follows. The basic principles are introduced in Sect. 2 including the theory of the SAM and its improvements. Section 3 summarizes the calculation steps of the proposed method. The simulation and the experimental signals are employed, respectively, in Sects. 4 and 5 to demonstrate the effectiveness of the proposed technique. At last, conclusions are drawn in Sect. 6.

2 The basic principles

2.1 The SAM method

SAM is constructed based on the theory of the cepstrum, it can preserve both the phase and amplitude characteristics

of the signal, which can reflect the signatures of the whole spectrum and can be applied to the fault detection of REBs. SAM can be used without any pre-defined parameters, such as the central frequency and bandwidth of the filter.

Suppose the signal $x(t)$ can be transformed from the time domain to the frequency domain using the Fourier transform (FT). The amplitude and phase information can be denoted as $A(f)$ and $\varphi(f)$, respectively, whose relationship can be expressed as:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt = A(f)e^{j\varphi(f)} \quad (1)$$

where j denotes the imaginary units, and $X(f)$ denotes the FT results of $x(t)$. The basic principle of SAM is that the phase contains more useful information than the amplitude, hence the phase $\varphi(f)$ needs to be kept unchanged during the signal analysis and reconstruction process, and only the amplitude $A(f)$ needs to be processed. The calculation process of the SAM can be expressed as follow:

$$x_m(t, MO) = \int_{-\infty}^{\infty} A(f)^{MO} e^{-j(\varphi(f)+2\pi ft)} df \quad (2)$$

where x_m denotes the improved signal which is obtained using the improved amplitude spectrum, the original phase spectrum and the inverse FT. According to the results shown in Ref.[27], the range of the weight MO is recommended to choose between -0.5 and 1.5 . It can be observed from Eq. (2) that when $MO=0$, the SAM is equal to the cepstrum pre-whitening; when $MO=1$, the signal will be unchanged throughout the whole process. Although the background interference in the signal is difficult to avoid, the fault characteristic embedded in the signal can be enhanced by selecting a suitable MO value.

For the recommended MO range, its effects on the signal can be divided into three kinds. When $MO > 1$, the components with larger amplitudes in the frequency domain will be enhanced and will be clearly distinguished from the components with smaller amplitudes, which can be considered as denoising. When $0 < MO < 1$, the frequency components with larger amplitudes weakened, while the components with smaller amplitudes will be relatively enhanced, and both the two parts will be retained, so it is also the commonly used MO range in fault diagnosis of REB. When $MO < 0$, the amplitude will have the opposite effect, and if the signal is dominated by noise, the smaller amplitudes frequency components will be effectively extracted through the SAM processing.

(1) To summarize the SAM method, the features enhanced signal is acquired using the SAM by applying the weight MO to the amplitude spectrum of the original signal, which has been proven to achieve good

results in bearing fault feature extraction. However, as the same MO is applied to the amplitude spectrum of the signal, it is likely to result in a significant energy enhancement of irrelevant interference components, which reveals that not all values of MO can achieve good results. Therefore, some shortcomings still need to be improved in the SAM.

- (2) Signal denoising in the frequency domain can be carried out before applying the MO to the signal which should not rely on too much prior knowledge of the signal characteristics. Then the effects of fault feature enhancement are improved, and the phenomenon of both defect features and noise components enhancement is avoided.
- (3) The values of MO have a certain influence on the analysis results of the signal, and different MO values should be selected for signals with different characteristics, thus, how to select the appropriate MO is also a problem to be considered.
- (4) Applying the envelope analysis directly to the analysis results of SAM will retain the signal characteristics of the full frequency band, and both the fault components and the interference parts will be reflected, therefore, the further selection of a proper analysis band is conducive to highlight the fault feature components.

2.2 The improved SAM method

To address the above problems in the SAM method, the following improvements are made.

- (1) The amplitude spectrum $A(f)$ of the signal $x(t)$ obtained by FT is analyzed based on the transmission path's influence suppression technique described in Ref.[15], and the signal distortion and interference components enhancement due to the influence of the transmission path is suppressed in advance. The Savitzky-Golay filtering (S-G filtering) is employed to simulate the trend line, the window length of the S-G filtering, and the polynomial order is set to 451 and 5, respectively.
- (2) The weight MO is applied to the transmission path's influence suppressed amplitude spectrum, the recommended range of MO is $-0.5-1.5$, as this paper is to analyze the actual bearing fault signals, and when $MO=0$, the signal is unchanged after analysis, only $MO > 0$ is taken into consideration. As the fault features in bearing vibration signals are often submerged by strong background interference, if the value of MO is too large, the noise components in the signal will also be significantly enhanced, which is not conducive to the effective extraction of defect signatures. Hence, the search range of MO is determined as $0 < MO \leq 2$ in this manuscript.

To obtain the optimal fault signatures extraction, the value of MO needs to be selected accurately. The harmonic significant index (HSI) can be used to detect the significance of harmonic components spaced at ω in the frequency spectrum. When ω is set to the fault characteristic frequency if the fault features are more obvious, the HSI value is larger, the HSI can be expressed as:

$$H(\omega) = \left[\frac{F_M(\omega)}{N(\omega)} \times \frac{F_M(2\omega)}{N(2\omega)} \times \dots \times \frac{F_M(K\omega)}{N(K\omega)} \right]^{1/K} = \left[\prod_{k=1}^K p(k\omega) \right]^{1/K} \tag{3}$$

$$F_M(K\omega) = \max \{F[K(\omega - 5) : K(\omega + 5)]\} \tag{4}$$

where $F(\omega)$ indicates the amplitude at frequency ω in the envelope spectrum, which is the amplitude at the FCF. $F_M(\omega)$ is employed to solve the situation that the actual characteristic frequency doesn't correspond to the theoretical value. $K=3$ is the maximum harmonic order. $N(\omega)$ is the average level of noise around frequency ω (the average of the spectral lines' amplitude around ω spaced at the rotating frequency). $P(\omega)$ denotes the significance of the characteristic frequency components.

- (3) The optimal resonance frequency band is selected in the SAM-processed signal, and then the envelope spectrum is calculated. Many approaches can determine the location of the resonance frequency band, whose aim is to select the proper central frequency and bandwidth. To ensure the valid fault feature extraction of bearing fault signals, the analysis frequency band is usually selected at least three times the FCF. In this paper, the bandwidth is determined as four times the FCF, and $m_{band} = \text{ceil}(4 f_g)$, f_g represents the FCF, ceil denotes round the elements to the nearest the integers greater than or equal to it. And the final envelope spectrum should ensure to exhibit the frequency range of at least $3 f_g$.

To determine the optimal resonance frequency band, it is still necessary to choose the appropriate target index. Considering that the HSI may be disturbed by noise components whose frequency is similar to fault feature components, which may lead to errors in analysis results. A sparsity indicator is introduced and calculated in the time domain, which has a strong sensitivity to irrelevant noise. The sparsity is calculated as:

$$S = \frac{\sqrt{1/N \sum_{i=1}^N x_i^2}}{1/N \sum_{i=1}^N |x_i|} \tag{5}$$

where x_i represents the i th data in signal x . The location of the optimal resonance frequency band can be then determined based on this indicator. It is found that when this paper's method is employed to decompose the signal, there exist several other resonance frequency bands that share similar sparsity values with that of the optimal resonance frequency band. Therefore, to further enrich the fault characteristics of single-band signals, five band signals are selected with the largest sparsity values for signal reconstruction, and then the bearing defect signatures can be detected in the reconstruction signal.

3 The whole structure of the proposed method

Based on the previous description, an ISAM method is proposed in this manuscript and is applied to real fault diagnosis of REB which achieves the suppression of irrelevant interference components and the enhancement of fault feature components in vibration signals. The flow chart of this technique is depicted in Fig. 1. The detailed steps are described below.

Step 1: The vibration signal of REB is acquired.

Step 2: The amplitude spectrum and phase spectrum of the original signal are obtained using the FT. The influence of the transmission path is suppressed in the amplitude spectrum.

Step 3: The search range of MO is decided, and then the optimal weight MO is acquired based on the maximum HSI.

Step 4: The optimal MO value is applied to the enhanced amplitude spectrum, and the final improved signal is acquired with the help of the original phase spectrum and the inverse FT.

Step 5: The search principles of the bandwidth and central frequency are decided.

Step 6: The whole frequency band of the signal acquired in step 4 is divided, and the reconstruction signal is obtained using the maximum sparsity values as the target.

Step 7: The envelope spectrum of the reconstruction signal is achieved, where the fault features of REB can be detected to accurately determine the fault type of bearing.

4 Simulation analysis

To verify the effectiveness of the proposed algorithm, a simulation bearing inner race fault signal is constructed in this paper, which consists of the periodic impacts, harmonics and white noise components, and the influence of the

Fig. 1 The flow chart of the approach in this manuscript

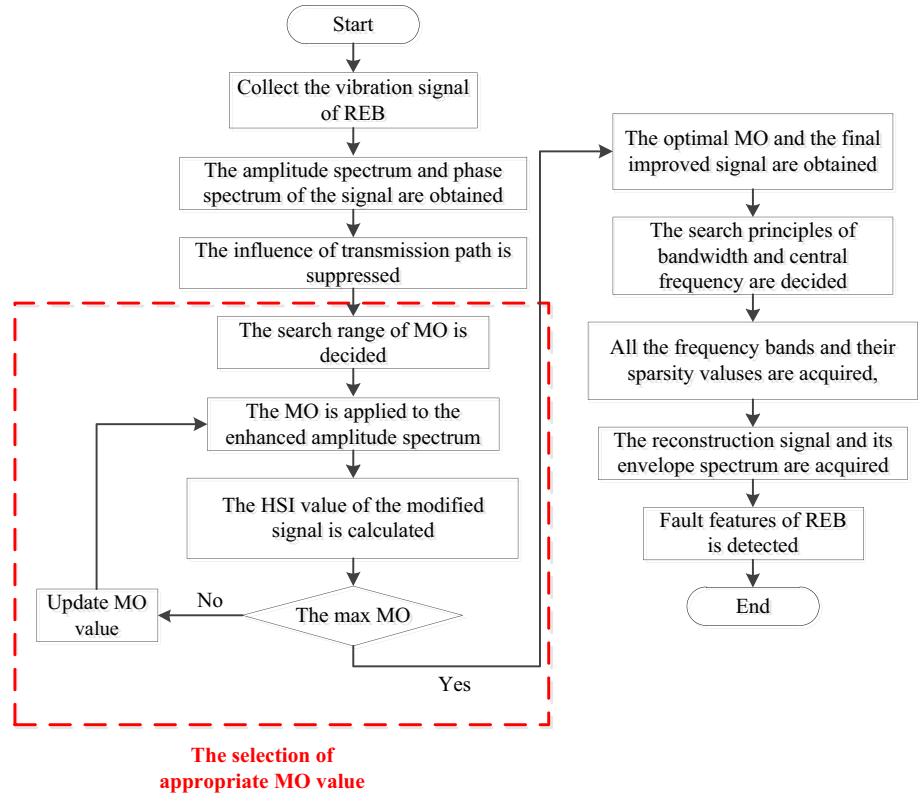


Table 1 Parameters in simulation signal

Letters	Meaning	values
A_i	Amplitude modulation	$1/f_r$
f_r	Rotating frequency	42 Hz
R	Damping coefficient of the system	0.05
f_n	Natural frequency of the system	3200 Hz
T	Interval of two adjacent impulses	1/185 s
t_i	Delay due to the slippage of rolling elements	0.01 $T \sim 0.02 T$
$n(t)$	Random white noise	

transmission path is further taken into consideration. This simulation signal can be expressed as:

$$\begin{cases}
 y(t) = x(t) * h(t) \\
 x(t) = \sum_{i=1}^N A_i s(t - iT - t_i) + B(t) + n(t) \\
 A_i = 4 \cos(2\pi f_r t + \pi/2) + 0.5 \\
 B(t) = \cos(200\pi t) + 0.8 \cos(100\pi t) + 0.5 \sin(220\pi t) \\
 s(t) = e^{-2\pi f_n r t} \sin(2\pi f_n t)
 \end{cases} \quad (6)$$

where * denotes the convolution operation, and the description of other parameters is presented in Table 1. $x(t)$ and $y(t)$ denote the excitation and response signals, respectively,

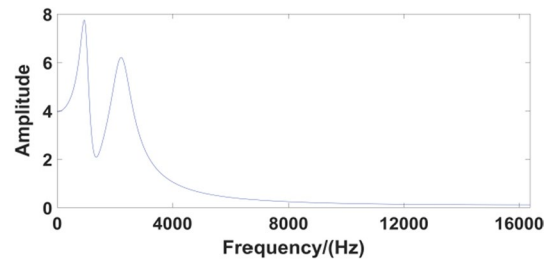


Fig. 2 Amplitude spectrum of the transfer function

and the response signal is used for subsequent verification. $B(t)$ denotes the irrelevant harmonic components. $S(t)$ is employed to simulate the single fault impact signal. $h(t)$ is the unit pulse response of the transfer function, which can be expressed as Eq. (7) and its amplitude spectrum is presented in Fig. 2.

$$H(s) = \frac{(4500\pi)^2}{s^2 + 4655s + (4500\pi)^2} + \frac{(2000\pi)^2}{s^2 + 1985s + (2000\pi)^2} \quad (7)$$

According to the values of all parameters employed in the simulation signal, the FCF of the signal is 185 Hz. The random noise is employed, and the signal-to-noise

ratio of this signal is determined as -15 db to make this simulation signal more complex. Based on the characteristics of the Fourier transform, and to facilitate subsequent analysis, the sampling frequency is determined as 2^N . If the sampling frequency is too large, the signal storage and processing pressure will increase, if the sampling frequency is too small, some high frequency signals will be ignored. Hence, the sampling frequency of all signals used in the simulation is 32768 Hz. The length of the signal for analysis is 1 s. Figure 3a and Fig. 3c display the simulation signal where only 0.2 s of the signal is shown to make the signals' characteristics more clear. Due to the interference of harmonics, noise and the transmission path, the periodic impact components cannot be detected in both excitation and response signals. The amplitude of the response signal in Fig. 3d is generally larger than that of the excitation signal in Fig. 3b by comparison. Significantly distortion also occurred in the response signal since the fluctuation trend is clearly displayed at the bottom of its amplitude spectrum, which is similar to that in Fig. 2. The envelope

spectrum of the response signal is exhibited in Fig. 3e, which fails to extract any useful information related to bearing fault. The ISAM approach is then applied to the simulation signal.

First, to suppress the signal distortion and interference components enhancement due to the influence of the transmission path in advance, the amplitude spectrum shown in Fig. 3d is processed using a trend-line-based transmission path's influence suppression technique. Then the weight MO is applied to the transmission path suppressed amplitude spectrum. To improve the enhancement of the fault features, the optimal MO is selected according to the principles in Sect. 2, with the maximum HSI of the processed signal as the target. The minimum value of the MO to be selected is 0.1, the maximum value is 2, and the MO is increased at the interval of 0.1. Figure 4a displays the relationship between different MO values and the corresponding HSI. Based on this curve, the maximum HSI 0.07 of the improved signal can be obtained when $MO = 1.7$, so this MO value is determined as the optimal one and the improved amplitude

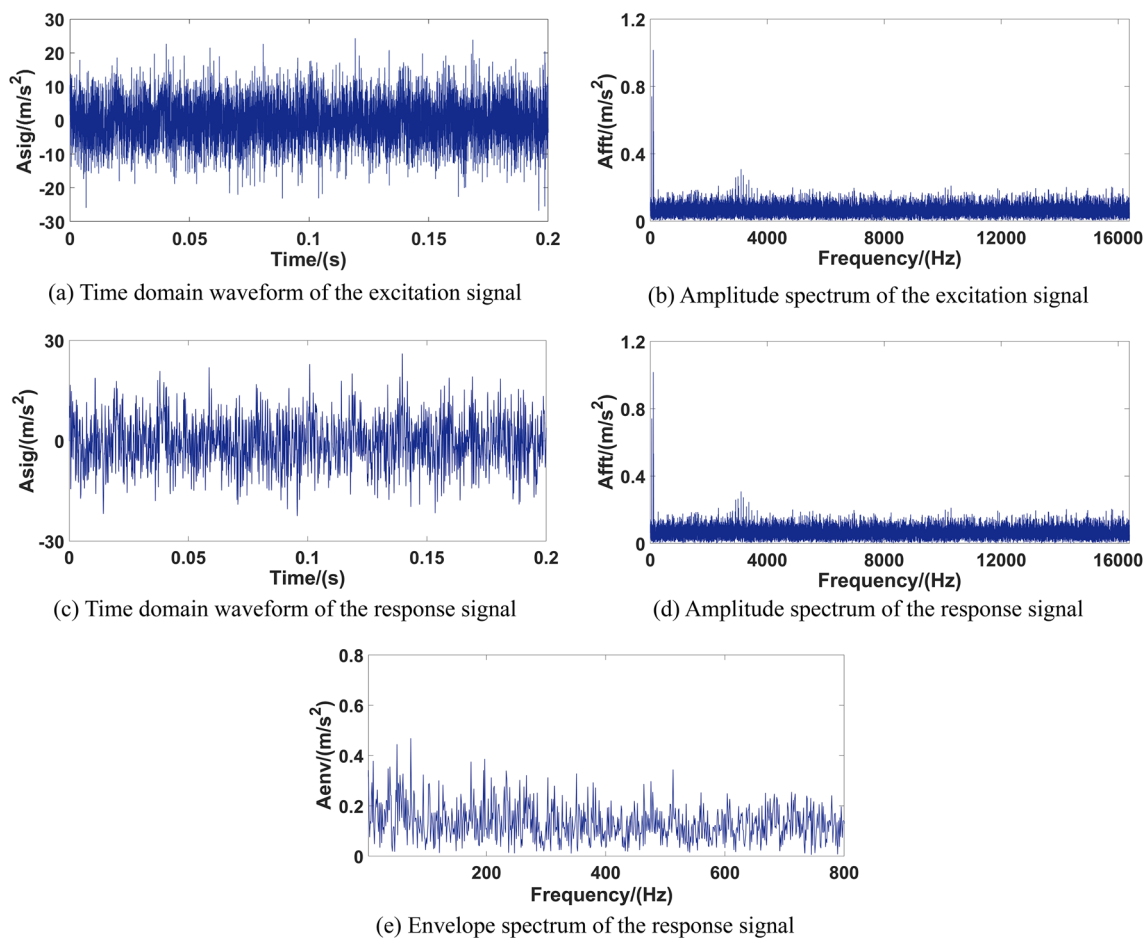


Fig. 3 The simulation signal. **a** Time domain waveform of the excitation signal

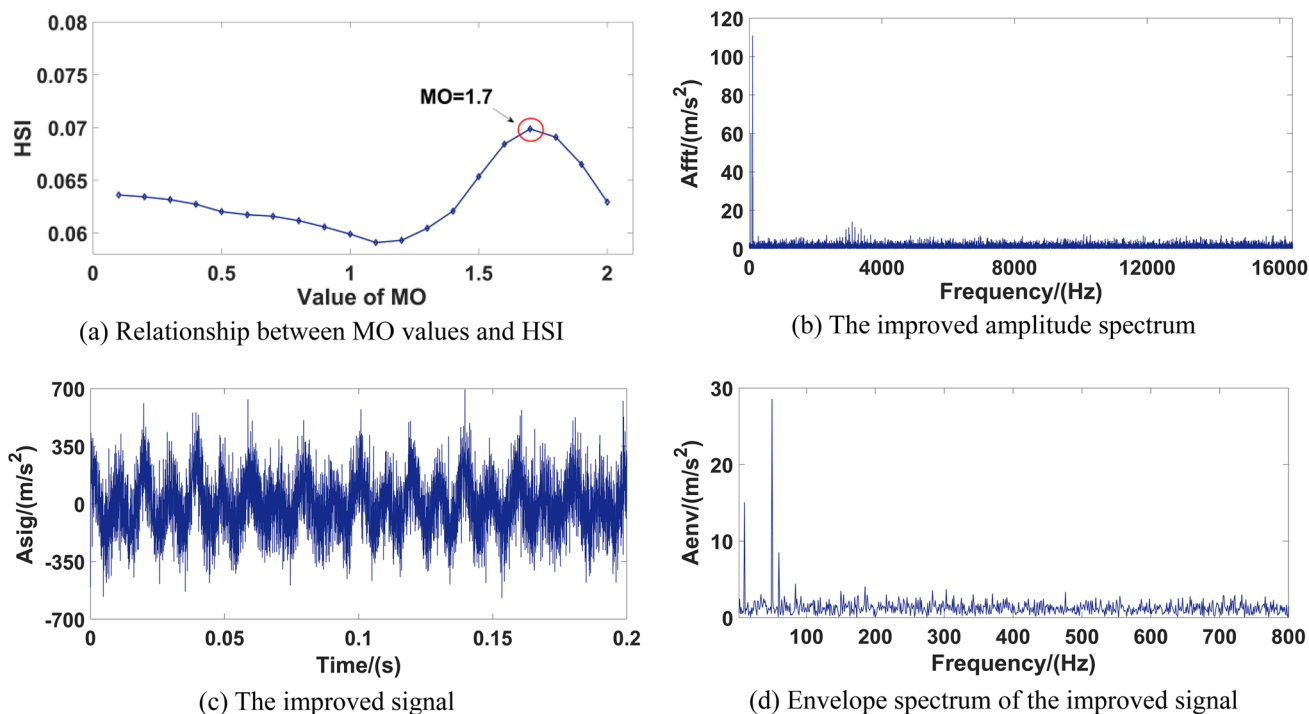


Fig. 4 Results of signal improvements

spectrum is exhibited in Fig. 4b. The fluctuation trend at the bottom of its amplitude spectrum can hardly be detected which indicates that the influence of the transmission path is effectively suppressed. With the help of the phase spectrum of the original signal and the inverse FT, the improved signal shown in Fig. 4c can be obtained, which shows obvious harmonic components. The envelope spectrum of this improved signal in Fig. 4d extracts obvious harmonic features, and no obvious bearing fault characteristics can be detected.

Subsequently, the improved signal is analyzed to determine the optimal resonance frequency band. The bandwidth of each analysis band for selection is determined as $185 \times 4 = 740$ Hz based on the FCF of REB. Figure 5a exhibits the sparsity values of signals in different frequency bands, based on the principle of maximum sparsity, five bands with central frequencies of 3160 Hz, 3253 Hz, 3067 Hz, 3346 Hz and 2974 Hz are selected for subsequent signal reconstruction, their corresponding sparsity values are 1.2793, 1.2787, 1.2786, 1.2774 and 1.2767, respectively, the reconstruction signal is obtained and presented in Fig. 5b. When comparing the original simulation signal and the reconstruction signal, it can be found that through a series of feature enhancement and noise suppression processes, the fault signatures in the signal are significantly enhanced and the impacts can be observed. Figure 5c presents the envelope spectrum of this reconstruction signal, where the rotating frequency components, the bearing inner race FCF 185 Hz and its harmonics

are detected with little interference. The fault features are abundant which shows the validity of the proposed approach, it can conclude that the bearing has inner race defect and the importance of determining an appropriate resonance frequency band is also demonstrated.

To highlight the importance of using SAM to enhance the fault features, the resonance frequency bands are selected directly using the original signal, and results are depicted in Fig. 6. The bandwidth for all the bands is still determined as 740 Hz, and the central frequencies of the five selected bands based on the principle of maximum sparsity are 649 Hz, 556 Hz, 742 Hz, 463 Hz and 835 Hz, with sparsity values are 1.271, 1.27, 1.2699, 1.2693 and 1.2676, respectively. Figure 6a and b shows the reconstruction signal and its envelope spectrum, where no valid fault information is identified. The envelope spectrum shows a complicated feature map with numerous irrelevant frequency components, which indicates that the use of the SAM method to enhance the feature components is of great meaning.

To further show the effectiveness of the technique, the Fast Kurtogram method is employed, and the optimal frequency band signal is selected for analysis, results are displayed in Fig. 7. As shown in Fig. 7(a), the optimal frequency band is determined with the central frequency, and bandwidth are 4096 Hz and 8192 Hz (highlight in red). Figure 7b presents the square envelope spectrum of the optimal band signal where no fault-related signatures are

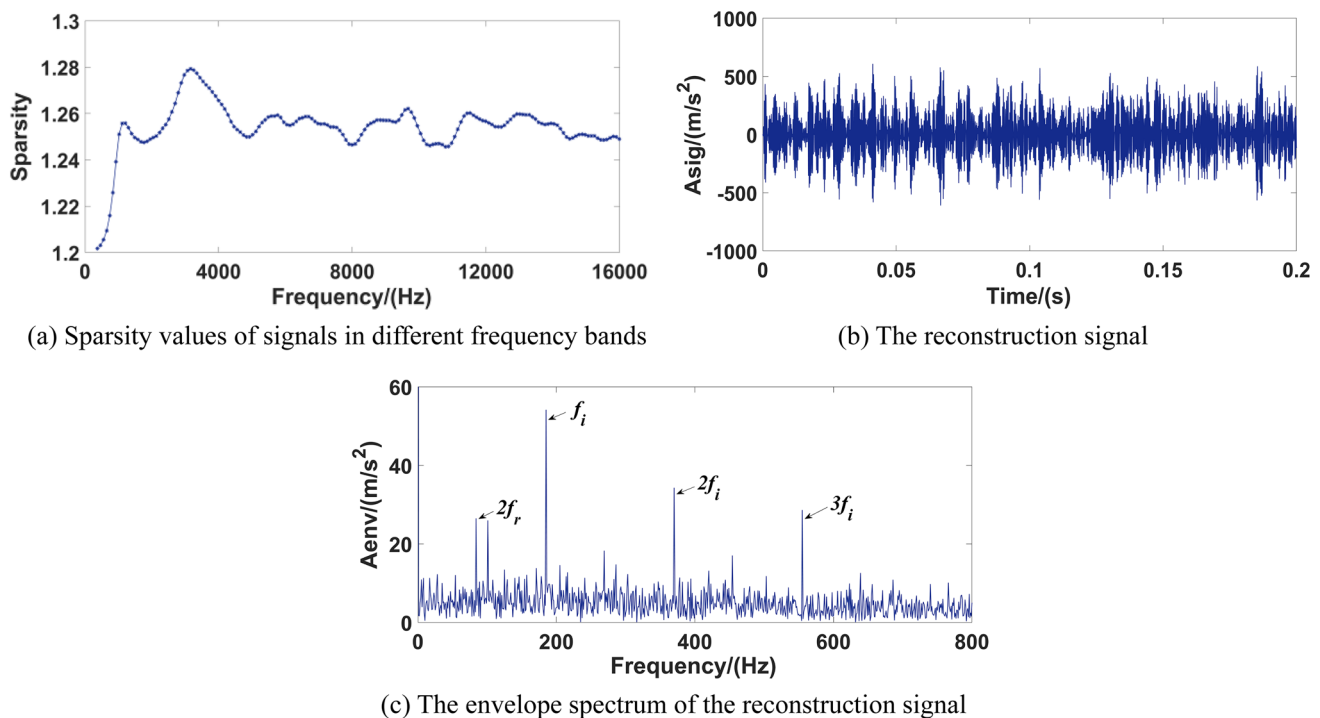


Fig. 5 Analyzed results of the simulation signal using the proposed technique

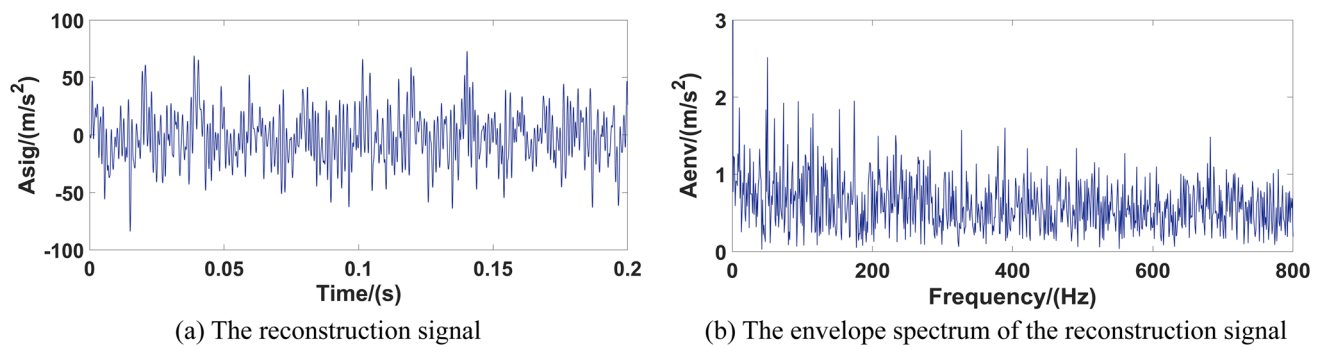


Fig. 6 Analyzed results of the simulation signal by choosing the optimal resonance band directly

successfully extracted, and the fault type of bearing is hard to judge. The phenomenon described previously reveals that the Fast Kurtogram method is hard to extract features from the signal buried by heavy noise.

5 Experimental analysis

Based on the results of the simulation analysis, the measured fault signals of REB in the laboratory are further employed to verify the effectiveness of the algorithm. Figure 8 exhibits the structure of the test rig, which can be used to simulate

the actual operation of different work conditions and fault types of REBs. The measuring point is selected far away from the fault bearing to increase the complexity of the measured signals. The signals generated by the bearing faults will pass through several mechanical interfaces and be affected by multiple structures before being collected by the sensors, which brings more difficulty to the extraction of bearing fault signatures.

During the tests, the type of bearing is selected as 6010, and the parameters of this bearing are shown in Tab.2. To simulate the existence of inner race and outer race fault, the laser wire-electrode cutting is employed and a defect

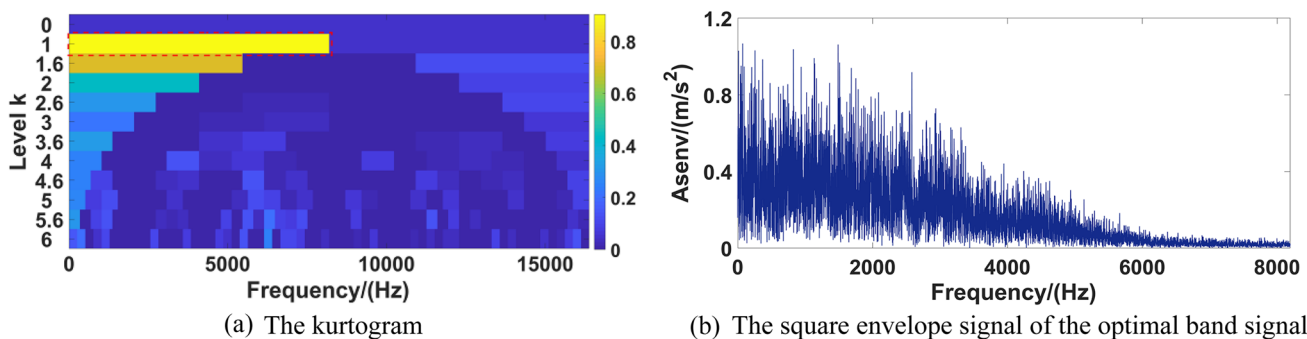


Fig. 7 Analyzed results of the simulation signal using the Fast Kurtogram

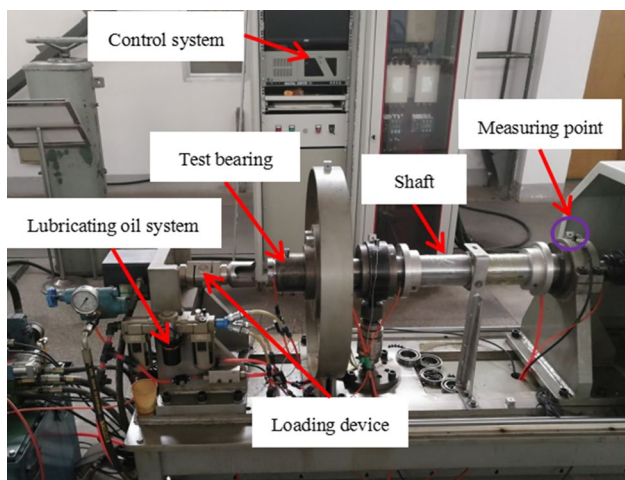


Fig. 8 Structure of the test rig

Table 2 Parameters of the test bearing

Inner diameter/mm	Outer diameter/mm	Pitch diameter D/mm	Ball diameter d/mm	Number of balls Z	Contact angle $\alpha/^\circ$
50	80	65	9	13	0

is generated (the width and depth of the defect are both 0.2 mm). The rotating speed is 3000r/min (rotational frequency $f_r = 50$ Hz), the radial and axial loading are both 1kN, and the sampling frequency is 32768 Hz. The FCF of inner (f_i) and outer (f_o) race defect can be expressed as:

$$\begin{cases} f_i = \frac{1}{2}Z \left(1 + \frac{d}{D} \cos a \right) f_r \\ f_o = \frac{1}{2}Z \left(1 - \frac{d}{D} \cos a \right) f_r \end{cases} \quad (8)$$

By calculation, the theoretical fault characteristic frequency for this REB with inner race and outer race defect are $f_i = 370$ Hz and $f_o = 280$ Hz.

5.1 Analysis of REB with inner race fault

Figure 9 presents the time domain waveform, envelope spectrum and amplitude spectrum of the measured inner race defect signal. The noise interference is obvious, and the fault impact characteristics are not able to be detected in Fig. 9a. In Fig. 9b, only one of the FCF components can be barely identified, the interference of irrelevant components is obvious and the fault signatures are not prominent and rich enough, which does not achieve an ideal result. As a result, the signal is analyzed using the method in this paper.

Similar to the simulation analysis, according to the calculation process in Sect. 2, the influence of the transmission path in the amplitude spectrum displayed in Fig. 9c is first suppressed. Then the weight MO is applied to the transmission path suppressed amplitude spectrum, the HSI values are calculated for the case of applying different MO values, the search range of MO values are still 0.1–2 with the interval of 0.1, and the analyzed results are depicted in Fig. 10a. According to the curve, the maximum HSI value of 0.211 of the improved signal can be acquired when $MO = 1.5$, so this value is determined as the optimal MO value for subsequent analysis. Figure 10b and c shows the amplitude spectrum and time domain waveform of the improved signal. The fluctuation trend at the bottom of Fig. 10b can be hardly detected which shows the effect of the transmission path is eliminated. The improved signal is obtained by combining the improved amplitude spectrum, the phase spectrum of the original measured signal and the inverse FT, which shows obvious harmonic features. Figure 10d shows the envelope spectrum of the improved signal, the amplitude of the spectral lines at the rotating frequency components is prominent,

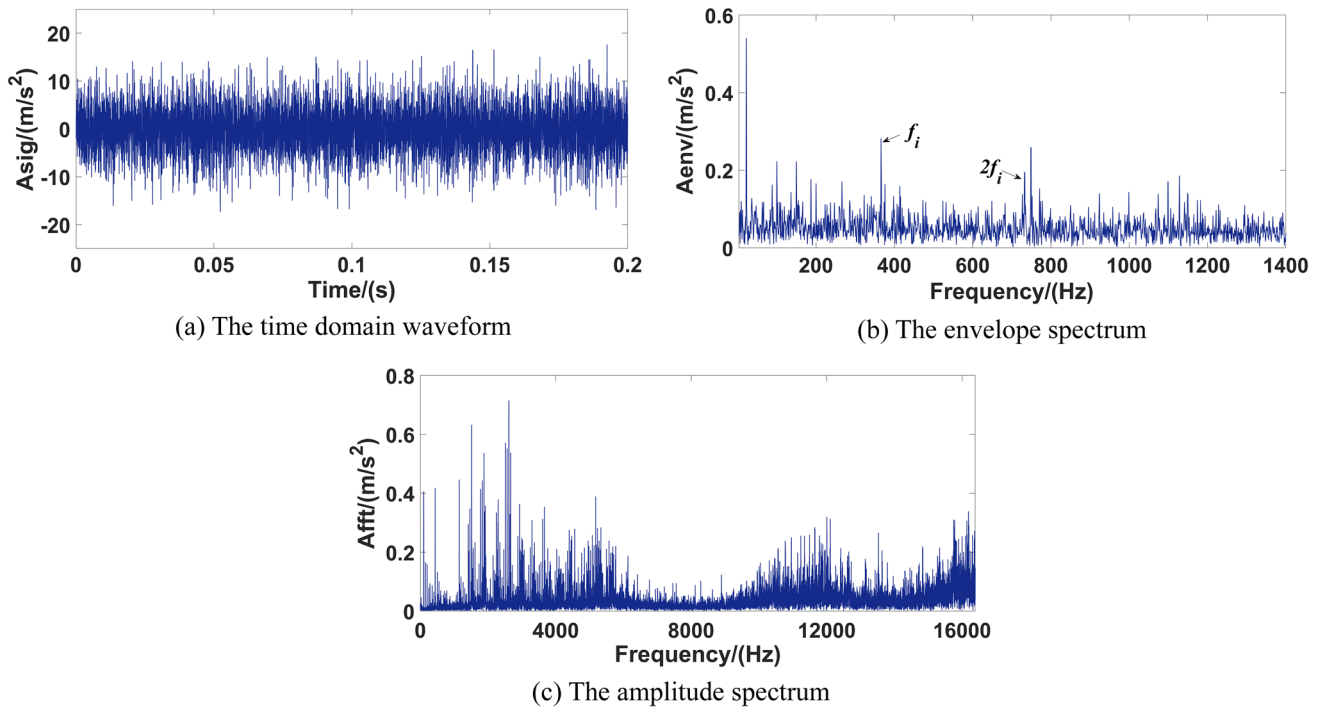


Fig. 9 The experimental bearing inner race fault signal

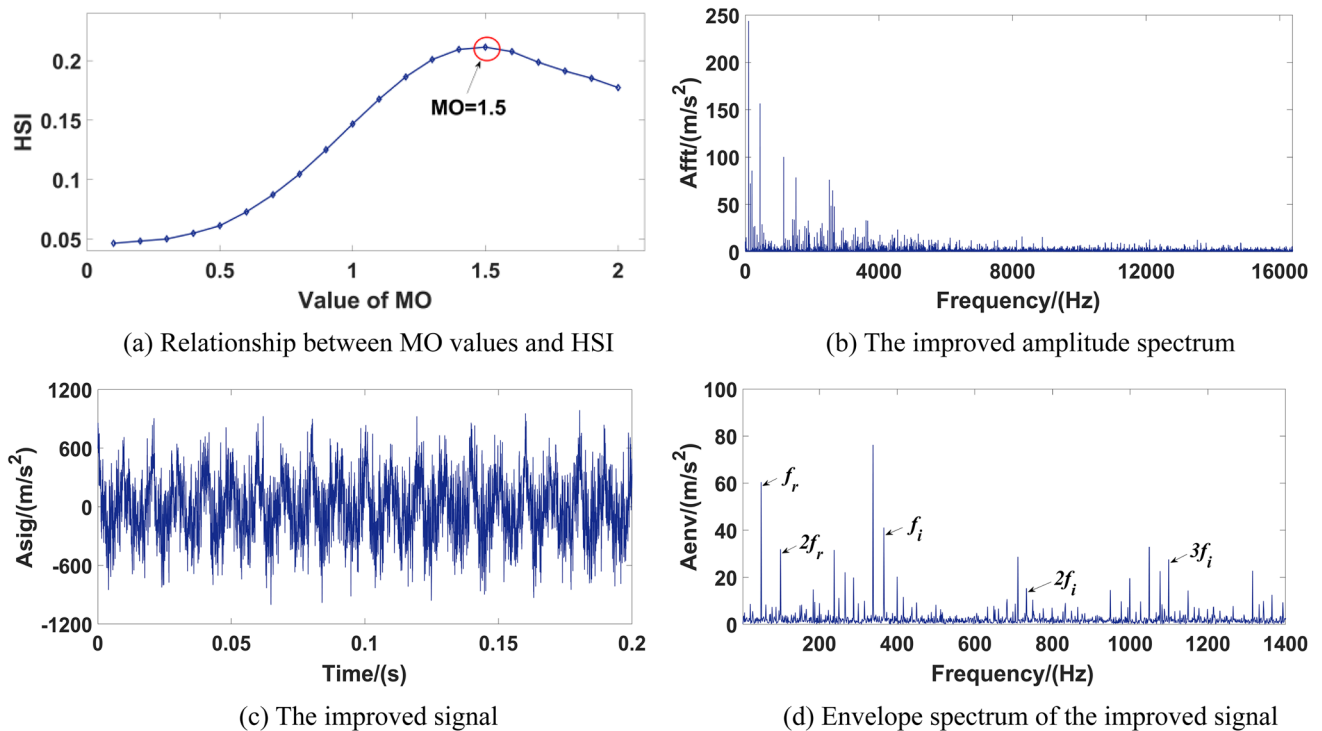


Fig. 10 Improvements of the bearing inner race fault signal

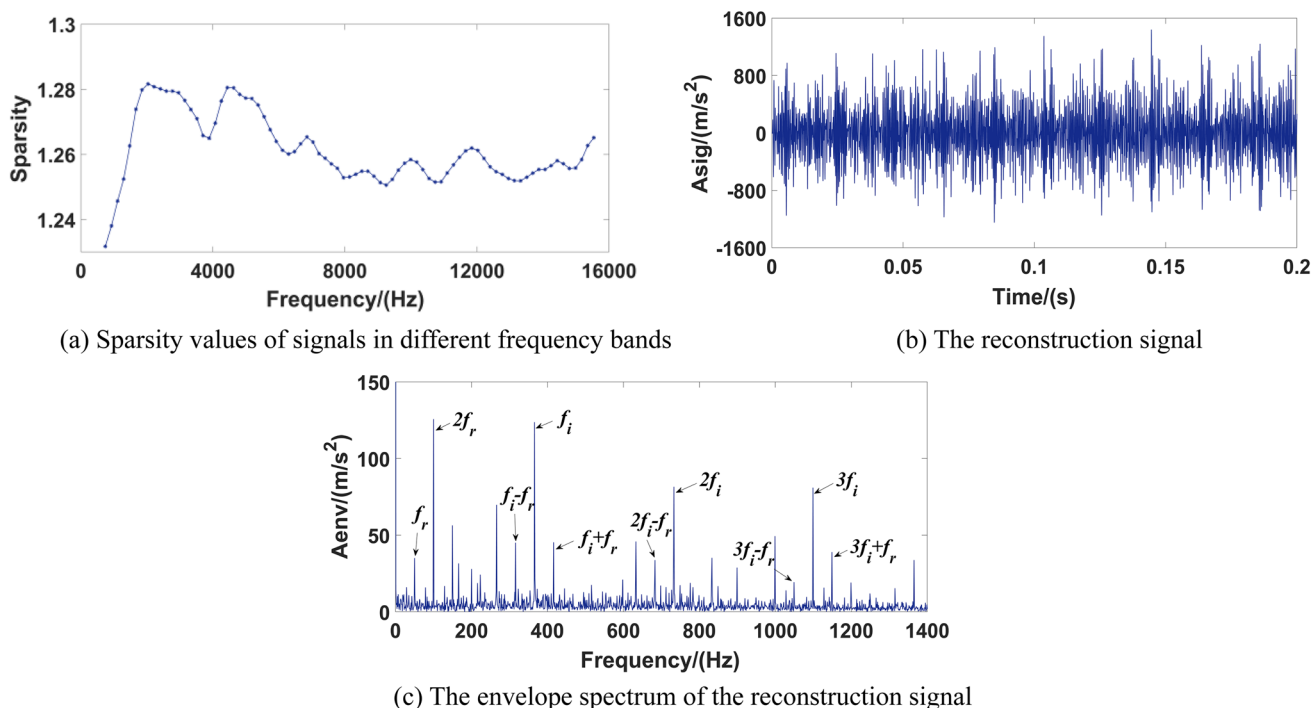


Fig. 11 Analyzed results of the bearing inner race fault signal using the proposed technique

the bearing inner race FCF f_i and its harmonics can be identified but not obvious, the influence of irrelevant frequency components is significant, hence some further analysis is needed to highlight the fault signatures.

Subsequently, the optimal resonance frequency band is determined, and results are depicted in Fig. 11. The bandwidth for the resonance frequency band selection is decided as $370 \times 4 = 1480$ Hz. Figure 11a shows the sparsity values of the signals in different bands, and the five optimal resonant frequency bands are selected based on the maximum sparsity index, whose central frequencies are 2035 Hz, 2220 Hz, 4440 Hz, 4625 Hz and 2405 Hz, respectively, their sparsity values are 1.282, 1.281, 1.2805, 1.2804 and 1.2801. The reconstruction signal is acquired using the five bands signal and shown in Fig. 11b, which exhibits certain components of impacts that may result from the bearing defect. Its envelope spectrum in Fig. 11c shows abundant fault features, where the rotating frequency components, the bearing inner race FCF and its harmonics are all prominent. The side bands around the FCF components are also successfully extracted, and the influence of irrelevant frequency components is limited. The phenomenon described above reveals that the bearing has inner race defect, which also shows that this method can deal with the actual fault feature extraction of REB.

Two comparisons are made to highlight the effectiveness of the method in this paper. First, the optimal resonance

frequency bands are selected directly in the original signal without feature enhancement by the SAM, the bandwidth for band selection is still 1480 Hz, and results are exhibited in Fig. 12. Based on the principle of maximum sparsity value, five optimal frequency bands are selected with the central frequencies are 4440 Hz, 4995 Hz, 2775 Hz, 4625 Hz and 4810 Hz, and the sparsity values of the corresponding band signals are 1.2778, 1.2775, 1.2775, 1.2774 and 1.2769, respectively. The optimal band selection results are different from the results using the proposed technique. The signal reconstruction is conducted using the five band signals, and the result is shown in Fig. 12a. Figure 12b shows the envelope spectrum of the reconstruction signal, though both the rotating frequency components and the FCF components are able to be observed, the fault signatures are not abundant enough when compared with results in this paper, and the interference components are also more obvious.

The analyzed results using the Fast Kurtogram are shown in Fig. 13. According to the kurtogram in Fig. 13a, the central frequency and bandwidth of the optimal analysis frequency band are determined as 15018 Hz and 2730 Hz, respectively. The squared envelope spectrum of this band signal is shown in Fig. 13b, where no useful information related to the bearing fault can be identified. Based on the previous comparisons, the effectiveness of the method in this paper is further highlighted.

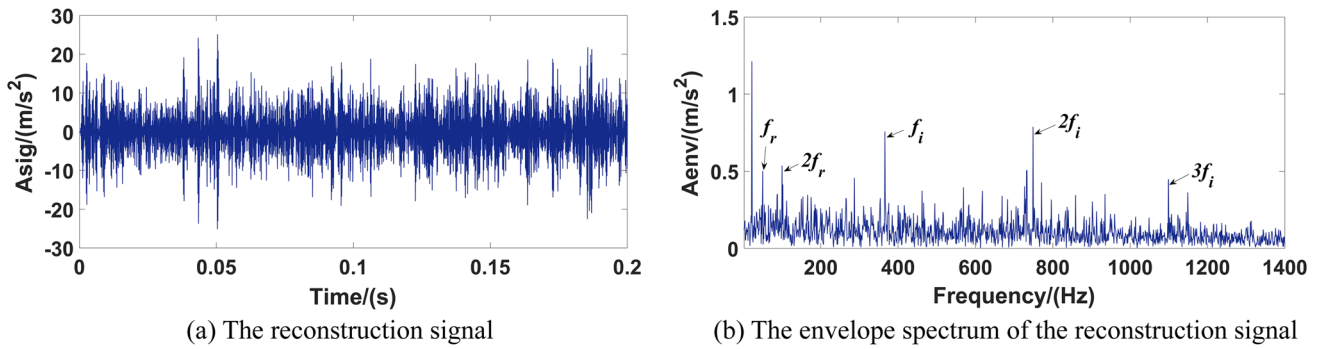


Fig. 12 Analyzed results of the bearing inner race fault signal by choosing the optimal resonance band directly

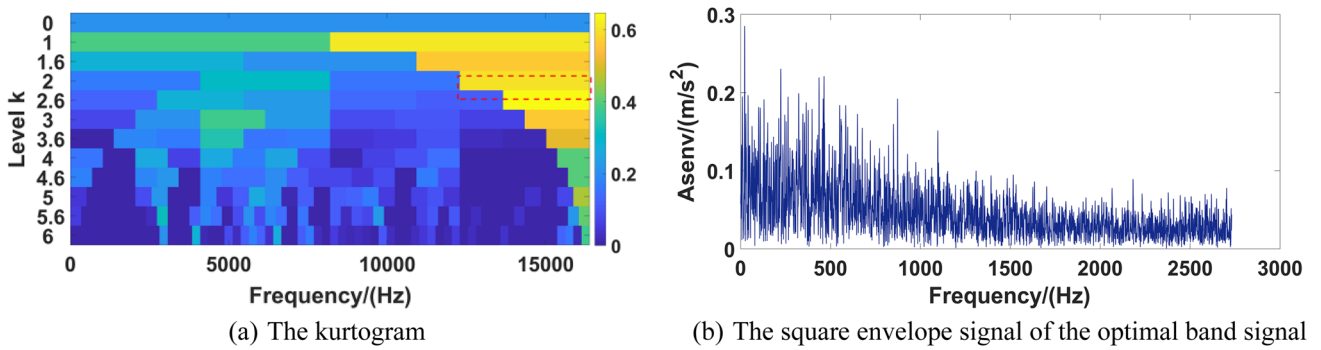


Fig. 13 Analyzed results of the bearing inner race fault signal using the Fast Kurtogram

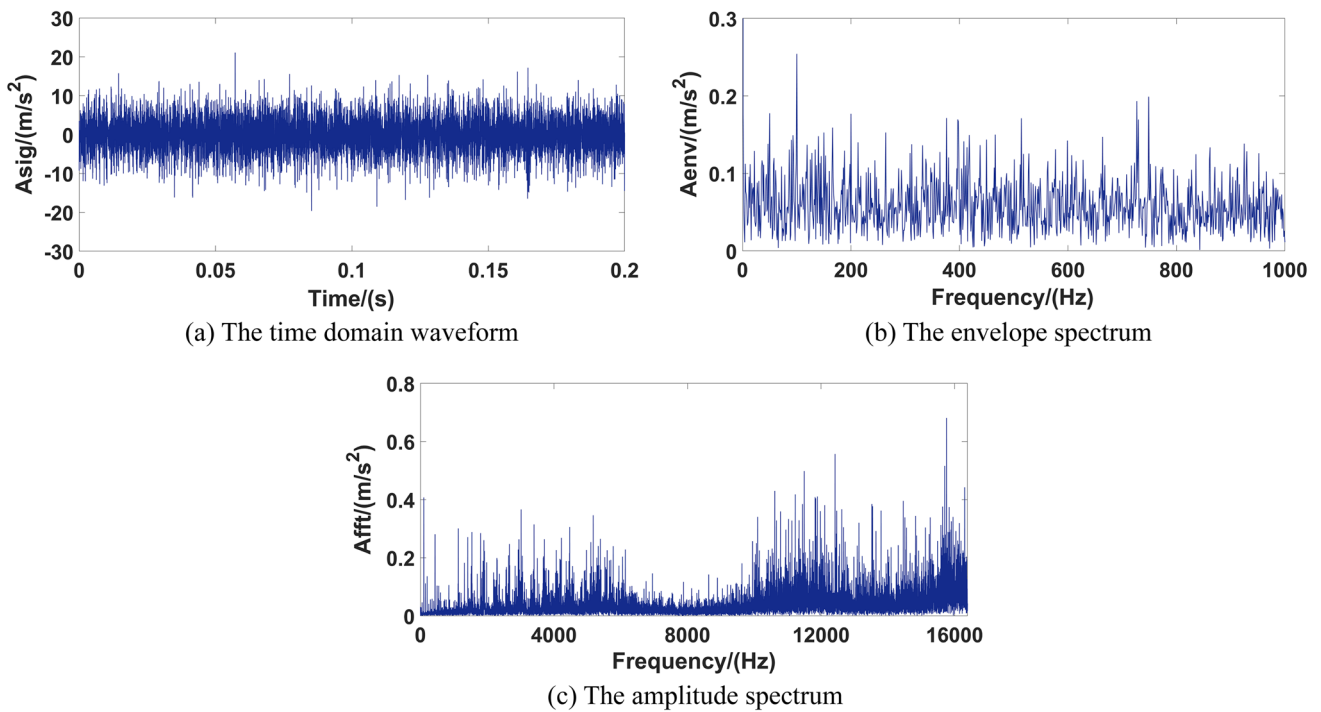


Fig. 14 The experimental bearing outer race fault signal

5.2 Analysis of REB with outer race fault

To demonstrate the wide applicability of the proposed algorithm, the measured bearing outer race fault signal is employed. Figure 14 shows its time domain waveform, envelope spectrum and amplitude spectrum of the signal. Due to the effects of the complex transmission path and the multi-structures of the equipment, the periodic impact components caused by bearing faults are unable to be distinguished. The frequency components in the envelope spectrum and amplitude spectrum are complex while no valid fault information can be observed.

The influence of the transmission path in Fig. 14c is suppressed in advance, and then the weight MO is applied to the transmission path suppressed amplitude spectrum. The optimal MO value is determined based on the curve presented in Fig. 15a. According to the curve, when MO=2, the improved signal has the maximum HSI value, so this MO is determined as the proper weight value for analysis. Figure 15b shows the improved amplitude spectrum where the fluctuation trend is eliminated, and the improved signal shown in Fig. 15c can be obtained with the help of the phase spectrum of the original signal and the inverse FT. The harmonic components are prominent in this improved signal, but no valid bearing fault information can be extracted from its envelope spectrum, and the rotating frequency components may result from the unbalance of the system.

Subsequently, the optimal resonance frequency band of the improved signal is selected, and results are depicted in Fig. 16. The analysis bandwidth is determined to be $280 \times 4 = 1120$ Hz, and the sparsity values for each band's signals are shown in Fig. 16a. Five optimal frequency bands are chosen with the maximum sparsity value as the target, which are 1.276, 1.275, 1.27, 1.269 and 1.268, respectively, their central frequencies are 1820 Hz, 1960 Hz, 1680 Hz, 2100 Hz and 4200 Hz. The above five band signals are employed for signal reconstruction, which is exhibited in Fig. 16b. Figure 16c shows the envelope spectrum of this reconstruction signal, where the outer race FCF and its harmonics can be identified, the influence of irrelevant noise components is able to be detected but are significantly suppressed when compared with those in the envelope spectrum of the original signal. The experimental bearing can be judged to have an outer race fault by the features introduced above.

Analogously, two comparison analysis is carried out to further demonstrate the validity of the proposed technique. The results of selecting the resonance frequency bands directly in the original experimental outer race fault signal are shown in Fig. 17. The bandwidth of each band is still fixed at 1120 Hz. Five optimal resonance frequency bands are determined, whose central frequencies are 7840 Hz, 7420 Hz, 7560 Hz, 7700 Hz and 8680 Hz, their corresponding sparsity values are 1.2743, 1.2741, 1.2738,

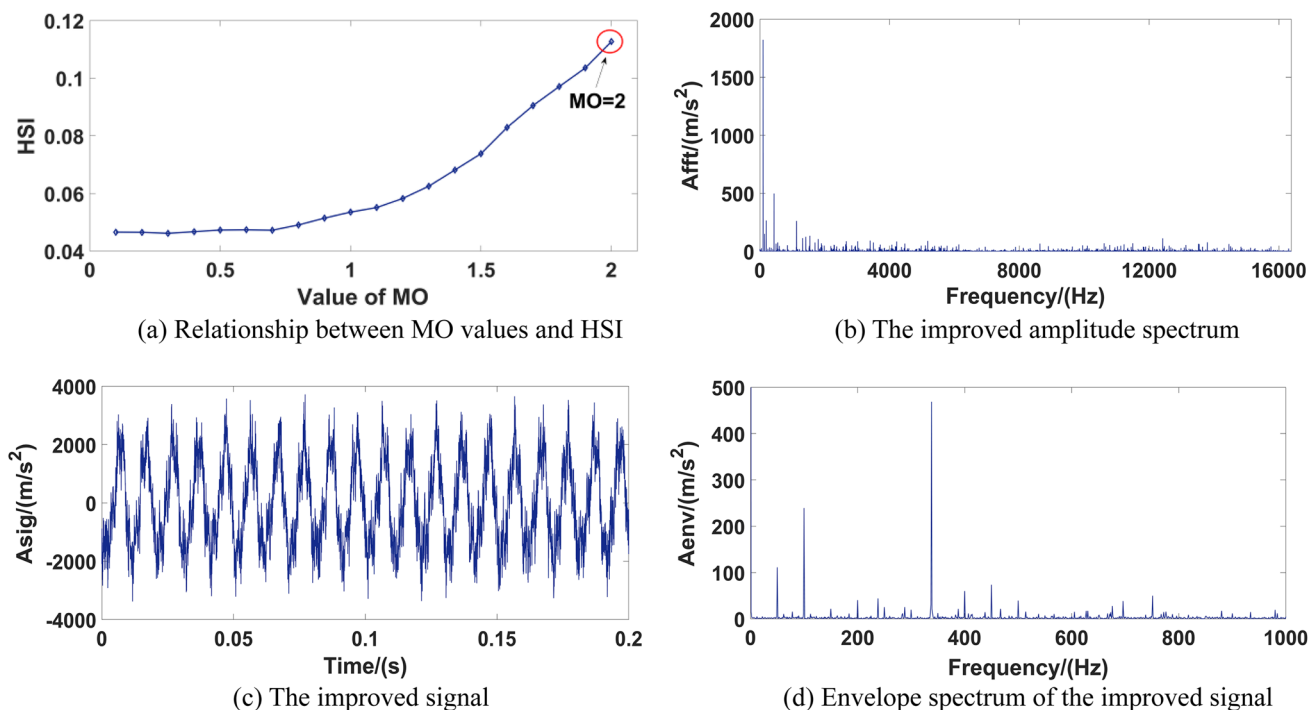


Fig. 15 The improvements of the bearing outer race fault signal

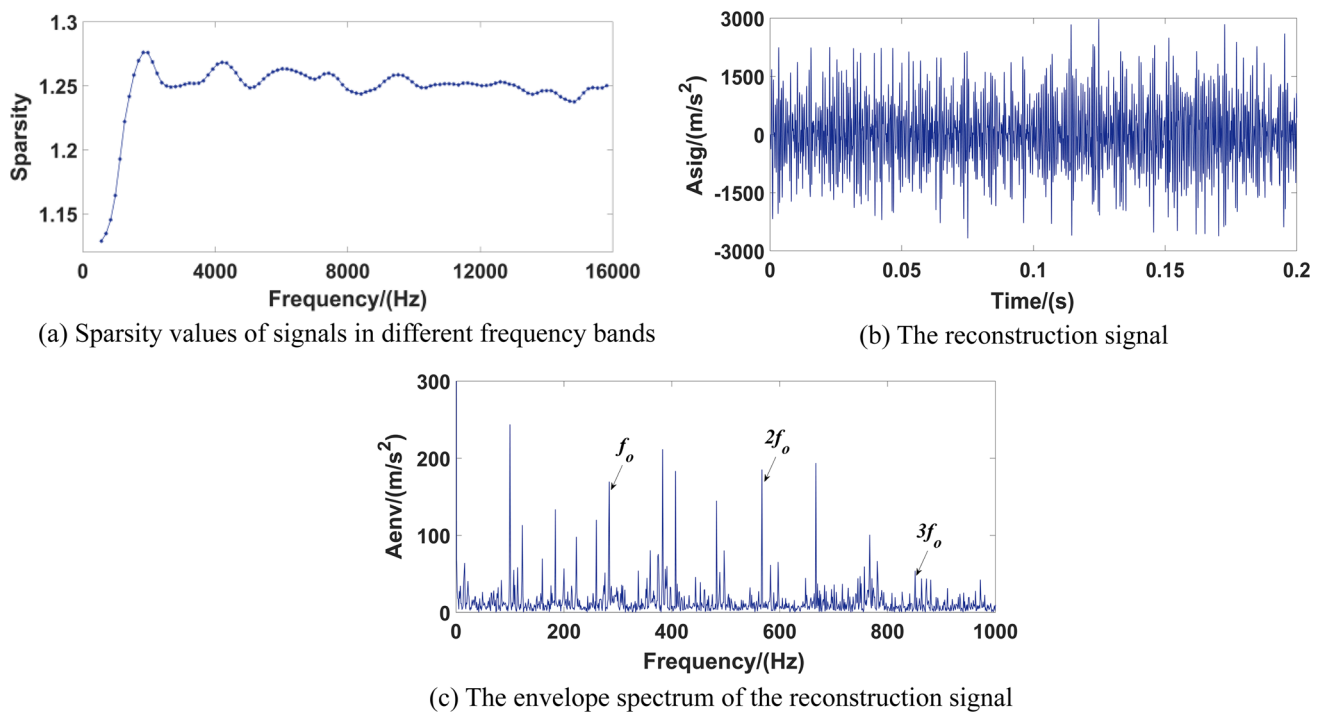


Fig. 16 Analyzed results of the bearing outer race fault signal using the proposed technique

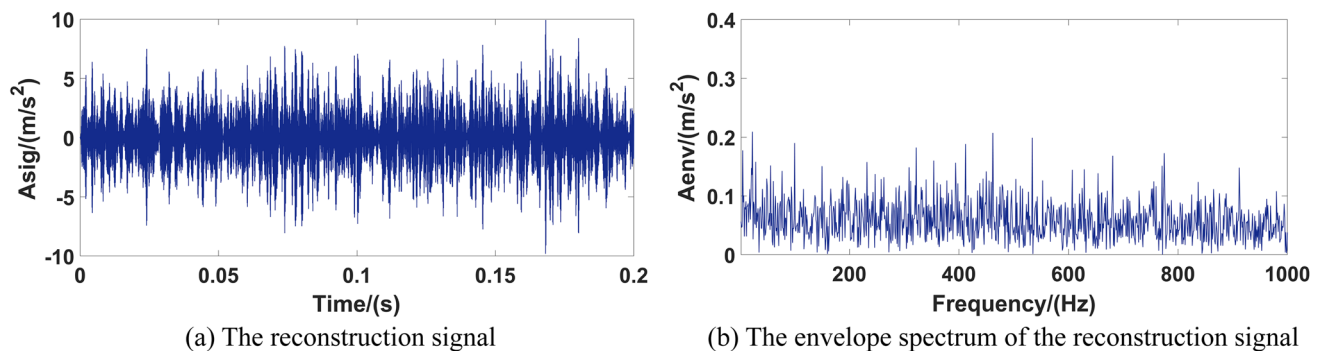


Fig. 17 Analyzed results of the bearing outer race fault signal by choosing the optimal resonance band directly

1.2737 and 1.2736. These five band signals are used to construct the reconstruction signal presented in Fig. 17a, and its envelope spectrum is presented in Fig. 17b. This feature map fails to extract any fault feature information, and the interference components are obvious.

Then the measured bearing outer race defect signal is analyzed using the Fast Kurtogram method. According to the results in Fig. 18, the central frequency and bandwidth

of the optimal analysis band are 8533 Hz and 682 Hz respectively, but no fault information is identified from the squared envelope spectrum of the optimal band signal, the strong background interference components result in large impacts. Based on the previous analysis, the effectiveness of the method in this paper is further verified.

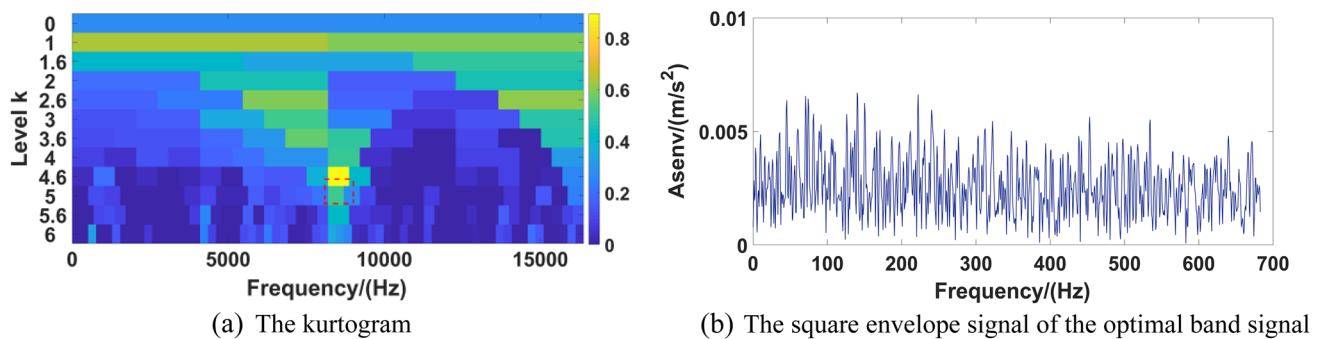


Fig. 18 Analyzed results of the bearing outer race fault signal using the Fast Kurtogram

6 Conclusion

In this paper, an ISAM-based fault diagnosis method for REB is proposed to detect the bearing faults seriously affected by strong background interference, and both the feature components enhancement and the interference components suppression are taken into consideration, which has certain practical engineering value. Through the research some conclusions can be drawn as follows.

- (1) The signal distortion and interference enhancement are suppressed first, and then the optimal MO selection based on the maximum HSI index can ensure the effect of fault feature enhancement and reduce the blindness of the analysis.
- (2) The optimal resonance frequency band selection and signal reconstruction is conducted with the maximum sparsity value as the target, which can further reduce the interference of irrelevant components, and highlight the fault signatures of REB.
- (3) Simulation bearing inner race defect signal, experimental bearing inner and outer race defect signals are all employed to verify the effectiveness of the method. Results show that this method can effectively suppress the interference of strong background noise and extract the abundant bearing fault characteristics. A series of comparisons further highlight the validity of this paper's method in weak fault feature extraction.

References

1. Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. *Mech Syst Signal Process* 21(1):108–124
2. Wang XL, Zheng JD, Ni Q, Pan HY, Zhang J (2022) Traversal index enhanced-gram (TIEgram): a novel optimal demodulation frequency band selection method for rolling bearing fault diagnosis under non-stationary operating conditions. *Mech Syst Signal Process* 172:109017
3. Zhang K, Chen P, Yang MR, Song LY, Xu YG (2022) The Harmonogram: a periodic impulses detection method and its application in bearing fault diagnosis. *Mech Syst Signal Process* 165:108374
4. Ni Q, Ji JC, Feng K, Halkon B (2021) A novel correntropy-based band selection method for the fault diagnosis of bearings under fault-irrelevant impulsive and cyclostationary interferences. *Mech Syst Signal Process* 153:107498
5. Xu YG, Deng YJ, Ma CY, Zhang K (2021) The Enfigram: a robust method for extracting repetitive transients in rolling bearing fault diagnosis. *Mech Syst Signal Process* 158:107779
6. Zhang YX, Huang BY, Xin Q, Chen H (2022) Ewtfergram and its application in fault diagnosis of rolling bearings. *Measurement* 190:110695
7. Liu ZC, Yang SP, Liu YQ, Lin JH, Gu XH (2021) Adaptive correlated Kurtogram and its applications in wheelset-bearing system fault diagnosis. *Mech Syst Signal Process* 154:107511
8. Wang L, Liu ZW, Cao HR, Zhang X (2020) Subband averaging kurtogram with dual-tree complex wavelet packet transform for rotating machinery fault diagnosis. *Mech Syst Signal Process* 142:106755
9. Gu J, Peng YX (2021) An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis. *Digit Signal Process* 113:103050
10. Yin C, Wang YL, Ma GC, Wang Y, Sun YX, He Y (2022) Weak fault feature extraction of rolling bearings based on improved ensemble noise-reconstructed EMD and adaptive threshold denoising. *Mech Syst Signal Process* 171:108834
11. Han MH, Wu YM, Wang YM (2021) Liu W (2021) Roller bearing fault diagnosis based on LMD and multi-scale symbolic dynamic information entropy. *J Mech Sci Technol* 35(5):1993–2005
12. Pang B, Nazari M, Tang GJ (2022) Recursive variational mode extraction and its application in rolling bearing fault diagnosis. *Mech Syst Signal Process* 165:108321
13. Li HD, Xu Y, An D, Zhang LX, Li SH, Shi HT (2020) Application of a flat variational modal decomposition algorithm in fault diagnosis of rolling bearings. *J Low Freq Noise Vib Act Control* 39(2):335–351
14. Deng LF, Zhang AH, Zhao RZ (2022) Intelligent identification of incipient rolling bearing faults based on VMD and PCA-SVM. *Adv Mech Eng* 14(1):16878140211072990
15. Zhu DC, Chen JH, Yin BL (2021) Fault feature extraction of rolling element bearing based on TPE-EVMD. *Measurement* 183:109880
16. Li H, Liu T, Wu X, Chen Q (2020) A bearing fault diagnosis method based on enhanced singular value decomposition. *IEEE Trans Ind Inform* 17(5):3220–3230

17. Dong SJ, Li Y, Zhu P, Pei XW, Pan XJ, Xu XY, Liu LH, Xing B, Hu XL (2022) Rolling bearing performance degradation assessment based on singular value decomposition-sliding window linear regression and improved deep learning network in noisy environment. *Meas Sci Technol* 33(4):045015
18. Liu QQ, Yang JT, Zhang K (2022) An improved empirical wavelet transform and sensitive components selecting method for bearing fault. *Measurement* 187:110348
19. Chegini SN, Bagheri A, Najafi F (2019) Application of a new EWT-based denoising technique in bearing fault diagnosis. *Measurement* 144:275–297
20. Mo ZL, Zhang H, Shen Y, Wang JY, Fu HY, Miao Q (2022) Conditional empirical wavelet transform with modified ratio of cyclic content for bearing fault diagnosis. *ISA Trans* 133:597–611
21. Duan RK, Liao YH, Yang L, Xue JT, Tang MJ (2021) Minimum entropy morphological deconvolution and its application in bearing fault diagnosis. *Measurement* 182:109649
22. He ZY, Chen G, Hao TF, Liu XY, Teng CY (2021) An optimal filter length selection method for MED based on autocorrelation energy and genetic algorithms. *ISA Trans* 109:269–287
23. Deng W, Li ZX, Li XY, Chen HY, Zhao HM (2022) Compound fault diagnosis using optimized MCKD and sparse representation for rolling bearings. *IEEE Trans Instrum Meas* 71:1–9
24. McDonald GL, Zhao Q, Zuo MJ (2012) Maximum correlated Kurtosis deconvolution and application on gear tooth chip fault detection. *Mech Syst Signal Process* 33:237–255
25. Zhang BY, Miao YH, Lin J, Yin YG (2021) Adaptive maximum second-order cyclostationarity blind deconvolution and its application for locomotive bearing fault diagnosis. *Mech Syst Signal Process* 158:107736
26. Wang ZJ, Zhou J, Du WH, Lei YG, Wang JY (2022) Bearing fault diagnosis method based on adaptive maximum cyclostationarity blind deconvolution. *Mech Syst Signal Process* 162:108018
27. Moshrefzadeh A, Fasana A, Antoni J (2019) The spectral amplitude modulation: a nonlinear filtering process for diagnosis of rolling element bearings. *Mech Syst Signal Process* 132:253–276
28. Jiang ZH, Zhang K, Xiang L, Xu YG (2022) Differential spectral amplitude modulation and its applications in rolling bearing fault diagnosis. *Measurement* 201:111755
29. Jiang ZH, Zhang K, Xiang L, Gang Y, Xu YG (2023) A time-frequency spectral amplitude modulation method and its applications in rolling bearing fault diagnosis. *Mech Syst Signal Process* 185:109832

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.