



Sparsity-guided multi-scale empirical wavelet transform and its application in fault diagnosis of rolling bearings

Kun Zhang¹ · Weikang Tian² · Peng Chen¹ · Chaoyong Ma² · Yonggang Xu^{2,3}

Received: 4 February 2021 / Accepted: 19 July 2021 / Published online: 26 July 2021
© The Brazilian Society of Mechanical Sciences and Engineering 2021

Abstract

Empirical wavelet transform (EWT) is a signal decomposition method that distinguishes signals from the frequency domain. When processing non-stationary and strong noise signals, a large number of invalid components may be obtained, or modal aliasing may occur. The biggest contribution of sparsity-guided multi-scale empirical wavelet transform (SMSEWT) is that it can optimize the segmentation method and extract useful frequency band, reduce the number of invalid components, and suppress modal aliasing. In order to divide frequency bands containing similar information into final components, Fourier spectrum will be divided equally and used to calculate kurtosis. Frequency bands with similar kurtosis are considered to contain the same kind of information, which will be combined to achieve adaptive segmentation of the spectrum. Subsequently, empirical wavelet filters will be constructed and the time-domain waveforms of each frequency band can be obtained. Using sparsity to select envelope components containing abundant periodic pulses can diagnose bearing faults.

Keywords Empirical wavelet transform · Sparsity · Multi-scale · Rolling bearing · Fault diagnosis

1 Introduction

Rotating machinery is an indispensable part of industrial society. Most major accidents in rotating machinery occur due to bearing failure. Bearings, considered as one of the most stressed components in rotating machinery, have a significant pulse on the performance and productivity of rotating machinery [1, 2]. Repeated collisions and periodic pulses occur when the bearing fails due to the position deformation of the fault [3]. Early bearing defects first excite resonance frequencies in high-frequency domain [4]. Traditional fault diagnosis techniques mainly include demodulation and

envelope-based methods, because they can identify the frequency caused by defects in bearings. However, the main challenge in applying this technology is to correctly select the center frequency and bandwidth to perform the envelope spectrum.

Spectral kurtosis (SK) is widely used as an important tool for effectively selecting the resonant frequency band in the fault diagnosis of bearing vibration signals. Antoni proposed Fast Kurtogram (FK) [5], using 1/3 and two binary forked tree structures, based on much systematic research on spectral kurtosis [6, 7]. Later, Randall and Sawalhi [8] combine the minimum convolution entropy with FK to amplify the pulse signal, making it easier to detect periodic components in the vibration signal. Combet and Gelman [9] improved FK through time-domain synchronous mean for early fault diagnosis of gears. Lei [10, 11] improved FK for early bearing fault diagnosis which has achieved good results. However, FK itself has certain limitations. For example, the segmentation of spectrum is not adaptive for finding the most suitable bandwidth and center frequency. It is worth mentioning that FK does not have an immune effect on accidental pulse. Hence, Protrugram was proposed by Barszcz [12] to find the optimal center frequency of resonance band. Wang [13] proposed adaptive SK and applied it to the mechanical fault diagnosis in [14,

Technical Editor: Monica Carvalho.

✉ Yonggang Xu
xyg_1975@163.com

¹ Graduate School of Environmental Science and Technology, Mie University, Mie 514-0001, Japan

² Beijing Engineering Research Center of Precision Measurement Technology and Instruments, Faculty of Materials and Manufacturing, Beijing University of Technology, Beijing 100124, China

³ Key Laboratory of Advanced Manufacturing Technology, Faculty of Materials and Manufacturing, Beijing University of Technology, Beijing 100124, China

[15], which can adaptively estimate the bandwidth. Xu and Tian [16] proposed empirical scanning spectrum kurtosis for the selection of optimal frequency band. Sparsogram [17] is proposed which can avoid the effects of interference. In order to processing signal with low SNR, Autogram [18] based on kurtosis of square envelope unbiased autocorrelation was proposed. Tse [19] improved the Sparsogram by combining genetic algorithms to adaptively determine the optimal center frequency and bandwidth. Unfortunately, it sacrifices the calculation time due to the genetic algorithm. Time–frequency analysis has a very important position in signal decomposition and feature extraction [20]. Similarly, methods such as synchrosqueezed wavelet transform (SWT) and synchrosqueezed Fourier transform (SFT) are widely used for fault detection and classification in bearings [21, 22]. Empirical mode decomposition (EMD) proposed by Huang [23] can adaptively process the non-stationary signal. Inspired by EMD, Gilles developed a new method named empirical wavelet transform (EWT) [24]. The method divides the signal into a plurality of IMFs by a local maximum rule in the Fourier spectrum. The main idea of EWT is to divide the spectrum of the signal by designing multiple filters with indefinite bandwidth for mode decomposition. Kedadouche [25] compared the advantages and disadvantages between EMD and EWT and provided a more reliable basis for the follow-up research of EWT. Then, EWT has been improved by experts and scholars in the direction of fault diagnosis and applied in rotor fault diagnosis [26, 27]. The adaptive parameterless EWT method is proposed in [28] combining quadrature derivative-based normalized Hilbert transform for adaptive mode separation. Wang [29, 30] mathematically formulate EWT as a constrained optimization problem for bearing fault diagnosis. Chen [31] extracts modulated signals by EWT processing orthogonal basis decomposition components and uses wavelet spatial neighboring coefficient denoising with data-driven threshold to improve SNR. Jiang and Zhu [32] presented a decoupling diagnosis method for compound faults of rolling bearings based on empirical wavelet transform-duffing oscillator. The components with mixed modes obtained by EWT are directly observed as chaotic motions through the processing of duffing oscillator, so single fault components are identified from empirical modes one by one. Hu [33, 34] proposed an enhanced EWT (EEWT) to improve the spectrum segmentation by order statistical filter (OSF), which is able to select the optimal resonance frequency band of modulated signal. OSF was also used by Zhang [35] and Xu [36] to construct a new boundary distribution diagram similar to FK. Huang [37] proposed the frequency phase space empirical wavelet transform to divide the Fourier spectrum with the Teager energy distribution as a reference. Zhao [38] uses maximum–minimum length curve method to transform the original vibration signal spectrum to scale space in order

to find meaningful clusters. The new method is named MSCEWT. Hsueh [39] et al. proposed a model based on deep CNN and EWT to detect and induce the motor operating state [40]. Ou [41] combines improved particle swarm algorithm, majorization-minimization-based total variation denoising algorithm and EWT to decompose bearing fault signals and extract fault features. Liu and Chen [42] present a general overview of the recent advancements made in research on the EWT algorithm and its state-of-the-art applications in a wide range of areas and providing some comprehensive references for reader concerning with EWT.

In this paper, a novel method named sparsity-guided multi-scale empirical wavelet transform method (SMSEWT) is proposed which can optimize the segmentation method and extract useful frequency band, reduce the number of invalid components, and suppress modal aliasing. In order to divide frequency bands containing similar information into final components, Fourier spectrum will be divided equally and used to calculate kurtosis. Frequency bands with similar kurtosis are considered to contain the same kind of information, which will be combined to achieve adaptive segmentation of the spectrum. Subsequently, empirical wavelet filters will be constructed and the time-domain waveforms of each frequency band can be obtained. Using sparsity to select envelope components containing abundant periodic pulses can diagnose bearing faults. The rest of this paper is structured as follows. Section 2 simply introduced the basic theory of EWT. Section 3 proposed multi-scale EWT that optimizes the spectral segmentation for further construction of EWT filters. In Sect. 4, SMSEWT is proposed for extraction of useful envelope components. Experimental validation and comparisons are performed in Sect. 5.

2 Empirical wavelet transform

2.1 Basic theory of EWT

In empirical wavelet transform, the frequency domain $[0, \pi]$ is adaptively divided into N continuous parts and expressed by Λ_n . According to the characteristics of Fourier spectrum, it can establish a filter bank to extract different components, as shown in Fig. 1. ω_n is defined as the boundary of continuous parts:

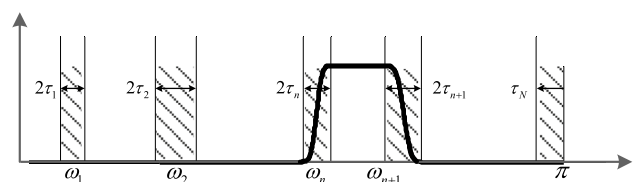


Fig. 1 Empirical wavelet transform filtering method

$$\begin{cases} \Lambda_n = [\omega_{n-1}, \omega_n] \\ \bigcup_{n=1}^N \Lambda_n = [0, \pi] \end{cases} \quad (1)$$

EWT consists of an empirical scaling function $\hat{\varphi}_1(\omega)$ and several empirical wavelet functions $\hat{\Psi}_n(\omega)$. The empirical scaling function and empirical wavelets function are expressed by:

$$\hat{\varphi}_n(\omega) = \begin{cases} 1; & |\omega| \leq (1 - \gamma)\omega_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1 - \gamma)\omega_n)\right)\right]; & (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n \\ 0; & \text{others} \end{cases} \quad (2)$$

$$\hat{\Psi}_n(\omega) = \begin{cases} 1; & (1 + \gamma)\omega_n \leq |\omega| \leq (1 - \gamma)\omega_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{(|\omega| - (1 - \gamma)\omega_{n+1})}{2\gamma\omega_{n+1}}\right)\right]; & (1 - \gamma)\omega_{n+1} \leq |\omega| \leq (1 + \gamma)\omega_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{(|\omega| - (1 - \gamma)\omega_{n+1})}{2\gamma\omega_n}\right)\right]; & (1 - \gamma)\omega_n \leq |\omega| \leq (1 + \gamma)\omega_n \\ 0; & \text{others} \end{cases} \quad (3)$$

where the transition function $\beta(x)$, the coefficient γ , and the transition phase τ_n are:

$$\begin{cases} \beta(x) = x^4(35 - 84x + 70x^2 - 20x^3) \\ \gamma < \min\left(\frac{\omega_{n+1} - \omega_n}{\omega_{n+1} + \omega_n}\right) \\ \tau_n = \gamma\omega_n, 0 < \gamma < 1 \end{cases} \quad (4)$$

Set the Fourier transform as $F(\bullet)$, the inverse Fourier transform is $F^{-1}(\bullet)$. The detail coefficients W_f^e can be defined as:

$$W_f^e(n, t) = \langle f, \Psi_n \rangle = \int f(\tau)\overline{\Psi_n(\tau - t)}d\tau \quad (5)$$

Calculate the approximation coefficients $W_f^e(0, t)$:

$$W_f^e(0, t) = \langle f, \varphi_1 \rangle = \int f(\tau)\overline{\varphi_1(\tau - t)}d\tau \quad (6)$$

The empirical modes could be given by:

$$f_0(t) = W_f^e(0, t) * \varphi_1(t) \quad (7)$$

$$f_n(t) = W_f^e(n, t) * \Psi_n(t) \quad (8)$$

2.2 The shortcomings of EWT

Gilles proposed some rules to find ω_n , dividing the whole Fourier spectrum into N parts. The most common rule is to use the local maxima of the spectrum to determine the segments. However, N is set by the first several local maxima, where ω_n is taken as the median of two adjacent maxima.

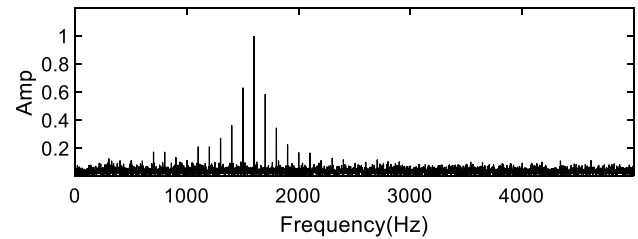
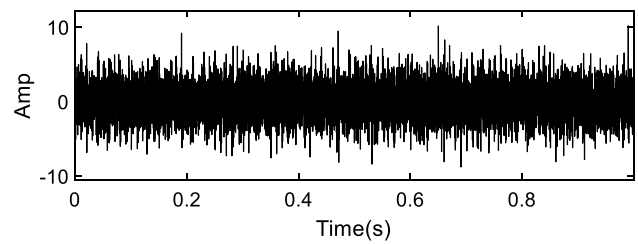


Fig. 2 The waveform of the simulation signal

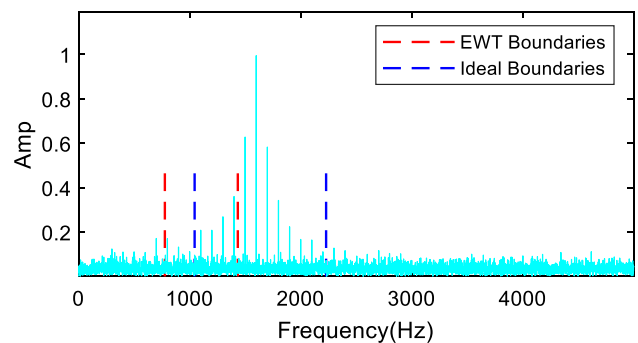


Fig. 3 Ideal boundaries and EWT boundaries

This may separate the resonance sideband into different parts due to the relative concentration of several maxima, so the periodic pulses component is not the best performance. A periodic pulses signal with a center frequency of 1500 Hz and a width of the sideband of 100 Hz is simulated by Eq. 9 and shown in Fig. 2.

$$s_0(t) = \sum_{i=1}^M A e^{-g2\pi f_n^i t} \sin\left(2\pi f_n^i t \sqrt{1 - g^2}\right) \quad (9)$$

where the amplitude $A=5$, the nature frequency $f_n=1600$ Hz, the damping coefficient $g = 0.05$, $M = 100$, and noise is $SNR = -5$ dB.

In Fig. 3, a resonance band appears in the spectrum of the simulation signal. It is necessary to extract the sideband for analysis in order to achieve the purpose of fault diagnosis. The two points of local maxima in frequency need to be selected according to the rule of EWT, which is highlighted by the red point. According to the local maxima rule, the first boundary is located near 777 Hz and the second one is around 1434 Hz, as shown in Fig. 3. Unfortunately, the

resonance frequency band deviates greatly from our ideal result, and the ideal resonance frequency band is divided into two parts, weakening the fault feature information.

Figure 4 shows the filters constructed by original EWT which divide the whole frequency domain into three parts. Both $\hat{\Psi}_1$ and $\hat{\Psi}_2$ contain a part of the ideal fragment. As shown in Fig. 5, the periodic pulses are seen vaguely in the second empirical mode with too much noise. Meanwhile, the periodic shock in the third empirical mode has been submerged in the strong high frequency. Obviously, the local maximum rule cannot extract the appropriate resonance sideband in the spectrum due to the limitation of its fragment division, which affects the final extraction effect.

3 Optimized empirical wavelet transform

The resonance band in the frequency domain can be seen as the generation of a series of pulses. Therefore, by dividing the spectrum into several scales and calculating the kurtosis of each scale, the trend of the spectrum can be depicted, which is a good prerequisite for effectively cutting the frequency domain. In this section, multi-scale empirical wavelet transform (MSEWT) was proposed to introduce the concept of multi-scale kurtosis into frequency domain. The proposed method is shown in Fig. 6, and the detail procession can be described as follows:

Step 1: Cutting the frequency domain using a given scale. In order to facilitate program editing, η is introduced to represent scale (scale = ηf_q). When $1 \leq \eta \leq 3$, the scale ($f_q \leq \text{scale} \leq 3f_q$) is seen as a suitable range.

Step 2: Calculate the kurtosis of each fragment. Since kurtosis is sensitive to accidental impulses, the corresponding kurtosis of the segment will increase when there is an obvious frequency peak in the scale segment. The local kurtosis transformation in the spectrum can reflect the strength of the frequency components in the spectrum. Extract the frequency concentrated part may be able to obtain the desired fault feature components.

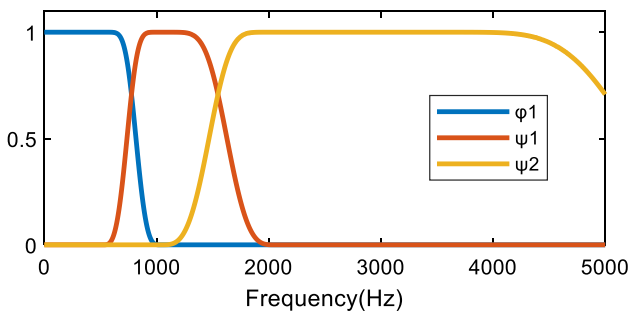


Fig. 4 Filters of EWT

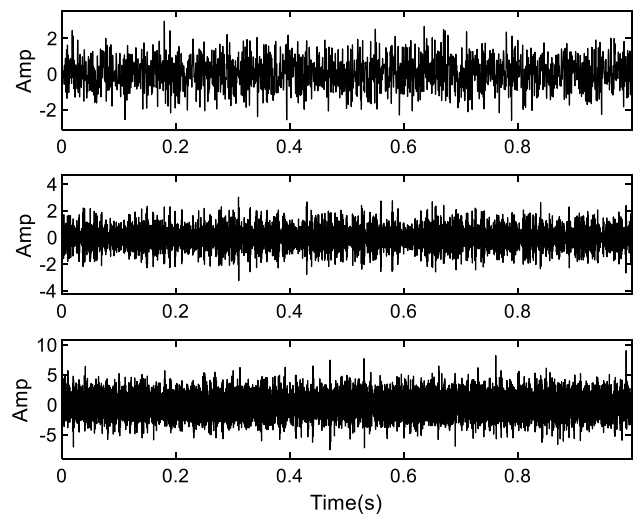


Fig. 5 Results decomposed by EWT

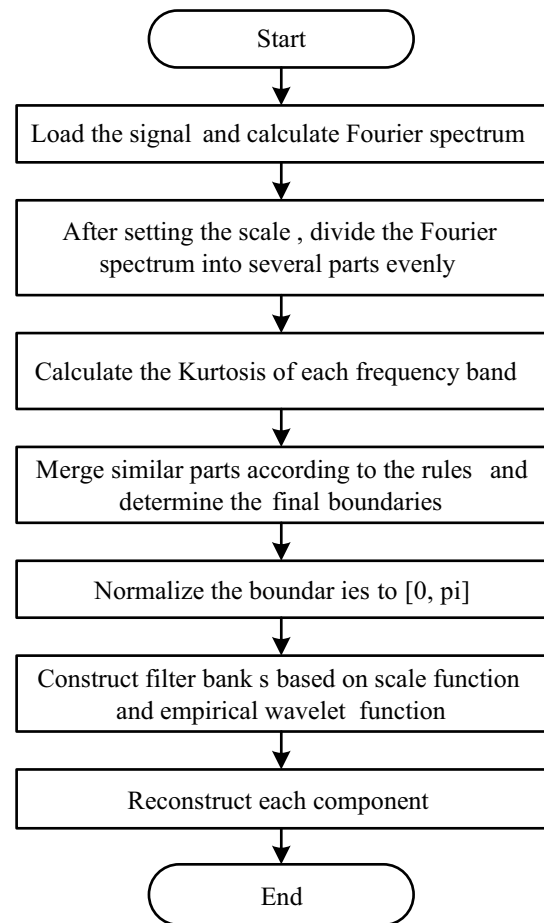


Fig. 6 Flowcharts of multi-scale empirical wavelet transform

Step 3: Fragments merging. The artificial averaging rule is defined by the following formula:

$$(K_i - \bar{K}_n)(K_{i+1} - \bar{K}_n) \leq 0 \tag{10}$$

where K_i represents the kurtosis of the i -th scale fragment.

If the scale kurtosis of two adjacent satisfies Eq. 10, the middle boundary between the two adjacent scale segments will be retained. If the formula is not satisfied, the two dimensions will be fused together. The two task segments contain the same components and need not be segmented.

Step 4: Construct filter bank based on empirical scale function and empirical wavelet function. Each individual filter in the filter bank represents an independent component. All components make up the original signal. Reconstruct the information in each filter. The original signal will be decomposed into several components located in different frequency bands.

3.1 Case 1 of simulation experiment

When the rolling balls go through the fault position of the bearing outer ring, the pulse will occur. Periodic pulses are often used to simulate bearing failures.

For the signal mentioned in Eq. 9, the spectrum will be divided into 25 parts of which red are the initial boundaries. Kurtosis is used to describe the spectrum trend characteristics to emerge the resonance frequency band. The useless boundaries will be merged by the artificial averaging rule. The blue line in Fig. 7 is the boundaries after merging. After removing the redundant boundaries, only two boundaries are finally determined, which separates the resonance sideband from the whole spectrum. Figure 8 shows the filters constructed by MSEWT.

As shown in Fig. 8, the empirical scale function and empirical wavelet function are constructed according to the final boundaries. Three components are obtained by MSEWT and shown in Fig. 9, with the second empirical mode containing rich characteristic frequency. The optimized frequency division method can effectively separate the resonance sideband from the signals including noise.

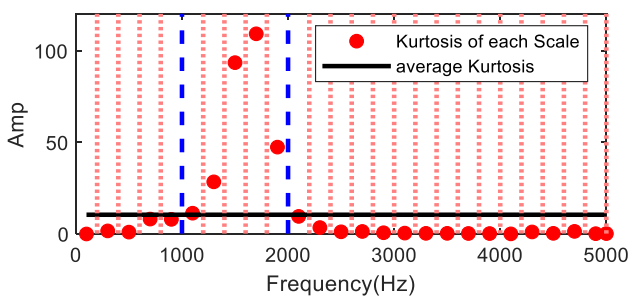


Fig. 7 Boundaries and kurtosis of the spectrum

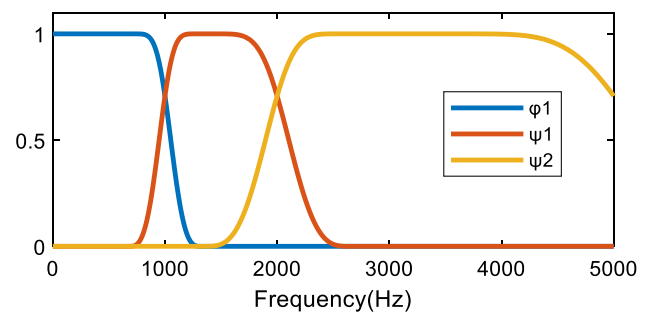
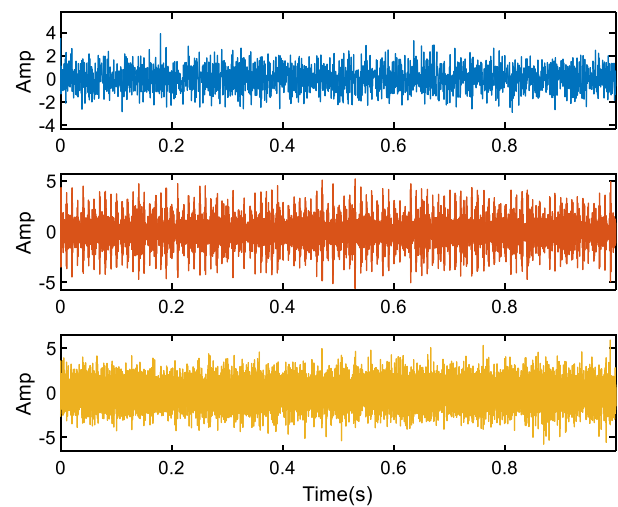
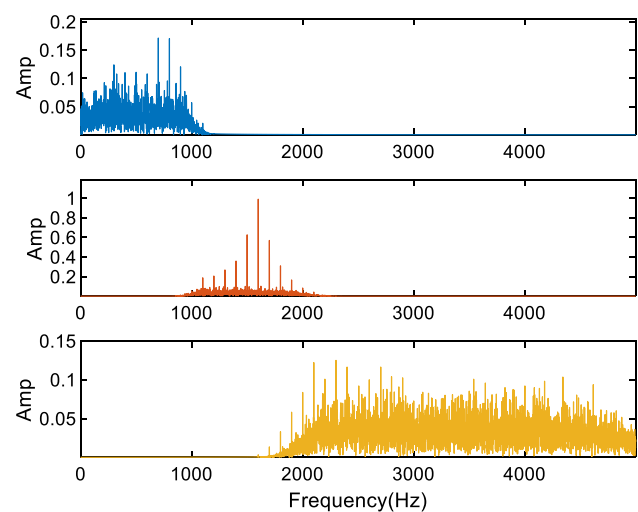


Fig. 8 Filters of MSEWT



(a) Components decomposed by MSEWT



(b) Spectra of the components

Fig. 9 The results decomposed by MSEWT

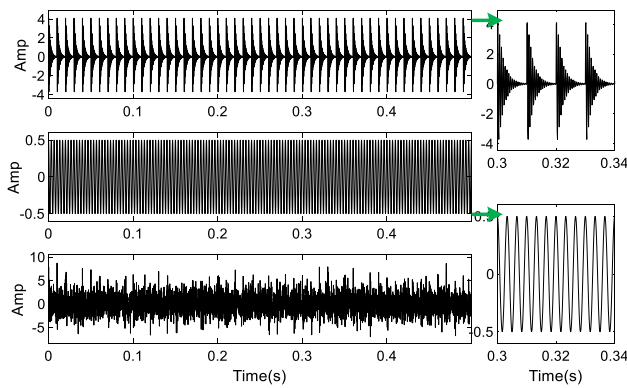


Fig. 10 Simulated signal

3.2 Case 2 of simulation experiment

In this section, a cosine signal is added in the periodic pulses, as shown in Fig. 10. The composition of signal 2 is as follows:

$$\begin{cases} s_{c1}(t) = \sum_{i=1}^M A e^{-g2\pi f_n^i t} \sin(2\pi f_n^i t \sqrt{1-g^2}) \\ s_{c2}(t) = 0.5 \sin(2\pi ft) \\ s_2(t) = s_{c1}(t) + s_{c2}(t) + \zeta \end{cases} \quad (11)$$

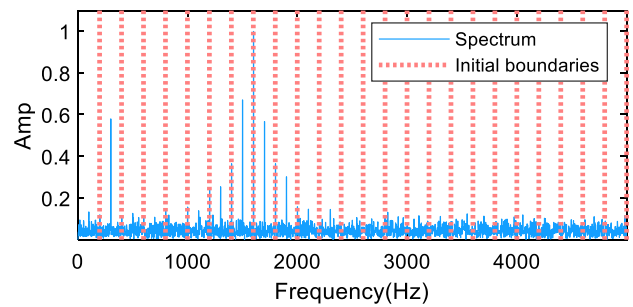
where the amplitude $A=5$, the nature frequency $f_n=1600$ Hz, the damping coefficient $g = 0.05$, $M = 100$, frequency $f = 300$ Hz, and noise is $\zeta = (\text{SNR} - 3 \text{ dB})$.

The spectrum separated by scaled boundaries is shown in Fig. 11. It can be seen that the frequency of 300 Hz is quarantined into the second fragment, which should be extracted for further analysis. The kurtosis is high in the second fragment regions because of its sensitivity to protrusions, shown in Fig. 11b. The dotted blue lines divide the frequency domain into five parts, exactly separating the single frequency component and the periodic pulse component according to the artificial averaging rule.

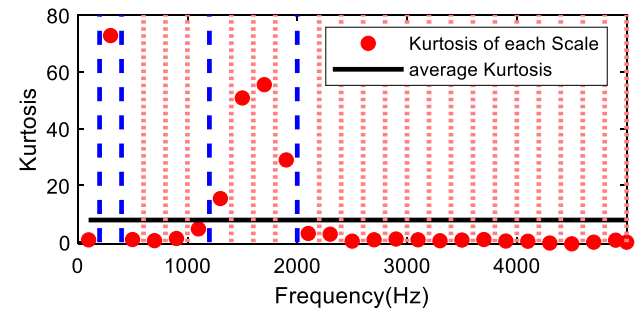
Then, one scale function and four empirical wavelet functions are constructed, as shown in Fig. 12a. The components are displayed, and the single frequency component and the periodic pulse component are effectively extracted in the second and fourth part. The result shows that the proposed spectrum partition method can effectively divide the frequency band containing a single frequency component.

3.3 Case 3 of simulation experiment

In this section, the simulation signal 3 (shown in Fig. 13) including periodic impulse signal, modulation signal, and accidental impulse signal is constructed to further verify



(a) Spectrum and boundaries



(b) Kurtosis of each Scale

Fig. 11 Spectrum and initial boundaries of signal

the effectiveness of the method in segment segmentation. The composition of the simulation signal is as follows:

$$\begin{cases} s_{c1}(t) = \sum_{i=1}^M 5e^{-g_1 2\pi f_{n1}^i t} \sin(2\pi f_{n1}^i t \sqrt{1-g_1^2}) \\ s_{c2}(t) = 0.8 \sin(2\pi \cdot 3000t) (1 + \cos(2\pi \cdot ft)) \\ s_{c3}(t) = 150e^{-g_2 2\pi f_{n2}^i t} \sin(2\pi f_{n2}^i t \sqrt{1-g_2^2}) \\ s_2(t) = s_{c1}(t) + s_{c2}(t) + s_{c3}(t) + \zeta \end{cases} \quad (12)$$

where the nature frequency is $f_{n1} = 1600$ Hz, $f_{n2} = 4000$ Hz. The damping coefficient $g_1 = 0.05, g_2 = 0.1$, $M = 100$, frequency $f=150$ Hz, and noise is $\zeta = (\text{SNR} - 2 \text{ dB})$.

The spectrum of signal 3 is shown in Fig. 14a, which was separated into 25 parts. The results decomposed by the proposed method are shown in Fig. 14b. MSEWT obtained four blue dotted lines as the boundaries. Then, filters are obtained according to the boundaries which divide the spectrum into five parts. Figure 15 shows the components and their spectrum. The periodic pulse information is extracted. The periodic characteristics can be well restored in extracted component after eliminating noise interference. The modulation signal is separated by the third band-pass filter. The waveforms of the two signals are basically identical, and obvious modulation frequency spectrum lines appear in the spectrum. Signal 3 demonstrates that the MSEWT can effectively separate the periodic pulse component, the accidental pulse component, and the modulation component (Fig. 16).

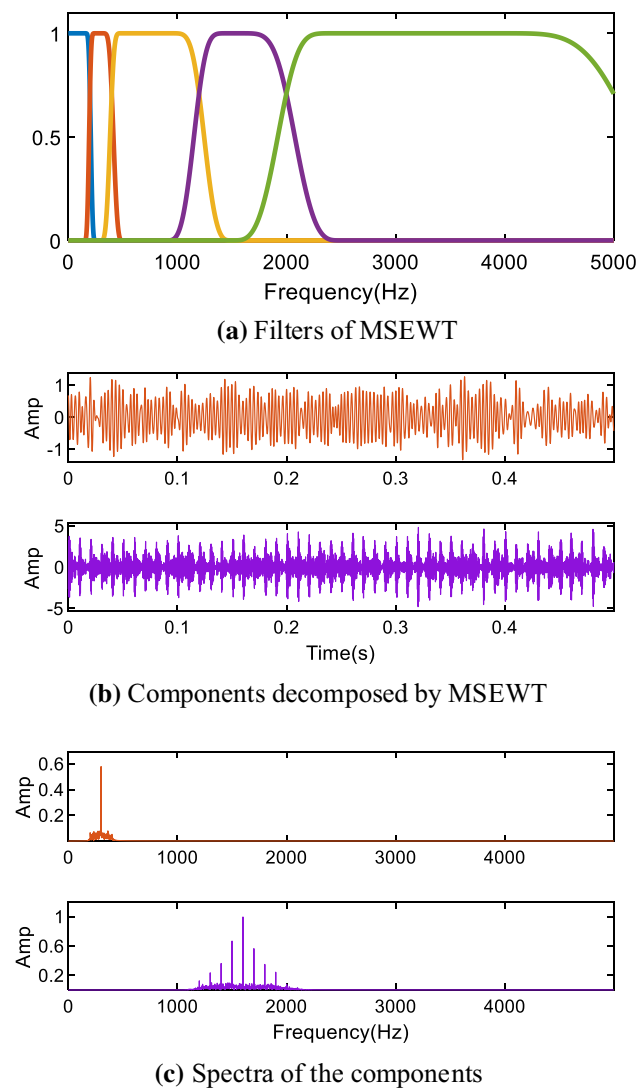


Fig. 12 The results decomposed by MSEWT

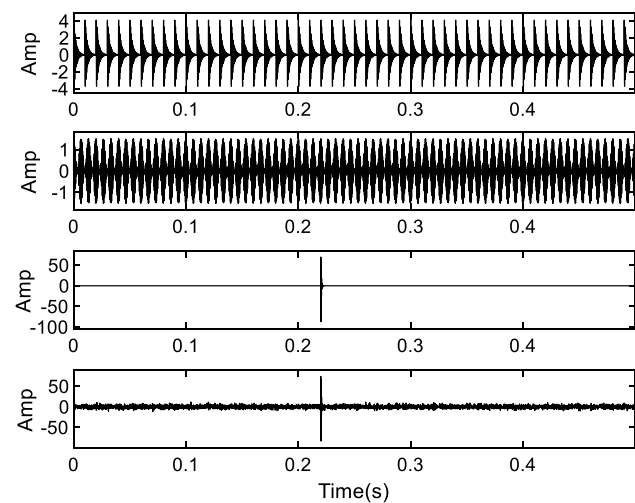


Fig. 13 Simulated signal

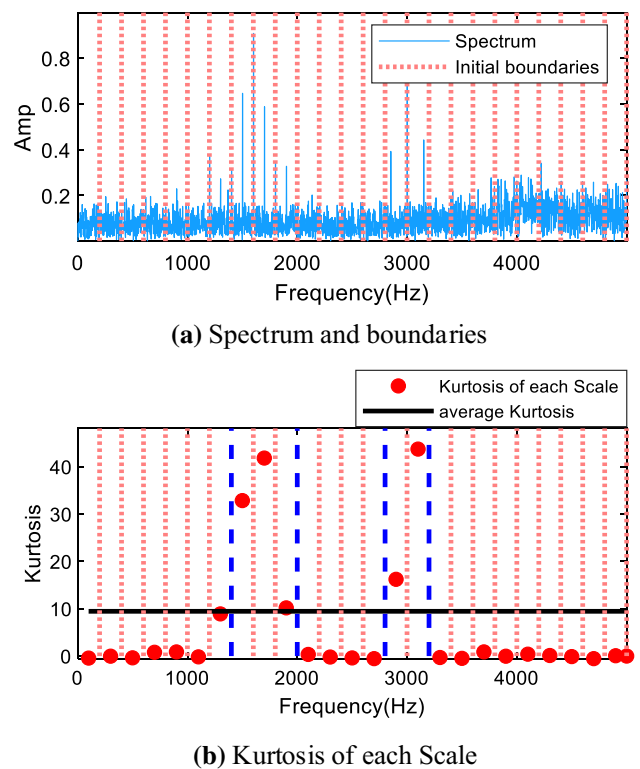


Fig. 14 Spectrum and initial boundaries of signal

4 Fault feature extraction method based on SMSEWT

The MSEWT method proposed in Sect. 3 obtains more reasonable boundaries and components, but it is necessary to extract required information from them. The analysis of envelope spectrum is normally used to diagnose bearing faults in engineering applications according to its characteristic frequency and harmonics. The bearing fault characteristic frequency and its harmonics shown in the envelope spectrum can be regarded as a few spikes with large amplitudes that reflect bearing fault signatures. Hence, [17] defines these spikes in the envelope spectrum as the sparsity representation of a bearing fault signal in a frequency domain. Sparsity measurement used in envelope spectrum shows good performance, which avoids the accidental interference in spectrum. The paper proposed sparsity-guided MSEWT (SMSEWT) to extract information for bearing fault diagnosis. More importantly, in order to obtain the envelope of the signal directly, we incorporate Hilbert transform into the filtering structure of SMSEWT.

4.1 The theoretical basis of sparsity

Sparsity has been applied in many fields. The basic expression can be described as:

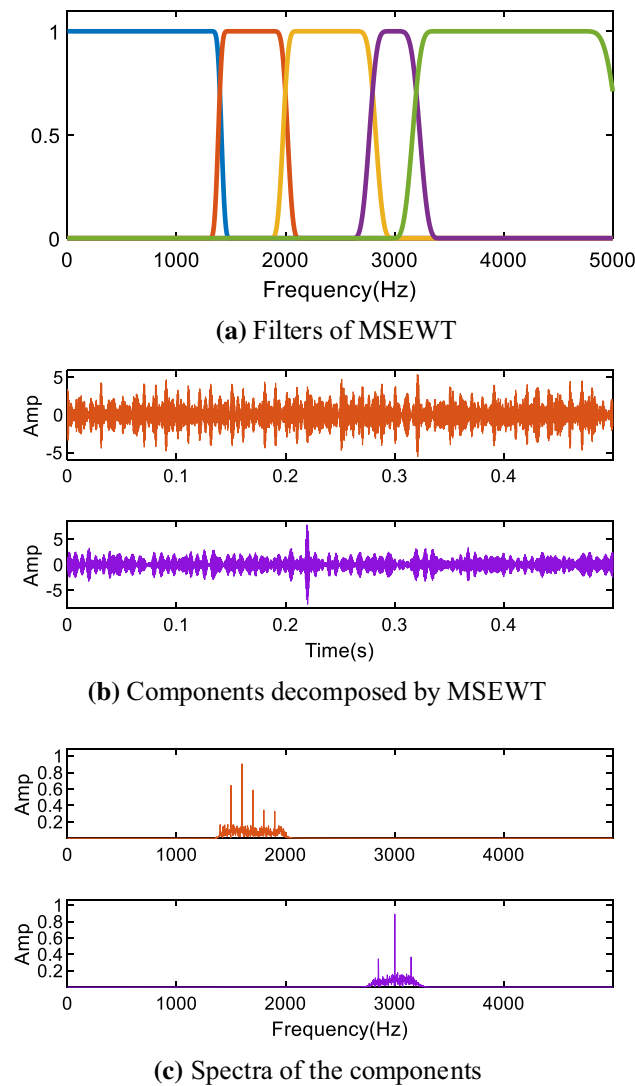


Fig. 15 The results decomposed by MSEWT

$$S = \frac{\sqrt{\sum_{f=0}^{F_s/2-1} |d(f)|^2}}{\sum_{f=0}^{F_s/2-1} |d(f)|} = \frac{\|d(f)\|_2}{\|d(f)\|_1} \tag{13}$$

where $\|d(f)\|_2$ and $\|d(f)\|_1$ are L2 norm and L1 norm, and $d(f)$ is the corresponding power spectrum of the envelope. It is worth mentioning that the sparsity of bearing signal needs to be calculated in the envelope spectrum, which can be regarded as a sparsity signal. However, the proposed method uses Hilbert transform algorithm to obtain the envelope spectrum.

4.2 The theoretical basis of Hilbert transform

Much mechanical fault information exists in the form of modulation in vibration and noise. For the processing of

modulated signal, there are two main purposes: obtaining the envelope and phase demodulation. In digital signal processing, this processing is particularly convenient with the help of Hilbert transform. The modulation signal is assumed to be in the form of:

$$x(t) = A(t)\cos(2\pi f_0 t + \varphi(t)) \tag{14}$$

Its Hilbert transformation can be expressed as:

$$\hat{x}(t) = \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{15}$$

Thus the envelope estimation of $x(t)$ can be obtained by:

$$\bar{A}(t) = |x(t) + \hat{x}(t)| \tag{16}$$

Then, the envelope spectrum can be obtained by calculating the FFT of $\bar{A}(t)$.

4.3 Connection of two methods

The sparsity cannot be calculated without the corresponding power spectrum of the envelope $d(f)$ obtained by the FFT of $\bar{A}(t)$'s autocorrelation. Interestingly, $d(f)$ can be worked out by EWT because of the coexist calculation of FFT. It will reduce a cycle time of FFT by combining Hilbert transform when the component is extracted by filter. Therefore, the empirical envelope signal is obtained from EWT.

$$f_{e0}(t) = H[W_f^e(0, t) * \varnothing_1(t)]^V \tag{17}$$

$$f_{en}(t) = H[W_f^e(n, t) * \Psi_n(t)]^V \tag{18}$$

where $[\bullet]^V$, $H[\bullet]$ denotes Fourier transform and Hilbert transform.

4.4 The proposed sparsity-guided multi-scale EWT

In this paper, a new method named sparsity-guided multi-scale EWT (SMSEWT) is proposed. Firstly, the vibration acceleration signals need to be collected by sensors and saved as digital signals. Then, set the scale and divide the entire spectrum equally. In the sample shown in Fig. 6, the initial boundaries divide the spectrum into 25 parts. After calculating the kurtosis for each frequency band, the average kurtosis can be obtained. It can be found that the kurtosis of most frequency bands is lower than average kurtosis, and a small number of continuous frequency bands have large values. These distinctive contiguous frequency bands will be merged according to the rules, and the final boundaries will be determined. After constructing the filter bank with empirical scale function and empirical wavelet function, a set of components will be obtained. Figure 6 shows the sparsity of

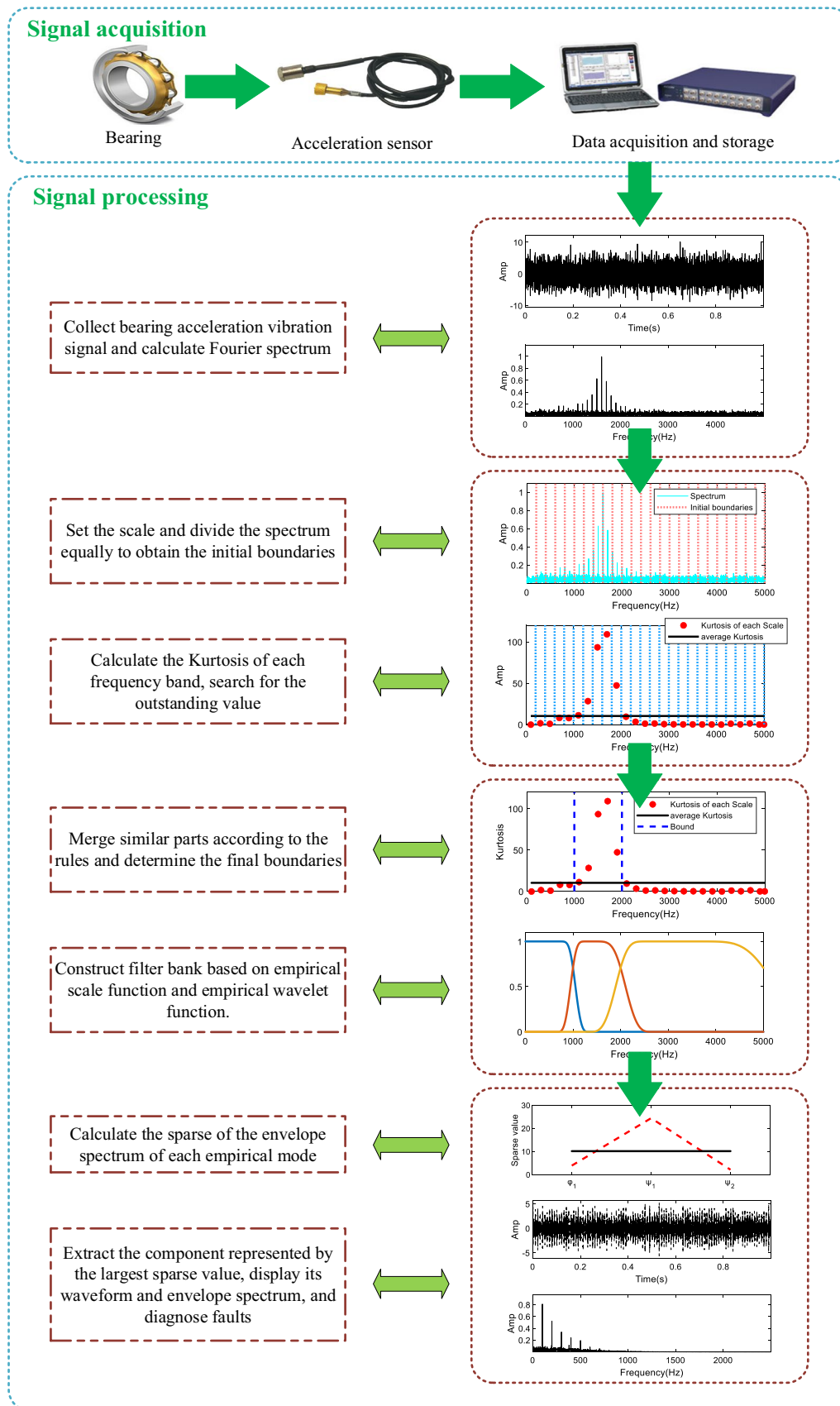


Fig. 16 The SMSEWT method

the envelope spectrum of each component, and the component with the largest value will be extracted. There are periodic pulses in the time-domain waveform, and the envelope spectrum contains characteristic frequencies and harmonics. The fault diagnosis is realized by further analysis.

5 Experimental Verification

5.1 Experimental verification of Case 1

As shown in Fig. 17, it is a schematic diagram of a bearing failure test bench, which can simulate the bearing fault type of the inner race, outer race, and balling. In the first case, SMSEWT is applied to the diagnosis of bearing faults in the rotor test rig. The collected signal shown in Fig. 18 comes from 6397 deep groove ball bearing with outer ring fault accompanied by characteristic frequencies of 76.08 Hz.

The bearing type and fault characteristic frequency is detailed in Table 1. The sampling frequency is 15360 Hz. An obvious resonance sideband near 2000 Hz needs to be extracted for bearing fault analysis.

As shown in Fig. 19, boundaries obtained by SMSEWT divide the whole frequency domain into seven parts according to the kurtosis of each part. The empirical wavelet filters are constructed based on the obtained boundaries. According to the proposed method, the sparsity of envelope signal is calculated to judge the fault information contained in this part. The maximum sparsity of envelope decomposition appears in the fourth part. Extracting this part, the envelope spectrum is obtained by FFT plotted in Fig. 19d. The extracted envelope signal embodies obvious periodic pulse in the time domain. At the same time, the fault characteristic frequency 76.88 Hz and its harmonics are obtained clearly in the envelope spectrum. The proposed method can effectively extract the bearing outer ring fault characteristic and successfully realize the bearing fault diagnosis.

In order to further verify the effectiveness of this method, Sparsogram, Fast Kurtogram (FK), and Protrugram are used for comparison. Figure 20 shows the maximum sparsity found for the Sparsogram at different decomposition levels (level=6 and 7). When the number of decomposition levels is 6, the

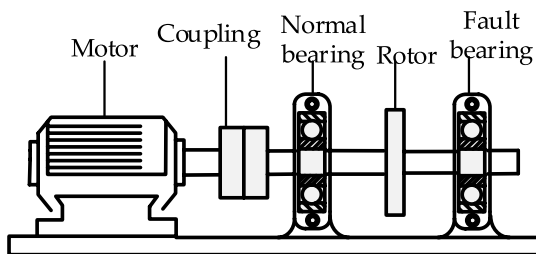


Fig. 17 Bearing fault test rig

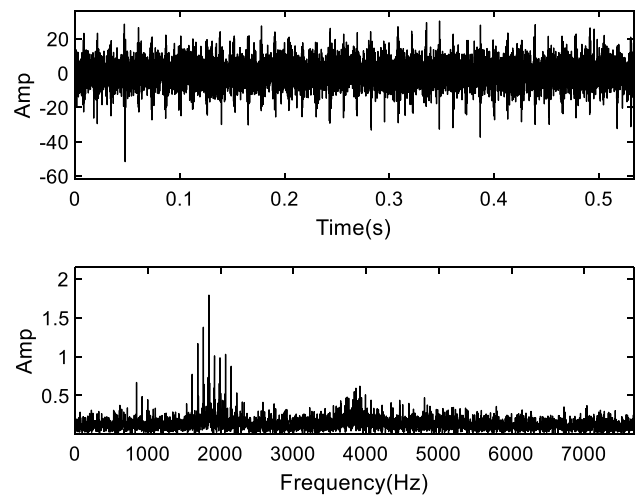


Fig. 18 Vibration signal of bearing failure of outer race

center frequency and bandwidth corresponding to the resonance frequency band obtained by the maximum sparsity are 1980 Hz and 120 Hz. Figure 20b shows the extracted fault feature envelope component. It is found that the envelope spectrum of the component has the pulse of the periodic interval but only the fault characteristic frequency and its second harmonics can be detected. The center frequency of the finally determined resonance frequency band is located at 7530 Hz, and the bandwidth is 60 Hz, when the level is 7. The result is completely different from the range determined when the level is 6. However, the fault signature frequency cannot be found in the envelope spectrum, and the extracted envelope signal has no regularity. The Sparsogram always goes to a narrower bandwidth in the frequency domain segmentation. As the level increases, the bandwidth will become smaller, and at the same time, the accuracy of the center frequency of the resonance band will decrease.

The filter-based FK is shown in Fig. 21 with a center frequency of 2080 Hz and a bandwidth of 320 Hz. Because of the deviation of the determined center frequency and the narrow bandwidth, the envelope component extracted contains only the fault characteristic frequency and its second harmonics, and the fault information is obviously less than that extracted by SMSEWT. The same result occurs in STFT-based FK. Considering comprehensively, the frequency segment extracted by Fast Kurtogram contains less fault feature information than SMSEWT because of the limitation of frequency band division.

Table 1 Characteristics frequency of 6307 bearing fault

Rotating speed	Inner race	Outer race	Rolling element	Cage
1496 rpm	123.39 Hz	76.08 Hz	49.61 Hz	9.51 Hz

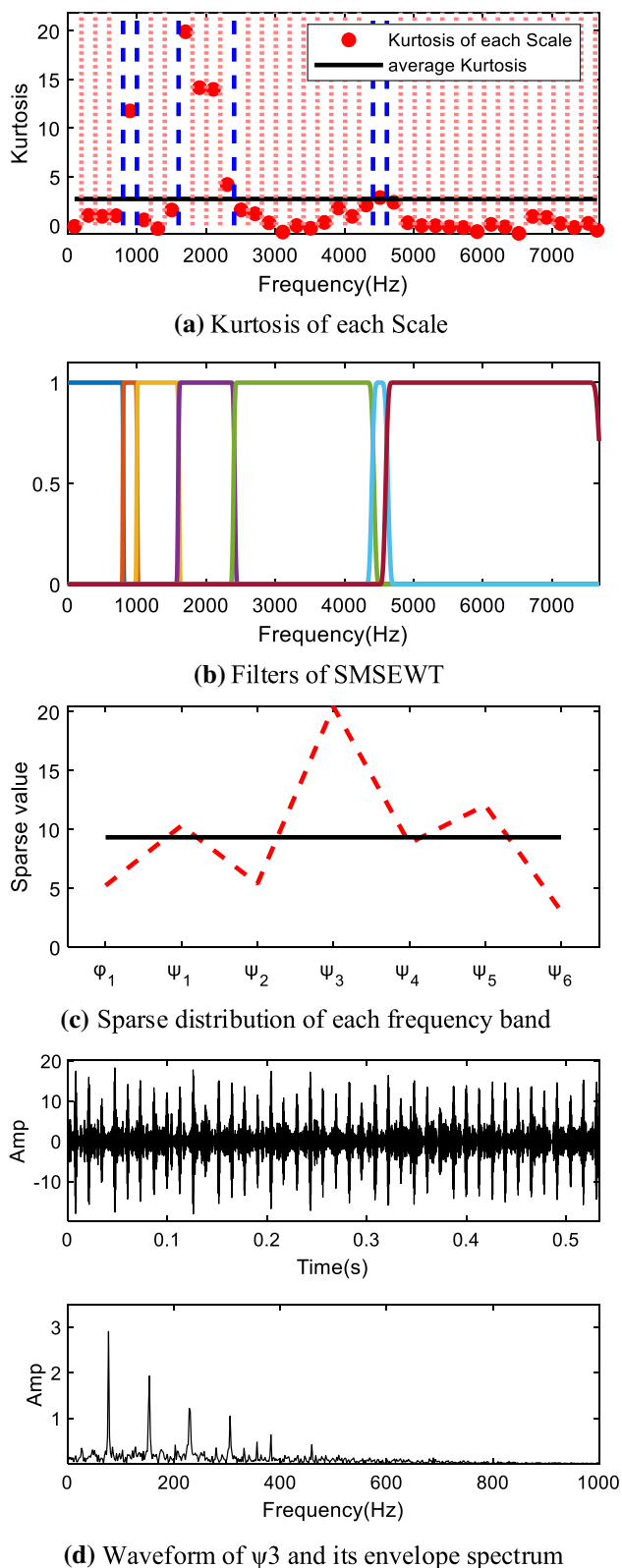


Fig. 19 The results decomposed by SMSEWT

The results of Protrugram are plotted in Fig. 22a, and it can be seen that the maximum value appears at 1795 Hz with the bandwidth which is 250 Hz. The envelope component extracted in Fig. 22b also contains inadequate fault feature information. Although this method can obtain more accurate central frequency, it sacrifices a lot of computing time with the scanning process.

5.2 Experimental verification of Case 2

In actual industrial production, some accidental effects caused by man-made or mechanical collisions are regularly mixed in vibration signals, which cause great difficulties for bearing fault diagnosis results. In this part, a fault vibration signal of bearing outer race containing accidental effects is obtained from HZXT-008, the small rotor rolling bearing test bench, shown in Fig. 23a. Figure 23b shows the location of the vibration sensor's measuring points in the actual acquisition: horizontal and vertical directions on the side near the planetary gearbox. In order to obtain vibration signals with more obvious fault characteristic frequencies, the signals collected by vertical sensors are analyzed to verify the effectiveness of the proposed method. The fault characteristic frequency of the measured bearing with the type of NSK HPS 6200 is shown in Table 2. The fault vibration signal of the bearing outer ring is obtained with the characteristic frequency of 76.20 Hz. The speed of motor is set to 1500 rpm, the sampling frequency is 12000 Hz, and the length of sampling sequence is 24000 (2 s sampling time). Figure 24 shows the collected vibration signal and its spectrum from the HZXT-008 rotor test bench. In the time domain, it can be found that there is a large accidental interference at 1.9 s, which will have a great influence on the kurtosis value. Affected by large disturbances, the value of the frequency in the frequency domain fluctuates within the range of 2500 Hz to 4000 Hz. If the segment filtered signal is obtained for analysis, an unsatisfactory analysis result may be obtained to make the diagnosis result fail. Four methods, SMSEWT, Sparsogram, FK, and Protrugram, were applied to analyze the signal with interference, and the results of the analysis were compared as follows.

Figure 25 shows the SMSEWT process. The sparsity of the two parts is higher, which correspond to 0–100 Hz filtered by \varnothing_1 and 900–1200 Hz filtered by Ψ_2 . The sparsity of the envelope component obtained by Ψ_2 is the largest, indicating the most faulty feature information. In the frequency domain obtained by Ψ_2 , the frequency 76.25 Hz and its second, third harmonics can be found which are close to the outer race fault characteristic frequency. These characteristic frequencies are sufficiently significant to diagnose the bearing outer ring failure. However, the filter \varnothing_1 extracts the envelope component of the low-frequency component containing more abundant power frequency information.

Fig. 20 Results decomposed by Sparsogram

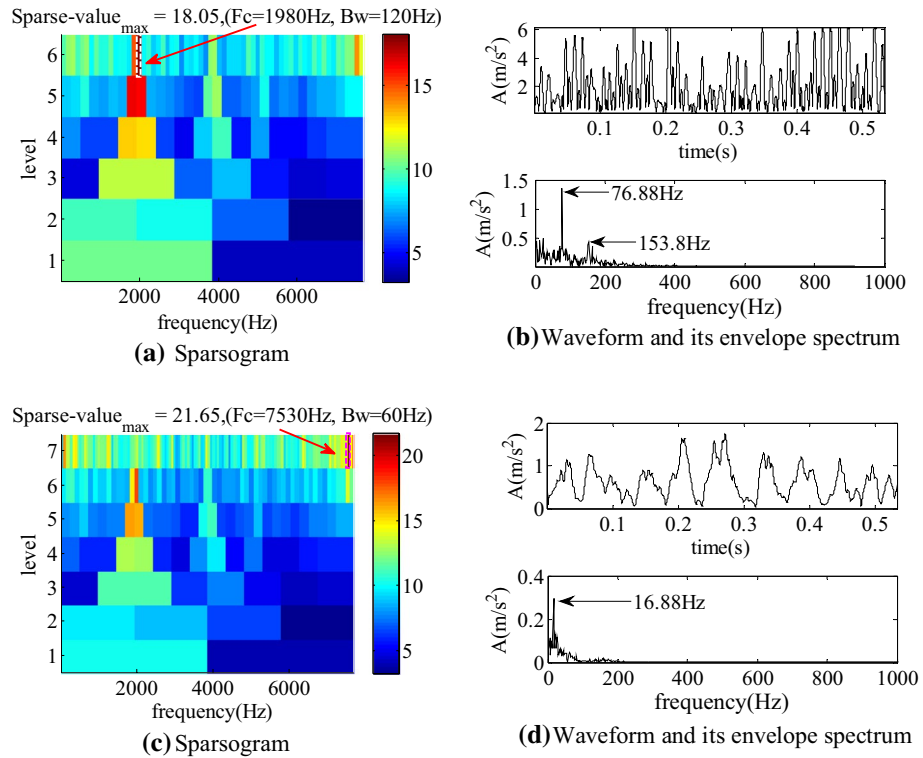


Fig. 21 Results decomposed by Fast Kurtogram

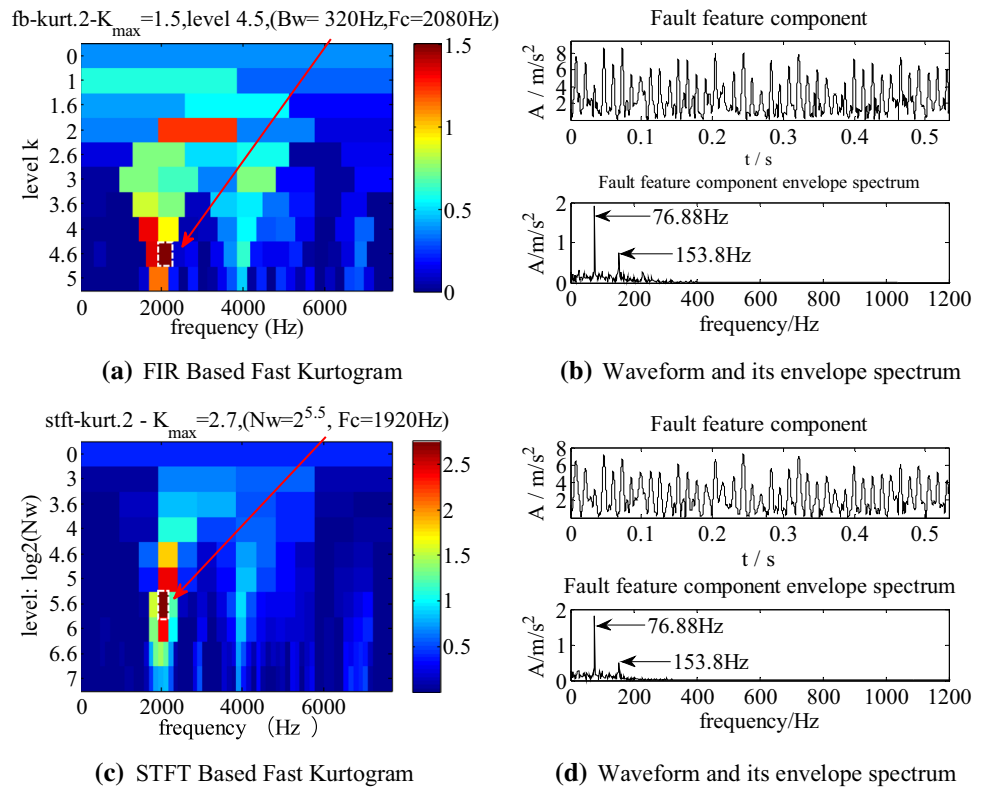


Fig. 22 Results decomposed by Protrugram

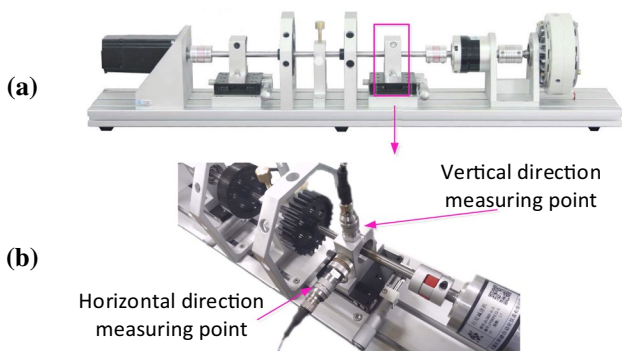
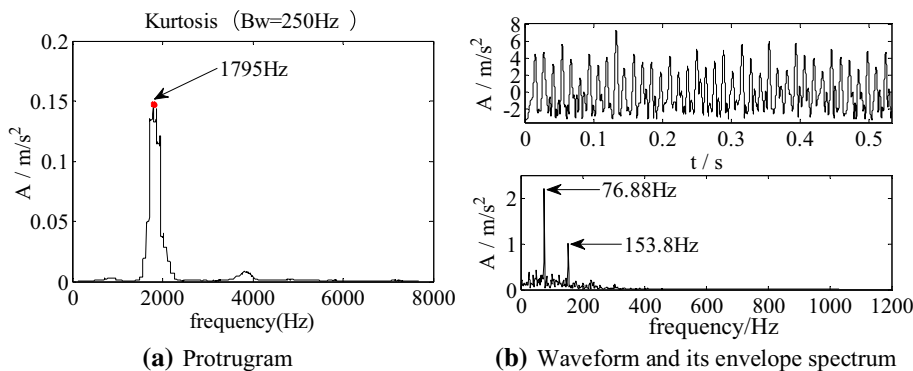


Fig. 23 HZXT-008 small rotor test bench

Table 2 Characteristics frequency of NSK HPS 6200 bearing fault

Rotating speed	Inner race	Outer race	Rolling element	Cage
1500 rpm	123.80 Hz	76.20 Hz	49.55 Hz	9.53 Hz

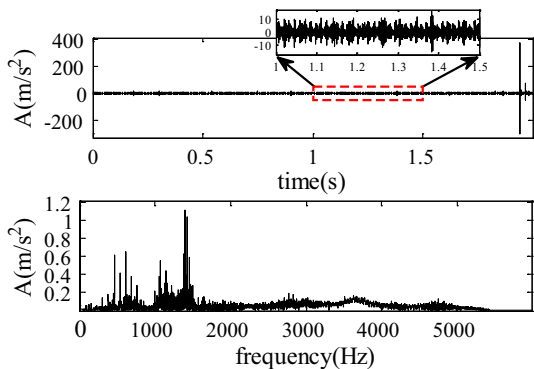


Fig. 24 Vibration signal of bearing failure of outer race with accidental pulse

The same as the former, although the envelope contains the interference component in the time domain, the envelope spectrum can show the periodicity. This fully proves that the proposed method can effectively avoid the accidental

interference and successfully extract the bearing fault information to realize the bearing fault diagnosis.

As shown in Fig. 26, when level is equal to 4, the center frequency and bandwidth of the resonance band corresponding to the maximum value obtained by Sparsogram are 937.5 Hz and 375 Hz. The resonance band range acquired is slightly lower than that obtained by SMSEWT. However, the obtained fault characteristic frequencies are found without any harmonic frequencies in the envelope spectrum. It is worth mentioning that when level is equal to 7, the whole spectrum was divided into smaller segments. The position of resonance band is determined to a lower frequency band. No periodic impulse information in the envelope waveform is obtained without any obvious fault characteristic frequency found in the envelope spectrum. The result cannot provide reliable data basis for bearing fault diagnosis.

The results of the FK are plotted in Fig. 27. By dividing of filter-based FK, the maximum kurtosis corresponding to the band range is determined in the middle of the entire frequency domain (2000 Hz~4000 Hz). For STFT-based FK, the obtained frequency domain is located in 3000 Hz~4500 Hz. Figure 27b shows the envelope signal and its envelope spectrum obtained by the filter-based FK. There is a prominent accidental interference in the envelope signal of the time domain. Since the amplitude of the interference is large, the vibration information of bearing fault cannot be clearly expressed. The frequency of the envelope spectrum has a trend of fluctuate due to the influence of interference, though the power frequency and the characteristic frequency appear in envelope spectrum. Obviously, the extraction performance of this method is unsatisfactory. The same situation also appears in Fig. 27d, the envelope spectrum of components obtained by STFT-based FK. Therefore, FK loses its ability to determine the resonance sideband for the bearing vibration signal with accidentally disturbed. At the same time, it confirms that kurtosis has strong sensitivity to accidental pulse.

The analysis results of the Protrugram are described in Fig. 28, with the filter bandwidth of 230 Hz and the step size of 1. The center frequency found by this method

Fig. 25 Results decomposed by SMSEWT

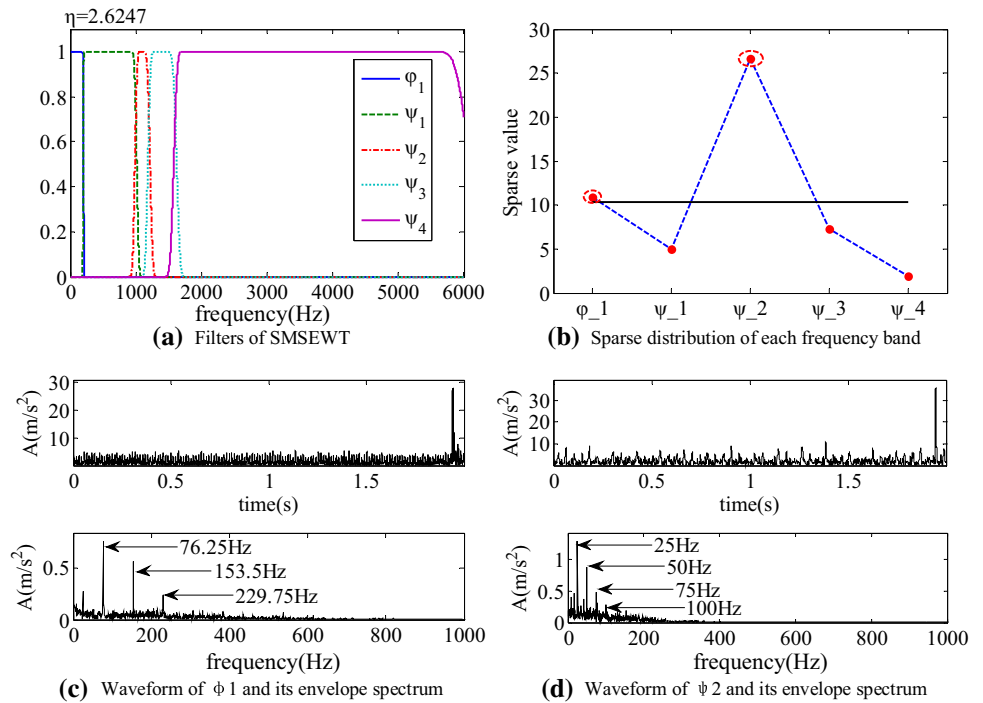
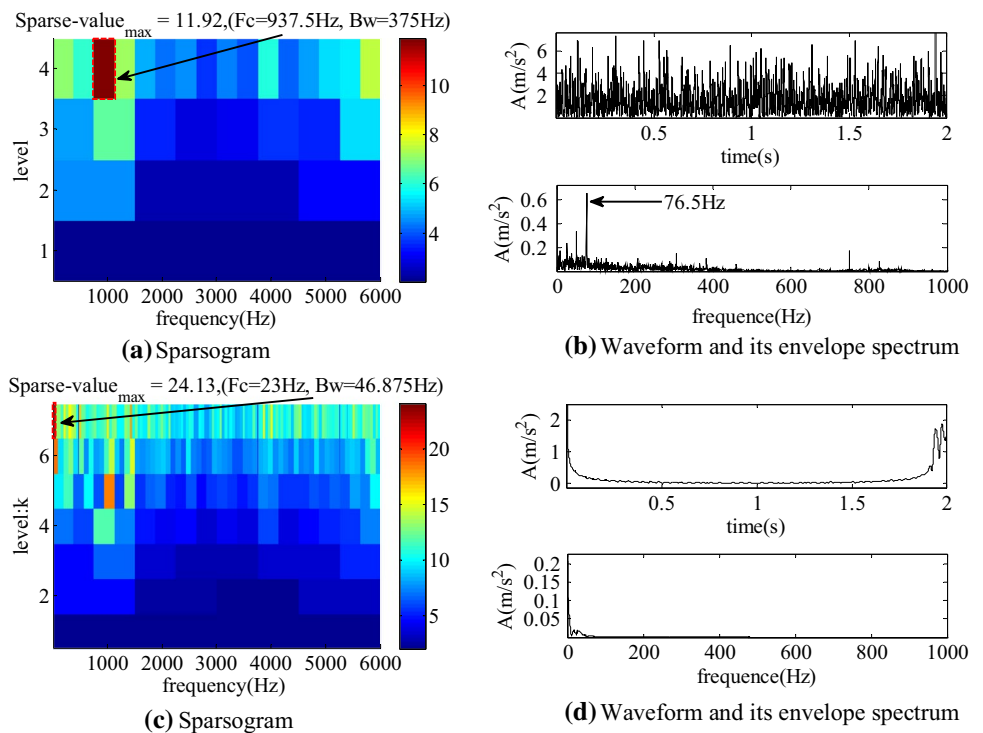


Fig. 26 Results decomposed by Sparsogram



is 1322 Hz. Protrugram estimates the center frequency by calculating the spectral kurtosis of the filtered signal, avoiding the effects of accidental interference. However, due to the fixed bandwidth, only fault characteristic frequency and the power frequency with its second harmonics

appear in the obtained envelope spectrum, and the extraction effect is obviously inferior to SMSEWT.

In engineering applications, the collected vibration signal is often mixed with accidental interference. By analyzing the bearing vibration signal with accidental pulse, it is fully

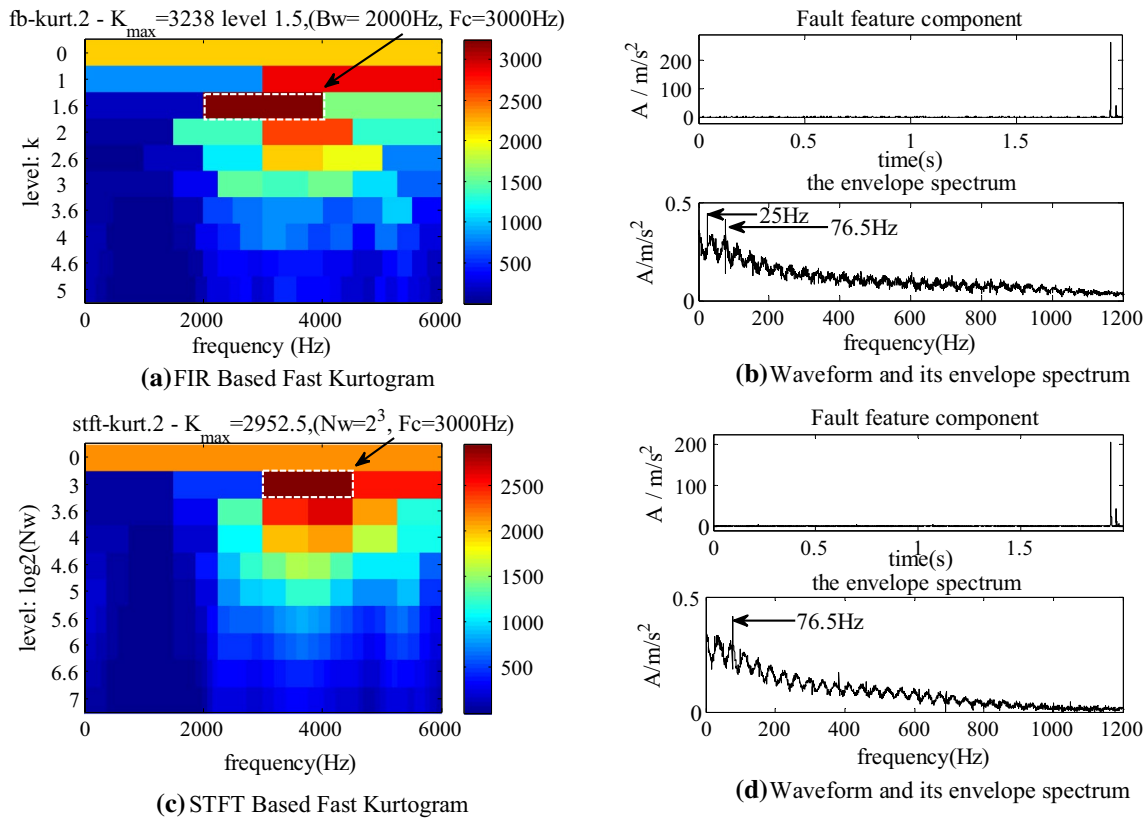
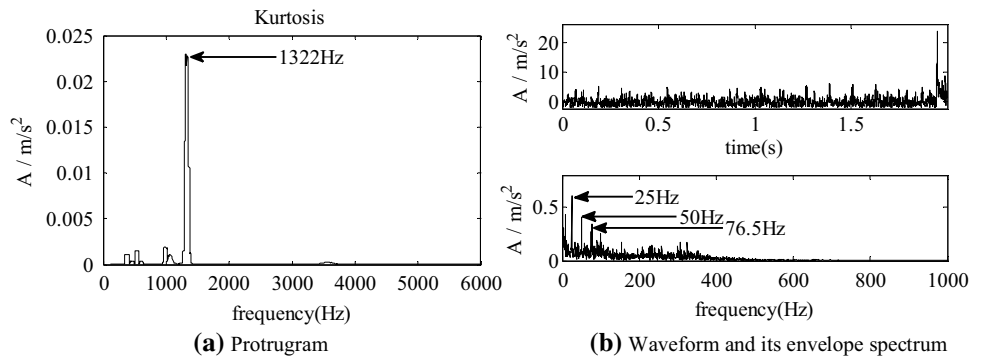


Fig. 27 Results decomposed by FK

Fig. 28 The results decomposed by Protrugram: a kurtosis curve by Protrugram; b extracted component of maximum kurtosis in spectrum



verified that the proposed SMSEWT can effectively avoid the influence of accidental interference and successfully extract the bearing fault characteristic frequency.

6 Conclusion

In this paper, a new method named sparsity-guided multi-scale empirical wavelet transform (SMSEWT) is proposed for rolling element bearing fault features extraction. Two major parts are included in the proposed algorithm. The first part is to improve the segmentation of the empirical wavelet

method. By defining the concept of multi-scale kurtosis, the spectrum was divided into several sub-parts and the kurtosis of each sub-portion was obtained. Then, the boundary merging was carried out by the artificial averaging rule, and the boundary adaptive rational division was realized. Three simulated signals were used to verify the validity of the boundary partitioning method. The second part is the extraction of the special effective component of the fault. The sparsity was selected as the basis for judging the regularity component of the fault feature to determine whether the segment was extracted. The highlight of this part is that the Hilbert transform and EWT are combined properly to

extract the envelope information of the fault feature component directly, which optimizes the calculation time. In this study, two experimental signals were used to verify the effectiveness of the proposed method.

This method optimizes the segmentation of EWT and makes it more suitable for analyzing bearing vibration signals. With adaptability, it avoids the influence of traditional methods. This method can be extended to fault diagnosis for gear, providing a novel and effective diagnostic method for fault diagnosis of rotor machinery.

Funding: This work is supported by the National Natural Science Foundation of China (Grant No. 51775005).

Declaration

Conflict of interest The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Song L, Wang H, Chen P (2018) Vibration-based intelligent fault diagnosis for roller bearings in low-speed rotating machinery. *IEEE Trans Instrum Meas* 67(8):1887–1899
- Zhang K, Ma C, Xu Y, Chen P, Du J (2021) Feature extraction method based on adaptive and concise empirical wavelet transform and its applications in bearing fault diagnosis. *Measurement* 172:108976
- Zhang F, Huang J, Chu F (2020) Mechanism and method for outer raceway defect localization of ball bearings. *IEEE Access* 8:4351–4360
- Xu Y, Zhang K, Ma C (2019) Adaptive Kurtogram and its applications in rolling bearing fault diagnosis. *Mech Syst Signal Pr* 130(1):87–107
- Antoni J (2007) Fast computation of the Kurtogram for the detection of transient faults. *Mech Syst Signal Pr* 21(1):108–124
- Antoni J (2006) The spectral kurtosis: a useful tool for characterising non-stationary signals. *Mech Syst Signal Pr* 20(2):282–307
- Antoni J, Randall RB (2006) The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. *Mech Syst Signal Pr* 20(2):308–331
- Sawalhi N, Randall RB, Endo H (2007) The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mech Syst Signal Pr* 21(6):2616–2633
- Combet F, Gelman L (2009) Optimal filtering of gear signals for early damage detection based on the spectral kurtosis. *Mech Syst Signal Pr* 23(3):652–668
- Lei Y (2011) Application of an improved Kurtogram method for fault diagnosis of rolling element bearings. *Mech Syst Signal Pr* 25(5):1738–1749
- Jia F (2015) Early fault diagnosis of bearings using an improved spectral kurtosis by maximum correlated kurtosis deconvolution. *Sensors* 15(11):29363–29377
- Barszcz T, Jabłoński A (2011) A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram. *Mech Syst Signal Pr* 25(1):431–451
- Wang Y, Liang M (2011) An adaptive SK technique and its application for fault detection of rolling element bearings. *Mech Syst Signal Pr* 25(5):1750–1764
- Wang Y, Liang M (2012) Identification of multiple transient faults based on the adaptive spectral kurtosis method. *J Sound Vib* 331(2):470–486
- Liu H (2014) Adaptive spectral kurtosis filtering based on Morlet wavelet and its application for signal transients detection. *Signal Process* 96:118–124
- Xu Y, Tian W (2019) Application of an enhanced Fast Kurtogram based on empirical wavelet transform for bearing fault diagnosis. *Meas Sci Technol* 30(3):035001
- Tse PW, Wang D (2013) The design of a new Sparsogram for fast bearing fault diagnosis: Part 1 of the two related manuscripts that have a joint title as “Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement - Parts 1 and 2.” *Mech Syst Signal Pr* 40(2):520–544
- Moshrefzadeh A, Fasana A (2018) The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mech Syst Signal Pr* 105:294–318
- Tse PW, Wang D (2013) The automatic selection of an optimal wavelet filter and its enhancement by the new Sparsogram for bearing fault detection. *Mech Syst Signal Pr* 40(2):520–544
- Varanis M, Silva AL (2021) A Tutorial Review on Time-Frequency Analysis of Non-Stationary Vibration Signals with Non-linear Dynamics Applications. *Braz J Phys* 51:859–877
- Daubechies I, Lu J, Wu H (2011) Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl Comput Harmon Anal* 30(2):243–261
- Oberlin T, Meignen S, Perrier V, The Fourier-based Synchrosqueezing Transform. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing, hal-00916088f.
- Huang NE (1971) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proce Mathe Phys Eng Sci* 1998(454):903–995
- Gilles J (2013) Empirical wavelet transform. *IEEE Trans Signal Process* 61(16):3999–4010
- Kedadouche M, Thomas M, Tahan AJMS (2016) A comparative study between empirical wavelet transforms and empirical mode decomposition methods: application to bearing defect diagnosis. *Mech Syst Signal Process* 81:88–107
- Cao H (2016) Wheel-bearing fault diagnosis of trains using empirical wavelet transform. *Measurement* 82:439–449
- Kedadouche M, Liu Z, Vu V (2016) A new approach based on OMA-empirical wavelet transforms for bearing fault diagnosis. *Measurement* 90:292–308
- Zheng J (2017) Adaptive parameterless empirical wavelet transform based time-frequency analysis method and its application to rotor rubbing fault diagnosis. *Signal Process* 130:305–314
- Wang D, Tsui K, Qin Y (2019) Optimization of segmentation fragments in empirical wavelet transform and its applications to extracting industrial bearing fault features. *Measurement* 133:328–340
- Wang D (2018) Sparsity guided empirical wavelet transform for fault diagnosis of rolling element bearings. *Mech Syst Signal Process* 101:292–308
- Chen J (2016) Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. *Renewable Energy* 89:80–92
- Jiang Y, Zhu H, Li Z (2016) A new compound faults detection method for rolling bearings based on empirical wavelet transform and chaotic oscillator. *Chaos, Solitons Fractals* 89:8–19
- Hu Y (2017) An adaptive and tacholess order analysis method based on enhanced empirical wavelet transform for fault detection of bearings with varying speeds. *J Sound Vib* 409:241–255

34. Hu Y, Li F, Li H (2017) An enhanced empirical wavelet transform for noisy and non-stationary signal processing. *Digit Signal Process* 60:220–229
35. Zhang K, Xu Y, Liao Z, Song L, Chen P (2021) A novel Fast Entrogram and its applications in rolling bearing fault diagnosis. *Mech Syst Signal Pr* 154(1):107582
36. Xu Y, Deng Y, Zhao J (2020) A novel rolling bearing fault diagnosis method based on empirical wavelet transform and spectral trend. *IEEE Trans Instrum Meas* 69(6):2891–2904
37. Huang X, Wen G, Liang L (2019) Frequency phase space empirical wavelet transform for rolling bearings fault diagnosis. *IEEE Access* 7:86306–86318
38. Zhao H, Zuo S, Hou M (2018) A novel adaptive signal processing method based on enhanced empirical wavelet transform technology. *Sensors* 18(10):3323
39. Hsueh YM, Ittangihal VR, Wu W (2019) Fault diagnosis system for induction motors by CNN using empirical wavelet transform. *Symmetry* 11(10):1212
40. Varanis M, Pederiva R (2018) Statements on wavelet packet energy-entropy signatures and filter influence in fault diagnosis of induction motor in non-stationary operations. *J Braz. Soc Mech Sci Eng* 40(2):98
41. Ou Y, He S, Hu C, Bao J (2020) Research on rolling bearing fault diagnosis using improved majorization-minimization-based total variation and empirical wavelet transform. *Shock and Vibration* 2020
42. Liu W, Chen W (2019) Recent advancements in empirical wavelet transform and its applications. *IEEE Access* 7:103770–103780

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