



Fault feature extraction of rolling element bearing based on EVMD

Danchen Zhu¹ · Guoqiang Liu¹ · Wei He² · Bolong Yin¹

Received: 10 April 2021 / Accepted: 16 November 2021 / Published online: 29 November 2021
© The Brazilian Society of Mechanical Sciences and Engineering 2021

Abstract

Aiming at the problem that the bearing fault signal is weak and usually interfered by the strong background noise, which makes the fault feature extraction very difficult, an enhanced variational mode decomposition (EVMD) technique is proposed. First, the autoregressive (AR) model was utilized to eliminate the stationary components in the signal in advance to reduce the noise interference and the maximum kurtosis of the residual signal was set as the target. Second, the maximum frequency-domain correlated kurtosis was adopted as the fitness value, and the decomposition modes K and quadratic penalty factor α in the VMD approach were adaptively selected by the whale optimization algorithm. Third, the reconstruction signal was acquired, then the enhanced envelope spectrum was employed to weaken the interference of irrelevant frequency components and the fault features of rolling element bearing could be extracted accurately. The results of simulation and experimental analysis show that the proposed algorithm can significantly reduce the noise interference and avoid the blindness selection of VMD parameters. The comparison with fix-parameter VMD and fast kurtogram approaches shows that the proposed technique can improve the effectiveness of defect signature extraction, which has a certain value for engineering application.

Keywords AR model · Whale optimization algorithm · Variational mode decomposition · Rolling element bearing · Fault feature extraction

1 Introduction

Rolling element bearing (REB) is an important part of the rotating machinery, which plays an important role in ensuring the performance of the whole machinery. However, REB is usually working under high-speed, high-pressure and heavy-load condition, and the probability of REB's failure is very high, which greatly threatens the safe and stable operation of the machinery. Therefore, it is of great urgency and practical significance to carry out researches on fault diagnosis approaches of REB. Recently, vibration-based monitoring techniques have become one of the most commonly used methods for the fault diagnosis of REB due to

their superior characteristics, but there still some problems need to be solved.

- (1) The fault features of REB are relatively weak compared with other irrelevant noise components. How to suppress the influence of background noise and extract the weak fault features accurately?
- (2) How to select the appropriate measuring points based on the structure characteristics and actual working conditions of the machinery to reduce the influence of the vibration caused by machinery itself?
- (3) Considering that some drawbacks still exist in the already proposed fault diagnosis approaches, how to improve the effectiveness of these approaches?

To address the above problems, many researches have been conducted and various achievements have been obtained. Resonance demodulation is widely used due to its convenient application. However, the proper selection of resonance frequency band needs to determine the central frequency and bandwidth by human experience in advance, which may be easily affected. Therefore, in the classic fast kurtogram [1] proposed by Antoni, the optimal frequency

Technical Editor: Samuel da Silva.

✉ Danchen Zhu
bryantzdc@sina.com

¹ Naval Petty Officer Academy, Bengbu 233012, China

² Equipment Condition Detection Institution, Guangzhou Mechanical Engineering Research Institute Co., Ltd., Guangzhou 510700, China

band is quickly determined based on the principle of maximum spectral kurtosis. On this basis, to improve the effectiveness of this approach, many improved methods have been proposed; for example, the Infogram [2], the correlated kurtosis-based kurtogram [3], the Autogram [4], the empirical wavelet transform-based fast kurtogram [5, 6], the subband average kurtogram [7] and so on. These methods further improve the accuracy and effectiveness of resonance frequency band selection.

To suppress the influence of transmission path, the blind deconvolution algorithms have been widely applied and developed rapidly in recent years, whose purpose is acquiring the approximate original fault impulse signal based on the maximization of the characteristic index, including the minimum entropy deconvolution (MED) [8], the maximum correlated kurtosis deconvolution (MCKD) [9, 10], the multipoint optimal minimum entropy deconvolution adjusted (MOMEDA) [11, 12] and the maximum second-order cyclostationarity blind deconvolution (CYCBD) [13]. These methods have their own characteristics, but the analysis results are also restricted by some parameters, such as the shift number M in MCKD and the selection of cycle frequency in CYCBD. Moreover, the results of blind deconvolution algorithms are significantly affected by the filter length, and to solve this problem, many studies are conducted [14–16] and the effectiveness of blind deconvolution algorithms is enhanced but is still not perfect.

The signal decomposition techniques are employed to eliminate the irrelevant components in the original signal. As a classic signal decomposition technique, empirical mode decomposition (EMD) has been widely utilized in fault diagnosis of REB [17, 18]. EMD can decompose the vibration signal into a series of intrinsic mode functions (IMFs), and the reconstruction signal is obtained by selecting the appropriate IMFs, which can suppress the irrelevant noise components and improve the signal-to-noise ratio. However, end effects and mode mixing seriously influence the performance of EMD. To address these problems, the ensemble empirical mode decomposition (EEMD) [19, 20] is proposed which solves the problem of mode mixing, but its performance is also affected by the number of ensemble and the variance of the added white noise. Moreover, the analysis efficiency is lower than the EMD. To further enhance the performance of EEMD, the complementary ensemble empirical mode decomposition (CEEMD) [21] and the complementary ensemble empirical mode decomposition with adaptive noise (CEEMDAN) [22] are also constructed. Variational mode decomposition (VMD) is a new signal decomposition technique, which is different from the recursive decomposition pattern of EMD [23]. VMD has the characteristics of high precision and fast convergence speed, which is widely employed in REB's fault diagnosis in recent years [24]. To improve

the effectiveness of VMD, researchers have carried out many discussions, whose focus are on two aspects: one is the optimization selection of quadratic penalty term α , the decomposition level K , and the improvement of signal reconstruction approach [25–27], the other is to optimize the decomposition steps of VMD [28, 29]. Though numerous researches have been conducted on VMD, how to determine the parameters more effectively according to the signal characteristics is still a difficult problem.

As a result, in order to solve the problem of parameters' selection in VMD and extract the fault features of REB under strong background noise, an enhanced variational mode decomposition (EVMD) technique is proposed. Firstly, the AR model is adopted to eliminate the stationary components in the original signal, and the maximum kurtosis value of the residual signal is set as the target to determine the order of AR model accurately. Secondly, as the noise still exists in the residual signal, the whale optimization algorithm (WOA) and frequency-domain correlated kurtosis (FDCK) are introduced to achieve the optimal selection of VMD's parameters; then, the signal with more obvious fault characteristics is extracted. Finally, based on the fault characteristic frequency (FCF) and envelope spectrum's characteristics of REB, an enhanced envelope spectrum is proposed to extract the fault signatures of REB accurately and suppress the interference of irrelevant frequency components. The main work of this paper is described as follows.

- (1) The AR model is utilized to eliminate the stationary components in the original signal, and the defect features are enhanced.
- (2) The whale optimization-based VMD is employed, and the FDCK is used as the fitness function, which improve the effects of signal decomposition.
- (3) An enhanced envelope spectrum (EES) is used to highlight the fault feature components.

The remainder of this paper is organized as follows: The theoretical background of relevant techniques employed in this study is reviewed in Sect. 2, including the principle of AR model, the steps of whale optimized VMD and the calculation approach of EES. The proposed technique is summarized in Sect. 3. Section 4 and Sect. 5 exhibit the results when the simulation and experimental signals are used to demonstrate the effectiveness of the proposed approach, respectively. Conclusions are presented in Sect. 6.

2 The basic principles

2.1 The AR model

The output of the AR model is related to the weighted sum of the input and the past n outputs. Hence, the signal can be predicted linearly by AR model. Suppose a signal x , whose length and mean value are N and zero, respectively. Then, the p order AR model can be expressed as:

$$y_t = \sum_{i=1}^n a_i x_{t-i} + \eta_t \tag{1}$$

where a_i represents the weighted parameters of autoregression, η_t denotes the zero mean white noise with the variance of σ^2 , which can also be called residual. Residual mainly represents the error between the real and the predicted values; for the vibration signal of REB, the residual mainly contains non-stationary components (fault feature components) which cannot be obtained by linear prediction.

The least square method can be utilized to estimate the parameters a_i in real application, and the steps are:

$$Y = Xa + \eta \tag{2}$$

where $Y = [x_{n+1} \ x_{n+2} \ \dots \ x_N]^T$, $a = [a_1 \ a_2 \ \dots \ a_n]^T$,

$$\eta = [\eta_{n+1} \ \eta_{n+2} \ \dots \ \eta_N]^T, X = \begin{bmatrix} x_n & x_{n-1} & \dots & x_1 \\ x_{n+1} & x_{n+2} & \dots & x_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-1} & x_{N-1} & \dots & x_{N-n} \end{bmatrix}$$

Then, the least square estimation can be expressed as:

$$Y = Xa + \eta \tag{3}$$

According to Eq. (1), besides determining the parameters of AR model, the order of model may also have a great influence. Hence, it is necessary to select the appropriate order of the AR model to remove the stationary components in the signal more effectively.

Many approaches can determine the appropriate order, as AR model is adopted to eliminate the stationary components in this paper, maximum kurtosis of residual signal is selected as the target. Therefore, the process of using the AR model can be described as following steps:

- (1) The search range of the order is determined as 1 to 100, AR models with different orders are established and the residual signals are acquired.
- (2) Kurtosis of all the residual signals is calculated.
- (3) The optimal order of AR model is determined based on the maximum of kurtosis value.

2.2 Variational mode decomposition

VMD is a newly proposed signal decomposition technique, which can adaptively decompose the signal into K mode components u_k , called intrinsic mode function (IMF). The principle of VMD is the process of constructing and solving a constrained variational problems, which is described as follows [25]:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\partial(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \tag{4}$$

$$\sum_k u_k = f \tag{5}$$

where $\omega_k = \{\omega_1, \omega_2, \dots, \omega_k\}$ denote the central frequency of each IMFs.

To solve the above problem, Lagrangian multiplier $\lambda(t)$ and quadratic penalty factor α are introduced, which are employed to ensure reconstruction accuracy and the performance of constraints, respectively. Then, the augmented Lagrangian is acquired and described as:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \tag{6}$$

The alternating direction method of multiplier (ADMM) is introduced to solve Eq. (6). On the basis of determining the number of decomposition modes K , the sub-modes \hat{u}_k^1 , its central frequency ω_k^1 and Lagrangian multiplier λ^1 should be initialized. The update principle of each sub-modes \hat{u}_k and its central frequency ω_k are described in Eqs. (7) and (8) [25]:

$$\hat{u}_k^{n+1} = \frac{\hat{f}(\omega) \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{7}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \tag{8}$$

The Lagrangian multiplier can also be updated.

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \tag{9}$$

where τ denotes the noise tolerance. The iteration of VMD is stopped when Eq. (10) is satisfied.

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (10)$$

where ε represents the convergence error. Based on the calculation process of VMD, four parameters K , α , τ and ε may influence the decomposition results and should be determined in advance, which brings a great challenge for the appropriate selection. Researches show that τ and ε have less effect than the other two parameters; hence, they are set as default, and the determination of the first two parameters is optimized in this paper.

2.3 Whale optimization algorithm

Whale optimization algorithm [30] is presented to simulate the hunting behavior of humpback whales; the position of each whale represents a candidate solution. In the algorithm, three behaviors of whales during hunting are simulated, which are encircling prey, search for prey and spiral updating position, respectively. These three behaviors can be expressed as:

(1) Encircling prey

The position $x(t+1)$ of the candidate solution in the $(t+1)$ th iteration can be updated by Eqs. (11) and (12); $X^*(t)$ represents the optimal position in the t th iteration:

$$D = |CX^*(t) - X(t)| \quad (11)$$

$$X(t+1) = |X^*(t) - AD| \quad (12)$$

$$A = |2ar - a| \quad (13)$$

$$C = 2r \quad (14)$$

where a decreases linearly from 2 to 0 according to the number of iterations, and r is randomly selected between 0 and 1.

(2) Search for prey

Search for prey is very similar to encircling prey, the difference is that search for prey no longer uses the optimal solution $X^*(t)$ of the current iteration, but randomly selects an candidate solution $X_{\text{random}}(t)$:

$$D = |CX_{\text{random}}(t) - X(t)| \quad (15)$$

$$X(t+1) = |X_{\text{random}}(t) - AD| \quad (16)$$

(3) Spiral updating position

Spiral updating position is employed to simulate the helix-shaped movement of whales based on the position of the whale and prey, which can be expressed as:

$$\begin{cases} X(t+1) = D' e^{bl} \cos(2\pi l) + X^*(t) \\ D' = |X^*(t) - X(t)| \end{cases} \quad (17)$$

where D' denotes the distance between the optimal position and any candidate position in the t th iteration. b is a constant value, and l is randomly selected between -1 and 1 .

Based on the previous introduction, the process of whale optimization can be described as follows:

- (1) The position of whale population X_i ($i = 1, 2, 3, \dots, N$) is initialized, The fitness value of each candidate positions is calculated, and $X^*(t)$ is the position corresponding to the optimal fitness value.
- (2) The values of parameters a , A , C , l and p are determined. In this paper, b is set to be 1.
- (3) When $p < 0.5$, if $|A| < 1$, Eqs. (11) and (12) are utilized to update the current position of the whale; else, Eqs. (15) and (16) are utilized to update the current position of the whale. When, $0.5 < p < 1$, the current position of the whale is updated according to Eq. (17).
- (4) a , A and C are updated based on the new position of the whale, and the fitness value of each candidate position is also recalculated; then, the best position $X^*(t)$ is determined again.
- (5) Repeat steps (3) and (4) until the maximum iteration is reached.

2.4 Enhanced envelope spectrum

Since the influence of strong background noise, the denoised signal obtained by using the AR model and VMD approach may still contain strong interference components. Hence, various irrelevant frequency components may still exist in its envelope spectrum, which affects the extraction and judgment of the fault signatures. To address the problem, as the interval of the fault feature components in the envelope spectrum is usually the FCF, some improvements are conducted to the classic envelope spectrum and the EES is proposed.

- (1) For discrete signal x , its envelope spectrum is expressed as $Ev(x)$, whose frequency range is 1 to $fs/2$, fs denotes the sampling frequency of signal x .
- (2) The noise level of the whole envelope spectrum is evaluated, and the spectral lines' amplitude at the FCF

components is also employed, and the noise level of $Ev(x)$ can be expressed as:

$$MEv(x) = \text{mean}[Ev(x)] \tag{18}$$

where mean denotes the average value.

- (3) The amplitude of spectral line at any frequency f_1 is recalculated, and the EES is acquired.

$$EEv(x)_{f_1} = \frac{Ev(x)_{f_1} * Ev(x)_{f_1+f_d}}{MEv(x)^2} \tag{19}$$

where $EEv(x)_{f_1}$ and $Ev(x)_{f_1}$ denote the amplitude of spectral line at frequency f_1 in envelope spectrum and EES, respectively, and f_d denotes the FCF. Therefore, by using Eq. (19), if f_1 is the FCF in the envelope spectrum of signal x , its amplitude will significantly be enhanced in the EES. Otherwise, if f_1 represents the irrelevant components, its amplitude in the EES will not be obviously improved. As a result, the enhanced envelope spectrum can further highlight the defect signatures in the envelope spectrum and reduce the interference of irrelevant frequency components.

3 Whole structure of the EVMD approach

3.1 The fitness function

As the characteristics of bearing fault signal are shown as prominent amplitude at the defect characteristic frequency components in frequency domain, based on the advantage of correlated kurtosis, the frequency-domain correlated kurtosis is utilized as the fitness function of whale optimization-based VMD approach. When the envelope spectrum of the decomposed IMFs has obvious fault features, the FDCK is larger. It should be noted that to facilitate the calculation, the envelope spectrum is obtained by using Hilbert transform and fast Fourier transform directly, which avoids the problem of selecting the parameters of band-pass filter in advance. The FDCK can be expressed as:

$$ECK_M(T) = \frac{\sum_{n=1}^N \left(\prod_{m=0}^M E(x)_{n-mT} \right)^2}{\left(\sum_{n=1}^N E(x)_n^2 \right)^{M+1}} \tag{20}$$

where $E(x)_n$ denotes the envelope of the signal x_n , T is the number of signal points corresponding to the FCF, N and M represent the length of the signal and shift number, respectively. In this paper, M is determined as 3.

3.2 The principle to determine the effective IMF

For each IMF u_k decomposed by the VMD approach, the FDCK $FDCK_i$ ($i = 1, 2, 3, \dots, K$) is calculated, and then, the average value of all the IMFs' $FDCK_i$ is acquired.

$$ME = \frac{1}{K} \sum_{i=1}^K FDCK_i \tag{21}$$

where ME denotes the mean value of all $FDCK_i$. If $FDCK_i > ME$, the corresponding IMF is considered as the effective IMF; else, the IMF is abandoned, and the final reconstruction signal is the sum of all effective IMFs.

3.3 The process of the fault feature extraction of REB using the EVMD

According to the theory described above, this paper presents the implementation of a new fault feature extraction technique for REB named EVMD. The flowchart of this technique is exhibited in Fig. 1, and the detailed steps are described below.

Step 1. The vibration signal of REB is acquired, the AR model is utilized to preprocess the signal and the stationary components are eliminated to reduce the interference of irrelevant components.

Step 2. The search ranges of K and α are set.

Step 3. The maximized $FDCK$ of a single IMF is selected as the target to optimize the parameters of VMD by WOA; the optimal parameters and the optimal decomposition results are acquired.

Step 4. The enhanced envelope spectrum of the reconstruction signal is obtained.

Step 5. Based on the enhanced envelope spectrum, the fault signatures and the fault type of REB are determined.

4 Simulation verification

The effectiveness of the proposed technique is demonstrated using the simulation signal. The construction of the simulation signal simulates the bearing with inner race defect, where periodic impacts, harmonic components and white noise interference are taken into consideration. The simulation signal is expressed as follows:

$$\begin{cases} x(t) = \sum_{i=1}^N A_i s(t - iT - t_i) + B(t) + n(t) \\ A_i = 2 \cos(2\pi f_r t + \pi/2) + 0.5 \\ B(t) = 0.5 \cos(2\pi f_m t + \pi/2) \\ s(t) = e^{-2\pi f_n t} \sin(2\pi f_n t) \end{cases} \tag{22}$$

Fig. 1 Flowchart of the proposed technique

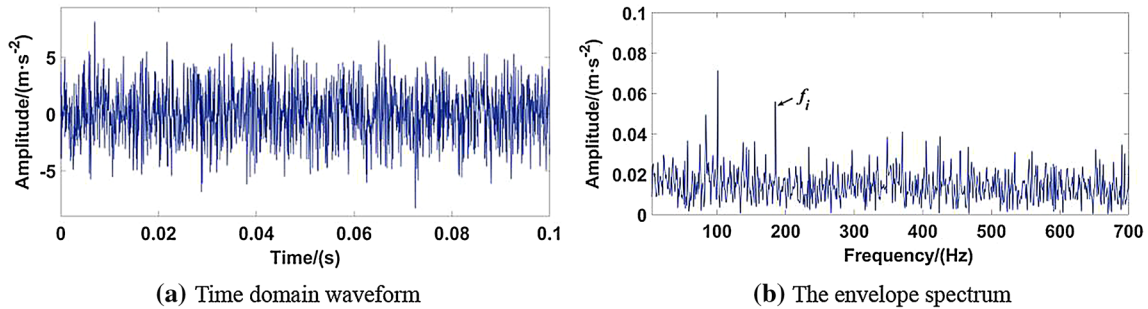
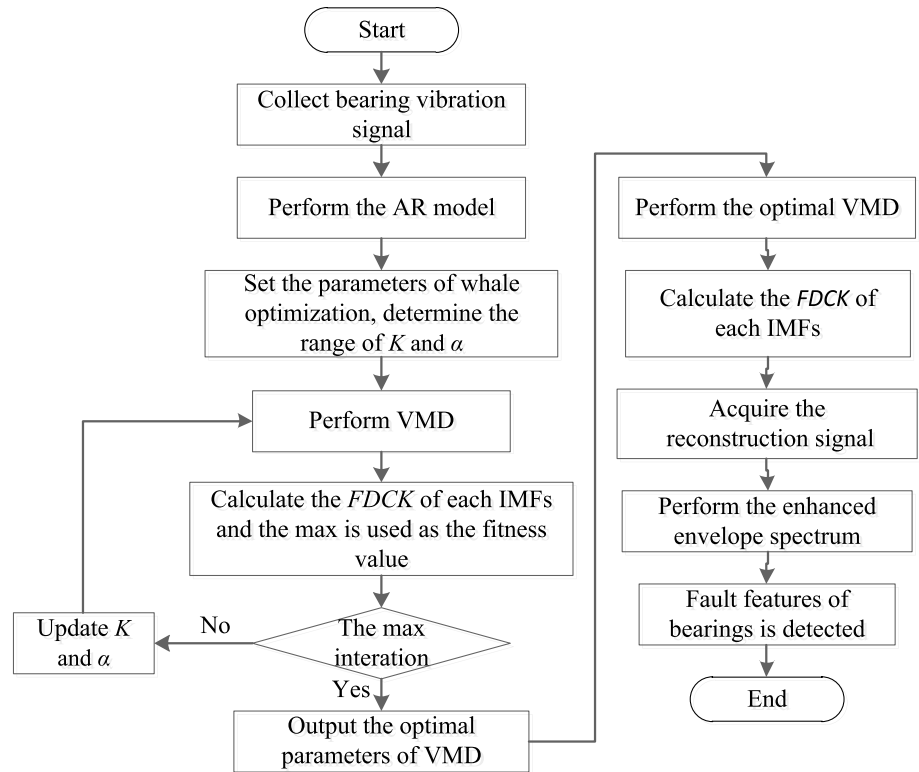


Fig. 2 The simulation signal

where $f_r = 42$ Hz represents the rotating frequency and $A_i = 1/f_r$ is adopted to simulate the amplitude modulation, $f_n = 3200$ and $r = 0.05$ denote the natural frequency of the system and the damping coefficient, respectively. The interval of two adjacent impulses is set as $T = 1/185$ s; hence, the FCF is 185 Hz. $t_i = 0.01 T \sim 0.02 T$ simulates the delay due to the slippage of rolling elements. $B(t)$ denotes the harmonic which is used to simulate the interference components existing in the signal and $f_m = 100$. $n(t)$ is adopted to represent the random noise and the signal-to-noise ratio is set to be -10db in the simulation. The sampling frequency is 16384 Hz, and the length of signal is 20 s in this simulation.

The signal $x(t)$ and its envelope spectrum are presented in Fig. 2. It can be observed from Fig. 2a that the periodic components are submerged by strong background noise, which

can hardly be distinguished. Meanwhile, only the FCF is presented in the envelope spectrum which means to achieve good results only by envelope demodulation under strong noise interference is very difficult.

Subsequently, the simulation signal is processed by the proposed method. First, AR model is employed to remove the stationary components of simulation signal in advance. According to Fig. 3a, the maximum kurtosis value is achieved when the order of AR model is determined as 51, the residual signal is obtained which is exhibited in Fig. 3b.

Then, the whale optimization-based VMD approach is utilized to process the residual signal. In order to ensure the complete decomposition of the signal and the efficiency of the calculation, the search ranges of K and α are set as 2 to 10 and 100 to 5000, respectively. The population size

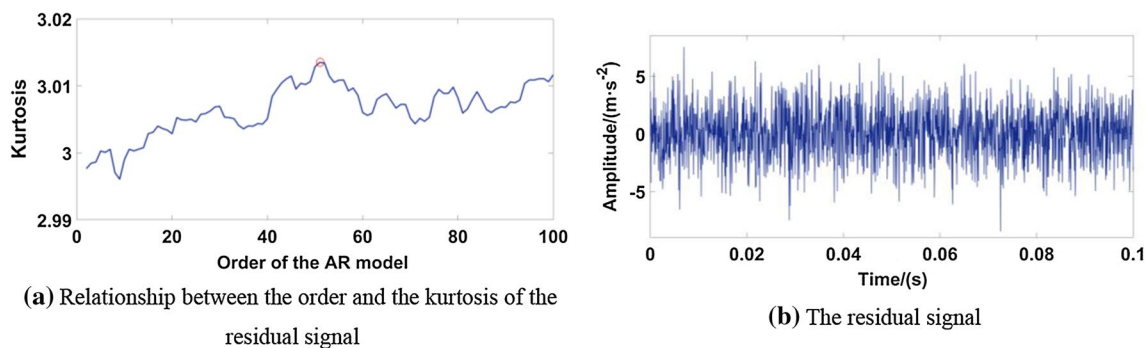


Fig. 3 Analyzed results of the simulation signal using the AR model

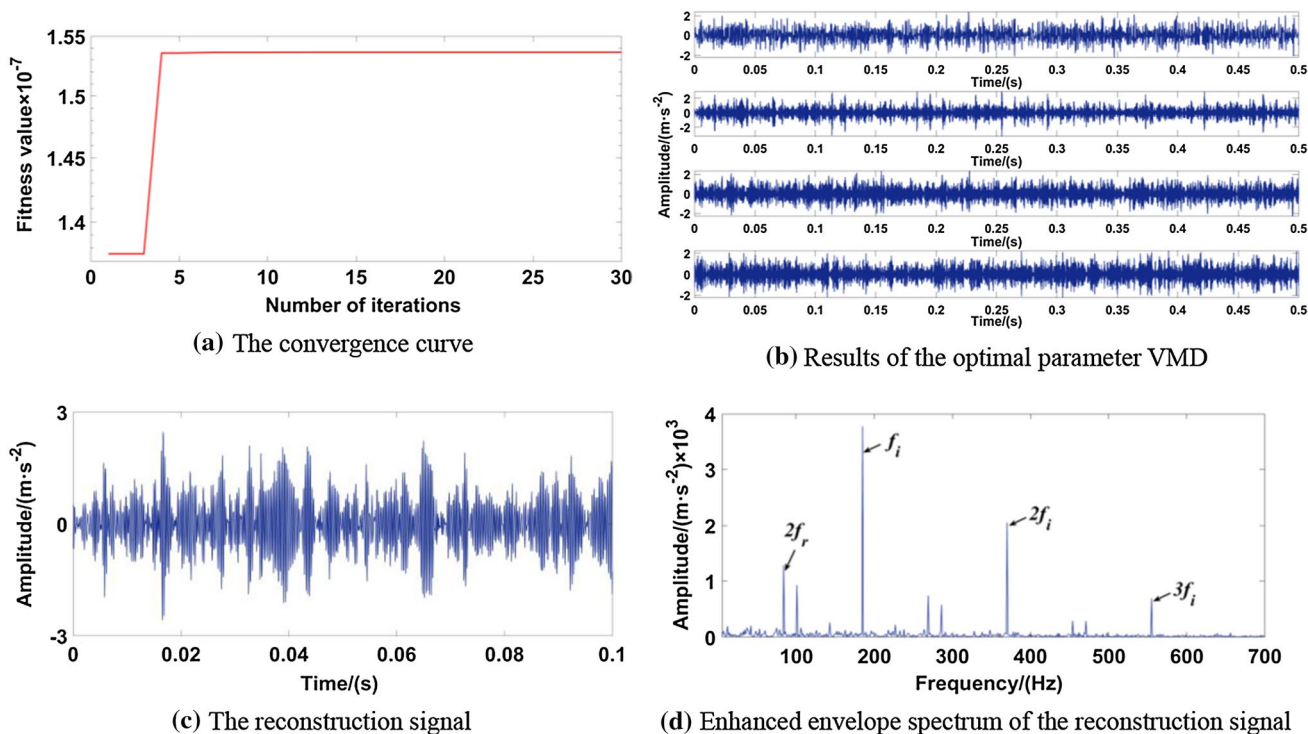


Fig. 4 Analyzed results of the proposed technique

and maximum number of iteration are determined as 20 and 30. Other parameters are selected as default. In the subsequent analysis, the parameters of the optimization algorithm remain unchanged. Based on the convergence curve in Fig. 4a, the optimal fitness value 1.54×10^{-7} is acquired after four iterations, while the optimal position is [1068, 4]; then, $\alpha = 1068$ and $K = 4$ are adopted as the optimal parameters for VMD, the decomposition results are exhibited in Fig. 4b. Based on the principle in Sect. 2.2, the FDCK of these four IMFs is calculated, which are 4.55×10^{-13} , 1.54×10^{-7} , 1.34×10^{-12} and 4.17×10^{-12} ; their average value is also obtained. By comparison, the IMF2 is determined as the effective IMF, which is then applied for the signal

reconstruction, and the reconstruction signal is depicted in Fig. 4c. It can be found that the interference components have been significantly removed in the reconstruction signal, and the fault impact components can be recognized. The enhanced envelope of the reconstruction signal in Fig. 4d shows the frequency range from 0 to 700 Hz, where rotating frequency $2f_r$, the FCF f_i and two of its harmonic components are prominent, the fault signatures of bearing defect are extracted accurately. Therefore, it can conclude that the weak fault feature of REB can be accurately extracted under the interference of strong background noise by the proposed technique.

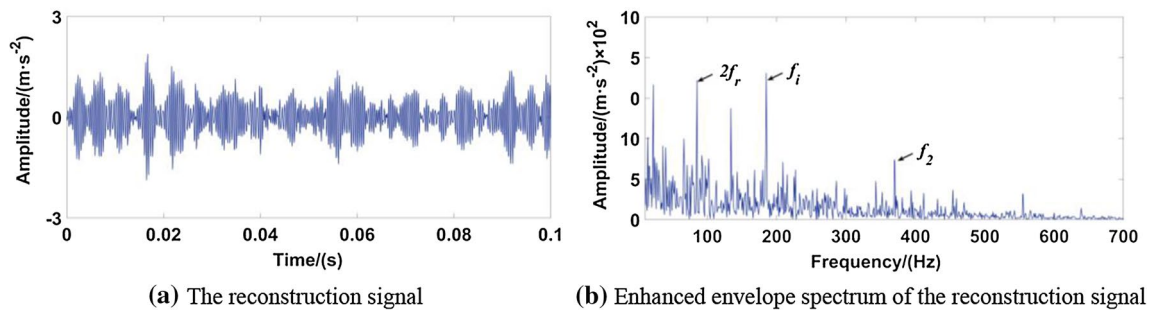


Fig. 5 Analyzed results of the simulation signal using fix-parameter VMD

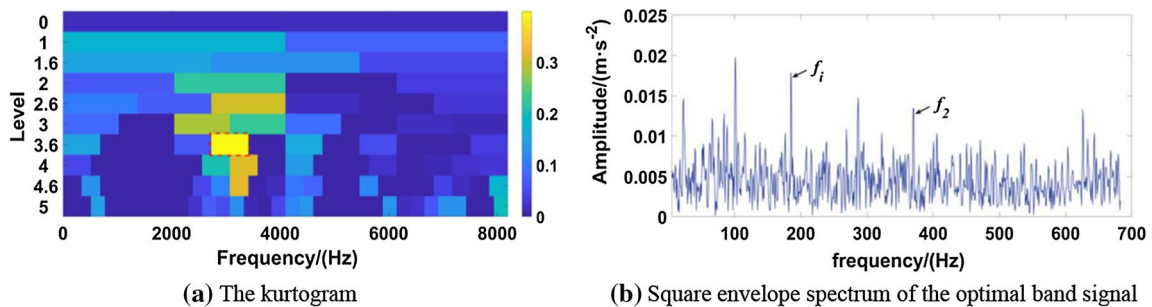


Fig. 6 Analyzed results of the simulation signal using the fast kurtogram

To highlight the importance of parameters' optimization, the parameters of VMD are determined as [3650, 5] randomly in advance, and the fix-parameter VMD is then applied to the residual signal displayed in Fig. 3b, the enhanced envelope spectrum is also calculated and the results are depicted in Fig. 5. Although the noise components are also suppressed according to the reconstruction signal in Fig. 5a, only the rotating frequency f_r , the FCF f_i and one of its harmonic $2f_i$ are able to be detected. What's more, the amplitude of the spectral line at $2f_i$ is not obvious, the noise components are more obvious than those in Fig. 4d, which verifies the necessity of selecting the optimal parameters for VMD.

To further show the effectiveness of this algorithm, the classic fast kurtogram method [1] is employed. The central frequency and bandwidth of the optimal analysis frequency band are 3072 Hz and 682 Hz according to the kurtogram exhibited in Fig. 6a. The envelope spectrum of the optimal band signal is depicted in Fig. 6b, where only the FCF and one of its harmonic are successfully extracted but not very prominent, and the influence of noise is very obvious. The phenomenon describes above further to demonstrate the effectiveness of the proposed technique in noise suppressing and weak fault features extracting.

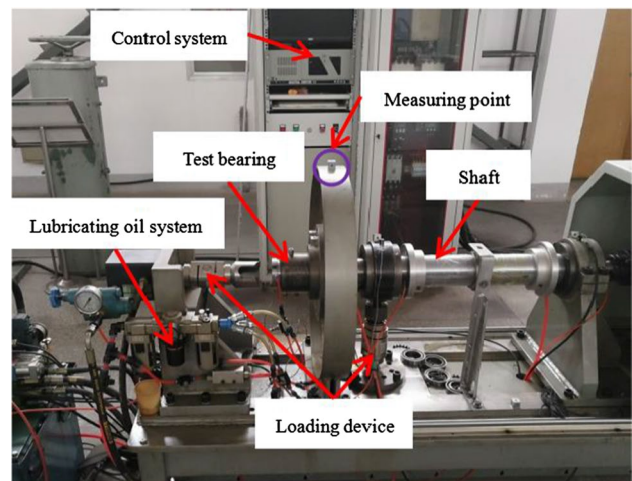


Fig. 7 The test rig

5 Experimental verification

The bearing vibration signal collected using a test rig is adopted to further demonstrate the effectiveness of the proposed approach for the fault feature extraction of REB. The test rig exhibited in Fig. 7 consists of bearing support structure, main shaft, experimental bearing, lubricating oil system, servo-driven motor, radial loading device, axial loading

Table 1 Parameters of the test bearing

Inner diameter/mm	Outer diameter/mm	Pitch diameter D/mm	Ball diameter d/mm	Number of ball Z/↑	Contact angle α/°
50	80	65	9	13	0

device and control system, which is able to simulate the common working state along with the defect types of REB.

For real test, the bearing may be arranged inside the machine, to avoid damaging the whole structure, and the measuring points for vibration capturing can only be selected near the bearing locations, but not on the bearing house directly. There may exist a complex transmission path between the defective bearing and the measuring point, which contains multiple connection interfaces and influenced by multistructures of the machine. And the measured signal can also be more complex, which brings great challenge for the accurate fault detection of REB. Hence, according to the above description, to simulate the complex transmission path, the measuring points are selected on the case instead of at the center position of the bearing support structure, shown in Fig. 7.

The experimental bearing type is 6010 with a defect on the inner race (outer race) generated by laser wire-electrode

cutting (the width and depth of the defect are both 0.2 mm). Table 1 presents the parameters of the experimental bearing. During the test, the rotating speed is set as 3000r/min ($f_r = 50$ Hz) and the sampling frequency is 16384 Hz, the radial and axial loads are 1KN and 2KN, respectively. The FCF of inner (f_i) and outer (f_o) race defect can be expressed as:

$$f_i = \frac{1}{2}Z\left(1 + \frac{d}{D} \cos a\right)f_r \tag{23}$$

$$f_o = \frac{1}{2}Z\left(1 - \frac{d}{D} \cos a\right)f_r \tag{24}$$

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

5.1 The REB with inner race defect

As presented in Fig. 8a, the repetitive transient impulses are not able to be detected in the experimental signal and the influence of noise is very obvious. In the envelope spectrum given in Fig. 8b, the rotating frequency f_r and its harmonic $2f_r$ can be detected. The FCF f_i can be barely distinguished and its harmonic components can be hardly observed;

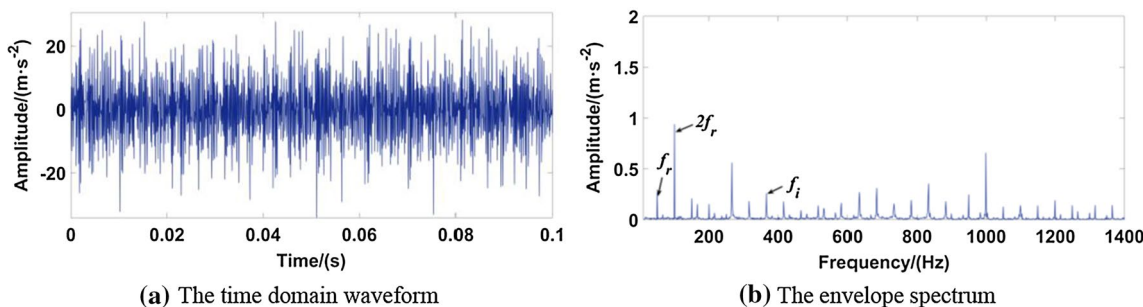


Fig. 8 The experimental signal with inner race defect

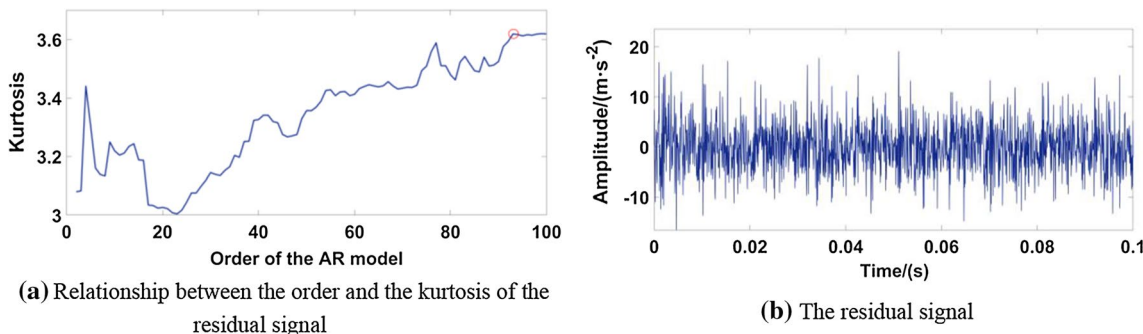


Fig. 9 Analyzed results of the measured inner race fault signal using the AR model

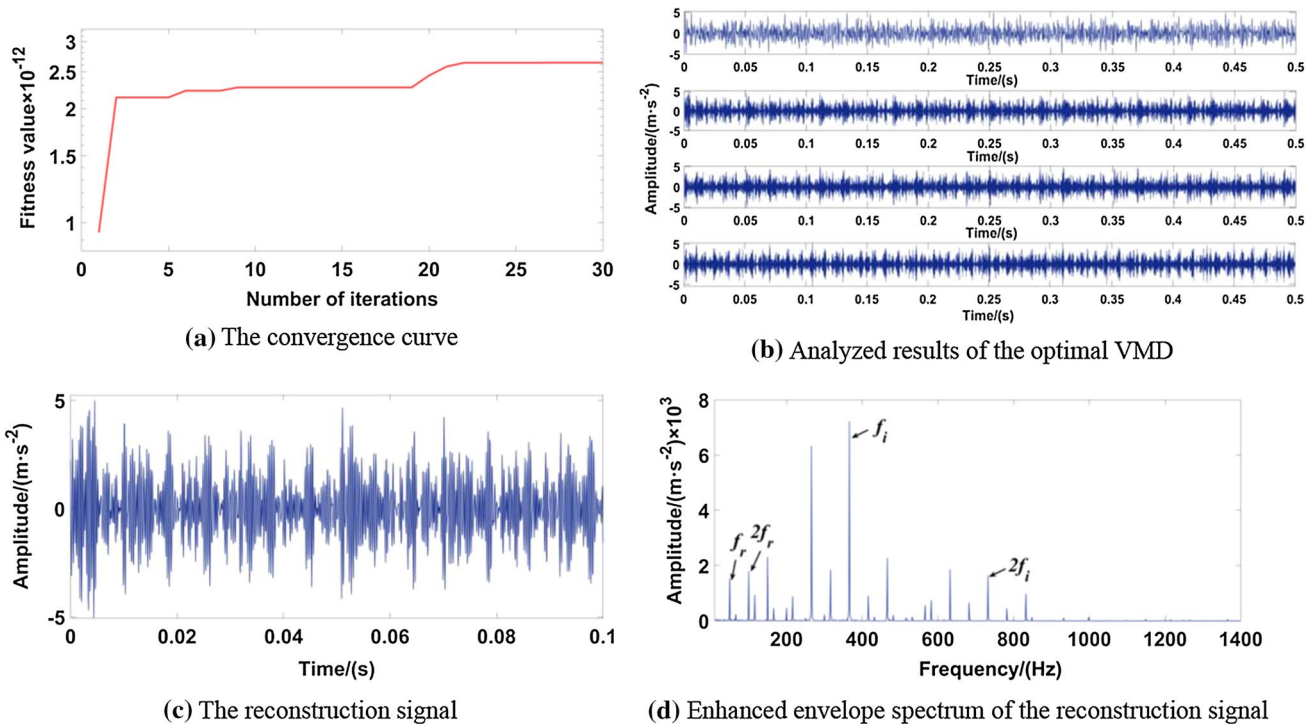


Fig. 10 Analyzed results of the measured inner race fault signal using the proposed technique

meanwhile, some irrelevant frequency components are also prominent.

Then, the technique proposed in this paper is applied to the experimental signal. Based on the relationship displayed in Fig. 9a, the optimal order of AR model is determined as 99, and the residual signal in Fig. 9b exhibits the suppression of irrelevant components compared with the original signal but not enough. Therefore, the optimized VMD approach is utilized to further analyze the residual signal, and the results are depicted in Fig. 10.

According to the convergence curve displayed in Fig. 10a, the maximum fitness value 2.64×10^{-12} is obtained after 22 iterations, and the optimal parameters for VMD are determined as $K=4$ and $\alpha=1011$, respectively. The

decomposition results using the optimal VMD are shown in Fig. 10b, and the FDCK of the four IMFs are calculated, which are 6.44×10^{-14} , 2.88×10^{-15} , 3.11×10^{-15} and 7.51×10^{-14} . Based on the average of the four FDCK values, the IMF1 and IMF4 are determined as the effective IMF, and the reconstruction signal is depicted in Fig. 10c where the fault impact components are able to be observed. The above phenomenon indicates that the optimized VMD approach can further enhance the fault signatures, while the irrelevant components are suppressed. The enhanced envelope spectrum of the reconstruction signal is displayed in Fig. 10d, where abundant fault signatures are successfully extracted. The rotating frequency components, FCF, along with one harmonic are prominent. What's more, the

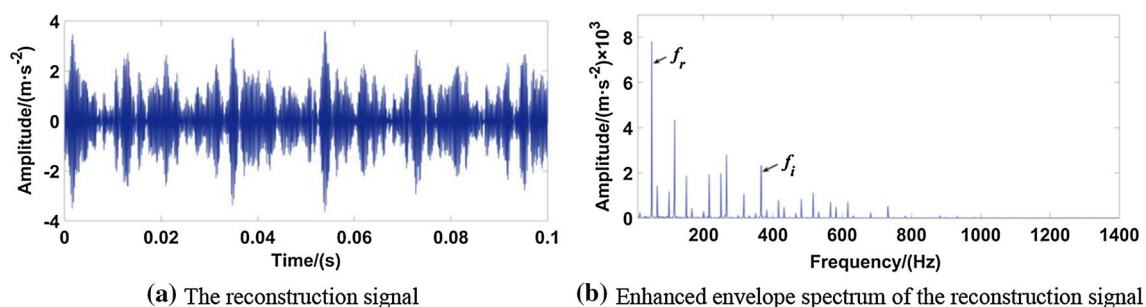


Fig. 11 Analyzed results of measured inner race fault signal using the fix-parameter VMD

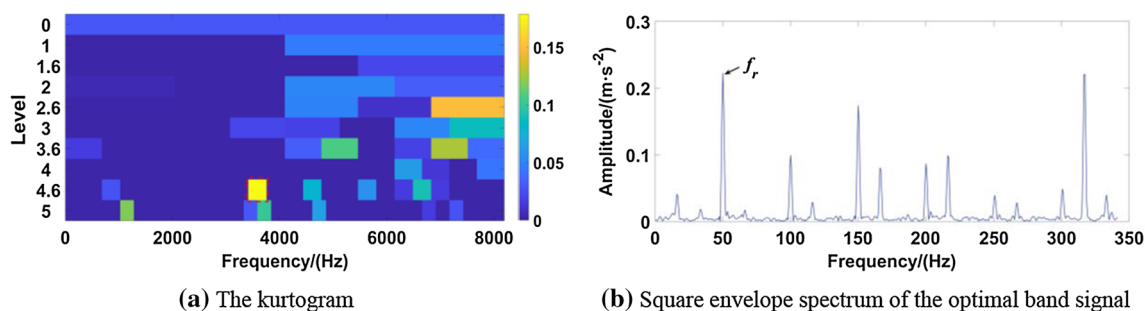


Fig. 12 Analyzed results of measured inner race fault signal using the fast kurtogram

sidebands modulated by rotating frequency around the FCF components can also be distinguished, based on the features mentioned above; it can conclude that the REB has inner race defect. The effectiveness of the proposed technique in fault diagnosis of REB is also verified when compared with the envelope spectrum shown in Fig. 8b.

Similar to the simulation analysis, two comparisons are conducted. The analyzed result of fix-parameter VMD approach is displayed in Fig. 11, the parameters are chosen in advance as $K=6$ and $\alpha=2650$, respectively. Although the impact components can also be identified in Fig. 11a, it is not as obvious as those in Fig. 10c. Meanwhile, in the envelope spectrum of the reconstruction signal, only the rotating frequency f_r and the FCF f_i are extracted successfully, and the amplitude of these spectral lines is lower than those in Fig. 10d, and fault features are not abundant enough compared with the result acquired by the proposed technique. The above description indicates that the random selection of VMD's parameters may not achieve the optimal defect signature extraction.

Besides, the fast kurtogram is employed and results are depicted in Fig. 12. Based on the kurtogram in Fig. 12a, the central frequency and bandwidth of the optimal frequency band are 3584 Hz and 341 Hz, the square envelope spectrum of the optimal band signal is exhibited in Fig. 12b, where useful fault signatures are unable to be identified except for the rotating frequency f_r , in which amplitude is very small.

The phenomenon described above indicates that it is difficult for the fast kurtogram to accurately determine the location of the optimal analysis frequency band and extract fault signatures effectively under strong background noise condition. By comparison, the effectiveness of the proposed technique is further demonstrated.

Since the optimization algorithm is employed to improve the effectiveness of the fault feature extraction, the proposed approach is more time-consuming, and the fast kurtogram method has high calculation efficiency.

5.2 The REB with outer race defect

To demonstrate the applicability of the technique, the measured signal with outer race defect is utilized for analysis, and the results are exhibited in Fig. 13. In Fig. 13a, the periodic impact components are unable to be observed due to the influence of strong background noise. The envelope spectrum in Fig. 13b shows a complex feature map where only the FCF and one of its harmonic are able to observe, which means using the envelope analysis directly cannot achieve ideal performance. Therefore, the proposed technique is applied to the signal.

To ensure the maximum of kurtosis, the optimal order of AR model is selected as 84 according to the curve exhibited in Fig. 14a, and the stationary components in the original signal are eliminated. Figure 14b displays the residual signal

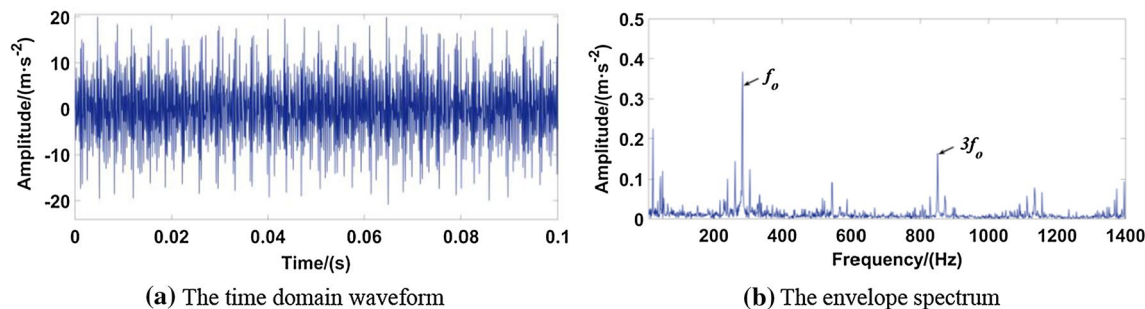


Fig. 13 The experimental signal with outer race defect

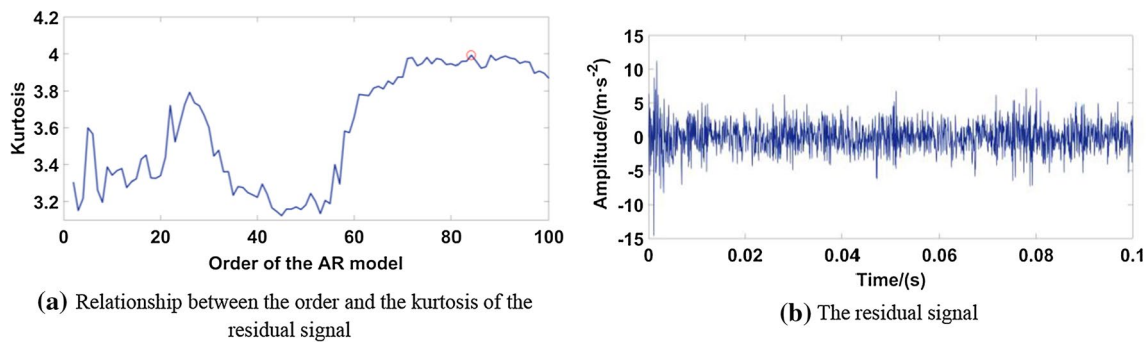


Fig. 14 Analyzed results of the measured outer race fault signal using the AR model

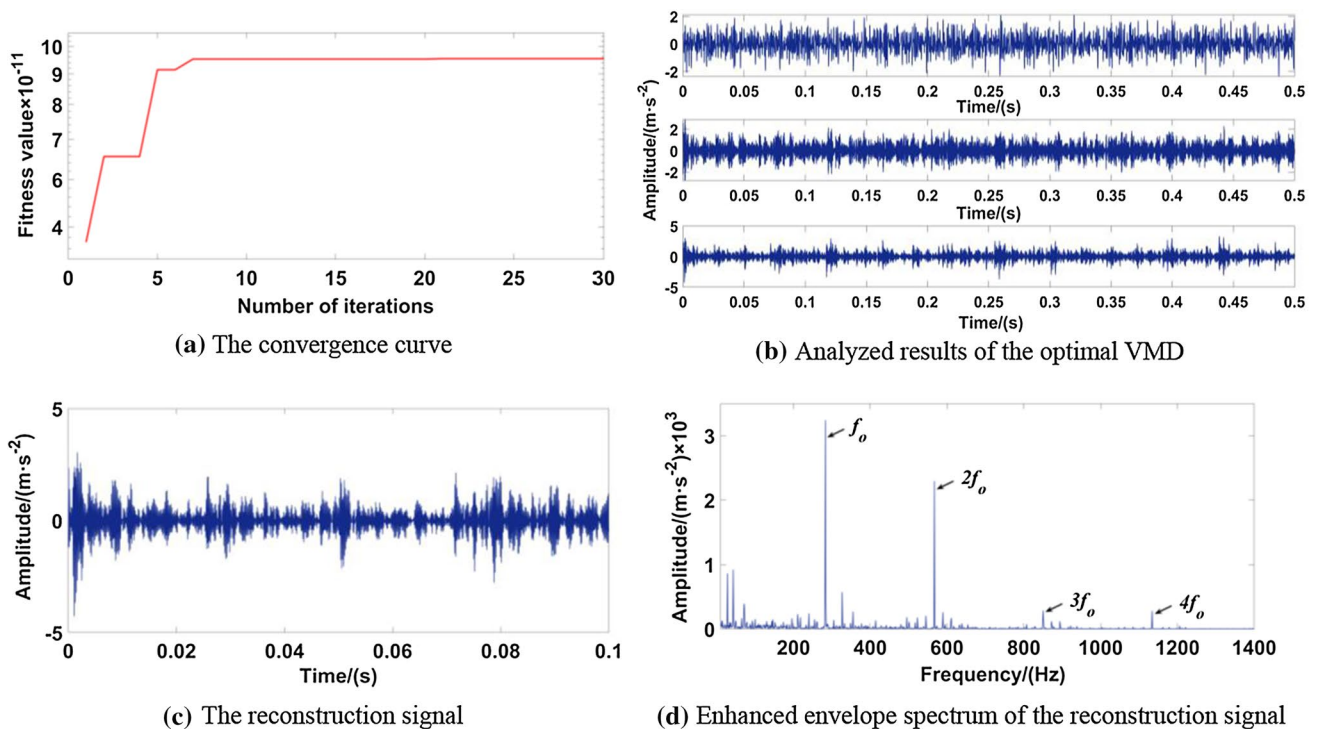


Fig. 15 Analyzed results of the measured outer race fault signal using the proposed technique

where the irrelevant components are weakened obviously compared with the original signal, which is convenient for the subsequent analysis.

Then, the residual signal is processed by the optimized VMD approach. The maximum value of the fitness function 9.528×10^{-11} is acquired after seven iterations according to the convergence curve shown in Fig. 15a, and the corresponding optimal parameters for VMD are $K=3$ and $\alpha=1000$, respectively. The decomposition results using the optimal VMD are exhibited in Fig. 15b and three IMFs are obtained. Three IMFs' FDCK are calculated, which

are 3.14×10^{-13} , 3.47×10^{-13} and 9.54×10^{-11} . Based on the mean of the three FDCK values, the effective IMF is selected as IMF3, the reconstruction signal is then acquired and depicted in Fig. 15c where the irrelevant noise is further suppressed. The enhanced envelope spectrum is finally calculated and displayed in Fig. 15d, where abundant fault signatures are successfully extracted. The FCF along with three harmonics is prominent, while the noise almost has no interference. The features mentioned above indicate that the REB has outer race defect.

Analogously, two comparisons are conducted to highlight the effectiveness of the proposed technique. Figure 16

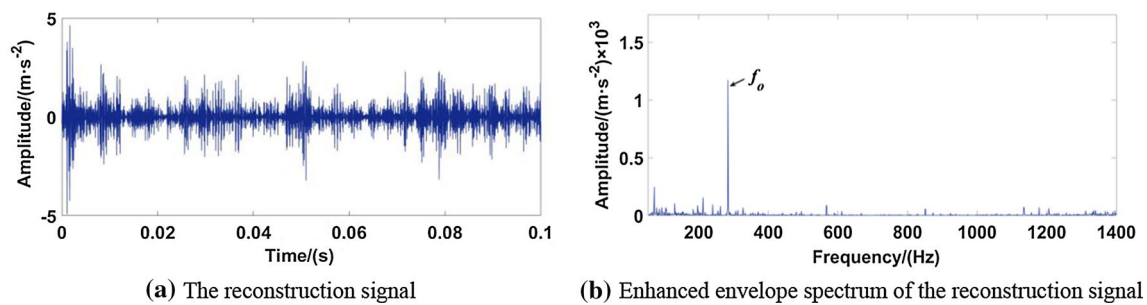


Fig. 16 Analyzed results of measured outer race fault signal using the fix-parameter VMD

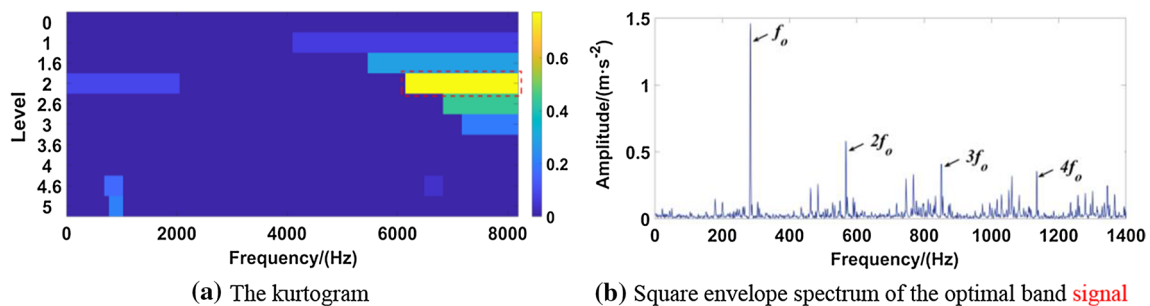


Fig. 17 Analyzed results of measured outer race fault signal using the fast kurtogram

displays the analyzed results using the fix-parameter VMD. In Fig. 16a, the impact components can be detected but in its enhanced envelope spectrum of Fig. 16b, only the FCF f_o is prominent.

At last, Fig. 17 shows the analyzed results using the fast kurtogram; the location of the optimal frequency band is determined by the kurtogram shown in Fig. 17a, whose central frequency and bandwidth are 7168 Hz and 2048 Hz, respectively, and the optimal band signal can be acquired. Figure 17b displays the square envelope spectrum of the optimal band signal, where the FCF and three harmonics are able to be detected.

When comparing Figs. 15d, 16b, 17b, by using the proposed method, abundant fault features can be detected with less interference frequency components in the enhanced envelope spectrum, and the amplitude of FCF components is relatively large. Two comparisons' result of REB with outer race defect further demonstrates that the proposed algorithm can eliminate the strong background noise and the fault features of REB can be extracted effectively. Moreover, similar to the results in Sect. 5.1, the calculation time for the proposed approach is longer.

6 Conclusions

To address the problem that the fault features of REB are weak and often submerged by background noise, an enhanced variational mode decomposition technique is proposed in this paper, which combines the AR model, the optimized VMD and enhanced envelope spectrum together; thus, the fault signatures of REB can be extracted accurately. Some conclusions are obtained during the research of this paper.

- (1) The FDCK is utilized as the fitness function and the parameters K and α are selected automatically, which avoid the problem that the parameters of VMD need to be determined by human experience in advance. The whole optimization process is simple, which can achieve good results.
- (2) With the help of AR model, the signal can be preprocessed to enhance the fault features in advance. The enhanced envelope spectrum is employed to further highlight the fault characteristic components, and the interference of irrelevant components can be suppressed.
- (3) The effectiveness of the proposed technique is verified by using the simulation signal with signal-to-noise ratio of -10 dB, the measured bearing inner race fault and

outer race fault signals. The results indicate that this method can accurately extract the fault features of REB under the interference of strong background noise. The comparison results with the fix-parameter VMD and fast kurtogram further demonstrate the effectiveness of the technique.

References

- Antoni J (2007) Fast computation of the kurtogram for the detection of transient faults. *Mech Syst Signal Process* 21(1):108–124
- Antoni J (2016) The infogram: entropic evidence of the signature of repetitive transients. *Mech Syst Signal Process* 74:73–94
- Gu X, Yang S, Liu Y et al (2016) Rolling element bearing faults diagnosis based on kurtogram and frequency domain correlated kurtosis. *Meas Sci Technol* 27:125019
- Moshrefzadeh A, Fasana A (2018) The Autogram: An effective approach for selecting the optimal demodulation band in rolling element bearings diagnosis. *Mech Syst Signal Process* 105:294–318
- Xu Y, Tian W, Zhang K et al (2018) Application of enhanced fast kurtogram based on empirical wavelet transform for bearing fault diagnosis. *Meas Sci Technol* 30:035001
- Zhang C, Liu Y (2020) A two-step denoising strategy for early-stage fault diagnosis of rolling bearings. *IEEE Trans Instrum Meas* 99:1–12
- Wang L, Liu Z, Cao H et al (2020) Subband averaging kurtogram with dual-tree complex wavelet packet transform for rotating machinery fault diagnosis. *Mech Syst Signal Process* 142:106755
- Gao Y, Karimi M, Kudreyko AA et al (2017) Spare optimistic based on improved ADMM and the minimum entropy deconvolution for the early weak fault diagnosis of bearings in marine systems. *ISA Trans* 78:98–104
- Zhang L, Cai B, Xiong G et al (2020) Multistage fault feature extraction of consistent optimization for rolling bearings based on correlated kurtosis. *Shock Vib* 7:1–16
- Yang R, Li H, Wang C et al (2018) Rolling element bearing weak feature extraction based on improved optimal frequency band determination. *ARCHIVE Proc Inst Mech Eng C J Mech Eng Sci* 2018:1–12
- Li Y, Cheng G, Liu C (2020) Research on bearing fault diagnosis based on spectrum characteristics under strong noise interference. *Measurement* 169:108509
- Xu Y, Cai Z, Cai X et al (2019) An enhanced multipoint optimal minimum entropy deconvolution approach for bearing fault detection of spur gearbox. *J Mech Sci Technol* 33:3–4
- Buzzoni M, Antoni J, D'Elia G (2018) Blind deconvolution based on cyclostationarity maximization and its application to fault identification. *J Sound Vib* 432:569–601
- Wang X, Yan X, He Y (2020) Weak fault detection for wind turbine bearing based on ACYCBD and IESB. *J Mech Sci Technol* 34(4):1399–1413
- Miao Y, Zhao M, Lin J et al (2017) Application of an improved maximum correlated kurtosis deconvolution method for fault diagnosis of rolling element bearings. *Mech Syst Signal Process* 92:173–195
- Zhu D, Zhang Y, He W et al (2020) Compound faults diagnosis of rolling element bearing using adaptive CYCBD and cross-correlation spectrum, shock and vibration. *J Vib Shock* 39(11):116–122
- Grover C, Turk N (2020) Rolling element bearing fault diagnosis using empirical mode decomposition and Hjorth parameters. *Proc Comput Sci* 167:1484–1494
- Xu Y, Cai Z, Ding K (2018) An enhanced bearing fault diagnosis method based on TVF-EMD and a high-order energy operator. *Meas Sci Technol* 29:095108
- Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adva Adapt Data Anal* 1(1):1–41
- Chen W, Xiao Y (2019) An improved ABC algorithm and its application in bearing fault diagnosis with EEMD. *Algorithms* 12(4):72
- Han T, Liu Q, Li Z et al (2019) Fault feature extraction of low speed roller bearing based on Teager energy operator and CEEMD. *Measurement* 138:400–408
- Bouhalais ML, Djebala A, Ouelaa N et al (2017) CEEMDAN and OWMRA as a hybrid method for rolling bearing fault diagnosis under variable speed. *Int J Adv Manuf Technol* 94:2475–2489
- Zhao H, Li L (2016) Fault diagnosis of wind turbine bearing based on variational mode decomposition and teager energy operator. *IET Renew Power Gener* 11(4):453–460
- Li H, Xu Y, An D et al (2019) Application of a flat variational modal decomposition algorithm in fault diagnosis of rolling bearings. *J Low Freq Noise Vib Active Control* 39(2):335–351
- Yan X, Jia M (2019) Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech Syst Signal Process* 122:56–86
- Gu R, Chen J, Hong R et al (2020) Incipient fault diagnosis of rolling bearings based on adaptive variational mode decomposition and Teager energy operator. *Measurement* 149:106941
- Wang H, Jiang X, Guo W et al (2020) An enhanced VMD with the guidance of envelope negentropy spectrum for bearing fault diagnosis. *Complexity* 2020:1–23
- Hua L, Tao L, Xing W et al (2020) An optimized VMD method and its applications in bearing fault diagnosis. *Measurement* 166:108185
- Shankar KP, Annamalai KL, Kumar LS (2018) Selecting effective intrinsic mode functions of empirical mode decomposition and variational mode decomposition using dynamic time warping algorithm for rolling element bearing fault diagnosis. *Trans Inst Meas Control* 41:1923–1932
- Mirjalili S, Lewis A et al (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67

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