Adaptive feature mode decomposition: a fault-oriented vibration signal decomposition method for identification of multiple localized faults in rotating machinery

Xiuzhi He · Xiaoqin Zhou · Jieli Li · Chris K. Mechefske · Rongqi Wang · Guofeng Yao · Qiang Liu

Abstract Identification of multiple mechanical faults from vibration signals has always been one of the most challenging tasks in the field of condition monitoring and fault diagnosis. This study proposes a new fault-oriented vibration signal decomposition method, called adaptive feature mode decomposition (AFMD), to identify multiple localized faults in rotating machines interfered by strong periodic harmonics in a robust and effective manner. The autoregressive model is first introduced as a preprocessing technique for initially reducing deterministic components of the raw signal. Then, for signal decomposition, an adaptive finite impulse response (FIR) filter bank is designed utilizing the blind deconvolution theory. The filter coefficients of the FIR filter bank are iteratively updated to make each filtered sub-signal infinitely approach their deconvolution objective functions based on the correlated kurtosis. Meanwhile, the filter length is adaptively determined using the developed evaluation indicator termed as the weighted squared envelope harmonic-to-noise ratio. Finally, several decomposed modes focusing on fault signatures can be acquired automatically, using the newly proposed mode selection strategy that considers signal similarity across multiple domains. The proposed AFMD method can significantly reduce the likelihood of incorrect diagnoses, as demonstrated by two simulated and two experimental datasets with multiple localized bearing and gear faults. The analysis results show that the proposed method outperforms over the state-of-the-art feature mode decomposition and the most popular variational mode decomposition in multi-fault feature extraction and weak fault detection, when interfered by strong periodic harmonics as well as other background noise.

Keywords Multiple faults · Fault diagnosis · Rotating machinery · Adaptive feature mode decomposition (AFMD) · Correlated kurtosis (CK) · Vibration signal processing

X. He · X. Zhou · J. Li · R. Wang · G. Yao · Q. Liu (✉)
School of Mechanical and Aerospace Engineering, Jilin University, Changchun 130025, China
e-mail: liuqiang2012@jlu.edu.cn

X. He
State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310027, China

X. He · X. Zhou · J. Li · R. Wang · G. Yao · Q. Liu
Key Laboratory of CNC Equipment Reliability, Ministry of Education, Jilin University, Changchun 130025, China

C. K. Mechefske
Department of Mechanical and Materials Engineering, Queen's University, Kingston K7L 3N6, Canada

1 Introduction

Due to harsh and complicated working conditions, the components of rotating machines are prone to damage

during their service lives, which may lead to equipment collapse and even major safety accidents [1]. It is of great significance to develop effective fault diagnosis methods to ensure the safe operation of mechanical equipment and prevent accidents. Vibration-based signal processing techniques [2] have been widely applied for rotating machinery fault diagnosis in the past few decades, with a series of remarkable contributions made by many researchers [3–6]. Benefiting from a solid mathematical foundation, excellent interpretability and smaller amounts of captured data, advanced signal processing methods for condition monitoring and fault diagnosis still have significant research value and irreplaceable engineering application requirements in the era of deep learning.

Gears and rolling element bearings are the most used components in rotating machines. When a localized fault occurs on the contact surface of such a rotating component, the defect position will result in periodic sudden impacts during operation that stimulate the damped free vibration of the system according to its natural frequency, thus generating a series of consecutive impulses in the vibration signal [7]. However, besides fault-induced impulses, periodic harmonics caused by shaft rotation and gear mesh as well as complex background noise and random shocks are also contained in the measured vibration signal, especially in modern equipment with higher integration whose complexity significantly increases the difficulty of fault feature extraction. This has brought more challenges to traditional feature extraction methods such as statistical characteristic analysis and frequency spectrum analysis.

Using signal decomposition methods to divide the measured vibration signal into a series of sub-signals with explicit physical meaning is a reasonable and effective way for further extraction of hidden fault features. It should be mentioned that the adaptive signal decomposition method represented by the empirical mode decomposition (EMD) proposed by Huang et al. [8] has attracted much attention in recent years, achieving many beneficial results in the field of mechanical fault diagnosis [9, 10]. Different from wavelet transform and other traditional signal decomposition methods through basis function expansion, EMD is completely data-driven. It does not need to construct any prior basis to match the characteristic structure of the signal, nor does it need to impose any constraints on the time domain, frequency domain, or

time–frequency domain, thus allowing it to realize the adaptive decomposition of arbitrary signals. Any multi-component signal could be adaptively decomposed by EMD into several sub-signals called intrinsic mode functions (IMFs), which is also referred to as "mode." After that, Wu and Huang [11] further proposed the ensemble empirical mode decomposition (EEMD) which uses the addition of white noise with a certain amplitude to the signal. It takes the ensemble mean of multiple decomposition results of the signal assisted by noise as the final decomposition result, further enhancing the stability of EMD to a certain extent. In order to improve the accuracy of signal decomposition and reconstruction, Yeh et al. [12] proposed the complementary ensemble empirical mode decomposition (CEEMD), which achieved the purpose of reducing residual noise within the modes by adding auxiliary noise to the original signal in the form of complementary addition and subtraction. Moreover, Torres et al. [13, 14] introduced two improved CEEMD methods to further solve the mode aliasing issue. Inspired by EMD, a great quantity of adaptive decomposition algorithms has emerged, such as local mean decomposition (LMD) [15], intrinsic time-scale decomposition (ITD) [16], local characteristic-scale decomposition (LCD) [17], and adaptive local iterative filtering decomposition (ALIFD) [18]. These algorithms mainly decompose the signal through iterative calculation in the time domain with low operating efficiency. They are prone to accumulation of errors, which significantly affects the final mode decomposition accuracy.

Some frequency-band-based decomposition methods such as empirical wavelet transform (EWT) [19] and variational mode decomposition (VMD) [20] were also proposed. EWT aims to construct a group of orthogonal filters based on the adaptive wavelet basis to divide the frequency spectrum of the signal into several frequency bands. However, EWT highly relies on spectrum segmentation based on robust peak detection preprocessing. For VMD, it transforms the decomposition problem into a constrained optimization problem. The ultimate goal of VMD is to decompose the signal with multiple frequency components into a series of discrete quasi-orthogonal bandwidth limited modes and minimize the sum of bandwidths of all modes. Benefiting from the solid mathematical foundation, non-recursive characteristic, and high computational efficiency, VMD has

become a popular and effective tool. Nevertheless, the proper selection of main parameters (mode number and balancing parameter) of VMD is always the key to its decomposition performance. More importantly, although the above-mentioned algorithms have been applied to fault diagnosis and have achieved good results in some applications [21–23], it should be noted that they are not designed for it. In other words, these methods do not consider the inherent characteristics of the faulty signal, i.e., impulsiveness and periodicity.

Recently, the feature mode decomposition algorithm (FMD) was proposed by Miao et al. [24], which has been proven to outperform VMD on fault signal decomposition. FMD is inspired by the blind deconvolution theory and combines it into the signal decomposition process. It is known that blind deconvolution is an effective filtering tool that can enhance the useful information with the noise and interferences suppressed for the complicated signal with unknown transfer function and unpredictable noise [25]. This is exactly appropriate for processing mechanical fault signals containing complex components [26]. The core of blind deconvolution for fault diagnosis is to use a fault-related indicator as the deconvolution objective function to orient the deconvolution process and design an inverse filter to convolute with the faulty signal to be processed, thus achieving the filtering of fault-related components. Motived by this, FMD designs such a FIR filter bank to divide the signal into several sub-signals and uses correlated kurtosis (CK) considering both impulsiveness and periodicity of the signal as the deconvolution objective function to deconvolute each sub-signal. The FIR filter can remove the limitations of filter shape and bandwidth of some algorithms such as VMD and EWT, which means more fault information can be extracted. The correlation coefficient (CC) of every two sub-signals serves as the mode selection criterion in which the one with lower CK value from the two sub-signals with maximum CC value is discarded, and the selection process will end when the number of reserved sub-signals equals the desired mode number. Therefore, benefiting from the advancements of blind deconvolution theory and the FIR filter bank without restrictions of filter shape and bandwidth, FMD has capability of extracting impulsive fault features even for the cases with multiple faults.

Nevertheless, the robustness and ability of multi-fault identification of FMD still needs to be enhanced. For instance, the filter length and the desired mode number are two essential input parameters of FMD which currently need to be pre-determined through trial and error, which increases the uncertainty of parameter selection. Furthermore, many redundant iterations are generated by the current mode selection principle of FMD, and more importantly, the useless or redundant modes cannot be thoroughly removed relying solely on temporal coefficient correlation. These issues might further lead to misdiagnosis or missed diagnosis especially for multiple faults under strong interferences, which limit the effectiveness and robustness of FMD in industrial applications.

To overcome these drawbacks, a new fault-oriented adaptive feature mode decomposition method, abbreviated as AFMD, is explored in this study for multi-fault identification in rotating machinery. Firstly, the autoregressive (AR) model is implemented as a preprocessing technique to reduce the periodic harmonics, in which the model order is determined according to an evaluation of the maximum kurtosis of the AR residue. Then, the adaptive FIR filter bank covering the whole frequency band of the faulty signal through a 50%-overlapping frequency band division method with initialization by Hanning window is designed for providing a fuzzy decomposition direction. After that, filter coefficients of the adaptive FIR filter bank are iteratively updated to make each filtered sub-signal infinitely approach their deconvolution objective functions on the basis of the CK. Finally, the decomposition results can be automatically achieved through a proposed mode selection strategy.

The main work of this study is as follows:

(1) A new indicator called weighted squared envelope harmonic-to-noise ratio (weighted SEHNR) is introduced to determine the filter length of the adaptive FIR filter bank in an automatic way instead of by manual experience. It is more suitable for dealing with multi-fault signals since it can take the fault information within all the useful modes into account.

(2) A new mode selection strategy that considers similarity of modes in the time domain, frequency domain and envelope domain along with the consistency of the estimated fault period of modes through autocorrelation is proposed. It

aids avoidance of mode redundancy and mode mixing, and more importantly, solves the problem of manually pre-setting the desired mode number.

(3) A new fault-oriented vibration signal decomposition called AFMD method is developed for adaptive identification of multiple localized faults in rotating machinery.

The main advancements of the proposed AFMD method can be briefly summarized below:

(1) The two key parameters of AFMD, i.e., the filter length of the adaptive FIR filter bank and the desired mode number, can be selected adaptively rather than preset by trial and error;

(2) AFMD has superior capability of concentrating on the multi-fault impulses contaminated by random shocks, background noise and strong deterministic harmonics caused by gear meshing and shaft rotation.

The reminder of this study is organized as below. Section 2 introduces the basic theory of FMD and its limitations. Section 3 elaborates the proposed AFMD method. Section 4 and Sect. 5 present the performance validation of the proposed AFMD over FMD and VMD through two simulated and two experimental scenarios. Discussion and conclusion are given in Sect. 6.

2 Feature mode decomposition (FMD) and its limitations

2.1 Basic theory of FMD

FMD is proposed to decompose a signal into several sub-signals (i.e., modes) non-recursively by using a FIR filter bank, which is inspired by the blind deconvolution theory [24]. To be specific, the FIR filter bank is designed by iteratively updating filter coefficients to make each mode infinitely approach the deconvolution objective function. Since CK takes impulsiveness and periodicity of a signal into account simultaneously, it is ideal for characterization of mechanical fault signals. FMD uses it as the deconvolution objective function to orientate the deconvolution process of each mode and further achieves the decomposition.

Supposing a mechanical vibration signal $x(t)$ with length N, FMD can be regarded as a constrained problem expressed as below:

$$\underset{\{f_k(l)\}}{\arg\max} \left\{ \mathrm{CK}_M(\mathbf{u}_k) = \sum_{n=1}^{N} \left(\prod_{m=0}^{M} u_k(n-mT_s) \right)^2 / \left(\sum_{n=1}^{N} u_k(n)^2 \right)^{M+1} \right\}$$

$$\mathrm{s.t.} \, u_k(n) = \sum_{l=1}^{L} f_k(l)x(n-l+1)$$

(1)

where $u_k(n)$ is the k-th mode, $f_k(l)$ is the k-th FIR filter with length L, and M is the shift order of the CK. T_s is the period presented by sampling number that is denoted as:

$$T_s = f_s \cdot T \tag{2}$$

where f_s is the sampling frequency and T is the fault period of interest. In the process of FMD, T_s is estimated by iteratively finding the local maximum value at the period position after the zero-crossing point of the signal's autocorrelation spectrum [24], which will be further described in the following section.

As to the FIR filter bank, it is achieved by uniformly dividing the whole frequency band of the signal into K segments with Hanning window initialization for better filter performance [24]. Each mode's lower and upper cut-off frequencies, i.e., f_l and f_u, can be illustrated as:

$$\begin{cases} f_l = k \cdot f_s/2K \\ f_u = (k+1) \cdot f_s/2K \end{cases} \quad k = 0, 1, 2, ..., K-1 \tag{3}$$

where K is the total number of filters in the FIR filter bank.

The constrained problem of Eq. (1) can be solved by the eigenvector algorithm [24, 27], and the mode will be expressed in matrix form:

$$\mathbf{u}_k = \mathbf{X}\mathbf{f}_k \tag{4}$$

where

$$\mathbf{u}_k = \begin{bmatrix} u_k(1) \\ \vdots \\ u_k(N-L+1) \end{bmatrix}, \mathbf{X} = \begin{bmatrix} x(1) & \cdots & x(L) \\ \vdots & \ddots & \vdots \\ x(N-L+1) & \cdots & x(N) \end{bmatrix}, \mathbf{f}_k = \begin{bmatrix} f_k(1) \\ \vdots \\ f_k(L) \end{bmatrix}$$

(5)

And CK of the mode will be represented as:

$$CK_M(\mathbf{u}_k) = \frac{\mathbf{u}_k^H \mathbf{W}_M \mathbf{u}_k}{\mathbf{u}_k^H \mathbf{u}_k} \tag{6}$$

where $[\cdot]^H$ represents the conjugate transpose operation, and $\mathbf{W}_M$ is a weighted matrix as:

$$\mathbf{W}_M = \begin{bmatrix} \left(\prod_{m=0}^{M} u_k(1-mT_s)\right)^2 & 0 & \cdots & 0 \\ 0 & \left(\prod_{m=0}^{M} u_k(2-mT_s)\right)^2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \left(\prod_{m=0}^{M} u_k(N-L+1-mT_s)\right)^2 \end{bmatrix} \\ \times \frac{1}{\sum_{n=1}^{N-L+1} u_k(n)^{M-1}} \tag{7}$$

Therefore, the expression of Eq. (6) can be rewritten as:

$$CK_M(\mathbf{u}_k) = \frac{\mathbf{f}_k^H \mathbf{X}^H \mathbf{W}_M \mathbf{X} \mathbf{f}_k}{\mathbf{f}_k^H \mathbf{X}^H \mathbf{X} \mathbf{f}_k} = \frac{\mathbf{f}_k^H \mathbf{R}_{XWX} \mathbf{f}_k}{\mathbf{f}_k^H \mathbf{R}_{XX} \mathbf{f}_k} \tag{8}$$

where $\mathbf{R}_{XWX}$ and $\mathbf{R}_{XX}$ are the weighted correlation matrix and correlation matrix, respectively. Then, it can be seen that Eq. (8) is a generalized Rayleigh quotient [24], which means its maximization problem with respect to $\mathbf{f}_k$ is equivalent to the eigenvector associated with the maximum eigenvalue λ of the following generalized eigenvalue problem:

$$\mathbf{R}_{XWX} \mathbf{f}_k = \mathbf{R}_{XX} \mathbf{f}_k \lambda \tag{9}$$

where λ corresponds to maximum CK. After solving Eq. (9) by means of an iterative algorithm whose process is similar to that illustrated in Ref. [28], the filter coefficients $\mathbf{f}_k$ will be continually updated to approach the final solution with maximum CK.

It should be noticed that the FIR filter bank initialized by Hanning window only provides a fuzzy decomposition direction and the shape and frequency bandwidth of filters are unlimited, which benefits from the blind deconvolution theory along with the above illustrated period estimation and filter iterative updating process [24]. However, in FMD, the total number of filters K and the desired mode number K_d need to be pre-defined according to the value relationship of $K \geq K_d$, thus indicating there might be many mixing or redundant modes containing the same or similar information under the updating principle of filter initialization when K is far greater than K_d. Hence, the correlation coefficient (CC) is applied as a measurement index for similarity comparison of all modes. The two modes with maximum CC value are locked

firstly and the mode with higher CK is reserved while the one with lower CK is abandoned until the number of reserved modes equals to the desired mode number K_d, which is regarded as the mode selection principle of FMD.

The flowchart of FMD is shown in Fig. 1, and its procedure can be concluded as below [24]:

(1) Input the raw vibration signal $\mathbf{x}$ and set initial parameters, i.e., total number of filters K, desired mode number K_d, filter length L, and pre-iteration number;

(2) Initialize the FIR filter bank by Hanning window with K filters (K is set as 5–10) and start iteration $i = 1$ (i goes up to the pre-iteration number);

(3) Decompose the raw signal $\mathbf{x}$ into a series of modes by $\mathbf{u}_k^i = \mathbf{x} * \mathbf{f}_k^i$, where $k = 1, 2, \ldots, K$, and $*$ denotes the convolution operation;

(4) Estimate fault period T_k^i by iteratively finding the point with local maximum value after the zero-crossing point of the autocorrelation spectrum of mode $\mathbf{u}_k^i$. Update filter coefficients using $\mathbf{x}$, $\mathbf{u}_k^i$, and T_k^i, and achieve one complete iteration cycle. Set $i = i + 1$;

(5) Determine whether i reaches the pre-iteration number. If yes, continue to the next step, otherwise return to Step 4;

(6) Calculate correlation coefficients of all modes and form a $K \times K$ matrix $\mathbf{CC}_{(K \times K)}$. Calculate and compare CK values of the two modes that have the maximum CC value of $\mathbf{CC}_{(K \times K)}$, and then abandon the one with lower CK. Set $K = K - 1$;

(7) Determine whether K reaches the desired mode number K_d. If yes, end the iteration and continue to the next step, otherwise return to Step 3;

(8) Output the results and select all the reserved modes as the final decomposition modes.

2.2 Limitations of FMD

By virtue of the FIR filter bank updating principle and CK as the deconvolution objective function as well as the period estimation without prior fault information, FMD is capable of extracting machinery impulsive fault features even for some cases with multiple faults. It expands the application range of conventional
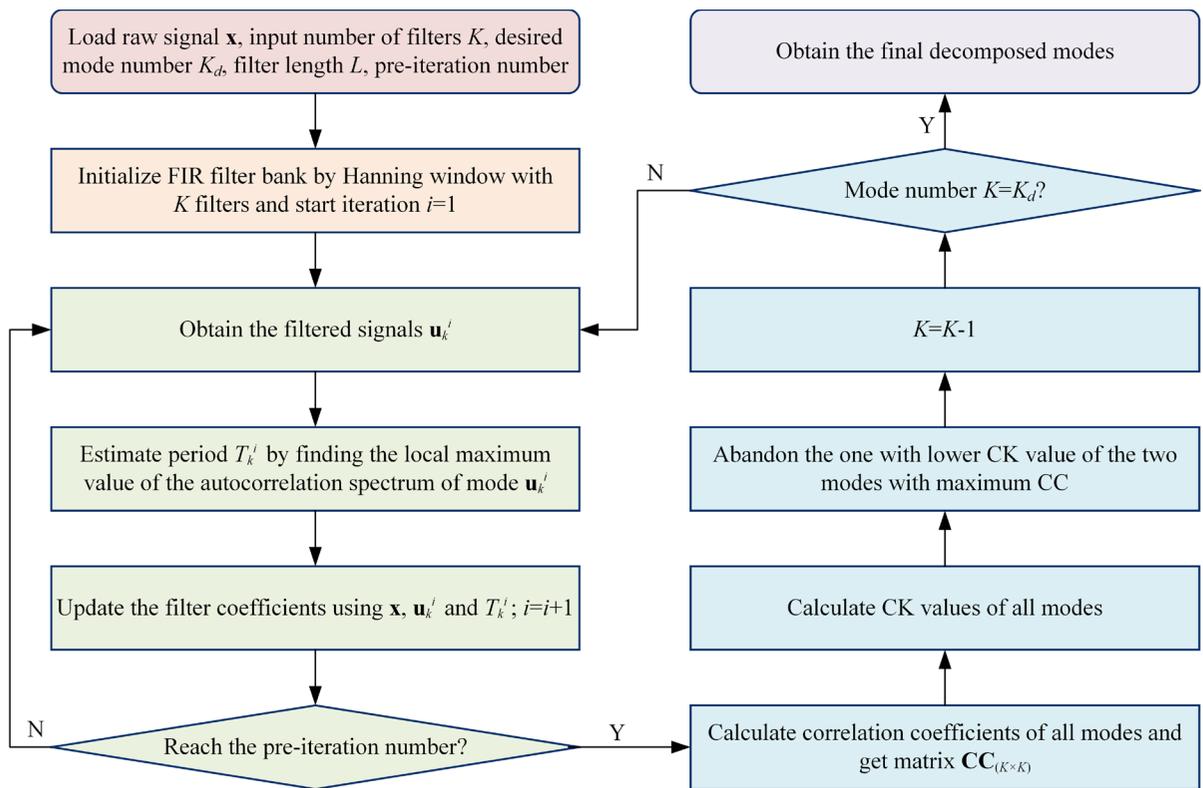
Fig. 1 Flowchart of the original FMD method

deconvolution methods and is superior to the decomposition methods based on frequency narrow-band filtering such as VMD, which has been verified through simulation and experiment scenarios by Miao et al. [24]. Despite the advantages, FMD still has limitations that need to be perfected for better performance, especially for feature extraction of multiple faults under periodic harmonic interference.

Taking an example for the further explanation, the numerical vibration signal with two bearing localized faults simulated in Ref. [24] is applied in this study for a fair comparison. The basic information of the simulation signal is introduced as below: The resonant frequencies of the two faults, i.e., inner race fault and outer race fault, are set as 3605 and 2000 Hz, respectively. Their faulty feature frequencies are 28 Hz (termed as f_2) and 22 Hz (termed as f_1), respectively. Besides, periodic harmonics, random shock interference and background Gaussian noise are also included in the synthetic signal. Detailed information of the synthetic signal can be found in Ref. [24]. One thing to point out is that the amplitudes of

the two periodic harmonics in the original signal used in Ref. [24] are both set as 0.025 and are far less than the amplitudes of the faulty impulses set as 1, which might not be appropriate on the occasion that the faulty bearing faces strong periodic harmonic interference such as in gearboxes. Therefore, the amplitudes of the two periodic harmonics used in this study are set as 0.5 instead of 0.025 to better investigate the performance of FMD under harmonic interference. The final simulated vibration signal containing two localized faults along with its frequency spectrum is displayed in Fig. 2, where f_r (7 Hz) and $2f_r$ (14 Hz) denote feature frequencies of the above-mentioned two periodic harmonics, respectively.

The initial input parameters are chosen by the recommendation of Ref. [24]. The filter length L and the number of filters K are set as 40 and 10, respectively, and the desired mode number K_d is suggested to be empirically selected from 3, 2, 1. The pre-iteration value of the first iteration is set as $20 - 2 \times (K - K_d)$ with the values of all other following iterations set as 2, according to the
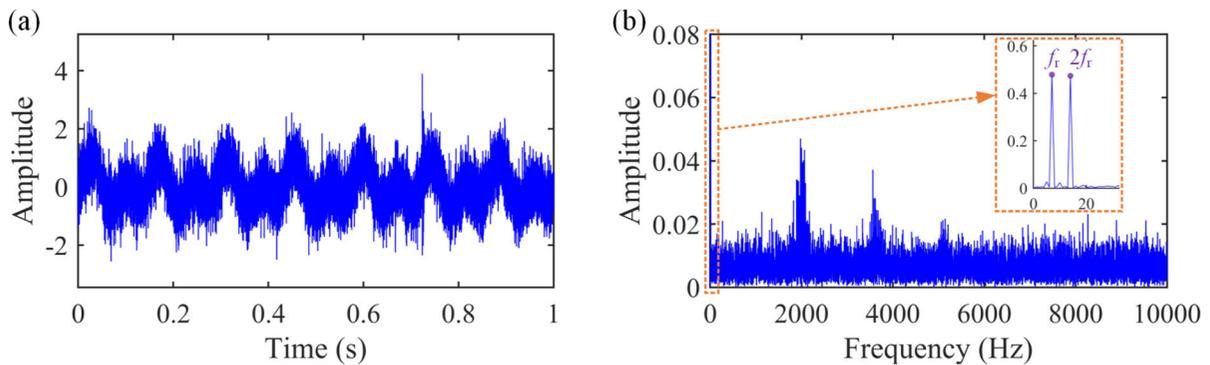
Fig. 2 **a** Multi-fault vibration signal simulated in Ref. [24] with adding amplitudes of periodic harmonics to 0.5 and **b** its frequency spectrum

MATLAB code shared by the authors. All the parameter settings remain the same unless otherwise specified.

The mode number is firstly set as 3 in FMD. As shown in Fig. 3, Mode #1 and Mode #3 are the periodic harmonic component and the random shock component, respectively, while only Mode #2 consists of the inner race fault feature frequency f_2 and its harmonics ($2\times$, $3\times$, etc.), without any information of the outer race fault f_1 throughout all these three decomposed modes. This is different from the result obtained in Ref. [24] where the three modes are respectively dominated by outer race fault, inner race fault and the random shock.

When the mode number is then set as 2, the multi-fault features are extracted into two separated modes as shown in Fig. 4, which is consistent with the result obtained in Ref. [24]. Based on the comparison results, one can find that the decomposition performance of FMD is highly sensitive to the desired mode number K_d. Using the current parameter settings might not be competent for the fault feature extraction task under strong periodic harmonic interference which is common in the field of mechanical fault diagnosis.

The filter length L is another important parameter to be pre-determined. In the original FMD, a suitable range of [30, 100] is recommended for filter length selection, which is verified based on a simulation case containing a single impulsive fault in Ref. [24]. $L = 40$ is then used in all simulation and experiment scenarios. Nevertheless, when dealing with more complex signals under different conditions,

the current empirical selection principle through trial and error might fail to meet the command of accurate feature extraction, especially always setting a fixed value as the filter length. To illustrate this, the decomposed results of FMD with $K_d = 2$ and $L = 41$ are displayed in Fig. 5, where Mode #1 and Mode #2 are dominated by periodic harmonics and inner race fault impulses, with a misdiagnosis of the outer race fault. Compared with Fig. 4, although the two filter lengths differ by only 1, i.e., $L = 40$ and $L = 41$, the diagnosis results are not the same.

Another finding to be mentioned is that the amplitudes of modes decomposed by the original FMD might be distorted due to the usage of filter designing functions embedded in MATLAB. Figure 6 shows the magnitude response with respect to normalized frequency of the fourth filter of the FIR filter bank after the first complete iteration of FMD.

It can be seen that the maximum value of the amplitude-frequency curve achieves 12.94 dB rather than 0 dB, which means the amplitudes of the obtained modes after the first complete iteration are all magnified by 4.44 times (corresponds to the filter gain of 12.94 dB). Furthermore, when the mode number $K_d = 3$ and the number of filters $K = 10$ are applied, according to the mode selection principle of FMD illustrated in Sect. 2.1, FMD involves 7 (calculated by $K - K_d$) times of redundant iteration for mode selection, leading to the amplitudes of final decomposed results magnified by 668.84 times. (Specific magnifications of all iterations are not shown here.)
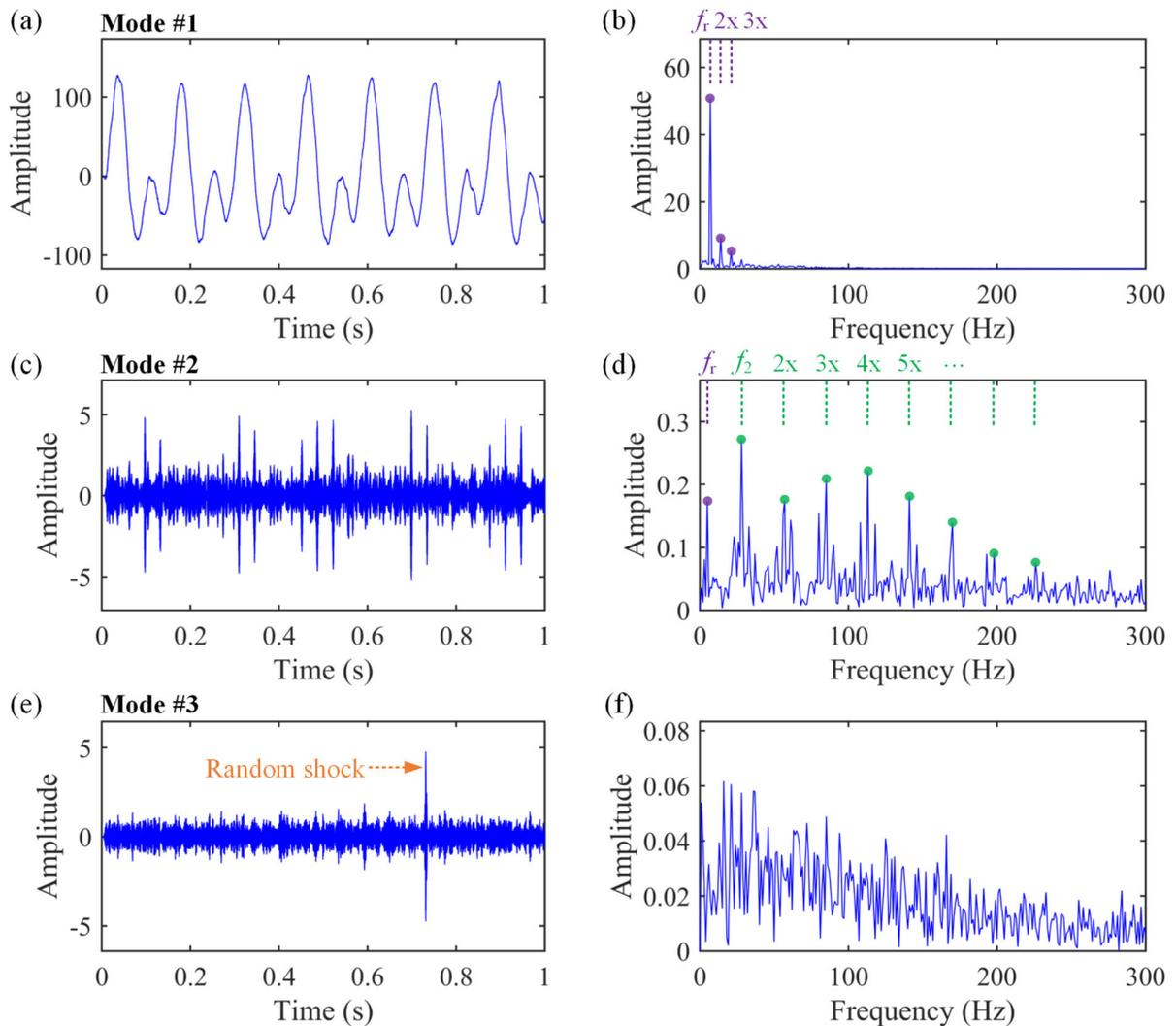
Fig. 3 Decomposed results of original FMD with $K_d = 3$ and $L = 40$: **a, c, e** time domain; **b, d, f** envelope spectra

This obviously limits the further development of FMD in quantitative diagnosis.

To sum up, the limitations of FMD can be concluded as follows:

(1) The filter length L and the desired mode number K_d are the two most important input parameters of FMD and need to be pre-determined by trial and error, which increases the uncertainty of parameter selection and further results in potential misdiagnosis or missed diagnosis particularly for multiple faults.

(2) FMD still faces the challenge caused by harmonic interference especially when choosing inappropriate input parameters, although it uses CK as the deconvolution objective function to take both impulsiveness and periodicity of a faulty vibration signal into consideration.

(3) The current mode selection principle of FMD has a lot of redundant iterations, and more importantly, the useless or redundant modes cannot be thoroughly removed relying solely on temporal coefficient correlation.

(4) The amplitudes of obtained results decomposed by FMD are significantly distorted, which is not conducive to further quantitative fault diagnosis.
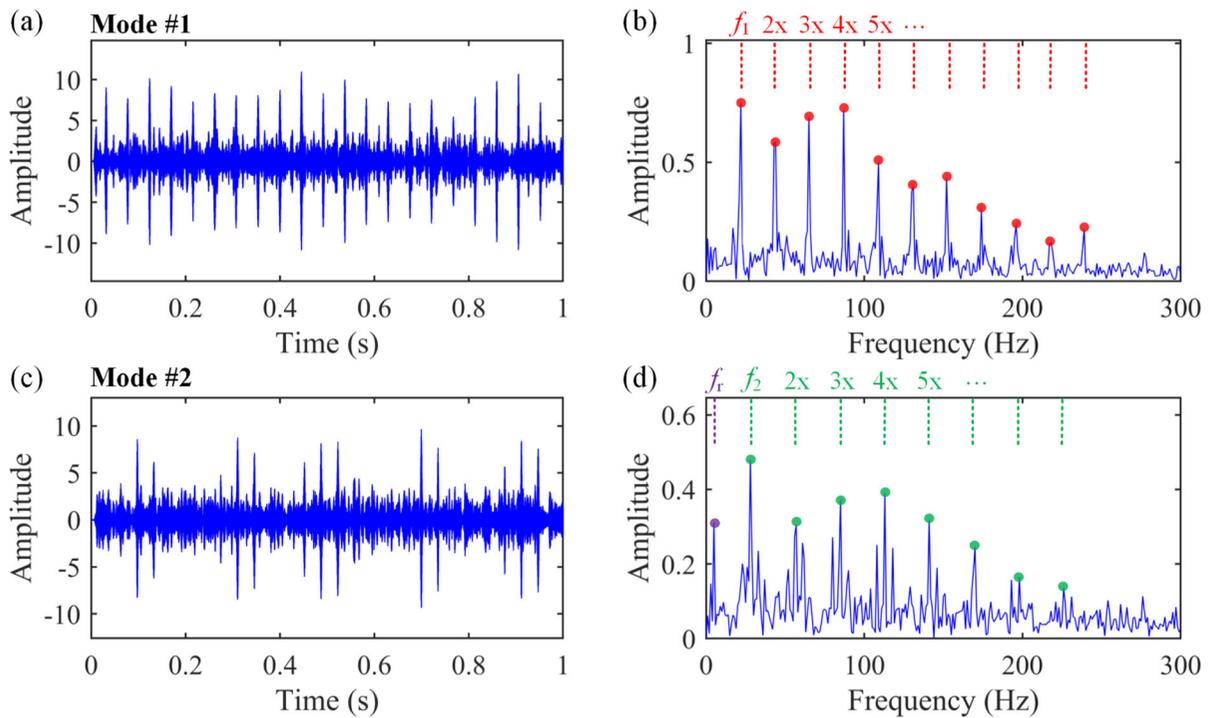
Fig. 4 Decomposed results of original FMD with $K_d = 2$ and $L = 40$: **a, c** time domain; **b, d** envelope spectra
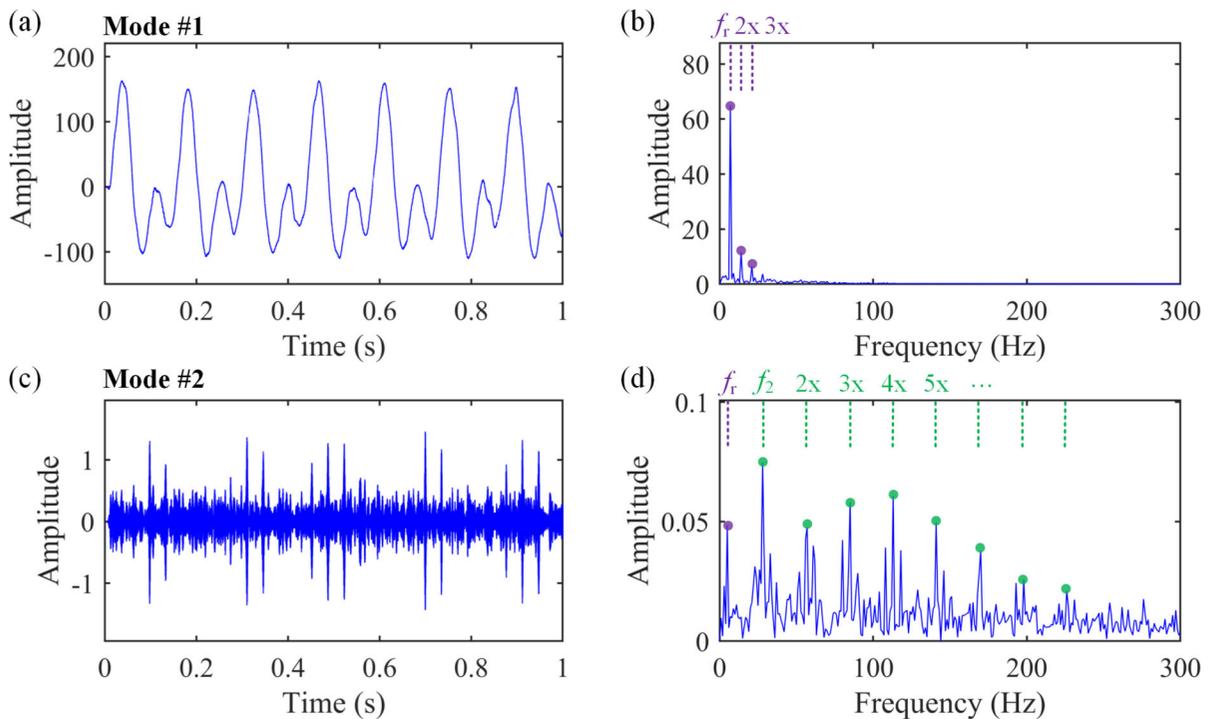
Fig. 5 Decomposed results of original FMD with $K_d = 2$ and $L = 41$: **a, c** time domain; **b, d** envelope spectra
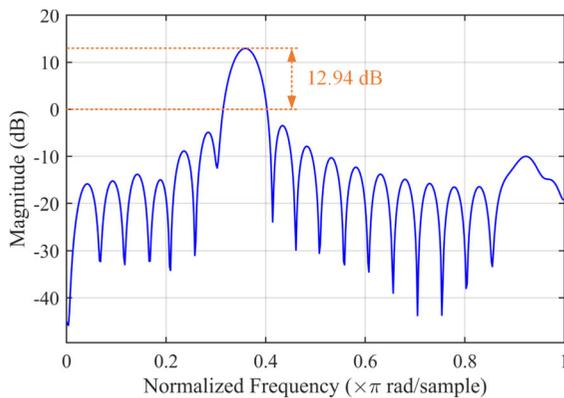
Fig. 6 Magnitude response with respect to normalized frequency of the fourth filter of the FIR filter bank after the first complete iteration of FMD

3 The proposed method: adaptive feature mode decomposition (AFMD)

To overcome shortcomings of the state-of-the-art FMD method and further robustly broaden its application to multi-fault diagnosis, the adaptive feature mode decomposition (AFMD) method is developed in this study. A new mode selection strategy is proposed to reduce redundant and noisy modes, aiding to achieve the decomposition task in an automatic way instead of by manual experience. It considers the similarity of modes in time domain, frequency domain and envelope domain along with the consistency of estimated fault period of modes through autocorrelation. Meanwhile, a new adaptive way for filter length selection that is more suitable for multi-fault feature detection is also introduced based on a new evaluation indicator called weighted squared envelope harmonic-to-noise ratio (weighted SEHNR). Besides, to separate the fault-induced signatures from the periodic harmonic interference, the autoregressive (AR) model is used to pre-whiten the raw signal and initially remove deterministic (discrete frequency) components, i.e., periodic harmonics. The proposed AFMD is elaborated as follows.

3.1 Autoregressive (AR) model as a preprocessing technique

For bearings and gears, the two most common parts in rotating machines, their collected faulty vibration signals with localized defects mainly contain the fault-induced impulses, periodic harmonics caused by rotating shafts or gear meshing, random shocks and background Gaussian white noise. The periodic harmonics are generally deterministic and consist of a series of discrete frequency components. On the contrary, the faulty impulses can be regarded as the random-like part because of their broadband characteristics in frequency spectrum which is similar to Gaussian white noise [29]. Since AR model is a popular method to estimate the deterministic periodic components, the residue after AR process will be the desired signal for FMD to bring into better play the role of CK on extraction of consecutive faulty impulses.

The AR model can be expressed by a linear regression on itself plus an additive noise term $e(n)$, whose mathematical formula for a real discrete-time vibration signal $x(n)$ is given as below:

$$x(n) = -\sum_{p=1}^{q} a(p)x(n-p) + e(n) \qquad (10)$$

where $a(p)$ is pth coefficient in the AR model and q is the model order.

The coefficient $a(p)$ can be acquired by solving the Yule–Walker equations through the Levinson–Durbin recursion (LDR) algorithm [30]. The order q serves as the key parameter and highly affects the final performance of AR model, which can be determined by manual experience or using the Akaike information criterion (AIC) that is the most commonly used approach. However, AIC is more suitable for the condition that the noise term $e(n)$, i.e., the AR residue, strictly conforms to a Gaussian distribution, which cannot directly reflect faulty information of the AR residue with faulty impulses.

Since kurtosis has been widely applied for measuring the impulsiveness of vibration signals without any prior fault information, it helps with the issue of model order selection in this study. In other words, the model order q is decided by finding the one corresponding to the maximum kurtosis of AR residue $e(n)$, which can be illustrated as follows:

$$\arg\max_{\{q\}} \left\{ \text{kurtosis}(q) = \frac{\frac{1}{N}\sum_{n=1}^{N}\left(e_q(n) - \overline{e}_q\right)^4}{\left(\frac{1}{N}\sum_{n=1}^{N}\left(e_q(n) - \overline{e}_q\right)^2\right)^2} \right\}$$
$$(11)$$

where $\overline{e}_q$ is the mean value of $e_q(n)$, and the range of order q can vary with different conditions rather than

being a fixed value. However, to the best of authors' knowledge, there has been no unified standard for selecting the range of order q. In this study, it is recommended from 20 to 200 according to experiments, which is competent for processing the current data length of all implemented scenarios in the following sections.

3.2 Adaptive FIR filter bank and its updating along with the period estimation

The so-called adaptive FIR filter bank proposed in Ref. [24] can be divided into two main steps: at first, design a FIR filter bank with several uniform segments initialized by Hanning window, and then iteratively update coefficients of the filter bank based on the blind deconvolution theory, i.e., to make each filtered signal closer to the deconvolution objective function with CK as the core.

As for the FIR filter bank, the superiority of Hanning window initialization over the ordinary simple initialization for locking the fault period rapidly and robustly has been clarified and verified in Ref. [24] and the detailed information will not be repeated in this study for the sake of brevity. But remarkably, as illustrated in Fig. 7a, the FIR filter bank of the original FMD is accomplished by dividing the whole frequency band of the raw signal into K segments equally with consecutive lower and upper cut-off frequencies, which can be expressed by Eq. (3) in Sect. 2.1.

In this study, another frequency band division method is proposed, as shown in Fig. 7b, which is more advantageous to make each segment contain more fault information through covering the whole frequency band with 50% overlapping between two adjacent segments. The lower and upper cut-off frequencies of the filter bank are then expressed as below:

$$\begin{cases} f_l = k \cdot f_s/2(K+1) \\ f_u = (k+2) \cdot f_s/2(K+1) \end{cases} \quad k = 0, 1, 2, \ldots, K-1 \tag{12}$$

where f_l and f_u are the lower and upper cut-off frequencies of segments, respectively, and f_s is the sampling frequency. The segment number K is selected as an odd number. By comparing these two division methods with the same number of filters, the proposed method shown in Fig. 7b could, to a large
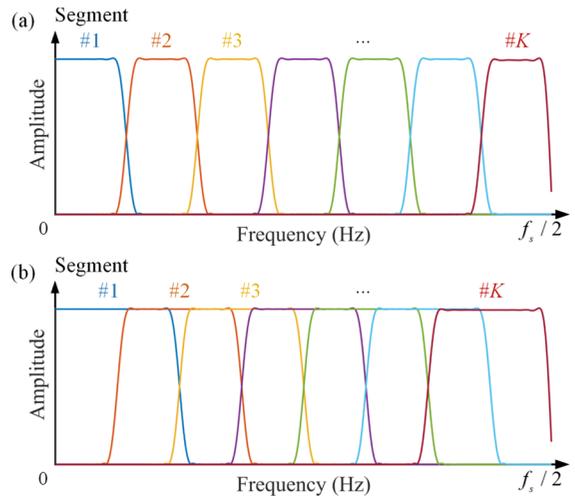
Fig. 7 FIR filter bank initialized by Hanning window of **a** the original FMD and **b** the proposed AFMD

extent, prevent an entire fault-induced frequency band from being separated into two or more segments by the ordinary one shown in Fig. 7a in which the filter banks have consecutive lower and upper cut-off frequencies.

It should be noted that the FIR filter bank is used for providing a fuzzy decomposition direction, and the FIR filter bank with initialization only gives a much coarser denoising process, which means the fault features might still be corrupted by noise and other non-faulty components. Therefore, although the proposed 50%-overlapping frequency band division method is better for each filter to contain more fault information, it still needs the further signal processing procedure to finish the fault diagnosis task.

To this end, the filter coefficient updating process inspired by the blind deconvolution theory is implemented subsequently to concentrate on the fault signatures. In brief, a series of modes and filter coefficients can be iteratively updated by filtering the signal using the FIR filter bank, and the updating iteration will be ended when the filtered signals within each mode reach their own maximum CK values under a given iteration cycle, whose procedure is elaborated according to Eqs. (4)–(9) in Sect. 2.1.

It should be noted that CK requires input fault period T_s as the prior knowledge which is a decisive factor in the filter coefficient updating process. The traditional way is to pre-define it in the light of the fault feature frequency of the candidate machine part. Nevertheless, it is impractical to acquire accurate T_s
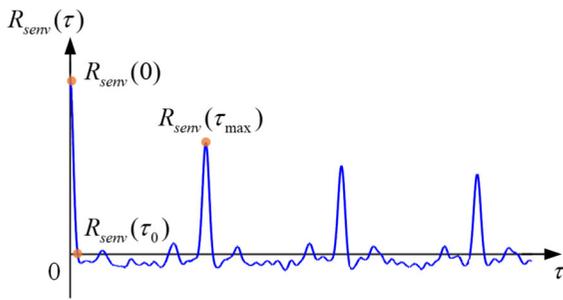
Fig. 8 Schematic diagram of the autocorrelation spectrum of the squared envelope of a faulty vibration signal

under unknown or fluctuating speed conditions in engineering practice. More importantly, when dealing with the signal within multiple faults, it is also unrealistic to assign all potential fault period values in the signal processing, since the specific fault localization of the machine component is unknown in advance. Hence, a practical and accurate approach for fault period estimation is highly desirable. Miao et al. [24] employed an autocorrelation-based solution by means of finding the local maximum value at the period position after the zero-crossing point of the signal's autocorrelation spectrum. Furthermore, the autocorrelation spectrum of signal's squared envelope, instead of the raw signal or its ordinary envelope, is applied in this study due to the fact that the squared envelope is more beneficial to enhancing fault characteristics [31].

The squared envelope of the vibration signal $x(n)$, written as $x_{\text{senv}}(n)$, can be calculated by

$$x_{\text{senv}}(n) = |x(n) + j\mathcal{H}\{x(n)\}|^2 \tag{13}$$

where $\mathcal{H}\ominus\{\cdot\}$ denotes the Hilbert transformation and j represents the imaginary unit. Then, the autocorrelation of the squared envelope signal with respect to a time lag τ is defined as below:

$$R_{\text{senv}}(\tau) = \sum_{n=1}^{N} x_{\text{senv}}(n)x_{\text{senv}}(n+\tau) \tag{14}$$

where $\tau = 0, 1, 2, \ldots, N-1$, with N the total sampling number.

Figure 8 displays the autocorrelation spectrum of the squared envelope of a single-fault vibration signal. $R_{\text{senv}}(0)$ is the global maximum of the autocorrelation spectrum, while $R_{\text{senv}}(\tau_{\max})$ represents the local maximum where $\tau_{\max}$ generally equals to the fault period T_s. In order to extract the location information of

$R_{\text{senv}}(\tau_{\max})$ easily in practice, the zero-crossing point is then named as the point τ_0 when $R_{\text{senv}}(\tau_0) = 0$. Consequently, the abscissa position of the point that reaches maximum value after the zero-crossing point is represented as the estimated fault period without any prior fault feature information, i.e., $\tau_{\max} = T_s$, which is applicable to the industrial scenarios.

3.3 Proposed mode selection strategy based on signal similarity in multiple domains

According to Sect. 3.2, the vibration signal can be divided into K segments that are regarded as the candidate modes through the FIR filter bank, and it is highly likely that many modes contain the same or similar components due to the operation of the 50%-overlapping division, filter updating, and automatic period estimation. To eliminate the redundant modes, a new mode selection strategy considering similarity of modes in multiple domains, i.e., time domain, frequency domain, and squared envelope domain, along with the identity of estimated faulty period of each mode, is proposed in this study.

Taking two modes $u_p(n)$ and $u_q(n)$ as an example, the similarity of these two modes in the time domain, also called the temporal correlation coefficient CC [24], is written by

$$CC_{pq} = \frac{\sum\limits_{n=1}^{N} \left(u_p(n) - \overline{u}_p\right)\left(u_q(n) - \overline{u}_q\right)}{\sqrt{\sum\limits_{n=1}^{N} \left(u_p(n) - \overline{u}_p\right)^2}\sqrt{\sum\limits_{n=1}^{N} \left(u_q(n) - \overline{u}_q\right)^2}} \tag{15}$$

where $\overline{u}_p$ and $\overline{u}_q$ denote the mean value of modes $u_p(n)$ and $u_q(n)$, respectively. Generally, larger CC values (higher than 0.7) could indicate a higher similarity between the two compared modes, and the same to the similarity indicators of other domains. Then, the mode similarity in the frequency domain can be evaluated by the spectral orthogonality (SO) [7] expressed as

$$SO_{pq} = \frac{\sum\limits_{m} S_p(m)S_q(m)}{\sqrt{\sum\limits_{m} S_p(m)^2}\sqrt{\sum\limits_{m} S_q(m)^2}} \tag{16}$$

where $S_p(m)$ and $S_q(m)$ represent the spectral amplitudes of modes $u_p(n)$ and $u_q(n)$, respectively, which can be acquired through discrete Fourier transform (DFT):

$$S(m) = \frac{1}{N} \sum_{n=0}^{N-1} u(n) e^{-j2\pi mn/N} \qquad (17)$$

The similarity of modes in the squared envelope spectrum (SES) domain, termed as SESO (orthogonality of squared envelope spectrum), is defined as:

$$\text{SESO}_{pq} = \frac{\sum_m \text{SES}_p(m)\text{SES}_q(m)}{\sqrt{\sum_m \text{SES}_p(m)^2}\sqrt{\sum_m \text{SES}_q(m)^2}} \qquad (18)$$

where $\text{SES}(m)$ is the squared envelope spectrum of modes and can be calculated by substituting Eq. (13) into Eq. (19):

$$\text{SES}(m) = \frac{1}{N} \sum_{n=0}^{N-1} x_{\text{senv}}(n) e^{-j2\pi mn/N} \qquad (19)$$

Besides, the consistency of estimated period of the two modes is presented by comparing their difference:

$$\varepsilon_T = \left| T_s^p - T_s^q \right| / T_s^p \leq 0.01 \qquad (20)$$

where the less the value ε_T is, the better the consistency of the two modes with respect to T_s is. Note that the threshold of value ε_T is set as 0.01 which satisfies the cases used in this study and could be further decreased upon different working conditions.

Before further elaborating the mode selection strategy, the squared envelope harmonic-to-noise ratio (SEHNR) should be illustrated first. Inspired by the harmonic-to-noise ratio (HNR) presented in Refs. [32, 33] for characterizing the periodicity of the faulty impulses, the SEHNR benefiting the characteristic of enhancing the fault features is proposed in this study as an enhanced version over HNR, which is defined as below:

$$\begin{aligned} \text{SEHNR} &= \frac{R_{\text{senv}}(\tau_{\max})/R_{\text{senv}}(0)}{1 - R_{\text{senv}}(\tau_{\max})/R_{\text{senv}}(0)} \\ &= \frac{R_{\text{senv}}(\tau_{\max})}{R_{\text{senv}}(0) - R_{\text{senv}}(\tau_{\max})} \end{aligned} \qquad (21)$$

where $R_{\text{senv}}(0)$ and $R_{\text{senv}}(\tau_{\max})$ can be computed by Eq. (14) and respectively denote the total energy of the squared envelope signal and the energy of the periodic fault-induced impulses. In this study, SEHNR = 0.5 is used as a reasonable threshold to evaluate the fault component whether it could be presented effectively, since the energy of noise component is two times of the energy of the desired fault component under this condition. In other words, the mode with SEHNR value lower than 0.5 is considered as a noisy one containing limited fault information.

Therefore, the proposed mode selection strategy can be illustrated as follows: First, the CK and SEHNR values of all modes along with their mean CK value are calculated; then, abandon the modes with CK lower than the mean CK value and the modes with SEHNR lower than 0.5; after that, calculate the multi-domain similarity of all reserved modes in pairs based on Eqs. (15), (16) and (18), and remove the modes with all similarity values in the time domain, frequency domain and squared envelope spectrum domain higher than 0.7 along with the modes with similar estimated period values based on Eq. (20); finally, the retained modes will be the desired decomposed modes.

3.4 Adaptive determination of filter length

Filter length L is the crucial factor in the filtering operation that is generally pre-defined by trial and error. However, an inappropriate filter length might result in misdiagnosis or missed diagnosis results even though a small difference, for example, $L = 40$ and $L = 41$, as shown in Figs. 4 and 5. Hence, an adaptive approach for selecting filter length is highly desired in practice.

For this purpose, a new index called weighted SEHNR is further established on the basis of the definition of SEHNR illustrated in Sect. 3.3, and is expressed as:

$$\text{Weighted SEHNR} = \frac{\text{SEHNR}_k}{\sum_{k=1}^{K_d} \text{SEHNR}_k} \times \text{SEHNR}_k \qquad (22)$$

where SEHNR_k denotes the SEHNR value of k-th mode and the ratio of SEHNR_k to the sum SEHNR value of all the final desired K_d modes is taken as its weight. This index could consider the fault information within all the useful modes rather than only focusing on the mode with maximum SEHNR value, which is more suitable for multi-fault feature detection.

Then, the filter length could be adaptive determined by searching the optimal one with maximum weighted SEHNR in a specific range $[L_{\text{lower}}, L_{\text{upper}}]$ where L_{lower} and L_{upper} are denoted as the lower and upper filter length, respectively, written as:

$$L_{\text{opt}} = \underset{L \in [L_{\text{lower}}, L_{\text{upper}}]}{\arg\max} \{\text{Weighted SEHNR}_L\} \qquad (23)$$

3.5 Procedure of the proposed AFMD method

Based on the above investigation, the AFMD method for multi-fault diagnosis is proposed in this paper. The overall flowchart of this approach is shown in Fig. 9, while its detailed implementation is illustrated as follows:

(1) Input the raw vibration signal **x** and set initial parameters, i.e., total number of filters K and pre-iteration number;

(2) Implement the AR model to preprocess the raw signal and obtain the AR residual signal **e** according to Sect. 3.1

(3) Initialize the FIR filter bank by Hanning window with K filters (K is set as 11) based on Eq. (12) and start iteration $i = 1$ (i goes up to the pre-iteration number);

(4) Set the filter length search range from 1 to 200, and start the first loop with $l = 1$

(5) Decompose the AR residual signal **e** into a series of modes by $\mathbf{u}_k^i = \mathbf{e} * \mathbf{f}_k^i$, where $k = 1, 2, \ldots, K$, and * denotes the convolution operation;

(6) Estimate fault period T_k^i by iteratively finding the point with local maximum value after the zero-crossing point of the autocorrelation spectrum of mode $\mathbf{u}_k^i$. Update filter coefficients using **e**, $\mathbf{u}_k^i$, and T_k^i, and achieve one complete iteration cycle. Set $i = i + 1$;

(7) Determine whether i reaches the pre-iteration number. If yes, continue to the next step, otherwise return to Step 5;

(8) Calculate CK and SEHNR values of all modes along with their mean CK value and abandon the modes with CK lower than the mean CK value and the modes with SEHNR lower than 0.5. Then, calculate similarity values of temporal correlation coefficient (CC), spectral orthogonality (SO), and orthogonality of the squared envelope spectrum (SESO) of all reserved modes in pairs based on Eqs. (15), (16) and (18), and remove the modes with all similarity values in the time domain, frequency domain, and squared envelope spectrum domain higher than 0.7 along with the modes with similar estimated period values based on Eq. (20);

(9) Obtain the retained modes and calculated the weighted SEHNR value of these modes. Determine whether l reaches the upper filter length. If yes, continue to the next step, otherwise return to Step 4;

(10) Obtain the optimal filter length with maximum weighted SEHNR value and the filtering results. Select all the reserved modes as the final desired decomposition modes.

Note that when using the filter designing functions embedded in MATLAB, the amplitudes of decomposed modes need to be corrected by evaluating the maximum value (unit: dB) of the amplitude-frequency curve of each corresponding filter and then reducing their amplitudes by corresponding times, which should be conducted in Step 6 of each iteration i.

4 Simulation analysis

In this section, a phenomenological vibration model for multiple localized faults of a two-stage gearbox is presented at first. Besides this vibration model, the simulated multi-fault signal for verifying the original FMD in Ref. [24] is also used here to validate the performance of the proposed method in terms of extracting multiple faults.

4.1 Vibration model for multiple localized faults of a two-stage gearbox

Supposing a parallel two-stage gearbox that consists of three shafts and two pairs of gears, where the driving gear installed in the input shaft of the first stage and a supporting rolling element bearing both contain a localized fault, then the vibration signal $x(t)$ could be modeled as follows [34]:

$$x(t) = x_{\mathrm{gmesh}}(t) + x_{\mathrm{fault}}(t) + x_{\mathrm{random}}(t) \\ + x_{\mathrm{grotation}}(t) + n(t) \tag{24}$$

The first part $x_{\mathrm{gmesh}}(t)$ denotes the gear meshing vibrations of two stages, given as:

$$x_{\mathrm{gmesh}}(t) = \sum_{m=1}^{M} A_{\mathrm{gmesh},1,m}[1 + a_m(t)] \\ \cos\left[2\pi m f_{\mathrm{gmesh},1} t + \phi_{\mathrm{gmesh},1,m} + b_m(t)\right] \\ + \sum_{m=1}^{M} A_{\mathrm{gmesh},2,m} \cos\left(2\pi m f_{\mathrm{gmesh},2} t + \phi_{\mathrm{gmesh},2,m}\right) \tag{25}$$

where $A_{\mathrm{gmesh},1,m}$, $f_{\mathrm{gmesh},1}$ and $\phi_{\mathrm{gmesh},1,m}$ represent the m-th order meshing amplitude of the first two mating
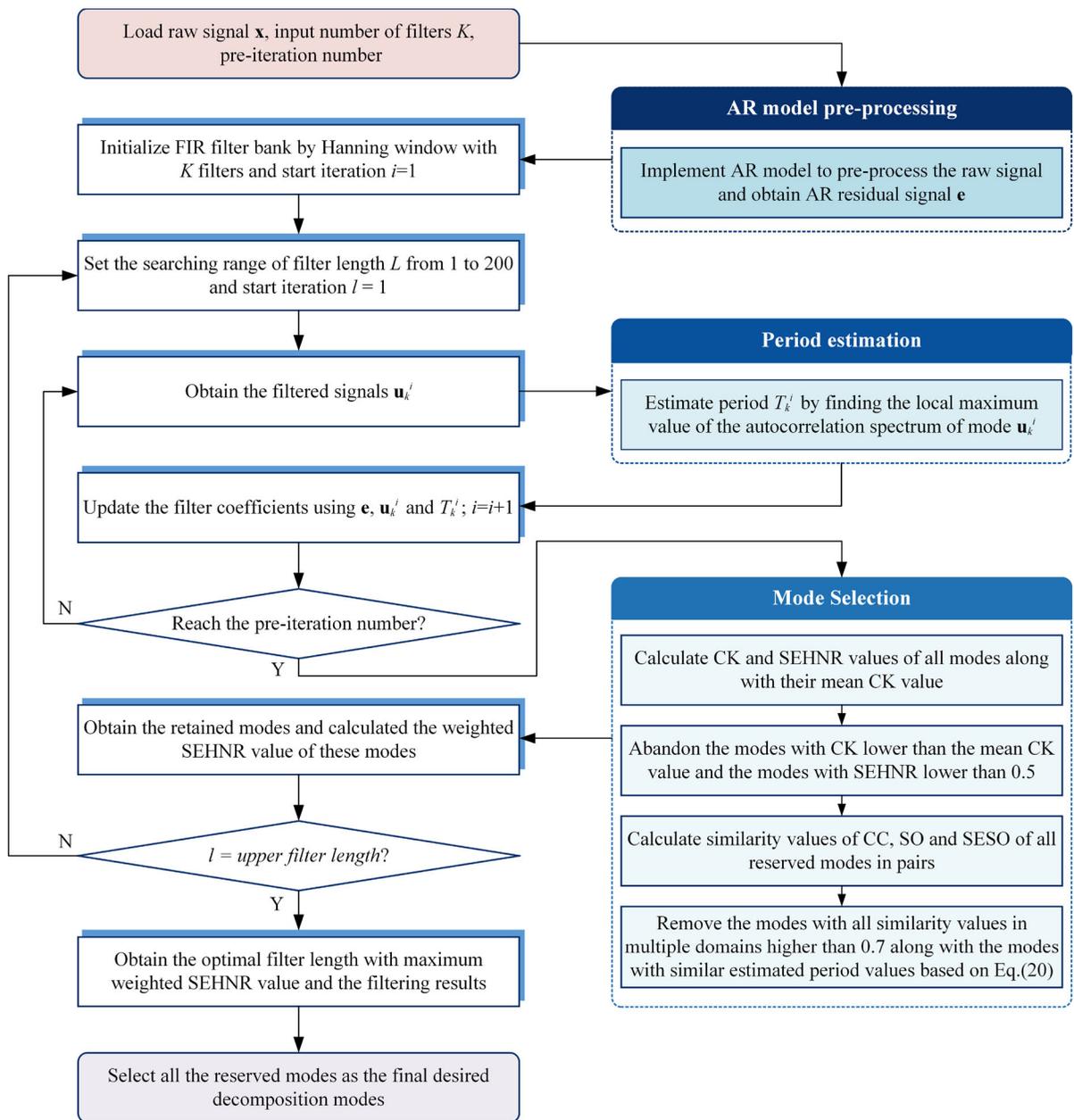
Fig. 9 Flowchart of the proposed AFMD method

gears, the meshing frequency and the m-th order meshing phase, respectively, with the similar definitions to $A_{gmesh,2,m}, f_{gmesh,2}$ and $\phi_{gmesh,2,m}$ of the second pair of mating gears; a_m and b_m are the m-th order amplitude modulation and frequency modulation functions of the first stage, respectively, which can be expressed as follows:

$$a_m(t) = \sum_{q=1}^{Q} A_{m,q} \cos\left(2\pi q f_d t + \alpha_{m,q}\right) \qquad (26)$$

$$b_m(t) = \sum_{q=1}^{Q} B_{m,q} \cos\left(2\pi q f_d t + \beta_{m,q}\right) \qquad (27)$$

where f_d denotes the rotating frequency of the driving gear mounted on the input shaft; $A_{m,q}$ and $B_{m,q}$ are the q-th order amplitudes of the m-th meshing harmonics of the two mating gears, respectively; $\alpha_{m,q}$ and $\beta_{m,q}$ are the phases.

The second part $x_{\text{fault}}(t)$ represents the transient impulses induced by localized defects

$$
\begin{aligned}
x_{\text{fault}}(t) = \sum_{gb=1}^{2} \sum_{p=1}^{P} & A_{\text{fault},gb} e^{-\beta_{n,gb}(t-pT_{\text{fault},gb})} \\
& \times \sin\left(2\pi f_{n,gb}\left(t - pT_{\text{fault},gb}\right) + \phi_{n,gb}\right)
\end{aligned}
\tag{28}
$$

where $A_{fault,gb}$ are the amplitudes of excited transient impulses of the faulty driving gear ($gb = 1$) and bearing ($gb = 2$) and $p = 1, 2, 3, \ldots, P$ is the total number of these impulses; $\beta_{n,gb}$ and $f_{n,gb}$ denote the damping coefficient and the resonance frequency of transient impulses, respectively; $T_{\text{fault},gb}$ is the time interval between two adjacent transient impulses with defining $\phi_{n,gb}$ as the phase.

The third part $x_{\text{random}}(t)$ represents the random shock interference whose definition expressed by Eq. (29) is much similar to that of Eq. (28):

$$
x_{\text{random}}(t) = A_{\text{random}} e^{-\beta_{rn}(t-t_r)} \sin(2\pi f_{rn}(t - t_r) + \phi_{rn})
\tag{29}
$$

where t_r and Δt_r are the trigger time and duration of the random shock, respectively.

Besides the Gaussian white noise $n(t)$, the remaining part $x_{\text{grotation}}(t)$ is modeled as the periodic harmonics caused by shaft rotation, written as follows:

$$
x_{\text{grotation}}(t) = \sum_{d=1}^{3} A_{\text{grotation},d} \cos\left(2\pi f_d t + \phi_{\text{grotation},d}\right)
\tag{30}
$$

where $A_{\text{grotation},d}$ are the vibration amplitudes caused by rotating shafts, and the f_d and $\phi_{\text{grotation},d}$ represent the shaft rotating frequency and phase, respectively. Note that in the x_{fault} term shown in Eq. (28), subscript $d = 1$ represents the driving gear, while in the $x_{\text{grotation}}$ term shown in Eq. (30), subscript $d = 1, 2, 3$ represents the ordinal number of the three shafts. Table 1 gives the detailed parameter values of the synthetic signal modeled by Eq. (24).

4.2 Simulation case #1: two localized faults in a bearing

In this subsection, a bearing compound fault signal is used for testing the proposed AFMD method. As mentioned in Sect. 2.2, this simulated signal is modeled by Miao et al. [24] to verify the original FMD algorithm, thus being applied in this study for a fair performance comparison with the proposed AFMD. Detailed information of the simulated fault signal is given in Sect. 2.2 and in Ref. [24], which will not be repeated here for the sake of brevity.

To better demonstrate performance of the proposed method, the very popular decomposition method VMD [20] is applied in this study. To make a fair comparison, the desired mode number of VMD is set as 2 and leave the default settings. Figure 10 shows the decomposed two modes of VMD. In the first mode, i.e., the first band-limited instinct mode function (abbreviated as BLIMF), the periodic harmonic components are dominated, as displayed in Fig. 10a, b, while the random shock component predominates BLIMF #2 in Fig. 10c, d. It can be concluded that VMD fails to identify the multiple faults within the signal in simulation case #1, thus leading to a missed diagnosis.

According to the flowchart of the proposed AFMD method presented in Fig. 9, the AR model preprocessing technique is firstly implemented with the optimal model order of 61, and the filtered residue termed as the input signal is used for the following decomposition procedures, as shown in Fig. 11. It can be found that compared with Fig. 2, the periodic harmonic interference components, i.e., f_r and $2f_r$, have been drastically reduced by about ten times.

The filter length of AFMD is then adaptively selected as $L = 181$ that corresponds to the maximum weighted SEHNR value of 3.023 as presented in Fig. 12. For comparison, $L = 40$ with the weighted SEHNR value of 1.984 is also marked in Fig. 12, which indicates that the recommended final filter length of original FMD is not the optimal one. This is one of the main factors leading to its unsatisfactory decomposition results as illustrated from Figs. 3, 4 and 5.

Figure 13 presents the decomposed results of the proposed AFMD. Two modes with different fault

Table 1 Parameters of the simulated multi-fault signal

Parameters	Values	Parameters	Values
Gear mesh part x_{gmesh}		Localized fault part x_{fault}	
$A_{gmesh,1,m}$ ($m = 1, 2, 3$)	0.3, 0.6, 0.5	$A_{fault,gb}$ ($gb = 1, 2$)	5, 7
$A_{m,q}$ ($q = 1, 2, 3$)	0.3, 0.75, 0.6	$\beta_{n,gb}$ ($gb = 1, 2$)	900, 700
$B_{m,q}$ ($q = 1, 2, 3$)	0.5, 1.5, 1	$T_{fault,gb}$ ($gb = 1, 2$)	1/20 s, 1/33 s
$A_{gmesh,2,m}$ ($m = 1, 2, 3$)	0.25, 0.45, 0.5	$f_{n,gb}$ ($gb = 1, 2$)	4800 Hz, 2800 Hz
$f_{gmesh,1}$	520 Hz		
$f_{gmesh,2}$	393.17 Hz		
Random shock part x_{random}		Gear rotation part $x_{grotation}$	
A_{random}	20	$A_{grotation,d}$ ($d = 1, 2, 3$)	0.2, 0.15, 0.1
β_{rn}	200	f_d ($d = 1, 2, 3$)	20 Hz, 12.68 Hz, 8.02 Hz
t_r; Δt_r	0.6937 s; 0.05 s		
f_{rn}	8000 Hz		

For simplicity, all the phase values included in Eq. (24) are set as zero
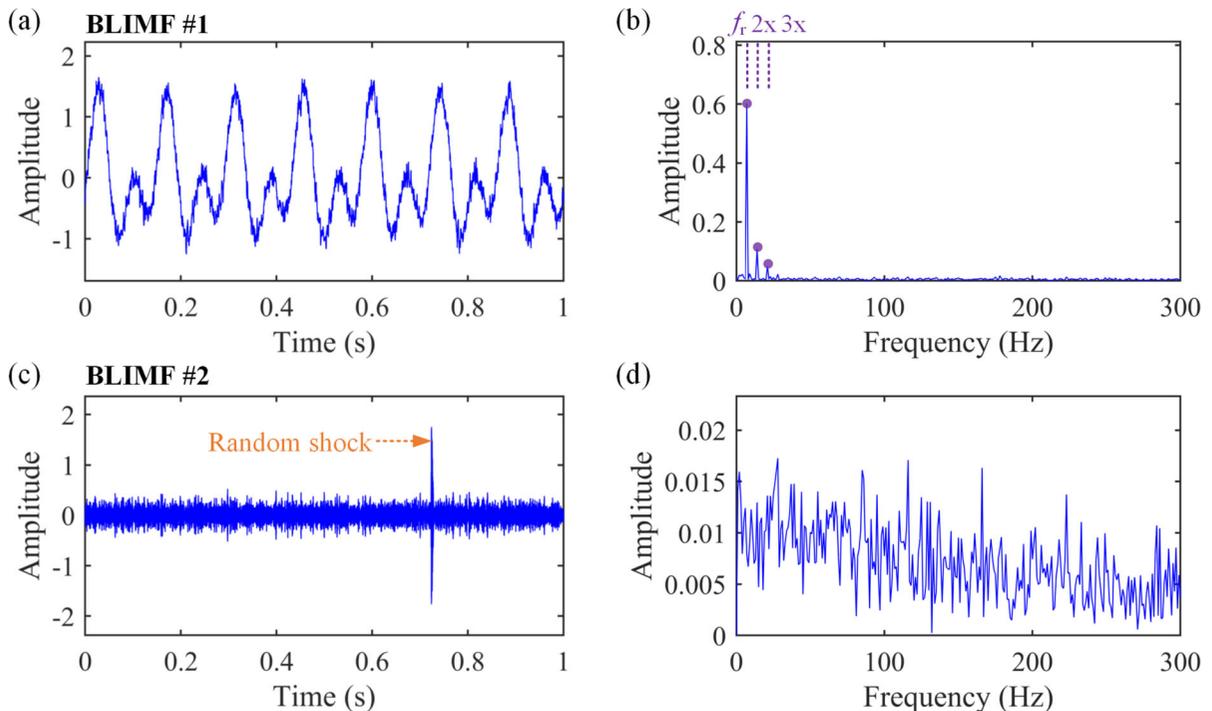
Fig. 10 Decomposed results of the multi-fault signal in simulation case #1 by VMD: **a, c** time domain; **b, d** envelope spectra

information are adaptively obtained, in which Mode #1 and Mode #2 are dominated by fault features f_1 with its harmonics ($2\times$, $3\times$, etc.) and f_2 with its harmonics, respectively, without any redundant modes and misdiagnosis results. That means the proposed method outperforms the original FMD under sever periodic harmonic interference.

4.3 Simulation case #2: two localized faults in a two-stage gearbox

In this subsection, a mixed signal including two localized faults, periodic harmonics, random shock and Gaussian background noise with SNR = − 8 dB is simulated through the vibration model illustrated in

(a)

(b)
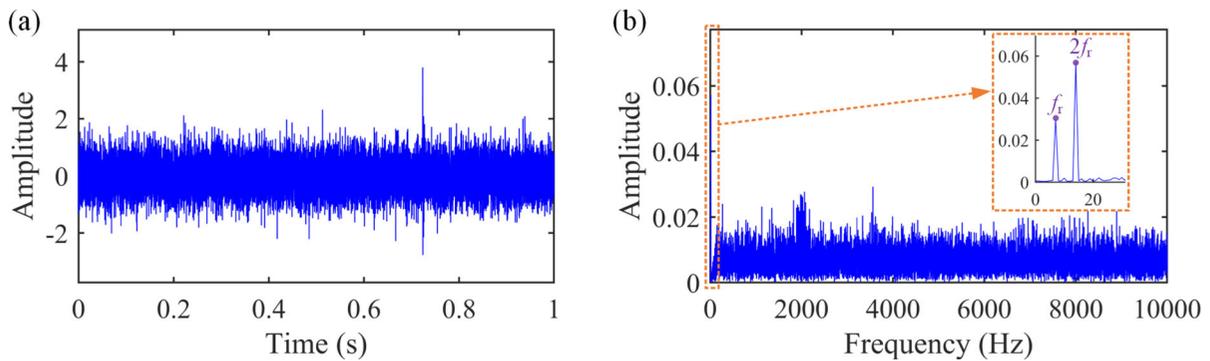
Fig. 11 a Temporal waveform and b its frequency spectrum of the multi-fault signal in simulation case #1 after AR model preprocessing
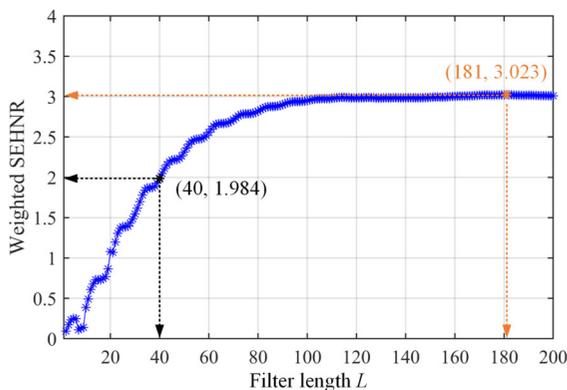
Fig. 12 Weighted SEHNR values of modes decomposed by the proposed AFMD with the change of filter length in simulation case #1

Sect. 4.1 to further validate performance of the proposed AFMD method on multi-fault detection. The sampling rate and time length of the signal are 20,000 Hz and 1 s, respectively. As shown in Fig. 14b, the exciting frequency resonance bands of the two faults are separately located around 4800 Hz and 2800 Hz with their amplitudes much less than the harmonics concentrated in the frequency range below 2000 Hz. It is more obvious in Fig. 14a that the fault-induced impulses are buried in background noise and affected by the harmonics along with the random shock.

The original FMD with filter length $L = 40$ and desired mode number $K_d = 3$ is firstly applied to simulation case #2. As shown in Fig. 15, the fault feature frequency f_2 and its harmonics ($2\times$, $3\times$, etc.) are clearly observed in the envelope spectrum of Mode #2. However, the random shock component (Mode #3)

and a redundant noisy mode (Mode #1) with useless information are also extracted as two final decomposed modes, without any evidence of fault features related to Fault #1.

When setting the desired mode number as $K_d = 2$, similar results appear in the final decomposition modes in Fig. 16, in which only Mode #2 is the meaningful one dominated by Fault #2 and noisy component is displayed in Mode #1 rather than the fault information of Fault #1. To sum up, FMD fails to detect the multiple localized faults of the complex simulated signal of simulation case #2 corrupted by random shock and periodic harmonics.

The VMD method is then applied to simulation case #2. From Fig. 17, the fault feature frequency f_2 and its harmonics ($2\times$, $3\times$, etc.) can also be detected in the envelope spectrum of BLIMF #2, similar to that of FMD in Figs. 15 and 16. As for BLIMF #1, it can be seen that the meshing frequency (20 Hz $\times$ 26 = 520 Hz) of the first pair of mating gears appears in Fig. 17b, which means VMD is also affected by a periodic component. Unfortunately, there is no evidence of fault features related to Fault #1, thus proving poor performance of VMD on multi-fault diagnosis on such a complex signal.

By contrast, the proposed AFMD method is finally used for simulation case #2. The filtered residual signal after AR model process with the optimal model order of 114 is shown in Fig. 18. Compared with Fig. 14b, the amplitudes of harmonics caused by gear meshing have been reduced by half or even more. However, the fault-related impulses are still buried in strong background noise in both the time domain and its frequency spectrum.
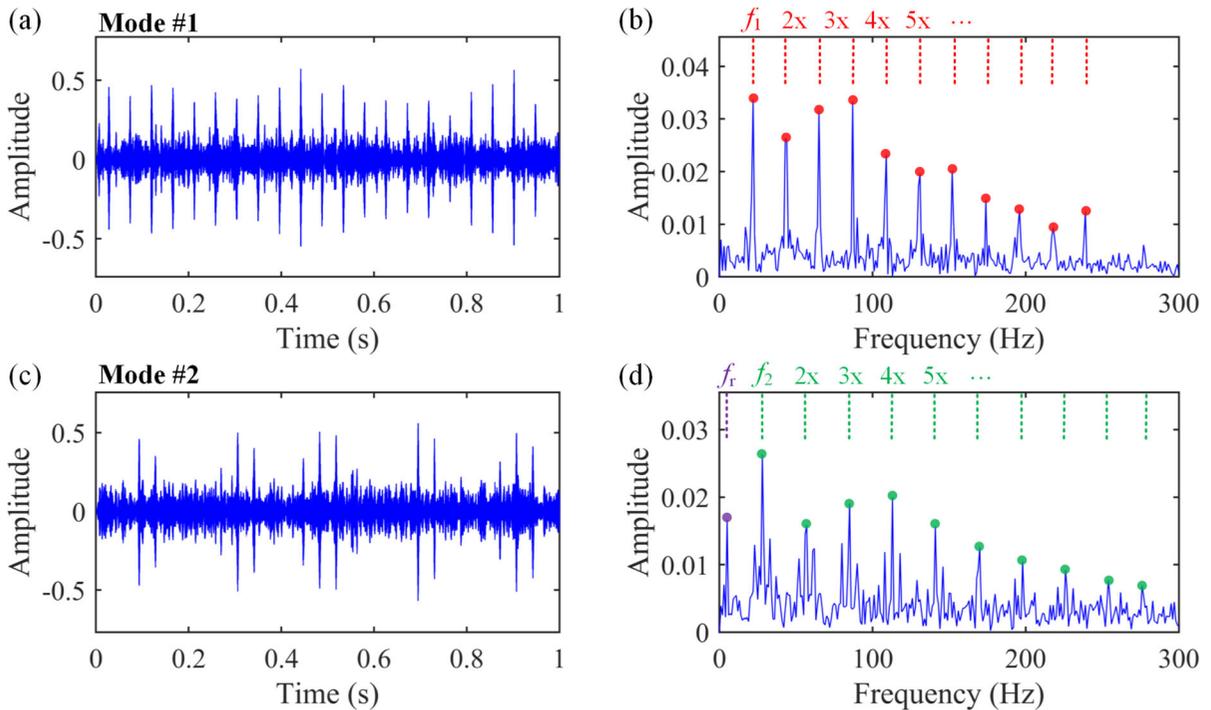
Fig. 13 Decomposed results of the multi-fault signal in simulation case #1 by the proposed AFMD: **a, c** time domain; **b, d** SES
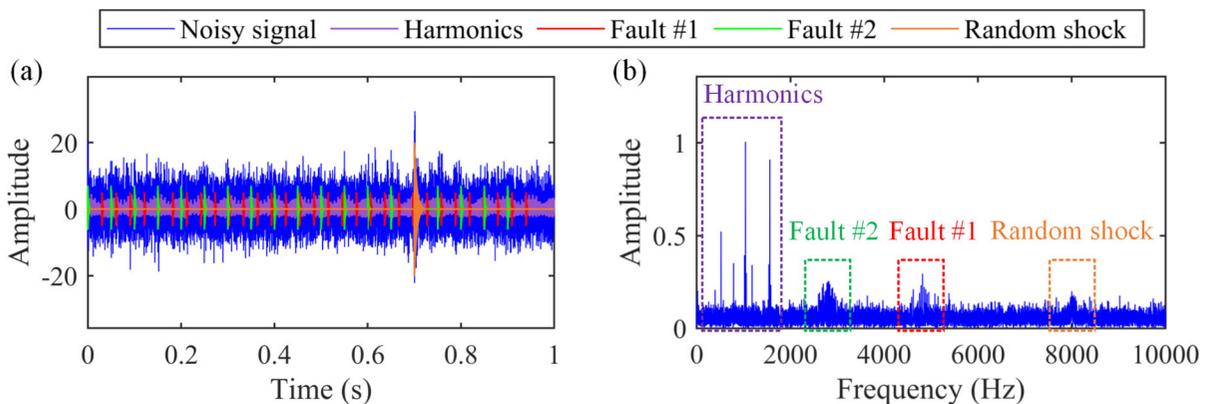
Fig. 14 a Temporal waveform and **b** its frequency spectrum of the multi-fault signal in simulation case #2

The filter length of the proposed AFMD is adaptively chosen as $L = 169$ which achieves the maximum weighted SEHNR value of 1.425. According to Fig. 19, one can find that a small filter length difference might lead to a large decomposition result, which also verifies the finding that filter length could highly affect the final decomposition performance. Nevertheless, the overall trend presented in Fig. 19 is consistent with the diagram in Fig. 12, which might indicate that the decomposition performance will not gradually

increase with a linear trend in a specific selection range. Therefore, the proposed strategy for filter length selection that balances computation burden and effectiveness of decomposition can achieve a good result demonstrated by the two simulation cases in Sect. 4 and will be further tested by two experimental scenarios in the following section.

Figure 20 depicts the decomposed result of the proposed AFMD with the selected filter length of 169. Without pre-defining the desired mode number, the
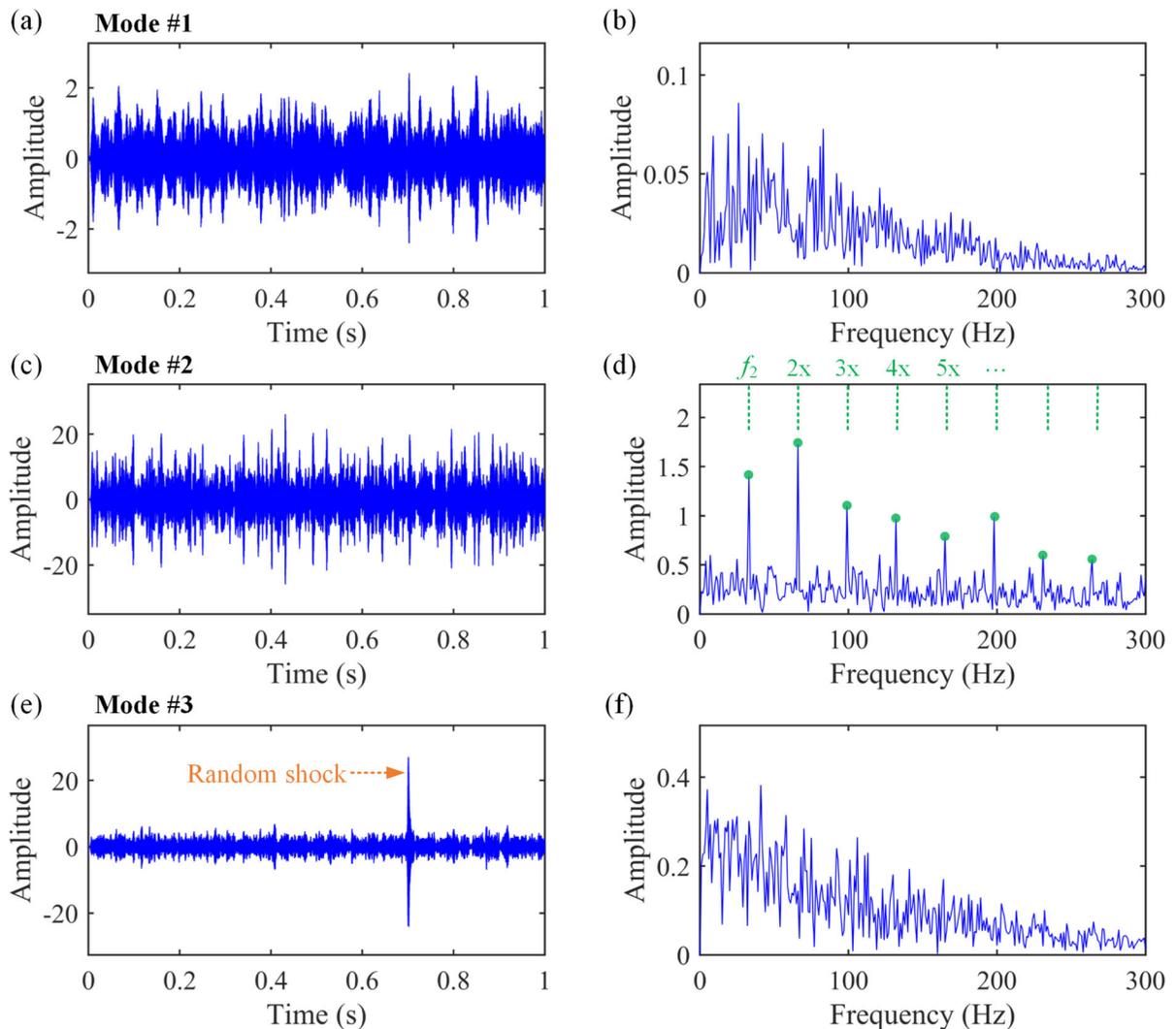
Fig. 15 Decomposed results of the multi-fault signal in simulation case #2 by original FMD with $K_d = 3$ and $L = 40$: **a, c, e** time domain; **b, d, f** envelope spectra

two localized fault features, i.e., f_1 with its multiples ($2\times$, $3\times$, etc.) and f_2 with its multiples ($2\times$, $3\times$, etc.), are clearly extracted into the obtained two modes with high SNR. Based on the two simulation cases, it can be concluded that the proposed AFMD method has remarkable capability of detecting and extracting multiple localized faults of rotating machinery even under strong periodic harmonic interference.

Table 2 presents the consumed time of the original VMD, FMD, and the proposed AFMD methods. The adopted platform is a desktop with a 2.10 GHz Intel I Core I i7-12700 CPU and 32.0 GB RAM, and all algorithms are coded and tested via MATLAB

software (R2022a). From the calculation results of the two simulation cases, VMD with mode number of 2 achieves the best computational efficiency. The calculation time of FMD with two modes (7.48 s for simulation case #1 and 6.63 s for simulation case #2) is comparable to that of FMD with three modes (8.40 s for simulation case #1 and 6.56 s for simulation case #2), but higher than that of VMD. The proposed AFMD respectively consumes 41.19 s and 20.11 s for these two simulation cases, which is a higher computational load compared to FMD. This is because the filter length is a key factor affecting the calculation efficiency, and the filter length of the original FMD is
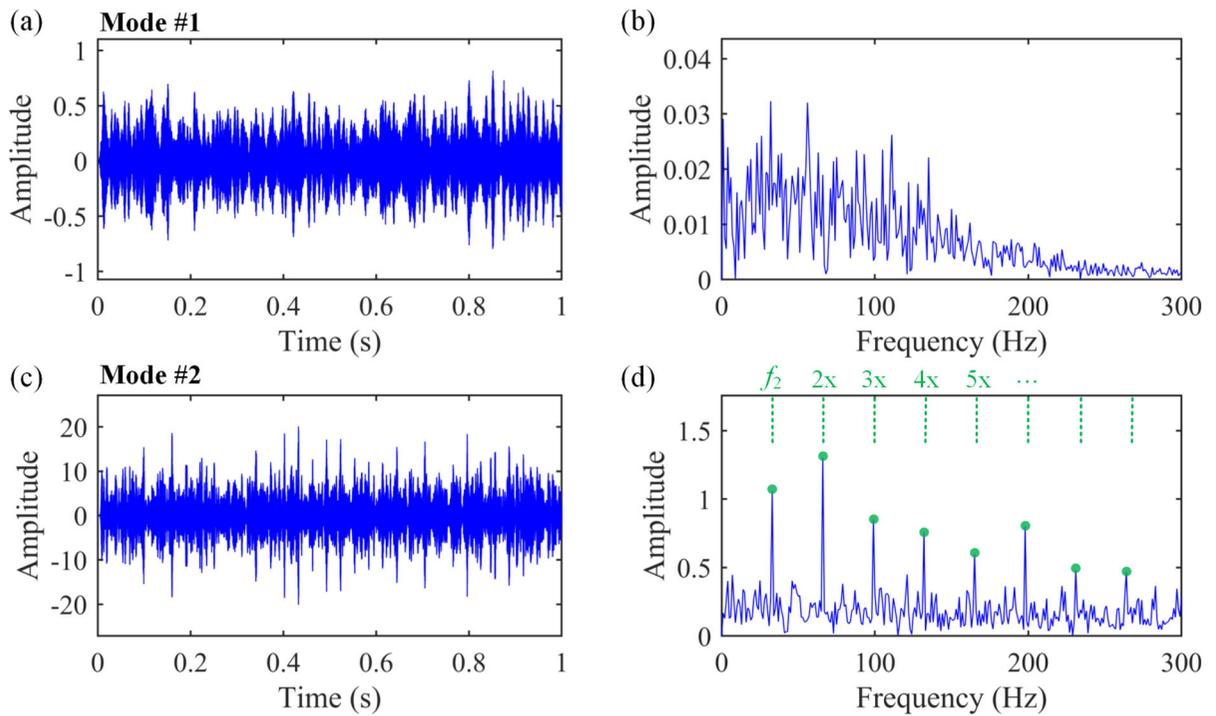
Fig. 16 Decomposed results of the multi-fault signal in simulation case #2 by original FMD with $K_d = 2$ and $L = 40$: **a, c** time domain; **b, d** envelope spectra
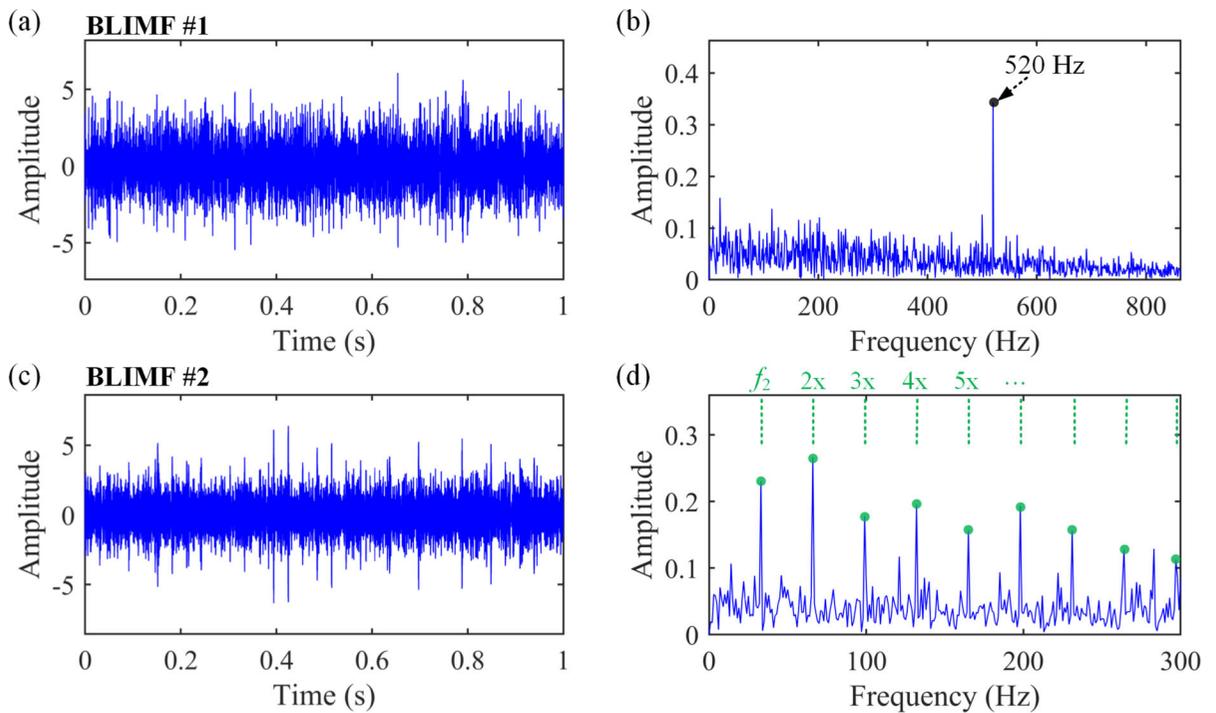
Fig. 17 Decomposed results of the multi-fault signal in simulation case #2 by VMD: **a, c** time domain; **b, d** envelope spectra
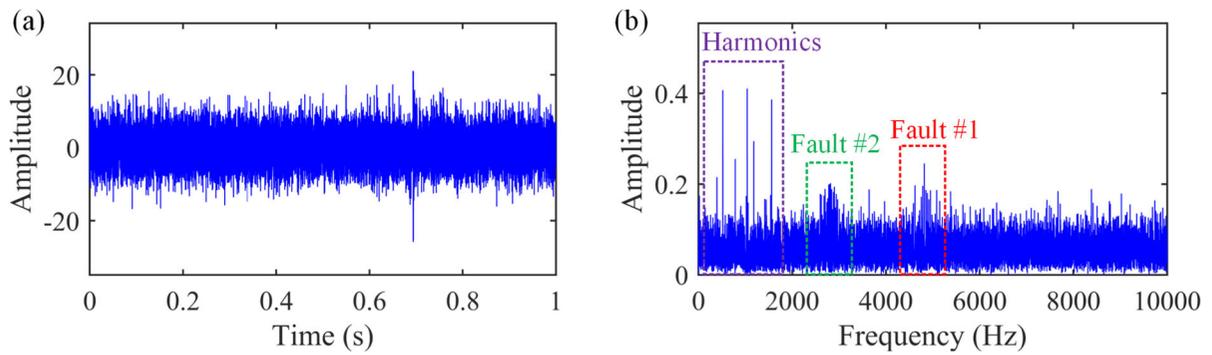
Fig. 18 a Temporal waveform and **b** its frequency spectrum of the multi-fault signal in simulation case #2 after AR model preprocessing
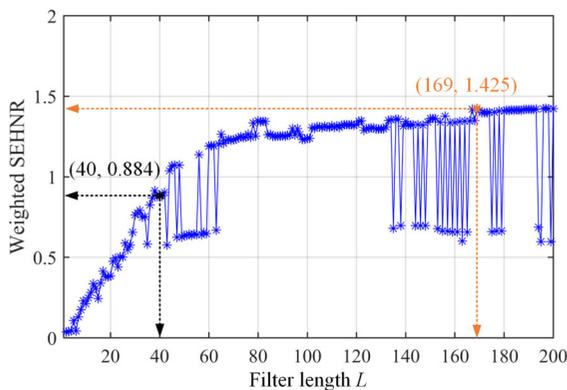
Fig. 19 Weighted SEHNR values of modes decomposed by the proposed AFMD with the change of filter length in simulation case #2

set as 40, while AFMD requires a large number of attempts (the maximum value in this study is 200) to select the filter length. Nevertheless, since the mode number of FMD is selected manually through trial and error, the total time of FMD with more than one attempt is comparable in the same magnitude with that of the proposed AFMD, which could be acceptable for off-line signal processing.

5 Experimental validation

In this section, two experimental vibration signals with multiple localized faults are used for further validating the effectiveness, generality, and superiority of the proposed AFMD method. A multi-fault vibration signal collected from a two-stage gearbox of our own laboratory is firstly used for validation in Sect. 5.1. Then, a run-to-failure bearing signal with

outer race and cage faults from XJTU-SY bearing dataset [35] is applied in Sect. 5.2.

5.1 Experimental case #1: two localized gear faults in a two-stage gearbox

To further validate the effectiveness of the proposed method on weak fault extraction affected by strong fault and periodic harmonics as well as background noise, this subsection introduces a complex vibration signal within two localized gear faults collected from a two-stage gearbox in our own experimental setup. The layout of the experimental apparatus and details of the faulty gears are both shown in Fig. 21. The driving gear mounted on the input shaft was seeded in a spall defect (specific size: length × width × depth = 45 mm × 1.8 mm × 1.5 mm) around the pitch line area of one tooth that is chosen as the severe fault, while a crack with 1-mm depth across an entire tooth face of the driven gear supported by the middle shaft was artificially made by electric discharge machining as a weak fault. An accelerometer is attached on the housing of the support bearing on the input shaft near the driving gear. According to design parameters of the test rig and the rotating speed of the input shaft set as 1200 rpm, the fault feature frequencies of both driving gear and driven gear can be calculated as 20 and 12.68 Hz, respectively. The used faulty signal plotted in Fig. 22 is acquired by sampling rate of 25.6 kHz with 64,000 samples (i.e., 2.5 s).

The original FMD method with filter length $L = 40$ and mode number $K_d = 3$ is first applied. From Fig. 23, the fault features of the driving gear (f_1, 2×, 3×, etc.) can be apparently observed in all three modes. By contrast, the fault information of the driven
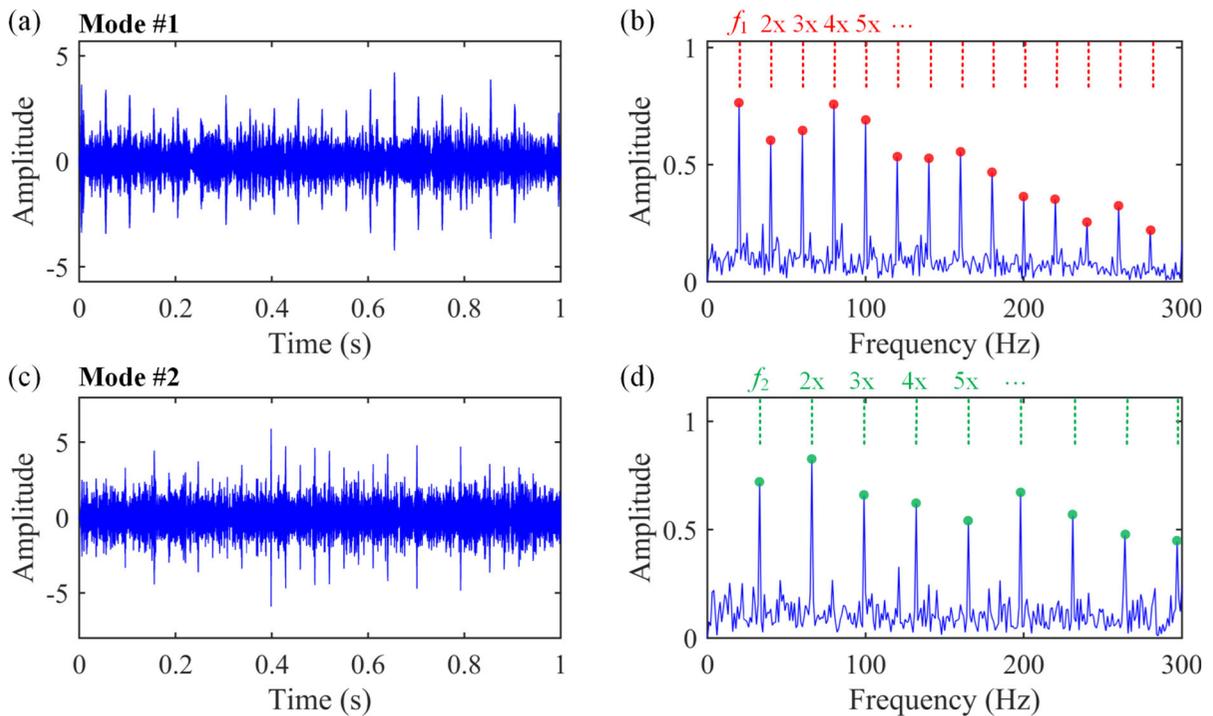
Fig. 20 Decomposed results of the multi-fault signal in simulation case #2 by the proposed AFMD: **a, c** time domain; **b, d** SES

Table 2 Consumed time of the simulation and experimental cases

Methods	Consuming time (s)			
	Simulation case #1	Simulation case #2	Experimental case #1	Experimental case #2
VMD (2)	0.47	0.52	2.23	0.71
FMD (2)	7.48	6.63	37.75	18.31
FMD (3)	8.40	6.56	39.68	17.74
FMD (2) + FMD (3)	15.88	13.18	77.43	36.05
AFMD	41.19	20.11	118.89	42.54

VMD (2) means that the corresponding mode number is set as 2, with similar definitions for FMD (2) and FMD (3)

gear (f_2, 2×, 3×, etc.) is relatively weak. In Mode #1, the fault characteristic frequencies f_2 and $4f_2$ are obvious, which might be regarded as the fault evidence. However, only the basic fault feature frequency f_2 is detected in the envelope spectrum of Mode #2 without any consecutive harmonics, while a similar result appears in Mode #3 with only $4f_2$ presented, thus being hard to be used as significant fault symptoms. Besides, one of the last two modes is redundant because of the high similarity of contents.

When setting mode number $K_d = 2$, the decomposed components shown in Fig. 24 are similar to

those of using $K_d = 3$, where the main difference is the number of modes. It can be found that although the FMD is capable of extracting the severe localized fault features, the consecutive weak fault signature cannot be detected in the decomposed modes. Moreover, FMD still faces the challenge of eliminating redundant modes effectively in practical applications.

VMD is then conducted on the same multi-fault signal. A very similar result to that of FMD has been acquired through VMD, as shown in Fig. 25b, in which the fault characteristic frequencies f_2 and $4f_2$ as well as the first two fault feature harmonics of Fault #1

Fig. 21 Entire experimental apparatus and layout of the two-stage gearbox
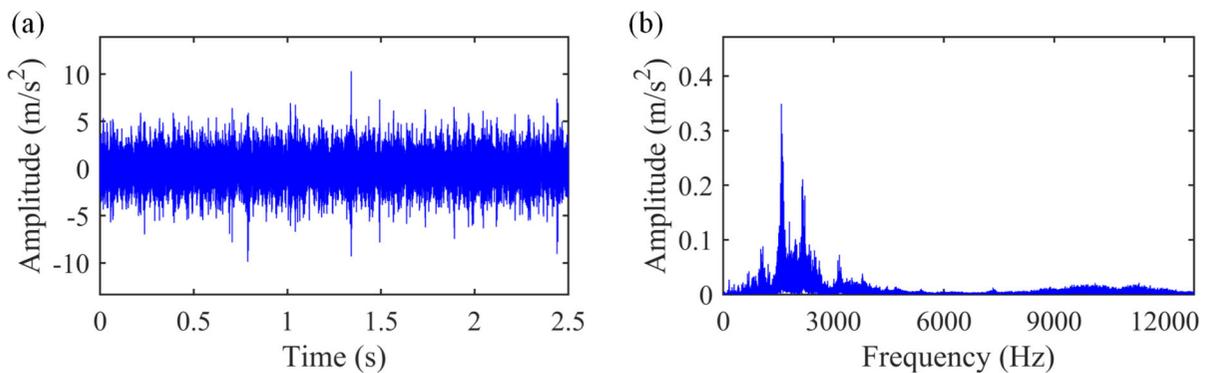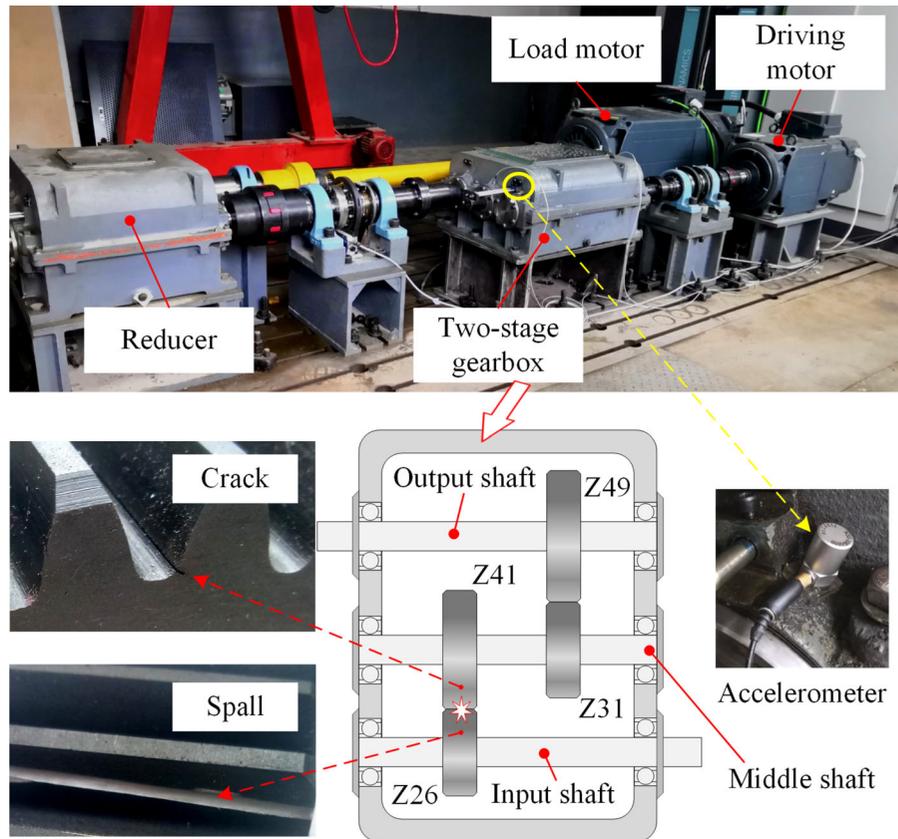
Fig. 22 **a** Temporal waveform and **b** its frequency spectrum of the multi-fault signal in experimental case #1

are obvious. Different from Figs. 23 and 24, the fault features of Fault #1 (f_1, 2×, 3×, 4×) appear in BLIMF #2 with strong background noise, as shown in Fig. 25d.

Figure 26 presents the residual signal after reducing periodic harmonics by means of AR model with model order of 170. By comparison with the raw vibration signal shown in Fig. 22, the fault-induced

impulses are more obvious in the time domain and the amplitudes of spectral lines centered around 2500 Hz are significantly decreased from over 0.3 to around 0.04, which further proves the effectiveness of the AR model process.

Similar to all the aforementioned cases, the filter length of the proposed AFMD method is also determined adaptively as $L = 199$ that corresponds to the
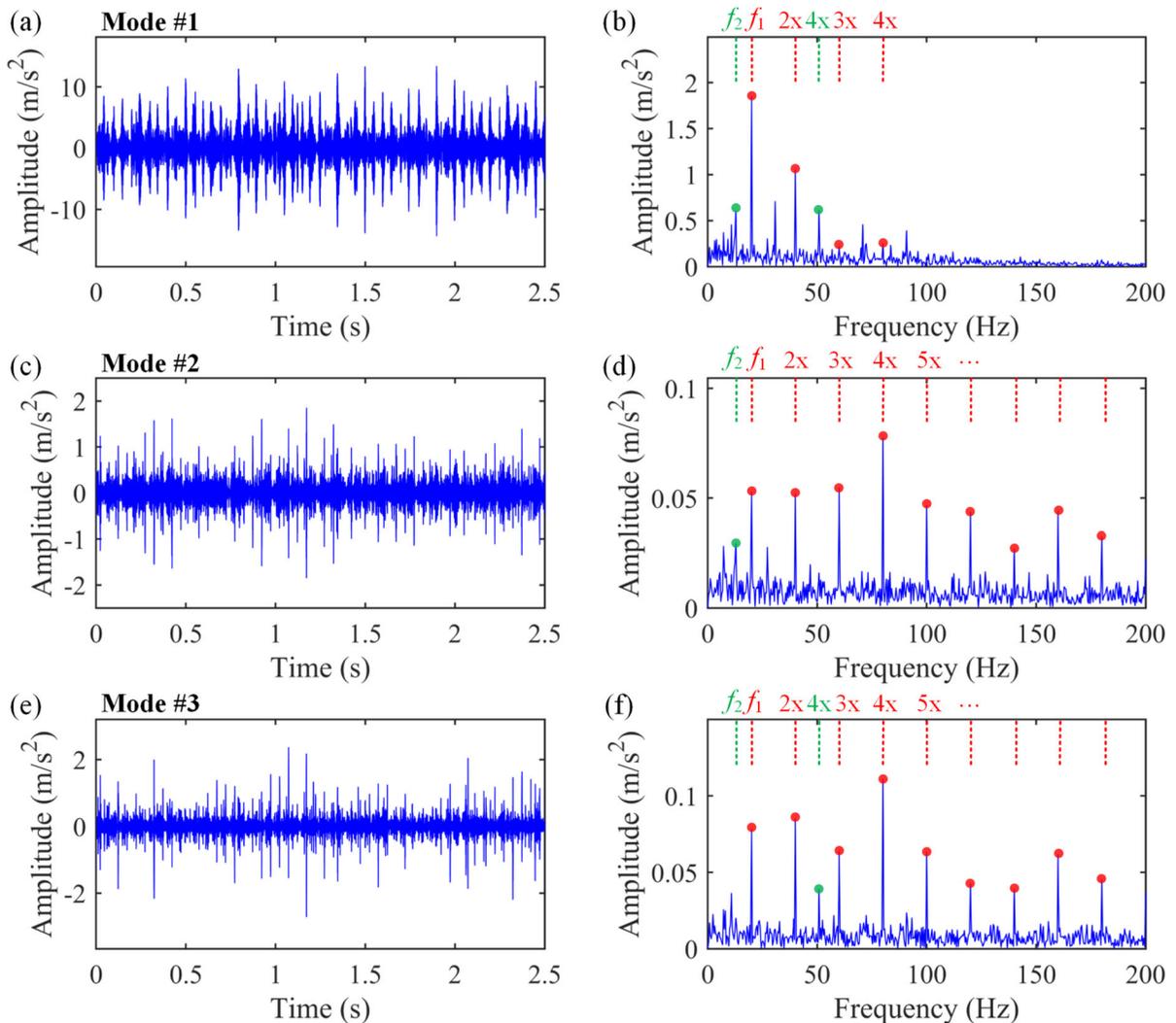
Fig. 23 Decomposed results of the multi-fault signal in experimental case #1 by original FMD with $K_d = 3$ and $L = 40$: **a, c, e** time domain; **b, d, f** envelope spectra

maximum weighted SEHNR of 1.177 of the decomposed modes, as shown in Fig. 27. Figure 28 shows two modes decomposed via the proposed AFMD method. The fault features of the driving gear f_1 and its multiples ($2\times$, $3\times$, etc.) are clearly extracted in the SES of Mode #2 and also obvious in the time domain, as shown in Fig. 28c, d. In Mode #1, from Fig. 28a, the fault features in the time domain are not significant, and seem to be affected by random impulses. In the SES displayed in Fig. 28b, besides the fault symptoms of the driving gear, the fault characteristic frequencies of the driven gear f_2 and $2\times$, $3\times$, etc., could be found. Nevertheless, the unexpected interference lines also

appear in the SES diagram. This case could verify the effectiveness of the proposed method on multi-fault signature extraction and weak fault detection from complicated vibration signals, but its ability on weak fault feature extraction still needs to be further enhanced, as well as the ability on multi-fault feature separation.

5.2 Experimental case #2: two localized faults in a run-to-failure bearing

In this subsection, a bearing vibration signal with two localized faults from the publicly available XJTU-SY
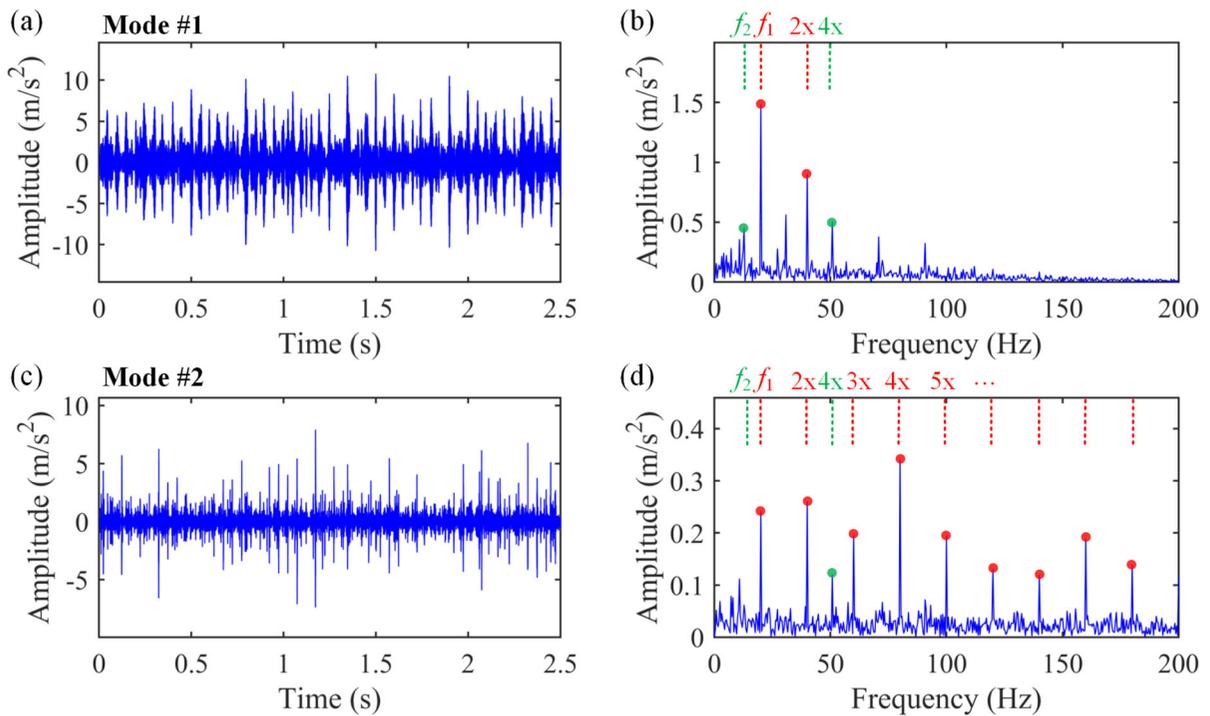
Fig. 24 Decomposed results of the multi-fault signal in experimental case #1 by original FMD with $K_d = 2$ and $L = 40$: **a, c** time domain; **b, d** envelope spectra
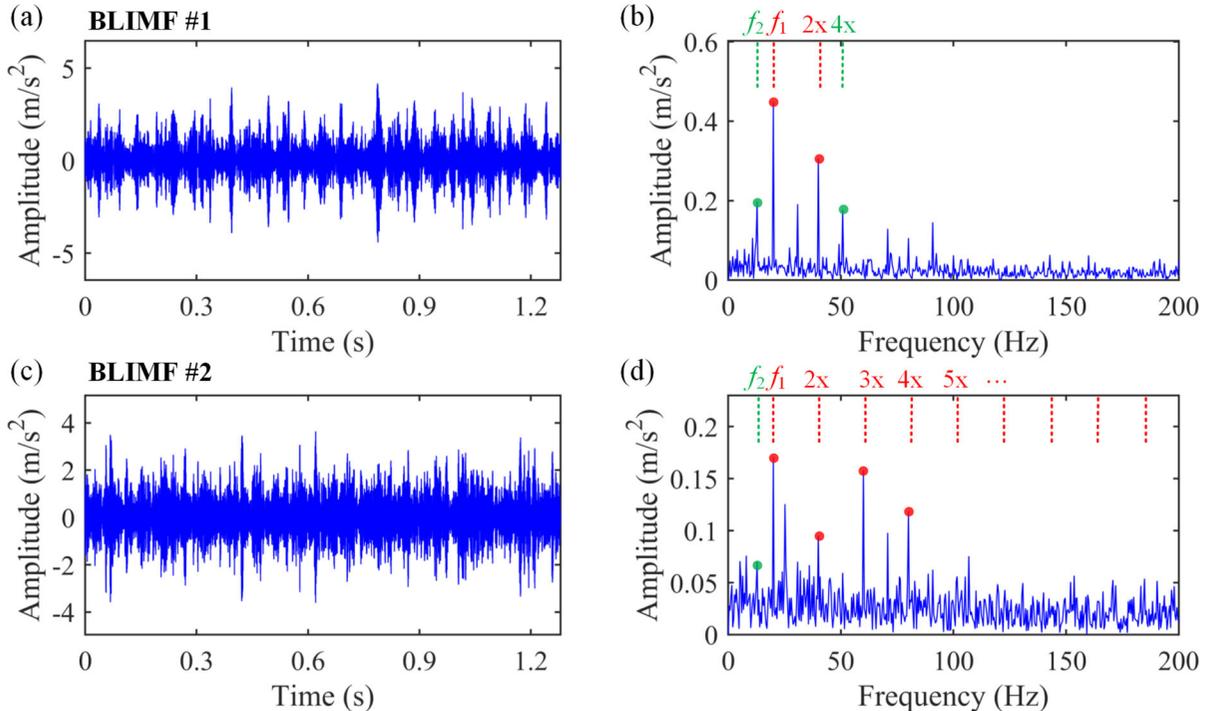
Fig. 25 Decomposed results of the multi-fault signal in experimental case #1 by VMD: **a, c** time domain; **b, d** envelope spectra
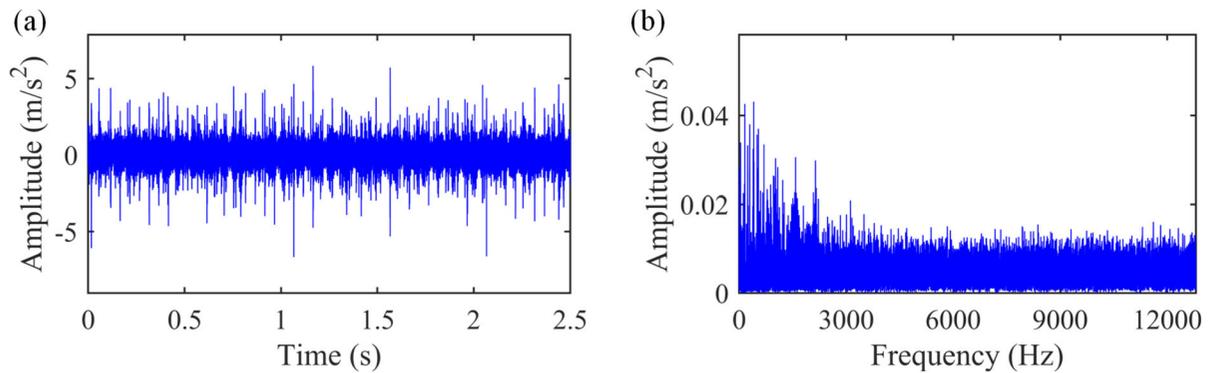
Fig. 26 a Temporal waveform and **b** its frequency spectrum of the multi-fault signal in experimental case #1 after AR model preprocessing
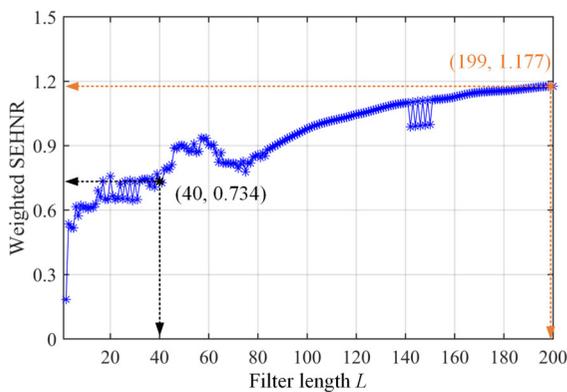
Fig. 27 Weighted SEHNR values of modes decomposed by the proposed AFMD with the change of filter length in experimental case #1

dataset [35] is introduced for performance evaluation of the proposed AFMD. The run-to-failure vibration signals were collected from an accelerometer vertically attached on the housing of the tested bearing, with sampling frequency set as 25.6 kHz and sampling time of 1.28 s for one record epoch. The rotating speed of the tested bearing is 2250 rpm and the radial force applied to the tested bearing housing is 11 kN. The 127th file of Bearing 2_3 dataset is selected as the multi-fault signal. It should be noted that only cage fault occurs in the Bearing 2_3 dataset which is described from the dataset provider [35]; however, López et al. [36] found evidence of the outer race fault which appears from epoch 127 to the end. Therefore, in this study, the vibration signal containing outer race fault and cage fault of the file number 127 in Bearing 2_3 dataset is used for further analysis. Note that according to designing parameters of the test bearing,

feature frequencies of the outer race fault and the cage fault are computed as 116.41 and 14.84 Hz, respectively.

Figure 29 depicts the temporal waveform and frequency spectrum of the used multi-fault signal with 1.28 s. It is difficult to remarkably observe regular impulses with periodicity from Fig. 29a. In Fig. 29b, the rotating frequency f_r and its multiples are clearly found and could be termed as the periodic harmonic interference, which corresponds to the simulation cases in Sect. 4.

The original FMD method is implemented to decompose the multi-fault signal of experimental case #1 at first, with filter length $L = 40$ and desired mode number $K_d = 3$. As shown in Fig. 30, periodic harmonics of shaft rotation, i.e., f_r and its two harmonics 2× and 3×, are predominant in Mode #1, while fault characteristics of cage f_2 along with its inconsecutive multiples appear in the envelope spectra of both Mode #2 and Mode #3.

When changing the mode number as $K_d = 2$, the information contained in decomposed Mode #1 and Mode #2 shown in Fig. 31 are close to Mode #1 and Mode #3 of Fig. 30. However, the fault signatures of outer race could not be extracted in the decomposition modes by means of FMD. From this observation, it can be seen that the original FMD is still sensitive to periodic harmonics induced by shaft rotation and is prone to missed diagnosis, consistent with the conclusion drawn through the simulation cases in Sect. 4.

VMD is also applied to the experimental case #2. As shown in Fig. 32b, one cannot find any fault evidence in the envelope spectrum of BLIMF #1. Meanwhile, only fault feature frequency f_1 clearly
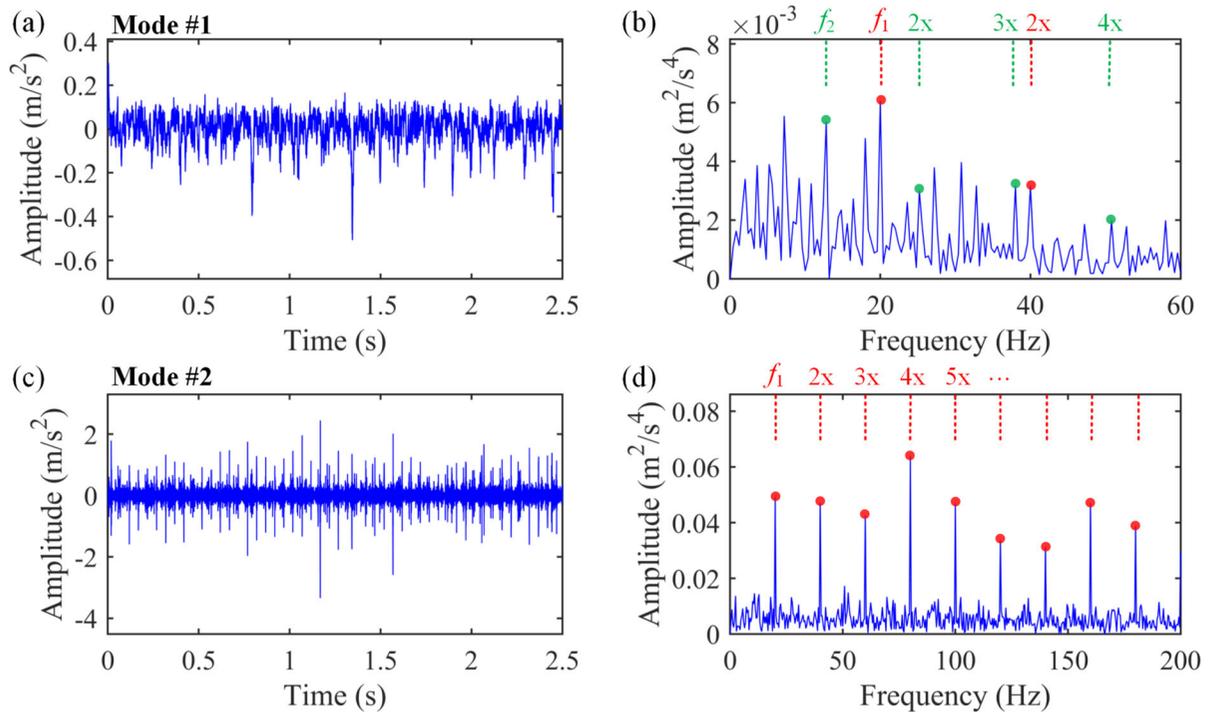
Fig. 28 Decomposed results of the multi-fault signal in experimental case #1 by the proposed AFMD: **a, c** time domain; **b, d** SES
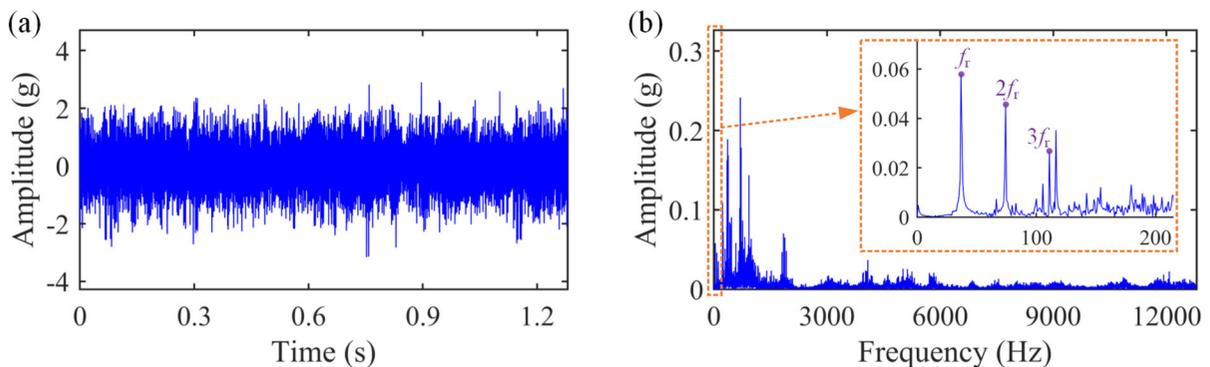
Fig. 29 **a** Temporal waveform and **b** its frequency spectrum of the multi-fault signal in experimental case #2

appears in Fig. 32d. However, the fault signatures of cage (f_2 and its harmonics) could not be extracted in the decomposition modes by VMD. In other words, VMD fails to compete the bearing multi-fault diagnosis task.

Figure 33 plots the temporal waveform and frequency spectrum of the residue of the multi-fault signal preprocessed by AR model with model order of 66. By comparison with Fig. 29b, amplitudes of the periodic harmonics induced by the rotating shaft are decreased by around ten times from near 0.3 to 0.03, as

presented in Fig. 33b. Besides, from the perspective of frequency spectrum, the overall magnitudes of the raw signal and the filtered residue before and after AR pretreatment are also significantly different with removal of the deterministic harmonics, thus aiding to enhance the fault-related signatures.

The filter length of 156 that achieves the maximum weighted SEHNR value of the decomposition results is adaptively selected for the proposed AFMD method, according to Fig. 34. As shown in Fig. 35, two modes are adaptively given by AFMD. The fault
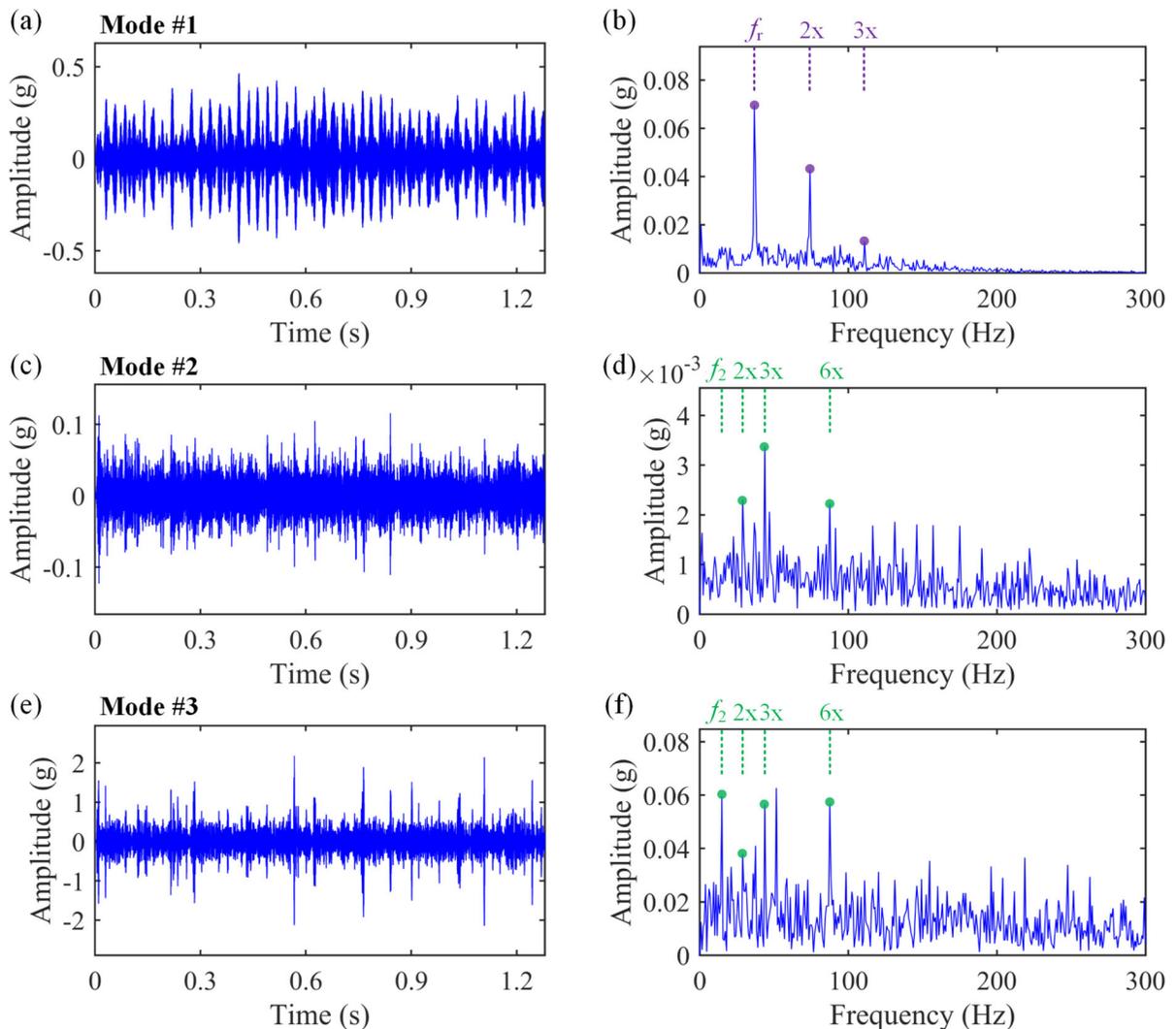
Fig. 30 Decomposed results of the multi-fault signal in experimental case #2 by original FMD with $K_d = 3$ and $L = 40$: **a, c, e** time domain; **b, d, f** envelope spectra

characteristic frequency f_2 and its harmonics are clearly decomposed into Mode #2. Different from the results acquired by the original FMD shown in Figs. 30 and 31, the outer race fault signatures (f_1, 2×, 3×, etc.) are clearly extracted with high SNR. Besides that, the cage fault frequency f_2 and its multiples as well as the side frequency $f_1 + f_2$ are also obvious in Mode #1, demonstrating the capability of the proposed AFMD method on multi-fault feature extraction over the original FMD and VMD.

Regarding the computational burden of all the algorithms, from the results of two experimental cases shown in Table 2, a similar conclusion to that based on

simulation cases could be drawn. VMD has significant computational efficiency, while AFMD takes the most time to complete the calculation but achieves a similar (experimental case #1) or same (experimental case #2) magnitude level with the total time of FMD that requires more attempts for mode number selection. Considering the fault diagnosis results of all the methods, it could be found that the proposed AFMD indeed needs more time but can ensure the accuracy of final diagnosis results, especially for multiple faults. At the same time, when using the original FMD by manually attempting more times to select the mode number, the time advantage of FMD to AFMD would
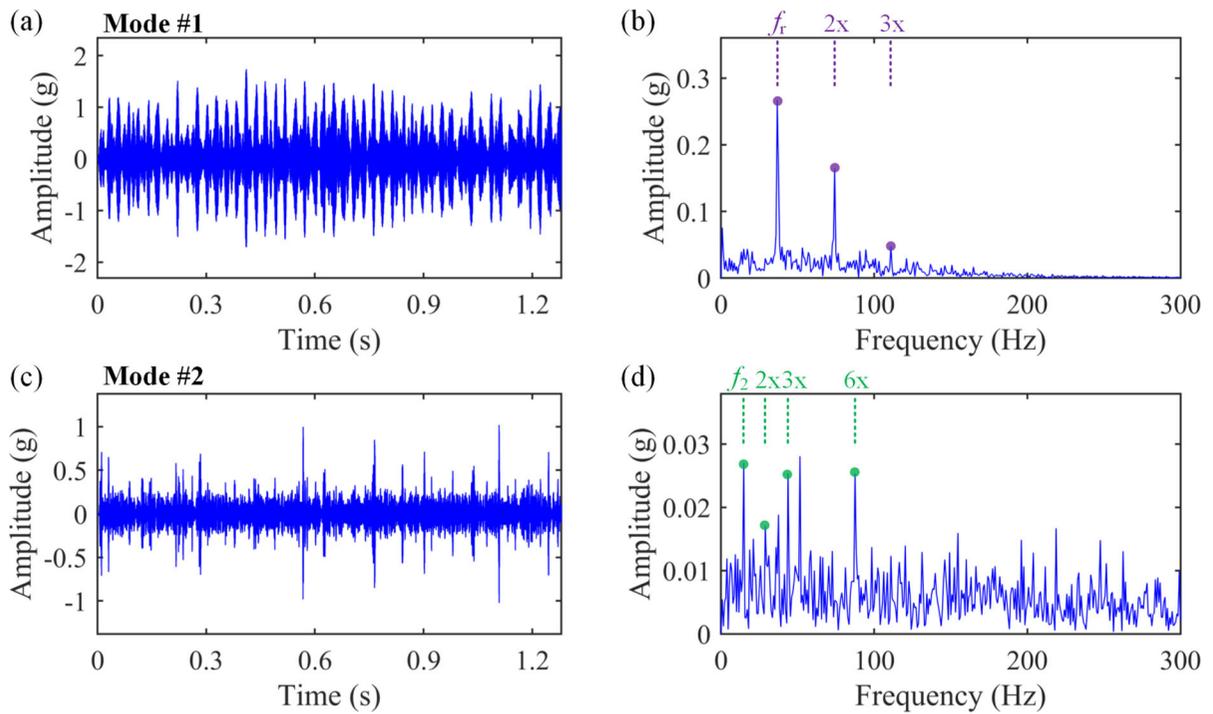
Fig. 31 Decomposed results of the multi-fault signal in experimental case #2 by original FMD with $K_d = 2$ and $L = 40$: **a, c** time domain; **b, d** envelope spectra
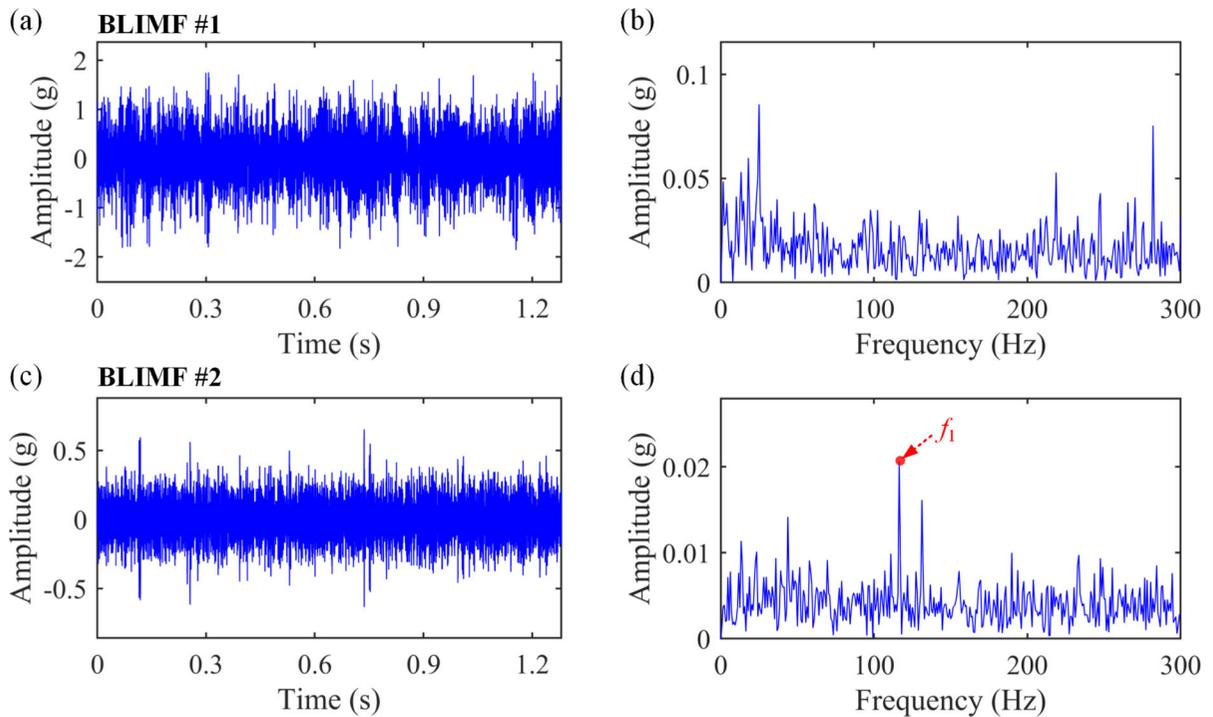
Fig. 32 Decomposed results of the multi-fault signal in experimental case #2 by VMD: **a, c** time domain; **b, d** envelope spectra
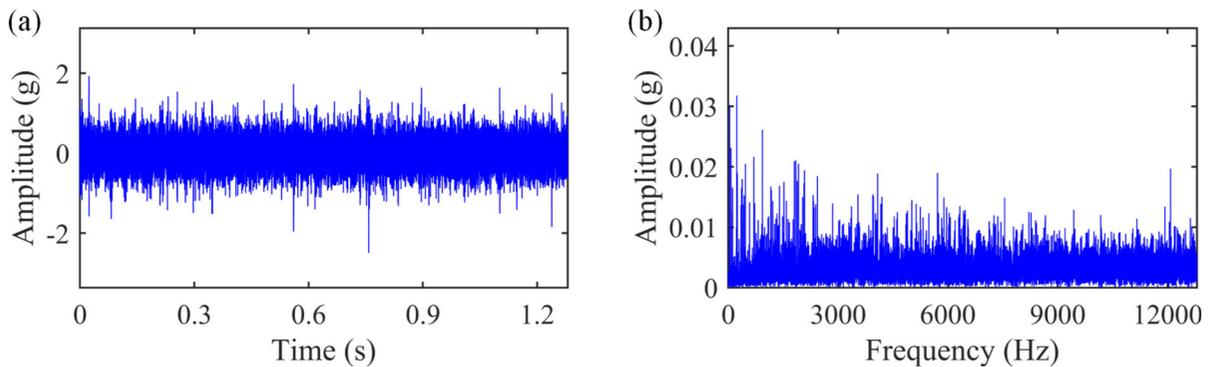
Fig. 33 a Temporal waveform and **b** its frequency spectrum of the multi-fault signal in experimental case #2 after AR model preprocessing
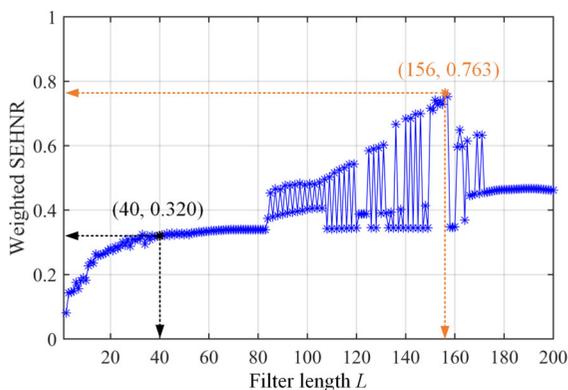
Fig. 34 Weighted SEHNR values of modes decomposed by the proposed AFMD with the change of filter length in experimental case #2

easily disappear. Besides, although VMD is faster, it is not initially designed for extracting the impulsiveness and periodicity of the faulty signal, and the fault diagnosis performance also relies on the proper parameter selection.

6 Discussion and conclusion

As a latest decomposition method, FMD has been already tested by Miao et al. [24] through simulation and experimental datasets and achieved better results over the popular VMD method. However, FMD still faces the challenging task of multi-fault diagnosis under interferences of deterministic harmonics induced by gear meshing and shaft rotation that are rather common components in the vibration signals collected from rotating machinery. Misdiagnosis and

missed diagnosis might appear when conducting FMD to such noisy signals with multiple localized faults on gears or bearings through the simulation case #1 and case #2 as well as experimental case #2. Besides, the main input parameters (i.e., filter length and desired mode number) that unfortunately need to be pre-defined by means of trial and error affect the final decomposition performance of FMD, even with small difference of selection value, which is verified by simulation case #1. At the same time, the current principle for mode selection still cannot thoroughly eliminate the redundant or useless modes, which has been found by simulation case #2, experimental case #1 and case #2. The limitations of FMD are summarized in Sect. 2.2 in detail.

The proposed AFMD is therefore introduced to diagnosis multiple localized faults of gears and bearings robustly. Benefiting from the AR model preprocessing procedure, AFMD initially removes the periodic harmonics and exploits the merit of CK focusing on fault-related impulses. *The core of AFMD is the "adaptive" merit, which is embodied in the aspect of focusing on multi-fault feature extraction adaptively, not being trapped in the dilemma of misdiagnosis or missed diagnosis caused by manual parameter selection.* The adaptive FIR filter bank and the newly proposed mode selection strategy, along with the adaptive determination of the filter length and the desired mode number, all aid to effectively accomplish the adaptive extraction of multiple faults with only useful modes reserved in the final decomposition results. The proposed new mode selection strategy has been proven to be more beneficial to thoroughly removing redundant and useless modes,
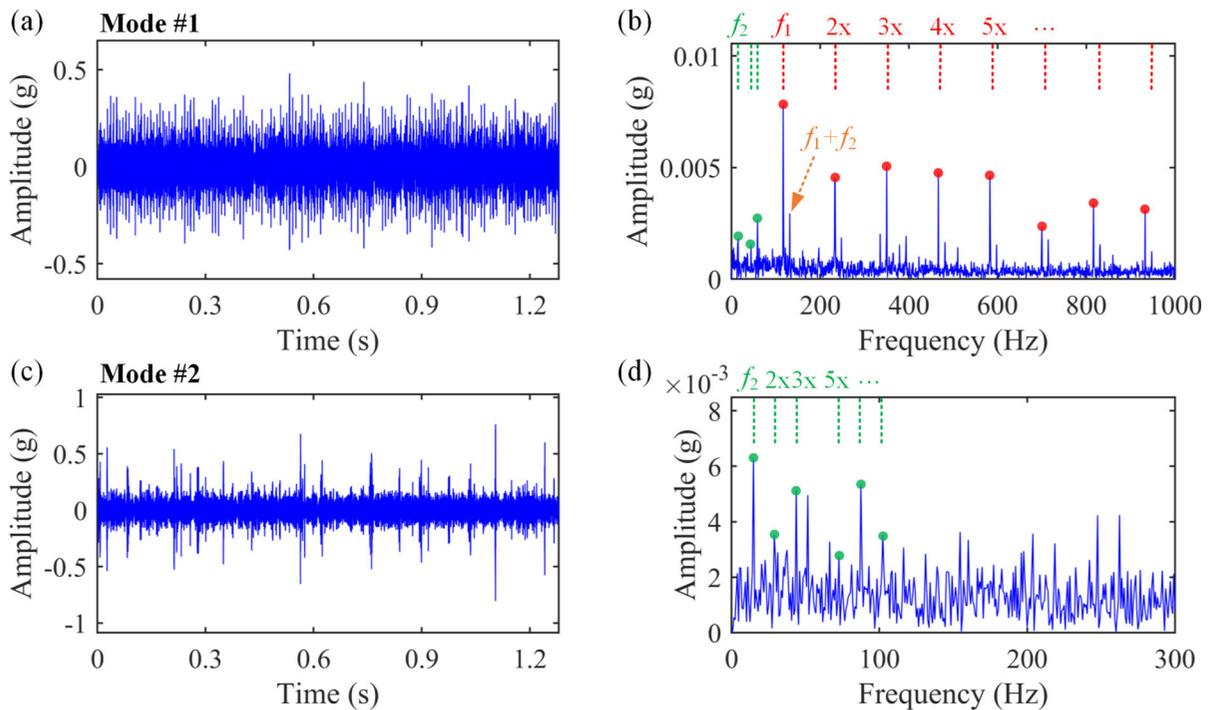
Fig. 35 Decomposed results of the multi-fault signal in experimental case #2 by the proposed AFMD: **a, c** time domain; **b, d** SES

since it considers the signal similarity across multiple domains (i.e., time domain, frequency domain and envelope spectrum domain) as well as the consistency of estimated faulty periods rather than single temporal correlation coefficient. The effectiveness of the proposed AFMD has been validated through two simulated and two experimental scenarios involving gear and bearing multiple localized faults.

Some thoughts on selection of the filter length need to be further clarified. As discussed in Sect. 2.2, different filter length values may significantly affect the final decomposition results, even with a very small difference such as $L = 40$ and $L = 41$ depicted in Figs. 4 and 5. This phenomenon occurs not only in FMD, but also in the proposed AFMD, owing to the intrinsic property of blind deconvolution theory. One can find that the obtained results of the weighted SEHNR values with the change of filter length in the four used validation scenarios exist the similar overall trend. The weighted SEHNR value has a rapid increase as the filter length increases from zero to a relatively small value. While it then gradually climbs to a relatively high level, but may fluctuate to a certain extent, as shown in Figs. 12, 19, 27 and 34. However, the specific weighted SEHNR values and their

fluctuation amplitudes in different simulation and experiment cases are not the same, since the processed signals have different components and complexities. The plots of the weighted SEHNR value with changing filter length are mainly utilized to present that the filter length has a significant impact on the decomposition results. In this study, for simplicity, the maximum values of the filter lengths are all set as 200 in the four validation cases, which can well illustrate the effects of the filter length. One can choose a greater maximum filter length value for further observing its effects on decomposition results. However, it should be noted that the greater the filter length is, the greater the computational burden and distortion probability [24]. This study only provides a practical solution for selection of the filter length, whose effectiveness has been demonstrated, but might not be the best one. Based on the proposed indicator weighted SEHNR, the intelligent optimization algorithms such as particle swarm optimization (PSO) and genetic algorithm (GA) could also be applied to choose the proper filter length value, while how to select the hyper-parameters of these algorithms may become another difficult problem to be solved.

The contributions of this study can be concluded as follows:

(1) A newly proposed AFMD fusing the merits of blind deconvolution theory into the decomposition process is presented as an efficient vibration signal decomposition method for adaptive multi-fault feature extraction of rotating machinery under complex interferences including strong periodic harmonics and random shocks.

(2) A new indicator called weighted SEHNR that is capable of evaluating the multi-fault information of decomposed modes is proposed to guide the selection of filter length in an automatic way, thus guaranteeing the decomposition performance of the proposed AFMD method.

(3) A new mode selection strategy that can evaluate the signal similarity across multiple domains and the consistency of estimated faulty periods is proposed.

Benefiting from the proposal of the mode selection strategy and the adaptive determination of filter length and mode number, the proposed AFMD method avoids the redundant iterations in the mode selection process of the state-of-the-art FMD and the inappropriate tunable parameter selection issue of VMD. The superiority of AFMD on multi-fault feature extraction by using appropriate key parameters was well-demonstrated in this study.

It is also noticed that the two main parameters, filter length and mode number, are automatically selected in this study, while some initial parameters such as the threshold of value ε_T shown in Eq. (20), and the searching range of filter length still need to be preset in advance by users according to various working conditions. The applied values corresponding to the initial parameters in this study could be taken as default values for practical cases with similar operation conditions as this study.

In the future, the adaption of the proposed method will concentrate on the multi-fault quantitative diagnosis under variable working conditions from industrial applications, which will have wide application prospects in practice and needs further investigations. In addition, the ability of the proposed AFMD method on extraction of weak fault features and separation of multi-fault features without mixtures should also be further enhanced in the following work.

Data availability Data of this study will be made available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declared no potential conflicts of interest concerning the research, authorship, and/or publication of this paper.

References

1. Salameh, J.P., Cauet, S., Etien, E., Sakout, A., Rambault, L.: Gearbox condition monitoring in wind turbines: a review. Mech. Syst. Sig. Process. **111**, 251–264 (2018)
2. Wang, T., Han, Q., Chu, F., Feng, Z.: Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: a review. Mech. Syst. Sig. Process. **126**, 662–685 (2019)
3. Goyal, D., Pabla, B.S.: The vibration monitoring methods and signal processing techniques for structural health monitoring: a review. Arch. Comput. Methods Eng. **23**(4), 585–594 (2016)
4. Zhang, J., Zhang, Q., Qin, X., Sun, Y., Zhang, J.: Gearbox compound fault diagnosis based on a combined MSGMD–MOMEDA method. Meas. Sci. Technol. **33**(6), 065102 (2022)
5. Jin, Z., He, D., Lao, Z., Wei, Z., Yin, X., Yang, W.: Early intelligent fault diagnosis of rotating machinery based on IWOA-VMD and DMKELM. Nonlinear Dyn. (2022)
6. Wang, X., Si, S., Li, Y.: Hierarchical diversity entropy for the early fault diagnosis of rolling bearing. Nonlinear Dyn. **108**(2), 1447–1462 (2022)
7. Zhou, X., He, X., Peng, D., Hou, Y., Liu, Q.: A practical methodology for enhancement and detection of transient faults in a gearbox without prior fault feature information. Meas. Sci. Technol. **32**(3), 035116 (2021)

8. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **454**(1971), 903–95 (1998)

9. Singh, D.S., Zhao, Q.: Pseudo-fault signal assisted EMD for fault detection and isolation in rotating machines. Mech. Syst. Sig. Process. **81**, 202–218 (2016)

10. Lei, Y., Lin, J., He, Z., Zuo, M.J.: A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech. Syst. Signal Process. **35**(1–2), 108–126 (2013)

11. Wu, Z., Huang, N.E.: Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv. Adapt. Data Anal. **01**(01), 1–41 (2009)

12. Yeh, J.-R., Shieh, J.-S., Huang, N.E.: Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv. Adapt. Data Anal. **02**(02), 135–156 (2010)

13. Torres, M.E., Colominas, M.A., Schlotthauer, G., Flandrin, P.: A complete ensemble empirical mode decomposition with adaptive noise. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4144–7 (2011)

14. Colominas, M.A., Schlotthauer, G., Torres, M.E.: Improved complete ensemble EMD: a suitable tool for biomedical signal processing. Biomed. Signal Process. Control **14**(1), 19–29 (2014)

15. Smith, J.S.: The local mean decomposition and its application to EEG perception data. J. R. Soc. Interface **2**(5), 443–454 (2005)

16. Frei, M.G., Osorio, I.: Intrinsic time-scale decomposition: time–frequency–energy analysis and real-time filtering of non-stationary signals. Proc. R. Soc. Ser. A. **463**(2078), 321–342 (2007)

17. Zheng, J., Cheng, J., Yang, Y.: A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. Mech. Mach. Theory **70**, 441–453 (2013)

18. Cicone, A., Liu, J., Zhou, H.: Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis. Appl. Comput. Harmon. Anal. **41**(2), 384–411 (2016)

19. Gilles, J.: Empirical wavelet transform. IEEE Trans. Signal Process. **61**(16), 3999–4010 (2013)

20. Dragomiretskiy, K., Zosso, D.: Variational mode decomposition. IEEE Trans. Signal Process. **62**(3), 531–544 (2014)

21. Zhang, X., Miao, Q., Zhang, H., Wang, L.: A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery. Mech. Syst. Signal Process. **108**, 58–72 (2018)

22. He, X., Zhou, X., Yu, W., Hou, Y., Mechefske, C.K.: Adaptive variational mode decomposition and its application to multi-fault detection using mechanical vibration signals. ISA Trans. **111**, 360–375 (2021)

23. Kedadouche, M., Thomas, M., Tahan, A.: A comparative study between empirical wavelet transforms and empirical mode decomposition methods: application to bearing defect diagnosis. Mech. Syst. Signal Process. **81**, 88–107 (2016)

24. Miao, Y., Zhang, B., Li, C., Lin, J., Zhang, D.: Feature mode decomposition: new decomposition theory for rotating machinery fault diagnosis. IEEE Trans. Ind. Electron. **70**(2), 1949–1960 (2023)

25. Miao, Y., Zhang, B., Lin, J., Zhao, M., Liu, H., Liu, Z., Li, H.: A review on the application of blind deconvolution in machinery fault diagnosis. Mech. Syst. Signal Process. **163**, 108202 (2022)

26. Sawalhi, N., Randall, R.B., Endo, H.: The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. Mech. Syst. Signal Process. **21**(6), 2616–2633 (2007)

27. Lee, J.Y., Nandi, A.K.: Extraction of impacting signals using blind deconvolution. J. Sound Vib. **232**(5), 945–962 (2000)

28. Buzzoni, M., Antoni, J., D'Elia, G.: Blind deconvolution based on cyclostationarity maximization and its application to fault identification. J. Sound Vib. **432**, 569–601 (2018)

29. Randall, R.B., Antoni, J.: Rolling element bearing diagnostics—a tutorial. Mech. Syst. Signal Process. **25**(2), 485–520 (2011)

30. Wang, D., Tse, P.W., Tsui, K.L.: An enhanced Kurtogram method for fault diagnosis of rolling element bearings. Mech. Syst. Signal Process. **35**(1), 176–199 (2013)

31. Ho, D., Randall, R.B.: Optimisation of bearing diagnostic techniques using simulated and actual bearing fault signals. Mech. Syst. Signal Process. **14**(5), 763–788 (2000)

32. Miao, Y., Zhao, M., Lin, J., Xu, X.: Sparse maximum harmonics-to-noise-ratio deconvolution for weak fault signature detection in bearings. Meas. Sci. Technol. **27**(10), 105004 (2016)

33. Xu, X., Zhao, M., Lin, J., Lei, Y.: Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis. Measurement **91**, 385–397 (2016)

34. He, X., Liu, Q., Yu, W., Mechefske, C.K., Zhou, X.: A new autocorrelation-based strategy for multiple fault feature extraction from gearbox vibration signals. Measurement **171**, 108738 (2021)

35. Wang, B., Lei, Y., Li, N., Li, N.: A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. IEEE Trans. Reliab. **69**(1), 401–412 (2020)

36. López, C., Wang, D., Naranjo, Á., Moore, K.J.: Box-cox-sparse-measures-based blind filtering: understanding the difference between the maximum kurtosis deconvolution and the minimum entropy deconvolution. Mech. Syst. Signal Process. **165**, 108376 (2022)