An Enhanced VMD with the Guidance of Envelope Negentropy Spectrum for Bearing Fault Diagnosis

Haien Wang, Xingxing Jiang, Wenjun Guo, Juanjuan Shi, and Zhongkui Zhu

School of Rail Transportation, Soochow University, Suzhou 215131, China

Correspondence should be addressed to Xingxing Jiang; jiangxx@suda.edu.cn

Received 9 October 2019; Revised 30 December 2019; Accepted 7 January 2020; Published 31 January 2020

Copyright © 2020 Haien Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Currently, study on the relevant methods of variational mode decomposition (VMD) is mainly focused on the selection of the number of decomposed modes and the bandwidth parameter using various optimization algorithms. Most of these methods utilize the genetic-like algorithms to quantitatively analyze these parameters, which increase the additional initial parameters and inevitably the computational burden due to ignoring the inherent characteristics of the VMD. From the perspective to locate the initial center frequency (ICF) during the VMD decomposition process, we propose an enhanced VMD with the guidance of envelope negentropy spectrum for bearing fault diagnosis, thus effectively avoiding the drawbacks of the current VMD-based algorithms. First, the ICF is coarsely located by envelope negentropy spectrum (ENS) and the fault-related modes are fast extracted by incorporating the ICF into the VMD. Then, the fault-related modes are adaptively optimized by adjusting the bandwidth parameters. Lastly, in order to identify fault-related features, the Hilbert envelope demodulation technique is used to analyze the optimal mode obtained by the proposed method. Analysis results of simulated and experimental data indicate that the proposed method is effective to extract the weak faulty characteristics of bearings and has advantage over some advanced methods. Moreover, a discussion on the extension of the proposed method is put forward to identify multicomponents for broadening its applied scope.

1. Introduction

The rolling bearing is a key part of rotating machines, whose performance determines the operation efficiency and reliability of the whole mechanical equipment [1], such as aircraft engines, high-speed trains, water-pumps, electric fans, and so on. When the rolling bearing is in early failure, the fault information is very weak. Furthermore, many noises and other environmental factors make the fault information difficult to be detected. If this early failure continues to develop, it could reduce the operating stability of the machines, and even cause the catastrophic failures [2, 3]. Therefore, the accurate and reliable fault diagnosis of early local damage for the rolling bearings is of particular significance.

In recent decades, various methods based on vibration signal have been introduced by researchers for the diagnosis of faulty rolling bearings, such as deconvolution analysis [4, 5], wavelet transform [6], spectral kurtosis-based method [7–10], time-frequency analysis [11, 12], and the adaptive signal decomposition methods [13, 14]. Among these methods, the adaptive signal decomposition methods have attracted more and more attention and are the hot topic for bearing fault diagnosis. Huang et al. [15] first proposed an adaptive decomposition algorithm to deal with the nonstationary signals, i.e., empirical modal decomposition (EMD), which can decompose the signal into several monocomponents. However, many drawbacks of the EMD method and its variants, including noise-sensitive and mode mixing [16], limit their effectiveness in extracting the weak bearing fault characteristics. To address these problems, Wu and Huang [17] developed an ensemble EMD (EEMD), while the more computational burden needs to be paid [18]. Afterwards, various advanced adaptive decomposition methods have been developed to process nonstationary signals, such as empirical wavelet transform-based methods...
and local mean decomposition-based methods [21, 22], which have also been widely used in the field of bearing fault diagnosis.

Recently, Dragomiretskiy and Zosso [23] innovatively proposed an adaptive signal decomposition technique, variational mode decomposition (VMD), which can decompose the signals with multicomponents into mono-components. Compared with the traditional adaptive decomposition methods, the VMD method has some obvious advantages as follows: noise suppression, nonrecursive sifting, and adaptively chosen bands. However, the effectiveness of VMD is greatly affected by the initial parameters in practical applications. Therefore, selecting the suitable initial decomposition parameters has become a critical issue in the study of VMD. Many researches [24–29] mainly focused on how to determine the bandwidth parameter and the number of decomposition modes for the fault diagnosis of bearings. However, the current types for initializing the parameters in VMD subject to some prior experience. Hence, further exploration on the decomposing strategy of the VMD method needs to be desired for its better application in bearing fault diagnosis.

Wang and Markert [25] found that the initial center frequency (ICF) has a significant impact on the filter bank and processing results, but its influence on bearing fault diagnosis results is rarely discussed in the accessible literature. Besides, most of the current methods employ only one balance parameter to match the bandwidths of all extracted modes, but each latent mode has its individual theoretical bandwidth. In order to solve these problems, our team [30] has proposed a novel ICF-guided VMD method to improve the ability of weak damage feature extraction of rotating machines. It was found that a fine ICF would help to perform the VMD easily without much prior experience. As a result, an energy fluctuation spectrum (EFS) index is used to locate the ICF and then optimize the balance parameter to match the bandwidth of fault-related component by incorporating the center frequency (CF). However, how to locate the ICF is significant for the successful application of VMD. Through research, we found that the reliability of the EFS index is limited by the parameters used in the time-frequency representation (TFR). In this paper, we propose an enhanced VMD using the envelope negentropy spectrum (ENS) as a guidance to mitigate the influence of the parameters used in TFR and strengthen the adaptability of the VMD. First, the TFR of the raw mechanical signal is obtained by short-time Fourier transform (STFT). Then, the ENS is constructed based on the results of TFR. Third, use the VMD method to extract the faulty modes on the basis of ICF located by ENS. Finally, the extracted modes are optimized to maximize the faulty information by adjusting the bandwidth parameter, and the Hilbert envelope demodulation technique is carried out on the optimal mode, from which the bearing faulty characteristics can be accurately detected. Compared with the original ICF-guided VMD method, the proposed method has an advantage of effectively avoiding the parameter influence of TFR. Simulated and experimental studies on bearing fault are conducted to validate the proposed method and its enhanced performance over some advanced methods. Moreover, to broaden the applications of the enhanced VMD method with the guidance of ENS, an interesting extension is proposed to identify the multifault-related bands of bearing and a detailed discussion is given for this idea.

The rest of this paper is outlined as follows. Section 2 reviews the basic theory of the VMD. The ENS is constructed and the enhanced VMD with the guidance of ENS is proposed in Section 3. In Section 4, the simulated and experimental cases are used to validate the superiority of the method. A discussion on the extension of the proposed method is given in Section 5 and the conclusions are finally drawn in Section 6.

2. Variational Mode Decomposition

VMD is a new technique for adaptive signal decomposition with nonrecursively sifting structure, which can divide a real-valued signal \( f(t) \) into \( K \) meaningful modes \( u_k(t), k \in (1, 2, \ldots, K) \). The overall framework of VMD is a variational constraint model, which minimizes the bandwidth of each estimated mode. Assume that each mode has a finite bandwidth with its individual CF, the optimization algorithm of alternating direction multiplier method [31] is adopted to solve the variational constraint model. The model of VMD is constructed as

\[
L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[ (\delta(t) + \frac{i}{\pi t}) \ast u_k(t) \right] e^{-i\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle,
\]

in which \( \partial_t (\cdot) \) is the partial derivative with \( t \); \( \delta \) is the Dirac distribution; \( \ast \) denotes the convolution operator; \( \alpha \) represents the bandwidth parameter; and \( \lambda(t) \) denotes the coefficient of Lagrangian multiplier. The last two terms in equation (1) are used to render the reconstructing requirement. As a result, \( L(\{u_k\}, \{\omega_k\}, \lambda) \) is solved by the alternate direction multiplier method to search its saddle point. The optimization process mainly contains that the meaningful modes and their CFs are updated by modulating each mode to the corresponding baseband. Specially, the detailed procedure of VMD is written as follows.

1. Initialize \( \{u_k^1(t)\} \) and its CF \( \{\omega_k^1\} \); \( \alpha; \lambda(t) \).
2. Update \( u_k^{n+1}(t) \) and \( \omega_k^{n+1}(t) \) via the following formula:

\[
u_k^{n+1}(\omega) = \frac{f(\omega) - \sum_{i \neq k} u_i^n(\omega)}{1 + 2\alpha(\omega - \omega_k^n)^2},
\]

where \( f(\omega) \) is the Fourier transform of \( f(t) \).
Complexity

\[ \omega_k^{n+1} (\omega) = \frac{\int_0^\infty \omega |u_k^{n+1} (\omega)|^2 d\omega}{\int_0^\infty |u_k^{n+1} (\omega)|^2 d\omega} \]  
\[ (3) \]

(3) Update \( \lambda \) according to the following formula:

\[ \lambda_{n+1} (\omega) \leftarrow \lambda_n (\omega) + \tau \left( f (\omega) - \sum_k u_k^{n+1} (\omega) \right) \]  
\[ (4) \]

(4) Execute iteratively equations (2)–(4) until the convergent is met as

\[ \sum_{k=1}^K \|u_k^{n+1} - u_k^n\|^2_2 < \varepsilon, \]  
\[ (5) \]
in which \( \varepsilon \) is the tolerance value of the converging judgment that is set as \( 10^{-7} \) typically.

3. An Enhanced VMD Method with the Guidance of Envelope Negentropy Spectrum

Applications of the original VMD need to preset some initial parameters, including the number of decomposed modes and the bandwidth parameter. The effectiveness of original VMD is mainly subjected to these initial parameters. In our previous study [30], it was found that the above two parameters did not require to be given accurately via incorporating the effective ICF into VMD. However, the way to select the effective ICF is easily influenced by the parameters in the TFR. In this section, the ENS is firstly constructed to robustly indicate the latent IC s of faulty mode. Then, an enhanced VMD method with the guidance of ENS is proposed to extract the fault-related modes, thereby maximizing the advantages of VMD in bearing fault diagnosis.

3.1. Envelope Negentropy Spectrum. Entropy could be regarded as a good measure to detect out-of-equilibrium perturbations in a system [32]. In time domain, the local fluctuations of analysis signal will come with a decrease of its entropy. Various researches have proved that entropy is effective to evaluate the impulsive feature of mechanical signal [33]. The ENS proposed in this paper, which combines the concept of envelope with the definition of entropy, inherits the advantages of entropy and has stability under different conditions.

Generally, the information entropy \( H(x) \) of signal \( x(j) \) with length \( N \) is defined as

\[ H(x) = -\sum_{j=1}^N p(x(j)) \ln (p(x(j))), \]  
\[ (6) \]

where \( p(x(j)) \) represents the probability density of \( x(j) \). Hence, we attempt to use the information entropy to observe the frequency range of latent faulty mode coarsely. The TFR techniques can reveal the local fluctuations of waveform with time in the whole range of analysis frequency. As a consequence, TFR is used to transform the one-dimensional signal into the plane distribution in the time-frequency domain. Subsequently, the information entropy in different frequency ranges can be calculated based on the result of TFR. As one of TFR methods, the short-time Fourier transform (STFT) is selected as the TFR tool in this paper due to its simplicity and rapidity. For a real value signal \( x(t) \), its STFT is denoted as [34]

\[ \text{TFR}(t, f) = \int_{-\infty}^{\infty} x(t)w(t-t)e^{-2j\pi ft} dt, \]  
\[ (7) \]

where \( w(t) \) is the window function. Then, the envelope waveform of the analysis signal is obtained to describe the envelope negentropy at different frequencies as follows.

\[ \text{Env}(t_i, f_j) = \left( \text{Re}\left(\text{TFR}(t_i, f_j)\right) \right)^2 + \left( \text{Im}\left(\text{TFR}(t_i, f_j)\right) \right)^2. \]  
\[ (8) \]

According to the envelope information, the entropy at frequency \( f_j \) can be derived based on equation (6). However, the early local bearing defect, which might cause the repetitive feature buried in the analysis signal, would decrease the entropy of analysis at a specific frequency band. For convenience, the opposite of entropy, i.e., negentropy, is used to indicate the latent faulty mode as equation (9). This is to make the highest impulsiveness to agree with the largest value of negentropy.

\[ \text{Henv}(f_j) = \sum_{t=1}^{N} \text{Env}(t_i, f_j) / \sum_{t=1}^{N} \text{Env}(t_i, f_j) \left( \text{Env}(t_i, f_j) \right)^2 \]  
\[ (9) \]

Then, the concept of ENS is constructed by drawing the \( \text{Henv}(f_j) \) with \( f_j \). Although the form of ENS is affected by the result of TFR as expressed in equations (8) and (9), the applications will verify that the protruding frequency for indicating the faulty mode is relatively stable with the change of TFR parameters.

3.2. The Proposed VMD Method with the Guidance of ENS. In the original VMD algorithm, there are four initial decomposition parameters, i.e., Lagrangian multiplier coefficient, bandwidth parameter, the number of the decomposed modes, and its corresponding ICFs. As the strong noise component exists in the collected signal with the early stage defect of bearing, there is no need to strictly enforce the constraint, i.e., the Lagrangian multiplier coefficient should be set to zero. One of the critical links to the successful applications of VMD is the judgment of the number of decomposition modes. Actually, the number of meaningful modes embedded in the analysis data is difficult to be known in advance. Moreover, their corresponding ICFs which are close to the convergent property should also be careful to preset. At present, there are three ways for us to set ICFs.

(1) Zero initialization

\[ \omega_k^0 = 0, \quad k = 1, \ldots, K. \]  
\[ (10) \]

(2) Uniformly spaced distribution
where \( \omega_k^0 \) is the ICF of kth mode; \( \text{sort}(\cdot) \) is a rank function and \( \text{rand}(\cdot) \) is the function of random number. However, these ways based on experience might lead to the ICFs far away from the real ones and cause an increase in computational burden, even divergence, in the optimization procedure. In fact, not all of the modes buried in the measured signal have to be extracted, and only the fault-related modes are required to be identified. Therefore, the ENS is introduced to define the ICF of the faulty mode and the decomposing strategy of VMD can be reconstructed with the guidance of ENS. The bandwidth parameter cannot also be predicted in advance because the effective bandwidth changes with the different mechanical systems. Nevertheless, the bandwidth parameter is first given coarsely, and then it can be adjusted based on the sensitive index of faulty modes. According to the suggestion of ref [30, 5000]. So the following study uniformly sets the initial bandwidth parameter can be set between 30, the initial bandwidth parameter can be set between

\[
\omega_k^0 = \frac{k - 1}{2K}, \quad k = 1, \ldots, K.
\]  
(11)

(3) Randomly spaced distribution

\[
\omega_k^0 = \text{sort}(\text{e}^{\text{rand}(1/K) + \text{rand}(0.5) - \text{e}^{\text{rand}(1/K)}}), \quad k = 1, \ldots, K,
\]  
(12)

Step 6: perform the Hilbert envelope demodulation analysis on the optimal mode.

4. Applications of the Proposed Method

4.1. Simulation Study. To verify the performance of the proposed method, in this section, according to the vibration characteristics of a local fault bearing, a simulation is constructed as follows:

\[
y(t) = x(t) + e(t) = \sum_k \exp \left( -\frac{\xi_0}{1 - \xi_0^2} \right) \cdot 2\pi f_{re}(t - \tau_0 - kT_0) \\
\times \sin 2\pi f_{re}(t - \tau_0 - kT_0) + e(t),
\]  
(15)

where the periodical impact signal is \( x(t) \); the second part \( e(t) \) is the noise composition simulated by Gauss white noise and the function awgn \((x(t), \text{SNR} \text{‘measured’}) \) is used to adjust the signal to noise ratio (SNR) of the simulated signal; \( \xi_0 \) is the damping coefficient; \( f_{re} \) denotes the resonance frequency; \( \tau_0 \) is the initial time of the impulsive feature; and \( T_0 \) is the time interval of two adjacent impulsive features. All of the parameters used in equation (15) are listed in Table 1. The sampling frequency of the simulated signal is 2000 Hz and the simulated signal has a time length of 1 s.

The simulated signal with the SNR of -5 dB is first analyzed by the proposed method to study its effectiveness and robustness. Figure 2 shows the waveform of the simulated signal. As shown in Figure 2(c), it is impossible to observe the fault-related component from the envelope spectrum. Then the different parameters of the TFR (window length changed from 128 to 512 and overlap rate changed from 50% to 90%) are applied to analyze the results of ENS. Figure 3(a) demonstrates that the proposed method can well indicate the ICF of latent faulty mode. For comparing with the proposed method, the advanced indicator, i.e., EFS defined in ref. [30], is utilized to process the simulated signal. The EFS is given as

\[
EFS(f_j) = \left( \frac{\sum_{i=1}^{N} \text{Env}(t_i, f_j) - \sum_{i=1}^{N} \text{Env}(t_i, f_j)/N}{N - 1} \right)
\]  
(16)

The ICFs indicated by the EFS under different TFR parameters are demonstrated in Figure 3(b). It can be observed that the ICFs located by the EFS [30] have the obvious fluctuation with the change of the parameters of TFR. To further evaluate the stability of the enhanced VMD method with the guidance of ENS in this article, the simulated signal is repeated ten times randomly. As shown in Figure 4, the success rates of the proposed method are always 100% at different simulations, in which the ICF is located in the range of [290, 310 Hz]. However, some fluctuation exists in the success rates of EFS at multirun simulation. As reported in
Adjust the bandwidth parameters by the following optimization strategy

Measure the vibration data from the rotating machines

Compute the ENS based on the TFR of the test data

Locate the ICF of latent faulty mode from the ENS

Extract the expected mode by incorporating the ICF into VMD

Let optimal mode (OM) = \( u_0 \), \( i = 1 \),
\[ a_1 = a_0 - \Delta \alpha \] and \( a_i = a_0 + i\Delta \alpha \), extracted the mode \( u_i \) and \( u_r \) via VMD and calculate the kurtosis values \( K_{u_i} \), \( K_{u_r} \)

\[ K_{OM} < K_{u_i} \]

\[ K_{u_i} \leq K_{OM} \geq K_{u_r} \]

\[ i = i + 1, \quad a_i = a_0 - i\Delta \alpha \]

and calculate \( K_{u(i-1)} \)

\[ K_{u_i} \leq K_{u(i-1)} \]

\[ \text{Yes} \rightarrow \quad OM = u(i-1) \]

\[ \text{OM = ur(i-1)} \]

\[ K_{OM} \leq K_{u_r} \]

\[ i = i + 1, \quad a_i = a_0 + i\Delta \alpha \]

and calculate \( K_{u(i-1)} \)

\[ K_{u_i} \leq K_{u(i-1)} \]

\[ \text{Yes} \]

Execute Hilbert envelope demodulation analysis on the OM

Figure 1: Flow chart of the enhanced VMD method with the guidance of ENS.

Table 1: Parameter setting of simulation signal.

<table>
<thead>
<tr>
<th>( \xi_0 )</th>
<th>( f_{re} )</th>
<th>( \tau_0 )</th>
<th>( T_0 )</th>
<th>( f_{outer} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>300 Hz</td>
<td>0.05 s</td>
<td>0.02 s</td>
<td>50 Hz</td>
</tr>
</tbody>
</table>

Table 2, the statistical values (means and variances denoted in equations (17) and (18)) also prove that the proposed method has a good stability for indicating the ICF of latent faulty mode.

\[
\mu = \frac{\sum A_i}{n} \tag{17}
\]

\[
D = \frac{\sum (A_i - \mu)^2}{n - 1} \tag{18}
\]

where \( \mu \) and \( D \) are the mean and variance, respectively, \( A_i \) represents the success rate of each simulation, and \( n \) denotes the number of simulations.

Specifically, an example of simulated signal with \(-5\) dB is given as follows. As illustrated in Figure 5(a), the TFR of the simulated signal is obtained by the STFT with a window length of 384 and overlap rate of 90%. Then, the ENS based on TFR is calculated as shown in Figure 5(b). For comparison, the spectra obtained by the EFS are also drawn in Figure 5(c). As shown in the figure, although EFS can find the ICF of the latent faulty mode under certain parameters, the spectra obtained by these indices are not obvious over the proposed method.

To further verify the antinoise ability of the proposed method, the simulated signal with SNR \(-10\) dB is analyzed as before. The waveform of the simulated signal with SNR \(-10\) dB is drawn in Figure 6. Figure 7(a) illustrates the ICFs extracted by the proposed method with different parameters of the TFR and Figure 7(b) shows the ICFs indicated by the EFS. It can be seen that a more steady result is obtained by the proposed method than the EFS. Furthermore, the simulated signal is also repeated ten times randomly.

Figure 8 shows the success rates of the proposed method are also always 100% at multirun simulations. However, the remarkable fluctuation exists in the success rates of EFS. Table 3 lists the means and variances of success rates obtained by the two different indexes. It also proves that the proposed method has a better performance to locate the ICF than the EFS.

Similarly, an example of simulated signal with \(-10\) dB is given as follows. As shown in Figure 9(a), the STFT with a window length of 384 and overlap of 90% is used to obtain the TFR of the simulated data. Figures 9(b) and 9(c) show the spectra indicated by the proposed method and EFS. We can see that the ICF indicated by the proposed method is closer to the real one. Meanwhile, EFS lost the ability to locate the correct ICF. Based on the guidance of ENS and EFS, the optimal mode is extracted by adjusting the bandwidth parameter, respectively. Then, the Hilbert envelope demodulation analysis is carried out on the optimal modes illustrated in Figures 10(a) and 10(b). As shown in Figures 10(c) and 10(d), the envelope spectrum of optimal mode based on ENS exhibits the faulty frequency 50 Hz and its harmonics clearly, while the envelope spectrum of optimal mode based on EFS shows nothing. The analysis results show that the proposed method is not susceptible to the parameters and has a good antinoise ability by comparing with the EFS from different views.

As a powerful analytical technique for bearing fault feature extraction, spectral kurtosis (SK) [8] is used to analyze the simulated signal for further comparison. The paving of kurtosis values of SK and the resulting signal filtered by frequency band (expressed by a solid red circle) are illustrated in Figure 11, and it can be clearly seen the kurtosis-dominant frequency band with CF of 718.75 Hz. The resulting bandpass filtering signal is further processed using the envelope analysis. And no fault feature appears in its envelope spectrum, as shown in Figure 11(c). The results mean that the proposed method outperforms SK in extracting the weak faulty features.
To further verify the superiority of the proposed method over some adaptive signal decomposition methods, the simulated noisy signal is analyzed by the traditional VMD method with fixed decomposition parameters ($K = 3$, $\alpha = 2500$) and the particle swarm optimization (PSO)-based

![Image](a)

![Image](b)

![Image](c)

Figure 2: The wave of simulated signal (SNR = -5 dB): (a) pure signal $x(t)$, (b) noisy signal $y(t)$, (c) envelope spectrum.

![Image](a)

![Image](b)

Figure 3: ICFs located by the proposed method and EFS of the simulated signal in Figure 2(b). (a) Envelope negentropy. (b) Energy fluctuation.

![Image](c)

Figure 4: Success rates of the proposed method and EFS on the simulated signal in Figure 2(b).

<table>
<thead>
<tr>
<th>Index</th>
<th>ENS (%)</th>
<th>EFS (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>100</td>
<td>55.29</td>
</tr>
<tr>
<td>$D$</td>
<td>0</td>
<td>13.19</td>
</tr>
</tbody>
</table>

Table 2: The mean and variance of success rate for different indices with SNR = -5 dB.

To further verify the superiority of the proposed method over some adaptive signal decomposition methods, the simulated noisy signal is analyzed by the traditional VMD method with fixed decomposition parameters ($K = 3$, $\alpha = 2500$) and the particle swarm optimization (PSO)-based
Figure 5: Spectral analysis of the simulated signal in Figure 2(b) at window length 384 and overlap rate 90%: (a) TFR, (b) ENS, (c) EFS.

Figure 6: The wave of simulated signal (SNR = −10 dB): (a) pure signal $x(t)$, (b) noisy signal $y(t)$, (c) envelope spectrum.
VMD method [26]. In the PSO-based VMD method, to guarantee consistency of the comparison, the number of modes $K$ varies from 3 to 8, and the bandwidth parameter $\alpha$ varies from 400 to 5000. The waveform and envelope spectra of the optimal target mode obtained by the fixed-parameter VMD are provided in Figures 12(a) and 12(b), from which we cannot see any information related to fault characteristic frequency. The bandwidth parameter and number of decomposed modes optimized by PSO search algorithm are a collection of $(2400, 8)$, the optimal target mode and its envelop spectrum are shown in Figures 12(c) and 12(d). It can be found that the optimal mode obtained by the PSO-based VMD method can barely detect the fault characteristic frequency, and there are too many interference terms. Furthermore, due to the complex iterative process of the PSO-based method, it takes more time to find the optimal mode than the proposed method in this paper.

### 4.2. Experimental Verification

#### 4.2.1. Experiment 1

In this section, we validate the proposed method using the experimental signals. As shown in Figure 13, the test rig consists of a motor, a tachometer, a gearbox, a tested rolling bearing, and control electronics. In this experiment, the rotating speed of the driving motor is 1740 r/min and the sampling frequency is 51.2 kHz. The accelerometer is mounted on the top of the tested bearing. The geometric parameters and characteristic frequencies of the tested bearings are listed in Table 4, where the ball pass frequency outer race (BPFO), ball pass frequency inner race (BPFI), and ball spin frequency (BSF) are denoted by $f_{ao}$, $f_{di}$, and $f_{BS}$, respectively.

The time waveform of the experimental signal and its envelope spectrum are illustrated in Figure 14. There is no obvious fault characteristic frequency in Figure 14(b). Then, the experimental data are analyzed by the proposed method. As shown in Figure 15, the ICFs located by the proposed method with different parameters of the TFR are more stable than the results indicated by the EFS. Speciality, the STFT with a window length of 384 and overlap of 90% is used to obtain the TFR of the experimental data, as shown in Figure 16(a). Then, the ENS can be obtained based on the TFR. Comparing with the proposed ENS, it can be seen from Figure 16 that the inaccurate ICF is indicated by the spectra constructed by the EFS. Based on the guidance of ENS, the optimal mode is extracted by adjusting the bandwidth parameter. As shown in Figures 17(a) and 17(c), the envelop spectrum of optimal mode exhibits the outer race faulty frequency $f_{ao}$ and its harmonics clearly, which is in good agreement with the actual situation. Furthermore, the optimal mode based on EFS and its envelop spectrum are illustrated in Figures 17(b) and 17(d) for comparison, from which we cannot extract any fault information from the experimental data. The above analysis shows that the proposed method can well capture the fault-related feature.

As comparison, SK is performed on the same vibration signal. The kurtogram and the resulting signal filtered by frequency band (indicated by a solid red circle) are illustrated in Figures 18(a) and 18(b). The envelop analysis is then carried out on the resulting bandpass filtering signal. From Figure 18(c), the SK is also inferior to the proposed method in fault diagnosis of rolling bearing.

For further comparison, the raw signal is also analyzed by the fixed-parameter VMD method and PSO-based VMD
method, and the results are shown in Figure 19. Similar to the results of simulation analysis, the fixed-parameter VMD ($K = 3, \alpha = 2500$) method cannot extract fault features. The optimal mode and its envelop spectrum determined by PSO-based VMD with the optimal solution ($K = 4, \alpha = 4800$) are displayed in Figures 19(c) and 19(d). Although the fault characteristic frequency $f_0$ is identified in the envelope spectrum, its harmonic cannot be found compared with the proposed method in Figure 17(c), and the PSO-based method is time consuming.

4.2.2. Experiment 2. To further confirm the enhanced performance of the proposed method in bearing defect identification, the faulty bearing data of the Case Western Reserve University (CWRU) Bearing Data Center [35, 36] is
Figure 11: Simulated signal in Figure 6(b) processed by the SK: (a) the paving of kurtosis values and (b) the signal filtered by frequency band (expressed by a solid read circle); (c) envelope spectrum of the filtered signal.

Figure 12: Results of the simulation signal in Figure 6(b) by different methods: (a) optimal mode extracted by the fixed-parameter VMD and (b) its envelope spectrum; (c) optimal mode extracted by PSO-based VMD and (d) its envelope spectrum.
Figure 13: SpectrQuest machinery fault simulator test rig.

Table 4: Geometry and characteristic frequencies of the rolling bearing.

<table>
<thead>
<tr>
<th>MB ER-16K</th>
<th>Ball numbers</th>
<th>Ball diameter (mm)</th>
<th>Pitch diameter (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fault type</td>
<td>$f_o$</td>
<td>$f_i$</td>
<td>$f_B$</td>
</tr>
<tr>
<td>Fault characteristic frequency (Hz)</td>
<td>103.88</td>
<td>157.41</td>
<td>67.28</td>
</tr>
</tbody>
</table>

Figure 14: The experimental bearing vibration data of the test rig: (a) waveform; (b) envelope spectrum.

Figure 15: ICFs located by the proposed method and EFS with the experimental signal in Figure 14(a). (a) Envelope negentropy. (b) Energy fluctuation.
also analyzed in this study. The CWRU bearing failure test device, which consists of a 2 hp electric motor (left), a torque transducer/encoder (center), and a dynamometer (right), is illustrated in Figure 20. The rotational frequency of the driving motor is 29.95 Hz. In this experiment, the accelerometers with a sampling frequency of 12 kHz are utilized to collect the vibration signal. In order to validate the ability of the proposed method for early weak fault feature extraction, the vibration data collected from test bearing with smallest faulty sizes (width 0.1778 mm and depth 0.2794 mm) is selected for analysis. The geometric parameters and characteristic frequencies of the test bearing are listed in Table 5. The waveform and envelope spectrum of the measured signal are shown in Figure 21. The fault-induced impulses are so weak to be identified from the background noise in the waveform, and the envelope spectrum also contains many interference components such as noise. Then the proposed method is applied to process the experimental data.
Experimental data in Figure 14(a) processed by the SK: (a) the paving of kurtosis values and (b) the signal filtered by frequency band (expressed by a solid red circle); (c) envelope spectrum of the filtered signal.

Results of the experimental signal in Figure 14(a) by different methods: (a) optimal mode extracted by the fixed-parameter VMD and (b) its envelope spectrum; (c) optimal mode extracted by PSO-based VMD and (d) its envelope spectrum.
Figure 22 displays the ICFs located by the proposed ENS under different parameters of TFR. Similarly, as shown in Figure 23, the proposed ENS can be obtained based on the TFR with a window length of 384 and overlap of 90%. Based on the guidance of ENS, the optimal mode is extracted by adjusting the bandwidth parameter. As shown in Figure 24, the envelope spectrum of optimal mode exhibits the outer race faulty frequency $f_o$ and its harmonics clearly, which is in good agreement with the actual situation. Comparing with the proposed method, ICF-guided VMD method based on EFS identified the rotation frequency $f_r$ and cannot extract any fault information. The above analysis shows that the proposed method can well capture the fault-related feature.

For comparison, the experimental bearing vibration data of CWRU Bearing Data Center is also analyzed by SK. As
Figure 22: ICFs located by the proposed method and EFS with the experimental signal in Figure 21(a). (a) Envelope negentropy. (b) Energy fluctuation.

Figure 23: Spectral analysis of the experimental data in Figure 21(a) at window length 384 and overlap rate 90%: (a) TFR, (b) ENS, (c) EFS.

Figure 24: Continued.
shown in Figure 25(c), we can see that the dominant frequency in the envelope spectrum obtained by the SK is not the fault-related frequency. So the SK is inferior to the proposed method in fault diagnosis of rolling bearing.

In addition, the fixed-parameter VMD ($K = 3$, $\alpha = 2500$) and PSO-based VMD method are also used to analyze the experimental signal for comparisons. The PSO-based VMD method finally determined the optimal parameters are $K = 4$.

Figure 24: Results of the experimental data in Figure 21(a) by different ICF locating methods: (a) optimal mode extracted by the proposed method and (c) its envelope spectrum; (b) optimal mode extracted based on EFS and (d) its envelope spectrum.

Figure 25: Experimental data in Figure 21(a) processed by the SK: (a) the paving of kurtosis values and (b) the signal filtered by frequency band (expressed by a solid read circle); (c) envelope spectrum of the filtered signal.
and \( \alpha = 600 \) through an iterative search process. The results of those two methods are shown in Figure 26. Although the fault characteristic frequency can be found in the spectra, all of them display many interference components and the amplitude of its harmonic is too small to be distinguished. This case verifies the enhanced performance of the proposed method in bearing fault detection again.

5. Discussions

When analyzing the simulation and experimental signals, it has proved that the proposed method has a good performance in bearing fault diagnosis. However, the situation of multicomponents about faulty information or compound faults of bearing usually exists in on-site application. Hence, an extension for identifying multicomponents is essential to broaden the applied scope of the proposed method. Although the enhanced VMD with the guidance of ENS focuses on extracting the mono-component currently, some excellent properties of the proposed method make it possible to be further utilized to extract the multifault information. (1) All bearing fault information can be revealed by the ENS due to its fine frequency resolution and sensitiveness to the bearing faulty feature. (2) In the enhanced VMD, the meaningful mode is extracted individually and the bandwidth parameter can be adjusted independently. According to these properties, it can be found that the iteratively decomposing strategy can be easily applied to the enhanced VMD with the guidance of ENS, which is given in Algorithm 1. By using the extension of the enhanced VMD, the procedure of bearing fault diagnosis can be concluded as follows. First, the ICFs of latent bearing faulty modes can be located by detecting the ENS and sorting the ICFs in ascending order. Second, iteratively extract the latent bearing faulty modes and adjust the bandwidth parameter of individual mode to enrich the fault information. Third, use the sensitive index to evaluate the faulty modes.

An experimental case is conducted on this extension to explain its procedure. The experimental dataset [37] is collected from an Aero engine rotor-bearing fault simulator and the test rig is illustrated in Figure 27. Three accelerometers are installed to collect the vibration signal with the sampling rate 10 kHz. The driving motor has a rotating speed of 1840 r/min. Table 6 lists the geometric parameters and characteristic frequencies of the tested bearings. Figure 28(a) shows the time waveform of the experimental signal collected by the channel #3 accelerometer. The envelope spectrum of the experimental signal is performed in Figure 28(b), in which it is difficult to detect the faulty characteristic frequency. Then, we use the extension for the enhanced VMD with the guidance of ENS to process the experimental signal. As shown in Figure 29(a), the STFT with a window length of 384 and an overlap of 90% is used to obtain the TFR of the experimental signal. Subsequently, the four obvious “hills” are clearly observed based on the ENS as plotted in Figure 29(b).

With the guidance of ENS, the four meaningful modes are iteratively extracted by the extension method. Figures 30(a)–30(e) show the time waveforms and Fourier spectra of the extracted meaningful modes. To remove the interference of noise component or inherent component, the kurtosis is selected as sensitive index to further evaluate the faulty modes. As shown in Figure 31, the last two modes have the larger kurtosis. Figure 32 illustrates that the typical fault 

Figure 26: Results of the experimental signal in Figure 21(a) by different methods: (a) optimal mode extracted by the fixed-parameter VMD and (b) its envelope spectrum; (c) optimal mode extracted by PSO-based VMD and (d) its envelope spectrum.
Construct the ENS: \( \text{Henv}(f_j) = \sum_{i=1}^{N} (\text{Env}(t_i, f_j)/\sum_{i=1}^{N} \text{Env}(t_i, f_j)) \ln[(\text{Env}(t_i, f_j)/\sum_{i=1}^{N} \text{Env}(t_i, f_j))] \)

Locate the first \( M \) ICFs based on the ENS: \( \{k_{m1}^{\alpha_j}, k_{m2}^{\alpha_j}, \ldots, k_{mM}^{\alpha_j}\} = \text{findpeaks}(\text{Henv}(f_j)) \)

for \( m = 1, 2, \ldots, M \) do

initialize: \( n \leftarrow 0, j \leftarrow 0, \alpha \leftarrow \alpha_0, \{\omega_m^{\alpha_j}\} \)

repeat

\( u_m^{\alpha_j+1}(\omega) = f(\omega)/\left(1 + 2\alpha(\omega - \omega_m^{\alpha_j})^2\right), \omega_m^{\alpha_j+1}(\omega) = \int_{\omega_m^{\alpha_j}}^{\infty} \omega_m^{\alpha_j+1}(\omega)^2 d\omega / \int_{\omega_m^{\alpha_j}}^{\infty} |\omega_m^{\alpha_j+1}(\omega)|^2 d\omega \),

\( n \leftarrow n + 1 \)

until convergence: \( \|u_m^{\alpha_j} - u_m^{\alpha_j+1}\|^2 / \|u_m^{\alpha_j}\|^2 < \epsilon \)

Let: \( \Omega_m = u_m^{\alpha_j} \) and its CF \( \Omega_m^\alpha = \omega_m^{\alpha_j}(\omega) \)

while \( \epsilon_r > 0 \) do

initialize: \( n \leftarrow 0, j \leftarrow 0, \alpha \leftarrow \alpha_0 + j\Delta\alpha, \omega_m^0 \leftarrow \text{CF}_m^j \)

repeat

\( u_m^{\alpha_j+1}(\omega) = f(\omega)/(1 + 2\alpha(\omega - \omega_m^{\alpha_j})^2), \omega_m^{\alpha_j+1}(\omega) = \int_{\omega_m^{\alpha_j}}^{\infty} \omega_m^{\alpha_j+1}(\omega)^2 d\omega / \int_{\omega_m^{\alpha_j}}^{\infty} |\omega_m^{\alpha_j+1}(\omega)|^2 d\omega \),

\( n \leftarrow n + 1 \)

until convergence: \( \|u_m^{\alpha_j} - u_m^{\alpha_j+1}\|^2 / \|u_m^{\alpha_j}\|^2 < \epsilon \)

\( j \leftarrow j + 1 \)

obtain mode: \( \Omega_m = u_m^j \) and its CF \( \text{CF}_m^j = \omega_m^j(\omega) \)

\( \epsilon_r = \text{kurtosis}(\Omega_m^\alpha) - \text{kurtosis}(\Omega_m^{\alpha-1}) \)

end while

Let: \( \Omega_m = \text{OM}_m, \text{CF}_m = \text{CF}_m^j \) and \( \alpha_m = \alpha_0 + (j - 1)\Delta\alpha \)

end for

Return \( \{\Omega_1, \Omega_2, \ldots, \Omega_M\}; \{\text{CF}_1, \text{CF}_2, \ldots, \text{CF}_M\}; \{\alpha_1, \alpha_2, \ldots, \alpha_M\} \)

Algorithm 1: An extension for the enhanced VMD with the guidance of ENS.

Figure 27: Aero engine rotor-bearing fault simulator.

Table 6: Geometry and characteristic frequencies of the test bearing.

<table>
<thead>
<tr>
<th>HRB6304</th>
<th>Ball numbers</th>
<th>Ball diameter</th>
<th>Pitch diameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7</td>
<td>9.6 mm</td>
<td>36 mm</td>
</tr>
<tr>
<td>Fault type</td>
<td>( f_i )</td>
<td>( f_o )</td>
<td>( f_B )</td>
</tr>
<tr>
<td>Fault characteristic frequency (Hz)</td>
<td>135.8</td>
<td>78.7</td>
<td>53.4</td>
</tr>
</tbody>
</table>

Characteristics of the being inner race fault can be detected from the envelope spectrum of the two modes, including the inner race fault frequency and its harmonic. Thus, it can be derived that the test rig has a bearing inner race fault, which is in line with the actual situation. Analysis results of this experimental signal demonstrate that the extension of the proposed method well works on the extraction of the multicomponents fault information.

For comparing purpose, SK, the fixed-parameter VMD (\( K = 3, \alpha = 2500 \)), and the PSO-based VMD method (optimal parameters: \( K = 8, \alpha = 5000 \)) are used to analyze the multicomponents data, as shown in Figures 33 and 34. The fault
Figure 28: The wave of multicomponents data: (a) waveform; (b) envelope spectrum.

Figure 29: Spectral analysis of the multicomponents data in Figure 28(a) at window length 384 and overlap rate 90%: (a) TFR, (b) ENS.

Figure 30: Continued.
characteristic frequency $f_i$ and one coarsely fault-related band can be observed from the envelope spectrum, but it is hard to detect its harmonics and the sidebands. This also proves that the method is superior to some advanced methods like SK, traditional VMD, and PSO-based VMD method in extracting the fault feature of multicomponents data.

6. Conclusions

In this paper, we presented an enhanced VMD with the guidance of ENS for bearing fault diagnosis. The main contributions and advantages of the enhanced VMD method are that it can overcome the difficulty in the initial parameters selection of the traditional VMD, and it is immune
Figure 33: Multicomponents data in Figure 28(a) processed by the SK: (a) the paving of kurtosis values and (b) the signal filtered by frequency band (expressed by a solid read circle); (c) envelope spectrum of the filtered signal.

Figure 34: Continued.
to the change of the parameters in TFR; thus, the reliability of the bearing fault diagnosis can be significantly improved. The simulation and experimental signal studies confirmed the effectiveness of the proposed method in the diagnosis of local damage of rolling bearings. The comparisons to VMD under the guidance of EFS demonstrated the advantage of the enhanced VMD in the robustness to the parameters in TFR and the excellent resistance to strong noise component. Moreover, the results validated the improved performance of the proposed method compared with some advanced methods including SK, fixed-parameter VMD, and PSO-based VMD in extracting the weak faulty feature of rolling bearing.

In order to adapt to the complex work conditions, an extension on the enhanced VMD is discussed to broaden the applications of this method. We found that the proposed method can be easily extended to extract the multicomponents fault information buried in mechanical signal and a practical case is used to introduce the procedure of the extension method. Comparison between the proposed method and other methods indicate that the proposed method outperforms some advanced methods in extracting the fault feature of multicomponents data.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research was supported by the National Natural Science Foundation of China (Grant nos. 51705349, 51875376, and 51875375), the China Postdoctoral Science Foundation (Grant no. 2019T120456), the Suzhou Prospective Research Program (Grant nos. SYG201932 and SYG201802), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (Grant no. KYCX18_2502), which are highly appreciated by the authors.

References


