Research Article

A Parameter-Optimized Variational Mode Decomposition Investigation for Fault Feature Extraction of Rolling Element Bearings

Guoping An,[1] **Qingbin Tong** **,**[1] **Yanan Zhang,**[2] **Ruifang Liu,**[1] **Weili Li,**[1] **Junci Cao,**[1] **Yuyi Lin,**[3] **Qiang Wang,**[1] **Ying Zhu,**[1] **and Xiaowen Pu**[1]

[1]*School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China*
[2]*State Grid JIBEI Electric Power Co., Ltd. Maintenance Branch State Grid, Beijing 102488, China*
[3]*Department of Mechanical and Aerospace Engineering, University of Missouri, Columbia MO 65211, USA*

Correspondence should be addressed to Qingbin Tong; qbtong@bjtu.edu.cn

Reliable fault diagnosis of the rolling element bearings highly relies on the correct extraction of fault-related features from vibration signals in time-frequency analysis. However, considering the nonlinear, nonstationary characteristics of vibration signals, the extraction of fault features hidden in the heavy noise has become a challenging task. Variable mode decomposition (VMD) is an adaptive, completely nonrecursive method of mode variation and signal processing. This paper analyzes the advantages of VMD compared with EMD in robustness of against noise, overcoming the end effect and mode aliasing. The signal decomposition performance of VMD algorithm largely depends on the selection of mode number k and bandwidth control parameter α. To realize the adaptability of influence parameters and the improvement of decomposition accuracy, a parameter-optimized VMD method is presented. The random frog leaping algorithm (SFLA) is used to search the optimal combination of influence parameters, and the mode number and bandwidth control parameters are set according to the search results. A multiobjective evaluation function is constructed to select the optimal mode component. The envelope spectrum technique is used to analyze the optimal mode component. The proposed method is evaluated by simulation and practical bearing vibration signals under different conditions. The results show that the proposed method can improve the decomposition accuracy of the signal and the adaptability of the influence parameters and realize the effective extraction of the bearing vibration signal.

1. Introduction

Rolling element bearings, as a very important component of rotating machinery, has been widely used in modern industry such as engineering machinery and aerospace [1, 2]. The working state of rolling element bearings is directly related to the safety of the rotating machinery. Rolling element bearings are easily damaged under the long-term operation of harsh environment with high speed, heavy load, strong impact, and high temperature. The developed mechanical faults may cause the deterioration of machine operating conditions, resulting in serious economic losses and casualties [3–5]. The vibration signal detected by the sensor is always related to the important physical information that a series of shock pulses will occur when the rolling element bearing is subjected to a local fault [6, 7]. However, the defect-induced impulses in practice are too weak to distinguish well from vibration signal corrupted by a large amount of background noise. Therefore, it is critical to remove noise and extract intrinsic fault features from the measured original signal for the fault diagnosis of rolling element bearing.

Many vibration analysis methods have been proposed in the literature for bearing fault detection in the time domain, the frequency domain, and the time-frequency domain, respectively [8, 9]. However, the vibration signal of rolling

element bearings is the nonstationary and nonlinearity signal. It is very difficult to identify the fault characteristics of the rolling element bearing only in the analysis of time domain or the frequency domain. To effectively analyze the fault features from the vibration signals, some traditional time-frequency analysis methods have been widely used, such as short-time Fourier transform (STFT) [10], Wigner–Ville distribution (WVD) [11], and wavelet transform (WT) [12]. However, due to the limitation of Heisenberg's uncertainty principle, STFT method cannot get high resolution in the time domain and frequency domain simultaneously when dealing with the nonstationary signals. The disadvantage of the WVD method is that it cannot guarantee nonnegativity and produce serious cross-term interference for multicomponent signals or signals with complex modulation laws. The WT method decomposes the signal by performing scaling and translation operations on the wavelet basis and can effectively obtain time-frequency information from the measured signal. It has good localization properties in the time domain and frequency domain and has multiresolution analysis features [13, 14]. However, the WT cannot accurately split the high-frequency band where the modulation information of machine fault always exists.

Compared with the traditional analysis methods, empirical mode decomposition (EMD) offers a different analysis approach to signal processing in the time-frequency domain. The EMD provides more realistic signal representations without artifacts imposed by the nonadaptive limitations of both Fourier and wavelet transform-based time-frequency analysis methods and is suitable for the analysis of the nonlinear and nonstationary signals [15–17]. It is based on the local characteristic time scales of a signal and can self-adaptively decompose the complicated signal into a limited number of intrinsic mode functions (IMFs) through automatically performing a series of recursive calculations. The IMFs represent the fundamental oscillatory modes embedded in the signal, from which the instantaneous time-frequency features of interest are deemed to be observed. This enables the EMD-based methods to have potential as promising tools for dealing with the engineering problems associated with the analysis of nonstationary signals [18]. Therefore, the EMD and its extension forms (such as Ensemble Empirical Mode Decomposition (EEMD)) have attracted the attention of many researchers and are widely applied in the fault diagnosis and recognition of rolling element bearings [19–23]. In practical applications, although the EMD and its improved method have advantages in the processing of the nonstationary signals, the method itself still has the following inherent defects:

Weak Robustness of against Noise. The EMD-based methods are sensitive to the complex noise in the vibration signal. A little change in the signal-to-noise ratio (SNR) can lead to the different signal decomposition results [18].

Mode Aliasing. The local mean is defined by the upper and lower envelopes of the signal in the EMD. Based on this definition, different modal components can be distinguished through the characteristic scale of the signal. The IMF is no longer limited to the narrowband signal, and it can also show amplitude modulation and frequency modulation at the same time. However, when there is a jump change in the time scale of the signal, the direct screening process will produce mode aliasing issues. Intuitively, it is impossible to effectively separate the different modal components according to the characteristic scale, which makes the existing IMFs contain the different time-scale components and cannot clearly reflect the intrinsic properties of the signal.

End Effect. The upper and lower envelopes of the signal are interpolated by the cubic spline interpolation in the EMD. The cubic spline interpolation needs two adjacent points. As a result, the divergence occurs at both ends of the data, and the divergent results gradually "pollute" the whole data sequence during the data decomposition process, which leads to the serious distortion and energy leakage.

Variational mode decomposition (VMD) method has been proposed and developed recently, which is an alternative nonrecursive signal decomposition method that can adaptively determine the relevant frequency bands and the corresponding mode simultaneously [24–26]. The VMD method decomposes a signal into a series of band-limited modes. These modes can be continuously updated with Wiener filtering, and the central frequency of each mode can be gradually demodulated to the corresponding baseband. The nonrecursive signal decomposition of VMD is more efficient than the EMD and its extension forms in computation. At the same time, the application of Wiener filtering makes the VMD method robust to the background noise. Due to the application of Wiener filters, the narrowbanded function of VMD resultant modes not only reduces the mode mixing issues existing in the EMD but also helps to accurately extract the fault characteristics of the signal through the Hilbert transform. However, the decomposition accuracy of the VMD method is usually affected by the number of modes k and the bandwidth control parameter α. The original VMD method used the default values to implement the signal analysis, which largely limits its decomposition precision and the capability of feature extraction to a certain extent.

In this paper, we firstly analyze the advantages of VMD compared with EMD in robustness of against noise, overcoming the end effect and mode aliasing. To realize the adaptability of influence parameters and the improvement of decomposition accuracy, a parameter-optimized VMD method is presented. The random frog leaping algorithm (SFLA) is used to search the optimal combination of influence parameters, and the mode number and bandwidth control parameters are set according to the search results. A multiobjective evaluation function is constructed to select the optimal mode component. The envelope spectrum technique is used to analyze the optimal mode component. The proposed method is evaluated by simulation and practical bearing vibration signals under different conditions. The remaining section of the paper is organized as

follows: Section 2 introduces the fundamental theory of the VMD. The superiorities of the VMD over the EMD are analyzed in Section 3. The parameter-optimized VMD algorithm is presented in Section 4. The fault feature extraction based on the parameter-optimized VMD is given in Section 5. Section 6 will present the experimental results and analysis. Finally, the conclusion is drawn in Section 7.

2. Brief Introduction to VMD

The VMD algorithm is an adaptive, quasiorthogonal, and completely nonrecursive signal processing method. It decomposes the input signals composed of multicomponents into several inherent modes with limited bandwidth, and most of these modes are closely around their corresponding central frequencies [24]. By solving the optimal solution of constrained variational problem, the central frequency and band limit of each mode can be decided. An input signal $f(t)$ can be expressed as follows:

$$f(t) = \sum_{k=1}^{K} u_k(t), \tag{1}$$

where the number of modes k is defined in advance and $u_k(t)$ is the narrowband mode function. It can be written as

$$u_k(t) = A_k(t)\cos(\phi_k(t)), \tag{2}$$

where $A_k(t)$ is the instantaneous amplitude of $u_k(t)$, $\phi_k(t)$ is the instantaneous phase, and $\phi_k(t)$ is the reduction function that instantaneous frequency $\omega_k(t) = d\phi_k(t)/dt \geq 0$. Compared to $\phi_k(t)$, the variation in $A_k(t)$ and $\omega_k(t)$ is more gradual that can be regarded as a harmonic signal of constant amplitude and frequency in a smaller time horizon.

The VMD decomposes the input signal into a certain number of modes, and the decomposed modes have specific sparsity property while reproducing the input signal. It is assumed each mode is closely integrated around the center frequency. To assess the bandwidth of a mode, the following scheme is needed: the VMD method decomposes the input signal into a certain number of modes, which make them reappear the input signal and have specific sparsity properties. It is assumed each mode is closely integrated around the center frequency. To assess the bandwidth of a mode, the following scheme is needed: (1) compute the analyzed signal by means of the Hilbert transform to get a one-sided frequency spectrum for the mode; (2) transform the frequency spectrum of each mode to the baseband by mixing with an exponential tuned to the estimated center frequency; (3) estimate the bandwidth through the H^1 Gaussian smoothness of the demodulated signal, that is, the squared L^2-norm of the gradient. The constrained variational problem would be expressed as follows:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t_2^2} \right\},$$

$$\text{s.t.} \sum_k u_k(t) = f(t), \tag{3}$$

where $\{u_k\}(k = 0, 1, 2, \ldots, K)$ represents the k-th mode component obtained by decomposition and $\{\omega_k\}$ represents the corresponding central frequencies of the k-th mode component.

To solve the constrained variational problem, the augmented Lagrange is introduced and the unconstrained variational problem is gotten by

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle, \tag{4}$$

where α represents the quadratic penalty factor, which can guarantee the accuracy of signal reconstruction in the presence of Gauss noise, and λ represents the Lagrange operator, which can be used to maintain the strictness of constraints. The saddle point of the augmented Lagrange L is the optimal solution of original minimization problem, which can be solved using alternate direction method of multipliers (ADMM). All the modes can be obtained from (5) in the frequency domain through updating each mode:

$$\widehat{u}_k^{n+1}(\omega) = \frac{\widehat{f}(\omega) - \sum_{i \neq k} \widehat{u}_i(\omega) + \widehat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2}, \tag{5}$$

where $\widehat{u}_k^{n+1}(\omega)$ can be equivalent to the Wiener filter of the current residual signal and the full spectrum of the real mode can be obtained by conjugate symmetry. Thus, the real part $\{u_k(t)\}$ can be achieved through utilizing the inverse Fourier transform of $\{\widehat{u}_k^{n+1}(\omega)\}$.

Similarly, to obtain the minimum value of ω_k^{n+1}, the central frequency updating problem can be transformed into the corresponding frequency domain, and the solutions of the central frequencies can be given as follows:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\widehat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\widehat{u}_k(\omega)|^2 d\omega}. \tag{6}$$

Therefore, the new value of ω_k can be set to the center of gravity of the corresponding modal power spectrum.

To update the Lagrange operator λ, the following expression is given:

$$\widehat{\lambda}^{n+1}(\omega) = \widehat{\lambda}^n(\omega) + \tau \left(\widehat{f}(\omega) - \sum_k \widehat{u}_k^{n+1}(\omega) \right). \tag{7}$$

According to the above theoretical description, the detailed procedures of VMD algorithm are given as follows:

(1) Define the number of modes k;

(2) Initialize: $\{\hat{u}_k^1\}\{\hat{\omega}_k^1\}$, $\hat{\lambda}^1$ and $n = 0$;

(3) Update u_k and ω_k according to Equations (5) and (6);

(4) Update λ according to Equation (7);

(5) Set the error $\varepsilon > 0$, if the inequality $(\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^{n^2}\|/\|\hat{u}_k^{n^2} < \varepsilon\|)$ holds, then the iteration stops, or else go back to step 2.

3. Superiorities of VMD over EMD

The VMD algorithm transfers the signal decomposition process to the variational framework and achieves adaptive signal decomposition by searching the optimal solution of the constrained variational model. Compared with the EMD method, the VMD method has more advantages in the noise robustness, the mode aliasing, and the end effect. The superiorities of the VMD over the EMD will be investigated in this section.

3.1. Noise Robustness and Elimination of Mode Aliasing. To verify the advantages of VMD in the noise robustness, the mode aliasing, and the end effect, a simulated signal $x(t)$ is designed in Figure 1, which is the sum of a harmonic signal and white noise η. The signal is modulated by 30 Hz and 56 Hz. The purpose of this simulation is to simulate the phenomenon that fault vibration signal is a multicomponent modulation signal. The noise existing in the original signal often appears as white noise in the practical applications, which covers the entire frequency domain. White noise with a signal-to-noise ratio of 2 db is added to the signal. The simulated signal is expressed as follows:

$$x(t) = \sin(2\pi \times 30t) + \frac{1}{4}\sin(2\pi \times 56t) + \eta. \quad (8)$$

The time waveform of the simulated signal collected by using a sampling frequency of 1000 Hz is shown in Figure 1. It can be clearly seen from Figure 1 that the harmonic signal has been seriously distorted by noise. Then, the EMD method is applied to process the simulated signal, and the corresponding results are shown in Figure 2.

From the spectrum diagram of the EMD decomposition shown in Figure 2, it can be seen that the extracted IMF4 mainly includes 30 Hz frequency components, but not 56 Hz frequency components. For the multicomponent simulation signal, the decomposed results also show that the decomposition effect of the EMD algorithm is not ideal, and there is mode aliasing the first three-order intrinsic modal function (IMFs). This is because some useful weak signals are submerged in the decomposed signal. The cubic spline fitting in the decomposition process of the EMD method leads to the deviation of decomposition. In addition, the first three-order IMFs also contain other components of the mode function, and the decomposition of EMD has pseudocomponents.

The VMD method is more effective for the decomposition of simulation signal; not only it can effectively remove pseudocomponents but also each IMF component shows a certain scale of modalities, and there is no mode aliasing between each other. VMD can realize the multiscale

representation of the simulated signal. Compared with Figures 2 and 3, the VMD method has stronger ability in the noise filtering than EMD method. This method successfully suppresses the noise distributed in different frequency bands, and its decomposition effect is better than the EMD decomposition method. EMD cannot effectively remove the noise, especially in the high frequency band. This means that noise still exists in the IMFs generated by EMD.

3.2. Suppression of End Effect. The decomposition results of EMD and VMD are orthogonal to the signal. That is, the energy sum of the decomposed mode functions is equal to the signal energy before the decomposition. However, if the decomposition result has an end effect, it will affect the decomposition accuracy of the signal and produce false components, which will result in a change in the energy sum of the modal function after decomposition. By analyzing the changes in the energy value before and after decomposition, we can understand the inhibitory effect on the signal end effect when the two methods are used for the signal analysis.

The energy expression of the mode function generated after the signal decomposition by the EMD and VMD can be expressed as follows:

$$E = \sqrt{\frac{\sum_{i=1}^{n} x^2(i)}{n}}, \quad (9)$$

where E is the energy of the original signal or the energy of the mode functions after the decomposition, $x(i)$ is the signal sequence, and n is the number of sample points for the signal.

Comparing the deviation between the energy of all mode functions and the energy of the original signal, the evaluation index ξ can be defined as

$$\xi = \frac{\left|\sqrt{\sum_{k=1}^{m} E_k^2} - E_i\right|}{E_i}, \quad (10)$$

where E_i is the energy of the original signal, E_k is the energy of the k-th modal function, and m is the total number of modal functions.

From the definition of the evaluation index depicted above, it can be seen that when ξ is the larger, the energy of the decomposed mode functions will be smaller. That is, the energy leakage after the signal decomposition becomes larger, and the end effect will become stronger. The simulated signals $X_2(t)$ and $X_3(t)$ are constructed as follows:

$$\begin{aligned} X_2(t) = {} & \sin(2\pi \times 3t) + 0.8\sin(2\pi \times 15t) \\ & + 0.4\sin(2\pi \times 45t))0.6\sin(2\pi \times 63t) + \sin(2\pi \times 90t), \end{aligned} \quad (11)$$

$$\begin{aligned} X_3(t) = {} & [1 + 0.5\sin(2\pi \times 3t)] \\ & \times \sin[2\pi \times 5t + \sin(2\pi \times 50t)] \\ & + [1 + \sin(2\pi \times 6t)] \\ & \times \sin[2\pi \times 8t + (0.6 \times \sin(2\pi \times 5t))]. \end{aligned} \quad (12)$$

The equations (8), (11), and (12) of the simulated signal are used to calculate the energy E_i, respectively. The energy E_k of
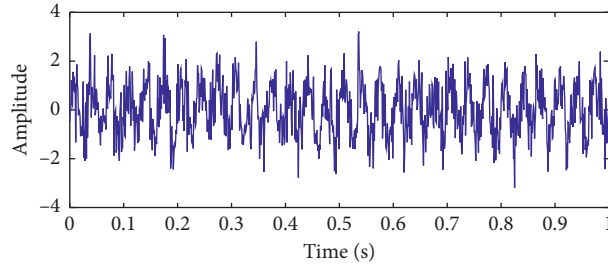
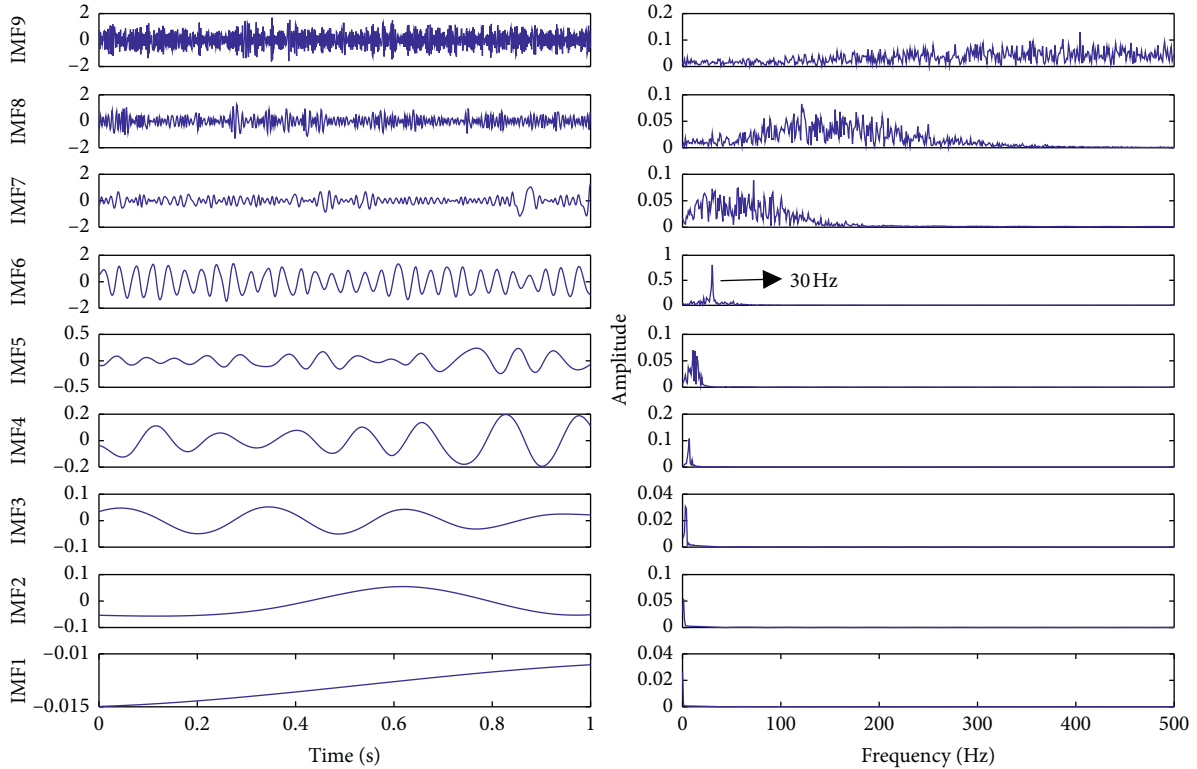FIGURE 1: The time domain plot of the simulated signal.

FIGURE 2: The decomposition results and corresponding frequency spectrum of IMFs with EMD.

each modal function and the evaluation parameter ξ of energy deviation after the decomposition of EMD and VMD are also calculated. The calculation results of the evaluation parameter ξ are shown in Table 1. It can be seen from Table 1 that the values of the evaluation parameter calculated after VMD decomposition are small, which indicates that compared with the EMD method, the energy leakage calculated by the VMD decomposition is smaller and the end effect is not obvious.

4. The Proposed Parameter-Optimized VMD Algorithm

The number of modes k and bandwidth control parameter α affect the accuracy of decomposition in the VMD algorithm. A large number of modes will lead to the redundant information in the result of signal decomposition, while a small number of modes will result in the phenomenon of

mode mixing. On the contrary, a wider filter bandwidth will introduce more noise and interference items into the decomposition result. The narrow filter bandwidth will cause important information missing in the signal decomposition. Therefore, how to choose the optimal parameter combination is the key to eliminate the noise and mode aliasing and extract the feature information accurately in the VMD algorithm. In this section, shuffled frog leaping algorithm (SFLA) is introduced into the algorithm to achieve the combination of optimal influence parameters [27–30]. A multiobjective evaluation function is constructed to select the optimal mode component in the VMD algorithm.

4.1. Shuffled Frog Leaping Algorithm. The SFLA simulates the thought transfer process of frogs in searching for the food according to their population. It combines global
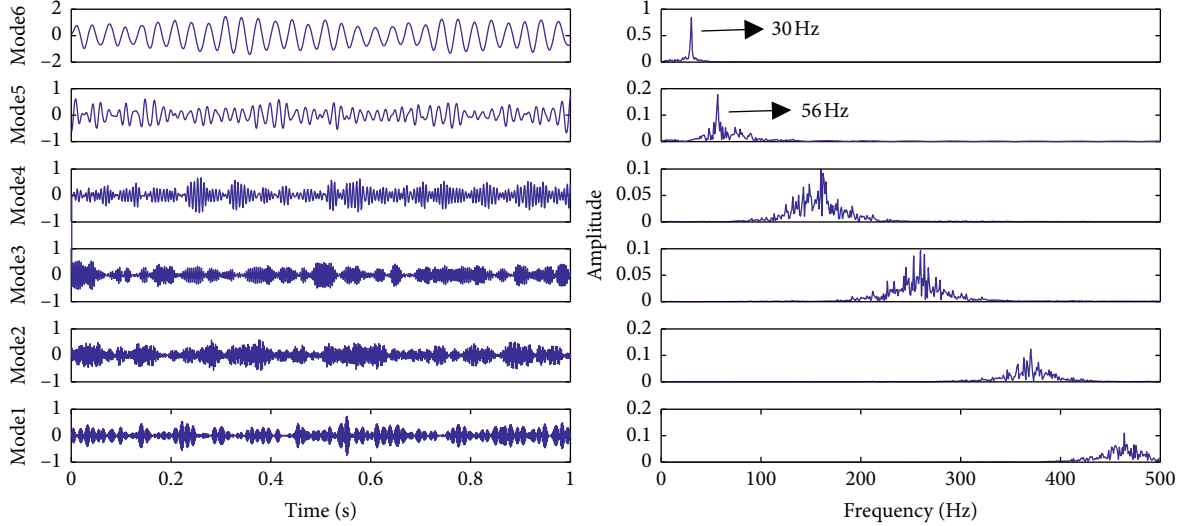
FIGURE 3: The decomposition results and corresponding frequency spectrum of modes with VMD.

TABLE 1: Energy leakage evaluation parameter.

Simulated signal	ξ	
	EMD	VMD
$X_1(t)$	0.2964	0.1620
$X_2(t)$	0.3648	0.0618
$X_3(t)$	0.3643	0.0425

information exchange and local deep search. Local search enables thoughts to be transmitted between local individuals, and hybrid strategies enable the exchange of local thoughts. Through this global information exchange and local depth exploration, the algorithm can jump out of the local extreme points and move towards the global optimum.

An initial population of P frogs is randomly generated within the S-dimensional space. The i-th frog is represented by S variables as $X_i = (x_{i1}, x_{i2}, \ldots, x_{iS})$. In each evolutionary iteration process, all frogs are arranged in a descending order according to the fitness value of the frogs. The population is divided into m subsets. The subset is referred to as memeplexes, and each contains n frogs. The method of allocation is the first frog enters the first memeplex, the second frog goes to the second memeplex, the m frog goes to the m-th memeplex, and the $m+1$ frog goes back to the first memeplex and so forth. Assuming that M^k is a set of frogs for the k-th memeplex, the allocation process can be described as follows:

$$M^k = \left\{ X_{k+m(l-1)} \in P \mid 1 \le l \le n \right\}, \quad 1 \le l \le m. \tag{13}$$

Within each memeplex, the frogs with the best and the worst fitness are identified as X_b and X_ω, respectively. Also, the frog with the global best fitness is identified as X_g. Then, an evolution process is applied to improve only the frog with the worst fitness (i.e., not all frogs) in each cycle. Accordingly, the position of the frog with the worst fitness is adjusted as follows:

$$D_i = \text{Rand} \times (X_b - X_\omega), \tag{14}$$

$$X'_\omega = X_\omega + D_i, D \le D_{\max}, \tag{15}$$

where $D_{\max}$ is the maximum allowed change for the position of the frog.

If the evolution process produces a better frog (solution), it replaces the worst frog. Otherwise, the calculations in equations (14) and (15) are repeated with respect to the global best frog (i.e., X_g replaces X_b). There is no improvement in this situation, and a new solution will be randomly generated, that is, to replace the worst frog with another frog with any fitness. The calculation will continue for a specific number of evolutionary iterations in each memeplex. Therefore, SFLA uses a process similar to the PSO algorithm to simultaneously perform independent local searches in each memeplex.

A predetermined number of memetic evolution steps are performed in each memeplex, and the solution of the evolved memeplexes $\{X_1, X_2, \ldots, X_P\}$ is replaced with a new population, which is called the shuffling process. The shuffling process facilitated the global exchange of information among frogs. Then, the population is sorted in the descending order of fitness value, the position X_g of the best frog of the population is updated, and the frog group is redivided into the memeplexes and evolved in each memeplex until the conversion criterion is met. Generally, the convergence criterion can be defined as follows:

The relative change in the fitness of the global frog within a number of consecutive shuffling iterations is less than a prespecified tolerance.

The maximum predefined number of shuffling iterations has been obtained.

4.2. Parameter Optimization by Using SFLA.

The SFLA is a metaheuristic intelligent optimization algorithm that has good capabilities of global optimization and fast convergence speed. The SFLA combines the advantages of gene-based memetic algorithm (MA) and the social behavior-based particle swarm optimization (PSO) algorithm. Therefore, the SFLA is used to optimize the influencing parameters of VMD, can avoid the intervention of subjective factors, and automatically screen out the best combination of influencing parameters.

Suppose that the population composed of N_{pop} frogs is X in D dimension space, and N_{pop} frogs are divided into N_m subgroups through the descending order. The best individual p_b and the worst individual p_w in the subgroup can be calculated. Group optimal solution s_1 in the maximum number of iterations M can be expressed as follows:

$$s_1 = A \times (p_b - p_w). \tag{16}$$

When using SFLA to optimize the mode number k and bandwidth control parameter α, the fitness function needs to be determined. Each update of the frog is achieved by comparing the fitness values.

Shannon entropy is a good indicator for evaluating signal sparsity. The size of entropy reflects the uniformity of probability distribution. The most uncertain probability distribution (equal probability distribution) has the largest entropy value. In order to reflect the sparseness of the measured signal, the concept of envelope entropy is proposed. The demodulated envelope signal is processed into a probability distribution sequence. The calculated entropy value reflects the sparsity of the original measurement signal [31]. The envelope entropy of the signal can be expressed as follows:

$$\left. \begin{array}{l} E_p = -\sum_{i=1}^{N} p_i \lg p_i \\ p_i = a(i) / \sum_{i=1}^{N} a(i) \end{array} \right\}. \tag{17}$$

In order to search the global optimal component, that is, to extract the mode component with the most abundant feature information from the bearing fault signal, the multiobjective evaluation function is constructed for the selection of the optimal mode component and the calculation of fitness value, which is based on the envelope entropy, the kurtosis, and the correlation coefficients. When the i-th frog is located in the position j (corresponding to a set of parameters α_j and k_j), the kurtosis, the correlation coefficient, and the envelope entropies of all mode components obtained by VMD processing are all calculated. The components with the largest kurtosis value, the highest

correlation, and the smallest envelope entropy are selected and reconstructed as the fitness value in the optimization processing. The optimization method of influencing parameters is briefly described below:

(1) Initialize the parameters: total number of frogs N_{pop}, number of subgroups N_m, number of each group frogs N_f, maximal number of iterations M, random initialization of frog individuals, and initialize the population.

(2) Implement VMD and obtain a set of IMFs.

(3) Construct the global fitness function based on the envelope entropy, the kurtosis, and the correlation coefficients.

(4) Calculate the fitness value of each frog.

(5) Rank the frogs according to their fitness values.

(6) Divided the sorted frogs N_{pop} into N_m subgroups according to the descending order of the objective function. The first frog goes to the first memeplex, the second frog goes to the second memeplex, frog m goes to the m-th memeplex, and frog $m + 1$ goes to the first memeplex.

(7) Determine the best individual of the subgroup p_b, the worst individual p_w, and the optimal solutions in the population S_1; the worst solution is improved by equation (16) in evolutional iteration M.

(8) Update the worst individual and descend the order to the individual to form a new group.

(9) Judge whether the algorithm satisfies the terminating condition and outputs the optimum solution when the algorithm satisfies the termination condition and otherwise moves on to step 6.

5. The Fault Feature Extraction by Parameter-Optimized VMD

The periodic impact energy caused by the failure of the rolling element bearing is weak, and it is relatively difficult to extract fault features due to the effects of noise and signal attenuation. When there is a fault for the rolling bearing, the useful characteristic components usually have very little energy, and it is submerged by the background noise. It is difficult to extract useful fault features. In order to extract the fault feature effectively and realize the fault diagnosis, the parameter-optimized VMD is presented to extract the useful fault features for the fault diagnosis of the rolling element bearing. The vibration signal is decomposed into a series of intrinsic mode functions by the parameter-optimized VMD algorithm. The envelope spectrum technique is utilized to analyze the best signal component. The fault features of the rolling bearing would be easily detected and extracted. The fault features extraction procedure of the parameter-optimized VMD method is briefly described as follows:

(1) Initialize population and parameters: the numbers of subgroup N_m, the numbers of each group frogs N_f, the numbers of iteration within a group N_e, and the numbers of evolutional iteration M.

(2) Optimize VMD parameters by applying SFLA and obtain global optimal parameters k and α.

(3) Decompose the original vibration signal into a set of the IMFs by the improved VMD.

(4) Calculate the envelope entropy, kurtosis, and correlation coefficients of all IMF components.

(5) Select the reconstructed IMF component with the largest kurtosis value, the highest correlation, and the smallest envelope entropy as the optimal component.

(6) Implement the spectrum analysis and compare the fault feature frequency in the envelope spectrum with the theoretical value of the bearing fault and determine the fault.

6. Experimental Results and Analysis

6.1. Simulation Analysis Using the Parameter-Optimized VMD. To quantitatively evaluate the effectiveness of the parameters-optimized VMD method, the simulation signal of rolling element bearings is constructed because the faults of rolling element bearings produce a series of shocks. Therefore, the simulation signal is mainly composed of the impact signal and noise signal generated by a bearing fault. The signal is sampled at 12 kHz. The simulated fault frequency f_i is set to 80 Hz. The resonance frequency f_n is set to 3 kHz. The rotating frequency f_r is 20 Hz. The simulated signal is expressed as follows:

$$\left. \begin{array}{l} x(t) = s(t) + n(t) = \sum_i A_i h(t - iT - \tau_i) + n(t) \\ A_i = 1 + A_0 \sin(2\pi f_r t) \\ h(t) = e^{-Ct} \sin(2\pi f_n t) \end{array} \right\}, \quad (18)$$

where $h(t)$ is the generated waveform of a single impact; A_i is the amplitude of the i-th impact force and considers possible periodic modulations, and $n(t)$ is the white Gaussian white noise; signal-to-noise ratio R_{SNR} is -1 dB; T is the mean spacing among impacts; the attenuation coefficient C is 700; and τ_i is an independent and identically distributed random variable.

The time domain plot and the envelope spectrum of the simulated signal are shown in Figure 4. It can be seen form Figure 4 that the impact signal is submerged in the strong background noise. The resonance frequency band in the spectrum is also not obvious. The frequency and period of the signal cannot be found, and the characteristic frequency of the fault signal cannot be accurately found from the envelope spectrum.

The simulated signal is decomposed by EEMD, and the corresponding frequency spectrum is shown in Figure 5. Obviously, the useful frequency components could not be distinguished from the decomposed IMFs, and they are contaminated with noise. Many parts of IMF 1 are replaced by the intermittent pulse signal. The replaced parts of IMF 1 are shifted to IMF 2 resulting in the phenomenon of mode mixing in the second and the following IMFs. In addition, as noted in Figure 4, the first three IMFs provided more information than the other IMFs, and the rest of the IMFs contain many redundant low-frequency components. In other words, the first three IMFs could be regarded as valid components of the signal, while the other IMFs were the low-frequency pseudocomponents that can mislead the analysis of the signal. The optimal mode component corresponding to the EEMD is IMF1. The envelope spectrum of the optimal mode component decomposed by the EEMD method is shown in Figure 6. It can be seen from Figure 6 that the impact characteristics associated with faults could not be identified, and the feature frequency of fault signals could not be extracted.

The VMD method is used to decompose the simulated signal, and the decomposed simulated signal has 5 mode components. The waveform and the corresponding frequency spectrum are shown in Figure 7. From the decomposition results, the VMD method can realize the adaptive segmentation of each component in the frequency domain, effectively overcome the mode aliasing phenomenon in EEMD, and has stronger noise robustness and weaker end effect than EEMD. The mode component corresponding to the minimum envelope entropy is mode 4, which is selected as the best component, and the envelope analysis is further done. The envelope spectrum of the signal is shown in Figure 8. It can be seen that the characteristic frequency of the fault signal cannot be accurately extracted by the original VMD method.

The parameter-optimized VMD method is implemented to analyze the simulation signal. The decomposition results and corresponding frequency spectrum of modes are shown in Figure 9. According to the decomposition results, the mode component corresponding to the smallest envelope entropy is IMF2, the mode component corresponding to the largest kurtosis is IMF5, and the mode component corresponding to the largest correlation is IMF5. The three mode components are reconstructed and used as the optimal component. The envelope spectrum of the reconstructed signal is shown in Figure 10. It can be seen that the spectral amplitude is prominent at the characteristic frequency 80 Hz, and the corresponding frequency doubling can also be obtained, which means that the parameter-optimized VMD can effectively decompose the fault signal and accurately extract the characteristic frequency of the fault signal.

6.2. Actual Vibration Signal Analysis. To further verify the effectiveness of the proposed parameter-optimized VMD method, the fault feature extraction of the actual experiment is implemented. The vibration data of rolling bearings are provided by Case Western Reserve University bearing data center [32]. The test stand consists of a 2 hp, three-phase induction motor, a torque transducer/encoder, and a dynamometer. The test bearings support the motor shaft at the drive end. Single point faults were introduced to the test bearings. The deep groove ball bearing with the type of 6205-2RS JEM SKF was used in the test. The locations of fault cover inner raceway, outer raceway, and rolling element. The tests are carried out under the four different motor loads with the motor speed. The vibration data were acquired at the sampling frequency of 12 kHz by using the
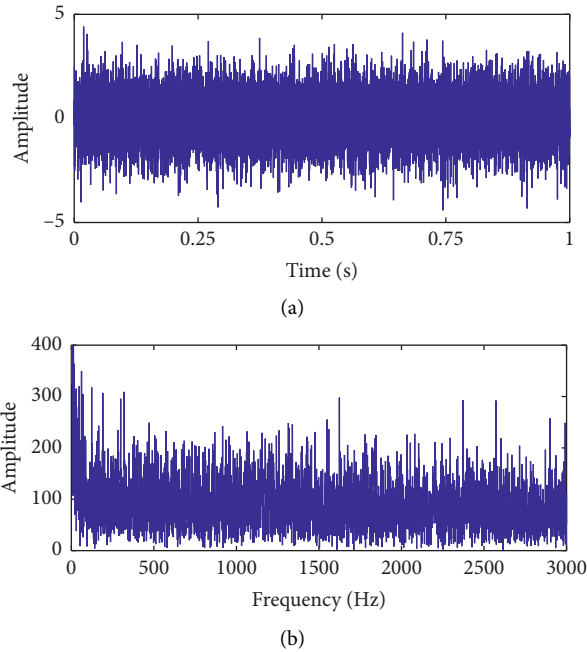
(a)

(b)

FIGURE 4: The time domain plot of the simulated signal and envelope spectrum.
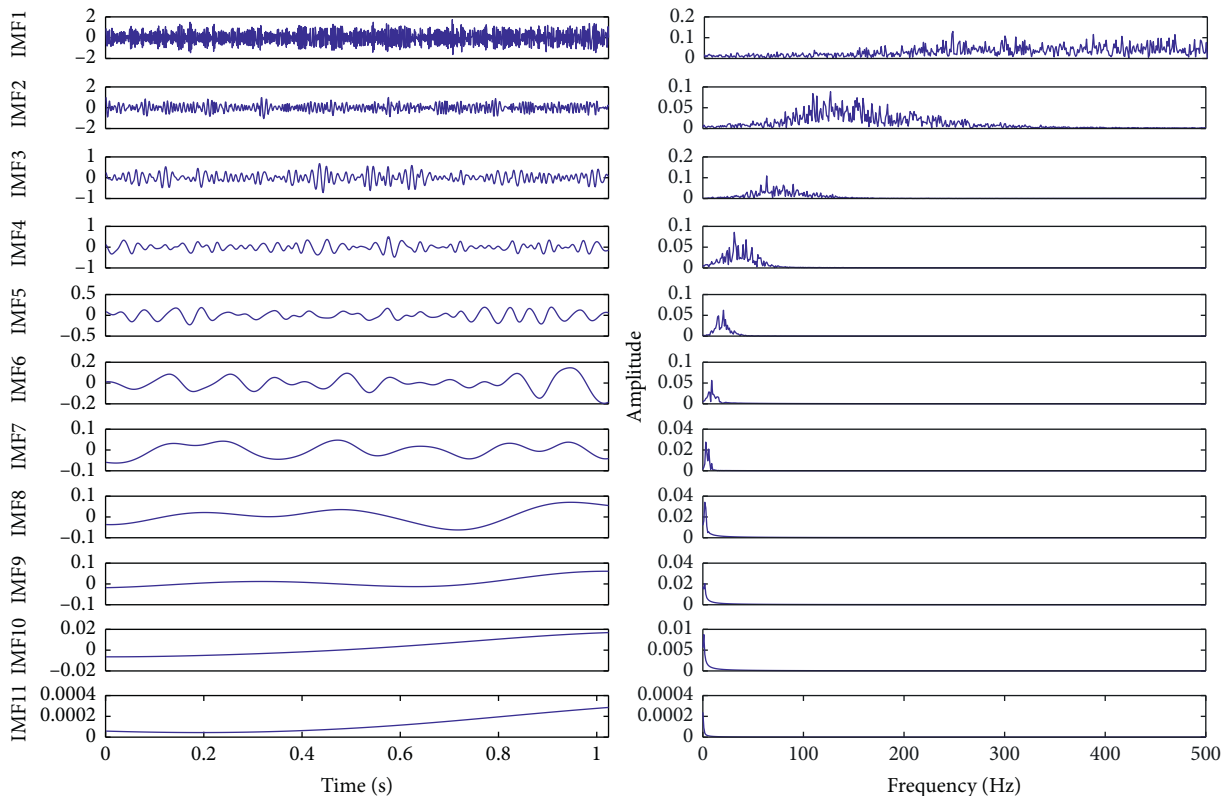

FIGURE 5: The decomposition results and corresponding frequency spectrum of IMFs with EEMD.

accelerometers, which are mounted at the drive end of the motor. The vibration signals of outer race defect with the motor load 0 hp and the fault diameters 7 mills are chosen to extract the fault feature. The characteristic frequency of the outer race defect signal is calculated to be at 107.37 Hz.

The time domain plot of the fault signal with outer race is shown in Figure 11, and Figure 12 shows the decomposition results of EEMD and corresponding the demodulated spectrum of IMFs. The first IMF component decomposed by EEMD contains abundant fault feature information and is
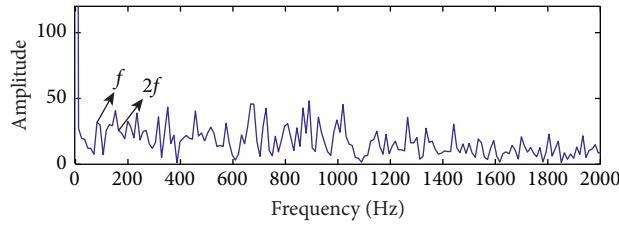
FIGURE 6: The envelope spectrum of optimal component by using EEMD.
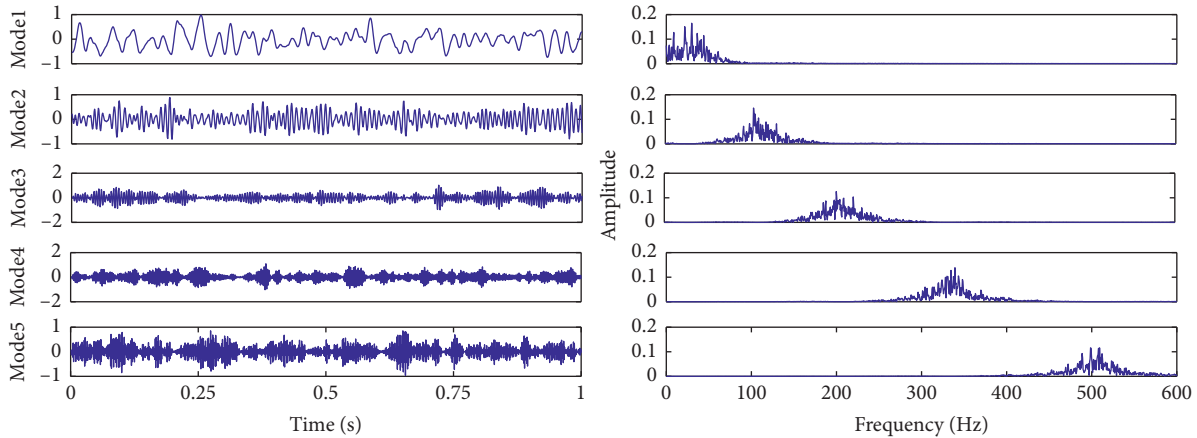
FIGURE 7: The decomposition results and corresponding frequency spectrum of modes with VMD.
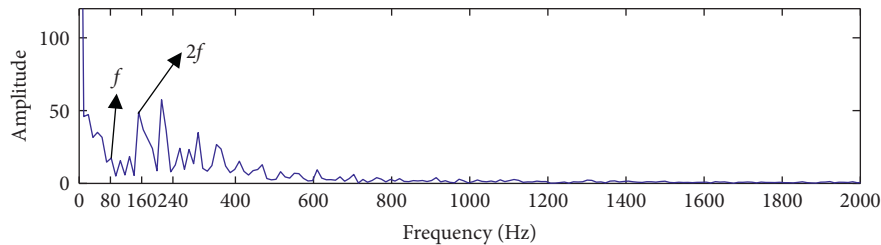
FIGURE 8: The envelope spectrum of optimal component by using VMD.

selected as the optimal feature component. The envelope spectrum of optimal component is shown in Figure 13. It can be seen from Figure 13 that the fault-rated impact features can be perceived from the time-frequency maps of the signals, and the characteristic frequency of the fault signal can be extracted. However, there are still many redundant components, and considerable background noise is also present in the figures, which smears the fault features and consequently leads to the risk of either false alarm or the failure of fault detection.

The VMD method is used to decompose the outer ring defect signal. The decomposition results and the corresponding frequency spectrum of IMFs are shown in Figure 14. According to the calculation results of decomposition, the mode component corresponding to the smallest envelope entropy is IMF6, the mode component corresponding to the largest kurtosis is IMF6, and the mode component corresponding to the largest correlation is IMF3. The three mode components are reconstructed

and used as the optimal component. The envelope spectrum of the signal is shown in Figure 15. It can be seen from Figure 15 that when the default values of the mode number and bandwidth control parameter were adopted, the fault-rated impact features can be perceived from the time-frequency maps of the signals, and the characteristic frequency of the fault signal can be extracted. Compared with the EEMD method, the traditional VMD method has more superiorities than the EEMD method in the noise robustness and the elimination of mode aliasing. However, compared with EEMD, the fault feature extracted from the optimal mode component reconstructed by decomposing the signal with the default value of the influencing parameters is not good.

The proposed parameter-optimized VMD method is utilized to analyze the practical bearing vibration signal. The decomposition results and corresponding frequency spectrum of modes are shown in Figure 16. According to the calculation results of decomposition, the mode
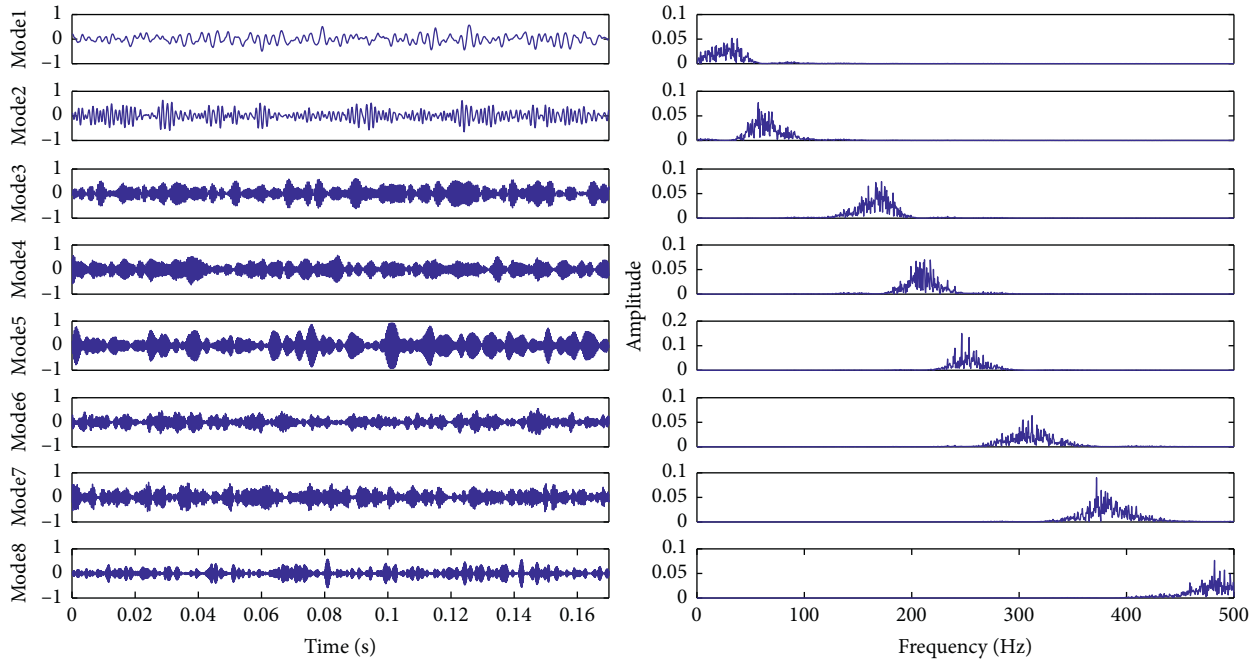
FIGURE 9: The decomposition results of parameter-optimized VMD and corresponding frequency spectrum of modes.
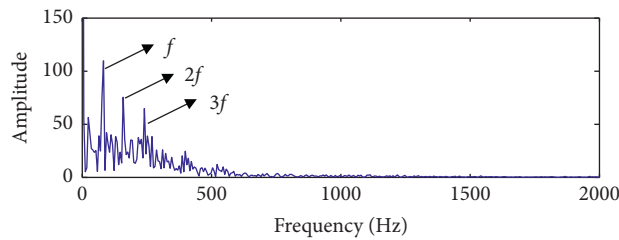

FIGURE 10: The envelope spectrum of optimal component by using the parameter-optimized VMD.
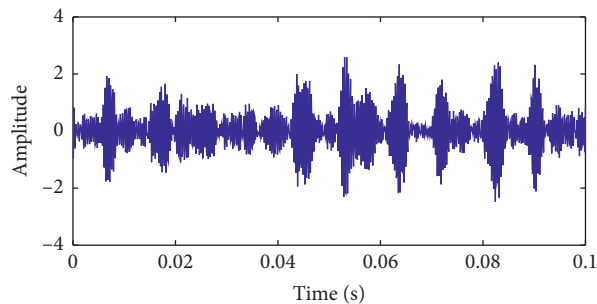

FIGURE 11: The time domain plot of the signal with outer race defect.

component corresponding to the smallest envelope entropy is IMF4, the mode component corresponding to the largest kurtosis is IMF5, and the mode component corresponding to the largest correlation is IMF3. The envelope spectrum of the reconstruction signal is shown in Figure 17. It can be seen that the spectral amplitude is prominent at the characteristic frequency 107.37 Hz,

which means that the parameter-optimized VMD can correctly decompose the fault signal and accurately extract the characteristic frequency of the fault signal. Compared with the VMD method without the optimization and the EEMD method, the fault frequency extracted by the proposed method is more prominent and the noise is also suppressed.
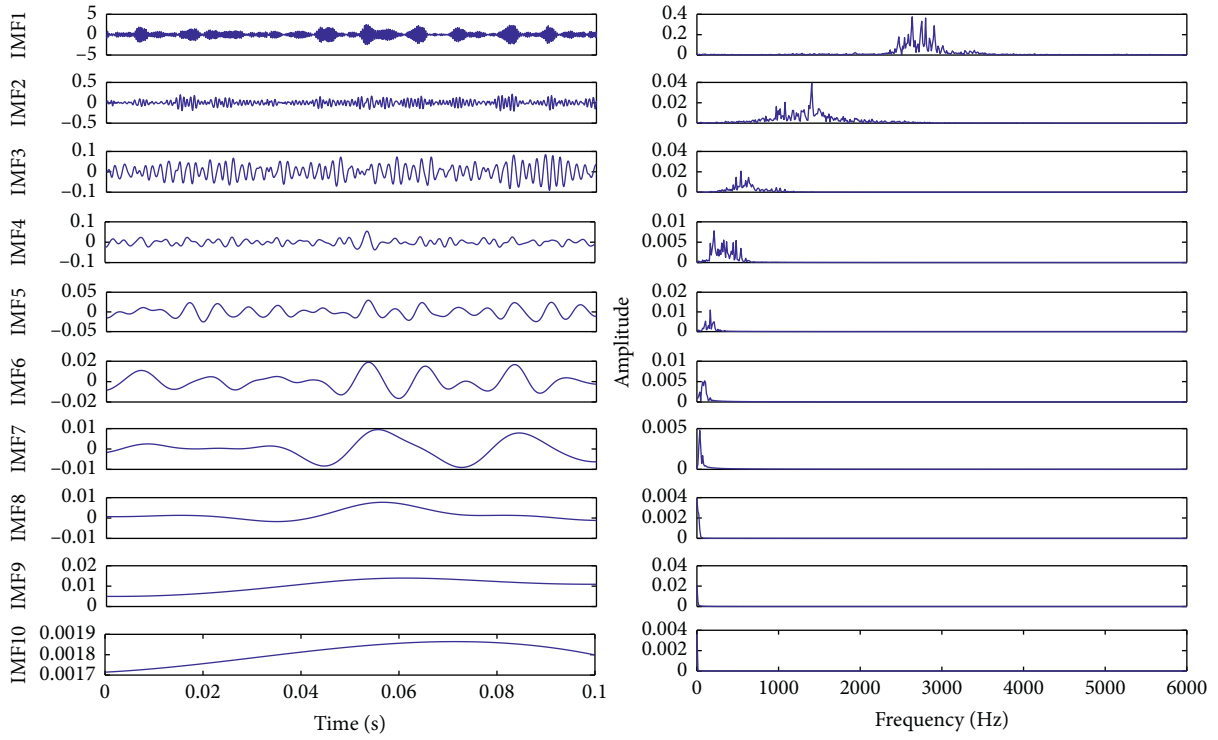
FIGURE 12: The decomposition results of EEMD and corresponding frequency spectrum of IMFs for outer ring defect.
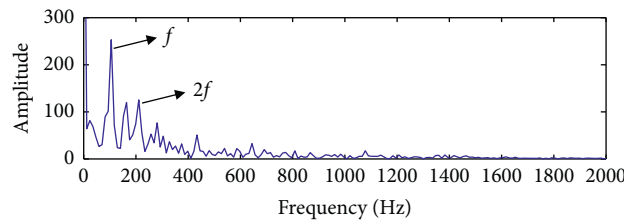

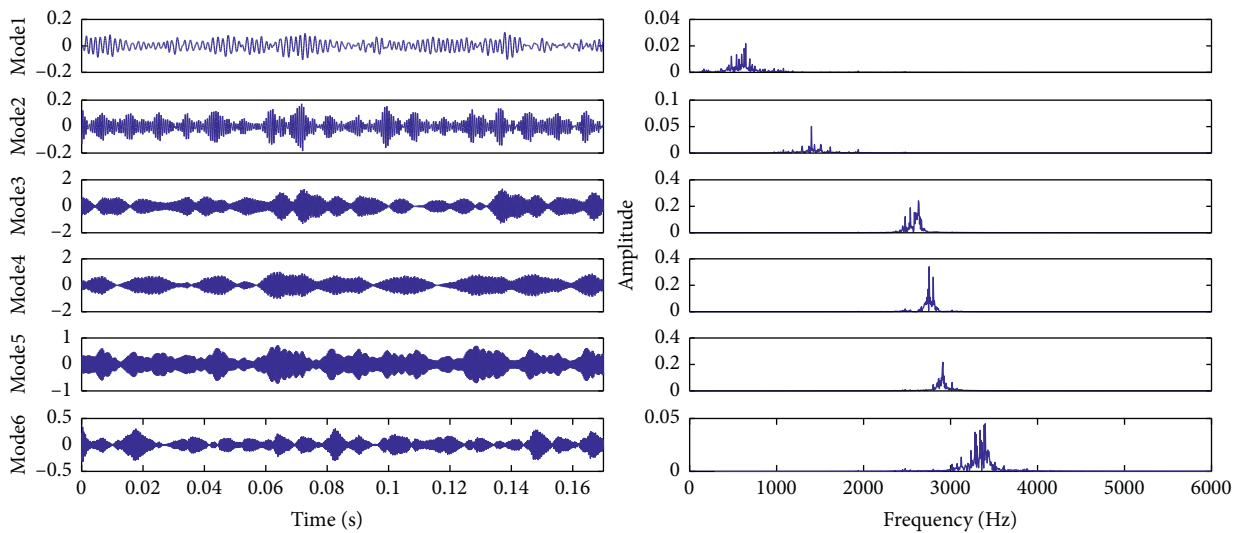FIGURE 13: The envelope spectrum of the optimal component for outer ring defect by using EEMD.

FIGURE 14: The decomposition results and corresponding frequency spectrum of modes for outer ring defect with VMD.
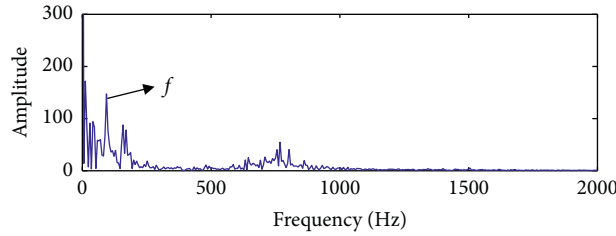
FIGURE 15: The envelope spectrum of optimal component for outer ring defect by using VMD.
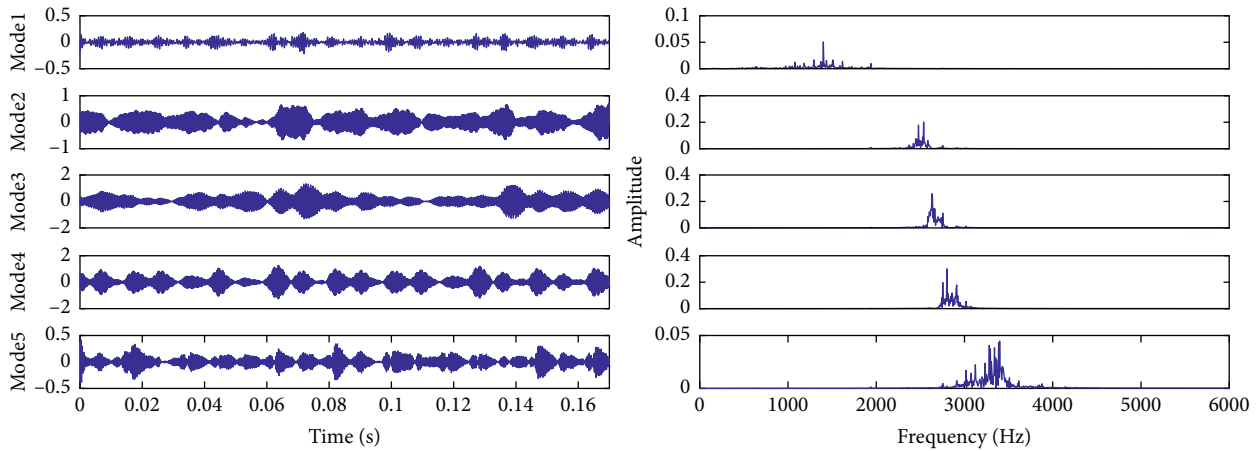

FIGURE 16: The decomposition results and corresponding frequency spectrum of modes by using the parameter-optimized VMD.
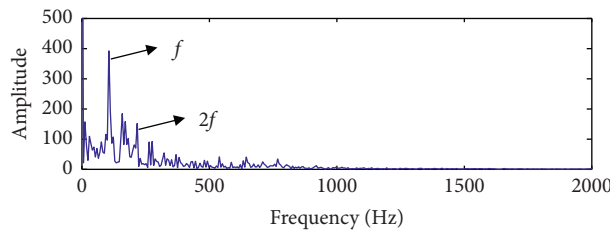

FIGURE 17: The envelope spectrum of the optimal component by using the parameter-optimized VMD.

7. Conclusion

The completely nonrecursive signal modal variational nature of the VMD method makes it more advantageous than EMD in terms of robustness against noise, overcoming end effects, and mode aliasing. This paper analyzes these three aspects. The decomposition accuracy of VMD method is affected by the choice of mode number k and bandwidth control parameter α. The parameter-optimized variational mode decomposition is developed to achieve the accurate decomposition of fault signal and adaptive control of influence parameters. Shuffled frog leaping algorithm is used to implement the optimization the influence parameters. The multiobjective evaluation function is constructed to select the optimal mode component. The envelope spectrum technique is used to analyze the optimal mode component. According to the characteristics of the vibration signal, we build the simulation signal to verify the feasibility and effectiveness of the signal and also use the vibration data of Western Reserve University to verify the proposed method. The experimental results show that the proposed parameter-optimized VMD method can correctly decompose the fault signal and accurately extract the characteristic frequency of the fault signal. Compared with the VMD method without the optimization and the EEMD method, the fault frequency extracted by the proposed method is more prominent and the noise is also suppressed. The proposed method also provides a new way to solve the problem for the analysis of vibration signal.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

Acknowledgments

References

[1] H. Zhao, S. Zuo, M. Hou et al., "A novel adaptive signal processing method based on enhanced empirical wavelet transform technology," *Sensors*, vol. 18, no. 10, Article ID 3323, 2020.

[2] H. Zhao, H. Liu, J. Xu, and W. Deng, "Performance prediction using high-order differential mathematical morphology gradient spectrum entropy and extreme learning machine," *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4165–4172, 2019.

[3] X. Y. Zhang and J. Z. Zhou, "Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines," *Mechanical Systems and Signal Processing*, vol. 41, no. 1-2, pp. 127–140, 2013.

[4] J. Qu, Z. Zhang, and T. Gong, "A novel intelligent method for mechanical fault diagnosis based on dual-tree complex wavelet packet transform and multiple classifier fusion," *Neurocomputing*, vol. 171, pp. 837–853, 2016.

[5] Q. Tong, Z. Sun, Z. Nie, Y. Lin, and J. Cao, "Sparse decomposition based on ADMM dictionary learning for fault feature extraction of rolling element bearing," *Journal of Vibroengineering*, vol. 18, no. 8, pp. 5204–5216, 2016.

[6] F. Jiang, Z. Zhu, W. Li, G. Chen, and G. Zhou, "Robust condition monitoring and fault diagnosis of rolling element bearings using improved EEMD and statistical features," *Measurement Science & Technology*, vol. 25, no. 2, pp. 1–14, 2014.

[7] Y. Lei, J. Lin, Z. He, and Y. Zi, "Application of an improved kurtogram method for fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1738–1749, 2011.

[8] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *Institute of Electrical and Electronics Engineers Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7319–7327, 2020.

[9] W. Deng, J. Xu, Y. Song, and H. Zhao, "Differential evolution algorithm with wavelet basis function and optimal mutation strategy for complex optimization problem," *Applied Soft Computing*, vol. 10, Article ID 106724, 2020.

[10] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics-A tutorial," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485–520, 2011.

[11] G. Dong and J. Chen, "Noise resistant time frequency analysis and application in fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 33, pp. 212–236, 2012.

[12] S. Abbas, A. Rafsanjani, A. Farshidianfar, and N. Irani, "Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine," *Mechanical Systems and Signal Processing*, vol. 21, no. 7, pp. 2933–2945, 2007.

[13] W. J. Wang and P. D. Mcfadden, "Application of wavelets to gearbox vibration signals for fault detection," *Journal of Sound and Vibration*, vol. 192, no. 5, pp. 927–939, 1996.

[14] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199–221, 2004.

[15] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[16] J. Cheng, D. Yu, J. Tang, and Y. Yang, "Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery," *Shock and Vibration*, vol. 16, no. 1, pp. 89–98, 2009.

[17] V. K. Rai and A. R. Mohanty, "Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2607–2615, 2007.

[18] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 01, no. 01, pp. 1–41, 2009.

[19] J. S. Smith, "The local mean decomposition and its application to EEG perception data," *Journal of the Royal Society Interface*, vol. 2, no. 5, pp. 443–454, 2005.

[20] B. J. Chen, Z. J. He, X. F. Chen et al., "A demodulating approach based on local mean decomposition and its applications in mechanical fault diagnosis," *Measurement Science and Technology*, vol. 22, no. 5, Article ID 055704, 2011.

[21] C. Park, D. Looney, M. M. Van Hulle, and D. P. Mandic, "The complex local mean decomposition," *Neurocomputing*, vol. 74, no. 6, pp. 867–875, 2011.

[22] W. Y. Liu, W. H. Zhang, J. G. Han, and G. F. Wang, "A new wind turbine fault diagnosis method based on the local mean decomposition," *Renewable Energy*, vol. 48, pp. 411–415, 2012.

[23] M. G. Frei and I. Osorio, "Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals," *Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences*, vol. 463, no. 2078, pp. 321–342, 2007.

[24] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *Institute of Electrical and Electronics Engineers Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[25] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 243–251, 2015.

[26] Y. Liu, G. Yang, M. Li, and H. Yin, "Variational mode decomposition denoising combined the detrended fluctuation analysis," *Signal Processing*, vol. 125, pp. 349–364, 2016.

[27] M. Eusuff, K. Lansey, and F. Pasha, "Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization," *Engineering Optimization*, vol. 38, no. 2, pp. 129–154, 2006.

[28] M. M. Eusuff and K. E. Lansey, "Optimization of water distribution network design using the Shuffled Frog Leaping Algorithm," *Journal of Water Resources Planning and Management*, vol. 129, no. 3, pp. 210–225, 2003.

[29] R. Chen, S.-K. Guo, X.-Z. Wang, and T.-L. Zhang, "Fusion of multi-RSMOTE with fuzzy integral to classify bug reports with an imbalanced distribution," *Institute of Electrical and*

Electronics Engineers Transactions on Fuzzy Systems, vol. 27, no. 12, pp. 2406–2420, 2019.

[30] B. Amiri, M. Fathian, and A. Maroosi, "Application of shuffled frog-leaping algorithm on clustering," *The International Journal of Advanced Manufacturing Technology*, vol. 45, no. 1-2, pp. 199–209, 2009.

[31] J. Sun, Q. Xiao, J. Wen, and F. Wang, "Natural gas pipeline small leakage feature extraction and recognition based on LMD envelope spectrum entropy and SVM," *Measurement*, vol. 55, pp. 434–443, 2014.

[32] Seeded Fault Test Data from Bearing Data Center of Case Western Reserve University, 2016, http://csegroups.case.edu/bearingdatacenter.