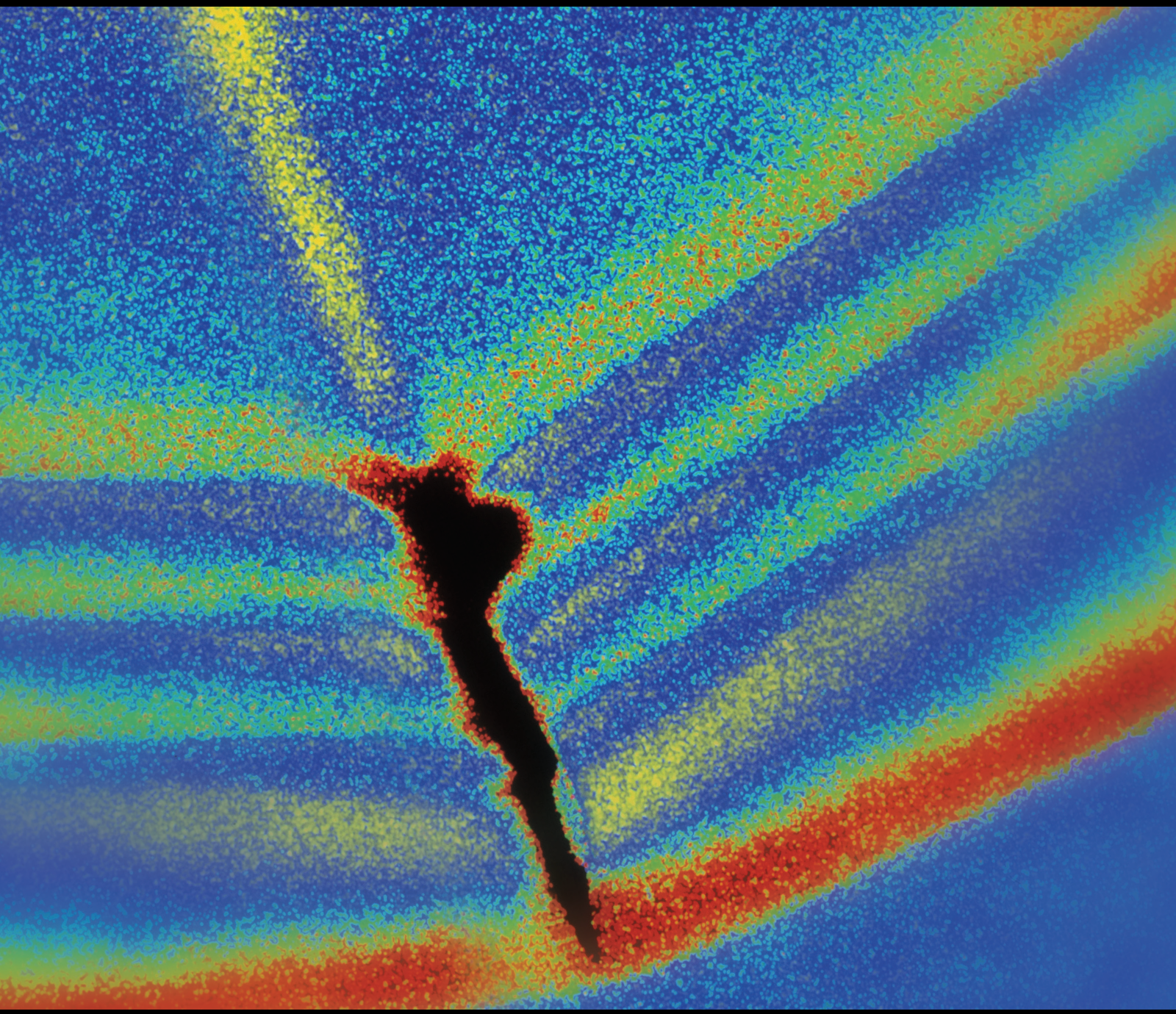
Shock and Vibration

Adaptive Mode Decomposition-Based Fault Diagnosis

Lead Guest Editor: Jinde Zheng
Guest Editors: Ke Feng, Haiyang Pan, and Jun-Sheng Cheng



Adaptive Mode Decomposition-Based Fault Diagnosis

Shock and Vibration

Adaptive Mode Decomposition-Based Fault Diagnosis

Lead Guest Editor: Jinde Zheng
Guest Editors: Ke Feng, Haiyang Pan, and Jun-Sheng Cheng

Chief Editor

Associate Editors

Academic Editors

Contents

Hindawi
Shock and Vibration
Volume 2022, Article ID 9925681, 13 pages
https://doi.org/10.1155/2022/9925681

Research Article

Rolling Bearing Fault Diagnosis Based on Component Screening Vector Local Characteristic-Scale Decomposition

Tengfei Guan, Shijun Liu, Wenbo Xu, Zhisheng Li, Hongtao Huang, and Qi Wang

Zhengzhou Research Institute of Mechanical Engineering Co., Ltd, Zhengzhou 45000, China

Correspondence should be addressed to Tengfei Guan; guantengfei1314@126.com

Received 9 April 2021; Revised 1 June 2021; Accepted 23 December 2021; Published 10 January 2022

Academic Editor: Ke Feng

The fault vibration signal of a bearing has nonstationary and nonlinear characteristics and can be regarded as the combination of multiple amplitude- and frequency-modulation components. The envelope of a single component contains the fault characteristics of a bearing. Local characteristic-scale decomposition (LCD) can decompose the vibration signal into a series of multiple intrinsic scale components. Some components can clearly reflect the running state of a bearing, and fault diagnosis is conducted according to the envelope spectrum. However, the conventional LCD takes a single-channel signal as the research object, which cannot fully reflect the characteristic information of the rotor, and the analysis results based on different channel signals of the same section will be inconsistent. To solve this problem, based on full vector spectrum technology, the homologous dual-channel information is fused. A vector LCD method based on cross-correlation coefficient component selection is given, and a simulation analysis is completed. The effectiveness of the proposed method is verified by simulated signals and experimental signals of a bearing, which provides a method for bearing feature extraction and fault diagnosis.

1. Introduction

Rotating machinery is developing in the direction of high speeds, heavy loads, and high reliability, which places higher requirements on mechanical transmission equipment [1]. The operational state of mechanical equipment is changing, and its safe, stable, and reliable operation must be ensured. Rolling bearings are widely used in mechanical equipment, and their working condition greatly affects its operation [2]. Owing to complex operating conditions and changing external environment, rolling bearings are prone to failure [3]. It is of great significance to monitor their working status and diagnose their fault degree [4, 5]. Some studies have focused on the fault features of rotating machinery through modern signal processing methods [6–8].

Fault vibration signals of bearings are usually weak nonstationary signals with complex frequency components. The key to fault diagnosis of rolling bearings is to extract effective feature information from vibration signals containing complex frequencies [9]. Vibration analysis and fault diagnosis have received considerable attention [10–14] and have been adopted to process nonstationary and nonlinear vibration signals. Among them, signal decomposition methods contribute much. Tiwari [15] described a self-adaptive signal decomposition technique, concealed component decomposition (CCD), as the basis of a precise bearing fault diagnosis model. Ying [16] introduced a novel permutation entropy-based improved uniform phase empirical mode decomposition (PEUPEMD) method and obtained better analysis than comparative methods about empirical mode decomposition (EMD) in decomposing accuracy and mode mixing suppression. Patel [17] applied variational mode decomposition (VMD) to filter out nonstationarities due to variable speed conditions and provided a complete diagnostic solution for the spur gear systems. Li [18] presented a local mean decomposition (LMD) method based on an improved compound interpolation envelope, whose effect was comparable to or slightly better than that of other methods. Zheng [19] proposed local characteristic-scale decomposition (LCD), a nonstationary signal analysis method that adaptively decomposes a signal to a series of intrinsic scale components in different scales. With good

compatibility, LCD methods have seen new applications, such as the local characteristic-scale decomposition-Teager energy operator (LCD-TEO) [20], improved local characteristic-scale decomposition (ILCD) [21], and piecewise cubic Hermite interpolating polynomial-local characteristic-scale decomposition (PCHIP-LCD) [22].

However, the conventional LCD methods focus on single-channel signals, which probably cause incomplete fault feature extraction. In the fault diagnosis of rotating machinery, the sensor information collected by a multi-sensor system is related to the same or different sides of rotating machinery in the same environment. There is an inevitable connection between various types of information, which are different in time, space, credibility, and expression, whose focuses and uses are not exactly the same, and whose requirements for information processing and management are different. If the information collected by each source is considered in isolation, then their internal connections and characteristics are lost. Multi-source information fusion can solve this problem [23]. Multi-sensor data fusion methodologies include the holospectrum [24–27], full spectrum [28–31], and full vector spectrum [32]. Proposed by Han [32], the full vector spectrum has been widely studied and applied in engineering [33–35] and has formed the basis of many compound methods. Chen [36] applied full-vector signal acquisition and information fusion to fault prediction. Gong [37] combined the full vector spectrum with ensemble empirical decomposition and applied it to the diagnosis of gear faults. Yu [38] introduced the empirical wavelet transform and variance contribution rate to the full vector spectrum, which improved the adaptability and accuracy of full vector information fusion.

Based on the above analysis, the main contributions of this paper are as follows:

(1) A signal processing method, vector LCD, is proposed, which fully considers homologous signals and intrinsic scale components (ISCs)

(2) Vector LCD can simplify the analysis of ISCs by taking the cross-correlation coefficient in screening components

(3) The fusion of optimal components can obtain more complete and accurate fault features

The remainder of this article is arranged as follows. Section 2 shows the calculation of LCD, presents the theory of the cross-correlation coefficient, describes the principles of the conventional full vector spectrum, and introduces a method for bearing feature extraction and fault diagnosis based on the correlation coefficient vector LCD. In Section 3, the proposed methodology is verified through application to the homologous simulation signals of a rolling bearing. In Section 4, the rolling bearing experimental data from two directions of sensors are used to validate vector LCD. Conclusions are given in Section 5.

2. Theoretical Description of Vector LCD

The proposed fault diagnosis method using vector local characteristic-scale decomposition (Vector LCD) is presented in Figure 1. Vibration signals were obtained from a test stand by a signal acquisition module. Then, the vector LCD method was used to compute and analyze the data. Next, the fusion data were enveloped into the spectrum. Lastly, the fault frequency features were matched with the specific fault type and the failure reason was located.

2.1. Local Characteristic-Scale Decomposition. According to the extreme value of a signal, the LCD can adaptively decompose nonlinear and nonstationary signals to a series of ISCs satisfying the following conditions:

(1) The length between any two adjacent extreme points of the data sample is monotonic

(2) If the extreme point in a data sample is X_k ($k = 1, 2, ..., M$) and τ_k is the corresponding time, then any two maximum (or minimum) value points (τ_k, X_k), (τ_{k+2}, X_{k+2}) can be connected to form a line segment. τ_{k+1} is the corresponding time of the maximum (or minimum) value point (τ_{k+1}, X_{k+1}) in the middle of the line segment. The corresponding function value at this moment is

$$A_k = X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\left(X_{k+2} - X_k\right). \tag{1}$$

The ratio of the function value to the maximum (or minimum)

$$a\left[X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\left(X_{k+2} - X_k\right)\right] + (1-a)X_{k+1} = 0, \tag{2}$$

remains unchanged, where $a \in (0, 1)$ is a constant, and $a = 1/2$ for frequency modulated, amplitude modulated, amplitude-frequency modulated, and sine-cosine signals.

On the basis of the ISC, LCD can decompose any signal $x(t)$ to a series of ISCs, as follows [39]:

(1) Find all extreme points of $x(t)$ and their corresponding moments τ_k ($k = 1, 2, ..., M$), set $a = 1/2$, and make a linear transformation for $x(t)$ between any two extreme points,

$$P_1 = L_k + \frac{L_{k+1} - L_k}{X_{k+1} - X_k}\left(x_t - X_k\right),$$

$$L_{k+1} = a\left[X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}\left(X_{k+2} - X_k\right)\right] + (1-a)X_{k+1}, \tag{3}$$

where $t \in (\tau_k, \tau_{k+1})$.

(2) Subtract $P_1(t)$ from the original signal $x(t)$ to get a new signal,

$$I_1(t) = x(t) - P_1(t). \tag{4}$$

(3) Ideally, $I_1(t)$ can be used as the first ISC. At this moment L_{k+1} is equal to zero; in practice, assuming a variable $\triangle e$, the iteration ends when $|L_{k+1}| \le \triangle e$. If

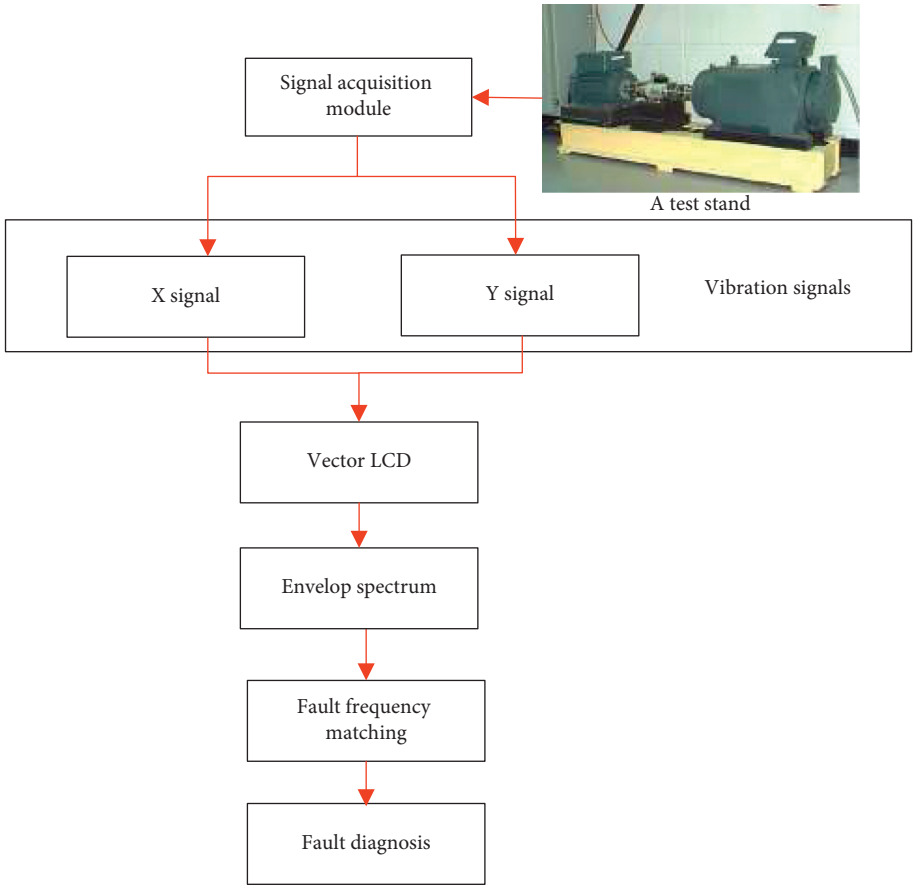


Figure 1: Block diagram of the proposed fault diagnosis method using vector LCD.

$I_1(t)$ does not meet the two conditions of ISC, the above steps are repeated k times until $I_k(t)$ satisfies the conditions and denote $I_k(t)$ as the first ISC $c_1(t)$ of $x(t)$.

(4) Subtract $c_1(t)$ from $x(t)$ to get a new signal, r_1. Take r_1 as the original data and repeat steps 1–3 to get the second ISC component, $c_2(t)$, of $x(t)$. Repeat n times to get n ISCs of the signal $x(t)$. The function does not terminate until r_n is monotonic:

$$x(t) = \sum_{p=1}^{n} c_p(t) + r_n(t). \tag{5}$$

It can be seen from equation (5) that the signal $x(t)$ can be reconstructed by n ISCs and a monotonic signal.

2.2. Correlation Coefficient. Correlation is a kind of nondeterministic relationship, and the correlation coefficient measures the degree of linear correlation between variables. The correlation coefficient between sequences $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ can be calculated as

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}, \tag{6}$$

where $\overline{x}$ and $\overline{y}$ are the mean values of sequences x-and y, respectively.

Suppose x_{nor} and y_{nor} are normal operating signals perpendicular to each other, and x_{abn}, y_{abn} are the homologous signals when faults occur. After the decomposition of LCD, we obtain four ISCs, ISC_{xnork}, ISC_{ynork}, ISC_{yabnk}, and ISC_{xabnk}, where $k = 1, 2, \ldots, N$ is the order of an ISC. The cross-correlation coefficient is calculated as follows:

(1) Calculate the correlation coefficients between ISC_{xnork} and x_{nor}, ISC_{ynork}, and y_{nor}, and find their average, U_k;

(2) Calculate the correlation coefficients between ISC_{xabnk} and x_{abn}, ISC_{yabnk}, and ISC_{yabnk}, and find their average, V_k;

(3) Calculate the correlation coefficients between ISC_{xnork} and ISC_{ynork}, ISC_{xabnk}, and ISC_{yabnk}, and find their average, W_k;

(4) Calculate the sensitivity factor,

$$S_k = \frac{U_k + V_k}{2} - W_k, \tag{7}$$

where U_k and V_k indicate the degree of correlation between the decomposed ISC and the initial signal. The larger the value, the more similar are ISC and the initial signal. The W_k indicates the correlation between ISCs of the same order. The smaller the value, the greater the change in the ISC. Overall, the larger the S_k, the more sensitive the ISC of this order, and the more it can reflect spectral changes.

2.3. Full Vector Spectrum. To overcome limitations due to incomplete and inaccurate sensor information, two orthogonal sensors are usually fixed on the same section of the rotor in the field test of large rotating machinery. The full vector spectrum technology meets the accuracy and reliability requirements of condition monitoring and fault diagnosis. The vortex phenomenon of the rotor is the combined effect of each harmonic frequency, and the vortex intensity at each harmonic frequency is the basis for fault judgment and identification. The space rotation trajectory of each harmonic is an ellipse, and the maximum vibration vector is in its long-axis direction [40].

Suppose the cross section channel signals ($\{x_k\}$ and $\{y_k\}$) are perpendicular to each other and form them into a complex signal ($\{z_k\} = \{x_k\} + j\{y_k\}$), only a single Fourier transform (FT) of the complex signal is needed to obtain the characteristic information required by the full vector spectrum under each harmonic frequency. The algorithm is robust, it greatly reduces calculation, and it is compatible with conventional analysis methods. When processing a single-channel signal, the algorithm is still valid and can meet real-time requirements. The characteristic information includes the main vibration vector R_{Lk}, assistant vibration vector R_{sk}, angle $\varnothing_{\alpha k}$ between the main vibration vector and the x-axis, and the elliptical trajectory's initial phase angle α_k, which are described by

$$\begin{cases} R_{Lk} = \dfrac{1}{2N}\left[\left|Z_k\right| + \left|Z_{N-k}\right|\right], \\ R_{sk} = \dfrac{1}{2N}\left[\left|Z_k\right| - \left|Z_{N-k}\right|\right], \\ \tan 2\varnothing_{\alpha k} = \dfrac{Z_{Ik}Z_{R(N-k)} - Z_{Rk}Z_{I(N-k)}}{Z_{Ik}Z_{I(N-k)} + Z_{Rk}Z_{R(N-k)}}, \\ \tan \alpha_k = \dfrac{Z_{Ik} + Z_{I(N-k)}}{Z_{Rk} + Z_{R(N-k)}}, \end{cases} \tag{8}$$

where $k = 1, 2, \ldots, N/2–1$. The characteristic information of each harmonic trace from equation (8) is the main characteristic information under each harmonic trace of full vector spectrum technology.

Figure 2 is a flowchart of the proposed method. At first, the LCD is applied to the homology information acquired from two orthogonal sensors. The cross-correlation coefficient is selected to choose the ISCs of orthogonal signals, the next optimal ISC of orthogonal signals can be obtained, the vector ISCs are formed through the optimal ISC component fusion, and the vector ISCs are enveloped and demodulated. The proposed method can identify the nonlinear characteristics of fault signals for fault diagnosis.

3. Simulation Analysis

Analog signals were analyzed to validate the effectiveness of vector local characteristic-scale decomposition in processing homologous signals. For the rolling bearing signal, the vibration signal at the time of failure presents a modulation phenomenon. The vibration signal of a rolling bearing with an outer ring fixed structure is

$$x(t) = \alpha \sin\left(2\pi f_b\right)\left(1 + \beta \sin\left(2\pi f_r t\right)\right), \tag{9}$$

where f_b is the passing frequency of the inner ring of the rolling bearing and f_r is the rotation frequency of the rotor.

Based on this, the following analog acceleration signal is constructed:

$$\begin{cases} x_{\text{nor}}(t) = \sin(40\pi t) + \sin(600\pi t), \\ y_{\text{nor}}(t) = \cos(40\pi t) + \cos(600\pi t), \\ x_{\text{abn}}(t) = \sin(40\pi t) + 0.8\sin(600\pi t) + 0.9(1 + \sin(40\pi t))\sin(200\pi t), \\ y_{\text{abn}}(t) = \cos(40\pi t) + 1.2\cos(600\pi t) + (1 + \sin(40\pi t))\cos(200\pi t), \end{cases} \tag{10}$$

where $x_{\text{nor}}(t)$ and $y_{\text{nor}}(t)$ are two initial vibration signals that are perpendicular to each other under normal operating conditions, $x_{\text{abn}}(t)$ and $y_{\text{abn}}(t)$ are two initial vibration signals that are perpendicular to each other under abnormal operating conditions, the sampling frequency $f_s = 1024Hz$, and there are $N = 2048$ sampling points.

The LCD is applied to decompose the two-channel signals corresponding to normal operating conditions. The time-domain waveforms and the ISCs are shown in Figure 3 for x-channel signals and Figure 4 for y-channel signals.

The LCD is then applied to decompose the two-channel signals corresponding to abnormal operating conditions, for which the time-domain waveforms and the ISCs of x- and y-channel signals are shown in Figures 5 and 6, respectively.

The correlation coefficients between ISCs and initial signals can be calculated, i.e., between $\text{ISC}_{\text{xnori}}$ and x_{nor}, $\text{ISC}_{\text{ynori}}$ and y_{nor}, $\text{ISC}_{\text{xabni}}$ and x_{abn}, and $\text{ISC}_{\text{yabni}}$ and y_{abn}, as shown in Table 1.

Figures 5 and 6 show that the initial signals are decomposed to six ISCs with different frequency bands. There is a great difference in amplitude between the x- and y-directions. If the signal is analyzed in a certain direction alone, the analysis results will also be quite different. Therefore, information fusion is necessary. The full vector fusion is introduced. Considering the value of the correlation coefficient, the first-order ISC_{xnor} and third-order ISC_{ynor} are closer to the initial signals than the others. The second-order ISC_{xabn} and second-order ISC_{yabn} have more

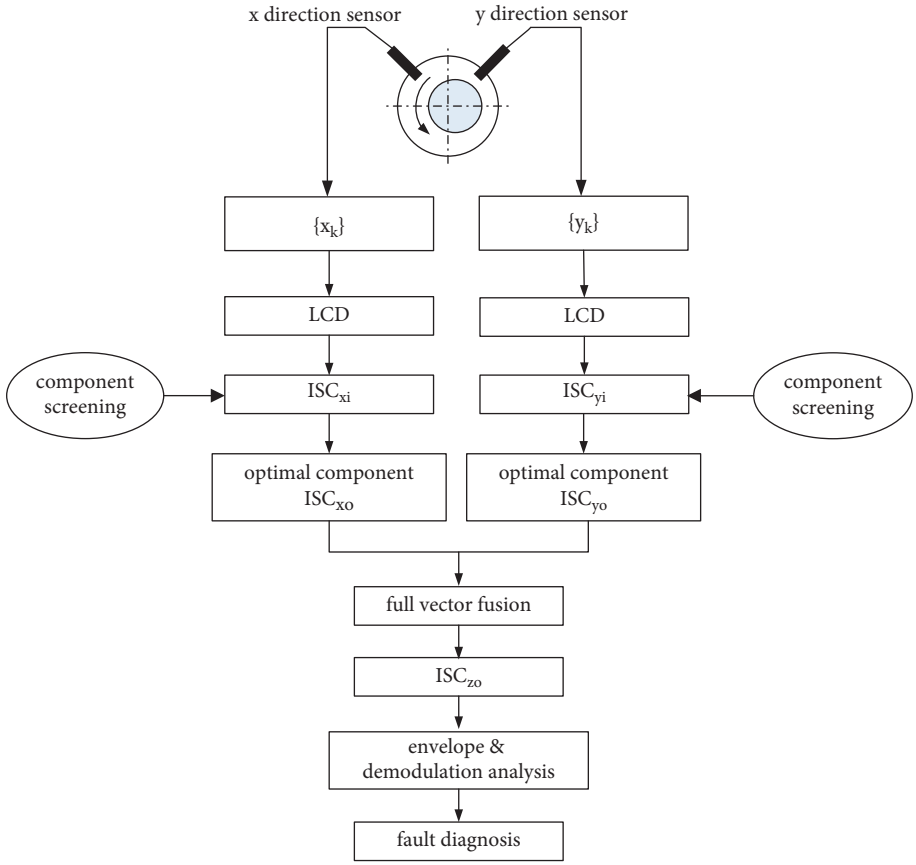


Figure 2: Flowchart of vector LCD.

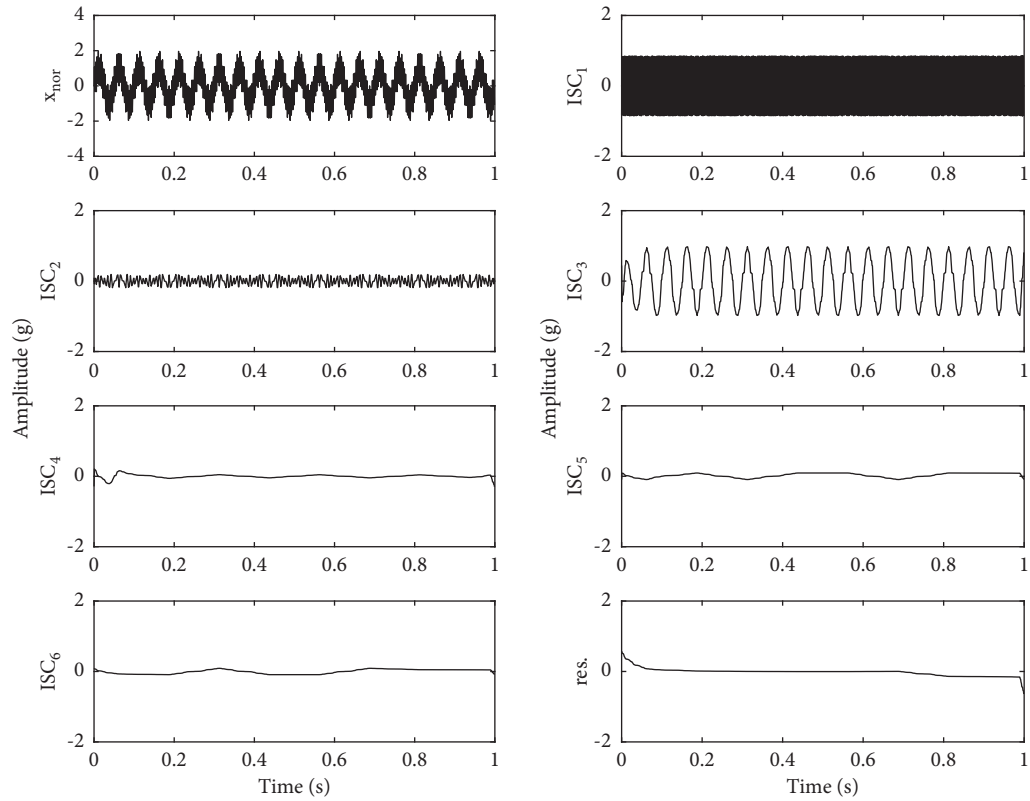


Figure 3: LCD decomposition results of x-channel under normal conditions.

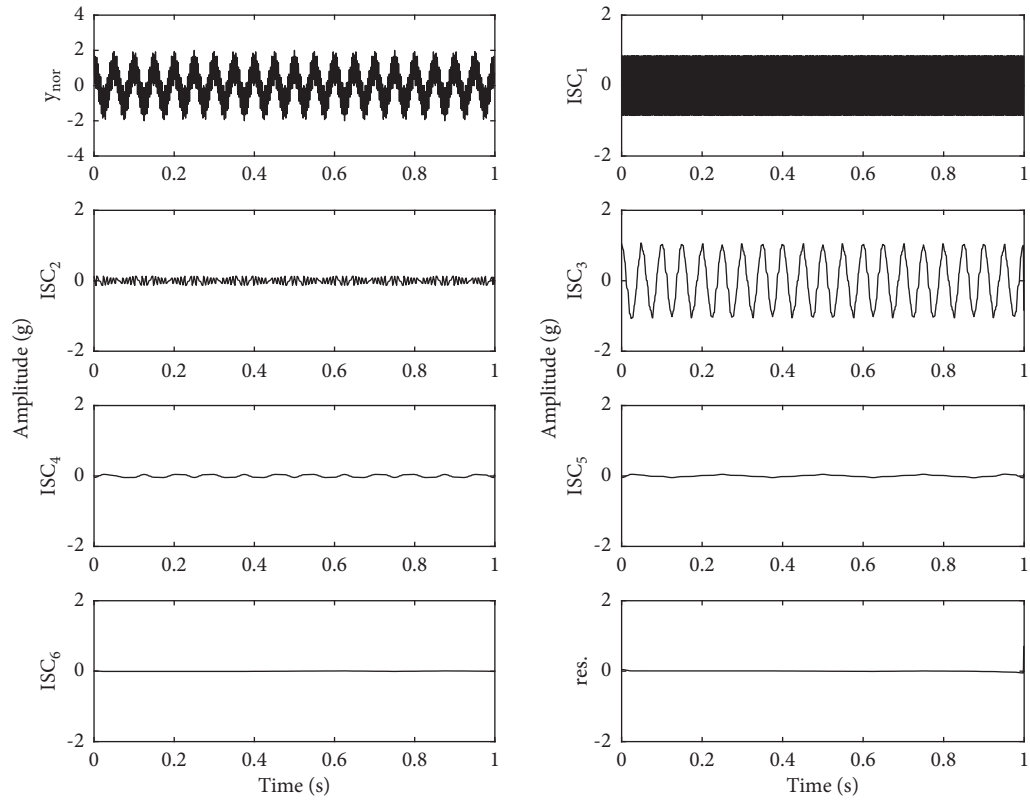


Figure 4: LCD decomposition results of y-channel under normal conditions.

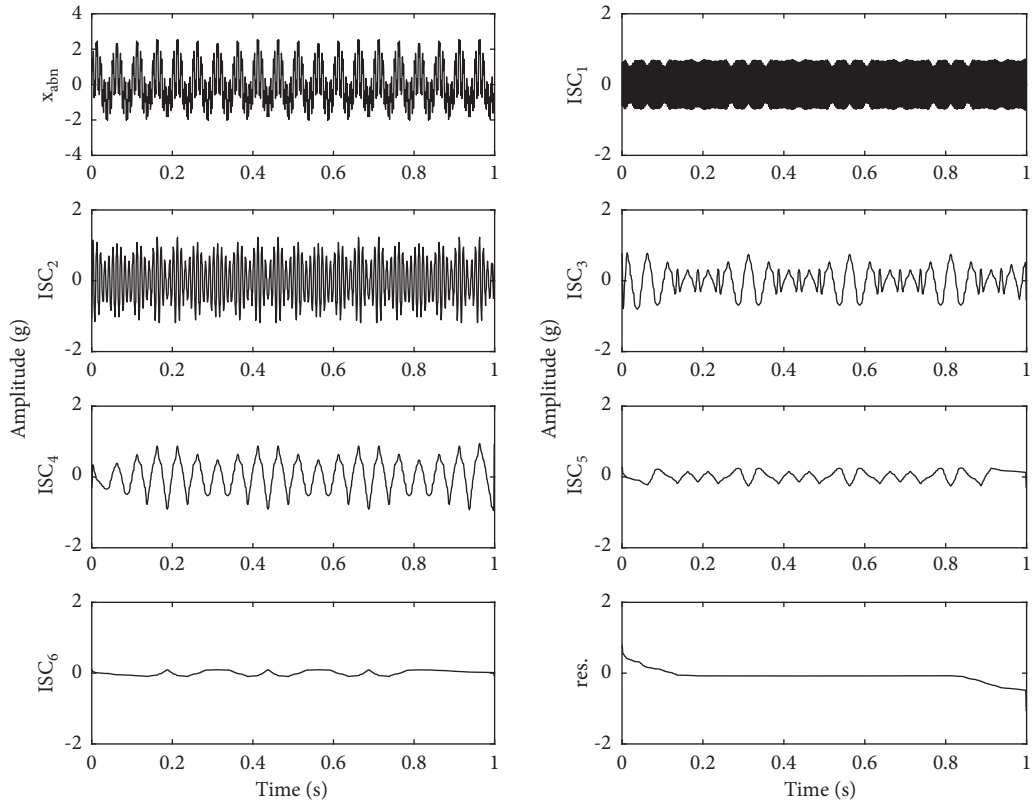


Figure 5: LCD decomposition results of x-channel under abnormal conditions.

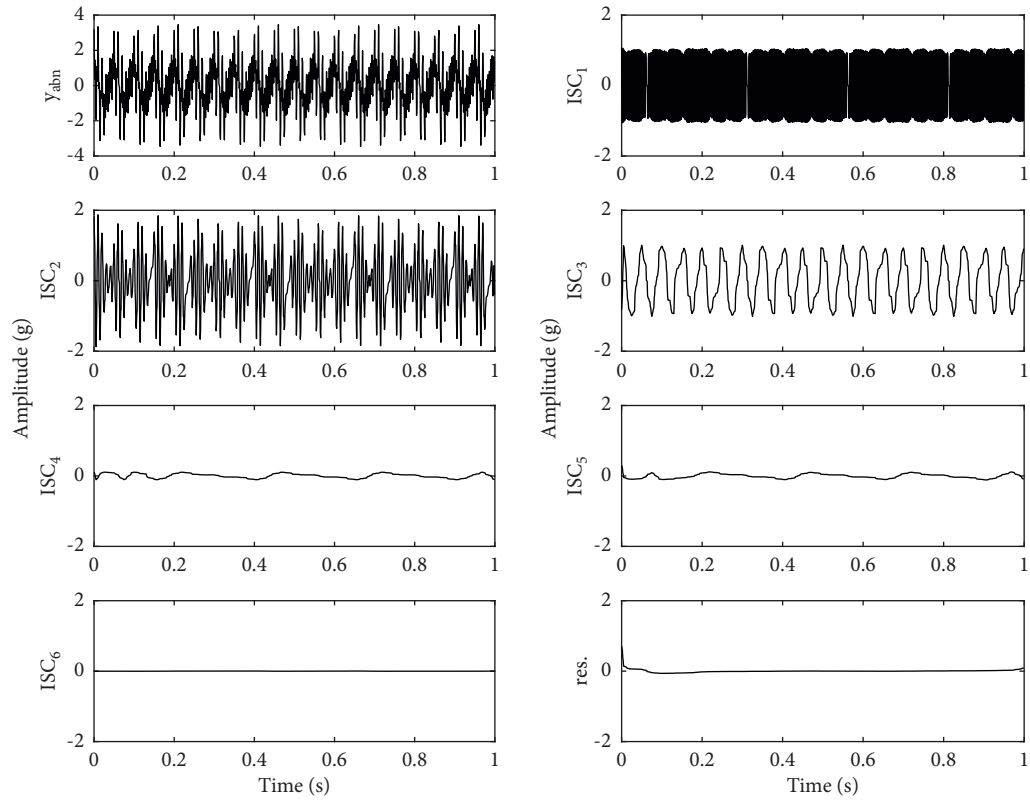


Figure 6: LCD decomposition results of y-channel under abnormal conditions.

Table 1: Correlation coefficients of ISCs of each order between two channel signals under two conditions.

ISC order		1	2	3	4	5	6
Correlation coefficient	x_{nor}	**0.7035**	0.1648	0.6998	0.0869	0.0032	0.0062
	y_{nor}	0.6983	0.1851	**0.7045**	−0.1065	0.11	−0.0203
	x_{abn}	0.5308	**0.6293**	0.4658	0.6119	0.0043	−0.0628
	y_{abn}	0.6385	**0.6576**	0.5345	0.0147	−0.0027	0.0125

The bold value is the maximum value of all ISCs (between the order 1 to 6).

fault information than others. Thus, we apply full vector fusion to the first-order ISC_{xnor}, third-order ISC_{ynor}, second-order ISC_{xabn}, and second-order ISC_{yabn}. However, the fault frequency information cannot be acquired just from time-domain waveforms. Finally, envelope and demodulation are applied to fusion signals, as shown in Figure 7. It is clear in Figure 7(a) that the rotation frequencies (20 Hz and 300 Hz) are consistent with the preset value under the normal operating condition. In Figure 7(b), the amplitudes of the fault frequency (100 Hz) and sidebands (20 Hz) show the existence of modulation under the abnormal operating condition.

From the simulation analysis, the vector LCD has a good analysis result when applied to fault signals with frequency or amplitude modulation, which can enhance the accuracy of fault diagnosis. The adoption of the cross-correlation coefficient avoids repeated analysis between multiple components and simplifies the analysis, for a unique and accurate conclusion.

4. Application

The validity and advantage of vector LCD in fault diagnosis of a rolling bearing were examined through experimental data from the Case Western Reserve University bearing data center. Figure 8 shows the layout of the test stand, which consisted of a 2-hp motor, torque transducer/encoder, dynamometer, and control electronics. The test bearings supported the motor shaft, and they had single-point faults from electro-discharge machining. Accelerometers collected vibration data and were attached to the support of the drive end bearing, fan end bearing, and motor supporting base. The technical parameters of fault diagnosis are listed in Table 2.

Normal baseline data, outer race fault data, and inner race fault data were adopted from the experimental signals. The ball fault data frequency is not matched with the usual value for the slip adjusting to lock onto a dominant frequency [41]. Therefore, ball fault data are not discussed here.

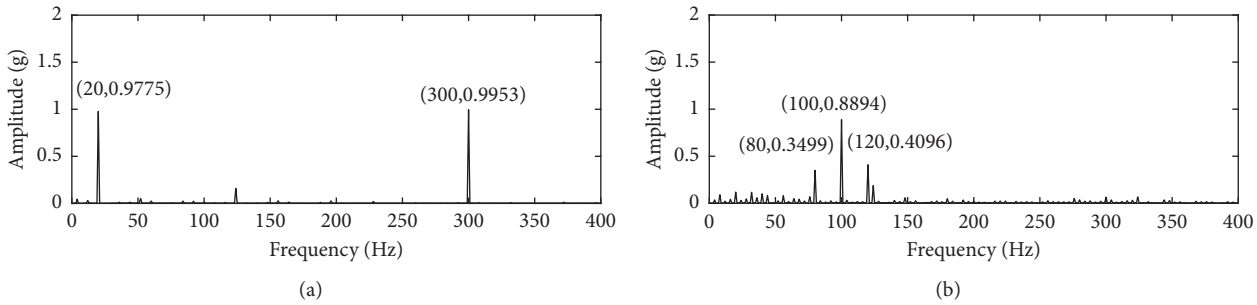


Figure 7: Vector LCD spectra of simulation signals: (a) normal condition; (b) abnormal condition.

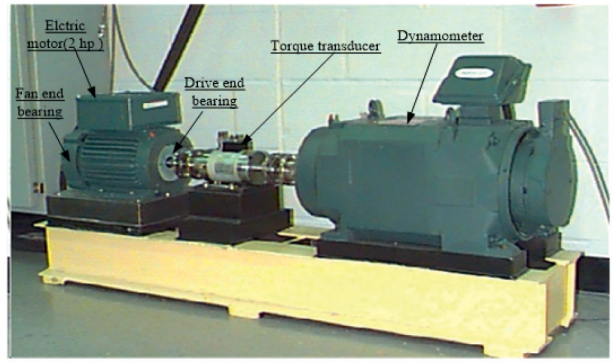


Figure 8: Case Western Reserve University bearing test stand.

Table 2: Technical parameters.

Fault location	Fault diameter (inches)	Fault depth (inches)	Motor load (hp)	Sampling frequency (Hz)
Normal	0	0	0	12000
Outer race	0.007	0.011	0	12000
Inner race	0.007	0.011	0	12000
Fault location	Motor speed (rpm)	Defect frequencies (Hz)	Sample points	Data file name
Normal	1797	—	8192	97.mat
Outer race	1797	107.3	8192	130.mat/144.mat
Inner race	1797	162.1	8192	105.mat

4.1. Normal Baseline Data Analysis. Because the normal baseline data only have one channel of drive end section, processing the single-sensor data is just a special case of a full vector spectrum. According to the flowchart of vector LCD, we first applied LCD to the drive-end single-channel signal to obtain 10 ISCs. Considering that the higher the order of ISC, the smaller the amplitude, only the first six orders of ISCs are displayed in Figure 9. We compare the values of ISC cross-correlation coefficients in Table 3, where the first is the optimal component. Finally, the envelope and demodulation are applied to the first ISC. From Figure 10, the rotation frequency f_r (29.95 Hz) and $2f_r$ (59.9 Hz) account for the main components without other fault frequencies, which is a normal operating condition.

4.2. Outer Race Fault Data Analysis. Outer race fault data were obtained by three homologous sensors in the same section, and channels of centered 6 : 00 clock and orthogonal 3 : 00 clocks were employed to vector LCD analysis. Similar to normal baseline data processing, ISC components were acquired after LCD. The ISC components of the x-and y-channels are shown in Figures 11 and 12, respectively. As seen in Table 4, the first ISC components of the x-and y-signals are optimal. From Figures 11 and 12, the time-domain waves of the first ISC components of the x- and y-signals have different amplitudes and frequencies. If only one direction of the x- or y-signal is analyzed, then different results will be obtained, which is contrary to fault diagnosis. Therefore, we apply full vector fusion to the two optimal ISC components to obtain a new combined signal. Finally, the envelope and demodulation are applied to the combined signal. The outer race fault frequency f_{OR} (107.3 Hz) can be easily distinguished from other spectra, and the amplitude of f_{OR} is the highest among all spectra (as seen in Figure 13).

4.3. Inner Race Fault Data Analysis. Inner race fault data were recorded in a single channel for the drive end, and the inner race fault signal was treated in the same way as normal

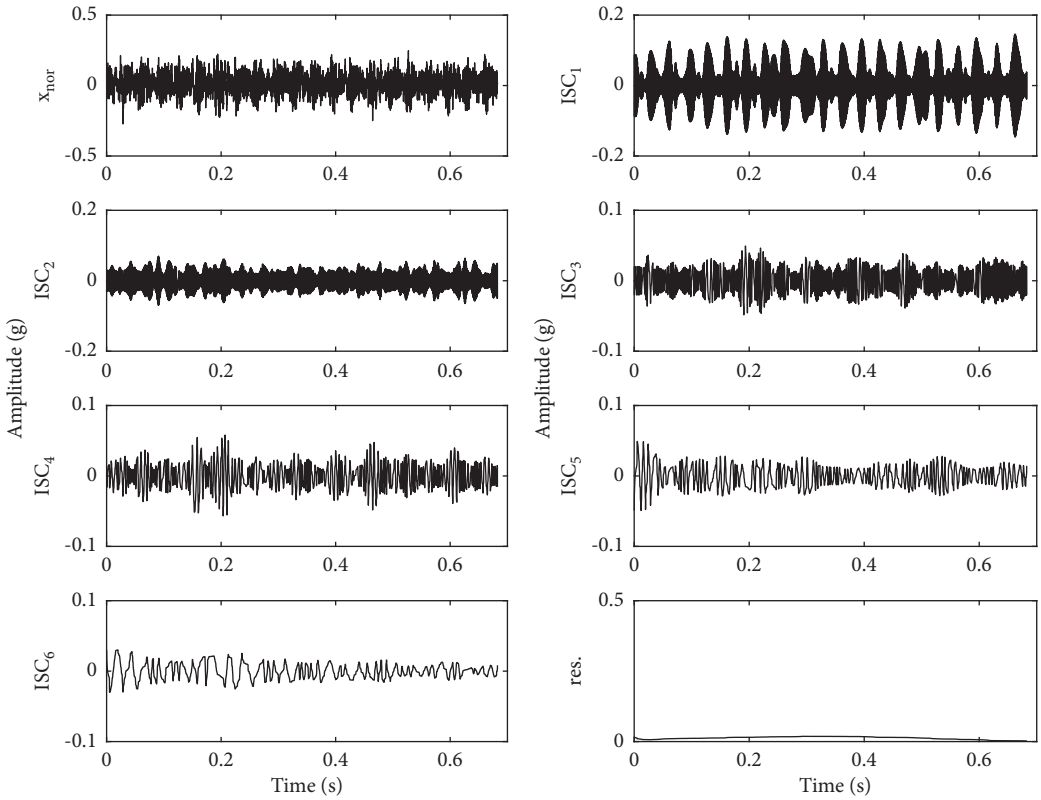


Figure 9: LCD decomposition results of x-channel under normal conditions.

Table 3: Correlation coefficients of ISCs of each order under normal conditions.

ISC order	1	2	3	4	5
Correlation coefficient	**0.7994**	0.5678	0.3904	0.4009	0.3037
ISC order	6	7	8	9	10
Correlation coefficient	0.1167	0.0773	0.0052	0.0077	0.0042

The bold value is the maximum value of all ISCs (between the order 1 to 10).

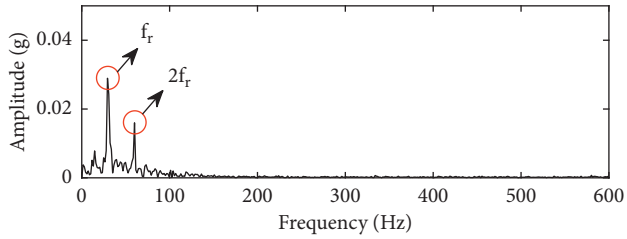


Figure 10: Vector LCD spectra of x-channel under normal conditions.

baseline data. The vector LCD spectra can be displayed after the optimal ISC is screened out, with decomposition results as shown in Figure 14. From Table 5, the optimal ISC (the first ISC component) can be easily obtained. From the vector LCD spectra in Figure 15, the inner race fault frequency f_{IR} (162.1 Hz) and rotation frequency harmonics f_r (29.95 Hz), $2f_r$ (59.9 Hz), and $4f_r$ (119.8 Hz) can be acquired from the spectra, which agrees with the inner race fault features.

This bearing experimental application shows that the proposed method can be applied to rolling bearing fault diagnosis. The screening of the optimal ISC can simplify fault diagnosis and clearly display the typical features. The full vector fusion between optimal ISCs of the x- and y-signals gives an accurate and unique conclusion for fault diagnosis. The vector LCD provides an easy way to extract fault features.

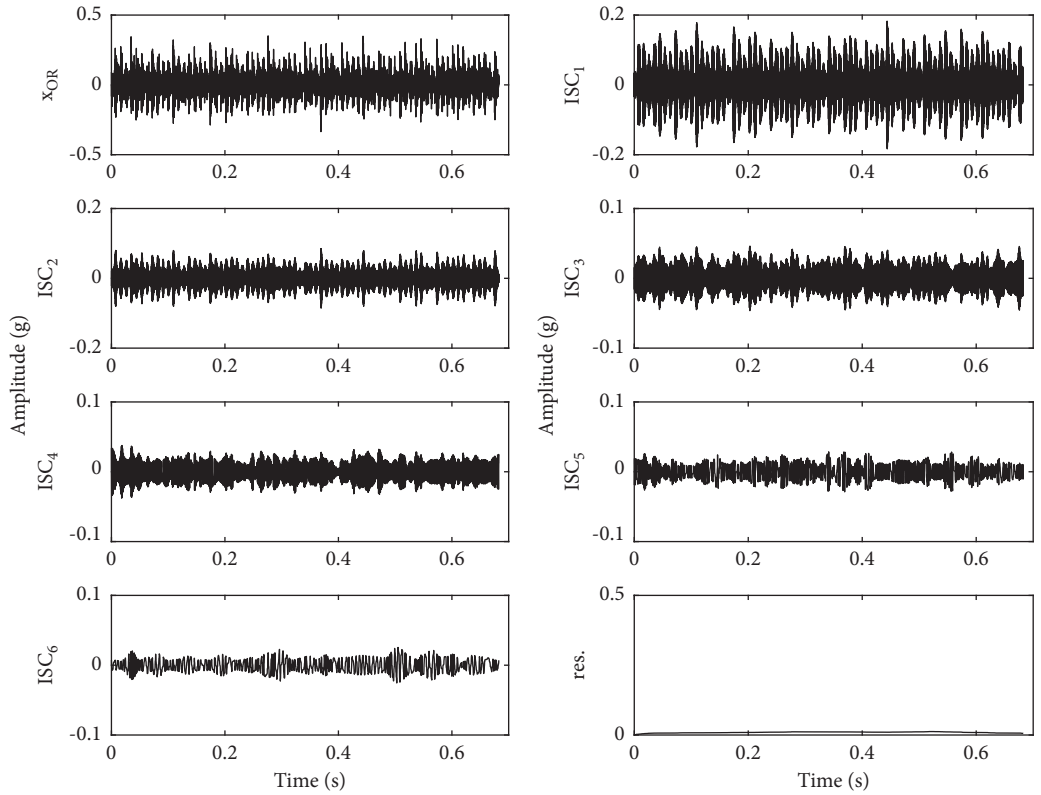


Figure 11: LCD decomposition results of x-channel under outer race fault condition.

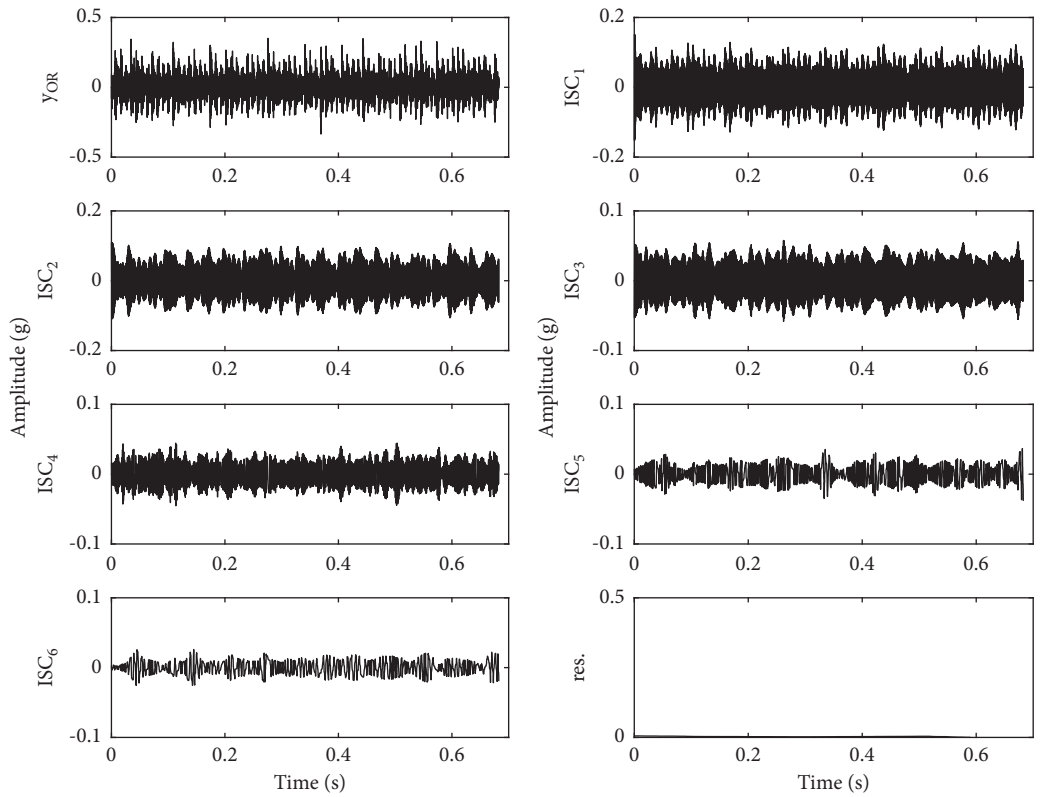


Figure 12: LCD decomposition results of y-channel under outer race fault condition.

Table 4: Correlation coefficient of ISCs of each order under outer race fault condition.

ISC order	Direction	1	2	3	4	5	6
Correlation	x	**0.8375**	0.5643	0.2833	0.2490	0.1737	0.1280
coefficient	y	**0.7529**	0.6911	0.3957	0.2071	0.1321	0.0858
ISC order	Direction	7	8	9	10	11	12
Correlation	x	0.0712	0.0243	0.0089	0.0001	0.002	—
coefficient	y	0.0426	−0.001	0.0054	0.0017	0.0004	0.0008

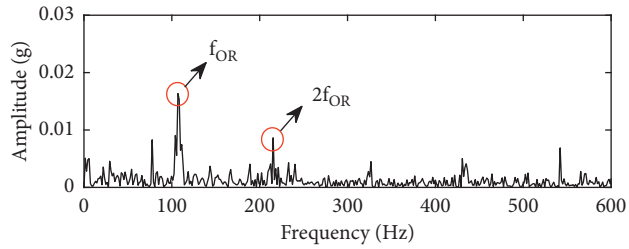


Figure 13: Vector LCD spectra of vector signal under outer race fault condition.

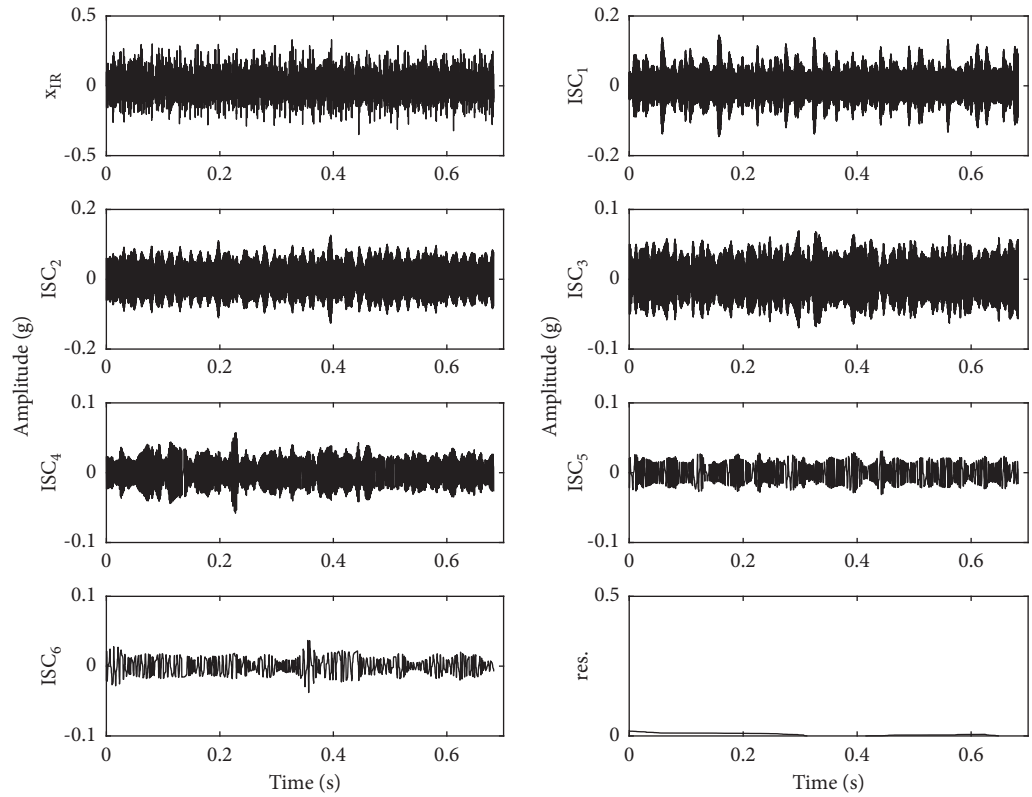


Figure 14: LCD decomposition results of x-channel under inner race fault condition.

Table 5: Correlation coefficient of ISCs of each order under inner race fault condition.

ISC order	1	2	3	4	5	6
Correlation coefficient	**0.7563**	0.7399	0.4803	0.3197	0.1993	0.1183
ISC order	7	8	9	10	11	12
Correlation coefficient	0.0300	0.0230	0.0005	0.0005	0.0013	0.0005

The bold value is the maximum value of all ISCs (between the order 1 to 12).

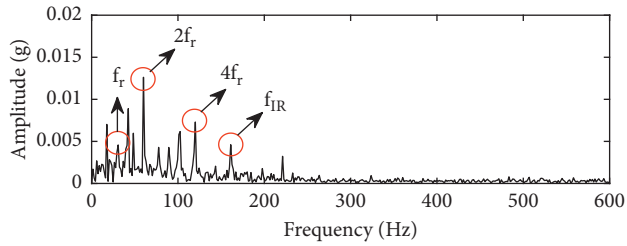


FIGURE 15: Vector LCD spectra of x-signal under inner race fault condition.

5. Conclusion

A multiple signal processing method of rolling bearing fault diagnosis was described in this paper. By combining the full vector spectrum with local characteristic-scale decomposition, the vector LCD can fully consider homologous signals and ISCs. This method is used to synchronously handle multiple signals. The cross-correlation coefficient was introduced to choose ISCs, which simplifies the analysis of ISCs. The vector ISC displays a more complete and precise fault frequency than a single ISC. The simulation analysis and rolling bearing experimental fault diagnosis verified the effectiveness of vector LCD. With the good compatibility of vector LCD, various types of faults and machines could be diagnosed in our future work. Moreover, the signals from different types of sensors will be combined to improve the accuracy of fault diagnosis.

Data Availability

The experimental data were taken from the Case Western Reserve University bearing data center.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the National Key Research and Development Program of China (2016YFF0203100) and the Henan Provincial Key Science and Technology Research Project of China (202102210075). The authors also thank the Case Western Reserve University bearing data center for the bearing fault test data.

References

[1] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[2] B. Zhao, X. Zhang, H. Li, and Z. Yang, "Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions," *Knowledge-Based Systems*, vol. 199, Article ID 105971, 2020.

[3] Y. Gao and D. Yu, "Intelligent fault diagnosis for rolling bearings based on graph shift regularization with directed graphs," *Advanced Engineering Informatics*, vol. 47, Article ID 101253, 2021.

[4] J. Cheng, Y. Yang, X. Li, and J. Cheng, "Adaptive periodic mode decomposition and its application in rolling bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 161, Article ID 107943, 2021.

[5] W. Zhu, G. Ni, Y. Cao, and H. Wang, "Research on a rolling bearing health monitoring algorithm oriented to industrial big data," *Measurement*, vol. 185, Article ID 110044, 2021.

[6] A. Glowacz, "Ventilation diagnosis of angle grinder using thermal imaging," *Sensors*, vol. 21, no. 8, p. 2853, 2021.

[7] P. Luo, N. Hu, L. Zhang, J. Shen, and Z. Cheng, "Improved phase space warping method for degradation tracking of rotating machinery under variable working conditions," *Mechanical Systems and Signal Processing*, vol. 157, Article ID 107696, 2021.

[8] X. Li, Y. Yang, H. Shao, X. Zhong, J. Cheng, and J. Cheng, "Symplectic weighted sparse support matrix machine for gear fault diagnosis," *Measurement*, vol. 168, Article ID 108392, 2021.

[9] J. Yang, D. Huang, D. Zhou, and H. Liu, "Optimal IMF selection and unknown fault feature extraction for rolling bearings with different defect modes," *Measurement*, vol. 157, Article ID 107660, 2020.

[10] A. Klausen, H. V. Khang, and K. G. Robbersmyr, "Multi-band identification for enhancing bearing fault detection in variable speed conditions," *Mechanical Systems and Signal Processing*, vol. 139, Article ID 106422, 2020.

[11] Y. Ma, J. Cheng, P. Wang, J. Wang, and Y. Yang, "Rotating machinery fault diagnosis based on multivariate multiscale fuzzy distribution entropy and Fisher score," *Measurement*, vol. 179, Article ID 109495, 2021.

[12] A. Glowacz, R. Tadeusiewicz, S. Legutko et al., "Fault diagnosis of angle grinders and electric impact drills using acoustic signals," *Applied Acoustics*, vol. 179, Article ID 108070, 2021.

[13] M. Hosseinpour-Zarnaq, M. Omid, and E. Biabani-Aghdam, "Fault Diagnosis of Tractor Auxiliary Gearbox Using Vibration Analysis and Random forest Classifier," *Information Processing in Agriculture*, 2021, In press.

[14] H. Pan, H. Xu, J. Zheng, J. Su, and J. Tong, "Multi-class fuzzy support matrix machine for classification in roller bearing fault diagnosis," *Advanced Engineering Informatics*, vol. 51, Article ID 101445, 2022.

[15] P. Tiwari and S. H. Upadhyay, "Novel self-adaptive vibration signal analysis: concealed component decomposition and its application in bearing fault diagnosis," *Journal of Sound and Vibration*, vol. 502, Article ID 116079, 2021.

[16] W. Ying, J. Zheng, H. Pan, and Q. Liu, "Permutation entropy-based improved uniform phase empirical mode decomposition for mechanical fault diagnosis," *Digital Signal Processing*, vol. 117, Article ID 103167, 2021.

[17] A. Patel and P. Shakya, "Spur gear crack modelling and analysis under variable speed conditions using variational mode decomposition," *Mechanism and Machine Theory*, vol. 164, Article ID 104357, 2021.

[18] X. Li, J. Ma, X. Wang, J. Wu, and Z. Li, "An improved local mean decomposition method based on improved composite interpolation envelope and its application in bearing fault feature extraction," *ISA Transactions*, vol. 97, pp. 365–383, 2020.

[19] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism and Machine Theory*, vol. 70, pp. 441–453, 2013.

[20] H. Liu, X. Wang, and C. Lu, "Rolling bearing fault diagnosis based on LCD-TEO and multifractal detrended fluctuation analysis," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 273–288, 2015.

[21] L. Wang and Z. Liu, "An improved local characteristic-scale decomposition to restrict end effects, mode mixing and its application to extract incipient bearing fault signal," *Mechanical Systems and Signal Processing*, vol. 156, Article ID 107657, 2021.

[22] C. Yang, J. Ma, X. Wang, X. Li, Z. Li, and T. Luo, "A Novel Based-Performance Degradation Indicator RUL Prediction Model and its Application in Rolling Bearing," *ISA Transactions*, 2021.

[23] P. Zhang, T. Li, G. Wang et al., "Multi-source information fusion based on rough set theory: a review," *Information Fusion*, vol. 68, pp. 85–117, 2021.

[24] S. Liu, "A modified low-speed balancing method for flexible rotors based on holospectrum," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 348–364, 2007.

[25] S. Liu and L. Qu, "A new field balancing method of rotor systems based on holospectrum and genetic algorithm," *Applied Soft Computing*, vol. 8, no. 1, pp. 446–455, 2008.

[26] L. Qu, Y. Liao, J. Lin, and M. Zhao, "Investigation on the subsynchronous pseudo-vibration of rotating machinery," *Journal of Sound and Vibration*, vol. 423, pp. 340–354, 2018.

[27] L. Qu, X. Liu, G. Peyronne, and Y. Chen, "The holospectrum: a new method for rotor surveillance and diagnosis," *Mechanical Systems and Signal Processing*, vol. 3, no. 3, pp. 255–267, 1989.

[28] A. Muszynska and P. Goldman, "Application of full spectrum to rotating machinery diagnostics," *Orbit*, vol. 1, no. 20, pp. 17–21, 1999.

[29] T. H. Patel and A. K. Darpe, "Application of full spectrum analysis for rotor fault diagnosis,"Springer, New York, NY, USA.

[30] T. Patel and A. Darpe, "Use of full spectrum cascade for rotor rub identification," *Advances in Vibration Engineering*, vol. 8, pp. 139–151, 2009.

[31] X. Zhao, T. H. Patel, and M. J. Zuo, "Multivariate EMD and full spectrum based condition monitoring for rotating machinery," *Mechanical Systems and Signal Processing*, vol. 27, pp. 712–728, 2012.

[32] J. Han and L. Shi, *Full Vector Spectrum Technology and its Enginnering Application*, China Machine Press, Beijing,China, 2008.

[33] L. Chen, J. Han, W. Lei, Z. Guan, and Y. Gao, "Prediction model of vibration feature for equipment maintenance based on full vector spectrum," *Shock and Vibration*, vol. 2017, Article ID 6103947, 8 pages, 2017.

[34] X. Gong, L. Ding, and W. Du, "Application of the full vector spectrum to local rub-impact fault diagnosis in rotor systems," *Journal of Residuals Science & Technology*, vol. 13, no. 8, 2016.

[35] H. Li, X. M. Dong, W. S. Hao, A. G. Liu, X. D. Yin, and A. J. Wang, "Applying full vector spectrum for electric hoist gearbox fault diagnosis," *Applied Mechanics and Materials*, vol. 365-366, pp. 725–728, 2013.

[36] L. Chen, J. Han, W. Lei, Y. Cui, and Z. Guan, "Full-vector signal acquisition and information fusion for the fault prediction," *International Journal of Rotating Machinery*, vol. 2016, Article ID 5980802, 7 pages, 2016.

[37] X. Gong, L. Ding, W. Du, and H. Wang, "Gear fault diagnosis using dual channel data fusion and EEMD method," *Procedia Engineering*, vol. 174, pp. 918–926, 2017.

[38] H. Yu, H. Li, Y. Li, and Y. Li, "A novel improved full vector spectrum algorithm and its application in multi-sensor data fusion for hydraulic pumps," *Measurement*, vol. 133, pp. 145–161, 2019.

[39] J. Cheng, J. Zheng, and Y. Yang, "A new non-stationary signal analysis approach-the local characteristic-scale decomposition method," *Journal of Vibration Engineering*, vol. 25, no. 2, pp. 215–220, 2012.

[40] J. Han and L. Shi, "Study on space precession and vibration characteristic of high-speed rotary shaft," *Journal of Vibration Engineering*, vol. 17, no. 3, pp. 69–72, 2004.

[41] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study," *Mechanical Systems and Signal Processing*, vol. 64-65, pp. 100–131, 2015.

Hindawi
Shock and Vibration
Volume 2021, Article ID 8456991, 13 pages
https://doi.org/10.1155/2021/8456991

Research Article

Rolling Bearing Fault Diagnosis Method Based on Multisynchrosqueezing S Transform and Faster Dictionary Learning

Guodong Sun,[1] **Ye Hu**,[1] **Bo Wu**,[2] **and Hongyu Zhou**[1]

[1]*School of Mechanical Engineering, Hubei University of Technology, Wuhan, Hubei, China*
[2]*Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China*

Correspondence should be addressed to Guodong Sun; sgdeagle@163.com and Hongyu Zhou; engzhy@126.com

Received 9 April 2021; Revised 28 April 2021; Accepted 4 May 2021; Published 19 May 2021

Academic Editor: Haiyang Pan

Addressing the problem that it is difficult to extract the features of vibration signal and diagnose the fault of rolling bearing, we propose a novel diagnosis method combining multisynchrosqueezing S transform and faster dictionary learning (MSSST-FDL). Firstly, MSSST is adopted to transform vibration signals into high-resolution time-frequency images. Then, the local binary pattern (LBP) operator is introduced to extract the low-dimensional texture features of time-frequency images, which improves the speed of fault recognition. Finally, nonnegative matrix factorization (NMF) with only one hyperparameter and nonnegative linear equation are used to solve the dictionary learning and feature coding, respectively. The feature coding is input into the classifier for training and recognition. Experiments show that our method performs well on the rolling bearing dataset of Case Western Reserve University (CWRU) and the Society for Machinery Failure Prevention Technology (MFPT). Further, the proposed method is applied to the loudspeaker pure-tone detection dataset, and the loudspeaker anomaly diagnosis is achieved. The diagnosis results verify that our method can meet the needs of practical engineering.

1. Introduction

Rotating machinery plays an increasing role in electric manufacturing, transportation, power, and other industries. As the core component of rotating machinery, rolling bearings directly affect the operation of the entire equipment. However, the damage of rolling bearings is inevitable, which may cause serious economic losses and even safety accidents [1–3]. Thus, it is of great significance to detect bearing faults in time and take appropriate maintenance measures according to the diagnosis results [4, 5].

Vibration signal, which contains rich information and can reflect the running state of rotating machinery well, has become the most commonly used signal source in fault diagnosis of rotating machinery. The vibration signals of rolling bearings have obvious nonlinear and nonstationary characteristics so that it is difficult to recognize their faults directly. By using the appropriate time-frequency analysis method to process the vibration signal, we can obtain the variation law of its spectrum with time. The idea of time-frequency analysis is originated from the Gabor transform [6]. Thereafter, short-time Fourier transform (STFT) [7], continuous wavelet transform (CWT) [8], and S transform (ST) [9] appear successively. Although these methods are easy to implement, the limitations of Heisenberg's uncertainty principle [10, 11] prevent them from improving both time and frequency resolutions. To obtain time-frequency images with better energy concentration of vibration signals, Daubechies et al. [12] proposed synchrosqueezed wavelet transform (SSWT). In essence, it is a time-frequency analysis method of energy rearrangement. Namely, based on CWT, the spectrum energy is redistributed and concentrated on the instantaneous frequency [13]. Based on this idea, Huang et al. [14] proposed the synchrosqueezing S transform (SSST) and Yu et al. [15] proposed the multisynchrosqueezing transform (MSST). SSST is an energy

rearrangement algorithm based on ST, and ST has better performance than STFT. MSST is a multiple energy rearrangement algorithm based on STFT. In other words, multiple iterations of synchrosqueezing transform are performed. The more iterations, the better the time-frequency energy concentration. Combining the advantages of both SSST and MSST, multisynchrosqueezing S transform (MSSST) [16] is proposed as a new time-frequency analysis method for rolling bearing vibration signals.

Although the time-frequency image of the vibration signal of the rolling bearing is more intuitive than the raw signal, the dimension of the time-frequency image is too large to be directly inputted into the classifier for training and recognition. Therefore, it is necessary to extract the low-dimensional and valuable features of the time-frequency image. Over recent years, many scholars have studied this problem. In reference [17–19], a convolutional neural network (CNN) was used to extract the features of CWT, STFT, and HHT time-frequency images, respectively. Li et al. [20] proposed a convolution sparse self-learning (CSSL) to extract the defective bearing morphological feature. Wang et al. [21] designed a one-dimensional vision ConvNet (VCN) to extract multiscale sensitive features of bearings in complex operating environments. These deep learning-based methods are relatively new, but they have three major problems as follows: (1) they require a large number of training samples which are expensive and difficult to obtain; (2) it is challenging and time-consuming to train an excellent CNN from scratch with only a small number of samples; (3) plenty of hyperparameters have to be predetermined for CNN, such as activation functions, epoch number, learning rate, momentum, kernel sizes, and numbers of layers.

Therefore, the traditional machine learning method based on feature engineering is still worth further study. In reference [22], two-dimensional nonnegative matrix factorization (2DNMF) is developed to extract more informative features from the ST time-frequency images for accurate fault recognition. In reference [23], the features of ST time-frequency images were first extracted by nonnegative matrix factorization (NMF) [24, 25], and then nondominated sorting genetic algorithms (NSGA-II) were proposed to make secondary the selection of features. Yu et al. [26] proposed a rolling bearing fault diagnosis method based on Hilbert–Huang transform (HHT) and supervised sparse coding (SSC). This method adopted SSC to obtain a sparse representation of the marginal spectrum generated by HHT and used the support vector machine (SVM) to achieve fault recognition. Although these methods can obtain high recognition precision on their own datasets, they may not be suitable for simultaneous diagnosis of fault location and damage degree.

In this regard, Sun et al. [27] proposed a method combining MSST and sparse feature coding based on dictionary learning (SFC-DL). Li et al. [28] designed a symplectic weighted sparse SMM (SWSSMM) model with the sparsity constraint and low-rank constraint, and Li et al. [29] developed the discriminative manifold random vector functional link neural network (DMRVFLNN) model. The above fault diagnosis methods cannot overcome the problem of how to select the optimal parameters for their models. More importantly, the method proposed by Sun et al. would consume many computer resources and running time because all the elements of time-frequency images are taken for nonnegative matrix factorization with sparseness constraints (NMFSCs) [30] and sparse coding.

To overcome the abovementioned problems, we propose to first extract the texture features of time-frequency images by the local binary pattern (LBP) [31, 32] operator and then use these texture features for dictionary learning. The LBP algorithm with linear order time complexity and space complexity is simple to calculate. Meanwhile, the texture features of time-frequency images have rich information and low dimension, which can greatly improve the performance of dictionary learning and feature coding. In addition, we use NMF instead of NMFSC for dictionary learning and reduce the hyperparameter to only one, which makes the optimal fault diagnosis model easier to be obtained. We name these optimizations faster dictionary learning (FDL).

In summary, the main contributions of this article are as follows:

(1) A new method for rolling bearing fault diagnosis is proposed by combining MSSST and FDL, which is named MSSST-FDL

(2) MSSST is adopted to obtain high-resolution time-frequency images of vibration signals, which can promote the accuracy of fault diagnosis

(3) To improve the speed of feature extraction from time-frequency images, we design the FDL algorithm by introducing LBP and NMF

(4) Experiment results on two rolling bearing datasets and one loudspeaker dataset show that the proposed method performs well and has the potential to be applied to different types of equipment

The remaining of this article is mainly described as follows. Section 2 introduces the theory of MSSST and FDL. Section 3 presents the experimental comparisons. An extended application of the proposed method is shown in Section 4. Finally, the conclusions are presented in Section 5.

2. The Proposed Method

In this paper, a new method based on time-frequency analysis and improved dictionary learning is proposed for fault diagnosis of rolling bearing. The main procedures are described as follows:

Step 1. Raw vibration signals of rolling bearing are collected under different working conditions, and their states are noted.

Step 2. The raw signals are segmented and ensured that each sample signal contains one complete period at least.

Step 3. MSSST is performed for sample signals to obtain the time-frequency images with high resolution.

Step 4. FDL is used to process the time-frequency images, and we can get the effective feature coding of each sample quickly.

Step 5. The feature coding set is divided into the training, validation, and testing sets. Then, the diagnosis model is obtained by cross-validation on the training set and the validation set.

Step 6. Testing set is input into the diagnosis model for fault diagnosis.

Figure 1 shows the overall framework of the proposed method, and the following subsections provide details of MSSST and FDL.

2.1. Multisynchrosqueezing S Transform. In practice, raw vibration signals are always nonlinear and nonstationary. It is necessary to process these complex signals. Fortunately, time-frequency analysis is an effective approach to reveal the frequency components and time-variation features of vibration signals. In various time-frequency analysis methods, MSSST combines the advantages of SSST and MSST to generate better energy concentration and suppress the cross-terms over the time-frequency plane.

Let the vibration signal be $s(t)$, then the expression of its ST is as follows:

$$\mathrm{ST}(t,\omega)=\frac{|\omega|}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}s(u)e^{-(u-t)^2\omega^2/2}e^{-i2\pi\omega u}\mathrm{d}u, \tag{1}$$

where ω is the frequency, t represents the time axis displacement parameter, i is an imaginary unit, and $e^{-(u-t)^2\omega^2/2}$ is the window function.

It can be seen from equation (1) that the window function of ST is flexible. Its window width can change according to the change of the frequency ω. The window width is wider in the low-frequency part and narrower in the high-frequency part. This not only improves the shortcomings of STFT but also inherits the multiresolution characteristics of CWT. Therefore, combining the advantages of ST and the idea of iterative compression of MSST, the MSSST can be expressed as

$$\begin{aligned}
\mathrm{MSSST}^{[1]}(t,\omega)&=\int_{-\infty}^{+\infty}\mathrm{ST}(t,\eta)\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta,\\
\mathrm{MSSST}^{[2]}(t,\omega)&=\int_{-\infty}^{+\infty}\mathrm{MSSST}^{[1]}(t,\eta)\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta,\\
&\vdots\\
\mathrm{MSSST}^{[N]}(t,\omega)&=\int_{-\infty}^{+\infty}\mathrm{MSSST}^{[N-1]}(t,\eta)\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta,
\end{aligned} \tag{2}$$

where $\mathrm{ST}(t,\eta)$ is the ST of signal $s(t)$, $\mathrm{MSSST}^{[N]}(t,\omega)$ is the MSSST after N iterations, $\widehat{\omega}_S(t,\eta)$ is the instantaneous frequency (IF) estimate based on ST, and its expression is defined as follows [14]:

$$\widehat{\omega}_s(t,\omega)=\omega+[i2\pi\mathrm{ST}(t,\omega)]^{-1}\frac{\partial\mathrm{ST}(t,\omega)}{\partial t}. \tag{3}$$

We substitute $\mathrm{MSSST}^{[1]}(t,\eta)$ into $\mathrm{MSSST}^{[2]}(t,\eta)$, and then the $\mathrm{MSSST}^{[2]}(t,\eta)$ can be expressed as follows:

$$\begin{aligned}
\mathrm{MSSST}^{[2]}(t,\omega)&=\int_{-\infty}^{+\infty}\mathrm{MSSST}^{[1]}(t,\eta)\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta\\
&=\int_{-\infty}^{+\infty}\left[\int_{-\infty}^{+\infty}\mathrm{ST}(t,\xi)\delta\left(\omega-\widehat{\omega}_s(t,\xi)\right)\mathrm{d}\xi\right]\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta\\
&=\int_{-\infty}^{+\infty}\mathrm{ST}(t,\xi)\int_{-\infty}^{+\infty}\delta\left(\omega-\widehat{\omega}_s(t,\xi)\right)\delta\left(\omega-\widehat{\omega}_s(t,\eta)\right)\mathrm{d}\eta\mathrm{d}\xi\\
&=\int_{-\infty}^{+\infty}\mathrm{ST}(t,\xi)\delta\left(\omega-\widehat{\omega}_s\left(t,\widehat{\omega}_s(t,\xi)\right)\right)\mathrm{d}\xi\\
&=\int_{-\infty}^{+\infty}\mathrm{ST}(t,\eta)\delta\left(\omega-\widehat{\omega}_s\left(t,\widehat{\omega}_s(t,\eta)\right)\right)\mathrm{d}\eta.
\end{aligned} \tag{4}$$

The rest can be done in the same manner, and then the $\mathrm{MSSST}^{[N]}(t,\omega)$ can be expressed as follows:

$$\mathrm{MSSST}^{[N]}(t,\omega)=\int_{-\infty}^{+\infty}\mathrm{ST}(t,\eta)\delta\left(\omega-\widehat{\omega}_s^{[N]}(t,\eta)\right)\mathrm{d}\eta \tag{5}$$

where $\widehat{\omega}_s^{[N]}(t,\omega)=\widehat{\omega}_s\left(t,\widehat{\omega}_s^{[N-1]}(t,\omega)\right)$.

The MSSST uses the S transform of the SSST to obtain time-frequency coefficients with better energy concentration. At the same time, combined with the idea of multiple iterations in the MSST, the time-frequency results can be further sharpened. After one iteration, MSSST will construct a new IF estimate to reassign the blurry ST energy. Therefore, after several iterations, the IF estimation in the MSSST will get closer and closer to the real IF of the vibration signal. Namely, the energy of the time-frequency distribution can be gradually concentrated.

2.2. Faster Dictionary Learning. The dimension of the time-frequency image generated by MSSST is 1600×800, which is too high. If all the elements of the time-frequency image are directly input into the classifier for training and recognition, serious overfitting will occur. Therefore, we propose the FDL algorithm to extract the effective features of time-frequency images. The procedures of FDL are shown in Figure 2. First of all, the texture feature vectors of MSSST time-frequency images are extracted by the LBP operator, and each feature vector is taken as one column of V. After that, one-tenth of the samples are uniformly and randomly selected to

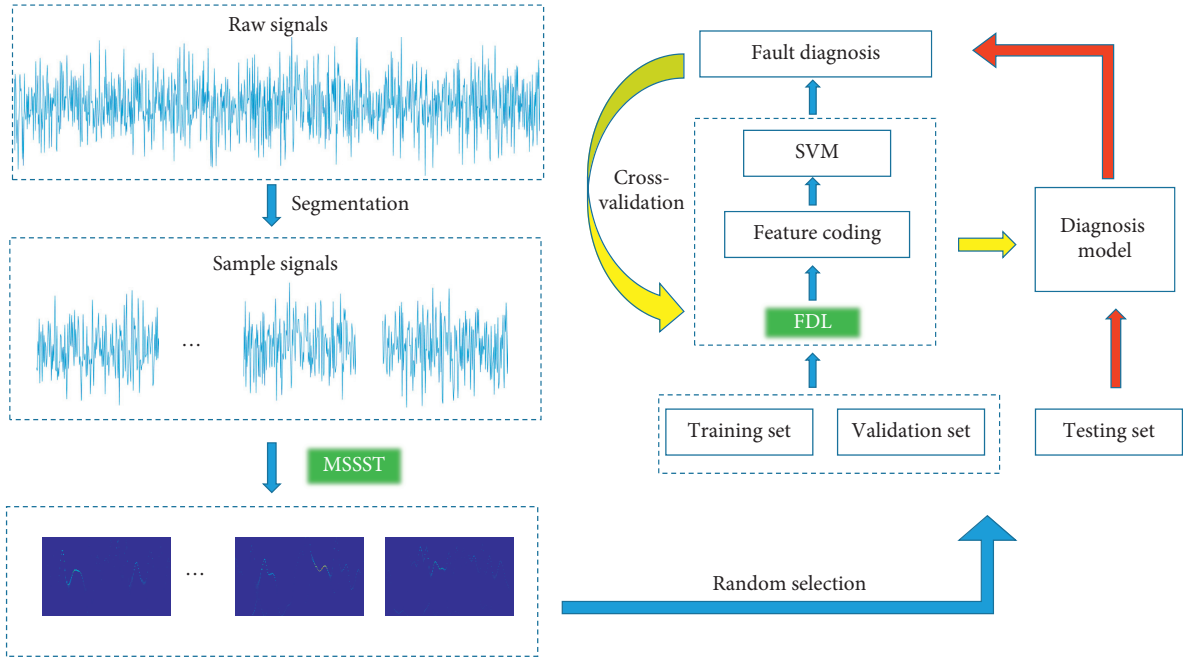


Figure 1: The overall framework of the presented method.

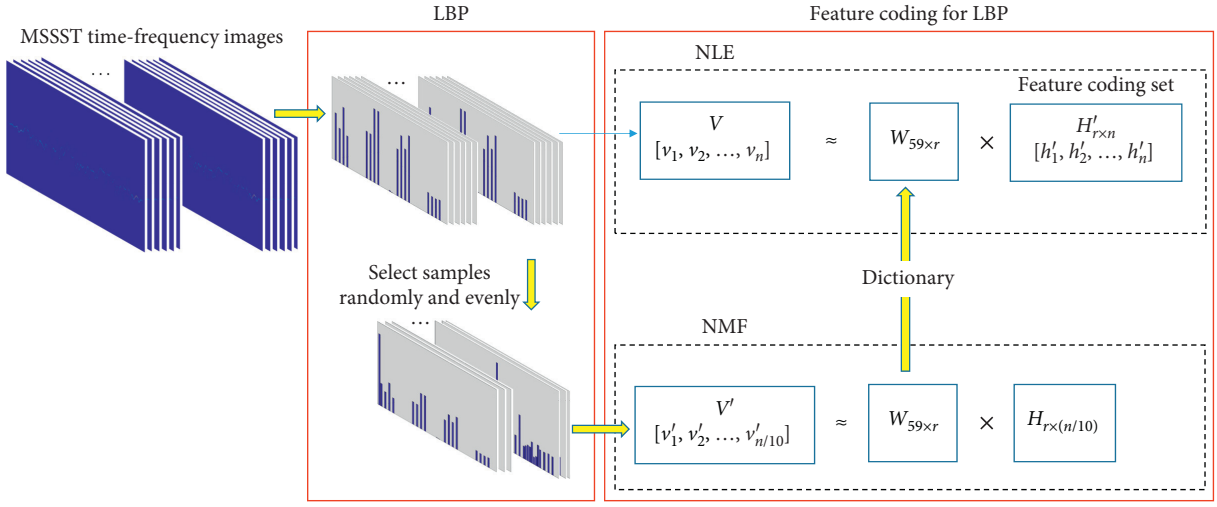


Figure 2: The procedures of FDL.

compose matrix V', which is decomposed by NMF to generate the dictionary W. At last, in combination with W, the feature coding set H' of all samples can be solved by NLE. More details are described in the following subsections.

2.2.1. Local Binary Pattern Operator. As shown in Figure 3, the original LBP operator is defined as a 3×3 square window, and the center point of the window is taken as the threshold value to compare it with 8 adjacent pixels. If the surrounding pixel value is greater than the value of the center point, this pixel is marked as 1; otherwise, it is 0. In the end, we can obtain 256 types of binary patterns.

To further reduce the number of binary patterns and improve the statistics, a uniform pattern [32] is designed, which recorded the jump times of binary numbers 0 and 1 of the LBP operator. If the number of jumps is less than or equal to 2, it is called uniform pattern, and all except the uniform pattern are classified into one class. As a result, the number of patterns has been reduced from 256 to 59. Namely, the dimension of the feature vector of the time-

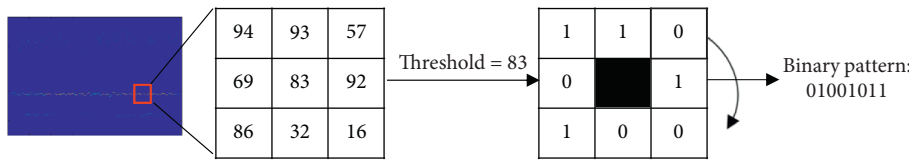


Figure 3: The original LBP operator.

frequency image is 59. The low-dimensional texture features of time-frequency images extracted by LBP can speed up dictionary learning and feature coding.

2.2.2. Feature Coding for LBP. By extracting the LBP features of the time-frequency images, the feature size can be reduced to 59. However, there are still some redundant features that affect the recognition accuracy. The optimal dimension can be further determined by feature coding so as to achieve the best recognition.

Before feature coding for LBP, it is necessary to find its basis dictionary. The method is NMF which is defined as follows:

$$V_{n\times m} \approx W_{n\times r} H_{r\times m}, \tag{6}$$

where $V_{n\times m}$, $W_{n\times r}$, and $H_{r\times m}$ are nonnegative matrices, V is the small-batch set of the LBP features, W is the basis dictionary, and H is the feature coding set. Each column of V is the LBP feature of one time-frequency image, and each column of H is corresponding to the feature coding of each LBP feature. Then, r can control the dimension of the feature coding.

NMF has only one hyperparameter, which can be used for dictionary learning to obtain the optimal model more efficiently. All the remaining samples can be represented by different linear combinations of column vectors of the basis dictionary. The linear combinations are called feature coding. We can solve the feature coding by nonnegative linear equation (NLE), which is expressed as follows:

$$\begin{aligned} &\min \left\| v_i - W h_i' \right\|^2 \\ &s.t. \quad h_{ij}' \geq 0, \end{aligned} \tag{7}$$

where v_i represents the ith sample, h_i' represents the corresponding feature coding of v_i, and h_{ij}' represents the jth element in the feature coding h_i'.

By equation (7), we can obtain the vector set $[h_1', h_2', \ldots h_n']$ in Figure 2. The set $[h_1', h_2', \ldots h_n']$ is the feature coding set of all LBP features, which is denoted as H' and input into the classifier for training and fault recognition.

3. Experimental Study

To verify the performance of the proposed method, the Case Western Reserve University (CWRU) [33] and the Machinery Failure Prevention Technology (MFPT) [34] rolling bearing vibration signal datasets are selected for experiments. All experiments are carried out with Windows 10 (64 bit), CPU Intel Xeon E5-2640@2.40 GHz, memory 64 GB, and MATLAB 2017b.

3.1. Description of Dataset. The CWRU test platform is mainly composed of a motor, torque sensor, dynamometer, and electronic control equipment. The designation of the tested bearing is 6205-2RS JEM SKF, which is located at the drive end. The EDM technology is adopted to set the faults at three different positions (inner race, outer race, and ball) of the bearing. Each fault location has three different degrees of damage (fault diameter of 0.007, 0.014, and 0.021 inches, respectively). Therefore, after adding the normal state, there are 10 kinds of health conditions of rolling bearings. Vibration signals under each health condition are collected at four different motor loads (0, 1, 2, and 3 hp) and speeds (1797, 1772, 1750, and 1730 r/min) with a sampling frequency of 12 kHz. To retain data features as much as possible and increase the number of samples, the 40 raw samples are obtained under 10 health conditions and 4 working conditions are further divided. As shown in Figure 4, each raw sample is continuously divided into 150 samples, and each sample has 800 sampling points. Finally, the total number of samples is 6000, including 600 samples for each health condition and 150 samples for each working condition, which are divided into the training set, validation set, and testing set according to the ratio of 6 : 2 : 2. To improve the robustness of the diagnosis method and meet the needs of practical engineering, the influence of working conditions on fault recognition is not considered. At the same time, to avoid contingency, 150 samples under each working condition are randomly divided into the training set (90), validation set (30), and testing set (30). The details of the CWRU dataset are shown in Table 1.

MFPT dataset includes inner race, outer race, and normal health conditions with a motor speed of 25 Hz. The sampling frequencies of the inner race and outer race data are both 97656 Hz. The inner race data are collected at 7 different motor loads (0, 50, 100, 150, 200, 250, and 300 lbs). The outer race data are collected at 7 different motor loads (25, 50, 100, 150, 200, 250, and 300 lbs). The sampling frequency of normal data is 97656 Hz, and the motor load is 270 lbs. MFPT data are segmented in the same way as CWRU. However, each segmented sample has 4000 sampling points. To reduce the redundant information of each sample and facilitate the subsequent operation, it is sampled down every 5 points to get the final sample length of 800 sampling points. In the end, a total of 2,100 samples are obtained, with 700 samples for each health condition. In the case of inner and outer race faults, 100 samples are corresponding to each motor load. The ratio of training set validation set and testing set is 6 : 2 : 2. They are selected at random in the same way as the CWRU. The details of the MFPT dataset are shown in Table 2.

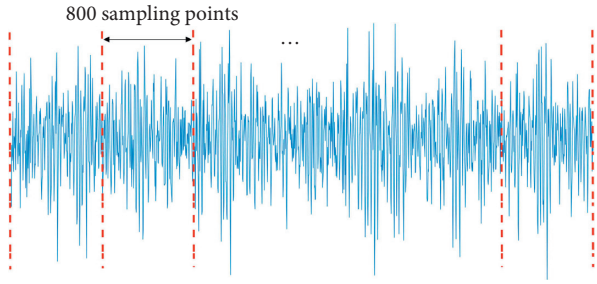


Figure 4: Segmentation of raw sample data of CWRU.

Table 1: Description of CWRU dataset.

Fault type	Fault diameter (inches)	Motor load (hp)	Motor speed (r/min)	Dataset			Fault label
				Training set	Validation set	Testing set	
Normal	—	0/1/2/3	1797/1772/1750/1730	360	120	120	1
Inner race	0.007	0/1/2/3	1797/1772/1750/1730	360	120	120	2
	0.014	0/1/2/3	1797/1772/1750/1730	360	120	120	3
	0.021	0/1/2/3	1797/1772/1750/1730	360	120	120	4
Ball	0.007	0/1/2/3	1797/1772/1750/1730	360	120	120	5
	0.014	0/1/2/3	1797/1772/1750/1730	360	120	120	6
	0.021	0/1/2/3	1797/1772/1750/1730	360	120	120	7
Outer race	0.007	0/1/2/3	1797/1772/1750/1730	360	120	120	8
	0.014	0/1/2/3	1797/1772/1750/1730	360	120	120	9
	0.021	0/1/2/3	1797/1772/1750/1730	360	120	120	10

Table 2: Description of MFPT dataset.

Fault type	Motor load (hp)	Motor speed (Hz)	Dataset			Fault label
			Training set	Validation set	Testing set	
Normal	270	25	420	140	140	1
Inner race	0/50/100/150/200/250/300	25	420	140	140	2
Outer race	25/50/100/150/200/250/300	25	420	140	140	3

3.2. Time-Frequency Analysis of Vibration Signals. To prove the superiority of MSSST, it is compared with several popular time-frequency analysis methods. Each type of faults in two datasets provides one sample randomly for our study. Table 3 shows the Rényi entropy (RE) [35, 36] of the samples. The time-domain waveform of the vibration signal of the CWRU inner race fault is shown in Figure 5(a). Figures 5(b)–5(g), respectively, show the time-frequency distribution of STFT, ST, Wigner–Ville transformation (WVT) [37], SSST, MSST, and MSSST.

It can be seen from Figure 5 that the energy concentration of time-frequency images by STFT and ST is poor, and WVT has a serious cross-term. Further comparing SSST, MSST, and MSSST, we find that MSST and MSSST have better energy concentration, which verifies that iterative compression can improve the time-frequency analysis algorithm, and we also find that the time-frequency image of MSSST has less redundant information than that of MSST. Meanwhile, it can be seen from Table 3 that the RE of MSSST is always the lowest, which further indicates that the energy of the time-frequency distribution by MSSST is more concentrated. Therefore, in this paper, MSSST is adopted as the time-frequency analysis method of the rolling bearing vibration signal.

3.3. Ablation Study

3.3.1. Hyperparameter r. From Section 2.2.2, it can be seen that the hyperparameter r directly determines the dimension of feature coding of the sample. In order to obtain the algorithm model, it is necessary to determine the value of the hyperparameter r. We change the value of r and combine different time-frequency analysis methods with FDL to carry out the experiments. To avoid contingency and particularity, each method is performed for ten repeated runs under different r. We determine the value of r by the average recognition accuracy of the validation set.

Detailed results are available in Figure 6. Table 4 records the highest average recognition accuracy and its corresponding standard deviation and r. According to Figure 6, from the overall trend, the average accuracy of MSSST is higher than that of other time-frequency analysis methods. From Table 4, it can be observed that the standard deviations

Table 3: Rényi entropy of different frequency analysis methods.

Dataset	Fault type	Method					
		STFT	ST	WVT	SSST	MSST	MSSST
CWRU	Normal	17.9889	16.9369	16.3525	12.6381	11.5885	10.6908
	Inner race	17.7992	16.8018	16.9188	12.7815	12.7316	11.8420
	Ball	17.1826	16.6566	16.4919	12.1864	11.2877	10.6046
	Outer race	17.6826	17.1549	16.3112	12.8543	12.2866	11.4123
MFPT	Normal	19.5311	18.5683	17.8579	14.8126	14.5465	12.4836
	Inner race	19.0094	18.3162	18.2588	14.6791	13.7552	11.1386
	Outer race	19.9868	19.3678	19.0506	14.4431	13.6602	11.9499

of diagnostic accuracy are relatively small, which indicates that the obtained models are relatively stable.

3.3.2. Module Combination. We carry out a series of experiments to study the effect of time-frequency analysis and feature extraction methods. The hyperparameters r applied to the testing set are set according to Table 4. The experimental results are shown in Figure 7. We have the following three findings. (1) Under the same conditions, the recognition accuracy of MSSST is always higher than that of MSST. Investigating its reason, MSST is based on STFT and the width of its window function is fixed. However, it is based on ST that MSSST has a variable window function, which improves the self-adaptability of spectrum analysis and can extract more detailed time-frequency characteristics in vibration signals. (2) The features extracted using NMF + NLE have higher recognition accuracy than those extracted directly using the LBP operator. That is because NMF + NLE gets the optimal feature dimension, while LBP has a fixed value of 59. (3) When LBP is combined with NMF + NLE, the recognition accuracy reaches the highest, which indicates that when NMF + NLE is directly used to extract the features of the time-frequency image, its dimension is too large. Therefore, too much redundant information is also extracted, which affects the recognition effect.

Taken together, these results suggest that whether in the CWRU dataset or the MFPT dataset, MSSST + FDL is the best combination.

3.3.3. Time for Feature Extraction. To verify the superiority of the time efficiency of the feature extraction algorithm in this paper, four different methods are designed to extract the features of the MSSST time-frequency image, and the time taken by them is recorded, respectively. We define T_d as the time spent in dictionary learning and T_f as the time spent in feature extraction of one sample.

As shown in Table 5, in CWRU and MFPT datasets, the dictionary learning time of texture feature vectors of time-frequency images is 0.037 and 0.012 hours, respectively. And yet, the time of dictionary learning on time-frequency images is 16.469 hours and 3.897 hours separately, which is very time-consuming. More importantly, in the process of fault recognition of the signal, although the time of only using LBP is the shortest, it can be seen from Section 3.3.2 that the recognition accuracy of this method is low. In addition, the feature coding of a time-frequency image takes 39,525 and 37,413.6 milliseconds, respectively.

By contrast, the proposed algorithm only consumes less than 100 milliseconds, saving more than 300 times of time, which can better meet the real-time requirements in practical engineering applications. At the same time, the feature coding algorithm in reference [27] is also very time-consuming, which further indicates that the proposed algorithm has high timeliness.

3.4. Comparisons with Other Methods. To prove the effectiveness of the proposed method, Table 6 shows the recognition accuracy of different fault diagnosis methods for bearing faults.

Reference [27] directly adopted NMFSC + NLE to obtain the sparse coding of MSST time-frequency images and trained SVM to diagnose bearing faults. The parameter sparsity was set to 0.7; the parameter rank was set to 25 and 100 in datasets CWRU and MFPT, respectively. In reference [38, 39], Hilbert–Huang transform and convolutional neural network (HHT + CNN) were combined to recognize the bearing state. The former input the CWRU time-frequency images of 32×32 pixels into CNN, while the latter input MFPT time-frequency images of 32×32 and 96×96 pixels. In reference [40], the wavelet packet energy features combined with multifractal features (WPE-MFs), which are of feature size 33, were used to train SVM.

It can be seen from the results that, compared with other methods, our method has only one hyperparameter and can achieve higher recognition accuracy with fewer features.

4. Extended Application

To demonstrate the practical engineering application value of the proposed method, we try to apply the proposed method to loudspeaker fault diagnosis.

4.1. Data Description and Analysis. The loudspeaker signal acquisition system consists of a microphone, acquisition card, sweep generator, and acquisition software. First, the sweep generator touches the signal collection points of the loudspeaker. Then, the signal acquisition software synchronously collects the data of the loudspeaker sound signal. Finally, the label of the collected sound signal is marked.

The sampling frequency of the loudspeaker sound signal is 8 kHz. To increase the number of samples and ensure that

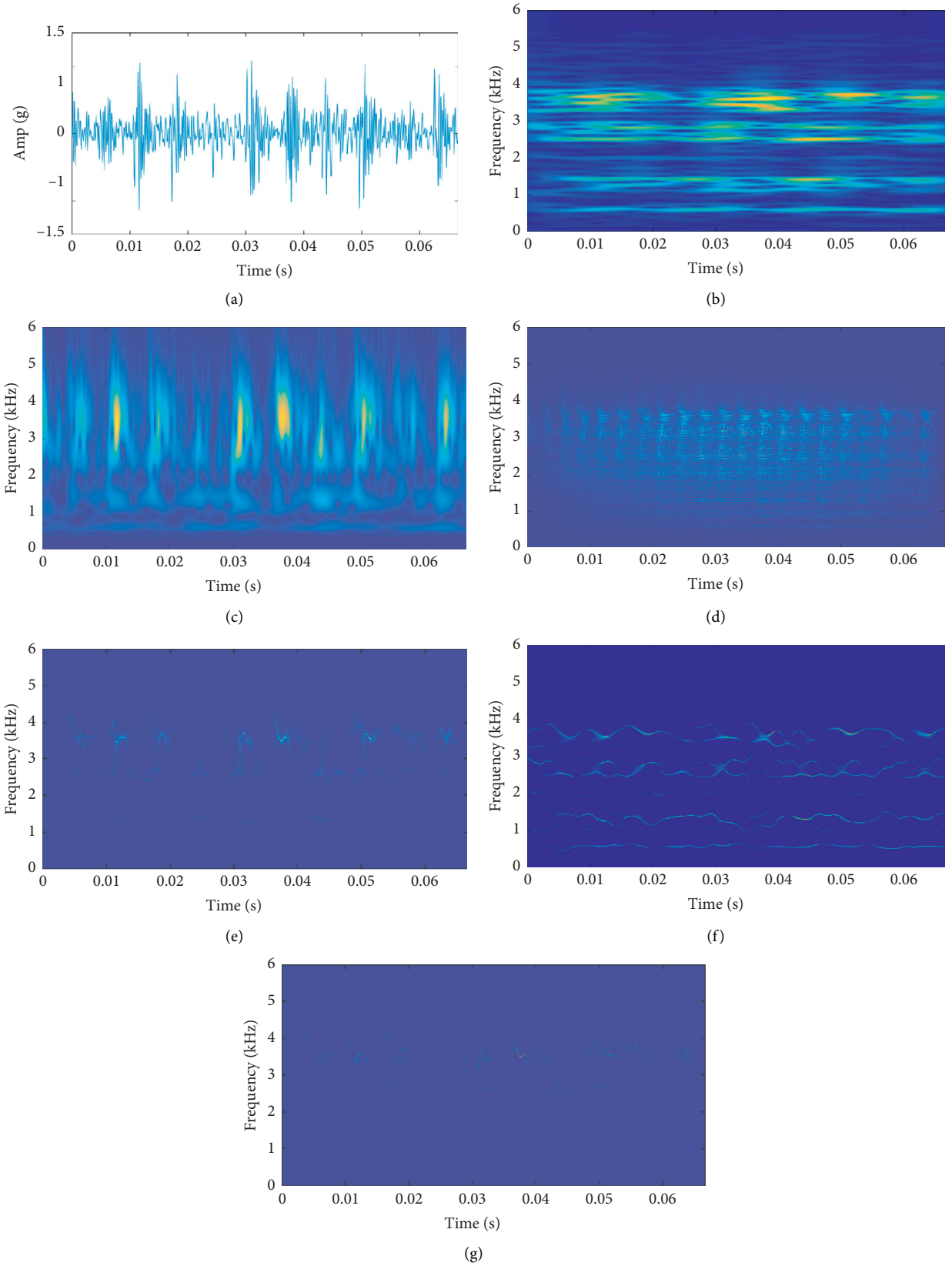


Figure 5: Time-frequency images of different frequency analysis methods: (a) raw signal; (b) STFT; (c) ST; (d) WVT; (e) SSST; (f) MSST; (g) MSSST.

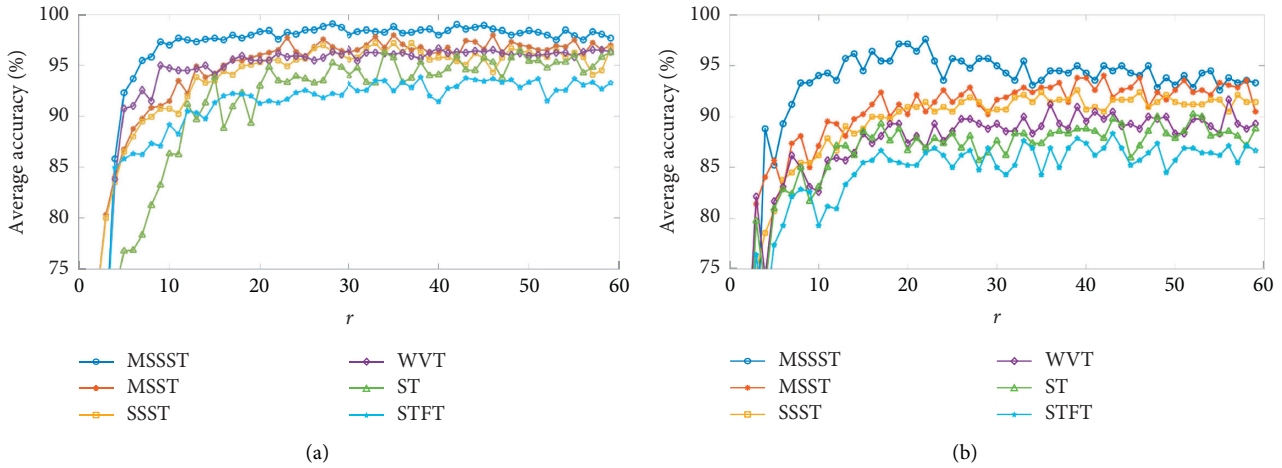
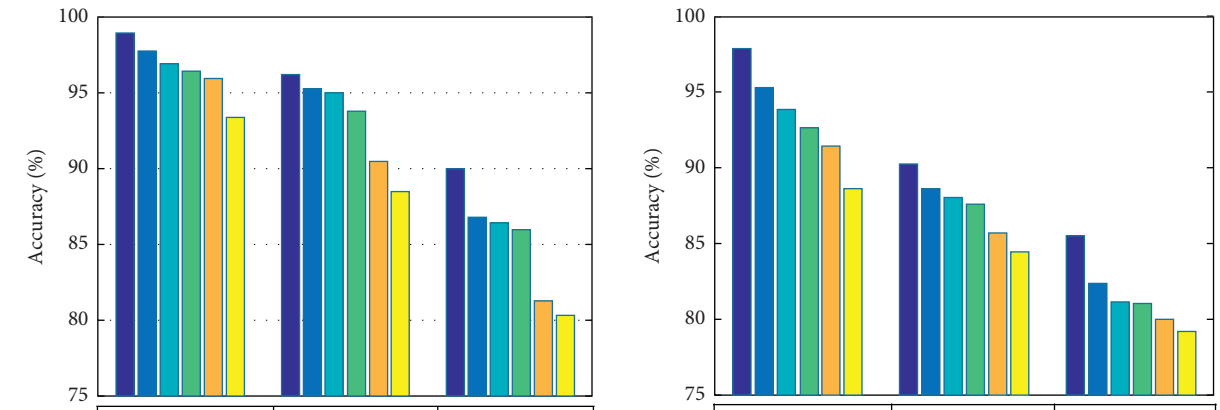


Figure 6: Average recognition accuracy of validation set under different r (a) in the CWRU dataset; (b) in the MFPT dataset.

Table 4: The highest average accuracy of validation set and its corresponding standard deviation and r.

Dataset	Data name	Methods					
		MSSST	MSST	SSST	WVT	ST	STFT
CWRU	Max average accuracy	99.08	98	97.16	96.66	96.41	93.83
	Standard deviation	0.32	0.42	0.46	0.43	0.51	0.42
	Optimal r	28	35	33	40	49	38
MFPT	Max average accuracy	97.61	94.04	92.61	91.66	90.24	88.33
	Standard deviation	0.36	0.44	0.42	0.48	0.53	0.50
	Optimal r	22	42	39	56	52	43



	LBP + NMF + NLE	NMF + NLE	LBP
MSSST	99	96.25	90
MSST	97.75	95.25	86.75
SSST	97	95	86.41
WVT	96.41	93.75	86
ST	96.00	90.42	81.25
STFT	93.33	88.5	80.33

	LBP + NMF + NLE	NMF + NLE	LBP
MSSST	97.85	90.23	85.55
MSST	95.23	88.55	82.30
SSST	93.80	88	81.1
WVT	92.61	87.55	81
ST	91.42	85.65	80
STFT	88.57	84.45	79.20

Figure 7: The recognition accuracy of testing set under different methods (a) in the CWRU dataset; (b) in the MFPT dataset.

Table 5: Time taken by different feature extraction methods.

Dataset	Feature extraction method	Feature size	Time	
			T_d (h)	T_f (ms)
CWRU	LBP	59	—	87.6
	NMF + NLE	28	16.469	39525.0
	LBP + NMF + NLE (FDL)	28	0.037	**88.8**
	NMFSC + NLE [27]	25	8.089	19413.0
MFPT	LBP	59	—	93.6
	NMF + NLE	22	3.897	37413.6
	LBP + NMF + NLE (FDL)	22	0.012	**95.4**
	NMFSC + NLE [27]	100	2.247	21573.6

Table 6: Comparison of the proposed method with other methods.

Dataset	Method	Sparsity	Feature size	Fault types	Accuracy (%)
CWRU	The proposed method	—	28	10	**99**
	SFC-DL + SVM [27]	0.7	25	10	98.03
	HHT + CNN [38]	—	32×32	10	95
	WPE-MF + SVM [40]	—	33	10	88.9
MFPT	The proposed method	—	22	3	**97.85**
	SFC-DL + SVM [27]	0.7	100	3	95.83
	HHT + CNN [39]	—	96×96	3	92.9
			32×32	3	75.9

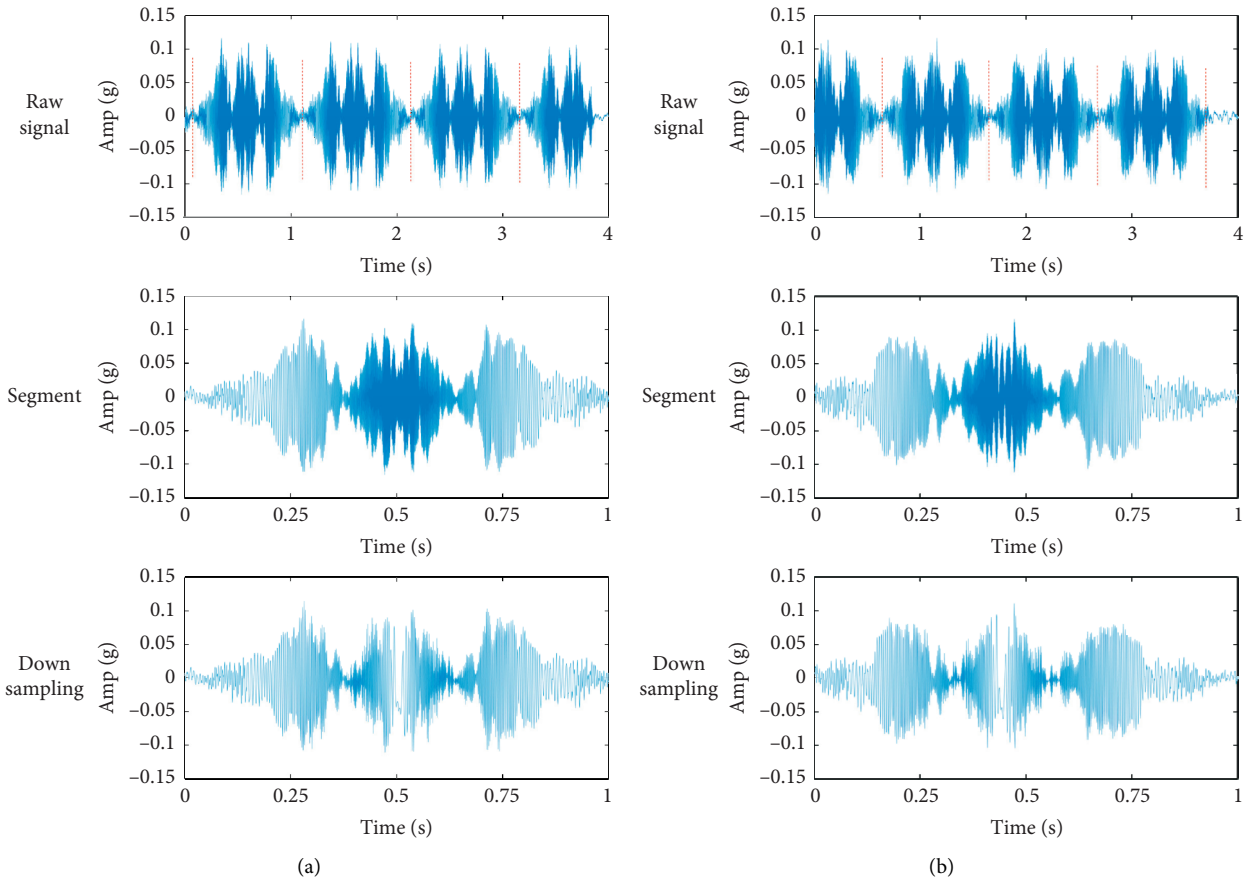


Figure 8: Time-domain waveform of loudspeaker signals: (a) normal; (b) abnormal.

Table 7: Description of loudspeaker pure-tone detection dataset.

Fault type	Training set	Validation set	Testing set	Fault label
Normal	600	200	200	1
Abnormal	600	200	200	2

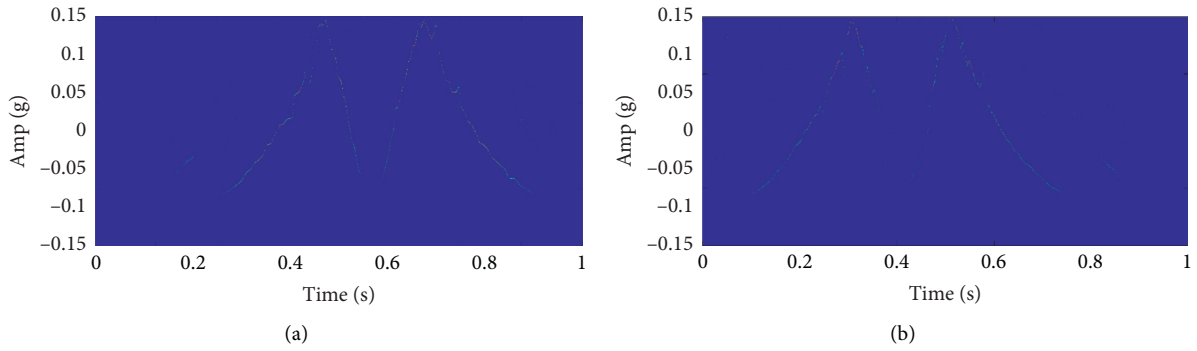


Figure 9: Time-frequency images of loudspeaker signals: (a) normal; (b) abnormal.

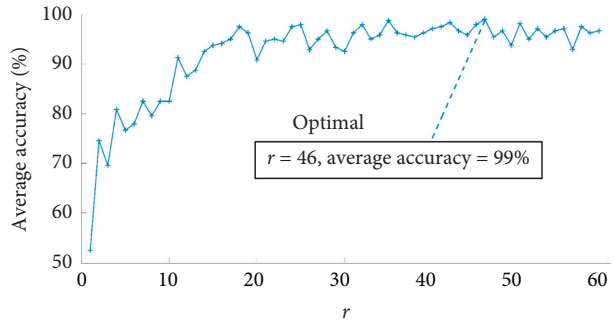


Figure 10: Average recognition accuracy of loudspeaker validation set under different r.

each sample contains one complete period at least, as shown in Figure 8, the raw sound signal is randomly segmented to obtain the segmented samples with 8000 sampling points. Then, the down-sampling process is performed on the segmented samples and each sample is of 1600 sampling points. And it can be seen from Figure 8 that the down-sampling sample not only reduces the computational complexity but also retains the waveform of the segmented sample. Finally, we can obtain a total of 2000 samples, which are divided into the training set, validation set, and testing set at a ratio of 6 : 2 : 2. More sample information is shown in Table 7.

4.2. Fault Diagnosis. The time-frequency images obtained by MSSST are shown in Figure 9. They have concentrated energy, but the dimension is 1600×3200, and there is a lot of irrelevant background information. Using the whole image as a raw sample for dictionary learning will consume a lot of computing resources and time, and redundant information will be added to the generated dictionary. Therefore, we can extract the optimal features of time-frequency images by FDL.

Figure 10 shows the average recognition accuracy of the validation set under different r values. It can be seen that the optimal value of r is 46. At this time, the average recognition accuracy of the validation set is 99%. This model is applied to the testing set. The experimental results show that 2 normal samples are misdiagnosed as the abnormal state and 1 abnormal sample is misdiagnosed as the normal state. Namely, the recognition accuracy is 98.5%.

In reference [41], the second-order time-reassigned multisynchrosqueezing transform was used to obtain the time-frequency images of the loudspeaker sound signals, and CNN was adopted to implement the feature extraction and fault recognition. The accuracy was 98.25%. By comparison, the proposed method is not inferior to the deep learning method.

5. Conclusions

In this paper, a new fault diagnosis method for the rolling bearing, which is called MSSST-FDL, is proposed to boost the speed and accuracy of recognition. Experiments show that the MSSST has better energy concentration than other time-frequency analysis methods; time-frequency images

with better energy concentration can improve the quality of fault diagnosis; the dictionary learning and feature coding of LBP feature vectors are faster than those of the whole time-frequency images, which can not only quickly determine the optimal hyperparameter but also meet the real-time requirement of fault diagnosis. The effectiveness of the proposed method is verified in CWRU and MFPT datasets, and the fault recognition accuracy is 99% and 97.85%, respectively. Furthermore, we apply the proposed method to loudspeaker anomaly diagnosis, and the recognition accuracy reaches 98.5%, which indicates that our method has the potential to be applied to other equipment.

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

This work was supported in part by the National Natural Science Foundation of China (grant no. 51775177).

References

[1] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: a review and roadmap," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106587, 2020.

[2] F. Zhou, S. Yang, H. Fujita, D. Chen, and C. Wen, "Deep learning fault diagnosis method based on global optimization gan for unbalanced data," *Knowledge-Based Systems*, vol. 187, Article ID 104837, 2020.

[3] Z. He, H. Shao, P. Wang, J. Lin, J. Cheng, and Y. Yang, "Deep transfer multi-wavelet auto-encoder for intelligent fault diagnosis of gearbox with few target training samples," *Knowledge-Based Systems*, vol. 191, Article ID 105313, 2020.

[4] M. Zeng, W. Zhang, and Z. Chen, "Group-based k-svd denoising for bearing fault diagnosis," *IEEE Sensors Journal*, vol. 19, no. 15, pp. 6335–6343, 2019.

[5] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2727–2736, 2018.

[6] D. Gabor, "Theory of communication. part 1: the analysis of information," *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946.

[7] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 3, pp. 235–238, 1977.

[8] J. Morlet, G. Arens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory-part I: complex signal and scattering in multilayered media," *Geophysics*, vol. 47, no. 2, pp. 203–221, 1982.

[9] R. G. Stockwell, L. Mansinha, and R. P. Lowe, "Localization of the complex spectrum: the S transform," *IEEE Transactions on Signal Processing*, vol. 44, no. 4, pp. 998–1001, 1996.

[10] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

[11] L. Cohen, *Time-Frequency Analysis: Theory and Applications*, Prentice-Hall, Upper Saddle River, NJ, USA, 1995.

[12] I. Daubechies, J. Lu, and H. Wu, "Synchrosqueezed wavelet transforms: a tool for empirical mode decomposition," 2009, http://arxiv.org/abs/0912.2437.

[13] F. Auger, P. Flandrin, Y.-T. Lin et al., "Time-frequency reassignment and synchrosqueezing: an overview," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 32–41, 2013.

[14] Z.-l. Huang, J. Zhang, T.-h. Zhao, and Y. Sun, "Synchrosqueezing s-transform and its application in seismic spectral decomposition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 2, pp. 817–825, 2016.

[15] G. Yu, Z. Wang, and P. Zhao, "Multisynchrosqueezing transform," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 7, pp. 5441–5455, 2019.

[16] K. Lin, *Research on fine-grained fault identification algorithm of rotating machinery vibration signal based on time-frequency image and multi-level sparse coding*, Ph.D. thesis, Hubei University of Technology, Wuhan, China, 2020.

[17] S. Guo, T. Yang, W. Guo, and C. Zhang, "A novel fault diagnosis method for rotating machinery based on a convolutional neural network," *Sensors*, vol. 18, no. 5, 2018.

[18] M.-F. Guo, N.-C. Yang, and W.-F. Chen, "Deep-Learning-based fault classification using hhilbert-huang transform and convolutional neural network in power distribution systems," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6905–6913, 2019.

[19] S. Duan, H. Zheng, and J. Liu, "A novel classification method for flutter signals based on the CNN and STFT," *International Journal of Aerospace Engineering*, vol. 2019, 8 pages, Article ID 9375437, 2019.

[20] Q. Li, X. Ding, Q. He, W. Huang, and Y. Shao, "Manifold sensing-based convolution sparse self-learning for defective bearing morphological feature extraction," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3069–3078, 2021.

[21] Y. Wang, X. Ding, Q. Zeng, L. Wang, and Y. Shao, "Intelligent rolling bearing fault diagnosis via vision convnet," *IEEE Sensors Journal*, vol. 21, no. 5, pp. 6600–6609, 2021.

[22] B. Li, P.-l. Zhang, D.S. Liu, S.-S. Mi, G.-Q. Ren, and H. Tian, "Feature extraction for rolling element bearing fault diagnosis utilizing generalized s transform and two-dimensional non-negative matrix factorization," *Journal of Sound and Vibration*, vol. 330, no. 10, pp. 2388–2399, 2011.

[23] B. Li, P.-l. Zhang, H. Tian, S.-s. Mi, D.-s. Liu, and G.-q. Ren, "A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10000–10009, 2011.

[24] P. Paatero and U. Tapper, "Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[25] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[26] X. Yu, E. Ding, C. Chen, and L. Li, "Rolling bearing fault state recognition method based on hht and supervised sparse coding," *Journal of China Coal Society*, vol. 40, no. 11, 2015.

[27] G. D. Sun, Y. Gao, K. Lin, and Y. Hu, "Fine-grained fault diagnosis method of rolling bearing combining multi-synchrosqueezing transform and sparse feature coding based on dictionary learning," *Shock and Vibration*, vol. 2019, 13 pages, Article ID 1531079, 2019.

[28] X. Li, Y. Yu, H. D. Shao, X. Zhong, J. Cheng, and J. s. Cheng, "Symplectic weighted sparse support matrix machine for gear fault diagnosis," *Measurement*, vol. 168, Article ID 108392, 2020.

[29] X. Li, Y. Yu, N. Q. Hu, Z. Cheng, and J. Cheng, "Discriminative manifold random vector functional link neural network for rolling bearing fault diagnosis," *Knowledge-Based Systems*, vol. 211, Article ID 106507, 2021.

[30] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.

[31] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.

[32] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[33] K. A. Loparo and K. A. Loparo, *Bearing Data Center*, Case Western Reserve University, Cleveland, OH, USA, 2013, http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-datacenter-website..

[34] E. Bechhoefer, "A quick introduction to bearing envelope analysis, MFPT Data," 2016, http://www.mfpt.org/FaultData/FaultData.htm.

[35] S. Aviyente and W. J. Williams, "Minimum entropy time-frequency distributions," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 37–40, 2005.

[36] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "measuring time-frequency information content using the renyi entropies," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1391–1409, 2001.

[37] T. A. Classen and M. Wolfgang, "The wigner distribution-a tool for time-frequency signal analysis, part ii: discrete-time signals," *Philips Journal of Research*, vol. 35, no. 3, pp. 276–300, 1980.

[38] Z. Yuan, L. B. Zhang, L. X. Duan, and T. Li, "Intelligent fault diagnosis of rolling element bearings based on hht and cnn," in *Proceedings of the Prognostics and System Health Management Conference*, pp. 292–296, Qingdao, China, 2019.

[39] W. L. Du, J. F. Tao, Y. M. Li, and C. L. Liu, "Wavelet leaders multifractal features based fault diagnosis of rotating mechanism," *Mechanical Systems and Signal Processing*, vol. 43, no. 1-2, pp. 57–75, 2014.

[40] V. David, A. Ferrada, E. L. Droguett, V. Meruane, and M. Modarres, "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings," *Shock and Vibration*, vol. 2017, Article ID 5067651, 17 pages, 2017.

[41] G. D. Sun, Y. Gao, Y. Xu, and W. Feng, "Data-driven fault diagnosis method based on second-order time-reassigned multisynchrosqueezing transform and evenly mini-batch training," *IEEE Access*, vol. 8, no. 99, p. 1, 2020.

Hindawi
Shock and Vibration
Volume 2021, Article ID 5561417, 24 pages
https://doi.org/10.1155/2021/5561417

Research Article

Root Crack Identification of Sun Gear in Planetary Gear System Combining Fault Dynamics with VMD Algorithm

Hongwei Fan,[1,2] **Yiqing Yang**,[1] **Hongwei Ma**,[1,2] **Xuhui Zhang**,[1,2] **Xiang Wan**,[1,2] **Xiangang Cao**,[1,2] **Qinghua Mao**,[1,2] **Chao Zhang**,[1] **and Qi Liu**[1]

[1]*School of Mechanical Engineering, Xi'an University of Science and Technology, No. 58 Yanta Road, Xi'an 710054, China*
[2]*Shaanxi Key Laboratory of Mine Electromechanical Equipment Intelligent Monitoring, Xi'an University of Science and Technology, No. 58 Yanta Road, Xi'an 710054, China*

Correspondence should be addressed to Hongwei Fan; hw_fan@xust.edu.cn

Received 21 February 2021; Revised 6 April 2021; Accepted 30 April 2021; Published 17 May 2021

Academic Editor: Jinde Zheng

Planetary gearbox is widely used in various low-speed machines due to its large transmission ratio. However, the gears in a planetary gearbox are prone to the mechanical faults due to the complex dynamic heavy load. Vibration frequencies caused by an early tooth root crack of sun gear are usually difficult to accurately extract, so its fault diagnosis is one of the main challenges of planetary gearbox reliability. In this paper, a simplified tooth root crack model of sun gear is proposed, and then a rigid-flexible coupled dynamics model of the whole planetary gear system is constructed. By the numerical simulation, the fault frequencies caused by a tooth root crack of sun gear are obtained. A Variational Mode Decomposition (VMD) algorithm for the vibration frequency extraction is proposed. The measured vibration signals are decomposed into the sparse Intrinsic Mode Functions (IMFs) by the VMD, and then the IMFs are further analyzed by the spectral method to accurately extract the crack-induced frequency components. The experimental results show that the proposed dynamics model and VMD method are feasible; an error between the characteristic frequencies from the tested signal analysis and the theoretical calculation is less than 1%.

1. Introduction

Because of large transmission ratio and compact structure, the planetary gearbox is widely used in various low-speed heavy machines. However, in the actual operation process, due to the complex dynamic load, the planetary gearbox is prone to various kinds of mechanical failure [1]. When the gearbox is subjected to a heavy or impact load, it is very easy to cause the tooth root crack. Because the early crack has a small size, the vibration energy caused by the crack is weak, so the crack is difficult to identify in time, and further the crack can develop into a serious fault, such as tooth breaking. Therefore, it is significant to study the feature analysis technology of gear crack in a planetary gearbox for the crack diagnosis.

At present, there are two kinds of tooth root crack modeling methods [2, 3], one is the parabolic crack model and the other is the penetrating crack model with simultaneous propagation along both tooth width and thickness directions. Chaari et al. [4, 5] put forward a penetrating tooth root crack model. Chen and Shao [6, 7] proposed a parabolic crack model with simultaneous propagation along both tooth width and thickness directions. Ma et al. [8] established a crack model with propagation only along the thickness direction, in which the crack depth is simulated by a parabola whose starting point is on the tooth transition circle and symmetric about the center line of tooth. Based on the tooth crack model, a whole dynamics model of planetary gear system can be built. There are three kinds of dynamics model of planetary gear system: lumped parameter model [9, 10], finite element model [11], and rigid-flexible coupled model [12,13]. The lumped parameter model of a planetary gearbox was established, and the influence of the installation error of planet carrier on the vibration response was studied by Zhai et al. [14]. Jiang et al. [15] built a lumped parameter model of planetary gear

system to solve the excitation forces that generate vibrations. Shen et al. [16] were devoted to analyzing fault mechanism and dynamic characteristics of gear wear and proposed a torsional dynamics model of planetary gear system with tooth wear. Liu et al. [17] established a 3D finite element model of planetary gearbox and found that the effect of crack on the vibration response is mainly reflected in the frequency domain. Liu et al. [18] proposed a dynamics finite element model of planetary gear system, in which there is a localized fault to be considered for the outer race of the planet bearing, and Liu et al. [19] also presented a rigid-flexible coupled dynamics model for planetary gear system, both the flexible ring gear, flexible supports of the ring gear and sun gear, and a rectangular local fault in the planet bearing are formulated in the model. Duan et al. [20] carried out a system transmission error investigation of gearbox by proposing a rigid-flexible coupled dynamics model, and the modal test was launched to verify its effectiveness. Fan et al. [21] developed a rigid-flexible coupled dynamics model of the planetary gearbox considering the flexibility of ring gear and sun gear shaft based on the shell theory and Timoshenko beam theory, respectively. The lumped parameter model cannot take into account the influence of gear tooth deformation on the system, and the whole finite element model is inefficient. Therefore, the rigid-flexible coupled model is adopted in this paper.

There are two kinds of common vibration signal analysis methods, including frequency-domain and time-frequency analysis. The frequency-domain signal analysis based on the Fourier transform is suitable for the periodic signals [22]. Time-frequency analysis is appropriate for the nonstationary signals. Typical time-frequency analysis methods include the short-time Fourier transform (STFT) [23], wavelet transform (WT) [24, 25], and Hilbert–Huang transform (HHT) [26, 27]. Huang et al. [28] proposed an Empirical Mode Decomposition (EMD) method based on the whole decomposition of signals, which avoid the selection of wavelet functions. Wang et al. [29] studied a signal analysis method based on EMD and extracted the effective signal components from the background noise. Yan and Liu [30] constructed an improved weak signal detection method to reduce the boundary effect of EMD. Ensemble Empirical Mode Decomposition (EEMD) method is proposed by Wu and Huang [31] to solve the mode confusion problem of EMD. Ma et al. [32] proposed a weak signal detection method by combining EMD with singular spectrum analysis. Due to the low computational efficiency of EEMD, Dragomiretskiy et al. [33] proposed a new signal decomposition method in 2014, i.e., Variational Mode Decomposition (VMD); it can avoid the mode confusion and has high efficiency, antinoise ability, and high-frequency resolution. Li and Zhu [34] applied the VMD for the mechanical fault diagnosis; the results show that the VMD is better than EEMD in signal decomposition effect. Wang et al. [35] used the VMD to realize the feature extraction of gear tooth crack in a single-stage gearbox, which verified the effectiveness of the VMD method for gear fault feature extraction.

In order to reduce the modeling error and the machining difficulty of the cracked sun gear, a simplified gear tooth crack model is proposed in this paper. Combined with the crack model, a rigid-flexible coupled model of planetary gear system with tooth crack of sun gear is constructed, and the dynamics simulation is designed to find the crack-induced signal frequency characteristics. In order to identify the crack from the tested vibration signals, a VMD-based signal analysis method is proposed to extract the featured frequency of tooth crack. The structure of this paper is as follows. Section 1 introduces the fault diagnosis principle based on vibration analysis and the theoretical characteristic frequency calculation. In Section 2, the crack modeling is introduced and the dynamics simulation is completed. Section 3 introduces the VMD-based fault feature extraction method, and in Section 4, the proposed model and analysis method are verified by experiments.

2. Frequency-Domain Diagnosis Basis for Planetary Gear System

In this paper, the crack mechanism of a planetary gear system is analyzed by dynamics simulation, and the crack type and its degree are judged by the characteristic frequencies [36].

2.1. Working Frequency of Planetary Gear System. A planetary gearbox consists of the sun gear, planet gear, ring gear, and planet carrier. If its basic components include two central gears Z and a planet carrier X, the planetary gear system type is $2Z$-X, as shown in Figure 1.

The dynamics response of a $2Z$-X planetary gear system is more complex than that of a fixed-axis gear system because of the relative motion among all of the gears. Therefore, the calculation of transmission ratio and characteristic frequencies cannot directly refer to that of a fixed-axis gear system.

The transmission ratio of a planetary gear system is

$$i = \frac{f_s}{f_c} = \left(\frac{Z_r}{Z_s}\right) + 1, \tag{1}$$

where f_c is the rotational frequency of planet carrier, Z_r is the tooth number of ring gear, and Z_s is the tooth number of sun gear.

Since the rotational speed of sun gear is equal to the input speed of transmission shaft, the rotational frequency of sun gear is

$$f_s = \frac{n}{60}, \tag{2}$$

where f_s is the rotational frequency of sun gear and n is the input rotational speed of transmission shaft.

The fundamental difference between the planetary and fixed-axis gear systems is that the planetary gear system has a planet carrier rotating around the main axis, and the planet gear installed on the planet carrier has both rotation and revolution. According to the relative motion principle, a public frequency can be added to the whole planetary gear system, which is equal to the rotational frequency of planet carrier and has an opposite direction, while the relative

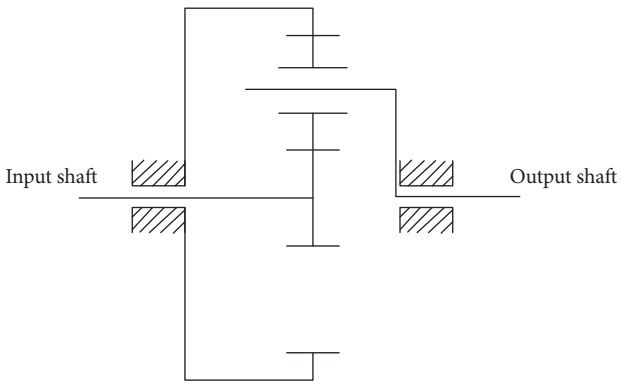


Figure 1: Schematic diagram of 2*Z*-*X* planetary gear system.

motion among the components of planetary gear system remains constant.

$$\frac{f_s - f_c}{f_r - f_c} = -\frac{Z_r}{Z_s}, \tag{3}$$

where f_r is the rotational frequency of ring gear.

As can be seen from equation (3), the rotational frequency of planet carrier is

$$f_c = \frac{Z_s}{Z_S + Z_r} \cdot f_s. \tag{4}$$

Similarly, the rotational frequency of planet gear is

$$f_p = (f_s - f_c) \cdot \left(\frac{Z_s}{Z_P}\right) - f_c, \tag{5}$$

where f_p is the rotational frequency of planet gear and Z_p is the tooth number of planet gear.

In the transmission process of a gear pair, the meshing frequencies of two gears engaged with each other are equal. Therefore, in a planetary gear system, the meshing frequency of each gear pair is equal, i.e.,

$$f_m = f_c \cdot Z_r = \frac{Z_s \cdot Z_r}{Z_s + Z_r} \cdot f_s, \tag{6}$$

where f_m is the meshing frequency of planetary gear system.

2.2. Fault Frequency of Planetary Gear System. If there is a crack in one of the teeth of sun gear in a planetary gear system, the cracked gear meshes with all the gears in a rotational period relative to the planet carrier. Therefore, the fault frequency of sun gear is

$$f_{sg} = \frac{f_m}{Z_s} \cdot n, \tag{7}$$

where f_{sg} is the fault frequency of sun gear.

Considering the modulation of fault frequency and fundamental frequency, the vibration signal model of a planetary gear system can be simply expressed as [37]

$$x(t) = h(t)c\left[1 + A\cos\left(2\pi f_g t + \phi\right)\right] \cdot \cos\left[2\pi f_m t + B\sin\left(2\pi f_g t + \varphi\right) + \theta\right], \tag{8}$$

where A and B are the amplitude and frequency modulation intensity caused by gear fault, A, $B > 0$; Φ, φ, and θ are the initial phase; f_m is the meshing frequency; f_{sg} is the fault frequency; c is a dimensionless constant, usually $c = 1$; and $h(t)$ describes an amplitude modulation of vibration transfer path.

If the vibration sensor is installed at the top of ring gear, the distance from the fault position to the sensor changes with the rotation of sun gear when a local fault occurs.

$$h(t) = \left[1 - \cos\left(2\pi f_s t\right)\right]. \tag{9}$$

The amplitude modulation of vibration signal can be written as [37]

$$\begin{aligned} h(t)a(t) &= \left[1 - \cos\left(2\pi f_s t\right)\right] \cdot \left[1 + A\cos\left(2\pi f_{sg} t + \phi\right)\right] = 1 - \cos\left(2\pi f_s t\right) + A\cos\left(2\pi f_{sg} t + \phi\right) \\ &\quad - \frac{1}{2}A\left\{\cos\left[2\pi\left(f_s + f_{sg}\right)t + \phi\right] + \cos\left[2\pi\left(f_s - f_{sg}\right)t - \phi\right]\right\}. \end{aligned} \tag{10}$$

The Fourier transform is applied to equation (10):

$$H(f) * A(f) = \delta(f) - \delta\left(f - f_s\right) + A\delta\left(f - f_s\right)\exp\left(j\phi\right) - \frac{1}{2}A\left\{\delta\left[f - \left(f_{sg} + f_s\right)\right]\exp\left(j\phi\right) + \delta\left[f - \left(f_s - f_{sg}\right)\right]\exp\left(-j\phi\right)\right\}. \tag{11}$$

As can be seen from equation (11) [37], when a crack occurs in the sun gear, the peak value in the amplitude spectrum appears in the rotational frequency f_s and its multiple frequencies mf_s, the fault frequency f_{sg} and its multiple frequencies nf_{sg}, and the combined frequencies $f_m \pm f_s$ and $nf_{sg} \pm mf_m$. If mf_s and nf_{sg} are considered as the modulation frequencies, the other peaks mainly appear at $1/4f_s$ and $n/4f_s$, f_m and mf_m, and the combined frequencies $n/4f_{sg} \pm mf_m(m, n = 1, 2, \ldots)$.

3. Crack Modeling and Dynamics Analysis

3.1. Geometric Model of Planetary Gear System. The 2Z-X planetary gear system used is composed of a sun gear, a planet carrier, a ring gear, and four planet gears. A 3D model of the planetary gear system is shown in Figure 2; the design parameters are shown in Table 1. On the premise of not affecting the analysis results, the auxiliary parts are simplified; for example, both the bearings and bolts used to connect the planet gears and the planet carrier are treated as cylinders.

3.2. Rigid-Flexible Coupled Model with Crack. In the above planetary gear system, the sun gear contacts with four planet gears to transmit motion and torque. When the sun gear is used as an input, it is simultaneously subjected to forces and torques from four planet gears. Therefore, the sun gear is prone to a mechanical fault. During gear transmission, the bending stress is concentrated at the tooth root, and the crack begins and extends from the tooth root [38], as shown in Figure 3.

In Figure 3, φ is the included angle between the tooth center line and the tangent line of root fillet curve, and the area between $\varphi = 30^\circ$ and $\varphi = 34^\circ$ is the stress concentration area [6]. The included angle between the crack extension line and tooth center line is defined as the crack extension angle α. In order to approach the crack growth law under the real working condition, $\varphi = 32^\circ$ and $\alpha = 74^\circ$ are taken in the following dynamics simulation. The chord of the approximate arc of the tooth root crack growth path is taken as the crack width growth line, i.e., w, shown in Figure 3, and the tooth width direction is called the crack length extension line, i.e., l, shown in Figure 3.

In order to study the frequency characteristics of different-scale cracks, three kinds of crack were designed, as shown in Table 2. It is known that the tooth width of sun gear is $l = 10$ mm, and the maximum crack width is $w = 1$ mm. The crack width and length were expanded simultaneously in a proportion of 20%, 40%, and 80% relative to the chord of tooth root arc and tooth width, respectively.

In order to reduce the difficulty of making crack, the crack width propagation path is treated approximately as an arc chord, as shown in Figure 4.

In order to obtain the crack-induced vibration frequency characteristics of sun gear, the finite element model is used to treat the sun gear with flexibility. The flexibilization steps are shown in Figure 5.

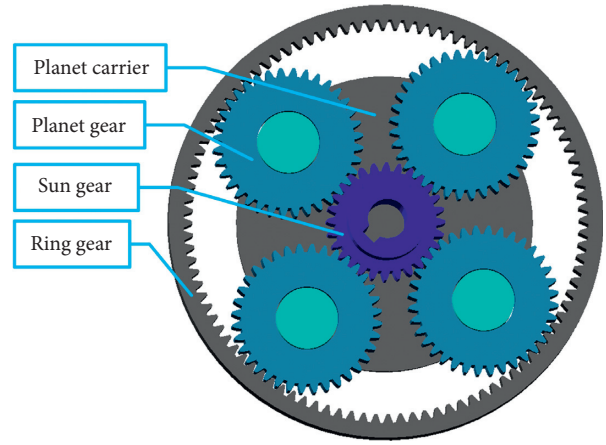


Figure 2: 3D model of planetary gear system.

As shown in Figure 5, for a finite element model, the generation quality of mesh is closely related to the element type and quantity. In this paper, ANSYS is used to make sun gear flexible. Considering the crack shape, a tetrahedron element is suitable for the gear with crack. For the tetrahedral element, the SOLID187 uses an optimized algorithm compared to the SOLID92 to achieve a high-quality model. When generating the mesh, the part without crack is firstly divided, and then the crack area is refined separately. The meshed sun gear is shown in Figure 6. Figure 6(a) is the entire mesh model; Figure 6(b) is the crack mesh model.

When a finite element model is introduced into a rigid-body dynamics system and connected to other parts, there is a region that does not deform, i.e., the rigid region. In this paper, ADAMS is used for the dynamics simulation; the rigid region is composed of the nodes of two interfaces and the surrounding nodes. Select those nodes as the master nodes and the nodes of end circles of sun gear as the slave nodes; the rigid region is created, as shown in Figure 7(a); the rigid-flexible coupled model of planetary gear system is obtained by connecting the flexible sun gear with other rigid bodies through rigid regions and interface nodes, as shown in Figure 7(b).

3.3. Dynamics Analysis. In dynamics simulation, the input shaft of torque and speed is the shaft where the sun gear is located, and the output shaft is that connected by the planet carrier. The parameters are set as shown in Table 3.

The input rotational frequency is 30 Hz, the torque is 5 N·m, and the direction of torque is opposite to the direction of rotational frequency. In order to avoid the sudden loading, the STEP function was adopted, which is expressed as STEP (time,0,0,0.1,10800D), the initial speed is 0, the speed is 10800°/s at 0.1 s, i.e., the frequency is 30 Hz, and the torque gradually increased to 5 N·m in the form of STEP (time,0,0,0.1,5000). In the model settings, the rigid-rigid contact was used between four planet gears and the ring gear, the rigid-flexible contact was used between four planet gears and the sun gear, and the other parameters were considered in terms of [39].

Table 1: Design parameters of planetary gear system.

Types of gears	Ring gear	Planet gear	Sun gear
Number of teeth	100	36	28
Module (mm)	1	Pressure angle (°)	20
Tooth width (mm)	10	Material	Steel
Young's modulus (Pa×10^{11})	2.07	Siméon Denis Poisson	0.29
Input speed (rpm)	1800		

In the simulation, the ring gear was fixed and the contact force between the gears was added to simulate the meshing. The contact force between the gears changed the angular acceleration with the meshing process. Therefore, the characteristic frequencies of the tooth crack were analyzed by the angular acceleration signal of sun gear.

The dynamics simulation was carried out and the results of angular acceleration in the time and frequency domains were obtained under different crack cases, as shown in Figure 8. In Figures 8(a) and 8(b), four subfigures from top to bottom are the angular acceleration results in the time and frequency domains under no crack, $1^{\#}$ crack, $2^{\#}$ crack, and $3^{\#}$ crack.

As can be seen from Figure 8(a), the amplitude of the angular acceleration is obviously smaller when the sun gear has no crack than a crack fault, and the amplitude increases with the increase of the crack size. As can be seen from Figure 8(b), the amplitude peaks in the frequency spectrum mainly appear in the meshing frequency f_m and its multiple frequencies mf_m, the fault frequencies f_{sg} of sun gear and its multiple frequencies nf_{sg}, and the combined frequencies $mf_m + nf_{sg}$.

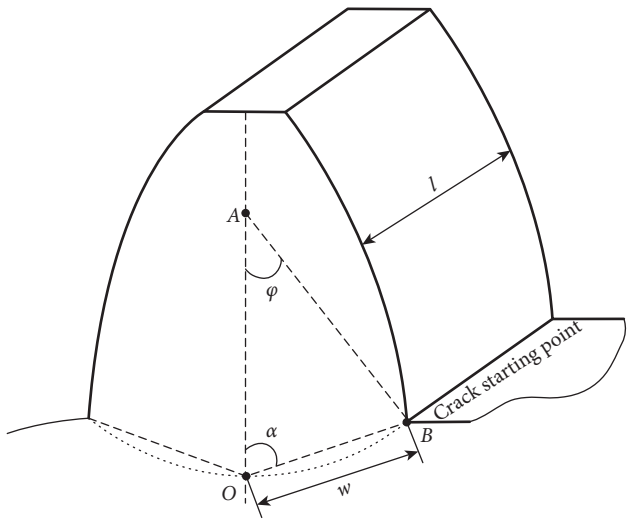


Figure 3: Tooth root crack model of sun gear.

Table 2: Crack parameters of sun gear.

Crack case	No crack	$1^{\#}$ crack	$2^{\#}$ crack	$3^{\#}$ crack
Crack length, l (mm)	0	2	4	8
Crack width, w (mm)	0	0.2	0.4	0.8

4. VMD Time-Frequency Analysis Method of Vibration Signals

The frequency resolution of the spectrum in Figure 8(b) is low, so it is difficult to accurately find the fault frequency values. Therefore, the VMD algorithm [40] is proposed for the time-frequency analysis of complex vibration signals.

4.1. Principle of VMD Algorithm. VMD is the decomposition of signal $x(t)$ into a number of sparse IMFs $m_k(t)$, each of which fluctuates around the central frequency ω_k, i.e., the bandwidth is limited around the central frequency. The bandwidth can be obtained by estimating the square norm of the gradient of the frequency shift signals.

First, the HHT transforms the real mode m_k into an analytic signal m_k^+ with a unilateral spectrum with non-negative frequencies:

$$m_k^+(t) = \left(\delta(t) + \frac{j}{\pi t}\right) * m_k(t). \tag{12}$$

Then, by multiplying with an exponential harmonic wave whose frequencies are adjusted to their respective central frequencies, the spectrum is shifted to the base band:

$$m_k^B(t) = \left(\delta(t) + \frac{j}{\pi t}\right) * m_k(t) e^{-j\omega_k t}. \tag{13}$$

Finally, the bandwidth is estimated by the square norm of the gradient:

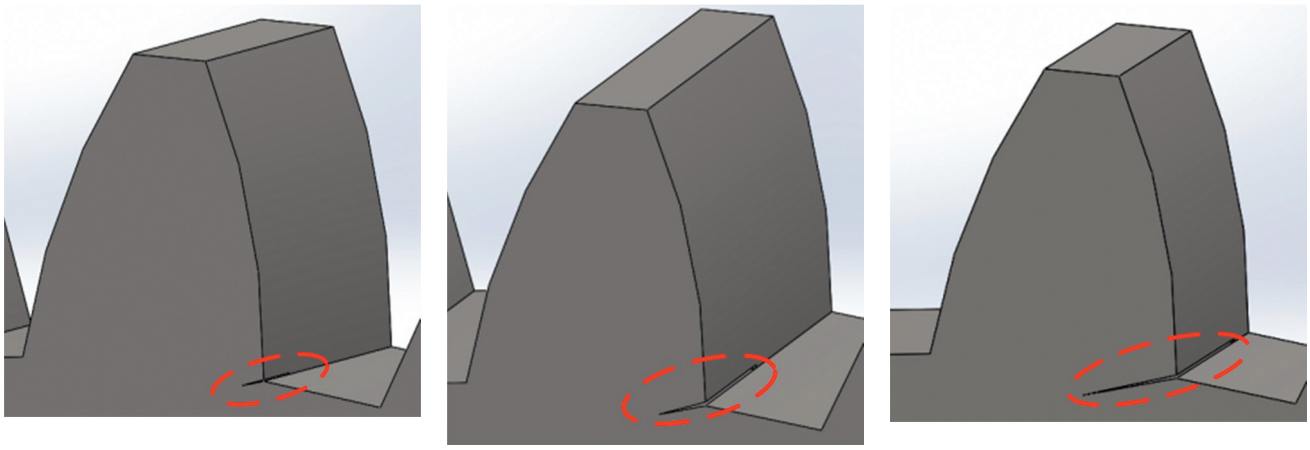
$$\triangle\omega = \left\| \partial t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * m_k(t) e^{-j\omega_k t} \right] \right\|_2^2. \tag{14}$$

Thus, all the estimated mode component bandwidths can be expressed as a constrained variational model:

$$\begin{aligned} &\min_{\{m_k(t)\},\{\omega_k\}} = \sum_{k=1}^{K} \left\| \partial t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * m_k(t) e^{-j\omega_k t} \right] \right\|_2^2, \\ &s.t. \sum_{k=1}^{K} m_k(t) = x(t), \end{aligned} \tag{15}$$

where $\delta(\cdot)$ is the Dirac Delta function, $*$ is the convolution operation, and K is the number of the IMFs.

In order to transform the optimization problem expressed by the above formula into an unconstrained form, a penalty factor α and a Lagrange multiplier are added to accelerate the convergence and strengthen the constraint. The objective function that needs to be minimized is transformed into an augmented Lagrange function:



(a) (b) (c)

Figure 4: 3D crack model of sun gear. (a) 1# crack. (b) 2# crack. (c) 3# crack.

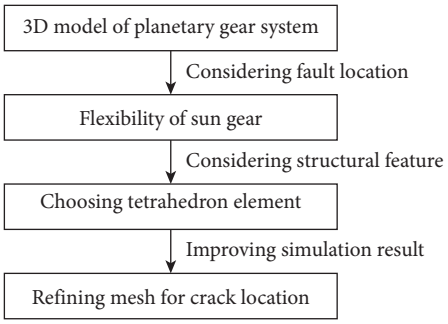
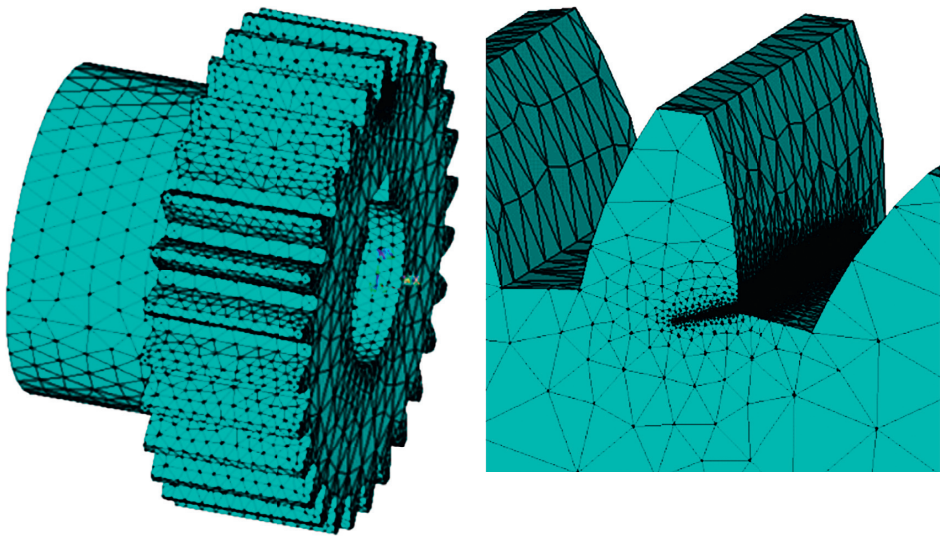
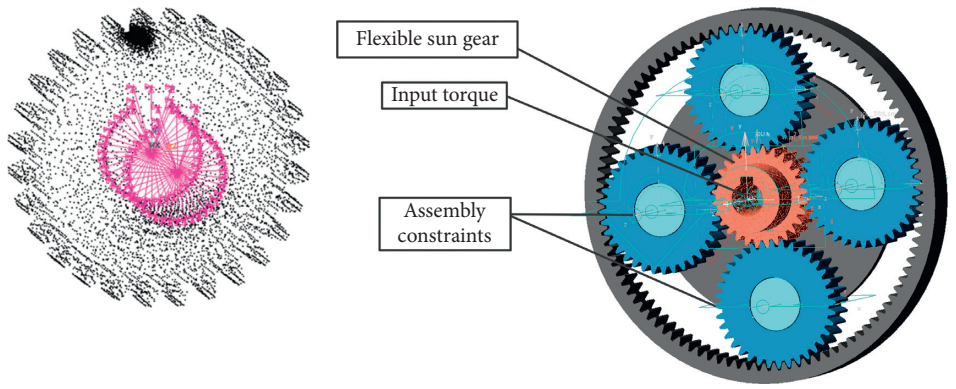


Figure 5: Flexibilization steps of sun gear.



(a) (b)

Figure 6: Mesh model of sun gear. (a) Entire mesh. (b) Crack mesh.



(a) (b)

Figure 7: Dynamics model of planetary gear system. (a) Rigid region model. (b) Rigid-flexible coupled model.

Table 3: Parameter setup used for dynamics simulation.

Parameter	Setup
Input frequency (Hz)	30
Input torque (N·m)	5
Function type	STEP

$$L[\{m_k(t)\},\{\omega_k\},\lambda(t)] = \alpha\sum_{k=1}^{K}\left\|\partial t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*m_k(t)e^{-j\omega_k t}\right]\right\|_2^2 + \left\|x(t)-\sum_{k=1}^{K}m_k(t)\right\|_2^2 + \langle\lambda(t),x(t)-\sum_{k=1}^{K}m_k(t)\rangle, \quad (16)$$

where $\lambda(t)$ is the Lagrange multiplier and $<\cdot,\cdot>$ is the inner product operation.

The minimization problem represented by equation (15) is solved by an Alternate Direction Multiplier (ADM) method to find the saddle point of an augmented Lagrange problem represented by equation (16). The equivalent minimization problem of equation (16) is solved, and each IMF is updated by the optimal solution $m_k(t)$:

$$m_k(t) = \arg\min_{m_k} L[\{m_k(t)\},\{\omega_k\},\lambda(t)] = \arg\min_{m_k}\left\{\alpha\sum_{k=1}^{K}\left\|\partial t\left\{\left[\delta(t)+\frac{j}{\pi t}\right]*m_k(t)e^{-j\omega_k t}\right\}\right\|_2^2 + \left\|x(t)-\sum_{k=1}^{K}m_k(t)+\frac{\lambda(t)}{2}\right\|_2^2\right\}. \quad (17)$$

In the frequency domain, the solution of equation (17) is

$$\hat{m}_k(\omega) = \hat{x}(\omega) - \frac{\sum_{i\neq k}\hat{m}_i(\omega) + 1/2\hat{\lambda}(\omega)}{1+2\alpha(\omega-\omega_k)^2}. \quad (18)$$

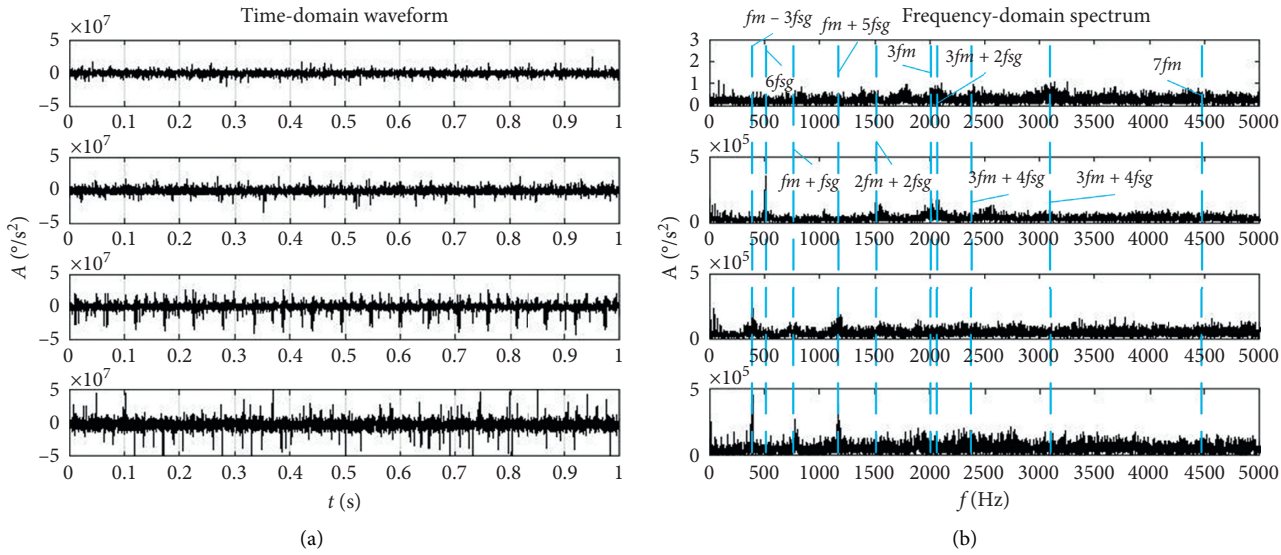
The IMF in the time domain can be obtained by Fourier inverse transform of equation (18).

The equivalent minimization problem of equation (16) is solved, and the central frequency ω_k of each IMF $m_k(t)$ is updated with the optimal solution:

$$\omega_k = \arg\min_{\omega_k} L[\{m_k(t)\},\{\omega_k\},\lambda(t)] = \arg\min_{\omega_k}\sum_{k=1}^{K}\left\|\partial t\left\{\left[\delta(t)+\frac{j}{\pi t}\right]*m_k(t)e^{-j\omega_k t}\right\}\right\|_2^2. \quad (19)$$

The central frequency ω_k is the centroid of the IMF power spectrum $\hat{m}_k(\omega)$, i.e.,

$$\omega_k = \frac{\int_0^\infty \omega|\hat{m}_k(\omega)^2|d\omega}{\int_0^\infty|\hat{m}_k(\omega)^2|d\omega}. \quad (20)$$



(a) (b)

Figure 8: Angular acceleration of sun gear for different crack cases. (a) Time-domain waveform. (b) Frequency-domain spectrum.

According to the solution process of the above constrained variational problem, the VMD can divide the frequency band according to the frequency characteristic of the signal and realize the adaptive decomposition of the signal by updating each mode and its corresponding central frequency.

The flowchart of the VMD algorithm is shown in Figure 9.

Step 1. Set $\{\hat{m}_k^0(t)\}$ $\{\hat{\omega}_k^0\}$ $\left\{\hat{\lambda}^0(t)\right\}$, and $n=0$; define the convergence threshold and number K of IMFs to be separated.

Step 2. $k = 1, 2, \ldots, K$ and $\omega \geq 0$; update each IMF $m_k(t)$ and its central frequency ω_k:

$$\hat{m}_k^{n+1}(\omega) = \hat{x}(\omega) - \sum_{i<k} \hat{m}_i^{n+1}(\omega) - \frac{\sum_{i>k} \hat{m}_i^n(\omega) + 1/2\hat{\lambda}^n(\omega)}{1+2\alpha(\omega-\omega_k^n)^2}, \quad \omega_k^{n+1} = \frac{\int_0^\infty \omega|\hat{m}_k^{n+1}(\omega)^2|d\omega}{\int_0^\infty |\hat{m}_k^{n+1}(\omega)^2|d\omega}. \tag{21}$$

Step 3. $\omega \geq 0$; update the Lagrange multiplier:

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau\left[\hat{x}(\omega) - \sum_k \hat{m}_k^{n+1}(\omega)\right], \tag{22}$$

where τ is the updating parameter of Lagrange multiplier.

Check the convergence condition:

$$\sum_k \frac{\left\|\hat{m}_k^{n+1}(t) - \hat{m}_k^n(t)\right\|_2^2}{\left\|\hat{m}_k^n(t)\right\|_2^2} < \varepsilon. \tag{23}$$

If the formula is true, let $m_k(t) = \hat{m}_k^{n+1}(t)$, $\omega_k(t) = \omega_k^{n+1}$, terminate the decomposition, and get K IMFs with the finite bandwidth. Otherwise, let $n = n+1$; return to Step 2.

4.2. Frequency-Domain Feature Extraction Using VMD. In the VMD algorithm, the value of K determines the number of IMFs in the decomposed signal. Taking 3# crack as an example, when K takes different values, the signal is decomposed by the VMD, and the frequency of each IMF is shown in Table 4. From Table 4, the central frequency of IMF6 is 4439 Hz, which is close to the maximum value of 4499 Hz in the case of $K=7$, and the central frequency of IMF4 is close to that of IMF5 in the case of $K=7$. In addition, since the planetary gear system consists of six moving parts, $K=6$ is reasonable.

When $K=6$, the angular acceleration signal for 3# crack is decomposed by the VMD, as shown in Figure 10, the VMD decomposition results for other crack cases are shown in Figure 11.

Figure 10 shows that the IMF frequencies correspond one to one with the central frequencies in Table 4 and are all the characteristic frequencies for sun gear crack. The central frequencies and amplitudes of the IMFs in Figure 11 are listed in Table 5. According to Table 5, Figure 12 is obtained.

As can be seen from Figure 12 and Table 5 above, the amplitude shown by the black line marked by rhombic dot is lower than others. When the cracks appear in the sun gear, the amplitudes corresponding to the central frequencies of each IMF are higher than those with no crack, and the errors of the central frequencies of each IMF are small.

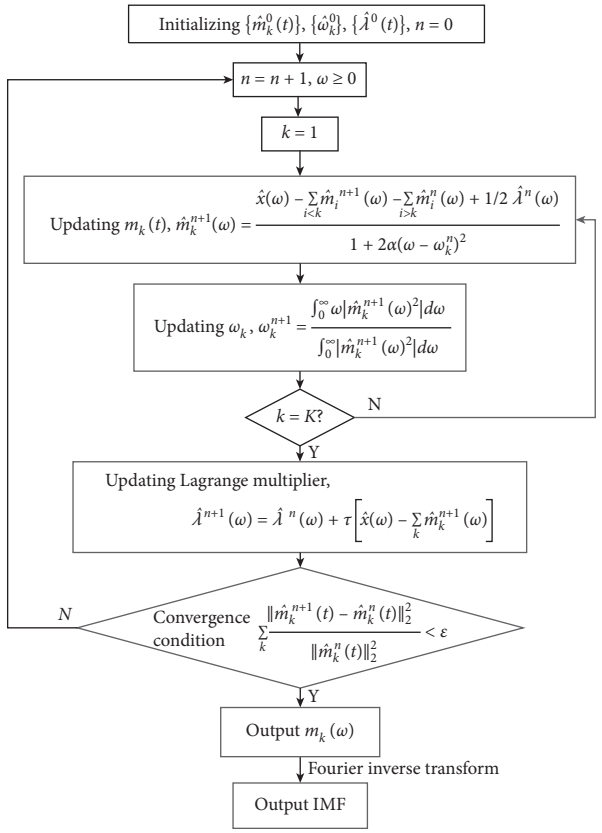


Figure 9: Flowchart of VMD algorithm.

Table 4: Central frequency of each IMF for 3# crack.

K	Central frequency of each IMF (Hz)						
	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
2	1182	2735	—	—	—	—	—
3	1182	1954	3457	—	—	—	—
4	390.8	1122	2725	3767	—	—	—
5	390.8	1182	1954	2846	4439	—	—
6	390.8	1182	1954	2735	3457	4439	—
7	390.8	1182	1954	2475	2846	3767	4499

In the case of 3# crack, the errors between the IMF central frequencies and their theoretical values are shown in Table 6.

As shown in Figure 13 and Table 6, the frequencies of each IMF correspond to the rotational frequency f_s of sun gear, the meshing frequency f_m, the fault frequencies f_{sg} and $n/4f_{sg}$ of sun gear, and the combined results of the above frequencies, respectively. The errors between the characteristic frequencies from dynamics simulation and theoretical calculation are less than 2%. It can be seen that the central frequencies of IMF components obtained by VMD can be used as the featured parameters identifying cracks in a planetary gear system.

5. Experimental Study

5.1. Test Setup. The test platform of a planetary gear system was established, as shown in Figure 14.

The test platform is composed of motor, frequency converter, planetary gearbox (one-stage transmission, deceleration ratio is 4.57 : 1), parallel-axle spur gearbox (two-stage transmission, deceleration ratio is 8.63 : 1), magnetic powder brake, power supply (adjustable current source), vibration sensor, signal collector, and upper computer.

For easy machining, three kinds of sun gear crack were made by Electrical Discharge Machining (EDM), and the

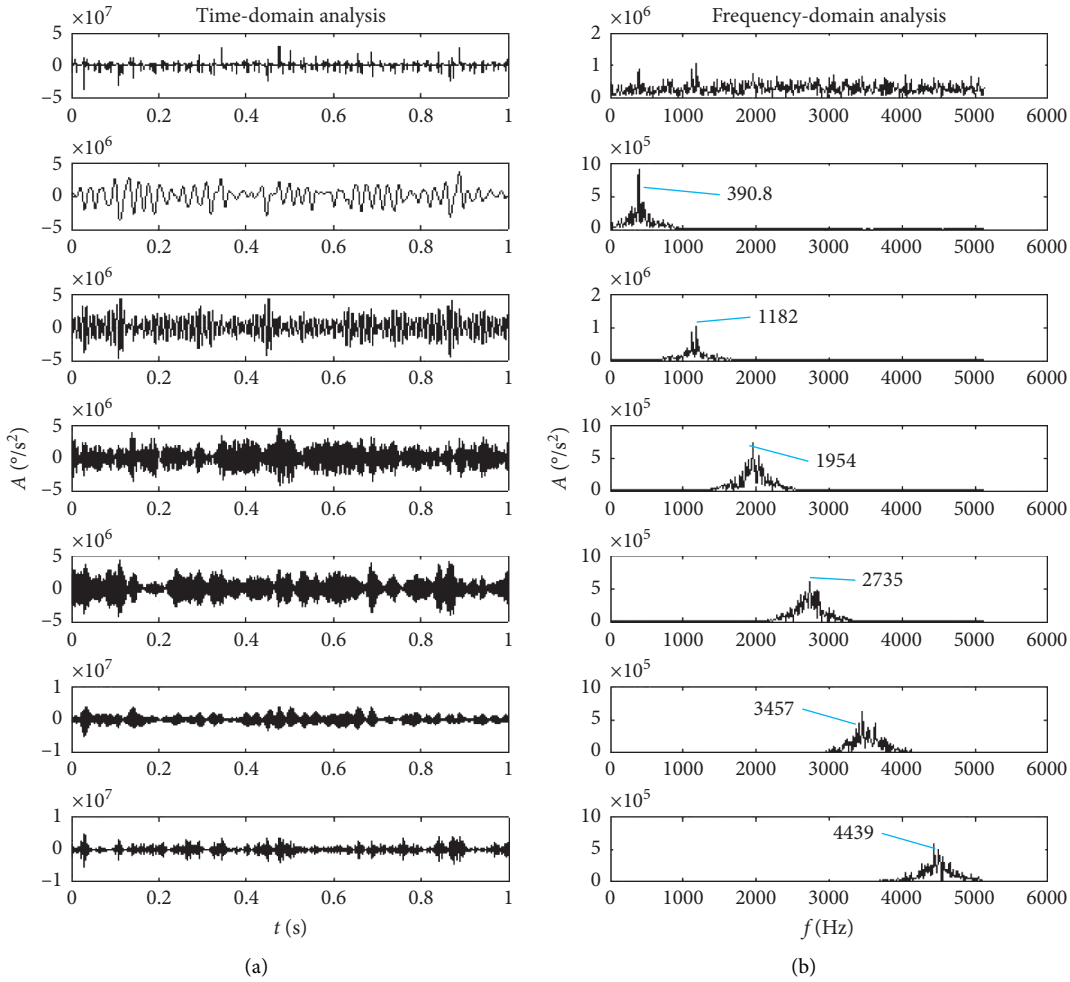


Figure 10: Each IMF of angular acceleration signal for 3# crack. (a) Time-domain waveform. (b) Frequency-domain spectrum.

actual "rectangle" is used instead of the ideal "trilateral" cracks shown in Figure 4 due to the limit of fabrication process, which does not affect the characteristic frequency extraction. Crack parameters are shown in Table 7.

Set the motor speed to 1800 rpm and the torque to 5 N·m. The 608A11-type accelerometer with a sensitivity of 0.103 V/g was used to measure the vertical vibration. The Spectral Quest data acquisition instrument was used, its sampling frequency is 10.24 kHz, and the sampled signal length is 6 s to 7 s.

6. Results and Discussion

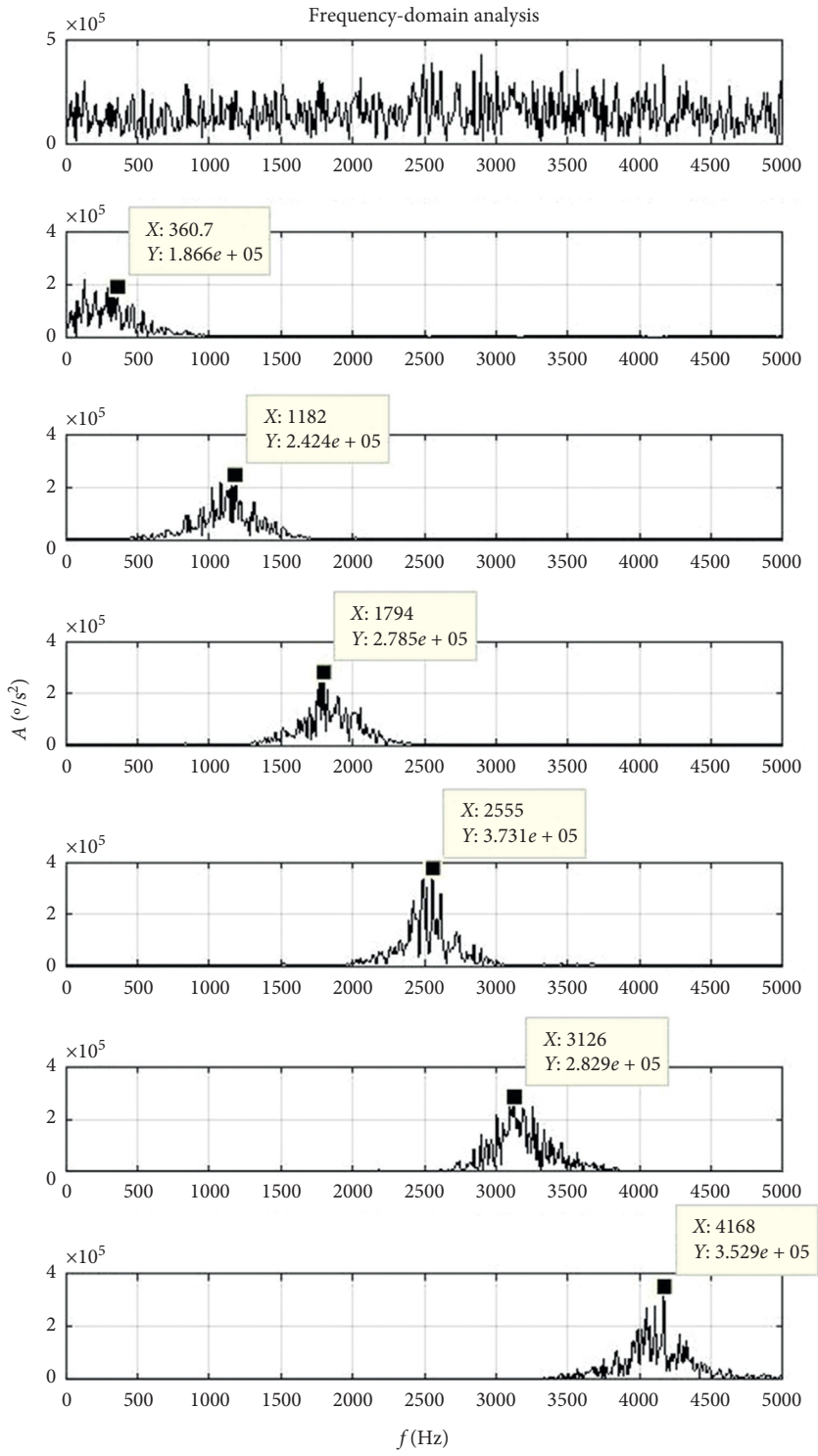
The vibration signals of sun gear under three kinds of actual crack were measured. The time-domain waveform and frequency-domain spectrum are shown in Figure 15.

The results of no crack and 1# to 3# cracks are listed from top to bottom. It can be seen that the amplitude of time-domain waveform increases with the increase of crack size. The peak value appears in the meshing frequency f_m and its multiple frequencies mf_m, local fault frequency f_{sg} of sun gear and its multiple frequencies nf_{sg}, $1/4f_{sg}$ and its multiple frequencies $n/4f_{sg}$, and combined frequencies $mf_m + nf_{sg}$ and $mf_m + n/f_{sg}$.

The VMD results of vibration signals under 1# to 3# cracks at $K = 6$ are shown in Figure 16; the central frequencies and amplitudes of each IMF are shown in Table 8.

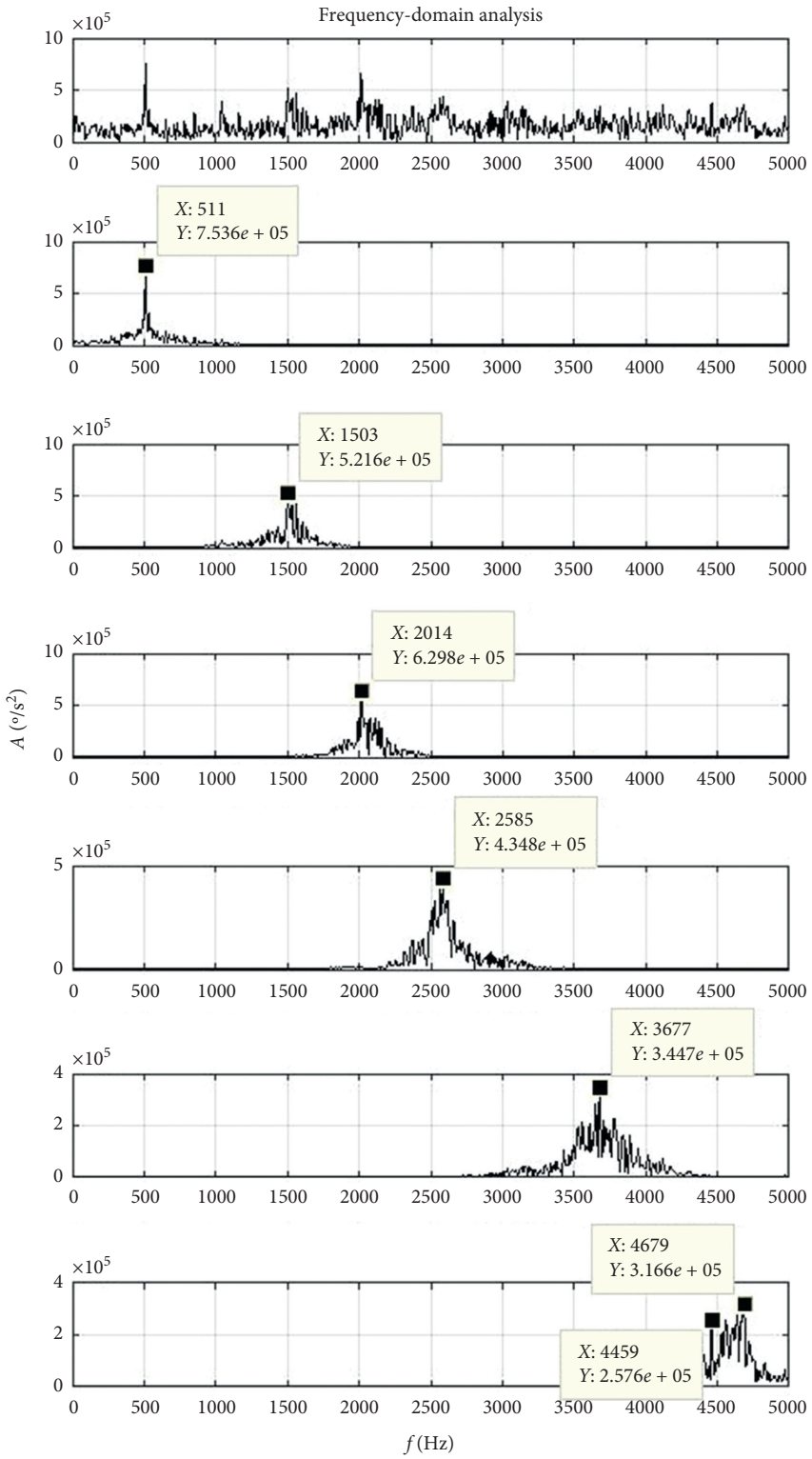
From Figure 17, it can be seen that with the increase of crack size, the total vibration energy increases, which proves that this experimental design is correct. However, due to the sun gear change and the complex gap distribution among the gear elements in an actual planetary gear system, these nonlinear factors inevitably affect the value of each IMF. The amplitudes corresponding to all the IMF central frequencies except the IMF2 of 1# to 3# cracks and the IMF3 of 1# to 2# cracks are larger than those with no crack.

In the case of 3# crack, the errors between the VMD and the theoretical frequency values of each IMF are shown in Table 9 and Figure 18. The frequencies of each IMF correspond to the rotational frequency fs of sun gear, the meshing frequency f_m, the fault frequency of sun gear f_{sg} and



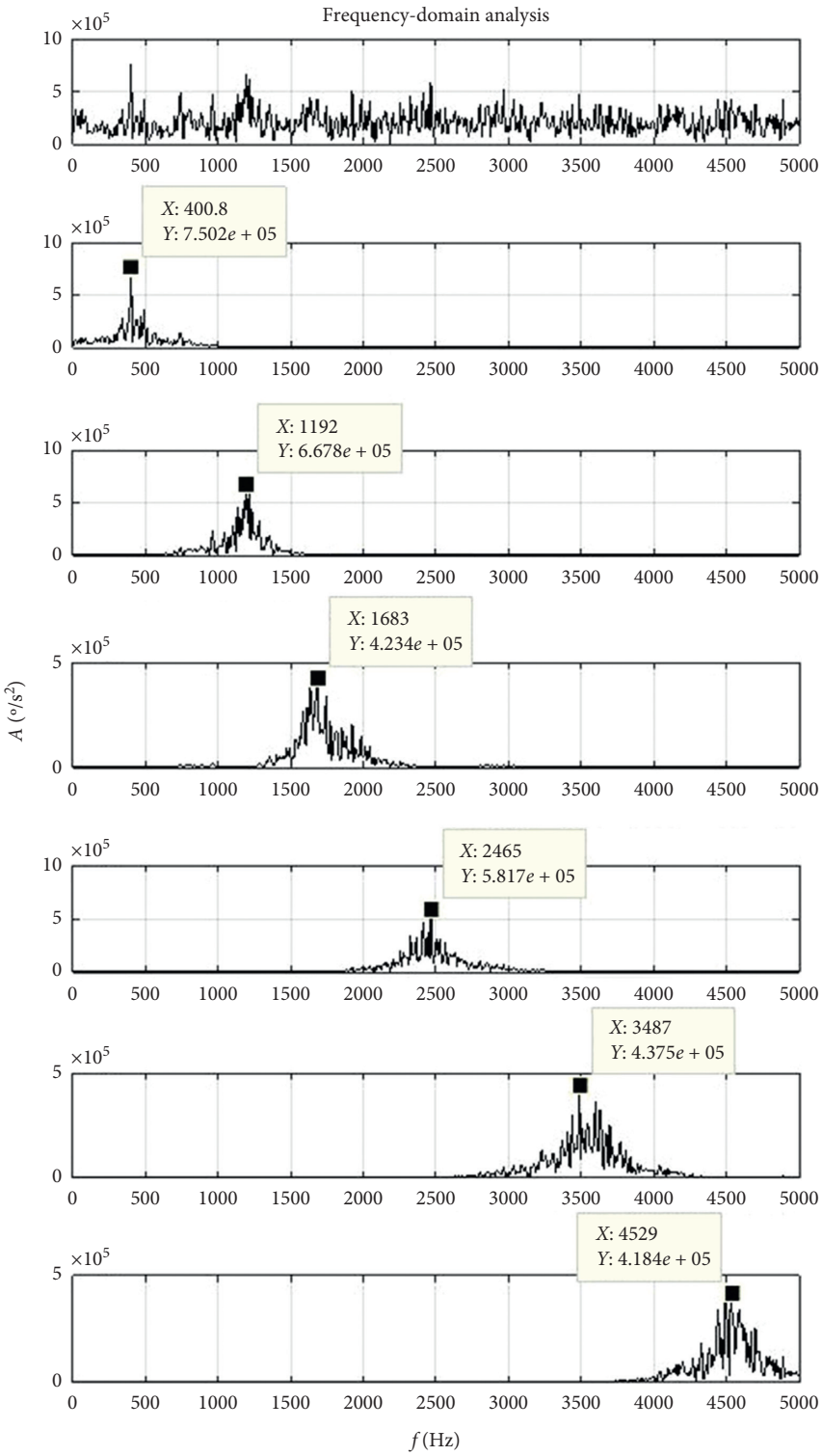
(a)

Figure 11: Continued.



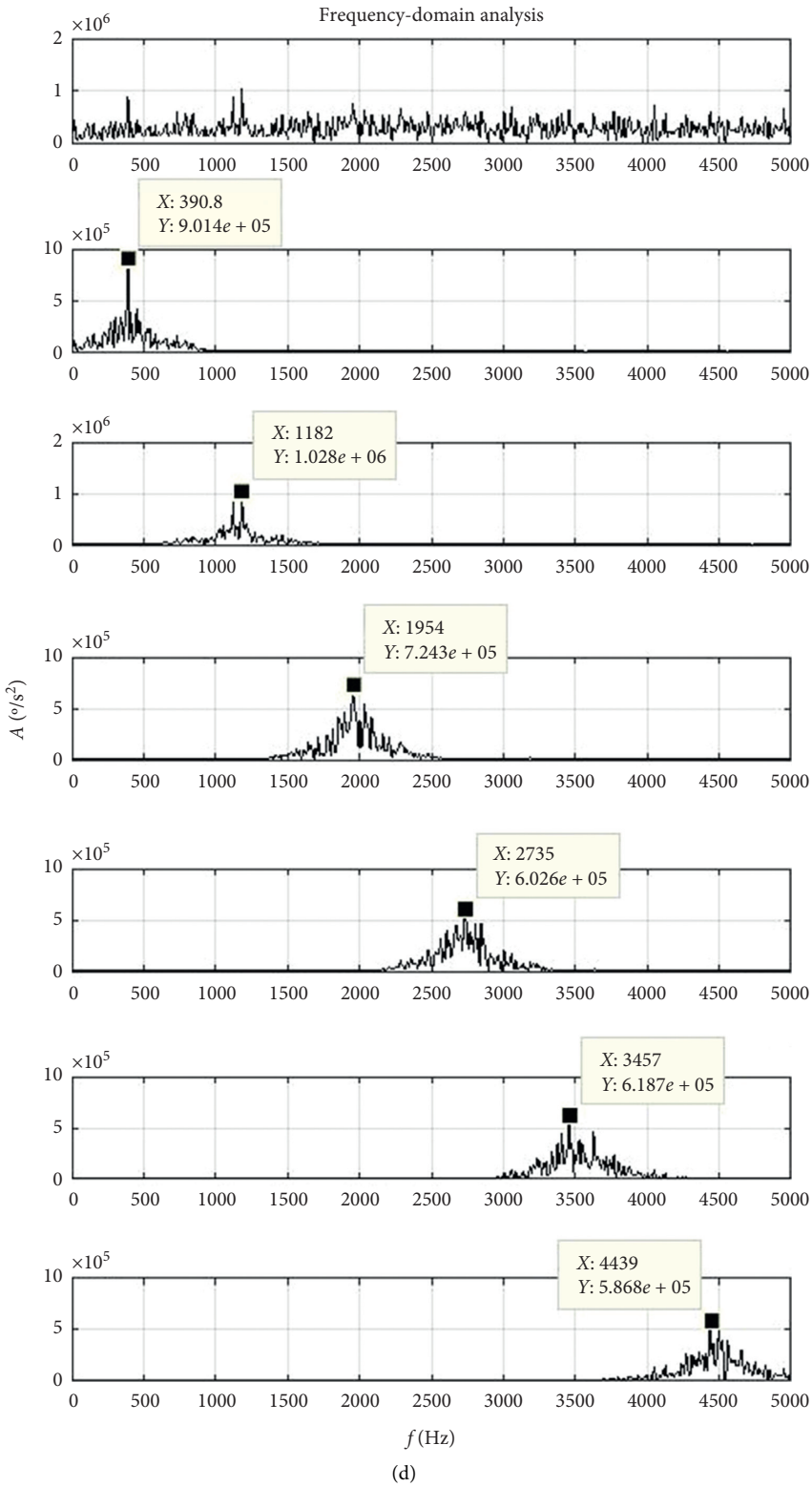
(b)

Figure 11: Continued.



(c)

Figure 11: Continued.



(d)

Figure 11: Each IMF of angular acceleration signal for no crack and 1# to 3# crack. (a) No crack. (b) 1# crack. (c) 2# crack. (d) 3# crack.

Table 5: Central frequency and amplitude of each IMF.

IMF	No crack		1# crack		2# crack		3# crack	
	f (Hz)	A (10^5)	f (Hz)	A (10^5)	f (Hz)	A (10^5)	f (Hz)	A (10^5)
IMF1	360.7	1.866	511	7.536	400.8	7.502	390.8	9.014
IMF2	1182	2.424	1503	5.216	1192	6.678	1182	10.280
IMF3	1794	2.785	2014	6.298	1683	4.234	1954	7.243
IMF4	2555	3.731	2585	4.348	2465	5.817	2735	6.026
IMF5	3126	2.829	3677	3.447	3487	4.375	3457	6.187
IMF6	4168	3.529	4679	3.166	4529	4.184	4439	5.868
Total	—	17.164	—	30.011	—	32.79	—	44.618

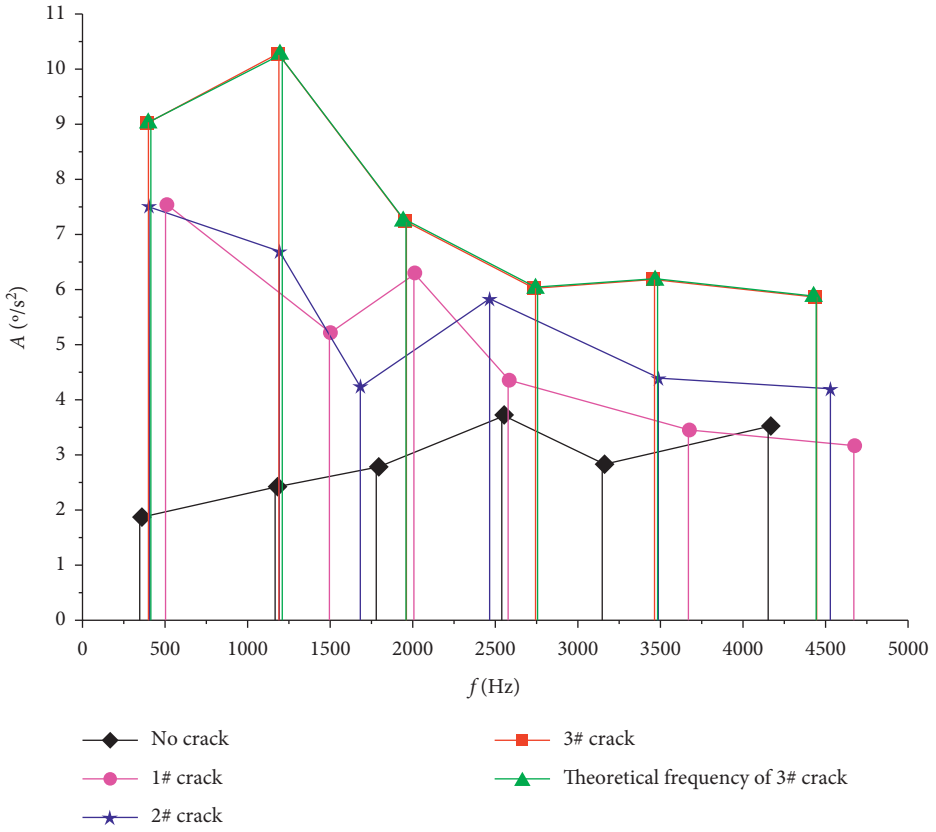


Figure 12: Central frequency and amplitude of each IMF.

Table 6: Error between the simulated VMD IMF central frequency and theoretical value for 3# crack.

IMF	VMD frequency (Hz)	Theoretical frequency (Hz)	Corresponding feature	Error (%)
IMF1	390.8	398.2	$f_m - 11/4f_{sg}$	1.86
IMF2	1182	1194.8	$2f_m - 5f_{sg}$	1.07
IMF3	1954	1944.6	$3f_m - 1/4f_{sg}$	0.48
IMF4	2735	2741.2	$4f_m + 5/4f_{sg}$	0.23
IMF5	3457	3467.8	$4f_m + 9f_{sg}$	0.31
IMF6	4439	4427.9	$7f_m - 7/4f_{sg}$	0.25

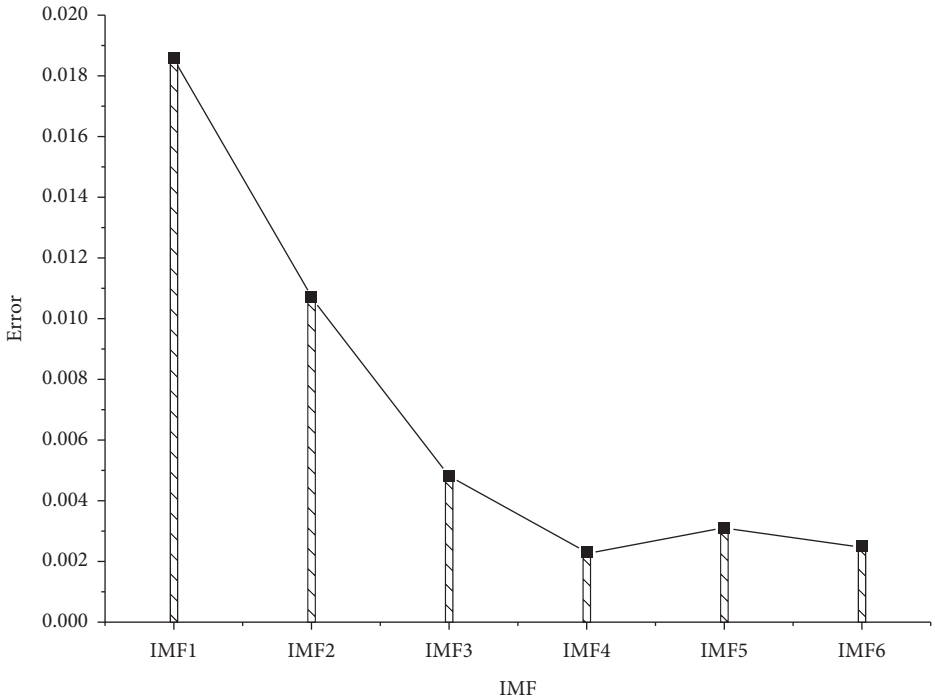


Figure 13: Error between the simulated VMD IMF central frequency and theoretical value for 3# crack.

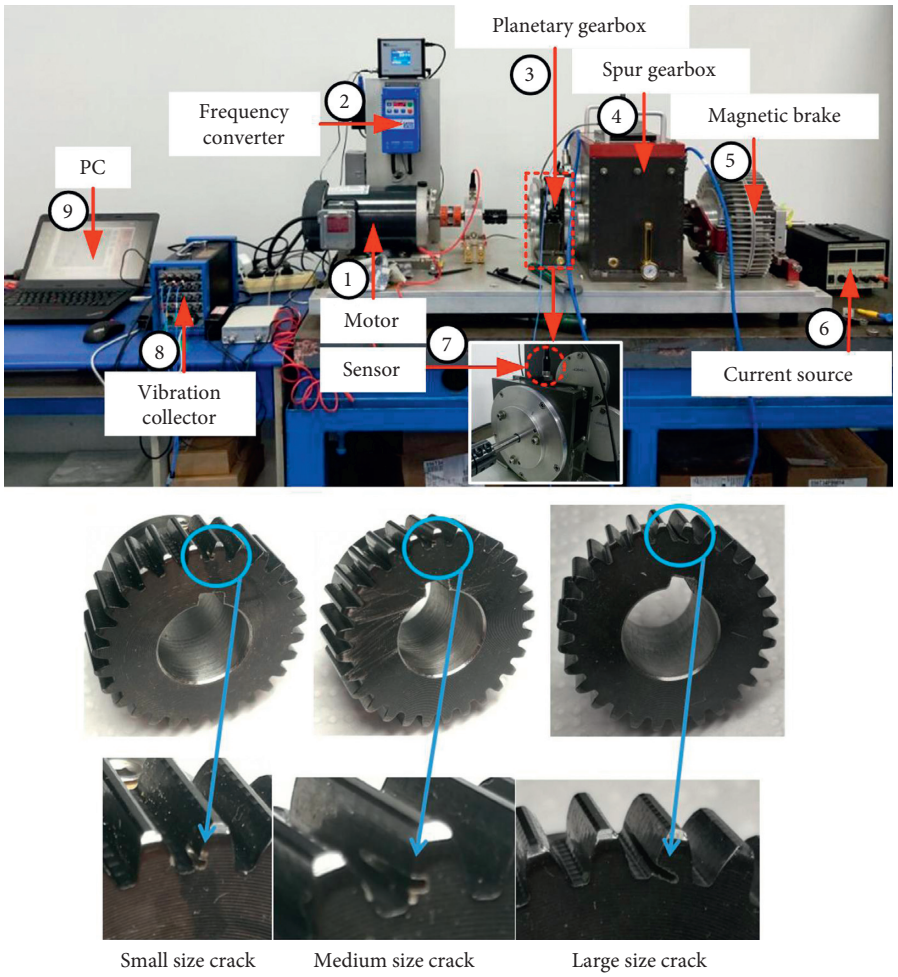


Figure 14: Test platform.

Table 7: Actual crack parameters of tested gear.

Crack number	1# crack	2# crack	3# crack
Crack length (mm)	1.43	3.92	8.47
Crack width (mm)	0.54	0.69	1.34
Crack height (mm)		0.50	

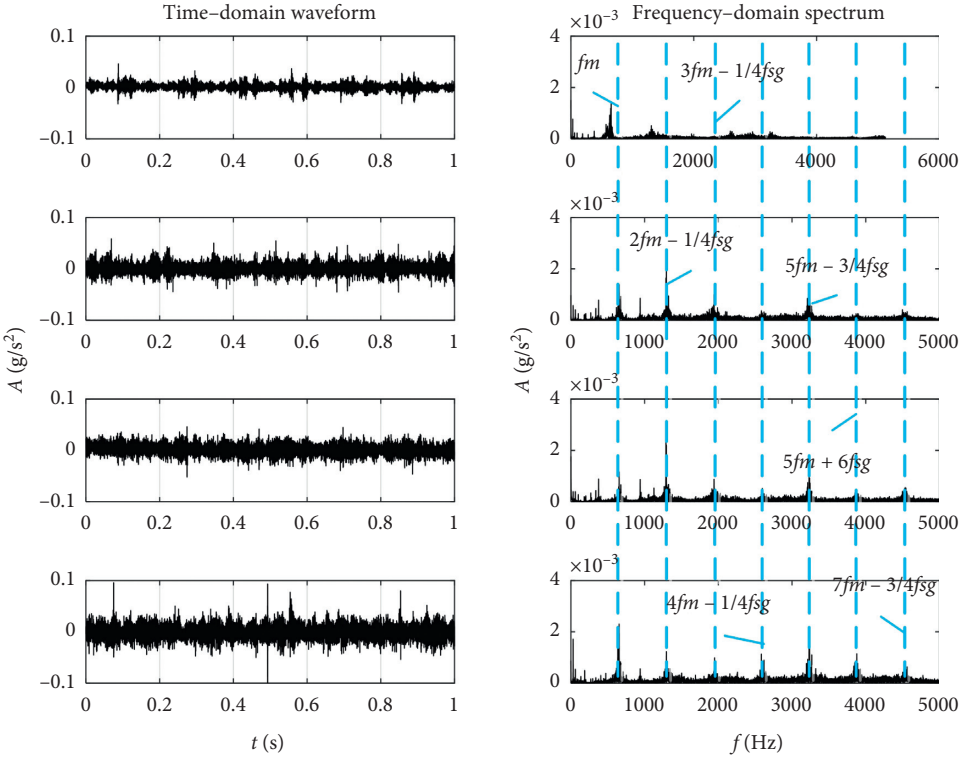


Figure 15: Time-domain waveform and frequency-domain spectrum for no crack and three crack cases.

$n/4f_{sg}$, and the combined results of the above frequencies, respectively. It can be seen that the errors between the VMD and the theoretical values are less than 1%.

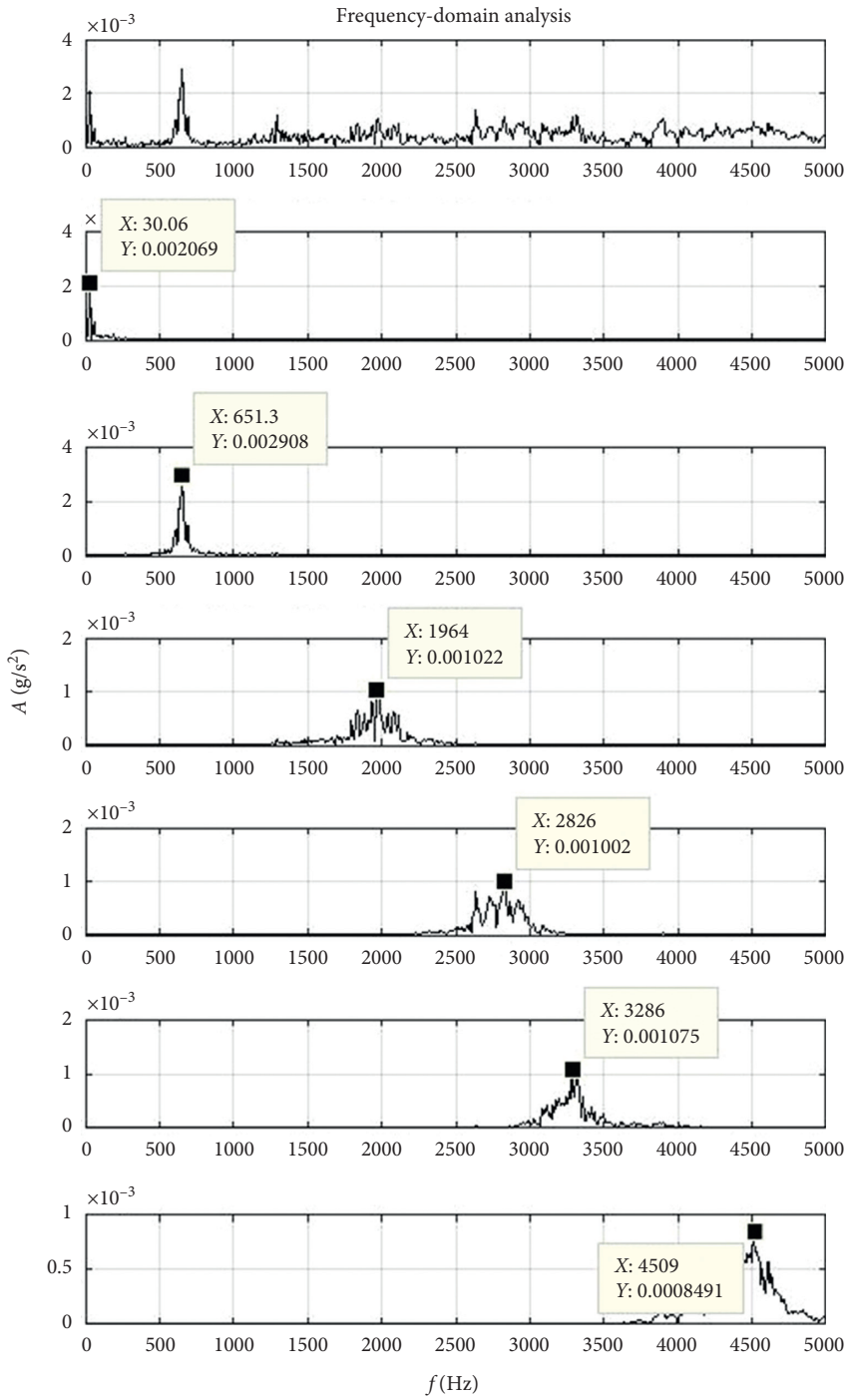
7. Conclusions

This paper proposes a tooth root crack identification method of sun gear in a planetary gear system combining the fault dynamics with VMD algorithm, which provides a positive contribution to the fault diagnosis of a complex mechanical transmission system.

(1) In order to overcome the shortcomings of the existing tooth crack models, a gear crack model with the simultaneous propagation in both length and width directions is proposed. And a rigid-flexible coupled model of planetary gear system is established by treating the sun gear as a flexible body and connecting it with other rigid components. The proposed model is appropriate for the fault frequency determination considering the computational accuracy and efficiency.

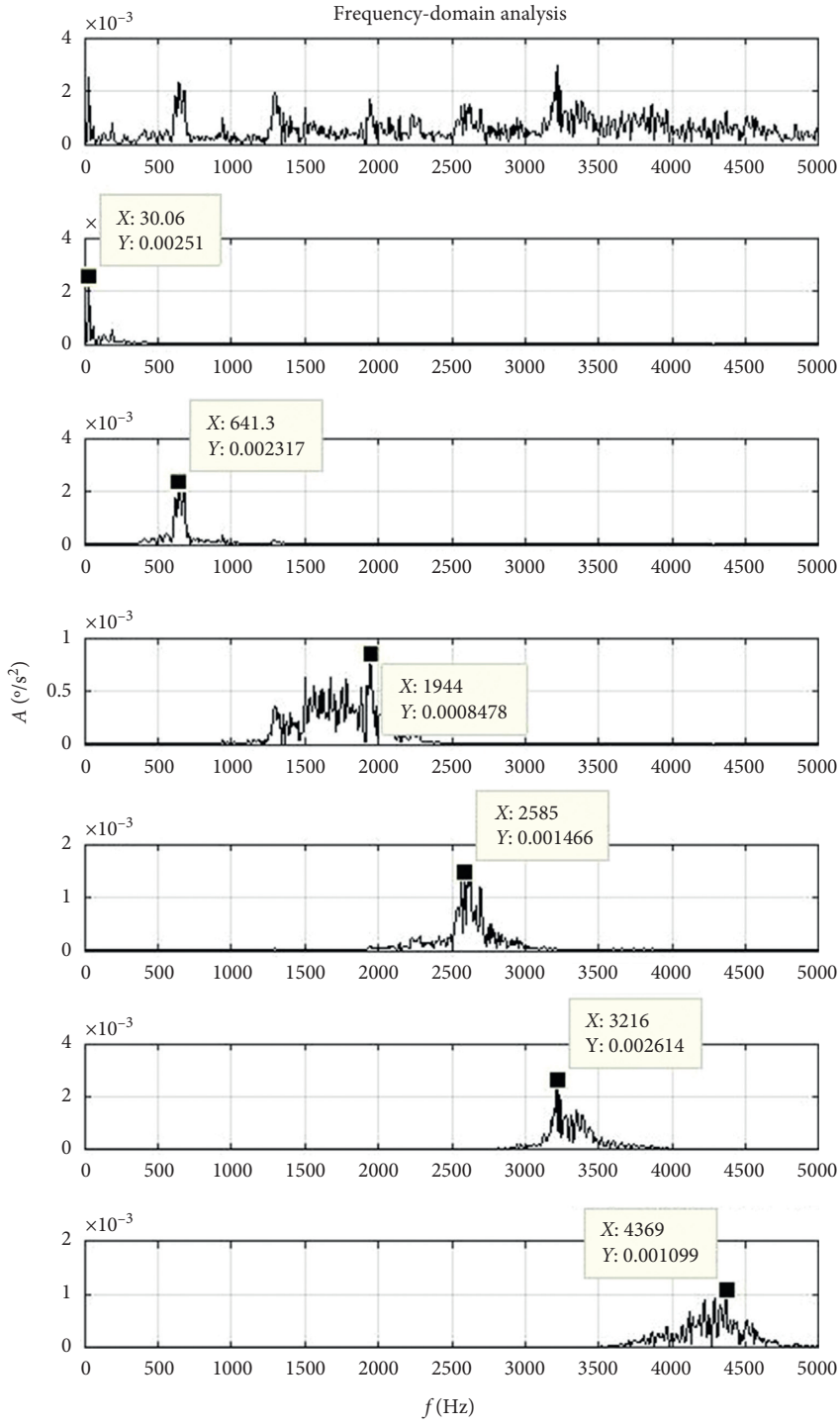
(2) In order to overcome the shortcomings of the existing EMD algorithms, the VMD method for the nonstationary weak signal is used. Through the analysis of simulated and tested signals, the total amplitude corresponding to each IMF central frequency obtained by VMD can reflect the level change of different crack scales, and the error between the tested VMD frequency and the theoretical value is less than 1%, which proves the proposed VMD algorithm effective.

(3) The test platform of a planetary gear system was established, and the tooth root crack of sun gear was made. The tested results show that the meshing



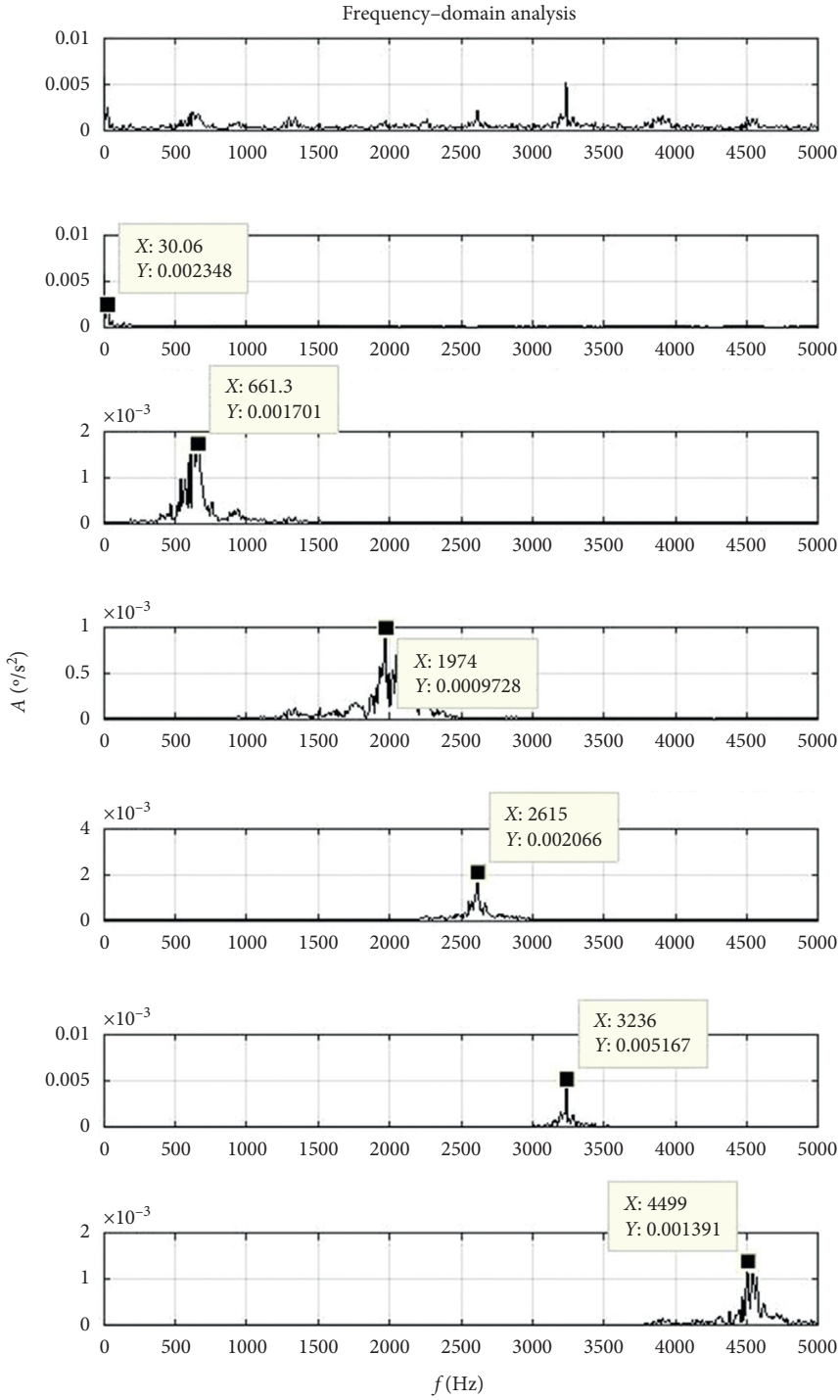
(a)

Figure 16: Continued.



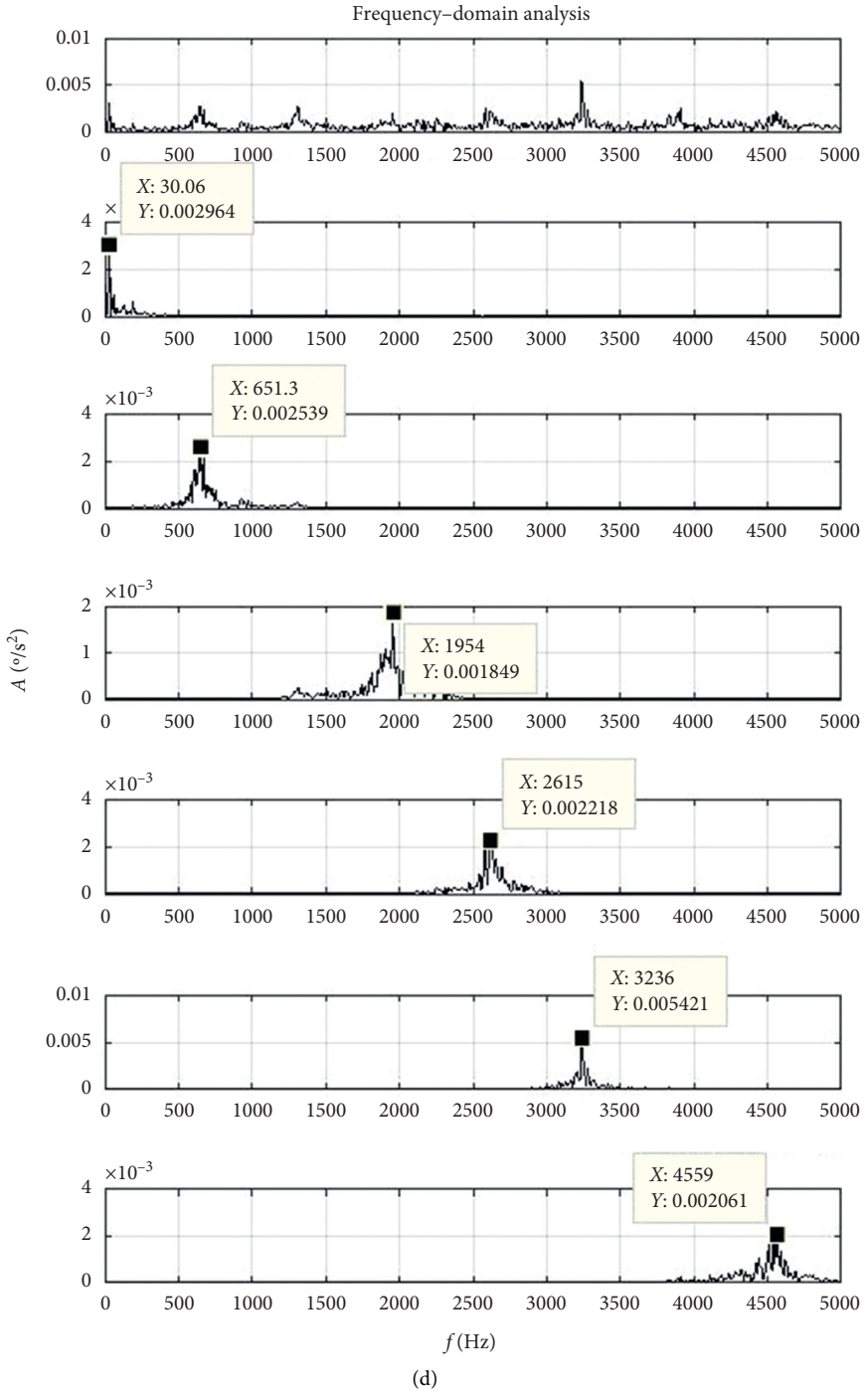
(b)

Figure 16: Continued.



(c)

Figure 16: Continued.



(d)

Figure 16: VMD results for no crack and three crack cases. (a) No crack. (b) 1# crack. (c) 2# crack. (d) 3# crack.

Table 8: Central frequency and amplitude of each IMF for 3# crack.

	No crack		1# crack		2# crack		3# crack	
	f (Hz)	A (10^{-3})	f (Hz)	A (10^{-3})	f (Hz)	A (10^{-3})	f (Hz)	A (10^{-3})
IMF1	30.06	2.07	30.06	2.51	30.06	2.35	30.06	2.96
IMF2	651.3	2.91	641.3	2.32	661.3	1.70	651.3	2.54
IMF3	1964	1.02	1944	0.85	1974	0.97	1954	1.85
IMF4	2826	1.00	2585	1.47	2615	2.07	2615	2.22
IMF5	3286	1.08	3216	2.61	3236	5.17	3236	5.42
IMF6	4509	0.85	4369	1.10	4499	1.39	4559	2.06
Total	—	8.93	—	10.86	—	13.65	—	17.05

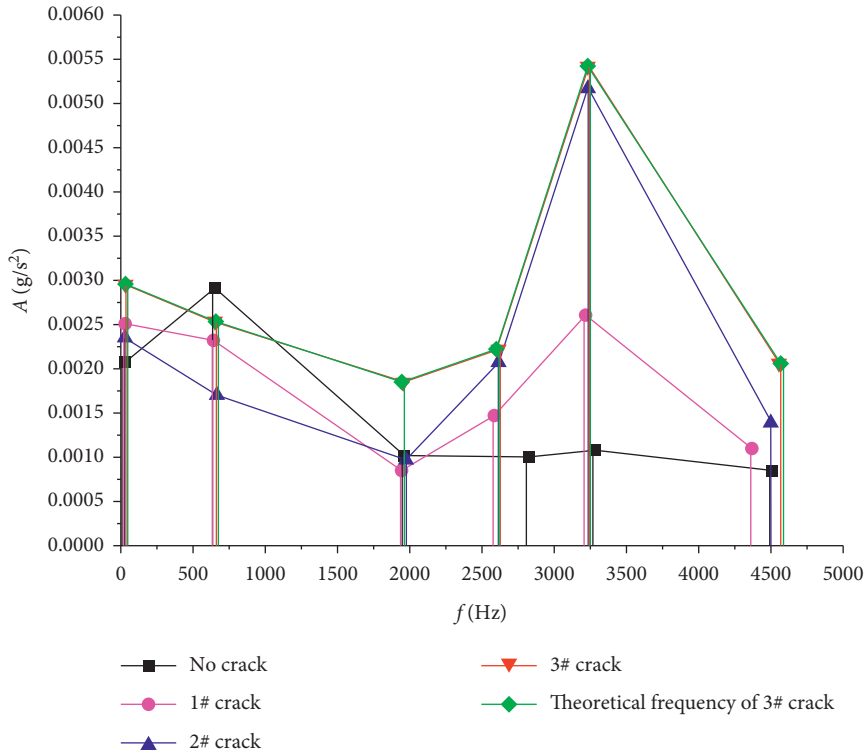


Figure 17: Central frequency and amplitude of each IMF.

Table 9: Error between the tested IMF central frequency and theoretical value for 3# crack.

IMF	Theoretical frequency (Hz)	VMD frequency (Hz)	Corresponding feature	Error (%)
IMF1	30	30.06	f_s	0.20
IMF2	656	651.3	f_m	0.72
IMF3	1944.6	1954	$3f_m-1/4f_{sg}$	0.48
IMF4	2600.6	2615	$4f_m-1/4f_{sg}$	0.55
IMF5	3233.1	3236	$5f_m-2/4f_{sg}$	0.09
IMF6	4568.6	4559	$7f_m-1/4f_{sg}$	0.21

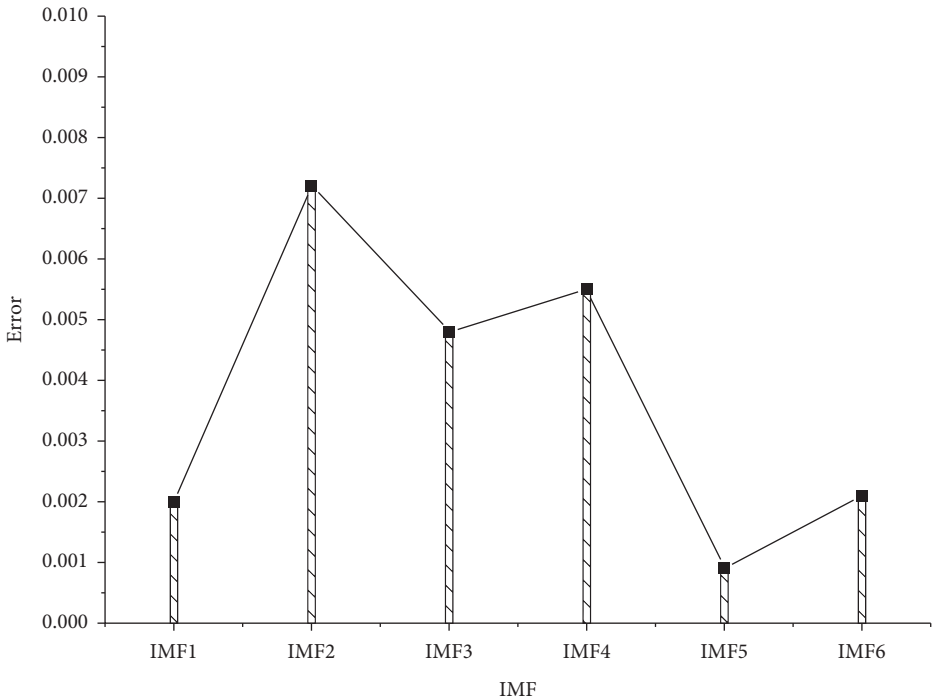


Figure 18: Error between the tested VMD IMF central frequency and theoretical value for 3# crack.

frequency, fault frequency, and their combined frequency are prominent in the amplitude spectrum of vibration signal with crack, so the gear crack model and rigid-flexible coupled system model established in this paper are feasible and the VMD method is also reliable, which provides an effective approach for the crack diagnosis of sun gear in a planetary gear system.

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 no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (no. 51834006), Natural Science Basic Research Program of Shaanxi Province in China (no. 2021JM-391), and Key R&D Project of Shaanxi Province in China (no. 2019GY-093).

References

[1] Y. Lei, "Research advances of fault diagnosis technique for planetary gearboxes," *Journal of Mechanical Engineering*, vol. 47, no. 19, pp. 59–67, 2011.

[2] O. D. Mohammed, M. Rantatalo, and U. Kumar, "Analytical crack propagation scenario for gear teeth and time-varying gear mesh stifmess," in *Procedings of the International Conference on Applied Mechanics ICAM*, pp. 1213–1218, Paris, France, August 2012.

[3] H. Ma, X. Pang, J. Zeng, Q. Wang, and B. Wen, "Effects of gear crack propagation paths on vibration responses of the perforated gear system," *Mechanical Systems and Signal Processing*, vol. 62-63, no. 2, pp. 113–128, 2015.

[4] F. Chaari, T. Fakhfakh, and M. Haddar, "Dynamic analysis of a planetary gear failure caused by tooth pitting and cracking," *Journal of Failure Analysis and Prevention*, vol. 6, no. 2, pp. 73–78, 2006.

[5] F. Chaari, T. Fakhfakh, and M. Haddar, "Analytical modelling of spur gear tooth crack and influence on gearmesh stiffness," *European Journal of Mechanics - A/Solids*, vol. 28, no. 3, pp. 461–468, 2009.

[6] Z. Chen and Y. Shao, "Dynamic simulation of spur gear with tooth root crack propagating along tooth width and crack depth," *Engineering Failure Analysis*, vol. 18, no. 8, pp. 2149–2164, 2011.

[7] Z. Chen and Y. Shao, "Mesh stiffness calculation of a spur gear pair with tooth profile modification and tooth root crack," *Mechanism and Machine Theory*, vol. 62, no. 4, pp. 63–74, 2013.

[8] H. Ma, R. Song, X. Pang, and B. Wen, "Time-varying mesh stiffness calculation of cracked spur gears," *Engineering Failure Analysis*, vol. 44, no. 5, pp. 179–194, 2014.

[9] C. Liu, D. Qin, T. C. Lim, and Y. Liao, "Dynamic characteristics of the herringbone planetary gear set during the variable speed process," *Journal of Sound and Vibration*, vol. 333, no. 24, pp. 6498–6515, 2014.

[10] G. He, K. Ding, W. Li, and Y. Li, "Frequency response model and mechanism for wind turbine planetary gear train vibration analysis," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 425–432, 2017.

[11] Z. Chen, Z. Zhu, and Y. Shao, "Fault feature analysis of planetary gear system with tooth root crack and flexible ring gear rim," *Engineering Failure Analysis*, vol. 49, pp. 92–103, 2015.

[12] J. Helsen, F. Vanhollebeke, B. Marrant, D. Vandepitte, and W. Desmet, "Multibody modelling of varying complexity for modal behaviour analysis of wind turbine gearboxes," *Renewable Energy*, vol. 36, no. 11, pp. 3098–3113, 2011.

[13] X. Jin, L. Li, W. Ju, Z. Zhang, and X. Yang, "Multibody modeling of varying complexity for dynamic analysis of large-scale wind turbines," *Renewable Energy*, vol. 90, pp. 336–351, 2016.

[14] H. Zhai, C. Zhu, C. Song, H. Liu, and H. Bai, "Influences of carrier assembly errors on the dynamic characteristics for wind turbine gearbox," *Mechanism and Machine Theory*, vol. 103, pp. 138–147, 2016.

[15] F. Jiang, K. Ding, G. He, Y. Sun, and L. Wang, "Vibration fault features of planetary gear train with cracks under time-varying flexible transfer functions," *Mechanism and Machine Theory*, vol. 158, Article ID 104237, 2021.

[16] Z. Shen, B. Qiao, L. Yang, W. Luo, Z. Yang, and X. Chen, "Fault mechanism and dynamic modeling of planetary gear with gear wear," *Mechanism and Machine Theory*, vol. 155, 2021.

[17] X. Liu, Y. Yang, and J. Zhang, "Effects of tooth-crack-induced mesh stiffness on fault signals of a planetary gear train," *Procedia Computer Science*, vol. 109, pp. 785–792, 2017.

[18] J. Liu, Y. Xu, Y. Shao, H. Xiao, and H. Li, "The effect of a localized fault in the planet bearing on vibrations of a planetary gear set," *Journal of Strain Analysis for Engineering Design*, vol. 53, no. 5, pp. 313–323, 2018.

[19] J. Liu, R. Pang, S. Ding, and X. Li, "Vibration analysis of a planetary gear with the flexible ring and planet bearing fault," *Measurement*, vol. 165, Article ID 108100, 2020.

[20] T. Duan, J. Wei, A. Zhang, Z. Xu, and T. C. Lim, "Transmission error investigation of gearbox using rigid-flexible coupling dynamic model: Theoretical analysis and experiments," *Mechanism and Machine Theory*, vol. 157, Article ID 104213, 2021.

[21] Z. Fan, C. Zhu, and C. Song, "Dynamic analysis of planetary gear transmission system considering the flexibility of internal ring gear," *Iranian Journal of Science and Technology Transactions of Mechanical Engineering*, vol. 44, no. 3, pp. 695–706, 2020.

[22] L. Cohen, *Time-frequency Analysis theory and Applications*, Prentice-Hall Inc., Hoboken, NJ, USA, 1995.

[23] Z. Feng, M. Liang, and F. Chu, "Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples," *Mechanical Systems and Signal Processing*, vol. 38, no. 1, pp. 165–205, 2013.

[24] K. Hadad, M. Pourahmadi, and H. Majidi-Maraghi, "Fault diagnosis and classification based on wavelet transform and neural network," *Progress in Nuclear Energy*, vol. 53, no. 1, pp. 41–47, 2011.

[25] N. Li, R. Zhou, Q. Hu, and X. Liu, "Mechanical fault diagnosis based on redundant second generation wavelet packet

transform, neighborhood rough set and support vector machine," *Mechanical Systems and Signal Processing*, vol. 28, pp. 608–621, 2012.

[26] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[27] X. Xiong, S. Yang, and C. Gan, "A new procedure for extracting fault feature of multi-frequency signal from rotating machinery," *Mechanical Systems and Signal Processing*, vol. 32, pp. 306–319, 2012.

[28] 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 of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[29] G. F. Wang, P. Deng, F. Q. Zhang et al., "Noise reduction of the transient electromagnetic weak signal under strong noise based on power detection of EMD," *Applied Mechanics and Materials*, vol. 110, pp. 1606–1612, 2012.

[30] J. Yan and L. Lu, "Improved Hilbert-Huang transform based weak signal detection methodology and its application on incipient fault diagnosis and ECG signal analysis," *Signal Processing*, vol. 98, pp. 74–87, 2014.

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

[32] R. Ma, Y. Chen, and H. Sun, "Detection of weak signals based on empirical mode decomposition and singular spectrum analysis," *IET Signal Processing*, vol. 7, no. 4, pp. 269–276, 2013.

[33] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[34] Z. N. Li and M. Zhu, "Research on mechanical fault diagnosis method based on variational mode decomposition," *Acta Armamentarii*, vol. 38, no. 3, pp. 593–599, 2017.

[35] H. M. Wang, H. H. Liu, and P. Xue, "Application of variational mode decomposition in fault feature extraction of gears," *Journal of Northeast Forestry University*, vol. 47, no. 8, pp. 97–102, 2019.

[36] Z. P. Feng, L. L. Zhao, and F. L. Chu, "Vibration spectral characteristics of localized gear fault of planetary gearboxes," *Proceedings of the CSEE*, vol. 33, no. 5, pp. 119–127, 2013.

[37] P. F. Zhi, L. Z. Fu, and J. Z. Ming, *Vibration Fault Diagnosis Method of Planetary Gearbox*, Science Press, Beijing, China, 2015.

[38] J. Kramberger, M. Sraml, S. Glodez, J. Flašker, and I. Potrč, "Computations & structures model for the analysis of bending fatigue in gears," *Computations & Structures*, vol. 82, no. 23-26, pp. 2261-2262, 2004.

[39] C. Cheng, *Dynamic Modeling of Two-Stage Planetary Gearbox with Tooth Cracks and its Dynamic Response Analyses*, Chengdu University of Electronic Science and Technology of China, Chengdu, China, 2017.

[40] P. F. Zhi, L. Z. Fu, and J. Z. Ming, *Principle and Fault Diagnosis Application of Complex Non-stationary Signal Analysis Method for Mechanical System*, Science Press, Beijing, China, 2018.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6650932, 12 pages
https://doi.org/10.1155/2021/6650932

Research Article

Research on Feature Extraction Method of Engine Misfire Fault Based on Signal Sparse Decomposition

Canyi Du,[1] Fei Jiang ,[2] Kang Ding,[2] Feng Li,[1] and Feifei Yu[1]

[1]*School of Automobile and Transportation Engineering, Guangdong Polytechnic Normal University, Guangzhou 510450, China*
[2]*School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China*

Correspondence should be addressed to Fei Jiang; 201910100398@mail.scut.edu.cn

Received 22 December 2020; Revised 9 March 2021; Accepted 17 March 2021; Published 25 March 2021

Academic Editor: Jinde Zheng

Engine vibration signals are easy to be interfered by other noise, causing feature signals that represent its operating status get submerged and further leading to difficulty in engine fault diagnosis. In addition, most of the signals utilized to verify the extraction method are derived from numerical simulation, which are far away from the real engine signals. To address these problems, this paper combines the priority of signal sparse decomposition and engine finite element model to research a novel feature extraction method for engine misfire diagnosis. Firstly, in order to highlight resonance regions related with impact features, the vibration signal is performed with a high-pass filter process. Secondly, the dictionary with clear physical meaning is constructed by the unit impulse function, whose parameters are associated with engine system modal characteristics. Afterwards, the signals that indicate the engine operating status are accurately reconstructed by segmental matching pursuit. Finally, a series of precise simulation signals originated from the engine dynamic finite element model, and experimental signals on the automotive engine are used to verify the proposed method's effectiveness and antinoise performance. Additionally, comparisons with wavelet decomposition further show the proposed method to be more reliable in engine misfire diagnosis.

1. Introduction

Because of its simple structure, output power stability, and convenient installation, the engine is widely utilized as the power output device in many mechanical systems, such as automotive, steamship, and spacecraft [1]. However, after the long-term harsh operation condition, the engine and other transmission system are prone to fault which will cause a sudden power interruption, thus affecting the reliability of the whole transmission system [2–4]. Furthermore, the vibration signals acquired from the sensor contain not only engine operating information but also various other strong noise signals, which causes useful feature signals to be submerged and makes it hard to diagnose engine fault. Therefore, it is of great significance to extract engine feature signals that indicate its health state and further monitor their operation status timely.

In terms of fault diagnosis in engine condition monitoring, many researchers have proposed various effective methods ranging from traditional signal processing to artificial intelligent algorithm. On the aspect of traditional signal processing methods, Bi et al. [5] proposed a novel fault diagnosis method of diesel engine valve clearance using the improved variational mode decomposition (VMD) and bispectrum algorithm. Vernekar et al. [6] applied empirical mode decomposition (EMD) to decompose the original vibration into a finite number of intrinsic mode functions and then used the Naïve Bayes algorithm as the classifier to detect the engine fault. Figlus et al. [7] applied a wavelet decomposition, while filtering the internal combustion engine's acoustic signal in order to diagnose an excessive valve clearance. On the aspect of diagnosis methods based on the artificial intelligent algorithm, it has been greatly developed and attracted more and more scholar's attention [8–11]. Deng et al. [8] proposed an adaptive method for choosing the parameters in the deep belief network (DBN) based on the improved quantum-inspired differential evolution algorithm, which could obtain higher classification

accuracy on detecting rolling bearing fault. Du et al. [9] directly used the vibration signal as the input of the probabilistic neural network to simplify the diagnosis process and combined the swarm intelligent algorithm to achieve efficient parallel search for the best diagnosis effect. Jafarian et al. [10] employed various intelligent diagnosis methods, such as the Artificial Neural Networks (ANN) and Support Vector Machines (SVM), to monitor the engine health state and highlighted their superiorities. Tao et al. [11] proposed a novel extreme gradient boosting-based misfire fault diagnosis approach utilizing the high-accuracy time frequency information of vibration signals and achieved higher diagnosis accuracy. Although the above methods could achieve certain effects, most methods are easily influenced by other noises and cannot obtain good results under the low signal-to-noise ratio (SNR).

Recently, due to its superiority of feature extraction ability under low SNR, sparse decomposition is widely applied in the field of image processing [12], compressed sensing [13], and fault diagnosis [14, 15]. Engine vibration signals have similar characteristics with these signals. Therefore, the sparse decomposition method can be theoretically transferred to the engine signal analysis and employed to extract its features for accurate diagnosis.

Additionally, like it is described in Reference [16], most of the reference simply used numerical simulation, which usually consists of sine and modulation components and noise part from the environment, to simulate the engine vibration signal and further to verify the proposed feature extraction method. However, the numerical simulation signal still has huge differences with the real-world complex engine signal. The algorithms validated by it may face trouble in handling the real engine vibration signal. Hence, to narrow the gap between the simulation signal and real signal as much as possible, a more concise dynamic finite element model should be constructed to generate the simulation signal rather than simply employ numerical simulation.

To address these problems, the proposed method combines the priority of signal sparse decomposition and engine finite element model to research a novel feature extraction method for engine misfire diagnosis in this paper. The organization of the rest paper is arranged as follows. Section 2 introduces the fundamental principle of the proposed method; Section 3 briefly illustrates the construction steps of the engine model and verifies the proposed feature extraction method; Section 4 is experimental verifications and comparison with wavelet analysis; Section 5 comes to conclusions.

2. Feature Extraction of Engine Misfire Fault

In the actual operating condition, the vibration signal collected from the hood is complex always contains unrelated signals generated by other components, which easily makes the feature signal represent the status of the engine submerged in strong noise. Additionally, different from other types of engine fault, misfire fault occurs with higher probability and its signal feature is relatively weak [17]. In order to detect engine misfire fault timely, it is necessary to find a method that can effectively extract its fault features under low SNR circumstances. Relying on its excellent feature extraction capabilities, the signal sparse decomposition method is therefore selected in this paper.

2.1. Basic Principle of MP. Matching pursuit (MP) is a sparse algorithm which utilizes linear combinations of atoms in the redundant dictionary to approximately represent a signal [14]. Due to its powerful signal sparse representation ability, MP is now widely applied in feature signal reconstruction to better detect imperceptible fault. Theoretically, the original signal x can be decomposed into a linear superposition of atoms $\mathbf{d}$ ($\|\mathbf{d}= 1\|$) in the dictionary $\mathbf{D} \in R^{n\times q}$, as shown in the following equation:

$$x = \left|\langle x, \mathbf{d}_0\rangle\right|\mathbf{d}_0 + R_1 x, \tag{1}$$

where $\mathbf{d}_0$ denotes the best matching atom in this procedure and $R_1 x$ is the residual signal after first MP. Because of the orthogonality between $\mathbf{d}_0$ and $R_1 x$, $R_1 x$ can further decompose in the same way as x. After applying N times to the residual signal, x can be represented as follows:

$$x = \sum_{n=0}^{N-1} \left|\langle R_n x, \mathbf{d}_n\rangle\right|\mathbf{d}_n + R_N x, \tag{2}$$

In order to avoid falling into the infinite loop, the iterative procedure stops until $R_N x$ is less than a preset threshold δ, which is a small value. From the above introduction, it can be known that the extraction accuracy of the feature signal mainly depends on the atoms in the dictionary. The more similar atoms with the feature signal, the better feature extraction result can be obtained.

2.2. Construction of the Dictionary with Physical Meaning. As we know, the unit response function has an attenuated oscillation waveform. The feature signals of engine misfire fault are continuous impact signals, which have a similar waveform with the unit response function in the time domain. Moreover, according to the principle of mechanical vibration, engine system's modes that are determined by modal parameters (f_d, ξ) would be effectively excited by series of impulse forces generated from the cylinder. Furthermore, its parameters could influence structure physical vibration characteristics. The natural frequency f_d and damping ratio ξ determine the speed of vibration and attenuation in the time domain, respectively. And, these parameters are also connected with the physical structure of the engine system. Therefore, according to the similarity principle between the original signal and atom, it is reasonable to choose the unit response function as an atom in the sparse dictionary [14]. Its specific expression is presented as follows:

$$d(t) = \exp\left(\frac{-2\pi\xi}{\sqrt{1-\xi^2}} f_d t\right) \sin 2\pi f_d t, \tag{3}$$

where f_d and ξ are the natural frequency and damping ratio, respectively. Equation (3) indicates that the atom is associated with the modal parameters of the engine again, which not only has high similarity with the original signal but also contains clear physical meaning.

2.3. Detailed Procedures of the Proposed Method. Inspired by MP and the engine vibration signal characteristics, the proposed method can be illustrated as shown in Figure 1. It mainly contains three parts: signal preprocessing, dictionary construction, and feature extraction. Due to the low-frequency noise caused by other machinery rotational components in the engine transmission system, the first part is utilized to improve the SNR of the obtained signal. The second is to construct a dictionary with clear physical meaning. Finally, the third is to apply MP to extract the engine feature signal for diagnosis. The detail procedures of the algorithm are implemented as follows:

(1) The signal $x(t)$ is obtained with a sampling frequency f_s from the vibration acceleration sensor installed in the hood. Afterwards, a high-pass filtering process is applied to the collected signal to remove the low-frequency noise. This procedure can further improve the SNR of the signal and minimize interferences from other irrelevant signals as much as possible.

(2) According to the theory proposed in Reference [15], the multiorder frequencies and damping ratios $(f_{dj}, \xi_j)(j = 1, \ldots, J)$ are obtained by edited cepstrum and rational fractional polynomial fitting method. Therefore, the atoms in the dictionary are constructed by equation (3). In order to expand the atoms and achieve the redundancy property of the dictionary, the time shift factor $\tau \in (0{:}\ \Delta t{:}\ T_s)$ is introduced to find out impact time of the signal. At last, the dictionary $\mathbf{D}$ consists of all the expanded atoms, with the shape of $p \times q$. $p = J \times \text{round}((T_s + 1)/\Delta t)$ and $q = T_s f_s$, where J is the maximum order of extracted modal parameters, Δt represents the time shift interval, and T_s is the ignition impact interval, and the notion of *round* presents for rounding off.

(3) To accelerate the calculation speed of MP, the signal is performed in N segmental. Each segment of the filtered signal $x_p(t)$ is divided by a duration of T_s, which is corresponding to the ignition impact interval, namely, the reciprocal of main exciting frequency in the automotive engine. $T_s = (1/f_{\text{main}}) = ((60 \times 2)/nA)$, where n is the rotation speed of the engine and A is the amount of the engine cylinder.

(4) Set the iteration value δ of the MP algorithm, and apply it to solve sparse coefficients of the n^{th} segment signal. Hence, the feature signal of the n^{th} segment signal can be represented by $x_{fn}(t)$.

(5) Based on the dividing time sequential of the signal, the reconstructed feature signal of the engine can be expressed by equation (4). It should be mentioned that it is a quick algorithm to reconstruct the engine feature signal, and the times of inner product equals to round$(T/T_s)Np$, where N is the maximal times of MP:

$$x_f(t) = \sum_{n=1}^{N} x_{fn}(t). \tag{4}$$

(6) Finally, the feature of engine operating status could be represented by the reconstructed signal $x_f(t)$ and the misfire fault diagnosis is carried out.

3. Simulation Analysis

To verify the effectiveness of the proposed method, a precise engine finite element model is firstly applied to generate more realistic engine simulation signals. Then, white noises are added into the original signals to research its feature extraction ability under the low SNR condition. Comparison analysis illustrates the advantages of the proposed method in terms of feature extraction ability.

3.1. Generation of the Engine Simulation Signals. Different from other fault signals, such as the bearing fault signal or broken gear signal, the engine signals are more complex whether under the health state or misfire fault state. Therefore, in order to generate more realistic engine signals for analysis, a high-quality four-cylinder engine multibody dynamic model is constructed by utilizing the AVL EXCITE platform [18]. Its main procedures are as follows:

(1) Build the 3-dimensional model of the four-cylinder engine power unit and crankshaft in CAD software, and import it into MSC Patranto to perform finite element mesh division, as shown in Figure 2. This step is to set the engine as a flexible body and make each part of the vibration signal obtainable, especially the hood part.

(2) As we know, the dynamic system analysis with huge degrees of freedom is time consuming. Hence, the matrix reduction method is accomplished by MSC Nastran to reduce the degrees of freedom of the constructed model and then improve the calculation efficiency.

(3) Import the reduced finite element model into the EXCITE power unit, and each component can be connected by springs, dampers, bearings, and other nonlinear elements to form the final coupled engine dynamic model.

(4) Set combustion pressure in the EXCITE PU module as the external excitation to make the whole system vibrate. Its variation law along with the crank angle under the normal state is illustrated in Figure 3(a), which is similar to actual engine working conditions. Meanwhile, the engine misfire fault is simulated by setting the combustion pressure of the corresponding faulty cylinder into a low value, as shown in Figure 3(b). It should be mentioned that the crank

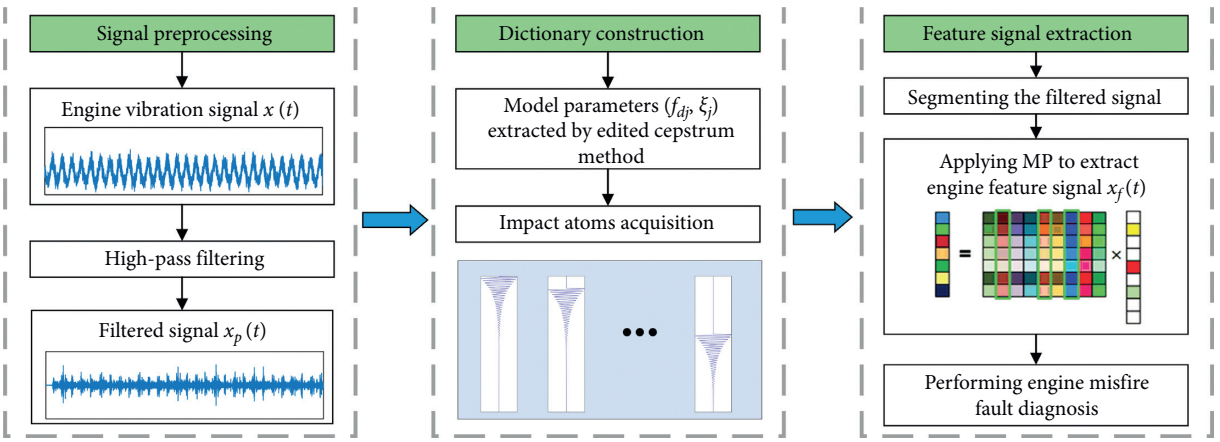
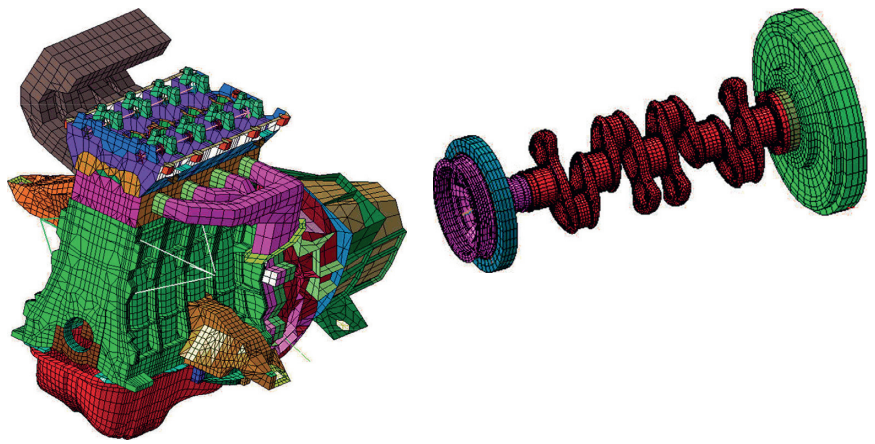
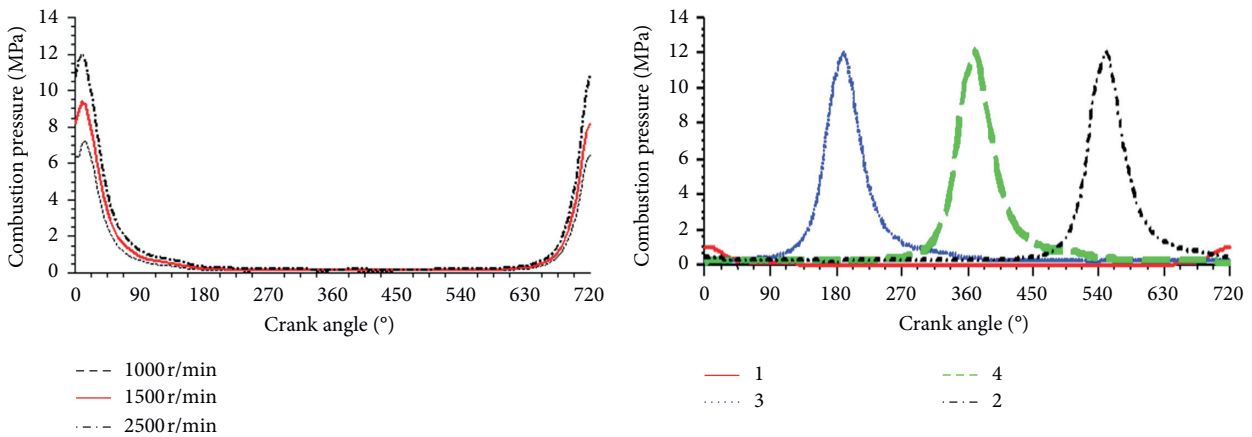


Figure 1: Flow chart of the engine signal feature extraction.



(a) (b)

Figure 2: Finite element models. (a) Engine power unit. (b) Crankshaft.



(a) (b)

Figure 3: Combustion pressure variation along with crank angle. (a) Normal state. (b) First cylinder misfire fault in the four-cylinder engine model.

angle represents the rotational angle of the crankshaft.

3.2. Normal State. Based on the dynamic model constructed in Section 3.1, the normal state signal is acquired on the hood with a sample frequency f_s of 18748 Hz. From Figure 4(a), it can be hardly seen of any features generated by engine impact forces both in the time domain and frequency domain. Even in the noiseless simulation condition, the features indicating the health state of the engine are still obscured by other low-frequency components. According to the signal analysis theory, impact forces will excite the system to vibrate at the high-frequency region. In order to eliminate the influence of low frequency as much as possible, combined with the spectrum of the original signal, a high-pass filter with the cutoff frequency $f_{\text{cut}} = 2000$ Hz is applied. The filtered signal $x_p(t)$ is displayed in Figure 4(b), where the SNR of impact signals is improved significantly.

Then, multiorder model parameters that highly correlated with the engine impact feature signal are identified to construct the sparse dictionary, and MP is further utilized to extract impact features from the filtered signal, and the iteration value δ is set as 0.001, which could make the residual signal as small as enough and ensure the accuracy of the reconstructed signal according to Reference [14]. Here, the results are illustrated in Table 1 and Figure 5. Almost all identified frequencies from Table 1 are located at corresponding resonance regions in Figure 4(b), which demonstrates the correctness of the identification results and further provides high quality parameters for the dictionary design. As can be seen in Figure 5, nearly all the amplitudes and moments of the impact features in the extracted signal match well with those of the filtered noiseless signal. The results strongly verify the effectiveness of the proposed method in the feature extraction aspect. Moreover, each impact feature that is generated by the healthy engine cylinder is clearly exposed, which reveals the health state of the engine.

To further research the antinoise performance of the proposed method, a Gaussian white noise with SNR = 5 dB is added into the simulation signal of Figure 4(a). The filter parameters are the same as the noiseless condition, and corresponding results are shown in Figure 6. Compared with Figure 4(b), the impact signals in Figure 6(b) are overwhelmed by noise after removing low-frequency interferences, which increases the difficulty in extracting the engine feature signal. Although, few extracted impact features exist slightly different, the vast majority of extracted signals obtained from the proposed method are still in accordance with the filtered noiseless signal, as shown in Figure 7. In sum, the proposed method can extract engine features well under low SNR conditions.

In order to further illustrate the advantage of the proposed method under low SNR, wavelet decomposition is chosen to process the same simulation signals in terms of its good feature extraction ability. In wavelet decomposition, db10 is selected as the basic function of the wavelet, and the decomposition layer is set as 4. FS represents the filtered signal; d1, d2, d3, and d4 denote the decomposition results. In the noiseless condition, it can be clearly seen that the wavelet decomposition method is able to extract engine impact features as effectively as the proposed method, as shown in Figure 8, when compared with the results in Figure 7 in the noise condition; though the wavelet decomposition method could denoise the filtered signal to some extent and make a few impact features clearer, the majority of the impact features are still submerged in strong noise, which reduces the diagnosis accuracy of engine misfire fault. In summary, the proposed method cannot only effectively extract and highlight engine features in the noiseless condition but also in the low SNR condition.

3.3. Single-Cylinder Misfire Fault. Engine misfire fault is mainly due to its cylinder failing to ignite properly, namely, under the unfired state. When the engine has misfire fault in the cylinder, corresponding impact signals will disappear theoretically due to its unsuccessful ignition in the faulty cylinder. Hence, the misfire fault feature can be described as a series of impact feature signal in the time domain; among the signals, some of the impact waves will miss due to the faulty cylinder.

Like the normal state, the single-cylinder misfire fault signal obtained from Section 3.1 also processed in the noiseless condition and 5 dB noise condition. As shown in Figure 9(a), it is still difficult to diagnose engine misfire fault by simply observing the original signal even in the noiseless condition. After removing interferences caused by low-frequency components, the impact features have appeared out much more clearly in Figure 9(b). Moreover, it can be found that one impact feature is missing in the four consecutive impact intervals by comparing with the filtered signal under the normal state. This phenomenon just intuitively reveals the features of engine misfire fault and demonstrates the correctness of the simulation signal generation model in turn. Then, the feature signal extracted by the proposed method is illustrated in Figure 10, which shows a high similarity to the filtered signal, and its diagnosis features are also clearly exposed.

Under 5 dB noise condition, signals are more complex to analyse and basically hard to see any impact features both in the original signal and filtered signal, as shown in Figure 11. Hence, the proposed method is utilized to extract important impact features indicating engine misfire fault. From Figure 12, all the engine ignition impact features with high amplitude are well extracted, and some weak impact features are also exposed slightly because of noise influence. In general, although the unavoidable noise coming from the environment may cause certain weak pseudoimpact features, the proposed method still can effectively extract engine misfire fault features with much higher amplitude to make correct diagnosis under the noise condition.

Similar to the normal state, the misfire signals are also performed by the wavelet decomposition method. As the results shown in Figure 13, engine misfire fault features can be well extracted under the noiseless condition. When it comes to the low SNR situation, majority of the impact

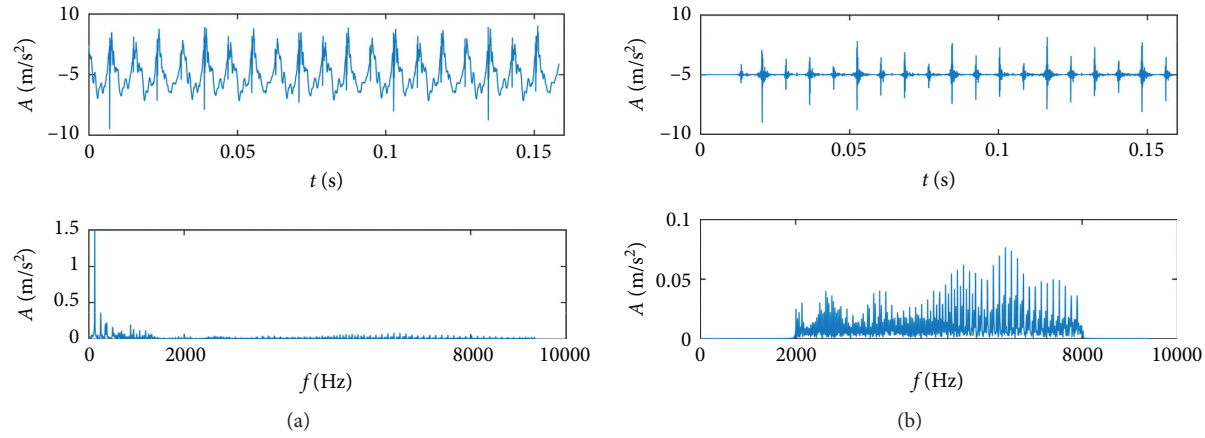


Figure 4: Vibration noiseless signal of the time domain and frequency domain under the normal state. (a) Original signal. (b) Filtered signal.

Table 1: The simulation results of multiorder model parameters' identification.

Parameter	1	2	3	4	5	6
Frequency f_{dj} (Hz)	2189	3996	5483	5969	7022	7901
Damping ratio ξ_j	0.0737	0.0286	0.0475	0.0095	0.0992	0.0125

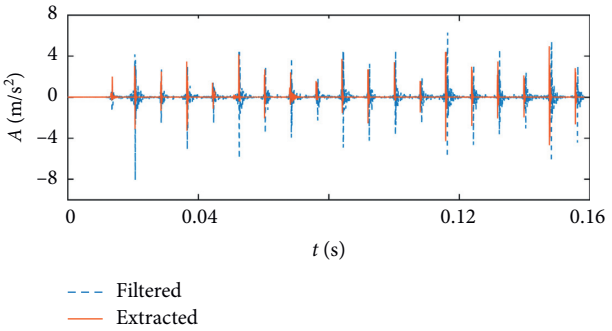


Figure 5: Filtered noiseless signal and extracted feature signal under the noiseless condition.

features are still covered by noise interferences, which hinders the diagnosis of engine misfire fault and further illustrates the superiority of the proposed method.

4. Experiment Verification

To further verify the effectiveness of the proposed method, a four-stroke in-line four-cylinder multipoint injection gasoline engine with a displacement of 1.6 L was operated under the normal state and single-cylinder misfire fault state, respectively. As shown in Figure 14, the vibration signal was obtained by the PCB vibration acceleration sensor mounted on the hood. The signal acquisition equipment contains the MKII signal collector and PAK acquisition and computer analysis system. The sampling frequency was set as 10240 Hz. Misfire fault of the engine was accomplished by cutting off the corresponding cylinder fuel supply. Additionally, the experiment was carried out under no load because of the limited conditions. Comparison analysis with the experimental signal illustrates the advantages of the proposed method in terms of feature extraction ability.

4.1. Normal State. As can be seen in Figure 15(a), the experimental signal acquired on the hood also includes low-frequency components' interference, while the important high-frequency regions containing impact features are suppressed largely. This also verifies the correctness of the constructed dynamic model from another aspect by observing the similarity of simulation signals and experimental signals. Therefore, a high-pass filter is applied to highlight the impact features of the original signal as in simulation. The results are presented in Figure 15(b); compared with the unfiltered signal, the SNR of the filtered signal is indeed improved, and some weak impact features can be observed. However, it is still difficult to diagnose the health state of the engine by the current signal because of the complex noise interferences.

The multiorder model parameters affecting impact vibration characteristics are identified and listed in Table 2. The filtered signal is then utilized to exhibit the feature extraction superiority of the proposed method, as shown in Figure 16. Most of the impact features have been effectively extracted as well as the impact amplitudes. Compared with the filtered signal, the extracted signal is clearer and can better represent the engine operation condition. Furthermore, although there are few weak impact features influenced by noise, the amplitudes of most of the reconstructed signals are the same as those of the filtered signals and there is no periodic low-amplitude impact, which are consistent with the normal simulations. Hence, the engine health state can be concluded to the normal state, and the results of the experiment further verify the effectiveness of the proposed method.

4.2. Single-Cylinder Misfire Fault. The misfire fault experiment was also conducted on the same condition with the normal state. As illustrated in Figure 17(a), the original vibration signal obtained from the hood is disorderly and

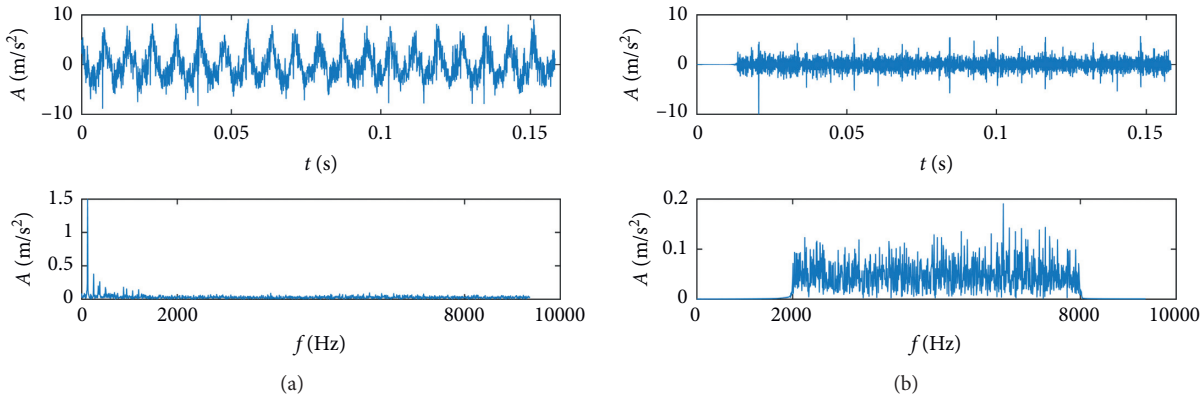


Figure 6: Vibration signal of the time domain and frequency domain under the normal state with 5 dB white noise. (a) Original signal. (b) Filtered signal.

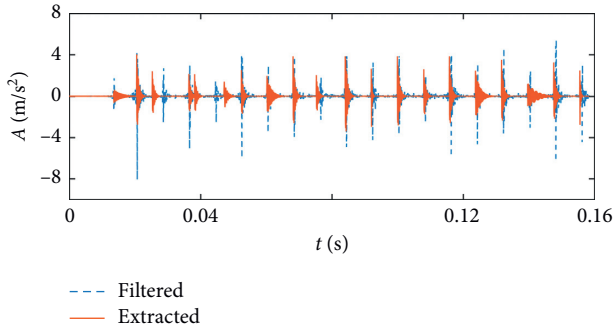


Figure 7: Filtered noiseless signal and extracted feature signal under 5 dB noise condition.

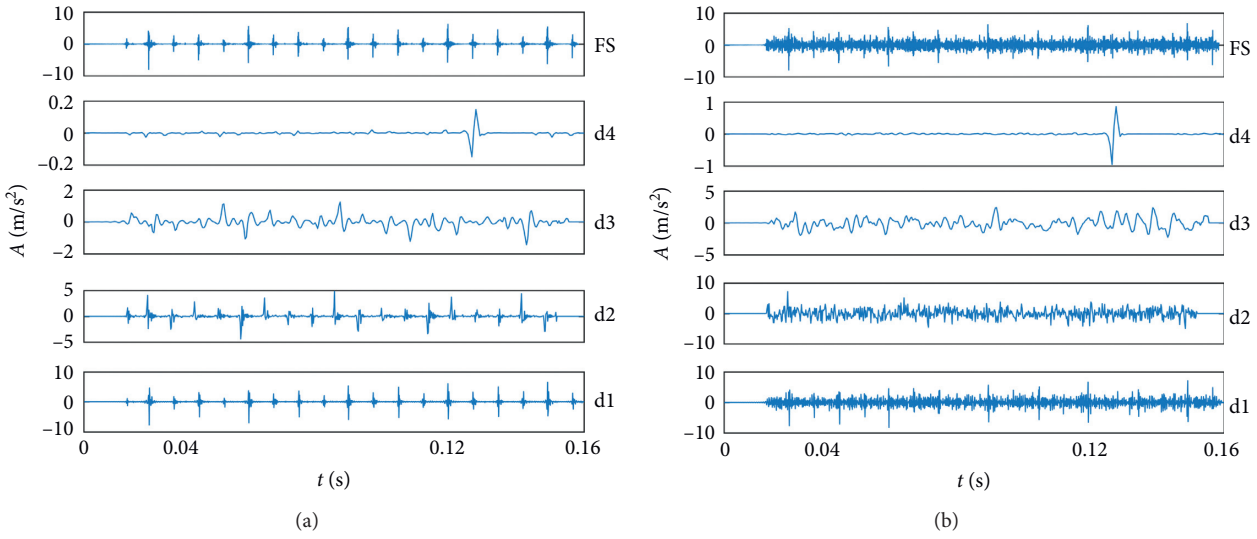


Figure 8: The results of the normal simulation signal processed by wavelet decomposition. (a) Noiseless. (b) Under 5 dB noise.

cannot be observed as useful information. After filtering the low-frequency noise, impact features located in the resonance area are highlighted, but still with some noise disturbances influencing engine health diagnosis.

Therefore, the impact features indicating the engine health state are further extracted by the proposed method. According to the misfire fault simulation results under noiseless and noise conditions, it can be known from that when single-cylinder misfire fault occurs, the extracted signal will still show equally spaced impacts rather than directly missing one impact. More importantly, the impact feature related with the faulty cylinder

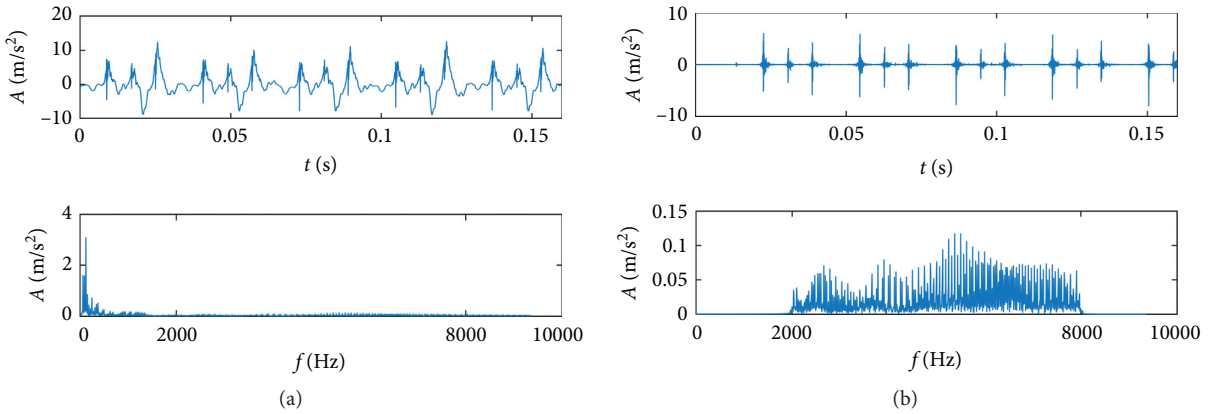


Figure 9: Vibration signal of the time domain and frequency domain under the misfire state. (a) Original signal. (b) Filtered signal.

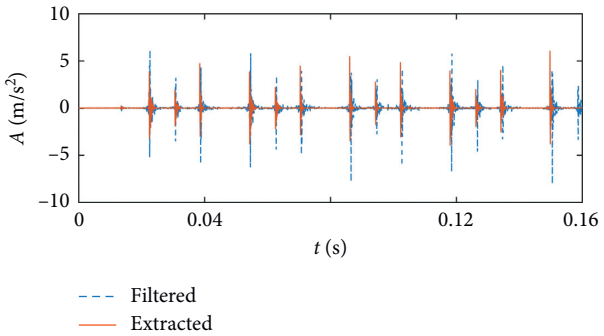


Figure 10: Filtered noiseless signal and extracted feature signal under the noiseless condition with single-cylinder misfire fault.

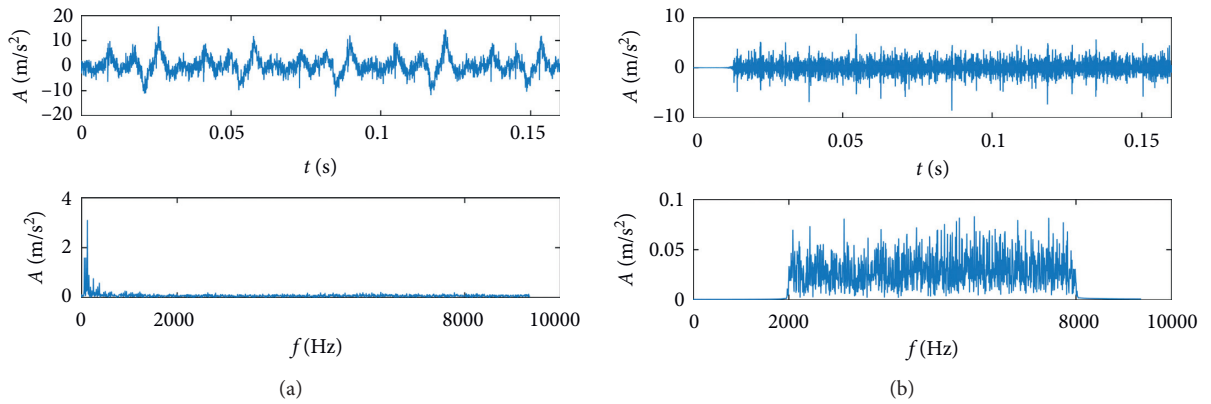


Figure 11: Vibration signal of the time domain and frequency domain under the misfire state with 5 dB white noise. (a) Original signal. (b) Filtered signal.

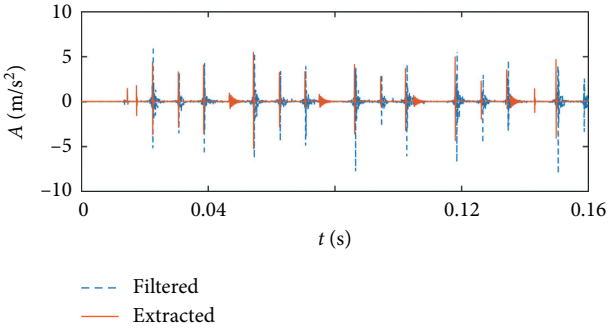


Figure 12: Filtered noiseless signal and extracted feature signal under 5 dB noise condition with single-cylinder misfire fault.

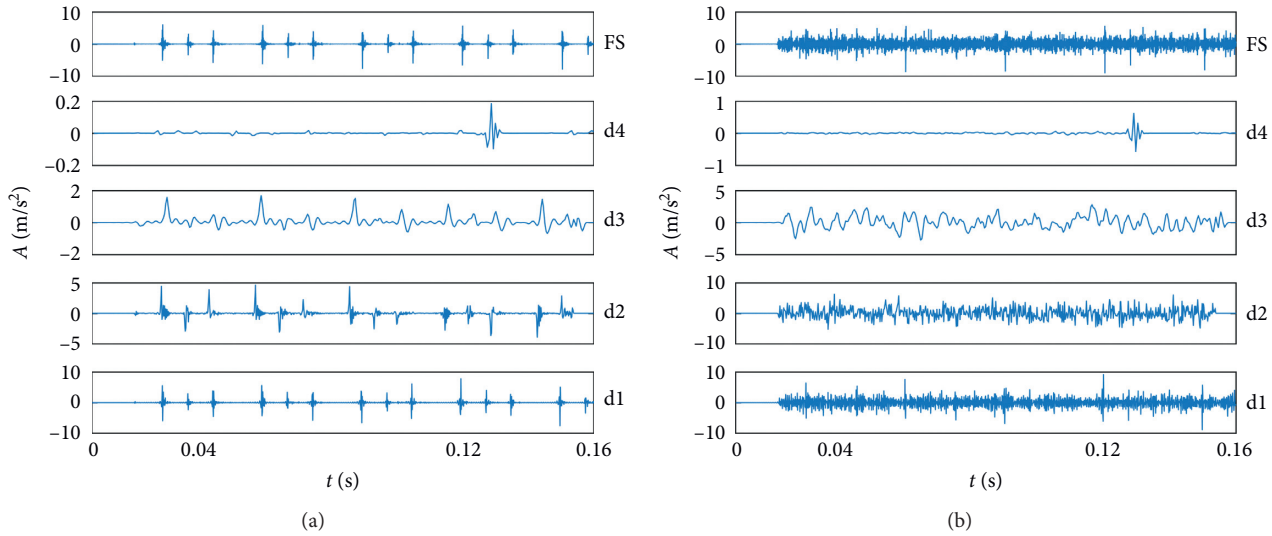


Figure 13: The results of the engine single-cylinder misfire fault signal processed by wavelet decomposition. (a) Noiseless. (b) Under 5 dB noise.

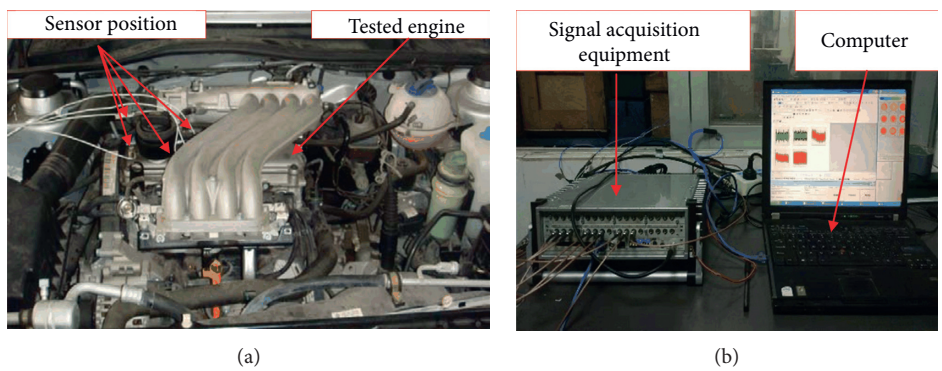


Figure 14: Experimental devices. (a) Tested engine. (b) Signal acquisition equipment.

Table 2: The experimental results of multiorder model parameters' identification.

Parameters	1	2	3	4
Frequency f_{dj} (Hz)	3089	3453	4099	4891
Damping ratio ξ_j	0.0132	0.0416	0.1730	0.0150

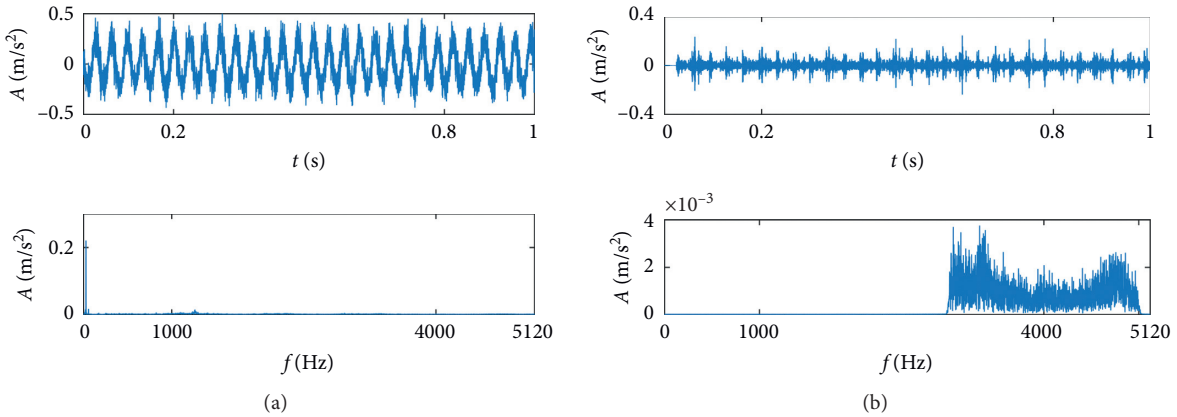


Figure 15: Experimental signal of the time domain and frequency domain under the normal state. (a) Original signal. (b) Filtered signal.

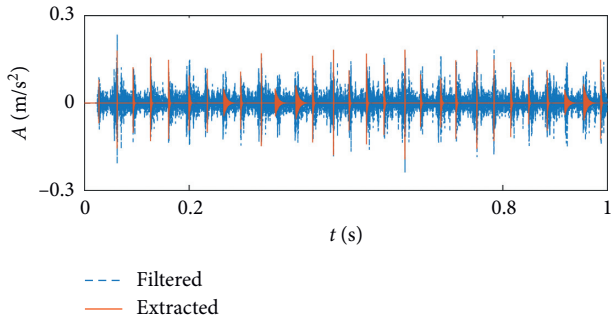


Figure 16: Experimental normal signal processed by the proposed method.

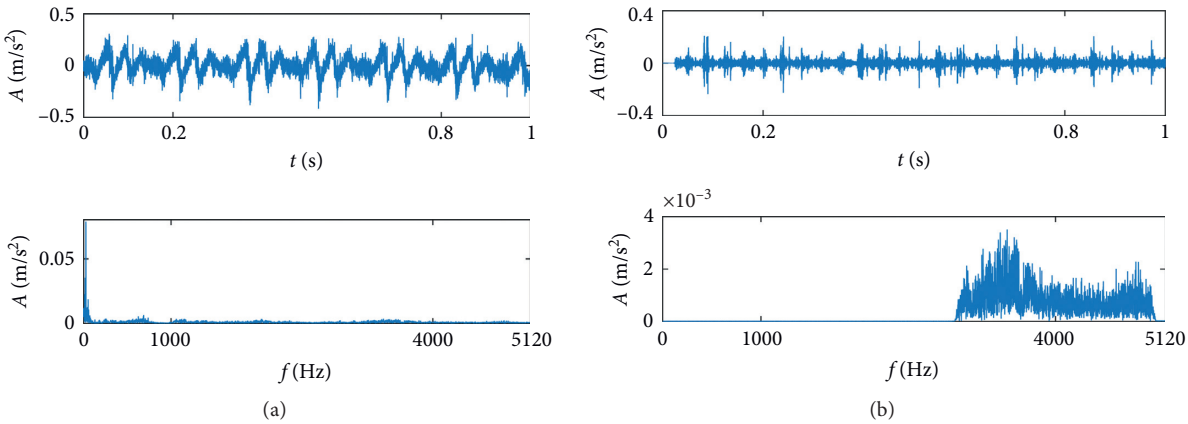


Figure 17: Experimental signal of the time domain and frequency domain under misfire fault. (a) Original signal. (b) Filtered signal.

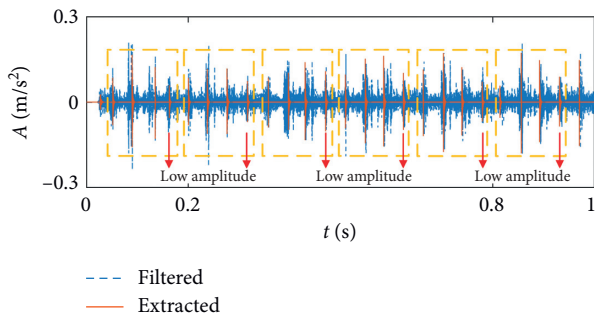


Figure 18: Experimental misfire fault signal processed by the proposed method.

has lower amplitude than others and shows quasi-periodicity. As shown in Figure 18, the results from the experimental signal are similar to the simulation results and therefore the engine misfire fault is diagnosed, which further verifies the correctness of the proposed method.

The results processed by wavelet decomposition are illustrated in Figure 19. It can be clearly seen that both extracted signals under the normal state and misfire state are submerged by strong noise, which hinders the engine misfire fault diagnosis. The comparison results further show the feature extraction advantage of the proposed method.

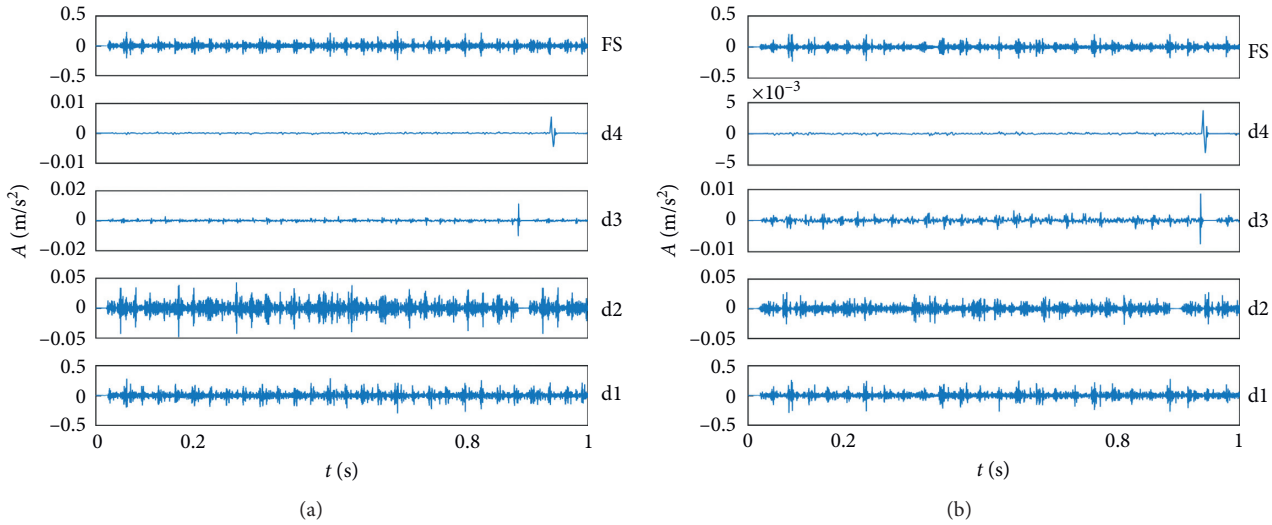


Figure 19: The results of the experimental signal processed by wavelet decomposition. (a) Normal state. (b) Single-cylinder misfire fault state.

5. Conclusions

Based on signal sparse decomposition theory, a feature extraction method is proposed to accurately extract impact features indicating the engine health state. The unit response function that reflects the system modal characteristic is used as the atom to construct the sparse dictionary with clear physical meaning. Moreover, the segmental MP algorithm is applied to reconstruct the impact features from the original signal. Simulations and experimental tests both verify the effectiveness of the proposed method. Some important conclusions are as follows:

(1) A four-cylinder engine multibody dynamic model constructed by the AVL EXCITE platform can generate more realistic signals than oversimplified numerical simulation. Furthermore, it cannot only simulate the engine normal signal but also the engine misfire fault signal by adjusting the combustion pressure of the corresponding cylinder.

(2) Original simulation signals contain many noise interferences, in which exists low-frequency components generated by other rotational parts. In the normal state, the proposed method can effectively extract all impact features in the noiseless condition. Moreover, under the low SNR condition, it can still make submerged impact features exposed clearly.

(3) In the engine misfire fault state, the faulty cylinder will not produce impact features in the corresponding ignition moment theoretically. Hence, impact features can be well extracted but with missing one impact by the proposed method in the noiseless condition. However, in the low SNR condition, besides the normal state impact features can be extracted, the pseudoimpacts with lower amplitude also appeared out due to the noise interferences.

(4) In real automotive engine tests, the proposed method still makes full use of the fault information that is buried in the original signal and highlights impact features to accurately diagnosis the health state of the engine, especially detect engine misfire fault.

(5) Though the wavelet decomposition method could obtain good extraction results under the noiseless condition as the proposed method, when it comes to the low SNR condition, only the proposed method can effectively extract the feature signals that indicate the engine health state.

Data Availability

The primary 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

This work was partly supported by the Natural Science Foundation of Guangdong Province, China (nos. 2018A030313947, 2019A1515011779, and 2020A1515010972), Fundamental Research Funds for the Central Universities (no. 2019MS054), and Key Project of Guangzhou Science and Technology Program, China (no. 201904010387).

References

[1] Y. Xu, B. Huang, Y. Yun, R. Cattley, F. Gu, and A. D. Ball, "Model based IAS analysis for fault detection and diagnosis of IC engine powertrains," *Energies*, vol. 13, no. 3, p. 565, 2020.

[2] C. Gu, X. Qiao, Y. Jin, and Y. Liu, "A novel fault diagnosis method for diesel engine based on MVMD and band energy,"

Shock and Vibration, vol. 2020, Article ID 8247194, 17 pages, 2020.

[3] N. Dayong, J. Yuhua, S. Hongyu et al., "Separation method of impulsive fault component for gasoline engine based on acoustic signal analysis," *Shock and Vibration*, vol. 2019, Article ID 8573479, 15 pages, 2019.

[4] F. Jiang, K. Ding, G. He et al., "Vibration fault features of planetary gear train with cracks under time-varying flexible transfer functions," *Mechanism and Machine Theory*, vol. 158, Article ID 104237, 2021.

[5] X. Bi, S. Cao, and D. Zhang, "Diesel engine valve clearance fault diagnosis based on improved variational mode decomposition and bispectrum," *Energies*, vol. 12, no. 4, p. 661, 2019.

[6] K. Vernekar, H. Kumar, and K. V. Gangadharan, "Engine gearbox fault diagnosis using empirical mode decomposition method and Naïve Bayes algorithm," *Sādhanā*, vol. 42, no. 7, pp. 1143–1153, 2017.

[7] T. Figlus, Š. Liščák, A. Wilk, and B. Łazarz, "Condition monitoring of engine timing system by using wavelet packet decomposition of a acoustic signal," *Journal of Mechanical Science and Technology*, vol. 28, no. 5, pp. 1663–1671, 2014.

[8] W. Deng, H. Liu, J. Xu, H. Zhao, and Y. Song, "An improved quantum-inspired differential evolution algorithm for deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, p. 7319, 2020.

[9] C. Du, W. Li, F. Yu et al., "Misfire fault diagnosis of automobile engine based on time domain vibration signal and probabilistic neural network," *International Journal of Performability Engineering*, vol. 16, no. 9, 2020.

[10] K. Jafarian, M. Mobin, R. Jafari-Marandi, and E. Rabiei, "Misfire and valve clearance faults detection in the combustion engines based on a multi-sensor vibration signal monitoring," *Measurement*, vol. 128, pp. 527–536, 2018.

[11] J. Tao, C. Qin, W. Li, and C. Liu, "Intelligent fault diagnosis of diesel engines via extreme gradient boosting and high-accuracy time-frequency information of vibration signals," *Sensors*, vol. 19, no. 15, p. 3280, 2019.

[12] Y. Liu, X. Chen, R. K. Ward, and Z. Jane Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.

[13] C. A. Riofrío, D. Gross, S. T. Flammia et al., "Experimental quantum compressed sensing for a seven-qubit system," *Nature Communications*, vol. 8, no. 1, pp. 1–8, 2017.

[14] G. He, K. Ding, and H. Lin, "Fault feature extraction of rolling element bearings using sparse representation," *Journal of Sound and Vibration*, vol. 366, pp. 514–527, 2016.

[15] F. Jiang, K. Ding, G. He et al., "Sparse dictionary design based on edited cepstrum and its application in rolling bearing fault diagnosis," *Journal of Sound and Vibration*, vol. 490, Article ID 115704, 2021.

[16] Y. Liu, J. Zhang, and L. Ma, "A fault diagnosis approach for diesel engines based on self-adaptive WVD, improved FCBF and PECOC-RVM," *Neurocomputing*, vol. 177, pp. 600–611, 2016.

[17] S. Li, Y. Zhang, L. Wang et al., "A CEEMD method for diesel engine misfire fault diagnosis based on vibration signals." in *Proceeding of the 39th Chinese Control Conference (CCC)*, pp. 6572–6577, Shenyang, China, July 2020.

[18] T. Parikyan, T. Resch, and H. H. Priebsch, "Structured model of crankshaft in the simulation of engine dynamics with AVL/excite," *Internal combustion engine division fall technical conference*, vol. 80159, pp. 105–114, 2001.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6690966, 15 pages
https://doi.org/10.1155/2021/6690966

Research Article

Early Weak Fault Diagnosis of Rolling Bearing Based on Multilayer Reconstruction Filter

Quanfu Li,[1,2] Yuxuan Zhou,[3] Gang Tang,[3] Chunlin Xin,[3] and Tao Zhang[4]

[1]*Southwest Jiaotong University, Chengdu 611756, China*
[2]*Shenhua Railway Equipment Co. Ltd., Beijing 100120, China*
[3]*Beijing University of Chemical Technology, Beijing 100029, China*
[4]*AECC Hunan Aviation Powerplant Research Institute, AECC Key Laboratory of Aero-engine Vibration Technology, Zhuzhou 412002, Hunan, China*

Correspondence should be addressed to Gang Tang; tanggang@mail.buct.edu.cn

Received 11 December 2020; Revised 5 January 2021; Accepted 21 January 2021; Published 2 March 2021

Academic Editor: Jinde Zheng

The early weak fault characteristics of rolling bearings are extremely weak and are easily overwhelmed by other noises. In order to effectively extract the characteristics of the early weak faults of the rolling bearings and draw on the multilayer wavelet decomposition idea, a method for diagnosing the early weak faults of the rolling bearing based on the multilayer reconstruction filter is proposed. As we all know, empirical wavelet transform (EWT) makes full use of wavelet filter bank, and variational mode decomposition (VMD) uses Wiener filter bank. This paper fully combines the advantages of the above two methods, adaptively determines the number of modes through empirical wavelet decomposition and divides the original signal, extracts the frequency band that contains the fault characteristic information, and effectively eliminates noise interference. These steps are repeated until the optimal component of the condition is obtained. In the output layer, the weak fault impact components are further separated by the strong filtering and signal decomposition capability of VMD. The advantages of the proposed method are proved by the experiment of weak fault of rolling bearing and the accelerated failure experiment of full life. The proposed method has the advantages of reducing noise influence and adaptive estimation of decomposed modes, which can be applied more efficiently in practice.

1. Introduction

Timely and effective identification of early weak faults for rolling bearings has an important significance to ensure the safety of equipment operation [1]. The small impact of early weak faults leads to weak fault characteristics. At the same time, the signal usually contains a large number of noise interferences caused by the mutual impact of other components, which brings about great challenges to relevant studies [2, 3].

Data-driven fault diagnosis methods for rolling bearings have been the most widely studied, most of which are based on vibration signals and acoustic signals, which are essentially the same, but each has its own advantages and disadvantages [4]. The advantages of vibration signal used in fault diagnosis are as follows: the cost of the sensor is relatively low, the vibration signal is easy to be measured, and the measured vibration signal can contain more fault information. Its disadvantages are as follows: this is a kind of contact measurement, and the sensor needs to be installed in the position close to the workpiece. The advantages of using acoustic signals for fault diagnosis are as follows: the requirements for the installation position are not very strict, they do not need to be attached to the workpiece under test, and the measurement cost is relatively low. The disadvantage is that the measured acoustic signal may contain sound other than the target workpiece, so the SNR is relatively small [5]. Adam Glowacz used acoustic signals for early bearing and stator of the single-phase induction motor

fault diagnosis [6]. Parey and Singh proposed a continuous wavelet transform and adaptive neural fuzzy reasoning system to process the acoustic signal of gearbox and complete fault diagnosis [7]. Adam Glowacz proposed a fault diagnosis method of single-phase motor based on acoustic signal [5]. Zhang and Stewart proposed a deep graph convolutional network for roller bearings based on acoustic signal [8]. Liu and Pei achieved fault detection on belt conveyor idlers by processing acoustic signals using machine learning method [9]. In the application of vibration signals for fault diagnosis, Huang et al. extracted frequency band entropy from vibration signals for fault diagnosis [10]. Wang et al. used time-frequency curves of vibration signals for fault detection under variable speed [11].

In the past decades, researchers have proposed many methods for the rolling bearing fault diagnosis, among which the adaptive signal decomposition method is quite popular [12]. Common adaptive decomposition methods include empirical mode decomposition (EMD), EWT, and VMD [13]. EMD has been widely used in various fields [14, 15]. Due to its difficulty in mathematical modeling, sensitivity to noise, and endpoint effect, scholars have developed a lot of improved EMD, but it still cannot achieve ideal results [16]. In order to overcome the limitations of EMD, Jerome Gilles and Konstantin Dragomiretskiy proposed EWT [17] and VMD [18], respectively. The EWT constructs an adaptive empirical wavelet filter bank and the variational modal decomposition fully borrows the Wiener filter. Both of them have been fully studied in the field of rolling bearing fault diagnosis.

Some researchers made a detailed comparative study on EWT and EMD and concluded that EWT has better performance than EEMD and EMD in model estimation [19]. EWT has been widely used in the fields of medical signal analysis, seismic signal analysis, meteorological prediction, and fault diagnosis [20–23]. In the field of fault diagnosis, some papers have improved EWT from the aspects of spectrum segmentation, so that it can be better applied to practical engineering problems [17, 24]. VMD obtains all modalities from the signal at the same time through a joint optimization scheme, so it has higher resolution [25]. Compared with EWT, this method has been widely used. A large number of scholars have also improved the VMD. Some scholars use the dimensionality-increasing feature of VMD to apply it to underdetermined blind source separation [26]. Jiang et al. [27] proposed a VMD decomposition strategy from rough to fine, which was well applied to the fault diagnosis of rotating equipment. Many scholars have also studied the parameter optimization of VMD to improve its adaptive ability [28, 29]. Xu et al. [30] proposed a method based on VMD which can reduce noise in the propagation path. Although EWT and VMD have been widely studied in the field of rolling bearing fault diagnosis, they still have some room for improvement.

Due to extremely weak fault characteristics and large amount of noise interference in vibration signal and acoustic signal, it is a great challenge to extract fault characteristic information effectively. Some scholars have combined the mode decomposition method with other methods. For example, Fan [2] combined with the advantages of EMD to improve VMD and achieved certain results.

This method utilizes EWT's great capability of mode number estimation [17] and its excellent filtering capability of the empirical wavelet filter bank and combines the kurtosis criterion to effectively extract the optimal mode from the original signal. Multilayer reconstruction filter can remove most of the noise components and highlight the fault impact components. However, due to the weak fault features, it is usually not possible to directly extract the fault characteristics from the denoised signal. By virtue of the great band-pass filtering capability and signal decomposition capability of VMD, it is used as the output layer to reconstruct and filter the signal after noise reduction. Because the multilayer reconstruction filter reduces the interference of redundant information to the VMD, the fault feature information can be separated and extracted better, and the early weak fault diagnosis of rolling bearing can be realized.

The structure of this paper is as follows: Section 2 introduces the theoretical basis of the method and introduces the process of the proposed method; Section 3 verifies the proposed method by using the acoustic signal of rolling bearing fault experiment and the vibration signal of accelerated fatigue experiment, respectively. The results and discussion are given in Section 4.

2. Methodology

2.1. Empirical Wavelet Transform. The EWT divides the Fourier spectrum of a signal and constructs a wavelet filter bank to extract the intrinsic modes and can be divided into several main components.

Assuming that the Fourier support $[0, \pi]$ is segmented into contiguous N segments, a total number of $N+1$ boundaries are needed. According to the local maximum values, the Fourier spectrum of a signal is divided into N segments. Denote ω_n to be the limits between each segment (where $\omega_0 = 0$ and $\omega_n = \pi$). Each segment is represented as $\Lambda_n = [\omega_{n-1}, \omega_n]$. A transition phase T_n of width $2\tau_n$ is then defined, as shown in Figure 1.

The empirical wavelets are defined as band-pass filters on each Λ_n. The empirical scale function $\hat{\phi}(\omega)$ and empirical wavelet function $\hat{\psi}(\omega)$ can be calculated by the two following equations, respectively:

$$\hat{\phi}(\omega)=\begin{cases}1, & \text{if } |\omega|\le(1-\gamma)\omega_n,\\ \cos\left[\frac{\pi}{2}\left(\frac{1}{2\gamma\omega_n}\left(|\omega|-(1-\gamma)\omega_n\right)\right)\right], & \text{if } (1-\gamma)\omega_n\le|\omega|\le(1+\gamma)\omega_n,\\ 0, & \text{otherwise,}\end{cases} \tag{1}$$

$$\hat{\psi}(\omega)=\begin{cases}1, & \text{if } (1+\gamma)\omega_n\le|\omega|\le(1-\gamma)\omega_{n+1},\\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}\left(|\omega|-(1-\gamma)\omega_{n+1}\right)\right)\right], & \text{if } (1-\gamma)\omega_{n+1}\le|\omega|\le(1+\gamma)\omega_{n+1},\\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}\left(|\omega|-(1-\gamma)\omega_n\right)\right)\right], & \text{if } (1-\gamma)\omega_{n+1}\le|\omega|\le(1+\gamma)\omega_{n+1},\\ 0, & \text{otherwise.}\end{cases} \tag{2}$$

The ratio γ in equations (1) and (2) is restricted to a small value as $0\le\gamma\le\min_n[(\omega_{n+1}-\omega_n)/(\omega_{n+1}+\omega_n)]$ to ensure the empirical scaling function and the empirical wavelets are a tight frame of ω_n: $\tau_n=\gamma\omega_n$, $0<\gamma<1$. The function $\beta(x)$ is an arbitrary $C^k([0,1])$ function, and $\beta(x)$ is defined as follows:

$$\beta(x)=x^4\left(35-84x+70x^2+20x^3\right). \tag{3}$$

The approximation coefficients $w_f^{\varepsilon}(0,t)$ and the detail coefficients $w_f^{\varepsilon}(n,t)$ can be calculated by the two following equations, respectively:

$$w_f^{\varepsilon}(0,t)=f,$$
$$\phi_1=\int f(\tau)\overline{\phi_1(\tau-t)}\mathrm{d}\tau=\left(\hat{f}(\omega)\overline{\hat{\phi}_1(\omega)}\right), \tag{4}$$

$$w_f^{\varepsilon}(n,t)=f,$$
$$\psi_n=\int f(\tau)\overline{\psi_n(\tau-t)}\mathrm{d}\tau=\left(\hat{f}(\omega)\overline{\hat{\psi}_n(\omega)}\right)^{\vee}, \tag{5}$$

where $\hat{\phi}_1(\omega)$ and $\hat{\psi}_n(\omega)$ are defined by equations (1) and (2), respectively.

The inverse empirical wavelet transformation is carried out by the following equation:

$$\begin{aligned}f(t)&=w_f^{\varepsilon}(0,t)*\phi_1(t)+\sum_{n=1}^{N}w_f^{\varepsilon}(n,t)*\psi_n(t),\\&=\left(\hat{w}_f^{\varepsilon}(0,\omega)*\hat{\phi}_1(\omega)+\sum_{n=1}^{\hat{N}}\hat{w}_f^{\varepsilon}(n,\omega)*\hat{\psi}_n(\omega)\right)^{\vee}.\end{aligned} \tag{6}$$

2.2. Variational Mode Decomposition. Variational modal decomposition is an adaptive time-frequency analysis tool, which decomposes the signal into multiple BLIMFs (band limited intrinsic mode function) through iterative solution. The detailed steps are described below.

First, it is assumed that the signal $x(t)$ can be decomposed into a finite number of inherent modes, each of which has a different center frequency and ω_k, a finite bandwidth.

The specific formula of the unilateral spectrum obtained by Fourier transform is as follows:

$$\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t). \tag{7}$$

In the above formula, $\delta(t)$ stands for Dirichlet function, and $*$ stands for convolution.

The spectrum of each component is obtained by the following equation:

$$\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{j\omega_k t}. \tag{8}$$

The variational constraint mode function of VMD is as follows:

$$\begin{aligned}&\min_{\{u_k\},\{w_k\}}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-jw_k t}\right\|_2^2\right\},\\&\text{s.t.}\quad\sum_k u_k=x.\end{aligned} \tag{9}$$

In the above formula, $\{u_k\}=\{u_1,\ldots,u_K\}$ represents the K components obtained after decomposition; $\{\omega_k\}=\{\omega_1,\ldots,\omega_K\}$ represents the center frequency corresponding to each component.

In order to obtain the optimal solution of the above variational constraint model, the VMD algorithm introduces the Lagrangian method, including the secondary punishment factor α and the Lagrangian multiplier λ. The formula is as follows:

$$L = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \left\| x(t) - \sum_k u_k(t) \right\|_2^2 + \lambda(t), x(t) - \sum_k u_k(t). \tag{10}$$

Iteratively update $\{u_k\}, \{w_k\}$, and finally obtain the saddle point of Lagrange equation, which is the optimal solution.

The specific steps of the VMD procedure are shown in Figure 2.

2.3. Proposed Method: Multilayer Reconstruction Filter. The VMD method has a reliable theoretical basis, with bandpass filtering capability similar to wavelet packet decomposition and excellent signal decomposition capability. However, weak fault pulses and strong background noise can seriously affect the accuracy of diagnosis. Therefore, drawing on the idea of multilayer wavelet decomposition, this paper introduces a powerful signal processing tool, EWT, as a preprocessing. EWT has an outstanding ability to detect the number of modes. This method divides the signal in the Fourier domain by looking for the local maximum and divides the signal into multiple frequency bands adaptively. By constructing an empirical wavelet filter bank with excellent performance, each frequency band is reconstructed and filtered to effectively eliminate the interference of nonimpulsive noise. As a fourth-order high-order statistic equation (11), kurtosis is very sensitive to early failures. The correlation coefficient can choose the most containing fault information frequency band. Therefore, with the help of the above two coefficients, a relatively optimal component can be well selected. Using this component as the input signal of the next layer, perform the above process again until the set number of decomposition layers is reached to obtain the optimal mode.

$$K = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{x_i(t) - \mu}{\sigma} \right]^4. \tag{11}$$

In the above formula, N is the length of the signal, x_i is the i-th EWT component, and μ and σ are the mean value and standard deviation of the signal, respectively.

Compared with the original signal, the selected optimal component has very low noise interference and no excessive redundant information. Taking this component as the input of the output layer, it can better exert the powerful signal filtering and decomposition ability of VMD, effectively decompose the fault pulse information, and realize accurate fault diagnosis. The flow chart of the proposed method is shown in Figure 3.

3. Experiment and Results

The method proposed in this paper aims at the early weak faults of rolling bearings with extremely weak fault characteristics and extremely high background noise. This chapter uses two related experiments for analysis. In experiment 1, weak faults were made in the inner and outer rings of the bearing ring by using laser cutting technology, and acoustic signals were collected by microphone sensor. The acoustic signals collected by this experiment met the premise of weak fault characteristics and great background noise. In experiment 2, the bearing accelerated degradation experiment of NSF I/UCR Intelligent Maintenance Systems was adopted. The experiment collected data on the entire life cycle of the bearing, which truly reflects the entire stage of failure from early initiation and development to destruction and shutdown, which is well in line with the needs of experimental research. Two experiments are presented below.

3.1. Simulation Experiment Analysis of Early Fault of Rolling Bearing

3.1.1. Experiment Design. In this paper, a rolling bearing fault test rig is designed, as shown in Figure 4. The type of bearing is NSK NU205EW, whose parameters are shown in Table 1. Laser cutting method is adopted to cut through 0.1 mm wide and 0.05 mm deep grooves in the inner and outer rings of bearing. Acquisition card was used to record the experimental data, MPA416 microphone was used to collect the acoustic signal, the sensor was installed in front of the faulty bearing, the distance was 30 mm, and the photoelectric encoder was used to record the speed. Experimental bench is with rated power of 1.5 kW, rated voltage of 380 V, rated frequency of 50 Hz, rated current of 3.4 A, rated speed of 2840 r/min of three-phase asynchronous motor drive, transverse load of 2 kN, sampling frequency of 20 kHz, and speed of 1309.2 r/min (f = 21.82 Hz). The inner and outer ring fault frequencies of the bearings are 169.1 Hz and 114.6 Hz, respectively. Equations (12)–(14), respectively, give the calculation formula of fault characteristic frequency of bearing's inner ring, outer ring, and rolling element.where f_I, f_O, f_B, respectively, represent the characteristic frequency of inner ring fault, the characteristic frequency of outer ring fault, and the characteristic frequency of rolling element fault (unit: Hz); r is the speed of rotation of the bearing (unit: (r/min)); n denotes ball number; d denotes the diameter of the rolling element (unit: mm); D denotes bearing pitch diameter (unit: mm); and α denotes contact Angle (unit: ∘).

$$f_I = \frac{n}{2} \frac{r}{60} \left(1 + \left(\frac{d}{D} \cos \alpha \right) \right), \tag{12}$$

$$f_O = \frac{n}{2} \frac{r}{60} \left(1 - \left(\frac{d}{D} \cos \alpha \right) \right), \tag{13}$$

$$f_B = \frac{n}{2} \frac{r}{60} \frac{D}{d} \left(1 - \left(\frac{d}{D} \cos \alpha \right)^2 \right), \tag{14}$$

3.1.2. Experimental Results and Analysis. First, the time domain waveform and Hilbert envelope spectrum of the original signal were made, as shown in Figures 5 and 6,

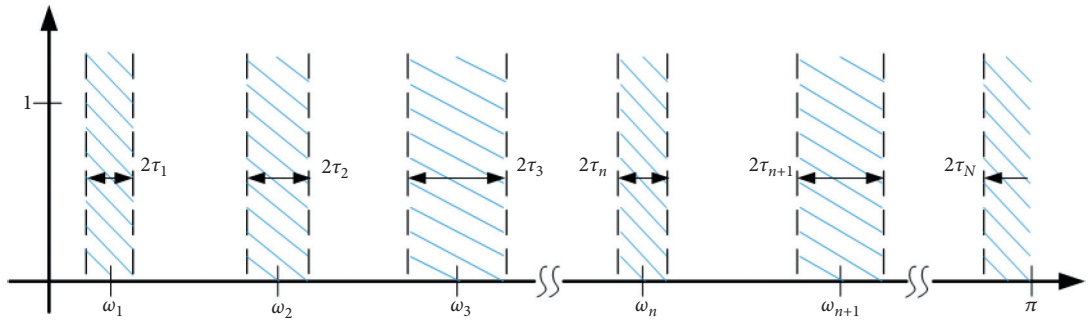


Figure 1: Partitioning of the Fourier axis.

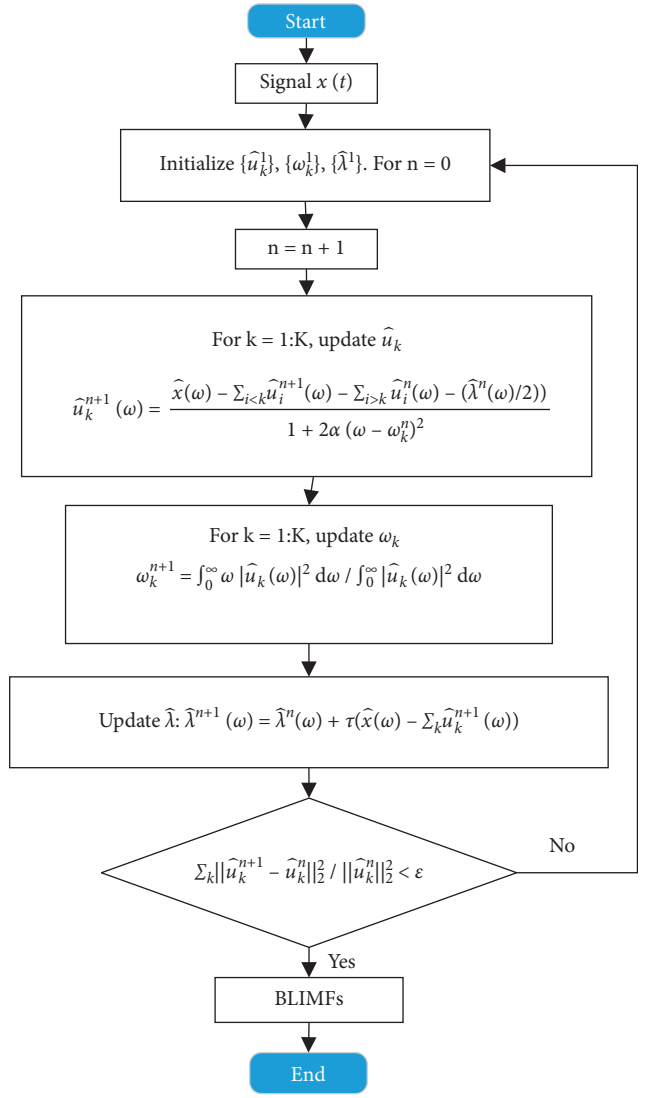


Figure 2: Variational mode decomposition algorithm flow.

respectively. The time domain diagram shows that the signal amplitude is low and unobvious periodic impact is found. Further observation of the signal's Hilbert envelope spectrum revealed a large number of irrelevant frequencies, especially in the low frequency region, indicating that the original signal contained a large number of noise interferences. From the spectrogram, only the inner ring fault characteristic frequency with less obvious amplitude can be found, but the inner ring fault characteristic frequency cannot be found, which is enough to prove the premise of

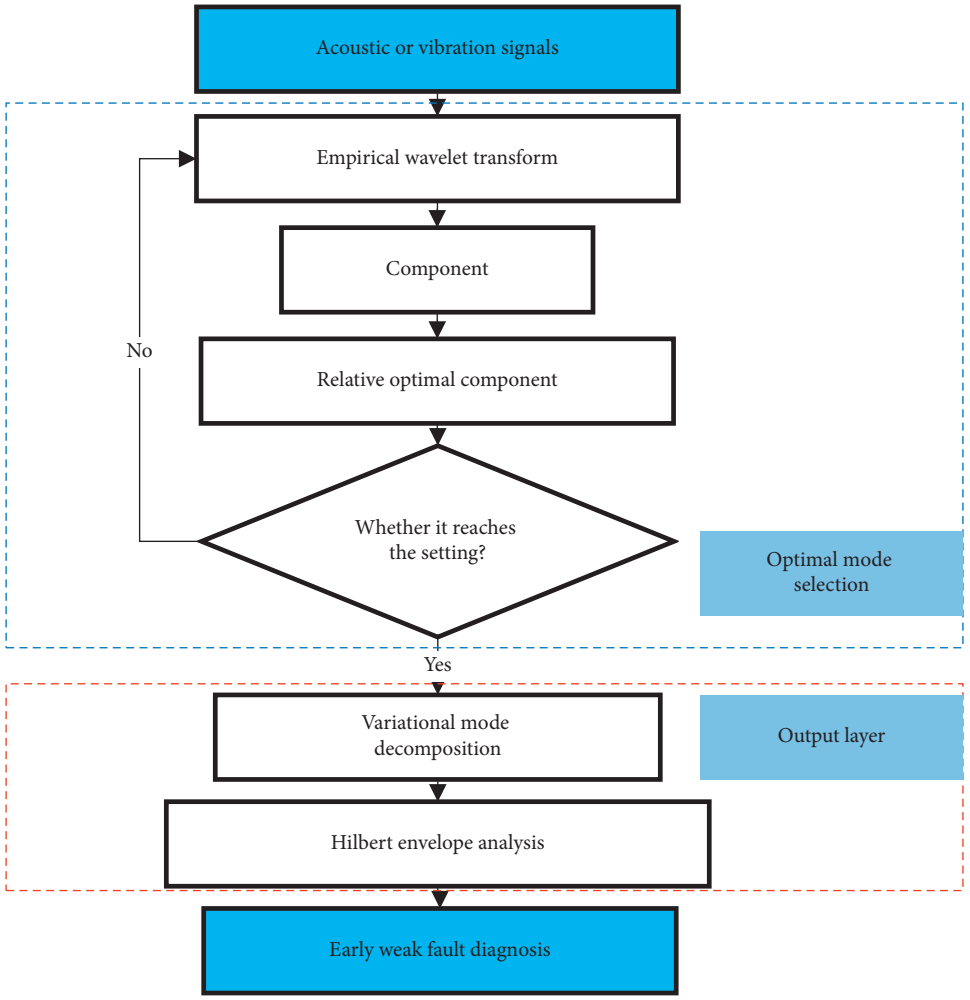


Figure 3: The flow chart of the proposed method.

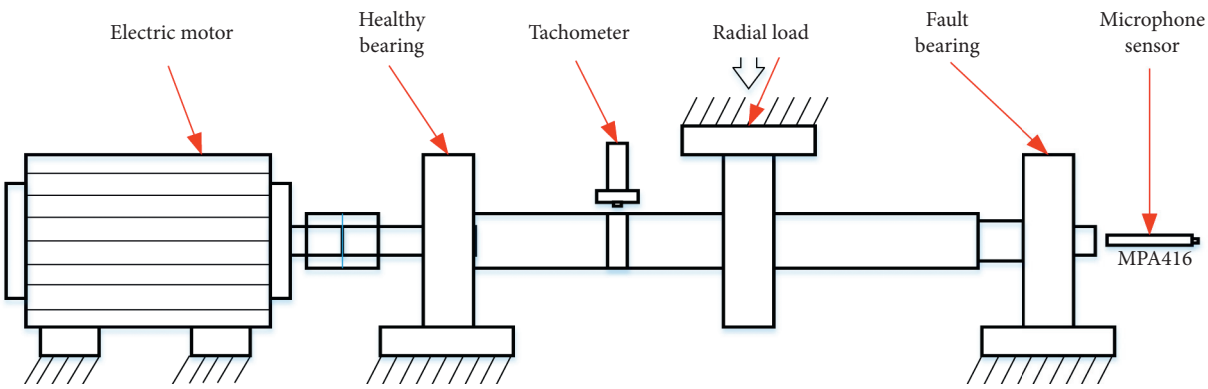


Figure 4: Rolling bearing fault test rig.

Table 1: Basic parameters of NSK NU205EW bearing.

Bearing pitch diameter	Roller diameter	Ball number	Contact angle
39	7.5	12	0°

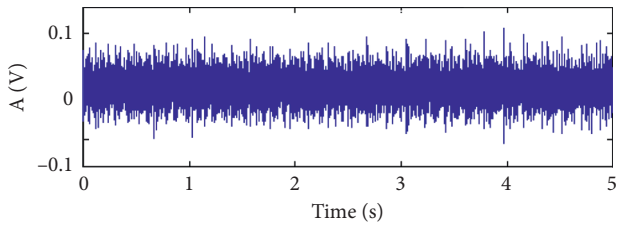


Figure 5: Time domain waveform of the original signal of experiment 1.

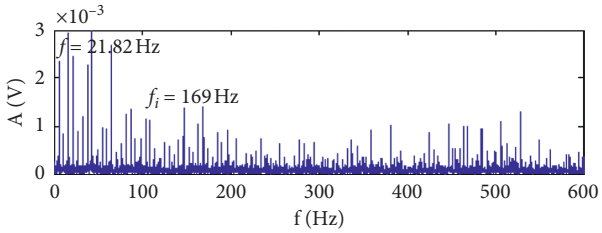


Figure 6: The original signal envelope of experiment 1.

weak fault characteristics and large noise interference. The following describes the specific process of signal processing using the method proposed in this paper.

The first reconstruction filter is performed. EWT is used to decompose the signal, and Figure 7 is the boundary detection graph. From the figure, EWT can detect 4 boundaries in the Fourier spectrum and divide the signal into 5 parts. By constructing an empirical wavelet filter bank, band-pass filtering is carried out on these 5 parts, and 5 components are finally obtained. Figure 8 shows the envelope spectra corresponding to the 5 components from top to bottom, and the corresponding correlation coefficients and kurtosis are given in Table 2.

According to the correlation coefficient, we can get that component 5 contains the richest information, which can also be seen from its corresponding envelope spectrum. At the same time, its envelope spectrum can clearly find the inner ring fault characteristic frequency and its multiplier, but it still cannot find the outer ring fault characteristic frequency. It also proves that primary filtering is insufficient for this early weak fault diagnosis. Meanwhile, the kurtosis of component 5 also reflects that it contains some fault pulse information.

Component 2 has the highest kurtosis, and the corresponding envelope spectrum can also find the inner ring fault characteristic frequency, rotation frequency, and its doubling frequency more clearly, so it can be judged that component 2 contains valid fault pulse information.

Based on the above two points, we choose the combination of components 2 and 5 as the optimal mode component. Through the above steps, part of the interference noise in the original signal has been removed, and a part of the fault features have been highlighted, laying a good foundation for the next processing.

Then input the reconstructed signal to the output layer after filtering, and set $K = 5$ and $\alpha = 2000$ for VMD [31]. Figure 9 shows the envelope spectrum of the BLIMF component obtained after the decomposition of VMD.

Careful analysis of the envelope spectra of the above 5 components shows that, in component 5, the fundamental frequency and frequency multiplication of the rotation frequency, the characteristic frequency of the inner ring fault, the characteristic frequency of the outer ring fault, and the theoretical calculation can be clearly found. There is almost no difference from the theoretically calculated frequency. Corresponding side bands are also found near the inner ring fault frequency, and the entire frequency spectrum contains almost no other noise interferences. The experimental results fully reflect the accuracy of the proposed method for early weak fault diagnosis.

3.2. Experimental Analysis of Full-Life Acceleration of Rolling Bearings

3.2.1. Experiment Design. Experiment 2 adopts the NSF I/UCR intelligent maintenance system center bearing life accelerated damage experiment data, which is used by many scholars for related research and is an internationally recognized data set. This data covers the whole process from the initiation of the fault to the failure, so it can reflect the damage characteristics of the actual early failure of the bearing and better reflect the effectiveness of the proposed method.

The motor transmits power to the shaft through a belt. Four bearings are mounted on the shaft. At the same time, a constant radial load of 6000 lbs is applied to the shaft. The speed is 2000 r/min, the bearing type is Rexnord ZA-2115 double row bearing, and the basic parameters are shown in Table 3. During operation, the bearing will be worn under heavy load and produce metal debris. A magnetic screw plug is installed on the test bench to collect metal debris. When

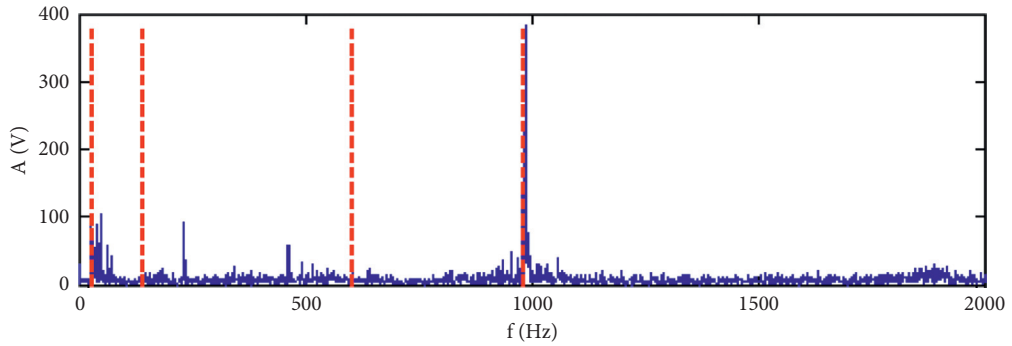


Figure 7: EWT boundary detection diagram of experiment 1.

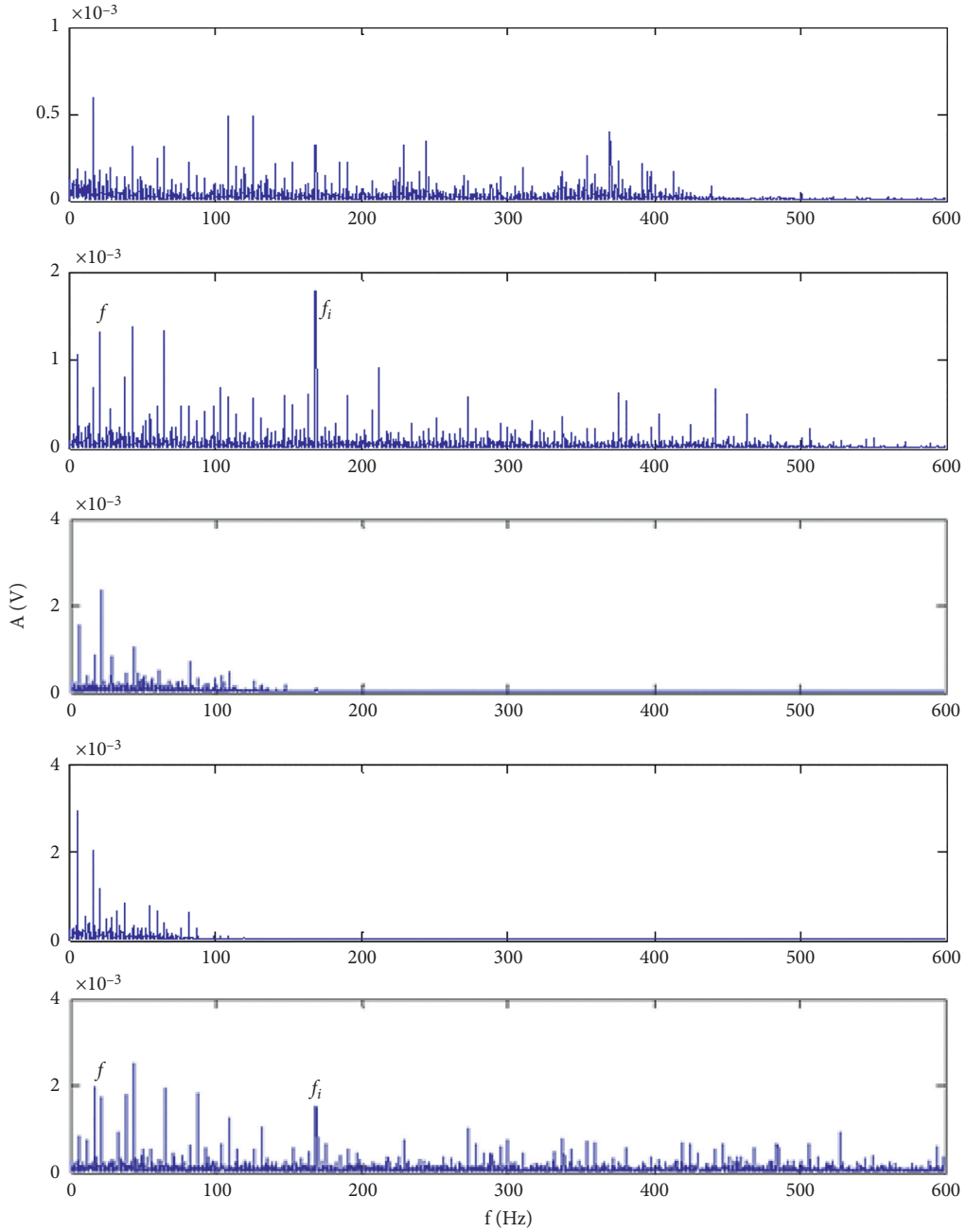


Figure 8: EWT component envelope spectra of experiment 1.

Table 2: Correlation coefficient and kurtosis of EWT component of experiment 1.

	Correlation coefficient	Kurtosis
Component 1	0.19	2.93
Component 2	0.36	4.54
Component 3	0.33	3.21
Component 4	0.41	3.51
Component 5	0.83	3.84

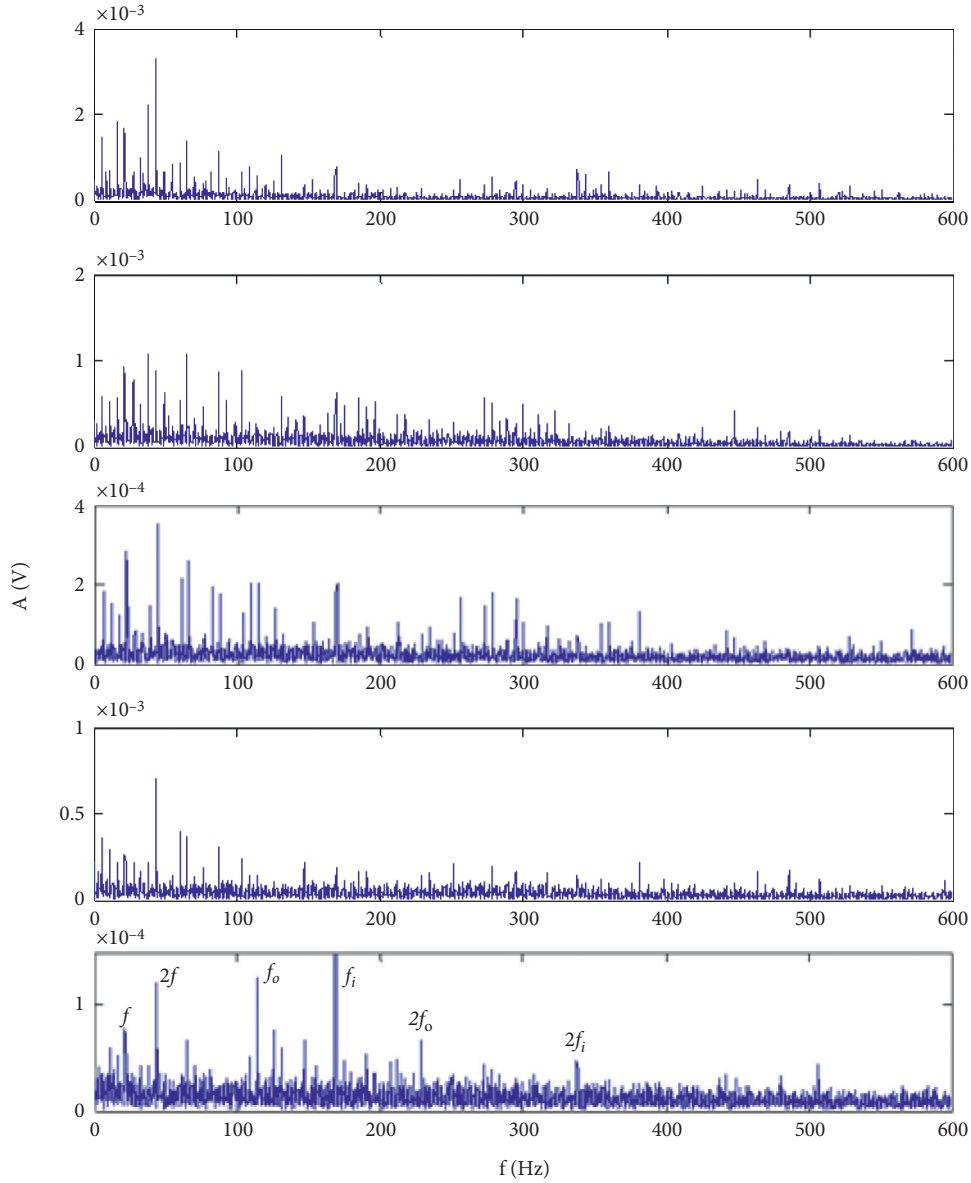


Figure 9: VMD component envelope spectrum of experiment 1.

the accumulation of metal debris exceeds a certain level, the electrical switch is automatically controlled to stop.

The experiment consists of three sets of data sets, and each set of data sets contains the complete degradation process data from the beginning of the bearing to the damage. In this part, the second data set is adopted. The sampling frequency is set at 20 kHz and the sampling time is 1s. A total of 20480 acceleration data points are collected in each sampling time, with a sampling interval of 10 min.

3.2.2. Experimental Results and Analysis. The rolling bearing fault test rig is shown in Figure 10. The change curve of the root mean square value of each collection point of the

Table 3: Basic parameters of Rexnord ZA-2115 double-row bearing.

Bearing pitch diameter (mm)	Roller diameter (mm)	Ball number	Contact angle
71.5	8.4	16	15.17°

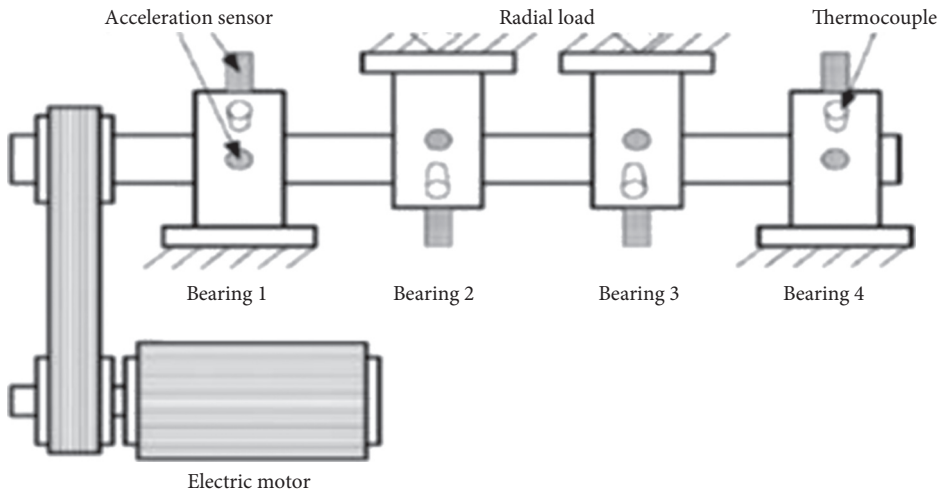


Figure 10: Rolling bearing accelerated degradation rig.

whole life data is shown in Figure 11. Since the RMS of the fault state and the normal state will be different under normal circumstances, most scholars regard the mutation points at T1 (about 5320 min) and T2 (about 6460 min) as the fault occurrence point. However, through the research in this article, it is found that the failure may have occurred earlier.

First of all, it can be found in Figure 11 that the signal RMS has not changed significantly around the 4520th minute. Analyzing its time domain waveform and envelope spectrum in Figures 12 and 13, no obvious periodic shock nor obvious fault characteristic information appears, which indicates that even if a fault occurs at this point, it is an extremely weak early fault, or no fault has occurred.

The experimental data were analyzed using the proposed method. The result of EWT boundary detection is shown in Figure 14. First, the signal is decomposed into 5 components in the first layer. Figure 15 shows the envelope spectrum of each component. In the envelope spectrum of these 5 components, no relevant fault characteristic information can be found, and a large number of interference frequencies are included. In component 3, a frequency component similar to the fault frequency of the outer ring was found, but it was not accurate (the fault frequency of the outer ring was 230.7 Hz, while the frequency was 232.5 Hz). Therefore, a layer of filter analysis could not effectively identify whether the bearing fault occurred and the type of fault. This also indicates that the actual fault at 4520 minutes in experiment 2 is weaker and more difficult to identify than the simulated fault in experiment 1.

The correlation coefficients and kurtosis corresponding to each component are given in Table 4. Through the comprehensive analysis of correlation coefficient, kurtosis, and envelope spectrum, component 5 is selected as the reconstructed signal after a layer filtering. Component 5 has a very high correlation coefficient, and it can be seen from its envelope spectrum that it contains sufficient frequency components. Meanwhile, its kurtosis value also reflects that the component may contain fault pulse information. Although component 1 has the highest kurtosis value, its correlation coefficient is too small, which may be the noise interference component. By removing component 1, some interference components are also removed correspondingly, which is also the purpose of the first layer of filtering.

Since the first layer of reconstruction filter analysis did not obtain very effective fault information, multilayer reconstruction filter analysis can be performed. After analyzing the one-to-three-layer filtering, it is found that the multilayer filter in this experiment is a certain improvement, but the most basic two-layer reconstruction filter can be used for accurate fault diagnosis, which will be described in detail below.

The filter reconstruction signal is input to the output layer. Figure 16 shows the envelope spectrum of the components obtained, and the corresponding kurtosis of each component is given in Table 5. In component 5, there is an extremely prominent fault characteristic frequency component; that is, the characteristic frequency of the outer ring fault is 230.7 Hz. At the same time, there is a relatively less obvious component of double frequency. At the same time, the kurtosis value of component 5 is 4.18, much larger than 3, which can also reflect that the bearing may have a fault impact. Figure 17 shows the envelope spectrum diagram of the 7050th minute T3 point in Figure 11. At this point, it can be clearly seen that the bearing has an obvious outer ring fault, and the fault characteristic frequency is 230.7 Hz, which is exactly the same as the fault characteristic frequency obtained by the method proposed in this paper after data

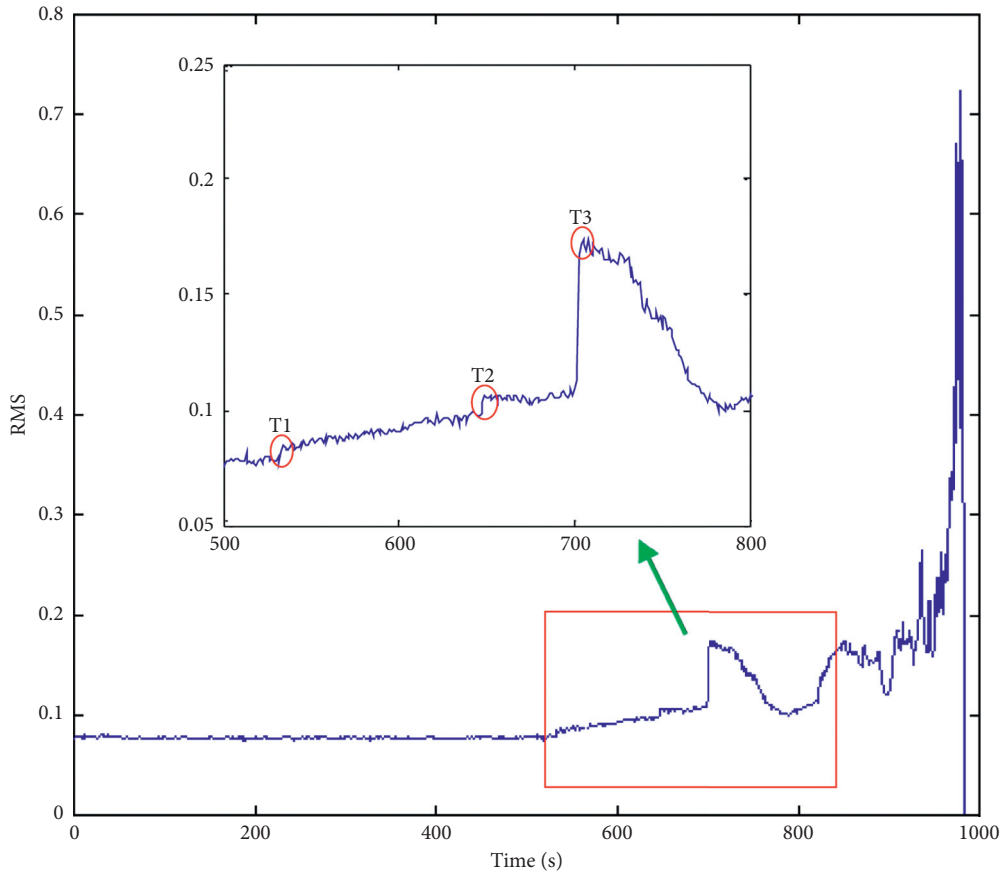


Figure 11: Root mean square value diagram of bearing life cycle.

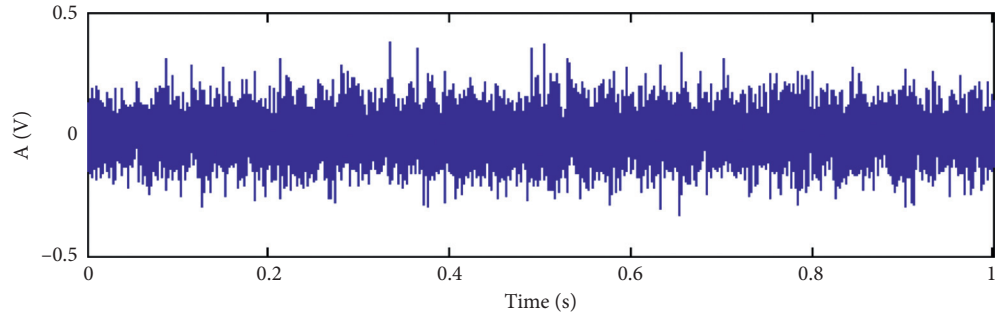


Figure 12: Data time domain waveform of the 4520th minute.

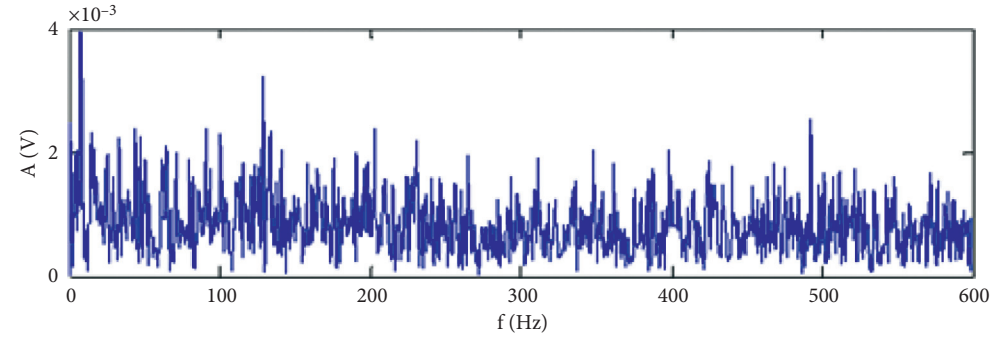


Figure 13: The data envelope spectrum at the 4520th minute.

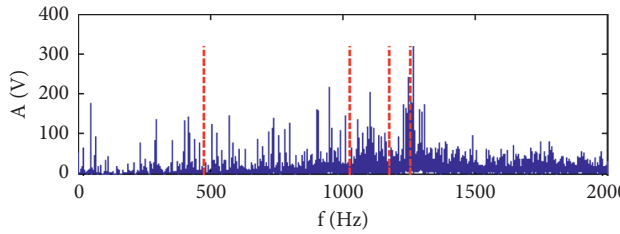


Figure 14: EWT boundary detection diagram of experiment 2.

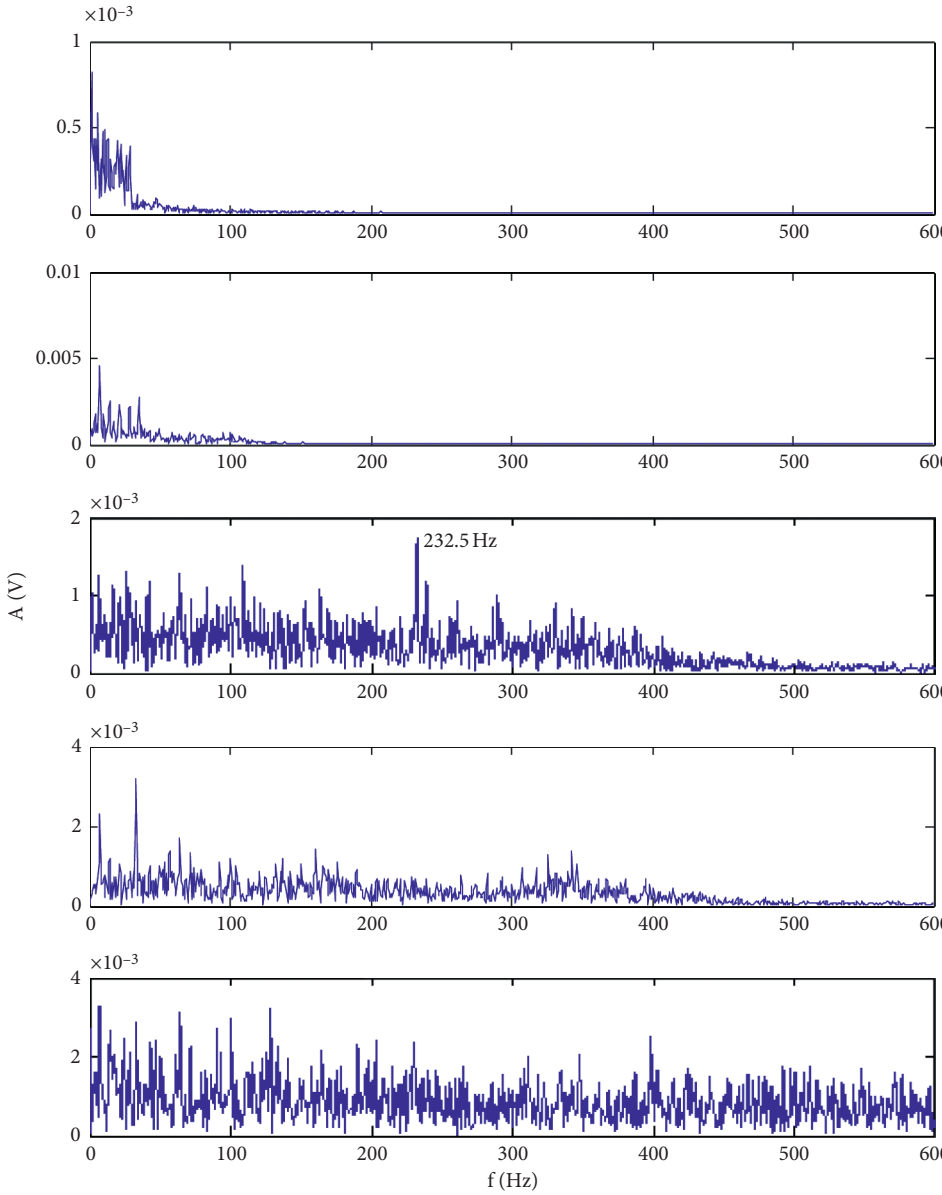


Figure 15: EWT component envelope spectrum of experiment 2.

Table 4: Correlation coefficient and kurtosis of EWT component of experiment 2.

	Correlation coefficient	Kurtosis
Component 1	0.06	4.60
Component 2	0.20	2.92
Component 3	0.23	2.98
Component 4	0.45	2.22
Component 5	0.94	3.55

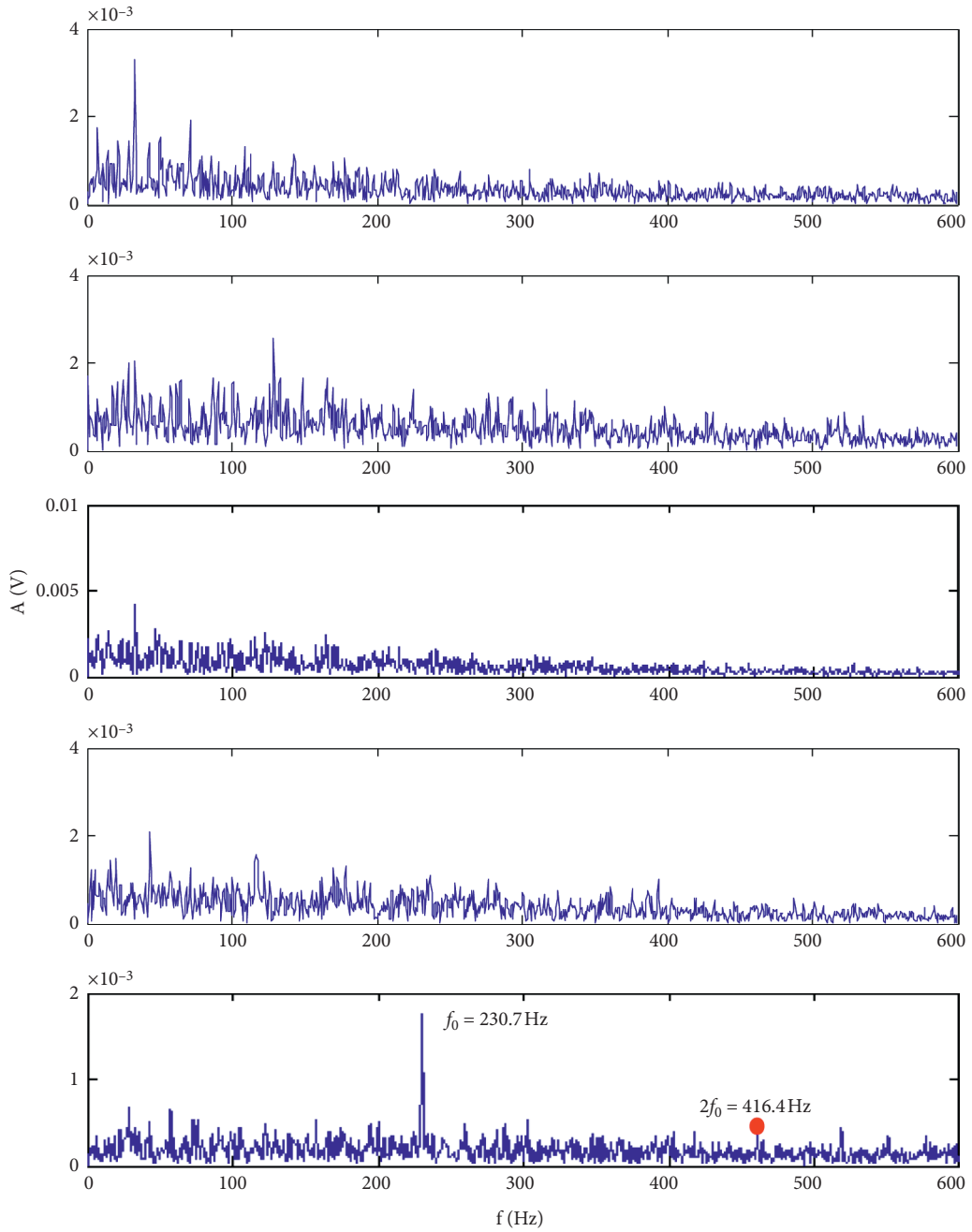


Figure 16: VMD component envelope spectrum of experiment 2.

Table 5: Correlation coefficient and kurtosis of EWT component of experiment 2.

	Component 1	Component 2	Component 3	Component 4	Component 5
Kurtosis	2.24	3.94	3.54	3.20	4.18

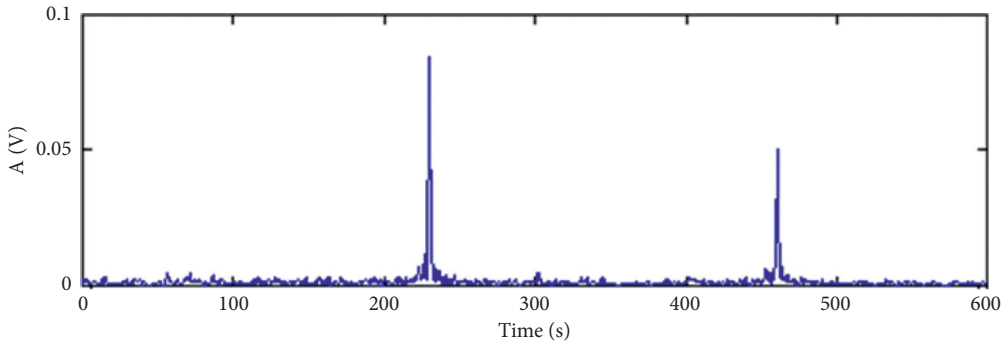


Figure 17: The data envelope spectrum at the 7050th minute.

processing of the 4520th minute. Based on the above information, it can be concluded that the bearing has had an early weak failure of the outer ring.

At the same time, we also analyzed the data at other time points. The data after 4520 minutes can effectively identify the fault features, so we will not make a detailed introduction here.

4. Conclusions

Aiming at the problem that the early weak fault characteristics of rolling bearing are weak, which is difficult to diagnose in time, this paper proposes a method for early weak fault diagnosis of rolling bearing based on multilayer reconstruction filter, which can effectively extract the fault characteristics from the acoustic and vibration signals, detect the fault earlier, and avoid serious consequences.

The early weak fault diagnosis method of rolling bearing based on multilayer reconstruction filter is proposed, which combines the characteristics of two advanced time-frequency analysis methods EWT and VMD, to gradually remove the interference information, highlight the fault pulse information, extract the early weak fault features, and accurately diagnose the fault. The proposed multilayer reconstruction filter method can be used to decompose the measured signal layer by layer to detect weaker fault characteristics, which means that the rolling bearing fault can be found earlier in the practical application, which is of great significance to the safe operation of the equipment.

In this paper, the validity of the proposed method is verified by the rolling bearing fault experiment and the full life cycle acceleration experiment data. Experiment results show that the proposed method can overcome the challenges brought about by low SNR and weak fault features, extract fault feature information earlier, and avoid the further evolution of the fault, resulting in incalculable losses. At the end of the paper, the future research work is prospected. In the proposed method, the parameters of some steps should be obtained by preprocessing the signal. Therefore, other parameter optimization methods can be integrated into the future work, so that the whole method can be adaptively selected to complete the early fault diagnosis.

Data Availability

The experimental data used in this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] G. Tang, Y. Wang, Y. Huang, N. Liu, and J. He, "Compound bearing fault detection under varying speed conditions with virtual multichannel signals in Angle domain," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 8, pp. 5535–5545, 2020.

[2] J. Fan and Z. C. Zhu, "An improved VMD with empirical mode decomposition and its application in incipient fault detection of rolling bearing," *IEEE Access*, vol. 6, no. 1, pp. 44483–44493, 2018.

[3] C. Qin, D. Wang, Z. Xu, and G. Tang, "Improved empirical wavelet transform for compound weak bearing fault diagnosis with acoustic signals," *Applied Sciences*, vol. 10, no. 2, pp. 682–692, 2020.

[4] J. J. Li, "*Research and application of the fault diagnosis of rolling bearing based on the sound signal*," M.S. thesis, Tiedao University, Shijiazhuang, China, 2017.

[5] A. Glowacz, "Fault diagnosis of single-phase induction motor based on acoustic signals," *Mechanical Systems and Signal Processing*, vol. 117, no. 1, pp. 65–80, 2019.

[6] A. Glowacz, W. Glowacz, Z. Glowacz et al., "Early fault diagnosis of bearing and stator faults of the single-phase induction motor using acoustic signals," *Measurement*, vol. 69, no. 8, pp. 5535–5545, 2018.

[7] A. Parey and A. Singh, "Gearbox fault diagnosis using acoustic signals, continuous wavelet transform and adaptive neuro-fuzzy inference system," *Applied Acoustics*, vol. 147, no. 1, pp. 133–140, 2019.

[8] D. C. Zhang and E. Stewart, "Intelligent acoustic-based fault diagnosis of roller bearings using a deep graph convolutional network," *Measurement*, vol. 156, no. 1, 2020.

[9] X. W. Liu and D. L. Pei, "Acoustic signal based fault detection on belt conveyor idlers using machine learning," *Advanced Powder Technology*, vol. 31, no. 7, 2020.

[10] G. Tang, Y. J. Huang, and Y. T. Wang, "Fractional frequency band entropy for bearing fault diagnosis under varying speed conditions," *Measurement*, vol. 171, no. 1, 2021.

[11] G. Tang, Y. Wang, Y. Huang, and H. Wang, "Multiple time-frequency curve classification for tacho-less and resampling-less compound bearing fault detection under time-varying speed conditions," *IEEE Sensors Journal*, vol. 21, no. 4, p. 1, 2020.

[12] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 108–126, 2013.

[13] S. Chen, Y. Yang, Z. Peng, S. Wang, W. Zhang, and X. Chen, "Detection of rub-impact fault for rotor-stator systems: a novel method based on adaptive chirp mode decomposition," *Journal of Sound and Vibration*, vol. 440, no. 1, pp. 83–99, 2019.

[14] S. Q. Wang, C. Y. Wang, Q. Sun, C. Y. Gao, and S. G. Yang, "Radar emitter signal intra-pulse feature extraction based on empirical mode decomposition," *Procedia Computer Science*, vol. 154, no. 1, pp. 504–507, 2019.

[15] S. D. C. Kiat and E. Y. K. Ng, "A computational intelligence tool for the detection of hypertension using empirical mode decomposition," *Computers in Biology and Medicine*, vol. 118, 2020.

[16] S. Chen, Y. Yang, Z. Peng, X. Dong, W. Zhang, and G. Meng, "Adaptive chirp mode pursuit: algorithm and applications," *Mechanical Systems and Signal Processing*, vol. 116, no. 1, pp. 566–584, 2019.

[17] D. Wang, K.-L. Tsui, and Y. Qin, "Optimization of segmentation fragments in empirical wavelet transform and its applications to extracting industrial bearing fault features," *Measurement*, vol. 133, no. 1, pp. 328–340, 2019.

[18] T. Gong, X. Yuan, X. Lei, and X. Wang, "Application of tentative variational mode decomposition in fault feature detection of rolling element bearing," *Measurement*, vol. 135, no. 1, pp. 481–492, 2019.

[19] K. Mourad and M. Thomas, "A comparative study between empirical wavelet transforms and empirical mode decomposition methods: application to bearing defect diagnosis," *Mechanical Systems and Signal Processing*, vol. 81, no. 1, pp. 88–107, 2016.

[20] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.

[21] W. Liu, S. Cao, and Y. Chen, "Seismic time-frequency analysis via empirical wavelet transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 28–32, 2016.

[22] J. Hu and J. Wang, "Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression," *Energy*, vol. 93, no. 1, pp. 1456–1466, 2015.

[23] J. L. Chen and J. Pan, "Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals," *Renewable Energy*, vol. 89, no. 89, pp. 80–92, 2020.

[24] J. Pan, J. Chen, Y. Zi, Y. Li, and Z. He, "Mono-component feature extraction for mechanical fault diagnosis using modified empirical wavelet transform via data-driven adaptive Fourier spectrum segment," *Mechanical Systems and Signal Processing*, vol. 72-73, no. 1, pp. 160–183, 2016.

[25] S. Chen, X. Dong, Z. Peng, W. Zhang, and G. Meng, "Nonlinear chirp mode decomposition: a variational method," *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 6024–6037, 2017.

[26] G. Li, G. Tang, G. Luo, and H. Wang, "Underdetermined blind separation of bearing faults in hyperplane space with variational mode decomposition," *Mechanical Systems and Signal Processing*, vol. 120, no. 1, pp. 83–97, 2019.

[27] X. Jiang, J. Wang, J. Shi, C. Shen, W. Huang, and Z. Zhu, "A coarse-to-fine decomposing strategy of VMD for extraction of weak repetitive transients in fault diagnosis of rotating machines," *Mechanical Systems and Signal Processing*, vol. 116, no. 1, pp. 668–692, 2019.

[28] X. B. Wang and Z. X. Yang, "Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery," *IEEE-ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 68–79, 2018.

[29] Z. Wang, G. He, W. Du et al., "Application of parameter optimized variational mode decomposition method in fault diagnosis of gearbox," *IEEE Access*, vol. 7, no. 1, pp. 44871–44882, 2019.

[30] Z. Xu, C. Qin, and G. Tang, "A novel deconvolution cascaded variational mode decomposition for weak bearing fault detection with unknown signal transmission path," *IEEE Sensors Journal*, vol. 21, no. 2, p. 1746, 2021.

[31] H. D. Wang, S. E. Deng, J. X. Yang et al., "Incipient fault diagnosis of rolling bearing based on VMD with parameters optimized," *Journal of Vibration and Shock*, vol. 39, no. 23, pp. 38–46, 2020.

Hindawi
Shock and Vibration
Volume 2021, Article ID 5534702, 16 pages
https://doi.org/10.1155/2021/5534702

Research Article

The Piston Ring-Cylinder Bore Interface Leakage of Bent-Axis Piston Pumps Based on Elastohydrodynamic Lubrication and Rotation Speed

Lvjun Qing,[1,2] **Lichen Gu**,[1,3] **Yan Wang**,[1] **Wei Xue**,[4] **and Zhufeng Lei**[2]

[1]*National Engineering Laboratory for Highway Maintenance Equipment, Chang'an University, Xi'an 710064, China*
[2]*National Joint Engineering Research Center for Special Pump Technology, Xi'an Aeronautical University, Xi'an 710077, China*
[3]*Institute of Electromechanical Technology, Xi'an University of Architecture and Technology, Xi'an 710055, China*
[4]*Institute of Geology of the Fifth Oil Production Plant, PetroChina Changqing Oilfield Company, Xi'an 710000, China*

Correspondence should be addressed to Lichen Gu; gulichen@126.com

Received 19 January 2021; Revised 12 February 2021; Accepted 16 February 2021; Published 22 February 2021

Academic Editor: Ke Feng

The bent-axis piston pump is the core component of electrohydrostatic actuators (EHA) in aerospace applications, and its wear of key friction interfaces is greatly related to the healthy operation of pumps. The leakage of the piston ring-cylinder bore interface (PRCB), as the important part of the return oil flow of the pump house that commonly assesses the wear of key friction interfaces in piston pumps, is changed with the rotation speed. Thus, the wear of key friction interfaces is usually inaccurate by using the leakage of PRCB. In order to obtain the relationship between the PRCB leakage and the rotation speed, an elastohydrodynamic lubrication model is proposed. First, the proposed model includes a minimum film thickness model of PRCB to analyze the dynamic change of oil film of PRCB when subject to the elastohydrodynamic lubrication. After that, a mathematical model of PRCB is induced by combining the minimum film thickness model with the flow equation, which helps produce the effects of the oil film on the leakage of PRCB. The proposed model is verified by numerical simulation and experiment. The results show that the leakage of PRCB has a negative effect on the return oil flow of the pump case in the range of rotation speed of 700–1300 r/min and discharge pressure of 10–20 MPa. Furthermore, the leakage of PRCB is proportional to the rotation speed, but the return oil flow of the pump case is decreased. The effects of rotation speed are enhanced under the high discharge pressure conditions.

1. Introduction

Bent-axis piston pumps are widely used in electrohydrostatic actuators (EHA) in aerospace applications due to the advantage of miniaturization and lightweight through high speed and high pressure. Therefore, the bent-axis piston pump greatly affects the healthy operation of EHA [1–3].

Performance degradation prediction and fault diagnosis are of great significance to ensure the safe and stable operation of mechanical equipment and reduce maintenance cost [4, 5]. In hydraulic pumps, the wear of key friction pairs is a major cause of the degradation of hydraulic pump performance [6–9]. As the wear increases with pump working time, the lubricating interface gap increases and the lubricating interface leakage increases, which eventually leads to the increase of the return oil flow rate of the pump case [10, 11]. Hence, the return oil flow rate is often used to evaluate the wear state of the key friction pairs in the piston pump [12–14]. In engineering scenarios, however, the lubricating interface leakage is inevitably changed with the varying working conditions, such as the rotation speed, which affects the accurate evaluation of wear states.

The bent-axis piston pump is shown in Figure 1. The cylinder block moves synchronously with the drive shaft through the bevel gears, and the piston reciprocates in the cylinder bore. The piston ring with the spherical cross-section is usually installed on the end of the piston to reduce the leakage between the piston and the cylinder bore. There are two lubricating interfaces, that is, the valve plate-cylinder block interface (VPCB) and the piston ring-cylinder bore

interface (PRCB). As one of the two lubricating interface leakages, PRCB leakage directly affects the return oil flow rate in bent-axis piston pump.

The lubrication interface gap, as the main parameter affecting PRCB leakage, is influenced by multifactor interactions. Pelosi et al. developed the fully coupled multibody dynamics model to capture the complex fluid-structure interaction phenomena affecting the nonisothermal fluid film conditions and analyzed the influence of the squeeze film effect and elastic deformation on the piston-cylinder interface [15]. In order to improve energy efficiency of piston pump, Song et al. presented a numerical method to analyze influence of temperature on lubricant film characteristics of the piston-cylinder interface [16]. Qian et al. presented a nonisothermal fluid-structure mathematical model considering the piston eccentricity, elastic deformations, and the fluid physical properties to analyze the piston-cylinder bore interface leakage and the fluid pressure distribution [17]. Zhang et al. established the mathematical model of the oil film thickness by using the cosine theorem in the cross section of the piston and studied the piston-cylinder interface leakage and pressure field under the condition of ultrahigh pressure [18]. Nie et al. established a parameterized elastohydrodynamic lubrication model and studied the influence of the elastohydrodynamic behavior, viscosity temperature effect, and deep-sea environmental pressure on piston-cylinder interface leakage in seawater hydraulic axial piston pump [19]. Cho et al. analyzed the film thickness, the pressure distribution, and the friction force of piston ring-cylinder bore interface under elastic deformation of plunger ring in bent-axis piston pump with the tapered piston [20]. Kumar established the theoretical model of the piston ring-piston bore pair in the bent-axis piston pump with rectangular cross-section of piston ring and analyzed the leakage flow characteristics between piston ring and cylinder bore under different working conditions through numerical simulation and test [21]. Manhartsgruber analyzed the dynamics of the leakage of the bent-axis piston pump with the tapered piston and revealed the influence of the piston rotation inside the cylinder bore on the pressure fluctuation under low speed and low pressure condition by experiment [22]. Zhang et al. analyzed the effect of centrifugal force and reciprocation inertial force on the leakage between piston and cylinder in the bent-axis piston pump under high speed and show the leakage of piston-cylinder pair increases obviously at high speed [23].

From the literature review, the previous studies mentioned mainly focused on the influence of the piston eccentricity, elastic deformation, and thermal effect on the piston-cylinder bore interface leakage in the swash-plate piston pump. Some researchers studied the effects of the elastic deformation and micromotion of piston ring on the oil film thickness, pressure field, and friction force of PRCB in the bent-axis piston pump under certain rotation speed conditions. However, the variable speed and wide speed are the typical working conditions of bent-axis piston pump in EHA, and the influence of rotation speed on PRCB leakage considering elastohydrodynamic lubrication effect has not been fully elucidated.

Rotation speed and its fluctuation are very meaningful monitoring parameters, and the leakage analysis of the friction pairs is also related to the rotation speed. Our previous research has proved that the operating state information of the swash-plate piston machine can be extracted from the speed fluctuation signal [24–27]. In this study, an elastohydrodynamic lubrication model is proposed to discuss the influence of rotation speed on the leakage of PRCB. The proposed model includes a minimum oil film model of PRCB to analyze the dynamic change of minimum oil film of PRCB under the elastohydrodynamic lubrication. Combining the minimum oil film model with the flow equation, a mathematical model of PRCB is induced to help produce the effects of the oil film on the leakage of PRCB. The accuracy of the proposed model is verified by simulation and experiment, which provides guidance on the piston pump wear diagnosis based on the leakage. The layout of this paper is organized as follows. A theoretical model of PRBC is established and explained in detail in Section 2. The simulation results are discussed in Section 3. The experimental results are presented in Section 4, followed by the conclusions in Section 5.

2. Elastic Deformation and Leakage Mechanism of PRBC

2.1. Elastohydrodynamic Lubrication Equations. In order to reduce the leakage of PRCB, a ring groove is usually opened on the spherical surface of the large end of the piston, and a piston ring is installed in the ring groove. The boundary conditions between the piston ring and the cylinder bore are different because of the change of the angle position of the piston in the suction and discharge region, as shown in Figure 2.

Because of the elastic deformation of the contact surface in the piston ring at relative motion, the gap of PRCB is changed, which affects the leakage. Therefore, it is necessary to access elastohydrodynamic lubrication (EHL) to analyze the leakage between the piston ring and the cylinder bore. The analysis model between the piston ring and the cylinder bore is shown in Figure 3. V represents the sliding velocity of the piston and the piston ring relative to the cylinder bore, R is the radius of the piston ring section, W is the unit load on the piston ring. Figures 3(a) and 3(b) show the pressure distribution of lubricating oil film between the piston ring and the cylinder bore, and its typical characteristics are the primary peak pressure and the secondary peak pressure. P_1 represents the initial pressure of the oil film, and it equals the pressure in the displacement chamber, and P_2 represents the pressure at the end of the film, and it equals the pressure in the pump case.

The assumptions for the EHL analysis between the piston ring and the cylinder bore are listed as follows:

(1) The lubricating pressure remains unchanged along the direction of film thickness

(2) The lubricating oil film remains unchanged along the circumferential direction of piston bore

(3) The fluid inertia force is ignored

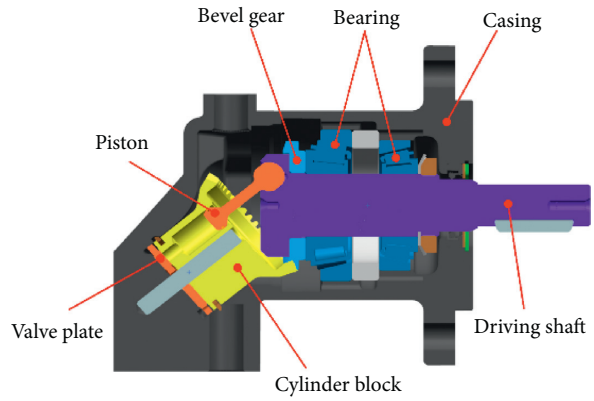


Figure 1: The configuration of a bent-axis piston pump.

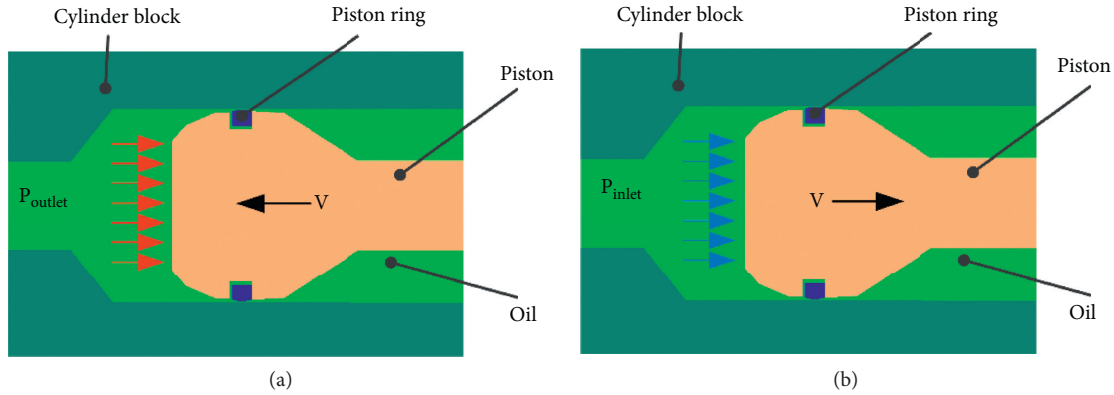


Figure 2: The behavior of PRCB. (a) The discharge region; (b) the suction region.

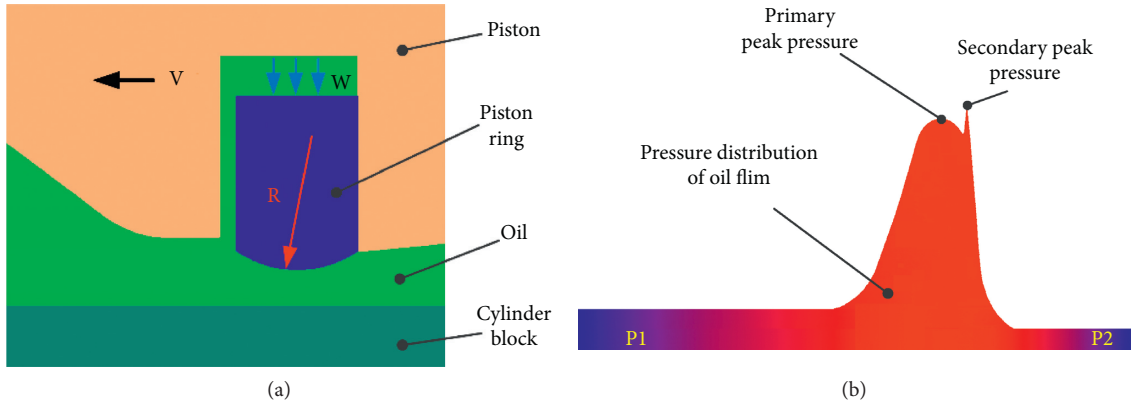


Figure 3: The analysis model of PRCB. (a) The configuration of PRCB; (b) the pressure distribution of PRCB.

(4) The compressibility and viscosity temperature characteristics of the lubricating oil film are ignored

(5) The only concern is the velocity gradient along the piston ring direction of motion

The Reynolds equation for one-dimensional isothermal elastohydrodynamic lubrication is shown as follows [28]:

$$\frac{d}{dx}\left(\frac{\rho h^3}{\eta}\frac{dp}{dx}\right) = 12u_s\frac{d(\rho h)}{dx}, \quad (1)$$

where ρ is oil film density, h is the oil film thickness, p is the pressure distribution of lubricating oil film, $p = p_1$ at $x = x_a$, $p = p_2$, $\partial p/\partial x = 0$ at $x = x_b$, u_s is the average oil film velocity, $u_s = v/2$, and η is the viscosity oil film.

The velocity of the piston relative to the cylinder bore can be expressed as follows:

$$v = -R_0\omega \sin\gamma \sin\varphi, \tag{2}$$

where R_0 is the nominal pitch circle radius of piston balls on the driving flange, ω is the angular velocity of the driving shaft, φ is the angular position of piston relative to the top dead point, γ is the angle between the axis of the cylinder and axis of the driving shaft.

The film thickness is expressed as follows [29]:

$$h(x) = h_0 + \frac{x^2}{2R} - \frac{2}{\pi E'} \int_{x_0}^{x_e} p(s) \ln (s-x)^2 ds + C, \tag{3}$$

where $1/E' = 1/2\,(1-\mu_1^2/E_1 + 1 - \mu_2^2/E_2)$, μ_1, $E^{'}$ is the equivalent Young's modulus, μ_1 and μ_2 are Poisson's ratio of the piston ring and the cylinder, E_1 and E_2 are Young's modulus of the piston ring and the cylinder, h_0 is the film thickness at $x = 0$, and C is the integral calculus constant.

The viscosity coefficient and density coefficient of lubricating oil are expressed as follows [30]:

$$\eta = \eta_0 \exp\left\{(\ln \eta_0 + 9.67)\left[-1 + \left(\frac{1+p}{p_0}\right)^z\right]\right\}. \tag{4}$$

The density coefficient of lubricating oil is expressed as follows [31]:

$$\rho = \rho_0\left(1 + \frac{0.6p}{1+1.7p}\right), \tag{5}$$

where η_0 is the viscosity of oil under atmospheric pressure, p_0 and z are the pressure-viscosity coefficients, and ρ_0 is the density of oil under atmospheric pressure.

The unit load on piston ring is expressed as follows:

$$\omega - \int_{x_a}^{x_b} p(x)dx = 0, \tag{6}$$

where x_a is the starting point of lubricating film in the analysis and x_b is the ending point.

2.2. PRCB Leakage Model. The PRCB leakage is composed of two parts. One is the Poiseuille flow by the pressure difference between the displacement chamber and the pump case, and the other is the Couette flow by the piston ring relative to the cylinder bore motion and the hydraulic oil viscosity. As the most important parameter of PRCB, the minimum oil film varies dynamically with the piston velocity due to the elastohydrodynamic lubrication effect, which will change the proportion of the Poiseuille flow and the Couette flow in the gap and then affect the leakage flow of PRCB. The dynamic model of the leakage flow of PRCB is established by combining the minimum oil film of PRCB with the flow equation, which is the theoretical basis for the subsequent study of the leakage flow of the plunger pair under the variable speed condition.

The instantaneous leakage flow rate between the piston ring and the cylinder bore is expressed as follows [23]:

$$q_{lpi} = \frac{\pi d_c \delta_p^3\left(1 + 1.5\varepsilon^2\right)}{12\mu l_i}\Delta p - \frac{\pi d_c \delta_p}{2} v, \tag{7}$$

where d_c is the diameter of piston, δ_p is the gap between piston ring and cylinder bore, Δp is the pressure difference between displacement chamber and pump case, μ is the oil viscosity, l_i is the width of piston ring, ε is the eccentricity of the piston ring, and v is the velocity of motion of piston ring in the cylinder bore.

The average leakage flow rate of PRCB is expressed as

$$\overline{q}_{lpi} = \frac{N}{T_0}\int_0^{T_0} \frac{\pi d_c \delta_p^3\left(1 + 1.5\varepsilon^2\right)}{12\mu l_i}\Delta p dt - \frac{N}{T_0}\int_0^{T_0} \frac{\pi d_c \delta_p v_{pi}}{2} dt, \tag{8}$$

where N is number of the pistons and T_0 is the time during one rotating speed of the bent-axis piston pump.

3. Simulation Results

3.1. Simulation Parameter Setting. The finite difference method is used to solve the theoretical model. In order to improve the calculation accuracy, the central difference scheme is used to discretize the Reynolds equation. The flow chart of the numerical calculation program for the elastohydrodynamic lubrication problem of PRCB is shown in Figure 4.

According to the given working parameters and the initial values of oil film thickness, distribution pressure, viscosity, density and other parameters, the Reynolds equation is calculated to obtain the new oil film distribution pressure, and then the elastic deformation, viscosity, density, and other parameters of the piston ring are calculated under the obtained pressure.

The Gauss–Seidel iterative method is used to solve the algebraic equations of Reynolds equation discretization, and the converging condition is expressed as follows:

$$p_i^{k+1} = (1-\alpha)p_i^k + \alpha\overline{p}_i^k, \tag{9}$$

where α is the relaxation factor and $\alpha = 0.8$ is chosen to obtain a fast convergence, p_i^k and p_i^{k+1} are the pressure values at steps k and $k+1$, and $\overline{p}_i^k$ is the temporary value at step $k+1$.

The whole process is repeated until the error tolerance of the Reynolds equation at each node is less than the relative precision requirement $\varepsilon = 1 \times 10^{-5}$; the error tolerance of the Reynolds equation ε is expressed as follows:

$$\left|\frac{p_i^{k+1} - p_i^k}{p_i^{k+1}}\right| \le \varepsilon. \tag{10}$$

MATLAB platform is used to calculate the leakage flow rate between piston ring and cylinder bore. In order to compare the characteristics of the leakage with different operating conditions, the parameters of simulation calculation process are set in Table 1.

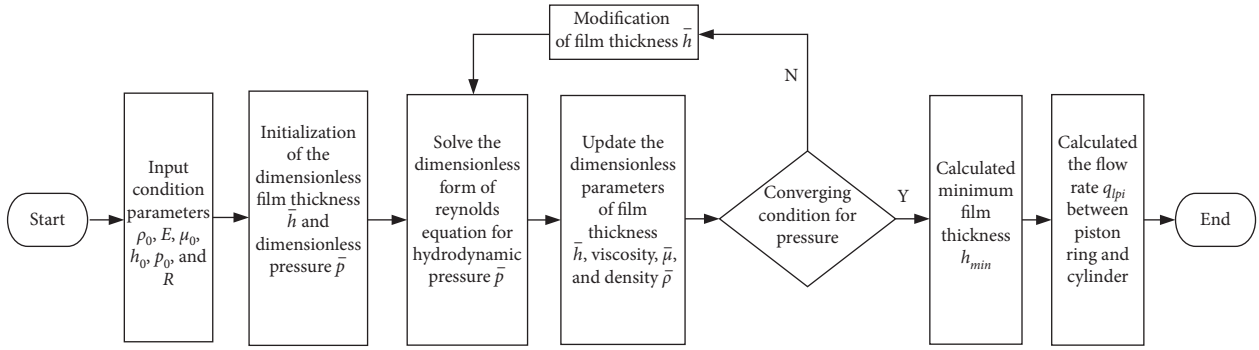


Figure 4: Flow chart of the calculation process.

Table 1: Simulation parameters setting.

Description	Numerical value	Unit
Suction pressure	0.1	MPa
Discharge pressure	10, 15, 20	MPa
Rotation speed	700, 800, 900, 1000, 1100, 1200, 1300	r/min
Oil viscosity (at 40°C)	0.046	Pa.s
Diameter of distribution circle in piston head	68.1	mm
Diameter of cylinder bore	16	mm
Number of pistons	7	
Swivel angle	40	deg
Width of piston ring	3.2	mm

3.2. Analysis of Simulation Results

3.2.1. The Pressure Distribution of PRCB. The pressure distribution of PRCB in the discharge region is shown in Figure 5.

As shown in Figure 5, the pressure of oil film changes in the lubrication zone at a certain angle position, and there is the maximum pressure value (the local maximum pressure). The local maximum pressure varies with the angular position of the discharge region, and there is also the maximum pressure (the global maximum pressure). As a whole, the local maximum pressure in the discharge region decreases first and then increases with the increase of angle position, and the global maximum pressure is 146 MPa at angle position $\varphi = 4.7^{\circ}$, while the angle position is near the top dead point 0° and the bottom dead point 180°; the local maximum pressure changes dynamically again. The reason is that the piston velocity is small near the top and bottom dead points, and the local maximum pressure is determined by the primary peak pressure, as shown in Figure 3(b). However, the local maximum pressure is determined by the secondary peak pressure in other region, and the secondary peak pressure decreases with the increase of the piston velocity.

The local maximum pressure distribution of PRBC in the discharge region at different speeds is shown in Figure 6.

As shown in Figure 6, when the speed increases from 700 r/min to 1300 r/min with a step of 200 r/min, the global maximum pressure is basically not affected by the change of speed, which is kept at 146 MPa, but the angular position of the global maximum pressure decreases from 4.7° to 2.5°.

The local maximum pressure distribution of PRBC in the discharge region at different discharge pressure is indicated in Figure 7.

As indicated in Figure 7, when the discharge pressure increases from 10 MPa to 20 MPa at $n = 700$ r/min, the local maximum pressure becomes larger as a whole, the global maximum pressure increases from 146 MPa to 205 MPa, and the angular position of the global maximum pressure increases from 4.7° to 15.3°.

The rotation speed has a weak effect on the value of the global maximum pressure, which has a significant effect on its the angular position. The discharge pressure has important influence on the value and the angle position of the global maximum pressure. The reason is that the global maximum pressure of PRBC in the discharge area is determined by the secondary peak pressure of the lubricating oil film. The secondary peak pressure is positively correlated with the discharge pressure. The angle position of the secondary peak pressure is affected by the rotation speed and the discharge pressure. The angular position of the secondary peak is positively correlated with the discharge pressure, and it is negatively correlated with the rotation speed.

3.2.2. The Minimum Thickness of PRCB. The minimum thickness of PRCB at different speed is displayed in Figure 8.

As displayed in Figure 8, the minimum film thickness increases first and then decreases with the increase of angle position in the discharge region and the suction region. However, the minimum film thickness of PRCB in the discharge region is obviously larger than that in the suction region as a whole. When the speed increases from 700 r/min to 1300 r/min with a step of 100 r/min, the minimum thickness increases in the discharge region and the suction region, and its increase in the discharge region is greater.

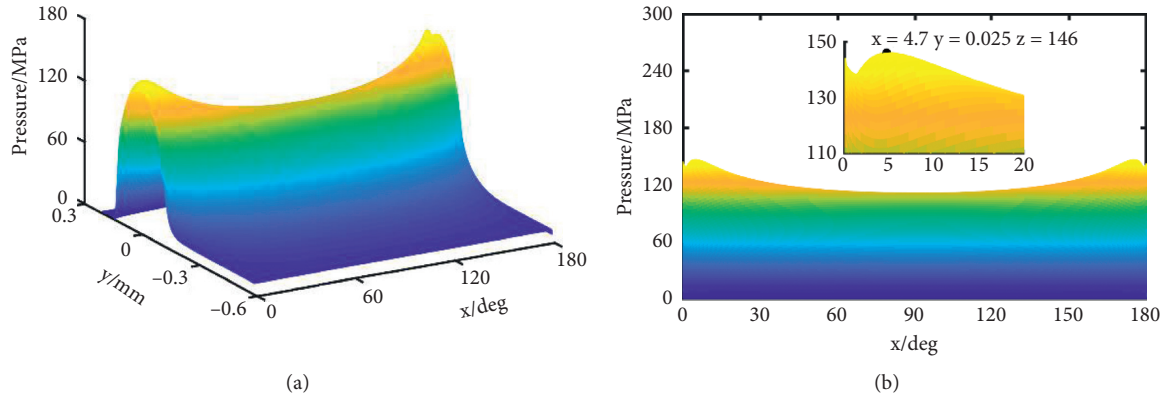


Figure 5: The pressure distribution of PRCB at $n = 700$ r/min $p = 10$ MPa. (a) 3D diagram of oil film pressure distribution. (b) Right view of oil film pressure distribution.

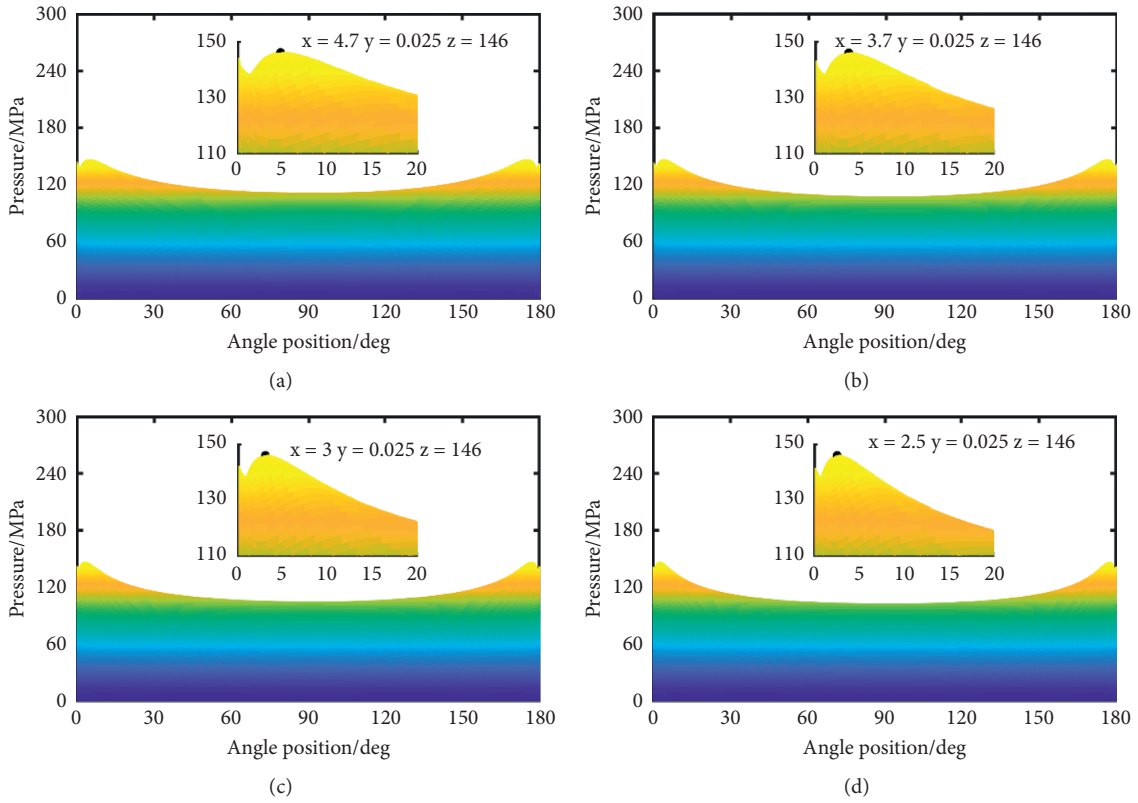


Figure 6: The effect of speed on the local maximum pressure distribution of the oil film at $p = 10$ MPa. (a) $n = 700$ r/min; (b) $n = 900$ r/min; (c) $n = 1100$ r/min; (d) $n = 1300$ r/min.

The effect of discharge pressure on the minimum thickness of PRCB is shown in Figure 9.

As shown in Figure 9, the minimum thickness of PRBC in the discharge region and the suction region increases first and then decreases with the increase of angular position, but the former is larger than the latter as a whole. When the discharge pressure increases gradually from 10 MPa to 20 MPa with a step of 5 MPa, the minimum oil film increases obviously in the oil drainage region, but only slightly in the suction region.

The reason is that the minimum thickness is determined by the velocity and the load of the piston ring. When the rotation speed and the discharge pressure are constant, the velocity of piston ring increases first and then decreases from the top dead point (TDC) to the bottom dead point (BDC) in the drainage region and is the same change trend from BDC to TDC in the suction region, so the change law of minimum thickness is the same as the velocity of piston ring. The discharge pressure mainly affects the load of the piston ring in the discharge region, and the discharge pressure increases,

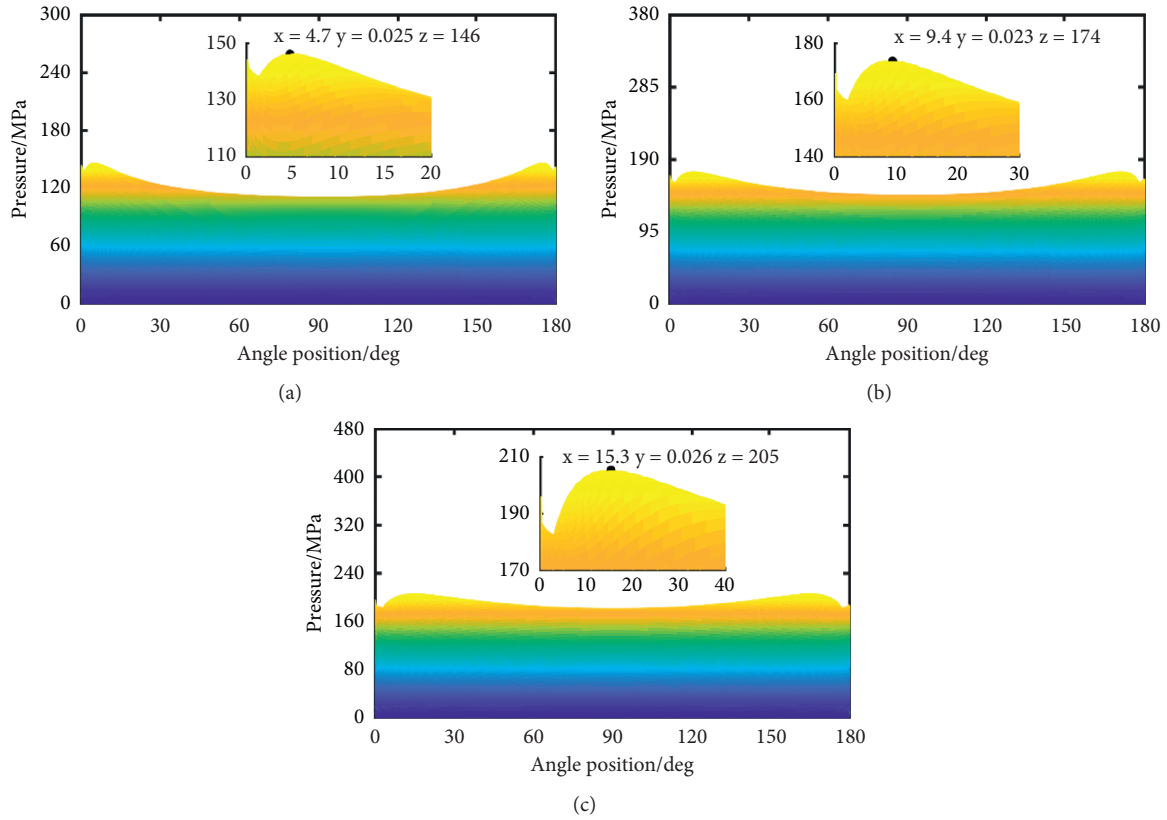


Figure 7: The effect of discharge pressure on the local maximum pressure distribution at $n = 700$ r/min. (a) $p = 10$ MPa; (b) $p = 15$ MPa; (c) $p = 20$ MPa.

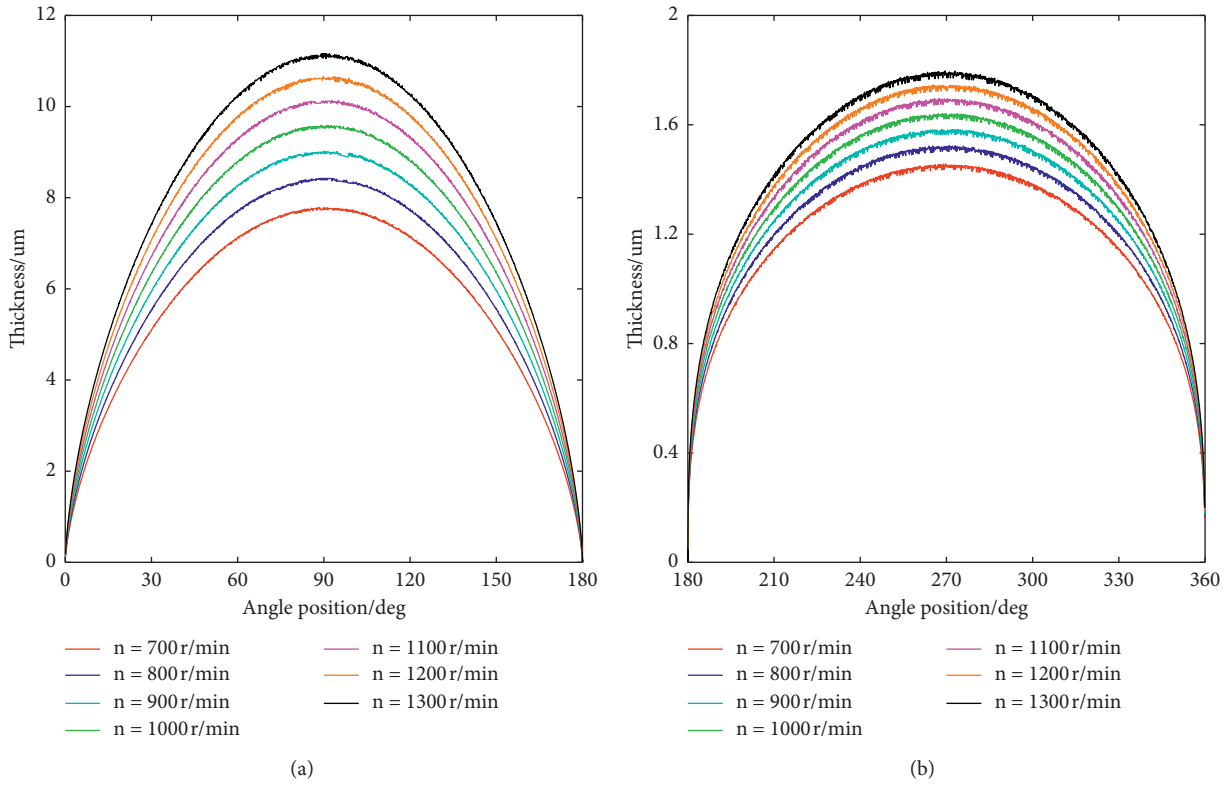


Figure 8: The minimum thickness of PRCB at different speeds. (a) The discharge region; (b) the suction region.

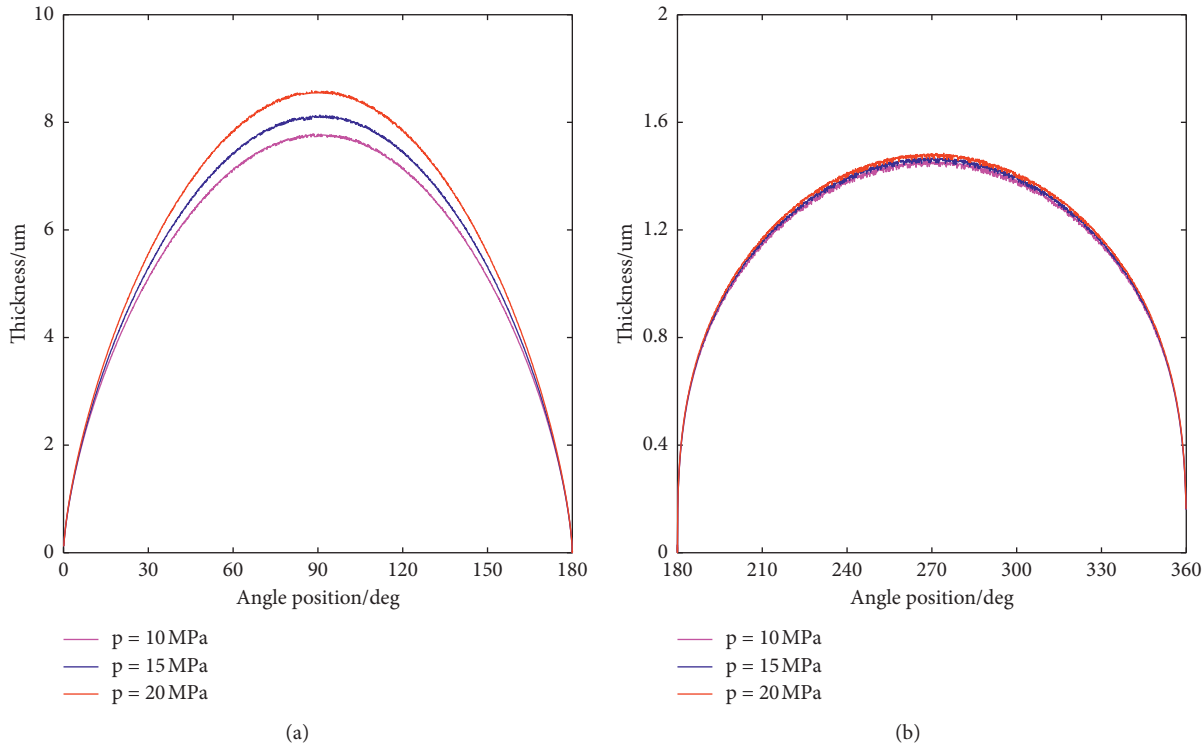


Figure 9: The minimum thickness at different pressure. (a) The discharge region; (b) the suction region.

and the load of the piston ring becomes larger, which leads to the increase of the elastic deformation of the piston ring, so the minimum thickness in the discharge region is greater than that in the suction region.

The mean value of minimum thickness of PRCB is shown in Figure 10.

As shown in Figure 10, it is found that the mean value of the minimum thickness in the discharge region and the suction region increases with the increasing rotation speed. When the discharge pressure is 20 MPa, with the rotation speed increasing from 700 r/min to 1300 r/min, the mean value of the minimum thickness changes the most, and the variation in the mean value in the discharge region is 2.87 um, while the variation in the mean value in the suction region is 0.28 um.

The speed has obvious influence on the mean value of minimum thickness in the discharge region and the suction region, and the mean increases with the increase of the speed. The discharge pressure has obvious influence on the mean in the discharge area, and the mean value increases with the increase of discharge pressure, but the discharge pressure has a weak effect on the mean in the suction area. The mean of minimum thickness in the discharge area is obviously higher than that in the suction area.

As the speed increases from 700 r/min, the variation of the mean minimum thickness increases with the increase of the variation of the speed in the discharge region and suction region, especially high discharge pressure, as described in Figure 11.

In this picture, the variation of mean minimum thickness in the discharge region is greater. Moreover, the variation in the discharge region is obviously higher than that in the suction region. The maximum of variation in the discharge region is 2.87 um, and the maximum of variation in the suction region is 0.28 um, and the former is 10 times the latter at $p = 20$ MPa.

3.2.3. The Leakage Flow Rate of PRCB per Piston. The leakage of PRCB is caused by the Poiseuille flow and the Couette flow. The leakage from the displacement chamber to the pump case is positive, whereas the leakage is negative. When the piston is in the suction region, the Poiseuille flow is not considered because the pressure difference between the displacement chamber and the pump case is very small.

The effect of rotation speed on the instantaneous flow of PRCB per piston is presented in Figure 12.

As presented in Figure 12, the Poiseuille flow first increases and then decreases with the increase of angular position in the discharge region, and the Poiseuille flow in the discharge pressure increases with the increase of speed, as shown in Figure 12(a). The Couette flow is negative in the discharge pressure region and is positive in the suction region. The value of the Couette flow first increases and then decreases with the increase of angular position, but the value of the Couette flow in the discharge region is significantly larger. When the speed increases from 700 r/min to 1300 r/min, the leakage flow in the discharge region and suction region increases, but the variation of the former is larger, as shown in Figure 12(b). The total flow rate of PRCB is negative in the discharge area and positive in the suction area. The value of the total leakage increases with the

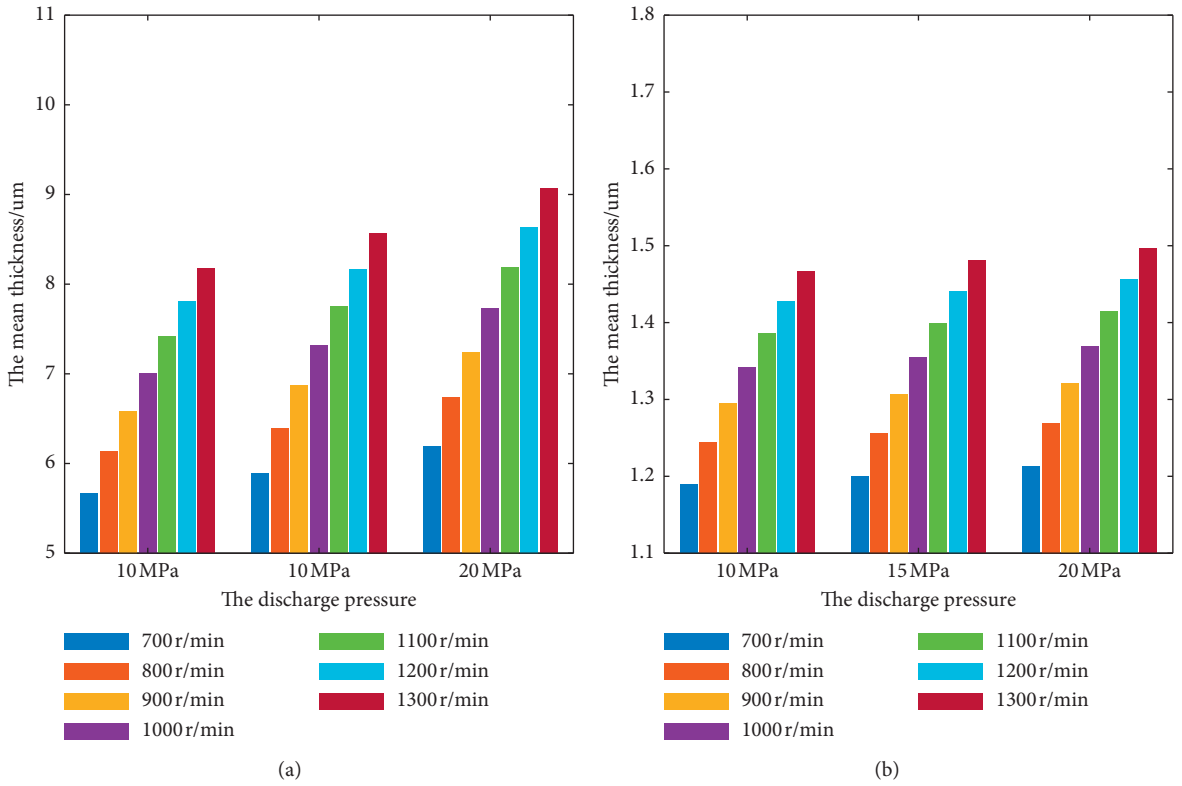


Figure 10: The mean value of the minimum thickness. (a) The discharge region; (b) the suction region.

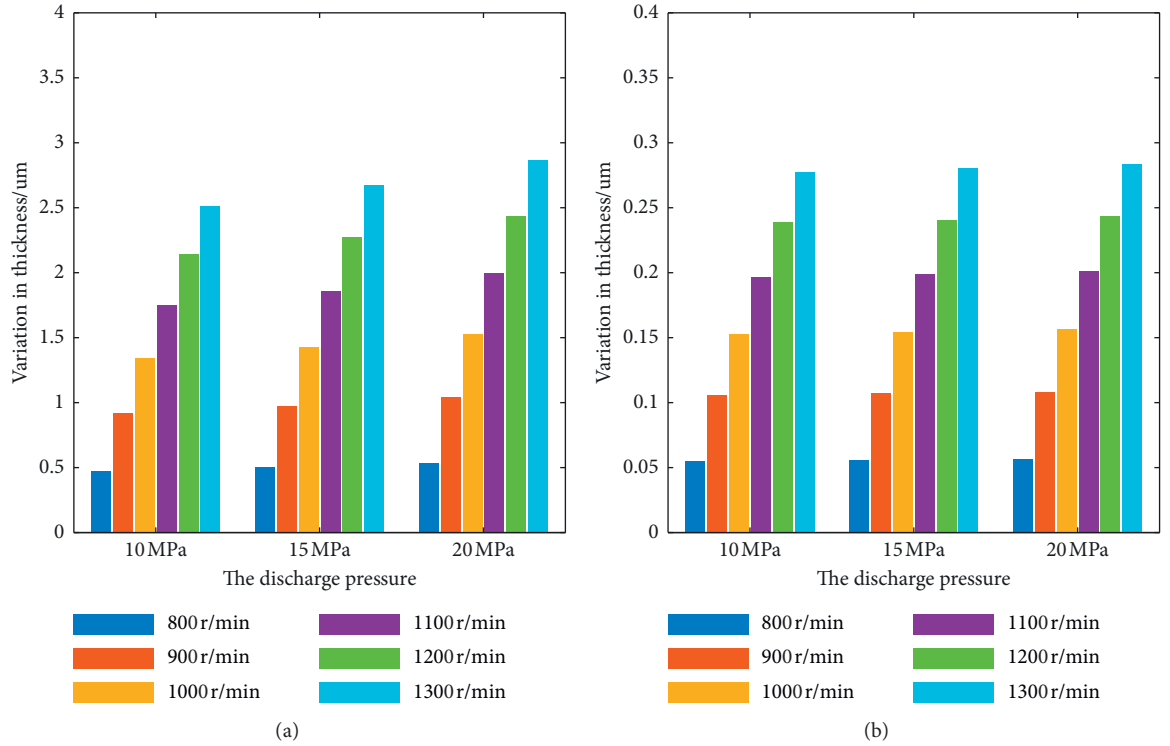


Figure 11: The variation of mean minimum film thickness. (a) The discharge region; (b) the suction region.

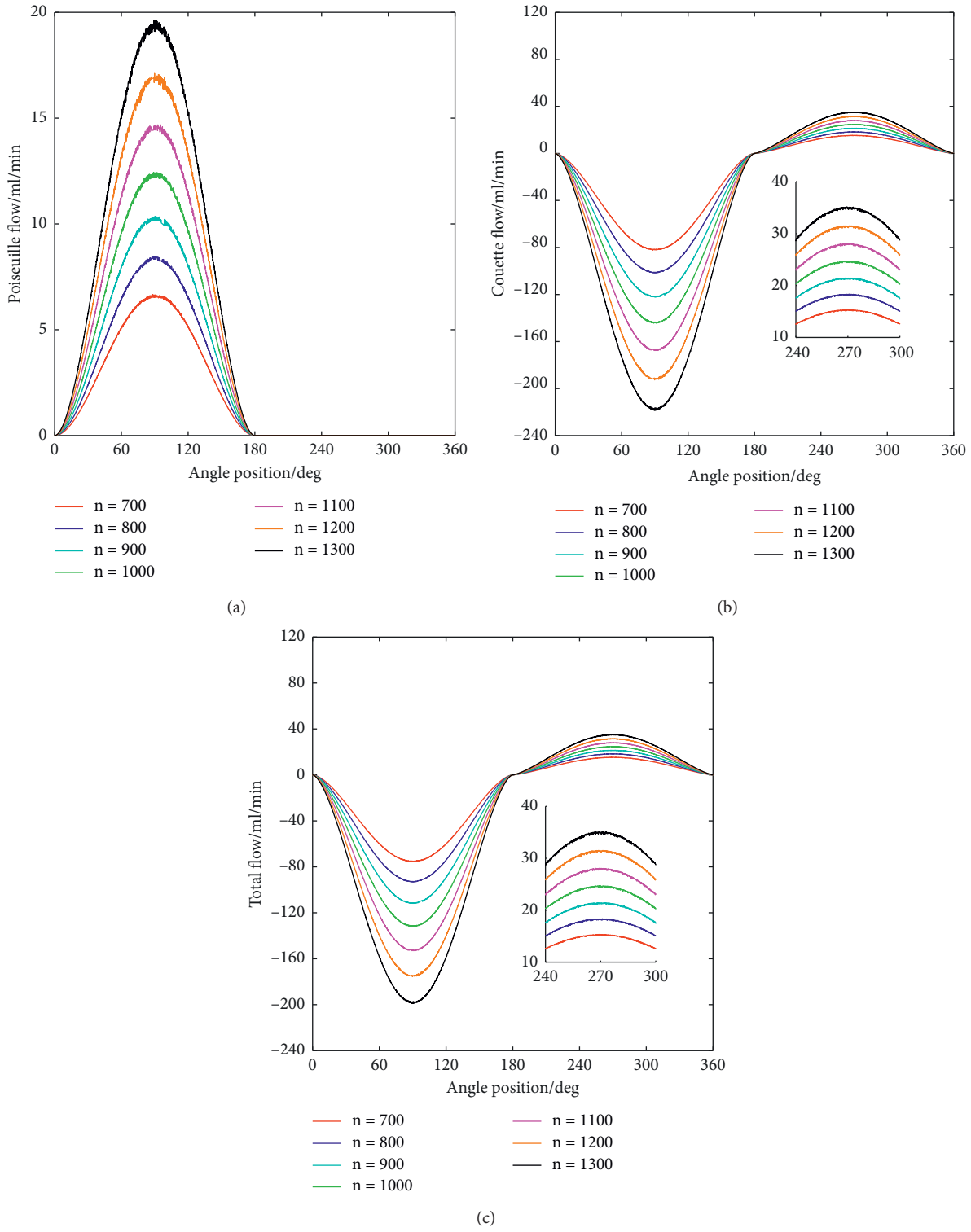


Figure 12: The leakage of PRCB per piston at different rotational speeds. (a) The Poiseuille flow; (b) the Couette flow; (c) the total flow rate.

increase of rotation speed, but the value in the discharge area is obviously larger than that in the suction region as shown in Figure 12(c).

The effect of the discharge pressure on the instantaneous flow rate of PRCB per piston is shown in Figure 13.

As shown in Figure 13, the Poiseuille flow first increases and then decreases with the increase of angular position in the discharge region and increases with the increase of speed as a whole, as shown in Figure 13(a). The Couette flow is negative in the discharge pressure region and is positive in

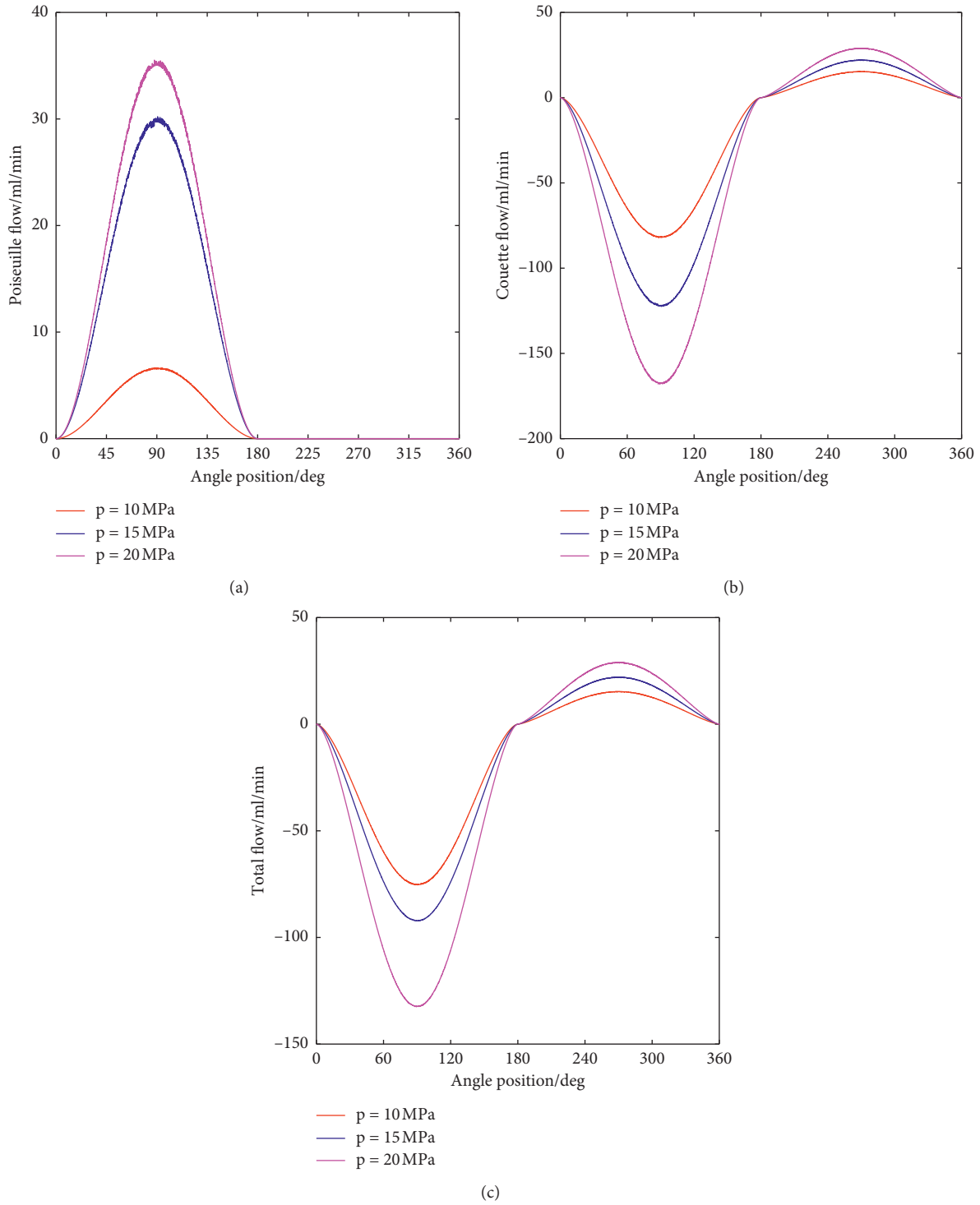


Figure 13: The leakage flow of PRCB per piston at different discharge pressures. (a) The Poiseuille flow; (b) the Couette flow; (c) the total flow rate.

the suction region. The value of the Couette flow first increases and then decreases with the increase of angular position, but the value of the Couette flow in the discharge region is significantly larger. When the speed increases from 700 r/min to 1300 r/min, the leakage flow in the discharge region and suction region increases, but the variation of the former is larger, as shown in Figure 13(b). The total flow rate is negative in the discharge region and is positive in the suction region. The total flow rate increases with the increase of the speed and the discharge pressure, but the value in the discharge region is obviously larger than that in the suction region, as shown in Figure 13(c).

The reason is that the Poiseuille flow is determined by the minimum thickness and the pressure difference between the displacement chamber and the pump case. The pressure difference and the minimum thickness increase with the increase of discharge pressure. Once the discharge pressure is determined, the pressure difference will basically remain unchanged. Therefore, the change trend of the Poiseuille flow consistent with the change trend of the minimum thickness in the discharge region under different the discharge pressure. The Couette flow is determined by the velocity of the piston ring and the minimum thickness between the piston ring and the cylinder, and the positive and negative of the Couette flow are determined by the direction of piston velocity. The leakage flow is caused by the Poiseuille flow and the Couette flow, and the two flows oppose in the discharge pressure region and support each other depending on in the suction region. According to Figures 12 and 13, the Couette flow plays a leading role in leakage flow of PRBC, especially the Couette flow in the discharge region.

The average leakage flow rate of PRCB per piston at different speeds is shown in Figure 14.

As shown in Figure 14, it is found that the average leakage flow is negative, which indicates the flow rate flows from the pump case to the displacement chamber. When the speed is constant, the average leakage flow rate increases with the increase of the discharge pressure. When the discharge pressure is constant, the average leakage flow rate increases with the increase of the speed. If the discharge pressure is 20 MPa, the mean increased from 27.9 ml/min to 75.6 ml/min with the increase of the speed from 700 r/min to 1300 r/min, which increased by 171 percent.

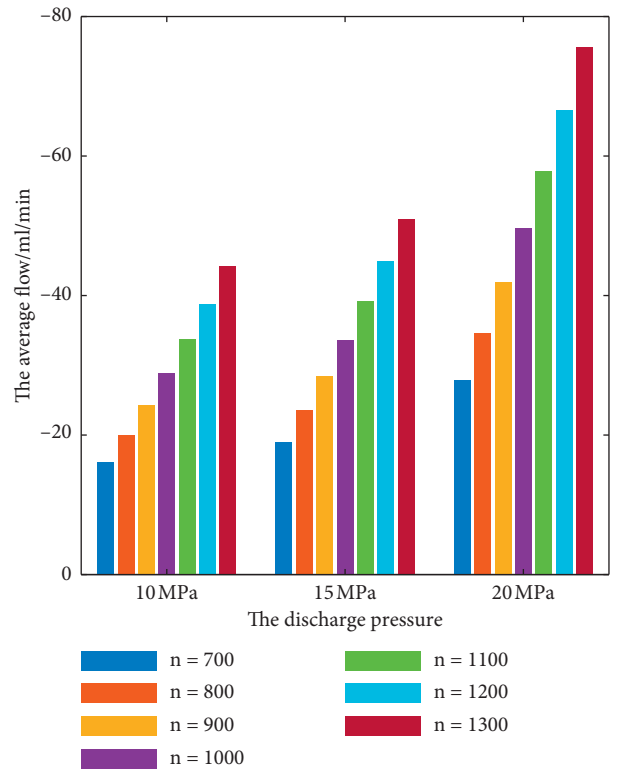


Figure 14: The average leakage flow of PRBC per piston at different speeds.

4. Experimental Results

4.1. Test Rig Description. The leakage flow rate of PRCB is difficult to measure directly. According to the previous works [21, 23], the leakage flow rate of PRCB is an important part of the return oil flow rate. Therefore, the return oil flow rate is measured to investigate the effects of rotation speed on the leakage flow of PRCB. The bent-axis piston pump experiments are implemented on a test rig, as shown in Figure 15.

As shown in Figure 15, the test rig consists of a bent-axis piston pump, a motor, a speed sensor, a pressure senor, a temperature senor, and a relief valve and so on. The bent-axis piston pump is continuously adjusted from 0 to 2000 r/min by the motor. The return oil flow rate is measured by the flow sensor which is installed on the outlet of the pump case in the bent-axis piston pump. The discharge pressure is controlled by adjusting the relief valve. The signals obtained from the main sensors are collected using NI acquisition card and LabVIEW software. The detailed descriptions of the main parameters of the test rig are shown in Table 2.

4.2. Data Processing. The data acquisition diagram of return oil flow rate in the bent-axis piston pump is presented in Figure 16. The sampling frequency was 1 Hz, and the time interval was 35 s under different working conditions. The empirical mode decomposition (EMD) was used to deal with the return oil flow rate under $p = 20$ MPa $n = 1300$ r/min to obtain its intrinsic mode functions (IMF), as shown in Figure 16(a). According to the noise reduction and smoothing of the collected the return oil flow rate data, the return oil flow rate diagram under different working conditions is shown in Figure 16(b). It can be seen from this figure that the overall trend of the return oil flow rate decreases with the increase of rotation speed.

4.3. Results Analysis. The simulated flow rate of PRCB and the measured the return oil flow rate in the bent-axis piston pump is shown in Figure 17.

In Figure 17, it is obvious that the simulated total flow rate of PRCB is negative, and its value increases with the increase of speed and discharge pressure, but the measured the return oil flow rate decreases with the increase of speed and increases with the increase of discharge pressure. The reason is that the measured the return oil flow rate is composed of the leakage of VPCB and the leakage of PRCB. The leakage of VPCB is the main part of the measured the return oil flow rate, and its values increases with the increase of the discharge pressure and is less affected by rotation speed. But the flow rate of PRCB is negative, which leads to the decrease of the measured return oil flow rate.

The variations of the simulated flow of PRCB and the measured return flow rate relative to the flow rate under the same discharge pressure and 700 r/min is shown in Figure 18.

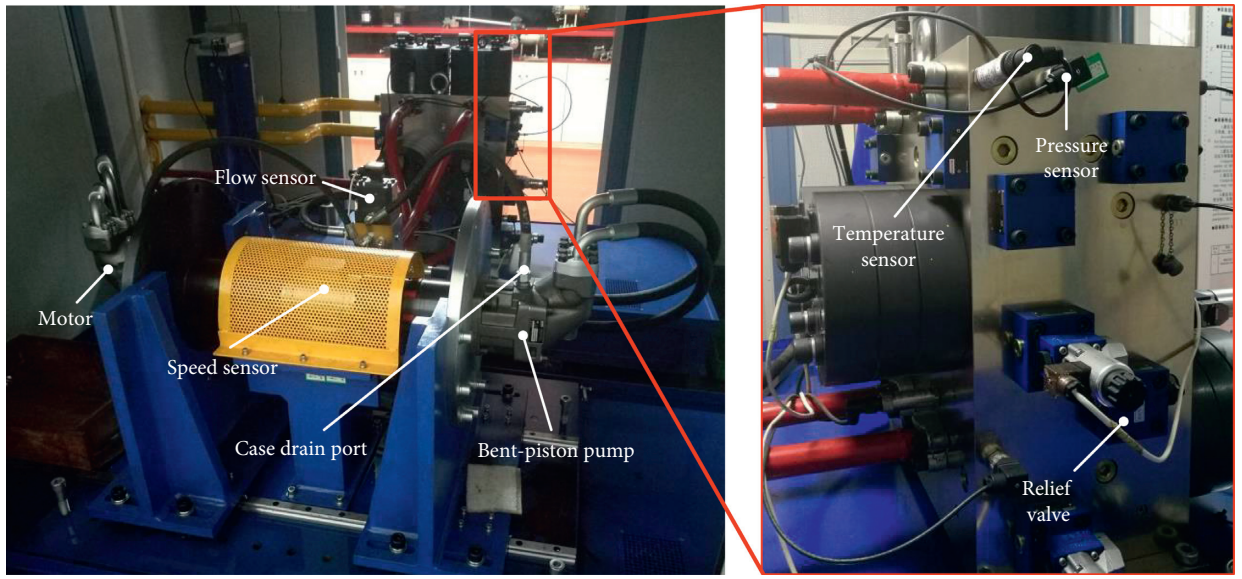


Figure 15: Photograph of the test rig.

Table 2: The main parameters of the test rig.

Devices	Descriptions
The bent-axis piston pump	Displacement: 80.4 ml/r; operating pressure: 0–42 MPa; max self-priming speed: 2300 r/min flow sensor: 0.02–4 l/min, ±0.3%; pressure sensor: 0–60 MPa, ±0.05%
The main sensors	Speed sensor: 0–6000 r/min, ±1 rpm; temperature sensor: −25-100°C, ±0.8%

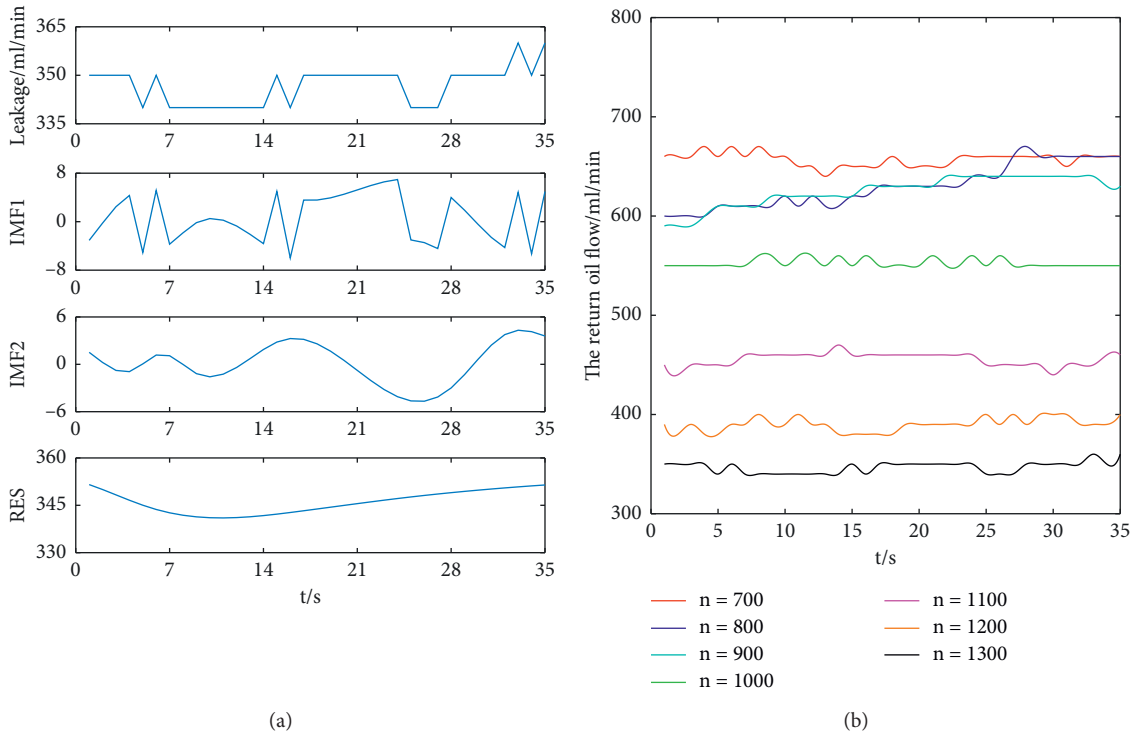


Figure 16: The data acquisition diagram of the return oil flow rate. (a) The return oil flow rate data EMD results under $p = 20$ MPa $n = 1300$ r/min; (b) the return oil flow rate under different working conditions.

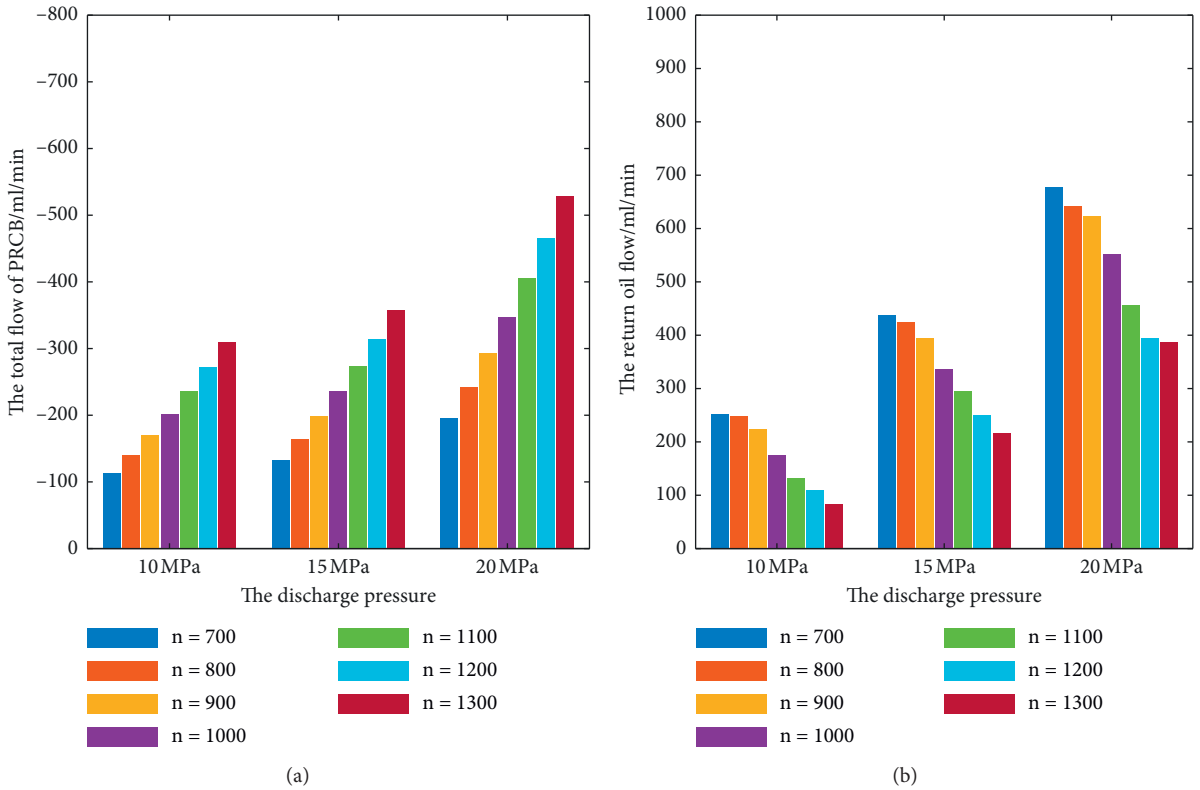


Figure 17: Simulated results and measured results. (a) Simulating the total flow rate of PRCB; (b) measuring the return oil flow rate.

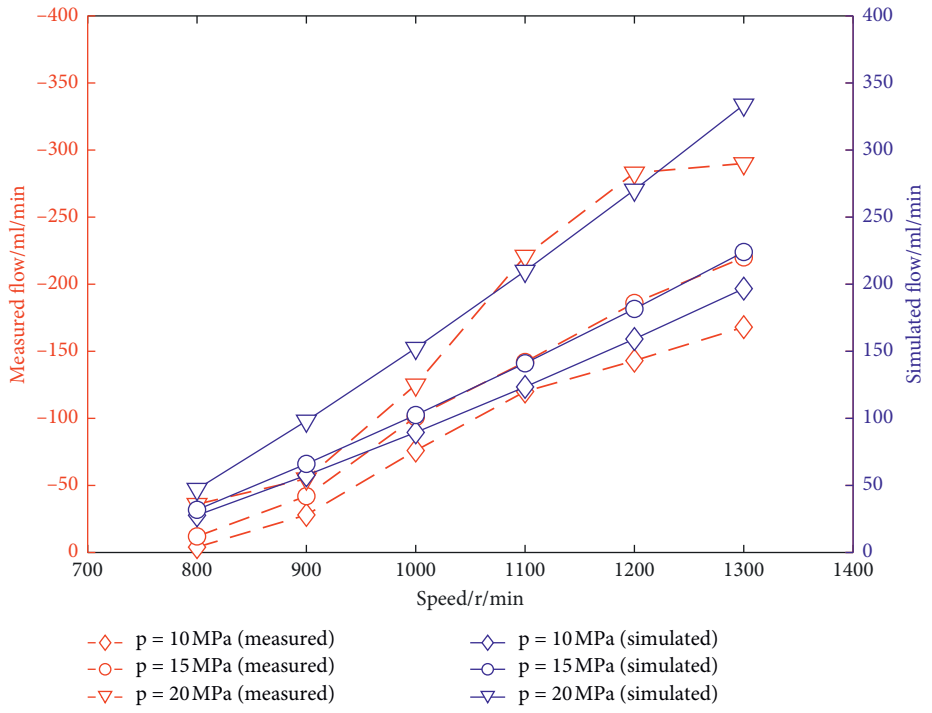


Figure 18: The variations of the simulated flow of PRCB and the measured return flow rate.

In this picture, it is obvious that the variations of the simulated flow of PRCB increase with the increase of rotation speed, indicating that the simulated flow of PRCB increases with the increase of speed. The variation of the measured return flow rate is negative, and it decreases with the increase of rotation speed, indicating that the measured

return flow rate decreases with the increase of rotation speed. Moreover, the trend of the two variations with rotation speed is basically the same at the different discharge pressure. However, there is a certain difference between the two; its maximum is 44 ml/min at $n = 1300$ r/min and $p = 20$ MPa. The main reason is that the measured result is by the leakage flow of PRCB but also by the leakage flow of VPCB.

5. Conclusions

The model is suitable for analyzing the dynamic change of the minimum oil film of PRCB under the action of elastohydrodynamic lubrication. Combining the minimum oil film model with the flow equation, a mathematical model of PRCB is introduced to help produce the effect of the oil film on the leakage of PRCB. The effect of rotation speed on the leakage flow of PRCB is analyzed through simulation and experiment, and the following conclusions are obtained.

(1) The lubrication of PRCB in the discharge of oil segment belongs to the elastohydrodynamic lubrication, and the angle position of the maximum oil film pressure is close to the top dead point 0° with the increase of rotation speed, but its maximum pressure value is basically not affected by the change of the rotation speed.

(2) The minimum thickness of PRCB increases with the increase of rotation speed, and its value in the discharge of oil segment is greater than that in the suction of oil segment, and the former varies more with the rotation speed.

(3) The average flow rate of PRCB is from the pump case to the displacement chamber, which reduces the return oil flow rate in the pump case. The return oil flow rate decreases with the increase of rotation speed, and the effect of rotation speed on the average flow rate of PRCB and the return oil flow rate is also remarkable under the high discharge pressure.

The return oil flow rate in the bent-axis piston pump should be the synthesis of the two friction pairs, which are related to the rotation speed. In this paper, the relationship between the elastohydrodynamic lubrication characteristics of PRCB and the rotation speed is analyzed. The relationship between leakage of the valve plate-cylinder block interface and rotation speed in the bent-axis piston pump will be the focus of further work.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. These data were derived from the test rig built by the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (no. 51675399), Special Research Project of Education Department of Shaanxi Province (no. 18JK0408), and the Science Foundation of Xi'an Aeronautical University (no. 2020KY0223).

References

[1] S. Cun, W. Shaoping, W. Xingjian, and Y. Zhang, "Variable load failure mechanism for high-speed load sensing electro-hydrostatic actuator pump of aircraft," *Chinese Journal of Aeronautics*, vol. 31, pp. 949–964, 2018.

[2] J. Du, S. Wang, L. Han, S. Zhao, and C. Guo, "Prognostic management verification system of aircraft hydraulic power supply system," in *Proceedings of 10th IEEE International Conference on Industrial Informatics (INDIN)*, Beijing, China, July 2012.

[3] Y. Hong and Y. Doh, "Analysis on the friction losses of a bent-Axis type hydraulic piston pump," *Journal of Mechanical Science and Technology*, vol. 18, pp. 1668–1679, 2004.

[4] K. Feng, P. Borghesani, W. A. Smith et al., "Vibration-based updating of wear prediction for spur gears," *Wear*, vol. 426-427, pp. 1410–1415, 2019.

[5] F. Zhao, Z. Tian, X. Liang, and M. Xie, "An integrated prognostics method for failure time prediction of gears subject to the surface wear failure mode," *IEEE Transactions on Reliability*, vol. 67, no. 1, pp. 316–327, 2018.

[6] B. Sun, Y. Li, Z. Wang, Y. Ren, Q. Feng, and D. Yang, "An Improved Inverse Gaussian Process with Random Effects and Measurement Errors for RUL Prediction of Hydraulic Piston Pump," *Measurement*, vol. 173, Article ID 108604, 2020.

[7] F. Lyu, J. Zhang, G. Sun et al., "Research on wear prediction of piston/cylinder pair in axial piston pumps," *Wear*, vol. 456-457, pp. 203338–203410, 2020.

[8] Z. Ma, S. Wang, J. Shi, T. Li, and X. Wang, "Fault diagnosis of an intelligent hydraulic pump based on a nonlinear unknown input observer," *Chinese Journal of Aeronautics*, vol. 31, no. 2, pp. 385–394, 2018.

[9] J. Ma, J. Chen, J. Li, Q. Li, and C. Ren, "Wear analysis of swash plate/slipper pair of axis piston hydraulic pump," *Tribology International*, vol. 90, pp. 467–472, 2015.

[10] N. Kumar, B. K. Sarkar, and S. Maity, "Leakage based condition monitoring and pressure control of the swashplate Axial piston pump," in *Proceedings of ASME 2019 Gas Turbine India Conference*, Chennai, Tamil Nadu, India, 2019.

[11] D. Bensaad, A. Soualhi, and F. Guillet, "A new leaky piston identification method in an axial piston pump based on the extended Kalman filter," *Measurement*, vol. 148, Article ID 106921, 2019.

[12] Z. Ma, S. Wang, H. Liao, and C. Zhang, "Engineering-driven performance degradation analysis of hydraulic piston pump based on the inverse Gaussian process," *Quality and Reliability Engineering International*, vol. 35, pp. 2278–2296, 2019.

[13] R. Guo, Z. Zhao, S. Huo, Z. Jin, J. Zhao, and D. Gao, "Research on state recognition and failure prediction of axial piston pump based on performance degradation data," *Processes*, vol. 8, no. 5, p. 609, 2020.

[14] X. Wang, S. Lin, S. Wang, Z. He, and C. Zhang, "Remaining useful life prediction based on the Wiener process for an aviation axial piston pump," *Chinese Journal of Aeronautics*, vol. 29, no. 3, pp. 779–788, 2016.

[15] M. Pelosi, *An Investigation of the Fluid-Structure Interaction of Piston/cylinder Interface*, Purdue University, West Lafayette, Indiana, 2012.

[16] Y. Song, J. Ma, and S. Zeng, "A numerical study on influence of temperature on lubricant film characteristics of the piston/cylinder interface in axial piston pumps," *Energies*, vol. 11, no. 7, p. 1842, 2018.

[17] D. Qian and R. Liao, "A nonisothermal fluid-structure interaction analysis on the piston/cylinder interface leakage of high-pressure fuel pump," *Journal of Tribology*, vol. 136, 2014.

[18] J. Zhang, B. Liu, R. Lü, Q. Yang, and Q. Dai, "Study on oil film characteristics of piston-cylinder pair of ultra-high pressure axial piston pump," *Processes*, vol. 8, no. 1, p. 68, 2020.

[19] S. Nie, M. Guo, F. Yin et al., "Research on fluid-structure interaction for piston/cylinder tribopair of seawater hydraulic axial piston pump in deep-sea environment," *Ocean Engineering*, vol. 219, Article ID 108222, 2021.

[20] I. Cho, I. Beak, J. Jo, J. Park, S. Oh, and J. Jung, "Lubrication characteristics of dual piston ring in bent-axis type piston pumps," *Journal of Mechanical Science and Technology*, vol. 24, no. 6, pp. 1363–1368, 2010.

[21] R. Kumar, *A Study of the Piston Ring - Cylinder Bore Interface for Bent axis Axial Piston Pump Using an Advanced Computer Model*, ETD Collection for Purdue University, West Lafayette, Indiana, 2009.

[22] B. Manhartsgruber, "The dynamics of leakage in bent Axis units without timing gear," in *Proceedings of BATH/ASME 2016 Symposium on Fluid Power and Motion Control*, Bath, UK, October 2016.

[23] J. Zhang, Y. Li, D. Zhang, B. Xu, F. Lv, and Q. Chao, "A centrifugal force interaction analysis on the piston/cylinder interface leakage of bent-Axis type piston pumps," in *Proceedings of 2016 Ieee/Csaa International Conference on Aircraft Utility Systems (Aus)*, Beijing, China, June 2016.

[24] J.-M. Liu, L.-C. Gu, and B.-L. Geng, "A practical signal processing approach for fault detection of axial piston pumps using instantaneous angular speed," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 234, no. 19, pp. 3935–3947, 2020.

[25] L. Gu, Q. Tian, and Z. Ma, "Extraction of the instantaneous speed fluctuation based on normal time-frequency transform for hydraulic system," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 234, no. 6, pp. 1196–1211, 2020.

[26] L. Gu, R. Xu, and N. Wang, "A novel reduced order dynamic model of axial piston motors with compression flow losses and Coulomb friction losses," *Industrial Lubrication and Tribology*, vol. 72, no. 5, pp. 567–573, 2019.

[27] Y. Liu, L. C. Gu, B. Yang, S. H. Wang, and H. B. Yuan, "A new evaluation method on hydraulic system using the instantaneous speed fluctuation of hydraulic motor," *Proceedings of the Institution of Mechanical Engineers Part C Journal of Mechanical Engineering Science*, vol. 232, no. 15, pp. 2674–2684, Article ID 095440621772257, 2018.

[28] B. J. Hamrock and B. O. Jacobson, "Elastohydrodynamic lubrication of line contacts," *A S L E Transactions*, vol. 27, no. 4, pp. 275–287, 1984.

[29] J. Pei, X. Han, and Y. Tao, "An Improved Stiffness Model for Line Contact Elastohydrodynamic Lubrication and its Application in Gear Pairs," *Industrial Lubrication and Tribology*, vol. 72, no. 5, pp. 703–708, 2020.

[30] Y. Zhang, H. Liu, C. Zhu, C. Song, and Z. Li, "Influence of lubrication starvation and surface waviness on the oil film stiffness of elastohydrodynamic lubrication line contact," *Journal of Vibration and Control*, vol. 24, no. 5, pp. 924–936, 2018.

[31] N. A. H. Tsuha and K. L. Cavalca, "Finite line contact stiffness under elastohydrodynamic lubrication considering linear and nonlinear force models," *Tribology International*, vol. 146, Article ID 106219, 2020.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6615761, 10 pages
https://doi.org/10.1155/2021/6615761

Research Article

A Proposed Bearing Load Identification Method to Uncertain Rotor Systems

Wengui Mao, Nannan Zhang, Dan Feng, and Jianhua Li

Hunan Provincial Key Laboratory of Wind Generator and Its Control, College of Mechanical Engineering, Hunan Institute of Engineering, Xiangtan 411104, China

Correspondence should be addressed to Wengui Mao; mwglikai@163.com

Received 14 December 2020; Revised 10 January 2021; Accepted 22 January 2021; Published 3 February 2021

Academic Editor: Haiyang Pan

Bearings are considered as important mechanical components in rotating machines. Bearing load is used as an indication of monitoring rotor system health, but there are interval and probability uncertain parameters in the process of obtaining bearing load from the rotor system. A bearing load strip enclosed by two bounding distributions is then formed, rather than a single distribution that we usually obtain through the load identification method for a deterministic rotor system. In this paper, a computational inverse approach that combines the interval and perturbation analysis method with regularization is presented to stably identify bearing load strip. Using an interval analysis method, a calculated transient response of the rotor structure only subjecting to the bearing load can be approximated as a linear function of the interval parameters in the rotor system. The perturbation analysis method based on Taylor expansion is used to transform the problem of the bearing load identification involving in probability parameters into two kinds of certain inverse problem, namely, the bearing load identification combining the mean value of uncertain parameters with calculated transient response function and the sensitivity identification of bearing load to each probability parameter. Regularization is used to overcome ill-posedness of bearing load identification arising from the noise-contaminated observed response. A rotor system with two bearings is investigated to demonstrate the effectiveness and accuracy of the presented method.

1. Introduction

Safety and reliability of the high-speed rotating mechanical equipment has attracted much attention with the increase in rotational speed and power [1–3]. In order to reduce the loss caused by mechanical equipment faults and achieve efficient fault diagnosis, it is necessary to perceive and manage the running status and health level of its key systems and components in real time [4]. Bearing is the main basic supporting structure of rotating machinery, and the running state of the rotor system can be assessed by bearing load [5]. At present, the theoretical system mostly uses the sensor to collect the fault characteristic signal of the bearing under specific working conditions and uses it as the basic data of the fault diagnosis model [6, 7]. Any deviations between the referenced bearing load and the observed value can be inferred as an indication of a change or damage in the rotor system. The approach for reliably monitoring the rotor system health is to create a real bearing load from the rotor system in its undamaged state [8]. However, in the actual work site, the acquisition system to determine the real bearing load may be affected by the noise in the environment or the uncertain factors caused by the manufacturing and service environment of the rotating machinery structure [9]. As a result, the bearing load identified by the collected signals do not fully reflect the real operation of the rotor system. Because of the existence of uncertain factors, the bearing load distribution will form a strip enclosed by two bounding distributions, rather than a single distribution that we usually obtain through a deterministic rotor system. If the observed bearing load is within the boundary of the referenced bearing load, the rotor system runs without fault.

In general, the problem to determine the bearing load strip from the collected output signals considering the

uncertain factors is called an uncertain inverse problem. This type of inverse problem has recently attracted more and more attention. As the most popular uncertainty modeling strategy, probability [10, 11], interval [12, 13], and combined method [14, 15] have been proposed to applied in engineering for uncertain parameter identification and load identifications. However, the probability method is a quantitative description, which describes the uncertainty parameters accurately but requires a lot of information coming from many expensive experimental tests to construct the probability density function; the interval method is a qualitative description, which handles the uncertainty with limited information, but it will lose the actual value of the project due to the expansion of the interval estimation. Thus, while proposing a new uncertainty inverse analysis framework for gaining the bearing load strip to evaluate uncertain rotor system fault, a more efficient computing method needs to be developed together.

In this paper, a hybrid method based on interval theory [16], perturbation theory [17], and regularization method [18] is proposed to construct the inverse analysis framework to deal with the uncertainty with imprecise information; utilizing Taylor expansion, the problem of bearing load identification of the uncertain rotor system is transformed into two kinds of deterministic bearing load identification problems. One kind is the middle bearing load identification taking the uncertain parameters as the mean value, and other is the reverse calculation of the sensitivity of the bearing load with respect to the uncertain parameters; the regularization method is used to treat the unstable solution coming from the noise in the measurement response.

2. Bearing Load Identification Problem

The transient response [19] coming from vibration analysis of the bearing-rotor system is shown in the following:

$$\begin{bmatrix} \mathbf{M}_{II} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}_{BB} \end{bmatrix} \{\ddot{q}_I\ \ddot{q}_B\} + \begin{bmatrix} \mathbf{C}_{II} & \mathbf{O} \\ \mathbf{O} & \mathbf{C}_{BB} + \mathbf{C}_B \end{bmatrix} \begin{Bmatrix} \dot{q}_I \\ \dot{q}_B \end{Bmatrix} + \begin{bmatrix} \mathbf{K}_{II} & \mathbf{K}_{IB} \\ \mathbf{K}_{IB} & \mathbf{K}_{BB} + \mathbf{K}_B \end{bmatrix} \begin{Bmatrix} \mathbf{q}_I \\ \mathbf{q}_B \end{Bmatrix} = \omega^2 \begin{Bmatrix} \mathbf{F}_{eI}^T \\ \mathbf{F}_{eB}^T \end{Bmatrix}, \tag{1}$$

where $\mathbf{O}$ is the zero matrix; $\mathbf{M}$, $\mathbf{K}$, and $\mathbf{C}$, respectively, represent the mass, stiffness, and damping matrices of the separate rotor structure; the index B denotes the node of the rotor with the bearing support and the index I denotes the node of the rotor without the bearing support; $\mathbf{K}_B$ and $\mathbf{C}_B$ represent the bearing stiffness and damping matrices, respectively; the transient displacement $\mathbf{q}$, velocity $\dot{q}$, and acceleration $\ddot{q}$ represent the structural response of the bearing-rotor system under unbalanced force $\mathbf{F}_e$ with the rotational speed ω.

Denoting $\mathbf{F}^T = \mathbf{C}_B\dot{q}_B + \mathbf{K}_B\mathbf{q}_B$ and moving $\mathbf{F}^T$ from the left side of equation (1) to the right side, the equivalent kinetic equation of the rotor structure subjecting to the unbalanced force $\mathbf{F}_e$ and the bearing load $\mathbf{F}^T$ is shown in the following:

$$\begin{aligned} &\begin{bmatrix} \mathbf{M}_{II} & \mathbf{O} \\ \mathbf{O} & \mathbf{M}_{BB} \end{bmatrix} \{\ddot{q}_I\ \ddot{q}_B\} + \begin{bmatrix} \mathbf{C}_{II} & \mathbf{O} \\ \mathbf{O} & \mathbf{C}_{BB} \end{bmatrix} \begin{Bmatrix} \dot{q}_I \\ \dot{q}_B \end{Bmatrix} + \begin{bmatrix} \mathbf{K}_{II} & \mathbf{K}_{IB} \\ \mathbf{K}_{IB} & \mathbf{K}_{BB} \end{bmatrix} \begin{Bmatrix} \mathbf{q}_I \\ \mathbf{q}_B \end{Bmatrix} \\ &= \omega^2 \begin{Bmatrix} \mathbf{F}_{eI}^T \\ \mathbf{F}_{eB}^T \end{Bmatrix} - \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & C_B \end{bmatrix} - \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \mathbf{O} & K_B \end{bmatrix} \begin{Bmatrix} q_I \\ q_B \end{Bmatrix} = \omega^2 \begin{Bmatrix} \mathbf{F}_{eI}^T \\ \mathbf{F}_{eB}^T \end{Bmatrix} \begin{Bmatrix} \mathbf{O} \\ \mathbf{F}^T \end{Bmatrix}, \end{aligned} \tag{2}$$

or simplified to

$$[\mathbf{M}]\{\ddot{q}\} + [\mathbf{C}_2]\{\dot{q}\} + [\mathbf{K}_2]\{\mathbf{q}\} = \omega^2\{\mathbf{F}_e\} - \{\mathbf{F}^T\}. \tag{3}$$

The measurable structural response $\mathbf{q}$ can be thought to be the superposition of responses of the rotor separately subjecting to F_e and $-\mathbf{F}^T$. The damping matrix $\mathbf{C}_2$ and stiffness matrix $\mathbf{K}_2$ of the separate rotor in equation (3) do not combine the effects of the bearing.

As the rotor structure is linear time invariant, the following relationship is obtained according to the superposition principle

$$[\mathbf{M}]\{\ddot{q}_T\} + [\mathbf{C}_2]\{\dot{q}_T\} + [\mathbf{K}_2]\{\mathbf{q}_T\} = \{\mathbf{F}^T\}, \tag{4}$$

where $\mathbf{q}_T = \mathbf{q}^e - \mathbf{q}$, $\mathbf{q}_T$ represent the transient response of the rotor structure only subjecting to the bearing load $\mathbf{F}^T$. The unbalance response $\mathbf{q}^e$ only subjecting to the unbalance load F_e can be numerically and accurately calculated by combining the information of the unbalance load F_e together with the matrices $\mathbf{M}$, $\mathbf{C}_2$, and $\mathbf{K}_2$ into the forward solver [20]. The calculated unbalance response $\mathbf{q}^e$ together with the measurable structural response q is combined into Green's function method and regularization operation [21] to reconstruct the bearing load $\mathbf{F}^T$.

3. Bearing Load Identification Method for Uncertain Rotor Systems

When the parameters in the rotor system, such as material properties and geometric structure, cannot be determined completely, the Green kernel function in the load identification used to characterize the dynamic characteristics of the

rotor system will be uncertain. When this convolution integral in the time domain is discretized, the whole concerned time period is separated into equally spaced intervals, and the response $\mathbf{q}_T$ only subjecting to the bearing load $\mathbf{F}^T$ can be expressed by a matrix form:

$$\mathbf{q}_T = \mathbf{H}(\lambda, \eta) \cdot \mathbf{F}^T(\lambda, \eta), \tag{5}$$

where $\mathbf{G}$ is the Green function matrix coming from the bearing load $\mathbf{F}^T$ to the response $\mathbf{q}_T$.

In general, the Green function matrix $\mathbf{H}$ of equation (5) is ill-conditioned, and the response $\mathbf{q}_T$ calculated by the measurable structural response $\mathbf{q}$ according to $\mathbf{q}_T = \mathbf{q}^e - \mathbf{q}$ inevitably carries noise. If the inverse of the matrix $\mathbf{H}$ exists, the bearing load $\mathbf{F}^T$ can be obtained by using the following equation based on regularization:

$$\begin{aligned}\mathbf{F}^T &= \mathbf{H}^{\alpha} \cdot \mathbf{q}_{T\delta} = \mathbf{V}\mathrm{Diag}\left(f(\alpha, \sigma_i)\sigma_i^{-1}\right)\mathbf{U}\mathbf{q}_{T\delta} \\ &= \sum_{i=1}^{m} f(\alpha, \sigma_i)\sigma_i^{-1}\left(\mathbf{U}_i^T \mathbf{q}_{T\delta}\right)\mathbf{V}_i,\end{aligned} \tag{6}$$

where $\mathbf{q}_{T\delta}$ represents the response with noise, $\mathbf{U} = [u_1, u_2, \ldots, u_m]$ is the left singular vector of $\mathbf{H}$ and $\mathbf{V} = [v_1, v_2, \ldots, v_m]$ is the right singular vector of $\mathbf{H}$, which are two normalized orthogonal matrices, and $f(\alpha, \sigma_i)$ denotes a filter function to attenuate the amplification effect of small singular value on noise.

The Green function matrix $\mathbf{H}(\lambda, \eta)$ is expressed as a set of uncertain parameters containing interval and probability parameters, so the bearing load is no longer a solution, but a solution set. In rotor system health monitoring, the upper and lower boundaries of the solution set are often concerned, and it is not necessary to solve all the possible values. According to the interval mathematics theory, when the uncertain parameter λ is in the interval form, the corresponding response $\mathbf{q}_T$ is also in the interval form. In order to obtain the maximum and minimum values of the bearing load, all the possible values of the interval uncertain parameter λ should be selected in turn, the bearing load should be identified based on the response $\mathbf{q}_T$ distribution strip, and the maximum and minimum values of the bearing load should be searched in the identification results. The recognition process with a traditional Monte Carlo simulation (MCS) [22] will involve complex multilayer nesting solution, and the forward problem calculation model needs to be called repeatedly, which will inevitably lead to the inefficiency of the solution. In this paper, the bearing load identification problem involving interval and probability uncertain parameters in equation (5) is transformed into a series of deterministic problems utilizing the method of interval analysis and matrix perturbation.

3.1. Boundary of Transient Response. According to the monotonicity analysis theory [23], the maximum and minimum values of bearing load must correspond to the boundary of interval uncertain parameters. The lower boundary λ^L and upper boundary λ^R of n-dimensional interval uncertain parameters can be described by the midpoint λ^c and radius λ^w of the interval. When the uncertainty level is small, the first-order Taylor expansion is carried out at the midpoint λc of the interval. The minimum and maximum values of the response $\mathbf{q}_T$ can be obtained directly and explicitly.

$$\begin{cases} q_{T\min}(t) = \min\limits_{\boldsymbol{\lambda}\in\Gamma} q_T(t, \boldsymbol{\lambda}) = q_T(t, \boldsymbol{\lambda}^c) - \sum\limits_{j=1}^{n}\left|\dfrac{\partial q_T(t, \boldsymbol{\lambda}^c)}{\partial \boldsymbol{\lambda}_j}\right|, \\ q_{T\max}(t) = \max\limits_{\boldsymbol{\lambda}\in\Gamma} q_T(t, \boldsymbol{\lambda}) = q_T(t, \boldsymbol{\lambda}^c) - \sum\limits_{j=1}^{n}\left|\dfrac{\partial q_T(t, \boldsymbol{\lambda}^c)}{\partial \boldsymbol{\lambda}_j}\right|, \\ \boldsymbol{\lambda} \in \boldsymbol{\lambda}^I = \left[\boldsymbol{\lambda}^L, \boldsymbol{\lambda}^R\right] = \left[\boldsymbol{\lambda}^c - \boldsymbol{\lambda}^w, \boldsymbol{\lambda}^c + \boldsymbol{\lambda}^w\right], \\ \lambda_j \in \lambda_j^I = \left[\lambda_j^L, \lambda_j^R\right] = \left[\lambda_j^c - \lambda_j^w, \lambda_j^c + \lambda_j^w\right], \\ \lambda = \lambda^c + \delta\lambda, \quad \delta\lambda \in [-1, 1]\lambda^w, \\ \lambda_j \in [-1, 1]\lambda_j^w, \quad j = 1, 2, \ldots, n. \end{cases} \tag{7}$$

3.2. Boundary of Bearing Load. The k-dimensional probabilistic uncertainty parameter η can be expressed as the mean value η_d and disturbance part $\Delta\eta_r$. According to the perturbation theory [24], the Green kernel function matrix of the corresponding rotor system and the bearing load to be identified contain disturbance parts.

$$\begin{aligned} \eta &= \eta_d + \Delta\eta_r, \\ \eta_i &= \eta_{di} + \Delta\eta_{ri}, \quad i = 1, 2, \ldots, q, \\ \mathbf{q}_T &= \left(\mathbf{H}_d + \Delta\mathbf{H}_r\right) \cdot \left(\mathbf{F}_d^T + \Delta\mathbf{F}_r^T\right), \end{aligned} \tag{8}$$

where the subscript d and r represent the mean value and disturbance part of the probabilistic uncertain parameters, respectively.

$$\begin{aligned} \mathbf{q}_T &= \mathbf{H}_d\mathbf{F}_d^T, \\ -\Delta\mathbf{H}_r\mathbf{F}_d^T &= \mathbf{H}_d\Delta\mathbf{F}_r^T. \end{aligned} \tag{9}$$

The Green kernel function matrix $\mathbf{H}_d$ in equation (8) can be obtained when the probabilistic uncertain parameters are taken as the mean value. Based on the response $\mathbf{q}_T$, the mean value $\mathbf{F}_d^T$ of the bearing load can be inversely calculated by the deterministic bearing load identification method. The disturbance part $\Delta\mathbf{F}_r$ of the bearing load can be identified through the reverse calculation of the sensitivity of the bearing load with respect to the probabilistic uncertain parameters. Among the solving the mean value $\mathbf{F}_d^T$ and the disturbance part $\Delta\mathbf{F}_r$, the Green function matrix $\mathbf{H}_d$ in equation (9) is the same, and only one matrix singular value decomposition operation is needed in regularization. Through the above matrix perturbation analysis method, the bearing load identification problem with k-dimensional probabilistic uncertain parameters can be transformed into the following $k+1$ deterministic problem. The bearing load with interval and probability uncertainty can be described

using the Taylor expansion of first order regarding near their means (λ^c, η_d).

$$\begin{cases} \mathbf{F}_L^T = \min\limits_{\boldsymbol{\lambda}\in\Gamma,\boldsymbol{\eta}\in\Omega} F_d^T(t,\boldsymbol{\lambda},\boldsymbol{\eta})|\mathbf{q}_{Tm} - \sum\limits_{i=1}^{n}\left|\dfrac{\partial \mathbf{F}^T(t,\boldsymbol{\eta}_d)}{\partial \boldsymbol{\eta}_i}\right|\Delta\boldsymbol{\eta}_{ri}|\mathbf{q}_{T\min}q_{T\max}, \\ \mathbf{F}_R^T = \max\limits_{\boldsymbol{\lambda}\in\Gamma,\boldsymbol{\eta}\in\Omega} F_d^T(t,\boldsymbol{\lambda},\boldsymbol{\eta})|\mathbf{q}_{Tm} - \sum\limits_{i=1}^{n}\left|\dfrac{\partial F^T(t,\boldsymbol{\eta}_d)}{\partial \boldsymbol{\eta}_i}\right|\Delta\boldsymbol{\eta}_{ri}|\mathbf{q}_{T\min}q_{T\max}, \end{cases} \tag{10}$$

where the transient response $\mathbf{q}_{Tm} = \mathbf{q}^e - \mathbf{q}^m$ for identifying the mean value $\mathbf{F}_d^T$ is obtained from the measurable structural response $\mathbf{q}^m$ to means (λ^c, η_d) together with the calculated unbalance response $\mathbf{q}^e$, and he disturbance part $\Delta\mathbf{F}_r$ centering on the $\mathbf{q}_{T\min}$ and $q_{T\max}$ from equation (7) using the deterministic bearing load identification method is computed. The first-order partial derivative $(\partial F^T(t,\eta_d)/\partial\eta_i)$ to the probabilistic parameters can be dealt by the difference method [25], and the difference scheme is used to transform the partial differential equation into an algebraic equation to solve, for example, $(\partial\mathbf{F}^T(t,\eta_d)/\partial\eta_i) = (\Delta\mathbf{F}^T(t,\eta_d)/\Delta\eta_i)$, and the increment $\Delta\mathbf{F}^T(t,\eta_d)$ of the bearing load is obtained by small disturbance $\Delta\eta_{ri}$ at the midpoint. With the mean value $\mathbf{F}_d^T$ and the disturbance part $\Delta\mathbf{F}_r$, the minimum $\mathbf{F}_L^T$ and maximum $\mathbf{F}_R^T$ of the bearing load can easily calculated based on equation (10).

4. Bearing Load Identification Process of Uncertain Rotor Systems

Based on the above discussion, the solution procedure towards a bearing load identification of the rotor system involving interval and probability uncertain parameters can be described as follows (see Figure 1):

(1) Construct a forward solver to calculate the unbalance response $\mathbf{q}^e$ or structural response q utilizing the mass, stiffness, and damping matrices and the corresponding load

(2) Construct an inverse problem solver for the deterministic load identification algorithm, combining the calculated or measured structural response $\mathbf{q}^m$ and the unbalance response $\mathbf{q}^e$ with the known mass, stiffness, and damping matrices into the inverse problem solver to calculate the bearing load

(3) Assume the uncertain parameters (λ, η) from engineering experience, combining the mean of uncertain parameters (λ^c, η_d) together with the measured structural response $\mathbf{q}^m$ into the inverse problem solver to calculate the mean value $\mathbf{F}_d^T$

(4) Calculate $\mathbf{q}_{T\min}$ and $\mathbf{q}_{T\max}$ from the forward solver using the obtained $\mathbf{F}_d^T$ and the known interval parameters λ based on the interval analysis method

(5) Combine the probabilistic parameters η together with the obtained $q_{T\min}$ and $q_{T\max}$ into the inverse problem solver to calculate the disturbance part $\Delta\mathbf{F}_r$ based on the difference method

(6) Utilize the mean value $\mathbf{F}_d^T$ and the disturbance part $\Delta\mathbf{F}_r$ to calculate $\mathbf{F}_L^T$ and $\mathbf{F}_R^T$, as shown in equation (10)

5. Examples

The numerical model in reference [26, 27] is provided to verify the feasibility of the proposed uncertain bearing load identification algorithm. Figure 2 depicts a rotor system with three discs supported by two fault predicted journal bearings, the geometrical parameters of the rotor structure are partly given in Figure 2(a), the transfer matrix model of the rotor structure is established as shown in Figure 2(b), and the rotor system is divided into 33 elements consisting of 34 nodes with three discs at element 9, 17, and 28. The bearing acts on the nodes 7 and 24. The angular speed is 4000 rpm. The elastic modulus E of the rotor material is 210 GPa, the shear elastic modulus G is 80 GPa, and the density is 7650 kg/m^3. The disc weight is 14 kg, and the unbalance mass to construct unbalanced force F_e of 14 kg at 0° is placed at 1 mm radii from the center of the middle disc. The mass, stiffness, and damping matrices of the rotor system are obtained by the transfer matrix method.

5.1. Identification of Bearing Load under Determined Structure. The structural response q can be easily obtained by using a forward solver (such as TMM) reported by Mao et al. [27] with the information of the mass, stiffness, damping matrices, and the corresponding unbalance load and bearing load. For testing the inverse problem solver for the deterministic load identification algorithm, the structural response at the nodes 3, 8, 20, and 25 used to identify the bearing load is gained by noise-contaminated response (a 3% Gaussian noise is directly added to the computer-generated response coming from the assumed value of bearing stiffness and damping coefficient in Table 1 into the TMM forward solver), as shown in Figure 3. With the help of regularization, the bearing loads are stably obtained and shown in Figure 4. It is known that the load time histories of

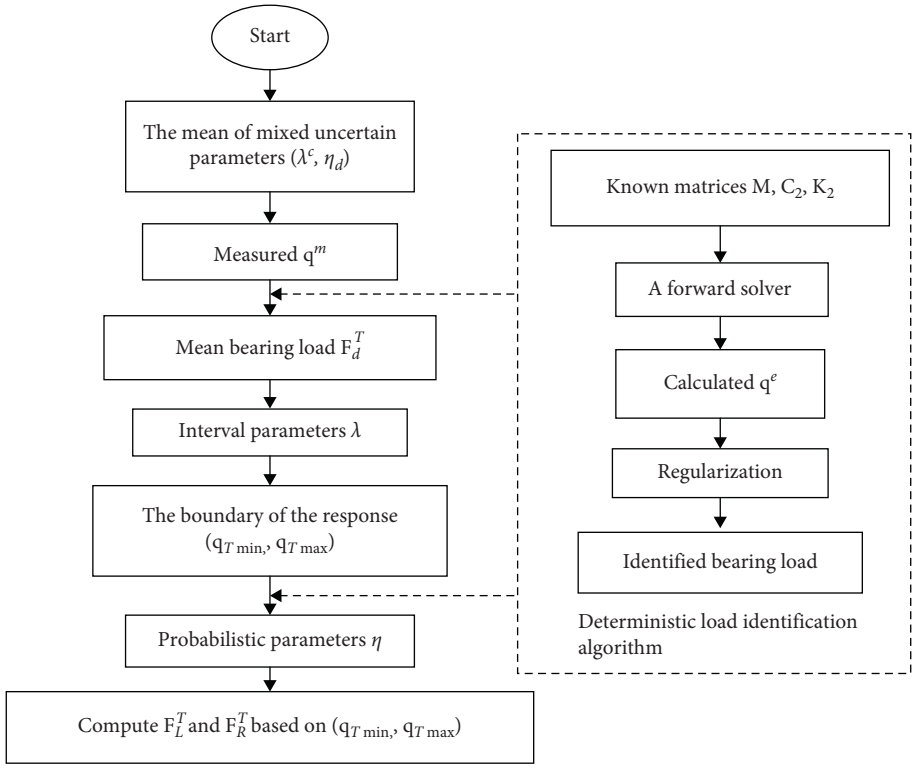


Figure 1: Solution procedure towards bearing load identification with uncertainty.

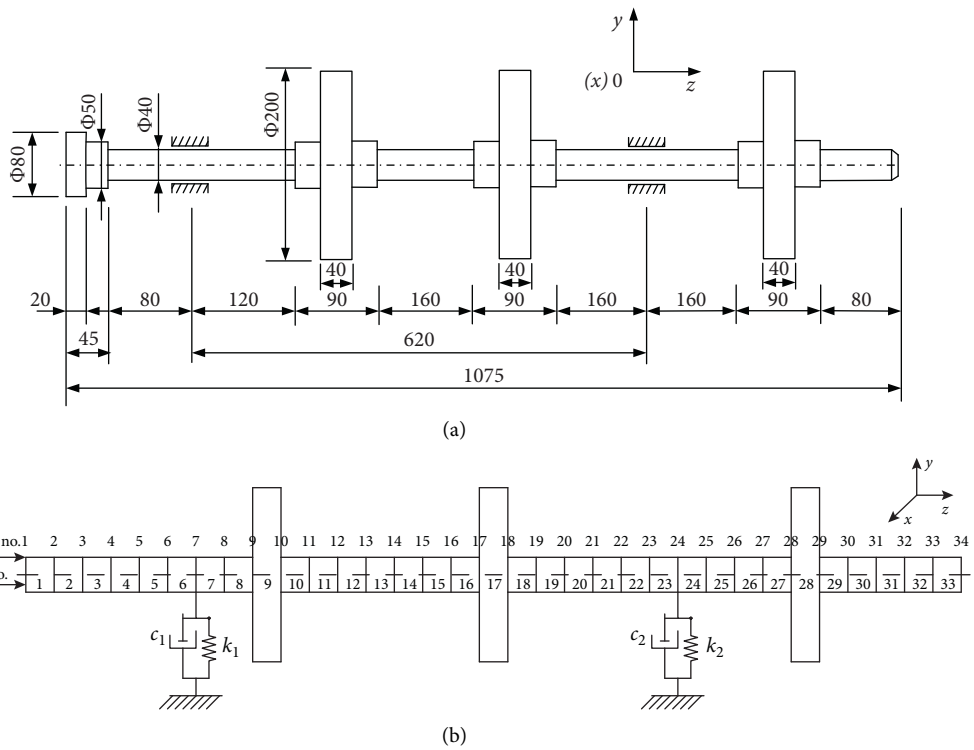


Figure 2: The model of a rotor system: (a) a bearing-rotor system with three discs; (b) the TMM model.

Table 1: Bearing dynamic parameters.

Bearing	Stiffness coefficients ($MN \cdot m^{-1}$)				Damping coefficients ($KN \cdot s \cdot m^{-1}$)			
	K_{xx}	K_{xy}	K_{yx}	K_{yy}	C_{xx}	C_{xy}	C_{yx}	C_{yy}
Left	46.36	83.4	−64.33	41.27	70.28	71.63	71.63	88.57
Right	13.38	29.16	−21.78	8.36	69.67	19.98	19.98	79.93

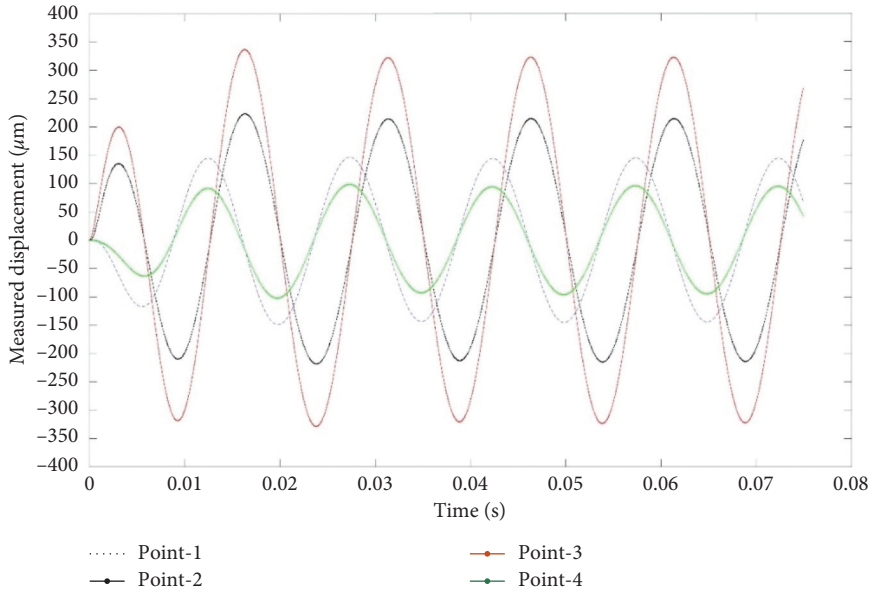


Figure 3: The structural response at four observed point.

the three cases are in good agreement, and in the acceptable range, it shows that the accuracy of the deterministic bearing load identification algorithm is feasible.

5.2. Bearing Load Identification under 3% Uncertainty. For testing the proposed uncertain bearing load identification algorithms, the unbalance parameters $\{m, e, \varphi\}$ are interval uncertain and the rotor material parameters $\{E, G\}$ are probability uncertain. Assumed the probabilistic uncertainty parameters are taken as the mean (that is, $\{E = 210\,\text{GPa}, G = 80\,\text{GPa}\}$), adjusting the interval uncertainty parameters under 3% uncertainty (that is, 3% off from the mean $\{m = 14\,\text{kg}, e = 1\,\text{mm}, \phi = 0^{\circ}\}$), and combining them together with the obtained mean bearing load as shown in Figure 4 into the forward solver (TMM) to calculate the response $\mathbf{q}_T$, the sensitivity curves of the response $\mathbf{q}_T$ with respect to the interval uncertainty parameters $\{m, e, \phi\}$ are obtained, as shown in Figure 5(a). It can be seen from the diagram that the response $\mathbf{q}_T$ is more sensitive to the unbalance parameters $\{m, e, \phi\}$ and the unbalance parameters have great influence on the identification of bearing load. Combining with the first-order derivatives to the unbalance parameters together with the response $\mathbf{q}_T(t, \lambda)$ coming from the obtained mean bearing load $\mathbf{F}_d^T$ into equation (7), the bounds of the response, $\mathbf{q}_{T\min}$ and $q_{T\max}$, can be obtained, shown in Figure 5(b).

Utilizing the obtained response, $\mathbf{q}_{T\min}$ and $q_{T\max}$, and the probability uncertainty parameters under 3% uncertainty into the inverse problem solver for the deterministic load identification algorithm, the bearing load with respect to each probability uncertain parameters is calculated and the results are plotted in Figure 6(a). It shows that the sensitivity curves vary greatly in the whole-time history, and the probability parameters will have a relatively huge single impact on the accuracy of the identified boundaries of the bearing load. Combining with the disturbance part $\Delta\mathbf{F}_r$ coming from the probabilistic parameters η together with the obtained $\mathbf{q}_{T\min}$ and $q_{T\max}$ together with the obtained mean bearing load $\mathbf{F}_d^T$ into equation (10), the bounds of bearing load, $\mathbf{F}_L^T$ and $\mathbf{F}_R^T$, can be obtained, shown in Figure 6(b). The identified bearing load boundary can effectively describe the time history of the bearing load change, the upper and lower boundaries of the bearing load can better contain the middle load, and the middle load is within the upper and lower boundaries. In addition, at the peak of the bearing load, the boundary of the identified bearing load is wide, so the interval and probability uncertain parameters have a great influence on the identification results of the bearing load currently.

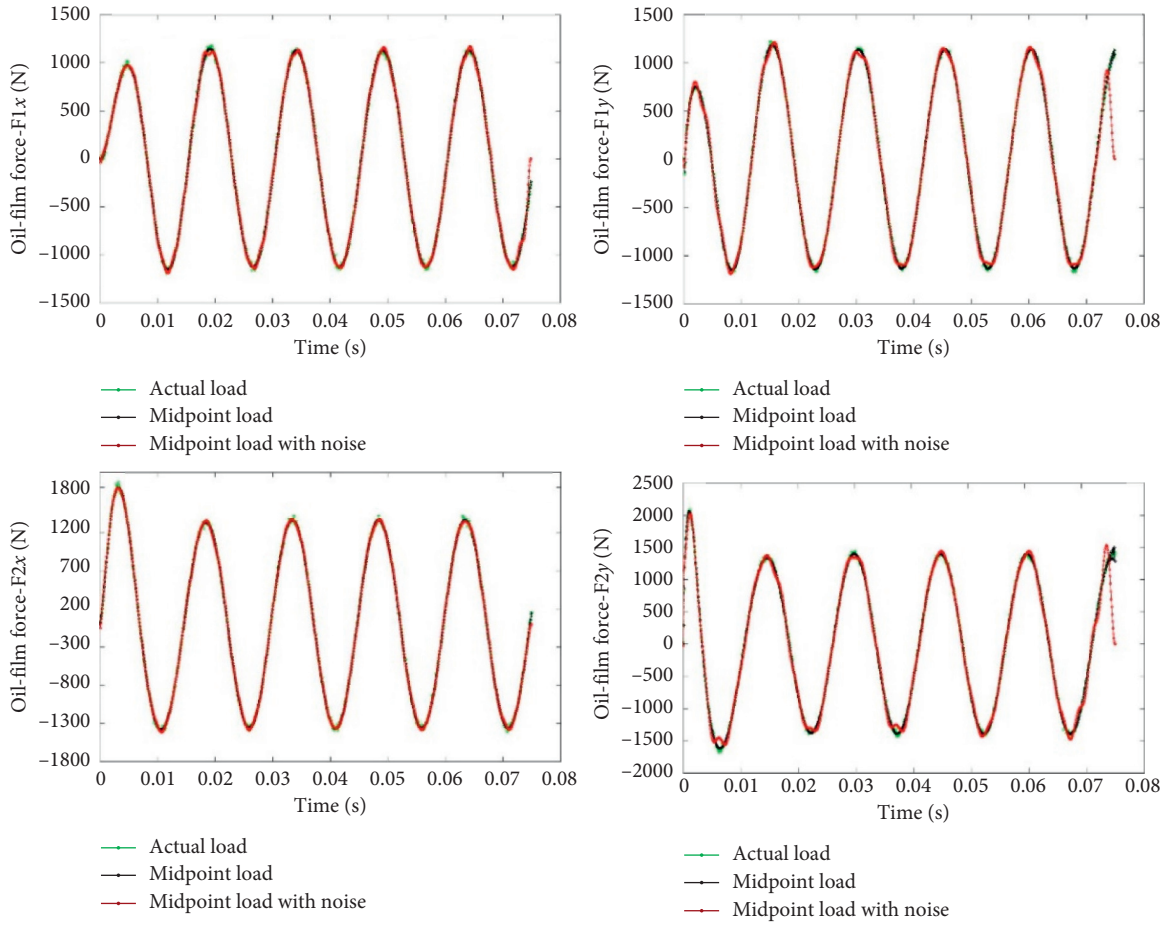
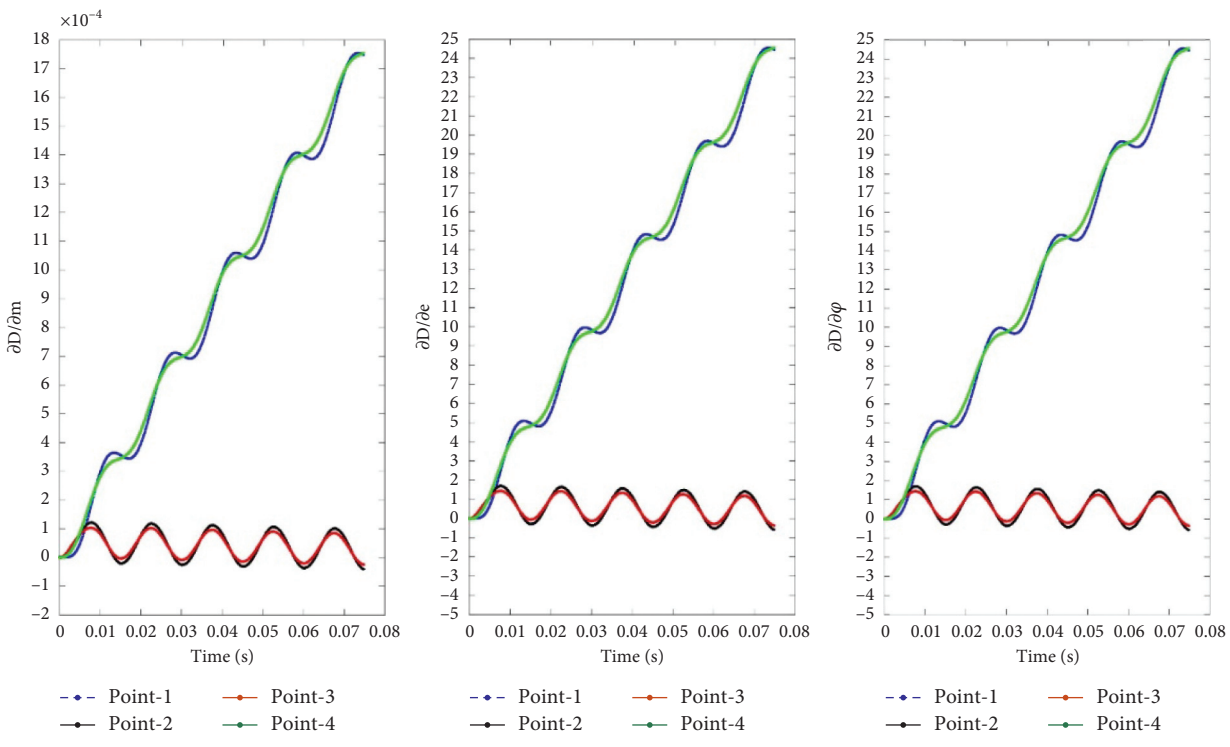
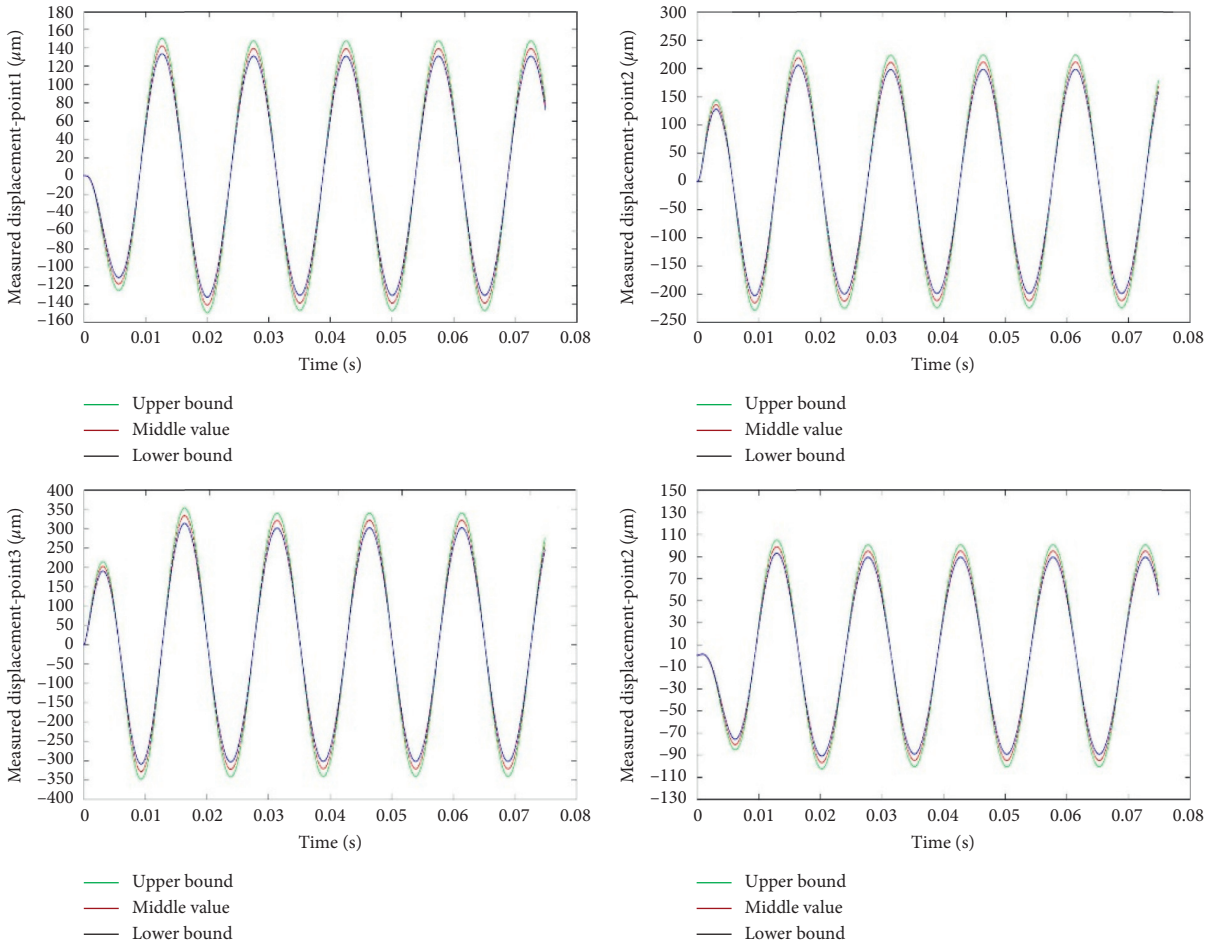


Figure 4: The identified bearing loads.



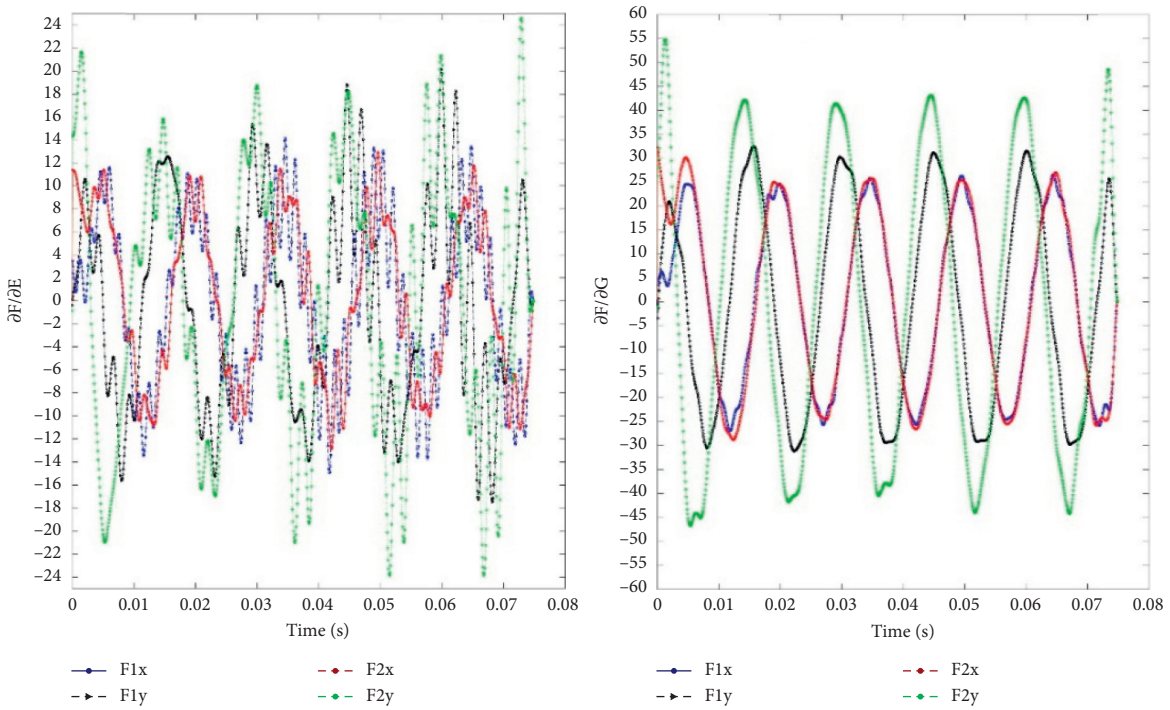
(a)

Figure 5: Continued.



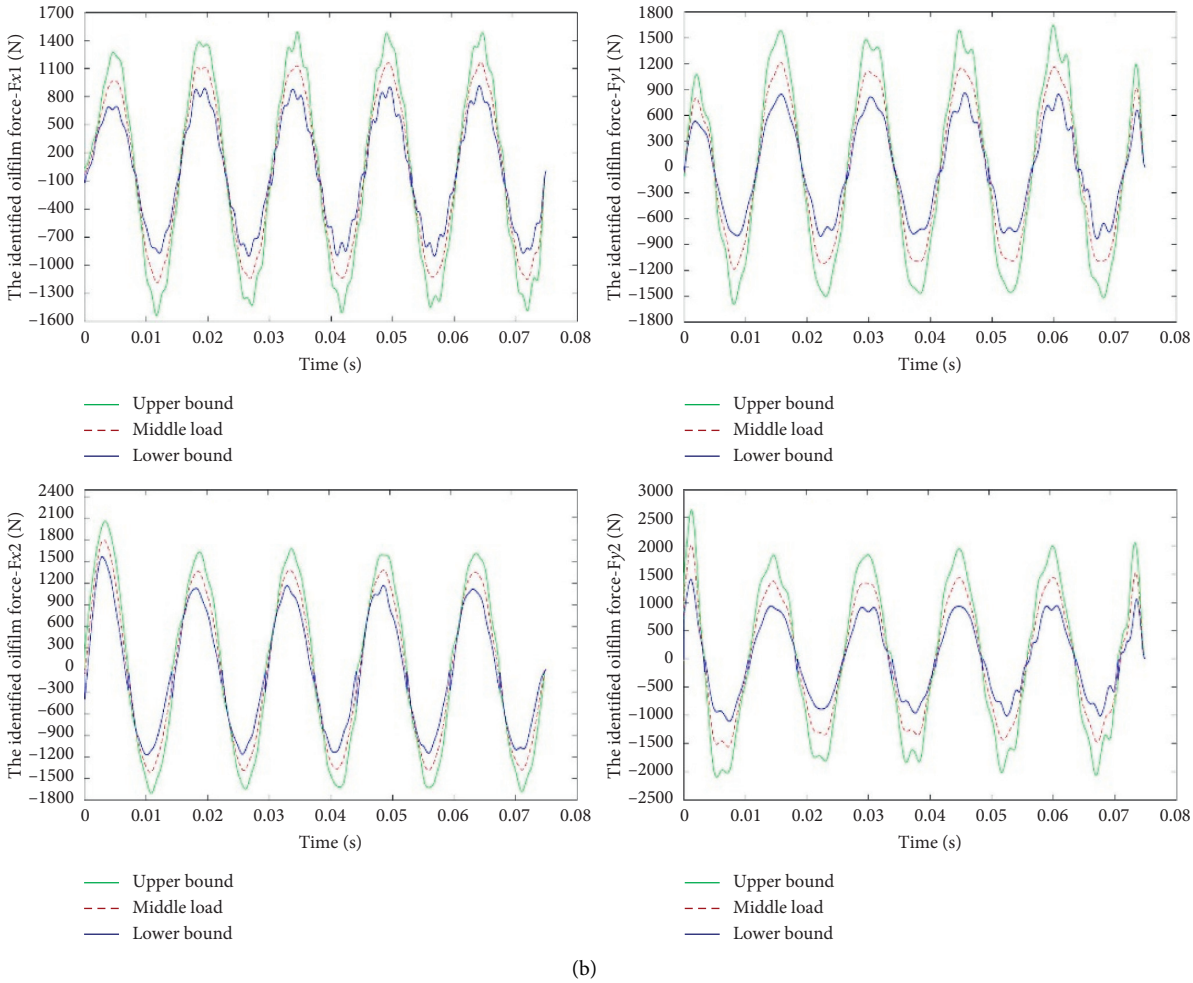
(b)

Figure 5: The rotor structure transient response: (a) the sensitivity curve of the transient response to the interval parameters; (b) the bounds of the rotor structure transient response.



(a)

Figure 6: Continued.



(b)

Figure 6: The identified uncertain bearing load: (a) the sensitivity curve of the bearing load to probabilistic parameters; (b) identified bearing load boundary at 3% uncertainty level.

6. Conclusions

Bearing load is the key factor that determines the lifetime and reliability of the rotor system. Bearing load plays a great role in health monitoring and fault diagnosis. Due to the existence of uncertain parameters, bearing load is a distribution strip. Based on the interval and perturbation theory and the improved regularization method, this paper proposes a new bearing load identification framework for the bearing-rotor system with uncertainty and measurement noise. The problem of bearing load identification is transformed into two kinds of certain inverse problems, namely, the bearing load identification on the mean value of uncertain parameters and the sensitivity identification of bearing load with respect to each uncertain parameter. The improved regularization method can overcome the ill-posedness of bearing load reconstruction for a deterministic rotor system. In the numerical example, the present method can stably identify the bounds of the bearing load only by knowing the bounds of the interval parameters and the statistical characteristics of the probability parameters, which has a certain practical value in engineering.

Data Availability

The data supporting the conclusion of the article are included in the relevant figures and tables in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 51775180) and Hunan Provincial Natural Science Foundation of China (grant no. 2019JJ60036).

References

[1] Y. Xiao and Z. Hua, "Misalignment fault prediction of wind turbines based on combined forecasting model," *Algorithms*, vol. 56, no. 13, pp. 1–21, 2020.

[2] O. Matsushita, M. Tanaka, H. Kanki, M. Kobayashi, and P. Keogh, "Stability problems in rotor systems," in *Vibrations of Rotating Machinery, Mathematics for Industry*, vol. 16, pp. 287–319, Springer, Berlin, Germany, 2017.

[3] V. A. Golubkov, A. U. Gulevitskiy, A. A. Ovodenko, V. F. Shishlakov, and A. S. Smirnova, "Structural analysis of vibrational characteristics of rotor systems," *Journal of Physics: Conference Series*, vol. 1515, Article ID 022088, 2020.

[4] X. Pang, J. Shao, X. Xue, and W. Jiang, "Application of axis orbit image optimization in fault diagnosis for rotor system," *International Journal of Rotating Machinery*, vol. 2020, Article ID 9540791, 13 pages, 2020.

[5] S. Abdenour, M. Kamal, C. Guy, and R. Hubert, "Prediction of bearing failures by the analysis of the time series," *Mechanical Systems and Signal Processing*, vol. 139, pp. 1–28, 2020.

[6] S. H. Liu, G. L. Wang, Q. L. Zhu et al., "Bearing loads of dual rotor-bearing system under stationary and transient conditions," *Journal of Beijing Institute of Technology*, vol. 21, no. 4, pp. 460–465, 2012.

[7] G. Y. Ying, W. J. Wu, S. L. Liu et al., "Research on bearing load sensitivity for multi-support rotor system," *Zhejiang Electric Power*, vol. 36, no. 1, pp. 39–42, 2017.

[8] H. Deng, Y. Diao, W. Wu, J. Zhang, M. Ma, and X. Zhong, "A high-speed D-CART online fault diagnosis algorithm for rotor systems," *Applied Intelligence*, vol. 50, no. 1, pp. 29–41, 2020.

[9] M. Cen, X. F. Liu, and C. Y. Fu, "Combined method of uncertain measurement fusion," in *Proceedings of the 2008 Chinese Control and Decision Conference*, pp. 1584–1587, Yantai, China, 2008.

[10] J. Du, "Optimized state estimation of uncertain linear time-varying complex networks with random sensor delay subject to uncertain probabilities," *IEEE Access*, vol. 7, pp. 113005–113016, 2019.

[11] J. X. Li, J. Hu, D. Y. Chen, and Z. H. Wu, "Distributed extended Kalman filtering for state-saturated nonlinear systems subject to randomly occurring cyberattacks with uncertain probabilities," *Advances in Difference Equations*, vol. 437, pp. 1–26, 2020.

[12] J. Liu, X. Han, C. Jiang, H. M. Ning, and Y. C. Bai, "Dynamic load identification for uncertain structures based on interval analysis and regularization method," *International Journal of Computational Methods*, vol. 8, no. 4, pp. 667–683, 2011.

[13] W. Zhang, J. Liu, C. Cho, and X. Han, "A hybrid parameter identification method based on Bayesian approach and interval analysis for uncertain structures," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 853–865, 2015.

[14] Z. Q. Liu and Y. W. Bao, "Vertex combination method for heat transfer analysis of structures with uncertain parameters," in *Proceedings of the 2016 3RD International Conference on Mechanics and Mechatronics Research*, Chongqing, China, 2016.

[15] J. M. Cheng, L. Yan, C. Zhang, and Z. Pei, "A combined method to deal with uncertain data in fuzzy *k*–nearest neighbor classifier," *in Proceedings of the 9th International FLINS Conference*, vol. 4, pp. 282–287, Computational Intelligence: Foundations and Applications, EMei, Chengdu, China, 2010.

[16] C. Hu, B. D. Youn, and P. Wang, "Fundamentals of probability theory," in *Engineering Design under Uncertainty and Health Prognostics, Springer Series in Reliability Engineering*, pp. 11–51, Springer, Cham, Switzerland, 2019.

[17] J. Jiang, H. Tang, M. S. Mohamed, S. Luo, and J. Chen, "Augmented tikhonov regularization method for dynamic load identification," *Applied Science*, vol. 6348, no. 10, pp. 1–23, 2020.

[18] L. Wang, Y. Huang, Y. Xie, and Y. Du, "A new regularization method for dynamic load identification," *Science Progress*, vol. 103, no. 3, pp. 1–15, 2020.

[19] M. I. Friswell, J. E. T. Penny, S. D. Garvey et al., *Dynamics of Rotating Machines*, Cambridge University Press, Cambridge, MA, USA, 2010.

[20] W. Mao, X. Han, G. Liu, and J. Liu, "Bearing dynamic parameters identification of a flexible rotor-bearing system based on transfer matrix method," *Inverse Problems in Science and Engineering*, vol. 24, no. 3, pp. 372–392, 2016.

[21] K. Li, J. Liu, X. Han, C. Jiang, and H. Qin, "Identification of oil-film coefficients for a rotor-journal bearing system based on equivalent load reconstruction," *Tribology International*, vol. 104, pp. 285–293, 2016.

[22] J. H. Yang, "An efficient method for Bayesian system identification based on Markov chain Monte Carlo simulation," *International Journal of Lifecycle Performance Engineering*, vol. 3, no. 1, pp. 20–34, 2019.

[23] T. W. Sherson, R. Heusdens, and W. B. Kleijn, "Derivation and analysis of the primal-dual method of multipliers based on monotone operator theory," *IEEE Transactions on Signal and Information Processing Over Networks*, vol. 5, no. 2, pp. 334–347, 2019.

[24] X. Sun, "Identification method of dynamic loads for stochastic structures based on matrix perturbation theory," *Journal of Mechanical Engineering*, vol. 50, no. 13, pp. 148–156, 2014.

[25] G. R. Liu, M. Chen, and M. Li, "Lower bound of vibration modes using the node-based smoothed finite element method (NS-FEM)," *International Journal of Computational Methods*, vol. 14, no. 2, Article ID 17500361, 2017.

[26] B. C. Wen, J. L. Gu, S. B. Xia et al., *The Higher Rotor Dynamics*, China Machine Press, Beijing, China, 2000.

[27] W. Mao, J. Li, Z. Huang, and J. Liu, "Bearing dynamic parameters identification for a sliding bearing-rotor system with uncertainty," *Inverse Problems in Science and Engineering*, vol. 26, no. 8, pp. 1094–1108, 2018.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6640387, 14 pages
https://doi.org/10.1155/2021/6640387

Research Article

A Novel Parameter-Adaptive VMD Method Based on Grey Wolf Optimization with Minimum Average Mutual Information for Incipient Fault Detection

Wang Xu[1] and Jinfei Hu[2]

[1]*School of Marine Engineering Equipments, Zhejiang Ocean University, Zhoushan, Zhejiang 316022, China*
[2]*National Engineering Research Center for Marine Aquaculture, Institute of Innovation and Application, Zhejiang Ocean University, Zhoushan, Zhejiang 316022, China*

Correspondence should be addressed to Jinfei Hu; hujinfei@zjou.edu.cn

Received 4 December 2020; Revised 18 January 2021; Accepted 22 January 2021; Published 2 February 2021

Academic Editor: Jinde Zheng

Recently, variational mode decomposition (VMD) has attracted wide attention on mechanical vibration signal analysis. However, there are still some dilemmas in the application of VMD, such as the determination of the number of mode decomposition K and quadratic penalty term α. In order to acquire appropriate parameters of VMD, an improved parameter-adaptive VMD method based on grey wolf optimizer (GWO) is developed by taking the minimum average mutual information into consideration (GWOMI). Firstly, the parameters (K, α) are adaptively determined through GWOMI. Then, the vibration signal is decomposed by the developed method and effective modes are extracted according to the maximum kurtosis. Finally, the extracted modes are processed by Hilbert envelope analysis to acquire the incipient fault features. With the simulation and experimental analysis, it is clearly found that the developed method is effective and performs better than some existing ones.

1. Introduction

Rotating machinery has been extensively employed in the manufacturing, traffic and transportation, marine vessel, etc. Rolling bearings are one of the most crucial parts of them and their unexpected failures usually lead to great production loss and high repair cost. Therefore, it is of great significance for fault detection at their incipient faulty stage to ensure the reliable and safely running of the machinery [1–4]. If a local fault occurs in a rolling bearing, some transient pulses will appear in the vibration signal. These transient pulses analysis can provide a possible solution for fault detection and diagnosis. However, these pulse signals caused by the early fault are usually very difficult for detection. Meanwhile, they are nonstationary, nonlinear, and easily covered by strong background noise, which makes it almost impossible to extract the incipient fault features from the raw signals [5, 6].

To address these problems, many methods have been developed based on signal processing. Time-frequency analysis is one of the most widely used signal processing methods to extract the fault feature. Wavelet transform is a typical time-frequency analysis method, which can decompose vibration signal into some wavelets and provide some significant local information. However, it is a non-adaptive signal analysis method as wavelet basis functions are determined in advance [7–11]. An empirical mode decomposition (EMD) method was developed by Huang et al. [12]. EMD could decompose a signal into some modes and a residue. And EMD has strong adaptive decomposition ability to time-varying signals [7]. Lu et al. [13, 14] investigated a kind of method based on an improved genetic algorithm and EMD, which could extract the fault features. However, there are some weaknesses of the methods based on EMD, such as poor antinoise capability and mode aliasing [15]. To address these above problems, Zheng et al.

proposed a modified EMD approach, which was employed to process the nonstationary signal [16]. Ensemble empirical mode decomposition (EEMD) method was developed by Wu et al. [17]. In the work [18], this method was successfully used for the fault diagnosis of bearing signals. However, it would introduce new noises and reduce the computational efficiency. Zheng et al. put forward a partly EEMD method to overcome these problems [19].

VMD is a novel signal analysis method developed by Dragomiretskiy [20]. This method can decompose signals into some intrinsic mode functions, which have limited bandwidths and different center frequencies. It can overcome the shortcoming of mode aliasing and has been widely used in the fault diagnosis of rotating machinery. However, the effect of VMD relies on the choice of its parameters (K, α). In practical applications, it is difficult to determine these parameters, and problems overdecomposition or underdecomposition may happen when parameters are not selected correctly [21, 22]. Shen et al. developed a modified VMD based on initial center frequency, which was successfully used in the fault diagnosis of rotating machinery [23]. Zhang et al. utilized the grasshopper optimization algorithm to improve the parameter adaptiveness of VMD and this method was used to extract fault features successfully [24]. Zhu et al. developed an adaptive VMD method, which decomposed a complex signal into some band-limited intrinsic mode functions [25]. Zhao et al. utilized a single-objective salp swarm algorithm to optimize VMD parameters, which could reduce mode aliasing and kept the fidelity of complex vibration signals [26]. Gu et al. reported an adaptive VMD based on GWO. In this method, the minimum average envelope entropy was used as the objective function [3]. However, signal decomposition, whose parameters are obtained by the minimum average envelope entropy, may result in the loss of some fault features.

Inspired by the above research results, an improved parametric-adaptive VMD based on grey wolf optimizer with mutual information algorithm (GWOMI-VMD) is developed in this work. The GWO algorithm has a great global searching ability with high convergence accuracy. However, it is noted that the optimization objective function has a great impact on the optimization results. And there are still some overdecomposition or underdecomposition problems in the VMD method when the most commonly used objective functions are applied in GWO, such as smoothness index, sparsity measurement, and correlation coefficient [27]. Mutual information (MI) is a valuable information measure in information theory, which is used to represent the mutual dependence between two variables. In this paper, minimum average mutual information (MAMI) is selected as the optimization objective function to reflect the decomposition effect of VMD. Then, the parameters are optimized by GWOMI and the signal is decomposed by our proposed method. Finally, effective modes are extracted according to the maximum kurtosis and envelope spectrum analysis is applied to the effective modes to acquire fault features.

The remainder of this paper is organized as follows: the main principles of VMD and GWO are briefly introduced in Section 2. In Section 3, an improved parametric-adaptive VMD method based on GWO with mutual information is given in detail. In Section 4 and Section 5, both simulation and experimental signals are utilized to demonstrate the validity of the GWOMI-VMD, and some comparisons between the proposed method and traditional VMD, and particle swarm optimization (PSO) optimized VMD and EMD methods are provided. In Section 6, some conclusions are drawn.

2. The Principle of VMD and GWO

2.1. Variational Mode Decomposition. VMD is a novel method of signal processing, which adaptively decomposes signals into K modes and these modes have different center frequencies ω_k. The crucial issue of the VMD algorithm is how to acquire the solution of the constrained variational problem, which is formulated in equation (1) as follows:

$$\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \quad \text{subject to} \sum_k u_k = f, \tag{1}$$

where $\{\mathbf{u}_k\} = \{u_1, u_2, \ldots, u_k\}$ and $\{\omega_k\}$ represent each mode component and center frequency, respectively. We introduce Lagrange multiplier λ and balance parameter α to transform equation (1) into an unconstrained one, which is formulated in the following equation:

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

The alternating direction method of multipliers is employed to acquire the saddle point of equation (2). As shown in equations (3) and (4), $\hat{u}_k^{n+1}$ and ω_k^{n+1} are updated to obtain the optimal values:

$$\widehat{u}_k^{n+1}(\omega) \longleftarrow \frac{\widehat{f}(\omega) - \sum_{i \neq k} u_i(\omega) + (\widehat{\lambda}(\omega)/2)}{1 + 2\alpha(\omega - \omega_k)^2}, \tag{3}$$

$$\omega_k^{n+1} \longleftarrow \frac{\int_0^\infty \omega \left\| \widehat{u}_k^{n+1}(\omega) \right\|^2 \mathrm{d}\omega}{\int_0^\infty \left\| \widehat{u}_k^{n+1} \right\|^2 \mathrm{d}\omega}, \tag{4}$$

where n is the number of iterations.

After each update, the modes and corresponding center frequencies are obtained. After that, the Lagrange multiplier is updated as shown in the following equation:

$$\widehat{\lambda}^{n+1}(\omega) \longleftarrow \widehat{\lambda}^n(\omega) + \tau\left(\widehat{f}(\omega) - \sum_k \widehat{u}_k^{n+1}(\omega)\right). \tag{5}$$

When equation (6) is satisfied, the above iteration is terminated:

$$\sum_k \frac{\left\| \widehat{u}_k^{n+1} - \widehat{u}_k^n \right\|_2^2}{\left\| \widehat{u}_k^n \right\|_2^2} < \varepsilon. \tag{6}$$

We can see from equation (3) that the modes are recognized as some Wiener filters. In equation (4), ω_k is the gravity of the corresponding mode power spectrum. And, in equation (2), the quadratic penalty term α can suppress mode aliasing, which is one of the critical parameters in VMD.

2.2. Grey Wolf Optimizer. Compared with other bionic intelligent algorithms, the grey wolf optimization algorithm has faster convergence and better global search capability. First, a group of grey wolves is randomly generated and they are divided into four social hierarchies according to their objective function values from best to worst, namely, α, β, δ, and ω. Next, the hunting (optimization) is guided by α, under the cooperation with β and δ by the way of circling the prey. Then, ω and other wolves move towards the prey, periodically updating their position, gradually reducing the distance between them and the prey, and finally successfully hunting [28]. The main procedure of the algorithm is given as follows.

The distances between the wolves and their prey during the hunting are formulated in equations (7) and (8), respectively:

$$\overrightarrow{D} = \left| \overrightarrow{h} \cdot \overrightarrow{X}_P(t) - \overrightarrow{X}_t \right|, \tag{7}$$

$$\overrightarrow{X}_p(t+1) = \overrightarrow{X}_p(t) - \overrightarrow{A} \cdot \overrightarrow{D}, \tag{8}$$

where $\overrightarrow{X}_p(t)$, $\overrightarrow{X}(t)$, and t stand for prey position, wolf position, and iteration number, respectively. The coefficients $\overrightarrow{A}$ and $\overrightarrow{h}$ are calculated as follows:

$$\begin{aligned} \overrightarrow{A} &= 2\overrightarrow{a} \cdot \overrightarrow{n}_1 - \overrightarrow{a}, \\ \overrightarrow{h} &= 2 \cdot \overrightarrow{n}_2, \end{aligned} \tag{9}$$

where $\overrightarrow{a}$ decreases gradually from 2 to 0 and $\overrightarrow{n}_1$ and $\overrightarrow{n}_2$ are random vectors in [0, 1].

During the hunting, α, β, and δ are obtained for the first three optimal solutions in the optimization space, and the positions of β and δ are changed according to α. The specific process is shown as follows:

$$\begin{aligned} \overrightarrow{D}_F &= \left| \overrightarrow{h}_i \cdot \overrightarrow{X}_F(t) - \overrightarrow{X}(t) \right|, \\ \overrightarrow{X}_i &= \overrightarrow{X}_F - \overrightarrow{A}_i \cdot \overrightarrow{D}_F, \\ \overrightarrow{X}_P(t+1) &= \frac{\overrightarrow{X}_1 + \overrightarrow{X}_2 + \overrightarrow{X}_3}{3}, \end{aligned} \tag{10}$$

where $F = \alpha, \beta, \delta$ and $i = 1, 2, 3$.

3. The Proposed Method

3.1. Minimum Average Mutual Information. Mutual information (MI) is a nonparametric and nonlinear metric in information theory and can be used to reflect the mutual dependence between two variables. The expression is as follows:

$$MI(a,b) = \sum_{a \in a} \sum_{b \in b} f(a,b) lg \frac{f(a,b)}{f(a) f(b)}, \tag{11}$$

where $f(a,b)$ represents the joint probability density function of a and b and $f(a)$ and $f(b)$ are edge probability density functions of a and b, respectively. If there is no overlapping information between a and b, in other words, a and b are independent variables, the MI value between them is 0. On the contrary, if there is more overlapping information between a and b, the MI value is closer to 1 [29].

In this paper, the minimum average mutual information (MAMI) is selected as the optimization objective function and its expression is given by

$$\text{objective}_{\text{GWO}} = \text{Minimum} \frac{\sum_{j=2}^K \mathbf{V}_{\text{MI}(u(j-1),u(j))}^{\mathbf{u_k}}}{K-1}, \tag{12}$$

where K is the number of mode decomposition and $\sum_{j=2}^K \mathbf{V}_{MI(u(j-1),u(j))}^{\mathbf{u_k}}$ represents the sum of MI between each neighboring mode. The smaller the MI is, the better the two modes are decomposed. Similarly, the smaller the average MI is, the better the overall effect of VMD will be. When $((\sum_{j=2}^K \mathbf{V}_{\text{MI}(u(j-1),u(j))}^{\mathbf{u_k}})/(K-1))$ is minimized, the signal can be properly decomposed into a series of modes by VMD. At this moment, the corresponding parameters (K, α) are optimal.

3.2. Mode Extraction Based on Maximum Kurtosis. After the signal is processed by the GWOMI-VMD method, the effective mode is extracted according to the maximum kurtosis. Kurtosis describes the peak value of waveform and can reflect the numerical statistics of vibration signal distribution features [30]. The expression is as follows:

$$K = \frac{\sum_{i=1}^{n} (x_i - \mu)^4}{n\sigma^4}, \tag{13}$$

where μ is the mean of signal $[x_1, x_2, \ldots, x_n]$ and σ and n represent the standard deviation of the signal and the length of the signal, respectively.

As we all know, the more the impact components in the signal, the greater the corresponding kurtosis. Therefore, the effective mode with the maximum kurtosis is selected for the fault feature extraction.

3.3. GWOMI-VMD. In this paper, we develop an incipient fault characteristics extraction method from the rolling bearing vibration signal, named GWOMI-VMD. The steps of GWOMI-VMD are described briefly as follows:

(1) Input the vibration signal, set the parameter range of VMD parameters K and α, and initialize the parameters of GWO.

(2) Decompose the signal by VMD and calculate the average mutual information of the decomposed modes.

(3) Update the objective function value and record the position of the wolves.

(4) Judge whether the number of iterations exceeds the maximum number of iterations. If so, the iteration will be terminated. If not, go back to step 2 to continue the iteration.

(5) Obtain the minimum value of the optimization objective function and save the optimal position (K, α).

(6) The signal is decomposed by VMD with the optimized (K, α), and the effective mode is extracted based on the maximum kurtosis.

(7) The envelope spectrum of the extracted mode is calculated and the fault characteristics are correspondingly obtained.

The flowchart of the developed method is illustrated in Figure 1.

4. Simulation Analysis

In order to demonstrate the performance of the developed method, a simulation signal y is configured as shown in the following equation:

$$\begin{cases} y_1(t) = \cos(2 * pi * f_1 * t), \\ y_2(t) = \frac{1}{4}(\cos(2 * pi * f_2 * t)), \\ y_3(t) = \frac{1}{16}(\cos(2 * pi * f_3 * t)), \\ y = y_1 + y_2 + y_3 + \text{noise}, \end{cases} \tag{14}$$

where $f_1 = 2\,\text{Hz}$, $f_2 = 24\,\text{Hz}$, and $f_3 = 288\,\text{Hz}$, the noise signal is random numbers distributed in $[-2, 2]$, and the sampling frequency is 1000 Hz. Figure 2 shows the composite signal y in the time domain and its frequency spectrum. Figure 3 shows each component of the composite signal y. We can clearly see the frequency components f_1 and f_2 from the frequency spectrum, but the frequency component f_3 has been completely submerged by the random noises in Figure 2.

The proposed method is performed to search the optimal parameters of VMD. In the optimization process, the optimization objective function adopts the minimum average mutual information (MAMI). And the convergence diagram of the optimization objective function is illustrated in Figure 4. As seen from Figure 4, the objective function value is 0.017 when the iteration time reaches 14, and, at this moment, the curve of the objective function converges. Besides, the corresponding parameters of VMD are optimized when the iteration time reaches 14. Then, the optimized parameters (9, 4721) are substituted into VMD to decompose the composite signal y, and Figure 5 shows the resulting modes.

From Figure 5, we can see that three frequency components f_1, f_2, and f_3 are extracted, respectively, which are consistent with the frequencies of the simulation signal y. It is worth mentioning that there is no serious aliasing between these modes.

In addition, the simulation signal is processed by the traditional VMD with parameters (3, 2000) as shown in Figure 6. We can clearly see that both frequency components f_1 and f_2 are extracted in the first mode. However, the frequency component f_3 cannot be extracted from these modes. Lastly, the particle swarm optimization (PSO) is utilized to improve the traditional VMD and optimal parameters (10, 4883) are obtained correspondingly. Then, the composite signal y is decomposed and the decomposition results are illustrated in Figure 7. From Figure 7, we can see that frequency components f_1 and f_2 are also extracted, but frequency component f_3 cannot be successfully extracted.

Through the above comparison analysis, we can find that the three frequency components f_1, f_2, and f_3 can be extracted accurately through the proposed method. However, the traditional VMD and PSO-VMD could not accurately extract the frequency component f_3. So, it is obvious that our proposed method GWOMI-VMD has a better performance than the traditional VMD and PSO-VMD.

5. Experiment Analysis

The validity and superiority of the GWOMI-VMD in weak fault diagnosis of rolling bearing are demonstrated through analyzing the experimental signal. The experimental data comes from the Case Western Reserve University Bearing Data Center [31]. Figure 8 shows the basic layout of the bearing test table.

As seen from Figure 8, the test table consists of the following parts: a dynamometer, an electric motor, and a torque transducer/encoder. The tested bearing supports the motor spindle shaft, and the tested bearing has single point faults. The vibration signal is collected through an

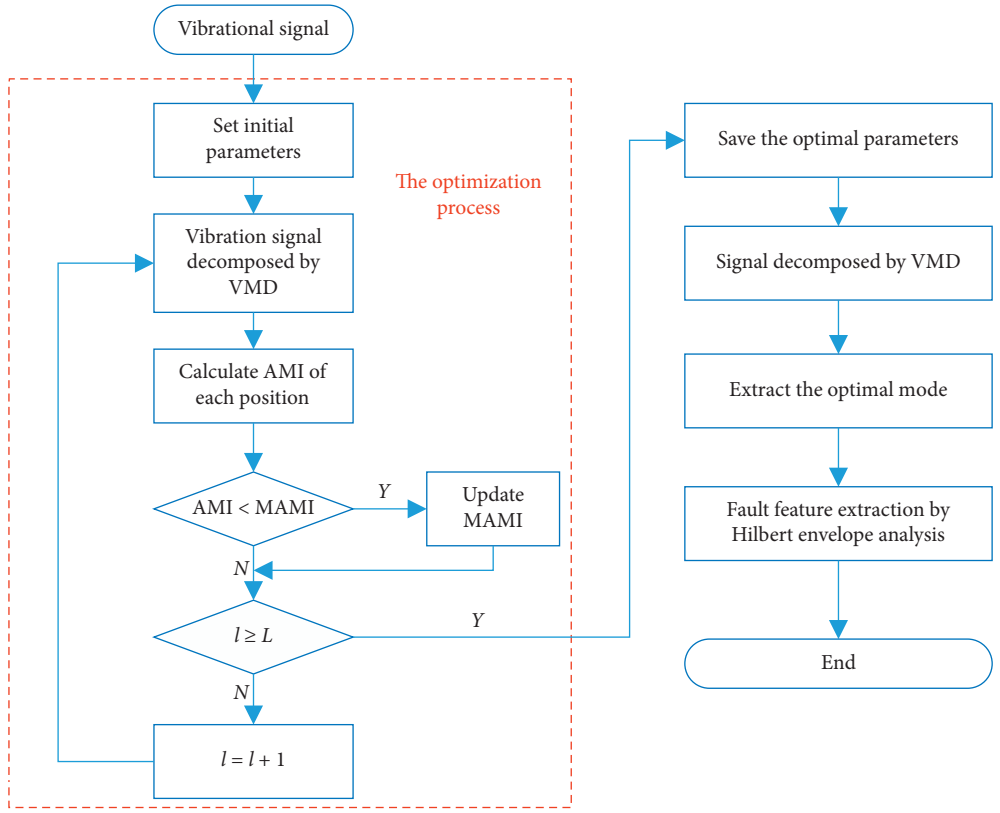


Figure 1: Procedure of the GWOMI-VMD.

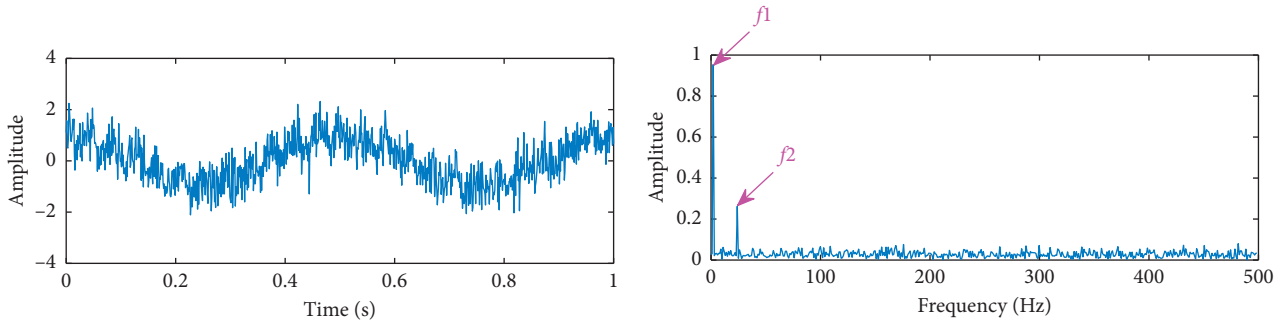


Figure 2: Simulation signal (y): (a) time-domain waveform and (b) frequency spectrum.

accelerometer. The specific technical parameters are listed in Table 1 and 6000 data points are sampled for analysis.

5.1. Bearing Inner Race Fault Feature Extraction. Figure 9 shows the experimental signal and it is processed based on the GWOMI-VMD method. Firstly, the optimal parameters (K, α) are determined according to the GWOMI algorithm. Figure 10 shows the GWOMI and PSO convergence curves for VMD parameter optimization. When the number of iterations is 13, PSO starts to converge at the objective function value of 0.032896. When the number of GWOMI iterations is 7, the objective function value is 0.032876, and the corresponding parameters of VMD are optimal with a value of (10, 5612). We can find that the convergence accuracy and speed of GWOMI are obviously better than those of PSO.

Then, the optimized parameters are substituted into VMD, and the signal is decomposed. Figure 11 shows the modes acquired by decomposition using GWOMI-VMD. The mode, which contains the most useful fault information, cannot be judged intuitively from these modes. Therefore, an index needs to be determined to help extract the effective mode.

As described above, the maximum kurtosis is used as an indicator for mode extraction in this work. The kurtosis diagram of the ten modes is shown in Figure 12. We can see that the kurtosis of the 10^{th} mode is the highest out of them, with a value of 3.113. Therefore, we choose the 10^{th} mode for the following analysis.

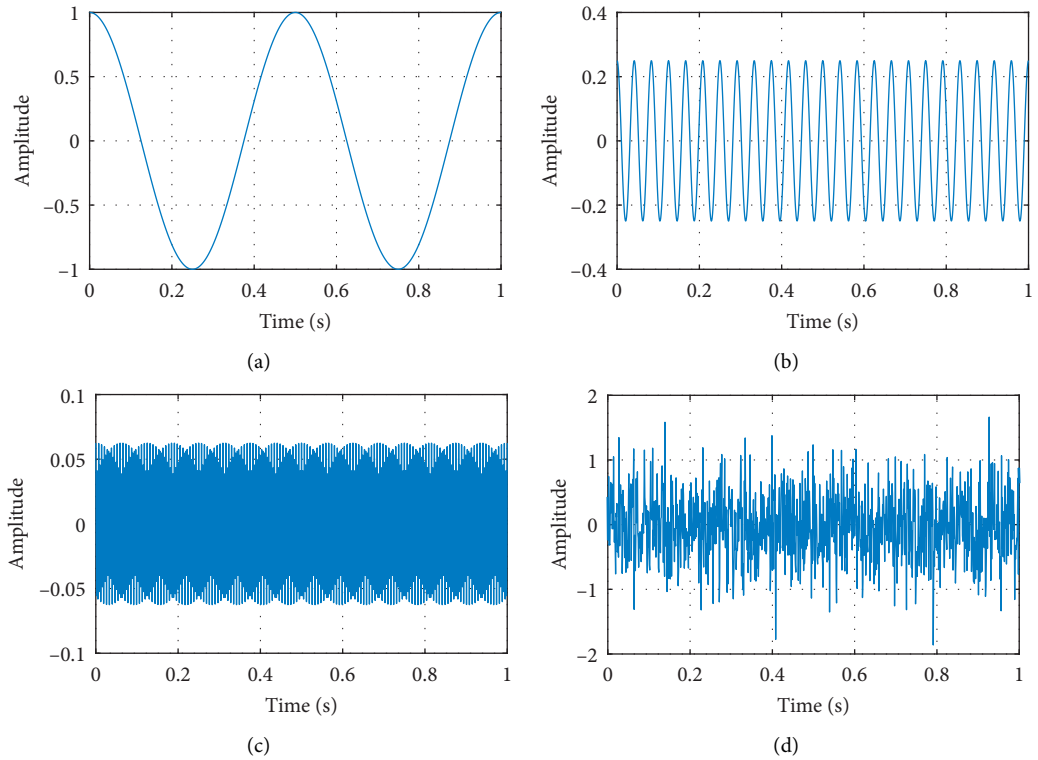


Figure 3: Each part of the simulation signal: (a) harmonic signal y_1, (b) harmonic signal y_2, (c) harmonic signal y_3, and (d) noise signal.

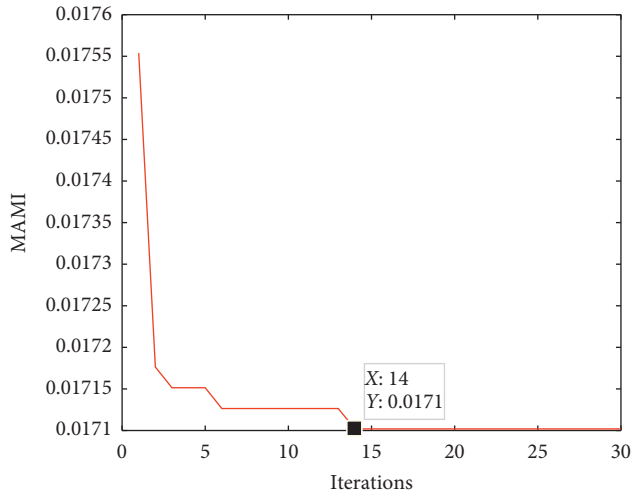


Figure 4: Convergence curve of GWOMI.

Envelope spectrum analysis is performed on the 10th mode and Figure 13 shows the results of the analysis. From Figure 13, the fault frequency *f*, rotation frequency fr, and its second harmonic frequency *2fr* have emerged obviously. What is more, compared with the frequency spectrum in Figure 9, we can find that the background noise is basically eliminated in Figure 13.

Comparison analysis: a performance comparison between maximum kurtosis and maximum correlation coefficient as the mode index is conducted. Figure 14 shows the envelope spectrum of the effective mode, which was extracted based on the maximum correlation coefficient. We find that although fault frequency *f* is extracted, the rotational speed frequency fr is not extracted as shown in Figure 13. In addition, there is a newly generated interference frequency, whose amplitude in the spectrum exceeds the fault frequency *f*.

In addition, the GWOMI-VMD is further compared with traditional VMD. In this paper, the parameters of traditional VMD are set as (5, 2500) based on experience. The

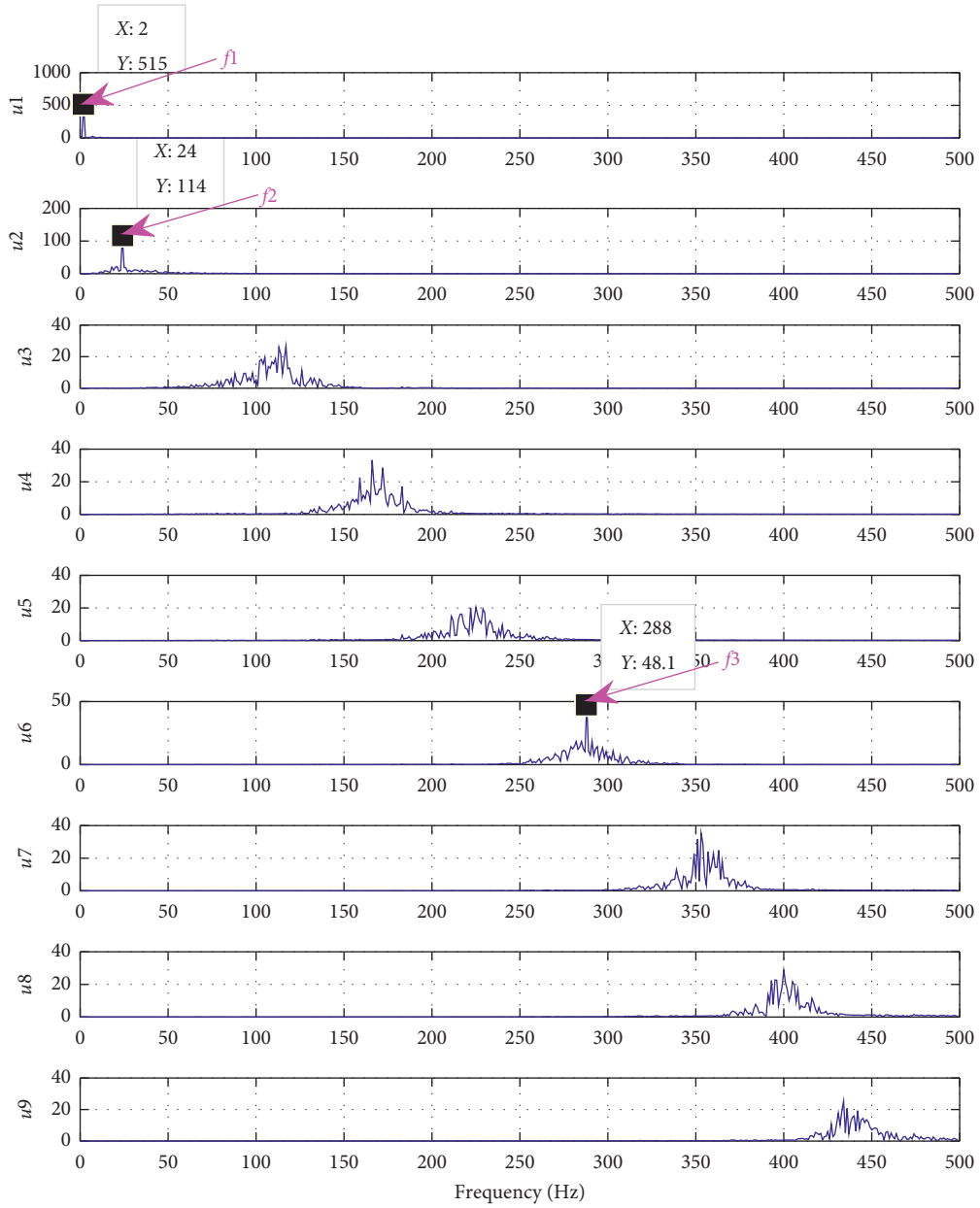


Figure 5: The spectrum of each mode decomposed by the proposed method.

decomposition results of traditional VMD are shown in Figure 15. Then, GWOMI-VMD method is compared with the EMD method and Figure 16(a) shows the first five modes obtained by using the GWOMI-VMD method. Figure 16(b) shows the first five modes obtained by using the EMD method. As seen from Figures 15 and 16(b), there is a certain degree of aliasing between these modes in the case of traditional VMD and EMD.

5.2. Bearing Outer Race Fault Feature Extraction. As demonstrated in the inner race, the optimal parameter of VMD is first obtained with the value of (9, 4340), and then the effective mode is extracted according to the maximum kurtosis. The kurtosis of the 9^{th} mode is the highest, with a value of 6.1024. Therefore, the 9^{th} mode is extracted out and Figure 17(a) shows its envelope spectrum. The frequency components in Figure 17(a) are basically consistent with the calculated fault frequencies listed in Table 1.

A performance comparison between maximum kurtosis and maximum correlation coefficient both as the mode index is conducted. Figure 17(b) gives the spectrum of the effective mode extracted by the maximum correlation coefficient. From Figure 17(b), we can find that the second rotation harmonic frequency 2fr is extracted. However, it is obvious that the fault frequency f is not extracted successfully, compared with Figure 17(a). Therefore, the effective mode extracted by the maximum kurtosis utilized in this paper has certain advantages over the correlation coefficient.

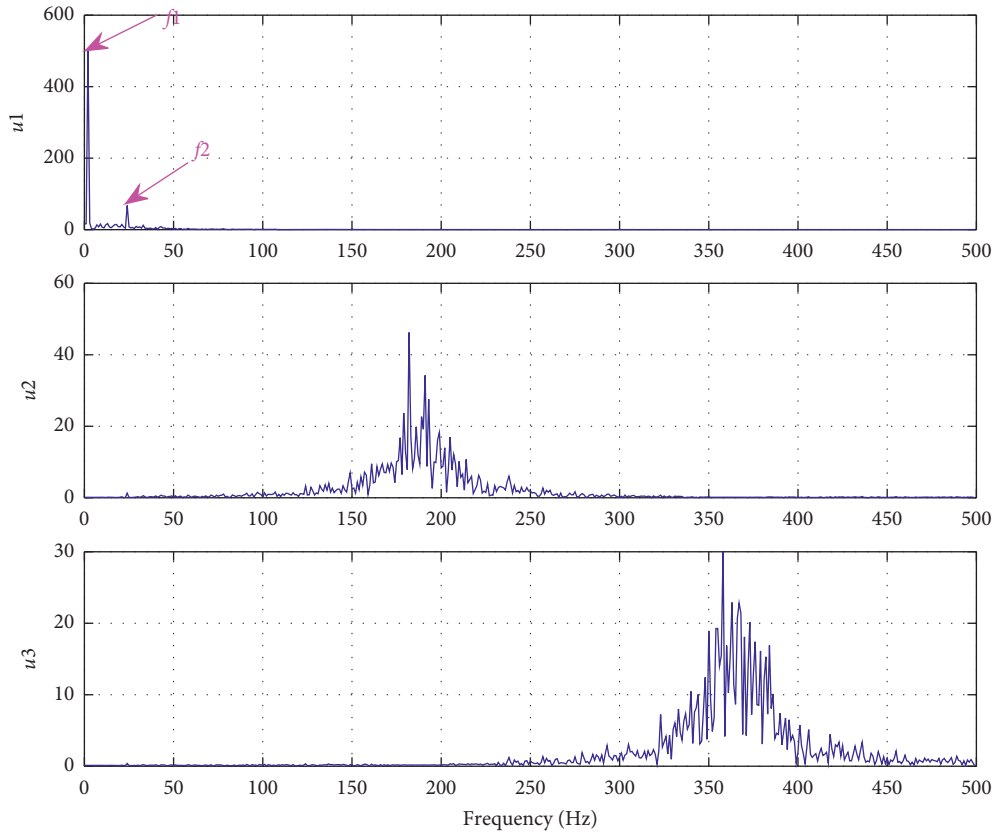


Figure 6: The spectrum of each mode decomposed by traditional VMD.

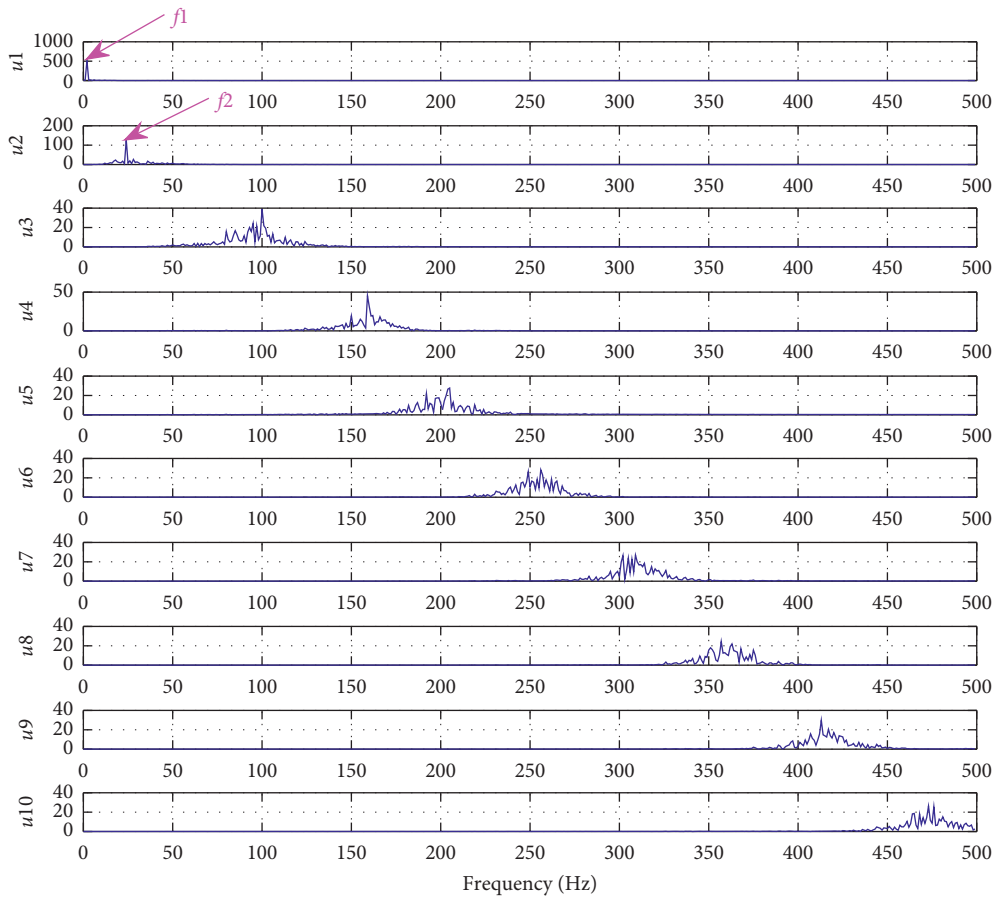


Figure 7: The spectrum of each mode decomposed by PSO-VMD.

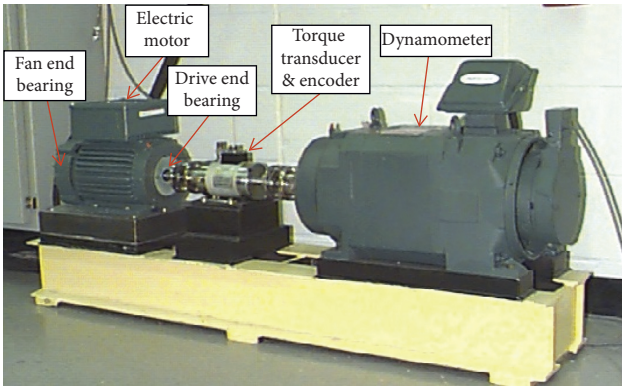


Figure 8: CWRU bearing test stand.

Table 1: Technical parameters.

Test bearing (inner race)		Test bearing (outer race)	
Parameters	Values	Parameters	Values
Fault size	0.007″	Fault size	0.007″
Sampling frequency fs	12000 Hz	Sampling frequency fs	12000 Hz
Fault characteristic frequency f	162 Hz	Fault characteristic frequency f	107 Hz
Rotational frequency fr	29.95 Hz	Rotational frequency fr	29.95 Hz

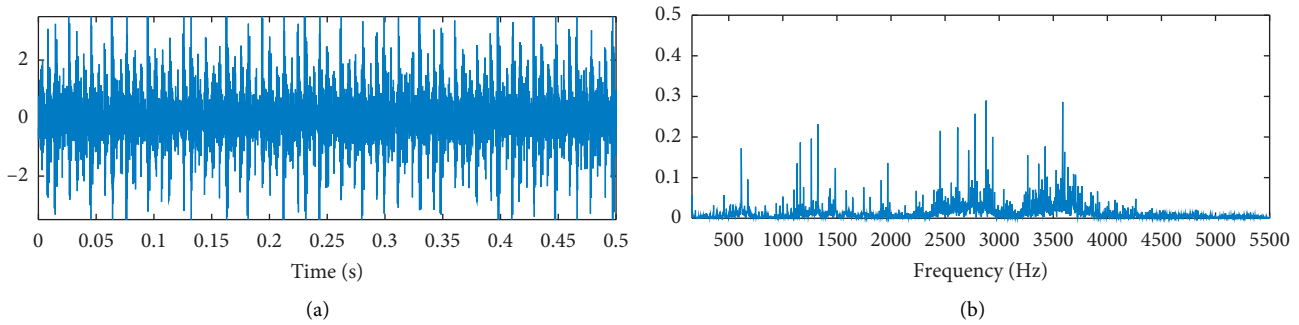


Figure 9: The original signal (a) time domain and (b) frequency domain.

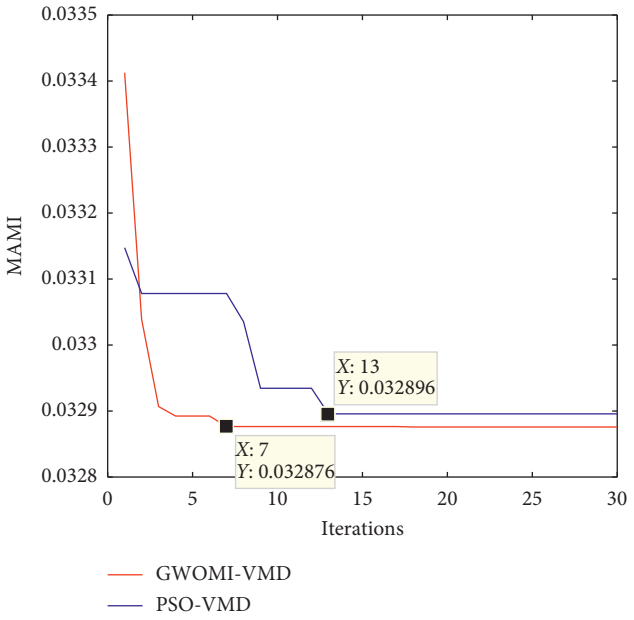


Figure 10: Convergence curves of GWOMI and PSO.

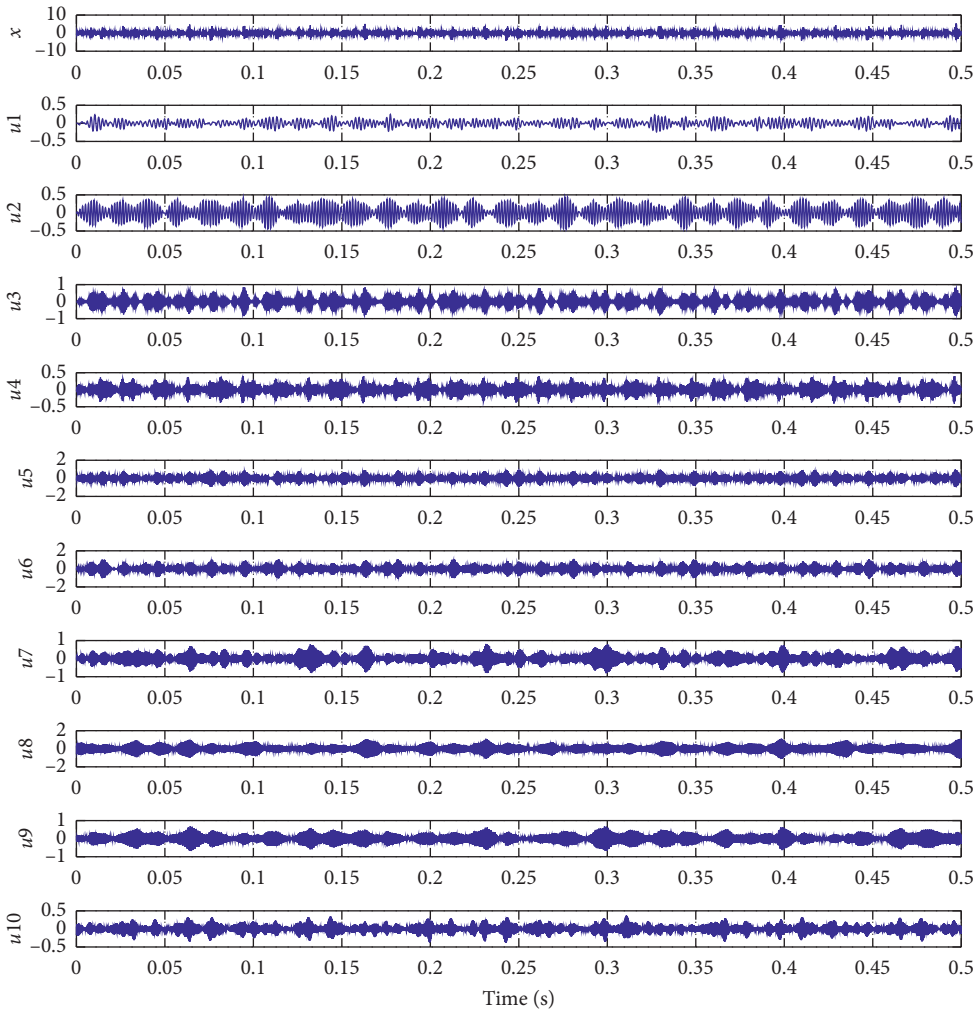


Figure 11: Results of the proposed method.

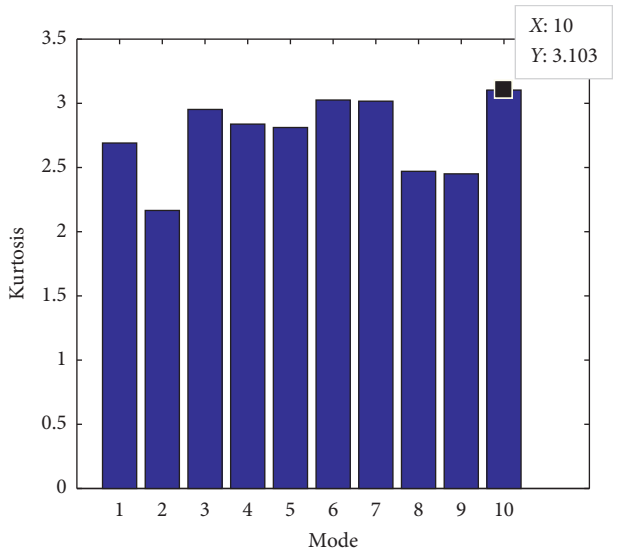


Figure 12: Kurtosis values of each mode.

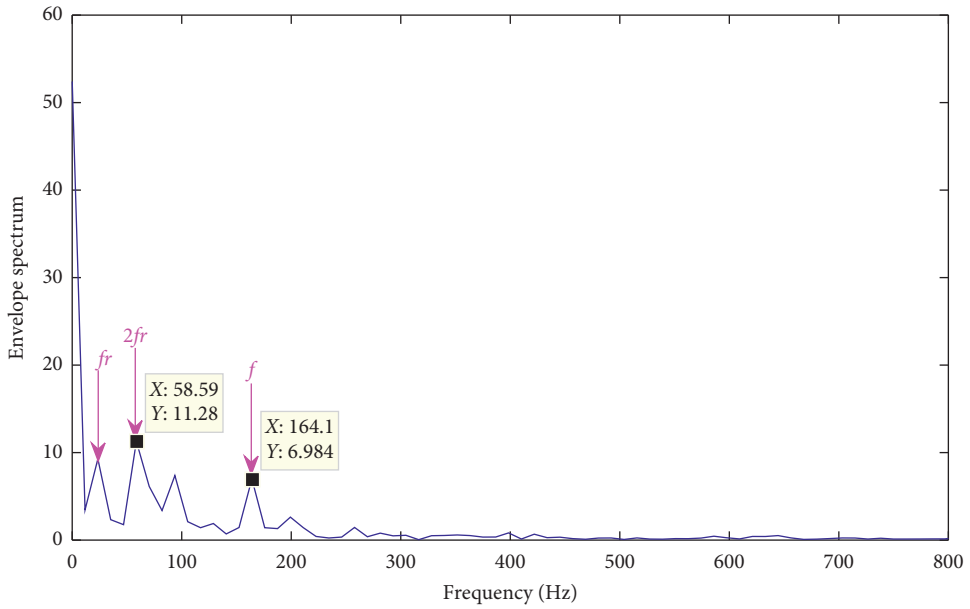


Figure 13: Envelope spectrum of effective mode extracted based on the maximum kurtosis.

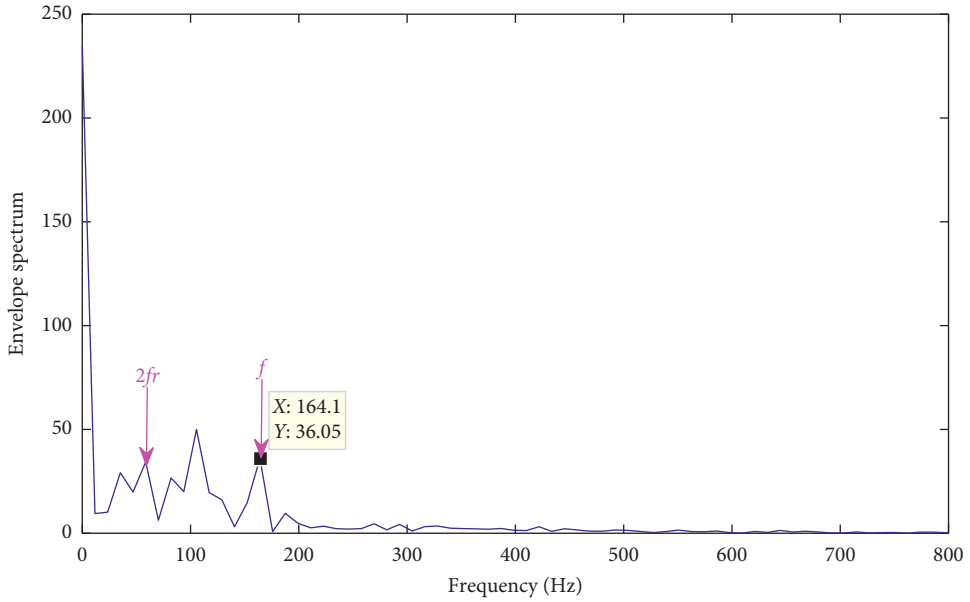


Figure 14: Envelope spectrum of effective mode extracted according to the maximum correlation coefficient.

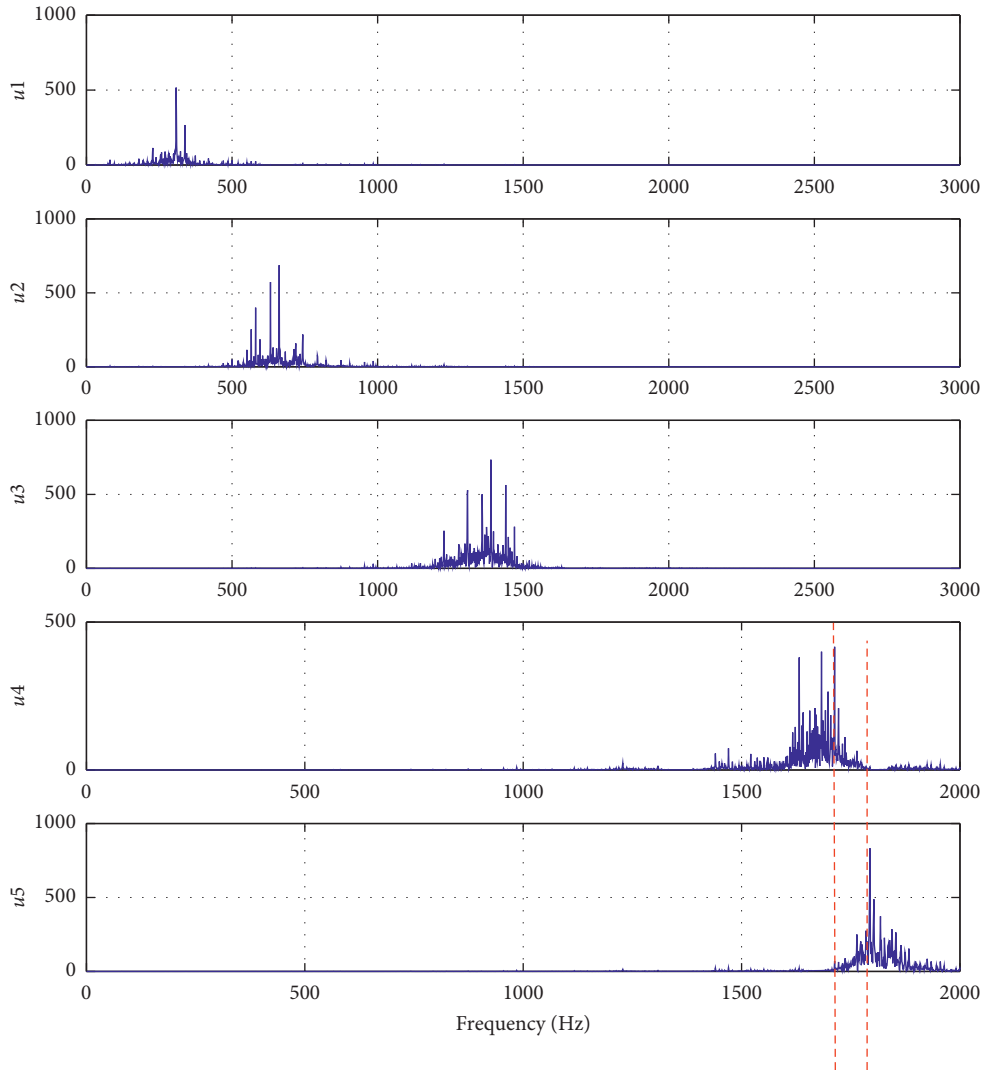


Figure 15: Spectrums of each mode decomposed by traditional VMD.

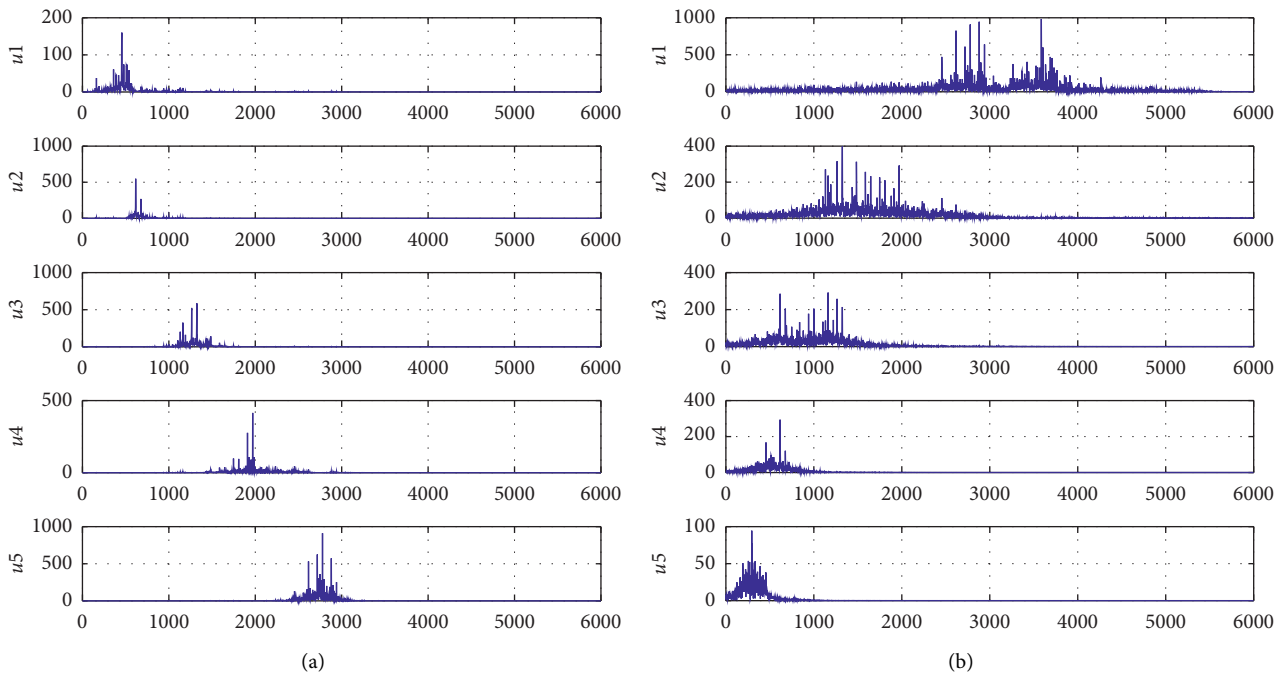


Figure 16: Frequency spectrums of $u1$, $u2$, $u3$, $u4$, and $u5$ decomposed by (a) GWOMI-VMD and (b) EMD.

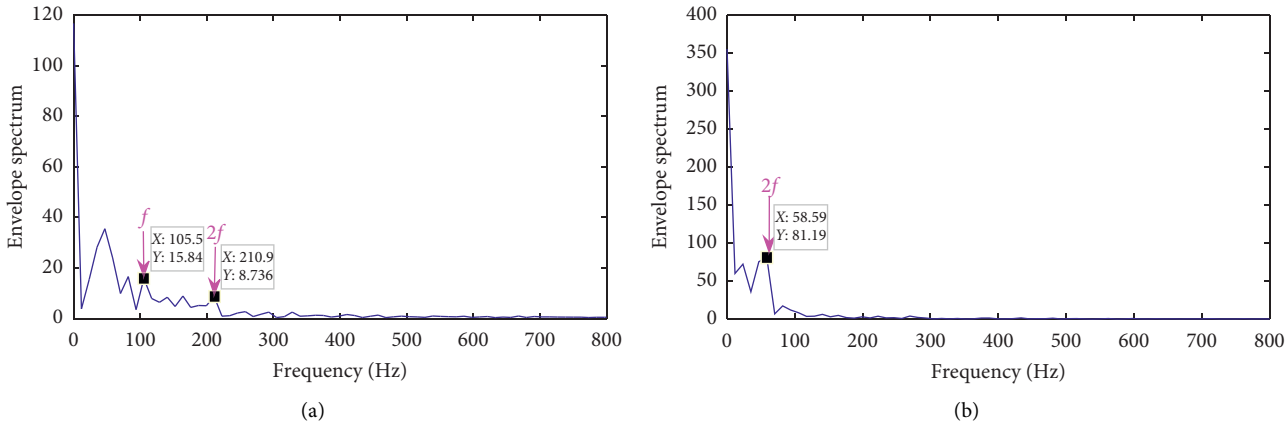


Figure 17: Envelope spectrum of effective mode extracted by (a) the maximum kurtosis and (b) the maximum correlation coefficient.

6. Conclusions

A parameter-adaptive VMD method based on GWOMI for fault diagnosis of rolling bearing is proposed in this work. It greatly enhances the parameter-adaptive ability of VMD. Firstly, the minimum mean mutual information is used as the optimization objective function of the GWOMI-VMD. After that, the maximum kurtosis is set as the effective mode index. Finally, envelope spectrum analysis is implemented on the effective mode to extract the fault feature. Through simulation and experimental analysis, the validity and feasibility of GWOMI-VMD are demonstrated. In addition, comparison results between GWOMI-VMD and the traditional VMD, PSO-VMD, and EMD are provided, from which the superiority of the GWOMI-VMD is further verified. Therefore, this paper has some certain values for incipient fault detection and diagnosis of rotating machinery.

Data Availability

The experimental data come from the Case Western Reserve University Bearing Data Center.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported in part by the Natural Science Foundation of Zhejiang Province, China, under Grant LQ18F010009. And this research was also supported by the Fundamental Research Funds for Zhejiang Provincial Universities and Research Institutes (no. 2019J00046).

References

[1] J. Jiao, J. H. Yue, D. Pei, and Z. Q. Hu, "Application of feature fusion using coaxial vibration signal for diagnosis of rolling element bearings," *Shock and Vibration*, vol. 2020, Article ID 8831723, 14 pages, 2020.

[2] C. Q. Shen, Y. M. Qi, J. Wang et al., "An automatic and robust features learning method for rotating machinery fault diagnosis based on contractive autoencoder," *Engineering Applications of Artificial Intelligence*, vol. 76, pp. 170–184, 2018.

[3] R. Gu, J. Chen, R. J. Hong et al., "Incipient fault diagnosis of rolling bearings based on adaptive variational mode decomposition and Teager energy operator," *Measurement*, vol. 149, Article ID 106941, 13 pages, 2020.

[4] X. Yan and M. Jia, "Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 122, pp. 56–86, 2019.

[5] Y. Pan, R. Hong, J. Chen, Z. Qin, and Y. Feng, "Incipient fault detection of wind turbine large-size slewing bearing based on circular domain," *Measurement*, vol. 137, pp. 130–142, 2019.

[6] H. Cao, F. Fan, K. Zhou, and Z. He, "Wheel-bearing fault diagnosis of trains using empirical wavelet transform," *Measurement*, vol. 82, pp. 439–449, 2016.

[7] X. Zhang, Q. Miao, Z. Liu, and Z. He, "An adaptive stochastic resonance method based on grey wolf optimizer algorithm and its application to machinery fault diagnosis," *ISA Transactions*, vol. 71, no. 2, pp. 206–214, 2017.

[8] X. Chen, Y. Yang, Z. Cui, and J. Shen, "Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy," *Energy*, vol. 174, pp. 1100–1109, 2019.

[9] H. Yu, H. Li, and Y. Li, "Vibration signal fusion using improved empirical wavelet transform and variance contribution rate for weak fault detection of hydraulic pumps," *ISA Transactions*, vol. 107, pp. 385–401, 2020.

[10] A. Kurmar, Y. Q. Zhou, C. P. Gandhi et al., "Bearing defect size assessment using wavelet transform based Deep Convolutional Neural Network (DCNN)," *Alexandria Engineering Journal*, vol. 29, no. 2, pp. 999–1012, 2020.

[11] J. Chen, Z. Li, J. Pan et al., "Wavelet transform based on inner product in fault diagnosis of rotating machinery: a review," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 1–35, 2016.

[12] Z. Shen, Q. A. Zheng, M. L. Wu 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.

[13] L. Lu, J. Yan, and C. W. De Silva, "Dominant feature selection for the fault diagnosis of rotary machines using modified

genetic algorithm and empirical mode decomposition," *Journal of Sound and Vibration*, vol. 344, pp. 464–483, 2015.

[14] Y. Xue, D. Y. Dou, and J. G. Yang, "Multi-fault diagnosis of rotating machinery based on deep convolution neural network and support vector machine," *Measurement*, vol. 156, Article ID 107571, 7 pages, 2020.

[15] X. Yuan, Q. Tan, X. Lei, Y. Yuan, and X. Wu, "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine," *Energy*, vol. 129, pp. 122–137, 2017.

[16] J. D. Zheng and H. Y. Pan, "Mean-optimized mode decomposition: an improved EMD approach for non-stationary signal processing," *ISA Transactions*, vol. 106, pp. 392–401, 2020.

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

[18] M. E. Torres, M. A. Colominas, and G. Schlotthauer, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proceedings of the IEEE International Conference on Acoustics IEEE, Prague Congress Centre (PCC)*, pp. 4144–4147, Prague, Czech Republic, 2011.

[19] J. Zheng, J. Cheng, and Y. Yang, "Partly ensemble empirical mode decomposition: an improved noise-assisted method for eliminating mode mixing," *Signal Processing*, vol. 96, pp. 362–374, 2014.

[20] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[21] M. Zhang, Z. Jiang, and K. Feng, "Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump," *Mechanical Systems and Signal Processing*, vol. 93, pp. 460–493, 2017.

[22] Z. Li, J. Chen, Y. Zi, and J. Pan, "Independence-oriented VMD to identify fault feature for wheel set bearing fault diagnosis of high speed locomotive," *Mechanical Systems and Signal Processing*, vol. 85, pp. 512–529, 2017.

[23] X. Jiang, C. Shen, J. Shi, and Z. Zhu, "Initial center frequency-guided VMD for fault diagnosis of rotating machines," *Journal of Sound and Vibration*, vol. 435, pp. 36–55, 2018.

[24] X. Zhang, Q. Miao, H. Zhang, and L. Wang, "A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery," *Mechanical Systems and Signal Processing*, vol. 108, pp. 58–72, 2018.

[25] J. Zhu, C. Wang, Z. Hu, F. Kong, and X. Liu, "Adaptive variational mode decomposition based on artificial fish swarm algorithm for fault diagnosis of rolling bearings," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 231, no. 4, pp. 635–654, 2017.

[26] X. L. Zhao, P. F. Wu, and X. X. Yin, "A quadratic penalty item optimal variational mode decomposition method based on single-objective salp swarm algorithm," *Mechanical Systems and Signal Processing*, vol. 138, Article ID 106567, 12 pages, 2020.

[27] T. Zhang, W. Chen, and M. Li, "AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier," *Biomedical Signal Processing and Control*, vol. 31, pp. 550–559, 2017.

[28] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[29] A. Khan, S. Rehman, M. Abbas, and A. Ahmad, "On the mutual information of relaying protocols," *Physical Communication*, vol. 30, pp. 33–42, 2018.

[30] L. R. Quitadamo, R. Lucia, F. Gozzo et al., "Kurtosis-based detection of intracranial high-frequency oscillations for the identification of the seizure onset zone," *International Journal of Neural Systems*, vol. 28, no. 7, Article ID 1850001, 12 pages, 2018.

[31] S. Wade and R. Robert, "Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study," *Mechanical Systems and Signal Processing*, vol. 64-65, pp. 100–131, 2015.

Hindawi
Shock and Vibration
Volume 2021, Article ID 6688420, 9 pages
https://doi.org/10.1155/2021/6688420

Research Article

Research on Fault Diagnosis Method of Electro-Hydrostatic Actuator

Lei Zhufeng,[1] **Qin Lvjun**,[1] **Wu Xiaodong**,[2] **Jin Wen**,[1] **and Wang Caixia**[1]

[1]*National Joint Engineering Research Center for Special Pump Technology, Xi'an Aeronautical University, 259 West Second Ring Road, Xi'an, Shaanxi 710077, China*
[2]*The 49th Research Institute of China Electronics Technology Group Corporation, 969 Longsheng Road, Harbin, Heilongjiang 150028, China*

Correspondence should be addressed to Lei Zhufeng; leizhufeng@qq.com

Received 3 December 2020; Revised 10 January 2021; Accepted 18 January 2021; Published 27 January 2021

Academic Editor: Ke Feng

Electrohydrostatic actuator is a type of actuator that uses hydraulic energy as the energy transmission carrier, which has the advantages of small size and high power. Since it is commonly used in harsh conditions such as strong vibration, high pressure, and heavy loads, condition monitoring and fault diagnosis of its hydraulic system are particularly important. This paper proposed a novel fault feature extraction method and applied to fault diagnosis of electrohydrostatic actuator. Firstly, the pressure signal of the hydraulic system is decomposed at multiple scales to obtain the center frequency of its maximum energy intrinsic mode component, and the feature data set is constructed based on the statistical features of the time domain. Then, a fault identification model of hydraulic system based on support vector machine is established. Finally, the fault classification and identification results of the hydraulic system are outputted. After a variety of method comparisons, the method proposed in this paper has a fault time ratio accuracy of 96.7%, which provides a basis and a new way for the fault diagnosis of the hydraulic system.

1. Introduction

The hydraulic transmission has high power density, compact structure, and good dynamic performance. It is widely used in construction machinery, marine vessels, aerospace, and robotics. Long pipeline, poor reliability, large volume, and heavy weight limit the application of traditional centralized electrohydraulic control systems. Distributed electrohydrostatic actuator is small in size, light in weight, and large in output power, which is an important development direction. However, the electrohydrostatic actuator is an important subsystem related to the reliability of the aircraft. How to increase the reliability and service capability of the system while reducing power consumption is a huge challenge it faces.

The hydraulic system is the core component of the electrohydrostatic actuator. The failure of the hydraulic system seriously affects the reliability of the electrohydrostatic actuator and has a great impact on the output power quality. The hydraulic system has problems such as fluid leakage, hydraulic valve failure, and insufficient cooling, which are also important factors affecting system stability. Timely and accurate fault diagnosis and repair of the hydraulic system is a prerequisite for the stable output power of the electrohydrostatic actuator.

Many scholars have conducted a lot of research on the diagnosis of hydraulic system faults. Linaric D et al. [1] applied the neural network model based on the mathematical model of the entire electrohydraulic servo system to perform model-based fault detection and isolation, verified the algorithm by simulation, and achieved significant results. Liu and Jiang [2] improved the application research of fuzzy neural network model in defect detection and classification. Tang et al. [3] through information fusion extracted fault features, through PSO-SVM performed fault classification, and proposed a combined information fusion and PSO-SVM multifault diagnosis method for piston pumps, which can effectively identify multiple faults of piston pumps. Gao

et al. [4] performs the adaptive matching of the noise threshold through median absolute deviation, and a noise suppression method based on Walsh transform is proposed. The method has been verified by numerical simulation and experimental verification of normal pumps and faulty pumps. The method provides a new idea for pump fault diagnosis. Deng et al. [5] in the process of motor bearing fault diagnosis, based on empirical model decomposition and fuzzy information entropy, improved particle swarm optimization algorithm and least squares support vector machine method; an early fault diagnosis method for rotating machinery is proposed and verified by experiments. The multiscale analysis method is an effective signal processing method. The commonly used multiscale processing methods include singular value decomposition (SVD), empirical mode decomposition (EMD), and wavelet transform (WT). In recent years, many scholars have done much research on multiscale analysis method, and they have been applied to the signal processing. Zhang et al. [6] proposed a rotating machinery fault diagnosis method based on Fourier-transform multifilter decomposition. This method uses Fourier-transform multifilter decomposition to decompose the original signal and obtain multiscale frequency domain information. After the feature selection and feature ranking of subsequent subsignals, furthermore, a fault diagnosis method for rotating machinery with higher accuracy and fewer features is obtained. Liang et al. [7] improved the local average decomposition (LMD) by bidirectional long and short-term memory model (Bi-LSTM), which realizes the adaptive expansion of the LMD endpoints, eliminates the LMD endpoint effects, and applies it to the feature extraction of rolling bearings under different loads. In 1998, Huang et al. [8] proposed EMD. EMD is an adaptive decomposition method, and it has been widely used in biomedicine [9–11], speech recognition [12, 13], system modeling [14–16], and process control [17–19]. EMD can decompose the original signal into different frequency bands and enrich the feature information of the original signal by extracting feature information in different frequency bands. EMD has been widely used in production practice. Hu et al. [20] decomposed the children's EEG into various brain wave components by EMD, and 11 different physical quantities are extracted as features in the intrinsic mode function (IMF). Finally, the random forest is used for activity recognition. The experiment verifies the effect of this method on the feasibility of activity identification. Vargas et al. [21] use EMD to decompose the ECG signal into several IMFs and process these IMFs based on the Viterbi algorithm and discrete wavelet transform to realize the noise removal of the ECG signal. Wavelet technology [22] is also an effective multiscale signal processing technology. Therefore, scholars have carried out a lot of research on wavelet technology. Jimenez et al. [23] filter and reduce the noise of wind turbine signals by wavelet transform and perform feature extraction and fault classification. Through the comparison of 6 schemes, the novelty of the method is verified. Based on the wavelet packet transform, a concept of the main frequency band of blasting is defined, which provides an important basis for accurately describing the frequency characteristics of the blasting process [24]. In 2014, Dragomiretskiy et al. [25] proposed the variational mode decomposition method. VMD can separate the components effectively, which is a completely nonrecursive variational mode decomposition model. It finds the center frequency of each IMF by iteratively searching for the optimal solution of the variational model.

Kumar et al. [26] proposed a VMD based on genetic algorithm (GA) and kernel-based mutual information (KEMI) fitness function, which can effectively extract the weak features of multiple defects in bearings. The advantages of VMD in local damage of beam structures are discussed by Mousavi et al. [27], and the peak synchronization between the instantaneous frequency and instantaneous amplitude of the first IMF is considered, which verifies the superiority of VMD in local damage detection.

Although a large number of scholars have conducted in-depth research on the fault diagnosis of hydraulic systems, with the continuous development of human society, hydraulic systems continue to be integrated into new use environments, various faults occur frequently, and even multiple faults appear at the same time. With the continuous advancement of computer science and information science, a data-driven hydraulic system fault diagnosis method using hydraulic system monitoring data has emerged. Finding fault characteristics from hydraulic system monitoring signals and accurately identifying fault types is the key study of this article.

This paper proposed a feature extraction method based on multiscale signal processing and applied it to the fault diagnosis of the hydraulic system to improve the fault identification accuracy and efficiency. Firstly, a new type of variational modal decomposition method is given. Secondly, we introduced the hydraulic system test platform on which the data collection of this article is based, introduced the composition and hydraulic principle of the hydraulic system test platform, and conducted data acquisition of hydraulic system pressure signal under four failure modes, and performed multiscale decomposition of the original signal; the maximum energy intrinsic mode component central frequency of pressure signal is obtained, the feature data set is constructed with the statistical features based on the time domain, and the pattern recognition is performed. Finally, the influence of multiple feature data set construction methods on the fault recognition accuracy of hydraulic system is discussed.

2. Establishment of Fault Diagnosis Model for Hydraulic Power Generation System

2.1. Variational Modal Decomposition. VMD is a completely nonrecursive variational modal decomposition model. In the process of iteratively searching for the optimal solution of the variational model, the center frequency and bandwidth of each decomposition and separation are determined, so each component of the signal can be effectively separated and the adaptive frequency domain division of the signal can be realized. Since the signal is decomposited into nonrecursive and variational modal components by VMD, it

shows better noise and sampling rate robustness than the recursive screening modal method of EMD.

VMD is based on the mixing-frequency variational problem-solving process, Hilbert transform, and classical Wiener filtering. In the process of iterative search for the optimal solution of the variational model, the center frequency and bandwidth of each decomposition and separation can be determined, which can effectively separate the components of the signal and realize the adaptive frequency domain division of the signal, so the signal can be decomposed into the sparsity component, which is a new signal decomposition estimation method. The variational modal decomposition algorithm redefines an AM-FM as an intrinsic modal function, and its expression is

$$u_k(t) = A_k(t)\cos\left(\phi_k(t)\right), \tag{1}$$

where $A_k(t)$ is amplitude of $u_k(t)$, $A_k(t) \geq 0$. $\omega_k(t) = \varphi_k'(t)$, $\omega_k(t)$ is the frequency of $u_k(t)$. The phase $\varphi_k(t)$ is a nondecreasing function. In the interval $[t-\delta, t+\delta]$, $u_k(t)$ is a harmonic signal with amplitude $A_k(t)$, where $\delta = 2\pi/\phi_k'(t)$.

VMD is different from the iterative screening process used in the empirical mode decomposition method when obtaining the IMF. It is a nonrecursive method based on the principle of the variational model. The signal is decomposed into sparse components; at the same time, the center frequency and bandwidth of each IMF are determined in the process of iteratively solving the optimal solution of the constrained variational model.

The variational constraint model is shown as follows:

$$\begin{aligned} &\min_{\{u_k\},\{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times u_k(t) \right] e^{-j\omega k t} \right\|_2^2 \right\}, \\ &s.t. \sum_k u_k = f, \end{aligned} \tag{2}$$

where $\{u_k\} := \{u_1, u_2, \ldots, u_K\}$ represents K IMF components, $\{\omega_k\} := \{\omega_1, \omega_2, \ldots, \omega_K\}$ represents the center frequency of each IMF, and $\sum_k := \sum_{k=1}^{K}$ represents the sum of all modes.

After the original signal is decomposed by VMD, the center frequency of the maximum energy of the IMF can be obtained according to the following formula:

$$\mathrm{CIMF} = \max_{1 \leq i \leq n-1} \left[E\left(\mathrm{IMF}_i, \mathrm{IMF}_{i+1}\right) \right], \tag{3}$$

where CIMF is the maximum energy of the IMF, IMF_i is the i^{th} IMF, and E (IMF_i) is energy of the i^{th} IMF.

The maximum energy of the IMF is the component with the largest energy in the process of multiscale analysis of the original signal. It occupies a large proportion of the information implicit in the original signal and is the main component in the original signal. It can be seen in Figure 1 [28].

As shown in Figure 1, IMF8 is maximum energy of the IMF.

2.2. Support Vector Machine. SVM is an effective tool for solving classification and regression problems. The linear classifier with the largest interval defined in the feature space is the basic model of SVM [29]. Nonlinear SVM is an improvement and general case of linear separable SVM. Linearly separable SVM is a linear classifier that uses hard interval maximization to learn when the training data are linearly separable. Linear SVM is a linear classifier that uses soft interval to maximize learning when the training data are approximately linearly separable. Nonlinear SVM is a nonlinear classifier that uses nuclear techniques to learn when the training data are linearly inseparable. Therefore, the SVM introduced with the kernel technique has become a substantial nonlinear classifier.

Suppose the training set of a feature space is $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$; among them, $x_i \in \chi = \Re^n$, $y_i \in y = \{+1, -1\}$, $i = 1, 2, \ldots, N$, x_i is a feature vector and y_i is a label of x_i.

The classification hyperplane is

$$h(x) = \omega \cdot x + b, \tag{4}$$

where x is the input vector, ω is the normal vector, and b is the intercept.

The classification decision function is

$$\begin{aligned} &\mathrm{Sign}\ (h((x)), \\ &\begin{cases} h(x) > 0, y_i = 1, \\ h(x) < 0, y_i = -1. \end{cases} \end{aligned} \tag{5}$$

SVM is to find an optimal hyperplane in the feature space so that the data set is divided into different classes, and the interval between the closest points is maximized. There are infinitely many solutions for its classification hyperplane. Using interval maximization to solve the classification hyperplane, the solution at this time is unique.

A cost factor ξ_i is paid for each slack variable ξ_i, and then formula 15 is transformed into

$$\frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N} \xi_i, \tag{6}$$

where $C > 0$ is the penalty parameter.

The function of C is to reconcile the coefficients of formula 6. When C increases, the penalty for misclassification increases, and on the contrary, the penalty for misclassification decreases.

Formula 6 is to make $(1/2)\|\omega\|^2$ as small as possible, that is, the interval is maximized while the points of misclassification are minimized.

Construct the Lagrangian function as follows:

$$L(\omega, b, \alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^{N} \alpha_i \left[y_i(\omega \cdot x_i + b) - 1 + \xi_i \right] + C\sum_{i=1}^{N} \alpha_i, \tag{7}$$

where $\alpha_i \geq 0$ is Lagrange multiplier, $i = 1, 2, \ldots, N$, N is the total number of samples in the sample set.

According to the KarushKuhnTucker condition, the partial differentiation of ω, b, and ξ is obtained and we set

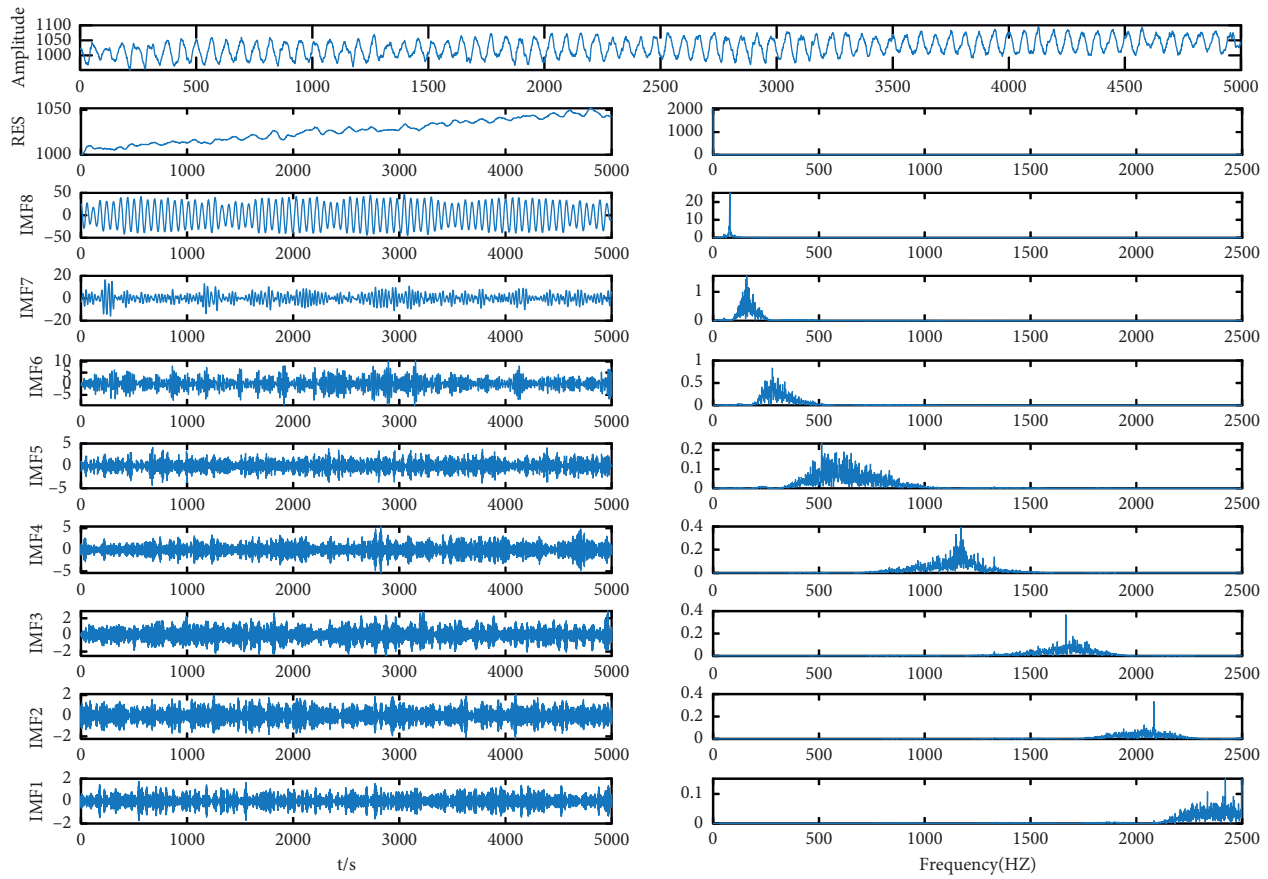


Figure 1: VMD decomposition example.

them equal to 0, and thus the following formula can be obtained:

$$\omega = \sum_{i=1}^{N} y_i \alpha_i x_i, \tag{8}$$

$$-\sum_{i=1}^{N} y_i \alpha_i = 0, \tag{9}$$

where $\alpha_i (y_i (\omega \cdot x_i + b) - 1) = 0$, $i = 1, 2, \ldots, N$, $y_i (\omega \cdot x_i + b) - 1 \geq 0$, $i = 1, 2, \ldots, N$, $\alpha_i \geq 0$, $i = 1, 2, \ldots, N$.

Thus, the separating hyperplane can be written as

$$\sum_{i=1}^{N} \alpha_i y_i (x \cdot x_i) + b = 0. \tag{10}$$

The classification decision function can be written as

$$f(x) = \text{sign}\left[\sum_{i=1}^{N} \alpha_i y_i (x \cdot x_i) + b\right], \tag{11}$$

where $b = y_j - \sum_{i=1}^{N} \alpha_i y_i (x \cdot x_i)$, y_j is support vector.

For linear classification problems, linear support vector machines are an effective method. However, in actual data, most of the data are linear and inseparable. In order to transform nonlinear problems into linear problems and linear inseparable problems into linear separable problems, linear support vector machines are suitable for the linear inseparable data problems that are often encountered in practical problems.

Let input space χ be a subset or discrete set of Euclidean space $\mathfrak{R}_n$ and feature space (Hilbert space) $\mathscr{H}$, if there is a mapping from χ to $\mathscr{H}$:

$$\varphi(x): \chi \longrightarrow \mathscr{H}. \tag{12}$$

Such that for all $x_i, x_j \in \chi$, the function $k(x_i, x_j)$ satisfies the following condition:

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j). \tag{13}$$

Then, $k(x_i, x_j)$ is the kernel function.where $\varphi(x)$ is the mapping function and $\varphi(x_i) \cdot \varphi(x_j)$ is the inner product of $\varphi(x_i)$ and $\varphi(x_j)$.

At this point, the objective function becomes

$$W(\alpha) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{N} \alpha_i, \tag{14}$$

where $\alpha_i > 0$ is Lagrangian multiplier and N is total number of samples in the sample set.

The inner product in the classification decision function can be replaced by the kernel function, and then the classification decision function becomes

$$f(x) = \text{sign}\left[\sum_{i=1}^{N} a_i y_i K(x_i, x) + b\right], \quad (15)$$

$$K(x_i, x) = \exp\left(\frac{-\|x_i - x\|^2}{2g^2}\right). \quad (16)$$

The following formula is a Gaussian radial basis function classifier, and its classification decision function is

$$f(x) = \text{sign}\left[\sum_{i=1}^{N} a_i y_i \exp\left(\frac{-\|x_i - x\|^2}{2g^2}\right) + b\right], \quad (17)$$

where g is the kernel function parameter.

2.3. Integrated Algorithm Flow and Steps. As is shown in Figure 2, the pressure signal collected in the hydraulic system is first used for signal noise reduction, and the noise-reduced signal is extracted by two methods. One is the feature extraction based on time series signal statistical features, including peak-to-peak value, standard deviation, variance, mean, peak value, skewness, and kurtosis; the second is to perform VMD on the noise-reduced signal to obtain the center frequency of maximum energy intrinsic mode component. Taking the center frequency of the maximum energy intrinsic mode component as the eigenvalue, combine the statistical features to combine the feature data set. Then, the combined feature data set is divided into a training set and a test set. Through the training set, a fault recognition model which is based on the support vector machine-based hydraulic system is constructed, and the model parameters are optimized. The test set is used to verify the accuracy and efficiency of the optimized fault recognition model. Finally, the classification and recognition results are output.

3. Experimental Analysis

3.1. Hydraulic Test Rig. In order to verify the reliability and robustness of the method described in this article, we developed a hydraulic test rig, as shown in Figure 3; it consists of hydraulic pump station, filter temperature control system, test console, hydraulic component test platform, test data acquisition system, and electrical control system, etc. The test rig uses the pump as the drive and the motor as the load. The performance of the hydraulic components can be tested by replacing different components on the hydraulic test rig. The performance parameters of the test rig are shown in Table 1.

The principle of the hydraulic test rig is shown in Figure 4.

As shown in Figure 4, 1 is the hydraulic pump group, 2 is the cooling system, 3 is the flow meter, 4 is the one-way valve, 5 is the pressure sensor, 6 is the torque speed tester, 7 is the loading pump, 8 is the motor, 9 is the temperature sensor, 10 is the overflow valve, 11 is the reversing valve, and 12 is the accumulator. The pump 1 is driven by an electric motor, and the system pressure is adjusted through a one-way valve group. The state of different hydraulic components can be monitored through the pressure and temperature values of different positions.

3.2. Failure Category of Hydraulic Power Generation System. In the hydraulic system, different hydraulic components will have different types of failures under different working conditions, even the same hydraulic components will have different types of failures under different working conditions, and the solutions are also different for different types of failures.

The fluid characteristics in the hydraulic system are different from the mechanical vibration characteristics, so it is impossible to perform fault diagnosis of the hydraulic system by monitoring the vibration of the system components.

This paper simulates the fluid leakage, hydraulic valve failure, accumulator failure, and other failure forms of the hydraulic system in the unstable state and the stable state of the system as shown in Table 2. Among them, 1 indicates the best state and 0 indicates the failure state.

As shown in Figure 5, using a hydraulic test rig, the inlet pressure of the check valve 3 was collected while simulating these failures. The sampling frequency was 100 Hz and the sampling time was 8 seconds. The fault diagnosis model of hydraulic system is verified through four fault types.

3.3. Fault Identification Based on the Maximum Energy Center Frequency. Firstly, the collected pressure signal noise reduction is made, and the maximum energy intrinsic mode component of each signal is calculated. The center frequency of the maximum energy intrinsic mode component is selected as the feature, and the statistical feature of the original signal based on support vector machine is combined with for classification calculation. Statistical features include peak-to-peak value, standard deviation, variance, mean, peak value, skewness, and kurtosis. 60% of the feature set is defined as the training set, and 40% is defined as the test set. According to the method of document 28, the original signal in this paper is decomposed by VMD to obtain 9 IMFs. The parameter C in SVM is 13.368 and g is 0.01. The classification result is shown in Figure 6.

It can be seen from Figure 6 that the prediction accuracy of the hydraulic power generation system fault diagnosis model based on the maximum energy center frequency can reach 96.7%.

4. Discussion

In order to verify the feasibility and effectiveness of using the multiscale features of variational modal decomposition to construct the feature data set proposed in this paper, a variety of methods are used for comparative analysis, including only using statistical features for fault identification, only using the maximum energy intrinsic mode component center frequency for fault identification and classification model, and using Gaussian kernel function and Sigmoid kernel function for fault recognition. The recognition results of multiple methods are

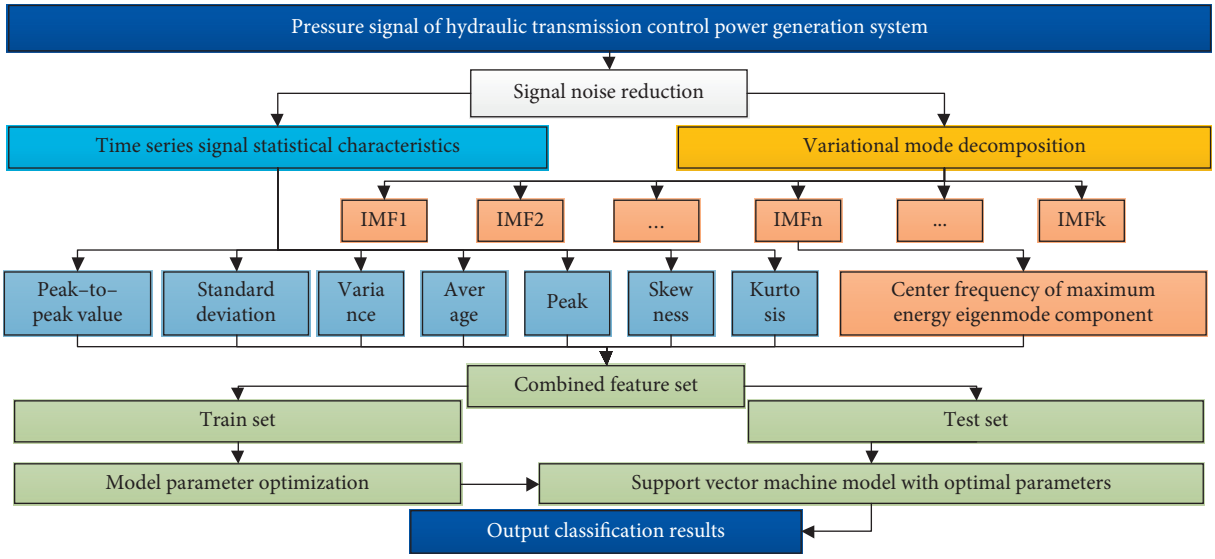


Figure 2: Integrated algorithm flow.

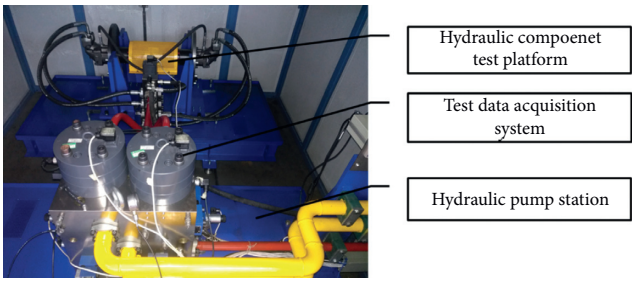


Figure 3: Hydraulic test rig.

Table 1: The performance parameters of the test rig.

PerformanceO	Operating temperature (°C)	Flow meter		Pressure sensor	Torque meter (N•m)
		Measurement accuracy	Linearity		
Target	50	±3%FS	±1.5%FS	±5%FS	
Performance	Operating pressure (bar)	Rated flow (L/min)	Total power (KW)	Noise (dB)	≤2000
Target	0–420	320	215	≤85 dB	

compared so that the best model for fault diagnosis of hydraulic power generation system is selected.

The comparison result is shown in Table 3, as shown in Figure 7.

It can be seen from Table 3 and Figure 7 that the accuracy of fault identification only using the statistical feature of the fault information feature data set is only 56.7%, and the accuracy of only using the central frequency of the maximum energy intrinsic mode component for fault identification reaches 93.3%. The accuracy of fault identification by integrating the feature data sets of the two feature combinations has been greatly improved, reaching 96.7%.

Since the Gaussian kernel function maps the feature data set to high dimensions, the indivisible data become separable and the Sigmoid kernel function makes the support vector machine model a multilayer perceptron neural network. Therefore, the classification performance of the Gaussian kernel function used in the classification model is obviously worse than the performance of the Sigmoid kernel function.

The center frequency of the maximum energy intrinsic mode component is a kind of multiscale spectral information, which is essentially different from the statistical information extracted based on time-domain signals. Using the center frequency of the maximum energy intrinsic mode component as a feature to construct a feature data set has a positive impact on the accuracy of electrohydrostatic actuator fault recognition.

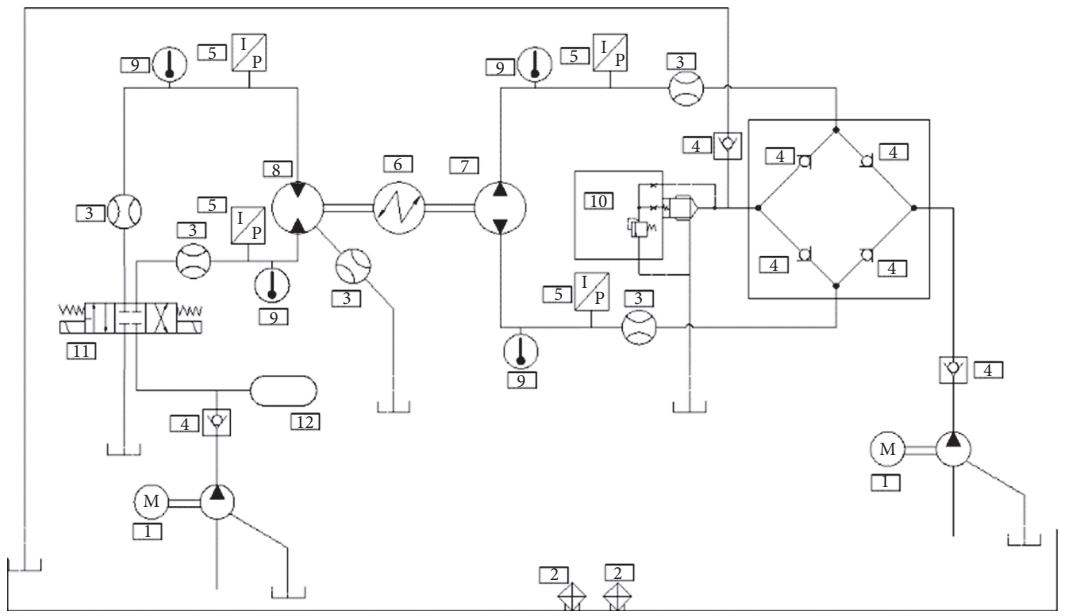


Figure 4: The principle of the hydraulic test rig.

Table 2: Simulated fault information.

Fault type	Valve status	Pump leak	Accumulator failure	System status
Valve status	0	1	1	1
Pump leak	1	0	1	1
Accumulator failure	1	1	0	1
System status	1	1	1	0

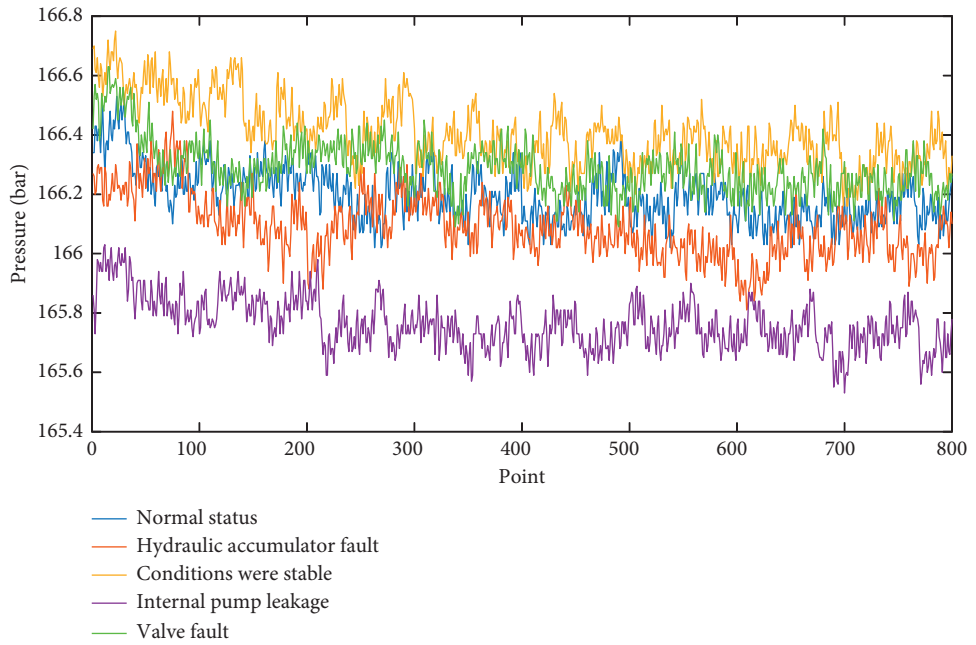


Figure 5: The pressure signal curve under five different fault conditions collected by the hydraulic system.

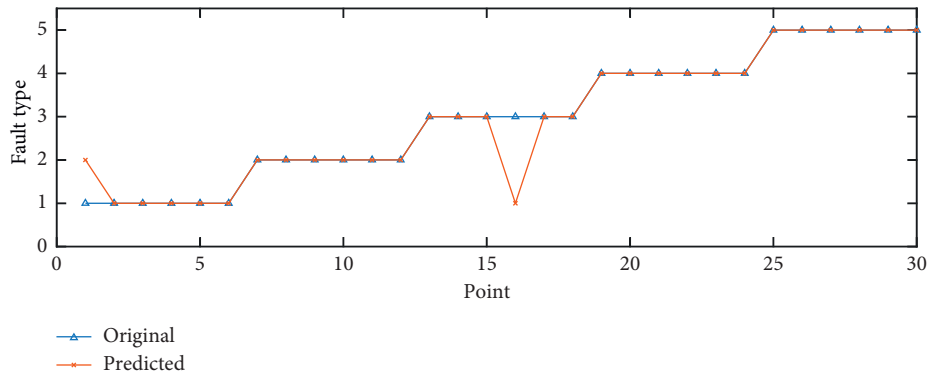


Figure 6: Result of fault identification.

Table 3: Various methods of fault identification results.

Method	Comprehensive feature data set		Center frequency feature data set		Statistical feature data set	
Kernel function type	Sigmoid	Gauss	Sigmoid	Gauss	Sigmoid (%)	Gauss (%)
Results	96.7	85.6	93.3	81.3	56.7	49.2

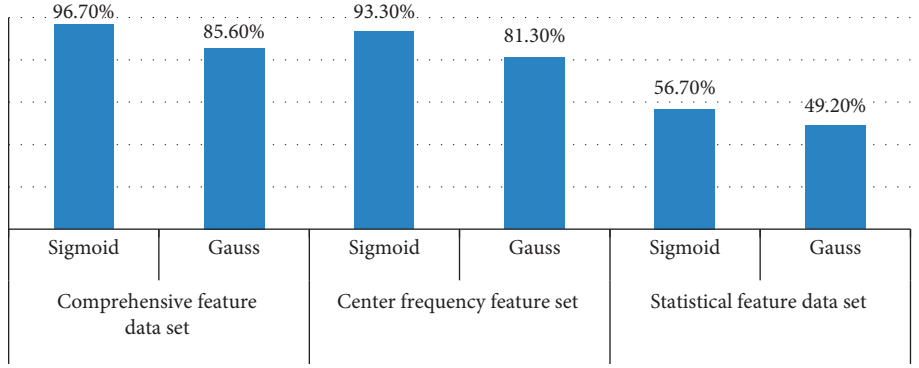


Figure 7: Fault identification results of various methods.

5. Conclusions

This paper proposes a multi-scale feature extraction method based on variational modal decomposition. The central frequency of the maximum energy intrinsic mode component and the statistical features based on the time-domain signal are combined to construct a feature data set and establish the fault diagnosis model of the hydraulic system, which greatly improves the accuracy of the model for fault identification of the electro-hydrostatic actuator. It can be seen that the feature data set constructed with the multiscale features obtained by the variational modal decomposition method described in this article can more clearly reflect the characteristics of the fault information, which provides a basis for the fault diagnosis of the hydraulic system and provides a new way for electro-hydrostatic actuator fault diagnosis.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Jin Wen was involved in the project administration; Lei Zhufeng wrote and prepared the original draft; Wu Xiaodong investigated the study; Qin Lvjun and Wang Caixia wrote, reviewed, and edited the manuscript.

Acknowledgments

Wei Hongbo and Zhang Yifei are acknowledged for their valuable technical support. This work was financially supported by the Science Foundation of Xi'an Aeronautical University (no. 2020KY0223) and the National Natural Science Foundation of China (no. 11902236).

References

[1] D. Linaric and V. Koroman, "Fault diagnosis of a hydraulic actuator using neural network," in *Proceedings of the IEEE*

International Conference on Industrial Technology, pp. 108–111, Maribor, SI, Slovenia, December 2003.

[2] C. Liu and H. Jiang, "A seventh-order model for dynamic response of an electro-hydraulic servo valve," *Chinese Journal of Aeronautics*, vol. 27, no. 6, pp. 1605–1611, 2014.

[3] H. Tang, Z. Wang, and Y. Wu, "A multi-fault diagnosis method for piston pump in construction machinery based on information fusion and PSO-SVM," *Journal of Vibro Engineering*, vol. 21, pp. 1904–1916, 2019.

[4] Q. Gao, H.-S. Tang, J.-W. Xiang, and Y. Zhong, "A multi-sensor fault detection strategy for axial piston pump using the Walsh transform method," *International Journal of Distributed Sensor Networks*, vol. 14, no. 4, Article ID 155014771877253, 2018.

[5] W. Deng, R. Yao, H. Zhao, X. Yang, and G. Li, "A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm," *Soft Computing*, vol. 23, no. 7, pp. 2445–2462, 2019.

[6] C. Zhang, L. Kong, Q. Xu, K. Zhou, and H. Pan, "Fault diagnosis of key components in the rotating machinery based on Fourier transform multi-filter decomposition and optimized LightGBM," *Measurement Science and Technology*, vol. 32, no. 1, Article ID 015004, 2021.

[7] J. Liang, L. Wang, J. Wu, Z. Liu, and G. Yu, "Elimination of end effects in LMD by Bi-LSTM regression network and applications for rolling element bearings characteristic extraction under different loading conditions," *Digital Signal Processing*, vol. 107, Article ID 102881, 2020.

[8] 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 of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[9] M. Li, X. Sun, and W. Chen, "Patient-specific seizure detection method using nonlinear mode decomposition for long-term EEG signals," *Medical & Biological Engineering & Computing*, vol. 58, no. 12, pp. 3075–3088, 2020.

[10] S. Yang, B. Li, Y. Zhang et al., "Selection of features for patient-independent detection of seizure events using scalp EEG signals," *Computers in Biology and Medicine*, vol. 119, Article ID 103671, 2020.

[11] S. Xie and S. Krishnan, "Feature extraction of epileptic EEG using wavelet power spectra and functional PCA," in *Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2551–2554, Berlin, Germany, July 2019.

[12] R. Hemavathi and R. Kumaraswamy, "A study on unsupervised monaural reverberant speech separation," *International Journal of Speech Technology*, vol. 23, no. 2, pp. 451–457, 2020.

[13] R. Tejus, Y. Nishant, S. Moin et al., "Role of source separation using combined RPCA and block thresholding for effective speaker identification in multi source environment," in *Proceedings of the 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, pp. 1253–1256, Bangalore, India, January 2017.

[14] Y. Xu and G. Yang, "A short-term wind speed forecasting hybrid model based on empirical mode decomposition and multiple kernel learning," *Complexity*, vol. 2020, Article ID 8811407, 13 pages, 2020.

[15] S. Li, X. Liu, and A. Lin, "Fractional frequency hybrid model based on EEMD for financial time series forecasting," *Communications in Nonlinear Science and Numerical Simulation*, vol. 89, Article ID 105281, 2020.

[16] W. Ding and F. Meng, "Point and interval forecasting for wind speed based on linear component extraction," *Applied Soft Computing*, vol. 93, Article ID 106350, 2020.

[17] K. H. Choi, "Low-frequency pressure fluctuations in an external-loop airlift reactor," *Korean Chemical Engineering Research*, vol. 58, pp. 665–674, 2020.

[18] Z. Lei and W. Su, "Mold level predict of continuous casting using hybrid EMD-SVR-GA algorithm," *Processes*, vol. 7, no. 3, p. 177, 2019.

[19] Y. Zhang, X. Yang, F. Zhu et al., "Characteristics of non-linear dynamics and energy transfer in a vibration gas-solid fluidized bed by using Hilbert-Huang transform," *Powder Technology*, vol. 344, pp. 970–980, 2019.

[20] L. Hu, K. Zhao, X. Zhou, B. W.-K. Ling, and G. Liao, "Empirical mode decomposition based multi-modal activity recognition," *Sensors*, vol. 20, no. 21, p. 6055, 2020.

[21] R. N. Vargas and A. C. P. Veiga, "Empirical mode decomposition, viterbi and wavelets applied to electrocardiogram noise removal," *Circuits System Signal Process*, vol. 28, 2020.

[22] X. Chiementin, B. Kilundu, L. Rasolofondraibe, S. Crequy, and B. Pottier, "Performance of wavelet denoising in vibration analysis: highlighting," *Journal of Vibration and Control*, vol. 18, no. 6, pp. 850–858, 2012.

[23] A. A. Jimenez, L. Zhang, C. Q. G. Munoz et al., "Maintenance management based on Machine Learning and nonlinear features in wind turbines," *Renewable Energy*, vol. 146, pp. 316–328, 2020.

[24] G. Chen, Q.-Y. Li, D.-Q. Li, Z.-Y. Wu, and Y. Liu, "Main frequency band of blast vibration signal based on wavelet packet transform," *Applied Mathematical Modelling*, vol. 74, pp. 569–585, 2019.

[25] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[26] A. Kumar, Y. Zhou, and J. Xiang, "Optimization of VMD using kernel-based mutual information for the extraction of weak features to detect bearing defects," *Measurement*, vol. 168, Article ID 108402, 2021.

[27] M. Mousavi, D. Holloway, J. C. Olivier, and A. H. Gandomi, "Beam damage detection using synchronisation of peaks in instantaneous frequency and amplitude of vibration data," *Measurement*, vol. 168, Article ID 108297, 2021.

[28] W. Su and Z. Lei, "Mold-level prediction based on long short-term memory model and multi-mode decomposition with mutual information entropy," *Advances in Mechanical Engineering*, vol. 11, no. 12, Article ID 168781401989443, 2019.

[29] X. Zhang and J. 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.