IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 19, 2018, accepted June 11, 2018, date of publication July 6, 2018, date of current version September 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2851374

An Improved VMD With Empirical Mode Decomposition and Its Application in Incipient Fault Detection of Rolling Bearing

FAN JIANG[1,2], ZHENCAI ZHU[1], AND WEI LI[1]

[1]Jiangsu Key Laboratory of Mine Mechanical and Electrical Equipment, School of Mechatronic Engineering, China University of Mining and Technology, Xuzhou 221116, China

[2]Postdoctoral Research Station of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Fan Jiang (jiangfan25709@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51605478, Grant 51475455, and Grant U1510205, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20160251, in part by the China Postdoctoral Science Foundation under Grant 2016M590513, in part by the Fundamental Research Funds for the Central Universities under Grant 2017QNA17, and in part by the Priority Academic Program Development of the Jiangsu Higher Education Institutions.

ABSTRACT Transient impulse analysis is an effective way to detect the bearing fault at its early stage. However, it is hard to precisely extract these so-called transient impulses because these collected vibration signals usually are non-stationary, nonlinear, and drowned by heavy background noise. Variational mode decomposition (VMD) can play the role as an adaptive signal processing tool to reveal the weak transient impulses from complex vibration signals. However, its reasonable mode number is difficult to pre-set and this would make the loss of useful transient impulses. To solve this issue, an improved VMD strategy is presented in this paper . For this method, it can not only utilize the advantages of traditional VMD and empirical mode decomposition (EMD) but also adaptively select sensitive intrinsic mode function (IMF) components for fault component analysis by proposed indexed values. EMD is first used to process the collected vibration signal into a series of IMFs, and the so-called useful IMFs are then evaluated by a sensitive IMF evaluation index which is based on the conjoint analysis of relatedness and kurtosis. Afterward, VMD is further improved to effectively decompose the denoised signal reconstructed from these selected useful IMFs of traditional EMD. Finally, the improved VMD is used for incipient fault diagnosis by a defined transient impulse monitoring index and Hilbert envelope analysis. Experiments are performed to demonstrate the effectiveness of the proposed method. The experimental results confirm that the proposed method can accurately extract the features of an incipient fault of a bearing.

INDEX TERMS Improved variational mode decomposition, incipient fault detection, empirical mode decomposition, rolling bearings.

I. INTRODUCTION

Rolling bearings are one of the widely used elements in rotating machinery and their failure is one of the most common causes of machine breakdown and accidents [1], [2]. Consequently, it is of great significance and increasingly attract attention in bearing fault diagnosis to increase the reliability of requirements and decrease loss owing to bearing faults. Vibration signal-based feature extraction has been verified to be an effective way to solve the problem of fault diagnosis of rotating machinery because the vibration signals collected from sensors include abundant fault-related information [3]– [5]. Our work employs vibration signals to diagnose faults of rotating machinery. To precisely extract the so-called features from complex vibration signals, the first step is to analyze the vibration signals by adopting a suitable signal processing method. In the past, many traditional techniques have been put forward to perform signal analysis in the time domain, frequency domain, and time-frequency domain [6]. Signal analysis in the time domain is the simplest method used for mechanical fault diagnosis. Frequency domain methods employ Fourier transform-based signal processing techniques, such as fast Fourier transform (FFT) [7] and interpolated discrete Fourier transform (IpDFT) [8]. However, the time domain and frequency domain methods used separately can only extract a portion of the useful information from the vibration signals and information in each of the other domains is lost.

As time-frequency methods can perform signal analysis in both the time and frequency domains simultaneously, many techniques have been introduced for processing the collected signals for condition monitoring, fault diagnosis, and mechanical life prediction, such as short-time Fourier transform (STFT) [9], Wigner-Ville distribution (WVD) [10], wavelet transform (WT) based methods [11], [12], empirical mode decomposition (EMD) [13], ensemble empirical mode decomposition (EEMD) [14], and variational mode decomposition (VMD) [15]. Among these, the last three methods employ adaptive time-frequency analysis algorithms. In [16], a suitable merit index was introduced to improve traditional EMD and HT and this method can automatically select sensitive IMFs for fault diagnosis. Du and Yang [17] substituted the conventional envelope mean with the average mean based on EMD for the diagnosis of bearing faults. In [18], a new fault diagnosis method for low-speed rolling bearings was presented using the parameter estimate of alpha-stable distribution (ASD) and an EMD-based vibration signal processing approach for filtering the trend and noise components. A fault detection method for roller-bearing systems was constructed using a wavelet denoising scheme, and proper orthogonal values (POV), of an intrinsic mode function (IMF) covariance matrix [19]. However, traditional EMD suffers from the mode mixing problem when analyzing vibration signals. To solve this problem, a new improved version, named EEMD, was proposed by Wu and Huang [14] and it also has been employed for fault diagnosis. In [20], EEMD and Hilbert marginal spectrum analysis were combined for multifault diagnosis of axle bearings based on an IMF confidence index for adaptive self-selection of the useful IMFs. In [21], EEMD and multiscale fuzzy entropy were introduced to extract features for the diagnosis of motor bearing faults. An effective fault diagnosis method was presented using improved EEMD and Hilbert square demodulation (HSD) [22]. In [23], a EEMD-based multiscale independent component analysis (ICA) method was proposed for slewing bearing fault detection and diagnosis. Lei and Zuo [24] presented an improved Hilbert-Huang transform (HHT) based on EEDM and sensitive IMFs for the fault diagnosis of rotating machinery. However, the decomposition results of EEMD do not meet the definition of a strictly IMF [14].

Recently, an adaptive signal decomposition algorithm, called variational mode decomposition (VMD), was proposed by Dragomiretskiy and Zosso [15]. VMD can adaptively find the frequency center and bandwidth of each component by iteratively searching the optimal solution of variational modes. Over the past few years, many scholars used it to analyze complex signals. Tang *et al.* [25] used VMD to design an underdetermined blind source separation method and it was successfully introduced for the compound fault diagnosis of roller bearings. In [26], a novel vibration signal processing method based on VMD and the Teager energy operator was presented for the fault diagnosis of wind turbines. Zhang *et al.* [27] constructed a fault diagnosis method for multistage centrifugal pumps by using VMD and the comparison results of VMD and EMD showed that the former can accurately extract the principal mode for fault feature extraction. Lv *et al.* [28] combined VMD and a multikernel support vector machine (MKSVM) to diagnose mechanical faults and the immune genetic algorithm (IGA) was used to optimize the MKSVM. These methods testify that VMD is an effective signal processing method for complex signals.

Using appropriate models or algorithms can accurately detect system abnormality and component fault to identify the root causes of these failures in time [29]. In this work, a novel method based on EMD and VMD, called improved VMD, is presented to precisely extract the transient impulse information of bearing faults from the nonlinear and nonstationary vibration signals. First, EMD is used to process the collected vibration signal into several IMFs. The number of so-called useful IMFs are then evaluated for denoised signal by the conjoint analysis of relatedness and kurtosis. Second, VMD is further improved to effectively decompose the denoised signal according to its vibration characteristics. Finally, the improved VMD is used for incipient fault diagnosis with a defined transient impulse monitoring index (TIME) and Hilbert envelope analysis, and experiments are then performed to demonstrate the effectiveness of the proposed method.

The remainder of this paper is structured as follows. Section II introduces theories, and includes the principles of EMD and VMD, and simulation analysis. Incipient fault diagnosis with IVMD is provided in Section III that covers four subsections: denoising with EMD based on sensitive IMF selection, the principle of IVMD, and transient impulse monitoring components. In Section IV, an explanation of experiments performed on a multifunction mechanical fault simulator. Conclusions are made in Section V.

II. THEORIES

A. EMD

The self-adaptive time-frequency analysis method, named EMD, was developed by Huang *et al.* [13] with three assumptions: (1) the target signal must have at least two extrema (one maximum and one minimum); (2) the characteristic time scale is defined by the time lapse between the extrema, and (3) if the data were totally devoid of extrema but contained only inflection points, then they can be differentiated one or more times to reveal the extrema. The process of EMD is shown as follows [13], [16]– [19].

(1) All local extrema points are found, and their upper and lower envelopes are obtained by cubic spline line.

(2) The mean of the upper and lower envelopes is calculated and marked as m_1, and then the difference between original signal $x(t)$ and m_1 defined h_1 as, i.e.,

$$h_1 = x(t) - m_1. \tag{1}$$

Determine whether h_1 is an IMF. if yes, the result of (1) is defined as the first IMF; otherwise, h_1 is regarded as the original signal $x(t)$, and above steps are repeated until h_{1k} becomes an real IMF. After this, the first IMF is decomposed by

$$c_1 = h_{1k}. \tag{2}$$

(3) Separate the first IMF c_1 from $x(t)$ by

$$r_1 = x(t) - c_1, \tag{3}$$

where r_1 is defined as the residue signal.

Repeat the above steps K times until the stop condition takes place. Then, K IMFs will be obtained and they satisfy

$$\begin{cases} r_1 - c_2 = r_2; \\ \vdots \\ r_{K-1} - c_K = r_K. \end{cases} \tag{4}$$

Finally, the original signal $x(t)$ can be decomposed into

$$x(t) = \sum_{k=1}^{K} c_k(t) + r_K(t). \tag{5}$$

B. VMD

VMD is a new self-adaptive and quasi-orthogonal signal processing method based on Wiener filtering, one-dimensional Hilbert transform, and heterodyne demodulation. Unlike EMD and EEMD, VMD can decompose a complex multicomponent signal into a series of sub-signals, which are mostly compact around a center pulsation, with a limited frequency bandwidth. To evaluate the bandwidth of a mode, the following constrained variational problem should be solved.

$$\begin{aligned} &\min_{\{u_k\},\{\omega_k\}} \{\sum_k \|\alpha_t[(\delta(t) + \frac{j}{\pi t}) * u_k(t)] \exp^{-j\omega_k t} \|_2^2\}, \\ &s.t. \sum_k u_k = f \end{aligned} \tag{6}$$

where t is the time script and δ is the Dirac distribution; u_k and ω_k are shorthand notations for the set of all modes and their center frequencies, respectively.

To render the problem unconstrained, a quadratic penalty term and Lagrangian multipliers are employed in [15] and a new solution expression can be obtained as follows:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = \eta \sum_k \{\sum_k \|\alpha_t[(\delta(t) + \frac{j}{\pi t}) * u_k(t)] \\ \times \exp^{-j\omega_k t} \|_2^2\} + \|f(t) - \sum_k u_k(t)\|_2^2 \\ + \langle \lambda(t), f(t) - \sum_k u_k(t)\rangle \end{aligned} \tag{7}$$

where η is the data-fidelity constraint parameter and λ is the Lagrangian multiplier.

Here, a sequence of iterative sub optimizations called alternate direction method of multipliers (ADMM), as can be seen in [15], is introduced to solve the above formula, and hence, the solution to the original minimization problem of (6) is now found as the saddle point of (7). Then, the modes u_k and their corresponding center frequency ω_k can be updated as

$$u_k^{n+1} \leftarrow \arg\min_{u_k} L(u_{i<k}^{n+1}, u_{i\geqslant k}^{n}, \omega_i^n, \lambda^n) \tag{8}$$

and

$$\omega_k^{n+1} \leftarrow \arg\min_{\omega_k} L(u_i^{n+1}, \omega_{i<k}^{n+1}, \omega_{i\geqslant k}^{n}, \lambda^n). \tag{9}$$

Through continuous iteration updates, all the sub-signals, called IMF modes, can be decomposed from the solution and are described as follows:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\eta(\omega - \omega_k^n)^2}. \tag{10}$$

C. SIMULATION ANALYSIS

Before constructing IVMD to decompose vibration signals, the performances of EMD, EEMD, and VMD are studied by simulation analysis and their results used to guide the parameter setting of VMD. The simulation signals are expressed as

$$\begin{cases} y_1 = 2sin(10\pi t) \\ y_2 = 1.5cos(80\pi t) \\ y_3 = \sum_i sin(200\pi t) \exp^{(-T_1(t - i*T_2)^2)} \\ y_{com} = y_1 + y_2 + y_3, \end{cases} \tag{11}$$

where $T_1 = 1000$, $T_2 = 0.25$, and $i = 1, 2, 3$.

Equation (11) shows that the simulation signal y_{com} is made up of oscillation signal y_3, and low-frequency signals y_1 and y_2. Usually, an actual sensor signal is accompanied by noise. Therefore, we defined $y_{com,e(t)} = y_{com} + e(t)$ and $e(t) = 0.2 * std(y_{com}) * rand(L)$, where $e(t)$ is the noise and L is the length of the simulation signal. The waves of these simulation signals are plotted in Fig. 1. This figure shows that identifying these components of the composite signal by directly observing it is difficult. Fig. 2 shows the decomposition results of simulation signal $y_{com,e(t)}$ with traditional

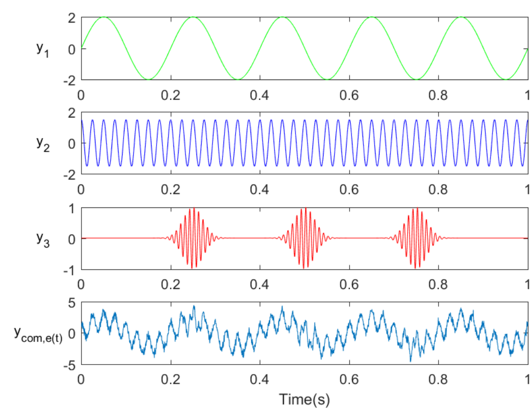


FIGURE 1. The waves of simulation signals.

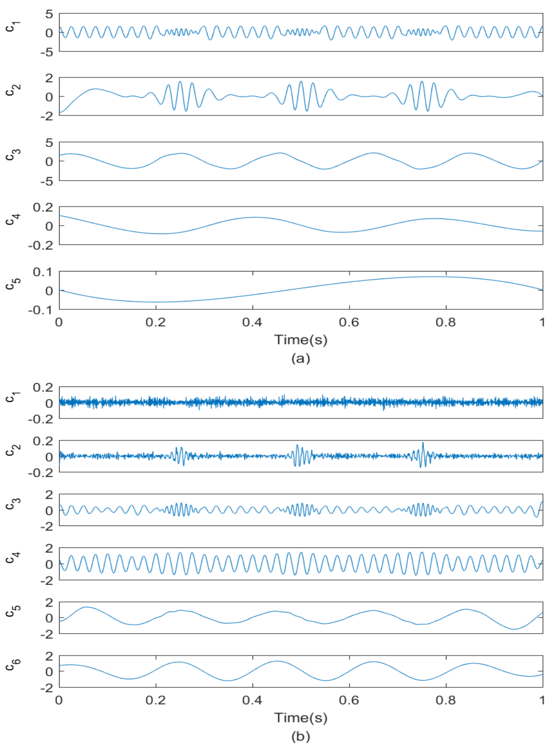


FIGURE 2. The decomposition results of simulation signal $y_{com,e(t)}$ by using: (a) EMD and (b) EEMD with the default parameters defined in [23].

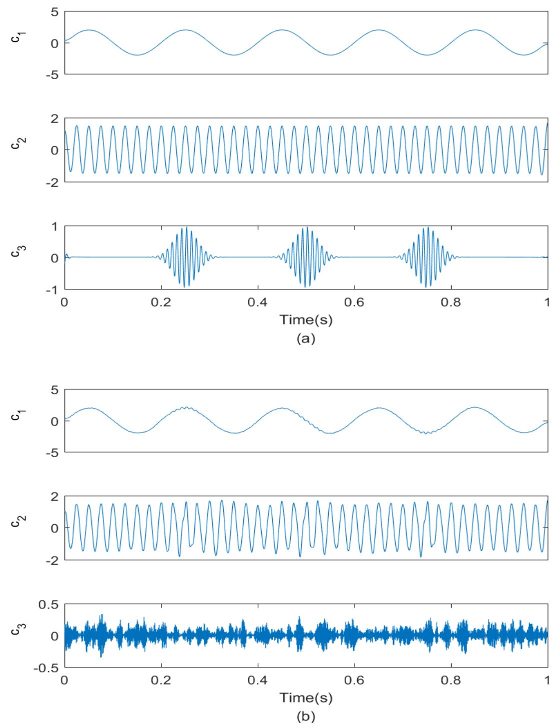


FIGURE 3. The processing results of simulation signals by VMD (mode number is 3 and the balancing parameter is 1500). (a) simulation signal y_{com}; (b) simulation signal $y_{com,e(t)}$.

EMD and EEMD. From this figure, the mode mixing problem of EMD is seen to be serious when it is used to deal with composite signal $y_{com,e(t)}$. In fact, no one completed sub-signal has been decomposed successfully by EMD. From the plots shown in Fig. 2 (b), it can be seen than the most information of y_1, y_2, and y_3 are decomposed into c_6, c_4, and c_3. However, part of the simulation signal y_1 is falsely decomposed into IMF c_5, and the reason is that the parameter settings are not optimal. Meanwhile, these decomposed components of EEMD are not strictly IMFs.

According to the principle of VMD, before the signals are processed, the mode number, balancing parameter of the data fidelity constraint, and the time step of the dual ascent should be preset, and among these the mode number is the most important one. Although the VMD algorithm can extract the characteristic information hidden in vibration signals, its decomposition effect depends on accurate parameter selection. That is, improper parameter setting will reduce its ability in signal decomposition. In theory, we should first determine the mode number and then select suitable balancing parameters of the data fidelity constraint and the time step of the dual ascent. Therefore, in this section, we discuss the effect of the mode number on decomposition results of VMD when the balancing parameter of the data fidelity constraint is set as 1500 suggested in [27]. Fig. 3 shows the processing results of simulation signal y_{com} and $y_{com,e(t)}$ by VMD whose mode number and balancing parameter are set as 3 and 1500, respectively. In Fig. 3 (a), the simulation is without noise, and hence we can easily determine that its real mode number is equal to those of its components. That is, here, the mode number of VMD is 3. From the plots shown in Fig. 3 (a), we can see that the simulation sub-signals are all successfully decomposed into three IMFs. However, if there are some noise components in the final simulated composite signal. This means that the real mode number of VMD is larger than that under the simulation without noise. Here, we still set the mode number of VMD as 3 and its decomposition results are plotted in Fig. 3 (b). In this figure, it can be noticed that the main simulation sub-signals are divided into the first two IMFs and the first IMFs include two simulation sub-signals with signals y_1 and y_3, i.e., these sub-signals with various structures are not successfully separated. Through the above comparative analysis, it can be found that traditional EMD, EEMD, and VMD all have some problems that need to be directly addressed before they can be used to deal with complex signals. In this paper, a new signal processing approach combining the best characteristics of EMD and VMD is proposed for incipient fault detection of rolling bearing.

III. INCIPIENT FAULT DIAGNOSIS WITH IVMD

VMD is a novel signal decomposition method that is theoretically well founded and can deal with nonlinear and nonstationary signals. However, it is hard to select a reasonable mode number for VMD. For the EMD method, the number of its IMFs can be obtained adaptively. Therefore, this work combines the best features of EMD and VMD to construct an approach for signal processing and weak fault extraction. Meanwhile, transient impulse analysis can reveal the characteristic information related to bearing faults. Therefore, the proposed method includes two parts: vibration signals processing with the proposed IVMD method and envelope analysis-based transient impulse information extraction; its main flow chart is shown in Fig. 4.

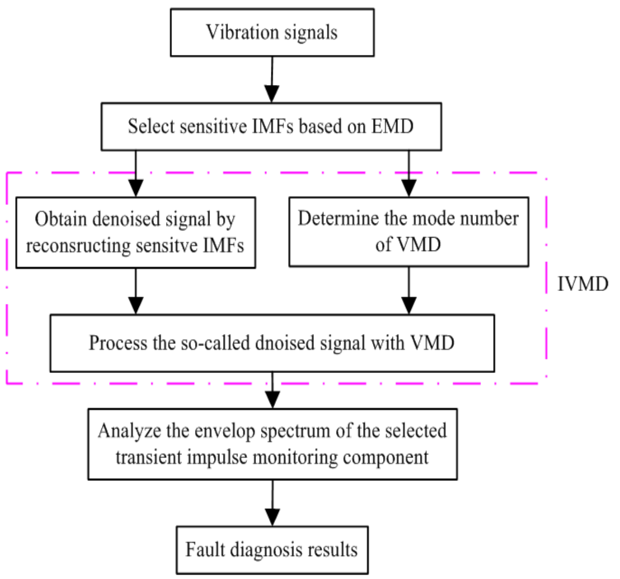


FIGURE 4. Flow chart of the proposed method.

A. DENOISING WITH EMD BASED ON SENSITIVE IMF SELECTION

A set of IMFs can be obtained by using EMD to process the collected vibration signals and the fault-related information may be decomposed into several IMF components. Therefore, how to select the sensitive IMFs is important for constructing a precise fault diagnosis method. For a mechanical system, a structure failure make its resilience metric or impact performance will be different from this under normal condition [30]. In this work, an index combining kurtosis and the correlation analysis between the decomposed IMFs with the original vibration signal is designed to select the sensitive IMFs, and the procedure is presented in the following steps.

(1) Compute correlation coefficient $r_{c_i,x}$ between the ith IMF and its original vibration signal $x(t)$ by

$$r_{c_i,x} = \| \frac{\sum_{t=1}^{n}(x(t)-\overline{x})(c_i(t)-\overline{c_i})}{\sigma_{c_i}\sigma_x} \| \tag{12}$$

where $c_i(t)$ is the ith decomposed IMF of EMD; σ_{c_i} and σ_x are the deviations of c_i and $x(t)$; and $\overline{c_i}$ and $\overline{x}$ are the mean values of c_i and $x(t)$, respectively.

(2) Calculate the kurtosis of the IMFs with

$$k_{c_i} = \frac{\sum(c_i(t)-\overline{c_i})^4}{(N-1)\sigma_{c_i}^4} \tag{13}$$

where N is the length of signal $c_i(t)$.

(3) Design a sensitive IMF evaluation index (SIEI) with

$$I_{SIEI}(i) = \frac{r_{c_i}}{\sum r_{c_i}} + \frac{k_{c_i}}{\sum k_{c_i}}. \tag{14}$$

(4) Rank these SIEI values from the greatest to the least. The K' IMFs with the biggest SIEI values are selected as sensitive IMFs to reconstruct the clear or denoising signal $x_{clear}(t)$ as

$$x_{clear}(t) = \sum_{k=1}^{K'} c_k(t). \tag{15}$$

According to the procedure for reconstructing the denoising signal, in fact, searching the sensitive IMFs is a process to determine whether an IMF is fault-related information or noise. Here, these sensitive IMFs are regarded as useful information and the remaining IMFs are noise components. At the same time, the fault characteristic frequencies (FCFs) of a bearing and their harmonics mainly occur in the case of high-frequency components, and hence these IMFs with the main frequency lower than the FCF can also be regarded as noise. After the selection of sensitive IMFs with SIEI and the frequency analysis, the noise in the vibration signals can be removed by reconstructing new vibration signals just with these selected IMFs.

B. THE PRINCIPLE OF IVMD

As we know, EMD, a self-adaptive time-frequency analysis algorithm, can be used to decompose a complex signal into a set of IMFs based on the characteristic of the signal. Although the EMD method suffers from the mode mixing problem when it is employed to analyze vibration signals, the number of its IMFs can be obtained adaptively. Meanwhile, the IMFs of EMD are nearly orthogonal functions and each IMF represents a simple oscillation mode with a physical meaning [13]. For EMD, if there is no so-called mode mixing problem, each IMF can be regarded as an independent mode. Even though a mode is decomposed into two or more IMFs at the same time, as long as no two or more IMFs exist that only belong to a mode, this is easy to be satisfied, and the number of IMFs can be used to set the mode number of VMD. Therefore, in this work, the signal reconstructed just with these selected sensitive IMFs of EMD is decomposed by VMD, and hence its modes can be regarded as these selected sensitive IMFs. This is the principle of IVMD.

To testify the effectiveness of the IVMD method, four sensitive IMFs of these plots in Fig. 2 are selected to reconstruct the denoising signal. By the proposed sensitive IMFs assessment method shown in above subsection, the first four IMFs are selected. For this reconstructed signal, it is made up of four IMFs, and the mode number of VMD is set as 4 and the decomposition results are shown in Fig. 5. From this figure, we can see that these simulation signals y_1, y_2, and y_3 are successfully decomposed into c_1, c_2, and c_3. Contrastive analysis between the plots of Fig. 2 (a) and the plots of Fig. 5 reveals that the proposed IVMD provides

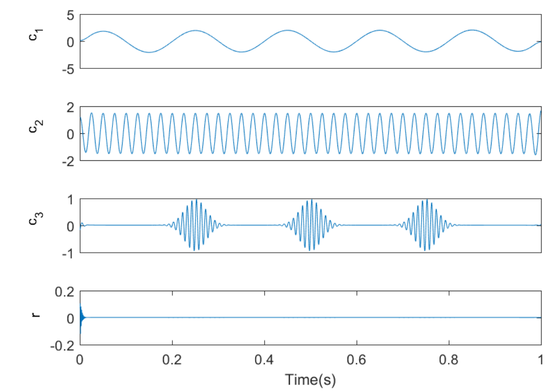


FIGURE 5. **The decomposition results of simulation signals shown in Fig. 1 with the proposed IVMD.**

better decomposition than traditional EMD. In Fig. 2 (b), the most information of simulation signals y_1, y_2, and y_3 are decomposed into components c_6, c_4, and c_3. Therefore, here, these three IMFs are selected to be the final decomposition results of the simulation signal with EEMD. Although the EEMD is able to decompose these simulation components just by judging from the general shape of these IMFs, it still needs to evaluate the matching degree of the decomposition signal and the real signal. The decomposition error is used to as the index to indicate the decomposition effect.

Fig. 6 shows the decomposition errors of simulation signals with EEMD and the proposed IVMD, in which the decomposition errors of simulation signal y_1, y_2, and y_3 are defined as $y_{1,erro}$, $y_{2,erro}$, and $y_{3,erro}$, respectively. From Fig. 6 (a), it can be easily seen that the maximum error is no more than 0.5 and the errors of almost all the points of the decomposition components are close to zero. However, the decomposition error of EEMD is obviously bigger than that of the proposed method. Therefore, we can conclude that the proposed IVMD achieves better decomposition precision than traditional EEMD. Furthermore, computational efficiency is also an important performance parameter for a signal processing method. Here, in the same compute, the calculation times of EEMD with the parameters defined in [23] and IVMD are 7.6798 s and 0.3651 s, respectively. That is, the proposed IVMD is also more time-saving than EEMD.

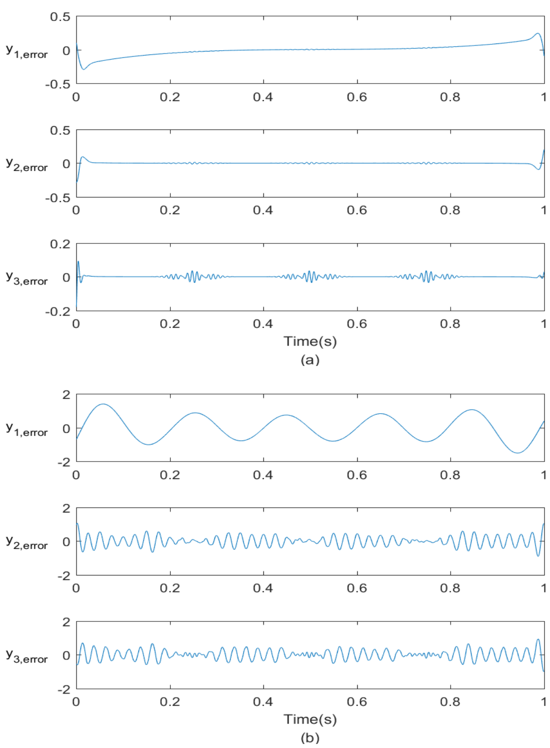


FIGURE 6. **Decomposition error of the simulation signals with (a) IVMD and (b) EEMD.**

C. TRANSIENT IMPULSE MONITORING COMPONENTS

If one of the bearing surfaces has a local defect, it will cause a set of impacts when the rolling element passes through the faulty surface. The FCFs of these impulses belonging to various bearing faults are determined by the motor rotational speed, fault location, and bearing geometric dimensioning. Therefore, we can identify a bearing fault by monitoring the FCFs and the FCFs can be formulated as follows [31]– [33].

$$\begin{cases} f_c = \dfrac{f_r}{2}(1 - \dfrac{B_d}{P_d}\cos(\theta)) \\ f_o = \dfrac{Nf_r}{2}(1 - \dfrac{B_d}{P_d}\cos(\theta)) \\ f_i = \dfrac{Nf_r}{2}(1 + \dfrac{B_d}{P_d}\cos(\theta)) \\ f_b = \dfrac{B_d f_r}{P_d}(1 - \dfrac{B_d^2}{P_d^2}\cos(\theta)) \end{cases} \tag{16}$$

where f_c, f_o, f_i, and f_b are the fundamental cage frequency, outer race fault frequency, inner race fault frequency, and rolling element fault frequency, respectively; N is the number of rolling elements; θ is the angle of the load from the radial plane; B_d and P_d are the ball and pitch diameter, respectively; and f_r is the rotating frequency.

A harmonic will be made in the spectrum of vibration signal when a defect appears on each part of the bearing. Generally, these FCFs include the main information related to various bearing faults and it can be extracted for bearing fault diagnosis. The vibration signals collected from faulty bearings include not only the resonance frequency in high frequency range excited by impacts but also the fundamental rotating frequency and its harmonics in low frequency range [34]. Therefore, the denoised signals obtained by the proposed IVMD are then carried out with envelope demodulation analysis to search the FCFs for bearing fault diagnosis.

As we know, some IMFs can be obtained by the proposed IVMD method, and not all of them are suitable for bearing fault monitoring or detecting. For a kind of bearing faults, such as outer race fault, inner race fault, or rolling element fault, the first thing is to monitoring the changes in the corresponding fault characteristic frequency. That is to say, we should select a IMF whose FCFs are easy to be found and the procedure is shown as follows:

(1) The envelop analysis of final IMFs, produced by the proposed IVMD method, is executed by Hilbert transform.

(2) The f_o, f_i, and f_b of bearing are calculated by (16).

(3) Define the frequency interval $[f_{low}, f_{up}]$ for searching the FCF. Here, we define that: $f_{low} = f_{center} - \Delta f$, and $f_{up} = f_{center} + \Delta f$. In this paper, we set $f_{center} = f_o$. Then the selection criteria for Δf is to make f_i, f_o and f_b can fall within the interval $[f_{low}, f_{up}]$.

(4) Search the frequency component with the maximum value in interval $[f_{low}, f_{up}]$ defined in step (1), and which is marked as f_{max}. At the same time, the mean value of these frequencies among $[f_{low}, f_{up}]$ is calculated, and which is denoted as f_{mean}.

(5) Set

$$I_{fre,i} = \frac{f_{max,i}}{f_{mean,i}}, \tag{17}$$

where $I_{fre,i}$ is defined as transient impulse monitoring index (TIMI), and i means the ordinal number of final IMFs of the proposed IVMD method.

(6) Finally, this IMF with the biggest $I_{fre,i}$ is defined as the sensitive monitoring component for detecting incipient bearing faults by using envelop analysis and fast Fourier transform.

IV. EXPERIMENTAL VERIFICATION

A. EXPERIMENTAL SETTING

The data of most fault diagnosis methods for rolling element bearings come from simulation mode or "seeded" detect, and is hard to reflect the real incipient failure or the natural damage peculiarity of bearings in the early stage. Therefore, the data collected from the full life cycle, i.e., normal-to-failure experiment, of bearings should be provided for the incipient fault diagnosis of bearings.

For a new fault detection method, it is important to collected reliability data sets before performing fault detection assessment [35]. Therefore, this work designs an experiment based on the standard prognostic vibration data collected by the Intelligent Maintenance Systems (IMS) of University of Cincinnati [36]. This test bench, shown in Fig. 7, includes an AC motor, several rub belts, a shaft, and four bearings. The type of the used bearings is Rexnord ZA-2115 double row bearings that were installed on the shaft. In this system, the rotation speed was set a constant value with 2000 RPM through the AC motor coupled to the shaft via rub belts. Meanwhile, a spring mechanism was used to simulate a radial load of 6000 lbs onto the shaft and bearings. As shown in Fig. 7, PCB 353B33 High Sensitivity Quartz ICP accelerometers were installed on the bearing housing to collected test data by NI DAQ Card 6062E with the sampling rate set at 20 kHz. Final, an outer race fault occurred after exceeding designed life time of the bearing which is more than 100 million revolutions. The test bearing has 16 rollers in each row and its pitch diameter and roller diameter are 71.501 mm

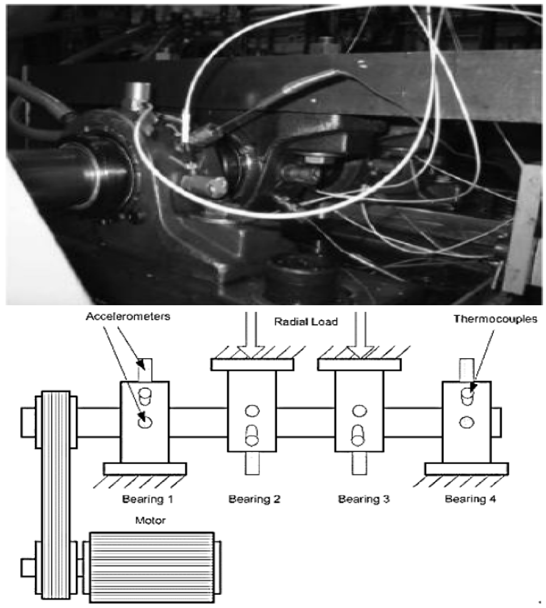


FIGURE 7. Bearing test rig and sensor placement illustration [36].

and 8.4074 mm, respectively. Therefore, the FCF of outer race fault is 236.4 Hz. Each file consists of 20,480 points and the recording interval time of each file is 10 minutes. In the normal-to-failure experiment, total 984 data files were collected in the duration of February 12, 2004 to February 19, 2004 and outer race defect was found in bearing 1 at last.

B. EXPERIMENTAL RESULTS AND ANALYSIS

Fig.8 illustrate the root mean square (RMS) of the vibration signal of bearing 1 under entire life-cycle. From this figure, the range of incipient fault probability is from 5000th minute to 65000th minute. Considering the RMS values of incipient bearing fault and normal condition will be different, so the mutation points at T1 (around 5320th minute) and T2 (around 6460th minute) are estimated to be the points that the bearing

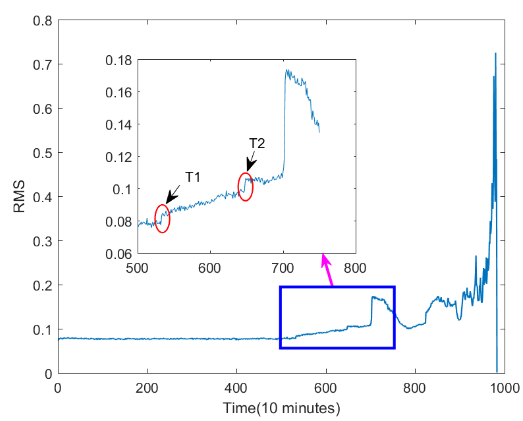


FIGURE 8. RMS for the entire life cycle of bearing 1.

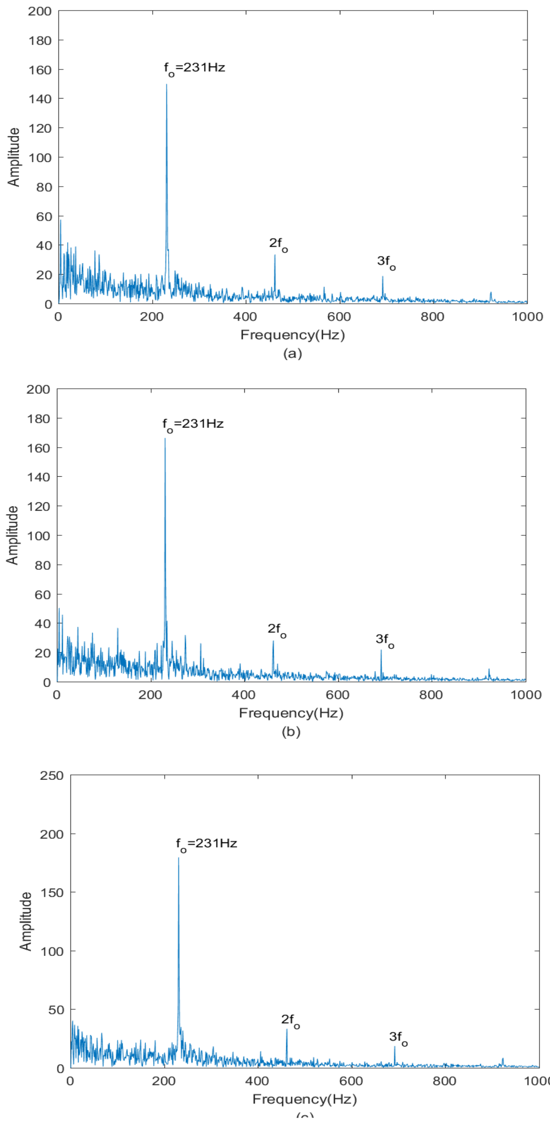


FIGURE 9. The analyzed results of the proposed method with the data collected at: (a) 6470th minute; (b) 6460th minute; (c) 6450th minute.

TABLE 1. SIEI values of vibration signals in case 1.

Time(minute)	$IMF1$	$IMF2$	$IMF3$	$IMF4$	$IMF5$	$IMF6$	$IMF7$	$IMF8$
6450th	0.5097	0.2462	0.2369	0.1376	0.1071	0.1141	0.1212	0.1674
6460th	0.5064	0.2457	0.2295	0.1421	0.1164	0.1428	0.1252	0.0903
6470th	0.4946	0.2468	0.2366	0.1460	0.1283	0.1062	0.1414	0.1316

become into fault from normal condition. Therefore, the data sets nearby the mutation points are processed by the proposed method, and which are: case 1 nearby 6460th minute; case 2 nearby 5320th minute.

First, we employ the proposed IVMD method to processing the vibration signal defined in case 1 and the corresponding SIEI and TIMI values are listed in Table 1 and Table 2, respectively. For EMD processing stage, the first six IMFs with biggest SIEI are selected for further processing by IVMD. When calculating the TIEI values, the Δf is set as 100 Hz.

TABLE 2. TIMI values of vibration signals in case 1.

Time(minute)	$IMF1$	$IMF2$	$IMF3$	$IMF4$	$IMF5$	$IMF6$	$IMF7$
6450th	0.6691	0.6641	0.8822	0.9115	0.9366	0.7330	0.7767
6460th	0.6299	0.6859	0.8888	0.9159	0.9311	0.7346	0.7470
6470th	0.6283	0.6472	0.8849	0.8874	0.9268	0.7127	0.7127

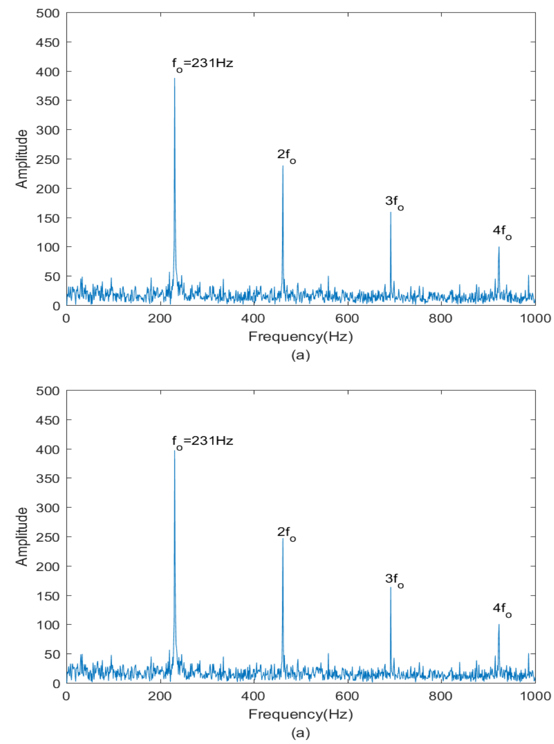


FIGURE 10. The envelop analysis results of the data collected at 6460th minute with (a) EMD and (b) EEMD.

From these TIMI values in Table 1, it can be seen that transient impulse monitoring components of these data collected at 6470th minute, 6460th minute and 6450th minute are all the fifth IMFs of the IVMD. Fig. 9 shows the analyzed results of the proposed method under case 1. In Fig. 9, (a), (b) and (c) mean the envelope spectrums of the data collected at 6470th minute, 6460th minute and 6450th minute with the proposed method, respectively. According to this figure, we can see that the characteristic frequency at 231 Hz (it is roughly regarded as the FCF of outer race fault) and its harmonic are successfully extracted in Fig. 9 (a) to (c), this means there is a defect in the outer race and the mutation points T2 may not be the time that the bearing become faulty condition from normal. Fig. 10 plots the envelop analysis results of the data collected at 6450th minute with EMD and EEMD. In this figure, the FCF and its harmonics of outer race fault are obvious. That is to say, at mutation points T2, the outer race fault of bearing also can be easily detected by using EMD and EEMD.

Table 3 shows the SIEI values of these IMFs of EMD. According to this table, the first six IMFs with biggest SIEI are also selected for further processing. After IVMD processing, seven IMFs can be obtained and their ITME values are

TABLE 3. SIEI values of vibration signals in case 2.

Time(minute)	$IMF1$	$IMF2$	$IMF3$	$IMF4$	$IMF5$	$IMF6$	$IMF7$	$IMF8$
5310th	0.4232	0.2518	0.2542	0.1506	0.1283	0.1050	0.1323	0.1063
5320th	0.4298	0.2573	0.2564	0.1483	0.1064	0.1050	0.1754	0.1069
5330th	0.4310	0.2652	0.2465	0.1647	0.1476	0.1100	0.1315	0.0925

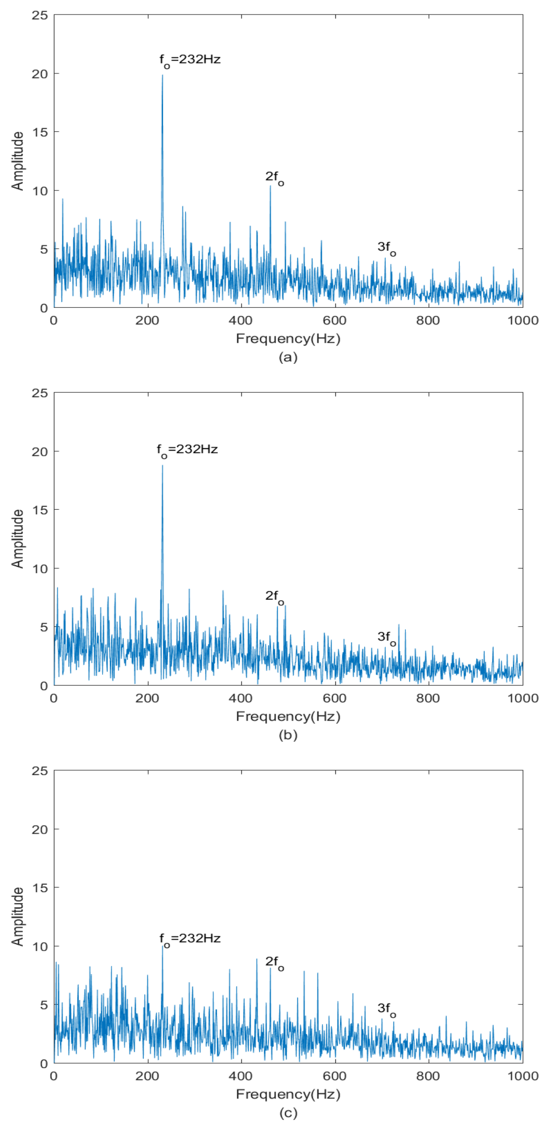


FIGURE 11. The analyzed results of the proposed method with the data collected at: (a) 5330th minute; (b) 5320th minute; (c) 5310th minute.

TABLE 4. TIMI values of vibration signals in case 2.

Time(minute)	$IMF1$	$IMF2$	$IMF3$	$IMF4$	$IMF5$	$IMF6$	$IMF7$
5310th	0.5935	0.6861	0.6188	0.6589	0.6854	0.5794	0.7084
5320th	0.5756	0.7046	0.7529	0.6052	0.7050	0.6914	0.8278
5330th	0.6607	0.6798	0.6889	0.6719	0.7729	0.6434	0.8379

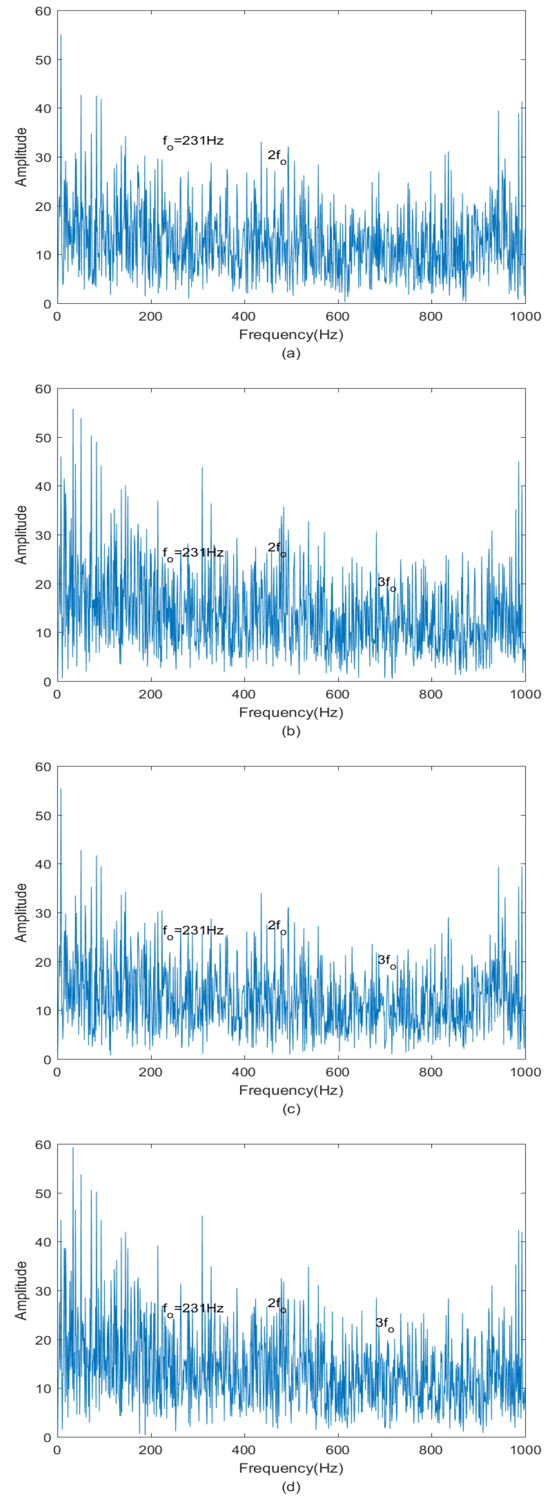


FIGURE 12. The analyzed results of the comparison methods: (a) directly envelope analysis; (b) EMD and envelope analysis; (c)denoising signal with EMD and envelope analysis;and (d) EEMD and envelope analysis.

shown in Table 4. From this Table, it can be seen that transient impulse monitoring components of these data collected at 6470th minute, 6460th minute and 6450th minute are all the seventh IMFs of the IVMD. Fig. 11 shows the analyzed results of the proposed method under case 2. Meanwhile, Fig. 9 (a), (b) and (c) mean the envelope spectrums of the data collected at 5330th minute, 5320th minute and 5310th minute with the proposed method, respectively. For case 2, the so-called FCF of outer race fault can be easily observed in Fig. 11 (a) and (b) and this reveals the outer race fault occurs at 5330th minute and 5320th minute. However, the analysis of these plots in Fig. 11 (c) reveals that the FCF of outer race fault cannot be found obviously because many other frequency peaks are also very highlight, so it can get a conclusion that the outer race fault may not occur. Comparative analysis of Fig. 9 and Fig. 11 also indicates that the mutation point nearby 5320th minute is the time that the outer race develops from normal condition to incipient fault and the mutation point nearby 6460th minute may be the time that the incipient fault develops further.

To prove the superiority of the proposed method, comparison experiments are carried out and which are: directly envelope analysis; EMD and envelope analysis; denoising signal with EMD and envelope analysis; EEMD and envelope analysis. From Fig. 11, we can see that the incipient fault can be found at the 5320th minute with the proposed method. Therefore, the data collected at the 5320th minute is analyzed with these considered comparison methods. Fig. 12 shows the results of these considered experiments with other analyzing methods. For these plots in Fig. 12 (a), the peak frequency nearby the FCF of outer race fault is 231 Hz and which is not so obvious to be found. At the same time, the outstanding peaks are all located away from FCF and its harmonic. That is, the incipient fault in outer race is hard to be identified by direct envelope analysis with vibration signals. EMD and EEMD are famous methods for signals processing, here, their analysis for incipient fault diagnosis are also developed. In their analysis, this one of the first five IMFs with the biggest kurtosis is selected for envelop analysis. From Fig. 12 (b), we can see that the frequency peak with highest possibility to be the FCF of outer race fault is 231 Hz and which is almost overwhelmed by other frequency ingredients. That is, it is also hard to detect the outer race fault with EMD and envelope analysis. The similar situations occur in Fig. 12 (c) and Fig. 12 (d), the FCF components are also hard to be observed. Therefore, according to the contrastive analysis of the plots in Fig. 11 and Fig. 12, we can get a conclusion that the proposed IVMD can decomposed out the FCF information from complex vibration signal for bearing fault diagnosis and has better effect than traditional envelop analysis, EMD and EEMD do. Moreover, the time of the proposed method for analyzing the data collected at 5320th minute is only 20.5307 s, but the traditional EEMD with the parameters recommended as a case in [23] needs 69.2145 s. That is, the proposed method is also a time-saving approach for incipient fault detection than EEMD does.

V. CONCLUSION

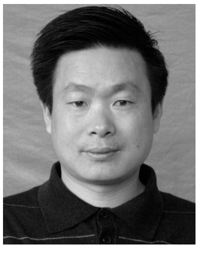
Transient impulse analysis is an effective way to detect the bearing fault at its early stage because these impulse compositions include abundant information regarding bearing faults. Variational mode decomposition (VMD) is able to reveal the weak transient impulses from complex vibration signals if its reasonable mode number can be pre-set. To extract the transient impulses and detect the incipient mechanical fault by the nonlinear and nonstationary vibration signals with heavy background noise, this paper proposes an improved VMD by combining the best characteristics of traditional VMD and EMD. First, EMD is used to obtained denoised signals the so-called useful IMFs which are evaluated by a sensitive IMF evaluation index (SIEI) which is based on the conjoint analysis of relatedness and kurtosis. Afterwards, VMD is further improved to effectively decompose the denoised signal and a transient impulse monitoring index (TIME) was designed to obtain the transient impulse monitoring component for incipient fault diagnosis with Hilbert envelope analysis. Experimental results shown that the proposed method can accurately extract the transient impulses of bearing fault in its early stage, and the detection results are better than those of traditional envelope analysis, EMD, EEMD and VMD. However, the so-called improved VMD proposed in this paper only focused on selecting reasonable mode number, so the next work should deal with the optimization of the balancing parameter of the data fidelity constraint or optimize both of these parameters to further improve the signal decomposition effect. Meanwhile, a real-time fault diagnosis can immediately eliminate the fault once it occurs [37], so the online incipient fault diagnosis method also is one of our next work.

REFERENCES

[1] C. Li and M. Liang, "Continuous-scale mathematical morphology-based optimal scale band demodulation of impulsive feature for bearing defect diagnosis," *J. Sound Vib.*, vol. 331, no. 26, pp. 5864–5879, 2012.
[2] B. Yao, P. Zhen, L. Wu, and Y. Guan, "Rolling element bearing fault diagnosis using improved manifold learning," *IEEE Access*, vol. 5, pp. 6027–6035, 2017.
[3] Y. Li, X. Liang, and M. J. Zuo, "Diagonal slice spectrum assisted optimal scale morphological filter for rolling element bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 85, pp. 146–161, Feb. 2017.
[4] S. Zhang, S. Lu, Q. He, and F. Kong, "Time-varying singular value decomposition for periodic transient identification in bearing fault diagnosis," *J. Sound Vib.*, vol. 379, pp. 213–231, Sep. 2016.
[5] Z. Feng, H. Ma, and M. J. Zuo, "Vibration signal models for fault diagnosis of planet bearings," *J. Sound Vib.*, vol. 370, pp. 372–393, May 2016.
[6] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Process.*, vol. 20, no. 7, pp. 1483–1510, 2006.
[7] Y. Liu, L. Guo, Q. Wang, G. An, M. Guo, and H. Lian, "Application to induction motor faults diagnosis of the amplitude recovery method combined with FFT," *Mech. Syst. Signal Process.*, vol. 24, no. 8, pp. 2961–2971, 2010.
[8] Q. Miao, L. Cong, and M. Pecht, "Identification of multiple characteristic components with high accuracy and resolution using the zoom interpolated discrete Fourier transform," *Meas. Sci. Technol.*, vol. 22, no. 5, p. 55701, 2011.
[9] J. Antoni and R. Randall, "The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines," *Mech. Syst. Signal Process.*, vol. 20, no. 2, pp. 308–331, 2006.
[10] Y. Zhou *et al.*, "Wigner–Ville distribution based on cyclic spectral density and the application in rolling element bearings diagnosis," *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, vol. 225, no. 12, pp. 2831–2847, 2011.
[11] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Rolling element bearing fault diagnosis using autocorrelation and continuous wavelet transform," *J. Vib. Control.*, vol. 17, no. 14, pp. 2081–2094, 2011.
[12] G.-S. Zhong, L.-P. Ao, and K. Zhao, "Influence of explosion parameters on wavelet packet frequency band energy distribution of blast vibration," *J. Central South Univ.*, vol. 19, no. 9, pp. 2674–2680, 2012.

[13] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
[14] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2008.
[15] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
[16] R. Ricci and P. Pennacchi, "Diagnostics of gear faults based on EMD and automatic selection of intrinsic mode functions," *Mech. Syst. Signal Process.*, vol. 25, no. 3, pp. 821–838, 2011.
[17] Q. Du and S. Yang, "Application of the EMD method in the vibration analysis of ball bearings," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2634–2644, 2007.
[18] Q. Xiong, Y. Xu, Y. Peng, W. Zhang, Y. Li, and L. Tang, "Low-speed rolling bearing fault diagnosis based on EMD denoising and parameter estimate with alpha stable distribution," *J. Mech. Sci. Technol.*, vol. 31, no. 4, pp. 1587–1601, 2017.
[19] J.-H. Ahn, D.-H. Kwak, and B.-H. Koh, "Fault detection of a roller-bearing system through the EMD of a wavelet denoised signal," *Sensors*, vol. 14, no. 8, pp. 15022–15038, 2014.
[20] C. Yi, J. Lin, W. Zhang, and J. Ding, "Faults diagnostics of railway axle bearings based on IMF's confidence index algorithm for ensemble EMD," *Sensors*, vol. 15, no. 5, pp. 10991–11011, 2015.
[21] H. Zhao, M. Sun, W. Deng, and X. Yang, "A new feature extraction method based on EEMD and multi-scale fuzzy entropy for motor bearing," *Entropy*, vol. 19, no. 1, p. 14, 2017.
[22] H. Chen, P. Chen, W. Chen, C. Wu, J. Li, and J. Wu, "Wind turbine gearbox fault diagnosis based on improved EEMD and Hilbert square demodulation," *Appl. Sci.*, vol. 7, no. 2, p. 128, 2017.
[23] M. Žvokelj, S. Zupan, and I. Prebil, "EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis," *J. Sound Vib.*, vol. 370, pp. 394–423, May 2016.
[24] Y. Lei and M. J. Zuo, "Fault diagnosis of rotating machinery using an improved HHT based on EEMD and sensitive IMFs," *Meas. Sci. Technol.*, vol. 20, no. 12, p. 125701, 2009.
[25] G. Tang, G. Luo, W. Zhang, C. Yang, and H. Wang, "Underdetermined blind source separation with variational mode decomposition for compound roller bearing fault signals," *Sensors*, vol. 16, no. 6, p. 897, 2016.
[26] H. Zhao and L. Li, "Fault diagnosis of wind turbine bearing based on variational mode decomposition and Teager energy operator," *IET Renew. Power Gener.*, vol. 11, no. 4, pp. 453–460, 2016.
[27] M. Zhang, Z. Jiang, and K. Feng, "Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump," *Mech. Syst. Signal Process.*, vol. 93, no. 15, pp. 460–493, 2017.
[28] Z. Lv, B. Tang, Y. Zhou, and C. Zhou, "A novel method for mechanical fault diagnosis based on variational mode decomposition and multikernel support vector machine," *Shock Vib.*, vol. 2016, Oct. 2016, Art. no. 3196465. [Online]. Available: https://www.hindawi.com/journals/sv/2016/3196465/
[29] B. Cai, L. Huang, and M. Xie, "Bayesian networks in fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 5, pp. 2227–2240, Oct. 2017.
[30] B. Cai, M. Xie, Y. Liu, Y. Liu, and Q. Feng, "Availability-based engineering resilience metric and its corresponding evaluation methodology," *Rel. Eng. Syst. Saf.*, vol. 172, pp. 216–224, Apr. 2018.
[31] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011.
[32] Y. Wang, G. Xu, Q. Zhang, D. Liu, and K. Jiang, "Rotating speed isolation and its application to rolling element bearing fault diagnosis under large speed variation conditions," *J. Sound Vib.*, vol. 348, pp. 381–396, Jul. 2015.
[33] P. D. McFadden and J. D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *J. Sound Vib.*, vol. 96, no. 1, pp. 69–82, Sep. 1984.
[34] R. B. Randall, J. Antoni, and S. Chobsaard, "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals," *Mech. Syst. Signal Process.*, vol. 15, no. 5, pp. 945–962, 2001.
[35] Y. Liu, Y. Pan, Z. Sun, and D. Huang, "Statistical monitoring of wastewater treatment plants using variational Bayesian PCA," *Ind. Eng. Chem. Res.*, vol. 53, no. 8, pp. 3272–3282, 2014.
[36] H. Qiu, J. Lee, and J. Lin, "Wavelet filter-based weak signature detection method and its application on roller bearing prognostics," *J. Sound Vib.*, vol. 289, pp. 1066–1090, 2006.
[37] B. Cai, H. Liu, and M. Xie, "A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks," *Mech. Syst. Signal Process.*, vol. 80, pp. 31–44, Dec. 2016.

FAN JIANG received the bachelor's degree in measurement and control technology and instrument and the Ph.D. degree in mechanical and electrical engineering from the China University of Mining and Technology, Xuzhou, China, in 2010 and 2015, respectively. His current research interests include machinery condition monitoring, fault diagnosis, and time-frequency signal analysis.



ZHENCAI ZHU received the Ph.D. degree in mechanical design and theory from the China University of Mining and Technology, Xuzhou, China, in 2000. He is currently a Professor with the School of Mechatronic Engineering, China University of Mining and Technology. His current research interests include mechanical design, dynamics analysis, tribology, mechanical condition monitoring, and reliability assessment.

WEI LI received the Ph.D. degree in automation engineering from the Institute for Automatic Control and Complex Systems, University of Duisburg-Essen, Duisburg, Germany. He is currently a Professor with the School of Mechatronic Engineering, China University of Mining and Technology. His current research interests include mechanical condition monitoring, fault diagnosis, life prediction, and reliability assessment.