

Article

Empirical Variational Mode Decomposition Based on Binary Tree Algorithm

Huipeng Li ^{1,2} , Bo Xu ^{1,2,*} , Fengxing Zhou ¹, Baokang Yan ¹ and Fengqi Zhou ¹

¹ Engineering Research Center for Metallurgical Automation and Measurement Technology, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; yehuip@hgnu.edu.cn (H.L.); zhoulengxing@wust.edu.cn (F.Z.); yanbaokang@wust.edu.cn (B.Y.); zhoulengqi@wust.edu.cn (F.Z.)

² School of Physics and Electronic Information, Huanggang Normal University, Huanggang 438000, China

* Correspondence: xubo@hgnu.edu.cn

Abstract: Aiming at non-stationary signals with complex components, the performance of a variational mode decomposition (VMD) algorithm is seriously affected by the key parameters such as the number of modes K , the quadratic penalty parameter α and the update step τ . In order to solve this problem, an adaptive empirical variational mode decomposition (EVMD) method based on a binary tree model is proposed in this paper, which can not only effectively solve the problem of VMD parameter selection, but also effectively reduce the computational complexity of searching the optimal VMD parameters using intelligent optimization algorithm. Firstly, the signal noise ratio (SNR) and refined composite multi-scale dispersion entropy (RCMDE) of the decomposed signal are calculated. The RCMDE is used as the setting basis of the α , and the SNR is used as the parameter value of the τ . Then, the signal is decomposed into two components based on the binary tree mode. Before decomposing, the α and τ need to be reset according to the SNR and MDE of the new signal. Finally, the cycle iteration termination condition composed of the least squares mutual information and reconstruction error of the components determines whether to continue the decomposition. The components with large least squares mutual information (LSMI) are combined, and the LSMI threshold is set as 0.8. The simulation and experimental results indicate that the proposed empirical VMD algorithm can decompose the non-stationary signals adaptively, with lower complexity, which is $O(n^2)$, good decomposition effect and strong robustness.

Keywords: non-stationary signal; empirical variational mode decomposition; binary tree; least square mutual information; information entropy



Citation: Li, H.; Xu, B.; Zhou, F.; Yan, B.; Zhou, F. Empirical Variational Mode Decomposition Based on Binary Tree Algorithm. *Sensors* **2022**, *22*, 4961. <https://doi.org/10.3390/s22134961>

Academic Editor: Len Gelman

Received: 23 May 2022

Accepted: 28 June 2022

Published: 30 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Many physical semaphores in real life are composed of multi-components, which have nonlinear and nonstationary characteristics. It is an essential way to explore the system characteristic by analyzing the inherent information contained in the components. Due to the interference of the external environment, these characteristic components are difficult to effectively identify. Consequently, the effective extraction of these signal modes becomes very important for the research of corresponding systems. The methods of signal or data processing have attracted more and more attention in various fields.

In recent years, many non-stationary signal processing methods have been proposed by scholars. Short time Fourier transform (STFT) is a traditional time-frequency transform (TF) method [1]. It provides a graphic display in the TF domain and has been successfully applied to the evaluation of mechanical fault characteristics [2]. STFT is restricted by the time-frequency resolution of window function, and the time-frequency resolution is fixed. Wavelet transform (WT) [3] has high-precision time-frequency resolution, and the selection of the WT basis function has a huge impact on performance and has strong experience [4].

Wigner Ville Distribution (WVD) [5] is a method to calculate the time–frequency distribution. It can reflect the instantaneous time–frequency relationship of a signal [6]. However, it is seriously affected by cross terms. S-transform (ST) [7] combines the advantages of STFT and WT, but its spectrum is rough and restricted by the Heisenberg’s uncertainty principle [8]. In theory, high-order statistics (HOS) [9] can completely suppress Gaussian noise, but it has poor effect on non-Gaussian noise and interferes with the high-order spectrum of the signal. Empirical mode decomposition (EMD) [10] is a method to decompose signals into a set of intrinsic mode functions (IMF). It is an excellent adaptive signal processing method and has been widely used in engineering [11,12]. However, there are some problems such as fitting overshoot, endpoint effect and modal aliasing, which seriously restrict its practical application [13]. Local mean decomposition (LMD) [14] adaptively decomposes the non-stationary multi-component signal into the sum of several product functions (PF) with physical meaning of instantaneous frequency [15]. It has the disadvantages of signal mutation and large amount of calculation caused by demodulation [16]. Inherent time scale decomposition (ITD) [17] suffers from the problem that the waveform of PF component fluctuates locally, resulting in signal distortion [18]. Although many scholars put forward a large number of improvement methods to the above method, they also gained good results and successfully applied them to mechanical fault diagnosis. Nevertheless, confined by the theoretical framework, these problems can be suppressed to a certain extent, but cannot be fundamentally excluded.

VMD [19] is a new adaptive signal processing method, which decomposes the signal into a set of modal functions with limited bandwidth by iteratively solving the variational problem. It realizes the frequency separation of each signal component and overcomes the problems of endpoint effect, modal mixing and waveform fluctuation existing in EMD, LMD and ITD. This strategy is very suitable for analyzing nonlinear and non-stationary vibration signals, and has been widely used in the engineering field [20,21]. However, the performance of VMD is affected by the inherent decomposition parameters i.e., the total number of modes K , the quadratic penalty parameter α , update step τ . Moreover, these parameters must be preset and have strong artificial experience. VMD converts signal decomposition into a constrained variational problem and adaptively decomposes the signal into the sum of several Intrinsic Mode Function (IMF) components, which is essentially different from previous signal processing methods [22]. However, the key parameters must be artificially set in advance, and these parameters have a great influence on the decomposition results. The most important thing is that there is no standard to measure the result of decomposition. Therefore, in practical application, the process of artificially setting parameter values in advance indicates that it is not a fully adaptive model [23,24]. In recent years, many scholars have conducted relevant research on VMD parameters setting. Xiao et al. [25] proposed an optimal value search method for the decomposition parameters (α and K) of VMD. However, the search method relies on personal experience and intuitive search mechanism, and lacks the basis of mathematical theoretical framework. Kaur et al. [26] used discrete wavelet transform (DWT) and wavelet packet transform (WPT) to set the VMD decomposition mode number. However, under the strong noise interference, it is a great challenge to obtain the modal number. Long et al. [27] applied the center frequency observation method to set the number of modes, and modified the parameter value τ according to the residual index to verify the influence of the parameter on the signal decomposition. However, the method of screening single parameter value based on single index does not take into account the interaction between parameters. Considering the complexity and variability of actual vibration signals, the application of this observation method in engineering may be limited. Recently, some studies have proposed optimization algorithms to adjust the parameters. In Ref. [28], particle swarm optimization (PSO) was adopted to select the optimal combination value of decomposition parameters K and τ of VMD. This method alleviates the experience of setting parameters manually to a certain extent. Wang et al. [29] obtained the optimal selected parameters, namely the mode number and penalty parameter of VMD, by using a

PSO optimization algorithm through the appropriate fitness function. Xu et al. [30] applied the variable dimension composite chaotic algorithm to adaptively select parameters of VMD and obtained excellent performance. However, the complexity of the algorithm and the feasibility of practical application need to be verified. A new method termed variational mode extraction (VME) extracts the natural mode function by knowing its approximate center frequency [31], which can adaptively extract the modal components in the signal. However, the residual signal in its decomposition has no strict mathematical definition and physical significance. The improved VME method is named Successive VMD (SVMD) in [32,33]. The SVMD achieves good adaptive effect, but the residual signal still lacks strict physical definition. The mentioned methods have improved VMD performance and achieved satisfactory effect to some extent. However, the synergistic influence of VMD key parameters has not been fully considered.

In view of the above issues, this paper proposes an empirical VMD method based on the binary tree model. Its main innovations and contributions can be summarized as follows:

(1) The number of modes K in a traditional VMD needs to be manually set. In this paper, K is set as a fixed value, that is, $K = 2$, which can effectively avoid empirically setting the value of K . Then the decomposition is iteratively executed according to the binary tree model until the single component of the decomposition is duplicated. Lastly, according to the mutual information between each component, the components with larger mutual information value are added to obtain the new IMF component.

(2) The α of traditional VMD still needs to be manually set. This paper presents a calculation equation of α , namely $\alpha = \text{round}(\text{RCMDE} \times (f_s/2) * \log(K))$. Where f_s is the sampling frequency of the signal, $\text{round}(\cdot)$ is the rounding function, RCMDE is the refined composite multi-scale dispersion entropy. RCMDE can measure the complexity of signal well and adjust the value of α dynamically.

(3) The τ critically affects the convergence of VMD. Hence, it is required to select an appropriate τ value according to the noise level of the signal to guarantee the optimal convergence of the algorithm. In this paper, the signal to noise ratio (SNR) is proposed to dynamically set the τ value, which can effectively guarantee the convergence and convergence speed of the algorithm.

(4) The algorithm proposed in this paper fully considers the common influence of several key parameters of VMD and is adaptive. Compared with the improved method based on the intelligent optimization algorithm, its computational complexity is lower.

Lastly, the effectiveness and superiority of the proposed method are verified by analyzing the simulation signal and the measured vibration signal.

The rest of this paper is organized as follows. The Section 2 introduces the relevant basic theoretical knowledge in detail. The Section 3 introduces the empirical VMD algorithm structure based on binary tree model in detail. In the Section 4, the effectiveness of the proposed algorithm is experimentally verified. Finally, the conclusion is drawn in the Section 5.

2. Related Works

2.1. Brief Introduction of VMD

VMD is an adaptive signal decomposition method based on Wiener filter, Hilbert transform and heterodyne demodulation. Its purpose is to decompose a real-valued input signal $x(t)$ into a set of sub-modes $\{u_k\}$ with a particular sparsity. Each of the limited bandwidth sub-modes is tightly centered around a central frequency ω_k . In the process of VMD, each submode u_k is transformed by Hilbert transform to obtain the corresponding unilateral spectrum. The analytical signal of each component is mixed with a pre-estimated center frequency $e^{-j\omega_k t}$ and the spectrum of each mode is modulated to the corresponding base-band. In order to obtain the bandwidth of the sub-modes, the constraint variational problem is introduced to calculate the square L^2 -norm of the gradient of the demodulated

signals and estimate the bandwidth of each mode. The optimized variational model constructed is shown in Equation (1).

$$\begin{cases} \min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_{k=1}^K \|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \quad \sum_{k=1}^K u_k(t) = x(t) \end{cases} \quad (1)$$

where, K is the number of mode components, $\{u_k\} = \{u_1, u_2, \dots, u_K\}$ and $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_K\}$ are components and corresponding central frequency respectively. $\sum_k := \sum_{k=1}^K$ is equivalent to the sum of all band components.

2.1.1. Solution of Variational Problem

By introducing the quadratic penalty factor α and the Lagrange multiplication operator $\lambda(t)$, the constrained variational problem is transformed into a non-constrained variational problem. The quadratic penalty factor can guarantee the signal reconstruction accuracy under the condition of Gaussian noise, and the Lagrange multiplier emphasizes the strictness of constraints.

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) := & \alpha \sum_k \|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ & + \|f(t) - \sum_k u_k(t)\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \end{aligned} \quad (2)$$

Then, the $\{u_k^{n+1}\}$, $\{\omega_k^{n+1}\}$ and λ^{n+1} are updated alternately by using the alternate direction method of multipliers (ADMM). By seeking the 'saddle point' of the extended Lagrange expression, that is, meeting the stop condition of iteration, mutually independent frequency band components $\{u_k\}$ are finally concluded. The decomposition process of the variational model is summarized in Algorithm 1.

Algorithm 1: Complete optimization of VMD

Initialize: $\{\hat{u}_k^1\}, \{\hat{\omega}_k^1\}, \hat{\lambda}^1, n \leftarrow 0$

repeat:

$n \leftarrow n + 1$

for $k = 1 : K$ **do**

Update \hat{u}_k for all $\omega \geq 0$:

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^2)} \quad (3)$$

Update ω_k :

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (4)$$

end for

Dual ascent for all $\omega \geq 0$:

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (5)$$

until convergence: $\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2 / \|\hat{u}_k^n\|_2^2 < \epsilon$ (6)

More detailed description of the VMD algorithm can refer to Ref. [19].

2.1.2. Parameter Influence Analysis

The process of solving the variational model shows that the performance of VMD is closely related to the intrinsic parameters [34], such as the total number of modes K , the quadratic penalty α , the update step τ , and the convergence fault tolerance threshold ϵ .

The influence of each parameter on the decomposition performance of VMD is analyzed as follows.

The performance of VMD is very sensitive to the value of K . If K is set too small, the signal will be under-segmented and some components will be included in other modes. On the contrary, the high value of K will cause mode duplication and other problems.

Parameter α is related to the performance of suppressing noise interference. A large value of α may result in a narrow bandwidth of modal components, and some information may be lost in the original signal. An α value that is too small will lead to too large of a bandwidth of modal components, some components will be included in other modes or extra noise will be captured.

When the noise level of the signal is low, the Lagrangian multiplier can ensure the optimal convergence by choosing the appropriate value of τ ($\tau > 0$). Accordingly, when the noise level of the signal is high, the Lagrangian multiplier will seriously hinder the convergence of the algorithm if $\tau > 0$. Setting $\tau = 0$ can effectively turn off the Lagrangian multiplier to ensure the effective convergence of the algorithm.

The value of convergence tolerance ε will affect the reconstruction accuracy of VMD decomposition. The reconstruction error (RSE) can be controlled by reducing the convergence of stop criterion to a certain extent.

From the above analysis, the above four key parameters seriously affect the performance of VMD. In addition, the interaction between parameters will affect the algorithm's performance. Therefore, choosing the right combination of VMD parameters is the key factor to determine its performance.

2.2. Least Squares Mutual Information

Mutual information [35–37] is a nonparametric and nonlinear measure index in information theory that can quantitatively express the correlation between two random variables and is more accurate than the correlation coefficient method [38]. According to the principle of irrelevance and orthogonality equivalence between zero mean random signals, mutual information can measure the coupling degree between Intrinsic Mode Function (IMF) components and residual information obtained in VMD decomposition. In other words, mutual information can measure whether modal mixing occurs and the determine its degree. Mutual information is defined as follows:

$$MI = \frac{1}{2} \int \sum_{y=1}^c p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx \quad (7)$$

where $p(x, y)$ is the joint probability density function of signals x and y , $p(x)$ and $p(y)$ are marginal probabilities distribution, respectively. In Equation (7), the response of the logarithm function in mutual information to outliers fluctuates greatly, which affects the accuracy of estimation. Therefore, in order to overcome this problem, in this paper, the square loss mutual information is induced to replace the logarithmic function, so as to reduce the interference of outliers and obtain more accurate mutual information estimator. The definition of substitution is as follows:

$$MI = \frac{1}{2} \int \sum_{y=1}^n p(x)p(y) \left(\frac{p(x, y)}{p(x)p(y)} - 1 \right)^2 dx \quad (8)$$

Avoiding the calculation of joint probability $p(x, y)$, marginal probability $p(x)$, and marginal probability $p(y)$, the least square estimation method is introduced to calculate the mutual information of the square loss, and the combined density ratio function is learned directly. The density ratio function is defined as follows:

$$\omega(x, y) = \frac{p(x, y)}{p(x)p(y)} \quad (9)$$

By taking the Gaussian radial basis kernel (RBF) model related to the parameters, the density ratio function can be approximated as:

$$\omega_{\alpha}(x, y) = \sum_{j=1}^n \alpha_j \exp\left(-\frac{\|x_j - y_j\|_2^2}{2h^2}\right) \quad (10)$$

where $\alpha = (\alpha_1, \dots, \alpha_n)^T$ is the parameter vector. Gaussian radial basis kernel is selected as the basis function ψ . Then, the least square learning is performed for the parameter α corresponding to the minimum of the following $J(\alpha)$.

$$\begin{aligned} J(\alpha) &= \frac{1}{2} \int \sum_{y=1}^n (\omega_{\alpha}(x, y) - \omega(x, y))^2 p(x) p(y) dx \\ &= \frac{1}{2} \int \sum_{y=1}^n \alpha^T \psi(x, y) \psi(x, y)^T \alpha p(x) p(y) dx \\ &\quad - \int \sum_{y=1}^n \alpha^T \psi(x, y) p(x, y) dx + C \end{aligned} \quad (11)$$

where the third term $C = \frac{1}{2} \int \sum_{y=1}^n \omega(x, y) p(x, y) dx$ is a constant independent of parameters, which can be ignored in the calculation. By using the sample average approximation of the expected values contained in the first and second terms of Equation (11) and introducing the L^2 regularization term, the learning rule can be derived as:

$$\min_{\alpha} \left[\frac{1}{2} \alpha^T \hat{G} \alpha - \alpha^T \hat{h} + \frac{\lambda}{2} \|\alpha\|^2 \right] \quad (12)$$

$$\begin{cases} \hat{G} = \frac{1}{n^2} \sum_{i,i'} \psi(x_i, y_{i'}) \psi(x_i, y_{i'})^T \\ \hat{h} = \frac{1}{n} \sum_{i=1}^n \psi(x_i, y_i) \end{cases} \quad (13)$$

where \hat{G} is a matrix with $n \times n$ order and \hat{h} is a n dimensional vector. The learning rule is a convex quadratic form related to α , and its optimization problem is as follows:

$$\hat{\alpha} := \operatorname{argmin}_{\alpha} \left[\frac{1}{2} \alpha^T \hat{G} \alpha - \alpha^T \hat{h} + \frac{\lambda}{2} \|\alpha\|^2 \right] \quad (14)$$

The analytical solution can be obtained by taking the derivatives of Equation (14) and solving it equal to zero, namely,

$$\hat{\alpha} = (\hat{G} + \lambda I)^{-1} \hat{h} \quad (15)$$

where λ is the regularization parameter and I is the identity matrix. The density ratio estimator obtained by the above method is substituted into Equation (16), which is equivalent to the square loss mutual information.

$$\text{MI} = \frac{1}{2} \int \sum_{y=1}^n \omega(x, y) p(x, y) dx - \frac{1}{2} \quad (16)$$

Then, the least squares mutual information (LSMI) estimation is obtained as follows:

$$\text{LSMI} = \frac{1}{2} \hat{h}^T (\hat{G} + \lambda I)^{-1} \hat{h} - \frac{1}{2} \quad (17)$$

where the regularization parameter λ and the parameters contained in the basis function ψ can be determined by the optimization algorithm related to rule J .

2.3. Refined Composite Multi-Scale Dispersion (RCMDE)

Information entropy [39] is an indicator that measures the uncertainty of information quantity, which represents the average uncertainty of a signal. In information theory, information entropy is used to measure the amount of information. The larger the information entropy, the greater the uncertainty and the complexity of the signal present. At present,

information entropy is widely used in the field of mechanical fault diagnosis and medical diagnosis, and has achieved fruitful research results. The commonly used information entropy primarily includes approximate entropy, sample entropy, permutation entropy and so on. Many achievements have been made in the application of these techniques in signal nonlinear feature extraction. Sample entropy is characterized by large computation and slow computation, and permutation entropy does not take into account the differences between vibration amplitudes. Rostaghi and Azami [40] proposed a new method to measure the complexity of time series, namely dispersion entropy (DE), which solves the shortage of sample entropy and permutation entropy to some extent. Furthermore, Azami [41] proposed the fine composite multi-scale dispersion entropy (RCMDE) method, which has the characteristics of good stability in multi-scale process.

2.3.1. Dispersion Entropy

(1) The normal distribution function (NDF) is used to map the time series $x = \{x_j | j = 1, 2, \dots, N\}$ to $y = \{y_j | j = 1, 2, \dots, N\}$ $y \in (0, 1)$. N denotes the length of the sequence. The mapping function is as follows:

$$y_j = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x_j} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \quad (18)$$

where μ and σ represent the expectation and standard deviation respectively.

(2) Map y to integers in the range $[1, C]$ using a linear transformation.

$$z_j^c = \text{round}(cy_j + 0.5) \quad (19)$$

where $\text{round}(\cdot)$ is the rounding function, and z_j^c is the j -th element of classification sequence z^c .

(3) Calculate the embedding vector as follows:

$$z_i^{m,c} = [z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c], i = 1, 2, \dots, N - (m-1)d \quad (20)$$

$$z_i^c = v_0, z_{i+d}^c = v_1, z_{i+(m-1)d}^c = v_{m-1} \quad (21)$$

where m is the embedded dimension. d is the time delay.

(4) The dispersion pattern corresponding to each $z_i^{m,c}$ is $r_{v_0, v_1 \dots v_{m-1}} \cdot z_i^{m,c}$. It contains m digits in total, and each digit has c values. Therefore, the total number of dispersion modes of $z_i^{m,c}$ is c^m .

(5) The probability of each dispersion mode t can be defined as:

$$p(r_{v_0, v_1 \dots v_{m-1}}) = \frac{N(r_{v_0, v_1 \dots v_{m-1}})}{N - (m-1)d} \quad (22)$$

where $N(r_{v_0, v_1 \dots v_{m-1}})$ is the number of dispersion modes corresponding to $z_i^{m,c}$.

According to the definition of information entropy, the dispersion entropy can be expressed as:

$$DE(x, m, c, d) = -\sum_{r=1}^{c^m} p(r_{v_0, v_1 \dots v_{m-1}}) \ln(p(r_{v_0, v_1 \dots v_{m-1}})) \quad (23)$$

where r represents the type of dispersion mode corresponding to $z_i^{m,c}$. From the calculation method of dispersion entropy, it can be seen that when the probabilities of all dispersion modes are equal, the dispersion entropy has the maximum value $\ln(c^m)$. The greater the dispersion entropy is, the higher the unpredictable degree of time series denotes, and vice versa.

2.3.2. Multi-Scale Dispersion Entropy

The multi-scale dispersion entropy (MDE) proposed on the basis of dispersion entropy can reflect the complexity of time series at different scales. The calculation method of multi-scale dispersion entropy is as follows:

Firstly, The original signal $u = \{u_1, u_2, \dots, u_L\}$ with length L is roughened to obtain N sequences with length τ , and the coarse-grained signal is obtained by calculating the average value of each sequence.

$$x_j^\tau = \frac{1}{\tau} \sum_{b=(j-1)\tau+1}^{j\tau} u_b, 1 \leq j \leq \left\lfloor \frac{L}{\tau} \right\rfloor \tag{24}$$

where $\lfloor L/\tau \rfloor$ represents the length of each coarse-grained time series.

Secondly, calculate the dispersion entropy $DE(x^\tau, m, c, d)$ of coarse-grained signal under each scale factor τ .

Finally, the MDE index can be obtained, as shown in the following:

$$MDE(x, m, c, d, \tau) = \frac{1}{\tau} \sum_{i=1}^{\tau} DE(x^\tau, m, c, d) \tag{25}$$

2.3.3. Refined Composite Multi-Scale Dispersion Entropy (RCMDE)

For time series with different scales, parameter τ actually corresponds to different starting points of coarsening process. RCMDE value is defined as the average value of dispersion entropy of coarsening sequence. The k -th coarsening sequence of signal $u = \{u_1, u_2, \dots, u_L\}$ is:

$$x_k^\tau = \{x_1^\tau, x_2^\tau, \dots, x_L^\tau\} \tag{26}$$

$$x_k^\tau = \frac{1}{\tau} \sum_{b=k+\tau(j-1)}^{k+\tau j-1} u_b, 1 \leq j \leq \left\lfloor \frac{L}{\tau} \right\rfloor, 1 \leq k \leq \tau \tag{27}$$

The RCMDE value under scale τ . is calculated as follows:

$$E_{RCMD}(x, m, c, d, \tau) = - \sum_{r=1}^{c^m} \bar{p}(r_{v_0 v_1 v \dots v_{m-1}}) \ln(\bar{p}(r_{v_0 v_1 v \dots v_{m-1}})) \tag{28}$$

where $E_{RCMD}(x, m, c, d, \tau)$ is the RCMDE value under scale τ . $\bar{p}(r_{v_0 v_1 v \dots v_{m-1}})$ is the average value of dispersion mode probability corresponding to coarsening sequence, as follows:

$$\bar{p}(r_{v_0 v_1 v \dots v_{m-1}}) = \frac{1}{\tau} \sum_{k=1}^{\tau} p_k^\tau \tag{29}$$

where p_k^τ is the probability of the dispersion pattern corresponding to the k -th coarsening sequence under scale τ .

2.4. Efficient Signal Evaluation Index

Generally speaking, the observation signal consists of effective signal and noise components. To verify the noise robustness of the proposed method, a kind of metric, namely, the signal noise ratio (SNR), was applied and may be defined as follows:

$$SNR = 10 \log \left\{ \frac{\sum_{i=1}^L x^2(i)}{\sum_{i=1}^L ((x(i) - \hat{x}(i))^2)} \right\} \tag{30}$$

where $x(i)$ is the observed signal, $\hat{x}(i)$ is the mean value of $x(i)$, and L denotes the length of $x(i)$.

2.5. Analysis of Intrinsic Mode Function

At present, many signal decomposition methods take the orthogonality between modal components as the stop condition of decomposition, including many improved VMD algorithms. However, the orthogonality between IMF components could not guarantee the unity of IMF component characteristics. According to the decomposition principle of standard VMD, each IMF component has a single characteristic. A test was executed to demonstrate the orthogonality between IMF components and its own singleness.

As shown in Figure 1a, three harmonic signals are selected, and their corresponding frequencies are 2 Hz, 24 Hz and 288 Hz, respectively. The three harmonics are linearly superimposed to obtain the mixed signal, as shown in Figure 1b. Then, the original VMD is used to decompose the mixed signal. The preset parameters of VMD are $K = 2$, $\alpha = 2000$. The decomposition result is shown in Figure 1c. The LSMI between the two IMF components (IMF1 and IMF2) in Figure 1c is 1.0×10^{-6} , which is approximately orthogonal. However, it can be clearly observed that the IMF1 component contains two harmonic components (2 Hz and 24 Hz). It is not a single modal function. Finally, the IMF1 and IMF2 components are combined to reconstruct the representation signal, and the results are shown in Figure 1d. Compared with the original mixed signal, the reconstructed error is 1.0×10^{-10} . Experimental results indicate that, although the orthogonal coefficient and reconstruction error index between IMF components are very small, it cannot fully ensure the unity of IMF components. On the one hand, it shows that the ideal built-in parameters of VMD have a great influence on the decomposition. On the other hand, VMD is needed to deal with the under decomposition and over decomposition of components adaptively. Therefore, this paper proposes a decomposition method based on the binary tree mechanism as shown in Figure 2, which can effectively ensure the unity of the IMF component.

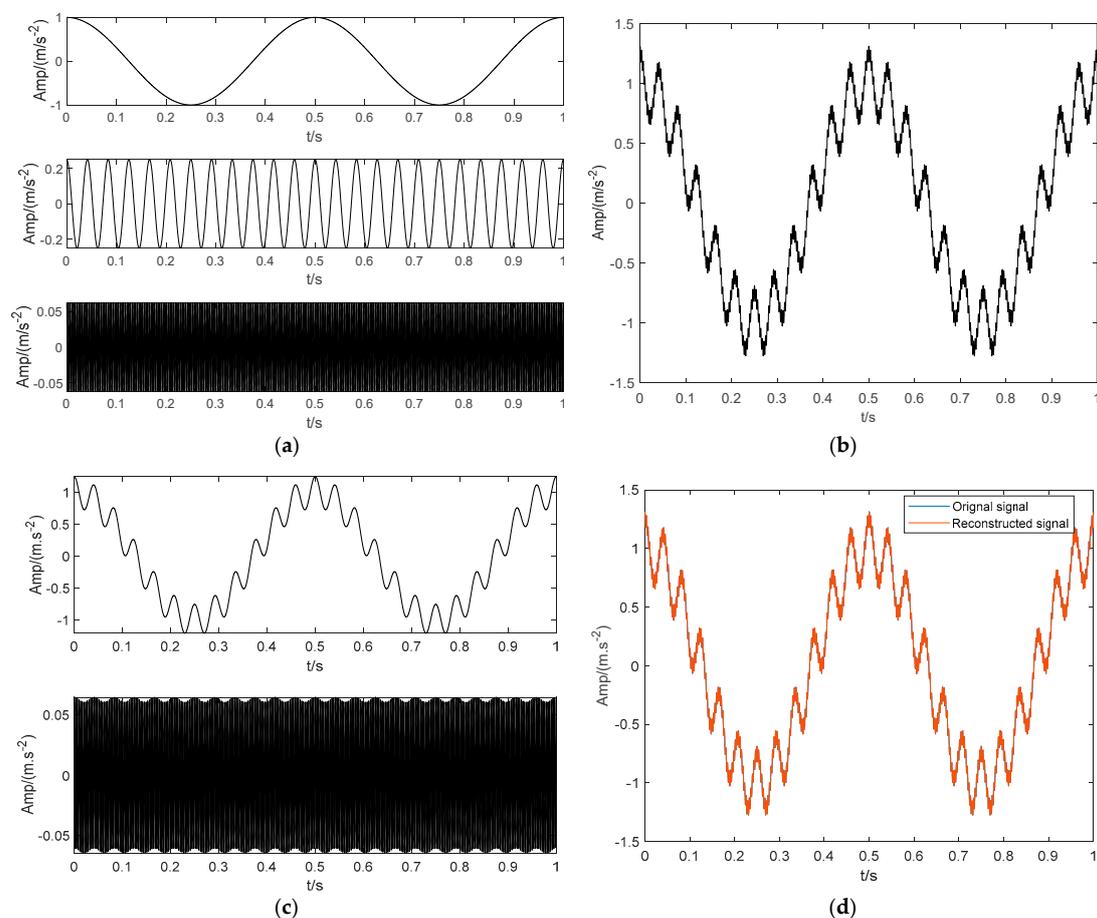


Figure 1. The singularity analysis of IMF components. (a) Sinusoidal harmonics. (b) Mixed signal. (c) Modes. (d) The reconstructed signal.

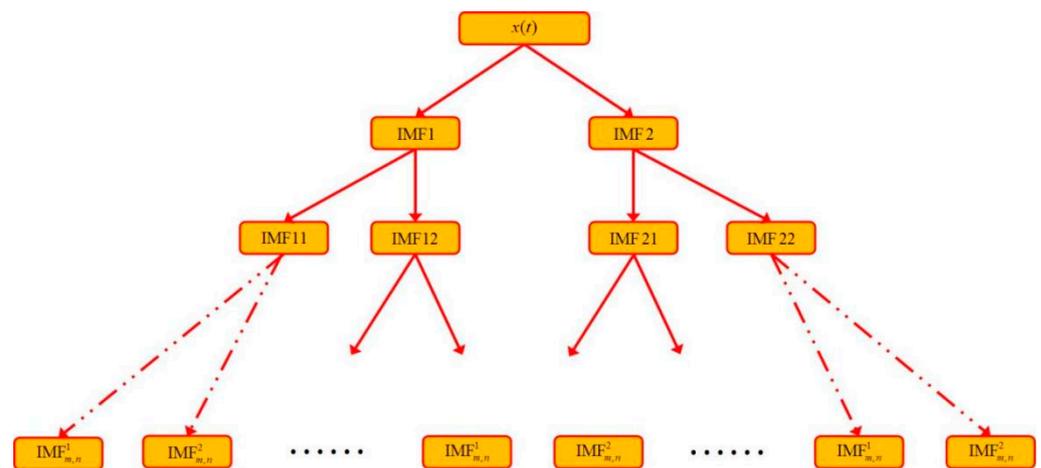


Figure 2. Diagram of binary tree decomposition.

3. Proposed Algorithm Framework

In this research work, an empirical VMD algorithm based on a binary tree model (BT-EVMD) is presented, which can effectively solve the problem of selecting the key parameters of VMD and make the decomposition process of VMD completely an adaptive process. The detailed procedure is as follows:

Step 1: Calculate the signal to noise ratio (SNR) and $RCMDE$ of the original non-stationary signal $x(t)$, and initialize key parameters of VMD as $K = 2$, $\alpha = RCMDE \times \text{round}((f_s/2) * \log(K))$ (f_s is the sampling frequency and $\text{round}(\cdot)$ is the rounding function), $\tau = \text{SNR}$ and $\varepsilon = 1 \times 10^{-7}$. The signal is decomposed by VMD to obtain two IMF components, called IMF1 and IMF2.

Step 2: Initializes the kernel parameters of the Gaussian radial basis function of LSMI. Preset the threshold δ of LSMI estimator $LMSI_E$ and the threshold ρ of reconstruction error (RSE, the error between the sum of decomposed modes and input signal).

Step 3: Compute the LSMI of IMF1 and IMF2 (if $LMSI_E = 0$, there is no similar information between IMF components, if $LMSI_E = 1$, the information between IMF components is exactly the same, $0 \leq LMSI_E \leq 1$). Determine whether $LMSI_E$ is greater than the threshold. If yes, end the decomposition. Else, calculate the reconstruction error ρ . If $\rho > 1 \times 10^{-7}$, terminate the process. Else, take the decomposed IMF components as two new signals, and repeat the Step 1 to continue the iteration.

Step 4: The LSMI between each IMF is computed. A new IMF component is obtained by adding the IMFs that meet $LMSI_E > \delta$.

After the decomposition of BT-EVMD, the multi-component and non-stationary signal can be adaptively decomposed into several sub-components.

The schematic diagram of decomposition is shown in Figure 2.

4. Experiment Validations

4.1. Simulation Analysis

In order to more clearly verify the effectiveness of the proposed method. This section applies the BT-EVMD algorithm to a typical analog signal and compares its performance with other decomposition algorithms. The signal is similar to those in references [15,29], but its composition is more complex. It contains high-frequency weak signal with intermittent time, periodic impulse signal and combined components with similar frequency. The signal can be calculated via Equation (31).

$$\begin{cases} x(t) = x_1(t) + x_2(t) + x_3(t) + x_4(t) \\ x_1(t) = \sin(2\pi f_1 t) \\ x_2(t) = 0.6 \times \sin(2\pi f_2 t) \\ x_3(t) = 0.4 \times \sin(2\pi f_3 t) \\ x_4(t) = \sum_i A_i h(t - iT - v_i) \end{cases} \quad t \in [0.1, 0.2] \cup [0.8, 0.9] \quad (31)$$

The simulated signal $x(t)$ is composed of three sinusoidal signals $x_1(t)$, $x_2(t)$ and $x_3(t)$ with different center frequencies, and a high frequency intermittent signal $x_4(t)$, as shown in Figure 3. Here, the sampling frequencies are $f_1 = 20$ Hz, $f_2 = 35$ Hz and $f_3 = 200$ Hz, respectively. $x_4(t)$ is a sinusoidal signal with periodic pulse attenuation and a frequency of 8 Hz, and is formed as:

$$\begin{cases} x_4(t) = \sum_i A_i h(t - iT - v_i) \\ h(t) = e^{-Ct} \sin(2\pi f_n t) \\ A_i = 1 + A_0 \sin(2\pi f_r t) \end{cases} \quad (32)$$

where A_0 is the initial amplitude of the impulse signal and A_i is the amplitude of shock signal after the i -th attenuation. T is the cyclic period, v_i is the random tiny slippage during each T , usually considered as $0.01T$ – $0.02T$. C denotes the damping coefficient. f_r is the rotation frequency of the simulated transmission shaft and f_n is the resonance frequency. Here, the simulated signal parameters are set as follow: $C = 750$, $f_r = 8$ Hz, $f_n = 3000$ Hz, $A_0 = 0.5$.

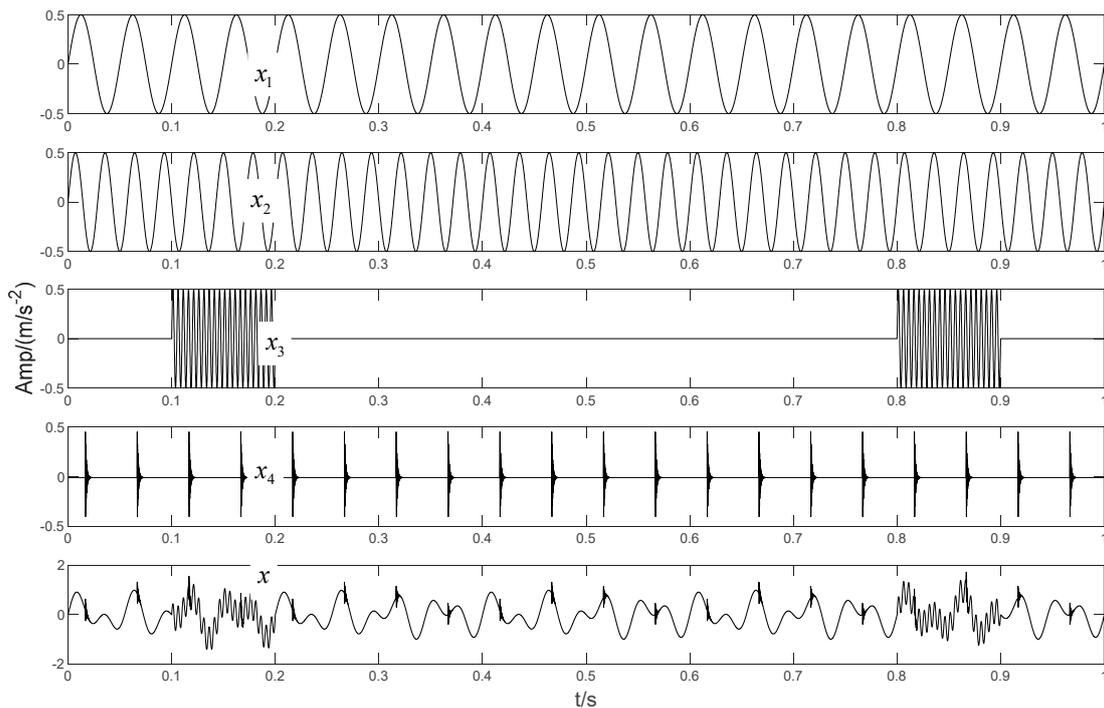


Figure 3. The simulated signal.

In order to illustrate the process of the algorithm, the binary tree empirical VMD algorithm is used to decompose the simulation signal presented in Figure 3, and the experimental results are shown in Figures 4–11. The simulated signal is firstly decomposed into two modes: IMF1 and IMF2. According to experimental experience, the threshold of decomposition termination is set as $\delta = 0.1$ and $\rho = 0.1$, respectively. Then, the LSMI and RSE between IMF1 and IMF2 are computed, and the results are shown in Table 1. The parameters listed in Table 1 are LSMI [IMF1, IMF2] = 0.0329 and RSE = 0.0328, respectively. The result of the comparison indicates that the first VMD decomposition of the simulated signal does not meet the stop condition. Therefore, components IMF1 and IMF2 need to be further decomposed, that is, the second-layer VMD decomposition. The decomposition

results are shown in Figures 5 and 6, in which the waveforms of components IMF12 and IMF22 are actually very similar. In order to quantify the similarity between the two components, the LSMI is computed as $\text{LSMI} [\text{IMF12}, \text{IMF22}] = 0.8414$, which satisfies the stop condition. At this time, the components IMF12 and IMF22 do not need to be further decomposed, and the first component IMF1' is obtained by adding the modes IMF12 and IMF22, as shown in Figure 7. The remaining components will then be decomposed by VMD. The third level decomposition results are shown in Figures 8 and 9. The component IMF11 is decomposed into sub-components IMF111 and IMF112, and the sub-modes IMF211 and IMF212 are obtained by decomposing the IMF21. The corresponding computation parameters of RSE and LSMI between sub-components can be found in Table 1. The Components IMF112, IMF211, and IMF212 are selected as their values of LSMI satisfy the stop condition. Thus, IMF112, IMF211 and IMF212 are added to construct a new mode IMF2', and the resultant signal is shown in Figure 10. Then, the remnant component IMF111 is decomposed, and the decomposition result is shown in Figure 11. It can be seen from Figure 11 that the decomposed components IMF1111 and IMF1112 are two single wave modes, so the decomposition is stopped and the two components are regarded as the third component IMF3' and the fourth component IMF4' respectively. Finally, four modal components are obtained, as shown in Figure 12. It can be observed from the experimental results that the number of modal components obtained by the proposed method is the same as that of the original simulation signal, and the waveform features are highly similar. To further explain the quantitative relationship between each IMF component obtained and its corresponding original simulated signal, the LSMI is computed between them. The specification of the LSMI is listed in Table 1. From the measured data, except that the LSMI between the component IMF4' and the simulated signal x_4 is slightly less than 1, other IMF components are almost the same as their corresponding simulated signals. Experimental results indicate that the proposed empirical VMD algorithm is adaptive and effective.

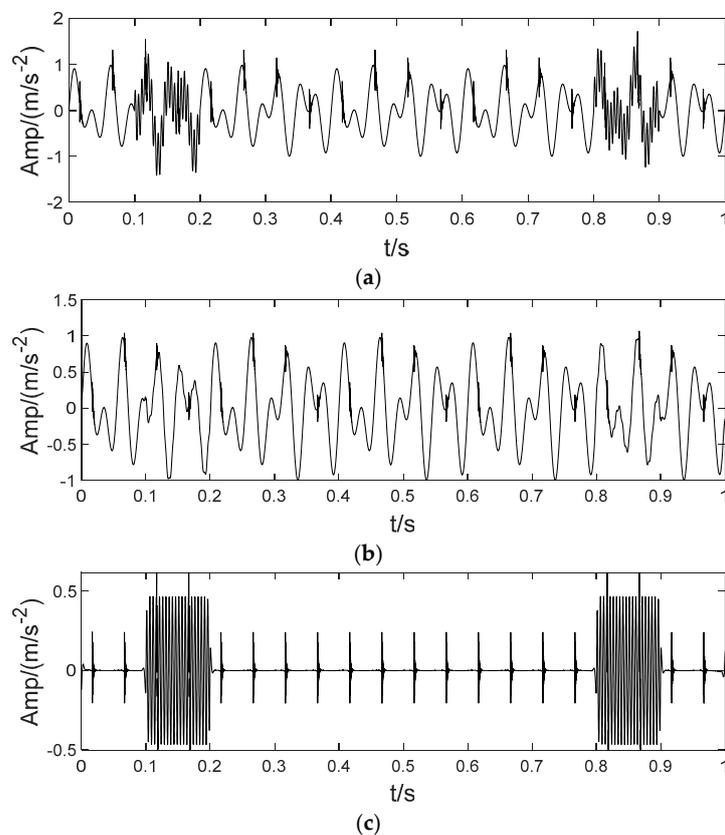


Figure 4. The first decomposition of empirical VMD based on binary tree. (a) Original simulation signal. (b) denotes the IMF1 component obtained by the first decomposition and (c) denotes IMF2.

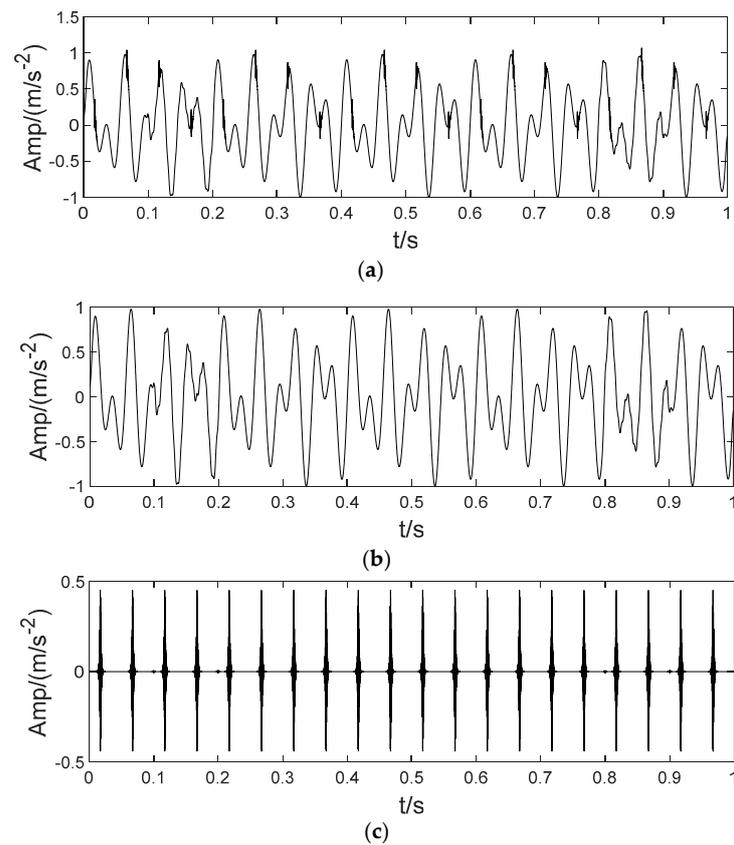


Figure 5. The second decomposition of IMF1. (a) The IMF1, (b) The IMF11 obtained by decomposition and (c) IMF12.

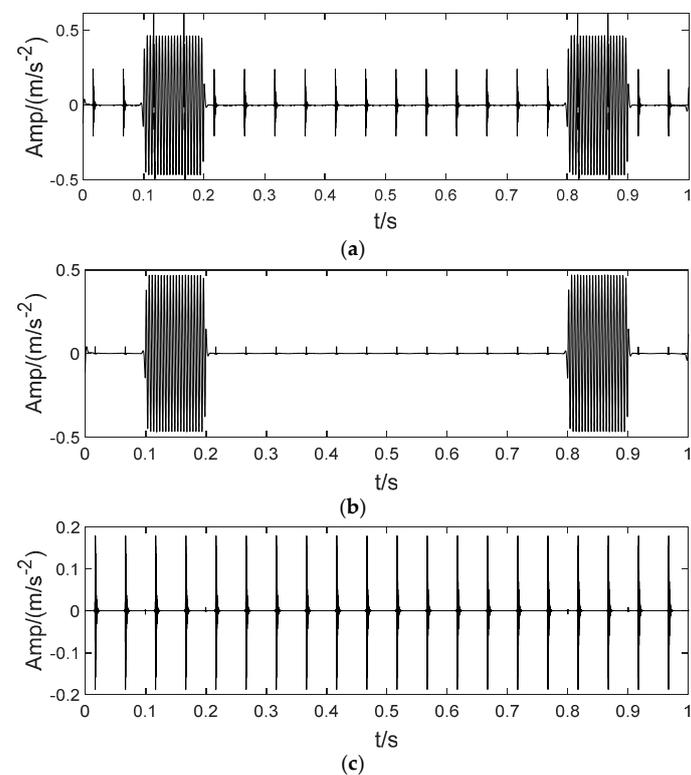


Figure 6. The second decomposition of IMF2. (a) IMF2, (b) The IMF21 obtained by decomposition and (c) IMF22.

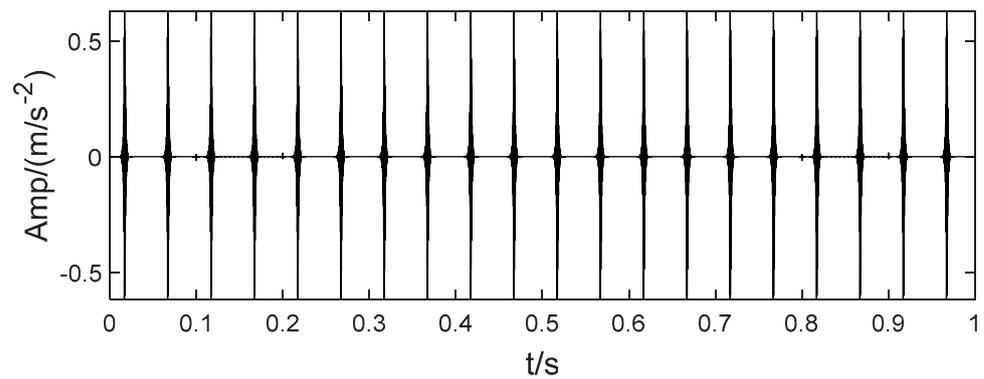
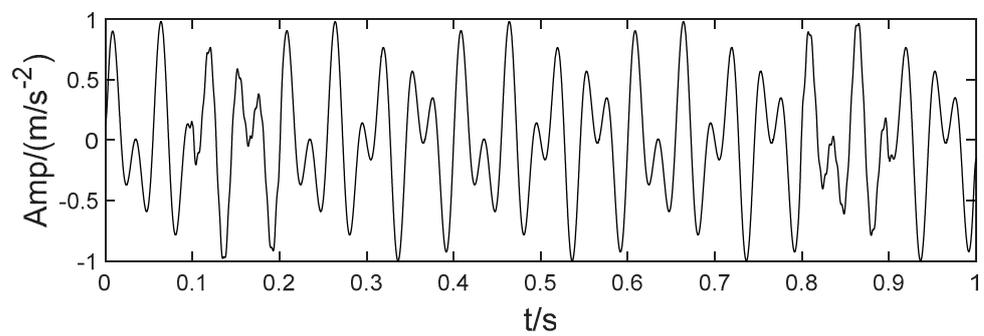
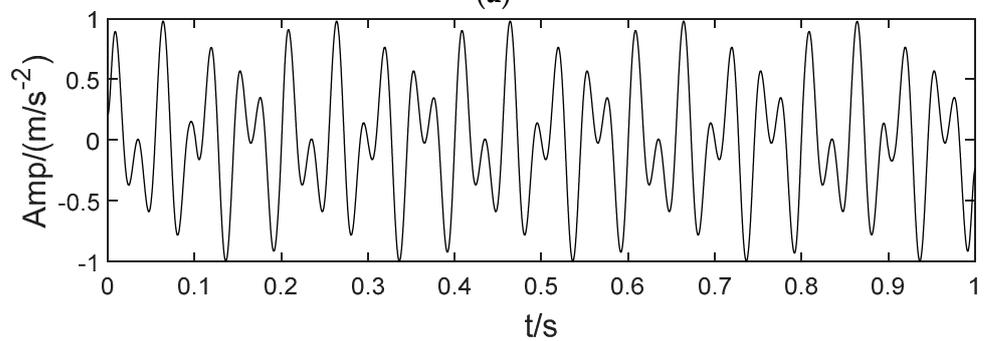


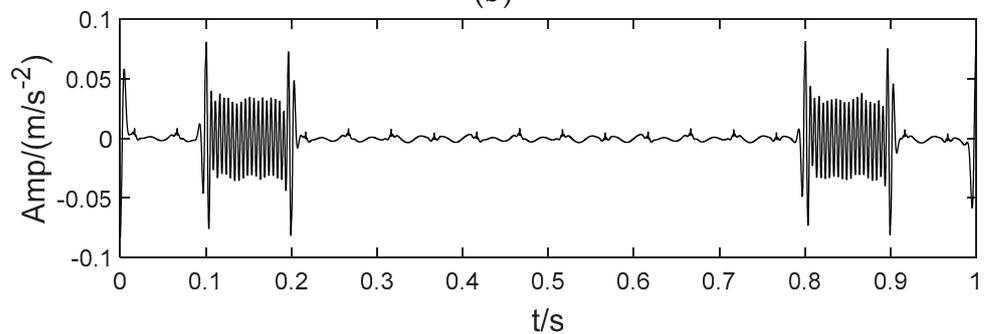
Figure 7. The first obtained IMF component.



(a)



(b)



(c)

Figure 8. The Third decomposition of IMF11. (a) IMF11, (b) The IMF111 obtained by decomposition and (c) IMF112.

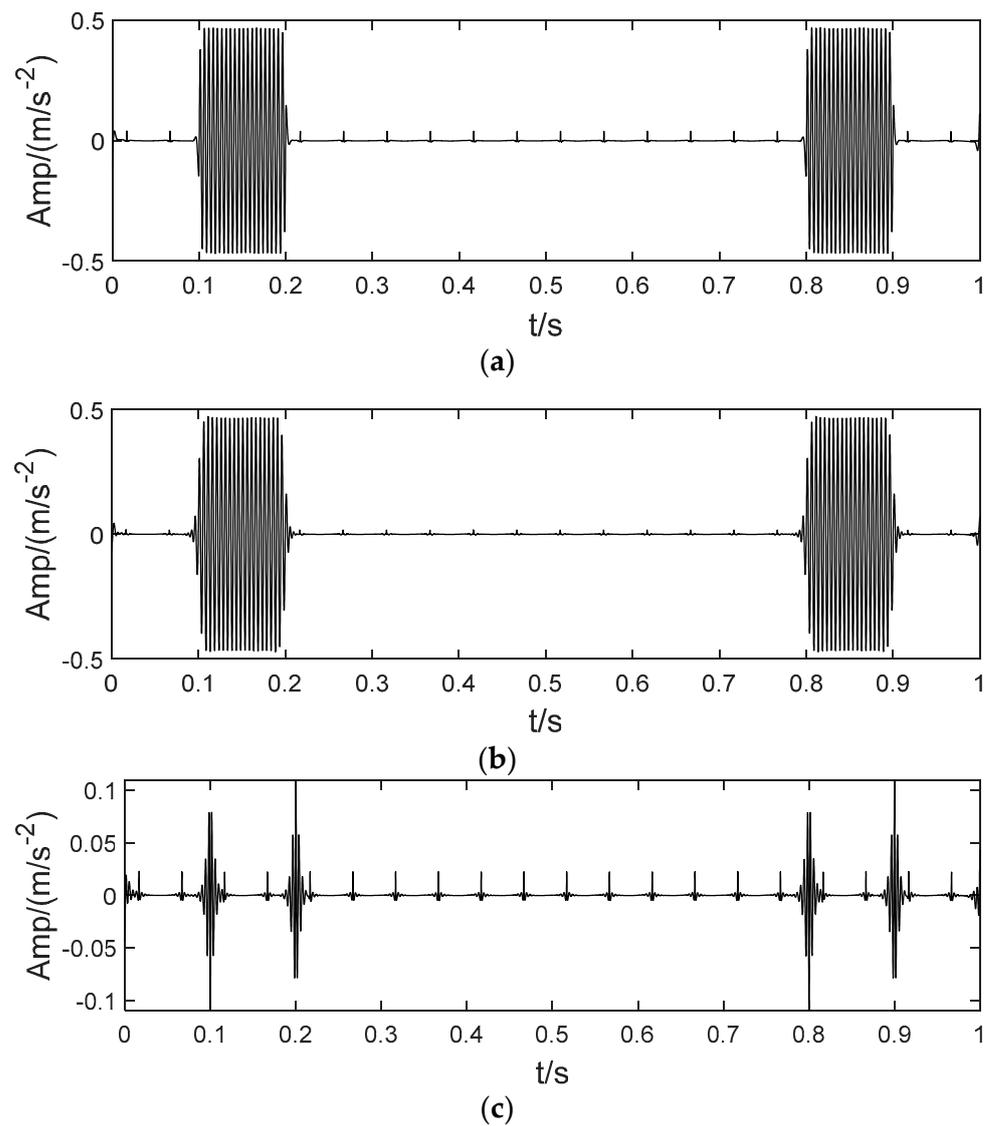


Figure 9. The third decomposition result. (a) IMF21, (b) IMF211 and (c) IMF212.

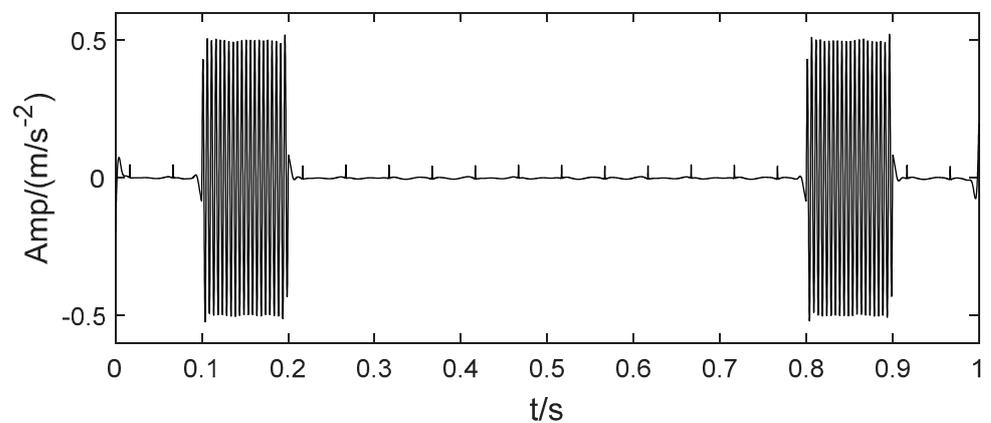


Figure 10. The second IMF component obtained by decomposition.

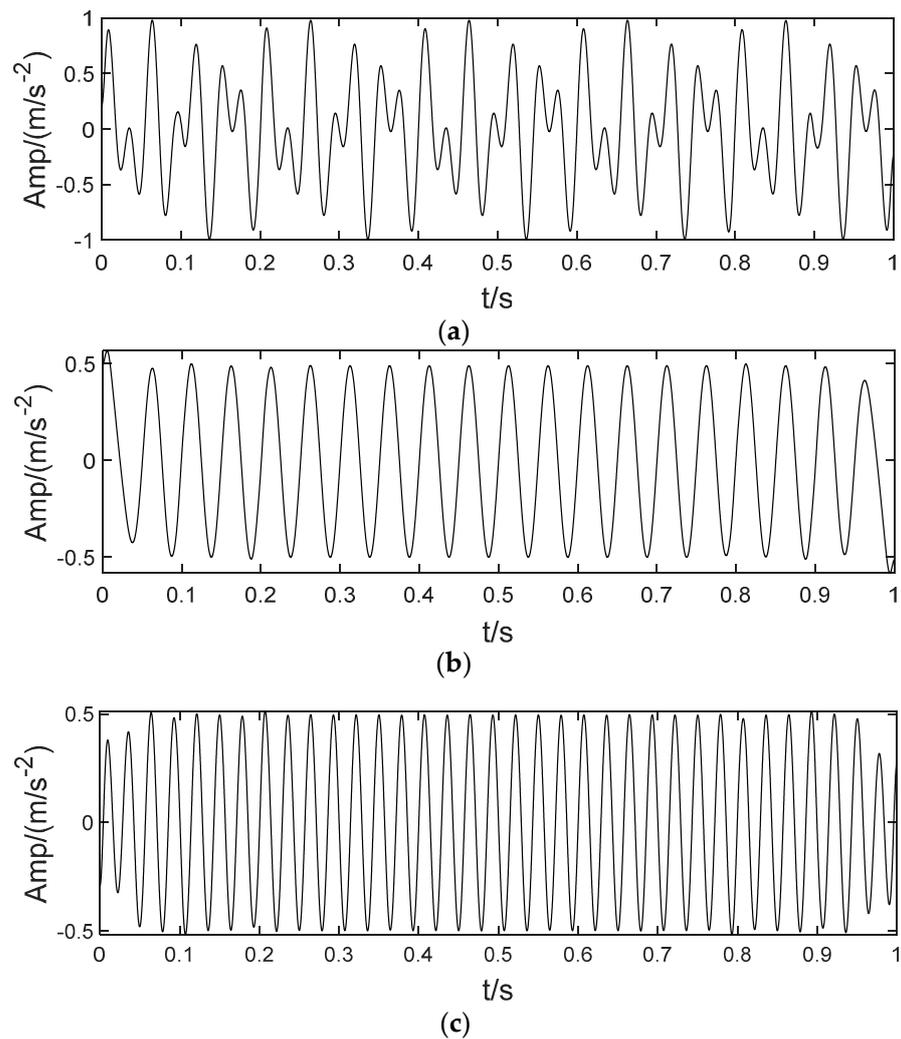


Figure 11. The fourth decomposition result. (a) IMF111, (b) IMF1111 and (c) IMF1112.

Table 1. The LSMI and RSE obtained by BT-EVMD.

Decomposition Hierarchy	Modes	LSMI Index	Reconstruction Error
The first layer	[IMF1, IMF2]	LSMI [IMF1, IMF2] = 0.0329	RSE = 0.0328
The second layer	[IMF11, IMF12] [IMF21, IMF22]	LSMI [IMF11, IMF12] = 0.0002	RSE = 0.0414
		LSMI [IMF21, IMF22] = 0.0016	
		LSMI [IMF11, IMF21] = 0.0028	
		LSMI [IMF11, IMF22] = 0.0004	
		LSMI [IMF12, IMF21] = 0.0012	
The third layer	[IMF111, IMF112] [IMF211, IMF212]	LSMI [IMF111, IMF112] = 0.1441	RSE = 0.0327
		LSMI [IMF211, IMF212] = 0.0849	
		LSMI [IMF111, IMF211] = 0.0052	
		LSMI [IMF111, IMF212] = 0.0228	
		LSMI [IMF112, IMF211] = 0.7181	
The fourth layer	[IMF1111, IMF1112]	LSMI [IMF1111, IMF1112] = 0.1325	RSE = 0.0112
		LSMI [IMF1111, IMF1112] = 0.0185	

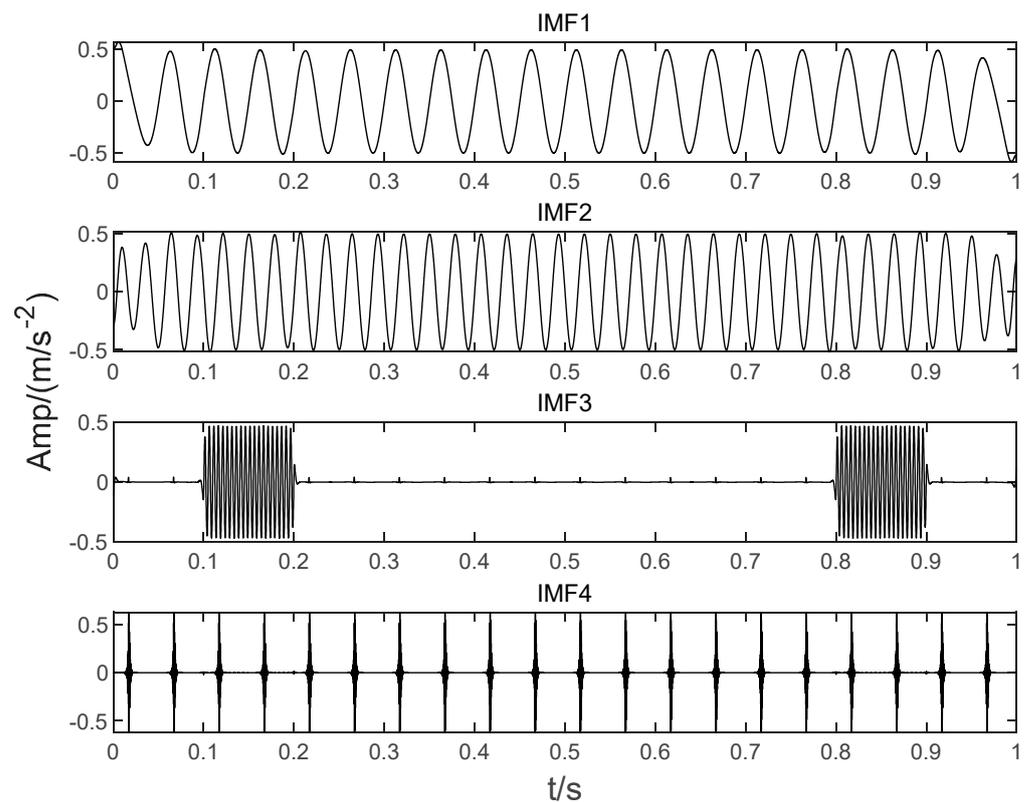


Figure 12. Decomposition of the mixed signal by proposed BT-EVMD method.

In order to verify the superiority of the proposed algorithm, this paper continues to utilize common signal processing methods such as LDM, ITD, CEEMDAN and SVMMD to process the simulation signal. The decomposition mode results are shown in Figure 13. Figure 13a denotes the result of the LMD decomposition. The signal is decomposed into four independent components, the number of which corresponds to that of the simulation signal without obvious aliasing, but the components corresponding to the simulation signal have no similarity at all. Figure 13b is the result of the ITD decomposition. The number of components decomposed by this method is inconsistent with that of the simulation signal, and there is no similarity with the original components. The number of components decomposed by the CEEMDAN method far exceeds the number of original signals, and there is a certain amount of mode mixing. Some waveform trends have certain similarities with the components of the original signal (such as IMF1–IMF3), but there are many irrelevant terms due to the over decomposition phenomenon, and the results are shown in Figure 13c. Figure 13d is the decomposition result of SVMMD. Comparing the waveform of the decomposed components with the original components, it can be observed that the decomposed IMF1 is a mixed signal of the $x_1(t)$ and $x_2(t)$ in the simulation signal, that is, the first IMF component is not a single component. In addition, the waveforms of the IMF3 and the IMF4 are similar, indicating that over decomposition occurs. Therefore, comparing the decomposition results of Figures 12 and 13, it can be observed that the components obtained based on the method proposed in this paper are not only consistent with the simulated signal in quantity, but also the characterization of the components is almost the same as that of the simulated signal. Compared with the above methods, the experimental results indicate that the performance of the proposed method has obvious advantages.

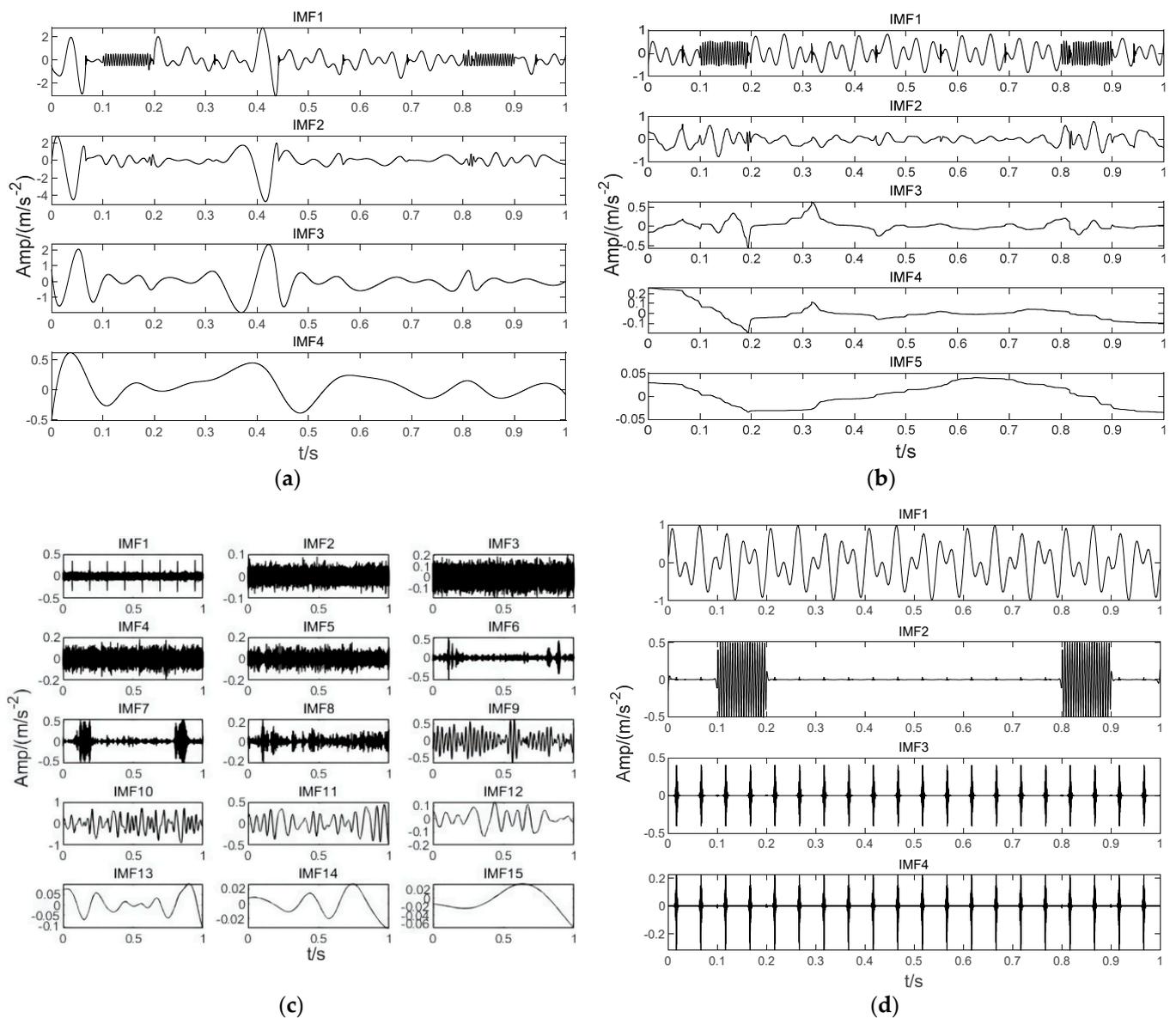


Figure 13. Performance evaluation of other state-of-art methods. (a) The components obtained by LMD. (b) The components obtained by ITD. (c) The components obtained by CEEMDAN. (d) The components obtained by SVMd.

4.2. Analysis of Test Data

In order to further verify the effectiveness of the proposed method in actual signal analysis, in this paper, the bearing fault signal of Western Reserve University was applied as the verification data [42]. As shown in Figure 14, the bearing fault test rig mainly consists of an induction motor (2 HP), a torque transducer, a dynamometer, and several units. Three accelerometers are mounted on the housing at 3, 6 and 12 o'clock positions of the motor drive ports. The vibration signals of rolling bearings were collected by a 16-channel data recorder. The platform used electrical discharge machining to arrange a single point of failure on the bearing (SKF6205).

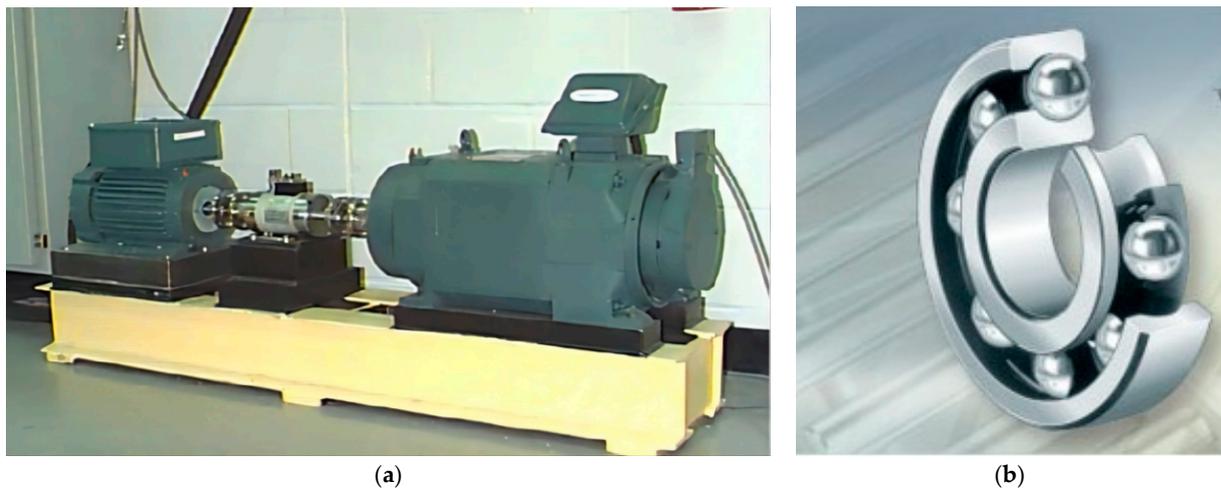


Figure 14. Apparatus & procedures: (a) Experimental platform. (b) Skf6205 deep groove ball bearing.

In the experiment, the fault data of the inner ring and outer ring of the drive end bearing were collected for analysis. According to Ref. [43], the specification parameters are listed in Table 2.

Table 2. The test platform parameters.

Bearing Fault Type	Parameters				
	Speed (r/min)	Characteristic Frequency (Hz)	Load (HP)	Data Length (L)	Sampling Frequency (kHz)
Inner ring fault	1797	162.1852	0	12,000	12
Outer ring fault	1796	107.3647	0	12,000	12

According to the method applied in [24], the inner and outer ring bearing fault signals are superimposed to obtain a mixed fault signal, as shown in Figure 15. Figure 15a shows the time domain waveforms of inner and outer ring bearing fault signals and the mixed signal. Figure 15b corresponds to their envelope spectrum analysis, respectively. The inner ring fault characteristic frequency f_{inner} , outer ring fault characteristic frequency f_{outer} and its frequency doubling items can be observed in the spectrum of the mixed signal. Then, the mixed signal is analyzed by using SVM, CEEMDAN, LDM, ITD and the proposed BT-EVMD methods. The decomposition results are shown in Figures 16–20, respectively.

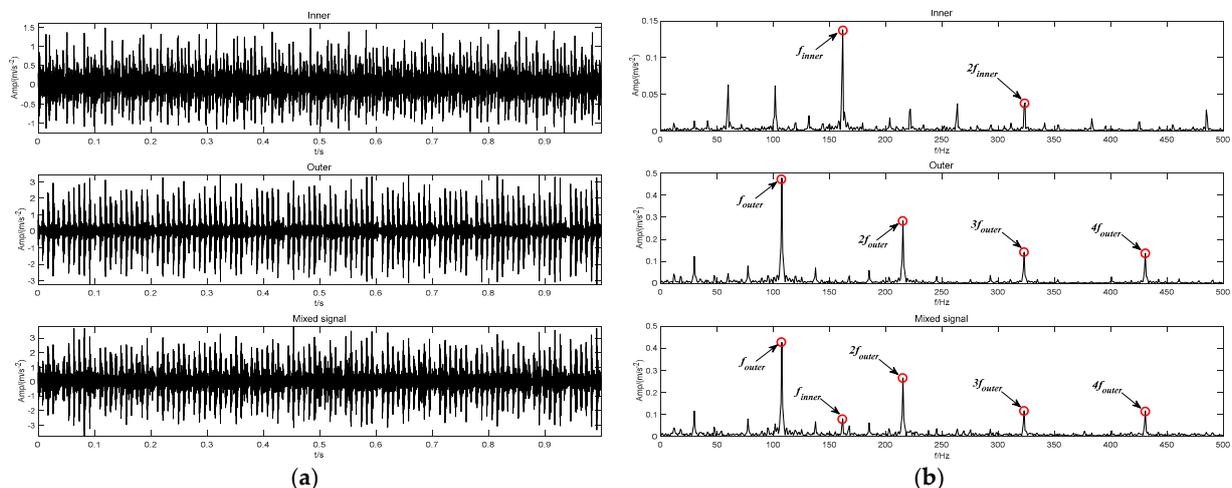


Figure 15. Original fault vibration signal: (a) Time-domain fault waveform. (b) Envelope spectrum.

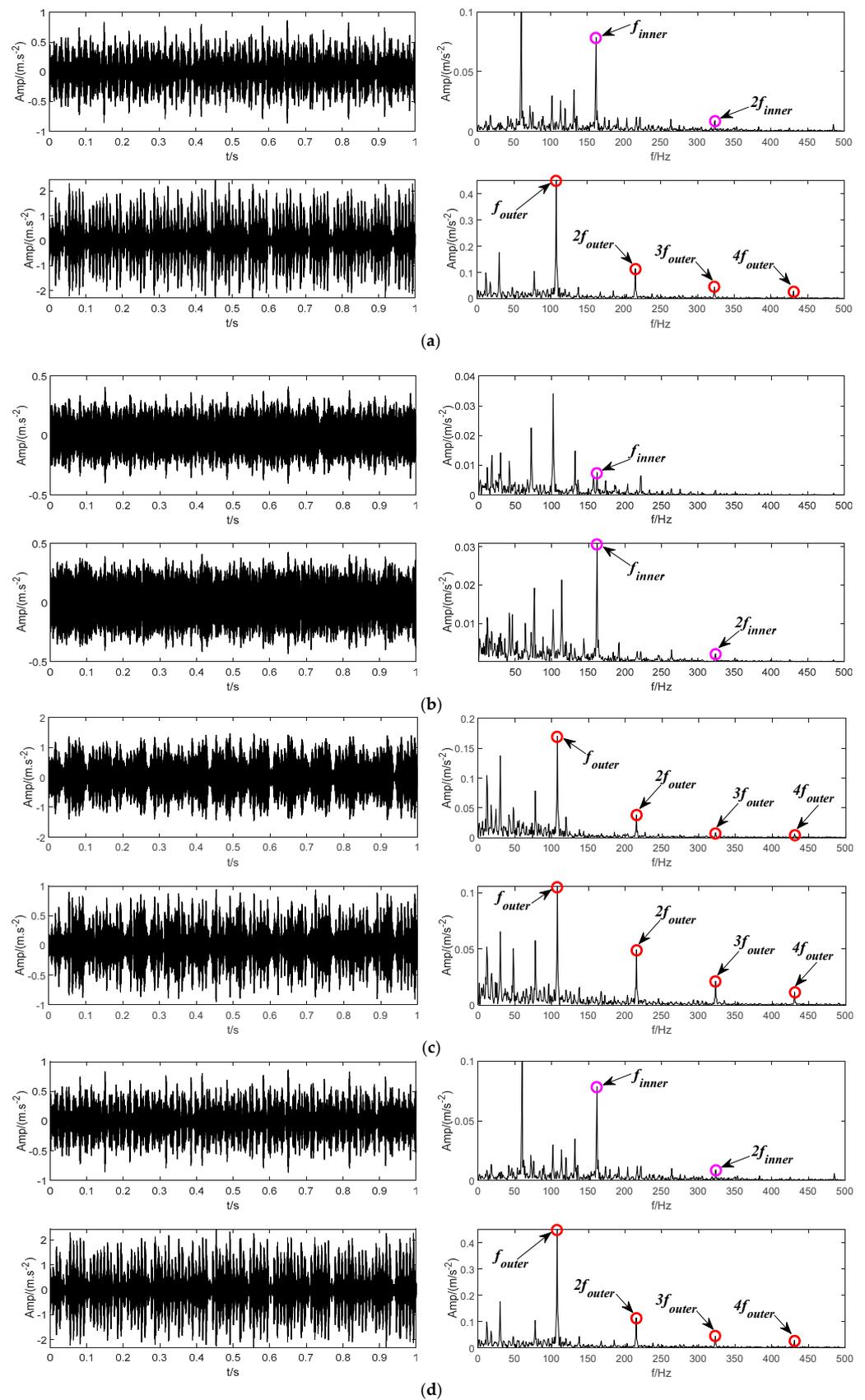


Figure 16. The modes and the corresponding envelope spectrum obtained by the proposed BT-EVMD method. (a) The first decomposition. (b) The second decomposition. (c) The third decomposition. (d) The final result.

The decomposition result of the proposed BT-EVMD is shown in Figure 16. The relevant decomposition parameters are shown in Table 3. The decomposition results of the first layer are shown in Figure 16a, in which the left part is the two IMF components obtained, and the right part is the envelope spectrums corresponding to the IMF component. From the envelope spectrum corresponding to IMF1 and IMF2 components, it can be observed that the envelope spectrum of IMF1 component includes the fault characteristic frequency f_{inner} and its doubling frequency $2f_{inner}$ of the bearing inner ring. Correspondingly, the envelope spectrum of IMF2 contains the fault characteristic frequency of the bearing outer ring (fundamental frequency, harmonic components $2f_{outer}$, 3 and $4f_{outer}$). The decomposition results of the first layer show that the fault signals representing the inner and outer rings of the bearing can be significantly separated. The mutual information results in Table 3 also verify the effectiveness of this decomposition. Although the mutual information between the two IMFs is very small, it cannot be proved that each mode is a mono-component signal. According to the decomposition process of BT-EVMD, the termination condition needs to be further verified. Thus, the two IMFs are further decomposed, and the results are shown in Figure 16b,c, respectively.

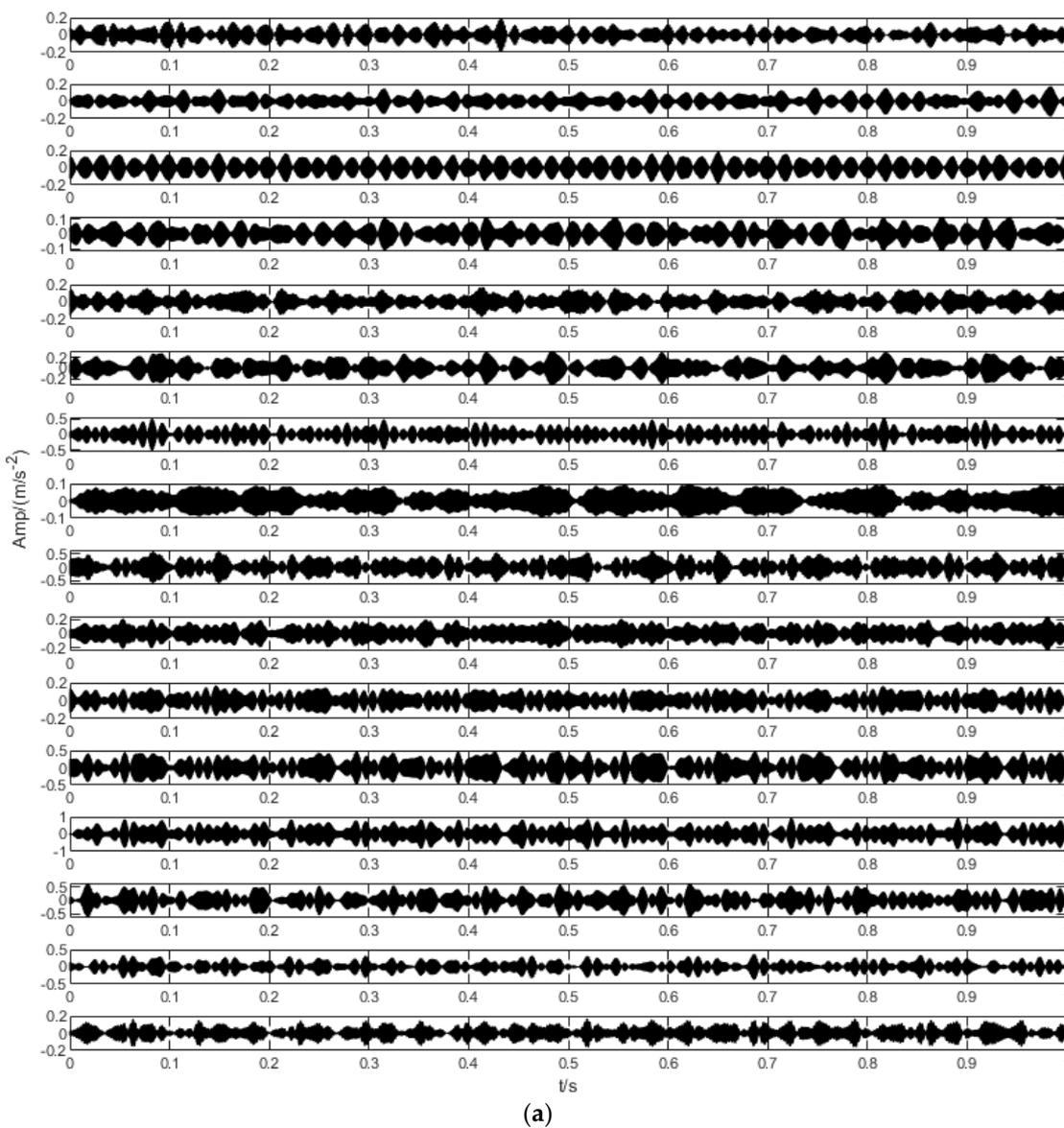


Figure 17. Cont.

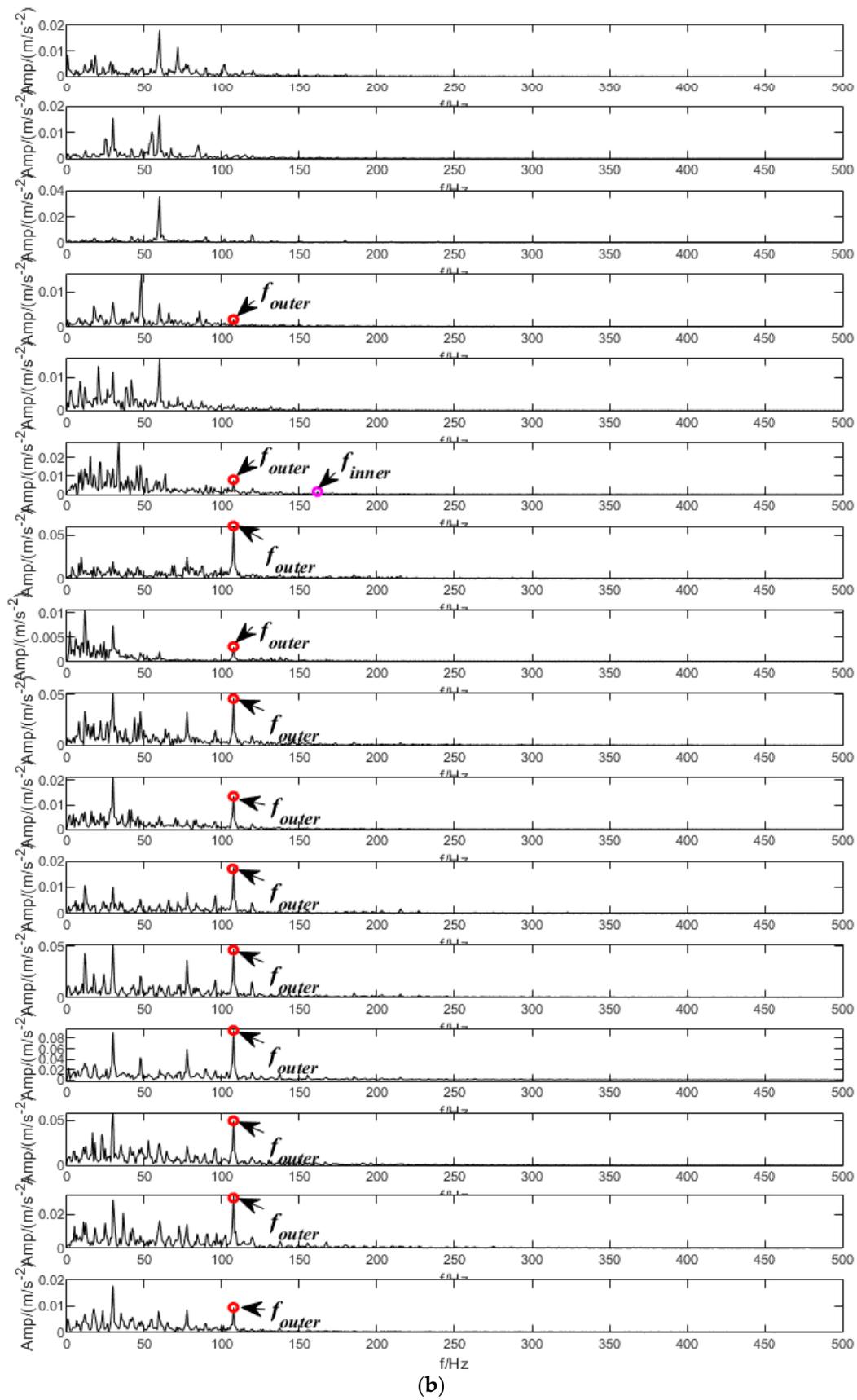


Figure 17. The IMF components and envelope spectrums using SVMD. (a) Decomposed modes. (b) Envelope spectrums.

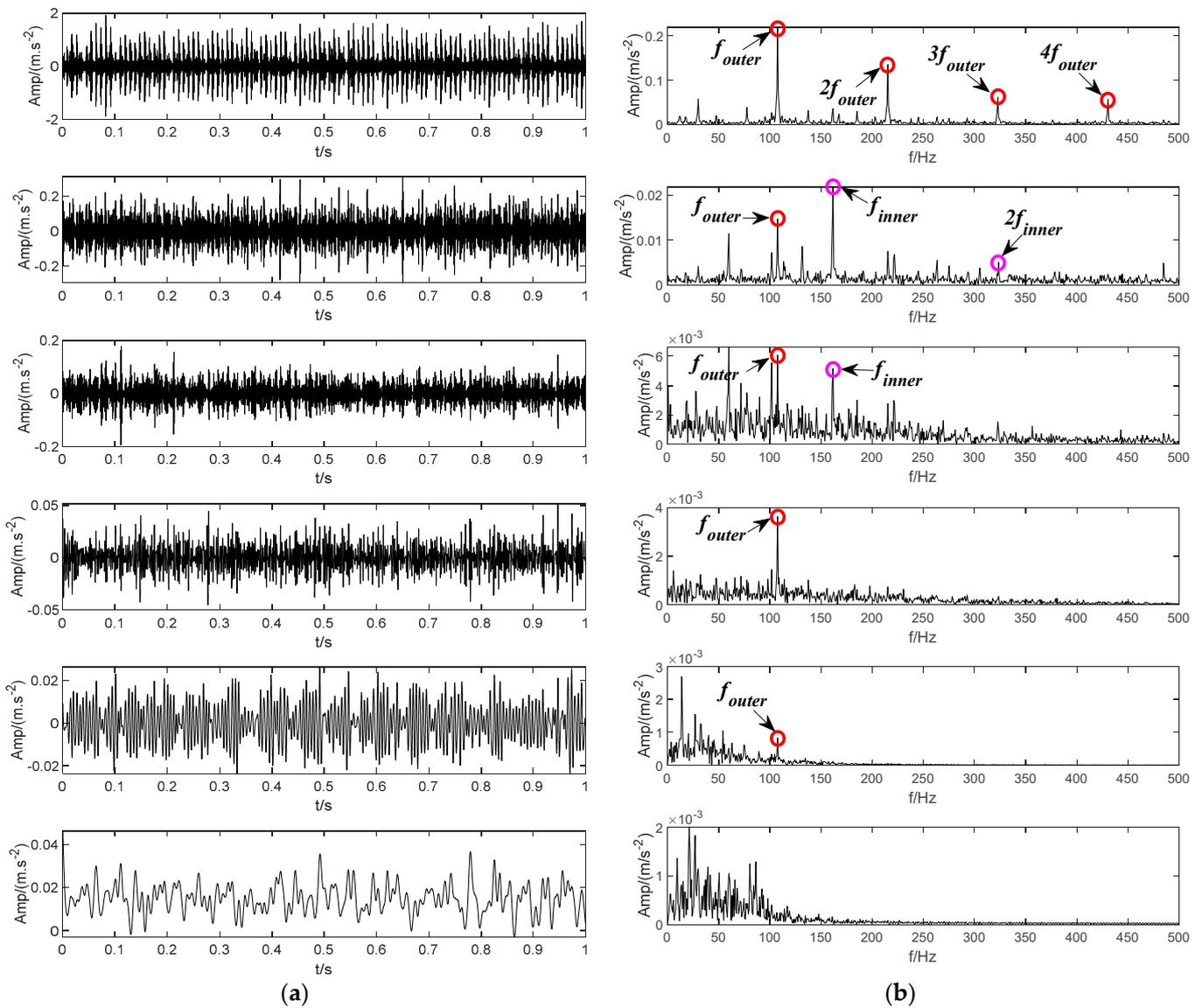


Figure 18. Decomposition result of CEEMDAN method. (a) The IMF components. (b) Envelope spectrums.

From the time domain components decomposed by IMF1 and IMF2 and their corresponding envelope spectra, it can be observed that the spectrum of IMF11 and IMF12 contain the fault characteristic frequency of the bearing inner ring, and the envelope spectra of IMF21 and IMF22 primarily contain the fault characteristic frequency of the bearing outer ring. The results in Table 2 show that the mutual information of the signals decomposed by IMF1 and IMF2 is relatively large. The situation indicates that the components are duplicated in the process of signal decomposition and there is no need to decompose. Afterwards, the components of the second layer decomposition are superimposed, that is, the original IMF1 and IMF2 components are restored. The final decomposition results are shown in Figure 16d.

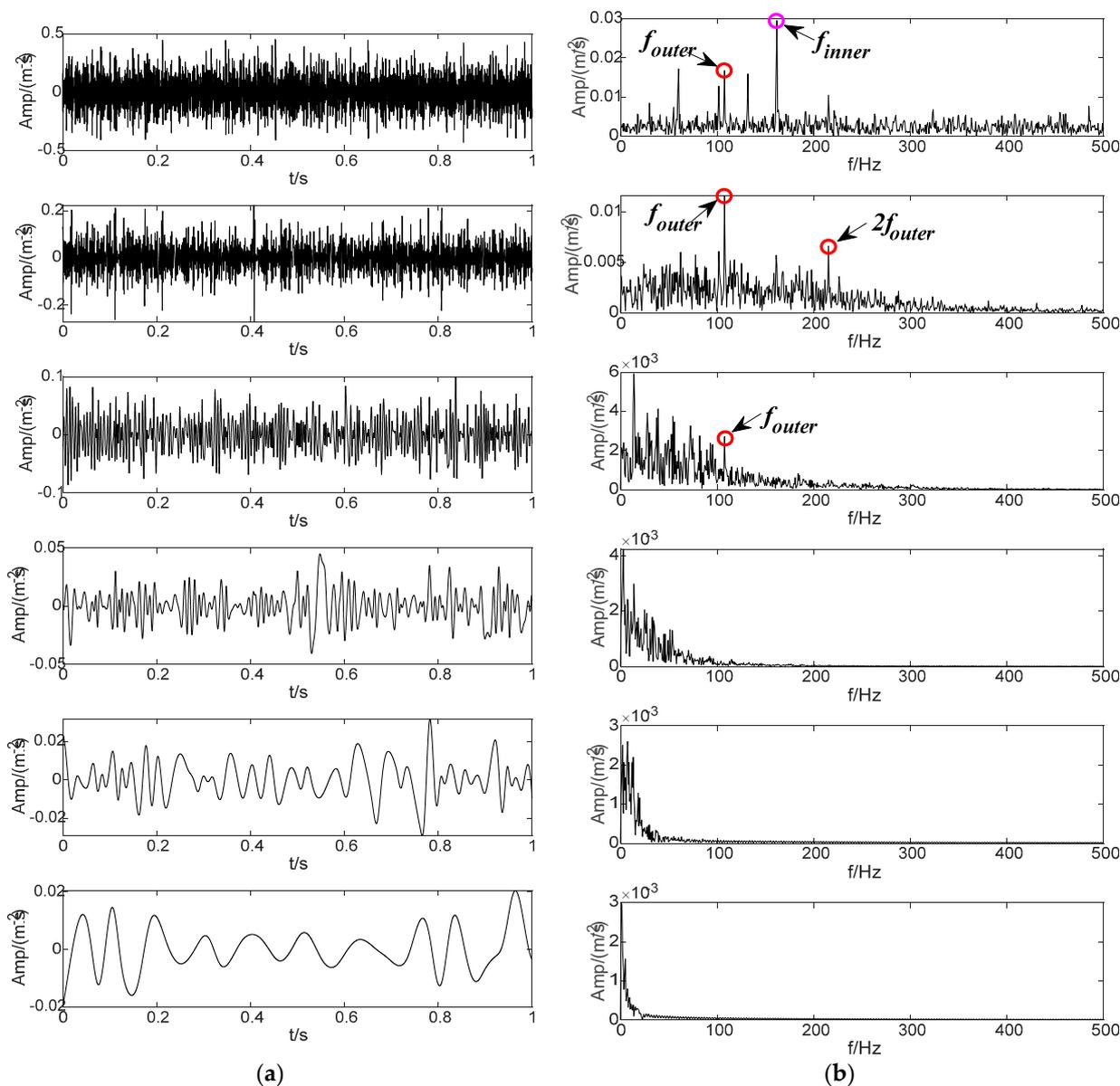


Figure 19. Decomposition result of LMD method. (a) The PF components obtained by LMD and (b) their envelope spectrums.

To evaluate the decomposition effect, some commonly used signal decomposition methods are used for signal processing and comparative analysis. Figure 17 shows the decomposition result based on the SVMD method, which finally obtains 16 IMF components. By observing the envelope spectrum corresponding to IMF components, it can be seen that many IMF components have common frequency items, indicating that they have copied themselves. The envelope spectrum of IMF6 component includes the fault characteristic frequencies of the inner and outer rings of the bearing at the same time. It shows that this method cannot effectively decompose the fault signals of the inner and outer rings in the signal. Figure 18 shows the decomposition results based on the CEEMDAN method, and six independent IMF components are obtained by this method. From the envelope spectrum analysis corresponding to each IMF component, it can be observed that some IMF components replicate each other. In particular, the envelope spectra of IMF2 and IMF3 components include the fault characteristic frequencies of the inner and outer rings of the bearing at the same time. In other words, there is a certain mixing feature. It shows that this method lacks the ability to effectively distinguish the internal and external fault

characteristics in the signal to some extent. Based on the decomposition of the LMD method, six PF components are obtained, and the results are shown in Figure 19. The observation shows that some PF components are duplicated, and the envelope spectrum of the PF1 component includes the fault characteristic frequencies of the inner and outer rings of the bearing. The results show that the LMD method still underperforms in the effective differentiation of fault items in the signal. The decomposition result of the ITD method is shown in Figure 20, which finally obtains four PRC components. In the component envelope spectrums shown in Figure 20b, PRC1 includes the fault characteristic frequency of the inner ring bearing, and the frequency spectrum of PRC4 contains the fault characteristic frequency of the bearing outer ring. However, the PRC2 and the PRC3 contain both inner and outer ring fault characteristic frequencies of the bearing. The results show that this method does not effectively distinguish the fault items in the signal. The comparative experimental results show that the BT-EVMD is more effective and robust than the above four methods.

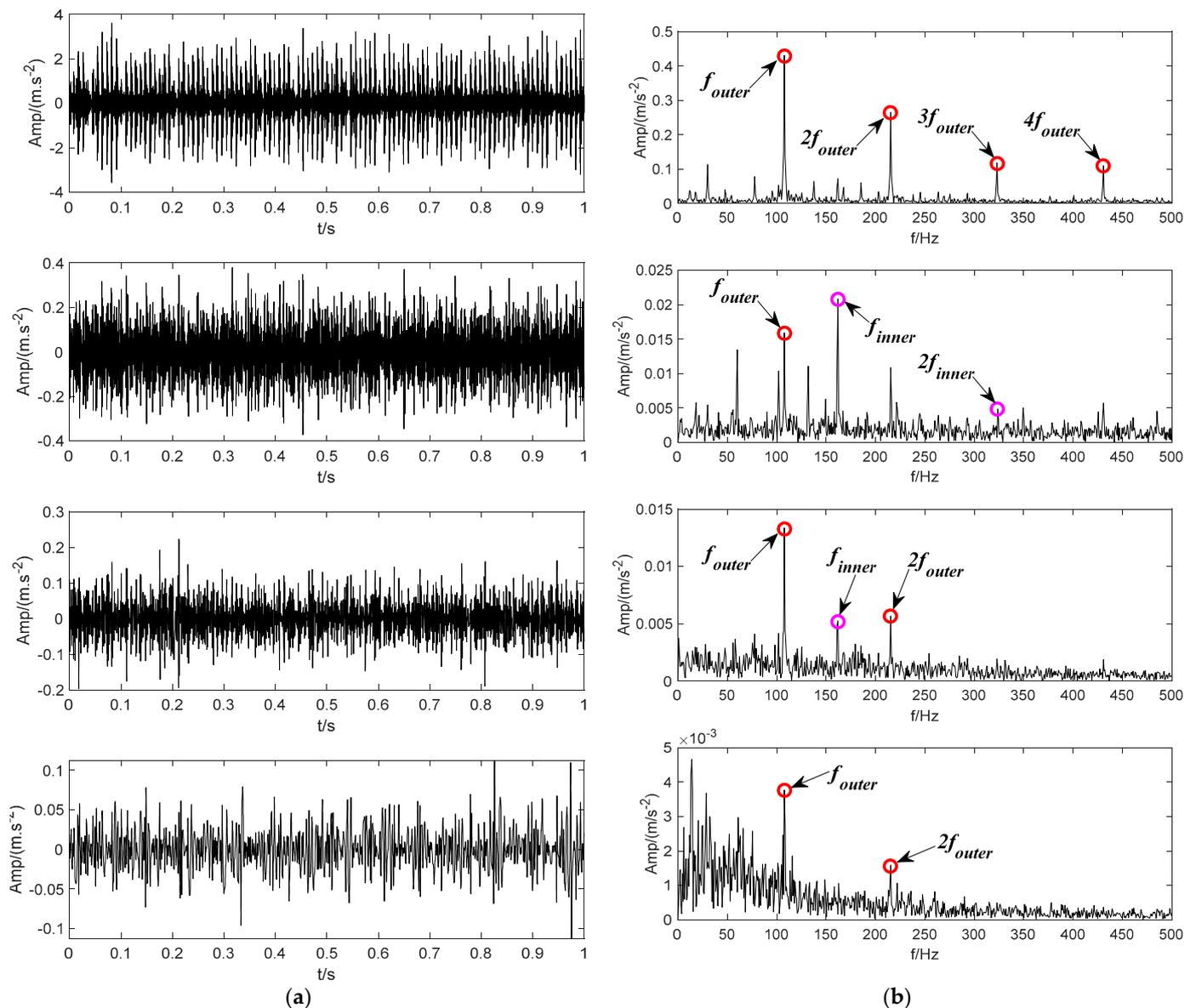


Figure 20. Performance evaluation of ITD method. (a) The PRC components obtained by ITD and (b) their envelope spectrums.

Table 3. The measurements obtained by BT-EVMD.

Decomposition Level	Parameters	Components	LSMI	Reconstruction Error
The first layer	$K = 2, \alpha_1 = 16,363, \tau = 0.019$	[IMF1, IMF2]	LSMI [IMF1, IMF2] = 0.0058	RSE = 0.0008
The second layer	$K_1 = 2, \alpha_1 = 21,386, \tau_1 = 0.011$	[IMF11, IMF12]	LSMI [IMF11, IMF12] = 0.1849	RSE1 = 0.0004
	$K_2 = 2, \alpha_2 = 5258, \tau_2 = 0.010$	[IMF21, IMF22]	LSMI [IMF21, IMF22] = 0.1916	RSE2 = 0.0002

5. Discussion and Conclusions

Novelties and Contributions: In this paper, an empirical VMD algorithm based on binary tree is proposed. By introducing the binary tree model and merging it into the VMD algorithm, the problem of selecting the total modal K of VMD can be effectively avoided. The binary tree model is a type of ergodic model that needs to select suitable stop conditions. In this paper, the least square mutual information is introduced as one of the stopping conditions of the VMD algorithm, which can effectively measure the nonlinear coupling degree of information between two IMF components and guarantee the orthogonality between IMF components. In addition, the performance of VMD is also affected by the value of the parameter $[\alpha, \tau]$. By analyzing the influence of parameter $[\alpha, \tau]$ on the performance of the VMD algorithm, the RCMDE of the signal is empirically introduced to dynamically adjust the value of parameter α , and the SNR of the signal is taken as the value of parameter τ . These improvements not only effectively avoid the experiential, blind and accidental problems of artificially setting parameters, they also avoid the rationality of designing fitness functions when searching VMD key parameters with an intelligent optimization algorithm, and the problem of excessive computational complexity. Lastly, simulation and measured signals are compared and analyzed with LMD, ITD, CEEMDAN and excellent SVMD methods. Experimental results indicate that the proposed method has better decomposition performance and robustness, and has full adaptability.

Further Work: The binary tree model requires a certain iterative process and certain judgment conditions, so it has increased certain computational complexity compared to traditional VMD. In addition, the value of the VMD parameter $[\alpha, \tau]$ is set through the computation of the signal time domain method, which is very empirical and lacks rigorous mathematical derivation and proof. Therefore, further research is required. In addition, there is a drawback when using RCMDE to dynamically adjust parameter α . That is, when the number of components in the signal is tiny, and their center frequency is too near, the value of RCMDE will lead to the value of α being too small, and the components whose center frequency is too close will be divided into the same frequency band, hence generating mode aliasing. Although this defect can be resolved through the iterative process of the binary tree, it will expend a large amount of computation. As a result, the value of parameter α and its solution need to be further researched.

Author Contributions: Conceptualization, B.X., F.Z. (Fengxing Zhou) and H.L.; methodology, H.L.; software, B.X.; validation, B.Y., B.X. and H.L.; formal analysis, B.Y.; investigation, F.Z. (Fengxing Zhou); resources, B.Y. and F.Z. (Fengqi Zhou); data curation, H.L. and B.X.; writing—original draft preparation, H.L. and B.X.; writing—review and editing, F.Z. (Fengqi Zhou); funding acquisition, F.Z. (Fengxing Zhou) and B.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant number 51975433 and Grant number 51975430); the Natural Science Foundation of Hubei Province (Grant number 2019CFB133).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors acknowledge the editors and reviewers for their constructive comments and all the supports on this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. He, M.; He, D. Deep Learning Based Approach for Bearing Fault Diagnosis. *IEEE Trans. Ind. Appl.* **2017**, *53*, 3057–3065. [[CrossRef](#)]
2. Wang, H.; Peng, C. Fuzzy Diagnosis Method for Rotating Machinery in Variable Rotating Speed. *IEEE Sens. J.* **2010**, *11*, 23–34. [[CrossRef](#)]
3. Chen, J.; Pan, J.; Li, Z.; Zi, Y.; Chen, X. Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. *Renew. Energy* **2016**, *89*, 80–92. [[CrossRef](#)]
4. Yan, R.; Gao, R.X.; Chen, X. Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Process.* **2014**, *96*, 1–15. [[CrossRef](#)]
5. Liu, X.; Jia, Y.X.; He, Z.W.; Zhou, J. Application of EMD-WVD and particle filter for gearbox fault feature extraction and remaining useful life prediction. *J. Vibroeng.* **2017**, *19*, 1793–1808. [[CrossRef](#)]
6. Faysal, A.; Ngui, W.K.; Lim, M.H.; Leong, M.S. Noise Eliminated Ensemble Empirical Mode Decomposition Scalogram Analysis for Rotating Machinery Fault Diagnosis. *Sensors* **2021**, *21*, 8114. [[CrossRef](#)] [[PubMed](#)]
7. Wang, Y.H.; Rao, Y.; Xu, D. Multichannel maximum-entropy method for the Wigner-Ville distribution. *Geophysics* **2020**, *85*, V25–V31. [[CrossRef](#)]
8. Zhang, X.H.; Zhao, J.M.; Bajrić, R.; Wang, L.L. Application of the DC Offset Cancellation Method and S Transform to Gearbox Fault Diagnosis. *Appl. Sci.* **2017**, *7*, 207. [[CrossRef](#)]
9. Peppas, K.P.; Mathiopoulou, P.T.; Yang, J.; Zhang, C.; Sasase, I. High-Order Statistics for the Channel Capacity of EGC Receivers Over Generalized Fading Channels. *IEEE Commun. Lett.* **2018**, *22*, 1740–1743. [[CrossRef](#)]
10. Guo, T.; Deng, Z.M. An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing. *Appl. Acoust.* **2017**, *127*, 46–62. [[CrossRef](#)]
11. Li, C.W.; Zhan, L.W.; Shen, L.Q. Friction Signal Denoising Using Complete Ensemble EMD with Adaptive Noise and Mutual Information. *Entropy* **2015**, *17*, 5965–5979. [[CrossRef](#)]
12. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Signal Process. Control* **2014**, *14*, 19–29. [[CrossRef](#)]
13. Osman, S.; Wang, W. A Morphological Hilbert-Huang Transform Technique for Bearing Fault Detection. *IEEE Trans. Instrum. Meas.* **2016**, *27*, 2646–2656. [[CrossRef](#)]
14. Gao, S.P.; Xu, Z.X.; Song, G.B.; Shao, M.X.; Jiang, Y.Y. Fault Location of Hybrid Three-terminal HVDC Transmission Line Based on Improved LMD. *Electr. Pow. Syst. Res.* **2021**, *201*, 107550. [[CrossRef](#)]
15. Song, E.Z.; Gao, F.; Yao, C.; Ke, Y. Research on Rolling Bearing Fault Diagnosis Method Based on Improved LMD and CMWPE. *J. Fail. Anal. Prev.* **2021**, *21*, 1714–1728. [[CrossRef](#)]
16. Priyadarshini, L.; Dash, P.K. Detection of Islanding and Non-islanding Fault Disturbances in Microgrid Using LMD and Deep Stacked RVFLN Based Auto-encode. *Electr. Eng.* **2021**, *103*, 2747–2767. [[CrossRef](#)]
17. Yu, M.; Pan, X. A Novel ITD-GSP-based Characteristic Extraction Method for Compound Faults of Rolling Bearing. *Measurement* **2020**, *159*, 107736. [[CrossRef](#)]
18. Voznesensky, A.; Kaplun, D. Adaptive Signal Processing Algorithms Based on EMD and ITD. *IEEE Access* **2019**, *7*, 171313–171321. [[CrossRef](#)]
19. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [[CrossRef](#)]
20. Li, K.; Su, L.; Wu, J.J.; Wang, H.Q.; Chen, P. A Rolling Bearing Fault Diagnosis Method Based on Variational Mode Decomposition and an Improved Kernel Extreme Learning Machine. *Appl. Sci.* **2017**, *7*, 1004. [[CrossRef](#)]
21. Zhang, M.; Jiang, Z.N.; Feng, K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. *Mech. Syst. Signal Process.* **2017**, *93*, 460–493. [[CrossRef](#)]
22. Mei, L.; Li, S.Y.; Zhang, C.; Han, M.X. Adaptive Signal Enhancement Based on Improved VMD-SVD for Leak Location in Water-Supply Pipeline. *IEEE Sens. J.* **2021**, *21*, 24601–24612. [[CrossRef](#)]
23. Liu, C.F.; Zhu, L.D.; Ni, C.B. Chatter detection in milling process based on VMD and energy entropy. *Mech. Syst. Signal Process.* **2018**, *105*, 169–182. [[CrossRef](#)]
24. He, X.Z.; Zhou, X.Q.; Yu, W.N.; Hou, Y.X.; Mechefske, C.K. Adaptive Variational Mode Decomposition and Its Application to Multi-fault Detection Using Mechanical Vibration Signals. *ISA Trans.* **2021**, *111*, 360–375. [[CrossRef](#)] [[PubMed](#)]
25. Xiao, H.S.; Wei, J.C.; Liu, H.S.; Li, Q.Q.; Shi, Y.L. Identification method for power system low-frequency oscillations based on improved VMD and Teager-Kaiser energy operator. *IET Gener. Transm. Distrib.* **2017**, *11*, 4096–4103. [[CrossRef](#)]
26. Kaur, C.; Bisht, A.; Singh, P.; Joshi, G. EEG Signal denoising using hybrid approach of Variational Mode Decomposition and wavelets for depression. *Biomed. Signal Process. Control* **2021**, *65*, 102337. [[CrossRef](#)]
27. Long, J.C.; Wang, X.P.; Dai, D.D.; Tian, M.; Zhu, G.W.; Zhang, J. Denoising of UHF PD signals based on optimised VMD and wavelet transform. *IET Sci. Meas. Technol.* **2017**, *11*, 753–760. [[CrossRef](#)]

28. Li, Y.; Xu, F.Y. Acoustic emission sources localization of laser cladding metallic panels using improved fruit fly optimization algorithm-based independent variational mode decomposition. *Mech. Syst. Signal Process.* **2022**, *166*, 108514. [[CrossRef](#)]
29. Wang, X.B.; Yang, Z.X.; Yan, X.A. Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery. *IEEE/ASME Trans. Mechatron.* **2017**, *23*, 68–79. [[CrossRef](#)]
30. Xu, B.; Zhou, F.X.; Li, H.P.; Yan, B.K.; Liu, Y. Early fault feature extraction of bearings based on Teager energy operator and optimal VMD. *ISA Trans.* **2019**, *86*, 249–265. [[CrossRef](#)]
31. Nazari, M.; Sakhaei, S.M. Variational Mode Extraction: A New Efficient Method to Derive Respiratory Signals from ECG. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 1059–1067. [[CrossRef](#)] [[PubMed](#)]
32. Nazari, M.; Sakhaei, S.M. Successive Variational Mode Decomposition. *Signal Process.* **2020**, *174*, 107610. [[CrossRef](#)]
33. Liu, S.S.; Yu, K.P. Successive Multivariate Variational Mode Decomposition Based on Instantaneous Linear Mixing Model. *Signal Process.* **2022**, *190*, 108311. [[CrossRef](#)]
34. Chen, Q.M.; Chen, J.M.; Lang, X.; Xie, L.; Rehman, N.U.; Su, H.Y. Self-tuning Variational Mode Decomposition. *J. Frankl. Inst.* **2021**, *358*, 7825–7862. [[CrossRef](#)]
35. Pahlavani, P.; Bigdeli, B. A mutual information Dempster-Shafer based decision ensemble system for land cover classification of hyperspectral data. *Front. Earth Sci.* **2017**, *11*, 774–783. [[CrossRef](#)]
36. Valdez, M.A.; Jaschke, D.; Vargas, D.L.; Carr, L.D. Quantifying complexity in quantum phase transitions via mutual information complex networks. *Phys. Rev. Lett.* **2017**, *119*, 225–301. [[CrossRef](#)] [[PubMed](#)]
37. Ball, K.R.; Grant, C.; Mundy, W.R.; Shafer, T.J. A multivariate extension of mutual information for growing neural networks. *Neural Netw.* **2017**, *95*, 29–43. [[CrossRef](#)]
38. Tani, N.; Ohta, M.; Higuchi, Y.; Akatsu, J.; Kumashiro, M. Lifestyle and subjective musculoskeletal symptoms in young male Japanese workers: A 16-year retrospective cohort study. *Prev. Med. Rep.* **2020**, *20*, 101171. [[CrossRef](#)]
39. Sukriti; Chakraborty, M.; Mitra, D. Automated detection of epileptic seizures using multiscale and refined composite multiscale dispersion entropy. *Chaos Solitons Fractals* **2021**, *146*, 110939. [[CrossRef](#)]
40. Rostaghi, M.; Azami, H. Dispersion Entropy: A Measure for Time-Series Analysis. *IEEE Signal Process. Lett.* **2016**, *23*, 610–614. [[CrossRef](#)]
41. Azami, H.; Rostaghi, M.; Abasolo, D.; Escudero, J. Refined Composite Multiscale Dispersion Entropy and its Application to Biomedical Signals. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2872–2879. [[CrossRef](#)] [[PubMed](#)]
42. CWRU Bearing Data Center. Available online: <http://csegroups.case.edu/bearingdatacenter/home> (accessed on 17 October 2020).
43. Smith, W.A.; Randall, R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. *Mech. Syst. Signal Process.* **2015**, *64*, 100–131. [[CrossRef](#)]