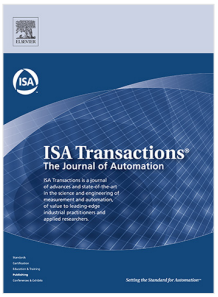
Accepted Manuscript



Early fault feature extraction of bearings based on Teager energy operator and optimal VMD

Bo Xu, Fengxing Zhou, Huipeng Li, Baokang Yan, Yi Liu

PII: S0019-0578(18)30445-2
DOI: https://doi.org/10.1016/j.isatra.2018.11.010
Reference: ISATRA 2958

To appear in: *ISA Transactions*

Received date : 13 June 2018
Revised date : 4 November 2018
Accepted date : 8 November 2018

Please cite this article as: B. Xu, F. Zhou, H. Li et al. Early fault feature extraction of bearings based on Teager energy operator and optimal VMD. *ISA Transactions* (2018), https://doi.org/10.1016/j.isatra.2018.11.010

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*Title page sh ACCEPTED MANUSCRIPT

Early fault feature extraction of bearings based on Teager energy operator and optimal VMD

Bo Xu[a,b], Fengxing Zhou[a,*], Huipeng Li[a,b], Baokang Yan[a], Yi Liu[b,c]

[a] *Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, Hubei 430081, PR China*
[b] *School of Electronic Information, Huang gang Normal University, Huang gang, Hubei, 438000, PR China*
[c] *School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, PR China*

Abstract: As the fault shock component in vibration signals is extremely sparse and weak, it is difficult to extract the fault features when large-scale, low-speed and heavy-duty mechanical equipment is in the early stage of failure. To solve this problem, an early fault feature extraction method based on the Teager energy operator, combined with optimal variational mode decomposition (VMD) is presented in this study. First, the Teager energy operator was used to strengthen the weak shock component of the original signal. Next, a logistic-sine complex chaotic mapping with variable dimensions was constructed to enhance the global search ability and convergence speed of the pigeon-inspired optimization (PIO) algorithm, which is named the variable dimension chaotic pigeon-inspired optimization (VDCPIO) algorithm. Then, the VDCPIO algorithm is used to search for the optimal combination value of key parameters of VMD. The enhanced vibration signal is decomposed into a set of intrinsic mode functions (IMFs) by the optimized VMD, and then kurtosis for every IMF and mean kurtosis of all IMFs are extracted. According to the average kurtosis, several IMFs, whose kurtosis value is greater than the average kurtosis value, are selected to reconstruct a new signal. Then, envelope spectrum analysis of the reconstructed signal is carried out to extract the early fault features. Finally, experimental verification of the method was performed using the simulated signal and measured signal from a rolling bearing; the experimental results indicate that the method presented in this paper is more effective to extract the early fault features of this kind of mechanical equipment.

Keywords: low-speed and heavy load; early fault; VMD; VDC; PIO; Teager energy operator

1 Introduction

Large-scale, low-speed and heavy-load machinery such as radar turrets and converters, generally has rotational speed lower than 60 rpm. The key components of these devices can be easily damaged. Faults can cause sudden and serious equipment accidents, and even endanger the safety of operators. Therefore, early fault diagnosis is of great theoretical significance and practical value, to ensure safe operation and avoiding heavy economic losses. However, early fault signals of this type of equipment have low energy and long periods, which are affected by strong background noise and attenuation of the signal. The fault features are extremely weak. Extracting fault information from early fault signals is the key to fault diagnosis.

The vibration signals of rotating machinery exhibit obvious nonlinear and non-stationary characteristics. Commonly used signal processing methods include the Short-time Fourier Transform (STFT) [1,2], Wavelet Transform (WT) [3,4], Wigner-Vile Distribution (WVD) [5,6], S Transformation (ST) [7,8], Higher Order Statistics (HOS) [9], Empirical Mode Decomposition (EMD) [10-13], Local Mean Decomposition (LMD) [14-16], and Intrinsic Time Scale Decomposition (ITD) [17,18]. The STFT is constrained by the time-frequency resolution of the window function, and the time-frequency resolution is fixed; WT needs to select the wavelet basis and decomposition layer, which makes it essenually a non-adaptive signal processing method; the ST transform combines the advantages of the STFT and the WT, but its spectrum is erratic, and constrained by the unqualified principle; WVD is heavily affected by the cross term; HOS can theoretically completely suppress Gaussian noise, but it cannot reduce the non-Gaussian noise, and can easily interfere with the higher-order spectrum of the signal. EMD, as an adaptive signal processing method, has many problems such as fitting-overshoot, the end-effect, and mode-mixing; therefore, its practical application is severely limited. LMD is likely to cause signal mutation in the process of demodulation and the computation burden is heavy; ITD uses the linear transformation method to divide the frequency domain of the vibration signal, which results in the local fluctuation of the waveform amplitude of the proper rotation component (PRC), and then leads to signal distortion. Many scholars have proposed a large number of improvement schemes based on the above methods. These studies presented fairly good results, which have been practically applied to mechanical fault diagnosis. However, due to the limitations of the theoretical framework, the problems discussed can only be suppressed to some extent by existing approaches, but cannot be fundamentally eliminated.

VMD is a new adaptive signal processing method [19]. By solving the variation problem iteratively, the signal is decomposed into a set of finite band modal functions. This method realizes the separation of each frequency component of the signal, and overcomes the problems, such as the end-effect, mode-mixing, and wave-fluctuation. It is very suitable for analyzing nonlinear and non-stationary vibration signals, and is widely used in the field of mechanical fault diagnosis [20-23]. However, the performance of VMD is affected by the total number of modal parameters K, quadratic penalty parameter α, update parameter τ, and convergence tolerance parameter ε. Moreover, these parameters must be preset and require a high level of experience and intuition. Therefore, VMD is not suitable as an adaptive model in practical engineering applications.

ACCEPTED MANUSCRIPT

At present, little research focuses on the parameter setting of VMD. Liu [24] proposed a traditional grid search method for searching the optimal values of the decomposition parameters K and α of VMD, but it was based on personal experience and intuition, lacking strong theoretical support. Xiao [25] proposed a parameter search method based on the peak number and minimum frequency of the FFT spectrum to set the value K and α of VMD. In [26], the FFT's spectral envelope is adopted to identify the effective frequency components contained in the signal, and the optimal value of VMD parameter K is determined according to the frequency components. However, the FFT method is not suitable when dealing with nonlinear and non-stationary vibration signals. The PSO optimization algorithm was used to search the optimal values of the parameters K and α of VMD in [27, 28]. However, it can suffer from the problem of local optimal solution, and the obtained values are not the global best values

With the significant development of bionic intelligent optimization technology, this paper attempts to use a new population-based swarm intelligence optimization algorithm to optimize the key parameters of VMD. Inspired by the homing behavior of doves in nature, Duan [29] proposed a new bionic intelligent optimization algorithm PIO (Pigeon-Inspired Optimization) in 2014. It mainly focuses on the optimization of the parameters of actual problems, and can be applied in areas including UAV formation, parameter optimization, automatic control, and image processing. [30-34]. Compared with other swarm intelligence algorithms, the pigeon swarm optimization algorithm is easy to realize, easy to program, has less control parameters, and is easy to embed into specific search steps. However, the standard PIO is still a heuristic optimization algorithm, and can easily suffer from the problem of finding the local optimal solution. In [35], a pigeon swarm optimization algorithm based on the quantum rotary gate and Chebyshev chaotic search is proposed, which can solve this problem to some extent. However, the traditional quantum rotary gate updates the phase of the quantum bit by tediously comparing the rotary angle in the query table, and the angle values in the table are discrete and fixed. Therefore, it has limited ability in increasing the population diversity because of the limited-state mutation operator.

Large-scale, low-speed and heavy-load machinery has very weak early fault shock features and extremely low signal-to-noise ratio (SNR). In order to extract the early fault features more effectively, a novel Teager-VMD based fault feature extraction method is investigated in this paper, which takes advantage of the variable dimensional chaotic PIO. The contributions of this work can be summarized as follows:

(1) Using the Teager energy operator to preprocess the original vibration signal, enhancing the shock component in the signal and the signal-to-noise ratio of the signal.

(2) Constructing a variable dimension logistic-sine complex chaotic mapping by using logistic mapping and sine mapping according to the unique search mechanism of PIO, which is also used to solve the problem of local optimization. That is, in the PIO initialization stage, the one-dimensional logistic-sine complex chaotic mapping was used to initialize the pigeon population and increase the diversity of the population. Then, in the search stage based on the map and compass operator, the random update mechanism of PIO was replaced by the two-dimensional logistic-sine complex chaotic mapping. Lastly, in the search stage based on the landmark operator, the random updating mechanism of PIO was replaced by the one-dimensional logistic-sine complex chaotic map again. In this way, the global search capability and convergence speed of the PIO can be effectively improved.

(3) Using the improved PIO to search for the optimal combination value of VMD's parameters K and α, which can help to construct an optimal VMD to achieve the best processing in fault feature extraction.

(4) Using the optimal VMD to decompose the enhanced signals into a set of IMFs, calculating the kurtosis of each IMF and the average kurtosis of all IMFs. Next, according to the average kurtosis, several IMFs, whose kurtosis is greater than average kurtosis, are screened out, and a new signal is reconstructed. Then, envelope spectrum analysis of the reconstructed signal is carried out to extract the features of the early faults. Finally, the effectiveness of the proposed method is verified by analyzing the simulated and measured bearing fault signals.

The remainder of this paper is organized as follows. In Section 2, the VMD method and the influence of its parameters are introduced in detail. In Section 3, the PIO algorithm is studied. In Section 4, our algorithm is presented and its main steps are explained. In Section 5, simulations of the algorithm are conducted and the results are discussed. In Section 6, case studies of fault diagnoses are provided to test the effectiveness of our approach. Finally, conclusions and future work are described in Section 7.

2. Variational mode decomposition

The decomposition process of VMD is the iterative solution process for the variational problem, which can be divided into two steps: establishing the variational model and solving the model. It is an adaptive signal decomposition method based on Wiener filtering, Hilbert transform, and heterodyne demodulation.

2.1. Establishment of variational models

We assume that the input signal $x(t)$ is composed of several signal components with different center frequencies and limited bandwidth. The signal decomposition problem can be transformed into the variational model. When the constraint that the sum of the components is equal to the input signal $x(t)$ is satisfied, the sum of the aggregating bandwidths of the components is considered to be the smallest. The construction steps are as follows:

ACCEPTED MANUSCRIPT

Step 1: In terms of Hilbert transform, the analytical signal of each modal function $u_k(t)$ is obtained to calculate the single-side spectrum of $u_k(t)$: $\left[\delta(t)+\frac{j}{\pi t}\right]*u_k(t)$, where $\delta(t)$ is the Dirichlet function, and $*$ is the convolution symbol.

Step 2: Frequency mixing is used to mix the analytical signals of each component with a pre-estimated center frequency $e^{-j\omega_k t}$, and then modulate the spectrum of each mode to the corresponding baseband: $\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}$.

Step 3: The square L^2 norm of the demodulation signal gradient is calculated, and the bandwidth of each mode signal is estimated. The optimal variational model can be constructed by introducing constraint conditions:

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

where, K is the number of modal components; $\{u_k\}:=\{u_1,u_2,\cdots u_K\}$ and $\{\omega_k\}:=\{\omega_1,\omega_2,\cdots,\omega_K\}$ are shorthand notations for the set of all modes and their center frequencies, respectively. $\sum_k:=\sum_{k=1}^{K}$ is understood as the summation over all modes.

2.2. Variational model solution

The second-order penalty factor α and the Lagrangian multiplication operator $\lambda(t)$ are introduced to transform the constraint variation problem into a non-binding variational problem. The second penalty factor can guarantee the reconstruction accuracy of the signal in the presence of Gaussian noise. The Lagrange multiplier can ensure the strictness of the model constraint conditions. The extended Lagrange can be written as:

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

On this basis, the alternate direction method of multipliers (ADMM) is used for the alternate iterative update $\{u_k^{n+1}\}$, $\{\omega_k^{n+1}\}$ and λ^{n+1} to search for the "saddle point" of an extended Lagrange expression, that is, to satisfy the stopping condition of the iteration. Finally, K mutually independent frequency band components are obtained. The solution procedure for the variational model is as follows:

Algorithm 1: Complete optimization of VMD

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

repeat:

$n\leftarrow n+1$

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

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

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

Update ω_k:

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

ACCEPTED MANUSCRIPT

end for

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

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

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

2.3. Influence of variational model parameters

It is known from the variational model solution process that the performance of VMD is closely related to the decomposition parameters, such as the total number of modalities K, the secondary penalty α, the update step length τ, and the convergence fault tolerance threshold ε. Various parameters affect the decomposition performance of VMD, as follows:

1) The total number of modal components K. The performance of VMD is very sensitive to the value of K. If the value of K is too small, the data will be under-segmented, and some components will be contained in other modalities. If the value K is too large, problems such as modal copying will occur.

2) The secondary penalty parameter α. If the value of α is too small, the bandwidth of the modal component will be too large; some components will be included in other modal components, or additional "noise" will be captured. If the value of α is too large, the bandwidth of the modal component will be too small, and some components in the original signal will be lost.

3) Update the step size τ. The VMD algorithm does not guarantee convergence to a global minimum. When the noise level of the signal is low, the Lagrangian multiplier ensures optimal convergence when the appropriate value of τ is chosen, and $\tau > 0$. When the signal has a high noise level and $\tau > 0$, the Lagrangian multiplier will severely hinder the convergence of the algorithm. Setting τ to 0 can effectively turn off the Lagrangian multiplier to ensure that the algorithm converges effectively.

4) Convergence tolerance ε. The value of parameter ε affects the reconstruction accuracy of the VMD decomposition. The reconstruction error can be controlled by reducing the convergence of the stopping criterion to a specified degree.

Therefore, the design of an optimal VMD should be focused on how to obtain the optimal combination value of key parameters of the VMD.

3. Variable dimension chaotic PIO

The autonomous homing behavior of pigeons has been an unsolved mystery for a long period of time, which has attracted many researchers to find out exactly how pigeons can accurately return home in long-distance conditions. The researchers proposed that the long-range homing behavior of pigeons is based on the dual help of the geomagnetic field and geomorphological landscape. The geomagnetic field is used to determine the general direction, and the actual direction is corrected based on the geomorphological landscape for the purpose of accurately locking the position.

3.1. Mathematical model of the pigeon swarm optimization algorithm

Imitating the mechanism that pigeons use as navigation tools to find targets, the PIO proposes two different operator models.

1) Map and compass operator. Pigeons can use magnetic objects to sense geomagnetic fields and then form maps in their minds. Meanwhile, they use the sun's altitude as a compass to orient their flight, and as they approach their destination, their dependence on the sun and magnetic objects decreases.

2) Landmark operator. Landmark operators are used to simulate the effects of landmarks on pigeons in navigation tools. When pigeons fly close to their destination, they will rely more on nearby landmarks. If pigeons are familiar with landmarks, they will fly directly to their destination. Otherwise, they will follow pigeons familiar with landmarks.

In the PIO model, virtual pigeons are used to simulate the navigation process. According to the principle of the map and compass operator, the pigeon's position and velocity are initialized, and in the multi-dimensional search space, the pigeon's position and velocity are updated in each iteration. Its position and velocity are recorded separately, as follows:

$$\begin{cases} X_i = [x_{i1}, x_{i2}, \cdots, x_{iD}] \\ V_i = [v_{i1}, v_{i2}, \cdots, v_{iD}] \end{cases} \tag{7}$$

Where, i =1, 2, …, N, and based on formula (8), each pigeon updates its position $\boldsymbol{X_i}$ and speed $\boldsymbol{V_i}$.:

ACCEPTED MANUSCRIPT

$$\begin{cases} V_i^{N_c} = V_i^{N_c-1} * e^{-R\times N_c} + rand \times (X_{gbest} - X_i^{N_c-1}) \\ X_i^{N_c} = X_i^{N_c-1} + V_i^{N_c} \end{cases} \tag{8}$$

where R is the map and compass factor, which has values between [0,1]. *rand* is a random number in the range [0,1]; N_c is the current number of iterations; $\boldsymbol{X_{gbest}}$ is the global optimal position after N_c-1 iterations, obtained by comparing the positions of all the pigeons. When the number of iterations reaches the required number, the calculation of the map and the compass operator is stopped, and the landmark operator is then calculated in the next iteration.

In the landmark operator, half the number of pigeons will be reduced after each iteration. As the pigeons far from the destination are unfamiliar with the landmarks, they will no longer have the ability to distinguish routes and are therefore discarded. $\boldsymbol{X_{center}}$ is the central position of the remaining pigeons, which will be treated as a landmark, and the reference direction for the flight. Based the following equation, the pigeon's position $\boldsymbol{X_i}$ can be updated:

$$X_{center}^{N_c-1} = \frac{\sum_{i=1}^{N^{N_c-1}} X_i^{N_c-1} \times F(X_i^{N_c-1})}{N^{N_c-1} \sum_{i=1}^{N^{N_c-1}} F(X_i^{N_c-1})}, \quad N^{N_c} = \frac{N^{N_c-1}}{2} \tag{9}$$

$$X_i = X_i^{N_c-1} + rand \times (X_{center}^{N_c-1} - X_i^{N_c-1}) \tag{10}$$

where:

$$F\left(X_i^{N_c-1}\right) = \begin{cases} \dfrac{1}{fitness(X_i^{N_c-1}) + \varepsilon} = fitness_{\min}(X_i^{N_c-1}) \\ fitness(X_i^{N_c-1}), \text{and } fitness_{\max}(X_i^{N_c-1}) > 0 \end{cases} \tag{11}$$

In Eq. (11), $\boldsymbol{fitness}\ (\cdot)$ is a fitness function, $\boldsymbol{fitness_{min}}\ (\cdot)$ addresses the minimization problem, and $\boldsymbol{fitness_{max}}\ (\cdot)$ addresses the maximization problem. Similarly, when the iterations reach the maximum number of iterations, the landmark operator also stops working. From the PIO algorithm processing, we can understand that the algorithm is divided into three stages: the ①initialization stage; ②map and compass operation stage; and ③landmark operation stage. In stage ①, the *rand* operator is used to initialize the position and velocity of the pigeon group, and the diversity of the individual is limited as well. In stage ②, the position and velocity of the pigeon population are updated by using the *rand* and *R* operator, which makes it easy for the algorithm to only obtain the local optimal solution. In stage ③, the *rand* operator is used to update the location of the remaining pigeons, which may also yield the local optimal solution. In addition, although the definition and function of the R operator are given in [29], and its range of values is set to be [0,1], the basis of setting the value of R is not given. Furthermore, once the value of R is given, it will be fixed in the search process of the second stage of the PIO, which means that every individual has to adjust their own speeds to be the same. Therefore, the search space in the second stage is limited, to a certain extent, by the size and type of the *R* operator, which is more likely to suffer from precocity problems.

3.2. Improvement of the algorithm

PIO has simple structure and strong convergence ability. However, it still has the inherent weakness of traditional bionic intelligent algorithms, which is that it can easily obtain the local optimal solution [35]. Therefore, in accordance with the unique search mechanism of PIO, in this section, an improved algorithm that combines a variable dimensional composite chaotic map with PIO, is proposed to improve the searching ability of the global optimal solution and convergence speed of PIO.

3.2.1. Complex dimensional chaotic mappings

Chaotic mapping [36] is a mathematical function that shows some chaotic behavior over time. Different values of a chaotic function form chaotic sequences. The population is initialized by the chaotic characteristics of the sequence generated by the chaotic mapping, which makes the population have better diversity. At present, there are many different chaotic mapping functions, among which logistic mapping and sine mapping are commonly used. They have the advantages of simple structure and good chaotic characteristics [37,38]. However, the mapping structure of these chaotic functions is simple, which also leads to the disadvantages of simple chaotic behavior and a fragile time interval. The construction process is as follows:

The definition of a known logistic mapping is as follows:

$$x_{n+1} = \mu x_n (1 - x_n) \tag{12}$$

where μ is the chaos control parameter and $\mu = 4\alpha, \alpha \in [0,1]$; x_n is the variable of the logistic map, and $x_n \in [0,1]$.

The definition of a known sine mapping is as follows:

ACCEPTED MANUSCRIPT

$$x_{n+1} = \beta \sin(\pi x_n) \tag{13}$$

where β is the chaos control parameter, and $\beta \in [0,1]$; x_n is the variable of the sine map $x_n \in [0,1]$.

The logistic map and sine map have the disadvantages of simple behavior and weak chaotic time intervals, and their weakness may negatively affect some chaos-based applications. Therefore, the logistic and sine maps are combined, and Eq. (12) is added into Eq. (13), which can obtain two completely different chaotic mappings, as shown below.

①1D logistic-sine complex chaotic mapping:

$$x_{n+1} = \sin(\pi(4\theta x_n(1-x_n))) \tag{14}$$

where θ is the chaos control parameter and $\theta \in [0,1]$; x_n is the variable of the composite map, and $x_n \in [0,1]$. It is important to note here that the two chaotic maps have the same definitional domain and value domain, and their composite maps are not necessarily chaotic. Thus, it is possible to determine whether the composite map is a chaotic map or not by drawing its bifurcation diagram, as shown in Fig. 1. As can be seen from Fig. 1, the logistic map and sine map have differences in their mathematical formula, but their chaotic behaviors are very similar and both have good chaotic characteristics. In addition, the composed maps not only exhibit chaotic behavior, but also have more complex chaotic behavior and better chaotic properties. In order to further illustrate the chaotic characteristics of these three different maps, the corresponding Lyapunov exponents are calculated here, where the Lyapunov exponent is defined as follows:

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \ln\left| f'(x_n) \right| \tag{15}$$

The Lyapunov exponent can be used to effectively describe the speed of separation of adjacent points in maps, or the sensitivity of the orbit to initial conditions in the singular attractor, and the larger the value is, the more sensitive the mapping is to the initial conditions, and the better the chaotic characteristics are. By calculating the Lyapunov exponent of three different maps (as shown in Fig. 2), the chaotic behavior of the 1D complex chaotic map, logistic map and sine map appears in different control intervals. It can be observed from Fig. 1 that the logistic-sine composite map exhibits hyperchaotic behavior when the control parameter $\theta \in [0.42, 0.57] \cup [0.61, 0.82] \cup [0.84, 1]$; the sine composite map exhibits hyperchaotic behavior when the control parameter $\beta \in [0.28, 0.36] \cup [0.61, 0.78] \cup [0.8, 1]$; and the sine composite map demonstrates hyperchaotic behavior when the control parameter $\alpha \in [0.85, 1]$. By comparing the chaotic behavior of the above three maps, it can be seen that the logistic-sine composite map has the best chaotic properties and the largest Lyapunov exponent. It also shows that the logistic-sine composite mapping has better initial sensitivity and stronger ergodicity. Therefore, the 1D logistic-sine chaotic map shows better chaotic characteristics.

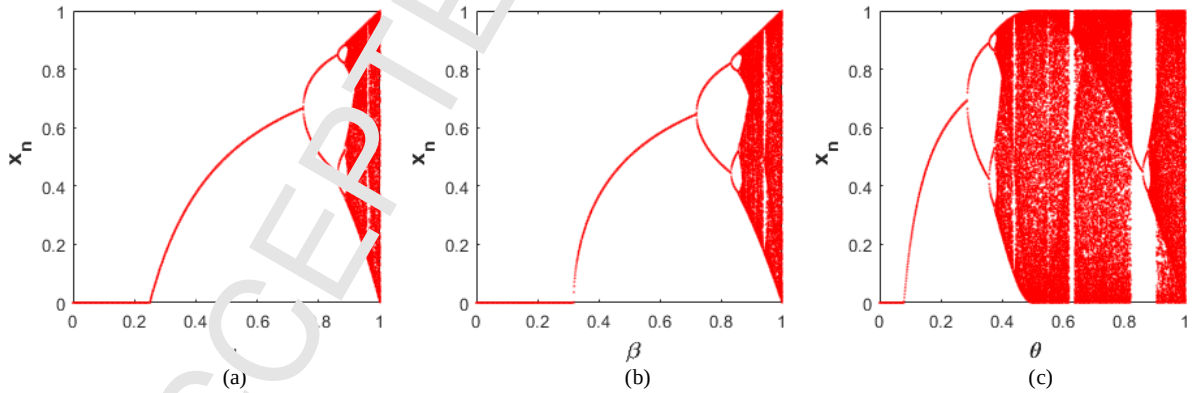
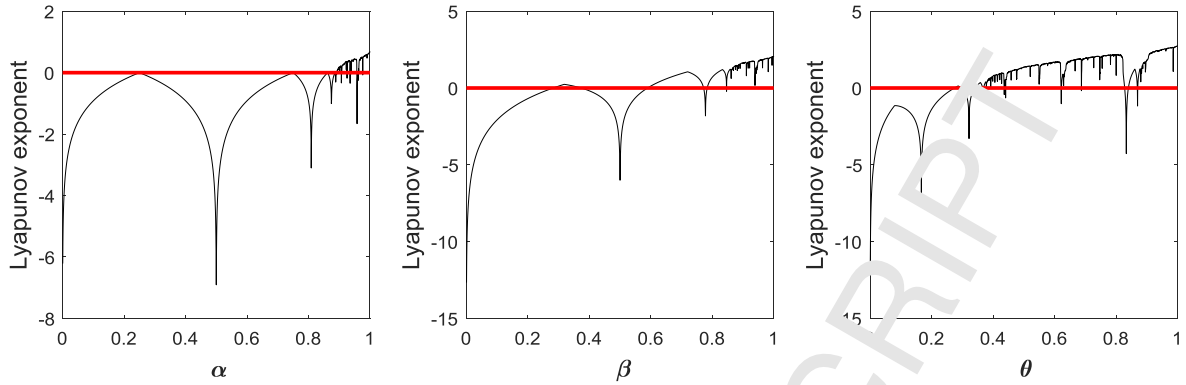


Fig. 1. Bifurcation diagrams corresponding to different chaotic mappings: (a) logistic mapping; (b) sine mapping; (c) 1D logistic-sine complex mapping

ACCEPTED MANUSCRIPT



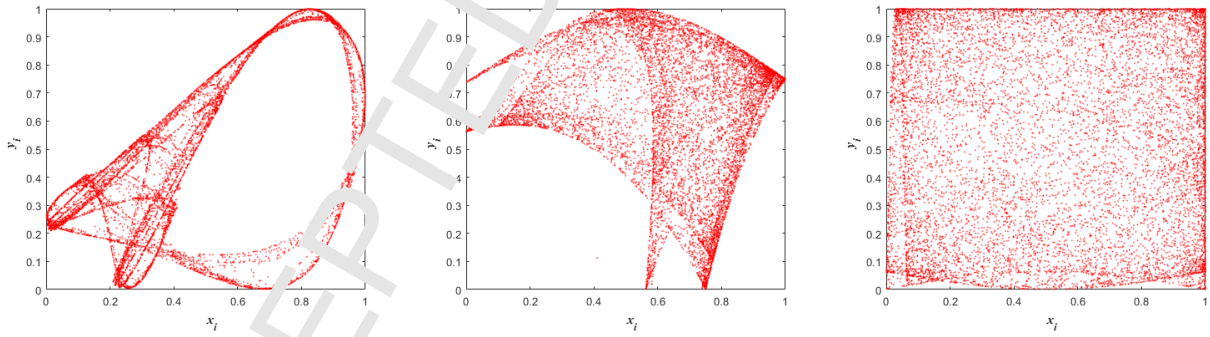
(a) (b) (c)

Fig. 2. The Lyapunov exponent corresponding to different chaotic mappings: (a) logistic mapping; (b) sine mapping; (c) 1D logistic-sine complex mapping.

②2D logistic-sine complex chaotic mapping:

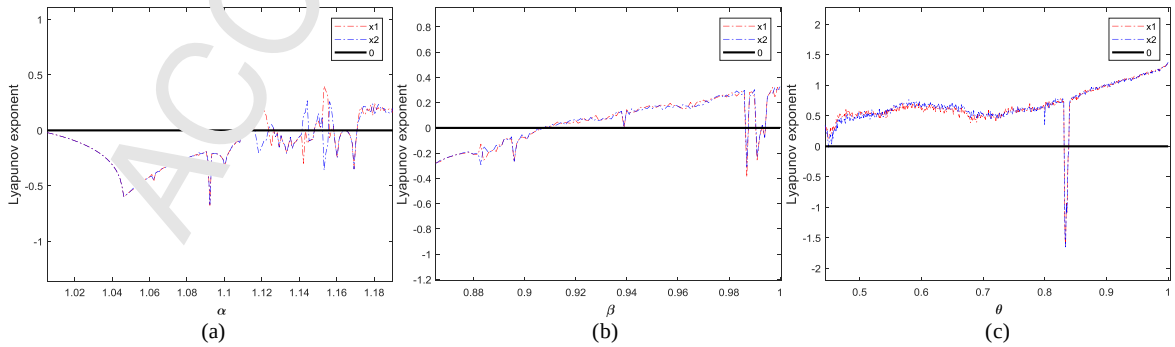
$$\begin{cases} x_{i+1} = \sin(\pi(4\theta x_i(1-x_i) + (1-\theta)\sin(\pi y_i))) \\ y_{i+1} = \sin(\pi(4\theta y_i(1-y_i) + (1-\theta)\sin(\pi x_{i+1}))) \end{cases} \tag{16}$$

where, θ is the chaos control parameter, and $\theta \in [0,1]$; x_n and y_n are the variables of the complex mapping, and $x_n \in [0,1]$, $y_n \in [0,1]$. In the 2D logistic-sine complex chaotic mapping whose two input parameters (x_i, y_i) are affected by each other, the output pairs (x_{i+1}, y_{i+1}) are distributed in the two-dimensional phase plane. Compared with the logistic mapping, the sine mapping, and 1D logistic-sine complex chaotic mapping, the 2D logistic-sine complex chaotic mapping has a more complex structure and has a stronger traversal. In addition, when $y_i = 0$ or $x_i = 0$, the 2D logistic-sine complex chaotic mapping does not degrade to a 1D logistic-sine complex chaotic mapping. It is shown that the 1D and 2D logistic-sine complex mappings are two different logistic-sine complex mappings. In addition, to further verify the performance of the 2D logistic-sine complex chaotic map, we continue to compare the chaotic orbit and Lyapunov exponents of chaotic maps with the 2D logistic map, 2D SLMM, and 2D LSCM, as shown in Fig. 3 and Fig. 4. It can be seen from Fig. 3 that the 2D-LSCM map has more complex chaotic orbits; further, from Fig. 4, we can see that the 2D-LSCM mapping has a larger Lyapunov exponent. Therefore, from the comparison results, the 2D-LSCM has better chaotic characteristics in the initial value sensitivity and ergodic property.



(a) (b) (c)

Fig. 3. The chaotic orbit of a complex mapping: (*a*) 2D logistic mapping; (*b*) 2D SLMM; (*c*) 2D LSCM.



(a) (b) (c)

ACCEPTED MANUSCRIPT

Fig.4. The Lyapunov exponent corresponding to different complex chaotic mappings: (*a*) 2D logistic mapping; (*b*) 2D SLMM; (*c*) 2D LSCM

3.2.2. Improved PIO based on variable-dimensional complex chaotic mappings

To solve the problem that the standard PIO algorithm easily obtains only the local optimal solution, three core steps of PIO are improved by using the variable dimensional chaotic mapping method. That is, in the first stage of PIO, the initial position and initial velocity of the pigeons are initialized with 1D logistic-sine complex mapping. In the second stage of PIO, the 2D logistic-sine complex mapping is used to update the location and speed of individual pigeons, and to get a better next-generation population. In the third stage of PIO, the 1D logistic-sine complex mapping is used to update the position of the individuals in the excellent population. Then, the global optimal solution will be obtained. The detailed algorithm process is shown as follows:

Algorithm 2: PIO implementation procedure

Input:

***N_P*:** number of individuals in pigeon swarm
***D*:** dimension of the search space
***R*:** the map and compass factor
***θ*:** the control parameters of LSCM
***X_1D*:** the parameters of 1D LSCM
***X_2D*, *Y_2D*:** the parameters of 2D LSCM
***S*:** the borders of the search space
***Nc_1max*:** the maximum number of generations that the map and compass operation is carried out
***Nc_2max*:** the maximum number of generations that the landmark operation is carried out.

Initialize:

$Nc_{1\max}=T_1$, $Nc_{2\max}=T_2$, $Np=p$, $D=d$, $X_{1D}=X_{2D}=Y_{2D}=rand$, $\theta=0.99$ **and** $S=[s_1,s_2]$

for N_p=1 **to** N_p **do**
  for D=1 **to** D **do**

$$X_i = s_1 + (\sin(\pi(4\theta X_{1D}^i(1-X_{1D}^i))) \times (s_2 - s_1)$$

$$V_i = \sin(\pi(4\theta X_{1D}^i(1-X_{1D}^i))$$

  end for
end for

$$X_p = X_i, N_c = 1$$

$$f(X_p) = fitness(X_p)$$

$$X_g := \arg\min[f(X_p)]$$

Map and compass operations:

for Nc=1 **to** $Nc_{1\max}$ **do**
  for i=1 **to** N_p **do**
    while $X_i > S$ **do**

$$\begin{cases} X_{2D}^{i+1}=\sin(\pi(4\theta X_{2D}^i(1-X_{2D}^i)+(1-\theta)\sin(\pi Y_{2D}^i)) \\ Y_{2D}^{i+1}=\sin(\pi(4\theta Y_{2D}^i(1-Y_{2D}^i)+(1-\theta)\sin(\pi X_{2D}^{i+1})) \\ V_i^{N_c}=V_i^{N_c-1}*\alpha^{X_{2D}^{i+1}\times N_c}+r_{2D}^{i+1}\times(X_{gbest}-X_i^{N_c-1}) \\ X_i^{N_c}=X_i^{N_c-1}+V_i^{N_c} \end{cases}$$

    end while
  end for
  evaluate X_i, and update X_p **and** X_g
end for

Landmark operations:

$X_{1D} = rand$

for Nc=$Nc_{1\max}$+1 **to** $Nc_{2\max}$ **do**
  while $X_p > S$ **do**

$$N_p = N_p / 2$$

$$X_i = X_i^{N_c-1} + (\sin(\pi(4\theta X_{1D}^i(1-X_{1D}^i))) \times (X_{center}^{N_c-1} - X_i^{N_c-1})$$

ACCEPTED MANUSCRIPT

end while

evaluate X_i, and update X_p and X_g

end for

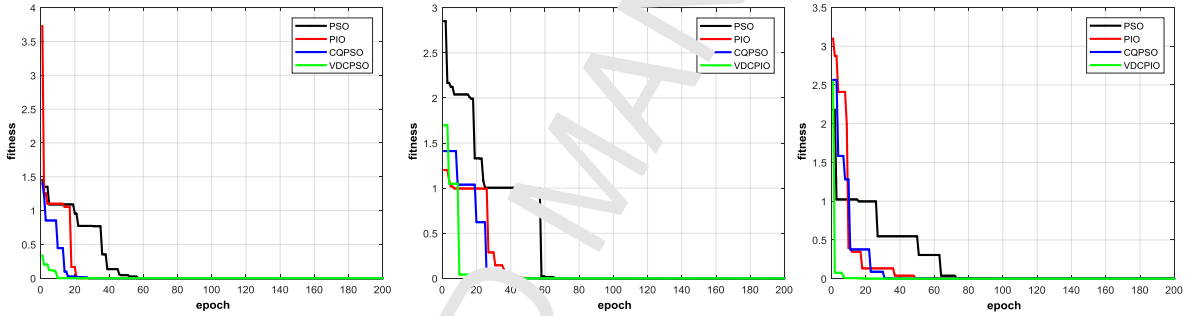
Output: X_{gbest}.

3.2.3. Performance comparison

To verify the convergence and optimization performance of VDCPIO, different algorithms, including the standard PIO, the standard PSO, CQPSO, and VDCPIO algorithms, are tested to compare their optimization performance. The objective functions are as follows:

$$\begin{cases} f_1(x) = 100(x_1^2 - x_2)^2 + (1 - x_1^2) \\ f_2(x) = 0.5 + \dfrac{\sin 2\sqrt{x_1^2 + x_2^2} - 0.5}{\left[1.0 + 0.001(x_1^2 + x_2^2)\right]^2} \\ f_3(x) = 4x_1^2 - 2.1x_1^4 + x_1^6 / 3 + x_1 x_2 - 4x_2^2 + x_2^4 \end{cases} \tag{17}$$

The initial population is set to 30. The maximum number of iterations is set to 200. The optimization result is shown in Fig. 5. From the figures, it can be observed that, for all three objective functions in Eq. (17), the evolution generation and convergence speed of the standard PIO algorithm are better than the standard PSO algorithm. The evolution generation and convergence speed of CQPSO algorithm is better than the standard PSO and standard PIO. The evolution generation and convergence speed of VDCPIO, proposed in this paper, are better than the others. Therefore, the experimental results indicate that the VDCPIO algorithm has better optimization performance.



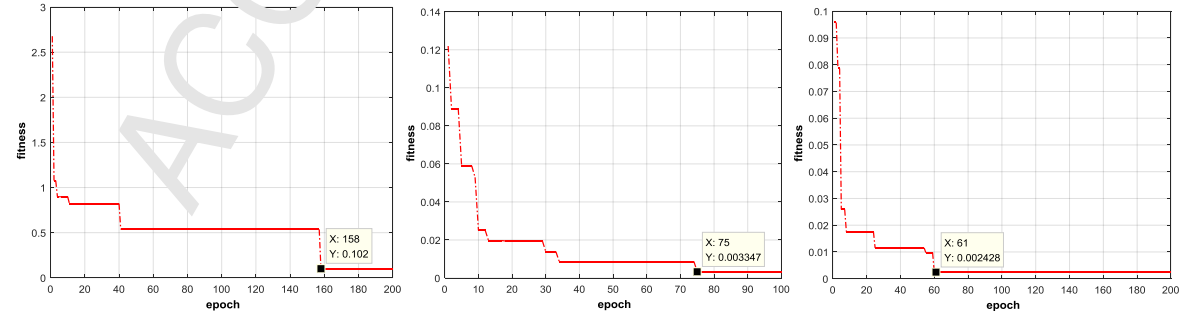
(a) (b) (c)

Fig. 5. Performance comparison of VDCPIO with standard PSO, standard PIO, CQPSO and other optimization algorithms for different test functions: (a) test functions $f_1(x)$; (b) test functions $f_2(x)$; (c) test functions $f_3(x)$.

Furthermore, to verify the effectiveness of the VDCPIO algorithm in solving the problem of local optimal solution, the $f_2(x)$ in Eq. (17) is taken as the test function. The test conditions are shown in Table 1, and the experimental results are shown in Fig. 6. The experimental results indicate that the standard PIO algorithm can easily obtain the local optimal solution when the dimension of the test function is large, the iterations of the PIO are insufficient, or the population of the PIO is insufficient. Meanwhile, the VDCPIO is used to test the same conditions. The experimental results are shown in Fig. 7, which indicate that the proposed VDCPIO can effectively overcome the local optimal solution.

Table 1
The local optimal solution of PIO under three different test conditions.

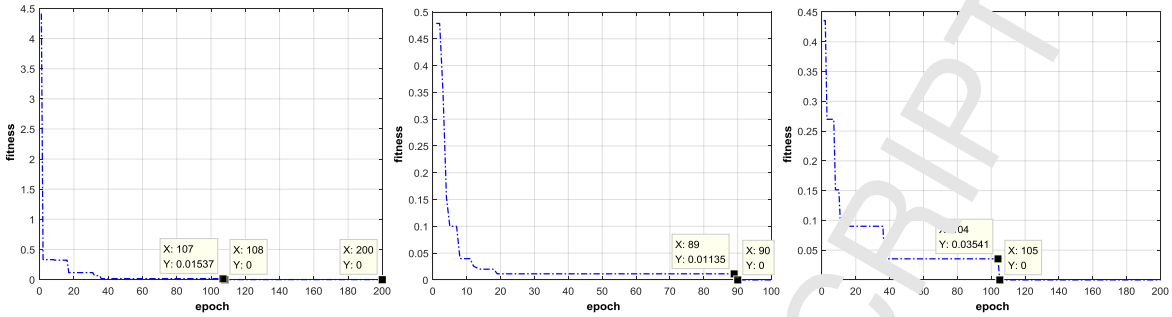
Test function	Dimension	Populations	Maximum iterations	Global optimum	Local optimum	Iterations
$f_2(x)$	30	30	200	0	0.1020	158
	0	30	100		0.0033	75
	0	10	200		0.0024	61



ACCEPTED MANUSCRIPT

(a) (b) (c)

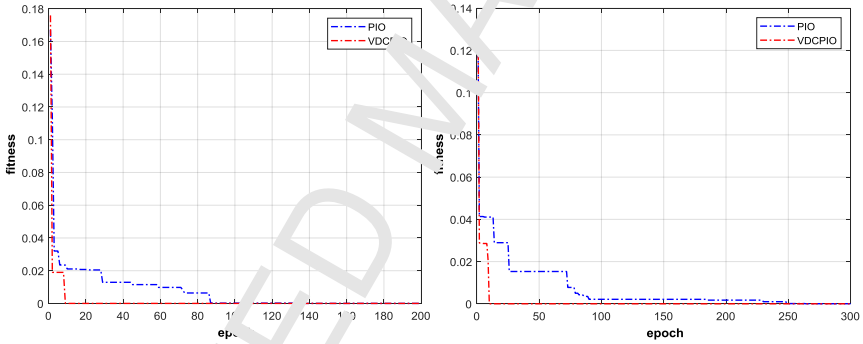
Fig. 6. Three cases of the PIO algorithm falling into the local optimal solution: (a) higher dimension of test function; (b) iteration is insufficient; (c) population is insufficient.



(a) (b) (c)

Fig. 7. Three cases of the VDCPIO algorithm falling into the local optimal solution: (a) higher dimension of test function; (b) iteration is insufficient; (c) population is insufficient.

According to the experimental results of PIO, it is easy obtain the local optimal solution when the parameter values of PIO are not properly set. Therefore, the local optimal solution can be avoided to some extent by adjusting the relevant parameter values of PIO rationally. Here, the parameter values of the PIO are adjusted to be further validated in two test conditions, as follows: ① the dimension of the test function is set to 30, populations are set to 60, and number of iterations is set to 200; ② the dimension of the test function is set to 30, populations are set to 30, and number of iterations is set to 300. Then, both PIO and VDCPIO are tested under these two conditions, and the results are shown in Fig. 8. The experimental results indicate that the local optimal solution can be effectively avoided by setting appropriate parameter values of PIO. Moreover, the convergence rate of the VDCPIO is significantly superior to that of PIO.



(a) (b)

Fig. 8. Optimization results of PIO and VDCPIO in two different test conditions: (a) test condition 1; (b) test condition 2.

Finally, the experimental results of the local optimal solution indicate when the parameter values of PIO are not suitable. It leads to the local optimal solution, and the VDCPIO does not obtain the local optimum under the same condition. In addition, the convergence speed of VDCPIO is significantly better than that of PIO when both VDCPIO and PIO do obtain the local optimal solution under the same conditions. By analyzing the optimization mechanism of PIO, the main reason for this is that the random number generated by PIO cannot complete the traversal search of the continuous variable space without repetition. Because the search space is limited, it is possible to obtain the local optimal solution. The goodness of pseudo-randomness, ergodicity, and initial sensitivity of chaos can overcome this defect of PIO. To further verify the validity and reliability of the proposed method, the experiment in Fig. 8 is repeated 100 times, and the experimental results are shown in Table 2. The experimental results from Table 2 indicate that in 100 times of repeated solution, the PIO obtains the local optimal solution 12 and 18 times, and the VDCPIO does so only 1 and 2 times. The experimental results indicate that there are many reasons for PIO to obtain the local optimal solution. The proposed VDCPIO method can effectively overcome this limitation.

Table 2

The number of times that local optimal solutions of PIO and VDCPIO were found under 100 repeated tests.

Test function	Algorithm	Local optimum	Global optimum	Iterations	Repetitions
$f_2(x)$	PIO	12	88	200	100
	VDCPIO	1	99		
	PIO	18	82	300	

ACCEPTED MANUSCRIPT

4 Fault feature extraction based on VDCPIO's optimal Teager-VMD algorithm

The vibration signals of large-scale, low-speed, and heavy-duty machinery are sparse: vibration signals with early faults include only a few signal components, such as frequency multiplication, frequency division, micro-impact, or even a single micro-impact component. This type of micro-impact or even a single micro-impact component is precisely better in reflecting the early fault features of the equipment. Because the impact components of early faults are extremely sparse and weak, it is not always possible to extract these fault features effectively if the VMD is used directly for decomposition. Therefore, before using VMD to decompose the vibration signal, the Teager energy operator is employed to enhance the shock component of the vibration signal and improve the signal-to-noise ratio of the signal. Then, the fitness function of VDCPIO is determined, and the combination value of the key parameters of VMD is searched. Finally, several IMFs are selected according to the average kurtosis, which is used to reconstruct a new signal. The characteristic frequency of the fault is extracted by analyzing the envelope spectrum of the reconstructed signal.

4.1. Teager energy operator

The equation of motion of a linear oscillator with no attenuation free oscillation:

$$m\ddot{x} + k\dot{x} = 0 \tag{18}$$

The general solution is a cosine function:

$$x(t) = A\cos(\omega t + \theta) \tag{19}$$

A simple mathematical method to analyze and track the energy of narrowband signals is the nonlinear energy tracking operator, referred to as "energy operator" and denoted as ψ. For a continuous signal $x(t)$, the definition of energy operator is given by

$$\psi_c[x(t)] = \left(\frac{dx(t)}{dt}\right)^2 - x(t)\frac{d^2x(t)}{dt^2} = [\dot{x}(t)]^2 - x(t)\ddot{x}(t) \tag{20}$$

Substituting $x(t)$ into the Eq. (20) yields $\psi_c[x(t)] = A^2\omega^2$, It can reflect and track changes in energy. The output of the Teager energy operator is the product of the instantaneous amplitude of vibration and the square of the instantaneous frequency. Compared with the traditional definition of energy, the product with the square of the frequency is increased. Because the vibration frequency of instantaneous impact is higher, the output of the Teager energy operator can effectively enhance the instantaneous impact composition.

4.2. Fitness function

As described in Section 2, the performance of VMD susceptible to the total number of mode K, the secondary penalty α, the updated step-size τ, and the convergence fault tolerance threshold ε. Because the actual vibration signal is complex and changeable, the relevant parameters of VMD are often difficult to determine. It is found that the performance of VMD in vibration signal decomposition is mainly influenced by two key parameters: the total number of modalities K and secondary penalty α. Therefore, the VDCPIO optimization algorithm is used to search the optimal combination values of K and α of VMD. Before optimization, the PIO algorithm needs to determine a suitable fitness function. Ideally, the VMD algorithm can decompose the complex vibration signals into a set of IMFs that are strictly pairwise orthogonal. At the same time, the summation of these IMFs can accurately reconstruct the original signal. However, in practical applications, the following conditions may be encountered. ① There is good orthogonality between each IMF, but the reconstruction error of the reconstructed signal is very large. The main reasons for this are as follows: when the value of α is low, the IMFs will capture additional "noise" information; when the value of α is high, the IMFs will lose some information. ② The accuracy of the reconstructed signal is very high, but the orthogonality between the IMFs is very poor. The main reasons for this are as follows: the value of K is low, which leads to the lack of signal segmentation, that is, information coupling between the IMF is serious. It is indicated that the orthogonality between the IMFs and the accuracy of signal reconstruction have some influence on each other. It is apparent that using the orthogonality of IMFs or the reconstruction accuracy of the signal as evaluation index of VMD performance cannot guarantee that the combined value of [K, α] of the VMD is optimal because of the problems mentioned above.

To address this problem, the ratio of $\sum_{i=1}^{n} V_{MI(\mathrm{IMF}(i),\mathrm{IMF}(i+1))}^{\mathrm{IMF}}$ to V_{MI} is used as the fitness function of VDCPIO. Its expression is given by

ACCEPTED MANUSCRIPT

$$fitness_{\mathrm{VDCPIO}} := \sum_{i=1}^{n} V_{MI(\mathrm{IMF}(i),\mathrm{IMF}(i+1))}^{\mathrm{IMF}} / V_{MI} \tag{21}$$

where V_{MI} is the mutual information of the reconstructed and original signals. The greater the value of mutual information, the higher the precision of the reconstructed signal and the smaller the reconstruction error. $\sum_{i=1}^{n} V_{MI(\mathrm{IMF}(i),\mathrm{IMF}(i+1))}^{\mathrm{IMF}}$ is the sum of the mutual information of each adjacent IMF. The smaller the mutual information, the smaller the correlation between the neighboring IMFs that is, the combined components of the signal can be well decomposed into a single IMF. Mutual information is a non-parametric and non-linear measure index in information theory, which can measure the correlation means to measure the amount of information shared between two variables. The measurement is not limited to a simple linear relationship, but the nonlinear relationship between variables can also be evaluated. Its expression is as follows:

$$\mathrm{MI}(\boldsymbol{x},\boldsymbol{y}) = \sum_{x\in x}\sum_{y\in y} p(\boldsymbol{x},\boldsymbol{y}) \lg \frac{p(\boldsymbol{x},\boldsymbol{y})}{p(\boldsymbol{x})p(\boldsymbol{y})} \tag{22}$$

where $p(\boldsymbol{x}, \boldsymbol{y})$ denotes joint distribution and $p(\boldsymbol{x})$ and $p(\boldsymbol{y})$ are marginal probabilities. If there is no overlapping information between $\boldsymbol{x}$ and $\boldsymbol{y}$, that is, x and y are independent, the mutual information value is equal to 0. Conversely, the higher the correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$, the closer the value of mutual information is to 1 [39].

Therefore, based on the above analysis, the ratio of $\sum_{i=1}^{n} V_{MI(\mathrm{IMF}(i),\mathrm{IMF}(i+1))}^{\mathrm{IMF}}$ to V_{MI} is used as the fitness function of PIO. When $\sum_{i=1}^{n} V_{MI(\mathrm{IMF}(i),\mathrm{IMF}(i+1))}^{\mathrm{IMF}}$ is minimum and V_{MI} is maximum, that is, their ratio is minimal, the signal can be correctly decomposed into a set of IMFs. Therefore, the combined value of [K, α] corresponding to the fitness value at this time is optimal, which can ensure that the VMD has the best decomposition performance.

4.3. Signal reconstruction based on kurtosis criterion

Large-scale, low-speed and heavy-load machinery operates at large loads and very low rotational speeds. When such machinery has an early fault, the energy of the shock component that characterizes the fault in the vibration signal is extremely sparse and weak. Thus, it is very difficult to directly analyze the spectrum of the vibration signal. Therefore, the kurtosis criterion is used to reconstruct a new signal according to the original vibration signal. Kurtosis reflects the numerical statistics of the distribution characteristics of the vibration signal, and it is a dimensionless parameter of the normalized fourth order central moment and the description of the waveform peak [40]. Its mathematical expression is given by

$$K = \frac{\sum_{i=1}^{n}(x_i - \mu)^4}{n\sigma^4} \tag{23}$$

where μ is the mean value of the signal x_i, o is the standard deviation of the signal x_i, and n is the length of the signal.

When the vibration signal has many shock components, the corresponding kurtosis also increases significantly. The greater the amount of shock, the greater the corresponding kurtosis. Thus, the sensitivity of kurtosis to shock signals can be used to detect the proportions or intensities of the shock components in IMF. In [41], the IMF with the greatest kurtosis is selected as the final analytical signal. However, the shock component that characterizes the fault in the vibration signal is generally distributed in different IMFs. Although the IMF with the greatest kurtosis contains more shock components or stronger shock components. But these shock components cannot fully characterize the features of early faults in large-scale, low-speed, and heavy-load machinery, such as the harmonic component formed by the fault shocking. Moreover, fault characteristic information, such as frequency doubling information, which is present in other IMF components, will be lost. Therefore, the average kurtosis of all IMFs is proposed as a screening indicator in this paper. The IMF components that do not contain or contain a very small number of shock components are eliminated based on the average kurtosis. Then, a new signal with the greatest shock component ratio or intensity is reconstructed using the remaining IMF components.

4.4. Fault feature extraction based on VDCPIO-based optimal VMD algorithm and Teager energy operator

Figure 9 shows the process of extracting fault features of the large-scale, low-speed and heavy-load machinery based on VDCPIO's optimal VMD algorithm and Teager energy operator. The detailed steps are listed as follows:

(1) Obtain the original vibration signal, and enhance the impact component of the vibration signal by using the Teager energy operator.

(2) Initialize the parameters of the variable dimension complex chaotic map, and determine the fitness

ACCEPTED MANUSCRIPT

function in the optimization process.

(3) Initialize the parameters of the pigeon population, use the VMD parameter combination [K, α] as the individual position of the pigeon group, and use the 1D-LSCM to generate a certain number of parameter combinations [K, α] as the initial position and velocity of the individual pigeon group.

(4) Under the condition of different pigeon individual positions, perform the VMD operation of the signal and calculate the corresponding fitness value of each pigeon position.

(5) Update the optimal values of the local optimum and population global by comparing the fitness value.

(6) Update the individual speed and position of pigeons by using formula (5).

(7) Repeat step 4 until the iteration number reaches the predefined set value, and obtain the best fitness value and pigeon individual position.

(8) After the optimization of the pigeon group is finished, a set of optimal parameter combination [K, α] is obtained. The combination value is used to set the parameters K, α of the VMD algorithm, and the optimized VMD algorithm is used to process the early fault signal.

(9) The IMF component is filtered according to the average kurtosis, and the most sensitive IMF component of the fault feature is obtained.

(10) The average kurtosis of the IMF components is calculated, and the IMF component that is larger than the average value is selected according to the average kurtosis.

(11) The filtrated IMF component is used to reconstruct the signal, and the reconstructed signal is demodulated by envelope demodulation. The corresponding envelope spectrum is obtained and matched with the fault characteristic frequency to judge the fault type.

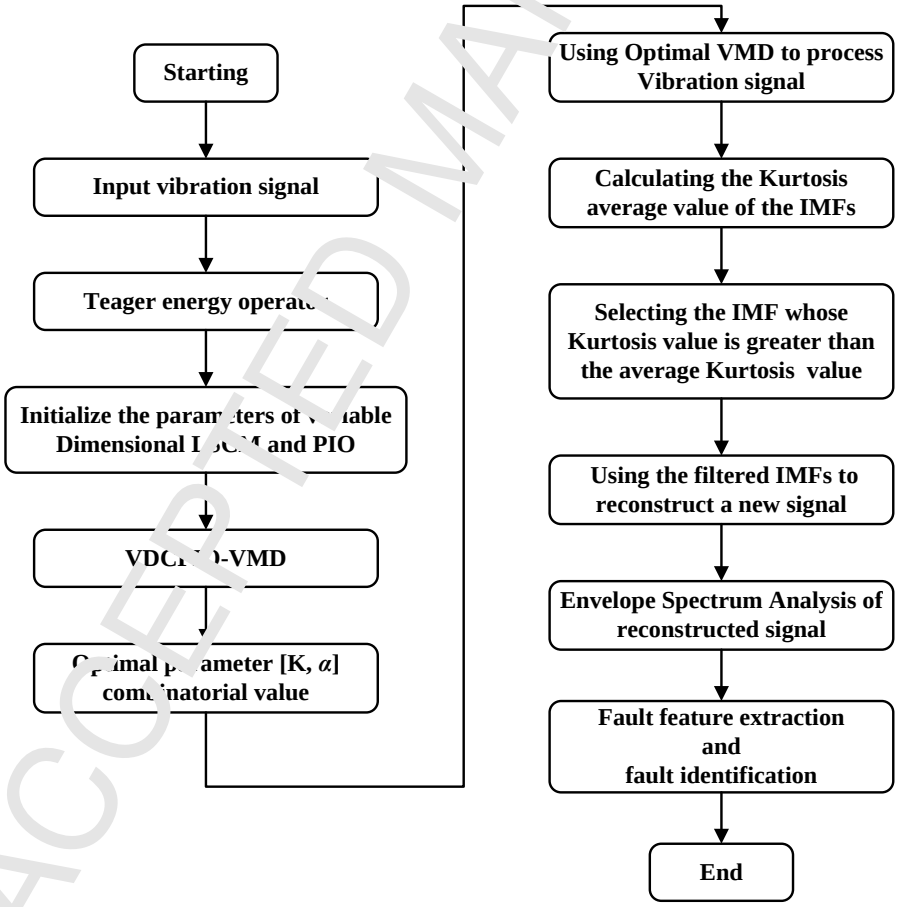


Fig. 9. Fault diagnosis process based on VDCPIO's Teager-VMD

5. Simulation and comparative analysis

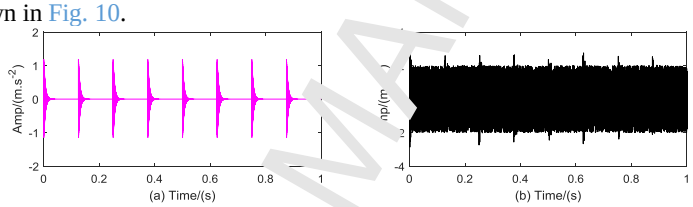
When pitting or defective faults occur in the bearings of large-scale, low-speed, and heavy-load machinery,

ACCEPTED MANUSCRIPT

periodic shock signals will be produced because of collision between the parts. In the middle or late stage of the fault, the fault frequency and its multiplier can be easily obtained using the traditional spectrum analysis because of the large amplitude of the shock signal. However, in the early stages of the fault, the shock signal of the fault is extremely sparse and weak, which is easily overwhelmed by strong noise. It is difficult to determine the fault characteristic frequency by using the traditional method. The vibration signal collected by the accelerometer will reveal that the shock signal caused by the fault produces a sinusoidal signal with exponential attenuation. Therefore, the vibration response caused by the simulated pulse is used as the fault signal of the bearing, and it is analyzed and verified. It is defined as follows:

$$\begin{cases} x(t) = s(t) + n(t) \\ s(t) = \sum_i A_i h(t - iT - \tau_i) \\ h(t) = e^{-Ct} \sin(2\pi f_n t) \\ A_i = 1 + A_0 \sin(2\pi f_r t) \end{cases} \tag{24}$$

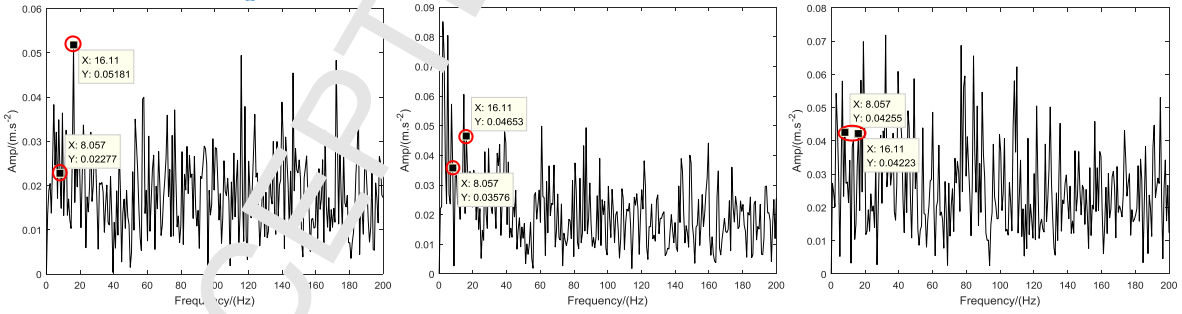
where $x(t)$ is a shock simulation signal with noise, $s(t)$ is a pure periodic shock signal, $n(t)$ is Gaussian white noise, T is the average period of the pulse, τ_i is the small fluctuation of the i-th shock relative to T, $h(t)$ is a damped exponential sinusoidal signal, C is the attenuation/damping coefficient, f_r is used to simulate the rotation frequency of the drive shaft, f_n is used to simulate the natural frequency of the system, A_0 is the initial amplitude of the shock signal, and A_i is the amplitude of the shock signal after i-th damping. Then the fault signal of the bearing can be simulated by setting the appropriate values for the relevant parameters of the simulation signal. The parameters are set as follows: C = 750, f_r = 1 Hz, f_n = 3000 Hz, A_0 = 0.5. In addition, f_{inner} denotes the simulated characteristic frequency of the bearing fault, and f_{inner} = 1/T = 8 Hz. The signal-to-noise ratio of $x(t)$ is set to be -11 dB. f_s is the sampling frequency and is set as 12 kHz. L is the data length of $x(t)$ and is set to be 12000. The simulated signal for the bearing fault is shown in Fig. 10.



(a) (b)

Fig. 10. Simulation signal of the fault of the inner ring of the bearing: (a) pure shock signal; (b) noisy shock signal.

As shown in Fig. 10, the amplitude of the pulse attenuation signal with noise is small, and it is basically submerged in the noise, which can simulate the weak vibration signal of the bearing's early fault. First, the simulation signal is processed using the traditional EEMD, LMD, and VMD algorithms, and the parameters are set (K = 4; α = 2000; τ = 0.3; ϵ = 1e-7) based on personal experience. Then, the corresponding envelope spectrum is obtained, as shown in Fig. 11.



(a) (b) (c)

Fig. 11. Envelope spectrum of the fault simulation signal of the inner ring of the bearing based on:(a) VMD; (b) LMD; (c) EEMD.

It can be observed from Fig. 11 that, in the envelope spectrum based on the corresponding VMD, LMD, and EEMD, the fault characteristic frequency (8 Hz) is completely submerged by other interference lines. Only the second harmonic frequency of the characteristic frequency of the simulated fault (16 Hz) in the envelope spectrum obtained by VMD is effectively extracted, and the spectral lines of the other two methods are almost completely submerged by the other interference lines. The results indicate that when the signal-to-noise ratio of the signal is low, which also means the shock signal is very weak, the simulation signal cannot be directly processed using the above signal method, and the characteristic frequency of the simulated fault cannot be effectively extracted.

The Teager energy operator is a nonlinear differential operator. The instantaneous value of the signal, as well as the non-linear combination of its differential, are used to estimate the total energy needed to generate the dynamic

ACCEPTED MANUSCRIPT

signal. It can enhance the transient characteristics of the signal, which also enhances the shock component of the signal. For this reason, before processing the signal, we use the Teager energy operator to enhance the shock components of the simulated signal in Fig. 12. Continue to use the standard VMD, LMD and EEMD to process the Teager energy signal is shown in Fig. 12. The corresponding envelope spectrum is shown in Fig. 13.

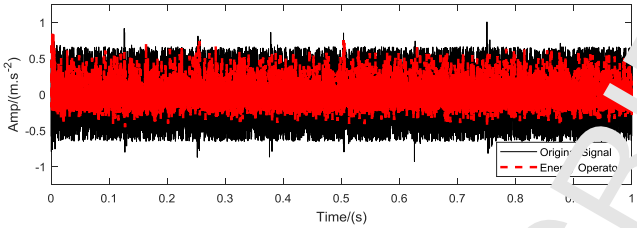
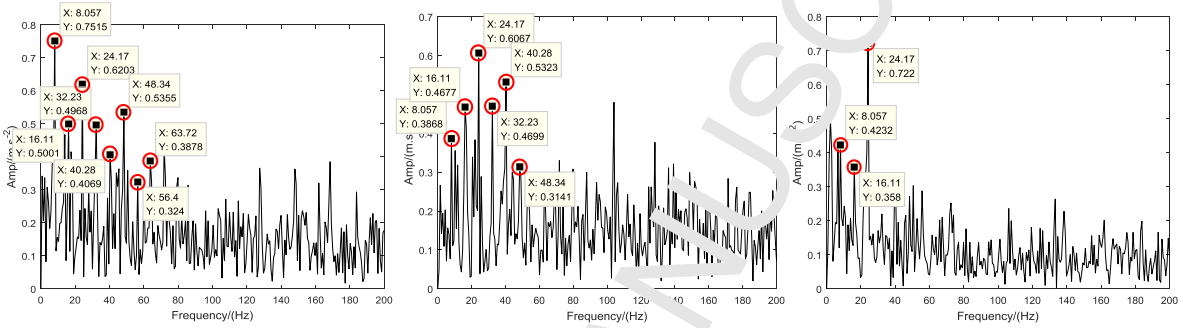


Fig. 12. Teager energy signal of the simulated signal of the fault of the inner ring of the bearing.



(a) (b) (c)

Fig. 13. Envelope spectrum of the fault simulation signal of the bearing based on the Teager energy operator: (a) VMD; (b) LMD; (c) EEMD.

Compared with the envelope spectrums shown in Fig. 11 and Fig. 13, it can be seen that the envelope spectrum of the fault simulation signal of the bearing pretreated by the Teager energy operator, its fundamental frequency, and the amplitude of the second harmonic frequency of the fault feature are larger than the envelope spectrum of the original simulation signal, and the frequency doubling component of the signal is richer. In addition, by comparing and analyzing the envelope spectrums corresponding to EEMD, LMD, and VMD, it can be seen that the fundamental frequency and amplitude of the second harmonic frequency of the fault characteristic frequency obtained by the VMD method are obviously larger than those obtained by the EEMD method, and slightly larger than that of the LMD method. Moreover, the multiple frequencies obtained by the VMD method are more abundant than those obtained by the LMD and EEMD methods. Therefore, analyzing and comparing the fault simulation signals proves that the VMD method has a better performance than the LMD and EEMD methods with respect to processing vibration signals. However, it can be observed from Fig. 11 and Fig. 13 that there are still many interference lines in the envelope spectrum obtained by the VMD method. By analyzing the principle of VMD, we know that the performance of VMD is closely related to its parameters, especially K and α, which have the greatest influence on VMD. Using a manual method to set the combination value of [K, α] of VMD has some blindness and contingency; hence, it is not possible to guarantee the best performance. Therefore, the VDCPIO algorithm is used to search the optimal combination value of [K, α] of VMD. Then, an optimal VMD algorithm is obtained. The detailed steps are described below:

(1) Input the simulation signal and use the Teager energy operator to enhance the shock component in the signal.

(2) Initialize the parameters of variable dimensional complex chaotic mappings, such as $\boldsymbol{\theta}$ = 0.99, $\boldsymbol{X_{1D}}$ = *rand*; $\boldsymbol{X_{2D}}$ = *rand*, and $\boldsymbol{Y_{1D}}$ = *rand*; then, determine the fitness function $\boldsymbol{fitness}_{\text{VDCPIO}}$ in the optimization process.

(3) Initialize the parameters of the pigeon populations, such as $\boldsymbol{Nc_{1max}}$ = 120, $\boldsymbol{Nc_{2max}}$ = 80, $\boldsymbol{N_P}$ = 30, D = 10, S = [2,4000], and use 1D-LSCM;

(4) Use the parameter combination of VMD [K, α] as the pigeon's individual position. Use 1D-LSCM to produce a certain number of parameter combination values of [K, α] to use these values as the initial position and velocity of individuals in the pigeon group. According to the standard deviation $\boldsymbol{Std}_{\text{signal}}$ of the simulation signal, set the updating step size of the VMD algorithm as τ = 0.003 and the convergence fault-tolerant threshold value as ϵ = 1e-7. Then, use 2D-LSCM to update the position and speed of pigeon swarm individuals in the global search stage. Finally, use 1D-LSCM to update the position of pigeon swarm individuals in the local search stage.

(5) Calculate the corresponding fitness value of each pigeon's position by VMD operation under different position conditions of the pigeon. Then, compare the fitness values, and update the individual local optimal value and the global optimum value of the population. Then, obtain a set of optimal combination value of [K, α].

ACCEPTED MANUSCRIPT

According to the optimal combination value f [K, α], reset parameters K and α of VMD.

(6) Process the simulation signal by the optimized VMD algorithm and obtain a number of IMF components. Then, reconstruct the new signal based on the average kurtosis of the IMFs to screen out the IMF components that are greater than the mean. Finally, analyze the envelope spectrum of the reconstructed signal.

(7) Compare the theoretical calculation value of the fault characteristic frequency of the simulation signal with the spectral line in the envelope spectrum.

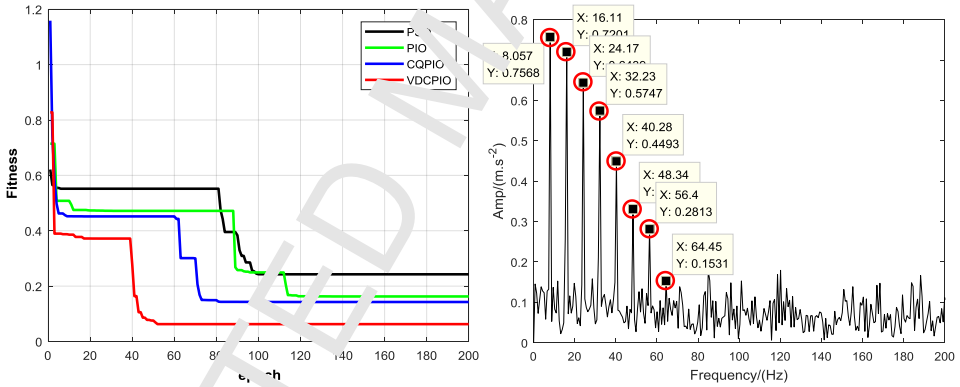
(8) Experimental validation.

Table 3 and Fig. 14(a) show the different intelligent optimization algorithms used to search the optimal parameter combination value of [K, α] of VMD. The results show that the proposed VDCPIO algorithm is superior to other algorithms with respect to iteration times and convergence speed. Then, the optimal combination value of [K, α] obtained using the VDCPIO algorithm is used to reset the corresponding parameters of VMD, and the simulated signal of the bearing fault is processed. The results are shown in Fig. 14(b). The amplitude of the fault characteristic frequency (8 Hz) and the amplitude of the multiple frequencies (2X-8X) of the spectral lines of fault features in Fig. 14(b) are greater than those in Fig. 13(a). Furthermore, the spectral lines are clearer and more prominent, and interference lines are less. Therefore, we can conclude that the characteristic frequency is extracted effectively, and that the proposed method is more effective and has a better performance.

Table 3

The iterations and fitness values are generated by using the different intelligent optimization algorithms to search for the optimal combination values of [**K, α**].

Parameters	Methods			
	PSO	PIO	CQPSO	**VDCPIO**
[K, α]	[6,1904]	[7,2275]	[9,2561]	[12,3169]
Iterations	102	119	82	52
Minimum of fitness	0.2427	0.1734	0.1431	0.0627

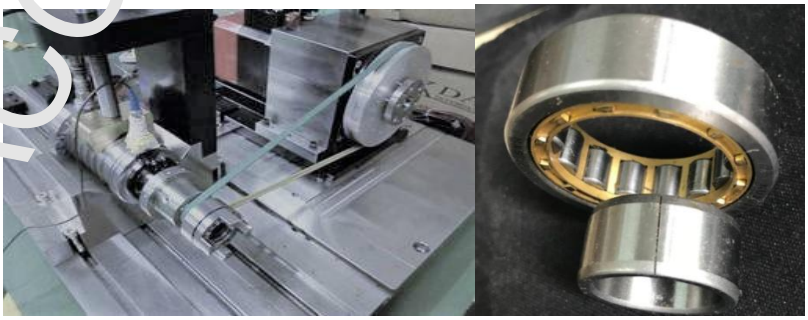


(a) (b)

Fig. 14. Envelope spectrum of the fault simulation signal based on VDCPIO's Teager-VMD algorithm: (a) search process of the optimal combination value of [**K**, $\boldsymbol{\alpha}$] of VMD; (b) envelope spectrum of the reconstructed signal.

6 Analysis and comparison of fault measured signals

To further verify the effectiveness of the proposed method in practical fault diagnosis, the vibration signals of low-speed and heavy-load machinery are simulated by controlling the load capacity and bearing speed of the fault diagnosis test-platform, as shown in Fig. 15.



ACCEPTED MANUSCRIPT

Fig. 15. Rolling bearing fault test platform

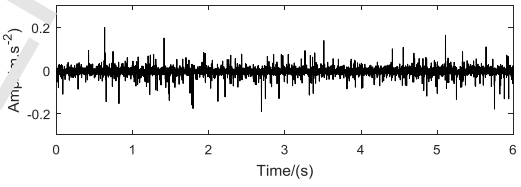
The pitch diameter of the bearing of the test platform is 39.5 mm, the diameter of the rolling body is 7.5 mm, the number of rolling bodies is 12, and the radial load force is 4000 N. The fault vibration signal of an inner ring rotating at 60 rpm is collected under heavy load., and the sampling frequency is 10 kHz. The theoretically calculated value of the fault characteristic frequency listed in Table 4.

Table 4
Theoretical calculation value of bearing fault characteristic frequency.

Speed(rpm)	Characteristic frequency of inner ring fault (f_{inner}/Hz)
60	7.14

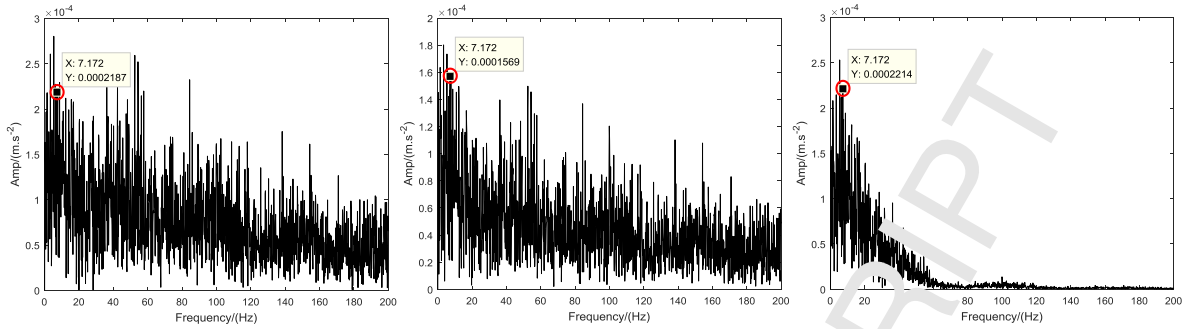
Now, we analyze the inner ring fault of the bearing. The fault signal of the inner ring of the bearing inner ring is processed by EEMD, LMD, and VMD, respectively, and the corresponding envelope spectrum is obtained, as shown in Fig. 16. It can be observed from Fig. 16 that the spectrum lines of the fault characteristic frequencies (the actual frequency is 7.17 Hz, slightly deviated from the theoretical calculation) in the envelope spectrums obtained using the standard EEMD, LMD, and VMD methods are almost submerged by other interference lines. The results show that the frequency of the bearing fault cannot be extracted directly using the above signal processing methods as the fault signal of a bearing operating under a large load and low speed is weak. Therefore, the shock components in the fault signal of the inner ring of the bearing are enhanced using the Teager energy operator. Then, EEMD, LMD, and VMD are used to process the enhanced fault signals, and the corresponding envelope spectrum was obtained, as shown in Fig. 17. It can be seen from Fig. 17 that the amplitude of the envelope spectrums obtained by the standard EEMD, LMD, and VMD methods are obviously increased, and the spectral lines of the fault characteristic frequency can be extracted effectively. In addition, the amplitude of the spectral line of the fault characteristic frequency extracted by the VMD method is the largest, and its interference spectral line is significantly less than those shown in Fig. 17(a) and Fig. 17(b). The results show that the fusion method based on the Teager energy operator and VMD algorithm can effectively extract the fault characteristic frequency of the inner ring of the bearing. However, there are several interfering spectral lines in the spectrum envelope, and the fault characteristic frequency is not obvious.

In this case, the VDCPIO algorithm is used to search the optimal combination value of [K, α] of VMD, and its search steps are consistent with the search steps mentioned in section 5. Except setting the updated step-size of VMD in accordance with the standard deviation of the signal, all the correlation parameters of chaotic mappings and pigeon swarm optimization algorithm remain unchanged. First, the optimal combination value of [K, α] = [13, 2217] is obtained, and the parameters of VMD are set. Then, VMD was used to process signal in Fig. 17(a). As the result, 13 IMF components (as shown in Fig. 18(a)) and their corresponding kurtosis [15.8, 20.5, 21.0, 21.3, 22.1, 20.3, 22.5, 18.1, 19.2, 22.0, 23.8, 21.9, 30.1] are obtained. The average kurtosis is 21.4. The IMF components having a kurtosis greater than the average kurtosis are selected to reconstruct a new signal (as shown in Fig. 18(b)), and the envelope spectrum of the reconstructed signal is analyzed (as shown in Fig. 18(c)). It can be observed from the envelope spectrum shown in Fig. 18(c) that this method can effectively extract not only the base frequency of the fault feature frequency (7.17 Hz) but also the double frequency (14.34 Hz) and triple frequency (21.51 Hz) of the fault feature frequency, and the interference spectral line is obviously reduced. Therefore, the experimental results show that the proposed method can effectively extract the early fault feature frequency of bearing under the low-speed and heavy-load condition. It provides an effective diagnosis basis for actual large-scale, low-speed, and heavy-load machinery.



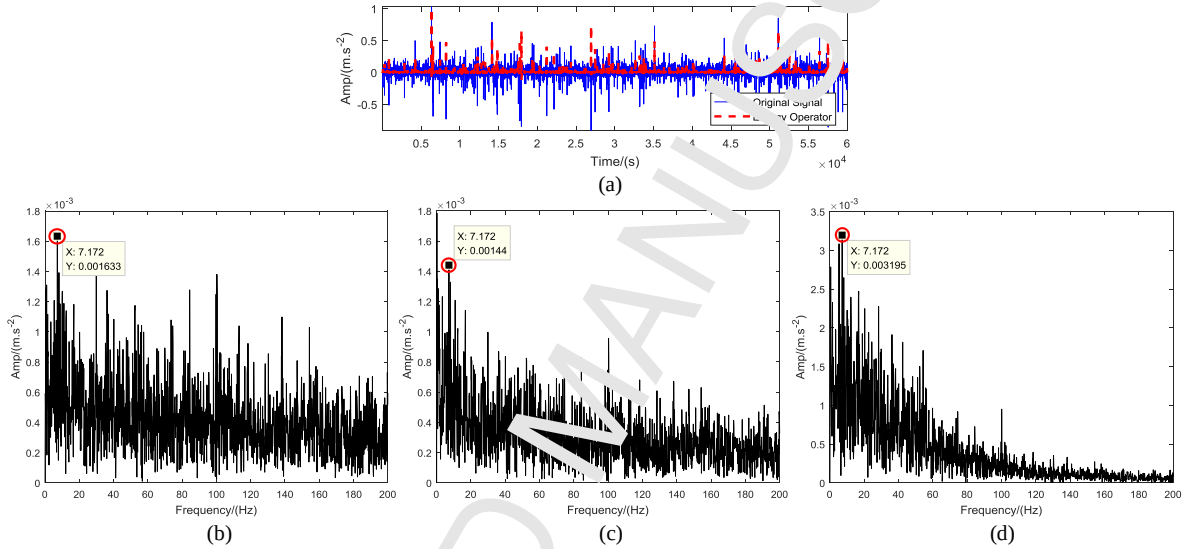
(a)

ACCEPTED MANUSCRIPT



(b) (c) (d)

Fig. 16. Waveform and envelope spectrum of fault signal of bearing inner ring: (a) Time domain waveform of the fault signal of the inner ring of the bearing; (b) Envelope spectrum based on EEMD; (c) Envelope spectrum based on LMD; (d) Envelope spectrum based on VMD.



(a)

(b) (c) (d)

Fig. 17. Fault waveform and envelope spectrum of the inner ring of the bearing are enhanced by the Teager energy operator: (a) Enhancing the waveform diagram of the fault signal of the inner ring of the bearing based on the Teager energy operator; (b) EEMD; (c) LMD; (d) VMD.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

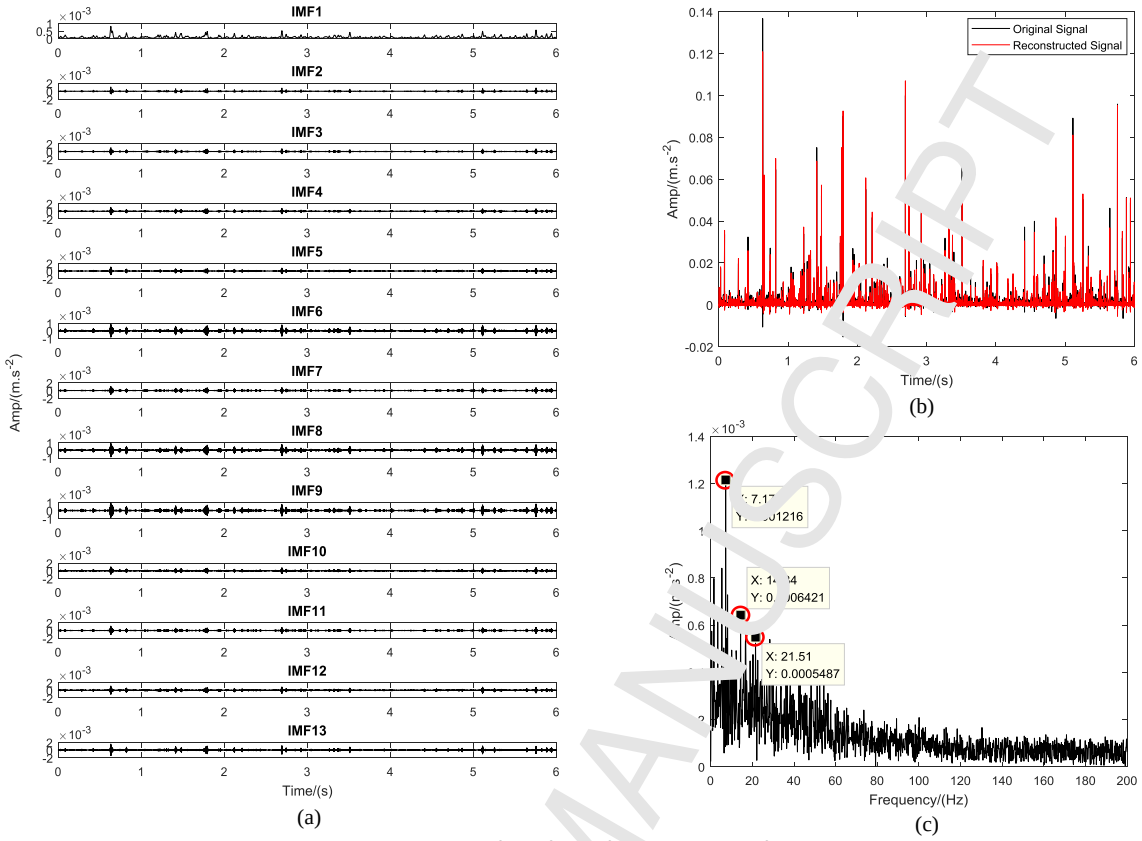


Fig. 18. Reconstructed signal and envelope spectrum of the fault of the inner ring of the bearing based on optimal VMD: (a) IMF components; (b) Reconstructed signal; (c) Envelope spectrum of reconstructed signal.

7. Summary

In this paper, the Teager energy operator and optimal variable mode decomposition (VMD) are combined as a hybrid method to extract the early fault features of large-scale, low-speed, and heavy-load machinery. The results of the experiment and application demonstrate the superiority of the proposed method. The conclusions are summarized as follows:

(1) A variable dimensional logistic-sine composite chaotic map method is proposed based on the logistic map and sine map. It is used to solve the problem that PIO algorithm is prone to trip in the local optimal solution. The experimental results indicate that it can effectively improve the global search ability and convergence speed of the PIO algorithm. Then, the improved PIO algorithm is used to search the optimal combination value of the key parameters [K, α] of the VMD to ensure that the VMD algorithm can obtain the best performance adaptively. To verify the effectiveness of the proposed optimal VMD method, the simulated signal of the bearing fault is further decomposed by the optimal VMD method, and the fault signal is reconstructed based on the average kurtosis criterion. The experimental results indicate that the proposed optimal VMD method has better performance in improving the signal-to-noise ratio (SNR), and can extract fault feature frequency more efficiently.

(2) It should be noted that the feature signal of the early fault of large-scale, low-speed and heavy-load machinery is extremely sparse and weak. This may cause the VMD to misjudge the fault signals as “noise”, rather than properly decomposing the fault signal into the corresponding IMF components. Using the Teager energy operator to directly process the original signal, the shock components in the signal can be enhanced. This helps the VMD to reasonably decompose the shock signals that characterize the fault features into the corresponding IMF components. The experiment results and applications indicate that the method combining the Teager energy operator and optimal VMD has better performance and advantages in extracting the early fault features of this kind of machinery.

(3) In the actual operation of large-scale, low-speed and heavy-load machinery, in addition to its low speed and heavy load, it may also operate intermittently. In such cases, the validity of the proposed method is to be verified.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 61174106), National Natural Youth Science Foundation of China under Grant (No. 51707079), and the Project Fund of the Hubei Provincial Department of Education (No. B2016006, 2016CFB463).

ACCEPTED MANUSCRIPT

References

[1] He, M. and He, D. Deep Learning Based Approach for Bearing Fault Diagnosis. IEEE Trans Ind App 2017; 53(3): 606-612.
[2] Wang, H. Q. and Chen, P. Fuzzy Diagnosis Method for Rotating Machinery in Variable Rotating Speed. IEEE Sens. Jour 2011; 11(1): 23-34.
[3] Md. Mizanur Rahman, Mohammad Nasir Uddin. Online Unbalanced Rotor Fault Detection of an IMD rive Based on Both Time and Frequency Domain Analyses. IEEE Trans Ind App 2017; 53(4): 4087-4096.
[4] Kedadouche, Mourad, Liu, ZH, Vu Viet-Hung. A new approach based on OMA-empirical wavelet transforms for bearing fault diagnosis. Measurement 2016; 90: 292-308.
[5] Liu X, Jia YX, He ZW, et al. Application of EMD-WVD and particle filter for gearbox fault feature extraction and remaining useful life prediction. Journal of Vibroengineering 2017; 19(3): 1793-1808.
[6] Cai JH, Hu WW. Feature extraction of gear fault signal based on Sobel operator and WHT. Shock and Vibration 2013; 20: 551–559
[7] Yang WX, Little Christian.; Court Richard. S-Transform and its contribution to wind turbine condition monitoring. Renewable Energy 2014; 62: 137-146.
[8] Zhang XH, Zhao JM, Bajrić Rusmir, et al. Application of the DC Offset Cancellation Method and S Transform to Gearbox Fault Diagnosis. MDPI 2017; 7(2): 1-22.
[9] Barbosa Tássio S, Ferreira Danton D, Pereira Daniel A. Fault Detection and Classification in Cantilever Beams Through Vibration Signal Analysis and Higher-Order Statistics. J Control Autom Electr Syst 2016; 27:535–541.
[10] Guo T, Deng ZM. An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing. Elsevier Ltd 2017; 127: 46-62.
[11] LI CW, Zhan LW, Shen LQ. Friction Signal Denoising Using Complete Ensemble EMD with Adaptive Noise and Mutual Information. Molecular Diversity Preservation International (MDPI) 2015; 17(9). 5965-5979.
[12] Colominas Marcelo A, Schlotthauer Gastón, Torres María E. Improved complete ensemble EMD A suitable tool for biomedical signal processing. Biomedical Signal Processing and Control 2014; 14: 19-29.
[13] Bacchetta Renato, Maran Barbara, Marelli Marcello, et al. A Morphological Hilbert-Huang Transform Technique for Bearing Fault Detection. Environmental Research 2016; 27(2): 2646-2656.
[14] Li YB, Xu MQ, Wang RX, et al. A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy. Shock and Vibration 2016; 360: 277-299.
[15] Guo W, Huang, LJ, Chen C. Elimination of end effects in local mean decomposition using spectral coherence and applications for rotating machinery. Digital Signal Processing 2016; 55: 52-63.
[16] Liu ZW, He ZJ Guo W, et al. A hybrid fault diagnosis method based on second generation wavelet denoising and local mean decomposition for rotating machinery. ISA Transactions 2016; 61: 211-220.
[17] Xing ZQ, Qu JF, Chai Y, et al. Gear fault diagnosis under variable conditions with intrinsic time-scale decomposition-singular value decomposition and support vector machine. Korean Society of Mechanical Engineers 2017; 31(2): 545-553.
[18] Duan LX, Yao MC, Wang JJ, et al. Integrative intrinsic time-scale decomposition and hierarchical temporal memory approach to gearbox diagnosis under variable operating conditions. Advances in Mechanical Engineering 2016; 8(8): 1-14.
[19] Konstantin Dragomiretskiy, Dominique Zosso. Variational Mode Decompositio. IEEE Trans Sign Prog 2014; 62(3): 531-544.
[20] Li K, Su L, Wu JJ, et al. A Rolling Bearing Fault Diagnosis Method Based on Variational Mode Decomposition and an Improved Kernel Extreme Learning Machine. MDPI 2017; 7(10): 1-20.
[21] Li Q, Liang Steven Y. Incipient Fault Diagnosis of Rolling Bearings Based on Impulse-Step Impact Dictionary and Re-Weighted Minimizing Nonconvex Penalty Lq Regularization Technique. MDPI 2017; 19(8): 1-20.
[22] ZHANG Ming, JIANG Zhinong, FENG Kun. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. Elsevier Ltd 2017; 93: 460-493.
[23] Li YB, Li GY, Wei Y. Health condition identification of planetary gearboxes based on variational mode decomposition and generalized composite multi-scale symbolic dynamic entropy. ISA Transactions 2018.
[24] Liu CF, Zhu LD, Ni CB. Chatter detection in milling process based on the energy entropy of VMD and WPD. Mechanical Systems and Signal Processing 2018; 105: 169-182.
[25] Xiao HS, Wei JC, Liu HS. Identification method for power system low-frequency oscillations based on improved VMD and Teager-Kaiser energy operator. The Institution of Engineering and Technology 2017; 11(16): 4096-4103.
[26] Shi P, Yang WX. Precise feature extraction from wind turbine condition monitoring signals by using optimized variational mode decomposition. IET Renewable Power Generation 2017; 11(3): 245-252.
[27] Long JC, Wang XP, Dai DD, et al. Denoising of UHF PD signals based on optimized VMD and wavelet transform. The Institution of Engineering and Technology 2017; 11(6): 753-760.
[28] Yi CC, Lv Y, Dang Z. A Fault Diagnosis Scheme for Rolling Bearing Based on Particle Swarm Optimization in Variational Mode Decomposition. Shock and Vibration 2016; 2016: 1-10.
[29] Duan HB, Qiao PX. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning. International Journal of Intelligent Computing and Cybernetics 2014; 7(1): 24-37.
[30] Wu Y, Yao JY, Qu XJ, et al. Adaptive reentry guidance method considering the influence of blackout zone. Acta Astronautica, 2018, 142: 253-267.
[31] Xue Q, Duan HB. Robust Attitude Control for Reusable Launch Vehicles Based on Fractional Calculus and Pigeon-inspired Optimization. IEEE/CAA Journal of Automatica Sinica 2017; 4(1): 89-97.
[32] Zhang G, Xie T, Zhang L. "In-Process Type" Dynamic Muskingum Model Parameter Estimation Method. Water 2017; 9(11): 849.
[33] Duan HB, Wang DB. Echo State Networks With Orthogonal Pigeon-Inspired Optimization for Image Restoration. IEEE Trans Neu Net and Lea Sys 2016; 27(11): 2413-2425.
[34] Gangireddy Srushnigdha, Ashok Joshi. Evolutionary method based integrated guidance strategy for reentry vehicles. Engineering Applications of Artificial Intelligence 2018; 69: 168-177.
[35] Xu XB, Deng YM. UAV Power Component-DC Brushless Motor Design with Merging Adjacent-Disturbances and Integrated-Dispatching Pigeon-Inspired Optimization. IEEE Trans Mag 2018; 54(8):
[36] Li CC, Luo GC, Qin K. An image encryption scheme based on chaotic tent map. Springer Netherlands 2017; 87: 127-133.
[37] Hua ZY, Zhou YC, Pun CM. 2D Sine Logistic modulation map for image encryption. Information Sciences 2015; 297: 80-94.

ACCEPTED MANUSCRIPT

[38] Hua ZY, Zhou YC. Image encryption using 2D Logistic-adjusted-Sine map. Information Sciences 2015; 339: 237-25.
[39] Khan Aimal, Rehman Saad, Abbas Muhammad, Ahmad Ayaz. On the mutual information of relaying protocols. Physical Communication 2018; 30: 33-42.
[40] Quitadamo Lucia Rita, Mai Roberto, Gozzo Francesca, et al. Kurtosis-Based Detection of Intracranial High-Frequency Oscillations for the Identification of the Seizure Onset Zone. International Journal of Neural Systems 2018; 28(7): 1850001-1850018.
[41] Zhao HS, Li L. Fault diagnosis of wind turbine bearing based on variational mode decomposition and Teager energy operator. IET Renewable Power Generation 2016; 11(4): 453-460.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

*Highlights (f

Research highlights:

1. The proposed variable dimensional chaotic PIO algorithm (VDCPIO) is constructed based on the variable dimensional composite chaotic mapping and the pigeon-inspired optimization (PIO) algorithm. Compared with the traditional PIO and PSO algorithm, it has better global searching ability and convergence speed.

2. Compared with wavelet, EEMD, LMD, ITD and other signal processing methods. Variational mode decomposition (VMD) is more suitable for analyzing nonlinear and non-stationary vibration signals. However, VMD's performance is greatly affected by the parameters [K, α]. The optimal VMD is obtained by using the VDCPIO algorithm to search the optimal combination values of the parameters [K, α].

3. The feature signal of early fault of large-scale low-speed and heavy-load machinery is extremely sparse and weak. The Teager energy operator is used to enhance the shock components in the original vibration signal, which helps the optimal VMD method to properly decompose the weak shock signal characterizing the fault feature into the corresponding IMF components.

4. Compared with the traditional EEMD, LMD and VMD methods. The Teager-VMD method based on VDCPIO parameter optimization, which has a distinct advantage in the feature extraction of early fault of large-scale low-speed and heavy-load machinery.

ACCEPTED MANUSCRIPT