

# Rolling bearing fault diagnosis based on adaptive smooth ITD and MF-DFA method

Journal of Low Frequency Noise,  
Vibration and Active Control  
2020, Vol. 39(4) 968–986  
© The Author(s) 2019  
DOI: 10.1177/1461348419867012  
journals.sagepub.com/home/lfn



Zhe Yuan<sup>1</sup>, Tingting Peng<sup>1</sup>, Dong An<sup>1,2</sup> , Daniel Cristea<sup>3</sup> and Mihai Alin Pop<sup>3</sup>

## Abstract

To effectively utilize a feature set to further improve fault diagnosis of a rolling bearing vibration signal, a method based on multi-fractal detrended fluctuation analysis (MF-DFA) and smooth intrinsic time-scale decomposition (SITD) was proposed. The vibration signal was decomposed into several proper rotation components by applying this new SITD method to overcome noise effects, preserve the effective signal, and improve the signal-to-noise ratio. Wavelet analysis was embedded in iteration procedures of intrinsic time-scale decomposition (ITD). For better results, an adaptive threshold function was used for signal recovery from noisy proper rotation components in the wavelet domain. Additionally, MF-DFA was used to reveal the multi-fractality present in the instantaneous amplitude of the proper rotation components. Finally, linear local tangent space alignment was applied for feature dimension reduction and to obtain fault characteristics of different types, further improving identification accuracy. The performance of the proposed method is determined to be superior to that of the ITD-MF-DFA method.

## Keywords

Smooth intrinsic time-scale decomposition, linear local tangent space alignment, multi-fractal detrended fluctuation analysis, vibration signal, bearing fault diagnosis

## Introduction

Because of complex operating environmental and objective factors, rolling bearing failures frequently occur in industrial processes. Safety insurance in production, improved quality, and the overall economic benefits of the product are very critical issues of modern industrial processes. A good solution to these problems is the application of effective performance monitoring and fault diagnosis processes.

Feature extraction characterizes the running state of rolling bearings and, thus, is one of the key research methods of rolling bearing fault diagnosis. According to the analysis of the vibration mechanism, the functional model corresponding to the classification of fault severity was established.<sup>1</sup> The non-linear and non-stationary characteristics of the signal and the interference of external factors on the obtained vibration signal are all factors that affect the extraction of features from complex vibration signals.<sup>2-5</sup>

For this purpose, a large number of articles on feature extraction methods can be found in the literature. As a classical time–frequency analysis method, wavelet transform (WT) is applied to mechanical fault diagnosis for its multi-scale analysis features, high time–frequency resolution, and rigorous mathematical foundation. Lou and

<sup>1</sup>School of Mechanical Engineering, Shenyang Jianzhu University, Shenyang, China

<sup>2</sup>Research Center for Analysis and Detection Technology, Shenyang Jianzhu University, Shenyang, China

<sup>3</sup>Materials Science Department, Transilvania University, Brasov, Romania

## Corresponding author:

Dong An, School of Mechanical Engineering, Shenyang Jianzhu University, Shenyang 110168, China.

Email: andong@sjzu.edu.cn



Loparo used WT to process accelerometer signals,<sup>6</sup> in which the standard deviation of wavelet coefficients was used to generate eigenvectors. Then, a neural fuzzy reasoning system was used as a classifier for bearing fault diagnosis. Liu and Ling proposed a fault diagnosis method based on wavelet packet theory.<sup>7</sup> The coefficients were taken as the pattern features of a ball-bearing fault, which greatly improved the effectiveness of pattern recognition. An improved Hilbert-Huang transform method based on WT was presented by Peng et al.,<sup>8</sup> which overcame the inevitable defects of the two methods when used separately for rolling bearing fault detection. Researchers have achieved much progress in fault detection and diagnosis methods based on WT.<sup>9-12</sup> With the purpose of improving the adaptability of wavelet applications, an adaptive threshold function is presented in this study, which was used to improve the efficiency of the denoising procedure. Because of the addition of shape adjustment factor  $m$ , the function becomes a comprehensive threshold function.

Recently, researchers have proposed an improved matching pursuit algorithm which has greater diagnosis precision in rolling bearing fault diagnosis.<sup>13</sup> New adaptive time–frequency analysis methods have been developed and have become a topic of concern in this field, namely, the empirical mode decomposition (EMD) and intrinsic time-scale decomposition (ITD).<sup>14</sup> An et al. proposed a fault diagnosis method based on ITD to determine the fault types of the wind turbine bearing.<sup>15</sup> In this case, the frequency center of the main proper rotation component was regarded as the fault feature vector. Yu and Liu applied a sparse coding shrinkage method based on ITD to fault diagnose for rolling bearings and effectively extracted the impulse component of the bearing vibration signal.<sup>16</sup> The ITD method is superior to the EMD method in reducing invalid components and mode mixing,<sup>14</sup> and the calculation time of the ITD method is much lesser than that of the EMD method, which significantly improves the work efficiency in practical applications. Nevertheless, its disadvantages should not be ignored, the ITD method is based on the linear transformation of signals to extract the baseline signal, which may lead to distortion, the production of burr, and end effects. To solve this drawback, many schemes have been proposed.<sup>17,18</sup> Hu et al. proposed a new approach named ensemble intrinsic time-scale decomposition (EITD) to deal with end effects and avoid distortion.<sup>19</sup> The steps included using cubic spline interpolation to fit baseline control points and mirror extension.

However, these improved ITD methods cannot meet the requirement of fault diagnosis perfectly. First, these methods use predictive models to improve classical ITD methods by spreading each single side of the signal, while a signal expansion method, such as that used with images, is imperfect in eliminating boundary distortion. In addition, prediction errors will occur during the screening process and destroy all proper rotation components (PRCs). Second, the mechanical fault signal in the initial period tends to exhibit a low signal-to-noise ratio (SNR) and small amplitude, and the traditional ITD method based on signal expansion is verified by analyzing the pure signal, whereby the weak signal with low SNR is not considered. Third, the portion of the signal that is extended by experience cannot reflect the true characteristics of the original weak signal. To solve the problem of noise interference and improving SNR, a method combining ITD and WT is proposed in this study. A correlation coefficient is used in the sifting process to show the iterative error and correlation relationship.

The fractal features hidden in the signal reflect the dynamic mechanism of the non-linear system under different states. Based on multi-fractal detrended fluctuation analysis (MF-DFA), a number of fault diagnoses and detection methods have been developed.<sup>20-22</sup> In this study, MF-DFA is applied to reveal the intrinsic multi-fractality in every extracted PRC. Finally, linear local tangent space alignment (LLTSA)<sup>23</sup> is implemented to reduce the value of dimension of the features, which will improve the accuracy of diagnosis.

In this paper, the concepts of ITD and MF-DFA are reviewed, and an adaptive threshold function is shown. The proposed methods of eliminating noise and restraining the end effects are discussed in detail, and a comparison of the proposed method and the original ITD method is shown. Moreover, rolling bearing signal decomposition based on smooth intrinsic time-scale decomposition (SITD) is studied, while MF-DFA and LLTSA are utilized to obtain the feature vector.

## Description of ITD method and MF-DFA method

The ITD method, a new self-adaptive signal decomposition method, is especially useful to analysis non-stationary or non-linear signals. MF-DFA is able to eliminate trend interference effectively and represent the intrinsic multi-fractal property. In this study, MF-DFA is applied to reveal the intrinsic multi-fractality in every extracted PRC.

### The ITD data decomposition method

The ITD method decomposes any non-stationary vibration complex signal  $X_t$  into a series of PRCs with practical physical meaning and a monotonous trend signal.  $\mathcal{L}t$  is a baseline extraction operator, used to extract a baseline signal from  $X_t$  and make the residual be in a proper rotation, as follows

$$X_t = \mathcal{L}X_t + (1 - \mathcal{L})X_t = L_t + H_t \quad (1)$$

where  $L_t = \mathcal{L}X_t$  is the baseline signal, and  $H_t = (1 - \mathcal{L})X_t$  indicates a proper rotation factor.

Suppose  $\{X_t, t \geq 0\}$  is a real-valued signal, and  $\{\tau_t, k = 1, 2, \dots\}$  are set to denote the local extrema of  $X_t$ ,  $\tau_k$  is the corresponding moment, and  $\tau_0 = 0$ . For the purpose of simplifying the definition, let  $X_k$  and  $L_k$  represent  $X(\tau_k)$  and  $L(\tau_k)$ , respectively.

It is assumed that  $X_t$  is available on  $[0, \tau_{k+2}]$ , and  $L_t$  and  $H_t$  have been defined on  $[0, \tau_k]$ . Intentionally, on the interval  $[\tau_k, \tau_{k+1}]$ , a piece-wise linear baseline-extracting operator  $\mathcal{L}$  can be expressed as follows

$$\mathcal{L}X_t = L_t = L_k + \frac{L_{k+1} - L_k}{X_{k+1} - X_k}(X_t - X_k), \quad t \in [\tau_k, \tau_{k+1}] \quad (2)$$

where

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

and  $\alpha$  is the linear gain in the interval  $(0, 1)$ , with a typical value of 0.5.

For signal  $X_t$ , the baseline-extracting operation is performed, and the first baseline signal,  $L_t^1 = \mathcal{L}X_t$ , is obtained. Then, the first PRC,  $H_t^1$ , is obtained by subtracting the baseline signal from the signal  $X_t$ , that is

$$H_t^1 = hX_t = (1 - \mathcal{L})X_t = X_t - L_t^1 \quad (4)$$

The baseline signal is set as the input signal when it is not a monotonic function. The above process is repeated  $P$  times until a drab trend signal is obtained. More precisely

$$\begin{aligned} X_t &= hX_t + \mathcal{L}X_t = hX_t + (h + \mathcal{L})\mathcal{L}X_t = \{h(1 + \mathcal{L}) + \mathcal{L}^2\}X_t = \left[ h \sum_{k=0}^{p-1} \mathcal{L}^k + \mathcal{L}^p \right] X_t \\ &= H_t^1 + H_t^2 + \dots + H_t^p + L_t^p \end{aligned} \quad (5)$$

where  $L_t^p$  denotes the final monotonic trend and  $H_t^k$  ( $k=0,1,2,\dots,p$ ) stands for a set of proper rotation components with a frequency range from high to low.

The ITD method represents a significant advancement in EMD, because the linear interpolation operation and selection process in EMD decomposition are eliminated. It improves upon the conventional method of EMD in terms of efficiency and accuracy, completing in a single step what the EMD may require numerous iterations to achieve. Instantaneous amplitude and frequency of PRC reflect the time–frequency information of the contaminated signals in real time.

The components decomposed by the ITD method have a certain physical significance. They are suitable for analyzing signals containing Amplitude Modulation-Frequency Modulation components, because the characteristics of the original signal are well reflected. Therefore, it is more suitable for vibration signal feature extraction and fault diagnosis of bearings and gearboxes.

### MF-DFA—A method for eliminating trends

It overcomes the deficiencies of DFA which must be used for stationary time series.

The MF-DFA procedure is shown in detail:

1. Assume that  $x_k$  is a time series with a length of  $N$ , and “profile” is defined as

$$Y(i) = \sum_{k=1}^i |x_k - \langle x \rangle|, \quad i = 1, \dots, N \quad (6)$$

$$\langle x \rangle = \frac{1}{N} \sum_{k=1}^N x_k \quad (7)$$

The profile  $Y(i)$  is divided into  $N_s = \text{int}(\frac{N}{s})$  disjoint segments with the same length  $s$ . Generally, the length  $N$  of the series cannot be divisible by segment length. Thus, to take advantage of the entire data set, the same procedure should be repeated from the opposite end. Hence,  $2N_s$  segments are obtained.

2. The local trend for every segment is fitted by applying the least-squares algorithm. Then, the mean square error is calculated, which is determined as

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^s \{ Y[(v-1)s + i] - y_v(i) \}^2 \quad (8)$$

for the  $v$ -th segment,  $v = 1, \dots, N_s$  and

$$F^2(s, v) = \frac{1}{s} \sum_{i=1}^s \{ Y[(N - (v - N_s))s + i] - y_v(i) \}^2 \quad (9)$$

In the  $v$ -th segment,  $y_v(i)$  is the fitting polynomial. Linear, quadratic, cubic, or higher order polynomials are employed in the fitting procedure (defined as DFA1, DFA2, DFA3, or etc., respectively).

Different orders of DFA may yield different detrended results that show the different abilities of eliminating trends in the series.

3. Define the  $q$ -th order fluctuation function by calculating the average of all segments

$$Fq(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(s, v)]^{\frac{q}{2}} \right\}^{\frac{1}{q}} \quad (10)$$

when  $q = 2$ , the procedure is the same as the standard DFA.

For different timescales  $s$ , repeat steps 2–3 to get the  $Fq(s)$ . Thus, by implementing the fitting polynomial of each segment, the trend of that segment will be eliminated efficiently.

4. Analyze log–log plots  $Fq(s)$  versus  $s$  for each value of  $q$  so as to determine the scaling behavior of the fluctuation functions  $Fq(s)$ . If the series  $x_i$  is long-range power-law correlated,  $Fq(s)$  and  $s$  are related as follows

$$Fq(s) \sim s^{H(q)} \quad (11)$$

$H(q)$  is defined as the generalized Hurst exponent and may be related to  $q$ . Normally,  $H(q)$  is used to describe the impact of past time series on current and future ones because of its ability of long-range correlation. For rolling bearings, owing to external factors, the status may change from normal to fault status. However, the fault status will persist all the time for its self-irreparability. Therefore, the variable corresponding to Hurst exponent

under different statuses can distinguish between normal status and fault status with high performance for vibration fault diagnosis.

In nature, multi-fractals exist in common circumstances. Oppositely, mono-fractals are rare. Therefore, the multi-fractal analysis is generally applicable to reveal intrinsic fractal characteristics. However, noise interference problems will influence the MF-DFA method. The proposed method shown herein effectively overcomes this problem.

## SITD method

With the purpose of effectively using the sensitive features contained in feature set for fault diagnosis and overcoming the noise interference problem, a new SITD decomposition method based on ITD was proposed.

Although the ITD approach has many advantages, we cannot ignore its problems, such as signal distortion, end effects, and redundant PRCs. These problems will influence feature extraction. One of the causes of these problems is the algorithm itself, while another cause is noise interference. As we all know, the fault vibration signal includes the resonant component, the impulse signal produced by the faulty bearing and noise. Impulse signals and noise are high-frequency signals, and their noise affects the extraction of fault signals. Therefore, noise elimination is the key step in feature extraction. For the purpose of eliminating the noise signal, the SITD method was proposed, in which wavelet analysis was embedded in the iteration procedures of ITD, which could preserve the effective signal and improve the SNR. To get better results, an adaptive threshold function was analyzed in the wavelet domain.

### An adaptive threshold function

To maintain the effective signal details as much as possible in the denoising process, a threshold function that is superior to both hard and soft threshold functions was designed. The proposed threshold function is presented as follows

$$\eta(x, th, m, n) = \begin{cases} x - 0.5 \operatorname{sign}(x) \frac{th^m}{|x|^{m-1}}, & |x| > th \\ 0.5 \operatorname{sign}(x) \frac{x^n}{th^{n-1}}, & |x| \leq th \end{cases} \quad (12)$$

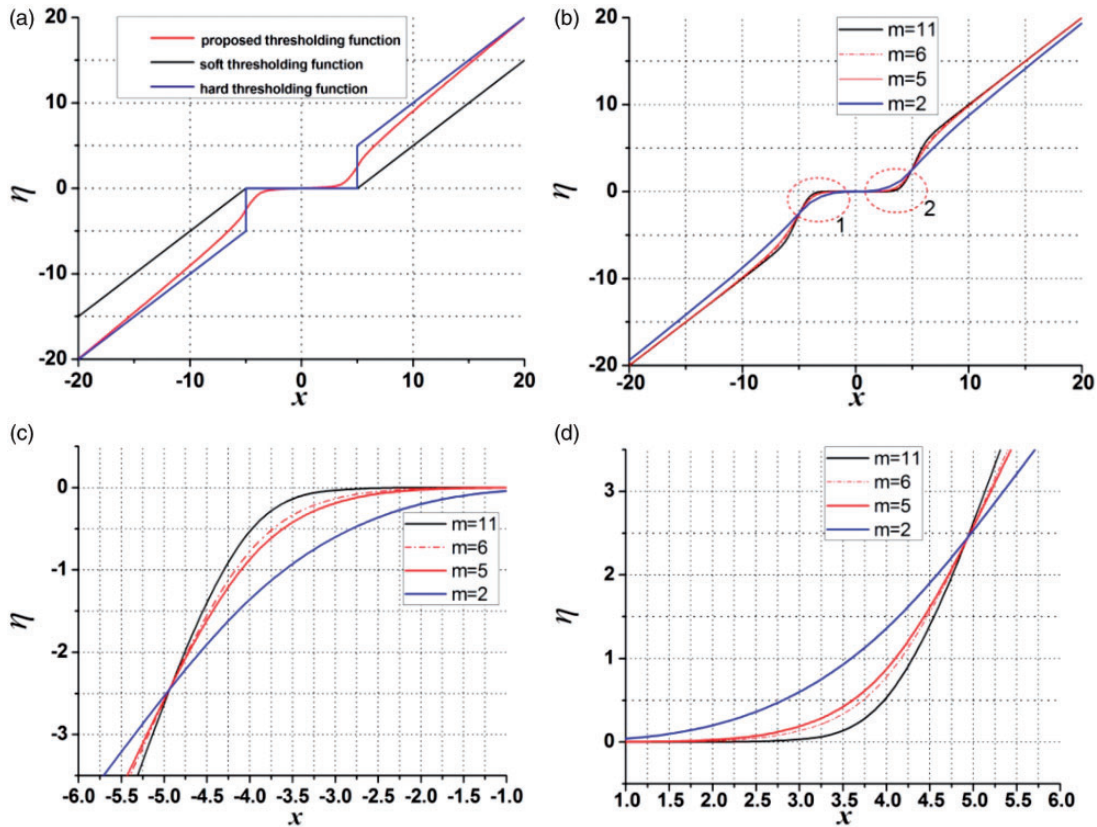
where  $th$  is the threshold, while parameters  $n$  and  $m$  can be adjusted and used to determine the shape of the function for coefficients that are smaller and larger than the absolute threshold value, respectively. To realize the first-order derivability of the threshold function, equation (12) needs to satisfy the following condition

$$\left. \frac{\partial \eta(x, th, m, n)}{\partial x} \right|_{x=th^-} = \left. \frac{\partial \eta(x, th, m, n)}{\partial x} \right|_{x=th^+} \quad (13)$$

Meanwhile,  $n = m + 1$ ; therefore, the expression of the threshold function (12) after adding the first-order derivable property is

$$\eta(x, th, m) = \begin{cases} x - 0.5 \operatorname{sign}(x) \frac{th^m}{|x|^{m-1}}, & |x| > th \\ 0.5 \operatorname{sign}(x) \frac{x^{m+1}}{th^m}, & |x| \leq th \end{cases} \quad (14)$$

When  $m$  is transformed, zone 1 and zone 2 from Figure 1(b) can be represented as shown enlarged in Figure 1 (c) and (d). When  $m = 1$ , the curve of the function and the soft threshold function in Figure 1(a) almost coincide. When  $m > 10$ , the curve is very close to the hard threshold function. The threshold function could nullify, shrink, or maintain the wavelet coefficients in different regions, unlike the soft and hard threshold functions. The region that could shrink the wavelet coefficient is the critical region in which the WT coefficient is composed of signal and noise. To achieve the purpose of maximally retaining the details of the signal while removing noise, a more



**Figure 1.** Proposed threshold function with different  $m$  values. (a) Proposed threshold function. (b) Proposed threshold function with hard and soft threshold function different  $m$  value. (c) Region one. (d) Region two.

realistic initial signal for target recognition is needed. The size of the contraction region could be adjusted, and the removal ratio of the noise signal could be controlled, resulting in the retention ratio of the signal details.

In conclusion, the threshold function could achieve a smooth transition of wavelet coefficient attenuation in the critical region. The size of the critical region was determined by adjusting the parameter  $m$ . The higher  $m$  is, the more similar the process is to the process of the hard threshold function. Simultaneously, the wavelet coefficient in the critical area could be contracted on a large scale. Therefore, when the value of  $m$  is relatively large, it is suitable to process a noisy signal with a low SNR. Conversely, the smaller the value of  $m$  is, the more wavelet coefficients are in the critical region, in which the wavelet coefficients with signal detail could be better preserved in the case of shrinking noise factor, thereby maintaining the original local singularity of the signal.

To more effectively preserve the signal singularity in the signal denoising process, the threshold function  $\eta(x, th, m)$  applicable to wavelet coefficients at different scales could be selected by adjusting the parameter  $m$  in equation (14). The corresponding denoising strategy is as follows: use the partial hard threshold function to denoise for the coefficient on a small scale, eliminating most of the noise; use the partial soft threshold function to denoise for the coefficient on a large scale and realize coefficient contraction in the critical area. For adaptive selection of the threshold function for specific signals, the mathematical model is established as follows

$$m_j = 1 + 10 \frac{E_{n_j}}{E_{d_j}} \quad (15)$$

where  $E_{d_j} = \sum_{k=0}^{N-1} d_{j,k}^2$ ,  $m_j \in (1, 11]$  when  $j = 1$ ,  $m_j$  reaches a maximum of 11; when  $j$  increases,  $m_j$  decreases, realizing the change in the threshold function from hard to soft. By substituting equation (15) into equation (14), the corresponding threshold function on each scale could be obtained.

The selection of suitable wavelet parameters is indispensable for signal analysis. The adaptive threshold function is used to improve the denoising efficiency and provide a more real input signal for the subsequent target

recognition process. Considering the applicability of the function, a comprehensive threshold function is generated by adding shape tuning factor  $m$ .

### Smooth intrinsic time-scale decomposition

The flow chart of the proposed wavelet-analysis-embedded ITD method is shown in Figure 2, while the transformation of PRCs into wavelet domain is performed through the equation as follows

$$W_{j,k} = \int_{-\infty}^{+\infty} PRC_s(t) \psi_{j,k}(t) dt \quad (16)$$

Here,  $\psi_{j,k}(t)$  is the mother wavelet,  $W_{j,k}$  represents the  $k$ -th wavelet coefficient at the  $j$ -th level,  $j, k \in \mathbb{Z}$ . The threshold denoising is used to retain the effective signal components in each PRC. New wavelet coefficients are obtained through denoise processing, and subsequently, a new PRC is reconstructed. The correlation coefficient ( $c$ ) is used as a stopping criterion to remove undesirable PRCs. The fact that the correlation coefficient value is smaller than the predefined threshold should be considered in the PRC reconstruction.

The evaluation function is described as follows

$$c = \frac{\sum (c_k(t) - \overline{c_k(t)}) (x(t) - \overline{x(t)})}{\sqrt{\sum (c_k(t) - \overline{c_k(t)})^2} \sqrt{\sum (x(t) - \overline{x(t)})^2}} \quad (17)$$

where  $c_k(t)$  is the  $k$ -th PRC, and  $\overline{c_k(t)}$  is the mean value.

Unlike previously, the new purified  $PRC_0$  is subtracted from the original signal to obtain a new baseline signal, which is set as the input signal.

$$L_t' = X_t - PRC_s(t) \quad (18)$$

Repeating the above-mentioned process is necessary in every sifting stage until a final monotonic trend is achieved.

The operation procedures of SITD are as follows:

1. Compute the  $PRC_s$  based on the procedure discussed previously.
2. Calculate the wavelet coefficients by the following equation

$$W_{j,k} = \int_{-\infty}^{+\infty} PRC_s(t) \psi_{j,k}(t) dt \quad (19)$$

3. Estimate the new denoised wavelet coefficients by presetting a positive value for threshold  $th$

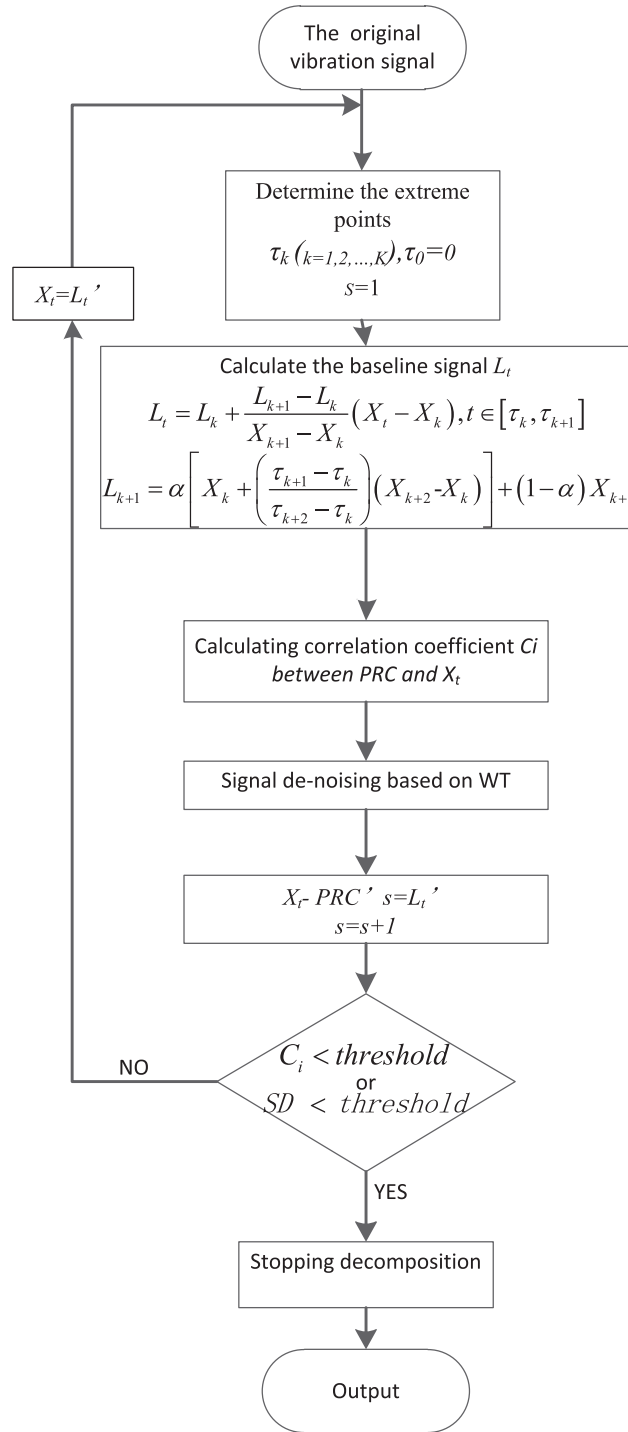
$$d_t(W_{j,k}) = \begin{cases} W_{j,k} - 0.5 \text{sign}(W_{j,k}) \frac{th^m}{|x|^{m-1}}, & |W_{j,k}| > th \\ 0.5 \text{sign}(W_{j,k}) \frac{W_{j,k}^{m+1}}{th^m}, & |W_{j,k}| \leq th \end{cases} \quad (20)$$

4. Reconstruct  $PRC_s$  from the denoised coefficients and approximate the coefficients by using the inverse WT.
5. Subtract the new  $PRC_s$  from  $X_t$  and obtain the residual as the input signal.

$X_t$  will be replaced with the residual signal, then return to step 1, the iteration process should be repeated and continued until residual  $L_t'$  becomes too small or a monotonic function from which no more  $PRC_s$  can be extracted.

### Comparison between the efficiency of ITD and SITD

To verify the effectiveness of this method in eliminating end effects and noise interference based on the analysis of widely used signals, comparative studies with traditional ITD methods are performed. The simulated signal

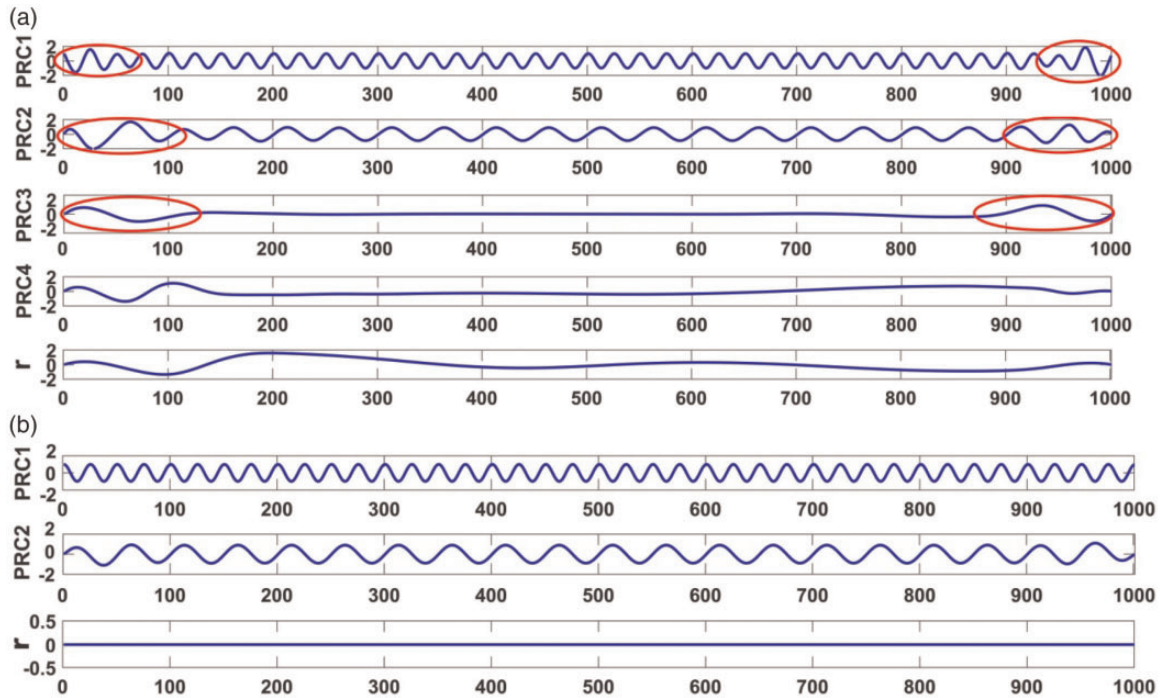


**Figure 2.** Flow chart of SITD. WT: wavelet transform; SD: standard deviation; PRC: proper rotation component.

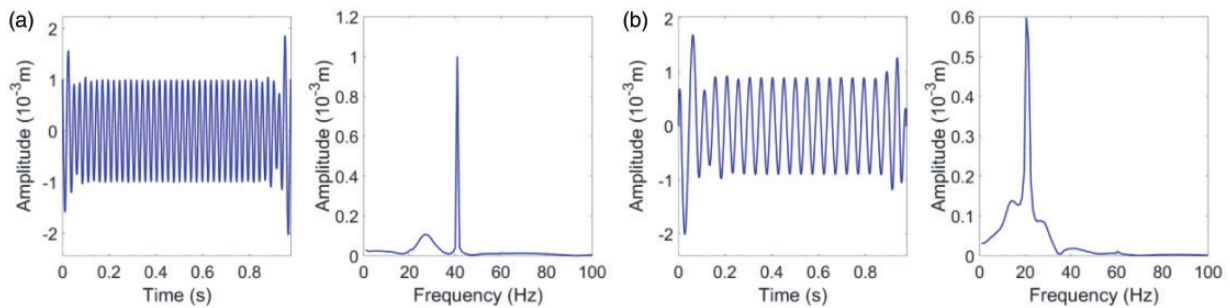
includes two components and Gaussian noise

$$y(t) = \sin(40\pi t) + \cos(80\pi t) \tag{21}$$

Gaussian noise information is added by simulation software. The two components with frequencies of 40 Hz and 20 Hz are defined as the first-order PRC and second-order PRC, respectively. When there is no noise,



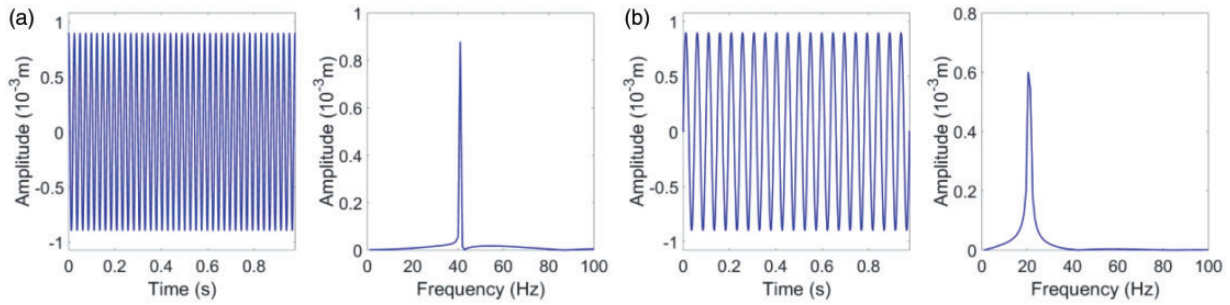
**Figure 3.** Signal decomposition. (a) Traditional ITD method. (b) WT combined with ITD method. PRC: proper rotation component.



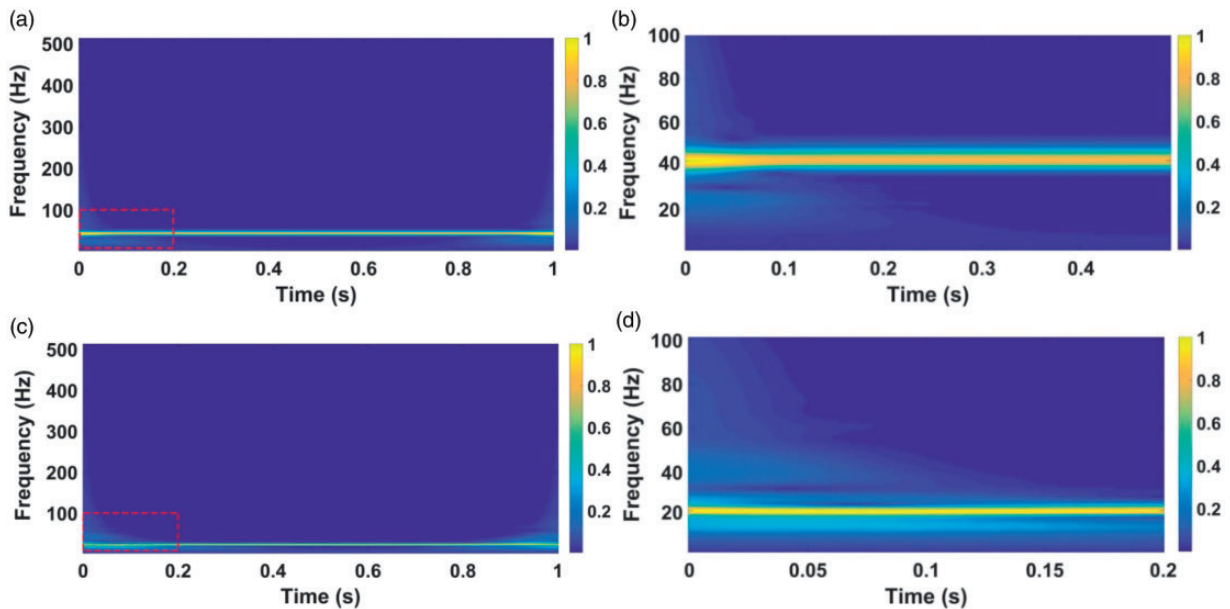
**Figure 4.** Time-domain and frequency-domain representation of PRC1 and PRC2, decomposed by the traditional ITD method. (a) PRC1. (b) PRC2. PRC: proper rotation component.

as shown in Figure 3, end effects occurred in Figure 3(a) by using the traditional ITD method. On the contrary, the wavelet-analysis-embedded ITD method could not only eliminate the end effects but also effectively reduce redundant PRCs (Figure 3(b)). As shown in Figures 4 and 5, a pure signal was obtained with the proposed method, and the proposed method has more advantages than the traditional ITD method in addressing the end effect problem. As illustrated in Figures 6 and 7, the enlargement of the (a) and (c) regions was shown in (b) and (d). Frequency–time representation is used to analyze the differences of PRC under the two methods. Figure 6 shows PRC1 and PRC2 with the corresponding error frequencies due to the end effect. The reason for this phenomenon is that the ITD method extracts the baseline signal based on the linear transformation, which may result in signal distortion, burr, and end effects. The comparative results in Figure 7 show that the wavelet-analysis-embedded ITD method was superior to the traditional ITD method in processing the distortion of signal and end effects effectively.

After adding the noise, the traditional ITD method was applied to perform a comparison analysis with the SITD approach in terms of denoising and feature recognition, resulting in frequency aliasing in every PRC, as shown in Figure 8(a). Figure 8(b) is the representation after denoising with a hard threshold function, and the noise is basically filtered out, but over-smoothing leads to distortion in the signal details. Figure 8(c) shows the



**Figure 5.** Time-domain and frequency-domain representation of PRC1 and PRC2, decomposed by proposed WT combined with the ITD method. (a) PRC1. (b) PRC2. PRC: proper rotation component.

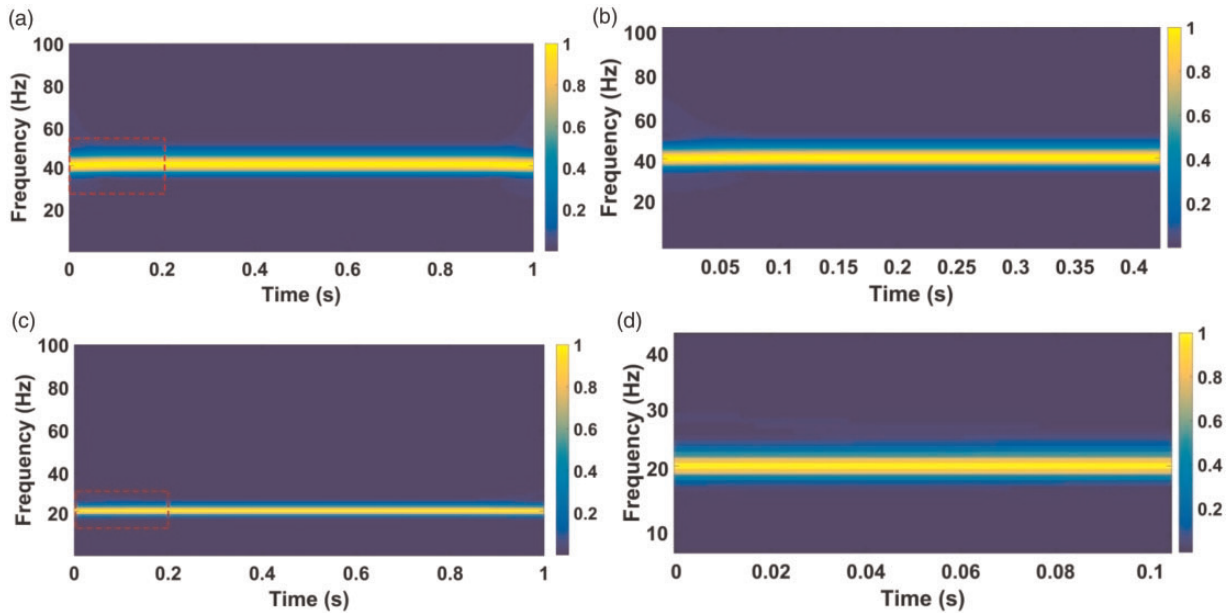


**Figure 6.** Frequency–time representation of PRC1 and PRC2 decomposed by the ITD method. (a) PRC1. (b) Red region of PRC1. (c) PRC2. (d) Red region of PRC2. PRC: proper rotation component.

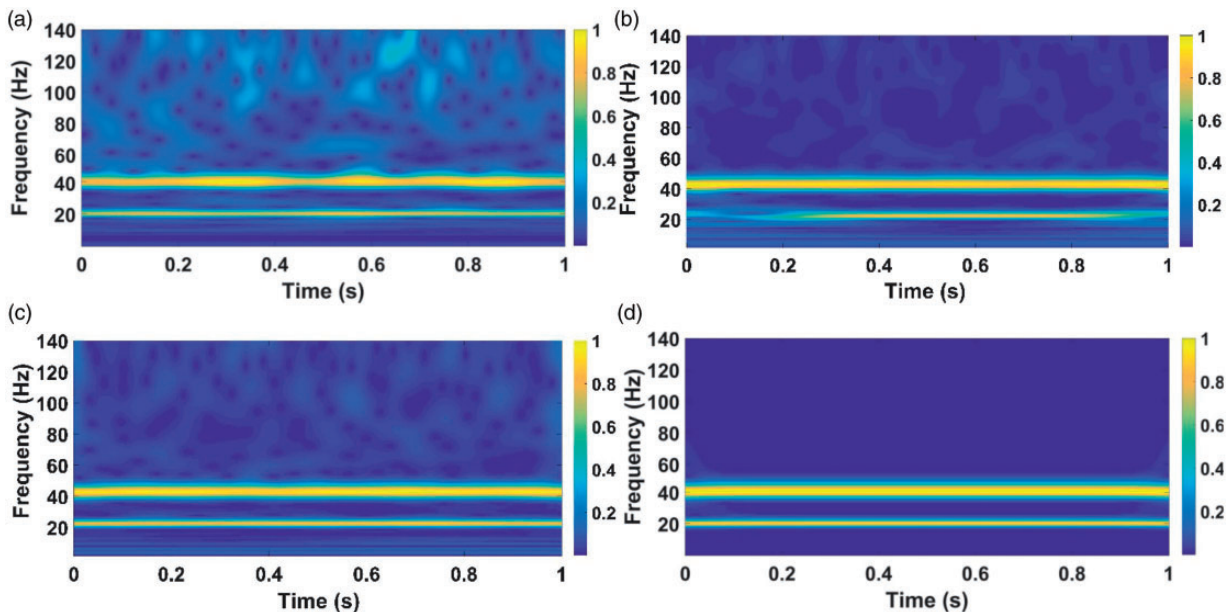
effect after denoising with a soft threshold function; the selected threshold value is half the threshold value used in Figure 8(b). The details of the signal are more evident than the result when using a hard threshold function for denoising; however, more noise is also retained. To preserve the signal details in the denoising procedure, it is necessary to design a new threshold function between the hard threshold function and the soft threshold function. The SITD algorithm filtered the background noise and extracted two weak PRCs from the investigated signal effectively (Figure 8(d)). The cross-correlation coefficients of the extracted PRCs and the real modes obtained by using the proposed method with different threshold function methods—soft threshold,<sup>24</sup> hard threshold,<sup>25</sup> and proposed adaptive threshold—are 0.7064, 0.7506, and 0.8654, respectively. The larger the coefficient value is, the better the extracted PRCs match the real signal mode.

## Experimental validation

The fault diagnosis approach for rolling bearings based on SITD and MF-DFA is presented in this part. The frame for fault identification is shown in Figure 9. Firstly, the original signal was decomposed into several PRCs by applying SITD method. Secondly, the instantaneous amplitude in each PRC was analyzed by MF-DFA, and the multi-fractal features were represented by the generalized Hurst exponents. The generalized Hurst index of different  $q$  was obtained. Thirdly, LLTSA was used to decrease the value of dimension. The feature vector was formed by applying the major influencing components got from LLTSA. Fault diagnosis of rolling bearing was realized by verifying the extracted fault feature vectors.



**Figure 7.** Frequency–time representation of PRC1 and PRC2, decomposed by the proposed WT combined with ITD method. (a) PRC1. (b) Red region of PRC1. (c) PRC2. (d) Red region of PRC2. PRC: proper rotation component.



**Figure 8.** Frequency–time representation. (a) Decomposed by the traditional ITD method. (b) Decomposed by WT (hard threshold function) combined with the ITD method. (c) Decomposed by WT (soft threshold function) combined with the ITD method. (d) Decomposed by the proposed WT combined with the ITD method.

### Feature extraction by applying SITD and MF-DFA

The proposed SITD method was used to extract the fault characteristics of the vibration signal from Case Western Reserve University. The experiment platform is shown in Figure 10(a) 2 hp Reliance Electric motor, torque encoder, and dynamometer), and the bearing information is listed in Table 1. The motor drives input and output shaft drive loads of 1, 2, 3, and 4 hp. EDM (electrical discharge machining) technology is used to produce single-point faults on the test bearings with different fault diameters (0.1778, 0.3556, and 0.5334 mm). The sampling frequency is 12 kHz.

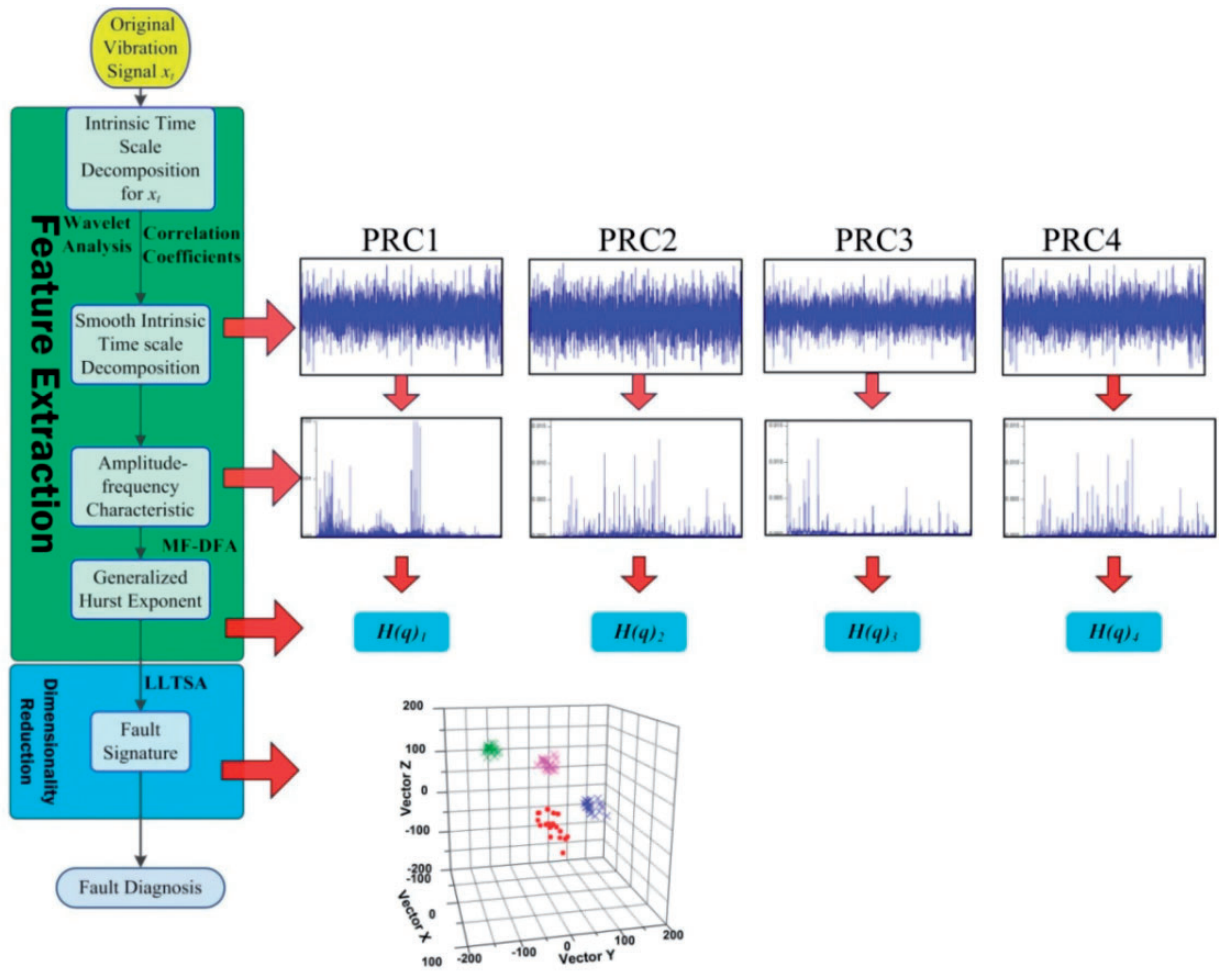


Figure 9. The scheme for fault identification. PRC: proper rotation component; LLTSA: linear local tangent space alignment; MF-DFA: multi-fractal detrended fluctuation analysis.

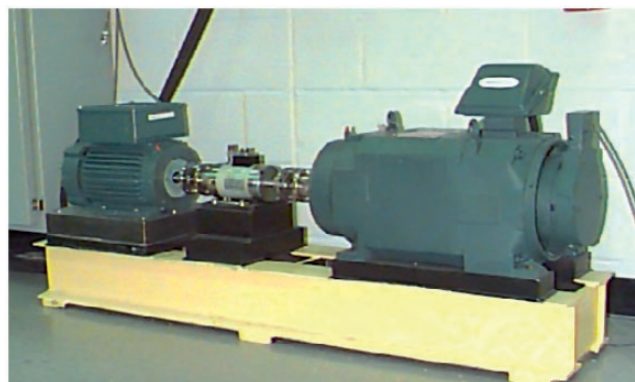


Figure 10. Test platform of rolling bearings.

Table 1. Test rolling information.

Ball number	Ball diameter (mm)	Pitch diameter (mm)	Inside diameter (mm)	Outside diameter (mm)	Thickness (mm)
9	7.95	39	25	52	15

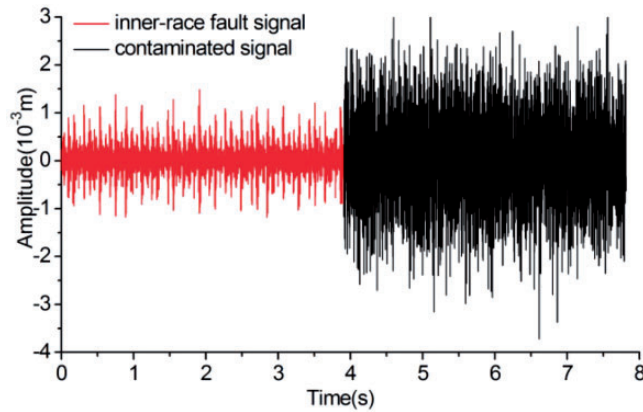


Figure 11. Time-domain diagram.

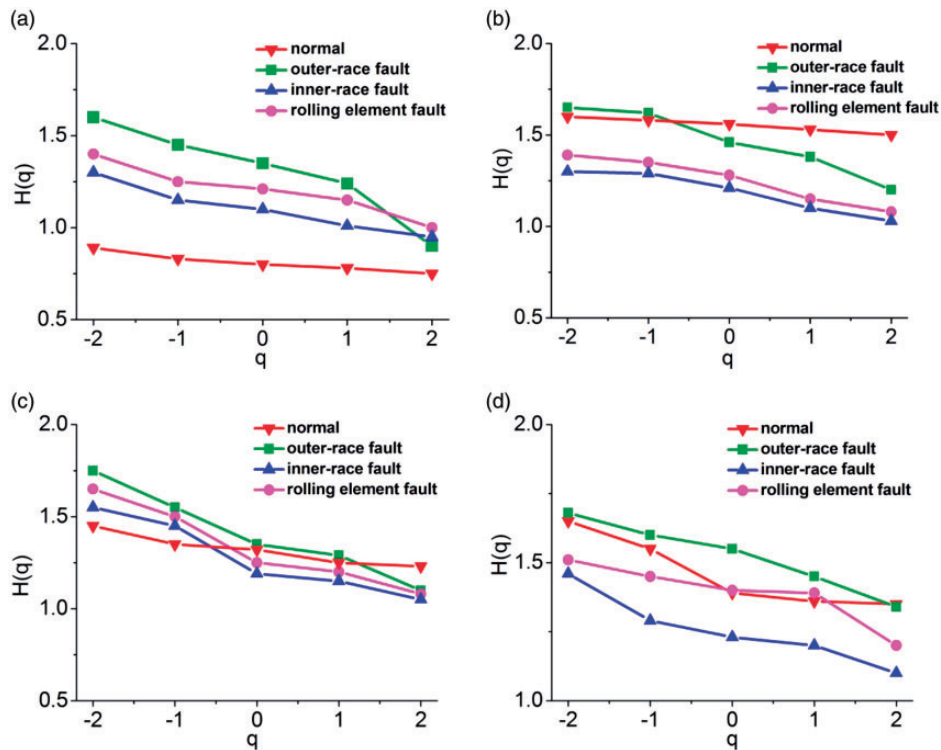
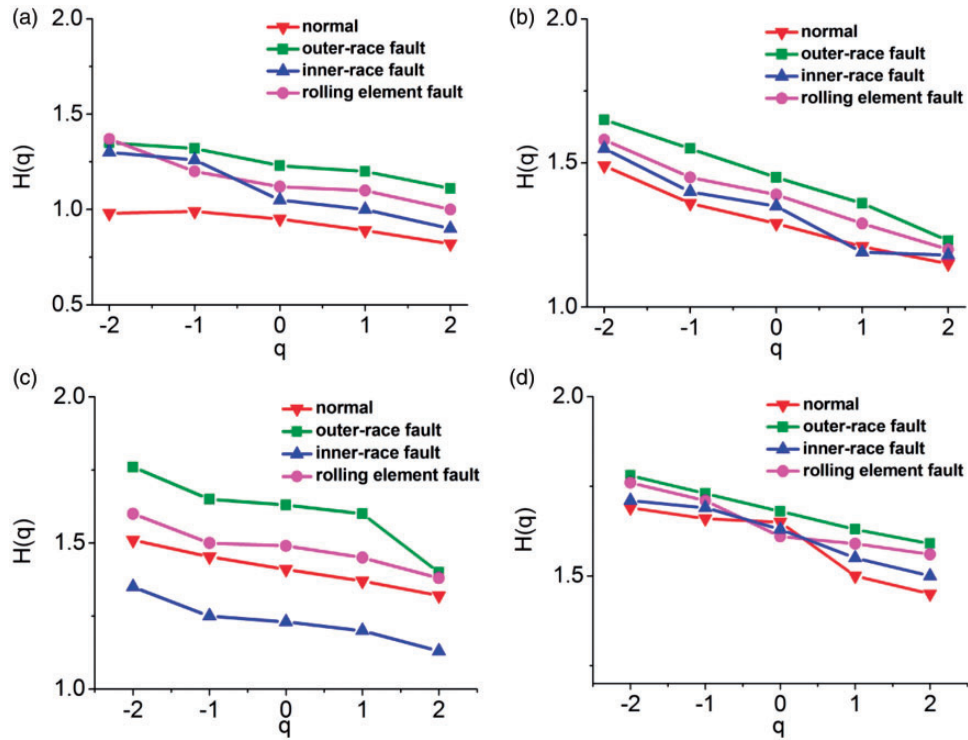


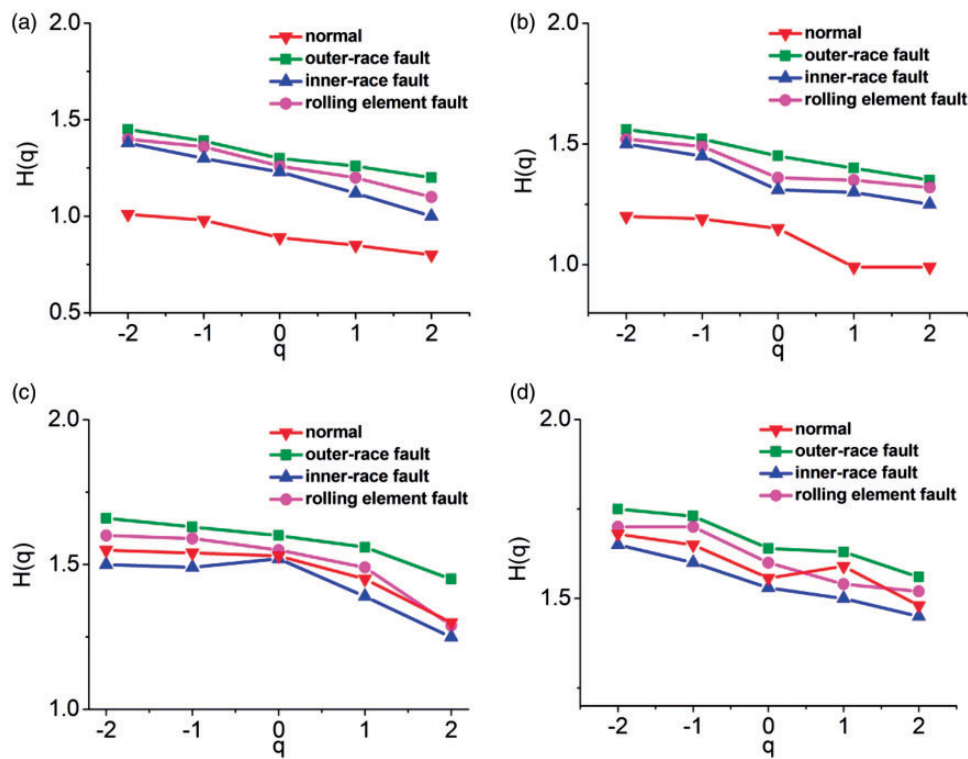
Figure 12. Generalized Hurst exponent of first four PRCs with a fault diameter of 0.1778 mm. (a) PRC1. (b) PRC2. (c) PRC3. (d) PRC4. PRC: proper rotation component.

The original signal was decomposed into multiple PRC signals by applying SITD. According to SITD decomposition, the first few PRCs exhibit the largest energy and frequency. In other words, the first few PRCs contain most of the effective signal information. With regard to this point, only the first four PRCs were preserved in this study.

To verify the applicability of the proposed method of the actual working situation, Gaussian white noise was added. Figure 11 depicts the time-domain representation of the inner-race fault signal and contaminated signal, which is composed of an inner-race fault signal and Gaussian white noise. The multi-fractal characteristics of each instantaneous amplitude matrix were analyzed by using MF-DFA. The scale index  $q$  of MF-DFA used in this study is  $-2, -1, 0, 1,$  and  $2$ .



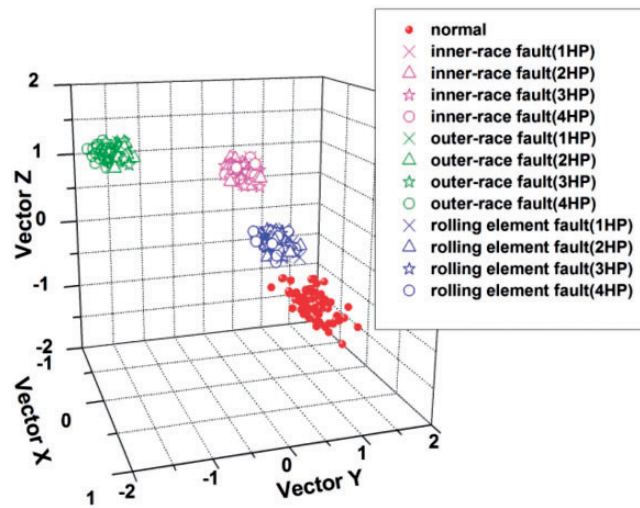
**Figure 13.** Generalized Hurst exponent of first four PRCs with a fault diameter of 0.3556 mm. (a) PRC1. (b) PRC2. (c) PRC3. (d) PRC4. PRC: proper rotation component.



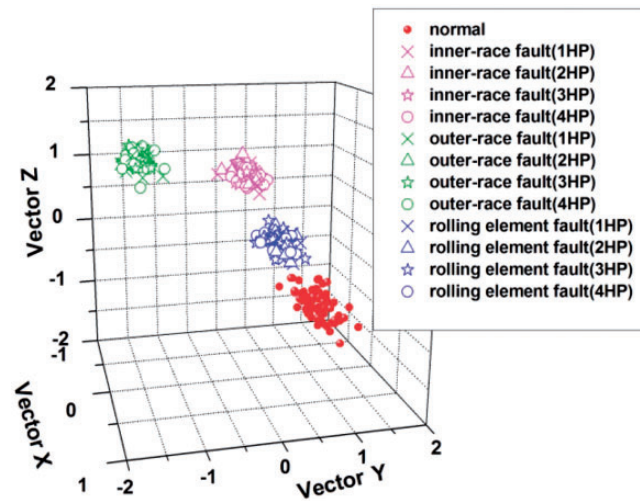
**Figure 14.** Generalized Hurst exponent of first four PRCs with a fault diameter of 0.5334 mm. (a) PRC1. (b) PRC2. (c) PRC3. (d) PRC4. PRC: proper rotation component.

**Table 2.** Information of the data set.

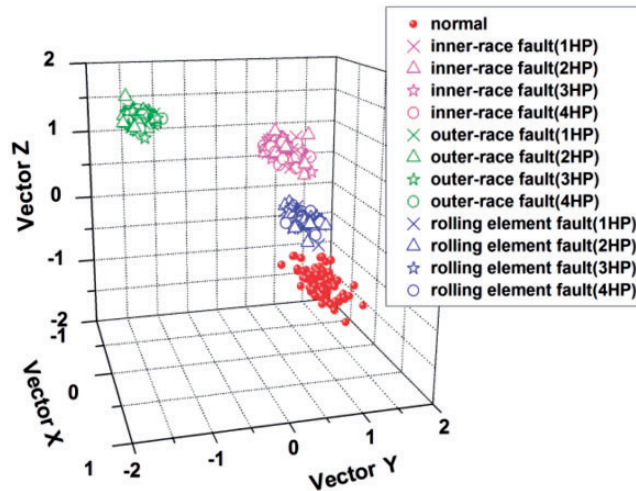
Label	Status of the rolling bearing	Fault diameter (mm)	Working condition and quantity (group)			
			1HP	2HP	3HP	4HP
1	Normal		80			
	Inner-race	0.1778	20	20	20	20
	Outer-race	0.1778	20	20	20	20
	Rolling element	0.1778	20	20	20	20
2	Normal		80			
	Inner-race	0.3556	20	20	20	20
	Outer-race	0.3556	20	20	20	20
	Rolling element	0.3556	20	20	20	20
3	Normal		80			
	Inner-race	0.5334	20	20	20	20
	Outer-race	0.5334	20	20	20	20
	Rolling element	0.5334	20	20	20	20



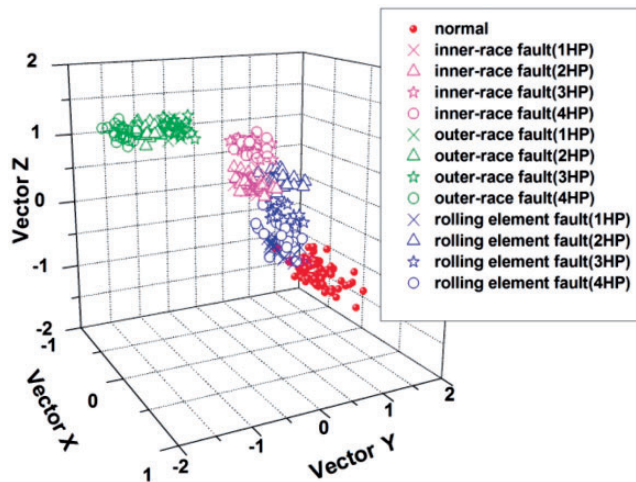
**Figure 15.** Diagram of the clustering result of Data set 1 (SITD + MF-DFA).



**Figure 16.** Diagram of the clustering result of Data set 2 (SITD + MF-DFA).



**Figure 17.** Diagram of the clustering result of Data set 3 (SITD + MF-DFA).



**Figure 18.** Diagram of the clustering result of Data set 1 (ITD + MF-DFA).

Figures 12 to 14 describe the generalized Hurst exponents for PRC1, PRC2, PRC3, and PRC4 in four different states (normal, outer-race fault, inner-race fault, rolling-element fault) with three different failure diameters (0.1778, 0.3556, and 0.5334 mm, respectively).

As can be seen from the figures, the generalized Hurst components of various states were distinguished. In summary, the various rolling bearing states have different multi-fractal features.

### *Application of LLTSA to rolling bearing dimension reduction*

In this study, the first four PRCs presented in the previous section were chosen and analyzed by the MF-DFA method. The number of multi-fractal features  $H(q)$  is 5 for the instantaneous amplitude corresponding to each PRC ( $q = [-2, -1, 0, 1, 2]$ ). Its total number is 20 for each set of data.

For rolling bearing diagnostic systems, the large number of features may decrease their robustness and may even result in a decrease in diagnostic accuracy. LLTSA is a type of algorithm used for dimensionality reduction.<sup>26–29</sup> Its purpose is to extract the major influencing factors. In this study, LLTSA projected the multi-fractal features into a three-dimensional space.

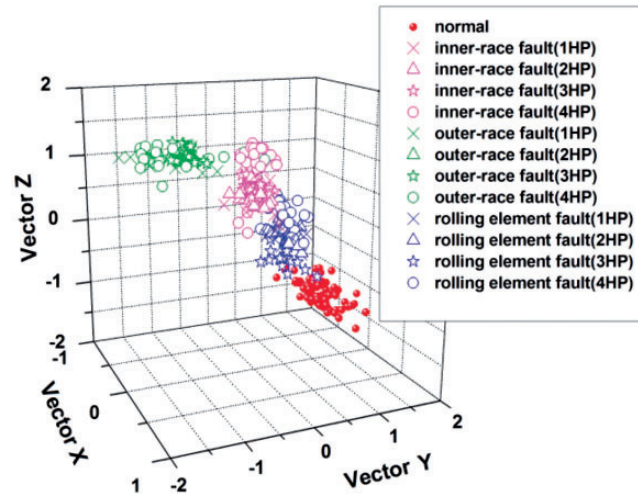


Figure 19. Diagram of the clustering result of Data set 2 (ITD + MF-DFA).

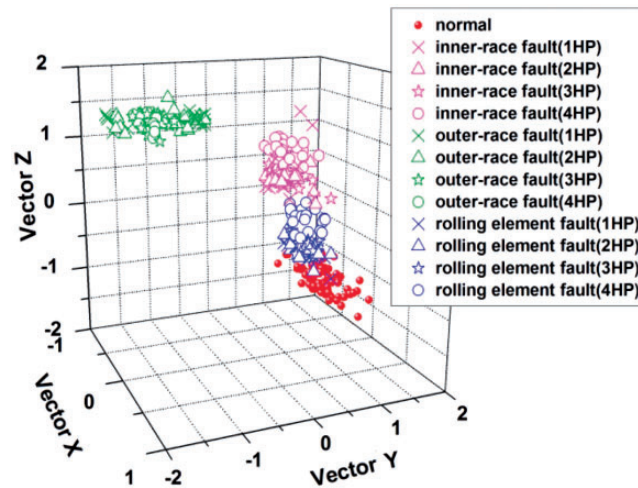


Figure 20. Diagram of the clustering result of Data set 3 (ITD + MF-DFA).

### Diagnostic performance under different operating conditions

A change in velocity will affect the fault characteristic frequency. The characteristics extracted by the traditional method fluctuate greatly under variable working conditions, which may lead to the mixing of feature vectors of different fault states and reduce the diagnostic accuracy. Considering these statements, it is indispensable to study the diagnostic characterization of the method under different circumstances (motor speeds of 1797, 1772, 1750, and 1720 r/min, respectively).

The normal state data set contains 80 groups. Every single fault data set is separated into 12 subsets under four different working conditions (load of output shaft drive is 1, 2, 3, and 4 horsepower) and three different faults diameters (0.1778, 0.3556, and 0.5334 mm); each subset contains 20 groups, and each group contains 4096 points. The data set information is shown in Table 2. Figures 15 to 17 exhibit the scatter plots of the three vectors.

As can be seen from the figures, there is no mixing of feature vectors of different fault states (normal, inner-race fault, outer-race fault, rolling element fault). In other words, a fluctuation in working conditions has no effect on the extracted feature vectors that were verified in the situations of different fault diameters. The signals processed by the ITD-MF-DFA method are shown in Figures 18 to 20, in which the extracted feature vectors in different fault states are combined with each other.

## Conclusions

To improve the accuracy of feature extraction in fault diagnosis, a new fault diagnosis approach involving rolling bearing decomposition based on SITD and MF-DFA was proposed. Comparative studies indicate that the SITD approach is superior to traditional ITD methods in treating the noise interference problem. After signal processing by SITD, the first four PRCs were selected to implement MF-DFA to exhibit the intrinsic multi-fractality in the instantaneous amplitude of the extracted PRCs. The multi-fractal features were represented by the generalized Hurst exponents. Finally, LLTSA was used to decrease the value of the dimensions, and the main multi-components are screened to construct feature vectors. The diagnostic performance under different operating conditions is studied. Experimental results show that the proposed approach is feasible for recognizing the rolling bearing fault types.

This method has the following characteristics: (1) the proposed method effectively eliminates noise interference and inhibits end effects; (2) the algorithm increases the SNR, improving the accuracy of feature extraction when the MF-DFA method is used in fault diagnosis; and (3) the method can be applied to the fault diagnosis of rolling bearings and the accurate identification of the fault types effectively.

## Acknowledgements

The authors greatly appreciate the constructive comments provided by the anonymous reviewers and the editors.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is sponsored by Natural Science Foundation of Liaoning Province, China (Grant Nos. 20180550927 and 20180550002); National Natural Science Foundation of China (Grant No. 51705342); and National Key Research and Development Program of China (2017YFC0703903).

## ORCID iD

Dong An  <https://orcid.org/0000-0002-7329-8704>

## References

1. Cui LL, Jin Z, Huang JF, et al. Fault severity classification and size estimation for ball bearing based on vibration mechanism. *IEEE Access* 2019; 7: 56107–56116.
2. Zhao LY, Yu W, Yan RQ. Rolling bearing fault diagnosis based on CEEMD and time series modeling. *Math Prob Eng* 2014; 2014: 1–13.
3. Abd-el-Malek MB and Hanna SS. Using filter bank property to simplify the calculations of empirical mode decomposition. *Commun Nonlinear Sci Numer Simul* 2018; 62: 429–444.
4. Moore KJ, Kurt M, Eriten M, et al. Wavelet-bounded empirical mode decomposition for measured time series analysis. *Mech Syst Signal Process* 2018; 99: 14–29.
5. Cui, LL, Wang J and Lee S. Matching Pursuit of an Adaptive Impulse Dictionary for Bearing Fault Diagnosis. *J Sound Vib.* 2014; 333: 2840–2862.
6. Lou X and Loparo KA. Bearing fault diagnosis based on wavelet transform and fuzzy inference. *Mech Syst Signal Process* 2004; 18: 1077–1095.
7. Liu B, Ling SF and Meng Q. Machinery diagnosis based on wavelet packets. *J Vib Control* 1997; 3: 5–17.
8. Peng ZK, Peter WT and Chu FL. A comparison study of improved Hilbert–Huang transform and wavelet transform: application to fault diagnosis for rolling bearing. *Mech Syst Signal Process* 2005; 19: 974–988.
9. Yuan HD, Chen J and Dong GM. Machinery fault diagnosis based on time–frequency images and label consistent K-SVD. *Proc Inst Mech Eng Part C: J Mech Eng Sci* 2018; 232: 1317–1330.
10. Yan R, Gao RX and Chen X. Wavelets for fault diagnosis of rotary machines: a review with applications. *Signal Process* 2014; 96: 1–15.
11. Van M and Kang HJ. Two-stage feature selection for bearing fault diagnosis based on dual-tree complex wavelet transform and empirical mode decomposition. *Proc Inst Mech Eng Part C: J Mech Eng Sci* 2016; 230: 291–302.

12. Chen J, Li Z, Pan J, et al. Wavelet transform based on inner product in fault diagnosis of rotating machinery: a review. *Mech Syst Signal Process* 2016; 70: 1–35.
13. Cui LL, Wang X, Wang HQ, et al. Improved fault size estimation method for rolling element bearings based on concatenation dictionary. *IEEE Access* 2019; 7: 22710–22718.
14. Frei MG and Osorio I. Intrinsic time-scale decomposition: time-frequency-energy analysis and real-time filtering of non-stationary signals. *Proc Math Phys Eng Sci* 2007; 463: 321–342.
15. An X, Jiang D, Chen J, et al. Application of the intrinsic time-scale decomposition method to fault diagnosis of wind turbine bearing. *J Vib Control* 2012; 18: 240–245.
16. Yu J and Liu H. Sparse coding shrinkage in intrinsic time-scale decomposition for weak fault feature extraction of bearings. *IEEE Trans Instrum Meas* 2018; 67: 1579–1592.
17. Lin JS. Improved intrinsic time-scale decomposition method and its simulation. *Appl Mech Mater* 2011; 121: 2045–2048.
18. Zeng JX, Wang GF, Zhang FQ, et al. The de-noising algorithm based on intrinsic time-scale decomposition. *Adv Mater Res* 2011; 422: 347–352.
19. Hu A, Yan X and Xiang L. A new wind turbine fault diagnosis method based on ensemble intrinsic time-scale decomposition and WPT-fractal dimension. *Renew Energy* 2015; 83: 767–778.
20. Lin J and Chen Q. Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and Mahalanobis distance criterion. *Mech Syst Signal Process* 2013; 38: 515–533.
21. Liu H, Wang X and Lu C. Rolling bearing fault diagnosis based on LCD-TEO and multifractal detrended fluctuation analysis. *Mech Syst Signal Process* 2015; 60: 273–288.
22. Telesca L and Lovallo M. Analysis of the time dynamics in wind records by means of multifractal detrended fluctuation analysis and the Fisher–Shannon information plane. *J Stat Mech* 2011; 2011: p07001.
23. Li F, Tang B and Yang R. Rotating machine fault diagnosis using dimension reduction with linear local tangent space alignment. *Measurement* 2013; 46: 2525–2539.
24. Donoho DL and Johnstone JM. Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 1994; 81: 425–455.
25. Gao HY. Wavelet shrinkage denoising using the non-negative garrote. *J Comput Graph Stat* 1998; 7: 469–488.
26. Wang J, Sun X and Du J. Local tangent space alignment via nuclear norm regularization for incomplete data. *Neurocomputing* 2018; 273: 141–151.
27. Li Y, Luo D and Liu S. Orthogonal discriminant linear local tangent space alignment for face recognition. *Neurocomputing* 2009; 72: 1319–1323.
28. Li F, Chu MK, Wang JX, et al. Life grade recognition of rotating machinery based on supervised orthogonal linear local tangent space alignment and optimal supervised Fuzzy C-Means clustering. *Measurement* 2015; 73: 384–400.
29. Su Z, Tang B, Liu Z, et al. Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine. *Neurocomputing* 2015; 157: 208–222. <