

Article

Tensor Singular Spectrum Decomposition Algorithm Based on Permutation Entropy for Rolling Bearing Fault Diagnosis

Cancan Yi ^{1,2}, Yong Lv ^{1,2,*}, Mao Ge ^{1,2}, Han Xiao ^{1,2} and Xun Yu ^{1,2,3}

¹ Key Laboratory of Metallurgical Equipment and Control Technology (Wuhan University of Science and Technology), Ministry of Education, Wuhan 430081, China; meycancan@wust.edu.cn (C.Y.); ge1656372625@gmail.com (M.G.); xiaohan@wust.edu.cn (H.X.); xyu13@nyit.edu (X.Y.)

² Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering (Wuhan University of Science and Technology), Wuhan 430081, China

³ Department of Mechanical Engineering, New York Institute of Technology, Old Westbury, NY 11568, USA

* Correspondence: lvyong@wust.edu.cn; Tel.: +86-27-6886-2857; Fax: +86-27-6886-2212

Academic Editors: Jose C. Principe and Badong Chen

Received: 13 March 2017; Accepted: 21 March 2017; Published: 23 March 2017

Abstract: Mechanical vibration signal mapped into a high-dimensional space tends to exhibit a special distribution and movement characteristics, which can further reveal the dynamic behavior of the original time series. As the most natural representation of high-dimensional data, tensor can preserve the intrinsic structure of the data to the maximum extent. Thus, the tensor decomposition algorithm has broad application prospects in signal processing. High-dimensional tensor can be obtained from a one-dimensional vibration signal by using phase space reconstruction, which is called the tensorization of data. As a new signal decomposition method, tensor-based singular spectrum algorithm (TSSA) fully combines the advantages of phase space reconstruction and tensor decomposition. However, TSSA has some problems, mainly in estimating the rank of tensor and selecting the optimal reconstruction tensor. In this paper, the improved TSSA algorithm based on convex-optimization and permutation entropy (PE) is proposed. Firstly, aiming to accurately estimate the rank of tensor decomposition, this paper presents a convex optimization algorithm using non-convex penalty functions based on singular value decomposition (SVD). Then, PE is employed to evaluate the desired tensor and improve the denoising performance. In order to verify the effectiveness of proposed algorithm, both numerical simulation and experimental bearing failure data are analyzed.

Keywords: tensor-based singular spectrum analysis; convex optimization; permutation entropy; fault diagnosis

1. Introduction

Rolling bearing is one of the most widely used and easily damaged rotating machine elements, the operation status of which is directly related to the operation of machinery and production efficiency. Thus, it has a very high practical value to achieve accurate diagnosis and recognition of rolling bearing fault. To maintain smooth operation, the diagnostic object should be changed from the traditional super-fault to early weak fault. The diagnostic methods should also be changed from linear analysis methods to non-traditional analysis methods. Many scholars have proposed different methods to accurately identify and extract the fault characteristics in frequency domain. For instance, Short-time Fourier transform (STFT) [1] is proposed to make up the shortcomings of Fourier transform so that it can be used to analyze non-stationary signal, but the time window width of STFT is fixed and cannot be adapted to the changes in the signal. Other non-stationary signal

processing techniques also have a variety of shortcomings. For instance, Empirical Mode Decomposition (EMD) is utilized in the analysis of non-stationary and nonlinear signal [2,3], which easily causes the phenomenon of modal aliasing [4] and edge effect [5]. The selection of decomposition level and wavelet basis function has a great impact on analytical results for Wavelet Transform (WT) [6–8], which lacks in adaption to signal processing. Hilbert–Huang transform (HHT) operator [9] owns windowing effects and the demodulated results will be non-transient, so maybe not accurate. Wigner–Ville distribution [10] for multi-component signal will produce crosstalk and spurious frequency components. In addition, there are some time–frequency analysis methods applied to speed signals in rotating systems [11–17]. In conclusion, these methods have some limitations in non-stationary, non-periodic, non-linear signal processing.

High-dimensional signal contains a lot of information, which is hard to be identified and may not be displayed in one-dimensional space. Therefore, aiming to decompose original signal to extract useful components, high-dimensional signal processing technology may be a breakthrough with great practical engineering significance. Traditional data representational models, such as vectors and matrices, are one-dimensional time series. It is clear that there is some structure in the observed signal allowing one to model it much more compactly. Commonly, the observed signals can be compressible and be expressed as parsimonious models. A higher-order model such as tensor is served as one way to achieve it, which is the most natural representation of high-dimensional data and maintains the intrinsic structural properties of the data. Theoretically, tensors can be seen as higher-order generalizations of vectors and matrices. Tensorization of the observed signal using segmentation and decomposition of the generated tensor are two main aspects of our approach. Tensor decomposition can dig out the potential high-dimensional data structure in the form of low-dimensional subspace, thus be able to reflect the essential characteristics. Generally, the concept of tensor was first proposed by Hitchcock [18]. In 1970, to the psychometrics community, the tensor is developed to the form of CANDECOMP (canonical decomposition) by Carroll [19] and PARAFAC (parallel factors) by Harshman [20]. Then, the research of tensor has gradually been extended to other areas such as signal processing, computer vision, numerical analysis and data mining. CANDECOMP/PARAFAC tensor decomposition, namely CP decomposition model, is a commonly used method. The CP decomposition factorizes a tensor into a sum of R-component rank-one tensors [21]. By the proposed model, three factor matrixes representing the combination of the vectors can be obtained from the rank-one components. Recently, tensor-based singular spectrum algorithm (TSSA) was proposed by Saeid [22] and has been applied to the field of EEG signal processing, which provides an effective way for solving the above problem. The main idea is the one-dimensional times series can be segmented as a matrix using a non-overlapping window. Then each row of matrix can be expressed as a reconstructed attractor matrix through phase space reconstruction [23]. The reconstructed attractor matrix forms the corresponding slice of the tensor, thus we have a 3D tensor to be decomposed. After that, the above-mentioned CP tensor decomposition method is used here. The key step is performed by the alternating least squares method (ALS) [24] to obtain the three factor matrices. TSSA method combines the advantages of phase space reconstruction, singular spectrum analysis [25] as well as tensor decomposition. However, the rank of tensor is difficult to estimate and the optimal selection among the decomposed tensor for reconstruction has a great influence on the final result.

In this paper, an improved TSSA decomposition method is proposed by applying the convex optimization for the rank estimation and permutation entropy (PE) for desired tensor selection. The rank of tensor in TSSA is usually identified by the number of non-zero singular value, while the singular value decomposition (SVD) algorithm is easily affected by the noise interference [26,27]. Consequently, a novel SVD algorithm using convex optimization framework is introduced in this paper. A parameterized non-convex penalty function is put forward to the singular value decomposition and the number of non-zero singular values is used to determine the rank of tensor [28–31]. Bandt [32] proposed the concept of average entropy parameter in the application of measuring the complexity of one-dimensional time series, namely PE. It is an algorithm study about describing irregular and nonlinear systems [33], which cannot or hard to be quantitatively described,

in a relatively simple way. The PE has the advantages of simple calculation, strong anti-noise ability, and high sensitivity to signal change. It can also be applied to the detection of the weak signal [34], and the dynamic mutation of complex system. When an abnormal failure or different kinds of faults occur in gear and bearing parts during operation process, the influence of the nonlinear factors and signal complexity are different. Hence, the vibration signals obtained from the mechanical system will also change, resulting in different values of PE. Thus, PE is employed as novel approach to select the desired tensor for reconstruction in this paper. In order to verify the validity of the proposed method, bearing fault vibration signal and fault simulation test bed are used for analysis. The proposed method is also compared with the traditional singular spectrum analysis (SSA) and EMD.

The structure of this paper is arranged as follows: In Section 2, tensor singular spectrum decomposition algorithm based on PE is briefly described. Then, the performance of abnormal signal detection using PE is focused on. The analysis results of numerical simulation signal and bearing fault signal are, respectively, described in the Sections 3 and 4. Section 5 introduces the conclusions.

2. Theory Description

2.1. Tensor Singular Spectrum Analysis

The proposed TSSA method in this paper mainly contains two stages. The first stage includes an embedding operation and the second stage is to decompose a 3D tensor. In the stage of embedding progress, a one-dimensional time series $x(i), i = 1, 2, \dots, n$ with length n is mapped into a 3D tensor \underline{X} .

There are two steps in the embedding stage. Firstly, $x(i)$ is segmented by a non-overlapping window with the size L . A $[n/L] \times L$ matrix T is obtained from $x(i)$ (where $I = n/L$):

$$T = \begin{pmatrix} X(1) & X(2) & \dots & X(L) \\ X(L+1) & X(L+2) & \dots & X(2L) \\ \dots & \dots & \dots & \dots \\ X(I-1)*L & X(I-1)*L+1 & \dots & X_n \end{pmatrix} \tag{1}$$

Then, the tensor \underline{X} is obtained by matrix T as demonstrated in Figure 1. Each slice of the tensor can be regarded as a reconstructed attractor matrix by a windowed version of matrix T . The segmentation is executed in one direction during this procedure.

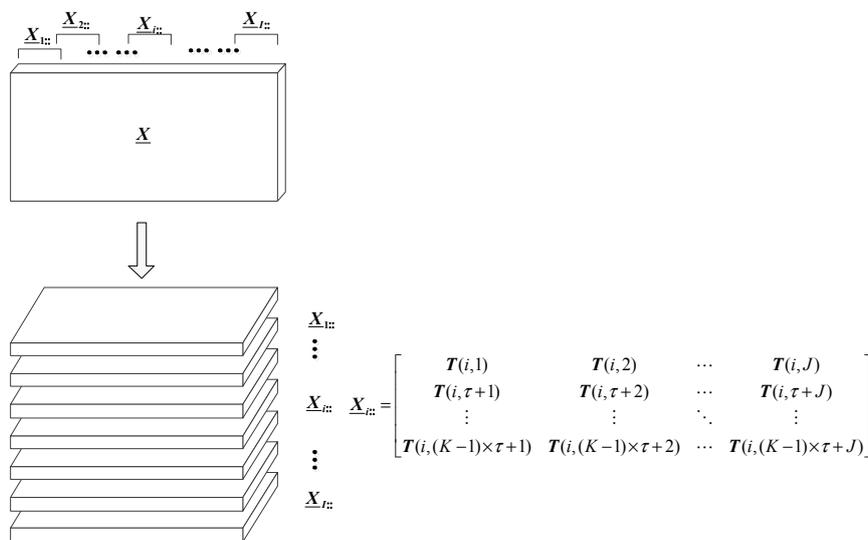


Figure 1. The construction process of the slices of tensor \underline{X} .

The slice $\underline{X}_{i,:}$ of tensor \underline{X} in Figure 1 is formed from the i -th row of matrix T using phase space reconstruction, where J is the length of reconstructed window, K is reconstructed embedding dimension, τ is the delay time. Consequently, the relationship of $L=(K-1)\times\tau+J$ can be obtained. Our way of converting a matrix to tensor can be described as follows:

$$\underline{X}_{ijk}=T(i,(j+(k-1)\times\tau)) \quad i=1,2,\dots,I; j=1,2,\dots,J; k=1,2,\dots,K \quad (2)$$

Thus, we can obtain a 3D tensor of size $I\times J\times K$. In the second stage, the CP tensor decomposition method for a 3D tensor is performed, which factorizes a tensor into a sum of component rank-one tensors. Naturally, the generalized bilinear principal component analysis is similar with the tensor decomposition [35]. Generally, the detailed CP model can be expressed as [36,37]:

$$\underline{X}=\sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r + \underline{e} \quad (3)$$

where R is corresponding to the rank of tensor \underline{X} . $\mathbf{a}_r \in R^{I \times 1}$, $\mathbf{b}_r \in R^{J \times 1}$ and $\mathbf{c}_r \in R^{K \times 1}$ are the vector elements of factor matrix $\mathbf{A} \in R^{I \times R}$, $\mathbf{B} \in R^{J \times R}$ and $\mathbf{C} \in R^{K \times R}$. The tensor of $\underline{e} \in R^{I \times J \times K}$ is the residual term. Hence, the CP model can be approximately explained by factor matrix $\llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$:

$$\underline{X} \approx \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket = \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r + \underline{e} \quad (4)$$

The above mentioned decomposition model is shown in Figure 2.

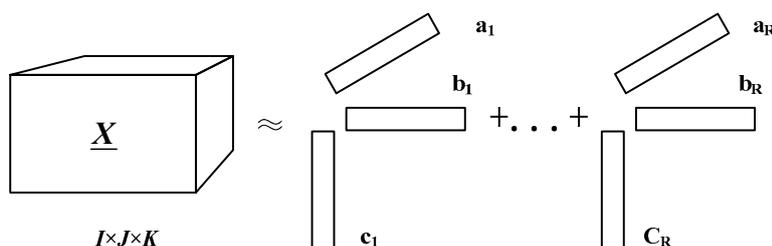


Figure 2. The illustration of CP model for a 3D tensor.

Since TSSA algorithm uses the iterative least squares method to seek factor matrix; we can define an error function as follows:

$$f(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \frac{1}{2} \|\underline{X} - \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket\|^2 = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (\underline{X}_{ijk} - \sum_{r=1}^R a_{ir} b_{jr} c_{kr})^2 \quad (5)$$

First, the random factor matrix is set as an initial matrix to \mathbf{A} , \mathbf{B} , \mathbf{C} . Then, \mathbf{B} and \mathbf{C} are fixed to solve for \mathbf{A} . Similarly, \mathbf{A} and \mathbf{C} are fixed to solve for \mathbf{B} , and next \mathbf{A} and \mathbf{B} are fixed to solve for \mathbf{C} in an alternating algorithm until reaching the convergence condition.

2.2. The Rank Estimation of Tensor Based on Convex Optimization

Similar to the matrix rank, the problem of rank determination has an important significance for tensor decomposition. This paper firstly proposed an algorithm based on SVD using convex optimization to estimate the rank of tensor. SVD is a classic subspace decomposition algorithm that constructs trajectory matrix based on the observed time series. The singular value is obtained by the trajectory matrix, which represents the different dynamic characteristics about each component in the original time series. For the one-dimensional time series $x(i)$, based on the theory of phase space reconstruction by selecting the appropriate embedding dimension m and the delay time τ , the neighborhood matrix or trajectory matrix can be expressed as:

$$Y = \begin{pmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ x(1) & x(2+\tau) & \cdots & x(2+(m-1)\tau) \\ x(j) & x(j+\tau) & \cdots & x(j+(m-1)\tau) \\ \vdots & \vdots & & \vdots \\ x(K) & x(K+\tau) & \cdots & x(K+(m-1)\tau) \end{pmatrix} \quad (6)$$

where $j=1,2,\dots,K$, $K=n-(m-1)\tau$. Generally, the trajectory matrix Y consists of the useful signal and noise component. Then, SVD is applied to Y by the form of $Y=U\Sigma V^T$, where $\Sigma = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$ is a diagonal matrix, $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, U and V represent the left and right feature vector. The singular value of $\sigma_1, \sigma_2, \dots, \sigma_r$ can be used to identify the useful signals and noise, respectively. If the singular value is arranged in accordance with the descending order, the larger singular values mainly reflect the useful signal and the smaller singular values mainly reflects the noise. The rank is estimated by the eigenvalues of the neighborhood matrix calculated by SVD in the conventional TSSA method. It is common fact that the result provided by SVD is susceptible to noise especially under the strong noise environment. Without loss of generality, it can be considered the problem of estimating a low-rank matrix X from its noisy observation Y .

$$Y = X + G, \quad X, Y, G \in R^{m \times n} \quad (7)$$

where Y represents a noisy neighborhood matrix, X is the expected low rank matrix, and G is the zero-mean additive white Gaussian noise matrix. Note that for unitary matrices U and V , $\Phi(X) = \Phi(UXV)$. Using the unitary invariant property of the Frobenius norm, we can define a typical low-rank matrix approximation (LRMA) issue as follows:

$$\begin{aligned} \hat{X} &= \arg \min_X \left\{ \frac{1}{2} \|Y - X\|_F^2 + \lambda \Phi(X) \right\} \\ &= \arg \min_X \left\{ \frac{1}{2} \|\Sigma - U^T X V\|_F^2 + \lambda \Phi(U^T X V) \right\} \\ &= U \cdot \arg \min_X \left\{ \frac{1}{2} \|\Sigma - X\|_F^2 + \lambda \Phi(X) \right\} \cdot V^T \end{aligned} \quad (8)$$

where $\Phi(X) = \sum_{i=1}^k \phi(\sigma_i(X); a)$, $\sigma_i(X)$ is the singular value of matrix X , $k = \min(m, n)$, $0 \leq a < 1/\lambda$, ϕ is a sparsity promotion function. When it meets $\phi(x) = |x|$, LRMA issue becomes a typical nuclear norm minimization problem (NNM). Note that $\Theta(\Sigma; \lambda, a) = \arg \min_X \left\{ \frac{1}{2} \|\Sigma - X\|_F^2 + \lambda \Phi(X) \right\}$. Therefore, the Equation (8) can be converted to:

$$\hat{X} = U \cdot \Theta(\Sigma; \lambda, a) \cdot V^T \quad (9)$$

It is proved that parametric non-convex penalty function can be more accurate in estimating non-zero singular value than the nuclear norm under strong noise background [29]. Therefore, we use the non-convex penalty function defined by Equation (10) to achieve the optimized estimation of singular value.

$$\Theta(\Sigma; \lambda, a) = \min \left\{ |\Sigma|, \max \left\{ (|\Sigma| - \lambda) / (1 - a\lambda), 0 \right\} \right\} \text{sign}(\Sigma) \quad (10)$$

In this paper, the number of non-zero singular values obtained by the convex optimization framework is equivalent to the rank R .

2.3. The Desired Tensor Selection Based on PE

After the above tensor decomposition and selecting the proper rank of the desired source signal, it is necessary to return it to single channel data, which is likely to be our desired signal components. The desired single channel data from the tensor is obtained by the same hankelisation procedure. Firstly, it is done by hankelizing matrices across the first slab or block hankelizing the tensor. Then, the obtained tensor is hankelized again. Finally, the one-dimensional signal is reconstructed by one-to-one correspondence. It is worth pointing out that the desired components are not a simple superposition of these reconstruction signal. Thus, an effective way to accurately evaluate the useful components should be taken into consideration. Permutation Entropy (PE) is one of the most effective ways to measure the randomness and dynamic mutation of time sequence. Similar to the above operations defined in Section 2.2, the trajectory matrix corresponding to the reconstruction signal is obtained by phase space reconstruction. The matrix in each row is rearranged in ascending order and j_1, j_2, \dots, j_m indicates the index of each element in the column of reconstruction components. A set of symbols sequence can be obtained from each row in the matrix.

$$S(l) = \{j_1, j_2, \dots, j_m\} \quad (11)$$

where $l = 1, 2, \dots, k$, and $k \leq m$. The probability of each symbol sequence was calculated and named as P_1, P_2, \dots, P_k . The value of PE corresponding to different symbol sequences in reconstruction signal is defined as the form of information entropy as follows:

$$H_p(d) = \sum_{j=1}^k P_j \ln(P_j) \quad (12)$$

It is also noted that the value of $H_p(d)$ represents the degree of randomness of time series, namely, the larger value of $H_p(d)$ declares the more randomness of time series, and vice versa [32]. Commonly, the permutation entropy of harmonic signal and the modulation signal is small, while the permutation entropy for random noise is large. Hence, PE can be used to select the desired tensor for signal reconstruction. Ultimately, the flowchart of the proposed method is shown in Figure 3.

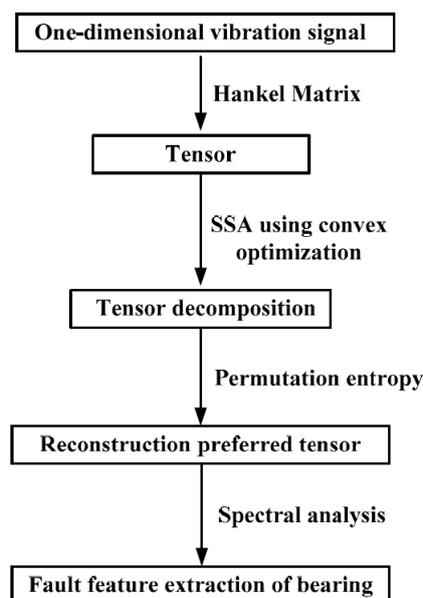


Figure 3. A flow chart of the method presents in this paper.

3. Test with Numerical Simulation Signal

3.1. The Performance of Abnormal Signal Detection Using PE

In the process of permutation entropy calculation, there are two parameter values that need to be determined: the embedding dimension m and delay time τ . Under the condition of the above parameters, the following representative signal is considered:

$$x_1 = [1 + 0.7 \sin(2\pi f_1 t)] \sin(2\pi f_2 t) \tag{13}$$

$$x_2 = 0.8 \sin(2\pi f_3 t) \tag{14}$$

where $f_1 = 5$, $f_2 = 10$, $f_3 = 15$. x_3 is the standard Gaussian white noise. Theoretically, the embedding dimension and delay time are chosen as $m = 6$ and $\tau = 1$ according to the False Nearest Neighbor algorithm (FNN) [38]. According to the definition of PE, the normalized PE values of three simulated signals are calculated as 0.2909, 0.4451 and 0.9644, respectively. The result demonstrated that the PE value corresponding to noisy signal is obviously larger than normal modulating signal and harmonic signal. It is proved that PE is sensitive to noisy signal and can be used to detect the abnormal signal. For the proposed method, the value of PE corresponding to the undesired tensor by hankelisation procedure is greater than the threshold value, which is defined as 0.6.

3.2. The Feature Extraction Result Provided by Proposed Method

Rolling bearings are mainly used to support the rotating part in the mechanical equipment and its vibration signal always contains a lot of information such as fault characteristic and noise component. The key step in fault diagnosis is to extract vibration signal feature in frequency domain. There are many kinds of the simulation signal model about bearing fault and the most typical one is proposed by Randall [39,40]. Without loss of generality, the numerical simulation signal is expressed as follows:

$$x(t) = \sum_{i=0}^M [A_0 \cdot \cos(2\pi f_m t + \phi_A) + C_A] \cdot e^{-B(t-iT-\tau_i)} \cdot \cos[2\pi f_n(t-iT-\tau_i) + \phi_w] + n(t) \tag{15}$$

where A_0 is the amplitude of resonance; f_m is the modulation frequency; ϕ_A , ϕ_w , and C_A are selected as arbitrary constants; B is the attenuation coefficient; T is defined as the average time between two impacts with $T=1/f_p$; f_p is fault character frequency; τ_i is regarded as the time lag from its mean period due to the presence of slip; f_n is the resonance frequencies of the bearing system; and $n(t)$ is the additive white Gaussian noise. The three common types of faults on outer race, inner race and rolling elements can be modeled as $f_m = 0$, $f_m = f_r$, $f_m = f_{re}$, respectively. It should also be noted that f_r and f_{re} are denoted as rotational frequency and retainer frequency. The fault character frequencies of three common types as f_i, f_o, f_{ro} are described in Table 1. Firstly, the outer race fault feature extraction under noise environment is researched and the parameter selection is listed as Table 2. Moreover, $n(t)$ is Gaussian white noise with variance 0.5 in the section of simulation signal analysis.

Table 1. The fault character frequency of three common types.

Inner Race f_i (Hz)	Outer Race f_o (Hz)	Rolling Element f_{ro} (Hz)
156	103.9	59.7

Table 2. The parameter selection of outer race fault simulation signal.

A_0	f_r (Hz)	f_m (Hz)	τ_i	f_n (Hz)	C_A	ϕ_A	ϕ_w	B
0.0003	17	0	0.01	2000	1	0	0	800

The sampling frequency and the sampling points are set as 10,000 Hz and 2000, respectively. Original signal without noise and noisy synthesis simulation signal in time-domain are shown in Figure 4a,b, respectively. The frequency spectrum about $x(t)$ is conducted in Figure 5.

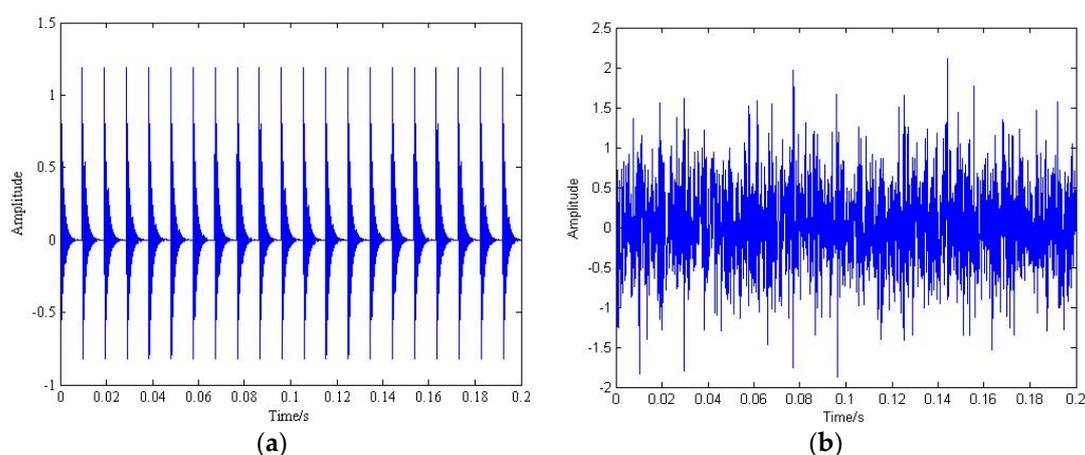


Figure 4. The original signal and the noisy signal in time domain: (a) the original signal without noise; and (b) the noisy signal.

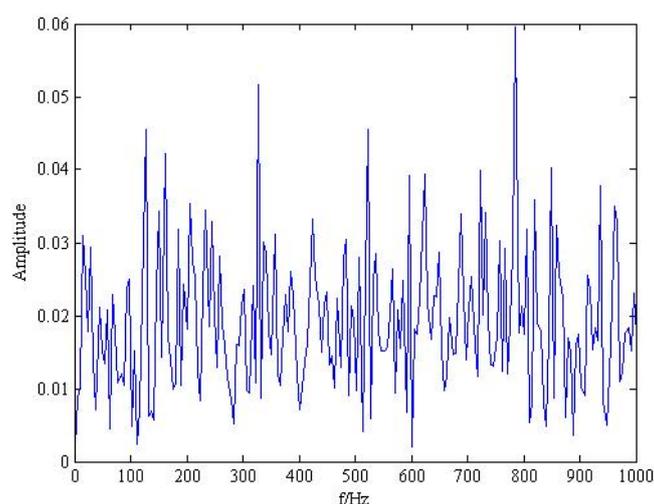


Figure 5. The simulation signal with noise in frequency domain.

In all the simulations and experiments, the value of L and J is set to 200 and 50, respectively. Commonly, the delay time τ is selected as 1 in order to keep the computational complexity low. Then, the parameter of I and K can be obtained by the relation of $I = n/L$, $L = (K-1)\times\tau + J$. In addition, the rank of tensor is determined by the convex optimization framework introduced in Section 2.2, and the desired tensor selection is performed by the index of PE in Section 3.1. According to Figure 4b, it is acceptable that the outer fault characteristics of the original simulation signal in the time domain cannot be clearly indicated in the strong background noise. For the frequency analysis in Figure 5, the outer race fault characteristic frequency is still concealed by noisy components.

In the simulation signal analysis section, in order to accurately evaluate the proposed method in signal reconstruction under noisy conditions, the proposed method, the conventional TSSA method and wavelet transform (WT) are used to the comparative analysis. Figure 6 is the result provided by conventional TSSA model and only the third harmonic ($3f_o$) can be inspected. It is important to note that the “wden” function is chosen as the automatic 1-D denoising function and the wavelet basis function is set as “sym3” for wavelet denoising in the rest of this paper. Figure 7 is a diagram of wavelet denoising in time domain and frequency domain. It is obvious that both the conventional TSSA and WT cannot completely identify the outer race fault characteristic frequency. Figure 8 is the

result provided by the proposed method based on PE in frequency domain, which indicates the outer race feature frequency f_o and its frequency multiplication such as $2f_o$, $3f_o$, $4f_o$ can be identified. Then, the proposed method is applied to characteristic frequency extraction for inner race fault and rolling elements fault. The result provided by the proposed method is drawn in Figure 9 and the modulation phenomenon is in keeping with the inner ring and rolling elements fault characteristics. Thus, we can make a conclusion that the proposed method has better performance in denoising and characteristic frequency extraction.

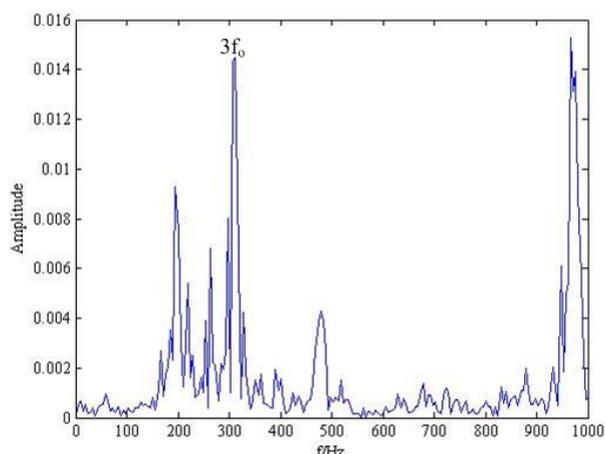


Figure 6. The result provided by conventional TSSA model in frequency-domain.

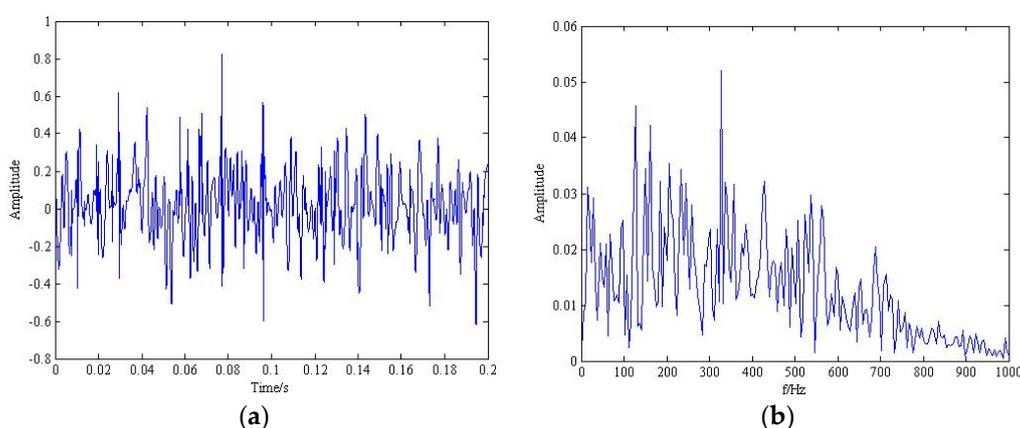


Figure 7. The result provided by wavelet transform in time and frequency domain: (a) wavelet denoising in time domain; and (b) wavelet denoising in frequency domain.

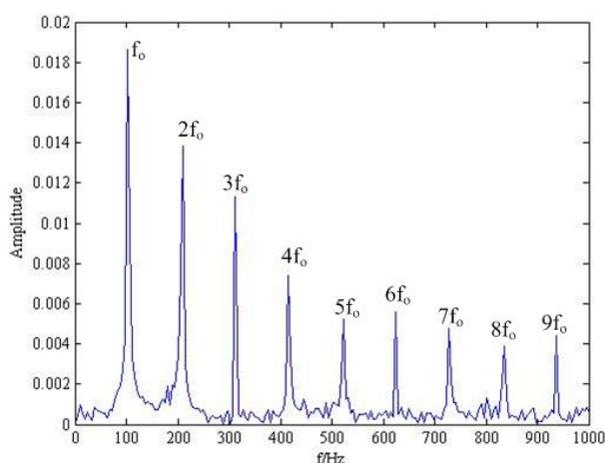


Figure 8. The analysis result provided by the proposed method for outer race fault.

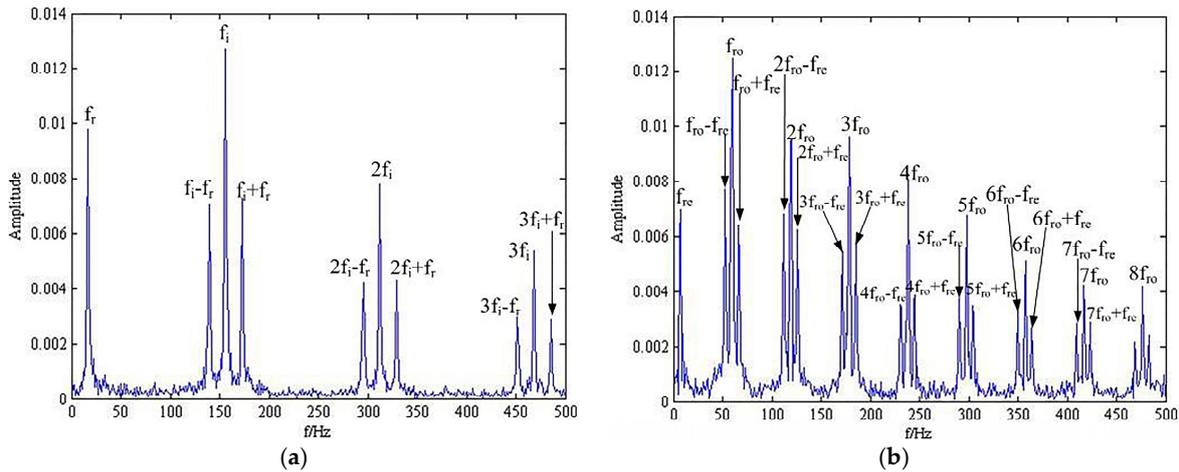


Figure 9. The analysis result provided by the proposed method for inner race and rolling elements: (a) inner race fault; and (b) rolling elements fault.

It can be seen from the above analysis result provided by different methods that conventional TSSA and WT cannot completely separate the noisy component from the original signal. However, the proposed tensor singular spectrum algorithm based on PE is more effective in removing noise component, better in reflecting the fault characteristics of the original signal. Through multiple simulations, the improved TSSA method has higher convergence and better reconstruction precision than TSSA.

For a more detailed description of the degree of noise reduction, the paper selects signal-to-noise ratio (SNR) as an index for evaluation, which is critical for assessing the degrees of noise contamination. The SNR of the noisy signal about outer race fault simulation was varied between 2.17 dB and 0.32 dB, while the corresponding reconstruction SNR of the simulation signal was then calculated and the result is plotted in the following Figure 10. The smaller values always indicate the poor reconstruction accuracy. As can be seen in Figure 10, the results also indicate that the proposed method performs well in denoising and signal reconstruction.

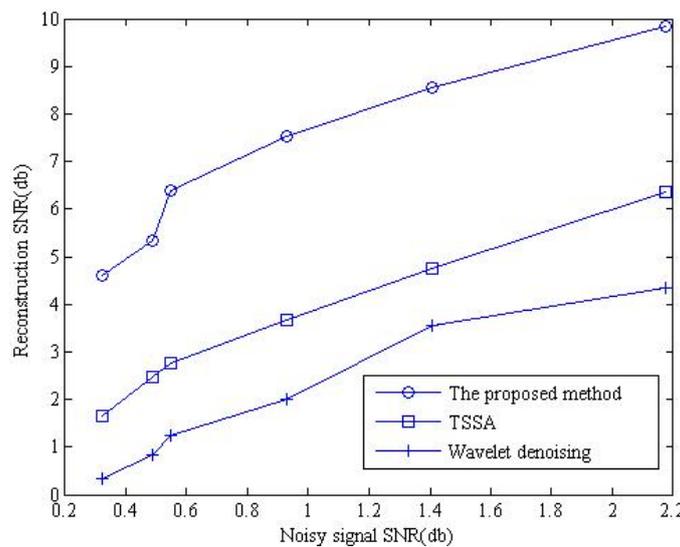


Figure 10. Reconstruction error (in SNR) using the proposed method, conventional TSSA and wavelet denoising algorithms.

4. Applications to Rolling Bearing Fault Feature Extraction

In industrial practice, the measured vibration signal is more complicated than the simulated signal. In order to verify the effectiveness of this approach in the real environment, the experimental device analog failure signal was measured to be analyzed. Simulation apparatus is shown in Figure

11, where bench consists of a 550 W (220 V/50 Hz) AC motor driven drive shaft through the coupling operation. The electric spark machining method was used to carry out pitting treatment on the outer ring of replaceable bearing to simulate the outer faults and the depth of pitting corrosion pit is close to 1.0 mm. The acceleration signal of the experiment was collected in the vertical direction of the bearing on the right side of the experimental platform using a CSI2130 data analyzer (Emerson, Sao Luis, Brazil) and vibration acceleration sensor PCB-352C33 (PCB, New York, NY, USA). The bearing type is deep groove ball bearing with model number of 6207. The medium diameter of removable bearing is $D = 53.5$ mm, the number of rolling elements is $Z = 9$, and the roller diameter is $d = 11.1$ mm. The detailed parameters of the experiment are shown in Table 3.

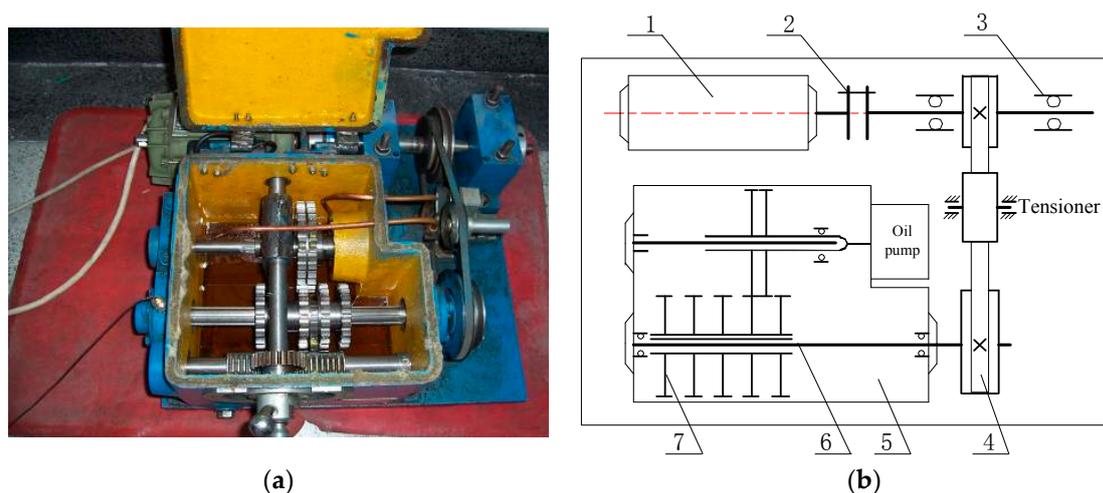


Figure 11. Fault simulation experimental table of roll bearing: (a) the physical map of test rig; and (b) the structure diagram of test rig: 1—electric motor, 2—coupler, 3—replaceable rolling bearing with 6207, 4—belt pulley, 5—gearbox, 6—transmission shaft, 7—fault gear sets.

Table 3. The experimental parameters and fault frequency.

Rotating Speed r/min	Rotating Frequency/Hz	Sampling Frequency/Hz	Sampling Time/s	Outer Fault Frequency/Hz
1450	24.17	16384	1	87.01

It is worth pointing out that the frequency of outer ring fault signal of rolling bearing is $f_o = 87.01$ Hz by calculating, and the rotation frequency is $f_r = 24.17$ Hz. The original signal in time domain and frequency domain is shown in Figure 12a,b, respectively.

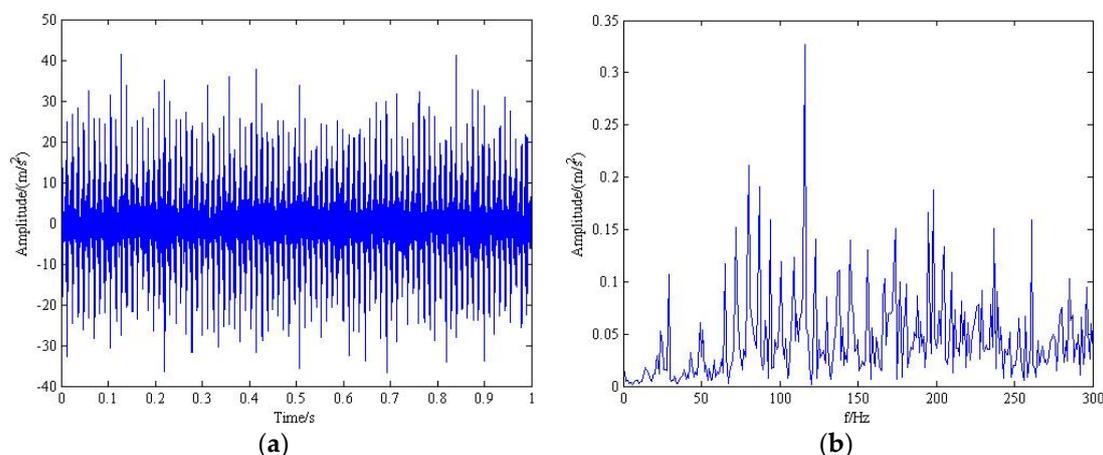


Figure 12. The original vibration signal in time domain and frequency domain: (a) the measured signal in time-domain; and (b) the measured signal in frequency-domain.

From time-domain waveform in Figure 12a,b, it cannot be determined where the failure occurred. Meanwhile, in order to illustrate the effectiveness the proposed method, the conventional noise reduction method such as wavelet denoising algorithm is applied and the results are shown in Figure 13.

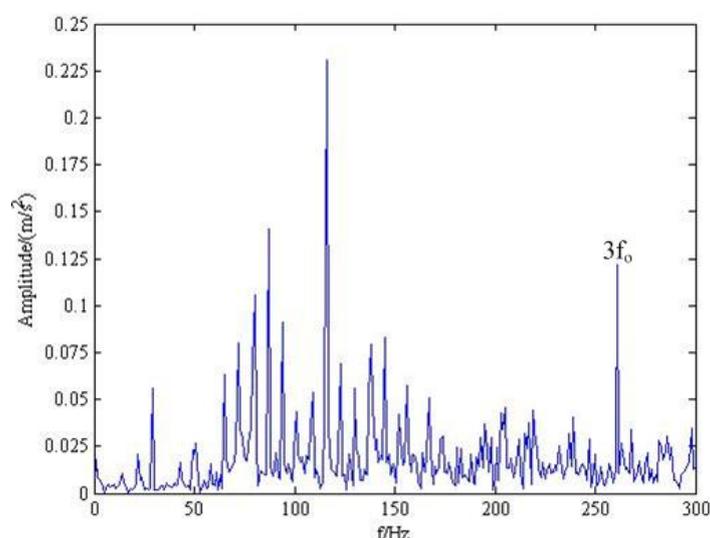


Figure 13. The result provided by wavelet denoising.

After the wavelet denoising, only the three multiple of outer fault feature frequency $3f_0$ can be inspected, while the other useful components are still difficult to be accurately identified.

Based on the False Nearest Neighbor algorithm (FNN) [38], the embedding dimension and delay time is determined as $m=12$ and $\tau=1$ in the experimental case studies. Finally, the proposed method based on PE is utilized to signal denoising and feature extraction, and the result is shown in Figure 14.

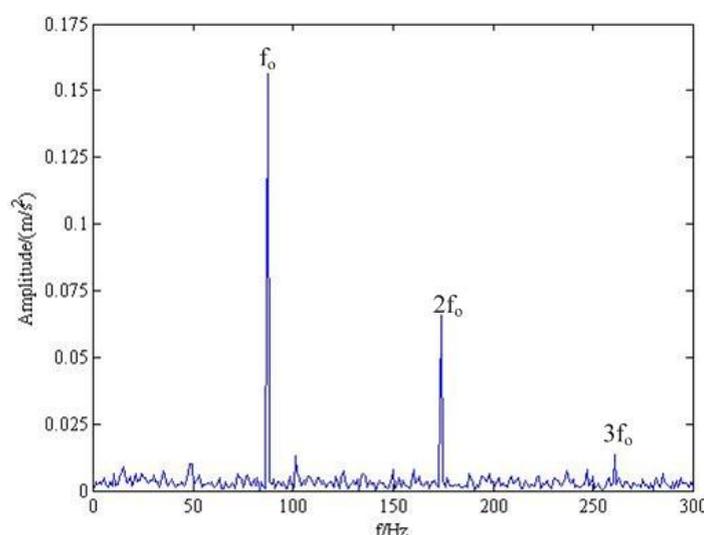


Figure 14. The result provided by the proposed method.

Through the tensor decomposition using PE, the unwanted signal components have been deducted and the feature information can be enhanced. It is obvious that the outer ring fault frequency f_0 and its harmonics ($2f_0$, $3f_0$) all can be identified in Figure 14. Hence, we can make a conclusion that the outer ring of the rolling bearing has occurred failure, which is consistent with the actual situation.

5. Conclusions

This paper improves the tensor singular spectrum analysis (TSSA) algorithm to address its problems of inestimable tensor rank and lacking in criterion to achieve the desired tensor selection. The main work of this paper was reflected in the following aspects: (1) the accurate estimation of the rank of tensor was solved by SVD using convex optimization and non-convex penalty function; (2) the desired tensor selection using PE was conducted to improve the denoising performance of the proposed method; and (3) in the processing of measured rolling bearing fault signal, the proposed improved TSSA method based on PE outperformed other methods in extracting weak fault characteristics, which demonstrated that the proposed method was feasible in the future real applications.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (No. 51475339); the Natural Science Foundation of Hubei province (No. 2016CFA042); and the State Key Laboratory of Refractories and Metallurgy, Wuhan University of Science and Technology (No. ZR201603).

Author Contributions: Cancan Yi, Yong Lv and Han Xiao conceived and designed the experiments; Cancan Yi and Mao Ge performed the experiments; Cancan Yi and Yong Lv analyzed the data; Cancan Yi and Han Xiao contributed reagents/materials/analysis tools; Cancan Yi and Xun Yu wrote the paper.

Conflict of Interests: The authors declare no conflict of interest.

Reference

1. Veer, K.; Agarwal, R. Wavelet and short-time Fourier transform comparison-based analysis of myoelectric signals. *Journal Appl. Stat.* **2015**, *42*, 1591–1601.
2. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.-C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*; The Royal Society: London, UK, 1998; Volume 454, pp. 903–995.
3. Lv, Y.; Yuan, R.; Song, G. Multivariate empirical mode decomposition and its application to fault diagnosis of rolling bearing. *Mech. Syst. Signal Process.* **2016**, *81*, 219–234.
4. Yang, Z.; Zhong, J. A hybrid EEMD-based SampEn and SVD for acoustic signal processing and fault diagnosis. *Entropy* **2016**, *18*, 112.
5. Wu, F.; Qu, L. An improved method for restraining the end effect in empirical mode decomposition and its applications to the fault diagnosis of large rotating machinery. *J. Sound Vib.* **2008**, *314*, 586–602.
6. Chen, J.; Dou, Y.; Li, Y.; Li, J. Application of Shannon Wavelet Entropy and Shannon Wavelet Packet Entropy in Analysis of Power System Transient Signals. *Entropy* **2016**, *18*, 437.
7. Wang, D.; Tsui, K.L. Dynamic Bayesian wavelet transform: New methodology for extraction of repetitive transients. *Mech. Syst. Signal Process.* **2017**, *88*, 137–144.
8. Wang, D. An extension of the infograms to novel Bayesian inference for bearing fault feature identification. *Mech. Syst. Signal Process.* **2016**, *80*, 19–30.
9. Huang, N.E. *Hilbert-Huang Transform and Its Applications*; World Scientific: Singapore, 2014.
10. Abeysekera, S.S.; Boashash, B. Methods of signal classification using the images produced by the Wigner-Ville distribution. *Pattern Recognit. Lett.* **1991**, *12*, 717–729.
11. Odgaard, P.F.; Stoustrup, J. Gear-box fault detection using time-frequency based methods. *Annu. Rev. Control* **2015**, *40*, 50–58.
12. Nejad, A.R.; Odgaard, P.F.; Gao, Z.; Moan, T. A prognostic method for fault detection in wind turbine drivetrains. *Eng. Fail. Anal.* **2014**, *42*, 324–336.
13. Odgaard, P.F.; Nejad, A.R. Frequency based wind turbine gearbox fault detection applied to a 750 kW wind turbine. In *Proceedings of the IEEE Conference on Control Applications, Antibes, France, 8–10 October 2014*; 1383–1388.
14. Fadlallah, B.; Chen, B.; Keil, A.; Principe, J.C. Weighted-permutation entropy: A complexity measure for time series incorporating amplitude information. *Phys. Rev. E* **2013**, *87*, 022911.
15. Chen, B.; Wang, J.; Zhao, H.; Principe, J.C. Insights into entropy as a measure of multivariate variability. *Entropy* **2016**, *18*, 196.

16. Chen, B.; Yuan, Z.; Zheng, N.; Principe, J.C. Kernel minimum error entropy algorithm. *Neurocomputing* **2013**, *121*, 160–169.
17. Chen, B.; Zhu, Y.; Hu, J. Mean-square convergence analysis of ADALINE training with minimum error entropy criterion. *IEEE Trans. Neural Netw.* **2010**, *21*, 1168–1179.
18. Hitchcock, F.L. The expression of a tensor or a polyadic as a sum of products. *J. Math. Phys.* **1927**, *6*, 164–189.
19. Carroll, J.D.; Chang, J.J. Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition. *Psychometrika* **1970**, *35*, 283–319.
20. Harshman, R.A. Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis. *UCLA Work. Pap. Phon.* **1970**, *16*, 1–84.
21. Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. *SIAM Rev.* **2009**, *51*, 455–500.
22. Kouchaki, S.; Sanei, S.; Arbon, E.L.; Dijk, D.-J. Tensor based singular spectrum analysis for automatic scoring of sleep EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2015**, *23*, 1–9, doi: 10.1109/TNSRE.2014.2329557.
23. Han, L.; Romero, C.E.; Yao, Z. Wind power forecasting based on principle component phase space reconstruction. *Renew. Energy* **2015**, *81*, 737–744.
24. Tomasi, G.; Bro, R. PARAFAC and missing values. *Chemom. Intell. Lab. Syst.* **2005**, *75*, 163–180.
25. Golyandina, N.; Zhigljavsky, A. *Singular Spectrum Analysis for Time Series*; Springer: Berlin/Heidelberg, Germany, 2013.
26. Nocedal, J.; Wright, S.J. Penalty and augmented Lagrangian methods. In *Numerical Optimization*; Springer: New York, NY, USA, 2006; pp. 497–528.
27. Rufai, M.A.; Anbarjafari, G.; Demirel, H. Lossy image compression using singular value decomposition and wavelet difference reduction. *Digit. Signal Process.* **2014**, *24*, 117–123.
28. Parekh, A.; Selesnick, I.W. Enhanced low-rank matrix approximation. *IEEE Signal Process. Lett.* **2016**, *23*, 493–497.
29. Selesnick, I.W.; Bayram, İ. Enhanced Sparsity by Non-Separable Regularization. *IEEE Trans. Signal Process.* **2015**, *64*, 2298–2313.
30. Yi, C.; Lv, Y.; Dang, Z.; Xiao, H. A Novel Mechanical Fault Diagnosis Scheme Based on the Convex 1-D Second-Order Total Variation Denoising Algorithm. *Appl. Sci.* **2016**, *6*, 403.
31. Yi, C.; Lv, Y.; Dang, Z.; Xiao, H.; Yu, X. Quaternion singular spectrum analysis using convex optimization and its application to fault diagnosis of rolling bearing. *Measurement* **2017**, *103*, 321–332.
32. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102.
33. Theodossiades, S.; Natsiavas, S. Non-linear dynamics of gear-pair systems with periodic stiffness and backlash. *J. Sound Vib.* **2000**, *229*, 287–310.
34. Zhao, H.; Sun, M.; Deng, W.; Yang, X. A New Feature Extraction Method Based on EEMD and Multi-Scale Fuzzy Entropy for Motor Bearing. *Entropy* **2016**, *19*, 14.
35. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52.
36. Harshman, R.A.; Berenbaum, S.A. Basic concepts underlying the PARAFAC-CANDECOMP three-way factor analysis model and its application to longitudinal data. In *Present and Past in Middle Life*; Academic Press: New York, NY, USA, 1981.
37. Burdick, D.S. An introduction to tensor products with applications to multiway data analysis. *Chemom. Intell. Lab. Syst.* **1995**, *28*, 229–237.
38. Kennel, M.B.; Brown, R.; Abarbanel, H.D.I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **1992**, *45*, 3403.
39. Randall, R.B.; Antoni, J.; Chobsaard, S. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mech. Syst. Signal Process.* **2001**, *15*, 945–962.
40. Antoni, J.; Randall, R.B. A stochastic model for simulation and diagnostics of rolling element bearings with localized faults. *J. Vib. Acoust.* **2003**, *125*, 282–289.

