

## Article

# Hierarchical Amplitude-Aware Permutation Entropy-Based Fault Feature Extraction Method for Rolling Bearings

Zhe Li <sup>1</sup> , Yahui Cui <sup>1</sup>, Longlong Li <sup>1,\*</sup>, Runlin Chen <sup>1</sup>, Liang Dong <sup>1</sup> and Juan Du <sup>2</sup>

<sup>1</sup> School of Mechanical and Precision Instrument Engineering, Xi'an University of Technology, Xi'an 710048, China; 1200211012@stu.xaut.edu.cn (Z.L.); cyhxut@xaut.edu.cn (Y.C.); chenrunlin@xaut.edu.cn (R.C.); 1200213020@stu.xaut.edu.cn (L.D.)

<sup>2</sup> Department of Basic, Air Force Engineering University, Xi'an 710051, China; juandu@stu.xidian.edu.cn

\* Correspondence: 205173@xaut.edu.cn

**Abstract:** In order to detect the incipient fault of rolling bearings and to effectively identify fault characteristics, based on amplitude-aware permutation entropy (*AAPE*), an enhanced method named hierarchical amplitude-aware permutation entropy (*HAAPE*) is proposed in this paper to solve complex time series in a new dynamic change analysis. Firstly, hierarchical analysis and *AAPE* are combined to excavate multilevel fault information, both low-frequency and high-frequency components of the abnormal bearing vibration signal. Secondly, from the experimental analysis, it is found that *HAAPE* is sensitive to the early failure of rolling bearings, which makes it suitable to evaluate the performance degradation of a bearing in its run-to-failure life cycle. Finally, a fault feature selection strategy based on *HAAPE* is put forward to select the bearing fault characteristics after the application of the least common multiple in singular value decomposition (LCM-SVD) method to the fault vibration signal. Moreover, several other entropy-based methods are also introduced for a comparative analysis of the experimental data, and the results demonstrate that *HAAPE* can extract fault features more effectively and with a higher accuracy.



**Citation:** Li, Z.; Cui, Y.; Li, L.; Chen, R.; Dong, L.; Du, J. Hierarchical Amplitude-Aware Permutation Entropy-Based Fault Feature Extraction Method for Rolling Bearings. *Entropy* **2022**, *24*, 310. <https://doi.org/10.3390/e24030310>

Academic Editor: Luca Faes

Received: 5 January 2022

Accepted: 21 February 2022

Published: 22 February 2022

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



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

**Keywords:** hierarchical amplitude-aware permutation entropy; rolling bearing; performance trend state assessment; fault feature extraction

## 1. Introduction

The condition monitoring of rotating machinery is a fundamental task in prognostics and health management systems [1]. As an essential part of rotating machinery, rolling bearings are critical and more easily damaged components compared to others [2]. If a rolling bearing has a local defect on the matching surface, a series of shocks are generated due to periodic collisions when the rolling element rolls over the defect area. These abnormal shocks may result in an unexpected failure of large-scale equipment if without timely inspection and maintenance [3,4]. To ensure the reliable operation of rolling bearings, a degradation evaluation of its running state and a fault diagnosis, two indispensable aspects of modern advanced rotating equipment, can be employed. Additionally, with these two techniques, considerable maintenance costs can be saved. Therefore, the performance evaluation and fault diagnosis of rolling bearings have become a major topic of research.

Nowadays, in the condition monitoring of rotating machinery, vibration-signal-based processing techniques are one of the most used methods for fault diagnosis, which can be classified into three categories: time-domain analysis (e.g., root mean square (RMS) and kurtosis), frequency-domain analysis (e.g., fast Fourier transform and envelope spectrum demodulation), and time–frequency analysis [5]. In addition, time–frequency analysis has become a popular method for rolling bearing fault feature extraction, and wavelet transform (WT) [6], empirical mode decomposition (EMD) [7], ensemble empirical mode decomposition (EEMD) [8], and variational mode decomposition (VMD) [9] are several typical time–frequency analysis algorithms. However, the parameter selection of the above

methods has some influence on their performance [10–12]. By contrast, as a matrix decomposition method, singular value decomposition (SVD) is a nonparametric algorithm that has been widely used in the noise reduction field [13,14]. The denoising method based on SVD eliminates the noise and defect unrelated components by setting the singular values (SVs) corresponding to the noise-related components to 0; thus, the informative components can be reconstructed from the persevered SCs [15]. The critical issue of reconstruction is to determine the dimension of the Hankel matrix. Generally, the dimension of the Hankel matrix can be reconstructed using the maximum dimension method, which depends on the length of the signal but ignores the signal frequency characteristics. The least common multiple in SVD (LCM-SVD) method has been introduced to address this problem, and it determines the reconstructed matrix's dimension based on frequency factor analysis [16].

Regarding the nonlinear dynamic analysis method of rolling bearing signals, entropy is developed into an effective approach to measure the dynamic change in complex time series in fault diagnosis [17]. Sample entropy (SE), fuzzy entropy (FE), and permutation entropy (PE) [18] algorithms are several entropy-based indexes that are extensively employed. Compared with the SE and FE approaches, PE has many advantages, such as clean theory, fast calculation efficiency, and robust anti-noise performance [19]. PE is always used to estimate the different patterns of orderings that are scattered in a time series and plays a significant role in feature extraction [20]. However, the direct application of PE in feature extraction does not ensure the data's denoising effect; hence, it is often used after a noise removal method [21]. Apart from this, PE is disabled to assess the amplitude information of the signal. To settle this issue, Azami et al., proposed an entropy algorithm named amplitude-aware permutation entropy (AAPE) based on PE, which improves its sensitivity to the amplitude and frequency of the time series by using the counting rule for each sorting mode [22,23]. However, single-scale entropy, such as SE, FE, PE, and AAPE, may lose key information during time series analysis. To solve this problem, multi-scale entropy (MSE) was first proposed by Costa M. et al., to measure the complexity of a time series [24]. Subsequently, multi-scale fuzzy entropy (MFE) [25], multi-scale permutation entropy (MPE) [26], and multi-scale amplitude-aware permutation entropy (MAAPE [27]) and its variants (IMAAPE [28]) were proposed sequentially. However, multi-scale entropy analysis ignores the effect of high-frequency components. To obtain more helpful information, Jiang et al., put forward a novel multi-scale analysis method called hierarchical analysis, which can reveal the inherent fault characteristics of a vibration signal in both low- and high-frequency components [29,30]. Derived from hierarchical analysis, hierarchical sample entropy (HSE) [31], hierarchical fuzzy entropy (HFE) [32], and hierarchical permutation entropy (HPE) [33] were proposed to extract fault characteristics.

Inspired by the advantages of hierarchical analysis and the AAPE algorithm, the HAAPE method for rolling bearings is proposed, and it includes the following two aspects: performance degradation assessment and fault diagnosis. In the performance degradation assessment stage, HAAPE is employed as an index to quantify the running state of the bearing in its run-to-failure life cycle, and Chebyshev's inequality theory is applied to establish a health threshold to judge the bearing performance trend between the healthy state and abnormal state. In the fault diagnosis stage, HAAPE is used after the LCM-SVD procedure to identify the bearing failure characteristics. Two experimental cases demonstrate that the proposed HAAPE can effectively extract the fault characteristics, and the experimental results show that HAAPE has superiority over existing methods in the prediction of performance degradation trend and fault diagnosis of rolling bearings.

The remainder of this article is constructed as follows: Section 2 introduces the related theoretical basis and the fundamentals of HAAPE. Simulation signals are used to verify the effectiveness of the proposed method in Section 3. The experimental results are arranged in Section 4, which includes two experimental cases, namely, case 1 and case 2. Finally, several conclusions are drawn in Section 5.

## 2. Methodologies

### 2.1. Amplitude-Aware Permutation Entropy (AAPE)

A one-dimensional time series  $X(i), i = 1, 2, \dots, N$  is denoted as the original signal. At any time point  $j$ , the  $m$  dimensional reconstruction vector  $X_j^{m,\tau}$  with the delay time  $\tau$  can be obtained as follows:

$$X_j^{m,\tau} = \begin{bmatrix} x_{(1)} & x_{(1+\tau)} & \cdots & x_{[1+(m-1)\tau]} \\ x_{(2)} & x_{(2+\tau)} & \cdots & x_{[2+(m-1)\tau]} \\ x_{(j)} & x_{(j+\tau)} & \cdots & x_{[j+(m-1)\tau]} \\ \vdots & \vdots & \ddots & \vdots \\ x_{(k)} & x_{(k+\tau)} & \cdots & x_{[k+(m-1)\tau]} \end{bmatrix}, j = 1, 2, \dots, k = N - (m - 1)\tau \quad (1)$$

The elements in the  $j$ -th reconstructed vector are arranged in ascending order as shown in Equation (2).

$$x[i + (j_1 - 1)\tau] \leq x[i + (j_2 - 1)\tau] \leq \dots \leq x[i + (j_m - 1)\tau] \quad (2)$$

where  $j_1, j_2, \dots, j_m$  are the index of the column of each element in the reconstructed vector. Hence, there are  $m!$  potential ordinal patterns, of which the  $d$ -th permutation is marked as  $\pi_d$ . The occurrence times of each  $\pi_d$  are calculated to obtain their occurrence probability. Finally, based on the definition of Shannon entropy, PE can be defined.

$$PE(x, m, \tau) = - \sum_{\pi_d=1}^{m!} p(\pi_d) \ln p(\pi_d) \quad (3)$$

However, there are two main shortcomings in the application process: (1) PE only considers the sequence number structure of the time series rather than the amplitude difference between sequential samples, which results in the loss of some valuable characteristics. (2) Under the appearance of equal amplitude, the difference in permutation patterns cannot be distinguished using PE. The AAPE based on PE addresses the two problems mentioned above. A flowchart of AAPE is given in Figure 1.

Supposing that the initial value of  $p(\pi_d^{m,\tau})$  is 0, for the original signal  $X$ , as  $j$  increases from 1 to  $N - m + 1$ ,  $p(\pi_d^{m,\tau})$  should be renewed whenever  $\pi_d^{m,\tau}$  appears.

$$p(\pi_d^{m,\tau}) = p(\pi_d^{m,\tau}) + \left( \frac{A}{m} \sum_{k=1}^m |x_{[j+(k-1)\tau]}| + \frac{1-A}{m-1} \sum_{k=2}^m |x_{[j+(k-1)\tau]} - x_{[j+(k-2)\tau]}| \right) \quad (4)$$

where  $A$  ( $A \in [0,1]$ ) stands for the correlation adjustment coefficient, which is related to the mean value and deviation between consecutive amplitudes.

Then, the probability of  $p(\pi_d^{m,\tau})$  of the time series is given by Equation (5).

$$p(\pi_d^{m,\tau}) = \frac{p(\pi_d^{m,\tau})}{\sum_{j=1}^{N-m+1} \left( \frac{A}{m} \sum_{k=1}^m |x_{[j+(k-1)\tau]}| + \frac{1-A}{m-1} \sum_{k=2}^m |x_{[j+(k-1)\tau]} - x_{[j+(k-2)\tau]}| \right)} \quad (5)$$

The AAPE calculation of the original signal  $X$  after normalization can be calculated by Equation (6).

$$AAPE(x, m, \tau, A) = \frac{- \sum_{\pi_d^{m,\tau}=1}^{m!} p(\pi_d^{m,\tau}) \ln p(\pi_d^{m,\tau})}{\ln(m!)} \quad (6)$$

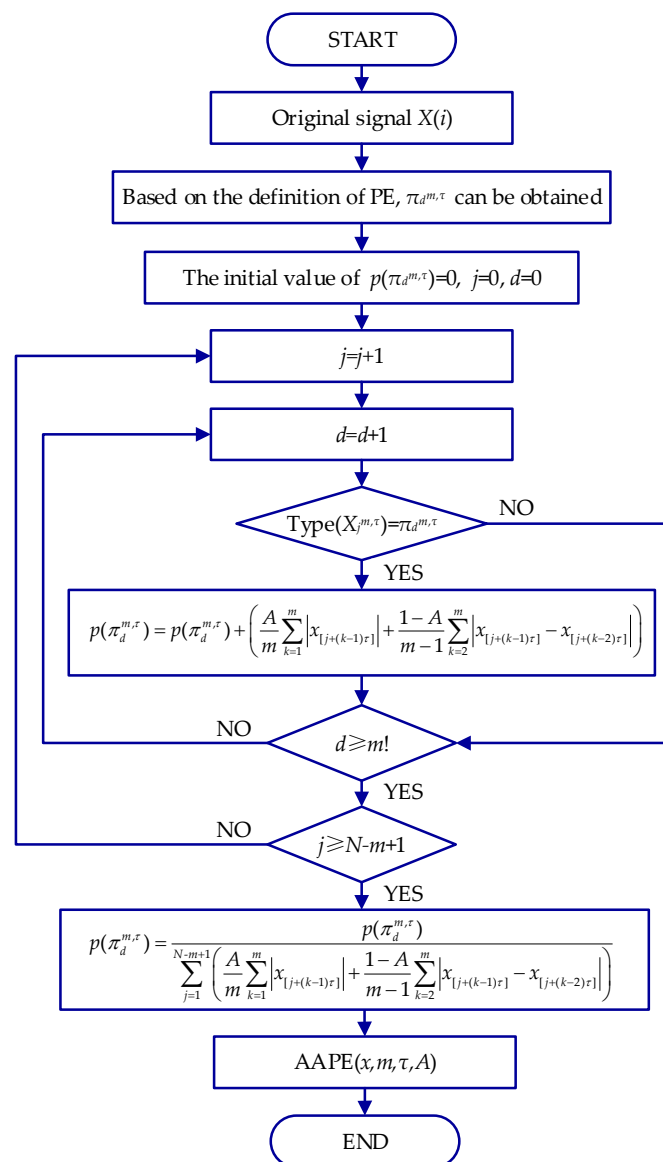


Figure 1. Flowchart of the AAPE.

2.2. The Least Common Multiple in Singular Value Decomposition (LCM-SVD)

The SVD, a matrix orthogonalization decomposition method, has been widely used in fault diagnosis analysis. A one-dimensional vibration time series needs to be reconstructed into a matrix, and the Hankel matrix is commonly applied in various formats. Given the original signal  $X(i), i = 1, 2, \dots, N$ , the Hankel matrix  $H_{m \times n}$  can be reconstructed as

$$H_{m \times n} = \begin{bmatrix} X(1) & X(2) & \dots & X(n) \\ X(2) & X(3) & \dots & X(n+1) \\ \vdots & \vdots & \ddots & \vdots \\ X(m) & X(m+1) & \dots & X(N) \end{bmatrix} \tag{7}$$

where  $1 < n < N, m = N - n + 1$ .

The SVD of  $H_{m \times n}$  can be defined as follows:

$$H_{m \times n} = U \Sigma V^T = [u_1, u_2, \dots, u_q] \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_q \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_q \end{bmatrix} \quad (8)$$

where the left singular vectors  $U_{m \times m}$  and right singular vectors  $V_{n \times n}$  are orthogonal matrices, and  $\Sigma_{m \times n}$  is a diagonal matrix composed of singular values ( $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_q \geq 0, q = \min(m, n)$ ).

For a noisy signal, the matrix dimension affects the denoising accuracy of SVD in fault identification. Under the matrix dimension determined by the least common multiple in SVD (LCM-SVD) method, the waveform error of the denoising results is much smaller than that of the traditional maximum dimension method [16]. Thus, we can obtain a well denoised signal via the LCM-SVD method. A flowchart of the LCM-SVD analysis process is shown in Figure 2, and the three steps are presented below.

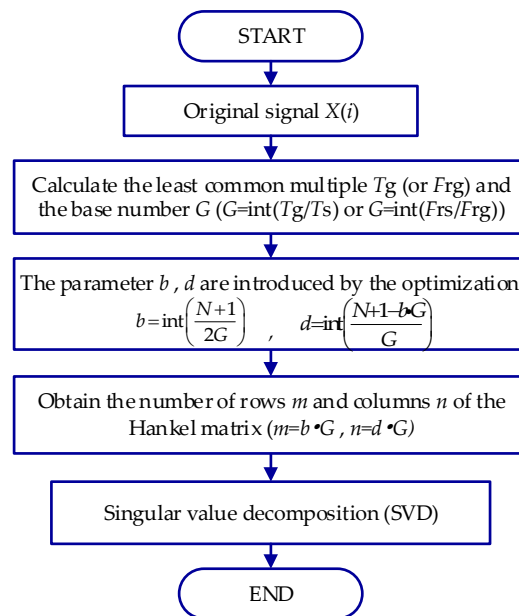


Figure 2. Flowchart of LCM-SVD.

Firstly, the least common multiple  $T_g$  is calculated by the periods of all frequency components. In the envelope spectrum analysis of the rotor system, each frequency of higher harmonics has a prominent frequency and is an integer multiple of the rotational frequency. Therefore, the least common multiple  $F_{rg}$  of each frequency component in the signal is the period of rotational frequency.

Then, a base number  $G$  is the ratio of the sampling period  $T_s$  (or sampling frequency  $F_{rs}$ ) and the least common multiple  $T_g$  (or  $F_{rg}$ ). To maximize the dimension of the Hankel matrix, the parameters  $b$  and  $d$  are introduced by the optimization computation.

Finally, the suitable row  $m$  and column  $n$  of the Hankel matrix form are determined to prepare for SVD.

### 2.3. Hierarchical Amplitude-Aware Permutation Entropy (HAAPE)

HAAPE is derived from the hierarchical analysis method and the AAPE algorithm, and its flowchart is shown in Figure 3. The detailed calculation procedure of HAAPE can be described as follows:

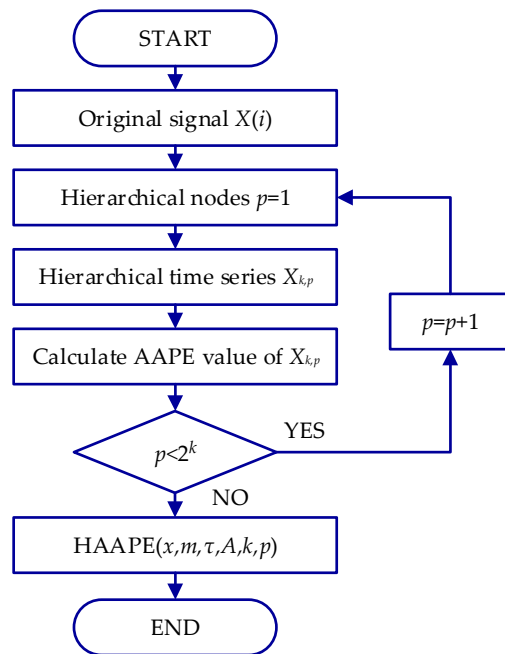


Figure 3. Flowchart of the HAAPE method.

(1) For a one-dimensional time series  $X(i), i = 1, 2, \dots, N$ , two operators  $Q_0$  and  $Q_1$  are defined as

$$\begin{cases} Q_0(x) = \frac{x(2j)+x(2j+1)}{2} \\ Q_1(x) = \frac{x(2j)-x(2j+1)}{2} \end{cases}, j = 1, 2, \dots, 2^{n-1} \quad (9)$$

where  $N = 2^n$ , and  $n$  is a positive integer, which represents the  $n$ -th level of decomposition. The averaging operator  $Q_0$  and difference operator  $Q_1$  describe the low- and high-frequency components of the signal, respectively. Accordingly, the original time series can be rebuilt by  $Q_0$  and  $Q_1$ .

$$X = \left\{ \left( Q_0(x)_j + Q_1(x)_j \right), \left( Q_0(x)_j - Q_1(x)_j \right) \right\}, j = 1, 2, \dots, 2^{n-1} \quad (10)$$

(2) The matrix  $Q_j$  operator is defined when  $j = 0$  or  $j = 1$  as follows:

$$Q_j(x) = \begin{bmatrix} \frac{1}{2} & \frac{(-1)^j}{2} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{(-1)^j}{2} & \dots & 0 & 0 \\ & & & & \ddots & & \\ 0 & 0 & 0 & 0 & \dots & \frac{1}{2} & \frac{(-1)^j}{2} \end{bmatrix}_{2^{n-1} \times 2^n} \quad (11)$$

(3) To conduct a hierarchical analysis of a time series, the operators  $Q_0$  and  $Q_1$  are applied repeatedly in step 2. To describe each node of the  $k$ -th layer in the hierarchical analysis, we introduce the positive integer parameter  $p$  and a vector  $[\zeta_1, \zeta_2, \dots, \zeta_k]$ . The relationship between  $p$  and the vector is calculated as follows:

$$p = \sum_{i=1}^k 2^{k-i} \zeta_i \quad (12)$$

That is, when  $k \in \mathbb{Z}_N^+$  is fixed,  $p$  corresponds to the sole vector  $[\zeta_1, \zeta_2, \dots, \zeta_k] \in \{0, 1\}$  by Equation (12). For example, when  $k = 3$  and  $p = 7$ , the unique vector  $[\zeta_1, \zeta_2, \zeta_3] = [1, 1, 1]$ .

(4) The sub-signal with node  $p$  component of the  $k$ -th layer decomposition in time series  $X_{k,p}$  can be obtained by Equation (13). Therefore, the high-frequency components  $X_{3,7}$  can be denoted as  $Q_1^3 \cdot Q_1^2 \cdot Q_1^1 \cdot X(i)$ .

$$X_{k,p} = Q_{\zeta_k}^k \cdot Q_{\zeta_{k-1}}^{k-1} \cdot \dots \cdot Q_{\zeta_1}^1 \cdot X(i) \tag{13}$$

(5) Finally, under the hierarchical layer  $k$ , the  $HAAPE$  value of the original signal  $X(i)$  can be expressed by the  $AAPE$  value of each layer node component.

$$HAAPE(x, m, \tau, A, k, p) = AAPE(x_{k,p}, m, \tau, A) \tag{14}$$

where  $m$  is the embedding dimension, and  $\tau$  is the time delay. The schematic diagram of the hierarchical analysis contains three layers, as shown in Figure 4.

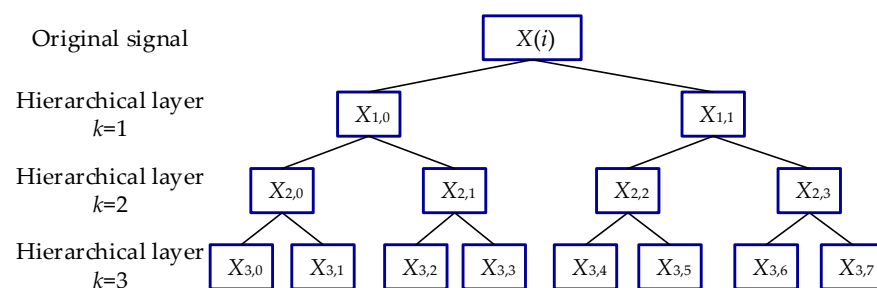


Figure 4. Hierarchical decomposition of original signal with three layers.

Five parameters need to be determined before calculating the  $HAAPE$  algorithm: signal length  $N$ , embedding dimension  $m$ , time delay  $\tau$ , correlation adjustment coefficient  $A$ , and hierarchical decomposition layer  $k$ . Specifically, the embedding dimension  $m$  is associated with the signal length  $N$  ( $N = 2048$ ). If the value of  $m$  is set too small, the phase space reconstructed vector will lose its usefulness and effectiveness. Conversely, if  $m$  is set too large, the phase space reconstructed vector will fail to reveal the change between time series distinctly. Generally, the embedding dimension  $m$  is 3–7 [34]. Time delay  $\tau$  is normally set to 1. Correlation adjustment coefficient  $A$  is usually set to 0.5 since the significance of the average amplitude is equal to the difference of the amplitude values. However,  $A \ll 0.5$  is recommended when the difference between two consecutive time series is more important than the average amplitude [22]. For the selection of the hierarchical decomposition layer  $k$ , the length of the time series is shortened with an increase in the  $k$  value. However, if the hierarchical decomposition layer  $k$  is smaller, some critical low- and high-frequency characteristics cannot be obtained, so the layer number  $k = 3$  is usually selected [35].

### 3. Numerical Simulation Analysis

In this section,  $HAAPE$ ,  $IMAAPE$ ,  $MAAPE$ , and  $HPE$  are selected for performance comparison. The 50 sets' independent random WGN and  $1/f$  noise containing 2048 data points are adopted to observe four entropy trends. Figure 5a,b show the randomly selected temporal waveforms of WGN and  $1/f$  noise, respectively. It can also be seen in Figure 5a that WGN is a random distribution and has a stable complexity in the period. In contrast,  $1/f$  noise in Figure 5b has a more complex structure and contains more mode information. Consequently,  $1/f$  noise is more complex than WGN regarding structure, whereas WGN is greater than  $1/f$  noise in terms of the quality of irregularities. Theoretically, the entropy value of WGN will be stable and higher than that of  $1/f$  noise.

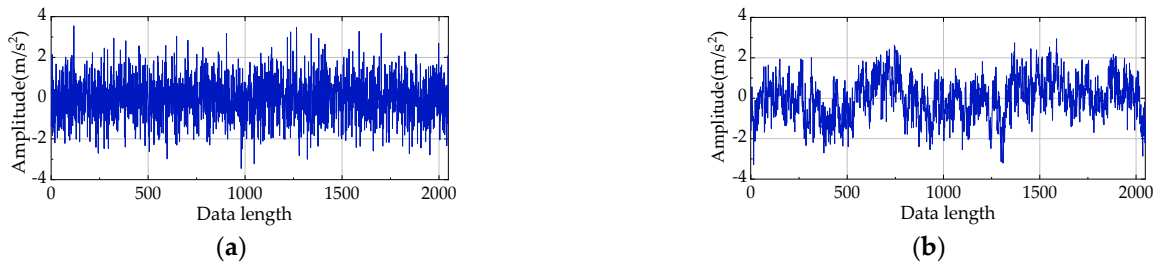


Figure 5. (a) Temporal waveform of WGN, and (b) temporal waveform of 1/f noise.

To ensure comparability and effectiveness, the parameters of the four various entropy algorithms are defined as the following: embedding dimension  $m = 3$ , time delay  $\tau = 1$ , adjustment coefficient  $A = 0.5$ , scale factor  $t = 8$ , and layer number  $k = 3$  (contains eight nodes). Then, the four entropies of WGN and 1/f noise are normalized, and the averages of 50 sets of values are displayed in Figure 6. In Figure 7, to eliminate the influence of the measurement scale, the coefficient of variation (CV) is used to judge the distribution degree of the four entropies, where the CV value is equal to the standard deviation divided by the mean value. The simulation results are defined as follows:

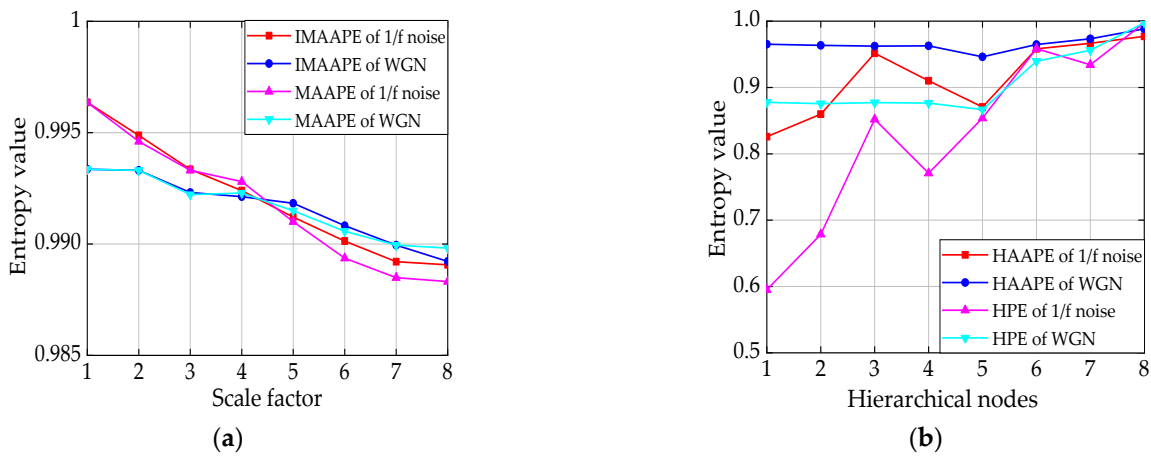


Figure 6. (a) Average of 50 sets of values of IMAAPE and MAAPE, and (b) average of 50 sets of values of HAAPE and HPE.

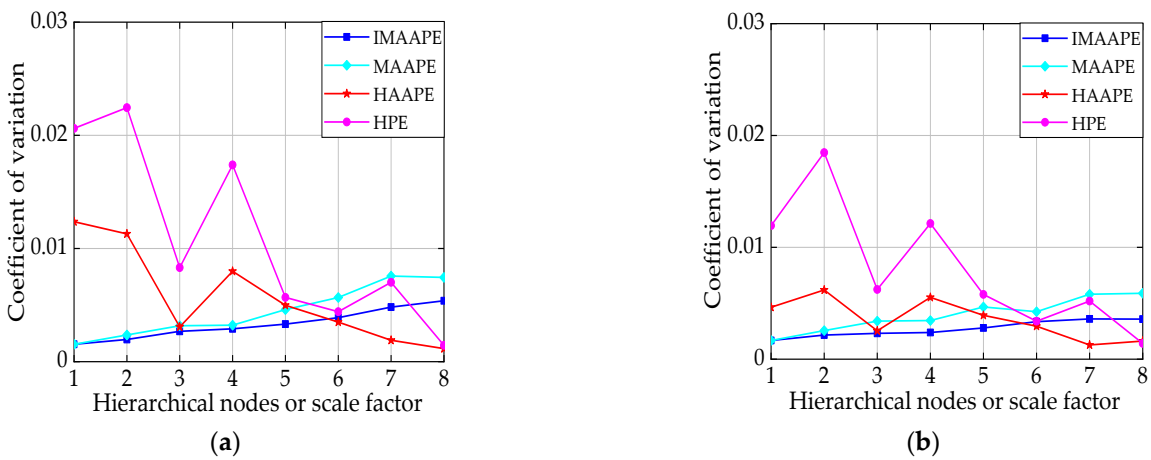


Figure 7. (a) CV values of four entropies for 1/f noise, and (b) CV values of four entropies for WGN.

(1) In Figure 6, the HAAPE and HPE values of WGN are greater than those of 1/f noise, which is in accordance with the theoretical analysis. The IMAAPE and MAAPE values of WGN are greater than those of 1/f noise in the last four scales (scale = 5,6,7,8).



(2) The *HAAPE* and *HPE* value curves can separate *WGN* noise and  $1/f$  noise signals well, while both the *IMAAPE* and *MAAPE* value curves overlap on most scales and have poor separability in Figure 6.

(3) As can be seen in Figure 7, whether *WGN* or  $1/f$  noise, the *CV* value of *IMAAPE* is lower than that of *MAAPE* on all scales. This indicates that *IMAAPE* is more stable than *MAAPE*. Meanwhile, the *CV* value of *HAAPE* is smaller than that of *HPE* in whole scales, which illustrates that the *AAPE* algorithm is superior to feature extraction.

(4) In Figure 7, as the scale factor (or hierarchical nodes) increases, the *CV* value of *IMAAPE* steadily increases, while *HAAPE* gradually decreases, indicating that the stability is significantly improved. Hierarchical analysis has an advantage over the traditional multi-scale analysis because the hierarchical analysis can simultaneously utilize the signal’s low- and high-frequency components.

Accordingly, compared with the other three methods (*IMAAPE*, *MAAPE*, and *HPE*), *HAAPE* can extract entire information and detect the dynamic trend of complex time series.

#### 4. Experimental Data Analysis

##### 4.1. Experimental Setup

The data sets in this section were published by the Center on Intelligent Maintenance Systems (IMS), the University of Cincinnati [36]. The data sets are used to conduct two cases, one for trend analysis of rolling bearing performance degradation to explicate the effectiveness of the *HAAPE* algorithm and the other one for a feature selection strategy based on *HAAPE* after *LCM-SVD* identifies valid fault characteristics.

The experimental bench’s schematic diagram and physical diagram are shown in Figure 8; four ZA-2115 type double-row roller bearings are installed on the single shaft. The radial load of the middle bearings, namely, bearings 2 and 3, is 6000 lbs, and bearings 1 and 4 play the supporting role. The shaft is driven by a motor, and its rotating speed is stable at 2000 r/min. During operation, the vibration data are collected by the PCB 353B33 high-precision ICP vibration sensor at a sampling frequency of 20 kHz (each sampling time is 1 s) every 10 min, collecting 20,480 data points. The experiment obtains three experimental data sets, ranging from the installation of the new bearing to complete failure. We use the second set for experimental analysis. The second set has 984 data sets, and each data set contains the vibration signals of the bearings in one direction. The experiment runs for 163.83 h, and the results show that the outer ring of bearing 1 fails as a result of disassembling and inspecting bearing 1 [37]. Table 1 presents the structural characteristics of ZA-2115 and its ball pass frequency outer (BPFO).

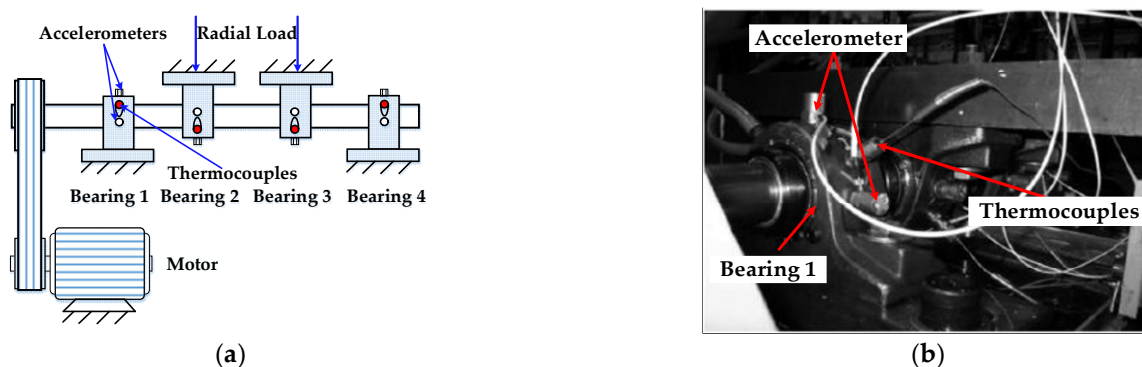


Figure 8. (a) The experimental bench’s schematic diagram and (b) its physical diagram.

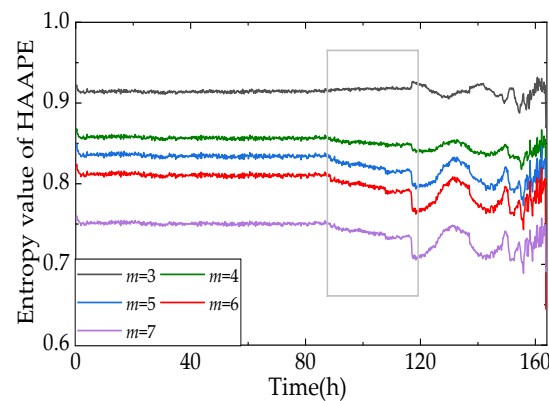
Table 1. The structural characteristic of ZA-2115 and its ball pass frequency outer ring.

Contact Angle	Pitch Diameter	Roller Diameter	Roller Number	BPFO
15.17°	71.5 mm	8.4 mm	16	236.4 Hz

#### 4.2. Case 1: Rolling Bearing Performance Degradation Trend Analysis

The trend analysis of rolling bearing performance degradation investigates the correlation between the healthy operational state and the whole life process of test signals. The RMS, kurtosis, HAAPE, IMAAPE, HPE, and PE methods are selected for comparison. The common analysis indexes used to evaluate the whole life cycle are the RMS and kurtosis values. The experimental data are divided into ten non-overlapping groups, and the average entropy value of the ten groups is used as the output result.

The amplitude value between two consecutive sampling points is more valuable than the average amplitude when employing performance degradation trend analysis, so the correlation adjustment coefficient  $A$  is set to 0.02 in this section. The scale factor should be as small as possible to avoid losing features in the multi-scale analysis [38]. Thus, the hierarchical decomposition layer  $k$  is fixed to 2, and scale factor  $t = 4$ . Figure 9 depicts the effects of the embedding dimension on HAAPE. With an increase in the embedding dimension  $m$ , the performance degradation trend notably grows in the gray rectangle. Nevertheless, there is no significant difference in effect between  $m = 6$  and  $m = 7$ . Meanwhile, in the computation process, we find that  $m = 6$  takes less time to calculate than  $m = 7$ , so  $m = 6$  is best.



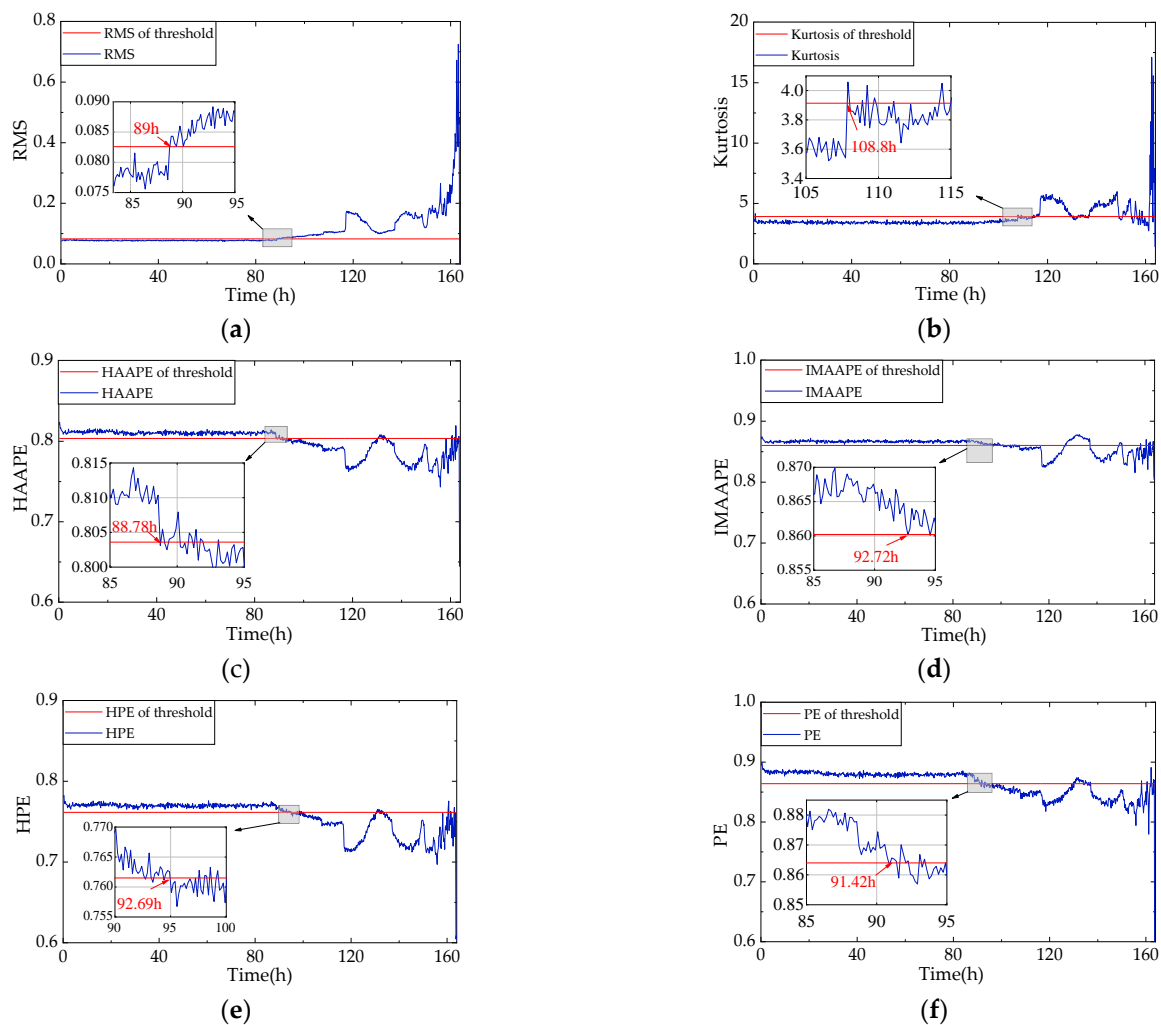
**Figure 9.** The influences of embedding dimensions on HAAPE.

The whole life of a rolling bearing can be divided into two conditions: a healthy state and an abnormal one. In practical engineering applications, we monitor the two conditions to obtain the degradation trend of the rolling bearing. Chebyshev’s inequality theory is introduced to determine the effective health threshold, which is defined as follows:

$$P\{|X - \mu_h| \geq \varepsilon_h\} \leq \frac{\sigma_h^2}{\varepsilon_h^2} \text{ or } P\{|X - \mu_h| < \varepsilon_h\} \geq 1 - \frac{\sigma_h^2}{\varepsilon_h^2} \quad (15)$$

where  $X$  stands for the series of HAAPE values under healthy conditions, and  $\mu_h$  and  $\sigma_h$  represent the mean value and standard deviation of  $X$ , respectively. According to Equation (15), when  $\varepsilon_h$  is set to  $5\sigma_h$ , the entropy value of a bearing is about  $\mu_h \pm 5\sigma_h$  in the healthy condition. In general, the test bearing remains healthy during the first one-quarter of the time period [39,40]. Therefore, we apply the first 50 h of indicator data to calculate the health threshold through Equation (15) to minimize errors. The HAAPE value of the rolling bearing data series exceeds the health threshold, which confirms that the rolling bearing is under abnormal conditions with 96% confidence probability.

RMS and the kurtosis value are sensitive to bearing degradation [39], so RMS and kurtosis are shown as evaluation indexes in Figure 10a,b, respectively. It can be seen in Figure 10a that when early failure occurs, the RMS values surpass the health threshold at nearly 89 h, and the change range is relatively small. Meanwhile, the kurtosis values indicate that the health threshold and curve intersect occur approximately 20 h later than those of the RMS index at 108.8 h in Figure 10b. This also reveals that kurtosis is insensitive to early bearing performance degradation.



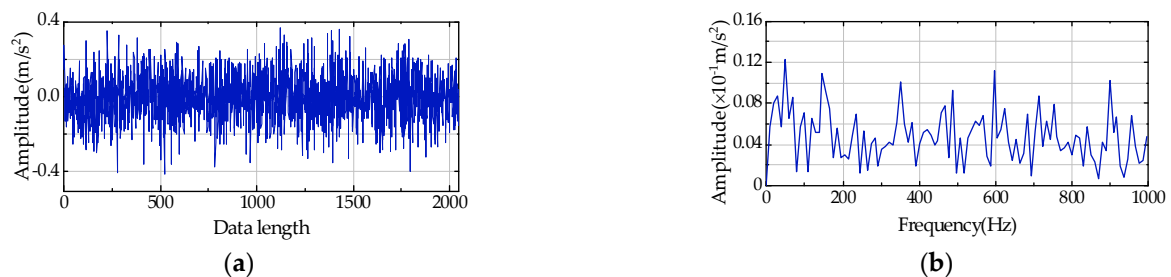
**Figure 10.** Six indicators with performance degradation trend analysis over whole lifetime. (a) RMS; (b) Kurtosis; (c) *HAAPE*; (d) *IMAAPE*; (e) *HPE* and (f) *PE*.

To investigate the trend analysis of the *HAAPE* algorithm, the algorithms *IMAAPE*, *HPE*, and *PE* are utilized for the comparison experiment. As shown in Figure 10c, the *HAAPE* values exceed the health threshold at 88.78 h and display a sharp drop, indicating a sudden change in the bearing running state. This demonstrates that the *HAAPE* algorithm successfully extracts the fault features of the rolling bearings and detects the early performance decline trend, which occurs earlier than the *RMS* index at about 13 min. In addition, the value of *HAAPE* maintains a stable state in the healthy stage; the entropy retains small values; and the curve varies when a bearing fault occurs, which means that the vibration signals of the healthy state are more complex and random than that of the abnormal state. Therefore, the *HAAPE* algorithm is sensitive to early bearing fault detection, and it draws the degradation trend of bearings effectively and accurately. In Figure 10d,e, both *IMAAPE* and *HPE* values indicate the occurrence of degradation at 93 h, about 3 h later than *HAAPE*. In addition, the range of entropy change is relatively smaller when detecting early failure, which shows that the above two algorithms have poor sensitivity to early failure. Figure 10f shows that the intersection of the health threshold and the *PE* curve is at 91 h, which is earlier than *IMAAPE* and *HPE* but later than *HAAPE*.

In summary, *HAAPE* can extract fault characteristics of rolling bearings and assess performance degradation process, which has a profound significance for the life prediction of rolling bearings.

#### 4.3. Case 2: The Feature Selection Strategy Based on HAAPE after LCM-SVD

In practical conditions, the operating environment of the bearing is often buried in noise. To verify the effectiveness of the proposed method, the 533rd, 534th, and 700th data groups are taken as analysis objects in this case and random Gaussian noise with SNR = 1 dB is added to the original vibrational signals. The 533rd and 534th data groups are the bearing vibration data when an early fault occurs, while the 700th data group is collected when bearing failure aggravates. Figure 11 shows the time-domain waveform and envelope spectrum of the 533rd data group. Figure 11a exhibits a random group signal time-domain waveform with a data length of 2048. Figure 11b displays the confusing and complex Hilbert envelope spectrum, and the fault frequency cannot be extracted because of noise in the time series. Consequently, this case requires envelope spectrum analysis to be carried out after LCM-SVD denoising.

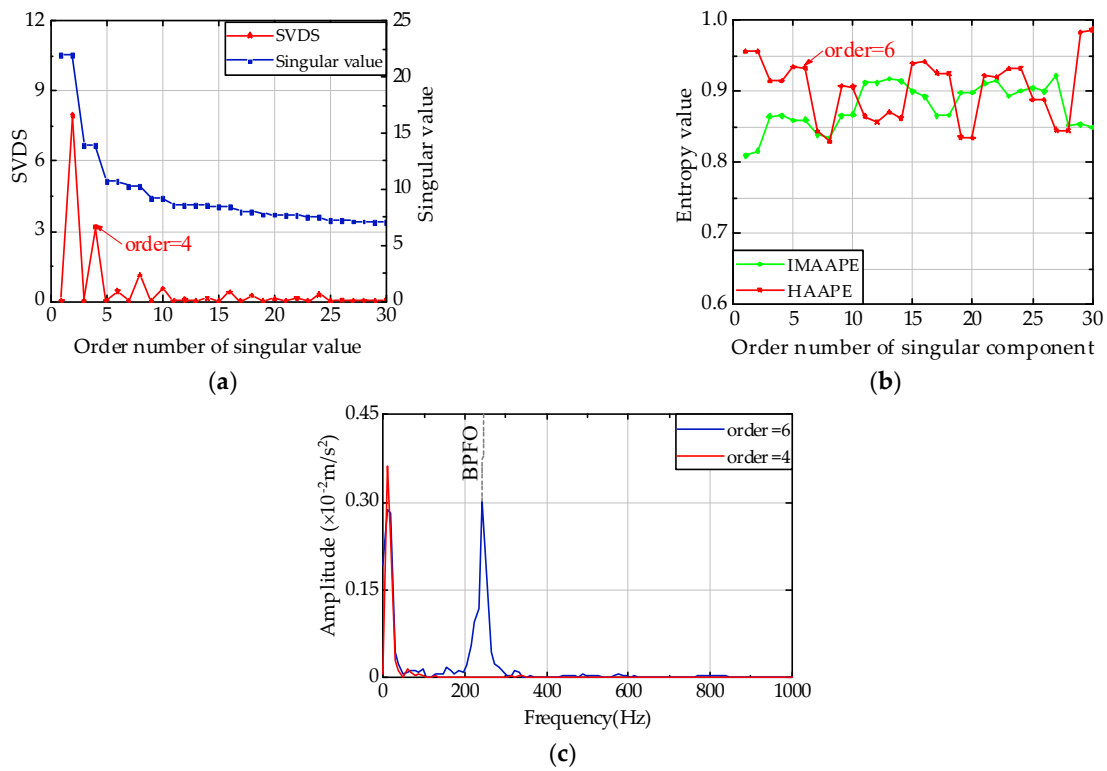


**Figure 11.** (a) Time-domain waveform, and (b) Hilbert envelope spectrum of the 533rd data group with Gaussian noise.

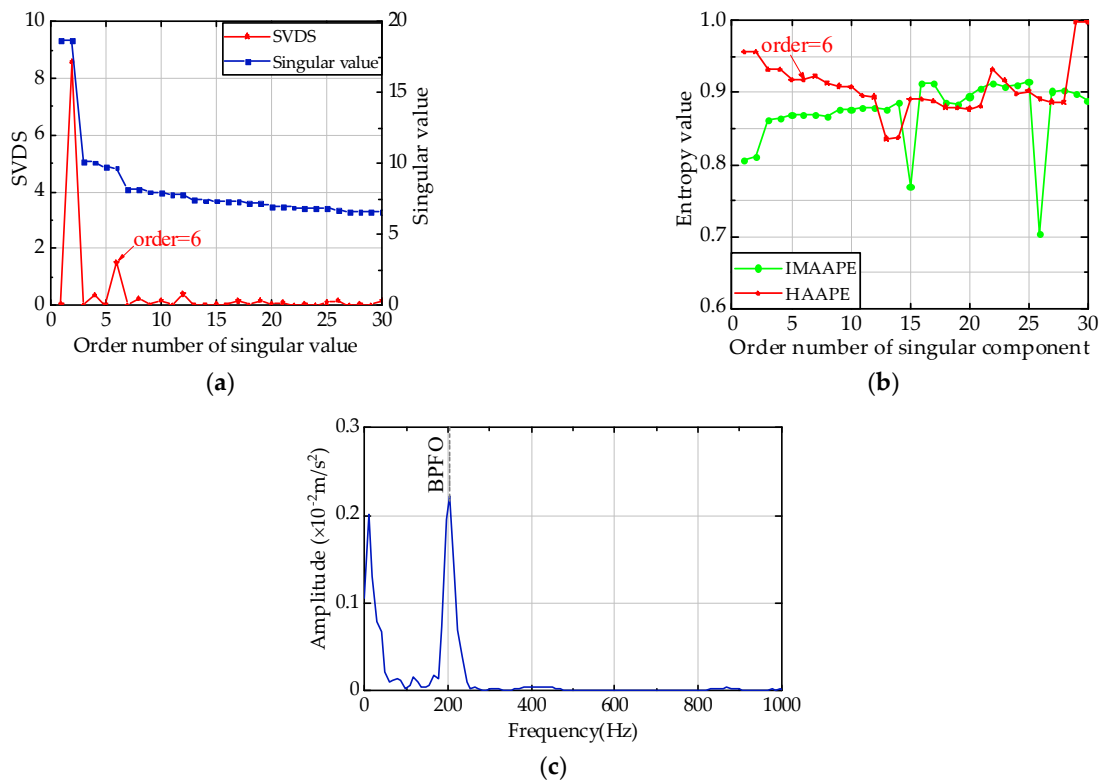
The difference spectrum unilateral maxima of singular value (DSUMSV) determines the effective rank order number by, from right to left, selecting the coordinate value corresponding to the highest difference of the peak compared with its neighbor one. This case adopts the DSUMSV to describe HAAPE's accuracy of identification. To validate the superiority of the HAAPE algorithm, the IMAAPE algorithm is compared in this case to a related analysis, and the selection of parameters ( $N$ ,  $m$ ,  $\tau$ ,  $A$ ,  $k$ ) refers to those presented in Section 2.3.

As shown in Figure 12a, there are three peaks in total in the singular value difference spectrum (SVDS), and the second maximum peak is in the fourth coordinate using the DSUMSV principle. Furthermore, its value is the highest difference from the maximum peak, which means that the maximum sudden change in the singular value happens in the fourth coordinate. This change shows the boundary of singular values between the signal and noise, so the effective rank order is four. In Figure 12b, the HAAPE values of the first 30 singular components are calculated, and every two values in the first six entropies nearly equal each other. Therefore, the effective rank order is set to six. The values of IMAAPE have no rules at all in Figure 12b. In Figure 12c, the rebuilt signal is analyzed by envelope analysis, where the Fourier transform of the envelope signal is multiplied by the Hann window function. Less useful frequency information can be extracted when the rank order is equal to four, while a frequency position stands out on the curve at 244.1 Hz when the order equals six, which is approximately equal to the fault frequency of the outer ring.

In Figure 13a,b, the DSUMSV method and HAAPE algorithm results show that the first six coordinates are identified as the effective rank order number. Meanwhile, with difficulty, we can observe the regularity of the IMAAPE values change in Figure 13b. As shown in Figure 13c, when the rank order is equal to six, the envelope analysis curve has a prominent peak at 205.1 Hz, which further suggests the possible potential damage to the outer ring. Therefore, the comparative analysis reveals that HAAPE is better than IMAAPE in characteristic frequency recognition.



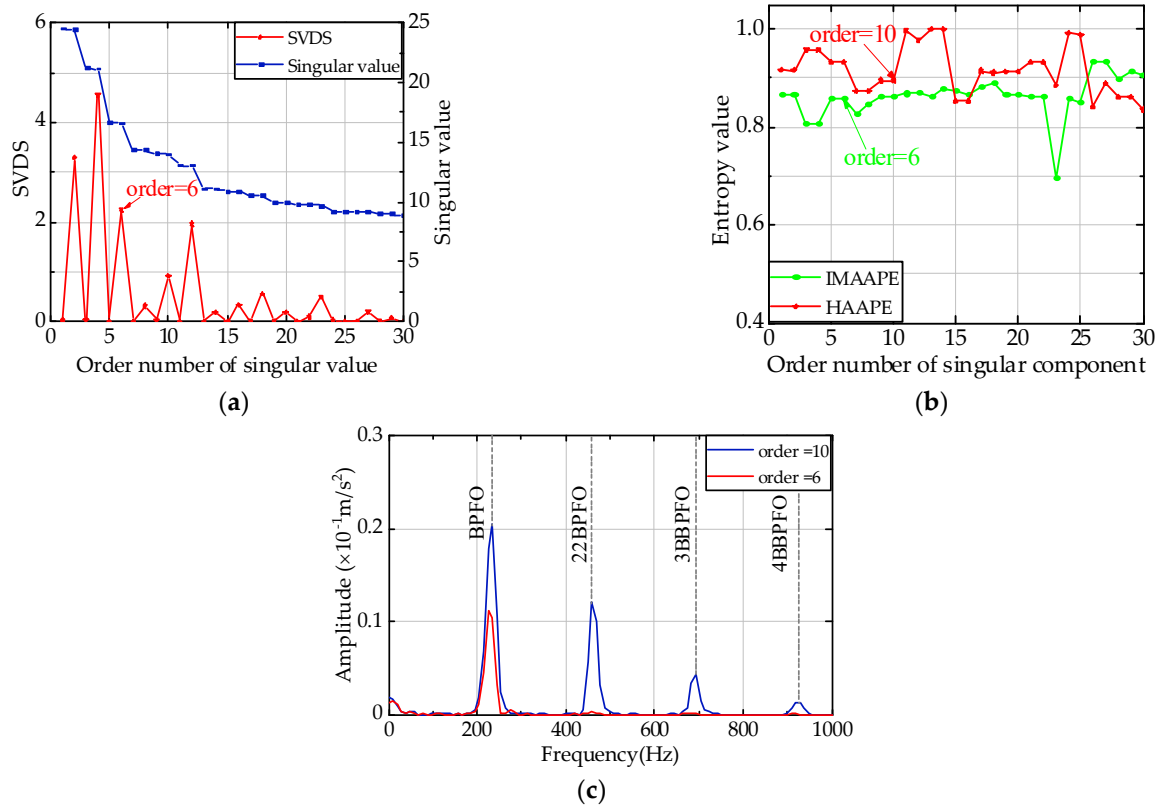
**Figure 12.** Analysis diagram of 533rd data group with Gaussian noise. (a) DSUMSV, (b) HAAPE and IMAAPE, and (c) Hilbert envelope spectrum.



**Figure 13.** Analysis diagram of 534th data group with Gaussian noise. (a) DSUMSV, (b) HAAPE and IMAAPE, and (c) Hilbert envelope spectrum.

Figure 10 shows the aggravation of the bearing failure, especially in the 700th data group. Figure 14a indicates that the second maximum peak is the highest difference from

the maximum peak, while the second maximum peak is located in the sixth coordinate. Hence, the effective rank order is set to six. According to Figure 14b, the *HAAPE* values of the first ten coordinates show a specific tendency, while the *IMAAPE* algorithm presents regularity in the first six coordinates. Consequently, the effective rank order is set to 10. In Figure 14c, when the effective rank order is six, only the fault frequency of the outer ring is found, and there is no more multiplier information about the fault frequency. Under the effective rank order of 10, the envelope curve has 3 prominent peaks at the fault frequency of 234.4 Hz, dual-frequency of 459 Hz, and tri-frequency of 693.4 Hz. Based on the above analysis, the failure of the bearing outer ring can be confirmed.



**Figure 14.** Analysis diagram of 700th data group with Gaussian noise. (a) DSUMSV, (b) *HAAPE* and *IMAAPE*, and (c) Hilbert envelope spectrum.

### 5. Conclusions

To evaluate the performance degradation trend of rolling bearings and to improve the accuracy of extracting useful fault features, a nonlinear dynamic signal analysis method called *HAAPE* based on *AAPE* is proposed to weigh the complexity of the time series composed of the fault vibration signal. Numerical simulation signals and experimental data verify the effectiveness of *HAAPE*. The following several conclusions can be drawn:

(1) Compared with the existing algorithms *IMAAPE*, *MAAPE*, and *HPE*, *HAAPE* has better stability and accuracy. It solves the issues arising from multi-scale entropy, which only considers the low-frequency components of the signal and ignores some information in the high-frequency components. Accordingly, *HAAPE* can extract more useful fault characteristics at both low- and high-frequencies from the vibration signal of rolling bearings.

(2) The results of experimental case 1 indicate that the proposed *HAAPE* can efficiently reflect the bearing performance degradation trend. Additionally, this approach is more sensitive to the early identification of failures compared to the other five methods. Hence, the *HAAPE* algorithm is an effective index for assessment performance degradation.

(3) Case 2 verifies that the *HAAPE* algorithm accurately decides the effective rank order. Especially after LCM-SVD denoising, *HAAPE* can perfectly identify the fault characteristics of the bearing using a few feature vectors, which reduces information redundancy and calculative burden.

**Author Contributions:** Conceptualization, Z.L. and L.L.; methodology, Z.L.; software, Z.L.; validation, R.C., L.L. and J.D.; formal analysis, L.D. and Z.L.; resources, Y.C.; data curation, Z.L.; writing—original draft preparation, Z.L.; writing—review and editing, L.L. and J.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research is supported by the National Key R&D Program of China (Grant No. 2018YFB2000505), the National Natural Science Foundation of China (Grant No. 51175419) and the University-level Doctoral Start-up Fund of XUT (No. 102-451120017).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The experimental data is available from the IMS Bearing Data deposited in <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/> (accessed on 2 January 2022).

**Acknowledgments:** The authors would like to thank the anonymous reviewers and the editor for their valuable and insightful suggestions.

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

## References

1. Wang, D.; Tsui, K.-L.; Miao, Q. Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators. *IEEE Access* **2018**, *6*, 665–676. [[CrossRef](#)]
2. Wan, L.; Gong, K.; Zhang, G.; Yuan, X.; Li, C.; Deng, X. An Efficient Rolling Bearing Fault Diagnosis Method Based on Spark and Improved Random Forest Algorithm. *IEEE Access* **2021**, *9*, 37866–37882. [[CrossRef](#)]
3. Wei, Y.; Li, Y.; Xu, M.; Huang, W. A Review of Early Fault Diagnosis Approaches and Their Applications in Rotating Machinery. *Entropy* **2019**, *21*, 409. [[CrossRef](#)] [[PubMed](#)]
4. Yan, X.; Xu, Y.; She, D.; Zhang, W. A Bearing Fault Diagnosis Method Based on PAVME and MEDE. *Entropy* **2021**, *23*, 1402. [[CrossRef](#)] [[PubMed](#)]
5. Chen, Y.; Zhang, T.; Luo, Z.; Sun, K. A Novel Rolling Bearing Fault Diagnosis and Severity Analysis Method. *Appl. Sci.* **2019**, *9*, 2356. [[CrossRef](#)]
6. Zhao, B.; Li, Q.; Lv, Q.; Si, X. A Spectrum Adaptive Segmentation Empirical Wavelet Transform for Noisy and Nonstationary Signal Processing. *IEEE Access* **2021**, *9*, 106375–106386. [[CrossRef](#)]
7. Wang, J.; Wei, Q.; Zhao, L.; Yu, T.; Han, R. An improved empirical mode decomposition method using second generation wavelets interpolation. *Digit. Signal Processing* **2018**, *79*, 164–174. [[CrossRef](#)]
8. Barbosh, M.; Singh, P.; Sadhu, A. Empirical mode decomposition and its variants: A review with applications in structural health monitoring. *Smart Mater. Struct.* **2020**, *29*, 093001. [[CrossRef](#)]
9. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Processing* **2014**, *62*, 531–544. [[CrossRef](#)]
10. Cheng, J.; Yang, Y.; Li, X.; Cheng, J. Adaptive periodic mode decomposition and its application in rolling bearing fault diagnosis. *Mech. Syst. Signal Processing* **2021**, *161*, 107943. [[CrossRef](#)]
11. Gong, T.; Yuan, X.; Yuan, Y.; Lei, X.; Wang, X. Application of tentative variational mode decomposition in fault feature detection of rolling element bearing. *Measurement* **2019**, *135*, 481–492. [[CrossRef](#)]
12. Lian, J.; Liu, Z.; Wang, H.; Dong, X. Adaptive variational mode decomposition method for signal processing based on mode characteristic. *Mech. Syst. Signal Processing* **2018**, *107*, 53–77. [[CrossRef](#)]
13. Fang, L.; Sun, H. Study on EEMD-Based KICA and Its Application in Fault-Feature Extraction of Rotating Machinery. *Appl. Sci.* **2018**, *8*, 1441. [[CrossRef](#)]
14. Li, L.; Cui, Y.; Chen, R.; Liu, X. Optimal SES Selection Based on SVD and Its Application to Incipient Bearing Fault Diagnosis. *Shock. Vib.* **2018**, *2018*, 8067416. [[CrossRef](#)]
15. Li, L.; Cui, Y.; Chen, R.; Chen, L.; Wang, L. Research on Rolling Bearing Fault Feature Extraction Based on Singular Value Decomposition considering the Singular Component Accumulative Effect and Teager Energy Operator. *Shock. Vib.* **2019**, *2019*, 3742512. [[CrossRef](#)]
16. Tongjian, C.; Bangyan, Y.E.; Qipeng, S.; Xuezhi, Z. Matrix Dimension Considering Frequency Factor in Singular Value Decomposition. *J. Mech. Eng.* **2019**, *55*, 1–14.

17. Voss, A.; Schulz, S.; Schroeder, R.; Baumert, M.; Caminal, P. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. A Math. Phys. Eng. Sci.* **2009**, *367*, 277–296. [[CrossRef](#)]
18. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)]
19. Sharma, S.; Tiwari, S.K.; Sukhjeet, S. Integrated approach based on flexible analytical wavelet transform and permutation entropy for fault detection in rotary machines. *Measurement* **2021**, *169*, 108389. [[CrossRef](#)]
20. Landauskas, M.; Cao, M.; Ragulskis, M. Permutation entropy-based 2D feature extraction for bearing fault diagnosis. *Nonlinear Dyn.* **2020**, *102*, 1717–1731. [[CrossRef](#)]
21. Vashishtha, G.; Chauhan, S.; Singh, M.; Kumar, R. Bearing defect identification by swarm decomposition considering permutation entropy measure and opposition-based slime mould algorithm. *Measurement* **2021**, *178*, 109389. [[CrossRef](#)]
22. Azami, H.; Escudero, J. Amplitude-aware permutation entropy: Illustration in spike detection and signal segmentation. *Comput. Methods Programs Biomed.* **2016**, *128*, 40–51. [[CrossRef](#)] [[PubMed](#)]
23. Yang, H.; Zhang, A.; Li, G. A New Singular Spectrum Decomposition Method Based on Cao Algorithm and Amplitude Aware Permutation Entropy. *IEEE Access* **2021**, *9*, 44534–44557. [[CrossRef](#)]
24. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2005**, *71*, 021906. [[CrossRef](#)] [[PubMed](#)]
25. Zhou, R.; Wang, X.; Wan, J.; Xiong, N. EDM-Fuzzy: An Euclidean Distance Based Multiscale Fuzzy Entropy Technology for Diagnosing Faults of Industrial Systems. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4046–4054. [[CrossRef](#)]
26. Li, H.; Huang, J.; Yang, X.; Luo, J.; Zhang, L.; Pang, Y. Fault Diagnosis for Rotating Machinery Using Multiscale Permutation Entropy and Convolutional Neural Networks. *Entropy* **2020**, *22*, 851. [[CrossRef](#)]
27. Chen, Y.; Zhang, T.; Zhao, W.; Luo, Z.; Sun, K. Fault Diagnosis of Rolling Bearing Using Multiscale Amplitude-Aware Permutation Entropy and Random Forest. *Algorithms* **2019**, *12*, 184. [[CrossRef](#)]
28. Chen, Y.; Zhang, T.; Zhao, W.; Luo, Z.; Lin, H. Rotating Machinery Fault Diagnosis Based on Improved Multiscale Amplitude-Aware Permutation Entropy and Multiclass Relevance Vector Machine. *Sensors* **2019**, *19*, 4542. [[CrossRef](#)]
29. Jiang, Y.; Peng, C.K.; Xu, Y. Hierarchical entropy analysis for biological signals. *J. Comput. Appl. Math.* **2011**, *236*, 728–742. [[CrossRef](#)]
30. Song, E.; Ke, Y.; Yao, C.; Dong, Q.; Yang, L. Fault Diagnosis Method for High-Pressure Common Rail Injector Based on IFOA-VMD and Hierarchical Dispersion Entropy. *Entropy* **2019**, *21*, 923. [[CrossRef](#)]
31. Li, W.; Shen, X.; Li, Y. A Comparative Study of Multiscale Sample Entropy and Hierarchical Entropy and Its Application in Feature Extraction for Ship-Radiated Noise. *Entropy* **2019**, *21*, 793. [[CrossRef](#)] [[PubMed](#)]
32. Shi, X.; Qiu, G.; Yin, C.; Huang, X.; Chen, K.; Cheng, Y.; Zhong, S. An Improved Bearing Fault Diagnosis Scheme Based on Hierarchical Fuzzy Entropy and Alexnet Network. *IEEE Access* **2021**, *9*, 61710–61720. [[CrossRef](#)]
33. Li, Y.; Li, G.; Yang, Y.; Liang, X.; Xu, M. A fault diagnosis scheme for planetary gearboxes using adaptive multi-scale morphology filter and modified hierarchical permutation entropy. *Mech. Syst. Signal Processing* **2018**, *105*, 319–337. [[CrossRef](#)]
34. Liu, W.; Yang, X.; Jinxing, S.; Duan, C. An Integrated Fault Identification Approach for Rolling Bearings Based on Dual-Tree Complex Wavelet Packet Transform and Generalized Composite Multiscale Amplitude-Aware Permutation Entropy. *Shock. Vib.* **2020**, *2020*, 8851310. [[CrossRef](#)]
35. Li, Y.; Xu, M.; Zhao, H.; Huang, W. Hierarchical fuzzy entropy and improved support vector machine based binary tree approach for rolling bearing fault diagnosis. *Mech. Mach. Theory* **2016**, *98*, 114–132. [[CrossRef](#)]
36. Qiu, H.; Lee, J.; Lin, J.; Yu, G. Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J. Sound Vib.* **2006**, *289*, 1066–1090. [[CrossRef](#)]
37. Hoseinzadeh, M.S.; Khadem, S.E.; Sadooghi, M.S. Quantitative diagnosis for bearing faults by improving ensemble empirical mode decomposition. *ISA Trans.* **2018**, *83*, 261–275. [[CrossRef](#)]
38. Song, Y.; Wang, W. A Fault Feature Extraction Method for Rolling Bearings Based on Refined Composite Multi-Scale Amplitude-Aware Permutation Entropy. *IEEE Access* **2021**, *9*, 71979–71993. [[CrossRef](#)]
39. Qian, Y.; Yan, R.; Hu, S. Bearing Degradation Evaluation Using Recurrence Quantification Analysis and Kalman Filter. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 2599–2610. [[CrossRef](#)]
40. Zhu, K.; Song, X. Cross-fuzzy entropy-based approach for performance degradation assessment of rolling element bearings. *Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng.* **2017**, *232*, 173–185. [[CrossRef](#)]