



Hui Li<sup>1</sup>, Fan Li<sup>1</sup>, Rong Jia<sup>1,\*</sup>, Fang Zhai<sup>2</sup>, Liang Bai<sup>3</sup> and Xingqi Luo<sup>3</sup>

- <sup>1</sup> School of Electrical Engineering, Xi'an University of Technology, Xi'an 710054, China; lihui@mail.xaut.edu.cn (H.L.); lifan\_dqgc@163.com (F.L.)
- <sup>2</sup> School of Humanities and Foreign Languages, Xi'an University of Technology, Xi'an 710054, China; xautzf@126.com
- <sup>3</sup> Institute of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an 710054, China; xaut\_bl@126.com (L.B.); luoxq@xaut.edu.cn (X.L.)
- \* Correspondence: Ja.rong@163.com

**Abstract**: Symplectic geometric mode decomposition (SGMD) is a newly proposed signal processing method. Because of its superiority, it has gained more and more attention in the field of fault diagnosis. However, the similar component reorganization problem involved in this method has not been clearly stated. Aiming at this problem, this paper proposes the SGMD-CS method based on the SGMD method and the cosine similarity (CS) and has been compared and verified on the simulation signal and the actual rolling bearing signal. In addition, in order to realize the intelligent diagnosis of the wind turbine bearing fault, the symplectic geometric entropy (SymEn) is extracted as the fault feature and input it into the AdaBoost classification model. In summary, this paper proposes a new wind turbine fault feature extraction method based on the SGMD-CS and AdaBoost framework, and the validity of the method is verified by the rolling bearing vibration data of the Electrical Engineering Laboratory of Case Western Reserve University.

**Keywords:** rolling bearings; symplectic geometric mode decomposition; cosine similarity; symplectic geometric entropy; AdaBoost

# 1. Introduction

As a clean energy source, wind power has developed rapidly in recent years because of its many advantages such as renewable energy and low pollution. With the continuous increase in the capacity of wind turbine machines around the world, the structure of these units has become increasingly complicated and coupled with their long-term use under harsh conditions, this places higher requirements on the fault diagnosis of wind turbines. Especially wind turbine bearings, as a key component for converting wind energy into electrical energy, have high failure rates and maintenance costs. In order to increase the output of wind turbines and reduce operation and maintenance costs, bearing status monitoring and fault diagnosis are essential to ensure the reliable and stable operation of wind turbines [1]. Because the fault will produce vibration, the vibration signal is collected by using the vibration sensor and analyzed to extract the hidden fault information to provide a basis for identifying the type of bearing fault. Regarding the placement of the sensors, literature [2] and [3] gave a new method to diagnose the damage of the transmission system bearings, that is, to measure vibration at the tower instead of the gearbox, which makes the measurement easier. And through the discussion of actual cases to study the reliability and effectiveness of this method for bearing fault detection.

No matter where the sensors are arranged, due to the randomness of wind energy, the load and speed of the bearing of the wind turbine will change accordingly, so the collected fault signals have the characteristics of nonlinear and nonstationary. The sampled vibration signal can be regarded as a time series, and a time series usually consists of different



Citation: Li, H.; Li, F.; Jia, R.; Zhai, F.; Bai, L.; Luo, X. Research on the Fault Feature Extraction of Rolling Bearings Based on SGMD-CS and the AdaBoost Framework. *Energies* **2021**, *14*, 1555. https://doi.org/10.3390/ en14061555

Received: 4 January 2021 Accepted: 5 March 2021 Published: 11 March 2021

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



**Copyright:** © 2021 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/). patterns. By decomposing it into a series of simple component sequences, judging and reconstructing these simple components, a large amount of information about the original state can be obtained. To obtain simple component sequences and reduce complexity, an effective signal decomposition method is needed that can separate components from the original signal effectively.

In the early days, the commonly used signal processing methods were fast Fourier transform and wavelet transform. The latter is based on wavelet basis functions, which can be regarded as adding a scale factor to the window function of the short-time Fourier transform (STFT), which makes up for the shortcomings of STFT. Therefore, it has been widely used in processing non-stationary signals. With the development of wavelet, wavelet packet, second-generation wavelet, multiwavelet, empirical mode wavelet, and flexible wavelet are proposed and applied one after another. Literature [4] compared the fast Fourier transform and wavelet transform comprehensively when monitoring the wind turbine bearing, reflecting the advantages of wavelet transform.

Since the choice of wavelet basis function in wavelet transform is based on experience and does not have adaptability, it has an important influence on the signal analysis results. Empirical mode decomposition (EMD), a new type of adaptive signal time-frequency processing method proposed by N.E. Huang and others in 1998 was based on the timescale characteristics of the data itself to decompose the complex signal into a finite number of intrinsic mode functions (IMF) [5]. Once this method was proposed, it has been widely used, but modal aliasing and end effects have appeared in the process of application. Therefore, many relevant scholars have done a lot of research here [6–8]. In addition, the local mean decomposition (LMD) proposed based on EMD has also been applied. It compensates for the shortcomings of the EMD method to a certain extent, mainly by adaptively decomposing the multi-component signal into several product functions (PF) [9]. This method is also widely used as an adaptive decomposition method [10].

The symplectic geometric spectrum method is a relatively new method of time series decomposition based on symplectic space compared to the traditional time series method based on Euclidean space. In symplectic geometry, since the symplectic matrix similar transformation is a normal transformation, it will not destroy the characteristics of the original time series, so when dealing with nonlinear problems, it has a better processing effect than singular spectral decomposition [11,12]. Regarding this method, relevant scholars have done a lot of research. Literature [13] estimated the embedding dimension of the symplectic geometric spectrum method and discussed the robustness of the method to noise, sequence length, and sampling interval in the study of time series. Literature [14] combined the symplectic geometry method and the principal component analysis method to study a chaotic time series. Literature [15] introduced the symplectic geometric spectral regression technique for the prediction of a nonlinear time series. The literature [16] and [17] decomposed the time series into the sum of some independent components based on symplectic geometry theory and derived a framework of this method. The literature [18] and [19] introduced the symplectic geometry algorithm to fault diagnosis and obtained good results. In addition, the method was introduced to modal parameter identification [20], vibration feature parameter identification for structural health monitoring [21], and the analysis of athlete's surface EMG signals [22].

In the process of decomposing the original time series using the symplectic geometry method, the initially obtained components are not completely independent, so components with the same characteristics need to be reorganized. Regarding how to reorganize, most of the literature does not give a detailed description. Literature [23] used an iterative approach, starting with the first component and looking for similar components with it. When the normalized mean square error (NMSE) of the residual signal and the original signal is less than the given threshold, the iteration stops, and the final recombination components are determined. However, there is still no clear description on how to determine similar components. In this paper, the cosine similarity (CS) is introduced to measure the similarity between components for the problem of component reorganization. As a method of

measuring similarity, cosine similarity is widely used in file comparison and text mining but rarely used in signal processing [24,25]. This paper introduces cosine similarity for the first time in the decomposition process of the symplectic geometric spectrum and constructs a cosine similarity matrix to seek components with the same characteristics and reorganize them.

The traditional fault diagnosis method based on the aforementioned signal processing technology has some shortcomings, and with the development of artificial intelligence technology, this brings new ideas to fault identification [26]. In this paper, based on the SGMD-CS method and machine learning theory, through calculating the symplectic geometric entropy as a feature vector and transporting it to the constructed AdaBoost classifier, so as to obtain intelligent fault diagnosis [27]. The main contributions and novelty of this paper are as follows:

- (1) Aiming at the problem of recombination of similar components obtained by using the SGMD algorithm, the cosine similarity is introduced into the SGMD algorithm to obtain the SGMD-CS algorithm. The effectiveness of the method is verified by simulation signals and actual rolling bearing fault signals.
- (2) Based on the SGMD-CS algorithm, SymEn is constructed as the extracted fault feature vector.
- (3) Using Adaboost algorithm to realize automatic identification of bearing failure modes.
- (4) A complete fault diagnosis flowchart of rolling bearings is given, and experimental research and comparative analysis are carried out.

The structure of this paper is as follows: In Section 2, the theoretical part of the symplectic geometry algorithm is introduced, and combining the symplectic geometry algorithm with cosine similarity is discussed. In Section 3, by verifying the simulation signal, the superiority of the SGMD-CS method in terms of signal decomposition is affirmed. In Section 4, we introduce the symplectic geometric entropy and AdaBoost algorithm based on decision trees. In Section 5, we adopt the data of Case Western Reserve University's Electrical Engineering Laboratory implemented fault diagnosis, highlighting the validity of the proposed method. Finally, in Section 6, we summarize the entire article.

### 2. Symplectic Geometry Algorithm

The core concept of the symplectic geometry algorithm uses the symplectic matrix to perform QR decomposition to find the eigenvalues of the Hamiltonian matrix. The algorithm is mainly divided into three steps: phase space reconstruction, symplectic QR decomposition, and diagonal averaging transformation. Where,  $J = J_{2n} = \begin{bmatrix} 0 & +I_n \\ -I_n & 0 \end{bmatrix}$ .

#### 2.1. Symplectic Theory

Before introducing the symplectic geometry method, some basic definitions and theorems are given.

**Definition 1.** *S* as a valid matrix; if  $JSJ^{-1} = S^{-T}$ , then *S* is a symplectic matrix.

**Definition 2.** *H* as a valid matrix; if  $JHJ^{-1} = -H^{-T}$ , then *H* is a Hamiltonian matrix.

**Theorem 1.** For any symplectic matrix  $A_{n \times n}$ , a new matrix  $M = \begin{pmatrix} A & 0 \\ 0 & -A^T \end{pmatrix}$  can be constructed, and M is also a Hamiltonian matrix.

**Theorem 2.** Assuming the Household matrix H, where  $H = H(k, w) = \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}$ ,  $P = I_n - \frac{2\overline{ww}^T}{\overline{w}^T\overline{w}}$ ,  $\overline{w} = (0, \dots, 0; w_k, \dots, w_n)^T \neq 0$ , H is a symplectic unitary matrix.

Next, the symplectic geometry method is introduced in detail.

### 4 of 19

### (1) Phase space reconstruction

For an original time series such as  $x = x_1, x_2, \dots x_n$ , *n* is the length of the signal. Based on Takens embedding theory [28], a one-dimensional signal can be reconstructed into a multidimensional signal by means of the method of delay. Thereby, a trajectory matrix *X* is constructed, which contains all dynamic information of the time series:

$$X = \begin{bmatrix} x_1 & x_{1+\tau} & \dots & x_{1+(d-1)\tau} \\ \vdots & \vdots & & \vdots \\ x_m & x_{m+\tau} & \cdots & x_{m+(d-1)\tau} \end{bmatrix}$$
(1)

where *d* is the embedding dimension and  $\tau$  is the delay time,  $n = m + (d-1)\tau$ . Obviously, choosing a different embedding dimension and delay time will result in a different trajectory matrix. Here, we refer to the method in [29] to determine the d value by calculating the power spectral density (PSD) of the time series and set  $\tau = 1$ . To be precise, the frequency of the maximum peak is estimated by the PSD method. If the normalized frequency is lower than the given threshold  $10^{-3}$ , d is set to  $\frac{n}{3}$ ; otherwise, the value is set to  $d = 1.2 \times \frac{F_s}{f_{max}}$ , and  $F_s$  is the sampling frequency.

(2) Symplectic QR decomposition

This step is the core of the symplectic geometry method [30].

To obtain the Hamiltonian matrix, the covariance matrix *A* is first obtained by performing an autocorrelation analysis of the trajectory matrix:

$$A = X^T X \tag{2}$$

Then, the Hamiltonian matrix *M* will be constructed from matrix *A*:

$$M = \left(\begin{array}{cc} A & 0\\ 0 & -A^T \end{array}\right) \tag{3}$$

It is proven that there is a Householder matrix,  $H = \begin{pmatrix} Q & 0 \\ 0 & Q \end{pmatrix}$ , where the matrix Q can be composed of a real symmetric matrix A, and the matrix H is also a symplectic geometric orthogonal matrix. Then:

$$HMH^{T} = \begin{pmatrix} Q & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & -A^{T} \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & Q \end{pmatrix}^{T}$$
$$= \begin{pmatrix} QAQ^{T} & 0 \\ 0 & -QA^{T}Q^{T} \end{pmatrix}$$
$$= \begin{pmatrix} B & 0 \\ 0 & -B^{T} \end{pmatrix}$$
(4)

where *B* is an upper Hessenberg matrix,  $(b_{ij} = 0, i > j + 1)$ ; it is easy to obtain the following:  $\lambda(A) = \lambda(B) = \lambda^2(X)$ .

The eigenvalues of B can be calculated as  $\lambda_1, \lambda_2, \dots, \lambda_d$ , according to the properties of the Hamiltonian matrix, and the eigenvalues of the matrix *X* are obtained as:

$$\sigma_i = \sqrt{\lambda_i} (i = 1, 2, \cdots, d) \tag{5}$$

where  $\lambda_1 > \lambda_2 > \cdots > \lambda_d$ ,  $\lambda_i$  in descending order, and the distribution of  $\lambda_i$  represents the spectral distribution of matrix A, where the smaller value is usually considered as the noise components.  $Q_i(i = 1, 2, \cdots, d)$  is the eigenvector corresponding to the eigenvalues of matrix A, and the transformation coefficient is calculated by the following formula:

$$S_i = Q_i^T X^T (i = 1, 2, \cdots, d) \tag{6}$$

Then:

$$Z_i = Q_i S_i (i = 1, 2, \cdots, d) \tag{7}$$

The corresponding reconstruction matrix  $X_i = Z_i^T$  can be obtained, and then trajectory matrix *X* can be expressed as:

$$X = X_1 + X_2 + \dots + X_d \tag{8}$$

## (3) Diagonal averaging transformation

The dimension of the obtained initial component matrix is  $m \times d$ . Through diagonal averaging, the reconstruction matrix  $X_i$  can be transformed into d sets of time series with length n, and the sum of the d sets of time series with length n is the original time series x. The specific implementation method is as follows:

For any element  $x_{ij}(1 \le i \le d, 1 \le j \le m)$  in the matrix  $X_i$ , let  $d^* = \min(m, d)$ ,  $d^* = \min(m, d)$ . If m < d, get  $x_{ij}^* = x_{ij}$ ; otherwise,  $x_{ij}^* = x_{ji}$  Then the elements  $y_k(k = 1, 2, \dots, n)$  in the corresponding time series  $Y_i$  are calculated as shown in Equation (9):

$$y_{k} = \begin{cases} \frac{1}{k} \sum_{p=1}^{k} x_{p,k-p+1}^{*} & 1 \le k \le d^{*} \\ \frac{1}{d^{*}} \sum_{p=1}^{d^{*}} x_{p,k-p+1}^{*} & d^{*} \le k \le m^{*} \\ \frac{1}{n-k+1} \sum_{p-k-m^{*}+1}^{n-m^{*}+1} x_{p,k-p+1}^{*} & m^{*} < k \le n \end{cases}$$
(9)

Based on Formula (9), the matrix  $X_i$  is converted into a series of  $Y_i(y_1, y_2, \dots, y_n)$ . Therefore, through diagonal averaging, we can transform the trajectory matrix X into a series of length n:

$$Y = Y_1 + Y_2 + \dots + Y_d \tag{10}$$

#### 2.2. SGMD-CS Algorithm

During the decomposition of symplectic geometric models, we have obtained  $\{Y_1, Y_2, \ldots, Y_d\}$ ; in these d sets of initial components, some components may have the same periodic component, frequency component, and characteristics. Therefore, these components are not completely independent of each other, and these initial components with the same characteristics need to be reconstructed by other methods. This paper introduces cosine similarity here during the reconstruction process. The definition is as follows:

$$Y_{ij} = \cos \theta_{ij} = \frac{Y_i \cdot Y_j}{\|Y_i\| \cdot \|Y_j\|} \quad 1 \le i \le d, 1 \le j \le d$$

$$(11)$$

Due to the interference of environmental factors, the signal often contains many noise components, so we should separate the useful components from the noise components. The specific method is as follows:

First, the noise components are separated. In the previous step, we can obtain d sets of initial components, and the new *d* sets of constructed components linebreak  $N_f = \sum_{i=1}^{f} Y_i (f = 1, 2, \dots, d)$  can be obtained by adding the first *f* sets of components. By calculating the cosine similarity  $N_{k,k+1} = \cos \theta_{k,k+1} = \frac{N_k \cdot N_{k+1}}{\|N_k\| \cdot \|N_{k+1}\|} (k = 1, 2, \dots d - 1)$  of two adjacent constructed components, when the value of the cosine similarity reaches a threshold and changes slowly, it can be approximated that the threshold corresponding to the demarcation point *k* is a turning point for reconstructing the effective components. Before this point, the component contains most of the information of the original signal, which should be retained for subsequent analysis; after this point, it can be regarded as noise. Second, the remaining effective *k* sets of components are recombined between similar components. By constructing a cosine similarity matrix, similar components are combined

to determine the final symplectic geometric components. The cosine similarity matrix constructed is as follows:

$$CSM = \begin{bmatrix} 1 & Y_{12} & \dots & Y_{1k} \\ Y_{21} & 1 & \dots & Y_{2k} \\ \vdots & \vdots & \dots & \vdots \\ Y_{k1} & Y_{k2} & \cdots & 1 \end{bmatrix}$$
(12)

Obviously, the cosine similarity matrix (CSM) is a symmetric matrix, and the diagonal elements are all the same. Through this matrix, the components with higher similarity are recombined.

In summary, the flowchart of the SGMD-CS algorithm is shown in Figure 1.



**Figure 1.** Flowchart of the symplectic geometric mode decomposition-cosine similarity (SGMD-CS) method.

## 3. Simulation Analysis

To verify the effectiveness of the proposed method, we refer to paper [20] to construct the complex amplitude modulation and frequency modulation signal to verify the SGMD-CS method. The constructed simulation signal is shown as Equation (13):

$$\begin{cases} x_1(t) = 2\sin(60\pi t) \times (1 + 0.5\sin(2\pi t)) \\ x_2(t) = \sin(120\pi t) \\ x_3(t) = 0.5\cos(10\pi t) \\ x(t) = x_1(t) + x_2(t) + x_3(t) \end{cases}$$
(13)

x(t) includes an amplitude modulation and frequency modulation component, a sine component, and a cosine component. The time-domain waveform of the simulation signal is shown in Figure 2a, and the frequency-domain waveform of the simulation signal is shown in Figure 2b:

Using the proposed SGMD-CS method to decompose the simulation signal, the decomposition results are shown in Figure 3. Here, taking the decomposition of the simulation signal as an example, each parameter in the decomposition process is explained. First, for the given simulation signal, we set the sampling frequency to 1000 and convert it into an original time series with a length of 1000. By calculating the power spectral density,  $f_{max}$  is estimated to be 29.7852, so the value of d is equal to 40. That is, after a series of transformations, 40 sets of initial components are finally decomposed. We need to reorganize the 40 sets of initial components with similar characteristics. To reduce the amount of calculation, we need to construct the signals  $N_f = \sum_{i=1}^f Y_i (f = 1, 2, \dots, 40)$  and calculate the cosine similarity  $\cos \theta_{k,k+1} (k = 1, 2, \dots 39)$  of two adjacent constructed signals, which is shown in Figure 4. It can be seen that  $\cos \theta_{k,k+1} = 1 (k \ge 6)$ ,  $\cos \theta_{6,7} = 1$  shows that the information between  $N_6$ and  $N_7$  is similar roughly. In other words, the information contained in the sum of the first 6 sets of components can be regarded that they already contain most of the information of the original series. So, the demarcation point k = 6 is a turning point for reconstructing the effective components, only the first 6 components need to be reconstructed.



Figure 2. Simulation signal. (a) Time domain; (b) Frequency domain.



Figure 3. The decomposition results of SGMD-CS. (a) Time domain; (b) Frequency domain.



Figure 4. Cosine similarity of adjacent constructing components.

Next, component reconstruction is performed by constructing a cosine similarity matrix for the first six components. Since the elements in the matrix represent the similarity between any two components, it is a symmetric matrix. For simplicity, only the elements above the main diagonal are listed. The color block diagram drawn according to the cosine similarity matrix is shown in Figure 5.



Figure 5. Matrix graph of cosine similarity.

It can be clearly seen from Figure 5 that the cosine similarity of the first component and second component is 0.9901, which can be considered that the two components are similar, and they are reorganized as  $SGC_1$ . Similarly, the cosine similarity between the third component and the fourth component is 0.9970, and they are regrouped into  $SGC_2$ . The fifth component and the sixth component, with a similarity of 0.9713, are recombined into  $SGC_3$ . The final decomposed components are shown in the first three waveforms of Figure 3. The fourth waveform represents the remaining component, that is, the sum of the initial components from the seventh to the fortieth component. The sum of these four parts is equal to the original simulated signal.

For comparison, the constructed simulation signals are decomposed with the LMD and EMD algorithms. The results of the decomposition are shown in Figures 6 and 7.

Decomposing the simulation signal by these three methods and comparing the obtained time-domain and frequency-domain waveforms, it can be seen that the SGMD-CS method can separate the trend components of the original simulation signal very well, and the components of the decomposition are almost equivalent to the components of the original signal. There is no doubt that the decomposition result is quite good. Moreover, the LMD method decomposes the original simulated signal into four components. It can be clearly seen in Figure 5 that the components  $PF_1$  and  $PF_2$  produce modal aliasing. The 30 Hz and 60 Hz components appear simultaneously in  $PF_1$ ; this situation is called underdecomposition. In addition, the 30 Hz component appears in  $PF_1$  and  $PF_2$ , and the 5 Hz component appears in  $PF_2$  and  $PF_3$ ; this situation is called over-decomposition. We can also find that the amplitude of some components is reduced, so the decomposition result is not ideal compared to the previous method. Similarly, the EMD method is used to decompose the simulated signal, and the decomposition result is shown in Figure 6. It can be seen that the decomposition obtains 8 components. Similar to LMD, under-decomposition and over-decomposition also occur during the decomposition process, and the decomposition result is not ideal. After analysis, regardless of whether the LMD or EMD is used, these two methods are not as effective as the SGMD-CS method for decomposing the simulation signal.



Figure 6. The decomposition results of local mean decomposition (LMD). (a) Time domain; (b) Frequency domain.



Figure 7. The decomposition results of empirical mode decomposition (EMD). (a) Time domain; (b) Frequency domain.

To further verify the similarity between the components decomposed by the SGMD-CS method and the components of the original simulated signal, the waveforms decomposed by the two approaches are drawn in Figure 8. It can be seen from the figure that the error of the decomposition result is smaller when using the proposed method, so this method can better strip the trend components of the complex signals. Therefore, in the wind turbine fault diagnosis process, the SGMD-CS method is first used to decompose the vibration signal collected by the sensor, thereby extracting the fault information hidden in the original signal, providing necessary materials for subsequent fault diagnosis, and improving the accuracy of fault diagnosis correspondingly.



Figure 8. Time domain waveforms of real components and SGMD-CS decomposition components.

## 4. Feature Classification

## 4.1. AdaBoost Theory

Boosting, as a meta-algorithm framework, is an important integrated learning technique that can reinforce weak learning classifiers, whose prediction accuracy is only slightly higher than that of random guessing, into strong predictors with high prediction accuracy. These weak classifiers can be any classifiers, such as decision trees, simple linear logistic classifiers, simple SVM classifiers, etc. This approach has been successfully applied to solve problems such as object detection, text analysis, and data mining. The most widely used boosting algorithm is AdaBoost, which was proposed by Freund and Schapire in 1996 [31]. As one of the best-supervised classification algorithms, it has been widely used to solve complex classification problems because of its simple concept [32,33].

Considering decision trees are resistant to overfitting since trees often have large edges and limited complexity, therefore, in this paper, we choose decision trees as the basic classifiers [34].

The specific algorithm steps are as follows:

- Step 1: Calculate the input.
  - (1) Given training set  $(x_1, y_1), \dots, (x_m, x_m)$ , where  $x_i \in X, y_i \in \{-1, +1\}$ .
  - (2) Weak learning algorithm.
- Step 2: Calculate the output  $H_{final}(x)$ .
  - (1) Initial weight distribution of training data.

$$D_1(i) = \frac{1}{m} (\forall i), \ i = 1, 2, \cdots, m$$
 (14)

(2) For  $t = 1, \dots, T$ , using the training data set with weight distribution  $D_t$  to learn, we obtain a weak classifier.

$$h_t(x): X \to \{-1, +1\}$$
 (15)

(3) Calculate the classification error rate on the training data set  $h_t(x)$ .

$$\varepsilon_t = \Pr_{i \sim D_t}[h_t(x_i) \neq y_i] \tag{16}$$

(4) Calculate the coefficient of  $h_t(x)$ .

$$\alpha_t = \frac{1}{2} \ln(\frac{1 - \varepsilon_t}{\varepsilon_t}) \tag{17}$$

It is worth noting that  $\alpha_t > 0$ .

(5) Update the weight distribution of the training data set.

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } y_i = h_t(x_i) \\ e^{\alpha_t} & \text{if } y_i \neq h_t(x_i) \end{cases}$$

$$= \frac{D_t(i)}{Z_t} \exp(-\alpha_t y_i h_t(x_i))$$
(18)

where  $Z_t$  is the normalization factor.

(6) Construct a linear combination of basic classifiers to obtain the final classifier.

$$H_{\text{final}}(x) = sign\left(\sum_{t} \alpha_t h_t(x)\right)$$
(19)

### 4.2. Feature Vector Selection

To obtain as much fault information as possible and improve the accuracy of fault diagnosis, many scholars have excelled in extracting feature vectors, trying to extract the fault features from various angles. Multidimensional feature vectors are selected frequently; however, due to the redundant information, the accuracy of fault diagnosis is reduced, which has forced scholars to seek various dimensionality reduction methods to further handle the selected feature vectors. Literature [35] extracted the fault information in the time domain, frequency domain, and time-frequency domain and then reduced the data through PCA to extract the fault features more comprehensively and realize the accurate identification of faults. In this article, unlike the concepts adopted by previous scholars, we do not consider how to select multidimensional feature vectors. Instead, we consider whether we can choose low-dimensional feature vectors.

Here, we adopt entropy to construct a low-dimensional feature vector. Entropy can usually be used to describe and quantify the degree of confusion in the system. At present, there are many estimation methods about entropy, such as approximate entropy, sample entropy, fuzzy entropy, etc., and they are widely used in various fields to measure the complexity of time series [36–38]. In this paper, calculate the symplectic geometric entropy by combining the SGMD-CS algorithm proposed above with the entropy, and select this index as the feature vector. The specific implementation steps are as follows: first, decompose the collected signal. Then, select the first two components decomposed to calculate the corresponding symplectic geometric entropy. Finally, the two-dimensional feature vector  $SymEn = [e_1, e_2]$  is input into the subsequent classifier for fault classification and recognition.

In the second section, we introduced the symplectic geometry algorithm. Through this algorithm, the eigenvalues of the real symmetric matrix A constructed by the trajectory matrix X can be obtained, and the values of  $\lambda_1, \lambda_2, \dots, \lambda_d$  decrease in sequence. The distribution of  $\lambda_i$  represents the spectral distribution of matrix A. Then, the probability of energy distribution in different directions can be defined as  $p_1, p_2, \dots, p_d$ , which are calculated as follows:

$$p_i = \frac{\lambda_i}{\sum\limits_{i=1}^d \lambda_i}$$
(20)

where *d* is the embedding dimension,  $0 \le p_i \le 1$ ,  $\sum_{i=1}^{d} p_i = 1$ , and  $p_i$  describes the uncertainty of entropy in different directions.

Then, the symplectic geometric entropy can be defined as follows:

$$SymEn = -\sum_{i=1}^{d} p_i \log(p_i)$$
(21)

Through formula (20) and formula (21), we can calculate symplectic geometric entropy as a feature vector to characterize the fault information. Finally, Figure 9 shows the complete fault diagnosis process.



Figure 9. Flowchart of fault diagnosis.

## 5. Experimental Analysis

## 5.1. Experimental Arrangement and Data Description

In this paper, the vibration data of rolling bearings at the Electrical Engineering Laboratory of Case Western Reserve University are selected, and the experimental platform is shown in Figure 10. Motor bearings were seeded with faults using electro-discharge machining (EDM). Faults ranging from 0.007 inches in diameter to 0.021 inches in diameter were introduced separately at the inner raceway, rolling element (i.e., ball), and outer raceway. The specifications of the rolling bearing tested in the experiment and related characteristic frequencies are shown in Table 1.



**Figure 10.** The test stand: 1. Fan end bearing; 2. Electric motor; 3. Drive end bearing; 4. Torque transducer/encoder; 5. Dynamometer.

 Table 1. Bearing specifications.

Class	Deep Groove Ball Bearing		
Туре	6205-2RS JEM SKF		
Position	Drive end		
Sampling frequency $f_s$ (Hz)	12,000		
Inside diameter (inches)	0.9843		
Outside diameter (inches)	2.0472		
Thickness (inches)	0.5906		
Ball diameter (inches)	0.3126		
Pitch diameter (inches)	1.537		
Rotation frequency (Hz)	fr		
Inner ring defect frequency $f_i$ (Hz)	$5.4152 \times f_r$		
Outer ring defect frequency $f_o$ (Hz)	$3.5848 \times f_r$		
Rolling element frequency $f_b$ (Hz)	$4.7135 \times f_r$		
Number of rolling elements	9		

#### 5.2. Signal Preprocessing

The proposed SGMD-CS method is applied to actual rolling bearing signals to prove its effectiveness and feasibility, so it can be used to realize the fault diagnosis of wind turbine bearings. The specific technical parameters of the analyzed bearing are shown in Table 2. Extract 12,000 data points for analysis.

Table 2. the specific technical parameters of analyzed bearing.

Parameters	Value		
Data	105.mat		
Fault diameter	0.007"		
Approximate motor speed	1797 rpm		
Rotation frequency	29.95 Hz		
Inner ring defect frequency	162.19 Hz		
Motor load	0 hp		

The time-domain waveform of the analyzed inner ring fault rolling bearing signal is shown in Figure 11. It is decomposed by the SGMD-CS method, and the first two symplectic geometric components are spectrum analyzed, as shown in Figure 12. Taking the amplitude spectrum of  $SGC_2$  as an example, the spectrum line at 161.9 Hz can be clearly seen from the figure. Because 161.9 Hz is close to the theoretical value 162.19 Hz, which is the inner ring fault frequency of rolling bearing, it can be considered that the inner ring fault has been identified. In addition, the sidebands of the high-order harmonics of the inner ring fault frequency are prominent and can be clearly observed, whose amplitude are modulated by  $f_r$ .



Figure 11. Time domain analysis result of rolling bearing with inner fault.



Figure 12. The decomposition results of SGMD-CS: (a) Time domain; (b) Frequency domain.

For comparison, the LMD and EMD methods are used to process the fault signal in the same way. The decomposition results are shown in Figures 13 and 14. It can be seen from the figure that there are so many interference components with large amplitude that the characteristic frequency of the fault is almost submerged in the noise signal. Characteristic frequency cannot be accurately identified, which brings certain difficulties to the judgment of the fault of the rolling bearing.



Figure 13. The decomposition results of LMD: (a) Time domain; (b) Frequency domain.



Figure 14. The decomposition results of EMD: (a) Time domain; (b) Frequency domain.

## 5.3. Classification of Different Fault Types

In the previous section, a series of comparative experiments were conducted to verify the effectiveness of the proposed SGMD-CS method for processing rolling bearing fault signals. According to the flow chart of fault diagnosis process shown in Figure 8, this method is combined with machine learning to realize the intelligent diagnosis of bearing faults.

In order to classify the different fault categories of the bearing, the original vibration signal data of each fault type is taken to intercept the first 50 windows, each window contains 1024 points, so as to obtain 50 fault samples for every type of fault. For each sample, the SGMD-CS method is used to decompose, the first two SGC components are selected for analysis, and the symplectic geometric entropy is calculated. Therefore, for four different fault types, the entire data set contains 200 samples, and the classification accuracy is verified by the five-fold cross-validation method. At the same time, in order to verify the validity of the symplectic geometric entropy as the extracted feature, approximate entropy, sample entropy, and fuzzy entropy are used for comparison. The specific data set description is shown in Table 3.

Table 3. Descri	iption of ex	operimental	dataset.
-----------------	--------------	-------------	----------

Fault Type	Data	Fault Diameter (Inches)	Motor Load (HP)	Rotation Frequency (r/min)	Class Label
Inner ring fault (IRF)	171.mat	0.014	2	1750	1
Outer ring fault (ORF)	199.mat	0.014	2	1750	2
Ball element fault (BF)	187.mat	0.014	2	1750	3
Normal (NOR)	099.mat	0.014	2	1750	4

Take one of the experiments as an example for specific analysis. Figure 15 shows the classification results of the model for different extracted feature vectors in the form of a confusion matrix. For the classification model that uses symplectic geometric entropy as the feature vector, the classification accuracy rate is 97.5%, and one misjudgment occurs because the rolling element fault is misjudged as normal. In terms of the approximate entropy is used as the feature vector, the classification accuracy rate is 95%, 2 misjudgments occur between the rolling element fault and the normal state. Using sample entropy as the

feature vector model, the classification accuracy rate is 90%. There are 4 misjudgments, and 3 misjudgments are relevant to rolling element fault. For the model that uses fuzzy entropy as the feature vector, the classification accuracy rate is only 70%, and there are 12 misjudgments, of which 8 misjudgments are related to the rolling element fault. To evaluate the performance of the classifier only from the index of accuracy, it can be considered that the symplectic geometric entropy as the extracted fault feature has a certain validity, while the fuzzy entropy calculated as the extracted fault feature under this condition performs poorly. In addition, we can also find that most of the misjudgments are related to the rolling element fault.



Figure 15. The confusion matrix: (a) SymEn; (b) ApEn; (c) SampEn; (d) FuzzyEn.

In order to comprehensively evaluate the classification effect of different models, four common evaluation indicators are used as shown in Figure 16. They are accuracy, precision, recall, and F1-score. Based on the four indicators, it is not difficult to draw the same conclusion as the previous one, that is, extracting the symplectic geometric entropy of the bearing vibration signal as a feature vector is extremely superior and more stable. The specific values of each evaluation index are shown in Table 4.

Similarly, the comparative results of other experiments are analyzed to avoid obvious deviations in a single experiment. In all experiments, calculate the mean and standard deviation of the experimental results under different evaluation standards, as shown in Table 5. On the whole, the accuracy of model that uses symplectic geometric entropy as the feature vector is the highest, followed by approximate entropy and sample entropy, and the accuracy of fuzzy entropy is the lowest. Other evaluation criteria also roughly conform to this law. And no matter which evaluation standard is used, the value of rolling element fault is lower than that of the other three fault types, so it is prone to misjudgment. Even under this kind of unfavorable conditions, the symplectic geometric entropy still has advantages than the other three types of entropy, reflecting its advantages in feature extraction.

The above experimental results show that the wind turbine fault feature extraction method based on the SGMD-CS and AdaBoost framework proposed in this paper is correct and effective, which has a certain significance for the bearing fault identification of wind turbines and then provide a corresponding basis for a technician to carry out maintenance work.



(**d**)

Figure 16. Evaluation criteria: (a) Accuracy; (b) Precision; (c) Recall; (d) F1-score.

		SymEn	ApEn	SampEn	FuzzyEn
	Accuracy	97.50	95.00	90.00	70.00
	Precision	100.00	100.00	91.67	63.64
Inner ring	Recall	100.00	100.00	100.00	63.64
	F1-score	100.00	100.00	95.65	63.64
	Precision	100.00	100.00	100.00	61.54
Outer ring	Recall	100.00	100.00	100.00	100.00
0	F1-score	100.00	100.00	100.00	76.19
Ball element	Precision	100.00	88.89	100.00	33.33
	Recall	88.89	88.89	66.67	11.11
	F1-score	94.12	88.89	80.00	16.67
Normal	Precision	92.31	91.67	78.57	92.31
	Recall	100.00	91.67	91.67	100.00
	F1-score	96.00	91.67	84.62	96.00

Table 4. Result of classification.

Recall

Table 5. Result of statistics.

		SymEn	ApEn	SampEn	FuzzyEn
	Accuracy	$93.00\pm5.12$	$91.00\pm4.54$	$88.50\pm7.83$	$67.50\pm3.95$
Inner ring	Precision Recall F1-score	$\begin{array}{c} 89.67 \pm 10.83 \\ 100.00 \pm 0.00 \\ 94.27 \pm 6.13 \end{array}$	$\begin{array}{c} 88.14 \pm 11.39 \\ 97.78 \pm 4.97 \\ 92.32 \pm 7.08 \end{array}$	$\begin{array}{c} 88.14 \pm 11.39 \\ 97.78 \pm 4.97 \\ 92.32 \pm 7.08 \end{array}$	$\begin{array}{c} 51.90 \pm 21.18 \\ 49.83 \pm 14.68 \\ 49.21 \pm 13.09 \end{array}$
Outer ring	Precision Recall F1-score	$\begin{array}{c} 96.00 \pm 8.94 \\ 92.35 \pm 8.39 \\ 94.08 \pm 8.27 \end{array}$	$\begin{array}{c} 96.57 \pm 4.80 \\ 92.57 \pm 12.99 \\ 93.99 \pm 7.24 \end{array}$	$\begin{array}{c} 93.70 \pm 8.80 \\ 96.57 \pm 4.80 \\ 94.88 \pm 5.09 \end{array}$	$\begin{array}{c} 69.31 \pm 13.17 \\ 88.70 \pm 7.29 \\ 76.88 \pm 7.74 \end{array}$
Ball element	Precision Recall F1-score	$\begin{array}{c} 88.95 \pm 12.08 \\ 88.82 \pm 10.58 \\ 88.71 \pm 10.31 \end{array}$	$\begin{array}{c} 85.99 \pm 14.17 \\ 77.10 \pm 15.24 \\ 80.52 \pm 12.79 \end{array}$	$\begin{array}{c} 86.48 \pm 16.34 \\ 67.84 \pm 25.46 \\ 74.57 \pm 21.12 \end{array}$	$\begin{array}{c} 50.33 \pm 17.89 \\ 36.72 \pm 21.71 \\ 41.10 \pm 20.40 \end{array}$
Normal	Precision Recall F1-score	$\begin{array}{c} 98.46 \pm 3.44 \\ 89.27 \pm 6.67 \\ 93.43 \pm 2.20 \end{array}$	$\begin{array}{c} 94.40 \pm 5.48 \\ 95.83 \pm 5.89 \\ 95.09 \pm 5.45 \end{array}$	$\begin{array}{c} 85.48 \pm 11.12 \\ 90.00 \pm 10.87 \\ 87.44 \pm 9.99 \end{array}$	$\begin{array}{c} 98.46 \pm 3.44 \\ 92.35 \pm 7.48 \\ 95.10 \pm 3.37 \end{array}$

# 6. Conclusions

This paper proposes a fault feature extraction method for wind turbines based on the SGMD-CS and AdaBoost framework, which has been improved and optimized as follows:

- 1. To address the problem of similar component recombination in the symplectic geometric decomposition process, the cosine similarity is introduced into this method, the SGMD-CS method is proposed, and a block diagram of the method is given. The effectiveness of this method is verified by constructing a complex AM-FM signal and comparing it with the decomposition results of the LMD and EMD methods. The results show that the decomposition error of this method is small, and the trend components of the original signal can be better stripped, so this method is suitable for the analysis of nonlinear time series. In addition, the characteristics of this method have also been compared and verified on actual rolling bearing fault signals.
- 2. When addressing the problem of using high-dimensional feature vectors when extracting fault feature information, there will be data redundancy, and the diagnosis accuracy will be reduced. In this paper, based on the SGMD-CS method, the symplectic geometric entropy is calculated as a low-dimensional feature vector and sent to the AdaBoost classification framework based on decision trees. According to the given fault diagnosis flow chart, taking the rolling bearing vibration data of Case Western Reserve University's Electrical Engineering Laboratory as an example, a high classification accuracy rate is obtained by discriminating and classifying the fault type. At the same time, compared with sample entropy, approximate entropy, and fuzzy entropy, symplectic geometric entropy is highlighted as a measure that can effectively extract fault information, thereby making the diagnosis more accurate.

**Author Contributions:** Conceptualization, H.L. and F.L.; methodology, R.J.; validation, H.L., F.L., and R.J.; formal analysis, L.B.; investigation, X.L.; resources, X.L.; writing—original draft preparation, F.L.; writing—review and editing, H.L.; writing—translation and sentence checking, F.Z.; visualization, L.B. and F.Z.; supervision, R.J.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Natural Science Foundation of China (grant number 51779206) and Scientific Research Program funded by the Shanxi Provincial Education Department (Program No.17JK0570).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

## References

- 1. Liu, Z.; Zhang, L. A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. *Measurement* **2020**, *149*, 107002. [CrossRef]
- 2. Ehsan, M.; David, W.; Qiao, S. Indicative Fault Diagnosis of Wind Turbine Generator Bearings Using Tower Sound and Vi-bration. *Energies* **2017**, *10*, 1853.
- 3. Francesco, C.; Luigi, G.; Alessandro, P.D.; Davide, A.; Francesco, N. Diagnosis of Faulty Wind Turbine Bearings Using Tower Vibration Measurements. *Energies* **2020**, *13*, 1474.
- 4. Daniel, S.; Pär, M.; Kim, B.; Per-Erik, L. Bearing monitoring in the wind turbine drivetrain: A comparative study of the FFT and wavelet transforms. *Wind Energy* **2020**, *23*, 1381–1393.
- 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. *Proc. R. Soc. A* 1998, 454, 903–995. [CrossRef]
- Wu, Z.; Huang, N.E. Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Adv. Adapt. Data Anal.* 2009, 1, 1–41. [CrossRef]
- Liu, B.; Zheng, P.; Dai, Q.; Zhou, Z. The Measurement and Elimination of Mode Splitting: From the Perspective of the Partly Ensemble Empirical Mode Decomposition. *Complexity* 2018, 2018, 4230649. [CrossRef]
- 8. Guo, T.; Deng, Z. An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing. *Appl. Acoust.* **2017**, 127, 46–62. [CrossRef]
- 9. Zhang, C.; Li, Z.; Hu, C.; Chen, S.; Wang, J.; Zhang, X. An optimized ensemble local mean decomposition method for fault detection of mechanical components. *Meas. Sci. Technol.* **2017**, *28*, 035102. [CrossRef]

- 10. Wang, Z.; Wang, J.; Cai, W.; Zhou, J.; Du, W.; Wang, J.; He, G.; He, H. Application of an Improved Ensemble Local Mean Decomposition Method for Gearbox Composite Fault Diagnosis. *Complexity* **2019**, 2019, 1564243. [CrossRef]
- 11. Xie, H.; Wang, Z.; Huang, H. Identification determinism in time series based on symplectic geometry spectra. *Phys. Lett. A* 2005, 342, 156–161. [CrossRef]
- 12. Fassbender, H.; Kressner, D. Structured Eigenvalue Problems. GAMM Mitt. 2006, 29, 297–318. [CrossRef]
- 13. Lei, M.; Wang, Z.; Feng, Z. A method of embedding dimension estimation based on symplectic geometry. *Phys. Lett. A* 2002, 303, 179–189. [CrossRef]
- 14. Lei, M.; Meng, G. Symplectic Principal Component Analysis: A New Method for Time Series Analysis. *Math. Probl. Eng.* 2011, 2011, 1–14. [CrossRef]
- 15. Xie, H.-B.; Dokos, S.; Sivakumar, B.; Mengersen, K. Symplectic geometry spectrum regression for prediction of noisy time series. *Phys. Rev. E* 2016, *93*, 93. [CrossRef]
- 16. Xie, H.-B.; Dokos, S. A symplectic geometry-based method for nonlinear time series decomposition and prediction. *Appl. Phys. Lett.* **2013**, *103*, 054103. [CrossRef]
- 17. Xie, H.-B.; Guo, T.; Sivakumar, B.; Liew, A.W.-C.; Dokos, S. Symplectic geometry spectrum analysis of nonlinear time series. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 2014, 470, 20140409. [CrossRef]
- 18. Lei, M.; Meng, G.; Dong, G. Fault Detection for Vibration Signals on Rolling Bearings Based on the Symplectic Entropy Method. *Entropy* 2017, 19, 607. [CrossRef]
- Zheng, Z.; Xin, G. Fault Feature Extraction of Hydraulic Pumps Based on Symplectic Geometry Mode Decomposition and Power Spectral Entropy. *Entropy* 2019, 21, 476. [CrossRef] [PubMed]
- 20. Jin, H.; Lin, J.; Chen, X.; Yi, C. Modal Parameters Identification Method Based on Symplectic Geometry Model Decomposition. *Shock. Vib.* **2019**, 2019, 1–26. [CrossRef]
- 21. Li, X.; Li, D. Structural damage identification based on symplectic geometric spectrum analysis method under the influence of environmental factors. *J. Water Resour. Archit. Eng.* **2016**, *14*, 154–160, 176.
- 22. Niu, X.; Qu, F.; Wang, N. Analysis and Evaluation of Surface EMG Signals of Athletes Based on EMD and Symplectic Geometry. J. Ocean Univ. China Nat. Sci. Ed. 2005, 1, 125–129.
- 23. Pan, H.; Yang, Y.; Li, X.; Zheng, J.; Cheng, J. Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis. *Mech. Syst. Signal Process.* **2019**, *114*, 189–211. [CrossRef]
- 24. Kalhori, H.; Alamdari, M.M.; Ye, L. Automated algorithm for impact force identification using cosine similarity searching. *Measurement* **2018**, *122*, 648–657. [CrossRef]
- 25. Al-Anzi, F.S.; AbuZeina, D. Toward an enhanced Arabic text classification using cosine similarity and Latent Semantic Indexing. J. King Saud Univ. Comput. Inf. Sci. 2016, 29, 189–195. [CrossRef]
- Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* 2018, 108, 33–47. [CrossRef]
- 27. Lei, M.; Meng, G.; Zhang, W.; Wade, J.; Sarkar, N. Symplectic Entropy as a Novel Measure for Complex Systems. *Entropy* **2016**, *18*, 412. [CrossRef]
- 28. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980;* Springer: Berlin/Heidelberg, Germany, 1981; pp. 366–381.
- 29. Bonizzi, P.; Karel, J.M.; Meste, O.; Peeters, R.L. Singular spectrum decomposition: A new time series decomposition. *Adv. Adapt. Data Anal.* **2014**, *6*, 107–109. [CrossRef]
- 30. Van Loan, C. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.* **1984**, *61*, 233–251. [CrossRef]
- Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. In Proceedings of the International Conference on International Conference on Machine Learning (ICML), Bari, Italy, 3–6 July 1996; Morgan Kaufmann Publishers Inc.: Bari, Italy, 1996; Volume 96, pp. 148–156.
- 32. Baig, M.M.; Awais, M.M.; El-Alfy, E.S.M. AdaBoost-based artificial neural network learning. *Neurocomputing* **2017**, 248, 120–126. [CrossRef]
- 33. Lee, W.; Jun, C.-H.; Lee, J.-S. Instance categorization by support vector machines to adjust weights in AdaBoost for imbalanced data classification. *Inf. Sci.* 2017, 381, 92–103. [CrossRef]
- 34. Schapire, R.E.; Freund, Y. Boosting: Foundations and Algorithms; MIT Press: Cambridge, MA, USA, 2013.
- 35. Li, H.; Fan, B.; Jia, R.; Zhai, F.; Bai, L.; Luo, X. Research on Multi-Domain Fault Diagnosis of Gearbox of Wind Turbine Based on Adaptive Variational Mode Decomposition and Extreme Learning Machine Algorithms. *Energies* **2020**, *13*, 1375. [CrossRef]
- 36. Gao, X.; Yan, X.; Gao, P.; Gao, X.; Zhang, S. Automatic detection of epileptic seizure based on approximate entropy, recurrence quantification analysis and convolutional neural networks. *Artif. Intell. Med.* **2020**, *102*, 101711. [CrossRef]
- Wu, H.; Zhou, J.; Xie, C.; Zhang, J.; Huang, Y. Two-dimensional time series sample entropy algorithm: Applications to rotor axis orbit feature identification. *Mech. Syst. Signal Process.* 2021, 147, 107123.
- Harezlak, K.; Kasprowski, P. Application of Time-Scale Decomposition of Entropy for Eye Movement Analysis. *Entropy* 2020, 22, 168. [CrossRef]