

A novel feature extraction method based on symbol-scale diversity entropy and its application for fault diagnosis of rotary machines

Shun Wang¹ , Yongbo Li¹ , Jiacong Zhang¹,
 Zheng Liu² and Zichen Deng¹

Abstract

Multiscale entropy-based methods have made great progress in the field of health condition monitoring and fault diagnosis of machines due to their powerful feature representation capabilities. However, existing multiscale entropy methods suffer from three major obstacles: high fluctuation under large scale-factor, loss of high-frequency information, and poor robustness to noises. Thus, this work proposes a symbol-scale analysis method to deal with the above problems. In one aspect, to capture fault features from the time series over multiple time scales, time-delay process of different intervals is utilized to obtain long-term features and short-term features. In the other aspect, symbol-scale analysis introduces a symbolization procedure and maps time series into a corresponding sequence of symbols to overcome the limitation of weak fault extraction under a low-signal-to-noise ratio environment. Moreover, the symbol-scale entropy approach is developed by integrating with diversity entropy, called symbol-scale diversity entropy. The effectiveness of the proposed strategy is intensively validated using two simulated signals and experimental cases. Results demonstrate its advantages in dynamic change tracking ability and calculation efficiency by comparing it with other state-of-the-art entropy methods. Apart from diversity entropy, the versatility of incorporating the proposed symbol-scale analysis and other entropy methods is also verified using experimental data.

Keywords

Symbol-scale analysis, time-delay analysis, diversity entropy, feature extraction, fault diagnosis, multiscale entropy

Introduction

Fault diagnosis plays a significant role in the health management of rotating machinery, which aims to ensure safe operation and reduction of unnecessary economic losses. Vibration-based diagnostics have garnered considerable attention from researchers in both academia and industry due to their effectiveness in identifying faults and potential issues in machinery.^{1–3} However, the vibration signals with partial damage contains much noise and thus represent nonlinear and nonstationary characteristics, which pose difficulties in rotating machinery fault diagnosis.^{4,5}

Fortunately, the nonlinear dynamic indicators can establish the bridge between the measured vibration signal and fault stages by measuring their dynamic change.⁶ Quantitative analysis of rotating machinery-generated time series signals based on nonlinear dynamic methods clarifies the law of system dynamics.

Entropy measure, with powerful feature representation capability, has made great progress in structural health monitoring^{7,8} and fault diagnosis of rotating machinery.^{9–12}

Several well-known entropy measures have been developed, including approximate entropy,¹³ sample entropy (SE),¹⁴ permutation entropy (PE),¹⁵ fuzzy entropy (FE),¹⁶ dispersion entropy,¹⁷ and so on.^{18–20}

¹MIIT Key Laboratory of Dynamics and Control of Complex System, School of Aeronautics, Northwestern Polytechnical University, Xian, China

²School of Engineering, University of British Columbia, Kelowna, BC, Canada

Corresponding author:

Yongbo Li, MIIT Key Laboratory of Dynamics and Control of Complex System, School of Aeronautics, Northwestern Polytechnical University, No. 127, Youyi Road (West), Xian, Shaanxi 710072, China.
 Email: yongbo@nwpu.edu.cn

These entropy techniques offer valuable insights into the underlying patterns and complexity of signals. In recent years, a new entropy measure called diversity entropy²¹ has been proposed. To further enhance the capability of entropy-based methods for signal analysis, symbolic dynamic filtering (SDF)²² has been incorporated with entropy methods.^{23–27} These approaches combine the advantages of entropy analysis with symbolic dynamics, allowing for more effective extraction of meaningful features from signals. By transforming the continuous signals into symbolic sequences and applying entropy measures to these symbolic sequences, SDF enhances the detection and characterization of specific patterns or abnormalities in the signals.

These entropy methods quantify the complexity through only a single scale, which neglects the correlations between the different time scales. Considering this problem, a multiscale analysis based on coarse-graining analysis was developed by Costa.²⁸ With the help of the coarse-graining analysis, multiscale-based entropy methods, such as multiscale SE, multiscale fuzzy entropy (MFE), multiscale permutation entropy (MPE), and multiscale diversity entropy (MDE) with modified multiscale analysis have been developed for comprehensive complexity evaluation of time series.^{29–31} Unfortunately, the data length can be shortened rapidly as the scale increases, resulting in inaccurate estimation at larger scales for traditional multiscale analysis. To avoid this shortcoming, composite multiscale entropy⁹ and refined composite multiscale entropy^{32,33} have been developed via refining the coarse-grained time series strategy to overcome the limitation of data length. In addition, D. Pham³⁴ proposed an improved multiscale analysis using the time-delay process to solve this problem. These methods have been applied in mechanical system fault diagnosis^{35,36} and structural health monitoring^{7,8} by extracting information from complex and nonlinear time series.

However, there are still problems in the traditional coarse-graining multiscale analysis. First, coarse-graining analysis used in many multiscale-based entropy methods generally leads to large fluctuations and deviations in a large scale factor. Second, traditional multiscale entropy methods based on coarse-grained analysis aim to obtain sequences under different scales using averaging. However, the averaging step is essentially a smoothing process of the original signal, which may result in the loss of information embedded in the high-frequency component of signals.³⁷ Lastly, in real acquisition data, the signals can be distorted by interference or noise. Finally, in the real acquisition data, the signals may be disturbed by interferences or noises. The direct use of entropy without noise

reduction can heavily influence the accuracy of feature extraction and further fault type identification.

In this work, a novel scale-analysis-based model named symbol-scale analysis is proposed to measure complexity under different scales via time-delay process^{34,38} and SDF.²² On the one hand, to relieve the limits of the above situations, inspired by a method to calculate the fractal dimension of time series,³⁹ a time-delay process with different intervals is introduced as the multiscale method. Unlike the past coarse-graining multiscale-based entropy methods, by doing this, the long-term features and short-term features can be obtained. The long-term features indicate the overall trend of signals, while the short-term features represent subtle variations in localized regions, where both long-term features and short-term features can be potentially vital to measure the complexity of signals and accomplish the classification tasks. On the other hand, unlike the coarse-graining-based multiscale approach that will lose high-frequency information due to the smoothing process, the multiscale time series generated here are completely based on the amplitude information of original data. Hence, the multiscale analysis based on time-delay process can effectively retain the important constructing information of original time series without information loss. Moreover, SDF is used to reduce the noise induced by the environment and enhance robustness in a low signal-to-noise ratio (SNR) environment. The original time series will be discretized into symbolic sequences by SDF so that the noise-related components will be reduced and the accuracy of fault feature extraction can be improved. Finally, the symbolic data will decrease the calculation complexity, thereby resulting in high computation efficiency.

In addition, symbol-scale analysis is further integrated with diversity entropy, and then a novel metric called symbol-scale diversity entropy (SSDE) is developed. In view of the advantages of SSDE, an intelligent diagnosis scheme is proposed by incorporating SSDE and support vector machine (SVM). The performance of the proposed SSDE method is firstly verified by simulation analysis of the bearing and gearbox. To establish the versatility of SSDE approach, this is followed by two different experimental case studies of rolling bearing and planetary gearbox. Compared with MFE, MPE, and MDE, the proposed SSDE method can achieve the highest recognition accuracy with the best robustness against noise and excellent calculation efficiency.

The main contributions of this work can be summarized as follows:

- (1) To capture the fault characteristics from the time series over multiple time scales, the time-delay

- process of different intervals is utilized to obtain long-term features and short-term features.
- (2) Symbol-scale analysis introduces a symbolization procedure and maps time series into a corresponding sequence of symbols to reduce the noise-related fluctuations of signals and overcome the limitation of weak fault extraction under a low-SNR environment.
 - (3) A novel metric called SSDE is developed, and the effectiveness of SSDE approach in fault feature extraction of rotating machines is verified by both simulated and experimental case studies of bearing and gearbox.

The remainder of this paper is organized as follows. The basic concepts of SSDE and the main steps of the proposed SSDE-based fault diagnosis approach are introduced in section “Methodology.” In section “Verification using simulated signals,” the effectiveness of the proposed method is demonstrated by simulated gearbox signals and bearing signals. Section “Experimental case studies” assesses the performance of the proposed SSDE-based method in fault condition recognition using two experimental case studies. Lastly, in section “Conclusion,” we conclude this article.

Methodology

Problem statement and motivation

In recent years, although various entropy methods and their multiscale extensions have been widely used in mechanical signals analysis, some deficiencies remain to be solved. A novel multiscale analysis strategy is an urgent need to solve the dilemma of information loss and poor robustness to noises. Based on the above analysis, the deficiencies of existing multiscale entropy methods can be concluded as two aspects:

- (1) Intrinsic limit of traditional coarse-graining multiscale method. In traditional multiscale entropy calculation, coarse-graining processing is a necessary step, where a typical coarse-graining strategy in MPE or MFE is a non-overlapping extraction averaging process. In fact, as can be seen from Figure 1(a), the average operation in coarse-graining multiscale analysis is similar to smoothing the raw signal, which is linear mean filtering. Therefore, the linear mean filtering will discard the spectral response characteristics that contain important information.³⁷ In addition, coarse-

graining analysis generally leads to large fluctuation and deviation in large scale factors, resulting in an inaccurate estimation of complexity and poor accuracy.

- (2) Drawbacks about robustness in realistic signals. The disturbances and noises could affect the estimation of entropy-based analysis by changing the frequency distribution and increasing the standard deviation of the data. The direct use of entropy without noise reduction can seriously influence the accuracy of feature extraction and fault type identification. Hence, when the entropy or multiscale entropy method is applied to the actual mechanical signal analysis, the noise reduction process is necessary to ensure more reliable analysis results.

To relieve the limits of the above situations, a novel multiscale model named symbol-scale analysis is first proposed to measure complexity under different scales via time-delay process and SDF. First, SSDE utilizes the time-delay process of different intervals to capture the fault characteristics from the time series over multiple time scales. Second, SSDE combines the merits of SDF in noise reduction and calculation efficiency, where symbolization is designed to overcome the limitation of weak fault extraction under a low-SNR environment. It can be indicated that SSDE successfully overcomes drawbacks of traditional multiscale entropy, and improves estimation accuracy and robustness significantly.

To visually illustrate the limitations of existing multiscale methods, Figure 1(b) presents a comparative analysis of information extraction and restoration abilities between the traditional coarse-graining multiscale method and the proposed symbol-scale analysis method. The analysis was conducted using a simulated gear signal from section “Simulated gearbox signals,” focusing on the spectral response characteristics at scale factors: 1, 2, and 3.

This analysis aims to highlight the differences in spectral characteristics obtained from the two methods. Specifically, the coarse-grained sequence obtained from the traditional method exhibits weak spectral components at the gear mesh frequency (f_m) and its harmonics, as illustrated in Figure 1(a). On the contrary, as demonstrated in the left column of Figure 1(b), the symbol-scale sequence obtained from the proposed method shows notable amplitudes at f_m and its harmonics. By showcasing these distinct spectral responses, the figure underscores the improved performance of

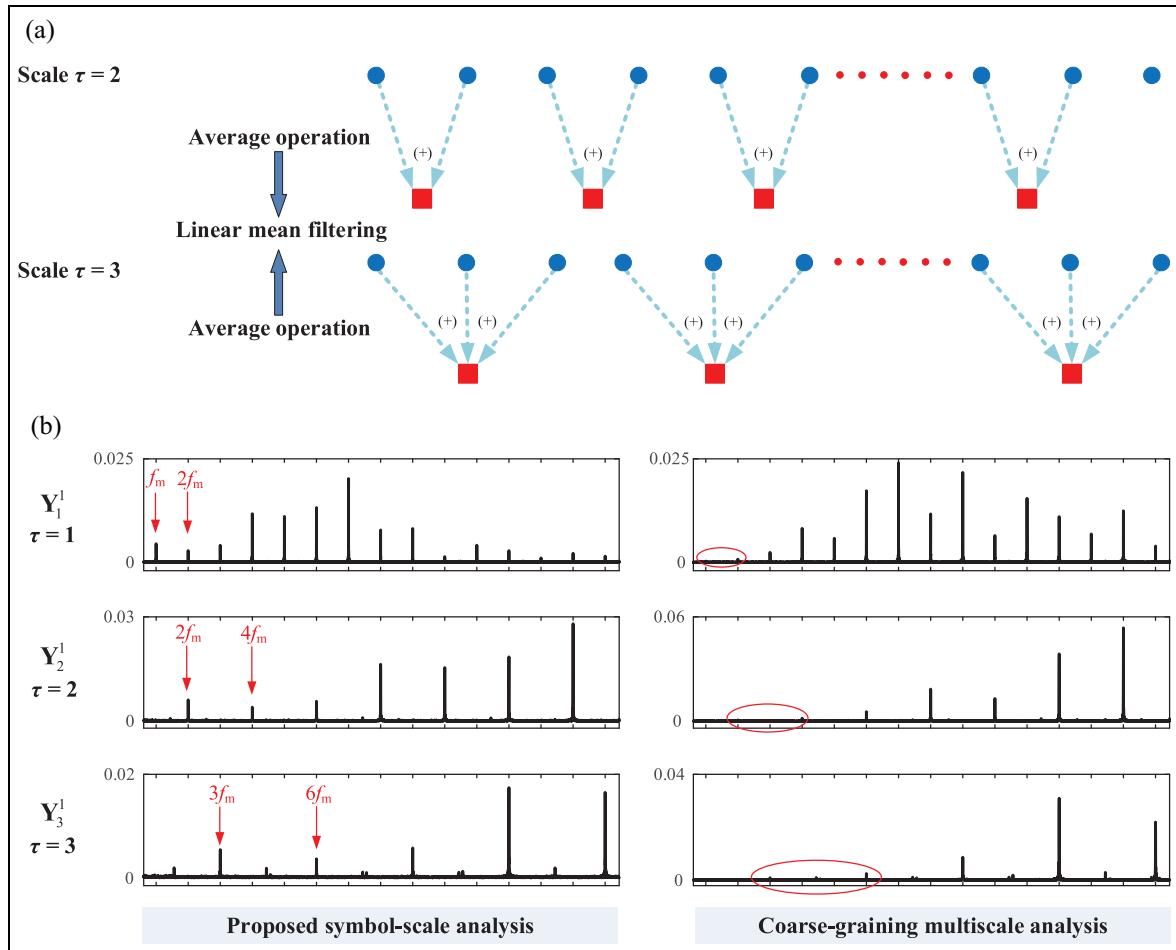


Figure 1. Limits summary of existing multiscale method: (a) coarse-graining multiscale method and (b) normalized spectrum of coarse-graining sequences and symbol-scale sequences at scale factors: 1, 2, and 3. f_m is gear mesh frequency.

the symbol-scale analysis method in capturing important frequency components that may be masked or weakened by traditional multiscale approaches.

Symbol-scale diversity entropy

In this part, we proposed a novel SSDE method. Overall, there are two main steps in our SSDE algorithm: (1) the symbol-scale algorithm aiming to generate multiple time series under different scales; (2) entropy calculation using the probability distribution-based diversity entropy.

Symbol-scale analysis. For a given time series $X = \{x_1, \dots, x_i, \dots, x_N\}$ with length N and scale parameter τ , the defined symbol-scale analysis can be accomplished following three steps:

Step 1: The original series is reconstructed by time-delay analysis of different intervals. For time series $X = \{x_1, \dots, x_i, \dots, x_N\}$, α new time series can be

constructed using Equation (1):

$$Y_\alpha^\beta = \left(x_\beta, x_{\beta+\alpha}, x_{\beta+2\alpha}, \dots, x_{\beta + \lceil \frac{N-\beta}{\alpha} \rceil \alpha} \right) \quad (1)$$

where $\beta = 1, 2, \dots, \alpha$. β and α are positive integers, where β is the initial point, α represents the step size of time-delay analysis and $\lceil N - \beta/\alpha \rceil$ indicates the nearest integer that does not exceed $N - \beta/\alpha$. In the case of $\alpha = 3$ and $N = 100$, three time series obtained by the above process can be described as follows:

$$\begin{aligned} Y_3^1 &: x_1, x_4, x_7, \dots, x_{97}, x_{100} \\ Y_3^2 &: x_2, x_5, x_8, \dots, x_{98} \\ Y_3^3 &: x_3, x_6, x_9, \dots, x_{99} \end{aligned} \quad (2)$$

Step 2: The obtained α time series Y_α^β , $\beta = 1, 2, \dots, \alpha$ are transformed into α symbol sequences S_α^β , $\beta = 1, 2, \dots, \alpha$ with ε symbols, where ε represents

the potential number of symbols. These α symbol sequences are the symbol-scale sequences. Here, maximum entropy partitioning method²² is utilized to accomplish the symbolization process, which is achieved by inducing the symbols to present a uniform probability distribution.

Step 3: Repeat Steps (1) and (2) for $\alpha = 1, 2, 3, \dots, \tau$, and obtain new symbol-scale sequences with different α values.

Calculation procedure of SSDE. Based on the symbol-scale analysis, the symbol-scale entropy is proposed to measure complexity under different symbol scales by integrating with diversity entropy. For a given time series $X = \{x_1, \dots, x_i, \dots, x_N\}$ with length N , the defined SSDE can be accomplished by following six steps:

Step 1: Reconstruct the original time series by symbol-scale analysis and obtain α symbol-scale sequences \mathbf{S}_α^β , $\beta = 1, 2, \dots, \alpha$ referring to section “Symbol-scale analysis.”

Step 2: Conduct the phase-space reconstruction for α symbol-scale sequences \mathbf{S}_α^β , $\beta = 1, 2, \dots, \alpha$. Based on Taken’s embedding theorem, construct the phase space matrix $\mathbf{S}(m)$ with a dimension of m as follows:

$$\begin{aligned} \mathbf{M}_\alpha^\beta(m) &= \left[\mathbf{M}_{\alpha,1}^\beta(m), \dots, \mathbf{M}_{\alpha,i}^\beta(m), \dots, \mathbf{M}_{\alpha,N-(m-1)}^\beta(m) \right]^T \\ &= \begin{bmatrix} \mathbf{S}_\alpha^\beta(1) & \mathbf{S}_\alpha^\beta(2) & \dots & \mathbf{S}_\alpha^\beta(1+(m-1)) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{S}_\alpha^\beta(i) & \mathbf{S}_\alpha^\beta(i+1) & \dots & \mathbf{S}_\alpha^\beta(i+(m-1)) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_\alpha^\beta(N-(m-1)) & \mathbf{S}_\alpha^\beta(N-m+2) & \dots & \mathbf{S}_\alpha^\beta(N) \end{bmatrix} \quad (3) \end{aligned}$$

where $\mathbf{s}_{\alpha,i}^\beta(m)$ represents the i -th embedding vector of length m for phase space matrix.

Step 3: Compute cosine similarity between adjacent vectors and obtain the cosine similarity matrix $\mathbf{Sim} = \{Sim_1, \dots, Sim_i, \dots, Sim_{N-m}\}$. Here, the cosine similarity between two vectors $\mathbf{M}_{\alpha,i}^\beta(m)$ and $\mathbf{M}_{\alpha,i+1}^\beta(m)$ can be expressed as Equation (4):

$$\begin{aligned} Sim_i &= Sim(\mathbf{M}_{\alpha,i}^\beta(m), \mathbf{M}_{\alpha,i+1}^\beta(m)) \\ &= \frac{\langle \mathbf{M}_{\alpha,i}^\beta(m), \mathbf{M}_{\alpha,i+1}^\beta(m) \rangle}{\|\mathbf{M}_{\alpha,i}^\beta(m)\|_2 \cdot \|\mathbf{M}_{\alpha,i+1}^\beta(m)\|_2} \quad (4) \end{aligned}$$

where $\langle x, y \rangle$ is the inner product between the vectors x and y ; and $\|x\|_2$ is the Euclidean norm of x .

Step 4: State pattern probability: the histogram approach is applied to estimate the empirical probability density function of the cosine similarity matrix \mathbf{Sim} . If the histogram has λ bins, P_k ($k = 1, 2, \dots, \lambda$) is used to denote the probability of each bin. Obviously,

$$\sum_{n=1}^{\lambda} P_n = 1.$$

Step 5: Calculate the diversity entropy H_{DE} value, which is expressed as:

$$H_{DE}(\mathbf{S}_\alpha^\beta, m, \lambda) = -\frac{1}{\ln \lambda} \sum_{n=1}^{\lambda} P_n \ln P_n \quad (5)$$

where \mathbf{S}_α^β represents symbol-scale sequence, m indicates the embedding dimension, and λ means the number of bins. Then, the SSDE value is defined as the means of αH_{DE} ; that is,

$$H_{SSDE}^\alpha(X, \alpha, \varepsilon, m, \lambda) = \frac{1}{\alpha} \sum_{\beta=1}^{\alpha} H_{DE}(\mathbf{S}_\alpha^\beta, m, \lambda) \quad (6)$$

Lastly, repeat Steps (1)–(5) for $\alpha = 1, 2, 3, \dots, \tau$, and then obtain τH_{SSDE}^α values, which can be expressed as:

$$H_{SSDE}^{1:\tau} = \{H_{SSDE}^1; H_{SSDE}^2; \dots; H_{SSDE}^{\tau-1}; H_{SSDE}^\tau\} \quad (7)$$

For a better explanation of the SSDE method, a flowchart of SSDE is illustrated in Figure 2. In addition, the pseudo-code of SSDE is illustrated in Algorithm 1.

Parameter analysis of proposed SSDE

Before usage of proposed SSDE method, there are four parameters that need to be considered, including number of intervals λ , number of symbols ε , embedding dimension m and scale factor τ . Here, the influence of the parameters will be discussed by observing their performance in evaluating the signal with known complexity levels. In this paper, we discuss the influence of parameter selection of SSDE through distinguishing various signals generated by mixture (MIX) (p) with different p . In essence, the MIX (p) process is in nature that $N \times p$ randomly chosen points of sinusoid signal of length N are replaced by random noise, where the spectrum distribution of signals generated by MIX (p) process with different p is different so that there are different intrinsic structures.¹⁶

In the study, $p = 0, 0.1, 0.3, 0.5, 0.7$, and 1 were applied to conduct five MIX (p) processes with different complexity levels. It is noticed that the signal of $p = 0$ represents the periodic signal and a value of

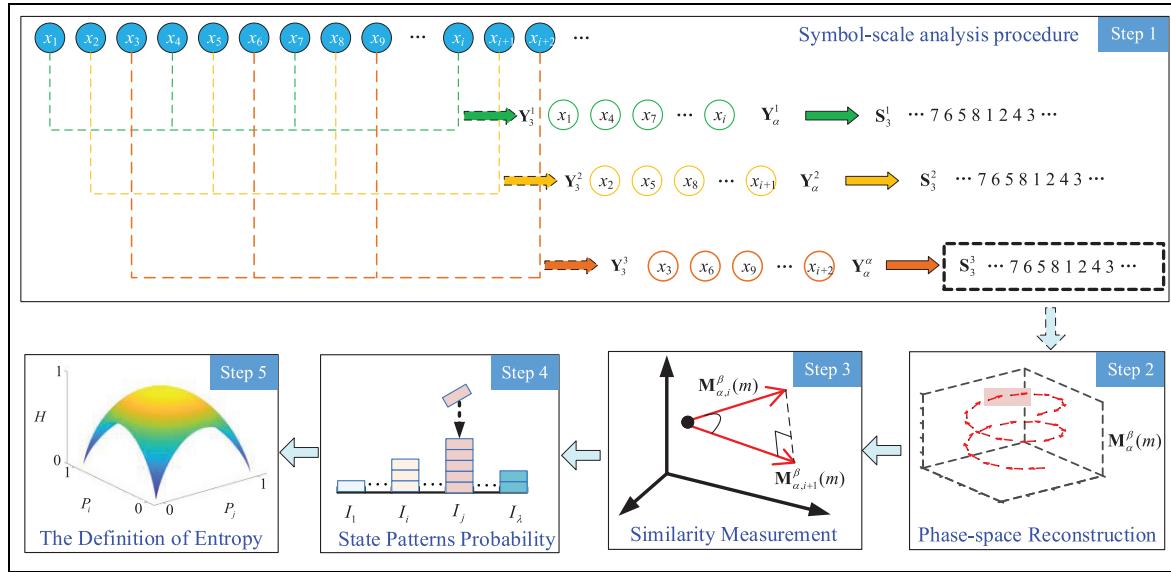


Figure 2. The illustration of symbol-scale diversity entropy with $\alpha = 3$.

Algorithm 1 Symbol-scale diversity entropy.

Input: Time series X , embedding dimension m , number of bins λ , number of symbols ε , scale factor τ

Output: SSDE value

- 1 For $\alpha = 1, 2, \dots, \tau$ do
- 2 Conduct time-delay analysis and obtain α new series according to Equation (1).
- 3 Encode the α sequences into the symbol time series of ε symbols.
- 4 Compute diversity entropy of each symbol-scale sequence S_α^β using Equation (4) and (5).
- 5 Compute the α -th H_{SSDE}^α value according to Equation (6).
- 6 Augment the data $H_{\text{SSDE}}^{1:\alpha} = \{H_{\text{SSDE}}^{1:\alpha-1}; H_{\text{SSDE}}^\alpha\}$.
- 7 End for

$p = 1$ represents pure white Gaussian noise. Here, one hundred independent realizations for each MIX (p) process were performed, and then the average value was computed to eliminate random factors. As mentioned above, their complexity levels should correspond to their irregularity levels. Therefore, we gradually reduce the value of p from MIX (1) to MIX (0), and entropy value decreases gradually. The results are illustrated in the Figure 3.

First, the influence of different m was examined. The parameter m of SSDE means the vector length of phase space. In the case of large dimension m , a longer time series is required. However, a larger m allows for a more detailed phase-space reconstruction of the dynamic process. Here, m was set to seven different values from two to eight with a step size of one. The mean and standard deviation curves of SSDE of five series generated by MIX (p) processes were computed and illustrated in Figure 3(a). As can be seen from Figure 3(a), when m is large (such as 7), the standard

deviation is also significant. Hence, generally, we set m as a small integer of 3 or 4 and m is fixed to 4 in the paper.

Then, the dependence of SSDE on the parameter ε is shown, which represents the number of symbols. Here, ε is set at eight different values from 5 to 40 with equal steps, and the analysis results are depicted in Figure 3(b). It can be found from Figure 3(b) that the SSDE value changes only slightly. As the number of symbols increases, the robustness to noise can be reduced, and the computation efficiency may decrease.⁴⁰ Meanwhile, as can be seen from Figure 3(b), a larger ε will result in larger error bar of SSDE curves. However, the fewer the symbols, the less fault information extracted by SSDE. Therefore, ε is recommended to set to 15–30, and we chose $\varepsilon = 20$ in the study.

The parameter λ represents the number of intervals. We set λ at eight different values chosen from 10 to 80 with equal steps. Figure 3(c) shows the dependence of SSDE on parameter λ . As can be seen from Figure 3(c),

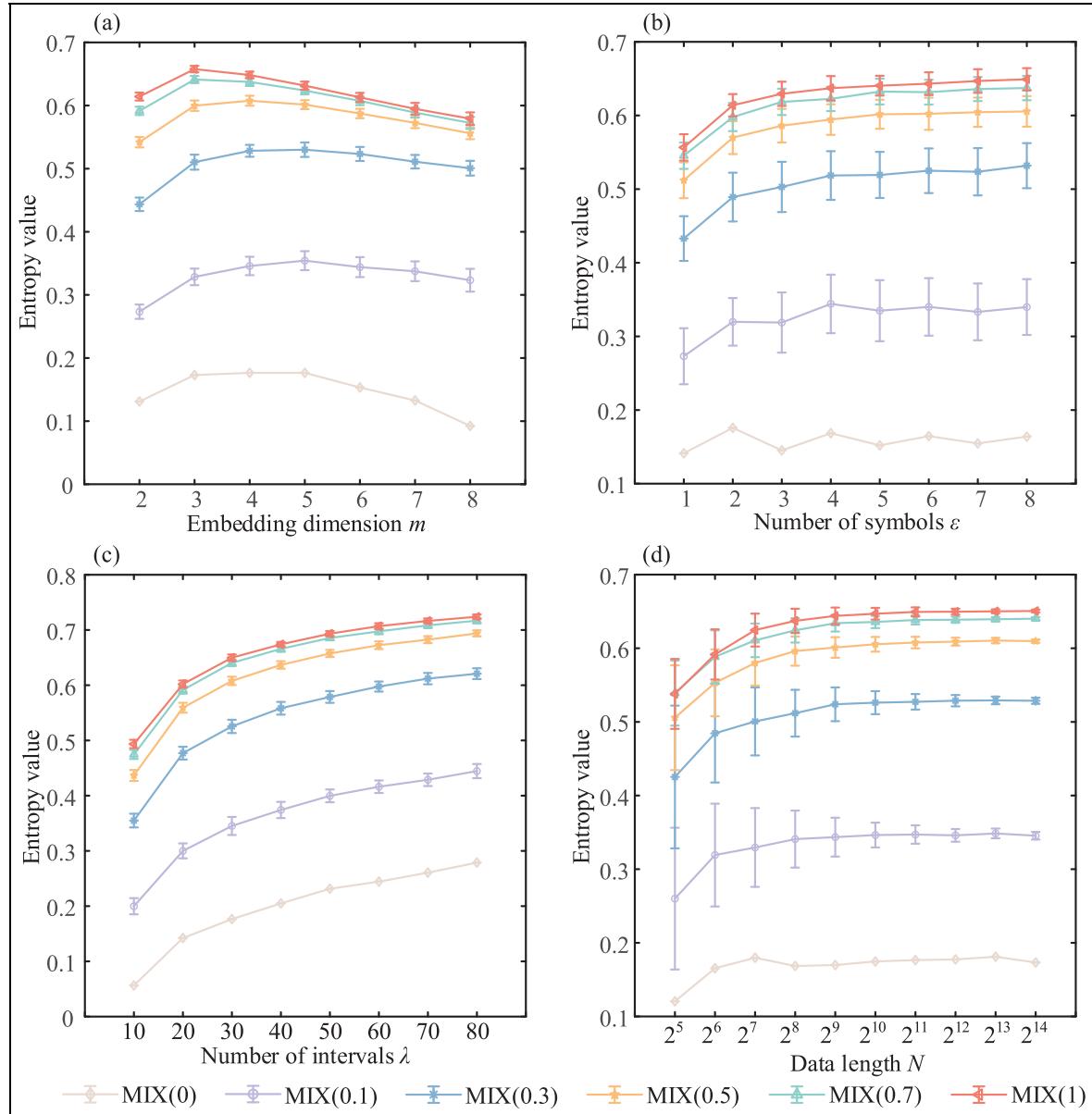


Figure 3. SSDE values of five series are shown as functions of (a) embedding dimension m , (b) number of symbols ε , (c) number of intervals λ and (d) data length N . (Error bar indicates the standard deviation of 100 independent realizations.) SSDE: symbol-scale diversity entropy.

SSDE curves are all in accordance with the complexity arrangement of different p values for MIX (p) process so that SSDE is much free to select parameters due to its strong relative consistency. In this study, interval number λ is set to 30 according to Wang et al.²¹ As for the scale factor τ , it represents the dimension of entropy features. A smaller scale number will fail to capture the relevant fault features adequately, leading to ineffective fault diagnosis. On the other hand, a large scale number can result in dimensionality issues and hinder the extraction of discriminative fault information, leading to

subpar recognition results. Furthermore, a larger scale number can also increase the computation time required for feature extraction, making the process less efficient. Therefore, here, it is recommended to set $\tau = 15$ to 30, and it has been set as 20 in this paper according to Wang et al.²¹

In addition, to evaluate the sensitivity of SSDE to data length, we assessed SSDE in the five MIX (p) process as a function of data length N , where the data length was set at 10 different values from 32 to 16,384. We chose $\varepsilon = 20$, $\lambda = 30$, $m = 4$ in all calculations of

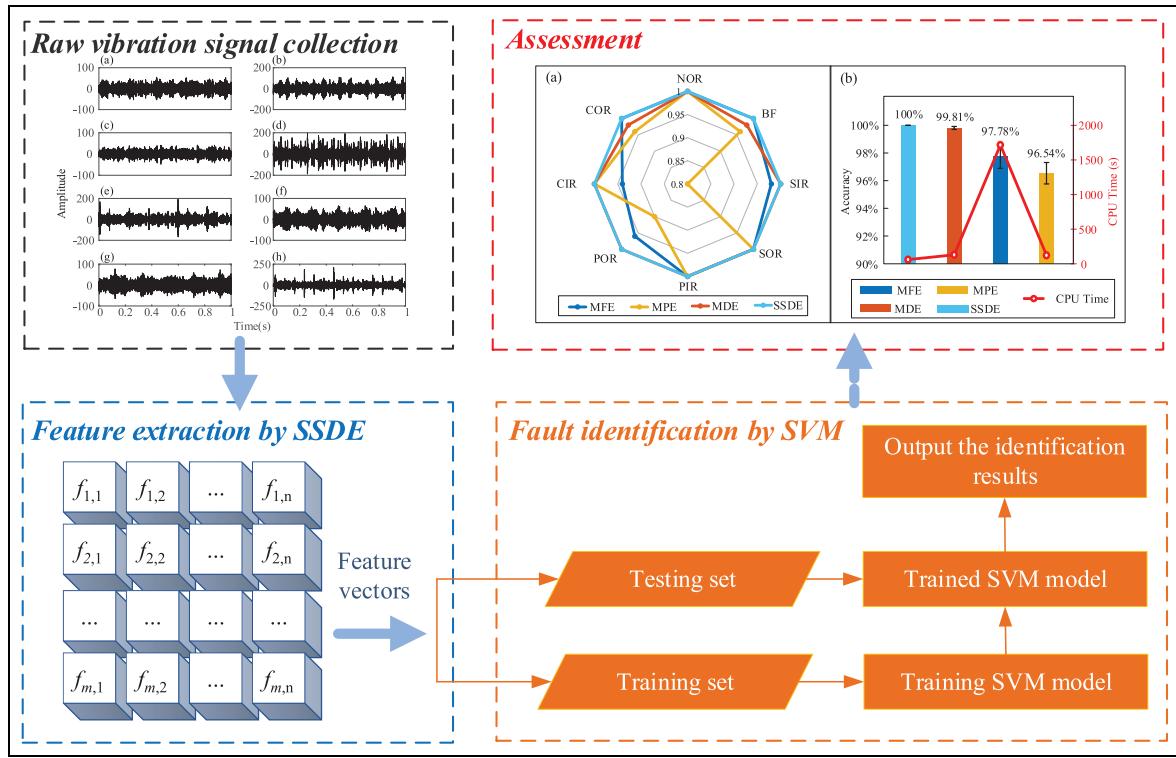


Figure 4. The scheme of proposed model for fault diagnosis.

SSDE, and the results are illustrated in Figure 3(d). It can be seen from Figure 3(d) that the error bar of SSDE curves decreases with the increase of N . However, when N falls below a certain value, such as N equals 32 and 64, it is difficult to characterize the complexity of the five signals correctly. Here, we recommend setting $N > 128$ and it is set as 2048 in this paper.

Strategy for the SSDE-based fault diagnosis model

In this section, a new diagnostic approach for rotating machinery is proposed by combining SSDE and SVM.⁴¹ The main steps of the SSDE–SVM approach are as follows. Also, the corresponding flowchart is shown in Figure 4.

Step 1: Acquisition of vibration signals of rotating machinery under different health conditions.

Step 2: Construction of feature vectors for different state samples by the SSDE method.

Step 3: Division of the training and testing sets for the obtained SSDE-based features.

Step 4: SVM model training and classification model construction.

Step 5: Identification of various fault patterns using the trained SVM model.

It is noted that after feature extraction using SSDE, there would be τ entropy features as the feature vector for the SVM classifier. In this paper, LIBSVM package (LIBSVM is an open source machine learning library, developed at the National Taiwan University, Taipei, Taiwan) is utilized to implement SVM.⁴¹ Gaussian kernel is adopted, the tradeoff parameter C is set to 1, and the kernel parameter γ is set to 0.5 throughout this work. In this paper, to verify the feature performance of the proposed method, existing entropy methods, including MPE, MFE, and MDE are also utilized to extract the weak fault characteristics. For comparison, we set $\tau = 20$ in all multiscale-based entropy methods for MDE, MFE, and MPE; $m = 2$, $r = 0.15$ in MFE; $m = 5$ in MPE; $m = 4$, $\lambda = 30$ in MDE according to Yan et al.¹⁵ and Wang et al.⁴² In summary, the parameters setting of the four entropy-based methods are shown in Table 1.

Verification using simulated signals

Simulation not only helps to evaluate fault conditions that are difficult to realize in real scenarios, but it also makes it simple to simulate different levels of failure. In addition, because the simulation does not cause false validation caused by the actual assembly and other environmental noise, it is an effective research method

Table 1. Parameter setting of entropy-based methods.

Method	Scale factor τ	Embedding dimension m	Tolerance threshold r	Number of intervals λ	Number of symbols ε
MFE	20	2	0.15	—	—
MPE	20	5	—	—	—
MDE	20	4	—	30	—
SSDE	20	4	—	30	20

MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy.

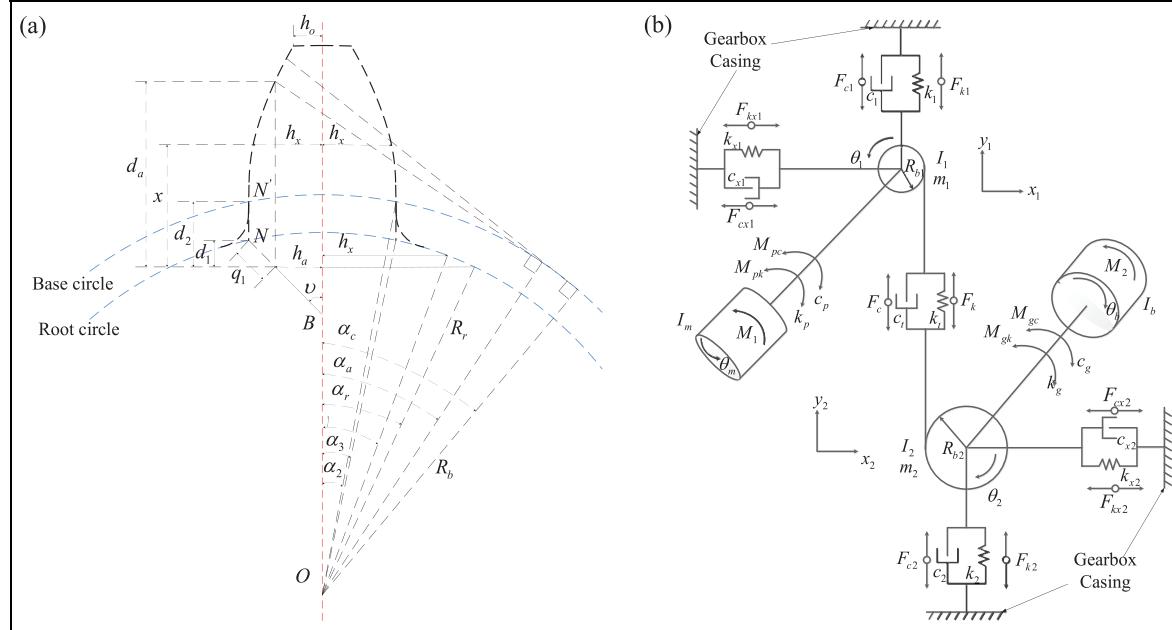


Figure 5. (a) Beam model of an external gear tooth and (b) dynamic model of one-stage gearbox.⁴³

with generality. Therefore, in this section, two simulated case studies are conducted to assess the feature extraction capability of the SSDE method, including simulation of different crack severity levels in gearbox and simulation of different fault types for bearing.

Simulated gearbox signals

Signal generation with simulated gearbox. To evaluate the performance of SSDE in fault feature extraction, signals corresponding to gear fault are simulated. This paper simulates the fixed-shaft gear system where the root circle is smaller than the base circle. First, the potential energy approach is utilized to solve the mesh stiffness of the fixed-shaft gear system.⁴³ Here, a cantilever beam model is utilized to simulate the external gear tooth, as shown in Figure 5(a). For details on the model and the calculation of mesh stiffness, readers can refer to the study of Liang et al.⁴³ It is worth noting that we consider the nonlinearity of the tooth contact stiffness and then utilize time-varying mesh stiffness as

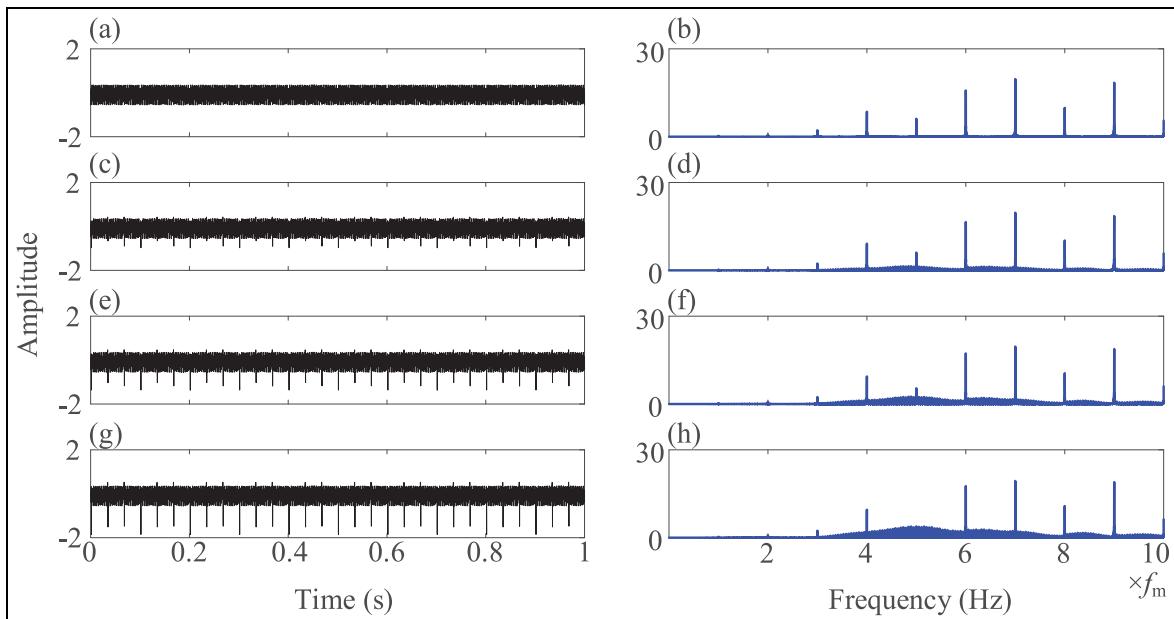
an input into the gear dynamic model. Second, a one-stage gear dynamic model is utilized to simulate the vibration response of three kinds of local tooth faults, as illustrated in Figure 5(b). Here, the dynamic gear system model considers the time-varying mesh stiffness and lateral and torsional vibrations.^{44,45}

According to the motion equations finally established in Appendix 1 of this article, the response of the gear can be solved. Table 2 provides the parameter values used for simulating the gearbox. In this study, four health conditions are considered in this paper including perfect condition and three types of crack severity levels (slight crack, moderate crack and severe crack). The time wave of the simulated gearbox signals is plotted in Figure 6(a), (c), (e), and (g), respectively. Furthermore, the corresponding spectrums are illustrated in Figure 6(b), (d), (f), and (h), respectively.

Analysis results and discussion. In this work, the simulated signals are added with noise to simulate the real working condition with different SNR levels. Here, the SNR

Table 2. Physical parameters of a spur gearbox.

Parameter	Pinion (driving)	Gear (driven)
Number of teeth	$N_1 = 19$	$N_2 = 31$
Module (mm)	$m = 3.2$	$m = 3.2$
Pressure angle	$\alpha = 20^\circ$	$\alpha = 20^\circ$
Mass (kg)	$m_1 = 0.7$	$m_2 = 1.822$
Mass moment (kg m^2)	$I_1 = 4.25 \times 10^{-4}$	$I_2 = 2.72 \times 10^{-3}$
Base circle radius (mm)	$R_{b1} = 28.3$	$R_{b2} = 46.2$
Root circle radius (mm)	$R_{r1} = 26.2$	$R_{r2} = 45.2$
Face width (m)	$L = 0.0381$	$L = 0.0381$
Young's modulus (GPa)	$E = 206.8$	$E = 206.8$
Poisson's ratio	$\nu = 0.3$	$\nu = 0.3$
Bearing stiffness (N/m)	$k_1 = k_2 = 5.0 \times 10^8$	
Bearing damping (kg/s)	$c_1 = c_2 = 4.0 \times 10^5$	
Torsional stiffness (Nm/rad)	$k_p = k_g = 4.0 \times 10^7$	
Torsional damping (Nms/rad)	$c_p = c_g = 3.0 \times 10^4$	
Rotational frequency (Hz)	$f_s = 30$	

**Figure 6.** Time-domain waveform and spectrum of the simulated gearbox signals: (a) waveform of perfect condition, (b) spectrum of perfect condition, (c) waveform of slight crack, (d) spectrum of slight crack, (e) waveform of moderate crack, (f) spectrum of moderate crack, (g) waveform of severe crack, and (h) spectrum of severe crack. (f_m is gear mesh frequency).

value ranges from 5 to 20 dB, and the stride is 5 dB. The definition of SNR can be expressed as follows:

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \left(\frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \quad (8)$$

where P_{signal} and P_{noise} represent the power of signal and noise, respectively.

The time waves of the simulated perfect condition and faults under different SNR levels are illustrated in Figure 7, respectively.

After acquiring the simulated signals, a total of 400 samples can be obtained, with one hundred samples for each health condition using non-overlapped cutting. Out of these samples, 50% are utilized for training, and the remaining samples are used for performance testing. Four entropy-based methods, namely MFE, MPE, MDE, and SSDE, are applied to extract fault features from the signals.

Figure 8 illustrates the fault diagnosis accuracies of these four methods under different SNR levels for the

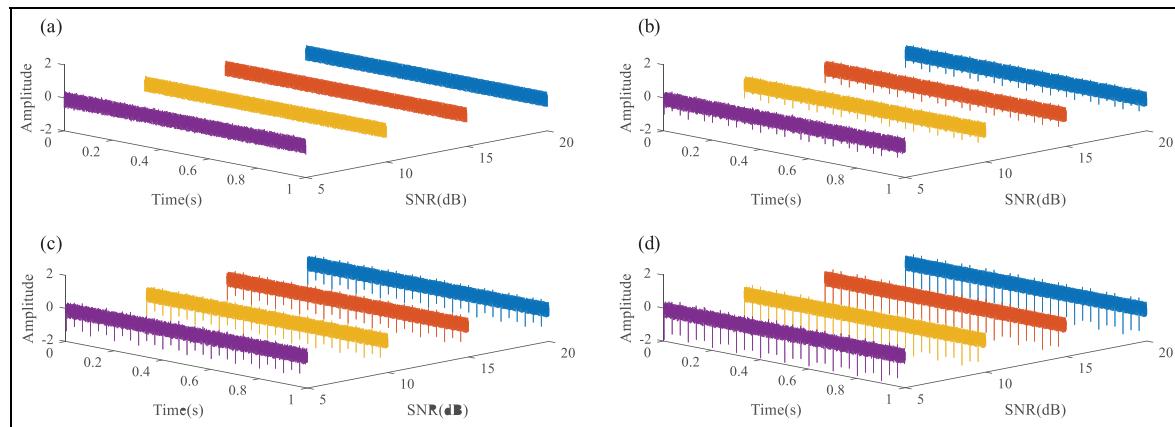


Figure 7. The time-domain signals of the simulated gear conditions under different SNR values: (a) time-domain signals of perfect condition, (b) time-domain signals of slight crack, (c) time-domain signals of moderate crack, and (d) time-domain signals of severe crack.

SNR: signal-to-noise ratio.

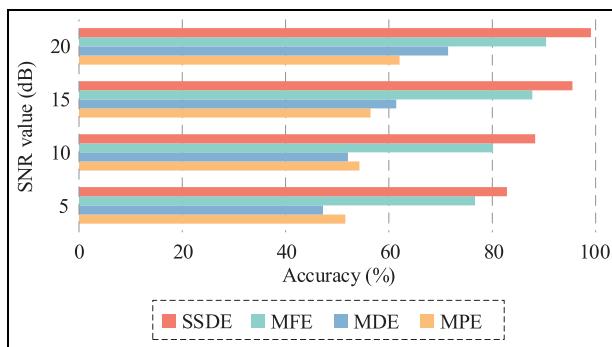


Figure 8. The recognition accuracies under different SNR values for the simulated gear fault.

SNR: signal-to-noise ratio.

simulated gear system. It can be observed that as the noise level increases, the diagnostic accuracies of all four methods show a decreasing trend. Importantly, our proposed SSDE method consistently achieves the highest accuracy among the four methods, demonstrating superior diagnostic performance. For instance, at an SNR of 20 dB, both SSDE and MFE methods achieve accuracies exceeding 90%. Even at a lower SNR of 10 dB with stronger noise, the accuracies of MPE, MDE, and MFE drop below 80%, while the proposed SSDE method maintains an accuracy of 88%. Furthermore, the accuracy of SSDE remains above 80% even in a strong-noise environment with an SNR of 5 dB. These comparison results clearly indicate that the SSDE method exhibits superior feature extraction ability compared to the MFE, MPE, and MDE methods, consistently achieving higher diagnostic accuracy in various SNR scenarios.

Table 3. Parameter setting of rolling bearing simulation.

Parameter	Value
Diameter of pitch circle (mm)	35.5
Diameter of rolling element (mm)	6.5
Rotational speed (rpm)	3000
Number of rolling elements	12
Sampling frequency (Hz)	10,240
Natural frequency of bearing (Hz)	4000
Angle of contact (°)	0

Simulated bearing signals

Signal generation with simulated bearing. The second simulated case is a rolling bearing, and the specific bearing model number is N205. The simulated rotational speed is set to 3000 rpm, and the detailed bearing-related parameter settings are listed in Table 3.

The diagrammatic sketches of bearing models are illustrated in Figure 9. For simplicity, it is assumed in the model that the sensor is at the maximum load density,⁴⁶ as shown in Figure 9(a). For the outer ring failure in Figure 9(a), since the relative positional relationship between the sensor and the defect does not change with time, the impulsive force is an ideal force, which does not consider the changes of the contact position. For the inner ring failure shown in Figure 9(b), the contact position of the rolling element in contact with the local defect changes as the local defect rotates with the inner ring. For the roller failure shown in Figure 9(c), when the rolling element rotates with the shaft, the local defect on the rolling element makes continuous contact with the outer and inner rings to generate impulses.

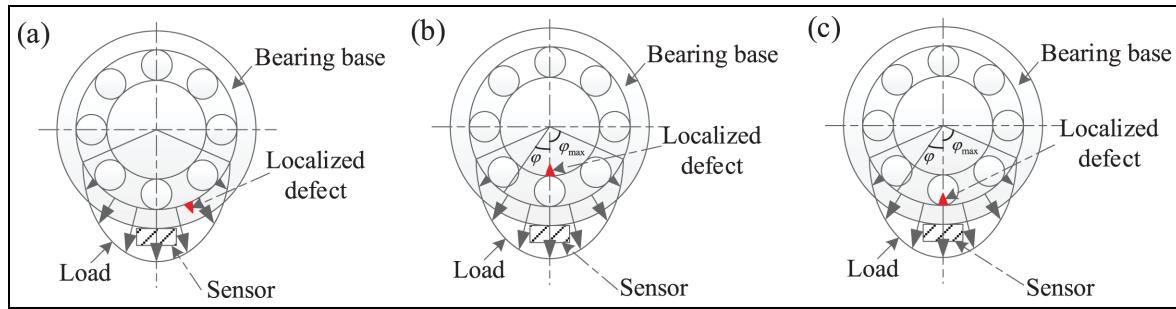


Figure 9. The schematic diagram of the simulated bearing model under different fault types: (a) outer ring failure, (b) inner ring failure, and (c) rolling element fault.

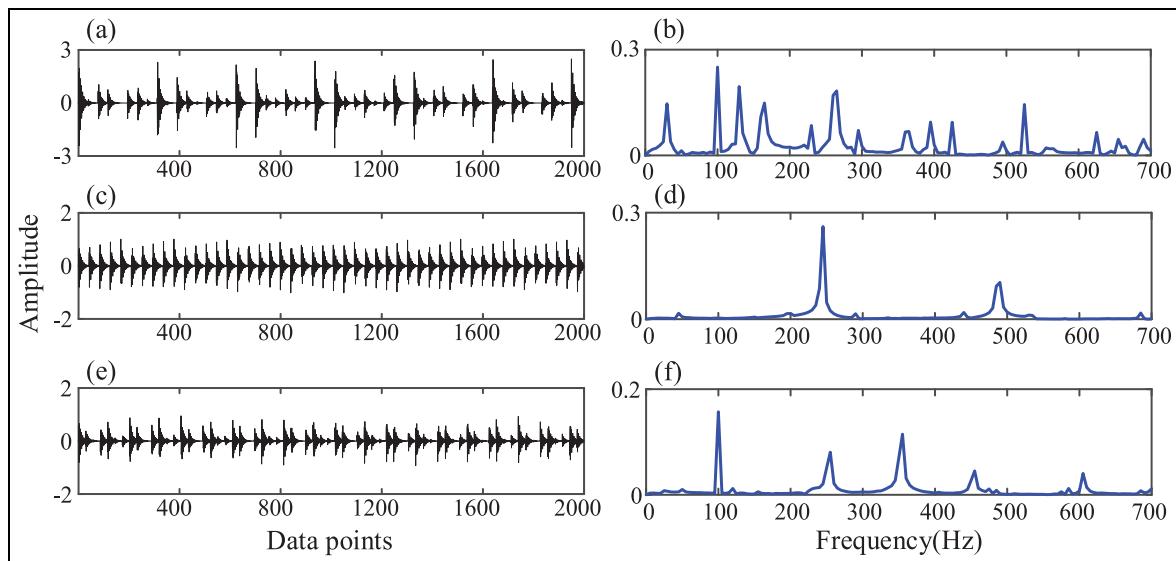


Figure 10. Time-domain signals and envelope spectrum of the simulated faults: (a) waveform of the rolling element failure, (b) envelope spectrum of the rolling element failure, (c) waveform of the outer ring failure, (d) envelope spectrum of the outer ring failure, (e) waveform of the inner ring failure, and (f) envelope spectrum of the inner ring failure.

Similarly, according to the equations established in the Appendix 2 of this article, the response signals of the simulated bearing can be obtained. For details on the bearing model, readers can refer to Wang et al.⁴² and McFadden and Smith.⁴⁶

The time domain signals of the simulated outer ring failure, inner ring failure, and rolling element failure are plotted in Figure 10(a), (c), and (e), respectively. Meanwhile, the corresponding envelope spectrum is illustrated in Figure 10(b), (d), and (f), respectively.

Analysis results and discussion. In this study, the simulated fault is added with noise with different SNR values, where the SNR value ranges from -5 to 20 dB, and the stride is 5 dB. The time waves of the simulated

three faults with different SNR values are illustrated in Figure 11, respectively.

For each failure type, a sliding window with 2048 points is used to extract a slice from the original signal as one sample. Each failure type contains 100 samples. Thus, there are a total of 300 samples. Next, the dataset is divided into a training dataset and a testing dataset, with 50% of the samples allocated for training and the remaining samples designated for performance testing. Afterward, fault features are extracted from the datasets using four different methods: MFE, MPE, MDE, and the proposed SSDE. The fault diagnosis accuracies of four entropy methods under different SNR values is demonstrated in Figure 12.

It can be observed from Figure 12 that the diagnostic accuracies of the four methods show a decreasing

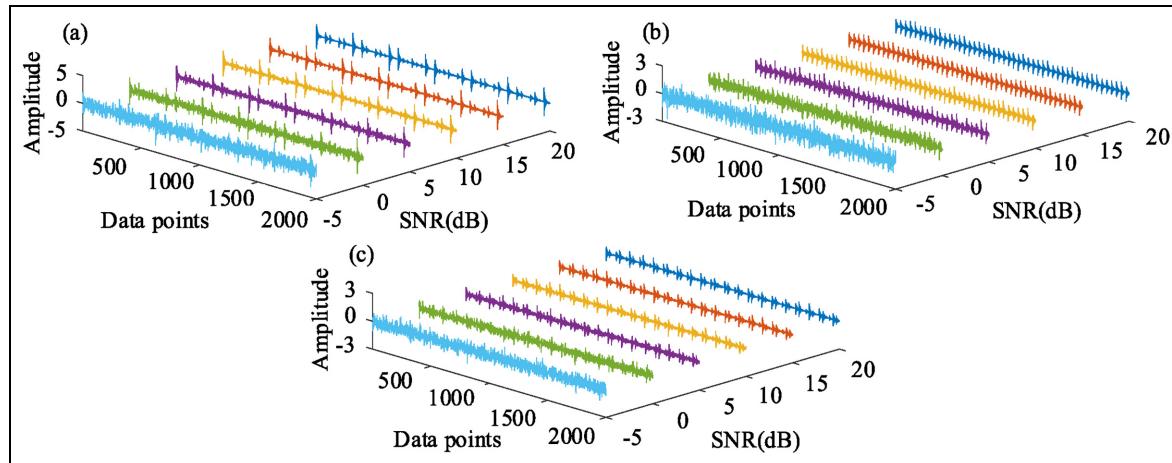


Figure 11. The time-domain signals of the simulated faults with different SNR values: (a) time-domain signals of the simulated rolling element failure, (b) time-domain signals of the simulated outer ring failure, and (c) time-domain signals of the simulated inner ring failure.

SNR: signal-to-noise ratio.

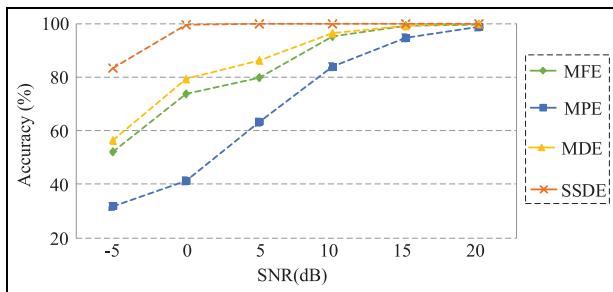


Figure 12. The recognition accuracies under different SNR values for the simulated bearing fault.

SNR: signal-to-noise ratio.

trend as the noise enhancement. In addition, our proposed SSDE method always obtains the highest accuracy among the four methods with the best diagnostic performance. For instance, when $\text{SNR} \geq 15 \text{ dB}$ with a

low-noise environment, the accuracies of all methods exceed 95%, and when $\text{SNR} = 0 \text{ dB}$ with stronger noises, the accuracies of MDE, MFE, and MPE are all below 80%, while that of the proposed SSDE still remains at 99%. Also, the accuracy of SSDE still remains higher than 80% even under strong-noise environment with $\text{SNR} = -5 \text{ dB}$. The comparison results suggest that the proposed SSDE is more robust to noise than the three state-of-the-art entropy methods.

In addition, to testify to the necessity of scale analysis, scale factors vary from 1 to 20, and the remaining parameters of the method are unchanged. Table 4 illustrates the diagnostic results. The results of SSDE are illustrated in Table 4, which compares identification accuracy trained with different scale factor values under different levels of noise. It can be observed that when SNR value is high, all models achieved relatively high accuracy, even on a small scale factor. For example, when $\text{SNR} \geq 15 \text{ dB}$, the accuracies with $\tau=5$ are

Table 4. The mean accuracies of SSDE with different scale factors and SNR values.

Scale factor	SNR (dB)					
	-5	0	5	10	15	20
1	45.23%	64.13%	71.63%	67.83%	67.87%	57.73%
2	47.43%	64.57%	72.37%	68.33%	71.73%	72.47%
5	48.67%	70.27%	75.97%	83.87%	93.17%	96.77%
10	55.43%	81.07%	98.77%	100%	100%	100%
15	71.07%	98.73%	100%	100%	100%	100%
20	83.47%	99.63%	100%	100%	100%	100%
25	90.17%	100%	100%	100%	100%	100%
30	92.63%	100%	100%	100%	100%	100%

SNR: Signal-to-noise ratio; SSDE: Symbol-scale diversity entropy.

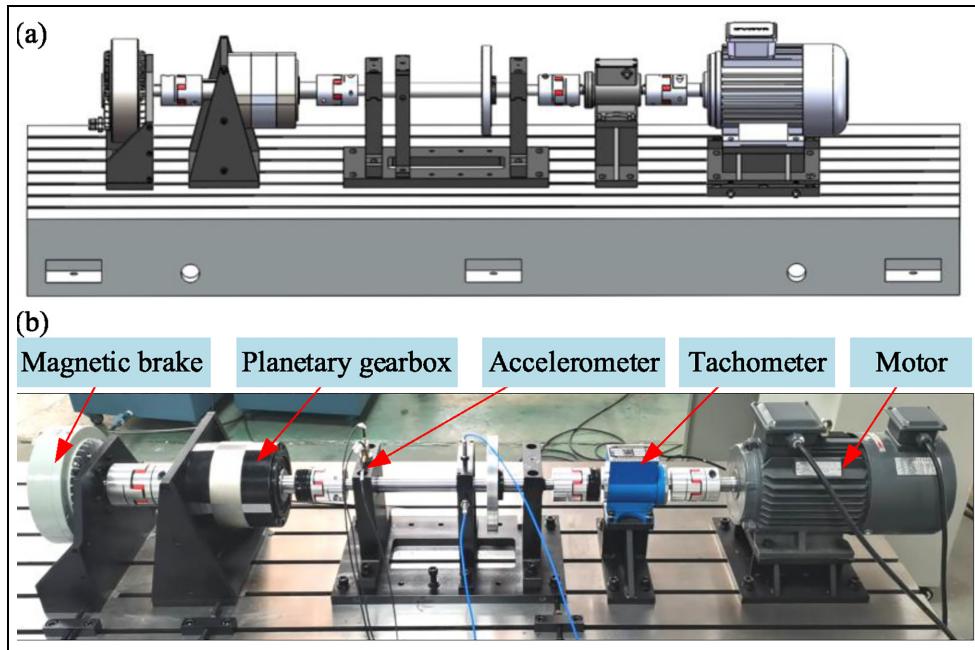


Figure 13. The planetary gearbox system of Case I: (a) the schematic layout of the test rig and (b) the real test rig.

above 90%. In contrast, a larger scale factor is needed to achieve higher accuracy when SNR value is low. For instance, a model with a scale factor of 15 can achieve 100% accuracy when SNR equals 5 dB. However, when SNR is equal to -5 dB, the model with a scale factor of 25 can obtain 90% accuracy, which needs a larger scale factor.

Experimental case studies

The simulation simulates the fault signals in an ideal environment. Still, there is a large amount of interference in the practical working environment. Hence, the effectiveness of the proposed method needs to be further verified in the real environment, which can be achieved in the experimental case study. Therefore, in this section, two different experimental case studies, including planetary gearbox and rolling bearing, are conducted to evaluate the benefits of SSDE for feature extraction and fault diagnosis of rotating machinery.

Case study I: fault diagnosis of planetary gearbox

Description of planetary gearbox test and gear fault. The first experimental case study was carried out on a planetary gearbox system produced by WuXi HouDe Automation Meter, as shown in Figure 13, which consists of a motor, planetary gearbox, tachometer, and magnetic brake. In order to acquire vibration signals,

Table 5. The parameters of test planetary gearbox.

Parameter	Value
Rotating speed (rpm)	1200
Load (Nm)	5
Sample frequency (Hz)	16,384
Number of teeth for sun gear	21
Number of teeth for planet gear	31
Number of teeth for ring gear	84

an accelerometer was installed on the top of the bearing casing, and one 5 Nm load was used to simulate the real application scenario. Here, the sampling frequency was set as 16,384 Hz, and the rotation speed of the motor was set as 1200 rpm. In summary, Table 5 lists the parameters of the planetary gearbox system.

In the first experimental case, eight health conditions of planetary gearbox were considered, including normal condition (NOR), planet gear with a missing tooth (PGMT), planet gear with a broken tooth (PGBT), planet gear with a spalling tooth (PGST), fracture of bearing cage (FBC), sun gear with a spalling tooth (SGST), sun gear with a broken tooth (SGBT), sun gear with a cracked tooth (SGCT), as shown in Figure 14. It is worth noting that each type of fault signals is sliced into 100 sub-signals using non-overlapped slicing method. Each sub-signal contains 2048 sample points, and there are 800 samples in total for this case study. The detail description of experimental data and sample

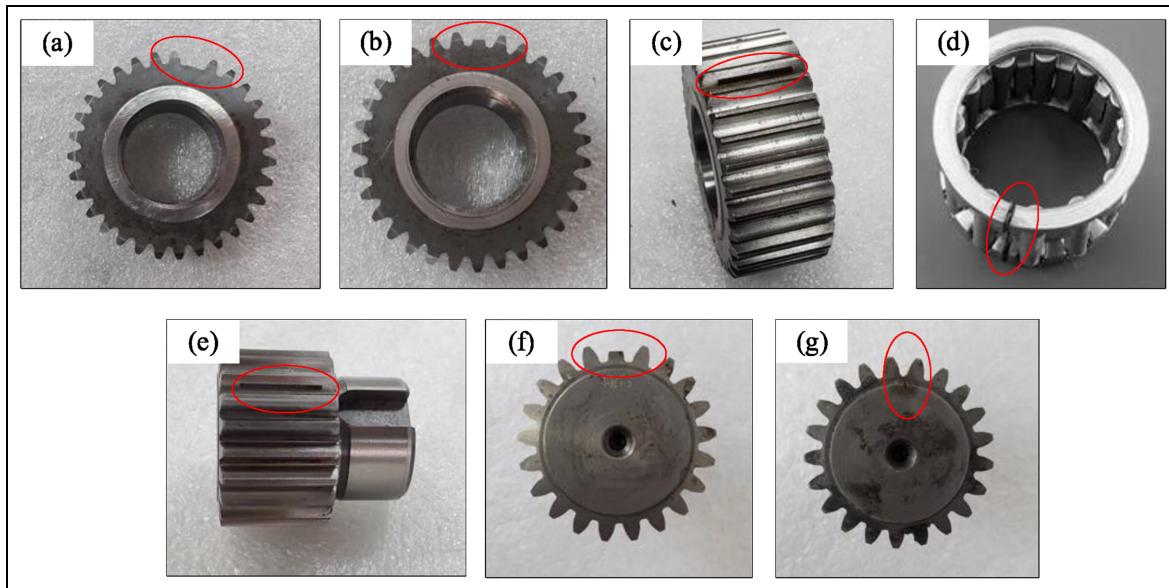


Figure 14. Planetary gearbox with artificial defect: (a) PGMT, (b) PGBT, (c) PGST, (d) FBC, (e) SGST, (f) SGBT, and (g) SGCT. PGMT: planet gear with a missing tooth; PGBT: planet gear with a broken tooth; PGST: planet gear with a spalling tooth; FBC: fracture of bearing cage; SGST: sun gear with a spalling tooth; SGBT: sun gear with a broken tooth; SGCT: sun gear with a cracked tooth.

number can be seen from Table 6. Figure 15 illustrates time-domain vibration signals of planetary gearbox under different fault categories and health condition.

Compared with state-of-the-art entropy methods. In the first experimental case, for comparison, feature extraction is conducted by SSDE, MDE, MFE, and MPE, respectively. Meanwhile, 20 trials are performed to decrease random effects for each method. The performance comparison results of four entropy methods are illustrated in Figure 16.

Figure 16(a) illustrates the diagnosis recognition rates for all health conditions under one of the trials. As can be observed from the radar diagram in Figure 16(a), the curve of SSDE is farthest from the center. Furthermore, as can be seen from Figure 16(b), the proposed SSDE obtains an average classification accuracy of 98.91% compared to the other methods, with the highest diagnostic result. The results indicate that the proposed method achieved the best diagnostic performance and displayed excellent stability compared to the other entropy methods. As for the calculation time in Figure 16(b), it also can be observed that SSDE performs best in calculation efficiency. The phenomenon indicates that SSDE performs best in feature extraction with the highest diagnostic accuracy and has high calculation efficiency, which can meet the online detection requirements and provide an attractive way for entropy-based feature extraction.

Evaluation metric and visualized interpretation. To visually analyze the feature space, a visual distribution representation of all entropy features is illustrated in Figure 17. Here, *t*-distributed stochastic neighbor embedding (*t*-SNE) algorithm⁴⁷ is applied to map entropy features onto a two-dimensional feature space. In Figure 17, the clustering effect represents the strength of the feature extraction capability: the lower the intra-class distance among samples for the same class and the larger the inter-class distance among different classes, the more powerful the feature extraction capability of the approach. From Figure 17(a), it can be observed that SSDE features of the eight types of health conditions can be clearly separated, and each cluster has a clear boundary. On the contrary, the feature distributions of the other three entropy methods after dimensionality reduction are mixed, where the FBC, SGCT, SGST, and SGBT conditions cannot be distinguished.

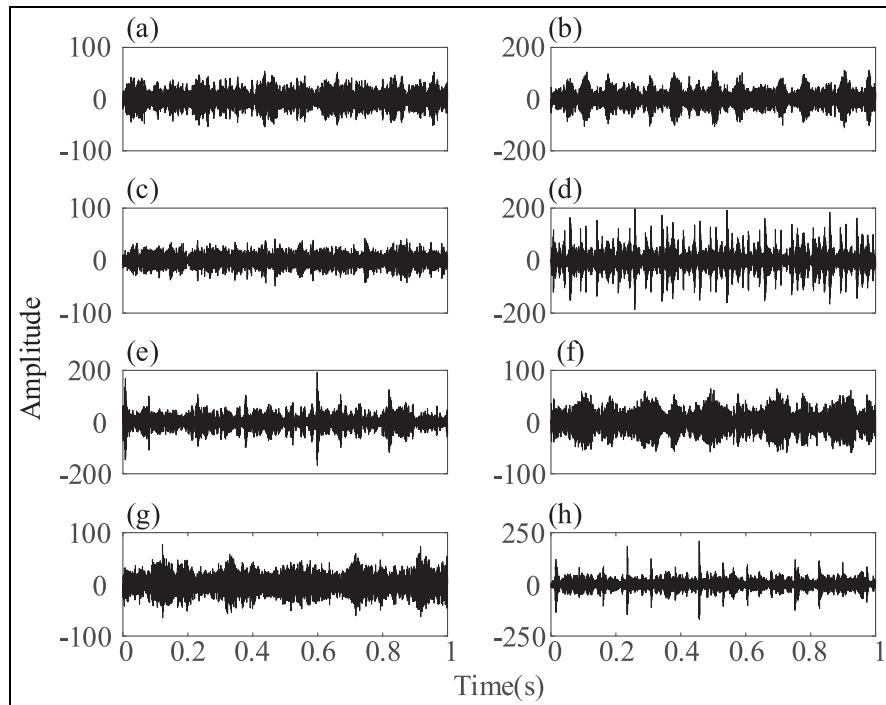
Meanwhile, in this study, the separable criterion based on distance is applied to compare the useful fault information for four entropy-based approaches fairly. The separable criterion, Fisher score J , can be expressed by Equation (9).⁴⁸

$$J = \frac{\text{tr}(\tilde{\mathbf{S}}_b)}{\text{tr}(\tilde{\mathbf{S}}_t)} \quad (9)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix, $\tilde{\mathbf{S}}_b$ indicates between-class scatter matrix, $\tilde{\mathbf{S}}_t$ represents the total scatter matrix, which can be expressed as⁴⁹:

Table 6. The introduction of experimental data for case study I.

Label	Fault location	Fault type	Fault size (mm)	Training sample number	Testing sample number
C1	—	Normal	—	50	50
C2	Planet gear	Missing fault	—	50	50
C3	Planet gear	Broken fault	1	50	50
C4	Planet gear	Spalling fault	15*0.75	50	50
C5	Bearing cage	Fracture fault	1	50	50
C6	Sun gear	Spalling fault	15*0.75	50	50
C7	Sun gear	Broken fault	0.2	50	50
C8	Sun gear	Crack fault	3	50	50

**Figure 15.** The time-domain signals of eight health conditions for case study I: (a) NOR, (b) SGST, (c) SGCT, (d) SGBT, (e) PGMT, (f) PGBT, (g) FBC, and (h) PGST.

NOR: normal condition; PGMT: planet gear with a missing tooth; PGBT: planet gear with a broken tooth; PGST: planet gear with a spalling tooth; FBC: fracture of bearing cage; SGST: sun gear with a spalling tooth; SGBT: sun gear with a broken tooth; SGCT: sun gear with a cracked tooth.

$$\begin{aligned} \tilde{\mathbf{S}}_b &= \sum_{k=1}^c n_k (\tilde{\mathbf{\mu}}_k - \tilde{\mathbf{\mu}})(\tilde{\mathbf{\mu}}_k - \tilde{\mathbf{\mu}})^T \\ \tilde{\mathbf{S}}_t &= \sum_{k=1}^n (\mathbf{z}_k - \tilde{\mathbf{\mu}})(\mathbf{z}_k - \tilde{\mathbf{\mu}})^T \end{aligned} \quad (10)$$

where $\tilde{\mathbf{\mu}}_k$ and n_k represent the mean vector and size of the k -th class, respectively, $\tilde{\mathbf{\mu}} = \sum_{k=1}^c n_k \tilde{\mathbf{\mu}}_k$ is the overall mean vector. In short, the better the clustering effect, the smaller the between-class distance, and the greater J the value.

To ensure fairness, we projected the features from 20-dimension to 2-dimension using t-SNE algorithm and calculated the value of its Fisher score J . The results are illustrated in Table 7 and show that the J value obtained by the SSDE algorithm is higher than the other three algorithms, indicating that the proposed algorithm is helpful for clustering. As a result, both clustering performance and Fisher score for the SSDE method achieve excellent results compared with other entropy-based methods, further validating the fault information extraction ability of SSDE.

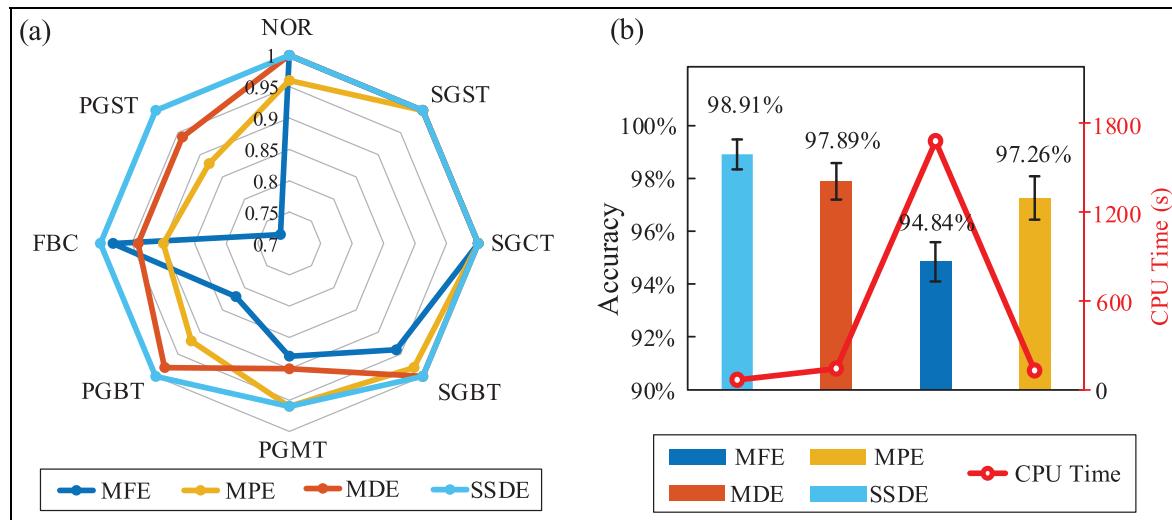


Figure 16. The diagnosis results of four entropy-based methods for the planetary gearbox system. (a) Radar chart of recognition accuracy for each health condition and (b) performance comparison of four entropy algorithms.

Table 7. The Fisher score of two experimental case studies.

Fisher score	SSDE	MDE	MFE	MPE
Case study I	0.9960	0.9946	0.9864	0.9949
Case study II	0.9976	0.9085	0.5830	0.5867

MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy.

Case study II: fault diagnosis of rolling bearing

Description of rolling bearings test and bearing fault. The second experimental case was conducted on the test rig manufactured by WuXi HouDe Automation Meter, as illustrated in Figure 18, which mainly includes a motor, tachometer, rolling bearing, and magnetic brake. The load was simulated to be generated by magnetic damping by the magnetic damping and the load was set as 5 Nm. A vertical accelerometer mounted on the bearing case was used to collect the vibration signals, and the sampling frequency was set as 10,240 Hz. Here, the motor speed was set to 3000 rpm. The test object is a cylindrical roller bearing and the specific model number is N205.

In this experiment case, different bearing failure types were implemented by replacing the test bearings, and eventually, eight health conditions, in total, were designed. The fault bearings are shown in Figure 19, including roller fault (RF), spalling of inner race (SIR), pitting of inner race (PIR), crack of inner race (CIR), pitting of outer race (POR), spalling of outer race (SOR), and crack of outer race (COR). In addition, each type of fault signal is sliced into 100 sub-signals using non-overlapped slicing, and each sub-

signal contains 2048 sample points. The time domain signals and the corresponding frequency spectrum of the collected vibration signals are illustrated in Figure 20. In addition, the experimental data setup of Case Study II is described in Table 8.

Compared with state-of-the-art entropy methods. Like Case study I, four entropy methods are conducted for comparisons, and 20 trials are performed on each method to decrease random effects. Figure 21 illustrates the comparison results and computation time. Similarly, according to the radar chart in Figure 21(a), it can be observed that the light blue curve of the SSDE approach is farthest from the center, and the SSDE method obtains the highest diagnosis result of 100% for all health conditions. In addition, from Figure 21(b), it can be observed that the SSDE approach not only performs best with the highest diagnosis result of 100%, but also takes the shortest time cost. Here, the calculation efficiency of SSDE is almost 30 times faster than that of MFE, and twice faster than that of MDE and MPE. The comparison results of Case study II further highlight the merits of the proposed SSDE method in both fault feature extraction and calculation

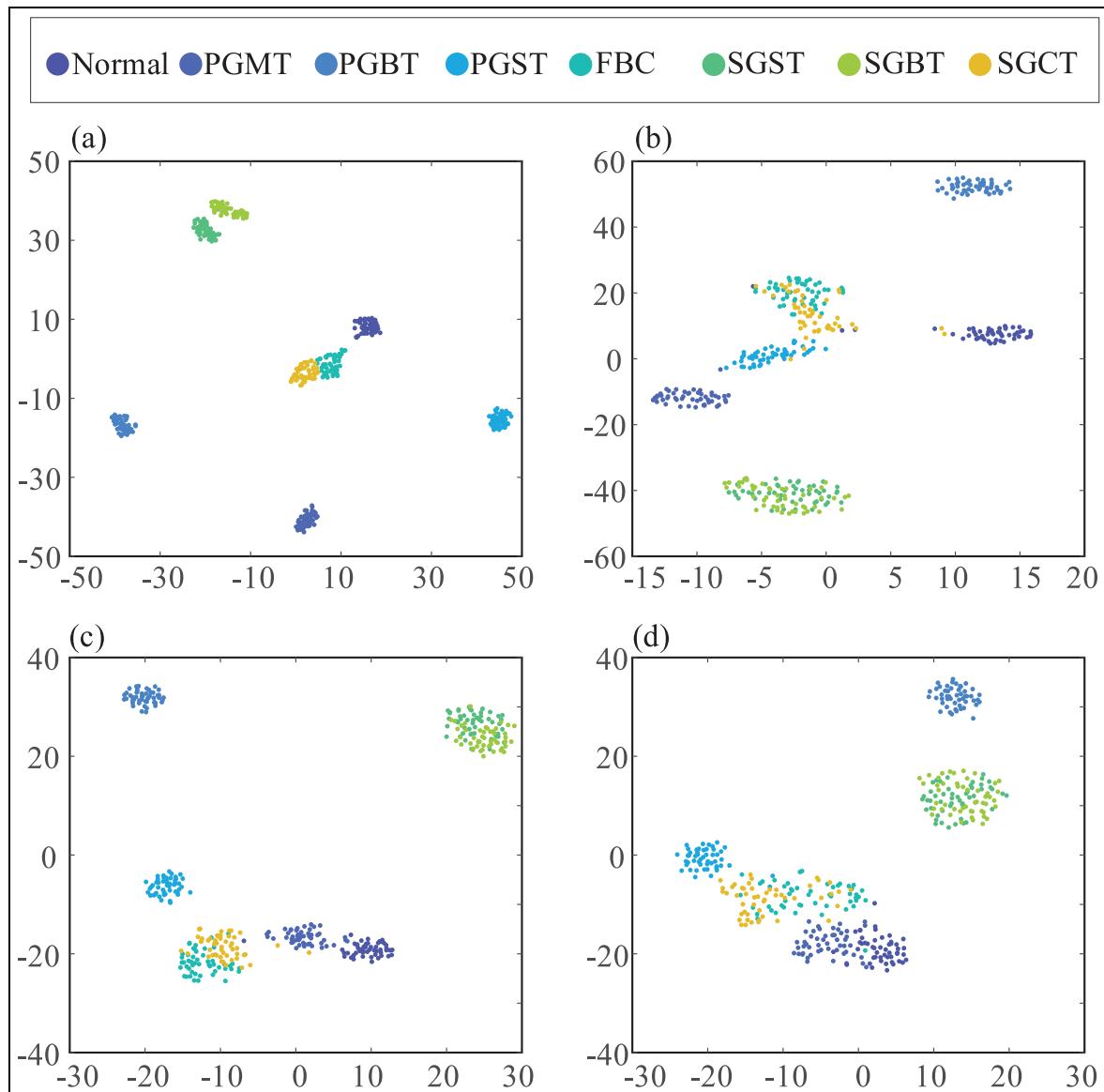


Figure 17. Feature visualization via t-SNE for four entropy methods: (a) SSDE, (b) MDE, (c) MPE, and (d) MFE.

MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy; SNE: stochastic neighbor embedding.

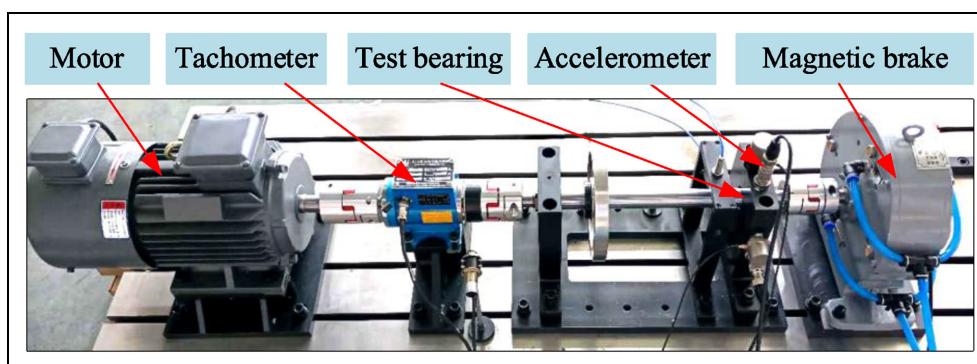


Figure 18. The test rig of rolling bearing system.

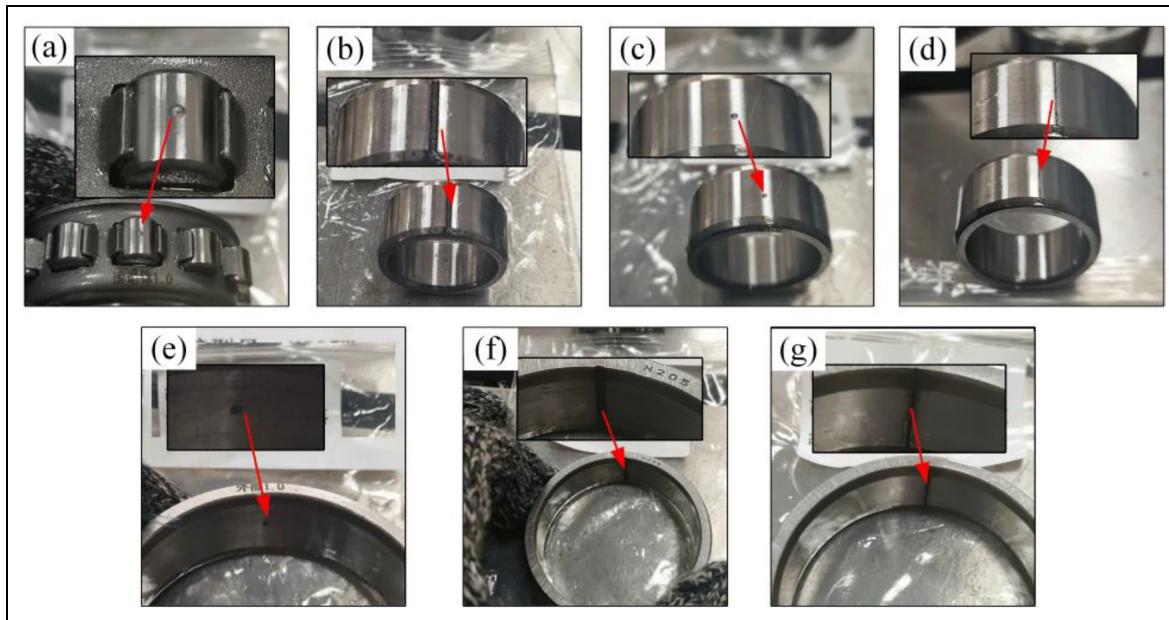


Figure 19. Bearings with different fault types: (a) roller fault, (b) spalling of inner race, (c) pitting of inner race, (d) crack of inner race, (e) pitting of outer race, (f) spalling of outer race, and (g) crack of outer race.

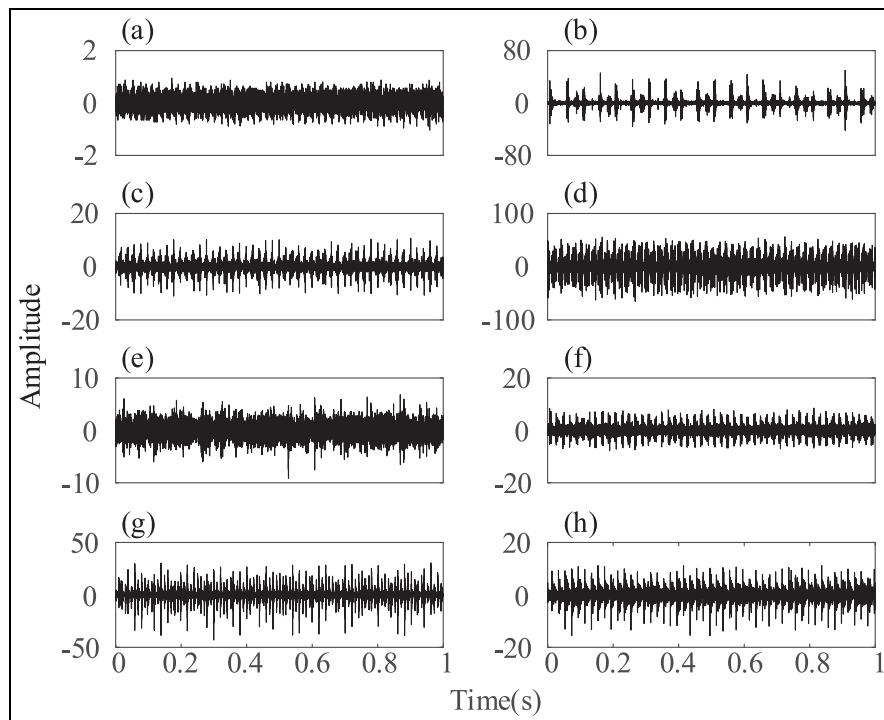


Figure 20. The time-domain signals of eight health conditions for case study II: (a) Normal condition, (b) roller fault, (c) spalling of inner race, (d) spalling of outer race, (e) pitting of inner race, (f) pitting of outer race, (g) crack of inner race, and (h) crack of outer race.

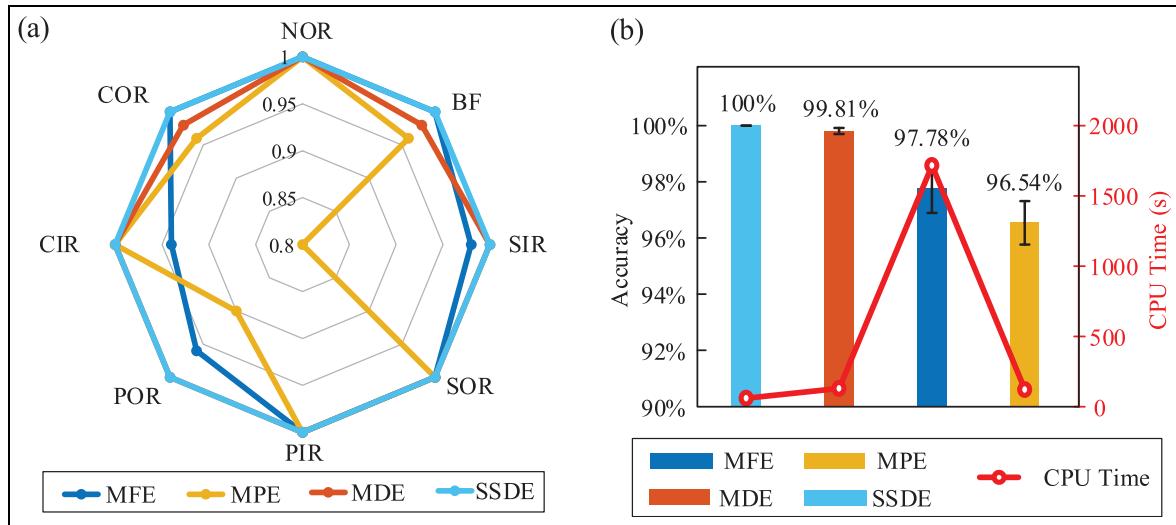


Figure 21. The diagnosis results of four entropy-based methods for bearing system. (a) Radar chart of recognition accuracy for each health condition and (b) performance comparison of four entropy algorithms.

Table 8. The introduction of experimental data for case study II.

Label	Fault location	Fault type	Fault size (mm)	Training sample number	Testing sample number
C1	—	Normal	—	50	50
C2	roller	Pitting	1	50	50
C3	Inner race	Spalling	1	50	50
C4	Outer race	Spalling	1	50	50
C5	Inner race	Pitting	1	50	50
C6	Outer race	Pitting	1	50	50
C7	Inner race	Crack	0.2	50	50
C8	Outer race	Crack	0.2	50	50

efficiency, which reduces the computation time of feature extraction and extracts more reliable information for fault diagnosis.

Robustness against noisy conditions. We also discuss the diagnosis accuracy of four entropy methods under a noisy environment using experimental data. The additional noise was added to the raw signals to compose signals with different SNR values, which can be expressed as Equation (18). Four entropy methods under different SNR values are studied, ranging from 30 to -5 dB, and the stride is 5 dB. Here, the remaining setting is unchanged, and Table 9 and Figure 22 illustrate the diagnosis results.

It can be found that all four algorithms achieve pretty high diagnostic accuracy with high SNR values. For example, when $\text{SNR} \geq 15$ dB, the diagnosis accuracies of four entropy methods are above 90%. Whereas, as the SNR of test samples decreases, so does the diagnosis accuracy of algorithms. For example, when SNR

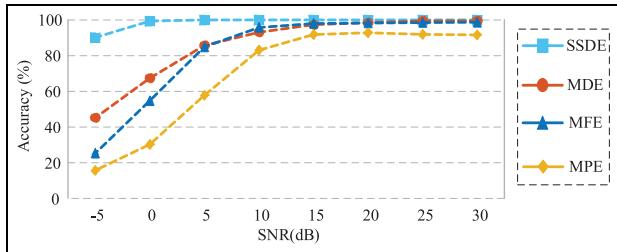
is equal to 0 dB, the diagnosis accuracy of MFE and MPE suffers from a remarkable decrease, where the accuracy of MFE decreases to 54.74%, and the accuracy of MPE decreases to 30.21%. In contrast, in the case of low SNR value, SSDE could still perform well and obtain relatively high diagnostic accuracy. For instance, when the SNR of test samples is equal to 0 dB, the accuracy of SSDE is higher than 99%, which is much higher than 54.74% and 30.21%, the diagnosis accuracy of MFE and MPE, respectively. In addition, the proposed SSDE still achieves a higher diagnostic accuracy than 90%, even with an SNR of -5 dB. However, the diagnosis accuracy of MFE and MPE is below 30%, and the accuracy of MDE is below 50%. The above diagnosis results show that SSDE is more robust to noise than state-of-the-art entropy methods.

Furthermore, to ensure a fair comparison, we have provided more detailed discussion in the paper regarding the impact of parameter settings on the comparison results. This information allows readers to better understand the parameter choices used in both the

Table 9. The classification results with different SNR values for case study II.

SNR (dB)	SSDE		MDE		MFE		MPE	
	Acc (%)	SD (%)						
30	100	0	99.62	0.17	98.75	0.43	91.63	1.05
25	100	0	99.51	0.38	98.58	0.52	91.9	1.26
20	100	0	98.54	0.51	98.33	0.52	92.81	1.72
15	100	0	97.36	0.9	98.04	0.54	91.82	1.23
10	100	0	93.14	1.01	95.84	0.83	83.11	1.55
5	100	0	85.76	1.81	84.91	1.23	57.8	2.59
0	99.35	0.28	67.41	1.87	54.74	1.41	30.21	1.85
-5	90.15	1.06	45.19	3.88	25.28	1.54	15.65	2.38

Acc: accuracy; MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy; SNR: signal-to-noise ratio; SD: standard deviation.

**Figure 22.** The recognition accuracies with different SNR values for four entropy methods.

SNR: signal-to-noise ratio.

proposed approach and the established entropy methods, enabling them to replicate the results and make a fair assessment of the comparison. The entropy features obtained with different parameter selections are input into the SVM classifier, following the same procedure as in previous case studies. The dimensionality of the input features for SVM classifier is fixed at 20, and the classification results are presented in Figure 23.

The results in Figure 23 clearly demonstrate the significant influence of parameter settings on the performance of the MPE and MFE methods. While MFE achieves a recognition rate of 90% when appropriate parameters are chosen, the recognition rate for MPE remains consistently below 90%. In contrast, the MDE method exhibits a high recognition rate of 98% when suitable parameters are selected, with the majority of recognition rates falling within the range of 90%–94%. Remarkably, the proposed SSDE method consistently outperforms the other entropy methods, surpassing a recognition rate of 98% across various parameter selections.

These findings emphasize the critical role of parameter selection in established entropy methods and highlight the superior performance of the proposed SSDE method. By providing a comprehensive analysis

of the impact of parameter settings and showcasing the consistently high recognition rate achieved by the SSDE method, our study bolsters the effectiveness and reliability of the proposed approach for fault diagnosis.

Evaluation metric and visualized interpretation. In addition, the t-SNE algorithm is applied to obtain two-dimensional visualization of entropy feature distribution with SNR = 0 dB as illustrated in Figure 24. It can be found from Figure 24(a) that each cluster of eight health conditions is clearly separated, and SSDE features are clustered well. On the contrary, a large number of features of the other three entropy methods are mixed. Like Case study I, the separable criterion is introduced to compare the useful feature information, and the results are illustrated in Table 7. Results verify that more useful fault information can be extracted by the SSDE method incorporating time-delay analysis and SDF process, thus generating higher diagnostic accuracy.

Integration with other entropy methods. In the subsection, apart from diversity entropy, we also integrate symbol-scale analysis with dispersion entropy and FE to further demonstrate the superiority of the proposed symbol-scale analysis method. The experiment of rolling bearing under 5 dB SNR are carried out to investigate the diagnosis performance of symbol-scale dispersion entropy, symbol-scale FE. Also, the original multiscale dispersion entropy, MFE are both used for comparison. The obtained classification results are listed in Figure 25.

It can be found that in this case study, symbol-scale dispersion entropy and symbol-scale FE also improved the recognition accuracy and achieved desired diagnostic performance improvement compared with the multiscale entropy approach. It can be indicated that the

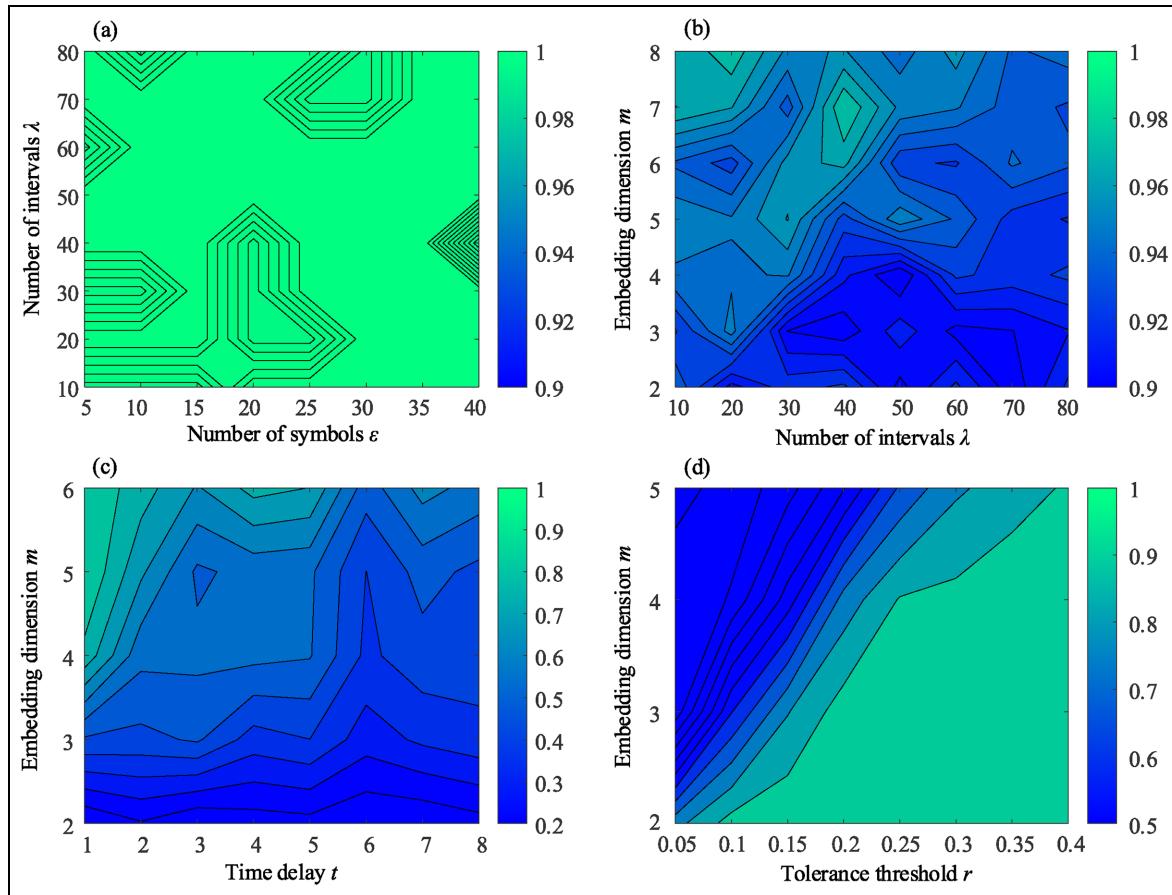


Figure 23. Classification accuracies when selecting different parameters for four entropy methods under 5 dB case: (a) SSDE, (b) MDE, (c) MPE, and (d) MFE.
MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy.

combination of the proposed symbol-scale and other entropy methods is also useful. The proposed symbol-scale method can successfully overcome the shortcomings of traditional multiscale analysis methods and significantly improve the accuracy of feature extraction and the robustness to noise.

Conclusion

In this work, SSDE is proposed to solve two main problems of coarse-graining-based multiscale approach. On the one hand, SSDE utilizes the time-delay process of different intervals to capture the fault characteristics from the time series over multiple time scales. On the other hand, SSDE combines the merits of SDF in noise reduction and calculation efficiency to overcome the limitation of weak fault extraction under a low-SNR environment. Simulations and experimental case studies

are applied to validate the superiority of SSDE method by comparing it with original multiscale entropy methods, including MDE, MFE, and MPE. Results demonstrate that the proposed SSDE can effectively recognize various bearing and gearbox fault types with the highest recognition of 98.91% and 100%, respectively. Moreover, the calculational efficiency of the proposed SSDE is the highest, which is nearly 30 times faster than that of MFE and twice faster than that of MDE, respectively.

In this paper, the concept of symbol-scale analysis is proposed and combined with diversity entropy, which can enhance the fault detection ability and improve the calculation efficiency. In our future work, the incorporation of symbol-scale analysis and other entropy methods will be discussed in more detail. Moreover, the symbol-scale entropy methods will be studied in other fields, such as acoustical signals.

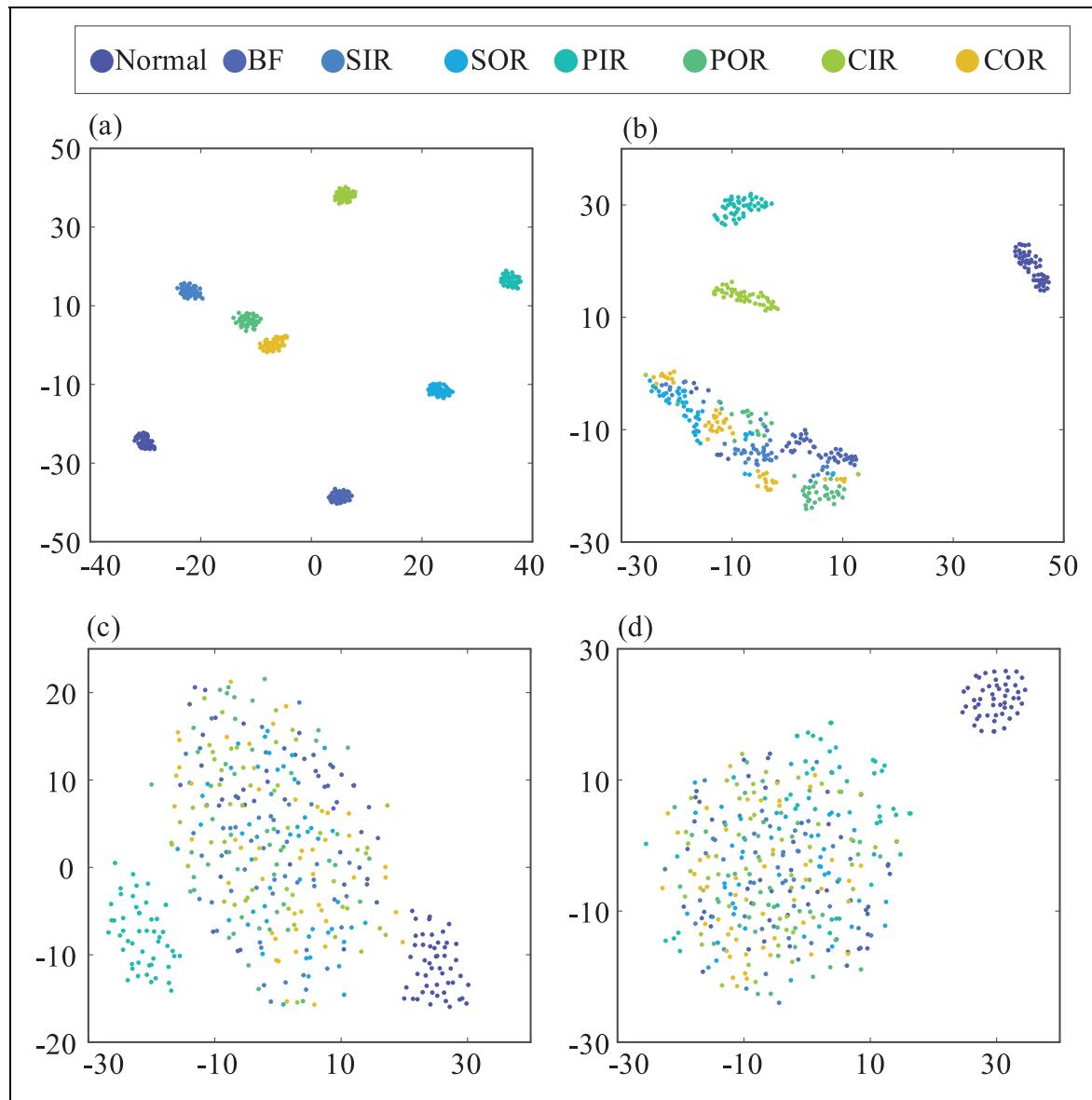


Figure 24. Feature visualization via t-SNE for four entropy methods under -2 dB case: (a) SSDE, (b) MDE, (c) MFE, and (d) MPE. MDE: multiscale diversity entropy; MFE: multiscale fuzzy entropy; MPE: multiscale permutation entropy; SSDE: symbol-scale diversity entropy; SNE: stochastic neighbor embedding.

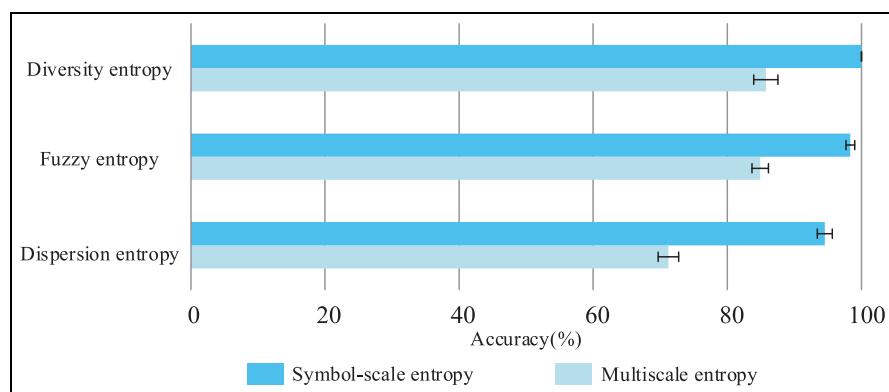


Figure 25. The versatility of incorporation of proposed symbol-scale approach and other entropy methods.

Declaration of conflicting interests

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

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research is supported by the National Natural Science Foundation of China under Grant 12172290 and 52250410345.

ORCID iDs

Shun Wang  <https://orcid.org/0000-0002-5766-6220>
 Yongbo Li  <https://orcid.org/0000-0003-2699-9951>

References

- Wei Y, Wang X, Xu Y, et al. Intelligent fault diagnosis of rotating machinery using composite multivariate-based multi-scale symbolic dynamic entropy with multi-source monitoring data. *Struct Health Monit* 2023; 22(1): 56–77.
- Sharma V and Parey A. Gearbox fault diagnosis using RMS based probability density function and entropy measures for fluctuating speed conditions. *Struct Health Monit* 2017; 16: 682–695.
- Wang Z, Yang N, Li N, et al. A new fault diagnosis method based on adaptive spectrum mode extraction. *Struct Health Monit* 2021; 20: 3354–3370.
- Liu R, Yang B, Zio E, et al. Artificial intelligence for fault diagnosis of rotating machinery: a review. *Mech Syst Signal Proc* 2018; 108: 33–47.
- Cerrada M, Sanchez R-V, Li C, et al. A review on data-driven fault severity assessment in rolling bearings. *Mech Syst Signal Proc* 2018; 99: 169–196.
- Huo Z, Martínez-García M, Zhang Y, et al. Entropy measures in machine fault diagnosis: insights and applications. *IEEE Trans Instrum Meas* 2020; 69: 2607–2620.
- Civera M and Surace C. Instantaneous spectral entropy: an application for the online monitoring of multi-storey frame structures. *Buildings* 2022; 12: 310.
- Civera M and Surace C. An application of instantaneous spectral entropy for the condition monitoring of wind turbines. *Appl Sci* 2022; 12: 1059.
- Wang Z, Yao L, Chen G, et al. Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals. *ISA Trans* 2021; 114: 470–484.
- Sharma S, Tiwari SK and Sukhejeet S. Integrated approach based on flexible analytical wavelet transform and permutation entropy for fault detection in rotary machines. *Measurement* 2021; 169: 108389.
- Wang S, Li Y, Si S, et al. Enhanced hierarchical symbolic sample entropy: efficient tool for fault diagnosis of rotating machinery. *Struct Health Monit* 2023; 22(3): 1927–1940.
- Wang X and Liu L. Concentric diversity entropy: a high flexible feature extraction tool for identifying fault types with different structures. *Mech Syst Signal Process* 2022; 171: 108934.
- Yan R and Gao RX. Approximate entropy as a diagnostic tool for machine health monitoring. *Mech Syst Signal Proc* 2007; 21: 824–839.
- Kedadouche M, Thomas M, Tahan A, et al. Nonlinear parameters for monitoring gear: comparison between Lempel-Ziv, approximate entropy, and sample entropy complexity. *Shock Vibr* 2015; 2015: e959380.
- Yan R, Liu Y and Gao RX. Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines. *Mech Syst Signal Process* 2012; 29: 474–484.
- Chen W, Zhuang J, Yu W, et al. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med Eng Phys* 2009; 31: 61–68.
- Li Y, Jiao S and Geng B. Refined composite multiscale fluctuation-based dispersion Lempel-Ziv complexity for signal analysis. *ISA Trans* 2023; 133: 273–284.
- Omidvarnia A, Mesbah M, Pedersen M, et al. Range entropy: a bridge between signal complexity and self-similarity. *Entropy* 2018; 20: 962.
- Zhang T, Chen W and Li M. Fuzzy distribution entropy and its application in automated seizure detection technique. *Biomed Signal Process Control* 2018; 39: 360–377.
- Rostaghi M, Khatibi MM, Ashory MR, et al. Fuzzy dispersion entropy: a nonlinear measure for signal analysis. *IEEE Trans Fuzzy Syst* 2021; 30(9): 3785–3796.
- Wang X, Si S and Li Y. Multiscale diversity entropy: a novel dynamical measure for fault diagnosis of rotating machinery. *IEEE Trans Ind Inf* 2020; 17(8): 5419–5429.
- Rajagopalan V and Ray A. Wavelet-based space partitioning for symbolic time series analysis. In: *Proceedings of the 44th IEEE conference on decision and control*, Seville, Spain, 2005, pp. 5245–5250.
- Li Y, Wang S, Yang Y, et al. Multiscale symbolic fuzzy entropy: an entropy denoising method for weak feature extraction of rotating machinery. *Mech Syst Signal Proc* 2022; 162: 108052.
- Li Y, Wang S, Li N, et al. Multiscale symbolic diversity entropy: a novel measurement approach for time-series analysis and its application in fault diagnosis of planetary gearboxes. *IEEE Trans Ind Inf* 2021; 18(2): 1121–1131.
- Wang Y, Xu Y, Liu M, et al. Cumulative residual symbolic dispersion entropy and its multiscale version: methodology, verification, and application. *Chaos, Solitons Fractals* 2022; 160: 112266.
- Rostaghi M and Azami H. Dispersion entropy: a measure for time-series analysis. *IEEE Signal Process Lett* 2016; 23: 610–614.
- Li C, Noman K, Liu Z, et al. Optimal symbolic entropy: an adaptive feature extraction algorithm for condition monitoring of bearings. *Inf Fusion* 2023; 98: 101831.
- Costa M, Goldberger AL and Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett* 2002; 89: 068102.

29. Zheng J, Cheng J, Yang Y, et al. A rolling bearing fault diagnosis method based on multi-scale fuzzy entropy and variable predictive model-based class discrimination. *Mech Mach Theory* 2014; 78: 187–200.
30. Zhou R, Wang X, Wan J, et al. EDM-Fuzzy: an euclidean distance based multiscale fuzzy entropy technology for diagnosing faults of industrial systems. *IEEE Trans Ind Inf* 2021; 17: 4046–4054.
31. Costa M, Goldberger AL and Peng C-K. Multiscale entropy analysis of biological signals. *Phys Rev E* 2005; 71: 021906.
32. Wu SD, Wu CW, Lin SG, et al. Analysis of complex time series using refined composite multiscale entropy. *Phys Lett* 2014; 378: 1369–1374.
33. Zheng J, Pan H, Yang S, et al. Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis. *Mech Syst Signal Process* 2018; 99: 229–243.
34. Pham TD. Time-shift multiscale entropy analysis of physiological signals. *Entropy* 2017; 19: 257.
35. Sharma S and Tiwari SK. A novel feature extraction method based on weighted multi-scale fluctuation based dispersion entropy and its application to the condition monitoring of rotary machines. *Mech Syst Signal Process* 2022; 171: 108909.
36. Zheng J, Ying W, Tong J, et al. Multiscale three-dimensional Holo-Hilbert spectral entropy: a novel complexity-based early fault feature representation method for rotating machinery. *Nonlinear Dyn* 2023; 111: 10309–10330.
37. Jiang Y, Peng C-K and Xu Y. Hierarchical entropy analysis for biological signals. *J Comput Appl Math* 2011; 236: 728–742.
38. Zhu X, Zheng J, Pan H, et al. Time-shift multiscale fuzzy entropy and laplacian support vector machine based rolling bearing fault diagnosis. *Entropy* 2018; 20: 602.
39. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D* 1988; 31: 277–283.
40. Li Y, Yang Y, Li G, et al. A fault diagnosis scheme for planetary gearboxes using modified multi-scale symbolic dynamic entropy and mRMR feature selection. *Mech Syst Signal Process* 2017; 91: 295–312.
41. Chang C-C and Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2011; 2: 27:1–27:27.
42. Wang X, Si S and Li Y. Hierarchical diversity entropy for the early fault diagnosis of rolling bearing. *Nonlinear Dyn* 2022; 108: 1447–1462.
43. Liang X, Zuo MJ and Pandey M. Analytically evaluating the influence of crack on the mesh stiffness of a planetary gear set. *Mech Mach Theory* 2014; 76: 20–38.
44. Tian X. *Dynamic simulation for system response of gearbox including localized gear faults*. MS Thesis, University of Alberta, 2004.
45. Liang X-H, Liu Z-L, Pan J, et al. Spur gear tooth pitting propagation assessment using model-based analysis. *Chin J Mech Eng* 2017; 30: 1369–1382.
46. McFadden PD and Smith JD. Model for the vibration produced by a single point defect in a rolling element bearing. *J Sound Vibr* 1984; 96: 69–82.
47. Zheng J, Jiang Z and Pan H. Sigmoid-based refined composite multiscale fuzzy entropy and t-SNE based fault diagnosis approach for rolling bearing. *Measurement* 2018; 129: 332–342.
48. Yang J, Liu YL, Feng CS, et al. Applying the Fisher score to identify Alzheimer's disease-related genes. *Genet Mol Res*. Epub ahead of print 27 June 2016. DOI: 10.4238/gmr.15028798.
49. Gu Q, Li Z and Han J. Generalized fisher score for feature selection. *arXiv:12023725 [cs, stat]*, <http://arxiv.org/abs/1202.3725> (2012, accessed 30 July 2021).

Appendix I

For the gear simulation, the motion equations in the x -direction can be expressed as:

$$m_1\ddot{x}_1 = -k_{x1}x_1 - c_{x1}\dot{x}_1 \quad (A1)$$

$$m_2\ddot{x}_2 = -k_{x2}x_2 - c_{x2}\dot{x}_2 \quad (A2)$$

The motion equations in the y -direction can be expressed as:

$$\begin{aligned} m_1\ddot{y}_1 = & -k_{y1}y_1 - c_y\dot{y}_1 + k_t(R_{b1}\theta_1 - R_{b2}\theta_2 + y_1 - y_2) \\ & + c_t(R_{b1}\dot{\theta}_1 - R_{b2}\dot{\theta}_2 + \dot{y}_1 - \dot{y}_2) \end{aligned} \quad (A3)$$

$$\begin{aligned} m_2\ddot{y}_2 = & -k_{y2}y_2 - c_2\dot{y}_2 + k_t(R_{b1}\theta_1 - R_{b2}\theta_2 + y_1 - y_2) \\ & + c_t(R_{b1}\dot{\theta}_1 - R_{b2}\dot{\theta}_2 + \dot{y}_1 - \dot{y}_2) \end{aligned} \quad (A4)$$

The motion equations of the rotation can be expressed as:

$$I_m\ddot{\theta}_m = M_1 - k_p(\theta_m - \theta_1) - c_p(\dot{\theta}_m - \dot{\theta}_1) \quad (A5)$$

$$I_b\ddot{\theta}_b = k_g(\theta_2 - \theta_b) + c_g(\dot{\theta}_2 - \dot{\theta}_b) - M_2 \quad (A6)$$

For the rotary motion of the pinion and gear, the motion equations can be expressed as:

$$\begin{aligned} I_1\ddot{\theta}_1 = & k_p(\theta_m - \theta_1) + c_p(\dot{\theta}_m - \dot{\theta}_1) - R_{b1}[k_t(R_{b1}\theta_1 \\ & - R_{b2}\theta_2 + y_1 - y_2) + c_t(R_{b1}\dot{\theta}_1 - R_{b2}\dot{\theta}_2 + \dot{y}_1 - \dot{y}_2)] \end{aligned} \quad (A7)$$

$$\begin{aligned} I_2\ddot{\theta}_2 = & -k_g(\theta_2 - \theta_b) - c_g(\dot{\theta}_2 - \dot{\theta}_b) + R_{b2} \\ & [k_t(R_{b1}\theta_1 - R_{b2}\theta_2 + y_1 - y_2) \\ & + c_t(R_{b1}\dot{\theta}_1 - R_{b2}\dot{\theta}_2 + \dot{y}_1 - \dot{y}_2)] \end{aligned} \quad (A8)$$

where:

k_t -mesh stiffness;

c_t -mesh damping coefficient;

m_1/m_2 -mass of the pinion/the gear;

k_{x1}/k_{x2} -horizontal radial stiffness of the input bearings/the output bearings;

c_{x1}/c_{x2} -horizontal radial viscous damping coefficient of input bearings/output bearings;

k_1/k_2 -vertical radial stiffness of the input bearings/the output bearings;

c_1/c_2 -vertical radial damping coefficient of the input bearings/the output bearings;

y_1/y_2 -vertical radial damping coefficient of the input bearings/the output bearings;

R_{b1}/R_{b2} -base circle radius of pinion/gear;

M_1/M_2 -input motor torque/output torque from load;

k_p/k_g -torsional stiffness of the input flexible coupling/the output flexible coupling;

c_p/c_g -damping coefficient of the input flexible coupling/the output flexible coupling;

$\theta_m/\theta_1/\theta_2/\theta_b$ -angular displacement of motor/ the pinion/ the gear/the load;

$I_m/I_1/I_2/I_b$ -mass moment of inertia of the motor/the pinion/ the gear/the load.

Appendix 2

The simulated outer ring failure signal can be indicated as:

$$v_o(t) = \left[\sum_{k=-\infty}^{+\infty} d_o \delta(t - kT_o) \right] * e(t) \quad (\text{A9})$$

where d_o is influence coefficient of outer race by defects. T_o is the reciprocal of outer race fault frequency. $e(t)$ represents damping function, $\delta(t)$ represents impulse function, and k represents impulse number. The damp function can be expressed as:

$$e(t) = \begin{cases} \exp(-t/T), & t > 0 \\ 0, & t \leq 0 \end{cases} \quad (\text{A10})$$

The simulated inner ring failure signal can be indicated as:

$$\left\{ \left(\sum_{t=-\infty}^{+\infty} (d_i \delta(t - kT_i)) \cdot q(2\pi f_r t) \cdot p(2\pi f_r t) \right) * e(t) \right\} \quad (\text{A11})$$

where d_i is influence coefficient of inner race by defects. T_i is the reciprocal values of fault characteristic frequency of inner race. $p(\varphi) = \cos \varphi$ represents the

influence coefficient of localized defect. $q(\varphi) = q_{\max} [1 - 1/2\sigma(1 - \cos \varphi)]^n$ represents radial load distribution, where σ means the coefficient of load distribution. In this paper $\sigma = 0.5$, $n = 1.1$ for cylindrical roller bearing. In addition, f_r is the rotating frequency.

The simulated rolling element failure signal can be indicated as Equation (12).

$$\left\{ \left(\sum_{t=-\infty}^{+\infty} \left(d_{bo} \delta(t - kT_b) + d_{bi} \delta \left(t - kT_b - \frac{1}{2} T_b \right) \right) \right) \cdot q(2\pi f_r t) \cdot p(2\pi f_r t) \right\} * e(t) \quad (\text{A12})$$

where d_{bo} , d_{bi} are influence coefficients of rolling elements when hitting the outer ring, and rolling element when hitting the inner ring, respectively. T_b is the reciprocal value of fault characteristic frequency of rolling elements.

Notation

RF	Roller fault
CIR	Crack of inner race
COR	Crack of outer race
DE	Diversity entropy
FBC	Fracture of bearing cage
FE	Fuzzy entropy
MDE	Multiscale diversity entropy
MFE	Multiscale fuzzy entropy
MPE	Multiscale permutation entropy
PGMT	Planet gear with a missing tooth
PGST	Planet gear with a spalling tooth
PGBT	Planet gear with a broken tooth
PE	Permutation entropy
PIR	Pitting of inner race
POR	Pitting of outer race
SDF	Symbolic dynamic filtering
SGBT	Sun gear with a broken tooth
SGST	Sun gear with a spalling tooth
SGCT	Sun gear with a cracked tooth
SIR	Spalling of inner race
SNR	Signal-to-noise ratio
SOR	Spalling of outer race
SSDE	Symbol-scale diversity entropy
SVM	Support vector machine