



Cumulative spectrum distribution entropy for rotating machinery fault diagnosis

Shun Wang^a, Yongbo Li^{a,*}, Khandaker Noman^b, Dong Wang^c, Ke Feng^d, Zheng Liu^e, Zichen Deng^a

^a School of Aeronautics, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

^b School of Civil Aviation, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

^c State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University, Shanghai 200240, China

^d Department of Industrial Systems Engineering and Management, National University of Singapore, 117576, Singapore

^e School of Engineering, University of British Columbia, Kelowna V1V 1V7, Canada

ARTICLE INFO

Communicated by W.-X. Ren

Keywords:

Cumulative spectrum distribution entropy

Entropy

Nonlinear dynamics

Feature extraction

Fault diagnosis

ABSTRACT

Entropy-based methods have shown promise in detecting dynamic changes in non-linear signals and have been widely applied in fault diagnosis for rotating machinery. However, these methods have limitations when it comes to capturing frequency-domain information of fault features, as they are primarily based on time-domain signals. To address this issue, this paper proposes a new entropy measure called cumulative spectrum distribution entropy (CSDEn), which is based on the cumulative distribution of the spectrum and considers both frequency probability and frequency values in the spectrum domain. The proposed method is evaluated using synthetic signals and experimental data from different bearing and gear working states. The results show that CSDEn outperforms other widely used entropy measures in detecting dynamic changes and measuring signal complexity with low noise sensitivity and high computing efficiency. Nonparametric Mann–Whitney U tests reveal significant differences between different working states for proposed CSDEn method, and compared with other entropy methods, CSDEn achieves the highest recognition rates in diagnosing different bearing and gear working states. Moreover, proposed CSDEn method demonstrates its effectiveness in addressing the challenges of small sample datasets and strong noise interference, making it highly competitive in real industrial applications.

1. Introduction

Rotating machinery is widely used in a range of modern industries, such as transportation, power equipment, aerospace, vehicles, etc. [1,2]. Due to the harsh working environments, rotating machinery is prone to failure, which will lead to high maintenance costs and even serious accidents. Therefore, it is of great significance for condition monitoring and fault diagnosis of rotating machinery [3].

Recently, entropy measure has drawn substantial attention in non-linear signal analysis and fault diagnosis due to its powerful feature representation capabilities [4–7]. This has resulted in significant development in this subject, making the entropy measure a valuable tool in researching nonlinear dynamical systems and fault feature extraction through time series [8]. As a nonlinear

* Corresponding author.

E-mail addresses: wangshun@mail.nwpu.edu.cn (S. Wang), yongbo@nwpu.edu.cn (Y. Li), khandakernoman93@nwpu.edu.cn (K. Noman), dongwang4-c@sjtu.edu.cn (D. Wang), ke.feng@outlook.com.au (K. Feng), zheng.liu@ubc.ca (Z. Liu), dweifan@nwpu.edu.cn (Z. Deng).

<https://doi.org/10.1016/j.ymssp.2023.110905>

Received 20 June 2023; Received in revised form 27 September 2023; Accepted 31 October 2023

Available online 15 November 2023

0888-3270/© 2023 Elsevier Ltd. All rights reserved.

Nomenclature

ApEn	Approximate entropy
CRDE	Cumulative residual distribution entropy
CRE	Cumulative residual entropy
CSDEn	Cumulative spectrum distribution entropy
DispEn	Dispersion entropy
DistEn	Distribution entropy
DivEn	Diversity entropy
FuzzyEn	Fuzzy entropy
PerEn	Permutation entropy
SampEn	Sample entropy
SD	Standard deviation
SNR	Signal-to-noise ratio
SpEn	Spectral entropy
WGN	White Gaussian noise

measure, entropy can quantify the regularity or orderliness of time series. When a fault or damage occurs in rotating machinery, the amplitude and frequency modulation phenomena will happen, resulting in a complexity change. This indicates that the entropy-based methods can be utilized to detect the dynamic changes of rotating machinery and further to identify different health conditions [9,10].

Since its introduction by Shannon in 1948 [11], researchers have been utilizing entropy-based metrics to analyze real-time series. Entropy can characterize the original signal's complexity or uncertainty, with higher entropy values indicating a more complex time series [12]. In recent years, there has been significant progress in the development of entropy-based methods for analyzing time series, including approximate entropy (ApEn) [13], sample entropy (SampEn) [14], permutation entropy (PerEn) [15], fuzzy entropy (FuzzyEn) [16], distribution entropy (DistEn) [17], dispersion entropy (DispEn) [18], diversity entropy (DivEn) [19] and so on [20–22]. To enhance the capability of entropy-based methods for fault diagnosis in low-SNR environments, symbolic dynamic filtering has been integrated with entropy methods [23–25]. Furthermore, Costa et al. [26] developed a multiscale analysis based on coarse-graining analysis, which led to the development of multiscale-based methods [27–29]. These multiscale-based entropy methods have been widely applied in mechanical system fault diagnosis by extracting information from complex and non-linear time series [30].

All of entropy techniques listed above are based on time domain signals. However, these methods have limitations in capturing changes in the frequency domain, particularly when it comes to detecting fault-related information. Therefore, in addition to the time domain, in the frequency domain, spectral entropy (SpEn) [31] can be calculated, which treats the normalized power distribution in the frequency domain as a probability distribution and calculates its Shannon entropy. SpEn can be interpreted as measuring the complexity or uncertainty in the frequency domain from energy perspective. A high value of SpEn means a flat, uniform spectrum, which implies the signal is complex, with a wide range of spectral content. On the contrary, a low SpEn value means that almost the power of spectrum is concentrated in a very small number of frequency components, indicating that the signal is less complex and more predictable [32].

Spectral entropy, as a measure of signal complexity in the frequency domain, has found widespread use in signal processing and feature extraction in fault diagnosis [33,34], speech recognition [35], biomedical signal processing [36], complex network [37], chaos analysis [38] and so on [39–42]. In addition, Mao et al. introduced SpEn into a complexity-entropy curve and validated that it can characterize time series efficiently [43]. Zhang et al. introduced a novel approach for tracking changes in financial time series by developing a complexity-entropy causality plane based on multiscale power spectrum entropy [44].

While SpEn and its multiscale version have found numerous applications in health monitoring and fault diagnosis [30,45], certain shortcomings still exist. Due to the algorithm's simplicity, SpEn loses some critical information about the time series. On the one hand, the signals can be distorted by interference or noise in actual acquisition data. The original SpEn calculates the frequency distribution directly, which counts the probability density of each spectral component so that changes in individual frequencies lead to changes in the entropy value, which is then sensitive to noises. On the other hand, the original SpEn only considers the frequency distribution in the reconstructed spectrum space, disregarding the values of frequencies. Specifically, assume that there are two spectrum distributions $D_1 = \{1, 2, 3, 4\}$ and $D_2 = \{4, 3, 2, 1\}$, where the information behind two spectrum distributions is different, the spectrum probability for them is $P_1 = \{0.1, 0.2, 0.3, 0.4\}$ and $P_2 = \{0.4, 0.3, 0.2, 0.1\}$. The same entropy value, which is $En = -(0.1 \log_2 0.1 + 0.2 \log_2 0.2 + 0.3 \log_2 0.3 + 0.4 \log_2 0.4) / \log_2 4 = 0.9232$, can be obtained. In other words, the classical spectral distribution probability fails to differentiate between two different spectral distributions or spectral shapes that represent distinct dynamic systems, resulting in the same entropy value.

Therefore, this paper presents a novel approach for estimating entropy in the spectral domain, which addresses the limitations of existing methods by considering both the probability and values of frequencies in the spectrum. Drawing inspiration from the concepts of distribution entropy and cumulative residual entropy [46], two established information measurement techniques, we

apply the concept of bin distribution to the probability statistics of spectrum to ensure better stability and robustness, and then introduce the idea of cumulative distribution function, instead of density function, to frequency-domain, to acquire more information contained in the frequency distribution of the time series.

The proposed CSDEn method leverages the inherent information present in the frequency domain by means of bin distribution and cumulative distribution function. On the one hand, the concept of bin distribution is introduced to the spectrum domain to detect the distribution of temporal and spatial structures in series. On the other hand, the cumulative distribution function is applied to the complexity measure of spectral distribution, which allows a very fine distinction between the chaotic complexity of the time series.

The results obtained from synthetic signals demonstrate the effectiveness of the proposed CSDEn method in characterizing time series generated from both stochastic systems (white noise and $1/f$ noise), and deterministic systems (chaotic and periodic series). Moreover, the CSDEn method exhibits low sensitivity to noise and high computational efficiency, making it a promising tool for analyzing and processing various types of signals. The proposed CSDEn method was also applied to the analysis of real-world mechanical signals, showing its potential in fault diagnosis. The experimental results demonstrated the superiority of the CSDEn method in detecting dynamic changes of time series and its excellent performance in diagnosing fault states for bearing and gearbox. The method effectively distinguished the fault signals from normal signals, and the results showed that the CSDEn method can be used as a powerful tool for fault diagnosis in mechanical systems. Overall, the proposed method provides a new perspective for analyzing signals in the frequency domain and has excellent potential for a wide range of applications.

The main contributions of this work can be summarized as follows: (1) To capture the characteristics from spectral domain in the time series, the concept of bin distribution is introduced to the probability statistics of spectrum to measure the complexity of the frequency distribution. (2) The cumulative distribution function instead of density function is introduced to acquire more information in the frequency distribution of the time series from a frequency domain perspective. (3) The effectiveness of the proposed CSDEn method is systematically validated through comparative studies using both synthetic and real-world signals. Statistical analysis and machine learning techniques are used to compare the performance of CSDEn with other entropy methods in terms of dynamic change detection, robustness to noise, stability to signal length, and computational efficiency.

The remainder of this paper is organized as follows. Section 2 defines original spectral entropy, cumulative residual entropy and proposes cumulative spectrum distribution entropy. In Section 3, the effectiveness of the proposed CSDEn method is demonstrated through synthetic signals. The real-world mechanical signals are applied to verify the performance of fault diagnosis applications in Section 4. Lastly, in Section 5, the conclusions of this work are summarized.

2. Theory

2.1. Spectral entropy

In signal processing, SpEn has been introduced to measure the distribution features of a signal spectrum [32,33]. For a given time series $\{x(n)\}$ with length N , $n = 0, 1, \dots, N-1$, the defined SpEn can be accomplished by the following steps:

Step 1 Obtain the spectrum $X(k)$, $k = 0, 1, \dots, N-1$ of signal $\{x(n)\}$ using the discrete Fourier transform, as follows:

$$X(k) = \text{DFT}[x(n)] = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi}{N}nk} \quad (1)$$

where k represents the k th frequency component, $k = 0, 1, \dots, N-1$.

Step 2 The power spectral density distribution function of signal $x(n)$ is obtained as

$$p(k) = \frac{|X(k)|^2}{\sum_{k=0}^{\frac{N}{2}-1} |X(k)|^2} \quad (2)$$

where $k = 0, 1, \dots, N/2-1$. It can be observed that $\sum_{k=0}^{\frac{N}{2}-1} p(k) = 1$, which meets the definition of Shannon entropy. It is noticed that when employing discrete Fourier transform, the summation runs from $k = 0$ to $k = N/2-1$ [42,47].

Step 3 The classical spectral entropy, SpEn, is estimated by computing the Shannon entropy from the obtained spectral density distribution, i.e.,

$$SpEn = - \sum_{k=0}^{\frac{N}{2}-1} p(k) \log_2 p(k) \quad (3)$$

SpEn may vary with the length of the spectral sequence. For comparison purposes, it is usually normalized by the length to obtain the normalized spectral entropy $SpEn_n$ as follows:

$$SpEn_n = - \frac{\sum_{k=0}^{\frac{N}{2}-1} p(k) \log_2 p(k)}{\log_2 \left(\frac{N}{2} \right)} \quad (4)$$

2.2. Cumulative residual entropy

Cumulative residual entropy (CRE), as a generalization of traditional Shannon entropy, uses the cumulative distribution of random variables, which allows for a more comprehensive measurement of the complexity of a system or signal and has been shown to be more robust and applicable in a variety of contexts [46,48,49]. The definitions of CRE for continuous and discrete versions are given as follows:

Mathematically, for a random vector $X \in \mathcal{R}^N$, the continuous CRE for X can be expressed as [46]:

$$CRE(X) = - \int_{\mathcal{R}_+^N} P(|X| > \lambda) \log P(|X| > \lambda) d\lambda \quad (5)$$

where $X = \{X_1, X_2, \dots, X_N\}$, $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$, and $|X| > \lambda$ represents $|X_i| > \lambda_i$ and $\mathcal{R}_+^N = \{x_i \in \mathcal{R}^N; x_i \geq 0\}$.

Then, the CRE for discrete situation is given as follow [46]:

Let X_1, X_2, \dots, X_N be positive and independent and identically distributed (i.i.d.) with distribution F . Let p_N be the empirical distribution of X_1, X_2, \dots, X_N and $F_N(X) = \frac{1}{N} \sum_{i=1}^N p_i$ is the cumulative density function, and $G_N(X) = 1 - F_N(X)$, the CRE of the empirical distribution F is defined as:

$$CRE(X) = - \int_0^\infty G_N(X) \log G_N(X) dx \quad (6)$$

2.3. Cumulative spectrum distribution entropy

Even though SpEn has made great applications in signal processing and analysis, a severe drawback still exists, resulting in poor stability and loss of information. Therefore, the introduction of the cumulative distribution function and bin distribution to SpEn allows for a more refined characterization of the chaotic complexity of time series.

For a given time series $\{x(n)\}$ with length N , $n = 0, 1, \dots, N-1$, the defined cumulative spectrum distribution entropy can be accomplished by the following steps:

Step 1 The discrete Fourier transform $X(k)$ of signal $x(n)$ can be obtained according to Eq. (1).

Step 2 The normalized power $p(k)$, $k = 0, 1, \dots, N/2 - 1$, of each frequency component is calculated according to Eq. (2).

Step 3 The spectral components of $N/2$ points are divided into M bins, and then the probability pdf_m is calculated, which is the power probability of each bin. Mathematically, it can be expressed as:

$$pdf_m = \sum_{k \in Bin_m} |p(k)| \quad (7)$$

where $m = 1, 2, \dots, M$, k represents the k th frequency component, and Bin_m represents the set of frequency components contained in the m th bin.

Step 4 Based on the probability density function, cumulative distribution function for the spectrum bin, denoted as F_d , can be obtained.

$$F_d(m) = \sum_{j=1}^m pdf_j \quad (8)$$

Step 5 The utilization of the normalized cumulative residual function (NCRF) enables effective retention of distribution information and amplification of differences between frequency bins. Mathematically, it can be expressed as:

$$ncrf_m = \frac{1 - F_d(m)}{\sum_{m=1}^M (1 - F_d(m))} \quad (9)$$

Step 6 The proposed CSDEn method applies the cumulative residual probability distribution $ncrf_m$ to the Shannon entropy theoretical framework. Mathematically, it can be expressed as:

$$CSDEn(x, M) = - \frac{1}{\log_2 M} \sum_{m=1}^M ncrf_m \log_2 (ncrf_m) \quad (10)$$

It can be observed that $\sum_{m=1}^M ncrf_m = 1$, which also meets the definition of Shannon entropy. Thus, the resulting entropy value should fall within the range $[0, 1]$. Meanwhile, it is a common practice to display and analyze only the positive half of the spectrum (i.e., the first $N/2$ points) because it contains all the necessary information for our entropy calculations [42,47].

Algorithm 1 Cumulative spectrum distribution entropy (CSDEn)**Input:** Time series $\{x(n)\}$ with length N , number of bin M **Output:** The value of $CSDEn(x, M)$

- 1: Obtain the spectrum $X(k)$ of time series using discrete Fourier transform according to Eq.(1).
- 2: Calculate the normalized power of each frequency component $p(k)$, $k = 0, 1, \dots, N/2 - 1$.
- 3: Apply the histogram approach to obtain the power probability pdf_m , $m = 1, 2, \dots, M$, of each spectral bin with M bins.
- 4: Based on the probability pdf_m , calculate cumulative density function $F_d(m)$ for the spectral bin.
- 5: Calculate the normalized cumulative residual function $ncrf_k$ according to Eq.(9).
- 6: Compute $CSDEn(x, M)$ according to Eq.(10).

Table 1
Parameter setting of entropy-based methods.

Method	Embedding dimension m	Tolerance threshold r	Bin number M	Class number c
CSDEn	–	–	64	–
SpEn	–	–	–	–
DistEn	2	–	128	–
DispEn	2	–	–	5
FuzzyEn	2	0.15	–	–
PerEn	4	–	–	–

The pseudocode of CSDEn is illustrated in Algorithm 1. CSDEn provides a measure of the complexity or uncertainty of a dynamic system based on the spectral content of a signal. As described in Introduction, the two spectrum distributions $D_1 = \{1, 2, 3, 4\}$ and $D_2 = \{4, 3, 2, 1\}$ have different spectrum probability for them, which is $P_1 = \{0.1, 0.2, 0.3, 0.4\}$ and $P_2 = \{0.4, 0.3, 0.2, 0.1\}$, respectively. Here, they have the same entropy value $En = 0.9232$, while the spectrum and information behind two spectrum distributions are different. In other words, the two spectral distributions or spectral shapes that represent different dynamic systems cannot be distinguished by the classical spectral distribution probability, which yields the same entropy value. On the contrary, through normalized cumulative residual function, the updated cumulative distribution $ncrf_1 = \{0.45, 0.35, 0.2, 0\}$ and $ncrf_2 = \{0.6, 0.3, 0.1, 0\}$ can be computed. Finally, we can obtain different entropy values: $CSDEn_1 = 0.7564$ and $CSDEn_2 = 0.6477$. Therefore, the introduction of the cumulative residual function to frequency-domain allows for a more refined characterization of the chaotic complexity of time series.

3. Performance verification using synthetic signals

In this section, we conducted a comprehensive performance evaluation of the proposed CSDEn method using synthetic signals. The evaluation included various aspects, including accuracy in estimating system dynamics, robustness to different data lengths and noise levels, and reliability. Here, we compared CSDEn with original SpEn and four widely used entropy methods: PerEn, DispEn, DistEn, and FuzzyEn.

Among them, SpEn does not require any parameters to be chosen, while PerEn involves embedding dimension m , DispEn involves embedding dimension m and class number c , DistEn involves embedding dimension m and bin number M , and FuzzyEn involves embedding dimension m and boundary width r . In this paper, we set the parameters of PerEn ($m = 4$), DispEn ($m = 2, c = 5$), DistEn ($m = 2, M = 128$), FuzzyEn ($m = 2, r = 0.15 * SD$) as suggested in [48,50,51]. As for CSDEn, the parameter bin number M is suggested to be set a value as the integer power of 2, and in this paper, it is set as 64. In summary, the parameter setting of the six entropy-based methods is shown in Table 1. All computations were performed on a computer equipped with Intel Core i5-9400F and 16 GB RAM, utilizing MATLAB R2018b software.

3.1. Sensitivity to signal length

In this subsection, six entropy methods were compared regarding their sensitivity to signal length. The evaluation was conducted using two types of noise signals, white Gaussian noise (WGN) and $1/f$ noise, with different sample points N . Signal lengths ranged from 200 to 4000 with a step of 200, and 100 independent signals were created for every length N . The mean and standard deviation (SD) of the results obtained from the six entropy methods were plotted against signal length, as illustrated in Fig. 1.

Overall, the results indicate that the entropy values become more robust with increasing signal length N , as evidenced by the decreasing SDs of the results. Moreover, as the signal length N increases, the mean entropy values tend to stabilize.

On the one aspect, as can be seen from Fig. 1, SpEn, DistEn, DispEn and FuzzyEn exhibit unstable behavior in small data lengths for both WGN and $1/f$ noise signals, as indicated by large error bars. This phenomenon indicates that SpEn, DistEn, DispEn, and FuzzyEn are highly dependent on the data length, which limits their use in analyzing short time series. In contrast, PerEn and CSDEn exhibit smaller error bars with less effect on data length, demonstrating greater stability. On the other aspect, it can be observed that

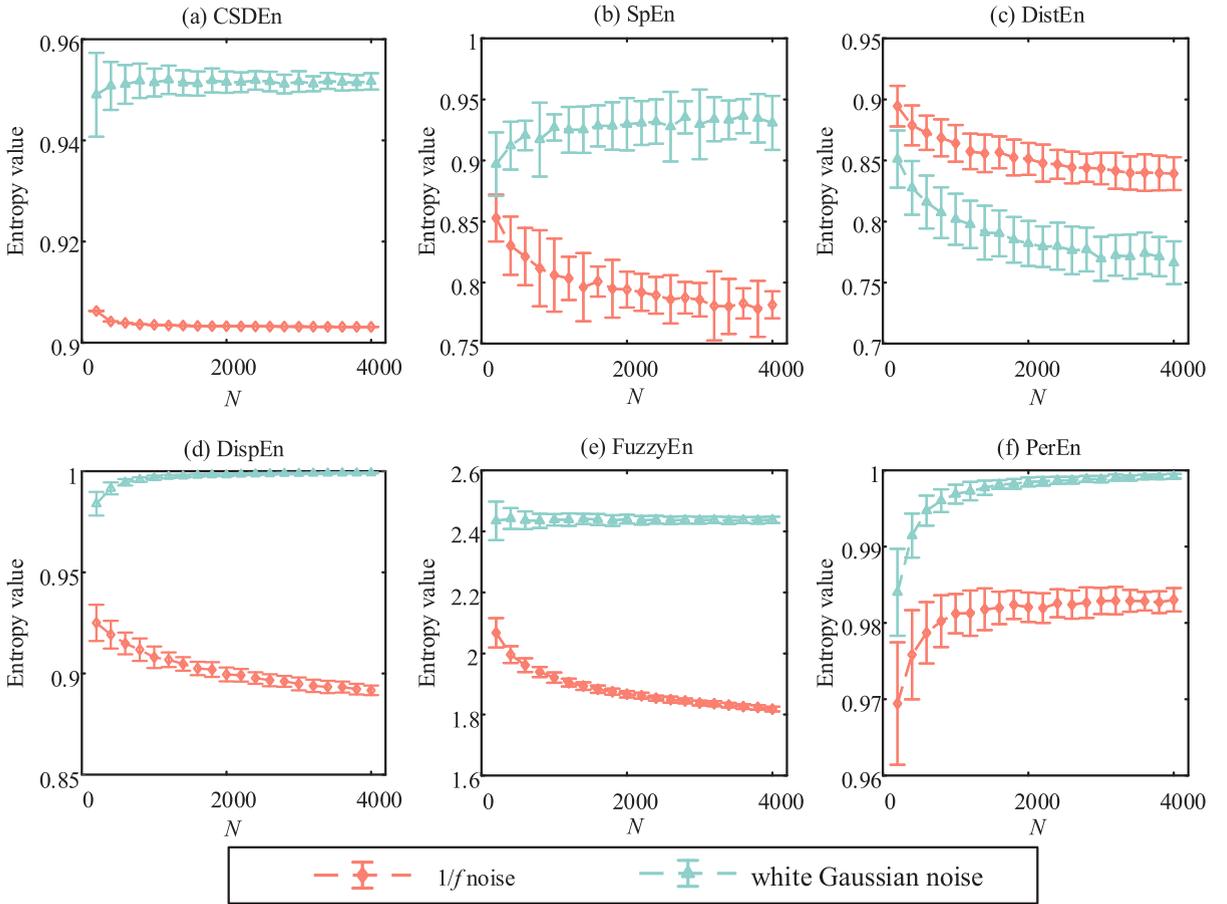


Fig. 1. Sensitivity analysis results of 100 independent signals for two types of noise signals with different sample points N using (a) CSDEn, (b) SpEn, (c) DistEn, (d) DispEn, (e) FuzzyEn, and (f) PerEn.

for data length $N < 1000$, the mean SpEn, DistEn, DispEn, and FuzzyEn values show a clear decreasing trend with poor consistency. Similarly, PerEn also displays an increasing trend. By contrast, CSDEn is much more stable, approximating a constant line, as shown in Fig. 1(a), and CSDEn values stabilize when the data length is larger than 400.

In addition, the coefficient of variation (CV) [49], which is the ratio of the standard deviation to the mean value, was also calculated to compare the dispersion of entropy values. This enables us to compare the degree of variation of one data series to another, even if the mean values are drastically different. The results are illustrated in Fig. 2. From Fig. 2(a), it can be observed that CSDEn obtains the smallest CV values, indicating that it can provide stable and consistent estimations for short signal analysis. Similarly, in Fig. 2(b), CSDEn obtains small CV values, stabilizing when the length is greater than 1000. These results suggest that as the data length increases, CSDEn, PerEn, and DispEn reach a more stable state of results earlier, whereas DistEn and SpEn have the largest CV values.

In summary, the results reveal that the proposed CSDEn approach has excellent stability with small SD and CV values among the six entropy methods, making it robust and consistent for short-time series analysis. PerEn and DispEn also show relatively small CV values, but CSDEn performs better regarding stability across different data lengths. This phenomenon suggests that CSDEn is a reliable method for analyzing short-time series data.

3.2. Dynamic change detection capability

In this subsection, the MIX process, an autoregressive model, and a chirp signal, which have been used in previous entropy studies [20], were employed to compare the dynamic change detection capability of time series.

The MIX(p) process, as a widely used dynamical system, is a combination of a sinusoidal signal of length N and independent identically distributed random noise. Specifically, $N \times p$ randomly chosen points in the sinusoidal signal are replaced with random noise, resulting in a mixed signal. Mathematically, the MIX process can be expressed as:

$$MIX(p)_i = (1 - Z_i) \times X_i + Z_i \times Y_i \quad (11)$$

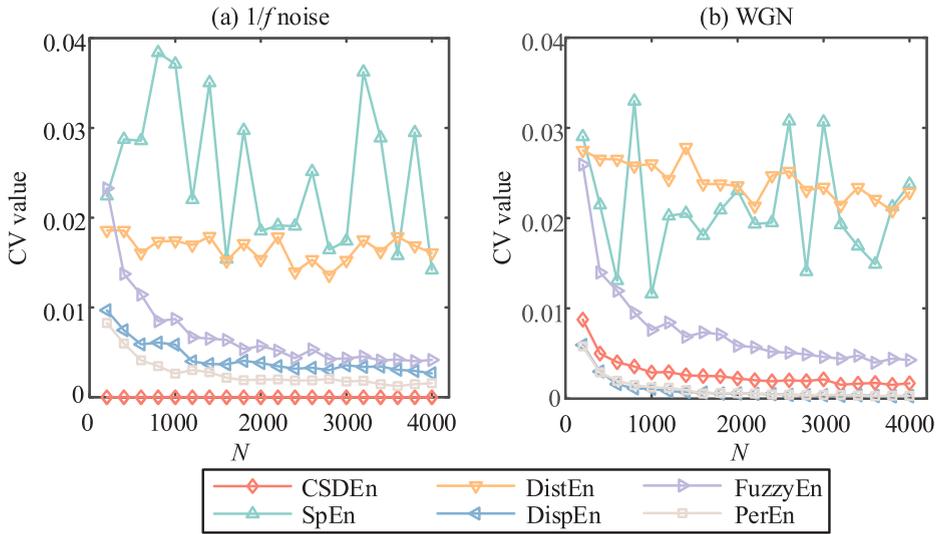


Fig. 2. The coefficient of variation (CV) values for verification results of sensitivity to signal length for (a) $1/f$ noise and (b) WGN signals.

where $X_i = \sqrt{2} \sin(2\pi i/12)$ is a sinusoidal signal, Y_i is a value uniformly distributed from $[-\sqrt{3}, \sqrt{3}]$, and Z_i is a random value taking 1 with probability p and 0 with probability $(1 - p)$. The MIX process refers to a combination of sinusoidal periodic signals and noisy signals. The process can be applied to control the complexity of the signal by adjusting the value of p . When p is small, the process is more periodic and regular, while a larger value of p leads to more stochastic and irregular behavior. The MIX process serves as a simulated scenario where the performance of entropy-based methods can be evaluated in the presence of both periodic signals and noise. Here, p was set to increase linearly from 0.01 to 0.5, which investigates the characterization capability when the periodic deterministic signal progressively turns into the stochastic signal.

The second simulated model is the first-order auto-regressive (AR(1)) model. In this study, a series of AR(1) processes with different control parameters were used to investigate the relationship between entropy and the spectral content of colored noise [49]. Mathematically, the AR(1) model can be expressed as:

$$X_{AR,i} = \varphi \times X_{AR,i-1} + \varepsilon_i \quad (12)$$

where φ is the parameter of the model, and ε_i is white noise. In this paper, $X_1 = 0.3$ and φ was set as a linearly varying parameter, increasing from 0 to 1.

Next, a chirp signal X_{chirp} was designed to study the signal frequency change detection capability, which can be expressed as:

$$X_{chirp} = \sin(2\pi \times (0.25t + 0.5) \times t) \quad (13)$$

It is noticed that the three synthetic models have a length of 20 480 points, and a sliding window of 2048 points with 75% overlap was moved along both signals so that there are total 37 samples for each synthetic model. Meanwhile, since the upper limit of FuzzyEn is not 1, FuzzyEn is normalized in this section for a better comparative analysis. The corresponding synthetic signals and analysis results of different entropy methods are shown in Fig. 3. The synthetic signals for three models are depicted in the upper row of Fig. 3. The corresponding entropy values for three models are shown in the lower row of Fig. 3.

Fig. 3(d) illustrates that all entropy methods are capable of detecting dynamic changes, with entropy increasing as the control parameter p varies for the MIX(p) process. This observation suggests a commonality in the entropy change patterns across all six methods as the MIX process undergoes a transition from periodicity to randomness. Among these methods, SpEn exhibits the least stability, characterized by pronounced fluctuations, and FuzzyEn displays a decline in some samples.

Similarly, alterations in the parameter φ induce modifications in the fundamental characteristics of the AR(1) process, encompassing transformations in its spectral attributes and autocorrelation structure. This phenomenon becomes apparent when examining Fig. 3(b), wherein parameter variations manifest as substantial fluctuations in signal amplitude. Importantly, Fig. 3(e) provides empirical evidence that all entropy methodologies, with the exception of DistEn, exhibit an aptitude for detecting and characterizing this phenomenon. It is noteworthy that, in contrast to the other entropy techniques, DistEn displays marked fluctuations without a discernible, consistent downward trend.

Moreover, Fig. 3(f) serves as an illustrative depiction of the signal's frequency modulation detection capabilities across six distinct entropy methodologies. Notably, SpEn and DistEn exhibit a limited capacity to capture variations associated with signal frequency in the chirp signal. This limitation stems from SpEn's exclusive consideration of the frequency distribution within the spectral domain, thereby disregarding specific frequency values. In contrast, when examining DispEn, there is an overall ascending trend, albeit with slight fluctuations. In the case of CSDEn, FuzzyEn, and PerEn, these entropy methods display a remarkable sensitivity to frequency-related variations within the chirp signal, outperforming SpEn, DistEn, and DispEn. Additionally, CSDEn, FuzzyEn, and PerEn curves

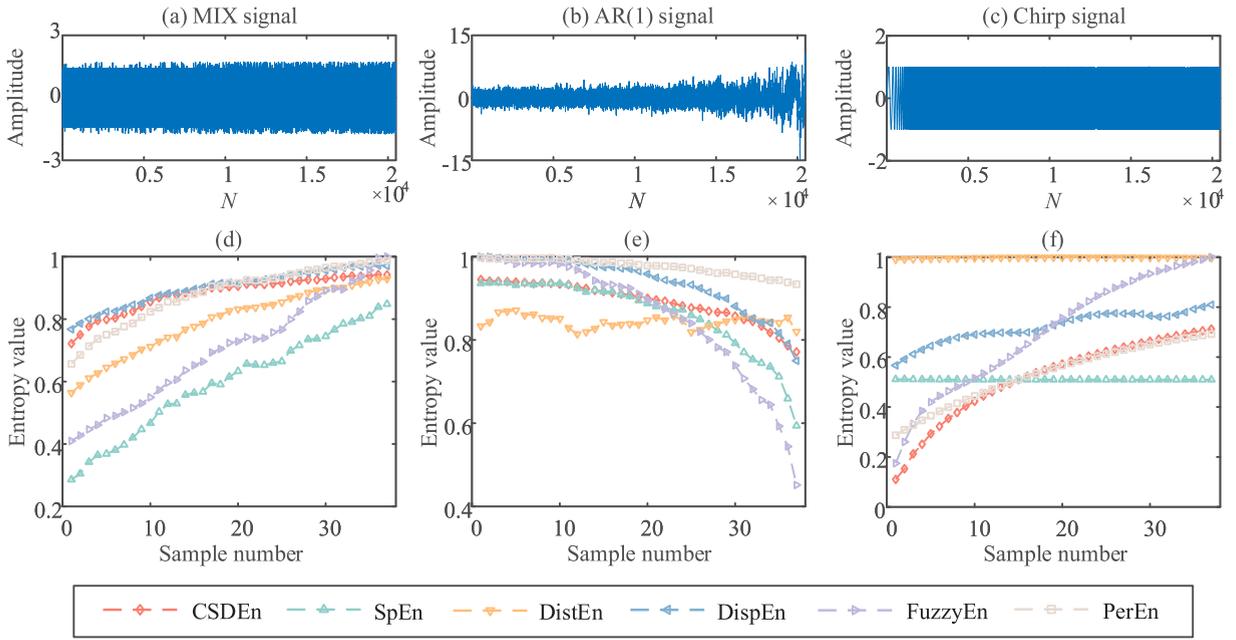


Fig. 3. Synthetic signals (upper row) and comparison of entropy values (lower row) obtained by CSDEn, SpEn, DistEn, DispEn, FuzzyEn, and PerEn using a sliding window of 2048 points with 75% overlap.

display an impressive trait of stability, evident in its consistently smoother curve throughout the progression. Consequently, CSDEn excels in its capacity to effectively characterize and delineate changes in signal frequency.

3.3. Robustness to noise

To verify the superior robustness of CSDEn to noise in non-linear and periodic signals, WGN with different signal-to-noise ratios (SNRs) is added to four MIX signals, including MIX(0.3), MIX(0.5), MIX(0.7), and MIX(1) with the signal length of 1000. The SNRs were varied from 40 dB to 5 dB, and 100 implementations of PerEn, DispEn, DistEn, FuzzyEn, SpEn, and CSDEn were performed under each SNR condition to minimize the effects of randomness. The resulting mean values and standard deviations of the entropy methods are shown in Fig. 4.

Fig. 4 offers a comprehensive illustration of our study. Entropy, serving as a metric for stochasticity, should logically exhibit higher values for MIX(p) signals with higher p values. Therefore, in theory, the entropy ranking for the four signals is as follows: $En_{MIX(0.3)} < En_{MIX(0.5)} < En_{MIX(0.7)} < En_{MIX(1)}$. Furthermore, even after the introduction of noise, this order of complexity should ideally remain intact.

At an SNR of 40 dB, the curves for DispEn, FuzzyEn, and PerEn are not consistent with the complexity arrangement of different p values. By contrast, the curves for CSDEn, SpEn, and DistEn align reasonably well with the expected complexity ranking of the MIX process. However, as noise levels escalate and SNR decreases, the distinction between the MIX process datasets diminishes. For example, in Fig. 4(c), with a noise level of 25 dB, the DistEn method exhibits a phenomenon of signal overlap. At an SNR of 5 dB, DistEn, DispEn, FuzzyEn, and PerEn lose their ability to differentiate between system complexities.

In summary, the insights drawn from Fig. 4 strongly support the assertion that CSDEn and SpEn excel in combating noise interference. These methods consistently maintain a high level of consistency across the entire SNR spectrum, ranging from 40 dB to 5 dB. These comparisons confirm that the proposed CSDEn method possesses a distinct advantage in terms of noise robustness.

3.4. Comparison of the computational complexity

The computational complexity of entropy methods is a critical aspect to consider when assessing their performance. To evaluate this, a common white noise series was used, and the signal length was varied from 100 to 1000 sample points. In each length condition, 100 implementations of entropy calculation were conducted, and the mean time cost was depicted in Fig. 5. The results indicate that the computational cost of CSDEn is comparable to PerEn, DispEn, and SpEn with the same order of magnitude, but considerably less expensive than DistEn and FuzzyEn. This finding is consistent with the fact that the computational cost of DistEn and FuzzyEn is $O(N^2)$, while the computational cost of CSDEn, PerEn, DispEn, and SpEn algorithms is $O(N)$.

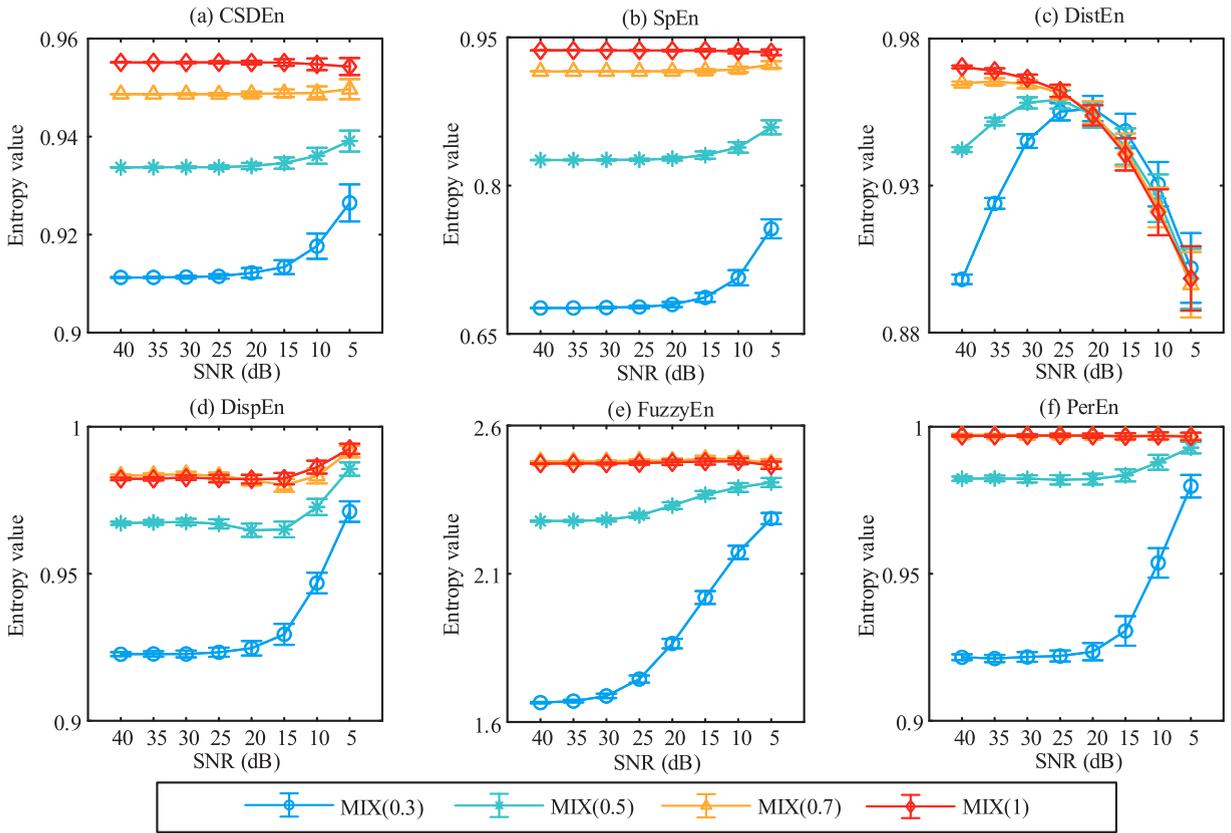


Fig. 4. Results of robustness to noise: entropy values calculated by (a) CSDEn, (b) SpEn, (c) DistEn, (d) DispEn, (e) FuzzyEn, and (f) PerEn under different SNR values.

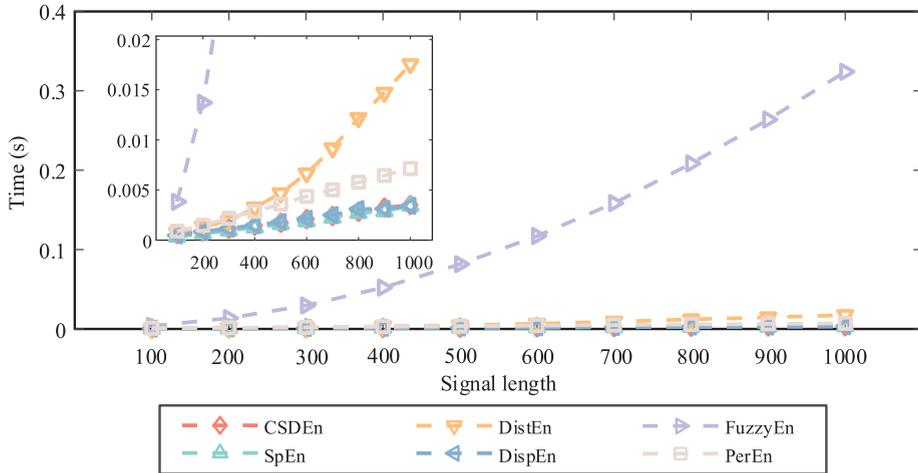


Fig. 5. Comparison of the computational time of CSDEn, SpEn, DistEn, DispEn, FuzzyEn, and PerEn for white noise series with different lengths.

4. Applications of real-world signals

The synthetic signals have been designed for performance verification in Section 3, but there is a large amount of interference in the practical working environment, so the effectiveness of the proposed method needs to be further verified in the real-world environment, which can be achieved in the experimental case study. Therefore, this section presents two case studies to demonstrate the practical applications of CSDEn in signal processing, including planetary gearbox and rolling bearing.

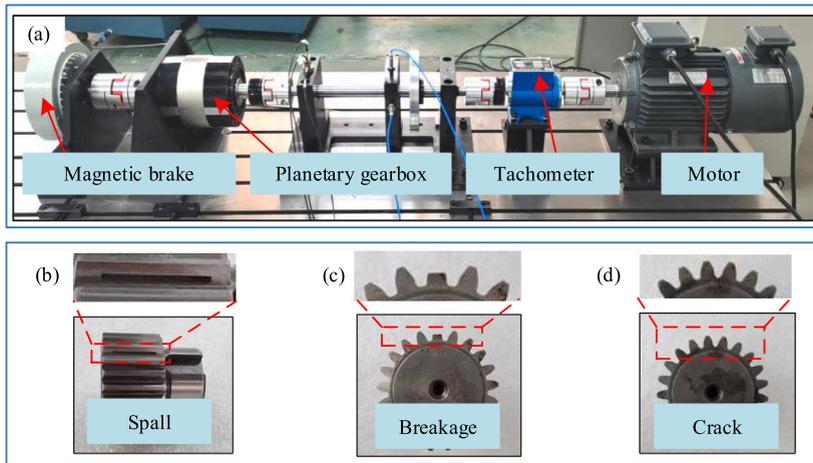


Fig. 6. The planetary gearbox system and fault gears: (a) real gearbox test rig, (b) sun gear with a spall, (c) sun gear with a breakage and (d) sun gear with a crack.

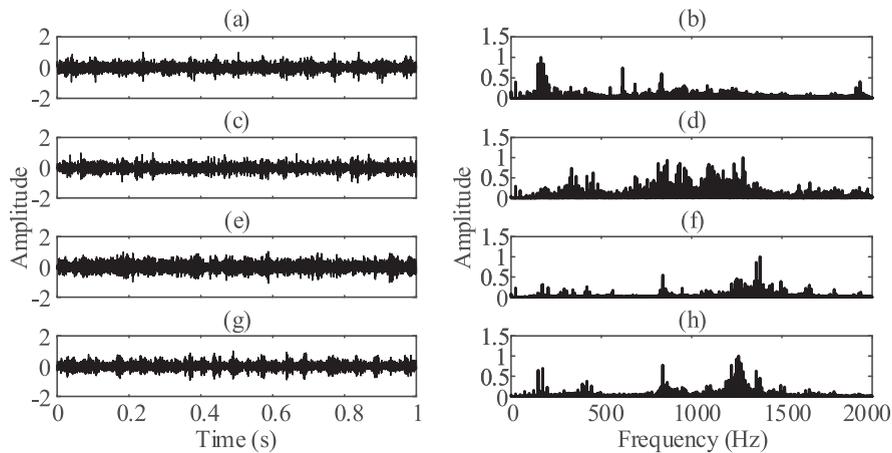


Fig. 7. The time-domain signals and corresponding normalized spectrum of four conditions for gears: (a) Normal condition, (b) spectrum of normal condition, (c) breakage fault, (d) spectrum of breakage fault, (e) crack fault, (f) spectrum of crack fault, (g) spall fault, (h) spectrum of spall fault.

4.1. Case study I: fault diagnosis of planetary gearbox

4.1.1. Data description of planetary gearbox

The first case study was carried out on a planetary gearbox system produced by WuXi HouDe Automation Meter, as shown in Fig. 6(a), which consists of a motor, planetary gearbox, tachometer, and magnetic brake. Vibration signals were acquired using an accelerometer installed on top of the gearbox casing, with a sampling frequency of 16 384 Hz. Here, the load was set as 5 N m and the rotation speed of motor was set as 2400 rpm.

In the first experimental case, four conditions of the planetary gearbox were considered, including normal condition and three typical gear failure types: sun gear with a spall, sun gear with a breakage, and sun gear with a crack. The visual representation of three failure types is displayed in Fig. 6(b)–(d). It is worth noting that each type of fault signal was normalized and sliced into 100 sub-signals. Each sub-signal contained 4096 sample points, and there were 400 samples for this case study. Fig. 7 illustrates the time-domain signals and corresponding normalized spectrum of the planetary gearbox under different fault categories and health conditions.

4.1.2. Results and analysis

In the experimental application, for comparison, seven entropy methods were conducted. It is noted that we also applied the cumulative residual distribution entropy (CRDE) [48] for feature extraction and made comparisons to highlight the strengths of cumulative spectrum distribution in proposed CSDEn method. CRDE is an enhanced variant of DistEn, drawing inspiration from cumulative residual entropy. Through this comparison, we aim to accentuate the advantages in CSDEn. The comparison results are

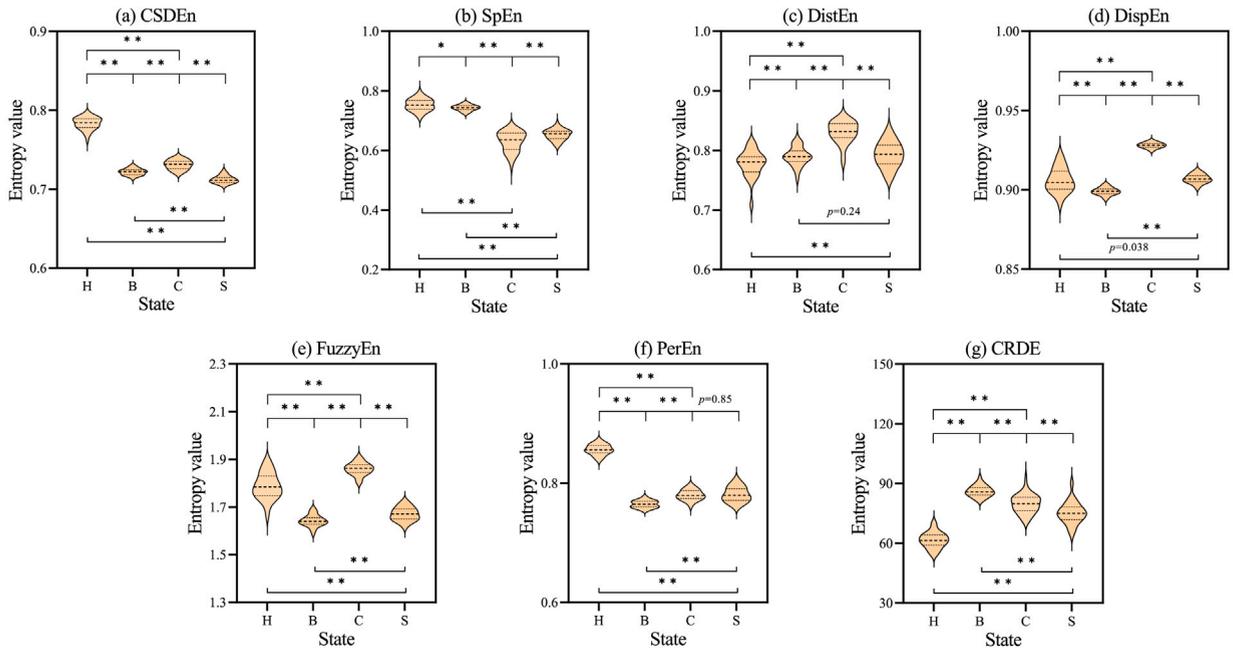


Fig. 8. Violin plots of entropy values obtained by (a) CSDEn, (b) SpEn, (c) DistEn, (d) DispEn, (e) FuzzyEn, (f) PerEn, and (g) CRDE for gear signals in different states. (H represents healthy state, B represents breakage fault, C represents crack fault and S represents spall fault). Statistical differences between any two states are analyzed using the Mann–Whitney U -test, and * means $p < 1E-3$, ** means $p < 1E-4$).

Table 2

Differences between entropy values using Mann–Whitney U -test for gear data. (H represents healthy state, B represents breakage fault, C represents crack fault and S represents spall fault).

p -value	Entropy method						
	CSDEn	SpEn	DistEn	DispEn	FuzzyEn	PerEn	CRDE
H vs. B	2.6E–34	1.8E–04	1.6E–05	3.3E–14	1.8E–33	2.6E–34	2.6E–34
H vs. C	2.6E–34	2.6E–34	4.2E–30	2.6E–34	2.5E–19	2.6E–34	2.6E–34
H vs. S	2.6E–34	2.6E–34	7.1E–06	3.8E–02	4.4E–31	2.6E–34	4.2E–32
B vs. C	8.2E–21	2.6E–34	8.2E–27	2.6E–34	2.6E–34	2.9E–25	1.4E–21
B vs. S	2.0E–25	2.6E–34	2.4E–01	1.1E–32	3.4E–11	4.9E–20	6.2E–29
C vs. S	4.7E–33	2.1E–06	5.6E–23	2.6E–34	2.6E–34	8.5E–01	3.4E–10

presented as violin plots in Fig. 8, which displays the full distribution of the entropy values and shows the median values of different entropy methods for gear signals in four states.

From Fig. 8(a), it is evident that the median values of healthy and faulty states in CSDEn exhibit significant differences, facilitating the discrimination between gear conditions. Additionally, CSDEn values in the normal state are greater than those in the faulty states, which is consistent with the fact that after fault there are periodic faulty impulses and the signal becomes less complex [9]. Furthermore, the median entropy values for the three fault states display substantial disparities, indicating that CSDEn can not only detect faults but also effectively discriminate between different types of gear faults.

Similarly, in Fig. 8(b), (e) and (g), SpEn, FuzzyEn and CRDE display discernible differences across various states. However, as illustrated in Fig. 8(c), DistEn exhibits similarities in median entropy values between spall fault and breakage fault, posing challenges for differentiation. Additionally, Fig. 8(d), (f) and (g) reveal analogous trends for DispEn and PerEn. In DispEn, the median entropy values for healthy state and spall fault are close, making distinction difficult. In the case of PerEn, the median entropy values for crack fault and spall fault exhibit similarity.

Moreover, in order to quantitatively assess the dissimilarities in entropy values among different states, we subjected all entropy values to statistical analysis using the Mann–Whitney U -test to obtain p -values for each method. A significance level of $p < 1E-3$ was employed to denote statistically significant differences between states, with more significant differences denoted by $p < 1E-4$. The statistical results are presented in Fig. 8 and Table 2.

From Fig. 8 and Table 2, it can be found that CSDEn, FuzzyEn, and CRDE demonstrate highly significant differences in distinguishing any two states, indicated by p -values far below $1E-4$. This implies a high degree of dissimilarity between any two conditions. SpEn delivers moderately strong results, with significant differences ($p \ll 1E-4$) between any two states except for H vs. B. In contrast, DistEn, DispEn, and PerEn exhibit less favorable performance, with DistEn unable to distinguish significant differences

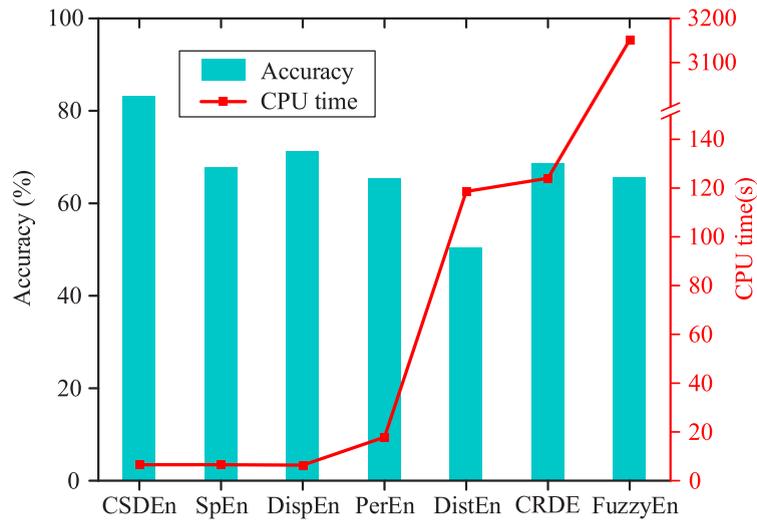


Fig. 9. Performance comparison of seven entropy algorithms for case study I.

between breakage fault and spall fault, DispEn failing to distinguish the healthy state from spall fault, and PerEn struggling to distinguish crack fault from spall fault.

Additionally, machine learning was employed to compare the recognition rates of different entropy methods. In particular, the support vector machine (SVM) algorithm [52] was utilized for classification and fault diagnosis. To assess the performance of each entropy method, we randomly selected 50 samples from each state as the training set (4×50 samples), while the remaining 50 samples were used as the test set (4×50 samples). Therefore, seven entropy methods were conducted for comparisons, and twenty trials were performed on each method to decrease random effects. The average classification results and computation time are presented in Fig. 9.

The results in Fig. 9 demonstrate that our proposed CSDEn method achieved the highest classification accuracy compared to the other methods, showcasing its superior feature extraction capability. CSDEn outperformed SpEn by incorporating the concept of bin distribution and cumulative distribution function, effectively capturing the spectrum characteristics. Additionally, the introduction of the cumulative distribution function in CRDE improved recognition rates compared to DistEn, highlighting the effectiveness of this function.

In addition to its exceptional classification accuracy, the proposed CSDEn method demonstrated notable advantages in terms of computational efficiency. It achieved the highest test classification accuracy while requiring less CPU time for feature extraction, where CPU time refers to the total time taken for the feature extraction process. The computational cost of CSDEn was found to be comparable to that of DispEn and SpEn, falling within the same order of magnitude. However, CSDEn exhibited significantly lower computational complexity compared to CRDE, PerEn, DistEn and FuzzyEn.

Specifically, CSDEn was approximately twenty times faster than DistEn and CRDE in terms of calculation efficiency, and three times faster than PerEn. This significant improvement in computational efficiency indicates that CSDEn not only outperforms other methods in feature extraction and diagnostic accuracy but also offers high-speed processing capabilities. These favorable characteristics make CSDEn well-suited for online detection requirements and present a promising approach for entropy-based feature extraction.

4.2. Case study II : fault diagnosis of rolling bearing

4.2.1. Data description of rolling bearing

The second experimental case was also conducted on the test rig manufactured by WuXi HouDe Automation Meter, as illustrated in Fig. 10(a), 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 N m. Also, 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.

In this experiment case, different bearing failure types were implemented by replacing the test bearings, and eventually, four conditions, in total, were designed. The fault bearings are shown in Fig. 10(b)–(d). In addition, each type of fault signal was sliced into 100 sub-signals and each sub-signal contains 4096 sample points. The time domain signals and the corresponding frequency spectrum of the collected vibration signals are illustrated in Fig. 11.

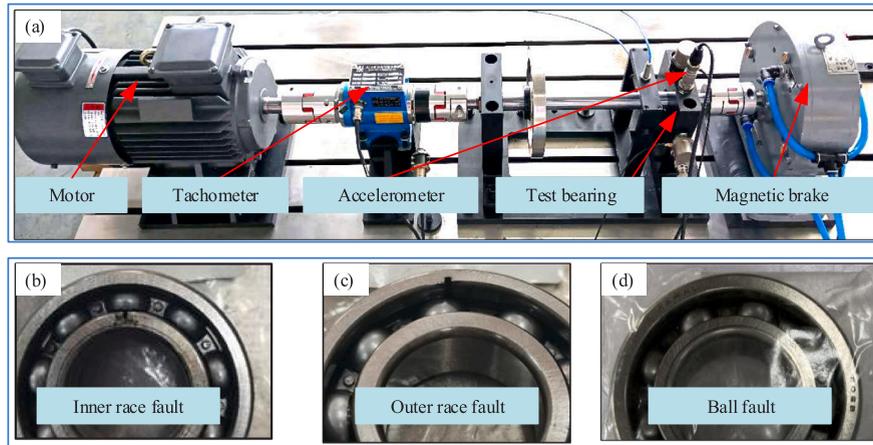


Fig. 10. The test rig of rolling bearing system and bearings with different fault types: (a) bearing test rig, (b) bearing with inner race fault, (c) bearing with outer race fault, (d) bearing with ball fault.

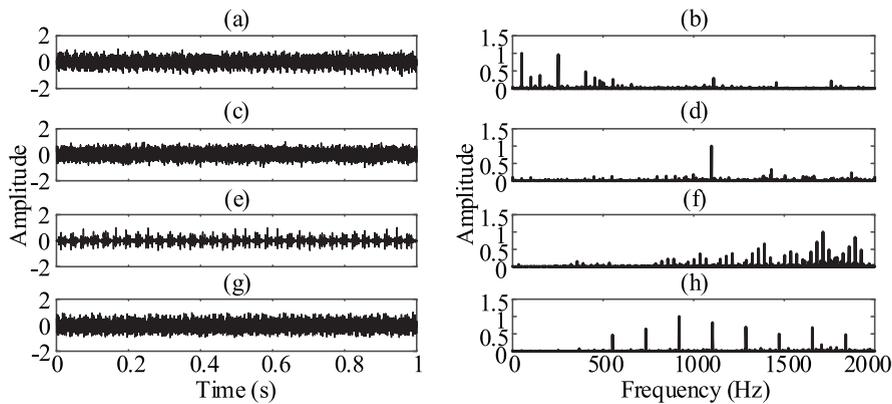


Fig. 11. The time-domain signals and corresponding normalized spectrum of four conditions for bearings: (a) Normal condition, (b) spectrum of normal condition, (c) ball fault, (d) spectrum of ball fault, (e) inner race fault, (f) spectrum of inner race fault, (g) outer race fault, (h) spectrum of outer race fault.

4.2.2. Results and analysis

Like Case study I, we conducted a comparative analysis employing the CSDEn, SpEn, DistEn, DispEn, FuzzyEn, PerEn, and CRDE methods for signal analysis. The resulting entropy values of the seven different methods under four different states are presented as violin plots in Fig. 12. Consistently, we subjected the entropy values to the Mann–Whitney U -test to obtain p -values for each method. The resultant differentiation outcomes are visually depicted in Fig. 12 and quantitatively summarized in Table 3.

As depicted in Fig. 12(a), the CSDEn method effectively distinguishes various operational states of bearings, demonstrating statistically significant differences. The associated p -values, detailed in Table 3, confirm that CSDEn exhibits significant distinctions between any two states, with p -values significantly below $1E-4$. This underscores a high degree of differentiation between any pair of conditions. Concurrently, the other comparative entropy methods also showcase notable proficiency in distinguishing between different states of bearings, yielding statistically significant distinctions with p -values well below $1E-4$.

In case study II, we conducted a similar comparison of different entropy methods in diagnostic accuracy, and to mitigate the influence of randomness, we performed twenty trials for each method. Fig. 13 provides a visual representation of the comparison results and computation time. Consistent with case study I, the CSDEn method achieved the highest diagnosis accuracy, exceeding 98%. Moreover, the line chart illustrates that the CSDEn approach not only outperformed other methods in terms of diagnostic accuracy but also exhibited excellent computational efficiency. The computational cost of CSDEn was found to be comparable to that of DispEn and SpEn, falling within the same order of magnitude. However, CSDEn demonstrated significantly lower computational complexity compared to the other entropy methods, like PerEn, DistEn and CRDE.

These results emphasize the merits of the proposed CSDEn method in both fault feature extraction and calculation efficiency, as observed in case study II. The CSDEn method reduces the computation time required for feature extraction while extracting more reliable information for fault diagnosis. By striking a balance between computational efficiency and diagnostic performance, the proposed CSDEn method offers valuable contributions to fault diagnosis in practical applications.

Table 3

Differences between entropy values using Mann–Whitney U -test for bearing data. (H represents healthy state, BF represents ball fault, IF represents inner race fault and OF represents outer race fault).

p -value	Entropy method						
	CSDEn	SpEn	DistEn	DispEn	FuzzyEn	PerEn	CRDE
H vs. BF	2.6E-34	6.7E-32	2.6E-34	2.6E-34	6.1E-31	2.6E-34	2.6E-34
H vs. IF	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34
H vs. OF	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34
BF vs. IF	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	1.1E-26	2.6E-34
BF vs. OF	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	3.2E-06	2.6E-34
IF vs. OF	2.6E-34	2.6E-34	2.6E-34	2.6E-34	2.6E-34	5.9E-27	2.6E-34

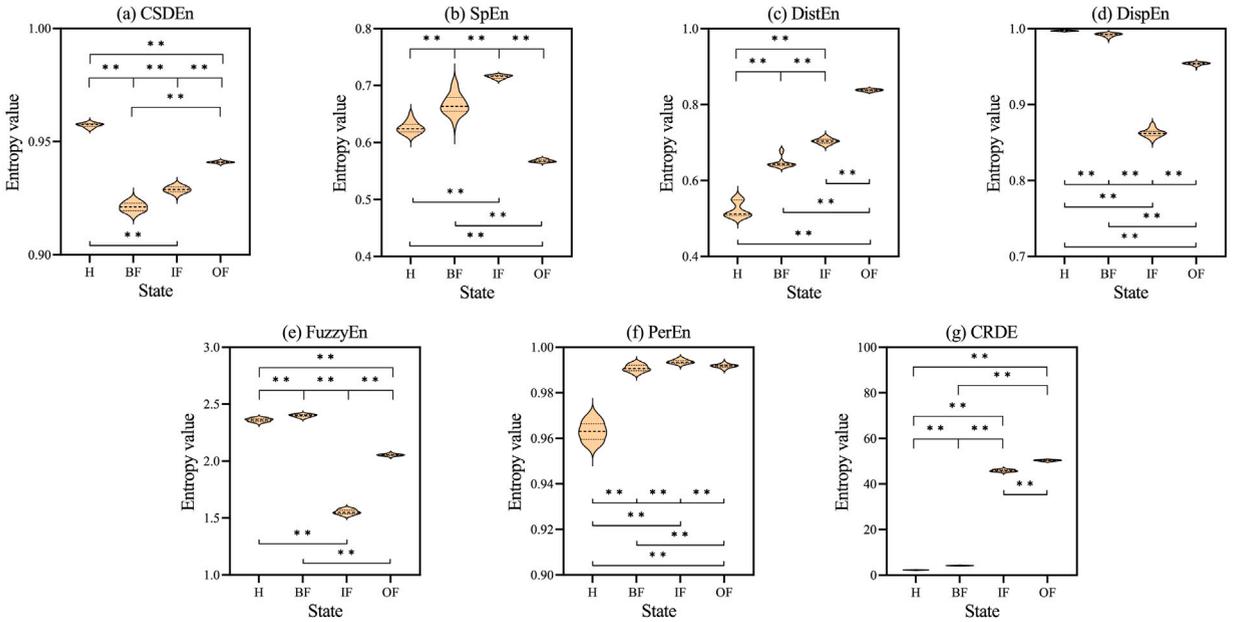


Fig. 12. Violin plots of entropy values obtained by (a) CSDEn, (b) SpEn, (c) DistEn, (d) DispEn, (e) FuzzyEn, (f) PerEn, and (g) CRDE for bearing signals in different states. (H represents healthy state, BF represents ball fault, IF represents inner race fault and OF represents outer race fault. Statistical differences between any two states are analyzed using the Mann–Whitney U -test, and ** means $p < 1E-4$).

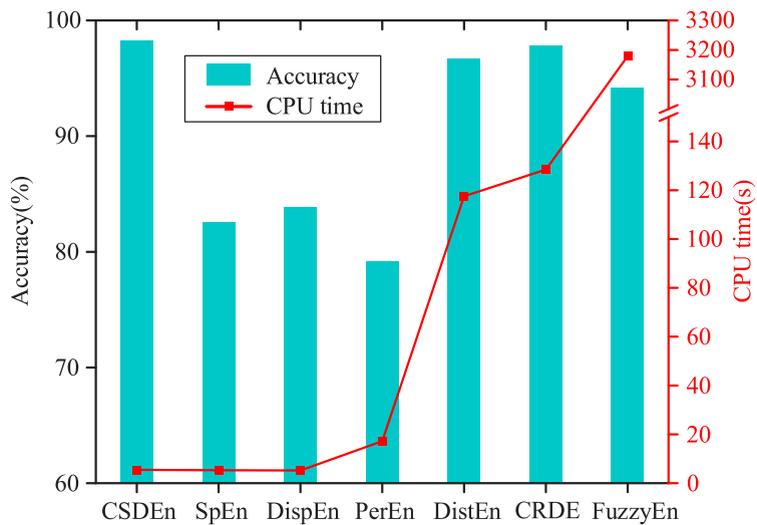


Fig. 13. Performance comparison of seven entropy algorithms for case study II.

Table 4
Parameters of five classification methods.

Method	Parameters	Value
SVM	Kernel function	RBF
	Kernel parameter	0.5
	Tradeoff parameter	1
RF	Estimation mode	Out-of-bag
	Number of trees	50
k-NN	Number of nearest neighbors	5
	Distance metric	Euclidean
ELM	Activation function	Sigmoidal function
	Number of hidden layer neurons	100
LR	Regularization parameter	0.001

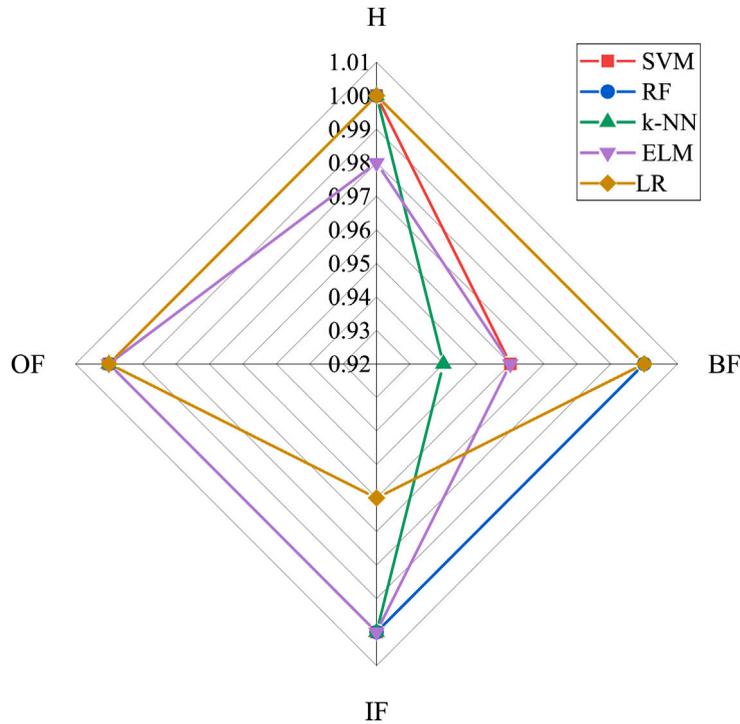


Fig. 14. Classification results of five different classifiers for case study II.

In addition, to assess the versatility of the proposed method, we also employed five different classifiers: support vector machine (SVM), random forests (RF), k-nearest neighbors algorithm (k-NN), extreme learning machine (ELM), and logistic regression (LR). The parameter values for each classifier are listed in Table 4. The feature sets used for classification were obtained using the CSDEn method. The final verification results are presented in Fig. 14.

The radar diagram depicted in Fig. 14 distinctly demonstrates that features derived from CSDEn consistently yield recognition rates exceeding 94% across a range of classifiers. The findings further establish the effectiveness and robustness of the proposed CSDEn-based features in classification tasks. The consistently high recognition rates obtained across multiple classifiers demonstrate the generalizability and discriminative power of these features. This aspect holds significant practical implications, indicating that the CSDEn-based features can be successfully applied to diverse classification algorithms, thus yielding reliable and consistent results in various real-world scenarios.

It is worth noting that in practical machine fault diagnosis applications, the choice of classifier may depend on specific dataset characteristics and the nature of the problem. Therefore, the generalizability of CSDEn-based features to different classifiers enhances their suitability for a wide range of diagnostic tasks, allowing practitioners to select the most appropriate classifier based on the specific requirements and characteristics of their applications.

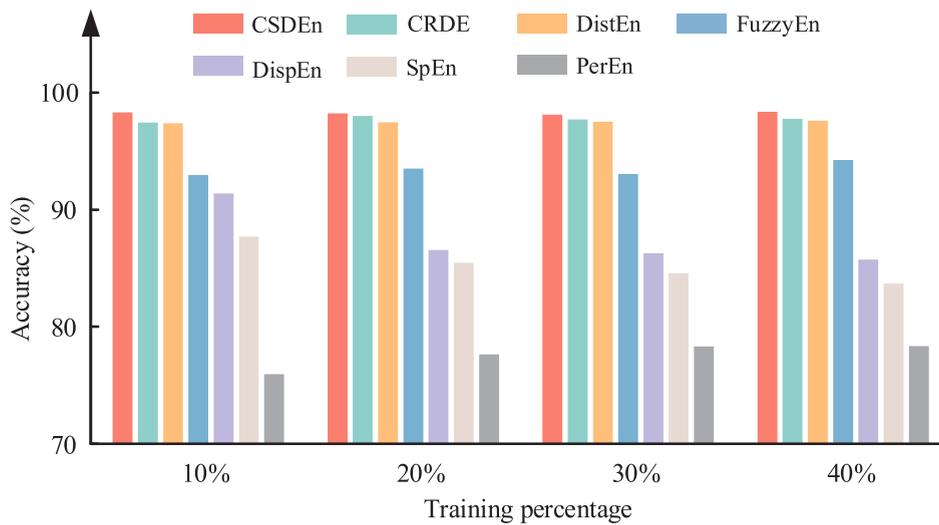


Fig. 15. Diagnostic accuracies for entropy methods with different training percentage.

4.3. Diagnosis performance under challenges

4.3.1. Fault diagnosis with small sample

To assess the performance of the proposed method with a limited sample size, we conducted a comparative analysis using different proportions of samples for training. The dataset from case study II was utilized for this evaluation. Proportions of 10%, 20%, 30%, and 40% of the samples were selected for training, while the remaining samples were used for testing. It is noted that SVM was used for classification, following the same procedure as in previous case studies. The classification results for different methods and varying proportions of the training set are presented in Fig. 15.

The results in Fig. 15 clearly show that the proposed CSDEn method outperforms the other five entropy methods in bearing fault diagnosis. Remarkably, even with only 10% of training samples, the CSDEn method achieves a recognition rate of 98% or higher. This demonstrates the method's robustness and effectiveness in fault diagnosis, even with a limited number of training samples. The ability to achieve high recognition rates with such a small sample size showcases the efficiency and accuracy of the proposed method in handling data scarcity scenarios.

4.3.2. Robustness against noises

In real-world industrial applications, noise has a significant impact on the performance of feature extraction and diagnosis models. To evaluate the robustness of the proposed CSDEn method against noise, we injected varying degrees of Gaussian white noise into the signals from case study II. The diagnosis results under different noise levels are summarized in Fig. 16.

Fig. 16 clearly demonstrates the robustness of the proposed CSDEn method in real industrial applications with noise. As the signal-to-noise ratio (SNR) decreases, the diagnosis accuracy of all methods also decreases. However, the proposed CSDEn consistently outperforms the other methods in terms of diagnosis accuracy across different levels of Gaussian white noise. Even with decreasing SNR values, CSDEn achieves high average identification accuracies, reaching 82% at SNR levels of -1 dB. This can be attributed to the comprehensive information captured through spectrum distribution, enabling CSDEn to provide more accurate fault diagnosis even in low SNR conditions. The robustness and resilience of the proposed method demonstrate its ability to maintain excellent diagnostic performance in challenging noisy environments. The results also highlight the significant improvement of CSDEn over the original SpEn, particularly when signals are contaminated by noise. This further validates the effectiveness and superiority of CSDEn in enhancing diagnostic accuracy and robustness, making it a reliable choice for fault diagnosis in noisy environments.

Moreover, random impulse noise is a common occurrence in measured signals, often attributed to external knocks on the bearing housing, electromagnetic interferences, measurement errors, or defects in the hardware system [53,54]. In response to this, we also conducted experiments where we introduced varying numbers of random impulses into the signals from case study II. It is noted that the amplitude of random impulses is set to twice the maximum value of the original signal. The results are presented in Fig. 17.

It is evident from the figure that both CSDEn and FuzzyEn exhibit robustness against random impulse noise, maintaining a high recognition rate even as the level of impulse noise increases. Both methods consistently achieve recognition rates above 90%, showcasing their resilience to this type of interference. In contrast, the recognition rates of CRDE and DistEn show a notable decline as the level of impulse noise increases. These methods are less robust in the presence of such noise, with their recognition rates decreasing significantly. While the recognition rates of SpEn, DispEn and PerEn in our study remain relatively stable, it is important to note that their overall performance is not as strong as CSDEn and FuzzyEn.

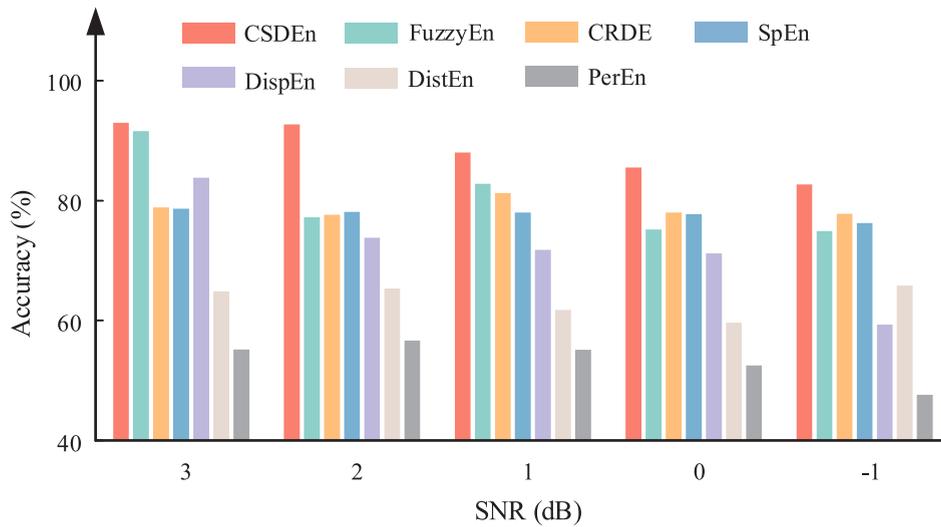


Fig. 16. Diagnostic accuracies for entropy methods under different degrees of white noise in signals.

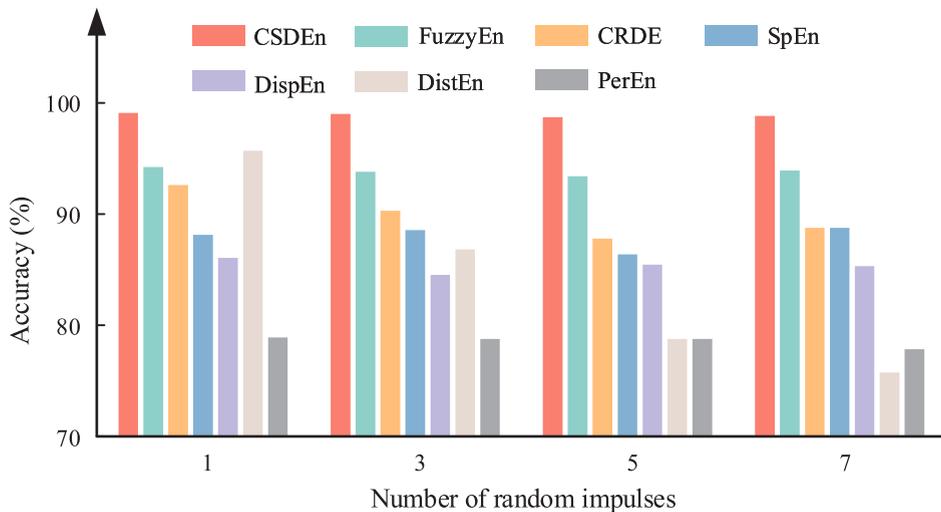


Fig. 17. Diagnostic accuracies for entropy methods under different degrees of random impulse noise in signals.

4.4. Parameter sensitivity analysis

In this subsection, we delved into the impact of the parameter M on the performance of CSDEn. We conducted a series of experiments using data from both case studies I and II while varying the parameter M . The entropy features generated with different M values were fed into the SVM classifier, following the same procedure as in previous case studies. Each parameter setting was subjected to twenty random trials, and the resulting average accuracies and calculational time for a single sample are presented in Table 5.

Analyzing the data from case study I, we observed a minor fluctuation in the recognition rate, oscillating between 82% and 85%. In contrast, the case study II data exhibited a similarly slight fluctuation, with recognition rates varying within the range of 97% to 99%. These empirical findings highlight a crucial conclusion: the choice of bin number M exerts a restricted influence on the ultimate diagnostic results. However, it is worth noting that as the parameter M increases, there is a modest increase in computational time. This analysis reinforces CSDEn’s exceptional stability, demonstrating its capacity to consistently maintain diagnostic performance across a spectrum of M values, further underscoring its robustness.

Table 5
Diagnostic performance under different parameter M for CSDEn.

Indexes	Case study I					Case study II				
	M = 32	M = 64	M = 128	M = 256	M = 512	M = 32	M = 64	M = 128	M = 256	M = 512
Accuracy (%)	84.07	83.15	82.93	83.95	82.65	97.82	98.25	98.17	98.60	98.45
Time (s)	0.0133	0.0133	0.0134	0.0138	0.0153	0.0131	0.0131	0.0139	0.0142	0.0152

5. Conclusions

This paper represents a novel entropy measure, called cumulative spectrum distribution entropy (CSDEn) that overcomes the limitations of traditional entropy-based methods, especially spectral entropy, for fault diagnosis in rotating machinery. By considering both frequency probability and frequency values in the spectrum domain, CSDEn is able to capture the frequency-domain information of fault features and detect dynamic changes in non-linear signals with high computing efficiency and low noise sensitivity. Moreover, experimental results show that CSDEn outperforms other widely used entropy measures in diagnosing different bearing and gear working states, demonstrating its superior performance in fault diagnosis of rotating machinery. Our CSDEn method offers unique contributions and advantages, particularly in terms of high computational efficiency, strong robustness to noise, and its almost-parameter-free nature.

In this preliminary study, we have evaluated the effectiveness of the proposed CSDEn method in diagnosing faults under constant load torque and speed conditions. However, its performance under variable speed and load torque conditions remains unknown. Therefore, as part of our future work, we plan to investigate and test the effectiveness of the CSDEn method in variable working conditions, specifically by combining it with order-tracking techniques. This will provide insights into the method's applicability in scenarios with varying operating conditions. Furthermore, we will expand our discussion on the incorporation of cumulative spectrum distribution entropy and other signal preprocessing methods. By exploring the combination of these approaches, we aim to further improve the accuracy and robustness of fault diagnosis in complex systems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgment

The research is supported by the National Natural Science Foundation of China under Grant 12172290 and 52250410345.

References

- [1] J. Zheng, W. Ying, H. Pan, K. Feng, Holo-hilbert square spectral analysis: A new fault diagnosis tool for rotating machinery health management, *Mech. Syst. Signal Process.* 189 (2023) 110069, <http://dx.doi.org/10.1016/j.ymssp.2022.110069>.
- [2] Q. Qian, Y. Qin, J. Luo, Y. Wang, F. Wu, Deep discriminative transfer learning network for cross-machine fault diagnosis, *Mech. Syst. Signal Process.* 186 (2023) 109884, <http://dx.doi.org/10.1016/j.ymssp.2022.109884>.
- [3] S. Jia, Y. Li, X. Wang, D. Sun, Z. Deng, Deep causal factorization network: A novel domain generalization method for cross-machine bearing fault diagnosis, *Mech. Syst. Signal Process.* 192 (2023) 110228, <http://dx.doi.org/10.1016/j.ymssp.2023.110228>.
- [4] C. Yang, M. Jia, Z. Li, M. Gabbouj, Enhanced hierarchical symbolic dynamic entropy and maximum mean and covariance discrepancy-based transfer joint matching with Welsh loss for intelligent cross-domain bearing health monitoring, *Mech. Syst. Signal Process.* 165 (2022) 108343, <http://dx.doi.org/10.1016/j.ymssp.2021.108343>.
- [5] M. Diaz, P.-E. Charbonnel, L. Chamoin, Merging experimental design and structural identification around the concept of modified constitutive relation error in low-frequency dynamics for enhanced structural monitoring, *Mech. Syst. Signal Process.* 197 (2023) 110371, <http://dx.doi.org/10.1016/j.ymssp.2023.110371>.
- [6] S. Sharma, S. Tiwari, 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.* 171 (2022) 108909, <http://dx.doi.org/10.1016/j.ymssp.2022.108909>.
- [7] A.S. Minhas, P. Kankar, N. Kumar, S. Singh, Bearing fault detection and recognition methodology based on weighted multiscale entropy approach, *Mech. Syst. Signal Process.* 147 (2021) 107073, <http://dx.doi.org/10.1016/j.ymssp.2020.107073>.
- [8] K. Noman, Y. Li, G. Wen, A.U. Patwari, S. Wang, Continuous monitoring of rolling element bearing health by nonlinear weighted squared envelope-based fuzzy entropy, *Struct. Health Monit.* (2023) 14759217231163090, <http://dx.doi.org/10.1177/14759217231163090>.
- [9] Y. Li, S. Wang, Y. Yang, Z. Deng, Multiscale symbolic fuzzy entropy: An entropy denoising method for weak feature extraction of rotating machinery, *Mech. Syst. Signal Process.* 162 (2022) 108052, <http://dx.doi.org/10.1016/j.ymssp.2021.108052>.
- [10] K. Noman, B. Hou, D. Wang, Y. Li, S. Wang, Weighted squared envelope diversity entropy as a nonlinear dynamic prognostic measure of rolling element bearing, *Nonlinear Dynam.* 111 (7) (2023) 6605–6620, <http://dx.doi.org/10.1007/s11071-022-08157-0>.
- [11] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.* 27 (3) (1948) 379–423, <http://dx.doi.org/10.1002/j.1538-7305.1948.tb01338.x>.

- [12] Y. Li, B. Geng, S. Jiao, Dispersion entropy-based Lempel-Ziv complexity: A new metric for signal analysis, *Chaos Solitons Fractals* 161 (2022) 112400, <http://dx.doi.org/10.1016/j.chaos.2022.112400>.
- [13] R. Yan, R.X. Gao, Approximate Entropy as a diagnostic tool for machine health monitoring, *Mech. Syst. Signal Process.* 21 (2) (2007) 824–839, <http://dx.doi.org/10.1016/j.ymssp.2006.02.009>.
- [14] M. Kedadouché, M. Thomas, A. Tahan, R. Guilbault, Nonlinear parameters for monitoring gear: Comparison between Lempel-Ziv, approximate entropy, and sample entropy complexity, *Shock Vib.* 2015 (2015) <http://dx.doi.org/10.1155/2015/959380>.
- [15] R. Yan, Y. Liu, R.X. Gao, Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines, *Mech. Syst. Signal Process.* 29 (2012) 474–484, <http://dx.doi.org/10.1016/j.ymssp.2011.11.022>.
- [16] W. Chen, J. Zhuang, W. Yu, Z. Wang, Measuring complexity using FuzzyEn, ApEn, and SampEn, *Med. Eng. Phys.* 31 (1) (2009) 61–68, <http://dx.doi.org/10.1016/j.medengphy.2008.04.005>.
- [17] P. Li, C. Liu, K. Li, D. Zheng, C. Liu, Y. Hou, Assessing the complexity of short-term heartbeat interval series by distribution entropy, *Med. Biol. Eng. Comput.* 53 (1) (2015) 77–87, <http://dx.doi.org/10.1007/s11517-014-1216-0>.
- [18] Y. Li, S. Jiao, B. Geng, Refined composite multiscale fluctuation-based dispersion Lempel–Ziv complexity for signal analysis, *ISA Trans.* (2022) <http://dx.doi.org/10.1016/j.isatra.2022.06.040>.
- [19] X. Wang, S. Si, Y. Li, Multiscale diversity entropy: A novel dynamical measure for fault diagnosis of rotating machinery, *IEEE Trans. Ind. Inform.* (2020) 1, <http://dx.doi.org/10.1109/TII.2020.3022369>.
- [20] M. Rostaghi, M.M. Khatibi, M.R. Ashory, H. Azami, Fuzzy dispersion entropy: A nonlinear measure for signal analysis, *IEEE Trans. Fuzzy Syst.* (2021) 1, <http://dx.doi.org/10.1109/TFUZZ.2021.3128957>.
- [21] B. Zhang, P. Shang, Transition permutation entropy and transition dissimilarity measure: Efficient tools for fault detection of railway vehicle systems, *IEEE Trans. Ind. Inform.* (2021) 1, <http://dx.doi.org/10.1109/TII.2021.3088415>.
- [22] X. Wang, L. Liu, Concentric diversity entropy: A high flexible feature extraction tool for identifying fault types with different structures, *Mech. Syst. Signal Process.* 171 (2022) 108934, <http://dx.doi.org/10.1016/j.ymssp.2022.108934>.
- [23] Y. Li, S. Wang, N. Li, Z. Deng, Multiscale symbolic diversity entropy: a novel measurement approach for time-series analysis and its application in fault diagnosis of planetary gearboxes, *IEEE Trans. Ind. Inform.* 18 (2) (2021) 1121–1131, <http://dx.doi.org/10.1109/TII.2021.3082517>.
- [24] S. Wang, Y. Li, S. Si, K. Noman, Enhanced hierarchical symbolic sample entropy: Efficient tool for fault diagnosis of rotating machinery, *Struct. Health Monit.* 22 (3) (2023) 1927–1940, <http://dx.doi.org/10.1177/14759217221116417>.
- [25] C. Li, K. Noman, Z. Liu, K. Feng, Y. Li, Optimal symbolic entropy: An adaptive feature extraction algorithm for condition monitoring of bearings, *Inf. Fusion* (2023) 101831, <http://dx.doi.org/10.1016/j.inffus.2023.101831>.
- [26] M. Costa, A.L. Goldberger, C.K. Peng, Multiscale entropy analysis of complex physiologic time series, *Phys. Rev. Lett.* 89 (6) (2002) 068102, <http://dx.doi.org/10.1103/PhysRevLett.89.068102>.
- [27] R. Zhou, X. Wang, J. Wan, N. Xiong, EDM-fuzzy: An euclidean distance based multiscale fuzzy entropy technology for diagnosing faults of industrial systems, *Ieee Trans. Ind. Inform.* 17 (6) (2021) 4046–4054, <http://dx.doi.org/10.1109/TII.2020.3009139>.
- [28] Z. Wang, L. Yao, G. Chen, J. Ding, Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals, *ISA Trans.* 114 (2021) 470–484, <http://dx.doi.org/10.1016/j.isatra.2020.12.054>.
- [29] J. Zheng, H. Pan, S. Yang, J. Cheng, Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis, *Mech. Syst. Signal Process.* 99 (2018) 229–243, <http://dx.doi.org/10.1016/j.ymssp.2017.06.011>.
- [30] Y. Li, X. Wang, Z. Liu, X. Liang, S. Si, The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review, *IEEE Access* 6 (2018) 66723–66741, <http://dx.doi.org/10.1109/ACCESS.2018.2873782>.
- [31] I. Rezek, S. Roberts, Stochastic complexity measures for physiological signal analysis, *IEEE Trans. Biomed. Eng.* 45 (9) (1998) 1186–1191, <http://dx.doi.org/10.1109/10.709563>.
- [32] P.-Y. Xiong, H. Jahanshahi, R. Alcaraz, Y.-M. Chu, J.F. Gómez-Aguilar, F.E. Alsaadi, Spectral entropy analysis and synchronization of a multi-stable fractional-order chaotic system using a novel neural network-based chattering-free sliding mode technique, *Chaos Solitons Fractals* 144 (2021) 110576, <http://dx.doi.org/10.1016/j.chaos.2020.110576>.
- [33] Y.N. Pan, J. Chen, X.L. Li, Spectral entropy: A complementary index for rolling element bearing performance degradation assessment, *Proc. Inst. Mech. Eng. C* 223 (5) (2009) 1223–1231, <http://dx.doi.org/10.1243/09544062JMES1224>.
- [34] Z. Huo, M. Martínez-García, Y. Zhang, R. Yan, L. Shu, Entropy measures in machine fault diagnosis: Insights and applications, *IEEE Trans. Instrum. Meas.* 69 (6) (2020) 2607–2620, <http://dx.doi.org/10.1109/TIM.2020.2981220>.
- [35] J.-l. Shen, J.-w. Hung, L.-s. Lee, Robust entropy-based endpoint detection for speech recognition in noisy environments, in: *ICSLP*, Vol. 98, 1998, pp. 232–235, <http://dx.doi.org/10.21437/ICSLP.1998-527>.
- [36] A. Vakkuri, A. Yli-Hankala, P. Talja, S. Mustola, H. Tolvanen-Laakso, T. Sampson, H. Viertiö-Oja, Time-frequency balanced spectral entropy as a measure of anesthetic drug effect in central nervous system during sevoflurane, propofol, and thiopental anesthesia, *Acta Anaesthesiol. Scand.* 48 (2) (2004) 145–153, <http://dx.doi.org/10.1111/j.0001-5172.2004.00323.x>.
- [37] H. Su, D. Chen, G.-J. Pan, Z. Zeng, Identification of network topology variations based on spectral entropy, *IEEE Trans. Cybern.* 52 (10) (2021) 10468–10478, <http://dx.doi.org/10.1109/TCYB.2021.3070080>.
- [38] Y. Peng, K. Sun, S. He, A discrete memristor model and its application in Hénon map, *Chaos Solitons Fractals* 137 (2020) 109873, <http://dx.doi.org/10.1016/j.chaos.2020.109873>.
- [39] C. Yang, K. Zhou, J. Liu, Q. Xu, Power spectral entropy-based graph construction for rotating machinery diagnosis using multi-sensor, in: *2021 Global Reliability and Prognostics and Health Management (PHM-Nanjing)*, 2021, pp. 1–5, <http://dx.doi.org/10.1109/PHM-Nanjing52125.2021.9612863>.
- [40] Y. Mi, A. Lin, D. Gu, X. Zhang, X. Huang, Bubble transfer spectral entropy and its application in epilepsy EEG analysis, *Commun. Nonlinear Sci. Numer. Simul.* 110 (2022) 106294, <http://dx.doi.org/10.1016/j.cnsns.2022.106294>.
- [41] H. Yu, H. Li, B. Xu, Rolling bearing degradation state identification based on LCD relative spectral entropy, *J. Fail. Anal. Prev.* 16 (4) (2016) 655–666, <http://dx.doi.org/10.1007/s11668-016-0133-y>.
- [42] P.P.A. Staniczenko, C.F. Lee, N.S. Jones, Rapidly detecting disorder in rhythmic biological signals: A spectral entropy measure to identify cardiac arrhythmias, *Phys. Rev. E* 79 (1) (2009) 011915, <http://dx.doi.org/10.1103/PhysRevE.79.011915>.
- [43] X. Mao, P. Shang, J. Wang, Y. Ma, Characterizing time series by extended complexity-entropy curves based on Tsallis, Renyi, and power spectral entropy, *Chaos* 28 (11) (2018) 113106, <http://dx.doi.org/10.1063/1.5038758>.
- [44] Y. Zhang, P. Shang, The complexity-entropy causality plane based on multiscale power spectrum entropy of financial time series, *Chaos* 28 (12) (2018) 123120, <http://dx.doi.org/10.1063/1.5054714>.
- [45] J. Zheng, W. Ying, J. Tong, Y. Li, Multiscale three-dimensional Holo–Hilbert spectral entropy: a novel complexity-based early fault feature representation method for rotating machinery, *Nonlinear Dynam.* (2023) 1–22, <http://dx.doi.org/10.1007/s11071-023-08392-z>.
- [46] M. Rao, Y.M. Chen, B.C. Vemuri, F. Wang, Cumulative residual entropy: A new measure of information, *Ieee Trans. Inf. Theory* 50 (6) (2004) 1220–1228, <http://dx.doi.org/10.1109/TIT.2004.828057>.
- [47] P.-Y. Xiong, H. Jahanshahi, R. Alcaraz, Y.-M. Chu, J. Gómez-Aguilar, F.E. Alsaadi, Spectral entropy analysis and synchronization of a multi-stable fractional-order chaotic system using a novel neural network-based chattering-free sliding mode technique, *Chaos Solitons Fractals* 144 (2021) 110576.

- [48] X. Mao, P. Shang, A.C. Yang, C.-K. Peng, Multiscale cumulative residual distribution entropy and its applications on heart rate time series, *Nonlinear Dynam.* 101 (4) (2020) 2357–2368, <http://dx.doi.org/10.1007/s11071-020-05934-7>.
- [49] Y. Wang, Y. Xu, M. Liu, Y. Guo, Y. Wu, C. Chen, W. Chen, Cumulative residual symbolic dispersion entropy and its multiscale version: Methodology, verification, and application, *Chaos Solitons Fractals* 160 (2022) 112266, <http://dx.doi.org/10.1016/j.chaos.2022.112266>.
- [50] M. Rostaghi, H. Azami, Dispersion Entropy: A measure for time-series analysis, *IEEE Signal Process. Lett.* 23 (5) (2016) 610–614, <http://dx.doi.org/10.1109/LSP.2016.2542881>.
- [51] W. Chen, Z. Wang, H. Xie, W. Yu, Characterization of surface EMG signal based on fuzzy entropy, *IEEE Trans. Neural Syst. Rehabil. Eng.* 15 (2) (2007) 266–272, <http://dx.doi.org/10.1109/TNSRE.2007.897025>.
- [52] Z. Wang, L. Yao, Y. Cai, J. Zhang, Mahalanobis semi-supervised mapping and beetle antennae search based support vector machine for wind turbine rolling bearings fault diagnosis, *Renew. Energy* 155 (2020) 1312–1327, <http://dx.doi.org/10.1016/j.renene.2020.04.041>.
- [53] X. Xu, M. Zhao, J. Lin, Y. Lei, Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis, *Measurement* 91 (2016) 385–397.
- [54] Y. Miao, J. Wang, B. Zhang, H. Li, Practical framework of Gini index in the application of machinery fault feature extraction, *Mech. Syst. Signal Process.* 165 (2022) 108333.