

Article

A Novel Blind Signal Detector Based on the Entropy of the Power Spectrum Subband Energy Ratio

Han Li ^{1,2} , Yanzhu Hu ¹ and Song Wang ^{1,*}

¹ School of Modern Post, Beijing University of Posts and Telecommunications, Beijing 100876, China; lihan_60@bupt.edu.cn (H.L.); bupt_automation_safety_yzhu@bupt.edu.cn (Y.H.)

² School of Intelligent Equipment, Shandong University of Science and Technology, Taian 271019, China

* Correspondence: wongsang@bupt.edu.cn; Tel.: +86-010-13366036076

Abstract: In this paper, we present a novel blind signal detector based on the entropy of the power spectrum subband energy ratio (PSER), the detection performance of which is significantly better than that of the classical energy detector. This detector is a full power spectrum detection method, and does not require the noise variance or prior information about the signal to be detected. According to the analysis of the statistical characteristics of the power spectrum subband energy ratio, this paper proposes concepts such as interval probability, interval entropy, sample entropy, joint interval entropy, PSER entropy, and sample entropy variance. Based on the multinomial distribution, in this paper the formulas for calculating the PSER entropy and the variance of sample entropy in the case of pure noise are derived. Based on the mixture multinomial distribution, the formulas for calculating the PSER entropy and the variance of sample entropy in the case of the signals mixed with noise are also derived. Under the constant false alarm strategy, the detector based on the entropy of the power spectrum subband energy ratio is derived. The experimental results for the primary signal detection are consistent with the theoretical calculation results, which proves that the detection method is correct.



Citation: Li, H.; Hu, Y.; Wang, S. A Novel Blind Signal Detector Based on the Entropy of the Power Spectrum Subband Energy Ratio. *Entropy* **2021**, *23*, 448. <https://doi.org/10.3390/e23040448>

Academic Editor: Ercan Kuruoglu

Received: 1 March 2021

Accepted: 8 April 2021

Published: 11 April 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: power spectrum subband energy ratio; sample entropy; interval entropy; PSER entropy; sample entropy variance; multinomial distribution; signal detection

1. Introduction

With the rapid development of wireless communication, spectrum sensing has been deeply studied and successfully applied in cognitive radio (CR). Common spectrum sensing methods [1,2] can be classified as matched filter detection [3], energy-based detection [4], and cyclostationary-based detection [5], among others. Matched-filtering is the optimum detection method when the transmitted signal is known [6]. Energy detection is considered to be the optimal method when there is no prior information of the transmitted signal. Cyclostationary detection is applicable to signals with cyclostationary features [2]. Energy detection can be classified into two categories: time domain energy detection [4] and frequency domain energy detection [7,8]. The power spectrum subband energy ratio (PSER) detector [9] is a local power spectrum energy detection technology. In order to enable PSER to be used in full spectrum detection, a new detector-based PSER is proposed in this paper, whose detection performance is higher than that of time domain energy detection.

In information theory, entropy is a measure of the uncertainty associated with a discrete random variable, and differential entropy is a measure of uncertainty of a continuous random. As the uncertainty of noise is higher than that of the signal, the entropy of the noise is higher than that of the signal, which is the basis of using entropy to detect signals. Information entropy has been successfully applied to signal detection [10–12]. The main entropy detection methods can be classified into two categories: the time domain and the frequency domain.

In the time domain, a signal with a low signal-noise ratio (SNR) is annihilated in the noise, and the estimated entropy is actually the entropy of the noise; therefore, the time domain entropy-based detector does not have an adequate detection performance. Nagaraj [12] introduced a time domain entropy-based matched-filtering for primary users (PU) detection and presented a likelihood ratio test for detecting a PU signal. Gu et al. [13] presented a new cross entropy-based spectrum sensing scheme that has two time-adjacent detected data sets of the PU.

In the frequency domain, the spectral bin amplitude of the signal is obviously higher than that of the noise. Therefore, the frequency domain entropy-based detector is widely applied in many fields. Zhang et al. [10,14] presented a spectrum sensing scheme based on the entropy of the spectrum magnitude, which has been used in many practical applications. For example, Jakub et al. [15] used this scheme to assist in blind signal detection, Zhao [16] improved the two-stage entropy detection method based on this scheme, and Waleed et al. [17] applied this scheme to maritime radar network. So [18] used the conditional entropy of the spectrum magnitude to detect unauthorized user signals in cognitive radio networks. Guillermo et al. [11] proposed an improved entropy estimation method based on Bartlett periodic spectrum. Ye et al. [19] proposed a method based on the exponential entropy. Zhu et al. [20] compared the performances of several entropy detectors in primary user detection. In these papers, the mean of the test statistics based on entropy could be calculated by differential entropy, however the variance of the test statistics based on entropy was not given in these papers, that is, its calculation formula was unknown. In detection theory, the variance of test statistics plays a very important role, therefore the detection based on entropy has a major drawback.

In the above references, none of the entropy-based detectors made use of the PSER. The PSER is a common metric used to represent the proportion of signal energy in a single spectral line. It has been extensively applied in the fields of machine design [21], earthquake modeling [22], remote communication [23,24], and geological engineering [25]. The white noise power spectrum bin conforms to the Gaussian distribution, and its PSER conforms to the Beta distribution [9]. Compared with the entropy detector based on spectral line amplitude, the PSER entropy detector has special properties.

This paper is organized as follows. In Section 2, the theoretical formulas of PSER entropy under pure noise and mixed signal case are deduced. The statistical characteristics of PSER entropy are summarized and the computational complexity of the main statistics are analyzed. Section 3 describes the derivation process of the PSER entropy detector under the constant false alarm strategy in detail. In Section 4, experiments are carried out to verify the accuracy of the PSER entropy detector, and the detection performance is compared with other detection methods. Section 5 provides additional details concerning the research process. The conclusions are drawn in Section 6.

2. PSER Entropy

PSER entropy takes the form of a classical Shannon entropy. Theoretically, the Shannon entropy can be approximated by the differential entropy; therefore, this method is adopted in the existing entropy-based signal detection methods [10]. However, this method has the disadvantage that the actual value is not the same as the theoretical value in some cases. Therefore, on the basis of analyzing the statistical characteristics of PSER entropy, this paper proposes a new method to calculate PSER entropy without using differential entropy. Firstly, the range of PSER [0,1] is divided into several equally spaced intervals. Then, under pure noise, the PSER entropy and its variance are derived using the multinomial distribution. Under signal mixed with noise, the PSER entropy and its variance are derived using the mixed multinomial distribution.

2.1. Probability Distribution for PSER

The signal mixed with additive Gaussian white noise (GWN) can be expressed as

$$s(n) = \begin{cases} z(n) & H_0 \\ x(n) + z(n) & H_1 \end{cases}, n = 0, 1, \dots, L-1 \quad (1)$$

where L is the number of sampling points; $s(n)$ is the signal to be detected; $x(n)$ is the signal; $z(n)$ is the GWN with a mean of zero and a variance of σ^2 ; H_0 represents the hypothesis corresponding to “no signal transmitted”; and H_1 corresponds to “signal transmitted”. The single-sided spectrum of $s(n)$ is

$$\vec{S}(k) = \sum_{n=0}^{L-1} s(n) e^{-j\frac{2\pi}{L}kn}, k = 0, 1, \dots, \frac{L}{2} - 1. \quad (2)$$

where j is the imaginary unit, and the arrow superscript denotes a complex function. The k th line in the power spectrum of $s(n)$ can be expressed as

$$P(k) = \frac{1}{L} |\vec{S}(k)|^2 = \frac{1}{L} \left((X_R(k) + Z_R(k))^2 + (X_I(k) + Z_I(k))^2 \right). \quad (3)$$

$X_R(k)$ and $X_I(k)$ represent the real and imaginary parts of the signal, respectively. $Z_R(k)$ and $Z_I(k)$ represent the real and imaginary parts of the noise, respectively.

The PSER $B_{d,N}(k)$ is defined as the ratio of the sum of the adjacent d bins from the k th bin in the power spectrum to the entire spectrum energy, i.e.,

$$B_{d,N}(k) = \frac{\sum_{l=k}^{k+d-1} P(l)}{\sum_{i=0}^{N-1} P(i)}, 1 \leq d < N-1, k = 0, 1, \dots, N-d, \quad (4)$$

where $N = L/2$, $\sum_{i=0}^{N-1} P(i)$ represents the total energy in the power spectrum and $\sum_{l=k}^{k+d-1} P(l)$ represents the total energy of the adjacent d bins. When there is noise in the power spectrum, it is clear that $B_{d,N}(k)$ is a random variable.

The probability distribution for $B_{d,N}(k)$ is described in detail in [9]. Under H_0 , $B_{d,N}(k)$ follows a beta distribution with parameters $d, N-d$. Under H_1 , $B_{d,N}(k)$ follows a doubly non-central beta distribution with parameters $d, N-d$ and non-centrality parameters λ'_k , and $\sum_{i=0}^{N-1} \lambda_i - \lambda'_k$, i.e., [9] (p. 7),

$$B_{d,N}(k) = \begin{cases} \beta(d, N-d) & H_0 \\ \beta_{d,N-d}(\lambda'_k, \sum_{i=0}^{N-1} \lambda_i - \lambda'_k) & H_1 \end{cases}.$$

where $\lambda_i = (X_R^2(k) + X_I^2(k)) / (N\sigma^2)$, i.e., the SNR of the k th spectral bin; $\lambda'_k = \sum_{l=k}^{k+d-1} \lambda_l$, the SNR of the d spectral lines starting from the k th spectral line; $\sum_{i=0}^{N-1} \lambda_i - \lambda'_k$ is the SNR of the spectral lines not contained in the selected subband. The probability density function (PDF) for $B_{d,N}(k)$ [9] (p. 7) is

$$f_{B_{d,N}(k)}(x) = \begin{cases} \frac{x^{d-1}(1-x)^{N-d-1}}{B(d, N-d)} & H_0 \\ e^{-(\delta_{k,1} + \delta_{k,2})} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\delta_{k,1}^j \delta_{k,2}^l (1-x)^{N+l-d-3} x^{j+d-1}}{j!l! B(j+d, N-d+l)} & H_1 \end{cases}, x \in [0, 1]. \quad (5)$$

where $\delta_{k,1} = \lambda'_k / 2$, $\delta_{k,2} = (\sum_{i=0}^{N-1} \lambda_i - \lambda'_k) / 2$. The cumulative distribution function (CDF) for $B_{d,N}(k)$ [9] (p. 7) is

$$F_{B_{d,N}(k)}(x) = \begin{cases} I_x(d, N-d) & H_0 \\ e^{-(\delta_{k,1} + \delta_{k,2})} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\delta_{k,1}^j \delta_{k,2}^l}{j!l!} I_x(j+d, N+l-d) & H_1 \end{cases}, x \in [0, 1]. \quad (6)$$

The subband used in this paper only contains one spectral bin, i.e., $d = 1$; therefore, $B_{1,N}(k)$ follows the distribution

$$B_{1,N}(k) = \begin{cases} \beta(1, N-1) & H_0 \\ \beta_{1,N-1}(\lambda_k, \sum_{i=0}^{N-1} \lambda_i - \lambda_k) & H_1 \end{cases} \quad (7)$$

The probability density function (PDF) for $B_{1,N}(k)$ is

$$f_{B_{1,N}(k)}(x) = \begin{cases} (N-1)(1-x)^{N-2} & H_0 \\ e^{-(\delta_{k,1}+\delta_{k,2})} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\delta_{k,1}^j \delta_{k,2}^l}{j!l!} \frac{(1-x)^{N+l-4} x^j}{B(j+1, N-1+l)} & H_1 \end{cases}, x \in [0, 1]. \quad (8)$$

where $\delta_{k,1} = \lambda_k/2$, $\delta_{k,2} = (\sum_{i=0}^{N-1} \lambda_i - \lambda_k)/2$.

The cumulative distribution function (CDF) for $B_{1,N}(k)$ is

$$F_{B_{1,N}(k)}(x) = \begin{cases} I_x(1, N-1) & H_0 \\ e^{-(\delta_{k,1}+\delta_{k,2})} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\delta_{k,1}^j \delta_{k,2}^l}{j!l!} I_x(j+1, N+l-1) & H_1 \end{cases}, x \in [0, 1]. \quad (9)$$

For the convenience of description, $B_{1,N}(k)$ is replaced by X_k .

2.2. Basic Definitions

X_k is in the range of $[0, 1]$. The range is divided into m intervals at $1/m$, i.e., $[0, 1/m)$, $[1/m, 2/m)$, $\dots$, $[(m-1)/m, 1]$. Then the i th interval is $[i/m, (i+1)/m)$, where $i = 0, 1, 2, \dots, m-1$. The probability that X_k falls into the i th interval [26] is

$$p_i = \int_{i/m}^{(i+1)/m} f_{X_k}(x) dx, \quad (10)$$

where $p_0 + p_1 + \dots + p_{m-1} = 1$. p_i is called the interval probability.

The PSER values of all spectral lines are regarded as a sequence $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})$. Let the number of times the data in $\mathbf{X}$ falls into the i th interval be t_i , which is a random variable, and $t_i = 0, 1, \dots, N$, $\sum_{i=0}^{m-1} t_i = N$. Let $\mathbf{t} = (t_0, t_1, \dots, t_{m-1})$. The frequency of the data in sequence $\mathbf{X}$ falling into the i th interval is denoted as T_i , i.e., $T_i = t_i/N$, and $\sum_{i=0}^{m-1} T_i = 1$.

The random variable Y_i is equal to $-T_i \log T_i$, which is called the sample entropy of PSER. The mean of the sample entropy of PSER in the i th interval is $H_i = E(Y_i)$, and H_i is the interval entropy. Notice that any two t_i are not independent of each other; therefore, any two Y_i are not independent of each other as well. Let $Z(\mathbf{t}; \mathbf{X}) = \sum_{i=0}^{m-1} Y_i$, i.e.,

$$Z(\mathbf{t}; \mathbf{X}) = - \sum_{i=0}^{m-1} \frac{t_i}{N} \log \left(\frac{t_i}{N} \right). \quad (11)$$

$Z(\mathbf{t}; \mathbf{X})$ is called sample entropy of PSER. $Z(\mathbf{t}; \mathbf{X})$ can be abbreviated as $Z(\mathbf{t})$ or Z .

By the definition of entropy, entropy is a mean, and it has no variance. However, the sample entropy is the entropy of a PSER sequence sample; therefore the sample entropy is a random variable, and it has a mean and a variance. The mean of the sample entropy is

$$H(m, N) = E(Z), \quad (12)$$

where m and N are the numbers of intervals and spectral bins, respectively. $H(m, N)$ can be called the total entropy or PSER entropy. The variance of the sample entropy is denoted as $Var(Z)$. In signal detection, $E(Z)$ and $Var(Z)$ are very important, and the calculation of $E(Z)$ and $Var(Z)$ is discussed in the following sections.

2.3. Calculating PSER Entropy Using the Differential Entropy

2.3.1. The Differential Entropy for X_k

Under H_0 , by the definition of differential entropy, the differential entropy for X_k is

$$\begin{aligned} h(B_{1,N}(k)) &= -\int_{-\infty}^{+\infty} p(x) \log p(x) dx \\ &= -\int_0^{+\infty} p(x) \log \left[(N-1)(1-x)^{N-2} \right] dx \\ &= -\int_0^{+\infty} p(x) [\log(N-1) + (N-2) \log(1-x)] dx \\ &= -\log(N-1) - (N-1)(N-2) \int_0^1 (1-x)^{N-2} \log(1-x) dx \\ &= \frac{N-2}{(N-1) \ln a} - \log(N-1), \end{aligned}$$

where a is the base of the logarithm. When $a = e$,

$$h(B_{1,N}(k)) = \frac{N-2}{N-1} - \ln(N-1). \quad (13)$$

2.3.2. PSER Entropy Calculated Using Differential Entropy

According to the calculation process of the spectrum magnitude entropy presented by Zhang in [10] and the equation given in [27] (p. 247), when $1/m \rightarrow 0$, the PSER entropy under H_0 is

$$\begin{aligned} H(m, N) &= -\sum_{i=0}^{m-1} p_i \log p_i = -\sum_{i=0}^{m-1} f_{X_k}(x) \frac{1}{m} \ln \left(f_{X_k}(x) \frac{1}{m} \right) \\ &\simeq h(B_{1,N}(k)) - \ln \left(\frac{1}{m} \right) = \frac{N-2}{N-1} - \ln \left(\frac{N-1}{m} \right). \end{aligned} \quad (14)$$

If $(N-1)/m < e$, then $H(m, N)$ is negative.

2.3.3. The Defect of the PSER Entropy Calculated Using Differential Entropy

PSER entropy is the mean of the sample entropy, and the sample entropy is nonnegative, therefore, the PSER entropy is nonnegative too. However, the PSER entropy calculated by Equation (14) is not always nonnegative. Especially when $(N-1)/m < e$, the PSER entropy is negative. The reason why Equation (14) is not always nonnegative is that the differential entropy is not always nonnegative.

The difference between the real PSER entropy calculated by simulation experiment and that calculated by differential entropy is shown in Figure 1. At least 10^4 Monte Carlo simulation experiments were carried out under $N = 256$ and different values of m .

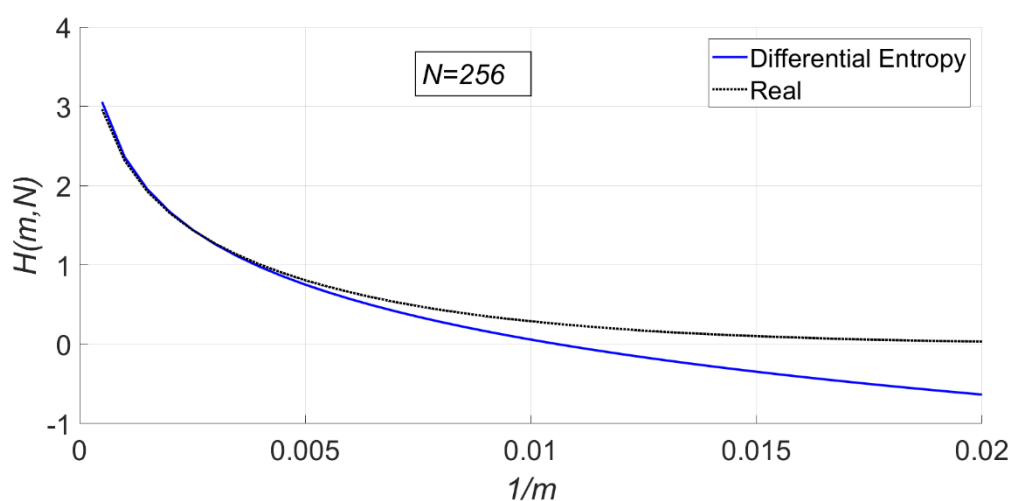


Figure 1. Comparison between the power spectrum subband energy ratio (PSER) entropy calculated using differential entropy and the real PSER entropy.

In Figure 1, the solid line is the PSER entropy calculated by the differential entropy, while the dotted line is the PSER entropy. When m is very large ($1/m < 0.005$), the results are close. However, when m is small ($1/m > 0.005$), the difference between the two methods is large, and even the total entropy calculated by the differential entropy is negative, which is inconsistent with the actual result. Therefore, in this paper a more reasonable method to calculate the PSER entropy is.

2.4. PSER Entropy under H_0

2.4.1. Definitions and Lemmas

Under H_0 , all X_k obey the same beta distribution. According to (10), the interval probability in i th interval is

$$p_i = \int_{i/m}^{(i+1)/m} (N-1)(1-t)^{N-2} dt = \left(1 - \frac{i}{m}\right)^{N-1} - \left(1 - \frac{i+1}{m}\right)^{N-1}. \quad (15)$$

p_i is essentially the area of the i th interval on the probability density map. Figure 2 shows p_0 and p_1 when m is 200, and N is 128.

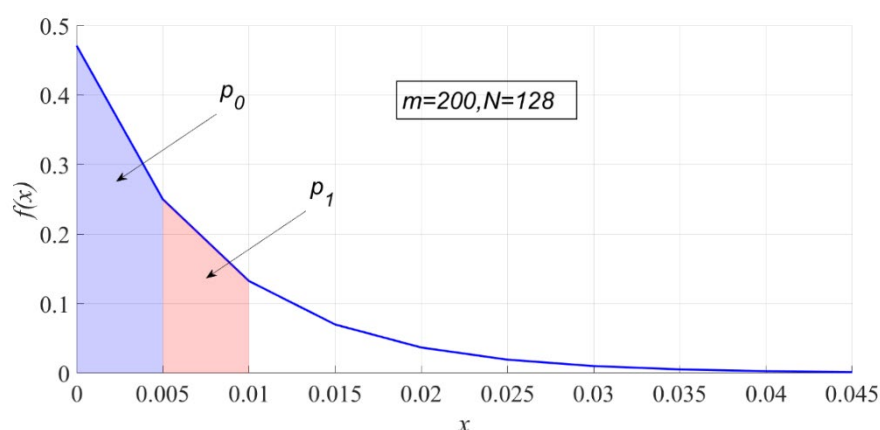


Figure 2. The interval probability under H_0 .

Let $T_{m,N} = \{\mathbf{t} = (t_0, t_1, \dots, t_{m-1}) : t_i \in \{0, 1, 2, \dots, N\}, t_0 + \dots + t_{m-1} = N\}$. It is a typical multinomial distribution problem that N data fall into m intervals. By the probability formula of multinomial distribution, the probability of $\mathbf{t} = (t_0, t_1, \dots, t_{m-1})$ is

$$\Pr(\mathbf{t}) = \Pr(t_0, t_1, \dots, t_{m-1}) = \binom{N}{t_0, t_1, \dots, t_{m-1}} p_0^{t_0} p_1^{t_1} \dots p_{m-1}^{t_{m-1}}, \quad (16)$$

where

$$\binom{N}{t_0, t_1, \dots, t_{m-1}} = \frac{N!}{t_0! t_1! \dots t_{m-1}!},$$

which is the multinomial coefficient. The following lemmas are used in the following analysis.

Lemma 1. If k is a non-negative integer, then [28] (p. 183)

$$\Pr(t_i = k, \mathbf{t} \in T_{m,N}) = \frac{N!}{k!(N-k)!} p_i^k (1-p_i)^{N-k}. \quad (17)$$

Lemma 2. If j is a non-negative integer, then

$$\sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) = \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}). \quad (18)$$

Proof. The left side of Equation (18) is $\sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) = 1$. From Lemma 1, the right side of Equation (18) is

$$\sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) = \sum_{j=0}^N \frac{N!}{j!(N-j)!} p_i^j (1-p_i)^{N-j} = 1.$$

□

Lemma 3. If k and l are a non-negative integer, and $k + l \leq N$, then [28] (p. 183)

$$\Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N}) = \frac{N!}{k!l!(N-k-l)!} p_i^k p_j^l (1-p_i-p_j)^{N-k-l}. \quad (19)$$

Lemma 4. If k and l are a non-negative integer, and $k + l \leq N$, then

$$\sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) = \sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N}). \quad (20)$$

Proof. From Lemma 3, the right side of Equation (20) is

$$\begin{aligned} & \sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N}) \\ &= \sum_{k=0}^N \sum_{l=0}^{N-k} \frac{N!}{k!l!(N-k-l)!} p_i^k p_j^l (1-p_i-p_j)^{N-k-l} \\ &= \sum_{k=0}^N \frac{N!}{k!(N-k)!} p_i^k (1-p_i)^{N-k} \sum_{l=0}^{N-k} \frac{(N-k)!}{l!(N-k-l)!} \left(\frac{p_j}{1-p_i}\right)^l \left(\frac{1-p_i-p_j}{1-p_i}\right)^{N-k-l} \\ &= \sum_{k=0}^N \frac{N!}{k!(N-k)!} p_i^k (1-p_i)^{N-k} = 1. \end{aligned}$$

□

2.4.2. Statistical Characteristics of T_i

The mean of T_i [28] (p. 183) is

$$E(T_i) = \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) \frac{j}{N} = \sum_{j=0}^N \frac{N!}{j!(N-j)!} p_i^j (1-p_i)^{N-j} \frac{j}{N} = p_i. \quad (21)$$

The mean-square value of T_i [28] (p. 183) is

$$E(T_i^2) = \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) \left(\frac{j}{N}\right)^2 = \frac{Np_i^2 + p_i - p_i^2}{N}. \quad (22)$$

The variance of T_i [28] (p. 183) is

$$\text{Var}(T_i) = E(T_i^2) - E^2(T_i) = \frac{p_i(1-p_i)}{N}. \quad (23)$$

2.4.3. Statistical Characteristics of Y_i

For the convenience of description, let

$$h_N(l) = -\frac{l}{N} \log\left(\frac{l}{N}\right), l = 0, 1, \dots, N.$$

The mean of the Y_i is the mean of all the entropy values of Y_i in the i th interval, that is,

$$\begin{aligned} H_i &= E(Y_i) = \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) h_N(j) \\ &= \sum_{j=0}^N \binom{N}{j} p_i^j (1 - p_i)^{N-j} h_N(j). \end{aligned} \quad (24)$$

H_i is the interval entropy. The following definition is not used to define the interval entropy in this paper:

$$H_i = E(Y_i) = - \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) \log(\Pr(t_i = j, \mathbf{t} \in T_{m,N})),$$

as this definition is the entropy of all probabilities of Y_i in the i th interval.

The mean-square value of Y_i is

$$\begin{aligned} E(Y_i^2) &= \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) h_N^2(j) \\ &= \sum_{j=0}^N \binom{N}{j} p_i^j (1 - p_i)^{N-j} h_N^2(j). \end{aligned} \quad (25)$$

The variance of Y_i is

$$\text{Var}(Y_i) = E(Y_i^2) - E^2(Y_i). \quad (26)$$

when $i \neq j$, the joint entropy of two interval is

$$\begin{aligned} H_{i,j} &= E(Y_i Y_j) = \sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) h_N(t_i) h_N(t_j) \\ &= \sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N}) h_N(k) h_N(l) \\ &= \sum_{k=0}^N \sum_{l=0}^{N-k} \frac{N!}{k!l!(N-k-l)!} p_i^k p_j^l (1 - p_i - p_j)^{N-k-l} h_N(k) h_N(l). \end{aligned} \quad (27)$$

when $i = j$, $H_{i,i} = E(Y_i^2)$, i.e., the mean-square value of Y_i .

2.4.4. Statistical Characteristics of $Z(\mathbf{t})$

The total entropy with m intervals and N spectral bins is

$$H(m, N) = E(Z) = \sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) Z. \quad (28)$$

Theorem 1. The PSER entropy is equal to the sum of all the interval entropy, i.e.,

$$H(m, N) = \sum_{i=0}^{m-1} H_i. \quad (29)$$

Proof. According to Lemma 2,

$$\begin{aligned}
 H(m, N) &= \sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) Z = \sum_{\mathbf{t} \in T_{m,N}} \left(\Pr(\mathbf{t}) \times \sum_{i=0}^{m-1} h_N(t_i) \right) \\
 &= \sum_{i=0}^{m-1} \sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N}) h_N(t_i) \\
 &= \sum_{i=0}^{m-1} \sum_{j=0}^N \frac{N!}{j!(N-j)!} p_i^j (1-p_i)^{N-j} h_N(t_i) \\
 &= \sum_{i=0}^{m-1} H_i.
 \end{aligned}$$

□

Theorem 2. The mean of the mean-square value of the sample entropy is equal to the sum of all the joint entropy of two intervals, i.e.,

$$E(Z^2) = \sum_{i=0}^{m-1} E(Y_i^2) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} H_{i,j}. \quad (30)$$

Proof. From Lemmas 2 and 4,

$$\begin{aligned}
 E(Z^2) &= \sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) Z^2 = \sum_{\mathbf{t} \in T_{m,N}} \left(\Pr(\mathbf{t}) \times \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} h_N(t_i) h_N(t_j) \right) \\
 &= \sum_{\mathbf{t} \in T_{m,N}} \left(\Pr(\mathbf{t}) \times \left(\sum_{i=0}^{m-1} h_N^2(t_i) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} h_N(t_i) h_N(t_j) \right) \right) \\
 &= \sum_{i=0}^{m-1} \sum_{\mathbf{t} \in T_{m,N}} (\Pr(\mathbf{t}) h_N^2(t_i)) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} \sum_{\mathbf{t} \in T_{m,N}} \Pr(\mathbf{t}) h_N(t_i) h_N(t_j) \\
 &= \sum_{i=0}^{m-1} E(Y_i^2) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} H_{i,j}.
 \end{aligned}$$

□

Theorem 3. The variance of the sample entropy is

$$\text{Var}(Z) = E(Z^2) - E^2(Z) = \sum_{i=0}^{m-1} E(Y_i^2) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} H_{i,j} - \left(\sum_{i=0}^{m-1} H_i \right)^2. \quad (31)$$

For the convenience of description, $H(m, N)$ and $\text{Var}(Z)$ under H_0 are replaced by μ_0 and σ_0^2 , respectively.

2.4.5. Computational Complexity

The calculation time of each statistic is mainly consumed in factorial calculation and the traversal of all cases.

Factorial calculation involves two cases: $\frac{N!}{j!(N-j)!}$ and $\frac{N!}{k!l!(N-k-l)!}$. They both have a time complexity of $O(N)$.

There are two methods to traverse all cases: the traversal of one selected interval and the traversal of two selected intervals. The corresponding expressions for these two

methods are $\sum_{j=0}^N \Pr(t_i = j, \mathbf{t} \in T_{m,N})$ and $\sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N})$, respectively.

The traversal of one selected interval requires that j take all the values from 0 to N . Considering the time spent computing the factorial, its time complexity is $O(N^2)$. Therefore, the time complexity of calculating the interval entropy is $O(N^2)$.

The traversal of two selected intervals requires firstly selection of k spectral bins from all N spectral bins, and then select l spectral bins from the remaining $N - k$ spectral bins. If k and l are fixed, then the time complexity is

$$O\left(N \frac{N!}{k!l!(N-k-l)!}\right).$$

$\sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t} \in T_{m,N})$ requires listing all the combinations of k and l , and its time complexity is

$$N \left(\binom{N}{0,0,N} + \binom{N}{0,1,N-1} + \cdots + \binom{N}{N,0,0} \right) = N3^N,$$

i.e., $O(N3^N)$. The time complexity of calculating the interval joint entropy $H_{i,j}$ is $O(N3^N)$.

Calculating the total entropy $H(m, N)$ requires computation of all m interval entropies, so its time complexity is $O(mN^2)$.

It takes the most time to calculate the variance of the sample entropy $Var(Z)$. As all $m(m-1)/2$ interval joint entropies have to be computed, the time complexity of calculating $Var(Z)$ is $O(Nm^23^N)$. In the following experiments, in order to ensure better detection performance, the values of m and N should not be too small, such as $m \geq 500$ and $N \geq 256$, and therefore the calculation time will be very long.

2.5. PSER Entropy under H_1

2.5.1. Definitions and Lemmas

Under H_1 , X_k obey $\beta_{1,N-1}(\lambda_k, \sum_{i=0}^{N-1} \lambda_i - \lambda_k)$, and different X_k have different non-centrality parameters, and therefore the calculation of total entropy and sample entropy variance under H_1 is much more complicated than that under H_0 . According to Equation (10), the interval probability in i th interval is

$$p_{k,i} = e^{-(\delta_{k,1} + \delta_{k,2})} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \frac{\delta_{k,1}^j \delta_{k,2}^l}{j!l!} \left(I_{(i+1)/m}(j+1, N+l-1) - I_{i/m}(j+1, N+l-1) \right). \quad (32)$$

where $\sum_{i=0}^{m-1} p_{k,i} = 1$. The subscript k stands for the label of X_k . Figure 3 shows p_0 and p_1 when $m = 200$, $N = 128$, $\delta_{k,1} = 1$, and $\delta_{k,2} = 2$.

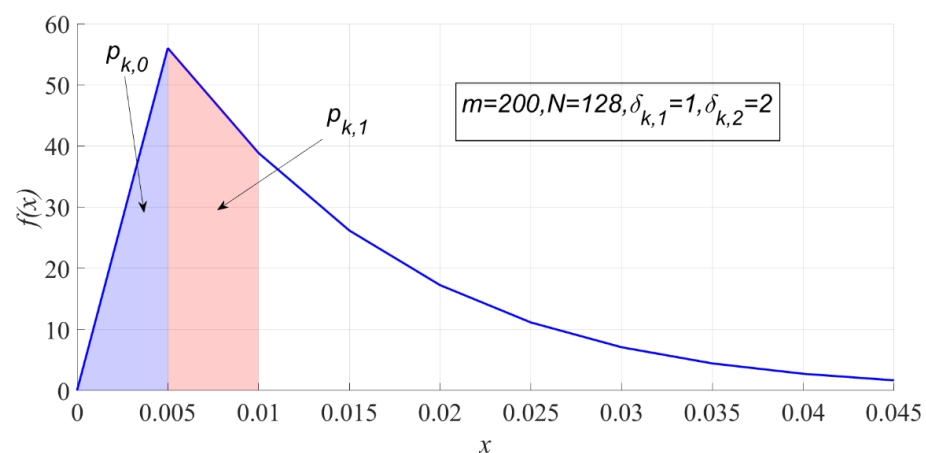


Figure 3. The interval probability of PSER under H_1 .

Under H_1 , different X_k obey different probability distributions, therefore this is a multinomial distribution problem under mixture distributions. Let

$$T'_{m,N} = \left\{ \mathbf{t}' = (t_{0,0}, \dots, t_{0,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,m-1}) : t_{k,i} \in \{0, 1, 2, \dots, N\}, t_i = \sum_{k=0}^{N-1} t_{k,i}, \sum_{i=0}^{m-1} t_i = N \right\},$$

where $t_{k,i}$ is the number of times of X_k falls into the i th interval. t_i represents the times of all X_k falls into the i th interval. In a sample, since X_k can only fall into one interval, $t_{k,i}$ only can be 0 or 1, and $\sum_{i=0}^{m-1} t_{k,i} = 1$.

By the probability formula of multinomial distribution, the probability when $\mathbf{t}' = (t_{0,0}, \dots, t_{0,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,m-1})$ is

$$\begin{aligned} \Pr(\mathbf{t}') &= \binom{N}{t_{0,0}, \dots, t_{0,m-1}, \dots, t_{k,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,m-1}} p_{0,0}^{t_{0,0}} \dots p_{0,m-1}^{t_{0,m-1}} \dots p_{k,0}^{t_{k,0}} \dots p_{k,m-1}^{t_{k,m-1}} \dots p_{N-1,0}^{t_{N-1,0}} \dots p_{N-1,m-1}^{t_{N-1,m-1}} \\ &= N! p_{0,0}^{t_{0,0}} \dots p_{0,m-1}^{t_{0,m-1}} \dots p_{k,0}^{t_{k,0}} \dots p_{k,m-1}^{t_{k,m-1}} \dots p_{N-1,0}^{t_{N-1,0}} \dots p_{N-1,m-1}^{t_{N-1,m-1}} = N! \prod_{k=0}^{N-1} \prod_{i=0}^{m-1} p_{k,i}^{t_{k,i}}. \end{aligned} \quad (33)$$

The following lemmas are used in the following analysis.

Lemma 5. If j is a non-negative integer, then

$$\begin{aligned} \Pr(t_i = \sum_{k=0}^{N-1} t_{k,i} = j, \mathbf{t}' \in T'_{m,N}) \\ = \frac{N!}{(N-j)!} \sum_{t_i=j} p_{0,i}^{t_{0,i}} \dots p_{k,i}^{t_{k,i}} \dots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \dots (1-p_{k,i})^{1-t_{k,i}} \dots (1-p_{N-1,i})^{1-t_{N-1,i}}. \end{aligned} \quad (34)$$

Proof.

$$\begin{aligned} \Pr(t_i = \sum_{k=0}^{N-1} t_{k,i} = j, \mathbf{t}' \in T'_{m,N}) &= \sum_{\mathbf{t}' \in T'_{m,N}, t_i = \sum_{k=0}^{N-1} t_{k,i}} \Pr(\mathbf{t}') \\ &= \sum_{\mathbf{t}' \in T'_{m,N}, t_i = j} \binom{N}{t_{0,0}, \dots, t_{0,m-1}, \dots, t_{k,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,m-1}} p_{0,0}^{t_{0,0}} \dots p_{0,m-1}^{t_{0,m-1}} \dots p_{k,0}^{t_{k,0}} \dots p_{k,m-1}^{t_{k,m-1}} \dots p_{N-1,0}^{t_{N-1,0}} \dots p_{N-1,m-1}^{t_{N-1,m-1}} \\ &= \sum_{\mathbf{t}' \in T'_{m,N}, t_i = j} \binom{N}{t_{0,i}, \dots, t_{k,i}, \dots, t_{N-1,i}, N-j} p_{0,i}^{t_{0,i}} \dots p_{k,i}^{t_{k,i}} \dots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \dots (1-p_{k,i})^{1-t_{k,i}} \dots (1-p_{N-1,i})^{1-t_{N-1,i}} \times \\ &\quad \binom{N-j}{t_{0,0}, \dots, t_{0,i-1}, t_{0,i+1}, \dots, t_{0,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,i-1}, t_{N-1,i+1}, \dots, t_{N-1,m-1}} \times \left(\frac{p_{0,0}}{1-p_{0,i}} \right)^{t_{0,0}} \dots \left(\frac{p_{0,i-1}}{1-p_{0,i}} \right)^{t_{0,i-1}} \\ &\quad \left(\frac{p_{0,i+1}}{1-p_{0,i}} \right)^{t_{0,i+1}} \dots \left(\frac{p_{0,m-1}}{1-p_{0,i}} \right)^{t_{0,m-1}} \dots \left(\frac{p_{N-1,0}}{1-p_{N-1,i}} \right)^{t_{N-1,0}} \dots \left(\frac{p_{N-1,i-1}}{1-p_{N-1,i}} \right)^{t_{N-1,i-1}} \left(\frac{p_{N-1,i+1}}{1-p_{N-1,i}} \right)^{t_{N-1,i+1}} \dots \left(\frac{p_{N-1,m-1}}{1-p_{N-1,i}} \right)^{t_{N-1,m-1}} \\ &= \frac{N!}{(N-j)!} \sum_{t_i=j} p_{0,i}^{t_{0,i}} \dots p_{k,i}^{t_{k,i}} \dots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \dots (1-p_{k,i})^{1-t_{k,i}} \dots (1-p_{N-1,i})^{1-t_{N-1,i}}. \end{aligned}$$

□

Lemma 6. If j is a non-negative integer, then

$$\sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') = \sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N}). \quad (35)$$

Proof. The left side of Equation (35) is $\sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') = 1$. From Lemma 5, the right side of Equation (35) is

$$\begin{aligned} &\sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N}) \\ &= \sum_{j=0}^N \frac{N!}{(N-j)!} \sum_{t_i=j} p_{0,i}^{t_{0,i}} \dots p_{k,i}^{t_{k,i}} \dots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \dots (1-p_{k,i})^{1-t_{k,i}} \dots (1-p_{N-1,i})^{1-t_{N-1,i}} = 1. \end{aligned}$$

Lemma 7. If g and l are a non-negative integer, and $g + l \leq N$, then

$$\Pr(t_i = g, t_j = l, t' \in T'_{m,N}) = \frac{N!}{(N-g-l)!} \sum_{t_i=g, t_j=l} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} p_{0,j}^{t_{0,j}} \cdots p_{k,j}^{t_{k,j}} \cdots p_{N-1,j}^{t_{N-1,j}} \times \\ (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}} (1-p_{0,j})^{1-t_{0,j}} \cdots (1-p_{k,j})^{1-t_{k,j}} \cdots (1-p_{N-1,j})^{1-t_{N-1,j}}. \quad (36)$$

Proof.

$$\Pr(t_i = g, t_j = l, t' \in T'_{m,N}) = \sum_{t' \in T'_{m,N}, t_i=g, t_j=l} \Pr(t') \\ = \sum_{t' \in T'_{m,N}, t_i=g, t_j=l} \binom{N}{t_{0,0}, \dots, t_{0,m-1}, \dots, t_{k,0}, \dots, t_{k,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,m-1}} p_{0,0}^{t_{0,0}} \cdots p_{0,m-1}^{t_{0,m-1}} \cdots p_{k,0}^{t_{k,0}} \cdots p_{k,m-1}^{t_{k,m-1}} \cdots p_{N-1,0}^{t_{N-1,0}} \cdots p_{N-1,m-1}^{t_{N-1,m-1}} \\ = \sum_{t' \in T'_{m,N}, t_i=g, t_j=l} \binom{N}{t_{0,i}, \dots, t_{k,i}, \dots, t_{N-1,i}, t_{0,j}, \dots, t_{k,j}, \dots, t_{N-1,j}, N-g-l} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} p_{0,j}^{t_{0,j}} \cdots p_{k,j}^{t_{k,j}} \cdots p_{N-1,j}^{t_{N-1,j}} \times \\ (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}} (1-p_{0,j})^{1-t_{0,j}} \cdots (1-p_{k,j})^{1-t_{k,j}} \cdots (1-p_{N-1,j})^{1-t_{N-1,j}} \\ \left(\binom{N-g-l}{t_{0,0}, \dots, t_{0,i-1}, t_{0,i+1}, \dots, t_{0,j-1}, t_{0,j+1}, \dots, t_{0,m-1}, \dots, t_{N-1,0}, \dots, t_{N-1,i-1}, t_{N-1,i+1}, \dots, t_{N-1,j-1}, t_{N-1,j+1}, \dots, t_{N-1,m-1}} \right) \times \left(\frac{p_{0,0}}{1-p_{0,i}} \right)^{t_{0,0}} \cdots \left(\frac{p_{0,i-1}}{1-p_{0,i}} \right)^{t_{0,i-1}} \\ \left(\frac{p_{0,i+1}}{1-p_{0,i}} \right)^{t_{0,i+1}} \cdots \left(\frac{p_{0,m-1}}{1-p_{0,i}} \right)^{t_{0,m-1}} \cdots \left(\frac{p_{N-1,0}}{1-p_{N-1,i}} \right)^{t_{N-1,0}} \cdots \left(\frac{p_{N-1,i-1}}{1-p_{N-1,i}} \right)^{t_{N-1,i-1}} \cdots \left(\frac{p_{N-1,i+1}}{1-p_{N-1,i}} \right)^{t_{N-1,i+1}} \cdots \left(\frac{p_{N-1,m-1}}{1-p_{N-1,i}} \right)^{t_{N-1,m-1}} \\ = \sum_{t_i=g, t_j=l} \binom{N}{t_{0,i}, \dots, t_{k,i}, \dots, t_{N-1,i}, t_{0,j}, \dots, t_{k,j}, \dots, t_{N-1,j}, N-g-l} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} p_{0,j}^{t_{0,j}} \cdots p_{k,j}^{t_{k,j}} \cdots p_{N-1,j}^{t_{N-1,j}} \times \\ (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}} (1-p_{0,j})^{1-t_{0,j}} \cdots (1-p_{k,j})^{1-t_{k,j}} \cdots (1-p_{N-1,j})^{1-t_{N-1,j}} \\ = \frac{N!}{(N-g-l)!} \sum_{t_i=g, t_j=l} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} p_{0,j}^{t_{0,j}} \cdots p_{k,j}^{t_{k,j}} \cdots p_{N-1,j}^{t_{N-1,j}} \times \\ (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}} (1-p_{0,j})^{1-t_{0,j}} \cdots (1-p_{k,j})^{1-t_{k,j}} \cdots (1-p_{N-1,j})^{1-t_{N-1,j}}.$$

□

2.5.2. Statistical Characteristics of T_i

The mean of T_i is

$$E(T_i) = \sum_{j=0}^N \Pr(t_i = j, t' \in T'_{m,N}) \frac{j}{N} \\ = \sum_{j=0}^N \sum_{t_i=j} \binom{N}{t_{0,i}, \dots, t_{k,i}, \dots, t_{N-1,i}, N-j} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}} \frac{j}{N} \\ = \sum_{j=0}^N \frac{j}{N} \frac{N!}{(N-j)!} \sum_{t_i=j} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}}. \quad (37)$$

The mean-square value of T_i is

$$E(T_i^2) = \sum_{j=0}^N \Pr(t_i = j, t' \in T'_{m,N}) \left(\frac{j}{N} \right)^2 \\ = \sum_{j=0}^N \left(\frac{j}{N} \right)^2 \frac{N!}{(N-j)!} \sum_{t_i=j} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}}. \quad (38)$$

The variance of T_i is

$$\text{Var}(T_i) = E(T_i^2) - E^2(T_i).$$

2.5.3. Statistical Characteristics of Y_i

The mean of the Y_i is

$$H_i = E(Y_i) = \sum_{j=0}^N \Pr(t_i = j, t' \in T'_{m,N}) h_N(j) \\ = \sum_{j=0}^N \frac{j}{N} \frac{N!}{(N-j)!} h_N(j) \sum_{t_i=j} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}}. \quad (39)$$

The mean-square value of Y_i is

$$\begin{aligned} E(Y_i^2) &= \sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N}) h_N^2(j) \\ &= \sum_{j=0}^N \frac{j}{N} \frac{N!}{(N-j)!} h_N^2(j) \sum_{t_i=j} p_{0,i}^{t_{0,i}} \cdots p_{k,i}^{t_{k,i}} \cdots p_{N-1,i}^{t_{N-1,i}} (1-p_{0,i})^{1-t_{0,i}} \cdots (1-p_{k,i})^{1-t_{k,i}} \cdots (1-p_{N-1,i})^{1-t_{N-1,i}}. \end{aligned} \quad (40)$$

The variance of Y_i is

$$\text{Var}(Y_i) = E(Y_i^2) - E^2(Y_i).$$

when $i \neq j$, the joint entropy of two interval is

$$\begin{aligned} H_{i,j} &= E(Y_i Y_j) = \sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') h_N(t_i) h_N(t_j) \\ &= \sum_{g=0}^N \sum_{l=0}^{N-g} \Pr(t_i = g, t_j = l, \mathbf{t}' \in T'_{m,N}) h_N(g) h_N(l) \end{aligned} \quad (41)$$

when $i = j$, $H_{i,i} = E(Y_i^2)$, i.e., the mean-square value of Y_i .

2.5.4. Statistical Characteristics of $Z(\mathbf{t}')$

Under H_1 , the PSER entropy is

$$H(m, N) = \sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') Z(\mathbf{t}').$$

Theorem 4. Under H_1 , the PSER entropy is equal to the sum of all the interval entropy, i.e.,

$$H(m, N) = \sum_{j=0}^{m-1} H_j. \quad (42)$$

Proof.

$$\begin{aligned} \sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') Z &= \sum_{\mathbf{t}' \in T'_{m,N}} \left(\Pr(\mathbf{t}') \times \sum_{i=0}^{m-1} h_N(t_i) \right) \\ &= \sum_{i=0}^{m-1} \sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') h_N(t_i) = \sum_{i=0}^{m-1} \sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N}) h_N(t_i) \\ &= \sum_{i=0}^{m-1} H_i. \end{aligned}$$

□

The mean of the mean-square value of $Z(\mathbf{t}')$ is

$$\begin{aligned} E(Z^2) &= \sum_{\mathbf{t}' \in T'_{m,N}} \Pr(\mathbf{t}') Z^2 = \sum_{\mathbf{t}' \in T'_{m,N}} \left(\Pr(\mathbf{t}') \times \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} h_N(t_i) h_N(t_j) \right) \\ &= \sum_{i=0}^{m-1} E(Y_i^2) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} H_{i,j}. \end{aligned} \quad (43)$$

The variance of $Z(\mathbf{t}')$ is

$$\text{Var}(Z) = E(Z^2) - E^2(Z) = \sum_{i=0}^{m-1} E(Y_i^2) + 2 \sum_{i=0}^{m-2} \sum_{j=i+1}^{m-1} H_{i,j} - \left(\sum_{i=0}^{m-1} H_i \right)^2. \quad (44)$$

For the convenience of description, $H(m, N)$ and $\text{Var}(Z)$ are denoted as μ_1 and σ_1^2 under H_1 .

2.5.5. Computational Complexity

Under H_1 , many statistics take a much longer time than under H_0 . The calculation time is mainly consumed in two aspects: calculation of $p_{k,i}$, and the factorial calculation and the traversal of all cases.

1. Calculation of $p_{k,i}$

As seen from Equation (32), $p_{k,i}$ is expressed by infinite double series under H_1 , and its value could only be obtained by numerical calculation. Since the number of calculation terms is set to be large, it will take a significant amount of calculation time.

2. Factorial calculation

Similar to the analysis under H_0 , the time complexity of the factorial calculation is $O(N)$.

3. The traversal of all cases

There are two methods to traverse all cases: the traversal of one selected interval and the traversal of two selected intervals. The corresponding expressions for these two methods are $\sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N})$ and $\sum_{g=0}^N \sum_{l=0}^{N-g} \Pr(t_i = g, t_j = l, \mathbf{t}' \in T'_{m,N})$, respectively.

Calculating $\Pr(t_i = j, \mathbf{t}' \in T'_{m,N})$ is a process of choosing j of N lines, and its time complexity is $O(NC_N^j)$. Similar to the analysis under H_0 , the computational complexity of $\sum_{j=0}^N \Pr(t_i = j, \mathbf{t}' \in T'_{m,N})$ is $O(N2^N)$. Therefore, the time complexity of calculating the interval entropy is $O(N2^N)$.

The computational complexity of $\sum_{k=0}^N \sum_{l=0}^{N-k} \Pr(t_i = k, t_j = l, \mathbf{t}' \in T_{m,N})$ is $O(N^3 3^N)$. Therefore the time complexity of calculating the interval joint entropy $H_{i,j}$ is $O(N^3 3^N)$. The time complexity of calculating the PSER entropy is $O(mN^3 3^N)$. The time complexity of calculating the variance of the sample entropy $\text{Var}(Z)$ is $O(m^2 N^3 3^N)$.

3. Signal Detector Based on the PSER Entropy

In this section, a signal detection method based on PSER entropy is deduced under the constant false alarm (CFAR) strategy according to the PSER entropy and sample entropy variance derived in Section 2. This method is also called full power spectrum subband energy ratio entropy detector (FPSSED), because it detects on the full power spectrum.

3.1. Principle

Signal detection based on PSER entropy takes the sample entropy as a detection statistic to judge whether there is a signal in the whole spectrum. The sample entropy is

$$Z = \sum_{i=0}^{m-1} Y_i$$

The PSER entropy of GWN is obviously different from that of the mixed signal. In general, the PSER entropy of the mixed signal will be less than that of GWN, but sometimes it will also be greater than that of GWN. This can be seen in Figure 4.

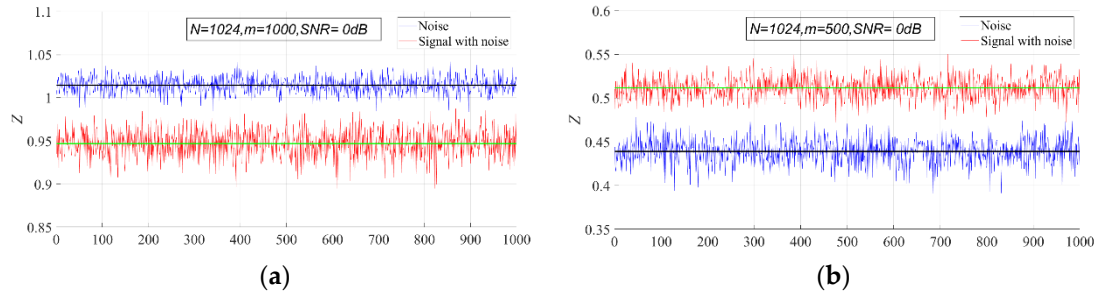


Figure 4. Comparison of the PSER entropy between GWN and the mixed signal ($N = 1024$, signal–noise ratio (SNR) = 0 dB). (a) $m = 1000$; (b) $m = 500$. The blue broken line is the sample entropy of GWN, and the red broken line is the sample entropy of the Ricker signal. The black line is the PSER entropy of GWN, and the blue line is the PSER entropy of the Ricker signal.

In Figure 4a, the PSER entropy of GWN is higher than that of the noisy Ricker signal. However, the PSER entropy of GWN is lower than that of the noisy Ricker signal in Figure 4b. Therefore, when setting the detection threshold of the PSER entropy detector, the relationship between the PSER entropy of the signal and that of noise should be considered.

3.1.1. The PSER Entropy of a Signal Less Than That of GWN

When the PSER entropy of signal is less than that of GWN, let the threshold be η_1 which tests the decision statistic. If the test statistic is less than η_1 , the signal is deemed to be present, and it is absent otherwise, i.e.,

$$\begin{cases} Z > \eta_1 & H_0 \\ Z < \eta_1 & H_1 \end{cases} \quad (45)$$

The distribution of Z is regarded as Gaussian in this paper, so

$$Z \sim \begin{cases} \mathcal{N}(\mu_0, \sigma_0^2), & H_0 \\ \mathcal{N}(\mu_1, \sigma_1^2), & H_1 \end{cases} \quad (46)$$

Under the CFAR strategy, the false alarm probability P_f can be expressed as follows:

$$P_f = \Pr(Z < \eta_1 | H_0) = 1 - Q\left(\frac{\eta_1 - \mu_0}{\sigma_0}\right). \quad (47)$$

η_1 can be derived from Equation (47)

$$\eta_1 = Q^{-1}(1 - P_f)\sigma_0 + \mu_0. \quad (48)$$

The detection probability P_d can be expressed as follows:

$$P_d = \Pr(Z < \eta_1 | H_1) = 1 - Q\left(\frac{\eta_1 - \mu_1}{\sigma_1}\right). \quad (49)$$

Substituting Equation (48) into Equation (49), P_d can then be evaluated as follows:

$$P_d = 1 - Q\left(\frac{Q^{-1}(1 - P_f)\sigma_0 + \mu_0 - \mu_1}{\sigma_1}\right). \quad (50)$$

3.1.2. The PSER Entropy of a Signal Larger Than That of GWN

When the PSER entropy of a signal is larger than that of GWN, let the threshold be η_2 . If the test statistic is larger than η_2 , the signal is deemed to be present, and it is absent otherwise, i.e.,

$$\begin{cases} Z < \eta_2 & H_0 \\ Z > \eta_2 & H_1 \end{cases} . \quad (51)$$

The false alarm probability P_f is

$$P_f = \Pr(Z > \eta_2 | H_0) = Q\left(\frac{\eta_2 - \mu_0}{\sigma_0}\right), \quad (52)$$

and

$$\eta_2 = Q^{-1}(P_f)\sigma_0 + \mu_0. \quad (53)$$

The detection probability P_d is

$$P_d = \Pr(Z > \eta_2 | H_1) = Q\left(\frac{\eta_2 - \mu_1}{\sigma_1}\right). \quad (54)$$

Substituting Equation (53) into Equation (54), P_d can then be evaluated as follows:

$$P_d = Q\left(\frac{Q^{-1}(P_f)\sigma_0 + \mu_0 - \mu_1}{\sigma_1}\right). \quad (55)$$

3.2. Other Detection Methods

In the following experiment, the PSER entropy detector compare with the commonly used full spectrum energy detection (FSED) [29] and matched-filtering detector (MFD) methods, under the same condition. In this section, we introduce these two detectors.

3.2.1. Full Spectrum Energy Detection

The performance of FSED is exactly the same as that of classical energy detection (ED). The total spectral energy is measured by the sum of all spectral lines in the power spectrum, that is,

$$T_{FD} = \sum_{k=0}^{N-1} P(k). \quad (56)$$

Let γ be the SNR, that is, $\gamma = \frac{1}{N^2\sigma^2} \sum_{k=0}^{N-1} (X_R^2(k) + X_I^2(k))$. When the detection length N is large enough, T_{FD} obeys a Gaussian distribution:

$$T_{FD} \sim \begin{cases} \mathcal{N}(N\sigma^2, N\sigma^4) & H_0 \\ \mathcal{N}((1+\gamma)N\sigma^2, (1+2\gamma)N\sigma^4) & H_1 \end{cases} . \quad (57)$$

Let the threshold be η_{FD} . The false alarm probability and detection probability can be expressed as follows:

$$P_f = \Pr(T_{FD} \geq \eta_{FD}) = Q\left(\frac{\eta_{FD} - N\sigma^2}{\sqrt{N}\sigma^2}\right), \quad (58)$$

$$P_d = \Pr(T_{FD} > \eta_{FD} | H_1) = Q\left(\frac{Q^{-1}(P_f) - \gamma\sqrt{N}}{\sqrt{1+2\gamma}}\right). \quad (59)$$

where $\eta_{FD} = (Q^{-1}(P_f)\sqrt{N} + N)\sigma^2$.

3.2.2. Matched Filter Detection

The main advantage of matched filtering is the short time to achieve a certain false alarm probability or detection probability. Hence, it requires perfect knowledge of the signal. In the time domain, the detection statistic of matched filtering is

$$T_{MFD} = \sum_{n=0}^{L-1} s(n)x(n), \quad (60)$$

where $s(n)$ is the transmitted signal, and $x(n)$ is signal to be detector. Let $E = \sum_{n=0}^{L-1} x^2(n)$, i.e., all energy of the signal, and η_{MFD} is the threshold. The false alarm probability and detection probability can be expressed as follows:

$$P_f = \Pr(T_{MFD} \geq \eta_{MFD} | H_0) = Q\left(\frac{\eta_{MFD}}{\sigma\sqrt{E}}\right), \quad (61)$$

$$P_d = \Pr(T_{MFD} \geq \eta_{MFD} | H_1) = Q\left(\frac{\eta_{MFD} - E}{\sigma\sqrt{E}}\right). \quad (62)$$

and $\eta_{MFD} = Q^{-1}(P_f)\sigma\sqrt{E}$.

4. Experiments

For this section, we verified and compared the detection performances of the FPSED, FSED, and MTD discussed in Section 3 through Monte Carlo simulations. The primary signal was a binary phase shift keying(BPSK) modulated signal with symbol ratio 1kbit/s, carrier frequency 1000 Hz, and sampling frequency 10^5 Hz.

We performed all Monte Carlo simulations for at least 10^4 independent trials. We set P_f to 0.05. We used mean-square error (MSE) to measure the deviation between the theoretical values and actual statistical results. All the programs in the experiment were run in MATLAB set up on a laptop with a Core i5 CPU and 16GB RAM.

Since the PSER entropy μ_1 , and the sample entropy variance σ_0^2 and σ_1^2 cannot be calculated, a large number of simulation data were generated in the experiment to obtain $\hat{\mu}_1$, $\hat{\sigma}_0^2$, and $\hat{\sigma}_1^2$, which replace μ_1 , σ_0^2 , and σ_1^2 , respectively.

4.1. Experiments under H_0

This section verifies whether the calculation results of each statistic is correct by comparing the statistical results with the theoretical calculation result. The effects of noise intensity, the number of spectral lines and the number of intervals on interval probability, interval entropy, PSER entropy, and the variance of sample entropy are analyzed.

4.1.1. Influence of Noise

According to the probability density function of PSER, PSER has nothing to do with noise intensity. Therefore, noise intensity has no effect on each statistic under H_0 . In the following experiment, the variance of noise has 10 values ranging from 0.1 to 1 at 0.1 intervals, $N = 512$, and $m = 500$.

In Figures 5 and 6, the theoretical and actual values of interval probability and interval entropy under different noise intensities are compared. The results show that the theoretical values are in good agreement with the statistical values, and the noise intensity has no effect on interval probability. In the first few intervals ($i < 7$), the interval probabilities are large, so the interval entropies are also large, contributing more to the total entropy.

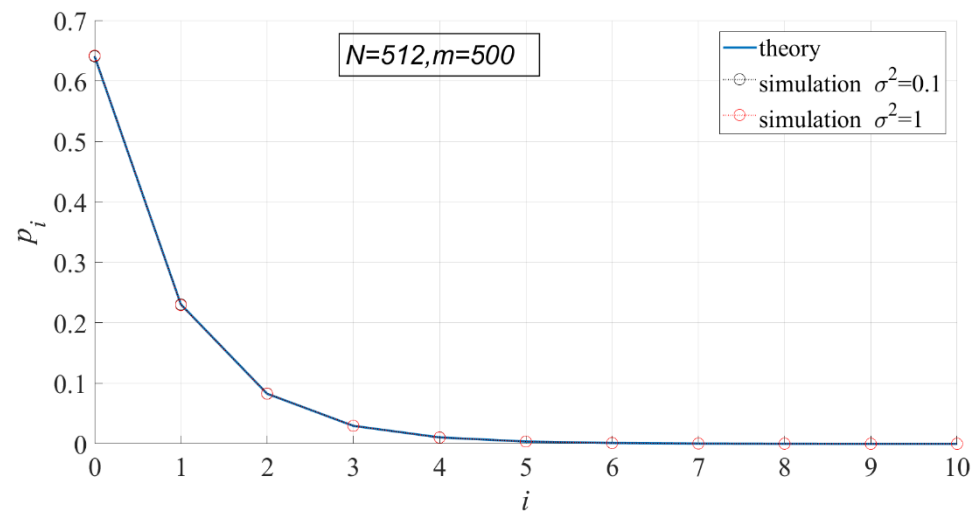


Figure 5. Comparison of the theoretical interval probability and the actual interval probability.

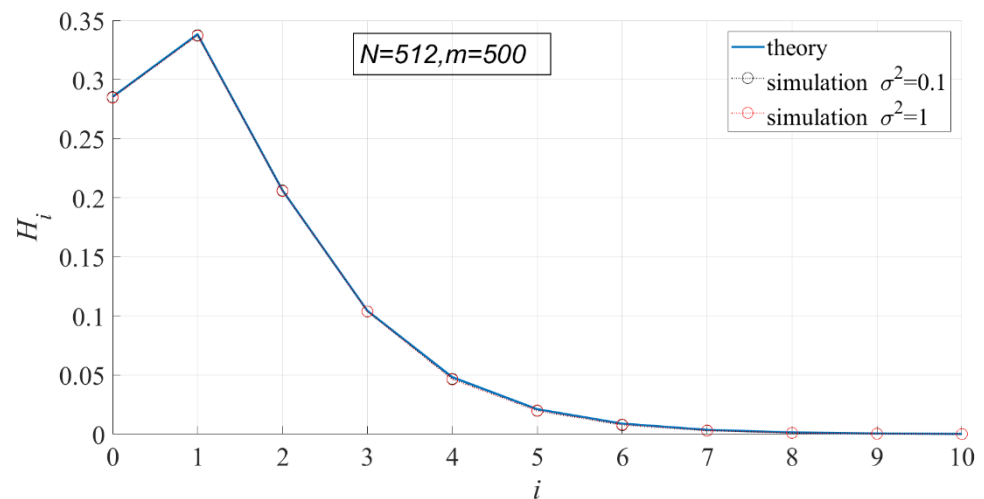


Figure 6. Comparison of the theoretical interval entropy and the actual interval entropy.

In Figure 7, the theoretical values of PSER entropies are compared with the actual values. It can be seen that the theoretical values are basically consistent with the actual values, but the actual values are slightly smaller than the theoretical values. The actual values do not change with noise intensity, indicating that noise intensity has no effect on PSER entropy.

Since the theoretical variance of sample entropy cannot be calculated, only the variation of the actual variance of sample entropy with noise intensity is shown in Figure 8. It can be seen that the variance of sample entropy does not change with the noise intensity.

The above experiments show that the actual statistical results are consistent with the theoretical calculation results, indicating that the calculation methods of interval probability, interval entropy, PSER entropy, and sample entropy variance determined in this paper are correct.

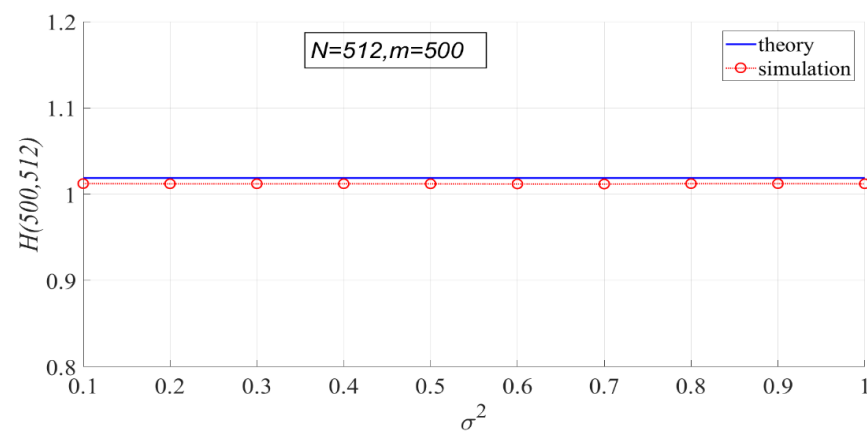


Figure 7. Comparison of theoretical PSER entropy and actual PSER entropy.

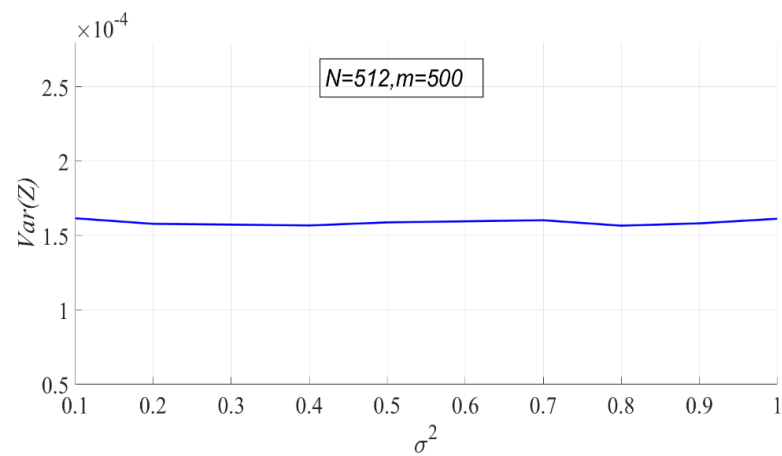


Figure 8. The actual variance of sample entropy.

4.1.2. Influence of N

In Figure 9, the effect of N on the interval probability is shown. When m is fixed, the larger N is, the larger p_0 is, and the interval probabilities in other intervals will be smaller. The reason for that is that the larger N is, the smaller the energy ratio of each line to the entire power spectrum.

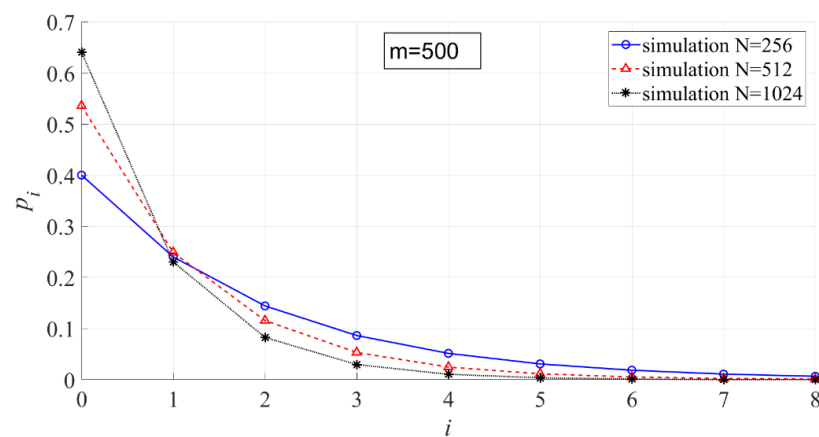


Figure 9. The effect of N on the interval probability ($m = 500$, $N = 256, 512, 1024$).

In Figure 10, the effect of N on the interval entropy is shown. The larger N is, the smaller H_i .

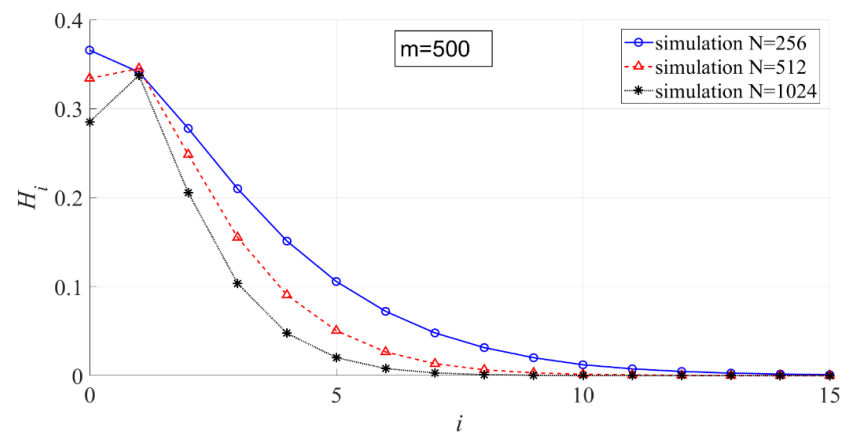


Figure 10. The effect of N on the interval entropy ($m = 500$, $N = 256, 512, 1024$).

In Figure 11, the effect of N on the interval entropy is shown. When m is fixed, the larger N is, the smaller $H(m, N)$ becomes. When N is the same, the larger m is, the larger the PSER entropy is.

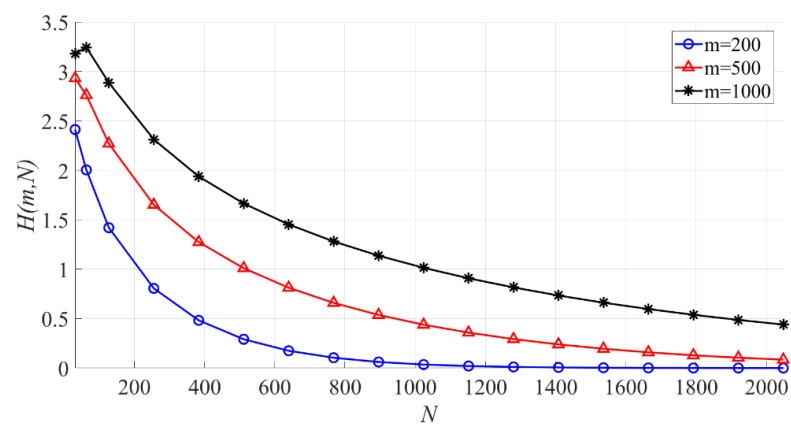


Figure 11. The effect of N on the PSER entropy ($m = 200, 500, 1000$).

In Figure 12, the effect of N on the variance of the sample entropy is shown. When m is fixed, the larger N is, the smaller $Var(Z)$ becomes.

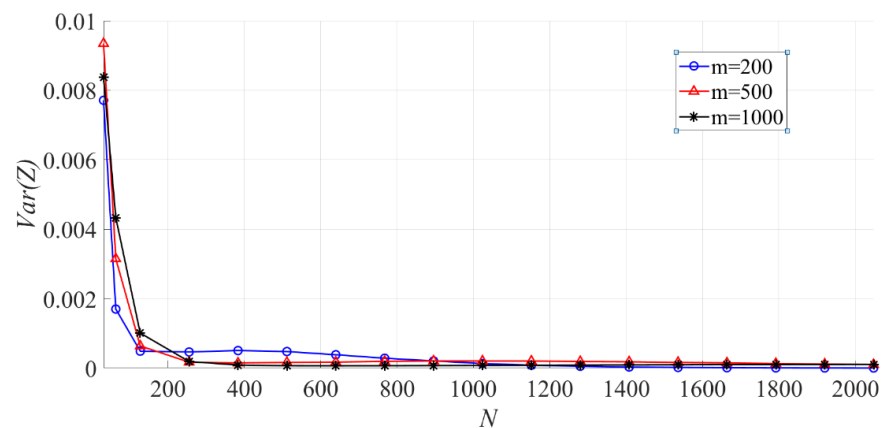


Figure 12. The effect of N on the variance of the sample entropy ($m = 200, 500, 1000$).

4.1.3. Influence of m

In Figure 13, the effect of m on the PSER entropy is shown. When N is fixed, the larger m is, the larger $H(m, N)$ becomes.

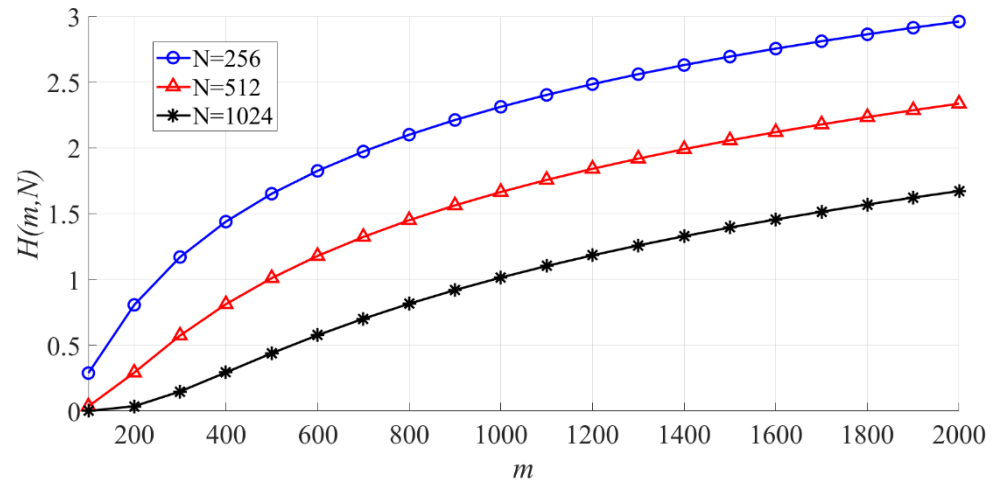


Figure 13. The effect of m on the PSER entropy ($N = 256, 512, 1024$).

In Figure 14, the effect of N on the variance of the sample entropy is shown. When N is fixed, the larger m is, the variance of sample entropy increases first, then decreases, and then increases slowly.

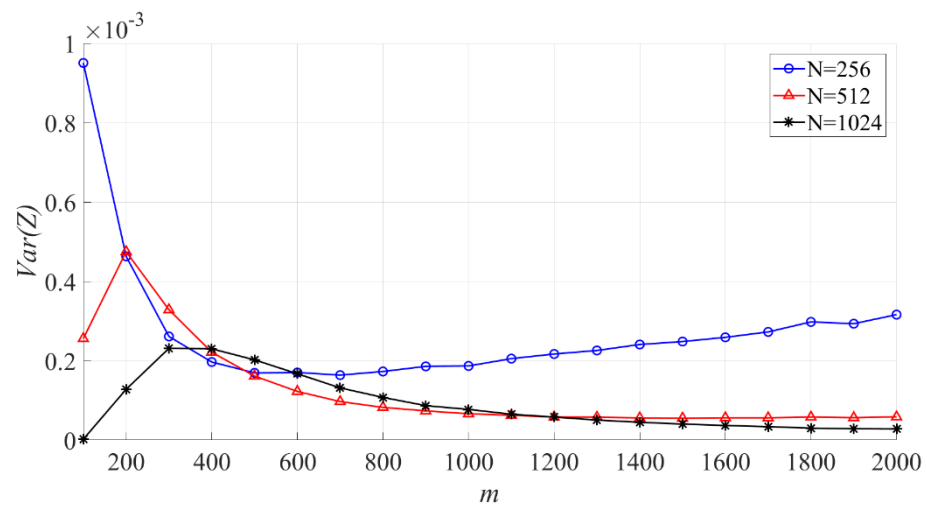


Figure 14. The effect of m on the variance of the sample entropy ($N = 256, 512, 1024$).

4.1.4. The Parameters for the Next Experiment

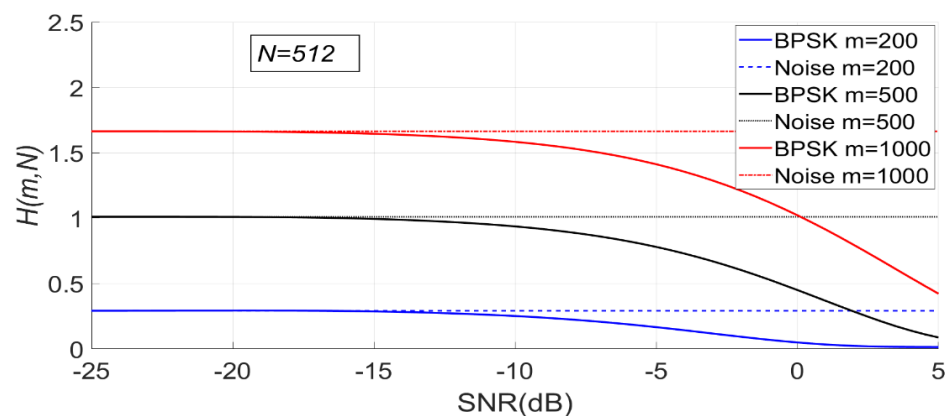
After theoretical calculations and experimental statistical analysis, the PSER entropy and sample entropy variance used in the following experiments were obtained, as shown in Table 1.

Table 1. The parameters under H_0 .

N \ m	200	500	1000
256	$\mu_0 = 0.809521,$ $\hat{\sigma}_0^2 = 0.000453$	$\mu_0 = 1.652905,$ $\hat{\sigma}_0^2 = 0.000169$	$\mu_0 = 2.313999,$ $\hat{\sigma}_0^2 = 0.000192$
512	$\mu_0 = 0.291461,$ $\hat{\sigma}_0^2 = 0.000466$	$\mu_0 = 1.011005,$ $\hat{\sigma}_0^2 = 0.000159$	$\mu_0 = 1.665232,$ $\hat{\sigma}_0^2 = 6.54 \times 10^{-5}$
1024	$\mu_0 = 0.035906,$ $\hat{\sigma}_0^2 = 0.000130$	$\mu_0 = 0.439132,$ $\hat{\sigma}_0^2 = 0.000202$	$\mu_0 = 1.014594,$ $\hat{\sigma}_0^2 = 7.7 \times 10^{-5}$

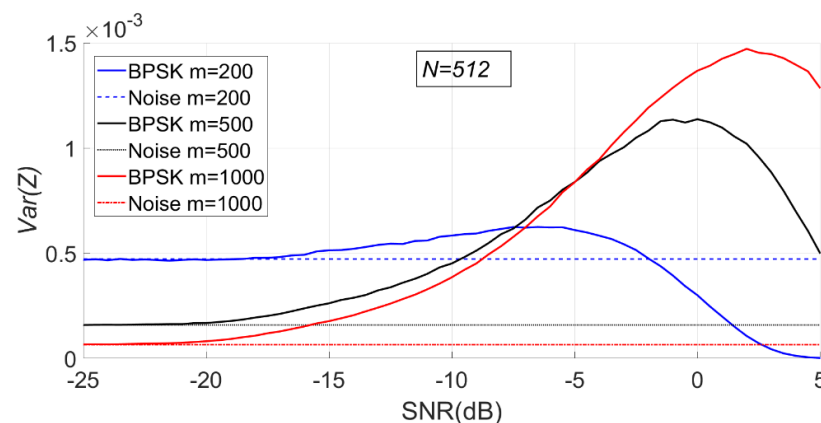
4.2. Experiments under H_1

When N is fixed, the change of PSER entropy of the BPSK signal with noise under H_1 is shown in Figure 15. It can be seen that the PSER entropy of BPSK signal decreases gradually with the increase of SNR. When the SNR is less than -15 dB, the PSER entropy of noise and that of the BPSK signal are almost the same; therefore, it is impossible to distinguish between noise and the BPSK signal.

**Figure 15.** The change of PSER entropy of the binary phase shift keying (BPSK) signal with noise when N is fixed ($m = 200, 500, 1000, N = 512$).

Since the PSER entropy of the BPSK signal is always less than that of noise, the threshold η_1 should be used in BPSK signal detection.

As can be seen from Figure 16, with the increase of SNR, the sample entropy variance of the BPSK signal first increases and then gradually decreases.

**Figure 16.** The change of sample entropy variance of BPSK signal with noise when N is fixed ($m = 200, 500, 1000, N = 512$).

When m is fixed, the change of PSER entropy and sample entropy variance of BPSK signal with noise under H_1 is shown in Figures 17 and 18. It can be seen that the PSER entropy of BPSK signal decreases gradually with the increase of SNR.

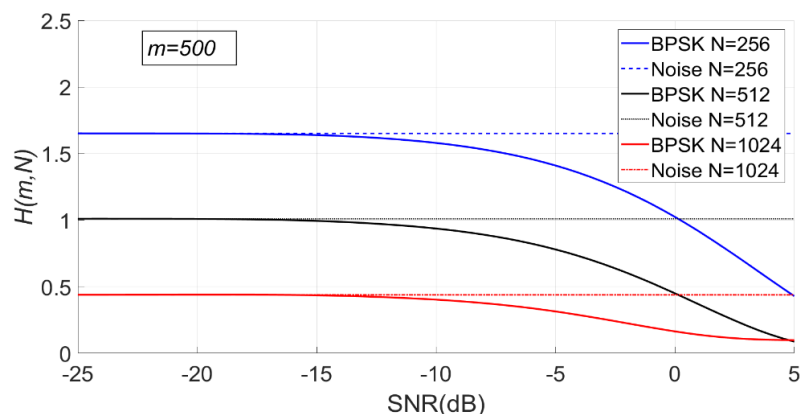


Figure 17. The change of PSER entropy of the BPSK signal with noise when m is fixed ($m = 500$, $N = 256, 512, 1024$).

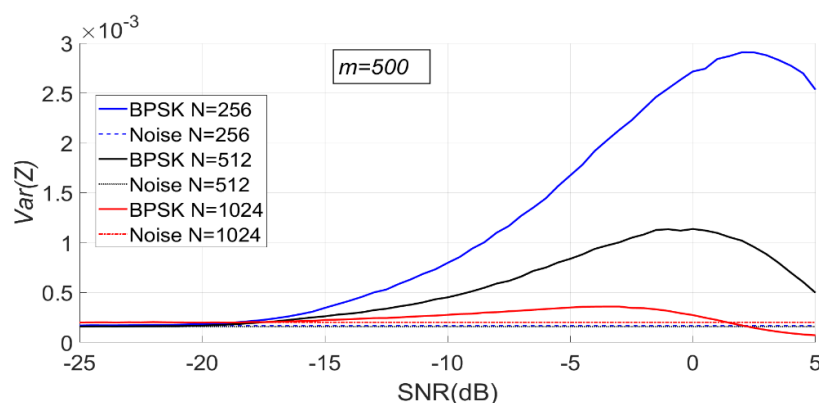


Figure 18. The change of sample entropy variance of the BPSK signal with noise when m is fixed ($m = 500$, $N = 256, 512, 1024$).

As can be seen from Figure 18, with the increase of SNR, the sample entropy variance of the BPSK signal first increases and then gradually decreases.

4.3. Comparison of Detection Performance

When N is 512, m is 200, 500 and 1000, respectively. The detection results of the BPSK signal are shown in Figures 19–21.

In Figure 19, the actual false alarm probabilities fluctuate slightly and do not change with the SNR, which is consistent with the characteristics of constant false alarm.

It can be seen from Figures 20 and 21 that the detection probability of the PSER entropy detector is obviously better than that of the FSED method when m is 1000. However, when m is 200, the detection performance is lower than that of FSED. There is no doubt that the detection performance of matched filtering is the best.

The MSEs of the actual and theoretical probabilities of these experiments are given in Table 2.

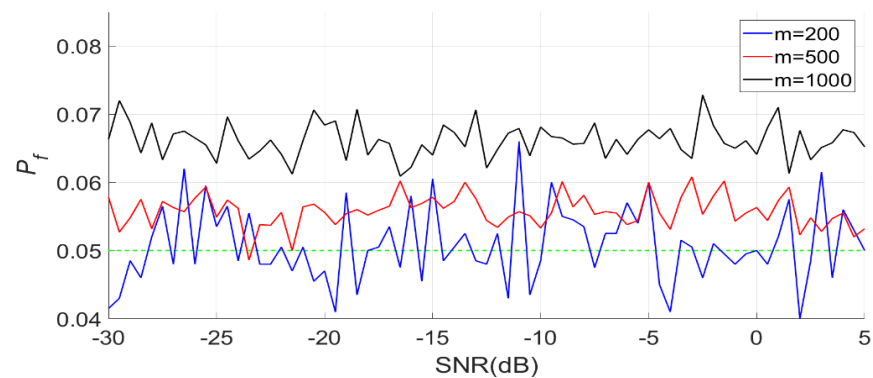


Figure 19. Actual false alarm probabilities of the PSER entropy detector ($m = 200, 500, 1000, N = 512$).

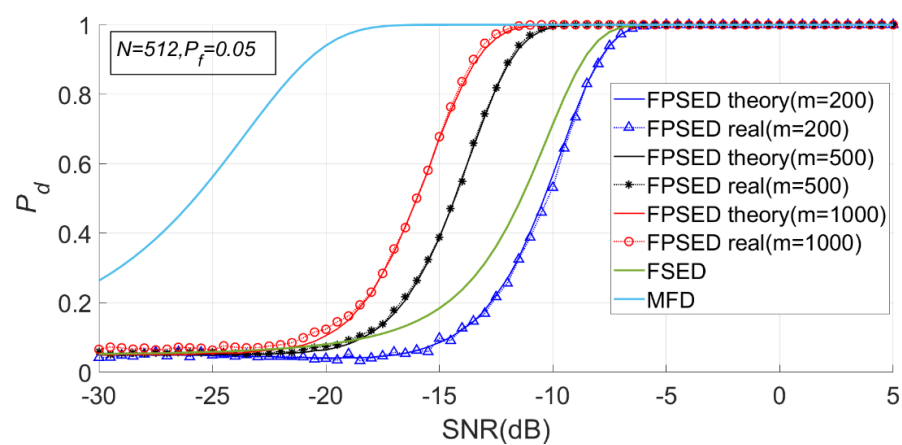


Figure 20. Detection probabilities of full spectrum energy detection (FSED), matched-filtering detector (MFD) and PSER entropy detectors ($m = 200, 500, 1000, N = 512$).

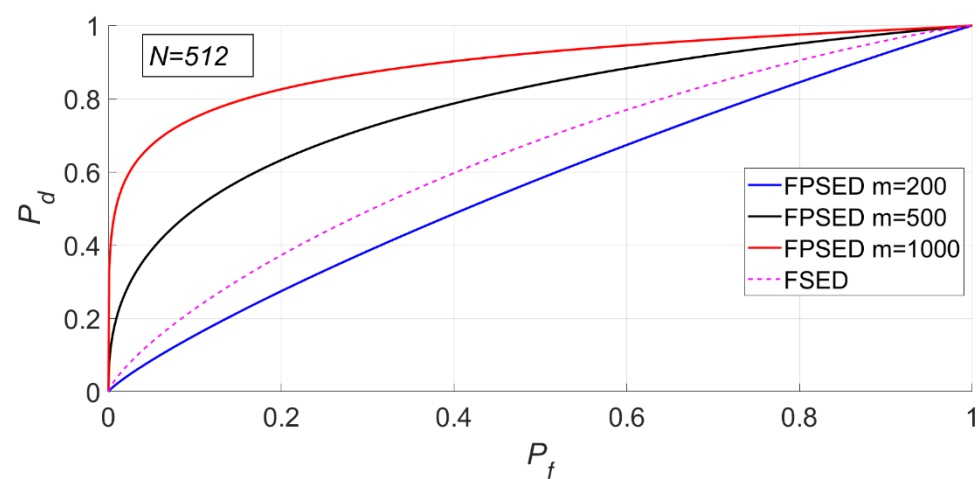


Figure 21. Receiver operating characteristic (ROC) curve of FSED and PSER entropy detectors ($m = 200, 500, 1000, N = 512$).

Table 2. Mean-square errors (MSEs) between actual and theoretical probabilities ($N = 512$).

Probability	$m = 200$	$m = 500$	$m = 1000$
P_f	0.3059×10^{-4}	0.3887×10^{-4}	2.6785×10^{-4}
P_d	0.5723×10^{-4}	0.2777×10^{-4}	1.1280×10^{-4}

The deviation between the actual and theoretical detection probabilities was very small, which indicated that the PSER entropy detector was accurate.

When m is 500, N is 256, 512, and 1024, respectively. The detection results for the BSPK signal are shown in Figures 22–24.

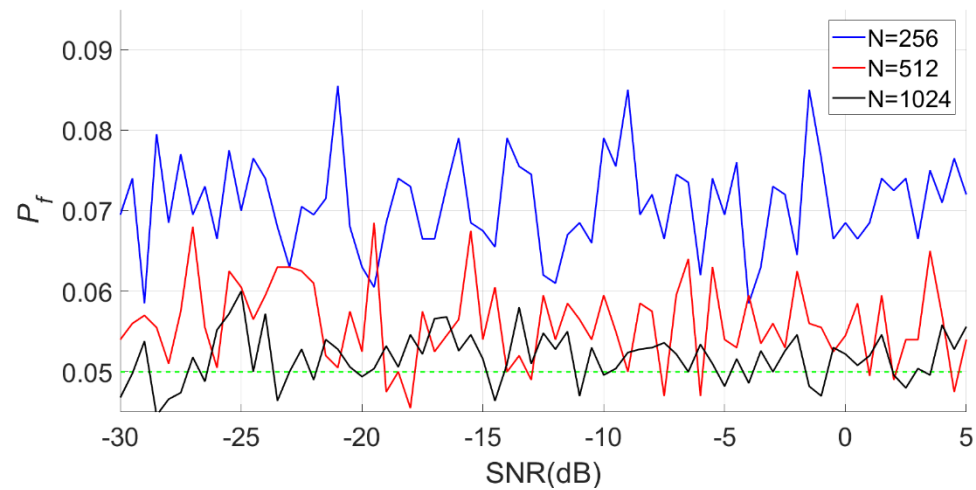


Figure 22. Actual false alarm probabilities of the PSER entropy detectors ($m = 500$, $N = 256, 512, 1024$).

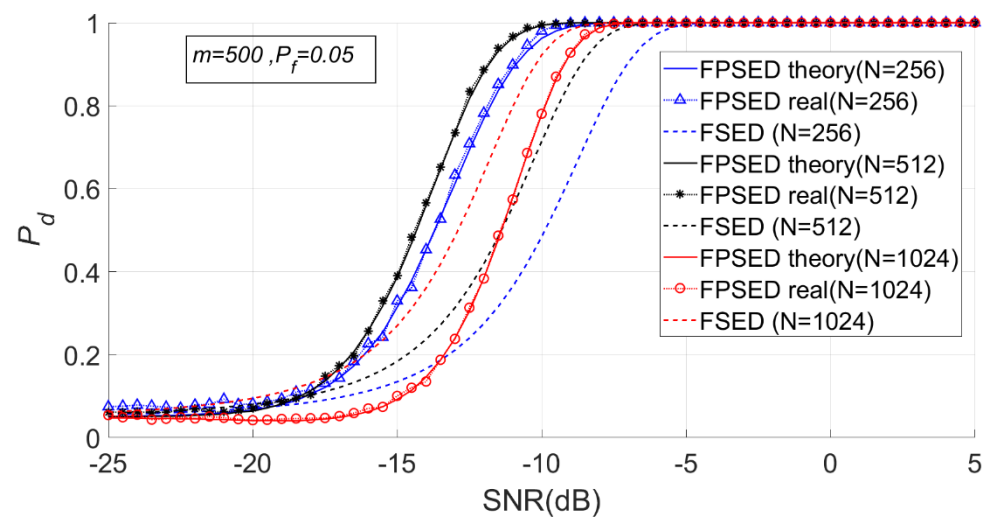


Figure 23. Detection probabilities of FSED and PSER entropy detectors ($m = 500$, $N = 256, 512, 1024$).

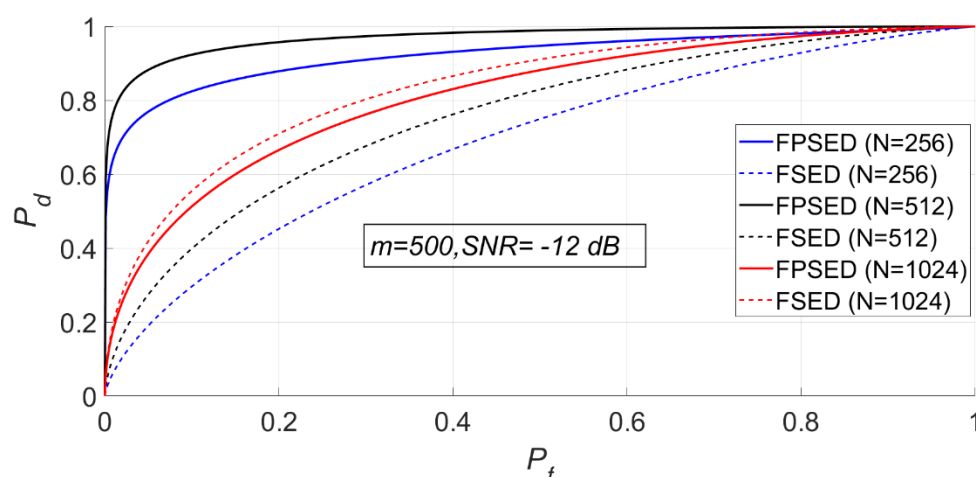


Figure 24. ROC of FSED and PSED entropy detectors ($m = 500$, $N = 256, 512, 1024$).

It can be seen from Figures 23 and 24 that when m is fixed, a larger N does not necessarily imply a better detection performance. The detection probability when N is 1024 is lower than that when N is 512. However, the detection performance of the full spectrum energy detection method will improve with the increase of N .

The MSEs of the actual and theoretical probabilities of these experiments are given in Table 3. The deviations between the actual and theoretical detection probabilities was very small when was 512 or 1024, however were higher when N was 256.

Table 3. MSEs between actual and theoretical probabilities ($m = 500$).

Probability	N = 256	N = 512	N = 1024
P_f	4.687958×10^{-4}	0.622887×10^{-4}	0.125177×10^{-4}
P_d	2.067099×10^{-4}	0.398899×10^{-4}	0.145410×10^{-4}

5. Discussions

5.1. Theoretical Calculation of Statistics

In Section 2, we analyzed the computational time complexity of each statistic. When m and N are large, the theoretical values of some statistics, such as $\text{Var}(Z)$ under H_0 , H_i , $H_{i,j}$, $H(m, N)$, and $\text{Var}(Z)$ under H_1 , cannot be calculated. This restricts the further analysis on the detection performance of the PSED entropy detector.

There are two ways to solve this problem: reducing the computational complexity and finding an approximate solution. Which way is feasible requires further study.

5.2. Experience of Selecting Parameters

The detection probability P_d of PSED entropy detection is related to the number of intervals m , the number of power spectrum lines N , and the SNR of the spectrum lines. Since the mathematical expressions of many statistics are too complex, the influence of the three factors on P_d cannot be accurately analyzed at present. Based on a large number of experiments, we summarize the following experiences as recommendations for setting parameters.

- (1) m cannot be too small. It can be seen from Figure 20, that, if m is too small, then the detection performance of the PSED entropy detector will be lower than that of energy detector. We suggest that $m \geq 500$.
- (2) N must be close to m . It can be seen from Figure 23, that N is not better when bigger. A large number of experiments show that when N is close to m , the detection probability is good.
- (3) When N is fixed, m can be adjusted appropriately through experiments.

5.3. Advantages of the PSER Entropy Detector

When using PSER entropy detection, the noise intensity does not need to be estimated in advance, and prior information of the signal to be detected is not needed. Therefore, the PSER entropy detector is a typical blind signal detection method.

5.4. Further Research

The detection performance of the PSER entropy detector will be further improved if some methods are used to improve the SNR of signals. In future research, a denoising method can be used to improve the SNR, the Welch or autoregressive method can be used to improve the quality of power spectrum estimation, and multi-channel cooperative detection can be used to increase the accuracy of detection.

6. Conclusions

In this paper, the statistical characteristics of PSER entropy are derived through strict mathematical analysis, and the theoretical formulas for calculating the PSER entropy and sample entropy variance from pure noise and mixed signals are obtained. In the process of derivation, we do not use the classical method of approximating PSER entropy using differential entropy, but use interval division and the multinomial distribution to calculate PSER entropy. The calculation results of this method are consistent with the simulation results, which shows that this method is correct. This method is not only suitable for a large number of intervals, but also suitable for a small number of intervals. A signal detector based on the PSER entropy was created according to these statistical characteristics. The performance of the PSER entropy detector is obviously better than that of the classical energy detector. This method does not need to estimate the noise intensity or require any prior information of the signal to be detected, and therefore it is a complete blind signal detector.

The PSER entropy detector can not only be used in spectrum sensing, but also in vibration signal detection, seismic monitoring, and pipeline safety monitoring.

Author Contributions: Conceptualization, Y.H. and H.L.; formal analysis, Y.H. and H.L.; investigation, Y.H. and H.L.; methodology, Y.H. and H.L.; validation Y.H., S.W. and H.L.; writing—original draft, H.L.; writing—review and editing, Y.H. and S.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Technology Nova Plan of Beijing City (No. Z201100006820122) and Fundamental Research Funds for the Central Universities (No. 2020RC14).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data is simulated.

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

References

1. Gupta, M.S.; Kumar, K. Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues. *J. Netw. Comput. Appl.* **2019**, *143*, 47–76. [\[CrossRef\]](#)
2. Arjoune, Y.; Kaabouch, N. A Comprehensive Survey on Spectrum Sensing in Cognitive Radio Networks: Recent Advances, New Challenges, and Future Research Directions. *Sensors* **2019**, *19*, 126. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Kabeel, A.A.; Hussein, A.H.; Khalaf, A.A.M.; Hamed, H.F.A. A utilization of multiple antenna elements for matched filter based spectrum sensing performance enhancement in cognitive radio system. *AEU Int. J. Electron. C* **2019**, *107*, 98–109. [\[CrossRef\]](#)
4. Chatziantoniou, E.; Allen, B.; Velisavljevic, V.; Karadimas, P.; Coon, J. Energy Detection Based Spectrum Sensing Over Two-Wave with Diffuse Power Fading Channels. *IEEE Trans. Veh. Technol.* **2017**, *66*, 868–874. [\[CrossRef\]](#)
5. Reyes, H.; Subramaniam, S.; Kaabouch, N.; Hu, W.C. A spectrum sensing technique based on autocorrelation and Euclidean distance and its comparison with energy detection for cognitive radio networks. *Comput. Electr. Eng.* **2016**, *52*, 319–327. [\[CrossRef\]](#)
6. Yucek, T.; Arslan, H. A survey of spectrum sensing algorithms for cognitive radio applications. *IEEE Commun. Surv. Tutor.* **2009**, *11*, 116–130. [\[CrossRef\]](#)

7. Gismalla, E.H.; Alsusa, E. Performance Analysis of the Periodogram-Based Energy Detector in Fading Channels. *IEEE Trans. Signal Process.* **2011**, *59*, 3712–3721. [\[CrossRef\]](#)
8. Dikmese, S.; Ilyas, Z.; Sofotasios, P.C.; Renfors, M.; Valkama, M. Sparse Frequency Domain Spectrum Sensing and Sharing Based on Cyclic Prefix Autocorrelation. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 159–172. [\[CrossRef\]](#)
9. Li, H.; Hu, Y.; Wang, S. Signal Detection Based on Power-Spectrum Sub-Band Energy Ratio. *Electronics* **2021**, *10*, 64. [\[CrossRef\]](#)
10. Zhang, Y.L.; Zhang, Q.Y.; Melodia, T. A frequency-domain entropy-based detector for robust spectrum sensing in cognitive radio networks. *IEEE Commun. Lett.* **2010**, *14*, 533–535. [\[CrossRef\]](#)
11. Prieto, G.; Andrade, A.G.; Martinez, D.M.; Galaviz, G. On the Evaluation of an Entropy-Based Spectrum Sensing Strategy Applied to Cognitive Radio Networks. *IEEE Access* **2018**, *6*, 64828–64835. [\[CrossRef\]](#)
12. Nagaraj, S.V. Entropy-based spectrum sensing in cognitive radio. *Signal Process.* **2009**, *89*, 174–180. [\[CrossRef\]](#)
13. Gu, J.; Liu, W.; Jang, S.J.; Kim, J.M. Spectrum Sensing by Exploiting the Similarity of PDFs of Two Time-Adjacent Detected Data Sets with Cross Entropy. *IEICE Trans. Commun.* **2011**, *E94B*, 3623–3626. [\[CrossRef\]](#)
14. Zhang, Y.; Zhang, Q.; Wu, S. Entropy-based robust spectrum sensing in cognitive radio. *IET Commun.* **2010**, *4*, 428. [\[CrossRef\]](#)
15. Nikonowicz, J.; Jessa, M. A novel method of blind signal detection using the distribution of the bin values of the power spectrum density and the moving average. *Digit. Signal Process.* **2017**, *66*, 18–28. [\[CrossRef\]](#)
16. Zhao, N. A Novel Two-Stage Entropy-Based Robust Cooperative Spectrum Sensing Scheme with Two-Bit Decision in Cognitive Radio. *Wirel. Pers. Commun.* **2013**, *69*, 1551–1565. [\[CrossRef\]](#)
17. Ejaz, W.; Shah, G.A.; Ul Hasan, N.; Kim, H.S. Optimal Entropy-Based Cooperative Spectrum Sensing for Maritime Cognitive Radio Networks. *Entropy* **2013**, *15*, 4993–5011. [\[CrossRef\]](#)
18. So, J. Entropy-based Spectrum Sensing for Cognitive Radio Networks in the Presence of an Unauthorized Signal. *KSII Trans. Internet Inf.* **2015**, *9*, 20–33.
19. Ye, F.; Zhang, X.; Li, Y. Collaborative Spectrum Sensing Algorithm Based on Exponential Entropy in Cognitive Radio Networks. *Symmetry* **2016**, *8*, 112. [\[CrossRef\]](#)
20. Zhu, W.; Ma, J.; Faust, O. A Comparative Study of Different Entropies for Spectrum Sensing Techniques. *Wirel. Pers. Commun.* **2013**, *69*, 1719–1733. [\[CrossRef\]](#)
21. Islam, M.R.; Uddin, J.; Kim, J. Acoustic Emission Sensor Network Based Fault Diagnosis of Induction Motors Using a Gabor Filter and Multiclass Support Vector Machines. *Adhoc Sens. Wirel. Netw.* **2016**, *34*, 273–287.
22. Akram, J.; Eaton, D.W. A review and appraisal of arrival-time picking methods for downhole microseismic data. *Geophysics* **2016**, *81*, 71–91. [\[CrossRef\]](#)
23. Legese Hailemariam, Z.; Lai, Y.C.; Chen, Y.H.; Wu, Y.H.; Chang, A. Social-Aware Peer Discovery for Energy Harvesting-Based Device-to-Device Communications. *Sensors* **2019**, *19*, 2304. [\[CrossRef\]](#)
24. Pei-Han, Q.; Zan, L.; Jiang-Bo, S.; Rui, G. A robust power spectrum split cancellation-based spectrum sensing method for cognitive radio systems. *Chin. Phys. B* **2014**, *23*, 537–547.
25. Mei, F.; Hu, C.; Li, P.; Zhang, J. Study on main Frequency precursor characteristics of Acoustic Emission from Deep buried Dali Rock explosion. *Arab. J. Geosciences* **2019**, *12*, 645. [\[CrossRef\]](#)
26. Moddemeijer, R. On estimation of entropy and mutual information of continuous distributions. *Signal Process.* **1989**, *16*, 233–248. [\[CrossRef\]](#)
27. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2006; p. 737.
28. Chung, K.L.; Sahlia, F.A. *Elementary Probability Theory, with Stochastic Processes and an Introduction to Mathematical Finance*; Springer: New York, NY, USA, 2004; p. 402.
29. Sarker, M.B.I. *Energy Detector Based Spectrum Sensing by Adaptive Threshold for Low SNR in CR Networks*; Wireless and Optical Communication Conference: New York, NY, USA, 2015; pp. 118–122.