entropy

Article

Low Complexity Estimation Method of Rényi Entropy for Ergodic Sources [†]

Young-Sik Kim

Department of Information and Communication Engineering, Chosun University, 309 Pilmoondae-ro Dong-gu, Gwangju 61452, Korea; iamyskim@chosun.ac.kr; Tel.: +82-62-230-7032

[†] This paper is an extended version of my paper published in the 2014 International Symposium on Information Theory and Its Applications, Melbourne, VIC, Australia, 26–29 October 2014.

Received: 11 July 2018; Accepted: 30 August 2018 ; Published: 31 August 2018

Abstract: Since the entropy is a popular randomness measure, there are many studies for the estimation of entropies for given random samples. In this paper, we propose an estimation method of the Rényi entropy of order α. Since the Rényi entropy of order α is a generalized entropy measure including the Shannon entropy as a special case, the proposed estimation method for Rényi entropy can detect any significant deviation of an ergodic stationary random source's output. It is shown that the expected test value of the proposed scheme is equivalent to the Rényi entropy of order α. After deriving a general representation of parameters of the proposed estimator, we discuss on the particular orders of Rényi entropy such as $\alpha \to 1$, $\alpha = 1/2$, and $\alpha = 2$. Because the Rényi entropy of order 2 is the most popular one, we present an iterative estimation method for the application with stringent resource restrictions.

Keywords: entropy estimation; Shannon entropy; Rényi entropy; quadratic entropy; random number generation; nearest neighbor distance; security

1. Introduction

Since the entropy is a popular randomness measure, many studies are devoted to the efficient estimation of the Shannon or Rényi entropy for given random samples. In particular, one of the important applications for entropy estimator is random number generators (RNGs). RNG is one of the fundamental cryptographic primitives and a good RNG can be modelled as an ergodic random source. For block ciphers and public key cryptography, an RNG also can be used as a key-stream generator. In addition, the digital signature algorithm (DSA) requires a random number for its computation [1]. Since a statistical bias in random numbers can be exploited to reduce the computational complexity of the exhaustive search by an attacker, the entire security of the crypto-systems usually depends on the statistically quality of RNG output.

In order to obtain the unpredictability of random output, many crypto-systems require a true (physical) random number generator (TRNG) [2–4] as well as pseudo-random number generators (PRNG). However, a TRNG output can be easily influenced by environments such as temperature, electro-magnetic wave, and so on. Therefore, an on-the-fly statistical test scheme, as known as online test, is requested to guarantee the statistical quality of RNG output in cryptographic standards [5]. In particular for the applications with stringent resource constraints such as sensor nodes, smart cards, etc., an online test scheme for RNG should both have compact size in hardware/software and detect a various range of statistical bias to ensure the security of the systems [6]. That is, we need a good and efficient test scheme of the statistical quality of random sources. Therefore, it is highly desirable to construct low-cost and reliable entropy estimation methods of the random output of TRNG.

When it comes to the randomness measure, the Shannon entropy is one of the widely used measures. Let $F_2 = \{0, 1\}$ be the finite field with two elements. Let X be a random variable for L-bit random symbols in F_2^L from the random source S and the probability of occurring random symbol b from the random source S is denoted by $\Pr(b)$. Then, the Shannon entropy of L-bit blocks from the random source S is defined as

$$H_L(X) = -\sum_{b \in F_2^L} \Pr(b) \log_2 \Pr(b). \tag{1}$$

In various literature, there have been developed efficient estimators of the Shannon entropy [7–11]. In addition, the complexity of estimating the Shannon entropy of a distribution on k elements from independent samples also has been developed [10,12,13].

In 1961, another generalized entropy measure is defined by Rényi [14]. The Rényi entropy is popularly used in a number of signal processing and pattern recognition applications [15–17]. For example, Rényi entropy has been used in cryptography, in the study of bio-informatics, and in the bio-medical applications [18]. Sometimes, the Rényi entropy can provide more strict randomness measure for the cryptographic applications such as the privacy amplification [19]. Rényi entropy of order α for L-bit blocks from the random source S is defined as

$$R_{\alpha,L}(X) = \frac{1}{1-\alpha} \log_2 \sum_{b \in F_2^L} \Pr(b)^{\alpha}, \tag{2}$$

where $\alpha > 0$ and $\alpha \neq 1$. Note that Rényi entropy is defined as the log base 2 of the expectation of $\Pr(b)^{\alpha-1}$ normalized by $1 - \alpha$.

Generally, a random source used in a cryptographic protocol should have a maximum entropy. As a result, Shannon entropy (or Rényi entropy) is recognized as one important measure of randomness. For example, standards such as NIST STS (Statistical Test Suits) [20] or AIS.31 [5], a widely used standard for evaluating TRNGs, include Entropy Estimation items. The proposed method estimates the actual Rényi entropy value very accurately, as can be confirmed from the simulation results. Therefore, if the estimation result shows a lower value (from the maximum), it can be interpreted as a signal that the random sources are generating a significant deviated output from the perfect.

The Rényi entropy is a generalization of the Shannon entropy since it contains the definition of the Shannon entropy when α approaches to one [21]. In particular, notice that it is easy to prove the Rényi entropy is always less than or equal to Shannon entropy by using Jensen's inequality [22] for $\alpha > 1$, i.e.,

$$R_{\alpha,L}(X) \leq H_L(X). \tag{3}$$

Here, the equality holds for the equiprobable random source. Therefore, Rényi entropy can be used as the lower bound of the Shannon entropy for a random source S.

There are some studies for the estimation of Rényi entropy [15,23,24]. In particular, many applications such as machine learning [16], blind deconvolution of linear channels [25], information flows in financial data [26] and cryptography [19] have paid attention to the Rényi entropy of order 2, also called the *quadratic entropy* or *collision entropy* [15]. The most straightforward approach for entropy estimation is the direct calculation of entropy based on the probability mass function (pmf) or the probability density function (pdf) of empirical data. For example, Erdogmus and Principe proposed the Rényi entropy or order estimator by using the non-parametic estimation of the pdf of a random variable [16]. In order to estimate the pdf of a given sample distribution, they used the Parzen windowing method, in which the pdf is approximated by the sum of kernels such as the Gaussian function. In addition, there are many non-parametric approaches to estimate entropies, which are usually based on the data compression [27] or the nearest neighbor distance [7,8,28,29]. There are results on the estimation of Rényi entropy rate of Markov chains [30]. This can be also considered for randomness measure for RNG because a skewed RNG can be modelled as Markov chain with vaying transition probability. However, many of them are not suitable in constrained devices with stringent resource restriction due to the computational complexity.

In this paper, we propose a low complexity estimation method of the Rényi entropy of order α, where α is a real number. This method does not require any initialization phase contrary to the previous Maurer's universal statistical test [8] and Coron's refined test [7] which are widely used for estimating Shannon entropy especially in cryptographic applications [5,20]. In the proposed scheme, we can estimate the Rényi entropy of order α without introducing complicated computations such as logarithms or divisions. We show that the expected value of the proposed test function is equivalent with the Rényi entropy of order α. That is, the output of the proposed estimation method almost surely converges to the Rényi entropy of order α values with large samples. Because the requirement of a large sample size for accurate estimation can be a drawback of the proposed scheme in some applications, we also propose an iterative algorithm for the Rényi entropy of order 2, which requires a relatively short sample size and shows accurate test results. Using the simple counting method, we can efficiently implement test module of RNGs based on the Rényi entropy of order 2. Therefore, it can be used as a statistical tester of RNGs for many embedded security systems such as smart cards. This is an extended version of the conference paper [31]. In this paper, we generalized the order from integer to real number. Also, iterative estimation method is included for more constrained environments.

The remainder of this paper is organized as follows: in Section 2, previous entropy estimation schemes are reviewed. In particular, for the comparison, Maurer's universal statistical test and Coron's refined scheme are presented in detail since they are the popular nearest neighbor distance based schemes, which are exploited by the proposed scheme. In Section 3, the proposed estimation scheme for the Rényi entropy of order α is presented. Firstly, we describe the estimation method and show that the expected value of the proposed test function is equivalent with the Rényi entropy of order α. We suggest the iterative estimation algorithm, which requires a relatively small sample size. Then, numerical results are given in Section 4. For the three block sizes and two sample sizes, the estimated values are compared with the Rényi entropy of order 2 or order $1/2$. For the iterative algorithm, we can check that the proposed algorithm is more stable with less sample size. Finally, we conclude this paper in Section 5.

2. Previous Works

2.1. The Nearest Neighbor Distance

In 1992, Maurer proposed the universal statistical test for evaluating statistical quality of random number generators [8]. By universal, this method can detect various kinds of statistical defects in random data. Maurer conjectured that the test result of his estimation method is related to Shannon entropy of L-bit blocks. Later, this conjecture was proved by Coron and Naccache [32]. The Maurer's universal statistical test is included in the statistical test suite by NIST for evaluating RNGs in cryptographic applications [20].

Let $N = (Q + K)L$. Let s^N denote the generated random sequence with length N. In the Maurer's universal statistical test, an initialization phase is required for the first Q L-bit blocks. In order to make it so each of 2^L blocks occurs at least once during the initialization phase with high probability, the size of Q should be greater than 10×2^L [8]. Then, in the evaluation phase, the next K L-bit blocks are used for the entropy estimation.

Let $b_n(s^N) = [s_{L(n-1)+1}, \cdots, s_{Ln}]$ be the n-th L-bit block of s^N. Then, the Maurer's universal statistical test is based on the following test function:

$$f_M(s^N) = \frac{1}{K} \sum_{n=Q+1}^{K+Q} \log_2 D_n(s^N),$$

where

$$D_n(s^N) = \begin{cases} n, & \text{if } \forall i < n, b_n(s^N) \neq b_{n-i}(s^N), \\ \min\{i \mid i \geq 1, b_n(s^N) = b_{n-i}(s^N)\}, & \text{otherwise.} \end{cases} \tag{4}$$

That is, $D_n(s^N)$ is a distance between the index of the current pattern b and the nearest previous index of the same pattern b. Note that this distance $D_n(s^N)$ is conversely proportional to the probability of occurring the pattern b. That is, if the probability of occurring the pattern b is small, the expected value of the distance $D_n(s^N)$ is large, and vice versa. In fact, Coron and Naccache proved that Maurer's universal statistical test is closely related to the Shannon entropy for a source emitting the sequence of binary random variables, $U^N = U_1, \cdots, U_N$ as follows [7]:

$$\lim_{L\to\infty} \left[E[f_M(U^N)] - H_L(U^N)\right] = \int_0^\infty e^\zeta \log_2 \zeta d\zeta \cong -0.8327462. \tag{5}$$

Later, Coron refined the Maurer's universal statistical test as an exact entropy estimator without the numerical discrepancy presented in (5) [7].

Therefore, Coron's refined test is adopted as an estimating method of Shannon entropy in AIS.31 specification, the German standard for cryptographic TRNGs [5]. Coron modified test function as

$$f_C(s^N) = \frac{1}{K} \sum_{n=Q+1}^{K+Q} g(D_n(s^N)), \tag{6}$$

where

$$g(i) = \frac{1}{\ln(2)} \sum_{k=1}^{i-1} \frac{1}{k} \tag{7}$$

and K and Q are given as the same parameters of the Maurer's universal statistical test. Then, he proved that the expected value of the test function $f_C(s^N)$ is equal to the Shannon entropy of L-bit blocks of the random source as follows [7]:

$$E[f_C(U^N)] = H_L(U^N).$$

Note that, in Coron's test, a logarithm is substituted by a summation of a series of integer divisions. In his paper, he proposed the approximated method in order to reduce computational complexity as follows [7]:

$$\sum_{k=1}^{i-1} \frac{1}{k} \approx \frac{1}{\ln(2)} \ln(i-1) + \frac{1}{2(i-1)} - \frac{1}{12(i-1)^2} + \mathcal{O}(\frac{1}{(i-1)^4}) - 0.577216. \tag{8}$$

2.2. Previous Entropy Estimation Approach

One of the widely used estimators for Shannon or Rényi entropy is the "plug-in" estimator. The "plug-in" approach estimates parameters and then substitutes them into the entropy function, (1) or (2). For example, let N_k be the frequency of the symbol k and N the total number of samples. Then, we have probability mass function (pmf) $p_k = N_k/N$. Using this pmf, we can directly calculate entropy of the given sample using (1) or (2).

Since the pdf of a random sample is unknown a priori, the estimation of pdf of a random variable is complicated and usually contains some complex functions. For example, Erdogmus and Principe proposed the estimation of entropy based on the non-parametic direct estimation of pdf, that is, the Parzen window method in the context of minimizing error entropy [16]. The Parzen estimator of the error pdf $f_e(\zeta)$ is given by

$$\hat{f}_e(\zeta) = \frac{1}{N}\sum_{i=1}^{N} \kappa(\zeta - e_i, \sigma^2), \tag{9}$$

where N is the sample size and κ is the kernel function, usually implemented by using the multidimensional Gaussian function with a radially symmetric variance σ^2. Then, we can directly calculate the Rényi entropy or Shannon entropy using the estimated pdf in (9). However, the computation of the sum of the Gaussian functions is usually infeasible in most constrained devices.

3. New Estimation Method of the Rényi Entropy of Order α

In this section, we derive the parameters for the estimation of the Rényi entropy of order α, for a real number α. In this derivation, we assume that an ergodic random source S and random sequences from the source S are over F_2 and consecutive and distinct L symbols are treated as a basic element of test function. Thus, the maximum value of Rényi entropy (and also Shannon entropy) will be L-bit. For the estimation of Rényi entropy, we firstly focus on the estimation of inner summation in (2). Then, to obtain exact value of Rényi entropy, the logarithm and division by $1 - \alpha$ will be applied to the result of the estimation. The test function is given as

$$f(s^N) = \frac{1}{K}\sum_{n=1}^{K} g(D_n(s^N)), \tag{10}$$

where $D_n(s^N)$ is the index distance defined in (4). Now, for given real number α and the index distance $D_n(s^N) = k$, we are going to find the values of $g(k)$ for each $k \geq 1$ which is closely related to the inner summation (2) for the estimation of Rényi entropy of order α. The following theorem gives us the general representation of $g(k)$ for given α.

Main Result: Proposed Test Function of Rényi Entropy of Order α

For the estimation of Rényi entropy of order α, the parameters $g(k)$ of estimator for given index distance k in (10) are given as

$$g(k) = \begin{cases} 1, & \textit{if } k = 1, \\ (-1)^{k-1} P_{k-1}^{\alpha-2}, & \textit{if } k \geq 2, \end{cases} \tag{11}$$

where

$$P_{k-1}^{\alpha-2} = \binom{\alpha-2}{k-1} = \frac{(\alpha-2)(\alpha-3)\cdots(\alpha-k)}{(k-1)!}. \tag{12}$$

The derivation of the proposed test function can be justifed as in the following proof.

Proof. We start from the the expectation of the test function $f(s^N)$ given as

$$E[f(U_S^N)] = \sum_{k=1}^{\infty} \Pr[D_n(U_S^N) = k] g(k), \tag{13}$$

where U_S^N is a vector of random variables for random sequence s^N of L-bit symbols and $g(k)$ is the k-th parameter for the estimation. Then, the probability $\Pr[D_n(U_S^N) = k]$ can be represented as

$$\Pr[D_n(U_S^N) = k] = \sum_{b \in B^L} \Pr[b_n = b, b_{n-1} \neq b, \cdots, b_{n-k+1} \neq b, b_{n-k} = b].$$

If the random variable is stationary, we have

$$\Pr[D_n(U_S^N) = k] = \sum_{b \in B^L} \Pr[b]^2 (1 - \Pr[b])^{k-1}. \tag{14}$$

From (13) and (14), the expectation can be represented as

$$E[f(U_S^N)] = \sum_{b \in B^L} \Pr[b] \gamma(\Pr[b]),$$

where $\gamma(\cdot)$ is defined as

$$\gamma(x) = x \sum_{k=1}^{\infty} (1-x)^{k-1} g(k).$$

Here, we are going to find the representation of $g(k)$ that satisfies the expected value of $E(f(U_S^N)) = \sum_{b \in F_2^L} \Pr(b)^{\alpha}$. Then, we have

$$\gamma(x) = x \sum_{k=1}^{\infty} (1-x)^{k-1} g(k) = x^{\alpha-1}.$$

By removing x at both sides, the equation is simplified as

$$\sum_{k=1}^{\infty} (1-x)^{k-1} g(k) = x^{\alpha-2}.$$

By substituting $x = 1 - t$, we have

$$\sum_{k=1}^{\infty} t^{k-1} g(k) = (1-t)^{\alpha-2}. \tag{15}$$

From (15), for $\alpha = 2$, we have $\sum_{k=1}^{\infty} t^{k-1} g(k) = 1$. That is, $g(1) = 1$, otherwise $g(k) = 0$. For $\alpha \neq 2$, the Tayler series at $t = 0$ of the right hand side of (15), $(1-t)^{\alpha-2}$, is given as

$$(1-t)^{\alpha-2} = \sum_{k=0}^{\infty} \binom{\alpha-2}{k} t^k (-1)^k, \tag{16}$$

where

$$\binom{\alpha-2}{k} = P_k^{\alpha-2} = \frac{(\alpha-2)(\alpha-3)\cdots(\alpha-1-k)}{k!}$$

and $P_0^{\alpha-2} = 1$. Note that the combination in (16) is a generalized binomial expansion for the real number α and a positive integer k [33]. Thus, we have

$$\sum_{k=0}^{\infty} t^k g(k+1) = \sum_{k=0}^{\infty} (-1)^k P_k^{\alpha-2} t^k.$$

Finally, the parameter $g(k)$ of the estimator for the exact Rényi entropy of order α for a real number α is given as

$$g(k) = \begin{cases} 1, & \text{if } k = 1, \\ (-1)^{k-1} P_{k-1}^{\alpha-2}, & \text{if } k \geq 2. \end{cases}$$

□

Table 1 shows examples of parameters of the proposed estimator for some cases of α. For the integer $\alpha \geq 3$, we can see the negative values of $g(k)$. This means that the test function in (10) may be negative after accumulation of parameters for given random samples. Therefore, in this case, we need to take absolute value of test result before applying logarithm to calculate actual Rényi entropy of order α.

Table 1. Values of the parameter $g(k)$ of the proposed estimator for some α's.

$\alpha \backslash k$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	$\cdots$	n
$\frac{1}{2}$	1	$\frac{3}{2}$	$\frac{15}{8}$	$\frac{35}{16}$	$\frac{315}{128}$	$\frac{693}{256}$	$\frac{3003}{1024}$	$\frac{6435}{2048}$	$\frac{109395}{32768}$	$\frac{230945}{65536}$	$\cdots$	$\frac{(2n-1)!}{4^{n-1}((n-1)!)^2}$
$\alpha \to 1$	1	1	1	1	1	1	1	1	1	1	$\cdots$	1
2	1	0	0	0	0	0	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{1}{n-1}$
3	1	-1	0	0	0	0	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{2}{n-1}$
4	1	-2	1	0	0	0	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{3}{n-1}$
5	1	-3	3	-1	0	0	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{4}{n-1}$
6	1	-4	6	-4	1	0	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{5}{n-1}$
7	1	-5	10	-10	5	-1	0	0	0	0	$\cdots$	$(-1)^{n-1}\binom{6}{n-1}$
8	1	-6	15	-20	15	-6	1	0	0	0	$\cdots$	$(-1)^{n-1}\binom{7}{n-1}$
9	1	-7	21	-35	35	-21	7	-1	0	0	$\cdots$	$(-1)^{n-1}\binom{8}{n-1}$
10	1	-8	28	-56	70	-56	28	-8	1	0	$\cdots$	$(-1)^{n-1}\binom{9}{n-1}$

Now we are going to derive representation of $g(k)$ for some particular orders, $\alpha \to 1$, $\alpha = 1/2$, and $\alpha = 2$ in the next subsections. First, let us start from the case of α approaching to 1, where Rényi entropy converges to Shannon entropy.

3.1. Convergence of Rényi Entropy and Shannon Entropy

The proposed estimation method converges to the same estimator of Shannon entropy by Coron [7] as in the following theorem.

Theorem 1. *The proposed test function of Rényi entropy converges to the test function of Shannon entropy by Coron when α goes to 1.*

Proof. If $\alpha \to 1$, from (11), the parameter converges

$$g(k) \to (-1)^{k-1}\frac{(-1)(-2)\cdots(-(k-1))}{(k-1)!} = 1. \tag{17}$$

This means that every case will be counted and the test function $f(s^N)$ in (10) will always converge to 1. To obtain actual value of Rényi entropy, the test function $f(s^N)$ should be applied by the logarithm of base 2 and divided by $(1-\alpha) \to 0$. To obtain the converged value for $\alpha \to 1$, we can use L'Hospital's theorem to $\frac{\log_2 f(s^N)}{1-\alpha}$. Then, we have

$$\begin{aligned}
\lim_{\alpha\to 1}\frac{\log_2 f(s^N)}{1-\alpha} &= \lim_{\alpha\to 1}\frac{d(\log_2 f(s^N))}{d\alpha}\frac{1}{\frac{d(1-\alpha)}{d\alpha}} \\
&= \lim_{\alpha\to 1} -\frac{1}{f(s^N)\ln 2}\frac{df(s^N)}{d\alpha} \\
&= \lim_{\alpha\to 1} -\frac{1}{f(s^N)\ln 2}\frac{1}{K}\sum_{n=1}^{K}\frac{dg(D_n(s^N))}{d\alpha}.
\end{aligned} \tag{18}$$

From (11), we can obtain the derivative of $g(k)$ with respect to α as

$$\begin{aligned}\lim_{\alpha\to 1}\frac{dg(D_n(s^N))}{d\alpha} &= (-1)^{i-1}\left[\frac{(\alpha-3)\cdots(\alpha-i)}{(i-1)!}+\frac{(\alpha-2)(\alpha-4)\cdots(\alpha-i)}{(i-1)!}\right.\\ &\quad \left.+\cdots+\frac{(\alpha-2)\cdots(\alpha-(i-1))(\alpha-i)}{(i-1)!}\right]_{\alpha=1}\\ &= -1-\frac{1}{2}-\frac{1}{3}-\cdots-\frac{1}{i-1}=-\sum_{k=1}^{i-1}\frac{1}{k}. \qquad (19)\end{aligned}$$

From (17), we also have $f(s^N)=\frac{1}{K}\sum_{n=1}^{K}1=1$ when α goes 1. Therefore, from (18) and (19), we have

$$\lim_{\alpha\to 1}\frac{\log_2 f(s^N)}{1-\alpha}=\frac{1}{K}\sum_{n=1}^{K}G(D_n(s^N)),$$

where $G(D_n(s^N))=\frac{1}{\ln 2}\sum_{k=1}^{i-1}\frac{1}{k}$ for $D_n(s^N)=i$. It is exactly the same result by Coron given in (6) and (7) [32]. □

That is, the proposed estimator includes the previous result as a special case and it can be considered as evidence that the proposed approach is valid for the entropy estimation.

3.2. Proposed Test Function for Rényi Entropy of Order $\frac{1}{2}$

In this subsection, we will derive a simplified test function for Rényi entropy of order 1/2. This order of Rényi entropy is closely related to the exponent of the average growth rate of average guesswork [34,35]. From the main result, for $k>1$ and $\alpha=\frac{1}{2}$, we have

$$\begin{aligned}g(k) &= (-1)^{k-1}\binom{-\frac{3}{2}}{k-1}=(-1)^{k-1}\frac{(-\frac{3}{2})(-\frac{5}{2})\cdots(-\frac{2k-1}{2})}{(k-1)!}\\ &= (-1)^{2(k-1)}\frac{1\times 3\times\cdots\times(2k-1)}{2^{k-1}(k-1)!}\\ &= \frac{1\times 2\times 3\times 4\times\cdots\times(2k-2)(2k-1)}{2^{2(k-1)}[(k-1)!]!}\\ &= \frac{(2k-1)!}{4^{k-1}[(k-1)!]}. \qquad (20)\end{aligned}$$

However, the calculation of the factorial is a complicated task and takes a long time even for the moderate size of integer. Therefore, we need to simplify (20) for the practical implementation. First, we can use the Stirling's approximation given as

$$k!\approx k^k e^{-k}\sqrt{2\pi k}.$$

Then, we have

$$\begin{aligned}\frac{(2k-1)!}{4^{k-1}[(k-1)!]^2} &= \frac{\sqrt{2\pi}\times(2k-1)^{2k-1}e^{-(2k-1)}\sqrt{2k-1}}{4^{k-1}\times 2\pi\times(k-1)^{2(k-1)}e^{-2(k-1)}(k-1)}\\ &= \frac{\sqrt{2k-1}}{e\sqrt{2\pi}\times 4^{k-1}}\times\left(\frac{2k-1}{k-1}\right)^{2k-1}\\ &= \frac{2\sqrt{2k-1}}{e\sqrt{2\pi}}\times\left(\frac{2k-1}{2k-2}\right)^{2k-1}\\ &= \frac{2\sqrt{2k-1}}{e\sqrt{2\pi}}\times\left(1+\frac{1}{2(k-1)}\right)^{2k-1}.\end{aligned}$$

Note that, for large enough k, the right-most factor in the last equality converges to the natural number e as follows:

$$\lim_{k\to\infty}\left(1+\frac{1}{2(k-1)}\right)^{2k-1} \approx \lim_{k\to\infty}\left(1+\frac{1}{k}\right)^{k} = e.$$

Thus, we have for large k

$$g(k) = \frac{(2k-1)!}{4^{k-1}[(k-1)!]^2} \approx \frac{2e\sqrt{2k-1}}{e\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}(2k-1)}. \tag{21}$$

When it comes to a big enough size of k, if $k \geq 10$, the error rate of the original value of $g(k)$ and its approximation in (21) is less than 1.31%. In Section 4, we use only the first five $g(k)$ for Rényi entropy of order $1/2$ such as $g(1) = 1$, $g(2) = 1.5$, $g(3) = 1.875$, $g(4) = 2.1875$, and $g(5) = 2.4609$. The remainder terms are estimated as $\sqrt{\frac{2}{\pi}(2n-1)} = 0.7979\sqrt{2n-1}$ for $n \geq 6$:

$$\begin{aligned} f(s^N) &= \frac{1}{K}\sum_{n=1}^{K} g(D_n(s^N)) \\ &= \frac{1}{K}\sum_{n=1}^{5} A_n g(n) + \frac{1}{K}\sum_{n=6}^{K} A_n\sqrt{\frac{2}{\pi}}\sqrt{2\times n-1}, \end{aligned}$$

where A_n ($n \geq 1$) is the number of symbol k of random samples with length N. In Section 4, we will see the small block size such as $L = 4$ or $L = 6$, the number of uses of exact values of $g(k)$ should be large to obtain more exact estimation results. However, for $L = 8$, only the first five exact values of $g(k)$ is enough to obtain good results.

3.3. Estimation of Collision Entropy

In this section, we discuss the estimation method of the Rényi entropy of $\alpha = 2$ for L-bit blocks. This case is both one of the widely used Rényi entropy orders and we can very efficiently implement the estimator for this case. We will see that this case is based on a simple counting of consecutive occurrence of the same L-bit random samples. The test value eventually converges to the Rényi entropy of order 2 with increasing sample size.

Assume an ergodic random source S. Then, from (12), the test function for 'collision entropy' is given as

$$f_R(s^K) = -\log_2 \frac{1}{K}\sum_{n=1}^{K} g(D_n(s^K)), \tag{22}$$

where

$$g(k) = \begin{cases} 1, & \text{if } k = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The proposed scheme can be classified as the entropy estimator based on the nearest neighbor distance, which is also used in the Maurer's and Coron's tests.

From the main result, it can be readily proved that the expected value of the proposed test function in (22) is equivalent to the Rényi entropy of order 2 as in the following proposition.

Proposition 1. *The expected value of the test function in (22) is equivalent with the Rényi entropy of order 2 (collision entropy) of L-bit sample from an ergodic random source S.*

Notice that the test function in (22) can be efficiently implemented as in the following description. Let V_{th} be the threshold to accept a given sample as random. The threshold V_{th} can be determined according to applications and a relevant statistical significant level. For implementation, the division and log base 2 in (22) are not essential for the decision and it is enough to only count the number of occurrences such that $D_n(s^K) = 1$. For example, in order to accept the given sample as random, the test function in (22) should be greater than the specified threshold of estimated entropy value as follows:

$$f_R(s^K) = \log_2 K - \log_2 \sum_{n=1}^{K} G_n(s^K) > V_{th}. \tag{24}$$

Then, (24) can be converted into the following relation:

$$\sum_{n=1}^{K} G_n(s^K) < K \cdot 2^{-V_{th}}. \tag{25}$$

Since the right-hand side (RHS) is fixed in (25) when the sample size K is also fixed, it is enough to check the number of times that the specified event on the left-hand side (LHS) is less than the pre-determined value in the RHS. This testing function will be referred to as the *basic test* of the iterative estimation algorithm presented in the following subsection.

Now, let us compare the computational complexity required by the proposed Rényi entropy estimation method with the complexity required by another entropy estimation methods based on the neareast neighbor distance. The required number of operations for three entropy estimation methods based on the nearest neighbor distance is listed in Table 2.

Table 2. Comparison of required number of operations for given N samples.

Method	Required Number of Operations
Maurer [8]	$(N-1)L + (N-1)S + D$
Coron [7]	$(N-1)L + 3(N-1)S + (3N+1)D$
Proposed ($\alpha = 2$)	$L + (N-1)S + D$

L: logarithm, S: summation, and D: division.

As you can seed in Table 2, the proposed method can minimize the number of logarithms and divisions for estimation. Note that the logarithm or division is much more complicated than summation. Therefore, the proposed method has the lowest computational complexity when it is compared with the other nearest neighbor distance based estimations, Maurer's method in NIST STS [20] and Coron's method in AIS.31 [5].

3.4. Iterative Estimation Algorithm for Collision Entropy

For the accurate estimation, the proposed test scheme requires a large sample size, which can be a drawback of the proposed scheme in some applications since it takes much time to collect enough samples for a single estimation. In order to mitigate this drawback, we propose an iterative testing scheme, which will always watch the generated random samples on-the-fly and continuously update the test value with a new counting result for a shorter sample size. The proposed iteration algorithm is presented in Algorithm 1.

In Algorithm 1, N_S is the sample size for the basic test and w $(0 \le w < 1)$ is the weight of the previously accumulated value. Algorithm 1 consists of basic tests, which are continuously carried out when the test is running. The inside statements of for-loop in Algorithm 1 correspond to the basic test that is explained in Section 3.3. Let N_I be the number of iterations. Algorithm 1 can be justified as in the following proposition.

Algorithm 1: Iterative Estimation Algorithm for Rényi Entropy of Order 2 (Collision Entropy)

```
Input  : Random sample $r$
Output: Accumulated value $S$ for Rényi entropy of order 2
begin
    $S := 0$   // Initialization;
    while Test is Running do
        $C := 0$   // Counter;
        $p := 0$   // Previous sample;
        for $i = 1$ to $N_S$ do
            $t := r$   // Get new random sample;
            if $t == p$ then $C++$   // If current sample is the same as the previous, increase counter;
            $p := t$   // Store current random sample;
        $S := w \times S + C$   // Accumulation;
        $R_E := [S/(N_S \times \frac{1}{1-w})]$   // Test value.;
    return $-\log_2 R_E$
```

Proposition 2. *For the stationary random source, after sufficiently large number of iterations, the test value in Algorithm 1 will converge to the Rényi entropy of order 2 (collision entropy).*

Proof. For convenience, let us introduce indices to the counted value C and the accumulated value S in Algorithm 1 such as C_k and S_k where $1 \leq k \leq N_I$. Then, the final accumulated value in Algorithm 1 is given as

$$S_{N_I} = \sum_{k=1}^{N_I} w^{N_I - k} C_k.$$

Suppose that the bias of random sample is stationary. That is, the bias level which can be represented as $\Pr(1)$ in the binary representation is fixed for several consecutive iterations, namely N_I iterations. Then, we can substitute the counted values C_k with the average of them, $\overline{C}$ where

$$\overline{C} = \frac{1}{N_I} \sum_{k=1}^{N_I} C_k. \tag{26}$$

Then, the S_{N_I} can be represented as

$$S_{N_I} = \overline{C} \sum_{k=1}^{N_I} w^{N_I - k} = \overline{C} \times \frac{1 - w^{N_I}}{1 - w}.$$

For a large integer N_I, we have $S_{N_I} \cong \frac{\overline{C}}{1-w}$. That is, $R_E := \left[S/(N_S \times \frac{1}{1-w})\right]$ in Algorithm 1 can be rewritten as

$$R_E = \frac{S_{N_I}}{N_S \times \frac{1}{1-w}} \cong \frac{\overline{C} \times \frac{1}{1-w}}{N_S \times \frac{1}{1-w}} = \frac{\overline{C}}{N_S}.$$

That is, the output R_E of Algorithm 1 is the average of counted values from the basic tests over the sample size N_S. Due to the time average in (26), the proposed algorithm can give us more stable estimated entropy values of the given random samples. □

The weight w will determine a trade-off between converging speed and reducing fluctuation, which will be shown in the next section. If w is close to 1, the estimated value shows less fluctuation at the cost of the sensitivity to the bias changes. In addition, if we choose $w = \frac{2^m - 1}{2^m}$, then the multiplication in the final step corresponds to m-bit left shift. Moreover, $wS = (2^m S - S)/2^m$ in Algorithm 1 can be implemented using m-bit left shift, $\lceil \log_2 S \rceil$-bit subtraction, and m-bit right shift.

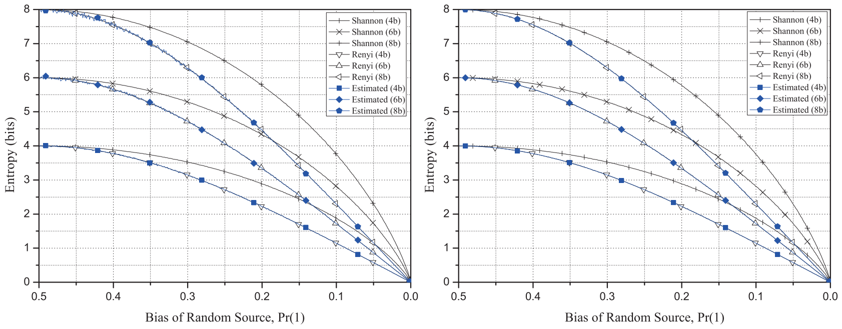
4. Numerical Results

In this section, we present simulation results for the proposed entropy estimator with distinct sample sizes. Simulation results are presented in three ways. First, we present the estimation performance of two sample sizes for Rényi entropy of order 2. Second, we show estimation performance for Rényi entropy of order 1/2. Finally, we present the result of estimating Rényi entropy in an iterative manner.

4.1. Simulation for Rényi Entropy of Order 2

In this simulation, we use $L = 4$, 6, and 8-bit blocks as a single input to the estimator. Therefore, the maximum entropies are also 4, 6 and 8-bit, respectively. For the simulation of the proposed estimator for the Rényi entropy of order 2, we choose two sample sizes; K_1 = 256,000 and K_2 = 10,240,000 which will be called the moderate sample size and the large sample size in the subsequent discussion, respectively. The moderate sample size is the same as the sample size specified in the entropy test for $L = 8$ of AIS.31 for physical random number generators [5]. For the simulation, we generate 500 random sequences with two distinct lengths K_1 and K_2. Each random sequence has a specified bias, which is represented as probability of occurring 1, $\Pr(1)$, in a binary random sequence with a range from 0.001 to 0.5.

Figure 1a shows the simulation results for the Rényi entropy of order 2 with the moderate sample size K_1 = 256,000. Notice that the test results are more accurate for high bias case (close to 0) than the low bias case (close to 0.5) clearly presented in Figure 1a for the sample size K_1. The simulation results for the Rényi entropy of order 2 with the large sample size K_2 = 10,240,000 are depicted in Figure 1b. In this figure, it is easy to see that the Rényi entropy of order 2 is less than or equal to the Shannon entropy as represented in (3). In addition, the test values of the proposed scheme are almost close to the Rényi entropy of order 2 as asserted in Proposition 1.



(**a**) Sample size: K_1 = 256,000.

(**b**) Sample size K_2 = 10,240,000.

Figure 1. Entropy calculation and estimation for 4-, 6-, and 8-bit blocks of two sample sizes. The estimated Rényi entropy of order 2 closely follows the real entropy value for high statistical bias, while it fluctuates for low bias ($\Pr(1) \to 0.5$).

Let us evaluate the amount of fluctuation for a given sample size K. Denote the probability of occurring block b as p_b. The amount of the fluctuation of the test values can be represented using the standard deviation σ_U of the number of occurrences of each block as in the following equation:

$$f_R(s^K) = \log_2 K - \log_2(m_U \pm k\sigma_U) = \log_2 K - \log_2 m_U - \log_2 \left(1 \pm \frac{k\sigma_U}{m_U}\right),$$

where m_U is the mean of the number of occurrences of each block and k is the number of standard deviations as which the test value is allowed to be away from the mean value. The mean and variance can be represented as

$$m_U = \sum_{b \in F_2^L} K\left[\Pr(b)\right]^2,$$

$$\sigma_U = \sum_{b \in F_2^L} \Pr(b)\sqrt{K\Pr(b)(1-\Pr(b))}.$$

Since the amount of the fluctuation is maximized at the no bias case (i.e., $Pr(1) = 0.5$), we can write the test function at that case as follows:

$$f_R(s^K) = \log_2 \frac{1}{p_b} - \log_2\left(1 \pm k\sqrt{\frac{1-p_b}{Kp_b}}\right).$$

Then, the amount of fluctuations at no bias case for K_1 can be evaluated as in the following example.

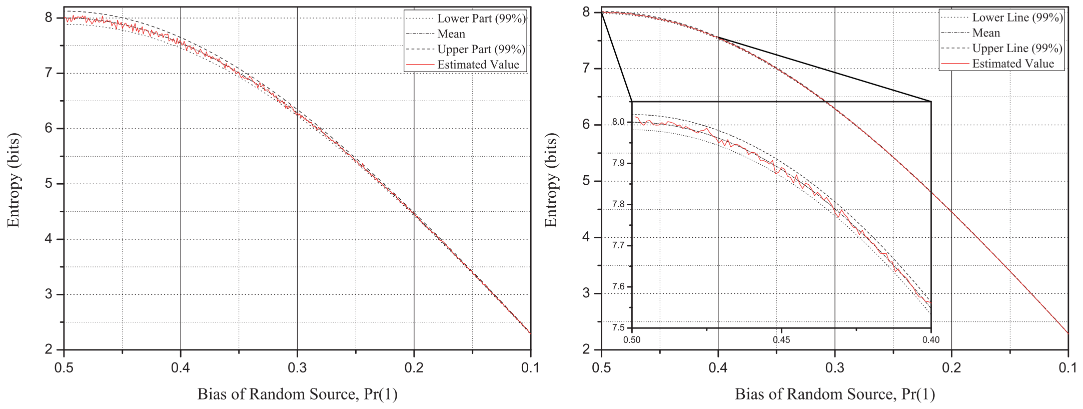
Example 1. *For example, suppose that* $K = 256,000$, $L = 8$, $k = 2.58$ *(for 99% confidence), and the random sample has no bias. Then,* $p_b = 1/256$, $m_U = Kp_b = 1000$, *and* $2.58\sigma_U = 2.58\sqrt{Kp_b(1-p_b)} = 81.42$. *Therefore, we have the test function of the Rényi entropy of order 2 given as*

$$8 - 0.1129 < f_R(s^K) < 8 + 0.1225$$

with 99% confidence.

In Figure 2a, the center line (dashed-dot) is the mean value of test function. The upper (dot) and lower (dashed) lines correspond $2.58\sigma_U$ and $-2.58\sigma_U$ lines, respectively. That is, with 99% confidence, we can say that the test value of the Rényi entropy of order 2 will be between the upper and lower lines. In fact, the estimated value line (solid) is located between the upper and the lower lines in Figure 2a. For the Rényi entropy of order 2 with the large sample size, Figure 2b shows that the three lines are almost merged even in the no bias case. As we can check in the enlarged box on the left side of Figure 2b, the deviation from the real entropy value is small.

Note that, in the randomness test, we are more interested in checking whether the given sample is random or not, rather than in identifying the exact test value. Thus, it is enough that the test value is accurate within the around of the specified threshold. Therefore, we can find a suitable sample size according to the application and accuracy of the test. In particular, when the post-processed TRNGs are available, it is more important to detect a low entropy value because the post-processing method can reduce some statistical bias in the random samples [36].



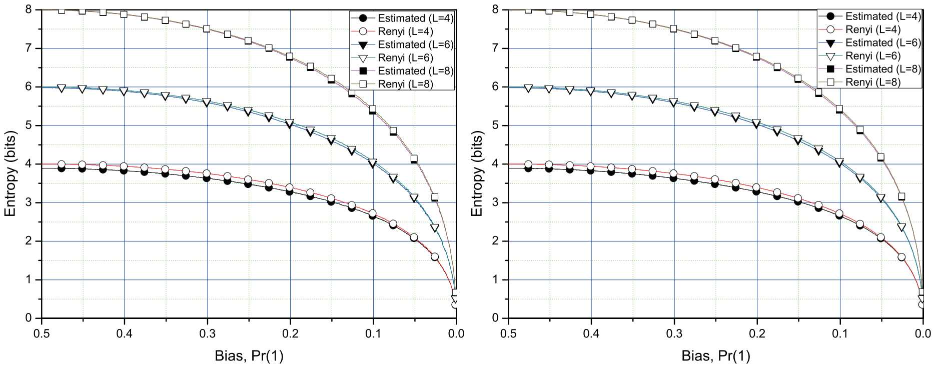
(**a**) Sample size: $K_1 = 256,000$. (**b**) Sample size $K_2 = 10,240,000$.

Figure 2. Deviation range of the test function for the given statistical bias and two sample sizes.

4.2. Simulation for Rényi Entropy of Order 1/2

In this simulation, we use $L = 4$, 6, and 8-bit blocks as a single input to the estimator of Rényi entropy of order $1/2$. For the case of the proposed estimator for the Rényi entropy of order $1/2$, we also choose two sample sizes; K_1 = 256,000 and K_3 = 1,024,000. For the simulation, we generate 500 random sequences with two distinct lengths K_1 and K_2. Each random sequence has a specified bias, which is represented as probability of occurring 1, $\Pr(1)$, in a binary random sequence with range from 0.001 to 0.5.

In Figure 3a,b, the simulation results for the Rényi entropy of order $1/2$ with the moderate sample size K_1 = 256,000 and large sample size K_3 = 1,024,000. Note that, for $L = 8$, the estimated entropy is almost matched with the actual values of the Rényi entropy of order $1/2$ except for the bias range from 0.2 to 0.05. However, for $L = 6$, the deviation of the estimated entropy from the exact Rényi entropy of order $1/2$ is slightly greater than that of the case for $L = 8$. For $L = 4$, there exists the greater deviation between the estimated entropy values and the exact entropy values. This is because we only use the first five exact values of $g(k)$ in (20) in the simulation. If we increase the number of uses of exact values of $g(k)$ instead of approximated values in (21), we can improve the quality of the estimation as in Figure 4. In Figure 4, we compare the results obtained by the uses of the first five exact values and the first 40th exact values, respectively.



(**a**) Sample size: $K_1 = 256,000$. (**b**) Sample size $K_3 = 1,024,000$.

Figure 3. Rènyi entropy of order $1/2$ calculation and estimation for 4-, 6-, and 8-bit blocks of two sample sizes. The first five values of exact $g(k)$ are used. Remainder values are approximated using (21).

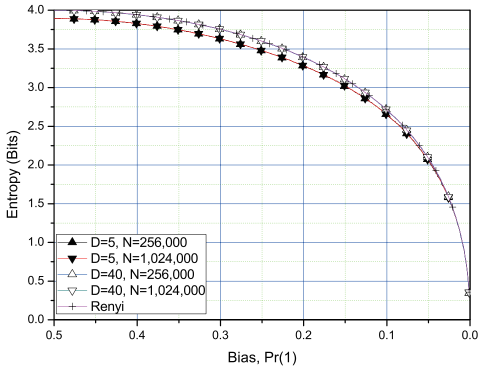
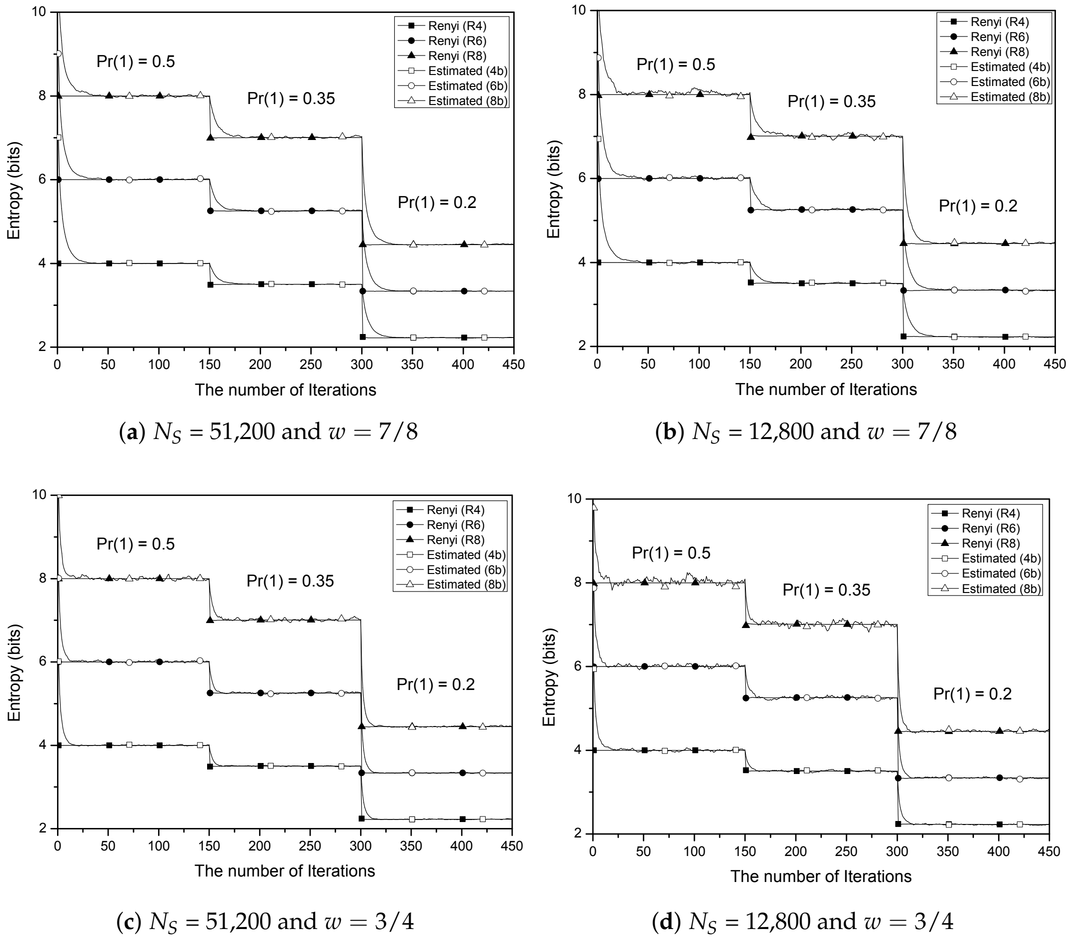


Figure 4. Entropy estimations the Rényi entropy of order $1/2$ with the different number of uses of the exact parameters of $g(k)$ for 4-bit blocks. of the sample size. D is the number of the exact values of parameters of $g(k)$.

4.3. Simulation for Iterative Estimation Scheme of Rényi Entropy of Order 2

Finally, since increasing sample size usually involves tangible cost, collecting of a large number of random samples is not suitable for the constrained devices. In that case, we can apply the iterative estimation scheme presented in Section 3.4 with the smaller size of random sample instead of collecting large random samples for a single entropy estimation. Figure 5 shows the estimation results using Algorithm 1 with weight value 7/8 for the sample sizes 51,200 and 12,800, respectively. In each subfigure, three block sizes such as four, six, and eight bits are tested. In Figure 5, the first 150 iterations show the accumulated test results for the no bias case, i.e., the probability of one in the binary representation of random samples $\Pr(1) = 0.5$. Due to the initialization of Algorithm 1, the first few iterations show relatively big estimated value. For instance, if $N_I = 1$, we have $S = C_1$ and $R_E = \frac{C}{8N_S}$. That is, the estimated value is increased by three bits such that $-\log_2 R_E = 3 - \log_2 \frac{C}{N_S}$. However, after about 25 iterations, the accumulated value converges to the Rényi entropy of order 2. Then, after the first 150 iterations, $\Pr(1)$ is suddenly changed from 0.5 to 0.35. In that situation, the accumulated test value smoothly converges to the new Rényi entropy value of order 2 for $\Pr(1) = 0.35$ within about 25 iterations again. Finally, after the first 300 iterations, the probability of one is abruptly changed again from 0.35 to 0.2. Similar to the previous bias change (0.5 → 0.35), the accumulated test value accordingly converges to new entropy value. Since the number of possible alphabets are exponentially increasing according to the block size L, for a given sample size N_S, the smaller block size (i.e., $L = 4$) shows less fluctuating test results.

Figure 5c,d show the simulation results of Algorithm 1 with weight 3/4. The overall tendency is similar to the results in Figure 5a,b except that the converging speed becomes faster at the cost of higher fluctuations.



(**a**) N_S = 51,200 and $w = 7/8$

(**b**) N_S = 12,800 and $w = 7/8$

(**c**) N_S = 51,200 and $w = 3/4$

(**d**) N_S = 12,800 and $w = 3/4$

Figure 5. Iterative Rényi entropy estimation based on the basic test with two distinct sample sizes and weights. At the 150th and 300th iterations, the statistical bias is changed from $\Pr(1) = 0.5$ to $\Pr(1) = 0.35$, and to $\Pr(1) = 0.2$, respectively.

It is also interesting to check the convergence speed. How quickly the proposed method converges to the actual entropy value when the environment changes depends on the sample size N_S and weight w. When we carefully observe Figure 5, it can be discovered that the larger the sample size and weight, the smaller the fluctuation. However, if the weight is large, the average value is reached at a slower rate. That is, the convergence speed is more related to weight than the sample size. However, it is not trivial to determine when the test value reached the real entropy of the random source.

5. Conclusions

In this paper, we proposed a new estimating method of the Rényi entropy of order α. After presenting the general representation of parameters of the proposed estimator, we investigate the simplified form of three particular orders, such as $\alpha \to 1$, $\alpha = 1/2$, and $\alpha = 2$ in detail. It turned out that the proposed estimator of the Rényi entropy of order 2 which is the widely applicable order can be efficiently implemented by using counting and comparison logics for random samples. The main motivation for this research is to develop a lightweight randomness test method that does not require complex computations to be applicable to systems with limited computational environments such as in the various IoT (Internet of Things) devices. The proposed scheme has a useful and interesting property such that the higher statistical bias in the random sequences, the more accurate detection of that bias for moderate sample size. Because the detection of high bias cases is more critical for the TRNG evaluation, the proposed scheme is acceptable as an on-the-fly entropy estimator with a moderate sample size. However, for the accurate estimation over the wide range of biases, we should test a large amount of random samples. Therefore, we propose an iterative algorithm that continuously carries out the basic tests for the relatively short sample size and updates the accumulated test value. Although it is demonstrated that the proposed method can estimate Rényi entropy of order α, more research on accuracy and convergence speed of the proposed method is also required. We keep this problem as further work.

Funding: This work was supported by the Institute for Information and Communications Technology Promotion (IITP) grant, which is funded by the Korean government (MSIT) (2017-0-00441, Development of Core Technologies of Intrusion Tolerance System for Autonomous Vehicles).

Acknowledgments: The author would like to thank anonymous reviewers and the associate editor for their valuable suggestions and comments that helped to improve the quality of this paper.

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

References

1. NIST. Digital signature standard (DSS). In Proceedings of the Federal Information Processing Standard (FIPS PUB 186), Gaithersburg, MD, USA, 19 May 1994.
2. Bucci, M.; Luzzi, R. Design of testable random bit generators. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 147–156.
3. Dichtl, M.; Golić, J.D. High-speed true random number generation with logic gates only. In *Cryptographic Hardware and Embedded Systems-CHES 2007*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 45–62.
4. Vasyltsov, I.; Hambardzumyan, E.; Kim, Y.-S.; Karpinskyy, B. Fast digital TRNG based on metastable ring oscillator. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 164–180.
5. Killman, W.; Schindler, W. A Proposal for: Funtionality Classes and Evaluation Methodology for True (Physical) Random Number Generators. Available online: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Zertifizierung/Interpretationen/AIS_31_Functionality_classes_evaluation_methodology_for_true_RNG_e.pdf?__blob=publicationFile&v=1 (accessed on 28 August 2018).
6. Schindler, W. Efficient online tests for true random number generators. In *International Workshop on Cryptographic Hardware and Embedded Systems*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 103–117.
7. Coron, J.-S. On the security of random sources. In *International Workshop on Public Key Cryptography*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 29–42.

8. Maurer, U. A universal statistical test for random bit generators. *J. Cryptol.* **1992**, *5*, 89–105. [CrossRef]
9. Wachowiak, M.P.; Smolikova, R.; Tourassi, G.D.; Elmaghraby, A.S. Estimation of generalized entropies with sample spacing. *Pattern Anal. Appl.* **2005**, *8*, 95–101. [CrossRef]
10. Wu, Y.; Yang, P. Optimal entropy estimation on large alphabets via best polynomial approximation. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 824–828.
11. Jiao, J.; Venkat, K.; Han, Y.; Weissman, T. Maximum likelihood estimation of information measures. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 839–843.
12. Acharya, J.; Orlitsky, A.; Suresh, A.T.; Tyagi, H. The complexity of estimating Rényi entropy. In Proceedings of the 26th Annual ACM-SIAM Symposium Discrete Algorithms, San Diego, CA, USA, 4–6 January 2015.
13. Valiant, G.; Valiant, P. Estimating the unseen: An $n/log(n)$-sample estimator for entropy and support size, shown optimal via new CLTs. In Proceedings of the 43rd annual ACM symposium on Theory of Computing, San Jose, CA, USA, 6–8 June 2011; pp. 685–694.
14. Rényi, A. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; pp. 547–561.
15. Erdogmus, D.; Principe, J.C.; Kim, S.-P.; Sanchez, J.C. A recursive Rényi's entropy estimator. In Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing, Martigny, Switzerland, 6 September 2002.
16. Erdogmus, D.; Principe, J.C. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. *IEEE Trans. Signal Process.* **2002**, *50*, 1780–1786. [CrossRef]
17. Xu, J.-W.; Erdogmus, D.M.; Ozturk, C.; Principe, J.C. Recursive Rényi's entropy estimator for adaptive filtering. In Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology, Darmstadt, Germany, 17 December 2003; pp. 134–137.
18. Maynou, J.; Gallardo-Chacón, J.-J.; Vallverdú, M.; Caminal, P.; Perera, A. Computational detection of transcription factor binding sites through differential Rényi entropy. *IEEE Trans. Inf. Theory* **2010**, *56*, 734–741. [CrossRef]
19. Bennett, C.H.; Brassard, G.; Crépeau, C.; Maurer, U. Generalized privacy amplification. *IEEE Trans. Inf. Theory* **1995**, *41*, 1915–1923. [CrossRef]
20. NIST. *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications*; NIST Special Publication 800-22; NIST: Gaithersburg, MD, USA, 2001.
21. Ben-Bassat, M.; Raviv, J. Rényi's entropy and the probability of error. *IEEE Trans. Inf. Theory* **1978**, *24*, 324–331. [CrossRef]
22. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2006.
23. Leonenko, N.; Pronzato, L.; Savani, V. A class of Rényi information estimators for multidemensional densities. *Ann. Stat.* **2008**, *36*, 2153–2182. [CrossRef]
24. Obremski, M.; Skorski, M. Renyi entropy estimation revisited. In *LIPIcs-Leibniz International Proceedings in Informatics*; Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik: Dagstuhl, Germany, 2017.
25. Erdogmus, D.; Hild, K.E.; Principe, J.C.; Lazaro, M.; Santamaria, I. Adaptive blind deconvolution of linear channels using Rényi's entropy with Parzen window estimation. *IEEE Trans. Signal Process.* **2004**, *52*, 1489–1498. [CrossRef]
26. Jizba, P.; Kleinert, H.; Shefaat, M. Renyi's information transfer between financial time series. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 2971–2989. [CrossRef]
27. Grassberger, P. Estimating the information content of symbol sequences andefficient codes. *IEEE Trans. Inf. Theory* **1989**, *35*, 669–675. [CrossRef]
28. Kaltchenko, A.; Timofeeva, N. Entropy estimators with almost sure convergence and an $O(n^{-1})$ variance. *Adv. Math. Commun.* **2008**, *2*, 1–13.
29. Nilsson, M.; Kleijn, W.B. On the estimation of differential entropy from data located on embedded manifolds. *IEEE Trans. Inf. Theory* **2007**, *53*, 2330–2341. [CrossRef]
30. Kamath, S.; Verdu, S. Estimation of entropy rate and Renyi entropy rate for Markov chains. In Proceedings of the 2016 IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–15 July 2016.
31. Kim, Y.-S. Estimation of Rényi entropy of order α based on the nearest neighbor distance. In Proceedings of the 2014 International Symposium on Information Theory and Its Applications, Melbourne, Australia, 26–29 October 2014; pp. 125–129.

32. Coron, J.-S.; Naccache, D. An accurate evaluation of Maurer's universal test. In *International Workshop on Selected Areas in Cryptography*; Springer: Berlin/Heidelberg, Germany, 1999.
33. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics*, 2nd ed.; Addison-Wesley: Boston, MA, USA, 1994; pp. 153–256.
34. Arikan, E. An inequality on guessing and its applcation to sequential decoding. *IEEE Trans. Inf. Theory* **1996**, *42*, 99–105. [CrossRef]
35. Beirami, A.; Calderbank, R.; Duffy, K.; Medard, M. Quantifying computational security subject to source constraints, guesswork and inscrutability. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 2757–2761.
36. Kim, Y.-S.; Jang, J.-W.; Lim, D.-W. Linear corrector overcoming minimum distance limitation for secure TRNG from (17, 9, 5) quadratic residue code. *ETRI J.* **2010**, *32*, 93–101. [CrossRef]