entropy

MDPI

Article

The Eigenvalue Complexity of Sequences in the Real Domain

Lingfeng Liu [1,*], Hongyue Xiang [1], Renzhi Li [1] and Hanping Hu [2]

[1] School of Software, Nanchang University, Nanchang 330031, China; xhyforwhat@163.com (H.X.); lerezz@163.com (R.L.)

[2] School of Automation, Huazhong University of Science & Technique, Wuhan 430074, China; hphu@hust.edu.cn

[*] Correspondence: vatanoilcy@163.com

Received: 5 November 2019; Accepted: 3 December 2019; Published: 5 December 2019

Abstract: The eigenvalue is one of the important cryptographic complexity measures for sequences. However, the eigenvalue can only evaluate sequences with finite symbols—it is not applicable for real number sequences. Recently, chaos-based cryptography has received widespread attention for its perfect dynamical characteristics. However, dynamical complexity does not completely equate to cryptographic complexity. The security of the chaos-based cryptographic algorithm is not fully guaranteed unless it can be proven or measured by cryptographic standards. Therefore, in this paper, we extended the eigenvalue complexity measure from the finite field to the real number field to make it applicable for the complexity measurement of real number sequences. The probability distribution, expectation, and variance of the eigenvalue of real number sequences are discussed both theoretically and experimentally. With the extension of eigenvalue, we can evaluate the cryptographic complexity of real number sequences, which have a great advantage for cryptographic usage, especially for chaos-based cryptography.

Keywords: eigenvalue; real number sequences; complexity

1. Introduction

Sequence complexity can be regarded as a series of measures that depicts the different characteristics of sequences. For cryptographic uses, the most important complexity measures of sequences are linear complexity, Lempel–Ziv (LZ) complexity, eigenvalue, and nonlinear complexity. The nonlinear complexity of a sequence y is an important measure, and it is defined as the length of the shortest Feedback Shift Register (FSR) that generates y. For the shortest Linear Feedback Shift Register (LFSR), it is referred to as the linear complexity. These two measures have been studied for many decades [1–6]. In addition, Lempel and Ziv proposed another well known complexity measure for a given sequence, which is called the LZ complexity [7]. The complexity is related to the number of distinct phrases and the rate of their occurrence along the sequence. In the same study, the eigenvalue was provided from a similar aspect as well, while the eigenvalue profile more closely reflected the rate of vocabulary growth than the LZ complexity. The relationship between LZ complexity and nonlinear complexity was studied in [8], which shows that these two complexity measures are converse in a sense.

For all these complexities, there exists a premise, which is that the measured sequences should be on the finite field. This implies that the cardinality of the state set of the sequences should be finite. For linear complexity and nonlinear complexity, the cardinality of the state set is always set to be two, which corresponds to a binary sequence. The Lempel–Ziv complexity and the eigenvalue can measure the sequences with N symbols, where N is finite. [9] studied the relationship between the eigenvalue and Shannon's entropy of finite symbol sequences. The authors of [10] studied the relationship between the nonlinear complexity and Shannon's entropy of random binary sequences. Moreover, the authors of [11] investigated a method to construct finite length sequences with the large nonlinear complexity on the finite field. In addition, the authors of [12] used the Lempel–Ziv complexity as a nonlinear analysis tool on the characterization of the effects of sleep deprivation on the electroencephalogram, etc. To the best of our knowledge, most of the studies in these complexity measures, including theoretical analyses and practical applications, were subjected to this constraint.

Nowadays, many physical systems can be used in cryptography for its complex dynamical properties, such as chaos-based cryptography [13–20]. However, chaotic systems are based on the real number field $\mathbf{R}^n$, and the cardinality of the state variable is infinite. Currently, we always prove a chaos-based secure cryptographic algorithm based on its high dynamical complexity. However, the dynamical complexity is not completely equal to the cryptographic complexity. Thus, the security of chaos-based cryptographic algorithms is not guaranteed by cryptography researchers [21].

In order to overcome this weakness of chaos-based cryptography, we should evaluate the complexity of chaotic sequences in a cryptographic way. However, all the cryptographic complexity measures currently used are only available for finite symbol sequences. Thus, we should extend the cryptographic complexity from the finite field to the real number field. In this paper, we mainly focused on the eigenvalue complexity, extending this measure from the finite field to the real number field to evaluate the cryptographic properties of real number sequences. The probability distribution, expectation, and variance of the eigenvalue of real number sequences are discussed both theoretically and experimentally.

The rest of this paper is organized as follows. In Section 2, a brief introduction for the eigenvalue is presented and its extension to the real number field is proposed. The eigenvalue of two kinds of real number sequences are discussed in Section 3, including the uniformly distributed random sequence and the logistic chaotic sequence. In Section 4, four kinds of chaotic sequences are evaluated and compared by using the extended eigenvalue measure. Finally, Section 5 concludes the whole paper.

2. Eigenvalue for Real Number Sequences

2.1. Eigenvalue for Binary Sequences

The eigenvalue was first proposed by Lempel and Ziv in [7], which described the number of words occurring from a particular parsing procedure of the sequence. Here, we simply summarized the definition of eigenvalue.

Let $\mathbf{F}_2$ denote the binary field and $x^N = x_0x_1x_2 \ldots x_{N\text{-}1}$ be s binary sequence with length N. x_i^j is denoted as the tuple $x_i \ldots x_j$ in the sequence, $i \leq j$. The prefix and suffix of sequence x^N is defined as x_0^j and $x_j^{N\text{-}1}$, respectively. When $j < N$-1, they refer to proper prefix and proper suffix, respectively. The vocabulary of a sequence x^N is the set consisting of all tuples. If a tuple x_i^j does not belong to the vocabulary of a proper prefix of x^N, it is called an eigenword. The eigenvalue of a sequence equals the total number of eigenwords. The eigenvalue profile of x^N is the integer-valued sequence determined by $k(y^i)$, $i = 1, 2, \ldots, N$.

Proposition 1 ([7]). *The eigenvalue $k(y^N)$ of sequence y^N equals the least l such that y^N is reproducible by y^l.*

Example 1. *Consider a binary sequence x^6 = 010110. The vocabulary of this sequence is {0, 1, 01, 10, 11, 010, 101, 011, 110, 0101, 1011, 0110, 01011, 10110, 010110}. The eigenwords set is {110, 0110, 10110, 010110}. Hence, the eigenvalue of x^6 is 4, and the eigenvalue profile of x^6 is 122244.*

Based on the definition of the eigenvalue that we have, the eigenvalue of x^N equals to k if, and only if, the following two conditions hold.

(1) The tuple $x_{k\text{-}1}x_k \ldots x_{N\text{-}1}$ does not belong to the vocabulary of a proper prefix of y^N.
(2) The tuple $x_k x_{k+1} \ldots x_{N\text{-}1}$ belongs to the vocabulary of a proper prefix of y^N.

Obviously, the definition of the eigenvalue is available for binary sequences, and can be extended to finite symbol sequences at most. However, the chaotic signals are defined on the real number domain, which cannot be measured by this index. Therefore, extending the eigenvalue measure to the real number domain is beneficial to chaos-based cryptography and many other aspects as well.

2.2. Eigenvalue of Sequences in the Real Domain

As we know, the eigenvalue measures the rate of growth of its vocabulary of a sequence. However, strictly speaking, for a real random or chaotic sequence, there will not exist a tuple that occurs more than once. Thus, we cannot judge a tuple based on whether it belongs or does not belong to the vocabulary of a proper prefix.

In the real number field, the Euclidean distance is always used to judge whether two points are close or not. Furthermore, in the dynamics analysis, many measures are based on the Euclidean distance, such as the Lyapunov exponent, Kolmogorov entropy, and embedding dimension. Therefore, in this paper, we used the Euclidean distance d to judge whether the new generated signal was repeated or not. Assume that the sequence $y^N = y_0y_1y_2 \ldots y_{N\text{-}1}$ is a sequence in the real domain, $y_i \in \mathbf{R}$. The state y_j is regarded to be identical with y_i if $|y_j \text{ - } y_i| < d$, where $i < j$, x_i is a state in the prefix $y_0{}^j$. d is defined as the undifferentiated distance that is used to judge whether two real numbers can be regarded as the same. On this basis, we can judge whether the tuple in a sequence belongs to the vocabulary of a proper prefix or not, and whether the eigenvalue can be used in the real number sequences.

Consider a random real number sequence $x^N = x_0x_1x_2 \ldots x_{N\text{-}1}$, where $\mathrm{x_i} \in (a, b)$. Assume that the distribution function of this sequence is $p(x)$. The probability $P(d)$ of the distance of two states x_i and x_j, being larger than d, can be calculated as

$$\begin{array}{c} P(d) = P(\left|x-y\right| > d) = \int_a^{a+d} p(x) \int\limits_{x+d}^{b} p(y)dydx + \int_{b-d}^{b} p(x) \int\limits_{a}^{x-d} p(y)dydx \\ + \int_{a+d}^{b-d} p(x)\left(\int\limits_{a}^{x-d} p(y)dy + \int\limits_{x+d}^{b} p(y)dy \right)dx \end{array} \quad (1)$$

As shown in Equation (1), the probability $P(d)$ can be calculated as the sum of three probabilities. One is the probability of $x_i \in (a, a+d)$ and $x_j \in (x_i+d, b)$; one is the probability of $x_i \in (b\text{-}d, b)$ and $x_j \in (a, x_j\text{-}d)$; and one is the probability of $x_i \in (a+d, b\text{-}d)$ and $x_j \in (a, x_i\text{-}d)$ or (x_i+d, b). Obviously, the probability $P(d)$ is influenced by the undifferentiated distance d. The states x_i and x_j in the sequence are regarded to be identical with the probability 1-$P(d)$. Thus, the probability of tuple $x_ix_{i+1} \ldots x_{N\text{-}1}$ belongs to the vocabulary M of its proper prefix, which can be written as

$$P(x_ix_{i+1}\ldots x_{N-1} \in M) = 1 - \left(1 - (1 - P(d))^{N-i}\right)^i \quad (2)$$

According to conditions (1) and (2), the probability of $k(x^N) = k$ can be written as

$$\begin{aligned} P\left(k(y^N) = k\right) &= P(x_kx_{k+1}\ldots x_{N-1} \in M, x_{k-1}x_k\ldots x_{N-1} \notin M) \\ &= 1 - P(x_kx_{k+1}\ldots x_{N-1} \in M, x_{k-1}x_k\ldots x_{N-1} \in M) \\ &\quad -P(x_kx_{k+1}\ldots x_{N-1} \notin M, x_{k-1}x_k\ldots x_{N-1} \in M) \\ &\quad -P(x_kx_{k+1}\ldots x_{N-1} \notin M, x_{k-1}x_k\ldots x_{N-1} \notin M) \end{aligned} \quad (3)$$

As we know, if the tuple $x_{k\text{-}1}x_k \ldots x_{N\text{-}1} \in M$, the tuple $x_kx_{k+1} \ldots x_{N\text{-}1}$ must belong to M as well. Thus, we have

$$P(x_kx_{k+1}\ldots x_{N-1} \in M, x_{k-1}x_k\ldots x_{N-1} \in M) = P(x_{k-1}x_k\ldots x_{N-1} \in M) \quad (4)$$

$$P(x_k x_{k+1} \ldots x_{N-1} \notin M, x_{k-1} x_k \ldots x_{N-1} \in M) = 0 \tag{5}$$

Furthermore, once the tuple $x_k x_{k+1} \ldots x_{N\text{-}1}$ does not belong to M, the tuple $x_{k\text{-}1} x_k \ldots x_{N\text{-}1}$ will not belong to M either. Thus, we have

$$P(x_k x_{k+1} \ldots x_{N-1} \notin M, x_{k-1} x_k \ldots x_{N-1} \notin M) = P(x_k x_{k+1} \ldots x_{N-1} \notin M) \tag{6}$$

According to Equations (4)–(6), Equation (3) can be simplified as

$$\begin{aligned} P\Big(k(y^N) = k\Big) &= 1 - P(x_{k-1} x_k \ldots x_{N-1} \in M) - 0 - P(x_k x_{k+1} \ldots x_{N-1} \notin M) \\ &= P(x_k x_{k+1} \ldots x_{N-1} \in M) - P(x_{k-1} x_k \ldots x_{N-1} \in M) \end{aligned} \tag{7}$$

When Equation (2) is put into Equation (7), the probability of $k(x^N) = k$ can be written as

$$P\Big(k(y^N) = k, d\Big) = \Big(1 - (1 - P(d))^{N-k+1}\Big)^{k-1} - \Big(1 - (1 - P(d))^{N-k}\Big)^{k} \tag{8}$$

Based on Equation (8), the expectation and variance of the eigenvalue of the real number random sequence x^N can be written as

$$\begin{aligned} E\Big(k(x^N), d\Big) = \sum_{k=1}^{N} k \cdot P\Big(k(x^N) = k\Big) = &\sum_{k=1}^{N} k\bigg(\Big(1 - (1 - P(d))^{N-k+1}\Big)^{k-1} - \Big(1 - (1 - P(d))^{N-k}\Big)^{k}\bigg) \\ &= 1 + \sum_{k=1}^{N-1} \Big(1 - (1 - P(d))^{N-k}\Big)^{k} \end{aligned} \tag{9}$$

and

$$D\Big(k\Big(x^N\Big), d\Big) = \sum_{k=1}^{N} \Big(k - E\Big(k\Big(x^N\Big), d\Big)\Big)^2 \cdot P\Big(k(x^N) = k, d\Big) \tag{10}$$

respectively.

Next, we use the extended eigenvalue to measure the complexity of uniformly distributed random sequences and logistic chaotic sequences.

3. Two Examples

3.1. Eigenvalue of Uniformly Distributed Random Sequence

Consider a uniformly distributed random sequence, whose distributed function is

$$f(x) = \frac{1}{b - a}, \; a < x < b \tag{11}$$

Without loss of generality, we can limit the region from (a, b) into (0, 1). The corresponding distributed function is $f(x) = 1$, $0 < x < 1$. When the distribution function is brought into Equation (1), we have

$$P(d) = P(\big|x - y\big| > d) = d^2 - 2d + 1 \tag{12}$$

Therefore, the probability of $k(x^N) = k$ can be depicted as

$$P\Big(k(y^N) = k, d\Big) = \Big(1 - \Big(2d - d^2\Big)^{N-k+1}\Big)^{k-1} - \Big(1 - \Big(2d - d^2\Big)^{N-k}\Big)^{k} \tag{13}$$

In order to have a more intuitive understanding, the probability distribution of the eigenvalue is depicted in Figure 1 with different undifferentiated distances. The length N is set to be 1000.

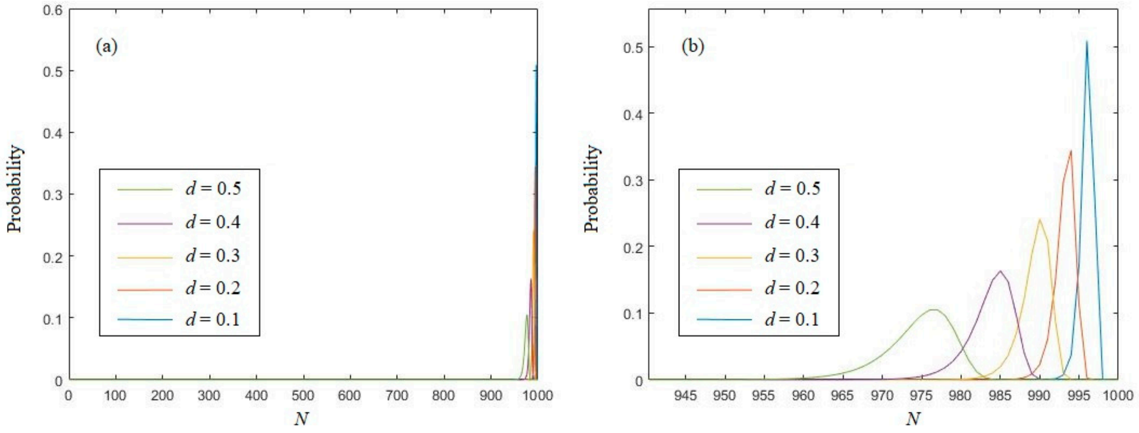

Figure 1. (**a**) The probability distribution of the eigenvalue of random real number sequences; (**b**) the enlargement of (**a**).

In Figure 1, we can see that most of the sequences' eigenvalue are located in a relatively narrow interval. Obviously, with different distances d, the probabilities of $k(x^N) = k$ are different. The peak will move left with the growth of d, and the peak value will be gradually decreased. Thus, to evaluate the eigenvalue of a real number sequence, the choice of distance d is crucial. Consider that [9] has studied the eigenvalue probability of n-symbols' random sequences with uniformly distributed sequences. In order to keep consistency, we should choose $2d - d^2 = 1/n$, and then the distance d should be chosen by

$$d = 1 - \sqrt{1 - 1/n} \tag{14}$$

Therefore, we can compare with the eigenvalue of binary random sequence by choosing $d = 0.2929$, and we can compare with the eigenvalue of 3-symbols random sequence by choosing $d = 0.1835$, and we can compare with the eigenvalue of 4-symbols random sequence by choosing $d = 0.1340$, etc.

Based on the distribution of eigenvalues, the expectation of the eigenvalue for random real number sequences can be approximately calculated as

$$\begin{aligned} E\Big(k(x^N), d\Big) &= \sum_{k=1}^{N} k \cdot P\Big(k(x^N) = k, d\Big) \\ &= \sum_{k=1}^{N} k\Bigg(\Big(1 - \big(2d - d^2\big)^{N-k+1}\Big)^{k-1} - \Big(1 - \big(2d - d^2\big)^{N-k}\Big)^{k}\Bigg) \\ &= 1 + \sum_{k=1}^{N-1}\Big(1 - \big(2d - d^2\big)^{N-k}\Big)^{k} \\ &\approx N - \log_{1/(2d-d^2)} N \end{aligned} \tag{15}$$

for moderate–large N. Set $N = 10{,}000$, a uniformly distributed random sequence is randomly generated. Figure 2 shows the eigenvalue of this sequence. In Figure 2, we can see that all the numerical results are near the theoretical curve we derived in Equation (15), which indicates that the expectation of eigenvalue of random real number sequences is correct.

Based on Equation (10), the variance of the eigenvalue of random real number sequences can be approximately written as

$$D\Big(k\Big(x^N\Big), d\Big) = \sum_{k=1}^{N}\Big(k - N + \log_{1/(2d-d^2)} N\Big)^2 \cdot P\Big(k(x^N) = k, d\Big) \tag{16}$$

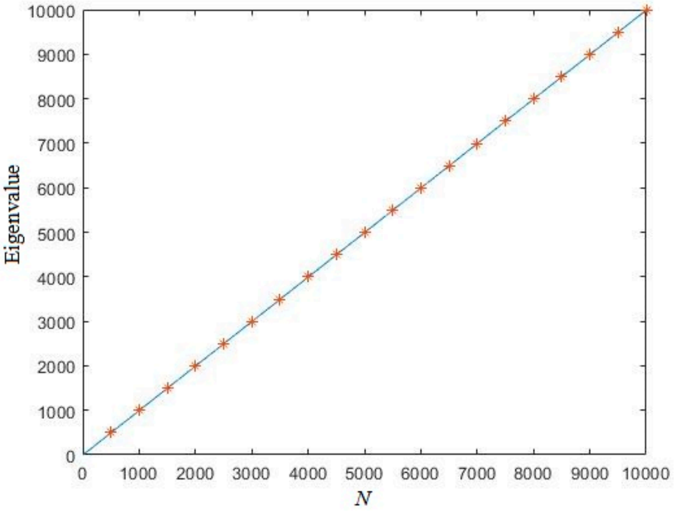

Figure 2. The Eigenvalue of random sequences (the solid line denotes the theoretical curve of the expectation of the eigenvalue for random real number sequences; symbol '*' denotes the eigenvalue of the randomly generated sequence).

Set length N from 1000 to 50,000. Figure 3 shows that for different distances d, the variances of the eigenvalue are all quite stable with the growth of length N.

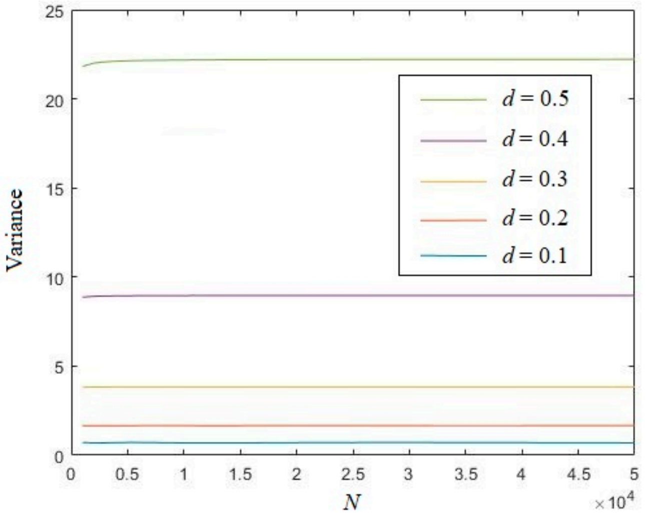

Figure 3. Variances of the eigenvalue of a random real number sequence.

3.2. Eigenvalue of Logistic Chaotic Sequence

Consider the following logistic chaotic map,

$$y_{i+1} = 1 - 2y_i^2 \tag{17}$$

where $y_i \in (-1, 1)$ is the state variable. For an initial condition y_0, we can generate a chaotic sequence $y_0 y_1 \ldots y_{n\,s}$ according to the iteration. The distribution function of Equation (17) is [22]

$$f(y) = \frac{1}{\pi\sqrt{1-y^2}},\ -1 \le y \le 1 \tag{18}$$

According to Equation (1), the probability $P(d)$ of the distance of two states, y_i and y_j, is larger than d and can be calculated as

$$\begin{aligned} P(d) = \int_{-1}^{-1+d} f(y_i) & \int_{x+d}^{1} f(y_j)dy_jdy_i + \int_{1-d}^{1} f(y_i)\int_{-1}^{x-d} f(y_j)dy_jdy_i \\ & +\int_{-1+d}^{1-d} f(y_i)\left(\int_{-1}^{x-d} f(y_j)dy_j + \int_{x+d}^{1} f(y_j)dy_j\right)dy_i \\ & = \frac{1}{4} + \frac{\arcsin(1-d)}{\pi} + \frac{\arcsin^2(1-d)}{\pi^2} \end{aligned} \tag{19}$$

When Equation (19) is brought into Equation (8), the probability of $k(x^N) = k$ for the logistic chaotic sequence can be easily calculated. Figure 4 depicts the probability distribution of the eigenvalue of logistic chaotic sequences with different d values. In Figure 4, we can see that, as with random sequences, the eigenvalue are also located in a relatively narrow interval, and the peak will move left with the growth of d. The peak value will gradually be decreased as well.

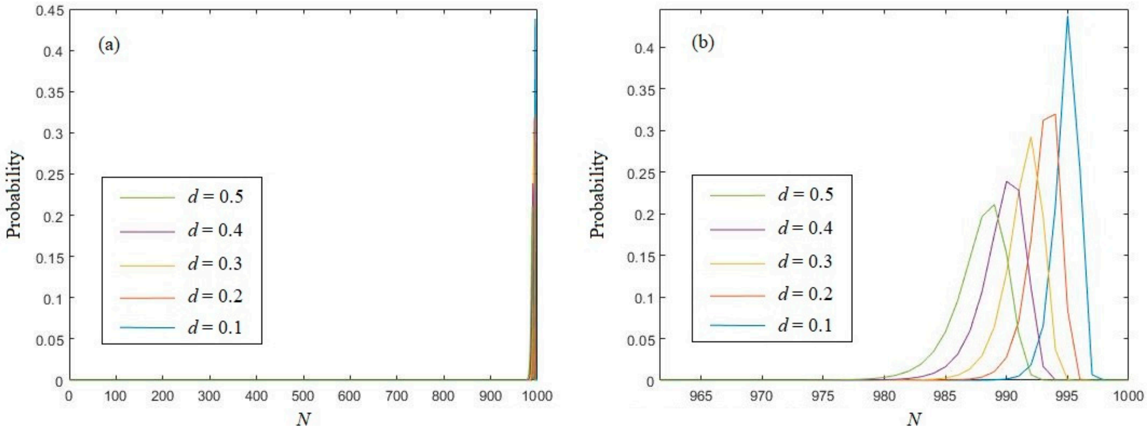

Figure 4. (**a**) The probability distribution of the eigenvalue of logistic chaotic sequences; (**b**) the enlargement of (**a**).

Based on Equation (9), the expectation of the eigenvalue of the logistic chaotic sequence can be written as

$$\begin{aligned} E\left(k(x^N), d\right) &= \sum_{k=1}^{N} k \cdot P\left(k(x^N) = k, d\right) \\ &= \sum_{k=1}^{N} k\left(\begin{array}{l} \left(1-\left(\frac{3}{4} - \frac{\arcsin(1-d)}{\pi} - \frac{\arcsin^2(1-d)}{\pi^2}\right)^{N-k+1}\right)^{k-1} \\ -\left(1-\left(\frac{3}{4} - \frac{\arcsin(1-d)}{\pi} - \frac{\arcsin^2(1-d)}{\pi^2}\right)^{N-k}\right)^{k} \end{array}\right) \\ &= 1 + \sum_{k=1}^{N-1}\left(1-\left(\frac{3}{4} - \frac{\arcsin(1-d)}{\pi} - \frac{\arcsin^2(1-d)}{\pi^2}\right)^{N-k}\right)^{k} \\ &\approx N - \log_{4\pi^2/(3\pi^2-4\pi\arcsin(1-d)-4\arcsin^2(1-d))} N \end{aligned} \tag{20}$$

for moderate–large N. When we randomly select an initial condition, Figure 5 shows the eigenvalue of this generated logistic sequence. Obviously, the eigenvalues of this sequence are all around the theoretical curve we derived in Equation (20).

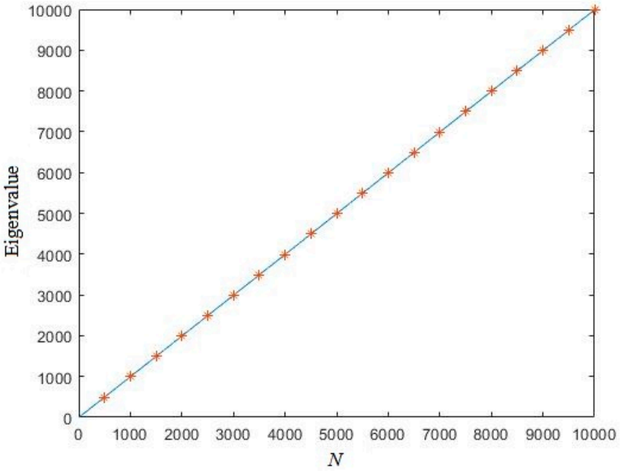

Figure 5. The eigenvalue of the logistic sequence (the solid line denotes the theoretical curve of expectation of eigenvalue for random real number sequences; symbol '*' denotes the eigenvalue of the generated logistic sequence).

Figure 6 depicts the comparison of the expectation of the eigenvalue of logistic sequences and uniformly distributed random sequences under the same undifferentiated distance $d = 0.1$. In Figure 6, we can see that there are almost no differences among the expectation eigenvalue of the logistic sequences and random sequences. After enlarging, we can see that the eigenvalue of logistic sequences is just a little lower than the random sequence, which implies that the logistic sequence cannot be regarded as a perfect random sequence in this sense. For other undifferentiated distances, the results are similar. Therefore, we omit them here to avoid redundancy.

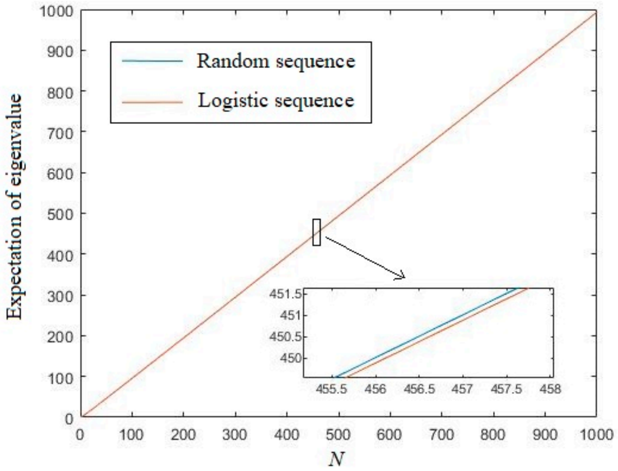

Figure 6. The expectation of the eigenvalue of logistic sequences and random sequences.

Correspondingly, the variance of the eigenvalue of logistic sequences can be approximately written as

$$D\Big(k\Big(x^N\Big),d\Big) = \sum_{k=1}^{N}\Big(k - N + \log_{4\pi^2/(3\pi^2-4\pi\arcsin(1-d)-4\arcsin^2(1-d))} N\Big)^2 \cdot P\Big(k\big(x^N\big) = k, d\Big) \tag{21}$$

The variances of the eigenvalue of logistic sequences with different distances d are depicted in Figure 7. This figure indicates that for every distance, the eigenvalue of logistic sequences are all stable with the growth of length N.

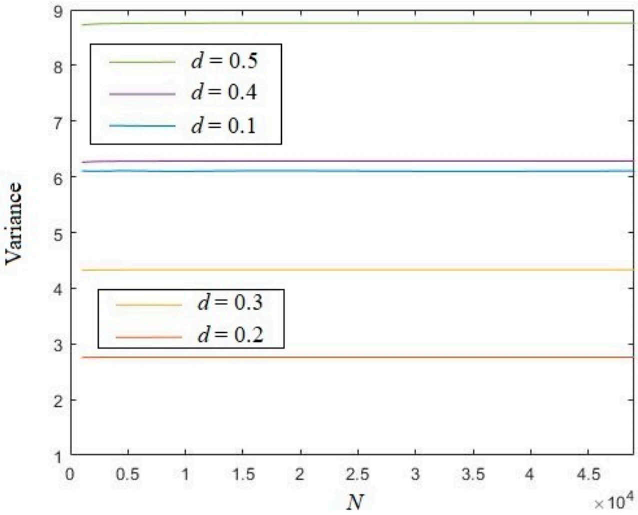

Figure 7. The variance of the eigenvalue of logistic sequences.

4. Measure the Complexity of Chaotic Sequences

With the extension of eigenvalue, we can use this complexity measure to evaluate the cryptographic characteristics of different chaotic sequences. Here, the following four kinds of 1-D chaotic sequences are generated and compared.

A. Chebyshev map

Chebyshev map can be written as

$$x_{i+1} = \cos(a \cdot \arccos(x_i)) \tag{22}$$

where $x_i \in (-1, 1)$ is the state variable, a is the control coefficient. The Chebyshev map will be chaotic since $a \geqq 2$. In this test, we always set $a = 3$.

B. Sine map

Sine map can be mathematically described as

$$x_{i+1} = r \sin(\pi x_i) \tag{23}$$

where $r \in (0, 1]$ is the control parameter. In this test, we set $r = 2$ to make the Sine map chaotic.

C. Tent map

Tent map is a kind of piece-wise function, which can be described as

$$x_{i+1} = \begin{cases} x_i/p, & x_i \in [0, p) \\ (1 - x_i)/(1 - p), & x_i \in [p, 1] \end{cases} \tag{24}$$

where $p \in (0, 1)$ is the control parameter. Particularly, when $p = 0.5$, the generated sequence will quickly fall into a short cycle. Therefore, we always set $p = 0.49$ in this test.

D. Logistic map

The Logistic map has already been described in Equation (17), which we omitted here to avoid redundancy.

Since the state variables of these four maps are in different domains, for consistency, we first compressed them to the identical interval (0, 1). When $d = 0.1$, the eigenvalue of these four kinds of chaotic sequences are depicted in Figure 8. Figure 8 shows that the chaotic sequences generated by the sine map have the largest eigenvalue, whereas the chaotic sequences generated by the Tent map have the lowest eigenvalue. For other distances d, the results are similar.

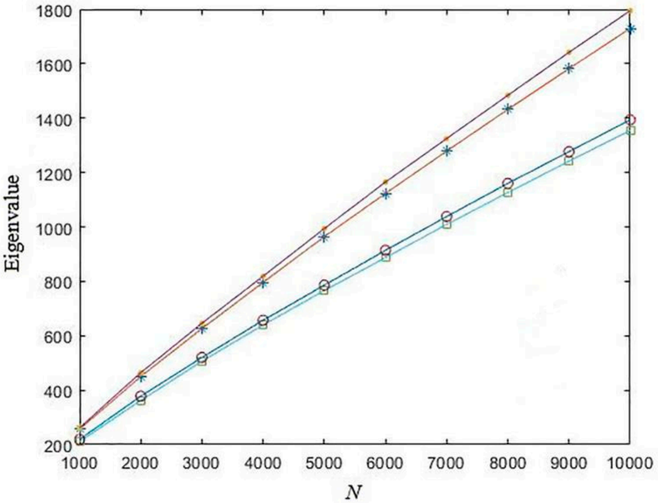

Figure 8. The eigenvalue of different chaotic sequences. (The asterisk represents the eigenvalue of the Chebyshev chaotic sequence; the solid point represents the eigenvalue of the sine chaotic sequence; the square represents the eigenvalue of the Tent chaotic sequence; the hollow point represents the eigenvalue of the logistic chaotic sequence).

Thus, it can be seen that with the extended eigenvalue, we can evaluate the cryptographic complexity of real number sequences effectively. However, it should be noted that this result does not imply that the Sine map is better than other chaotic maps in cryptographic application. On the one hand, the eigenvalue is only one of the cryptographic complexity measures; on the other hand, the eigenvalue value is influenced by the control parameter of chaotic maps. For example, the eigenvalue of the Sine chaotic sequence will be lower than the eigenvalue of the Chebyshev chaotic sequence when $r = 1$.

5. Conclusions

In order to evaluate the cryptographic complexity of real number sequences, in this paper, we extended the so-called eigenvalue from the binary field to the real number field. The extended eigenvalue was influenced by the undifferentiated distance, and we gave an exact value of this distance corresponding to the N-symbol sequences. Both uniformly distributed random sequences and logistic sequences were used as examples. The probability distribution, expectation and variance of these two kinds of real number sequences were discussed both theoretically and experimentally. With the extension of eigenvalue, we could evaluate the cryptographic complexity of real number sequences, which has a great advantage for cryptographic usage, especially for chaos-based cryptography. Furthermore, four kinds of chaotic sequences were evaluated by this extended complexity measure, which indicates that our study is effective and of great interest.

Author Contributions: Conceptualization, L.L.; Data curation, R.L.; Formal analysis, L.L. and H.X.; Funding acquisition, H.H., L.L.; Investigation, H.H.; Methodology, L.L.; Resources, L.L.; Supervision, L.L.

Funding: This research was funded by National Key R&D Program of China: 2017YFB0802000; National Natural Science Foundation of China: 61862042; 61601215; Postgraduate Innovation Project of Jiangxi Province: YC2019S101; Cryptography Theoretical Research of National Cryptography Development Fund: MMJJ20170109.

Acknowledgments: This work is supported by the National Natural Science Foundation of China (61862042, 61601215); Postgraduate Innovation Project of Jiangxi Province (YC2019S101); Cryptography Theoretical Research of National Cryptography Development Fund (MMJJ20170109).

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

References

1. Massey, J.L.; Serconek, S. A Fourier transform approach to the linear complexity of nonlinearly filtered sequences. In *Annual International Cryptology Conference*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 332–340.
2. Kolokotronis, N.; Kalouptsidis, N. On the linear complexity of nonlinearly filtered PN-sequences. *IEEE Trans. Inf. Theory* **2003**, *49*, 3047–3059. [CrossRef]
3. Limniotis, K.; Kolokotronis, N.; Kalouptsidis, N. New results on the linear complexity of binary sequences. In *2006 IEEE International Symposium on Information Theory*; IEEE: New York, NY, USA, 2006; pp. 2003–2007.
4. Erdmann, D.; Murphy, S. An approximate distribution for the maximum order complexity. *Des. Codes. Cryptogr.* **1997**, *10*, 325–339. [CrossRef]
5. Rizomiliotis, P. Constructing periodic binary sequences of maximum nonlinear span. *IEEE Trans. Inf. Theory* **2006**, *52*, 4257–4261. [CrossRef]
6. Rizomiliotis, P.; Kolokotronis, N.; Kalouptsidis, N. On the quadratic span of binary sequences. *IEEE Trans. Inf. Theory* **2005**, *51*, 1840–1848. [CrossRef]
7. Lempel, A.; Ziv, J. On the complexity of finite sequences. *IEEE Trans. Inf. Theory* **1976**, *22*, 75–81. [CrossRef]
8. Limniotis, K.; Kolokotronis, N.; Kalouptsidis, N. On the nonlinear complexity and Lempel-Ziv complexity of finite length sequences. *IEEE Trans. Inf. Theory* **2007**, *53*, 4293–4302. [CrossRef]
9. Liu, L.; Miao, S.; Hu, H.; Deng, Y. On the Eigenvlaue and Shannon's Entropy of Finite Length Random Sequences. *Complexity* **2015**, *21*, 154–161. [CrossRef]
10. Liu, L.; Miao, S.; Liu, B. On nonlinear complexity and Shannon's entropy of finite length random sequences. *Entropy* **2015**, *17*, 1936–1945. [CrossRef]
11. Tosun, P.; Abásolo, D.; Stenson, G.; Winsky-Sommerer, R. Characterisation of the effects of sleep deprivation on the electroencephalogram using permutation Lempel-Ziv complexity, a non-linear analysis tool. *Entropy* **2017**, *19*, 673. [CrossRef]
12. Peng, J.; Zeng, X.; Sun, Z. Finite length sequences with large nonlinear complexity. *Adv. Math. Commun.* **2018**, *12*, 215–230. [CrossRef]
13. Li, P.; Wang, Y.C.; Wang, A.B.; Wang, B.J. Fast and tunable all-optical physical random number generator based on direct quantization of chaotic self-pulsations in two-section. *IEEE J. Sel. Top. Quantum Electron.* **2013**, *19*, 0600208.
14. Li, R.; Liu, Q.; Liu, L. Novel image encryption algorithm based on improved logistic map. *IET Image Process.* **2019**, *13*, 125–134. [CrossRef]
15. Huang, X.; Liu, L.; Li, X.; Yu, M.; Wu, Z. A new pseudorandom bit generator based on mixing three-dimensional Chen chaotic system with a chaotic tactics. *Complexity* **2019**, *2019*, 6567198. [CrossRef]
16. Kanso, A.; Smaoui, N. Logistic chaotic maps for binary numbers generations. *Chaos Solitons Fract.* **2009**, *40*, 2557–2568. [CrossRef]
17. Larger, L.; Dudley, J.M. Optoelectronic chaos. *Nature* **2010**, *465*, 41–42. [CrossRef] [PubMed]
18. Bahi, J.M.; Fang, X.; Guyeux, C.; Wang, Q. On the design of a family of CI pseudo-random number generators. In Proceedings of the 2011 7th International Conference on Wireless Communications, Networking and Mobile Computing, Wuhan, China, 23–25 September 2011; pp. 1–4.
19. Masuda, N.; Aihara, K. Cryptosystems with discretized chaotic maps. *IEEE Trans Circuits Syst. I* **2002**, *49*, 28–40. [CrossRef]
20. Li, P.; Wang, Y.C.; Wang, A.B.; Yang, L.Z.; Zhang, M.J.; Zhang, J.Z. Direct generation of all-optical random numbers from optical pulse amplitude chaos. *Opt. Express* **2012**, *20*, 4297–4308. [CrossRef] [PubMed]
21. Kocarev, L. Chaos-based cryptography: A brief overview. *IEEE Circ. Syst. Mag.* **2001**, *1*, 6–21. [CrossRef]
22. Liu, L.; Miao, S.; Hu, H.; Cheng, M. N-phase Logistic chaotic sequence and its application for image encryption. *IET Signal Process.* **2016**, *10*, 1096–1104. [CrossRef]