Linear Complexity versus Pseudorandomness: On Beth and Dai's Result

Yongge Wang*

Center for Applied Cryptographic Research, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada,
ygwang@cacr.math.uwaterloo.ca

Abstract. Beth and Dai studied in their Eurocrypt paper [1] the relationship between linear complexity (that is, the length of the shortest Linear Feedback Shift Register that generates the given strings) of strings and the Kolmogorov complexity of strings. Though their results are correct, some of their proofs are incorrect. In this note, we demonstrate with a counterexample the reason why their proofs are incorrect and we prove a stronger result. We conclude our note with some comments on the use of the LIL test (the law of the iterated logarithm) for pseudorandom bits generated by pseudorandom generators.

1 Introduction

Feedback shift registers, in particular linear feedback shift registers, are the basic components of many keystream generators. And the linear complexity of strings is an important tool to study different kinds of linear or nonlinear feedback registers.

Kolmogorov [6], and Chaitin [2] introduced the notion of Kolmogorov-Chaitin complexity of strings, which measures the minimum size of the input to a fixed universal Turing machine to generate the given string. While this complexity measure is of theoretical interest, there is no algorithm for computing it (it is equivalent to the problem of deciding whether a Turing machine halts on a given input, whence it is unsolvable).

Beth and Dai studied in their Eurocrypt paper [1] the relationship between linear complexity (that is, the length of the shortest Linear Feedback Shift Register that generates the given string) of strings and the Kolmogorov complexity of strings. Though their results are correct, some of their proofs are incorrect. In this note, we demonstrate the reason why their proofs are incorrect and we prove a stronger result.

Many stream ciphers utilize deterministically generated "random" sequences to encipher the message stream. Since the security of the system is based on the "randomness" of the key sequences, the criterion to determine the degree of "randomness" is crucial. In cryptographic community, the "randomness postulate"

* Most of this work was done when the author was a post-docs at the University of Wisconsin–Wilwaukee.

by Golomb [5] and the "universal test" by Maurer [10] have gained widespread popularity. In Rueppel [13], [14, Ch.4], and Niederreiter [12], linear complexity profiles are proposed as a test for randomness. However, these tests cannot detect certain weaknesses in "pseudorandom sequences". For example, it should not be the case that the number of 1's in the initial segments of a random sequence is always larger (or equal) than the number of 0's. And it is easy to see that none of the above mentioned popular "tests" can detect this "weakness" in a sequence. At the end of this note, we suggest the use of the LIL test (the law of the iterated logarithm) as an additional test for pseudorandom sequences. The LIL test can detect several weaknesses in a pseudorandom sequence including the one mentioned above.

We close this section with some notation we will use. $\{0,1\}^*$, $\{0,1\}^n$, and $\{0,1\}^\infty$ are the set of finite binary strings, the set of binary strings of length n, and the set of infinite binary sequences respectively. The length of a string s is denoted by $|s|$. For a sequence $s \in \{0,1\}^* \cup \{0,1\}^\infty$ and an integer number $n \geq 1$, $s[1..n]$ denotes the initial segment of length n of s ($s[1..n] = s$ if $|s| \leq n$) while $s[n]$ denotes the nth bit of s, i.e., $s[1..n] = s[1] \dots s[n]$.

For a set $\mathbf{C} \subseteq \{0,1\}^\infty$ of infinite sequences, $Prob[\mathbf{C}]$ denotes the probability that $s \in \mathbf{C}$ when s is chosen by a random experiment in which an independent toss of a fair coin is used to decide the value of each bit in s. This probability is defined whenever $\mathbf{C}$ is measurable under the usual product measure on $\{0,1\}^\infty$ (which can also be considered as the unit interval on the real number line).

2 Definitions and basic results

2.1 Linear feedback shift registers

Linear feedback shift registers (LFSR) are the basic components of many keystream generators (see, e.g., Menezes et al. [11]). There are several reasons for this: LFSRs are well-suited to hardware implementation; they can produce sequences of large period; and they can produce sequences of good statistical properties.

A *linear feedback shift register* (LFSR) of length l is a sequence of 0-1 bits $(s_1, \dots, s_l, c_1, \dots, c_l)$ with $c_1 = 1$. The *output* of the LFSR is the infinite sequence $s_1 s_2 s_3 \dots$ where s_i for $i > l$ is defined by the following equation:

$$s_i = \sum_{j=1}^{l} c_j s_{i-l-1+j} \mod 2.$$

An LFSR $L(s_1, \dots, s_l, c_1, \dots, c_l)$ is said to *generate* an infinite sequence $s = s_1 s_2 \dots$ if s is the output sequence of $L(s_1, \dots, s_l, c_1, \dots, c_l)$. The *linear complexity* of an infinite sequence s, denoted $L(s)$, is defined as follows:

1. If s is the zero sequence $000\dots$, then $L(s) = 0$.
2. If no LFSR generates s then $L(s) = \infty$.
3. Otherwise $L(s)$ is the length of the shortest LFSR that generates s.

For a finite string $s \in \{0,1\}^n$, the *linear complexity* $L(s)$ of s is defined as the length of the shortest LFSR that generates a sequence having s as its first n bits.

Theorem 1. *(see, e.g., [9, 11])*

1. *For $s \in \{0,1\}^n$, $0 \le L(s) \le n$.*
2. *For $s \in \{0,1\}^n$, $L(s) = 0$ if and only if $s = 0 \dots 0$.*
3. *For $s \in \{0,1\}^n$, $L(s) = n$ if and only if $s = 0 \dots 01$.*
4. *If s is periodic with period n, then $L(s) \le n$.*

Theorem 2. *For any strings $s_1, s_2, s_3 \in \{0,1\}^*$, $L(s_1 s_2 s_3) \ge L(s_2)$.*

Proof. This is straightforward from the definitions. □

Theorem 3. *(Massey [9]) For any given string $s \in \{0,1\}^n$, the Berlekamp-Massey algorithm will compute $L(s)$ in $O(n^2)$ bit operations.*

Theorem 4. *For any $s \in \{0,1\}^n$, either $L(s0) \ge n/2$ or $L(s1) \ge n/2$.*

Proof. In the proof of Theorem 3 (see [9, 11]), it has been shown that if the shortest LFSR for generating s is $L(s_1, \dots s_l, c_1, \dots, c_l)$, then there are two cases:

1. $L(s0) = l$. Then
$$L(s1) = \begin{cases} l & \text{if } l > n/2, \\ n+1-l & \text{otherwise.} \end{cases}$$
2. $L(s1) = l$. Then
$$L(s0) = \begin{cases} l & \text{if } l > n/2, \\ n+1-l & \text{otherwise.} \end{cases}$$

This completes the proof of the theorem. □

Theorem 5. *(Rueppel [14]) Let $k \le n$ and $N_n(k) = |\{s \in \{0,1\}^n : L(s) = k\}|$. Then*
$$N_n(k) = \begin{cases} 2^{\min(2n-2k, 2k-1)} & \text{if } n \ge k > 0, \\ 1 & \text{if } k = 0. \end{cases}$$

2.2 Kolmogorov complexity

Kolmogorov complexity, as developed by Chaitin [2] and Kolmogorov [6] gives a satisfactory theoretical description of the complexity of individual finite strings and infinite sequences. In this section, we review the fundamentals of Kolmogorov complexity theory that we will use in this paper. For more details, it is referred to Li and Vitanyi [7]. Let U be a universal Turing machine. Then the Kolmogorov complexity of a string $s \in \{0,1\}^n$ is defined by

$$K(s) = \min\{|x| : U(x) = s, x \in \{0,1\}^*\}.$$

We are also interested in the self-delimiting Turing machines. A *self-delimiting* Turing machine is a deterministic Turing machine M such that the program set

$$PROG_M = \{x \in \{0,1\}^* : M \text{ halts on input } x \text{ after finitely many steps}\}$$

is prefix-free, i.e., a set of strings with the property that no string in it is a proper prefix of another. Let U_c be a universal self-delimiting Turing machine, then the Chaitin-Kolmogorov complexity of a string $s \in \{0,1\}^n$ is defined by

$$H(s) = \min\{|x| : U_c(x) = s, x \in \{0,1\}^*\}.$$

Theorem 6. *(see, e.g., [2, 6, 7])*

1. *There is a constant $c > 0$ such that $K(s) \leq H(s) + c$ for all $s \in \{0,1\}^*$.*
2. *There is a constant $c > 0$ such that $H(s) \leq K(s) + 2\log|s| + c$ for all $s \in \{0,1\}^*$.*
3. *For each Turing machine M and each string $s \in \{0,1\}^*$, let $K_M(s) = \min\{|x| : M(x) = s, x \in \{0,1\}^*\}$. Then there is a constant $c_M > 0$ such that $K(s) \leq K_M(s) + c_M$ for all $s \in \{0,1\}^*$.*
4. *There is a constant $c > 0$ such that $K(s) \leq |s| + c$ for all $s \in \{0,1\}^*$.*
5. *For any constant $c > 0$, $|\{s \in \{0,1\}^n : K(s) \geq n - c\}| \geq 2^{n-c-1}$.*

An infinite sequence $s \in \{0,1\}^\infty$ is Martin-Löf *random* (see, e.g., [7, 8]) if and only if there is a constant $c > 0$ such that $H(s[1..n]) \geq n - c$ for almost all n.

Lemma 1. *Let $s \in \{0,1\}^\infty$ be a Martin-Löf random sequence. Then there is a constant $c > 0$ such that $n - 2\log n - c \leq K(s[1..n]) \leq n + c$ for almost all n.*

Proof. It follows from Theorem 6 and the definition of a Martin-Löf random sequence. □

Theorem 7. *(see, e.g., [7]) $Prob\{s : s \text{ is Martin-Löf random}\} = 1$.*

3 Linear complexity versus Kolmogorov complexity

Beth and Dai [1] proved the following theorem on the relationship between linear complexity and Kolmogorov complexity.

Theorem 8. *(Beth and Dai [1]) For all $\varepsilon (0 < \varepsilon < 1)$*

$$P_{\varepsilon,n} = Prob\{s \in \{0,1\}^n : (1-\varepsilon) \cdot 2L(s) \leq K(s) \leq (1+\varepsilon) \cdot 2L(s)\} \to 1$$

when $n \to \infty$.

After proving the above result, Beth and Dai [1] "proved" the following result:

Theorem 9. *([1]) With probability 1, a sequence $s \in \{0,1\}^\infty$ satisfies the following property:*

$$\lim_{n\to\infty} \frac{K(s[1..n])}{L(s[1..n])} = 2. \tag{1}$$

Their proof is as follows:

"Proof". Apply the Borel-Cantelli lemma to the independent cylinder sets

$$A_{k,\varepsilon} = \{s \in \{0,1\}^\infty : (1-\varepsilon) \cdot 2L(s[2^{k-1}..2^k-1]) \leq$$

$$K(s[2^{k-1}..2^k-1]) \leq (1+\varepsilon) \cdot 2L(s[2^{k-1}..2^k-1])\}$$

for the positive integers k and $\varepsilon > 0$. From Theorem 8, we conclude that

$$\sum_{k=1}^{\infty} Prob[A_{k,\varepsilon}] = \infty$$

for the positive integers k and $\varepsilon > 0$. Thus the assertion. □

In the following we will show that the above "proof" is incorrect. Note that the second Borel-Cantelli lemma states as follows.

Theorem 10. *(The second Borel-Cantelli Lemma [4]) Let* $\mathbf{C}_1, \mathbf{C}_2, \ldots \subseteq \{0,1\}^\infty$ *be a sequence of independent, Lebesgue measurable sets, i.e.,*

$$Prob[\mathbf{C}_i] \cdot Prob[\mathbf{C}_j] = Prob[\mathbf{C}_i \cap \mathbf{C}_i]$$

for $i \neq j$*, such that* $\sum_{k=1}^{\infty} Prob[\mathbf{C}_k]$ *diverges. Then*

$$\mathbf{C} = \{s \in \{0,1\}^\infty : s \in \mathbf{C}_k \textit{ for only infinitely many } k\}$$

has probability 1*.*

By the Borel-Cantelli Lemma and the assertion $\sum_{k=1}^{\infty} Prob[A_{k,\varepsilon}] = \infty$, we can only infer in the above "proof" of Beth and Dai that

$$\mathbf{A} = \{s \in \{0,1\}^\infty : s \in A_{k,\varepsilon} \text{ for infinitely many } k\}$$

has probability 1. However, from this result we cannot infer that the equation (1) holds for all $s \in \mathbf{A}$, since there is an infinite sequence s such that $s \in A_{k,\varepsilon}$ for infinitely many k, but the equation (1) does not hold for s. Whence, Beth and Dai's "proof" is incorrect.

Indeed, we can construct an infinite sequence s such that $s \in A_{k,\varepsilon}$ for almost all k, but the equation (1) does not hold for s. In order to construct such a sequence s, we first prove two preliminary results.

Lemma 2. *Let* k *and* n *be two positive integers. Then for any string* $s \in \{0,1\}^{kn}$ *such that* $L(s) = n$*,* $K(s) \leq 2n + 2\log k + c$ *for some constant* $c > 0$*.*

Proof. Define a Turing machine M_L as follows:

$$M_L(x) = \begin{cases} t[1..il] & \text{if } x = (i, s_1 \ldots s_l c_1 \ldots c_l) \text{ and the LFSR} \\ & L(s_1, \ldots, s_l, c_1, \ldots, c_l) \text{ generates the infinite sequence } t, \\ \text{undefined} & \text{otherwise.} \end{cases}$$

By Theorem 6, there is a constant $c > 0$ such that for all $s \in \{0,1\}^*$

$$K(s) \leq \min\{|x| : M_L(x) = s\} + c.$$

For any string $s \in \{0,1\}^{kn}$ such that $L(s) = n$, let $L_s(s_1, \ldots, s_n, c_1, \ldots, c_n)$ be the shortest LFSR which generates a sequence having s as its first kn bits. By the definition of M_L, we have $M_L(k, s_1 \ldots s_n c_1 \ldots c_n) = s$. That is,

$$K(s) \leq |(k, s_1 \ldots s_n c_1 \ldots c_n)| + c \leq 2n + 2\log k + c.$$

This completes the proof. □

Theorem 11. *There is a constant $k_0 > 0$ such that for each $k > k_0$, there is a string $s_k \in \{0,1\}^{2^{k-1}}$ with the following properties:*

1. *$2L(s_k) - c \leq K(s_k) \leq 2L(s_k) + c$ for some constant $c > 0$.*
2. *$L(s_k s_{k+1}[1]) \geq 2^{k-2}$.*

Proof. We define the strings s_k by induction on k. Let k_0 be a large enough constant, $|s_{k_0}| = 2^{k_0-2}$ be a string with $L(s_{k_0}) = 2^{k_0-3}$, and $k > k_0$. Assume that s_{k-1} has already been defined. Then, by Theorems 4, 5, and 6, there is a string $s'_k \in \{0,1\}^{2^{k-3}}$ such that

- $2L(s'_k) = 2^{k-3}$;
- $K(s'_k) \geq 2^{k-3} - c_0$ for some constant c_0;
- $L(s_{k-1} s'_k[1]) \geq 2^{k-3}$.

Let $s_k \in \{0,1\}^{2^{k-1}}$ be any string with the properties that $s'_k = s_k[1..2^{k-3}]$ and $L(s_k) = L(s'_k)$. By Lemma 2,

$$K(s_k) \leq 2L(s_k) + 2\log 4 + c_1$$

for some constant $c_1 > 0$. It is straightforward to check that

$$K(s_k) \geq K(s'_k) - c_2 = 2^{k-3} - c_2 - c_0 = 2L(s_k) - c_2 - c_0$$

for some constant $c_2 > 0$. This completes the proof of the theorem. □

Theorem 12. *There is an infinite sequence s such that $s \in A_{k,\varepsilon}$ for almost all k, but the equation (1) does not hold for s.*

Proof. Let k_0 and s_k be defined as in Theorem 11. Define an infinite binary sequence s as follows:

$$s[2^{k-1}..2^k - 1] = \begin{cases} 0 \ldots 0 & \text{if } k \leq k_0, \\ s_k & \text{otherwise.} \end{cases}$$

Then $s \in A_{k,\varepsilon}$ for all $k > k_0$. Now we show that the equation (1) does not hold for s. It is straightforward to check that for any $k > k_0$,

$$K(s[1..2^k]) \leq \sum_{i=k_0+1}^{k} K(s_k) + c \leq 2^{k-2} + c$$

for some constant $c > 0$. However, by the choice of s_k in Theorem 11, for any $k > k_0$,

$$L(s[1..2^k]) \geq L(s_k s_{k+1}[1]) \geq 2^{k-2}.$$

Whence

$$\limsup_{n\to\infty} \frac{K(s[1..n])}{L(s[1..n])} \leq \lim_{k\to\infty} \frac{2^{k-2}+c}{2^{k-2}} = 1 < 2.$$

Which implies that the equation (1) can not hold for s. □

Theorem 12 shows that Beth and Dai's "proof" of Theorem 9 is incorrect. In the following we will prove a stronger result which implies Theorem 9. We first prove several lemmas.

Lemma 3. *For any positive integer n,*

$$|\{s \in \{0,1\}^n : L(s) \leq \frac{n}{2} - \log n\}| \leq 2^{n-2\log n}$$

and

$$|\{s \in \{0,1\}^n : L(s) \geq \frac{n}{2} + \log n\}| \leq 2^{n-2\log n}.$$

Proof. By Theorem 5, we have

$$\begin{aligned} |\{s \in \{0,1\}^n : L(s) \leq \frac{n}{2} - \log n\}| &= \sum_{k \leq (n/2)-\log n} N_n(k) \\ &= \sum_{k \leq (n/2)-\log n} 2^{2k-1} \\ &\leq 2^{n-2\log n} \end{aligned}$$

In the same way, we can show that $|\{s \in \{0,1\}^n : L(s) \geq \frac{n}{2} + \log n\}| \leq 2^{n-2\log n}$. □

Lemma 4. *(see, e.g., [7]) Let $\mathbf{C}_1, \mathbf{C}_2, \ldots \subseteq \{0,1\}^\infty$ be a recursively presentable sequence of Lebesgue measurable sets, such that $\sum_{n=1}^{\infty} Prob[\mathbf{C}_n]$ converges to a finite number effectively. Then for any Martin-Löf random sequence $s \in \{0,1\}^\infty$, $s \in \mathbf{C}_n$ for only finitely many n.*

Theorem 13. *Let $s \in \{0,1\}^\infty$ be a Martin-Löf random sequence. Then the equation (1) holds for s.*

Proof. Define a recursively presentable sequence of Lebesgue measurable sets as follows (for more details about recursively presentable sequence of sets, it is referred to Martin-Löf [8]): for each positive integer n, let

$$\mathbf{C}_n = \{s \in \{0,1\}^\infty : L(s[1..n]) \leq \frac{n}{2} - \log n \text{ or } L(s[1..n]) \geq \frac{n}{2} + \log n\}.$$

Then, by Lemma 3,

$$Prob[\mathbf{C}_n] \leq 2 \cdot 2^{-2\log n} = \frac{2}{n^2}.$$

Whence, $\sum_{n=1}^{\infty} Prob[\mathbf{C}_n]$ converges to a finite number effectively. By Theorem 4, for any Martin-Löf random sequence s, we have $s \in \mathbf{C}_n$ for only finitely many n. That is, $n/2 - \log n \leq L(s[1..n]) \leq n/2 + \log n$ for almost all n. Now, by Theorem 1,

$$\limsup_{n\to\infty} \frac{K(s[1..n])}{L(s[1..n])} \leq \lim_{n\to\infty} \frac{n+c}{n/2 - \log n} = 2$$

and

$$\liminf_{n\to\infty} \frac{K(s[1..n])}{L(s[1..n])} \geq \lim_{n\to\infty} \frac{n - 2\log n}{n/2 + \log n} = 2.$$

This completes the proof of the theorem. □

Proof of Theorem 9. This follows from Theorems 7 and 13. □

Remark: Note that Niederreiter [12] has proved the following result: for any function f defined on the positive integers with the property that $\sum_{n=1}^{\infty} 2^{-f(n)} < \infty$, we have $Prob[\mathbf{C}_f] = 1$, where

$$\mathbf{C}_f = \left\{ s \in \{0,1\}^{\infty} : \left| L(s[1..i]) - \frac{i}{2} \right| \leq f(i) \text{ a.e.} \right\}.$$

Whence, by Theorem 7 and the above result of Niederreiter, Theorem 9 follows. However, Theorem 13 is stronger than Theorem 9.

4 Comments on statistical tests

We conclude our note with some comments on statistical tests for pseudorandom bits generated by pseudorandom generators.

Martin-Löf randomness concept [8] has been the most successful one defined in the literature (see, e.g., [7]). By the Chaitin-Kolmogorov complexity characterization of Martin-Löf's random sequences, a sequence is Martin-Löf random if and only if it is incompressible, that is, if and only if the sequence withstands the compressibility test.

A Martin-Löf random sequence is defined to withstand all computational statistic tests (with unlimited resource bound). For example (see, e.g., [7]), a Martin-Löf random sequence withstands the frequency test (that is, the law of large numbers), the gap test, the correlation test, and the law of the iterated logarithm test. The result (that is, Theorem 13) of this paper can be interpreted in the following sense: a Martin-Löf random sequence withstands the LFSR tests. Amongst others, many of the statistic tests used for testing true randomness have been used to test the quality of the pseudorandom bits generated by a pseudorandom generator (see, e.g., [10, 11]), for example, the frequency test, the gap test, the correlation test, and the Maurer universal test which is a kind of

compressibility test. However, a very strong test, the law of the iterated logarithm test (LIL test), has not been used in testing pseudorandom sequences. The celebrated iterated logarithm has been one of the most beautiful and profound discoveries (see, e.g., [3, 4]) of probability theory. Wang [16–18] has shown that the law of the iterated logarithm holds for infinite polynomial time pseudorandom sequences. We suggest that this law should also be used in testing the quality of finite pseudorandom strings. By a standard diagonalization argument of Ville [15] for constructing "Kollektiv" sequences, it can be shown that there are sequences which pass the "randomness postulate" test (by Golomb [5]), the "universal test" (by Maurer [10]), and the LFSR test (by Rueppel [12–14]), but do not pass the LIL test. It should be mentioned that the LIL test can be finished in $O(n^2)$ time, whence it is an attractive addition to the currently used tests.

In the following we will present more details on the LIL test. For a sequence s, let $S_n(s) = \sum_{i=1}^{n} s[i]$. Then a sequence is said to withstand the frequency test if the value of $\frac{S_n(s)}{n}$ is close enough to $\frac{1}{2}$. However, this test is not successful in detecting whether a string s always has more 1's than 0's in its initial segments. Obviously, a pseudorandom sequence has some deficiency if there is always more 1's (or 0's) than 0's (or 1's) in its initial segments. As we have mentioned in the previous paragraph, Ville's construction can be used to show that all popular statistical tests used for pseudorandomness in the literature (see, e.g., [10, 11]) cannot detect this deficiency which may have undesired effects in certain applications. We suggest the use of the law of the iterated logarithm (LIL). This test is "universal" in the sense that it covers many of the commonly used statistical tests such as the gap test and the frequency test, and in addition, it can detect the above mentioned deficiency in a sequence.

For a sequence s, let

$$S_n^*(s) = \frac{2 \cdot S_n(s) - n}{\sqrt{n}}$$

denote the *reduced number* of 1's in $s[1..n]$. Note that $S_n^*(s)$ amounts to measuring the deviations of $S_n(s)$ from $\frac{n}{2}$ in units of $\frac{1}{2}\sqrt{n}$. In probability theory, $S_n(s)$ is called the *number of successes* and $S_n^*(s)$ is called the *reduced number of successes*.

The law of large numbers says that, for a pseudorandom string s, the limit of $\frac{S_n(s[1..n])}{n}$ is $\frac{1}{2}$. But it says nothing about the reduced deviation $S_n^*(s[1..n])$. It is intuitively clear that, for a pseudorandom string s, $S_n^*(s[1..n])$ will sooner or later take on arbitrary large values. Moderate values of $S_n^*(s[1..n])$ are most probable, but the maxima will slowly increase. How fast? Can we give an optimal upper bound for the fluctuations of $S_n^*(s[1..n])$? The law of the iterated logarithm, which was first discovered by Khintchine for the classical cases, gives a satisfactory answer for the above questions.

Definition 1. *A sequence* $s \in \{0,1\}^\infty$ *satisfies* the law of the iterated logarithm *if*

$$\limsup_{n\to\infty} \frac{2\sum_{i=1}^{n} s[i] - n}{\sqrt{2n \ln \ln n}} = 1$$

and

$$\liminf_{n\to\infty} \frac{2\sum_{i=1}^{n} s[i] - n}{\sqrt{2n \ln\ln n}} = -1.$$

It has been shown that the law of the iterated logarithm holds for Martin-Löf random sequences (see [17]) and for infinite polynomial time pseudorandom sequences (see [16–18]). Since there is an efficient algorithm to compute the reduced number of successes in a string, a pseudorandom sequence $s \in \{0,1\}^n$ of high "quality" should have the following LIL property: for large enough $i \leq n$, the value of $\frac{2\sum_{j=1}^{i} s[j]-i}{\sqrt{2i \ln\ln i}}$ should lie in the interval $[-1 - f(i), 1 + f(i)]$ for some function $f(i) \in o(\frac{1}{i})$ and the value should "reach" both 1 and -1 "frequently".

In the above paragraphs, we have given an outline of the LIL test for cryptographic pseudorandomness. In order to implement this test in practice, much work still needs to be done. For example, the distribution of the random variable $X_i = \frac{2\sum_{j=1}^{i} s[j]-i}{\sqrt{2i \ln\ln i}}$ should be carefully analyzed. For more details, it is referred to Chow [3] and Feller [4].

Acknowledgment

The author would like to thank one anonymous referee for his valuable report which helps improve the presentation of this paper.

References

1. T. Beth and Z.-D. Dai. On the complexity of pseudo-random sequences — or: If you can describe a sequence it cannot be random. In: *Advances in Cryptology, Proc. of Eurocrypt '89*, pp. 533–543, LNCS 434, Springer Verlag, 1989.
2. G. J. Chaitin. On the length of programs for computing finite binary sequences. *J. Assoc. Comput. Mach.*, 13:547–569, 1966.
3. Y. S. Chow and H. Teicher. *Probability Theory*. Springer Verlag, 1997.
4. W. Feller. *Introduction to Probability Theory and Its Applications*. Volume I. John Wiley & Sons, Inc., 1968.
5. S. W. Golomb. *Shift Register Sequences*. Holden-Day, San Francisco, Calif., 1967.
6. A. N. Kolmogorov. Three approaches to the definition of the concept "quantity of information". *Problemy Inform. Transmission*, **1**:3–7, 1965.
7. M. Li and P. Vitanyi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, 1993.
8. P. Martin-Löf. The definition of random sequences. *Information and Control*, **9**:602–619, 1966.
9. J. L. Massey. Shift-register synthesis and BCH decoding. *IEEE Transactions on Information Theory*, **15**:122–127, 1969.
10. U. Maurer. A universal statistical test for random bit generators. *Journal of Cryptology*, **5**:89–105, 1992.
11. A. Menezes, P. van Oorschot, and S. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.

12. H. Niederreiter. The probability theory of linear complexity. In: *Advances in Cryptology, Proc. of Eurocrypt '88*, pp. 191–209, LNCS 330, Springer Verlag, 1989.
13. R. Rueppel. Linear complexity and random sequences. In: *Advances in Cryptology, Proc. of Eurocrypt '85*, pp. 167–188, LNCS 219, Springer Verlag, 1986.
14. R. Rueppel. *Analysis and Design of Stream Ciphers*. Springer, 1986.
15. J. Ville. *Ètude Critique de la Notion de Collectif*. Gauthiers-Villars, Paris, 1939.
16. Y. Wang. The law of the iterated logarithm for p-random sequences. In: *Proc. 11th Conference on Computational Complexity (formerly: Conference on Structure in Complexity Theory)*, pages 180-189. IEEE Computer Society Press, 1996.
17. Y. Wang. *Randomness and Complexity*. PhD thesis, Universität Heidelberg, 1996.
18. Y. Wang. Resource bounded randomness and computational complexity. To appear in: *Theoretical Computer Science*, 1999.