Applicable Algebra in Engineering, Communication and Computing (2022) 33:485–504
https://doi.org/10.1007/s00200-020-00467-3

ORIGINAL PAPER

On the stability of periodic binary sequences with zone restriction

Ming Su[1] · Qiang Wang[2]

Received: 31 March 2019 / Revised: 12 June 2020 / Accepted: 26 September 2020 /
Published online: 11 October 2020

Abstract

Traditional global stability measure for sequences is hard to determine because of large search space. We propose the k-error linear complexity with a zone restriction for measuring the local stability of sequences. For several classes of sequences, we demonstrate that the k-error linear complexity is identical to the k-error linear complexity within a zone, while the length of a zone is much smaller than the whole period when the k-error linear complexity is large. These sequences have periods 2^n, or $2^v r$ (r odd prime and 2 is primitive modulo r), or $2^v p_1^{s_1} \cdots p_n^{s_n}$ (p_i is an odd prime and 2 is primitive modulo p_i^2, where $1 \leq i \leq n$) respectively. In particular, we completely determine the spectrum of 1-error linear complexity with any zone length for an arbitrary 2^n-periodic binary sequence.

Keywords Stability · Linear complexity · k-error linear complexity · Zone restriction · Binary sequences

Mathematics Subject Classification 94A60 · 94A55 · 65C10 · 68P25

Ming Su is supported by the China Scholarship Council and partially by Tianjin Key Research and Development Project 19YFZCSF00900. The research of Qiang Wang is partially funded by NSERC of Canada.

✉ Ming Su
nksuker@gmail.com

Qiang Wang
wang@math.carleton.ca

[1] Department of Computer Science, Nankai University, Tianjin, China

[2] School of Mathematics and Statistics, Carleton University, Ottawa, Canada

1 Introduction

Let $S = (s_0, s_1, s_2, \dots)$ be an N-periodic sequence with terms in the finite field $\mathbb{F}_q$ of q elements. We note that N need not be the least period of the sequence. We denote $S = (s_0, s_1, \dots, s_{N-1})^\infty$ and define $S^N(x) = s_0 + s_1 x + \cdots + s_{N-1}x^{N-1}$. The *linear complexity* of a periodic sequence over $\mathbb{F}_q$ is the length of the shortest linear recurrence relation which the sequence satisfies. In algebraic terms the linear complexity of an N-periodic sequence is given by $L(S) = N - \deg(\gcd(1 - x^N, S^N(x)))$; see for example [3, p. 28].

For an integer k, $0 \le k \le N$, the minimum linear complexity of those sequences with not more than k term changes in a period N from the original sequence S is called the *k-error linear complexity* of S, denoted as $L_{N,k}(S)$, i.e.,

$$L_{N,k}(S) = \min_{W_H(T) \le k} \{L(S + T)\},$$

where T is an N-periodic sequence, $W_H(T)$ is the Hamming weight of T in one period, the addition "+" for two sequences is defined elementwise in $\mathbb{F}_q$. A sequence T reaching the $L_{N,k}(S)$ is called an *error vector* of the k-error linear complexity. When $N = 2^n$, denote the k-error linear complexity of S by $L_k(S)$.

In addition to the Berlekamp–Massey algorithm [11] for computing the linear complexity with computational complexity $O(N^2)$, there are efficient algorithms of several types of periodic sequences with computational complexity $O(N)$, such as the Games-Chan algorithm [7] for computing the linear complexity of a 2^n-periodic binary sequence; the algorithm due to Meidl [13] for computing the linear complexity of a $u2^n$-periodic binary sequence, where u is odd; the algorithm for computing the linear complexity of a sequence with period p^n over $\mathbb{F}_q$ [26], where p is an odd prime and q is a prime and a primitive root mod p^2; and the algorithm for computing the linear complexity of a sequence with period $2p^n$ over $\mathbb{F}_q$ [25], where p and q are odd primes, and q is a primitive root mod p^2. These algorithms work because the factorization of $X^N - 1$ is simple under these assumptions.

Correspondingly, there are also efficient algorithms of computing the k-error linear complexity for certain types of sequences such as the Stamp–Martin algorithm [20] for computing the k-error linear complexity of a 2^n-periodic binary sequence, the algorithm for computing the k-error linear complexity of p^n-periodic sequences over $\mathbb{F}_{p^m}$ [9], and the algorithm for computing the k-error linear complexity of a sequence with period $2p^n$ over $\mathbb{F}_q$ [27]. We also remark that there are some studies on the properties of k-error linear complexity of binary sequences, see [8, 23]. Earlier, Sălăgean et al. studied approximation algorithms for the k-error linear complexity of odd-periodic binary sequences by using DFT and some relaxation [1, 19]. However, there is no efficient general algorithm for calculating the k-error linear complexity for an arbitrary binary sequence, in particular, binary sequence with arbitrary even period.

A well-designed sequence should not only have a large linear complexity, but also large k-error linear complexities for cryptographic purpose. This means its linear complexity should not decrease a lot when k errors occur; see [20] and [4]. In order to measure the stability of a given periodic sequence, we have to consider k errors that can occur anywhere within the whole period N. This means the computational task is heavy because the capacity of search space for all possible binary errors is $\sum_{t=0}^{k}\binom{N}{t}$, which is very large for common N and moderate k. Indeed, it becomes exponential of N when k is large, resulting in infeasible computations. This motivates us to study k-error linear complexity with a zone restriction. Intuitively, there can be many error vectors that reach the k-error linear complexity. We show that for many sequences we can find a window of proper length Z containing at least one error vector, no matter where we start with. For this purpose, we first define the *k-error linear complexity with a zone of length Z starting at the position j*, denote by (N, k; Z, j)-error linear complexity, as the minimum of all linear complexities such that these errors occur in positions between j and $j + Z - 1$. That is,

$$L_{N,k;Z,j}(S) = \min_{\substack{W_H(T) \leq k \\ supp(T) \subseteq [j, j+Z \bmod N)}} \{L(S+T)\},$$

where $[j, j+Z \bmod N) := [j, N) \cup [0, (j+Z) \bmod N)$ if $N - Z < j < N$. Moreover, we define the *k-error linear complexity with a zone of length Z* as

$$\mathcal{L}_{N,k;Z}(S) = \min_{j\in[0,N)} \{L_{N,k;Z,j}(S)\},$$

Obviously, $L_{N,k;Z,j}(S)$ is easier to compute and this provides a natural upper bound of $L_{N,k}(S)$. By the definition, $L_{N,k;Z,j}(S)$ can be different for different choices of j's.

In this paper, we study the relation between $L_{N,k;Z,j}(S)$ and $L_{N,k}(S)$ and prove that for a zone length Z appropriately chosen, for any j we have $\mathcal{L}_{N,k;Z}(S) = L_{N,k;Z,j}(S) = L_{N,k}(S)$ for several classes of sequences, and Z can be very small compared to the period N. Accordingly, we can efficiently determine the global stability via a local stability. We focus on binary sequences with even period and a large k-error linear complexity, in particular, several classes of sequences with periods 2^n, or $2^v r$ (r odd prime and 2 is primitive modulo r), or $2^v p_1^{s_1} \cdots p_n^{s_n}$ (p_i is an odd prime and 2 is primitive modulo p_i^2, where $1 \leq i \leq n$) respectively.

Sequences with period 2^n have attracted a lot of attention [5]; one typical example is the de Bruijin sequence of maximal 2^n-periodic sequence generated by NFSR of stage n [2]. Despite that there is an efficient algorithm to compute the k-error linear complexity of these sequences, we still demonstrate our method by showing that there exists a small zone of length $Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$ containing the *support* (positions of nonzero entries) of an error vector reaching the k-error linear complexity for any 2^n-periodic binary sequence. This means that we can indeed reduce the global stability to a local stability. Furthermore, we completely describe the spectrum of 1-error linear complexity with any given zone length. This can help us to obtain the exact counting functions of the (N, k; Z, 0)-error linear complexity for any 2^n-periodic binary sequence.

Moreover, we found two more classes of binary sequences such that their global stability can be reduced to a local stability. The first class of sequences has a large linear complexity and a large k-error linear complexity with period $2^v r$, such that r is an odd prime and 2 is a primitive root modulo r. The length of a zone is $Z = 2^{\lceil \log_2(N - L_{N,k}(S)) \rceil}$. More details can be found in Theorem 3. We want to emphasize that our result applies to quite a lot of sequences. By Artin's conjecture, approximately 37% of all primes satisfy that 2 is a primitive root modulo r. We also justify that there is a high proportion of sequences who have the required a large linear complexity and a large k-error linear complexity, among those sequences with period $2^v r$ where 2 is primitive modulo r. In particular, we show that if 2^v is upper-bounded by a polynomial of r and $\sum_{t=0}^{k}\binom{2^v r}{t} < \frac{2^{r-1}}{2^v}$, then a large proportion of these sequences have desired properties so that their global stability can be reduced to local stability. The second class of sequences has the period $N = 2^v p_1^{s_1} p_2^{s_2} \dots p_n^{s_n}$, where p_i is odd prime and 2 is a primitive root modulo p_i^2 for all $1 \le i \le n$. For any such N-periodic binary sequence S such that $L(S) > L_{N,k}(S) \ge N - \min(2^v, p_1 - 2, \dots, p_n - 2)$, we show in Theorem 4 that there exists a zone of length $Z = 2^{\lceil \log_2(N - L_{N,k}(S)) \rceil}$ such that $L_{N,k}(S) = L_{N,k;Z,j}(S)$ for any j. We remark that the length Z of the properly chosen zone is dependent on the k-error linear complexity of the sequence. One may argue that we can not save on computation since we do not know apriori what Z should be. However, most sequences that are used in cryptography have a large k-error linear complexity, and thus the length Z is rather small. Therefore, with a reasonable effort and a few trials, the computations would stablize and an error vector can be found within the zone. In summary, our reduction method works best for those types of periodic sequences with the large k-error linear complexity, which is very common due to cryptographic requirement.

The rest of this paper is organized as follows. In Sect. 2, we study the $(k;\ Z,\ j)$-error linear complexity for any periodic binary sequence S with period 2^n, and find a proper zone of length $Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$ such that $L_{N,k}(S) = L_{N,k;Z,j}(S) = \mathcal{L}_{N,k;Z}(S)$ for any j. The larger $L_{N,k}(S)$, the smaller zone length Z. In Sect. 3, we study the linear complexity affected by 1-error occurrence within a zone of length Z, and give the exact counting functions of the 1-error linear complexity with a restriction on zone length Z for a random 2^n-periodic binary sequence. In Sect. 4, we prove Theorems 3 and 4.

2 Reduction from global stability to local stability with a zone restriction for any binary sequence of period 2^n

In this section, we show that the global stability can be reduced to local stability with zone restriction for any binary sequence of period 2^n. We denote the binary sequence with the only nonzero entry '1' at position j by $E_1(j)$, $0 \le j < 2^n$, in each period 2^n, and the expected $(k;\ Z,\ 0)$ linear complexity of N-periodic sequences by $\mathcal{E}_{N,k;Z,0}$. Without causing any confusion, we denote $E_{k;Z,0}$ the expected $(k;\ Z,\ 0)$ linear complexity of 2^n-periodic binary sequences, and $\mathcal{N}_k(c)$ the number of sequences achieving k-error linear complexity value c of 2^n-periodic binary sequences. Because the sequence S has period 2^n, we only need to consider the multiplicity of $x = 1$ as a root of $S^{2^n}(x)$ when we compute the linear complexity $L(S) = 2^n - deg(gcd(1 - X^{2^n}, S^{2^n}(X))$. It is straightforward to derive the following useful result.

Lemma 1 *For two 2^n-periodic sequences S, S', if $L(S') \neq L(S)$, then $L(S+S') = \max\{L(S), L(S')\}$. If $L(S) = L(S')$, then $L(S+S') < L(S) = L(S')$. In particular $L(E_1(j) + E_1(j+i2^s)) \leq 2^n - 2^s$, where $0 \leq s < n$.*

Proof Obviously, we can write $S^{2^n}(x) = (1-x)^{2^n - L(S)} g_S(x)$ for the sequence S, where $g_S(1) = 1$. Similarly, $S'^{2^n}(x) = (1-x)^{2^n - L(S)} g_{S'}(x)$, where $g_{S'}(1) = 1$. If $L(S') \neq L(S)$, then $S^{2^n}(x) + S'^{2^n}(x) = (1-x)^{2^n - \max\{L(S), L(S')\}} \tilde{g}(x)$, and $\tilde{g}(1) = 1$. Therefore we have $L(S+S') = \max\{L(S), L(S')\}$. If $L(S') = L(S)$, we obtain $S^{2^n}(x) + S'^{2^n}(x) = (1-x)^{2^n - L(S)}(g_S(x) + g_{S'}(x))$, and $g_S(1) + g_{S'}(1) = 0$. Therefore, $L(S+S') < L(S) = L(S')$. In particular, the degree of $\gcd(x^j + x^{j+i2^s}, x^{2^n} - 1)$ is at least 2^s because $x^j + x^{j+i2^s} = x^j(1+x^i)^{2^s}$. □

For the trivial case $L_k(S)$=0, there is only one error vector in a period. It is well known from [12] and [23] that $L_k(S) \neq 2^n - 2^s$ for any integer $s < n$, $k > 1$, and for S with odd Hamming weight, $L_1(S) = L_0(S)$. In particular, when $k = 1$, we can determine the number of error vectors $E_1(j)$ such that $L(S + E_1(j)) = L_1(S)$.

Next, we study the number of error vectors by Lemma 1.

Lemma 2 *For a sequence S satisfying $L(S) = 2^n$ and $2^n - 2^s < L_1(S) < 2^n - 2^{s-1}$, we have exact 2^{n-s} error vectors with Hamming weight 1 in one period achieving $L_1(S)$.*

Proof Suppose there is $E_1(j)$ such that $0 \leq j < 2^s$ and $L(S + E_1(j)) = L_1(S)$, we claim that we must have a set of error vectors at positions $j + i2^s$, where $0 \leq i < 2^{n-s}$ and the addition + is performed modulo 2^n, such that $L(S + E_1(j+i2^s)) = L\Big(\big(S + E_1(j)\big) + \big(E_1(j) + E_1(j+i2^s)\big)\Big) = L(S + E_1(j))$. Since $L(E_1(j) + E_1(j+i2^s)) \leq 2^n - 2^s < L_1(S) = L(S + E_1(j))$, we conclude the above claim by Lemma 1.

For any other error vector $E_1(j')$ such that $j' - j \equiv d\,(mod\, 2^s)$ and $0 < d < 2^s$, the largest nonnegative integer t such that $2^t \mid (j' - j)$ must satisfy $t < s$. Hence the degree of $\gcd(x^j + x^{j'}, x^{2^n} - 1)$ is exactly $2^t \leq 2^{s-1}$ and thus $L(E_1(j) + E_1(j')) = 2^n - 2^t \geq 2^n - 2^{s-1} > L_1(S)$. Therefore $L(S + E_1(j')) = L(S + E_1(j) + E_1(j) + E_1(j')) = L(E_1(j) + E_1(j')) > L_1(S)$. This shows that there is exactly one error vector $E_1(j)$ such that $0 \leq j < 2^s$ and $L(S + E_1(j)) = L_1(S)$. Hence there are exactly 2^{n-s} error vectors in the whole period achieving $L_1(S)$. □

For any positive integer m, we let E_m denote a binary vector of length 2^n with Hamming weight m. For example, assume E_m has '1' at positions $i_1, i_2, \dots, i_m$, where $0 \leq i_1 < i_2 < \dots < i_m \leq 2^n - 1$. Then we define the support of E_m as $Supp(E_m) = \{i_1, i_2, \dots, i_m\}$. Now we can show that there exists at least one error vector whose support is contained in a smaller zone, by adjusting entries of an error vector.

Lemma 3 *Let S be a binary sequence with period 2^n. Suppose $2^n - 2^s < L_k(S) < 2^n - 2^{s-1}$ for some integer s. For any $0 \leq j < N$, there exists* *at least one error vector E_m of weight m, $m \leq k$ such that $L(S+E_m) = L_k(S)$ and $supp(E_m) \subseteq [j, j+2^s \bmod N)$, where $[j, j+2^s \bmod N) := [j, N) \cup [0, (j+2^s) \bmod N)$ if $N - 2^s < j < N$.*

Proof According to the definition of the k-error linear complexity, there exists an error vector E_m, $m \leq k$ of Hamming weight m such that $L(S+E_m) = L_k(S)$. If $i_m \geq 2^s$, then we can define a new vector $E_m^{'} := E_m + E_1(i_m) + E_1(i_m \bmod 2^s)$ so that

$$L(S+E_m^{'}) = L\big(S + E_m + E_1(i_m) + E_1(i_m \bmod 2^s)\big).$$

By the assumption and Lemma 1, we have

$$L(S+E_m) = L_k(S) > 2^n - 2^s \geq L(E_1(i_m) + E_1(i_m \bmod 2^s)),$$

and thus $L(S+E_{m'}) = L_k(S)$, where $Supp(E'_m) = Supp(E_m) \setminus \{i_m\} \cup \{i_m \bmod 2^s\}$. Therefore, we can consecutively adjust those entries of E_m so that we can find $\bar{E}_m$ such that $supp(\bar{E}_m) \subseteq [0, 2^s)$ and $L(S+\bar{E}_m) = L_k(S)$.

Since $L(S) = N - \deg(\gcd(1-x^N, S^N(x))) = N - \deg(\gcd(1-x^N, x^j \cdot S^N(x)))$, the linear complexity of a periodic sequence S is invariant to the cyclic shift with respect to the offset j, and the cyclic-shift invariant property also holds for the k-error linear complexity. Therefore, the proof is complete. □

Because of the assumption $2^n - 2^s < L_k(S) < 2^n - 2^{s-1}$, we derive $2^{s-1} < 2^n - L_k(S) < 2^s$. Let $Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$. Then for any $0 \leq j < N$, $L_{k;Z,j}(S) \leq L_k(S)$ by Lemma 3. Conversely, it is always true that $L_k(S) \leq L_{k;Z,j}(S)$. Therefore we obtain the following theorem.

Theorem 1 *Let S be any 2^n-periodic binary sequence. For any positive integer k, there always exists $k' \leq k$ such that $k' \leq Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$ and $L_k(S) = L_{k'}(S) = L_{k';Z,j}(S) = \mathcal{L}_{k';Z}(S)$, for any $0 \leq j < 2^n$.*

Proof If $k \leq Z$, then $L_k(S) = L_{k;Z,j}(S)$. Otherwise if $k > Z$, then we can find an error vector $E_{k'}$ with the support in $[j, j+Z \bmod N)$ such that $L_k(S) = L(S+E_{k'})$ by adjusting the error positions into the zone, in which some mapped positions in the zone will collide. Accordingly, we have $L_k(S) = L_{k'}(S)$. Then, $L_{k'}(S) = L_{k';Z,j}(S)$ for any $0 \leq j < 2^n$, and the latter is equal to $\mathcal{L}_{k';Z}(S)$. □

Theorem 1 shows that we can efficiently verify the global stability of a binary sequence of period 2^n with large k-error linear complexity via a local stability. If $L_k(S)$ is big, then Z can be very small. If k is large, then we can reduce k so that it is bounded by the zone length as well. Of course, there is the degenerated case when $L_k(S) = 0$, in this case we have to set $Z = 2^n$. However, as we commented earlier, we focus on sequences with a large k-error linear complexity and thus the zone length is significantly reduced.

Stamp and Martin [20] proposed the efficient algorithm for computing the k-error linear complexity of 2^n-periodic binary sequence, which is an extension of the Games-Chan algorithm [7], by using a cost array recording the number of bits changes required in the original sequence without influence on the results of any previous steps. The error vectors were not investigated in [20]. Lauder and Paterson [10] studied the error linear complexity spectrum of 2^n-periodic binary sequence, where they can recover one error vector (they defined it as the *critical error sequence*) by a deterministic algorithm using the B-Pullup and L-Pullup [10, p. 275, 1 a), 2 a)].

Remark 1 Let $2^n - 2^s < L_k(S) < 2^n - 2^{s-1}$, then according to the Stamp-Martin algorithm, the left part and the right part of intermediate sequence are not equal for the initial $n - s$ steps. By the deterministic B-Pullup algorithm [10, p. 275, 1 a)] of an error sequence of length 2^s, the support of the expanded error sequence is located at the left part. Indeed, followed by $n - s$ B-Pullups we finally obtain an error vector, whose support is in the zone $[0, 2^s = 2^{\lceil \log_2(2^n - L_k(S)) \rceil})$.

Example 1 Let $S = (10000011001000110000000000000000)^\infty$ be a 32-periodic sequence. By the algorithm for computing the error linear complexity spectrum [10, p. 277] as well as the 'C' code implementation at http://www.isg.rhul.ac.uk/~kp/celcs.c, we obtain $L(S) = 30$, $L_2(S) = 23$, and deduce one error vector by the deterministic B-Pullup and L-Pullup [10, p. 275, 1 a), 2 a); p.278] as follows:

$$\begin{aligned} e = 0 &\overset{LPullup}{\rightarrow} 10 \overset{BPullup}{\rightarrow} 1000 \overset{BPullup}{\rightarrow} 10000000 \overset{LPullup}{\rightarrow} 0010000010000000 \\ &\overset{BPullup}{\rightarrow} 00100000100000000000000000000000. \end{aligned} \tag{1}$$

Therefore, from their method one error vector with support (2, 8) is derived. However, there still exists another error vector with support (0, 10) in the zone [0, 16), not discovered by the B-Pullup and L-Pullup.

By a deterministic algorithm using the B-Pullup and L-Pullup [10, p. 275, 1 a), 2 a)], one error vector is obtained. But there may exist many error vectors, even in the zone. Actually, the number of error vectors of $L_k(S)$ depends on the value of $L_k(S)$, and we can determine the complete set of error vectors of $L_k(S)$ of 2^n-periodic sequences from an error vector by some transformations for $k \le 4$, see [23]. In the following, we will provide an algorithm deriving an error error within the zone $[0, Z)$, $Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$ from any other error vector.

Algorithm 1 Computing an error vector within the zone $[0, Z)$, $Z = 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$

```
Input: Any error vector of $L_k(S)$ with the support $(i_1, i_2, \cdots, i_m)$, where $m \le k$
Output: An error vector within the zone $[0, Z)$
1: Computing $L_k(S)$ by the Stamp-Martin, or Lauder-Paterson Algorithm
2: $Z \leftarrow 2^{\lceil \log_2(2^n - L_k(S)) \rceil}$
3: for $l = 1 \to m$ do
4:     $j_l \leftarrow i_l \pmod Z$
5: end for
6: return An error vector with the support $(j_1, j_2, \cdots, j_m)$
```

Note that Algorithm 1 can be directly extended to any zone $[j, j + Z \bmod N)$ because of the cyclical-shift-invariant property on the k-error linear complexity of a periodic sequence. Moreover, the computational complexity of the steps 3–5 in Algorithm 1 is $O(k)$, which does not bring much extra cost in addition to the Stamp-Martin or Lauder-Paterson Algorithm.

Example 2 Let $S = (10000001100100011000000000000000)^{\infty}$ be a 32-periodic sequence, the same sequence in Example 1. Since $L_2(S) = 23$, we have $Z = 16$. We can verify that the vector with support (0, 26) is an error vector of $L_2(S)$, by Algorithm 1 we obtain an error vector within the zone $Z = 16$, whose support is (0, 10).

3 Spectrum of 1-error linear complexity with arbtrary zone length

In this section we assume $N = 2^n$ and $n \ge 4$. It is well known that the linear complexity of a 2^n-periodic sequence S is 2^n if and only if it has odd Hamming weight. The 1-error linear complexity of a 2^n-periodic sequence can be any integer between 0 and $2^n - 1$. However, if S has an odd Hamming weight, then $L_1(S)$ can not be any integer of the form $2^n - 2^s$ where $1 \le s \le n$. For more details we refer the reader to [6, 12, 21, 22, 24].

If S is a 2^n-periodic sequence with even Hamming weight, then $L_1(S) = L_0(S)$. In this case, $L_{1;Z,0}(S) = L_0(S)$ for any zone of length Z. In order to study the distribution of $L_{1;Z,0}$, we only need to consider 2^n-periodic sequences with odd Hamming weight.

Theorem 2 *Let S be a 2^n-periodic sequence with odd Hamming weight and L_1 be its 1-error linear complexity. Let $0 < Z \le 2^n$, and $a = \lfloor \log_2(Z) \rfloor$.*

1. *For $2^n - 2^s < L_1 < 2^n - 2^{s-1}$ with some integer s, we have the following*
   (i) *if $s \le a$, then $L_{1;Z,0}(S) = L_1$, and the number of such sequences S is $2^{L_1 - 1 + s}$;*
   (ii) *if $s > a$, then we have*

$$L_{1;Z,0}(S) \in \{L_1, 2^n - 2^{s-1}, \dots, 2^n - 2^a\}.$$

The number of all the sequences S that achieve these values equals

$$\begin{cases} Z \cdot 2^{L_1-1}, & \textit{if } L_{1;Z,0} = L_{1;Z,0}(S) = L_1; \\ Z \cdot 2^{L_1+s-t-2}, & \textit{if } L_{1;Z,0} = 2^n - 2^t, \textit{where } a+1 \le t \le s-1; \\ 2^{L_1-1+s} - Z \cdot 2^{L_1+s-a-2}, & \textit{if } L_{1;Z,0} = 2^n - 2^a. \end{cases}$$

2. *For* $L_1 = 0$, *we have*

$$L_{1;Z,0}(S) \in \{0, 2^n - 2^{n-1}, \dots, 2^n - 2^a\}.$$

The number of all the sequences S that achieve these values equals

$$\begin{cases} Z, & \textit{if } L_{1;Z,0} = L_1; \\ Z \cdot 2^{n-t-1}, & \textit{if } L_{1;Z,0} = 2^n - 2^t, \textit{ where } a+1 \le t \le n-1; \\ 2^n - Z \cdot 2^{n-a-1}, & \textit{if } L_{1;Z,0} = 2^n - 2^a. \end{cases}$$

Proof For any 2^n-periodic sequence S with odd Hamming weight, we have $2^n - 2^s < L_1(S) < 2^n - 2^{s-1}$ for some positive integer s or $L_1 = 0$. Let us first assume that $2^n - 2^s < L_1(S) < 2^n - 2^{s-1}$ for some positive integer s. From the proof of Lemma 2, there exists exactly one j, $0 \le j < 2^s$, such that $L_1(S) = L(S + E(j))$.

(i) if $s \le a$, then $2^s \le 2^a \le Z$ by the definition of $a = \lfloor \log_2(Z) \rfloor$. Hence, in the zone of length Z, there is at least one error vector reaching the $L_1(S)$, so $L_{1;Z,0}(S) = L_1(S)$.

(ii) if $s > a$, then $Z < 2^s$. Consider $0 \le j < 2^s$ such that $L_1(S) = L(S + E(j))$. Let $m \cdot 2^{a+1} \le j < (m+1)2^{a+1}$ for some nonnegative integer m. Let $\bar{j} = j - m2^{a+1}$ be the positive integer less than 2^{a+1} such that $\bar{j} \equiv j \bmod 2^{a+1}$. If $Z < \bar{j} < 2^{a+1}$, then we take $j' = \bar{j} - 2^a$, which satisfies $0 < j' < 2^a < Z$. Since $j' - j$ is an odd multiple of 2^a, we conclude that $L(E_1(j) + E_1(j')) = 2^n - 2^a$ because the multiplicity of the root 1 for the binomial $x^j + x^{j'}$ is exactly 2^a. Hence $L(S + E_1(j)) = L_1(S) < 2^n - 2^{s-1} \le 2^n - 2^a = L(E_1(j) + E_1(j'))$. Therefore $L(S + E_1(j')) = \max\{L(S + E_1(j)), L(E_1(j) + E_1(j'))\} = 2^n - 2^a$. And for any $\hat{j}$ satisfying $Z < \hat{j} < 2^{a+1}$, $\hat{j} \ne j'$, we have $L(E_1(j') + E_1(\hat{j})) > 2^n - 2^a$ because $|\hat{j} - j'| < 2^a$, accordingly we have $L(S + E_1(\hat{j})) = \max\{L(S + E_1(j')), L(E_1(\hat{j})) + E_1(j'))\} > 2^n - 2^a$. Thus, $L_{1;Z,0} = 2^n - 2^a$.

On the other hand, we assume $0 \le \bar{j} \le Z$. For $m = 0$, we must have $L_{1;Z,0} = L_1$ because $0 \le j < Z$. Let $m \ge 1$. We claim $L_{1;Z,0} = 2^n - 2^{a+1+v_2(m)}$ where $v_2(m)$ is the *largest integer such that* $2^{v_2(m)} \mid m$. Indeed, we take $j' = j - m \cdot 2^{a+1}$. which satisfies $0 < j' < Z$. Since $j - j' = m \cdot 2^{a+1}$, we conclude that $L(E_1(j) + E_1(j')) = 2^n - 2^{a+1+v_2(m)}$ because the multiplicity of the root 1 for the binomial $x^j + x^{j'}$ is exactly $2^{a+1+v_2(m)}$.

Table 1 $L_{1;Z,0}$ values at different intervals

	$0 \le j \pmod{2^{a+1}} \le Z$	$Z < j \pmod{2^{a+1}} < 2^{a+1}$
$0 \le j \le 2^{a+1}$	L_1	$2^n - 2^a$
$2^{a+1} \le j < 2^{a+2}$	$2^n - 2^{a+1}$	$2^n - 2^a$
$\vdots$	$\vdots$	$\vdots$
$m \cdot 2^{a+1} \le j < (m+1)2^{a+1}$	$2^n - 2^{a+1+v_2(m)}$	$2^n - 2^a$
$\vdots$	$\vdots$	$\vdots$
$2^s - 2^{a+1} \le j < 2^s$	$2^n - 2^{s-1}$	$2^n - 2^a$

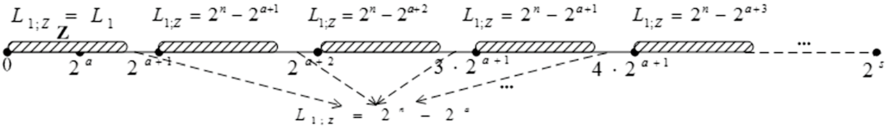

Fig. 1 $L_{1;Z,0}$ values for j at different intervals

We note that $a + 1 + v_2(m) \le s - 1$ because $j - j' = m2^{a+1} < 2^s$. Hence $L(S + E(j)) = L_1(S) < 2^n - 2^{s-1} \le 2^n - 2^{a+1+v_2(m)} = L(E_1(j) + E_1(j'))$. Therefore $L(S + E_1(j')) = \max\{L(S + E_1(j)), L(E_1(j) + E_1(j'))\} = 2^n - 2^{a+1+v_2(m)}$. And for any $\hat{j}$ satisfying $0 \le \hat{j} < Z$, $\hat{j} \ne j'$, we have $L(E_1(j') + E_1(\hat{j})) \ge 2^n - 2^a$ because $|\hat{j} - j'| < 2^{a+1}$, accordingly we have $L(S + E_1(\hat{j})) = \max\{L(S + E_1(j')), L(E_1(\hat{j})) + E_1(j'))\} \ge 2^n - 2^a$. In this case, $L_{1;Z,0} = 2^n - 2^{a+1+v_2(m)}$. Hence $L_{1;Z,0}(S) \in \{L_1, 2^n - 2^{s-1}, \dots, 2^n - 2^a\}$. We summarize these results in Table 1 and Fig. 1.

Now we count the number of all these sequences having the 1-error linear complexity L_1 and odd Hamming weight such that $L_{1;Z,0}(S) = 2^n - 2^t$ where $a + 1 \le t \le s - 1$. For each sequence S with 1-error linear complexity L_1 and odd Hamming weight, we need to count the number of error positions j's such that $L_{1;Z,0} = L(S + E_1(j)) = 2^n - 2^t$. We prove that the proportion of j's over an interval of length 2^{t+1} such that $L_{1;Z,0} = 2^n - 2^t$ is $Z/2^{t+1}$, where $t \ge a + 1$.

First we show that every sub-interval I of length 2^{t+1} in the interval $[0, 2^s)$ contains at least one interval of length Z for possible j's such that $L_{1;Z,0} = 2^n - 2^t$. We will construct an interval of length Z for j's within I. Because the length of I is 2^{t+1}, we can always choose an odd integer m' such that $[m'2^t, (m'+1)2^t) \subset I$. By the above proof, there exists an interval of length Z within the interval $m'2^t = (m'2^{t-a-1})2^{a+1} \le j < ((m'2^{t-a-1}) + 1)2^{a+1}$ such that $L_{1;Z,0} = 2^n - 2^{a+1+v_2(m'2^{t-a-1})} = 2^n - 2^{a+1+t-a-1} = 2^n - 2^t$.

Then we show that I can not contain more than one intervals of length Z for possible j's such that $L_{1;Z,0} = 2^n - 2^t$. We prove it by contradiction. Suppose there are j,j' such that $0 < |j' - j| < 2^{t+1}$ and $L(S + E_1(j)) = L(S + E_1(j')) = 2^n - 2^t$. In this case, $L(E_1(j) + E_1(j')$=$L(S + E_1(j) + S + E_1(j')) < 2^n - 2^t$. However, the root 1 of $x^j + x^{j'}$ is at most 2^t times, implying $L(E_1(j) + E_1(j')) \ge 2^n - 2^t$, a contradiction.

Therefore, the proportion of j's in each interval of length 2^{t+1} such that $L_{1;Z,0} = 2^n - 2^t$ is $Z/2^{t+1}$, for each sequence having the 1-error linear complexity L_1

and odd Hamming weight. Since there are 2^{L_1-1+s} sequences with odd Hamming weight such that $2^n - 2^s < L_1 < 2^n - 2^{s-1}$ (see [24][p. 2000, Theorem 3]), there are $Z2^{L_1-1+s}/2^{t+1} = Z2^{L_1+s-t-2}$ sequences having the 1-error linear complexity L_1 and odd Hamming weight such that $L_{1;Z,0}(S) = 2^n - 2^t$. Similarly, for $L_{1;Z,0} = L_1$, there is only one internal of length Z within the interval $[0, 2^s)$ and thus the proportion is $Z/2^s$. From Table 1 and Fig. 1, the proportion of j's giving $2^n - 2^a$ is $1 - \frac{Z}{2^{a+1}}$.

When $L_1 = 0$, the proof is similar and thus we omit the details. □

The distribution of k-error linear complexity is provided in [23] when $k \leq 4$. In particular, the number $N_0(c)$ of 2^n-periodic sequences with the linear complexity c is 2^{c-1}; see [14]. In Theorem 2, we have counted the number of sequences with odd Hamming weight achieving (1, Z)-error linear complexity values. In the following, we count the number $\mathcal{N}_{1;Z,0}(c)$ of all sequences achieving (1, Z)-error linear complexity value c, without emphasizing on their Hamming weights.

Corollary 1 *Let $a = \lfloor \log_2(Z) \rfloor$. The value $\mathcal{N}_{1;Z,0}(c)$ is equal to*

$$\begin{cases} 2^{c-1} + 2^{c-1+s}, & \textit{if } s \leq a, 2^n - 2^s < c < 2^n - 2^{s-1}; \\ (1+Z)2^{c-1}, & \textit{if } s > a, 2^n - 2^s < c < 2^n - 2^{s-1}; \\ \sum\limits_{s=a+1}^{n} \sum\limits_{L_1=2^n-2^s+1}^{2^n-2^{s-1}-1} (1 - \frac{Z}{2^{a+1}}) \cdot 2^{L_1-1+s} + (1 - \frac{Z}{2^{a+1}}) \cdot 2^n + 2^{2^n-2^a-1}, & \textit{if } c = 2^n - 2^a; \\ \sum\limits_{s=t+1}^{n} \sum\limits_{L_1=2^n-2^s+1}^{2^n-2^{s-1}-1} \frac{Z}{2^{t+1}} \cdot 2^{L_1-1+s} + \frac{Z}{2^{t+1}} \cdot 2^n + 2^{2^n-2^t-1}, & \textit{if } c = 2^n - 2^t, a < t \leq s-1; \\ 1 + Z, & \textit{if } c = 0. \end{cases}$$

Proof We note that every 2^n-periodic sequence S has a linear complexity $L(S) = c < 2^n$ if and only if it has an even Hamming weight. In this case, $L_1(S) = L(S)$. Hence the result follows immediately from Theorem 2. □

The exact expectation $\mathcal{E}_{1;Z,0}$ can be derived from the above counting functions, and may be used as a measure for determining the randomness of a 2^n-periodic binary sequence, with variations on Z. The exact formula is too complicated. Thus we omit all the details here. Instead, we provide a concrete example for the expected values of $L_{1;Z,0}$ for sequences with period $N = 2^8$ in Table 2. Note that the row $Z = N$ corresponds to the expected 1-error linear complexity of 2^8 -periodic binary sequences, where the expected 1-error linear complexity of 2^n -periodic binary sequences was provided in [12, 24].

4 Extension to sequences with other even periods

Now we consider stability of other periodic sequences with even period $N = 2^v r$ such that r, v are positive integer and r is odd. For some types of $2^v r$-periodic binary sequences, we can still find a proper zone of length Z so that $L_{N,k}(S) = L_{N,k;Z,j}(S)$.

Table 2 $\mathcal{E}_{N,1;Z,0}$ for $N = 2^8$

Z	$\mathcal{E}_{N,1;Z,0}$
1	254.0000
2	253.5000
3	253.2500
4	253.0000
5	252.9375
6	252.8750
7	252.8125
8	252.7500
N	252.7185

From the paper by Niederreiter [16, Theorem 1, P. 503], there exists $2^v r$-periodic binary sequence S with $L_{N,k}(S) \geq N - 2^v$ and $L(S) = N$, provided that

$$\sum_{j=0}^{k} \binom{N}{j} < 2^{\min_{2 \leq i \leq h} |C_i|},$$

where $C_2, \ldots, C_h$ are the different cyclotomic cosets modulo r. In the following, we will reveal that $L_{N,k}(S) = L_{N,k;Z,j}(S) = \mathcal{L}_{N,k;Z}(S)$ with certain $Z \ll N$ for some of these 'ideal' cryptographic sequences.

Theorem 3 *Let $N = 2^v r$, $v > 0$, r be an odd prime, and 2 be a primitive root modulo r. If S is an N-periodic binary sequence such that*

$$L(S) = N - c > L_{N,k}(S) \geq N - \min(2^v, r - 2), \tag{2}$$

for some nonnegative integer c, then for any $0 \leq j < N$ there exists at least one error vector E_m, $m \leq k$ such that

$$L(S + E_m) = L_{N,k}(S) \, and \, Supp(E_m) \subseteq [j, j + Z \bmod N),$$

where $Z = 2^{\lceil \log_2(N - L_{N,k}(S)) \rceil}$. In particular, if $L(S) = N > L_{N,1}(S) \geq N - \min(2^v, r - 2)$, then there exists exactly one error vector E_1 satisfying that $Supp(E_1) \subseteq [j, j + Z \bmod N)$ and $L_{N,1}(S) = L_{N,1;Z,j}(S)$.

Proof First we consider $j = 0$, i. e., the zone starts with 0. Because 2 is primitive root modulo the prime number r, the cyclotomic polynomial $\Phi_r(X)$ of the order r is irreducible over $\mathbb{F}_2$. Hence $X^{2^v r} - 1 = (X^r - 1)^{2^v} = ((X - 1)\Phi_r(X))^{2^v}$. Let α be a primitive r-th root of unity. That is, $\Phi_r(\alpha) = 0$.

If $2^v \leq r - 2$, then α can not be a root of $S^N(x) + E_t(x)$ for any error polynomial $E_t(x)$ of Hamming weight $t \leq k$ because of the assumption $L_{N,t}(S) \geq L_{N,k}(S) \geq N - 2^v$. Otherwise, $\Phi_r(X) \mid s^N(x) + E_t(x)$ and thus the greatest common divisor of $s^N(x) + E_t(x)$ and $X^N - 1$ has degree greater than 2^v, a contradiction.

If $2^v > r-2$, then $L_{N,k}(S) \geq N-r+2$. Similarly, α can not be a root of $S^N(x)+E_t(x)$ for any error polynomial $E_t(x)$ of Hamming weight $t \leq k$.

Now we only need to consider the multiplicity of root 1 when computing $L_{N,k}(S)$. As in the proof of Lemma 3, we can derive an error vector E_m such that $Supp(E_m) \subseteq [0,Z)$ and $L_k(S+E_m) = L_{N,k}(S)$. Indeed, suppose there exists an error vector E_m such that $L(S+E_m) = L_{N,k}(S)$, where $Supp(E_m) = \{i_1, i_2, \dots, i_m\}$. Note that $Z = 2^{\lceil \log_2(N-L_{N,k}(S)) \rceil} \geq N - L_{N,k}(S)$. If $i_m \geq Z$, then we can define a new vector $E'_m = E_m + E_1(i_m) + E_1(i_m \bmod Z)$. Because $L(S+E_m) = L_{N,k}(S) \geq N-Z \geq L(E_1(i_m)+E_1(i_m \bmod Z))$, we must have $L(S+E'_m) \leq L_{N,k}(S)$ by counting the multiplicity of 1's. Hence $L(S+E'_m) = L_{N,k}(S)$. Continuing this process, we can derive an error vector such that the support is contained [0, Z).

In particular, if $L(S) = N > L_{N,1}(S) \geq N - \min(2^v, r-2)$, then there exists $s>0$ such that $2^{s-1} < N - L_{N,1}(S) \leq 2^s$. From the previous discussion, $Z=2^s$ and there exists at least one error vector E_m such that $m \leq 1$ such that $L(S+E_m) = L_{N,1}(S)$ and $supp(E_m) \subseteq [0, 2^s)$. Because $L(S) > L_{N,1}(S)$, we must have $m=1$. Suppose there are E_1, E'_1 such that $L(S+E_1) = L(S+E'_1) = L_{N,1}(S)$ and $Supp(E_1), Supp(E'_1) \subseteq [0, 2^s)$. In this case, the multiplicity of root 1 in $S^N(x)+E_1(x)$ and $S^N(x)+E'_1(x)$ is greater than 2^{s-1} respectively, however the multiplicity of root 1 of the generating polynomial corresponding to $E_1+E'_1$ is not more than 2^{s-1}, a contradiction.

Because of the cyclic-shift invariant property of the (k-error) linear complexity of periodic sequence, we can extend the result to any $j \in [0,N)$. □

Theorem 3 says that if $2^v \geq r-1$ then $Z = 2^{\lceil \log_2(N-L_{N,k}(S)) \rceil} \leq 2^{\lceil \log_2(r-2) \rceil} < 2r$. On the other hand, if $2^v < r-1$ then Theorem 3 gives $Z \leq 2^v$ for any N-periodic binary sequence S such that $L(S) = N-c \geq L_{N,k}(S) \geq N-2^v$.

Example 3 Let S_1 be the following binary sequence with period 304 (= 16 $*$ 19).

S_1 = 11100011001111011010100011011100110000001100000001101110010100000
1101000110010010000100011001010010100110010111001001101011001000110100
0000100100001111001000101011000010010111110101001111110101101111000110
10011111011111101010111001000100010000110000111011111111111001011011011
01001010011010010001001101000.

The linear complexity is 304 and 1-error linear complexity is 301. The zone length is 4 and we have $L_{304,1}(S_1) = \mathcal{L}_{304,1;4}(S_1)$.

Example 4 Let S_2 be a random binary sequence with period 176 (= 11 $*$ 16),

S_2 = 111001111011001001101101111001001101010010101111111011011100010011
0001101001110010011101010100000100011111000001000011101011101101110011
10101000111010000100001101101001101110010.

Then $L(S_2) = L_1(S_2) = 175$, $L_2(S_2) = 169$, and the length of zone is 8. Indeed, we find an error vector of L_2 with two errors at positions 6 and 7 within the zone [0, 8).

We observe that the above result can be extended to any N-periodic binary sequence S such that $L(S) = N - c \geq L_{N,k}(S) = N - r + 1$ when $2^v < r - 1$. In this case, we can take $Z = r$.

Proposition 1 *Let $N = 2^v r$, $v > 0$, r be an odd prime, and 2 be a primitive root. If $2^v < r - 1$ and S is an N-periodic binary sequence such that $L(S) = N - c \geq L_{N,k}(S) = N - r + 1$ for some positive integer c, then there exists at least one error vector E_m, $m \leq k$ such that $L(S + E_m) = L_{N,k}(S)$ and $Supp(E_m) \subseteq [j, j + Z \pmod{N})$ for any $0 \leq j < N$, where $Z = r$. In particular, $L_{N,k}(S) = L_{N,k;Z,j}(S)$.*

Proof First, we consider $j = 0$. Let α be the primitive r-th root of unity. Since the multiplicity of 1 is at most $2^v < r - 1$ and $L_{N,k}(S) = N - r + 1$, there exists an error vector E_m, $m \leq k$ reaching $L_{N,k}$, such that the generating polynomial corresponding to $S + E_m$ will be divisible by $\Phi_r(X) = 1 + X + \cdots + X^{r-1}$. Suppose E_m has entry '1' at positions $i_1, i_2, \ldots, i_{m-1}, i_m$, where $i_1 < i_2 < \ldots < i_m$. If $i_m > r$, since the generating polynomial of $S + E_m + E_1(i_m) + E_1(i_m \bmod r)$ will be also divisible by $1 + X + \cdots + X^{r-1}$, implying $L\big(S + E_m + E_1(i_m) + E_1(i_m \bmod r)\big) \leq N - r + 1$. Moreover, $L(E_1(i_m) + E_1(i_m \bmod r) = N - r$ and $L_{N,k}(S) = L(S + E_m) \leq N - r + 1$ imply that $L\big(S + E_m + E_1(i_m) + E_1(i_m \bmod r)\big) = N - r + 1$. Thus we get an error vector $E_m + E_1(i_m) + E_1(i_m \bmod r)$ reaching $L_{N,k} = N - r + 1$. Consequently, we obtain an error vector reaching $L_{N,k}(S)$ with support in $[0, r)$. Finally, we extend to any $j \in [0, N)$ because of the cyclic-shift invariant property on the (k-error) linear complexity of periodic sequence. □

Remark 2 From Theorem 3 we obtain a small zone of length $Z \leq 2^v$ or $Z < 2r$ such that $L_{N,k}(S) = L_{N,k;Z,j}(S) = \mathcal{L}_{N,k;Z}(S)$ for the above classes of sequences with large linear complexity and k-error linear complexity. Our assumptions on these classes of sequences are not very restricted. By Artin's conjecture, there are approximately 37% of all primes having 2 as a primitive root [15]. By the following Corollary 2, we show that under certain conditions almost all random sequences have k-error linear complexity greater than or equal to $N - 2^v$ and about 50% of these sequences have linear complexity equal to the period. Therefore our result can be very useful to determine the stability of many random binary sequences with low computational cost.

Corollary 2 *Suppose r is a prime, 2 is a primitive root modulo r and $2^v \leq p(r)$, where $p(r)$ is a polynomial of the variable r. Let $N = 2^v r$. If*

$$\sum_{t=0}^{k} \binom{N}{t} < \frac{2^{r-1}}{2^v}, \tag{3}$$

for $0 \leq c < \min(2^v, r - 2)$, we have

$$\Pr(L(S) \geq N - c, L_{N,k}(S) \geq N - \min(2^v, r - 2)) \to 1 - 2^{-1-c}, asr \to \infty. \tag{4}$$

Proof For any positive intger k, we denote by $\mathcal{M}_{N,k}(c)$ the number of N-periodic sequences with the k-error linear complexity not more than c. Obviously,

$$\mathcal{M}_{N,k}(N-2^v-1) \le \min\left(2^N, \mathcal{M}_{N,0}(N-2^v-1)\sum_{t=0}^{k}\binom{N}{t}\right).$$

Then by Proposition 1 and Lemma 1 in [14] (page 2818), we have

$$\mathcal{M}_{N,0}(N-2^v-1) = 2^N - (2^{r-1}-1)^{2^v}\cdot 2^{2^v}.$$

If $\mathcal{M}_{N,0}(N-2^v-1)\sum_{t=0}^{k}\binom{N}{t} \le 2^N$, i. e.,

$$\sum_{t=0}^{k}\binom{N}{t} \le \frac{2^N}{\mathcal{M}_{N,0}(N-2^v-1)} = \frac{1}{1-(1-\frac{1}{2^{r-1}})^{2^v}}, \tag{5}$$

then we have $\mathcal{M}_{N,k}(N-2^v-1) \le \mathcal{M}_{N,0}(N-2^v-1)\sum_{t=0}^{k}\binom{N}{t}$.

Denote by ρ the ratio of the number of periodic sequences satisfying $L_{N,k}(S) \ge N-2^v$ over the number of all periodic sequences with period N. Hence

$$\begin{aligned}\rho &= \frac{2^N - \mathcal{M}_{N,k}(N-2^v-1)}{2^N}\\ &\ge 1-\frac{\mathcal{M}_{N,0}(N-2^v-1)}{2^N}\sum_{t=0}^{k}\binom{N}{t}\\ &= 1-\left(1-\left(1-\frac{1}{2^{r-1}}\right)^{2^v}\right)\sum_{t=0}^{k}\binom{N}{t}.\end{aligned} \tag{6}$$

Note that

$$\left(1-\frac{1}{2^{r-1}}\right)^{2^v} = \sum_{\substack{d=0\\ s=2d}}^{2^{v-1}-1}\left(\binom{2^v}{s}\cdot\frac{1}{2^{(r-1)\cdot s}} - \binom{2^v}{s+1}\cdot\frac{1}{2^{(r-1)\cdot(s+1)}}\right) + \frac{1}{2^{(r-1)\cdot 2^v}}.$$

For $d=0$ and $s=0$, we have $\binom{2^v}{s}\cdot\frac{1}{2^{(r-1)\cdot s}} - \binom{2^v}{s+1}\cdot\frac{1}{2^{(r-1)\cdot(s+1)}} = 1-\frac{2^v}{2^{r-1}}$. Then for $d>0$, we have $s\ge 2$ and

$$\frac{\binom{2^v}{s}\cdot\frac{1}{2^{(r-1)\cdot s}}}{\binom{2^v}{s+1}\cdot\frac{1}{2^{(r-1)\cdot(s+1)}}} = \frac{2^{r-1}(s+1)}{2^v-s} > \frac{2^r}{2^v} > 1.$$

The last inequality holds because we have $r>v$ by the assumption. Therefore, we obtain

$$\left(1-\frac{1}{2^{r-1}}\right)^{2^v} > 1-\frac{2^v}{2^{r-1}}. \tag{7}$$

If $\sum_{t=0}^{k}\binom{N}{t} < \frac{2^{r-1}}{2^v}$, then by (7) we have $\frac{2^{r-1}}{2^v} < \frac{1}{1-(1-\frac{1}{2^{r-1}})^{2^v}}$ and the condition (5) holds. Therefore, from (6) and (7) we derive

$$\rho > 1 - \frac{2^v}{2^{r-1}}\sum_{t=0}^{k}\binom{N}{t}.$$

For small k, we have $\sum_{t=0}^{k}\binom{N}{t} \le c_0N^k$ for some constant c_0, and thus $\frac{2^v}{2^{r-1}}\sum_{t=0}^{k}\binom{N}{t} \le 2^{1-r}c_0p(r)^{k+1}r^k)$. Hence we must have

$$\rho \to 1\, as\, r \to \infty.$$

This implies that almost all sequences of period 2^vr satisfy $L(S) \ge L_{N,k}(S) \ge N - 2^v$ as long as $r \to \infty$.

Next we prove that once $r \to \infty$, for $0 \le c < 2^v$ we have

$$\frac{|S\,:\,L(S) = N - c|}{|S\,:\,L(S) \ge N - 2^v|}\to\frac{1}{2^{c+1}}. \tag{8}$$

By the relationship between the linear complexity and Günther weight of the GDFT of sequences [14][p. 2818], and almost all sequences satisfy $L(S) \ge N - 2^v$, we only consider the first column of the GDFT matrix, and the contribution to the Günther weight of other columns are all 2^v. Additionally, the elements of the first column are over $\mathbb{F}_2$, and the pattern of the first column is the transpose of $\underbrace{0\ldots0}_{c}\ 1 *{*}*$, where '$*$' can be 0 or 1. Hence we have (8).

If $2^v \le r - 2$, then we obtain $\Pr(L(S) \ge N - c, L_{N,k}(S) \ge N - 2^v) = \Pr(L(S) \ge N - c, L(S) \ge N - 2^v) \to \frac{1}{2} + \cdots + \frac{1}{2^{c+1}} = 1 - 2^{-1-c},\ as\, r \to \infty$.

If $2^v > r - 2$, then almost all sequences of period 2^vr satisfy $L(S) \ge L_{N,k}(S) \ge N - 2^v$ and we obtain $Pr(L(S) \ge N - r + 2) \to 1 - 2^{1-r}$ similarly. Because

$$Pr(L_{N,k}(S) < N - r + 2) \le Pr(L(S) < N - r + 2)\sum_{t=0}^{k}\binom{N}{t},$$

for small k we have $Pr(L_{N,k}(S) < N - r + 2) = O(2^{1-r} \cdot N^k) = O(2^{1-r}p(r)^kr^k)$, then $Pr(L_{N,k}(S) \ge N - r + 2) \to 1$. Therefore, for $0 \le c < r - 2$ we obtain

$$\Pr(L(S) \ge N - c, L_{N,k}(S) \ge N - r + 2) \to 1 - 2^{-1-c},$$

analogously. □

Remark 3 According to the result in [18][p. 25; Theorem 1.2.8], for $0 < \frac{k}{N} \le 1/2$ we have

$$\sum_{t=0}^{k}\binom{N}{t} \le 2^{NH(\frac{k}{N})}, \tag{9}$$

where $H(\frac{k}{N}) := -\frac{k}{N}\log(\frac{k}{N}) - (1-\frac{k}{N})\log(1-\frac{k}{N})$ is the entropy function on the variable $\frac{k}{N}$, and the base of the $\log(\cdot)$ is 2.

Consequently, if k satisfies

$$NH\left(\frac{k}{N}\right) < r - 1 - v, \tag{10}$$

then the condition (3) holds by (9) and (10). Hence we have a weaker but explicit requirement of k for (4) holds. Note that the entrophy function $H(x)$ is non-decreasing when $0 < x \le 1/2$.

We provide the following example to demonstrate the usefulness of our results.

Example 5 Let $N = 30304 = 32 * 947$, i.e., $r = 947$, $v = 5$. Let us consider $k = 10$. The exhaustive search method of computing $L_{N,10}(S)$ is estimated as $\sum_{t=0}^{10}\binom{32*947}{t} > \binom{32*947}{10} \approx 2^{127}$, which is infeasible. But from (10) we require $H(\frac{k}{N}) < \frac{941}{32*947}$ and thus $k \le 96$. Hence the condition (10) holds for $k = 10$. From Corollary 2 we know that almost all sequences satisfy $L_{N,10} \ge N - 32$. We can take the zone length $Z = 32$.

If $c = 1$ and $L(S) \ge N - c$, then 3 out of 4 such random sequences satisfy the condition (2) in Theorem 3, and we have $L_{N,10}(S) = L_{N,10;32,0}(S)$. If $L(S) \ge N - 2$, then 7 out of 8 such random sequences satisfy (2), and we have $L_{N,10}(S) = L_{N,10;32,0}(S)$. The percentage of these sequences satisfying the condition (2) grows as c increases. Finally, if $L(S) \ge N - 31$, then almost any random sequence satisfy (2), and $L_{N,10}(S) = L_{N,10;32,0}(S)$.

Now we move to other types of sequences with period $N = 2^v r$, where r is a composite. More generally, we have

$$1 - X^N = (1 - X^r)^{2^v} = \left(\prod_{d|r}\Phi_d(X)\right)^{2^v}, \tag{11}$$

where $\Phi_d(X)$ is the d-th cyclotomic polynomial. We do not require that $\Phi_d(X)$ is irreducible over $\mathbb{F}_2$, which is required for existing fast algorithms of computing the (k-error) linear complexity. Then, we can similarly obtain the following result by analyzing the multiplicity of root 1 when the k-error linear complexity is large.

Theorem 4 *Suppose $N = 2^v p_1^{s_1} p_2^{s_2} \dots p_n^{s_n}$, $v > 0$, p_i is odd prime, and 2 is a primitive root modulo p_i^2 for all $1 \le i \le n$. For any N-periodic binary sequence S such that $L(S) > L_{N,k}(S) \ge N - \min(2^v, p_1 - 2, \dots, p_n - 2)$, and any $j \in [0, N)$ there exists a zone of length $Z = 2^{\lceil \log_2(N - L_{N,k}(S)) \rceil}$ such that $L_{N,k}(S) = L_{N,k;Z,j}(S) = \mathcal{L}_{N,k;Z}(S)$.*

Proof First, we suppose $N = 2^v \cdot p^s$, where $2^v < p - 1$ and 2 is a primitive root p^2. Then, 2 is a primitive root of p^s for any integer $s \ge 1$ (see [17, p. 348, Theorem 9. 10]). From (11) we derive

$$X^N - 1 = \left((X-1)\Phi_p(X)\Phi_{p^2}(X)\cdots\Phi_{p^s}(X)\right)^{2^v}.$$

Because the degree of each irreducible polynomial $\Phi_{p^i}(X)$ is $\phi(p^i) = p^i - p^{i-1} \geq p - 1$, we only need to consider the multiplicity of root 1 for estimating $L_{N,k}(S)$. The rest of proof is similar to the proof of Theorem 3.

Secondly, suppose $N = 2^v \cdot p \cdot q$, 2 is a primitive root modulo p and q. Let c be the least integer such that $2^c \equiv 1 \bmod (pq)$, then $\Phi_{pq}(X)$ can be factorized into $\frac{(p-1)(q-1)}{c}$ irreducible polynomials, each with degree c. In addition, because 2 is the primitive root modulo p, we have $(p-1) \mid c$ and thus $c \geq p - 1$. Similarly, $c \geq q - 1$. From (11) we derive $X^N - 1 = \left((X-1)\Phi_p(X)\Phi_q(X)\Phi_{pq}(X)\right)^{2^v}$, implying the degree of any irreducible factors except $X - 1$ is greater than or equal to $\min(p-1, q-1)$. Hence we only need to consider the multiplicity of root 1. The rest of proof follows.

Finally, if 2 is the primitive root of $p_1^{s_1}, \cdots, p_n^{s_n}$ for any integers $i_1, \dots, i_n$, then we obtain $X^N - 1 = \left((X-1)\prod_{\substack{0 \leq t_j \leq s_j \\ j = 1, 2, \dots, n}} \Phi_{p_1^{t_1}\cdots p_j^{t_j}\cdots p_n^{t_n}}(X)\right)^{2^v}$. Similarly, the degree of each irreducible factor of $\Phi_{p_1^{t_1}\cdots p_j^{t_j}\cdots p_n^{t_n}}(X)$ is no less than $\min(p_1 - 1, \dots, p_n - 1)$. Hence we only need to consider the multiplicity of the root 1 analogously. □

Remark 4 Actually, there is a large proportion of N satisfying Theorem 4. Heuristically, if 2 is a primitive root modulo a prime p, then 2 is also most likely a primitive root modulo p^2. Indeed, because $2^{p-1} \equiv 1 \pmod{p}$ and $2^{p(p-1)} \equiv 1 \pmod{p^2}$, 2 is a primitive root modulo p^2 if $2^{(p-1)} \not\equiv 1 \pmod{p^2}$. However, the prime satisfying $2^{(p-1)} \equiv 1 \pmod{p^2}$ is called *Wieferich Prime*, and the only known Wieferich Prime up to now is 1093 and 3511 (https://en.wikipedia.org/wiki/Wieferich_prime). In this case, 2 is not a primitive root of 1093 and 3511. It has been conjectured that infinitely many Wieferich primes exist, and that the number of Wieferich primes below x is approximately $\log(\log(x))$. By Artin's conjecture, 2 is the primitive root of 37% prime numbers. These prime numbers are most likely not Wieferich primes and therefore satisfy the conditions of Theorem 4.

Now let $N = 2^v M$, $v > 0$ and M be odd. If M has the factorization $p_1^{s_1}$ for certain prime number p_1, then the probability that p_1 satisfies the condition of Theorem 4 is a little bit over 1/3. Similarly, If M has a factorization $p_1^{s_1} \cdots p_r^{s_r}$, then it has at least $(1/3)^r$ probability satisfying the condition. As a result, M has the probability at least $1/3 + \cdots + (1/3)^r + \cdots = 1/2$ satisfying the condition of Theorem 4.

We don't estimate the probability of the condition $(L(S) > L_{N,k}(S) \geq N - \min(2^v, p_1 - 2, \dots, p_n - 2))$ since it becomes complicated when the number of prime factors increases, but we experimentally observe the property for randomly generated sequences of such periods from time to time.

Example 6 From computer experiments, there are many examples of period N satisfying Theorem 4. For example, $N = 1184\ (= 32 * 37), 484\ (= 4 * 11^2), 968\ (= 8 * 11^2), 1144\ (= 8 * 11 * 13)$;

$1936\ (= 16 * 11^2), 2288\ (= 16 * 11 * 13)$; $3344\ (= 11 * 16 * 19), 3952\ (= 13 * 16 * 19)$. The global stability can be effectively determined by local stability within a much smaller zone. For example, for the random generated $N = 4 * 11^2$-periodic sequence

$$
\begin{aligned}
S = {} & 010110001110010010101001101011100110010010001110011100010110001000\\
& 0010100100100101001000011110000101011100011110000000111010011001101110\\
& 0110101110100010110111000011101010100101101100001000101001111010010110\\
& 1110000000011010111010000000100100001011011100000100111001110101110111\\
& 0010101000100111010010100011110011110100110110001000110000010001110100\\
& 0110101010110001000100000000000000101101100111100001101000010011110000\\
& 001011100000000100000111011001100110100111101000000100100100110000100,
\end{aligned}
$$

we have $L(S) = L_1(S) = 482$, $L_2(S) = 480$, and the zone is $[0, 4)$. Indeed, we can find an error vector E_2 of L_2 such that error positions are 1 and 3 within the zone $[0, 4)$.

Acknowledgements We thank the anonymous referees and editors for their helpful suggestions. Ming Su expresses his sincere thanks for the hospitality during his visit to School of Mathematics and Statistics, Carleton University, Canada.

References

1. Alecu, A., Sălăgean, A.: An approximation algorithm for computing the k-error linear complexity of sequences using the discrete fourier transform. In: 2008 IEEE International Symposium on Information Theory, ISIT 2008, Toronto, ON, Canada, pp. 2414–2418 (2008)
2. Chen, L., Gong, G.: Communication System Security. CRC Press, Boca Raton (2012)
3. Cusick, T.W., Ding, C., Renvall, A.: Stream Ciphers and Number Theory. Elsevier, Amsterdam (1998)
4. Ding, C., Xiao, G., Shan, W.: The Stability Theory of Stream Ciphers. Springer, Heidelberg (1991)
5. Etzion, T., Kalouptsidis, N., Kolokotronis, N., Limniotis, K., Paterson, K.G.: Properties of the error linear complexity spectrum. IEEE Trans. Inf. Theory **55**(10), 4681–4686 (2009)
6. Fu, F., Niederreiter, H., Su, M.: The characterization of 2^n-periodic binary sequences with fixed 1-error linear complexity. In: Sequences and their Applications - SETA 2006, 4th International Conference, Beijing, China, Proceedings. pp. 88–103 (2006)
7. Games, R.A., Chan, A.H.: A fast algorithm for determining the complexity of a binary sequence with period 2^n. IEEE Trans. Inf. Theory **29**(1), 144–146 (1983)
8. Han, Y.K., Chung, J., Yang, K.: On the k-error linear complexity of p^m-periodic binary sequences. IEEE Trans. Inf. Theory **53**(6), 2297–2304 (2007)

9. Kaida, T., Uehara, S., Imamura, K.: An algorithm for the k-error linear complexity of sequences over $GF(p^m)$ with period p^n, p a prime. Inf. Comput. **151**(1–2), 134–147 (1999)
10. Lauder, A.G.B., Paterson, K.G.: Computing the error linear complexity spectrum of a binary sequence of period 2^n. IEEE Trans. Inf. Theory **49**(1), 273–280 (2003)
11. Massey, J.L.: Shift-register synthesis and BCH decoding. IEEE Trans. Inf. Theory **15**(1), 122–127 (1969)
12. Meidl, W.: On the stability of 2^n-periodic binary sequences. IEEE Trans. Inf. Theory **51**(3), 1151–1155 (2005)
13. Meidl, W.: Reducing the calculation of the linear complexity of u 2 ^{v}-periodic binary sequences to Games–Chan algorithm. Des. Codes Cryptogr. **46**(1), 57–65 (2008)
14. Meidl, W., Niederreiter, H.: On the expected value of the linear complexity and the k-error linear complexity of periodic sequences. IEEE Trans. Inf. Theory **48**(11), 2817–2825 (2002)
15. Moree, P.: Artin's primitive root conjecture—a survey. Integers **12**(6), 1305–1416 (2012)
16. Niederreiter, H.: Periodic sequences with large k-error linear complexity. IEEE Trans. Inf. Theory **49**(2), 501–505 (2003)
17. Rosen, K.H.: Elementary Number Theory and its Applications, 5th edn. Pearson, London (2005)
18. Roman, S.: Coding and Information Theory. Springer, Berlin (1992)
19. Sălăgean, A., Alecu, A.: An improved approximation algorithm for computing the k-error linear complexity of sequences using the discrete Fourier transform. In: Sequences and their Applications-SETA 2010-6th International Conference, Paris, France, 2010. Proceedings. pp. 151–165 (2010)
20. Stamp, M., Martin, C.F.: An algorithm for the k-error linear complexity of binary sequences with period 2^n. IEEE Trans. Inf. Theory **39**, 1398–1401 (1993)
21. Su, M.: The distribution of complexity measurments for periodic sequences. Ph. D. dissertation in Nankai University (2004)
22. Su, M.: Decomposing approach for error vectors of k-error linear complexity of 2^n-periodic binary sequences. In: WCC(Workshop on Coding and Cryptography) preproceeding, Ullensevang, Norway, May, 2009, preproceedings. pp. 399–415 (2009)
23. Su, M.: Decomposing approach for error vectors of k-error linear complexity of certain periodic sequences. IEICE Trans **97**(7), 1542–1555 (2014)
24. Su, M., Chen, L.: The properties of the 1-error linear complexity of p^n-periodic sequences over $\mathbb{F}_p$. In: IEEE International Symposium on Information Theory 2006, Seattle, USA, July 9-14, 2006. pp. 1998–2002. IEEE (2006)
25. Wei, S., Xiao, G., Chen, Z.: A fast algorithm for determining the minimal polynomial where of a sequence with period $2p^n$ over $GF(q)$. IEEE Trans. Inf. Theory **48**(10), 2754–2758 (2002)
26. Xiao, G., Wei, S., Lam, K., Imamura, K.: A fast algorithm for determining the linear complexity of a sequence with period p^n over $GF(q)$. IEEE Trans. Inf. Theory **46**(6), 2203–2206 (2000)
27. Zhou, J.: On the k-error linear complexity of sequences with period $2p^n$ over $GF(q)$. Des. Codes Cryptogr. **58**(3), 279–296 (2011)