Counting Functions for the k-Error Linear Complexity of 2^n-Periodic Binary Sequences*

Ramakanth Kavuluru and Andrew Klapper

Department of Computer Science, University of Kentucky,
Lexington, KY 40506, USA
{rvkavu2,klapper}@cs.uky.edu

Abstract. Linear complexity is an important measure of the cryptographic strength of key streams used in stream ciphers. The linear complexity of a sequence can decrease drastically when a few symbols are changed. Hence there has been considerable interest in the k-error linear complexity of sequences which measures this instability in linear complexity. For 2^n-periodic sequences it is known that minimum number of changes needed per period to lower the linear complexity is the same for sequences with fixed linear complexity. In this paper we derive an expression to enumerate all possible values for the k-error linear complexity of 2^n-periodic binary sequences with fixed linear complexity L, when k equals the minimum number of changes needed to lower the linear complexity below L. For some of these values we derive the expression for the corresponding number of 2^n-periodic binary sequences with fixed linear complexity and k-error linear complexity when k equals the minimum number of changes needed to lower the linear complexity. These results are of importance to compute some statistical properties concerning the stability of linear complexity of 2^n-periodic binary sequences.

Keywords: Periodic sequence, linear complexity, k-error linear complexity.

1 Introduction

The linear complexity of a sequence is the length of the shortest linear feedback shift register (LFSR) that can generate the sequence. The LFSR that generates a given sequence can be determined using the Berlekamp-Massey algorithm using only the first $2L$ elements of the sequence, where L is the linear complexity of the sequence. The typical assumption in the analysis of the security of stream ciphers is that the attacker has access to a part of the key stream and wants to use this to predict the remainder of the key stream. Thus the problem of designing a good stream cipher is reduced to the problem of designing a fast key stream

* This material is based upon work supported by the National Science Foundation under Grant No. CCF-0514660. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

R. Avanzi, L. Keliher, and F. Sica (Eds.): SAC 2008, LNCS 5381, pp. 151–164, 2009.

generator whose output is hard to predict from a prefix of the the output. Hence for cryptographic purposes sequences with high linear complexity are essential as an adversary would then need large initial segments of the sequences to recover the LFSRs that generate them using the Berlekamp-Massey algorithm.

A system is insecure if all but a few symbols of the key stream can be extracted. So for a cryptographically strong sequence, the linear complexity should not decrease drastically if a few symbols are changed. If it did, an attacker could modify the known prefix of the key stream and try to decrypt the result using the Berlekamp-Massey algorithm. If the resulting sequence differed from the actual key stream by only a few symbols, the attacker could extract most of the message. This observation gives rise to k-error linear complexity of sequences introduced in [7]. The k-error linear complexity of a periodic sequence is the smallest linear complexity achieved by making k or fewer changes per period. In addition to having large linear complexity, cryptographically strong sequences should, thus, also have large k-error linear complexity at least for small k.

Let $\mathbf{S} = (s_0, s_1, \cdots, s_{T-1})^\infty$ be a periodic binary sequence with period T. We associate the polynomial $\mathbf{S}(x) = s_0 + s_1 x + \cdots + s_{T-1}x^{T-1}$ and the corresponding T-tuple $\mathbf{S}^{(T)} = (s_0, s_1, \cdots, s_{T-1})$ to $\mathbf{S}$. The relationship between the linear complexity, denoted $L(\mathbf{S})$, of $\mathbf{S}$ and the associated polynomial $\mathbf{S}(x)$ is given by

$$L(\mathbf{S}) = T - \deg(\gcd(x^T - 1, \mathbf{S}(x))). \tag{1}$$

Let $w_H(\mathbf{S})$ denote the Hamming weight of the T-tuple $\mathbf{S}^{(T)}$. For $0 \leq k \leq T$, the k-error linear complexity of $\mathbf{S}$, denoted $L_k(\mathbf{S})$, is given by

$$L_k(\mathbf{S}) = \min_{\mathbf{E}} L(\mathbf{S} + \mathbf{E}), \tag{2}$$

where the minimum is over all T-periodic binary sequences $\mathbf{E}$ such that $w_H(\mathbf{E}) \leq k$. Since we consider only 2^n-periodic sequences, we use $T = 2^n$ and the observation

$$x^T - 1 = x^{2^n} - 1 = (x-1)^{2^n} \tag{3}$$

for the rest of the paper.

Let $merr(\mathbf{S})$ denote the minimum value k such that the k-error linear complexity of a 2^n-periodic sequence $\mathbf{S}$ is strictly less than its linear complexity. That is

$$merr(\mathbf{S}) = \min\{k : L_k(\mathbf{S}) < L(\mathbf{S})\}. \tag{4}$$

Kurosawa et al. [3] derived the formula for the exact value of $merr(\mathbf{S})$.

Lemma 1. *For any nonzero 2^n-periodic sequence $\mathbf{S}$, we have*

$$merr(\mathbf{S}) = 2^{w_H(2^n - L(\mathbf{S}))},$$

where $w_H(j)$, $0 \leq j \leq 2^n - 1$, denotes the Hamming weight of the binary representation of j.

The counting function of a sequence complexity measure gives the number of sequences with a given complexity measure value. Rueppel [6] determined the

counting function of linear complexity for 2^n-periodic binary sequences. Using equations (1) and (3) it is straightforward to obtain the number of 2^n-periodic binary sequences with fixed linear complexity. For the rest of the paper let $\mathcal{N}(L)$ and $\mathcal{A}(L)$ denote, respectively, the number of and the set of 2^n-periodic binary sequences with given linear complexity L, $0 \le L \le 2^n$. Rueppel [6] showed that

$$\mathcal{N}(0) = 1 \text{ and } \mathcal{N}(L) = 2^{L-1} \text{ for } 1 \le L \le 2^n. \tag{5}$$

Recently, using efficient algorithms to compute the linear complexity of p^n periodic sequences over $\mathbb{F}_p$, Meidl [4] obtained the counting function and the expected value for the 1-error linear complexity of 2^n-periodic binary sequences. Meidl and Venkateswarlu [5] extended these results to p^n-periodic sequences over $\mathbb{F}_p$. Fengxiang and Wenfeng [1] used Meidl's [4] approach of analyzing Games-Chan algorithm to obtain the counting functions and gave the exact expression for the expected value of the 2-error linear complexity of a random 2^n-periodic binary sequence with linear complexity $2^n - 1$.

In this paper we perform a more rigorous analysis of Games-Chan algorithm to enumerate all the possible values of k-error linear complexity of sequences in $\mathcal{A}(L)$ for $k = 2^{w_H(2^n - L)}$, that is when k is the minimum number of changes needed to lower the linear complexity below L. For certain sets of these values, we also derive the corresponding number of sequences in $\mathcal{A}(L)$ whose k-error linear complexity equals the values in those sets. For the rest of the paper by $k_{min}(L)$ denote the minimum number of changes needed to lower the linear complexity of sequences in $\mathcal{A}(L)$, that is $k_{min}(L) = 2^{w_H(2^n - L)}$.

2 Games-Chan Algorithm

In this section we describe the Games-Chan algorithm and list some results using its analysis.

By Lemma 1 for any 2^n-periodic sequence $\mathbf{S}$ with $merr(\mathbf{S}) = 2^m$, $m \in \{0, \cdots, n\}$, the linear complexity $L(\mathbf{S})$ can be uniquely expressed as

$$L(\mathbf{S}) = 2^n \text{ or } L(\mathbf{S}) = 2^n - \sum_{i=1}^{m} 2^{n-r_i},$$

where $0 < r_1 < \cdots < r_m \le n$.

The Games-Chan algorithm [2] is a fast algorithm for computing the linear complexity of a 2^n-periodic binary sequence. For any $\mathbf{S} \in \mathcal{A}(L)$ with period $\mathbf{S}^{(2^n)} = (s_0, \cdots, s_{2^n-1})$, denote the left and right halves of $\mathbf{S}^{(2^n)}$ by

$$\mathbf{S}_L^{(2^{n-1})} = (s_0, \cdots, s_{2^{n-1}-1}) \quad \text{and} \quad \mathbf{S}_R^{(2^{n-1})} = (s_{2^{n-1}}, \cdots, s_{2^n-1}).$$

Let $\mathbf{S}_L$ and $\mathbf{S}_R$ denote the 2^{n-1} periodic sequences

$$\mathbf{S}_L = (s_0, \cdots, s_{2^{n-1}-1})^\infty \quad \text{and} \quad \mathbf{S}_R = (s_{2^{n-1}}, \cdots, s_{2^n-1})^\infty. \tag{6}$$

Games-Chan Algorithm. Let $\mathbf{S}$ be 2^n-periodic binary sequence.

(i) If $\mathbf{S}_L^{(2^{n-1})} = \mathbf{S}_R^{(2^{n-1})}$, then $L(\mathbf{S}) = L(\mathbf{S}_L)$.
(ii) If $\mathbf{S}_L^{(2^{n-1})} \neq \mathbf{S}_R^{(2^{n-1})}$, then $L(\mathbf{S}) = 2^{n-1} + L(\mathbf{S}_L + \mathbf{S}_R)$.
(iii) Apply the above procedure recursively to the 2^{n-1}-periodic binary sequence $\mathbf{S}_L$ in (i), or the 2^{n-1}-periodic binary sequence $\mathbf{S}_L + \mathbf{S}_R$ in (ii).

We make some observations and establish notation we use for the rest of the paper. We note that the procedure of the Games-Chan algorithm as stated here is executed a total of n times to compute the linear complexity of any $\mathbf{S} \in \mathcal{A}(L)$. In the ith step, $i = 0, \cdots, n-1$, the algorithm computes the linear complexity of a 2^{n-i}-periodic binary sequence. Let $\psi^i(\mathbf{S})$, $i = 0, \cdots, n-1$, denote the first period of the 2^{n-i}-periodic binary sequence considered in the ith step of the algorithm when run with input sequence $\mathbf{S}$. Let $\psi_L^i(\mathbf{S})$ and $\psi_R^i(\mathbf{S})$ denote, respectively, the left and right halves of $\psi^i(\mathbf{S})$. Let $m^i(\mathbf{S})$ denote the total value contributed to $L(\mathbf{S})$ in the algorithm during the execution from the 0-th step to the i-th step of the algorithm. For any two finite binary sequences, $\mathbf{S}$ and $\mathbf{S}'$, of same length let $d_H(\mathbf{S}, \mathbf{S}')$ denote the Hamming distance between $\mathbf{S}$ and $\mathbf{S}'$. We slightly abuse the notation because we also use $d_H(\mathbf{S}, \mathbf{S}')$ to denote the Hamming distance between the first periods of $\mathbf{S}, \mathbf{S}' \in A(L)$. It is straightforward to derive the following lemma from the Games-Chan algorithm.

Lemma 2. *Let $\mathbf{S}$ be a 2^n-periodic binary sequence. For any t integers $r_1, \cdots, r_t$ such that $0 < r_1 < r_2 < \cdots < r_t \leq n$, we have*

$$L(\mathbf{S}) = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}) \tag{7}$$

if and only if

$$\psi_L^{u-1}(\mathbf{S}) = \psi_R^{u-1}(\mathbf{S}) \quad \textit{exactly when} \quad u \in \{r_1, \cdots, r_t\}.$$

For any $\mathbf{S} \in \mathcal{A}(L)$ where L is as in equation (7), the following properties of vectors $\psi^l(\mathbf{S})$, $0 \leq l \leq n$, are straightforward to obtain.

P1: If $l = r_i - 1$, for some $i \in \{1, \cdots, t\}$, then $w_H(\psi^l(\mathbf{S})) = 2 \cdot w_H(\psi^{l+1}(\mathbf{S}))$.
P2: For any $l \neq r_i - 1$, for all $i \in \{1, \cdots, t\}$, we have $w_H(\psi^l(\mathbf{S})) \geq w_H(\psi^{l+1}(\mathbf{S}))$.

By $\mathcal{P}_l$, $0 \leq l \leq n$, denote the number of distinct possibilities, over all sequences in $\mathcal{A}(L)$, for the 2^{n-l}-vector during the l-th step such that the 2^{n-l-1}-vector during the $(l+1)$-th step is fixed. It is straightforward to get the following properties.

P3: If $l = r_i - 1$, for some $i \in \{1, \cdots, t\}$, then $\mathcal{P}_l = 1$.
P4: For any $l \neq r_i - 1$, for all $i \in \{1, \cdots, t\}$, we have $\mathcal{P}_l = 2^{2^{n-l-1}}$.

We also use the following result in the next section. It can be proved using the procedure of Games-Chan algorithm and Lemma 2.

Lemma 3. *Let $\mathbf{S} \in \mathcal{A}(L)$ with $L \neq 0$ represented as*

$$L(\mathbf{S}) = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{8}$$

where $0 < r_1 < r_2 < \cdots < r_t \leq n$. *Let* $\mathbf{S}' \neq \mathbf{S}$ *be any other* 2^n*-periodic binary sequence such that* $m^{l-1}(\mathbf{S}) = m^{l-1}(\mathbf{S}')$ *for some* $l \in \{1, \cdots, n\}$. *If* $d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')) \neq 0$, *then*

$$d_H(\mathbf{S}, \mathbf{S}') \geq 2^b \cdot d_H(\psi^l(\mathbf{S}), \psi^l(\mathbf{S}')), \tag{9}$$

where b, $1 \leq b \leq t$, *is the unique integer determined by the inequality* $r_b \leq l < r_{b+1}$ *assuming* $r_0 = 0$ *and* $r_{t+1} = n + 1$.

3 Expression for $k_{min}(L)$-Error Linear Complexity

In this section we analyze the structure of the Games-Chan algorithm to derive an expression to enumerate all possible values of $k_{min}(L)$-error linear complexity of sequences in $\mathcal{A}(L)$ in terms the coefficients in the binary expansion of $2^n - L$. We handle the case when $1 < L < 2^n$ as the results are simple when $L = 0$ or 1 and as we already know the results when $L = 2^n$ [4]. We need the following generalization of [1, Lemma 2] whose proof is similar to that of Lemma 2 in [1].

Lemma 4. *For any sequence* $\mathbf{S} = (s_0, \cdots, s_{2^n-1})^\infty \in \mathcal{A}(L)$, *we have* $L \leq 2^n - 2^{n-r}$, $r = 1, \cdots, n$, *if and only if*

$$\sum_{i=0}^{2^r-1} s_{j+i \cdot 2^{n-r}} = 0 \text{ for } j = 0, \cdots, 2^{n-r} - 1.$$

We prove an auxiliary result that is used in the main result of this section.

Lemma 5. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *with* $1 < L < 2^n$. *Consider the representation of* L *as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{10}$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n + 1 = r_{t+1}$ *and* $1 \leq t \leq n - 1$. *Let* $\mathbf{S}'$ *be any* 2^n*-periodic binary sequence such that* $d_H(\mathbf{S}, \mathbf{S}') = k_{min}(L) = 2^t$ *and* $L(\mathbf{S}') = L_{2^t}(\mathbf{S})$. *Define the two integers*

$$l_1 = \min\{i : 0 \leq i \leq n - 1 \quad \textit{and} \quad m^i(\mathbf{S}') \neq m^i(\mathbf{S})\} \tag{11}$$

and

$$\begin{aligned} l_2 = \min\{i : 0 \leq i \leq n - 1 \quad & \textit{and} \quad d_H(\psi_L^i(\mathbf{S}), \psi_R^i(\mathbf{S})) = 2^{t-j} \\ & \textit{with} \quad r_j \leq i < r_{j+1}\}. \end{aligned} \tag{12}$$

Then we have $l_1 = l_2$.

Proof. From Lemma 1 we know $k_{min}(L) = 2^t$ which implies $L(\mathbf{S}') < L(\mathbf{S})$. We note that there exists at least one integer i, $0 \leq i \leq n - 1$, such that $m^i(\mathbf{S}') \neq m^i(\mathbf{S})$ since otherwise $L(\mathbf{S}) = L(\mathbf{S}')$. Hence the set on the right hand

side of equation (11) is not empty. From the procedure of the Games-Chan algorithm and using the fact $L(\mathbf{S}') < L(\mathbf{S})$ equation (11) implies

$$\psi_L^{l_1}(\mathbf{S}) \neq \psi_R^{l_1}(\mathbf{S}) \quad \text{and} \quad \psi_L^{l_1}(\mathbf{S}') = \psi_R^{l_1}(\mathbf{S}'). \tag{13}$$

From equation (13) we get

$$d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) \geq d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})). \tag{14}$$

Let b be the unique integer determined by the inequality $r_b \leq l_1 < r_{b+1}$. Since $\psi_L^{r_t-1}(\mathbf{S}) = \psi_R^{r_t-1}(\mathbf{S})$ and because the vectors considered during all the steps, except the last one, of the Games-Chan algorithm have nonzero Hamming weight, we have $w_H(\psi^{r_t-1}(\mathbf{S})) \geq 2$. So using properties P1 and P2 we get $w_H(\psi^{l_1+1}(\mathbf{S})) \geq 2^{t-b}$ and thus

$$d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) \geq 2^{t-b}. \tag{15}$$

Now we show that $d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) = 2^{t-b}$. If not, from equation (15) we have $d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) > 2^{t-b}$. By equation (14) this implies

$$d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) > 2^{t-b}. \tag{16}$$

But from Lemma 3 we know $d_H(\mathbf{S}, \mathbf{S}') \geq 2^b \cdot d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}'))$, which implies $d_H(\psi^{l_1}(\mathbf{S}), \psi^{l_1}(\mathbf{S}')) \leq 2^{t-b}$ since $d_H(\mathbf{S}, \mathbf{S}') = 2^t$. This contradicts inequality in (16). Thus we have

$$d_H(\psi_L^{l_1}(\mathbf{S}), \psi_R^{l_1}(\mathbf{S})) = 2^{t-b}. \tag{17}$$

From equation (17) we know that the set on the right hand side of equation (12) is not empty and $l_2 \leq l_1$. By a denote the unique integer determined by the inequality $r_a \leq l_2 < r_{a+1}$. Because there are a steps before the l_2-th step where the left and right halves are equal it is evident from equation (21) that altering $\psi^{l_2}(\mathbf{S})$ such that $\psi_L^{l_2}(\mathbf{S}) = \psi_R^{l_2}(\mathbf{S})$ and propagating these changes to the 0-th step of the Games-Chan algorithm will require exactly $2^a \cdot 2^{t-a} = 2^t$ changes in $\mathbf{S}^{(2^n)}$. But if $l_2 < l_1$, forcing $\psi_L^{l_2}(\mathbf{S}) = \psi_R^{l_2}(\mathbf{S})$ will result in a 2^n-periodic binary sequence $\mathbf{S}''$ such that $d_H(\mathbf{S}, \mathbf{S}'') = 2^t$ and $L(\mathbf{S}'') < L(\mathbf{S}')$. This contradicts the fact that $L(\mathbf{S}') = L_{2^t}(\mathbf{S})$. Thus we have $l_2 = l_1$. □

Theorem 1. *Let $\mathbf{S} \in \mathcal{A}(L)$ with $1 < L < 2^n$. Consider the representation of L as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}), \tag{18}$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ and $1 \leq t \leq n-1$. Define the integer $w = \min\{i : r_i = n + i - t, 1 \leq i \leq t+1\}$. Then $L_{k_{min}(L)}(\mathbf{S})$ is 0 or is in one of the two forms

$$\begin{aligned} L_{j,l,C} \quad := \quad & 2^n - \sum_{i=1}^{j-1} 2^{n-r_i} - 2^{n-l} + C, \quad 1 \leq j \leq w-1, \\ & r_{j-1} \leq l \leq r_j - 2, \quad \text{and} \quad 1 \leq C \leq 2^{n-l-1} - 1, \end{aligned} \tag{19}$$

or

$$\begin{aligned} L_{w,l,C} \quad &:= \quad 2^n - \sum_{i=1}^{w-1} 2^{n-r_i} - 2^{n-l} + C, \\ & \qquad r_{w-1} \le l \le r_w - 3 \quad \textit{and} \quad 1 \le C \le 2^{n-l-1} - 2^{t-w+1}. \end{aligned} \tag{20}$$

Proof. From Lemma 1 and equation (18) $merr(\mathbf{S}) = k_{min}(L) = 2^t$. The sequences in $\mathcal{A}(L)$ whose 2^t-error linear complexity is 0 are those with exactly 2^t 1s per period. For any other sequence $\mathbf{S}$ in $\mathcal{A}(L)$ we show that the 2^t-error linear complexity is in one of the forms as stated in the theorem.

Define the integer l as in equation (12). That is

$$\begin{aligned} l = \min\{i : 0 \le i \le n-1 \quad &\text{and} \quad d_H(\psi_L^i(\mathbf{S}), \psi_R^i(\mathbf{S})) = 2^{t-j} \\ &\text{with} \quad r_j \le i < r_{j+1}\}. \end{aligned} \tag{21}$$

We already know that the set on the right hand side of equation (21) is not empty due to the intermediate findings of Lemma 5. By b denote the unique integer determined by the inequality $r_b \le l < r_{b+1}$. From the proof of Lemma 5 we know that altering $\psi^l(\mathbf{S})$ such that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ and propagating these changes to the 0-th step of the Games-Chan algorithm will require exactly 2^t changes in $\mathbf{S}^{(2^n)}$. We also see that it is necessary to alter $\psi^l(\mathbf{S})$ so that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ to achieve the smallest linear complexity that can be obtained by making exactly 2^t errors in $\mathbf{S}^{(2^n)}$ since the remaining $n-l$ steps can only add a maximum of 2^{n-l-1} to the linear complexity of the modified sequence.

Note that $l \ne r_j - 1$, $j = 1, \cdots, t$, since $\psi_L^{r_j-1}(\mathbf{S}) = \psi_R^{r_j-1}(\mathbf{S})$, $j = 1, \cdots, t$. Next we show that

$$\forall \quad l+1 \le i \le n-1, \quad w_H(\psi^i(\mathbf{S})) = 2^{t-j} \quad \text{with} \quad r_j \le i < r_{j+1}. \tag{22}$$

If equation (22) does not hold, then let m be any integer such that $l+1 \le m \le n-1$ and $w_H(\psi^m(\mathbf{S})) \ne 2^{t-a}$ where a is uniquely determined by the inequality $r_a \le m < r_{a+1}$. Since $\psi_L^{r_t-1}(\mathbf{S}) = \psi_R^{r_t-1}(\mathbf{S})$, we have $w_H(\psi^{r_t-1}(\mathbf{S})) \ge 2$. So using properties P1 and P2 we get $w_H(\psi^m(\mathbf{S})) \ge 2^{t-a}$. This implies $w_H(\psi^m(\mathbf{S})) > 2^{t-a}$ since we assumed $w_H(\psi^m(\mathbf{S})) \ne 2^{t-a}$. Again, using P1 and P2 we have $w_H(\psi^{l+1}(\mathbf{S})) > 2^{a-b} \cdot 2^{t-a} = 2^{t-b}$ which contradicts the fact $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-b}$. Thus $w_H(\psi^m(\mathbf{S})) = 2^{t-a}$ and so equation (22) holds.

To obtain the form of $L_{2^t}(\mathbf{S})$ we consider two cases based on the value of w.

Case 1: $w \le t$

From the definition of w in the theorem statement it can be shown that $n - r_i = t - i$ for $i = w, \cdots, t$, which implies

$$L = 2^n - (2^{n-r_1} + \cdots + 2^{n-r_{w-1}} + 2^{t-w} + 2^{t-w-1} + \cdots + 2^0). \tag{23}$$

From equations (18), (23) and Lemma 2 this means that the left and right halves are equal from the $(r_w - 1)$-th step to $(n-1)$-th step of the execution of the Games-Chan algorithm. Using the fact that $n - r_w = t - w$, this implies that the 2^{t-w+1}-vector considered during the $(r_w - 1)$-th step

$$\psi^{r_w-1}(\mathbf{S}) = (\psi^{r_w-1}(\mathbf{S})_0, \cdots, \psi^{r_w-1}(\mathbf{S})_{2^{t-w+1}-1}) = (1, \cdots, 1) \tag{24}$$

is an all 1 vector. From the definition of w, equation (24) also implies that $w_H(\psi^{r_w-2}(\mathbf{S})) = 2^{t-w+1}$. That is

$$d_H(\psi_L^{r_w-3}(\mathbf{S}), \psi_R^{r_w-3}(\mathbf{S})) = 2^{t-w+1}. \tag{25}$$

By equation (25) and using the definition of l in equation (21) we have $l \leq r_w - 3$. We consider two cases based on the value of l.

Case 1a: $r_{w-1} \leq l \leq r_w - 3$

We first note that this case occurs only when the binary expansion of L as in equation (18) satisfies $r_{w-1} \leq r_w - 3$. Throughout this case we use the fact that $n - r_w = t - w$. From the definition of l in equation (21) we have $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-w+1}$. We already know that making 2^{t-w+1} changes in $\psi^l(\mathbf{S})$ so that $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ is necessary to achieve the the smallest linear complexity possible by making $k_{min}(L) = 2^t$ changes in $\mathbf{S}^{(2^n)}$. But we have to decide for each of the 2^{t-w+1} positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ, whether the change should be made in $\psi_L^l(\mathbf{S})$ or at the corresponding position in $\psi_R^l(\mathbf{S})$. In this case there is a unique of making these 2^{t-w+1} changes so that the linear complexity of the 2^{n-l-1}-periodic binary sequence with period equal to either of the equal halves obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ is as small as possible. Next we describe a unique way of making these changes.

Let $\psi^{l+1}(\mathbf{S}') = \psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ be the 2^{n-l-1}-vector obtained after forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ such that the linear complexity of the 2^{n-l-1}-periodic binary sequence with period $\psi^{l+1}(\mathbf{S}')$ is as small as possible. The left and right halves of the vectors considered are not equal from the r_{w-1}-th step to the $(r_w - 2)$-th step of the Games-Chan algorithm when executed with input sequence $\mathbf{S}$. From equation (24) $\psi^{r_w-1}(\mathbf{S})$ is a 2^{t-w+1}-vector with all 1s. Hence for all $v = r_{w-1}, r_{w-1}+1, \cdots, r_w - 2$ due to the procedure of the Games-Chan algorithm we have

$$\sum_{j=0}^{2^{r_w-v-1}-1} \psi^v(\mathbf{S})_{i+j2^{t-w+1}} = 1 \quad \text{for} \quad i = 0, \cdots, 2^{t-w+1} - 1. \tag{26}$$

Let p_i, $0 \leq p_i \leq 2^{n-l-1} - 1$, $i = 0, \cdots, 2^{t-w+1} - 1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. This means $w_H(\psi^{l+1}) = 2^{t-w+1}$ with 1s at positions p_i, $i = 0, \cdots, 2^{t-w+1} - 1$. As equation (26) is valid for $v = l + 1$, this implies that the mapping $p_i \mapsto p_i \bmod 2^{t-w+1}$ is one-one and onto since otherwise $w_H(\psi^{r_w-1}(\mathbf{S})) < 2^{t-w+1}$. Hence for each p_i, $i = 0, \cdots, 2^{t-w+1} - 1$, only one of the choices, that is, changing $\psi_L^l(\mathbf{S})_{p_i}$ or $\psi_R^l(\mathbf{S})_{p_i}$ results in the 2^{n-l-1}-vector $\psi^{l+1}(\mathbf{S}')$ that satisfies

$$\sum_{j=0}^{2^{r_w-l-2}-1} \psi^{l+1}(\mathbf{S}')_{i+j2^{t-w+1}} = 0 \quad \text{for} \quad i = 0, \cdots, 2^{t-w+1} - 1. \tag{27}$$

The contribution to $L(\mathbf{S})$ during the first $l - 1$ steps of the algorithm is

$$(2^{n-1} + 2^{n-2} + \cdots + 2^{n-l}) - \sum_{i=1}^{w-1} 2^{n-r_i} = 2^n - 2^{n-l} - \sum_{i=1}^{w-1} 2^{n-r_i}.$$

Thus the 2^t-error linear complexity of $\mathbf{S}$ is of the form

$$L_{2^t}(\mathbf{S}) = 2^n - 2^{n-l} - \sum_{i=1}^{w-1} 2^{n-r_i} + C, \tag{28}$$

where C is the linear complexity of the 2^{n-l-1}-periodic binary sequence with period $\psi^{l+1}(\mathbf{S}')$. By equation (27) and Lemma 4 the value C in equation (28) satisfies

$$C = L((\psi^{l+1}(\mathbf{S}'))^{\infty}) \leq 2^{n-l-1} - 2^{t-w+1}. \tag{29}$$

Also, $\psi^{l+1}(\mathbf{S}')$ is not the all zero vector from the definition of l in equation (21), which implies $C \geq 1$. Thus from equations (28) and (29) $L_{2^t}(\mathbf{S})$ is in the form $L_{w,l,C}$ given in equation (20).

Case 1b: $r_{j-1} \leq l \leq r_j - 2$, $1 \leq j \leq w-1$
From the definition of l in equation (21) we have $d_H(\psi_L^l(\mathbf{S}), \psi_R^l(\mathbf{S})) = 2^{t-j+1}$ Also, by equation (22) we have $w_H(\psi^{r_j-1}(\mathbf{S})) = 2^{t-j+1}$. Since $j \neq w$ we have $n - r_j > t - j$ and so $\psi^{r_j-1}(\mathbf{S})$ is not an all 1 vector. More specifically if

$$G = \{g : \psi^{r_j-1}(\mathbf{S})_g = 0, g = 0, \cdots, 2^{n-r_j+1} - 1\}$$

then

$$|G| = 2^{n-r_j+1} - 2^{t-j+1}. \tag{30}$$

Using a similar argument as that in Case 1a we have

$$L_{2^t}(\mathbf{S}) = 2^n - 2^{n-l} - \sum_{i=1}^{j-1} 2^{n-r_i} + C, \tag{31}$$

where C is the linear complexity of the 2^{n-l-1}-periodic binary sequence with period $\psi^{l+1}(\mathbf{S}')$, which is equal to either of the equal halves obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ such that the lowest possible linear complexity is achieved. The left and right halves of the vectors considered from the l-th step to the $(r_j - 2)$-th step are not equal. So by equation (30) due to the procedure of the Games-Chan algorithm we have

$$\sum_{f=0}^{2^{r_j-l-1}-1} \psi^l(\mathbf{S})_{i+f2^{n-r_j+1}} = 0 \quad \text{for} \quad i \in G \tag{32}$$

and

$$\sum_{f=0}^{2^{r_j-l-1}-1} \psi^l(\mathbf{S})_{i+f2^{n-r_j+1}} = 1 \quad \text{for} \quad i \in \{0, \cdots, 2^{n-r_j+1} - 1\} - G. \tag{33}$$

Let p_i, $0 \leq p_i \leq 2^{n-l-1} - 1$, $i = 0, \cdots, 2^{t-j+1} - 1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. This means $w_H(\psi^{l+1}(\mathbf{S})) = 2^{t-j+1}$. By equations (32)

and (33), this implies that the mapping $p_i \mapsto p_i \bmod 2^{n-r_j+1}$ is one-one since otherwise $w_H(\psi^{r_j-1}(\mathbf{S})) < 2^{t-j+1}$. We can see the mapping is not onto from equation (30). Also, each element in G does not occur as the inverse image of any element of the set $\{p_i : i = 0, \cdots, 2^{t-j+1}\}$. We split the summation in equation (32) into two separate summations involving terms exclusively from $\psi_L^l(\mathbf{S})$ or $\psi_R^l(\mathbf{S})$. For each $i \in G$ we have

$$\Sigma_L(l,i) = \sum_{f=0}^{2^{r_j-l-2}-1} \psi_L^l(\mathbf{S})_{i+f2^{n-r_j+1}}$$

and (34)

$$\Sigma_R(l,i) = \sum_{f=0}^{2^{r_j-l-2}-1} \psi_R^l(\mathbf{S})_{i+f2^{n-r_j+1}}.$$

For each $i \in G$, from equations (32) and (34) we know that $\Sigma_L(l,i)+\Sigma_R(l,i) = 0$ which implies $\Sigma_L(l,i) = \Sigma_R(l,i) = 0$ or $\Sigma_L(l,i) = \Sigma_R(l,i) = 1$. Note that none of the terms involved in the summations of equation (32) can be altered when forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$. Using these remarks it can be shown that by making appropriate changes at one of the positions p_i or $p_i + 2^{n-l-1}$, for each $i = 0, \cdots, 2^{t-j+1}$ in $\psi^l(\mathbf{S})$, we can only guarantee that $w_H(\psi^{l+1}(\mathbf{S}'))$ is even by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$. Thus the value C in equation (31) satisfies $1 \le C \le 2^{n-l-1} - 1$. Hence $L_{2^t}(\mathbf{S})$ is in the form $L_{j,l,C}$, $1 \le j \le w-1$, as in equation (19).

Case 2: $w = t + 1$
The proof in this case is similar to that for Case 1 and both forms in equations (19) and (20) are identical.

This completes the proof of the theorem. □

4 Counting Functions

In this section we derive expressions for the number of sequences in $\mathcal{A}(L)$ with fixed $k_{min}(L)$-error linear complexity. We need the following generalization of [1, Lemma 3].

Lemma 6. *Let* $\mathbf{S} \in \mathcal{A}(L)$ *such that* $1 \le L \le 2^n - 2^r$, $r = 1, \cdots, n-1$. *Let* $\mathbf{S}'$ *be a* 2^n*-periodic binary sequence corresponding to the polynomial*

$$\mathbf{S}'(x) = \mathbf{S}(x) + \sum_{t=0}^{g} x^{i_t},$$

where $0 \le g \le 2^r - 1$ *and* $i_t \in \{0, \cdots, 2^n - 1\}$, $t = 0, \cdots, g$. *If the mapping* $i_t \mapsto i_t \bmod 2^r$ *is one-one, then we have* $L(\mathbf{S}') > L(\mathbf{S})$.

Theorem 2. *Let $\mathcal{N}_{k_{min}(L)}(\mathcal{C})$ be the number of sequences in $\mathcal{A}(L)$, $1 < L < 2^n$, with fixed $k_{min}(L)$-error linear complexity $\mathcal{C}$. Let L be represented as*

$$L = 2^n - (2^{n-r_1} + 2^{n-r_2} + \cdots + 2^{n-r_t}),$$

where $r_0 = 0 < r_1 < r_2 < \cdots < r_t < n+1 = r_{t+1}$ and $1 \le t \le n-1$. Let $L_{j,l,C}$ be defined as in equations (19) and (20) and let $w = \min\{i : r_i = n+i-t, 1 \le i \le t+1\}$. Then for $1 \le j \le w$, if $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$ then

$$\mathcal{N}_{k_{min}(L)=2^t}(L_{j,l,C}) = 2^{\rho(j,l,C)},$$

where

$$\begin{aligned}\rho(j,l,C) = 2^n - 2^{n-l} - \sum_{i=1}^{j-1} 2^{n-r_i} + \sum_{i=0}^{w-j-1} (r_{w-i} - r_{w-i-1} - 1)2^{t-w+i+1} \\ + (r_j - l - 1)2^{t-j+1} + C - 1.\end{aligned} \tag{35}$$

Also, $\mathcal{N}_{k_{min}(L)=2^t}(0) = 2^{\rho(0)}$, where $\rho(0) = \sum_{i=0}^{w-2}(r_{w-i} - r_{w-i-1} - 1)2^{t-w+i+1} + (r_1 - 1)2^t$ and $\mathcal{N}_{k_{min}(L)=2^t}(\mathcal{C}) = 0$ for all $\mathcal{C}$ not in the form $L_{j,l,C}$ as in equations (19) and (20).

Proof. From equations (19) and (20) the $k_{min}(L)$-error linear complexity of $\mathbf{S} \in \mathcal{A}(L)$ is of the form

$$L_{j,l,C} = 2^n - \sum_{i=1}^{j-1} 2^{n-r_i} - 2^{n-l} + C \quad \text{for} \quad 1 \le j \le w \tag{36}$$

where $r_{j-1} \le l \le r_j - 2$ (For $l = r_w - 2$, there exist no positive values for C in equation (20) and hence no valid values for $L_{w,l,C}$). We determine the counting function for the number of sequences in $\mathcal{A}(L)$ with $k_{min}(L)$-error linear complexity equal to each of the values $L_{j,l,C}$ in equation (36) when $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$. From the definition of l in equation (21) and by equation (22), for any $\mathbf{S} \in \mathcal{A}(L)$ if $r_{j-1} \le l \le r_j - 2$ we know

$$w_H(\psi^{l+1}(\mathbf{S})) = w_H(\psi^{r_j-1}(\mathbf{S})) = 2^{t-j+1}. \tag{37}$$

We consider two cases based on the value of w.

Case 1: $w \le t$

From equation (24) for any $\mathbf{S} \in \mathcal{A}(L)$ the 2^{t-w+1}-vector $\psi^{r_w-1}(\mathbf{S})$ is an all 1 vector.

Let $\mathcal{D}^1(j,l,C)$ be the number of distinct 2^{n-l-1}-vectors $\psi^{l+1}(\mathbf{S})$ over all $\mathbf{S} \in \mathcal{A}(L)$ such that the 2^{n-r_w+1}-vector $\psi^{r_w-1}(\mathbf{S})$ is an all 1 vector. To determine $\mathcal{D}^1(j,l,C)$ we make the following observations.

(i) By equation (22) it is evident that during the execution of Games-Chan algorithm form the $l+1$-th step to the $(n-1)$-th step the Hamming weight of the vectors considered does not change between two consecutive steps except when going from the $(r_i - 1)$-th step to the r_i-th step for $i = j, \cdots, t$.

(ii) Using (i) the procedure of the Games-Chan algorithm also implies that over all sequences in $\mathcal{A}(L)$ for any integer a such that $l+1 \le a < r_j$ or $r_i \le a < r_{i-1}$ for some $i \in \{j, \cdots, t\}$, the number of distinct vectors in the a-th step that result in a fixed vector $\mathbf{v}$ in the $(a+1)$-th step is $2^{w_H(\mathbf{v})}$.
(iii) The definition of w implies $n - r_w = t - w$.

From these observations and by using property P1 recursively we obtain

$$\mathcal{D}^1(j,l,C) = \prod_{i=0}^{w-j-1} (2^{r_{w-i}-r_{w-i-1}-1})^{2^{t-w+i+1}} (2^{r_j-l-2})^{2^{t-j+1}}. \tag{38}$$

Recall that $\psi^{l+1}(\mathbf{S}')$ is the 2^{n-l-1}-vector obtained by forcing $\psi_L^l(\mathbf{S}) = \psi_R^l(\mathbf{S})$ so that the least linear complexity is achieved by making $k_{min}(L)$ errors in $\mathbf{S}^{(2^n)}$. Let $\mathcal{D}^2(j,l,C)$, $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$, be the number of choices for $\psi^{l+1}(\mathbf{S}')$ such that the linear complexity of the 2^{n-l-1}-periodic sequence with period $\psi^{l+1}(\mathbf{S}')$ is C. By equation (5), we have

$$\mathcal{D}^2(j,l,C) = 2^{C-1} \quad \text{for} \quad 1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}. \tag{39}$$

Over all $\mathbf{S} \in \mathcal{A}(L)$, for a fixed $\psi^{l+1}(\mathbf{S}) = \mathbf{v}$ with $w_H(\mathbf{v}) = 2^{n-r_w+1}$ and for a fixed choice of $\psi^{l+1}(\mathbf{S}')$ with $L((\psi^{l+1}(\mathbf{S}'))^\infty) = C$, the number of possibilities, denoted by $\mathcal{D}^3(w,l,C)$, for $\psi^l(\mathbf{S})$ such that $\psi_L^l(\mathbf{S}) + \psi_R^l(\mathbf{S}) = \mathbf{v}$ and $d_H(\psi^l(\mathbf{S}), \psi^{l+1}(\mathbf{S}') \mid \psi^{l+1}(\mathbf{S}')) = 2^{n-r_w+1}$ is

$$\mathcal{D}^3(w,l,C) = 2^{2^{t-j+1}}, \tag{40}$$

where $\psi^{l+1}(\mathbf{S}') \mid \psi^{l+1}(\mathbf{S}')$ is the 2^{n-l}-vector formed by concatenating two copies of $\psi^{l+1}(\mathbf{S}')$.

Let p_i, $0 \le p_i \le 2^{n-l-1} - 1$, $i = 0, \cdots, 2^{t-j+1} - 1$, be the positions where $\psi_L^l(\mathbf{S})$ and $\psi_R^l(\mathbf{S})$ differ. From Cases 1a and 1b of the proof of Theorem 1 the mapping $p_i \mapsto p_i \bmod 2^{n-r_j+1}$ is one-one. Using this mapping and the condition $1 \le C \le 2^{n-l-1} - 2^{n-r_j+1}$, by Lemma 6 for fixed $\psi^{l+1}(\mathbf{S})$ and $\psi^{l+1}(\mathbf{S}')$ each of the $2^{2^{t-j+1}}$ possibilities for $\psi^l(\mathbf{S})$ satisfies

$$L(\psi_L^l(\mathbf{S})) > C \quad \text{and} \quad L(\psi_R^l(\mathbf{S})) > C. \tag{41}$$

By equations (38)-(41), using properties P3 and P4 recursively we obtain

$$\mathcal{N}_{2^t}(L_{j,l,C}) = \mathcal{P}_0\mathcal{P}_1 \cdots \mathcal{P}_{l-1}\mathcal{D}^1(j,l,C)\mathcal{D}^2(j,l,C)\mathcal{D}^3(j,l,C). \tag{42}$$

We have

$$\begin{aligned} \mathcal{P}_0\mathcal{P}_1 \cdots \mathcal{P}_{l-1} &= \prod_{i=1}^{j-1} (\mathcal{P}_{r_{i-1}} \cdots \mathcal{P}_{r_i-2})(\mathcal{P}_{r_{j-1}} \cdots \mathcal{P}_{l-1}) \\ &= \left(\prod_{i=1}^{j-1} 2^{\sum_{z=1}^{r_i-r_{i-1}-1} 2^{n-r_i+z}} \right) 2^{\sum_{z=0}^{l-r_{j-1}-1} 2^{n-l+z}}. \end{aligned} \tag{43}$$

By equations (38)-(41) and (43) a straightforward algebraic simplification of the right hand side of equation (42) gives $\mathcal{N}_{2^t}(L_{j,l,C}) = 2^{\rho(j,l,C)}$ with $\rho(j,l,C)$ as in equation (35). We note that the condition in equation (41) is necessary to avoid double counting in determining the number of distinct possibilities for $\psi^l(\mathbf{S})$ over all $\mathbf{S} \in \mathcal{A}(L)$ such that $\psi^{l+1}(\mathbf{S})$ and $\psi^{l+1}(\mathbf{S}')$ are fixed.

Case 2: $w = t + 1$
In this case we note that the two possibilities for vectors in the $(n-1)$-th step of the Games-Chan algorithm are 01 and 10. Using this it can be shown that the expression for $\mathcal{D}^1(j,l,C)$ in equation (38) holds for $w = t + 1$. The remaining details are similar to those in Case 1.

To obtain $\mathcal{N}_{2^t}(0)$ we only have to count the number of $\mathbf{S} \in \mathcal{A}(L)$ with $w_H(\mathbf{S}) = 2^t$. By equation (22) and property P1 the expression for $\mathcal{N}_{2^t}(0)$ follows using an argument similar to that for finding $\mathcal{D}^1(j,l,C)$ as in equation (38).

This completes the proof of the theorem. □

5 Conclusion

In this paper we studied the k-error linear complexity of 2^n-periodic binary sequences by performing a rigorous analysis of the Games-Chan algorithm. We derived an expression for all the possible values of k-error linear complexities of 2^n-periodic binary sequences with fixed linear complexity when k is the minimum number of changes needed to the lower the linear complexity. For certain sets of these values, we obtained the corresponding number of sequences with fixed linear complexity and k-error linear complexity. Our results further research in analyzing the stability of linear complexity of 2^n-periodic binary sequences. These results, however, have limited importance for practical cryptography in part due to the restriction to 2^n-periodic sequences.

Acknowledgements

The first author thanks Dr. Zongming Fei for providing office space and resources while researching for this paper. The authors also thank anonymous reviewers for their helpful suggestions.

References

1. Fengxiang, Z., Wenfeng, Q.: The 2-error linear complexity of 2^n-periodic binary sequences with linear complexity $2^n - 1$. Journal of Electronics (China) 24(3), 390–395 (2007)
2. Games, R.A., Chan, A.H.: A fast algorithm for determining the complexity of a pseudo-random sequence with period 2^n. IEEE Trans. Inform. Theory 29(1), 144–146 (1983)
3. Kurosawa, K., Sato, F., Sakata, T., Kishimoto, W.: A relationship between linear complexity and k-error linear complexity. IEEE Trans. Inform. Theory 46(2), 694–698 (2000)

4. Meidl, W.: On the stability of 2^n-periodic binary sequences. IEEE Trans. Inform. Theory 51(3), 1151–1155 (2005)
5. Meidl, W., Venkateswarlu, A.: Remarks on the k-error linear complexity of p^n-periodic sequences. Design, Codes and Cryptography 42(2), 181–193 (2007)
6. Rueppel, R.A.: Analysis and Design of Stream Ciphers. Springer, Heidelberg (1986)
7. Stamp, M., Martin, C.F.: An algorithm for the k-error linear complexity of binary sequences with period 2^n. IEEE Trans. Inform. Theory 39(4), 1398–1401 (1993)