# Remarks on the $k$-error linear complexity of $p^n$-periodic sequences

Wilfried Meidl[1] and Ayineedi Venkateswarlu[2]

[1]Sabanci University, Orhanli, Tuzla, 34956 Istanbul, Turkey, wmeidl@sabanciuniv.edu

[2]Temasek Laboratories, National University of Singapore, 5 Sports Drive 2, Singapore 117508, Republic of Singapore, tslav@nus.edu.sg

**Abstract**

Recently the first author presented exact formulas for the number of $2^n$-periodic binary sequences with given 1-error linear complexity, and an exact formula for the expected 1-error linear complexity and upper and lower bounds for the expected $k$-error linear complexity, $k \geq 2$, of a random $2^n$-periodic binary sequence. A crucial role for the analysis played the Chan-Games algorithm. We use a more sophisticated generalization of the Chan-Games algorithm by Ding *et al.* to obtain exact formulas for the counting function and the expected value for the 1-error linear complexity for $p^n$-periodic sequences over $\mathbb{F}_p$, $p$ prime. Additionally we discuss the calculation of lower and upper bounds on the $k$-error linear complexity of $p^n$-periodic sequences over $\mathbb{F}_p$.

**keywords:** linear complexity, $k$-error linear complexity, Chan-Games algorithm, periodic sequences, stream cipher

**AMS Classification:** 94A55, 94A60, 11B50

## 1 Introduction

Let $S = s_1, s_2, \ldots$ be a sequence with terms in the finite field $\mathbb{F}_q$ (or shortly over $\mathbb{F}_q$). If, for a nonnegative integer $N$, the terms of $S$ satisfy $s_{i+N} = s_i$ for all $i \geq 1$, then we say that $S$ is $N$-periodic. The *linear complexity* of a periodic sequence $S$ over the finite field $\mathbb{F}_q$, denoted by $L(S)$, is the smallest positive integer $L$ for which there exist coefficients $d_0 = 1, d_1, d_2, \ldots, d_L$ in $\mathbb{F}_q$ such that

$$d_0 s_i + d_1 s_{i-1} + \ldots + d_L s_{i-L} = 0 \quad \text{for all } i \geq L + 1.$$

Trivially, the linear complexity of an $N$-periodic sequence can at most be $N$. The concept of linear complexity is very useful in the study of the security of stream ciphers (see [10, 11]). A necessary condition for the security of a keystream generator is that it produces a sequence with large linear complexity.

A cryptographically strong sequence should not only have a large linear complexity, but also altering a few terms should not cause a significant decrease of the linear complexity. According to this requirement, for an integer $k$, $0 \leq k \leq N$, in [12] Stamp and Martin defined the *$k$-error linear complexity* $L_k(S)$ of an $N$-periodic sequence $S$ with period $(s_1, s_2, \ldots, s_N)$ to be the smallest linear complexity that can be obtained by altering $k$ or fewer of the terms $s_i$, $1 \leq i \leq N$.

The concept of $k$-error linear complexity was built on the earlier concept of *sphere complexity* $SC_k(S)$ introduced in the monograph [1]. The sphere complexity $SC_k(S)$ of an $N$-periodic sequence over $\mathbb{F}_q$ can be defined by

$$SC_k(S) = \min_T L(T),$$

where the minimum is taken over all $N$-periodic sequences $T \neq S$ over $\mathbb{F}_q$ for which the period of $T$ differs from the period of $S$ at $k$ or fewer positions. Obviously, we have

$$L_k(S) = \min(SC_k(S), L(S)).$$

A lot of research has been done on the linear complexity and the $k$-error linear complexity of keystream sequences (for a recent survey we refer to [10]). However, for $k > 0$ we do not have formulas for the number of sequences with given $k$-error linear complexity or exact formulas for the expected $k$-error linear complexity of a random $N$-periodic sequence, not even for small $k$ such as $k = 1$. One exception is the rather simple case where $N$ is prime and $q$ is a primitive root modulo $N$. In this case the linear complexity can only attain the values $N$, $N-1$, 1 and 0. As a result, for this particular period it is possible to obtain exact values on the $k$-error linear complexity, $k > 0$ (cf. [8]).

In [8, 9] a technique to obtain lower bounds on the expected $k$-error linear complexity $E_k$ of a random $N$-periodic sequence over $\mathbb{F}_q$ has been presented. The technique of [8, 9] does not support the calculation of an upper bound for $E_k$. Solely for the rather simple case that $N$ is prime and $q$ is a primitive root modulo $N$, the technique of [8, 9] yields an exact formula for $E_k$ (cf. [8]).

We will consider $p^n$-periodic sequences over the finite field $\mathbb{F}_q$, $q = p^m$ for a prime $p$. For this class of sequences the technique of [8, 9] provides the lower bound

$$E_k \geq p^n - \log_q\left(\sum_{t=0}^{k}\binom{p^n}{t}(q-1)^t\right) - \frac{q}{q-1} \tag{1}$$

for the expected value $E_k$ of the $k$-error linear complexity.

$p^n$-periodic sequences over a finite field $\mathbb{F}_q$ with characteristic $p$ have been studied from several viewpoints. In [2] Games and Chan presented an algorithm that efficiently determines the linear complexity of a given $2^n$-periodic binary sequence. The Chan-Games algorithm has been generalized in [12] respectively [6] to an algorithm computing the $k$-error linear complexity of a $2^n$-periodic binary sequence for a fixed $k$ respectively for all $k$ simultaneously. These algorithms have been generalized in [1], [3] and [4] to more sophisticated algorithms applicable to $p^n$-periodic sequences over the finite field $\mathbb{F}_q$ with characteristic $p$.

In [7], elements of the algorithms in [2] and [12] have been used to obtain exact formulas for the counting function and the expected value for the 1-error linear complexity of $2^n$-periodic binary sequences. Moreover for $k \geq 2$ bounds for the expected $k$-error linear complexity of $2^n$-periodic binary sequences have been discussed. The question to which extent the more sophisticated algorithms in [1, 3] can be utilized to obtain related results on $p^n$-periodic sequences over $\mathbb{F}_q$ arises naturally. In Section 2, the main part, we obtain exact formulas for the number of $p^n$-periodic sequences over the prime field $\mathbb{F}_p$ with given 1-error linear complexity and for the expected 1-error linear complexity. In Section 3 we concentrate on the calculation of bounds on the $k$-error linear complexity of $p^n$-periodic sequences over $\mathbb{F}_p$.

## 2 Counting functions and expected values for $\mathbf{k = 1}$

In [9] it has been shown that the number $\mathcal{N}(L)$ of $p^n$-periodic sequences over $\mathbb{F}_q$, $q = p^m$, $p$ prime, with given linear complexity $L$, $0 \leq L \leq p^n$, is given by

$$\mathcal{N}(0) = 1 \quad \text{and} \quad \mathcal{N}(L) = (q-1)q^{L-1} \quad \text{for} \quad 1 \leq L \leq p^n. \tag{2}$$

In [5] Kurosawa *et al.* showed that the minimum value $k$ for which the $k$-error linear complexity of a $p^n$-periodic sequence $S$ over $\mathbb{F}_q$ is strictly less than the linear complexity $L(S)$ of $S$ is exactly determined by

$$k = Prod(p^n - L(S)), \tag{3}$$

where $Prod(C) := \prod_{j=0}^{m-1}(i_j + 1)$ if $C = i_0 + i_1 p + \cdots + i_{m-1}p^{m-1}$. In particular, the sequences with maximal possible linear complexity $p^n$ are the only sequences for which the 1-error linear complexity is less than the linear complexity. Hence it suffices to calculate the number of sequences with linear complexity $p^n$ and given 1-error linear complexity $L$, $0 \leq L < p^n$, in order to obtain the complete counting function for the 1-error linear complexity. As it is well known (see e.g. [5, Proposition 2.1]), the set of $p^n$-periodic sequences over $\mathbb{F}_q$, $q = p^m, p$ prime, with maximal possible linear complexity $p^n$ is exactly the set of sequences for which the sum of the elements of one period is not zero.

We will utilize the generalized Chan-Games algorithm presented in [1]. The algorithm can be described as follows:
Let $S$ be a $p^n$-periodic sequence over $\mathbb{F}_q$, $q = p^m$, $p$ prime, with period $(s_1, s_2, \ldots, s_{p^n})$ and $\mathcal{A} = (a_{i,j})$ the $(p-1) \times p$-matrix with $a_{i,j} = \binom{p-j}{i-1}$, then we define the matrix $\mathcal{B}$ to be the $(p-1) \times p^{n-1}$-matrix with $l$th column equal to $\mathcal{A}(s_l\, s_{l+p^{n-1}} \ldots s_{l+(p-1)p^{n-1}})^T$, $l = 1, 2, \ldots, p^{n-1}$. The linear complexity $L(S)$ of the sequence $S$ is then given by

$$(p-w)p^{n-1} + L(S_1),$$

where $w$ is the least integer such that the $w$th row of $\mathcal{B}$ is not the zero row, or $w = p$ if $\mathcal{B}$ is the zero matrix, and $S_1$ is the $p^{n-1}$-periodic sequence with the $w$th row of $\mathcal{B}$ as period if $\mathcal{B}$ is not the zero matrix, or $(s_1, s_2, \ldots, s_{p^{n-1}})$ as period if $\mathcal{B}$ is the zero matrix. The generalized Chan-Games algorithm is obtained by applying this result recursively, which is possible since the period length of $S_1$ is again a power of $p$. In the final step we will have a sequence with period $p^0 = 1$, i.e., a constant sequence $s_1, s_1, \ldots$. If $s_1 \neq 0$ we add 1 to the present value for the linear complexity of $S$.

The described algorithm motivates a mapping $\varphi_n$ from $\mathbb{F}_q^{p^n}$ into $\mathbb{F}_q^{(p-1)\times p^{n-1}}$, $n \geq 1$, defined by

$$\varphi_n((s_1, s_2, \ldots, s_{p^n})) = \mathcal{B},$$

where $\mathcal{B}$ is defined as above.

Let $H(\mathbf{v})$ denote the Hamming weight of a vector $\mathbf{v}$. Let $\mathbf{s}^{(n)}$ be any element of $\mathbb{F}_q^{p^n}$ and let $\boldsymbol{b}(u)$, $u = 0, \ldots, p-2$, be the $(u+1)$th row of the matrix $\mathcal{B}$. We collect some (obvious) properties of the matrix $\mathcal{A}$ and the mapping $\varphi_n$ respectively the matrix $\mathcal{B} = \varphi_n(\mathbf{s}^{(n)})$.

P1 The matrix $\mathcal{A}$ has rank $p-1$. Hence the linear system $\mathcal{A}\mathbf{x} = \mathbf{b}$ has $q$ different solutions in $\mathbb{F}_q^p$. In particular the vectors $c(1, 1, \ldots, 1)$, $c \in \mathbb{F}_q$, are the solutions of the homogenous system $\mathcal{A}\mathbf{x} = \mathbf{0}$.

P2 $H(\boldsymbol{b}(u)) \leq H(\mathbf{s}^{(n)})$ for $0 \leq u \leq p-2$.

P3 The sum of the elements of the first row $\boldsymbol{b}(0)$ of $\mathcal{B}$ equals the sum of the elements of $\mathbf{s}^{(n)}$.

P4 The set $\varphi_{t+1}^{-1} := \{\mathbf{v} \in \mathbb{F}_q^{p^{t+1}} \mid \varphi_{t+1}(\mathbf{v}) = \mathcal{B}\}$ for a given $(p-1) \times p^t$-matrix $\mathcal{B}$ over $\mathbb{F}_q$ has cardinality $q^{p^t}$.

We restrict ourselves to the case of the prime field $\mathbb{F}_p$. Then we can show the following lemma.

**Lemma 1** *Let $\mathcal{A}$ be the matrix defined as above and suppose that for $\mathbf{v} \in \mathbb{F}_p^p$ we have $\mathcal{A}\mathbf{v} = (u_1 \neq 0, u_2, \ldots, u_{p-1})$. Then we have $p$ vectors $\mathbf{v}_i$, $1 \leq i \leq p$, such that the first component of $\mathcal{A}\mathbf{v}_i$ is zero, i.e., $\mathcal{A}\mathbf{v}_i = (0, u_2', \ldots, u_{p-1}')$ for some $u_2', \ldots, u_{p-1}' \in \mathbb{F}_p$, and $\mathbf{v}_i$ differs from $\mathbf{v}$ at exactly one position. Moreover for each given $z \in \mathbb{F}_p$ there exists exactly one vector $\mathbf{v}_{i_z}$, $1 \leq i_z \leq p$, which differs from $\mathbf{v}$ at exactly one position and $\mathcal{A}\mathbf{v}_{i_z} = (0, z, \hat{u}_3, \ldots, \hat{u}_{p-1})$.*

*Proof.* Evidently, for $1 \leq i \leq p$, the vectors $\mathbf{v}_i := \mathbf{v} + \mathbf{e_i}$, where $\mathbf{e_i}$ is the vector with $i$th entry $-u_1$ and $H(\mathbf{e_i}) = 1$, satisfy $\mathcal{A}\mathbf{v}_i = (0, u_2', \ldots, u_p')$ for some $u_2', \ldots, u_p' \in \mathbb{F}_p$. Since the second row of $\mathcal{A}$ consists of all elements of the prime field $\mathbb{F}_p$, we will have $\mathcal{A}\mathbf{v}_{i_z} = (0, z, \hat{u}_3, \ldots, \hat{u}_{p-1})$ for exactly one $1 \leq i_z \leq p$ and for some $\hat{u}_3, \ldots, \hat{u}_{p-1} \in \mathbb{F}_p$. □

**Proposition 1** *Let $S$ be a $p^n$-periodic sequence over $\mathbb{F}_p$ with maximal possible linear complexity $L(S) = p^n$. Then the 1-error linear complexity of $S$ is 0 or of the form*

$$\begin{aligned} L_{r,w,C} \quad := \quad & p^n - wp^r + C, \quad 0 \leq r \leq n-1, \\ & 2 \leq w \leq p-1 \ \textit{and} \ 0 \leq C \leq p^r - 1, \quad \textit{or} \\ & w = p, r \neq 0 \ \textit{and} \ 1 \leq C \leq p^r - 1. \end{aligned} \tag{4}$$

*Proof.* Evidently the sequences $S$ with maximal linear complexity $p^n$ and 1-error linear complexity $L_1(S) = 0$ are exactly the sequences with one term different from 0 per period. We now show that the 1-error linear complexity of the remaining $p^n$-periodic sequences $S$ with period $\mathbf{s}^{(n)}$ and linear complexity $p^n$ is of the form (4). Since $L(S) = p^n$, the sequence $S$ does not have the zero sum property. With the property P3 for all $1 \leq m \leq n$ the first row of the matrix $\varphi_m \varphi_{m+1} \cdots \varphi_n(\mathbf{s}^{(n)})$ is not the zero vector. Suppose that $r$, $0 \leq r \leq n-1$, is the largest integer such that the first row $\boldsymbol{b}(0)$ of the $(p-1) \times p^r$-matrix $\mathcal{B} = \varphi_{r+1} \cdots \varphi_n(\mathbf{s}^{(n)})$ has Hamming weight 1. We want to change one term of the preimage of $\mathcal{B}$ so that the resulting linear complexity of the sequence is as small as possible. Since the linear complexity of the sequence corresponding to $\boldsymbol{b}(1)$ is lower than $p^r$ if and only if $\boldsymbol{b}(1)$ has the zero sum property, the optimal choice is to perform a term change such that we obtain the zero vector for $\boldsymbol{b}(0)$ and additionally a vector with zero sum property for $\boldsymbol{b}(1)$. According to Lemma 1 we have exactly one choice for the term change with this property. In the case where $r = 0$, the matrix $\mathcal{B}$ is a column matrix and hence $\boldsymbol{b}(0) \neq \mathbf{0}$. By changing one term we can make $\boldsymbol{b}(1)$ also zero. If after the term change $\boldsymbol{b}(w)$ is the first non zero entry in $\mathcal{B}$ then the 1-error linear complexity of $S$ is $p^n - w$, $2 \leq w \leq p-2$. Observe that after the term change, if the column matrix $\mathcal{B}$ becomes zero then the first row of $\varphi_2 \cdots \varphi_n(\mathbf{s}^{(n)})$ contains $p$ identical nonzero entries. Thus the 1-error linear complexity of $S$ is $p^n - p + 1$.

Now suppose $1 \leq r \leq n-1$ and $\boldsymbol{b}(1)$ is different from the zero vector after the term change, then the 1-error linear complexity of $S$ is $p^n - 2p^r + C$, $1 \leq C \leq p^r - 1$. If after the term change $\boldsymbol{b}(1)$ is the zero vector but $\boldsymbol{b}(2)$ is not, then the 1-error linear complexity of $S$ is $p^n - 2p^r$ if the linear complexity of the sequence with period $\boldsymbol{b}(2)$ is $p^r$ and $p^n - 3p^r + C$, $1 \leq C \leq p^r - 1$, if the linear complexity of the sequence with period $\boldsymbol{b}(2)$ is $1 \leq C \leq p^r - 1$. In general, if after the term change $\boldsymbol{b}(w)$, $3 \leq w \leq p-2$, is the first row in $\mathcal{B}$ not equal to the zero vector, then the 1-error linear complexity of $S$ is $p^n - wp^r$ if the linear complexity of the sequence with period $\boldsymbol{b}(w)$ is $p^r$ and $L_1(S) = p^n - (w+1)p^r + C$, $1 \leq C \leq p^r - 1$, if the linear complexity of the sequence with period $\boldsymbol{b}(w)$ is $1 \leq C \leq p^r - 1$. Finally if after the term change $\mathcal{B}$ is the zero matrix, then the 1-error linear complexity of $S$ is $p^n - p^{r+1} + p^r$ if the linear complexity of the sequence $S_1$ whose period consists of the first $p^r$ terms of the (altered) preimage of $\mathcal{B}$ is $p^r$ and $L(S) = p^n - p^{r+1} + C$, $1 \leq C \leq p^r - 1$, if the linear complexity of $S_1$ is $1 \leq C \leq p^r - 1$. Note that the 1-error linear complexity will never be $p^n - p^{r+1}$. $\square$

The next proposition presents the counting function for the 1-error linear complexity for $p^n$-periodic sequence over $\mathbb{F}_p$ with maximal possible linear complexity $L(S) = p^n$.

**Proposition 2** *Let $\bar{\mathcal{N}}_1(L)$ be the number of $p^n$-periodic sequences $S$ over $\mathbb{F}_p$ with maximal possible linear complexity $L(S) = p^n$ and 1-error linear complexity $L_1(S) = L$, and let $L_{r,w,C}$ be defined as in (4). Then*

$$\bar{\mathcal{N}}_1(L_{r,w,C}) = (p-1)^2 p^{p^n - wp^r + r + C},$$

$\bar{\mathcal{N}}_1(0) = (p-1)p^n$, *and* $\bar{\mathcal{N}}_1(L) = 0$ *if* $L \neq 0$ *is not of the form* (4).

*Proof.* Evidently we have $\bar{\mathcal{N}}_1(0) = (p-1)p^n$, which equals the number of $p^n$-periodic sequences $S$ over $\mathbb{F}_p$ with one term different from 0 per period. The identity $\bar{\mathcal{N}}_1(L) = 0$ if $L \neq 0$ is not of the form (4) immediately follows from Proposition 1.

The sequences with linear complexity $p^n$ and 1-error linear complexity $p^n - 2p^r + C$, $1 \leq C \leq p^r - 1$, are exactly those sequences for which the matrix $\mathcal{B} = \varphi_{r+1} \cdots \varphi_n(\mathbf{s}^{(n)})$ has a first row $\boldsymbol{b}(0)$ with $H(\boldsymbol{b}(0)) = 1$, and additionally after changing one term of the preimage of $\mathcal{B}$ in the unique way such that $\boldsymbol{b}(0)$ becomes the zero vector and $\boldsymbol{b}(1)$ has the zero sum property, the sequence with period $\boldsymbol{b}(1)$ (altered version) has linear complexity $C$. We have $(p-1)p^r$ possibilities to choose $\boldsymbol{b}(0)$ with $H(\boldsymbol{b}(0)) = 1$, $(p-1)p^{C-1}$ possibilities to choose a sequence with linear complexity $C$ for $\boldsymbol{b}(1)$, and initially the term of $\boldsymbol{b}(1)$ in the same column as the nonzero entry in $\boldsymbol{b}(0)$ can be chosen arbitrarily. The remaining rows of $\mathcal{B}$ are arbitrary. Hence we have $(p-1)^2 p^{r+C} p^{(p-3)p^r}$ different choices for $\mathcal{B}$. According to P4 the matrix $\mathcal{B}$ has $p^{p^r}$ preimages $\mathbf{s}^{r+1} \in \mathbb{F}_p^{p^{r+1}}$, which will be the first row of a certain $(p-1) \times p^{r+1}$-matrix $\mathcal{B}'$. Note that $H(\mathbf{s}^{r+1}) > 1$, else we would obtain the zero matrix for $\mathcal{B}$ with one term change. For exactly $p^{(p-1)p^{r+1}}$ vectors $\mathbf{s}^{r+2} \in \mathbb{F}_p^{p^{r+2}}$ the matrix $\mathcal{B}' = \varphi_{r+2}(\mathbf{s}^{r+2})$ has $\mathbf{s}^{(r+1)}$ as the first row. Recursively we get $p^{p^n - p^{r+1} + p^r}$ for the numbers of vectors $\mathbf{s}^{(n)} \in \mathbb{F}_p^{p^n}$ with $\varphi_{r+1} \cdots \varphi_n(\mathbf{s}^{(n)}) = \mathcal{B}$, which leads to the desired formula for the number of $p^n$-periodic sequences over $\mathbb{F}_p$ with 1-error linear complexity $p^n - 2p^r + C$, $1 \leq C \leq p^r - 1$.

To determine the number of sequences with linear complexity $p^n$ and 1-error linear complexity $L_{r,w,C}$, $3 \leq w \leq p-1$, $C \geq 1$, we have to consider the $(p-1) \times p^r$-matrices that can be transformed into a matrix for which $\boldsymbol{b}(w-1)$ is the first row different from the zero vector by changing exactly one term in the preimage. The first $w-1$ rows of $\mathcal{B}$ can have nonzero elements in exclusively one column, say the column with index $i$. The $i$th element of $\boldsymbol{b}(0)$ must of course be nonzero, the $i$th element of $\boldsymbol{b}(1)$ can be chosen arbitrarily. These two elements uniquely determine the term change that has to be performed in a preimage in order to obtain $\boldsymbol{b}(0) = \boldsymbol{b}(1) = \mathbf{0}$. For $2 \leq u \leq w-2$, the $i$th element of $\boldsymbol{b}(u)$ is uniquely determined such that $\boldsymbol{b}(u)$ is transformed into the zero vector after that uniquely determined term change. For $\boldsymbol{b}(w-1)$ we choose one of the $(p-1)p^{C-1}$ vectors with corresponding $p^r$-periodic sequence having linear complexity $C$. Note that the $i$th entry of $\boldsymbol{b}(w-1)$ is adapted according to the term change that has to be performed in the preimage. The remaining entries of $\mathcal{B}$ are again arbitrary. This yields $(p-1)^2 p^{C+r} p^{(p-1-w)p^r}$ different matrices with the desired properties. With the same argument as before we get the formula for $\bar{\mathcal{N}}_1(L_{r,w,C})$. Note that for $C = p^r$ we get the formula for $\bar{\mathcal{N}}_1(L_{r,w-1,0})$. In the case where $r = 0$ we always can make $\boldsymbol{b}(1) = \mathbf{0}$ by a single term change in the original sequence. Suppose $\boldsymbol{b}(w-1)$ is the first nonzero entry in $\mathcal{B}$ then we get $C = 1$, and so $\bar{\mathcal{N}}_1(L_{0,w,1}) = \bar{\mathcal{N}}_1(L_{0,w-1,0})$ for $3 \le w \le p-1$.

Finally according to P1, $\varphi_{r+1}(\mathbf{s}^{r+1}) = \mathcal{B}$ is the zero matrix if and only if $\mathbf{s}^{(r+1)}$ consists of $p$ identical copies of a vector $\mathbf{s}^{(r)} \in \mathbb{F}_p^{p^r}$. Let $M(r, C)$ be the number of vectors which have Hamming distance 1 to a vector in $\mathbb{F}_p^{p^{r+1}}$ that consist of $p$ identical copies of a vector $\mathbf{s}^{(r)} \in \mathbb{F}_p^{p^r}$ such that the corresponding $p^r$-periodic sequence has linear complexity $C$. Then the number $\bar{\mathcal{N}}_1(L_{r,p,C})$, $1 \le C \le p^r - 1$, is given by $M(r, C)p^{p^n - p^{r+1}}$. With simple combinatorial arguments we get $M(r, C) = (p-1)^2 p^{r+C}$, which yields the desired formula. Again with $C = p^r$ we get the formula for $\bar{\mathcal{N}}_1(L_{r,p-1,0})$. □

The construction of the integers $L_{r,\omega,C}$ in (4) reflects the operation mode of the Chan-Games algorithm. Evidently, the set of integers of the form (4) can also be described as the set of integers $L$, $0 < L < p^n$, which are not of the form $p^n - p^t$, $t = 0, 1, \ldots, n-1$. We observe that $r = \lfloor \log_p(p^n - L_{r,\omega,C}) \rfloor$ and combine Proposition 2 and the identity (2) to the following theorem, where we use the fact that $L_1(S) = L(S)$ if $L(S) < p^n$.

**Theorem 1** *Let $\mathcal{N}_1(L)$, $0 \le L \le p^n$, be the number of $p^n$-periodic sequences over $\mathbb{F}_p$, $p$ prime, with 1-error linear complexity equal to $L$. Then we have*

$$\begin{aligned}
\mathcal{N}_1(0) &= 1 + (p-1)p^n \\
\mathcal{N}_1(L) &= (p-1)p^{L-1} \quad \text{if } L = p^n - p^t, t = 0, 1, \ldots, n-1, \\
\mathcal{N}_1(L) &= (p-1)p^{L-1} + (p-1)^2 p^{L + \lfloor \log_p(p^n - L) \rfloor} \quad \text{if } L \ne p^n \text{ and} \\
& \quad\ L \ne p^n - p^t, t = 0, 1, \ldots, n, \text{ and} \\
\mathcal{N}_1(p^n) &= 0.
\end{aligned}$$

From Proposition 2 we can conclude that a large proportion of the $p^n$-periodic sequences with linear complexity $p^n$ still possesses a very high linear complexity after changing one of its terms. We use Proposition 2 to obtain an exact formula for the expected value of the 1-error linear complexity of a random $p^n$-periodic sequence over $\mathbb{F}_p$ with linear complexity $p^n$.

**Proposition 3** *The expected value $E_{1|L=p^n}$ of the 1-error linear complexity of a random $p^n$-periodic sequence $S$ over $\mathbb{F}_p$ with linear complexity $L(S) = p^n$, $n \ge 2$, is given by*

$$E_{1|L=p^n} = p^n - 1 - \frac{p}{p-1} + \frac{p^{n+1}}{(p-1)p^{p^n}} - \sum_{r=1}^{n-1} \frac{p^{r+1}}{p^{p^r}}.$$

*Proof.* From Proposition 2 we have

$$\begin{aligned}
p^{p^n-1}(p-1)E_{1|L=p^n} &= \sum_{r=1}^{n-1}\sum_{w=2}^{p}\sum_{C=1}^{p^r-1}\bar{\mathcal{N}}_1(L_{r,w,C})\cdot L_{r,w,C} \\
&\quad +\sum_{r=0}^{n-1}\sum_{w=2}^{p-1}\bar{\mathcal{N}}_1(L_{r,w,0})\cdot L_{r,w,0} \qquad (5)\\
&= \sum_{r=1}^{n-1}\sum_{w=2}^{p}\sum_{C=1}^{p^r-1}(p-1)^2p^{p^n-wp^r+r+C}(p^n-wp^r+C) \\
&\quad +\sum_{r=0}^{n-1}\sum_{w=2}^{p-1}(p-1)^2p^{p^n-wp^r+r}(p^n-wp^r) \\
&= (p-1)^2p^{p^n+n}\sum_{r=1}^{n-1}\sum_{w=2}^{p}p^{-wp^r+r}\sum_{C=1}^{p^r-1}p^C \\
&\quad -(p-1)^2p^{p^n}\sum_{r=1}^{n-1}\sum_{w=2}^{p}p^{-wp^r+r}wp^r\sum_{C=1}^{p^r-1}p^C \\
&\quad +(p-1)^2p^{p^n}\sum_{r=1}^{n-1}\sum_{w=2}^{p}p^{-wp^r+r}\sum_{C=1}^{p^r-1}Cp^C \\
&\quad +(p-1)^2p^{p^n+n}\sum_{r=0}^{n-1}\sum_{w=2}^{p-1}p^{-wp^r+r} \\
&\quad -(p-1)^2p^{p^n}\sum_{r=0}^{n-1}\sum_{w=2}^{p-1}p^{-wp^r+r}wp^r \\
&= T_1-T_2+T_3+T_4-T_5.
\end{aligned}$$

With a sequence of well known algebraic manipulations including expansion of some series one can obtain

$$\begin{aligned}
T_1 &= (p-1)p^{p^n+n-1}-(p-1)p^{2n}-T_4,\\
T_2 &= T_6-p^{p^n-p+1}+p^{p^n-1}(2p-1)-(p-1)p^{2n}-T_5, \text{ and}\\
T_3 &= T_6+p^n-(p-1)p^{p^n}\sum_{r=1}^{n-1}p^{-p^r+r}-p^{p^n-p+1}.
\end{aligned}$$

Combining the results we get

$$T_1-T_2+T_3+T_4-T_5 = (p-1)p^{p^n+n-1}-p^{p^n-1}(2p-1)$$

$$+p^n - (p-1)p^{p^n}\sum_{r=1}^{n-1} p^{-p^r+r},$$

and hence

$$(p-1)p^{p^n-1}E_{1|L=p^n} = (p-1)p^{p^n-1}\left(p^n - 1 - \frac{p}{p-1} + \frac{p^{n+1}}{(p-1)p^{p^n}} - \sum_{r=1}^{n-1}\frac{p^{r+1}}{p^{p^r}}\right),$$

which yields the desired formula. □

**Theorem 2** *The expected value $E_1$ of the 1-error linear complexity of a random $p^n$-periodic sequence over $\mathbb{F}_p$, $n \geq 2$, is given by*

$$E_1 = p^n - 2 - \frac{1}{p(p-1)} + \frac{1}{p^{p^n}}\left(p^n + \frac{1}{p-1}\right) - (p-1)\sum_{r=1}^{n-1}\frac{p^r}{p^{p^r}}.$$

*Proof.* With (2) and (3) we get the sum $p^{p^n}E_1$ by adding

$$\sum_{L=0}^{p^n-1}(p-1)p^{L-1}L = p^{p^n+n-1} - \frac{p^{p^n}}{p-1} + \frac{1}{p-1}$$

to (5), which will yield the result. □

## 3 On the expected k-error linear complexity, k ≥ 2

We start with a proposition which rules out several values for the $k$-error linear complexity. It is an analogue to [7, Proposition 1]

**Proposition 4** *Let $S$ be any $p^n$-periodic sequence over $\mathbb{F}_p$. Then for $k \geq 2$ the $k$-error linear complexity $L_k(S)$ of $S$ is different from $p^n - p^t$ for every integer $t$ with $0 \leq t < n$.*

*Proof.* If the Hamming weight of the period $\mathbf{s}^{(n)}$ of $S$ is at most $k$ then we have $L_k(S) = 0$. Else there is a largest integer $t$ such that the first row $\boldsymbol{b}(0)$ of $\mathcal{B} = \varphi_{t+1}\cdots\varphi_n(\mathbf{s}^{(n)})$ satisfies $H(\boldsymbol{b}(0)) \leq k$, and we can obtain $\boldsymbol{b}(0) = \mathbf{0}$ by at most $k$ term changes in $\mathbf{s}^{(n)}$. Thus we have $L_k(S) = p^n - wp^t + C$, $2 \leq w \leq p$. If $w = 2$, i.e., if we cannot obtain $\boldsymbol{b}(1) = \mathbf{0}$ by at most $k$ term changes, then we have $1 \leq C \leq p^t - 1$, since by Lemma 1 we are at least able to force $\boldsymbol{b}(1)$ to have the zero sum property. Consequently we have

$L_k(S) \le p^n - p^t - 1$. If $w = p$, i.e. with at most $k$ term changes in $\mathbf{s}^{(n)}$ the matrix $\mathcal{B}$ can be transformed into the zero matrix, then $L_k(S) = p^n - p^{t+1} + C$. We can exclude that $C = 0$ since then the first row of $\mathcal{B}' = \varphi_{t+2} \cdots \varphi_n(\mathbf{s}^{(n)})$ must have a smaller Hamming weight than $k + 1$, which is a contradiction to the definition of $t$. □

The following Proposition 5 and Corollary 1 are generalizations of [7, Proposition 2, Corollary 2] and [7, Theorem 3, Corollary 3], respectively. The proofs are similar to the proofs in [7], and therefore omitted.

**Proposition 5** *For $k \ge 2$ and $0 \le t \le n$, the number $\mathcal{M}_k(t)$ of $p^n$-periodic sequences $S$ over $\mathbb{F}_p$ with $k$-error linear complexity $L_k(S) > p^n - p^t$ is given by*

$$\mathcal{M}_k(t) = p^{p^n} - p^{p^n - p^t} \sum_{j=0}^{k} \binom{p^t}{j} (p-1)^j.$$

*The number $\mathcal{M}_k(t+1,t)$, $0 \le t \le n-1$, of $p^n$-periodic sequences $S$ over $\mathbb{F}_p$ satisfying $p^n - p^{t+1} < L_k(S) < p^n - p^t$ is given by*

$$\mathcal{M}_k(t+1,t) = p^{p^n - p^t} \sum_{j=0}^{k} \binom{p^t}{j} (p-1)^j - p^{p^n - p^{t+1}} \sum_{j=0}^{k} \binom{p^{t+1}}{j} (p-1)^j.$$

Observe that for $p^t \le k < p^{t+1}$ we have $\mathcal{M}_k(0) = \cdots = \mathcal{M}_k(t) = 0$ and $\mathcal{M}_k(t+1) > 0$. The partition $[p^n - p^{t+1}, p^n - p^t)$, $t = n-1, n-2, \ldots, 0$, of the interval $[0, p^n - 1)$ along with the above proposition yields the following bounds.

**Corollary 1** *For an integer $k \ge 2$ the expected value $E_k$ of the $k$-error linear complexity of a random $p^n$-periodic sequence over $\mathbb{F}_p$ satisfies*

$$p^n - p^{\lfloor \log_p k \rfloor + 1} + 1 - \frac{1}{p^{p^n}} \sum_{j=0}^{k} \binom{p^n}{j} (p-1)^j - \sum_{t=\lfloor \log_p k \rfloor + 1}^{n-1} \frac{p^t}{p^{p^t}} \sum_{j=0}^{k} \binom{p^t}{j} (p-1)^{j+1}$$

$$\le E_k \le p^n - p^{\lfloor \log_p k \rfloor} - 1 - \frac{p^n - p^{n-1} + 1}{p^{p^n}} \sum_{j=0}^{k} \binom{p^n}{j} (p-1)^j -$$

$$\sum_{t=\lfloor \log_p k \rfloor + 1}^{n-1} \frac{p^t}{p^{p^t+1}} \sum_{j=0}^{k} \binom{p^t}{j} (p-1)^{j+1}.$$

We emphasize that the technique used in [8, 9] yields only lower bounds. Hence the main improvement is that our method also yields an upper bound. We observe that if $k$ is a small proportion of the period then the upper and the lower bound given in Corollary 1 do not differ significantly.
As stated in [7], in the binary case the lower bound in Corollary 1 improves the lower bound (1). As experimental results demonstrate, it needs a refined analysis in order to obtain an appreciable improvement of (1). Though our approach yields complex formulas and becomes infeasible if $p$ is not very small, we find it worth to be discussed. We restrict ourselves to the ternary case.

We know that the $k$-error linear complexity of a ternary $3^n$-periodic sequence $S$ is less than $3^n - 3^t$ if and only if the Hamming weight of the first row $\boldsymbol{b}_t(0)$ of the $2 \times 3^t$-matrix $\mathcal{B} = \varphi_{t+1} \cdots \varphi_n(\mathbf{s}^{(n)})$ is at most $k$, i.e., we can obtain the zero vector for $\boldsymbol{b}_t(0)$ by changing just $k$ or fewer terms in the preimage of $\mathcal{B}$. If we additionally can obtain the zero vector for the second row of $\mathcal{B}$ by changing just $k$ or fewer terms in the preimage of $\mathcal{B}$, then the $k$-error linear complexity of $S$ is at most $3^n - 2 \cdot 3^t$. Let $\mathbf{c} = \binom{x}{y}$ be a column of $\mathcal{B}$. If $x \neq 0$ then we can transform $\mathbf{c}$ into the zero column by one (unique) term change in the preimage of $\mathcal{B}$. If $x = 0$ but $y \neq 0$ then we need 2 term changes in the preimage of $\mathcal{B}$ in order to obtain the zero column for $\mathbf{c}$ (we will have 3 different options to change two terms).
These observations lead to the following generalization of the Hamming weight.

**Definition 1** *The Chan-Games weight of a non zero column is* 1 *plus the number of zeros that lie above the first nonzero element of the column. The zero column has Chan-Games weight* 0*. The Chan-Games weight* $Wt(\mathcal{B})$ *of a matrix* $\mathcal{B}$ *is the sum of the Chan-Games weights of its columns.*

According to the above observations the $k$-error linear complexity of a $3^n$-periodic ternary sequence $S$ is at most $3^n - 2 \cdot 3^t$ if and only if $Wt(\mathcal{B}) \leq k$. With combinatorial arguments we get the following Lemma.

**Lemma 2** *The number of ternary* $2 \times 3^t$*-matrices* $\mathcal{B}$ *satisfying* $Wt(\mathcal{B}) \leq k$ *is given by*

$$\sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor \frac{k-j}{2} \rfloor} \binom{3^t - j}{i} 2^i.$$

*Proof.* For each choice of $0 \leq j \leq k$ columns with Chan-Games weight 1 we can choose at most $\lfloor (k-j)/2 \rfloor$ further columns with Chan-Games weight 2 in order that $Wt(\mathcal{B})$ does not exceed $k$. □

Lemma 2 and Proposition 5 yield the following results.

**Proposition 6** *For $k \geq 2$ and $0 \leq t \leq n-1$, the number of ternary $3^n$-periodic sequences $S$ with $k$-error linear complexity $L_k(S) > 3^n - 2\cdot 3^t$ is given by*

$$3^{3^n} - 3^{3^n - 2\cdot 3^t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor \frac{k-j}{2} \rfloor} \binom{3^t - j}{i} 2^i.$$

*The number of ternary $3^n$-periodic sequences $S$ with $k$-error linear complexity $3^n - 2\cdot 3^t < L_k(S) < 3^n - 3^t$ is given by*

$$S_{II} = 3^{3^n - 3^t} \sum_{j=0}^{k} \binom{3^t}{j} 2^j - 3^{3^n - 2\cdot 3^t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor \frac{k-j}{2} \rfloor} \binom{3^t - j}{i} 2^i,$$

*and the number of ternary $3^n$-periodic sequences $S$ with $k$-error linear complexity $3^n - 3^{t+1} < L_k(S) \leq 3^n - 2\cdot 3^t$ is given by*

$$S_I = 3^{3^n - 2\cdot 3^t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor \frac{k-j}{2} \rfloor} \binom{3^t - j}{i} 2^i - 3^{3^n - 3^{t+1}} \sum_{j=0}^{k} \binom{3^{t+1}}{j} 2^j.$$

With Proposition 6 we can improve (1) in the ternary case.

**Corollary 2** *The expected $k$-error linear complexity $E_k$ of a random $3^n$-periodic ternary sequence satisfies*

$$3^n - 3^{\lfloor \log_3 k \rfloor} - 1 - \sum_{t=\lfloor \log_3 k \rfloor + 1}^{n-1} 3^{-3^t}(3^{t-1}+1) \sum_{j=0}^{k} \binom{3^t}{j} 2^j -$$
$$\frac{3^{n-1}+2}{3^{3^n}} \sum_{j=0}^{k} \binom{3^n}{j} 2^j -$$
$$\sum_{t=\lfloor \log_3 k \rfloor}^{n-1} (3^t - 1) 3^{-2\cdot 3^t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor (k-j)/2 \rfloor} \binom{3^t - j}{i} 2^i \quad \geq$$

$$E_n \geq 3^n - 2\cdot 3^{\lfloor \log_3 k \rfloor} + 1 - \sum_{t=\lfloor \log_3 k \rfloor + 1}^{n-1} 3^{-3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 2^j - \frac{1}{3^{3^n}} \sum_{j=0}^{k} \binom{3^n}{j} 2^j -$$
$$\sum_{t=\lfloor \log_3 k \rfloor}^{n-1} 3^{-2\cdot 3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor (k-j)/2 \rfloor} \binom{3^t - j}{i} 2^i. \quad (6)$$

*Proof.* We solely prove the lower bound. If we put $\lfloor \log_3 k \rfloor = l$, then

$$\begin{aligned} 3^{3^n} E_k &\geq \sum_{t=l}^{n-1} S_I(3^n - 3^{t+1} + 1) + S_{II}(3^n - 2 \cdot 3^t + 1) = \\ &\sum_{t=l}^{n-1} (3^n - 3^{t+1} + 1)(S_I + S_{II}) + \sum_{t=l}^{n-1} 3^t S_{II} := A_1 + A_2. \end{aligned}$$

Since $S_I + S_{II} = \mathcal{M}(t+1, t)$, the term $A_1$ is exactly the term for the lower bound obtained in Corollary 1 for $q = 3$. For $A_2$ we get

$$A_2 = \sum_{t=l}^{n-1} 3^{3^n - 3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 2^j - \sum_{t=l}^{n-1} 3^{3^n - 2\cdot 3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor (k-j)/2 \rfloor} \binom{3^t - j}{i} 2^i.$$

Combining the terms we obtain

$$\begin{aligned} 3^{3^n} E_k &\geq 3^{3^n}(3^n + 1) - 3^{3^n} 3^{l+1} - \sum_{j=0}^{k} \binom{3^n}{j} 2^j + 3^{3^n} 3^{-3^l + l} 3^{3^l} \\ &\quad -3^{3^n} \sum_{t=l+1}^{n-1} 3^{-3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 2^j \\ &\quad -3^{3^n} \sum_{t=l}^{n-1} 3^{-2\cdot 3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor (k-j)/2 \rfloor} \binom{3^t - j}{i} 2^i \\ &= 3^{3^n}(3^n + 1 - 3^{l+1} + 3^l) - \sum_{j=0}^{k} \binom{3^n}{j} 2^j - 3^{3^n} \sum_{t=l+1}^{n-1} 3^{-3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 2^j \\ &\quad -3^{3^n} \sum_{t=l}^{n-1} 3^{-2\cdot 3^t + t} \sum_{j=0}^{k} \binom{3^t}{j} 6^j \sum_{i=0}^{\lfloor (k-j)/2 \rfloor} \binom{3^t - j}{i} 2^i, \end{aligned}$$

which yields the desired formula. □

Table 1: Example to the ternary case, $N = 243$: $k$ is given as absolute value and percentage of $N$, the bounds are given relative to the period length $N$. New Lower Bound (NLB) and New Upper Bound (NUB) refer to the bounds (6), Old Lower Bound (OLB) refers to the bound (1).

| $k$ | 2 | 3 | 6 | 10 | 15 | 20 | 25 | 30 | 40 | 50 |
|---|---|---|---|---|---|---|---|---|---|---|
| $k\%$ | 0.82 | 1.24 | 2.47 | 4.12 | 6.17 | 8.23 | 10.29 | 12.35 | 16.46 | 20.58 |
| NLB | 0.98 | 0.97 | 0.94 | 0.907 | 0.88 | 0.8 | 0.78 | 0.72 | 0.67 | 0.6 |
| NUB | 0.984 | 0.978 | 0.96 | 0.94 | 0.92 | 0.89 | 0.88 | 0.82 | 0.78 | 0.75 |
| OLB | 0.95 | 0.93 | 0.88 | 0.82 | 0.75 | 0.69 | 0.64 | 0.585 | 0.49 | 0.41 |

(Table, file plot.eps)

# 4 Conclusion

The linear complexity and the $k$-error linear complexity are important but still not completely understood quality measures for sequences over finite fields. Until now exact formulas for the number of $N$-periodic sequences with given $k$-error linear complexity and for the expected $k$-error linear complexity are basically just known for $k = 0$ (see [8, 9]). Specifically, $p^n$-periodic sequences over a finite field $\mathbb{F}_q$ with characteristic $p$ have been studied from several viewpoints (see [1]–[6], [12]). In this contribution we provide the exact counting function and the expected value for the 1-error linear complexity for the case that $N = p^n$ and $q = p$. The results are a generalization of the results on the binary case presented in [7]. We emphasize that this generalization is not straightforward. Instead of the Chan-Games algorithm which works for the binary case, the more sophisticated algorithm by Ding et *al.*, which generalized the Chan-Games algorithm to arbitrary finite fields has to be analyzed.

It seems to be very difficult to obtain exact results for larger $k$. Our method permits the calculation of lower and upper bounds for the $k$ error linear complexity of $p^n$-periodic sequences over $\mathbb{F}_p$, $p$ prime. Until now, only lower bounds have been known. Finally we indicate how a refined analysis can provide an improvement of the bounds. The fact that the calculations become infeasible if $p$ is not very small, points out that it may be difficult to obtain exact results for larger $k$.

## References


[1] C. Ding, G. Xiao, and W. Shan, The Stability Theory of Stream Ciphers, Lecture Notes in Computer Science 561, Springer-Verlag, Berlin-Heidelberg, New York (1991).

[2] R. A. Games, A. H. Chan, A fast algorithm for determining the complexity of a binary sequence with period $2^n$, IEEE Trans. Inform. Theory 29 (1983), pp. 144–146.

[3] T. Kaida, S. Uehara, and K. Imamura, A new algorithm for the $k$-error linear complexity of sequences over $GF(p^m)$ with period $p^n$, Sequences and Their Applications (C. Ding, T. Helleseth and H. Niederreiter, eds.), Springer-Verlag, London, 1999, pp. 284–296.

[4] T. Kaida, On the generalized Lauder-Paterson algorithm and profiles of the $k$-error linear complexity over $GF(3)$ with period 9, Proceedings (extended abstracts) of the international conference on Sequences and Their Applications 2004, Seoul, Oct. pp. 24–28.

[5] K. Kurosawa, F. Sato, T. Sakata, and W. Kishimoto, A relationship between linear complexity and $k$-error linear complexity, IEEE Trans. Inform. Theory 46 (2000), pp. 694–698.

[6] A. G. B. Lauder, K. G. Paterson, Computing the linear complexity spectrum of a binary sequence of period $2^n$, IEEE Trans. Inform. Theory 49 (2003), pp. 273–280.

[7] W. Meidl, On the stability of $2^n$-periodic binary sequences, IEEE Trans. Inform. Theory 51 (2005), pp. 1151–1155.

[8] W. Meidl and H. Niederreiter, Linear complexity, $k$-error linear complexity, and the discrete Fourier transform, J. Complexity 18 (2002), pp. 87–103.

[9] W. Meidl, H. Niederreiter, On the expected value of the linear complexity and the $k$-error linear complexity of periodic sequences, IEEE Trans. Inform. Theory 48 (2002), pp. 2817–2825.

[10] H. Niederreiter, Linear complexity and related complexity measures for sequences, Progress in Cryptology - Proceedings of INDOCRYPT 2003 (T. Johansson and S. Maitra, eds.), Lecture Notes in Computer Science, Springer-Verlag, Berlin, 2904 (2003), pp. 1–17.

[11] R.A. Rueppel, Analysis and Design of Stream Ciphers, Springer-Verlag, Berlin (1986).

[12] M. Stamp, C. F. Martin, An algorithm for the $k$-error linear complexity of binary sequences with period $2^n$, IEEE Trans. Inform. Theory 39 (1993), pp. 1398–1401.