Linear complexity of binary sequences with optimal autocorrelation magnitude of length $4N$

V A Edemskiy and A B Minin
Yaroslav-the-Wise Novgorod State University, ul. B. St. Petersburgskaya, 41
173003 Veliky Novgorod, Russia
E-mail: Vladimir.Edemsky@novsu.ru

Abstract. In this paper, we study the linear complexity of series of binary sequences with optimal autocorrelation magnitude of length $4N$. These sequences are obtained from almost-perfect binary sequences and binary sequences with optimal autocorrelation of length $N$. The construction of these sequences were presented E.I. Krengel and P.V. Ivanov. We show that considered sequences have the high linear complexity. Also we derive the 1-error linear complexity of these sequences.

1. Introduction
Balanced or almost balanced binary sequences are widely used in digital communication, radar systems and stream ciphering [1, 2]. The autocorrelation of a binary sequence $x = \{x_i\}$ of period $T$ at shift $\tau$ is defined by

$$C_x(\tau) = \sum_{i=0}^{T-1} (-1)^{x_i+x_{i+\tau}}.$$

It is well known that $C_x(\tau) \equiv 4 - T \pmod{4}$ (see for example [2]). In particular, if $T = 4N$ then $C_x(\tau) \equiv 0 \pmod{4}$. Here $N$ is a positive integer. The binary sequence of length $4N$ is called the sequence with optimal autocorrelation when its out-of-phase autocorrelation coefficients belong to the set $\{0, -4\}$ and the sequence with optimal autocorrelation magnitude when they belong to the set $\{0, \pm 4\}$.

An important problem in sequence design is to find sequences with optimal autocorrelation and with optimal autocorrelation magnitude. Sequences with optimal autocorrelation magnitude of odd period have been studied in a lot of papers last years.

Two constructions of new binary sequences with optimal autocorrelation magnitude of length $4N, N \equiv 2 \pmod{4}$ were presented by E.I. Krengel and P.V. Ivanov in [3]. The first construction of E.I. Krengel and P.V. Ivanov uses the binary Sidelnikov (Sidelnikov-Lempel-Cohn-Eastman [4, 5]) sequences whereas the second one uses the binary Ding-Helleseth-Martinsen sequences [6].

The linear complexity is also a significant characteristic of binary sequences [7]. It is considered as primary quality measure for periodic sequences. The concept of the linear complexity is very useful in studying the security of stream ciphers. Sequences with optimal autocorrelation property and with large linear complexity are needed in certain communication systems and cryptography. In [3], authors calculated the linear complexity of several new constructed sequences and they assume that considered sequences have high linear complexity. In this paper, we will try to prove the conjecture from [3]. Also we study the 1-error linear complexity of these sequences. The denotation of $k$-error linear complexity of binary sequences was introduced by Stamp M., and Martin C. F. in [8]. The linear complexity and the $k$-error linear complexity are important cryptographic characteristics of sequences and provide...
information on the predictability and thus unsuitability of a sequence for cryptography. The remainder of this paper is organized as follows. In Section 2 we introduce some basics and recall the definition of sequences. Sections 3 and 4 are dedicated to the study of the linear complexity of new sequences and 1-error linear complexity. Section 5 concludes the work in this paper.

2. Preliminaries

To that end, we first briefly recall the definitions from [3] and the definitions of sequences which will be considered here.

The definition of almost perfect sequences was introduced by J. Wolfmann [9]. A sequence is called an almost perfect if all of its out-of-phase autocorrelation coefficients except one are 0. It is worth pointing out that D. Jungnickel, A. Pott offered another definition of almost perfect sequences [10]. The sequence is called almost perfect if all of their off-peak autocorrelation coefficients are as small as theoretically possible - with exactly one exception. We will use the definition of Wolfmann here. In recent years, almost perfect sequences have attracted more attention and they have been applied successfully to the extremely low power over-the-horizon radar.

Almost perfect binary sequences of a length $2N = 2(p_1^m + 1)$ were studied by Wolfmann, Langevin, and Bradley and Pott [11] (see also references here).

Let $p_1$ be an odd prime and $m \geq 1$ be a positive integer. Throughout this paper, we will denote by $GF(p_1^m)$ the finite field of order $p_1^m$ for a positive integer $m \geq 1$. The multiplicative group of this finite field is cyclic. Since $GF^∗(p_1^m)$ is cyclic, we denote by $\alpha$ the primitive element of the finite field $GF(p_1^m)$ and $\beta = \alpha^{p_1^{m+1}}$. It is easy to prove that $\beta \in GF(p_1^m)$ and $\beta$ is the primitive element of $GF(p_1^m)$.

The almost perfect binary sequence $x$ can be defined as

$$x_i = \begin{cases} 1, & \text{if } \text{Tr}(\eta \alpha^i) \neq 0 \text{ and ind}_\beta \text{Tr}(\eta \alpha^i) \equiv 0 \pmod{2}, \\ 0, & \text{otherwise}, \end{cases} \quad (1)$$

where $\eta \in GF(p_1^m), \eta \neq 0$, $\text{ind}_\beta z$ is the discrete logarithm $z$ to the base $\beta$, and

$$\text{Tr} x = x + x^p + \cdots + x^{p^{m-1}}, \quad x \in GF(p_1^m)$$

is the trace function from $GF(p_1^m)$ in $GF(p_1^m)$. The linear complexity the almost perfect binary sequences of length $2(p_1^m + 1)$ was studied earlier in [12].

Now we recall a definition of Sidelnikov sequences. These sequences are also called of Sidelnikov-Lempel-Cohn-Eastman sequences. Let $p_2$ be an odd prime and $k \geq 1$ be a positive integer. Denote by $\gamma$ the primitive element of the finite field $GF(p_2^k)$. Then the balanced binary Sidelnikov sequence of length $p_2^k - 1$ is defined by

$$y_i = \begin{cases} 1, & \text{if } \gamma^i + 1 \text{ is a square in } GF(p_2^k), \\ 0, & \text{otherwise}, \end{cases} \quad (2)$$

where $i = 0, \ldots, p_2^k - 2$ [5]. Then $y$ has optimal autocorrelation.

Ding et al. give several new families of binary sequences of even period [6]. Their construction is based on cyclotomy.

Let $p_3$ be a prime of the form $p_3 \equiv 1 \pmod{4}$, and let $\theta$ be a primitive root modulo $p_3$. By definition, put

$$D_0 = \{\theta^{4s} \mod p_3; s = 1, \ldots, (p_3 - 1)/4\}$$

and $D_n = \theta^n D_0, n = 1, 2, 3$ where the arithmetic is as in $GF(p_3)$. Then $D_k$ are cyclotomic classes of order four modulo $p_3$ [13].

Let $\mathbb{Z}_m = \{0, 1, \ldots, m\}$ be the residue class ring modulo. It is well known that $\mathbb{Z}_{2p_3} \cong \mathbb{Z}_2 \times \mathbb{Z}_{p_3}$ relative to isomorphism $\phi(a) = (a \mod 2, a \mod p_3)$ [14].

Ding et al. considered sequences defined as

$$z_i = \begin{cases} 1, & \text{if } i \mod 2p_3 \in C, \\ 0, & \text{if } i \mod 2p_3 \notin C, \end{cases} \quad (3)$$
for $C = \phi^{-1} \left( \{ 0 \} \times (D_i \cup D_j) \cup \{ 1 \} \times (D_i \cup D_j) \right)$ where $i, j,$ and $l$ are pairwise distinct integers between $0$ and $3$ (also for $C^{(0)} = C \cup \{ 0 \}$) [6]. They are now known as sequences of Ding- Helleseth- Martinsen.

The autocorrelation of these sequences was studied in [6]. If $p_3 \equiv 1 \pmod{4}$ then $p_3$ can be expressed as the sum of squares of two integers. By [6], the sequence $z$ have an optimal autocorrelation value when $p_3 \equiv 5 \pmod{8}$ and $p_3 = 1 + 4b^2, (i, j, l) = (0, 1, 3); (0, 2, 1)$ or $p_3 = a^2 + 4, (i, j, l) = (1, 0, 3); (0, 1, 2)$. Here $a, b$ are integers and $a \equiv 1 \pmod{4}$. The linear complexity of Ding- Helleseth- Martinsen sequences was studied in [15].

As noted in [3] there exist $p_i, m, k$ such that $p_i^m + 1 = p_k^m - 1$ or $p_i^m + 1 = 2p_3$. For these values $p_i, m, k$ we form the sequence $v = y \cdot y$ of length $2N = 2(p_i^m + 1)$ using a concatenation $y$ when $p_i^m + 1 = p_k^m - 1$ and $v = z \cdot z$ when $p_i^m + 1 = 2p_3$. It means that, $v_i = y_{i-N}$ or $v_i = z_{i-N}$ when $i = 2N, 2N + 1, ..., 4N - 1$.

Krengel and Ivanov presented new construction interleaved binary sequence $w$ with period $4N$ which defined as

$$w_i = \begin{cases} v_i, & \text{if } i = 2k, \\ x_i, & \text{if } i = 2k + 1, \\ k = 0, 1, ..., 2N - 1. \end{cases}$$

By [3] $w$ has optimal autocorrelation magnitude. Krengel and Ivanov considered a few examples of these sequences for different values of $p_i, m, k$ and calculated their linear complexity. These examples showed that new constructed sequences have high linear complexity. So, authors of [3] assume that their new sequences have high linear complexity and they can be applied in cryptography. We proof here that the linear complexity of these sequences is equal to their length, i.e., these sequences really have high linear complexity.

3. The linear complexity

The linear complexity $L$ of a sequence is an important parameter in its evaluation as a keystream cipher for cryptographic applications as a measure of its predictability. A high linear complexity is necessary for a good cryptographic sequence.

The linear complexity (also called linear span) of $u$ over the finite field $GF(2)$ is defined to be the smallest positive integer $L$ such that there are constants $c_1, ..., c_L \in GF(2)$ satisfying

$$u_i = c_1u_{i-1} + c_2u_{i-2} + ... + c_Lu_{i-L}$$

for all $i \geq L$. The sequences satisfying this relation, are called linear recurring sequences. Linear recurring sequences over fields are well-known subjects of research in applied algebra and discrete mathematics, dating back to Fibonacci. The linear recurring sequences are used in radar-location, coding theory, generation of pseudo-random numbers, etc.

In engineering terms, $L$ is the length of the shortest linear feedback shift register that can generate sequence $u$. Knowledge of just $2L$ consecutive digits of the sequence is sufficient to enable the remainder of the sequence to be constructed. Thus, it is reasonable to suggest that ‘good’ sequences have $L$ bigger than a half of the period of the sequence [16].

Throughout this paper $S_u(t)$ denotes the polynomial of sequence $u$, i.e., $S_u(t) = \sum_{t=0}^{M-1} u_t t^i$ where $M$ is a length of sequence.

It is a familiar fact that the linear complexity $L$ and the minimal polynomial $m(t)$ of $w$ can be determined by the following formulas [16]:

$$L = 4N - \deg \left( \gcd (x^{4N} - 1, S_w(t)) \right),$$

$$m(t) = (x^{4N} - 1)/\gcd(x^{4N} - 1, S_w(t)).$$

So, in this case we need to study the greatest common divisor of two polynomial.

It is worth pointing out that the minimal polynomials of $m(t)$ defined here may be the reciprocals of the minimal polynomials defined in other references.

Let $x$ be the almost perfect binary sequence of length $2N$ defined by (1). By the definition of this sequence we see that $x_i + x_{i+N} = 1$ for $0 \leq i < N$ except two positions and there exist $j$ and $j + N$ such that $x_j = x_{j+N} = 0$. 

3
Let us introduce the auxiliary polynomial $T(t) = \sum_{i=0}^{N-1} x^i t^i$. Before we give the main result of this section, we establish the following lemma.

**Lemma 1.** Let $x$ be a binary sequence of length $2N$ defined by (1) and let $j$: $x_j = x_{j+N} = 0$. Then

$$S_x(t) = T(t)(1 - t^N) + t^{j+N} + t^N(t^N - 1)/(t - 1).$$

**Proof.** From our definition it follows that

$$S_x(t) = \sum_{i=0}^{N-1} x_i t^i = T(t) + \sum_{i=0}^{2N-1} x_i t^i.$$

By [12] we obtain

$$\sum_{i=0}^{2N-1} x_i t^i = t^N \sum_{i=0}^{N-1} (1 - x_i) t^i - t^{j+N} = t^N (1 + t + \cdots + t^{N-1}) - T(t)t^N - t^{j+N}.$$

This completes the proof of Lemma 1.

The following statement is the main result in this paper.

**Theorem 1.** Let $w$ be the binary sequence of length $4N$ defined by (1). Then the linear complexity of $w$ over $GF(2)$ is equal to $L = 4N$ and $m(t) = (t^N - 1)^4$.

**Proof.** It is straightforward to verify that $S_w(t) = S_y(t^2) + tS_z(t^2)$.

Since by definition

$$S_y(t) = (1 + t^N)S_y(t) \text{ or } S_y(t) = (1 + t^N)S_z(t),$$

it follows by Lemma 1 that

$$S_w(t) = (1 + t^{2N})S_y(t^2) + tT(t^2)(1 - t^{2N}) + t^{2j+2N+1} + \frac{t^{2N+1}(t^{2N+1} - 1)}{t^2 - 1}. \tag{6}$$

A proof of this equality can be easily carried out by computation of $S_w(t)$, which is omitted here.

By (6) we see that $S_w(1) = 1$ and the polynomial $S_w(t)$ is not divided by $1 + t + \cdots + t^{N-1}$. Hence, we obtain that $\gcd(t^{4N} - 1, S_w(t)) = 1$ over $GF(2)$. The conclusion of this theorem then follows from (5).

So, by Theorem 1 we see that sequence $w$ have high linear complexity. It is worth pointing out that properties of sequences $y, z$ are not used in the proof of this theorem. Hence, we can use for obtaining sequences with high linear complexity also Ding- Helleseth- Martinsen’s sequences when $p_3 \neq 1 + 4b^2$ or $p_3 \neq a^2 + 4$. But in this case the sequence $w$ will not be a sequence with the optimal autocorrelation magnitude.

4. The k-error linear complexity

For a sequence to be cryptographically strong, its linear complexity should be large, but not significantly reduced by changing a few terms. For example, the T-periodic sequence 0,0,0,...,0,1 has the highest linear complexity $T$. However, if the last bit of the sequence 0,0,0,...,0,1 is changed to 0, its linear complexity decreases to zero. Ding et al. noticed this problem first in their book, and proposed the weight complexity and sphere complexity. Stamp and Martin introduced the $k$-error linear complexity, which is similar to the sphere complexity, and put forward the concept of the $k$-error linear complexity profile.

The $k$-error linear complexity of a sequence $u$ of length $M$ is defined by $L_k(u) = \min_t L(t)$, where the minimum is taken over all binary $M$-periodic sequences $t$ for which the Hamming distance of the vectors $(t_0, t_1, \ldots, t_{M-1})$ and $(u_0, u_1, \ldots, u_{M-1})$ is at most $k$ [8]. It means that, the $k$-error linear complexity of a sequence $u$ is the smallest linear complexity over the finite field $GF(2)$ that is possible to obtain by changing at most $k$ terms of the sequence per period. Sequences that are suitable as keystreams should possess not only a large linear complexity but also the change of a few terms must not cause a significant decrease of the linear complexity.

**Lemma 2.** Let $w$ be a binary sequence of length $4N$ defined by (1). Then $L_1(w) = 2N + 2$.

**Proof.** Let $e$ be an error-sequence of $w$. It means that $e_t = 1$ if $w_t$ is changed when computing the $k$-error linear complexity of $w$, and otherwise $e_t = 0$.

It is well-known that the 1-error linear complexity of $w$ is computed by the following formula

$$L_1(w) = \min_{0 \leq \deg(e(t)) \leq 1} \{4N - \deg(\gcd(t^{4N} - 1, S_w + e(t)))\},$$

where $S_w(t)$ is the linear complexity of the $w$ sequence and $\deg(e(t))$ is the degree of the polynomial $e(t)$.
where $e(t) \in GF(2)[t]$ is the generating polynomial of an error-sequence and $wt(e(t))$ means the number of non-zero coefficients of a polynomial $e(t)$. To conclude the proof, it remains to note that $S_w(t) + t^{2j+2N+1} \equiv 0 (mod (t^{2N} - 1)/(t^2 - 1))$ by (6).

5. Conclusion
We proved the conjecture that the new binary sequences with optimal autocorrelation magnitude of length $4N, N \equiv 2 (mod 4)$ presented by E.I. Krengel and P.V. Ivanov in [3] really have high linear complexity. These sequences are obtained from almost-perfect binary sequences and Sidelnikov (Sidelnikov-Lempel-Cohn-Eastman) sequences or Ding-Helleseth-Martinsen binary sequences with optimal autocorrelation of length $N$. Also we derive the 1-error linear complexity of these sequences. Since the 1-error linear complexity of $w$ is close to the half of the period, it follows that these sequences should be selected carefully for use in stream cipher systems.

Acknowledgements
The reported research was funded by Russian Foundation for Basic Research and Novgorod region, grant no. 18-41-530001.

References
[1] Fan P and Darnell M 1996 Sequence design for communications applications (Research Studies Press Ltd., England, London)
[2] Golomb S W and Gong G 2005 Signal design for good correlation - for wireless communication, cryptography and radar (Cambridge University Press, USA)
[3] Krengel E I and Ivanov P V 2016 Two costructions of binary sequences with optimal autocorrelation magnitude Electron. Lett. 52 1457-59
[4] Sidelnikov V M 1969 Some k-valued pseudo-random sequences and nearly equidistant codes Probl. Inf. Transm. 5 12–16
[5] Lempel A, Cohn M and Eastman W 1977 A class of balanced binary sequences with optimal autocorrelation property IEEE Trans. Inf. Theory 23 (1) 38-42
[6] Ding C, Helleseth T and Martinsen H 2001 New families of binary sequences with optimal three-valued autocorrelation IEEE Trans. Inf. Theory 47 (1) 428-433
[7] Cusick T, Ding C and Renvall A 1998 Stream Ciphers and Number Theory (North-Holland Publishing Co., Amsterdam)
[8] Stamp M and Martin C F 1993 An algorithm for the k-error linear complexity of binary sequences with period $2^n$ IEEE Trans. Inform. Theory 39 1398-1401
[9] Wolfmann J 1992 Almost perfect autocorrelation sequences IEEE Trans. Inf. Theory 38 (4) 1412-18
[10] Jungnickel D and Pott A 1999 Perfect and almost perfect sequences Discrete Applied Mathematics 95 331–359
[11] Bradley S P and Pott P 1995 Existence and nonexistence of almostperfect autocorrelation sequences IEEE Trans. Inf. Theory 41 (1) 301–304
[12] Edemskiy V and Minin A 2016 About the linear complexity of the almost perfect sequences. International Journal of Communications 1 223–226
[13] Hall M 1975 Combinatorial Theory (Wiley, New York)
[14] Zhang J and Zhao C 2015 The linear complexity of a class of binary sequences with period $2p$ Appl. Algebra Eng. Commun. Comput. 26 (5) 475–491
[15] Ireland K and Rosen M 1982 Classical Introduction to Modern Number Theory (Springer)
[16] Lidl R and Niederreiter H 1983 Finite Fields (Addison-Wesley)