Arithmetic crosscorrelation of pseudorandom binary sequences of coprime periods

Zhixiong Chen[1], Zhihua Niu[1,2] and Arne Winterhof[3]

1. Key Laboratory of Applied Mathematics of Fujian Province University,
Putian University, Putian, Fujian 351100, P. R. China
ptczx@126.com
2. School of Computer Engineering and Science,
Shanghai University, Shanghai 200444, P. R. China
zhniu@shu.edu.cn
3. Johann Radon Institute for Computational and Applied Mathematics,
Austrian Academy of Sciences, Altenberger Straße 69, A-4040 Linz, Austria
arne.winterhof@oeaw.ac.at

March 23, 2022

Abstract

The (classical) crosscorrelation is an important measure of pseudorandomness of two binary sequences for applications in communications. The arithmetic crosscorrelation is another figure of merit introduced by Goresky and Klapper generalizing Mandelbaum's arithmetic autocorrelation.

First we observe that the arithmetic crosscorrelation is constant for two binary sequences of coprime periods which complements the analogous result for the classical crosscorrelation.

Then we prove upper bounds for the constant arithmetic crosscorrelation of two Legendre sequences of different periods and of two binary m-sequences of coprime periods, respectively.

1 Introduction

Sequences with good correlation properties are essential ingredients in a wide range of applications including cryptography, CDMA systems and radar ranging [3]. A great deal of research has gone into the design and generation of sequences and families of sequences with good correlation properties. For example, for CDMA we need large families of sequences with small pairwise correlations.

Let N be the common (minimal) period of two (periodic) binary sequences

$$\mathcal{S} = (s_i)_{i \geq 0} \quad \text{and} \quad \mathcal{T} = (t_i)_{i \geq 0}$$

over the binary field $\mathbb{F}_2 = \{0, 1\}$. The *(classical) periodic crosscorrelation* of $\mathcal{S}$ and $\mathcal{T}$ at lag τ, denoted by $\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau)$, is defined by

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau) = \sum_{0 \leq i < N} (-1)^{s_i + t_{i+\tau}}, \; 0 \leq \tau < N.$$

For $\mathcal{S} = \mathcal{T}$, it is called the *(classical) periodic autocorrelation of $\mathcal{S}$ at τ*, which is denoted by

$$\mathcal{A}_{\mathcal{S}}(\tau) = \mathcal{C}_{\mathcal{S},\mathcal{S}}(\tau), \quad 0 \leq \tau < N.$$

A different notion of autocorrelation is the arithmetic autocorrelation introduced by Mandelbaum [12] and later generalized to the arithmetic crosscorrelation by Goresky and Klapper [4]. (Note that Mandelbaum did not use the term arithmetic autocorrelation.) In the *arithmetic crosscorrelation*, a sequence is added to a shift of another one with carry, rather than bit by bit modulo 2. According to [4, Proposition 2] or the discussions in [9, 10], we can demonstrate the computation (of the arithmetic crosscorrelation of $\mathcal{S}$ and $\mathcal{T}$) as follows:

Write

$$S(2) = \sum_{0 \leq i < N} s_i 2^i, \quad T^{(\tau)}(2) = \sum_{0 \leq i < N} t_{i+\tau} 2^i.$$

We compute $S(2) - T^{(\tau)}(2)$ in $\mathbb{Z}$. If $S(2) - T^{(\tau)}(2) \geq 0$, we consider the unique binary expansion of $S(2) - T^{(\tau)}(2)$:

$$S(2) - T^{(\tau)}(2) = \sum_{0 \leq i < N} w_i 2^i, \; w_i \in \{0, 1\}.$$

If $S(2) - T^{(\tau)}(2) < 0$, we consider the binary expansion of $2^N - 1 + S(2) - T^{(\tau)}(2) \geq 0$:

$$2^N - 1 + S(2) - T^{(\tau)}(2) = \sum_{0 \leq i < N} w_i 2^i, \; w_i \in \{0, 1\}.$$

Then we compute the arithmetic crosscorrelation of $\mathcal{S}$ and $\mathcal{T}$ at τ, denoted by $\mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau)$:

$$\mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau) = N_0 - N_1 = 2N_0 - N = N - 2N_1, \tag{1}$$

where for $j \in \{0, 1\}$, N_j is the number of $i = 0, 1, \ldots, N-1$ with $w_i = j$. For $\mathcal{S} = \mathcal{T}$, this is the *arithmetic autocorrelation* of $\mathcal{S}$ at τ, denoted by

$$\mathcal{A}_{\mathcal{S}}^A(\tau) = \mathcal{C}_{\mathcal{S},\mathcal{S}}^A(\tau).$$

The reader is referred to the research papers by Goresky and Klapper [4–6] or their monograph [7] for more background and results on arithmetic auto-/crosscorrelation.

It is desirable that the absolute values of both the classical and the arithmetic cross-/autocorrelations are as small as possible for $1 \le \tau < N$. However, sequences with small $\max\limits_{1 \le \tau < N} |\mathcal{A}_{\mathcal{S}}(\tau)|$ may have large $\max\limits_{1 \le \tau < N} |\mathcal{A}_{\mathcal{S}}^A(\tau)|$ and vice versa.

For example, any m-sequence $\mathcal{S}$ produced by an n-order linear feedback shift register (LFSR), that is of period $2^n - 1$ satisfies

$$\mathcal{A}_{\mathcal{S}}(\tau) = -1 \quad \text{and} \quad |\mathcal{A}_{\mathcal{S}}^A(\tau)| \le 2^{n-1} - 1, \quad 1 \le \tau < 2^n - 1,$$

see [8, p.764] and [1]. Numerical examples in [1] support the conjecture

$$\max_{1 \le \tau < 2^n - 1} |\mathcal{A}_{\mathcal{S}}^A(\tau)| = 2^{n-1} - 1.$$

In Section 3.3 below, we will give a bound on the absolute value of the arithmetic autocorrelation function of m-sequences for small lags τ, which is not considered in [1].

Any ℓ-sequence $\mathcal{S}$ produced by a feedback with carry shift register (FCSR) with the prime connection number p, that is of period $p-1$, can be defined by

$$s_i = (a2^{-i} \mod p) \mod 2,$$

for some $a \not\equiv 0 \pmod{p}$ and satisfies

$$\mathcal{A}_{\mathcal{S}}((p-1)/2) = p - 1,$$
$$\max_{\substack{1 \le \tau < p-1 \\ \tau \ne (p-1)/2}} |\mathcal{A}_{\mathcal{S}}(\tau)| = EL_3(p),$$
$$\mathcal{A}_{\mathcal{S}}^A(\tau) = 0, \quad 1 \le \tau < p - 1,$$

where $EL_3(p)$ is the greatest even number less than $p/3$, see [13, Theorem 5] and [7, Theorem 13.3.1].

It is easy to see that the classical crosscorrelation $\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau)$ is constant if the periods of $\mathcal{S}$ and $\mathcal{T}$ are coprime. Indeed, if p is the period of $\mathcal{S}$ and q is the period of $\mathcal{T}$ with $\gcd(p, q) = 1$, then we have by the Chinese Remainder Theorem

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau) = \sum_{i=0}^{pq-1} (-1)^{s_i + t_{i+\tau}} = \sum_{i_1=0}^{p-1} (-1)^{s_{i_1}} \sum_{i_2=0}^{q-1} (-1)^{t_{i_2 + \tau}},$$

which is constant since the sum over i_2 is independent of τ. In particular, if both $\mathcal{S}$ and $\mathcal{T}$ are balanced with odd pq (which is true for example for Legendre sequences

3

of period $p \equiv 3 \pmod 4$ and m-sequences), we get $|\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau)| = 1$ for $0 \leq \tau < pq$. If one of the sequences $\mathcal{S}$ and $\mathcal{T}$ is balanced with even period, we get $\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau) = 0$ for $0 \leq \tau < pq$.

In this work, first we will prove that the arithmetic crosscorrelation $\mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau)$ is also constant under the same assumption that the periods of $\mathcal{S}$ and $\mathcal{T}$ are coprime, see Section 2. Then we consider two Legendre sequences of different periods and two binary m-sequences of coprime periods and derive upper bounds on their constant arithmetic crosscorrelation in Sections 3.1 and 3.2, respectively. Finally we provide some numerical data in Section 4.

Occassionally, we will use negative indices for N-periodic sequences $\mathcal{S} = (s_i)_{i \geq 0}$ defined in the obvious way,

$$s_{-i} = s_{N-i}, \quad i = 0, 1, \ldots.$$

We use the notation $f(n) = O(g(n))$ or $f(n) \ll g(n)$ if there is an absolute constant $c > 0$ such that $|f(n)| \leq cg(n)$.

2 Arithmetic crosscorrelation

In this section we prove the following result on the constant arithmetic crosscorrelation.

Theorem 1. *Let $\mathcal{S} = (s_i)_{i \geq 0}$ and $\mathcal{T} = (t_i)_{i \geq 0}$ be two sequences over $\mathbb{F}_2$ of periods $p > 1$ and $q > 1$, respectively. If p and q are coprime, then the arithmetic crosscorrelation $\mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau)$ has the same value for any $0 \leq \tau < pq$.*

Proof. We consider the unique binary expansion of a non-negative integer $W < 2^r - 1$:

$$W = w_0 + w_1 2 + \cdots + w_{r-1} 2^{r-1}, \quad w_i \in \{0, 1\}, \ 0 \leq i < r,$$

and define the *weight* of W, denoted by $wt(W)$, as

$$wt(W) = \sum_{i=0}^{r-1} w_i.$$

It is clear that

$$wt(2^k W) = wt(W) \quad \text{for any integer } k \geq 0.$$

Furthermore, for $k \geq 0$ let W_k be defined by

$$W_k \equiv 2^k W \pmod{2^r - 1}, \quad 0 \leq W_k < 2^r - 1.$$

Then we have

$$\begin{aligned} W_k &\equiv w_0 2^k + w_1 2^{k+1} + \cdots + w_{r-1-k} 2^{r-1} \\ &\quad + w_{r-k} + w_{r-k+1} 2^1 + \cdots + w_{r-1} 2^{k-1} \pmod{2^r - 1}. \end{aligned}$$

4

and thus
$$wt(W_k) = wt(W), \quad k \geq 0. \tag{2}$$

Since otherwise the result is trivial, we may assume that $\mathcal{T}$ is not constant. The common minimal period of $\mathcal{S}$ and $\mathcal{T}$ is $N = pq$. According to the definition of the arithmetic crosscorrelation, we need to determine the weight of $S(2) - T^{(\tau)}(2)$ or $2^{pq} - 1 + S(2) - T^{(\tau)}(2)$ for $0 \leq \tau < pq$. We remark that

$$-(2^{pq} - 1) < S(2) - T^{(\tau)}(2) < 2^{pq} - 1,$$

since $0 \leq S(2) \leq 2^{pq} - 1$ and $0 < T^{(\tau)}(2) < 2^{pq} - 1$.

Put
$$\lambda_k = \sum_{m=0}^{p-1} (s_m - t_{m+kp}) 2^m, \quad k \geq 0.$$

It is easy to see that
$$\lambda_{k+q} = \lambda_k, \quad k \geq 0.$$

Case 1. We assume $S(2) - T^{(\tau)}(2) \geq 0$.
Take $x \in \{0, 1, \ldots, q-1\}$ with $x \equiv \tau p^{-1} \pmod{q}$. We substitute

$$n = n_1 + n_2 p \quad \text{and} \quad n_3 = x + n_2$$

to get

$$
\begin{aligned}
2^{xp}(S(2) - T^{(\tau)}(2)) &\equiv 2^{xp} \sum_{n=0}^{pq-1} (s_n - t_{n+\tau}) 2^n, \\
&\equiv 2^{xp} \sum_{n_2=0}^{q-1} \left(\sum_{n_1=0}^{p-1} (s_{n_1} - t_{n_1+(x+n_2)p}) 2^{n_1} \right) 2^{n_2 p}. \\
&\equiv \sum_{n_3=x}^{q-1+x} \left(\sum_{n_1=0}^{p-1} (s_{n_1} - t_{n_1+n_3 p}) 2^{n_1} \right) 2^{n_3 p} \\
&\equiv \sum_{n_3=x}^{q-1+x} \lambda_{n_3} 2^{n_3 p} \\
&\equiv \sum_{n_3=0}^{q-1} \lambda_{n_3} 2^{n_3 p} \pmod{2^{pq} - 1},
\end{aligned}
$$

since $\lambda_{k+q} = \lambda_k$. We note that $\sum_{n_3=0}^{q-1} \lambda_{n_3} 2^{n_3 p}$ is independent of τ and we have a fixed number Ω with $0 \leq \Omega < 2^{pq} - 1$ such that

$$\Omega \equiv \sum_{n_3=0}^{q-1} \lambda_{n_3} 2^{n_3 p} \pmod{2^{pq} - 1}.$$

5

Then by (2), we derive for any τ with $S(2) - T^{(\tau)}(2) \geq 0$,

$$wt(S(2) - T^{(\tau)}(2)) = wt(2^{xp}(S(2) - T^{(\tau)}(2))) = wt(\Omega).$$

Case 2. We assume $S(2) - T^{(\tau)}(2) < 0$.
We need to determine the weight of $2^{pq} - 1 + S(2) - T^{(\tau)}(2)$, which is a non-negative number smaller than $2^{pq} - 1$. Since for $x \in \{0, 1, \ldots, q-1\}$ with $x \equiv \tau p^{-1} \pmod{q}$,

$$2^{xp}(2^{pq} - 1 + S(2) - T^{(\tau)}(2)) \equiv 2^{xp}(S(2) - T^{(\tau)}(2)) \pmod{2^{pq} - 1},$$

we follow the proof in Case 1 to get

$$2^{xp}(2^{pq} - 1 + S(2) - T^{(\tau)}(2)) \equiv \sum_{n_3=0}^{q-1} \lambda_{n_3} 2^{n_3 p} \equiv \Omega \pmod{2^{pq} - 1},$$

from which we derive

$$wt(2^{pq} - 1 + S(2) - T^{(\tau)}(2)) = wt(2^{xp}(2^{pq} - 1 + S(2) - T^{(\tau)}(2))) = wt(\Omega)$$

for any τ with $S(2) - T^{(\tau)}(2) < 0$.

Putting both cases together, we obtain by (1)

$$
\begin{aligned}
\mathcal{C}_{\mathcal{S},\mathcal{T}}^{A}(\tau) &= pq - \begin{cases} 2wt(S(2) - T^{(\tau)}(2)), & \text{if } S(2) - T^{(\tau)}(2) \geq 0, \\ 2wt(2^{pq} - 1 + S(2) - T^{(\tau)}(2)), & \text{otherwise,} \end{cases} \\
&= pq - 2wt(\Omega).
\end{aligned}
$$

So $\mathcal{C}_{\mathcal{S},\mathcal{T}}^{A}(\tau)$ is constant which completes the proof. $\qquad\square$

3 Upper bounds for special pairs of binary sequences

In this section, we will prove upper bounds on the arithmetic crosscorrelation for two binary Legendre sequences of different periods and two binary m-sequences of coprime periods, respectively. For results on their arithmetic autocorrelation, we refer the reader to [1, 10].

3.1 Arithmetic crosscorrelation of two binary Legendre sequences of different periods

For a prime $p > 2$ let (ℓ_n) be the *Legendre sequence* defined by

$$\ell_n = \begin{cases} 1, & \text{if } \left(\frac{n}{p}\right) = 1; \\ 0, & \text{otherwise,} \end{cases} \quad n \geq 0,$$

where $\left(\frac{\cdot}{\cdot}\right)$ is the Legendre symbol. Obviously, (ℓ_n) is p-periodic.

We need a preliminary result on the pattern distribution of two Legendre sequences.

Lemma 1. *Let p and q be primes with $2 < p < q$ and denote by $\mathcal{S} = (s_i)_{i \geq 0}$ and $\mathcal{T} = (t_i)_{i \geq 0}$ the Legendre sequences of periods p and q, respectively.*

For an integer $k \geq 0$ and any pattern $\underline{e} \in \{0,1\}^{2k+2}$, the number Σ_k of $i = 0, 1, \ldots, pq - 1$ with

$$(s_{i-k}, s_{i-k+1}, \ldots, s_i, t_{i-k}, t_{i-k+1}, \ldots, t_i) = \underline{e}$$

satisfies

$$\Sigma_k = \frac{pq}{2^{2k+2}} + O\left(2^{-k} k p^{1/2} q\right) \quad \text{for} \quad k \leq (0.5 \log p - \log \log p) / \log 2.$$

Proof. Write $\underline{e} = (e_0, \ldots, e_{2k+1})$ and put

$$\delta_{i,j,p} = 1 - (-1)^{e_{k-j}} \left(\frac{i-j}{p}\right) \quad \text{and} \quad \delta_{i,j,q} = 1 - (-1)^{e_{2k+1-j}} \left(\frac{i-j}{q}\right).$$

Note that for i and $j \leq k$ with $\gcd(i-j, pq) = 1$ we have

$$\delta_{i,j,p} \delta_{i,j,q} = \begin{cases} 4, & \text{if } (s_{i-j}, t_{i-j}) = (e_{k-j}, e_{2k+1-j}), \\ 0, & \text{otherwise}, \end{cases}$$

and thus

$$\Sigma_k = \frac{1}{2^{2k+2}} \sum_{i=0}^{pq-1} \prod_{j=0}^{k} \delta_{i,j,p} \delta_{i,j,q} + O(kq),$$

where the O-term comes from those i with $\gcd(i-j, pq) > 1$ for some $j = 0, 1, \ldots, k$. Expanding the product we get

$$\Sigma_k = \frac{1}{2^{2k+2}} \sum_{U,V \subseteq \{0,1,\ldots,k\}} \sum_{i=0}^{pq-1} \prod_{j \in U} (-1)^{e_{k-j}} \left(\frac{i-j}{p}\right) \prod_{j \in V} (-1)^{e_{2k+1-j}} \left(\frac{i-j}{q}\right) + O(kq).$$

The contribution to Σ_k of $U = V = \emptyset$ is trivially

$$\frac{pq}{2^{2k+2}}.$$

The contribution for $U = \emptyset$ and the $(2^{k+1} - 1)$ sets $V \neq \emptyset$ is bounded by

$$\frac{1}{2^{2k+2}} \cdot (2^{k+1} - 1) \cdot p \max_{V \neq \emptyset} \left| \sum_{i=0}^{q-1} \left(\frac{\prod_{j \in V} (i-j)}{q} \right) \right| = O\left(2^{-k} k p q^{1/2}\right)$$

by the Weil bound, see for example [11, Theorem 5.41]. Analogously, the contribution to Σ_k for $V = \emptyset$ and the $(2^{k+1} - 1)$ sets $U \neq \emptyset$ is

$$O\left(2^{-k} k p^{1/2} q\right).$$

The contribution of the $2^{2k+2} - 2^{k+2} + 1$ remaining (U, V) with $U \neq \emptyset$ and $V \neq \emptyset$ is bounded by

$$\frac{1}{2^{2k+2}} \cdot (2^{2k+2} - 2^{k+2} + 1) \cdot \max_{U,V \neq \emptyset} \left| \sum_{i=0}^{pq-1} \prod_{j \in U} \left(\frac{i-j}{p} \right) \prod_{j \in V} \left(\frac{i-j}{q} \right) \right|$$

$$\leq \max_{U,V \neq \emptyset} \left| \sum_{i_1=0}^{p-1} \left(\frac{\prod_{j \in U} (i_1 - j)}{p} \right) \sum_{i_2=0}^{q-1} \left(\frac{\prod_{j \in V} (i_2 - j)}{q} \right) \right|$$

$$= O(k^2 (pq)^{1/2})$$

by the Chinese Remainder Theorem and the Weil bound. Collecting everything and verifying that

$$\max\{kq, 2^{-k} k p^{1/2} q, k^2 (pq)^{1/2}\} = 2^{-k} k p^{1/2} q \quad \text{for} \quad k \leq (0.5 \log p - \log \log p) / \log 2$$

we get the result. $\qquad \square$

Theorem 2. *Let $\mathcal{S} = (s_i)_{i \geq 0}$ and $\mathcal{T} = (t_i)_{i \geq 0}$ be two Legendre sequences over $\mathbb{F}_2$ of prime periods $p > 2$ and $q > 2$, respectively. If $p < q$, then the arithmetic crosscorrelation $\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau)$ satisfies*

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau) \ll p^{1/2} q (\log p)^2, \quad 0 \leq \tau < pq.$$

Proof. By Theorem 1 we may assume $\tau = 0$. Without loss of generality we may assume $S(2) \geq T^{(0)}(2)$. We write

$$S(2) - T^{(0)}(2) = \sum_{i=0}^{pq-1} (s_i - t_i) 2^i = \sum_{i=0}^{pq-1} w_i 2^i \quad \text{with} \quad w_i \in \{0, 1\}.$$

Note that the w_i are unique and we have to estimate the number of $i = 0, 1, \dots, pq-1$ with $w_i = 1$.

Assume that for some $n \geq k \geq 1$ and $a \in \{0, 1\}$

$$(s_{n-k}, t_{n-k}) = (a, 1 - a),$$
$$s_{n-k+j} = t_{n-k+j}, \quad j = 1, \dots, k-1,$$
$$(s_n, t_n) \in \{0, 1\}^2.$$

For $a = 1$ we have

$$2^{n-k+1} > \sum_{i=0}^{n-k} (s_i - t_i) 2^i \geq 2^{n-k} - \sum_{i=0}^{n-k-1} 2^i > 0$$

and thus w_{n-k+1} depends only on (s_{n-k+1}, t_{n-k+1}). Obviously, we have

$$w_{n-k+j} = s_{n-k+j} - t_{n-k+j} = 0 \quad \text{for} \quad j = 1, \dots, k-1$$

8

and $w_n = 1$ if and only if $s_n \neq t_n$.

For $a = 0$ we have

$$0 < 2^{n-k+1} + \sum_{i=0}^{n-k}(s_i - t_i)2^i < 2^{n-k+1}$$

and thus

$$w_{n-k+j} = 1 + s_{n-k+j} - t_{n-k+j} = 1 \quad \text{for} \quad j = 1, \ldots, k-1$$

and $w_n = 1$ if and only if $s_n = t_n$.

Altogether there are 2^{k+1} different patterns $\underline{e} \in \{0,1\}^{2k+2}$ such that

$$(s_{n-k}, s_{n-k+1}, \ldots, s_n, t_{n-k}, t_{n-k+1}, \ldots, t_n) = \underline{e} \tag{3}$$

implies $w_n = 1$. For each of these 2^{k+1} patterns $\underline{e}$ there are

$$\frac{pq}{2^{2k+2}} + O\left(2^{-k}kp^{1/2}q\right)$$

different $n = k, k+1, \ldots, k+pq-1$ satisfying (3) by Lemma 1. Hence, for each $k = 1, 2, \ldots$ there are at least

$$2^{k+1}\frac{pq}{2^{2k+2}} + O\left(\frac{2^{k+1}}{2^k}kp^{1/2}q\right) = \frac{pq}{2^{k+1}} + O\left(kp^{1/2}q\right)$$

different n in each fixed interval of length pq with $w_n = 1$. Choose

$$M = \left\lfloor \frac{\log p}{2\log 2} - \frac{2\log\log p}{\log 2} \right\rfloor \leq \log p.$$

Summing up, the number N_1 of $n = 0, 1, \ldots, pq-1$ with $w_n = 1$ is at least

$$
\begin{aligned}
N_1 &\geq \sum_{k=1}^{M} \frac{pq}{2^{k+1}} + O\left(\sum_{k=1}^{M} kp^{1/2}q\right) \\
&= \frac{pq}{4}\sum_{k=0}^{M-1} 2^{-k} + O\left(M^2 p^{1/2}q\right) \\
&= \frac{pq}{2}\left(1 - \left(\frac{1}{2}\right)^M\right) + O(p^{1/2}q(\log p)^2) \\
&= \frac{pq}{2} + O(p^{1/2}q(\log p)^2),
\end{aligned}
$$

where in the last step we used

$$p^{-1/2}(\log p)^2 \leq 2^{-M} \leq 2p^{-1/2}(\log p)^2$$

9

by the choice of M.

Similar, we can show that the number N_0 of n with $w_n = 0$ is at least

$$N_0 \geq \frac{pq}{2} + O(p^{1/2}q(\log p)^2).$$

Since

$$N_0 = pq - N_1 \leq \frac{pq}{2} + O(p^{1/2}q(\log p)^2)$$

we get the result by (1). $\qquad\square$

Put $N = pq$. In the important case that p and q are of the same order of magnitude the bound is of order of magnitude $N^{3/4}(\log N)^2$.

3.2 Arithmetic crosscorrelation of two binary m-sequences of coprime periods

First note that

$$d = \gcd(n_1, n_2) = 1 \quad \text{if and only if} \quad t = \gcd(2^{n_1} - 1, 2^{n_2} - 1) = 1.$$

This can be easily verified. On the one hand if $d > 1$, then

$$2^{n_i} - 1 = \left(2^d - 1\right)\left(1 + 2^d + \cdots + 2^{(n_i/d - 1)d}\right), \quad i = 1, 2,$$

and $2^d - 1$ is a nontrivial divisor of t. On the other hand if $t > 1$, then there is a prime divisor $p > 2$ of t and the order of 2 modulo p divides d.

Let g_n be a primitive element of the finite field $\mathbb{F}_{2^n}$. Then the sequence of the form[1]

$$s_i = \mathrm{Tr}_n(g_n^i), \quad i = 0, 1, \ldots$$

is an m-sequence of period $2^n - 1$, where

$$\mathrm{Tr}_n(c) = c + c^2 + \cdots + c^{2^{n-1}}, \quad c \in \mathbb{F}_{2^n},$$

denotes the (absolute) trace of $\mathbb{F}_{2^n}$.

We need a result on the pattern distribution of two m-sequences.

Lemma 2. *Let $n_1 < n_2$ be two coprime positive integers and denote by $\mathcal{S} = (s_i)_{i \geq 0}$ and $\mathcal{T} = (t_i)_{i \geq 0}$ two m-sequences of periods $2^{n_1} - 1$ and $2^{n_2} - 1$, respectively.*

For an integer $k \geq 0$ and any pattern $\underline{e} \in \{0, 1\}^{2k+2}$, the number Σ_k of $i = 0, 1, \ldots, (2^{n_1} - 1)(2^{n_2} - 1) - 1$ with

$$(s_{i-k}, s_{i-k+1}, \ldots, s_i, t_{i-k}, t_{i-k+1}, \ldots, t_i) = \underline{e}$$

---

[1]In fact, any m-sequence can be defined as $\bar{s}_i = \mathrm{Tr}_n(ag_n^i)$, $i = 0, 1, \ldots$, for some $0 \neq a \in \mathbb{F}_{2^n}$, which is a shift of $s_i = \mathrm{Tr}_n(g_n^i)$.

satisfies

$$\left| \Sigma_k - \frac{(2^{n_1} - 1)(2^{n_2} - 1)}{2^{2k+2}} \right| \leq 2^{n_1-k-1} + 2^{n_2-k-1} + 1 \quad for \quad k < n_1.$$

Proof. Let

$$s_i = \text{Tr}_{n_1}(g_{n_1}^i), \quad i = 0, 1, \ldots$$

for a primitive element g_{n_1} of $\mathbb{F}_{2^{n_1}}$ and

$$t_i = \text{Tr}_{n_2}(g_{n_2}^i), \quad i = 0, 1, \ldots$$

for a primitive element g_{n_2} of $\mathbb{F}_{2^{n_2}}$, respectively.
   Put

$$N = (2^{n_1} - 1)(2^{n_2} - 1)$$

and

$$\delta_{i,j,n_1} = 1 + (-1)^{e_{k-j}} \psi_{n_1}(g_{n_1}^{i-j}) \quad \text{and} \quad \delta_{i,j,n_2} = 1 + (-1)^{e_{2k+1-j}} \psi_{n_2}(g_{n_2}^{i-j}),$$

where

$$\psi_n(c) = (-1)^{\text{Tr}_n(c)}, \quad c \in \mathbb{F}_{2^n},$$

is the additive canonical character of $\mathbb{F}_{2^n}$. Note that

$$\delta_{i,j,n_1} \delta_{i,j,n_2} = \begin{cases} 4, & \text{if } (s_{i-j}, t_{i-j}) = (e_{k-j}, e_{2k+1-j}), \\ 0, & \text{otherwise,} \end{cases}$$

and thus we have

$$\Sigma_k = \frac{1}{2^{2k+2}} \sum_{i=0}^{N-1} \prod_{j=0}^{k} \delta_{i,j,n_1} \delta_{i,j,n_2}.$$

Expanding the product we get

$$\Sigma_k = \frac{1}{2^{2k+2}} \sum_{U,V \subseteq \{0,1,\ldots,k\}} \sum_{i=0}^{N-1} \prod_{j \in U} (-1)^{e_{k-j}} \psi_{n_1}(g_{n_1}^{i-j}) \prod_{j \in V} (-1)^{e_{2k+1-j}} \psi_{n_2}(g_{n_2}^{i-j}).$$

The contribution of $U = V = \emptyset$ is

$$\frac{N}{2^{2k+2}}.$$

The contribution of $U = \emptyset$ and the $(2^{k+1} - 1)$ sets $V \neq \emptyset$ is bounded by

$$2^{n_1-k-1} \max_{V \neq \emptyset} \left| \sum_{i=0}^{2^{n_2}-2} \psi_{n_2} \left(\sum_{j \in V} g_{n_2}^{-j} g_{n_2}^i \right) \right|.$$

Since $k < n_1 < n_2$ and the minimal polynomial of $g_{n_2}^{-1}$ is of degree n_2, we have

$$\sum_{j \in V} g_{n_2}^{-j} \neq 0$$

and the sum over i equals -1. So the contribution of $U = \emptyset$ and $V \neq \emptyset$ is at most

$$2^{n_1 - k - 1}.$$

Similarly we see that the contribution of $V = \emptyset$ and $U \neq \emptyset$ is

$$2^{n_2 - k - 1}.$$

In the remaining case the contribution of the $2^{2k+2} - 2^{k+2} + 1$ pairs of sets (U, V) with $U \neq \emptyset$ and $V \neq \emptyset$ is at most 1.

Putting everything together, we complete the proof. $\qquad\square$

Theorem 3. *Let $\mathcal{S} = (s_i)_{i \geq 0}$ and $\mathcal{T} = (t_i)_{i \geq 0}$ be two binary m-sequences over $\mathbb{F}_2$ of periods $2^{n_1} - 1$ and $2^{n_2} - 1$, respectively. If $n_1 < n_2$ and $\gcd(n_1, n_2) = 1$, then the arithmetic crosscorrelation $\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau)$ satisfies*

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau) \ll n_1 2^{n_2}, \quad 0 \leq \tau < (2^{n_1} - 1)(2^{n_2} - 1).$$

Proof. As in the proof of Theorem 2 and using Lemma 2 we get

$$
\begin{aligned}
N_i &\geq \frac{(2^{n_1} - 1)(2^{n_2} - 1)}{2} \sum_{k=1}^{n_1 - 1} 2^{-k} - (n_1 - 1)(2^{n_1} + 2^{n_2} + 1) \\
&= \frac{(2^{n_1} - 1)(2^{n_2} - 1)}{2} + O(n_1 2^{n_2})
\end{aligned}
$$

for $i = 0, 1$. Hence,

$$|N_0 - N_1| \ll n_1 2^{n_2}$$

and the result follows. $\qquad\square$

The theorem above indicates that the arithmetic crosscorrelation of two m-sequences is quite small. In particular, if n_1 and n_2 are close and $N = (2^{n_1} - 1)(2^{n_2} - 1)$, then the bound is of order of magnitude $N^{1/2} \log N$ which is in good correspondence with the expected value of the absolute value of the arithmetic crosscorrelation of two random sequences of period N, see [7, Theorem 8.3.6].

However, the arithmetic autocorrelation of m-sequences is quite large. Numerical data indicates that its maximum value is the greatest number less than half of the period [1]. To improve the results in [1], in the following subsection, we will estimate the arithmetic autocorrelation function of m-sequences for small lags τ.

3.3 Arithmetic autocorrelation function of m-sequences for small lags τ

By [10, Proposition 2.1] we have for any N-periodic sequence

$$\mathcal{A}_{\mathcal{S}}^{A}(\tau) = -\mathcal{A}_{\mathcal{S}}^{A}(N-\tau), \quad \tau = 1, 2, \ldots, N-1. \tag{4}$$

Now we state a result on the pattern distribution of an m-sequence of period $2^n - 1$.

Lemma 3. *Let $\mathcal{S} = (s_i)_{i \geq 0}$ be an m-sequence of period $2^n - 1$. Choose τ with $1 \leq \tau < n$.*
(1) For $k \geq 0$ and any non-zero pattern $\underline{e} \in \{0,1\}^{2k+2} \setminus \{(0,0,\ldots,0)\}$, the number Σ_k of $i = 0, 1, \ldots, 2^n - 2$ with

$$\left(s_{i-k}, s_{i-k+1}, \ldots, s_i, s_{i-k+\tau}, s_{i-k+\tau}, \ldots, s_{i+\tau} \right) = \underline{e}$$

is

$$\Sigma_k = 2^{n-2k-2}, \quad k \leq \min\{\tau, n-\tau\} - 1.$$

(2) For $k \geq \tau$ and non-zero pattern $\underline{e} \in \{0,1\}^{k+\tau+1} \setminus \{(0,0,\ldots,0)\}$, the number σ_k of $i = 0, 1, \ldots, 2^n - 2$ with

$$\left(s_{i-k}, s_{i-k+1}, \ldots, s_{i+\tau} \right) = \underline{e}$$

is

$$\sigma_k = 2^{n-k-\tau-1}, \quad \tau \leq k \leq n - \tau - 1.$$

Proof. For $\ell = 1, 2, \ldots, n$, we see that each non-zero pattern of length ℓ occurs as $(s_i, s_{i+1}, \ldots, s_{i+\ell-1})$ for exactly $2^{n-\ell}$ different i with $0 \leq i < 2^n - 1$, see for example [3, Proposition 5.2]. We choose $\ell = k + \tau + 1 \leq n$ and (2) follows.

For (1) note that the choice of $(s_{i+1}, s_{i+2}, \ldots, s_{i-k+\tau-1})$ is free. Hence, we derive $2^{n-\ell+\tau-k-1} = 2^{n-2k-2}$ different i with the desired pattern property. $\square$

Theorem 4. *Let $\mathcal{S} = (s_i)_{i \geq 0}$ be an m-sequence of period $2^n - 1$. We have*

$$\left| \mathcal{A}_{\mathcal{S}}^{A}(\tau) \right| \leq 2^{\min\{n-1, \tau+1, 2^n - \tau\}} - 1, \quad 1 \leq \tau < 2^n - 1.$$

Proof. The bound

$$\left| \mathcal{A}_{\mathcal{S}}^{A}(\tau) \right| \leq 2^{n-1} - 1$$

follows from [1]. By (4) it remains to show

$$\left| \mathcal{A}_{\mathcal{S}}^{A}(\tau) \right| \leq 2^{\tau+1} - 1 \quad \text{for} \quad \tau \leq n - 3.$$

As before, let N_j be the number of digits equal to $j \in \{0,1\}$ in the binary expansion of the integer $S(2) - S^{(\tau)}(2)$. As in the above proofs, we get by Lemma 3

$$N_1 \geq \sum_{k=1}^{\min\{\tau, n-\tau\}-1} 2^{k+1} \Sigma_k + \sum_{k=\min\{\tau, n-\tau\}}^{n-\tau-1} 2^{\tau} \sigma_k = \sum_{k=1}^{n-\tau-1} 2^{n-k-1} = 2^{n-1} - 2^{\tau}$$

13

as well as $N_0 \geq 2^{n-1} - 2^\tau$ and thus

$$N_j \leq 2^{n-1} + 2^\tau - 1, \quad j = 0, 1.$$

Hence, $|N_0 - N_1| \leq 2^{\tau+1} - 1$ and (1) finishes the proof. □

4 Final remarks

Below we list some numerical data for the arithmetic crosscorrelations of two Legendre sequences of coprime periods in Table 1 and two binary m-sequences of coprime periods in Table 2.

p	q	$\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau)$
7	11	-1
7	13	5
7	17	-13
7	23	1
11	13	-3
11	19	-5
13	17	-7
17	29	9

Table 1: Arithmetic crosscorrelation of two Legendre sequences $\mathcal{S}$ and $\mathcal{T}$ of periods p and q

Let us denote by $M_{\mathcal{S}}(X) \in \mathbb{F}_2[X]$ the minimal polynomial of $\mathcal{S}$, see [2] for details.

$M_{\mathcal{S}}(X)$	$M_{\mathcal{T}}(X)$	$\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau)$
$X^3 + X^2 + 1$	$X^4 + X^3 + 1$	-1
$X^3 + X^2 + 1$	$X^5 + X^3 + 1$	1
$X^3 + X^2 + 1$	$X^7 + X^6 + 1$	-3
$X^3 + X^2 + 1$	$X^8 + X^6 + X^5 + X^4 + 1$	7
$X^5 + X^3 + 1$	$X^8 + X^6 + X^5 + X^4 + 1$	-1
$X^6 + X^5 + 1$	$X^7 + X^6 + 1$	-1
$X^7 + X^6 + 1$	$X^8 + X^6 + X^5 + X^4 + 1$	-3

Table 2: Arithmetic crosscorrelation of two binary m-sequences $\mathcal{S}$ and $\mathcal{T}$ with minimal polynomials $M_{\mathcal{S}}(X)$ and $M_{\mathcal{T}}(X)$

These tables indicate that the size of $\mathcal{C}_{\mathcal{S},\mathcal{T}}^A(\tau)$ can be quite different for different periods of similar size.

For two binary sequences $\mathcal{S}$ of period p and $\mathcal{T}$ of period q with $\gcd(p,q) > 1$, their classical and arithmetic crosscorrelations are both not constant. For example, if $\mathcal{S}$ is an m-sequence with minimal polynomial $X^4 + X^3 + 1$ and $\mathcal{T}$ is an m-sequence with minimal polynomial $X^4 + X + 1$, we compute that

$$\mathcal{C}_{\mathcal{S},\mathcal{T}}(\tau) \in \{-1, -5, 3, 7\}, \quad \mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau) \in \{-3, -7, -9, 1, 3, 5\}.$$

So it would be interesting to find an upper bound on $\mathcal{C}^A_{\mathcal{S},\mathcal{T}}(\tau)$ when the period of $\mathcal{S}$ is not coprime to that of $\mathcal{T}$. For m-sequences the method of this paper still provides non-trivial results for very small and very large lags τ but fails for most τ.

Acknowledgments

References

[1] Z. Chen, Z. Niu, Y. Sang and C. Wu, "Arithmetic autocorrelation of binary m-sequences," Nov. 2021. [Online]. Available: https://arxiv.org/abs/2111.11176.

[2] T. W. Cusick, C. Ding and A. Renvall, *Stream Ciphers and Number Theory.* Elsevier, Amsterdam, 2004.

[3] S. W. Golomb and G. Gong, *Signal Design for Good Correlation: For Wireless Communication, Cryptography and Radar.* Cambridge University Press, Cambridge, 2005.

[4] M. Goresky and A. Klapper, "Arithmetic crosscorrelations of feedback with carry shift register sequences," *IEEE Trans. Inf. Theory*, vol. 43, no.4, pp.1342-1345, 1997.

[5] M. Goresky and A. Klapper, "Some results on the arithmetic correlation of sequences," in Sequences and Their Applications 2018 (SETA 2008), Lecture Notes in Comput. Sci., vol. 5203, pp.71-80.

[6] M. Goresky and A. Klapper, "Statistical properties of the arithmetic correlation of sequences," Int. J. Found. Comput. Sci., vol. 22, no.6, pp.1297-1315, 2011.

[7] M. Goresky and A. Klapper, *Algebraic Shift Register Sequences*. Cambridge University Press, Cambridge, 2012.

[8] T. Helleseth, "Maximal-length sequences," in Encyclopedia of Cryptography and Security (2nd Ed.), pp.763-766, Springer, Boston, MA, 2011.

[9] R. Hofer, L. Merai and A. Winterhof, "Measures of pseudorandomness: Arithmetic autocorrelation and correlation measure," In: C. Elsholtz, P. Grabner (Hrsg.), Number Theory - Diophantine Problems, Uniform Distribution and Applications, pp.303-312, 2017.

[10] R. Hofer and A. Winterhof, "On the arithmetic autocorrelation of the Legendre sequence," Adv. Math. Commun., vol. 11, no.1, pp.237-244, 2017.

[11] R. Lidl and H. Niederreiter, *Finite Fields*. With a foreword by P. M. Cohn. Second edition. Encyclopedia of Mathematics and its Applications, vol. 20. Cambridge University Press, Cambridge, 1997.

[12] D. M. Mandelbaum, "Arithmetic codes with large distance," IEEE Trans. Inf. Theory, vol. 13, no.2, pp.237-242, 1967.

[13] T. Tian, W.-F. Qi, "Autocorrelation and distinctness of decimations of ℓ-sequences," SIAM J. Discret. Math., vol. 23, no.2, pp.805-821, 2009.