Interleaved Sequences Over Finite Fields

by

Jing He

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematics

Carleton University
Ottawa, Ontario

©2013
Jing He
Acknowledgements

First and foremost, I would like to take this opportunity to express my utmost gratitude to Professor Daniel Panario and Qiang Wang, my academic supervisors, who offered their continuous advice and encouragement throughout the course of this thesis. I thank them for the systematic guidance and great effort they put into training me in the scientific field. I have received a tremendous amount of mathematical knowledge and insight, support and patience. They have always been generous with their time. They also gave me many opportunities to take various conferences which broaden my outlook. They patiently provided the vision, encouragement and advise necessary for me to proceed through the doctoral program and complete my dissertation.

Special thanks to my committee, Dr. Guang Gong, Dr. Halim Yanikomeroglu, Dr. Saban Alaca, and Dr. Mike Newman for their support, guidance and helpful suggestions. Their guidance has served me well and I owe them my heartfelt appreciation.

I would like to thank Dr. Arne Winterhof. He gave me many great ideas. During collaborating with him in a project, I have learned a great deal from him as a mathematician. I would also like to thank Dr. Guang Gong. She gave me many good suggestions on my research both in theoretical study and in engineering applications.

Throughout the past years, my master supervisor Dr. Adam Van Tuyl has cared for my PhD study. He sometimes gives me advice on how to be a mathematician. In March 2012, he invited me to give a talk in the math colloquium at Lakehead University.

Members of staff team in our department also deserve my sincerest thanks, their friendship and assistance has meant more to me than I could ever express.
My friends, David Thomson and Georgios Tzanakis, I am very happy that, in many cases, my friendships with you have extended well beyond our shared study time at Carleton University.

I wish to thank my parents, and my aunt. Their love provided my inspiration and was my driving force. I owe them everything and wish I could show them just how much I love and appreciate them. My husband and my son whose love and encouragement allowed me to finish this journey.

Finally, I would like to dedicate this work to my lost grandma, who left us too soon. I hope that this work makes you proud.
Abstract

We study cryptographic properties such as correlation, correlation measure and linear complexity profile of the sequences defined by characters with interleaved structure.

Families of pseudorandom sequences with low cross correlation have important applications in communications and cryptography. Among several known constructions of sequences with low cross correlation, interleaved constructions proposed by Gong use two sequences of the same period with two-level autocorrelation. We generalize the interleaved algorithm and construct some families of interleaved sequences. We study the balance property and the cross correlation of interleaved sequences such that the base sequences may not have the same period, or they may not have two-level autocorrelation. In particular, we study the interleaved sequences of two Legendre sequences of periods $p$ and $q$, respectively, where $p$ and $q$ are odd prime numbers.

Then we interleave two $m$-ary sequences of integer periods $s$ and $t$, where the first sequence is an $m$-ary power residue sequence of prime period $s = p$ and the second one is any $m$-ary sequence of period $t$. We provide nontrivial lower bounds on the linear complexity profile for any periods $p$ and $t$. We study the power correlation measure of order $k$ (with bound $B$) of the interleaved $m$-ary sequences. As an example, we give more details on the correlation measure of order $k$ (with bound $B$) of interleaved Legendre sequences, a special case of interleaved $m$-ary sequences. We provide a nontrivial upper bound on the correlation measures of order $k$ for $k$ odd or $k$ even if $p \geq 2t$ or $p = t$. The nontrivial bound is achieved for all $k$, $p$ and $t$, when the correlation measure of order $k$ with lags is bounded by $B$. In the remaining cases we show by example that the correlation measures can approach the trivial bound. A relation between the linear complexity profile and the power correlation measure of order $k$ with bound $B$ is also provided.
For applications in communications sequences with a small alphabet size are often required. We define a new class of sequences over the set of complex numbers with unit magnitude (and 0) which have small alphabet size, long period and almost perfect autocorrelation. Sequences with low correlation property or zero correlation property around the origin can be used in such systems for reducing multiple-access interference. Such sequence sets are called low correlation zone (LCZ) or zero correlation zone (ZCZ) sequence sets (families), respectively. A new family of interleaved polyphase sequences is constructed using the newly defined sequences as the base sequences and the shift sequence being a Costas sequence. We find the ZCZ of this constructed family of interleaved polyphase sequences.

A family of sequences constructed by shift-adding one two-prime generator of order 2 sequence of period $pq$, is also studied. We show that the normalized maximal correlation magnitude of the family is $\frac{9}{\sqrt{pq}} + 4 \left( \frac{1}{p} + \frac{1}{q} \right)$, which is low when compared to the trivial bound.
Contents

Acknowledgements i

Abstract iii

Chapter 1. Introduction 1
  1.1. Motivation 1
  1.2. Overview of new results 5

Chapter 2. Preliminaries 9
  2.1. Sequences and cryptographic properties 9
  2.2. Maximal length sequences 18
  2.3. Construction of GMW sequences by interleaved approach 25
  2.4. Character sums over finite fields 25
  2.5. Costas sequences 31
  2.6. Some well-known signal sets (families) 32

Chapter 3. Interleaved sequences 35
  3.1. \((f(x), m)\)-interleaved sequences 35
  3.2. IRI-sequences 40
  3.3. Signal sets (families) generated by \((f(x), m)\)-sequences 43
  3.4. A new variant of the interleaved algorithm 45

Chapter 4. Correlation properties and linear complexity of interleaved binary sequences 49
  4.1. Correlation property 49
  4.2. The correlation values of binary interleaved sequences 51
  4.3. Linear complexity of binary interleaved sequences 63

Chapter 5. Linear complexity profile, correlation measures and nonlinearity 66
<table>
<thead>
<tr>
<th>Chapter 5.</th>
<th>Linear complexity profile of ( m )-ary interleaved sequences</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.</td>
<td>Relating linear complexity profile and (power) correlation measures</td>
<td>74</td>
</tr>
<tr>
<td>5.3.</td>
<td>Nonlinearity of the boolean function derived from binary interleaved sequences</td>
<td>89</td>
</tr>
<tr>
<td>Chapter 6.</td>
<td>Zero correlation zone of interleaved polyphase sequences</td>
<td>94</td>
</tr>
<tr>
<td>6.1.</td>
<td>A new class of almost perfect sequences</td>
<td>95</td>
</tr>
<tr>
<td>6.2.</td>
<td>Interleaved construction of polyphase sequences</td>
<td>98</td>
</tr>
<tr>
<td>6.3.</td>
<td>A Zero-Correlation Zone (ZCZ) signal set (family)</td>
<td>99</td>
</tr>
<tr>
<td>Chapter 7.</td>
<td>Shift-addition construction</td>
<td>105</td>
</tr>
<tr>
<td>7.1.</td>
<td>Shift and add property and shift-addition construction</td>
<td>106</td>
</tr>
<tr>
<td>7.2.</td>
<td>Shift-addition family from ( m )-ary power residue sequences</td>
<td>106</td>
</tr>
<tr>
<td>7.3.</td>
<td>Shift-addition family from two-prime generators of order 2</td>
<td>107</td>
</tr>
<tr>
<td>Chapter 8.</td>
<td>Future work</td>
<td>113</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>116</td>
</tr>
</tbody>
</table>
CHAPTER 1

Introduction

1.1. Motivation

Pseudorandom sequences are widely used in computer science and engineering including applications to spread spectrum communication systems, radar systems, signal synchronization, simulation, and cryptography [27]. A set of sequences of the same period is called a family (also called a signal set or a sequence set). The pseudorandom sequences in a good family should be easy to generate (possibly with hardware or software), have good distribution properties which make them appear statistically to be random, have low cross correlation values so that each sequence may be separated from the others in the family, and arise from some underlying algebraic structure so they can be analyzed using standard mathematical tools. In this thesis we seek families of binary, or more generally, $m$-ary sequence (sequences of $m$ symbols) with low correlation, good randomness, and large linear complexity (profile). Such kind of families of sequences have important applications in code-division multiple-access (CDMA) communications and cryptography [27].

Correlation is a measure of the similarity, or relatedness, between two phenomena. In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time-lag applied to one of them. When the two waveforms are the same one with different time-lags applied, the function is called the autocorrelation. The trivial value of the correlation function is equal to the period of the sequence. When the maximum magnitude of the nontrivial autocorrelation and cross correlation values of the sequence is smaller than the square root of the period, we say the sequence has low correlation. Such sequences with low correlation are employed in CDMA communications and can successfully combat interference from other users sharing a common channel.

Let us consider a family of $k$ sequences with same period $v$, let $C_{\text{max}}$ be the maximum magnitude of the nontrivial autocorrelation and cross correlation values
of the sequences in the family. It is known that $C_{\text{max}} \geq \sqrt{\frac{k-1}{v(k-1)}}$ which is called the Welch bound (see [79] or [90]). Families of sequences matching the Welch bound are used in communication systems.

In telecommunications, sequences with low correlation are used as signals in three spread spectrum communication systems: synchronous code division multiple access (CDMA), asynchronous code division multiple access, and quasi-synchronous code division multiple access. In synchronous CDMA Walsh Hadamard sequences are usually employed which have zero cross correlation (also called orthogonal cross correlation). However, the orthogonal property is lost at a non-trivial time shift. Gaudenzi, Elia and Viola [22] proposed the quasi-synchronous code-division multiple-access communication systems. The major feature of quasi-synchronous CDMA is that the duration of any signature sequence is significantly longer than the maximal signal delay with respect to one period. In contrast to completely asynchronous CDMA systems for which all aperiodic correlations must be taken into consideration, in the case of quasi-synchronous CDMA, it is sufficient to focus on a small window (the interference window), around the zero shift. This simplifies the sequence design and allows to drastically reduce both the internal-sequence interference and multiple access interference. This enhances the system performance. Moreover, this is important since constructions of sequences for asynchronous CDMA systems are difficult to find.

LCZ (Low Correlation Zone) or ZCZ (Zero Correlation Zone) are topics related to correlation. Long, Zhang, and Hu [58] propose families that have low-correlation values around the origin, which can be used as spreading sequences in the QS-CDMA systems. Such sequences are called Low-Correlation Zone (LCZ) sequences. If we have zero-correlation values around the origin we have Zero-Correlation Zone (ZCZ) sequences. For a prime $p$, Tang and Fan [82] propose $p$-ary LCZ families by extending the alphabet size of each sequence in [58]. Later, Kim, Jang, No and Chung [47] present a new construction method for quaternary LCZ sequence sets by using a binary sequence of the same period with ideal autocorrelation. These authors also calculate the correlation values of their sequence sets constructed from a maximal length sequence or a GMW sequence; see [47].
For zero correlation zone (ZCZ) sequence sets, Torii, Nakamura, and Suehiro [84] proposed two methods for construction ZCZ sequences based on perfect sequences and unitary matrices. Wang and Qi [87] introduce the concept of $D$-matrix and use an interleaved structure together with the $D$-matrix to get the ZCZ length for the interleaved ZCZ sequence sets. Tang and Udaya [83] provide a new binary LCZ sequence set derived from the interleaved technique and Hadamard matrices. Zhou, Tang and Gong [95] give a new construction of LCZ/ZCZ sequence sets with period $2N$ using two shift sequences of period $N$. Finally, Hu and Gong [39] generalize this LCZ/ZCZ sequence set of period $2N$ to any multiple of $N$ using the interleaving technique and complex Hadamard matrices.

In addition to studying correlation properties we also consider some other properties of sequences, such as linear complexity (profile). There are two concepts, linear complexity and linear complexity profile, which are important randomness measures of a sequence. The linear complexity of a sequence is the length of the shortest linear recurrence relation which equals the degree of the polynomial generating the sequence. A sequence of low linear complexity can be fully determined via the Berlekamp-Massey algorithm [61]. Hence, to resist algebraic attacks a high linear complexity is a prerequisite for cryptographic applications. Rueppel [72] introduced the linear complexity profile that describes how the linear complexity grows in terms of the length of a finite sequence. A high lower bound on the linear complexity profile is also a desirable feature of a sequence for cryptography. Linear complexity is a global concept while linear complexity profile is a local concept. Both are important cryptographic characteristics of sequences and provide information on the predictability and thus unsuitability for cryptography of the sequence.

There always exists recurrence relations for a periodic sequence. The feedback polynomial is derived from the shortest recurrence relation of a sequence. The standard approach to find the linear complexity of a sequence is to determine the degree of the feedback polynomial. To estimate the bounds of linear complexity profile we employ correlation measures of order $k$ of a sequence [62] and apply Weil’s inequality using a relation between the linear complexity profile and the correlation measure of order $k$ of a sequence given in [6].
The above are the cryptographic properties we study for the constructed sequences in this thesis. Our goal is to construct large-size families of sequences with long period and nice cryptographic properties.

A binary maximal length sequence meets the three randomness postulates proposed by Golomb in 1967 [26]: (1) balance property, (2) run property, and (3) two-level autocorrelation property. Given the low autocorrelation sidelobes of a maximal length sequence, it is natural to attempt to construct families of low correlation sequences starting from maximal length sequences. Some well-known constructions of this type are the families of Gold and Kasami sequences. Gong first proposed the interleaved construction in [29] and then gave more details and introduced families of interleaved sequences in [31]. The interleaved sequences include many previously better known sequences such as multiplexed sequences, clock-controlled sequences, Kasami sequences, GMW sequences, geometric sequences, and No sequences.

Compared to maximal length sequences, Legendre sequences also possess the two-level autocorrelation property [96]. We consider a generalization of the interleaved construction in [31] and construct new binary families by Legendre sequences. We first give new results about the upper bounds on the maximal magnitude of correlation for the new binary families. Then we extend Legendre sequences to power residue sequences and employ them in the generalized interleaved construction to get new families of interleaved $m$-ary sequences. We give estimates for the bounds on the linear complexity profile.

A character is a homomorphism which preserves the operation on elements from a group $G$ to the multiplicative group $U$ of complex numbers. A polyphase sequence is a sequence of complex numbers with unit magnitude. Power residue sequences [78] are polyphase sequence with low correlation and variable alphabet sizes, represented by multiplicative characters. The binary case of power residue sequences is the Legendre (or quadratic residue) sequences which are well studied and have several important properties including a low (periodic) autocorrelation [96].

When studying families constructed by power residue sequences, we aim at obtaining a new class of sequences, similar to power residue sequences, defined by characters over finite fields. We obtain a new class of almost perfect sequences, which means the autocorrelation of the sequence is zero almost everywhere but for
one exceptional shift in one period, by means of multiplicative characters. Then we observe that a zero correlation zone happens among all sequences in the new family constructed by these almost perfect sequences.

When studying the linear complexity profile, we also provide results about two cryptographic measures, called sparsity and nonlinearity of the boolean functions derived from every sequence in the family of binary interleaved Legendre sequences. Boolean functions are important in the theory of error-correcting codes, as well as in cryptography. An affine (or linear) Boolean function does not provide an effective confusion. To overcome this, we need functions which are as far as possible from being an affine function. The effectiveness of these functions is measured by a parameter called nonlinearity. The sparsity (also called algebraic thickness) of a boolean function is the minimal number of its nonzero coefficients in any polynomial representation. We connect sequences with boolean functions and give the bounds to the sparsity and the nonlinearity of the boolean functions derived from binary interleaved Legendre sequences based on the same approach used to provide the linear complexity profile of these sequences.

The new results in this thesis are mostly contained in Section 3.4 and Chapters 4, 5 and 6. Next, we give an overview of new results.

1.2. Overview of new results

The goal of this thesis is to construct families of sequences with low correlation and high linear complexity profile. Mainly we employ the interleaved technique. In Chapter 2, we give background knowledge about sequences and the cryptographic properties we consider. The required mathematics is also given. At the end this chapter we study maximal length sequences which are among the most important and widely used sequences. Studying them highlights the ideas from which the interleaved structure came from. In Chapter 3 we introduce the interleaved sequences based on the work by Gong [29].

We give a new algorithm in Section 3.4 which extends Gong’s interleaving algorithm in [29] to a general case. By this we get a new binary family of interleaved sequences using two binary sequences of integer periods $v$ and $w$. A formula to determine the number of zeros in one period of an interleaved sequence is given in Section 3.4.
In Chapter 4 we obtain the following results about cross correlation of interleaved sequences.

- The cross correlation bounds for the interleaved sequences from one base sequence with two-level autocorrelation and the other base sequence being balanced.
- The cross correlation bounds for the interleaved sequences arising from two two-level autocorrelated base sequences.
- The cross correlation bounds for the interleaved sequences from two base sequences with same length.
- The cross correlation bounds for the interleaved sequences from two Legendre sequences with any two odd prime periods $p$ and $q$. The result is in two cases: (1) $p \equiv 3 \pmod{4}$; (2) $p \equiv 1 \pmod{4}$.

The linear complexity of a sequence is the degree of the feedback polynomial of the sequence. For a sequence of period $T$, the linear complexity can be computed as

$$T - \deg \left( \gcd(x^T - 1, G(x)) \right),$$

where $G(x)$ is the generating function of the sequence [14]. We continue by studying the interleaved sequence constructed by the algorithm in Section 3.4. In particular, we give the generating functions of the interleaved sequences, and then we obtain bounds on the linear complexities of the interleaved sequences in Section 4.3.

Instead of considering only binary sequences, we extend the base sequences to be $m$-ary sequences in Chapter 5, and thus the interleaved sequences are of $m$ symbols. In Sections 5.1 and 5.2 we estimate the linear complexity profile of interleaved sequences, the first time that the linear complexity profile has been shown for interleaved sequences. We give two approaches. On one hand we use the shortest recurrence relation to find a lower bound of the linear complexity profile. On the other hand we estimate the correlation measure of order $k$ of the sequences and then obtain a lower bound of the linear complexity profile. Our second approach gives a tighter bound which we use to estimate the linear complexity profile using correlation measures in Section 5.2.

The results about correlation measures and linear complexity profile of the $m$-ary interleaved sequences are the following:
• Lower bounds for the linear complexity profile of binary interleaved sequences and \( m \)-ary interleaved sequences.

• A relation between linear complexity profile and correlation measures with bounded lags.

• Upper bounds for the correlation measures with bounded lags or unbounded lags.

• Improved upper bounds for the correlation measures in the binary case \((m = 2)\), that is, when the base sequences are Legendre sequences.

• Lower bound for the linear complexity profile of \( m \)-ary interleaved sequences using the relation between linear complexity profile and correlation measures with bounded lags.

For binary interleaved sequences we also derive a boolean function and show the bounds on the nonlinearity and the sparsity of this boolean function by the Weil’s inequality in Section 5.3.

In Chapter 6, using maximal length sequences and multiplicative characters, we construct a class of sequences with almost perfect autocorrelation. We interleave the defined sequences to get a ZCZ-signal set (family). In Section 6.2 we give an interleaved algorithm for complex-valued sequences which is different from the interleaved algorithm for \( m \)-ary sequences in Section 3.4. Shift sequences in which the differences of any two consecutive elements are in an arithmetic progression are employed in Section 3.4 and in Chapter 4 and 5. In Chapter 6 we employ several kinds of Costas arrays as the shift sequence. The basic background knowledge on Costas arrays is given in Section 2.5. Then we interleave two sequences in this class to construct a zero correlation zone (ZCZ) signal set (family) in Section 6.3 with large size.

The shift-addition construction, which can be viewed as a special case of interleaved construction, is applied on two-prime generator of order two (also called the generalized cyclotomic sequence of order two) to get a new family of sequences. In Chapter 7 we construct a new family using two-prime generator of order two under the shift-and-addition construction. The maximal correlation magnitude of this family is low and at most the order \( 9\sqrt{pq} + 4(1/p + 1/q) \), where \( pq \) is the period of every sequence in the family.
We finish with a series of open problems in Chapter 8. These unanswered questions provide ample motivation for future work.
CHAPTER 2

Preliminaries

The theory of finite fields, whose origins can be traced back to the works of Gauss and Galois, has played a part in various branches of mathematics. In recent years there has been a resurgence of interest in finite fields, and this is partly due to important applications in coding theory and cryptography. One area of applications of finite fields is sequences. In Section 2.1 we review basic knowledge about sequences and we provide cryptographic concepts related to sequences. These are important measures of randomness. Sequences with good randomness properties can be used in communication systems. In Section 2.2 we introduce maximal length sequences which are widely used in communication systems because of their good randomness properties. Maximal length sequences naturally possess interleaved structure which is also a motivation for the interleaved construction being proposed. In Section 2.4 we introduce character sums which are an important tool for us to estimate bounds on some cryptographic properties of the sequences generated by characters over finite fields. We also introduce GMW and Costas sequences.

Throughout this chapter we let $\ell$ be a power of a prime number $p$ and let $\mathbb{F}_{\ell}$ denote a finite field of characteristic $p$ with $\ell$ elements.

2.1. Sequences and cryptographic properties

Pseudo-random sequences with high linear spans and low correlation function values are widely used in communication and cryptology. Therefore, we should pick the sequences carefully according to their cryptographic properties. In this section we give the definitions for several cryptographic concepts.

2.1.1. Sequences. We study $m$-ary sequences in the thesis. Unless specified, we suppose the definitions in this section are over $\mathbb{Z}_m$, the residue class ring modulo $m$.

We denote a sequence $\mathbf{a} = (a_0, a_1, \ldots)$ or $(a_i)_{i \geq 0}$ or $\mathbf{a} = a(t)$ for $t \geq 0$. 

9
DEFINITION 2.1. Given a sequence \( a = (a_0, a_1, \ldots) \), if \( a_i \in \mathbb{Z}_m \), then we say that \( a \) is an \( m \)-ary sequence or a sequence over \( \mathbb{Z}_m \). If there exist integers \( T > 0 \) and \( u \geq 0 \) such that

\[ a_{i+T} = a_i, \text{ for all } i \geq u, \]

then the sequence is said to be *ultimately periodic* and \( T \) is called a period of the sequence. The smallest number \( T \) satisfying (1) is called a *(least) period* of the sequence. If \( u = 0 \), then the sequence is said to be *periodic*. When the context is clear, we simply say the period of \( a \) instead of the least period of \( a \), denoted \( \text{per}(a) \).

Let \( v \) be a positive integer and let \( a = (a_0, \ldots, a_{v-1}) \) be an \( m \)-ary sequence of period \( v \). For any integer \( i \geq 0 \), let the left shift operator act on \( a \) by \( L^i(a) = (a_i, a_{i+1}, \ldots, a_{i+v-1}) \). In particular, define \( L^\infty(a) = (0, \ldots, 0) \).

DEFINITION 2.2. Two sequences \( a = (a_0, \ldots, a_{v-1}) \) and \( b = (b_0, \ldots, b_{v-1}) \) of the same period \( v \) are called *(cyclically) shift equivalent* if there exists an integer \( k \) such that \( a_i = b_{i+k} \), for all \( i \geq 0 \). In this case, we write \( a = L^k(b) \), or simply \( a \sim b \). Otherwise, they are called *cyclically shift distinct*.

DEFINITION 2.3. For a sequence, \( k \) consecutive same elements preceded by and followed by different elements is called a *run* of length \( k \).

DEFINITION 2.4. We say that \( a \) is *balanced* if the difference of the numbers of any two elements appearing in one period of the sequence is not to exceed 1.

2.1.2. Correlation properties. Correlation is one of the most important measure for pseudorandom sequences. The pseudorandom sequences with low cross correlation employed in CDMA communications can successfully combat interference from the other users who share a common channel. On the other hand, the sequences with low cross correlation employed in either stream cipher crypto-systems as key stream generators or in digital signature algorithms as pseudorandom number generators can resist cross-correlation attacks.

DEFINITION 2.5. The *cross correlation function* \( C_{a,b}(\tau) \) of two \( m \)-ary sequences \( a \) and \( b \) of period \( v \) is defined in [26] as

\[ C_{a,b}(\tau) = \sum_{i=0}^{v-1} \omega^{a_i - b_{(i+\tau) \mod v}}, \quad \tau = 0, 1, \ldots, \]
where \( \omega \) is a primitive \( m \)th root of unity. If \( \overline{b} = \overline{a} \), then denote \( C_{\overline{a}}(\tau) = C_{\overline{a},\overline{b}}(\tau) \) as the autocorrelation of \( \overline{a} \).

The next definition shows the criteria for low correlation.

**Definition 2.6.** [37] Let \( s_j = (s_{j,0}, \ldots, s_{j,T-1}) \), \( 0 \leq j < r \), be \( r \) shift-distinct \( m \)-ary sequences of period \( T \). Let \( S = \{s_0, \ldots, s_{r-1}\} \), and let \( \delta = \max |C_{s_i,s_j}(\tau)| \) for any \( 0 \leq \tau < T \), \( 0 \leq i,j < r \), where \( \tau \neq 0 \) if \( i = j \). The set \( S \) is said to be a \( (T,r,\delta) \) signal set and \( \delta \) is referred to as the maximal correlation of \( S \). The set \( S \) has low correlation if \( \delta \leq c\sqrt{T} \) where \( c \) is a constant.

The square root of \( T \) above is based on Welch bound which is discussed in Chapter 3.

Besides considering correlation globally, we can also consider it through a window which is also called a zone.

**Definition 2.7.** [58] Let \( s_i = (s_{i,0}, s_{i,1}, \ldots, s_{i,T-1}) \), \( 0 \leq i < r \), be \( r \) shift-distinct \( m \)-ary sequences with period \( T \). The set \( S \) is called a low correlation zone signal set denoted by \( \text{LCZ} (T,r,L,\delta) \)-signal set if

\[
(2) \quad |C_{s_i,s_j}(\tau)| \leq \delta \text{ for any } |\tau| < L, \ 0 \leq i \neq j < r,
\]

and

\[
(3) \quad |C_{s_i}(\tau)| \leq \delta \text{ for any } 0 < |\tau| < L, \ 0 \leq i < r,
\]

where \( L \) is the length of LCZ signal set.

If \( \delta = 0 \), then \( S \) is a zero correlation zone signal set denoted by \( \text{ZCZ} (T,r,L) \)-signal set.

A signal set that satisfies (3) only is a low autocorrelation zone signal set \( S \); it is denoted by \( \text{LACZ} (T,r,L,\delta) \)-signal set.

Niederreiter pointed out [65] a deficiency of most used linear congruential method: “Other known regularities in sequences of linear congruential pseudorandom numbers are certain long range correlations. These regularities can be disruptive in simulations where random irregularities are desired.” With this in mind, Mauduit and Sarkozy introduced new measures of pseudorandomness.
Let $m \in \mathbb{N}$, $m \geq 2$, and $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ a finite set ("alphabet") of $m$ symbols ("letters"); consider a sequence with its elements taken from $\mathcal{A}$ of period $T$ (view as a $T$-tuple) $s = (s_1, s_2, \ldots, s_T) \in \mathcal{A}^T$.

**Definition 2.8.** [63] Let $\mathcal{E}_m = \{\varepsilon_1, \ldots, \varepsilon_m\}$ denote the set of $m$th roots of unity, and $\mathcal{F}$ denote the set of bijections $\varphi: \mathcal{A} \leftrightarrow \mathcal{E}_m$ (so that $|\mathcal{F}| = m!$). For $\phi = (\varphi_1, \ldots, \varphi_k)$ with $\varphi_i \in \mathcal{F}$ for $i = 1, \ldots, k$ and $D = (\tau_1, \ldots, \tau_k)$ with non-negative integers $\tau_1 < \cdots < \tau_k$, we write

$$G(s, \phi, M, D) = \sum_{n=1}^{M} \varphi_1(s_{n+\tau_1}) \varphi_2(s_{n+\tau_2}) \cdots \varphi_k(s_{n+\tau_k}).$$

Then the $\mathcal{E}_m$-correlation measure of order $k$ of $s$ is defined as

$$\Gamma_k(s) = \max_{\phi, M, D} |G(s, \phi, M, D)|$$

where the maximum is taken over all $\phi \in \mathcal{F}^k$, and $D = (\tau_1, \ldots, \tau_k)$ and $M$ such that $M + \tau_k \leq N$.

This definition, which is introduced in [63], coincides with the definition [62] for binary sequences.

**Definition 2.9.** [62] The correlation measure of order $k \geq 1$ of a binary sequence $s = (s_n)_{0 \leq n < T}$ of period $T$ is defined as

$$C_k(s) = \max_{M, D} \left| \sum_{n=0}^{M-1} (-1)^{s_n+\tau_1+s_{n+\tau_2}+\cdots+s_{n+\tau_k}} \right|,$$

where the maximum is taken over all $D = (\tau_1, \ldots, \tau_k)$ with non-negative integers $\tau_1 < \tau_2 < \cdots < \tau_k$ and $M$ is such that $M - 1 + \tau_k \leq T - 1$.

When $k = 2$, this is related to the classical autocorrelation.

Motivated by [6], a relation between the linear complexity profile and the correlation measures for $m$-ary sequences is given in [12]. If $m$ is a prime number, the result is a straightforward extension of the binary case in [6]. For a composite number, the definition of power correlation measure of order $k$ is introduced by Chen and Winterhof in [12].

**Definition 2.10.** [12] Let $m \geq 2$ and the complex number $\omega = \exp\left(\frac{2\pi\sqrt{-1}}{m}\right)$, and thus $\mathcal{E}_m = \{\omega^j : 0 \leq j \leq m - 1\}$. For $c \in \{1, 2, \ldots, m - 1\}$ we define
\( \mu_c : \mathbb{Z}_m \to \mathcal{E}_m \) by

\[
\mu_c(z) = \omega^{zc}, \quad z \in \mathbb{Z}_m.
\]

We note that \( \mu_c \in \mathcal{F} \) if and only if \( \gcd(c, m) = 1 \), where \( \mathcal{F} \) denotes the set of \( m! \) bijections \( \varphi : \mathbb{Z}_m \to \mathcal{E}_m \). Let

\[ \mathcal{F}' = \{ \mu_c : c = 1, 2, \ldots, m - 1 \}. \]

The \textit{power correlation measure of order} \( k \) \textit{of the \( m \)-ary sequence} \( s = (s_0, \ldots, s_{T-1}) \) with \( s_i \in \mathbb{Z}_m \) for \( i = 0, \ldots, T - 1 \) is

\[
\Gamma'_k(s) = \max_{\phi, M, D} |G(s, \phi, M, D)|,
\]

where the maximum is taken over all \( \phi = (\mu_{c_1}, \ldots, \mu_{c_k}) \) with \( \mu_{c_i} \in \mathcal{F}' \) for \( i = 1, \ldots, k \), the set \( D = (\tau_1, \ldots, \tau_k) \) of lags with non-negative integers \( 0 \leq \tau_1 < \cdots < \tau_k \) and for any non-negative integer \( M \) such that \( M + \tau_k \leq T \).

In particular, the \textit{correlation measure of order} \( k \geq 1 \) \textit{of a binary sequence} \( y = (y_n)_{0 \leq n < T} \) of period \( T \) is defined as

\[
C_k(y) = \max_{M, D} \left| \sum_{n=0}^{M-1} (-1)^{y_{n+\tau_1} + y_{n+\tau_2} + \cdots + y_{n+\tau_k}} \right|,
\]

with the same integer \( M \) and the set \( D \) of lags as before.

\subsection*{2.1.3. Linear complexity properties.}

Linear feedback shift registers are widely used in communications. A linear feedback shift register (LFSR) is a shift register if whose input bit is a linear function of its previous state. The initial value of the LFSR is called the seed or the initial state, and because the operation of the register is deterministic, the stream of values produced by the register is completely determined by its current (or previous) state. Likewise, because the register has a finite number of possible states, it must eventually enter a repeating cycle. However, an LFSR with a well-chosen feedback function can produce a sequence of bits which appears random and which has a very long cycle (period). Applications of LFSRs include generating pseudo-random numbers and pseudo-noise sequences.

For example, a linear function \( f(x_0, x_1, x_2) = x_0 + x_1 \) and the seed \( 1, 0, 0 \) generate the sequence \( 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, \ldots \). In fact, we input \( x_0 = 1, x_1 = 0, x_2 = 0 \) to the function to get the output \( x_3 = 1 \) and we input \( x_1 = 0, x_2 = 0, x_3 = 1 \) to get output \( x_4 = 0 \) and so on. This sequence has period 7.
The linear function also indicates a linear relation between an element and the fixed number of its previous consecutive elements. For this example, the relation is $x_n = x_{n-3} + x_{n-2}$. A few LFSR connection polynomials are defined on this relation. In the following we introduce the linear relation which is called linear recurrence relation and the polynomials defined on the relation.

**Definition 2.11.** For an integer $N \geq 2$ the $N$th linear complexity $L(s, N)$ of a sequence $s = (s_n)_{0 \leq n < T}$ of period $T$ over a ring $\mathbb{Z}_m$ is the smallest positive integer $L$ such that there are constants $c_1, \ldots, c_L \in \mathbb{Z}_m$ satisfying, for $0 \leq n < N - L$, the recursive relation

$$s_{n+L} = c_{L-1}s_{n+L-1} + c_{L-2}s_{n+L-2} + \cdots + c_0s_n.$$

If $s$ starts with $N - 1$ zeros, then we define $L(s, N) = 0$ if $s_{N-1} = 0$, and $L(s, N) = N$ if $s_{N-1} \neq 0$. The sequence $(L(s, N))_{1 \leq N < \infty}$ is the linear complexity profile of $s$. The linear complexity of $s$ is defined by

$$LC(s) = \sup_{N \geq 1} L(s, N).$$

**Definition 2.12.** [57] Let $s$ be a $m$-ary periodic sequence with period $T$. A polynomial of the form

$$f(x) = x^l - c_{l-1}x^{l-1} - \cdots - c_1x - c_0 \in \mathbb{Z}_m[x]$$

is called a characteristic polynomial of the sequence $s$ if

$$s_{n+l} = c_{l-1}s_{n+l-1} + c_{l-2}s_{n+l-2} + \cdots + c_1s_{n+1} + c_0s_n,$$

for all $0 \leq n < T - l$.

A characteristic polynomial of minimal degree $L$ is called the minimal polynomial, denoted $m(x)$, of $s$.

Obviously, the degree of the minimal polynomial is the linear complexity of the sequence, i.e.

$$\deg(m) = LC(s).$$

The minimal polynomial of a periodic sequence is uniquely determined. We can use the Berlekamp Massey algorithm to find the minimal polynomial of a linearly recurrent sequence. This algorithm is due to Massey (1969) who showed that the
iterative algorithm proposed in 1967 by Berlekamp for decoding BCH codes can be used for finding the shortest LFSR that generates a given sequence [61].

Given a sequence \( \mathbf{s} = (s_0, \ldots, s_{l-1}) \) of length \( l \), we want to find a shortest LSFR, such that it produces the sequence \( \mathbf{s} \) in the first \( l \) symbols. The main idea of the Berlekamp-Massey algorithm is to simulate the LFSR up to the \( n \)th symbol and see if the symbol is correct. If it is not, we alter the LFSR connection polynomial so that it produces a correct symbol. A very important remark is that our correction cannot cause previous symbols to be incorrect, because then we would have to start all over again. This makes the Berlekamp-Massey algorithm a so-called exact greedy algorithm. It produces an optimal solution that, however, needs not to be unique unless the length of the solution is less than or equal to half the length of the sequence. We adopt here the notation \( l \) and \( n \) as in [61].

**Theorem 2.13.** [61] If \( 2l < n \), then the LFSR synthesis algorithm with first \( n \) inputs produces the unique minimal-length solution.

This tells us that the sequence can be determined as long as at least two times the linear complexity of consecutive elements are known by the Berlekamp-Massey algorithm. It also shows large linear complexity and linear complexity profile are required. Therefore, the linear complexity profile is an important cryptographic characteristic of a sequence that provides information on the predictability and thus suitability for cryptography of the sequence.

Low linear complexity (profile) is undesirable in cryptography. More precisely, a good pseudorandom sequence should have linear complexity close to the period \( T \) and linear complexity profile close to \( N/2 \) for length \( N \leq 2T \) (see [72], p. 53) for given length \( N \).

The following lemma describes the computation of the minimal polynomial of a periodic sequence.

**Lemma 2.14.** [14, Proposition 2.3.2] Let \( \mathbf{s} = (s_n)_{0 \leq n < T} \) be a sequence of period \( T \) over \( \mathbb{Z}_m \) with generating polynomial be

\[
G_{\mathbf{s}}(x) := s_0 + s_1 x + \cdots + s_{T-1} x^{T-1}.
\]
Then the minimal polynomial \( m(x) \) of \( s \) is given by
\[
m(x) = \frac{x^T - 1}{\gcd(x^T - 1, G_s(x))}
\]
and the linear complexity of \( s \) is
\[
T - \deg(\gcd(x^T - 1, G_s(x))).
\]

2.1.4. Boolean functions.

**Definition 2.15.** Let \( \mathbb{F}_2^m = \{ (a_0, \ldots, a_{n-1}) \mid a_i \in \mathbb{F}_2 \} \) denote a vector space over \( \mathbb{F}_2 \) or dimension \( n \). A function with \( n \) binary inputs and one binary output is a boolean function of \( n \) variables. That is, \( f : \mathbb{F}_2^m \to \mathbb{F}_2 \), can be represented as:
\[
f(x_0, \ldots, x_{n-1}) = \sum c_{i_1 i_2 \ldots i_t} x_{i_1} x_{i_2} \cdots x_{i_t}, \quad c_{i_1 i_2 \ldots i_t} \in \mathbb{F}_2,
\]
where the sum runs through all subsets \( \{i_1, \ldots, i_t\} \) of \( \{0, 1, \ldots, n - 1\} \).

This shows that there are \( 2^{2^n} \) different boolean functions on \( n \) variables.

Equation (5) is called the **algebraic normal form** of \( f \). In Boolean logic, the algebraic normal form (ANF) provides a way for standardizing and normalizing logical formulas. The general ANF can be written as
\[
f(x_1, x_2, \ldots, x_n) = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n + \cdots + a_{1,2,\ldots,n} x_1 x_2 \cdots x_n,
\]
where coefficients \( a_0, a_1, \ldots, a_{1,2,\ldots,n} \in \{0,1\} \).

**Definition 2.16.** The **sparsity** \( \text{spr}(B) \) or **algebraic thickness** of a Boolean function is the minimal number of its nonzero coefficients in any polynomial representation, or equivalently, in its algebraic normal form.

Let \( F \) be a field. We say a function \( L : F^m \to F^n \) is linear if for any vectors \( x \) and \( y \) in \( F^m \), \( L(x + y) = L(x) + L(y) \) and for any scalar \( a \), \( L(ax) = aL(x) \). We say a function \( A : F^m \to F^n \) is affine if there is a linear function \( L : F^m \to F^n \) and a vector \( b \) in \( F^n \) such that \( A(x) = L(x) + b \) for all \( x \) in \( F^m \). In other words, an affine function is just a linear function plus a translation.

Any function from \( \mathbb{F}_2^m \) to \( \mathbb{F}_2 \) is a Boolean function. Boolean functions on the space \( \mathbb{F}_2^m \) are not only important in the theory of error-correcting codes, but also in cryptography where they occur, for example, in private key systems. We need functions that are as far as possible from being a linear function. Here the linear
functions are over a vector space. The nonlinearity is an important cryptographic parameter; see [10], [19], or [9]. In practice, it is useful to have boolean functions with high nonlinearity which is a measure of how far are the functions from being an affine function.

The **Hamming distance** between two Boolean functions is the number of positions in which their truth tables differ.

**Definition 2.17.** The **nonlinearity** of a Boolean function \( f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2 \) is the distance from \( f \) to the set of affine functions with \( m \) variables

\[
\mathfrak{NL}(f) = \min_{h \text{ is affine}} d(f, h),
\]

where \( d \) is the Hamming distance. Equivalently, nonlinearity is the number of bits which must be changed in the truth table of the Boolean function to reach the closest affine function.

**Definition 2.18.** [60] The **Fourier transform** of a Boolean function \( B \) is an integer valued function \( \hat{f} : \{0,1\}^n \rightarrow [-2^n, 2^n] \) defined as

\[
\hat{f}(b) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x)+b \cdot x},
\]

where \( b \in \mathbb{F}_2^n \) and the product \( b \cdot x \) denotes the conventional dot product in \( \mathbb{F}_2^n \).

The Hamming distance between two Boolean functions can be computed as

\[
d(f, h) := \sum_{x \in \mathbb{F}_2^n} (f(x) + h(x)) := 2^n - \sum_{x \in \mathbb{F}_2^n} \hat{f}(x) \hat{g}(x).
\]

The Hamming distance between two Boolean functions \( f \) and \( h \) can also be computed as

\[
d(f, h) = w_H(f + h),
\]

where \( w \) is the Hamming weight: the Hamming weight of a Boolean function is defined as

\[
w_H(f) = |\{x \in \mathbb{F}_2^n \mid f(x) = 1\}|,
\]

the number of nonzero element of the true table of the Boolean function.
The nonlinearity of a Boolean function $f$ is related to the maximum magnitude of the Fourier transform values, and given by

$$\mathcal{NL}(f) = 2^{n-1} - 2^{-1} \max_{h \in F_2^n} |\hat{f}(h)|.$$  

One can check that $\sum_{x \in F_2^n} (-1)^{f(x)+b \cdot x} = 2^n - 2w_H(f+b)$ for a Boolean function $f$, and (8) follows when substituting $d(f,b)$ to $w_H(f+b)$.

2.2. Maximal length sequences

Maximum length sequences (which are also called m-sequences, pseudo random sequences or pseudo noise sequences) naturally possess interleaved structure. The interleaved representation of maximal length sequences laid the foundation of the developments of new interleaved constructions. In this section we introduce maximal length sequences and their three different representations: polynomial representation, trace representation, and interleaved representation.

Maximum length sequences are certain $\ell$-ary sequences of length $T = \ell^n - 1$ that satisfy a linear recurrence given by the corresponding primitive polynomial (see Definition 2.20) of degree $n$ [64]. Although these sequences are not truly random because they can be predicted by a linear recurrence relation, nevertheless, they have many useful properties. Such maximal length sequences with good randomness properties are essential components in a wide variety of modern applications including radar, spread spectrum, error correction, cryptographic systems, and Monte Carlo simulations. Acceptable maximal length sequences should exhibit no statistical bias in the occurrence of individual symbols or small blocks of symbols. With these goals in mind, Golomb [26] defined a maximal length sequence to be a periodic $\ell$-ary sequence that passes three well known statistical tests for randomness: balance, run, and ideal two-level autocorrelation (see Section 2.1).

2.2.1. Polynomial representation. Maximum length sequences can be generated by primitive polynomials over a certain finite field rapidly. First we give the definition of a minimal and of a primitive polynomial.

Definition 2.19. The minimal polynomial of an element $\alpha$ over a field $F$ is the monic polynomial over $F$, of lowest degree, having $\alpha$ as a zero.
**Definition 2.20.** A *primitive polynomial* is the minimal polynomial of a primitive element. A primitive element is a generator of the multiplicative group of a finite field.

For a sequence \(a\) we recall that \(L^i(a) = (a_i, a_{i+1}, \ldots)\). Then

\[
a_{n+k} = \sum_{i=0}^{n-1} c_i a_{k+i}
\]

\[
L^n(a) = \sum_{i=0}^{n-1} c_i L^i(a)
\]

\[
0 = \left( L^n - \sum_{i=0}^{n-1} c_i L^i \right) (a).
\]

For \(f(x) = x^n - \sum_{i=0}^{n-1} c_i x^i\), define

\[
f(L) = L^n - \sum_{i=0}^{n-1} c_i L^i.
\]

Then

(9) \(f(L)(a) = 0\).

Given a sequence \(a\), the monic polynomial \(m(x)\) satisfying (9) with the lowest degree is called the *minimal polynomial* of \(a\). If the minimal polynomial \(m(0) \neq 0\), then \(a\) is a periodic sequence. The period of a polynomial \(m(x)\) is defined as the least positive integer \(r\) such that

\[
m(x) \mid (x^r - 1)
\]

denoted by \(\text{per}(m(x)) = r\).

**Proposition 2.21.** \(\text{per}(a) = \text{per}(m(x))\).

**Proof.** If \(\text{per}(m(x)) = r\), then \(m(x) \mid (x^r - 1)\). It follows \((L^r - 1)(a) = 0\). Therefore,

\[
\text{per}(a) \leq \text{per}(m(x)).
\]

If \(\text{per}(a) = r\), we have \((L^r - 1)(a) = 0\). It follows \(m(x) \mid (x^r - 1)\). We get

\[
\text{per}(m(x)) \leq \text{per}(a).
\]

\(\square\)
If the minimal polynomial $m(x)$ of a sequence $g$ is a primitive polynomial over $\mathbb{F}_\ell$ of degree $n$, then $g$ is a maximal length sequence. We give the definition below.

We let the primitive polynomial $p(x)$ of degree $n$ over $\mathbb{F}_\ell$ be as

$$p(x) = x^n - \sum_{i=0}^{n-1} a_i x^i$$

where, all the coefficients $a_i$, $i = 0, \ldots, n-1$ are in $\mathbb{F}_\ell$ with $a_0 = 1$ (since we consider periodic sequences).

A maximal length sequence is an $(\ell^n - 1)$-length sequence that satisfies a linear recurrence, defined over $\mathbb{F}_\ell$, given by a corresponding primitive polynomial $p(x)$ of degree $n$.

For any primitive polynomial $p(x)$, the reciprocal polynomial $p^*(x)$, defined as $p^*(x) = x^n p(x^{-1})$, is also a primitive polynomial, and the maximal length sequence generated by $p^*(x)$ is exactly the reverse of the maximal length sequence generated by $p(x)$. The polynomial $p(x)$ is called the characteristic polynomial and the reciprocal polynomial of $p(x)$ is called the feedback polynomial.

2.2.2. Trace representation. In this section we discuss the trace representations for maximal length sequences. The trace function is defined from $\mathbb{F}_{\ell^n}$ to $\mathbb{F}_\ell$ as

$$\text{Tr}_{\ell^n}^\ell(x) = x + x^\ell + \cdots + x^{\ell^{n-1}}.$$ 

Some properties that trace functions possess are:

- $\text{Tr}_{\ell^n}^\ell(x + y) = \text{Tr}_{\ell^n}^\ell(x) + \text{Tr}_{\ell^n}^\ell(y)$;
- $\text{Tr}_{\ell^n}^\ell(x^\ell) = \text{Tr}_{\ell^n}^\ell(x)$;
- $\text{Tr}_{\ell^n}^\ell(x) = 0$ has $\ell^{n-1}$ roots;
- $\sum_{x \in \mathbb{F}_{\ell^n}} (-1) \text{Tr}_{\ell^n}^\ell(x) = 0$.

Let $\alpha$ be a primitive element of $\mathbb{F}_{\ell^n}$. Then for nonzero $\beta \in \mathbb{F}_{\ell^n}$

$$(\text{Tr}_{\ell^n}^\ell(\beta), \text{Tr}_{\ell^n}^\ell(\beta\alpha), \ldots, \text{Tr}_{\ell^n}^\ell(\beta\alpha^{\ell^{n-2}}))$$

is a maximal length sequence.

The following properties of maximal length sequences can be found in [3], [20], and [59]:

- Balance property: Since $\text{Tr}_{\ell^n}^\ell(x) = 0$ has $q^{n-1}$ roots, a maximal length sequence of period $\ell^n - 1$ has $\ell^{n-1}$ copies for each nonzero element in $\mathbb{F}_\ell$
and \(\ell^{n-1} - 1\) 0’s.

• **Shift-and-add property:** Let \(T\) denote the period of a maximal length sequence \(a\). Then

\[
a + L^\tau(a) = L^\tau'(a)
\]

where \(0 < \tau' \neq \tau < T\) is an integer. We use the trace representation of a maximal length sequence to get

\[
a + L^\tau(a) = \left( \text{Tr}_{n1}(\beta) + \text{Tr}_{n1}(\beta \alpha^\tau), \text{Tr}_{n1}(\beta \alpha^{\tau+1}), \ldots, \text{Tr}_{n1}(\beta \alpha^{\tau+\ell^{n-2} \mod \ell^{n-1}}) \right).
\]

Since \(\alpha\) is primitive and \(1 + \alpha^\tau\) can also be expressed by a power of \(\alpha\), say \(\alpha^\tau'\), we get

\[
a + L^\tau(a) = \left( \text{Tr}_{n1}(\beta (1 + \alpha^\tau)), \text{Tr}_{n1}(\beta (1 + \alpha^\tau+1)), \ldots, \text{Tr}_{n1}(\beta (1 + \alpha^\tau+\ell^{n-2} \mod \ell^{n-1})) \right).
\]

• **Ideal two-level autocorrelation:** The autocorrelation \(C_a\) of any binary maximal length sequence \(a\) of length \(N\) is given by

\[
C_a(\tau) = \begin{cases} 
N, & \text{for } \tau = 0; \\
-1, & \text{for } 1 \leq |\tau| \leq N - 1.
\end{cases}
\]

If the maximal length sequence is binary, that is \(p = 2\), the proof is given as follows.

If \(A\) is the number of places where any sequence \((a_0, \ldots, a_{2^n-2})\) and its \(\tau\) cyclic shift \((a_\tau, \ldots, a_{\tau+2^n-2} \mod 2^n-1)\) agree and \(D\) is the number of places where they disagree so that \(A + D = N\), then, from the definition of autocorrelation,

\[
C_a(\tau) = \frac{A - D}{N}.
\]

From the Shift-and-Add property, the element-wise sum of the above two sequences must be a maximal length sequence. Then \(D\) would denote the number of 1’s in the resultant sequence and \(A\) would represent the number
of 0’s modulo 2. By the balance property, $A = 2^{n-1} - 1$ and $D = 2^{n-1}$ for any $\tau \neq 0$. For $\tau = 0$, the resultant sequence is an all zero sequence, leading to a value of autocorrelation 1.

If $q \neq 2$, we can use the properties of exponential sums to obtain the result; see more details in [27].

2.2.3. Interleaved representation. We use the concept of decimated sequence to give the interleaved representation of a maximal length sequence.

**Definition 2.22.** Let $a = (a_0, \ldots, a_{T-1})$ be a sequence and $0 < d < T$ be a positive integer. Define a new sequence $a^{(d)} = (a_0^{(d)}, \ldots, a_i^{(d)}, \ldots)$ as

$$a_i^{(d)} = a_{d \cdot i} \pmod{T}$$

which is called a $d-$decimated sequence of $a$.

There are important properties showing the relation between a sequence $a$ and its $d-$decimated sequence $a^{(d)}$:

1. $$\text{per}(a^{(d)}) = \frac{\text{per}(a)}{\gcd(\text{per}(a), d)}.$$

2. Assume $\gcd(\text{per}(a), d) = 1$. Then

$$C_{a^{(d)}}(\tau) = C_{a^{(d)}}(\tau).$$

Indeed,

$$C_{a^{(d)}}(\tau) = \sum_{i=0}^{T-1} (-1)^{a_i^{(d)} + a_{i+d\tau}} = \sum_{i=0}^{T-1} (-1)^{a_i + a_{i+d\tau}} = \sum_{i=0}^{T-1} (-1)^{a_i + a_{i+d\tau}}.$$

Assume that $a$ is a maximal length sequence over $\mathbb{F}_\ell$ of degree $n$. Let $f(x)$ be the minimal polynomial of $a$, $\alpha$ a root of $f(x)$ in the extension field $\mathbb{F}_{\ell^n}$, $d = \frac{\ell^n - 1}{\ell - 1}$, and $\beta = \alpha^d$. Then $\beta$ is a primitive element in $\mathbb{F}_\ell$. Using the trace representation of $a$, for $0 \leq k \leq \ell^n - 1$, we write $k = id + j$ where $0 \leq j < d$, $0 \leq i < \ell - 1$. Then

$$a_k = a_{id+j} = \text{Tr}(\eta \alpha^{id+j}) = \alpha^{id} \text{Tr}(\eta \alpha^j) = \beta^i a_j,$$

that is,

$$a_{id+j} = \beta^i a_j, \quad 0 \leq j < d, \quad 0 \leq i < \ell - 1.$$
We arrange \( \underline{a} \) into an \((\ell - 1) \times d\) array, \( A = (a_{ij}) \), whose entries are given by 
\[ a_{ij} = a_{id+j}. \]
Then \( A \) is the array form of the sequence \( \underline{a} \) that has the following structure:

- Let \( R \) be the first row vector denoted as \( R = (a_0, \ldots, a_{d-1}) \). The other row vectors are multiples of the first row. The array form of the maximal length sequence can be written as

\[
\begin{bmatrix}
R \\
\beta R \\
\vdots \\
\beta^{\ell-2} R
\end{bmatrix},
\]

which is called the projective property of the maximal length sequence.

- Each column vector of the array, that is, a \( d \)-decimated sequence given in Definition 2.22 of \( \underline{a} \), is either a maximal length sequence of length \( \ell - 1 \) or a zero sequence, depending on whether \( \text{Tr}(\alpha^j) \) is equal to zero for \( 0 \leq j < d \). Also all nonzero column sequences are shift-equivalent maximal length sequences. More precisely, the \((\ell - 1) \times d\) array is given by

\[
\begin{pmatrix}
a_0 & a_1 & \cdots & a_j & \cdots & a_{d-1} \\
\beta a_0 & \beta a_1 & \cdots & \beta a_j & \cdots & \beta a_{d-1} \\
\vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\
\beta^{\ell-2} a_0 & \beta^{\ell-2} a_1 & \cdots & \beta^{\ell-2} a_j & \cdots & \beta^{\ell-2} a_{d-1}
\end{pmatrix}
\]
or

\[
(A_0, A_1, \ldots, A_{d-1}),
\]

where \( A_j \) is the \( j \)th column of \( A \), which is given by

\[ A = (a_j, \beta a_j, \ldots, \beta^{\ell-2} a_j)^T, \quad 0 \leq j < d. \]

Each column is either a maximal length sequence over \( \mathbb{F}_\ell \) of length \( \ell - 1 \) or a zero sequence depending on whether or not \( a_j \) is zero.
• The shift sequence of $a$ is determined by

$$e_j = \begin{cases} 
  e, & \text{if } 0 \neq a_j = \beta^e; \\
  \infty, & \text{if } a_j = 0.
\end{cases}$$

If $n$ is a composite number, let $m$ be a proper factor of $n$ and $d = \frac{\ell^m - 1}{\ell - 1}$. Then any maximal length sequence over $\mathbb{F}_\ell$ of degree $n$ can be arranged into a $(\ell^m - 1) \times d$ array where each column sequence is either a maximal length sequence over $\mathbb{F}_\ell$ of degree $m$ or a zero sequence for which all the maximal length sequences are shift equivalent.

**Example 2.23.** We give an example of a maximal length sequence with interleaved structure. We use the primitive polynomial $f(x) = x^6 + x + 1$ over $\mathbb{F}_2$ to generate a maximal length sequence

$$a = (000100001100010100111010011010111011001101010111111)$$

of period $2^6 - 1$. We arrange this sequence into a 3 by 21 array, where each column is a maximal length sequence of period 3 or a zero sequence:

$$
\begin{array}{cccccccccccccc}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}
$$

We can view this 3 by 21 array as interleaving a binary maximal length sequence $a = (1, 0, 1)$ with the shift sequence

$$e = (1, 1, 1, \infty, 1, 0, \infty, \infty, 1, 1, 2, 0, \infty, 1, \infty, 0, 1, 0, 1, 1, 0).$$

It can also be arranged into a 7 by 9 array, where each column is a maximal length sequence of period 7 or a zero sequence:

$$
\begin{array}{cccccccccc}
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 . \\
0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}
$$
We can view this 7 by 9 array as interleaving a binary maximal length sequence \( \mathbf{a} = (1, 1, 1, 0, 1, 0, 0) \) with the shift sequence \( \mathbf{e} = (\infty, 3, 6, 5, 2, 3, 5, 3) \).

### 2.3. Construction of GMW sequences by interleaved approach

When the integer \( m \) is a composite, i.e. \( m = j \times k \), we define a GMW sequence \( \{b_n\} \) of period \((2^m - 1)\) as:

\[
b_n = \text{Tr}_{2^m}^2 ([\text{Tr}_{2^j}^m(\alpha^n)]^r)
\]

where \( n = 0, 1, \ldots, 2^m - 2 \), \( \alpha \) is a primitive element of \( \mathbb{F}_{2^m} \), and \( 0 < r < 2^j - 1 \) is any integer relatively prime to \( 2^j - 1 \), i.e. \( \gcd(r, 2^j - 1) = 1 \).

The GMW sequence has the maximal length \( 2^m - 1 \) and perfect autocorrelation \([32]\). Now we give the algorithm to obtain GMW sequences.

**Algorithm 2.24. (GMW Algorithm)**

**Input:** The minimum polynomial \( m_\alpha(x) \) of the primitive element \( \alpha \) of \( \mathbb{F}_{2^m} \) over \( \mathbb{F}_2 \). The values of \( j, k \) and \( r \).

1. **Step 1:** Compute \( \alpha^0, \alpha^1, \alpha^2, \ldots, \alpha^{2^m-2} \) from \( m_\alpha(x) \) over \( \mathbb{F}_2 \).
2. **Step 2:** For all \( n = 0, 1, \ldots, 2^m - 2 \), evaluate

\[
b_n = \sum_{u=0}^{j-1} \sum_{h_1=0}^{k-1} \cdots \sum_{h_w=0}^{k-1} \alpha^{n C(k,r)2^u}
\]

where \( C(k,r) = \sum_{i=1}^{w} 2^{jh_i+t_i} \).

**Example 2.25.** A GMW sequence of period 63 is

\[b_n = \text{Tr}_{2^3}^2 ([\text{Tr}_{2^2}^3(\alpha^n)]^3),\]

\( n = 0, 1, \ldots, 62 \) where \( m_\alpha(x) = x^6 + x^5 + x^2 + x + 1 \) over \( \mathbb{F}_2 \), \( m = 6 = 3 \times 2 = j \times k \) and \( r = 3 \) is chosen.

Applying the GMW algorithm yields the following sequence:

\[b_n = (0000010100100011110010110001100111100110011110011010000100110011110011).\]

### 2.4. Character sums over finite fields

Character sums can be used to estimate bounds of some cryptographic properties of power residue sequences. In this section we first give the definition of
characters of a finite field, then we introduce some sequences defined using characters over finite fields, and finally we give some well-known bounds on character sums.

2.4.1. Characters of a finite field.

Definition 2.26. Let $G$ be a finite abelian group (written multiplicatively). A character $\chi$ of $G$ is a homomorphism from $G$ into the multiplicative group $U$ of complex numbers of absolute value 1, that is, a mapping $\chi : G \rightarrow U$ with $\chi(g_1g_2) = \chi(g_1)\chi(g_2)$ for all $g_1, g_2 \in G$ satisfying $\chi(0) = 0$, $\chi(1) = 1$.

In a finite field $\mathbb{F}_\ell$ there are two finite abelian groups that are the additive group $(\mathbb{F}_\ell, +)$ and the multiplicative group $(\mathbb{F}_\ell^*, \cdot)$ of the field, where $\mathbb{F}_\ell^* = \mathbb{F}_\ell \setminus \{0\}$. Therefore, we have to make an important distinction between the characters pertaining to these two group structures.

Remark 2.27. The character defined on the additive group $(\mathbb{F}_\ell, +)$ mapping into the multiplicative group $U$ of complex numbers is the additive character $\psi$. An additive character $\psi$ satisfies $\psi(x + y) = \psi(x)\psi(y)$ for all $x, y \in \mathbb{F}_\ell$. Similarly, the character $\chi$ defined on $(\mathbb{F}_\ell^*, \cdot)$ and satisfying $\chi(x \cdot y) = \chi(x)\chi(y)$ is the multiplicative character.

We recall the definition of trace function. Let $p$ be a prime number and $\ell$ a power of $p$. For a positive integer $n$, the trace function from the finite field $\mathbb{F}_\ell^n$ to its subfield $\mathbb{F}_\ell$ is defined by

$$\text{Tr}_{\ell}^n(x) = x + x^\ell + x^\ell^2 + \cdots + x^\ell^{n-1}, \quad x \in \mathbb{F}_\ell^n,$$

and if $\ell = p^n$ the trace function from $\mathbb{F}_\ell$ to its prime field $\mathbb{F}_p$ is

$$\text{Tr}_{\ell/p}(x) = x + x^{p} + x^{p^2} + \cdots + x^{p^{m-1}}, \quad x \in \mathbb{F}_\ell.$$

Example 2.28. Consider the additive group of $\mathbb{F}_\ell$. Let $p$ be the characteristic of $\mathbb{F}_\ell$; then the prime field contained in $\mathbb{F}_\ell$ is $\mathbb{F}_p$, which we identify with $\mathbb{Z}_p$. Let $\text{Tr}_{\ell/p} : \mathbb{F}_\ell \rightarrow \mathbb{F}_p$ be the absolute trace function from $\mathbb{F}_\ell$ to $\mathbb{F}_p$. Then the function $\psi_1$ defined by $\psi_1(x) = \exp \left( \frac{2\pi \sqrt{-1} \text{Tr}_{\ell/p}(x)}{p} \right)$ for all $x \in \mathbb{F}_\ell$ is a character of the additive group of $\mathbb{F}_\ell$, since for $x_1, x_2 \in \mathbb{F}_\ell$ we have $\text{Tr}_{\ell/p}(x_1 + x_2) = \text{Tr}_{\ell/p}(x_1) + \text{Tr}_{\ell/p}(x_2)$, and so $\psi(x_1 + x_2) = \psi(x_1)\psi(x_2)$. 
We adopt the notation \( e_p(\ast) \) to denote \( \exp\left(\frac{2\pi\sqrt{-1}}{p}\right) \). The character above can be written as \( \psi_1(x) = e_p\left(\text{Tr}_{\ell/p}(x)\right) \).

For each \( b \in \mathbb{F}_\ell \), the function

\[
\psi_b(x) = \exp\left(\frac{2\pi\sqrt{-1}\text{Tr}_{\ell/p}(bx)}{p}\right), \quad \text{for all } x \in \mathbb{F}_\ell
\]
defines an additive character of \( \mathbb{F}_\ell \). When \( b = 0 \), \( \psi_0(x) = 1 \) for all \( x \in \mathbb{F}_\ell \), and is called the trivial additive character of \( \mathbb{F}_\ell \).

Let \( \alpha \) be a fixed primitive element of \( \mathbb{F}_\ell \). For each \( k = 0, 1, \ldots, \ell - 2 \), the function \( \chi_k \) with

\[
\chi_k(\alpha^j) = \exp\left(\frac{2\pi\sqrt{-1}jk}{\ell - 1}\right), \quad \text{for } j = 0, 1, \ldots, \ell - 2
\]
defines a multiplicative character of \( \mathbb{F}_\ell \). When \( k = 0 \), \( \chi_0(x) = 1 \) for all \( x \in \mathbb{F}_\ell^* \), and is called the trivial multiplicative character of \( \mathbb{F}_\ell \).

**Definition 2.29.** Let \( \chi_0 \) denote the trivial multiplicative character, that is \( \chi_0(c) = 1 \) for all \( c \in \mathbb{F}_\ell^* \). The order of a multiplicative character \( \chi \) is the least positive integer \( s \) such that \( \chi^s = \chi_0 \).

**2.4.2. Sequences defined by characters.** A polyphase sequence is a sequence whose terms are complex roots of unity of the form

\[
a_n = e^{\frac{2\pi\sqrt{-1}x_n}{N}} \text{ for some positive integer } N, \text{ where } x_n \text{ is an integer.}
\]

Multiplicative characters map the elements from a finite field to complex values. Therefore, multiplicative characters have very important applications on polyphase sequences, particularly, on one important class of sequences: power residue sequences.

**Definition 2.30.** Let \( p \) be an odd prime and \( m \) a divisor of \( p - 1 \). Let \( \alpha \) be a primitive element of \( \mathbb{F}_p \). The nonzero elements in \( \mathbb{F}_p^* \) can be partitioned into \( m \) cosets \( U_k := \{\alpha^n \mid n \equiv k \pmod{m}\} \) for \( k = 0, \ldots, m - 1 \) and \( n = 0, 1, \ldots, p - 1 \). The \( m \)-phase (\(-\)-ary) power residue sequence \( (y_n)_{0 \leq n < p} \) of period \( p \) is defined by

\[
y_n = \begin{cases} 
0, & \text{if } n = 0; \\
k, & \text{if } n \in U_k.
\end{cases}
\]
Remark 2.31. The above $m$–phase power residue sequence can also be defined as

$$y_n = \log_\alpha n \pmod{m} \text{ for } n \not\equiv 0 \pmod{m}.$$  

The complex-valued sequence $\omega_{m}^{y_n}$, where $\omega_m$ is the $m$th root of unity, can be represented by a multiplicative character

$$\psi(x) = \exp\left(\frac{2\pi\sqrt{-1}\log_\alpha x}{m}\right)$$  

of $\mathbb{F}_p$, that is for $n \not\equiv 0 \pmod{p}$,

(11)  

$$\omega_{m}^{y_n} = \psi(n).$$

As a special case of power residue sequences, we introduce Legendre symbols and Legendre sequences.

Definition 2.32. Let $p$ be an odd prime number. An integer $i$ is a quadratic residue modulo $p$ if it is congruent to a perfect square modulo $p$ and is a quadratic non-residue modulo $p$ otherwise. The Legendre symbol is defined as:

$$\left(\frac{i}{p}\right) = \begin{cases} 
0, & \text{if } i \equiv 0 \pmod{p}; \\
1, & \text{if } i \text{ is a quadratic residue modulo } p; \\
-1, & \text{if } i \text{ is a quadratic non-residue modulo } p.
\end{cases}$$

Definition 2.33. Consider $p = 7$. The Legendre sequence $s = \{s_i | i \geq 0\}$ of period $p$ is defined as

$$s_i = \begin{cases} 
1, & \text{if } i \equiv 0 \pmod{p}; \\
0, & \text{if } i \text{ is a quadratic residue modulo } p; \\
1, & \text{if } i \text{ is a quadratic non-residue modulo } p.
\end{cases}$$

Example 2.34. Let odd prime $p$ be 7. Observe that $1^2 = 1 \pmod{7}$, $2^2 = 4 \pmod{7}$, $3^2 = 2 \pmod{7}$, $4^2 = 2 \pmod{7}$, $5^2 = 4 \pmod{7}$, and $6^2 = 1 \pmod{7}$. Legendre symbols for $p = 7$ are

$$\left( \begin{array}{cccccc} 
0 & + & + & - & + & - \\
\end{array} \right).$$

The Legendre sequence is

$$\left( \begin{array}{ccccccc} 
1 & 0 & 0 & 1 & 0 & 1 & 1 \\
\end{array} \right).$$

A Sidelnikov sequence is defined as follows.
Definition 2.35. Let $p$ be an odd prime and $m$ be a positive integer. Let $\mathbb{F}_{p^m}$ be the finite field with $p^m$ elements, and $\alpha$ be a primitive element of $\mathbb{F}_{p^m}$. The Sidelnikov sequence $s = \{s(t) \mid t = 0, 1, 2, \ldots, p^m-2\}$ of period $p^m - 1$ is defined as [78]

$$s(t) = \begin{cases} 1, & \text{if } \alpha^{t+1} \in N; \\ 0, & \text{otherwise}, \end{cases}$$

where $N = \{\alpha^{2t+1} \mid t = 0, 1, \ldots, p^m-1\}$ is the set of quadratic non-residues over $\mathbb{F}_{p^m}$.

Definition 2.36. Let $p$ and $q$ be two distinct odd primes. Define $Q := \{q, 2q, \ldots, (p-1)q\}$, $Q_0 = Q \cup \{0\}$ and $P := \{p, 2p, \ldots, (q-1)p\}$.

The $pq$-periodic sequence $(s_i)_{0 \leq i < pq}$ over $\mathbb{F}_2$ is defined by

$$s_i = \begin{cases} 0, & \text{if } i \mod pq \in Q_0; \\ 1, & \text{if } i \mod pq \in P; \\ \frac{1-(\frac{i}{p})(\frac{i}{q})}{2}, & \text{otherwise}. \end{cases}$$

The sequence $(s_i)_{0 \leq i < n}$ is the two-prime generator (or generalized cyclotomic sequence) of order 2.

The sequences above are defined by multiplicative characters. The following sequence is defined by a multiplicative character and the trace function.

Definition 2.37. [54] Let $\ell$ be a prime power and $\alpha$ be a primitive element of $\mathbb{F}_{\ell^n}$. Let a maximal length sequence of period $\ell^n - 1$ over $\mathbb{F}_\ell$ be given by

$$u(t) = \text{Tr}_{1}^n(\alpha^{\ell^t}), \ t = 0, 1, \ldots, \ell^n - 2,$$

where $\text{Tr}_{1}^n(\cdot)$ denotes the trace mapping from $\mathbb{F}_{\ell^n}$ to $\mathbb{F}_\ell$ and $\gcd(s, \ell^n - 1) = 1$.

Define the sequence $a(t)$ over $\Omega_{\ell-1}^+=\{1, \omega_{\ell-1}, \omega_{\ell-1}^2, \ldots, \omega_{\ell-1}^{\ell-2}\} \cup \{0\}$ as

$$a(t) = \begin{cases} \omega_{\ell-1}^t \cdot \omega_{\ell-1}^{\text{ind}_a(u(t))}, & \text{if } u(t) \neq 0, \\ 0, & \text{otherwise}, \end{cases}$$
for $t = 0, 1, \ldots, \ell^n - 2$, where $\omega_{\ell^{-1}}$ is a $(\ell - 1)$th root of unity and $\text{ind}_\beta x$ denotes the discrete logarithm of nonzero $x$ with respect to $\beta = \alpha^{\ell^{-1}}$ of order $\ell - 1$. The sequence \{a(t)\} is the Lee sequence.

Later this definition was generalized as follows.

**Definition 2.38.** [55] Let $\ell$ be a prime power and $\alpha$ be a primitive element of $\mathbb{F}_{\ell^n}$. Let a maximal length sequence of period $\ell^n - 1$ over $\mathbb{F}_\ell$ be given by

$$u(t) = \text{Tr}_{\ell^n}^n(\alpha^{st}), \quad t = 0, 1, \ldots, \ell^n - 2,$$

where $\text{Tr}_{\ell^n}(\cdot)$ denotes the trace mapping from $\mathbb{F}_{\ell^n}$ to $\mathbb{F}_{\ell}$ and $\gcd(s, \ell^n - 1) = 1$. Define the sequence $a(t)$ over $\Omega_{\ell^{-1}}^+ := \{1, \omega_{\ell^{-1}}, \omega_{\ell^{-1}}^2, \ldots, \omega_{\ell^{-1}}^{\ell-2}\}$ as

$$a(t) = \begin{cases} 
\omega_{\ell^{-1}}^t \cdot \psi(u(t)), & \text{if } u(t) \neq 0, \\
0, & \text{otherwise},
\end{cases}$$

for $t = 0, 1, \ldots, \ell^n - 2$, where $\omega_{\ell^{-1}}$ is a $(\ell - 1)$th root of unity and $\psi$ is a multiplicative character of $\mathbb{F}_\ell$. The sequence \{a(t)\} is the generalized Lee sequence.

**2.4.3. Bounds on character sums.** A multiplicative character $\psi$ is a homomorphism from the multiplicative group $\langle \mathbb{F}_{p^m}^*, \cdot \rangle$ to the unit circle. In this section we give well-known bounds for the character sum $\sum x \psi(x)$ over a subset of $\mathbb{F}_{p^m}$.

The next is an estimate of the complete character sum of a polynomial using the well-known Weil’s bound on exponential sums.

**Theorem 2.39. (Weil’s Theorem)** Let $\psi$ be a multiplicative character of $\mathbb{F}_{p^n}$ of order $d > 1$. Suppose a monic polynomial $f \in \mathbb{F}_{p^n}[x]$ with positive degree has $m$ distinct roots and $f$ is not a $d$th power. Then for $n \geq 1$ we have

$$\left| \sum_{x \in \mathbb{F}_{p^n}} \psi(f(x)) \right| \leq (m - 1)p^{n/2}.$$

The so-called incomplete sums arise from a restriction of the range of summation, namely by summing only over “intervals” or “boxes”. Incomplete sums have been considered mostly for prime fields $\mathbb{F}_p$. For incomplete sums with non-trivial multiplicative character $\psi$ of $\mathbb{F}_p$ there is a classical inequality of Pólya and Vinogradov shown in the next theorem.
Theorem 2.40. Let $\psi$ be a non-trivial multiplicative character modulo $p$. Then

$$\left| \sum_{m=a+1}^{a+b} \psi(m) \right| < Cp^{1/2}(\log p),$$

where $C$ is a constant.

We observe that this bound is only nontrivial when $b > p^{1/2}(\log p)$.

Notation 2.41. We write $A \ll B$ or $A = O(B)$ if $|A| \leq kB$ for some constant $k$.

For a prime field $\mathbb{F}_p$ and when the sum is over an interval, Pólya and Vinogradov give the following estimate that follows from Weil’s inequality [77].

Theorem 2.42. If $\psi$ is a multiplicative character modulo $p$ of order $d$, and $f(x)$ is a polynomial that is not a $d$th power modulo $p$, then

$$\sum_{x=N+1}^{N+H} \psi(f(x)) \ll p^{1/2} \log p.$$

The following theorems give bounds on incomplete sums of Legendre symbols.

Theorem 2.43. Let $1 \leq N < p$ and $f(X) \in \mathbb{Z}[X]$ be a monic polynomial which is not a square in $\mathbb{Z}_p[X]$. Then we have

$$\left| \sum_{x=0}^{N-1} \left( \frac{f(x)}{p} \right) \right| < (\deg f)p^{1/2} \log p.$$

2.5. Costas sequences

In the thesis Costas sequences are employed in a construction. We introduce Costas sequences in this section.

Costas arrays were introduced by Costas in the context of SONAR detection [13]. Costas arrays are $n \times n$ arrays of dots and blanks with exactly one dot in each row and column, and with distinct vector differences between pairs of dots. Two well-known algebraic constructions of Costas arrays are the Welch-Costas arrays and the Lempel-Golomb Costas arrays. These methods successfully construct $n \times n$ Costas arrays for infinitely many, but not all, orders $n$. Welch-Costas arrays are constructed taking a primitive root $\alpha$ for a prime number $p$ and defining the $(i,j)$th entry of the array $A$ to be 1 if $i \equiv \alpha^j \pmod{p}$. Golomb [28] showed that finite fields can be used to generate Costas arrays.
Note that there exist Costas sequences not generated algebraically. However, the two types of algebraic-generated Costas sequences shown below will be used in our result.

**Definition 2.44.** Let \( \{a_1, a_2, \ldots, a_N\} \) be an ordering of the elements in the set \( \{1, 2, \ldots, N\} \). A difference triangle is an array containing in the \( j \)th row the differences \( a_{i+j} - a_i \) for all \( i = 1, \ldots, N-j \). A Costas array (sequence) is a difference triangle with no repeated values on any given row.

**Definition 2.45.** Given a prime number \( p \) and a primitive element \( \alpha \in \mathbb{F}_p \), \( \{\alpha^1, \alpha^2, \ldots, \alpha^{p-1}\} \) is a length \( p-1 \) Welch Costas sequence.

**Example 2.46.** In \( \mathbb{F}_{11} \), 2 is a primitive root, we get a length 10 Costas array (sequence) since \( \{2^1, 2^2, \ldots, 2^{10}\} = \{2, 4, 8, 5, 10, 9, 7, 3, 6, 1\} \).

**Definition 2.47.** [28] Given a prime power \( \ell = p^m \) and primitive elements \( \alpha, \beta \in \mathbb{F}_\ell \), \( \{\log_\beta(1-\alpha^1), \log_\beta(1-\alpha^2), \ldots, \log_\beta(1-\alpha^{\ell-2})\} \) is a length \( \ell - 2 \) Lempel-Golomb Costas sequence.

**Example 2.48.** Let \( q = 11 \) and \( \alpha = 2, \beta = 3 \). Then we have the Lempel-Golomb Costas sequence of length 9

\[
\{\log_3(1-2^1), \ldots, \log_3(1-2^9)\} = \{\log_3(10), \log_3(8), \log_3(4), \log_3(7), \log_3(2), \log_3(3), \log_3(5), \log_3(9), \log_3(6)\} = \{5, 7, 8, 3, 9, 2, 6, 4, 1\}.
\]

### 2.6. Some well-known signal sets (families)

In the previous sections we introduced many sequences with nice properties. In this section we briefly review some signal sets that are related to maximal length sequence which we introduced before. The signal sets introduced in this section are derived from a decimation and have found applications in communications. We recall that the \( d \)-fold decimation of \( \mathbf{a} = (a_0, a_1, \ldots) \) is the sequence \( \mathbf{b} = (b_0, b_1, \ldots) \) where \( b_i = a_{di} \), which consists of every \( d \)th element from \( \mathbf{a} \), denoted \( \mathbf{b} = \mathbf{a}^{(d)} \).

The constructed signal sets are binary. Certainly, they can be generalized to \( q \)-ary, where \( q \) is a prime power but here we consider the sequences only over \( \mathbb{F}_2 \).
Construction 2.49. [23] (Gold-pair construction)

Let \( n \) be an odd integer. Pick two maximal length sequences \( a \) and \( b \). Let 
\[
a = \text{Tr}(\alpha^i),
\]
where \( \alpha \) is a primitive element of \( \mathbb{F}_{2^n} \) and \( b = a^{(d)} \), the \( d \)-decimated sequence of \( a \).

Define a sequence
\[
z_j = L^j(a) + b, \quad 0 \leq j < 2^n - 1
\]
and the set
\[
S(d) = \{z_j \mid 0 \leq j < 2^n - 1\} \bigcup \{a, b\}
\]
consisting of \( 2^n + 1 \) sequences of period \( 2^n - 1 \).

Selecting \( d \) from the following list we can get different signal sets (families):

(1) to obtain Gold family pick \( d = 2^k + 1 \), \( \gcd(k, n) = 1 \) and \( k \leq \frac{n-1}{2} \);

(2) to obtain the Kasami (large set) family pick \( d = 2^{2k} - 2^k + 1 \), \( \gcd k, n = 1 \) and \( k \leq \frac{n-1}{2} \);

(3) to obtain the Welch family pick \( d = 2^{\frac{n-1}{2}} + 3 \);

(4) to obtain the Niho family pick \( d = 2^{2k} + 2^k - 1 \) where
\[
k = \begin{cases} 
\frac{n-1}{4}, & \text{if } n \equiv 1 \pmod{4}, \\
\frac{3n-1}{4}, & \text{if } n \equiv 3 \pmod{4}.
\end{cases}
\]

If \( n \) is even, say \( n = 2m \), then we have the following assertions:

- \( d^2 \equiv 2d \pmod{2^n - 1} \);

- if \( \beta = \alpha^d \), where \( \alpha \) is a primitive element of \( \mathbb{F}_{2^n} \), then \( \beta \) is a primitive element in \( \mathbb{F}_{2^m} \);

- for any \( y \in \mathbb{F}_{2^m}, x \in \mathbb{F}_{2^n}, \text{Tr}^{m}_n(xy) = y\text{Tr}^{m}_n(x) \).

With these assertions Kasami construct a signal set consists of \( 2^m \) sequences of period \( 2^n \).
Construction 2.50. [44] (Kasami (small set))

Let $\mathbf{s}_\lambda = (s_{\lambda,0}, s_{\lambda,1}, \ldots)$ be a binary sequence whose elements are given by

$$s_{\lambda,i} = f_\lambda(\alpha^i), \ i = 0, 1, \ldots$$

where $f_\lambda(x) = \text{Tr}^1_{m_1}(\text{Tr}^m_n(x^2) + \lambda x^d), \ \lambda \in \mathbb{F}_{2^m}, \ x \in \mathbb{F}_{2^n}$. A signal set $S$ consisting of $\mathbf{s}_\lambda$ for all $\lambda \in \mathbb{F}_{2^m}$ is a Kasami (small set) signal set

$$S = \{\mathbf{s}_\lambda \mid \lambda \in \mathbb{F}_{2^m}\}.$$
CHAPTER 3

Interleaved sequences

In this chapter we present the interleaved construction, which was first proposed by Gong in 1995. The interleaved sequence is defined associated with a polynomial \( f \) and an integer \( m \). Then we focus on a special case of interleaved sequences, which is called primitive-interleaved sequence (PI-sequence) corresponding to a primitive polynomial \( f \). The autocorrelation has been studied for PI-sequences.

The primitive-interleaved sequences include a large number of popular sequences, such as multiplexed sequences, clock-controlled sequences, Kasami sequences, GMW sequences, geometric sequences, and No sequences.

Section 3.1 gives the original definition of interleaved sequences that appeared in [29]. Then we study the primitive-interleaved sequence (PI-sequence), and give their autocorrelation. Section 3.2 generalizes the primitive-interleaved sequences to irreducible-interleaved sequences (IRI-sequences) when considering \( f \) as an irreducible polynomial. The cross correlation of two IRI-sequences is given. The autocorrelation of an IRI-sequence can be obtained when the two IRI-sequences are identical. After introducing interleaved sequences, Section 3.3 gives the sequence sets, also called signal sets or families, consisting of a number of sequences of the same period as and generated by the interleaved algorithm. In Section 3.4 we give a new interleaved algorithm and show the properties of a family of interleaved sequences.

3.1. \((f(x), m)\)-interleaved sequences

An \((f(x); m)\)-interleaved sequence over a finite field \( \mathbb{F}_\ell \) is a sequence where all its \( m \)-decimated sequences have \( f(x) \) as the common characteristic polynomial [29]. In the following we review the main results in [29].

**Definition 3.1.** [29] Let \( f(x) \) be a polynomial over \( \mathbb{F}_\ell \) of degree \( n \) with \( f(0) \neq 0 \). Let \( \underline{u} = \{u(k)\} \) be a periodic sequence over \( \mathbb{F}_\ell \) of period \( T \), and let \( 0 < m < T \)
be a positive integer. We write \( k = im + j \) for \( j = 0, \ldots, m - 1 \). For each \( j \), if the \( m \)-decimated sequence \( u_j = u(im + j) \) is generated by \( f(x) \), then \( u \) is called an interleaved sequence over \( \mathbb{F}_\ell \) associated with \( f(x) \) and of size \( m \). The sequences \( u_j = u(im + j) \) for \( j = 0, \ldots, m - 1 \) are called component sequences of \( u \).

Usually we only consider \( f \) to be irreducible or, particularly, primitive. In this way we can find roots of the polynomial to express the elements of the sequence.

Let \( f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 \) be an irreducible polynomial over \( \mathbb{F}_\ell \) of degree \( n \), and let \( \alpha \) be a root of \( f \) we have the following lemma.

**Lemma 3.2.** Let \( \underline{a} = a(i) \) be a sequence whose elements are given by
\[
a(i) = Tr(\beta \alpha^i), \ i \geq 0, \beta \in \mathbb{F}_\ell^n.
\]
Then \( f \) is the characteristic polynomial of \( \underline{a} \), denoted \( \underline{a} \in G(f) \), where \( G(f) \) is the set collecting all the sequences generated by the same characteristic polynomial \( f \).

**Proof.** If \( \beta = 0 \), then \( \underline{a} \) is a zero sequence, and so \( \underline{a} \in G(f) \). If \( \beta \neq 0 \), then
\[
a(k + n) + c_{n-1}a(k + n - 1) + \cdots + c_1a(k + 1) + c_0a(k)
= \text{Tr}(\beta \alpha^{k+n}) + c_{n-1}\text{Tr}(\beta \alpha^{k+n-1}) + \cdots + c_1\text{Tr}(\beta \alpha^{k+1}) + c_0\text{Tr}(\beta \alpha^k)
= \text{Tr}(\beta \alpha^k(\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0))
= \text{Tr}(\beta \alpha^k0) = 0,
\]
for all \( k \geq 0 \). This implies \( f(L)\underline{a} = 0 \) which means \( \underline{a} \in G(f) \).

In other words, if we write \( u \) as a matrix form, that is
\[
u = [(L^0(u))^{(m)}, \ldots, (L^{m-1}(u))^{(m)}],
\]
each column \( u_j \) is an \( m \)-decimation of the sequence \( L^j(u) \), i.e.
\[
u_j = (L^j(u))^{(m)}.
\]

**Proposition 3.3.** [29] Let \( u \) be an interleaved sequence over \( \mathbb{F}_\ell \) associated with \( f(x) \) of degree \( n \) and size \( m \), and \( h(x) \) be the minimal polynomial of \( u \) over \( \mathbb{F}_\ell \). Then

1. \( h(x)|f(x^m) \) so that \( LC(u) \leq nm \);
2. \( \text{per}(u) | m \cdot \text{per}(f) \).
Proof. By Definition 3.1 $u_j$ is generated by $f(x)$ and so we have

\[(12)\quad f(L)(u_j) = 0,\]

and $u_j$ is an $m$-decimated sequence of $u$, we obtain

\[(13)\quad f(L^m)(u) = 0.\]

Therefore $f(x^m)$ is a characteristic polynomial of $u$, and so the minimal polynomial $h(x)$ divides $f(x^m)$.

By Proposition 2.21 we have

$$\text{per}(u) = \text{per}(h(x)).$$

The result follows from

$$\text{per}(h(x))\mid \text{per}(f(x^m)) \quad \text{and} \quad \text{per}(f(x^m)) \mid m \cdot \text{per}(f(x)).$$

\[\square\]

Let $G(f, m) = \{ u | u(im + j) = \text{Tr}_{1}^n(\gamma_j \alpha^i) \}$ represent the set of all interleaved sequences over $\mathbb{F}_\ell$ associated with $f(x)$ and $m$, and let $G(f(x))$ denote the set of sequences with $f(x)$ as their characteristic polynomial, i.e. if $f(L)(s) = 0$, then $s$ is a sequence in $G(f(x))$. We can show

$$G(f, m) = G(f(x^m)).$$

Since Proposition 3.3 implies that the minimal polynomial of a sequence $u \in G(f, m)$ divides $f(x^m)$, it follows $G(f, m) \subseteq G(f(x^m))$. For a sequence $s \in G(f(x^m))$ it satisfies $f(L^m)(s)$. We arrange $s$ in an array which consists of $m$ columns. Since $L^m$ works as taking the $m$-decimal sequence of $s$, each column $L^m(s)$ of the array satisfies $f(L^m(s)) = 0$. Therefore, $s \in G(f, m)$ and the result follows.

Let $u \in G(f, m)$. When $f(x)$ is a special polynomial, like a primitive polynomial which generates all nonzero elements in the extension field of $\mathbb{F}_\ell$ or an irreducible polynomial, we introduce two new sequences which are special cases of $(f(x), m)$-interleaved sequences.
Definition 3.4. \cite{29} If \( f(x) \) is an irreducible polynomial over \( \mathbb{F}_\ell \) of degree \( n \), then \( f(x) \) is the minimal polynomial over \( \mathbb{F}_\ell \) of the component sequences \( u_j \) (\( 0 \leq j < m \)) of \( u \). We call \( u \) an irreducible interleaved sequence (IRI-sequence). The IRI sequence \( u \) has trace representation
\[
u(i m + j) = \text{Tr}_1^n(\gamma_j \alpha^i),
\]
where \( \gamma_j \in \mathbb{F}_{\ell^n} \) and \( \alpha \) is a root of \( f(x) \) in \( \mathbb{F}_{\ell^n} \).

Furthermore, if \( f(x) \) is primitive over \( \mathbb{F}_\ell \), then every component sequence \( u_j \) is a maximal length sequence over \( \mathbb{F}_\ell \) generated by \( f(x) \). In this case, we call \( u \) a primitive interleaved sequence (PI-sequence). A PI sequence \( u \) has the same trace representation with \( \alpha \) a primitive element of \( \mathbb{F}_{\ell^n} \).

A large number of popular sequences are included in the class of PI-sequences. The multiplexed sequences and the clock-controlled sequences introduced in \cite{2, 25, 40, 41} are PI-sequences over \( \mathbb{F}_2 \) with size \( m = 2^n - 1 \). Maximal length sequences over \( \mathbb{F}_\ell \) \cite{59, 96} and GMW-sequences \cite{74} are also PI-sequences.

Because the root \( \alpha \) of \( f(x) \) is primitive in PI-sequences, the \( \gamma_j \) in the trace representation of \( \nu(i m + j) = \text{Tr}_1^n(\gamma_j \alpha^i) \) can also be expressed by \( \alpha^{e_j} \).

The shift sequence of the PI-sequence was described in \cite{21}. It is defined as
\[
e_j = \begin{cases} 
\infty, & \text{for } \gamma_j = 0, \\
k, & \text{for } \gamma_j = \alpha^k,
\end{cases}
\]
and
\[
e_{im+j} = e_j + i,
\]
with \( 0 \leq i < \ell^n - 1 \) and \( 0 \leq j < m \). The sequence \( \xi = (e_i)_{0 \leq i < (\ell^n - 1)m} \) is the shift sequence of the interleaved sequence \( u \). Since \( \gamma_j = \alpha^{e_j} \) can be expressed by a power of \( \alpha \), the trace representation of \( u \) can be written as
\[
u(i m + j) = \text{Tr}_1^n(\alpha^{i+e_j}).
\]

Now we look at the autocorrelation of \((f(x), m)\) PI-sequences, when \( f(x) \) is a primitive polynomial. The canonical character of \( \mathbb{F}_{\ell^n} \) is defined by
\[
\chi(x) = \exp\left(2\pi \sqrt{-1} \text{Tr}_1^n(x)/\ell \right), \; x \in \mathbb{F}_{\ell^n}.
\]
In particular, we have
\[
\chi(x)\chi(y) = \exp \left(2\pi \sqrt{-1} \text{Tr}_1^n(x) / \ell\right) \exp \left(2\pi \sqrt{-1} \text{Tr}_1^n(y) / \ell\right) = \exp \left(2\pi \sqrt{-1} \text{Tr}_1^n(x - \text{Tr}_1^n(y)) / \ell\right) = \chi(x - y).
\]

Let \( u \in G(f(x), m) \) be a PI-sequence, where \( f \) is primitive over \( \mathbb{F}_\ell \). Then the periodic autocorrelation function of \( u \) is defined by
\[
C_u(\tau) = \sum_{k=0}^{T-1} \chi(u(k))\chi(u(k + \tau)), \quad T = m(\ell^n - 1).
\]

By the property of complete character sum [57, p. 192], we have, for \( \zeta \in \mathbb{F}_{\ell^n} \),
\[
\sum_{x \in \mathbb{F}_{\ell^n}} \chi(\text{Tr}_1^n(\zeta x)) = \begin{cases} 
0, & \text{for } \zeta \neq 0, \\
\ell^n, & \text{for } \zeta = 0.
\end{cases}
\]

We adapt to the notations in [29]. Let \( \varepsilon = (e_0, \ldots, e_{T-1}) \) be the shift sequence of \( u \) with \( 0 \leq j < m \). For \( 0 \leq s < m \), define
\[
S_0(s) = \{ e_j - e_{j+s} \mid 0 \leq j < m-s \}, \\
S_1(s) = \{ e_j - e_{j+s} \mid m-s \leq j < m \},
\]
where \( e_j \) may be \( \infty \), and since \( e_j \) can take \( \infty \) we denote the number of \( j \)'s such that \( e_j - e_{j+s} = \infty \) as
\[
N_\infty(s) = | \{ j \mid e_j - e_{j+s} = \infty, \ 0 \leq j < m \} |.
\]

The autocorrelation of the PI-sequence \( u \) is determined by the next theorem.

**Theorem 3.5. [29]** The autocorrelation function \( C_u(\tau) \) of the PI-sequence \( u \) is at most an \( (m+1) \)-valued function and satisfies
\[
C_u(\tau) = \begin{cases} 
m(\ell^n - 1), & \tau \equiv 0 \pmod{\ell^n - 1}, \\
-m + (N + N_\infty(s))\ell^n, & \tau \equiv rm + s \pmod{m(\ell^n - 1)};
\end{cases}
\]
where \( 0 \leq s < m \), \( r \) occurs \( N \) times in \( S_0(s) \cup S_1(s) \).

**Proof.** First, we change the 1-dimensional index \( 0 \leq k < T \), where \( T = m(\ell^n - 1) \) is the period, to the 2-dimensional index \( k = im + j \) and \( \tau = rm + s \) with \( 0 \leq j, s < m \). By
\[
u(im + j) = \text{Tr}_1^n(\alpha^{i+j}), \quad j = 0, \ldots, m-1,
\]
we express the correlation function by a character sum and change it to a double sum using the 2-dimensional index. By adding a term we get a complete sum over \( \mathbb{F}_{\ell^n} \). We show the details next:

\[
C_u(\tau) = \sum_{k=0}^{T-1} \chi(u(k))\overline{\chi(u(k+\tau))}
= \sum_{j=0}^{m-1} \sum_{i=0}^{\ell^n-2} \chi \left( \text{Tr}_{n}^1(\alpha^{i+j}) - \text{Tr}_{n}^1(\alpha^{i+r+j+s}) \right)
= -m + \sum_{j=0}^{m-1} \sum_{x \in \mathbb{F}_{\ell^n}} \chi \left( \text{Tr}_{n}^1(x(\alpha^{e_j} - \alpha^{r+j+s})) \right).
\]

(14)

Here \( S_0(s) \) and \( S_1(s) \) are collecting the values \( e_{j+s} - e_j \) takes. Since \( \alpha \) is a primitive element of \( \mathbb{F}_{\ell^n} \), \( \alpha^{e_j} - \alpha^{r+j+s} \not\equiv 0 \pmod{\ell^n-1} \) for \( s \neq 0 \) and \( 0 \leq j \leq m-1 \). We have two cases: the trivial case \( s = 0 \), that is, \( e_{j+s} - e_j = 0 \), yields \( S_0(0) \cup S_1(0) = \{0\} \). We only need to consider the nontrivial case.

For the case of \( s \neq 0 \), we consider

\[
\alpha^{e_j} - \alpha^{r+j+s} = 0
\]

which is equivalent to

\[
r \equiv e_j - e_{j+s} \pmod{\ell^n-1}, \text{ for } j + s < m,
\]

and

\[
r \equiv e_j - e_{j+s-m} - 1 \pmod{\ell^n-1}, \text{ for } j + s \geq m.
\]

Let \( N \) be the number of \( r \)'s occurring in \( S_0(s) \cup S_1(s) \). For each \( r \) the summand \( \chi(\text{Tr}_{n}^1(x(\alpha^{e_j} - \alpha^{r+j+s}))) \) reaches 1 once. Thus the double sum in (14) equals \((N + N_{\infty}(s))\ell^n\) and the result follows.

\[
\square
\]

3.2. IRI-sequences

This section is a further study of IRI-sequences. We give the cross correlation of two IRI-sequences, and thus autocorrelation of an IRI-sequence when setting the two sequences to be identical.
A different definition of IRI-sequences is given in [42]. With a fixed root $\alpha$ of $f(x)$, the $(f(x); m)$-interleaved sequence $u$, defined by

$$u(t) = u(im + j) = \text{Tr}_1^m(\gamma_j \alpha^i),$$

can be identified with the $m$-tuple $\gamma = (\gamma_0, \ldots, \gamma_{m-1})$, $\gamma_j \in \mathbb{F}_\ell^n$ for $0 \leq j < m$. Hence, the $m$-tuple $\gamma$ is uniquely determined by the sequence $u$.

**Definition 3.6.** [42] A sequence $u = u(t)_{t \geq 0}$ over $\mathbb{F}_\ell$ is called an $(f(x); m)$-interleaved sequence if there exists an $m$-tuple $\gamma$ defined by

$$\gamma = (\gamma_0, \ldots, \gamma_{m-1}), \quad \gamma_j \in \mathbb{F}_\ell^n \text{ for } 0 \leq j < m,$$

such that

$$u(t) = \text{Tr}_1^m(\gamma_j \alpha^i)$$

if $t = im + j$ and $0 \leq j \leq m - 1$ and $i \geq 0$ holds, where $\alpha$ is a root of $f(x)$. We can also write the sequence as $u_\gamma$.

The autocorrelation of PI-sequences has been studied in the previous section. Then a question arises: can the cross correlation and autocorrelation values of $u$ be expressed explicitly in terms of the control parameter, $m$-tuple $\gamma = (\gamma_0, \ldots, \gamma_{m-1})$, when $f(x)$ is irreducible?

A correlation function between two sequences $u_\gamma$ and $u_\beta$ for $\gamma \neq \beta$ is defined as

$$C_{u_\gamma, u_\beta}(\tau) = \sum_{k=0}^{\ell^n - 1} \chi(u_\gamma(k) - u_\beta(k + \tau)),$$

where $\chi(x)$ is the canonical character defined by

$$\chi(x) = \exp\left(2\pi \sqrt{-1} \text{Tr}_p^\ell(x)/p\right).$$

We denote $C_{u_\gamma, u_\beta}(\tau)$ shortly by $C_{\gamma, \beta}(\tau)$.

It is clear that

$$C_{\gamma, \beta}(\tau) = C_{\gamma, \beta}(\tau') \text{ if } \tau \equiv \tau' \pmod{m(\ell^n - 1)}$$

and

$$C_{\gamma, \beta}(\tau) = C_{\beta, \gamma}(-\tau').$$

The two interleaved sequences $u_\gamma$ and $u_\beta$ are identical if $\gamma = \beta$ and so $C_\gamma(\tau)$ is the autocorrelation.
We extend the original $m$–tuple $(\gamma_0, \ldots, \gamma_{m-1})$ to a $t$–tuple using

\begin{equation}
\gamma_t = \gamma_j \alpha^i, \text{ for } t = im + j, \ 0 \leq j < m.
\end{equation}

We note that when $0 \leq t < m$, i.e. $i = 0$, $\gamma_t$ is consistent to the original $\gamma_j$. The $t$th element $u(t)$ of the IRI-sequence $\mathcal{u}_\gamma$ is determined simply by the element $\gamma_t$. In fact, we have

$$u(t) = \text{Tr}_1^n(\gamma_t) = \sum_{k=0}^{n-1} \gamma_t^k.$$  

Denote

$$\varphi_\gamma = (\gamma_0, \ldots, \gamma_T), \quad T = m(t^n - 1)$$

and

$$\varphi^{\ell_k}_\gamma = (\gamma_0^\ell_k, \ldots, \gamma_T^\ell_k), \ 0 \leq k < n.$$  

We observe that $\mathcal{u}_\gamma$ is a sequence over $\mathbb{F}_\ell$ of period $T$ and $\varphi_\gamma$ a sequence over $\mathbb{F}_{\ell^n}$. It is clear that in one period $\mathcal{u}_\gamma$ corresponds to a $T$-tuple

$$\mathcal{u}_\gamma = \sum_{k=0}^{n-1} \varphi^{\ell_k}_\gamma.$$  

In discussing the correlation values we need the following operations on the $m$–tuples:

- $\rho(\gamma_0, \ldots, \gamma_{m-1}) = (\rho \gamma_0, \ldots, \rho \gamma_{m-1})$, for $\rho \in \mathbb{F}_{\ell^n}$;

- $(\gamma_0, \ldots, \gamma_{m-1}) + (\beta_0, \ldots, \beta_{m-1}) = (\gamma_0 + \beta_0, \ldots, \gamma_{m-1} + \beta_{m-1})$;

- $L^\tau(\gamma_0, \ldots, \gamma_{m-1}) = (\gamma_\tau, \ldots, \gamma_{m-1+\tau})$, where $\gamma_{i+\tau}$ is defined as in (16).

We denote $w_H(\gamma)$ the Hamming weight of $\gamma$ that is the number of nonzero elements in $\gamma$. Then we study some properties about the operators mentioned above.

Let $\gamma$ and $\beta$ be two $m$–tuples over $\mathbb{F}_{\ell^n}$. We have

- $\gamma_{m+\tau} = \alpha \gamma_\tau$ and $L^\tau(\varphi_\gamma) = \varphi_{L^\tau(\gamma)}$ for any $\tau$. This property can be easily obtained from Definition 3.6.

- $L^{dm+\tau}(\gamma) = \alpha^d(\gamma_\tau, \ldots, \gamma_{m-1+\tau})$, as can be seen from Definition 3.6.
• $w_H(L^\tau(\gamma)) = w_H(\gamma - L^d\gamma) = w_H(\gamma)$ for any $\tau$ and $0 < d < \text{ord}(\alpha)$.

One can easily get $w_H(L^\tau(\gamma)) = w_H(\gamma)$. So we only need to show the second identity. Since $\gamma - L^d\gamma = (1 - \alpha^d)\gamma$ and $0 < d < \text{ord}(\alpha)$, we have $1 - \alpha^d \neq 0$ and so the second identity holds.

\[ L^\tau(u_\gamma) = u_{L^\tau(\gamma)}. \]

The correlation can be derived from the properties above.

**Theorem 3.7.** [42] Let $\text{per}(f(x)) = \ell^n - 1$. For any $\tau$ we have

\[ C_{\gamma,\beta}(\tau) = -m + (m - w_H(\gamma - L^\tau(\beta)))\ell^n. \]

**Proof.** Let $\rho = (\rho_0, \ldots, \rho_{m-1}) = \gamma - L^\tau(\beta)$, that is $\rho_i = \gamma_i - \beta_i + \tau$ for $0 \leq i < m$. We have

\[
C_{\gamma,\beta}(\tau) = \sum_{t=0}^{T-1} \chi(u_\gamma(t) - u_\beta(t + \tau)) \\
= \sum_{t=0}^{T-1} \chi(u_\gamma - L^\tau(u_\beta))(t) \\
= \sum_{t=0}^{T-1} \chi(u_\gamma - L^\tau(\beta))(t) \text{ by (17)} \\
= \sum_{j=0}^{m-1} \sum_{i=0}^{\ell^n-2} \chi(\text{Tr}_{1}^{n}(\rho_j \alpha^i)) \\
= -\sum_{\rho_j \neq 0} 1 + \sum_{\rho_j = 0} (\ell^n - 1) \\
= -w_H(\rho) + (m - w_H(\rho))(\ell^n - 1) \\
= -m + (m - w_H(\gamma - L^\tau(\beta)))\ell^n.
\]

\[ \square \]

### 3.3. Signal sets (families) generated by $(f(x), m)$–sequences

In this section we study the signal sets (families) constructed by the interleaved algorithm due to Gong [31]. The algorithm generates an interleaved sequence introduced in Section 3.1, then add a sequence, say $b$, onto the interleaved sequence
to obtain more sequences in the signal set. Using the algorithm we introduce two
signal sets with low correlation proposed by Gong [31].

Interleaved Algorithm

- Choose
  \[ \mathbf{a} = (a_0, \ldots, a_{v-1}) \]
  and
  \[ \mathbf{b} = (b_0, \ldots, b_{v-1}) , \]
  two sequences over \( \mathbb{F}_p \) of period \( v \) with two-level autocorrelation.

- Choose \( \mathbf{e} = (e_0, \ldots, e_{v-1}) \), an integer sequence whose elements are taken from \( \mathbb{Z}_v \).

- Construct \( \mathbf{u} = (u_0, \ldots, u_{v^2-1}) \), a \((v,v)\) interleaved sequence whose \( j \)th column sequence is given by \( L_j^e(\mathbf{a}) \), \( j = 0, 1, \ldots, v-1 \).

- Set \( \mathbf{g}_j = (s_{j,0}, \ldots, s_{j,v^2-1}) \), \( 0 \leq j < v \) whose element are defined by
  \[ s_{j,i} = u_i + b_{j+i}, \quad 0 \leq i, j < v \]
  or equivalently,
  \[ \mathbf{g}_j = \mathbf{u} + L^j(\mathbf{b}). \]

- A signal set \( S = S(\mathbf{a}, \mathbf{b}, \mathbf{e}) \) is defined as
  \[ S = \{ \mathbf{g}_j : j = 0, \ldots, v-1 \}. \]

In particular, two signal sets with good parameters were proposed. Gong [29] uses two maximal length sequences of the same period \( v \) to construct a family of long sequences from interleaved structures with the desired properties. Then she generalizes this idea to use several types of two-level autocorrelation sequences of period \( v \) [31]. Also in [31] the criterion for choosing the shift sequence \( \mathbf{e} = (e_0, \ldots, e_{v-1}) \) to get maximum correlation value \( 2v + 3 \) is given like \(|\{e_j - e_{j+s} \mid 0 \leq j < v - s\}| = v - s \), for all \( 1 \leq s < v \) (Theorem 2 in [31]).

- Construction A. \((p^n - 1)^2, p^n - 1, 2p^n + 1\) signal set.
  1. Choose \( \mathbf{a} \) and \( \mathbf{b} \), two sequences over \( \mathbb{F}_p \) of period \( p^n - 1 \) with two-level autocorrelation.
  2. Choose \( \alpha \) and \( \beta \), primitive elements of \( \mathbb{F}_{p^{n^2}} \) and \( \mathbb{F}_{p^n} \), respectively.
(3) Compute \( e_j \)'s satisfying
\[
\beta^e_j = \text{Tr}_{2^n}(\eta \alpha^{j+1}), \quad 0 \leq j < p^n - 1
\]
where \( \eta \in \mathbb{F}_p \) with \( \text{Tr}_{2^n}(\eta) = 0 \).

- Construction B. \((v^2, v, 2v + 3)\) binary signal set where \( v \) is prime.

1. Choose \( a \) and \( b \), two two-level autocorrelation sequences over \( \mathbb{F}_2 \) with prime period \( v \).
2. Choose \( \alpha \), a primitive element of \( \mathbb{F}_v \).
3. Set \( e_j = \alpha^j, \quad 0 \leq j < v \).

### 3.4. A new variant of the interleaved algorithm

In this section we give a new result which extends Gong’s interleaved algorithm to the case of two base sequences of different periods. We also give the number of zeros of every sequence in the constructed family in one period. The results can also be found in [35].

**Definition 3.8.** [29, 31] Fix two positive integers \( s \) and \( t \) where both \( s \) and \( t \) are not equal to 1. Given an \( m \)-ary sequence \( a = (a_0, \ldots, a_{s-1}) \) of period \( s \) \((a) \) is called the base sequence and a sequence \( e = (e_0, \ldots, e_{t-1}) \), for each \( 0 \leq i \leq t - 2 \) such that \( e_i \in \mathbb{Z}_s \) and \( e_{t-1} = \infty \) \((\infty) \) is called the shift sequence), let \( u = (u_0, \ldots, u_{st-1}) \) be an \( m \)-ary sequence of period \( st \). We arrange the elements of the sequence \( u \) into an \( s \times t \) matrix as follows:

\[
A_u = \begin{bmatrix}
  u_0 & \cdots & u_{t-1} \\
  \vdots & \ddots & \vdots \\
  u_{(s-1)} & \cdots & u_{(s-1)t+t-1}
\end{bmatrix}
\]
satisfying that each column of \( A_u \) is a shift of \( a \). Let \( A_j \) be the \( j \)th column. Then \( A = [A_0, \ldots, A_{t-1}] \) and \( A_j = L^e_j(a) \) and \( L^\infty(a) = (0, \ldots, 0) \). The matrix \( A_u \) is called the matrix form of sequence \( u \), and \( u \) is called an interleaved sequence from the base sequence \( a \) and the shift sequence \( e \).

Given a base sequence \( a \) and a shift sequence \( e \), an interleaved sequence \( u \) is uniquely determined. So we also say \( u \) is an \((s, t)\)-interleaved sequence associated with \((a, e)\). Moreover, using another sequence \( b = (b_0, \ldots, b_{s-1}) \) of the same period
s, Gong [31] constructed a family of interleaved \((s,s)\)-sequences with the desired properties. Here we consider the case where \(s\) is not necessarily equal to \(t\).

**Algorithm 3.9.** Let \(s\) and \(t\) be two positive integers. Suppose that \(a = (a_0, \ldots, a_{s-1})\) and \(b = (b_0, \ldots, b_{t-1})\) are two \(m\)-ary sequences of periods \(s\) and \(t\), respectively.

1. Choose \(e = (e_0, \ldots, e_{t-1})\) as the shift sequence for which the first \(t - 1\) elements are over \(\mathbb{Z}_s\) and \(e_{t-1} = \infty\). Moreover, if we let \(d_{i-1} = e_i - e_{i-1}\), then we choose \(e\) such that \(d_0, d_1, \ldots, d_{t-3}\) are in an arithmetic progression with common distance \(d \neq 0\).
2. Construct an interleaved sequence \(u = (u_0, \ldots, u_{st-1})\), whose \(j\)th column in the matrix form is given by \(L^e_j(a)\).
3. For \(0 \leq i < st - 1, 0 \leq j \leq t\), define \(s_j = (s_{j,0}, \ldots, s_{j,st-1})\) as follows:
   \[
   s_{j,i} = \begin{cases} 
   u_i + b_{j+i}, & 0 \leq j \leq t - 1, \\
   u_i, & j = t.
\end{cases}
   \]
4. Define the family of sequences \(S = \mathcal{S}(a, b, e)\) as \(S = \{s_j| j = 0,1,\ldots,t\}\), where \(a\) is the first base sequence, \(e\) is the shift sequence, and \(b\) is the second base sequence.

**Remark 3.10.** The shift sequence \(e\) can be expressed by a polynomial in \(j\) of degree 2
\[
e_j = e_0 + j d_0 + \frac{j(j - 1)}{2} d, \quad 0 \leq j \leq t - 2.
\]

Because it is desirable to have more sequences in a family, without loss of generality, we can assume that \(s \leq t\). We also concentrate on binary sequences \((m = 2)\) in the next section. In particular, we are interested in Legendre sequences of prime periods since they are a nice class of sequences with good randomness properties.

**Remark 3.11.** In Algorithm 3.9 we can use two sequences with different periods \(v\) and \(w\) and restrict the shift sequence \(e\) to satisfy that \(e_1 - e_0, e_2 - e_1, \ldots, e_{w-2} - e_{w-3}\) are in an arithmetic progression.
3.4.1. Balance Property. The sequences in good signal sets should be hard to distinguish from a random sequence. The balance property provides information on the number of each element in one period. We need the distinct elements to be close to an uniform distribution. In this section we give the number of zeros in a period of every interleaved sequence constructed by the extended interleaved algorithm given in Algorithm 3.9. For a binary sequence, we want the number of ones and the number of zeros to be close.

For any sequence \( s \), let \( N_0(s) \) denote the number of zeros of the sequence \( s \).

**Theorem 3.12.** Let us choose \( a \) as the first base sequence with period \( v \) and \( b \) as the second base sequence with period \( w \). Then using Algorithm 3.9 we construct a family \( \mathcal{S}(a, b, e) = \{ s_j \mid j = 0, 1, \ldots, w \} \) with the property that the number \( N_0(s_j) \) of zeros in one period of each sequence \( s_j \) is:

\[
\begin{align*}
N_0(a) \cdot (N_0(b) - 1) + (v - N_0(a)) \cdot (w - N_0(b) - 1), & \quad b_{j+w-1} = 0, j \leq w - 1; \\
N_0(a) \cdot N_0(b) + (v - N_0(a)) \cdot (w - N_0(b) - 1), & \quad b_{j+w-1} = 1, j \leq w - 1. 
\end{align*}
\]

**Proof. Case 1.** \( j = w \).

In this case \( s_w = u \). Then we arrange the elements of \( u \) into the \( v \times w \) matrix

\[
A_u = \begin{bmatrix} L^{e_0}(a), & \ldots, & L^{e_{w-2}}(a), & 0 \end{bmatrix}.
\]

Since each of the first \( w - 1 \) columns is just a shift of \( a \) and \( a \) contains exactly \( N_0(a) \) zeros, then the number of zeros in \( u \) is

\[
N_0(u) = (w - 1) \cdot N_0(a) + v.
\]

**Case 2.** \( 0 \leq j \leq w - 1 \).

We denote

\[
B = \begin{bmatrix} b_j & b_{j+1} & \cdots & b_{j+w-1} \\
& b_j & b_{j+1} & \cdots & b_{j+w-1} \\
& & & & \\
b_j & b_{j+1} & \cdots & b_{j+w-1} 
\end{bmatrix}.
\]

Then, we have

\[
s_j = u + L^j(b) = A_u + B.
\]
If $b_{j+w-1} = 0$, then there are $N_0(b) - 1$ zeros in $b_j, \ldots, b_{j+w-2}$, and $w - N_0(b)$ ones in $b_j, \ldots, b_{j+w-2}$. So

$$N_0(s_j) = N_0(a) \cdot (N_0(b) - 1) + (v - N_0(a)) \cdot (w - N_0(b)) + v.$$ 

If $b_{j+w-1} = 1$, then there are $N_0(b)$ zeros in $b_j, \ldots, b_{j+w-1}$, and $w - N_0(b) - 1$ ones in $b_j, \ldots, b_{j+w-1}$. So

$$N_0(s_j) = N_0(a) \cdot N_0(b) + (v - N_0(a)) \cdot (w - N_0(b) - 1).$$

□

We observe that the number of zeros in each sequence from the constructed family is independent of the chosen shift sequence $e$. The purpose for adding strong conditions on $e$ is to get a desired correlation property. Wang and Qi [88] give the balance property of the interleaved construction with two Legendre sequences of twin prime periods $p$ and $p + 2$, respectively. However, there are some typos in their result. In the next corollary that follows immediately from the previous theorem, we give the balance property to the interleaved construction with two Legendre sequences of any two prime periods $p$ and $q$ which also corrects Wang and Qi’s result. We remark that a Legendre sequence of length $p$ is balanced and so it contains exactly $(p-1)/2$ zeros.

**Corollary 3.13.** Let $p$ and $q$ be two prime numbers and $\mathcal{S}(a, b, e) = \{s_j \mid j = 0, 1, \ldots, q\}$ be the family of interleaved sequences constructed by Algorithm 3.9, where the base sequences $a$ and $b$ are Legendre sequences of period $p$ and $q$ respectively. Then the number of zeros $N_0(s_j)$ in one period of $s_j$ $(0 \leq j \leq q)$ is

$$N_0(s_j) = \begin{cases} 
\frac{p(q+1)-q+1}{2}, & j = q; \\
\frac{p(q+1)+2}{2}, & b_{j+q-1} = 0, j \leq q - 1; \\
\frac{p(q-1)}{2}, & b_{j+q-1} = 1, j \leq q - 1.
\end{cases}$$
CHAPTER 4

Correlation properties and linear complexity of interleaved binary sequences

In this chapter we study correlation properties as well as linear complexity of a family of binary interleaved sequences given by Algorithm 3.9 when $m = 2$ in the algorithm. In Section 4.1 we overview some bounds on correlation values of several well-known signal sets. The cross correlation values of binary interleaved sequences coming from Algorithm 3.9 are shown in Section 4.2.

To find the linear complexity of a sequence, the basic idea is to find the minimal polynomial of the sequence whose degree equals the linear complexity. In Section 4.3 we use this method to obtain results on the linear complexity of the binary interleaved sequences associated with two sequences of integer periods $s$ and $t$, respectively.

However, this method to estimate a lower bound for the linear complexity only works for the interleaved sequence $u$ in Algorithm 3.9. Therefore, we will use a completely different method to estimate the lower bound of the linear complexity of sequences in the next chapter.

4.1. Correlation property

The volume of communication traffic over a variety of media has been steadily increasing over the past few decades, and will continue to do so. This increase has led to a need for methods that allow many users to share communication channels. It is essential to find large families of easily generated binary sequences with high linear complexities and low correlation function values. The smaller the pairwise cross correlation and the larger the family, the higher the capacity of the system. Also, the higher the linear span (complexity), the harder it is for an adversary to jam or intercept messages. These issues will be discussed in next chapter; here we focus on correlation.
Consider a set of \( M \) sequences

\[
\mathcal{S} = \{s_i : 0 \leq i \leq M - 1\},
\]

where each \( s_i \) is a sequence of period \( T \). We say \( \mathcal{S} \) is a sequence family (or a signal set) of period \( T \). The size \( M \), the maximum autocorrelation

\[
\theta_A(\mathcal{S}) := \max\{|C_{s_i}| : s_i \in \mathcal{S}, \ i = 0, \ldots, M - 1, \ \tau \not\equiv 0 \pmod{T}\},
\]

and the maximum cross correlation

\[
\theta_C(\mathcal{S}) := \max\{|C_{s_i,s_j}| : s_i, s_j \in \mathcal{S}, \ i \neq j, i, j = 0, \ldots, M - 1, \ \tau \not\equiv 0 \pmod{T}\}
\]

are key parameters of \( \mathcal{S} \) when \( \mathcal{S} \) is employed in a code-division multiple access (CDMA) system [37]. A large family size is required to support a large number of simultaneous users. Small autocorrelation \( \theta_A(\mathcal{S}) \) is required to ensure message synchronization and small cross correlation \( \theta_C(\mathcal{S}) \) is required to minimize interference among different users.

Many authors do not distinguish between autocorrelation and cross correlation and define

\[
\theta(\mathcal{S}) := \max\{\theta_A(\mathcal{S}), \theta_C(\mathcal{S})\}
\]

as the maximum correlation of \( \mathcal{S} \). The most well-known bounds are those due to Welch, Sidelnikov, and Levenshtein, see [37]. Welch bound [90] shows that

\[
(19) \quad \theta(\mathcal{S}) \geq T \sqrt{\frac{M - 1}{MT - 1}}.
\]

When \( M \) and \( T \) both tend to infinity, the bound (19) asserts that \( \theta(\mathcal{S}) \) must grow at least like \( \sqrt{T} \).

Bounds on the minimum possible value of \( \theta \) for given period \( T \), family size \( M \), and alphabet size \( \ell \) are available; they can be used to judge the merits of a particular sequence design. In CDMA systems, there is greatest interest in designs in which the parameter \( \theta \) is in the range \( \sqrt{T} \leq \theta \leq 2\sqrt{T} \).

For example, the Gold family and Kasami family are constructed as below.

Let \( r \) be odd and \( d = 2^{k} + 1 \) where \( k, 1 \leq k \leq r - 1 \), is an integer satisfying \( (k, r) = 1 \). Let \( \{s(t)\} \) be a cyclic shift of a maximum length sequence of period
$T = 2^r - 1$ that satisfies $s(dt) \neq 0$ and let $\mathfrak{G}$ be the Gold family of $2^r + 1$ sequences given by

$$\mathfrak{G} = \{s(t)\} \cup \{s(dt)\} \cup \{s(t) + s(d(t + \tau)) \mid 0 \leq \tau \leq T - 1\}.$$ 

Then each sequence in $\mathfrak{G}$ has period $T = 2^r - 1$ and $\theta(\mathfrak{G})$ of $\mathfrak{G}$ satisfies

$$\theta(\mathfrak{G}) \leq \sqrt{2^{r+1} + 1}.$$ 

The Gold family remains the best-known family of maximal length sequences having low cross correlation. Applications include the Navstar Global Positioning System whose signals are based on Gold sequences [24].

The family of Kasami sequences has a similar description. Let $r = 2v$ and $d = 2^v + 1$. Let $\{s(t)\}$ be a cyclic shift of a maximum length sequence of period $T = 2^r - 1$ that satisfies $s(dt) \neq 0$, and consider the family of Kasami sequences given by

$$\mathfrak{K} = \{s(t)\} \cup \{s(t) + s(d(t + \tau)) \mid 0 \leq \tau \leq 2^v - 2\}.$$ 

Then the Kasami family $\mathfrak{K}$ contains $2^v$ sequences of period $T = 2^r - 1$. In this case $\theta = 1 + 2^v$ [44].

Unfortunately, there are a limited number of known instances of such sequences. The correlation properties of sequences generated by various modified shift registers have been studied, including GMW sequences [32], geometric sequences [49], cascaded GMW sequences [1, 50]. The results are often closely related to results from coding theory [92]. In a few cases families of sequences with good correlation properties have been found, such as Kasami sequences [44, 45], bent function sequences [51, 56, 68, 71], and No sequences [67].

### 4.2. The correlation values of binary interleaved sequences

In Section 3.4 we give a variant of Gong’s interleaved algorithm and by Algorithm 3.9 we construct a new family of interleaved sequences of period $wv$, where $w$ and $v$ are positive integers ($w$ and $v$ correspond to $s$ and $t$ in Algorithm 3.9, respectively). Also a family of interleaved sequences from two Legendre sequences of any prime periods $p$ and $q$ is defined and it is shown that its cross correlations is low. We show the correlation values of every sequence in the family in this section.
Before we provide the proof of the cross correlation values we need the following results.

**Proposition 4.1.** (Proposition 3 in [31]) Let \( a \) and \( b \) be two sequences over \( \mathbb{F}_2 \) of period \( T \). For \( \tau \geq 0 \), we have

1. \( \langle a, b \rangle = C_{a,b}(0) \);
2. \( \langle a, L^\tau(b) \rangle = C_{a,b}(\tau) \);
3. \( \langle L^i(a), L^{i+\tau}(a) \rangle = C_{L^i(a), L^i(a)}(\tau) = C_a(j-i+\tau) \) where \( i,j \geq 0 \);
4. for \( c,d \in \mathbb{F}_2 \), \( \langle a+c, b+d \rangle = (-1)^{c+d} \langle a, b \rangle \).

**Remark 4.2.** For convenience, we introduce the extended sequence defined by \( e_{j+w} = 1 + e_j \). Hence, the extended shift sequence \( e \) is

\[
(e_0, \ldots, e_{w-2}, \infty, e_w = 1 + e_0, \ldots, e_{2w-2} = 1 + e_{w-2}, \infty).
\]

For the extended one, we still use the same notation \( e \) for notational convenience.

In the following when encountering an element of \( e \) out of the range of the original shift sequence, we just use the extended shift sequence, i.e. \( e_{j+w} = 1 + e_j \).

The following proposition studies the matrix form of an interleaved sequence. It is a modification of Proposition 4 in [31] from \( v = w \) to arbitrary \( v \) and \( w \).

**Proposition 4.3.** [31] Let \( u \) be a \((v, w)-\)interleaved sequence associated with \((a, e)\). We extend the sequence \((e_0, \ldots, e_{w-2}, \infty)\) to \( e = (e_0, \ldots, e_{w-1}) \) by defining \( e_{j+w} = 1 + e_j \), for \( j = 0, \ldots, w-1 \). For \( \tau \geq 0 \), let \([T_1, \ldots, T_j, \ldots, T_{w-1}]\) be the matrix form of \( L^\tau(u) \) consisting the columns \( T_j \) for \( j = 1, \ldots, w-1 \). If we write \( \tau = rw + s \), \( 0 \leq r < v \), \( 0 \leq s < w \) then

\[
T_j = L^{s+e_j}(a).
\]

**Proof.** We use one index \( k \) for \( 0 \leq k < vw - 1 \), or two indices \((i, j)\) for \( 0 \leq i < v \) and \( 0 \leq j < w \), to show the position of an element in the matrix form of an interleaved sequence. Let

\[
A_u = \begin{bmatrix}
  u_0 & \cdots & u_{w-1} \\
  \vdots & \ddots & \vdots \\
  u_{(v-1)w} & \cdots & u_{(v-1)w+w-1} \\
\end{bmatrix} = \begin{bmatrix}
  u_{0,0} & \cdots & u_{0,w-1} \\
  \vdots & \ddots & \vdots \\
  u_{v-1,0} & \cdots & u_{v-1,w-1} \\
\end{bmatrix}
\]
be the matrix form of \( u \). Let \( A_j \) be the \( j \)th column of \( A_u \). We observe that the first entry in the sequence \( L^\tau(u) \) is \( u_{r,s} \) in \( A_u \). From the definition of the interleaved sequences, we have \( u_{r,v+j} = u_{r+1,j} \) for each \( j \) with \( v - s \leq j < w \). So \([T_0, T_1, \ldots, T_{w-1}]\) has the following matrix form

\[
\begin{bmatrix}
    u_{r,s} & \cdots & u_{r,w-1} & u_{r+1,0} & \cdots & u_{r+1,s-1} \\
    u_{r+1,s} & \cdots & u_{r+1,w-1} & u_{r+2,0} & \cdots & u_{r+2,s-1} \\
    \vdots \\
    u_{v-1,s} & \cdots & u_{v-1,w-1} & u_{0,0} & \cdots & u_{0,s-1} \\
    \vdots \\
    u_{r-1,s} & \cdots & u_{r-1,w-1} & u_{r,0} & \cdots & u_{r,s-1}
\end{bmatrix}
\]

Therefore, for \( 0 \leq j < w - s \), we recall

\[(20)\]

\[T_j = L^\tau(A_{s+j}) = L^\tau(L^{e_{s+j}}(a)) = L^{r+e_{s+j}}(a).\]

For \( w - s \leq j < w \), we have

\[(21)\]

\[T_j = L^{r+1}(A_{j-(w-s)}) = L^{r+1}(L^{e_{j-(w-s)}}(a)) = L^{r+1+e_{j-(w-s)}}(a).\]

For \( 0 \leq j \leq w - 1 \), define

\[(22)\]

\[e_{j+w} = 1 + e_j.\]

Then, the sequence \( e \) of period \( w \) can be expanded to a sequence of period \( 2w \). For simplicity we still use the symbol \( e \) for that sequence.

Applying (22) for \( w - s \leq j \leq w - 1 \) we obtain

\[1 + e_{j-(w-s)} = e_{j-(w-s)+w} = e_{j+s}.\]

Substituting it into (21), we get that \( T_j = L^{r+e_{j+s}}(a) \). Together with (20), the result follows.

From Proposition 4.3, the following modification of Lemma 1 in [31] is immediate.

**Lemma 4.4.** Let \( \mathcal{S} \) be a family of sequences constructed using Algorithm 5.9, and let \( \tau = rw + s \) with \( 0 \leq s < w \) and \( 0 \leq r < v \). Then, for \( s_k \in \mathcal{S} \setminus \{s_w\} \), the \( j \)th column sequence of \( L^\tau(s_k) \) is given by

\[L^{r+e_{s+j}}(a) + b_{k+s+j}, \quad 0 \leq j < w.\]
Moreover, the $j$th column sequence of $s_w$ is given by $L^{r+e_{s+j}}(a)$.

**Remark 4.5.** Let $s_h, s_k \in \mathcal{S}$ be two sequences in the $(v, w)$–interleaved sequence family from Algorithm 3.9. Let $[S_0, \ldots, S_{w-1}]$ and $[T_0, \ldots, T_{w-1}]$ be the matrix forms of $s_h$ and $L^\tau(s_k)$, respectively, where $\tau \geq 0$. Proposition 4.1 (2) and Proposition 4.3 imply that the cross correlation between $s_h$ and $s_k$ can be computed as

$$C_{h,k}(\tau) = \sum_{j=0}^{w-1} <S_j, T_j>.$$ 

Similarly we have the following modification of Lemma 2 in [31].

**Lemma 4.6.** For $0 \leq h, k \leq w$, suppose that $s_h$ and $s_k$ are two sequences in $\mathcal{S}$. If $\tau = rw + s$, $0 \leq s < w$, $0 \leq r < v$, then the correlation function between $s_h$ and $s_k$ is

$$C_{h,k}(\tau) = \begin{cases} 
\sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+s+j}} C_2(e_{j+s} - e_j + r), & 0 \leq h, k \leq w - 1, \\
\sum_{j=0}^{w-1} (-1)^{b_{k+s+j}} C_2(e_{j+s} - e_j + r), & h = w.
\end{cases}$$

**Proof.** Let $[S_0, \ldots, S_{w-1}]$ and $[T_0, \ldots, T_{w-1}]$ be the matrix forms of $s_h$ and $L^\tau(s_k)$, respectively. According to Remark 4.5,

$$C_{h,k}(\tau) = \sum_{j=0}^{w-1} <S_j, T_j>.$$ 

First we consider $0 \leq h, k \leq w - 1$. From Lemma 4.4, for $0 \leq j < w$, we have

$$S_j = L^{e_j}(a) + b_{k+j}, \quad T_j = L^{r+e_{s+j}}(a) + b_{k+s+j}.$$ 

Applying Proposition 4.1 (4) and then Proposition 4.1 (3), we get

$$<S_j, T_j> = <L^{e_j}(a) + b_{k+j}, L^{r+e_{s+j}}(a) + b_{k+s+j}> = (-1)^{b_{h+j} + b_{k+s+j}} <L^{e_j}(a), L^{r+e_{s+j}}(a)> = (-1)^{b_{h+j} - b_{k+s+j}} C_2(e_{j+s} - e_j + r).$$ 

The case $h = w$ or $k = w$ can be shown similarly.

We recall that $e = (e_0, \ldots, e_{w-1})$, where $e_{w-1} = \infty$. In order to study $e_{j+s} - e_j$ for $0 \leq j, s < w$, we introduce three elements: $\infty$, $\infty_1$, and $\infty_2$, and define $\infty - \infty = \infty$, $k - \infty = \infty_1$ and $\infty - k = \infty_2$ for any integer $k$. 

Lemma 4.7. Let $d_i = e_{i+1} - e_i$ be in an arithmetic progression for $i = 0, \ldots, w - 3$. For $0 \leq r < v$, $1 \leq s < w$, let $N(r, s)$ be the number of $j$ with $0 \leq j < w$ such that $e_{j+s} - e_j + r \equiv 0 \pmod{v}$, then $N(r, s) \leq \lfloor \frac{w}{v} \rfloor + 1$.

Proof. For the shift sequence $e = (e_0, \ldots, e_{w-1})$, we need $d_0 = e_1 - e_0, d_1 = e_2 - e_1, \ldots, d_{w-2} = e_{w-1} - e_{w-2}$ to be in an arithmetic progression with a constant difference $d$. So we deduce that $e_i = e_0 + id_0 + \frac{i(i-1)d}{2}$. Since $j + s = w$ gives $e_{j+s} - e_j + r = \infty_2$, we consider the following two cases:

(1) $j + s < w$. We have
\[
e_{j+s} - e_j + r = (j + s)d_0 + \frac{(j + s)(j + s - 1)d}{2} - jd_0 - \frac{j(j - 1)d}{2} + r \n\]
This is a linear equation modulo $v$. It has no solution or one solution for $j$ when $j < v$. Thus in the extended $e$, there are at most $\lfloor \frac{w}{v} \rfloor + 1$ $j$’s satisfying $e_{j+s} - e_j + r \equiv 0 \pmod{v}$, and $j < w$.

(2) $j < w < j + s$. The above expression changes into
\[
e_{j+s} - e_j + r = 1 + e_{(j+s)} \pmod{w} - e_j + r \n\]
\[
= 1 + (j + s - w)d_0 + \frac{(j + s - w)(j + s - 1)d}{2} - jd_0 - \frac{j(j - 1)d}{2} + r \n\]
\[
= 1 + (s - w)d_0 + r + \frac{(s - w)^2 - (s - w)}{2}d + (s - w)dj. \n\]
This equation has at most one solution for $j < v$, and so has at most $\lfloor \frac{w}{v} \rfloor + 1$ solutions when $j < w$. \hfill \square

We comment that a different proof of the above result is in [88]. Lemma 4.6 shows that the cross correlation of any two sequences in the constructed family is related to $C_2(e_{j+s} - e_j + r)$. Let $a$ be a balanced sequence of period $v$ with $(v - 1)/2$ zeros. Then we can denote $C_2(\infty_1) = C_2(\infty_2) = -1$ and $C_2(\infty - \infty + k) = v$. Indeed, the sequence $a$ will turn into the zero sequence after shifting it for infinitely many times. Therefore, $C_2(\infty_1) = \sum_{i=0}^{v-1}(-1)^{ai} = -1$ since it has $(v - 1)/2$ zeros in any period. The same reason leads to $C_2(\infty_2) = -1$. Similarly, we get $C_2(\infty - \infty) = \sum_{i=0}^{v-1}(-1)^{0} = v$. We comment that these notations are convenient.
in the case \( j + s = w - 1 \) or \( j = w - 1 \), that is, when one of the terms in \( e_{j+s} - e_j \)
is \( \infty \).

**Theorem 4.8.** Let \( \mathbf{a} \) be a two-level autocorrelation sequence with period \( v \) and \( \mathbf{b} \) be a balanced low cross correlation sequence of period \( w \) with the maximal absolute value of nontrivial autocorrelation equal to \( \delta_b \). The family of sequences \( \mathcal{S} \) generated by Algorithm 3.9 is a \( (vw, w + 1, \delta_1) \) signal set, where

\[
\delta_1 = \max \left\{ \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right)(v + 1) + w, \delta_b v \right\}.
\]

**Proof.** We know that the autocorrelation of \( \mathbf{a} \) is \( C_\mathbf{a}(\tau) = \{-1, v\} \).

**Case 1.** \( \tau = 0 \). It follows \( \tau = 0 \cdot w + 0 \), that is, \( r = s = 0 \). By Lemma 4.6 we have

\[
C_{h,k}(0) = \begin{cases} \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+j}} C_\mathbf{a}(0), & 0 \leq h \leq w - 1, \\ \sum_{j=0}^{w-1} (-1)^{b_{h+j}} C_\mathbf{a}(0), & h = w. \end{cases}
\]

Since \( C_\mathbf{a}(\tau) = \{-1, v\} \), we have

\[
C_\mathbf{a}(0) \cdot \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+j}} = v \cdot \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+j}} = v \cdot C_\mathbf{b}(h - k).
\]

We want to find the nontrivial correlation value. So when \( \tau = 0 \), the two sequences \( s_h, s_j \) should be different. Hence \( |C_{h,k}(0)| \leq \delta_b v \).

When one of the sequences \( s_h \) or \( s_k \) is \( \mathbf{u} \), we have

\[
|C_{h,k}(\tau)| = v \cdot \sum_{j=0}^{w-1} (-1)^{b_{j+k}} \leq v,
\]

because the sequence \( \mathbf{b} \) is balanced.

**Case 2.** \( \tau = rw + 0 \) and \( 0 < r < v \).

In this case since \( s = 0 \) and \( r \neq 0 \), we have \( e_{j+s} - e_j + r = r \) (mod \( v \)) and so \( N(r, 0) = \{0\} \). Then, for \( 0 \leq h, k < w \), we have

\[
C_{h,k}(\tau) = \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+j}} C_\mathbf{a}(r) = C_\mathbf{a}(r) \cdot C_\mathbf{b}(h - k) = (-1) \cdot C_\mathbf{b}(h - k).
\]

Hence \( |C_{h,k}(\tau)| \leq w \).

When one of the sequences \( s_h \) or \( s_k \) is \( \mathbf{u} \), we have

\[
|C_{h,k}(\tau)| = \left| \sum_{j=0}^{w-1} (-1)^{b_{h+j}} C_\mathbf{a}(r) \right| = \left| C_\mathbf{a}(r) \cdot \sum_{j=0}^{w-1} (-1)^{b_j} \right| \leq \left| C_\mathbf{a}(r) \right| = 1.
\]
Case 3. $\tau = rw + s \ (0 \leq r < v, \ 0 < s < w)$.

For $0 \leq h, k \leq w$, we have

$$C_a(e_{j+s} - e_j + r) = \begin{cases} v, & e_{j+s} - e_j + r \equiv 0 \pmod{v}, \\ -1, & e_{j+s} - e_j + r \not\equiv 0 \pmod{v}, \\ -1, & e_{j+s} - e_j + r = \infty_1 \text{ or } \infty_2. \end{cases}$$

We only need to categorize the values of $N(r, s)$ to calculate $C_{h, k}(\tau)$. We consider the following cases:

**Case 3.1:** If $N(r, s) = 0$, there is no $j$ satisfying $e_{j+s} - e_j + r \equiv 0 \pmod{v}$ and so for $0 \leq h, k < w$ and $0 \leq j \leq w - 1$, we have $C_a(e_{j+s} - e_j + r) = -1$. Thus,

$$C_{h, k}(\tau) = (-1) \cdot C_a(k + s - h).$$

The maximal absolute value of the correlation values is

$$|C_{h, k}(\tau)| \leq |C_a(k + s - h)| = w.$$

When one of the sequences $s_h$ or $s_k$ is $\infty$, the correlation value is

$$\left| C_a(e_{j+s} - e_j + r) \cdot \sum_{j=0}^{w-1} (-1)^{b_j} \right| \leq 1.$$

**Case 3.2:** If $N(r, s) = 1$, there is one $j$, say $j_0$, satisfying $e_{j_0+s} - e_{j_0} + r \equiv 0 \pmod{v}$, then, for $0 \leq h, k < w$, we have

$$C_{h, k}(\tau) = \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+s+j}} \cdot C_a(e_{j+s} - e_j + r)$$

$$= (-1)^{b_{h+j_0} - b_{k+s+j_0}} \cdot C_a(0) + \sum_{j \neq j_0} (-1)^{b_{h+j} - b_{k+s+j}} \cdot (-1)$$

$$= (-1)^{b_{h+j_0} - b_{k+s+j_0}} (v + 1) + (-1) \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+s+j}}$$

$$= (-1)^{b_{h+j_0} - b_{k+s+j_0}} (v + 1) + (-1) \cdot C_a(k + s - h)$$

$$= (-1) \cdot C_a(k + s - h) + \{\pm (v + 1)\}.$$ 

Therefore

$$|C_{h, k}(\tau)| \leq (v + 1) + w.$$
When one of the sequences $s_h$ or $s_k$ is $u$, we get

$$
C_{h,k}(\tau) = \sum_{j=0}^{w-1} (-1)^{b_{k+j}} C_\alpha(e_{j+s} - e_j + r)
$$

$$
= (-1)^{b_{k+j_0}} C_\alpha(0) + \sum_{j \neq j_0} (-1)^{b_{k+j}} \cdot (-1)
$$

$$
= (-1)^{b_{k+j_0}} (v + 1) + (-1) \sum_{j=0}^{w-1} (-1)^{b_{k+j}}.
$$

Hence,

$$
|C_{h,k}(\tau)| \leq (v + 1) + 1.
$$

**Case 3.3:** If $N(r,s) = 2$, there are two $j$'s, say $j_0$ and $j_1$, satisfying $e_{j+s} - e_j + r \equiv 0 \pmod{v}$. Then, for $0 \leq h,k < w$,

$$
C_{h,k}(\tau) = \sum_{j=0}^{w-1} (-1)^{b_{h+j} - b_{k+j}} C_\alpha(e_{j+s} - e_j + r)
$$

$$
= (-1)^{b_{h+j_0} - b_{k+j_0}} C_\alpha(0) + (-1)^{b_{h+j_1} - b_{k+j_1}} C_\alpha(0)
$$

$$
+ \sum_{j \neq j_0, j_1} (-1)^{b_{h+j} - b_{k+j}} \cdot (-1)
$$

$$
= \{ \pm(v + 1) \} + \{ \pm(v + 1) \} + (-1) \cdot C_\alpha(k + s - h).
$$

Therefore,

$$
|C_{h,k}(\tau)| \leq 2(v + 1) + w.
$$

When one of the sequences $s_h$ or $s_k$ is $u$, we obtain

$$
C_{h,k}(\tau) = \sum_{j=0}^{w-1} (-1)^{b_{k+j}} C_\alpha(e_{j+s} - e_j + r)
$$

$$
= \{ \pm(v + 1) \} + \{ \pm(v + 1) \} + (-1) \sum_{j=0}^{w-1} (-1)^{b_{k+j}}.
$$

Thus,

$$
|C_{h,k}(\tau)| \leq 2(v + 1) + 1.
$$

**Case 3.4:** As we can see from the previous sub-cases, the maximum magnitude of the correlation value in our estimations grows as $N(r,s)$ increases. Hence we only give here the estimates for the largest value of $N(r,s)$.

If $N(r,s) = \left\lfloor \frac{w}{v} \right\rfloor + 1$, then for $0 \leq h,k < w$, we have

$$
C_{h,k}(\tau) = \{ \pm(v + 1) \} + \cdots + \{ \pm(v + 1) \} + (-1) \cdot C_\alpha(k + s - h),
$$
where the number of copies of \( \{ \pm (v + 1) \} \) is \( \left\lfloor \frac{w}{v} \right\rfloor + 1 \). Thus

\[
|C_{h,k}(\tau)| \leq \left| - \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right) \cdot (v + 1) - w \right| = \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right) \cdot (v + 1) + w.
\]

If one of the sequences \( s_h \) or \( s_k \) is \( u \),

\[
|C_{h,k}(\tau)| \leq \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right) \cdot (v + 1) + 1.
\]

Recall that \( \delta_1 \) denotes the maximum magnitude among all the cross correlation values and nontrivial autocorrelation values of \( \mathcal{S} \). Hence we obtain a \((vw, w + 1, \delta_1)\) signal set, where

\[
\delta_1 = \max \left\{ \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right) (v + 1) + w, \delta_b v \right\}.
\]

\[\Box\]

The previous theorem has full generality. We do not require that two sequences \( a \) and \( b \) have the same period, nor that the sequence \( b \) is two-level autocorrelated. In the next theorem we focus on an important case when both two base sequences have two-level autocorrelation (and thus they are balanced as well).

**Example 4.9.** Take two Legendre sequences of periods 19 and 7, respectively. The first base sequence is \( a = (1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1) \) and the second base sequence is \( b = (1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 1) \). For example, we pick the shift sequence \( \underline{e} = (1 \ 2 \ 5 \ 9 \ 1 \ 7 \ \infty) \). Then the family is a \((133, 8, 21)\)-signal set, where \( 21 < 2\sqrt{133} \).

Let us consider another example. Take two Legendre sequences of periods 11 and 7 as two base sequences respectively. Then we construct a \((77, 8, 13)\)-signal set using the shift sequence \( \underline{e} = (2 \ 4 \ 8 \ 5 \ 10 \ 9 \ \infty) \). We note that \( 13 < 2\sqrt{177} \) in this case.

**Theorem 4.10.** If both \( a \) and \( b \) are two-level autocorrelation sequences with periods \( v \) and \( w \), respectively, then the family of sequences constructed by Algorithm 3.9 is a \((vw, w + 1, \delta_2)\) signal set with

\[
\delta_2 = \left( \left\lfloor \frac{w}{v} \right\rfloor + 1 \right) (v + 1) + 1.
\]

**Proof.** The proof is similar to Theorem 4.8 but requires further refinements. In particular, in this case \( \delta_b = 1 \) because sequence \( b \) is two-level autocorrelated.
• When \( k + s - h \equiv 0 \pmod{w} \), the correlation function

\[
C_b(k + s - h) = \sum_{j=0}^{w-1} (-1)^{b_{k+s-j} - b_{k+j}} = w.
\]

It is equivalent to that the sequence \( L^{k+s}(b) - L^h(b) \) is the zero sequence.

By Lemma 4.6, we have

\[
|C_{h,k}(\tau)| = \sum_{j=0}^{w-1} (-1)^{b_{k+s-j} - b_{k+j}} b(e_{j+s} - e_j + r)
\]

\[
= \sum_{j=0}^{w-1} C_a((e_{j+s} - e_j + r)\mod v)
\]

\[
\leq \left(\left\lceil \frac{w}{v} \right\rceil + 1 \right) \cdot \sum_{j=0}^{v-1} C_a((e_{j+s} - e_j + r)\mod v)
\]

\[
\leq \left(\left\lceil \frac{w}{v} \right\rceil + 1 \right) \cdot v.
\]

The last inequality holds because there is at most one \( j \) such that \( e_{j+s} - e_j + r \equiv 0 \pmod{v} \).

• When \( k + s - h \not\equiv 0 \pmod{w} \), let \( j_0, \ldots, j_{\left\lfloor \frac{w}{v} \right\rfloor} \) be the solutions from 0 to \( w-1 \) such that \( e_{j+s} - e_j + r \equiv 0 \pmod{v} \). We observe that this is the case that gives the worst possible cross correlation value. Then the correlation function satisfies

\[
|C_{h,k}(\tau)| = \left| \sum_{j=0}^{w-1} (-1)^{b_{k+s-j} - b_{k+j}} \cdot C_a((e_{j+s} - e_j + r)\mod v) \right|
\]

\[
= \left| (-1)^{b_{k+s-j_0} - b_{k+s-j_0}} \cdot C_a(0) + \cdots + (-1)^{b_{k+s-j_{\left\lfloor \frac{w}{v} \right\rfloor}} - b_{k+s-j_{\left\lfloor \frac{w}{v} \right\rfloor}} \cdot C_a(0) \right|
\]

\[
+ \sum_{j \neq j_0, \ldots, j_{\left\lfloor \frac{w}{v} \right\rfloor}} (-1)^{b_{k+s-j} - b_{k+j}} \cdot (-1)
\]

\[
= \left| \{\pm(v+1)\} + \cdots + \{\pm(v+1)\} + (-1) \cdot C_a(k + s - h) \right|
\]

\[
\leq \left(\left\lceil \frac{w}{v} \right\rceil + 1 \right) \cdot (v+1) + 1.
\]

Hence, the maximal absolute value of the correlation values is

\[
|C_{h,k}(\tau)| \leq \left(\left\lceil \frac{w}{v} \right\rceil + 1 \right) \cdot (v+1) + 1.
\]

\(\square\)
We emphasize again that it is desirable to generate many sequences in a family of sequences and thus we can assume \( w \geq v \). This means that we use \( a \) as the first base sequence and \( b \) as the second base sequence. In particular, when \( v = w \) and the two base sequences are two-level autocorrelated, we recover Gong’s result (Theorem 2 in [31] or page 364 in [27]).

**Corollary 4.11.** When \( v \) and \( w \) are equal, the family of sequences generated by Algorithm 3.9 is a \((v^2, v + 1, 2v + 3)\) signal set.

Next we obtain a few results when both base sequences \( a \) and \( b \) are Legendre sequences with the period equal to prime numbers \( p \) and \( q \), respectively. The following is the known result for the autocorrelation of Legendre sequences.

**Proposition 4.12.** [96]. Let \( s \) be a Legendre sequence of prime period \( p \). Then, if \( p \equiv 3 \pmod{4} \), \( C_s(\tau) = \{-1, p\} \), and if \( p \equiv 1 \pmod{4} \), \( C_s(\tau) = \{1, -3, p\} \).

This shows that the Legendre sequence \( a \) of period \( p \equiv 3 \pmod{4} \) has ideal two-valued autocorrelation \( C_a(\tau) = \{-1, p\} \).

**Corollary 4.13.** Fix a prime number \( p \equiv 3 \pmod{4} \) and any other prime \( q \geq p \). The family of sequences \( S \) generated by Algorithm 3.9 from two Legendre sequences of periods \( p \) and \( q \) is a \((pq, q + 1, \delta)\) signal set, where

\[
\delta = \delta_1 = \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p + 1) + q.
\]

Furthermore, when both \( p \) and \( q \) are congruent to \( 3 \pmod{4} \) we obtain

\[
\delta = \delta_2 = \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p + 1) + 1.
\]

We recall that Wang and Qi’s result is the case when taking two Legendre sequences \( a \) and \( b \) with twin prime periods \( p \equiv 3 \pmod{4} \) and \( q = p + 2 \), respectively.

**Corollary 4.14.** [88] Let two Legendre sequences of twin prime periods \( p \) and \( p + 2 \), where \( p \equiv 3 \pmod{4} \) be the base sequences under the construction of the interleaved algorithm. The maximum magnitude of the nontrivial cross correlation values of this constructed family is \( 3p + 4 \).

If the Legendre sequence \( a \) has prime period \( p \equiv 1 \pmod{4} \), then \( a \) is not two-level autocorrelated (in fact, it is three-level correlated). In this case, we slightly
modify this Legendre sequence so that we have a two-level autocorrelation sequence \( a' \).

**Lemma 4.15.** [75] Let \( a \) be a Legendre sequence with prime period \( p \equiv 1 \pmod{4} \). If we let the entries \( a'_i = -\infty \) with \( i \equiv 0 \pmod{p} \) and \( a'_i = a_i \) with \( i \not\equiv 0 \pmod{p} \), then the modified Legendre sequence \( a' \) has two-valued autocorrelation.

Now we estimate the maximal cross correlation of the interleaved construction from \( a \) and \( b \) by using the maximal cross correlation of the interleaved construction from \( a' \) and \( b \).

**Theorem 4.16.** Fix a prime number \( p \equiv 1 \pmod{4} \) and any other prime \( q \geq p \). The family of sequences \( \mathfrak{S} \) generated by Algorithm 3.9 from two Legendre sequences of periods \( p \) and \( q \) is a \((pq, q+1, \delta_3)\) family, where \( \delta_3 = \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p+1) + 3q - 2 \).

**Proof.** Fix a prime \( p \equiv 1 \pmod{4} \), we use the modified Legendre sequence \( a' = (-\infty, a_1, \ldots, a_{p-1}) \) as the first base sequence in Algorithm 3.9. Then by Theorem 4.8, the family of sequence \( \mathfrak{S}' \) constructed from \( a' \) and \( b \) has the maximal correlation value

\[
|C_{h,k}(\tau)'| \leq \delta_1 = \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p+1) + q.
\]

For each sequence \( s_j' \in \mathfrak{S}' \), we have \( s_j' = u' + L_j(b) \) and every column of the matrix form of the interleaved sequence \( u' \) contains a \(-\infty\). Therefore, there are \( q-1 \) copies of \(-\infty\)'s in every sequence \( s_j' \in \mathfrak{S}' \) for \( j = 0, \ldots, q \). If the sequence \( s_j \) in the family of sequences is constructed by Algorithm 3.9 from ordinary Legendre sequences \( a \) and \( b \), then the difference between \( s_j \) and the corresponding \( s_j' \) happens exactly at these entries of \(-\infty\) in \( s_j' \). Then the correlation function of any two sequences \( s_j = (s_{j0}, s_{j1}, \ldots, s_{j(q-1)}) \), \( s_k = (s_{k0}, \ldots, s_{k(q-1)}) \) in the family \( \mathfrak{S} \) constructed from both \( a \) and \( b \) are ordinary Legendre sequences is

\[
|C_{h,k}(\tau)| \leq |C_{h,k}(\tau)'| + 2(q - 1) \\
\leq \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p+1) + q + 2(q - 1) \\
= \left( \left\lfloor \frac{q}{p} \right\rfloor + 1 \right) \cdot (p+1) + 3q - 2.
\]

\( \square \)
4.3. Linear complexity of binary interleaved sequences

We focus on binary sequences in this chapter and so in the following two theorems we set the prime field to be \( \mathbb{F}_2 \). Then we study the linear complexity of every sequence in the interleaved constructed family \( \mathcal{S} \) as in Algorithm 3.9 associated with two binary sequences of integer periods \( s \) and \( t \), respectively. First, we need to know the generating function of every sequence in \( \mathcal{S} \).

The generating function for the interleaved sequence \( u \) as defined in Algorithm 3.9 (i.e. \( s, t \in \mathcal{S} \)) is

\[
G_u(x) = \sum_{j=0}^{t-1} \sum_{i=0}^{s-1} a_i \gamma^j x^{it+j},
\]

since when \( j = t - 1 \) the value of \( a_i \gamma^{j+1} = a_i \gamma^1 = 0 \).

If we know the two base sequences \( a \) and \( b \), then the generating function in one period of \( u \) can be expressed in terms of \( G_a(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{s-1} x^{s-1} \).

**Theorem 4.17.** For the interleaved sequence \( u \) in Algorithm 3.9 constructed from two random binary sequences of periods \( s \) and \( t \), respectively, the linear complexity of \( u \) is at least \( t - 1 \) if \( G_a(1) \neq 0 \).

**Proof.** We have

\[
G_u(x) \equiv \sum_{j=0}^{t-2} x^{j-te} \cdot G_a(x^t) \pmod{x^{st} - 1}
\]

\[
\equiv G_a(x^t) \cdot \sum_{j=0}^{t-2} x^{j-te} \pmod{x^{st} - 1}.
\]

We denote \( \xi(x) \) as

\[
\xi(x) \equiv 1 + x^{1-td_0} + x^{2-t(2d_0+d)} + \cdots + x^{t-2-t[(t-2)d_0+2(t-3)(t-2)]} \pmod{x^{st} - 1}.
\]

Then

\[
G_u(x) \equiv G_a(x^t) \cdot x^{-te_0} \cdot \xi(x) \pmod{x^{st} - 1}.
\]

Let \( \gamma \) be a primitive \( t \)th root of unity. Then in \( \mathbb{F}_{2^t} \) we have

\[
G_u(\gamma) = G_a(\gamma^t) \cdot (\gamma^{-te_0}) \cdot \xi(\gamma) = G_a(1) \cdot 1 \cdot \xi(\gamma)
\]

\[
= G_a(1) \cdot (1 + \gamma + \gamma^2 + \cdots + \gamma^{t-2}) = G_a(1) \cdot \frac{\gamma^{t-1} - 1}{\gamma - 1}.
\]
By the fact that
\[
\frac{\gamma^{t-1} - 1}{\gamma - 1} \neq 0
\]
if and only if
\[
\gamma^{t-1} \neq 1
\]
and this is true for all \(\gamma^h\) with \(0 < h < t\). Thus at least \(t - 1\) \(t\)th roots of unity are non-zeros of \(G_u(x)\). So we have
\[
LC(u) \geq t - 1 \text{ if } G_u(1) \neq 0.
\]
\[\square\]

Let us consider a binary interleaved sequence \(s_r\) in the family \(\mathcal{S}\) in Algorithm 3.9 different from \(u\).

**Theorem 4.18.** The linear complexity for any sequence \(s_r \in \mathcal{S}\) is
\[
LC(s_r) \leq t \cdot LC(u) + t \delta,
\]
where
\[
\delta = \begin{cases} 
1, & \text{if } 1 \text{ is a root of } G_u(x); \\
0, & \text{if } 1 \text{ is not a root of } G_u(x).
\end{cases}
\]

**Proof.** The generating function of a typical interleaved sequence \(s_r \in \mathcal{S}\) is
\[
G_{s_r}(x) = \sum_{j=0}^{t-1} \sum_{i=0}^{s-1} (a_{i+e_j} + b_{j+r}) x^{it+j}
\]
\[
= \sum_{j=0}^{t-1} \sum_{i=0}^{s-1} a_{i+e_j} x^{it+j} + \sum_{j=0}^{t-1} \sum_{i=0}^{s-1} b_{j+r} x^{it+j}
\]
\[
= G_u(x) + x^{-r} \cdot \sum_{i=0}^{s-1} x^{it} \cdot \sum_{j=0}^{t-1} b_{j+r} x^{j+r} \pmod{x^{st} - 1}
\]
\[
= x^{-te_0} \cdot G_u(x^t) \cdot \xi(x) + x^{-r} \cdot \frac{x^{st} - 1}{x^t - 1} \cdot G_u(x) \pmod{x^{st} - 1}.
\]

Let \(\beta\) be a primitive \((st)\)th root of unity and denote \(\beta^t = \eta\) and \(\beta^s = \gamma\). Then
\[
G_{s_r}(\beta) = \eta^{-e_0} \cdot G_u(\eta) \cdot \xi(\beta) + \beta^{-r} \cdot 0 \cdot G_u(\beta)
\]
\[
= \eta^{-e_0} \cdot G_u(\eta) \cdot \xi(\beta),
\]
because
\[
\frac{x^{st} - 1}{x^t - 1} = 0.
\]
if and only if \( x \) is equal to a \((st)\)th root of unity but not a \(t\)th root of unity. So we set \( x = \beta^{js+i} \) with \( i \neq 0 \) and \( 0 \leq j \leq t-1 \) and such \( \beta^{js+i} \) are solutions of \( \frac{x^{st}-1}{x^t-1} = 0 \). We have
\[
G_{s_r}(\beta^{js+i}) = \eta^{-i\alpha} \cdot G_{\alpha}(\eta^i) \cdot \xi(\beta^{js+i}) + \beta^{-r(js+i)} \cdot 0 \cdot G_{\beta}(\beta^{js+i}) = \eta^{-i\alpha} \cdot G_{\alpha}(\eta^i) \cdot \xi(\beta^{js+i}).
\]
With the possibility of 1 being a root or not, the total number of roots of \( G_{\alpha}(x) \) is \( s - LC(\alpha) - \delta \), where \( \delta \) is defined by
\[
\delta = \begin{cases} 
1, & \text{if } 1 \text{ is a root of } G_{\alpha}(x); \\
0, & \text{if } 1 \text{ is not a root of } G_{\alpha}(x).
\end{cases}
\]
It means there are \( s - LC(\alpha) - \delta \) such \( \eta^i \) roots, and for all \( j = 0, 1, \ldots, t-1 \), there are in total \( t(s - LC(\alpha) - \delta) \) such \( \beta^{js+i} \) roots of \( G_{s_r} \). Therefore,
\[
LC(s_r) \leq st - t(s - LC(\alpha) - \delta) = t \cdot LC(\alpha) + t\delta.
\]
\[\square\]

We observe that if the base sequences are arbitrary binary sequences, we neither have enough information to find the minimal polynomial of the interleaved sequences nor the linear complexity. In the next chapter we present a way to find a lower bound for the linear complexity. In fact, we study the linear complexity profile, the supreme of which is the linear complexity.
Linear complexity profile, correlation measures
and nonlinearity

In this chapter we study linear complexity profile and power correlation measures of order $k$ of a signal set of $m$-ary interleaved sequences. Section 5.1 gives a lower bound for the linear complexity profile of the family of $m$-ary interleaved sequences. Section 5.2 focus on the power correlation measure of order $k$. We give several results in this section. A general relation between the linear complexity profile and the power correlation measures with lags bounded by $B$ is given. Then we study the power correlation measures with lags bounded by $B$ for the $m$-ary interleaved sequences. Finally, we use this relation to derive a lower bound on the linear complexity profile for interleaved sequences. We also consider a special family of interleaved sequences, which requires the first base to be a Legendre sequence. We study the correlation measures for interleaved Legendre sequences. We also show through examples that the power correlation measures can be trivial if there is no bound $B$ on the lags.

The above way to estimate power correlation measures can also be applied to find the nonlinearity of the boolean function derived from binary interleaved sequence. We get an upper bound on the nonlinearity in Section 5.3 using this idea.

5.1. Linear complexity profile of $m$-ary interleaved sequences

We recall that the linear complexity profile is the sequence consisting of the $N$th linear complexity for an integer $N \geq 2$. The $N$th linear complexity $L(s, N)$ of a sequence $s = (s_n)_{0 \leq n < T}$ of period $T$ over a ring $\mathbb{Z}_m$ is the smallest positive integer $L$ such that there are constants $c_1, \ldots, c_L \in \mathbb{Z}_m$ satisfying, for $0 \leq n < N - L$, the
recursive relation

\[ s_{n+L} = c_L^{-1}s_{n+L-1} + c_L^{-2}s_{n+L-2} + \cdots + c_0s_n. \]

In this and the next section we estimate some inequalities for the linear complexity profile using the bound given in Theorem 2.42. Later we use \( \gg, \ll \) notation which was introduced in Notation 2.41. If \( A \ll B \) or \( A = O(B) \), it means \( |A| \leq kB \) for some constant \( k \).

Let \( p \) be an odd prime and \( m \) a divisor of \( p-1 \). Let \( \alpha \) be a primitive element of \( \mathbb{F}_p \). The nonzero elements in \( \mathbb{F}_p^* \) can be partitioned into \( m \) cosets \( U_k := \{ \alpha^n \mid n \equiv k \pmod{m} \} \) for \( k = 0, \ldots, m-1 \) and \( n = 0, 1, \ldots, p-1 \). We recall that an \( m \)-phase (-ary) power residue sequence \( (a_n)_{0 \leq n < p} \) of period \( p \) is defined as

\[ a_n = \begin{cases} 
0, & \text{if } n = 0; \\
k, & \text{if } n \in U_k.
\end{cases} \]

We use \( m \)-ary power residue sequence as the first base sequence and any other sequence \( b \) over \( \mathbb{Z}_m \) as the second base sequence to construct a signal set of sequences by means of Algorithm 3.9.

Let us see first the binary case, that is \( m = 2 \). If we consider the two base sequences \( a \) and \( b \) to be Legendre sequences (binary) of odd prime periods \( p \) and \( q \), respectively, we get the following result, which gives a lower bound for the linear complexity.

**Theorem 5.1.** The \( N \)th linear complexity \( L \) of the interleaved sequence \( y_{\bar{r}} \) defined in Algorithm 3.9 associated with two Legendre sequences of odd prime periods \( p \) and \( q \), respectively, satisfies

\[ L \geq \frac{\min\{N, pq\}}{q \cdot p^{1/2} \log p + 4p + 4q + 1}. \]

**Proof.** Since \( L(y_{\bar{r}}, N) \) is monotonic increasing till reaching the supremum, the linear complexity, we may assume \( N \leq 2pq \). When \( N > 2pq \), the value of \( L(y_{\bar{r}}, N) \) is always equal to the linear complexity.

Put \( c_L = -1 = 1 \in \mathbb{F}_2 \) and let

\[ \sum_{l=0}^{L} c_l y_{n+l} = 0, \quad 0 \leq n \leq N - 1 - L, \]
be a linear recurrence sequence over $\mathbb{F}_2$ of $y_r$ of order $L$, which is equivalent to

$$1 = \prod_{l=0}^{L} (-1)^{y_{n+l}}, \quad 0 \leq n \leq N - 1 - L. \quad (23)$$

We write

$$n = iq + j, \quad 0 \leq j \leq q - 1$$

and get for $0 \leq r < q$

$$(-1)^{y_n} = \left( \frac{e_j + i}{p} \right) \left( \frac{j + r}{q} \right), \quad (24)$$

if

$$e_j + i \not\equiv 0 \pmod{p}, \quad j + r \not\equiv 0 \pmod{q}, \quad 0 \leq j \leq q - 2,$$

and for $r = q$

$$(-1)^{y_n} = \left( \frac{e_j + i}{p} \right)$$

if

$$e_j + i \not\equiv 0 \pmod{p}, \quad 0 \leq j \leq q - 2.$$

We present the proof only for the slightly more complicated case $0 \leq r < q$.

There are at most

$$(L + 1)p$$

different

$$0 \leq n \leq N - L - 1 \leq pq - 1$$

with

$$n = iq + j \text{ and } q - 1 - L \leq j \leq q - 1,$$

and at most

$$(L + 1)p$$

different $n$ in this range with $j + r + l \equiv 0 \pmod{p}$ for some $0 \leq l \leq L$. We recall that

$$e_j = e_0 + jd_0 + \frac{j(j - 1)}{2}d, \quad 0 \leq j \leq q - 2,$$

is a polynomial in $j$ of degree 2, for each $0 \leq i < p$, and there are at most 2 different $j$ modulo $p$ with $e_j + i \equiv 0 \pmod{p}$ giving at most

$$2(L + 1)p \left\lceil \frac{q}{p} \right\rceil$$
different
\[ n = iq + j \text{ with } e_{j+l} + i \equiv 0 \pmod{p} \]
for some \( 0 \leq l \leq L \). For \( i = \left\lfloor \frac{N-L}{q} \right\rfloor \) there are at most \( q \) different \( j \) such that \( iq + j \leq N - L - 1 \).

We sum (23) over \( n = iq + j \) and get by (24)
\[ N - L \leq \sigma + (L + 1)(2p \left\lceil \frac{q}{p} \right\rceil + p), \tag{25} \]
where \( \sigma \) is
\[ \left| \sum_{j=0}^{q-2-L} \sum_{i=0}^{\left\lfloor \frac{N-L}{q} \right\rfloor-1} \left( \prod_{l=0}^{L} (e_{j+l} + i)^{c_l} \right) \left( \frac{\prod_{l=0}^{L} (j + r + l)^{c_l}}{q} \right) \right|. \]

and
\[ \sigma \leq \sum_{j=0}^{q-2-L} \left| \sum_{i=0}^{\left\lfloor \frac{N-L}{q} \right\rfloor-1} \left( \prod_{l=0}^{L} (e_{j+l} + i)^{c_l} \right) \right|. \]

The last term \((L + 1)(2p \left\lceil \frac{q}{p} \right\rceil + p)\) comes from the argument that when \( 0 < n = iq + j < N - L < pq \) there are at most \( 2p \left\lceil \frac{q}{p} \right\rceil + p \) times for the Legendre symbols not defined corresponding to \((-1)^s_n\) and there are at most \( L + 1 \) copies in the sum.

We note that \( c_l \in \mathbb{F}_2 \). For \( c_l = 0 \), the Legendre symbols \( \left( \frac{e_{j+l} + i}{p} \right)^0 \) and \( \left( \frac{j + r + l}{q} \right)^0 \) in the sum are equal to 1. So in the sum there are only the Legendre symbols \( \left( \frac{e_{j+l} + i}{p} \right)^1 \) and \( \left( \frac{j + r + l}{q} \right)^1 \) left when \( c_l = 1 \), for some \( 0 \leq l < L \).

We rename such \( l's \) satisfying \( c_l = 1 \) by \( \tau_i' \)'s. We observe that in total there are at most \( l \) such \( \tau_i' \)'s. Under the same condition of \( 0 \leq l \leq L \), we set \( 0 \leq \tau_0 < \cdots \tau_i < \cdots < \tau_v \leq L \) with \( 0 \leq i \leq v \leq L \). When \( 0 \leq n < N - L \) and \( j \) runs from 0 to \( q - 1 \), the \( i \) is bounded by \( 0 \leq i \leq \left\lfloor \frac{N-L}{q} \right\rfloor \). We have that
\[ \sigma = \left| \sum_{j=0}^{q-1} \sum_{i=0}^{\left\lfloor \frac{N-L}{q} \right\rfloor-1} \left( \frac{e_{j+\tau_0} + i}{p} \right) \left( \frac{j + r + \tau_0}{q} \right) \cdots \left( \frac{e_{j+\tau_v} + i}{p} \right) \left( \frac{j + r + \tau_v}{q} \right) \right|. \]

The idea to estimate \( \sigma \) is following. First, to simplify and cancel \( \left( \frac{j + r + \tau_0}{q} \right) \cdots \left( \frac{j + r + \tau_v}{q} \right) \)
we estimate \( \frac{t^*+t_0}{q} \cdots \frac{t^*+t_n}{q} \) by the trivial bound 1. When only

\[
\left( \frac{e_{j+t_0} + i}{p} \right) \cdots \left( \frac{e_{j+t_n} + i}{p} \right)
\]

are left in the sum, to avoid the carryover of \( j + \tau_v \geq q \) we then split the sum and estimate the complete sum for \( j = 0 \) to \( q - 1 \) into an incomplete sum for \( j = 0 \) to \( q - 1 - L \) and \( L \) copies of the trivial bound 1 of \( \left( \frac{e_{j+t_0}+i}{p} \right) \cdots \left( \frac{e_{j+t_n}+i}{p} \right) \). Finally, we can apply the standard bound on the incomplete sum of \( \left( \frac{e_{j+t_0}+1}{p} \right) \cdots \left( \frac{e_{j+t_n}+1}{p} \right) \) since there are no carryovers, i.e. there are no squares in \( \left( \frac{e_{j+t_0}+1}{p} \right) \cdots \left( \frac{e_{j+t_n}+1}{p} \right) \). Then \( \sigma \) is estimated by

\[
\sigma \leq \sum_{j=0}^{q-1} \left| \sum_{i=0}^{N-L} \left( \frac{e_{j+t_0} + i}{p} \right) \cdots \left( \frac{e_{j+t_n} + i}{p} \right) \cdot 1 \right|
\]

\[
\leq \sum_{j=0}^{q-1-L} \left| \sum_{i=0}^{N-L} \left( \frac{e_{j+t_0} + i}{p} \right) \cdots \left( \frac{e_{j+t_n} + i}{p} \right) + L \right| \sum_{i=0}^{N-L} 1
\]

\[
< (q-L) \cdot \max_{j \in \mathbb{F}_q} \left| \sum_{i=0}^{N-L} \left( \frac{e_{j+t_0} + i}{p} \right) \cdots \left( \frac{e_{j+t_n} + i}{p} \right) + L \right| \cdot \frac{N-L}{q}
\]

Because there is no carryover in \( \left( \frac{e_{j+t_0}+i}{p} \right) \cdots \left( \frac{e_{j+t_n}+i}{p} \right) \) and the degree of it is at most \( L \), the inner sum above can be estimated by \( Lp^{1/2} \log p \) using the bound in Theorem 2.43. Hence,

\[
\sigma < q \cdot \max_{j \in \mathbb{F}_q} \left| \sum_{i=0}^{N-L} \left( \frac{e_{j+t_0} + i}{p} \right) \cdots \left( \frac{e_{j+t_n} + i}{p} \right) + L \right| \cdot \frac{N-L}{q}
\]

\[
< q \cdot \frac{Lp^{1/2} \log p + L \cdot (N-L)}{q}.
\]

Collecting everything, we have the inequality

\[
N - L \leq q \cdot \frac{Lp^{1/2} \log p + L \cdot (N-L)}{q} + (L + 1) \left( 2p \left( \left\lfloor \frac{q}{p} \right\rfloor \right) + p \right).
\]

We note that \( \frac{N-L}{q} \leq p \leq 2p \). We solve for \( L \) and get

\[
L \geq \frac{N - (2p \left\lfloor \frac{q}{p} \right\rfloor + p)}{1 + qp^{1/2} \log p + 2p \left\lfloor \frac{q}{p} \right\rfloor + 2p}.
\]
We simplify the result further in

\[
N - L \leq q \cdot L^{p^{1/2}} \log p + L \cdot \frac{N - L}{q} + (L + 1) \left( 2p \left\lceil \frac{q}{p} \right\rceil + p \right)
\]

\[
N \leq q \cdot L^{p^{1/2}} \log p + L \cdot 2p + (L + 1) (2q - 1 + p) + L
\]

\[
\leq q \cdot L^{p^{1/2}} \log p + 2pL + (L + 1) (2q + p)
\]

\[
\leq q \cdot L^{p^{1/2}} \log p + 2pL + 2L (2q + p)
\]

\[
= q \cdot L^{p^{1/2}} \log p + 4pL + 4L.
\]

Finally, we get

\[
L \geq \frac{N}{q \cdot p^{1/2} \log p + 4p + 4q + 1}.
\]

When \( N > 2pq \), the linear profile is equal to the linear complexity, i.e., \( L(\underline{a}, N) = LC(\underline{a}) \). So we have

\[
L \geq \frac{\min\{N, 2pq\}}{q \cdot p^{1/2} \log p + 4p + 4q + 1}.
\]

\[\square\]

The result can be easily generalized to \( m \)-ary interleaved sequences. We adopt the notations of \( \ll \) and \( \gg \) in the next theorem.

**Theorem 5.2.** Let \( p > 2 \) be a prime number. For \( 0 \leq r \leq t \) let \( y_r \) be an interleaved sequence as defined in Algorithm 3.9 of two \( m \)-ary sequences \( \underline{a} \) and \( \underline{b} \) of periods \( s = p \) and \( t \), where \( \underline{a} \) is an \( m \)-phase power residue sequence and \( \underline{b} \) can be any \( m \)-ary sequence. Then we have

\[
L(y_r, N) \gg \frac{\min\{N, pt\}}{tp^{1/2} \log p + p}.
\]

**Proof.** As in the previous theorem we can assume \( N \leq pt \).

Let \( m \) be an integer dividing \( p - 1 \). Put \( c_L = -1 \) and let

\[
\sum_{l=0}^{L} c_l y_{n+l} = 0, \quad 0 \leq n \leq N - 1 - L,
\]

be a linear recurrence sequence over \( \mathbb{Z}_m \) of \( y_r \) of order \( L \), which is equivalent to

\[
1 = \prod_{l=0}^{L} (\omega_m)^{c_l y_{n+l}}, \quad 0 \leq n \leq N - 1 - L,
\]

where \( \omega_m \) is a primitive \( m \)th root of unity.
Let $\psi$ denote the multiplicative character of $\mathbb{F}_p$

$$\psi(x) = \exp\left(\frac{2\pi\sqrt{-1}\log_\alpha x}{m}\right),$$

where $\alpha$ is a primitive root of $\mathbb{F}_p$. We write

$$n = it + j, \quad 0 \leq j \leq t - 1,$$

and for $0 \leq r < t$ since $y_n = a_{e_j+i} + b_{j+r}$ we have

$$(\omega_m)^{y_n} = \psi(e_j + i)\omega_m^{j+r},$$

if $e_j + i \not\equiv 0 \pmod{p}$. We note that since $\omega_m$ is an $m$th root of unity, $|\omega_m^j| = 1$ for all $j$.

For $r = t$ we get

$$(\omega_m)^{y_n} = \psi(e_j + i),$$

if $e_j + i \not\equiv 0 \pmod{p}$ and $0 \leq j \leq t - 2$. We present the proof only for the slightly more complicated case $0 \leq r < t$.

There are at most $(L + 1)p$ different

$$0 \leq n \leq N - L - 1 \leq pt - 1$$

with

$$n = it + j$$

and $t - 1 - L \leq j \leq t - 1$.

Since

$$e_j = e_0 + jd_0 + \frac{j(j-1)}{2}d, \quad 0 \leq j \leq t - 2,$$

is a polynomial modulo $p$ in $j$ of degree 2, for each $0 \leq i < p$, there are at most 2 different $j$ modulo $p$ with $e_j + i \equiv 0 \pmod{p}$ giving at most

$$2(L + 1)p\left\lfloor \frac{t}{p} \right\rfloor$$

different

$$n = it + j$$

with $e_{j+l} + i \equiv 0 \pmod{p}$ for some $0 \leq l \leq L$. For $i = \left\lfloor \frac{N-L}{t} \right\rfloor$ there are at most $t$ different $j$ such that $it + j \leq N - L - 1$. We estimate the error term $(L + 1)p + 2(L + 1)p\left\lfloor \frac{t}{p} \right\rfloor$ from above by $O(L(p + t))$.

We sum (26) over $n = it + j$ for $0 \leq n \leq N - 1 - L$ and get by (27)

$$(28) \quad N - L = \sigma + O(L(p + t)), $$
where
\[
\sigma = \left| \sum_{j=0}^{t-2-L} \sum_{i=0}^{\frac{N-L}{l}-1} \psi \left( \prod_{l=0}^{L} (e_{j+l} + i)^{c_l} \right) \cdot \sum_{m=0}^{L} (j+r+l)^{c_l} \right|.
\]

For each fixed \( L, 0 \leq l < L < p \) the difference
\[
e_{j+L} - e_{j+l} = (L - l)d_0 + \frac{L^2 - L - l^2 + l}{2}d + (L - l)d_j
\]
is a polynomial in \( j \) of degree 1 and there are at most \( \lceil t/p \rceil \) different \( j \) with
\[
e_{j+l} \equiv e_{j+L} \pmod{p},
\]
such that
\[
\prod_{l=0}^{L} (e_{j+l} + i)^{c_l}
\]
contains an \( m \)th power if \( c_l + c_L \geq m \). For these \( j \) we estimate the inner sum trivially by \( p \). For the remaining at most \( t \) different \( j \) we may apply Theorem 2.42, or see [8] or [77, Lemma 3.4] for incomplete character sums, which is applicable since \( c_L = -1 \) and
\[
\prod_{l=0}^{L} (e_{j+l} + X)^{c_l}
\]
is not an \( m \)th power, and we get
\[
\sigma \ll L(tp^{1/2} \log p + p).
\]
The bound in the statement of the theorem follows from (28) after simple calculations.

\[
N - L \ll L(tp^{1/2} \log p + p)
\]
\[
L(tp^{1/2} \log p + p) \gg N
\]
\[
L \gg \frac{N}{tp^{1/2} \log p + p}.
\]
5.2. Relating linear complexity profile and (power) correlation measures

There is an alternative method to find a lower bound on the linear complexity profile using a relation between the linear complexity profile and the (power) correlation measures. Power correlation measure is one of the possible randomness properties of interest. Our strategy is to find a nontrivial upper bound on the power correlation measure and derive a lower bound on linear complexity using the relation between them. We find through this method the lower bound of the linear complexity profile for interleaved sequences that is higher than the one obtained in Section 5.1 in some cases, like when the period $p$ of the first base sequence is large.

The correlation measure of order $k$ of a finite pseudo-random binary sequence was introduced by Mauduit and Sarkozy [62]. These authors then generalized the definition to $m$-ary sequences (sequences of $m$ symbols) [63]. Brandstätter and Winterhof [6] proved a general relation between linear complexity profile and the correlation measure of order $k$ for binary sequences from which strong lower bounds on the linear complexity profile can be derived whenever suitable upper bounds on the correlation measure of order $k$ are known; this is the case for Legendre sequences. Chen and Winterhof then considered a relation between the linear complexity profile and the correlation measures for $m$-ary sequences. If $m$ is a prime, the result is a straightforward extension of the binary case in [6]. For composite $m$ they introduced and studied the power correlation measure and gave this relation in [12].

We adopt the notation in Definition 2.10. As we will see later, to observe the correlation measures with bounded lags, we introduce below the definition of power correlation measure of order $k$ with bound $B$.

**Definition 5.3.** Let $m \geq 2$ and the complex number $\omega = \exp\left(\frac{2\pi i \sqrt{-1}}{m}\right)$, and thus $\mathcal{E}_m = \{\omega^j : 0 \leq j \leq m-1\}$. For $c \in \{1, 2, \ldots, m-1\}$ we define $\mu_c : \mathbb{Z}_m \to \mathcal{E}_m$ by

$$\mu_c(z) = \omega^{zc}, \quad z \in \mathbb{Z}_m.$$ 

We note that $\mu_c \in \mathcal{F}$ if and only if $\gcd(c, m) = 1$, where $\mathcal{F}$ denotes the set of $m!$ bijections $\varphi : \mathbb{Z}_m \to \mathcal{E}_m$. Let

$$\mathcal{F}' = \{\mu_c : c = 1, 2, \ldots, m-1\}.$$
The power correlation measure of order \( k \) with lags bounded by \( B \) of the \( m \)-ary sequence \( \mathbf{s} = (s_0, \ldots, s_{T-1}) \in \mathbb{Z}_m^T \) is

\[
\Gamma_k'(\mathbf{s}, B) = \max_{\phi, M, D} \left| \sum_{n=1}^{M} \prod_{i=1}^{k} \mu_{c_i}(s_{n+\tau_i}) \right|,
\]

where the maximum is taken over all \((\mu_{c_1}, \ldots, \mu_{c_k})\) with \( \mu_{c_i} \in \mathcal{F}' \) for \( i = 1, \ldots, k \), the set \( D = (\tau_1, \ldots, \tau_k) \) of lags with non-negative integers \( 0 \leq \tau_1 < \cdots < \tau_k \leq B \) and any non-negative integer \( M \) such that \( M + \tau_k \leq T \).

In particular, the correlation measure of order \( k \geq 1 \) with lags bounded by \( B \) of a binary sequence \( \mathbf{y} = (y_n)_{0 \leq n < T} \) of period \( T \) is defined as

\[
C_k(\mathbf{y}, B) = \max_{M, D} \left| \sum_{n=0}^{M-1} (-1)^{y_{n+\tau_1} + y_{n+\tau_2} + \cdots + y_{n+\tau_k}} \right|,
\]

with the same \( M, D \) and \( 0 \leq \tau_1 < \cdots < \tau_k \leq B \) as before.

Similar to the proof of [12, Theorem 1], we get a relation between the linear complexity profile and the bounded power correlation measure of order \( k \) with bound \( B \) for \( m \)-ary sequences.

We briefly overview our results in the following. We give a general relation between the linear complexity profile and the power correlation measures with bound \( B \) in Proposition 5.4. Then we use this relation to derive a lower bound on the linear complexity profile for interleaved sequences. First, in Theorem 5.8 we study the power correlation measures with bound \( B \) for interleaved sequences.

Next, we consider a special family of interleaved sequences, which require the first base to be a Legendre sequence. We study the correlation measures for interleaved Legendre sequences in Theorems 5.12, 5.13 and 5.16. At the end of this section we show some examples where the power correlation measures can be trivial if there is no bound \( B \) on the lags.

The following proposition shows a relation between linear complexity profile and power correlation measures.

**Proposition 5.4.** The linear complexity profile and the power correlation measure with lags bounded by \( B \) of an \( m \)-ary sequence \( \mathbf{s} \) satisfy

\[
L(\mathbf{s}, N) \geq \min\{N - \max_{1 \leq k \leq L(\mathbf{s}, N) + 1} \Gamma'_k(\mathbf{s}, B), B\}.
\]
Proof. We write \( L = L(g, N) \) and choose \( c_0, \ldots, c_{L-1} \in \mathbb{Z}_m \) such that

\[
s_{n+L} = c_{L-1}s_{n+L-1} + \cdots + c_0s_n, \quad 0 \leq n \leq N - L - 1.
\]

With \( c_L = -1 \), we get

\[
1 = \omega_m^{c_Ls_{n+L-1} + \cdots + c_0s_n}, \quad 0 \leq n \leq N - L - 1,
\]

where \( \omega_m \) is a primitive \( m \)th root of unity and hence

\[
N - L = \sum_{n=0}^{N-L-1} \omega_m^{\sum_{i=0}^{L} c_is_{n+i}}.
\]

Following the notation in Definition 5.3, we derive from (29) that

\[
N - L = \sum_{n=0}^{N-L-1} \mu_c(s_n)\mu_{c_1}(s_{n+1})\cdots\mu_{c_L}(s_{n+L}) \leq \max_{1 \leq k \leq L+1} G_k^r(g, B),
\]

and the result follows.

Next we give an upper bound for the power correlation measures (with the bound \( B \) on the lags) for some special families of interleaved sequences. A special case of the interleaved family \( \mathcal{G} \) happens when we employ a power residue sequence as the first base.

Notation 5.5. As a special case of \( \mathcal{G} \) defined as in Algorithm 3.9 we denote \( \mathcal{Y}_a \) the family of interleaved complex-valued sequences associated with the first base \( g \), an \( m \)-ary power residue sequence of odd prime period \( p \) where \( m \mid (p - 1) \), and the second base \( b \), an \( m \)-ary sequence of positive integer period \( t \). We note that in Algorithm 3.9 we put \( d_{i-1} = e_i - e_{i-1} \), and then we choose \( g \) such that \( d_0, d_1, \ldots, d_{t-3} \) is in an arithmetic progression with common distance \( d \neq 0 \) (mod \( s \)). It implies

\[
e_j = e_0 + jd_0 + \frac{j(j - 1)}{2}d, \quad 0 \leq j \leq t - 2.
\]

We remark that the sequences in \( \mathcal{G} \) in Algorithm 3.9 and so thus \( \mathcal{Y}_a \) can also be written in the following form:

- Let \( u_{it+j} \) denote the element in the \( i \)th row and the \( j \)th column in the matrix form of sequence \( u \). Then, for \( 0 \leq i < s \) and \( 0 \leq j < t \),

\[
u_{it+j} = a_{e_j+i}.
\]
• For a sequence \( y_r = (y_0, \ldots, y_{st-1}) \in \mathcal{G} \) where \( 0 \leq r \leq t - 1 \), the element in the \( i \)th row and \( j \)th column in its matrix form can be written as
\[
y_{i t+j} = a_{e_{j+i}} + b_{j+r}, \text{ for } 0 \leq i < s \text{ and } 0 \leq j < t.
\]

• We also define, for \( 0 \leq j < t \),
\[
e_{j+t} = 1 + e_j.
\]

We study the power correlation measure of order \( k \) and linear complexity profile of the sequences in the constructed family \( Y_e \).

**Theorem 5.6.** Let \( a \) and \( b \) be \( m \)-ary sequences of positive integer periods \( s \) and \( t \). For \( y_r \in \mathcal{G}(a,b,e) \) as defined in Algorithm 3.9 for \( r = 0, \ldots, t \), the power correlation measure of order \( k \) of \( y_r \) is upper bounded by
\[
\Gamma_k(y_r) \leq t \cdot \Gamma_k(a).
\]

For a positive integer \( B \) and \( k \leq B + 1 \), we also have
\[
\Gamma_k(y_r, B) \leq t \cdot \Gamma_k(a, B).
\]

**Proof.** Set \( n = it + j \) for \( 0 \leq j < t \). Following the notations \( M, D \) and \( \phi \) in Definition 5.3, we have
\[
\Gamma_k(y_r) = \max_{M, D, \phi} \left| \sum_{n=0}^{M-1} (\omega_m) \sum_{i=0}^{t-1} \sum_{h=1}^{l-1} c_h a_{e_{j+i} + \tau_h + i} + c_h b_{j+r} + \tau_h \right|,
\]
\[
\leq \max_{M, D, \phi} \left| \sum_{n=0}^{M-1} (\omega_m) \sum_{i=0}^{t-1} \sum_{h=1}^{l-1} c_h a_{e_{j+i} + \tau_h + i} + c_h b_{j+r} + \tau_h \right|,
\]
\[
\leq \max_{M, D, \phi} \left| \sum_{n=0}^{M-1} (\omega_m) \sum_{i=0}^{t-1} \sum_{h=1}^{l-1} c_h a_{e_{j+i} + \tau_h + i} + c_h b_{j+r} + \tau_h \right|,
\]
\[
= t \cdot \Gamma_k(a),
\]

Therefore,
\[
\Gamma_k(y_r) \leq t \cdot \Gamma_k(a).
\]
If $0 < \tau_1 < \cdots < \tau_k \leq B$, it implies $k \leq B + 1$ and the result also holds for the power correlation measure with bound $B$, that is,

$$\Gamma'_k(y_r, B) \leq t \cdot \Gamma'_k(a, B).$$

□

Proposition 5.4 immediately implies the following corollary.

**Lemma 5.7.** Fix an integer $B$. For $k \leq B + 1$ and $y_r \in \mathcal{S}$

$$L(y_r, N) \geq N - t \cdot \max_{1 \leq k \leq L(y_r, N) + 1} \Gamma'_k(a, B).$$

Theorem 5.6 shows that the power correlation measure of an interleaved sequence can be estimated by the power correlation measure of its first base sequence. We are interested in finding the power correlation measures of the family of interleaved sequences associated with the power residue sequences as the base sequences, particularly Legendre sequences. Mauduit and Sarkozy give the upper bound $kp^{1/2} \log p$ for the correlation measure of order $k$ for a Legendre sequence of period $p$ in [62, Theorem 1] without considering the case $i \equiv 0 \pmod{p}$ for Legendre symbol $\left( \frac{i}{p} \right)$. The exceptional case needs to be considered by adding up to $kp$. From Theorem 5.6 we can not get any information about the correlation measures for interleaved sequences since the upper bound is greater than the trivial bound. Hence, we prove the correlation measures of the sequences in the family $Y_e$ in Notation 5.5 in a different manner.

**Theorem 5.8.**

1. The power correlation measure of order $k$ of $y_r \in Y_e$ is bounded by

$$\Gamma'_k(y_r) \ll k(tp^{1/2} \log p + p)$$

if one of the conditions (i) $p \geq 2t$ or (ii) $p = t$ holds.

2. If $B = p$ and $k \leq B + 1$, the power correlation measure of order $k$ with lags bounded by $B$ of $y_r \in Y_e$ is upper bounded by

$$\Gamma'_k(y_r, B) \ll k(tp^{1/2} \log p + p).$$

**Proof.** We denote

$$S_M = \sum_{n=0}^{M-1} (\omega_m)^{c_1 y_{n+\tau_1} + c_2 y_{n+\tau_2} + \cdots + c_k y_{n+\tau_k}}.$$
In the following, \( \phi = (\omega^c_1, \ldots, \omega^c_m) \), \( M \) and \( D = (\tau_1, \ldots, \tau_k) \) are taken as in Definition 5.3, and \( k \) is a positive integer such that \( 0 \leq \tau_1 < \cdots < \tau_k \leq pt - M \). Then \( \Gamma'_{k}(y_r) \) can be expressed as

\[
\Gamma'_k(y_r) = \max_{M, \phi, D} |S_M|.
\]

Here \( \phi \) corresponds to a \( k \)-tuple \( (c_1, \ldots, c_k) \in \mathbb{Z}_m^k \). Therefore, the maximum is taken over all non-negative integers \( \tau_1 < \tau_2 < \cdots < \tau_k \), all \( (c_1, \ldots, c_k) \in \mathbb{Z}_m^k \) and all \( M \) such that \( M + \tau_k \leq pt \).

We write \( n = it + j \) for \( 0 \leq j < t \) and \( 0 \leq i < p \). For \( j \geq 0 \) we put

\[
e^j = 1 + e_j
\]

which is consistent with (31) and (32). We get

\[
|S_M| \leq \sum_j \left\lfloor \frac{2^j - 1}{p} \right\rfloor \sum_{i=0}^{2^j - 1} \psi(e_j + \tau_1 + i) c_1 \cdots \psi(e_j + \tau_k + i) c_k + O(k(p + t)),
\]

where \( \psi \) is defined in (41) and the outer sum runs through all \( 0 \leq j < t \) with \( j \not\equiv t - 1 - \tau_h \pmod{t} \) for \( 1 \leq h \leq k \), that is,

\[
e_{j+\tau_h} \not\equiv \infty.
\]

The argument yields the error term \( O(kp) \) since for each \( \tau_h \) there are at most one \( j \) such that \( e_{j+\tau_h} = \infty \) and \( p \) choices of \( i \) such that \( e_{j+\tau_h} + i = \infty \). We also have an error term \( O(kt) \) for the terms \( n = it + j \) such that \( e_{j+\tau_h} + i \equiv 0 \pmod{p} \). Indeed, by (30), \( e_{j+\tau_h} + i \equiv 0 \pmod{p} \) is a polynomial of \( j \) of degree 2. For each estimated \( \tau_h \), there are at most \( p \cdot 2^{\left\lfloor \frac{t}{p} \right\rfloor} \approx O(t) \) choices of \( n = it + j \) satisfying \( e_{j+\tau_h} + i \equiv 0 \pmod{p} \). Combining the two error terms we obtain \( O(k(p + t)) \) in (34).

For \( 1 \leq h \leq k \) each \( c_h \) is less than \( m \). But the sum of any two of the \( c_i \)'s can be greater than or equal to \( m \). The product \( \prod_{h=1}^{k} (e_{j+\tau_h} + i)^{c_h} \) in the character \( \psi \) can be an \( m \)th power if any two distinct factors are equal. Without loss of generality, we consider the case of \( e_{j+\tau_1} + i \equiv e_{j+\tau_h} + i \pmod{p} \) for \( 2 \leq h \leq k \).

We use the notation

\[
\tau_h = s_h + v_h t
\]
with, for $1 \leq h \leq k$,

$$0 \leq v_h \leq p \quad \text{and} \quad 0 \leq s_h < t.$$  \hspace{1cm} (35)

For $0 \leq j < t$ and $1 \leq h \leq k$, we put

$$\delta_h(j) = \begin{cases} 
0, & j < t - s_h, \\
1, & j \geq t - s_h.
\end{cases}$$

If for any $2 \leq h \leq k$, either

$$s_1 - s_h \not\equiv t(\delta_1(j) - \delta_h(j)) \pmod{p}$$

or

$$s_1 - s_h \equiv t(\delta_1(j) - \delta_h(j)) \pmod{p} \quad \text{and} \quad v_1 - v_h \not\equiv \delta_h(j) - \delta_1(j) \pmod{p},$$

we estimate

$$\sum_{j} \left| \sum_{i=0}^{\lfloor \frac{m}{k} \rfloor - 1} \psi \left( (e_{j+r_1} + i)^{c_1} \cdots (e_{j+r_k} + i)^{c_k} \right) \right| \ll k \left( tp^{1/2} \log p + p \max_{2 \leq h \leq k} F_h \right),$$

where $F_h$ is the number of $0 \leq j < t$ with

$$e_{j+r_1} \equiv e_{j+r_k} \pmod{p}$$

or equivalently

$$e_{j+s_1-t\delta_1(j)} + v_1 + \delta_1(j) - (e_{j+s_h-t\delta_h(j)} + v_h + \delta_h(j)) \equiv 0 \pmod{p}.$$  \hspace{1cm} (36)

We now put $d_{i-1} = e_i - e_{i-1}$ and choose $e$ such that $d_0, d_1, \ldots, d_{t-3}$ is in an arithmetic progression with common distance $d \not\equiv 0 \pmod{p}$. As given in (30) we get

$$e_j = e_0 + j d_0 + \frac{j(j-1)}{2} d, \quad 0 \leq j \leq t - 2,$$

and hence

$$\left( [s_1 - s_h] - t(\delta_1(j) - \delta_h(j))] d_0 + [v_1 - v_h + \delta_1(j) - \delta_h(j)] + f(j) d \equiv 0 \pmod{p},$$

where

$$f(j) = \frac{1}{2} [(s_1 - t\delta_1(j))^2 - (s_h - t\delta_h(j))^2 + (s_h - s_1) + t(\delta_1(j) - \delta_h(j))]$$

$$+ s_1 - s_h - t (\delta_1(j) - \delta_h(j)) j.$$  \hspace{1cm} (37)
First, if
\[ s_1 - s_h \not\equiv t(\delta_1(j) - \delta_h(j)) \pmod{p}, \]
we have modulo \( p \) exactly one \( j \) with
\[ e_{j+\tau_1} - e_{j+\tau_h} \equiv 0 \pmod{p}, \]
and thus
\[ \max_{2 \leq h \leq k} F_h = O(1 + t/p). \]

Second, if
\[ s_1 - s_h \equiv t(\delta_1(j) - \delta_h(j)) \pmod{p}, \]
Equation (37) simplifies to
\[ v_1 - v_h + \delta_1(j) - \delta_h(j) \equiv 0 \pmod{p} \]
and there is no \( 0 \leq j < t \) with
\[ e_{j+\tau_1} - e_{j+\tau_h} \equiv 0 \pmod{p} \]
if
\[ v_1 - v_h \not\equiv \delta_h(j) - \delta_1(j) \pmod{p}. \]

Now we consider the case that \( p \geq 2t \) or \( p = t \). First we consider the case \( p \geq 2t \) and note that (38) is not possible in this case. Otherwise (38) and (35) imply
\[ -t < s_1 - s_h = \epsilon t + kp < t \]
for \( \epsilon \in \{-1, 0, 1\} \) and for some integer \( k \). This also implies
\[ 0 < |k| < 2t/p \leq 1 \]
and such a \( k \) does not exist.

For \( p = t \), (38) implies \( s_1 = s_h \) and (39) must hold since otherwise \( v_1 = v_h \) and thus \( \tau_1 = \tau_h \) which is a contradiction.

Consider the power correlation measure of order \( k \) with bound \( B = p \). Since
\[ 0 \leq \tau_1 < \cdots < \tau_k \leq p \], we have \( 0 < s_1, s_h \leq p \) and \( v_1 = v_h = 0 \). The case of
\[ s_1 - s_h \equiv t(\delta_1(j) - \delta_h(j)) \pmod{p} \quad \text{and} \quad v_1 - v_h \equiv \delta_h(j) - \delta_1(j) \pmod{p} \]
is not possible since otherwise $\tau_1 = \tau_h$ which is a contradiction. To obtain the result we apply the same bound as in (36) on $\Gamma_k'(y_r, B)$ without restriction on $p$ and $t$.

\textbf{Example 5.9.} Let us construct a family of 24 interleaved sequences by picking the two base sequences to be two Legendre sequences of periods 13 and 23, respectively. The bound we obtained in the previous theorem is $13 \cdot 23^{1/2} \log 23$. However, we can directly calculate the maximum correlation measure of order two for all interleaved sequences in this family. The maximum is 48 which is much smaller than the bound given in the above theorem.

Similarly, for every interleaved sequence in the family constructed from two Legendre sequences of periods, 11 and 7, respectively, the maximum correlation measure of order 2 is 19 which is also much smaller than the bound given in the above theorem.

\textbf{Remark 5.10.} Because of the limit of the current mathematical tools, mainly character sums, we can only estimate the upper bound for power correlation measure by the result shown in Theorem 5.8. We will look for alternative methods to improve the bound.

Theorem 5.8 shows that if we give a proper bound $B$ on the lags $\tau_1, \ldots, \tau_k$, the correlation measure of a small order $k \leq B + 1$ with lags bounded by $B$ is upper bounded by a non-trivial bound without restrictions on the periods of the two bases.

By Proposition 5.4 and Theorem 5.8, we derive in the next theorem a lower bound for the linear complexity of interleaved sequences in terms of the correlation measure of order $k$ with bound $B = p$.

\textbf{Theorem 5.11.} A lower bound for the linear complexity profile of the interleaved sequence $y_r \in \mathcal{Y}_a$ in Notation 5.5 holds:

$$\log(L(y_r, N)) \geq \min \left\{ p, \frac{N}{tp^{1/2} \log p + p} \right\}.$$ 

\textbf{Proof.} We write $L = L(y_r, N)$. In Part (2) of Theorem 5.8 we showed that

$$\Gamma_k'(y_r, p) \ll k(tp^{1/2} \log p + p).$$
By Proposition 5.4 we have
\[ L \geq N - \max_{1 \leq k \leq L+1} \Gamma'_k(y_r, p) \]
\[ L \geq N - \max_{1 \leq k \leq L+1} k(tp^{1/2} \log p + p) \]
\[ \geq N - L(tp^{1/2} \log p + p). \]

The lags are bounded by \( B = p \) and
\[ \min\{N - L(tp^{1/2} \log p + p), p\} = N - L(tp^{1/2} \log p + p) \]

Then by Proposition 5.4
\[ L(y_r, N) \geq \min \left\{ p, \frac{N}{tp^{1/2} \log p + p} \right\}. \]

\[ \square \]

This lower bound is higher than the one obtained in Theorem 5.2 when \( p > t \).

Now we study the correlation measure of order \( k \) (with lags bounded by \( B \)) for the binary case, i.e. when \( m = 2 \). The result in Theorem 5.8 is improved in this case. The improvement is that the correlation measure of order \( k \) is non-trivially upper bounded if \( k \) is odd and the interleaved sequences are associated with one Legendre sequence and any other binary sequence. Moreover, the non-trivial upper bound is also achieved for \( k = 2 \) if the interleaved sequence is associated with two Legendre sequences. We compare the results on the correlation measures with and without bound \( B \) on the lags.

The corollary below follows from Part (2) in Theorem 5.8.

**Corollary 5.12.** Let \( a \) be a Legendre sequence of period \( p \), where \( p \) is an odd prime, and \( b \) be any binary sequence of positive integer period \( t \). For \( y_r \in Y_e(a, b, c) \) the correlation measure of order \( k \) with bound \( B = p \) for \( k \leq B + 1 \) is bounded by
\[ C_k(y_r, p) \ll k(tp^{1/2} \log p + p). \]

**Theorem 5.13.** (Binary case) Let \( y_r \) be a sequence in the family of interleaved sequences associated with a Legendre sequence \( a \) of odd prime period \( p \), a binary sequence \( b \) of positive integer period \( t \) and a shift sequence \( e \) satisfying (30). The correlation measure of order \( k \) of \( y_r \in Y_e(a, b, c) \) is bounded by
\[ C_k(y_r) \ll k(tp^{1/2} \log p + p) \]
if one of the following conditions holds:

(i) \( k \) is odd;
(ii) \( k \) is even and \( p \geq 2t \) or \( p = t \).

**Proof.** We write

\[
S_M = \sum_{n=0}^{M-1} (-1)^{y_n+\tau_1+y_n+\tau_2+\cdots+y_n+\tau_k},
\]

where \( k \) is a positive integer such that \( 0 \leq \tau_1 < \cdots < \tau_k \leq pt - M \). Then \( C_k(y_r) \) can be expressed as \( C_k(y_r) = \max_{M,D} |S_M| \), where \( D = (\tau_1, \ldots, \tau_k) \) and the maximum is taken over all non-negative integers \( \tau_1 < \tau_2 < \cdots < \tau_k \). We get

\[
|S_M| \leq \sum_j \left| \sum_{i=0}^{j-1} \left( \frac{e_{j+\tau_1+i}}{p} \right) \cdots \left( \frac{e_{j+\tau_k+i}}{p} \right) \right| + O(k(p+t)),
\]

where the outer sum runs through all \( 0 \leq j < t \) with \( j \not\equiv t - 1 - \tau_h \pmod{t} \) for \( 1 \leq h \leq k \). The error term comes from the same argument as in the proof of Theorem 5.8.

If \( k \) is odd, the polynomial

\[
\prod_{h=1}^{k} (e_{j+\tau_h} + X)
\]

has odd degree \( k \) and cannot be a square. Hence, we can apply Theorem 2.42 to estimate the inner sum by

\[
O(kp^{1/2} \log p).
\]

If \( k \) is even, the result follows Theorem 5.8. \( \square \)

**Remark 5.14.** In the case \( p < t \) we do not have a suitable bound for the correlation measure of order \( k \) when \( k \) is even.

The lower bound is the same as that we obtained in Theorem 5.11.

**Corollary 5.15.** (Binary case) Let \( y_r \) be a sequence in the family of interleaved sequences associated with a Legendre sequence \( a \) of odd prime period \( p \), a binary sequence \( b \) of positive integer period \( t \) and a shift sequence \( e \) satisfying (30). The linear complexity profile of \( y_r \in Y_s(a,b,e) \) is bounded by

\[
L(y_r, N) \gg \min \left\{ p, \frac{N}{tp^{1/2} \log p + p} \right\}.
\]
The result in Theorem 5.13 can be improved if the two bases are both Legendre sequences. In this case, the correlation measure of order \( k = 2 \) is also bounded by the non-trivial bound without restriction on the periods of the two bases.

**Theorem 5.16.** If the two bases are Legendre sequences of period \( p \) and \( q \), where \( p \) and \( q \) are odd prime numbers, the correlation measure of order 2 of every interleaved sequence \( y_r \in Y_r \) associated with the two Legendre sequences is bounded by

\[
C_2(y_r) \ll qp^{1/2} \log p + p.
\]

**Proof.** We use the notation \( \tau_h = s_h + v_h q \) for \( h = 1, 2 \), where \( q \) plays the role of \( t \).

For \( k = 2 \), we may assume \( p < 2q \) and \( p \neq q \). Denote

\[
A = \{ j \in \mathbb{Z} \mid 0 \leq j < q - \max\{s_1, s_2\} \text{ or } q - \min\{s_1, s_2\} \leq j < q \}
\]

and

\[
B = \{ j \in \mathbb{Z} \mid q - \max\{s_1, s_2\} \leq j < q - \min\{s_1, s_2\} \}.
\]

If \( j \in A \), we have \( \delta_1(j) = \delta_2(j) \). If \( j \in B \), we have \( \delta_1(j) \neq \delta_2(j) \).

The remaining case

\[
s_1 - s_2 \equiv q(\delta_1(j) - \delta_2(j)) \pmod{p}
\]

and

\[
v_1 - v_2 \equiv (\delta_2(j) - \delta_1(j)) \pmod{p}
\]

is equivalent to

\[
s_1 \equiv s_2 \pmod{p}, \ s_1 \neq s_2 \text{ and } v_1 = v_2 \text{ for } j \in A
\]

and

\[
s_1 \equiv s_2 \pm q \pmod{p} \text{ and } v_1 = v_2 \mp 1 \text{ for } j \in B.
\]

First, if

\[
s_1 \equiv s_2 \pmod{p},
\]

we have

\[
e_{j+\tau_1} \equiv e_{j+\tau_2} \pmod{p} \text{ for all } j \in A
\]
but

\[ e_{j + \tau_1} \not\equiv e_{j + \tau_2} \pmod{p} \]

for all but \(O(1 + q/p)\) different \(j \in B\). Then

\[
\begin{align*}
|S_M| & \leq \sum_{j \in B} \left| \sum_{i=0}^{\left\lfloor \frac{M}{q} \right\rfloor - 1} \left( \frac{e_{j + \tau_1} + i}{p} \right) \left( \frac{e_{j + \tau_2} + i}{p} \right) \right| + \\
& \quad \sum_{i=0}^{\left\lfloor \frac{M}{q} \right\rfloor - 1} \left| \sum_{j \in A} \left( \frac{j + r + s_1}{q} \right) \left( \frac{j + r + s_2}{q} \right) \right| + O(p + q) \\
& \ll qp^{1/2} \log p + pq^{1/2} \log q \\
& \ll qp^{1/2} \log p, \text{ since } p < 2q.
\end{align*}
\]

Second, if \(s_1 - s_2 \equiv \pm q \pmod{p}\), we have

\[ e_{j + \tau_1} \equiv e_{j + \tau_2} \pmod{p} \]

for all \(j \in B\), but

\[ e_{j + \tau_1} \not\equiv e_{j + \tau_2} \pmod{p} \]

for all but \(O(1 + q/p)\) different \(j \in A\). Then the bound for \(|S_M|\) is the same as in Theorem 5.13.

\[
\text{\textbf{Remark 5.17.}} \text{ Let the two bases be Legendre sequences of periods } p \text{ and } q. \text{ If } q \text{ is large with respect to } p \text{ we can pick the following lags to show that the correlation measure of order } 4h \text{ is close to } pq \text{ (provided that } h \text{ is small with respect to } q):}
\]

\[
\begin{align*}
\tau_1, \tau_2, & \tau_3, \tau_4, \ldots, \tau_{4h-3}, \tau_{4h-2}, \tau_{4h-1}, \tau_{4h} \\
= & (0, p, q, p + q, \ldots, \\
& 2(h-1)q, 2(h-1)q + p, 2(h-1)q + q, 2(h-1)q + p + q),
\end{align*}
\]

with \(\tau_{4h} \leq pq - M\). We note that every four lags in a group are in the pattern of

\[
2(v-1)q, 2(v-1)q + p, 2(v-1)q + q, 2(v-1)q + p + q,
\]

with \(1 \leq v \leq h\).

For \(0 \leq j \leq q - p - 1\), by definition of \(e_j\) we have

\[
\left( \frac{e_{j + \tau_1} + i}{p} \right) \left( \frac{e_{j + p + \tau_1} + i}{p} \right) = 1
\]
and also, clearly,

\[
\left( \frac{j + r}{q} \right) \left( \frac{j + r + q}{q} \right) = 1.
\]

Hence, the contribution of these \( j \) to the correlation measure of order 4 is \((q-p)p + O(q)\). For the remaining \( j \) with \( q - p \leq j < q \) we can estimate the corresponding sums by \( O(p^{3/2} \log p) \).

**Example 5.18.** We estimate the correlation measure of order 4 and show that the upper bound can reach the trivial maximal magnitude. We take \( \tau_1 = 0, \tau_2 = p, \tau_3 = q \) and \( \tau_4 = p + q \).

- When \( 0 \leq j \leq q - p - 1 \),

\[
C_4(s_k) = \left| \sum_{j=0}^{q-p-1} \sum_{i=0}^{\lfloor \frac{M}{q} \rfloor - 1} \left( \frac{e_j + i}{p} \right) \left( \frac{j + k}{q} \right) \left( \frac{e_{j+p} + i}{p} \right) \left( \frac{j + k + p}{q} \right) \right| \\
= \left| \sum_{j=0}^{q-p-1} \sum_{i=0}^{\lfloor \frac{M}{q} \rfloor - 1} \left( \frac{e_j + i + 1}{p} \right) \left( \frac{j + k}{q} \right) \left( \frac{e_j + i + 1}{p} \right) \left( \frac{j + k + p}{q} \right) \right| \\
= \sum_{j=0}^{q-p-1} \sum_{i=0}^{\lfloor \frac{M}{q} \rfloor - 1} 1 = p(q - p).
\]
When \( q - p \leq j < q \),

\[
\begin{align*}
C_4(s_k) &= | \sum_{j=q-p}^{q-1} \sum_{i=0}^{\lfloor \frac{j}{q} \rfloor - 1} \left( \frac{e_j + i}{p} \right) \left( \frac{j + k}{q} \right) \left( \frac{e_{j+p-q} + i + 1}{p} \right) \left( \frac{j + k + p}{q} \right) \\
& \quad \left( \frac{e_{j+q} + i}{p} \right) \left( \frac{j + k + q}{q} \right) \left( \frac{e_{j+p-q} + i + 2}{p} \right) \left( \frac{j + k + p}{q} \right) | \\
& = | \sum_{j=q-p}^{q-1} \sum_{i=0}^{\lfloor \frac{j}{q} \rfloor - 1} \left( \frac{e_j + i + 1}{p} \right) \left( \frac{j + k}{q} \right) \left( \frac{e_{j+p-q} + i + 1}{p} \right) \left( \frac{j + k + q}{q} \right) \\
& \quad \left( \frac{e_j + i + 1}{p} \right) \left( \frac{j + k}{q} \right) \left( \frac{e_j + i}{p} \right) \left( \frac{j + k + q}{q} \right) | \\
& \leq p \cdot \sum_{i=0}^{\lfloor \frac{q}{p} \rfloor - 1} \left( \frac{e_j + i}{p} \right) \left( e_j + i \right) \left( \frac{e_j + i + 1}{p} \right) \left( \frac{e_j + i + 1}{p} \right) | \\
& < p \cdot p^{1/2} \log p.
\end{align*}
\]

\[ \square \]

**Remark 5.19.** If \( q \) is large with respect to \( p \) again, for all but \( O(p^2 + q) \) many \( n \), \( 0 \leq n = iq + j < pq \) (with \( 0 \leq j < q - p \) and \( 0 \leq i < p \)), the following recurrence relation holds

\[
y_n + y_{n+p} + y_{n+q} + y_{n+p+q} = 0.
\]

Equivalently, we have

\[
\left( \frac{(e_j + i)(e_{j+p} + i)(e_{j+q} + i)(e_{j+p+q} + i)}{p} \right) = 1
\]

and

\[
\left( \frac{(j + r)(j + p + r)(j + q + r)(j + p + q + r)}{q} \right) = 1,
\]

which may be considered a *weak point* of correlation measure of order \( k \) since we have shown in Example 5.18 that the correlation measure approaches the trivial bound.

**Remark 5.20.** If \( q \) is large with respect to \( p \), the correlation measure of order \( 4h + 2 \) of the sequences constructed from two Legendre sequences of odd prime
periods $p$ and $q$ are also close to $pq$ (provided that $h$ is small with respect to $q$). Here we can choose the first $4(h - 1)$ lags with the same pattern as in Remark 5.17 and the last six lags as in the pattern
\[0, p, q, 2p + q, 2q + p, 2q + 2p\]

We have compared the correlation measure of order $k$ for the family of binary interleaved Legendre sequences with arbitrary lags or with bounded lags. From Remark 5.17 we find that the correlation measure with arbitrary lags can approach the trivial bound. So it is necessary to select proper bound on the lags in order to achieve a non-trivial upper bound for correlation measures.

5.3. Nonlinearity of the boolean function derived from binary interleaved sequences

We study the boolean function derived from binary interleaved sequences and two cryptographic properties: sparsity and nonlinearity. We recall Definition 2.15 that every boolean function in $n$ variables maps an $n$–tuple vector to 0 or 1. A binary sequence can be connected with a boolean function using the following definition.

**Definition 5.21.** Given a binary sequence $(s(0),\ldots,s(T−1)) \in \{0,1\}^T$ of length $T$, a *Boolean function derived from a binary sequence* $B(u_1,\ldots,u_n)$ in $n$ variables is obtained through setting
\[B(u_1,\ldots,u_n) = s(h), \ h = 0,\ldots,T−1,\]
and
\[B(u_1,\ldots,u_n) = 0, \ h = T,\ldots,2^n−1,\]
where $2^{n−1} \leq T < 2^n$ for some $n$ and $h = u_1 \cdot 2^0 + u_2 \cdot 2^1 + \cdots + u_n \cdot 2^{n−1}$.

We recall that the *Fourier coefficients* of a Boolean function $B = B(u_1,\ldots,u_n)$ in $n$ variables are defined as
\[\hat{B}(l) = 2^{-n} \sum_{h \in \mathbb{F}_2^n} (-1)^{f(h)+<h,l>},\]
where $l, h = 0, \ldots, 2^n - 1$ and $< h, l >$ is the conventional inner product of the two vectors taking the coefficients from the binary representation $h = u_1 \cdot 2^0 + u_2 \cdot 2^1 + \cdots + u_n \cdot 2^{n-1}$.

The nonlinearity $\mathfrak{N}(B)$ is defined as

\[(40) \quad \mathfrak{N}(B) = 2^n - 2^{n-1} \max_{l \in \mathbb{F}_2} |\hat{B}(l)|.\]

The sparsity $\text{spr}(B)$ or algebraic thickness of $B(u_1, \ldots, u_n)$ is the minimal number of its nonzero coefficients in its algebraic normal form.

In this section we study two cryptographic properties, sparsity and nonlinearity, of the boolean function derived from a binary interleaved sequence.

We get a lower bound of the sparsity of the Boolean function derived from a binary sequence of any length. This generalizes the result in [69, Theorem 1] which gives the lower bound of the sparsity to a binary sequence of period equal to a power of 2.

**Theorem 5.22.** Let $B(u_1, \ldots, u_r)$ be the Boolean function derived from a binary sequence $s(h)$ of length $N$, where $2^{r-1} < N < 2^r$ for some integer $r$. Then we have the bound

\[
\text{spr}(B) \geq \frac{N}{2} \max_{1 \leq j \leq 2^{\text{spr}(B)}} C_j(s(h)),
\]

where $C_j(s(h))$ is the correlation measure of order $j$ of the binary sequence $s(h)$ for $0 \leq h < N$.

**Proof.** Define $r$ by the inequalities $2^{r-1} \leq N < 2^r$, i.e. $r = \lceil \log_2 N \rceil$. There are $r$ variables in the Boolean function derived from the binary sequence of length $N$ after padding zeros if needed. We note that there is some $k$ such that $2^{k-1} \leq \text{spr}(B) < 2^k$. For each $i = 0, \ldots, 2^k - 1$ we consider the function

\[B_i(V_1, \ldots, V_{r-k}) = B(V_1, \ldots, V_{r-k}, e_1, \ldots, e_k),\]

where $i = (e_1, \ldots, e_k)$ is the binary representation of $i$. Assume the total number of distinct monomials in $V_1, \ldots, V_{r-k}$ occurring in all these functions does not exceed $\text{spr}(B)$. Therefore, one can find a non-trivial linear combination

\[
\sum_{i=0}^{2^k-1} c_i B_i(V_1, \ldots, V_{r-k}) = 0, \quad c_0, \ldots, c_{2^k-1} \in \mathbb{F}_2,
\]
where at least one of $c_0, \ldots, c_{2^k - 1}$ is nonzero. By definition, we have $B(u_1, \ldots, u_r) = s_h$ where $(u_1, \ldots, u_r)$ is the binary representation of $h$. Therefore,

$$
\prod_{i=0}^{2^k-1} \left( (-1)^{s(y \cdot 2^k + i)} \right) c_i = (-1)^{\sum_{i=0}^{2^k-1} c_i B_i(v_1, \ldots, v_{r-k})} = 1
$$

for all $0 \leq y \leq 2^{r-k} - 1$ and $i = 0, \ldots, 2^k - 1$, where $(v_1, \ldots, v_{r-k})$ is the binary representation of $y$. Hence,

$$
2^{r-k} = \sum_{y=0}^{2^{r-k}-1} \prod_{i=0}^{2^k-1} \left( (-1)^{s(y \cdot 2^k + i)} \right) c_i \leq \max_{1 \leq j \leq 2^k} C_j(s(h)).
$$

Since

$$
\frac{1}{2} \frac{2^r}{\text{spr}(B)} \leq \frac{2^r}{2^k},
$$

we have

$$
\text{spr}(B) \geq \frac{2^{r-1}}{\max_{1 \leq j \leq 2^{\text{spr}(B)}} C_j(s(h))}.
$$

□

The nonlinearity gives the smallest possible Hamming distance between the vector of values of $B$ and the vector of values of a linear function in $n$ variables over $\mathbb{F}_2$.

**Theorem 5.23.** The nonlinearity $\mathfrak{N}(B)$ of a Boolean function $B$ derived from a binary sequence $s = s(h)$, as defined in Definition 5.21, satisfies

$$
\mathfrak{N}(B) \geq 2^{n-1} \left( 1 - 8^{1/4} 2^{-n/4} C_2(s)^{1/4} \right).
$$

**Proof.** By the definition of correlation measure of order 2 we have $C_2(s(h)) < 2^n$. We fix some $m < n$ and write $h, l \in \mathbb{F}_2^n$ as

$$
h = k + 2^m j \quad \text{and} \quad l = s + 2^m t,
$$

with $0 \leq k, s < 2^m$ and $0 \leq j, t < 2^{n-m}$. In particular,

$$
\langle h, l \rangle = \langle k, s \rangle + \langle j, t \rangle.
$$
Hence,

\[
|\hat{B}(l)| = |\hat{B}(s + 2^m t)| = |2^{m-n} \sum_{k=0}^{2^m-1} \sum_{j=0}^{2^{n-m}-1} (-1)^{B(k+2^m j)+<k,s> + <j,t>})| \leq 2^{m-n} \sum_{k=0}^{2^m-1} |\sum_{j=0}^{2^{n-m}-1} (-1)^{B(k+2^m j)+<k,s> + <j,t>}).
\]

We factor out \((-1)^{<k,s>}\) which is a constant in the inner sum and by the Cauchy-Schwarz inequality we get

\[
2^{2n-m} |\hat{B}(l)|^2 \leq \sum_{k=0}^{2^m-1} \left| \sum_{j=0}^{2^{n-m}-1} (-1)^{B(k+2^m j)+<j,t>} \right|^2 \leq \sum_{k=0}^{2^m-1} \left| \sum_{j_1,j_2=0}^{2^{n-m}-1} (-1)^{B(k+2^m j_1)+B(k+2^m j_2)+<j_1,t> + <j_2,t>} \right|
\]

\[
\leq \sum_{j_1,j_2=0}^{2^{n-m}-1} \left| \sum_{k=0}^{2^m-1} (-1)^{B(k+2^m j_1)+B(k+2^m j_2)} \right| \leq \sum_{j_1,j_2=0}^{2^{n-m}-1} \left| \sum_{k=0}^{2^m-1} (-1)^{s(k+2^m j_1)+s(k+2^m j_2)} \right|.
\]

For \(j_1 \neq j_2\) the inner sum can be estimated by \(C_2(s)\) in Theorem 5.16; otherwise the inner sum is equal to \(2^m\). Hence

\[
|\hat{B}(l)|^2 \leq 2^{m-2n} \left(2^{n-m}2^m + 2^{2(n-m)}C_2(s)\right) = 2^{m-n} + 2^{-m}C_2(s).
\]

Then we pick \(m\) satisfying \(2^m \leq \sqrt{2}C_2(s) < 2^{m+1}\) so that we can conclude the proof.

This result is derived from unpublished notes [7]. In [7] the authors considered the nonlinearity of the Boolean function derived from binary sequences after cutting some elements to shorten the length to a power of 2. Here we consider extending the length by padding zeros to the sequence. We hope to find some other proper
paddings besides zero padding to significantly improve the bounds on sparsity and nonlinearity.
CHAPTER 6

Zero correlation zone of interleaved polyphase sequences

A polyphase sequence is a sequence of complex numbers with unit magnitude. For technical reasons polyphase sequences with a small phase alphabet are often required. We extend polyphase sequence to complex-valued sequences. The elements are taken from complex numbers with unit magnitude and zero.

In this chapter we construct a class of complex-valued sequences with almost perfect autocorrelation, in Section 6.1, using maximal length sequences and multiplicative characters. The sequence is of \( m + 1 \) symbols, i.e. elements taken from \( \{0, 1, \omega, \ldots, \omega^{m-1}\} \), where \( \omega \) is a primitive \( m \)th root of unity, by transforming a maximal length sequence through a multiplicative character of order \( m \).

In quasi-synchronous CDMA systems, zero correlation zone (ZCZ) signal sets have been employed as spreading sequences due to their good correlation property within the allowed time delay. To accommodate many multiple access users and to ease the synchronization requirement, it is desirable to construct a ZCZ signal set of a given length having both the set size and the ZCZ length as large as possible. However, the Tang-Fan-Matsufuji bound [81] implies that the product of the set size and the ZCZ length is no greater than the period of every sequence. As a result, it is interesting to construct optimal ZCZ sequence sets with respect to the Tang-Fan-Matsufuji bound.

In Section 6.2 we give the interleaved algorithm for complex-valued sequences. Then in Section 6.3 we interleave two almost perfect sequences defined in Section 6.1 to construct a zero correlation zone (ZCZ) signal set and show that the signal set reaches the optimal bound as the period approaches infinity.
6.1. A new class of almost perfect sequences

We denote by $\Omega_n$ the set of all the $n$th roots of unity. We work with complex-valued sequences that are defined using multiplicative characters over finite fields. Let $k$ be an integer. The co-domain of the multiplicative character of order $m$ is $\Omega_m := \{1, \omega, \omega^2, \ldots, \omega^{m-1}\}$, where $\omega$ is a primitive $k$th root of unity. It is natural to extend the co-domain to $\Omega_m^+ := \{1, \omega, \omega^2, \ldots, \omega^{m-1}\} \cup \{0\}$ by including 0; we can see the transmission of a zero as a “pause” in every period.

We define a sequence by transforming a maximal length sequence to a complex-valued sequence of $m + 1$ symbols by a multiplicative character of order $m$. Lee sequences are defined over $\Omega_{\ell-1}^+$ where $\ell$ is a prime or a prime power. We define a new class of sequences over $\Omega_m^+$ where $m$ is an integer and $m \mid \ell - 1$.

**Definition 6.1.** Let a maximal length sequence of period $\ell^{n_1} - 1$ over $\mathbb{F}_\ell$ be given by

$$u(t) = Tr_{1}^{n_1}(\alpha^{st}), \quad t = 0, 1, \ldots, \ell^{n_1} - 2$$

where $\gcd(s, \ell^{n_1} - 1) = 1$. Define the sequence $b(t)$ over $\Omega_m^+$ by

$$b(t) = \begin{cases} 
\psi(u(t)), & \text{if } u(t) \neq 0, \\
0, & \text{otherwise,}
\end{cases}$$

for $t = 0, 1, \ldots, \ell^{n_1} - 2$, where $\psi$ is a multiplicative character of order $m$ mapping $\mathbb{F}_\ell$ into $\Omega_m^+ = \{0, 1, \omega, \ldots, \omega^{m-1}\}$ by noting that $0 \in \mathbb{F}_\ell$ maps to $0 \in \mathbb{C}$, where $\omega$ is a primitive $m$th root of unity and $m \mid \ell - 1$.

We show later that the newly defined sequence has longer period than the generalized Lee sequence (see Definition 2.37 and Definition 2.38) and the zero autocorrelation can be achieved almost everywhere except at some points. As usual, by period we mean the smallest positive period.

**Theorem 6.2.** Let us assume that a multiplicative character $\psi$ has order $m$, with $m \mid (\ell - 1)$. Let $n_1$ be a positive integer. Let $d = (\ell^{n_1} - 1)/(\ell - 1)$. The sequence $b(t)$ as given in Definition 6.1 has period $T = dm$ and has zero autocorrelation at $\tau \neq 0 \pmod{d}$.

**Proof.** Let $T$ denote the period of $b(t)$. Obviously $T$ must divide the period $\ell^{n_1} - 1$ of the maximal length sequence $u(t)$. We write the periodic autocorrelation
function as

\[ C_b(\tau) = \frac{T}{\ell_{n_1} - 1} \sum_{t=0}^{\ell_{n_1} - 2} b(t)b(t + \tau) \]

\[ = \frac{T}{\ell_{n_1} - 1} \sum_{t=0}^{\ell_{n_1} - 2} \psi(u(t))\overline{\psi(u(t + \tau))} \]

where the terms in the last sum for which \( u(t + \tau)u(t) = 0 \) are discarded since they result in zero contributions.

1. When \( \tau \not\equiv 0 \pmod{d} \), by the 2-tuple balance property shown by Zieler [96] for the maximal length sequence \( u(t) \), every \((u(t), u(t + \tau))\) takes on any given pair \((x, y) \in \mathbb{F}_2^2 \setminus \{(0, 0)\}\) exactly \( \ell_{n_1} - 2 \) times. In this way \( x \) and \( y \) are independent and we get

\[ C_b(\tau) = \ell_{n_1} - 2 \cdot \frac{T}{\ell_{n_1} - 1} \sum_{x \in \mathbb{F}_2^*} \sum_{y=1}^{\ell - 1} \psi(x)\overline{\psi(y)} \]

\[ = \ell_{n_1} - 2 \cdot \frac{T}{\ell_{n_1} - 1} \sum_{x \in \mathbb{F}_2^*} \psi(x) \cdot \sum_{y=1}^{\ell - 1} \psi(y) = 0 \]

due to the fact that \( \sum_{y=1}^{\ell - 1} \overline{\psi(y)} = 0 \) for any multiplicative character \( \psi \).

2. When \( \tau = dj \) for some \( j \), also by the 2-tuple balance property, the pair \((u(t), u(t + \tau))\) appears \( \ell_{n_1} \) times for every \( u(t) \neq 0 \) and the pair \((0, 0)\) appears \( \ell_{n_1} - 1 \) times, where \( \alpha \) is a primitive element of \( \mathbb{F}_{\ell_{n_1}} \) and the integer \( s \) satisfies \( \gcd(s, \ell_{n_1} - 1) = 1 \). Then

\[ C_b(dj) = \ell_{n_1} \cdot \frac{T}{\ell_{n_1} - 1} \sum_{x \in \mathbb{F}_2^*} \psi(x)\overline{\psi(\alpha^{sdj}x)} \]

\[ = \ell_{n_1} \cdot \overline{\psi(\alpha^{sdj})} \frac{T}{\ell_{n_1} - 1} \sum_{x \in \mathbb{F}_2^*} \psi(x)\overline{\psi(x)} \]

\[ = \overline{\psi(\alpha^{sdj})}\ell_{n_1} - 1(\ell - 1) \]

since \( \sum_{x \in \mathbb{F}_2^*} \psi(x)\overline{\psi(x)} = \sum_{x \in \mathbb{F}_2^*} 1 = \ell - 1 \).

We state that the period of \( b(t) \) is \( d\ell \). Since \( \alpha \in \mathbb{F}_{\ell_{n_1}} \) is a primitive element and \( \alpha^d \in \mathbb{F}_\ell \), where \( d = \ell_{n_1} - 1 \), we have \( \text{Tr}_{n_1}(\alpha^d) = \alpha^d \) and we write \( \psi(u(t + d\ell)) \) in

\[ \psi(\text{Tr}_{n_1}(\alpha^{st}\alpha^{sdm})) = \psi(\alpha^{sdm}\text{Tr}_{n_1}(\alpha^{st})) = \psi(\alpha^{sdm})\psi(\text{Tr}_{n_1}(\alpha^{st})). \]
The above is just $\psi(u(t))$, since the order of $\psi$ is $m$. Therefore the period of $b(t)$ is a divisor of $dm$.

From (42), if $j$ is a multiple of $m$, we have $\psi(\alpha^m) = 1$ since $\psi$ is of order $m$. Therefore $C_\alpha(dj)$ is a function of $\tau$ that repeats itself with period $dm$. Hence, the period $T$ of $b(t)$ is exactly $dm$. 

□

**Example 6.3.** Let $\ell = 3$ and consider $f(x) = x^3 - 2x - 2$ which is a primitive polynomial over $\mathbb{F}_3$. Let $\alpha$ be a root of $f$, and therefore $\alpha$ is a primitive root of $\mathbb{F}_3$. Then $u_1(t) = \text{Tr}_1^3(\alpha^t) = (0, 0, 1, 1, 0, 2, 1, 2, 1, 0, 0, 2, 2, 2, 0, 1, 2, 1, 2, 0, 2)$ is a maximal length sequence of period $3^3 - 1$. By Definition 6.1 we define the sequence $a(t) = e^{2\pi i \text{ind}_{\beta^t u_1}(t)}$, by setting the primitive element 2 of $\mathbb{F}_3$ as $\beta$, which is

$$a(t) = (0, 0, +, +, 0, -, +, +, +, 0, +, +, 0, 0, -, -, -, 0, +, - , -, +, - , 0, -).$$

This sequence has almost perfect autocorrelation. The autocorrelation is

$$C_{a(t)}(\tau) = 0 \text{ for } \tau \not\equiv 0 \pmod{13}$$

and

$$C_{a(t)} = -18 \text{ or } 26 \text{ for } \tau \equiv 0 \pmod{13}.$$

The Lee sequence in [54] and the generalized Lee sequence in [55] were considered over a prime field $\mathbb{F}_p$, i.e. for the special case $\ell = p$. However, the result is also applicable to the (generalized) Lee sequence over $\mathbb{F}_\ell$ with $\ell$ a prime power (see Definition 2.37 and 2.38). It has been shown [54] that the period of the Lee sequence over $\mathbb{F}_\ell$ is $d = (\ell^{n_1} - 1)/(\ell - 1)$ with the notation as in Definition 2.37. The following theorem shows that Lee sequences have zero autocorrelation at all nonzero shifts in one period.

**Theorem 6.4.** [54] Let $n \equiv -2 \pmod{\ell}$. With the notation above in Definition 2.37, if the condition $d + vs \equiv 0 \pmod{\ell - 1}$ is satisfied for some integer $v$, then the Lee sequence $(a(t))_{t \geq 0}$ defined in Definition 2.37 has zero periodic autocorrelation, that is, for $\tau \not\equiv 0 \pmod{d}$, $C_a(\tau) = \sum_{t=0}^{d-1} a(t + \tau)a(t) = 0$. 

We compare the Lee sequence in Definition 2.38 and the sequence in Definition 6.1 in the following aspects. The Lee sequence has period \( d = (\ell^n - 1)/(\ell - 1) \), alphabet size \( \ell - 1 \) and perfect autocorrelation. The sequence we defined in Definition 6.1 has larger period \( m d \), where \( m \mid (\ell - 1) \), smaller alphabet size \( m \) and zero autocorrelation everywhere except at the point \( \tau \equiv 0 \) (mod \( d \)). Besides zero autocorrelation, in practice sequences with a large periods and small alphabet size are required.

6.2. Interleaved construction of polyphase sequences

Polyphase sequences are complex valued. Therefore the interleaved algorithm given in Section 3.4 needs to be modified. The following algorithm is a variant from [31] for complex-valued sequences.

**Algorithm 6.5.** Let \( s \) and \( t \) be two positive integers. Suppose that \( a = a(i)_{0 \leq i \leq s-1} \) and \( b = b(i)_{0 \leq i \leq t-1} \) are two complex-valued sequences of periods \( s \) and \( t \), respectively.

1. Choose \( e = (e_0, \ldots, e_{t-1}) \) as the shift sequence for which the first \( t - 1 \) elements form a Costas sequence with \( e_j \leq s \) for \( j = 0, \ldots, t - 2 \) and set \( e_{t-1} = \infty \).
2. Construct an interleaved sequence \( u = (u_0, \ldots, u_{st-1}) \). Arrange \( u \) as an \( s \times t \) matrix, called the matrix form of \( u \), as follows: the \( j \)th column in the matrix is given by \( L^j(a) \), where \( L^j \) is the left shift operator by shifting \( e_j \) units.
3. For \( 0 \leq i < st - 1 \), \( 0 \leq j \leq t \), define \( s_j = (s_{j,0}, \ldots, s_{j,st-1}) \) as
   \[
   s_{j,i} = \begin{cases} 
   u_i \cdot b_{j+i}, & 0 \leq j \leq t - 1, \\
   u_i, & j = t.
   \end{cases}
   \]
4. Define the family of sequences \( \mathfrak{S} = \mathfrak{S}(a, b, e) \) as \( \mathfrak{S} = \{ s_j \mid j = 0, 1, \ldots, t \} \), where \( a \) is the first base sequence, \( e \) is the shift sequence, and \( b \) is the second base sequence.

We need to rewrite Proposition 4.1 to consider inner products over a complex vector space. Let \( a \) and \( b \) be two sequences over \( \mathbb{C}^T \) of period \( T \). For \( \tau \geq 0 \), we have that for \( c, d \in \mathbb{C} \), \( \langle a \cdot c, b \cdot d \rangle = c\overline{d} \langle a, b \rangle \).
When $\tau > t$ the zero correlation of the interleaved sequence in Algorithm 6.5 can be destroyed. A discussion on the upper bound of this correlation for the case $\tau > t$ is given in [31]; see also [35]. Therefore, we restrict to $0 < |\tau| \leq t - 1$ and modify Lemma 4.6 to the following new lemma.

**Lemma 6.6.** For $0 \leq h, k \leq t$, suppose that $s_h$ and $s_k$ are two complex-valued sequences constructed using Algorithm 6.5. If $0 < |\tau| \leq t - 1$, then the correlation function between $s_h$ and $s_k$ is

\[
C_{h,k}(\tau) = \begin{cases} 
\sum_{j=0}^{t-1} b(h + j)\overline{b(k + \tau + j)}C_\alpha(e_{j+\tau} - e_j), & 0 \leq h, k \leq t - 1, \\
\sum_{j=0}^{t-1} b(k + \tau + j)C_\alpha(e_{j+\tau} - e_j), & h = t.
\end{cases}
\]

**Lemma 6.7.** For $0 \leq h, k \leq w$, suppose that $s_h$ and $s_k$ are two sequences in Algorithm 6.5. If $\tau = \tau_1 t + \tau_0$, $0 \leq \tau_0 < t$, $0 \leq \tau_1 < s$, then the correlation function between $s_h$ and $s_k$ is

\[
C_{h,k}(\tau) = \begin{cases} 
\sum_{j=0}^{t-1} b(h + j)\overline{b(k + \tau_0 + j)}C_\alpha(e_{j+\tau_0} - e_j + \tau_1), & 0 \leq h, k \leq t - 1, \\
\sum_{j=0}^{t-1} b(k + \tau_0 + j)C_\alpha(e_{j+\tau_0} - e_j + \tau_1), & h = t.
\end{cases}
\]

Algorithm 3.9 and the results on interleaved sequences obtained from Section 3.4 to Chapter 5 are for $m$-ary interleaved sequences. For the shift sequence in Algorithm 3.9 we put $d_{i-1} = e_i - e_{i-1}$, and then we choose $g$ such that $d_0, d_1, \ldots, d_{t-3}$ is in an arithmetic progression with common distance $d \not\equiv 0 \pmod{s}$, where $s$ and $t$ are the periods of the two base sequences, respectively.

Algorithm 6.5 is designed for complex-valued sequences. To get large zero-correlation zone we employ a Costas sequence as the shift sequence.

### 6.3. A Zero-Correlation Zone (ZCZ) signal set (family)

In this section we construct an interleaved complex-valued sequence family using the sequence defined in Definition 6.1 under the interleaved algorithm for complex-valued sequences in Algorithm 6.5. To achieve zero correlation, we require the following conditions.

1. The period of the first base is greater than or equal to that of the second base.
2. Welch Costas sequences are employed in the algorithm in [31]; we employ Lempel-Golomb Costas sequences or other kinds of Costas sequences as the shift sequence. We show that these two changes are crucial to achieve zero correlation zone in the constructed family of sequences.

**Construction 6.8.** Let \( \alpha \) be a primitive root of \( \mathbb{F}_{\ell^m} \), and \( \beta \) be a primitive root of \( \mathbb{F}_{\ell^{n_2}} \), for \( n_1 \geq n_2 \). We define two maximal length sequences of period \( \ell^{n_1} - 1 \) and \( \ell^{n_2} - 1 \), respectively, as

\[
u_1 = u_1(t) = T_1^{n_1}(\alpha^{s_1 t}), \text{ for } t = 0, 1, \ldots, \ell^{n_1} - 2, \text{ gcd}(s_1, \ell^{n_1} - 1) = 1,
\]

and

\[
u_2 = u_2(t) = T_2^{n_2}(\beta^{s_2 t}) \text{ for } t = 0, 1, \ldots, \ell^{n_2} - 2, \text{ gcd}(s_2, \ell^{n_2} - 1) = 1.
\]

For simplicity, we denote

\[d_1 = \frac{\ell^{n_1} - 1}{\ell - 1}, \quad d_2 = \frac{\ell^{n_2} - 1}{\ell - 1}, \quad T_1 = d_1 m, \quad T_2 = d_2 m.
\]

Let \( \psi : \mathbb{F}_\ell^* \to \Omega_m \) be a multiplicative character of order \( m \) mapping \( \mathbb{F}_\ell \) to \( \Omega_m^+ \) by noting that zero maps to zero. Let \( a = a(t) \) and \( b = b(t) \) be two sequences over \( \Omega_m^+ \) as in Definition 6.1

\[a(t) = \psi(u_1(t)) \text{ and } b(t) = \psi(u_2(t)).\]

By Theorem 6.2 the periods of \( a(t) \) and \( b(t) \) are \( T_1 \) and \( T_2 \), respectively.

Take a Costas sequence \( \epsilon = (c_0, \ldots, c_{R-1}) \) of period \( R \) with \( 0 \leq c_t < T_1 \). Let \( R \) satisfy

\[(44) \qquad T_2 \leq R < d_1.\]

Define the shift sequence \( \epsilon = (e_0, \ldots, e_{T_2-1}) \) of period \( T_2 \) where \( e_j = c_j \) for \( 0 \leq j \leq T_2 - 2 \) and \( e_{T_2-1} = \infty \). We use \( a(t) \) and \( b(t) \) in Algorithm 6.5 to construct a family \( \mathcal{S} = \mathcal{S}(a(t), b(t), \epsilon) \) of \( T_2 + 1 \) interleaved sequences of period \( T_1 T_2 \).

**Remark 6.9.** The extended \( \epsilon \) sequence was introduced in Remark 4.2 in Chapter 4 when the range of \( j + \tau \) goes beyond the period of \( \epsilon \). However, in this chapter the shift sequence is only taken from part of the Costas sequence since \( T_2 \leq R \). Moreover, the extended shift sequence is defined as \( e_{j+\tau} = c_{j+\tau} \) if \( R > j + \tau \geq T_2 \).
Recall the definition of the length of ZCZ, denoted as $L_{zc}$, in Definition 2.7 in Chapter 2.

**Theorem 6.10.** The length of ZCZ of the family in Construction 6.8 is

$$L_{zc} = \min\{T_2, R - T_2\}.$$ 

**Proof.** If $R - T_2 > \tau \geq 0$ we have that $e_j$ and $e_{j+\tau}$ are elements in the Costas sequence $c$, all whose elements are distinct, that is,

$$e_j \not\equiv e_{j+\tau} \pmod{R} \quad \text{if} \quad \tau \not\equiv 0 \pmod{R}.$$ 

We consider two cases for $|\tau| < \min\{T_2, R - T_2\}$:

1. If $e_{j+\tau} - e_j \not\equiv 0 \pmod{d_1}$ we have $C_a(e_{j+\tau} - e_j) = 0$ and thus $C_{h,k}(\tau) = 0$ for fixed $h, k, \tau$. This is also true in the autocorrelation case, that is when $h = k$. Since

$$|e_{j+\tau} - e_j| < \min\{T_2, R - T_2\} < R < d_1,$$

then $e_{j+\tau} - e_j \not\equiv 0 \pmod{d_1}$ holds for all $|\tau| < R - T_2$ but $\tau = 0$.

2. If $\tau = 0$, it follows $e_{j+\tau} - e_j = 0$. We use the 2-tuple balance property [96] for the maximal length sequence $u_2$ to conclude that $(u_2(h+j+\tau), u_2(k+j))$ takes on any given pair $(x, y) \in F^2_{\ell n_2} \setminus \{(0,0)\}$ exactly $\ell n_2 - 2$ times for fixed $h, k, \tau$. We can write (43) for $0 \leq h \leq T_2 - 1$ as

$$C_{h,k}(\tau) = \sum_{j=0}^{T_2-1} \psi(u_2(k+j)) \overline{\psi(u_2(h+j+\tau))} \cdot C_a(0)$$

$$= T_1 \cdot \sum_{j=0}^{T_2-1} \psi(u_2(k+j)) \overline{\psi(u_2(h+j+\tau))}$$

$$= T_1 \frac{T_2}{\ell n_2 - 1} \cdot \sum_{j=0, u_2(k+j) \neq 0, u_2(h+j+\tau) \neq 0}^{\ell n_2 - 1} \psi(u_2(k+j)) \overline{\psi(u_2(h+j+\tau))}$$

$$= T_1 T_2 \frac{\ell n_2}{\ell n_2 - 1} \sum_{x \in \mathbb{F}^2_{\ell n_2}} \psi(x) \overline{\psi(y)}$$

$$= T_1 T_2 \frac{\ell n_2}{\ell n_2 - 1} \sum_{x \in \mathbb{F}^2_{\ell n_2}} \psi(x) \sum_{y=1}^{\ell n_2 - 1} \psi(y) = 0.$$ 

Similarly, for (43) with $h = T_2$ we use the balance property of maximal length sequences to show that the cross correlation is zero as well. The
balance property implies that every nonzero element appears exactly $\ell^{n_2-1}$ times in one period. We have
\[
C_{h,k} = \sum_{j=0}^{T_2-1} \psi(u_2(k + j + \tau)) \cdot C_a(0)
\]
\[
= \frac{T_1 T_2}{\ell^{n_2} - 1} \ell^{n_1-1} \sum_{x \in \mathbb{F}_q^*} \psi(x) = 0.
\]
Therefore, the cross correlation of any two interleaved sequences and the nontrivial autocorrelation of every sequence in Construction 6.8 is zero for any $|\tau| < \min\{T_2, R - T_2\}$.

\begin{proof}\end{proof}

In order to get a larger window for the zero correlation zone, it is desirable to choose $R$ as large as possible within the range of $T_2$ and $d_1$. In many cases, the parameter $R$ is related to a prime number or prime power. For example if the Costas sequence in Construction 6.8 is a Welch Costas sequence, then $R = p - 1$ for a prime $p$. If it is a Lempel-Golomb Costas sequence, then $r$ is a prime power and $R = r - 2$.

We want a large $\min\{T_2, R - T_2\}$, so picking $T_2$ close to $R/2$ will optimize the minimum, that is, be away from zero.

**Example 6.11.** Take a Lempel-Golomb Costas sequence of period $R = \ell^{n_1-1} - 2$ over $\mathbb{F}_{\ell^{n_1-1}}$. Let $\mu, \eta$ be two primitive elements of $\mathbb{F}_{\ell^{n_1-1}}$. Set the shift sequence to
\[
e = (\log_\eta(1 - \mu^1), \log_\eta(1 - \mu^2), \ldots, \log_\eta(1 - \mu^{(\ell^{n_2-1})m-1}), \infty),
\]
with elements $0 < e_j < \ell^{n_1-1}$ for $j = 1, \ldots, \frac{(\ell^{n_2-1})m}{\ell - 1} - 1$. If $m$, the order of $\psi$ given in Construction 6.8, is a proper divisor of $\ell - 1$, the length of ZCZ of the family $\mathcal{S}$ of interleaved sequences in Construction 6.8 is
\[
L_{zcz} = T_2 = \frac{(\ell^{n_2} - 1)m}{\ell - 1}.
\]
The largest $T_2$ is obtained at $n_2 = n_1 - 1$ and $m = \frac{\ell - 1}{2}$. In this case the length of ZCZ is
\[
L_{zcz} = T_2 = \frac{\ell^{n_1-1} - 1}{2}.
\]
Remark 6.12. In [81], Tang, Fan and Matsufuji established a lower bound on LCZ and ZCZ sequence sets. Denote the size of sequence set to be $s$, the period of every sequence in the family to be $v$ and the length of ZCZ to be $L_{zcz}$. Then the Tang-Fan-Matsufuji bound is

$$s \leq \frac{v}{L_{zcz}} - 1. \quad (46)$$

Example 6.11 gives a $(T_1, T_2 + 1, T_2)$-ZCZ signal set, where $T_1 = \frac{\ell^{n_1} - 1}{2}$, $L_{zcz} = T_2 = \frac{\ell^{n_1} - 1}{n_1 - 1}$, the period of every sequence is $T_1 T_2$ and the number of sequences in the set is $T_2 + 1$. The maximum length of the zero-correlation zone under Construction 6.5 is $L_{zcz} = T_2 = \frac{\ell^{n_1} - 1}{2}$. The Tang-Fan-Matsufuji bound of the length of ZCZ in this case is

$$\min\{T_2, R - T_2\} - 1 = T_1 - 1 = \frac{\ell^{n_1} - 3}{2}. \quad (47)$$

The ratio of the largest length $L_{zcz}$ we get in (45) to the optimal length (47) is

$$\frac{\ell^{n_1} - 1}{\ell^{n_1} - 3}.$$  

Currently the best $L_{zcz}$-to-optimal ratio we get is $1/3$ when $\ell = 3$ and $n_1 = 2$. In the future we consider to reduce the number of columns of the matrix form given in Algorithm 6.5 to get a smaller period for the interleaved sequences. The reduced period would make (47) become smaller and so increase the $L_{zcz}$-to-optimal ratio.

The autocorrelation may be nonzero for $\tau > T_2$. Let us see an example.

Example 6.13. Let the shift sequence $e = (\alpha^1, \ldots, \alpha^{T_2})$ shown in Construction 6.8, where $\alpha$ is a primitive element of $F_p$ with an odd prime $p$. Let $T_1 = \frac{(p^{n_1} - 1)m}{p - 1}$ and $T_2 = \frac{(p^{n_2} - 1)m}{p - 1}$. When $\tau > T_2$, we consider Lemma 6.7. Let $\tau = \tau_1 T_2 + \tau_0$. If $e_{j + \tau_0} - e_j + \tau_1 \equiv 0 \pmod{p}$, we have

$$\alpha^{j+\tau_0} - \alpha^j + \tau_1 \equiv 0 \pmod{p}$$
$$\alpha^j (\alpha^{\tau_0} - 1) + \tau_1 \equiv 0 \pmod{p}$$
$$\alpha^j \equiv -\tau_1 (\alpha^{\tau_0} - 1)^{-1} \pmod{p}.$$  

We can always find a solution for $j$ for the equation $e_{j+\tau_0} - e_j + \tau_1 \equiv 0 \pmod{p}$ since $\alpha$ is a primitive element of $F_p$. One of the solutions satisfies $e_{j + \tau_0} - e_j + \tau_1 = 0$. This implies for some $j$ the autocorrelation $C_2(e_{j + \tau_0} - e_j + \tau_1) \neq 0$.
and it contributes a nonzero summand in the sum in Lemma 6.7. Therefore, the correlation of interleaved sequences is nonzero for \( \tau > T_2 \).
A special case of the interleaved construction occurs when the matrix form of the interleaved sequences is a column vector and the two base sequences are the same. If each time we shift the base sequence and we add to itself, then this is the shift-addition construction. Motivated from the shift-and-add property possessed by maximal length sequences (see Section 2.2), we consider shift-addition construction in this chapter. We still consider the use of the sequences defined from multiplicative characters, for example, two-prime generator of order 2 sequences. Under such shift-addition construction we can get a family of sequences of the same period as the original sequence.

Several authors construct new sequences with shift-and-addition structure and study their correlation magnitudes. Zhang and Zhou [94] construct binary sequences by using the shift-and-addition of a Legendre sequence of prime period $T = p$. Rushanan [73] applies the Weil bound to prove that the correlation magnitude of the sequences in [94] are bounded by $2\sqrt{p} + 5$. Han and Yang [33] then generalize the construction for $M$-ary sequence families from the shift and addition of power residue sequences, where $M$ is an integer. Also, Kim et al. [48] construct $M$-ary sequence families from the shift and addition of Sidelnikov sequences. Recently, Han and Yang [34] summarize the known constructions, and using the same shift-and-addition technique, they further construct $M$-ary sequence families with larger size than the known ones, but the same maximum correlation magnitude.

In this chapter we construct a new family using two-prime generator sequences of order two under the shift-and-addition construction. The maximal magnitude of this family is low and $9\sqrt{pq} + 4\left(\frac{1}{q} + \frac{1}{p}\right)$ as shown in Theorem 7.7, where $pq$ is the period of every sequence in the family.
7.1. Shift and add property and shift-addition construction

A sequence of elements from a finite field satisfies the *shift and add property* (SAP) if the component-wise sum of any two shifts of the sequence is either a shift of the sequence or the all zero sequence. If \( \mathbf{a} \) is a sequence possessing SAP, then there exist two shifts \( \tau_1 \) and \( \tau_2 \) such that \( L^{\tau_1}(\mathbf{a}) + L^{\tau_2}(\mathbf{a}) = 0 \).

Zierler [96, Theorem 6], gave a classification of periodic sequences with the SAP; he stated that a sequence \( \mathbf{s} \) had the SAP if and only if \( \mathbf{s} \) is a maximal length sequence. Unfortunately, his proof and results contained some errors, as was pointed out by Gong, Di Porto and Wolfowicz [30] where they give examples of sequences with the SAP which are not maximal length sequences. Zierler’s conclusion is true for sequences with SAP over a prime field \( \mathbb{F}_p \), where \( p \) is a prime.

**Theorem 7.1.** [3] If \( \mathbf{s} \) is a periodic sequence of elements from a finite field of prime order, then \( \mathbf{s} \) has the shift and add property if and only if \( \mathbf{s} \) is a maximal length sequence.

Next we give the shift-addition construction.

**Construction 7.2.** Let \( \mathbf{s} = s(t) \) for \( 0 \leq t \leq T - 1 \) be an \( m \)-ary sequence of period \( T \). Construct a family of sequences as below.

\[
\mathcal{I}_s = \{ cs(t) \mid 0 \leq c \leq m - 1 \},
\]

\[
\mathcal{A}_s^{\delta} = \{ c_0 \mathbf{s} + c_1 L^l(\mathbf{s}) \mid 0 \leq l \leq \lfloor T/2 \rfloor, \ c_0 + c_1 \equiv \delta \pmod{m} \},
\]

where \( \delta \) is a constant in \( 0 \leq \delta \leq m - 1 \) and \( 1 \leq c_0, c_1 \leq m - 1 \). The constructed family \( \mathcal{G}_s = \mathcal{I}_s \cup \mathcal{A}_s^{\delta} \) has period \( T \), with all its sequences cyclically distinct.

In the next section, we study the shift-adding family from two-prime generator sequences of order 2.

7.2. Shift-addition family from \( m \)-ary power residue sequences

The results in this section are given by Yu and Gong in [93]. Let \( \mathbf{s} = (s_i) \) be an \( m \)-ary \( p \)-periodic power residue sequence of the finite field \( \mathbb{F}_p \) with \( p \) an odd prime and \( m \) a divisor of \( p - 1 \). Let \( \alpha \) be a primitive root of \( \mathbb{F}_p \). Then \( \mathbf{s} \) can be written as \( s_n = \log_\alpha n \pmod{m} \) for \( n \not\equiv 0 \pmod{p} \).
The corresponding complex-valued sequence \( \omega_m \), where \( \omega_m \) is an \( m \)th root of unity, can be represented by a multiplicative character
\[
\psi(x) = \exp \left( \frac{2\pi \sqrt{-1} \log m}{m} x \right)
\]
of \( \mathbb{F}_p \), that is for \( x \neq 0 \) (mod \( p \)).

**Theorem 7.3.** [93] Let \( s(t) \) be an \( m \)-ary power residue sequence of odd prime period \( p \). For any \( \delta, 0 \leq \delta \leq m - 1 \) the maximum correlation magnitude of \( \mathcal{G}_2^\delta \) in Construction 7.2 is given by
\[
2 \sqrt{p} + 5.
\]
The maximum correlation of \( \mathcal{G}_2^\delta \) is achieved by applying the Weil’s theorem.

**Theorem 7.4.** [57] Let \( f_1(x), \ldots, f_l(x) \) be \( l \) monic, distinct, irreducible polynomials in \( \mathbb{F}_q[x] \) which have positive degrees \( d_1, \ldots, d_l \), respectively. Let \( d \) be the number of distinct roots of \( f(x) = \prod_{i=1}^l f_i(x) \) in its splitting field over \( \mathbb{F}_q \). Let \( \psi_1, \ldots, \psi_l \) be multiplicative characters of \( \mathbb{F}_q \), where \( \psi_i(0) = 0, 1 \leq i \leq l \). Assume that the product character \( \prod_{i=1}^l \psi_i(f_i(x)) \) is nontrivial, i.e. \( \prod_{i=1}^l \psi_i(f_i(x)) \neq 1 \) for some \( x \in \mathbb{F}_q \). Then, for every \( a_i \in \mathbb{F}_q^* \),
\[
\left| \sum_{x \in \mathbb{F}_q} \psi_1(a_1 f_1(x)) \cdots \psi_l(a_l f_l(x)) \right| \leq (d - 1) \sqrt{q}.
\]
In particular, if \( \prod_{i=1}^l \psi_i^{d_i}(x) = 1 \) for all \( x \in \mathbb{F}_q^* \), then
\[
\left| \sum_{x \in \mathbb{F}_q} \psi_1(a_1 f_1(x)) \cdots \psi_l(a_l f_l(x)) \right| \leq (d - 2) \sqrt{q} + 1.
\]

We consider a more complicated case. The sequence being shifted and added to construct a family is a two-prime generator of order two which can be expressed as a product of two Legendre symbols.

### 7.3. Shift-addition family from two-prime generators of order 2

The construction of the generalized cyclotomic sequence of order two (also called two-prime generator of order two) in Definition 2.36 is derived from a natural application of generalized cyclotomy by Whiteman [91], although he studied the generalized cyclotomy of order 2 only for the purpose of searching for residue
difference sets. The generalized cyclotomic sequence of order two has several good randomness properties [16]. Ding prove it has large linear span [17] and showed its autocorrelation values [18].

Let $p$ and $q$ be two distinct odd primes and put $n := pq$. Furthermore, we define

$$Q := \{q, 2q, \ldots, (p-1)q\}, \quad Q_0 = Q \cup \{0\}$$

and

$$P := \{p, 2p, \ldots, (q-1)p\}.$$

We consider the $n$–periodic sequence $(s_i)_{0 \leq i < n}$ over $\mathbb{F}_2$ defined by

$$s_i = \begin{cases} 
0, & \text{if } i \pmod{n} \in Q_0; \\
1, & \text{if } i \pmod{n} \in P; \\
1 - \left(\frac{i}{p}\right) \left(\frac{i}{q}\right), & \text{otherwise.}
\end{cases}$$

The sequence $(s_i)_{0 \leq i < n}$ is the two-prime generator (or generalized cyclotomic sequence) of order 2.

The autocorrelation of the two-prime generator of order two is the main result of [18]. The proof in [18] uses formulas for generalized cyclotomic numbers. Brandstätter and Winterhof present a slightly shorter proof without these formulas [5]. Their proof also do not need the restriction presented in [18] that $\gcd(p-1, q-1) = 2$.

**Theorem 7.5.** [5, 18] Let $s = s(i)$ for $0 \leq i \leq pq - 1$ be a two-prime generator of order two sequence associated with the two sets $Q$ and $P$. The autocorrelation of a two-prime generator of order two sequence is given by

$$C_s(\tau) = \begin{cases} 
q - p - 3, & \tau \in P; \\
p + 1 - q, & \tau \in Q; \\
-1, & \tau \in \mathbb{Z}_{pq}^*, \frac{p+q}{2} \text{ even}; \\
-1 + 2 \left(\frac{i}{p}\right) \left(\frac{i}{q}\right), & \tau \in \mathbb{Z}_{pq}^*, \frac{p+q}{2} \text{ odd};
\end{cases}$$

and

$$(-1)^{s(i)} = \begin{cases} 
1, & i \pmod{pq} \in Q_0; \\
-1, & i \pmod{pq} \in P; \\
\left(\frac{i}{p}\right) \left(\frac{i}{q}\right), & i \in \mathbb{Z}_{pq}^*.
\end{cases}$$
We use a two-prime generator of order 2 as the base sequence to construct a new family of sequences by means of the shift-addition construction. The binary case of Construction 7.2, that is, when \( m = 2 \), \( c_0, c_1 \in \{0, 1\} \), is shown below. For binary case we do not need to consider \( \delta \) appearing in Construction 7.2.

CONSTRUCTION 7.6. Let \( s = (s_i)_{0 \leq i \leq pq-1} \) be a two-prime generator of order 2 of period \( pq \), where \( 2 \mid (pq - 1) \). Let

\[
S_s = \{ s + cL^l(s) \mid 0 \leq l \leq \lceil pq/2 \rceil \}.
\]

The constructed family \( S_s \) is of period \( pq \), with all the sequences being cyclically distinct.

We now determine the maximum correlation magnitude of \( S_s \).

THEOREM 7.7. The maximum correlation magnitude of \( S_s \) in Construction 7.6 is given by \( 9\sqrt{pq} + 4(p + q) \).

PROOF. We derive the upper bound on the correlation magnitude of a pair of sequences \( a = (a_i) \) and \( b \) in \( A_s \) depending on several cases. In each case, we exclude \( \tau = 0 \) at \( a = b \).

- **Case 1.** \( a, b \in A_s \): Let \( a = b = s + L^l(s) \).

  For simplicity, let us write

  \[
  \sigma = \left( \frac{i}{p} \right) \left( \frac{i}{q} \right) \left( \frac{i + l_1}{p} \right) \left( \frac{i + l_1}{q} \right) \left( \frac{i + \tau}{p} \right) \left( \frac{i + \tau}{q} \right) \left( \frac{i + l_2 + \tau}{p} \right) \left( \frac{i + l_2 + \tau}{q} \right).
  \]

  We note that \( (-1)^{s_i} = \left( \frac{i}{p} \right) \left( \frac{i}{q} \right) \) for \( i \not\equiv 0 \mod p \). The correlation is given by

  \[
  |C_{a,b}(\tau)| = \left| \sum_{i=0}^{pq-1} (-1)^{s_i + s_{i+1} - s_{i+\tau} - s_{i+i+\tau}} \right|
  \]

  \[
  \leq \sum_{i=0}^{pq-1} |\sigma| + 4(p + q),
  \]

  \[
  \leq \sum_{i=0}^{pq-1} |\sigma| + 4(p + q),
  \]
where the error term $4(p + q)$ is given for the cases of $(-1)^{s_i} \neq \left(\frac{i}{p}\right) \left(\frac{i}{q}\right)$.

Then

$$|C_{a,b}(\tau)| \leq \left| \sum_{u \in \mathbb{F}_p} \left(\frac{u}{p}\right) \left(\frac{u + \tau}{p}\right) \left(\frac{u + l}{p}\right) \left(\frac{u + l + \tau}{p}\right) \right| \cdot \left| \sum_{v \in \mathbb{F}_q} \left(\frac{v}{q}\right) \left(\frac{v + \tau}{q}\right) \left(\frac{v + l}{q}\right) \left(\frac{v + l + \tau}{q}\right) \right| + 4(p + q)$$

$$\leq 3\sqrt{p} \cdot 3\sqrt{q} + 4(p + q)$$

$$= 9\sqrt{pq} + 4(p + q).$$

The result is obtained by applying Theorem 7.4.

• **Case 2.** $a, b \in A_2$: Let $a = s + L_1(s)$, $b = s + L_2(s)$, where $l_1 \neq l_2$.

The correlation is given

$$|C_{a,b}(\tau)| = \left| \sum_{i=0}^{pq-1} (-1)^{s_i + s_i + i_1 - s_{i+\tau} - s_{i+1} + \tau} \right|$$

$$\leq \left| \sum_{i=0}^{pq-1} \sigma \right| + 4(p + q)$$

$$= \left| \sum_{u \in \mathbb{F}_p} \left(\frac{u}{p}\right) \left(\frac{u + \tau}{p}\right) \left(\frac{u + l_1}{p}\right) \left(\frac{u + l_2 + \tau}{p}\right) \right| \cdot \left| \sum_{v \in \mathbb{F}_q} \left(\frac{v}{q}\right) \left(\frac{v + \tau}{q}\right) \left(\frac{v + l_1}{q}\right) \left(\frac{v + l_2 + \tau}{q}\right) \right| + 4(p + q)$$

$$\leq 3\sqrt{p} \cdot 3\sqrt{q} + 4(p + q)$$

$$= 9\sqrt{pq} + 4(p + q).$$

• **Case 3.** $a = s$, $b = s + L^l(s)$.  


The correlation is given

\[ |C_{a,b}(\tau)| = \left| \sum_{i=0}^{pq-1} (-1)^{s_i-s_{i+\tau}-s_{i+\tau}} \right| \]

\[
\leq \left| \sum_{i=0}^{pq-1} \left( \frac{i}{p} \right) \left( \frac{i}{q} \right) \left( \frac{i+\tau}{p} \right) \left( \frac{i+\tau}{q} \right) \right| + 3(p+q) \\
= \left| \left( \sum_{u \in \mathbb{F}_p} \left( \frac{u}{p} \right) \left( \frac{u+\tau}{p} \right) \right) \left( \sum_{v \in \mathbb{F}_q} \left( \frac{v}{q} \right) \left( \frac{v+\tau}{q} \right) \right) \right| + 3(p+q) \\
\leq 2\sqrt{p} \cdot 2\sqrt{q} + 3(p+q) \\
= 4\sqrt{pq} + 3(p+q).
\]

- **Case 4.** When \( a = \underline{s} \) and \( b = L^l(\underline{s}) \), the cross correlation is equal to the autocorrelation of \( \underline{s} \). The values are given by Theorem 7.5:

\[
C_{\underline{s}}(l+\tau) = \begin{cases} 
q - p - 3, & l + \tau \in P; \\
p + 1 - q, & l + \tau \in Q; \\
-1, & l + \tau \in \mathbb{Z}_{pq}^*, \frac{p+q}{2} \text{ even}; \\
-1 + 2 \left( \frac{l+\tau}{p} \right) \left( \frac{l+\tau}{q} \right), & l + \tau \in \mathbb{Z}_{pq}^*, \frac{p+q}{2} \text{ odd}. 
\end{cases}
\]

Besides the interleaved construction, the shift-addition construction is also widely used. A shift-addition family from power residue sequences has been shown in Section 7.2. We apply the shift-addition construction on a two-prime generator of order two and obtain a \( \left( pq, \left\lceil \frac{pq}{2} \right\rceil, 9\sqrt{pq} + 4(p+q) \right) \)-family. The period of each sequence is \( pq \), where \( p \) and \( q \) are two distinct odd prime numbers. The period of the shift-addition sequences in Theorem 7.3 is odd prime \( p \). Now we normalize the maximal correlation magnitude and get

\[
\frac{9}{\sqrt{pq}} + 4 \left( \frac{1}{p} + \frac{1}{q} \right).
\]

The normalized maximal correlation magnitude in Theorem 7.3 is

\[
\frac{2}{\sqrt{p}} + \frac{5}{p}.
\]
There are infinitely many choices of $p$ and $q$ such that the maximal correlation magnitude in (48) is smaller than (49). So our result improves Theorem 7.3. One simple example is for twin prime $p$ and $q = p + 2 > 36$. 
CHAPTER 8

Future work

This thesis naturally leads to several future works.

(1) **Continuing work on interleaved construction:** In this thesis, we have dealt with interleaving two base sequences. A natural extension is to consider three or more base sequences. We were not aware of any work in this direction till now. So we start by interleaving three maximum length sequences which are generated by primitive polynomials and have ideal two-level autocorrelation. Let \( a, b \) and \( c \) be three \( m \)-sequences and let \( e \) and \( f \) be two shift sequences. We have obtained a family of sequences with low correlation [35] which covers the case that the two base sequences are \( m \)-sequences originally given in [29]. We consider an obtained sequence \( s_k \in \Theta(a, b, e) \) to be our first base sequence and \( c \) to be our second base sequence. By using another shift sequence \( f \) and applying the interleaving construction again, we can obtain more families of sequences and the cross correlation of those families are still low.

I am also planning to shift-and-add interleaved \( \ell \)-ary sequences. Assume we get a family of interleaved sequences \( \Theta \) by interleaving construction. Let \( s_k = (s_0, \ldots, s_{pq-1}) \) be a sequence in \( \Theta \) over \( F_\ell \). If \( \delta \) is a fixed value in \( F_\ell \), then we shift-and-add \( s_k \) in the way that

\[
t_k = c_0 s_k + c_1 L^w(s_k),
\]

where \( c_0, c_1 \in F_\ell \) and \( c_0 + c_1 = \delta \). The set \( \{t_k \mid 0 \leq k \leq q, \ 1 \leq w \leq \frac{pq-1}{2}\} \) with a fixed \( \delta \) can be a potential family of sequences to study. We are interested in the correlation measures of such sequences and I want to explore if the correlation measures are low.

(2) **Sequences defined from multiplicative characters:** Sequences like power residue (particularly Legendre sequences), Sidelnikov and two-prime
generator of order two sequences are well-known polyphase sequences with low correlation and the terms of these sequences can be represented by multiplicative characters. We have introduced Legendre sequences, power residue sequences and two-prime generator of order two in Section 2.4.2.

A Sidelnikov sequence is defined as follows. Let $p$ be an odd prime and $m$ be a positive integer. Let $\mathbb{F}_{p^n}$ be the finite field with $p^n$ elements, and $\alpha$ be a primitive element of $\mathbb{F}_{p^n}$. Let $m$ be a divisor of $p^n - 1$. Let $D_k = \{\alpha^{mi+k} - 1 \mid 0 \leq i < \frac{p^n-1}{m}\}$ for $0 \leq k \leq m - 1$. The $m$-ary Sidelnikov sequence $s(t) = \{s(t) \mid t = 0, 1, 2, \ldots, p^n - 2\}$ of period $p^n - 1$ is defined as

$$s(t) = \begin{cases} 
0, & \text{if } \alpha^t = -1; \\
k, & \text{if } \alpha^t \in D_k.
\end{cases}$$

Equivalently, $s(t)$ is defined by $s(t) \equiv \log_\alpha (\alpha^t + 1) \pmod{m}$, $0 \leq t \leq p^n - 2$. In [78], it was shown that Sidelnikov sequences have optimal autocorrelation and balance property.

We plan to define new sequences based on the three types of polyphase sequences above. All the three types of sequences with low correlation can be studied by the known bounds on sums involving characters with polynomial arguments. These results trace back to Weil. A key tool for this future is the use of bounds on hybrid sums. The result is given in the next theorem.

**Theorem 8.1.** [66] Let $\psi$ be a nontrivial multiplicative character of $\mathbb{F}_p$ of order $d$, and let $\chi$ be a nontrivial additive character of $\mathbb{F}_p$. Suppose that $g(x) \in \mathbb{F}_p[x]$ has $m$ distinct roots in its splitting field and that $g(x)$ is not a $d$th power. Suppose further that $f(x) \in \mathbb{F}_p[x]$ has degree $n$. Then

$$\left| \sum_{x \in \mathbb{F}_p} \psi(g(x))\chi(f(x)) \right| \leq (m + n - 1)\sqrt{p}.$$ 

Let $p$ and $q$ be odd prime numbers and $n$ be an integer such that $\gcd(p, q^n - 1) = 1$. Let $m$ be a divisor of both $p - 1$ and $q^n - 1$. Put $T = p(q^n - 1)$ and define two sets $Q = \{\frac{p-1}{2} + j(q^n - 1) \mid j = 0, 1, \ldots, p-1\}$ and $P = \{0, p, 2p, \ldots, (q^n - 1)p\}$. We note that $P \cap Q = \{T/2\}$ and put $Q^* = Q \setminus \{T/2\}$ and $R = \{0, 1, 2, \ldots, T - 1\} \setminus (P \cup Q^*)$. Define
a multiplicative character $\psi$ of order $m$ over $\mathbb{F}_p$ and define a nontrivial additive character $\chi$ of order $m$ over $\mathbb{F}_{q^n}$.

Then we define an $m$–ary sequence $s = s(t)$ of period $T$ as

$$s(t) = \begin{cases} 0, & \text{if } t \equiv 0 \pmod{T}; \\ \chi(a_r p^r + a_{r-1} p^{r-1} + \cdots + a_0) \psi(\alpha^t + 1), & \text{otherwise}, \end{cases}$$

where $\alpha$ is a primitive element of $\mathbb{F}_p$.

To estimate the magnitude of the correlation and the power correlation measure we apply the bound on hybrid sum. Discussion on estimating $\chi(a_r p^r + a_{r-1} p^{r-1} + \cdots + a_0)$ can be found in [57, pp.218].

(3) Hu and Gong increase the number of shift sequences under some assumptions and the number of the second base sequences to obtain more sequences by interleaving techniques [39]. The family they give is $(s, \omega_j, e_i)$, where $s$ corresponds to the first base sequence in the interleaved algorithm, $\omega_j$ the second base sequences and $e_i$ the shift sequences. Assume there are $M$ shift sequences $e_i$ with $i = 1, \ldots, M$. Because more shift sequences are employed, the signal set [39] has a larger size and the cross correlation has been shown to be low. We expect that an increase on the number of shift sequences or an increase on the number of the second base sequences in Algorithm 6.5 may bring more sequences to the family while still keeping the low correlation.
Bibliography


